

Análisis de ruido con series de tiempo de la red GNSS de monitoreo continuo del Ecuador (REGME)

Gustavo Roberto León Lara¹, Marco Patricio Luna Ludeña²

¹ Universidad de las Fuerzas Armadas – ESPE, Departamento de Ciencias Exactas. Av. General Rumiñahui s/n. Sangolquí - Pichincha. Ecuador

² Universidad de las Fuerzas Armadas – ESPE Departamento de Ciencias de la Tierra y de la Construcción. Av. General Rumiñahui s/n. Sangolquí - Pichincha. Ecuador

*Autor para correspondencia: gustavo.leon@upec.edu.ec

Recibido: 2024/07/03 Aprobado: 2024/11/30

DOI: <https://doi.org/10.26621/ra.v1i31.956>

RESUMEN

En la presente investigación se analizó el tipo de ruido presente en la Red GNSS de Monitoreo Continuo del Ecuador (REGME) con la aplicación de técnicas de series temporales, como la serie de Fourier y modelos ARIMA. Se logró extraer y clasificar el ruido de 52 estaciones, para obtener una visión clara de las interferencias que afectan las mediciones GNSS; estos resultados permitieron generar recomendaciones para la optimización del monitoreo continuo y el aumento de la precisión de los datos geodésicos.

Tras un análisis inicial de la duración de los datos, se identificaron 16 estaciones viables para el estudio, descartando las demás debido a la insuficiencia de datos o a información que no cumplía con los requisitos para el tratamiento. Los datos seleccionados fueron depurados y se aplicó un análisis de descomposición lineal a las series temporales para calcular valores estadísticos, como tendencias, estacionalidades y ruido. Se concluyó que una estación presentaba ruido blanco, mientras que fue común encontrar ruido rosa o *flicker* en la mayoría, típico de estaciones de monitoreo continuo. Sin embargo, algunas estaciones, como GZEC y SIEC, mostraron un tipo de ruido rojo o *random walk*, con índices espectrales cercanos a -2, atribuido a la estabilidad del monumento o al tipo de suelo en el que se encuentran empotradas las estaciones. Así, se busca que los hallazgos encontrados a partir de esta investigación puedan aplicarse a otras áreas del conocimiento.

Palabras clave: ruido, serie de Fourier, series de tiempo, estaciones de monitoreo

ABSTRACT

In the present investigation, the type of noise present in the GNSS Continuous Monitoring Network of Ecuador (REGME) was analyzed, with the application of time series techniques, such as the Fourier series and ARIMA models. It was possible to extract and classify the noise of 52 stations, to obtain a clear view of the interferences that affect GNSS measurements. These results allowed for generating recommendations for the optimization of continuous monitoring and to increase the accuracy of geodetic data.

After an initial analysis of the duration of the data, 16 viable stations were identified for the study, discarding the others due to insufficient data or information that did not meet the requirements for treatment. The selected data was cleaned, and a linear decomposition analysis was applied to the time series to calculate statistical values such as trends, seasonality, and noise. It was concluded that one station presented white noise, and it was common to find pink noise or flicker in the majority, typical of continuous monitoring stations. However, some stations, such as GZEC and SIEC, showed a type of red noise or random walk, with spectral indices close to -2, attributed to the stability of the monument or the type of soil in which the stations are embedded. Thus, it is intended that the findings of this research can be applied to other areas of knowledge.

Keywords: noise, Fourier series, time series, monitoring stations

Gustavo Roberto León Lara  orcid.org/0009-0003-8690-4659

Marco Patricio Luna Ludeña  orcid.org/0000-0003-1433-2658



INTRODUCCIÓN

En 1974, surgieron los sistemas globales de navegación satelital (GNSS), herramientas fundamentales para la navegación terrestre y por satélite. Estos sistemas han evolucionado significativamente con el desarrollo de nuevos sensores, lo que ha permitido que la navegación por satélite impulse una revolución en la topografía moderna. Gracias a su metodología, características y aplicaciones específicas, se considera que esta tecnología es altamente efectiva (Bhardwaj, 2020). Además, los GNSS son de gran utilidad para la comunidad científica, pues su amplia gama de aplicaciones en la investigación y vigilancia de la Tierra facilita la comprensión del posicionamiento y la órbita terrestre, la meteorología, el clima espacial, la reflectometría y el monitoreo de terremotos (Jin et al., 2022).

Del mismo modo, los sistemas de posicionamiento global (GPS) proporcionan mediciones precisas en cualquier lugar y momento, ya sea en el espacio o en el tiempo. Sin embargo, estos sistemas se ven influenciados por diversos factores que pueden afectar a la precisión y fiabilidad de los datos. Entre estos factores, el ruido desempeña un papel fundamental (Ao et al., 2023). Por lo tanto, es esencial entender y caracterizar el ruido presente en las series de tiempo de las mediciones GNSS de la REGME para corregir los errores en las mediciones. Esta comprensión permitirá mejorar la calidad de los datos obtenidos y facilitará su aplicación en diversas áreas, como la planificación territorial y la gestión de desastres naturales.

En Ecuador, la Red GNSS de Monitoreo Continuo (REGME) tiene como objetivo monitorear el movimiento de la corteza terrestre y la provisión de información para aplicaciones geodésicas y geofísicas. Esta red abarca el territorio nacional con 33 estaciones permanentes; así, ofrece información crucial para la toma de decisiones en planificación territorial y gestión de desastres naturales, gracias a su robusta cobertura y alta densidad de puntos de monitoreo (Leiva, 2007).

En este contexto, el desarrollo de la presente investigación, a través del análisis de series de tiempo, permitirá identificar el ruido generado por las 33 estaciones que se encuentran ubicadas en terrenos, pilares y edificaciones. Por su ubicación, estas estaciones generan interferencias en las señales GPS. Los resultados de este análisis facilitarán la toma de decisiones por parte de las entidades responsables del desarrollo cartográfico y servirán a proyectos relacionados con la geodesia, así como a investigadores y otras organizaciones clave. Al identificar los ruidos que distorsionan las mediciones, el estudio ofrecerá la posibilidad de proponer mejoras que optimicen la calidad de los datos, lo que a su vez generará resultados positivos en futuras investigaciones y proyectos.

La investigación se fundamenta en un robusto marco teórico que respalda su desarrollo. En este contexto, se adopta la definición de modelo propuesta por Mauricio (2007): "Un modelo para un proceso estocástico es cualquier conjunto de hipótesis bien definidas sobre las propiedades estadísticas de dicho proceso". En este estudio se presentan modelos univariantes conocidos como "ruidos", que pueden clasificarse en tres tipos: ruido blanco, ruido rojo y ruido de salto (pág. 13).

El ruido blanco se caracteriza por presentar una media cero y una varianza constante a lo largo del tiempo. Por otro lado, el ruido rojo exhibe una correlación positiva entre observaciones cercanas en el tiempo, lo que implica que los valores tienden a ser similares a sus predecesores. Adicionalmente, existe el ruido de salto, que se refiere a cambios abruptos en la serie de tiempo y está comúnmente asociado con eventos externos o anomalías (Zhang et al., 2017).

Existen diferentes modelos que nos ayudan a caracterizar el ruido; uno de ellos es el modelado estadístico que utiliza modelos ARIMA (AutoRegressive Integrated Moving Average). En este enfoque, el ruido blanco generalmente presenta coeficientes autoregresivos con valores cercanos a cero, mientras que los ruidos rosa y rojo requieren de una estructura más compleja para su análisis y caracterización (Aggarwal et al., 2020).

Otro método relevante es el de Allan Variance, propuesto por Zhang et al. (2013), que se utiliza para caracterizar el ruido en relojes y sistemas de tiempo, y que también puede aplicarse a señales GNSS. Este método se basa en el cálculo de la varianza de las diferencias de frecuencia a diferentes escalas de tiempo, lo que permite distinguir entre el ruido blanco, rosa y rojo.

Asimismo, se puede emplear el análisis de la función de autocorrelación de la señal. En este análisis, el ruido blanco presenta una autocorrelación cercana a cero en todos los retardos, mientras que el ruido rosa muestra un decaimiento gradual con el tiempo. Por otro lado, el ruido rojo exhibe una autocorrelación significativa en retardos más largos (Benoist et al., 2020). En esta misma línea, el análisis de la desviación estándar residual calcula la desviación estándar de las diferencias entre los valores observados y los valores predichos por un modelo. Este análisis permite comparar la desviación residual con el modelo esperado para los diferentes tipos de ruido, proporcionando así información adicional sobre la calidad del modelo y la naturaleza del ruido presente en los datos (Dicker, 2012).

Finalmente, una herramienta efectiva para analizar, modelar y predecir los ruidos es el uso de series temporales, que nos permiten identificar y caracterizar los diferentes tipos de ruido de manera significativa. En este contexto, el análisis espectral se presenta como un método altamente confiable que emplea la transformada de Fourier para realizar la transición de la representación temporal a la frecuencial. Este método examina el espectro de potencia de la señal, en el que el ruido blanco exhibe una distribución de potencia constante a lo largo de todas las frecuencias. Por su parte, el ruido rosa presenta una caída de 3 dB por octava en el espectro de frecuencias, mientras que el ruido rojo muestra una mayor concentración de potencia en las frecuencias bajas (Kaczmarek y Kontny, 2018; Wang et al., 2012).

Por lo tanto, esta investigación llevará a cabo el análisis y la determinación del tipo de ruido mediante la observación del espectro de potencia, aproximado por una ley dependiente de la frecuencia, tal como fue propuesto por Agnew (1992). La periodicidad se obtiene a través del análisis espectral utilizando series de Fourier, con el objetivo de identificar el tipo de ruido presente en cada una de las estaciones de la red (REGME). Esto permitirá establecer si las estaciones están ubicadas en sitios apropiados o estables, contribuyendo así a proporcionar información valiosa que debe ser considerada en el análisis y la predicción de las series de tiempo. En caso contrario, se recomendará una nueva monumentación de aquellas estaciones de monitoreo continuo cuyos datos estén fuertemente afectados por el ruido.

MÉTODOS

Las estaciones de monitoreo continuo de la red REGME en un único archivo. Actualmente, la red está constituida por 52 estaciones; con toda la información se desarrolló la metodología para la obtención y el análisis del ruido de estas estaciones.

Homogenización de Marcos de Referencia

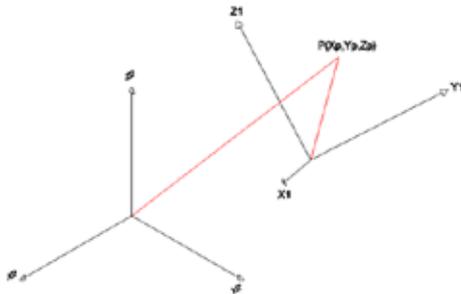
Se llevó a cabo la homogeneización de la base de datos de la forma más actual, ITRF2020, utilizando el método de Helmert. Este método es comúnmente empleado para transformar coordenadas entre dos sistemas de referencia geodésicos diferentes; así, se facilitan las transformaciones de posición, rotación y escala entre sistemas de coordenadas (Mataija et al., 2014; Tran et al., 2023).

La transformación de Helmert generalmente se expresa mediante una serie de parámetros de transformación:

- Desplazamiento en las direcciones X,Y,Z: Representa las diferencias de origen entre los sistemas de referencia en esas direcciones.
- Rotación alrededor de las direcciones X,Y,Z: Representa las diferencias de orientación entre los sistemas de referencia.
- Factor de escala: Representa la diferencia de escala entre los sistemas.

Estos parámetros se utilizan en una fórmula matemática para calcular las coordenadas transformadas en el sistema de referencia de destino a partir de las coordenadas originales en el sistema de referencia de origen.

Figura 1 Transformación Helmert



Nota. Proceso de transformación de coordenadas por el método de Helmert (Mataija et al., 2014).

La fórmula general de transformación de Helmert para coordenadas tridimensionales (X,Y,Z) se expresa de la siguiente manera:

$$\begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} + (1 + \delta) \begin{pmatrix} 1 & R_z & -R_y \\ -R_z & 1 & R_x \\ R_y & -R_x & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} \quad (1)$$

Donde:

- X_2, Y_2, Z_2 son las coordenadas transformadas.
- X_1, Y_1, Z_1 son las coordenadas originales.
- δ es el factor de escala.
- R_x, R_y, R_z son las cantidades de las matrices de rotación alrededor de los ejes X,Y,Z.
- T_x, T_y, T_z son las cantidades de desplazamiento en las direcciones X,Y,Z.

Los parámetros de transformación de los diferentes marcos de referencia al marco actual ITRF2020 se detallan en la siguiente tabla.

Tabla 1 Transformación Helmert ITRF2020

Unidad	Tx	Ty	Tz	d	Rx	Ry	Rz
	mm	mm	mm		mas	mas	mas
ITRF 2005 a ITRF 2020	-2,7	-0,1	1,4	-0,65	0,00	0,00	0,00
ITRF 2008 a ITRF 2020	-0,2	-1,0	-3,3	0,29	0,00	0,00	0,00
ITRF 2014 a ITRF 2020	1,4	0,9	-1,4	-0,42	0,00	0,00	0,00

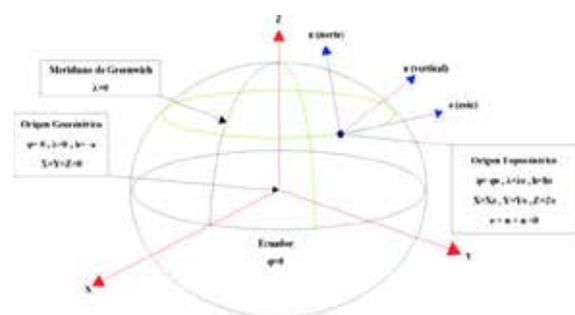
Fuente. Elaborado por el autor, adaptado de Zamora et al. (2021) y Altamimi et al. (2011)

Conversión a Coordenadas Topocéntricas

Las coordenadas cartesianas globales (X,Y,Z) son aquellas que se encuentran referidas al centro de masas de la tierra o geocentro, incluyendo la atmósfera y los océanos. El eje Z coincide con el eje de rotación terrestre, el eje X se sitúa en la dirección del plano meridiano que pasa por Greenwich y está contenido en el plano ecuador, internacionalmente considerado como meridiano origen o meridiano cero desde 1884, y el eje Y es aquel contenido en el plano ecuador y perpendicular al eje X; su sentido será tal que los tres ejes formen una triplete dextrógrafa.

Los valores de estas coordenadas son excesivamente altos, lo que dificulta la ubicación y orientación sobre la superficie terrestre. Para facilitar el análisis de las series temporales, se realiza la conversión a coordenadas locales en relación con un sistema topocéntrico, cuyo origen se establece en un punto específico sobre la superficie terrestre. Esta conversión se orienta de acuerdo con lo ilustrado en la Figura 2, a fin de realizar una mejor interpretación y análisis de los datos.

Figura 2 Sistema de coordenadas cartesianas globales (X,Y,Z) y topocéntricas locales (E,N,U)



Fuente. Sistema de coordenadas cartesianas globales (X,Y,Z) y topocéntricas locales (E,N,U) (Drewes y Sánchez, 2011)

Las coordenadas cartesianas globales o geocéntricas (X,Y,Z) se convierten a coordenadas topocéntricas (E,N,U), en las que (E,N) corresponden a la parte horizontal y (U) es la componente vertical. Esta conversión ofrece una interpretación más clara del comportamiento de cada serie en todos sus componentes; además, los valores son pequeños y centrados en cero, lo que facilita su estudio (Luna, 2017).

Detección de Valores Atípicos e Imputación de Datos Faltantes

Una vez generadas las series temporales de cada estación en coordenadas topocéntricas, se procede a la depuración mediante la imputación de los datos faltantes y a la detección de valores atípicos con un diagrama de caja y bigotes (o *boxplot*). Este enfoque gráfico proporciona una representación clara de la distribución estadística y permite identificar de manera visual las observaciones que se apartan significativamente de la norma.

Al crear un *boxplot*, se genera una caja que abarca el rango intercuartílico (IQR), definido como la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1). La mediana se representa como una línea en el interior de la caja. Los "bigotes" del *boxplot* se extienden hasta ciertos límites, 1.5 veces el IQR, marcando así la zona esperada para la mayoría de los datos.

Los datos que caen fuera de estos límites son considerados atípicos y se representan como puntos individuales, por lo que este método es particularmente valioso en la exploración inicial de conjuntos de datos, ya que permite identificar posibles anomalías que podrían requerir una atención adicional. Es fundamental tener en cuenta que, si bien el *boxplot* es una herramienta poderosa, la interpretación de datos atípicos debe considerar el contexto específico del problema y la naturaleza de los datos para así evitar conclusiones precipitadas.

A continuación, se reemplazan los mismos como NA, para luego volver a imputar con el método SDT (*Seasonal Decomposition of Time Series*), de manera que se mantenga el mismo número de datos originales. Para la imputación de datos faltantes se utiliza la librería "impute TS" del método SDT en el *software* R. En primer lugar, el algoritmo realiza una descomposición estacional de series temporales por Loess vía STL para estimar las componentes Tt (tendencia), St (estacionalidad) y et (ruido), con base en la regresión Loess desarrollada por Dudek (2023). Este enfoque implica dos aplicaciones iterativas de la regresión Loess. Para utilizar el método, se requiere especificar una frecuencia de muestreo relacionada con el periodo de la componente estacional. Esta frecuencia se define al declarar la variable de datos como un objeto "ts" con una frecuencia (52,12,4,1) para representar lo semanal, mensual, trimestral o anual, respectivamente.

Descomposición Aditiva de las Series Temporales

Las series temporales generadas por las estaciones de monitoreo continuo están compuestas de una frecuencia principal, así como de sus armónicos más altos; se asume que la serie de tiempo contiene tanto la parte determinista mediante el modelo funcional, que incluye la tendencia y la estacionalidad, así como el ruido de fondo (Amiri-Simkooei et al., 2007).

Una vez depuradas las series se realiza la descomposición de estas, tomando en consideración la tendencia, la estacionalidad y el ruido, de manera que la suma de todos estos componentes represente la serie temporal real. El modelo teórico es:

$$Y_t = T_t + S_t + N_t \quad (2)$$

La tendencia de las series temporales a largo plazo se aproxima a una línea recta que describe el comportamiento aproximado de una estación; para su obtención se utilizó una regresión lineal.

La estacionalidad se obtiene con base en el análisis espectral de las series temporales. El espectro es el promedio estadístico de una señal o

tipo específico que se analiza en términos de su contenido de frecuencia. El espectro de potencia de una serie temporal describe la distribución de potencias en componentes de frecuencia que componen esa señal (Luna, 2017a). La densidad espectral de potencia (PSD, *Power Spectral Density*) es uno de los términos más comunes para referirse al espectro de potencia. Para el análisis se utilizó la transformada Fourier Scargle (1982), dada por:

$$P(f_n) = \frac{1}{N} \left[\left(\sum_{i=1}^N v_i \cos(2\pi f_n t_i) \right)^2 + \left(\sum_{i=1}^N v_i \sin(2\pi f_n t_i) \right)^2 \right]$$

donde $f_n = n/T$, T es el periodo fundamental, v_i es el residuo y $n=1,2,\dots,N/2$.

Determinación del Ruido

Para poder identificar el tipo de ruido que afecta a cada estación es importante conocer el índice espectral; para ello, se realiza un análisis mediante el estudio del espectro de potencia propuesto por Agnew (1992), que está aproximado por una ley de potencia dependiente de la frecuencia de la forma:

$$P(f) = P_0 f / f_0^\alpha \quad (4)$$

Donde:

f es la frecuencia temporal.

P_0 y f_0 son constantes de normalización.

α es el índice espectral.

Los valores del índice espectral varían entre $-3 < \alpha < 0$, un patrón que aparece en la mayoría de los fenómenos geofísicos (Agnew, 1992). Para la obtención de este valor es necesario linealizar la función anterior, que da como resultado una recta $\alpha \log(P(f)) - \log(f)$, en la que la pendiente es el valor correspondiente al índice espectral. Cuando el índice espectral toma valores próximos a cero, se conoce como ruido blanco; si los valores están cercanos a -1, son del tipo de ruido de parpadeo o rosa (flicker) y, si los valores son próximos o mayores a -2, tienen las características de ruido browniano o de camino aleatorio (*random walk*).

Resultados y Discusión

Al trabajar con diferentes marcos de referencia, es esencial someter los datos a un proceso de homogenización con anterioridad al análisis. Esto se realiza mediante los parámetros de transformación entre cada uno de los marcos. De no llevarse a cabo este proceso, se podrían generar saltos en las series temporales en la fecha de cambio, tal y como señala Luna Ludeña (2017).

Una solución a este problema es el método de transformación de Helmert, cuyos resultados se presentan en la Tabla 2. En esta tabla, se puede apreciar la transformación de los diferentes marcos de referencia al marco actual ITRF20.

Tabla 2 Estación LJEC

Se- ma- na	X	Y	Z	X	Y	Z
	IGS05			IGS20		
1517	1192829,012	-6252161,654	-440799,181	1192829,012	-6252161,654	-440799,181
1518	1192829,013	-6252161,662	-440799,183	1192829,013	-6252161,662	-440799,183
1519	1192829,015	-6252161,663	-440799,181	1192829,015	-6252161,663	-440799,181
	IGS08			IGS20		
1632	1192829,012	-6252161,653	-440799,164	1192829,012	-6252161,653	-440799,164
1633	1192829,012	-6252161,655	-440799,165	1192829,012	-6252161,655	-440799,165
1634	1192829,014	-6252161,643	-440799,162	1192829,014	-6252161,643	-440799,162
	IGS14			IGS20		
1934	1192829,010	-6252161,664	-440799,117	1192829,010	-6252161,664	-440799,117
1935	1192829,011	-6252161,660	-440799,117	1192829,011	-6252161,660	-440799,117
1936	1192829,009	-6252161,667	-440799,118	1192829,009	-6252161,667	-440799,118
	IGS20			IGS20		
2238	1192828,968	-6252161,682	-440799,074	1192828,968	-6252161,682	-440799,074
2239	1192828,968	-6252161,678	-440799,075	1192828,968	-6252161,678	-440799,075
2240	1192828,969	-6252161,676	-440799,076	1192828,969	-6252161,676	-440799,076

Nota. La tabla muestra los diferentes IGS recopilados de diferentes semanas e IRTF y su respectiva transformada a ITRF2020. Elaborado por el autor.

Conversión a Coordenadas Topocéntricas

Como se mencionó anteriormente, la transformación de las coordenadas cartesianas globales a coordenadas topocéntricas facilita el análisis de las series temporales. Este cambio permite una mejor ubicación y orientación en la superficie terrestre, aspectos cruciales para el estudio. Además, se observa que los valores del sistema cartesiano son significativamente más altos en comparación con los del sistema topocéntrico, tal y como se detalla en la Tabla 3.

Tabla 3 Estación PREC

Se- ma- na	X	Y	Z	E	N	U
2247	1329685,22027	-6236040,73174	-188840,20143	-0,05603	0,08686	-0,02273
2248	1329685,22116	-6236040,73449	-188840,20307	-0,05547	0,08523	-0,02051
2249	1329685,22051	-6236040,73510	-188840,20200	-0,05568	0,08636	-0,01957
2250	1329685,22113	-6236040,73058	-188840,20294	-0,05485	0,08534	-0,02345

Nota. La tabla contiene un ejemplo de datos de la estación en ITRF20; con coordenadas XYZ y coordenadas locales. Elaborado por el autor.

Detección de Valores Atípicos e Imputación de Datos Faltantes

En la Tabla 4 se presentan los datos correspondientes a las estaciones de la red REGME. En ella, se puede observar que la mayoría de las estaciones presentan un porcentaje elevado de valores faltantes. Algunas estaciones, como CLEC, STEC, LPEC y POEC, cuentan con un número reducido de datos de observación, lo que impide realizar un análisis adecuado de las series temporales.

Tabla 4 Estación REGME

Antena	Observación	Datos Faltantes	%
ABEC	288	111	38.54 %
ALEC	437	55	12.59 %
BHEC	319	65	16.93 %
CEEC	117	35	23.03 %
CHEC	358	134	27.24 %
CLEC	181	311	63.21 %
COEC	479	13	2.64 %
CUEC	634	111	14.90 %
CXEC	469	0	0.00 %
DPEC	165	3	1.79 %
ECEC	482	10	2.03 %
EPEC	467	2	0.43 %
EREC	466	25	5.09 %
ESEC	35	2	5.41 %
FOEC	318	81	20.30 %
GQEC	117	0	0.00 %
GUEC	271	128	32.08 %
GYEC	285	466	62.05 %
GZEC	423	68	13.85 %
IBEC	434	58	11.79 %
ICEC	117	0	0.00 %
JNEC	37	3	7.50 %
LAEC	40	0	0.00 %
LJEC	673	61	8.31 %
LPEC	72	327	81.95 %
LREC	43	541	92.64 %
MAEC	556	75	11.89 %
MHEC	154	350	69.44 %
MTEC	151	433	74.14 %
MUEC	31	9	22.50 %
NJEC	407	97	19.25 %
ONEC	287	112	28.07 %
OREC	28	140	83.33 %
PAEC	166	2	1.19 %
PDEC	93	397	81.02 %
PEEC	34	365	91.48 %
PIEC	40	0	0.00 %
PJEC	139	352	71.69 %
PLEC	154	14	8.33 %
PMEC	281	36	11.36 %
POEC	165	3	1.79 %
PREC	433	58	11.81 %
PTEC	243	502	67.38 %
PVEC	3	358	99.17 %
QVEC	479	152	24.09 %
SCEC	161	307	65.60 %
SEEC	362	129	26.27 %
SIEC	398	1	0.25 %
STEC	162	422	72.26 %
TNEC	474	30	5.95 %
TPEC	138	30	17.86 %
TSEC	120	48	28.57 %

Nota. Listado de antenas con los respectivos datos faltantes. Adaptado de la base de datos de la ITRF (2020).

A partir de este primer análisis, se han seleccionado 16 estaciones de monitoreo continuo, que han registrado un tiempo de observación de aproximadamente 7 años, equivalente a más de 354 semanas de datos semanales, como se detalla en la Tabla 5. Esta duración de la serie temporal proporciona una base sólida para el análisis posterior de los datos.

Tabla 5 Estaciones seleccionadas

Antenas	Semanas	Datos existentes	Datos Faltantes	%
ALEC	358	307	51	14.25 %
BHEC	358	294	64	17.88 %
COEC	358	350	8	2.23 %
CUEC	358	288	70	19.55 %
CXEC	358	358	0	0.00 %
ECEC	358	348	10	2.79 %
EPEC	358	356	2	0.56 %
EREC	358	333	25	6.98 %
GZEC	338	299	39	11.54 %
IBEC	354	333	21	5.93 %
LJEC	358	326	32	8.94 %
MAEC	358	328	30	8.38 %
PMEC	317	281	36	11.36 %
PREC	358	323	35	9.78 %
SIEC	358	357	1	0.28 %
TNEC	358	329	29	8.10 %

Nota. Selección de antenas con menor porcentaje de datos perdidos. Adaptado de la base de datos de la ITRF (2020).

El porcentaje de datos faltantes en estas estaciones varía entre 0 % y 19.55 %. La estación CXEC se destaca como la más completa, mientras que la estación CUEC presenta el mayor porcentaje de datos faltantes, un 19.55 %. Sin embargo, es importante destacar que el porcentaje de datos faltantes en ambas estaciones se encuentra dentro de los parámetros aceptables. Según Li et al. (2023), los datos faltantes no deben exceder el 20 % del total para que sea viable realizar un proceso de relleno o imputación.

Es fundamental analizar los valores atípicos o anómalos de las estaciones seleccionadas, ya que estos influyen de manera significativa en la determinación del ruido. En la Tabla 6 se presentan estos valores junto con sus respectivos porcentajes, que oscilan entre 0.28 % y 10 %. Estos porcentajes son considerados elevados para estaciones de monitoreo continuo; las estaciones EREC y EPEC son las más afectadas. Una ventaja primordial de estas estaciones es el extenso periodo de tiempo durante el que se han recopilado datos, que permite corregir las anomalías a través del proceso de imputación de datos faltantes.

Tabla 6 Porcentajes de datos atípicos por estación

Antenas	Semanas	E	N	Up	%
ALEC	358	0	0	1	0.28 %
BHEC	294	0	1	0	0.34 %
COEC	358	0	0	2	0.56 %
CUEC	358	0	0	0	0.00 %
CXEC	358	12	0	7	5.31 %

ECEC	358	32	0	4	10.06 %
EPEC	358	24	0	12	10.06 %
EREC	358	1	0	2	0.84 %
GZEC	338	0	0	0	0.00 %
IBEC	354	7	0	2	2.54 %
LJEC	358	0	0	0	0.00 %
MAEC	358	0	0	0	0.00 %
PMEC	317	0	0	0	0.00 %
PREC	358	0	0	1	0.28 %
SIEC	358	0	0	0	0.00 %
TNEC	358	0	0	0	0.00 %

Nota. Datos atípicos por coordenadas. Adaptado de la base de datos de la ITRF (2020).

En el resto de las estaciones, el porcentaje de datos atípicos varía entre 0.28 % y 5 %, cifras que son coherentes con los porcentajes reportados en estudios similares. Por ejemplo, Nikolaidis (2002) alcanzó un nivel promedio de limpieza de datos de entre 1 % y 4 %. El método utilizado por este autor se basa en la aplicación de la mediana y el rango intercuartílico (IQR) como estadísticos para identificar los valores atípicos en las series temporales analizadas.

Estas estaciones presentan una ventaja significativa debido al extenso período de observación de datos que poseen, lo que permite llevar a cabo el proceso de imputación de datos en cada una de ellas. En la Tabla 7 se presenta un ejemplo ilustrativo del proceso de imputación de datos, que muestra cómo se ha aplicado este método para completar las series temporales de las estaciones analizadas.

Tabla 7 Estación PREC datos imputados

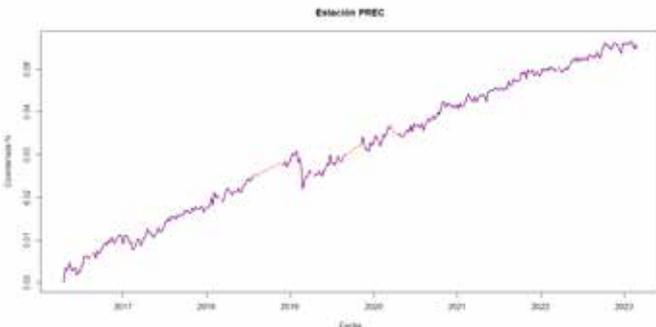
Se-mana	E			N			U		
	E	N	U	E	N	U	E	N	U
	Datos a imputar			Datos imputados					
2000	-0,0135480	0,0214260	0,0005628	-0,0135480	0,0214260	0,0005628			
2001	-0,0108230	0,0213740	0,0021686	-0,0108230	0,0213740	0,0021686			
2002	-0,0104060	0,0222150	0,0001867	-0,0104060	0,0222150	0,0001867			
2003	-0,0117360	0,0212160	-0,0031361	-0,0117360	0,0212160	-0,0031361			
2004	-0,0127140	0,0214100	-0,0033505	-0,0127140	0,0214100	-0,0033505			
2005	-0,0152960	0,0227840	-0,0008734	-0,0152960	0,0227840	-0,0008734			
2006	-0,0144620	0,0231670	-0,0047987	-0,0144620	0,0231670	-0,0047987			
2007	-0,0152960	0,0239840	-0,0009092	-0,0152960	0,0239840	-0,0009092			
2008	-0,0152960	0,0242830	-0,0009181	-0,0152960	0,0242830	-0,0009181			
2009	-0,0156480	0,0236900	-0,0040444	-0,0156480	0,0236900	-0,0040444			
2010	-0,0147350	0,0244480	0,0012425	-0,0147350	0,0244480	0,0012425			
2011	-0,0145260	0,0249190	0,0002500	-0,0145260	0,0249190	0,0002500			
2012				-0,0152967	0,0256096	0,0009131			
2013				-0,0152192	0,0249410	0,0002573			
2014				-0,0157603	0,0254164	0,0019170			
2015				-0,0160994	0,0254678	0,0032720			
2016				-0,0167291	0,0259351	0,0034426			
2017				-0,0166912	0,0258578	0,0030655			
2018				-0,0169022	0,0261076	0,0040264			
2019				-0,0168977	0,0261763	0,0033469			
2020				-0,0168274	0,0263685	0,0049256			
2021				-0,0171459	0,0266423	0,0043296			
2022				-0,0165898	0,0272103	0,0036871			
2023				-0,0165552	0,0270906	0,0058534			

Se- mana	E	N	U	E	N	U
2024				-0,0169618	0,0273041	0,0043042
2025				-0,0168974	0,0275517	0,0057649
2026				-0,0169861	0,0277625	0,0058349
2027				-0,0169009	0,0280456	0,0055094
2028				-0,0169566	0,0277023	0,0038567
2029	-0,0165470	0,0276570	0,0048520	-0,0165470	0,0276570	0,0048520
2030	-0,0161300	0,0275990	0,0028968	-0,0161300	0,0275990	0,0028968
2031	-0,0164820	0,0285040	-0,0002742	-0,0164820	0,0285040	-0,0002742
2032	-0,0164820	0,0271050	-0,0002325	-0,0164820	0,0271050	-0,0002325
2033	-0,0161300	0,0280980	0,0028819	-0,0161300	0,0280980	0,0028819
2034	-0,0161300	0,0285980	0,0028670	-0,0161300	0,0285980	0,0028670
2035	-0,0159210	0,0301680	0,0018417	-0,0159210	0,0301680	0,0018417
2036	-0,0135480	0,0297230	0,0003154	-0,0135480	0,0297230	0,0003154
2037	-0,0135480	0,0303220	0,0002975	-0,0135480	0,0303220	0,0002975
2038	-0,0159210	0,0308680	0,0018209	-0,0159210	0,0308680	0,0018209
2039	-0,0172520	0,0281690	-0,0014513	-0,0172520	0,0281690	-0,0014513
2040	-0,0164820	0,0290040	-0,0002891	-0,0164820	0,0290040	-0,0002891

Nota. La tabla contiene ejemplos de datos de estación en ITRF20, imputados mediante el método seasonally decomposed missing value imputation (SDM). Adaptado de la base de datos de la ITRF (2020).

Para facilitar la comprensión del proceso de imputación de datos mediante el método seasonally decomposed missing value imputation, se presentan a continuación las figuras correspondientes. Estas ilustraciones muestran el enfoque utilizado y los resultados obtenidos en el relleno de los datos faltantes.

Figura 3 Serie de tiempo de las coordenadas norte (N)



Nota. La figura muestra los datos transformados de la coordenada norte de cada semana desde el 2017 hasta el 2023. Adaptado de ITRF (2020).

Figura 5 Serie de tiempo de las coordenadas este (E)

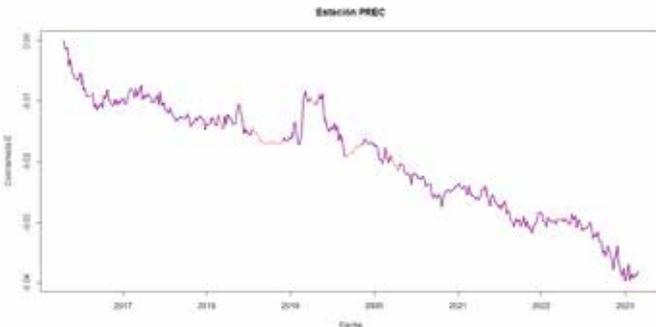
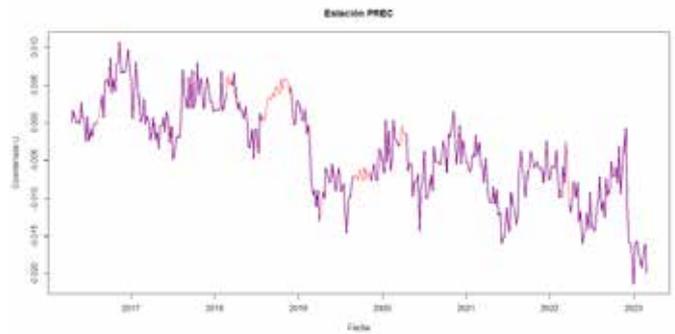


figura muestra los datos transformados de la coordenada este de cada semana desde el 2017 hasta el 2023. Adaptado de ITRF (2020).

Figura 6 Serie de tiempo de las coordenadas de altura (U)



Nota. La figura muestra los datos transformados de la coordenada altura de cada semana desde el 2017 hasta el 2023. Adaptado de ITRF (2020).

Las figuras presentadas ilustran los resultados del proceso de imputación de los datos faltantes, que ha permitido generar una serie temporal completa; así, se facilita la determinación del ruido con mayor precisión y mejora la calidad del análisis subsecuente.

Descomposición Aditiva de las Series Temporales

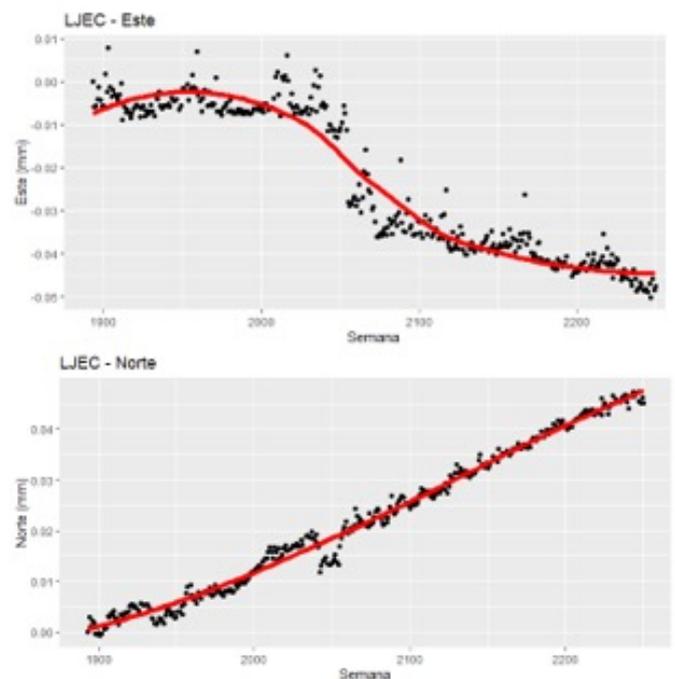
Con las series depuradas, se lleva a cabo la descomposición de estas, considerando los componentes de tendencia, estacionalidad y ruido, de modo que la suma de estos represente la serie temporal real. El modelo teórico utilizado es el siguiente:

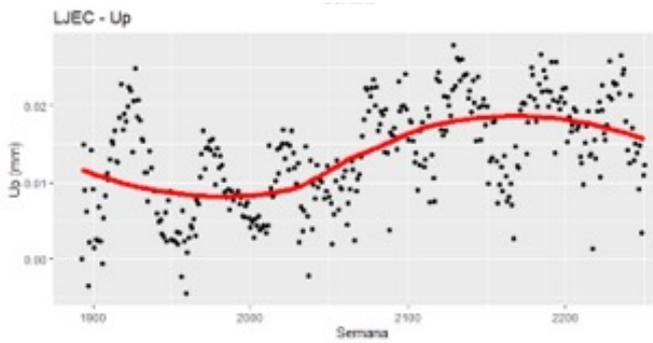
$$y_t = T_t + S_t + N_t(4)$$

En la Figura 6 se presenta el resultado de esta descomposición, donde se visualiza la tendencia lineal de la serie, obtenida a través de una regresión lineal. Asimismo, se incluye la estacionalidad basada en el análisis espectral de las series de tiempo, lo cual será fundamental para la obtención del componente de ruido en etapas posteriores del análisis.

Figura 7

Representación de la tendencia de la estación LJEC



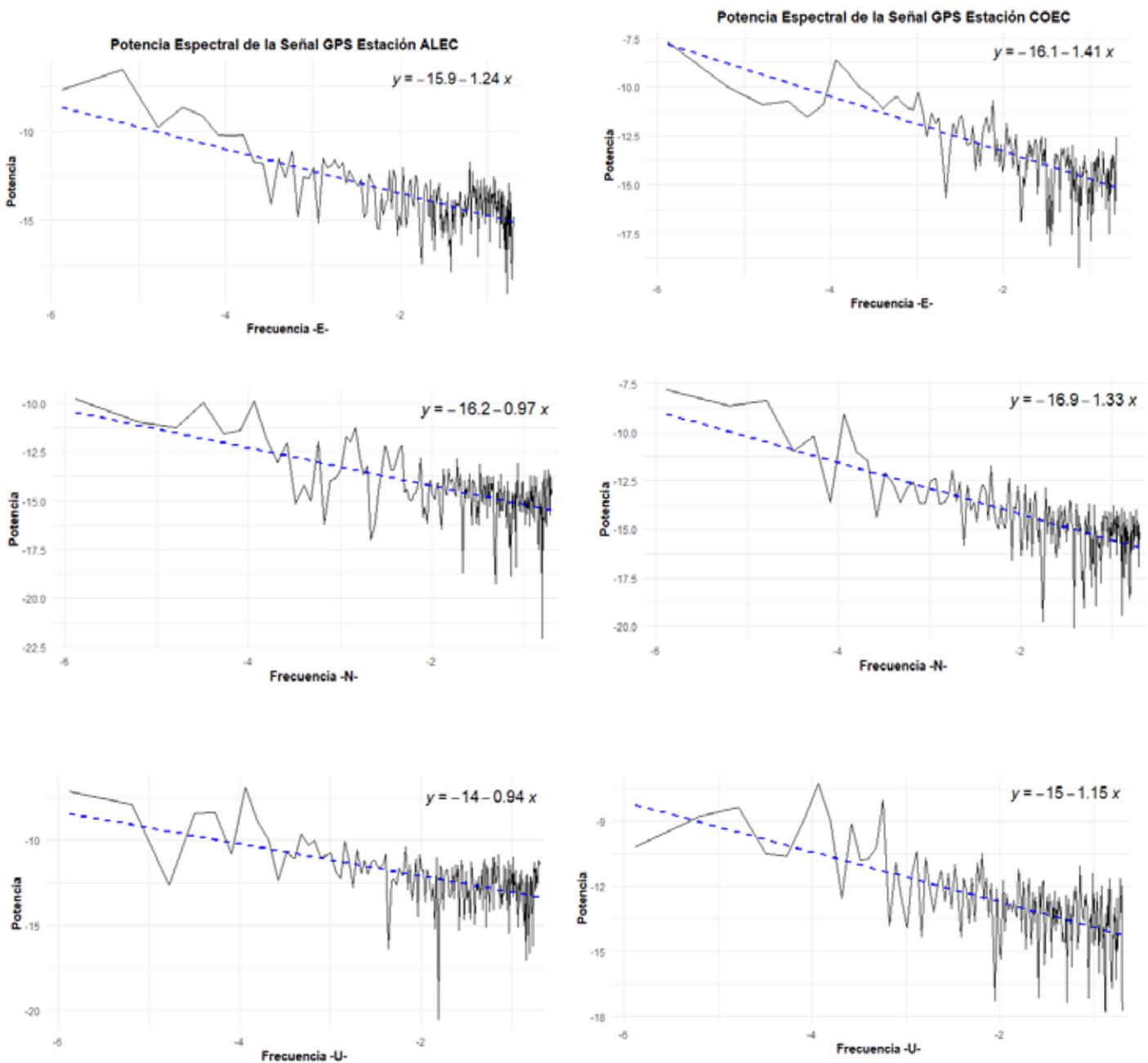


Nota. La figura muestra el gráfico de tendencia de las coordenadas norte, este y altura de la estación LJEC. Elaborado por el autor, adaptado de ITRF (2020).

Determinación del Ruido

En la Figura 8 se presenta un ejemplo de los valores de los índices espectrales de las estaciones ALEC y COEC para los componentes horizontal (e,n) y vertical (u).

Figura 8 Potencia espectral de la señal GPS estaciones ALEC y COEC



Nota. La figura muestra la potencia con la que reciben las antenas la señal del GPS. La descomposición de tendencia y estacionalidad mediante la transformada de la serie de Fourier permite identificar la cantidad de ruido. Elaborado por el autor, adaptado de ITRF (2020).

Los valores del índice espectral para las estaciones de monitoreo continuo de la red REGME seleccionadas para este estudio se muestran en la Tabla 7; estos varían entre -0.99 hasta -1,84, que corresponden a valores de ruido rosa (*flicker*) y ruido browniano o de camino aleatorio (*random walk*).

Tabla 8 Índice espectral en los componentes E, N, U

Antena	E	N	U
ALEC	-1,2371	-0,9704	-0,9399
BHEC	-1,5553	-1,3284	-1,5874
COEC	-1,4143	-1,3295	-1,1523
CUEC	-1,4950	-1,3498	-1,2700
CXEC	-1,3627	-1,1845	-0,9226
ECEC	-1,4515	-1,3269	-0,7443
EPEC	-1,4423	-0,8399	-1,0330
EREC	-1,4379	-1,1624	-1,2600
GZEC	-1,7678	-1,4540	-1,2697
IBEC	-1,4137	-1,1681	-1,1356
LJEC	-1,1308	-1,3284	-0,8665
MAEC	-1,3862	-1,1838	-1,1248
PMEC	-0,8023	-0,9588	-0,7034
PREC	-1,5287	-1,2827	-0,9626
SIEC	-1,6022	-1,2675	-0,8315
TNEC	-1,1997	-1,3182	-0,8039

Nota. Los datos representan el índice espectral de las antenas, obtenidos a partir de la descomposición de tendencia y estacionalidad mediante la transformada de la serie de Fourier, para identificar la cantidad de ruido. Elaborado por el autor, adaptado de ITRF (2020).

Los valores más altos de índice espectral corresponden a las estaciones GZEC y SIEC, con índices de -1.7678 y -1.6022, respectivamente. Ambas estaciones están situadas en las terrazas de edificios con centraje forzoso: la estación GZEC está situada sobre un pilar de concreto, mientras que la estación SIEC presenta una estructura piramidal de acero inoxidable. Los índices espectrales obtenidos para estas estaciones se acercan notablemente al ruido rojo o *random walk*, que se caracteriza por ser inversamente proporcional al cuadrado de la frecuencia. Este tipo de ruido es generalmente atribuido a la inestabilidad de los monumentos, tal y como han señalado Ward y Greenwood (2007).

La estación PMEC muestra características de ruido blanco, lo cual se refleja en un índice espectral cercano a cero. En contraste, el resto de las estaciones presentan índices espectrales cercanos a -1, lo que corresponde al ruido rosa o *flicker*, también conocido como ruido de parpadeo. Este tipo de ruido se caracteriza por un espectro de potencia inversamente proporcional a la frecuencia de la señal; se ubica entre el ruido blanco y el ruido rojo o *random wak*. Su denominación deriva de esta posición intermedia entre ambos tipos de ruido. Aunque el origen del ruido rosa no está completamente claro, es común en la mayoría de las estaciones GNSS, tal y como han señalado estudios previos (Ward y Greenwood, 2007), información que concuerda con los resultados obtenidos en esta investigación.

CONCLUSIONES

De las 52 estaciones que conforman la red REGME, 34 presentaron más del 20 % de datos faltantes o anómalos, lo que limitó su inclusión en el análisis. Por esta razón, se seleccionaron únicamente las 16 estaciones con datos más completos para el presente estudio. En este contexto, se recomienda que, en investigaciones futuras, se implementen medidas de mantenimiento y corrección en todas las estaciones, con el fin de garantizar la obtención de datos más fiables y de mayor calidad.

A partir de los análisis realizados en las 16 estaciones seleccionadas, se concluye que el 93.7 % de ellas presentan ruido rosa o *flicker*, característico de las estaciones de monitoreo continuo. Estos resultados coinciden con los hallazgos de estudios previos, por lo que se refuerza la validez de las observaciones en el contexto de este estudio.

Las estaciones GZEC y SIEC están ubicadas en las terrazas de edificios públicos. La estación GZEC está situada sobre columnas de hormigón con una altura de entre 2 y 3 metros, y registra un índice espectral de -1.7678. Por su parte, la estación SIEC se encuentra empotrada en una estructura piramidal de acero inoxidable de 6 metros, con un índice espectral de -1.6022. Estas estructuras, debido a la inestabilidad del tipo de suelo sobre el que se asientan, tienden a generar problemas como ruido de tipo rojo o *random walk*, lo que puede afectar a la precisión de las mediciones.

La estación EPEC, con un índice espectral de -1.4423, está ubicada sobre una edificación de dos plantas construida en un terreno con un nivel freático elevado. Esta característica del terreno podría influir en la calidad de los datos obtenidos, por lo que se recomienda que la estación sea objeto de una investigación adicional para evaluar en mayor profundidad los posibles efectos del nivel freático en sus mediciones y determinar así las medidas correctivas necesarias.

REFERENCIAS

- Aggarwal, K., Mukhopadhyay, S. y Tangirala, A. K. (2020). *Statistical characterization and time-series modeling of seismic noise* (arXiv:2009.01549; Número arXiv:2009.01549). arXiv. <http://arxiv.org/abs/2009.01549>
- Agnew, D. C. (1992). The time-domain behavior of power-law noises. *Geophysical Research Letters*, 19 (4), Article 4. <https://doi.org/10.1029/91GL02832>
- Altamimi, Z., Collilieux, X. y Métivier, L. (2011). ITRF2008: An improved solution of the international terrestrial reference frame. *Journal of Geodesy*, 85 (8), 457-473. <https://doi.org/10.1007/s00190-011-0444-4>
- Amiri-Simkooei, A. R., Tiberius, C. C. J. M. y Teunissen, P. J. G. (2007). Assessment of noise in GPS coordinate time series: Methodology and results. *Journal of Geophysical Research: Solid Earth*, 112 (B7), Article B7. <https://doi.org/10.1029/2006JB004913>
- Ao, M., Wang, X., Sun, Y., Wang, F., Zhang, H., Wei, L., Liu, S. y Yang, D. (2023). Large-Scale Land Deformation Monitoring over Southern California with Multi-Path SAR Data. *Remote Sensing*, 15 (1), Article 1. <https://doi.org/10.3390/rs15010143>

- Benoist, C., Collillieux, X., Rebeschung, P., Altamimi, Z., Jamet, O., Métivier, L., Chanard, K. y Bel, L. (2020). Accounting for spatiotemporal correlations of GNSS coordinate time series to estimate station velocities. *Journal of Geodynamics*, 135, 101693. <https://doi.org/10.1016/j.jog.2020.101693>
- Bhardwaj, A. (2020). Terrestrial and Satellite-Based Positioning and Navigation Systems—A Review with a Regional and Global Perspective. *Engineering Proceedings*, 2 (1), Article 1. <https://doi.org/10.3390/ecsa-7-08262>
- Dicker, L. H. (2012). *Residual variance and the signal-to-noise ratio in high-dimensional linear models* (arXiv:1209.0012; Número arXiv:1209.0012). arXiv. <http://arxiv.org/abs/1209.0012>
- Dudek, G. (2023). STD: A Seasonal-Trend-Dispersion Decomposition of Time Series. *IEEE Transactions on Knowledge and Data Engineering*, 35 (10), Article 10. <https://doi.org/10.1109/TKDE.2023.3268125>
- ITRF. (2023). ITRF I Homepage [<https://itrf.ign.fr/en/homepage>]. <https://itrf.ign.fr/en/homepage>
- Jin, S., Wang, Q. y Dardanelli, G. (2022). A Review on Multi-GNSS for Earth Observation and Emerging Applications. *Remote Sensing*, 14(16), Article 16. <https://doi.org/10.3390/rs14163930>
- Kaczmarek, A. y Kontny, B. (2018). Identification of the Noise Model in the Time Series of GNSS Stations Coordinates Using Wavelet Analysis. *Remote Sensing*, 10 (10), Article 10. <https://doi.org/10.3390/rs10101611>
- Leiva, C. (2007, junio). *SIRGAS - ECUADOR*. https://www.sirgas.org/fileadmin/docs/Boletines/Bol12/14_Actividades_SIRGAS_en_Ecuador_Leiva.pdf
- Li, Z., Lu, T., Yu, K. y Wang, J. (2023). Interpolation of GNSS Position Time Series Using GBDT, XGBoost, and RF Machine Learning Algorithms and Models Error Analysis. *Remote Sensing*, 15 (18), Article 18. <https://doi.org/10.3390/rs15184374>
- Luna Ludeña, M. (2017a). *Contribución a la definición de un modelo de velocidades de la corteza terrestre para Ecuador a partir de datos GNSS* [Tesis de doctorado, Universidad Politécnica de Madrid]. <https://doi.org/10.20868/UPM.thesis.48790>
- Luna Ludeña, M. (2017b). *Contribución a la definición de un modelo de velocidades de la corteza terrestre para Ecuador a partir de datos GNSS* [Tesis de doctorado, Universidad Politécnica de Madrid]. <https://doi.org/10.20868/UPM.thesis.48790>
- Mataija, M., Pogarčić, M. y Pogarčić, I. (2014). Helmert Transformation of Reference Coordinating Systems for Geodesic Purposes in Local Frames. *Procedia Engineering*, 69, 168-176. <https://doi.org/10.1016/j.proeng.2014.02.217>
- TM-2555.pdf. (s. f.).
- Tran, D. T., Nocquet, J.-M., Luong, N. D. y Nguyen, D. H. (2023). Determination of Helmert transformation parameters for continuous GNSS networks: A case study of the Géoazur GNSS network. *Geo-spatial Information Science*, 26 (1), Article 1. <https://doi.org/10.1080/10095020.2022.2138569>
- Wang, W., Zhao, B., Wang, Q. y Yang, S. (2012). Noise analysis of continuous GPS coordinate time series for CMONOC. *Advances in space research*, 49 (5), Article 5.
- Ward, L. M. y Greenwood, P. E. (2007). 1/f noise. *Scholarpedia*, 2 (12), 1537. <https://doi.org/10.4249/scholarpedia.1537>
- Zamora et al. (2021) y Altamimi et al. (2011). Parámetros de transformación entre los marcos geodésicos CR05 y CR-SIRGAS contemplando diferentes soluciones ITRF/Transformation parameters between the CR05 and CR-SIRGAS geodetic frames contemplating different ITRF solutions. *Ingeniería*, 31 (1), Article 1.
- Zhang, Q., Niu, X., Chen, Q., Zhang, H. y Shi, C. (2013). Using Allan variance to evaluate the relative accuracy on different time scales of GNSS/INS systems. *Measurement Science and Technology*, 24 (8), Article 8. <https://doi.org/10.1088/0957-0233/24/8/085006>
- Zhang, Z., Li, B., Shen, Y. y Yang, L. (2017). A noise analysis method for GNSS signals of a standalone receiver. *Acta Geodaetica et Geophysica*, 52 (3), Article 3. <https://doi.org/10.1007/s40328-016-0189-x>