**Análisis espacio temporal de la variable vapor de agua atmosférico en estaciones GNSS de América**

***B. Peralta 1\*****, V.I. Babolene 1, L.H. Peña 1, M.V. Mackern 1,2,3, P. Rosell1,2,3, M.F. Camisay2,3 y M. L. Mateo1,2*

*Centro de Ingeniería Mendoza Argentina*

*1 Universidad Nacional de Cuyo, Facultad de Ingeniería. Mendoza. Argentina*

*2 Universidad Juan Agustín Maza, Facultad de Ingeniería y Enología. Mendoza. Argentina*

*3 Consejo Nacional de Investigaciones Científicas y Técnicas*

**Resumen**

La variable atmosférica vapor de agua, es considerada una de las variables esenciales, ya que es reguladora de la temperatura en el planeta. Es una variable muy cambiante, tanto en el espacio como en el tiempo, por lo cual resulta fundamental medir, monitorear y analizar su variabilidad.

La técnica más utilizada para medir el vapor de agua atmosférico es mediante radiosondeos. Estos por su costo son utilizados generalmente en aeropuertos y respondiendo al requerimiento de la Asociación internacional de aviación civil, son lanzados a las 12 hs y en algunos casos a las 00 hs, permitiendo conocer el contenido de agua precipitable del perfil atmosférico sobre el sitio en esos horarios.

Desde las estaciones continuas de observación GNSS, es posible estimar el retardo cenital troposférico que se produce en la señal por la suma de dos factores, la componente hidrostática del aire seco y la componente húmeda causada por el contenido de vapor de agua de la troposfera. Contando con la componente húmeda del retardo es posible determinar el vapor de agua integrado (IWV) en la atmósfera de manera indirecta.

En América se dispone de más de 500 estaciones GNSS, que conforman la red SIRGAS, cuyo principal objetivo es brindar un marco de coordenadas único, preciso y global. Desde el Centro de procesamiento de Ingeniería Mendoza Argentina se calcula el vapor de agua en las 500 estaciones SIRGAS y se dispone de una serie de datos de más de 10 años (2012 a 2023).

En este trabajo se presenta una metodología de análisis de las series temporales del vapor de agua. La misma se basa en el modelado de la serie considerando el valor medio, la tendencia decadal, la variabilidad anual y semi-anual, como también la variabilidad diaria y los residuos respecto de los datos de IWV. Se presenta la regionalización de las estaciones considerando el análisis de los parámetros determinados en el modelo de ajuste aplicado.

El análisis fue realizado a través de scripts de python usando librerías como numpy, pickle entre otras. Estas bibliotecas nos permitieron realizar cálculos precisos y manipular datos de manera eficiente, brindando una comprensión más profunda de las relaciones entre las variables. Además, para facilitar la colaboración y el control de versiones en nuestro equipo, empleamos GitHub como plataforma central.

**Palabras clave:** Vapor de agua; retardo troposférico; ZTD; SIRGAS

**Introducción**

El vapor de agua atmosférico es una variable crucial que influye en la regulación térmica de nuestro planeta. Su naturaleza altamente cambiante en el espacio y el tiempo exige una medición, monitoreo y análisis minuciosos. Mientras los radiosondeos son la técnica predominante para su medición, las estaciones continuas de observación GNSS han surgido como una herramienta clave para estimar el contenido de vapor de agua en la atmósfera de manera indirecta.

En este contexto, se destaca la red SIRGAS con más de 500 estaciones en América, donde desde CIMA (Centro de procesamiento de Ingeniería en Mendoza, Argentina), se han calculado (Mackern et al., 2020) y recopilando datos sobre el vapor de agua durante más de una década (2012-2023). Este trabajo presenta una metodología de análisis de las series temporales del vapor de agua, incorporando modelado estadístico, tratando de comprender su comportamiento a lo largo del tiempo.

Este análisis se basa en la identificación de patrones temporales como el valor medio, la tendencia decadal, variabilidades anuales y semi-anuales, además de la variación diaria, junto con la regionalización de las estaciones según los parámetros extraídos del modelo aplicado.

**Materiales y métodos**

**Recopilación de Datos**

Los datos utilizados en este estudio corresponden a más de 500 estaciones GNSS pertenecientes a la red SIRGAS en América. Las mismas se encuentran ubicadas en diversas regiones del continente y han registrado estimaciones del vapor de agua atmosférico durante un periodo que abarca desde 2012 hasta 2023, por CIMA. Por lo cual los productos de CIMA fueron los datos utilizados en este modelado.

**Cálculo del Vapor de Agua Atmosférico**

La variable atmosférica de vapor de agua fue estimada a partir de las estaciones GNSS desde el retardo cenital troposférico, ZTD. Este retardo se produce debido a la suma de dos componentes: la componente hidrostática del aire seco y la componente húmeda causada por el contenido de vapor de agua en la troposfera. Esta componente húmeda del retardo permite determinar el vapor de agua integrado (IWV) en la atmósfera de manera indirecta (Rosell et al., 2023).

**Filtrado de estaciones**

Como el objetivo del trabajo es modelar una serie temporal, al igual que en Bianchi et al.,(2016) se decidió analizar solamente aquellas estaciones que tuvieran datos de al menos 7 años consecutivos. Por lo tanto, de un total de 553 estaciones con productos troposféricos han sido analizadas 345.

**Modelado de las Series Temporales de IWV**

Para analizar la variabilidad del vapor de agua a lo largo del tiempo, se aplicó un modelado estadístico de las series temporales. El modelado se realizó utilizando técnicas de mínimos cuadrados y la función polinómica según Bianchi et al, (2016) y Nilsson and Elgered (2008), aplicada a la variable IWV(Ec. 1) en este caso, estimando ocho parámetros para cada una de las estaciones seleccionadas:

(Ec. 1)

**Herramientas de Programación**

La implementación de este análisis se realizó utilizando Python, un lenguaje de programación versátil y ampliamente utilizado en el ámbito científico y de investigación. Entre las principales bibliotecas empleadas, numpy se destacó por su capacidad para manejar arreglos y matrices de datos numéricos de manera eficiente. Esta biblioteca fue clave para realizar operaciones matemáticas avanzadas, como el cálculo de desviaciones estándar y la manipulación de grandes conjuntos de datos sin sacrificar velocidad ni precisión.

Otra herramienta esencial fue pickle, una biblioteca que permite la serialización y deserialización de objetos en Python, facilitando el almacenamiento y recuperación de datos complejos. Pickle fue utilizada principalmente para guardar modelos y resultados intermedios, lo que permitió retomar el trabajo sin necesidad de recalcular procesos costosos en términos computacionales. Esto optimizó el flujo de trabajo al asegurar que los datos procesados estuvieran fácilmente accesibles para futuras consultas y análisis adicionales.

Además, se utilizó GitHub como plataforma de control de versiones. Esta herramienta permitió una colaboración fluida entre los miembros del equipo, facilitando el seguimiento de los cambios, la revisión de código y la gestión de diferentes ramas del proyecto. GitHub no solo sirvió como repositorio central, sino que también ofreció un entorno de desarrollo colaborativo que permitió realizar un seguimiento eficiente de las actualizaciones, garantizando la integridad y seguridad de los datos y el código a lo largo de todo el ciclo de vida del proyecto.

Por último, la biblioteca random de Python fue utilizada para la selección aleatoria de datos, lo que resultó fundamental en este análisis para dividir los conjuntos en muestras de entrenamiento y validación. Al generar selecciones aleatorias, se asegura una distribución equitativa y no sesgada de los datos, lo que permite resultados más representativos y evita la influencia de patrones específicos en las muestras, garantizando un análisis robusto y confiable.

**Validación del Modelo**

En esta fase de validación, se ha aplicado una comparación entre los valores modelados por la función “ajustada” y los datos observados en las estaciones GNSS de la red SIRGAS (fig. 1, 2 y 3). Para este propósito, el 70% de los datos fue seleccionado aleatoriamente utilizando la función random de Python, y se estimaron las diferencias entre los valores de vapor de agua IWVmodelado70% y los valores reales IWVSIRGAS obtenidos a partir de las estaciones. Este enfoque permitió identificar la precisión del modelo ajustado y cuantificar el nivel de desviación estándar para cada estación en particular.

Se calcularon los desvíos estándar de las diferencias entre los valores modelados y observados para todas las estaciones. Los resultados mostraron que la mayor parte de las estaciones GNSS tienen un desvío estándar que se concentra en torno a los 6 kg/m², lo cual indica un ajuste razonablemente adecuado. Se observó una distribución simétrica en los desvíos estándar, lo que sugiere una dispersión homogénea de las diferencias, con pocas estaciones mostrando comportamientos atípicos, lo que permitió concluir que el modelo tiene un buen desempeño en la mayoría de los casos.

Finalmente, se validó el valor medio del vapor de agua obtenido del modelo con los valores medios de las estaciones GNSS, y se encontró que las diferencias eran menores a ±1 kg/m². Esto asegura que el modelo es capaz de reproducir correctamente la media del IWV para la mayoría de las estaciones, brindando un alto grado de confianza en la metodología aplicada.

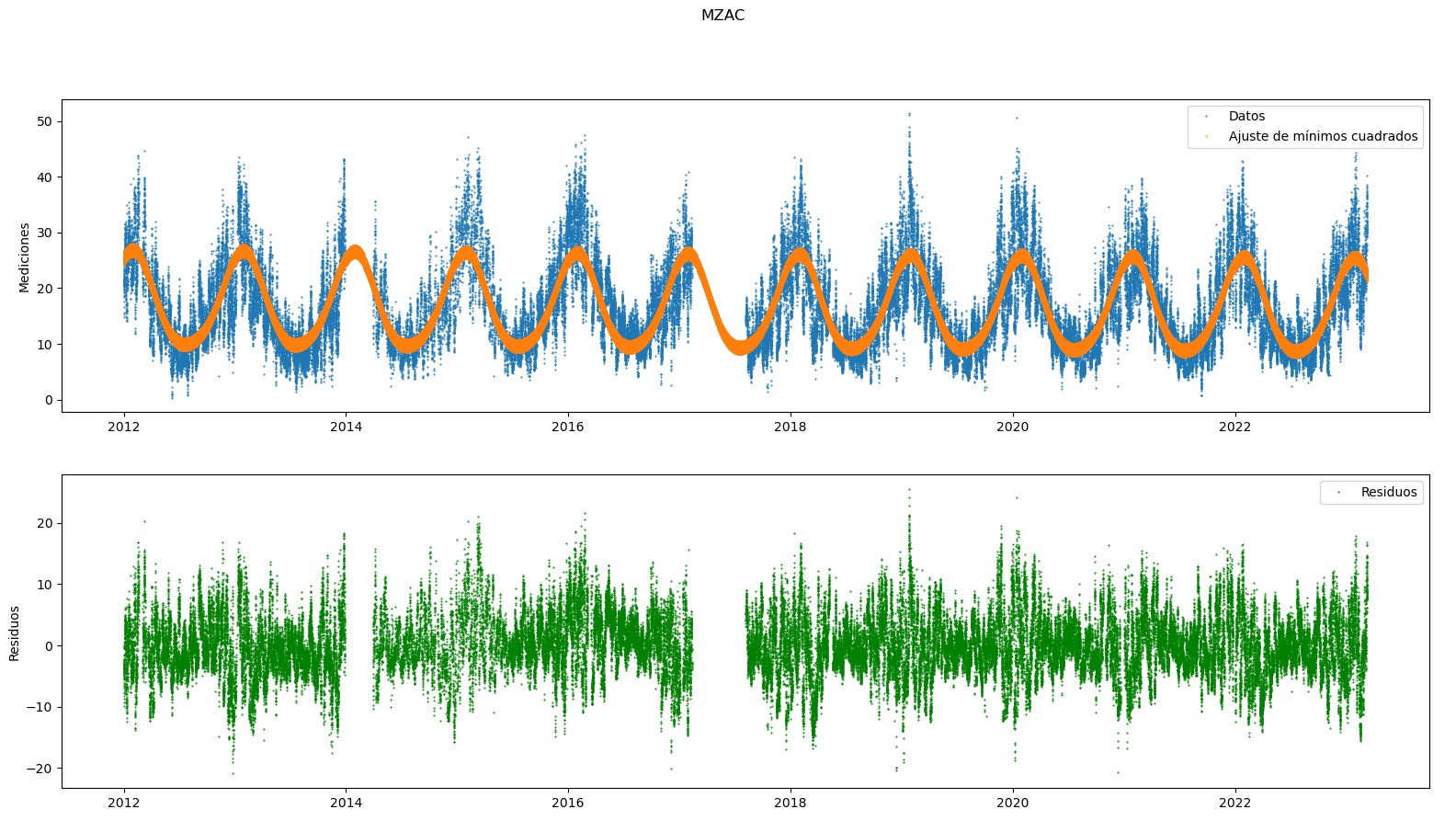
**Análisis Regional**

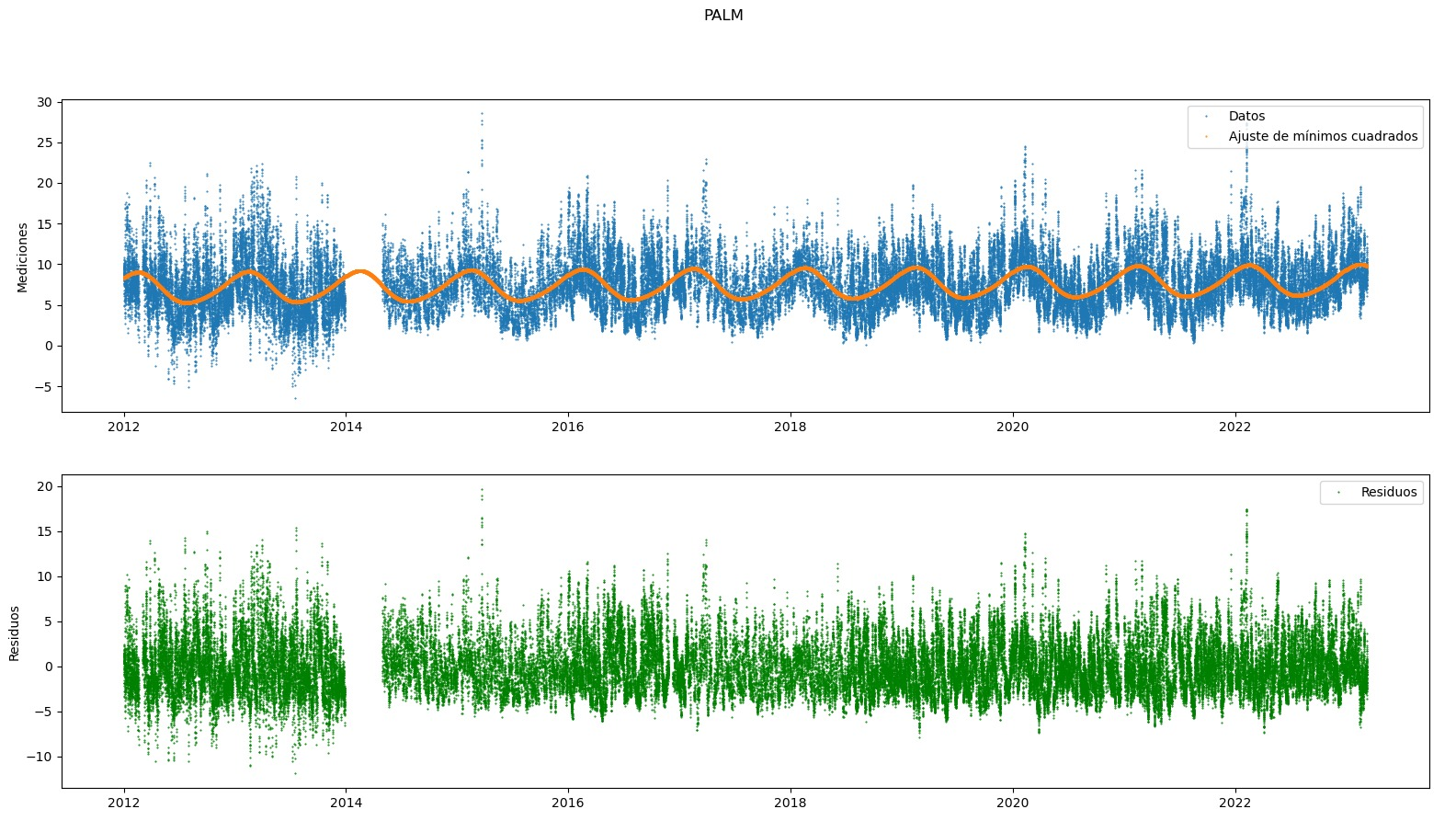
Se realizó una regionalización de las estaciones considerando los parámetros determinados en el modelo de ajuste aplicado. Esto permitió identificar patrones regionales en la variabilidad del IWV.

**Resultados y discusión**

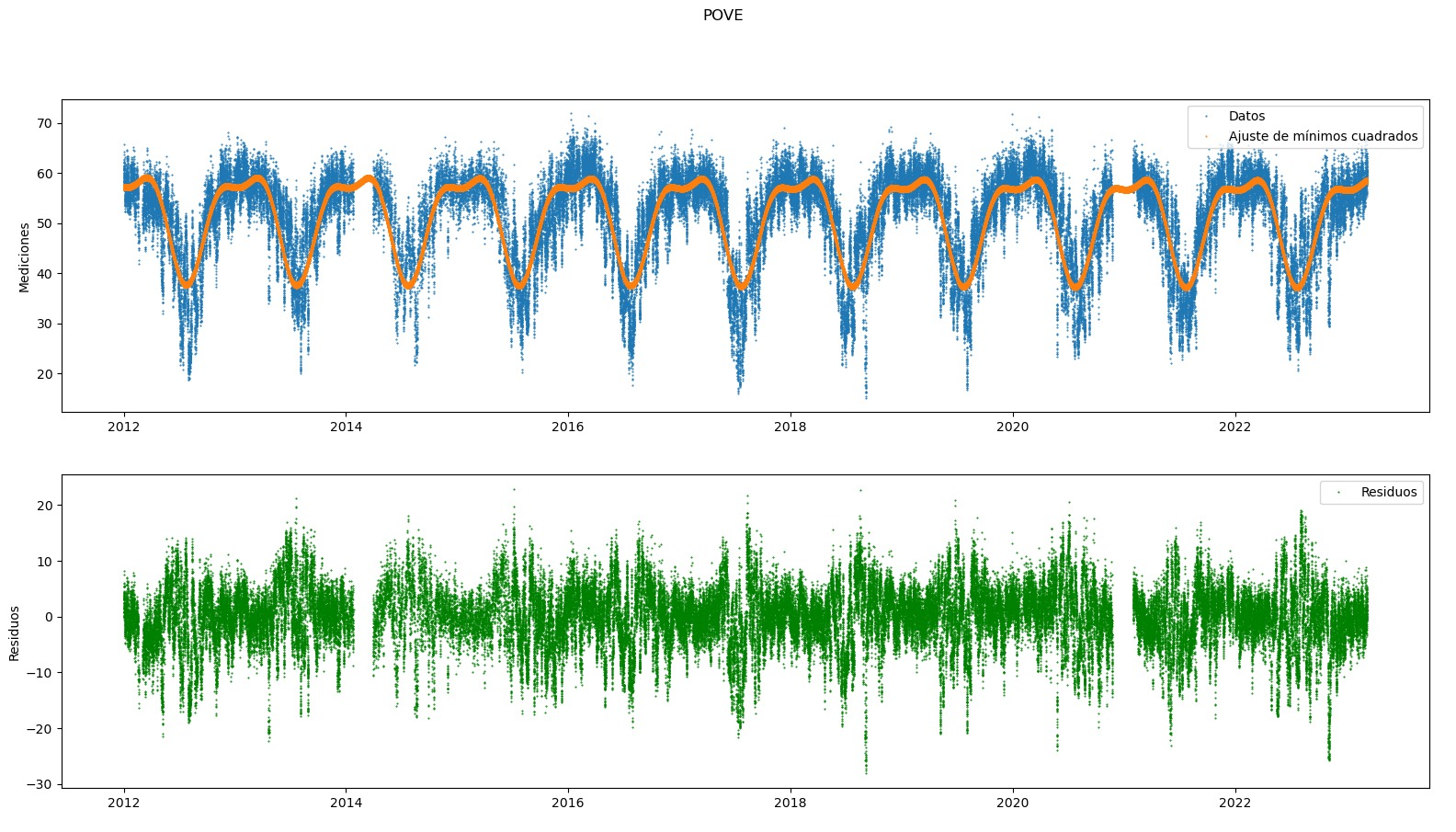
**Resultado I**

Se modeló la serie de IWV aplicando mínimos cuadrados y estimando 8 coeficientes (valor medio, tendencia decadal, variabilidad anual, semianual y diaria) de una función del IWV (Ec. 1), para cada una de las estaciones seleccionadas. Se pueden observar 3 ejemplos en las estaciones de Mendoza, Argentina (Fig. 1), Antártida (Fig. 2) y Porto Velho, Brasil (Fig.3).

**Figura 1**. Datos (IWVSIRGAS) y ajuste por mínimos cuadrados (IWVmodelado). Diferencias IWVSIRGAS - IWVmodelado. MZAC (Mendoza, Argentina)



**Figura 2**. Datos (IWVSIRGAS) y ajuste por mínimos cuadrados (IWVmodelado). Diferencias IWVSIRGAS - IWVmodelado. PALM (Antártida)

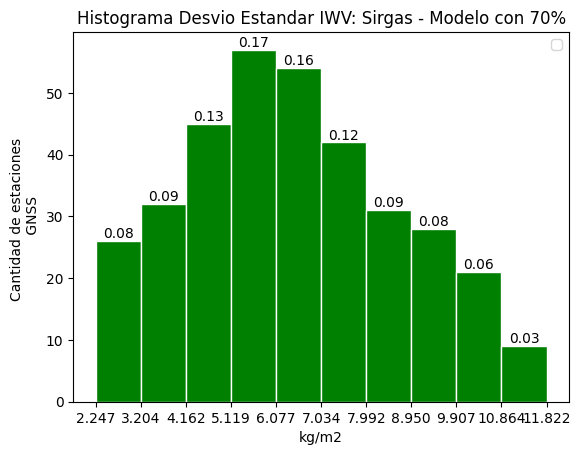


**Figura 3**. Datos (IWVSIRGAS) y ajuste por mínimos cuadrados(IWVmodelado). Diferencias IWVSIRGAS - IWVmodelado. POVE (Porto Velho, Brasil).

**Resultado II**

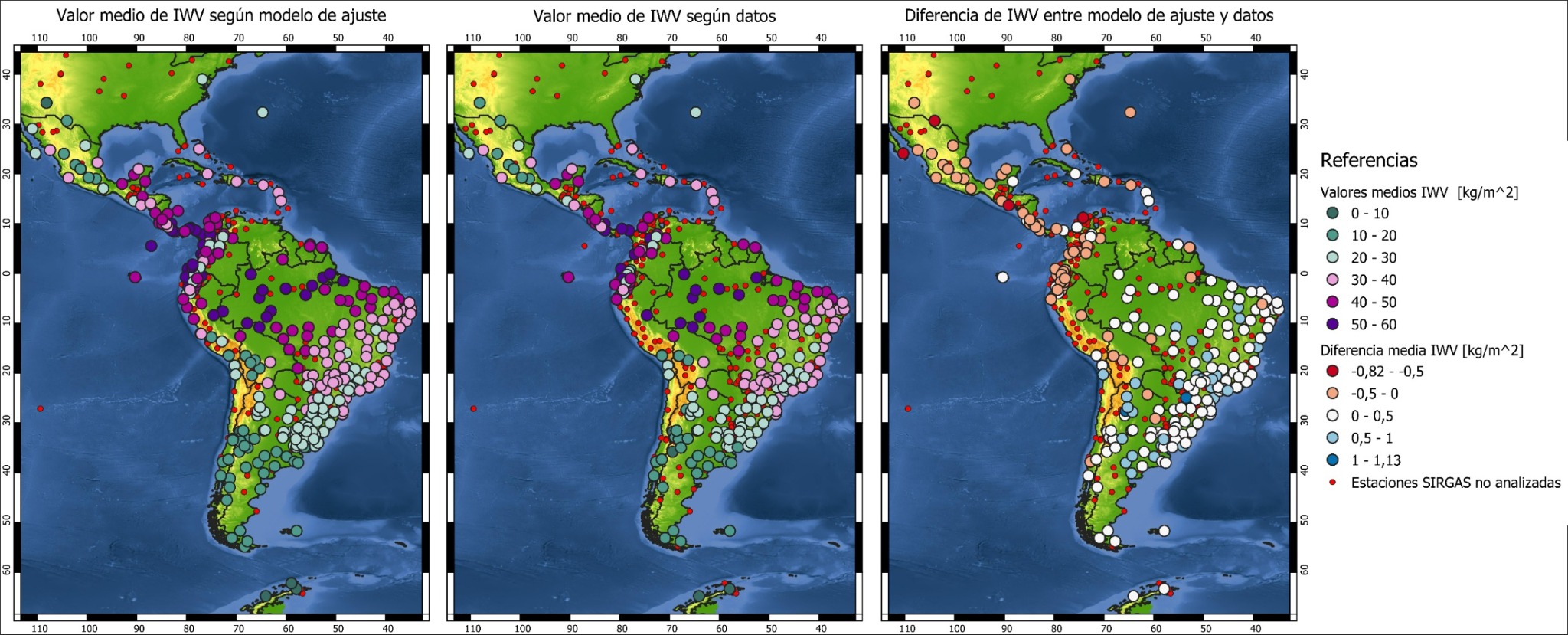
La función "Modelada"(IWVmodelado) se validó utilizando el 70% de los datos y con respecto a los valores dato de la serie, IWVSIRGAS. Se estimó el desvío estándar de las diferencias para cada estación. El histograma del desvío estándar de las diferencias entre el IWV modelado y los datos de la serie (IWVSIRGAS muestra que la mayor parte de las estaciones GNSS tienen un desvío estándar concentrado en torno a los 6 kg/m², con un pico en el intervalo de 6.077 kg/m², donde se encuentra el 17% de las estaciones. La distribución tiene una forma simétrica de tipo campana, lo que sugiere una dispersión homogénea de los desvíos estándar, con una media de alrededor de 6.07 kg/m². El rango del desvío estándar va desde 2.247 kg/m² hasta 11.822 kg/m², lo que indica que la variabilidad en las diferencias entre el modelo y los datos reales varía significativamente según la estación GNSS.

El percentil 50 (mediana) se sitúa cerca del intervalo de 6.077 kg/m², lo que confirma que la mitad de las estaciones tienen un desvío estándar menor o igual a este valor. El percentil 25 se sitúa alrededor de los 4.162 kg/m², mientras que el percentil 75 alcanza aproximadamente los 7.992 kg/m², lo que indica que el 50% central de las estaciones se encuentra dentro de este rango intercuartílico. Además, se observa que los desvíos más pequeños, por debajo de 3 kg/m², y los más grandes, superiores a 9 kg/m², corresponden a un número reducido de estaciones (alrededor del 8% y 3% respectivamente), lo que sugiere que estas estaciones presentan comportamientos atípicos en términos de variabilidad.

En resumen, el modelo logra una aproximación adecuada en la mayoría de las estaciones, con un comportamiento central bien definido, aunque con algunas estaciones fuera del rango típico de desvío estándar.

**Figura 4.** Histograma de los desvíos estándar de las diferencias IWVSIRGAS - IWVmodelado.

**Resultado III**

 Se validó el valor medio (parámetro a0) del IWVmodelado con respecto al valor medio del IWVSIRGAS calculado en cada estación desde los ZTDSIRGAS, por Rosell et al., (2023). En todas las estaciones comparadas (242) la diferencia resultó menor que ±1 [kg/m2]. Además se pudo observar una muy buena correspondencia regional del IWV medio calculado (Fig. 5).

**Figura 5**. Valor medio de IWV según el modelo (izq.) y los datos calculados por Rosell (centro). Diferencias de IWV medio(derecha).

**Discusión**

Si bien los resultados obtenidos son prometedores los residuos todavía muestran cierta periodicidad que el modelo de ajuste no está representando.

A priori hay 2 posibles resultados, el más esperado es poder encontrar estas frecuencias y que obtengamos una serie más representativa de los datos o que no haya una periodicidad y estos residuos se deban a fenómenos en las regiones o algún error de medición en las estaciones.

En trabajos futuros intentaremos, a través de varios modelos matemáticos como series de Fourier o series wavelet encontrar las altas frecuencias y los correspondientes parámetros que el modelo aplicado no ha logrado inferir, tratando de lograr un modelo de mejor ajuste a los datos. También analizaremos aplicar algoritmos de machine learning (Inteligencia artificial) para ver que resultados se obtienen y si las series pueden representar mejor las altas frecuencias.

**Conclusiones**

De un total de 553 estaciones GNSS de la red SIRGAS, que disponen de IWV, se seleccionaron 345 estaciones que cumplieron la condición de disponer de una serie temporal de 7 años o más, para modelar su variabilidad.

Se modeló la serie de IWV aplicando mínimos cuadrados y estimando 8 coeficientes (valor medio, tendencia decadal, variabilidad anual, semianual y diaria) aplicando una función similar a la aplicada por Bianchi et al, (2016), para cada una de las estaciones seleccionadas.

La función "Modelada" se validó calculando las diferencias del valor IWVmodelado con respecto a los valores dato de la serie, IWVSIRGAS. Se estimó el desvío estándar de las diferencias para cada estación concluyendo que para el 76% de las estaciones (263) el desvío estándar es menor que 8 kg/m2(≈ un 25% del valor medio de IWV). Resultando un sesgo de 0kg/m2(no mostrado).

Se validó el valor medio (coeficiente a0) del IWVmodelado con respecto al valor medio del IWV calculado en cada estación desde los productos troposféricos de SIRGAS (Rosell et al.,2023). Las diferencias resultaron menores que ± 1 [kg/m2]. Además, se observa una muy buena correspondencia regional del IWV medio (Fig. 5).

Si bien se estimó la tendencia decadal (coeficiente a1) en cada estación, sólo 231 estaciones cuentan con más de 10 años sobre las que hemos iniciado su análisis (no mostrado)

Aún nos falta modelar la variabilidad sub-diaria, quizás horaria, para lo cual estamos encaminando una nueva investigación en la selección de frecuencias a adoptar por estación.

**Agradecimientos**

Deseamos expresar nuestro más sincero agradecimiento a CIMA por proveernos de los IWVSIRGAS, a SIRGAS por brindar las observaciones con las cuales se estiman los ZTD y a la Facultad de Ingeniería de la Universidad Nacional de Cuyo por su invaluable apoyo y colaboración en la realización de este proyecto de investigación. Su respaldo ha sido fundamental para el éxito de estos desarrollos.

**Referencias**

Bianchi, C. E., Mendoza, L. P. O., Fernández, L. I., Natali, M. P., Meza, A. M., &

Moirano, J. F. (2016). Multi-year GNSS monitoring of atmospheric IWV

over Central and South America for climate studies. Annales

Geophysicae, 34(7), 623–639.

https://doi.org/10.5194/angeo-34-623-2016

Mackern, M. V., Mateo, M. L., Camisay, M. F., & Morichetti, P. V. (2020).

Tropospheric Products from High-Level GNSS Processing in Latin

America. In International Association of Geodesy Symposia book series.

https://doi.org/10.1007/1345\_2020\_121

Nilsson, T and Elgered, G. (2008). Long-term trends in the atmospheric water

vapor content estimated from groundbased GPS data, J.Geophys.

Res.-Atmos., 113, https://doi.org/10.1029/2008JD010110

Rosell, P. A., Mackern Oberti, M.V., Rivera, J.A., Euillades, P.A. (2023). 10-Year

assessment of GNSS integrated water vapour in the SIRGAS network,

Journal of South American Earth Sciences, Volume 130, 104539, ISSN

0895-9811, https://doi.org/10.1016/j.jsames.2023.104539.