

ANÁLISIS DE CALIDAD DE DATOS MENSUALES DE TEMPERATURA Y PRECIPITACIÓN EN ESTACIONES COSTERAS ARGENTINAS

Valentina Rabanal¹, Silvina Righetti^{1,2}, Natalia Pessacg³
valenrabanal@gmail.com

¹ Servicio Meteorológico Nacional (SMN)

² Departamento de Ciencias de la Atmósfera y los Océanos (DCAO, FCEyN-UBA)

³ Instituto Patagónico para el Estudio de los Ecosistemas Continentales (IPEEC CONICET)

Palabras clave: Observaciones, valores atípicos, datos faltantes

1) INTRODUCCIÓN

Las áreas costeras de Argentina se extienden a lo largo de un amplio gradiente latitudinal, desde el estuario exterior del Río de la Plata hasta el extremo sur del país. A lo largo de estos 20 grados de latitud, las condiciones climáticas varían notablemente, desde climas húmedos en el norte hasta ambientes áridos en el sur, lo que configura una gran diversidad de características ambientales (Pessacg, 2022).

En el contexto de calentamiento global, resulta fundamental contar con una caracterización climática precisa de estos entornos costeros. Para ello, se recopilaron datos observados de temperatura y precipitación proporcionados por el Servicio Meteorológico Nacional (SMN), abarcando un período de 60 años. El objetivo principal de este trabajo es presentar y discutir el análisis de calidad realizado sobre estas observaciones climáticas, como paso previo indispensable para su uso en estudios de tendencias, variabilidad y cambio climático en las zonas costeras del país.

2) METODOLOGÍA

Se analizaron las series mensuales de temperatura y precipitación acumulada, desde 1961 a 2020, de 22 estaciones provistas por el SMN. En este trabajo se identificaron en primer lugar los datos faltantes en las series mensuales de cada variable. En el caso que alguna estación presentara extensos periodos sin datos, superando el umbral del 30%, se descartó automáticamente siguiendo las recomendaciones de Penalba et al. (2014).

Por otro lado, se identificaron los outliers (valores atípicos), considerando aquellos valores que fueran mayores o iguales a 2 desvíos, en el caso de temperatura media mensual, y mayores o iguales a 4 desvíos, para los acumulados mensuales de precipitación. Una vez identificados los outliers, se buscó analizar la veracidad de los mismos a partir de distintas técnicas. Si se concluía que el dato era un valor posible, se lo conservó, y si se determinaba que era un valor erróneo (por ejemplo, por un mal tipeo del número), se lo descartó.

Para analizar la validez de un outlier X_o , en primera instancia, se localizó la estación más cercana (radio <200 km) y se comparó X_o con el valor de ese mes en la estación vecina. Para los datos de temperatura, si en ambas estaciones el comportamiento del registro mensual era similar (es decir, ambos valores por encima o por debajo de sus propias medias), el outlier X_o se consideró un valor posible y se lo mantuvo en la serie, en caso contrario se lo descartó. En el caso de la precipitación, si en ambas estaciones sus acumulados se encontraban por encima de sus propias medias, el dato se conservó, en caso contrario se lo desechó.

Para los datos de precipitación, donde no existe una estación vecina (o la misma no tuviera dato), se observaron los acumulados diarios de ese mes, para así detectar los días de mayor contribución. Por otra parte, para la temperatura, se graficaron las series medias diarias u horarias para analizar el comportamiento durante todo el mes. En el caso que una serie de

temperatura media diaria contara con más de 4 días consecutivos con valores mayores al outlier positivo o menores al outlier negativo, no se consideró al valor mensual como sospechoso. Para casos puntuales de outliers recientes de precipitación, los récords se confirmaron o rechazaron a partir de información de medios o notas técnicas publicadas.

3) RESULTADOS Y DISCUSIÓN

Al analizar los datos faltantes de las series, las estaciones de Villa Gesell y Puerto Santa Cruz se descartaron automáticamente. Los porcentajes de faltantes fueron de 53% y 43% para la temperatura y precipitación de Villa Gesell, mientras que para la localidad santacruceña los datos faltantes alcanzaron el 68% y 72%, respectivamente. Luego de este descarte inicial, se siguió trabajando con 20 estaciones.

En las Figuras 1 a) y b) se muestran la cantidad de datos disponibles para cada estación de temperatura y precipitación, respectivamente. Se puede apreciar que la mayor cantidad de datos faltantes se encuentran en los primeros años de las series para ambas variables, mientras que para fin del período prácticamente todas las estaciones analizadas presentan observaciones. Las series que presentan un periodo blanco extenso al comienzo se deben a que la fecha de inicio de mediciones es posterior a 1961. Esto se puede observar en San Julián (1963), Tres Arroyos (1965), La Plata (1967), Viedma (1970), Benito Juárez (1981), Las Flores (1988) y Puerto Madryn (1991).

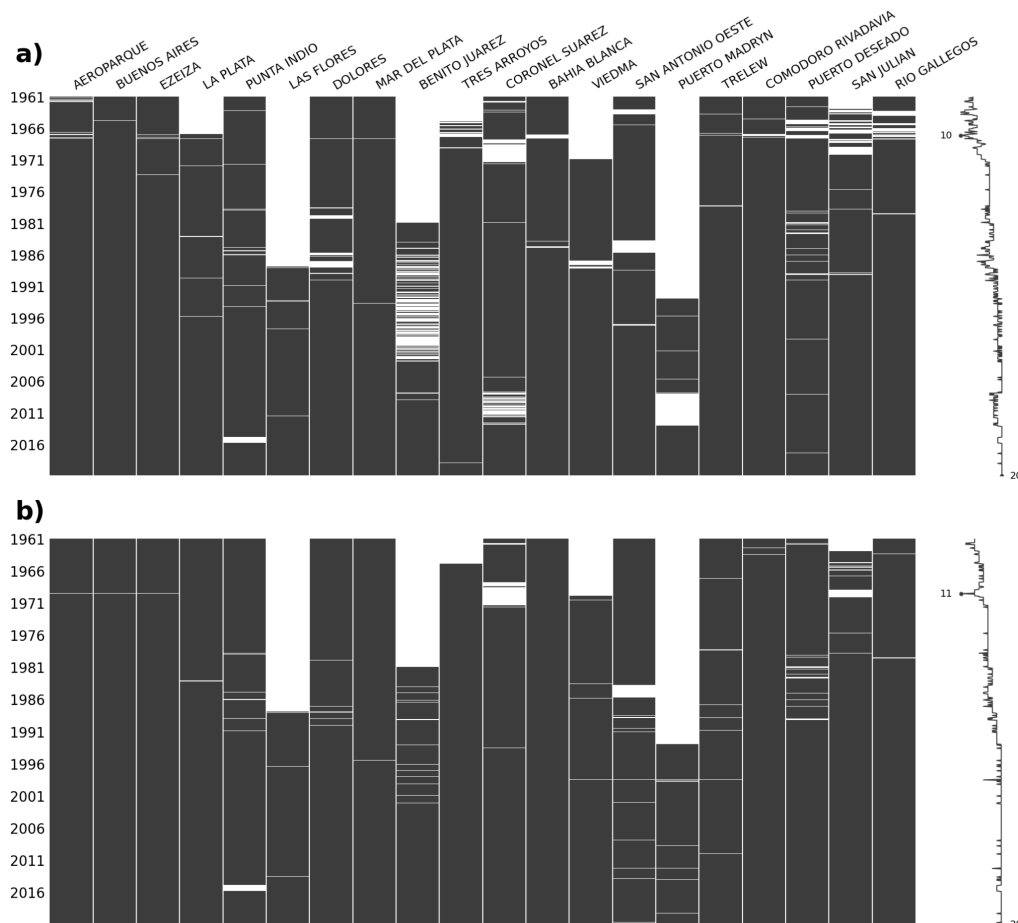


Figura 1: Cantidad de datos de a) temperatura media mensual y b) acumulados mensuales de precipitación. En el eje vertical se encuentran los meses, desde enero 1961 hasta diciembre 2020, mientras que en las columnas se presenta cada una de las estaciones. Las zonas grises indican la presencia de un dato y el blanco refiere a la ausencia. El gráfico de línea continua a la derecha indica la cantidad de datos para cada mes, integrando todas las estaciones.

Al comparar las Figuras 1 a y b, se ve que los mayores faltantes se encuentran en las series de temperatura media mensual. Sin embargo, la cantidad de datos faltantes de precipitación no incluye a los meses incompletos (es decir, aquellos en que no se midió todos los días). Un caso llamativo es el de Benito Juárez, donde en la Figura 1 a se observa una gran cantidad de datos faltantes (29%), mientras que en la Figura 1 b no se aprecia este comportamiento (menos del 5% de faltantes). No obstante, al analizar la cantidad de meses incompletos, esto totaliza el 32% de las observaciones registradas. Quedará para un análisis posterior determinar si se descarta o no a esta estación.

Respecto a los outliers, se identificaron 39 casos para temperatura y 40 para precipitación (Tabla 1). Se observa que 17 estaciones registraron al menos un outlier de temperatura, con San Antonio Oeste y San Julián como las localidades con mayor cantidad de valores que superan el criterio utilizado (6 y 5, respectivamente). En el caso de la precipitación, son 2 las estaciones que tienen más de 5 outliers, Puerto Madryn (6) y Trelew (8).

Luego de aplicar las técnicas descriptas en la metodología, se determinó que los 79 casos analizados eran posibles y no se los descartó de las series de estudio. Los datos finales que se considerarán para la caracterización climática de los entornos costeros (Tabla 1) abarcan más del 95% de los datos de precipitación, mientras que para temperatura se utilizarán más del 90% de los datos en todas las estaciones, excepto en Puerto Madryn, Benito Juárez y Coronel Suárez, donde los porcentajes de datos preservados es inferior.

En síntesis, este análisis permitió evaluar la calidad de las series climáticas disponibles, garantizando una base de datos robusta y confiable que servirá para la caracterización climática de los entornos costeros en estudios posteriores.

ESTACIÓN	TMP		PP		ESTACIÓN	TMP		PP	
	OUT	DAT (%)	OUT	DAT (%)		OUT	DAT (%)	OUT	DAT (%)
AEROPARQUE	1	720 (99)	0	720 (99)	CNEL. SUAREZ	1	720 (87)	0	720 (96)
BUENOS AIRES	1	720 (99)	0	720 (99)	B. BLANCA	1	720 (98)	3	720 (100)
EZEIZA	1	720 (99)	1	720 (99)	S. A. OESTE	6	682 (95)	5	686 (95)
LA PLATA	3	648 (99)	0	720 (99)	VIEDMA	3	600 (94)	2	612 (96)
PUNTA INDIO	1	720 (96)	0	720 (97)	PTO. MADRYN	2	336 (79)	6	360 (96)
DOLORES	0	720 (95)	0	720 (99)	TRELEW	2	720 (99)	8	720 (99)
LAS FLORES	1	396 (97)	1	408 (97)	C. RIVADAVIA	4	720 (99)	5	720 (99)
BENITO JUAREZ	0	480 (71)	0	480 (95)	PTO. DESEADO	1	720 (94)	5	720 (97)
MAR DEL PLATA	3	720 (99)	1	720 (99)	SAN JULIÁN	5	696 (90)	3	696 (94)
TRES ARROYOS	0	684 (96)	0	672 (98)	RÍO GALLEGOS	3	720 (96)	0	720 (99)

Tabla 1: Cantidad de outliers (OUT) y de datos (DAT) (y porcentajes) por estación.

REFERENCIAS

Penalba, O. C., J. A. Rivera, Pántano V. C. (2014). The CLARIS LPB database: constructing a long-term daily hydro-meteorological dataset for La Plata Basin, Southern South America. 1, 20–29, <https://doi.org/10.1002/GDJ3.7>

Pessacg, N., Blázquez, J., Lancelotti, J., Solman, S. (2022). Climate Changes in Coastal Areas of Patagonia: Observed Trends and Future Projections. In: Helbling, E.W., Narvarte, M.A., González, R.A., Villafañe, V.E. (eds) Global Change in Atlantic Coastal Patagonian Ecosystems. Natural and Social Sciences of Patagonia. Springer, Cham. https://doi.org/10.1007/978-3-030-86676-1_2