

Tarea 4

Matías Briceño

9 de junio de 2022

Parte 1. Análisis armónico para ciclo estacional

Para este análisis se utilizaron registros diarios de precipitación y caudal de la cuenca del río de Trancura en la región de la Araucanía. Se consideró el periodo 1979-2018. Se realizaron tratamientos previos al análisis en cada uno de los sets de datos que consistieron en la eliminación de días bisiestos para un correcto análisis armónico y el reemplazo de días faltantes por el valor medio de la serie.

A través de un análisis de Fourier se calcularon los tres armónicos principales con los cuales suele representarse ciclos estacionales ($T = 365$ días, $T/2$ y $T/3$). En cada uno de los casos el cálculo se hizo con una regresión lineal múltiple para conseguir las amplitudes respectivas. Además, se calcularon las medias para cada uno de los días de tal manera de encontrar el ciclo estacional medio. Los resultados se pueden ver en la **Figura 1**.

El ciclo estacional de precipitación muestra un pick esperable en el período de invierno que se puede apreciar en las dos técnicas utilizadas. La media muestra gran ruido, que está relacionado con la naturaleza de la distribución de esta variable dado que las precipitaciones cercanas a cero son más probables. El ciclo estacional del caudal presenta dos picks importantes, el primero y con mayor amplitud ocurre durante el mismo período que el pick del ciclo de precipitación que podría implicar una correlación entre ambas variables. El segundo pick podría estar relacionado a otros fenómenos como la temperatura, dado que ocurre en un período en el que típicamente las temperaturas comienzan a subir. Una explicación podría ser el deshielo.

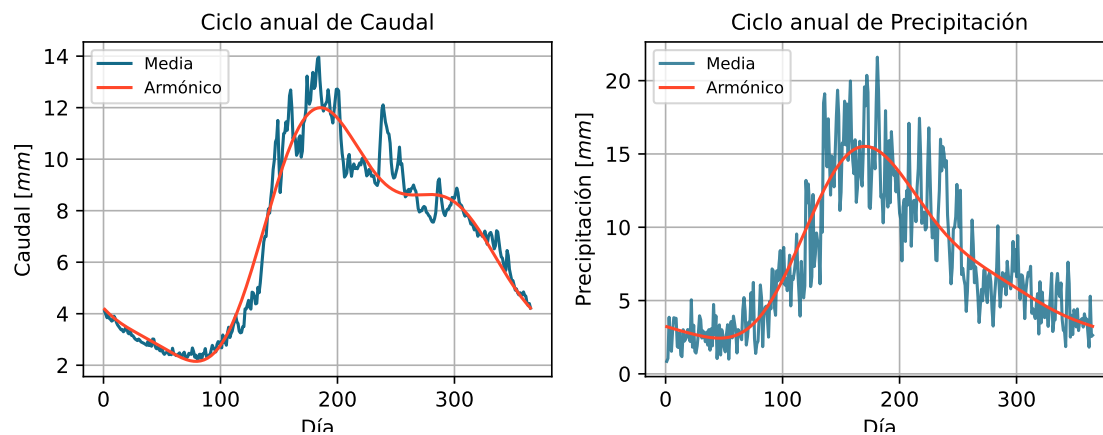


Figura 1: **Ciclos estacionarios de Caudal y Precipitación, Cuenca Trancura.** Los ciclos fueron calculados a partir de 40 años de muestra (1979-2018).

Parte 2. Análisis espectral de anomalías

Para esta sección se hizo un análisis armónico sobre las anomalías históricas de las mismas series trabajadas previamente. Para el cálculo de las anomalías se utilizó la diferencia entre el ciclo estacional dado por los tres armónicos principales y los registros curados. Se utilizó la Transformada de Fourier Rápida (*fft*) dado que el cómputo necesario para utilizar el método de regresión lineal es importante. La **Figura 2** muestra los periodogramas para cada espectro calculado. Se observa que el poder espectral en ambos espectros tienen un patrón similar.

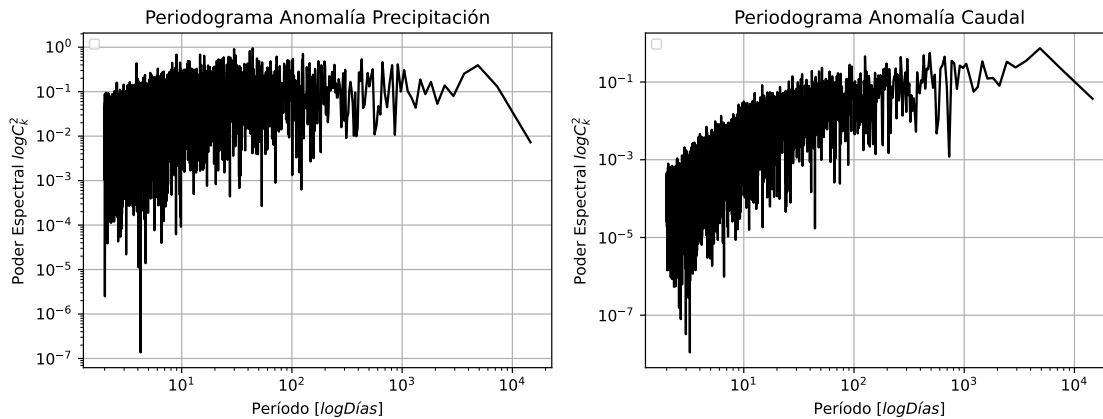


Figura 2: **Periodogramas de anomalías diarias de Caudal y Precipitación.** Se presenta el poder espectral para cada período utilizado.

Parte 3. Modelos lineales para predicción de Anomalía del Caudal

De la misma manera que en la primera parte de este informe, se calcularon los ciclos estacionales para conseguir las series de anomalías de caudal y precipitación, pero esta vez de los registros mensuales con el objetivo de construir un modelo que sea capaz explicar la variación de las anomalías del caudal a partir de las anomalías de precipitación. El primer modelo (*modelo 1*) consistió en una regresión lineal simple entre ambas series. Los resultados de este modelo así como las series anomalías se pueden ver en la **Figura 3**. El R^2 entre la serie de Anomalía del caudal y las predicciones del *modelo 1* fue de 0.48.

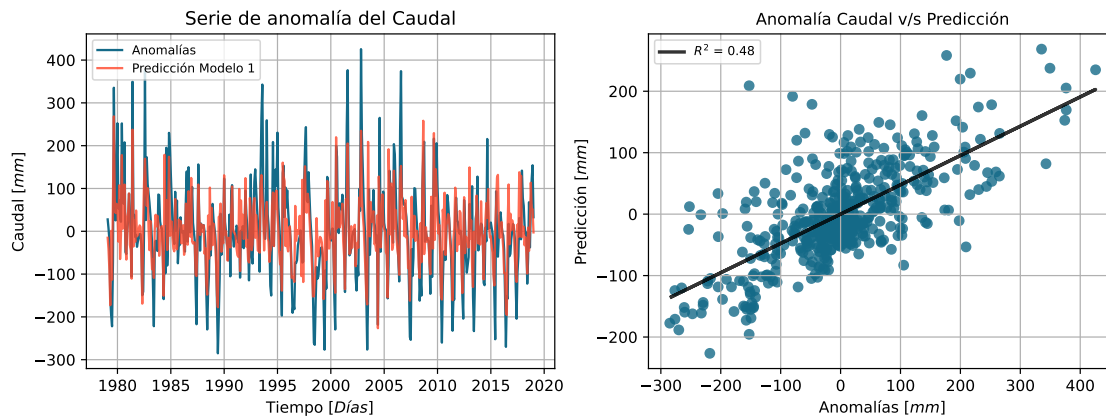


Figura 3: **Modelo 1 para anomalías de caudal, Cuenca Trancura.** La figura izquierda muestra la serie de anomalías del caudal junto con la estimada con el *modelo 1*. La figura derecha muestra un diagrama de dispersión entre ambas series.

Se construyeron otros tres modelos multivariados usando desfases de 1, 2 y 3 meses. El *modelo 2* agrega un desfase de 1 mes, el *modelo 3* agrega desfases de 1 y 2 meses y el *modelo 3* agrega desfases de 1, 2 y 3 meses. Los resultados muestran que al agregar variables al modelo este aumenta en 0.02 el R^2 entre la variables inicial y su predicción por lo que las nuevas variables no están aportando información demasiado relevante para explicar la variabilidad de la anomalía, quizás, por un problema de multicolinealidad (correlación entre las mismas variables de entrada). Una opción sería agregar mas variables para complejizar el modelo, pero esto podría llevar a un problema de sobreajuste que no sería bueno si se quiere utilizar el modelo para hacer predicciones con datos que no se utilizaron en el entrenamiento.

Para evaluar este patrón y en qué punto el modelo se sobre ajusta se iteró la técnica anterior 300 veces agregando cada vez mas variables de desfase. Los analisis se pueden ver en la **Figura 4**. Es interesante notar que el crecimiento del R^2 no es acelerado sino que casi constante, habiendo puntos en los que no crece (multicolinealidad). Antes de llegar a las 250 iteraciones en modelo explica a la perfección toda la varianza de la anomalía.

Evaluación de modelos multivariados

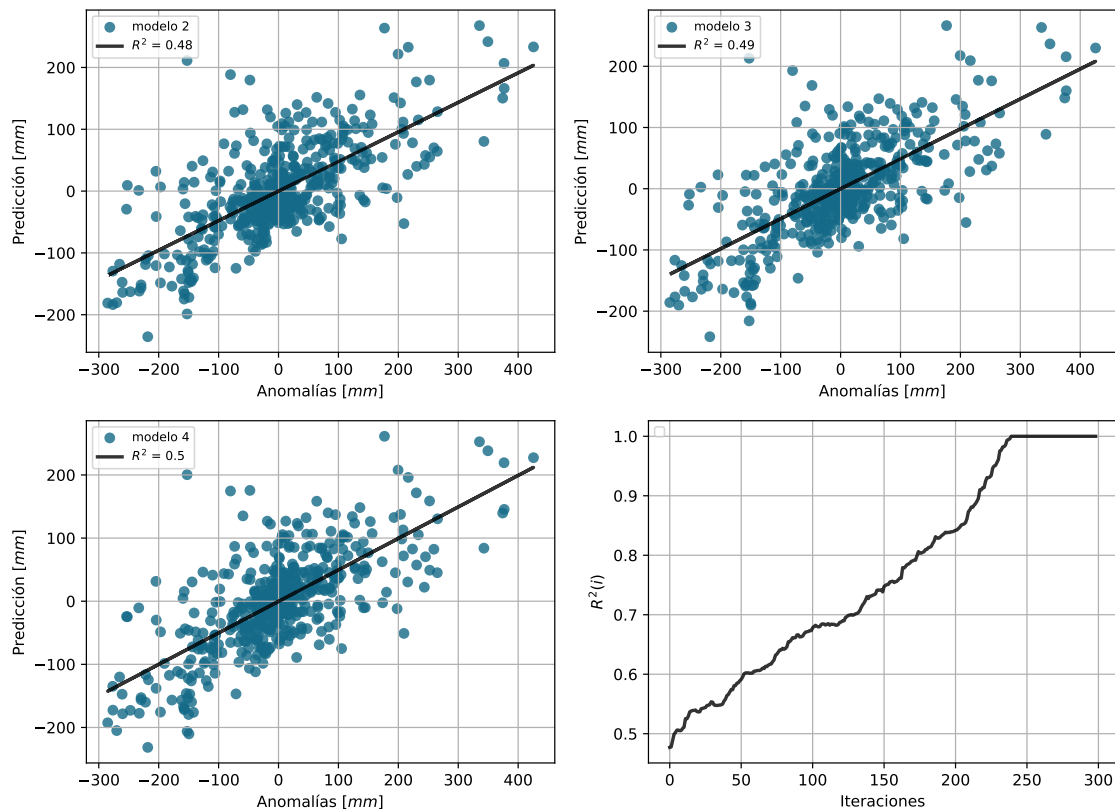


Figura 4: **Evaluación de modelos multivariados.** Los tres diagramas de dispersión muestran el desempeño de cada uno de los tres modelos (*modelo 2*, *modelo 3*, *modelo 4*). El gráfico inferior derecho muestra la evaluación de los primeros 300 modelos en los que se sumó una variable de desfase i .

Parte 4. Evaluación de modelos para cada mes

En esta ultima parte, se evaluó un modelo similar al *modelo 4* pero esta vez para cada mes de manera independiente. Los resultados se pueden ver en la *Figura 5*. Se observa que los modelos de los meses de verano tienen un menor rendimiento con respecto los de invierno y que el mejor rendimiento es el de octubre. Por otro lado, el rendimiento mejora con respecto al *modelo 4* en los meses de abril, mayo, julio y octubre. Se puede concluir que aplicar el *modelo 4* sería adecuado para análisis que abarquen todo el año, pero para análisis mas específicos, como solo el invierno, sería mas adecuado este tipo de modelos mas acotados.

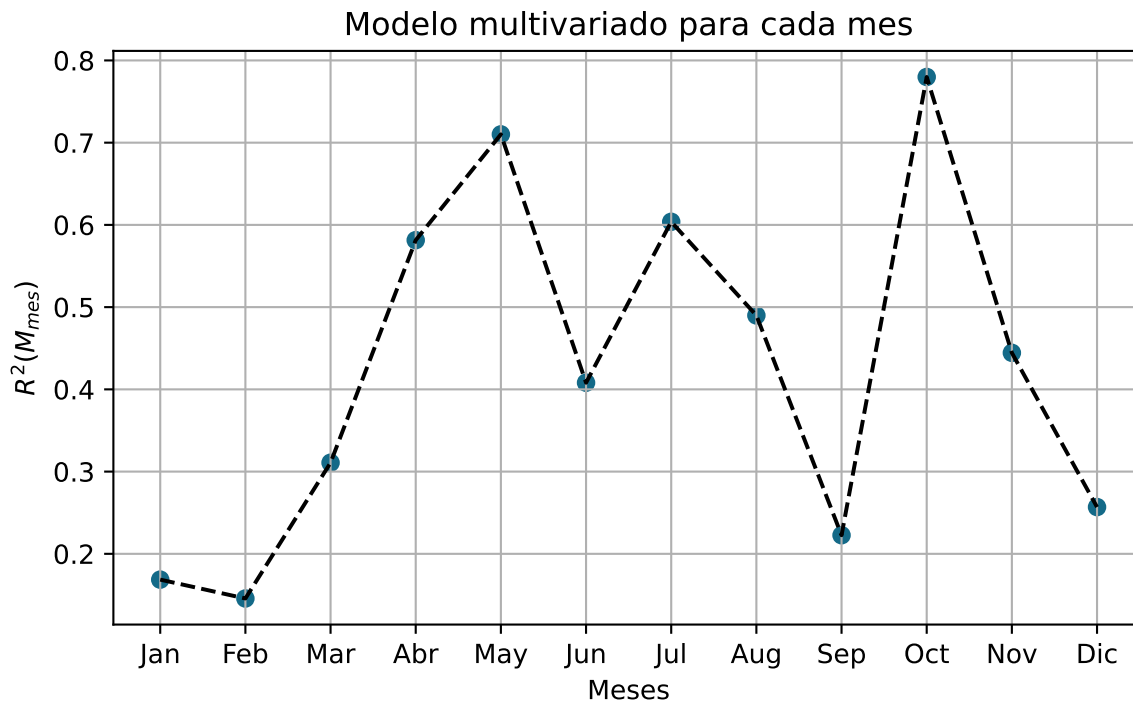


Figura 5: **Modelo multivariado para cada mes.** El gráfico muestra el desempeño de cada $modelo_{mes}$ medido en R^2 .

Referencias

- [1] M. H. De Groot M. J. Schervish. *Probabilidad y Estadística*. 1992.
- [2] D. S. Wilks. *Statistical Methods in the Atmospheric Sciences*. 2^a ed. 2006.