

Desarrollo de Modelos de Pronóstico Tratamiento de Errores

Abril 2025



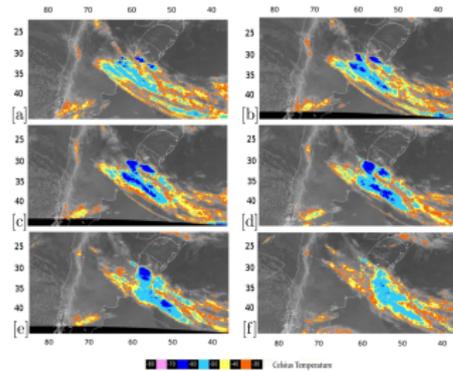
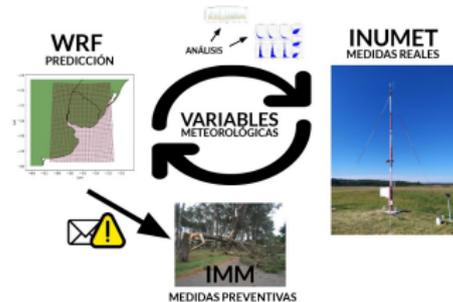
Contenido

- 1 Desarrollo de Modelos.
- 2 Modelos y Errores

Estrategia

Para el desarrollo de Modelos de pronóstico operativos.

- ¿ Que tipo de variable nos interesa pronosticar ?
- ¿ Cual es horizonte de tiempo en que necesitamos generar el pronóstico ?
- ¿ Que tipo de información histórica y operativa disponemos ?
- ¿ Que capacidad de cálculo operativa disponemos ?



¿ Que tipo de Variable ?

Que tipo de variable es la variable objetivo (target) a pronósticar.

Variables (\mathbb{R})



$$y \in \mathbb{R}$$

- **Ejemplo:** Pronóstico de la velocidad del viento (u):

$$u \in \mathbb{R}, \quad u \geq 0$$

Variables Discretas (\mathbb{Z})



$$y \in \mathbb{Z}$$

- **Ejemplo:** Pronóstico del número de aerogeneradores operativos (N):

$$N \in \mathbb{Z}, \quad 0 \leq N \leq N_{\max}$$

Donde N_{\max} es el número total de aerogeneradores.

Variables binarias toman solo dos valores, típicamente 0 o 1. Por ejemplo, si un aerogeneradore está operativo (1) o no (0).



$$y \in \{0, 1\}$$

- **Ejemplo:** Pronóstico del estado de una turbina (S):

$$S \in \{0, 1\}$$

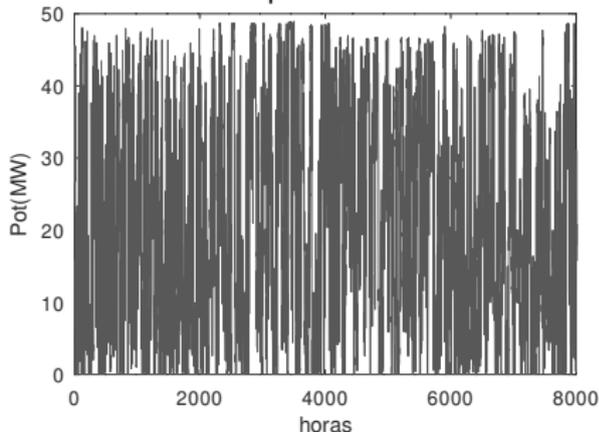
Donde:

- $S = 1$: Operativo.
- $S = 0$: No operativo.

Analizar la calidad de los datos

- Graficar datos y analizar valores.
- Graficar datos tanto de entrada al modelo como variable objetivo (target).
- Eliminar datos o periodos de tiempo.
- Correlacionar datos similares, ejemplos datos de velocidad de viento en torres cercanas.
- Por ejemplo analizar, valores máximos, valores mínimos, durante un periodo de tiempo.

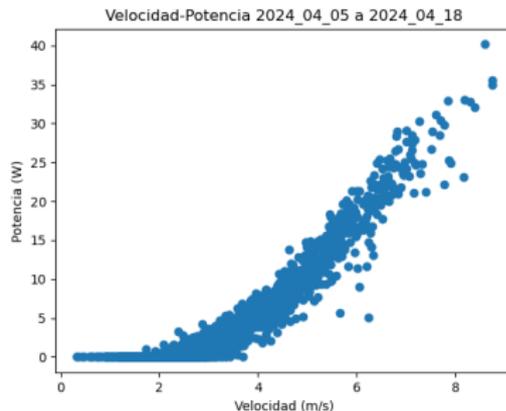
Ejemplo potencia generada por un parque eólico con 50 MW de potencia instalada.



Analizar la calidad de los datos

- Graficar scatter.
- Analizar puntos en el scatter.
- Analizar posible afectación física de las variables, por ejemplo afectación de la velocidad por obstáculos.

Potencia generada por un microarogenerador vs velocidad de viento en torre cercana.



Correlación

La correlación de Pearson, mide la relación lineal entre dos variables X e Y :

$$\rho = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \cdot \sum (Y_i - \bar{Y})^2}}$$

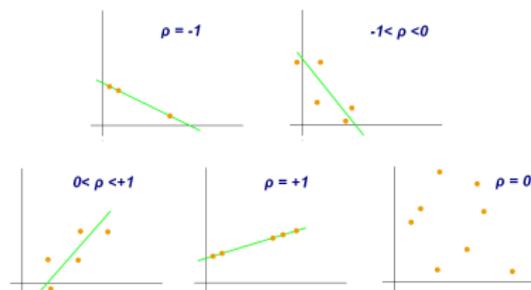
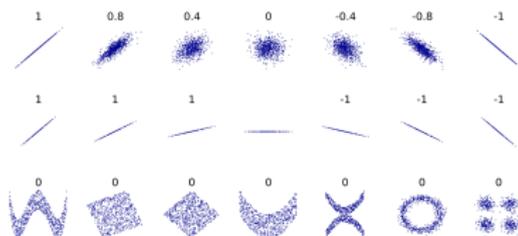
Donde:

- ρ es el coeficiente de correlación de Pearson.
- X_i e Y_i son los valores individuales de las variables X e Y .
- \bar{X} e \bar{Y} son los valores medios de X e Y , respectivamente.

Este coeficiente varía entre -1 y 1 :

- $r = 1$: Correlación positiva perfecta.
- $r = -1$: Correlación negativa perfecta.
- $r = 0$: No hay correlación lineal.

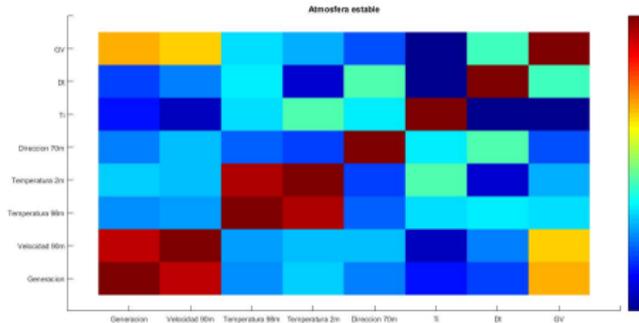
Correlación



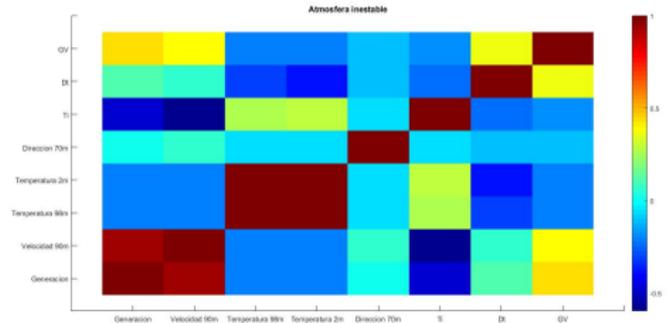
Analizar Correlación

Entre variables de entrada y variable objetivo (target)

Condiciones de estabilidad en CLA



Condiciones de inestabilidad en CLA



Análisis de errores.

Error en cada instante t , en función del horizonte de tiempo k .

Error, en función del horizonte de tiempo (k).

$$e(t+k|t) = P(t+k) - \hat{P}(t+k|t)$$

Error normalizado por la potencia nominal instalada, en función del horizonte de tiempo (k).

$$e(t+k|t)_n = \frac{1}{P_{inst}} (P(t+k) - \hat{P}(t+k|t))$$



Sesgo, BIAS.

Sesgo de error en función del horizonte de tiempo (k).

$$BIAS(k) = \frac{1}{N} \sum_{t=1}^{t=N} e(t + k|t)$$

Sesgo de error normalizado por la potencia nominal instalada en función del horizonte de tiempo (k).

$$BIAS_n(k) = \frac{1}{P_{inst}} \frac{1}{N} \sum_{t=1}^{t=N} e(t + k|t)$$



Error medio absoluto, MAE (Mean Absolute Error).

Error medio absoluto en función del horizonte de tiempo (k).

$$MAE(k) = \frac{1}{N} \sum_{t=1}^{t=N} |e(t + k|t)|$$

Error medio absoluto normalizado por la potencia nominal instalada en función del horizonte de tiempo (k).

$$MAE_n(k) = \frac{1}{P_{inst}} \frac{1}{N} \sum_{t=1}^{t=N} |e(t + k|t)|$$



Raíz del error Cuadrático Medio, RMSE (Root Mean Squared Error).

Raíz del error cuadrático medio en función del horizonte de tiempo (k).

$$RMSE(k) = \sqrt{\frac{1}{N} \sum_{t=1}^{t=N} (e(t+k|t))^2}$$

Raíz del error cuadrático medio normalizado por la potencia nominal instalada en función del horizonte de tiempo (k).

$$RMSE_n(k) = \frac{1}{P_{inst}} \sqrt{\frac{1}{N} \sum_{t=1}^{t=N} (e(t+k|t))^2}$$



Percentiles

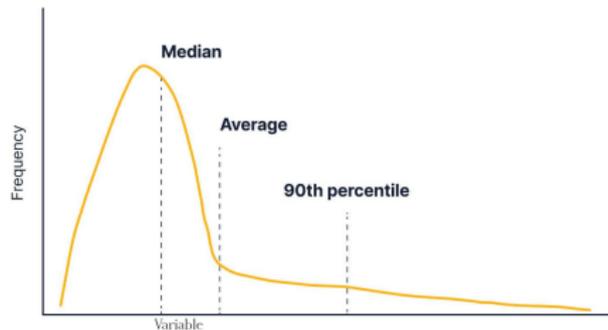
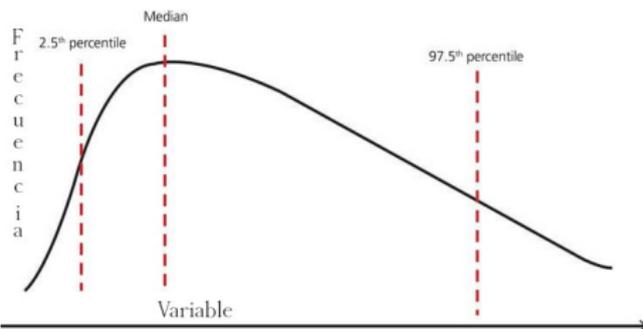
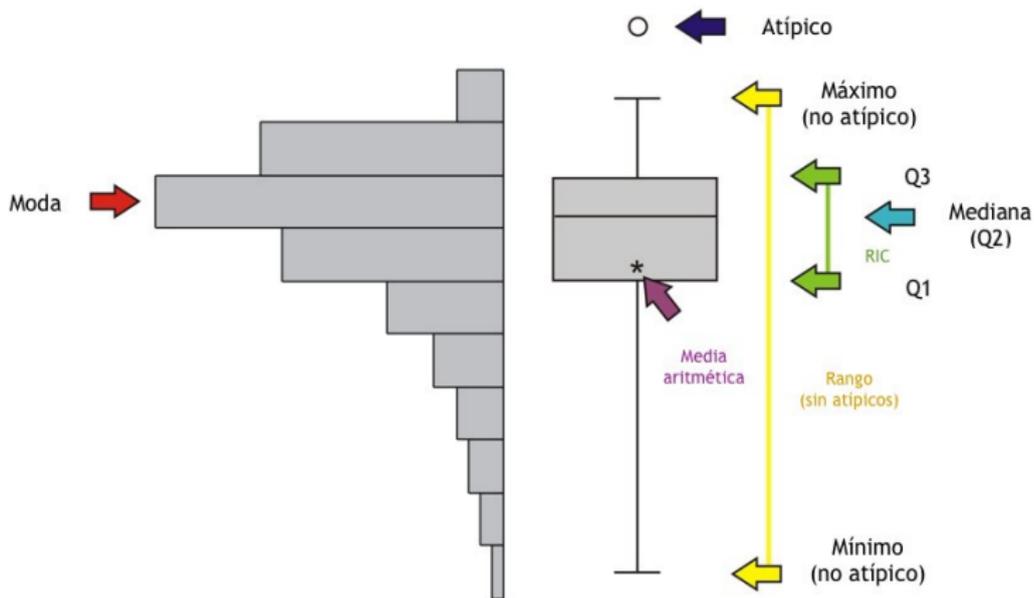
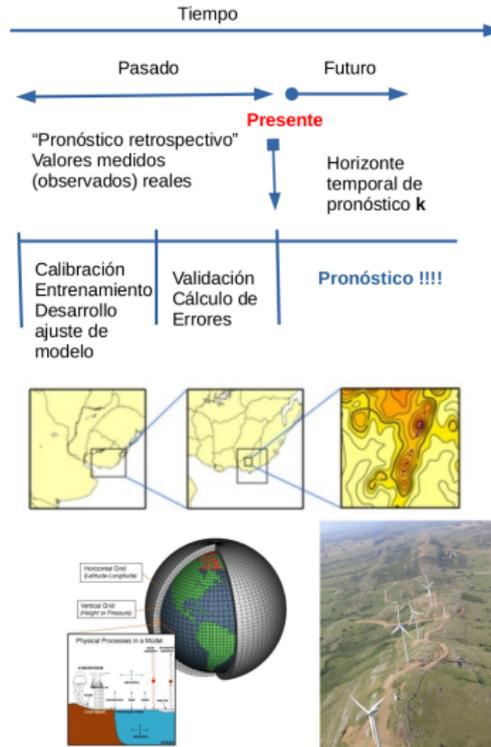


Diagrama de caja - Box plot

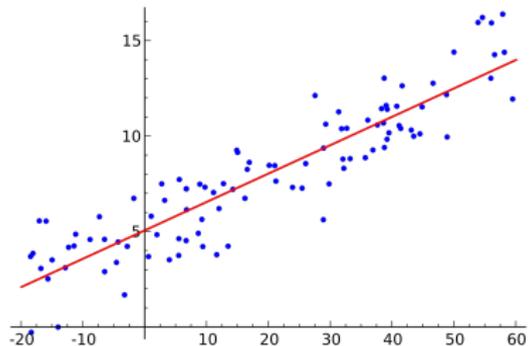


Desarrollo y Pronóstico



Regresión Lineal

$$X = a\hat{X} + b$$



Regresión Lineal Múltiple

$$X = a_1 \hat{X}_1 + a_2 \hat{X}_2 + \dots + a_n \hat{X}_n + b$$

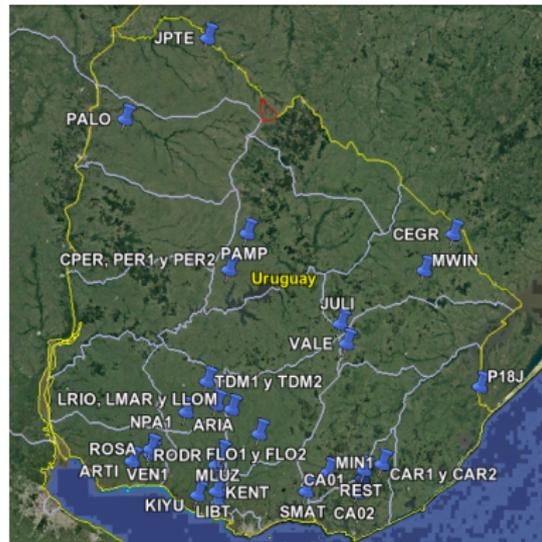
Análisis

- Condiciones de Borde: ECMWF (50 miembros, datos **cada 6hrs** grilla, 55 Km.
- Entrenamiento RNA: Con **ensembles**. Diferentes input y configuraciones de RNA.
- VARIABLES meteorológicas U, V, P, T100m, T2m.



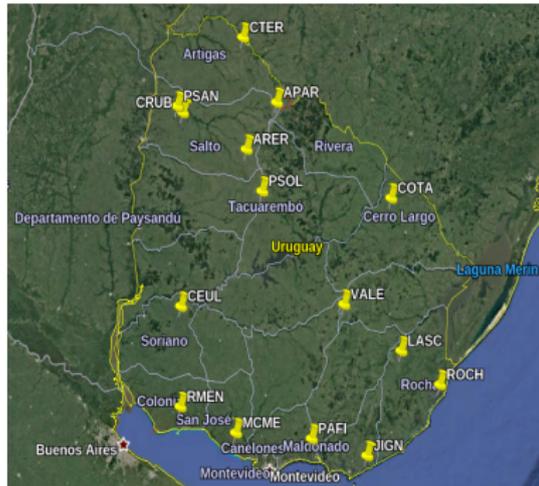
Corrección de variables meteorológicas con RNA

- Datos: ECMWF. Grilla 0.5(aprox. 55 km), paso temporal de 6 horas, inicialización 00 GMT, 50 miembros.
- Interpolación espacial: Se hace una interpolación en cada punto de interés (parques eólicos y torres anemométricas).
- Variabes: U, V, T2, T100m y P. En base a estas variables se calculan velocidad y dirección del viento, temperatura a 2 m y 100m, densidad del aire y presión atmosférica.

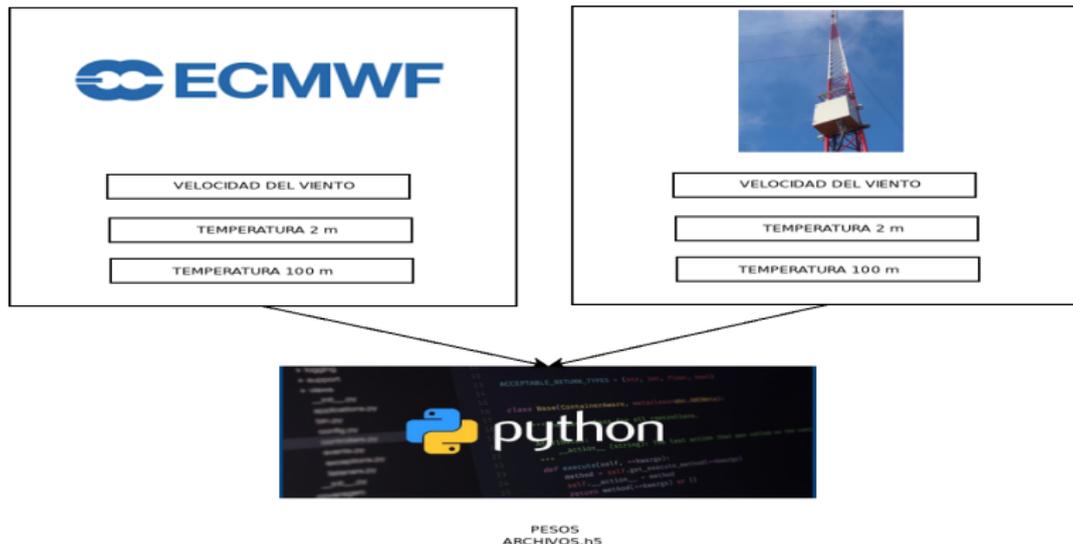


Corrección de variables meteorológicas con RNA

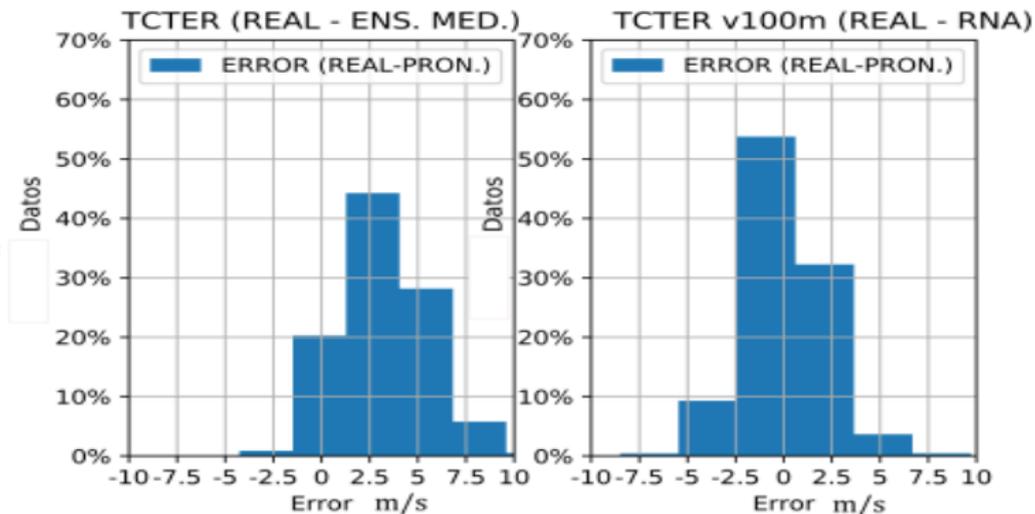
- Interpolación temporal: Se obtiene una serie horaria de los datos a partir de una interpolación lineal.
- Torres Meteorológicas: A partir de un conjunto de torres meteorológicas, se corrigen tres variables pronosticadas por el modelo: velocidad del viento, temperaturas a 2 m y 100 m.
- Aplicación de la Corrección: Dicha corrección se hace en un período de entrenamiento, donde al final del proceso obtenemos "pesos" para determinada variable.



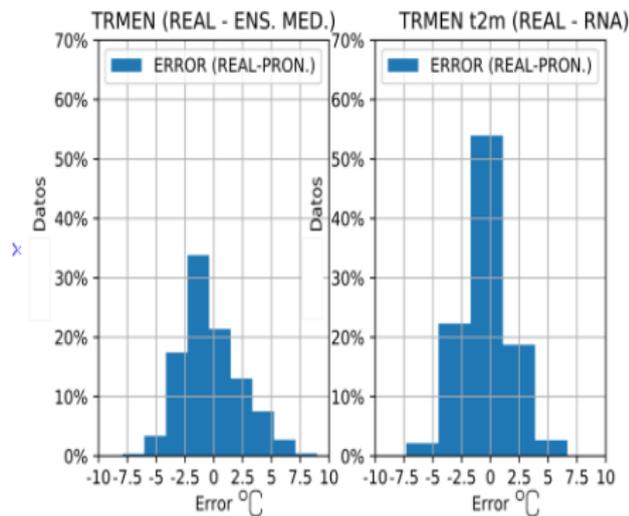
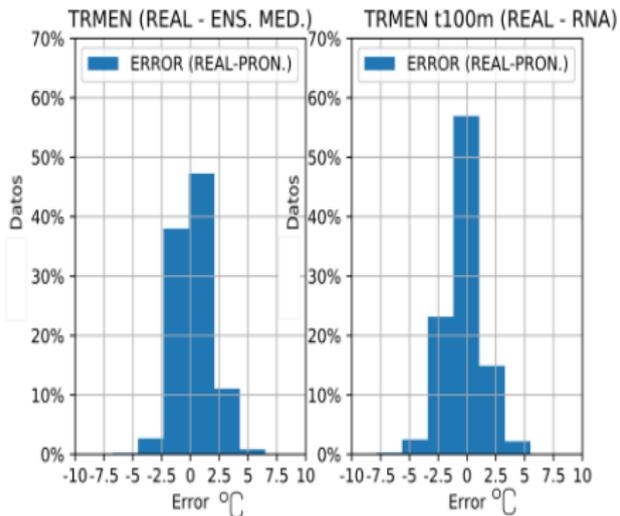
Proceso de Corrección de Variables Atmosféricas



Proceso de Corrección de Variables Atmosféricas

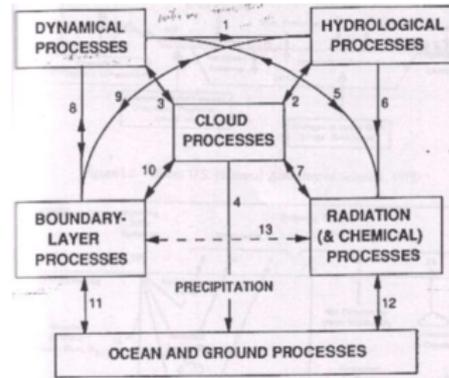


Proceso de Corrección de Variables Atmosféricas



Parametrizaciones del WRF

- Microphysics
- Cumulus parameterization
- Surface Layer
- Land-Surface Model
- Planetary Boundary Layer
- Atmospheric Radiation
- Physics Interactions

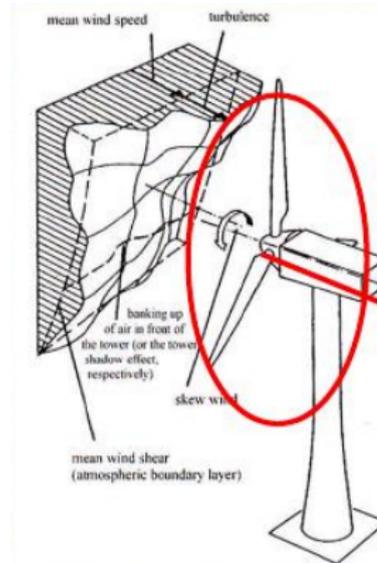


CLA → Energía Eólica

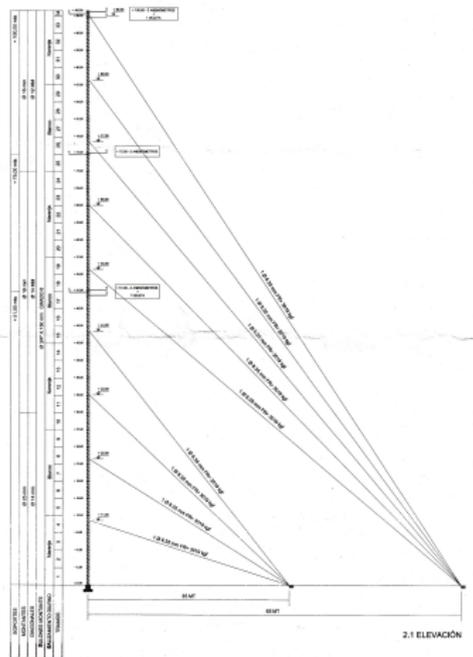
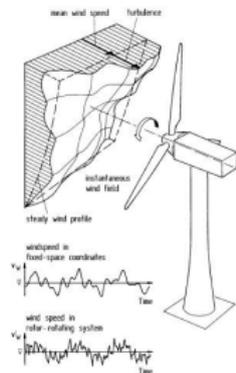
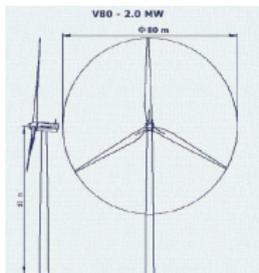
Importa describir el ciclo diario del perfil de velocidades en el área de barrido de palas de los aerogeneradores

Viento

$$\vec{V}(x, y, z, t), \text{ Estabilidad } \frac{\partial \theta}{\partial z}, \vec{V}^{\prime}(x, y, z, t), IT(t)$$



Perfil de velocidades en altura ciclo diario

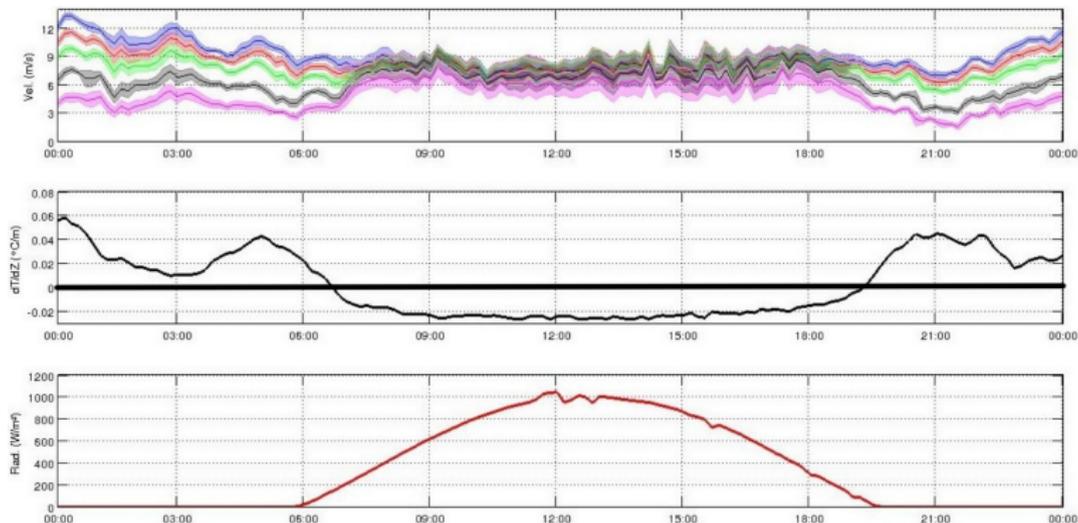


- Anemómetros: 101.8 m - 81.8 m - 60.8 m - 25.7 m - 10.1 m
- Termómetros: 100.8 m 3.4 m

Interacción entre $I_{Rad} \rightarrow \frac{\delta T}{\delta Z}_{10min} \rightarrow (\bar{V}_{10min}, \sigma_{10min})$

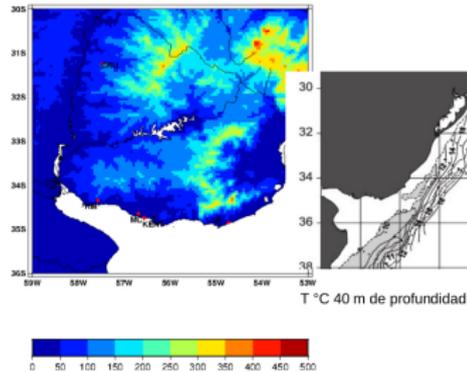
6 Diciembre de 2014

- Anemómetros: 101.8 m - 81.8 m - 60.8 m - 25.7 m - 10.1 m
- Termómetros: 100.8 m 3.4 m



Radiación solar y la Mesoescala

Tower measurement location in topography

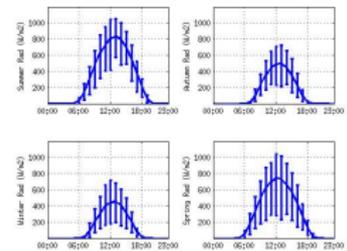
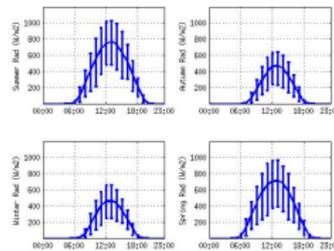
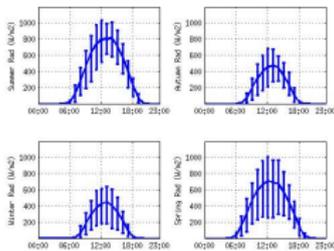


Section	T (°C)
(a) August 2003	
Mar del Plata	8.438 ± 1.106
Punta Médanos	8.963 ± 1.114
Río de la Plata	8.875 ± 1.263
Punta del Este	13.383 ± 2.556
Punta del Diablo	9.804 ± 0.743
Albardillo	15.687 ± 2.024
Río Grande	10.694 ± 0.309
Soldado	17.235 ± 1.014
	17.277 ± 0.948
	18.030 ± 1.113
	18.442 ± 0.630
(b) February 2004	
Mar del Plata	9.391 ± 3.221
Punta Médanos	9.934 ± 3.502
Río de la Plata	9.370 ± 3.580
Punta del Este	23.896 ± 0.139
Punta del Diablo	15.018 ± 0.266
Albardillo	24.318 ± 0.253
Río Grande	15.407 ± 0.089
Soldado	24.281 ± 2.191
	16.023 ± 0.502
	20.104 ± 3.720
	23.515 ± 2.174
	23.558 ± 1.669

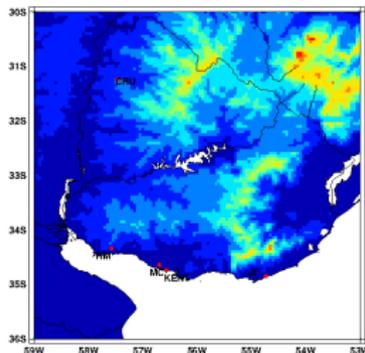
MC irradiance

Percentiles 16 % y 84 %
CR irradiance

J1 irradiance



Mesoescala $\frac{\delta T}{\delta Z}$

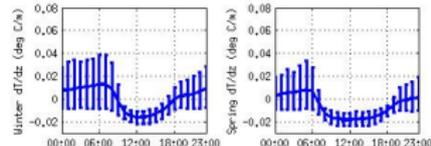
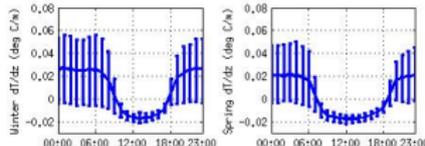
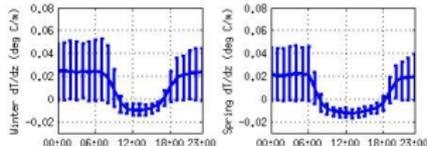
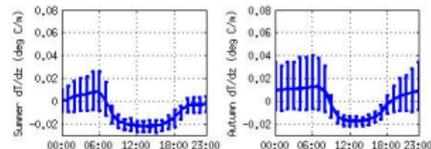
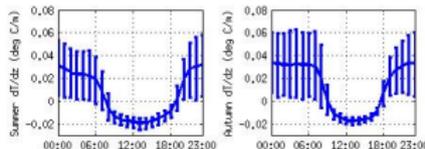
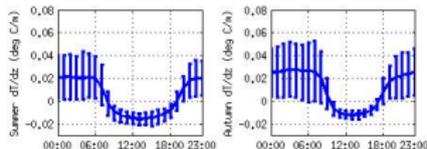


Percentiles 16% y 84 % $\frac{\delta T}{\delta Z}$ (°C/m)

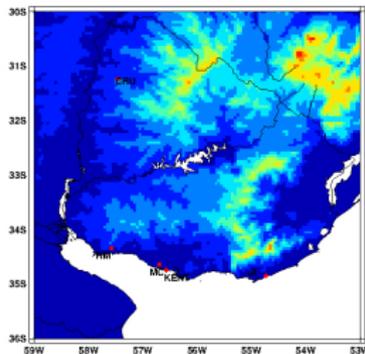
RM $\frac{\delta T}{\delta Z}$ 2 - 100 m

CR $\frac{\delta T}{\delta Z}$ 3 - 99 m

Jl $\frac{\delta T}{\delta Z}$ 12 - 98 m



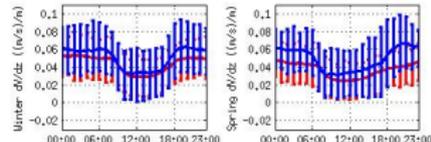
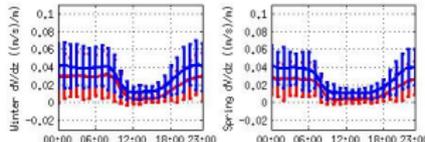
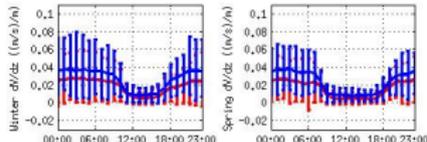
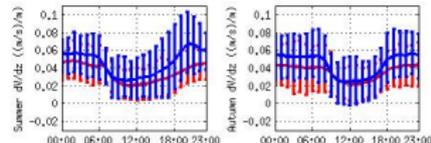
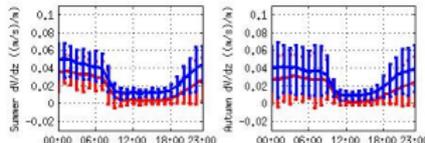
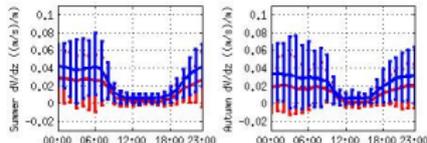
Mesoescala $\frac{\delta V}{\delta Z}$



Percentiles 16 % y 84 % $\frac{\delta V}{\delta Z}$ (1/s)
RM 63.5- 81 m (azul), and 81 - 101 m (rojo).

CR 60 - 80 m (azul) and 80 - 101 m (rojo)

Jl 12 - 40 m (azul) and 40 - 98 m (rojo)



Parametrizaciones de PBL

Different strategies for PBL parametrization (Stull, 1988) (Stensurd, 2007).- **Diferentes estrategias para resolver numéricamente la CLA “cierre”:**

- Friction in surface and scale of velocity, **tensión rasante en superficie y escala de velocidad**

$$u_*^2 = \frac{[\tau_{xz}^2 + \tau_{yz}^2]^{1/2}}{\rho} = [u'w'^2 + v'w'^2]^{1/2}$$

- A mixing length l , description of turbulence (Prandtl, 1925), mean wind velocity \bar{u} , surface layer temperature

$$\text{scale } \theta_* SL = \frac{-w'\theta'}{u_*} S \text{ a turbulent eddy } z'$$

reference level z , $u' = -\frac{\partial \bar{u}}{\partial z} z'$ and

$$w' = -c \left| \frac{\partial \bar{u}}{\partial z} \right| z' \text{ with } l^2 = cz'^2.$$

The most common assumption is $l = kz$ ($k = 0.40$, Von Karman constant) **longitud l como una escala de los vortices que producen la mezcla**

- Stability regime (Monin, 1954) (Busigner 1971)

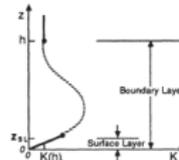
$$L = \frac{-\bar{\theta} u_*^3}{kg(w'\theta')}, \quad \zeta = \frac{z}{L} \text{ for a given height } (z).$$

($\zeta > 0$ stable, $\zeta < 0$ unstable, and $\zeta = 0$ is neutral)

$$\text{then } \phi_m(\zeta) = \frac{kz}{u_*} \frac{\partial \bar{u}}{\partial z} \text{ and } \phi_h(\zeta) = \frac{z}{\theta_*} \frac{\partial \bar{\theta}}{\partial z}$$

gradientes adimensionados en función de la estabilidad

- First-order closure * Cierre de Primer Orden: Nonlocal**



$$*** K_m = u_* kz \phi_m^{-1} \left(1 - \frac{z}{h}\right)^D ***$$

$$* K_m = l^2 \left| \frac{\partial \bar{u}}{\partial z} \right| * \text{Local}$$

- One-and-a-half-order closure * Cierre Orden 1.5**
 $e = (u'^2 + v'^2 + w'^2)/2$ (TKE)

$$\frac{\partial \bar{e}}{\partial t} = -u'w' \frac{\partial \bar{u}}{\partial z} - v'w' \frac{\partial \bar{v}}{\partial z} + \left(\frac{g}{\theta}\right) w'\theta' - \frac{\partial [w'((\rho'/\rho) + \theta)]}{\partial z} - \epsilon$$

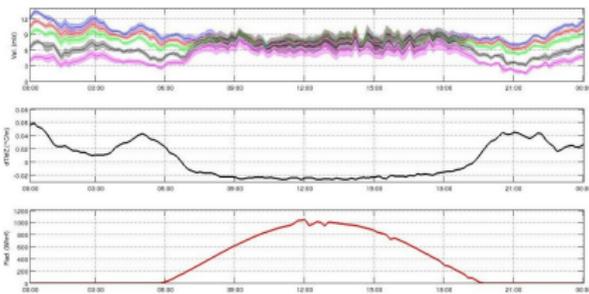
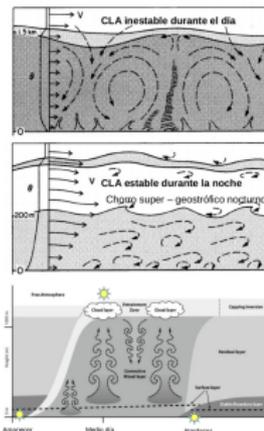
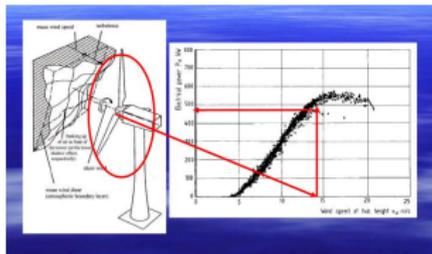
$$\frac{\partial (\theta'^2)}{\partial t} = -2w'\theta' \frac{\partial \bar{\theta}}{\partial z} - \frac{\partial w'\theta'^2}{\partial z} - 2\epsilon_\theta - \epsilon_R$$

Then:

$$*** w'u' = -K_m(\bar{e}, \bar{\theta}^2) \frac{\partial \bar{u}}{\partial z} ***$$

Local

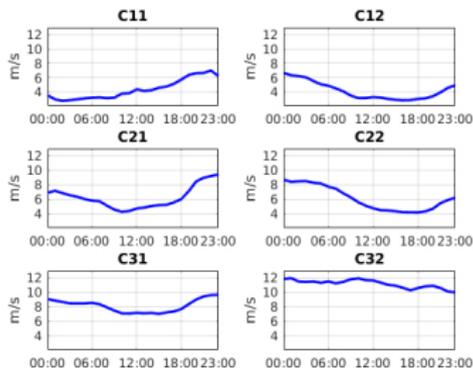
¿Desempeño de los modelos numéricos en la CLA?



¿Desempeño de los modelos numéricos en la CLA?

Esquema – CLA	Tipo	Orden de cierre	Altura de la CLA
YSU	No-local	1	Ri
MYJ	local	1.5	TKE
ACM2-Pleim	No-local inestabilidad; local estabilidad	1	Ri
Boulac	local	1,5	TKE
Bretherton-Park	local	1.5	Ri
GBM-TKE	local	1.5	Salto de la inversión térmica
Shin-Hong	No-local función de la resolución de la grilla	1	Ri
MYNN	local	1.5	TKE

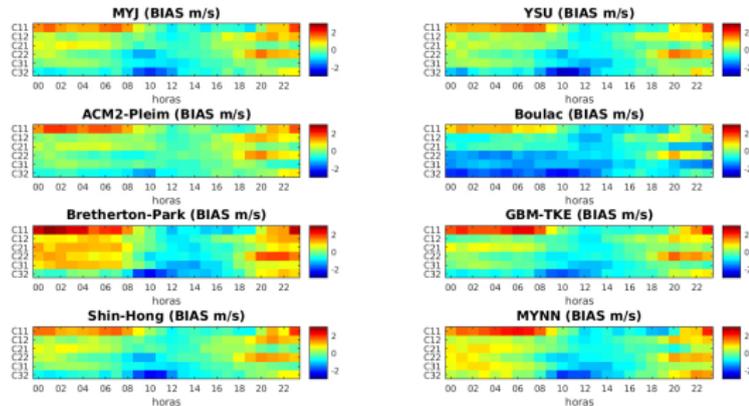
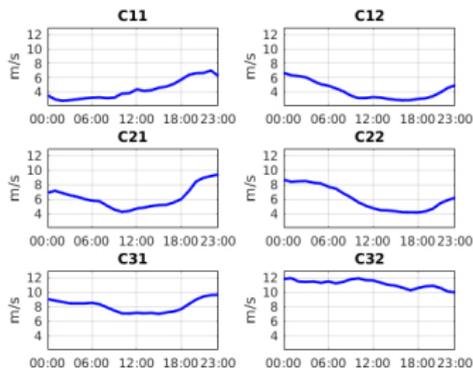
Cilios diarios medidos a 100 metros de altura



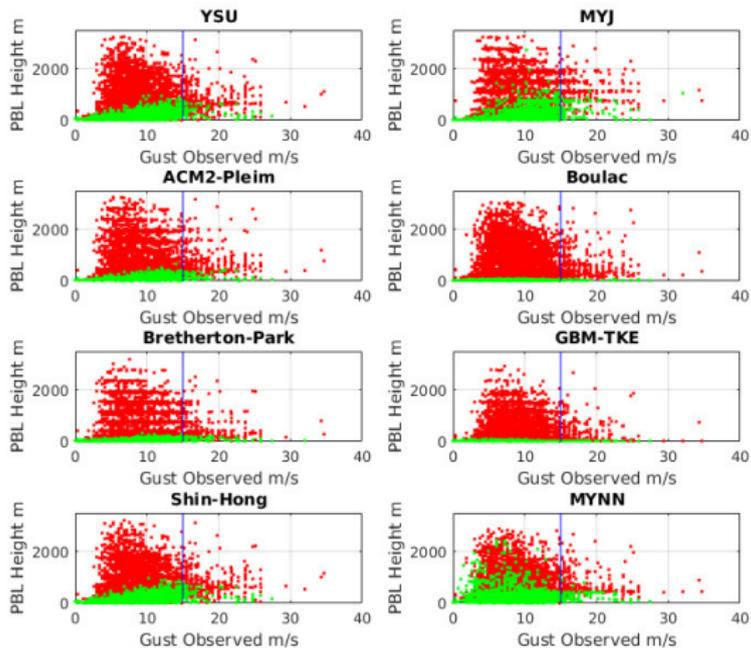
¿Desempeño de los modelos numéricos en la CLA?

Esquema – CLA	Tipo	Orden de cierre	Altura de la CLA
YSU	No-local	1	Ri
MYJ	local	1.5	TKE
ACM2-Pleim	No-local inestabilidad; local estabilidad	1	Ri
Boulac	local	1,5	TKE
Bretherton-Park	local	1.5	Ri
GBM-TKE	local	1.5	Salto de la inversión térmica
Shin-Hong	No-local función de la resolución de la grilla	1	Ri
MYNN	local	1.5	TKE

Cilos diarios medidos a 100 metros de altura

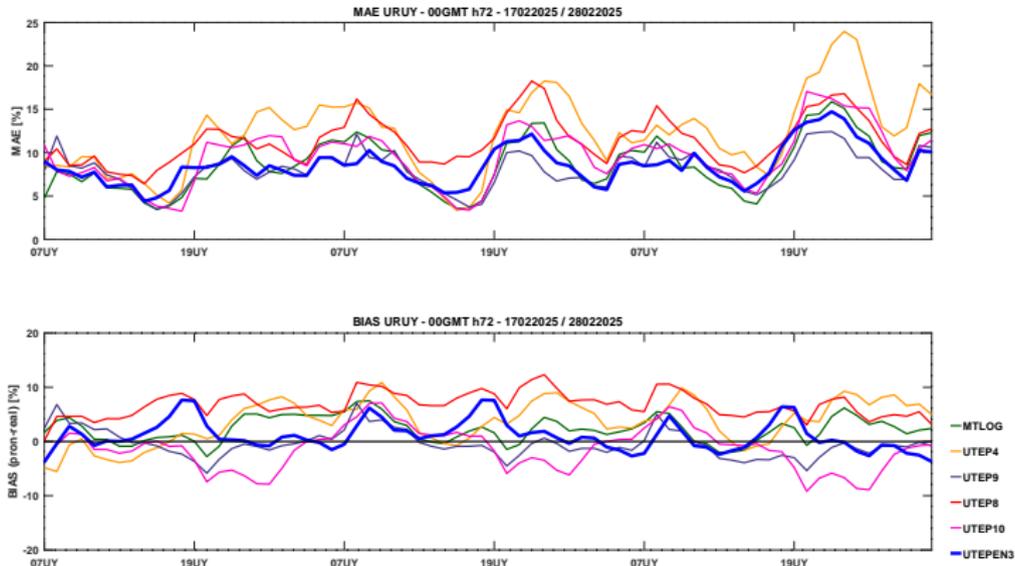


Plotear



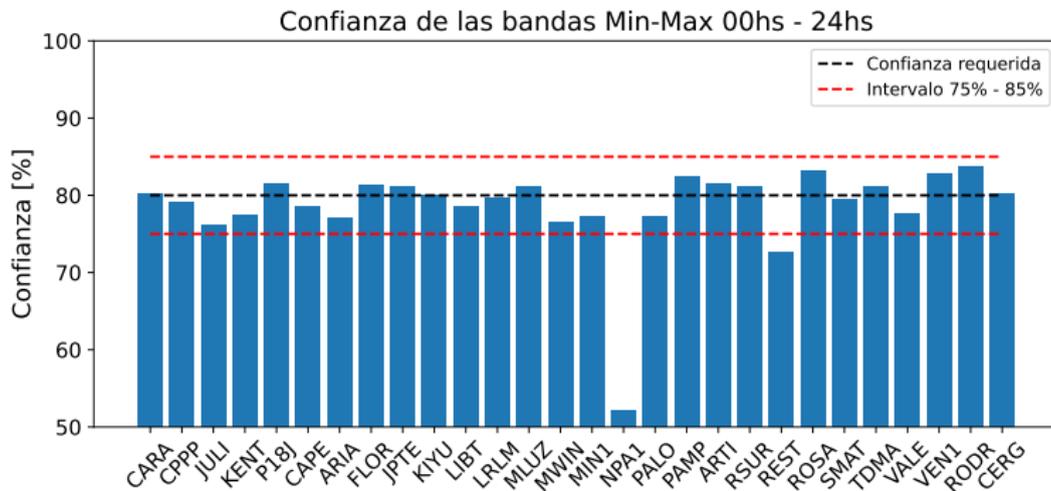
Evaluar el error en función del horizonte temporal k

Comparar modelos.



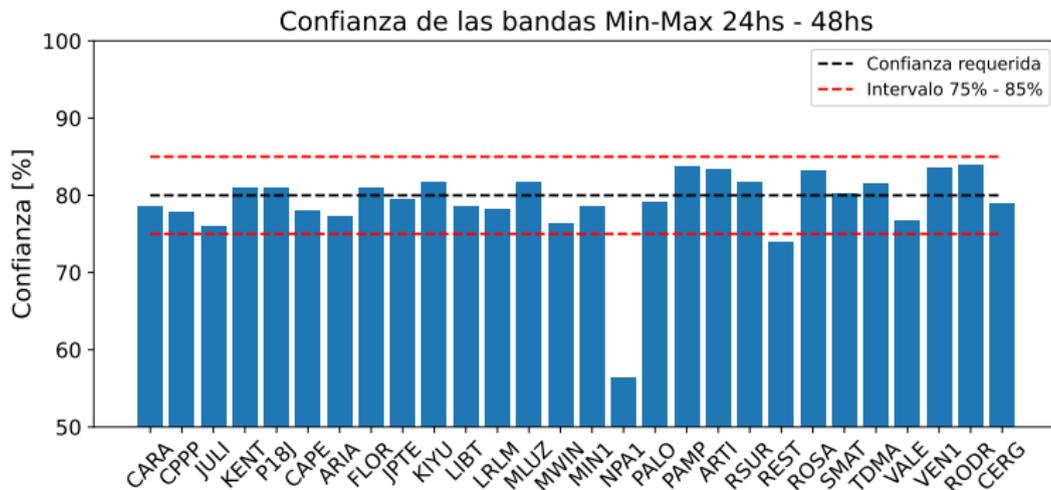
Evaluar Desempeño de bandas de confianza función del horizonte temporal k

Evaluar Desempeño por Parque Eólico entre 0 y 24 horas.



Evaluar Desempeño de bandas de confianza función del horizonte temporal k

Evaluar Desempeño por Parque Eólico entre 24 y 48 horas.



Evaluar Desempeño de bandas de confianza función del horizonte temporal k

Evaluar Desempeño por Parque Eólico entre 48 y 72 horas.

