

Jornadas de Automática

Estimación en tiempo real de frecuencia en sistemas eléctricos mediante kriging regularizado ℓ_1

Moreno-Blazquez, Carlos*, Fele, Filiberto, Alamo, Teodoro^a

^aDpto. de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Av. de los Descubrimientos s/n, 41092, Sevilla, España.

To cite this article: Moreno-Blazquez, C., Fele, F., Alamo, T. 2025. Real-time frequency estimation in power systems using ℓ_1 -regularized kriging. *Jornadas de Automática*, 46. <https://doi.org/10.17979/ja-cea.2025.46.12250>

Resumen

En este trabajo abordamos el problema de predicción en tiempo real de frecuencia en la red de distribución. Para esto, proponemos una formulación basada en datos utilizando el método kriging ordinario. Presentamos un oráculo de predicción no paramétrico que permite estimar la trayectoria futura de la frecuencia de red directamente a partir de datos históricos de intensidad y frecuencia. Demostramos la efectividad de nuestra metodología mediante un modelo simulado de un nodo de la red de distribución.

Palabras clave: Aprendizaje para el control, Métodos no paramétricos, Redes eléctricas inteligentes, Monitoreo y control de restricciones y seguridad, Control de recursos de energía renovable, Control basado en datos.

Real-time frequency estimation in power systems using ℓ_1 -regularized kriging

Abstract

We address the real-time forecast of frequency deviations in the grid. We take a data-driven approach and rely on ordinary kriging interpolation. We present a nonparametric prediction oracle which predicts the grid frequency from past frequency trajectories and current injections. We showcase our approach on a simulated distribution network node.

Keywords: Learning for control, Nonparametric methods, Smart grids, Constraint and security monitoring and control, Control of renewable energy resources, Data-based control.

1. Introducción

La frecuencia de la red eléctrica es un indicador del equilibrio instantáneo entre generación y demanda. Las desviaciones abruptas que pueden manifestarse en cuestión de segundos tras una perturbación reflejan la complejidad dinámica del sistema eléctrico interconectado y su vulnerabilidad frente a eventos disruptivos (Wang et al., 2020). Este hecho pone de manifiesto la necesidad de contar con predicciones de alta resolución temporal que posibiliten una detección temprana de inestabilidades, permitiendo así la activación preventiva de mecanismos de control, como la desconexión de carga o el ajuste dinámico de generación (North American Electric Re-

liability Corporation (NERC), 2020; Denninger et al., 2021).

La pertinencia de esta capacidad predictiva se ha visto subrayada por el apagón que afectó a la Península Ibérica el 28 de abril de 2025. A las 12:33 CET, la pérdida casi simultánea de generación en las provincias de Granada, Badajoz y Sevilla alcanzó cerca de 2,2 GW en apenas 20 segundos, sin que los mecanismos de defensa automáticos fueran capaces de detener la caída de frecuencia antes de alcanzar los 48,0 Hz, momento en el cual entró en funcionamiento el esquema de desconexión de carga (Wikipedia contributors, 2025; Diario de Sevilla, 2025). En los treinta minutos anteriores, ya se habían detectado dos periodos de oscilaciones de potencia y frecuencia (entre 12:03 y 12:07 CET, y entre 12:19 y 12:21 CET), si

*Autor para correspondencia: cmb@us.es

bien en ese momento los valores permanecían dentro de los rangos operativos y no dieron lugar a activaciones preventivas (ENTSO-E, 2025).

En este contexto, la capacidad para anticipar la evolución de la frecuencia con alta resolución temporal se convierte en una herramienta clave para mitigar eventos catastróficos. Tecnologías avanzadas de medida como las Unidades de Medida Fasorial (PMUs) o los Registradores de Perturbaciones de Frecuencia (FDRs), sincronizadas mediante GPS y operativas con frecuencias de muestreo de hasta 120 Hz, han permitido observar de forma detallada estos transitorios rápidos (Xu and Overbye, 2015; Dominguez et al., 2023). No obstante, el valor operativo de estos dispositivos depende de forma crítica de la existencia de modelos predictivos que interpreten sus datos en tiempo real, facilitando decisiones proactivas.

Para esto, los enfoques no paramétricos basados en datos emergen como una posible solución. Estos métodos extraen patrones directamente de las mediciones, capturando así la variabilidad inherente al sistema en condiciones reales Ekmwenrenren et al. (2023); Kruse et al. (2020).

Dentro de este marco, la metodología de interpolación kriging representa un enfoque prometedor. Esta técnica proporciona predicciones insesgadas y de varianza mínima a partir de datos dispersos mediante interpolaciones ponderadas (Cressie, 1993). Sin embargo, su implementación en entornos operativos presenta desafíos computacionales significativos, particularmente en redes de gran escala donde el número de muestras y ubicaciones puede ser elevado.

En consonancia con estos retos, este trabajo propone una reformulación dual del problema de kriging ordinario regularizado, combinada con métodos de optimización acelerada de primer orden con mecanismos de reinicio. Esta estrategia permite garantizar la convergencia rápida dentro de las restricciones temporales impuestas por la operación en tiempo real, sin sacrificar la precisión de las predicciones.

2. Formulación del problema

2.1. Hipótesis de trabajo

Se dispone de un conjunto de observaciones $\mathcal{D} = \{(y_1, z_1), \dots, (y_N, z_N)\}$ de un proceso $y(z) \in \mathbb{R}^{n_y}$, donde $z \in \mathbb{R}^n$ representa, en general, el dominio temporal y espacial del proceso. Siguiendo el método *kriging ordinario* (OK), consideramos para este proceso el modelo (Cressie, 1993, §3.2)

$$y(z) = \mathbb{E}[y(z)] + \delta(z),$$

donde $\delta(z)$ es un proceso aleatorio estacionario de media cero que capta la variabilidad del proceso. Este modelo se basa en una hipótesis de estacionariedad, según la cual $\mathbb{E}[y(z)] = \bar{c} \in \mathbb{R}^{n_y}$ para todo z , donde \bar{c} es una constante desconocida.

Utilizando la información incorporada en la base de datos \mathcal{D} , se quiere predecir el valor $y(z_0)$ que toma el proceso en un nuevo punto z_0 , como combinación lineal de las observaciones

$$\hat{y}(z_0) := \sum_{i=1}^N \lambda_i y_i. \quad (1)$$

En el resto de esta sección, describiremos como obtener la ponderación $\{\lambda_i\}_{i=1}^N$ según el enfoque estadístico kriging.

2.2. Variograma

Las fluctuaciones del residuo $\delta(z)$ se describen mediante el *semivariograma* $\gamma(h)$, el cual se define como

$$2\gamma(h) = \text{Var}(\delta(z+h) - \delta(z)) = \mathbb{E}[(\delta(z+h) - \delta(z))^2], \quad (2)$$

y cuantifica cómo disminuye la similitud entre dos puntos a medida que la separación vectorial h aumenta. La ecuación (2) admite la siguiente interpretación: cuando dos puntos están cerca ($\|h\|$ pequeño), la diferencia $\delta(z+h) - \delta(z)$ tiende a ser pequeña y, por tanto, $\gamma(h)$ es bajo; a medida que $\|h\|$ crece, las observaciones se hacen menos parecidas y $\gamma(h)$ aumenta, hasta estabilizarse en un valor (*sill* en inglés) equivalente a la varianza completa para un proceso estacionario (Cressie, 1993). Un salto en $h \rightarrow 0^+$ —denominado *nugget*—refleja la variabilidad a escalas tan pequeñas que no son captadas por la discretización del muestreo, o bien el ruido de medición. Este comportamiento creciente y asintótico de $\gamma(h)$ permite modelar la autocorrelación del proceso sobre su dominio.

Para estimar empíricamente el semivariograma, se agrupan las muestras por intervalos de distancia $\{h_1, h_2, \dots, h_M\}$ en su dimensión espacio-temporal. En concreto, para todo los pares de datos cuya distancia $\|z_i - z_j\|$ esté incluida en el intervalo $(h_{k-1}, h_k]$, calculamos el variograma de momentos

$$\hat{\gamma}(h_k) = \frac{1}{2|N(h_k)|} \sum_{(i,j) \in N(h_k)} (y_i - y_j)^2, \quad (3)$$

donde $N(h_k)$ agrupa los pares (i, j) que cumplen $h_{k-1} < \|z_i - z_j\| \leq h_k$. Este estimador proporciona un conjunto de pares dispersos $\{(h_k, \hat{\gamma}(h_k))\}_{k=1}^M$ que, al graficarlos contra la distancia, revelan la autocorrelación del proceso.

Sobre este conjunto se ajusta luego un modelo $\gamma(h; \eta, \theta, \alpha)$ parametrizado por el nugget η , el sill θ y el rango α (del inglés *range*, distancia a la que $\gamma(h)$ alcanza el sill). En el estudio numérico presentado en la sección 5 se hace uso de la formulación gaussiana (véase Fig. 2), que corresponde a:

$$\gamma(h; \eta, \theta, \alpha) := \eta + (\theta - \eta) \left(1 - \exp\left(-\frac{h^2}{\left(\frac{4}{7}\alpha\right)^2}\right) \right). \quad (4)$$

2.3. Predicción por kriging ordinario

Con el modelo de variograma obtenido, se procede a formular el problema de optimización relativo al problema OK, con el objetivo de obtener los coeficientes $\{\lambda_i\}_i^N$ en (1) que *minimicen la varianza del error de predicción*.

Bajo la hipótesis de estacionariedad del proceso, se tiene que $\mathbb{E}[y_i] = \mathbb{E}[y_j] = \bar{c}$ para todo $i, j \in \{1, \dots, N\}$. El método OK exige que $\sum_{i=1}^N \lambda_i = 1$. Por consecuencia, la estimación obtenida por (1) es insesgada, pues $\mathbb{E}[\hat{y}(z_0)] = \sum \lambda_i \mathbb{E}[y_i] = \sum \lambda_i \bar{c} = \bar{c}$. Entonces, el problema OK se expresa como

$$\{\lambda_i^*\}_{i=1}^N = \arg \min_{\lambda_1, \dots, \lambda_N} \frac{1}{2} \text{Var}[y(z_0) - \hat{y}(z_0)] \quad (5a)$$

$$\text{sujeto a } \sum_{i=1}^N \lambda_i = 1. \quad (5b)$$

Observamos que el error de predicción se define como

$$e := y(z_0) - \hat{y}(z_0) = y(z_0) - \sum_{i=1}^N \lambda_i y_i.$$

Este tiene esperanza nula, pues la igualdad $\mathbb{E}[e] = \mathbb{E}[y(z_0) - \hat{y}(z_0)] = 0$ se infiere directamente de la restricción $\sum_{i=1}^N \lambda_i = 1$ y de la suposición de media nula del ruido $\mathbb{E}[\delta_i] = 0$, para todo $i \in \{1, \dots, N\}$. Entonces, la varianza se puede escribir como

$$\text{Var}[e] = \mathbb{E}[e^\top e] = \lambda^\top \Gamma \lambda - 2\gamma_0^\top \lambda,$$

donde $\lambda = (\lambda_i^*)_{i=1}^N$, Γ es la matriz de (semi)varianzas entre observaciones, cuyas entradas son $\Gamma_{ij} = \gamma(\|z_i - z_j\|)$, y γ_0 es un vector columna cuya entrada i -ésima es $\gamma(\|z_i - z_0\|)$, respecto al nuevo punto de predicción z_0 . De esta forma, (5) se expresa como:

$$\min_{\lambda \in \mathbb{R}^N} \lambda^\top \Gamma \lambda - 2\gamma_0^\top \lambda, \quad (6a)$$

$$\text{sujeto a } \mathbf{1}^\top \lambda = 1. \quad (6b)$$

La solución de (6) se puede obtener imponiendo las condiciones de optimalidad, que resultan en el sistema de ecuaciones lineales

$$\begin{cases} \sum_{i=1}^N \lambda_i \gamma(z_i - z_j) + \mu = \gamma(z_j - z_0), & j = 1, \dots, N, \\ \sum_{i=1}^N \lambda_i = 1, \end{cases} \quad (7)$$

donde $\mu \in \mathbb{R}$ es el multiplicador de Lagrange asociado a la restricción (5b). Este sistema se recoge en formato matricial como

$$\begin{pmatrix} \Gamma & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma_0 \\ 1 \end{pmatrix}, \quad (8)$$

donde $\mathbf{1}$ indica el vector columna de unos, de dimensión adecuada.

2.4. OK con regularización de tipo lasso

Para introducir esparsidad y robustez en la solución del problema, evitando así el sobreajuste, añadimos un término de regularización ℓ_1 (Tibshirani and Wasserman, 2017):

$$\min_{\lambda \in \mathbb{R}^N} \lambda^\top \Gamma \lambda - 2\gamma_0^\top \lambda + \varrho \|\lambda\|_1, \quad (9a)$$

$$\text{sujeto a } \sum_{i=1}^N \lambda_i = 1, \quad (9b)$$

donde $\varrho > 0$ regula el nivel de esparsidad, permitiendo seleccionar automáticamente un subconjunto reducido de datos sin dejar de explotar la estructura de dependencia espacio-temporal. Sin embargo, este término de regularización dificulta la resolución del problema, el cual deja de ser diferenciable. Notamos que es posible en este caso obtener una versión diferenciable de (9) utilizando una formulación aumentada; sin embargo, esta se caracterizaría por un alto número de variables de decisión ($2N$ en su versión primal, $2N+1$ en su versión dual) que—junto a la estructura densa de la matriz Γ —puede resultar inviable para aplicaciones que requieran una solución en escalas de tiempo pequeñas.

3. Resolución eficiente del OK regularizado

Moreno-Blazquez et al. (2024) introducen un método para resolver de forma numéricamente eficiente una variante del problema kriging caracterizada por un enfoque estocástico

naive. Este enfoque propone construir una matriz Γ diagonal, asumiendo hipótesis análogas a las utilizadas en la regresión por mínimos cuadrados ordinarios.

Esta simplificación hace que la versión dual del problema de optimización asociado sea separable, y la solución presentada en la sección 3 de Moreno-Blazquez et al. (2024) logra tiempos de ejecución del orden de milisegundos incluso para bases de datos muy grandes. Sin embargo, al renunciar a las correlaciones cruzadas definidas por el semivariograma, el enfoque *naive* pierde las propiedades de coherencia y robustez estadística del kriging clásico.

En el caso del kriging ordinario, la matriz Γ es densa. Para facilitar la solución del problema de forma numéricamente eficiente, proponemos recurrir a la descomposición espectral de Γ

$$\Gamma = Q D Q^\top, \quad D = \text{diag}(d_1, \dots, d_N),$$

donde las columnas de $Q \in \mathbb{R}^{N \times N}$ son los autovectores de Γ y D contiene sus autovalores. Definiendo la nueva variable $v = Q^\top \lambda$, y $\xi = Q^\top \gamma_0$, obtenemos

$$\min_{v \in \mathbb{R}^N} v^\top D v - 2\xi^\top v, \quad (10a)$$

$$\text{sujeto a } \mathbf{1}^\top Q v = 1. \quad (10b)$$

La formulación así obtenida es equivalente a (6). Esta versión diagonalizada del problema permite aplicar métodos basados en gradiente muy eficaces, como el que se describe en Moreno-Blazquez et al. (2024).

Proponemos ahora inducir esparsidad en el espacio transformado, evitando así el sobreajuste de la solución. Incorporamos un término ℓ_1 sobre v . El problema regularizado queda

$$\min_{v \in \mathbb{R}^N} v^\top D v - 2\xi^\top v + \bar{\varrho} \|v\|_1, \quad (11a)$$

$$\text{sujeto a } \mathbf{1}^\top Q v = 1, \quad (11b)$$

donde $\bar{\varrho} > 0$ controla el grado de esparsidad en el nuevo marco de referencia. De esta forma, obtenemos un esquema computacionalmente eficiente, interpretable y robusto para OK regularizado. En las siguientes secciones, se explicará como explotar la separabilidad que ofrece el enfoque dual del problema (11) de manera que pueda llegar a utilizarse el algoritmo presentado en Moreno-Blazquez et al. (2024), computacionalmente eficiente, para resolver el nuevo problema en el espacio de coordenadas transformado.

3.1. Enfoque dual

Observamos que la solución de (11) se puede calcular eficientemente confiando en su reformulación dual, ya que el problema cumple las condiciones para la dualidad fuerte (Beck, 2017, Apend. A). En este caso, el número de variables de decisión duales es igual al número de restricciones de igualdad—en caso de OK solo existe una—, una cantidad significativamente menor que el número de variables primales N , el tamaño de la base de datos.

Ahora, sea $\mu \in \mathbb{R}$ la variable dual asociada a la única restricción de igualdad de (11). Esto conduce a la función dual

$$\phi(\mu) = -\mu + \min_{\lambda_1, \dots, \lambda_N} \sum_{i=1}^N d_i v_i^2 - 2\xi_i v_i + \varrho_i |v_i| + \mu v_i. \quad (12)$$

Para μ fijado, el problema (12) es separable, y los minimizadores pueden calcularse para $i = 1, \dots, N$ como

$$v_i^*(\mu) = \underset{v_i \in \mathbb{R}}{\operatorname{argmín}} d_i v_i^2 - 2\xi_i v_i + \varrho_i |v_i| + \mu v_i. \quad (13)$$

Este problema de optimización escalar posee la siguiente solución en forma cerrada (Moreno-Blazquez et al., 2024, Prop. 2):

$$v_i^*(\mu) = \frac{\operatorname{sign}(\mu - 2\xi_i)}{2d_i} \max\{0, \varrho - |\mu - 2\xi_i|\}. \quad (14)$$

4. Resolución algorítmica del enfoque dual

Es posible demostrar que $\phi(\mu)$ satisface

$$\phi(\mu + \Delta\mu) \geq \phi(\mu) + \Delta\mu^\top g(\mu) - \frac{1}{4} \Delta\mu^\top \mathbf{1}_Q D^{-1} \mathbf{1}_Q^\top \Delta\mu, \quad (15)$$

para cada $\mu, \Delta\mu \in \mathbb{R}$, donde $g(\mu) := -1 + \sum_{i=1}^N v_i^*(\mu)$ y $\mathbf{1}_Q = \mathbf{1}^\top Q$ (Beck, 2017). Utilizando esta propiedad de forma iterativa, partiendo de cualquier μ , podemos maximizar el lado derecho de (15) con respecto a $\Delta\mu$ en pasos sucesivos; esto nos permitirá recuperar el maximizador μ^* de (12) a medida que $\Delta\mu^* \rightarrow 0$. Este último puede obtenerse explícitamente al diferenciar el lado derecho en (15) y resolver para $\Delta\mu$, como

$$\Delta\mu^* = \Omega^{-1} g(\mu), \quad (16)$$

con $\Omega := \frac{1}{2} \mathbf{1}_Q D^{-1} \mathbf{1}_Q^\top$.

El Algoritmo 1 describe la implementación propuesta para la solución de (11) ($\psi(\cdot)$ expresa el lado derecho en (14)).

Algorithm 1

- 1: **Entradas** $\epsilon > 0$.
- 2: **Inicialización** $k = 0$, $\mu^0 = \mathbf{0}_{n+1}$.
- 3: **repeat**
- 4: $k = k + 1$.
- 5: **for** $i = 1, \dots, N$ **do**
- 6: $v_i^k = \psi(d_i, \varrho_i, r_i^\top \mu^{k-1} - 2\xi_i)$.
- 7: **end for**
- 8: $g^k = -1 + \sum_{i=1}^N v_i^k$.
- 9: $\mu^k = \mu^{k-1} + \Omega^{-1} g^k$.
- 10: **until** $\|g^k\| < \epsilon$
- 11: **Salidas** $\lambda^* \leftarrow Q v^k$

Utilizando el resultado de la Prop. 2 en Moreno-Blazquez et al. (2024), Algoritmo 1 calcula, para el valor de μ en cada iteración $k = 0, 1, 2, \dots$, los valores candidatos para $\{v_i^*(\mu)\}_{i=1}^N$ (pasos 5–7), utilizados para obtener el gradiente $g(\mu)$ (paso 8). Posteriormente, μ se actualiza según (16) (paso 9). Este algoritmo posee una tasa de convergencia de $\mathcal{O}(\frac{1}{k})$ (Moreno-Blazquez et al., 2024). Los tiempos de ejecución descritos en la sección 5 se han logrado utilizando una variante acelerada y con *restart* del Algoritmo 1 (Alamo et al., 2019).

5. Resultados numéricos

Evaluamos el método de kriging propuesto en un escenario simplificado que modela las inyecciones de potencia en un

punto de conexión común (PCC), donde la red de distribución se conecta con la red principal. Los resultados demuestran la eficacia de este enfoque para capturar dinámicas transitorias a escalas de fracciones de segundos, prediciendo las trayectorias de frecuencia en un horizonte temporal de 0,5 s. Las predicciones se obtienen a partir de datos históricos de sensores—en particular, las inyecciones de corriente en el marco dq y la frecuencia. También mostramos cómo los tiempos de cálculo asociados son compatibles con aplicaciones en tiempo real.

La simulación se implementa en MATLAB Simscape (Fig. 1); los parámetros eléctricos y de muestreo utilizados en este estudio de caso se enumeran en la Tabla 1. La frecuencia de muestreo seleccionada es $f_s = 80$ Hz, una década por encima de la frecuencia de corte del sistema, estimada mediante análisis FFT. Con esta frecuencia de muestreo, el horizonte temporal de 0,5 ms corresponde a $N_p = 40$ pasos discretos.

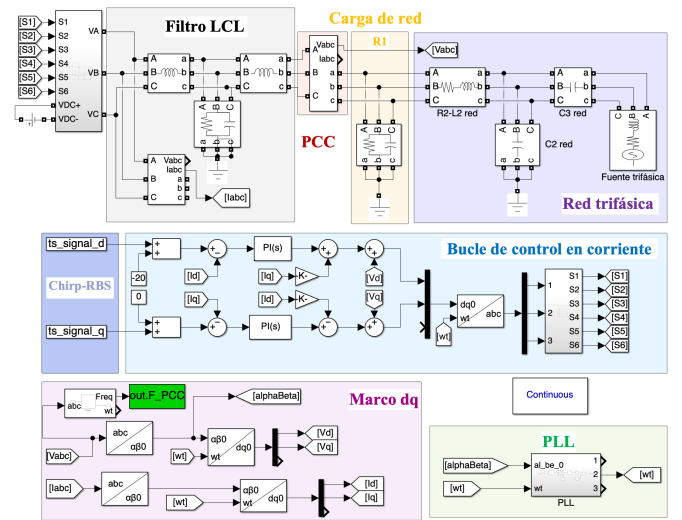


Figura 1: Modelo MATLAB Simscape del caso de estudio.

Tabla 1: Parámetros eléctricos del experimento numérico.

Parámetro	Símbolo	Valor
Valores base	V_b, S_b, f_b	380 V, 1,5 kVA, 50 Hz
Valores del filtro	L_{f1}, L_{f2}	0,027 p.u., 0,008 p.u.
Valores del filtro	C_f, R_f	0,727 p.u., 0,031 p.u.
Resistencia de carga	R_1, C_1	2 p.u., 0,05 p.u.
Línea 1	R_2, L_2, C_2	0,015 p.u., 0,15 p.u., 0,05 p.u.
Línea 2	R_3, L_3, C_3	0,015 p.u., 0,15 p.u., 10 p.u.
Frecuencia de muestreo	f_s	80 Hz
Horizonte de predicción	N_p	40 (500 ms)

Consideramos una descripción ARX de la dinámica del sistema (modelo autorregresivo con entradas exógenas):

$$y(t+1) = F(y(t), \dots, y(t-n_a), u(t), \dots, u(t-n_b)) + e(t), \quad (17)$$

donde $y(t)$ es la frecuencia en Hz en el instante t , $u(t) := (i_d(t), i_q(t))$ representa la t -ésima entrada de corriente en el marco dq , y $e(t)$ es el error de modelado. Siguiendo un enfoque no paramétrico, postulamos que la ecuación (17) se cumple para alguna función F cuyo único parámetro especificado es el orden del modelo $n_a, n_b \in \mathbb{N}$ (número de valores pasados de salida y de entradas, respectivamente).

Denotamos la trayectoria de frecuencia a N_p pasos predicha en el instante t como $\hat{y}(t+l+1|t)$, con $0 \leq l < N_p$. Esta se obtiene aplicando recursivamente un predictor de un paso adelante (véase, por ejemplo, Zhang and Ljung (2004)), que se expresa como:

$$\hat{y}(t+l+1|t) = F(y(t+l|t), \dots, y(t-n_a|t), u(t+l|t), \dots, u(t+l-n_b|t)), \quad (18)$$

donde, para $0 \leq m \leq n_a$,

$$y(t+l-m|t) := \begin{cases} \hat{y}(t+l-m|t), & \text{si } t+l-m > t, \\ y(t+l-m), & \text{en caso contrario,} \end{cases}$$

y, de forma análoga, $u(t+l-m|t)$ representa valores de entrada planificados si $t+l-m > t$, y valores pasados en caso contrario.

Para identificar la dinámica de la red y obtener el conjunto de datos \mathcal{D} , excitamos el sistema mediante una combinación de señales chirp y de secuencia pseudoaleatoria (PRS), generadas en el marco referencial dq . Estas se aplicaron como perturbaciones aditivas a la referencia del lazo de control de corriente. Las señales se diseñaron para cubrir el intervalo de frecuencias [0, 30] Hz, abarcando la frecuencia de resonancia del sistema. Esta elección garantiza que la dinámica quede persistentemente excitada (Persis and Tesi, 2019).

Los parámetros de la señal de excitación se ajustaron de modo que la operación normal de la red no se viera afectada durante la adquisición de datos. En particular, nos aseguramos de cumplir los requisitos de calidad de señal establecidos por el IEEE (IEEE Power and Energy Society, 2021) y las normas europeas (CENELEC (European Committee for Electrotechnical Standardisation), 2011), que limitan la distorsión armónica total (THD) en tensión al 8% medido sobre 10 ciclos de la señal. Nuestro análisis de distintos segmentos de 10 ciclos confirmó que la THD de la señal excitada no superó el 7,36%, garantizando así el cumplimiento de la reglamentación y la integridad operativa de la red.

Mediante este experimento generamos 79792 pares de entrada-salida, que se dividieron aleatoriamente para obtener un conjunto de entrenamiento \mathcal{D} de cardinalidad $N = 71813$ y un conjunto de validación independiente de $N_v = 7979$ elementos. Para cada punto de consulta z_0 del conjunto de validación, restringimos el conjunto de datos de entrenamiento a los $N^\circ = 250$ elementos de \mathcal{D} más cercanos a z_0 (en norma euclídea). Esta selección localizada es coherente con las hipótesis estadísticas consideradas.

Para obtener la predicción de la frecuencia de red, se siguen los siguientes pasos de forma recursiva hasta conformar la trayectoria completa sobre el horizonte establecido (500 ms en este caso):

1. Identificar el conjunto de datos reducido a los N° puntos “vecinos” al estado actual (punto de consulta z_0);
2. Ajustar la matriz Γ de (semi)varianzas entre los puntos de la vecindad reducida haciendo uso del variograma teórico ajustado con antelación (véase Fig. 2);

3. Realizar el cambio de variable para proyectar el problema en el nuevo espacio dimensional, obteniendo el problema (11);
4. Determinar los pesos kriging $\{\lambda_i^*\}_{i=1}^N$ ejecutando el Algoritmo 1 hasta cumplir con la tolerancia numérica establecida;
5. Calcular la predicción en el siguiente instante utilizando la ecuación (1).

Cabe destacar que, con el enfoque propuesto, cada predicción de un solo paso requiere un tiempo de ejecución medio de 0,67 ms, pudiendo generar predicciones de N_p pasos en 20–60 ms en la mayoría de los casos.

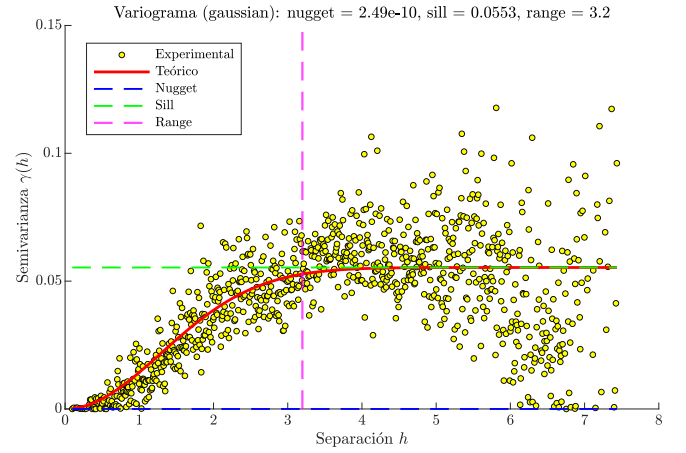


Figura 2: Variograma empírico (puntos amarillos) y ajuste teórico gaussiano (línea roja), junto con las referencias al nugget (línea azul), al sill (verde) y al rango (magenta).

A continuación, se compara el algoritmo propuesto con el enfoque *naive* de Moreno-Blazquez et al. (2024) y con el OK estándar (no regularizado). Para ello, se ha fijado una penalización uniforme $\varrho_i = 1$ en el algoritmo en Moreno-Blazquez et al. (2024), y $\bar{\varrho}_i = 0,01$ en Algoritmo 1 presentado aquí, para todos los $i = 1, \dots, N$.¹ Las pruebas han sido obtenidas con regresores de $n_a = 5$ y $n_b = 5$, correspondiente a un regresor de dimensión total 18.

Para evaluar la precisión del predictor, el error de cada trayectoria se cuantifica como el área bajo la curva del error absoluto (IAE) entre la señal real y la señal estimada, expresada en porcentaje. Sea $y_v^{(j)}(t)$ la trayectoria real y $\hat{y}^{(j)}(t)$ la trayectoria predicha para la muestra $j \in \{1, \dots, N_v\}$ (en el conjunto de entrenamiento) en tiempo continuo $\tau \in [0, T]$:

$$\zeta^{(k)} = 100 \int_0^T |y_v^{(j)}(\tau) - \hat{y}^{(j)}(\tau)| d\tau.$$

En la implementación discreta, con instantes $t_i = i \cdot T_s$ y paso constante $T_s = 1/f_s$, se aproxima esta integral mediante la regla del trapecio:

$$\zeta_d^{(k)} = 100 \sum_{i=1}^{N_p-1} \frac{|y_v^{(k)}(t_i) - \hat{y}^{(k)}(t_i)| + |y_v^{(k)}(t_{i-1}) - \hat{y}^{(k)}(t_{i-1})|}{2} T_s,$$

dando lugar a resultados como los visibles en la Fig. 3.

¹La diferencia en los valores seleccionados para $\bar{\varrho}_i$ y $\varrho_i = 1$ se debe a la distinta acción que estos coeficientes tienen sobre la respectiva función de coste. Por esta razón, dichos valores se han escogido por métodos de validación cruzada, hasta observar resultados de esparsidad comparables.

Tabla 2: Errores de predicción obtenidos para los N_v puntos de validación por los distintos algoritmos.

Algoritmo	Error ζ_d medio	Error ζ_d mediano
OK regularizado (Alg. 1)	0,798 %	0,766 %
OK estándar	4,769 %	0,968 %
Naive kriging	0,845 %	0,779 %

Los errores de predicción calculados sobre los N_v puntos de validación se muestran en la tabla 2. El enfoque propuesto, basado en kriging ordinario con proyección en un nuevo espacio transformado, aprovecha de la estructura estadística completa del variograma, reduciendo así el error medio en aproximadamente un 5,7 % y la mediana en un 1,8 % frente al método de Moreno-Blazquez et al. (2024). En comparación con el kriging ordinario clásico, el enfoque aquí propuesto logra una disminución del 83,3 % en el error medio y del 21,0 % en la mediana. Estos resultados cuantifican la precisión ganada al incorporar en el predictor el término de regularización ℓ_1 , incrementando su robustez y reduciendo significativamente el efecto de valores atípicos. Podemos comprobar como el término lasso, tanto en el enfoque *naive* como en el nuevo, mejora la estabilidad numérica ante autovalores pequeños de la matriz de covarianza, produciendo mejores estimaciones.

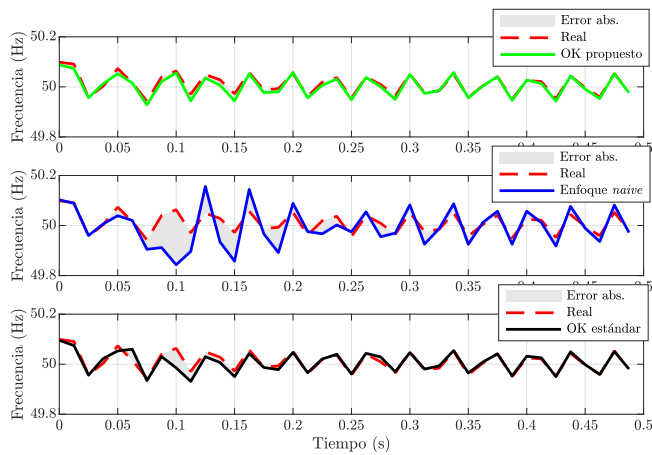


Figura 3: Comparativa de una trayectoria de validación (línea discontinua roja) con la predicción obtenida por: el enfoque OK lasso de este artículo (verde), el método en Moreno-Blazquez et al. (2024) (azul) y el OK estándar (negro). En gris se indica el área del error absoluto.

6. Conclusiones

Hemos desarrollado un predictor de frecuencia basado en kriging regularizado con norma ℓ_1 que conserva las fortalezas estadísticas del kriging (insesgidez y mínima varianza). Demostramos numéricamente que el término *lasso* contribuye a la robustez de la solución ante muestras con valores atípicos en la base de datos.

El algoritmo propuesto reduce drásticamente el coste computacional de la predicción. Todos los pasos (cálculo de semivarianzas, proyección y optimización iterativa) se realizan en milisegundos, lo que hace esta propuesta compatible con la operación en línea. Gracias a su escalabilidad y rapidez, nuestro enfoque se demuestra adecuado para aplicaciones de control y monitorización proactiva en redes eléctricas.

Agradecimientos

Esta publicación es parte del proyecto de investigación e innovación aplicada 2024/00000800, cofinanciado por la UE - Ministerio de Hacienda y Función Pública - Fondos Europeos - Junta de Andalucía - Consejería de Universidad, Investigación e Innovación, y del proyecto PID2022-142946NA-I00 financiado por MICIU/AEI /10.13039/501100011033 y por FEDER, UE. F. Fele también agradece la ayuda RYC2021-033960-I financiada por MICIU/AEI /10.13039/501100011033 y por la Unión Europea NextGenerationEU/PRTR.

Referencias

- Alamo, T., Krupa, P., Limon, D., 2019. Gradient based restart FISTA. In: 2019 IEEE 58th Conference on Decision and Control (CDC). pp. 3936–3941.
- Beck, A., 2017. First-order methods in optimization. SIAM.
- CENELEC (European Committee for Electrotechnical Standardisation), 2011. Voltage characteristics of electricity supplied by public distribution systems. European Norm EN 50160.
- Cressie, N., 1993. Statistics for spatial data. John Wiley & Sons.
- Denninger, R., Probst, L., Burger, B., Schultz, A., Dressler, Y., 2021. Visualisierung des potentials netzbildender wechsellrichter zur bereitstellung von reserveleistung. <https://www.energy-charts.info/charts/frequency/chart.htm?l=en&c=DE×lider=1&hour=11&datetimerpicker=28.04.2025>.
- Diario de Sevilla, 2025. Pérdidas de generación precedieron apagón granada badajoz sevilla. https://www.diariodesevilla.es/economia/perdidas-generacion-precidieron-apagon-granada-badajoz-sevilla_0_2003930695.html.
- Dominguez, X., Prado, A., Arboleya, P., Terzija, V., 2023. Evolution of knowledge mining from data in power systems: The big data analytics breakthrough. Electric Power Systems Research 218, 109193.
- Ekowenrenren, E., Simpson-Porco, J. W., Farantatos, E., Patel, M., Haddadi, A., Zhu, L., 2023. Data-driven fast frequency control using inverter-based resources. IEEE Transactions on Power Systems.
- ENTSO-E, 2025. Entso-e expert panel initiates the investigation into the causes of iberian blackout. <https://www.entsoe.eu/news/2025/05/09/entso-e-expert-panel-initiates-the-investigation-into-the-causes-of-iberian-blackout/>.
- IEEE Power and Energy Society, 2021. IEEE draft standard for harmonic control in electric power systems. IEEE P519/D5.1, January 2021, 1–30.
- Kruse, J., Schäfer, B., Withaut, D., 2020. Predictability of power grid frequency. IEEE Access 8, 149435–149446.
- Moreno-Blazquez, C., Fele, F., Limon, D., Alamo, T., 2024. Predicción de voltajes en la red eléctrica por interpolación kriging. Revista Iberoamericana de Automática e Informática industrial.
- North American Electric Reliability Corporation (NERC), 2020. White paper: Fast frequency response concepts and bulk power system reliability needs. <https://www.nerc.com/comm/PC/Pages/Inverter-Based-Resource-Performance-Task-Force.aspx>.
- Persis, C. D., Tesi, P., 2019. On persistency of excitation and formulas for data-driven control. In: 2019 IEEE 58th Conference on Decision and Control (CDC). pp. 873–878.
- Tibshirani, R., Wasserman, L., 2017. Sparsity, the lasso, and friends. Lecture notes from “Statistical Machine Learning,” Carnegie Mellon University, Spring.
- Wang, W., Yao, W., Chen, C., Deng, X., Liu, Y., 2020. Fast and accurate frequency response estimation for large power system disturbances using second derivative of frequency data. IEEE Transactions on Power Systems 35 (3), 2483–2486.
- Wikipedia contributors, 2025. 2025 iberian peninsula Blackout. https://en.wikipedia.org/wiki/2025_Iberian_Peninsula_blackout.
- Xu, T., Overbye, T., 2015. Real-time event detection and feature extraction using pmu measurement data. In: 2015 IEEE International Conference on Smart Grid Communications (SmartGridComm). pp. 265–270.
- Zhang, Q., Ljung, L., 2004. Multiple steps prediction with nonlinear arx models. IFAC Proceedings Volumes 37 (13), 309–314, 6th IFAC Symposium on Nonlinear Control Systems 2004 (NOLCOS 2004), Stuttgart, Germany, 1-3 September, 2004.