

# Jornadas de Automática

## Modelado de potencia eólica mediante modelos físicos y procesos gaussianos

Martínez-Gutiérrez, Samuel<sup>a</sup>, Gutiérrez, Carlos<sup>a</sup>, Merino, Alejandro<sup>a</sup>, Sarabia, Daniel<sup>a,\*</sup>

<sup>a</sup> Departamento de Digitalización, Área de Ingeniería de Sistemas y Automática, Universidad de Burgos, Escuela Politécnica Superior, Avda. Cantabria, s/n, 09006 (Burgos).

**To cite this article:** Martínez-Gutiérrez, S., Gutiérrez, C., Merino, A., Sarabia, D. 2025. Wind power modelling using physical models and gaussian processes. *Jornadas de Automática*, 46.  
<https://doi.org/10.17979/ja-cea.2025.46.12218>

### Resumen

El modelado preciso de la producción de potencia en sistemas eólicos es esencial para optimizar su operación en tiempo real y cumplir objetivos técnicos o económicos. Este trabajo compara dos enfoques: uno paramétrico, basado en expresiones analíticas del coeficiente de potencia,  $C_P(\lambda, \beta)$ , y otro no paramétrico, que utiliza procesos gaussianos (GP) para modelar probabilísticamente la relación entre variables operativas y la potencia generada. Los modelos paramétricos son eficientes, interpretables y útiles cuando se conoce la dinámica del sistema, funcionando bien en distintos escenarios. En contraste, los modelos GP ofrecen mayor flexibilidad, integran la incertidumbre y se adaptan a patrones complejos en los datos. Se analiza el ajuste de ambos enfoques con datos reales de una turbina eólica y se evalúa su rendimiento mediante métricas como el Root Mean Square Error (RMSE) y el coste computacional. Los resultados muestran que ambos métodos son complementarios para modelar el comportamiento de turbinas eólicas en contextos diversos.

**Palabras clave:** Métodos no paramétricos, Optimización paramétrica, Modelado, Sistemas de potencia, Modelado y simulación de sistemas de potencia.

### Wind power modelling using physical models and gaussian processes

#### Abstract

Accurate modeling of power production in wind energy systems is essential for optimizing real-time operation and meeting technical or economic objectives. This work compares two approaches: a parametric one, based on analytical expressions of the power coefficient  $C_P(\lambda, \beta)$ , and a non-parametric one, which uses Gaussian Processes (GP) to probabilistically model the relationship between operating variables and generated power. Parametric models are efficient, interpretable and useful when the system dynamics are known, performing well under different scenarios. In contrast, GP models offer greater flexibility, integrate uncertainty and adapt to complex patterns in the data. We analyse the fit of both approaches with real wind turbine data and evaluate their performance using metrics such as Root Mean Square Error (RMSE) and computational cost. The results show that both methods are complementary for modelling the behaviour of wind turbines in different contexts.

**Keywords:** Nonparametric methods, Parametric optimization, Modelling, Power systems, Modelling and simulation of power systems

### 1. Introducción

Se conoce habitualmente como curva de potencia de un aerogenerador a la relación que existe entre la velocidad del viento que incide en el rotor y la potencia eléctrica producida por la turbina. Esta relación es fundamental para la evaluación del recurso eólico, el diseño y dimensionado de sistemas basados en energía eólica, así como para la

monitorización del rendimiento operativo de las turbinas. Habitualmente, los fabricantes proporcionan esta curva en forma tabular o gráfica, representando únicamente puntos discretos bajo condiciones estandarizadas y controladas. Sin embargo, la potencia producida en realidad por un aerogenerador va a diferir en muchas ocasiones de la proporcionada por el fabricante, a causa de las condiciones reales de operación. Por ello, ha surgido un campo de

investigación dedicado al modelado avanzado de estas curvas, con el objetivo de representar con mayor precisión el comportamiento de los aerogeneradores y sus variaciones bajo condiciones reales de operación.

En (Sohoni et al., 2016) se presenta una revisión crítica de las diferentes técnicas de modelado de curvas de potencia y sus aplicaciones en sistemas energéticos basados en viento. Los enfoques se agrupan en categorías tales como, modelos discretos, deterministas o probabilísticos, paramétricos o no paramétricos, modelos estocásticos y aquellos basados en datos especificados por el fabricante o datos de operación reales. A su vez, estas clasificaciones se pueden combinar, surgiendo así multitud de variedades y tipos de modelos. En particular, en los modelos paramétricos la relación entre las entradas y las salidas es fijada mediante un conjunto de ecuaciones matemáticas que contienen un número finito de parámetros, mientras que, en los modelos no paramétricos, no se hace ninguna asunción acerca de la forma funcional del fenómeno bajo observación.

Entre los enfoques paramétricos, existen multitud de modelos como los logísticos (Villanueva & Feijóo, 2018) o modelos polinomiales (Deshmukh & Deshmukh, 2008). Sin embargo, uno de los más completos es el basado en el modelo físico de conversión de la velocidad del viento en potencia, el cual incorpora en el coeficiente de potencia  $C_P$  (Castillo et al., 2023), (Reyes et al., 2015), una variable que representa la fracción de potencia del viento incidente que una turbina eólica es capaz de convertir en energía útil. Estos modelos asumen una expresión matemática predefinida del coeficiente de potencia y suelen ser eficientes cuando se dispone de poca información, como en etapas tempranas de diseño o para estimaciones generales.

Por otro lado, los modelos no paramétricos, como los basados en redes neuronales, máquinas de soporte vectorial (Veena et al., 2020) o procesos gaussianos (Zhou et al., 2014) y (Pandit et al., 2020) permiten capturar relaciones más complejas entre variables y se ajustan mejor a contextos operacionales reales con múltiples fuentes de incertidumbre y ruido. Sin embargo, a diferencia de los modelos con formulaciones predefinidas, requieren una cantidad considerable de datos para su entrenamiento.

El objetivo de este trabajo es comparar dos enfoques representativos para la estimación de la curva de potencia de un aerogenerador: un modelo paramétrico basado en el modelo físico de conversión eólica y un modelo no paramétrico de tipo gaussiano en el que no se hace ninguna suposición física ni de estructura matemática. El ajuste de ambos modelos se llevará a cabo con el mismo conjunto de datos reales que provienen de un sistema SCADA de un parque eólico real. La comparación se centrará no solo en la precisión de los modelos, sino también en su aplicabilidad práctica dependiendo del propósito final. Con este análisis se busca proporcionar una visión crítica de las ventajas y limitaciones de cada enfoque.

El artículo se organiza de la siguiente manera: en el apartado 2 se presenta la formulación de ambos modelos, en el apartado 3 se explica el procedimiento para ajustar cada uno de los modelos a los datos medidos y en el apartado 4 se proporciona información sobre los datos utilizados. En el apartado 5 y 6 se comparan y discuten los resultados

obtenidos, y en el apartado 7 se muestran unas breves conclusiones, así como futuras líneas de investigación.

## 2. Modelos

A continuación, se presenta la formulación de los dos modelos propuestos para estimar la potencia producida por un aerogenerador. En el subapartado 2.1 se expone el modelo paramétrico teórico de la curva de potencia de un aerogenerador y en el subapartado 2.2 se expone el uso procesos gaussianos para modelar la misma curva de potencia.

### 2.1. Modelo paramétrico basado en principios físicos

La potencia generada  $P$  por una turbina puede calcularse mediante (1).

$$P = \frac{1}{2} \rho \pi R^2 v^3 C_P \quad (1)$$

Donde  $\rho$  es la densidad del aire ( $\text{kg/m}^3$ ),  $R$  es el radio del rotor (m),  $v$  es la velocidad del viento (m/s) y  $C_P$  es el coeficiente de potencia adimensional.

En principio, todos los datos excepto el coeficiente de potencia  $C_P$  son conocidos: la potencia y la velocidad del viento las podemos extraer de un histórico de datos, el radio del rotor se puede obtener de la ficha técnica de la turbina y la densidad del aire en este caso la supondremos constante en  $1.225 \text{ kg/m}^3$  (densidad estándar a nivel del mar y a  $15^\circ\text{C}$ ) para no introducir una variable extra. Para turbinas con control de pitch como la que se utiliza para este estudio, el coeficiente de potencia depende fundamentalmente de dos variables: el ángulo de pitch  $\beta$  y el coeficiente de velocidad específica  $\lambda$ .

El ángulo de pitch  $\beta$  es el que forma la cuerda del perfil aerodinámico de la pala con el plano de rotación del rotor. Se utiliza para modificar las fuerzas aerodinámicas de sustentación y arrastre generadas por las palas, y con ello la potencia producida. Al aumentar el ángulo de pitch se reduce la potencia generada y viceversa.

El coeficiente de velocidad específica (2) es un número adimensional obtenido mediante el cociente entre la velocidad lineal en la punta de las palas  $\omega R$ , que depende a su vez de la velocidad de giro del rotor de la turbina  $\omega$  (rad/s) y la velocidad del viento  $v$  que incide en el rotor de la turbina.

$$\lambda = \omega R / v \quad (2)$$

La relación entre el coeficiente de potencia y las variables  $\beta$  y  $\lambda$  ha sido objeto de estudio en numerosos trabajos, siendo lo más habitual utilizar funciones polinómicas, sinusoidales o exponenciales. En nuestro caso se utilizará una función exponencial (3) dependiente de 9 parámetros  $C_k$ , ( $k = 1, \dots, 9$ ), ya que ha demostrado ser útil en el modelado de turbinas de velocidad variable con control de pitch como es nuestro caso (Manyonge et al., 2012).

$$C_P(\lambda, \beta) = C_1 \left( \frac{C_2}{\alpha} - C_3 \beta - C_4 \beta^{C_5} - C_6 \right) e^{\frac{-C_7}{\alpha}} \quad (3)$$

$$\frac{1}{\alpha} = \left( \frac{1}{\lambda + C_8 \beta} \right) - \left( \frac{C_9}{\beta^3 + 1} \right)$$

Por tanto, la estimación de la potencia con el modelo físico dependerá de las variables velocidad de viento ( $v$ ), ángulo de

pitch ( $\beta$ ) y velocidad específica ( $\lambda$ ).

De hecho, son los sistemas de control de la turbina quienes actuando sobre el pitch y decidiendo la velocidad de giro del rotor consiguen extraer la potencia deseada para cada velocidad de viento, y que los fabricantes muestran en la llamada *curva de potencia*.

## 2.2. Modelo no paramétrico basado en Proceso Gaussiano

Los procesos gaussianos (GP) son una técnica bayesiana no paramétrica ampliamente utilizada en problemas de aprendizaje supervisado. Sus principales ventajas son su capacidad para modelar relaciones complejas entre variables sin requerir una forma funcional específica además de proporcionar una estimación de la incertidumbre. Esta flexibilidad permite que los GP se adapten con facilidad a distintos tipos de datos, lo que los convierte en herramientas versátiles para una amplia variedad de aplicaciones.

Desde un enfoque más formal, un proceso gaussiano se define como una colección de variables aleatorias, tal que cualquier conjunto finito de ellas sigue una distribución conjunta gaussiana. Una de las formas más intuitivas de entender los GP es a través de la perspectiva del espacio de funciones (Function Space View), donde un GP representa una distribución de probabilidad sobre funciones continuas, en lugar de sobre los parámetros de un modelo determinado (Rasmussen & Williams, 2005). La formulación de un GP se expresa comúnmente con la ecuación (4).

$$f(x) \sim GP(m(x), k(x, x')) \quad (4)$$

Donde  $x$  representa un conjunto de variables de entrada,  $f(x)$  es la respuesta del proceso real que se quiere modelar,  $m(x)$ , es la función media y  $k(x, x')$  es la función de covarianza o *kernel*. La elección de estas dos funciones tiene un impacto significativo en el desempeño del proceso gaussiano. Por ello, en este estudio se ha decidido utilizar la configuración por defecto recomendada en la literatura como punto de partida. En concreto, se fija  $m(x) = 0$ , por ser una aproximación neutra y como kernel se emplea el Radial Basis Function (RBF), también conocido como Squared Exponential (SE), debido a su simplicidad y a su capacidad para modelar funciones suaves. Su expresión matemática se muestra en la ecuación (5), donde  $\sigma_f^2$ , es la varianza de proceso y  $\ell$  es el parámetro de longitud de escala.

$$k(x, x') = \sigma_f^2 \cdot \exp\left(-\|x - x'\|^2 / (2\ell^2)\right) \quad (5)$$

En la práctica, las observaciones del proceso no se corresponden directamente con  $f(x)$ , sino que están afectadas por ruido aditivo. Suponiendo que este ruido es gaussiano, independiente e idénticamente distribuido, el modelo observado se define según la ecuación (6).

$$y(x) = f(x) + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_n^2) \quad (6)$$

De este modo, el vector de observaciones  $y$  sigue siguiendo una distribución gaussiana multivariante (7) que depende de la matriz de covarianza total  $K_y$  (resultado de la suma de la matriz kernel  $K$ ) y obtenida a partir de  $k(x, x')$  y de la matriz  $\sigma_n^2 \cdot I$ , siendo  $I$  la matriz identidad.

$$y(x) \sim \mathcal{N}(0, K_y) \quad (7)$$

En el caso particular de este estudio, el proceso gaussiano busca predecir el comportamiento de la potencia generada por un aerogenerador. Para que sea comparable con el modelo físico descrito en el subapartado 2.1 se utilizarán sus mismas variables de entrada  $v, \beta$  y  $\lambda$ , creándose el vector  $x = [v, \beta, \lambda]^T$ , que junto a los hiperparámetros  $\theta = \{\ell, \sigma_f, \sigma_n\}$  definen el modelo final según la ecuación (8).

$$P \sim GP(0, k([v, \beta, \lambda]^T, [v', \beta', \lambda']^T)) + \varepsilon \quad (8)$$

Una buena práctica al trabajar con procesos gaussianos, especialmente cuando se emplean múltiples variables de entrada con diferentes unidades o escalas, es normalizar cada variable en un rango común, por ejemplo, entre 0 y 1. Esta normalización facilita el ajuste del modelo y garantiza que el proceso de inferencia no se vea sesgado por la magnitud relativa de las variables, preservando su influencia relativa en la función de covarianza. Posteriormente, cuando se desee conocer el valor real de la variable predicha, basta con aplicar el proceso inverso de desnormalización. En el presente trabajo, esta normalización se aplicará como parte del preprocesamiento de datos.

## 3. Metodología

En este apartado se formulan los problemas de optimización a resolver y el software utilizado para, en el primer enfoque, ajustar los parámetros del modelo y para, en el segundo enfoque, entrenar el proceso gaussiano.

### 3.1. Modelo físico. Mínimos cuadrados

Se plantea un problema de optimización con el objetivo de obtener el valor de los 9 parámetros  $C_k$  ( $k = 1, \dots, 9$ ) de la curva del  $C_P$ , tal que se minimice la función objetivo, que en este caso es el error cuadrático entre la potencia real medida  $P_i$  y la potencia estimada mediante el modelo  $\hat{P}_i$  en cada punto de operación  $i$  (9), donde  $N$  es el número total de muestras.

$$\min_{\{C_k\}} \sum_{i=1}^N (P_i - \hat{P}_i(\lambda_i, \beta_i))^2 \quad (9)$$

La potencia estimada  $\hat{P}_i$  se puede obtener sustituyendo  $\lambda$  por (2) en la función exponencial (3), dado que lo que se mide en realidad es la velocidad de giro de la turbina  $\omega$  y, a su vez, sustituyendo  $C_P$  por (3) en la expresión (1), obteniendo (10).

$$\hat{P}_i = \frac{1}{2} \rho \pi R^2 v_i^3 C_1 \left( \frac{C_2}{\alpha_i} - C_3 \beta_i - C_4 \beta_i^{C_5} - C_6 \right) e^{\frac{-C_7}{\alpha_i}} \quad (10)$$

$$\frac{1}{\alpha_i} = \left( \frac{1}{\omega_i R / v_i + C_8 \beta_i} \right) - \left( \frac{C_9}{\beta_i^3 + 1} \right)$$

Se ha planteado y resuelto este problema de optimización en Python 3.11.5 (Van Rossum & De Boer, 1991), utilizando la herramienta de modelado y optimización Pyomo 6.7.1, (Hart et al., 2011) y el solver IPOPT 3.11.1 (Wachter, 2002).

### 3.2. Proceso Gaussiano. Máxima verosimilitud

Para ajustar los hiperparámetros de un proceso gaussiano,  $\theta = \{\ell, \sigma_f, \sigma_n\}$ , se hace uso del principio de la máxima verosimilitud. Por las características del proceso gaussiano, es

posible inferir la distribución de densidad marginal de las observaciones, también conocida como verosimilitud marginal  $p(y|X, \theta)$ . Esta representa la probabilidad de observar los datos  $y$  dados los puntos de entrada  $X$  y los hiperparámetros  $\theta$ . Sin embargo, para estimar los hiperparámetros, es habitual maximizar la función log-verosimilitud (11), siendo  $N$  es el número total de muestras.

$$\log p(y|X, \theta) = \frac{y^T(K_y)^{-1}y}{2} - \frac{\log|K_y|}{2} - \frac{N}{2}\log(2\pi) \quad (11)$$

A diferencia del modelo físico, en el que fue necesario implementar manualmente en Python el problema de optimización para ajustar los parámetros, en el caso del proceso gaussiano se ha empleado la biblioteca GPflow 2.9.1 (Matthews et al., 2017). Esta herramienta facilita el ajuste automático de los hiperparámetros del modelo mediante funciones específicas que minimizan la log-verosimilitud marginal negativa, utilizando para ello el optimizador L-BFGS-B de la biblioteca SciPy (Virtanen et al., 2020).

#### 4. Datos experimentales

Los datos que se utilizarán en este estudio fueron obtenidos del repositorio de acceso abierto Zenodo (Duc & Simley, 2022) y pertenecen al proyecto francés SMARTEOLE, que estudia el parque Sole du Moulin Vieux (Simley et al., 2021). Los datos extraídos del SCADA del parque consisten en medidas realizadas entre el 17 de febrero y el 25 de mayo de 2020 de una serie de variables para la turbina 1 (modelo Senvion MM82 de 2050 kW de potencia nominal). El dataset consta de 134.661 puntos muestreados con una frecuencia de 1 minuto para medidas de la potencia eléctrica generada  $P$ , velocidad del viento  $v$ , ángulo de pitch  $\beta$  y velocidad de giro de la turbina  $\omega$ . Estos datos han sido filtrados de acuerdo a 4 criterios: i) El desalineamiento entre el rotor de la turbina y la dirección del viento debe ser menor que  $0.1^\circ$  para considerar únicamente situaciones en las que la turbina esté alineada completamente al viento. ii) La velocidad de giro del generador debe ser positiva. iii) La velocidad medida debe estar dentro del rango de operación de la turbina ( $v_{cut-in} < v < v_{cut-off}$ ). iv) La potencia generada debe ser siempre positiva. Además de este filtro, para evitar problemas con el cálculo del coeficiente de potencia, se ha reemplazado el valor del ángulo de pitch  $\beta$  por cero en todos los puntos en los que éste era negativo (valores ligeramente inferiores a cero debidos a errores de medida). Una vez filtrados, se disponen de 1131 puntos (Figura 1) ordenados por fechas, de los cuales se va a utilizar el 80% inicial para entrenar el modelo y el 20% restante para validarlo.

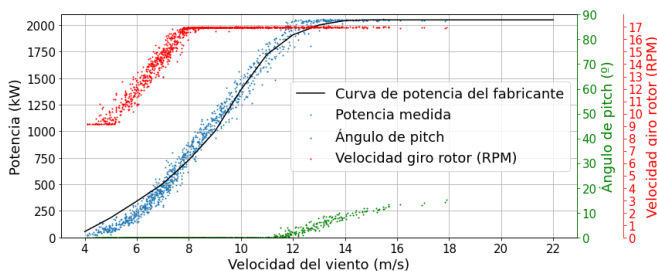


Figura 1. Valores medidos de potencia, ángulo de pitch y velocidad de giro del rotor. En negro, la curva de potencia suministrada por el fabricante.

## 5. Resultados

### 5.1. Modelo físico

Al resolver el problema de optimización se obtienen los parámetros  $C_k$  de la ecuación (3), recogidos en la Tabla 1.

Tabla 1. Parámetros obtenidos para la curva del coeficiente de potencia  $C_P$ .

Parámetro	Valor	Parámetro	Valor
$C_1$	0.000320415	$C_6$	18212.9
$C_2$	278954	$C_7$	19.6804
$C_3$	6.81025e-08	$C_8$	0
$C_4$	82.5864	$C_9$	2.35016e-05
$C_5$	1.72139		

En la Figura 2 se muestra la potencia estimada a partir de la ecuación (10) y los parámetros de la Tabla 1, respecto al viento incidente. En la misma gráfica se compara dicha potencia estimada con la potencia medida, que se ha clasificado en dos grupos: *Datos ajuste*, con los que se ha resuelto el problema de optimización y se han obtenido los parámetros de la curva de  $C_P$ ; y *Datos test*, los cuales no se han utilizado para la obtención del modelo y sirven para la validación del mismo.

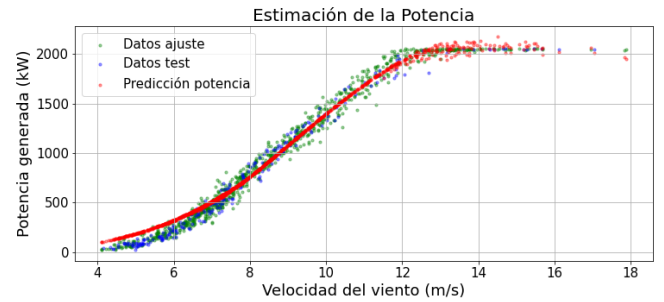


Figura 2. Predicción de la potencia producida a partir del modelo físico en función de la velocidad del viento.

En general, las estimaciones son adecuadas para la mayor parte de los puntos, si bien para el rango de velocidades de viento entre 4-7 m/s se aprecia mayor error. Esto puede deberse principalmente a dos motivos: por un lado, se aprecia una mayor concentración de datos test en ese rango de velocidades, por lo que es probable que nuestro modelo no tenga información suficiente para estimar la potencia en esas condiciones; por otro lado, el hecho de suponer la densidad del aire constante, puede generar errores en la estimación de la potencia, ya que la velocidad del viento y la densidad pueden tener cierto grado de correlación a través de factores ambientales como la temperatura o la presión atmosférica.

### 5.2. Proceso gaussiano

En la Tabla 2 se muestran los valores óptimos de los hiperparámetros obtenidos tras ajustar el proceso gaussiano con los 905 datos del conjunto de entrenamiento.

Tabla 2. Valores de los hiperparámetros del Proceso Gaussiano.

Hiperparámetro	$\sigma^2$	$\ell$	$\sigma_n^2$
Valor	0.593362	0.339067	0.000587091

En la Figura 3 se presentan las predicciones del valor medio del proceso gaussiano sobre el conjunto completo de datos. Las predicciones se visualizan en una gráfica potencia-velocidad, donde se forma una nube de puntos. Visualmente, se puede observar que, en términos generales, el proceso gaussiano predice mejor la dispersión en la potencia. No obstante, existe una zona comprendida aproximadamente entre los 8 y 12 m/s en la que el modelo no logra capturar toda la variabilidad presente. Esta limitación podría deberse a diversos factores, como, por ejemplo, la ausencia de variables relevantes, como la densidad del aire, que no se han considerado en el modelo. Asimismo, también podría estar relacionada con el tipo de *kernel* empleado.

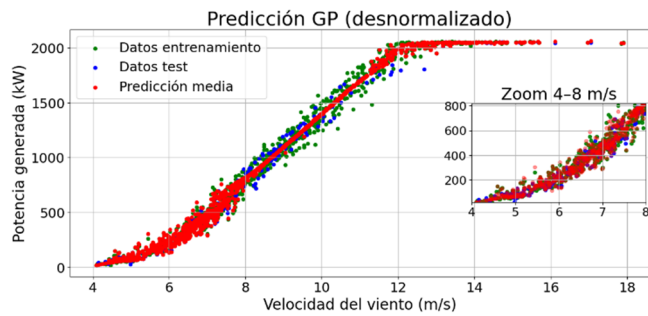


Figura 3. Predicción de la potencia del GP ajustado para los datos de entrenamiento y de test.

### 5.3. Comparación de modelos

En la Figura 4 se compara la predicción de la potencia de ambos modelos con los datos test. Se puede comprobar como de forma general, la predicción con el modelo gaussiano es mucho más precisa que con el modelo físico. Además, la predicción con el proceso gaussiano presenta la ventaja de proporcionar un intervalo de confianza de la predicción, que en el modelo físico es más complejo de obtener.

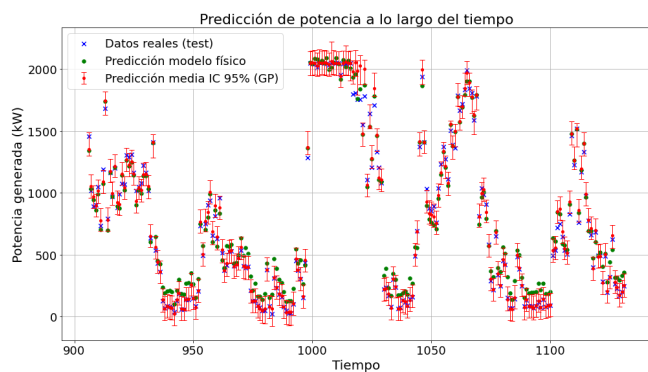


Figura 4. Comparación entre las predicciones a lo largo del tiempo, del modelo físico y del proceso gaussiano.

Como medida de comparación de la bondad de ajuste de ambos modelos, se ha utilizado la raíz del error cuadrático medio (RMSE), cuyo valor se ha calculado tanto para el conjunto de datos de entrenamiento como para el conjunto de test y se muestra en Tabla 3, junto con el tiempo de cálculo que se tarda en realizar ambos ajustes.

Las Figuras 5 y 6 muestran cómo varía la potencia generada en función de  $\lambda$  para distintas velocidades del viento, manteniendo constante el ángulo  $\beta$ , a modo de ejemplo, en cero. La Figura 5 representa el modelo físico y la

Figura 6 los resultados obtenidos con el proceso gaussiano. Este proceso reproduce bien la estructura del modelo físico en las zonas con datos de entrenamiento ( $\lambda$  entre 7.5 y 11), pero muestra diferencias notables en áreas sin datos medidos, como a altas velocidades del viento, donde también aumenta la incertidumbre. Esto destaca la principal limitación de los procesos gaussianos: cuando se extrapolan fuera del rango de entrenamiento, tienden al valor de la media y a aumentar la varianza predictiva.

Tabla 3. Comparación del RMSE y los tiempos de ajuste y entrenamiento.

Indicador	Modelo físico	Proceso gaussiano
RMS (Entrenamiento)	81.3522	48.7089
RMSE (Test)	86.6744	52.0309
Tiempo ajuste/entrenamiento	2.68 (s)	10.6 (s)

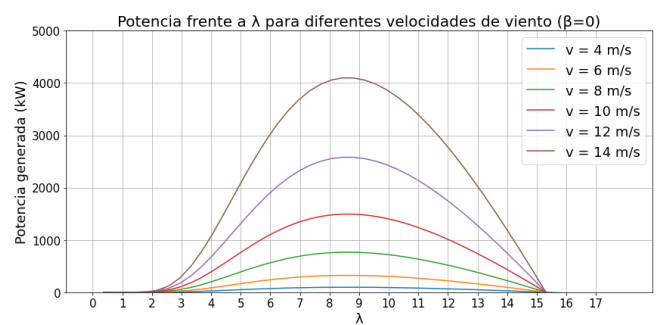


Figura 5. Representación de la potencia frente a  $\lambda$  para distintas velocidades de viento usando el modelo físico.

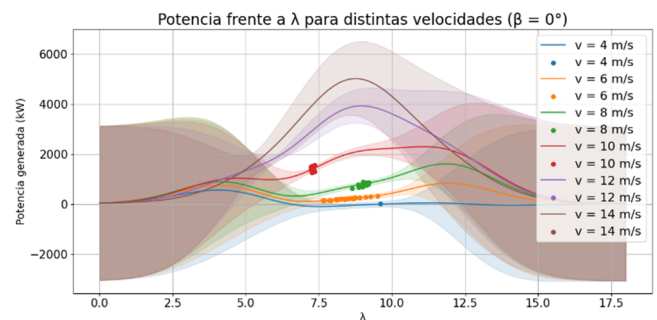


Figura 6. Representación de la potencia frente a  $\lambda$  para distintas velocidades de viento usando el proceso gaussiano.

## 6. Discusión

Se ha realizado el ajuste de parámetros de un modelo paramétrico físico y el ajuste de hiperparámetros de un proceso gaussiano a partir del mismo conjunto de datos. El modelo gaussiano ha resultado en un menor RMSE que el modelo paramétrico, lo que indica que sus predicciones se ajustan mejor a los datos reales bajo el criterio del error cuadrático medio, además de proporcionar información sobre el grado de incertidumbre de la predicción.

De manera general el proceso gaussiano no impone ninguna estructura física y predefinida al modelo, lo que le permite tener en cuenta una amplia variedad de efectos en la potencia y fenómenos entre variables, consiguiendo una mejor predicción de la potencia y su dispersión, frente al modelo físico que está condicionado a la estructura impuesta a priori. En particular, las suposiciones hechas en el modelo físico no

permiten capturar la complejidad de los fenómenos subyacentes, y se basan en: i) Considerar una densidad del aire constante en su valor estándar, está claro que la densidad cambia a lo largo del día y que está correlacionada de alguna manera con la velocidad del aire a través de la temperatura y la presión. ii) Suponer una eficiencia constante, ya que la potencia medida es la potencia eléctrica producida por el generador y el modelo físico representa la potencia mecánica. El efecto de la eficiencia se ha trasladado a la estimación del coeficiente de potencia, pero esta inclusión no permite tener en cuenta que la eficiencia pueda variar con la velocidad de giro de la turbina, que es un fenómeno habitual y que haría que se produzca una mayor o menor potencia respecto a la esperada para distintas velocidades. iii) Considerar un modelo estacionario de producción de potencia. Para una misma velocidad de viento medida, la potencia real generada puede variar si se está en mitad de un transitorio.

Ambas metodologías sufren de la problemática de la fiabilidad de las medidas experimentales, debido a posibles errores de calibración, *offset*, etc., que afectan al ajuste de los modelos. No está claro en qué medida estos problemas impactan a una u otra metodología.

En cuanto a los tiempos de cálculo, con la cantidad de datos usados en este estudio, el modelo físico se obtiene en una cuarta parte del tiempo necesario para entrenar el proceso gaussiano. Sin embargo, se sabe que este último se vuelve inviable con grandes volúmenes de datos por el aumento exponencial del tiempo de cálculo.

## 7. Conclusiones

A la vista de los resultados, si el objetivo es obtener un modelo dinámico para ensayar cómo se comporta el sistema ante variaciones en las diferentes entradas, que pueden ser distintas de las de la operación normal o, por ejemplo, ensayar distintas estrategias de control de potencia, creemos que el modelo físico es el más adecuado, ya que obtenemos con él la familia de curvas del coeficiente de potencia y conocemos exactamente cómo se relacionan las distintas variables entre sí en todo el rango. Por otro lado, el modelo gaussiano es muy bueno haciendo predicciones en situaciones similares a las de funcionamiento normal del proceso sobre las que se han realizado observaciones, pero no nos da información sobre la estructura física del sistema, por lo que sus aplicaciones serán otras: tareas de monitorización, diagnóstico de fallos o el apoyo a la operación en tiempo real. Como línea futura, se pretende combinar ambas metodologías para tratar de aprovechar los puntos fuertes de cada método.

## Agradecimientos

La publicación es parte del proyecto “Gestión Óptima en Tiempo Real del ciclo Power-to-H<sub>2</sub>-to-Power (OptiMaPH2P)”, TED2021-131220B-I00, financiado por MCIN/AEI y por la Unión Europea “NextGenerationEU” y del proyecto “Gestión óptima en tiempo real y bajo incertidumbre para gemelos digitales (OptiDit)”, PID2021-

1236540B-C33, financiado por MCIN y por la Unión Europea “FEDER”. Este trabajo también forma parte de la Tesis Doctoral de Samuel Martínez-Gutiérrez, financiada con un contrato predoctoral de Formación del Profesorado Universitario (FPU), convocatoria 2022, concedido por el MUNI de España.

## Referencias

- Castillo, O. C., Andrade, V. R., Rivas, J. J. R., & González, R. O. (2023). Comparison of Power Coefficients in Wind Turbines Considering the Tip Speed Ratio and Blade Pitch Angle. *Energies*, 16(6), 2774. <https://doi.org/10.3390/en16062774>
- Duc, T., & Simley, E. (2022, November 24). SMARTEOLE Wind Farm Control open dataset. Zenodo. <https://doi.org/10.5281/zenodo.7342466>
- Hart, W. E., Watson, J.-P., & Woodruff, D. L. (2011). Pyomo: modeling and solving mathematical programs in Python. *Mathematical Programming Computation*, 3(3), 219–260.
- Manyonge, A. W., Ochieng, R. M., Onyango, F. N., & Shichikha, J. M. (2012). Mathematical Modelling of Wind Turbine in a Wind Energy Conversion System: Power Coefficient Analysis. *Applied Mathematical Sciences*, 6(91), 4527–4536.
- Matthews, A. G. de G., van der Wilk, M., Nickson, T., Fujii, Keisuke., Boukouvalas, A., León-Villagrà, P., Ghahramani, Z., & Hensman, J. (2017). GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40), 1–6. <http://jmlr.org/papers/v18/16-537.html>
- Pandit, R. K., Infield, D., & Kolios, A. (2020). Gaussian process power curve models incorporating wind turbine operational variables. *Energy Reports*, 6, 1658–1669. <https://doi.org/10.1016/J.EGYR.2020.06.018>
- Rasmussen, C. E., & Williams, C. K. I. (2005). Gaussian Processes for Machine Learning. Gaussian Processes for Machine Learning. <https://doi.org/10.7551/MITPRESS/3206.001.0001>
- Reyes, V., Rodríguez, J. J., Carranza, O., & Ortega, R. (2015). Review of mathematical models of both the power coefficient and the torque coefficient in wind turbines. *IEEE International Symposium on Industrial Electronics*, 2015-September, 1458–1463. <https://doi.org/10.1109/ISIE.2015.7281688>
- Simley, E., Fleming, P., Girard, N., Alloin, L., Godefroy, E., & Duc, T. (2021). Results from a wake-steering experiment at a commercial wind plant: investigating the wind speed dependence of wake-steering performance. *Wind Energy Science*, 6(6), 1427–1453. <https://doi.org/10.5194/wes-6-1427-2021>
- Sohoni, V., Gupta, S. C., & Nema, R. K. (2016). A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems. *Journal of Energy*, 2016(1), 8519785. <https://doi.org/10.1155/2016/8519785>
- Van Rossum, G., & De Boer, J. (1991). Interactively testing remote servers using the Python programming language. *CWI Quarterly*, 4(4), 283–303.
- Veena, R., Mathew, S., & Petra, M. I. (2020). Artificially intelligent models for the site-specific performance of wind turbines. *International Journal of Energy and Environmental Engineering*, 11(3), 289–297. <https://doi.org/10.1007/S40095-020-00352-2/FIGURES/5>
- Villanueva, D., & Feijóo, A. (2018). Comparison of logistic functions for modeling wind turbine power curves. *Electric Power Systems Research*, 155, 281–288. <https://doi.org/10.1016/J.EPSR.2017.10.028>
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wachter, A. (2002). An interior point algorithm for large-scale nonlinear optimization with applications in process engineering. Carnegie Mellon University.
- Zhou, J., Guo, P., & Wang, X. R. (2014). Modeling of wind turbine power curve based on Gaussian process. *Proceedings - International Conference on Machine Learning and Cybernetics*, 1, 71–76. <https://doi.org/10.1109/ICMLC.2014.7009094>