

Comparación de mediciones convencionales y automáticas de temperatura: procedimientos metodológicos y casos de estudio

Nota Técnica SMN 2023-135

María Mercedes Poggi¹, María de los Milagros Skansi¹, Karina Flores¹, Norma Garay¹ y Hernán Veiga¹

¹ Dirección Central del Monitoreo del Clima, Dirección Nacional de Ciencia e Innovación de Productos y Servicios, Servicio Meteorológico Nacional.

Enero 2023

Información sobre Copyright

Este reporte ha sido producido por empleados del Servicio Meteorológico Nacional con el fin de documentar sus actividades de investigación y desarrollo. El presente trabajo ha tenido cierto nivel de revisión por otros miembros de la institución, pero ninguno de los resultados o juicios expresados aquí presuponen un aval implícito o explícito del Servicio Meteorológico Nacional.

La información aquí presentada puede ser reproducida a condición que la fuente sea adecuadamente citada.

Resumen

El Servicio Meteorológico Nacional de Argentina (SMN) se encuentra en proceso de automatización de su red de estaciones. Se están realizando mediciones automáticas y manuales en paralelo de distintas variables, las cuales deben ser contrastadas para determinar los efectos del cambio de instrumental. El presente trabajo aborda esta temática y describe los procedimientos estadísticos que utiliza el SMN para la comparación y la caracterización estadística de las series de temperatura horaria, y de temperatura máxima, mínima y media diaria. Dichos procedimientos incluyen la identificación de sesgos, la aplicación de tests de hipótesis y el análisis de presencia de tendencias y periodicidades en las series de las diferencias, junto con el de frecuencia estacional de eventos ("frío extremo", "frío", "medio", "cálido" y "cálido extremo"). Los estudios se llevaron a cabo considerando el período completo de datos y desagregando la información en escalas mensual y estacional. Aquí se muestran los resultados de pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque Buenos Aires. Para estos tres lugares y el año completo, se encontraron medias inferiores a los 0.5°C en las series de las diferencias de las temperaturas máximas, mínimas y medias diarias, y en las series de las temperaturas horarias en los tiempos de observación. Las diferencias medias de las variables diarias resultaron, en su mayoría, negativas, lo que sugiere una sobreestimación por parte de los sistemas automáticos de observación. Para todas las ubicaciones y a partir del test de Kolmogorov-Smirnov, no se encontraron evidencias que permitan concluir que las series de temperaturas diarias obtenidas automática y convencionalmente no se extraen de la misma distribución, con un 95% de confianza. En cuanto a la comparación estacional de frecuencia de eventos, se observó una señal clara de sobreestimación y subestimación del número de casos registrados por parte de las estaciones automáticas en las categorías cálidas y frías, respectivamente, en Pigüé.

Abstract

The National Meteorological Service of Argentina (NMS) is in the process of automating its network of stations. Automatic and manual measurements of different variables are being collected in parallel, which must be compared in order to determine the effects of the change of instrumentation. This work addresses this topic and describes the statistical procedures used by the NMS for the comparison and statistical characterization of the hourly and daily maximum, minimum, and mean temperature series. These procedures include the identification of biases, the application of hypothesis tests and the analysis of the presence of tendencies and periodicities in the series of differences, together with the seasonal frequency of events ("extreme cold", "cold", "medium", "warm" and "extreme warm"). The studies were carried out considering the entire data period and disaggregating the information on monthly and seasonal scales. The results of pairs of stations located in Villa Gesell, Pigüé and Aeroparque Buenos Aires are shown here. For the three places and the whole year, averages of less than 0.5°C were found in the series of the differences of the daily maximum, minimum and mean temperatures, and in the series of the hourly temperatures at the observation times. The average differences of the daily variables were, mostly, negative, suggesting an overestimation by automatic observation systems. For all the locations and based on the Kolmogorov-Smirnov test, no evidence was found that would allow concluding that the daily temperature series obtained automatically and conventionally are not extracted from the same distribution, with 95% confidence. Regarding the seasonal comparison of frequency of events, a clear sign of overestimation and underestimation of the number of cases registered by the automatic stations in the warm and cold categories, respectively, was observed in Pigüé.

Palabras clave: temperatura, automatización, comparación, procedimientos estadísticos.

Citar como:

Poggi, M.M., M.M.Skansi, K. Flores, N. Garay, H. Veiga, 2023: Comparación entre mediciones manuales y automáticas de temperatura: procedimientos metodológicos y casos de estudio. Nota Técnica SMN 2023-135.

1. INTRODUCCIÓN

La automatización de una red de observación ofrece múltiples ventajas: proporciona observaciones en tiempo real con alta resolución temporal a un costo relativamente bajo luego de su instalación; genera un suministro continuo de datos sin interrupciones; y permite obtener mediciones en lugares no densamente poblados, así como de difícil acceso. Sin embargo, presenta desafíos para la vigilancia del clima a largo plazo, dado que toda modificación realizada en un sistema de observación (en este caso, la transición de las técnicas de observación manuales a las automáticas) puede introducir sesgos sistemáticos y generar inhomogeneidades en el registro climático, las cuales deben evaluarse y, de ser necesario, ajustarse (WMO, 2017). Como las estaciones meteorológicas automáticas (EMAs) no necesariamente están instaladas en las mismas ubicaciones y alturas que las convencionales (EMCs), sumado a que los nuevos instrumentos tienen diferentes sensibilidades, precisión y tiempos de respuesta a las variables atmosféricas, probablemente se originarán discontinuidades en las series de tiempo que deben ser analizadas.

A la hora de reemplazar estaciones meteorológicas, la Organización Meteorológica Mundial (OMM) recomienda obtener y utilizar observaciones convencionales o manuales y automáticas en paralelo como práctica principal para el proceso de homogeneización (WMO, 2007). La “Guía de Prácticas Climatológicas” (OMM, 2018) sugiere llevar a cabo registros en un período común (período de superposición) de por lo menos un año, aunque preferentemente dos o más, para determinar los efectos de la sustitución de instrumental y de emplazamiento en los datos climatológicos. Esto posibilitará realizar comparaciones entre las mediciones, detectar sesgos, desarrollar relaciones matemáticas y derivar ajustes a partir de ellas. En el pasado, las observaciones en períodos superpuestos no eran frecuentes, por lo que también se empleaban métodos estadísticos utilizando datos de estaciones vecinas para inferir ajustes y reducir las discontinuidades por cambios en la instrumentación, procedimiento de observación o ubicación. Este enfoque también es útil en la actualidad, en especial cuando no se cuentan con observaciones en paralelo.

Numerosos trabajos abordan la identificación y el cálculo de sesgos entre series manuales y automáticas de temperatura (Hausfather y otros, 2016; Fiebrich y Crawford, 2009; y Davey y Pielke, 2005; entre otros). Dichos estudios muestran que los sesgos rara vez son constantes y dependen de la interacción entre las variables meteorológicas y las características geográficas locales de cada lugar, como la altitud, la exposición y la distancia a las masas de agua (Acquaotta y otros, 2019). Milewska y Vincent (2016), por ejemplo, analizaron la diferencia entre datos convencionales y automáticos de temperaturas máximas y mínimas diarias en función de la estación del año y de las condiciones de velocidad del viento. Encontraron que la mayoría de los sesgos variaba con las estaciones y pocos mostraban dependencia con el viento.

Entre otros trabajos de interés, también se destaca el de Kaspar y otros (2016), que presenta el análisis y la comparación de mediciones paralelas de temperatura para estaciones ubicadas en diferentes regiones climáticas de Alemania. Dicho estudio no sólo determinó las diferencias entre las mediciones manuales y automáticas en distintas escalas temporales (horarias y diarias), sino también analizó el impacto de los cambios en los sensores y del procesamiento de los datos en la homogeneidad de las series temporales. Para ello, se llevaron a cabo diversas acciones: se compararon las mediciones horarias automáticas y tradicionales de temperatura; se contrastaron diferentes métodos para calcular la temperatura media diaria; y se compararon las temperaturas medias y extremas diarias obtenidas con los dos sistemas de observación. La duración del período de superposición fue de al menos seis años.

Acquaotta y otros (2019), por su parte, identificaron los sesgos en las temperaturas máximas y mínimas diarias utilizando registros en paralelo en estaciones de Piamonte, Italia, y de la Península de Gaspé,

Canadá. Las estaciones se seleccionaron de forma tal de minimizar las diferencias en el entorno, la exposición, la distancia, y la elevación entre ambas mediciones, siendo el tipo de instrumento el factor de sesgo más importante. Con un período de superposición de seis años, el procedimiento para la comparación se dividió en cuatro partes: una caracterización estadística de las series diarias de temperatura; una comparación entre las series diarias; una comparación entre tipos de eventos (extremadamente fríos, fríos, normales, cálidos y extremadamente cálidos); y una verificación de la homogeneidad de las series de diferencia. Sus resultados mostraron una mayor frecuencia de días cálidos y extremadamente cálidos en los registros automáticos, en comparación con los manuales. Concluyeron, entonces, que el uso de sistemas automáticos podría sesgar significativamente el cálculo de eventos extremos.

En la actualidad, el Servicio Meteorológico Nacional (SMN) se encuentra en proceso de automatización de su red de EMCs, por lo que se están realizando mediciones automáticas y manuales en paralelo de diferentes variables meteorológicas, las cuales deben ser contrastadas. En este trabajo se abordan los procedimientos que utiliza el organismo para comparar y cuantificar las posibles diferencias entre los registros de temperatura de ambos sistemas de observación. Como casos de estudio, se consideraron pares de estaciones cuyas EMAs no sólo pertenecen al SMN, sino también a otras redes o propietarios. Estas se sitúan aproximadamente en el mismo lugar que las EMCs o a menos de 5 km de separación, con el objetivo último de garantizar la homogeneidad de los registros y la continuidad de las series de referencia de la red del SMN. Cabe destacar que, si bien preservar las series históricas de dicha red representa la principal finalidad de este trabajo, en un futuro, las metodologías aquí planteadas serán utilizadas para la comparación de pares de estaciones que se encuentren a una mayor distancia, de hasta 20 km. Estos pares de estaciones no podrán ser utilizados para análisis de homogeneización, pero sí para estudios exploratorios, en pos de determinar la calidad de las observaciones automáticas.

2. MEDICIONES CONVENCIONALES VS AUTOMÁTICAS

Siguiendo a Kaspar y otros (2016), las EMCs y EMAs presentan varias diferencias en la forma de obtención de sus datos, las cuales deben ser tenidas en cuenta en la comparación:

- En las EMCs, las temperaturas se miden con termómetros de líquido con lectura manual. Por el contrario, en las EMAs se utiliza un termómetro de resistencia de platino. También existen diferencias en los tiempos de calibración de los instrumentos.
- En general, las mediciones manuales se realizan cada hora, si bien la cantidad de observaciones diarias puede verse reducida dependiendo del plan de labor de la estación. Las mediciones automáticas, en cambio, son continuas y sus registros se almacenan en la base de datos cada cierta cantidad de minutos.
- En cuanto al cálculo de la temperatura media diaria, las EMCs lo hacen a partir de los cuatro tiempos de observación, de acuerdo con la Ecuación 1:

$$T_{media} = \frac{T_I + T_{II} + T_{III} + T_{IV}}{4} \quad (1)$$

donde T_I es la temperatura observada a las 03:00 h; T_{II} , a las 09:00 h; T_{III} , a las 15:00 h; y T_{IV} , a las 21:00 h. Las EMCs que no miden la temperatura en el primer tiempo de observación utilizan la Ecuación 2:

$$T_{media} = \frac{T_{II} + T_{III} + T_{IV}}{3} \quad (2)$$

Cabe destacar que, en estos casos, la temperatura media diaria sólo se utiliza para el cálculo de la temperatura media mensual, la cual se obtiene aplicando una corrección que depende de la estación y del mes del año. A nivel diario no es representativa, dado que hay una sobreestimación de la variable.

Para las EMAs y aquellas ECMs que tienen mediciones en todas las horas, los valores medios diarios también se pueden obtener como el promedio aritmético de los 24 valores horarios.

- Para los registros de las temperaturas máximas y mínimas diarias, las EMCs emplean termómetros de mercurio y alcohol, respectivamente. Sus lecturas se realizan a las 09:00 y 21:00 h. En las EMAs, todas las mediciones se obtienen a partir de termómetros eléctricos. Cabe destacar que, en base a los extremos térmicos diarios, se puede considerar otro enfoque para el cálculo de la temperatura media diaria, que viene dado por la Ecuación 3:

$$T_{media} = \frac{T_{min} + T_{max}}{2} \quad (3)$$

donde T_{min} es la mínima diaria y T_{max} , la temperatura máxima diaria. Esta última definición es muy utilizada en estudios de cambio climático global por su simplicidad y disponibilidad (New y otros, 1999, entre otros).

3. SELECCIÓN DE PARES DE ESTACIONES

Hasta el momento, las EMCs seleccionadas para el análisis pertenecen a la red del SMN, mientras que las EMAs, al SMN, al Sistema Nacional de Radares Meteorológicos (SINARAME), a la Bolsa de Cereales de Córdoba y a la Empresa Argentina de Navegación Aérea (EANA). EANA, junto con el SMN, poseen EMAs que componen la red AWOS (“Automatic Weather Observation System”). Existen otras instituciones que integran sus datos en la red de observación del SMN (la Universidad de la Punta de San Luis, el Ministerio de Producción de Corrientes y Chaco, la Bolsa de Cereales de Entre Ríos, etc.) que se analizarán para también utilizarlas. En esta instancia, la evaluación y el control de la calidad de los datos serán cruciales, ya que se desconoce el mantenimiento y el estado de los sensores de muchas de las EMAs que no son responsabilidad del SMN. Las EMAs de la red AWOS, ubicadas en las pistas de los principales aeropuertos, sí tienen un mantenimiento regular, debido a que sirven de soporte para la meteorología aeronáutica.

Los procedimientos metodológicos aquí presentados y aplicados están pensados para casos de EMCs y EMAs (o pares de estaciones) que se encuentran en el mismo predio o a una distancia máxima de 5 km, y aproximadamente a la misma elevación. Si bien no es una regla estricta ni tampoco hay un consenso único con ello, esta restricción de distancia máxima implicaría que, en un futuro, las EMAs podrán llegar a reemplazar a las EMCs en el proceso de automatización de la red de estaciones, manteniendo el mismo número identificador. Los registros automáticos y manuales podrán ser homogeneizados, garantizando así la continuidad de las series históricas de la red del SMN.

En un futuro, los procedimientos metodológicos considerados también se aplicarán en EMAs y EMCs que se encuentren a una distancia cercana pero que no cumplan con el apartamiento máximo de 5 km. Estas estaciones no podrán ser utilizadas para análisis de homogeneización, pero sí podrán ser tenidas en

cuenta para estudios exploratorios y comparativos, que empleen como base las mismas técnicas estadísticas aquí propuestas. Para tales propósitos, los pares de estaciones a considerar seguirán los criterios planteados y adoptados por Acquotta y otros (2019) que, a su vez, fueron formulados por diferentes autores:

- La distancia entre los pares de estaciones debe ser inferior a los 20 km (Isotta y otros, 2013).
- La diferencia de altura entre los pares de estaciones debe ser inferior a 50 m (Biancotti y otros, 2005).
- La exposición de los instrumentos y la topografía local deben ser similares en ambos sitios, dado que las fuertes diferencias en la topografía y la exposición influyen fuertemente en la distribución de la temperatura (Trewin y Trevitt, 1996).

Según Acquotta y otros (2019), estos criterios no sólo evitan la corrección de los datos por elevación y latitud, sino también reducen las diferencias debidas al microclima. La influencia del microclima en las mediciones de temperatura de estaciones cercanas es potencialmente mucho mayor que las influencias por las diferencias de altura (Gallo, 2005).

Además de los criterios presentados, existe uno adicional planteado por Acquotta y otros (2019) que no es tenido en cuenta: el período de solapamiento de las mediciones en paralelo debe ser superior a 5 años (Vincent y Mekis, 2009). En este trabajo, en cambio, se considera un período de al menos un año, siguiendo los lineamientos de la OMM (2018). En un futuro, esta flexibilización del criterio también permitirá incluir una mayor cantidad de estaciones para el análisis.

4. PROCEDIMIENTOS METODOLÓGICOS PARA LA COMPARACIÓN EN DISTINTAS ESCALAS TEMPORALES

Si bien la comparación entre las mediciones de las EMCs y las EMAs se centra en la escala diaria, también se considera la horaria. El enfoque a implementar se divide en cinco partes:

1. comparación de las series de temperatura horaria;
2. comparación y caracterización estadística de las series de temperatura máxima y mínima diaria;
3. cálculo, comparación y caracterización estadística de las series de temperatura media diaria;
4. detección de tendencia y periodicidades en las series de las diferencias de temperatura máxima, mínima y media diaria;
5. comparación estacional de frecuencia de eventos.

En todas las partes, las series y las distribuciones de las diferencias de temperatura se calculan restando las observaciones convencionales menos las automáticas, siguiendo la convención planteada por Milewska y Vicent (2016), y Acquotta y otros (2019). De esta manera, valores positivos indican una subestimación por parte de los datos automáticos, mientras que valores negativos, una sobreestimación. Además, para una comparación correcta entre las series de las EMCs y las EMAs, en caso de encontrarse un dato ausente en una serie, se elimina también el de su homóloga.

Cabe destacar que en el Apéndice 1 se presentan las consideraciones tenidas en cuenta y los procedimientos llevados a cabo para obtener las series horarias y diarias de las EMAs.

4.1. Comparación de las series de temperatura horaria

Como primera instancia y siguiendo el trabajo de Kaspar y otros (2016), se calculan las series y las distribuciones de las diferencias entre los datos convencionales y automáticos de temperatura en los tiempos de observación. Se computan los valores medios y los desvíos estándares de dichas series. También se grafican boxplots de las diferencias en función de los meses. Para todos los análisis, se relevan y excluyen aquellos días y horas con más de 5°C de diferencia en módulo, contrario a lo que consideran Kaspar y otros (2016), que toman el umbral de 2°C de diferencia para la remoción del dato en escala horaria.

Cabe resaltar que, en caso de existir un porcentaje mayor o igual a 30% de dato convencional faltante a las 03:00 h, se remueven los valores asociados a dicho horario, no interviniendo en las series de las diferencias horarias. Más aún, no se pueden calcular las temperaturas medias diarias convencionales derivadas de los cuatro tiempos de observación y la Ecuación 1.

4.2. Comparación y caracterización estadística de las series de temperatura máxima y mínima diaria

Se calculan las series y distribuciones de las diferencias entre los datos convencionales y automáticos de temperatura máxima y mínima diaria. Siguiendo a Kaspar y otros (2016), también se grafican boxplots de las diferencias para cada mes del año. Nuevamente, se trabaja con diferencias que oscilan entre $\pm 5^\circ\text{C}$. Los valores fuera de ese rango se relevan y excluyen del análisis por considerarse excesivos. Según Squintu y otros (2019), y Acquaotta y otros (2016), estos determinan un error en las mediciones diarias de temperatura, ya sean manuales o automáticas.

Para las series de las EMCs y las EMAs, y las series de las diferencias, se computan distintas estadísticas descriptivas: la media, la mediana, el primer y tercer cuartil, y los valores mínimos y máximos. Se identifica además el porcentaje de registros faltantes.

Para la comparación de cada par de series, se calculan el error cuadrático medio (RMSE, por sus siglas en inglés) y el coeficiente de correlación de rango de Kendall o Rank-Kendall, el cual también se testea. El RMSE se puede interpretar como un error promedio entre los dos registros, y el coeficiente de correlación mide la asociación entre las dos series. Como el RMSE está afectado por el bias de la media, también se considera la métrica RMSE centrado. Con respecto a la correlación, y dada la naturaleza de la variable a analizar, se podría optar por el coeficiente de Pearson, en lugar del de Rank-Kendall. En este sentido, por ejemplo, Vincent y otros (2018) computan la correlación de Pearson entre las temperaturas diarias de las observaciones en paralelo, previo y post eliminar la estacionalidad (la estacionalidad se obtiene promediando los valores diarios durante una ventana de 5 días centrada en cada día calendario y durante el período de superposición de 5 años). Aquí, sin embargo, se eligió el coeficiente de Rank-Kendall siguiendo el trabajo de Acquaotta y otros (2019). En un futuro, y si se cuenta con un período de superposición suficientemente largo, también se implementará lo planteado por Vincent y otro (2018), filtrando la estacionalidad.

Por otra parte, se llevan a cabo las pruebas o los tests t-Student y de Wilcoxon de rango con signo para la comparación de medias, y el de Kolmogorov-Smirnov para la comparación de distribuciones. Los dos primeros tests posibilitan explorar la similitud estadística de los valores medios, mientras que el último permite evaluar si ambas series provienen de la misma distribución. Los resultados de todos los tests se analizan a partir del p-valor, considerando un nivel de significancia $\alpha = 0.05$. Cabe destacar que estas pruebas de hipótesis se realizan siguiendo el trabajo de Acquaotta y otros (2019), si bien ellos sólo utilizan

los tests t-Student, de Wilcoxon y de Kolmogorov-Smirnov clásicos, sin considerar la dependencia serial o autocorrelación de las observaciones diarias dentro de cada serie. Una breve descripción de estos tres tests bajo dependencia serial se puede encontrar en el Apéndice II.

Como adicional, y en base a los resultados de Milewska y Vincent (2016), los análisis también se realizan desagregando la información según las estaciones del año: verano (enero, febrero y diciembre, DEF), otoño (marzo, abril y mayo, MAM), invierno (junio, julio y agosto, JJA), y primavera (septiembre, octubre y noviembre, SON).

4.3. Cálculo, comparación y caracterización estadística de las series de temperatura media diaria

Siguiendo el trabajo de Kaspar y otros (2016), las temperaturas medias diarias se calculan a partir de tres procedimientos: el tradicional (utilizando la Ecuación 1 para las EMCs y el promedio aritmético de los 24 valores horarios para las EMAs); uno basado en los registros en los tiempos de observación (empleando la Ecuación 1 tanto para las EMCs como para las EMAs); y el último como el valor medio entre los extremos diarios (definido por la Ecuación 3). Para aquellas EMCs que cuentan mediciones en todas sus horas, se considera el procedimiento adicional de calcular la temperatura media diaria como el promedio de los 24 registros horarios para ambos sistemas de observación. Luego, las series manuales y automáticas, y las series de las diferencias se comparan con los mismos procedimientos que se usan para las temperaturas máximas y mínimas diarias.

Cabe destacar que los análisis de esta sección posibilitan evaluar las diferencias de las temperaturas medias diarias no sólo por el cambio de instrumental, sino también por la modificación del procedimiento de cálculo.

4.4. Detección de tendencias y ciclos en las series de las diferencias de temperatura máxima, mínima y media diaria

Se analiza la presencia de componentes determinísticas y de tendencias lineales en las series de las diferencias de las temperaturas diarias. Para lo primero, se emplea el test de Mann-Kendall. Para lo segundo, se testean las pendientes de las rectas de regresión ajustadas por cuadrados mínimos, utilizando el test de Fisher para determinar su significancia. Luego, se realiza un análisis armónico para detectar periodicidades, ciclos perfectos, previo restar las tendencias lineales en caso de que estas sean significativas. La significancia de dichas periodicidades se estudia mediante el análisis espectral y el modelo de ruido rojo o blanco. La elección del modelo se determina en función del test de Anderson.

4.5. Comparación estacional de frecuencia de eventos

Como instancia final, las series de temperatura máxima, mínima y media diarias convencionales y automáticas se dividen en 5 clases o eventos, denominados: “frío extremo” (EX_F), “frío” (F), “medio” (M), “cálido” (C) y “cálido extremo” (EX_C), de acuerdo a los valores de sus percentiles (Tabla I). El análisis de los eventos se realiza a escala estacional. Para cada estación del año, los percentiles se estiman considerando una única distribución, creada a partir de los datos manuales y automáticos. Luego, para cada serie y temporada, se calcula la frecuencia de valores dentro de cada clase, y se obtiene la proporción correspondiente a cada sistema de observación. Nuevamente, esta sección se realiza siguiendo los procedimientos utilizados por Acquattrota y otros (2019).

Tabla I: Rangos de las cinco clases o categorías de eventos de temperatura. Los umbrales se estimaron para cada estación del año, a partir de los percentiles (P) de una única distribución, creada mediante la unión de las series diarias convencionales y automáticas. Con “ x ” se indica la serie de temperatura a clasificar.

Clase	Rango
EX_F	$x \leq P5$
F	$P5 < x \leq P20$
M	$P20 < x \leq P80$
C	$P80 < x \leq P95$
EX_C	$P95 < x$

5. CASOS DE ESTUDIO

En esta sección se presentan los resultados de los pares de estaciones de Villa Gesell, Pigüé y Aeroparque Buenos Aires (Figura 1, Tabla II), con sus EMAs pertenecientes al SMN, a SINARAME y a la red AWOS de EANA, respectivamente. Estas estaciones tienen menos de un 5% de datos faltantes en la escala horaria en el período de superposición analizado. Este es de al menos un año y puede ser, por distintos motivos, menor al período de registros en paralelo real.

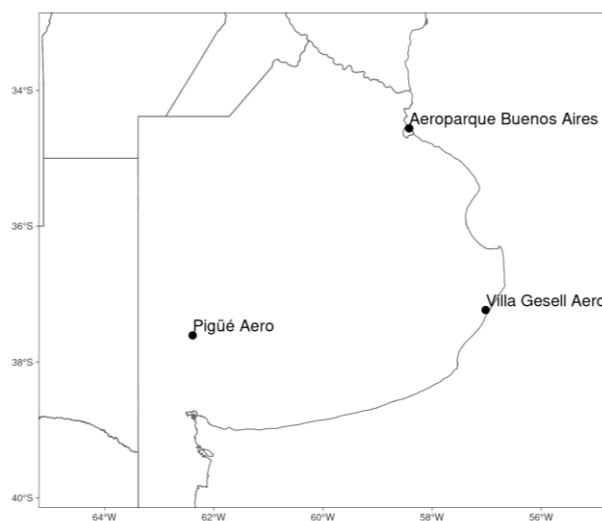


Figura 1: Ubicación de las EMCs y EMAs utilizadas para la comparación.

Tabla II: Pares de EMCs y EMAs utilizados para la comparación, junto con la distancia entre ellas, la diferencia de altura y el período de registros en paralelo analizado. Una diferencia de altura positiva indica que la EMC está a mayor elevación que la EMA, mientras que una diferencia negativa, lo contrario. Se presenta también el plan de labor de las EMCs y los pasos temporales de almacenamiento de datos en las EMAs.

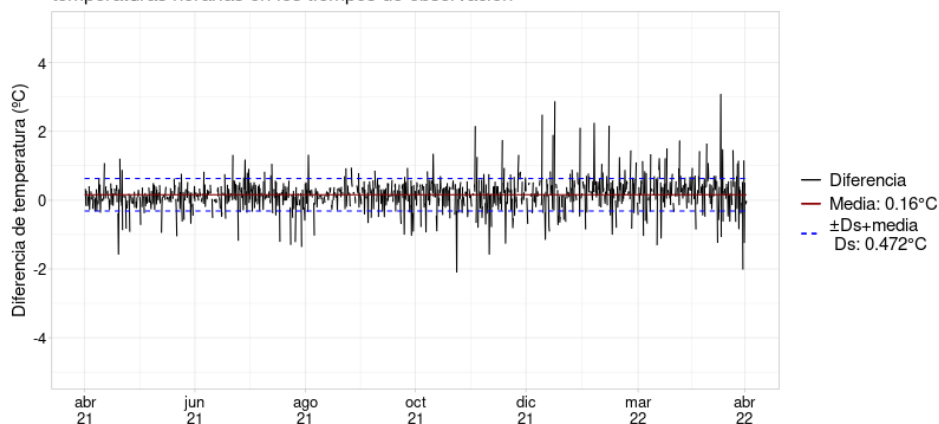
Nombre de EMC	Nombre de EMA	Red de EMA	Plan de labor en EMC	Paso temporal de datos en EMA (min)	Distancia entre EMC y EMA (km)	Diferencia de altura entre EMC y EMA (m)	Período de superposición analizado
Villa Gesell Aero	Villa Gesell	SMN	3 mediciones diarias	5	0.01	0	29-04-2021 al 30-04-2022
Pigüé Aero	RMA10 Pigüé	SINARAME	3 mediciones diarias	10	0	-3	01-01-2019 al 30-04-2022
Aeroparque Buenos Aires	Aeroparque	Red AWOS de EANA	4 mediciones diarias	1	0.35 La EMA se encuentra en la cabecera de la pista.	0	31-12-2018 al 16-09-2020

5.1. Comparación de las series de temperatura horaria

En las Figuras 2, 3 y 4 se presentan las series temporales y los histogramas de las diferencias entre las temperaturas horarias en los tiempos de observación medidas convencional y automáticamente para los pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque Buenos Aires, respectivamente. Se muestran también las diferencias medias y los desvíos estándares. En todos los casos se trabaja con diferencias que oscilan entre $\pm 5^{\circ}\text{C}$.

Estaciones Villa Gesell Aero (EMC) y Villa Gesell SMN (EMA)

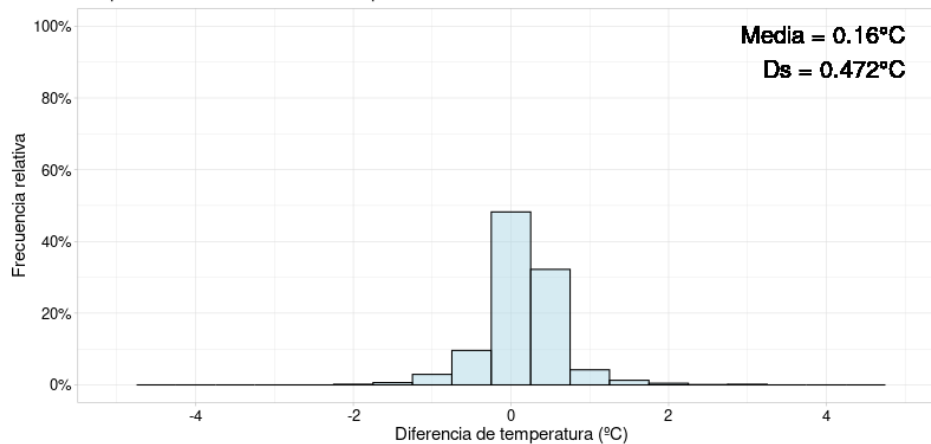
Serie de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación



Período: 29-04-2021 21:00 - 30-04-2022 21:00.
Cantidad de datos: 1093. Porcentaje de datos faltantes: 0.55%.

Estaciones Villa Gesell Aero (EMC) y Villa Gesell SMN (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación

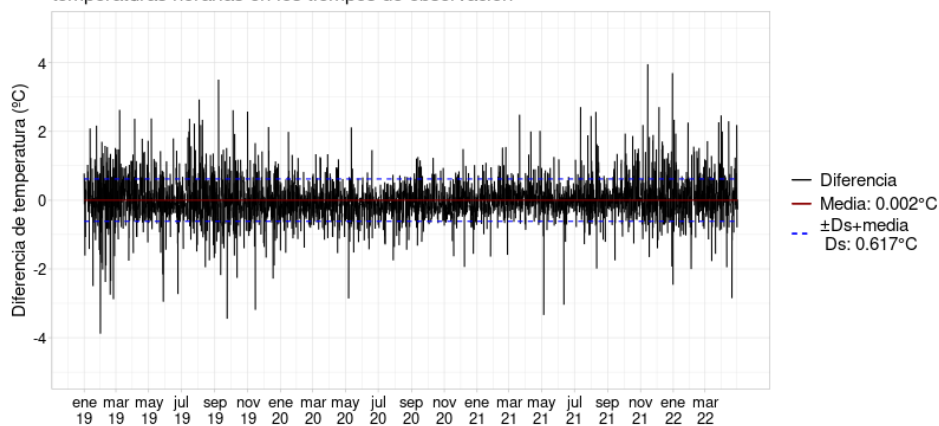


Período: 29-04-2021 21:00 - 30-04-2022 21:00.
Cantidad de datos: 1093. Porcentaje de datos faltantes: 0.55%.

Figura 2: Arriba: Serie temporal de las diferencias entre las mediciones convencionales y automáticas de temperatura en los horarios de observación (línea negra) para el par de estaciones ubicadas en Villa Gesell. En línea roja se representa la diferencia media y en líneas azules, los desvíos estándares sumados a esta. **Abajo:** Histograma de las diferencias entre las mediciones convencionales y automáticas de temperatura en los horarios de observación.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

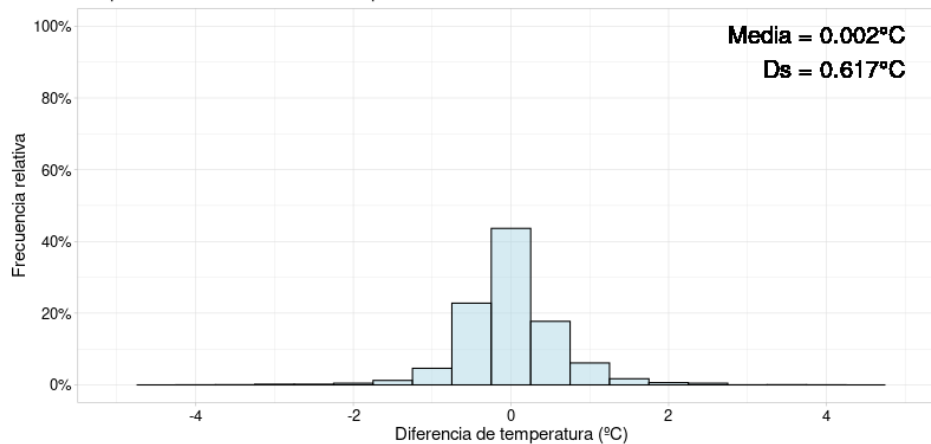
Serie de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación



Periodo: 01-01-2019 09:00 - 30-04-2022 21:00.
Cantidad de datos: 3612. Porcentaje de datos faltantes: 0.99%.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación

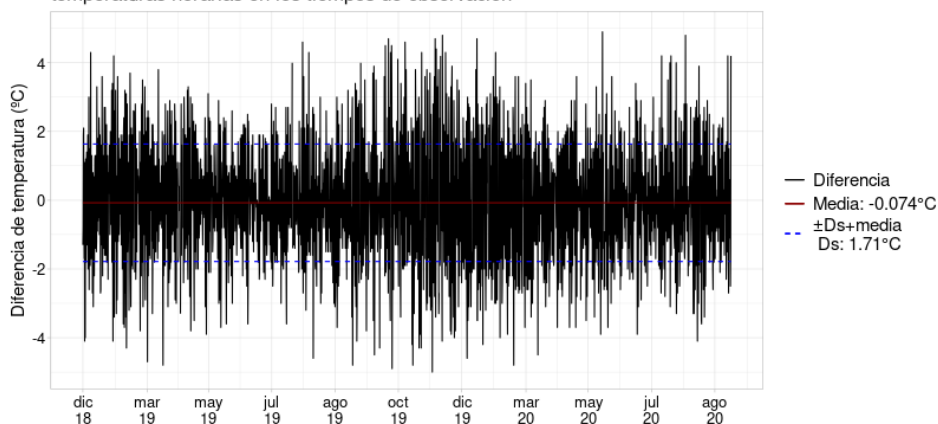


Periodo: 01-01-2019 09:00 - 30-04-2022 21:00.
Cantidad de datos: 3612. Porcentaje de datos faltantes: 0.99%.

Figura 3: Ídem Figura 2 pero para el par de estaciones ubicadas en Pigüé.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

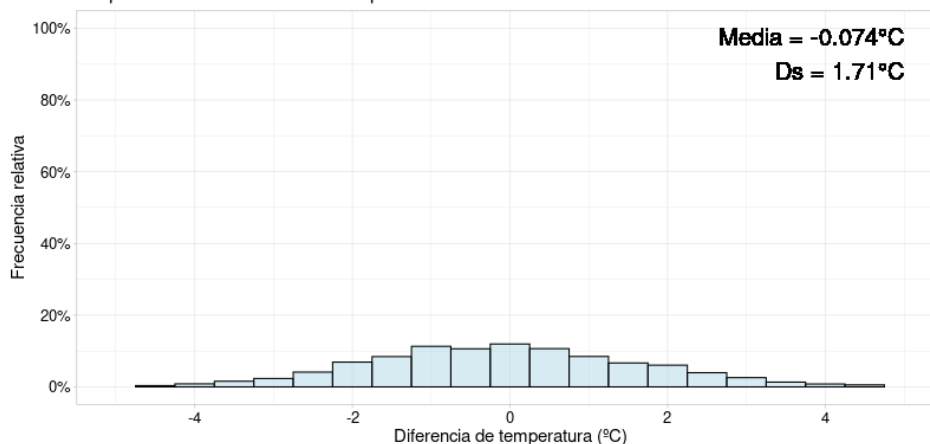
Serie de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación



Periodo: 31-12-2018 21:00 - 16-09-2020 09:00.
Cantidad de datos: 2475. Porcentaje de datos faltantes: 0.96%.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas horarias en los tiempos de observación



Periodo: 31-12-2018 21:00 - 16-09-2020 09:00.
Cantidad de datos: 2475. Porcentaje de datos faltantes: 0.96%.

Figura 4: Ídem Figura 2 pero para el par de estaciones ubicadas en Aeroparque.

Si bien en las tres ubicaciones se encuentran diferencias promedio inferiores a los 0.2°C en módulo, el menor valor se obtiene en Pigüé (Figura 3). La media de las diferencias allí es de 0.002°C. Analizando las series temporales, se observa que las diferencias en Villa Gesell presentan menor dispersión, seguido por las de Pigüé (Figuras 2 y 3 arriba). En cuanto a las distribuciones de frecuencia, la de Aeroparque revela una forma achatada (Figura 4 abajo), indicando poca concentración de los valores en torno a su media. Esto también se ve reflejado en la magnitud del desvío estándar. Para el resto de las ubicaciones, sí se observa una distribución empinada, con un marcado intervalo modal en torno a los 0°C.

En Villa Gesell no se registraron diferencias de temperaturas horarias superiores a los 5°C, en Pigüé sólo 2, mientras que en Aeroparque, más de 20. La máxima diferencia en esta última ubicación se observó el día 07-03-2019 a las 21 h y fue de 8.8°C.

5.2. Comparación y caracterización estadística de las series de temperatura máxima y mínima diaria

En las Figuras 5, 6 y 7 se presentan las series temporales y los histogramas de las diferencias entre los datos tradicionales y automáticos de temperatura mínima diaria para los tres pares de estaciones considerados. Se muestran también las diferencias medias y los desvíos estándares. En todos los casos se trabaja con diferencias que oscilan entre $\pm 5^{\circ}\text{C}$.

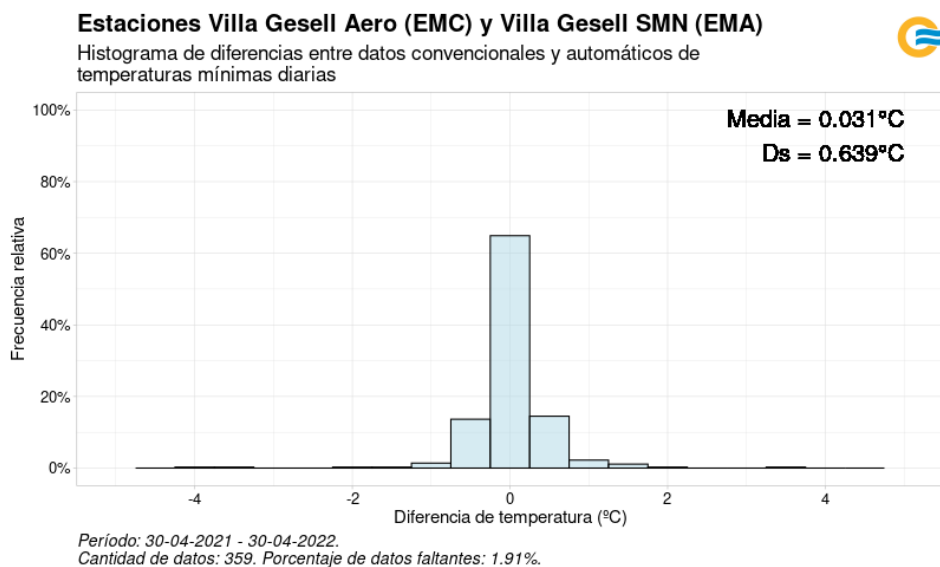
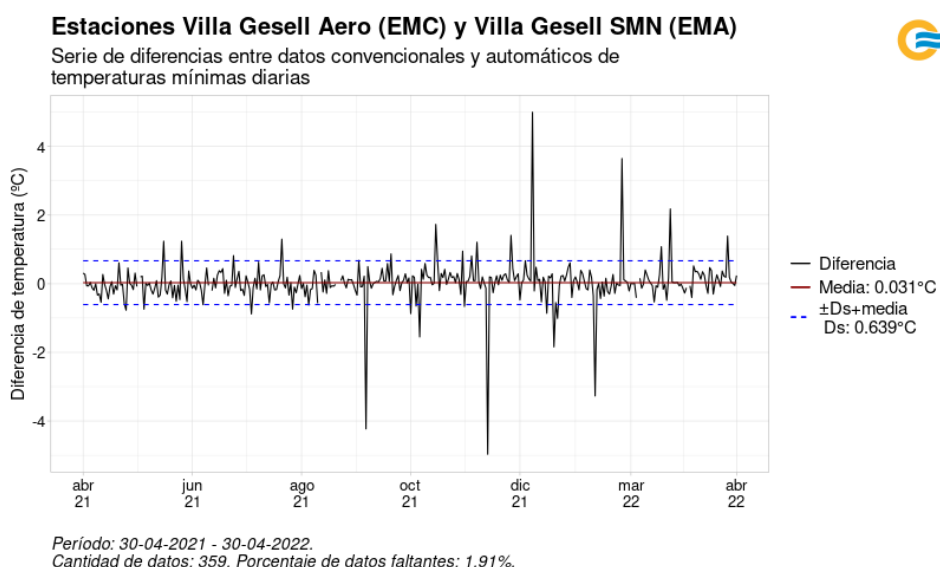
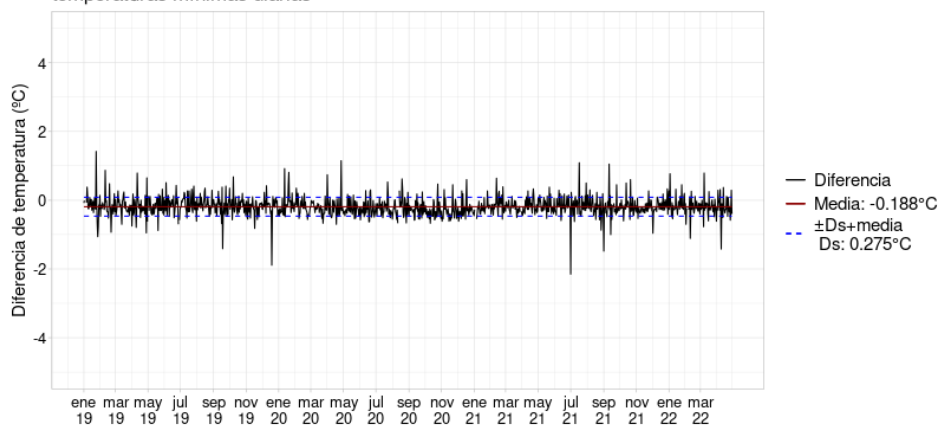


Figura 5: Arriba: Serie temporal de las diferencias entre las mediciones convencionales y automáticas de temperatura mínima diaria (línea negra) para el par de estaciones ubicadas en Villa Gesell. En línea roja se representa la diferencia media y en líneas azules, los desvíos estándares sumados a esta. **Abajo:** Histograma de las diferencias entre las mediciones convencionales y automáticas de temperatura mínima diaria.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

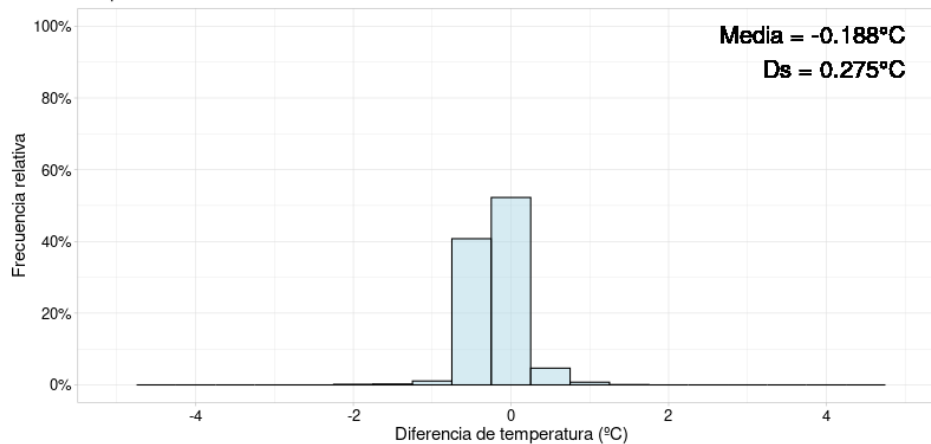
Serie de diferencias entre datos convencionales y automáticos de temperaturas mínimas diarias



Periodo: 02-01-2019 - 30-04-2022.
Cantidad de datos: 1200. Porcentaje de datos faltantes: 1.23%.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas mínimas diarias

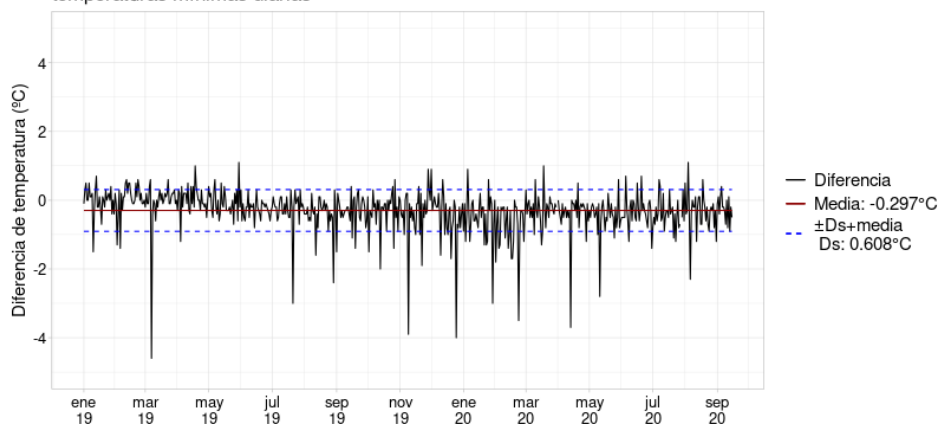


Periodo: 02-01-2019 - 30-04-2022.
Cantidad de datos: 1200. Porcentaje de datos faltantes: 1.23%.

Figura 6: Ídem Figura 5 pero para el par de estaciones ubicadas en Pigüé.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

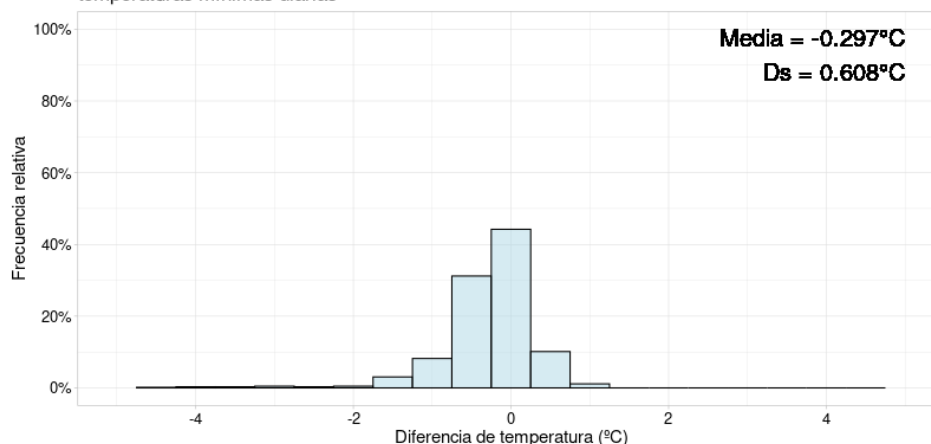
Serie de diferencias entre datos convencionales y automáticos de temperaturas mínimas diarias



Periodo: 01-01-2019 - 15-09-2020.
Cantidad de datos: 622. Porcentaje de datos faltantes: 0.32%.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas mínimas diarias



Periodo: 01-01-2019 - 15-09-2020.
Cantidad de datos: 622. Porcentaje de datos faltantes: 0.32%.

Figura 7: Ídem Figura 5 pero para el par de estaciones ubicadas en Aeroparque.

Para la temperatura mínima, la menor diferencia media se presenta en Villa Gesell (Figura 5), mientras que para la temperatura máxima, en Pigüé (no se muestra). En las tres ubicaciones se encuentran diferencias medias menores a los 0.3°C en valor absoluto para la primera variable y de 0.5°C para la segunda. El mayor desvío en las series de la temperatura mínima se registra en Villa Gesell, con 0.639°C, seguido por Aeroparque, con 0.608°C (Figura 7). En cuanto al comportamiento de las distribuciones de las frecuencias, los intervalos modales son en torno a los 0°C (Figuras 5, 6 y 7 der.) para todos los sitios. En el caso de Pigüé y Aeroparque, esto también se observa para las temperaturas máximas diarias.

En la Tabla III se presentan distintas medidas estadísticas de posición (la mediana, el primer y tercer cuartil, y los valores mínimos y máximos) que describen las distribuciones de temperatura mínima diaria obtenidas manual y automáticamente para los pares de estaciones analizados. En la Tabla IV se muestran los mismos resultados pero para la temperatura máxima diaria.

Tabla III: Valores del mínimo, primer cuartil, mediana, tercer cuartil y máximo de las distribuciones de la temperatura mínima diaria para los pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque.

Nombre de la estación	Tipo de estación	Mínimo (°C)	Primer cuartil (°C)	Mediana (°C)	Tercer cuartil (°C)	Máximo (°C)
Villa Gesell Aero	EMC	-1.6	5.0	9.0	13.2	21.0
Villa Gesell	EMA	-1.3	5.1	8.9	13.2	21.1
Pigüé Aero	EMC	-5.9	3.4	8.5	13.1	24.3
RMA10 Pigüé	EMA	-6.5	3.8	8.7	13.2	24.5
Aeroparque Buenos Aires	EMC	2.2	10.6	14.2	18.6	26.9
Aeroparque	EMA	2.4	10.9	14.6	19.1	26.4

Tabla IV: Ídem Tabla III pero para las distribuciones de la temperatura máxima diaria.

Nombre de la estación	Tipo de estación	Mínimo (°C)	Primer cuartil (°C)	Mediana (°C)	Tercer cuartil (°C)	Máximo (°C)
Villa Gesell Aero	EMC	8.7	15.5	20.0	24.0	42.3
Villa Gesell	EMA	9.5	15.6	20.6	24.3	41.9
Pigüé Aero	EMC	5.0	15.8	21.5	26.8	38.9
RMA10 Pigüé	EMA	5.3	16.0	21.7	27.0	39.5
Aeroparque Buenos Aires	EMC	8.6	16.7	21.5	26.3	37.5
Aeroparque	EMA	10.3	17.2	21.9	26.8	37.4

De las Tablas III y IV se observa que, salvo para el mínimo en la temperatura máxima diaria de Aeroparque, todas las diferencias de las medidas de posición analizadas son menores a los 1°C en módulo. Más aún, en la mayoría de los casos, las diferencias son inferiores al medio grado y negativas, lo que indicaría que, en general, hay sobreestimación de los valores por parte de las EMAs. Las máximas diferencias dependen de la ubicación y la variable.

En la Tabla V se encuentran los valores de RMSE resultantes de comparar las series manuales y las automáticas de temperatura mínima y máxima diaria para las tres ubicaciones analizadas, considerando el año completo y discretizando la información en las cuatro estaciones del año. En la Tabla VI se presentan los RMSE centrados.

Tabla V: Valores de RMSE entre las mediciones convencionales y automáticas de temperatura mínima y máxima diaria para los pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque, considerando todo el año completo y las cuatro estaciones del año.

RMSE (°C)	Temperatura mínima diaria			Temperatura máxima diaria		
	Villa Gesell	Pigüé	Aeroparque	Villa Gesell	Pigüé	Aeroparque
Año completo	0.638	0.333	0.676	0.516	0.323	0.927
DEF	0.989	0.338	0.806	0.610	0.386	1.082
MAM	0.392	0.317	0.642	0.507	0.319	0.781
JJA	0.368	0.329	0.582	0.473	0.299	0.740
SON	0.604	0.351	0.683	0.463	0.267	1.179

Tabla VI: Ídem Tabla V pero para los valores de RMSE centrado.

RMSE centrado (°C)	Temperatura mínima diaria			Temperatura máxima diaria		
	Villa Gesell	Pigüé	Aeroparque	Villa Gesell	Pigüé	Aeroparque
Año completo	0.638	0.275	0.607	0.406	0.275	0.811
DEF	0.987	0.291	0.725	0.448	0.315	0.936
MAM	0.388	0.254	0.606	0.398	0.276	0.697
JJA	0.368	0.285	0.484	0.369	0.270	0.625
SON	0.603	0.267	0.600	0.387	0.211	1.031

En casi la totalidad de los casos se observan los mínimos de RMSE y de RMSE centrado en Pigüé, y los máximos, en Aeroparque (Tablas V y VI). Comparando los resultados por variables y para Villa Gesell, se encuentra que los menores valores de las métricas corresponden a la temperatura máxima diaria, tanto para el año completo como para los trimestres DEF y SON. Lo mismo ocurre para Pigüé en el año completo, JJA y SON. Esto estaría indicando un menor error promedio entre los registros de temperatura máxima que entre las series de temperatura mínima en dichas temporadas. Para Aeroparque ocurre lo contrario: los menores RMSE y RMSE centrado tienen lugar en la temperatura mínima diaria en el año completo y en su desagregado estacional.

En cuanto al coeficiente de correlación de Rank-Kendall, su rango se extiende desde 0.83 hasta 0.98. Comparando los resultados por ubicaciones, los mayores valores se obtienen en Pigüé y los menores, en Aeroparque, para ambas variables. En Pigüé, la temperatura máxima presenta las mayores correlaciones para todos los trimestres y el año completo. Lo mismo sucede en Villa Gesell, a excepción de MAM. En Aeroparque, las correlaciones más altas se advierten en la temperatura mínima diaria. Ninguna estación del año presenta los mejores resultados en forma unificada. Los coeficientes de correlación de Rank-Kendall no se muestran.

De las pruebas de t-Student y de Wilcoxon se puede concluir que, a excepción de la temperatura mínima diaria de Villa Gesell, las medias de las temperaturas extremas diarias obtenidas de las EMCs y las EMAs son significativamente diferentes, con un 95% de confianza. Esto se observa para el año completo y las cuatro estaciones del año. Para todos los casos, sin embargo, los p-valores de las pruebas de Kolmogorov-Smirnov son mayores al umbral de 0.05. No se pueden rechazar las hipótesis nulas y, por lo tanto, no existen evidencias que permitan establecer que las temperaturas extremas diarias automáticas y

convencionales no pertenecen a la misma distribución con un 95% de confianza. Los resultados de las pruebas de t-Student, de Wilcoxon y de Kolmogorov-Smirnov no se muestran.

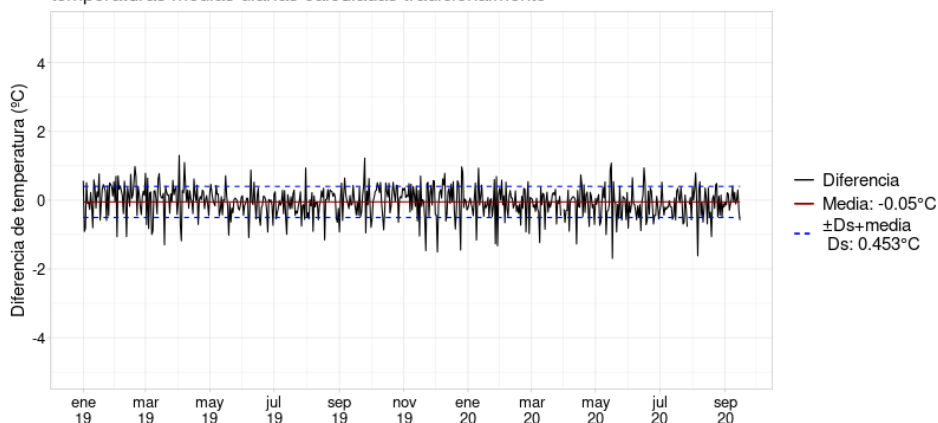
Con respecto a las diferencias de temperatura mínima diaria mayores a 5°C que resultaron excluidas, en Pigüé y Aeroparque no se registraron ningún valor. En Villa Gesell se relevaron dos valores. En esta última ubicación se observó la máxima diferencia de temperatura mínima el día 09/09/2021; fue de 15.1°C aproximadamente. Se debió a un dato automático erróneo. Para las temperaturas máximas diarias, se removieron dos datos de Pigüé con diferencias levemente superiores a los 5°C, uno en Villa Gesell y cuatro en Aeroparque.

5.3. Cálculo, comparación y caracterización estadística de las series de temperatura media diaria

En la Figuras 8 se presenta la serie temporal y el histograma de las diferencias entre los datos manuales y automáticos de temperatura media diaria obtenidas a partir del procedimiento tradicional para el par de estaciones ubicadas en Aeroparque. Se muestran también las diferencias medias y los desvíos estándares. Nuevamente, se trabaja con diferencias que oscilan entre $\pm 5^\circ\text{C}$. Para las EMCs de Villa Gesell y Pigüé no fue posible calcular la temperatura media diaria utilizando la Ecuación 1 porque no se cuentan con las mediciones en los cuatro tiempos de observación, sólo en tres. En estos casos, sería conveniente hacer la comparación de la temperatura media a nivel mensual. Sin embargo, muchas veces, el corto período de superposición de las observaciones hace que los análisis mensuales no sean significativos estadísticamente.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Serie de diferencias entre datos convencionales y automáticos de temperaturas medias diarias calculadas tradicionalmente

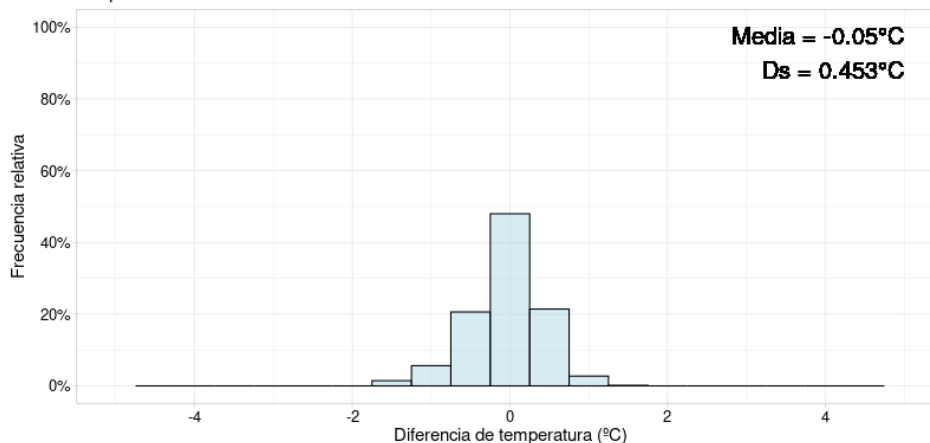


Período: 01-01-2019 - 15-09-2020.
Cantidad de datos: 621. Porcentaje de datos faltantes: 0.48%.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)



Histograma de diferencias entre datos convencionales y automáticos de temperaturas medias diarias calculadas tradicionalmente



Período: 01-01-2019 - 15-09-2020.
Cantidad de datos: 621. Porcentaje de datos faltantes: 0.48%.

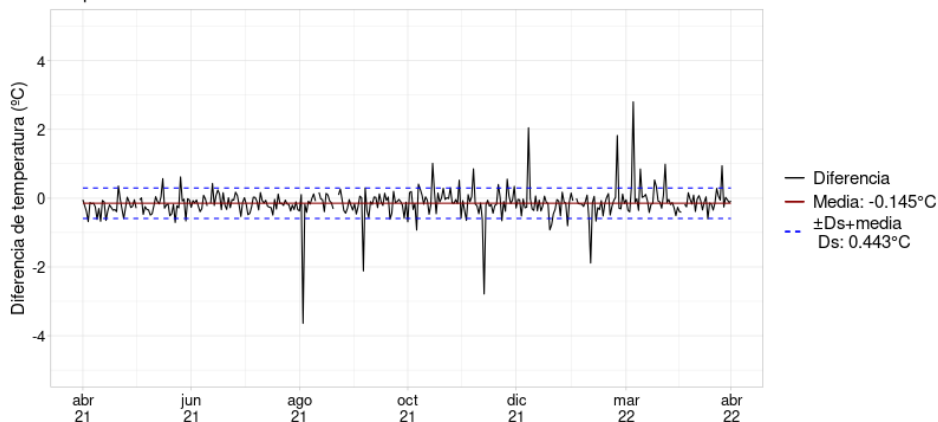
Figura 8: Arriba: Serie temporal de las diferencias entre las mediciones convencionales y automáticas de temperatura media diaria (línea negra) para el par de estaciones ubicadas en Aeroparque. La variable se obtuvo mediante el procedimiento tradicional. En línea roja se representa la diferencia media y en líneas azules, los desvíos estándares sumados a esta. **Abajo:** Histograma de las diferencias entre las mediciones convencionales y automáticas de temperatura media diaria.

En Aeroparque se observan bajos valores de media y de desvío estándar en la serie de las diferencias de temperatura media diaria calculada tradicionalmente (Figura 8 arriba), de -0.05°C y 0.453°C , respectivamente. Esto indica que, en promedio, el cambio de sensor y de procedimiento de cálculo para la temperatura media no genera marcadas diferencias en la temperatura media diaria de dicha locación. La distribución de la variable presenta una forma aproximadamente simétrica, con un marcado intervalo modal en torno a los 0°C (Figura 8 abajo).

En las Figuras 9, 10 y 11 se muestran las series temporales y los histogramas de las diferencias entre los datos tradicionales y automáticos de temperatura media diaria calculada a partir de los extremos diarios y la Ecuación 3, para los tres pares de estaciones considerados. Se muestran también las diferencias medias y los desvíos estándares. En todos los casos se trabaja con diferencias que oscilan entre $\pm 5^{\circ}\text{C}$.

Estaciones Villa Gesell Aero (EMC) y Villa Gesell SMN (EMA)

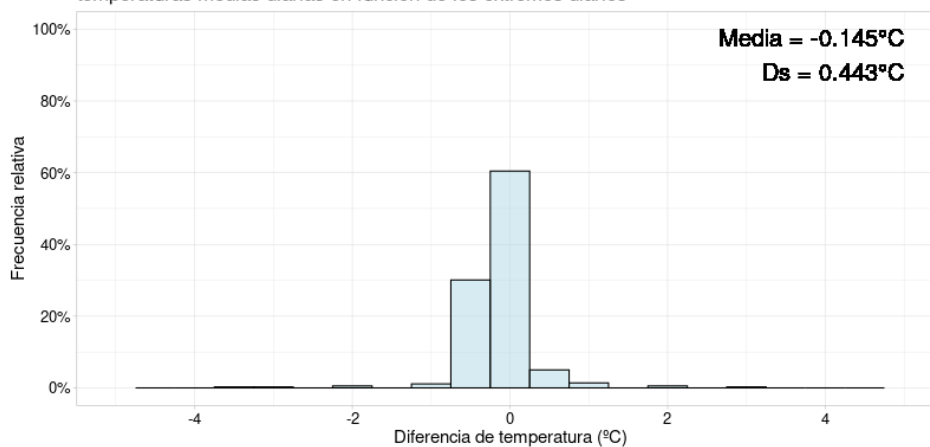
Serie de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



Periodo: 30-04-2021 - 30-04-2022.
Cantidad de datos: 359. Porcentaje de datos faltantes: 1.91%.

Estaciones Villa Gesell Aero (EMC) y Villa Gesell SMN (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios

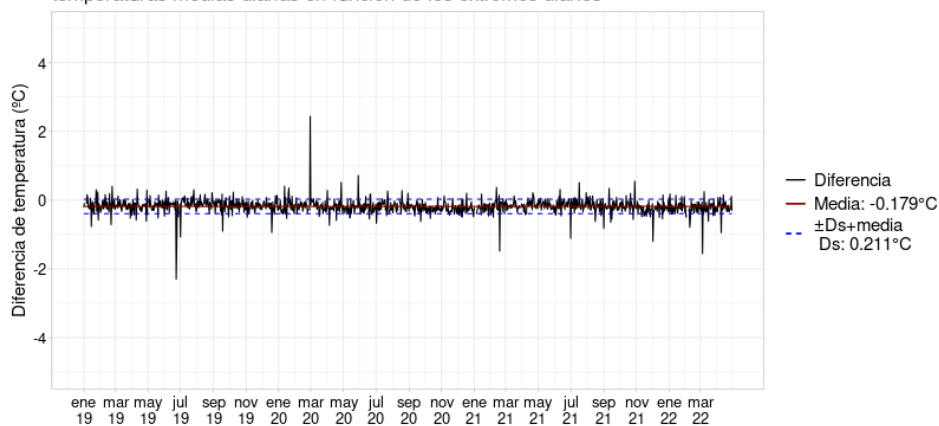


Periodo: 30-04-2021 - 30-04-2022.
Cantidad de datos: 359. Porcentaje de datos faltantes: 1.91%.

Figura 9: Arriba: Serie temporal de las diferencias entre las mediciones convencionales y automáticas de temperatura media diaria (línea negra) para el par de estaciones ubicadas en Villa Gesell. La variable se obtuvo mediante el procedimiento basado en los extremos diarios. En línea roja se representa la diferencia media y en líneas azules, los desvíos estándares sumados a esta. **Abajo:** Histograma de las diferencias entre las mediciones convencionales y automáticas de temperatura media diaria.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

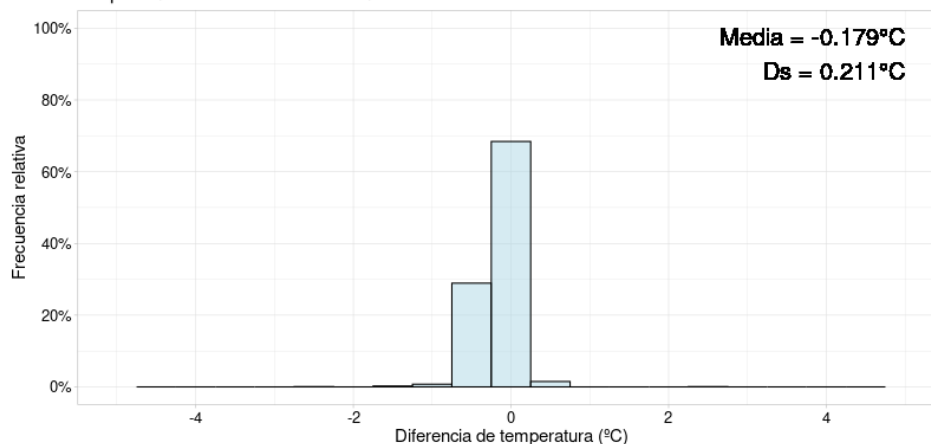
Serie de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



Periodo: 02-01-2019 - 30-04-2022.
Cantidad de datos: 1200. Porcentaje de datos faltantes: 1.23%.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios

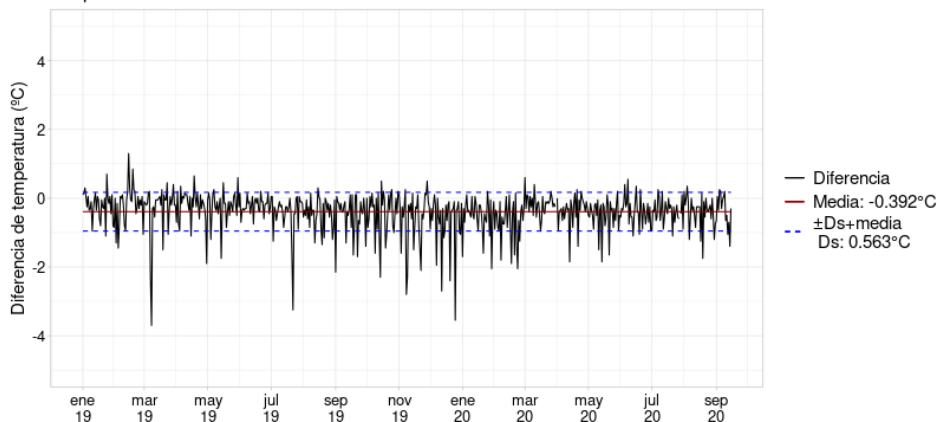


Periodo: 02-01-2019 - 30-04-2022.
Cantidad de datos: 1200. Porcentaje de datos faltantes: 1.23%.

Figura 10: Ídem Figura 9 pero para el par de estaciones ubicadas en Pigüé.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

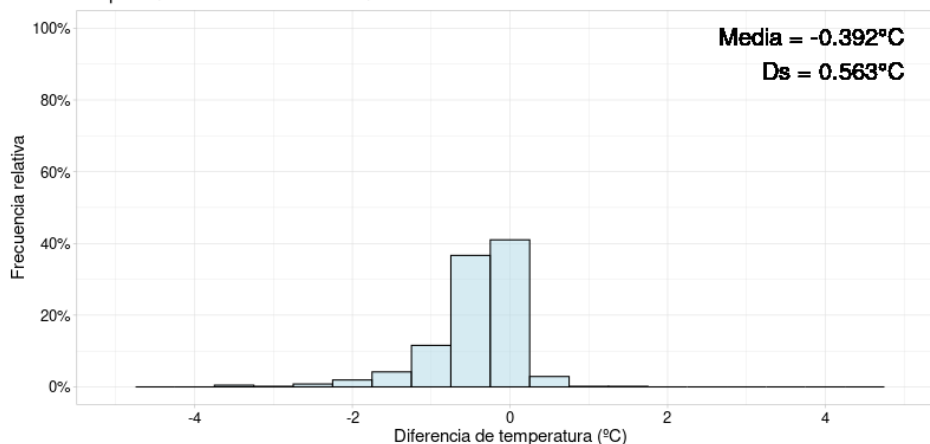
Serie de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



Periodo: 01-01-2019 - 15-09-2020.
Cantidad de datos: 622. Porcentaje de datos faltantes: 0.32%.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Histograma de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



Periodo: 01-01-2019 - 15-09-2020.
Cantidad de datos: 622. Porcentaje de datos faltantes: 0.32%.

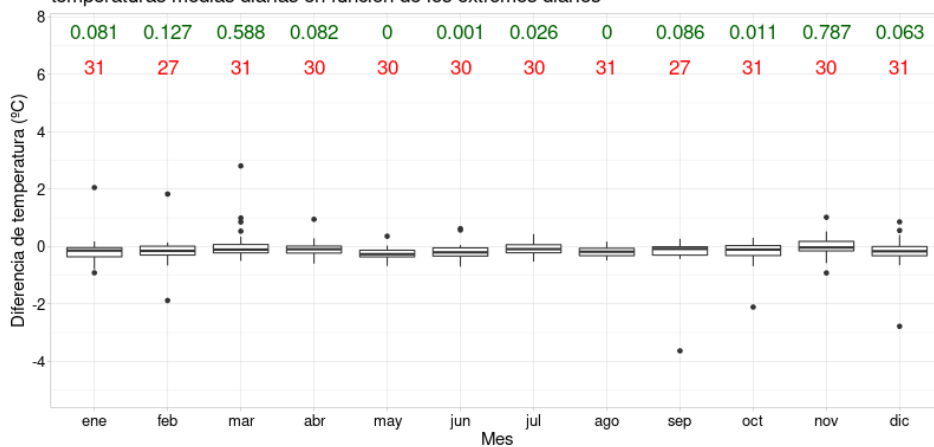
Figura 11: Ídem Figura 9 pero para el par de estaciones ubicadas en Aeroparque.

En Villa Gesell, Pigüé y Aeroparque, las medias de las diferencias de temperatura media diaria calculada en base a los extremos diarios resultaron negativas (Figuras 9, 10 y 11). Esto indica que, en promedio, los datos automáticos sobreestiman a los convencionales. Tanto la mayor dispersión de los datos, traducida en el mayor valor de desvío estándar, como la más alta diferencia de valor medio se observan en Aeroparque. En dicha ubicación, la distribución de las diferencias presenta una forma asimétrica hacia la izquierda (Figura 11 abajo).

En la Figura 12 se presentan los boxplots de las diferencias de temperatura media diaria calculada a partir de los valores extremos diarios para las EMCs y EMAs de Villa Gesell, Pigüé y Aeroparque, junto con los p-valores del test t-Student de igualdad de medias para cada uno de los meses del año. Para Aeroparque también se muestran los resultados de la temperatura media surgida del procedimiento tradicional. El desagregado mensual permite analizar si existe un ciclo anual presente en las series de dichas diferencias que tal vez no se detecte en el promedio total.

Estaciones Villa Gesell Aero (EMC) y Villa Gesell SMN (EMA)

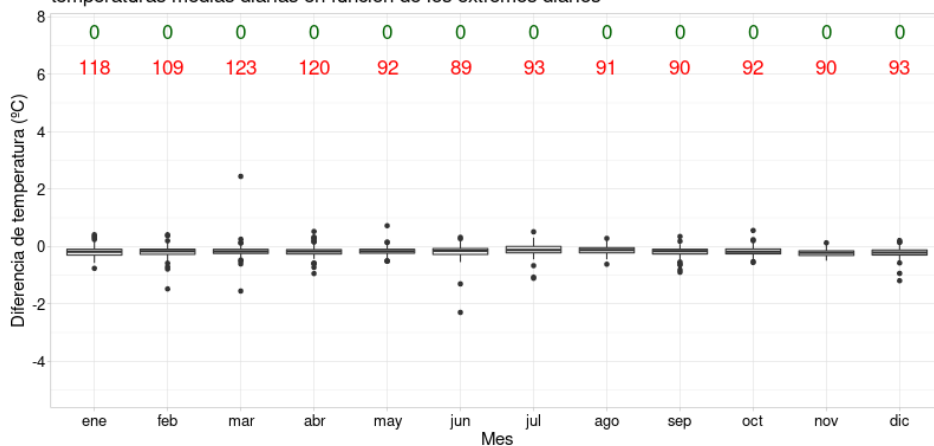
Boxplots mensuales de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



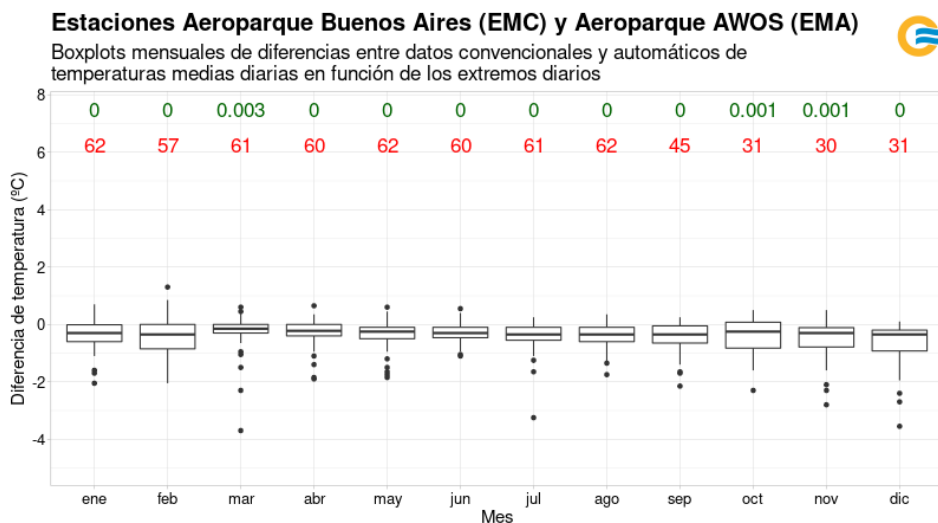
Período: 30-04-2021 - 30-04-2022.
Cantidad de datos: 359. Porcentaje de datos faltantes: 1.91%.

Estaciones Pigüé Aero (EMC) y Pigüé SINARAME (EMA)

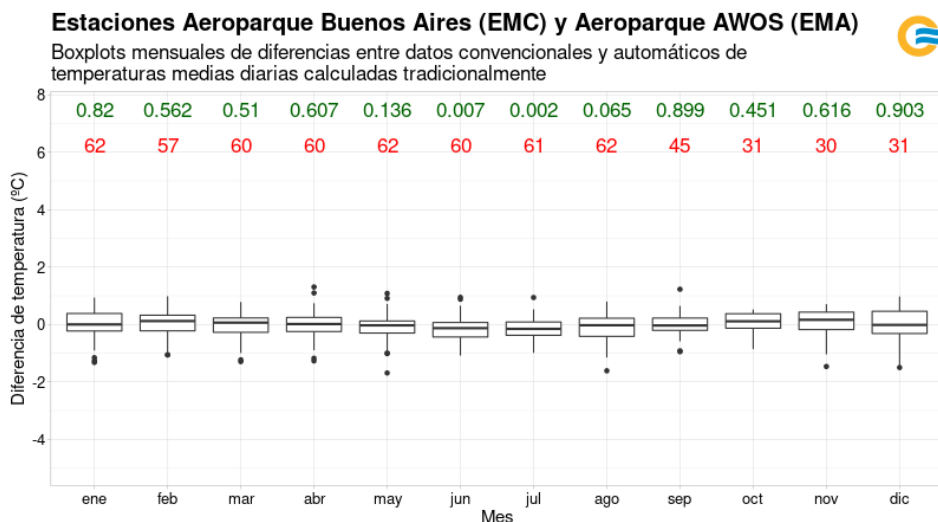
Boxplots mensuales de diferencias entre datos convencionales y automáticos de temperaturas medias diarias en función de los extremos diarios



Período: 02-01-2019 - 30-04-2022.
Cantidad de datos: 1200. Porcentaje de datos faltantes: 1.23%.



Período: 01-01-2019 - 15-09-2020.
Cantidad de datos: 622. Porcentaje de datos faltantes: 0.32%.



Período: 01-01-2019 - 15-09-2020.
Cantidad de datos: 621. Porcentaje de datos faltantes: 0.48%.

Figura 10: Boxplots de las diferencias entre las mediciones convencionales y automáticas de temperatura media diaria para cada mes del año y el par de estaciones ubicadas en Villa Gesell (**arriba**), Pigüé (**centro arriba**) y Aeroparque (**centro abajo**). La variable se obtuvo mediante el procedimiento basado en los valores extremos diarios. Para Aeroparque, también se consideró el procedimiento tradicional (**abajo**). En la parte superior, los números en verde corresponden a los p-valores del test t-Student para la comparación de medias, mientras que los números en rojo, a la cantidad de datos dentro de cada mes.

En Villa Gesell se observan p-valores mayores a 0.05 (Figura 10 arriba) en septiembre y de noviembre a abril, por lo que, para esos meses, no existen indicios que posibiliten establecer la desigualdad de medias de las temperaturas medias diarias de las EMCs y EMAs. De mayo a agosto se obtienen p-valores menores a 0.05, lo mismo que en Pigüé y Aeroparque en todos los meses del año (Figura 10 centro arriba y centro abajo). Esto permite inferir que las medias de la variable convencional y automática son significativamente diferentes, con un 95% de confianza. Cabe destacar que aquí se están analizando las temperaturas diarias calculadas a partir de la Ecuación 3. Si se considera el procedimiento tradicional, los resultados de Aeroparque revelan p-valores mayores a 0.05 y, por lo tanto, no hay evidencias para concluir la

desigualdad de medias a lo largo de todo año, a excepción de los meses junio y julio, con un 95% de confianza (Figura 10 abajo).

En la Tabla VII se presentan la mediana, el primer y tercer cuartil, y los valores mínimos y máximos, medidas estadísticas que permiten caracterizar las distribuciones de la temperatura media diaria registradas convencional y automáticamente para los pares de estaciones analizados. Dicha variable se obtuvo mediante el procedimiento basado en los valores extremos diarios.

Tabla VII: Valores del mínimo, primer cuartil, mediana, tercer cuartil y máximo de las distribuciones de la temperatura media diaria para los pares de estaciones ubicadas en Villa Gesell, Pigué y Aeroparque. La variable se obtuvo mediante el procedimiento basado en los valores extremos diarios.

Nombre de la estación	Tipo de estación	Mínimo (°C)	Primer cuartil (°C)	Mediana (°C)	Tercer cuartil (°C)	Máximo (°C)
Villa Gesell Aero	EMC	4.4	10.5	14.2	18.5	30.9
Villa Gesell	EMA	4.8	10.5	14.2	18.6	31.1
Pigué Aero	EMC	1	9.8	15.1	19.6	29.9
RMA10 Pigué	EMA	0.9	10.0	15.2	19.7	30.3
Aeroparque Buenos Aires	EMC	6.1	13.8	17.9	22.6	30.5
Aeroparque	EMA	6.5	14.0	18.2	23.0	30.7

Para la temperatura media diaria, todas las medidas de posición presentan diferencias menores a los 0.5°C en módulo (Tabla VII). Análogo a lo observado en las temperaturas extremas diarias, en general, los mayores valores de las medidas de posición se observan en las EMAs.

En la Tabla VIII se presentan los valores de RMSE resultantes de comparar las series manuales y automáticas de temperatura media diaria calculada, siempre que sea posible, a partir de los cuatro procedimientos: el tradicional y los basados en los cuatro tiempos de observación, en los extremos diarios y en los 24 valores horarios. Se considera la información del año completo y discretizada en las cuatro estaciones del año. En las Tablas IX se muestran los RMSE centrados para todos los casos. Cabe destacar que la EMC de Aeroparque es la única de las analizadas que mide no sólo en los cuatro tiempos de observación, sino también en las 24 h. Sólo en ella fue posible calcular la temperatura media diaria mediante todos los procedimientos considerados.

De entre los tres pares de estaciones analizadas y con excepción del trimestre JJA, en Pigué se observan los mejores resultados en cuanto al RMSE y al RMSE centrado para la temperatura media diaria calculada a partir de los valores extremos (Tablas VIII y IX). En Aeroparque, la variable definida mediante el procedimiento de 24 h evidencia los valores más bajos de ambas métricas, seguido por la variable derivada del método tradicional. En general, los peores resultados se obtienen con el procedimiento basado en los extremos diarios.

Tabla VIII: Valores de RMSE entre las mediciones convencionales y automáticas de temperatura media diaria para los pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque, considerando todo el año completo y las cuatro estaciones del año. Las temperaturas medias diarias se calculan, siempre que sea posible, a partir de los cuatro procedimientos: el tradicional y los basados en los cuatro tiempos de observación, en los extremos diarios y en los 24 valores horarios.

RMSE (°C)	Temperatura media diaria					
	A partir del procedimiento tradicional	A partir de los tiempos de observación	A partir de los valores extremos diarios			A partir de los 24 valores horarios
	Aeroparque	Aeroparque	Villa Gesell	Pigüé	Aeroparque	Aeroparque
Año completo	0.455	0.551	0.466	0.277	0.685	0.323
DEF	0.500	0.597	0.563	0.283	0.794	0.371
MAM	0.461	0.562	0.437	0.284	0.608	0.317
JJA	0.426	0.498	0.276	0.283	0.580	0.303
SON	0.428	0.552	0.536	0.255	0.804	0.295

Tabla IX: Ídem Tabla VIII pero para los valores de RMSE centrado.

RMSE centrado (°C)	Temperatura media diaria					
	A partir del procedimiento tradicional	A partir de los tiempos de observación	A partir de los valores extremos diarios			A partir de los 24 valores horarios
	Aeroparque	Aeroparque	Villa Gesell	Pigüé	Aeroparque	Aeroparque
Año completo	0.453	0.547	0.442	0.211	0.563	0.318
DEF	0.500	0.597	0.535	0.202	0.643	0.371
MAM	0.458	0.553	0.426	0.229	0.529	0.314
JJA	0.404	0.475	0.229	0.237	0.441	0.284
SON	0.427	0.549	0.513	0.163	0.651	0.288

El coeficiente de correlación de Rank-Kendall abarca el rango de 0.88 a 0.98. Para la temperatura media diaria basada en los extremos, los mayores coeficientes se obtienen en Pigüé y los menores, en Aeroparque, análogo a lo ocurrido en las temperaturas extremas diarias. En esta última ubicación, sin embargo, se obtienen más altos valores de correlación si se consideran el procedimiento tradicional y el basado en los 24 valores horarios. Los resultados de los coeficientes de correlación de Rank-Kendall no se muestran.

Tabla X: p-valores del test t-Student aplicado a las mediciones convencionales y automáticas de temperatura media diaria para el par de estaciones ubicadas en Villa Gesell, considerando todo el año completo y las cuatro estaciones del año. Las temperaturas medias diarias se calculan, siempre que sea posible, a partir de los cuatro procedimientos: el tradicional y los basados en los cuatro tiempos de observación, en los extremos diarios y en los 24 valores horarios. En negrita se marcan los p-valores menores al umbral de significancia de 0.05.

p-valor del test t-Student	Temperatura media diaria					
	A partir del procedimiento tradicional	A partir de los tiempos de observación	A partir de los valores extremos diarios			A partir de los 24 valores horarios
	Aeroparque	Aeroparque	Villa Gesell	Pigüé	Aeroparque	Aeroparque
Año completo	0.004	0.003	0	0	0	0
DEF	0.890	0.997	0.006	0	0	0.437
MAM	0.084	0.014	0.068	0	0	0.058
JJA	0	0	0	0	0	0
SON	0.508	0.269	0.004	0	0	0.030

Tabla XI: Ídem Tabla X pero para el test de Wilcoxon.

p-valor del test de Wilcoxon	Temperatura media diaria					
	A partir del procedimiento tradicional	A partir de los tiempos de observación	A partir de los valores extremos diarios			A partir de los 24 valores horarios
	Aeroparque	Aeroparque	Villa Gesell	Pigüé	Aeroparque	Aeroparque
Año completo	0.042	0.026	0	0	0	0
DEF	0.452	0.647	0	0	0	0.797
MAM	0.229	0.040	0	0	0	0.228
JJA	0	0	0	0	0	0
SON	0.153	0.034	0.001	0	0	0.199

En las Tablas X y XI se muestran los p-valores de los tests t-Student y de Wilcoxon, respectivamente, resultantes de comparar las medias de las series manuales y automáticas de temperatura media diaria para el par de estaciones ubicado en Villa Gesell, Pigüé y Aeroparque. Se considera la información del año completo y discretizada en las cuatro estaciones del año. Nuevamente, sólo para Aeroparque fue posible calcular las temperaturas medias diarias mediante todos los procedimientos considerados. Para el resto de las ubicaciones, estas se obtuvieron a partir de los extremos diarios.

Para las tres ubicaciones, el año completo y todos los trimestres, a excepción de MAM para Villa Gesell, las medias de las temperaturas medias diarias de las EMCs y las EMAs calculadas mediante los extremos diarios son significativamente diferentes, con un 95% de confianza. Esto se pudo concluir de las pruebas de t-Student y de Wilcoxon (Tablas X y XI). Para las temperaturas medias diarias computadas tradicionalmente y mediante los horarios de observación y las 24 h de Aeroparque se obtienen los mismos

resultados, tanto para el año completo como el período JJA. En cambio, para DEF, los p-valores resultantes de ambos tests sí son mayores al umbral de 0.05. Por lo tanto, no se pueden rechazar las hipótesis nulas de los tests con un 95% de confianza. No existen evidencias que permitan establecer la desigualdad de las medias en estos casos. Lo mismo ocurre para las temperaturas medias diarias calculadas a partir del procedimiento tradicional en MAM y SON, y de los 24 datos horarios en MAM, según ambos tests. Los diferentes resultados encontrados en el año completo y su desagregado estacional y mensual (Figura X) reflejan la necesidad de realizar los análisis en escalas intranuales.

En cuanto a la comparación de las distribuciones, los p-valores de las pruebas de Kolmogorov-Smirnov son mayores al umbral de 0.05 para todas las ubicaciones, para el año completo y las cuatro estaciones del año, y para todos los procedimientos de cálculo de las temperaturas medias diarias. No se pueden rechazar las hipótesis nulas con un 95% de confianza, por lo que no habrían indicios para concluir que las temperaturas medias diarias automáticas y convencionales no provienen de la misma distribución. Los resultados de la prueba de Kolmogorov-Smirnov no se muestran.

Con respecto a las diferencias de temperatura media diaria mayores a 5°C que resultaron excluidas, sólo se registró un valor basado en los extremos diarios en Villa Gesell, en el día 09/09/2021, y fue de 7.6°C. Esta diferencia está explicada por el dato automático erróneo de temperatura mínima.

Cabe destacar que, en casi todos los test aplicados de las secciones 5.2 y 5.3, se encontraron los mismos resultados al considerar las autocorrelaciones en las series diarias de las diferencias que sin contemplarlos.

5.4. Detección de tendencias y ciclos en las series de las diferencias de temperatura máxima, mínima y media diaria

En las Tablas XII, XIII y XIV se presentan los p-valores del test de Mann-Kendall aplicado en las series de las diferencias entre las temperaturas mínimas, máximas y medias diarias obtenidas convencional y automáticamente para el par de estaciones ubicado en Villa Gesell, Pigüé y Aeroparque, respectivamente. Se considera la información del año completo y discretizada en las cuatro estaciones del año. Las temperaturas medias diarias se calculan, siempre que sea posible, a partir de los cuatro procedimientos.

Para la ubicación Villa Gesell y el test de Mann-Kendall (Tabla XII) se encuentran p-valores menores al umbral de 0.05 en el año completo, tanto para la temperatura mínima como para la temperatura media diaria en función de los extremos. En este último caso también se hallan p-valores inferiores a 0.05 en el test de Fisher (no se muestran). Por lo tanto, se pueden rechazar las hipótesis nulas de ambos tests con un 95% de confianza y concluir que la serie de las diferencias presenta componentes determinísticas, con una tendencia lineal significativa. Lo mismo ocurre en Pigüé (Tabla XIII sólo para el test de Mann-Kendall) con las series de las diferencias de temperatura máxima y media en función de los extremos para el año completo y los trimestres DEF y SON. Situaciones análogas se observan en Aeroparque (Tabla XIV sólo para el test de Mann-Kendall), en las series correspondientes a las temperaturas mínimas y medias diarias calculadas a partir de los procedimientos basados en los extremos diarios y en los 24 datos horarios para el año completo y el trimestre DEF. Para el procedimiento tradicional y el que considera los tiempos de observación, se pueden establecer las mismas conclusiones pero sólo para el año completo. Se encuentran p-valores menores a 0.05 sólo en el test de Mann-Kendall en la serie de la temperatura mínima diaria de Aeroparque de JJA. Lo análogo se evidencia en las temperaturas medias diarias basadas en los extremos y en el procedimiento tradicional en MAM y DEF, respectivamente. Se puede concluir, entonces, que estas series no son al azar con un 95% de confianza.

Tabla XII: p-valores del test de Mann-Kendall aplicado a las diferencias entre las mediciones convencionales y automáticas de temperatura mínima, máxima y media diaria para el par de estaciones ubicadas en Villa Gesell, considerando todo el año completo y las cuatro estaciones del año. Las temperaturas medias diarias se calculan, siempre que sea posible, a partir de los cuatro procedimientos: el tradicional y los basados en los cuatro tiempos de observación, en los extremos diarios y en los 24 valores horarios. En negrita se marcan los p-valores menores al umbral de significancia de 0.05. Con asterisco se señalan los p-valores del test de Fisher también inferiores a dicho umbral.

p-valor del test de Mann-Kendall	Temperatura mínima diaria	Temperatura máxima diaria	Temperatura media diaria
			A partir de los valores extremos diarios
Año completo	0.001	0.767	0.030*
DEF	0.063	0.188	0.849
MAM	0.004*	0.116	0.017
JJA	0.413	0.837	0.741
SON	0.050	0.585	0.238

Tabla XIII: Ídem Tabla XII pero para el par de estaciones ubicadas en Pigüé.

p-valor del test de Mann-Kendall	Temperatura mínima diaria	Temperatura máxima diaria	Temperatura media diaria
			A partir de los valores extremos diarios
Año completo	0.095	0*	0*
DEF	0.797	0.002*	0.010*
MAM	0.807	0.777	0.455
JJA	0.236	0.993	0.595
SON	0.222	0.001*	0.001*

Tabla XIV: Ídem Tabla XII pero para el par de estaciones ubicadas en Aeroparque.

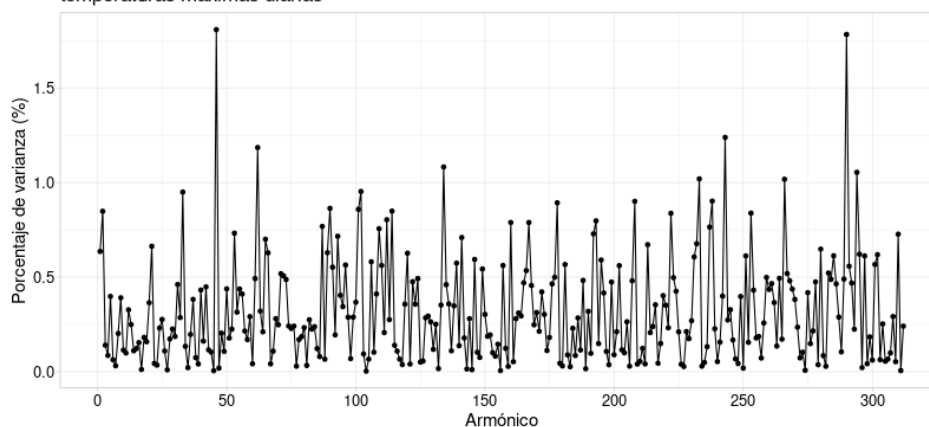
p-valor del test de Mann-Kendall	Temperatura mínima diaria	Temperatura máxima diaria	Temperatura media diaria			
			A partir del procedimiento tradicional	A partir de los tiempos de observación	A partir de los valores extremos diarios	A partir de los 24 valores horarios
Año completo	0*	0.918	0*	0.008*	0*	0*
DEF	0*	0.995	0.032	0.100	0*	0.002*
MAM	0*	0.779	0.006*	0.157	0.002	0*
JJA	0.045	0.363	0.938	0.809	0.129	0.551
SON	0.735	0.960	0.805	0.505	0.938	0.923

Para analizar la presencia y la significancia de ciclos inmersos en la serie de las diferencias se utilizan los análisis armónico y espectral. En la Figura 11 se presentan los resultados para la temperatura máxima diaria de Aeroparque. Dado que estas metodologías requieren la ausencia de datos faltantes y que existen pocos de estos, se puede optar por completar la serie usando el promedio de dos valores contiguos.

A partir del análisis armónico se observan dos ciclos que explican una mayor varianza que el resto, de aproximadamente un 1.8% de la varianza total de la serie de las diferencias de la temperatura máxima diaria (Figura 11 arriba). Sin embargo, el análisis espectral permite concluir que dichos ciclos no son significativos, dado que la densidad espectral asociada a ellos no supera la banda de significancia superior del modelo de ruido rojo (Figura 11 abajo). Cabe destacar que se utiliza el modelo de ruido rojo en función del resultado del test de Anderson. La tendencia lineal es no significativa según el test de Fisher, por lo que no es filtrada.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Análisis armónico de serie de diferencias entre datos convencionales y automáticos de temperaturas máximas diarias



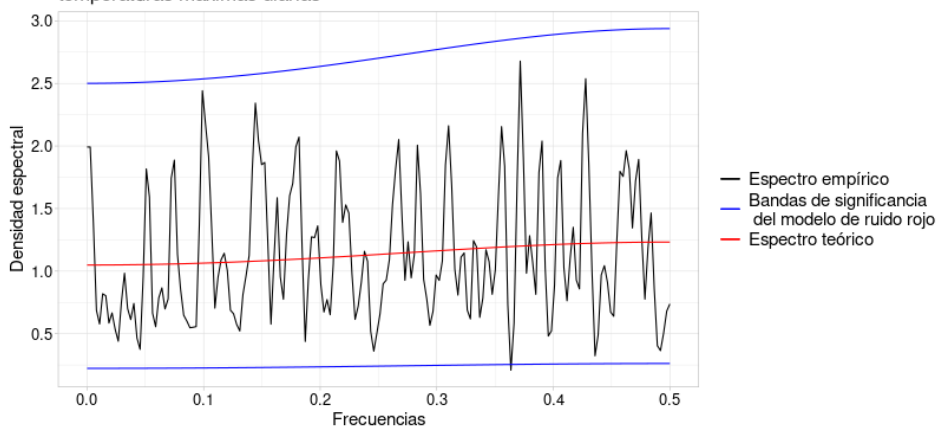
Período: 01-01-2019 - 15-09-2020.

Cantidad de datos: 618. Porcentaje de datos faltantes: 0.96%.

Datos faltantes completados para los análisis armónico y espectral.

Estaciones Aeroparque Buenos Aires (EMC) y Aeroparque AWOS (EMA)

Análisis espectral de serie de diferencias entre datos convencionales y automáticos de temperaturas máximas diarias



Período: 01-01-2019 - 15-09-2020.
Cantidad de datos: 618. Porcentaje de datos faltantes: 0,96%.
Datos faltantes completados para los análisis armónico y espectral.

Figura 11: *Arriba:* Análisis armónico de la serie de las diferencias entre las mediciones convencionales y automáticas de temperatura máxima diaria para el par de estaciones ubicadas en Aeroparque. *Abajo:* Análisis espectral de la misma serie considerando un modelo de ruido rojo.

5.5. Comparación estacional de frecuencia de eventos

En las Tablas XV y XVI se presentan los resultados del análisis de eventos de temperatura mínima y máxima diaria, respectivamente, por estación del año y la ubicación Pigüé. La elección de esta ubicación es porque es la que presenta la mayor cantidad de datos analizada para la comparación.

Tabla XV: Número de eventos de temperatura mínima diaria dentro de las categorías o clases “frío extremo” (EX_F), “frío” (F), “medio” (M), “cálido” (C) y “cálido extremo” (EX_C), en el par de estaciones ubicadas en Pigüé y para cada estación del año.

Número de eventos de temperatura mínima diaria	Eventos en	EX_F	F	M	C	EX_C
DEF	EMC	17	52	190	46	15
	EMA	15	44	194	51	16
MAM	EMC	18	51	201	48	17
	EMA	16	49	202	51	17
JJA	EMC	15	42	167	35	14
	EMA	13	40	162	44	14
SON	EMC	15	42	167	36	12
	EMA	13	40	161	43	15

Tabla XVI: Ídem Tabla XV pero para la temperatura máxima diaria.

Número de eventos de temperatura máxima diaria	Eventos en	EX_F	F	M	C	EX_C
DEF	EMC	17	50	191	49	13
	EMA	16	45	193	50	16
MAM	EMC	19	49	200	49	17
	EMA	15	51	200	51	17
JJA	EMC	14	42	163	39	14
	EMA	14	39	163	42	14
SON	EMC	15	40	166	38	13
	EMA	14	40	163	40	15

En cuanto a la comparación estacional de frecuencia de eventos de Pigüé, se observa, en general, una subestimación del número de eventos por parte de la EMA en la categoría EX_F y F, y una sobreestimación en los eventos EX_C y C, tanto para la temperatura mínima diaria (Tabla XV) como para la máxima diaria (Tabla XVI). Esto es coincidente con lo hallado por Acquaotta y otros (2019), cuyos resultados también mostraron un mayor porcentaje de días cálidos y extremadamente cálidos en los registros automáticos, en comparación con los convencionales. Para ambas variables, la EMA iguala la cantidad de eventos de la EMC en las clases EX_C en MAM y JJA. Lo mismo ocurre en la categoría M en la temperatura mínima diaria de dichos trimestres. Cabe destacar que, aunque las EMC y la EMA de Pigüé cuenten con la misma cantidad de eventos dentro de una categoría, puede que no se correspondan a los mismos días. Es por ello que también se analiza el número de días compartidos dentro una misma clase (no se muestra).

6. CONCLUSIONES

Actualmente el SMN se encuentra en proceso de automatización de su red de estaciones convencionales. Para ello, se están realizando mediciones automáticas y manuales en paralelo a lo largo de toda la Argentina, las cuales deben ser comparadas. En este trabajo se presentan los procedimientos seleccionados por el organismo para la comparación de los registros de temperatura, con el objeto de determinar el impacto que puede generar el cambio del sistema de observación sobre la futura homogeneidad de las series temporales.

Como primera instancia, se realizó una búsqueda bibliográfica de trabajos que permitieran definir métricas, pruebas estadísticas y procedimientos a seguir para llevar a cabo la comparación. Se identificaron los trabajos de Kaspar y otros (2016), y de Acquaotta y otros (2019), entre otros, que se utilizaron como guía, replicando sus cálculos y adaptándolos a las necesidades del SMN.

Los procedimientos de comparación realizados se enfocaron en la escala diaria, pero también se contempló la horaria. Como casos de estudio se utilizaron pares de estaciones ubicadas en Villa Gesell, Pigüé y Aeroparque Buenos Aires. Como resultados del análisis se encontró que Pigüé exhibió las menores diferencias medias en las series horarias de temperatura. Lo mismo se observó en la serie de la temperatura máxima diaria. Para las tres ubicaciones, se encontraron medias inferiores a los 0.2°C y a los 0.5°C en módulo en las series de las temperaturas horarias en los tiempos de observación y en las series de las diferencias de las temperaturas máximas y mínimas diarias, respectivamente. Analizando distintas

medidas de posición de las distribuciones de los extremos diarias (mínimo, máximo, media, mediana y primer y tercer cuartil) se halló que, en la mayoría de los casos, las diferencias fueron inferiores al medio grado y negativas, lo que indicaría que, en general, hay sobreestimación de los valores por parte de las EMAs. En base a los valores de RMSE y a los coeficientes de correlación de rango de Kendall se advirtió que, para Villa Gesell y Pigüé y el año completo, existe una mayor correspondencia entre las series de temperatura máxima que entre las de temperatura mínima. Para Aeroparque ocurrió lo contrario.

De las pruebas de t-Student y de Wilcoxon se pudo establecer que, a excepción de la temperatura mínima diaria de Villa Gesell, las medias de las temperaturas extremas diarias obtenidas de las EMCs y las EMAs son significativamente diferentes, con un 95% de confianza. Sin embargo, en función de los resultados del test de Kolmogorov-Smirnov, no se hallaron indicios que permitan concluir que las temperaturas extremas diarias automáticas y convencionales no proceden de la misma distribución.

Se exploraron distintos procedimientos para el cálculo de las temperaturas medias diarias, si bien no todos pudieron ser aplicados en Villa Gesell y Pigüé porque no miden en los cuatro tiempos de observación ni en las 24 h. Para el caso del procedimiento basado en los extremos diarios, las medias de las diferencias resultaron negativas en las tres ubicaciones, indicando que, en promedio, los datos automáticos sobreestiman a los convencionales. La mayor dispersión de los datos y la máxima diferencia de valor medio se observaron en Aeroparque. Al desagregar la información en escala mensual, sólo en Villa Gesell y en los meses de septiembre y de noviembre a abril no se encontraron indicios de desigualdad de medias de las temperaturas medias diarias de las EMCs y EMAs. En Pigüé, por el contrario, se advirtieron los mejores resultados en cuanto al RMSE, al RMSE centrado y a la correlación de rango de Kendall para el año completo y las cuatro estaciones del año. Para los tres sitios, el año completo y casi todos los trimestres, las medias de las temperaturas medias diarias de las EMCs y las EMAs calculadas mediante los extremos diarios resultaron significativamente diferentes, con un 95% de confianza. Esto se pudo concluir de las pruebas de t-Student y de Wilcoxon. En cuanto a la comparación de las distribuciones, no se hallaron evidencias de que las temperaturas medias diarias automáticas y convencionales no provienen de la misma distribución en todos los casos.

En Aeroparque y la temperatura media obtenida de manera tradicional, se encontraron mejores resultados que al considerar el promedio de los extremos diarios. Se observaron bajos valores de media y de desvío estándar en la serie de las diferencias, lo que indica que, en promedio, el cambio en el procedimiento de cálculo de la variable no genera marcadas diferencias en dicha locación. Los hallazgos de las pruebas de t-Student y de Wilcoxon también fueron más alentadores. No se encontraron evidencias para concluir la desigualdad de medias a lo largo de todo año, a excepción de los meses junio y julio, y para establecer la no procedencia de la misma distribución. Cabe destacar que, a partir de estos análisis, se estudió las diferencias de las temperaturas medias diarias tanto por el cambio de instrumental como por la modificación del procedimiento de cálculo.

Se investigó la presencia de tendencias y ciclos en todas las series de las diferencias de temperatura diaria. Los análisis armónico y espectral permitieron identificar dos ciclos inmersos en la serie de las diferencias de la temperatura máxima en Aeroparque que, si bien explican alrededor del 1.8% de la varianza total, no resultaron significativos.

Con respecto a la comparación estacional de frecuencia de eventos de Pigüé, se observó, en general, una mayor frecuencia de días cálidos y extremadamente cálidos en las mediciones automáticas, en comparación con las manuales, tanto para la temperatura mínima diaria como para la máxima diaria. Lo encontrado va en línea con las diferencias medias negativas encontradas en todas las series de temperaturas diarias, que reflejan una sobreestimación de los promedios diarios por parte de las EMAs.

Cabe destacar que los tres lugares presentan un período diferente de superposición de observaciones, así como una distinta cantidad de pares de datos para el análisis. A su vez, poseen diferente resolución temporal en sus datos automáticos. Esto debe ser tenido en cuenta a la hora de interpretar los resultados e inferir conclusiones.

A partir de los resultados aquí encontrados se permitió identificar y caracterizar estadísticamente las diferencias entre observaciones convencionales y automáticas, una etapa clave en el proceso de homogeneización. Se evidenció la necesidad de replicar los cálculos y las comparaciones no sólo en otras EMCs y EMAs del SMN y demás redes emplazadas en el mismo sitio, sino también en otras variables meteorológicas y en pares de estaciones con mayores separaciones, con el objeto de analizar la calidad de los datos automáticos. Todo esto se abordará en trabajos futuros.

7. REFERENCIAS

Acquaotta, F., Fratianni, S., Venema, V. 2016. Assessment of parallel precipitation measurements networks in Piedmont, Italy. *Int J Climatol* 36(12):3963-3974.

Acquaotta, F., Fratianni, S., Aguilar, E. 2019. Influence of instrumentation on long temperature time series. *Climatic Change* 156, 385-404.

Batista, W. 2020. *Introducción a la Inferencia Estadística Aplicada: Teoría, cálculo e interpretación*. Editorial Facultad de Agronomía.

Biancotti, A., Destefanis, E., Fratianni, S., Masciocco, L. 2005. On precipitation and hydrology of Susa Valley (Western Alps). *Geogr Fis Dinam Quat* VII:51-58.

Davey, C.A., Pielke, R.A. 2005. Microclimate exposures of surface based weather stations. *Bull Am Met Soc* 86:497-504.

Devore, J.L. 2008. *Probabilidad y Estadística para Ingeniería y Ciencias*. 7ma edición. Cengage Learning.

Fiebrich, C.A., Crawford, K.C. 2009. Automation: A step toward improving the equality of daily temperature data produced by climate observing networks. *J Atmos Ocean Technol* 26:1246-1260.

Gallo, K.P. 2005. Evaluation of temperature differences for paired stations of the U. S. Climate Reference Network. *J Clim Notes and Correspondence* 18:1629–1636.

Hausfather, Z., Cowtan, K., Menne, M.J., Claude, N.W. 2016. Evaluating the impact of U.S. Historical Climatology Network homogenization using the U.S. Climate Reference Network. *Geophys Res Lett*: 1695-1701.

Isotta, F., Frei, C., Weilguni, V., Tadic, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., Vertacnik, G. 2013. The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int J Climatol* 34:1657-1675.

Kaspar, F., Hannak, L., Schreiber, K.J. 2016. Climate reference stations in Germany: Status, parallel measurements and homogeneity of temperature time series, *Adv. Sci. Res.*, 13, 163-171.

Katz, R.W., 1982. Statistical evaluation of climate experiments with general circulation models: A parametric time series modeling approach. *Journal of the Atmospheric Sciences*, 39, 1446-1455.

Katz, R.W., 1985. Probabilistic models. In: A.H. Murphy and R.W. Katz, eds., *Probability, Statistics, and Decision Making in the Atmospheric Sciences*. Westview, 261-288.

Lanzante, J.R. 2021. Testing for differences between two distributions in the presence of serial correlation using the Kolmogorov–Smirnov and Kuiper's tests. *Int J Climatol* 41(14):6314-6323.

Laurmann, J., Gates, L. 1977. Statistical considerations in the evaluation of climatic experiments with atmospheric general circulation models. *Journal of the Atmospheric Sciences*, 34(8), 1187-1199.

Madden, R.A., 1979. A simple approximation for the variance of meteorological time averages. *Journal of Applied Meteorology*, 18, 703-706.

Milewska, E., Vincent, L. 2016. Preserving Continuity of Long-Term Daily Maximum and Minimum Temperature Observations with Automation of Reference Climate Stations using Overlapping Data and Meteorological Conditions, *Atmosphere-Ocean*, 54:1, 32-47, DOI: 10.1080/07055900.2015.1135784.

New, M., Hulme, M., Jones, P.D. 1999: Representing twentieth century space-time climate variability. Part 1: development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate* 12, 829-856.

Organización Meteorológica Mundial (OMM). 2018. Guía de prácticas climatológicas. OMM- No. 100, Organización Meteorológica Mundial, Ginebra, Suiza.

OMM, 1966. Nota Técnica No 79 OMM, Climatic Change.

Squintu, A., van der Schrier, G., Brugnara, Y., Klein Tank, A. 2019. Homogenization of daily temperature series in the European Climate Assessment & Dataset. *Int J Climatol* 39:1243–1261.

Trewin, B.C., Trevitt, A.C.F. 1996. The development of composite temperature records. *Int J Climatol* 16:1227–1242.

Vincent, L., Mekis, E. 2009. Discontinuities due to joining precipitation station observations in Canada. *J Appl Met Climatol* 48(1):156–166.

Vincent, L., Milewska, E., Wang, X., Hartwell, M. 2018. Uncertainty in homogenized daily temperatures and derived indices of extremes illustrated using parallel observations in Canada. *Int J Climatol*, 38: 692-707.

Von Storch, H., 1995. Misuses of statistical analysis in climate research. In: H. von Storch and A. Navarra, eds., *Analysis of Climate Variability*. Springer, 11–26.

Wilks, D.S. 2006. *Statistical Methods in the Atmospheric Sciences*. 2nd Edition, Academic Press, London.

WMO (World Meteorological Organization). 2007. Guidelines for managing changes in climate observation programmes. World Climate Data and Monitoring Programme Series, WCDMP-No. 62, WMO-TD No. 1378, World Meteorological Organization, Geneva, Switzerland.

WMO. 2017. Challenges in the Transition from Conventional to Automatic Meteorological Observing Networks for Long-term Climate Records. WMO- No. 1202, World Meteorological Organization, Geneva, Switzerland.

Yue, S., Wang, C. 2002. The influence of serial correlation on the Mann–Whitney test for detecting a shift in mean. *Advances in Water Resources*. 25. 325-333.

AGRADECIMIENTOS

Irene Barnatan, Elian Wolfram, Federico Verstraeten y Lucas Berengua.

APÉNDICE I

Como se expresó en la sección 2, los registros de las EMAs se almacenan en la base de datos cada cierta cantidad de minutos, intervalo o paso temporal. La EMA Aeroparque AWOS considera el intervalo de 1 min, mientras que Villa Gesell SMN y Pigüe SINARAME, de 5 y 10 min, respectivamente. Aún cuando dos EMAs compartan el mismo paso temporal, el horario de almacenamiento del dato no siempre es el mismo y tampoco es necesariamente constante en todo el período de medición. Por ejemplo, considerando un paso temporal de 10 min, algunas EMAs almacenan sus registros a las 12:05 h y 12:15 h, y otras, a las 12:00 h y 12:10 h.

Dado que los distintos intervalos y horarios de almacenamiento de las series automáticas generan inconvenientes a la hora del análisis, es necesario definir un formato común entre todos los datos obtenidos automáticamente. En este caso, se decide llevar toda la información al paso temporal de 10 min, considerando la secuencia de 12:00 h, 12:10 h, etc. Para ello, se debe tener en cuenta que, salvo las EMAs de la red AWOS, los termómetros electrónicos de las EMAs no reportan una única variable térmica, sino tres: temperatura instantánea (*TempInst*), y temperatura máxima y mínima en el intervalo de tiempo (*TempMáx* y *TempMín*, respectivamente). La *TempInst* se utiliza para obtener las temperaturas horarias, mientras que las *TempMáx* y *TempMín*, para definir las temperaturas máximas y mínimas diarias. Como las EMAs de la red AWOS sólo reportan *TempInst*, esta se emplea para definir todas las variables diarias.

Para definir el formato común en las EMAs con almacenamiento a escala minutil, por ejemplo, los diez datos de las 12:01 h hasta las 12:10 h inclusive son candidatos para ser asignados a los del horario de las 12:10 h. Para obtener la *TempInst* de las 12:10 h, se elige el último registro no faltante de *TempInst* dentro del paso de tiempo de 10 min. Para definir las *TempMáx* y *TempMín* de las 12:10 h, se seleccionan los máximos y mínimos valores de *TempMáx* y *TempMín* (o de *TempInst* para las EMAs de la red AWOS), respectivamente, dentro del paso temporal. Se considera sin información en el horario de las 12:10 h sólo en el caso en que los diez datos minutales sean faltantes.

Los valores de *TempInst* cada 10 min se utilizan para obtener las temperaturas horarias. Para las temperaturas horarias en los tiempos de observación no se seleccionan los datos instantáneos de las horas en punto, sino los de aproximadamente 10 min previos, para imitar los horarios reales de medición de los observadores en las EMCs, beneficiando la correspondencia de los datos entre los pares de estaciones. Para las temperaturas horarias en cada una de las 24 h, sí se toman los valores a las horas en punto. Las temperaturas máximas y mínimas diarias se obtienen a partir de los valores máximos y mínimos de cada hora, conseguidos en función de las series de 10 min. Para cada hora, se busca el máximo de *TempMáx* y el mínimo de *TempMín* alcanzado, y luego, a partir de ellos, se retienen los valores más altos y bajos de un día. Sólo cuando los seis *datos de TempMáx y/o TempMín* cada 10 min dentro de una hora sean faltantes, la *TempMáx y/o TempMín* de esa hora se toma como ausente y no es posible calcular la temperatura máxima y/o mínima diaria, según el caso. Otra cuestión a tener en cuenta es que, para adquirir los extremos diarios en las EMAs, se considera el rango de las 21 h del día anterior a las 20:50 h del propio día, para que, nuevamente, exista una correspondencia con las mediciones manuales. En las EMCs se toman los rangos de 9 a 21 h y de 21 a 9 h para obtener las temperaturas máxima y mínima diarias. Cabe destacar que los datos de las EMAs deben ser sometidos a un control de calidad antes de emplearse. Como sus registros son cada cierta cantidad de minutos, los cuales luego son agregados en escalas horarias y diarias, se debe establecer un criterio de máximo número de datos faltantes por minutos y/o hora. En la actualidad, el SMN se encuentra en etapa de implementación de un nuevo sistema de gestión de datos climáticos. En una nota técnica posterior se documentarán los procesos de control de calidad que se implementen o configuren tanto en tiempo real como en diferido.

APÉNDICE II

En esta sección se describe brevemente en qué consisten las pruebas o test de hipótesis utilizados para evaluar, con un cierto nivel de confianza, las diferencias entre una serie de la EMC y su correspondiente EMA. Se presentan las hipótesis nula y alternativa que supone cada test, qué implicación tiene el rechazo de la primera, y qué conclusiones se pueden inferir a partir de esto. No se pretende detallar las fórmulas de los estadísticos de prueba ni el cálculo de los p-valores, pero sí exponer las modificaciones a las ecuaciones de los test clásicos debidas a la autocorrelación existente en las series de la temperatura diaria. Son modificaciones propuestas por diversos autores en el ámbito de las ciencias atmosféricas o hidrológicas, lo que suponen un agregado a los trabajos relevados.

Cabe destacar que, para la formulación de todos los test, se considera que las series son conjuntos de datos o muestras extraídas de poblaciones. La serie de la EMC representa la muestra x_1 y la conforman los n_1 datos medidos manualmente. La serie de la EMA corresponde a la muestra x_2 con n_2 datos medidos automáticamente. Más aún, como los registros de ambas estaciones son tomados en forma simultánea, en paralelo, se tratan de muestras pareadas, siendo $n_1 = n_2 = n$.

Sobre los test de hipótesis

Este apartado sigue el capítulo 8 de Devore (2008).

En toda prueba de hipótesis hay dos hipótesis contradictorias sobre parámetros y/o distribuciones poblacionales que entran en juego: la hipótesis nula (H_0), la afirmación desde el inicio que se supone como verdadera, y la hipótesis alternativa (H_a), la negación de la anterior. El objetivo consiste en decidir, con base en información muestral, cuál de las dos hipótesis es la correcta. Sólo si la evidencia muestral sugiere que H_0 es falsa, esta será rechazada en favor de H_a . En caso contrario, se seguirá creyendo que H_0 es cierta, dado que no hay información suficiente para refutarla. A su vez, para toda prueba de hipótesis, se definen dos elementos adicionales: un estadístico de prueba, una expresión en función de los datos muestrales utilizada como base para la decisión del test, y un nivel de significancia α , que se emplea para determinar la confianza de la prueba.

El resultado de un test de hipótesis se puede analizar a partir del p-valor. Como definición general, dado los datos de una muestra aleatoria, el p-valor es la máxima probabilidad que existe, en caso de ser cierta H_0 , de obtener otra muestra al azar cuyos datos sean tanto o más desfavorables para dicha hipótesis como los de la muestra original o las muestras originales (Batista, 2020). El p-valor es una medida estadística que permite determinar el rechazo o no de H_0 , confrontando con el valor de α establecido, aquí $\alpha = 0.05$. En este caso, si el p-valor es menor al umbral de 0.05, se puede rechazar H_0 a favor de H_a , aún con el riesgo de cometer un error de tipo I (el rechazo erróneo de H_0). Existe una fuerte evidencia contra H_0 , dado que hay menos de un 5% de probabilidad de que esta sea correcta. En el caso del test t-Student para diferencias de medias y el test de Kolmogorov-Smirnov para la comparación de distribuciones, para el cálculo del p-valor que contemple la autocorrección de los datos, es necesario una corrección del tamaño de muestra efectivo, detallado más abajo.

Test de Kendall para la correlación

Este apartado sigue al capítulo 3 de Wilks (2006).

Para establecer la existencia de una asociación entre la serie de una EMC y su correspondiente EMA se utiliza el coeficiente de correlación de rango de Kendall o de Rank-Kendall. Dicho coeficiente permite resumir la asociación a través de los rangos de las observaciones de ambas muestras. A diferencia del de Pearson, es una medida no paramétrica, ya que no se basa en suposiciones sobre las distribuciones de los datos. Esto lo hace más adecuado para utilizarlo en datos no distribuidos normalmente, como generalmente ocurre en las series de variables atmosféricas. Sus valores también oscilan entre -1 y 1. Los valores 1 y -1 representan una correlación perfecta positiva y negativa, respectivamente, mientras el valor cero, la ausencia de correlación.

A partir de las dos muestras, se puede calcular el coeficiente de correlación de Rank-Kendall muestral. Para testear su significancia se utiliza el test de Kendall, una prueba de hipótesis que establece como H_0 que el coeficiente poblacional es igual a cero, es decir, que no hay asociación entre las variables a nivel poblacional. La H_a define lo contrario. Si el p-valor calculado es menor al umbral determinado por α , se rechaza H_0 , lo que posibilitaría concluir que hay asociación entre las temperaturas diarias obtenidas de una EMC y de su correspondiente EMA con un $(1-\alpha).100\%$ de confianza.

Test t-Student bajo dependencia serial para la comparación de medias

Este apartado sigue el capítulo 5 de Wilks (2006).

Para probar si los valores medios de las series de temperatura diaria de una EMC y de su correspondiente EMA son significativamente diferentes, se trabaja con el problema equivalente de probar si la diferencia de las medias de las dos muestras es significativamente distinta de cero. Luego, se considera el test t-Student para diferencias de medias bajo dependencia serial. La elección de este test está motivada por tratarse de un caso de muestras pareadas y porque los datos integrantes de cada una de las muestras no son independientes entre sí, sino que exhiben una fuerte correlación serial o autocorrelación. Esto es de esperarse, el dato de temperatura de un día presenta una fuerte dependencia con el de sus días anteriores, lo que genera una violación a una suposición clave para muchas pruebas estadísticas. Cabe destacar que, si bien este el test t-Student es paramétrico, se decide utilizarlo en base al Teorema Central del Límite y al número de muestra suficientemente grande. En el apartado siguiente se describe una alternativa no paramétrica a dicho test.

Debido a la falta de independencia entre las observaciones dentro de las muestras, el test considerado aplica correcciones de tamaño efectivo de la muestra, n' , a las ecuaciones del estadístico y del p-valor de su forma clásica, reemplazando a n , el tamaño real de la muestra de las diferencias. La utilización de n' en lugar de n representa una corrección de la varianza de la distribución muestral del promedio de las diferencias debido a la correlación serial que presentan los datos diarios de temperatura. Ignorar esta dependencia conduce a una subestimación de esta varianza que, a su vez, lleva a una sobreestimación del estadístico de prueba y a una excesiva tasa de rechazos de H_0 . Siguiendo a Laurmann y Gates (1977), n' se puede estimar en términos de n y del coeficiente de autocorrelación con desfase o lag-1 de la muestra de las diferencias (r_1), utilizando la aproximación dada por la Ecuación A.1:

$$n' \simeq n \frac{1-r_1}{1+r_1} \quad \text{A.1}$$

Cuando no hay autocorrelación, $r_1 = 0$ y $n' = n$. Si existe una autocorrelación y es positiva, $r_1 \neq 0$ y $n' < n$, y, a medida que r_1 aumenta, n' se vuelve progresivamente más pequeño. Cuando $r_1 < 0$, Lanzante (2021) considera el enfoque conservador de que $n' = n$, lo que previene que n' exceda a n . Cabe destacar que la estimación de n' a partir de la Ecuación A.1 supone que los datos subyacentes siguen un proceso autorregresivo de primer orden. Las autorregresiones de primer orden suelen ser aproximaciones razonables para representar la persistencia de los valores meteorológicos diarios (Wilks, 2006). Para otros casos, como el de los datos horarios, puede ser necesario un modelo más complicado para describir la persistencia y derivar una expresión apropiada para n' (Katz 1982, 1985).

Como H_o y H_a , el test t-Student para diferencias de medias bajo la dependencia serial establece que la muestra de las diferencias entre los pares de datos provienen de una población con media igual y distinta a cero, respectivamente. Si el p-valor obtenido es menor que α , se rechaza H_o , lo que permitiría concluir, con $(1 - \alpha).100\%$ de confianza, que las medias de las temperaturas diarias obtenidas de una EMC y de su correspondiente EMA son significativamente diferentes.

Un ejemplo del test t-Student para las diferencias de medias bajo dependencia serial se encuentra en el capítulo 5 de Wilks (2006), donde se compara las temperaturas máximas promedio en Ithaca y Canandaigua para enero de 1987. Allí se observa que el valor de autocorrelación con lag-1 de la serie de las diferencias de las temperaturas diarias pareadas es mucho más pequeña que las autocorrelaciones en las dos series de temperaturas diarias individuales. Como gran parte de la dependencia temporal se exhibe conjuntamente en las series individuales, esta se elimina al calcular las diferencias. Esto evidencia una de las ventajas de usar la serie de diferencias para realizar la prueba de hipótesis: la autocorrelación relativamente baja de la serie de diferencia se traduce en un tamaño de muestra efectivo mayor, lo que produce una prueba más sensible.

Test de Wilcoxon de rango con signo bajo dependencia serial para la comparación de medias

Este apartado sigue el capítulo 15 de Devore (2008) y al trabajo de Yue y Wang (2002).

La prueba Wilcoxon de rango con signo o de Wilcoxon Signed-Rank para observaciones pareadas es una alternativa no paramétrica al test t-Student para muestras pareadas. Si no se supone normalidad, el primer test puede utilizarse para evaluar las hipótesis del segundo, siempre que la distribución de las diferencias entre los pares de datos sea continua y simétrica. Si las dos muestras tienen distribuciones continuas que difieren sólo con respecto a sus medias (de modo que la distribución de una es la distribución de la otra pero con un desplazamiento), entonces la distribución de las diferencias será simétrica y continua (lo que no implica que las distribuciones de las muestras originales sean simétricas individualmente). Aunque la suposición de simetría pueda parecer restrictiva, es más débil que la suposición de normalidad.

La prueba de Wilcoxon de rango con signo se basa en rangos en lugar de los valores numéricos de los datos. Por lo tanto, no depende de la distribución de los datos subyacentes y es resistente a los valores atípicos (Wilks, 2006). Sin embargo, sí continúa exigiendo la independencia serial entre las observaciones dentro de las muestras. Como se expresó anteriormente, las series de temperaturas a contrastar violan este supuesto debido a la autocorrelación diaria. Para subsanar esto, se considera el enfoque optado por Yue y Wang (2002), quienes propusieron “pre-blanquear” (“*prewhitening*”) la serie a evaluar antes de aplicar el test, siguiendo al trabajo de von Storch (1995) para la prueba de Mann-Kendall. Según sus resultados, este enfoque permite remover la influencia de la correlación serial de la serie y eliminar su influencia en el test. Una serie autocorrelacionada se pre-blanquea a partir de la siguiente ecuación:

$$y_t = x_t - r_1 \cdot x_{t-1} \quad (\text{A.2})$$

Donde x_t es la serie original, x_{t-1} es la serie desfasada un lag-1, r_1 es ahora el coeficiente de autocorrelación de la muestra de lag-1, e y_t es la serie de los residuos.

Bajo la suposición de continuidad y simetría en la distribución de la muestra de las diferencias, la H_0 del test de Wilcoxon de rango con signo para observaciones pareadas establece que la media de la población de la que se extrae la muestra de las diferencias es igual a cero. La H_a plantea lo contrario. Si el p-valor obtenido es menor que α , se rechaza H_0 , lo que permite concluir, con $(1 - \alpha).100\%$ de confianza, que las medias de las temperaturas diarias obtenidas de una EMC y de su correspondiente EMA son significativamente diferentes.

Test de Kolmogorov-Smirnov bajo dependencia serial para la comparación de distribuciones

Este apartado sigue Lanzante (2021).

Si bien el test t-Student es la prueba de hipótesis más utilizada para comparar dos conjuntos de datos, en ocasiones es necesario comparar otros parámetros de las muestras, más allá de sus valores medios. Por ejemplo, en un contexto de automatización de la red de observaciones, un aumento en la cola superior o inferior de la distribución de las temperaturas medidas puede llegar a tener igual o mayor repercusión que un cambio en la media. En este caso se puede desear, inclusive, determinar las diferencias entre las distribuciones enteras de las dos muestras.

Para testear las diferencias entre las distribuciones de las series de temperaturas diarias obtenidas automática y convencionalmente se utiliza el test de Kolmogorov-Smirnov para dos muestras, la prueba más frecuentemente empleada para la comparación de distribuciones. Dicho test involucra la comparación de las funciones de distribución acumulada (CDFs, por sus siglas en inglés) derivadas de dos muestras diferentes, y se basa en la máxima distancia entre las dos CDFs. Al igual que el anterior test, esta es una prueba no paramétrica que supone que los datos dentro de cada muestra son estadísticamente independientes. Por ello y siguiendo a Lanzante (2021) bajo el supuesto de que las series de tiempo subyacentes siguen un proceso autorregresivo de primer orden, se utiliza el enfoque del tamaño de muestra efectivo definido en la Ecuación A.1 para estimar el p-valor de la prueba y así evitar la inflación del error de tipo I.

El test de Kolmogorov-Smirnov considerado establece como H_0 que las muestras pertenecen a la misma población (es decir, tienen la misma distribución) y como H_a , lo contrario. Si el p-valor de la prueba es menor que α , se rechaza H_0 y se puede inferir, con un $(1 - \alpha).100\%$ de confianza, que las distribuciones de las temperaturas diarias obtenidas de una EMC y de su correspondiente EMA son significativamente diferentes.

Test de Mann-Kendall bajo dependencia serial para componentes determinísticas

Para analizar si la serie de la diferencia de la temperatura diaria de la EMC y su correspondiente EMA presenta una componente determinística, se utiliza el test no paramétrico de Mann-Kendall y el enfoque planteado en von Storch (1995) de pre-blanquear la serie previo aplicar la prueba, definido en la Ecuación A.2.

El test de Mann-Kendall plantea como H_0 que la serie de las diferencias es el azar y como H_a , que no lo es. Si el p-valor de la prueba es menor que α , se puede rechazar H_0 , concluyendo que la serie de las diferencias tiene una componente determinística que, a priori, es una tendencia, con un $(1 - \alpha).100\%$ de significancia.

Test de Fisher para la tendencia lineal

Para determinar la existencia de una tendencia lineal en la serie de la diferencia de la temperatura diaria de la EMC y su correspondiente EMA, se realiza un ajuste lineal a partir de la metodología de cuadrados mínimos y la prueba de Fisher para testear la pendiente. El método de cuadrados mínimos permite encontrar la recta o el modelo de regresión lineal que mejor ajusta a los datos de la muestra, definida a partir de la ordenada de origen $\hat{\beta}_0$ y la pendiente $\hat{\beta}_1$. Dicha recta es una estimación de la verdadera recta poblacional con ordenada y pendiente β_0 y β_1 , respectivamente.

El test de Fisher plantea como H_0 que la pendiente de la recta poblacional es igual a cero y como H_a , que es distinto. Si el p-valor de la prueba es menor que α , se puede rechazar H_0 con $(1 - \alpha).100\%$ de confianza. Con el mismo nivel de confianza se puede concluir que la pendiente de la recta de regresión de la serie de las diferencias de las temperaturas diarias de una EMC y una EMA es significativamente distinta de cero y, por lo tanto, presenta una tendencia.

Análisis armónico, análisis espectral y test de Anderson

Para detectar periodicidades e identificar ciclos significativos dentro la serie de las diferencias de las temperaturas diarias de una EMC y su correspondiente EMA, se utilizan las metodologías de análisis armónico y de análisis espectral. El análisis armónico consiste en representar las fluctuaciones o variaciones en la serie original como el resultado de la suma de un conjunto de armónicos o de series de funciones senos y cosenos (Wilks, 2006). También permite ver cuáles son los armónicos que mayor varianza explican de la serie original y filtrarlos, para aislar efectos. Sin embargo, no permite testear la significancia de dichos armónicos. Para ello, se emplea la metodología de análisis espectral. El análisis espectral es una técnica de análisis en el dominio de frecuencias, que permite identificar ondas significativas presentes en una serie. En este caso se utiliza el análisis espectral simple según el método de Blackman Tukey (OMM, 1966) y la ventana de Hann. La significancia del espectro se estudia mediante el modelo de ruido blanco o de ruido rojo. El primer modelo considera que la serie es aleatoria y el segundo, que presenta persistencia. Para determinar qué modelo emplear, se realiza el test de Anderson. La H_0 de dicho test establece que la serie es al azar y la H_a , lo contrario, es decir, que presenta memoria. En caso de rechazar H_0 , se elige el modelo de ruido rojo y se calcula su espectro teórico y sus bandas de significancia. Más aún, se puede concluir que las ondas cuyas densidades espectrales superen la banda de significancia superior (previamente identificadas pero no evaluadas estadísticamente mediante el análisis armónico) son significativas con $(1 - \alpha).100\%$ de confianza.

Instrucciones para publicar Notas Técnicas

En el SMN existieron y existen una importante cantidad de publicaciones periódicas dedicadas a informar a usuarios distintos aspectos de las actividades del servicio, en general asociados con observaciones o pronósticos meteorológicos.

Existe no obstante abundante material escrito de carácter técnico que no tiene un vehículo de comunicación adecuado ya que no se acomoda a las publicaciones arriba mencionadas ni es apropiado para revistas científicas. Este material, sin embargo, es fundamental para plasmar las actividades y desarrollos de la institución y que esta dé cuenta de su producción técnica. Es importante que las actividades de la institución puedan ser comprendidas con solo acercarse a sus diferentes publicaciones y la longitud de los documentos no debe ser un limitante.

Los interesados en transformar sus trabajos en Notas Técnicas pueden comunicarse con Ramón de Elía (rdelia@smn.gov.ar), Luciano Vidal (lvidal@smn.gov.ar) o Martin Rugna (mrugna@smn.gov.ar) de la Gerencia de Investigación, Desarrollo y Capacitación, para obtener la plantilla WORD que sirve de modelo para la escritura de la Nota Técnica. Una vez armado el documento deben enviarlo en formato PDF a los correos antes mencionados. Antes del envío final los autores deben informarse del número de serie que le corresponde a su trabajo e incluirlo en la portada.

La versión digital de la Nota Técnica quedará publicada en el Repositorio Digital del Servicio Meteorológico Nacional. Cualquier consulta o duda al respecto, comunicarse con Melisa Acevedo (macevedo@smn.gov.ar).