Universidad de Buenos Aires

Facultad de Ciencias Exactas y Naturales

Departamento de Ciencias de la Atmósfera y los Océanos

Coherencia regional entre el clima y el régimen hidrológico en distintas cuencas hídricas de la Argentina. Implementación y comparación de pronósticos estadísticos de caudal.

Tesis presentada para optar por el título de Doctor de la Universidad de Buenos Aires en el área de Ciencias de la Atmósfera y los Océanos

Melanie Meis

Directores: Dra. María Paula Llano y Dra. Daniela Rodriguez
Consejero de Estudios: Dra. Marcela Hebe González
Lugar de trabajo: Departamento de Ciencias de la Atmósfera y los Océanos
Buenos Aires, 12 de diciembre de 2019

Resumen

El uso de los distintos cuerpos de agua de un territorio para el transporte fluvial, la generación de energía hidroeléctrica, el turismo, y el provicionamiento de agua, como así también lo que refiere a sistemas de alerta y riesgo requieren del trabajo conjunto de distintos sectores de una sociedad. Sin embargo, en las últimas décadas este trabajo ha requerido un mayor esfuerzo puesto que se ha visto afectado por la ocurrencia de eventos extremos tanto de crecidas como de bajantes, en ríos de gran interés, causando severas consencuencias socioeconómicas. En este sentido, la continúa contribución desde el área científica resulta ser necesaria para colaborar en el desarrollo adecuado de las distintas actividades económicas mencionadas, como así también con los sistemas de alerta y riesgo. Por lo tanto, en este trabajo de tesis se pretendió ahondar la relación entre la coherencia tanto hidrológica como climática dentro de un mismo río y entre distintas cuencas, en particular, la cuenca del río Paraná, Uruguay y la del río Atuel. Analizar las series de tiempo del caudal de un río constituye una síntesis de situaciones climáticas sobre una cuenca, permitiendo de ese modo realizar inferencias climatológicas en cada una de las regiones particulares de estudio. En este sentido se propuso estudiar el caudal como indicador de variabilidad climática e hidrológica, es decir, analizando la relación entre el caudal y distintas oscilaciones climáticas como el ENSO, entre otras. Asimismo, distintos modelos estadísticos fueron empleados para representar y pronosticar series temporales de caudal, donde diferentes métricas de evaluación de rendimiento fueron tenidas en cuenta, por ejemplo, varios de los resultados mostraron que podrían ser de utilidad en la prevención de eventos extremos. Es decir, los mismos permitirían a los tomadores de decisión contar con mayor cantidad de herramientas para evaluar una posible mitigación de los impactos negativos relacionados con el caudal.

Palabras clave: caudal, coherencia, eventos extremos, pronóstico, modelos estadísticos

Abstract

The use of the different water resources of a territory for goods or passengers transportation, hydroelectric energy generation, water supply, as well as warning and risk systems require a joint work from different sectors of the society. However, in recent years, this joint work has been considerably affected by the occurrence of extreme events in rivers of great interest. That is to say, both river floods and discharge reduction have generated serious socioeconomic consequences. In this sense, the continuous contribution of the scientific area results necessary to colaborate to the adequate development of the above economic activities, as well as to generate alert and risk systems. Thence, in this thesis work, we pretended to deepen the relation between hydrological and climate coherence within a particular river and between different basins, particularly, Paraná, Uruguay and Atuel rivers. Studying discharge time series involves a synthesis of climatic situations within a basin, which enable the possibility of carrying out climatological inferences in each particular region studied. As follows, it was proposed to study the streamflow as a climatic and hydrologic variability indicator, going deeper in the relation between discharge and several climatic oscillations such as ENSO, among others. In addition, different statistical models were employed to represent and forecast discharge time series, where multiple perfomance evaluation metrics were obtained. From these results, it was seen that them could be considered for extreme events prevention. Thus, the results obtained might be useful for the decision makers, by expanding their set of available tools for assessing options for mitigation of negative impacts related to discharges.

Keywords: discharge, coherence, extreme events, forecast, statistical models

Dedicaciones

A Maximiliano y Benjamín

Agradecimientos

- A mi Director de beca de doctorado Dr. Vargas por darme su apoyo en estos años y confiar en mí.
- A mis Directoras Dra. Llano y Dra. Rodriguez por acompañarme y apoyarme en este proceso de varios años.
- A mi consejera de estudios la Dra. Marcela González, gracias por apoyarme y acompañarme en estos años.
- A las personas que trabajan en el Departamento de Ciencias de la Atmósfera y los Océanos.
- A las personas que trabajan en el Instituto del Cálculo por abrirme sus puertas y brindarme su apoyo y conocimiento en todo momento.
- A Mariela Sued y Manuel Benjamín por el apoyo, conocimiento y el trabajo brindado.
- A Leandro Kazimierski y Ramiro Saurral por su ayuda y conocimiento brindado.
- A la Facultad de Ciencia Exactas y Naturales de la Universidad de Buenos Aires.
- Al CONICET por solventar mis estudios de doctorado.
- A los jurados de esta tesis Dr. Araneo, Dr. Grimson y Dr. Seoane por sus comentarios y dedicación brindada que enriquecieron este trabajo.
- A Maxi el amor de mi vida, que fue fundamental en el apoyo, en sus conocimientos y en levantarme más de una vez. Te amo. Y a Benjamín que en los últimos meses del doctorado llegó a nuestra vida para darme el último empujón y hacerme super feliz, te amo.

- A mis papás Beatriz y Rodolfo gracias por todo lo que hicieron por mi siempre, me ayudaron a ser quien soy hoy, una mejor persona y profesional. Los amo.
- A mis abuelos Gladis y Jorge gracias porque se que siempre me acompañan estén en el lugar que estén. Los amo.
- A mi familia heredada mis suegros Patricia y Eduardo, y mi cuñada Agustina por escucharme siempre, apoyarme y animarme. Gracias por su paciencia infinita.
- A mis amigos de la facu son incondicionales siempre estuvieron en las buenas y en las malas. Se que tengo el mejor grupo de amigos que alguien podría tener los adoro. Gracias por ser como son.
- A Roberta nuestra gata hermosa que me acompaño en mis crisis doctorales y me dió su amor constante aunque también recibí algún rasguño.

Tabla de contenidos

Introducción General **1**

Referencias **7**

1. Análisis y Selección de estaciones de referencia **8**

1.1. Resumen 8

1.2. Introducción 8

1.3. Metodologías 10

1.4. Cuenca del Plata 15

1.4.1. Río Paraná 15

1.4.2. Río Uruguay 27

1.5. Cuenca del Atuel 35

1.5.1. Río Atuel 35

1.6. Conclusiones 44

1.7. Referencias 46

2. Estudio de coherencia entre estaciones de referencia **51**

2.1. Resumen 51

2.2. Introducción 51

2.3. Datos y Metodologías 53

2.4. Resultados 53

2.5. Conclusiones 66

2.6. Referencias 67

3. Oscilaciones Climáticas y el Caudal **68**

3.1. Resumen 68

3.2. Introducción 68

3.3. Índices climáticos 71
3.4. Datos y Metodología 72
3.5. Resultados 77
3.5.1. Forzantes climáticos y su relación con el caudal 77
3.5.2. Estudio de las relaciones bivariadas mediante cópulas 95
3.5.3. Ajustes Marginales 99
3.5.4. Evaluación del modelo 108
3.5.5. Pronóstico desde la distribución conjunta 111
3.5.6. Evaluación del percentil 95 como medida de aviso temprano de evacuación 122
3.5.7. Comparación entre el modelo teórico y los datos observados . . . 130
3.5.8. Probabilidad estimada para eventos extremos 131
3.5.9. Periodos de retorno bivariados en eventos extremos 133
3.6. Conclusiones 141
3.7. Referencias 143

4. Modelado y Predicción de las series de Caudal en la Baja Cuenca del Plata 149

4.1. Resumen 149
4.2. Introducción 149
4.3. Datos y Metodología 151
4.4. Resultados 154
4.4.1. Análisis en la estación Túnel Subfluvial (río Paraná) 154
4.4.2. Pronóstico SARIMA para la estación Túnel Subfluvial a 6 y 12 meses158
4.4.3. Pronóstico SARIMA para la estación Corrientes a 6 y 12 meses . 159
4.4.4. Pronóstico SARIMA para la estación Paso de los Libres a 6 y 12 meses 159
4.4.5. Incorporar variable exógena mediante modelos SARIMAX 160
4.5. Conclusiones 164
4.6. Referencias 166

Apéndice A. Predicción del Caudal en el Paraná Medio 168

A.1. Resumen 168

A.2. Introducción 169
A.3. Datos y Metodología 171
A.4. Resultados 175
A.5. Conclusiones 178
A.6. Referencias 180

5. Conclusiones Generales **182**
5.1. Conclusiones 182
5.2. Referencias 189

Apéndice B. Tablas Capítulo 4 **190**
B.1. Túnel Subfluvial 190
B.2. Corrientes 195
B.3. Paso de los Libres 199

Introducción General

Distintos aspectos socioeconómicos de la Argentina son fuertemente dependientes de la conjunción entre el clima y los regímenes hidrológicos. En particular, el territorio argentino presenta una numerosa cantidad de cuencas hídricas, muchas de las cuales resultan de suma importancia para el desarrollo económico y social de la región, siendo las mismas varias veces afectadas por devastadoras inundaciones y/o por periodos extensos de sequía.

Por otra parte, la disponibilidad de los recursos hídricos en la Argentina se encuentra distribuida de manera no equitativa, 2/3 del territorio consiste en zonas áridas y semiáridas, y tan solo el 1/3 restante presenta cuerpos de agua (IANAS, 2015). Por lo que el aprovechamiento de los mismos para las actividades económicas se realiza de forma regional. En este sentido, como ejemplo se puede mencionar que la explotación de energía e irrigación en Argentina, en general, se lleva a cabo en ríos con gran caudal (Popescu et al., 2014). Estudios como el de Guerrero et al. (2013) analizaron la relación entre el cambio climático y la modificación del costo de drenaje para mantener el canal de navegación del río Paraná, donde a través de modelos numéricos evaluaciones estadísticas de escenarios climáticos fueron utilizados para calcular la probabilidad de fracaso en la navegación y el costo asociado del mantenimiento del canal. Más aún, el trabajo propuesto por Castex et al. (2015) muestra que la disponibilidad de recursos hídricos en la región de Mendoza es crucial para la actividad vitivinícola, y que la cantidad de caudal y la potencialidad de la recarga del acuífero es fuertemente dependiente de la contribución de la nieve y el derretimiento del hielo cercano a los Andes.

La profundización de estudios en cuencas hídricas tiene como objetivo colaborar en la mejora de las diversas actividades económicas tales como la agricultura, la ganadería, la pesca, la hidroelectricidad, el transporte, entre otras. Más aún, actividades como las mencionadas, en conjunto con el turismo resultan de suma importancia en lo que refiere

a los ingresos económicos de distintas provincias.

En base a los anterior, estudios estadísticos realizados en cuencas de gran importancia para la región como es la Cuenca del Plata (LPB, del inglés La Plata Basin) han mostrado que los regímenes hidrológicos presentaron cambios y diferentes tendencias (García y Vargas, 1998). Varias razones, entre ellas el clima, pueden generar diferentes regímenes de caudal, es decir, la variabilidad del caudal respondería a diversidad de fenómenos climáticos. En este sentido, varios autores como Robertson y Mechoso (1998) y Camilloni et al. (2004) distinguieron que los caudales de los principales ríos de la LPB presentan una variabilidad interanual e interdecadal causada por la variabilidad climática. También, Rivera et al. (2017) analizaron los aspectos regionales de las sequías en términos de las deficiencias del caudal en los ríos andinos de la Patagonia. Los autores encontraron que oscilaciones climáticas de onda larga presentaron un marcado comportamiento regional y temporal en modular la variabilidad de la sequía en los caudales.

Otro de los factores influyentes en los recursos naturales, específicamente en los recursos hídricos es el cambio climático, que no solo actúa como una fuente de impacto sino que también se encuentra ligado a un aumento en la vulnerabilidad de la población. En este sentido, el cambio climático pronosticado por varios científicos podría acarrear una disponibilidad insuficiente de energía hidroeléctrica durante los periodos secos. Sin embargo, es necesario destacar que el caudal medio en LPB como un todo ha aumentando un 35 % desde $1951-1970$ a $1980-1999$ (Berbery y Barros, 2002) dando lugar a un aumento en la producción hidroenergética, navegabilidad y otros usos a partir de los recursos hídricos. Es decir, la importancia en la detección temprana de cambios y/o singularidades en las variables hídricas da lugar a una posible planificación de los recursos hídricos de una región permitiendo así una minimización de impactos negativos. En este sentido, el trabajo de Lauro et al. (2016) ha estudiado saltos y tendencias en ríos cuyanos, encontrando ciertas disimilitudes con lo propuesto por autores anteriores en lo que refiere a la significancia de las tendencias. Por ejemplo, a diferencia de no haber evidencia significativa de tendencia en la estación La Angostura del río Atuel por parte de Boninsegna (2009), los autores hallaron una tendencia significativa positiva en el caudal anual.

Por otra parte, Camilloni et al. (2013) con el objetivo de estudiar el cambio climático en la LPB analizaron mediante un modelo hidrológico y datos provenientes de modelos climáticos proyecciones a largo plazo del caudal obteniendo como medidas de estudio la

frecuencia de eventos de inundación en los tramos inferiores de los ríos Paraná y Uruguay. Los autores si bien proyectan un aumento de cierta frecuencia, las incertidumbres son de una magnitud considerable. Asimismo, en Schawank et al. (2013) también consideraron proyecciones climáticas y otro modelo hidrológico para inferir pronósticos del caudal en el río Mendoza, arrojando como resultado una disminución del caudal, acarreando una merma de la disponibilidad del agua, y teniendo como principal impacto la agricultura, fuente principal de la economía regional.

Por lo tanto, se entiende que tanto la variabilidad como el cambio climático afectan el caudal, y de esta manera resulta de suma importancia profundizar los estudios para generar un mayor conocimiento del tema. Puesto que, una buena planificación y prevención socioeconómica frente a situaciones extremas es de relevancia para todos los sectores de un país.

Si bien ha quedado expuesto que gran cantidad de trabajos han estudiado el régimen hidrológico de diferentes ríos y su relación con el clima, todavía existen distintos acercamientos posibles para continuar explorando las variabilidades a largo plazo, junto con nuevos registros temporales de información que darían lugar a continuar comprendiendo su comportamiento. En este sentido, hay nuevas metodologías que no han sido aplicadas aún en nuestros ríos, y que podrían permitir obtener otra mirada de la información de los mismos, y de esta manera profundizar en la relación de los ríos con fenómenos climáticos.

Más aún, en el presente trabajo de tesis se pretende ahondar la relación entre la coherencia tanto hidrológica como climática dentro de una misma cuenca y entre distintas cuencas en particular la cuenca del río Paraná, Uruguay y la del río Atuel (las cuales presentan características fisiográficas diferentes). Pues, estudiar la serie de tiempo del caudal de un río constituye una síntesis de situaciones climáticas sobre una cuenca, permitiendo de ese modo realizar inferencias climatológicas en la región particular de estudio.

En concordancia con lo anterior, profundizar los estudios en ciertos ríos de la Cuenca del Plata y la cuenca del río Atuel resulta de gran utilidad por ser fuentes continuas de aprovechamiento de recursos hídricos (abastecimiento, navegabilidad, generación de energía y riego). Además, la Cuenca del Plata presenta algunas características negativas asociadas a situaciones extremas en lo que refiere a inundaciones y bajantes en las tomas de agua. Por lo cual, el análisis en estaciones de medición de referencia en periodos extensos permite obtener una evaluación para solventar los delicados problemas de decisión

que impone el pronóstico hidrológico.
Asimismo, estudiar el comportamiento de datos hidrológicos como es el caudal durante un largo periodo de tiempo resulta una herramienta importante y necesaria para el manejo y diseño asociado a los recursos hídricos de un país. Conforme a lo anterior, uno de los posibles lineamientos de investigación asociados a los recursos hídricos se encuentra ligado a la hidrología estocástica, la cual mediante la exploración y el análisis de los datos, por medio de distintos modelos estadísticos y probabilísticos permite en algún sentido ayudar a la interpretación física y estimación de probabilidades de ocurrencia de distintos fenómenos.
En base a estos antecedentes se motiva el siguiente objetivo general.

Objetivo General

Análisis sobre la coherencia regional entre diferentes regímenes hidrológicos y el clima para la evaluación de modelos estadísticos de pronóstico del caudal de los ríos Paraná, Uruguay y Atuel.

Objetivos Específicos

- Estudiar la coherencia climática dentro de cada cuenca hidrológica para la selección de estaciones de referencia.
- Analizar el comportamiento conjunto entre cuencas y con respecto a patrones característicos del clima de la región.
- Analizar el riesgo conjunto de extremos climáticos y su influencia en los ríos estudiados.
- Implementar modelos de pronóstico estadístico en los ríos considerados y comparación de los mismos.

Esta tesis se estructura en cinco capítulos. El primer capítulo tuvo como objetivo la exploración de las series de tiempo del caudal mensual y la selección de estaciones de referencia de cada cuenca estudiada. En el capítulo dos se pretendió estudiar, evaluar y analizar la coherencia entre las estaciones de referencia con el fin de evaluar si hay comportamientos similares o no. Un tercer capítulo que estudió el riesgo conjunto de eventos extremos de

caudal estacional de las estaciones de referencia y distintos índices climáticos a fin de obtener un modelo de pronóstico. Además, se cuenta con un capítulo cuatro donde se analizó para distintas estaciones del río Paraná a nivel mensual un modelo de pronóstico. Asimismo, se presenta un apéndice que involucró la exploración de una nueva metodología multivariada en series de caudal. Por último se incluye un capítulo cinco con las conclusiones generales del trabajo de tesis.

Referencias

Boninsegna, J., 2009. Impacto del Cambio Climático en la región Centro Oeste de Argentina y posibles medidas de adaptación. Foro: Escenarios climáticos y medidas de adaptación. Secretaria de Ambiente de la Nación. Buenos Aires. http://ambiente.gov.ar/archivos/web/AdCC/File/09_boninsegna.pdf/

Berbery, H., Barros, V.R., 2002. The hydrologic cycle of the La Plata basin in South America. Journal of Hydrometeorology. Vol. 3, 630-645.

Camilloni, I., Saurral, R., Montroull, N., 2013. Hydrological projections of fluvial floods in the Uruguay and Paraná basins under different climate change scenarios. International Journal of River Basin Management.Vol 11:4, 389-399.

Camilloni, I., Doyle, M., Barros, V.R., 2004. Interannual Variability of the South Atlantic High and Rainfall in Southeastern South America during Summer Months. En: XIII CONGRESSO BRASILEIRO DE METEOROLOGIA, 2004, Anais.

Castex, V., Morán Tejeda, E., Beniston, M., 2015. Water availability, use and governance in the wine producing region of Mendoza, Argentina. Environmental Science & policy. VOl 48, 1-11.

García, N., Vargas, W., 1998. The temporal climatic variability in the Rio de la Plata basin displayed by the rivers discharges. Climatic Change. Vol. 38, 359-379.

Guerrero, M., Re, M., Kazimierski, L., Menéndez, A., Ugarelli, R., 2013. Effect of climate change on navigation channel dredging of the Parana River. International Journal of

River Basin Management. Vol. 11, No. 4, 439–448.

IANAS The Inter-American Network of Academies of Sciences, 2015. Urban water challenges in the Americas. 1-622.

Lauro, C., Vich, A., Moreiras, S., 2016. Variabilidad del régimen fluvial en cuencas de la región de Cuyo. GEOACTA. Vol. 40(2), 28-51.

Popescu I., Brandimarte, L., Peviani, M., 2014. Effects of climate change over energy production in La Plata Basin. International Journal of River Basin Management. Vol. 12, 319–327.

Rivera, J., Araneo, D., Penalba, O., Villalba, R., 2017. Regional aspects of streamflow droughts in the Andean rivers of Patagonia, Argentina. Links with large-scale climatic oscillations. Hydrology research. En prensa.

Robertson, A.W. and Mechoso, C.R., 1998. Interannual and decadal cycles in river flows of southeastern South America. J. Climate. Vol. 11, 2570-2581.

Schawank, J., Escobar, R., Girón G., Morán Tejada, E., 2014. Modeling of the Mendoza river watershed as a tool to study climate change impacts on water availability. Environmental science & policy. 1-6.

Capítulo 1

Análisis y Selección de estaciones de referencia

1.1. Resumen

En este capítulo se analizaron las series temporales de las estaciones de medición localizadas en las cuencas de los ríos Paraná, Uruguay y Atuel. Estos ríos que poseen características fisiográficas e hidrológicas diferentes presentan caudales que son útiles para la generación de energía e irrigación. Una exploración de los datos de caudal mensual fue llevada a cabo mediante ciertas medidas estadísticas. Asimismo, se contemplaron técnicas espectrales para representar la variabilidad espacial y temporal de las estaciones. De esta manera, la coherencia climática entre pares de estaciones fue analizada dando lugar a un resumen de la información disponible a partir de la semejanza entre grupos de estaciones de un mismo río.

Mediante este análisis exploratorio para la red de estaciones seleccionada se sugirió la selección de ciertas estaciones de referencia con el objetivo siguiente de profundizar estudios hidroclimatológicos para los ríos considerados.

1.2. Introducción

La explotación de energía hidroeléctrica tanto como la irrigación en Argentina presentan su mayor concentración en ríos de gran caudal. Cuencas como la del río Paraná, Uruguay en primer lugar y en un tercer lugar la del Atuel cumplen la condición anterior. De es-

ta manera es necesario el estudio conjunto entre el clima y los factores hidrológicos, en particular en esos ríos.

Autores como García y Vargas (1995) y Collischonn et al. (2001) han estudiado cómo la precipitación impacta en el caudal evidenciando que los incrementos en el volumen del agua de los ríos han sido mayores con respecto al valor medio de la precipitación en la LPB. Estos cambios representativos podrían estar asociados con la deforestación y el cambio en el manejo de la agricultura causando una mayor escorrentía superficial (Tucci, 2001), y como consecuencia una variación en la tasa de infiltración. Asimismo, García y Vargas (1998) han explorado la relación precipitación-caudal en la Cuenca del Plata, donde identificaron un notable cambio en el valor medio de ambas variables en la década del setenta. Doyle y Barros (2011) también estudiaron la relación anterior destacando los distintos factores causales del aumento del caudal en las últimas décadas en LPB. En Vargas et al. (2010) se analizó la estructura de las anomalías en el caudal de los ríos Paraná y Uruguay para poder establecer un modelo de decisión con respecto a la ocurrencia de extremos climáticos. Asimismo, Antico et al. (2015) estudiaron la relación entre la contribución de distintas oscilaciones en las cuatro mayores inundaciones registradas en el río Paraná.

En otro orden, los estudios sobre el jet en capas bajas (Berbery y Collini, 2000; Berbery y Barros, 2002), la zona de convergencia del Atlántico Sur (Barros et al., 2000; Barros et al., 2002) y la presencia del anticiclón del Atlántico Sur (Prohaska, 1976) muestran que estos factores son prioritarios y determinantes significativos del clima y la precipitación en la LPB. De esta manera es importante entender un número de aspectos que afectan el comportamiento del caudal. Por otra parte, la cuenca del Atuel se encuentra regida por la fusión nival y la acumulación de nieve sobre las altas cordilleras de los Andes (Compagnucci y Vargas 1998). A su vez, la fluctuación del caudal del río Atuel se encuentra relacionada con la circulación atmosférica durante la estación fría, especialmente en las áreas oeste del Pasaje de Drake, el Pacífico Sur y sobre la zona adyacente a Sudamérica (Araneo y Compagnucci, 2008). Asimismo, algunos trabajos como el de Araneo y Villalba (2015) sugieren que el comprensivo comportamiento de la variabilidad hidroclimática en particular de regiones montañosas requieren de series temporales con un periodo de registro extenso de manera tal de caracterizar la variabilidad y tendencias en las mismas.

En este capítulo se consideró el comportamiento hidrológico de las series de tiempo de

la red de estaciones para la cuenca del río Paraná, Uruguay y Atuel en cada uno de sus máximos periodos comunes disponibles. Pues, mediante dichas series temporales se estudió si la ubicación de las estaciones era óptima, suficiente, baja o redundante, a fin de realizar una selección de estaciones como referencia, con un mínimo número de estaciones para la representación regional. Esto podría resultar relevante y de utilidad para la elaboración de proyectos de infraestructura relacionados con recursos hídricos y sus usos en diferentes estados climáticos (Oki y Kane, 2006).

1.3. Metodologías

El trabajo se inició con un análisis estadístico para caracterizar las series temporales de estaciones de medición: onda media anual, máximos y mínimos absolutos, y desvío estándar del caudal. Con el objetivo de estudiar la coherencia espacial y temporal en un periodo en común para cada cuenca en particular.

Datos faltantes

En cuanto a los datos faltantes que pudiesen presentar las estaciones de medición que se consideraron en la presente tesis se tuvo en cuenta el procedimiento que se detalla a continuación. En el caso de que se tuviese hasta dos datos contiguos faltantes se utilizó para completar los mismos una interpolación lineal. Por otra parte, si el periodo temporal faltante resultaba ser de mayor longitud se procedió a obtener la diferencia entre el resto de los meses cercanos al periodo en cuestión y sin faltantes. Luego, los años que presentaron menor diferencia con respecto al año faltante se tuvieron en cuenta para ser candidatos a completar los meses del periodo faltante. Más aún, los meses de los dos años que presentaron la menor diferencia fueron los que se promediaron para completar el periodo faltante.

Correlación de Spearman

Un análisis de correlación fue realizado a partir del coeficiente de Spearman (ρ) (Siegel, 1985) (Ec. 1.1). El test no asume que la distribución muestral se ajusta a una cierta distribución teórica particular. La correlación se calcula entre dos variables aleatorias (x e y). En el caso de este procedimiento, los datos son ordenados y reemplazados por su

rango, más precisamente se definen:

$$\hat{\rho} = 1 - \frac{6\sum D^2}{N(N^2-1)} \tag{1.1}$$

donde D es la diferencia entre los correspondientes estadísticos de orden de $x - y$. N es el número de parejas de datos.

Para evaluar la significancia del coeficiente de correlación empleamos el test de Student a dos colas con un 95 % de confianza (Ec. 1.2), (R Core Team, (2013)):

$$t = \hat{\rho}\sqrt{\frac{N-2}{1-\hat{\rho}^2}} \tag{1.2}$$

con $\hat{\rho}$ el coeficiente de correlación.

Cálculo de anomalías

Para las series temporales que registran periodo común inicial luego de mediados de la década del setenta, se filtra la onda anual media a las series temporales y luego la presencia de tendencia lineal en la serie es analizada a partir de la metodología de cuadrados mínimos con un 95 % de confianza. Por otra parte, puesto que García y Vargas (1998) documentaron un salto climático, como así también Jaime y Menéndez (2002) en la Cuenca del Plata en la década del setenta, se utilizó el Test de Yamamoto (Yamamoto et al., 1987) considerando ventanas de veinte años para detectar fehacientemente el salto característico de la década y obtener una mejor aproximación de las anomalías de caudal de las series temporales que tuviesen un periodo común que involucrase dicha década. Luego de analizar el test y obtener donde podría estar la ocurrencia del salto, se evaluó la presencia de tendencia lineal en la serie de datos mediante cuadrados mínimos.

Espectro simple

Para el estudio de ondas estacionales, un análisis espectral fue realizado mediante el método de Blackman Tukey (Nota Técnica N°79 de la Organización Meteorológica Mundial de Cambio Climático (OMM (1966) y Chatfield (1995)). La transformada de Blackman-Tukey transforma el dominio de tiempo de análisis a un dominio de frecuencia. Este primer acercamiento es útil para identificar ondas significativas que ocurren en las series. Sin embargo, la metodología sufre del compromiso (trade-off) entre varianza y resolución, y a veces solo es posible identificar energía significativa en ciertos intervalos del espacio

de la onda y no en ondas individuales, especialmente para ondas cortas. Sin embargo, existen algunas medidas técnicas que pueden ser utilizadas para mejorar la eficiencia del método. Es decir, para asegurar la estabilidad estadística se fija un punto de truncamiento (M), el 30 % de los datos, (Chatfield, 1995). Adicionalmente, para reducir la pérdida de potencia la ventana de Tukey-Hanning fue aplicada. En general, a través de esta metodología resulta posible analizar todas las ondas desde la más frecuente de dos meses a la menos frecuente asociada al periodo fundamental. A su vez, el procedimiento de Anderson (Anderson, 1942) fue aplicado para determinar si la persistencia ocurre en cada serie de tiempo. Es decir, si el coeficiente de correlación asociado con lag=1 (r_1) está dentro del intervalo $r_1 \in \frac{-1 \pm t_g \sqrt{n-2}}{n-1}$ la hipótesis nula no es rechazada y el modelo de espectro presenta un proceso de ruido blanco (serie aleatoria). Caso contrario, la significancia es determinada por un proceso Markoviano de ruido rojo. En la fórmula anterior el t_g correspondiente está asociado a la significancia del test de Student y n es la cantidad de datos considerados.

Wavelet

La metodología de wavelet es empleada para detectar ciertas señales que no resultan evidentes considerando la metodología de Blackman Tukey en el análisis espectral simple. Es decir, el método de Blackman-Tukey a pesar de ser eficiente para estimar la parte continua del espectro, resulta menos útil para la detección de la señal de componentes que son puramente sinusoidales o cerca de serlo (Ghil et al., 2002). En este sentido, el análisis mediante wavelet es incorporado como una herramienta más precisa para estimar el periodo de longitud de la onda.

Asumiendo que se posee una serie temporal X_n, con constante equispaciado temporal δt y $n = 0...N-1$, y que se tiene una función wavelet, $\Psi(\eta)$, que depende de un parámetro temporal no dimensional η. En este sentido se emplea Morlet wavelet, y al ser la misma compleja, presenta la posibilidad de extraer cierto tipo de información del proceso como la amplitud y la fase.

$$\Psi(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \tag{1.3}$$

con ω_0 la frecuencia no dimensional (Torrence y Ocampo, 1998). En este estudio para analizar la significancia de las ondas se consideró simular series temporales provenientes

de un modelo de ruido blanco. La simulación fue repetida cien veces (Schmidbauer y Roesch, 2018). Además, es necesario tener en cuenta que la transformada de wavelet en ciertos intervalos temporales contiene información de los puntos vecinos, y el número de vecinos influyentes dependerá del wavelet considerado y de la escala empleada. Más aún, los efectos de los bordes pueden ocurrir si la transformada es centrada cercana al principio o al final de la serie temporal (cono de influencia), (Maraun y Kurths, 2004).

Cross wavelet y wavelet coherencia

Tratar con procesos no estacionarios como los que se observan en series hidrológicas sugieren el uso de representación de las señales en el espacio tiempo-frecuencia. Para sobrellevar el problema asociado a series no estacionarias se introduce el método de wavelet coherencia (Gouhier et al., 2018). Wavelet coherencia es un wavelet cruzado, que revela áreas con gran potencia (energía) en común entre dos pares de datos. Esto significa qué tan coherente resulta la transformada del wavelet cruzado de dos series temporales en el espacio tiempo-frecuencia (Brebbia, 2013).

La transformada cruzada del espectro de wavelet puede ser definida de la siguiente manera (Ec. 1.4):

$$W_n^{xy}(s) = W_n^x(s)W_n^{y*}(s) \tag{1.4}$$

donde x e y son las series de tiempo, y $W_n^{y*}(s)$ la transformada de wavelet y^* indica la transformada compleja conjugada (Torrence y Compo, 1998).

Como el espectro cruzado de wavelet es complejo, la energia de la wavelet cruzada es definida como $W_n^{xy}(s)$. Dadas dos series temporales de igual longitud es posible computar la coherencia de wavelet cruzada para esta transformada. Esta herramienta es de suma importancia, pues dado un cierto nivel de significancia pre establecido, ondas en común en ambas series temporales pueden ser detectadas para ciertas ventanas temporales. Esta metodología es particularmente útil cuando uno espera la presencia de una onda sólo bajo ciertas circunstancias, que no necesariamente se encuentra en todo el periodo de análisis. Otra medida que se destaca del análisis es la fase, que es útil para estimar el desfasaje entre las ondas en común previamente mencionadas. Para este estudio, el procedimiento realizado por Grinsted et al. (2004) fue el que se tuvo en cuenta, donde el nivel de significancia del wavelet coherencia es estimado utilizando el método de Monte Carlo. Se genera una gran cantidad de ensamble de pares de datos subrogados con los mismos coeficientes

autoregresivos de orden uno (AR1) que los pares de datos a analizar. Para cada par se calcula el wavelet coherencia. Luego la significancia se testea considerando un test de χ^2 al 5 % de significancia.

Observaciones

Los datos de series de caudal fueron adquiridos a través de la Subsecretaría de Recursos Hídricos (SRH). Por otra parte, la Prefectura Naval Argentina registra el nivel del río en la regla hidrométrica, donde la conversión a caudal (Q[$\frac{m^3}{s}$]) (mínimo, máximo o medio) se realiza por medio de lo que se conoce como curva h-Q o Ley h-Q. Es decir, los datos de caudal son estimados a partir de curvas hidrométricas h-Q mediante mediciones de altura del río.

1.4. Cuenca del Plata

Área de estudio

La cuenca del Plata está localizada en el sur de Sudamérica, entre los 14-37°S y 43-67°O. Esta cuenca se expande sobre un vasto territorio geográfico (3,6 millones de km^2) comprendiendo cinco países: Argentina, Brasil, Uruguay, Bolivia y Paraguay. La mayor parte de su extensión se encuentra en Argentina y Brasil, 30 % y 46 % respectivamente. Debido a su gran dimensión horizontal LPB es una fuente significativa de agua para la agricultura, ganadería, hidroelectricidad y navegación. De esta manera, LPB es considerada la quinta cuenca más importante en el mundo.

En particular, este análisis se centró en el río Paraná y el río Uruguay. El primero presenta irregularidades en su curso, el trayecto antes de la descarga del río Paraguay es en pendiente, y luego se convierte en una llanura (Barros et al., 2006), destacando variaciones fisiográficas en la sección del río. Por otro lado, el río Uruguay presenta varios tributarios, una pendiente pronunciada, pero no variaciones de altura a lo largo de su curso. Asimismo, el río Uruguay presenta un caudal medio anual cinco veces menor que el observado en el río Paraná.

Por otra parte, LPB debido a su gran extensión se ve afectada por diferentes regímenes climáticos. Sin embargo, para el estudio se ha considerado el área de interés del Litoral Argentino y sus alrededores. En esta área en particular de la cuenca, los máximos de precipitación ocurren durante las estaciones de transición (máximo absoluto en otoño) con un mínimo marcado en invierno (sistemas frontales), y uno menos importante en verano (sistemas convectivos de mesoescala y pasajes de sistemas frontales), Barros et al. (2006) y Nesbitt et al. (2006).

1.4.1. Río Paraná

Datos

Los datos de caudal mensual de las estaciones de medición Corrientes, Túnel Subfluvial y Timbúes fueron empleados para el estudio hidrológico climático del río Paraná (Figura

1.1 y Tabla I).

Tabla I: *Ubicación, periodo y porcentaje de datos faltantes de las estaciones de medición del río Paraná.*

Cuenca del río Paraná				
Estación	Latitud	Longitud	Periodo	% Faltante
Corrientes	27°27′35″	58°49′60″	1904/01/01 – 2016/12/01	0 %
Túnel Subfluvial	31°43′11″	60°31′03″	1904/09/01 – 2016/12/01	0 %
Timbúes	32°39′30″	60°43′50″	1905/09/01 – 2016/12/01	0,54 %

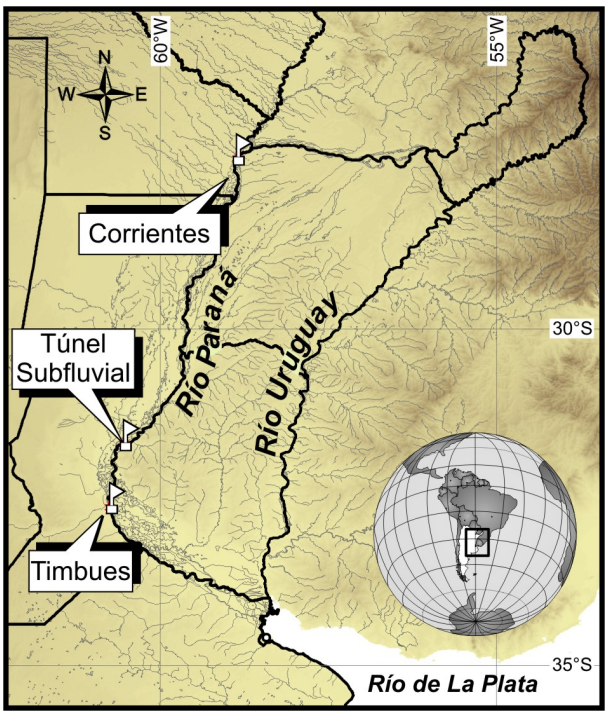


Figura 1.1: Ubicación de las estaciones de medición del río Paraná.

Resultados

En base a las estaciones de medición consideradas, el periodo en común resultó ser entre el 01/1906 y 12/2016. En la Figura 1.2 se muestra la serie temporal correspondiente a cada estación en el periodo de tiempo mencionado.

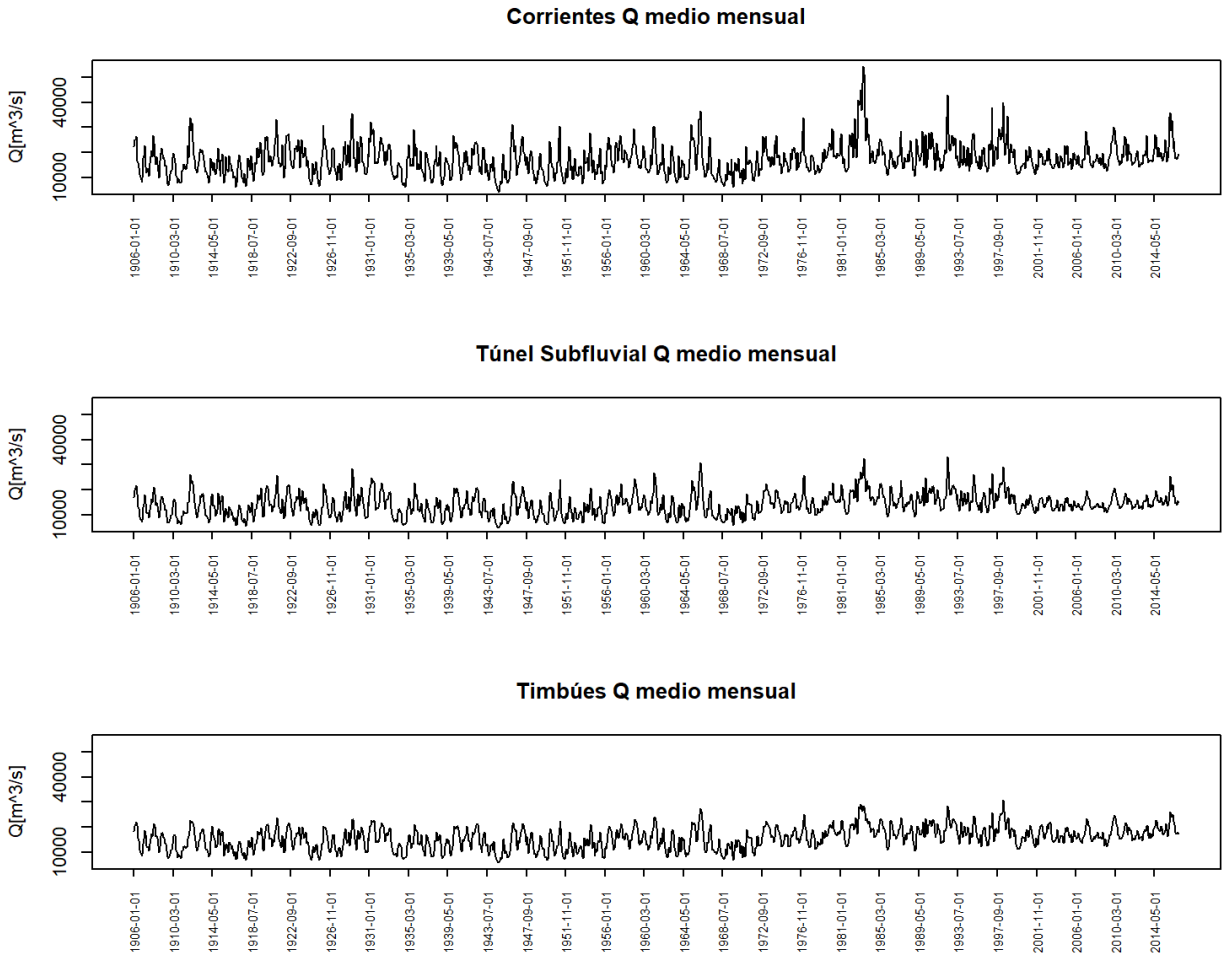


Figura 1.2: Serie de tiempo del caudal medio mensual (Q $[\frac{m^3}{s}]$) de las estaciones del río Paraná: superior Corrientes, medio Túnel Subfluvial e inferior Timbúes.

Primeramente, se realizó una comparación de las ondas medias anuales entre las estaciones de medición. También, se analizaron los valores extremos de cada serie temporal, teniendo en cuenta el desvío estándar. En la Figura 1.3 se presentan las ondas medias anuales de cada estación considerada del río Paraná. En la estación Corrientes, el máximo medio mensual de caudal estaría presente alrededor del mes de febrero, mientras que las estaciones restantes exhibieron la ocurrencia al mes siguiente. El valor mínimo de caudal

estaría ocurriendo al finalizar los meses fríos, alrededor del mes de septiembre para las tres estaciones de medición.

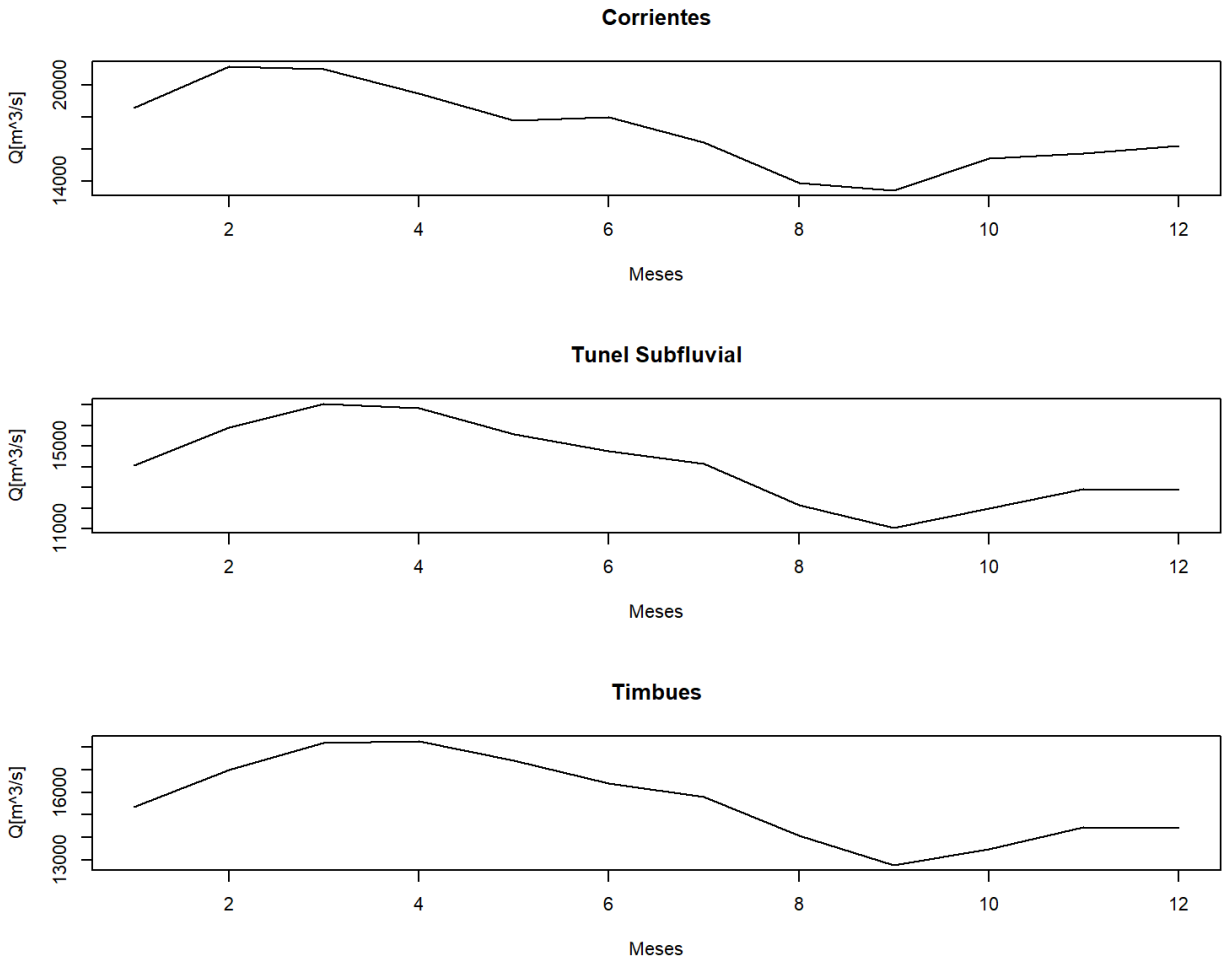


Figura 1.3: Ondas medias anuales del Q $[\frac{m^3}{s}]$ en las estaciones del río Paraná: superior Corrientes, medio Túnel Subfluvial e inferior Timbúes.

En la Tabla II se presentan los valores medios mensuales para cada estación junto con los desvíos estándar como medida de dispersión de los datos, más aún se exhiben los máximos y mínimos absolutos mensuales con su respectivo año de ocurrencia. En lo que refiere a los valores extremos a partir del caudal mensual podemos detallar la siguiente introducción exploratoria. En la estación Corrientes se destacó el mes de julio del año 1983 por ser el que mayor diferencia presentó con respecto a su valor medio (superior a seis desvíos). Para la estación Túnel Subfluvial, tanto junio como julio presentaron una distancia superior a los cuatro desvíos medido respecto de su valor medio en los años 1992 y 1983, respectivamente. En cuanto a la estación Timbúes, los valores medios mensuales

más alejados se correspondieron con tres desvíos estándar desde su valor medio para los meses de enero y febrero, ambos casos para el año 1983. Los valores mínimos si tenemos en cuenta un único valor medio mensual para todo el periodo de análisis solo se encontraron alejados respecto a su valor medio menos de dos desvíos para todas las estaciones con la excepción de Timbúes. En este último caso hay valores mensuales absolutos que superaron los dos desvíos estándar, algunos de los cuales se corresponden con el año 1944.

Tabla II: *Valores correspondientes a la onda media anual de caudal, desvíos estándar (sd), valor máximo y mínimo junto con su ocurrencia temporal, a) Corrientes, b) Túnel Subfluvial y c) Timbúes. Unidades:* $[\frac{m^3}{s}]$.

a) Estación Corrientes				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	18560,7	5778,9	38794/1983	7444/1945
Febrero	21148,5	5819,7	40782/1983	10842/1953
Marzo	20944,8	5832,0	45117/1983	10527/1915
Abril	19454,9	4987,9	40025/1998	10982/1968
Mayo	17796,4	5570,0	47790/1983	9844/1968
Junio	18001,1	6351,4	54468/1983	7955/1944
Julio	16422,8	5820,0	51307/1983	6706/1944
Agosto	13915,1	4727,4	34098/1983	5348/1944
Septiembre	13433,4	4789,7	28379/1989	4618/1944
Octubre	15423,1	5734,4	34456/1998	4092/1944
Noviembre	15747,5	4980,9	29621/1997	6136/1916
Diciembre	16237,7	5891,7	40956/1982	6562/1917

b) Estación Túnel Subfluvial				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	14064,7	4065,9	26079/1912	5995/1945
Febrero	15922,9	3881,2	27397/1966	8199/1953
Marzo	17076,1	4296,1	30938/1966	8537/1953
Abril	16845,9	3889,7	27164/1966	9618/1925
Mayo	15607,3	4027,0	29046/1998	8542/1944
Junio	14774,3	4316,8	33050/1992	6821/1944
Julio	14149,7	4452,6	32465/1983	6063/1944
Agosto	12145,7	3858,4	24610/1983	5260/1944
Septiembre	11040,1	3480,6	19785/1982	4795/1944
Octubre	11997,2	3785,2	23403/1983	4660/1944
Noviembre	12942,6	4000,3	23236/1983	5572/1944
Diciembre	12896,9	3992,7	22395/1982	5565/1917

c) Estación Timbúes				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	15338,2	4141,3	28240/1983	7040/1945
Febrero	16980,7	3683,5	27564/1983	9298/1953
Marzo	18187,5	3713,7	29078/1983	9872/1925
Abril	18285,7	3431,6	27238/1983	10540/1925
Mayo	17401,4	3927,3	30921/1998	9678/1968
Junio	16411,5	4084,1	28458/1992	8003/1944
Julio	15809,0	4238,7	28246/1983	7207/1944
Agosto	14088,4	4261,1	25801/1983	6369/1944
Septiembre	12761,4	3938,7	22635/1983	5873/1944
Octubre	13479,1	3883,7	22902/1998	5974/1944
Noviembre	14467,8	4141,5	23742/1998	6548/1944
Diciembre	14413,2	4227,5	24409/1997	6885/1917

Por otra parte se calcularon las correlaciones entre pares de estaciones. En la Figura 1.4 se muestran las magnitudes de las correlaciones entre las estaciones de medición teniendo en cuenta el coeficiente de correlación de *Spearman*. A partir de la misma se observa la relación notoria entre las tres estaciones con coeficientes cercanos a uno (significativos y positivos), destacando la disminución de dicho coeficiente al aumentar la distancia entre ellas si consideramos a Timbúes como estación de referencia.

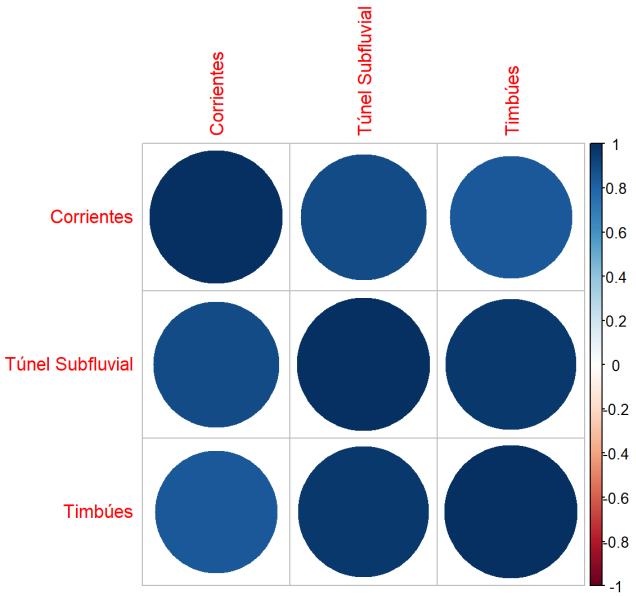


Figura 1.4: Correlación entre las estaciones de medición mediante el coeficiente de correlación de Spearman.

Previamente a obtener los valores de anomalías mensuales, debido a la bibliografía consultada (Berbery y Barros (2002), García y Vargas (1998)) donde destacaron cambios abruptos en el valor medio del caudal durante la década del setenta, se procedió a realizar el test de Yamamoto encontrando el año del salto entre 1974/75 para las estaciones del río Paraná. Por lo que se estimaron los valores medios para los periodos 1906-1974 y 1975-2016, los desvíos estándar, y los valores máximos y mínimos absolutos. Con el fin de que al realizar el cálculo de los valores de anomalías se introduzca el menor ruido a las series temporales.

En la Tabla III se presenta la información para el primer periodo y la Tabla IV para el segundo periodo propuesto. A partir de las mismas se observa al igual que en Berbery y Barros (2002) y otros autores, un incremento en la onda media anual para el segundo periodo entre el 20 % y el 22 % para las estaciones de Corrientes y Túnel Subfluvial. Sin embargo, para la estación Timbúes el incremento fue de un 30 %, mayor que en las estaciones anteriores.

En lo que refiere a los valores extremos, para la estación Corrientes se destacó para el periodo $1906-1974$ que los valores máximos más alejados de su respectiva media mensual se correspondían con los meses de agosto, septiembre y octubre, los dos primeros ocurrieron en el año 1957, mientras que para el tercero en 1935. En estos casos, el apartamiento fue superior a tres desvíos estándar. De acuerdo con documentación correspondiente al Instituto Nacional del Agua (INA), estos eventos estuvieron asociados a eventos de crecida en el río Paraná (Informe INA, 2003). En cuanto a los valores mínimos, para el primer periodo no hay eventos que destacar ya que teniendo en cuenta el desvío como apartamiento del valor medio ningún mes supera los dos desvíos. Para el segundo periodo, mayo, junio y julio del año 1983 presentaron el mayor apartamiento (superior a seis desvíos) respecto de su valor medio, siendo el más extremo el mes de julio (al igual que lo encontrado al considerar un único periodo de tiempo). Por otra parte, la serie de valores de la estación presentó meses con menor caudal de lo normal (superior a dos desvíos) en mayo y julio de los años 1972 y 1975, respectivamente.

En el caso de la estación Túnel Subfluvial en el primer periodo: enero/1912, febrero/1966, septiembre/1957 y octubre/1957 fueron los que resultaron con un mayor apartamiento, superiores a tres desvíos. Nuevamente estos meses y años fueron reportados por el INA como eventos de crecida en el río Paraná. Con respecto, al segundo periodo temporal julio/1983 se destaca con un apartamiento superior a cuatros desvíos respecto de su valor medio. Por otro lado, los valores mínimos absolutos tanto en el primer intervalo temporal como en el segundo no presentaron valores destacables. Para la estación Timbúes en el intervalo $1906-1974$ al igual que las estaciones previas, los meses de crecida septiembre y octubre de 1957 fueron los que exhibieron el mayor apartamiento, en cuanto al segundo intervalo de tiempo el periodo mayo/1998 y julio/1983 presentaron el mayor apartamiento, siendo distinto a lo encontrado anteriormente al considerar un único intervalo temporal.

Tabla III y IV: *Valores correspondientes a la onda media anual de caudal, desvíos estándar (sd), valor máximo y mínimo junto con su ocurrencia temporal para el periodo 1906-1974 y 1975-2016, a) Corrientes, b) Túnel Subfluvial y c) Timbúes. Unidades:* $[\frac{m^3}{s}]$.

III a) Estación Corrientes 1906-1974				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	17544,5	5385,4	33840/1912	7444/1945
Febrero	20342,4	5447,2	33583/1929	10842/1953
Marzo	21115,0	5971,5	36609/1966	10527/1915
Abril	19214,5	4720,7	30607/1961	10982/1968
Mayo	16585,7	4329,8	29009/1931	9844/1944
Junio	16497,7	4999,2	29156/1931	7955/1944
Julio	14552,5	4391,1	24641/1922	6706/1944
Agosto	11777,3	3803,1	24688/1957	5348/1944
Septiembre	11167,5	3758,9	26735/1957	4618/1944
Octubre	13264,5	5019,9	29134/1935	4092/1944
Noviembre	13886,6	4515,9	24951/1923	6136/1916
Diciembre	14395,6	5187,1	30333/1911	6562/1917

IV a) Estación Corrientes 1975-2016				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	20064,6	6196,3	38794/1983	10965/1986
Febrero	22389,6	6349,7	40782/1983	12599/2000
Marzo	20765,9	5935,9	45117/1983	13778/1990
Abril	19770,4	5494,3	40025/1998	13076/1978
Mayo	19751,1	6781,6	47790/1983	11032/1978
Junio	20300,2	7637,1	54468/1983	12036/1978
Julio	19332,6	6743,1	51307/1983	12323/1975
Agosto	17271,2	4251,9	34098/1983	11980/2001
Septiembre	16746,1	4320,7	28379/1989	11247/2001
Octubre	18576,6	5210,8	34456/1998	11816/1978
Noviembre	18619,1	4435,4	29621/1997	11248/1999
Diciembre	18985,5	5908,1	40956/1982	10583/1988

III b) Estación Túnel Subfluvial 1906-1974				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	13144,6	4021,6	26079/1912	5995/1945
Febrero	15238,1	3813,3	27397/1966	8199/1953
Marzo	16722,9	4493,8	30938/1966	8537/1953
Abril	16852,7	4304,1	27164/1966	9618/1925
Mayo	14908,4	4065,6	25270/1961	8542/1944
Junio	13686,2	3913,5	23557/1931	6821/1944
Julio	12955,2	4185,3	23025/1931	6063/1944
Agosto	10525,4	3287,8	18760/1922	5260/1944
Septiembre	9382,7	2812,7	18873/1957	4795/1944
Octubre	10450,8	3351,5	22407/1957	4660/1944
Noviembre	11466,8	3717,9	22822/1935	5572/1944
Diciembre	11522,9	3838,3	19730/1923	5565/1917

IV b) Estación Túnel Subfluvial 1975-2016				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	15520,1	3768,3	25478/2016	9273/1986
Febrero	17071,9	3807,5	24815/1977	10433/1986
Marzo	17693,8	3976,5	27239/1983	11577/2004
Abril	16871,5	3190,2	25629/1998	11915/2004
Mayo	16699,1	3782,6	29046/1998	9687/1978
Junio	16442,7	4426,5	33050/1992	10000/1978
Julio	16112,2	4218,1	32465/1983	10237/1978
Agosto	14807,5	3215,2	24610/1983	10903/1996
Septiembre	13762,7	2676,7	19785/1982	10282/1981
Octubre	14537,5	3028,9	23403/1983	10464/1981
Noviembre	15366,9	3212,7	23236/1983	10143/1999
Diciembre	15154,3	3152,4	22395/1982	9157/1988

III c) Estación Timbúes 1906-1974				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	13923,2	3606,8	22691/1912	7040/1945
Febrero	15804,2	3312,2	25201/1966	9298/1953
Marzo	17173,9	3520,5	27479/1966	9872/1925
Abril	17604	3497,1	26373/1966	10540/1925
Mayo	16291,2	3748,3	24064/1961	9678/1968
Junio	15008,5	3712	22708/1961	8003/1944
Julio	14338,7	3940,9	22560/1931	7207/1944
Agosto	12185,3	3631,9	20145/1922	6369/1944
Septiembre	10759,8	3010,3	19381/1957	5873/1944
Octubre	11593,2	3147,3	21300/1957	5974/1944
Noviembre	12586,6	3538,2	21079/1935	6548/1944
Diciembre	12583	3717,9	19993/1972	6885/1917

IV c) Estación Timbúes 1975-2016				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	17553,8	3993,1	28240/1983	10875/1986
Febrero	18894,9	3526,9	27564/1983	11945/1986
Marzo	19843,8	3498,3	29078/1983	13405/2000
Abril	19409,7	3085,6	27238/1983	14029/2009
Mayo	19149,9	3606	30921/1998	11324/1978
Junio	18565,4	3640,9	28458/1992	11479/1978
Julio	18224,4	3581,7	28246/1983	11956/1978
Agosto	17214,7	3275,3	25801/1983	12354/1975
Septiembre	16049,2	2962,5	22635/1983	11912/1977
Octubre	16577,2	2866,9	22902/1998	12087/1999
Noviembre	17558,3	3084,5	23742/1998	11255/1999
Diciembre	17419,8	3182	24409/1997	10655/1988

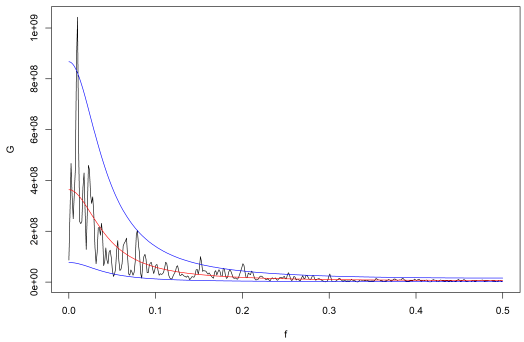
Un estudio de los valores de anomalías de caudal mensual fue llevado a cabo. Para ello,

se filtraron los valores medios mensuales de cada serie temporal restando una onda anual estimada para una primer parte del periodo analizado hasta 1974, y una segunda onda media anual estimada para el periodo entre $1975 - 2016$. Por otra parte, teniendo en cuenta la recta de regresión a través del método de cuadrados mínimos se analizó la presencia significativa de tendencia en cada una de las series de datos. Sin embargo, los ajustes a cada una de las series temporales no resultaron significativos, por lo tanto no se sustrajo dicha tendencia estimada.

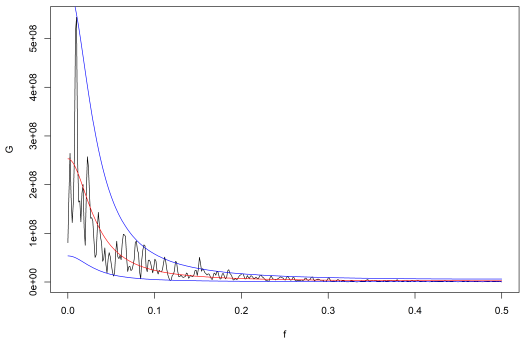
Al realizar nuevamente los cálculos de correlaciones simples entre tres pares de estaciones no se encontraron diferencias significativas con lo hallado anteriormente.

En lo que respecta al análisis univariado de cada serie temporal se llevaron a cabo dos metodologías distintas: el espectro simple (considerando el método de Blackman Tukey) y wavelet para estudiar ciclos característicos en las series de anomalías de caudal. En el caso del análisis espectral, la significancia (5 %) fue determinada por un proceso Markoviano de ruido rojo. Se observó en las tres estaciones, la significancia de ondas cortas entre los tres y cinco meses (0.3 y 0.2 de frecuencia, respectivamente) con un 95 % de confianza (Figura 1.5). A su vez ondas semianuales resultaron significativas, como así también se observó un remanente de la onda anual. Por otra parte, en la estación Corrientes se destaca a partir de la Figura 1.5a un máximo de energía significativo asociado a una onda de alrededor de los 8 años (0.01 de frecuencia). Si bien en las estaciones Túnel Subfluvial y Timbúes, la onda anterior no resultó significativa si presentó un máximo considerable de energía alrededor de ese valor de periodo.

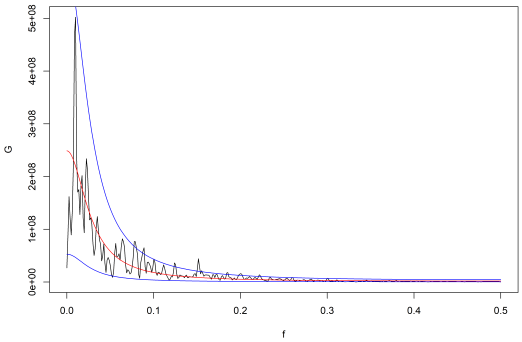
Debido a las limitaciones que presenta la metodología anterior, se empleó el método de wavelet, nuevamente la significancia se analizó con un 5 %, Figura 1.6. En cuanto a ondas largas se observó que en las tres estaciones dos centros de energía aproximadamente alrededor de la década del veinte hasta mediados de la década del treinta y entre principios de la década del setenta y finales de la del ochenta para ondas entre los 64 y 128 meses, siendo más intensos en las estaciones de Túnel Subfluvial y Timbúes para el primer intervalo temporal, en concordancia a lo hallado por Pasquini y Depetris (2007). Para las tres estaciones dichos centros de energía resultaron significativos. Asimismo, fueron destacables tres ó cuatro centros de máxima energía relativa en las ondas entre 32 y 64 meses en aproximadamente mismos periodos de temporales que los mencionados. Por otra parte, es necesario mencionar que no habría casi señal de ondas cercanas a la anual hasta



(a) Corrientes.

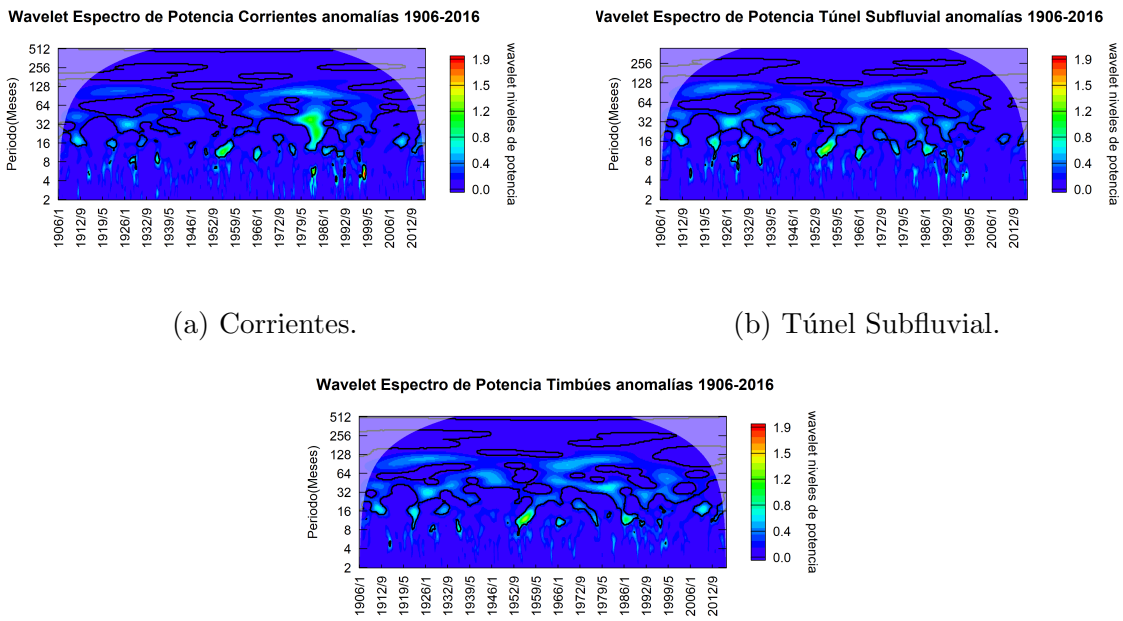


(b) Túnel Subfluvial.



(c) Timbúes.

Figura 1.5: Espectros siguiendo un modelo de ruido rojo con un 95 % de confianza. Línea negra el espectro estimado, línea roja el espectro teórico, y las bandas de confianza en líneas azules.



(a) Corrientes.

(b) Túnel Subfluvial.

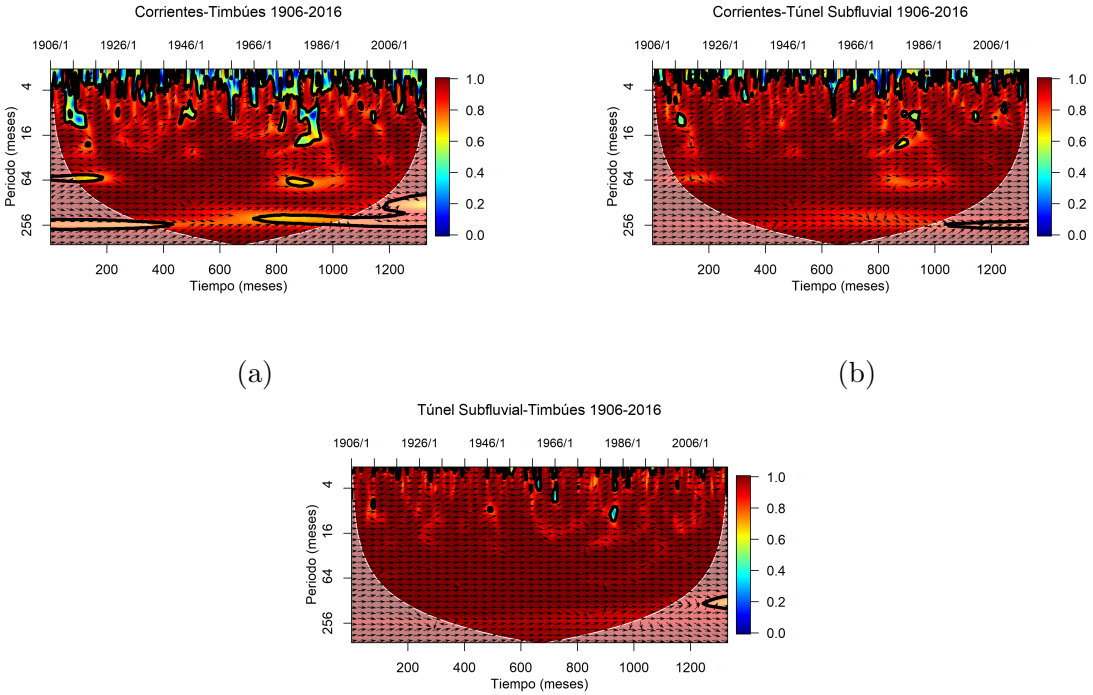
(c) Timbúes.

Figura 1.6: Morlet Wavelet con un 95 % de confianza. a) Corrientes, b) Túnel Subfluvial, c) Timbúes.

aproximadamente los años setenta (salvo un centro alrededor de fines de los cincuenta). Se observó una señal esporádica de energía en los centros alrededor de los 12 meses para el segundo periodo temporal, la cual estaría en concordancia con los espectros simples, lo que estaría indicando la presencia de un remanente de la onda anual a partir de la década del 70. Asimismo, las tres estaciones presentaron un máximo relativo de energía para las ondas entre 8 y 16 meses alrededor de fines de la década del 50 y principios del 60. Esto podría estar asociado con un cambio en la circulación entre la década del 50 y el 60 destacado por Vargas et. al (1995), generando como consecuencia, posiblemente, un cambio en la precipitación y de manera indirecta en el caudal.

Para profundizar el estudio de la relación entre las estaciones de medición se empleó la metodología de wavelet coherencia. A partir de la Figura 1.7 es posible notar que en todos los pares de estaciones todas las ondas resultaron significativas (tonos rojos) y en fase, a excepción de las ondas más rápidas menores a los cuatro meses. De todas maneras, el par Corrientes-Timbúes fue el que menor cantidad de ondas cortas no significativas presentó, siendo coherente con la mayor distancia entre estaciones, a diferencia de lo observado para el par Túnel Subfluvial-Timbúes donde prácticamente presentaron los mismos ciclos

de ondas y el desarrollo de los mismos al mismo tiempo.



(a) (b)

(c)

Figura 1.7: Wavelet coherencia con un 95 % de confianza. a) Corrientes-Timbúes, b) Corrientes-Túnel Subfluvial, c) Túnel Subfluvial-Timbúes.

1.4.2. Río Uruguay

Datos

Los caudales mensuales de las estaciones de medición de El Soberbio, Garruchos y Paso de los Libres fueron empleados para el estudio hidrológico climático del río Uruguay (Figura 1.8 y Tabla V).

Tabla V: *Ubicación, periodo y porcentaje de datos faltantes de las estaciones de medición del río Uruguay.*

Cuenca del río Uruguay				
Estación	Latitud	Longitud	Periodo	% Faltante
El Soberbio	27°17′54″	54°11′36″	1980/05/01 − 2016/12/01	1,56 %
Garruchos	28°10′32″	55°38′36″	1931/05/01 − 2016/12/01	1,29 %
Paso de los Libres	29°43′09″	57°05′02″	1909/01/01 − 2016/12/01	0 %

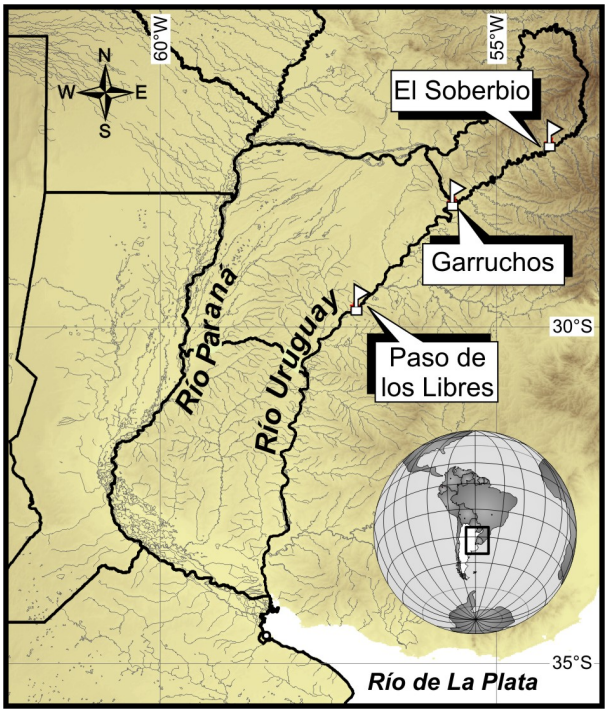


Figura 1.8: Ubicación de las estaciones de medición del río Uruguay.

Resultados

En base a las estaciones de medición consideradas, el periodo en común fue entre el 01/1981 y 12/2016. En la Figura 1.9 se exhiben las series temporales correspondientes a cada estación en el periodo de tiempo mencionado.

En lo que refiere a la estadística básica realizada se procedió a obtener la curva media

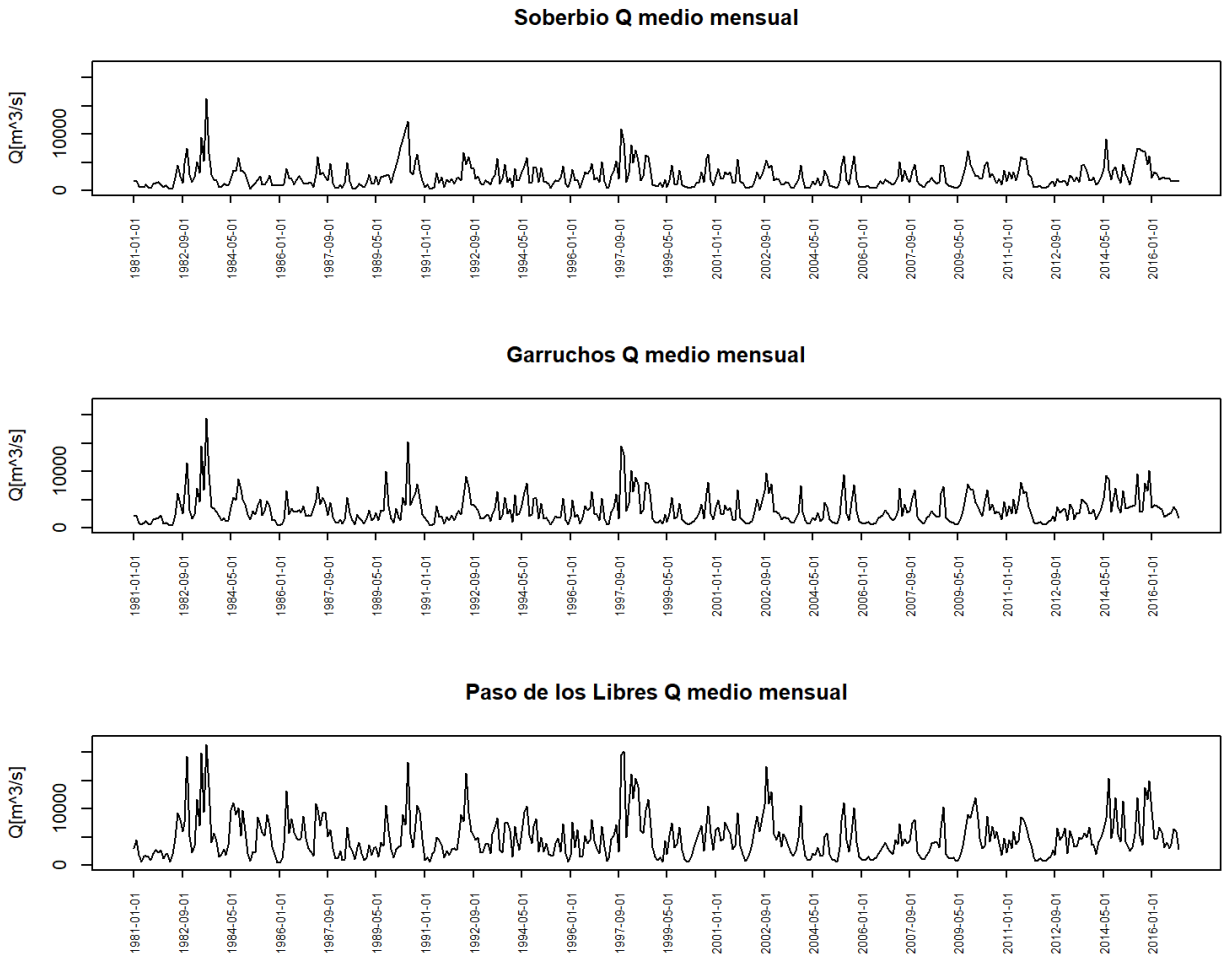


Figura 1.9: Series de tiempo del (Q $[\frac{m^3}{s}]$) medio mensual de las estaciones del rıo Uruguay: superior El Soberbio, medio Garruchos e inferior Paso de los Libres.

anual para cada estación. Además, se incluyó el desvío estándar, los máximos y mínimos absolutos de cada mes con su respectivo año. La onda media anual para cada estación se exhibe en la Figura 1.10. De la misma, se observó que en los meses fríos el caudal aumentaba alcanzando un máximo destacable en octubre, a diferencia de lo visto para las estaciones del río Paraná que presentaban una disminución del caudal en estos meses.

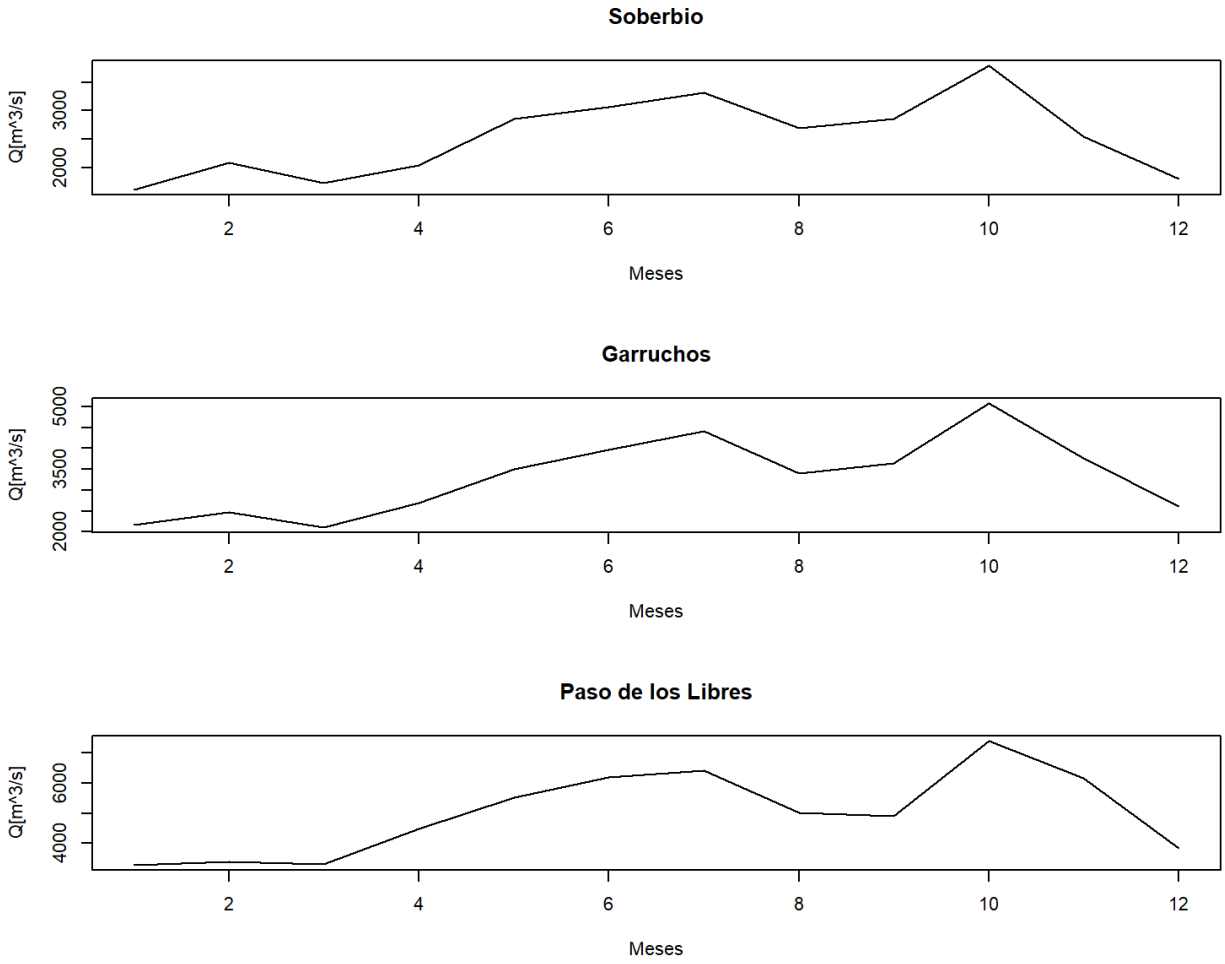


Figura 1.10: Ondas medias anuales del Q[$\frac{m^3}{s}$] en las estaciones del rıo Uruguay: superior El Soberbio, medio Garruchos e inferior Paso de los Libres.

En la Tabla VI se presentan los valores medios mensuales para cada estación junto con los desvíos estándar como medida de dispersión de los datos, más aún se exhiben los máximos y mínimos absolutos mensuales con su respectivo año de ocurrencia. En el caso de la estación Soberbio, los valores más apartados de su respectivo valor medio mensual se correspondieron con el mes de julio de 1983 coincidente con las estaciones del río Paraná. Mientras que las estaciones de Garruchos y Paso de los Libres, el mayor apartamiento se produjo en el mes de febrero de 1998 (superior a cuatro desvíos).

Por otra parte, en cuanto a las correlaciones simples de a pares entre estaciones se presentan en la Figura 1.11 teniendo en cuenta el coeficiente de correlación de Spearman. A partir de la misma se observó la relación notoria entre las tres estaciones con coeficientes cercanos a uno (significativos y positivos), los mismos disminuyen al aumentar la distan-

cia entre las estaciones si considerásemos Paso de los Libres como estación de referencia.

Tabla VI: *Valores correspondientes a la onda media anual de caudal, desvíos (sd), valor máximo y mínimo junto con su ocurrencia temporal, a) El Soberbio, b) Garruchos y c) Paso de los Libres. Unidades:* $[\frac{m^3}{s}]$.

Estación El Soberbio				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	1613,75	1087,99	4652,46/2015	374,5/1985
Febrero	2092,88	1699,06	8144,55/1998	506,68/2004
Marzo	1729,04	1495,43	7633,71/1990	370,85/1991
Abril	2047,56	1830,88	9179,14/1990	373,84/1982
Mayo	2866,04	2493,85	10724,57/1990	433,24/1982
Junio	3070,93	2481,93	12270/1990	504,50/2006
Julio	3314,13	2797,76	16220/1983	489,05/2006
Agosto	2690,98	1858,69	7327,56/2015	378/1988
Septiembre	2851,75	1849,67	7101,59/2015	533,1/2003
Octubre	3796,02	2062,75	10904/1997	891,3/1985
Noviembre	2551,57	1734,93	84651997	6887,42/1985
Diciembre	1807,10	1246,64	6106,20/2015	522,3/1987

Estación Garruchos				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	2167,75	1323,33	6639,22/2015	562,645/1986
Febrero	2476,41	1892,75	10127,57/1998	635,045/2000
Marzo	2100,45	1501,03	7105/1983	499,1/1991
Abril	2692,53	1942,11	8936/1998	434,8/1982
Mayo	3503,37	2891,39	14542/1983	572/1982
Junio	3968,82	3157,42	15246/1990	819,387/2006
Julio	4403,83	3406,62	19428/1983	744,2/1981
Agosto	3403,87	2235,82	9507/1983	578,3/1981
Septiembre	3643,22	2251,40	9983/1989	837,4/1991
Octubre	5090,19	2640,97	14429/1997	1457/1985
Noviembre	3760,93	2716,48	12996/1997	1370/1991
Diciembre	2602,96	2066,24	10131,25/2015	571,2/1985

Estación Paso de los Libres				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	3278,18	2921,63	11354,1/2015	425,045/1986
Febrero	3395,87	2844,84	16058,55/1998	588,855/2000
Marzo	3309,46	2562,45	11759,99/1998	621/1991
Abril	4491,10	3594,04	15444/1998	637/1982
Mayo	5518,99	4062,47	19827/1983	749,1964/2009
Junio	6189,18	4683,31	18252/1990	1097,025/2006
Julio	6405,45	3958,08	21344/1983	1408,677/2006
Agosto	5015,24	2955,97	13909/1983	1000/1981
Septiembre	4912,39	2625,46	11691/1998	1458/1991
Octubre	7400,92	3829,95	19500/1997	2574/1991
Noviembre	6166,44	4350,03	20136/1997	1807,808/2010
Diciembre	3847,09	3502,70	15007,52/2015	535/1985

La onda anual fue filtrada para el periodo considerado en las tres estaciones del río Uruguay. Más aún se analizó la presencia significativa de tendencia lineal en las series de tiempo. Sin embargo, en las tres series del río Uruguay no se rechazó la hipótesis nula (la serie no presenta tendencia lineal) con un 5 % de significancia.

Luego, se repitieron los cálculos de pares de correlaciones simples entre las tres estaciones

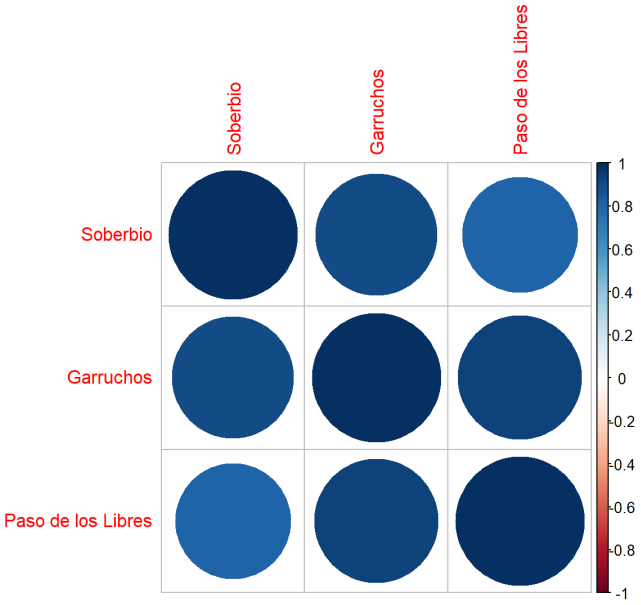
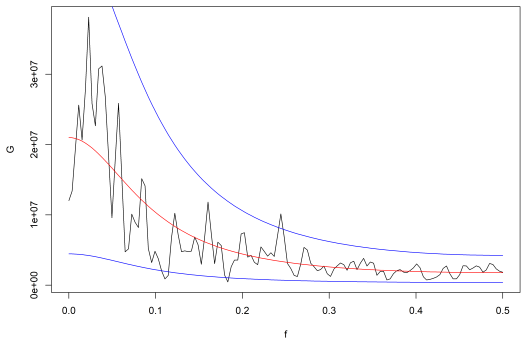


Figura 1.11: Correlación entre las series de tiempo de las estaciones de medición mediante el coeficiente de correlación de Spearman.

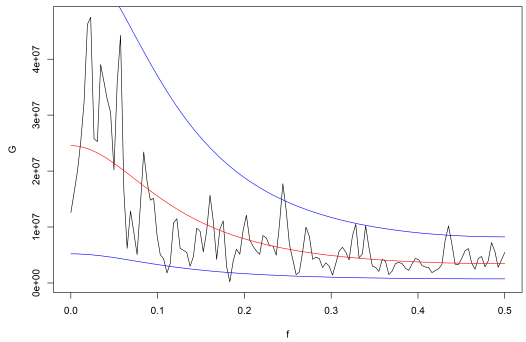
no presentando grandes diferencias a lo observado anteriormente. Es decir, las correlaciones continuaron siendo positivas y significativas, y teniendo en cuenta a Paso de los Libres como estación de referencia, la correlación disminuyó levemente al aumentar la distancia geográfica entre las estaciones.

En cuando al análisis de los ciclos característicos en las series de datos mediante la metodología por espectro simple (Figura 1.12) se encontró que las tres series temporales presentaron ondas cortas significativas entre 2 y 4 meses (0.5 y 0.25 de frecuencia, respectivamente). En particular, los resultados fueron similares a los encontrados en Meis y Llano (2018), presentando leves diferencias que podrían estar relacionadas con el periodo de dependencia de las series. Por otra parte, también fue posible observar los máximos de energía alrededor de los 8 años (alrededor de 0.01 de frecuencia) que no fueron significativos pero si fueron detectados en las estaciones del río Paraná.

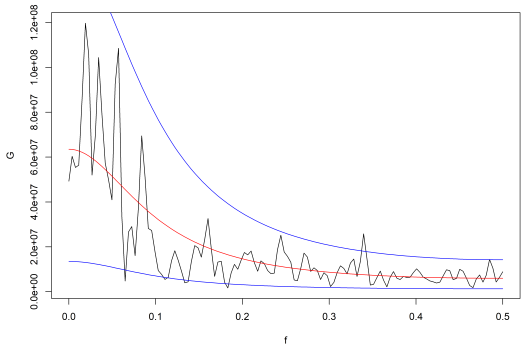
Profundizando el estudio mediante wavelet se reforzó lo visto en el análisis espectral, ondas cortas entre tres y cuatro meses fueron encontradas esporádicamente, Figura 1.13. Asimismo fue posible observar las mismas ondas significativas entre los dos y cinco años que fueron detectadas en las estaciones del río Paraná en el mismo intervalo temporal, alrededor de la década del noventa.



(a) El Soberbio.

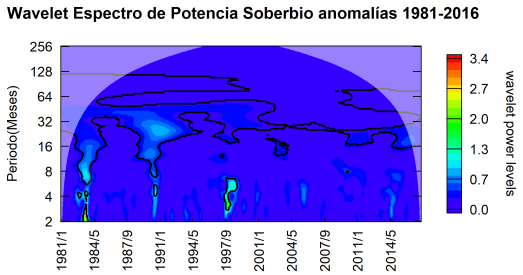


(b) Garruchos.

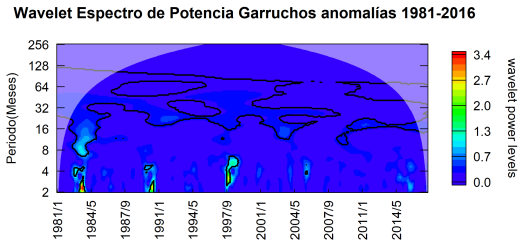


(c) Paso de los Libres.

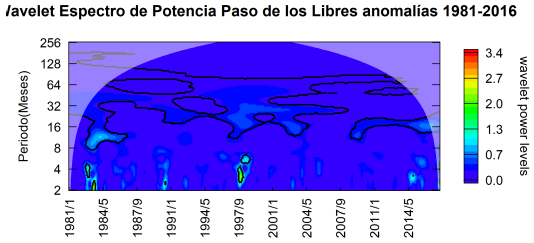
Figura 1.12: Espectros siguiendo un modelo de ruido rojo con un 95 % de confianza. Línea negra el espectro estimado, línea roja el espectro teórico, y las bandas de confianza en líneas azules.



(a) El Soberbio.



(b) Garruchos.

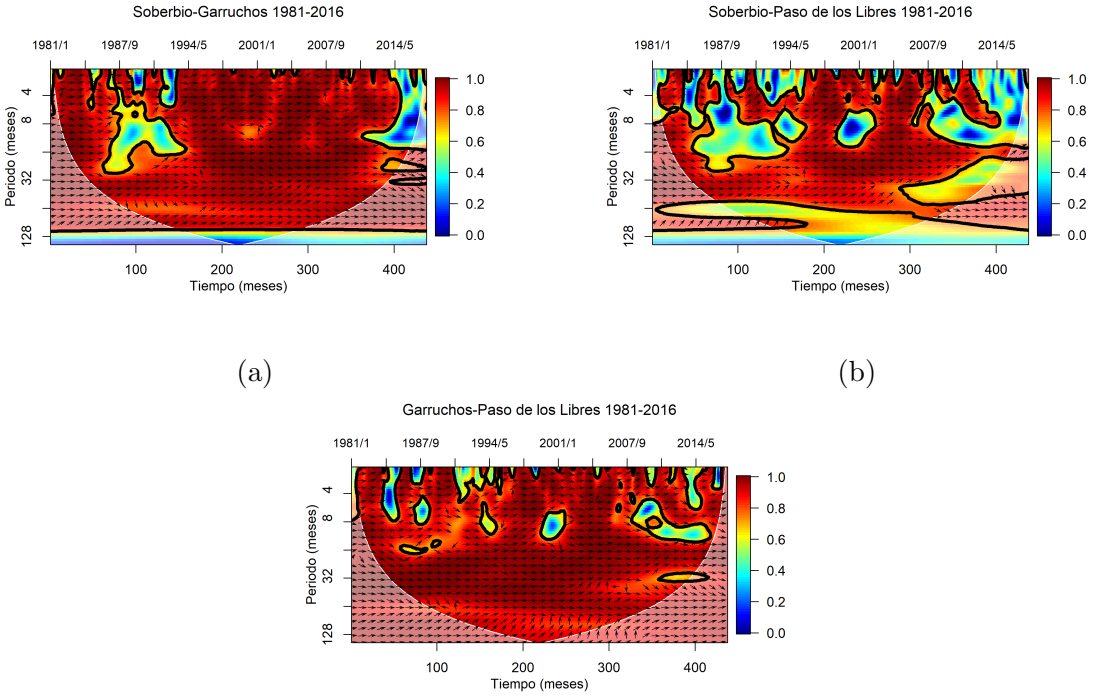


(c) Paso de los Libres.

Figura 1.13: Morlet Wavelet con un 95 % de confianza. a) El Soberbio, b) Garruchos, c) Paso de los Libres.

Si bien las ondas anteriores resultaron significativas, las mismas presentaron menor energía que en las estaciones del Paraná.

Nuevamente la metodología de wavelet coherencia fue utilizada para el estudio simultáneo de ondas características en pares de estaciones. Se destaca a partir de la Figura 1.14 que a grandes rasgos la significancia disminuye con la distancia, es decir, hay menor cantidad de ondas características que se dan en el par de estación considerado. Lo anterior fue más notable en ondas cortas, menores a un año. Por otra parte los pares de estaciones que se encontraron más próximos entre sí presentaron significancia prácticamente en todas las ondas largas, donde en promedio ocurren en fase en el par estudiado.



(a) (b)

(c)

Figura 1.14: Wavelet coherencia con un 95 % de confianza. a) El Soberbio-Garruchos, b) El Soberbio-Paso de los Libres, c) Garruchos-Paso de los Libres.

1.5. Cuenca del Atuel

1.5.1. Río Atuel

Área de estudio

La región central-oeste de la República Argentina conocida como cuyana, comprende las provincias de Mendoza, San Juan y San Luis. Esta región del país es considerada de suma importancia por las actividades agrícolas que se desarrollan en la misma, así como también por su capacidad de generar energía hidroeléctrica. Este es el motivo por el cual la realización de estudios hidroclimáticos en la región resultan de gran importancia pues colaboran con el desarrollo adecuado de dichas actividades socio-económicas.

Varios ríos se encuentran ubicados en esta región, donde el volumen de los mismos depende fundamentalmente de la precipitación nívea en los meses fríos sobre las altas cumbres y del posterior derretimiento en los meses cálidos de la nieve y el hielo acumulado (Compagnucci y Araneo 2007). Dentro de este análisis, el caudal del río Atuel (ubicado principalmente en la provincia de Mendoza) fue analizado.

Datos

Las estaciones de medición El Sosneado, Loma Negra, La Angostura y Carmensa fueron consideradas para el estudio hidrológico climático del río Atuel (Figura 1.15 y Tabla VII).

Tabla VII: *Ubicación, periodo y porcentaje de datos faltantes de las estaciones de medición del río Atuel.*

Cuenca del río Atuel				
Estación	Latitud	Longitud	Periodo	% Faltante
Carmensa	35°11′06″	67°36′11″	1985/07/01−2016/12/01	1,56 %
El Sosneado	35°05′02″	69°14′55″	1972/07/01−2016/12/01	1,29 %
Loma Negra	35°15′52″	68°52′25″	1980/11/01−2016/12/01	0 %
La Angostura	35°05′56″	67°43′33″	1906/07/01−2016/12/01	0,22 %

Resultados

En base a las estaciones de medición consideradas, el periodo en común resultó ser entre el 01/1986 y 12/2016. En la Figura 1.16 se exhiben las series temporales correspondientes

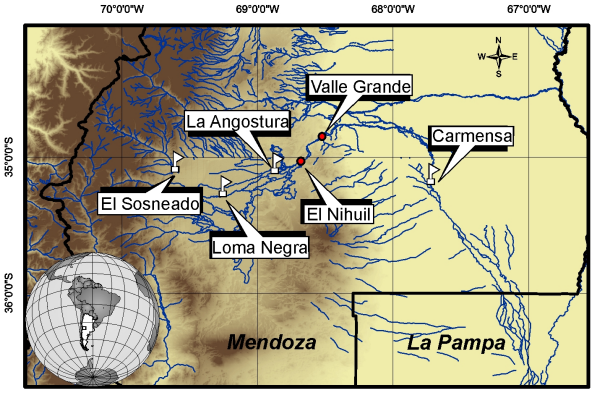


Figura 1.15: Ubicación de las estaciones de medición del río Atuel, en conjunto con las presas Valle Grande y el Nihuil.

a cada estación en el periodo de tiempo mencionado.

En lo que refiere a la estadística básica, se procedió a obtener la onda media anual para cada estación, incluyendo el desvío para cada mes y los máximos y mínimos absolutos de cada mes con su respectivo año. En las Figura 1.17 se presentan las ondas medias anuales observando que tres de las cuatro estaciones (El Sosneado, Loma Negra y La Angostura) presentaron un comportamiento semejante, es decir, valores mínimos en los meses fríos (levemente superiores a los 20 m^3/seg) y máximos en los meses cálidos (cercanos a los 60 m^3/seg en promedio). Por otra parte, la estación Carmensa se destacó por un comportamiento disímil a las anteriores, exhibiendo prácticamente una curva media anual en contrafase. Más aún, los valores mínimos resultaron ser de una magnitud menor a las estaciones anteriores.

En la Tabla VIII se presentan los valores medios mensuales para cada estación junto con los desvíos estándar como medida de dispersión de los datos, más aún se exhiben los máximos y mínimos absolutos mensuales con su año de ocurrencia. En el caso de la estación Sosneado, los valores máximos correspondientes a los meses de mayo y noviembre se encontraron alejados del valor medio con una distancia superior a tres desvíos estándar de su respectivo mes (no coincide el año de ocurrencia). Por otra parte, con respecto a los valores mínimos, el más extremo fue el del mes de abril con una distancia superior a dos desvíos estándar respecto a su valor medio siendo el 2014 su año de ocurrencia.

Con respecto a la estación Loma Negra, los meses de febrero y junio del año 2006 y el mes de julio del año 2002 presentaron valores alejados del valor medio superior a los tres desvíos estándar. En cuanto al valor mínimo, el más alejado ocurrió en el mes

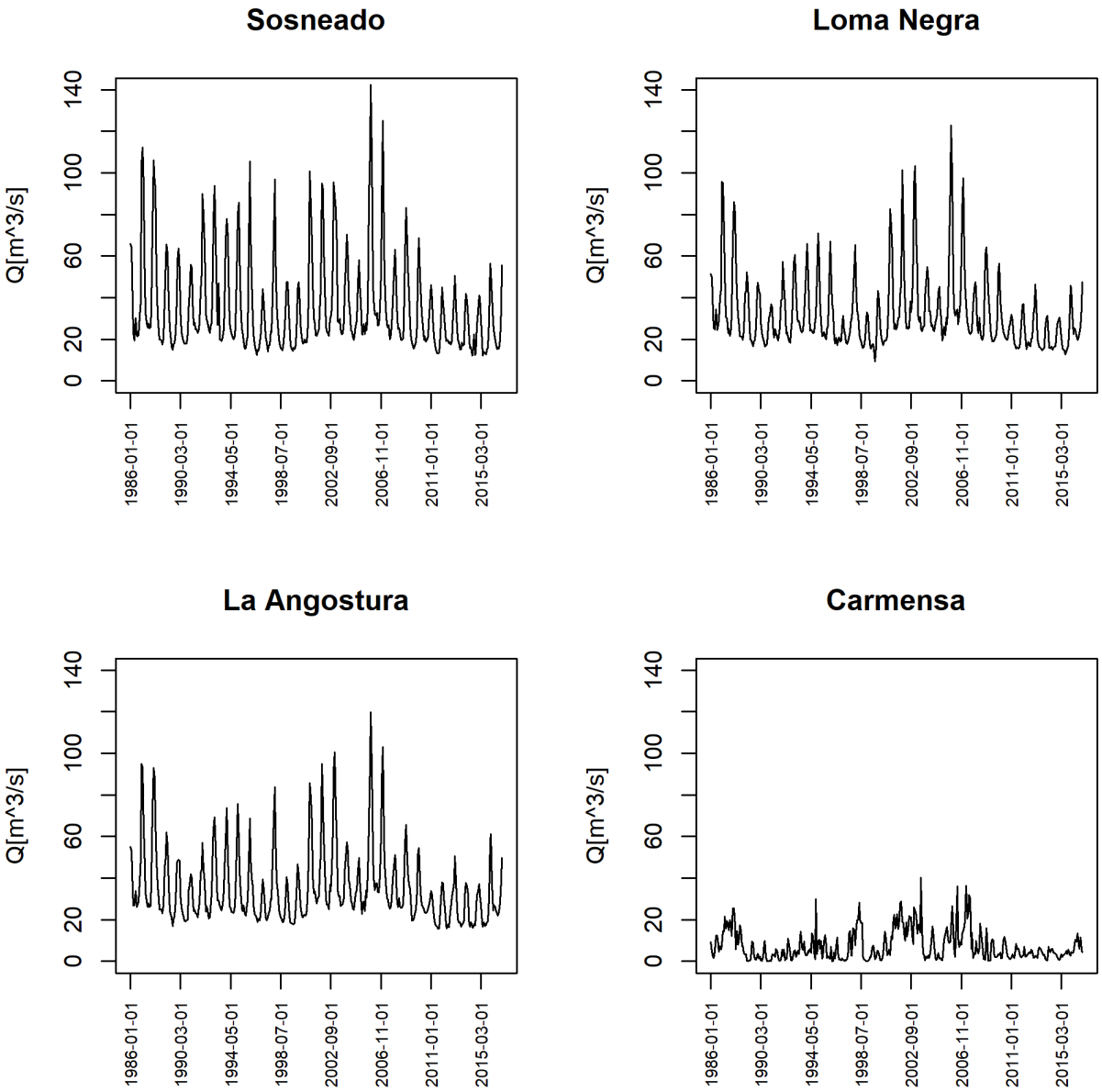


Figura 1.16: Series de tiempo del (Q[$\frac{m^3}{s}$]) medio mensual de las estaciones del río Atuel: Sosneado, Loma Negra, La Angostura y Carmensa.

de julio 2015. En la estación La Angostura, el mes de febrero, al igual que la estación anterior resultaría ser el más extremo, junto con el mes julio del mismo año, puesto que presentaron valores de caudal alejados de su respectivo valor medio alrededor de tres veces su desvío estándar. En cuanto a los valores mínimos se destaca el mes de abril del año 2014 y mayo del año 2012, en ambos casos la distancia al valor central se correspondió con valores cercanos a dos desvíos estándar. La estación Carmensa presentó varios meses que se alejaban de su respectivo valor medio, tres veces su desvío estándar (febrero/2003,

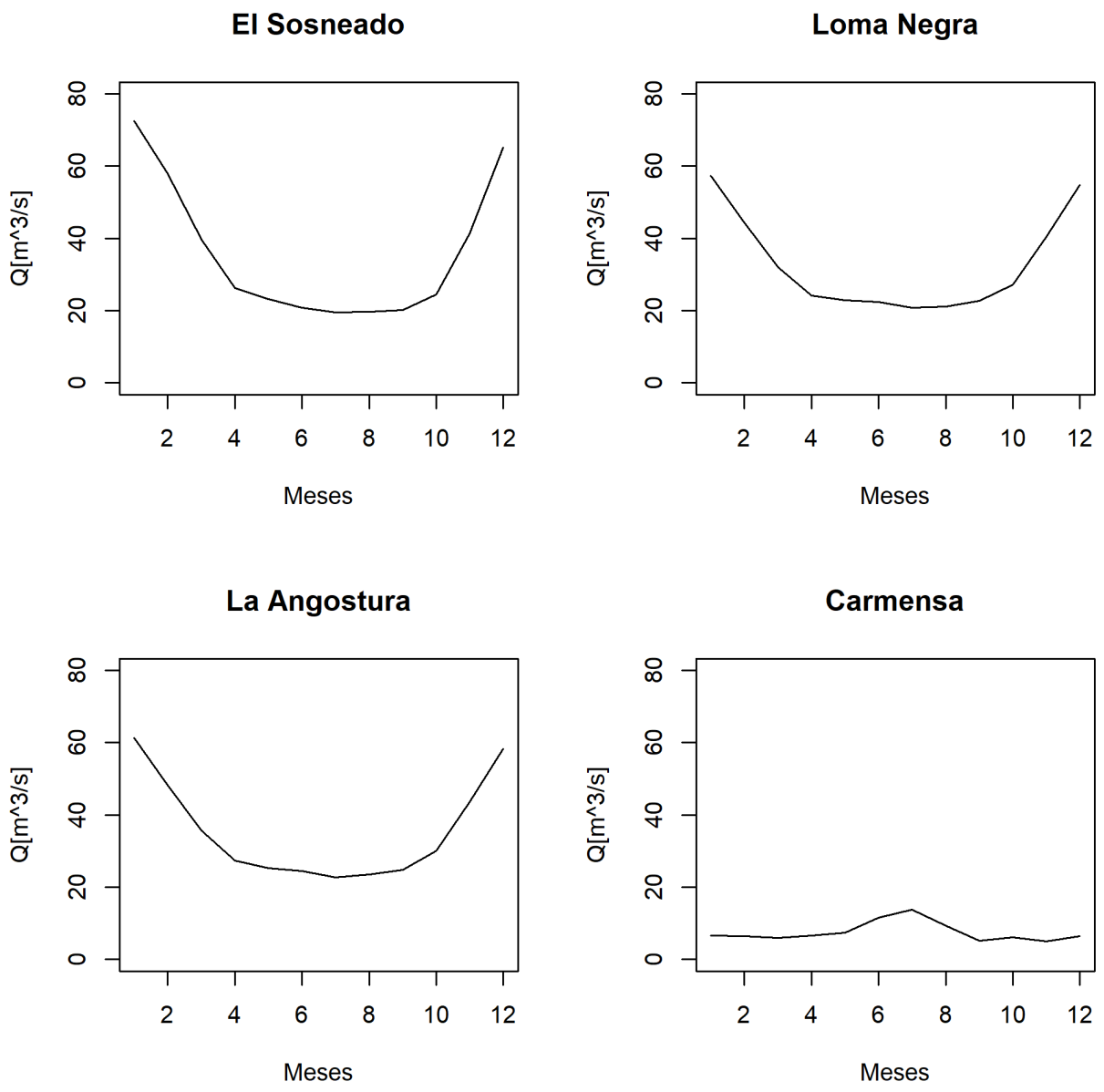


Figura 1.17: Ondas medias anuales del Q[$\frac{m^3}{s}$] en las estaciones del río Atuel: Sosneado, Loma Negra, La Angostura y Carmensa.

agosto/2007, octubre/1994, noviembre/2001). En lo que respecta a los valores mínimos podría mencionarse el del mes de julio del año 2015.

Tabla VIII: *Valores correspondientes a la onda media anual de caudal, desvíos (sd), valor máximo y mínimo junto con su ocurrencia temporal, a) El Sosneado, b) Loma Negra, c) La Angostura y d) Carmensa. Unidades:* [$\frac{m^3}{s}$].

Estación El Sosneado				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	72,41	25,74	142,33/2006	36,73/2012
Febrero	58,01	18,98	108,29/2006	30,13/2014
Marzo	39,55	10,94	65/1992	24,22/2014
Abril	26,20	5,42	39,21/2006	15,87/2014
Mayo	23,27	6,36	47/1993	12,23/2015
Junio	20,87	5,22	31,34/2006	11,99/2014
Julio	19,54	4,95	32,65/2006	12,37/1996
Agosto	19,66	4,12	29,87/2002	12,81/2015
Septiembre	20,16	4,53	30,36/2002	12,41/2014
Octubre	24,42	6,23	35,2/1987	13,25/2014
Noviembre	41,39	15,27	88,4/2005	23,45/1996
Diciembre	65,08	24,92	111,44/2005	30,73/1996

Estación Loma Negra				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	57,35	23,03	122,90/2006	30,50/2015
Febrero	44,40	16,27	98,82/2006	21,96/2014
Marzo	31,92	9,47	52,70/1987	16,16/2014
Abril	24,22	6,05	36,63/2007	15,31/2012
Mayo	22,83	4,94	31,30/2006	15,09/2015
Junio	22,34	4,88	34,50/1986	14,82/2015
Julio	20,86	4,32	34,41/2006	12,76/2015
Agosto	21,14	5,23	38,41/2002	14,57/2015
Septiembre	22,67	5,60	33,11/2002	9,44/1999
Octubre	27,18	7,39	41,27/2001	16,28/1998
Noviembre	40,34	13,24	67,80/1987	19,55/1996
Diciembre	54,77	23,01	101,54/2001	25,42/1996

Estación La Angostura				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	61,35	22,07	119,75/2006	33,91/2011
Febrero	48,24	14,41	93,03/2006	30,51/2012
Marzo	35,62	9,05	55,01/2003	20,06/2014
Abril	27,25	6,38	40,58/2006	16,56/2014
Mayo	25,28	5,03	34,24/2006	15,88/2012
Junio	24,43	5,02	37,53/2006	16,74/2014
Julio	22,75	4,63	37,57/2006	16,16/2014
Agosto	23,46	5,09	36,87/2002	15,75/2011
Septiembre	24,81	4,74	33,54/2002	15,86/2011
Octubre	29,97	7,48	44,00/2002	18,99/1998
Noviembre	43,63	12,80	73,00/1987	23,14/1998
Diciembre	58,22	21,76	96,40/2005	29,21/2010

Estación Carmensa				
Mes	$\bar{Q}$	sd	máx/ año	mín/ año
Enero	6,55	6,58	24,43/2003	0,03/1999
Febrero	6,39	6,64	26,57/2006	0,01/1999
Marzo	6,04	6,80	21,53/1987	0,13/1996
Abril	6,60	8,41	36,40/2007	0,2/1990
Mayo	7,38	6,80	28,29/1998	0,95/1997
Junio	11,58	5,15	23,27/2007	3,19/2015
Julio	13,85	8,88	40,33/2003	3,86/2015
Agosto	9,35	6,54	30,18/2007	2,07/1990
Septiembre	5,10	5,70	21,32/2002	0,261990
Octubre	6,18	7,30	29,88/1994	0,18/1991
Noviembre	5,04	6,79	28,86/2001	0,21/1990
Diciembre	6,50	7,09	26,16/2002	0,23/1998

Por otra parte se calcularon las correlaciones entre las estaciones. En la Figura 1.18 se muestran las magnitudes de las correlaciones entre las estaciones de medición. A partir de la misma se observa la notoria relación entre tres estaciones con coeficientes de a pares cercanos a uno (significativos y positivos), mientras que la relación con Carmensa resulta ser muy inferior.

Por otro lado, para el análisis respecto a los valores de anomalías de caudal mensual, se filtraron los valores medios mensuales de cada serie temporal, y se analizó para cada una de las series la presencia de tendencia lineal, mediante la recta de regresión a través del método de cuadrados mínimos. La varianza explicada por cada recta resultó ser aproximadamente un 2 %, motivo por el cual se decidió no filtrar la posible tendencia de manera

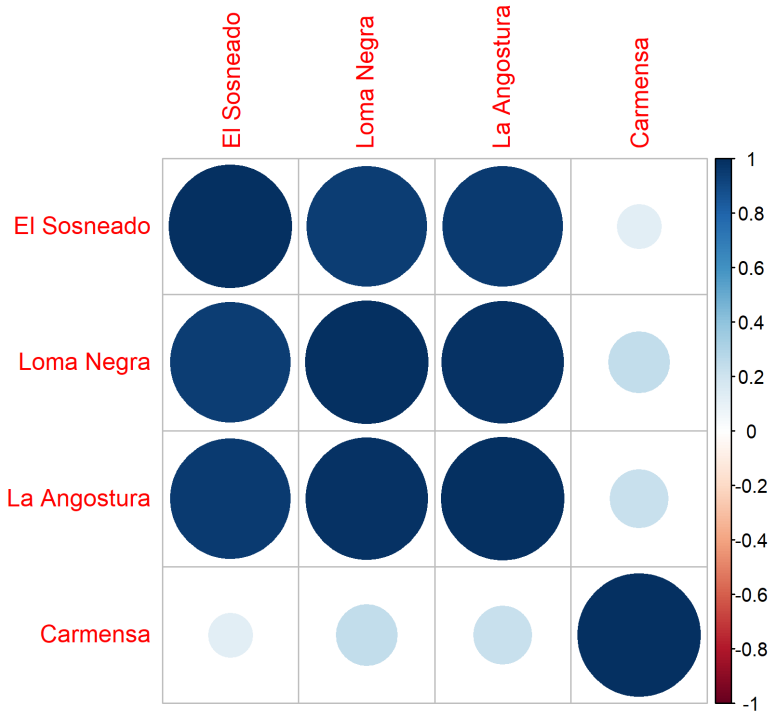
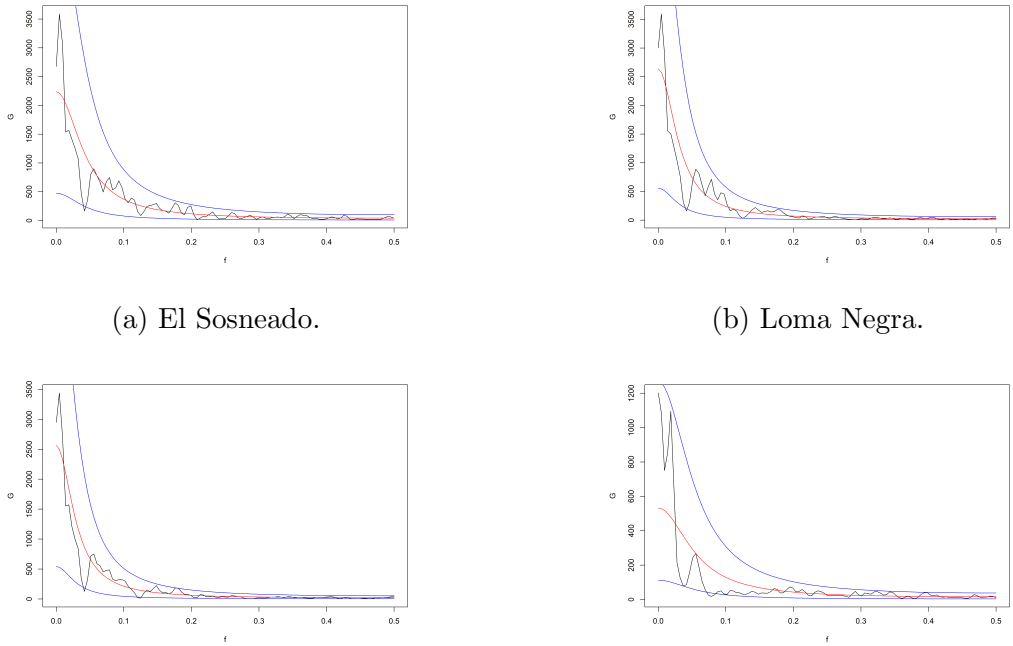


Figura 1.18: Correlación entre las series de tiempo de las estaciones de medición mediante el coeficiente de correlación de Spearman.

tal de no generar ruido en las series temporales.

Nuevamente se observó la estrecha relación que presentaron entre tres pares de estaciones al realizar correlaciones, con la salvedad que la relación de las series filtradas con Carmensa sería levemente más importante, pues hay un aumento en el valor del coeficiente de correlación.

En cuanto al análisis univariado de cada serie temporal, no se observó en el caso del análisis por medio del espectro simple ondas significativas a excepción de La Angostura que presentó un pico significativo alrededor de los 5 meses (0.2 de frecuencia). Por otra parte, si bien en el resto de las estaciones no se detectaron ondas significativas, si fue posible destacar dos máximos relativos de energía asociados a una onda con período alrededor de 18 años (0.004 de frecuencia) y una de un periodo más corto entre 8 y 10 años (0.01 y 0.008 de frecuencia, respectivamente). Los resultados previos expuestos son observables



(a) El Sosneado.

(b) Loma Negra.

(c) La Angostura.

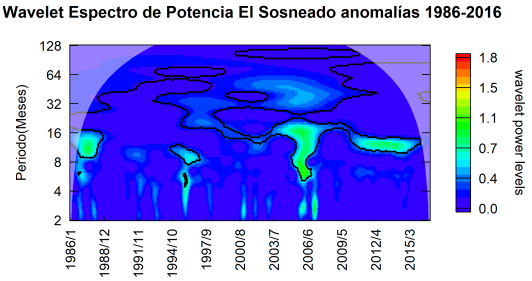
(d) Carmensa.

Figura 1.19: Espectros siguiendo un modelo de ruido rojo con un 95 % de confianza. Línea negra el espectro estimado, línea roja el espectro teórico, y las bandas de confianza en líneas azules.

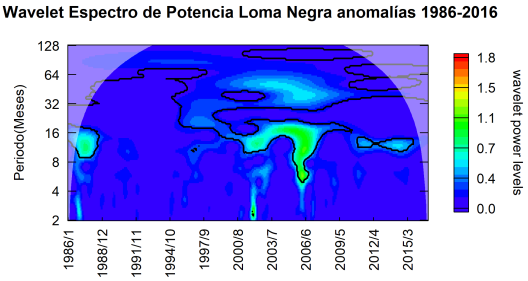
en la Figura 1.19.

Sin embargo, probablemente, debido a las limitaciones de la metodología anterior, a través del wavelet (Figura 1.20) se observó que las cuatro estaciones de medición presentaron ondas de energía significativa entre los 2 y los 5 años para el periodo temporal que abarca la primer década del siglo XXI, estas ondas podrían estar asociadas a los máximos de energía que presentaba la metodología anterior pero que de todas maneras no resultaban significativos. Además, en las estaciones de El Sosneado, Loma Negra y La Angostura se observaron ondas significativas (95 % de nivel de confianza) alrededor del año y medio para varios periodos temporales con una mayor energía alrededor de los años $2003-2006$.

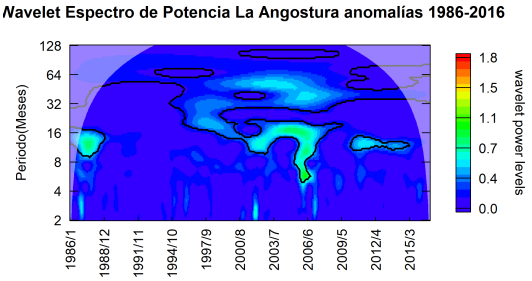
En lo que refiere a la relación entre las estaciones de El Sosneado y Loma Negra se observa en la Figura 1.21 que a fines climatológicos ondas significativas de igual magnitud afectan a ambas estaciones en fase. Por otro lado, se expone la presencia de ondas más cortas donde la energía resultó ser menor y no significativa para ambos puntos de medición. Esto es coherente con los espectros de wavelet realizados donde las ondas cortas no resultaban significativas. Además, las relaciones de coherencia entre El Sosneado y La



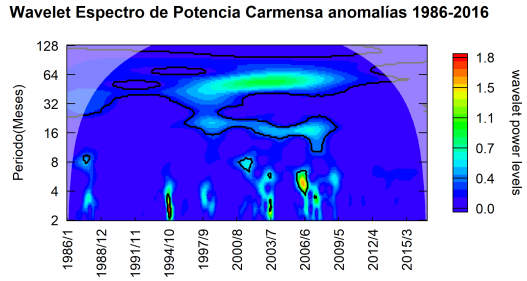
(a) El Sosneado.



(b) Loma Negra.



(c) La Angostura.



(d) Carmensa.

Figura 1.20: Morlet Wavelet con un 95 % de confianza. a) El Sosneado, b) Loma Negra, c) La Angostura, d) Carmensa.

Angostura; Loma Negra y La Angostura presentaron características similares al análisis previo (Figura 1.21). Sin embargo, como era de esperar en base a los resultados analizados hasta el momento se observó la presencia de ondas significativas en las estaciones Carmensa y La Angostura para ondas entre 64 y 128 meses en ambas estaciones ocurriendo en fase, siendo coherente con los resultados del espectro del wavelet. Es importante destacar que la presencia de dichas ondas largas asociadas a un periodo temporal entre 5 y 10 años podría ser adjudicadas al fenómeno del El Niño-Oscilación del Sur (ENSO). Estas relaciones se ampliaran en el capítulo tres.

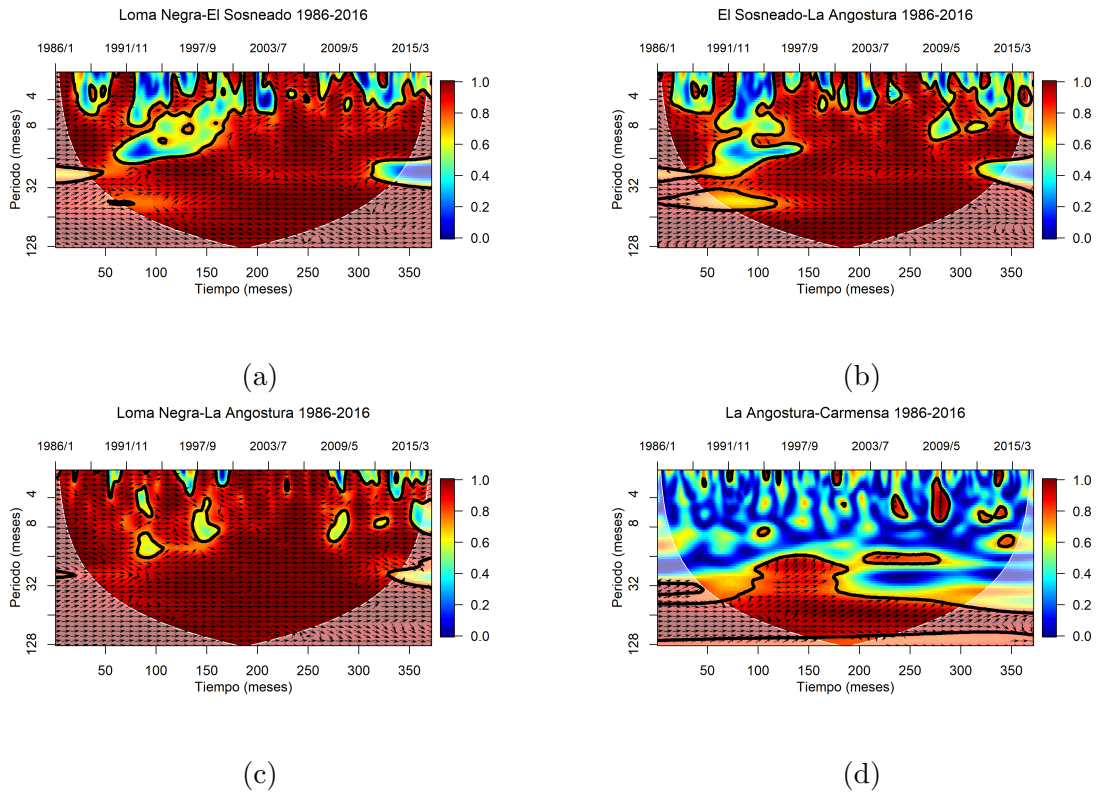


Figura 1.21: Wavelet coherencia con un 95 % de confianza. a) Loma Negra-El Sosneado, b) El Sosneado-La Angostura, c) Loma Negra-La Angostura, d) La Angostura-Carmensa.

1.6. Conclusiones

Aún no resulta del todo suficiente el entendimiento de los procesos hidrológicos en conjunto con que las componentes determinísticas de los datos observacionales son usualmente desconocidas (Sang, 2013). En este sentido, el estudio exploratorio del caudal tanto en la LPB como en la cuenca del río Atuel son necesarios para entender cómo es su interacción con el clima, para luego poder desarrollar y estudiar posibles relaciones estadísticas entre la precipitación (como principal factor climático) y el caudal.

Los ríos considerados son utilizados para la producción de energía, y todavía hay un gran potencial para la expansión de dicha actividad. Por ejemplo, en el caso del la LPB su potencial ha sido analizado por un gran número de científicos (Berbery y Barros (2002), Popescu et al. (2012), entre otros), así como también importantes proyectos de investigación como el proyecto CLARIS LPB donde científicos de varios países estudiaron en conjunto por un prolongado tiempo la cuenca en numerosos y diversos aspectos.

En particular, en este capítulo, la estructura estadística de la onda anual media de caudal para cada estación fue analizada. En el caso del río Paraná, puesto que el periodo en común entre las estaciones incluía prácticamente todo el siglo XX, fue posible observar los cambios en los valores medios a partir de la década del setenta coincidente con otros autores ya mencionados, así como también la detección de eventos de crecida considerable durante la primera mitad del siglo, junto con los eventos de mayor apartamiento del año 1983. Por otro lado, en el río Uruguay los eventos más extremos que se registraron se correspondieron con el mes de julio del año 1983 coincidente con las estaciones del río Paraná. En base a los resultados del río Atuel, se entiende que resultó posible detectar la influencia que ocasionan los embalses previos a la estación Carmensa afectando el régimen de caudal. Sin embargo, dado el año de registro de la estación Carmensa, posterior a la fecha de la inauguración de los embalses (El Nihuil: 1948 y Valle Grande: 1964), no es posible en un principio conocer la magnitud de influencia sobre la misma.

Además, a partir de los valores de anomalías de las series de tiempo, se emplearon técnicas espectrales para representar temporal y espacialmente la variabilidad de las series. En este caso, algunas estaciones no presentaron ninguna onda significativa como es el caso del río Atuel. En este sentido, la metodología acarrea el problema de trade-off entre resolución de frecuencia y varianza, dando lugar a que resulte difícil precisar e identificar ondas cortas sin sacrificar la exactitud de la energía estimada. A pesar que el método

de Blackman Tukey puede separar el ruido y las componentes caóticas de las periódicas, su resolución no resulta suficiente para estimar cada periodicidad individual (Yiou et al., 1996). Por esta razón, se decidió mejorar los resultados considerando la metodología de wavelet. En el caso del río Paraná, si bien el método de Blackman Tukey detectaba la significancia de ondas cercanas a los doce meses, a través del método de wavelet se pudo observar la presencia de las mismas más que nada en una segunda parte del periodo. En el caso del río Uruguay, la primera metodología si bien presentaba un máximo absoluto de energía el mismo no resultaba significativo, pudiendo encontrar a través de la segunda metodología la significancia de esas ondas largas en ciertos intervalos de tiempo. En lo que respecta al río Atuel el análisis espectral no permitiría encontrar ondas significativas a diferencia de si encontrarlas con la metodología de wavelet.

Los estudios anteriores realizados, en conjunto con la metodología de wavelet coherencia detectaron mismas ondas de baja frecuencia presentes en el grupo de estaciones de un mismo río. Más aún, el estudio mostró que las estaciones cercanas exhibieron la presencia de ondas que se producen de manera prácticamente en fase. En este sentido, se entiende que es posible realizar una selección de estaciones de referencia en cada río, en particular en lo que refiere a ondas más largas posiblemente relacionadas con la variabilidad interdecadal y decadal.

En conclusión, el análisis realizado de las series temporales sugirió la siguiente selección de estaciones de referencia para estudios con fines climatológicos. En el caso del río Paraná, sería conveniente seleccionar más de una estación de referencia, debido a sus características fisiográficas (Corrientes y Túnel Subfluvial). En este sentido, este estudio extiende la literatura actual, puesto que la mayoría de los análisis han estudiado el alto Paraná. Además, las ciudades de mayor concentración de población se encuentran en la parte media e inferior de dicho río. En el río Uruguay, Paso de los Libres podría ser representativa en el tramo dentro del territorio Argentino; y en el caso del río Atuel sería la estación La Angostura, debido a su registro temporal y la no influencia de los embalses siguientes como sucedería con Carmensa.

Por lo tanto, la selección de estación de referencias podría resultar útil para monitorear en menor cantidad de puntos, y de esta manera reducir el número de variables para la representación hidrológica tanto para estudios temporales extensos como en grandes escalas espaciales. De esta manera, regiones particulares de las cuencas fueron estudiadas

en detalle, para poder facilitar el uso de los recursos hídricos en las regiones mencionadas.

1.7. Referencias

Anderson, R. L, 1942. Distribution of the serial correlation coefficient. Ann. Math. Statist. 13, no. 1, 1-13.

Antico, A., Torres, M.E., Diaz, H.F., 2015. Contributions of different time scales to extreme Paraná floods. Climate Dynamics. Vol. 46. 3785 - 3792.

Araneo, D., Villalba R., 2015. Variability in the annual cycle of the rio Atuel streamflows and its relationship with tropospheric circulation. International Journal of Climatology. Vol. 35. 2948-2967.

Araneo, D., Compagnucci, R., 2008. Atmospheric circulation features associated to Argentinean Andean rivers discharge variability. Geophysical Research letters. Vol. 35. 1-6.

Barros, V.R., Clark, R., Silva Dias, P., 2006. Climate Change in the La Plata Basin. 1st ed. Buenos Aires: CIMA-CONICET-UBA.

Barros, V.R., Doyle, M., González, M., Camilloni, I., Bejarán, R., Caffera, R.M., 2002. Climate variability over subtropical South America and the South American Monsoon: A review. Meteorológica. Vol. 27 Nro. 1 y 2, 33-57.

Barros, V.R., Gonzalez, M., Liebmann, B., Camilloni, I., 2000. Influence of the South Atlantic Convergence Zone and South Atlantic Sea Surface Temperature on Interannual Summer Rainfall in Southeastern South America. Theory. and Appl. Climatol. Vol. 67. 123-133.

Berbery, H., Barros, V.R., 2002. The hydrologic cycle of the La Plata basin in South America. Journal of Hydrometeorology. Vol. 3, 630-645.

Berbery, H., Collini, E., 2000. Springtime precipitation and water vapor flux convergence over southeastern South America. Mon. Wea. Rev. Vol 128, 1328-1346.

Brebbia, C., 2013. Water and society II. WIT Transactions on Ecology and the Environment. Vol 178.

Chatfield, C., 1995. The analysis of time series: An introduction. 5th edition. Chapman & Hall/CRC.

CLARIS LPB. A Europe-South America network for climate change assessment and impact studies in La Plata Basin (CLARIS LPB), 2008-2012.

Collischonn, W., Tucci, C.E.M., Clarke, R.T., 2001. Further evidence of changes in the hydrological regime of the River Paraguay: part of a wider phenomenon of climate change? Journal of Hydrology. Vol.245, 218-232.

Compagnucci, R., Araneo, D., 2007. Alcances de El Niño como predictor del caudal de los ríos andinos argentinos, Ing. Hidrául. Méx. Vol. 23. 23 – 25.

Compagnucci, R.H., Vargas, W.M., 1998. Inter-annual variability of the Cuyo rivers streamflow in the argentinean andean mountains and ENSO events. International Journal of Climatology. Vol. 18, 1593-1609.

Doyle, M., Barros, V.R., 2011. Attribution of the river flow growth in the La Plata Basin. International Journal of Climatology. Vol. 31. 2234–2248.

García, N., Vargas, W., 1998. The temporal climatic variability in the "Rio de la Plata" basin displayed by the rivers discharges. Climatic Change. Vol. 38. 359-379.

García, N., Vargas, W., 1995. The spatial variability of runoff and precipitation in the Río de la Plata basin. Hydrological Sciences Journal. Vol. 41.

Ghill, M., Allen, M.R., Dettinger, M.D., Ide, K., Kondrashov, D., Mann, M.E., Robertson, A.W., Saunders, A., Tian, Y., Varadi, F., Yiou, P.,2002. Advanced spectral methods for climatic time series. Reviews of Geophysics. 40(1), 1003.

Gouhier, T., Grinsted, A., Simko, V., 2018. R package biwavelet: conduct univariate and bivariate wavelet analyses.

Grinsted, A., Moore, J., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics. Vol. 11, 561–566.

Informe INA, 2003. Jaime, P., Menéndez, A. Vinculación entre el caudal del río Paraná y el fenómeno de el Niño.

Jaime, P., Menéndez, A., 2002. Análisis del régimen hidrológico de los ríos Paraná y Uruguay. Informe INA-LHA 05-216-02, Comitente: Proyecto Freplata.

Maraun, D., Kurths,J., 2004. Cross wavelet analysis: significance testing and pitfalls. Nonlinear Processes in Geophysics. Vol.11, 505-514.

Meis, M., Llano, M.P., 2018. Hydrostatistical study of the Paraná and Uruguay Rivers. International Journal of River Basin Management. Vol. 17. 1-12.

Nesbitt S., Cifelli R., Rutledge S, 2006. Storm morphology and rainfall characteristics of TRMM precipitation features. Monthly Weather Review. Vol. 134, 2702–2721.

Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. Freshwater Resources. Vol. 313, 1068-1072.

OMM, 1966. Nota técnica N° 79 OMM, Climatic Change.

Pasquini, A., Depetris, P., 2007. Discharge trends and flow dynamics of South American rivers draining the southern Atlantic seaboard: an overview. Journal of Hydrology, 333, 385–399.

Popescu, I. Brandimarte, L., Perera, U., Peviani, M., 2012. Assessing residual hydropower potential of the La Plata Basin accounting for future user demands. Hydrology and Earth System Sciences. Vol. 16. 2813-2823.

Prohaska, H.E., 1976. Climates of central and south America. World survey of climatology. Elsevier Scientific Publishing Company, Amsterdam. Vol.12, 57-69.

R Core Team, 2013. R: A Language and environment for statistical Computing. R Foundation for Statistical Computing, Vienna.

Sang, Y., 2013. A review on the applications of wavelet transform in hydrology time series analysis. Atmospheric Research. Vol. 122, 8-15.

Schmidbauer, H., Roesch, A., 2018. WaveletComp 1.1: A guided tour through the R package.

Siegel, S.,1985. Nonparametric statistical: for the behavioral sciences. 1st ed. Tokyo: McGraw-Hill Book Company.

Torrence, C., Compo, G., 1998. A Practical guide to wavelet analysis. Bulletin of the American Meteorological Society. Vol.79, 61-78.

Tucci, C., 2001. Some scientific challenges in the development of South America's water resources. Hydrological Sciences Journal. Vol.46. 937-946.

Vargas, W.M., Bischoff, S., Naumann, G., Marcuzzi, E., 2010. Runoff properties of extreme discharges on Paraná and Uruguay rivers. Hydrology and Earth System Sciences Discussions. Vol. 7. 2949-2973.

Vargas, W.M., Minetti, J., Poblete, A., 1995. Statistical study of climatic jump in the regional zonal circulation over South America. Journal of the Meteorological Society of Japan. Vol.73, nro. 5. 849-856.

Yamamoto R, Iwashima T, Sanga N, Hoshiai M, 1987. Detection of the climatic jumps. XIX IUGG. General Assembly. August 9–22. Vancouver, Canada.

Yiou, P., Baert, E., Loutre, M., 1996. Spectral analysis of climate data. Surveys in Geophysics. Vol. 17, Issue 6, 619-663.

Capítulo 2

Estudio de coherencia entre estaciones de referencia

2.1. Resumen

Mediante la selección de estaciones de referencia obtenidas en el capítulo anterior se analizó la covariabilidad espacial y temporal del caudal en distintas cuencas de la Argentina. Diferentes escalas temporales fueron analizadas; mensual, estacional y anual. Más aún, las anomalías mensuales de caudal entre pares de estaciones también fueron estudiadas. El análisis mostró que existe una relación entre las estaciones para ciertos intervalos temporales, en particular, para ondas de baja frecuencia. Por otra parte, se remarca la necesidad de profundizar en estudios estacionales para suavizar las fluctuaciones de alta frecuencia. Es decir, filtrar la componente estocástica que presentan las series hidrológicas, y de esta manera ahondar su estudio en relación con la variabilidad climática.

2.2. Introducción

Diferentes regímenes climáticos rigen en Argentina, lo cual en parte está asociado a su gran extensión espacial. En el territorio argentino es posible encontrar distintos tipos de climas desde áridos hasta húmedos. En este sentido, el estudio de la variabilidad espacial y temporal de distintas variables, en particular las asociadas con los recursos hídricos, requiere ser continuamente evaluado. Pues, el mismo da origen a herramientas económicas y sociales fundamentales para nuestro país.

Los recursos hídricos pueden verse afectados por diversas causas, tanto antropogénicas como naturales, siendo una de ellas la variabilidad climática. La misma ocurre en una escala temporal y espacial muy variable, que a menudo impacta negativamente en la agricultura y en los ecosistemas naturales (Meinke y Stone, 2005). En este sentido, la realización de pronósticos geoespaciales del caudal en cualquier escala temporal puede llevar a la mejora del uso del suelo y del manejo de los recursos hídricos, dando lugar a una cierta advertencia en la disponibilidad del agua (Clarkson et al., 2001).

De esta manera, estudiar la simultaneidad y covariabilidad entre series temporales de distintas regiones puede ayudar a comprender el grado de impacto que la variabilidad climática puede ocasionar en un territorio. En particular, Meis y Llano (2018) aplicaron distintas técnicas espectrales y de coherencia para estudiar el comportamiento de diferentes estaciones del río Paraná y el Uruguay, así como también entre dichos ríos. Asimismo, Sleziak y Szolgay (2015) estudiaron la relación entre distintos ríos de Eslovaquía, a través de la metodología de wavelet coherencia, analizando así la relación entre las series temporales de caudal en un cierto intervalo de tiempo. Los autores encuentran que la mayoría de las series temporales presentaban ondas en fase de periodo entre 2 y 8 años. A su vez, en Alemania, Petrow y Mers (2007) analizaron estaciones de medición, para un periodo común, mediante distintos indicadores de inundación. Los mismos lograron determinar distribuciones espaciales, como así también patrones estacionales. Asimismo, los autores encontraron, entre otros resultados, que estaciones con tendencias significativas se encontraban espacialmente agrupadas.

De este modo, los estudios sobre la distribución espacial y temporal de las series de caudal permiten colaborar con la información de la disponibilidad hídrica de un territorio. Por otra parte, los mismos contribuyen a obtener un mejor planeamiento frente a situaciones adversas.

En este capítulo se procedió a analizar la ocurrencia conjunta de ciclos característicos en estaciones de referencia. De esta manera, se evaluó la relación entre cuencas de distinta localización y fisiografía, analizando la posible semejanza entre las mismas. Para lo anterior, se tuvo en cuenta diferentes escalas temporales, es decir, primeramente se realizó un estudio a nivel mensual teniendo en cuenta un periodo en común. En segundo lugar, se analizó la relación a nivel anual y estacional, y por último se estudiaron los valores anómalos mensuales de caudal.

2.3. Datos y Metodologías

De acuerdo al capítulo anterior las estaciones de Corrientes, Túnel Subfluvial, Paso de los Libres y La Angostura podrían ser consideradas como estaciones de referencia. Este grupo particular de estaciones presentó un extenso periodo en común que comprende desde 1/1909 a 12/2016.

Entre las particularidades que las series presentaron se puede mencionar el caso de Paso de los Libres donde el año 1945 fue faltante. La bibliografía consultada demostró que fue un año seco (Genta et al., 1998), por lo tanto, se decidió completarlo con el decil 10. Asimismo, en el caso de La Angostura, si bien anteriormente se habían rellenado sus datos faltantes como se menciona en el capítulo 1, puesto a que la estación Loma Negra es semejante a la estación en cuestión, se decidió suplantar los datos faltantes por los datos de la estación Loma Negra. La cantidad de datos faltantes de la estación La Angostura se correspondió con un 0.22 % (8/2012 a 10/2012)

En lo que se refiere a los valores anuales de caudal, se consideró un único dato por año, evaluando el caudal promedio del año. Para el estudio del caudal estacional se utilizó el valor medio para cada estación del año considerando al verano (dic-ene-feb), otoño (mar-abr-may), invierno (jun-jul-ago) y primavera (sep-oct-nov). Por otro lado, para el cálculo de las anomalías se tuvo en cuenta el procedimiento realizado en el capítulo anterior.

Las metodologías estadísticas que se emplearon para evaluar la semejanza y coherencia entre estaciones fueron descriptas en el capítulo anterior (correlación de Spearman y wavelet coherencia).

2.4. Resultados

Mensuales 1909-2016

En primer lugar se trabajó con las series mensuales para los 109 años de información sin filtrar ningún tipo de particularidad que la serie pudiera presentar (Figura 2.1). Primeramente, se realizó una correlación de Spearman para evaluar la relación entre las mismas, encontrando que la correlación entre los pares de estaciones del Paraná y el Uruguay resultó ser positiva y significativa (Figura 2.2). En este caso se observó que el mayor coeficiente de correlación lo presentó el par Corrientes-Túnel Subfluvial, lo cual era lo

esperado por ser un mismo río; seguido por el par Corrientes-Paso de los Libres. Con respecto a la estación La Angostura la relación con las estaciones del río Paraná fue positiva y significativa, mientras que con la estación Paso de los Libres resultó ser negativa (Figura 2.2). Esto estaría relacionado con el comportamiento de las respectivas ondas medias anuales de cada estación (Capítulo 1), donde tanto la estación de Cuyo como las del río Paraná presentan sus valores máximos en los meses de verano, mientras que Paso de los Libres presenta un comportamiento opuesto alcanzado un primer valor máximo en el mes de julio.

Continuando con el análisis de la semejanza entre las series temporales de las estacio-

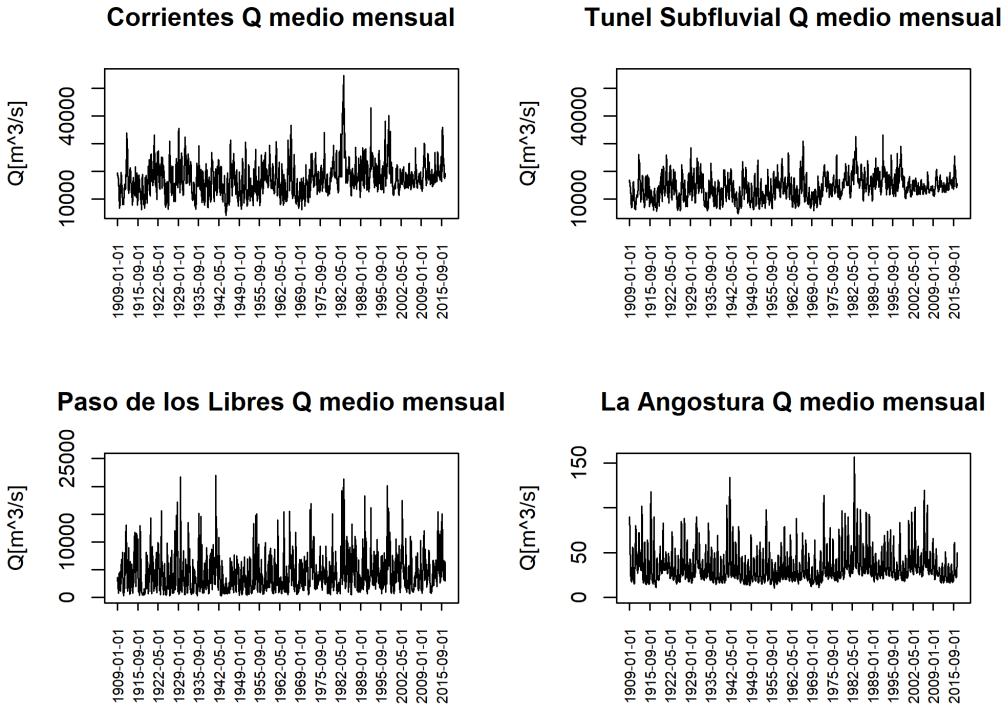


Figura 2.1: Series temporales de los datos centenarios de las estaciones Corrientes, Túnel Subfluvial, Paso de los Libres y La Angostura.

nes, se empleó el método de wavelet coherencia para determinar la presencia de ciclos característicos en pares de series temporales de estaciones y determinar así el desfasaje de ocurrencia de los mismos (flechas negras), Figura 2.3. Como se observó en el capítulo anterior, la relación entre la estación Corrientes y Túnel Subfluvial continuó siendo estrecha, en este extenso periodo temporal, presentando ciertas particularidades en ondas entre dos y tres meses donde la presencia de las mismas no resulta significativa en ambas estaciones.

Por otro lado, para el resto de las ondas la ocurrencia resultó ser prácticamente en fase para el par estudiado. Con respecto a la relación del caudal entre el río Paraná y Uruguay, encontramos que para el par Corrientes-Paso de los Libres tanto para comienzos del siglo XX como para finales de la segunda mitad del siglo se observó la presencia de ciertas ondas características entre 32 y 64 meses que presentarían una ocurrencia simultánea en ambas estaciones. En cuanto a ondas largas, también se destacaron las de entre los 64 y 128 meses con una significancia del 5 %, y con un cierto desfasaje menor a un año durante mediados de la década del setenta hasta principios de los noventa, aproximadamente. En lo que refiere a ondas cortas, se encontraron ciertas ondas esporádicas que ocurren en fase para el par analizado con ciclos característicos menores a los 8 meses. En lo que refiere a la relación Túnel Subfluvial y Paso de los Libres se observaron prácticamente las mismas características descriptas para el par anterior tanto en ondas cortas como en ondas largas.

En el estudio entre regiones alejadas donde su régimen fluvial resulta completamente distinto, se realizó el análisis entre las estaciones Corrientes y La Angostura; Túnel Subfluvial y La Angostura; y Paso de los Libres y La Angostura. En el primer par se destacó la presencia notoria de una onda anual presente en ambas estaciones con un desfasaje entre dos y tres meses (inclinación de las flechas en la Figura) que podría estar asociado a la ocurrencia en los valores mínimos de la onda media anual. Asimismo se observó para ondas más largas, durante la segunda mitad del siglo XX, la presencia en fase de ondas marcadas entre los 32 y 64 meses. En cuanto al segundo par, no se observaron diferencias notorias con respecto al análisis anterior, a excepción de que las ondas largas entre 32 y 64 meses si bien concentran un máximo relativo de energía, no es tan marcado como el caso anterior. En el estudio de coherencia entre La Angostura y Paso de los Libres se observa en la Figura 2.3 nuevamente la mayor energía en ondas significativas alrededor de la onda anual (entre los 8 y 16 meses). Las mismas presentaron un desfasaje de ocurrencia en promedio de seis meses, lo cual podría estar en relación con que La Angostura presenta su máximo medio anual en la estación cálida mientras que en Paso de los Libres es en la estación fría donde comienza aumentar su caudal medio. Por otra parte, en lo que refiere las ondas largas, se destaca la significancia de ondas entre los 32 y 64 meses en varios intervalos de tiempo que ocurren en fase para el par considerado.

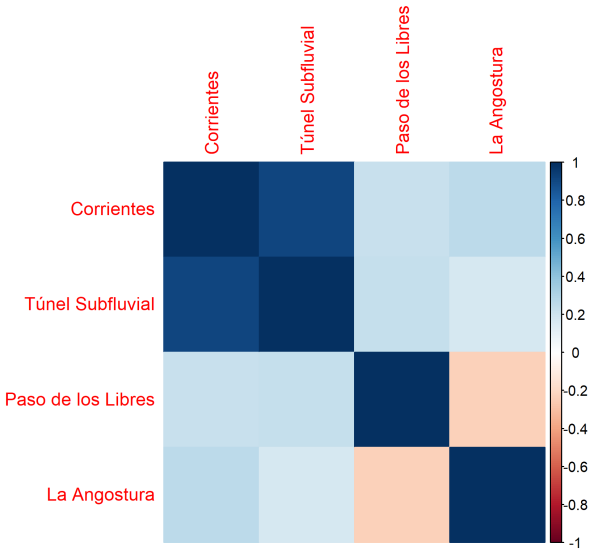


Figura 2.2: Correlaciones mensuales entre las series temporales de las estaciones de medición.

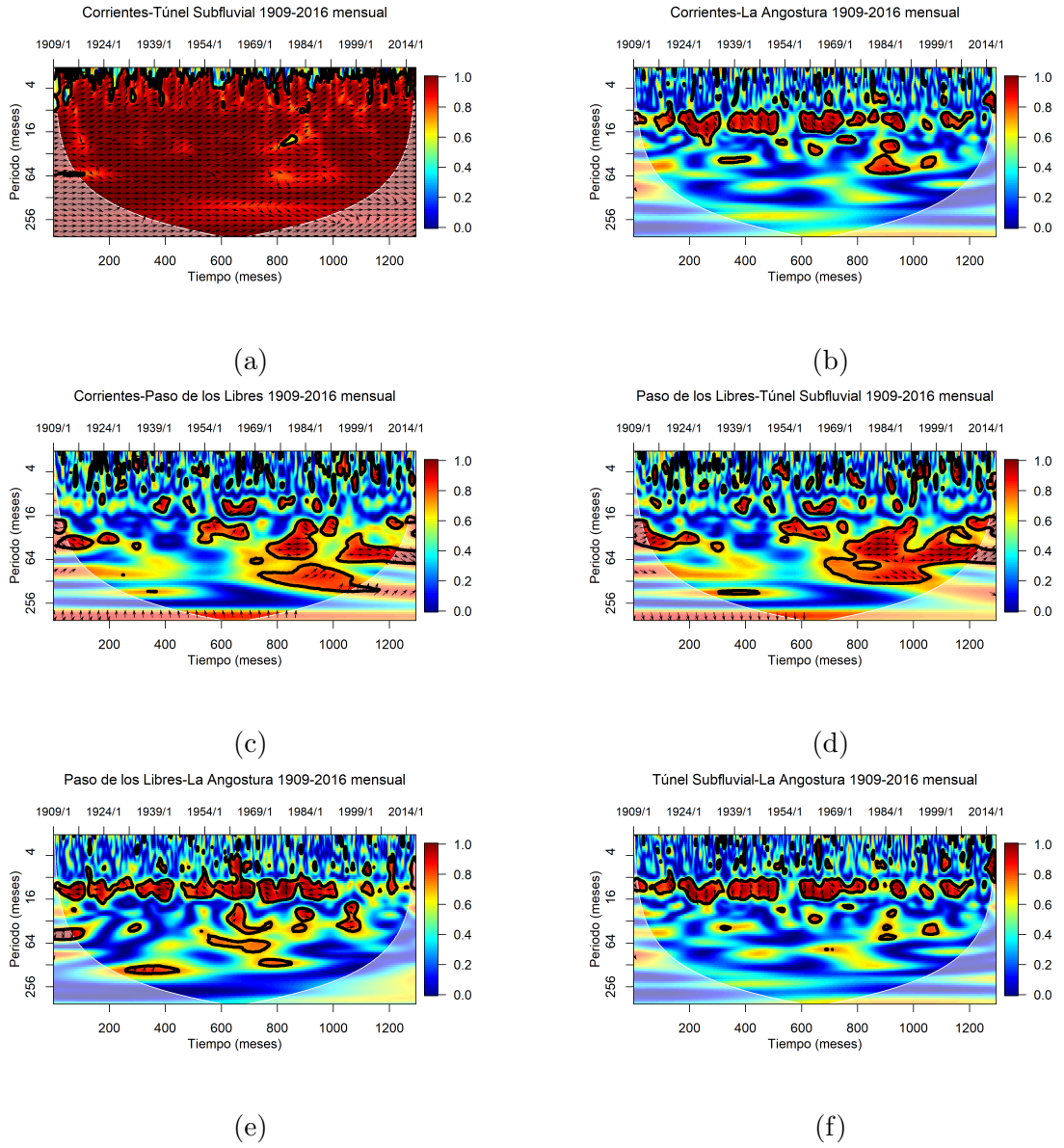


Figura 2.3: Wavelet coherencia de datos mensuales de caudal con un 95 % de confianza para el periodo 1909-2016. a) Corrientes-Túnel Subfluvial, b) Corrientes-La Angostura, c) Corrientes-Paso de los Libres, d) Paso de los Libres-Túnel Subfluvial, e) Paso de los Libres-La Angostura, f) Túnel Subfluvial-La Angostura.

Valores anuales medios de caudal 1909-2016

Para cada año de cada estación de medición se obtuvo un valor medio anual de caudal. Se realizó el análisis anterior de comparación entre las estaciones no solo considerando la correlación de Spearman (Figura 2.4), sino también teniendo en cuenta el procedimiento de wavelet coherencia. De acuerdo con la primer metodología, la mayor diferencia con respecto a lo hallado previamente con los valores mensuales es con el par de estudio La Angostura-Paso de los Libres, donde la correlación resulta positiva, esto estaría relacionado con el haber filtrado la onda anual. Es decir, el promedio anual suavizó la variabilidad de alta frecuencia, y esto es claramente observado en la Figura 2.5 entre Corrientes y Túnel Subfluvial, a excepción de ciertos años que no presentarían las mismas ondas significativas menores a tres años. Por otra parte, en Corrientes y Paso de los Libres, resultó bien clara la presencia de ondas relativamente en fase entre 4 y 8, y entre 8 y 10 años que suceden en ambas estaciones durante la segunda mitad del siglo XX, similar es lo que se observa también para la relación Paso de los Libres-Túnel Subfluvial. En lo que respecta con la estación La Angostura tanto Corrientes como Túnel Subfluvial, destacaron una banda de ondas entre los 3 y 6 años significativas y en fase que ocurren entre mediados de los setenta y mediados de los ochenta. En cambio, entre Paso de los Libres y La Angostura si bien también se encuentran ondas significativas entre los 3 y 6 años aproximadamente, existe un leve desfasaje de ocurrencia entre las estaciones menor a un año.

Valores estacionales medios de caudal 1909-2016

A nivel estacional se encontró para las correlaciones simples entre pares de estaciones, un resultado similar al hallado en el momento de realizar el estudio a los valores mensuales de caudal (ver Figura 2.6). En la Figura 2.7 se muestra el wavelet coherencia para cada par de estaciones. Es necesario mencionar que el par Corrientes y Túnel Subfluvial no presentó diferencias relevantes a lo hallado en los análisis previos. En el caso de Corrientes-Paso de los Libres donde en el análisis mensual presentaba cierto ruido, y en la escala temporal anual debido al suavizado se perdía gran cantidad de información, el análisis estacional resulta ser un balance temporal entre ambos. De esta manera, fue posible observar grupos de ondas significativas que ocurren relativamente en fase en ambas estaciones para varios intervalos temporales. Algunas de las ondas abarcan entre 4 y 8 Δt,

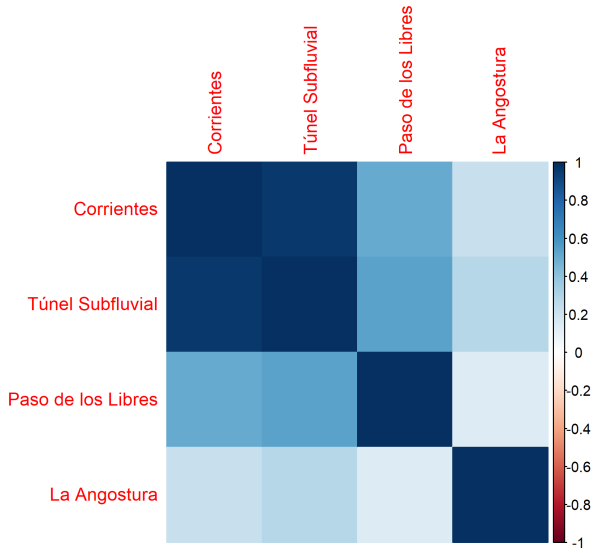


Figura 2.4: Correlación para datos anuales con un 95 % de confianza.

y entre 16 y 32 Δt (con $\Delta t = 3$ meses). El análisis previo resulta ser análogo para el par Túnel Subfluvial-Paso de los Libres. En el caso del par La Angostura-Corrientes tanto la onda anual ahora centrada en 4 Δt como también una onda aproximadamente de 16 Δt resultaron ser lo más destacable. En el caso de La Angostura-Túnel Subfluvial lo más notorio ocurre nuevamente para la onda anual con el mismo desfasaje establecido en el análisis mensual. Por otra parte, entre La Angostura y Paso de los Libres se observa que el suavizado de la variabilidad de alta frecuencia permitió obtener un análisis con menor ruido para este método donde la mayor parte de la energía estaría concentrada alrededor de la onda anual, pero, también fue posible hallar ondas esporádicas entre 8 y 16 Δt para distintos periodos temporales ocurriendo prácticamente en fase en ambas estaciones.

Valores de anomalía de caudal mensual 1975-2016

En general, las escalas temporales estudiadas, en la mayor parte de los pares destacaron la presencia significativa de ondas de baja frecuencia dentro de la segunda mitad del siglo XX. En conjunto con los cambios ocurridos en el valor medio durante el siglo pasado (García y Vargas, 1998; entre otros), se decidió aplicar el test de Yamamoto, et al. (1987) considerando ventanas de veinte años para comprobar el año del salto climático men-

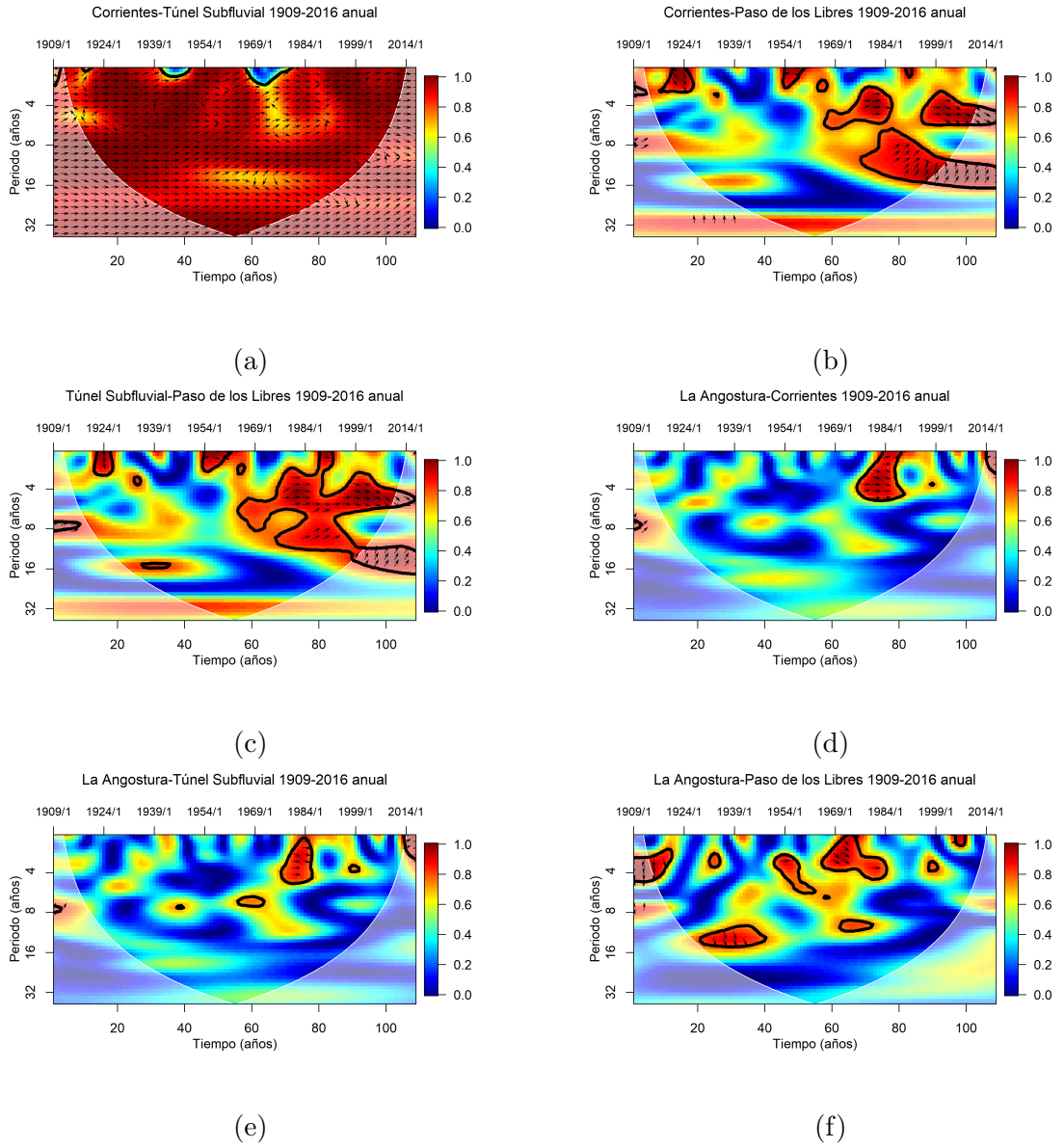


Figura 2.5: Wavelet coherencia para valores de caudal anual con un 95 % de confianza para el periodo 1909-2016. a) Corrientes-Túnel Subfluvial, b) Corrientes-Paso de los Libres, c) Túnel Subfluvial-Paso de los Libres, d) La Angostura-Corrientes, e) La Angostura-Túnel Subfluvial, f) Túnel Subfluvial-Paso de los Libres.

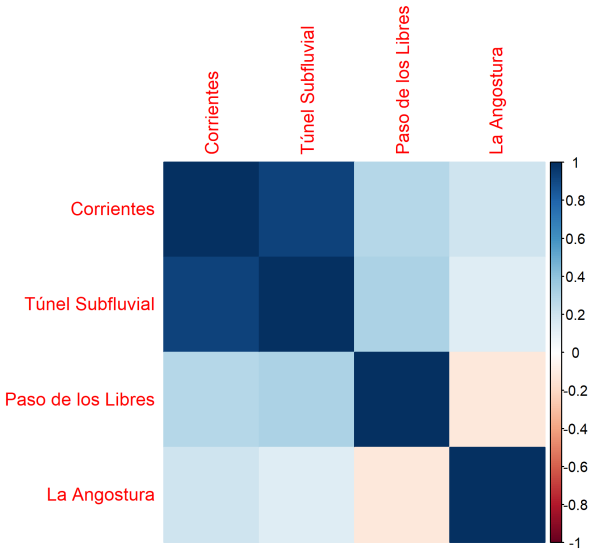


Figura 2.6: Correlación de valores estacionales de caudal con un 95 % de confianza.

cionado en los antecedentes previos. De esta manera, se encontró que dicho salto podría establecerse alrededor de la década del setenta en las cuatro estaciones de medición empleadas. En base al resultado anterior el periodo $1975 - 2016$ fue tenido en cuenta para evaluar las anomalías de caudal. De este modo, los valores de anomalías de caudal fueron obtenidos filtrando la onda media anual para dicho periodo. Luego, la tendencia lineal a partir de la metodología de cuadrados mínimos fue tenida en cuenta, encontrando que para las estaciones de medición o bien la tendencia no resultaba significativa o la varianza explicada resultaba ser menor al 1 %, por lo tanto, no fue filtrada.

En la Figura 2.8 observamos la correlación de a pares de estaciones en donde se destaca a diferencia del análisis anterior que la relación con La Angostura por parte del resto de las estaciones si bien resultan ser positivas para todos los casos es no significativa al tener en cuenta el coeficiente de correlación de Spearman.

El estudio de la wavelet coherencia entre las estaciones es mostrada en la Figura 2.9. A partir de la misma se observa que entre las estaciones del río Paraná no hay prácticamente diferencias con lo observado en el previo análisis. En el caso de Corrientes y Paso de los Libres, las ondas largas (entre 32 y 64 meses) ocurren en fase como se había observado en el análisis mensual. Asimismo, se observó la presencia de ondas entre los 64 y 128 meses

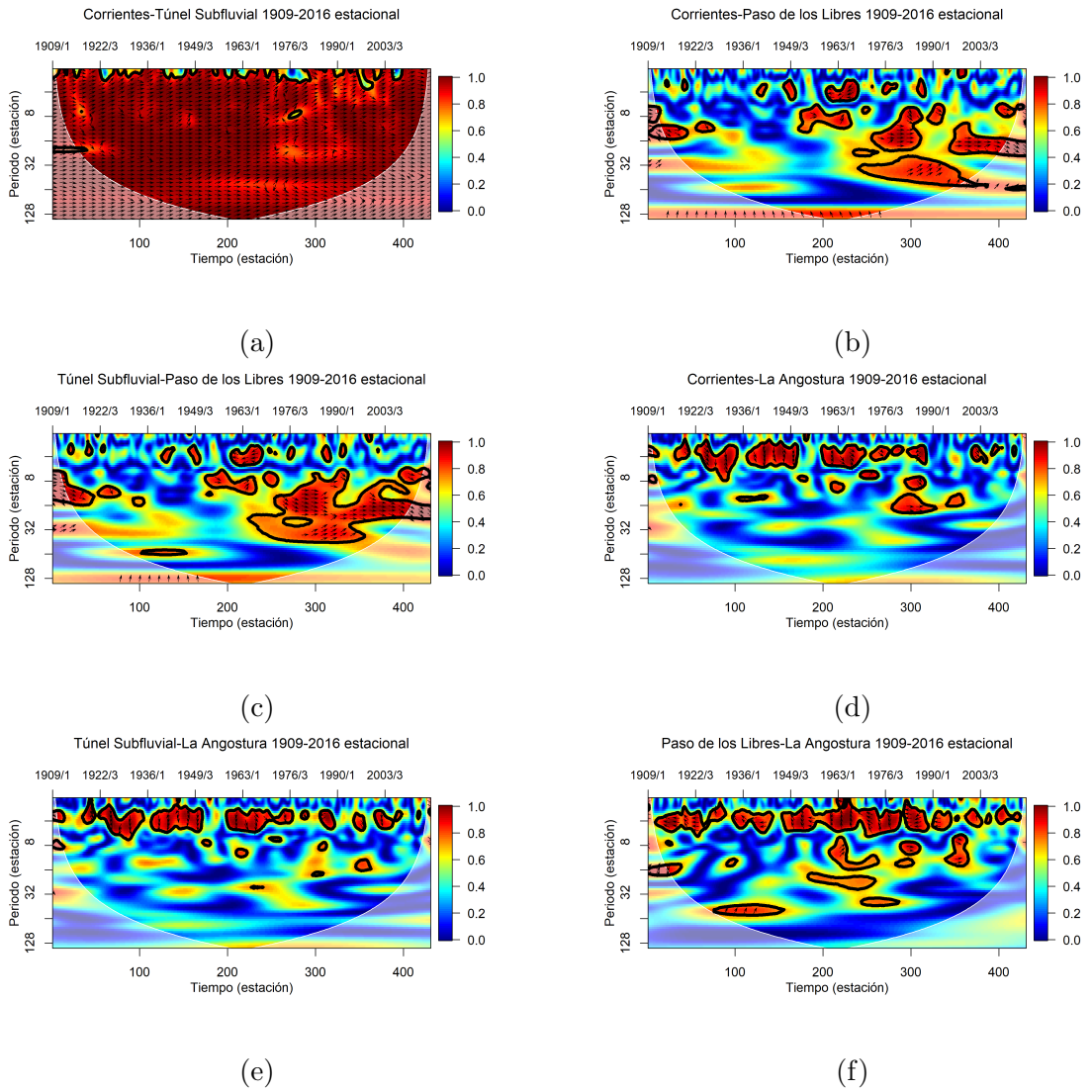


Figura 2.7: Wavelet coherencia estacional del caudal con un 95 % de confianza para el periodo 1909-2016. a) Corrientes-Túnel Subfluvial, b) Corrientes-Paso de los Libres, c) Túnel Subfluvial-Paso de los Libres, d) Corrientes-La Angostura, e) Túnel Subfluvial-La Angostura, f) Paso de los Libres-La Angostura.

en la década del 90 con un desfasaje menor a 32 meses entre las estaciones. También se destacó la presencia de ondas esporádicas entre los 4 y 8 meses que ocurren en promedio en forma simultánea en ambas estaciones. El par Túnel Subfluvial y Paso de los Libres presentó un comportamiento similar al par de estaciones anteriores, sin embargo, la presencia de las ondas largas descriptas anteriormente abarcan un periodo temporal mucho mayor siendo significativas y en fase en todo el intervalo temporal (teniendo en cuenta el cono de influencia). Este periodo de onda de alrededor de los 120 meses (aproximadamente 10 años ya había sido mencionado en Meis y Llano (2018), donde se observaban ondas significativas alrededor de los 10 años con un 5 % significancia, posiblemente relacionados con la modulación del ciclo del ENSO 10-15 años (Sun y Yu, 2009).

Con respecto al estudio con La Angostura notamos que tanto Paso de los Libres como Corrientes presentaron ondas significativas y en fase para ambos pares de estudio entre los 16 y 32 meses en intervalos temporales puntuales. Asimismo, en La Angostura-Corrientes volvió a ser notoria la presencia significativa de ondas entre los 32 y 64 meses que ocurren entre los años ochenta y finales de los noventa. En cuanto a La Angostura y Túnel Subfluvial si bien se observaron centros significativos y esporádicos entre los 4 y los 8 meses, junto con un centro de máxima energía relativo y significativo para ondas entre los 32 y 64 meses a la mitad del periodo, los mismos presentaron mayor ruido con respecto al par anterior. Por otro lado, La Angostura y Paso de los Libres no exhibieron grandes diferencias al análisis mensual a excepción de no notar la presencia de la onda anual que sí se destacaba en el primer estudio.

De lo anterior se desprende que, por ejemplo, en el caso de las correlaciones de a pares de estaciones con La Angostura si bien resultaban no significativas, la metodología de wavelet permitió establecer la presencia de ciertas ondas que ocurren tanto en las estaciones de la Cuenca del Plata y la del Atuel para igual periodos de tiempo. Además, también es importante remarcar que el filtro de la onda anual destacó que la energía que estaba concentrada alrededor de dicho intervalo ahora pueda ser resaltada en ondas de periodo más corto como las que se hizo mención entre los 4 y los 8 meses.

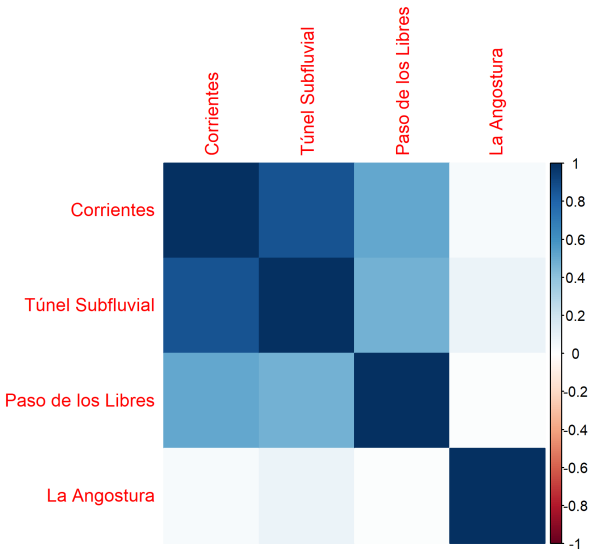


Figura 2.8: Correlación de anomalías de caudal con un 95 % de confianza.

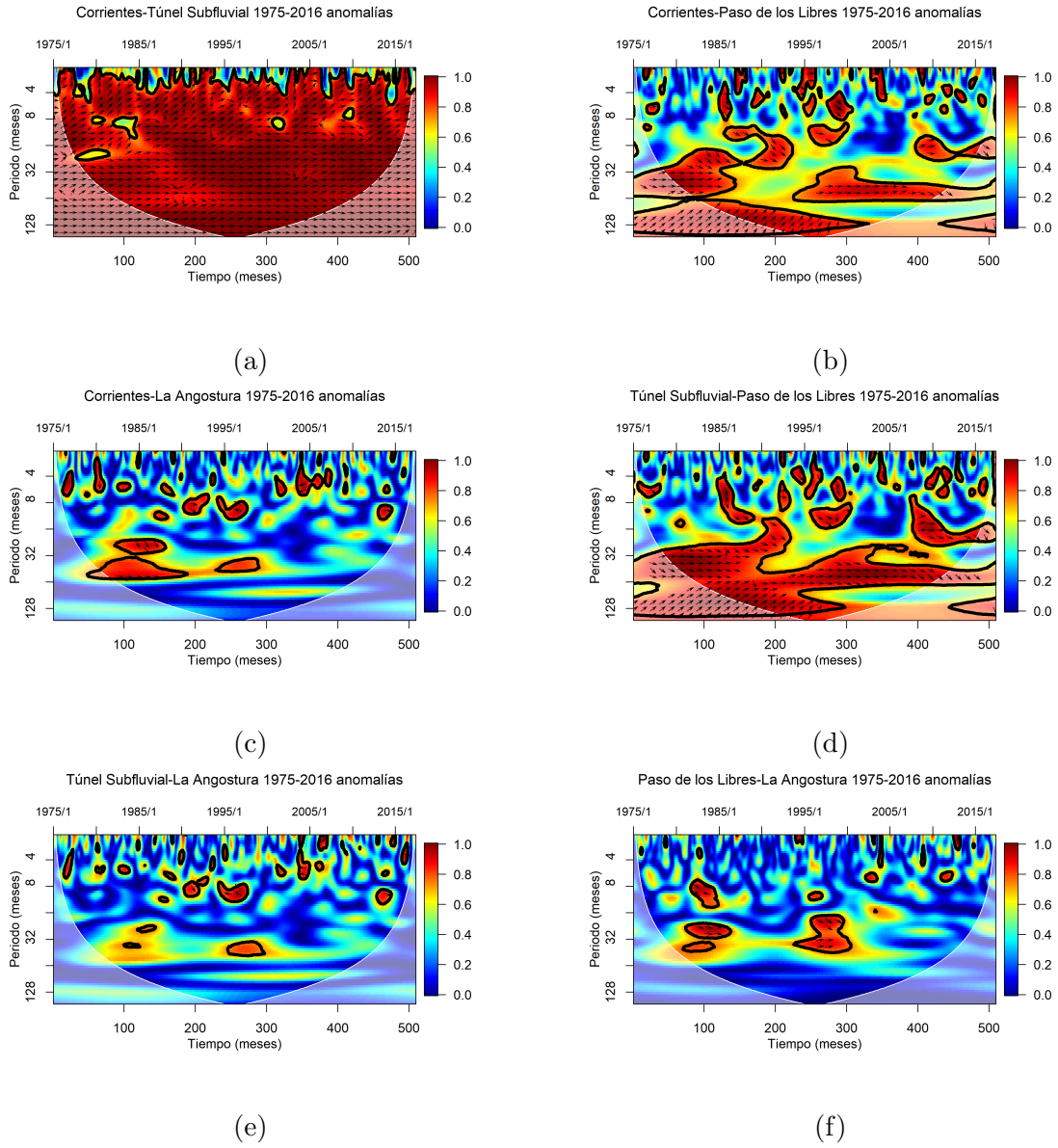


Figura 2.9: Wavelet coherencia entre las anomalías de caudal de las estaciones con un 95 % de confianza para el periodo 1975-2016. a) Corrientes-Túnel Subfluvial, b) Corrientes-Paso de los Libres, c) Túnel Subfluvial-Paso de los Libres, d) Corrientes-La Angostura, e) Túnel Subfluvial-La Angostura, f) Paso de los Libres-La Angostura.

2.5. Conclusiones

Siendo la Argentina un país agrícolo-ganadero, la disponibilidad hídrica presenta un rol fundamental para el desarrollo adecuado de dichas actividades. Por lo tanto, es necesario tener en cuenta el comportamiento y la relación que existe entre distintas cuencas, lo cual permitiría colaborar con la realización de medidas preventivas, en el caso de que fuese necesario realizarlas, para intentar subsanar eventos extremos tales como inundaciones y sequías. De esta manera, estudiar la variabilidad espacial y temporal de los recursos hídricos de un país permite un mejor aprovechamiento de los mismos.

Mediante distintas metodologías fue posible realizar un estudio de la covariabilidad espacial y temporal de la red de estaciones en cuencas de interés. Asimismo, evaluar diversas escalas temporales permitió analizar la presencia de ciclos característicos que ocurren en pares de estaciones, y con qué desfasaje se hace presente entre las mismas.

De acuerdo al análisis considerado, notamos por un lado que la relación entre las estaciones de la Cuenca del Plata resultó ser más estrecha que con la de la cuenca del Atuel. Sin embargo, en prácticamente todas las escalas temporales, en mayor o en menor medida, fue posible observar la presencia de ondas largas entre aproximadamente 2 y 5 años entre las estaciones de la Cuenca del Plata y la del Atuel consideradas. Dada su longitud, dichos ciclos podrían estar asociados a variabilidades climáticas conocidas. De esta manera en el siguiente capítulo se ahondará en dicha cuestión.

Asimismo, resulta importante destacar que para obtener un estudio más preciso de la variabilidad climática y su relación con el caudal, el análisis de datos estacionales podría resultar útil. Pues el mismo permite suavizar las fluctuaciones de alta frecuencia en el caudal, sin necesidad de ir a una escala temporal mayor. Además, como fue observado, dicho filtro no afecta las ondas de baja frecuencia que se quieren profundizar, no resultando necesario estudiar las anomalías de caudal.

2.6. Referencias

Clarkson, N.M., Clewett, J.F., George, D.A., 2001. Customised spatial climate forecasts to improve land and water management. Conference in Geospatial information in Agriculture 2001. Australia.

García, N., Vargas, W., 1998. The temporal climatic variability in the "Rio de la Plata" basin displayed by the rivers discharges. Climatic Change. Vol. 38. 359-379.

Genta, J., Perez-Iribarren, G., Mechoso, C.R., 1998: A recent increasing trend in the streamflow of rivers in southeastern South America. J. Climate, 11, 2858–2862.

Meinke, H., Stone, R., 2005. Seasonal and inter-annual climate forecasting: The new tool for increasing preparedness to climate variability and change in agricultural planning and operations. Climatic Change. Vol. 70. 221-253.

Meis, M., Llano, M.P., 2018. Hydrostatistical study of the Paraná and Uruguay Rivers. International Journal of River Basin Management. Vol. 17. 1-12.

Petrow, T., Mers, B., 2007. Trends in flood magnitude, frequency and seasonality in Germany in the period 1951–2002. Journal of Hydrology. Vol. 371. 129–141.

Sleziak, P., Szolgay, J., 2015. Analysis of the similarities between time series of discharges of selected Slovak rivers using wavelet coherence. Meteorologický Casopis, 18, 2015, 25 – 32

Sun, F., Yu, JY., 2009. A 10-15year modulation cycle of ENSO intensity. Journal of Climate. Vol. 22. 1718-1735.

Yamamoto R, Iwashima T, Sanga N, Hoshiai M, 1987. Detection of the climatic jumps. XIX IUGG. General Assembly. August 9–22. Vancouver, Canada.

Capítulo 3

Oscilaciones Climáticas y el Caudal

3.1. Resumen

Entender y monitorear eventos extremos de distintas variables meteorológicas es de suma importancia. En particular, resulta útil estudiar y pronosticar eventos que ocurren en cuencas de interés para la actividad económica y social de una región. En este capítulo se pretende cuantificar la relación entre índices climáticos y eventos extremos de caudal estacionales. El estudio comienza analizando la relación entre el clima y el caudal, y más aún se estima el desfajase de ocurrencia en las ondas del el índice y los valores de caudal medio estacional. Teniendo en cuenta el resultado, las series temporales son puestas en fase, y se utiliza el método de cópulas para ajustar una función de distribución de probabilidad conjunta. De esta manera, se obtiene un modelo que resulta particularmente útil para cuantificar la probabilidad de ocurrencia conjunta entre eventos extremos y permite monitorear sus periodos de retorno. Finalmente, se generan predicciones y se validan los modelos separando entre series de entrenamiento y de testeo. Los modelos que se obtienen pueden ser pensados como un acercamiento simple y efectivo para monitorear eventos extremos teniendo en cuenta ciertos índices climáticos.

3.2. Introducción

Predecir eventos extremos como crecidas y bajantes de ríos con cierta antelación resultaría de suma importancia para los tomadores de decisión. Tanto las sequías como las inundaciones afectan diversas actividades socio-económicas que se realizan a partir de

los recursos hídricos tales como, el manejo del agua potable, la navegación, el riego, la generación de energía, la ganadería y la agricultura.

Trabajos previos sugieren la relación entre grandes inundaciones o sequías influenciadas por oscilaciones climáticas (Grimm y Tedeschi, 2009; Penalba y Rivera, 2016). Camilloni y Barros (2003) destacaron la clara relación entre las fases del ENSO y las anomalías de caudal en el alto Paraná. Asimismo, la conexión entre el caudal medio y los eventos ENSO ha sido estudiado por Berri et al. (2002) para el alto Paraná mostrando evidencia del crecimiento del caudal medio durante los eventos NIÑO. Por otro lado, la mayor parte de las grandes inundaciones en la Cuenca del Plata se encuentran relacionadas con los eventos NIÑO. En este sentido, Antico et al. (2015) analizaron la contribución de distintas oscilaciones sobre las más grandes inundaciones registradas en el río Paraná. Más aún, el aumento de la frecuencia de eventos extremos en la región del sur de la LPB fue destacada por Cavalcanti et al. (2015). Por lo que el monitoreo de las variables hidrológicas y meteorológicas resulta ser clave para minimizar los impactos negativos que lo anterior involucra en la región.

Asimismo, el estudio del régimen hidrológico del río Atuel en la Argentina con relación a fenómenos climáticos ha sido expuesto por varios autores. Compagnucci y Vargas (1998) analizaron la variabilidad interanual en los ríos que comprenden la zona de Cuyo y su relación con los eventos ENSO, encontrando que la asociación entre dicho fenómeno y las anomalías del caudal en verano resulta evidente. Así también, Compagnucci et al. (2000) estudiaron el caudal del río Atuel a fin de establecer la posible influencia de los eventos ENSO en el mismo, hallando señales significativas en distintas bandas de frecuencia. Más aún, Araneo et al. (2015) lograron determinar los principales patrones de variabilidad de escala intra-estacional del caudal en el río Atuel asociado a una variabilidad intraestacional de anomalías de temperatura sobre la región de la cuenca. Por otra parte, Rivera et al. (2017) muestran la influencia que la región tropical del océano Pacífico presenta en el desarrollo de sequías hidrológicas, a nivel interanual relacionado con la fase fría del ENSO, y con la Oscilación Decadal del Pacífico (PDO) en las variaciones decadales.

Los extensos registros temporales de caudal en diversas cuencas han demostrado que las señales hidrológicas integran y reflejan oscilaciones climáticas interanuales, decadales y multidecadales (Labat, 2008). De esta manera la correlación entre el caudal y oscilaciones climáticas puede ser estudiada utilizando modelos hidrológicos estocásticos. Modelos

multivariados han sido empleados para estudiar la relación entre distintas variables meteorológicas e hidrológicas en diferentes regiones del mundo. AghaKouchak et al. (2014) han cuantificado el riesgo conjunto de anomalías de temperatura y precipitación en ciertas estaciones ubicadas en California, Estados Unidos, con el objetivo de colaborar con los tomadores de decisión, mientras que Khendun et al. (2014) modelaron la relación entre PDO, ENSO y las anomalías de precipitación en el estado de Texas para definir un modelo de predicción para eventos extremos. También, Cong y Brandy (2012) estimaron la relación conjunta usando un modelo de cópula en una región de Suecia para informar a las políticas de decisión en el sector de la agricultura en un contexto de cambio climático. Asimismo, Fang et al. (2018) han analizado las combinaciones de riesgo entre estaciones húmedas y secas contiguas a través del índice estandarizado de precipitación empleando modelos de cópulas. Ganguli y Reddy (2013) analizaron la relación entre ENSO y eventos de sequía en una región de la India, donde los resultados sugieren que incluir al ENSO en el modelo de cópula de variabilidad climática es útil para evaluar el riesgo de sequía y el manejo de los recursos hídricos en el área. En este sentido, Ward et al. (2014) han mostrado en una escala global que el ENSO ejerce una influencia fuerte y extensa tanto en el peligro como en el riesgo de inundación.

Como ha sido expuesto existe una fuerte relación entre ciertas oscilaciones climáticas y eventos extremos en varias partes del mundo. Por lo tanto cualquier esfuerzo que pueda ser realizado para cuantificar las relaciones anteriores es de gran interés. En este sentido, el objetivo principal de este capítulo fue estudiar eventos extremos del caudal en el río Paraná, Uruguay y Atuel en relación a forzantes climáticos, usando metodologías estocásticas como el método de cópulas. Se estimó la ocurrencia de eventos extremos de caudal y su relación con ciertas oscilaciones utilizando un modelo probabilístico, teniendo en cuenta un desfajase temporal entre el par analizado. Más aún, se cuantificaron las relaciones anteriores y se desarrollaron pronósticos.

Como ya se ha mencionado previamente, la cuantificación y el entendimiento de la variabilidad hidrológica es de suma importancia para la estimación del riesgo de inundación (Kiem et al., 2003), como así también para el riesgo de sequía. Ambas resultaron ser las principales motivaciones para la aplicación de una metodología que permitiese colaborar en el desarrollo de planes de mitigación.

3.3. Índices climáticos

Puesto que la realización de este capítulo involucró la utilización de ciertos índices climáticos, se realizará a continuación una breve explicación sobre los mismos.

El monitoreo del fenómeno del ENSO puede ser realizado en particular a través de distintas variables, una de ellas es la temperatura superficial del mar del océano Pacífico. Barnston et al. (1997) construyeron un índice de monitoreo conocido como NIÑO 3.4 que presentó una fuerte relación entre la presión a nivel del mar y la temperatura superficial. El índice climático se calcula a partir de la media móvil de tres meses de las anomalías de la temperatura superficial del mar para la región que comprende las latitudes 5°N-5°S y longitudes 120°-170°O.

Asimismo, otro indicador de monitoreo del fenómeno del ENSO es el índice de Oscilación del Sur (SOI). Este índice se calcula como la diferencia entre la presión atmosférica a nivel del mar en Tahití, dividido por su desvío estándar, y en Darwin también dividido por su desvío estándar. En eventos El Niño (respectivamente, La Niña) la presión en Tahití desciende (respectivamente, aumenta) con valor inferior al promedio, mientras que en Darwin aumenta (respectivamente, desciende) por encima del valor medio.

Por otra parte, el fenómeno del ENSO no es el único documentado, existen otro tipo de variabilidades como la Oscilación Decadal del Pacífico (PDO) que resultarían de utilidad para pronosticar eventos extremos en ciertas variables hidrometeorológicas. En particular, dicha oscilación al igual que el ENSO consiste en un calentamiento o enfriamiento de las aguas del Pacífico Norte (más allá de los 20°N), más aún resulta un patrón que modularía la variabilidad interanual del ENSO (Mantua et al., 1997). Por otra parte, el índice PDO fue definido como la componente principal de las anomalías mensuales de la temperatura superficial del mar en el océano Pacífico Norte (al norte de 20°N), removiendo primeramente el promedio mensual de temperatura superficial del mar.

La oscilación conocida como Dipolo del Océano Índico fue reconocida en primer lugar por Saji et al. (1999), donde los autores destacaron un patrón de variabilidad asociado a temperaturas anómalamente bajas de la superficie del mar frente a Sumatra, y altas temperaturas de la superficie del mar en el Océano Índico occidental (fase cálida). Asimismo, dicha oscilación estaría asociada a una influencia en la variabilidad de la precipitación interanual sobre la región central de Brasil y la Cuenca del Plata Subtropical en la prima-

vera austral (Chan et al., 2008). De igual manera en Garbarini (2016), también, se destacó la relación entre la precipitación estacional y la serie temporal del índice un mes previo a la precipitación encontrando relación negativa y significativa en la región subtropical de LPB en primavera y otoño. También, para la estación de verano fue posible establecer una relación significativa y negativa hacia el oeste de la Argentina, particularmente en la región de cuyo.

La Oscilación Antártica es un patrón de variabilidad atmosférica en el Hemisferio Sur fuera de los trópicos. La oscilación da lugar a cambios en el intensidad del viento zonal en latitudes altas afectando la circulación meridional. Para analizar este fenómeno se consideraron dos índices por un lado el índice medio mensual diario de la Oscilación Antártica (AAO) y por otro el índice Marshall del modo Anular del Sur (SAM, del inglés Southern Annular Mode). El primero de ellos se obtiene proyectando las anomalías mensuales de altura en 700 hPa. en el modo de la función empírica ortogonal principal (EOF), y luego es normalizada por el desvío estándar del índice mensual. En cuanto al índice SAM, el mismo se calcula como una diferencia zonal de presión entre las latitudes 40°S y 65°S, Marshall (2003).

El estudio de la influencia de la variabilidad de la Oscilación Antártica en las anomalías de precipitación en la región del Sudeste del Sudamérica ha mostrado que la señal de la AAO es significativamente fuerte durante las estaciones de invierno y finales de primavera (Silvestri y Vera, 2003). Más aún, Gillette et al. (2006) encontraron una región que disminuye la precipitación en verano y aumenta en invierno sobre el norte de Argentina, consistente con lo hallado por los previos autores mencionados.

3.4. Datos y Metodología

Para llevar a cabo el objetivo propuesto en este capítulo se consideró el caudal medio mensual para las estaciones de referencias seleccionadas (Capítulo 1): Corrientes, Túnel Subfluvial, Paso de los Libres y La Angostura. En este capítulo en particular se emplearon 45 años de estudio desde diciembre de 1974 hasta diciembre del 2016. Además seis índices climáticos fueron tenidos en cuenta: NIÑO 3.4, Southern Oscillation Index (SOI), Pacific Decadal Oscillation index (PDO), Antarctic Oscillation (AAO), Southern Annular Mode index (SAM),Indian Ocean Dipole index (DMI) (https://www.esrl.noaa.gov/). Si bien se

contaba con series temporales que presentaban un mayor registro temporal, la decisión de comenzar el estudio hacia finales del año 1974 radicó en la evidencia significativa de un cambio en el régimen de la circulación del Hemisferio Sur (García y Vargas, 1998).
Todas las series temporales fueron divididas estacionalmente, es decir, el caudal medio estacional (dic-ene-feb, mar-abr-may, jun-jul-ago y sep-oct-nov) fue tenido en cuenta. La Figura 3.1 muestra la evolución de las series temporales evaluadas en este capítulo. Además, los tests Aumentado de Dickey Fuller (Fuller, 1996) y Phillips-Perron (1998) fueron empleados para analizar la condición de estacionaridad en las series con un 5 % de significancia. Para cada par caudal estacional e índice climático se empleó el mismo periodo temporal. Como un primer acercamiento se estudió la relación entre la variable hidrológica y la climática considerando una correlación cruzada. Es importante tener en cuenta la fuerte suposición de normalidad de las series que asume la metodología. A su vez, la aproximación anterior fue acompañada por la metodología de wavelet coherencia para determinar la presencia de ondas significativas en conjunto en las series temporales de caudal y en las de los índices (Maraun y Kurths, 2004; Grinsted et al., 2004).
Puesto que uno de los objetivos del capítulo fue el de cuantificar la relación multivariada en eventos extremos, se estimaron funciones de distribución multivariadas usando la metodología de cópulas. De acuerdo al teorema de Sklar (1959), si $F_{(X,Y)}$ es la distribución conjunta de un par de variables aleatorias (X, Y), con F_X y F_Y sus distribuciones marginales, existe una única función cópula (Ec. 3.1) tal que:

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)) \tag{3.1}$$

para todo $x, y \in \mathbb{R}$.

Si F_X y F_Y son funciones continuas, existe una única C; de otra manera, C está definida solamente en el rango $F_X \times F_Y$.
Alternativamente, si C es una cópula, y F_X y F_Y sus funciones de distribución, la función F_{XY} dada por la ecuación anterior es una función de distribución conjunta con F_X y F_Y distribuciones marginales.
Para estimar distribuciones de probabilidad conjunta mediante la teoría de cópulas, se necesita obtener las distribuciones marginales que ajustan a cada serie de datos. De esta manera, se utilizó el método de máxima verosimilitud para identificar parámetros corres-

pondientes a distintas distribuciones (Normal, Lognormal, Gamma, Gumbel, Weibull, GEV) para cada estación de medición e índices. Asimismo, cada muestra fue comparada con las respectivas distribuciones teóricas usando el test de Kolmogorov-Smirnov (K-S). Un problema con este test es que la población de los parámetros es desconocida, entonces el test se encuentra influenciado por los parámetros estimados de la muestra (Jogesh y Rao, 2004). Por esta razón algunos autores tienen en cuenta la implementación de una metodología conocida como Bootstrap (Efron y Tibshirani, 1993) para estimar la distribución muestral del estadístico usando métodos de remuestreo aleatorios.

La metodología de remuestreo de Bootstrap es aplicada para la comparación entre la distribución de una muestra (en este caso una serie temporal) y una distribución teórica, para la cual sus respectivos parámetros son estimados. Primero, es necesario tener en cuenta que la muestra sigue una cierta distribución teórica, y que a partir de la misma los parámetros serán estimados utilizando el método de máxima verosimilitud. Con dichos parámetros el test K-S es realizado, y el estadístico obtenido (D) es considerado como "la verdad". Luego, una muestra de igual tamaño n que la serie original es generada con los mismos parámetros estimados de la muestra original, dependiendo de la distribución teórica bajo la cual se esté comparando. Posteriormente, el test K-S es aplicado teniendo en cuenta esos parámetros de la muestra teórica obtenidos utilizando el método de máxima verosimilitud.

Este procedimiento es repetido en este caso mil veces obteniendo mil estadísticos (D*). Finalmente, para determinar si la hipótesis nula debe ser rechazada (la muestra sigue una cierta distribución teórica) "se cuenta" la proporción que el estadístico D* es mayor que el estadístico conocido como "la verdad", D, ($\lambda = \#\{D* > D\}$, obteniendo un p-valor* definido como $\frac{\lambda}{N}$, con N igual a 1000). Si el p-valor* es mayor que 0.20 la hipótesis nula no es rechazada, y la muestra sigue la distribución teórica propuesta con un 95 % de confianza (Huaylla, 2015) . Además, es necesario mencionar que en el caso de que más de una distribución teórica no rechazara la hipótesis nula, se seleccionó el modelo que presentó el menor criterio de Akaike *(AIC)*.

Usando las distribuciones marginales teóricas que definieron el ajuste más adecuado, se procedió a seleccionar la función de cópula que mejor describe la relación bivariada entre cada par, índice climático y caudal estacional, utilizando el algoritmo propuesto por Schepsmeier et al. (2018). En este sentido, se concluye seleccionando una cópula

de acuerdo a diferentes familias (utilizando el criterio de AIC), y obteniendo la estimación de la función de probabilidad conjunta para cada caso. Dos tests de bondad de ajuste con un 95 % de confianza para cada familia de cópula fueron realizados utilizando "reesampleo" de Bootstrap (Schepsmeier et al., 2018): Cramer-Von Mises (CVM) y Kolmogorov-Smirnov (K-S).

Una vez que se ajustó una función cópula a los datos, y se obtuvo la correspondiente estimación de la función de distribución conjunta para el índice y el caudal de cada estación de medición, se utilizó el modelo para pronóstico. De esta manera, se propuso realizar muestras de los valores predichos para la variable respuesta *(Caudal)* de la distribución condicional obtenida después de observar un valor explicativo *(Índice Climático)*, para luego calcular métricas basadas en el valor esperado de la variable de respuesta dada una observación de la variable predictora, y en cuantiles de interés obtenidos de la distribución condicional.

En este sentido, lo que se quiere obtener es una muestra condicional de la variable Y *(caudal)* dada una observación $X = x$ *(índice)*. Para muestrear de la distribución condicional obtenida a partir del modelo de cópula se realizan los siguientes pasos:

- Construir $U_1 = F_X^{-1}(X)$ y $U_2 = F_Y^{-1}(Y)$, las variables uniformes correspondientes. Asimismo llamando $u_1 = F_X^{-1}(x)$ al valor observado de U_1.
- Obtener una muestra condicional de U_2 dado que $U_1 = u_1$, utilizando la función de distribución condicional de la cópula (Ec. 3.2).
- Transformar la muestra uniforme hacia el espacio de la variable original, aplicando F_Y.

Este procedimiento se llevo a cabo utilizando la implementación propuesta por Schepsmeier et al. (2018).

$$C(u_2|u_1) = P(U_2 \leq u_2|U_1 = u_1) = \frac{\partial^1 C(u_1, u_2)}{\partial u_1} \tag{3.2}$$

Asimismo, se consideró otra métrica para poder describir las funciones cópulas, con el objetivo de seleccionar el modelo más adecuado (Venter, 2001). En este sentido se tuvo en cuenta la media condicional acumulada (Ec. 3.3). De esta manera el valor esperado de U_2 dado $U_1 < z$ es:

$$E(U_2|U_1 < z) = \frac{\int_0^z \int_0^1 u_2 c(u_1, u_2) du_2 du_1}{z} \tag{3.3}$$

Como último punto, el percentil 95 obtenido del modelo de copula condicional fue analizado como predictor de eventos estacionales extremos. Para ello, se consideraron datos de altura de río [mts.] en el periodo diciembre-1974 a noviembre-2016 de las estaciones Túnel Subfluvial y Paso de los Libres, río Paraná y Uruguay, respectivamente. Los datos de altura fueron obtenidos de la SHR.

A partir de los datos de altura se observó si durante los trimestres de verano, otoño, invierno y primavera se superó al menos una vez el nivel de evacuación que se encuentra estipulado por Prefectura Naval (https://contenidosweb.prefecturanaval.gob.ar/alturas/). Para la estación Túnel Subfluvial el nivel de evacuación es de 5 [mts.], mientras que para la estación Paso de los Libres es de 8.5 [mts.]. De esta manera teniendo en cuenta dichos niveles de evacuación, los datos de altura se convirtieron en vectores de ceros y unos con un periodo de resolución trimestral. Donde cero se corresponde con el caso en el que durante un trimestre en particular no ocurrió ninguna evacuación, y uno a que al menos ocurrió una evacuación en el trimestre analizado.

Para evaluar el modelo se utilizó la metodología conocida como curva ROC, del inglés Receiver Operating Characteristic (Robin et al., 2011). La misma resulta ser una curva que mide la sensibilidad frente a la especificidad de un modelo para distintos puntos de corte de un cierto parámetro. Es decir, cada punto de la curva ROC representa la sensibilidad/especificidad de un par correspondiente a un punto particular de decisión (umbral). Asimismo otra medida de evaluación del rendimiento del modelo es el área bajo la curva ROC presentando valores entre cero y uno, cuanto más cercano a uno se encuentre dicha área mejor resulta ser el rendimiento del modelo.

Mediante la curva ROC se obtuvieron los umbrales de caudal esperado del percentil 95 con el objetivo de dar con un valor de caudal de evacuación. En particular, se consideraron dos valores de corte de caudal uno de manera visual a partir del gráfico de la curva ROC y otro teniendo en cuenta el F-score. El F-score resulta ser una métrica útil para evaluar el rendimiento del modelo que tiene en cuenta la precisión y el recall (exhaustividad) para estimar dicho valor. La precisión es el número de muestras positivas identificadas correctamente dividido por el número total de muestras que fueron clasificadas como positivas, y el recall es el número de los resultados positivos identificados correctamente divididos por el número de todas muestras positivas. El F-score es la media armónica entre ambas

cantidades.

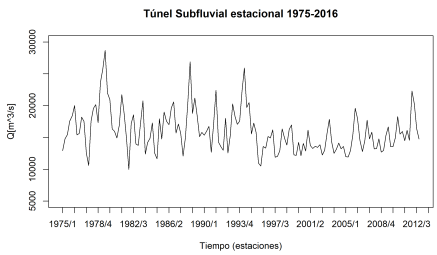
Finalmente para los umbrales que se identificaron a partir de los métodos anteriores se calcularon sus respectivas tablas de contingencia y mediante un test de chi-cuadrado se evaluó la hipótesis nula de independencia entre las variables con un 95 % de confianza.

3.5. Resultados

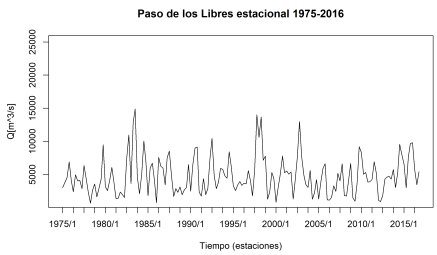
3.5.1. Forzantes climáticos y su relación con el caudal

En este capítulo se tuvieron en cuenta las series estacionales del caudal de la estación Túnel Subfluvial, Paso de los Libres y La Angostura. En cuanto a los índices climáticos se utilizó el índice NIÑO 3.4, SOI, PDO, SAM, AAO y DMI. Todas las series temporales mencionadas se observan en la Figura 3.1.

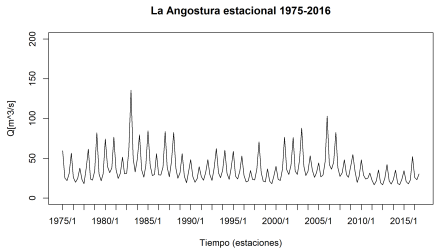
Asimismo teniendo en cuenta series estacionales tanto del caudal como de distintos índices climáticos se analizó la estacionaridad de las series. Mediante el test Aumentado de Dickey Fuller y Phillips-Perron se rechazó la hipótesis nula de no estacionaridad con un p-valor menor a 0.05. Esta suposición resulta clave para la mayoría de los métodos que se utilizaron a lo largo del capítulo. En esta primer parte se consideró el estudio entre distintos índices climáticos y las estaciones de referencia seleccionadas en el Capítulo 2. Puesto que, Corrientes y Túnel Subfluvial resultaban prácticamente semejantes, especialmente para periodos de baja frecuencia, en este capítulo solo se muestran los resultados con respecto a la estación Túnel Subfluvial. Para todos los pares de estudio, índice-caudal, se estudió como primera aproximación una correlación cruzada, sin embargo debido a la hipótesis de normalidad que este análisis presenta también se tuvo en cuenta una aproximación que no asume normalidad, el wavelet coherencia. A partir de estas metodologías se pretendió analizar la existencia o no de relación entre el caudal estacional y las distintas oscilaciones, y encontrar con qué tiempo de respuesta la segunda impacta sobre la primera. De manera tal de hallar una herramienta de pronóstico estacional entre las variables considerando un modelo probabilístico.



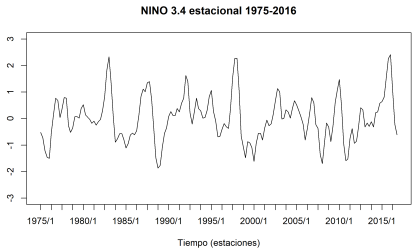
(a)



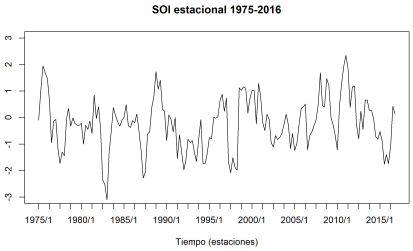
(b)



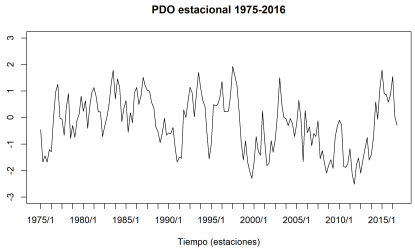
(c)



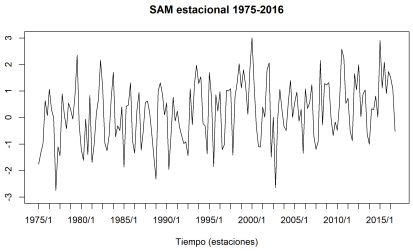
(d)



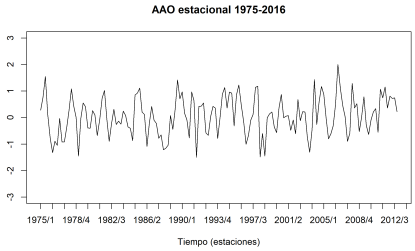
(e)



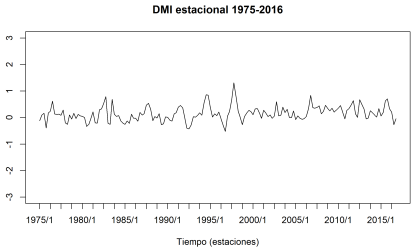
(f)



(g)



(h)

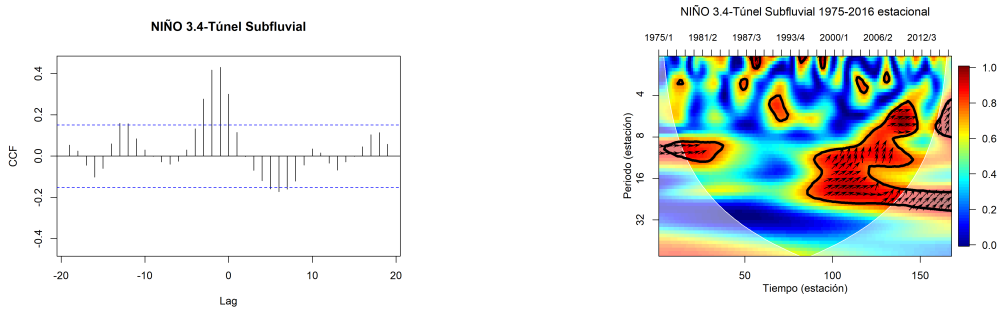


(i)

Figura 3.1: Series temporales trimestrales 1975-2016. a) Túnel Subfluvial, b) Paso de los Libres, c) La Angostura, d) NIÑO 3.4, e) SOI, f) PDO, g) SAM, h) AAO, i) DMI.

3.5.1.1. Índice NIÑO 3.4-Caudal

En el caso de estudio entre el NIÑO 3.4 y el caudal de la estación Túnel Subfluvial es posible observar a partir de la Figura 3.2a que al tener en cuenta la metodología de correlación cruzada, las mayores correlaciones se encuentran para uno y dos trimestres. Es decir, la mayor señal que el índice genera en el caudal se ve reflejado entre una y dos estaciones después (1 Δt y 2 Δt), con Δt igual a 1 trimestre. Con un valor de correlación significativo igual a 0.42. Por otra parte, la presencia de ondas significativas en el par de estudio es analizada mediante la metodología de wavelet coherencia. Este estudio conjunto entre el caudal estacional y el NIÑO 3.4 mostró ondas significativas en diferentes periodos, con un cierto desfajase entre las variables. La Figura 3.2b muestra la relación entre el NIÑO 3.4 y Túnel Subfluvial, a partir de la misma se observa un periodo de tiempo significativo entre principios del registro analizado y los años ochenta con un periodo estacional entre 8 Δt (8 trimestres) y 16 Δt (16 trimestres). Dada la inclinación de la fase (flecha negra en la Figura) se infiere que en promedio el desfasaje se encuentra cercano a 2 Δt. Además, ondas significativas fueron observadas en ambas variables a principios del siglo XXI, con periodos entre 4 y 8 Δt, y 8 a 16 Δt, y alrededor de 16 Δt.



(a) (b)

Figura 3.2: (a) Correlación Cruzada, (b) Wavelet coherencia entre el NIÑO 3.4 y Túnel Subfluvial.

También, la relación entre el caudal y el índice puede ser observada mediante un diagrama de dispersión, en el cual se tuvo en cuenta el resultado anterior de desfasaje temporal de una variable respecto a la otra. A partir de la Figura 3.3 se puede observar una relación positiva, es decir que para valores altos del índice, los valores de caudal también son mayores. Lo cual esta en concordancia con los antecedentes previos mencionados entre el ENSO y el caudal (Camilloni y Barros, 2003; Antico et al., 2015; entre otros).

Asimismo, se cuantificó el grado de relación entre las series temporales a partir del coeficiente no paramétrico de Spearman obteniendo una correlación de 0.29, siendo la misma significativa.

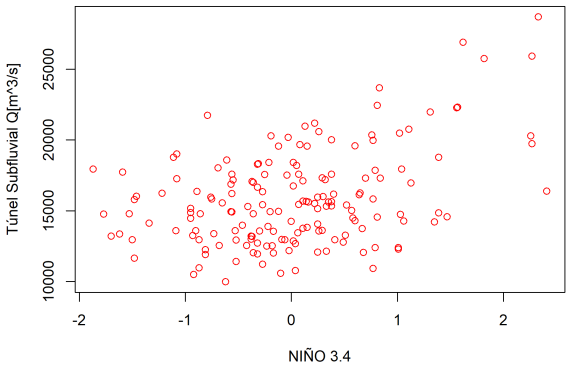
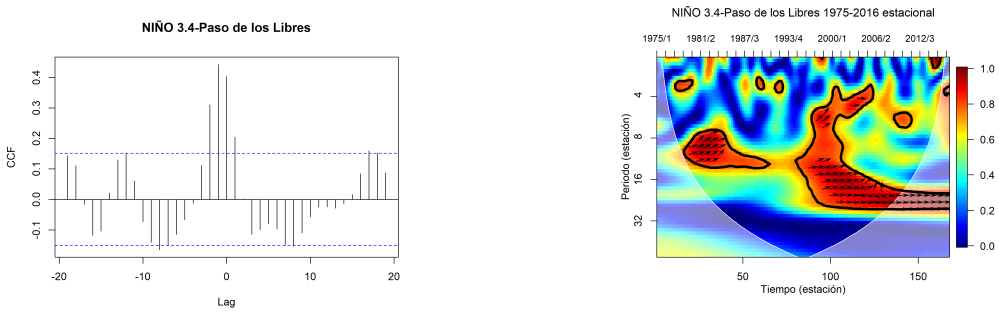


Figura 3.3: Diagrama de dispersión entre Túnel Subfluvial y NIÑO 3.4 con desfasaje igual a dos Δt.

Entre el índice y el caudal en Paso de los Libres, los resultados que se observaron fueron similares (Figura 3.4 y 3.5). También se destacó la relación positiva entre ambas variables, y se estableció un desfasaje de 2 Δt. Por otra parte, la correlación entre ambas variables teniendo en cuenta la fase estimada es de 0.2 (significativo con un 95 % de confianza), siendo menor la relación con esta estación que con Túnel Subfluvial.
En el caso de Túnel Subfluvial y Paso de los Libres las relaciones con el índice NIÑO 3.4 ya habían sido analizadas para un periodo de 100 años en Meis y Llano (2018), donde fue observada la relación entre el índice climático y su dominio sobre el caudal.



(a) (b)

Figura 3.4: (a) Correlación Cruzada, (b) Wavelet coherencia entre el NIÑO 3.4 y Paso de los Libres.

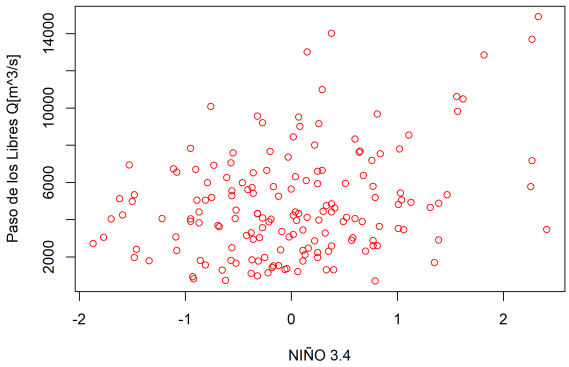
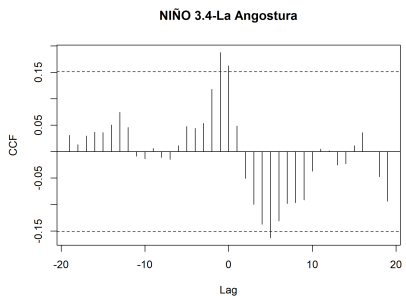


Figura 3.5: Diagrama de dispersión entre Paso de los Libres y NIÑO 3.4 con desfasaje igual a dos Δt.

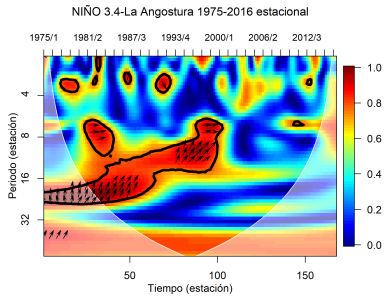
Es importante aclarar que si bien en ambos pares de estudios se encontró a partir de la correlación cruzada que la mayor correlación era para un retraso de un Δt, y del método de wavelet coherencia entre 1 y 2 Δt, se decidió considerar que el desfasaje es 2 Δt para los cálculos que se realizaron en las secciones posteriores. Esto es debido a que el objetivo del capítulo es la generación de un pronóstico con el mayor tiempo de prevención posible. Por otra parte, el estudio conjunto entre el índice y el caudal estacional de la estación La Angostura destacó a partir del estudio de la relación lineal entre las variables que la mayor señal se observa con un desfasaje de un trimestre (Figura 3.6a). En el caso del wavelet coherencia se observó a partir de la Figura 3.6b ondas significativas en ambas variables entre 8 y 16 Δt, donde se estimó una diferencia de fase que podría establecerse nula en algunos intervalos de tiempo, como así también con un desfasaje menor a dos estaciones. De esta manera a partir de las dos metodologías se estableció un desfasaje entre ambas variables de 1 Δt. Por último, a partir del gráfico de dispersión (Figura 3.7) se destaca una leve relación positiva entre las variables, donde además el coeficiente de correlación de Spearman al tener en cuenta dicho corrimiento temporal arrojó un valor alrededor de 0.15 (significativo a un nivel del 5 %).

3.5.1.2. Índice Oscilación del Sur-caudal

El análisis entre el índice SOI y el caudal estacional de la estación Túnel Subfluvial teniendo en cuenta una relación lineal presentó una correlación máxima en valor absoluto igual a dos trimestres (Figura 3.8a). Por otra lado, en el método no lineal resultó notable



(a)



(b)

Figura 3.6: (a) Correlación Cruzada, (b) Wavelet coherencia entre el NIÑO 3.4 y La Angostura.

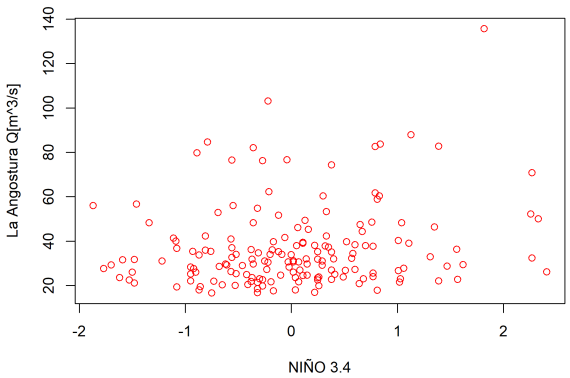
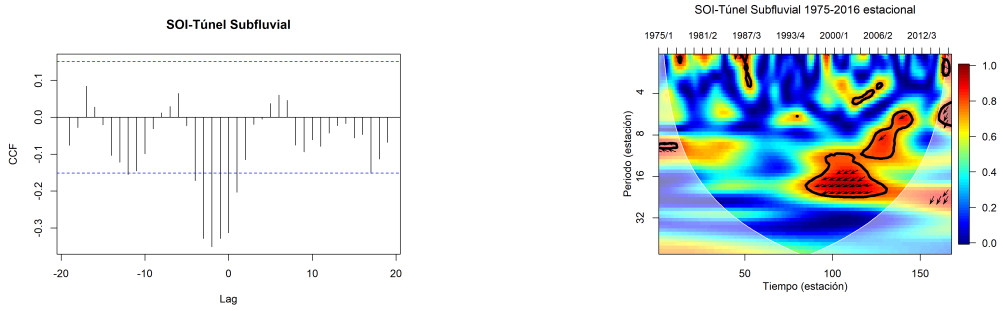


Figura 3.7: Diagrama de dispersión entre La Angostura y NIÑO 3.4 con desfasaje igual a un Δt.

la contrafase de ciertas ondas características significativas entre los 8 y los 16 Δt durante la segunda mitad del periodo de análisis, Figura 3.8b. Teniendo en cuenta las dos observaciones anteriores se puede establecer a partir de ambas metodologías un desfasaje entre ambas variables de aproximadamente 2 Δt, es decir dos estaciones temporales. Asimismo, en la Figura 3.9 se puede observar la relación negativa entre ambas variables, es decir, a menor valor del índice mayor es el valor de caudal involucrado. Más aún, el coeficiente de correlación de Spearman para estas dos variables teniendo en cuenta el desfasaje de 2 Δt presentó un valor igual a -0.24 (significativo con un 95 % de confianza).

En el caso con la estación de medición Paso de los Libres se observó por un lado valores significativos de correlación para más de un retraso Δt, particularmente, en 1 y 2 trimestres (Figura 3.10a). Sin embargo, mediante la Figura 3.10b se puede observar



(a) (b)

Figura 3.8: (a) Correlación Cruzada, (b) Wavelet coherencia entre el SOI y Túnel Subfluvial.

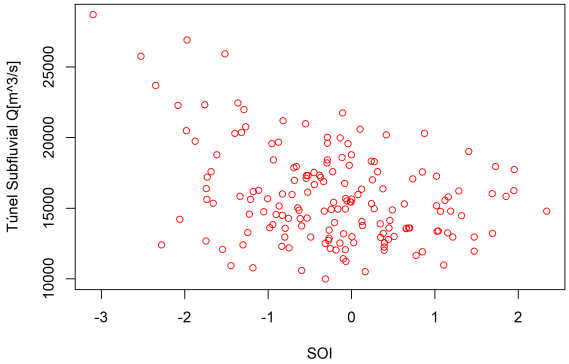
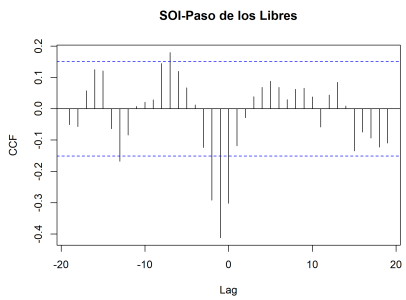


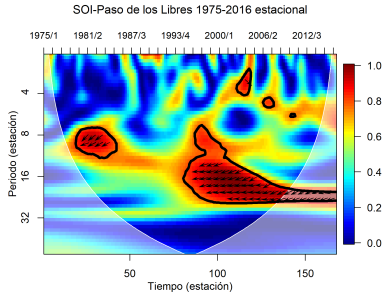
Figura 3.9: Diagrama de dispersión entre Túnel Subfluvial y SOI con desfasaje igual a dos Δt.

que el mayor intervalo temporal de ondas significativas para ambas variables resultó estar prácticamente en contrafase. Por lo tanto, estableciendo un compromiso entre ambas metodologías se decide considerar un desfasaje de 1 trimestre a fin de poder realizar un pronóstico del caudal a partir del índice SOI. También al igual que para la estación anterior la relación es negativa entre el par analizado (Figura 3.11). Sin embargo, Paso de los Libres exhibió una mayor señal presentado un coeficiente de correlación significativo igual a -0.34.

La relación entre el fenómeno ENSO a través del índice SOI y el caudal estacional de La Angostura presentó un claro desfasaje de 2 estaciones temporales (Figura 3.12). Al igual que los casos anteriores la relación es negativa la cual se puede apreciar mediante el diagrama de dispersión, Figura 3.13. Al tener en cuenta la correlación no paramétrica



(a)



(b)

Figura 3.10: (a) Correlación Cruzada, (b) Wavelet coherencia entre el SOI y Paso de los Libres.

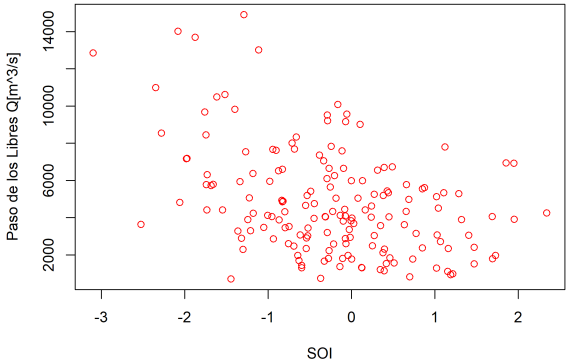
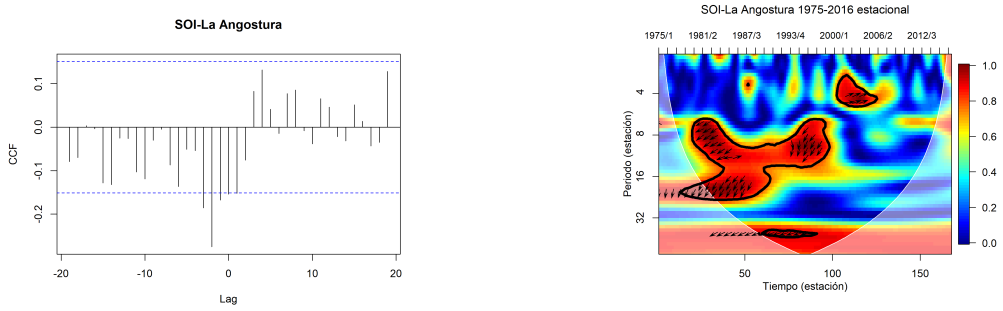


Figura 3.11: Diagrama de dispersión entre Paso de los Libres y SOI con desfasaje igual a un Δt.

de Spearman se obtuvo -0.27, significativa con un 95 % de confianza.

3.5.1.3. Índice Oscilación Decadal del Pacífico-caudal

En el caso particular de la exploración de la relación entre el índice PDO y el caudal estacional de Túnel Subfluvial se observa a partir de la correlación cruzada (Figura 3.14a) que el mayor valor de correlación se destaca para un trimestre. Sin embargo, al tener en cuenta el wavelet coherencia hay ondas entre los 4 y 8 Δt, como así también alrededor de los 16 Δt que presentarían un desfasaje alrededor de 2 estaciones (Figura 3.14b). Por lo tanto teniendo en cuenta ambas metodologías, no fue tan claro que desfasaje considerar para el análisis posterior, por lo que se tuvo en cuenta 1 y 2 estaciones. En la Figura 3.15 se destaca la relación positiva entre las variables, y los coeficientes de correlación de



(a) (b)

Figura 3.12: (a) Correlación Cruzada, (b) Wavelet coherencia entre el SOI y La Angostura.

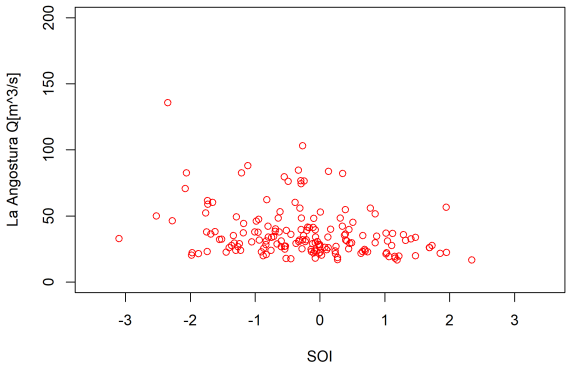
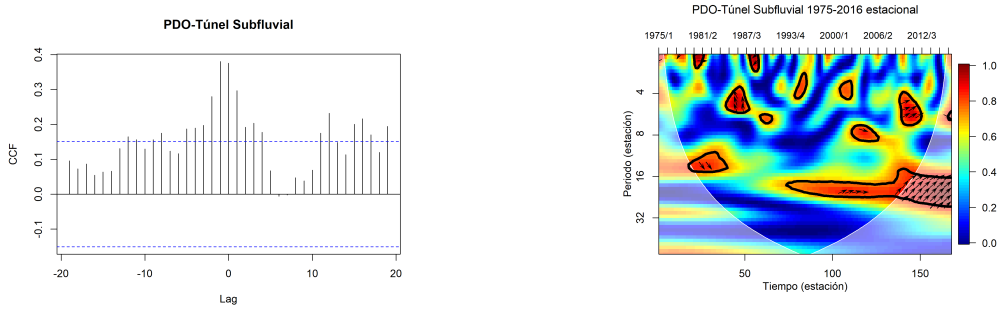


Figura 3.13: Diagrama de dispersión entre La Angostura y SOI con desfasaje igual a dos Δt.

Spearman sería de 0.37 para 1 estación, y de 0.25 para 2 estaciones.

Al igual que el estudio anterior, el índice PDO y la estación Paso de los Libres tampoco estaría claro que desfasaje considerar si una estación o dos. A juzgar por la correlación cruzada (Figura 3.16a) se entiende que la mayor correlación significativa se encuentra a un trimestre, sin embargo, al tener en cuenta la metodología de wavelet el desfasaje no es tan evidente como las relaciones anteriores (Figura 3.16b). Por lo tanto, también se procedió a tener en cuenta ambos desfasajes para el análisis posterior. La correlación de Spearman resultó ser de 0.22 y 0.20, para 1 Δt y 2 Δt, respectivamente. Se exhibe en la Figura 3.17 la relación positiva entre ambas variables para un desfasaje de 2 Δt.

Por último, en el estudio del par PDO-caudal estacional de La Angostura se destacó un desfasaje que se puede establecer en 3 estaciones. Si bien a través de la correlación



(a) (b)

Figura 3.14: (a) Correlación Cruzada, (b) Wavelet coherencia entre el PDO y Túnel Subfluvial.

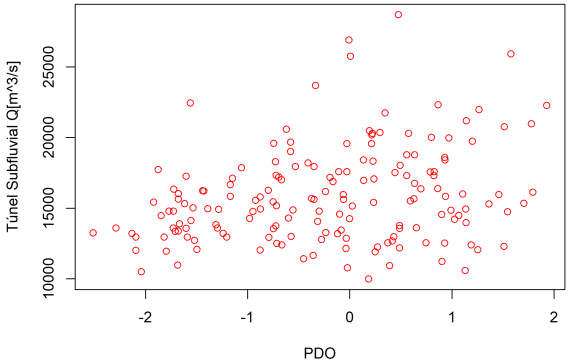
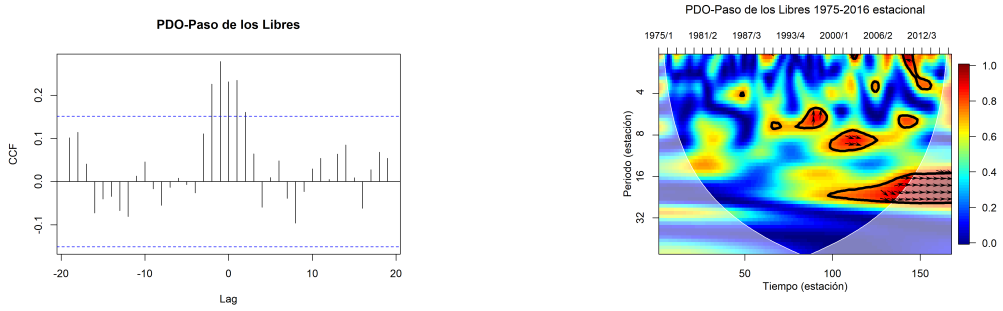


Figura 3.15: Diagrama de dispersión entre Túnel Subfluvial y PDO con desfasaje igual a dos Δt.

cruzada (Figura 3.18a) se observó una de las mayores correlaciones a lag 0, la misma no resulta útil como herramienta de pronóstico. Es por ello que se establecieron 3 estaciones como desfasaje entre las variables a partir de las dos metodologías. Puesto que a través de la metodología de wavelet coherencia se destacaron ondas de periodo de 4 Δt que presentan un desfasaje entre 3 y 4 estaciones (Figura 3.18b). El coeficiente de correlación de Spearman para este corrimiento presentó un valor igual a 0.22, significativo con un 95 % de confianza. Dicha relación es también observable a partir de la Figura 3.19.

3.5.1.4. Índice Modo Anular del Sur-caudal

La relación SAM y caudal estacional de Túnel Subfluvial muestra por un lado a partir del método lineal, una señal negativa y significativa para 9 y 10 trimestres (Figura 3.20a).



(a) (b)

Figura 3.16: (a) Correlación Cruzada, (b) Wavelet coherencia entre el PDO y Paso de los Libres.

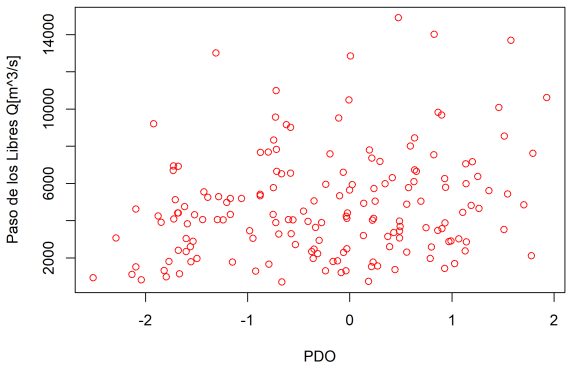
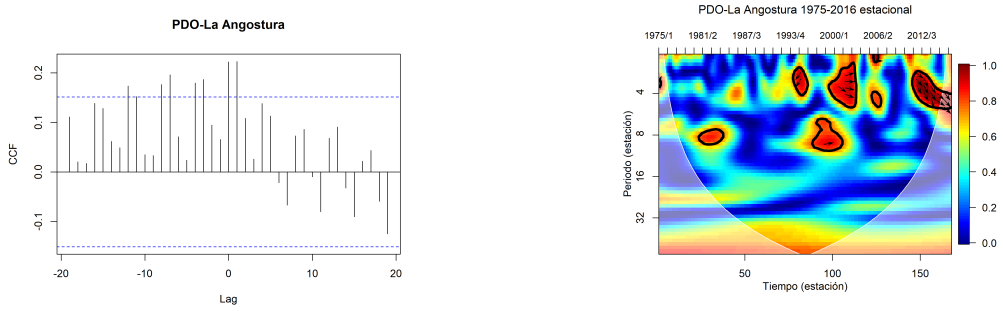


Figura 3.17: Diagrama de dispersión entre Paso de los Libres y PDO con desfasaje igual a dos Δt.

Asimismo, en el método no lineal es notable una onda de baja frecuencia entre 16 y 32 estaciones que abarca un periodo temporal menor a 50 estaciones (Figura 3.20b), que estaría presentando una diferencia de fase entre 6 y 7 estaciones. Por lo tanto, si bien hay un tipo de relación, dado que para ambos métodos el desfasaje exhibió una gran amplitud (superior al año y medio), y que solo se encontró señal en un periodo determinado de tiempo, se entiende que no sería del todo adecuado considerar este índice para un pronóstico estacional con tanto tiempo de antelación ya que se ha observado hasta ahora que otras oscilaciones tienen un impacto en más corto plazo, pudiendo cambiar así los resultados encontrados. Por lo que esta relación no será tenida en cuenta para los futuros cálculos.

Para el segundo par SAM-caudal Paso de los Libres, a diferencia del anterior, sí se en-



(a) (b)

Figura 3.18: (a) Correlación Cruzada, (b) Wavelet coherencia entre el PDO y La Angostura.

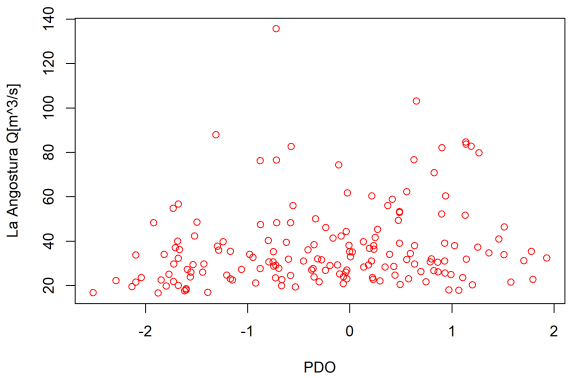
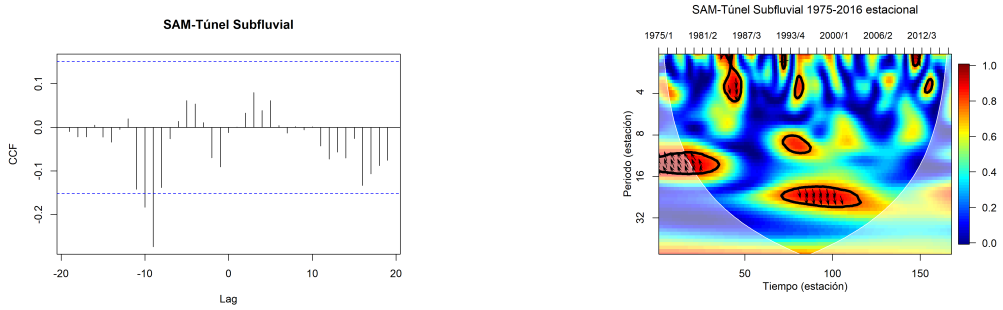


Figura 3.19: Diagrama de dispersión entre La Angostura y PDO con desfasaje igual a tres Δt.

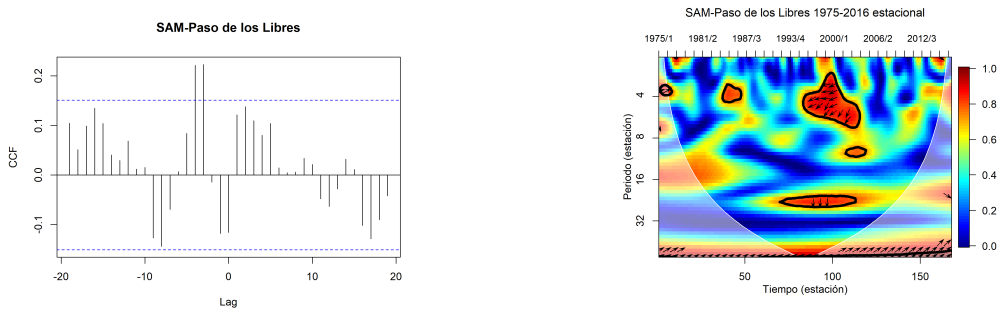
contró una señal más inmediata mediante el método lineal, en particular para 3 trimestres (Figura 3.21a), donde también es levemente apreciable mediante el wavelet coherence para periodos alrededor de 4 a 8 estaciones (Figura 3.21b). Más aún, se presenta el diagrama de dispersión entre estas dos variables (Figura 3.22), con dicho desfasaje, donde es posible apreciar una relación positiva, en conjunto con un coeficiente de correlación igual a 0.22 (significativo a un nivel del 5 %).

En el caso de La Angostura no fue posible establecer un tiempo de respuesta entre el SAM con respecto al caudal. Es decir, mediante la correlación cruzada no se encontró ningún tipo de relación, y mediante el segundo método si bien se destacó un pequeño intervalo temporal hacia final del periodo no se considera lo suficientemente razonable para establecer una diferencia de fase entre ambas variables (Figura 3.23).



(a) (b)

Figura 3.20: (a) Correlación Cruzada, (b) Wavelet coherencia entre SAM y Túnel Subfluvial.



(a) (b)

Figura 3.21: (a) Correlación Cruzada, (b) Wavelet coherencia entre SAM y Paso de los Libres.

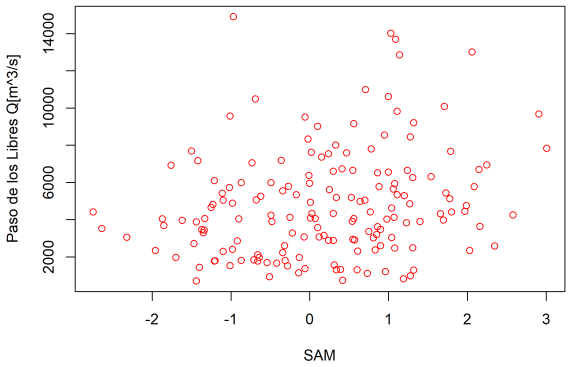
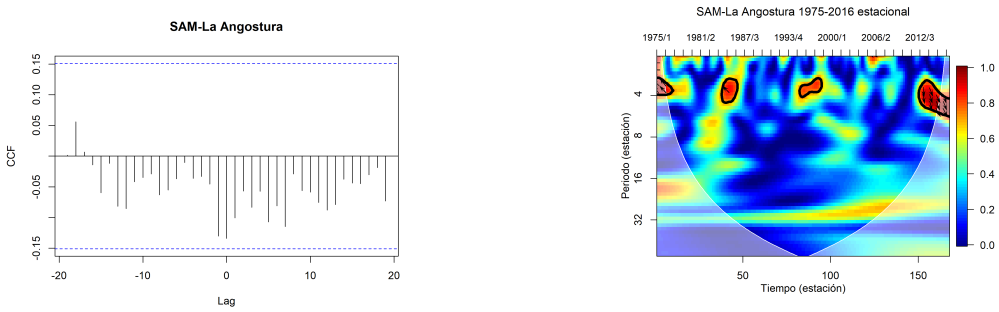


Figura 3.22: Diagrama de dispersión entre Paso de los Libres y SAM con desfasaje igual a tres Δt.



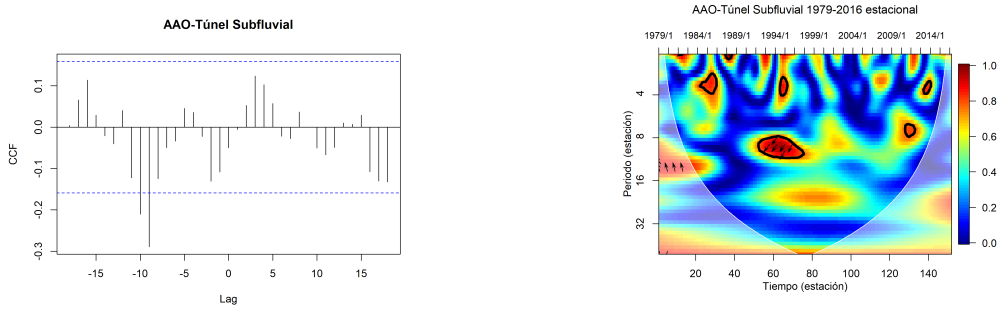
(a) (b)

Figura 3.23: (a) Correlación Cruzada, (b) Wavelet coherencia entre SAM y La Angostura.

3.5.1.5. Índice Oscilación Antártica-caudal

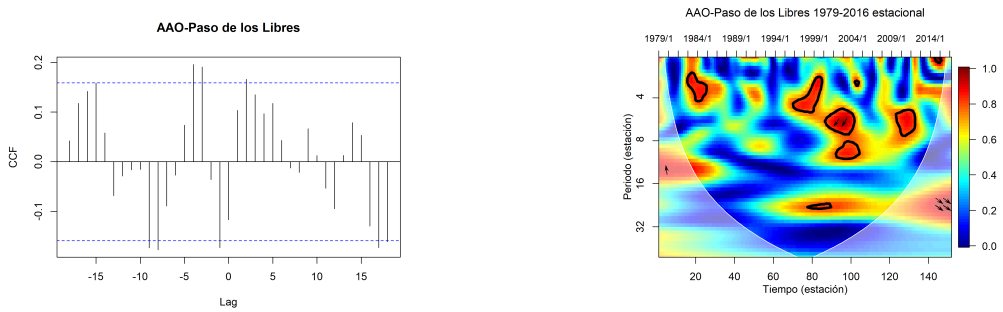
Al igual que con el índice SAM, el caudal estacional de la estación Túnel Subfluvial tampoco estaría presentando una relación en un plazo de tiempo próximo, y si bien el método no lineal destaca ciertos intervalos significativos no se estaría en condiciones de establecer un tiempo respuesta entre ambas variables (Figura 3.24).

Al igual que la estación anterior, con Paso de los Libres, no resultaría tan claro determinar un desfasaje entre ambas variables. Si bien, mediante la metodología de correlación cruzada se observaron ciertos desfasajes significativos que podrían indicar un tiempo de respuesta del caudal, al considerar wavelet coherencia no fue factible encontrar una diferencia de fase que permanezca durante un periodo prolongado de tiempo (Figura 3.25). Por lo tanto, se entiende que a fin de establecer un pronóstico estacional para eventos extremos la utilidad que podría dar este índice no es del todo clara.



(a) (b)

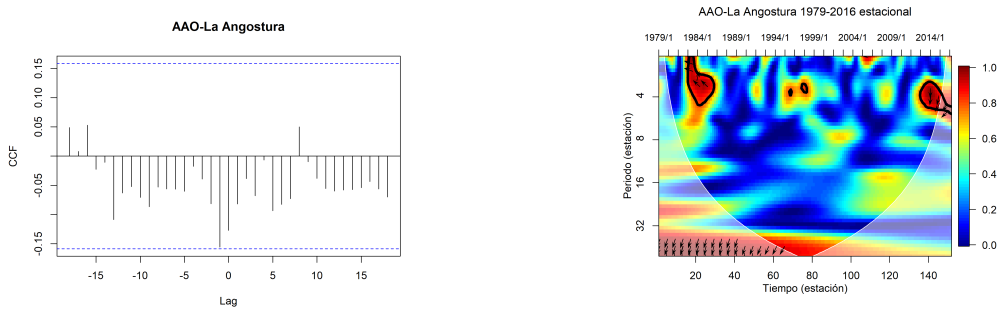
Figura 3.24: (a) Correlación Cruzada, (b) Wavelet coherencia entre AAO y Túnel Subfluvial.



(a) (b)

Figura 3.25: (a) Correlación Cruzada, (b) Wavelet coherencia entre AAO y Paso de los Libres.

No resultó posible establecer una diferencia de fase que resulte evidente entre el caudal de La Angostura y el índice. Es decir, mediante la correlación cruzada ningún trimestre resultó significativo, y mediante el wavelet coherencia no existió un periodo de ondas significativo que presentara una fase constante en un intervalo temporal considerable (Figura 3.26).



(a) (b)

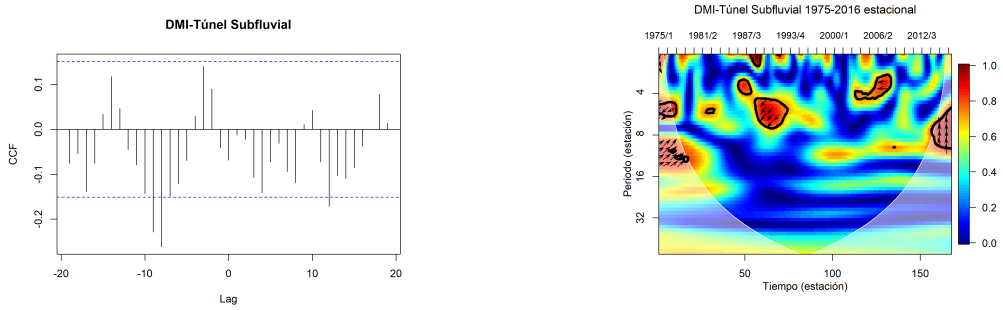
Figura 3.26: (a) Correlación Cruzada, (b) Wavelet coherencia entre AAO y La Angostura.

3.5.1.6. Índice Dipolo del Océano Índico-caudal

De la misma manera que el índice SAM y el AAO, el índice DMI y el caudal de Túnel Subfluvial no presentarían relación para trimestres cercanos (Figura 3.27a). Asimismo, el wavelet coherencia si bien destacó ciertos intervalos significativos (Figura 3.27b), los mismos ocurren solamente en ciertos intervalos de tiempo en conjunto con que la diferencia de fase no se mantiene constante. De acuerdo con lo anterior no sería este índice útil para poder realizar los objetivos propuestos en este capítulo.

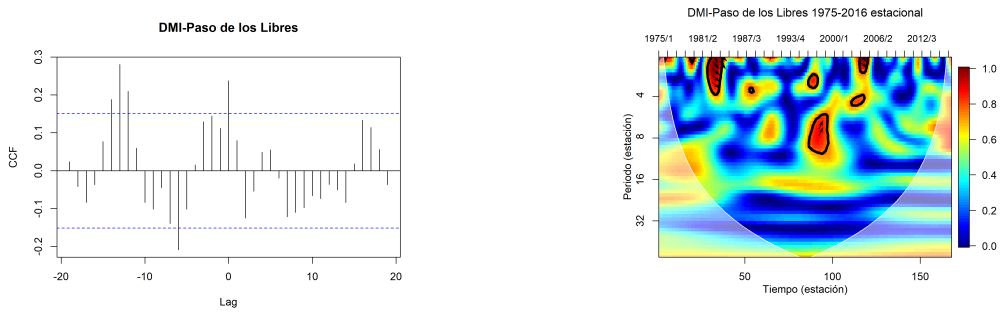
En la relación índice DMI y el caudal de Paso de los Libres se observó que el no desfasaje resultaba significativo, lo cual no resulta provechoso para realizar un pronóstico (Figura 3.28a). Además se destacaron otros lags significativos, sin embargo, se entiende que no resultarían útiles para el tipo de modelo que se pretende realizar. Asimismo, estos desfasajes que se observaron a partir de la metodología lineal, también fueron observables a través del wavelet coherencia para ciertos intervalos temporales significativos (Figura 3.28b).

Por último se analizó el par DMI-caudal La Angostura. A partir de la correlación cruzada se infiere una relación positiva y significativa a una estación y también para cuatro estaciones (Figura 3.29a). Mediante wavelet coherencia o bien podría pensarse que para



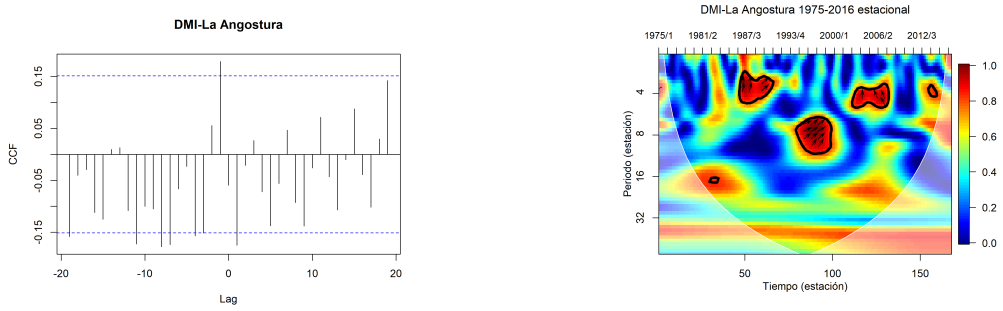
(a) (b)

Figura 3.27: (a) Correlación Cruzada, (b) Wavelet coherencia entre DMI y Túnel Subfluvial.



(a) (b)

Figura 3.28: (a) Correlación Cruzada, (b) Wavelet coherencia entre DMI y Paso de los Libres.



(a) (b)

Figura 3.29: (a) Correlación Cruzada, (b) Wavelet coherencia entre DMI y La Angostura.

ciertos periodos la diferencia de fase entre las variables sería alrededor de 1 estación o una relación en contra fase de aproximadamente 4 estaciones (Figura 3.29b). Más aún, teniendo en cuenta el coeficiente de correlación de Spearman con un corrimiento de un trimestre se obtuvo un valor que no resultó significativo. En el caso de considerar 4 estaciones el valor de correlación si bien en la metodología lineal resultaba significativo, en este caso para la correlación de Spearman no lo fue. Puesto que este índice con la estación La Angostura presentaron mayor evidencia de relación y posible pronóstico frente a los otras estaciones, se evaluará si resulta posible encontrar una función bivariada a través de la metodología de cópulas que describa la relación con este último desfasaje (Figura 3.30).

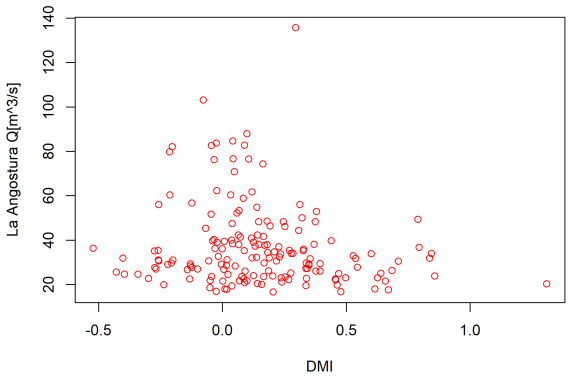


Figura 3.30: Diagrama de dispersión entre La Angostura y DMI con desfasaje igual a cuatro Δt.

A modo de resumen se presenta a continuación la Tabla I. En la misma se exhibe el desfasaje que se ha tenido en cuenta entre el par índice/caudal de cada estación de medición.

Resumen de índices:

Tabla I: *Resumen de desfasajes de los pares índice-caudal de cada estación de medición.*

Índice & Estación			
Índice/Estación	Túnel Subfluvial	Paso de los Libres	La Angostura
NIÑO 3.4	$2\Delta t$	$2\Delta t$	$1\Delta t$
SOI	$2\Delta t$	$1\Delta t$	$2\Delta t$
PDO	$1\Delta t/2\Delta t$	$1\Delta t/2\Delta t$	$3\Delta t$
SAM	$---$	$3\Delta t$	$---$
AAO	$---$	$---$	$---$
IOD	$---$	$---$	$4\Delta t$

3.5.2. Estudio de las relaciones bivariadas mediante cópulas

La principal ventaja que se obtiene a partir del método de cópulas, para representar la dependencia entre dos variables, es que el procedimiento puede ser realizado independientemente de las distribuciones marginales que cada variable presente (Genest y Favre, 2007). En este sentido, de acuerdo a los desfasajes previamente establecidos se procedió a analizar mediante el algoritmo propuesto por Schepsmeier et al. (2018) la selección de la mejor cópula entre varias familias (superior a treinta). El algoritmo emplea un test de hipótesis que evalúa la dependencia entre las variables a través de Kendall Tau.

De acuerdo a los índices y los caudales estacionales que presentaron alguna relación desfasada temporalmente se obtuvieron las familias de cópulas que a continuación se detallan. Es decir, en la relación entre el NIÑO 3.4 y el caudal estacional de las tres estaciones consideradas, Túnel Subfluvial, Paso de los Libres y La Angostura, mediante el criterio de AIC que considera el algoritmo, se seleccionó la cópula Joe (Ec. 3.4).

$$C_\theta(u,v) = 1 - [(1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta]^{\frac{1}{\theta}} \tag{3.4}$$

con u, v variables uniformes, y $\theta > 1$ el parámetro estimado.

Por otra parte, la relación entre el índice SOI y el resto de los caudales estacionales no

arrojó la misma familia de cópulas. En el caso de la estación Túnel Subfluvial, la familia rotada 90° de Joe (Ec. 3.5) fue la obtenida. Para Paso de los Libres la familia rotada 270° de Clayton (Ec. 3.6), mientas que para la estación La Angostura fue la rotada 270° de Gumbel (Ec. 3.7).

$$C_{\theta}^{90}(u,v) = 1 - [v^{\theta} + (1-u)^{\theta} - (v)^{\theta}(1-u)^{\theta}]^{\frac{1}{\theta}} \tag{3.5}$$

con u, v variables uniformes, y $\theta < -1$ el parámetro estimado.

$$C_{\theta}^{270}(u,v) = [\max v^{-\theta} + (1-u)^{-\theta} - 1; 0]^{-\frac{1}{\theta}} \tag{3.6}$$

con u, v variables uniformes, y $\theta < 0$ el parámetro estimado.

$$C_{\theta}^{270}(u,v) = [\exp[-((-\log(v))^{\theta} + (-\log(1-u))^{\theta})^{-\frac{1}{\theta}}] \tag{3.7}$$

con u, v variables uniformes, y $\theta \leq -1$ el parámetro estimado.

Al igual que con el índice climático anterior, el índice PDO tampoco presentó un único tipo de familia de cópula para todas las estaciones. Por ejemplo, en la estación Túnel Subfluvial se obtuvo la familia rotada 180° de tipo II de Tawn (Ec. 3.8). En el caso de Paso de los Libres se obtuvo la familia rotada BB7180°, Joe-Clayton cópula (Ec. 3.9), mientras que para La Angostura la familia rotada 180° de Joe (Ec. 3.10).

$$C_{\theta}^{180}(u,v) = \exp(log(1-u)^{1-\theta_1} - [(-\theta_1 log(1-u))^{\theta_2} + (-log(1-v))^{\theta_2}]^{\frac{1}{\theta_2}}) \tag{3.8}$$

con u, v variables uniformes, y $0 \leq \theta_1 \leq 1, \theta_2 \geq 1$ los parámetros estimados.

$$C_{\theta}^{180}(u,v) = 1 - (1 - [(1-u^{\theta})^{-\delta} + (1+v^{\theta})^{-\delta} - 1]^{-\frac{1}{\delta}})^{\frac{1}{\theta}} \tag{3.9}$$

con u, v variables uniformes; $\theta \geq 1$ y $\delta > 0$ los parámetros estimados.

$$C_{\theta}^{180}(u,v) = u + v - [u^{\theta} + v^{\theta} - u^{\theta}v^{\theta}]^{\frac{1}{\theta}} \tag{3.10}$$

con u, v variables uniformes, y $\theta > 1$ el parámetro estimado.

Como se había mostrado en la subsección anterior la relación entre el índice SAM y las distintas estaciones solo se había establecido para Paso de los Libres. En este sentido, la familia de cópula que se seleccionó fue la de Frank (Ec. 3.11).

$$C_\theta(u,v) = -\frac{1}{\theta} log[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1}] \tag{3.11}$$

con u, v variables uniformes, $\theta \in \mathbb{R} \setminus \{0\}$

Por otra parte, la selección de un modelo bivariado, a través de cópulas entre el índice DMI y el caudal estacional de La Angostura arrojó independencia entre las variables. Es decir, no fue posible establecer un modelo que relacione las dos variables a partir de lo establecido.

La Tabla II muestra el parámetro θ estimado para cada ajuste de la relación NIÑO 3.4-caudal, y el resultado de los test de bondad de ajuste para la función de cópulas. En ella se exhiben los p-valores para los tests Cramer-Von Mises (CVM) y Kolmogorov-Smirnov (K-S) obtenidos a través de reesampleo de Bootstrap. En el caso particular de la relación entre NIÑO 3.4 y el caudal de Túnel Subfluvial de acuerdo a los p-valores (0.21 y 0.39), la hipótesis nula no es rechazada, y la familia cópula de Joe ajusta a dicho par con un 95 % de confianza. El resto de los resultados obtenidos para los pares restantes de índices-caudales se exhiben en las Tablas III-V.

Tabla II: *Familia de cópula para el par NIÑO 3.4-caudal, parámetro estimado, p-valor de los test CVM y K-S.*

Familias de Cópulas para la relación NIÑO 3.4 y el caudal				
Función	Parámetro	Estación	p-valor CVM	p-valor K-S
Joe	1.55	Túnel Subfluvial	0.21	0.39
Joe	1.34	Paso de los Libres	0.31	0.24
Joe	1.2	La Angostura	0.57	0.62

Tabla III: *Familia de cópula para el par SOI-caudal, parámetro estimado, p-valor de los test CVM y K-S.*

Familias de Cópulas para la relación SOI y el caudal				
Función	Parámetro	Estación	p-valor CVM	p-valor K-S
Joe 90°	-1.50	Túnel Subfluvial	0.14	0.40
Clayton 270°	-0.64	Paso de los Libres	0.87	0.73
Gumbel 270°	-1.24	La Angostura	0.86	0.84

Tabla IV: *Familia de cópula para el par PDO-caudal, parámetro estimado, p-valor de los test CVM y K-S.*

Familias de Cópulas para la relación PDO y el caudal				
Función	Parámetro	Estación	p-valor CVM	p-valor K-S
Tawn2 180°	1.76 & 0.40	Túnel Subfluvial ($1\Delta t$)	–	0.25
Tawn2 180°	1.79 & 0.26	Túnel Subfluvial ($2\Delta t$)	–	0.14
BB7 180°	1.16 & 0.27	Paso de los Libres ($1\Delta t$)	0.87	0.91
Gumbel 180°	1.19	Paso de los Libres ($2\Delta t$)	0.26	0.25
Joe 180°	1.34	La Angostura	0.44	0.50

En el caso particular de la Familia Tawn2 el algoritmo para el cálculo del test de CVM no fue implementado. Razón por la cual no fue estimado.

Tabla V: *Familia de cópula para el par SAM-caudal, parámetro estimado, p-valor de los test CVM y K-S.*

Familias de Cópulas para la relación SAM y el caudal				
Función	Parámetro	Estación	p-valor CVM	p-valor K-S
Frank	1.34	Paso de los Libres	0.354	0.17

Asimismo, es importante notar que uno de los objetivos del capítulo era estudiar la probabilidad conjunta de eventos extremos en mayor profundidad, pero no todas las familias de cópulas son adecuadas para valores extremos. Por lo tanto sobre las cópulas que fueron seleccionadas hay que analizar con cierto cuidado si las mismas resultarían útiles para cumplir con uno de los objetivos propuestos en el presente capítulo.

3.5.3. Ajustes Marginales

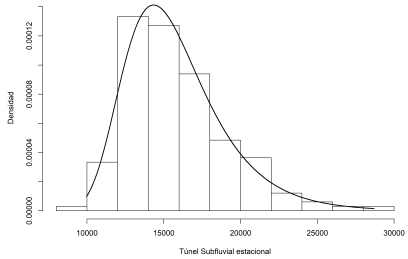
Para las series temporales que presentaron relaciones de a pares se realizaron los ajustes marginales individuales de cada serie. Como se describió en la sección de Datos y Metodologías varias funciones de distribución fueron tenidas en cuenta (Normal, Lognormal, Gumbel, Weibull, Gamma, GEV). Mediante el método de máxima verosimilitud fueron estimados los respectivos parámetros muestrales de cada distribución teórica propuesta, junto con un test de bondad de ajuste, K-S. A continuación se presentan las Tablas VI-XII, en ellas se muestran los valores de p-valor obtenidos para el test K-S mediante reesampleo de Bootstrap. Asimismo, para las funciones de distribución que no rechazan la hipótesis nula se exhibe el valor del criterio de AIC, de manera tal de seleccionar las funciones que menor valor presenten.

De acuerdo con la Tabla VI se observa que cuatro funciones de distribución teóricas de las propuestas no rechazaron la hipótesis nula del test de K-S de bondad de ajuste marginal del caudal estacional de Túnel Subfluvial (Lognormal, Gumbel, Gamma, GEV). De esta manera, para poder realizar una selección adecuada se empleó el criterio previamente mencionado en conjunto con un análisis de herramienta del tipo diagnóstico dada por el método de los cuantiles (QQ-PLOT). Por lo tanto, teniendo en cuenta un compromiso entre el criterio de selección de modelo AIC, y el método de cuantiles, el ajuste más adecuado sería mediante la función Gumbel, con ciertos parámetros estimados a partir del método de máxima verosimilitud (Tabla XIII). Además, en las Figuras 3.31a y 3.31b se graficó un histograma con los datos observacionales en conjunto con la función de densidad teórica propuesta, y la distribución teórica de cuantiles contra la distribución empírica de cuantiles, con un 5 % de significancia (Fox y Weisberg, 2018), respectivamente. En ambas figuras puede verificarse que el ajuste es coherente.

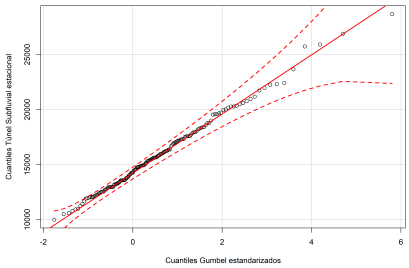
Tabla VI: *Funciones de distribución teóricas, p-valor Bootstrap y AIC para Túnel Subfluvial.*

Funciones de Distribución para Q estacional de Túnel Subfluvial		
Función	p-valor Bootstrap	AIC
Normal	0.003	− − −
Lognormal	0.299	3144.929
Gumbel	0.62	3113.537
Weibull	0	− − −
Gamma	0.20	3132.706
GEV	0.575	3123.522

A partir de la Tabla VII es posible observar que en el caso del ajuste marginal del caudal estacional de la estación Paso de los Libres tanto las funciones teóricas Gumbel, Weibull y GEV pueden ser consideradas, puesto que, su respectivo p-valor resulta ser superior a 0.20. Sin embargo, al realizar un compromiso entre el criterio de selección de modelos, y el análisis diagnóstico de cuantiles se decidió utilizar la distribución teórica de Gumbel para el ajuste marginal de los datos muestrales provenientes de esta estación de medición. Esta función fue la que se tuvo en cuenta para los cálculos posteriores, sus parámetros



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



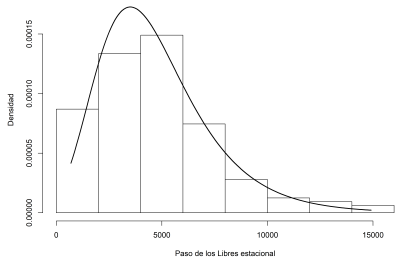
(b) QQ-PLOT.

Figura 3.31: Análisis de bondad de ajuste marginal del caudal estacional de Túnel Subfluvial.

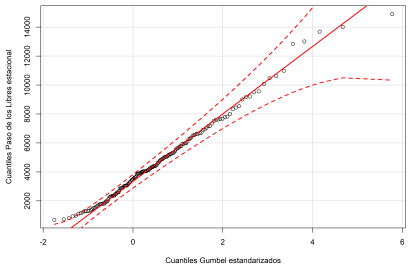
estimados se exhiben en la Tabla XIII. Al igual que en el análisis anterior se exhibe el histograma de los datos estacionales de caudal, junto con la función de densidad de Gumbel, como así también el gráfico de cuantiles (Figuras 3.32). También es posible verificar la coherencia del ajuste.

Tabla VII: *Funciones de distribución teóricas, p-valor Bootstrap y AIC, para Paso de los Libres.*

Funciones de Distribución para Q estacional de Paso de los Libres		
Función	p-valor Bootstrap	AIC
Normal	0.001	− − −
Lognormal	0.07	− − −
Gumbel	0.52	2996.281
Weibull	0.913	3002.703
Gamma	0	− − −
GEV	0.412	3005.628



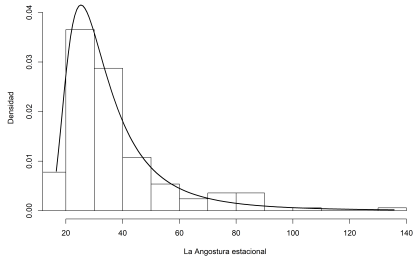
(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



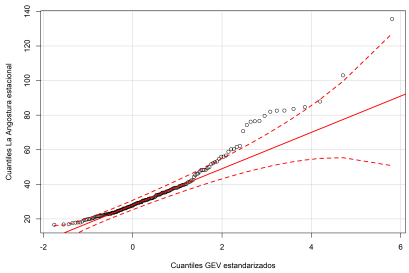
(b) QQ-PLOT.

Figura 3.32: Análisis de bondad de ajuste marginal del caudal estacional de Paso de los Libres.

Asimismo, el caudal asociado a la estación La Angostura, de las seis funciones de distribución solamente la función de GEV fue la que no rechazó la hipótesis nula del test K-S, es decir, la muestra proviene de la distribución propuesta con un 5 % de significancia (Tabla VIII), en la Tabla XIII se muestran los valores estimados de los parámetros. Por otra parte, si bien no se rechazó la hipótesis nula es posible observar a través de la Figura 3.33b (gráfico de cuantiles teóricos contra observados) que no estaría resultando un buen ajuste para los mayores cuantiles, pero sí para la cola inferior de la distribución. También se muestra el histograma de los datos observados en conjunto con la curva de densidad de probabilidad de la función teórica GEV (Figura 3.33a), donde también es posible observar lo recientemente mencionado.



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



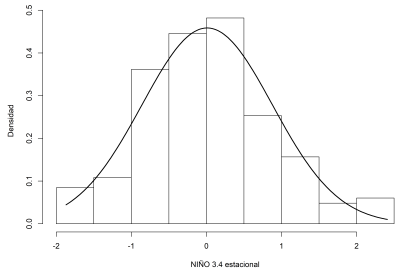
(b) QQ-PLOT.

Figura 3.33: Análisis de bondad de ajuste marginal del caudal estacional de La Angostura.

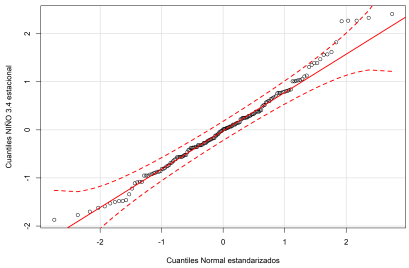
Tabla VIII: *Funciones de distribución teóricas, p-valor Bootstrap y AIC, para La Angostura.*

Funciones de Distribución para Q estacional de La Angostura	
Función	p-valor Bootstrap
Normal	0
Lognormal	0.031
Gumbel	0.014
Weibull	0
Gamma	0
GEV	0.88

Con respecto a los índices climáticos primeramente se realizó el ajuste marginal para el NIÑO 3.4. Los resultados se presentan en la Tabla IX, a partir de la misma es posible notar que no todas las funciones de distribución resultan posibles de utilizar, puesto que el índice climático presenta no solo valores positivos sino también negativos. Por lo tanto, las funciones teóricas como Normal, Gumbel y GEV fueron las que se tuvieron en cuenta para el ajuste marginal. Tanto la distribución teórica Normal y GEV no rechazaron la hipótesis nula con un 5 % de significancia, los respectivos valores de las estimaciones de los parámetros se presentan en la Tabla XIII. Sin embargo, teniendo en cuenta el criterio de selección de modelo, la distribución de GEV sería la más adecuada. De todas maneras se tendrá en cuenta ambos ajustes. Es decir, para la obtención de la función de densidad de probabilidad conjunta mediante el método de cópulas se analizarán dos modelos. Puesto que, si los resultados a obtener son similares resultaría más conveniente considerar la función de distribución Normal debido a menor cantidad de parámetros a estimar, en conjunto con que resulta ser una función más conocida y más rápida de interpretar en el caso de una posible implementación del tipo operativa. El histograma de los datos estacionales en conjunto con la función de densidad de probabilidad, y el gráfico de cuantiles teóricos versus los empíricos se exhiben en las Figuras 3.34 y 3.35, para la distribución Normal y GEV respectivamente. En las Figuras de ambos ajustes es posible notar que ciertos valores extremos no están dentro del intervalo de confianza del 95 %.

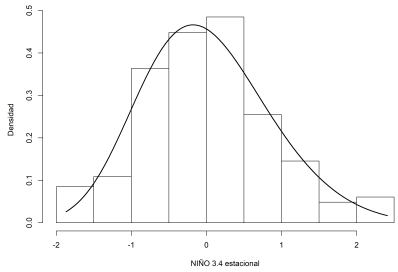


(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.

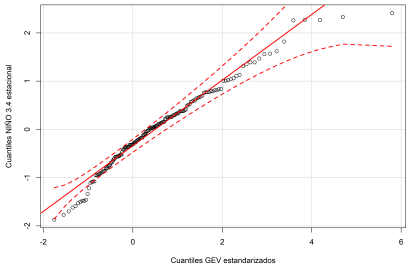


(b) QQ-PLOT.

Figura 3.34: Análisis de bondad de ajuste marginal Normal del índice NIÑO 3.4.



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



(b) QQ-PLOT.

Figura 3.35: Análisis de bondad de ajuste marginal GEV del índice NIÑO 3.4.

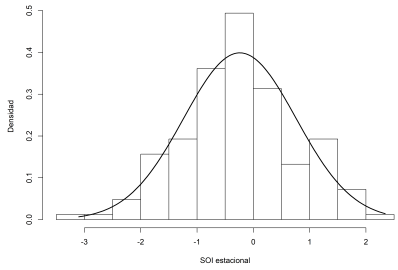
Tabla IX: *Funciones de distribución teóricas, p-valor Bootstrap y AIC para NIÑO 3.4.*

Funciones de Distribución para el índice NIÑO 3.4		
Función	p-valor Bootstrap	AIC
Normal	0.729	421.757
Lognormal	– – –	– – –
Gumbel	0	– – –
Weibull	– – –	– – –
Gamma	– – –	– – –
GEV	0.718	407.547

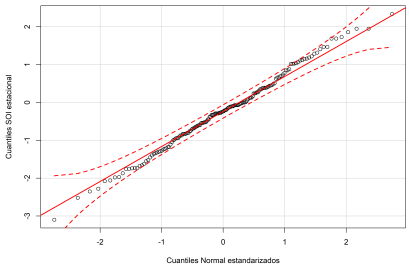
En el caso del ajuste marginal de los valores estacionales muestrales correspondientes al índice SOI, dos distribuciones se ajustarían con un 5 % de significancia, Normal y GEV (Tabla X). Sin embargo, teniendo en cuentas los elementos de selección de modelos ya previamente mencionados se consideró el modelo asociado a una distribución Normal el más adecuado. Los parámetros estimados de la distribución se presentan en la Tabla XIII. Además se exhibe en las Figuras 3.36a y 3.36b el histograma de frecuencia de los datos observados en conjunto con la curva de densidad de probabilidad de la distribución Normal con los parámetros correspondientes, y el gráfico de cuantiles de los datos empíricos contra los teóricos. De ambas herramientas es posible observar un ajuste adecuado.

Tabla X: *Funciones de distribución teóricas, p-valor Bootstrap y AIC, para SOI.*

Funciones de Distribución para el índice SOI		
Función	p-valor Bootstrap	AIC
Normal	0.729	475.1676
Gumbel	0.014	– – –
GEV	0.718	476.6847



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



(b) QQ-PLOT.

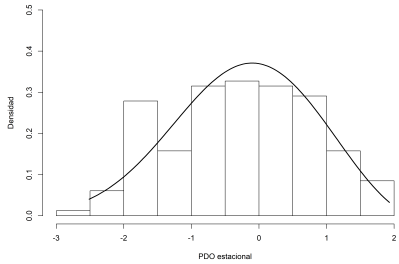
Figura 3.36: Análisis de bondad de ajuste marginal del índice SOI.

El ajuste marginal de los datos muestrales estacionales del índice PDO se observa en la Tabla XI. De los ajustes considerados, la única que no rechazo la hipótesis nula con un 5 % de significancia fue la distribución GEV. Es importante destacar que el método de

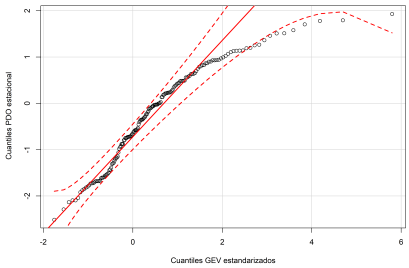
diagnóstico de cuantiles evidenció ciertas observaciones que caerían por fuera del intervalo de confianza del 95 % sobre todo en la cola superior (Figura 3.37a y 3.37b). Esto fue tenido en cuenta para los cálculos posteriores que infirieran sacar conclusiones sobre la cola superior del índice.

Tabla XI: *Funciones de distribución teóricas, p-valor Bootstrap y AIC, para PDO.*

Funciones de Distribución para el índice PDO		
Función	p-valor Bootstrap	AIC
Normal	0.005	– – –
Gumbel	0.037	– – –
GEV	0.22	– – –



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



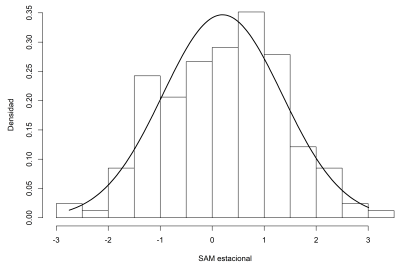
(b) QQ-PLOT.

Figura 3.37: Análisis de bondad de ajuste marginal del índice PDO.

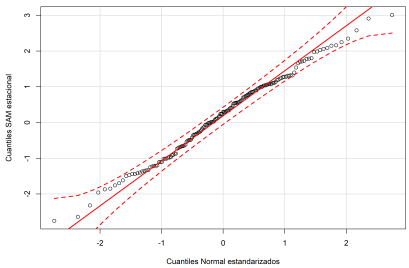
El análisis del ajuste marginal de los datos estacionales del índice SAM arrojó que la distribución Normal sería la más adecuada (Tabla XII). Si bien tanto la distribución Normal como la distribución GEV no rechazan la hipótesis nula del test K-S considerando reesampleo de Bootstrap, la distribución normal presentó un compromiso más óptimo frente al método de selección y diagnóstico. De este último se muestran el histograma de la distribución de la muestra en conjunto con la curva de densidad teórica, y los gráficos de cuantiles, coincidente con lo dicho previamente (Figura 3.38). También, los valores estimados de los parámetros de la distribución son expuestos en la Tabla XIII

Tabla XII: *Funciones de distribución teóricas, p-valor Bootstrap y AIC, para SAM.*

Funciones de Distribución para el índice SAM		
Función	p-valor Bootstrap	AIC
Normal	0.245	– – –
Gumbel	0.002	– – –
GEV	0.367	– – –



(a) Histograma de los datos observacionales en conjunto con la función de densidad de probabilidad teórica.



(b) QQ-PLOT.

Figura 3.38: Análisis de bondad de ajuste marginal del índice SAM.

Tabla XIII: *Valores de los parámetros estimados de los ajustes marginales, junto con sus desvíos estándar.*

Serie	Densidad teórica	Parámetros estimados	SE
Túnel Subfluvial	*Gumbel*	μ=14339.17 σ=2604.48	μ=212.60 σ=159.08
Paso de los Libres	*Gumbel*	μ=3505.55 σ=2129.72	μ=175.82 σ=134.36
La Angostura	*GEV*	μ=27.82 σ=9.31 ψ=0.33	μ=0.84 σ=0.72 ψ=0.08
NIÑO 3.4	*Normal*	μ=0.01 σ=0.86	μ=0.07 σ=0.05
NIÑO 3.4	*GEV*	μ=-0.33 σ=0.80 ψ=-0.17	μ=0.07 σ=0.04 ψ=0.04
SOI	*Normal*	μ=-0.23 σ=1	μ=0.08 σ=0.05
PDO	*GEV*	μ=-0.56 σ=1.07 ψ=-0.37	μ=0.09σ=0.07 ψ=0.06
SAM	*Normal*	μ=0.20 σ=1.15	μ=0.09 σ=0.06

Los parámetros μ y σ en el caso de la distribución teórica *Gumbel* representan el parámetro de locación y de forma. Por otro lado, los parámetros μ y σ para la distribución *Normal* representan el valor medio y desvío estándar, respectivamente. Finalmente, en el caso de la distribución teórica *GEV*, μ representa el parámetro de locación, σ el de escala y ψ el de forma.

3.5.4. Evaluación del modelo

Para validar los modelos obtenidos se consideraron diferentes procedimientos. A continuación, se mostrarán los resultados para Túnel Subfluvial-NIÑO 3.4 y se detallará para otras series temporales al final de la subsección. Primero, para evaluar el rendimiento del modelo fuera de los datos de calibración, la serie fue separada en dos grupos uno de entrenamiento y otro de testeo, un 70 % y un 30 % de los datos respectivamente. Se ajustó el algoritmo de selección de modelo de cópula con los datos de entrenamiento y luego se realizó un test de bondad de ajuste del modelo calibrado para ambos conjuntos de datos. Como se describió en la sección de metodología siempre se utilizaron los test de Cramer-Von Mises (CVM) y Kolmogorov-Smirnov (K-S) para bondad de ajuste. Para el procedimiento de evaluación se utilizaron dos enfoques distintos: uno el que se realizó en Khendun et al. (2014), separando las series de manera aleatoria utilizando el 70 % de

los datos para entrenamiento y el 30 % restante para testeo. Se calibró la cópula para los datos de entrenamiento, que resultó ser la familia de Joe, con θ=1.56, y se realizaron los tests de bondad de ajuste. En todos los casos la hipótesis nula no fue rechazada, indicando que el modelo propuesto ajusta a ambas series. Los p-valores obtenidos fueron 0.2 y 0.38 para el test de Cramer-Von Mises (CVM); 0.2 y 0.44 para el test de Kolmogorov-Smirnov (K-S) (conjunto de datos de entrenamiento y testeo, respectivamente). Asimismo, en el segundo método de evaluación se repitió el procedimiento de separar las series pero esta vez conservando su secuencia temporal. Es decir, el primer 70 % de los datos fue tenido en cuenta para entrenamiento, y se conservó el 30 % restante para testeo. Nuevamente, se obtuvo la familia cópula de Joe (con θ=1.63 en los datos de entrenamiento, y la hipótesis nula de los tests de bondad de ajuste no fue rechazada para ninguno de los casos, arrojando p-valores iguales a 0.2 y 0.37 para el test de CVM, mientras que para el test de K-S fueron de 0.25 y 0.36 (entrenamiento y testeo respectivamente). Es importante notar que este segundo enfoque da lugar a una evaluación fuera del periodo de entrenamiento. Los procedimientos anteriores fueron repetidos para la estación Paso de los Libres y La Angostura. En el caso de Paso de los Libres, la evaluación del modelo solo pudo ser realizada al separar a la serie de manera aleatoria (modelado por la familia de cópula Joe con θ=1.39). Esto puede estar relacionado con una menor señal del ENSO a través de este índice en la estación del río Uruguay en comparación a la cuenca del Paraná, a su vez combinado con una cantidad no suficiente de datos. En el caso de la separación temporal, el algoritmo propuesto no pudo encontrar una función de cópula adecuada para ajustar el conjunto de datos de entrenamiento. Por otra parte en el caso de la estación La Angostura no se logró validar el modelo que relacionaba esta estación de medición con el índice NIÑO 3.4, probablemente debido a una relación no tan estrecha y a un volumen no suficiente de datos.

En las Tablas que a continuación se presentan se detallan los p-valores obtenidos para el resto de los ajustes considerados para evaluar la relación entre los índices y el caudal tanto para el conjunto de datos de entrenamiento como de testeo en la secuencia aleatoria (Tabla XIV-XV), como así también de secuencia temporal (Tabla XVI-XVII). En el caso del índice SOI las familias fueron validadas para las tres estaciones teniendo en cuenta ambos métodos. En el caso de la relación entre la PDO y el caudal de Túnel Subfluvial y Paso de los Libres la evaluación de dichos modelos no pudo ser realizada porque se

encontró que las familias de cópulas o bien no fueron las mismas que fueron seleccionadas en un principio o bien se obtenía directamente independencia entre las variables estudiadas. Para la PDO y La Angostura el modelo si pudo ser validado de manera aleatoria y los resultados de los test se encuentran en la Tabla XIV-XV. En la relación entre SAM y Paso de los Libres tampoco pudo ser realizada la evaluación del modelo.

En conclusión, a partir de los resultados anteriores, solo algunas relaciones pudieron ser validadas teniendo en cuenta la secuencia temporal, que dado dicho resultado serán estudiadas con fines de pronóstico. Las que solo pudieron ser validadas de manera aleatoria se tendrán en cuenta para la obtención de ciertas métricas pero no con fines de predicción, principalmente, por encontrar que ciertas evaluaciones de los modelos no resultaron válidas temporalmente para las series. Es decir, esto indica que a pesar de que el modelo puede ser calibrado para todo el conjunto de datos, y que su calibración fue validada para particiones aleatorias, no se cuenta con gran cantidad de datos para realizar la evaluación que conserva el periodo temporal, y de esta manera sería aconsejable no generar pronóstico para estas series.

Tabla XIV: *Evaluación del conjunto de datos de Entrenamiento Aleatorio.*

Relación	Familia	Parámetro	p-valor CVM	p-valor K-S
SOI-Túnel Subfluvial	JOE 90	θ=-1.54	0.36	0.49
SOI-Paso de los Libres	CLAYTON 270	θ=-0.63	0.91	0.89
SOI-La Angostura	GUMBEL 270	θ=-1.24	0.84	0.91
PDO-La Angostura	JOE 180	θ=-1.4	0.56	0.59

Tabla XV: *Evaluación del conjunto de datos de Testeo Aleatorio.*

Relación	Familia	Parámetro	p-valor CVM	p-valor K-S
SOI-Túnel Subfluvial	JOE 90	θ=-1.54	0.20	0.23
SOI-Paso de los Libres	CLAYTON 270	θ=-0.63	0.77	0.60
SOI-La Angostura	GUMBEL 270	θ=-1.24	0.22	0.33
PDO-La Angostura	JOE 180	θ=-1.4	0.32	0.34

Tabla XVI: *Evaluación del conjunto de datos de Entrenamiento Secuencia Temporal.*

Relación	Familia	Parámetro	p-valor CVM	p-valor K-S
SOI-Túnel Subfluvial	JOE 90	θ=-1.49	0.20	0.13
SOI-Paso de los Libres	CLAYTON 270	θ=-0.64	0.91	0.88
SOI-La Angostura	GUMBEL 270	θ=-1.22	0.63	0.64

Tabla XVII: *Evaluación del conjunto de datos de Testeo Secuencia Temporal.*

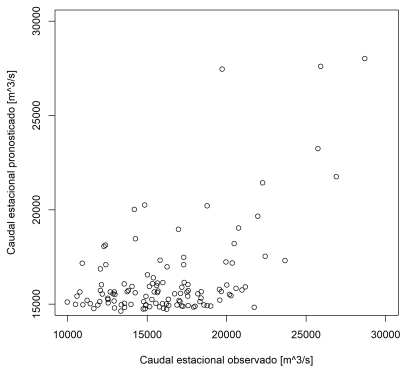
Relación	Familia	Parámetro	p-valor CVM	p-valor K-S
SOI-Túnel Subfluvial	JOE 90	θ=-1.49	0.40	0.21
SOI-Paso de los Libres	CLAYTON 270	θ=-0.64	0.75	0.61
SOI-La Angostura	GUMBEL 270	θ=-1.22	0.82	0.87

3.5.5. Pronóstico desde la distribución conjunta

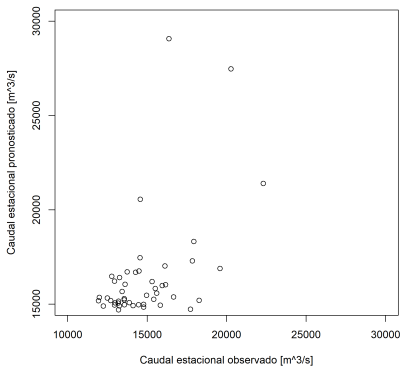
Mientras se realizaba el ajuste de la distribución bivariada a los pares de datos, siempre se consideró que el caudal estacional era la variable respuesta y los valores del índice la variable explicativa. Por lo tanto, el modelo propuesto tiene sentido utilizarlo como herramienta de pronóstico de la siguiente manera: una vez que se observa el valor del índice, se condiciona la distribución conjunta a ese valor, y se utiliza la distribución de caudal obtenida para predecir métricas de interés. Puesto que la serie temporal del índice se encuentra desfasada respecto a la serie de caudal, los resultados obtenidos resultan ser pronósticos con ese desfasaje como horizonte.

Una manera simple de predecir desde la distribución es calcular el valor esperado. Siguiendo esta idea, se generó el siguiente procedimiento para predicción: dado un valor del índice, se muestreó de manera aleatoria desde la distribución condicional obtenida del modelo y la observación. Para esto, a partir de la Ec.(3.2) se muestreó condicionalmente desde la cópula, y se consideraron las distribuciones marginales correspondientes para realizar la transformación del dominio de la serie, al dominio uniforme de la cópula o viceversa, como se encuentra descripto en la sección metodológica. Luego, la media condicional de caudal dada la observación del índice fue estimada promediando las muestras obtenidas de la distribución condicional. Se utilizó esta cantidad como un valor predictivo para el caudal. Para validar estas predicciones, se separó nuevamente a las series temporales, utilizando el primer 70 % (datos de entrenamiento) para calibrar el modelo.

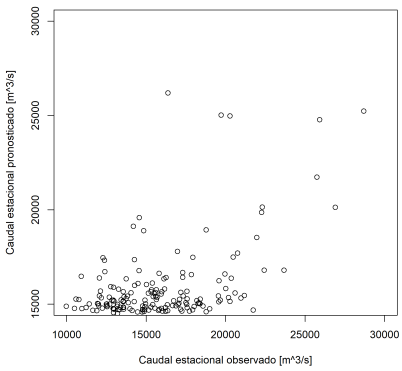
Se generaron valores predictivos del caudal considerando el procedimiento anterior, para toda la serie y se comparó con los valores observados de ambos conjuntos de datos (entrenamiento y testeo). En ambos casos se computó la correlación considerando Pearson y Spearman entre los valores observados y los pronosticados. Los resultados mostraron que la correlación en la serie de entrenamiento fue de 0.54 y 0.32, respectivamente (p-valores iguales a 3.789e-10 y 0.0005). En el caso de la serie de testeo los valores fueron de 0.52 y 0.37, respectivamente (p-valores iguales a 0.0002 y 0.01). Ambos casos para Túnel Subfluvial. Se exhiben gráficos de dispersión para los valores observados y pronosticados en la Figura 3.39. Más aún, el mismo análisis de correlación fue realizado para los datos que se correspondían con valores del NIÑO 3.4 por encima y por debajo de su mediana de manera separada. Los resultados mostraron una fuerte relación para la fase positiva del índice y prácticamente nula para la fase negativa.



(a) Datos de Entrenamiento.



(b) Datos de Testeo.



(c) Todo el conjunto de datos.

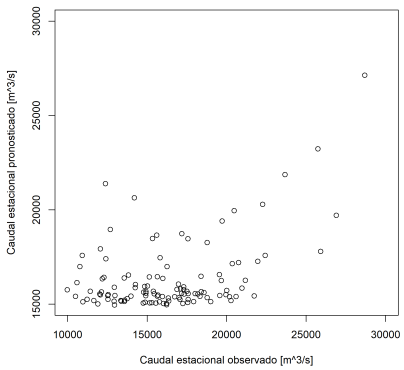
Figura 3.39: Diagramas de dispersión entre los pares observados y pronosticados para la estación Túnel Subfluvial (índice NIÑO 3.4).

Por otra parte, si se considera el procedimiento anterior pero teniendo en cuenta el ajuste marginal de distribución GEV para los datos del índice se destaca que los resultados de los valores de correlación son similares tanto para el conjunto de entrenamiento como de testeo en el estudio de la relación NIÑO 3.4 y Túnel Subfluvial. Con lo cual para modelar esta relación se consideró la distribución marginal Normal para el índice porque además de que los resultados fueron similares, resulta ser más simple esta distribución marginal que la distribución GEV.

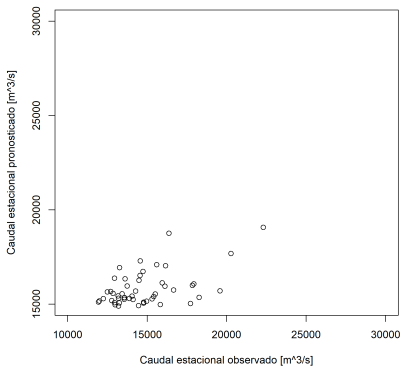
Asimismo, la correlación para el conjunto de datos de entrenamiento entre los valores pronosticados y observados para el caudal a partir del modelo entre el índice SOI y Túnel Subfluvial fue de 0.45 y 0.21, respectivamente, (p-valores iguales a 3.668e-07 y 0.02), mientras que para el caso de la muestra de testeo fue de 0.51 para el coeficiente de Pearson y 0.37 para el coeficiente de Spearman, con p-valores iguales a 0.0001 y 0.007, respectivamente. Considerando todo el conjunto de datos resultó ser de 0.49 y 0.25 con p-valores iguales a 3.296e-11 y 0.001, respectivamente. Todas las correlaciones fueron significativas con al menos un 5 % de significancia. Además, se exhibe en la Figura 3.40 dichos pares de datos estudiados, donde es posible notar la estrecha relación entre los datos de entrenamiento, testeo, y considerando todo el conjunto de datos.

En el caso de estudio entre el índice SOI y la estación Paso de los Libres se obtuvo una correlación de los datos de entrenamiento para Pearson igual a 0.48 y 0.32 para Spearman con p-valores iguales a 7.009e-08 y 0.0004, respectivamente. En el caso de los datos de testeo los valores de correlación fueron iguales a 0.43 y 0.38, con p-valores iguales a 0.001 y 0.00719, respectivamente. Si se tuviesen en cuenta todo el conjunto de datos el valor de correlación obtenida fue de 0.47 (p-valor igual a 1.345e-10) y 0.35 (p-valor igual a 3.054e-06). En la Figura 3.41 se exhiben dichas relaciones. Es importante destacar que si bien el índice NIÑO 3.4 y SOI son utilizados para el monitoreo del fenómeno ENSO en el caso particular de la estación Paso de los Libres este modelo pudo ser validado de manera aleatoria y manteniendo la secuencia temporal como se mostró al principio de la sección. Por lo que el estudio de pronóstico bivariado se realizó en este caso arrojando resultados favorables como los expuestos previamente.

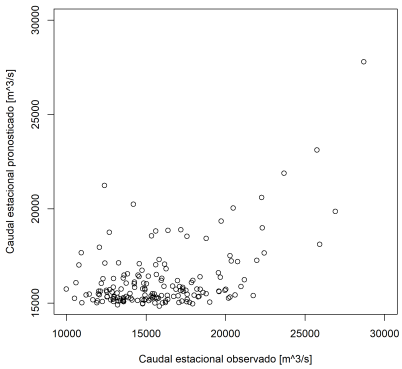
Para el estudio de La Angostura, los valores de correlación observados y esperados tanto para el conjunto de entrenamiento como de testeo presentaron menores valores que los resultados anteriores. En el caso de los datos de entrenamiento los valores fueron iguales



(a) Datos de Entrenamiento.

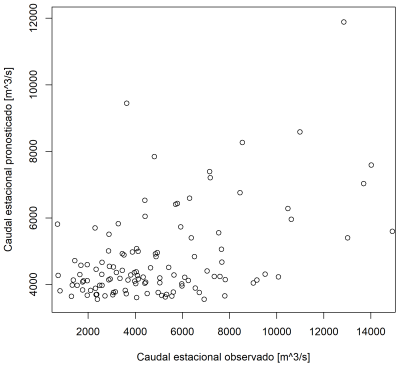


(b) Datos de Testeo.

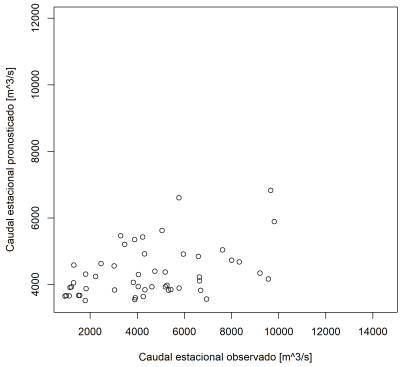


(c) Todo el conjunto de datos.

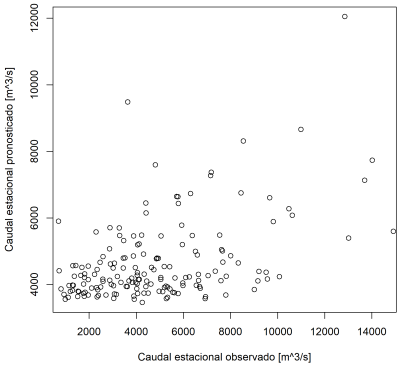
Figura 3.40: Diagramas de dispersión entre los pares observados y pronosticados para la estación Túnel Subfluvial (índice SOI)).



(a) Datos de Entrenamiento.



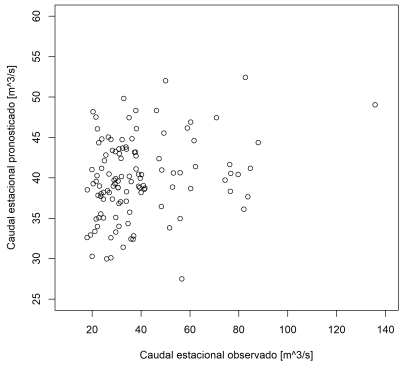
(b) Datos de Testeo.



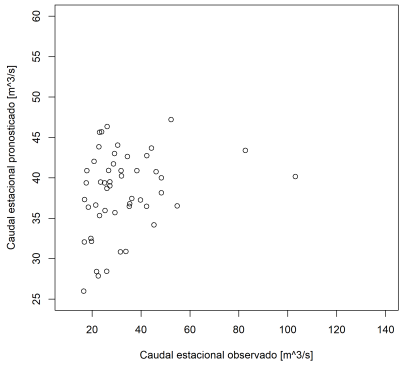
(c) Todo el conjunto de datos.

Figura 3.41: Diagramas de dispersión entre los pares observados y pronosticados para la estación Paso de los Libres.

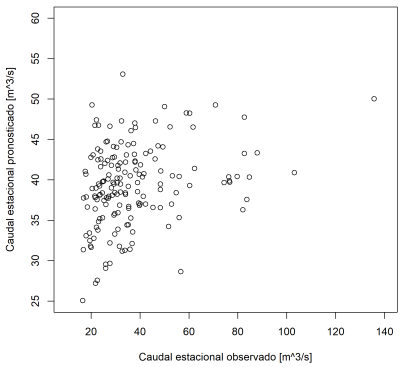
a 0.28 (p-valor igual a 0.004) para la correlación de Pearson, y de 0.24 (p-valor igual a 0.008) para la correlación de Spearman. Los datos de testeo presentaron valores de correlación igual a 0.26 y 0.28 con p-valores iguales a 0.04 y 0.04, respectivamente. Por otra parte, al tener en cuenta todo el conjunto de datos el valor de correlación obtenida fue de 0.28 y 0.27, con p-valores iguales a 0.0002 y 0.0004, respectivamente. En la Figura 3.42 se presentan las relaciones estudiadas donde es posible notar una relación levemente más estrecha entre menores valores de caudal que entre valores altos.



(a) Datos de entrenamiento.



(b) Datos de Testeo.



(c) Todo el conjunto de datos.

Figura 3.42: Diagramas de dispersión entre los pares observados y pronosticados para la estación La Angostura.

Además, un segundo enfoque fue explorado para el pronóstico, más focalizado en la predicción de eventos extremos para el caudal. Se siguió el mismo procedimiento anterior: una vez que se observó un valor del índice, se utilizó el modelo propuesto para estimar la distribución condicional para el correspondiente valor de caudal. En este caso, en vez de computar el valor esperado y utilizarlo como pronóstico, la idea fue estimar cuantiles de

la distribución condicional, y utilizar los mismos como límites de intervalo de confianza para el caudal. Por ejemplo, se puede esperar que el caudal no sea superior que el cuantil correspondiente al 95 % de la distribución condicional con un 95 % de confianza. En este caso, el modelo fue ajustado al periodo completo de las series, y siguiendo el mismo procedimiento que en el caso previo, se generaron muestras de la distribución condicional para cada observación del índice. Se estimaron los cuantiles 5 % y 95 % de la distribución condicional de esas muestras, generando intervalos de confianza (Figura 3.43) para la serie de Túnel Subfluvial. Se concluyó con umbrales por encima y por debajo para el caudal, inferidos a través de la serie del índice. El resultado teórico que debe quedar en claro es que el caudal debería estar por debajo del umbral superior con un 95 % de confianza, y por encima del umbral inferior también con un 95 % de confianza. Para poder evaluar la exactitud de la calibración de las series se ha realizado, para cada valor desde 5 % al 95 % en pasos de 5 % cual fue la proporción de las observaciones del caudal que estuvieron por debajo del cuantil estimado. Los resultados pueden ser observados en la Figura 3.44 para Túnel Subfluvial, junto con la proporción teórica esperada.

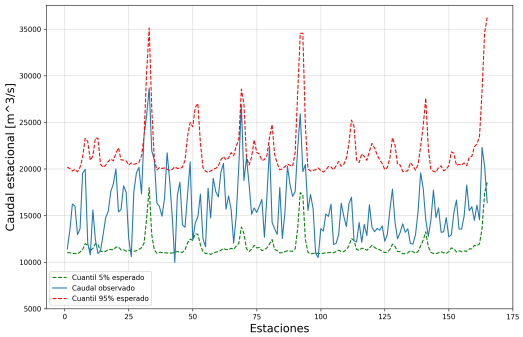


Figura 3.43: Series de cuantiles 5 % y 95 % para la distribución del caudal de Túnel Subfluvial condicionada al índice NIÑO 3.4 observado, junto con el caudal observado.

Mismo procedimiento fue realizado para las estaciones Túnel Subfluvial, Paso de los Libres y La Angostura pero teniendo en cuenta el índice SOI. En el caso de Túnel Subfluvial es posible notar a partir de las Figuras 3.45 y 3.46 resultados parecidos a los obtenidos por el índice NIÑO 3.4, notando una cierta sobreestimación en el cálculo del percentil empírico para valores extremos de cola derecha. Sin embargo, debido a este parecido en el resultado se empleará otra métrica hacia el final de la subsección en forma de compa-

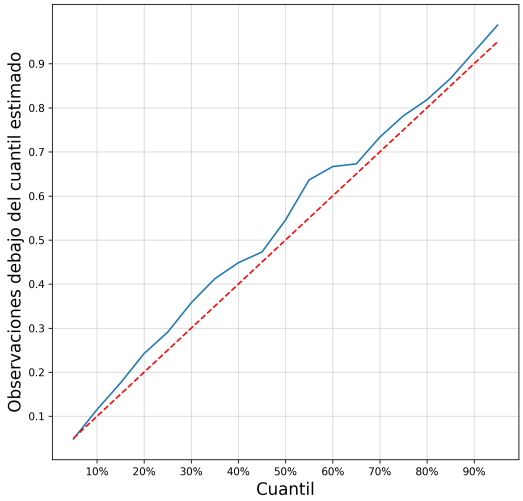


Figura 3.44: Proporción de observaciones que caen por debajo del cuantil estimado. Relación Túnel Subfluvial-NIÑO 3.4.

ración entre los dos modelos.

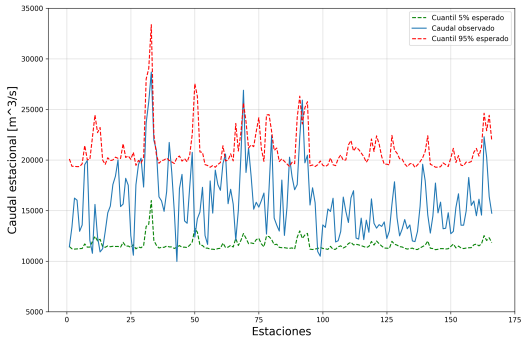


Figura 3.45: Series de cuantiles 5 % y 95 % para la distribución del caudal de Túnel Subfluvial condicionada al índice SOI observado, junto con el caudal observado.

En el caso de Paso de los Libres también resultaron interesantes los resultados obtenidos, puesto que a partir de este modelo es posible notar como eventos extremos superiores al percentil 95 pudieron ser modelados (Figura 3.47). Más aún, fue posible notar que la predicción de los cuantiles inferiores al 95 (Figura 3.48) debe ser analizada con cierto cuidado, destacando en este sentido la importancia de la relación índice-caudal para eventos extremos asociados a la cola derecha de la distribución de caudal.

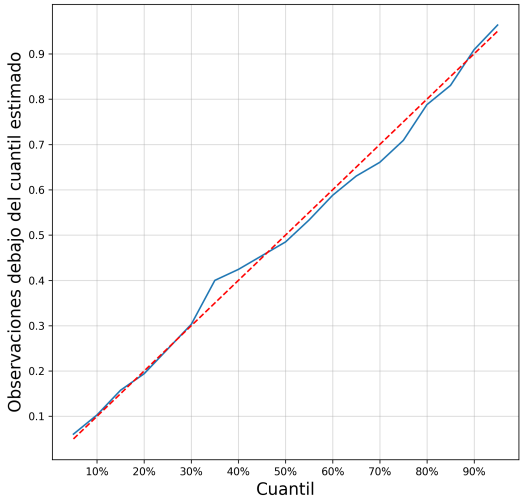


Figura 3.46: Proporción de observaciones que caen por debajo del cuantil estimado. Relación Túnel Subfluvial-SOI.

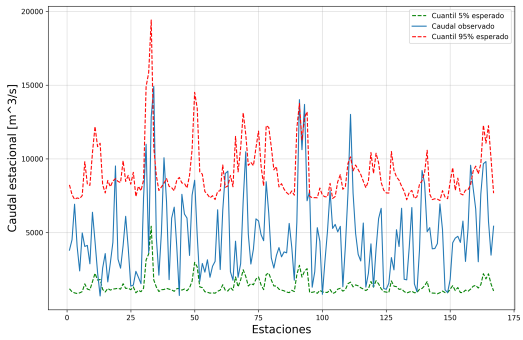


Figura 3.47: Series de cuantiles 5 % y 95 % para la distribución del caudal de Paso de los Libres condicionada al índice SOI observado, junto con el caudal observado.

Por último, el estudio para la estación La Angostura es interesante de acuerdo al modelo realizado para el pronóstico de extremos de caudal pero para la cola izquierda de la distribución, es decir, para predecir valores bajos de caudal. En este caso es posible notar a partir de las Figuras 3.49 y 3.50 la obtención de un modelo que resulta útil para la predicción de valores del caudal por debajo del percentil 10. Para destacar el modelo se realizó un zoom del intervalo de confianza del 5 % (Figura 3.51).

Como se mencionó previamente se empleó otra métrica entre la comparación de los mode-

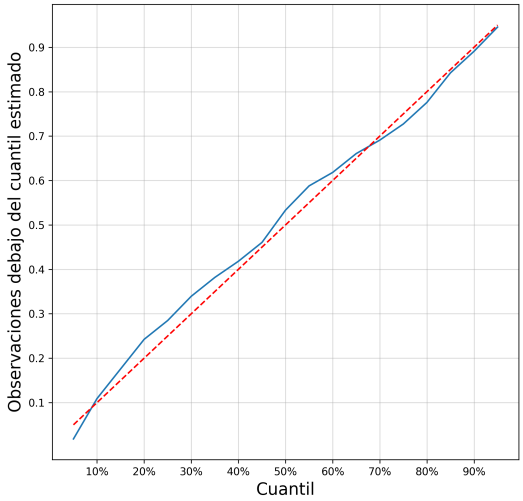


Figura 3.48: Proporción de observaciones que caen por debajo del cuantil estimado. Relación Paso de los Libres-SOI.

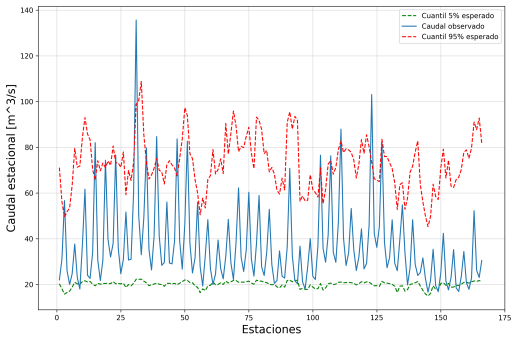


Figura 3.49: Series de cuantiles 5 % y 95 % para la distribución del caudal de La Angostura condicionada al índice SOI observado, junto con el caudal observado.

los propuestos para Túnel Subfluvial teniendo en cuenta tanto el índice NIÑO 3.4, como el índice SOI con el objetivo de comprender cual de los ajustes modela mejor la dependencia en los extremos. De acuerdo con la Ec.(3.3), se estimó el valor esperado de caudal condicional a un valor acumulado del índice, tanto a partir de los modelos teóricos que se obtiene al considerar las familias de cópulas propuestas por las dos relaciones anteriores, como así también a partir de los datos observados.

En el caso en el que la dependencia más fuerte se presenta en la cola izquierda del índice (en este caso, para el índice SOI), se estimó $E(U_2|U_1 < z)$ prestando especial atención al

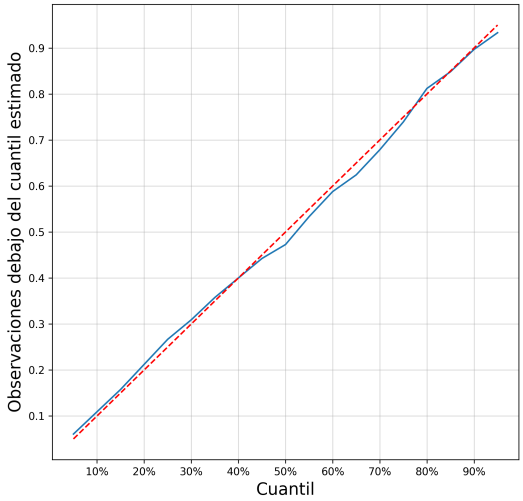


Figura 3.50: Proporción de observaciones que caen por debajo del cuantil estimado. Relación La Angostura-SOI.

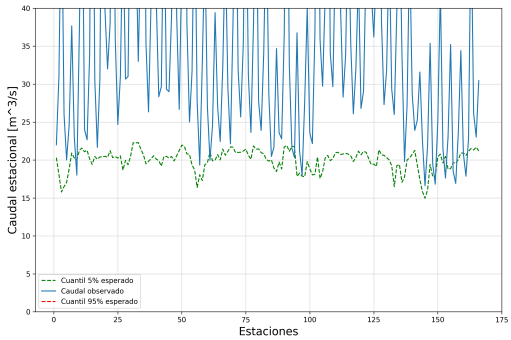
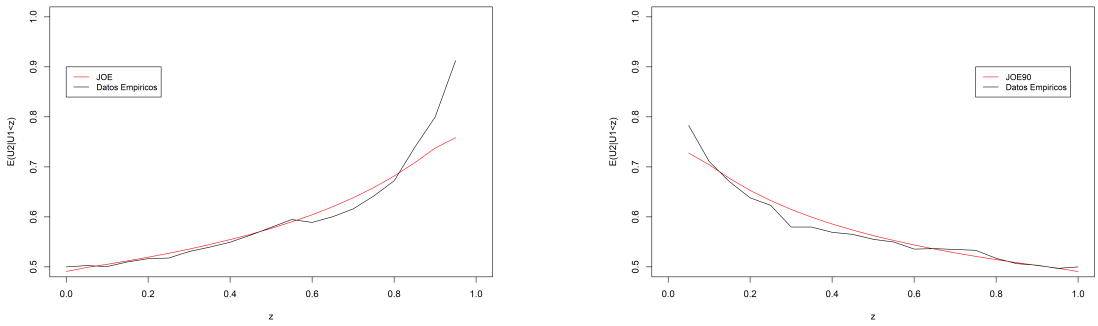


Figura 3.51: Series de cuantiles 5 % y 95 % (zoom) para la distribución del caudal de La Angostura condicionada al índice SOI observado, junto con el caudal observado.

comportamiento de las métricas cuando $z \to 0$, en la cola relevante de la variable predictora. Por otro lado, cuando la dependencia es en la cola derecha se estimó $E(U_2|U_1 > z)$, esperando coincidencia cuando $z \to 1$.

La métrica calculada para el modelo NIÑO 3.4-Túnel Subfluvial se observa en la Figura 3.52a, mientras que para el para el par SOI-Túnel Subfluvial se exhibe en la Figura 3.52b. A partir de las figuras anteriores es posible notar que al analizar los valores extremos del valor medio de la variable a predecir, el caudal, el modelo que involucra el índice SOI tiende a estar más cerca a los datos empíricos a diferencia del modelo que involucra al



(a) (b)

Figura 3.52: Estimación del valor esperado de caudal condicional a un valor acumulado del índice: a) NIÑO 3.4, b) SOI.

NIÑO 3.4 donde la diferencia que hay para el valor medio esperado entre la curva teórica y la que se obtiene a partir de los datos observados es más amplia. De acuerdo con la métrica anterior, y en conjunto con lo analizado a partir de la relación de los cuantiles (Figuras 3.45 y 3.46) se vislumbra que dos métricas distintas estarían no necesariamente seleccionando un único modelo. En este sentido, ambos modelos resultan de utilidad en el presente análisis de predicción de valores extremos de caudal a partir de un cierto índice climático.

3.5.6. Evaluación del percentil 95 como medida de aviso temprano de evacuación

En esta nueva sección del Capítulo 3 se planteó evaluar y cuantificar el posible sistema de alerta de la serie de percentil 95 para la distribución del caudal de Túnel Subfluvial y Paso de los Libres condicionadas a los índices NIÑO 3.4 y SOI observados.

Teniendo en cuenta que para la estación Túnel Subfluvial se consideraron dos índices climáticos con los que se pudo obtener un modelo adecuado se mostrará primero los cálculos realizados condicionando al índice NIÑO 3.4, y luego al índice SOI. De acuerdo a lo descripto en la sección anterior se realizó la curva ROC (Figura 3.53) entre la variable que toma valores ceros y unos que discrimina cuando la altura del río en la estación Túnel Subfluvial supera los 5mts. al menos una vez en el trimestre (variable respuesta) y el caudal estimado condicional del percentil 95 a partir del índice NIÑO 3.4 (variable predicha).

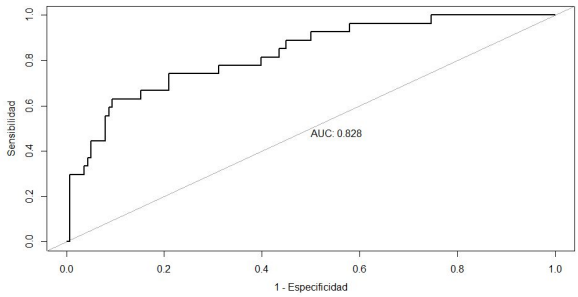


Figura 3.53: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice NIÑO 3.4.

A partir de la estimación anterior de la curva ROC se obtuvieron los umbrales de caudal correspondientes. Para cada valor del caudal estimado de manera condicional de la cópula se obtuvo una nueva variable binaria. Si la observación era mayor o igual que el umbral obtenido de la curva ROC se clasificará al mismo con un uno, caso contrario con cero, obteniendo en total 166 vectores.

Teniendo en cuenta la librería performanceEstimation de R (Torgo, 2016) se estimó el valor del F-score para cada par de vectores y se seleccionó a partir del mismo el umbral que presentó el mayor F-score. El umbral se correspondió con 23147.76 $[m^3/s]$ y presentó un F-score igual a 0.60. Cabe destacar que más allá de la elección del umbral, el área bajo la curva ROC obtenida fue de 0.83, valor que muestra que el percentil 95 % estimado es un predictor bastante razonable para la variable de respuesta considerada, es decir, el nivel de evacuación emitido por prefectura. En la Figura 3.54 se observan los valores de 1 - especificidad (razón de falsos positivos) y sensibilidad para el umbral que tiene asociado la maximización del F-score en la curva ROC. Se puede observar que en este punto parecería haber una disminución de la pendiente en la curva. Para ese valor de umbral se obtuvo la tabla de contingencia asociada, Tabla XVIII.

Teniendo en cuenta la Tabla XVIII, se realizó un test chi-cuadrado para evaluar la asociación entre ambas variables con un 95 % de confianza, de acuerdo al p-valor obtenido de 0.0004998 se rechaza la hipótesis nula con lo que se infiere una asociación entre ambas variables.

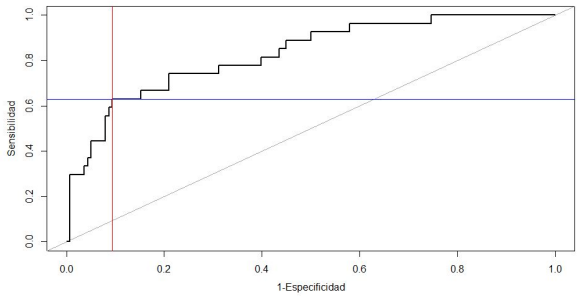


Figura 3.54: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice NIÑO 3.4. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que maximiza el F-score.

Tabla XVIII: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC que maximiza el F-score.*

	Respuesta		
		0	1
Predicho			
	0	125	10
	1	13	17

Por otro lado se estimó de manera visual un umbral que representa un cambio en la pendiente de la curva ROC y se obtuvo otro umbral asociado a un valor de 20708.24 $[m^3/s]$. Además, en la Figura 3.55 se muestra la curva ROC con línea vertical y horizontal asociados a los valores de 1-especificidad y sensibilidad respectivamente para ese umbral fijo. Así como observamos que el corte anterior correspondía a un primer cambio de pendiente en la curva, la elección de este segundo umbral responde a un efecto similar. Este segundo umbral corresponde a un método de alerta más conservador, en el sentido que emite más alertas, mejorando el recall (cercano al 90 %) al costo de generar mayor cantidad de falsos positivos (cercanos al 45 %).

De acuerdo a este nuevo umbral que se estima de manera visual se obtiene la siguiente tabla de contingencia, Tabla XIX. Además, teniendo en cuenta el test chi-cuadrado para evaluar la asociación entre ambas variables con un 95 % de confianza, se obtiene un p-valor igual a 2.188e-07 con lo que se infiere una asociación entre ambas variables.

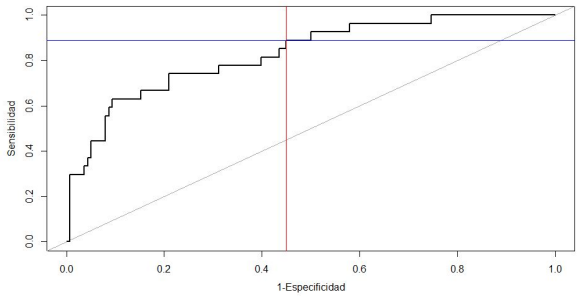


Figura 3.55: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice NIÑO 3.4. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que se encuentra con un cambio en el crecimiento de la curva ROC.

Tabla XIX: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC donde se destacó un cambio en la pendiente de la curva.*

	Respuesta		
		0	1
Predicho			
	0	76	3
	1	62	24

A continuación se presentan los resultados teniendo en cuenta el índice SOI para obtener la esperanza condicional del caudal. Nuevamente se obtiene la curva ROC (Figura 3.56) entre el vector de ceros y unos de la altura del río en la estación Túnel Subfluvial (variable respuesta) y el caudal estimado condicional del percentil 95 a partir del índice SOI (variable predicha).

De la estimación de la curva ROC se obtiene nuevamente los umbrales de caudal correspondientes y los respectivos vectores binarios, y se estimó el valor del F-score para cada par de vectores seleccionando al igual que en el proceso anterior el umbral que presente el mayor F-score. El umbral se correspondió con 22490.01 $[m^3/s]$ y presentó un F-score igual a 0.55. Por otra parte, para este nuevo modelo, el área bajo la curva ROC arrojó un valor igual a 0.76. En comparación con el modelo anterior el área es menor y por lo tanto este modelo presenta un rendimiento menor teniendo en cuenta esta métrica. En la Figura 3.57 se destaca la curva ROC con el umbral que tiene asociado la maximización

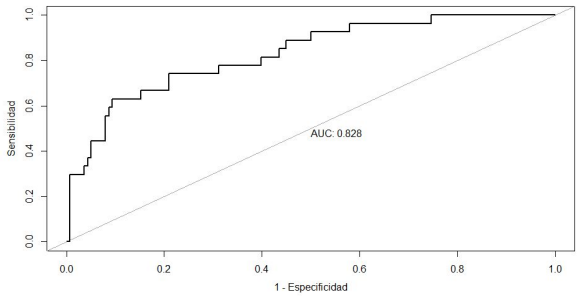


Figura 3.56: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI.

del F-score.

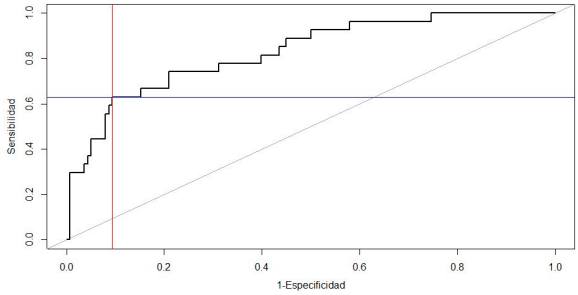


Figura 3.57: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que maximiza el F-score.

Se calculó la tabla de contingencia, Tabla XX, y se realizó nuevamente el test chi-cuadrado para evaluar la asociación entre ambas variables, y de acuerdo al p-valor obtenido (0.001783) se rechazó la hipótesis nula.

Tabla XX: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC que maximiza el F-score.*

	Respuesta		
		0	1
Predicho			
	0	127	12
	1	12	15

En cuanto a la estimación de manera visual del umbral que representa un cambio en la pendiente de la curva ROC se obtiene un valor igual a 20014.22 $[m^3/s]$. En la Figura 3.58 se muestra la curva ROC con línea vertical y horizontal asociados a los valores de 1-especificidad y sensibilidad, respectivamente, para ese umbral fijo.

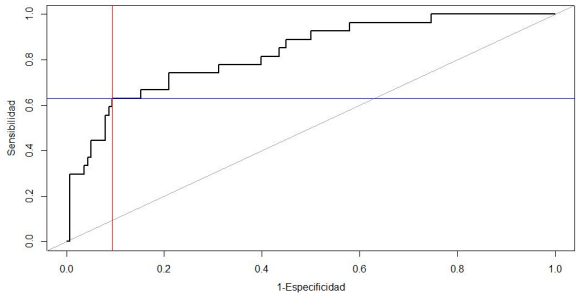


Figura 3.58: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Túnel Subfluvial y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que se encuentra con un cambio en el crecimiento de la curva ROC.

A partir del umbral anterior se obtiene la siguiente tabla de contingencia asociada, Tabla XXI. Mediante, el test chi-cuadrado se obtiene un p-valor igual a 0.001783, y por lo tanto se rechaza la hipótesis nula, y entonces hay una asociación entre ambas variables.

Tabla XXI: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC donde se destacó un cambio en la pendiente de la curva.*

		Respuesta	
		0	1
Predicho			
	0	70	4
	1	69	23

Teniendo en cuenta las métricas realizadas para la evaluación del sistema de alerta del percentil 95 esperado de caudal es destacable que se obtuvieron resultados razonables tanto cuando se tuvo en cuenta el índice NIÑO 3.4 como cuando se tuvo en cuenta el índice SOI. En ambos casos la comparación entre el sistema de aviso de evacuación y el caudal esperado del percentil 95 fue favorable, de acuerdo a los valores obtenidos para el

área debajo de las respectivas curvas ROC. Los resultados más satisfactorios se encontraron para el caso del índice NIÑO 3.4. Asimismo, en ambos estudios se propusieron dos umbrales de caudal esperado a partir de los cuales prevenir sobre una posible evacuación. De este modo no solo se estudia la relación general de la variable de caudal esperado con el sistema de aviso de evacuación sino que también se cuantifica el rendimiento de dos posibles sistemas de predicción de evacuación a partir de la variable construida.

En el caso de la estación Paso de los Libres se tuvo en cuenta el caudal esperado asociado al percentil 95 a partir del índice SOI. Al igual que los procedimientos anteriores, se realizó la curva ROC (Figura 3.59) entre la variable binaria obtenida de la altura del río en la estación Paso de los Libres que tiene en cuenta el nivel de evacuación (variable respuesta) y el caudal estimado condicional del percentil 95 a partir del índice SOI (variable predicha). Luego, mediante la estimación de la curva ROC se obtuvieron los umbrales de caudal correspondiente, y los respectivos vectores binarios. Asimismo, se estimó el valor del F-score para cada par de vectores y se seleccionó el umbral que presentó el mayor F-score. En este caso, el umbral se correspondió con 9141.305 $[m^3/s]$ y presentó un F-score igual a 0.49. El área bajo la curva ROC presentó un valor igual a 0.75. En la Figura 3.60 se destaca la curva ROC con el umbral que tiene asociado la maximización del F-score.

Por otra parte se obtiene la tabla de contingencia asociada al umbral anterior, Tabla XXII. Mediante el test chi-cuadrado se obtiene un p-valor igual a 4.637e-05 por lo tanto se rechaza la hipótesis nula y se infiere así que hay una asociación entre ambas variables. Finalmente, se estimó de manera visual el umbral que representa un posible cambio en la pendiente de la curva ROC encontrando un valor igual a 8332.985 $[m^3/s]$. Asimismo, en la Figura 3.61 se presenta la curva ROC con línea vertical y horizontal asociadas a los valores de 1-especificidad y sensibilidad respectivamente para ese umbral fijo.

Tabla XXII: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC que maximiza el F-score.*

		Respuesta	
		0	1
Predicho			
	0	100	14
	1	31	22

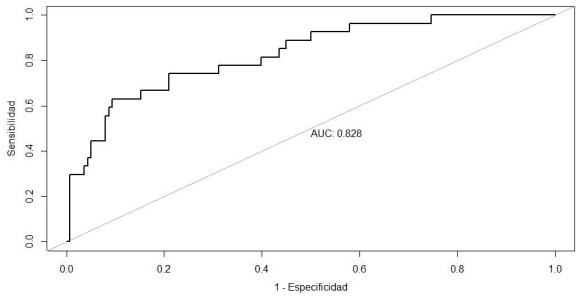


Figura 3.59: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Paso de los Libres y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI.

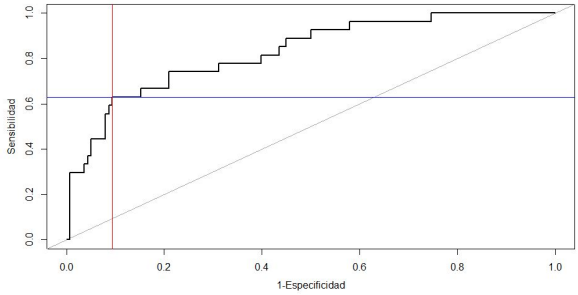


Figura 3.60: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Paso de los Libres y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que maximiza el F-score.

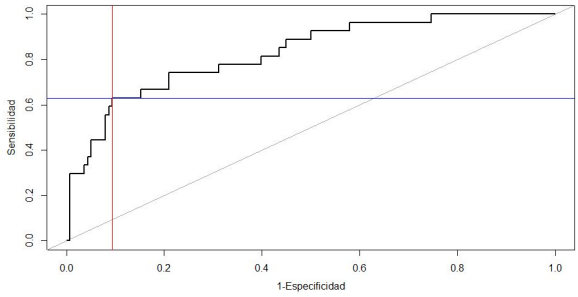


Figura 3.61: Curva ROC entre la variable respuesta binaria trimestral de altura del río en Paso de los Libres y el caudal esperado del percentil 95 teniendo en cuenta el índice SOI. La línea roja representa el valor de 1-especificidad y la línea azul el de sensibilidad para el umbral que se encuentra con un cambio en el crecimiento de la curva ROC.

De acuerdo a este nuevo umbral se obtiene la siguiente tabla de contingencia, Tabla XXIII, y teniendo en cuenta el test chi-cuadrado el p-valor obtenido fue igual a 0.0003283, y por lo tanto se rechazó la hipótesis nula y entonces hay una asociación entre ambas variables.

Tabla XXIII: *Tabla de contingencia entre la variable respuesta (altura binaria) y el caudal binario a partir del umbral obtenido de la curva ROC donde se destacó un cambio en la pendiente de la curva.*

	Respuesta		
		0	1
Predicho			
	0	72	7
	1	59	29

Al igual que en la estación anterior, para Paso de los Libres el área bajo la curva ROC indicó una buena relación entre la variable de caudal esperado construida y los niveles de evacuación por altura de río. Asimismo, también se examinó el rendimiento de dos posibles umbrales a partir de los cuales prevenir sobre posibles evacuaciones.

3.5.7. Comparación entre el modelo teórico y los datos observados

Se han realizado simulaciones utilizando el modelo de cópula para observar y comparar la función teórica con los datos medidos. Los resultados que se muestran se corresponden con los distintos modelos que pudieron ser validados de manera aleatoria y manteniendo la secuencia temporal o solo por la primera, Figura 3.62. En este sentido se muestran los resultados para el par Túnel Subfluvial-NIÑO 3.4, Paso de los Libres-NIÑO 3.4, mismas estaciones de medición pero con el índice SOI, y también el par La Angostura-SOI. Además, se exhibe la simulación para el par La Angostura-PDO. Nuevamente, es posible observar la relación positiva entre las variables en el comportamiento teórico (puntos negros), como también en los datos reales (datos rojos) en el caso de las mediciones que involucran el índice NIÑO 3.4. Esta relación nace de la estrecha correlación positiva entre el NIÑO 3.4 y la precipitación en la región de la LPB (como ya ha quedado expuesto en base a la bibliografía considerada). Adicionalmente, es importante mencionar que la distribución de los datos para valores extremos muestra un adecuado modelado teórico

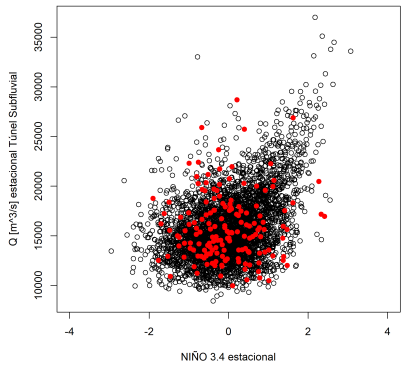
en la cola superior. En el caso del índice SOI y la estación del río Paraná se destaca para valores bajos del índice valores altos de caudal, caso contrario es la relación que se pretendió modelar en La Angostura que estuvo relacionada con los valores bajos de caudal y valores altos del índice. Por otra parte en el caso de la relación de esta última estación y la PDO se observó una simulación donde al no ser tan fuerte la relación no es tan apreciable la misma como los casos anteriores. Es necesario recordar en este caso que, la relación que se quiere estudiar está asociada con poder modelar eventos de poco caudal asociado a valores negativos del índice, puesto que dada la familia de cópula obtenida es lo que tiene sentido estudiar.

3.5.8. Probabilidad estimada para eventos extremos

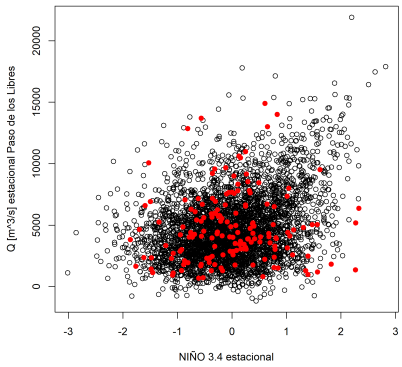
Para determinar la probabilidad conjunta de ocurrencia de eventos extremos, se ha estimado la probabilidad acumulada para la relación bivariada. Este análisis resulta útil para estimar la probabilidad de caudales extremos (por encima de los percentiles 90 y 95, y por debajo de los percentiles 10 y 5) y relacionar los mismos con ciertas oscilaciones climáticas. En este sentido, dado que se está estudiando la relación bivariada es también necesario fijar un umbral para el índice climático.

Debido a la ocurrencia de ciertos eventos extremos en el caudal estacional, y teniendo en cuenta los desfasajes establecidos previamente para el índice climático, se observa que los valores estacionales del NIÑO 3.4 asociado a esos eventos extremos se encontraban por encima del percentil 95. La Figura 3.63 muestra la función de densidad de probabilidad acumulada del modelo teórico para Túnel Subfluvial y NIÑO 3.4. Las líneas rojas resaltan el cuadrante que se encuentra asociado con esos eventos extremos, es decir, fijando los umbrales del percentil 95 para ambas variables. De acuerdo con este modelo teórico la probabilidad conjunta para eventos extremos fue del 2.2 %. La relación entre Paso de los Libres y el NIÑO 3.4 mostró resultados similares y por esta razón la figura no es mostrada.

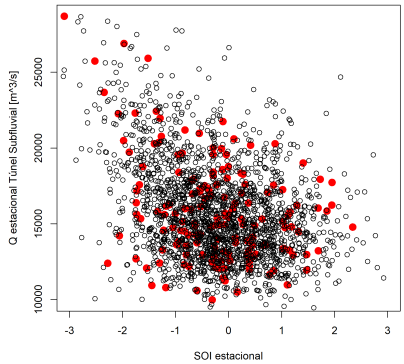
Por otra parte, los resultados obtenidos con el índice SOI para las estaciones anteriores en lo que refiere a los valores de probabilidad estimados, los mismos resultaron ser similares, es decir, alrededor del 2 % en ambos casos de estudio, con la diferencia en que en este caso particular es distinto el cuadrante del que se desea estimar dicha probabilidad Figuras 3.64a y 3.64b. En el caso de la estación La Angostura el valor de probabilidad



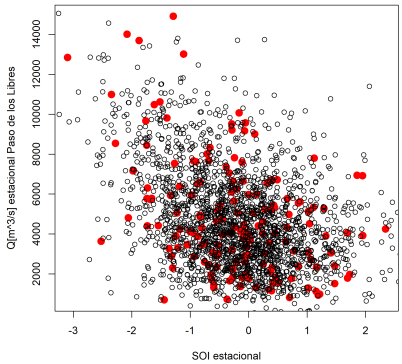
(a)



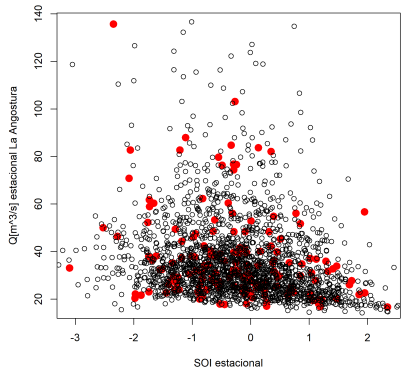
(b)



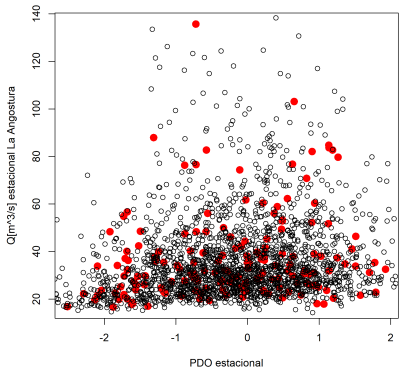
(c)



(d)



(e)



(f)

Figura 3.62: Diagramas de dispersión para la relación observada (puntos rojos), y la simulación que se obtendrían de una función cópula de cada caso particular, y sus correspondientes funciones de distribución teóricas marginales.

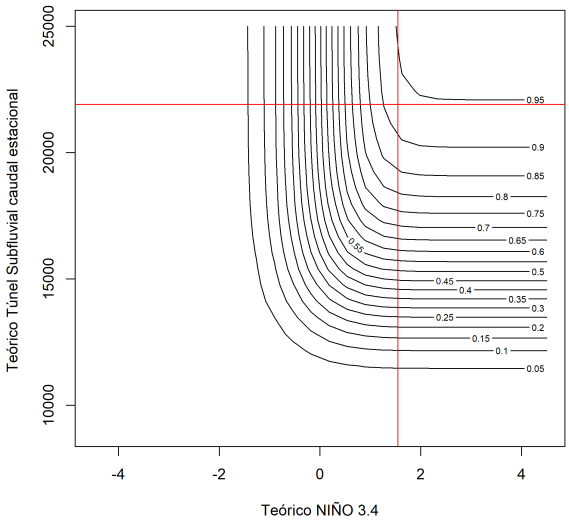
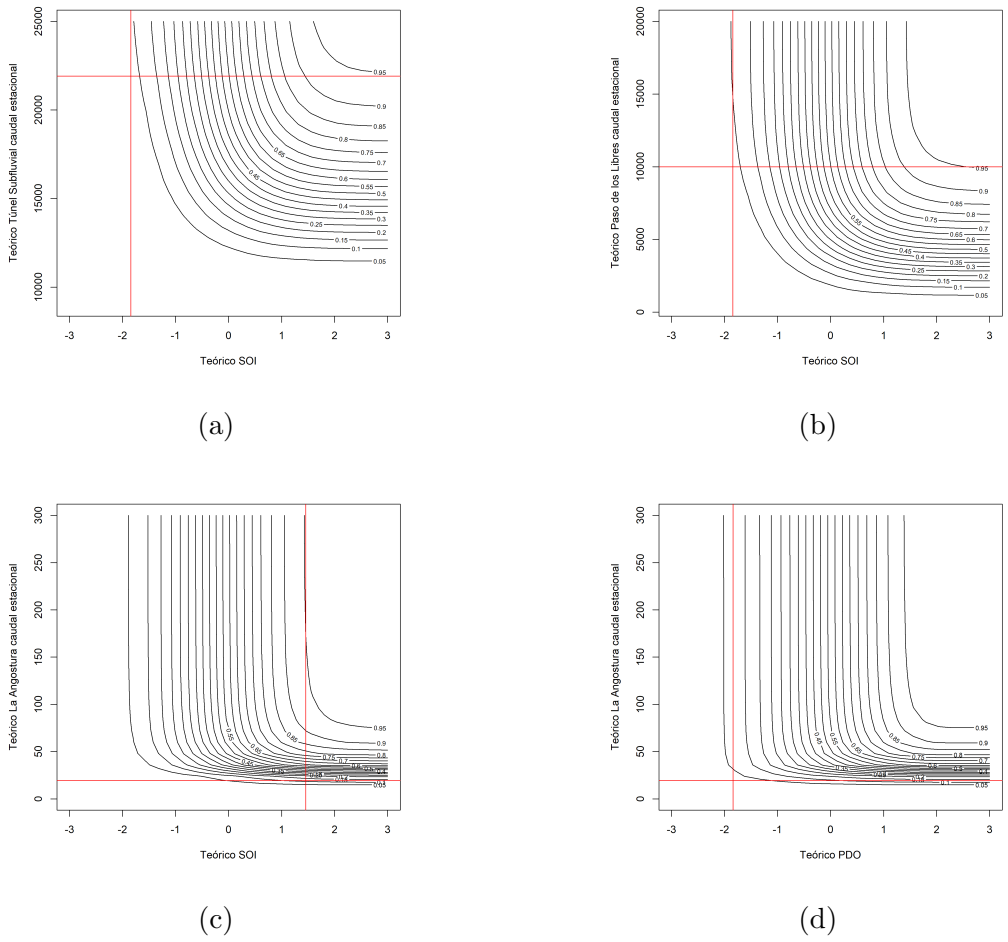


Figura 3.63: Función de densidad de probabilidad acumulada del modelo teórico, las líneas rojas representan los percentiles límites que se consideraron en el cálculo de probabilidad realizada.

para eventos extremos (percentil 5 y 10 del caudal; 95 y 90 para el índice) fue de del 1 % y del 3 %, respectivamente (Figura 3.64c). Por otra parte también se estimó la probabilidad de ocurrencia conjunta entre La Angostura y PDO en eventos extremos, donde el caudal presente valores por debajo de los percentiles 5 y 10 en conjunto con los valores del índice PDO (Figura 3.64d). Para estas condiciones los resultados obtenidos fueron de un 3 % para el primer caso, y de un 6 % para el segundo.

3.5.9. Periodos de retorno bivariados en eventos extremos

Una métrica que puede ser estimada a partir de la distribución conjunta del par de series es el periodo de retorno (Salvadori y De Michele, 2004). La Figura 3.65 muestra la distribución de los datos para Túnel Subfluvial-NIÑO 3.4 (puntos negros), y el periodo de retorno en isolíneas que fueron obtenidos de la función de densidad de probabilidad que se ha considerado en la subsección anterior. Si se analizan los eventos extremos del caudal estacional (por encima del percentil 95, línea roja) asociada con valores del NIÑO 3.4 por encima del percentil 95 (1.54) se observa un periodo de retorno de en promedio cada 15 años de acuerdo a los datos empíricos, y de 11.5 años de acuerdo a los datos simulados.



(a) (b)

(c) (d)

Figura 3.64: Función de densidad de probabilidad acumulada del modelo teórico, las líneas rojas representan los percentiles límites que se consideraron en el cálculo de probabilidad realizada.

Los eventos que se encuentran por encima de esos umbrales son el otoño/invierno 83, invierno 92 y otoño 98 para el caudal estacional, así como también los veranos 98 y 16. También es importante notar que el periodo de retorno para cada variable para el percentil 95 sería de 5 años, en este sentido si ambas variables fueran independientes el resultado sería muchísimo mayor, logrando un valor de 100 años de periodo de retorno para los umbrales conjuntos previos. En otras palabras, el último resultado revela una fuerte dependencia entre las variables. Más aún, el análisis muestra que eventos extremos para ambas variables parecerían estar fuertemente correlacionadas, por lo que es importante y necesario modelar la relación índice-caudal. Además, es importante notar que si bien había ciertos eventos cuyos valores se encontraban por encima del percentil 95 (21916.61 m^3/s), el valor del índice se encontraba por debajo; todavía esos eventos cuentan como eventos El Niño (verano 83, otoño 95). En efecto, eventos en la cola alta se corresponden con la fase cálida del ENSO, y condiciones del caudal no normales (se encuentran sustancialmente por encima del valor medio). Finalmente, en la Figura 3.65 es posible observar el evento estacional más extremo, invierno 83. La combinación de eventos extremos del caudal y extremos estacionales positivos del índice NIÑO 3.4 para este caso tiene un valor cercano a los 200 años de periodo de retorno. Este valor de periodo de retorno no es tan sorpresivo, puesto que, este evento ENSO es considerado excepcional (Camilloni y Barros, 2000).

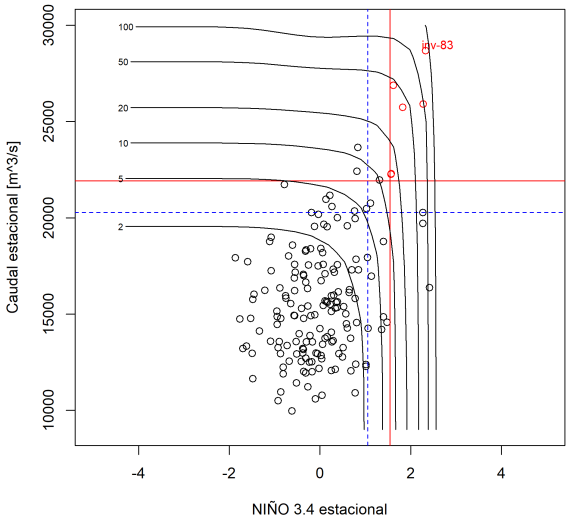


Figura 3.65: Distribución de los datos para Túnel Subfluvial-NIÑO 3.4 (puntos negros), y el periodo de retorno (líneas negras).

Con respecto a la relación Paso de los Libres NIÑO 3.4 (Figura 3.66), se ha observado eventos extremos con valores del caudal por encima del percentil 95, y por encima del percentil 95 del índice (1.39). Al computar el periodo de retorno, se observó que los valores para aquella magnitud ocurren, en promedio, cada 14.5 años. Entre las observaciones, se cuenta con cinco eventos en el periodo de estudio (otoño e invierno 83, invierno 92, verano y otoño 98). Además, había eventos por encima del percentil 95 para el caudal (10050.38 m^3/s), pero con valores del índice climático por debajo de aquel percentil (primavera 82, 97 y 2002). Adicionalmente, invierno del año 1984, se encontraba cercano al umbral del percentil 95 para el caudal estacional. Este fue un evento aislado, puesto que no se encuentra asociado con una fase cálida del ENSO. Bibliografía existente menciona inundaciones substanciales que ocurrieron en la región de la LPB para esa estación del año (Righi y Souza Robaina, 2010), donde los autores hacen énfasis en que la precipitación muestra valores extremos, posiblemente debido a causas físicas.

Ambos análisis no solo muestran los umbrales del percentil 95, sino también el umbral del percentil 90, ambos para el caudal estacional y el índice climático estacional (Figura 3.65 y 3.66). El análisis de periodo de retorno de Túnel Subfluvial-NIÑO 3.4 para este tipo de eventos fue en promedio de 5 años, mientras que para Paso de los Libres fue un levemente menor.

En el estudio respecto al índice SOI se encontró que para Túnel Subfluvial, los resultados de periodo de retorno tanto en la simulación como a través de los datos, los mismos fueron muy similares a los hallados por el índice NIÑO 3.4 (Figura 3.67). Para Paso de los Libres el valor obtenido de periodo de retorno a partir de los datos simulados fue menor que para el índice anterior pues fue de alrededor de cada 12 años en promedio. Más aún, se observa lo mismo en los resultados empíricos al compararlos (Figura 3.66 y Figura 3.68), sin embargo resultaría más prudente tener en cuenta con mayor prioridad este modelo puesto que este si pudo ser validado manteniendo la secuencia temporal.

El cálculo de probabilidad de eventos extremos de ocurrencia conjunta entre la estación La Angostura y el índice SOI, y más aún el periodo de retorno (Figura 3.69) relacionados con eventos NIÑA y bajo caudal estacional destacó una probabilidad de ocurrencia del 3 %, teniendo asociado un periodo de retorno de 7.8 años. Es decir, en promedio ocurren eventos superiores al percentil 90 del índice y por debajo del percentil 10 del caudal estacional cada 7.8 años. Dentro de los eventos observados que cumplen las condiciones anteriores

se enumeran los siguientes, con la ocurrencia temporal del índice (estación/año), y dos estaciones siguientes para el caudal, verano76 SOI - invierno76 caudal; verano99 SOI - invierno99 caudal; verano09 SOI - invierno09 caudal; verano11 SOI - invierno11 caudal; primavera11 SOI - otoño12 caudal; verano12 SOI - invierno12 caudal. Es importante destacar que las observaciones que cumplieron las condiciones propuestas están todas relacionadas con eventos NIÑA de intensidad leve a fuerte. Por otra parte, si las variables estudiadas (índice-caudal) fueran independientes los eventos extremos considerados tendrían asociado un periodo de retorno estimado de 25 años en promedio, pero ya se ha destacado a través de distintos procedimientos metodológicos la relación existente entre ambas.

En el estudio entre La Angostura y el índice PDO (Figura 3.70), el resultado del periodo de retorno conjunto considerando como umbrales los percentiles 5 para cada variable presentó un valor estimado de 7.8 años, mientras que al tener como umbrales los percentiles 10 el valor estimado fue de 3.80 años en promedio. De todas maneras en este caso de estudio al no haber sido validado de las dos maneras propuestas y solo en forma aleatoria los resultados deberían ser considerados con mayor cuidado que los modelos que si lograron ser validados de ambas formas. Asimismo, como herramienta de diagnóstico resulta interesante, puesto que ya otros autores como Araneo y Villalba (2015) destacaron la relación entre la fase negativa de la Oscilación Decadal del Pacífico y el caudal presentando valores por debajo de su respectivo valor medio.

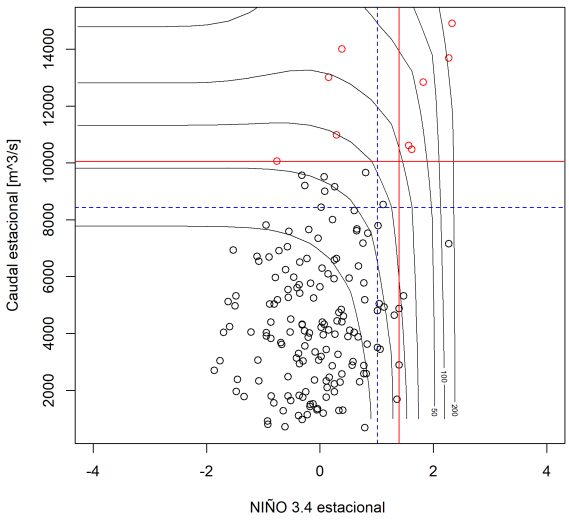


Figura 3.66: Distribución de los datos para Paso de los Libres-NIÑO 3.4 (puntos negros), y el periodo de retorno (líneas negras).

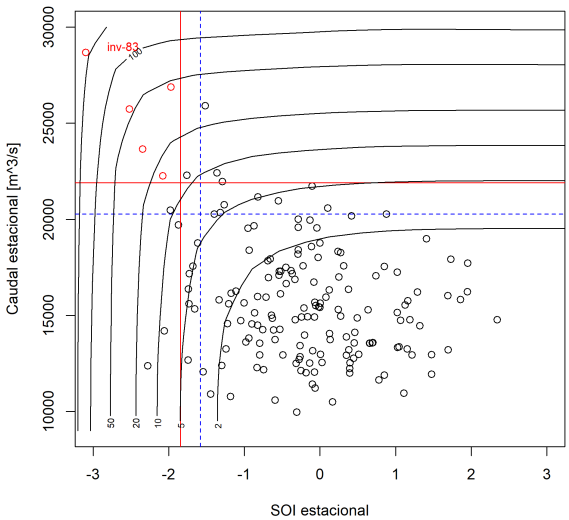


Figura 3.67: Distribución de los datos para Túnel Subfluvial-SOI (puntos negros), y el periodo de retorno (líneas negras).

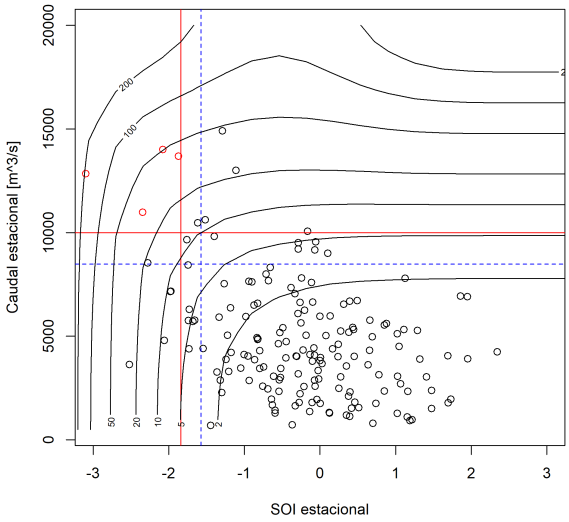


Figura 3.68: Distribución de los datos para Paso de los Libres-SOI (puntos negros), y el periodo de retorno (líneas negras).

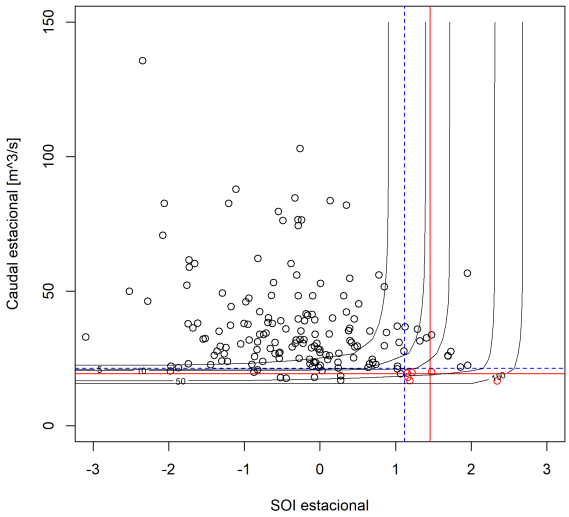


Figura 3.69: Distribución de los datos para La Angostura-SOI (puntos negros), y el periodo de retorno (líneas negras).

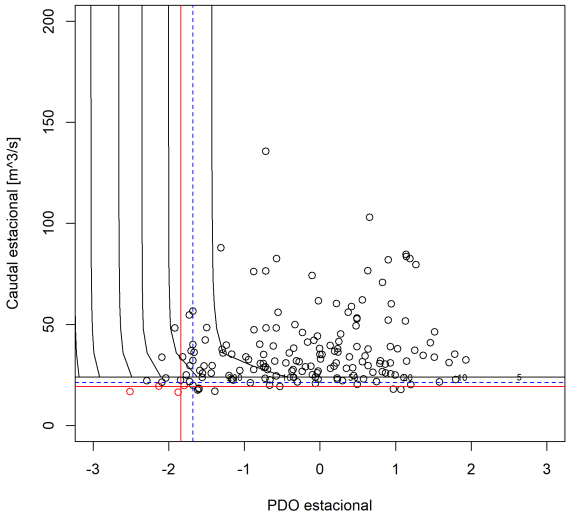


Figura 3.70: Distribución de los datos para La Angostura-PDO (puntos negros), y el periodo de retorno (líneas negras).

3.6. Conclusiones

Este análisis propone un método para estimar la función de densidad de probabilidad conjunta combinando un desfasaje para el índice climático y el caudal estacional. El principal objetivo del capítulo fue el de encontrar una herramienta útil para anticipar eventos extremos en el caudal, puesto que se estudió la interacción entre el índice observado y el caudal que ocurre entre 1, 2 o 3 trimestres después dependiendo del modelo establecido. Estimar la distribución conjunta de un par de variables respuesta/explicatoria puede ser útil en varios aspectos. En este capítulo, se ha derivado métricas que resultan particularmente útiles para cuantificar algunas relaciones conocidas, como así también para proponer simples ideas de pronóstico. Sin embargo, la ventana de posibles aplicaciones puede ser ampliada, ya que la distribución conjunta de un par de variables contiene toda la información probabilística en ella. Por ejemplo, la información sobre la distribución conjunta puede ser utilizada como un input para modelos más complejos. En este caso, esta herramienta habilitó a cuantificar probabilidades (y sus periodos de retorno asociados) para ciertos eventos extremos conjuntos que ocurrieron en los últimos años. Una métrica como la utilizada puede ser útil para monitorear por ejemplo, la ocurrencia de eventos extremos en la variable caudal y el índice climático relacionado con intensos eventos ENSO (Takahashi y Dewitte, 2016). Más aún, se ha realizado dos posibles enfoques para el pronóstico estacional de caudal. El primero, que sigue la línea de trabajo de Khendun et al. (2014) está relacionado con estimar el valor esperado de caudal dado el valor observado del índice, y utilizarlo como un pronóstico. El segundo es computando intervalos de confianza para el caudal dado los valores del índice climático. En el primer caso, se ha encontrado que hay señal en la predicción al estimar una correlación significativa entre los valores observados y los valores pronosticados. En el segundo caso, el pronóstico de los cuantiles es especialmente útil ya que provee una manera simple de obtener un valor que el caudal no superará con cierto intervalo de confianza (en el caso que se este mirando la cola derecha de la distribución de datos del caudal). En este sentido, el cuantil 95 % de la distribución condicional teórica parece ser una métrica de monitoreo razonable, ya que de acuerdo a los resultados obtenidos es un umbral que el caudal respeta con buena confianza, siendo al mismo tiempo suficientemente ajustado como para no levantar alarmas en todos los periodos. Esto resulta evidente en el estudio entre el índice NIÑO 3.4 y el caudal en la estación Túnel Subfluvial, como así también entre el índice SOI y las tres

estaciones consideradas (teniendo en cuenta que en la estación La Angostura el umbral de estudio era el cuantil 5 %). Más aún, se adicionó una métrica para evaluar la esperanza acumulada condicional a fin de comparar modelos para Túnel Subfluvial, mostrando que resulta válido tener más de un modelo en consideración.

Adicionalmente, los modelos obtenidos fueron evaluados por diferentes enfoques mostrando que pueden ser considerados para monitorear la relación entre el caudal, el ENSO y las estaciones de medición en la Cuenca del Plata y la del Atuel. Dichos modelos fueron validados temporalmente en el caso de estudio Túnel Subfluvial y el índice NIÑO 3.4 y SOI, al igual que el índice SOI con las estaciones Paso de los Libres y La Angostura, brindando cierta certeza que tanto las métricas monitoreadas como los pronósticos que se han realizado pueden ser significativos para su aplicación. Más aún, también se ha revisado la consistencia de los umbrales computados para los valores de caudal.

Por otra parte, con el uso de distintas métricas se amplió la evaluación del rendimiento de los modelos considerados tanto en la estación Túnel Subfluvial como en la estación Paso de los Libres. Particularmente, se tuvo en cuenta el área debajo de la curva ROC para evaluar la relación entre la variable dicotómica obtenida a partir de la altura del caudal y la variable numérica del percentil 95 de caudal esperado. De acuerdo a los resultados obtenidos se encontró que la variable obtenida a partir del modelo que utiliza el índice NIÑO 3.4 para obtener el caudal esperado del percentil 95 en la estación Túnel Subfluvial presenta los resultados más favorables a la hora de anticipar evacuaciones. Asimismo, se examinaron dos posibles umbrales de aviso en todos los casos de estudio, es decir valores de caudal esperados a partir de los cuales anticipar que se va a emitir un aviso de evacuación. Para ambos umbrales se realizaron las respectivas tablas de contingencia y mediante el test chi-cuadrado se evaluó la hipótesis nula de la falta de asociación entre ambas variables. En todo los casos la hipótesis fue rechazada.

En general, este capítulo de la tesis otorga un enfoque de estudio de la relación entre índices climáticos y el caudal más profunda en una manera que no se encontró en la literatura entre las variables consideradas y las regiones de estudio. Además, se propone una forma simple pero nueva de emplear los resultados obtenidos con esta metodología en el caso de variables meteorológicas.

Para concluir, es importante mencionar que el procedimiento empleado mostró que dado un cierto umbral para el índice climático fue posible cuantificar la probabilidad de eventos

extremos de caudal con una cierta anticipación. En este sentido, fue posible obtener una simple y válida manera probabilística climática de monitorear la ocurrencia de eventos extremos en tres ríos de suma importancia socio-económica tanto en el sur de la región de la Cuenca del Plata, como en la región de Cuyo más precisamente en el río Atuel de variables bien estudiadas y conocidas. En este sentido, por ejemplo, en el caso de Túnel Subfluvial, la metodología prevé la ocurrencia de eventos superiores al percentil 95 en ambas variables en promedio cada 11.5 años. Esto indica que uno espera aproximadamente la mitad de los eventos de caudal se corresponda con eventos extremos del índice. Una identificación temprana de este tipo de eventos podría ayudar a mitigar su impacto negativo.

3.7. Referencias

AghaKouchak, A., Cheng, Ll, Mazdiyasni, O., Farahmand, A., 2014. Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. Geophys. Res. Lett. Vol. 41. 8847-8852.

Antico, A., Torres, M.E., Diaz, H.F., 2015. Contributions of different time scales to extreme Paraná floods. Climate Dynamics. Vol. 46. 3785 - 3792.

Araneo, D., Villalba R., 2015. Variability in the annual cycle of the rio Atuel streamflows and its relationship with tropospheric circulation. International Journal of Climatology. Vol. 35. 2948-2967.

Araneo, D. Rivera, J., Villalba, R., 2015. Variabilidad intraestacional de las condiciones níveas en los Andes Centrales relacionadas con los cambios en el régimen hidrológico del río Atuel Acta Geológica Lilloana. Vol. 27 p. 77 - 86.

Barnston, A., Kumar, A., Goddard, L., Hoerling, M., 2005. Improving seasonal prediction practices thorogh attribution of climate variability. BAMS59-72.

Berri, G., Ghietto, M., García, NO., 2002. The influence of ENSO in the flows of the

upper Paraná river of South America over the past 100 years. American Meteorological Society. Vol. 3. 57-65.

Camilloni, I.A., Barros, V.R., 2003. Extreme discharge events in the Paraná River and their climate forcing. Journal of Hydrology. Vol. 278, 94-106.

Camilloni, I.A., Barros, V. R., 2000. The Paraná river response to El Niño 1982-83 and 1997-98 events. J. Hydrometeorology, 1, 412-430.

Cavalcanti, I.F.A., Carril, A.F., Penalba, O.C., Grimm, A.M., Menéndez, C.G., Sanchez, E., Cherchi, A., Sorensson, A., Robledo, F., Rivera, J., Pántano, V., Bettolli, L.M., Zaninelli, P., Zamboni I., Tedeschi, R.G., Dominguez, M., Ruscica, R., Flach, R., 2015. Precipitation extremes over La Plata Basin - Review and new results from observations and climate simulations. Journal of Hydrology. Vol. 523. 211-230.

Chan, S., Behera, S.K., Yamagata, T. 2008. Indian Ocean Dipole influence on South American rainfall. Geophysical Research Letter. 35, L14S12.

Compagnucci, R.H., Blanco, S.A., Figliola, M.A., Jacovkis, P.M., 2000. Variability in subtropical Andean Argentinean Atuel river: a wavelet approach. Environmetrics 11, 251–269.

Compagnucci, R.H., Vargas, W.M., 1998. Inter-annual variability of the Cuyo rivers streamflow in the argentinean andean mountains and ENSO events. International Journal of Climatology. Vol. 18, 1593-1609.

Cong, R.G, Brandy, M., 2012. The interdependence between rainfall and temperature: copula analyses. The Scientific World Journal. Vol. 2012. 1-11.

Efron, B., Tibshirani, R.J., 1993. An introduction to Boostrap. London: Springer-Science+Business Media, B.V.

Fang, W., Huang, S., Huang, G., Huang, Q., Wang, H., Wang, L., Zhang, Y., Li, P., Ma, L., 2018. Copulas-based risk analysis for inter-seasonal combinations of wet and dry conditions under a changing climate. International Journal of Climatology, DOI: 10.1002/joc.5929.

Fox, J., Weisberg, S., 2018. An R Companion to Applied Regression, Third Edition, Sage.

Fuller, W. A., 1996. Introduction to statistical time series, second ed., New York: John Wiley and Sons.

Ganguli, P., Reddy, M J., 2013. Analysis of ENSO-based climate variability in modulating drought risks over western Rajasthan in India. Journal of Earth System Science. Vol. 122. Issue 1. 253-269.

Garbarini, E., 2016. Algunos indicadores para la predicción estadística de la precipitación estacional en Argentina. Tesis de Licenciatura. Universidad de Buenos Aires.

García, NO., Vargas, WM., 1998. The temporal climatic variability in the 'Río de la Plata' basin displayed by the rivers discharges. Climatic Change. Vol. 38, 359-379.

Genest, C., Favre, A.C., 2007. Everything you always wanted to know about copula modeling but were afraid to ask. Journal of Hydrologic Engineering. Vol. 12. 347-368.

Gillett, N.P., Kell, T.D., Jones, P. D., 2006. Regional climate impacts of the Southern Annular Mode. Geophysical Research Letters. Vol. 33, L23704.

Grimm, A.M., Tedeschi, R.G., 2009. Enso and extreme rainfall events in South America. Journal of Climate, 22, 1589-1609.

Grinsted, A., Moore, J., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics. Vol. 11, 561–566.

Huaylla, C., 2015. Bondad de ajuste: Una revisión de algunos métodos existentes. Tesis de Licenciatura. Universidad de Buenos Aires.

Jogesh Babu, G., Rao, C.R., 2004. Goodness-of-fit tests when parameters are estimated. The Indian Journal of Statistics, 66 (1), 63–74.

Khendun, C.P, Mishra, A.K, Singh, V.P, Giardino, J.R., 2014. A copula-based precipitation forecasting model: Investigating the interdecadal modulation of ENSO's impacts on monthly precipitation. Water Resources Research, Vol. 50, 580–600.

Kiem, A., Franks, S., Kuczera, G., 2003. Multi-decadal variability of flood risk. Geophysical Research Letters, Vol. 30, no. 2, 1035.

Labat, D., 2008. Wavelet analysis of the annual discharge records of the world's largest rivers. Advances in water resources 2008 v.31 no.1, pp. 109-117.

Mantua, N. J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bull. Amer. Meteor. Soc. 78:1069-1079.

Maraun, D., Kurths, J., 2004. Cross wavelet analysis: significance testing and pitfalls. Nonlinear Processes in Geophysics. Vol.11, 505-514.

Marshall, G. J., 2003. Trends in the Southern Annular Mode from observations and reanalyses. J. Clim., 16, 4134-4143.

Meis, M., Llano, M.P., 2018. Hydrostatistical study of the Paraná and Uruguay Rivers. International Journal of River Basin Management. Vol. 17:1, 1-12.

Penalba O.C., Rivera J., 2016. Precipitation response to El Niño/La Niña events in Southern South America ? emphasis in regional drought occurrences. Advances in Geosciences.

vol. 42 p. 1 - 14.

Phillips, P.C.B, Perron, P., 1988, "Testing for a unit root in time series regression", Biometrika, 75, 335–346.

Righi, E., de Souza Robaina, L.E., 2010. Enchentes do rio Uruguai no río Grande Do Sul entre 1980 e 2005: uma análise geográfica. Sociedade & Natureza. Vol. 22. 35-54.

Rivera, J., Penalba, O., Villalba, R., Araneo, D., 2017. Spatio-temporal patterns of the 2010-2015 extreme hydrological drought across the Central Andes, Argentina. Water. Vol 9, 1-18.

Robin, X., Turck, N., Hainard, A. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. Vol. 12, 77.

Saji, N.H., Goswami, B.N., Vinayachandran, P.N., Yamagata, T., 1999. A dipole mode in the tropical Indian Ocean. Nature 401: 360-363.

Salvadori, G., De Michele, C., 2004. Frequency analysis via copulas: theoretical aspects and applications to hydrological events. Water resources research. Vol. 40. 1-17.

Schepsmeier, U., Stoeber, J. Brechmann, E.C., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., Killiches, M., Joe, H., Vatter, T., 2018. Statistical Inference of Vine Copulas. cran.r-project.

Silvestri, G. y Vera, C., 2003. Antarctic Oscillation signal on precipitation anomalies over southeastern South America. Geophys Res Letters 30, 21: 21-15.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris. 8: 229–231.

Takahashi, K., Dewitte, B., 2016. Strong and moderate nonlinear El Niño regimes. Clim.

Dyn. Vol. 46. 1627-1645.

Torgo, L., 2016. An Infra-Structure for Performance Estimation of Predictive Models. cran.r- project.

Venter,G., 2001. Tails of copulas. ASTIN Conference.

Ward, P., Jongman, B., Kummu, M., Dettinger, M., Weiland, F., Winsemius, F., 2014. Strong influence of El Niño Southern Oscillation on flood risk around the world. PNAS. Vol. 111, no. 44, 15659–15664.

Capítulo 4

Modelado y Predicción de las series de Caudal en la Baja Cuenca del Plata

4.1. Resumen

El objetivo principal de este capítulo fue continuar con la generación de resultados que permitiesen a los tomadores de decisión contar con una mayor cantidad de herramientas para evaluar una reducción de impactos negativos asociados a crecidas y/o bajantes en el volumen del caudal. Puesto que en el capítulo anterior, el factor climático estuvo estrechamente ligado con el comportamiento del caudal a nivel estacional, en este se propuso estudiar una escala mensual y evaluar la capacidad de modelar y predecir el caudal sin tener y teniendo en cuenta factores externos. Se empleó un modelado estadístico basado en el modelo estacional autoregresivo integrado de media móvil (SARIMA), y además un modelo que contempló la incorporación de una variable exógena (SARIMAX), con el principal objetivo de representar y pronosticar las series temporales del caudal mensual en distintas estaciones de la baja Cuenca del Plata.

4.2. Introducción

En las últimas décadas, la aplicación de modelos estadísticos en variables comprendidas en el área climatológica e hidrológica ha adquirido mayor demanda, no solo en el mode-

lado de las mismas sino también para su pronóstico. Los modelos estocásticos como los autoregresivos de media móvil (ARMA), los modelos autoregresivos integrados de media móvil (ARIMA) de Box y Jenkins (1976) o los modelos estacionales autoregresivos integrados de media móvil (SARIMA), son algunos de los utilizados debido a su bajo costo computacional como así también por su rápida implementación. Diversos autores han considerado la metodología estadística desarrollada por Box y Jenkins en el estudio de distintas variables meteorológicas como temperatura, viento, precipitación y caudal, entre otras. Torres et al. (2005) realizaron un pronóstico del promedio horario de la velocidad del viento en Navarra, España, y concluyeron que estos modelos resultaban útiles en un pronóstico a corto plazo. Por otra parte, Cadenas y Rivera (2007) obtuvieron resultados adecuados en el pronóstico de la velocidad del viento en el sur de la costa de Oaxaca, México, para permitir luego una optimización de las plantas de generación de energía eléctrica.

Con respecto a la precipitación se aplicaron los modelos ARMA para obtener pronósticos a corto plazo en la cuenca del río Sieve en Italia (Toth et al., 2000) y en estaciones meteorológicas localizadas en el estado de Colorado (EE.UU) y el río Arno (Italia), (Burlando et al., 1992). En la India se desarrolló un modelado estadístico de precipitación mediante modelos ARIMA, Chattopadhyay y Chattopadhyay (2010) realizaron estudios con el Monzón de verano y Narayanan et al. (2013) con el período pre- Monzónico (marzo, abril y mayo).

En los últimos años se han implementado diferentes modelos estadísticos en el continente africano. Estudios relacionados con los modelos SARIMA fueron realizados en distintas municipalidades de Ghana teniendo como objetivo obtener un modelo de precipitación mensual, como así también un pronóstico de la misma con el propósito de mitigar impactos negativos en el sector agricultor (Sampson et al., 2013). Afrifa-Yamohah et al. (2016), también, consideraron los modelos SARIMA para el pronóstico mensual de la precipitación en Brong Ahafo, otra región de Ghana. Los autores encontraron un modelo que representaba el comportamiento de un modo significativo para un período de dos años, teniendo en cuenta su validación. Otro de los países en donde se aplicaron los modelos SARIMA fue en una región de Kenia (Kibunja et al., 2014), donde los autores seleccionaron un modelo apropiado de pronóstico de precipitación y su precisión para pronosticar la misma. En Warri Tower, Nigeria, el modelado de la precipitación y su pronóstico fue

realizado con modelos estacionales ARIMA, encontrando un año de pronóstico adecuado (Eny y Adeyeye, 2015).

Por otro lado, el trabajo desarrollado por Mohammed (2015) podría ser un claro ejemplo de la aplicación de estos modelos estadísticos en el área hidrológica. En el mismo se pronosticó la cantidad de escorrentía para el año 2011 utilizando el promedio de todas las estaciones de medición dentro de cada estado de EE.UU, a partir de los modelos SARIMA y ARIMA. Por otro lado, en el mismo trabajo el autor sugiere que debido al cambio climático la escorrentía no puede ser pronosticada a más de un año.

Puesto que conocer el caudal de los ríos es de gran importancia para la navegación de los mismos, De Figueiredo y Blanco (2016) modelaron el nivel del agua del río Amazonas utilizando modelos SARIMA con un horizonte de pronóstico a 24 meses encontrando resultados satisfactorios. Asimismo, en el río Caroní, Venezuela, el modelado de la serie temporal mensual del caudal en el periodo 1950-2003 fue realizado mediante modelos SARIMA (Bello et al., 2010), encontrando mediante un análisis de los residuos un modelo adecuado. En Argentina, los modelos autoregresivos (AR) fueron implementados por Pierini et al. (2012) en el pronóstico del caudal diario del río Colorado, sin embargo, los resultados obtenidos por esta metodología no resultaron tan precisos al momento de la validación.

El objetivo principal de este capítulo fue seleccionar un modelo estocástico SARIMA (Wei, 2005), entre cien combinaciones de hiperparámetros, con el propósito de representar las series temporales del caudal mensual en tres estaciones de referencia: Túnel Subfluvial, Corrientes y Paso de los Libres. Además, con una adecuada selección de los hiperparámetros del modelo se pretendió desarrollar un pronóstico del caudal mensual con un horizonte a 6 y 12 meses. Finalmente, se analizó el caso particular de incorporar al modelo SARIMA el caudal esperado a seis meses que tenía en cuenta la oscilación climática del ENSO como variable regresora.

4.3. Datos y Metodología

Los datos medios mensuales (período $1975 - 2016$) de la Base de Datos Hidrológica Integrada (BDHI) publicada por la Subsecretaría de Recursos Hídricos de la República Argentina fueron considerados de las estaciones del río Paraná (Túnel Subfluvial y Co-

rrientes) y río Uruguay (Paso de los Libres). Un primer estudio exploratorio fue realizado y explicado en detalle para la estación Túnel Subfluvial en el período temporal $1975-2015$. El modelado estadístico empleado en este estudio consideró los modelos propuestos por Box y Jenkins (1976). Estos autores formularon modelos estocásticos para el estudio de las series de tiempo, como por ejemplo los modelos ARMA (Ec. 4.1):

$$Z_t - \phi_1 Z_{t-1} - ... - \phi_p Z_{t-p} = u_t - \theta_1 u_{t-1} - ... - \theta_q u_{t-q} \tag{4.1}$$

Donde $Z's$ es la serie temporal original, $\phi's$ y $\theta's$ representan los términos asociados a la parte autoregresiva (AR, del inglés Autoregresive) y de promedios móviles (MA, del inglés Moving Average), respectivamente; c es un término constante en el modelo. Los errores (u_t) siguen una distribución de ruido blanco con media cero y varianza σ^2, $u_t \sim N(0, \sigma^2)$. La combinación de los parámetros ϕ_i , con $0 \leq i \leq p$, y θ_i , con $0 \leq i \leq q$ da lugar a los modelos ARMA.

Sin embargo, debido al supuesto de estacionaridad en las series temporales que estos modelos presentan, los modelos autoregresivos integrados de medias móviles (ARIMA) fueron desarrollados. Estos modelos permiten transformar las series temporales no estacionarias a estacionarias.

$$\phi_i(B)(1-B)^d Z_t = \theta_q(B) u_t \tag{4.2}$$

con $u_t \sim N(0, \sigma^2)$.

En particular, en este trabajo se implementó el modelado estadístico considerando la presencia de cierta estacionalidad en las series temporales utilizando los modelos estacionales autoregresivos integrados de medias móviles (SARIMA). Los modelos SARIMA permiten estudiar series temporales que no presentan un proceso estacionario (cambios en el valor medio, varianza, o en la estructura de autocorrelación), como así también series temporales que presentan cierta estacionalidad (fluctuaciones periódicas). Los modelos SARIMA se representan de la siguiente manera: SARIMA$(p, d, q)(P, D, Q)_S$, siendo (p, d, q) la parte ordinaria, y $(P, D, Q)_S$ la parte estacional de la serie de tiempo, con s la duración de la fluctuación periódica (implementado por Hyndman, 2016). En general, el modelo puede ser expresado de la siguiente forma (Ec. 4.3):

$$\phi_{PS}(B_S)\phi_p(B)(1-B_S)^D(1-B)^d Z_t = \theta_{QS}(B_S)\theta_q(B)u_t \tag{4.3}$$

con $u_t \sim N(0, \sigma^2)$ y $D > 0$ al orden de la diferencia asociada a la parte estacional del modelo.

Es importante mencionar, que en todos los casos se consideró p y q menores o iguales a 4 y P y Q menores o iguales a 1 dando lugar a cien modelos posibles. Se buscó obtener modelos parsimoniosos (modelos sencillos), es decir, considerar un modelo con un orden mayor de parámetros podría significar un estudio más complejo y más costoso computacionalmente. Con el objetivo de modelar las series de caudal mediante el mejor modelo estadístico posible se consideraron métodos de identificación (función de autocorrelación (ACF), autocorrelación parcial (PACF)), estimación de los parámetros del modelo y métodos de diagnóstico. Para ello, el estudio de los residuos fue realizado mediante diferentes técnicas (ACF, PACF) y test estadísticos, como el de Ljung-Box con la hipótesis nula que los datos considerados se distribuyen de forma independiente (5 % de nivel de significancia). Para la selección del modelo, un cierto compromiso entre la ACF, PACF y el test de Ljung-Box fue establecido. Asimismo, para la elección del modelo se tuvieron en cuenta las siguientes dos métricas: el criterio de Akaike (AIC) y el coeficiente de eficiencia modelo de Nash–Sutcliffe (NSE). Este último índice resulta ser siempre menor que uno, donde valores cercanos a uno representan modelos adecuados mientras que valores negativos representan un mal rendimiento del modelo.
Por otra parte, dos pronósticos para el caudal mensual fueron desarrollados para el periodo 07/2016 a 12/2016 (6 meses), considerando el periodo de entrenamiento $1975 - 2016/6$, y un pronóstico de 12 meses para el periodo 01/2016 a 12/2016 considerando el periodo de entrenamiento $1975 - 2015$ en las estaciones de medición Túnel Subfluvial, Corrientes y Paso de los Libres, (Stoffer, 2016). (Ec. 4.4).

$$\phi_{is}(B_s)\phi_i(B)(1-B_s)^D(1-B)^d Z_{t+1} = \theta q(B_s)\theta q(B)u_{t+1} \tag{4.4}$$

Finalmente, para Túnel Subfluvial se consideró incorporar la variable de caudal esperado, obtenido en el Capítulo 3, al modelo correspondiente seleccionado como una variable exógena (Xie et al., 2013) para evaluar así el valor agregado de la misma en dicho modelo (Ec. 4.5). En este caso particular se tuvo en cuenta el caudal medio esperado estimado

a partir del índice NIÑO 3.4 ya que de acuerdo con los resultados finales anteriores del Capítulo 3 fue el modelo que obtuvo un mejor rendimiento. Es importante recordar que el caudal esperado se obtuvo de manera trimestral mientras que en este capítulo se evalúa un modelo mensual por lo cual se estimaron los datos mensuales del caudal esperado trimestral repitiendo el valor de cada mes dentro del trimestre, y luego se realizó un promedio móvil de orden tres para suavizar dicha serie. También se comparó el modelo univariado con el modelo que incluye una variable exógena utilizando una validación cruzada para cada modelo. Para ello, se truncó la serie en 200 puntos diferentes (comenzando con la serie completa y descartando de a un mes) y para cada serie truncada se generaron conjuntos de entrenamiento y testeo como se observa en la Figura 4.1, y se procedió a predecir valores para los tiempos del conjunto de testeo a partir de un modelo ajustado en el conjunto de entrenamiento y la variable regresora correspondiente al periodo de testeo. Para cada iteración se calcula la raíz del error cuadrático medio para el modelo con y sin variable exógena y finalmente se calcula el promedio de las diferencias relativas entre los errores de ambos. Es importante aclarar que como la variable exógena presenta como último dato el tercer trimestre del 2016, es decir los meses de junio-julio-agosto, se decidió realizar la validación hasta septiembre 2016.

$$\phi_P(B_s)\phi_p(B)(1-B_s)^D(1-B)^d Z_t = \theta_Q(B_s)\theta_q(B)u_t + \sum_{h=0}^{b} \beta_h X_{t-h} \tag{4.5}$$

con $u_t \sim N(0, \sigma^2)$, y $D > 0$ el orden de la diferencia asociada a la parte estacional del modelo, X_t es la variable externa.

4.4. Resultados

4.4.1. Análisis en la estación Túnel Subfluvial (río Paraná)

La serie de tiempo mensual correspondiente a la estación Túnel Subfluvial en el periodo mencionado es exhibida en la Figura 4.2. En la misma es posible notar un mayor número de valores extremos durante los primeros veinte años.

Por otro lado, es importante tener presente la estacionalidad y estacionaridad existente en la serie de tiempo de la estación considerada (análisis ya realizado en los capítulos

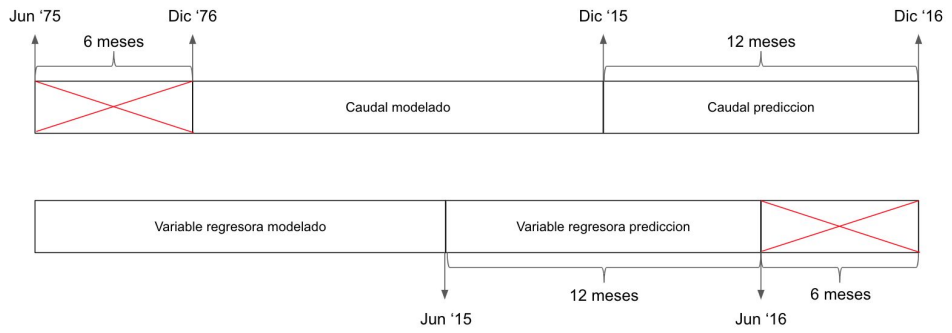


Figura 4.1: Ejemplo de generación de conjuntos de entrenamiento y testeo tanto para la serie de caudal como para la variable exógena. Caso particular de predicción a 12 meses para la serie terminada en Diciembre 2016. Notar que la variable regresora que se puede observar en los tiempos del conjunto de entrenamiento corresponde a caudal esperado 6 meses adelante en el tiempo.

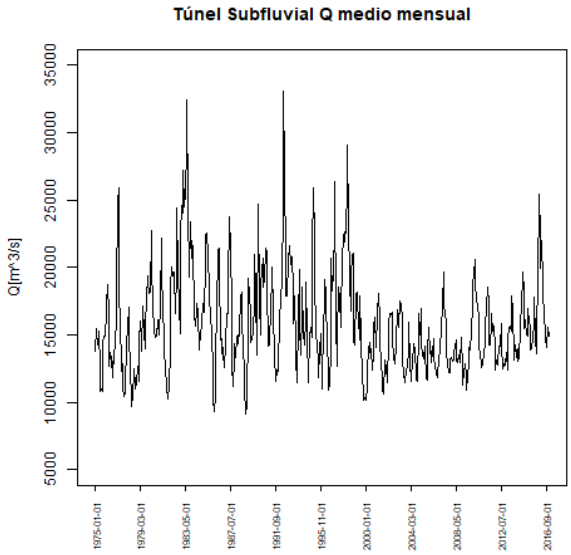


Figura 4.2: Serie temporal del caudal mensual de la estación Túnel Subfluvial para el periodo 1975-2015.

anteriores). En este sentido, habíamos notado claramente la presencia de una onda anual, con el máximo medio mensual de caudal entre los meses de marzo y abril, y el mínimo alrededor del mes de septiembre. Dentro de este estudio exploratorio del caudal mensual

de la estación Túnel Subfluvial se analizó la autocorrelación y la autocorrelación parcial (Figura 4.3).

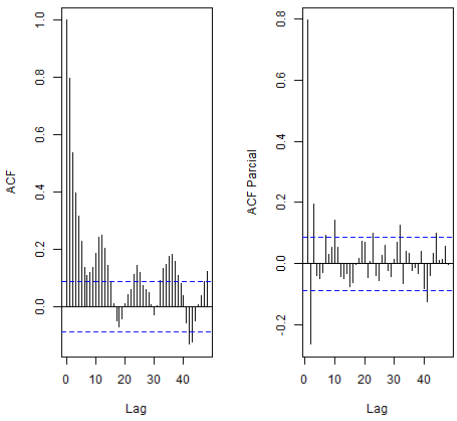


Figura 4.3: ACF (izq.) representa la función de autocorrelación y PACF (der.) representa la función de autocorrelación parcial de la serie de caudal mensual en la estación Túnel Subfluvial para el periodo 1975-2015.

En la figura anterior se exhiben los valores de la autocorrelación de la serie de caudal, donde se destaca a priori que los meses múltiplos de doce presentan una alta y significativa correlación. Motivo por el cual se decide nuevamente chequear la estacionalidad diferenciando a la serie de tiempo una vez. En la Figura 4.4 es posible ver con mayor detalle la estacionalidad de la serie (onda anual). Asimismo, de la función de autocorrelación parcial se observa que el lag 3 resulta ser significativo indicando en principio que el hiperparámetro p podría ser del orden de tres.

De acuerdo a las 100 posibles combinaciones de hiperparámetros se estableció de acuerdo al valor de AIC que el más adecuado era un modelo SARIMA$(3,0,0)(0,1,1)_{12}$ para los dos periodos de entrenamiento mencionados en la sección de metodología. En las Tablas I y II (ver Apéndice B.1) se presenta los valores obtenidos de AIC para las diferentes combinaciones, así como también el valor de NSE para cada uno de ellos para los periodos de entrenamiento considerados. En particular se destaca que para la combinación seleccionada encontramos un valor de NSE igual a 0.70. En el caso del periodo de entrenamiento $1975-junio/2016$ el valor de NSE fue también de 0.70. Para la combinación seleccionada, a modo de ejemplo se exhibe la Figura 4.5 de ACF, PACF y el test de Ljung-Box de los

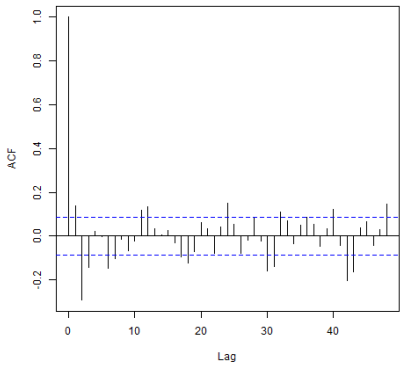
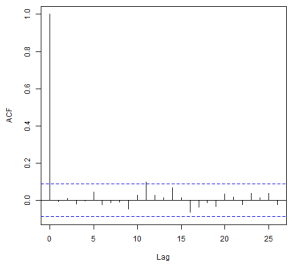
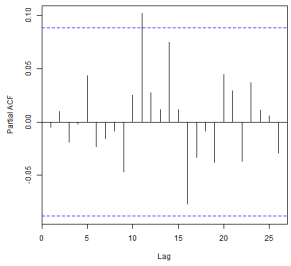


Figura 4.4: Función de autocorrelación (ACF) de la serie diferenciada en la estación Túnel Subfluvial del caudal mensual correspondiente al periodo 1975-2015.

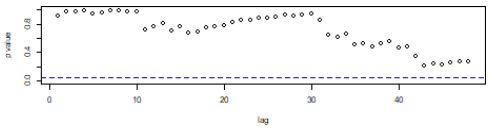
residuos para el primer periodo. A partir de la misma es posible observar que tanto en la ACF como la PACF se destaca el lag 11 significativo.



(a)



(b)



(c)

Figura 4.5: ACF, PACF, y la serie temporal de residuos estandarizados junto con los p-valores del test Ljung-Box correspondientes a la estación Túnel Subfluvial para el periodo de entrenamiento $1975 - 2015$.

4.4.2. Pronóstico SARIMA para la estación Túnel Subfluvial a 6 y 12 meses

Teniendo en cuenta el modelo anterior se calculó el pronóstico para ambas ventanas temporales (6 y 12 meses). Los resultados se presentan en las Figuras 4.6 y 4.7. Es posible observar a partir de las figuras anteriores que si bien para el pronóstico a 6 meses la variable a predecir se encuentra dentro del intervalo de confianza generado por el modelo, no se logra obtener la ubicación correcta de los puntos mínimos en esos tramos. Por otra parte, en el pronóstico a 12 meses lo más importante a notar es la subestimación de los máximos que se encuentran en enero y abril de 2016.

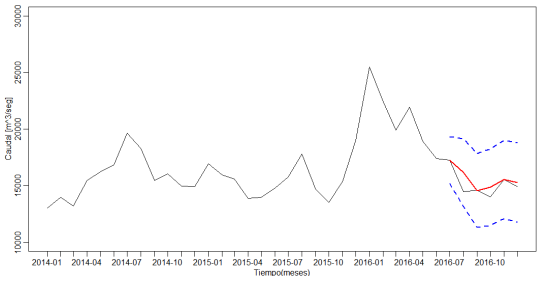


Figura 4.6: Pronóstico a 6 meses de la estación Túnel Subfluvial. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

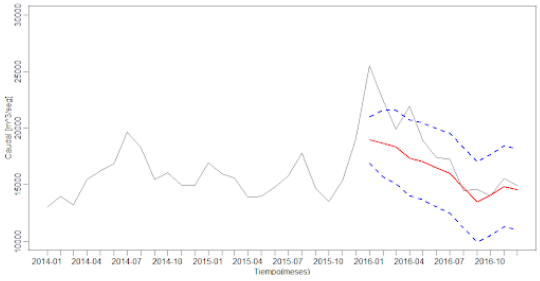


Figura 4.7: Pronóstico a 12 meses de la estación Túnel Subfluvial. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

4.4.3. Pronóstico SARIMA para la estación Corrientes a 6 y 12 meses

En el caso de Corrientes la combinación de los hiperparámetros para los periodos de entrenamiento dio lugar al modelo SARIMA$(3,0,3)(0,1,1)_{12}$ (Tablas III y IV, ver Apéndice B.2). Los valores del coeficiente de NSE resultaron ser en ambos periodos de entrenamiento igual a 0.63. Por otra parte, de acuerdo al análisis de los residuos de la ACF, PACF, y el test de Ljung-Box para el modelo propuesto se encontró únicamente que la ACF presentó el lag 11 significativo. Los resultados se presentan en las Figuras 4.8 y 4.9. En el pronóstico a 6 meses, si bien nuevamente el caudal se encuentra dentro del intervalo de confianza, el modelo presenta una sobrestimación en los primeros meses del pronóstico que luego da lugar a una subestimación del máximo relativo del mes de agosto. En el caso del pronóstico a 12 meses es interesante destacar las dificultades que presenta el modelo en los primeros meses de pronóstico. Sin embargo, es importante destacar que el año 2015/2016 resultó estar afectado en la región de la Cuenca del Plata por el fenómeno del ENSO.

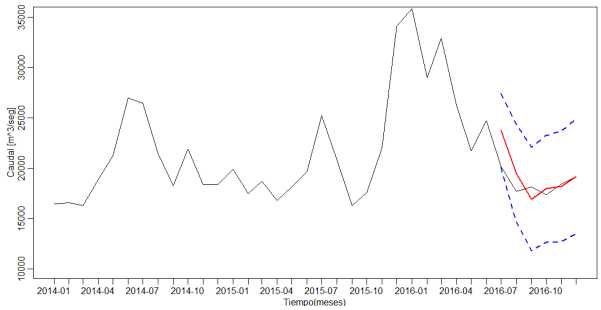


Figura 4.8: Pronóstico a 6 meses de la estación Corrientes. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

4.4.4. Pronóstico SARIMA para la estación Paso de los Libres a 6 y 12 meses

Para Paso de los Libres la combinación de los hiperparámetros para los periodos de entrenamiento dio lugar al modelo SARIMA$(3,0,2)(1,1,1)_{12}$ y el NSE para los dos periodos de entrenamiento fue igual a 0.36 (Tabla V y VI, ver Apéndice B.3). De acuerdo con el

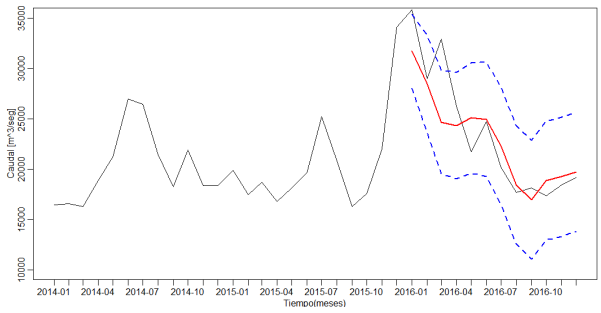


Figura 4.9: Pronóstico a 12 meses de la estación Corrientes. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

análisis de los residuos de la ACF, PACF, y el test de Ljung-Box para el modelo propuesto no se encontraron lag significativos.

Los resultados se presentan en las Figuras 4.10 y 4.11. En el pronóstico a 6 meses sí hay una mejora con respecto a los resultados de las estaciones anteriores donde se observa una mejor ubicación de los momentos máximos y mínimos. En el caso del pronóstico a 12 meses se destaca una sobrestimación del caudal para los primeros meses del pronóstico.

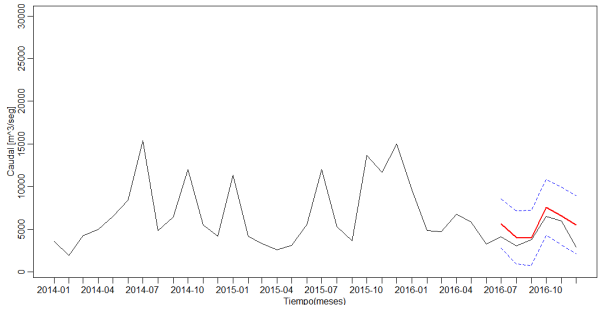


Figura 4.10: Pronóstico a 6 meses de la estación Paso de los Libres. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

4.4.5. Incorporar variable exógena mediante modelos SARIMAX

De acuerdo a los resultados anteriores hallados para la estación Túnel Subfluvial donde se observó particularmente que el pronóstico a 12 meses durante los primeros meses presenta una subestimación del caudal, posiblemente asociado a la influencia del fenómeno ENSO, se propone considerar la variable de caudal estimado del Capítulo 3 como variable exógena

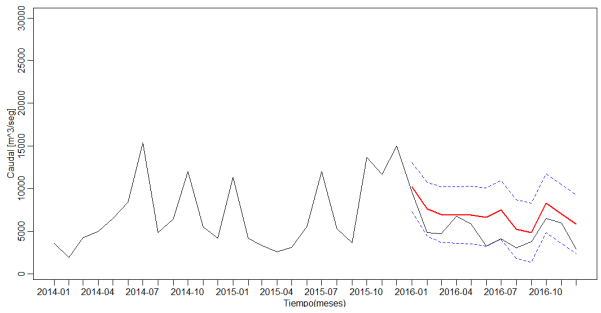


Figura 4.11: Pronóstico a 12 meses de la estación Paso de los Libres. La línea roja representa el pronóstico del modelo considerado, la línea negra es la serie observada, mientras que en azul representa un desvío del pronóstico realizado.

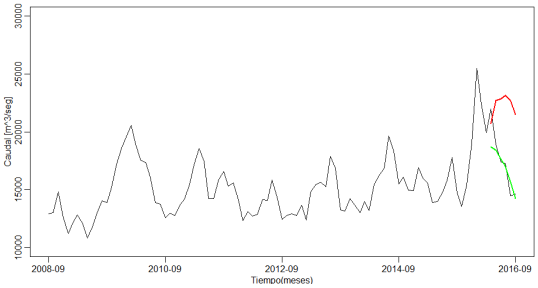
al modelo seleccionado en la sección anterior.

Primeramente, se comparó el modelo sin regresora con el SARIMAX mediante la raíz del error cuadrático medio por medio de una validación cruzada de doscientas iteraciones de a un paso tanto para el pronóstico de 12 meses como de 6 meses. Como resultado se encontró que para el pronóstico a 12 meses, el modelo SARIMAX es 2.44 % peor que el modelo univariado, mientras que para el pronóstico a 6 meses resultó ser 1 % mejor. Sin embargo, es interesante destacar que a pesar de que en promedio la incorporación de la variable exógena no parece mejorar el modelado, si se puede notar una diferencia a la hora de pronosticar bajo un evento extremo. Para evidenciar este fenómeno, reportamos el rendimiento relativo del modelo SARIMAX respecto al univariado en las primeras 12 iteraciones de la validación cruzada, correspondientes a la predicción de periodos que finalizan durante octubre 2015 - septiembre 2016, periodo que incluye un evento Niño. Asimismo se incluyen las Figuras 4.12 y 4.13 a modo de ejemplos de pronósticos de los últimos 6 y 12 meses que involucran un periodo donde el fenómeno ENSO afectó la región de estudio. Asimismo, en las Tablas I y II se muestra la diferencia entre los errores de ambos modelos para distintos cortes de la serie. Para cada mes se truncó la serie total en ese periodo y se separó en conjuntos de entrenamiento y testeo. Se calculó la raíz del error cuadrático medio en el conjunto de testeo para el modelo con y sin variable exógena y finalmente se computó la diferencia relativa entre los errores de ambos. Se reporta el cambio relativo del error obtenido en el modelo con variable exógena respecto al error obtenido sin variable exógena. De este modo, valores negativos implican que el modelo SARIMAX presentó un mejor rendimiento. Cabe destacar que en los casos reportados en

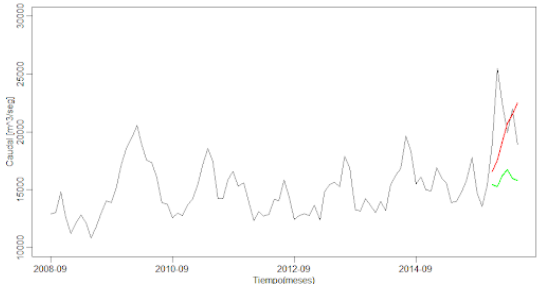
ambas tablas, el conjunto de testeo coincide total o parcialmente con un evento Niño.

Tabla I: *Porcentual de la raíz cuadrada del error cuadrático medio para seis meses de la estación Túnel Subfluvial.*

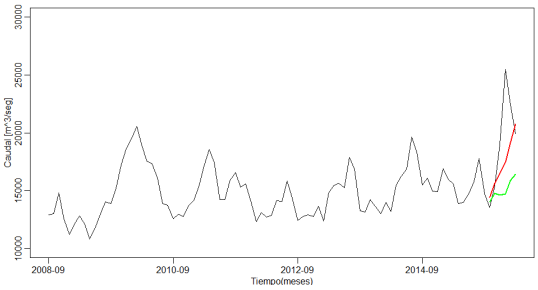
FIN FECHA TESTEO	6 MESES
SEPTIEMBRE 16	74.79 %
AGOSTO 16	68.42 %
JULIO 16	65.75 %
JUNIO 16	-0.54 %
MAYO 16	-50.22 %
ABRIL 16	-62.57 %
MARZO 16	-51.22 %
FEBRERO 16	-38.48 %
ENERO 16	-40.4 %
DICIEMBRE 15	-61.29 %
NOVIEMBRE 15	-47.55 %
OCTUBRE 15	-46.45 %



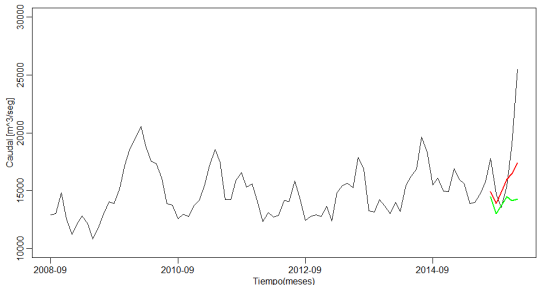
(a) Fecha fin testeo septiembre 2016.



(b) Fecha fin testeo mayo 2016.



(c) Fecha fin testeo abril 2016.



(d) Fecha fin testeo enero 2016.

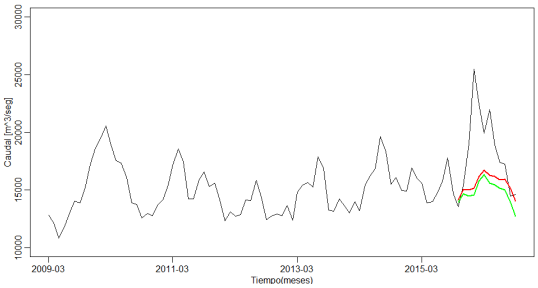
Figura 4.12: Ejemplos de pronóstico a seis meses del caudal mensual en la estación Túnel Subfluvial de a un paso de octubre 2015 hasta fecha fin de testeo septiembre 2016. La línea roja representa el modelo que incluye variable regresora mientras que la línea verde resulta ser el modelo simple.

De acuerdo con la Tabla I y la Figura 4.12 se observa que a medida que se pronostica cada vez más cerca de los valores máximos relativos que ocurren al inicio del 2016 el error disminuye. Lo anterior tiene sentido, puesto que la variable exógena incorporada al modelo es estimada a partir del índice NIÑO 3.4 y el periodo temporal en cuestión se encontró asociado con la influencia del fenómeno ENSO.

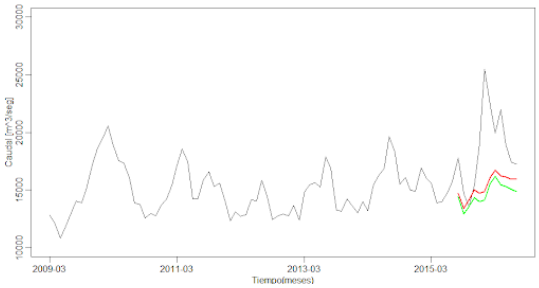
Tabla II: *Porcentual de la raíz cuadrada del error cuadrático medio para doce meses de la estación Túnel Subfluvial.*

FIN FECHA TESTEO	12 MESES
SEPTIEMBRE 16	-9.86 %
AGOSTO 16	-9.37 %
JULIO 16	-11.89 %
JUNIO 16	-13.04 %
MAYO 16	-12.97 %
ABRIL 16	-12.76 %
MARZO 16	-11.46 %
FEBRERO 16	-10.99 %
ENERO 16	-8.29 %
DICIEMBRE 15	-11.46 %
NOVIEMBRE 15	-10.99 %
OCTUBRE 15	-8.29 %

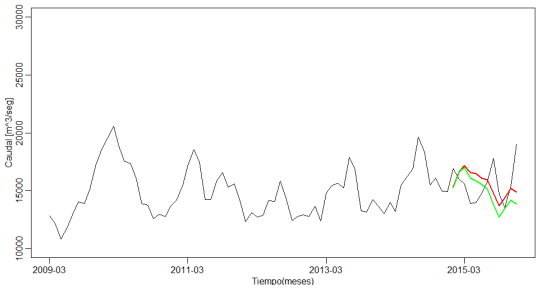
En el caso del pronóstico a doce meses se observa de la Tabla II y la Figura 4.13 que al incorporar una variable exógena se obtiene una mejora teniendo en cuenta el error porcentual cuadrático medio relativo al modelo simple. Sin embargo, es importante notar que el pronóstico a seis meses en ciertos pasos presentó una mejor estimación del caudal mensual esperado para el periodo en cuestión analizado.



(a) Fecha fin testeo septiembre 2016.



(b) Fecha fin testeo julio 2016.



(c) Fecha fin testeo enero 2016.

Figura 4.13: Ejemplos de pronóstico a doce meses del caudal mensual en la estación Túnel Subfluvial de a un paso desde octubre 2015 hasta fecha fin de testeo septiembre 2016. La línea roja representa el modelo que incluye variable regresora mientras que la línea verde resulta ser el modelo simple.

4.5. Conclusiones

El continuo análisis hidrológico en la Cuenca del Plata es necesario dado el potencial riesgo de valores extremos de caudal (inundaciones y bajantes) que provocan daños irreparables en regiones dentro de la cuenca. En este sentido, el monitoreo del caudal mensual es fundamental. Para lo cual, el modelado estadístico resultaría una herramienta útil para mejorar el sistema de alerta y control en el futuro cercano. De esta manera, en este trabajo se empleó el modelado estadístico de acuerdo a los modelos SARIMA, propuesto por Box y Jenkins, para ciertas estaciones localizadas en la parte baja de la Cuenca del Plata.

A través de herramientas de identificación, estimación y diagnóstico se propuso modelar las series temporales de 41 años de las estaciones Túnel Subfluvial, Corrientes y Paso de los Libres. Para ello, todas las posibles combinaciones de modelos SARIMA con p y q entre cero y cuatro, P y Q entre cero y uno fueron contempladas, y se seleccionó mediante la métrica AIC el mejor modelo para cada serie de tiempo. Asimismo, se estimó el

valor de NSE, para los dos periodos de entrenamiento considerados en este capítulo, que indicaron que el modelo no presentó un bajo rendimiento. Sin embargo, en la realización de los pronósticos a seis y doce meses mediante el método sin variable exógena se encontraron ciertas particularidades en lo que refiere a la sobreestimación y la subestimación del caudal mensual.

Como caso particular de estudio se realizó un análisis teniendo en cuenta como variable regresora el caudal medio estimado de manera condicional del Capítulo 3 de la estación Túnel Subfluvial. En este sentido, mediante el uso de validación cruzada se realizó una comparación entre modelos. Se encontró que en promedio la variable regresora no mejora el rendimiento de los modelos. Sin embargo se pudo observar ciertas mejoras en lo que refiere al pronóstico en los sectores de la serie que presentan valores extremos. Este último resultado es coherente con el desarrollo de la presente tesis, puesto que en el capítulo anterior encontramos que la relación entre el índice NIÑO 3.4 y el caudal es más fuerte cuando nos encontramos bajo la influencia del fenómeno. Más aún, la distribución conjunta encontrada funcionaba particularmente bien en la cola correspondiente a valores altos del índice y del caudal, sugiriendo que la señal que genera el índice puede no ser tan fuerte para valores medios.

Si bien en el estudio para Túnel Subfluvial se propuso seleccionar una combinación de hiperparámetros óptima para el modelo SARIMA, y luego se incorporó al mismo una variable regresora (caudal esperado), en este sentido observemos que también habría sido posible realizar una nueva selección de hiperparámetros utilizando la variable exógena desde un principio. Sin embargo, no se espera encontrar grandes cambios puesto que la relación entre la variable exógena y el caudal sólo es fuerte en ventanas de tiempo acotadas y relacionadas con eventos extremos de la fase Niño.

En base a los resultados obtenidos, los modelos estudiados debido a su rápida implementación podrían contribuir al monitoreo de los caudales. Sin embargo, resulta necesario profundizar en otras metodologías y en una escala temporal que permita considerar una variabilidad de mayor frecuencia. Es decir, se entiende que dicha consideración permitiría contemplar los eventos de crecida o bajantes. Por lo tanto, se propone en el próximo anexo profundizar en el estudio de las series temporales diarias de caudal con el fin de obtener un pronóstico más adecuado.

4.6. Referencias

Afrifa-Yamohah E., Bashiru I., Saeed, I., Azumah, K., 2016. Sarima modelling and forecasting of monthly rainfall in the BrongAhafo region of Ghana. WorldEnvironment 6(1), 1-9.

Bello, J., Vega González, C., Perez, E., 2010. Análisis estocástico de una serie cronológica de caudales mensuales. Caso de estudio: Río Caroní, Venezuela. Revista Ingeniería UC, Vol. 17, No. 1, 46-54.

Box, G., Jenkins, G., 1976. Time series analysis: forecasting and control. Holden Day, San Francisco, CA.

Burlando, P., Rosso, R., Cadavid, L., Salas, J., 1992. Forecasting of short-term rainfall using ARMA models. Journal of Hydrology 144, 193-211.

Chattopadhyay, S., Chattopadhyay, G., 2010.Univariate modelling of summer-monsoon rainfall time series: comparison between ARIMA and ARNN. C. R. Geoscience 342, 100–107.

Cadenas, E., Rivera, W., 2007. Wind speed forecasting in the SouthCoast of Oaxaca, México, Renewable Energy. Vol. 32, no. 12, pp 2116-2128.

De Figueiredo, N., Cavalcante Blanco, C., 2016. Water level forecasting and navigability conditions of the Tapajós River- Amazon-Brazil. La Houille Blanche. Vol. 3, 53-64.

Eny, D., Adeyeye, F., 2015. Seasonal ARIMA modeling and forecasting of rainfall in Warri Town, Nigeria. Journal of Geoscience and Environment Protection. Vol. 3, 91-98.

Hyndman, R., 2016. Forecasting functions for time series and linear models. Package 'forecast'. http://github.com/robjhyndman/forecast.

Kibunja, H., Kihoro, J., Orwa, G., Yodah, W., 2014. Forecasting precipitation using SA-

RIMA model: a case study of Mt. Kenya region. Mathematical Theory and Modeling. Vol.4, No.11.

Mohammed, V., 2015. Long-term runoff study using SARIMA and ARIMA models in the United States. Meteorol. Appl. 22, 592–598.

Narayanan, P., Basistha, A., Sarka,r S., Sachdeva, K. , 2013. Trend analysis and ARIMA modelling of pre-monsoon rainfall data for Western India. C. R. Geoscience 345, 22–27.

Pierini, J., Goméz, E., Telesca, L., 2012. Prediction of water flows in Colorado River, Argentina. Lat. Am. J. Aquat. Res. 40(4), 872-880.

Sampson, W., Suleman, N., Gifty, A., 2013. Proposed Seasonal Autoregressive Integrated Moving Average model for forecasting rainfall pattern in the Navrongo municipality of Ghana. Journal of Environment and Earth Science 3(12), 80-85.

Stoffer, D., 2016. Applied Statistical Time Series Analysis.Package 'astsa'. http://www.stat.pitt.edu/stoffer/tsa4/.

Torres, J.L., García, A., De Blasa, M., De Francisco, A., 2005. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Solar Energy. Vol. 79, Issue 1, 65-77.

Toth, E., Brath, A., Montanari, A., 2000. Comparison of short-term rainfall prediction models for real-time flood forecasting. Journal of Hydrology 239, 132–147.

Wei, W., 2005. Time series analyisis, univariate and multivariate methods, Second edition Addison Wesley Longman IncDiv Pearson, 1-624.

Apéndice A

Predicción del Caudal en el Paraná Medio

A.1. Resumen

Para este nuevo estudio se consideró la variabilidad diaria del caudal, puesto que en el capítulo anterior se había hallado que ciertos meses debido a crecidas o bajantes no del todo esperadas no pudieron ser captadas por el modelo SARIMA propuesto. De esta manera, en el presente anexo se propone la comparación entre 4 modelos estadísticos: VAR, SARIMA, SARIMAX, ODPC para analizar la predicción del caudal diario en la estación Corrientes (río Paraná). Asimismo, se tuvieron en cuenta variables input que brindan información temporal de estaciones que se encuentran hacia el norte y este de la estación Corrientes. Las series temporales input que presentaron el mejor modelo teniendo en cuenta el error cuadrático medio de predicción fueron las que se obtuvieron a partir de la técnica de componentes principales dinámicas. Se encontró que la capacidad de las componentes principales dinámicas resultó fundamental a la hora de considerar variables regresoras, y que utilizar dicha metodología permitió de alguna manera obtener un mejor resultado en la evaluación del pronóstico en comparación con el método univariado SARIMA.

A.2. Introducción

Un buen empleo de los recursos hídricos de un país implica proveer en consecuencia un adecuado manejo en el transporte fluvial, energía hidroeléctrica, turismo, y aprovisionamiento de agua. En este sentido, es esencial un conocimiento adecuado de los posibles cambios en el caudal de manera tal de contribuir con dicha información a los tomadores de decisión. Asimismo, la predicción del caudal se encuentra segmentado en dos categorías temporales importantes: pronóstico de corto plazo (horario y diario) siendo crucial para operaciones de inundaciones y mitigación de sistemas, y pronóstico de largo plazo (semanal, mensual y anual) que resulta sumamente importante para la generación de energía, el transporte de sedimentos, manejo de planes de irrigación y otras aplicaciones (Yaseen et al., 2015).

En cuanto a modelos estadísticos empleados en el área de recursos hídricos diversos ejemplos y modelos pueden mencionarse. En este sentido, Asaad Shamsehaul (2010) exploró el uso de redes neuronales (ANN) para el pronóstico del caudal del Nilo Azul en Sudán. Estos modelos utilizaron el índice de precipitación como un input externo, y además, los autores tuvieron en cuenta las recientes y ocurrentes lluvias otorgándole una mayor preponderancia a estos eventos. Asimismo, los modelos que utilizaron el caudal estacional esperado y el índice de precipitación como inputs externos presentaron el mejor desarrollo. Además, los resultados exhibieron la capacidad de las ANN en ser utilizadas para predecir el caudal en países en desarrollo. De esta manera el trabajo muestra como herramientas y técnicas hidro-informáticas pueden ser utilizadas para lidiar con los problemas de recursos hídricos en países en desarrollo. Asimismo, en el caso de países pobres y densamente poblados, Berz (2001) noto que las inundaciones continúan siendo la causa de grandes muertes en relación con otros desastres naturales. En países en desarrollo, las inundaciones también pueden tener impactos significativos en la economía. Sin embargo, la reducción de pérdidas por inundación también pueden ser logradas adoptando varias medidas no estructurales, como regular el uso del suelo en zonas de planicies inundables, mejorar el diseño de las viviendas, desarrollar modelos de pronóstico, educación pública, planes de emergencia y seguros de inundación. De todas maneras, la potencialidad por inundaciones rápidas y sus daños está en aumento en varias regiones debido al desarrollo social y económico. Además, las rápidas inundaciones son esperadas que aumenten en frecuencia y severidad debido al impacto del cambio global en el clima (Huntington,

2006; Borga et al., 2011)

En Dehghani et al. (2014) se estudiaron dos escenarios de pronóstico de sequía del río Karon, en Irán. Un escenario contempló convertir a las series mensuales de caudal en series de índice estandarizado hidrológico de sequía (SHDI), y las ANN fueron consideradas teniendo en cuenta como entrenamiento de las mismas estos nuevos conjuntos de datos. En el segundo escenario, las series fueron directamente utilizadas para pronóstico y luego a partir de las series de pronóstico obtenidas se convirtieron las mismas en series de SHDI. Los resultados destacaron que el primer procedimiento presentó un mejor desempeño. Además, los autores utilizaron el procedimiento de componentes principales (PCA) y Forward Selection (Stepwise) para remover dependencia y reducir el número de variables input. De esta manera, la realización de un pronóstico de sequía resulta necesario ya que permite generar una colaboración directa con los tomadores de decisión a la hora de tomar medidas en el contexto del manejo de recursos hídricos.

Por otra parte, las técnicas de modelos autoregresivos de promedios móviles han sido los modelos estadísticos más utilizados para el pronóstico del caudal, pero los mismos han mostrado ser deficientes en representar las estructuras no-lineales dinámicas inherentes en la transformación de los datos de caudal. Por lo que, Fashae et al. (2018) comparó un modelo de ANN con el modelo ARIMA en el río Opeki en Nigeria. El desempeño de los dos modelos fue sujeto a dos indicadores, por un lado, el coeficiente de correlación, y por otro, el error cuadrático medio de predicción. Los resultados obtenidos destacaron un mejor rendimiento para el modelo ARIMA.

Con una creciente demanda por los recursos de agua, incertidumbres en estimar la capacidad de agua, predecir etapas de inundación y áreas de inundación, áreas de bajo caudal y sequía hidrológica; la predicción del caudal resulta de suma importancia para una mitigación efectiva y manejo de inundaciones y bajantes, como así también de desastres naturales asociados a los recursos hídricos (Maher y Eyre, 2012).

Si bien en el capítulo anterior se implementaron los modelos SARIMA en el caudal mensual de estaciones de la Cuenca del Plata, fue notable como cierta variabilidad no pudo ser captada por el modelo. De esta manera, en el presente apéndice se propuso considerar una variedad de modelos, junto con la inclusión de una escala menor de tiempo a la empleada anteriormente, de manera de intentar captar esa variabilidad de mayor frecuencia que podría estar asociada a bajantes y crecidas abruptas en el caudal, observado en el

capítulo anterior.

En síntesis, la realización de este apéndice tuvo como principal objetivo la implementación de un pronóstico estadístico de caudal medio diario que no solo presentase un resultado adecuado, sino que además, el modelo resultase de una interpretación sencilla de manera tal que pueda ser aplicado y explicado a una comunidad no necesariamente científica. De esta manera, se realizó un pronóstico en la estación Corrientes del río Paraná, presentando como fronteras de pronóstico 1, 3 y 5 días. Asimismo, considerando la ubicación estratégica de la estación de medición en Corrientes, la generación de mayor información resulta esencial, puesto que una continúa contribución al pronóstico diario del caudal podría ser de utilidad tanto a fines de análisis de capacidad energética, como de almacenamiento de agua en la región, como así también para eventuales medidas de alerta de inundación y evacuación.

A.3. Datos y Metodología

Para la realización de este estudio se tuvieron en cuenta los datos medios diarios provenientes de la Base de Datos Hidrológica Integrada (BDHI) publicada por la Subsecretaría de Recursos Hídricos de la República Argentina de las estaciones de medición de Corrientes (río Paraná), Puerto Pilcomayo (río Paraguay) e Itatí (río Paraná). Los datos se corresponden con el período 26/11/1999 al 31/12/2015.

Datos faltantes

Los datos faltantes representaban menos del 10 % en todas las estaciones de medición. De todas maneras para subsanar el dato faltante se completó el mismo mediante una interpolación lineal o considerando el mes más parecido de otro año en el caso de que sean varios datos faltantes consecutivos (superior a dos) analizando diferencias de 4 meses previos y posteriores a los datos faltantes.

Posibles transformaciones a los datos que fueron analizadas:

- Considerar datos no más allá del percentil 95
- Evaluar el logaritmo de los datos y trabajar con estos datos transformados

Con respecto a la primera transformación, la misma sería pensar una idea más robusta. Por un lado, es probable que el modelo empleado presente un mejor desempeño al contar con menos extremos de los que en verdad deberían ser, sin embargo, se termina perdiendo el verdadero valor del dato.

Por otra parte, considerar el logaritmo de los datos presenta la desventaja de que para el usuario los valores con los que se enfrentaría serían distintos, por lo tanto habría que transformar y luego del pronóstico antitransformar nuevamente los datos, pero podría resultar útil para reducir la variabilidad de la serie. De todas maneras, esta última opción fue implementada, y teniendo en cuenta distintas fronteras de pronóstico 1 día, 3 y 5 días, se estimó la diferencia del valor del caudal de un cierto día con respecto al día de horizonte de pronóstico.

De esta manera, siendo el objetivo de este estudio la implementación de un pronóstico estadístico de caudal medio diario, que no solo presente un resultado adecuado sino que además el modelo resulte de una interpretación sencilla de manera tal que pueda ser aplicado y explicado a una comunidad no necesariamente científica, se consideraron los siguientes modelos de series de tiempo:

- SARIMA
- SARIMAX
- VAR
- ODPC

A continuación, se realizó una breve descripción de cada metodología empleada en el presente estudio.

SARIMA

Los modelos SARIMA como mencionamos en el capítulo anterior permiten estudiar series temporales que no presentan un proceso estacionario, como así también series temporales que presentan cierta estacionalidad. Los modelos SARIMA se representan de la siguiente manera: SARIMA$(p, d, q)(P, D, Q)_S$, siendo (p, d, q) la parte ordinaria, y $(P, D, Q)_S$ la parte estacional de la serie de tiempo (implementado por Hyndman, 2016). En general, el modelo puede ser expresado de la siguiente forma (Ec. A.1):

$$\phi_{PS}(B_S)\phi_p(B)(1-B_S)^D(1-B)^d Z_t = \theta_{QS}(B_S)\theta_q(B)u_t \tag{A.1}$$

SARIMAX

Incorporar variables externas a un modelo puede mejorar el pronóstico de los valores futuros (Xie et al., 2013). El adicionar una variable externa como input en un modelo SARIMA da lugar al siguiente modelo:

$$\phi_P(B_s)\phi_p(B)(1-B_s)^D(1-B)^d Z_t = \theta_Q(B_s)\theta_q(B)u_t + \sum_{h=0}^{b}\beta_h X_{t-h} \tag{A.2}$$

$D > 0$ el orden de la diferencia asociada a la parte estacional del modelo, X_t es la variable externa.

VAR

El modelo vector autoregresivo (VAR) es una extensión del modelo univariado autoregresivo, y resulta útil cuando uno está interesado en predecir múltiples series temporales utilizando un único modelo. En particular, para este análisis se utilizó el procedimiento establecido por Pfaff (2008), donde estima el VAR considerando métodos de regresión lineal para cada ecuación. El método presenta la siguiente ecuación (Ec. A.3):

$$Z_t = A_1 Z_{t-1} + ... + A_p Z_{t-p} + CX_t + u_t \tag{A.3}$$

Z_t es un vector de variables endógenas de $K \times 1$, y u_t representa una perturbación esférica de la misma dimensión. Las matrices $A_1, .., A_p$ son de dimensión $K \times K$. Asimismo, una constante o una tendencia puede ser incluída como regresor determinístico, como así también una variable dummy estacional o variables exógenas (término CX_t).

ODPC

La metodología tiene como objetivo principal encontrar un vector de una serie temporal de menor dimensión que logre representar la estructura dinámica de los datos (Peña et al., 2019). Es una técnica de reducción para series temporales multivariadas.

La metodología propone reducir la dimensión de la serie $Z_t \in R^m$, para ello dado $a \in R^{m(k+1)}$, $a = (a_o^{'}, ..., a_k^{'})$ con $\|a\| = 1$, $B \in R^{(k+1)m}$, con $(B)_{hj} = b_{h,j}$, de manera tal que

se define

$$MSE(a,B) = \frac{1}{T-(k_1+k_2)} \sum_{t=k_1+k_2+1}^{T} \sum_{j=1}^{m} (Z_{t,j} - \sum_{h=0}^{k_2} b_{h,j} f_{t-h}(a))^2, donde \tag{A.4}$$

$$f_t(a) = \sum_{h=0}^{k_1} a_n^{'} Z_{t-h}, t = k+1, ..., T \tag{A.5}$$

Y se busca $\hat{a}$ y $\hat{B}$ que minimizan $MSE(a,B)$. De esta manera se obtienen las componentes principales dinámicas como

$$\hat{Z}_t = \sum_{h=0}^{k} \hat{b}_h f_{t-h}(\hat{a}) \tag{A.6}$$

Las componentes principales dinámicas son definidas como una combinación lineal de los valores presente y pasado de las series temporales que minimicen la reconstrucción del error cuadrático medio.

El número de lags (k_i, con $i = 1, 2$) utilizados en la reconstrucción son parámetros a determinar y tienen que ser elegidos. Se denomina k_1 al número de lags usados para definir la componente principal dinámica y k_2 es el número de lags utilizados para reconstruir. En particular en este trabajo se consideraron tres tipos de combinaciones. Una primera combinación que consideró k_1 y k_2 igual a 2 (ODPC222), una segunda combinación k_1 y k_2 igual a 6 (ODPC662), y una tercera combinación k_1 y k_2 igual a 1 (ODPC112).

Validación Cruzada

Este procedimiento consiste en estimar el desempeño del modelo en diferentes conjuntos de entrenamiento y testeo. Para ello, se construyeron 1000 conjuntos de entrenamiento tomando una ventana deslizante de 5 años de longitud a lo largo de 1000 días consecutivos. Cada período de 5 años se utilizó para entrenar un modelo, con el que se pronosticó el siguiente valor de caudal correspondiente al horizonte de pronóstico (1, 3 o 5 días después de la finalización del período de entrenamiento).

En este sentido todos los modelos fueron entrenados con observaciones que ocurren previamente a la observación utilizada para el testeo. En la Figura A.1 se presenta de forma esquemática el procedimiento de elección de conjuntos para validación cruzada considerado en este trabajo, los puntos azules corresponden a las observaciones de entrenamiento mientras que el punto rojo es la observación de testeo.
Finalmente, se consideró como métrica el error cuadrático medio a lo largo de las 1000 iteraciones.

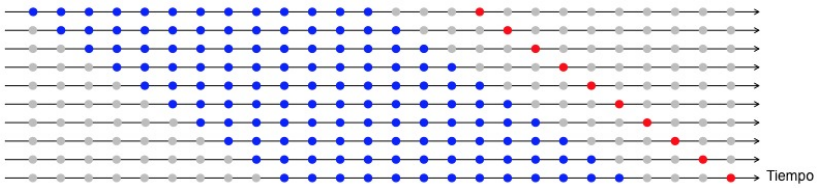
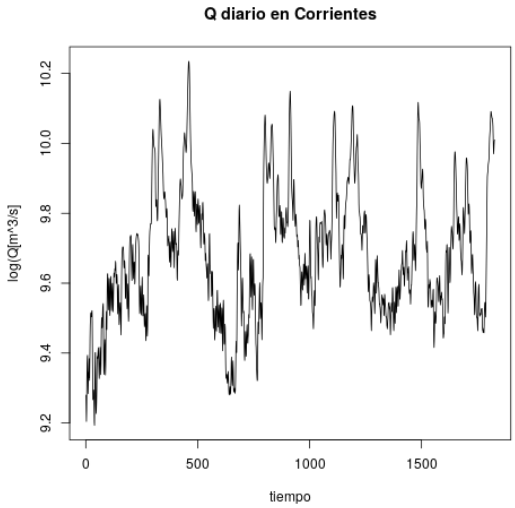


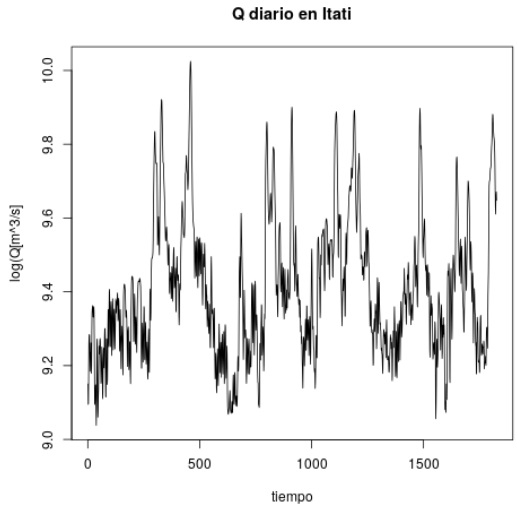
Figura A.1: Procedimiento de validación cruzada. Los puntos azules corresponden a las observaciones de entrenamiento mientras que el punto rojo es la observación de testeo.

A.4. Resultados

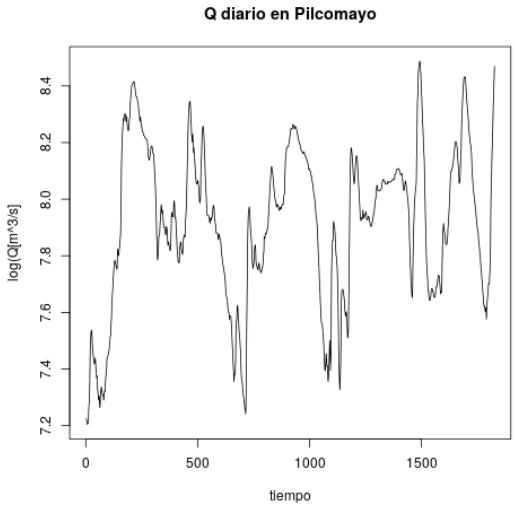
Primeramente se aplicó logaritmo a cada una de las series temporales para suavizar la variabilidad de las mismas, Figura A.2. Luego, las series de tiempo se diferenciaron con un cierto lag de acuerdo a la frontera de pronóstico que se quiso realizar.
Tres procedimientos distintos se realizaron para obtener el pronóstico diario de Corrientes utilizando los modelos estadísticos mencionados en la sección anterior. Un primer procedimiento consistió en modelar el conjunto de datos a partir de las metodologías VAR y ODPC. A partir de los mismos se consideraron los valores ajustados en las estaciones Pilcomayo e Itatí y se utilizaron como variables regresoras para el pronóstico de Corrientes utilizando un modelo SARIMAX. Un segundo procedimiento consistió en tener en cuenta las series asociadas a la estación Corrientes que se obtienen luego de realizar la aplicación del modelo VAR y ODPC. Con estas series, se realizó el mismo procedimiento anterior, es decir se empleó un modelo SARIMAX para pronosticar la serie de Corrientes ajustada por VAR y ODPC, utilizando como variables regresoras los ajustes de Pilcomayo e Itatí. Finalmente, el tercer procedimiento considerado consistió en no tener en cuenta variables regresoras y únicamente realizar cada ajuste de los modelos VAR y ODPC.



(a)



(b)



(c)

Figura A.2: Series de tiempo de caudal $[\frac{m^3}{s}]$ a) Corrientes, b) Itatí y c) Pilcomayo.

Asimismo, cada vez que se empleó el modelo ODPC se tuvieron en cuenta tres tipos de combinaciones asociada a los valores de k_1 y k_2. Todos los modelos se compararon con un modelo SARIMA donde se realiza un ajuste univariado a las serie temporal de Corrientes. De esta manera, a continuación se exhiben las Tablas I-III que contienen la diferencia relativa del error cuadrático medio de cada una de las metodologías y procedimientos realizados con respecto al modelo SARIMA univariado. Llamaremos al primer procedimiento mencionado como *Corrientes no ajustado*, *Corrientes ajustado* está asociado al segundo procedimiento, mientras que *Corrientes sin regresoras* al tercer procedimiento.

Tabla I: *ECM de 1 día.*

ECM en estación Corrientes, río Paraná					
	VAR	ODPC222	ODPC662	ODPC112	SARIMAX
Corrientes no ajustado	-0.03	-0.46	-0.46	-0.46	-0.44
Corrientes ajustado	0.54	0.62	0.8	0.71	-0.44
Corrientes sin regresoras	-0.48	-0.49	-0.5	-0.47	-0.44

Tabla II: *ECM de 3 días.*

ECM en estación Corrientes, río Paraná					
	VAR	ODPC222	ODPC662	ODPC112	SARIMAX
Corrientes no ajustado	0.17	0.08	0.1	0.02	0.01
Corrientes ajustado	0.67	0.22	0.53	0.32	0.01
Corrientes sin regresoras	0.87	0.39	0.38	0.33	0.01

Tabla III: *ECM de 5 días.*

ECM en estación Corrientes, río Paraná					
	VAR	ODPC222	ODPC662	ODPC112	SARIMAX
Corrientes no ajustado	0.18	0.05	0.07	0.05	0
Corrientes ajustado	0.42	0.02	0.05	0.22	0
Corrientes sin regresoras	1.14	0.12	0.14	0.05	0

A partir de las tablas anteriores se observa que el pronóstico a 1 día tanto para el primer como tercer procedimiento, el modelo ODPC presentó un menor error cuadrático medio frente al método univariado. Asimismo, es posible destacar que el modelo ODPC exhibió un mejor desempeño en la estimación de la diferencia del caudal a un día con respecto a cualquiera de las metodologías empleadas para esos dos procedimientos. En particular, se obtuvo una mejora entre un 46 % y un 50 % frente al modelo SARIMA utilizando dicha métrica mencionada.

Por otra parte, para los pronósticos a 3 y 5 días los resultados no fueron favorables puesto que se buscaba encontrar un mejor pronóstico frente al método univariado. Sin embargo, se destaca una variabilidad notable del error cuadrático medio entre los coeficientes utilizados en el ODPC. De esta manera, la realización de una validación de dichos parámetros podría resultar fundamental para obtener una mejora en el pronóstico.

A.5. Conclusiones

Encontrar uno o más modelos que resulten sencillos a la hora de explicar su funcionamiento resulta ser una tarea compleja. Varios modelos estadísticos pueden representar de manera eficiente el comportamiento de distintas series de tiempo pero carecen de posibilidad de una explicación comprensible de cómo es su desarrollo.

Asimismo, la comprensión por parte de los tomadores de decisión de los resultados de un modelo es fundamental para poder realizar las decisiones más óptimas en lo que refiere al manejo de los recursos naturales, con lo cual en este apéndice se propuso explorar distintos modelos de series de tiempo para evaluar la predicción del caudal diario en Corrientes

(estación del río Paraná) con horizonte de pronóstico de 1, 3 y 5 días. En particular, se quiso evaluar la predicción de una nueva metodología multivariada ODPC en series de caudal diario.

Más aún, mediante el uso de validación cruzada para series temporales se encontró que el modelo de ODPC resultó fundamental para lograr una reducción del error cuadrático medio con frontera de pronóstico a un día. Es decir, el rol de este modelo permitió de alguna manera destacar cierta variabilidad presente a lo largo de las series temporales. Más aún, el empleo de esta metodología dio lugar en cierta forma a filtrar el ruido de las series teniendo en cuenta la estructura temporal de los datos en conjunto. Sin embargo, los resultados no fueron satisfactorios para los pronósticos de 3 y 5 días. Por otro lado, a 5 días se encontraron mejores resultados para el modelo SARIMAX que no considera los ajustes de los modelos ODPC ni VAR, probablemente asociado a la menor variabilidad que presenta el modelo SARIMAX frente a los multivariados.

En conclusión, la combinación del modelo ODPC para obtener un suavizado en la estructura dinámica de los datos considerando, los mismos como variables regresoras en conjunto con el modelo SARIMA para predecir la estación Corrientes obtuvo el menor error cuadrático medio como pronóstico a un día, como así también no considerando variables regresoras. De esta manera, se logró encontrar un modelo que mejora un clásico modelo SARIMA que incluye variables regresoras a partir de un modelo multivariado como es el ODPC. Sin embargo, puesto que el resultado hallado no fue del todo favorable, ya que no se obtuvieron resultados óptimos de pronóstico a 3 y 5 días, en parte generó un nuevo conocimiento sobre las distintas aplicaciones que se podrían considerar sobre la metodología ODPC.

A.6. Referencias

Asaad Shamsehaul, 2010. Artificial neural network model for river flow forecasting in a developing country. Journal of Hydroinformatics. 12 (1), 22-35.

Berz, G., 2001. Discussion: flood disasters: lessons from the past—worries for the future. Proceedings of the Institution of Civil Engineers - Water and Maritime Engineering. Vol. 148 Issue 1, 57-58.

Borga, M., Anagnostou, E.N., Blöschl, G.. Creutine, J-D., 2011. Flash flood forecasting, warning and risk management: the HYDRATE project. Environmental Science & Policy Vol. 14, Issue 7, 834-844.

Dehghani, M., Saghafian, B., Saleh, F., Farokhnia, A., Noori, R., 2014. Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation. International Journal of Climatology. 34, 1169–1180.

Fashae, O.A., Olusola, A.O., Ndubuisi, I., Udomboso, C.G., 2018. Comparing ANN and ARIMA model in predicting the discharge of River Opeki from 2010 to 2020. River research and applications. 35, 169–177.

Hyndman, R., 2016. Forecasting Functions for Time Series and Linear Models.Package 'forecast'. http://github.com/robjhyndman/forecast.

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: review and synthesis. Journal of Hydrology. 319, 83–95.

Maher, D. y Eyre, B., 2012. Carbon budgets for three autotrophic Australian estuaries: Implications for global estimates of the coastal air-water CO2 flux. Global Biogeochemical Cycles, 26(1).

Pfaff, B., 2008. VAR, SVAR and SVEC Models: Implementation within R Package Vars. Journal of Statistical Software, 27, 1-32.

Peña, D., Smucler, E., Yohai, V., 2019. Forecasting multiple time series with one-sided dynamic principal components. Journal of the American Statistical Association. DOI: 10.1080/01621459.2018.1520117.

Xie, M., Sandels, C., Zhu, K. y Nordstrom, L., 2013. A Seasonal ARIMA model with exogenous variables for Elspot electricity prices in Sweden. 2013 10th International Conference on the European Energy Market (EEM). 1-4.

Yaseen, Z.M, , El-shafie, A., Jaafar, O., Afan, H., Sayl, K., 2015. Artificial intelligence based models for stream-flow forecasting: 2000–2015. Journal of Hydrology. 530, 829–844.

Capítulo 5

Conclusiones Generales

5.1. Conclusiones

En esta tesis doctoral se profundizó el estudio de la coherencia regional entre diferentes regímenes hidrológicos (ríos Paraná, Uruguay y Atuel) y el clima. Más aún, también se evaluaron distintos modelos estadísticos de pronóstico para el caudal estacional, mensual y diario en los diferentes ríos.

Un primer estudio exploratorio del caudal fue realizado en el primer capítulo de este trabajo tanto en la LPB como en la cuenca del río Atuel de forma tal de comenzar a comprender los distintos procesos involucrados en los regímenes de caudales de estas cuencas consideradas. Asimismo, en el primer capítulo fue posible corroborar los cambios en los valores medios en el río Paraná a partir de la década del setenta (García y Vargas, 1998; Jaime y Menéndez, 2002; entre otros), así como también la detección de eventos de crecida considerable durante la primera mitad del siglo. Por otro lado, en el río Uruguay los eventos más extremos que se registraron se corresponden con el mes de julio del año 1983 coincidente con las estaciones del río Paraná. En lo que refiere a los resultados obtenidos en valores medios mensuales de caudal en el análisis del río Atuel fue posible detectar un comportamiento totalmente disímil entre la estación Carmensa y el resto de las estaciones contempladas.

Además, a partir de los valores de anomalías de las series de tiempo técnicas espectrales fueron empleadas para representar temporal y espacialmente la variabilidad de las series. En este caso, algunas estaciones no presentaron ninguna onda significativa como es el caso del río Atuel. Lo anterior podría estar relacionado con que la metodología acarrea el

problema de trade-off entre resolución de frecuencia y varianza, dando lugar a que resulte difícil precisar e identificar ondas cortas sin sacrificar la exactitud de la energía estimada. En este sentido, a pesar que el método de Blackman Tukey puede separar el ruido y las componentes caóticas de las periódicas, su resolución no resulta suficiente para estimar cada periodicidad individual (Yiou et al., 1996). Por esta razón, se decide mejorar los resultados considerando la metodología de wavelet. En el caso del río Paraná, si bien el método de Blackman Tukey detectaba la significancia de ondas cercanas a los doce meses, a través del wavelet se pudo observar la presencia de las mismas más que nada en una segunda parte del periodo. En el caso del río Uruguay para la primer metodología se presentaba un máximo absoluto de energía, el cual no resultaba significativo, pudiendo encontrar a través de la segunda metodología la significancia de esas ondas largas en ciertos intervalos de tiempo. En lo que respecta al río Atuel, el análisis espectral no permitía encontrar ondas significativas a diferencia de si encontrarlas con wavelet.

Combinando los estudios anteriores realizados, se entiende que es posible realizar una selección de estaciones de referencia en cada río, en particular en lo que refiere a ondas más largas posiblemente relacionadas con la variabilidad interdecadal y decadal. Por otro lado, el estudio muestra que las estaciones cercanas exhiben la presencia de ondas que se producen de manera prácticamente en fase en el par de estación.

De esta manera, el primer estudio realizado en este trabajo de las series temporales de la red de estaciones de caudal y su extenso registro temporal que algunas de ellas presenta, sugiere la selección de estaciones de referencia para futuros estudios. En el caso del río Paraná, sería conveniente de seleccionar más de una estación de referencia, debido a sus características fisiográficas (Corrientes y Túnel Subfluvial). En este sentido, este estudio extiende la literatura actual, puesto que la mayoría de los análisis han estudiado el alto Paraná. Además, las ciudades de mayor concentración de población se encuentran en la parte media e inferior de dicho río. En el río Uruguay, Paso de los Libres podría ser representativa del río Uruguay en el tramo dentro del territorio Argentino; y en el caso del río Atuel la estación La Angostura, debido a su registro temporal y la no influencia de los embalses siguientes como sucedería con Carmensa.

Así, la selección de estación de referencias podría resultar útil para monitorear en menor cantidad de puntos, y de esta manera reducir el número de variables para la representación hidrológica tanto para estudios temporales extensos como en grandes escalas espaciales.

De esta manera, regiones particulares de las cuencas fueron estudiadas en detalle, para poder facilitar al uso de los recursos hídricos en las regiones mencionadas.
Asimismo, la selección de las estaciones de referencia permitió de alguna manera evaluar el comportamiento conjunto, es decir, se tuvo en cuenta el comportamiento y la relación que existe entre distintas cuencas lo que permitiría colaborar con la realización de medidas preventivas, en el caso de que fuese necesario realizarlas, para subsanar eventos extremos tales como inundaciones y sequías. De esta manera, estudiar la variabilidad espacial y temporal de los recursos hídricos de un país permite un mejor aprovechamiento de los mismos.
Mediante distintas metodologías fue posible realizar un estudio de la covariabilidad espacial y temporal de la red de estaciones en cuencas de interés. Asimismo, evaluar diversas escalas temporales permitió analizar la presencia de ciclos característicos que ocurren en pares de estaciones, y con qué desfasaje se hace presente entre las mismas.
De acuerdo al análisis considerado, se notó por un lado que la relación entre las estaciones de la Cuenca del Plata resultó ser más estrecha que con la de la cuenca del Atuel. Sin embargo, en prácticamente todas las escalas temporales, en mayor o en menor medida, fue posible observar la presencia de ondas largas entre aproximadamente 2 y 5 años entre las estaciones de la Cuenca del Plata y la del Atuel consideradas. Dada su longitud, dichos ciclos podrían estar asociados a variabilidades climáticas conocidas.
Asimismo, resulta importante destacar que para obtener un estudio más preciso de la variabilidad climática y su relación con el caudal, el análisis de datos estacionales podría resultar útil. Pues, el mismo permite suavizar las fluctuaciones de alta frecuencia en el caudal, sin necesidad de ir a una escala temporal mayor. Además, como fue observado, dicho filtro no afecta las ondas de baja frecuencia que se quieren profundizar, no resultando necesario estudiar las anomalías de caudal.
Dado que parte del trabajo involucraba profundizar el estudio del caudal de cuencas de la Argentina y su relación con el clima, se realizó un abordaje donde se propuso un método para estimar la función de densidad de probabilidad conjunta combinando un desfasaje para ciertos índices climáticos y el caudal estacional. En este caso, se tuvo como principal objetivo encontrar una herramienta útil para anticipar eventos extremos en el caudal, puesto que se está estudiando la interacción entre el índice observado y el caudal que ocurre entre 1, 2 o 3 trimestres después dependiendo del modelo establecido.

Estimar la distribución conjunta de un par de variables respuesta/explicatoria puede ser útil en varios aspectos. Asimismo, se han derivado métricas que resultan particularmente útiles para cuantificar algunas relaciones conocidas, como así también para proponer simples ideas de pronóstico. Sin embargo, la ventana de posibles aplicaciones puede ser ampliada, ya que la distribución conjunta de un par de variables contiene toda la información probabilística en ella. Por ejemplo, la información sobre la distribución conjunta puede ser utilizada como un input para modelos más complejos. En nuestro caso, esta herramienta habilitó a cuantificar probabilidades (y sus periodos de retorno asociados) para ciertos eventos extremos conjuntos que ocurrieron en los últimos años. Una métrica como la utilizada puede ser útil para monitorear por ejemplo, la ocurrencia de eventos extremos en la variable caudal y el índice climático relacionado con intensos eventos ENSO (Takahashi y Dewitte, 2016). Más aún, se han realizado dos posibles enfoques para el pronóstico estacional de caudal. El primero, que sigue la línea de trabajo de Khendun et al. (2014) está relacionado con estimar el valor esperado de caudal dado el valor observado del índice, y utilizarlo como un pronóstico. El segundo es computando intervalos de confianza para el caudal, dado los valores del índice climático. En el primer caso, se ha encontrado que hay señal en la predicción al estimar una correlación significativa entre los valores observados y los valores pronosticados. En el segundo caso, el pronóstico de los cuantiles es especialmente útil ya que provee una manera simple de obtener un valor que el caudal no superará con cierto intervalo de confianza (en el caso que se este mirando la cola derecha de la distribución de datos del caudal). En este sentido, el cuantil 95 % de la distribución condicional teórica parece ser una métrica de monitoreo razonable, ya que de acuerdo a los resultados obtenidos es un umbral que el caudal respeta con buena confianza, siendo al mismo tiempo suficientemente ajustado como para no levantar alarmas en todos los periodos. Esto resulta evidente en el estudio entre el índice NIÑO 3.4 y el caudal en la estación Túnel Subfluvial, como así también entre el índice SOI y las tres estaciones consideradas (teniendo en cuenta que en la estación La Angostura el umbral de estudio era el cuantil 5 %). Más aún, se adicionó una métrica para evaluar la esperanza acumulada condicional a fin de comparar modelos para Túnel Subfluvial, mostrando que resulta válido tener más de un modelo en consideración.

Adicionalmente, los modelos fueron validados por diferentes enfoques mostrando que pueden ser considerados para monitorear la relación entre el caudal, el ENSO y las estaciones

de medición en la Cuenca del Plata y la del Atuel. Los modelos fueron validados temporalmente en el caso de estudio Túnel Subfluvial y el índice NIÑO 3.4 y SOI, al igual que el índice SOI con las estaciones Paso de los Libres y La Angostura, dando cierta certeza que tanto las métricas monitoreadas como los pronósticos que se realizaron pueden ser significativas para aplicación. Más aún, también se ha revisado la consistencia de los umbrales computados para los valores de caudal.

En este sentido, con el uso de distintas métricas se evaluó el rendimiento de los modelos considerados tanto en la estación Túnel Subfluvial como en la estación Paso de los Libres. Particularmente, se tuvo en cuenta el área debajo de la curva ROC para evaluar la relación entre la variable dicotómica obtenida a partir de la altura del caudal y la variable numérica del percentil 95 de caudal esperado. De acuerdo a los resultados obtenidos se encontró que la variable obtenida a partir del modelo que utiliza el índice NIÑO 3.4 para obtener el caudal esperado del percentil 95 en la estación Túnel Subfluvial presenta los resultados más favorables a la hora de anticipar evacuaciones. Asimismo, se examinaron dos posibles umbrales de aviso en todos los casos de estudio, es decir valores de caudal esperados a partir de los cuales anticipar que se va a emitir un aviso de evacuación. Para ambos umbrales se realizaron las respectivas tablas de contingencia y mediante el test chi-cuadrado se evaluó la hipótesis nula de la falta de asociación entre ambas variables. En todo los casos la hipótesis fue rechazada.

También, es importante mencionar que el procedimiento muestra que dado un cierto umbral para el índice climático es posible cuantificar la probabilidad de eventos extremos de caudal con una cierta anticipación. En este sentido, fue posible obtener una simple y válida manera probabilística climática de monitorear la ocurrencia de eventos extremos en tres ríos de suma importancia socio-económica en tanto como en el sur de la región de la Cuenca del Plata, como en la región de Cuyo más precisamente en el río Atuel de variables bien estudiadas y conocidas. En este sentido, por ejemplo, en el caso de Túnel Subfluvial, la metodología prevé la ocurrencia de eventos superiores al percentil 95 en ambas variables en promedio cada 11.5 años. Esto indica que uno espera aproximadamente que la mitad de los eventos de caudal se corresponda con eventos extremos del índice. Una identificación temprana de este tipo de eventos podría ayudar a mitigar su impacto negativo.

Finalmente un distinto tipo de análisis fue llevado a cabo relacionado con la potencialidad

de predicción del caudal mensual y diario en la LPB. Puesto, que el continuo análisis hidrológico en la Cuenca del Plata es necesario dado el potencial riesgo de valores extremos de caudal (inundaciones y bajantes) que provocan daños irreparables en regiones dentro de la cuenca. En este sentido, el monitoreo del caudal mensual es fundamental. Para lo cual, el modelado estadístico resultaría una herramienta útil para mejorar el sistema de alerta y control en el futuro cercano. De esta manera, en este trabajo se empleó el modelado estadístico de acuerdo a los modelos SARIMA, propuesto por Box y Jenkins, para ciertas estaciones localizadas en la parte baja de la Cuenca del Plata.

A través de herramientas de identificación, estimación y diagnóstico se propuso modelar las series temporales de 41 años de las estaciones Túnel Subfluvial, Corrientes y Paso de los Libres. Para ello, cien posibles combinaciones de hiperparámetro del modelo SARIMA fueron contempladas incluyendo la onda anual, y se seleccionó mediante la métrica AIC el mejor modelo para cada serie de tiempo.

Asimismo, se estimó el valor de NSE para los dos periodos de entrenamiento considerados en este capítulo que indicaron que el modelo no presentó un bajo rendimiento. Sin embargo, en la realización de los pronósticos a seis y doce meses mediante el método sin variable exógena se encontraron ciertas particularidades en lo que refiere a la sobreestimación y la subestimación del caudal mensual.

Además, como caso particular de estudio se realizó un análisis teniendo en cuenta como variable regresora el caudal medio estimado de manera condicional del Capítulo 3 de la estación Túnel Subfluvial. En este sentido, mediante el uso de validación cruzada se realizó una comparación entre modelos. Se encontró que en promedio la variable regresora no mejora el rendimiento de los modelos. Sin embargo se pudo observar ciertas mejoras en lo que refiere al pronóstico en los sectores de la serie que presentan valores extremos. Este último resultado es coherente con el desarrollo de la presente tesis, puesto que en el capítulo anterior encontramos que la relación entre el índice NIÑO 3.4 y el caudal es más fuerte cuando nos encontramos bajo la influencia del fenómeno. Más aún, la distribución conjunta encontrada funcionaba particularmente bien en la cola correspondiente a valores altos del índice y del caudal, sugiriendo que la señal que genera el índice puede no ser tan fuerte para valores medios.

Si bien en el estudio para Túnel Subfluvial se propuso seleccionar una combinación de hiperparámetros óptima para el modelo SARIMA, y luego se incorporó al mismo una

variable regresora (caudal esperado), en este sentido observemos que también habría sido posible realizar una nueva selección de hiperparámetros utilizando la variable exógena desde un principio. Sin embargo, no se espera encontrar grandes cambios puesto que la relación entre la variable exógena y el caudal sólo es fuerte en ventanas de tiempo acotadas y relacionadas con eventos extremos de la fase Niño.

En base a los resultados obtenidos, los modelos estudiados debido a su rápida implementación podrían contribuir al monitoreo de los caudales. Sin embargo, resulta necesario profundizar en otras metodologías y en una escala temporal que permita considerar una variabilidad de mayor frecuencia. Es decir, se entiende que dicha consideración permitiría contemplar los eventos de crecida o bajantes. Por lo tanto, se propone en el próximo anexo profundizar en el estudio de las series temporales diarias de caudal con el fin de obtener un pronóstico más adecuado. Por lo tanto, se propuso como último análisis la exploración de datos diarios de caudal de forma tal de intentar de reducir la variabilidad del caudal que se había destacado en el análisis previo utilizando un método multivariado ODPC. En este sentido, se destacó la implicancia positiva del método para obtener una mejora en el pronóstico a un día, no encontrando resultados del todo favorables a 3 y 5 días. Por otra parte, se propone continuar explorando la potencialidad de la metodología multivariada para distintas aplicaciones en series temporales.

5.2. Referencias

Se encuentran consignadas en los capítulos anteriores.

Apéndice B

Tablas Capítulo 4

B.1. Túnel Subfluvial

Tabla I: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del modelo SARIMA del caudal medio mensual en el periodo 1975-2015 para la estación Túnel Subfluvial. En negrita se destaca el modelo seleccionado* $SARIMA(3,0,0)(0,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	9443.58	-0.46
2	1	0	0	0	9027.56	0.39
3	2	0	0	0	9006.55	0.42
4	3	0	0	0	8983.57	0.45
5	4	0	0	0	8985.41	0.45
6	0	1	0	0	9106.43	0.28
7	1	1	0	0	8992.40	0.43
8	2	1	0	0	8993.10	0.44
9	3	1	0	0	8985.46	0.45
10	4	1	0	0	8987.30	0.45
11	0	2	0	0	9026.42	0.39
12	1	2	0	0	8990.72	0.44
13	2	2	0	0	8991.41	0.44
14	3	2	0	0	8986.40	0.45
15	4	2	0	0	8964.34	0.48
16	0	3	0	0	9010.57	0.42
17	1	3	0	0	8989.87	0.44
18	2	3	0	0	8991.34	0.44
19	3	3	0	0	8970.56	0.47
20	4	3	0	0	8933.68	0.52
21	0	4	0	0	9007.89	0.42

22	1	4	0	0	8989.74	0.45
23	2	4	0	0	8993.87	0.44
24	3	4	0	0	8943.27	0.51
25	4	4	0	0	8953.44	0.50
26	0	0	1	0	9349.17	-0.19
27	1	0	1	0	8893.87	0.54
28	2	0	1	0	8872.20	0.57
29	3	0	1	0	8858.01	0.58
30	4	0	1	0	8859.30	0.58
31	0	1	1	0	9002.03	0.43
32	1	1	1	0	8862.45	0.57
33	2	1	1	0	8863.53	0.57
34	3	1	1	0	8859.39	0.58
35	4	1	1	0	8862.01	0.58
36	0	2	1	0	8911.20	0.53
37	1	2	1	0	8862.44	0.58
38	2	2	1	0	8862.67	0.58
39	3	2	1	0	8857.38	0.59
40	4	2	1	0	8852.45	0.59
41	0	3	1	0	8884.03	0.56
42	1	3	1	0	8861.54	0.58
43	2	3	1	0	8863.53	0.58
44	3	3	1	0	8856.79	0.59
45	4	3	1	0	8852.98	0.59
46	0	4	1	0	8881.42	0.56
47	1	4	1	0	8863.51	0.58
48	2	4	1	0	8855.90	0.59
49	3	4	1	0	8837.07	0.61
50	4	4	1	0	8854.17	0.59
51	0	0	0	1	9242.71	0.07
52	1	0	0	1	8763.79	0.67
53	2	0	0	1	8738.63	0.68
54	**3**	**0**	**0**	**1**	**8721.42**	**0.70**
55	4	0	0	1	8722.89	0.70
56	0	1	0	1	8887.32	0.56
57	1	1	0	1	8726.84	0.69
58	2	1	0	1	8728.22	0.69
59	3	1	0	1	8722.95	0.70
60	4	1	0	1	8724.88	0.70
61	0	2	0	1	8781.88	0.65
62	1	2	0	1	8727.36	0.69
63	2	2	0	1	8725.37	0.70
64	3	2	0	1	8724.75	0.70
65	4	2	0	1	8726.57	0.70
66	0	3	0	1	8752.54	0.68
67	1	3	0	1	8724.19	0.70
68	2	3	0	1	8725.95	0.70
69	3	3	0	1	8724.38	0.70
70	4	3	0	1	8727.56	0.70

71	0	4	0	1	8748.44	0.68
72	1	4	0	1	8725.75	0.70
73	2	4	0	1	8727.21	0.70
74	3	4	0	1	8724.81	0.70
75	4	4	0	1	8726.17	0.70
76	0	0	1	1	9239.99	0.08
77	1	0	1	1	8765.47	0.67
78	2	0	1	1	8740.46	0.69
79	3	0	1	1	8722.58	0.70
80	4	0	1	1	8724.12	0.70
81	0	1	1	1	8884.60	0.57
82	1	1	1	1	8728.39	0.70
83	2	1	1	1	8729.70	0.70
84	3	1	1	1	8724.19	0.70
85	4	1	1	1	8726.10	0.70
86	0	2	1	1	8781.89	0.66
87	1	2	1	1	8728.71	0.70
88	2	2	1	1	8727.07	0.70
89	3	2	1	1	8725.93	0.70
90	4	2	1	1	8728.18	0.70
91	0	3	1	1	8753.90	0.68
92	1	3	1	1	8725.61	0.70
93	2	3	1	1	8727.48	0.70
94	3	3	1	1	8727.07	0.70
95	4	3	1	1	8729.43	0.70
96	0	4	1	1	8749.85	0.68
97	1	4	1	1	8727.36	0.70
98	2	4	1	1	8728.80	0.70
99	3	4	1	1	8725.26	0.71
100	4	4	1	1	—	—

Tabla II: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del modelo SARIMA del caudal medio mensual en el periodo 1975-junio/2016 para la estación Túnel Subfluvial. En negrita se destaca el modelo seleccionado* $SARIMA(3,0,0)(0,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	9566.81	-0.45
2	1	0	0	0	9140.79	0.40
3	2	0	0	0	9119.92	0.43
4	3	0	0	0	9097.10	0.46
5	4	0	0	0	9098.99	0.46
6	0	1	0	0	9223.11	0.29
7	1	1	0	0	9105.82	0.44

8	2	1	0	0	9106.29	0.45
9	3	1	0	0	9099.03	0.46
10	4	1	0	0	9100.98	0.46
11	0	2	0	0	9142.31	0.40
12	1	2	0	0	9103.78	0.45
13	2	2	0	0	9104.62	0.45
14	3	2	0	0	9099.94	0.46
15	4	2	0	0	9086.99	0.47
16	0	3	0	0	9124.72	0.42
17	1	3	0	0	9103.32	0.45
18	2	3	0	0	9104.89	0.45
19	3	3	0	0	9083.83	0.48
20	4	3	0	0	9048.83	0.52
21	0	4	0	0	9122.24	0.43
22	1	4	0	0	9103.40	0.45
23	2	4	0	0	9086.37	0.48
24	3	4	0	0	9057.75	0.51
25	4	4	0	0	9059.13	0.51
26	0	0	1	0	9475.45	-0.19
27	1	0	1	0	9008.42	0.55
28	2	0	1	0	8986.42	0.57
29	3	0	1	0	8972.34	0.58
30	4	0	1	0	8973.84	0.58
31	0	1	1	0	9120.14	0.43
32	1	1	1	0	8976.52	0.58
33	2	1	1	0	8977.41	0.58
34	3	1	1	0	8973.90	0.58
35	4	1	1	0	8965.43	0.60
36	0	2	1	0	9028.39	0.53
37	1	2	1	0	8976.28	0.58
38	2	2	1	0	8976.74	0.58
39	3	2	1	0	8970.84	0.59
40	4	2	1	0	8966.38	0.60
41	0	3	1	0	8998.82	0.56
42	1	3	1	0	8975.88	0.58
43	2	3	1	0	8977.87	0.58
44	3	3	1	0	8970.35	0.59
45	4	3	1	0	8967.01	0.60
46	0	4	1	0	8996.29	0.56
47	1	4	1	0	8977.82	0.58
48	2	4	1	0	8970.60	0.59
49	3	4	1	0	8953.63	0.61
50	4	4	1	0	8951.83	0.62
51	0	0	0	1	9372.18	0.06
52	1	0	0	1	8882.12	0.66
53	2	0	0	1	8857.49	0.68
54	**3**	**0**	**0**	**1**	**8840.09**	**0.70**
55	4	0	0	1	8841.73	0.70
56	0	1	0	1	9007.31	0.56

57	1	1	0	1	8845.34	0.69
58	2	1	0	1	8846.29	0.69
59	3	1	0	1	8841.75	0.70
60	4	1	0	1	8843.75	0.70
61	0	2	0	1	8903.33	0.65
62	1	2	0	1	8845.07	0.69
63	2	2	0	1	8843.65	0.70
64	3	2	0	1	8843.69	0.70
65	4	2	0	1	8845.47	0.70
66	0	3	0	1	8871.28	0.67
67	1	3	0	1	8843.05	0.70
68	2	3	0	1	8844.79	0.70
69	3	3	0	1	8843.20	0.70
70	4	3	0	1	8845.44	0.70
71	0	4	0	1	8867.39	0.68
72	1	4	0	1	8844.51	0.70
73	2	4	0	1	8845.87	0.70
74	3	4	0	1	8842.40	0.70
75	4	4	0	1	8843.97	0.70
76	0	0	1	1	9366.34	0.09
77	1	0	1	1	8883.84	0.66
78	2	0	1	1	8859.36	0.68
79	3	0	1	1	8841.32	0.70
80	4	0	1	1	8843.03	0.70
81	0	1	1	1	9002.69	0.58
82	1	1	1	1	8846.95	0.69
83	2	1	1	1	8847.80	0.69
84	3	1	1	1	8843.05	0.70
85	4	1	1	1	8845.03	0.70
86	0	2	1	1	8902.92	0.66
87	1	2	1	1	8846.44	0.69
88	2	2	1	1	8845.35	0.70
89	3	2	1	1	8844.95	0.70
90	4	2	1	1	8846.76	0.70
91	0	3	1	1	8872.58	0.68
92	1	3	1	1	8844.50	0.70
93	2	3	1	1	8846.36	0.70
94	3	3	1	1	8849.32	0.70
95	4	3	1	1	8847.92	0.70
96	0	4	1	1	8868.81	0.68
97	1	4	1	1	8846.17	0.70
98	2	4	1	1	8847.51	0.70
99	3	4	1	1	8842.96	0.70
100	4	4	1	1	—	—

B.2. Corrientes

Tabla III: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del modelo SARIMA del caudal medio mensual en el periodo 1975-2015 para la estación Corrientes. En negrita se destaca el modelo seleccionado* $SARIMA(3,0,3)(0,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	9959.87	-0.66
2	1	0	0	0	9603.14	0.22
3	2	0	0	0	9600.34	0.22
4	3	0	0	0	9580.22	0.26
5	4	0	0	0	9579.83	0.26
6	0	1	0	0	9680.59	0.08
7	1	1	0	0	9595.80	0.23
8	2	1	0	0	9576.47	0.27
9	3	1	0	0	9580.63	0.26
10	4	1	0	0	9581.83	0.26
11	0	2	0	0	9630.87	0.17
12	1	2	0	0	9584.27	0.25
13	2	2	0	0	9586.25	0.25
14	3	2	0	0	9551.17	0.31
15	4	2	0	0	9566.43	0.29
16	0	3	0	0	9613.68	0.21
17	1	3	0	0	9586.22	0.25
18	2	3	0	0	9570.67	0.28
19	3	3	0	0	9552.75	0.31
20	4	3	0	0	9531.77	0.35
21	0	4	0	0	9612.66	0.21
22	1	4	0	0	9581.13	0.26
23	2	4	0	0	9541.21	0.33
24	3	4	0	0	9541.87	0.33
25	4	4	0	0	9533.09	0.35
26	0	0	1	0	9860.01	-0.33
27	1	0	1	0	9448.27	0.44
28	2	0	1	0	9447.18	0.44
29	3	0	1	0	9431.03	0.46
30	4	0	1	0	9431.18	0.47
31	0	1	1	0	9565.57	0.28
32	1	1	1	0	9444.65	0.45
33	2	1	1	0	9441.99	0.45
34	3	1	1	0	9431.50	0.47
35	4	1	1	0	9433.16	0.47
36	0	2	1	0	9497.77	0.38
37	1	2	1	0	9433.11	0.46
38	2	2	1	0	9435.01	0.46

39	3	2	1	0	9429.54	0.47
40	4	2	1	0	9426.29	0.48
41	0	3	1	0	9475.58	0.41
42	1	3	1	0	9434.92	0.46
43	2	3	1	0	9436.54	0.46
44	3	3	1	0	9427.30	0.47
45	4	3	1	0	9429.29	0.47
46	0	4	1	0	9467.82	0.42
47	1	4	1	0	9433.89	0.46
48	2	4	1	0	9427.04	0.48
49	3	4	1	0	9428.40	0.48
50	4	4	1	0	9434.74	0.47
51	0	0	0	1	9736.60	0.03
52	1	0	0	1	9319.43	0.60
53	2	0	0	1	9315.75	0.61
54	3	0	0	1	9299.09	0.63
55	4	0	0	1	9298.53	0.63
56	0	1	0	1	9433.75	0.50
57	1	1	0	1	9311.51	0.61
58	2	1	0	1	9309.46	0.62
59	3	1	0	1	9298.74	0.63
60	4	1	0	1	9300.39	0.63
61	0	2	0	1	9361.65	0.57
62	1	2	0	1	9301.04	0.62
63	2	2	0	1	9302.07	0.62
64	3	2	0	1	9298.03	0.62
65	4	2	0	1	9301.89	0.63
66	0	3	0	1	9338.57	0.59
67	1	3	0	1	9301.60	0.62
68	2	3	0	1	9303.44	0.62
69	**3**	**3**	**0**	**1**	**9294.39**	**0.63**
70	4	3	0	1	9300.22	0.62
71	0	4	0	1	9333.38	0.60
72	1	4	0	1	9302.94	0.63
73	2	4	0	1	9303.08	0.63
74	3	4	0	1	9299.38	0.63
75	4	4	0	1	9301.38	0.63
76	0	0	1	1	9730.29	0.07
77	1	0	1	1	9321.42	0.60
78	2	0	1	1	9317.74	0.60
79	3	0	1	1	9301.07	0.63
80	4	0	1	1	9300.52	0.63
81	0	1	1	1	9432.54	0.50
82	1	1	1	1	9313.51	0.61
83	2	1	1	1	9311.45	0.62
84	3	1	1	1	9300.73	0.63
85	4	1	1	1	9302.39	0.63
86	0	2	1	1	9363.14	0.57
87	1	2	1	1	9303.04	0.62

88	2	2	1	1	9304.06	0.62
89	3	2	1	1	9299.91	0.63
90	4	2	1	1	—	—
91	0	3	1	1	9340.38	0.59
92	1	3	1	1	9303.58	0.63
93	2	3	1	1	9305.43	0.63
94	3	3	1	1	9296.31	0.63
95	4	3	1	1	9318.01	0.62
96	0	4	1	1	9335.34	0.60
97	1	4	1	1	9304.94	0.63
98	2	4	1	1	9304.83	0.63
99	3	4	1	1	9301.38	0.63
100	4	4	1	1	9304.41	0.63

Tabla IV: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del modelo SARIMA del caudal medio mensual en el periodo 1975-junio/2016 para la estación Corrientes. En negrita se destaca el modelo seleccionado* $SARIMA(3,0,3)(0,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	10090.27	-0.64
2	1	0	0	0	9719.65	0.24
3	2	0	0	0	9716.89	0.25
4	3	0	0	0	9696.74	0.28
5	4	0	0	0	9696.80	0.28
6	0	1	0	0	9802.47	0.10
7	1	1	0	0	9712.42	0.25
8	2	1	0	0	9693.55	0.29
9	3	1	0	0	9697.43	0.28
10	4	1	0	0	9698.80	0.28
11	0	2	0	0	9750.16	0.19
12	1	2	0	0	9700.67	0.27
13	2	2	0	0	9702.65	0.27
14	3	2	0	0	9668.84	0.33
15	4	2	0	0	9670.46	0.33
16	0	3	0	0	9731.71	0.23
17	1	3	0	0	9702.62	0.27
18	2	3	0	0	9687.72	0.30
19	3	3	0	0	9666.34	0.34
20	4	3	0	0	9649.72	0.36
21	0	4	0	0	9729.92	0.23
22	1	4	0	0	9697.84	0.29
23	2	4	0	0	9658.77	0.35
24	3	4	0	0	9659.13	0.35

25	4	4	0	0	9651.15	0.36
26	0	0	1	0	9992.85	-0.33
27	1	0	1	0	9565.16	0.45
28	2	0	1	0	9563.87	0.46
29	3	0	1	0	9548.43	0.48
30	4	0	1	0	9549.20	0.48
31	0	1	1	0	9688.34	0.29
32	1	1	1	0	9561.31	0.46
33	2	1	1	0	9558.44	0.47
34	3	1	1	0	9549.37	0.48
35	4	1	1	0	9551.18	0.48
36	0	2	1	0	9617.79	0.39
37	1	2	1	0	9550.31	0.47
38	2	2	1	0	9552.22	0.47
39	3	2	1	0	9548.44	0.48
40	4	2	1	0	9545.12	0.49
41	0	3	1	0	9593.79	0.42
42	1	3	1	0	9552.16	0.47
43	2	3	1	0	9553.83	0.47
44	3	3	1	0	9545.86	0.49
45	4	3	1	0	9547.98	0.49
46	0	4	1	0	9584.73	0.44
47	1	4	1	0	9551.69	0.48
48	2	4	1	0	9545.75	0.49
49	3	4	1	0	9551.65	0.48
50	4	4	1	0	9536.16	0.51
51	0	0	0	1	9867.12	0.06
52	1	0	0	1	9437.97	0.60
53	2	0	0	1	9435.02	0.61
54	3	0	0	1	9418.16	0.62
55	4	0	0	1	9418.44	0.62
56	0	1	0	1	9556.30	0.51
57	1	1	0	1	9431.17	0.61
58	2	1	0	1	9428.21	0.61
59	3	1	0	1	9418.35	0.62
60	4	1	0	1	9420.24	0.62
61	0	2	0	1	9484.73	0.58
62	1	2	0	1	9419.37	0.62
63	2	2	0	1	9420.74	0.62
64	3	2	0	1	9418.97	0.62
65	4	2	0	1	9422.15	0.62
66	0	3	0	1	9459.85	0.59
67	1	3	0	1	9420.48	0.62
68	2	3	0	1	9422.34	0.62
69	**3**	**3**	**0**	**1**	**9414.76**	**0.63**
70	4	3	0	1	9421.82	0.62
71	0	4	0	1	9452.97	0.60
72	1	4	0	1	9421.88	0.62
73	2	4	0	1	9422.26	0.62

74	3	4	0	1	9418.44	0.63
75	4	4	0	1	9420.44	0.63
76	0	0	1	1	9861.29	0.07
77	1	0	1	1	9439.96	0.60
78	2	0	1	1	9437.01	0.60
79	3	0	1	1	9420.10	0.63
80	4	0	1	1	9420.39	0.63
81	0	1	1	1	9555.09	0.51
82	1	1	1	1	9433.16	0.61
83	2	1	1	1	9430.21	0.61
84	3	1	1	1	9420.32	0.63
85	4	1	1	1	9422.20	0.63
86	0	2	1	1	9485.93	0.58
87	1	2	1	1	9421.36	0.62
88	2	2	1	1	9422.74	0.62
89	3	2	1	1	9420.86	0.62
90	4	2	1	1	9424.30	0.63
91	0	3	1	1	9461.53	0.60
92	1	3	1	1	9422.48	0.62
93	2	3	1	1	9424.35	0.62
94	3	3	1	1	9416.75	0.63
95	4	3	1	1	9435.52	0.62
96	0	4	1	1	9454.89	0.60
97	1	4	1	1	9423.86	0.62
98	2	4	1	1	9423.93	0.63
99	3	4	1	1	9420.40	0.64
100	4	4	1	1	—	—

B.3. Paso de los Libres

Tabla V: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del modelo SARIMA del caudal medio mensual en el periodo 1975-2015 para la estación Paso de los Libres. En negrita se destaca el modelo seleccionado* $SARIMA(3,0,2)(1,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	9462.49	-0.63
2	1	0	0	0	9289.65	-0.13
3	2	0	0	0	9279.71	-0.10
4	3	0	0	0	9279.43	-0.10
5	4	0	0	0	9279.72	-0.09
6	0	1	0	0	9348.63	-0.28

7	1	1	0	0	9277.26	-0.09
8	2	1	0	0	9279.23	-0.09
9	3	1	0	0	9280.88	-0.09
10	4	1	0	0	9281.60	-0.09
11	0	2	0	0	9310.70	-0.17
12	1	2	0	0	9279.22	-0.09
13	2	2	0	0	9281.21	-0.09
14	3	2	0	0	9283.20	-0.09
15	4	2	0	0	9256.73	-0.03
16	0	3	0	0	9302.47	-0.15
17	1	3	0	0	9280.38	-0.09
18	2	3	0	0	9283.21	-0.09
19	3	3	0	0	9259.43	-0.02
20	4	3	0	0	9254.45	-0.01
21	0	4	0	0	9294.38	-0.13
22	1	4	0	0	9277.03	-0.08
23	2	4	0	0	9239.17	0.01
24	3	4	0	0	9240.84	0.01
25	4	4	0	0	9259.61	-0.03
26	0	0	1	0	9339.81	-0.25
27	1	0	1	0	9176.18	0.12
28	2	0	1	0	9164.14	0.14
29	3	0	1	0	9163.34	0.15
30	4	0	1	0	9162.63	0.15
31	0	1	1	0	9233.05	0.01
32	1	1	1	0	9160.64	0.15
33	2	1	1	0	9162.56	0.15
34	3	1	1	0	9164.31	0.15
35	4	1	1	0	9164.42	0.15
36	0	2	1	0	9196.08	0.08
37	1	2	1	0	9162.55	0.15
38	2	2	1	0	9164.64	0.15
39	3	2	1	0	9166.44	0.15
40	4	2	1	0	9162.15	0.16
41	0	3	1	0	9188.15	0.10
42	1	3	1	0	9164.19	0.15
43	2	3	1	0	9166.31	0.15
44	3	3	1	0	9158.01	0.18
45	4	3	1	0	—	—
46	0	4	1	0	9176.48	0.13
47	1	4	1	0	9161.81	0.16
48	2	4	1	0	9154.47	0.18
49	3	4	1	0	9150.65	0.19
50	4	4	1	0	9144.43	0.21
51	0	0	0	1	9206.90	0.11
52	1	0	0	1	9062.04	0.33
53	2	0	0	1	9056.69	0.34
54	3	0	0	1	9054.23	0.35
55	4	0	0	1	9054.96	0.35

56	0	1	0	1	9104.43	0.28
57	1	1	0	1	9053.41	0.35
58	2	1	0	1	9054.61	0.35
59	3	1	0	1	9055.54	0.35
60	4	1	0	1	9056.19	0.35
61	0	2	0	1	9081.23	0.31
62	1	2	0	1	9054.33	0.35
63	2	2	0	1	9056.12	0.35
64	3	2	0	1	9053.05	0.36
65	4	2	0	1	9058.15	0.35
66	0	3	0	1	9072.78	0.32
67	1	3	0	1	9055.59	0.35
68	2	3	0	1	9057.95	0.35
69	3	3	0	1	9058.19	0.35
70	4	3	0	1	9056.41	0.36
71	0	4	0	1	9063.02	0.34
72	1	4	0	1	9055.32	0.35
73	2	4	0	1	9057.42	0.35
74	3	4	0	1	9057.73	0.35
75	4	4	0	1	9058.48	0.36
76	0	0	1	1	9207.70	0.12
77	1	0	1	1	9060.43	0.34
78	2	0	1	1	9054.60	0.35
79	3	0	1	1	9052.35	0.36
80	4	0	1	1	9053.15	0.36
81	0	1	1	1	9104.38	0.28
82	1	1	1	1	9051.40	0.36
83	2	1	1	1	9052.82	0.36
84	3	1	1	1	9053.76	0.36
85	4	1	1	1	9054.68	0.36
86	0	2	1	1	9080.10	0.32
87	1	2	1	1	9052.60	0.36
88	2	2	1	1	9054.39	0.36
89	**3**	**2**	**1**	**1**	**9051.20**	**0.37**
90	4	2	1	1	—	—
91	0	3	1	1	9071.41	0.33
92	1	3	1	1	9053.66	0.36
93	2	3	1	1	9052.00	0.36
94	3	3	1	1	9056.04	0.36
95	4	3	1	1	9057.93	0.36
96	0	4	1	1	9061.65	0.35
97	1	4	1	1	9052.94	0.36
98	2	4	1	1	9053.58	0.37
99	3	4	1	1	9054.33	0.37
100	4	4	1	1	9055.94	0.37

Tabla VI: *AIC, NSE para las 100 posibles combinaciones de los hiperparámetros del mo-*

delo SARIMA del caudal medio mensual en el periodo 1975-junio/2016 para la estación Paso de los Libres. En negrita se destaca el modelo seleccionado $SARIMA(3,0,2)(1,1,1)_{12}$

Número de parámetros y métricas						
	p	q	P	Q	AIC	NSE
1	0	0	0	0	9576.40	-0.62
2	1	0	0	0	9405.78	-0.14
3	2	0	0	0	9396.02	-0.11
4	3	0	0	0	9395.45	-0.10
5	4	0	0	0	9395.04	-0.10
6	0	1	0	0	9463.15	-0.28
7	1	1	0	0	9393.17	-0.10
8	2	1	0	0	9395.10	-0.10
9	3	1	0	0	9396.66	-0.10
10	4	1	0	0	9396.93	-0.10
11	0	2	0	0	9425.77	-0.18
12	1	2	0	0	9395.08	-0.10
13	2	2	0	0	9397.09	-0.10
14	3	2	0	0	9399.07	-0.10
15	4	2	0	0	9370.92	-0.03
16	0	3	0	0	9418.59	-0.16
17	1	3	0	0	9396.22	-0.10
18	2	3	0	0	9399.08	-0.10
19	3	3	0	0	9374.11	-0.03
20	4	3	0	0	9367.18	-0.01
21	0	4	0	0	9410.06	-0.13
22	1	4	0	0	9392.13	-0.09
23	2	4	0	0	9353.68	0.00
24	3	4	0	0	9355.50	0.00
25	4	4	0	0	9373.71	-0.03
26	0	0	1	0	9452.41	-0.24
27	1	0	1	0	9288.46	0.12
28	2	0	1	0	9276.58	0.14
29	3	0	1	0	9275.81	0.15
30	4	0	1	0	9274.56	0.15
31	0	1	1	0	9344.88	0.01
32	1	1	1	0	9272.98	0.15
33	2	1	1	0	9274.88	0.15
34	3	1	1	0	9276.61	0.15
35	4	1	1	0	9276.35	0.15
36	0	2	1	0	9307.51	0.08
37	1	2	1	0	9274.86	0.15
38	2	2	1	0	9276.98	0.15
39	3	2	1	0	9278.73	0.15
40	4	2	1	0	9273.27	0.16
41	0	3	1	0	9300.48	0.10
42	1	3	1	0	9276.53	0.15

43	2	3	1	0	9278.57	0.15
44	3	3	1	0	9270.87	0.17
45	4	3	1	0	—	—
46	0	4	1	0	9288.48	0.13
47	1	4	1	0	9273.45	0.16
48	2	4	1	0	9266.02	0.18
49	3	4	1	0	9262.85	0.19
50	4	4	1	0	9255.86	0.21
51	0	0	0	1	9320.50	0.10
52	1	0	0	1	9170.43	0.33
53	2	0	0	1	9165.44	0.34
54	3	0	0	1	9162.81	0.35
55	4	0	0	1	9163.49	0.35
56	0	1	0	1	9213.58	0.28
57	1	1	0	1	9162.20	0.35
58	2	1	0	1	9163.30	0.35
59	3	1	0	1	9164.10	0.35
60	4	1	0	1	9164.74	0.35
61	0	2	0	1	9190.15	0.31
62	1	2	0	1	9162.96	0.35
63	2	2	0	1	9164.74	0.35
64	3	2	0	1	9161.16	0.36
65	4	2	0	1	9166.68	0.35
66	0	3	0	1	9181.60	0.32
67	1	3	0	1	9164.18	0.35
68	2	3	0	1	9163.73	0.35
69	3	3	0	1	9167.07	0.35
70	4	3	0	1	9164.58	0.36
71	0	4	0	1	9171.49	0.34
72	1	4	0	1	9163.63	0.35
73	2	4	0	1	9165.92	0.35
74	3	4	0	1	9165.86	0.36
75	4	4	0	1	9166.55	0.36
76	0	0	1	1	9321.04	0.11
77	1	0	1	1	9168.91	0.34
78	2	0	1	1	9163.54	0.35
79	3	0	1	1	9161.14	0.36
80	4	0	1	1	9161.84	0.36
81	0	1	1	1	9213.40	0.29
82	1	1	1	1	9160.39	0.36
83	2	1	1	1	9161.70	0.36
84	3	1	1	1	9162.50	0.36
85	4	1	1	1	9163.39	0.36
86	0	2	1	1	9188.93	0.32
87	1	2	1	1	9161.42	0.36
88	2	2	1	1	9163.19	0.36
89	**3**	**2**	**1**	**1**	**9159.94**	**0.37**
90	4	2	1	1	9163.29	0.36
91	0	3	1	1	9180.32	0.33

92	1	3	1	1	9162.44	0.36
93	2	3	1	1	9160.56	0.36
94	3	3	1	1	9164.17	0.36
95	4	3	1	1	9163.53	0.37
96	0	4	1	1	9170.21	0.35
97	1	4	1	1	9161.39	0.36
98	2	4	1	1	9162.56	0.36
99	3	4	1	1	9166.91	0.36
100	4	4	1	1	9164.82	0.37

En algunos de los casos no fue posible calcular el AIC y el NSE ya que por problemas de convergencia del algoritmo de estimación no se pudieron obtener los estimadores de los parámetros del modelo considerado.