

# Predicción de voltajes en la red eléctrica por interpolación Kriging

Moreno-Blazquez, Carlos\*, Fele, Filiberto, Limon, Daniel, Alamo, Teodoro<sup>a</sup>

<sup>a</sup>Dpto. de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Av. de los Descubrimientos s/n, 41092, Sevilla, España.

## Resumen

En este trabajo, abordamos el problema de la predicción en línea de las trayectorias de voltaje e intensidad nodales en la red de distribución. Para esto, proponemos una formulación basada en datos utilizando la interpolación Kriging, una técnica de aprendizaje automático que ha mostrado aplicaciones prometedoras en el campo del control basado en datos. Producimos un oráculo de predicción no paramétrico que permite inferir trayectorias futuras directamente a partir de medidas de voltaje e intensidad en tiempo real. Además, proporcionamos una implementación algorítmica simple pero efectiva basada en el conocido esquema ISTA. Demostramos la efectividad de nuestra metodología para la predicción rápida (subsegundos) de la dinámica del voltaje mediante simulaciones.

*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.

## Prediction of grid voltages by means of Kriging interpolation

### Abstract

We address the problem of the online prediction of nodal voltage and current trajectories in the distribution grid. For this, we propose a data-driven formulation based on the Kriging interpolation, a machine learning technique that recently showed promising applications in the field of data-based control. We produce a nonparametric prediction oracle which allows to infer future trajectories directly from real-time voltage and current measures. We further provide a simple but effective algorithmic implementation based on the well-known ISTA scheme. We showcase the effectiveness of our methodology for the fast (subsecond) prediction of voltage dynamics through simulations.

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

## 1. Introducción

En un mundo donde las fuentes de energía renovable y las tecnologías emergentes como los vehículos eléctricos están transformando el panorama energético, la capacidad de predecir el voltaje y la frecuencia de la red eléctrica se ha vuelto crucial. El uso creciente de fuentes de generación distribuida basadas en energías renovables ha introducido desafíos significativos para mantener la estabilidad de la red, especialmente debido a la naturaleza intermitente de estas fuentes y la carga adicional introducida por los vehículos eléctricos (Milano et al., 2018). Estos cambios dinámicos pueden causar fluctuaciones en el voltaje y la frecuencia de la red, afectando la calidad del suministro eléctrico (Kaur and Vaziri, 2006).

Para abordar estos desafíos, la predicción en tiempo real del voltaje y la frecuencia de la red eléctrica es esencial. Una predicción precisa no sólo mejora el monitoreo y el control de los sistemas de energía, sino que también juega un papel fundamental en la protección de estos sistemas (Dobbe et al., 2020). Esta capacidad permite optimizar el estado de carga de los sistemas de almacenamiento de energía y controlar microrredes y plantas de energía virtuales de manera óptima, mejorando la calidad y la estabilidad del suministro eléctrico (Zufferey et al., 2020).

En el ámbito de la ingeniería de control, los problemas de estimación y predicción se han abordado mediante métodos paramétricos, que requieren la definición de un modelo basado en principios físicos o datos observados. Sin embargo, estos pueden ser limitados por la necesidad de conocer y representar con precisión las leyes físicas subyacentes y los parámetros específicos del sistema (Car et al., 2021).

Recientemente, han surgido métodos no paramétricos que ofrecen una alternativa poderosa y flexible, siendo ideales para entornos con alta variabilidad. Estos no requieren una definición explícita de un modelo, pues se basan directamente en los datos observados para realizar predicciones y tomar decisiones de control (Nadales et al., 2023; Ordonez et al., 2021; Merchán-Riveros et al., 2024). Entre estos métodos, el Kriging se ha destacado en recientes aplicaciones en el campo de la clasificación y el control (Carnerero et al., 2023). El presente estudio se centra en la aplicación de Kriging para la predicción en tiempo real de voltaje y corriente en la red eléctrica. Este enfoque no solo facilita la gestión y control de redes eléctricas en tiempo real, sino que también ofrece una robustez superior frente a las fluctuaciones y el ruido en los datos, comparado con los métodos paramétricos tradicionales.

La rápida intervención en la red eléctrica ante eventos ines-

\* Autor para correspondencia: cmb@us.es

perados a menudo requiere acciones en escalas de tiempo inferiores a un segundo (Deakin et al., 2021). Por lo tanto, es fundamental contar con la capacidad de prever con precisión la evolución del voltaje de la red en intervalos de tiempo muy pequeños, del orden de décimas de segundo, como se ha destacado en estudios previos (Gupta and Milanovic, 2006). La predicción de trayectorias del voltaje de la red se vuelve fundamental para mantener su equilibrio dinámico, gestionando eficazmente la demanda y la inserción de nuevas cargas. Además, las directivas vigentes, como las limitaciones de tensión a corto plazo (EN Task Group on Statutory Voltage Limits, 2017), enfatizan la importancia de contar con pronósticos precisos en escalas de tiempo muy pequeñas para garantizar la estabilidad y fiabilidad del sistema eléctrico.

Para predecir la trayectoria de la red en pequeñas escalas de tiempo es crucial comprender su comportamiento de forma precisa. Esto requiere datos detallados de la dinámica del sistema, utilizando excitaciones que no alteren su funcionamiento normal. El método de barrido de frecuencia (Francis et al., 2011; Huang et al., 2009), que inyecta señales sinusoidales de diferentes frecuencias para estimar la impedancia de la red, es común en aplicaciones reales. Sin embargo, este método necesita un dispositivo adicional y es lento por requerir la inyección de múltiples señales. Otros métodos, que utilizan señales de banda ancha, reducen el tiempo de medición. Por ejemplo, los métodos de inyección de impulsos (Liu et al., 2020; Céspedes and Sun, 2012), aunque rápidos, pueden excitar respuestas no lineales y dañar la red debido a su alta perturbación.

Métodos avanzados, como el uso de ruido blanco limitado en banda (Xiao et al., 2007), PRBS (Martin et al., 2013; Rojnila et al., 2017), señales chirp (Shen et al., 2013) y señales multitonales (Xiao et al., 2022), minimizan la perturbación pero aún enfrentan problemas, como comentan (Haberle et al., 2023), para estimar de forma paramétrica la impedancia de la red. Estos métodos, aunque requieren perturbaciones secuenciales y varios ciclos de medición, permiten generar datos no lineales ricos al excitar el sistema de manera controlada en frecuencia y amplitud, sin alterar demasiado la red y capturando su comportamiento no lineal inherente.

Este trabajo se enfoca en la propuesta de un método de predicción no paramétrico, Kriging, para las variables de interés en la red de distribución. De acuerdo con este enfoque, la operación de la red eléctrica se perturba con pequeñas inyecciones de corriente en el punto de acoplamiento común (PCC), al fin de excitar su dinámica y obtener un conjunto de datos adecuado para la descripción del sistema. Los datos de operación recolectados en tiempo real se transforman de forma directa en predicciones de trayectorias futuras, permitiendo su uso en aplicaciones de monitoreo de restricciones y control con requisitos estrictos de tiempo de cálculo. Como segunda contribución, se describe una posible implementación práctica, produciendo una versión *ad hoc* del algoritmo ISTA (del inglés Iterative Shrinkage-Thresholding Algorithm), especializada para la estructura del problema Kriging (1).

El documento está organizado de la siguiente manera. En la sección 2 presentamos el enfoque del problema, y su formulación dual se detalla en la sección 3.1. El algoritmo propuesto se presenta en la sección 3.2. Posteriormente, se muestra la estructura de la base de datos utilizada, presentando los modelos

NARX en la sección 4 y destacando la importancia de utilizar el marco *dq* en la sección 4.1. El análisis numérico se presenta en la sección 5, antes de las observaciones finales.

## 2. Formulación del problema Kriging

El Kriging (Cressie, 1990) es una técnica avanzada dentro de la estadística espacial, que permite realizar predicciones óptimas en espacios geográficos a partir de datos dispersos. Esta metodología, introducida por Matheron en la década de 1960 (Matheron, 1967), se fundamenta en el concepto de interpolación ponderada, tal como fue refinado y popularizado por Krige en el contexto minero sudafricano (Krige, 1981). Dada su capacidad para modelar la variabilidad espacial con precisión ha sido útil no sólo en el campo de la minería, donde las estimaciones iniciales basadas en promedios simples eran insuficientes debido a la variabilidad local del depósito, sino también en meteorología, climatología y control basados en datos (Carnerero et al., 2023).

Según Hemyari and Nofziger (1987), Kriging es una forma de promediado ponderado en la cual los pesos se eligen de manera que el error asociado con el predictor sea menor que para cualquier otra suma lineal. Sea  $\mathcal{D} = \{z_1, z_2, \dots, z_N\}$  un conjunto de datos compuesto por las muestras  $z_i \in \mathbb{R}^n$ ,  $i = 1, \dots, N$ . Dado un punto  $\bar{z} \in \mathbb{R}^n$ , estamos interesados en determinar  $\bar{z}$  en función de una combinación lineal de las  $z_i$  del conjunto de datos  $\mathcal{D}$ . Con este objetivo, podemos definir una función  $J_\gamma : \mathbb{R}^n \times \mathcal{D} \rightarrow [0, \infty)$  que evalúa la *disimilitud* entre  $\bar{z}$  y todos los  $z_i \in \mathcal{D}$ . Esta función devolverá un valor alto cuando  $\bar{z}$  sea significativamente diferente del conjunto de datos  $\mathcal{D}$ ; por otro lado, valores pequeños indicarán alta similitud. Esto se traslada formalmente al siguiente problema (Carnerero et al., 2022):

$$J_\gamma(\bar{z}, \mathcal{D}) := \min_{\lambda_1, \dots, \lambda_N} \sum_{i=1}^N w_i \lambda_i^2 + \gamma_i |\lambda_i| \quad (1a)$$

$$\text{s.t. } \bar{z} = \sum_{i=1}^N z_i \lambda_i \quad (1b)$$

$$1 = \sum_{i=1}^N \lambda_i, \quad (1c)$$

donde  $\gamma_i \geq 0$  y  $w_i > 0$ ,  $i = 1, \dots, N$ , se utilizan para ponderar la información proporcionada por cada muestra de datos. Esto permite, por ejemplo, filtrar puntos de datos ruidosos o irrelevantes. Los términos  $\gamma_i |\lambda_i|_{i=1}^N$  promueven la esparsidad. Por otro lado, la restricción (1b) define  $\bar{z}$  como una combinación lineal de las muestras  $\{z_i\}_{i=1}^N$  ponderadas por  $\{\lambda_i\}_{i=1}^N$ ; la restricción (1c) normaliza la solución.

Ahora, supongamos que están disponibles  $N$  pares de entrada-salida observados  $(u_i, y_i)$ ,  $i = 1, \dots, N$ , y sea  $\bar{u}$  una nueva entrada, cuya salida  $\bar{y}$  asociada sea *desconocida*. Definiendo  $z_i = u_i$  (por lo tanto,  $\mathcal{D} = \{u_1, u_2, \dots, u_N\}$ ), y  $\bar{z} = \bar{u}$ , resolvemos (1) para obtener  $\lambda^* := \{\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*\}$ . Con esto, podemos obtener una estimación  $\hat{y}$  de  $\bar{y}$  mediante la interpolación Kriging (Carnerero et al., 2022; Matsui and Yamakawa, 2023)

$$\hat{y} = \sum_{i=1}^N \lambda_i^*(\bar{z}, \mathcal{D}) \cdot y_i. \quad (2)$$

La función (1) es un componente clave en muchos desarrollos novedosos en el contexto de enfoques basados en datos. Formulaciones similares se han utilizado para identificación de sistemas, como en el conocido *data-enabled predictive control* (véase por ejemplo la reseña bibliográfica de Markovsky et al. (2023)), donde estos métodos directos de identificación a partir de trayectorias observadas se fundamentan en el enfoque conductual, o *behavioural approach*, de Willems (1986). Del mismo modo, enfoques similares como los conocidos métodos de optimización directa de pesos (Roll et al., 2005), permiten encontrar la combinación lineal de puntos en la base de datos que se asemeje lo mejor posible a un punto específico.

### 3. ISTA para Kriging

La formulación (1) constituye un programa de optimización estrictamente convexo sujeto a restricciones convexas; como tal, tiene una solución única  $\lambda^* := (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*)$  entre todas las posibles combinaciones lineales de las muestras de datos en  $\mathcal{D}$  que producen el punto de consulta  $\bar{z}$ .

El problema (1) se enmarca en la estructura estándar de un problema cuadrático, tras la introducción de variables auxiliares para obtener una formulación lineal equivalente del valor absoluto en (1a), resultando en un problema cuya variable de decisión tiene dimensión  $2N$ . De esta manera la solución de (1) se puede calcular utilizando una amplia variedad de *solvers* disponibles comercialmente o en acceso abierto.

#### 3.1. Enfoque dual

La adopción del enfoque primal en problemas de predicción y control está limitada por el hecho de que el número de variables de decisión en (1) crece linealmente con el tamaño  $N$  de la base de datos. Observamos que la solución de (1) se puede calcular eficientemente confiando en su reformulación dual (Merhy et al., 2018; Matsui and Yamakawa, 2023). En este caso, el número de variables de decisión duales es igual al número de restricciones de igualdad, es decir,  $n + 1$ , una cantidad que puede ser significativamente menor que el número de variables primales  $N$ .

Definiendo  $\bar{r} = [\bar{z}^\top \mathbf{1}^\top]^\top$ ,  $r_i = [z_i^\top \mathbf{1}^\top]^\top$ ,  $i = 1, \dots, N$ , podemos reescribir (1) como

$$J_y^* = \min_{\lambda_1, \dots, \lambda_N} \sum_{i=1}^N w_i \lambda_i^2 + \gamma_i |\lambda_i| \quad (3a)$$

$$\text{s.t. } \bar{r} = \sum_{i=1}^N r_i \lambda_i, \quad (3b)$$

y consideramos la siguiente suposición a lo largo del artículo

**Suposición 1.** *El problema de optimización (3) es factible, es decir, la matriz  $R := [r_1, r_2, \dots, r_N]$  tiene rango de fila completo.*

Ahora, sea  $\mu \in \mathbb{R}^{n+1}$  la variable dual asociada a las restricciones de igualdad (3b). Esto conduce a la función dual

$$\phi(\mu) = -\mu^\top \bar{r} + \min_{\lambda_1, \dots, \lambda_N} \sum_{i=1}^N w_i \lambda_i^2 + \gamma_i |\lambda_i| + (\mu^\top r_i) \lambda_i. \quad (4)$$

Fijado  $\mu$ , el problema (4) es separable, y los minimizadores pueden calcularse para  $i = 1, \dots, N$  como

$$\lambda_i^*(\mu) = \operatorname{argmín}_{\lambda_i \in \mathbb{R}} w_i \lambda_i^2 + \gamma_i |\lambda_i| + (\mu^\top r_i) \lambda_i. \quad (5)$$

Este problema de optimización escalar posee una solución en forma cerrada (Beck, 2017, Example 6.8), como se afirma en la siguiente proposición.

**Proposición 2.** *Consideremos el problema de optimización*

$$z^* = \operatorname{argmín}_{z \in \mathbb{R}} w z^2 + \gamma |z| + cz, \quad (6)$$

con  $w > 0$ ,  $\gamma \geq 0$ , y  $c \in \mathbb{R}$ . Definamos  $\psi: \mathbb{R} \rightarrow \mathbb{R}$  como

$$\psi(w, \gamma, c) := \begin{cases} \operatorname{sign}(c) \left( \frac{\gamma - |c|}{2w} \right) & \text{si } |c| > \gamma, \\ 0 & \text{en otro caso.} \end{cases} \quad (7)$$

Entonces,  $z^* = \psi(w, \gamma, c)$ .

#### 3.2. Algoritmo K-ISTA

Bajo la Suposición 1, el problema (1) cumple las condiciones para la dualidad fuerte (Bertsekas, 2009, Prop. 5.3.3), lo que significa que  $J_y^* = \max_{\mu} \phi(\mu)$ . Además, definiendo  $W := \operatorname{diag}([w_1, w_2, \dots, w_N])$ , es posible demostrar, véase por ejemplo (Beck, 2017), que  $\phi(\mu)$  satisface

$$\phi(\mu + \Delta\mu) \geq \phi(\mu) + \Delta\mu^\top g(\mu) - \frac{1}{4} \Delta\mu^\top R W^{-1} R^\top \Delta\mu, \quad (8)$$

para cada  $\mu, \Delta\mu \in \mathbb{R}^{n+1}$ , donde

$$g(\mu) := -\bar{r} + \sum_{i=1}^N r_i \lambda_i^*(\mu). \quad (9)$$

Esta propiedad proporciona una forma intuitiva de maximizar la función dual  $\phi(\mu)$ . Partiendo de cualquier  $\mu$  fijo, podemos maximizar el lado derecho de (8) con respecto a  $\Delta\mu$  a lo largo de iteraciones sucesivas; esto nos permitirá recuperar el maximizador  $\mu^*$  de (4) a medida que  $\Delta\mu^* \rightarrow 0$ . Este último puede obtenerse explícitamente al diferenciar el lado derecho en (8) y resolver para  $\Delta\mu$ , como

$$\Delta\mu^* = \Omega^{-1} g(\mu), \quad (10)$$

con  $\Omega := \frac{1}{2} R W^{-1} R^\top$ . Teniendo en cuenta que  $W \in \mathbb{R}^{N \times N}$ , con conjuntos de datos grandes, una forma eficiente en términos de memoria para calcularlo es  $\Omega = \sum_{i=1}^N \frac{1}{2w_i} r_i r_i^\top$ .

La implementación algorítmica propuesta para la solución de (1) se basa en el conocido *algoritmo de contracción y umbral iterativo* (ISTA) (véase (Beck and Teboulle, 2009; Alamo et al., 2019) y sus referencias) que especializamos a la reformulación dual particular de (1) adoptada en este artículo, utilizando el resultado de la Proposición 2.

---

**Algorithm 1** K-ISTA

---

i: **Entradas**  $\epsilon > 0$ .  
ii: **Inicialización**  $k = 0$ ,  $\mu^0 = \mathbf{0}_{n+1}$ .  
iii: **repeat**  
iv:  $k = k + 1$ .  
v: **for**  $i = 1, \dots, N$  **do**  
vi:  $\lambda_i^k = \psi(w_i, \gamma_i, r_i^\top \mu^{k-1})$ .  
vii: **end for**  
viii:  $g^k = -\bar{r} + \sum_{i=1}^N r_i \lambda_i^k$ .  
ix:  $\mu^k = \mu^{k-1} + \Omega^{-1} g^k$ .  
x: **until**  $\|g^k\| < \epsilon$   
xi: **Salidas**  $\lambda^* \leftarrow \{\lambda_i^k\}_{i=1}^N$

---

En el Algoritmo 1, primero **i**) se define la tolerancia  $\epsilon > 0$  para establecer el criterio de parada con el cuasi cumplimiento de las restricciones de igualdad. **ii**) Se inicializan el contador de iteraciones  $k = 0$  y la variable dual  $\mu^0 = \mathbf{0}_{n+1}$ . **iii**) El algoritmo itera hasta que la norma del gradiente  $g^k$  sea menor que la tolerancia impuesta  $\epsilon$ . **iv**) En cada iteración, se incrementa el contador  $k$  en uno. **v**) Para cada  $i : i = 1, \dots, N$ , en **vi**) se actualiza  $\lambda_i^k$  usando la ecuación (7). **viii**) Se recalcula el gradiente  $g^k$  como en (9), actualizando seguidamente en **ix**)  $\mu^k$  con (10). Finalmente, en **xi**) el conjunto de coeficientes que satisface la condición de parada forma el vector solución  $\lambda^*$ .

Este algoritmo hereda la simplicidad del método ISTA, demostrando una tasa de convergencia de  $\mathcal{O}(\frac{1}{k})$ . Aunque el algoritmo propuesto llega a ser compatible con computaciones en tiempo real, es posible obtener variantes optimizadas que disminuyen el número de iteraciones y, por ende, el tiempo de cómputo. Esto se demuestra en la sección 5, donde los resultados han sido obtenidos utilizando K-ISTA y una versión del mismo basada en el esquema FISTA con *restart*, que acelera la tasa de convergencia a  $\mathcal{O}(\frac{1}{k^2})$  (véase (Alamo et al., 2019) para más detalle).

#### 4. Composición del conjunto de datos NARX

El sistema a controlar se caracteriza por sus entradas,  $u \in \mathcal{U} \subset \mathbb{R}^m$ , y sus salidas  $y \in \mathcal{Y} \subset \mathbb{R}^p$ , donde  $\mathcal{U}$  e  $\mathcal{Y}$  se suponen convexos. La única información disponible de este sistema es un conjunto histórico adecuado de  $N$  pares observados de entrada-salida  $\{(u_1, \tilde{y}_1), \dots, (u_N, \tilde{y}_N)\}$ , donde  $\tilde{y}$  denota la medida posiblemente ruidosa de la salida, con el ruido asumido acotado por  $\epsilon$ .

Asumimos que la dinámica del sistema puede describirse mediante un modelo autoregresivo no lineal con entradas exógenas (NARX) (Chen and Billings, 1989; Leontaritis and Billings, 1985) como

$$y(k+1) = f(y(k), \dots, y(k-n_a), u(k), \dots, u(k-n_b)), \quad (11)$$

donde  $n_a, n_b \in \mathbb{N}$  parametrizan el horizonte para los valores pasados de salida y entrada, respectivamente. Bajo suposiciones leves sobre la observabilidad del sistema, el modelo NARX puede describir la dinámica de un sistema no lineal para horizontes suficientemente grandes. Sin embargo, pueden existir errores de modelado debido a una descripción incompleta de la dinámica real del sistema representada por  $f$ .

Sea  $\bar{n}$  el orden del sistema. Como demostrado por Levin and Narendra (1997), se puede lograr una representación perfecta del sistema seleccionando los horizontes de memoria  $n_a$  y  $n_b$  como  $n_a = n_b = 2\bar{n}$ . En caso de que el orden del sistema sea desconocido, los valores de  $n_a$  y  $n_b$  deben estimarse mediante algún método de validación cruzada.

Nos referimos a los parámetros de entrada de  $f$  en (11) como el regresor  $z(k) := (x(k), u(k))$ , donde  $x(k) := (y(k), y(k-1), \dots, y(k-n_a), u(k-1), \dots, u(k-n_b))$  es el estado en el instante de muestra  $k$ . Además, sea  $\tilde{z}(k)$  el regresor *ruidoso* que incluye las observaciones corruptas, es decir,  $\tilde{z}(k) := (\tilde{x}(k), u(k)) = (\tilde{y}(k), \tilde{y}(k-1), \dots, \tilde{y}(k-n_a), u(k), u(k-1), \dots, u(k-n_b))$ , y sea  $\mathcal{D}$  el conjunto de regresores conformado por datos entrada-salida reorganizado de la siguiente manera:

$$\mathcal{D} = \{\tilde{z}(k), k = 1, \dots, N\}. \quad (12)$$

Asimismo, las trayectorias futuras a  $N_p$  pasos de la salida, correspondientes a cada uno de los regresores  $\tilde{z}(k)$ , se almacenan en otro conjunto  $\mathcal{D}^+ = \{(\tilde{y}(k+1), \dots, \tilde{y}(k+N_p)), k = 1, \dots, N\}$ .

##### 4.1. Marco $dq$

Una vez clara la estructura del modelo NARX, se va a emplear el marco  $dq$  para representar las medidas trifásicas. Este es un marco de referencia que considera dos ejes, llamados eje *directo* y eje en *cuadratura*, que giran a una frecuencia dada (por ejemplo, la frecuencia de red). Esto permite convertir las componentes de corriente y voltaje en constantes en estado estacionario, lo que simplifica significativamente el diseño y análisis de controladores.

En el marco  $dq$ , las componentes directas ( $d$ ) y en cuadratura ( $q$ ) pueden ser controladas de manera independiente. Esta capacidad de desacoplamiento es particularmente ventajosa, por ejemplo, para la implementación de controladores proporcionales-integrales (PI) (Francis et al., 2011). Además, las componentes de alta frecuencia, o armónicos, son más fáciles de filtrar en este marco, lo que mejora la calidad de la señal y la eficiencia del sistema (Saccomando and Svensson, 2001). La facilidad de las transformaciones directas e inversas entre los marcos  $abc$  y  $dq$  también contribuye a su preferencia, ya que estas transformaciones están bien definidas y pueden implementarse fácilmente en sistemas de control digital.

Por tanto, haciendo uso de las ventajas del marco  $dq$  y conformando el conjunto de datos como en (12), nuestro objetivo es inferir una aproximación  $F$  basada en aprendizaje (implícito) del modelo NARX en (11) que, dado un regresor del sistema  $\tilde{z} = (\tilde{x}, \bar{u})$ , devuelva un sucesor  $\hat{y}^+ \in \mathcal{Y}$ . En nuestro caso,

$$\hat{y}^+ = F(\tilde{z}, \mathcal{D}) = \sum_{i=1}^N \lambda_i^*(\tilde{z}, \mathcal{D}) \cdot \tilde{y}_i, \quad (13)$$

donde  $F$  toma la forma de (2). Esta expresión hace uso de los pesos  $\{\lambda_1^*, \dots, \lambda_N^*\}$  obtenidos al resolver (1), en función de  $\tilde{z}$  y  $\mathcal{D}$ , mediante el Algoritmo 1.

#### 5. Análisis numérico

Para probar el enfoque propuesto, consideramos un caso de simulación implementado en Matlab Simulink, como puede



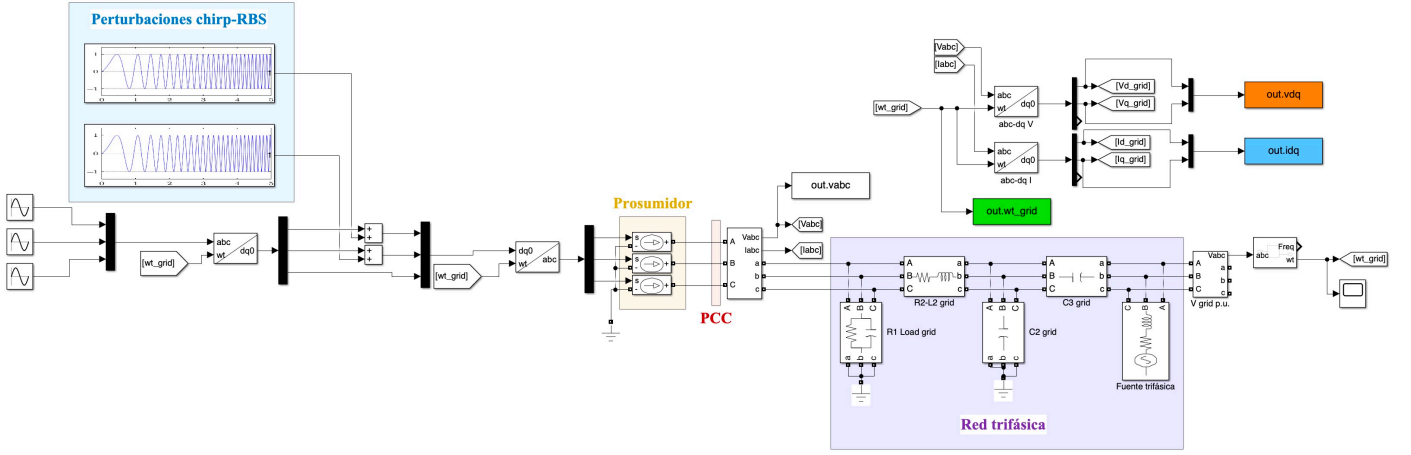


Figura 1: Modelo Matlab Simulink de un prosumer convencional conectándose a la red en el PCC. Se toman datos de corriente y voltaje en el marco  $dq$  con el objetivo de conformar el dataset  $\mathcal{D}$ . Así mismo, se recogen datos del ángulo de fase  $\omega_g t$  para poder utilizar las predicciones  $dq$  en el marco  $abc$ .

verse en la Figura 1. Lo que se pretende es estimar la trayectoria de voltaje de la red en el marco  $dq$  a partir de medidas de corriente y voltaje obtenidas de una serie de sensores situados en el PCC. Cabe comentar que los resultados se muestran en el marco  $abc$ , lo cual es posible realizando predicciones al mismo tiempo de la frecuencia de la red  $\omega_g$  haciendo uso del modelo utilizado para predecir el voltaje.

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
Carga	$R_1$	2 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	$f_s$	1 kHz
Horizonte de predicción	$N_p$	200

Los parámetros eléctricos y de medición del sistema representado en la Figura 1 pueden verse contenidos en la Tabla 1. Dado que el fin último es estimar trayectorias en escalas de tiempo inferiores al segundo, se ha elegido un tiempo de cálculo de la trayectoria de 0.2 s. Dada la frecuencia de muestreo  $f_s$ , se deduce el valor del horizonte de predicción  $N_p = 0,2f_s$ . Dado que el voltaje en  $dq$  tiene dos componentes, en este caso la  $\hat{y}^+ \subset \mathbb{R}^{2N_p}$ , siendo en este caso  $2N_p = 400$ , valores que deben ser predichos.

Durante el experimento de identificación de la red, consideramos condiciones de red constantes y estacionarias. Para excitar la red eléctrica bajo prueba y poder generar  $\mathcal{D}$ , el sistema se excitó en corriente en el marco  $dq$  mediante una combinación de señales chirp (oscilaciones con frecuencia y amplitud variable a lo largo del tiempo) y PRBS (Nelles, 2020) no correlacionadas, cada una con amplitud de 0,1 p.u. como máximo. El rango de frecuencias de excitación para la señal chirp y PRBS se ha escogido de modo que queden por debajo de la frecuencia de muestreo, y así, el pico de resonancia del sistema estuviera incluido; en este caso, el rango fue de  $[\frac{f_s}{2,5}, \frac{f_s}{1,8}]$  rad/s. De este experimento se obtuvieron 9201 pares de entrada-salida. De estos, se extrajo la base de datos de entrenamiento  $\mathcal{D}$ . Para validar el modelo, utilizamos un conjunto de datos de 4501 elementos,

obtenidos mediante señales chirp y PRBS del mismo modo.

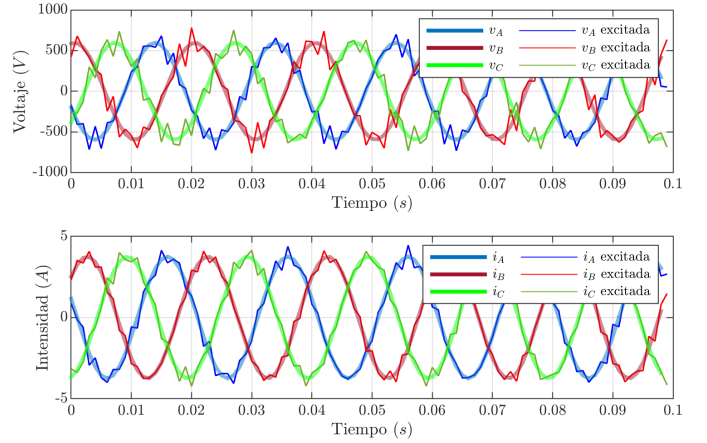


Figura 2: Voltajes y corrientes trifásicos perturbados para la obtención de la base de datos NARX utilizada en la regresión.

Las respuestas de corriente y voltaje resultantes en el PCC se muestran en la Figura 2. Se observa como el nivel de perturbación es bastante pequeño en comparación con el caso no excitado. Dado que la excitación también está limitada a unos pocos segundos, no deteriora la operación continua de la red. Analizamos este aspecto más en detalle, para asegurar que el funcionamiento normal de la red no se vea afectado y que se cumplen los límites establecidos por el estándar IEEE (IEEE Power and Energy Society, 2021) y el estándar Europeo (CENELEC (European Committee for Electrotechnical Standardisation), 2011). Éstos fijan un límite del 8 % de distorsión armónica total (THD) en voltaje, sobre 10 ciclos de la señal. Nuestro análisis sobre diversos segmentos de 10 ciclos ha verificado que el THD en la señal excitada no excede el 6,74 %, asegurando así el cumplimiento de las normativas vigentes y la integridad operativa de la red. Esta evaluación se ilustra gráficamente en la Figura 4, donde se presenta el espectro de frecuencias y el THD calculado utilizando el segmento de la señal de voltaje asociado con la máxima distorsión.

$\mathcal{D}$  está en la forma de (12), donde se ha impuesto que

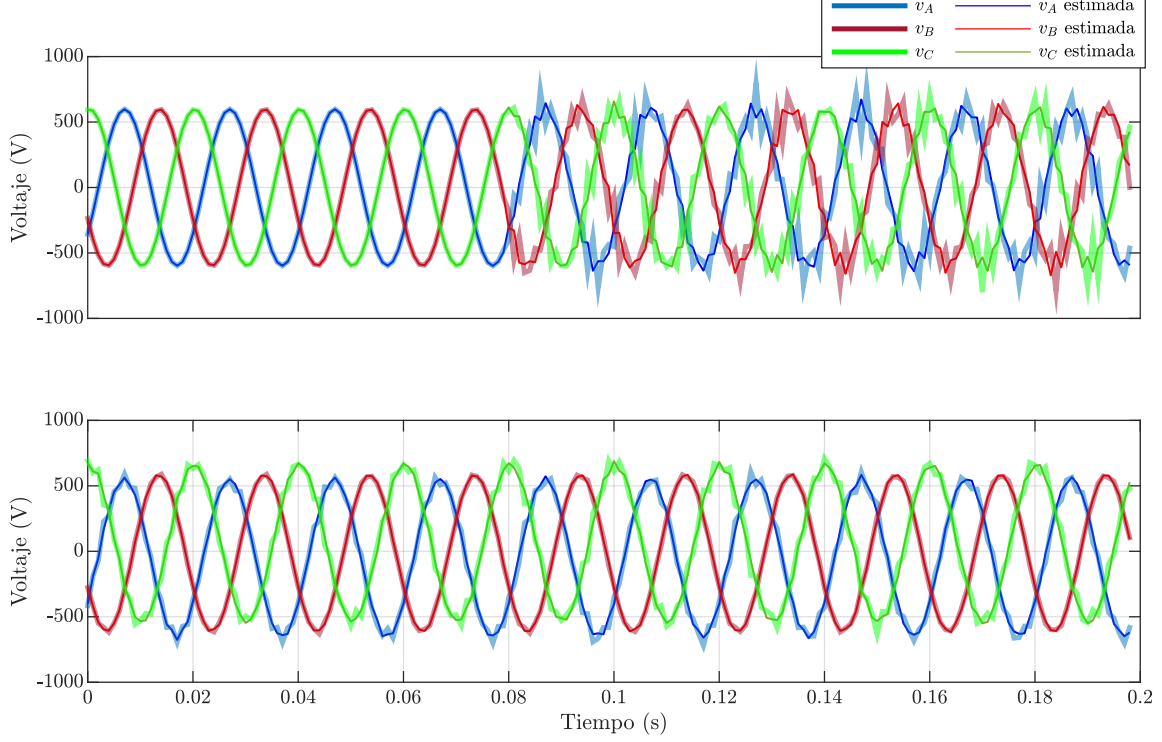


Figura 3: Comparación entre la trayectoria generada en simulación y la estimada por (1)-(2) haciendo uso del Algoritmo 1, K-ISTA, en dos escenarios diferentes utilizando la base de datos NARX con  $n_a = n_b = 4$ . La gráfica de arriba representa un escenario donde las tres componentes son excitadas al mismo tiempo, mientras que abajo únicamente las fases  $a$  y  $c$  son excitadas, dejando a  $b$  en todo momento sin excitar.

$n_a = n_b$  y se han tenido en cuenta diferentes valores, como puede verse en la Tabla 2. Cabe destacar que cada regresor  $\bar{z}$  tiene dimensión  $n = (n_a + 1)p + (n_b + 1)m$ .

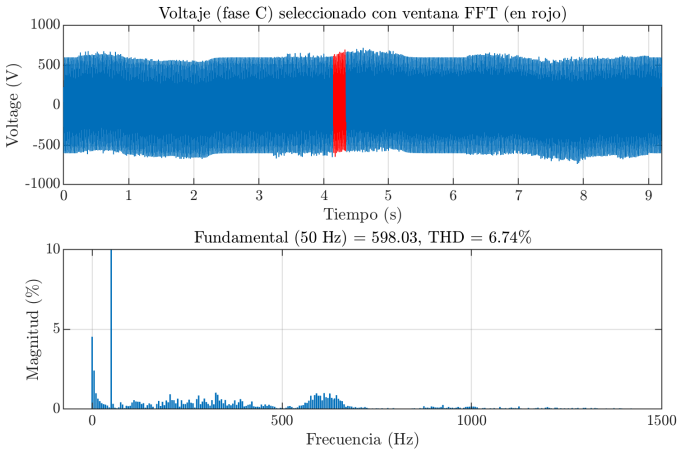


Figura 4: Análisis de la Distorsión Armónica Total (THD) en el voltaje de la fase  $c$  del sistema trifásico  $abc$ . En la gráfica superior se muestra la forma de onda temporal del voltaje, donde se marcan en rojo los 10 ciclos seleccionados para obtener el espectro de frecuencia con el cual se calcula el THD representado en la gráfica inferior, correspondiente al peor de los casos. Este análisis asegura que la señal cumple con el límite máximo permitido según los datos experimentales utilizados.

Para mejorar el tiempo de cálculo y la calidad de la predicción, se calcula una partición de  $\mathcal{D}$  llamada  $\mathcal{D}^\circ$ , la cual será la base de datos que se utilizará para realizar la predicción. Esta partición se conforma exclusivamente para cada punto de consulta  $\bar{z}$ , tomando los  $N^\circ = 1000$  puntos de la base de datos original  $\mathcal{D}$  que están más cercanos a  $\bar{z}$ ; la cardinalidad  $N^\circ$  se eligió basado en distintas pruebas de simulación, donde se evaluó la precisión de la predicción. El peso  $w_i$  asociado a cada uno de los puntos en  $\mathcal{D}^\circ$  se ha ajustado para asignar mayor importancia a los puntos más cercanos a  $\bar{z}$  (es decir, penalizando menos los coeficientes  $\lambda_i$  resultantes para estos puntos).

En este caso, se ha optado por definir  $w_i := \exp(\alpha \|\bar{z} - \bar{z}_i\|^s)$ , donde  $\alpha > 0$  y  $s \geq 1$  son parámetros ajustables. La modificación de estos parámetros permite controlar el peso de los puntos utilizados en la regresión, garantizando que el problema (1) esté siempre bien condicionado numéricamente. Los datos experimentales se han obtenido con  $s = 1$  y haciendo uso de normas euclídeas, tomando  $\alpha = \log(1,05)/r^\circ$ , donde  $r^\circ$  se define como la distancia máxima entre  $\bar{z}$  y cualquier punto en  $\mathcal{D}^\circ$ , es decir,  $r^\circ = \max_{i \in \{1, \dots, N^\circ\}} \|\bar{z}_i - \bar{z}\|_2$ .

Para validar el predictor, calculamos el error relativo normalizado como  $\zeta = \|\bar{y}_v - \hat{y}^+\|_\infty / \|\bar{y}_v\|_\infty$ , donde  $\bar{y}_v$  es una muestra de salida del conjunto de datos de validación, y  $\hat{y}^+$  la salida estimada por Kriging. En la Tabla 2 es posible ver una comparación de los tiempos de cálculo de la estimación con horizonte  $N_p = 200$ , teniendo en cuenta que se ha utilizado un ordenador equipado con un procesador Intel Core i7 de 8ª generación y 16 GB de RAM. La fila que se destaca en dicha tabla hace referencia a la

mejor elección de parámetros  $n_a = n_b = 4$ , escogida por producir el menor error medio de predicción obtenido con los datos de validación. Destacar que los tiempos mostrados en la Tabla 2 se han obtenido tanto con el algoritmo K-ISTA presentado en la sección 3.2, como con una versión acelerada de dicho algoritmo basada en el esquema FISTA con *restart* (Alamo et al., 2019), imponiendo una tolerancia  $\epsilon = 10^{-3}$  para establecer el criterio de parada.

Adicionalmente, en Figura 3 se muestra una comparativa de la calidad de la predicción realizada tras resolver (1)-(2) haciendo uso del Algoritmo 1, K-ISTA, en dos escenarios diferentes.

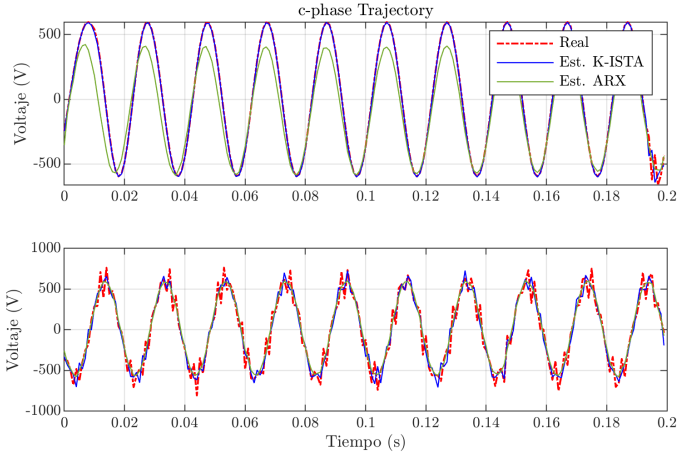


Figura 5: Comparación de dos trayectorias predichas para la fase  $c$ : modelo ARX (paramétrico) frente a predicción basada en Kriging (no paramétrico). En la gráfica superior, se evidencia que el modelo ARX, entrenado con la totalidad de la base de datos, no estima adecuadamente la señal no excitada. Por otro lado, en la gráfica inferior, el modelo basado en Kriging demuestra una mayor precisión en la captura de las excitaciones de la señal, superando al modelo ARX en la detección de dichas perturbaciones.

Como se discutió en la sección 2, esto se realiza en dos pasos: primero, se resuelve (1) para obtener el conjunto de coeficientes  $\{\lambda_1^*, \dots, \lambda_{N_o}^*\}$  que satisfacen la relación  $\bar{z} = \sum_{i=1}^{N_o} \lambda_i^* z_i$ ; luego, la trayectoria de voltaje  $\hat{y}^+$  se estima a través de (2), como  $\hat{y}^+ = \sum_{i=1}^{N_o} \lambda_i^* \tilde{y}_i$ . Del mismo modo, se predice la trayectoria de la frecuencia de la red  $\hat{\omega}_g^+$  usando los mismos coeficientes  $\lambda_i$  para poder pasar la predicción de voltaje del marco  $dq$  al marco  $abc$ . De esta forma,  $\hat{\omega}_g^+ = \sum_{i=1}^{N_o} \lambda_i^* \tilde{\omega}_{g,i}$ .

Asimismo, para evaluar la efectividad del predictor, se ha comparado con un modelo ARX ajustado sobre  $\mathcal{D}$  por mínimos cuadrados ordinarios, utilizando como métrica de evaluación la media de  $\zeta$  sobre todo el conjunto de validación. Con  $n_a, n_b = 4$ , el predictor basado en K-ISTA alcanza una precisión del  $\zeta = 3,78\%$ , comparado con el  $\zeta = 5,04\%$  del modelo ARX, pudiendo estimar de forma más exacta las excitaciones a las que está siendo sometido el modelo simulado. Esto puede verse en la Figura 5, donde se ha comparado la predicción de trayectorias para la fase  $c$ , mostrando que el modelo basado en Kriging tiene una mayor precisión en la captura de las excitaciones de la señal en comparación con el modelo ARX.

Tabla 2: Tiempo de cálculo (CT) y error relativo ( $\zeta$ ) para valores distintos de  $n_a$  y  $n_b$ , tomando  $\epsilon = 10^{-3}$ .

$n_a, n_b$	CT K-ISTA [ms]			CT K-ISTA acelerado [ms]			$\zeta$ [%]
	Max	Median	Min	Max	Median	Min	Mean
2	1135	25.80	0.47	28.40	5.97	0.26	3.92
4	715	18.80	0.58	41.24	5.04	0.28	3.78
8	919	19.39	0.74	64.41	5.52	0.50	3.95
16	3372	14.17	0.71	49.88	5.08	0.40	5.61
25	3686	12.96	0.35	74.79	4.81	0.37	11.3
50	4896	19.80	1.39	91.26	7.89	0.70	6.29

## 6. Conclusiones

En este trabajo se ha propuesto un método de predicción no paramétrico basado en Kriging para las variables de interés en la red de distribución. Los resultados demuestran que la precisión de la predicción puede llegar a ser inferior al 5%, con un tiempo de cálculo de menos de 200 ms en el peor de los casos, para grandes horizontes de predicción. Esto destaca su aplicabilidad en aplicaciones de la red que requieran maniobras en escalas de tiempo inferiores al segundo.

## Agradecimientos

Expresar nuestro agradecimiento a María Camila Merchán Riveros por su valiosa ayuda en el modelado de redes eléctricas para nuestro artículo.

Este trabajo ha sido realizado en el marco del Proyecto de investigación PID2022-142946NA-I00 financiado por MICIU/AEI /10.13039/501100011033 y por FEDER, UE, y del Proyecto 2023/00000487 financiado por el VIIPP-2022 de la Universidad de Sevilla. F. Fele y C. Moreno Blázquez también agradecen el soporte de 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.
- Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences 2 (1), 183–202.
- Bertsekas, D. P., 2009. Convex optimization theory. Athena Scientific.
- Car, M., Lešić, V., Vašak, M., 2021. Cascaded control of back-to-back converter dc link voltage robust to grid parameters variation. IEEE Transactions on Industrial Electronics 68 (3), 1994–2004.
- Carnerero, A. D., Ramirez, D. R., Alamo, T., 2022. Probabilistic interval predictor based on dissimilarity functions. IEEE Transactions on Automatic Control 67 (12), 6842–6849.
- Carnerero, A. D., Ramirez, D. R., Limon, D., Alamo, T., 2023. Kernel-based state-space Kriging for predictive control. IEEE/CAA Journal of Automatica Sinica 10 (5), 1263–1275.
- CENELEC (European Committee for Electrotechnical Standardisation), 2011. Voltage characteristics of electricity supplied by public distribution systems. European Norm EN 50160.
- Céspedes, M., Sun, J., 2012. Online grid impedance identification for adaptive control of grid-connected inverters. In: 2012 IEEE Energy Conversion Congress and Exposition (ECCE). IEEE, pp. 914–921.
- Chen, S., Billings, S. A., 1989. Representations of non-linear systems: the NARMAX model. International journal of control 49 (3), 1013–1032.
- Cressie, N., 1990. The origins of Kriging. Mathematical geology 22, 239–252.

- Deakin, M., Greenwood, D. M., Taylor, P. C., Armstrong, P., Walker, S., 2021. Analysis of network impacts of frequency containment provided by domestic-scale devices using matrix factorization. *IEEE Transactions on Power Systems* 36 (6), 5697–5707.
- Dobbe, R., van Westering, W., Liu, S., Arnold, D., Callaway, D., Tomlin, C., 2020. Linear single- and three-phase voltage forecasting and Bayesian state estimation with limited sensing. *IEEE Transactions on Power Systems* 35 (3), 1674–1683.
- ENA Task Group on Statutory Voltage Limits, 2017. Statutory voltage limits at customers' terminals in the UK and options for future application of wider limits at low voltage. Tech. Rep. ETR 140, Energy Networks Association.
- Francis, G., Burgos, R., Boroyevich, D., Wang, F., Karimi, K., 2011. An algorithm and implementation system for measuring impedance in the dq domain. In: 2011 IEEE Energy Conversion Congress and Exposition. IEEE, pp. 3221–3228.
- Gupta, C. P., Milanovic, J. V., 2006. Probabilistic assessment of equipment trips due to voltage sags. *IEEE Transactions on power delivery* 21 (2), 711–718.
- Haberle, V., Huang, L., He, X., Prieto-Araujo, E., Smith, R. S., Dörfler, F., 2023. MIMO grid impedance identification of three-phase power systems: Parametric vs. nonparametric approaches. In: 2023 62nd IEEE Conference on Decision and Control (CDC). IEEE, pp. 542–548.
- Hemyari, P., Nofziger, D., 1987. Analytical solution for punctual Kriging in one dimension. *Soil Science Society of America journal* 51 (1), 268–269.
- Huang, J., Corzine, K. A., Belkhaty, M., 2009. Small-signal impedance measurement of power-electronics-based ac power systems using line-to-line current injection. *IEEE Transactions on Power Electronics* 24 (2), 445–455.
- IEEE Power and Energy Society, 2021. IEEE draft standard for harmonic control in electric power systems. IEEE P519/D5.1, January 2021, 1–30.
- Kaur, G., Vaziri, M., 2006. Effects of distributed generation (DG) interconnections on protection of distribution feeders. In: 2006 IEEE Power Engineering Society General Meeting.
- Krige, D. G., 1981. Lognormal-de Wijsian geostatistics for ore evaluation. South African Institute of mining and metallurgy Johannesburg.
- Leontaritis, I. J., Billings, S. A., 1985. Input-output parametric models for nonlinear systems part i: deterministic non-linear systems. *International journal of control* 41 (2), 303–328.
- Levin, A., Narendra, K., 1997. Identification of nonlinear dynamical systems using neural networks. In: *Neural Systems for Control*. Elsevier, pp. 129–160.
- Liu, Z., Liu, J., Liu, Z., 2020. Analysis, design, and implementation of impulse-injection-based online grid impedance identification with grid-tied converters. *IEEE Transactions on Power Electronics* 35 (12), 12959–12976.
- Markovsky, I., Huang, L., Dörfler, F., 2023. Data-driven control based on the behavioral approach: From theory to applications in power systems. *IEEE Control Systems Magazine* 43 (5), 28–68.
- Martin, D., Nam, I., Siegers, J., Santi, E., 2013. Wide bandwidth three-phase impedance identification using existing power electronics inverter. In: 2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC). IEEE, pp. 334–341.
- Matheron, G., 1967. Kriging or polynomial interpolation procedures. *CIMM Transactions* 70 (1), 240–244.
- Matsui, H., Yamakawa, Y., 2023. Sparse estimation in ordinary Kriging for functional data. arXiv preprint arXiv:2306.15537.
- Merchán-Riveros, M. C., Albea, C., Seuret, A., 2024. Data-driven control design for power converters approximated as switched affine systems and experimental validation. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 1–1.
- Merhy, D., Alamo, T., Stoica Maniu, C., Camacho, E. F., 2018. Zonotopic constrained kalman filter based on a dual formulation. In: 2018 IEEE Conference on Decision and Control (CDC). pp. 6396–6401.
- Milano, F., Dörfler, F., Hug, G., Hill, D. J., Verbič, G., 2018. Foundations and challenges of low-inertia systems (invited paper). In: 2018 Power Systems Computation Conference (PSCC). pp. 1–25.
- Nadales, J., Carnerero, A., Moreno-Blazquez, C., Haes-Ellis, R., Limon, D., 2023. Learning-based NMPC on SoC platforms for real-time applications using parallel lipschitz interpolation. *IFAC-PapersOnLine* 56 (2), 6298–6303, 22nd IFAC World Congress.
- Nelles, O., 2020. *Nonlinear dynamic system identification*. Springer.
- Ordóñez, J. G., Nadales, J. M., Limon, D., Gordillo, F., 2021. Data-driven multirate predictive control of power inverters based on kinky inference. In: 2021 60th IEEE Conference on Decision and Control (CDC). pp. 4358–4363.
- Roinila, T., Messo, T., Santi, E., 2017. MIMO-identification techniques for rapid impedance-based stability assessment of three-phase systems in dq domain. *IEEE Transactions on Power Electronics* 33 (5), 4015–4022.
- Roll, J., Nazin, A., Ljung, L., 2005. Nonlinear system identification via direct weight optimization. *Automatica* 41 (3), 475–490.
- Saccomando, G., Svensson, J., 2001. Transient operation of grid-connected voltage source converter under unbalanced voltage conditions. In: *Conference Record of the 2001 IEEE Industry Applications Conference. 36th IAS Annual Meeting (Cat. No. 01CH37248)*. Vol. 4. IEEE, pp. 2419–2424.
- Shen, Z., Jaksic, M., Mattavelli, P., Boroyevich, D., Verhulst, J., Belkhaty, M., 2013. Three-phase ac system impedance measurement unit (IMU) using chirp signal injection. In: 2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC). IEEE, pp. 2666–2673.
- Willems, J. C., 1986. From time series to linear system—part i. finite dimensional linear time invariant systems. *Automatica* 22 (5), 561–580.
- Xiao, D., Hu, H., Chen, S., Song, Y., Pan, P., Molinas, M., 2022. Rapid dq-frame impedance measurement for three-phase grid based on interphase current injection and fictitious disturbance excitation. *IEEE Transactions on Instrumentation and Measurement* 71, 1–13.
- Xiao, P., Venayagamoorthy, G., Corzine, K., 2007. A novel impedance measurement technique for power electronic systems. In: 2007 IEEE Power Electronics Specialists Conference. IEEE, pp. 955–960.
- Zufferey, T., Renggli, S., Hug, G., 2020. Probabilistic state forecasting and optimal voltage control in distribution grids under uncertainty. *Electric Power Systems Research* 188, 106562.