Control de Calidad de Radiación Solar

Energía Solar Fotovoltaica

Oscar Perpiñán Lamigueiro

Universidad Politécnica de Madrid

Variable aleatoria y proceso estocástico

- Una **variable aleatoria** es una función que asigna un único numero real a cada resultado de un espacio muestral en un experimento.
- Un **proceso estocástico** es una variable aleatoria que evoluciona a lo largo del **tiempo** (p.ej. la radiación).

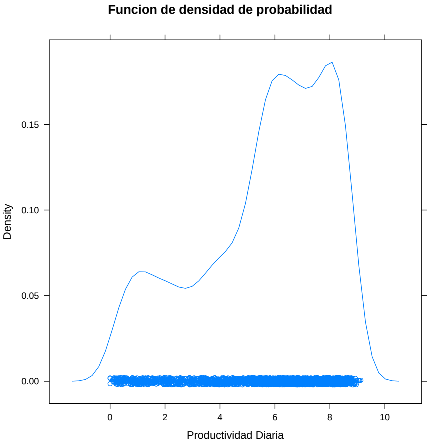
Función de densidad de probabilidad

La función de densidad de probabilidad, $f(X)$, de una variable aleatoria **asigna probabilidad** a un suceso:

$$P(a < X < b) = \int_a^b f(x)dx$$

$$P(X < b) = \int_{-\infty}^b f(x)dx$$

$$P(X > a) = \int_a^{\infty} f(x)dx$$



Media, varianza y desviación estándar

- La **media** de una variable aleatoria es el **centro de masas** de su función densidad de probabilidad:

$$\mu_X = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

- La **varianza** de una variable aleatoria es la **media del cuadrado de las desviaciones** respecto a la media:

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f(x)dx$$

- La **desviación estándar** es la raiz cuadrada de la varianza: $\sigma_X = \sqrt{\sigma_X^2}$

Combinación lineal de variables aleatorias

- La **media de la suma** de varias variables aleatorias **independientes** es la suma de las medias:

$$\mu_{X_1+\ldots+X_n} = \mu_{X_1} + \ldots + \mu_{X_n}$$

- La **varianza de la *suma o resta*** de varias variables aleatorias **independientes** es la **suma** de las varianzas:

$$\sigma^2_{X_1\pm\ldots\pm X_n} = \sigma^2_{X_1} + \ldots + \sigma^2_{X_n}$$

Media y varianza de la media muestral

- Una **muestra de una población** es un conjunto de variables aleatorias independientes $(X_1 ... X_n)$.
- Si se toma una muestra de una población, la media de la muestra es otra variable aleatoria (que es una suma de variables aleatorias)

$$\overline{X} = \frac{1}{i = n} \sum_{n} X_i$$

Media y varianza de la media muestral

Sea una población cuya media es μ y su varianza es σ^2:

- La **media de la media muestral** es la media poblacional:

$$\mu_{\overline{X}} = \frac{1}{i = n} \sum_{n} \mu_{X_i} = \mu$$

- La **varianza de la media muestral** es la varianza poblacional dividido por el número de muestras.

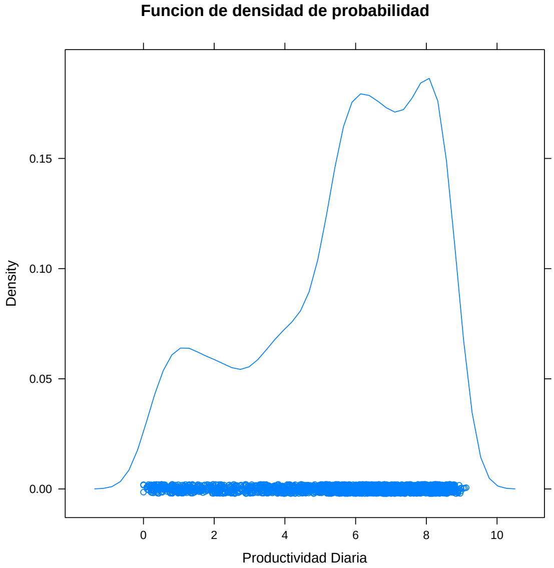
$$\sigma^2_{\overline{X}} = \sigma^2_{\frac{1}{n}X_1} + ... + \sigma^2_{\frac{1}{n}X_n} = \frac{\sigma^2}{N}$$

Por tanto, una forma de **reducir la incertidumbre** es realizar la **medida en repetidas ocasiones**.

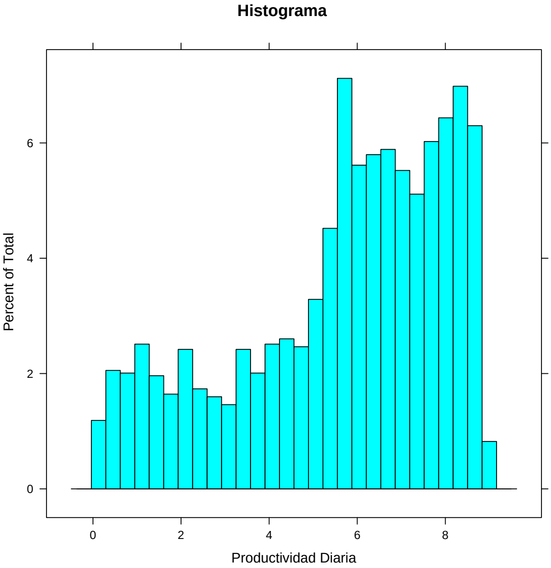
Mediana y cuartiles

- La **mediana** divide el conjunto de valores de la variable en **dos mitades** iguales (divide el area encerrada por la función densidad de probabilidad en dos partes iguales).
- Los **cuartiles** dividen este area en **cuatro** partes iguales.
- El area encerrada entre cada par de cuartiles es igual al 25% del total.
- La **mediana** es el **segundo cuartil**.
- La **distancia intercuartil** (definida entre los cuartiles 1 y 3) es una **medida de la dispersión** de la variable.

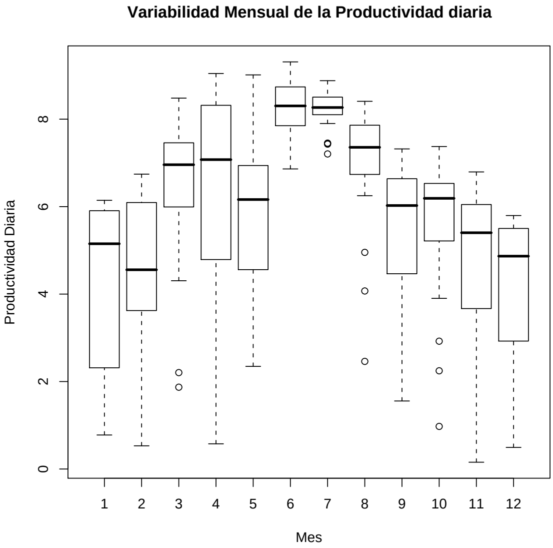
Función de Densidad de Probabilidad



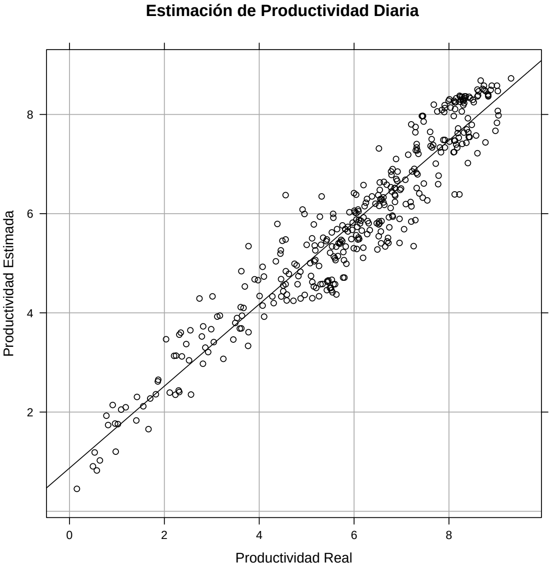
Histograma


Histograma
Percent of Total
0
2
4
6
0
2
4
6
8
Productividad Diaria

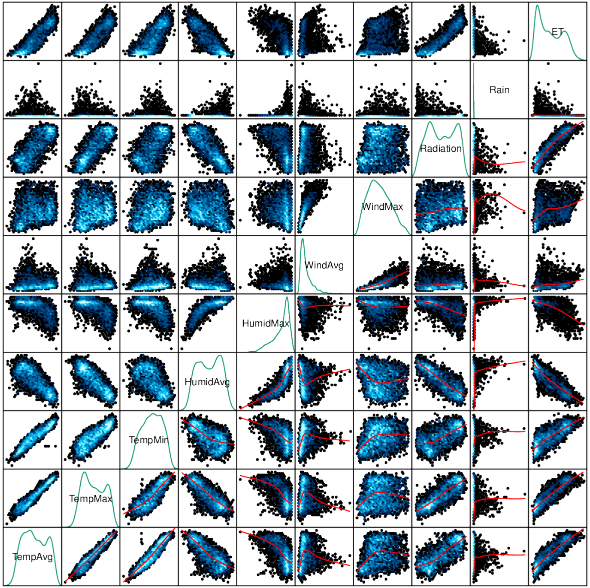
Gráficos boxplot



Gráficos de dispersión



Matrices de gráficos de dispersión



Introducción

Las medidas recogidas por estaciones meteorológicas se deben filtrar para eliminar datos erroneos.

- Límites Físicos
- Tests de persistencia
- Tests de rampas (irradiancia)
- Tests de envolvente (medida de varias componentes)
- Coherencia espacial
- Coherencia estadística

Límites físicos

Irradiación Diaria

- La radiación global en el plano horizontal debe ser inferior a la extraterrestre ($K_{td} \leq 1$)

$$G_d(0) \leq B_{od}(0)$$

- El índice de claridad debe ser superior a 0.03

$$K_{td} = \frac{G_d(0)}{B_{od}(0)} \geq 0.03$$

- La radiación global en el plano horizontal debe ser inferior a la de un modelo de cielo claro

Límites físicos

Irradiancia (intradiaria)

- El índice de claridad debe ser inferior a 1 cuando la altura solar es suficiente:

$$k_t < 1 \text{ si } \gamma_s > 2^\circ$$

- Límites inferiores para cielos cubiertos (baja transparencia atmosférica)

$$k_t \geq 10^{-4} \cdot (\gamma_s - 10^\circ) \text{ si } \gamma_s > 10^\circ$$

$$G \geq 0 \text{ si } \gamma_s \leq 10^\circ$$

Tests de variabilidad

Test de persistencia

$$\frac{1}{8}\overline{k_t} \leq \sigma_{k_t} \leq 0.35$$

La media y la desviación estándar se calculan con todas las muestras de un día completo.

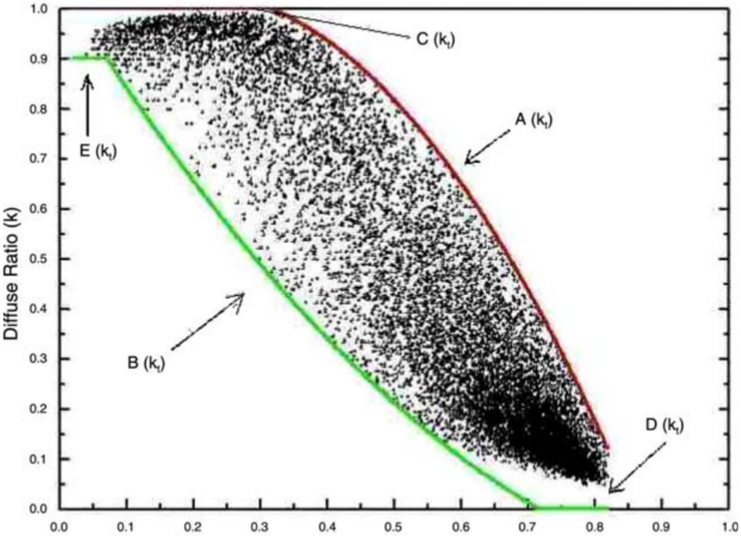
Test de rampas

$$|k_t(t) - k_t(t-1)| < 0.75 \quad \text{si} \quad \gamma_s(t) > 2^\circ$$

Límites a las variaciones de la irradiancia entre instantes sucesivos.

Tests de envolvente

Sólo para estaciones con medida simultánea de global y directa/difusa.



Coherencia espacial

- Las medidas de una estación se pueden comparar con las recogidas por estaciones cercanas.
- Esta comprobación debe realizarse con **datos agregados** (diarios) (la variabilidad espacial intradiaria puede ser alta)
- Esta comprobación debe realizarse con estaciones que tienen **clima y geografía similar**.

Coherencia espacial

Pasos

- Estimamos la irradiación en el lugar, x_0, con la interpolación espacial de las estaciones cercanas, x_i.
  - Los pesos w_i son una función inversa de la distancia (IDW).

$$\widehat{G}_d(x_0) = \frac{\sum_{i=1}^{N} w_i G_d(x_i)}{\sum_{i=1}^{N} w_i}$$

- Comparamos la irradiación estimada, $\widehat{G}_d(x_0)$, con la medida en la estación, $G_d(x_0)$.

$$\left|\widehat{G}_d(x_0) - G_d(x_0)\right|$$

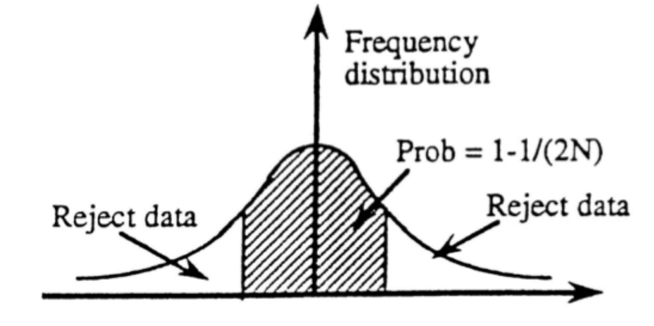
- La diferencia absoluta debe estar por debajo de un límite (p.ej. 50%)

Coherencia estadística

Una medida puede ser etiquetada como *outlier* si es poco probable que pertenezca a la misma distribución que el conjunto.

Método de Chauvenet

Una medida es un *outlier* si la probabilidad de obtener su desviación respecto de la media es inferior al inverso de 2 veces el número de elementos en el conjunto.



Método de Chauvenet

1. Sean $G_d(x_i)$ las medidas de radiación diaria del conjunto formado por N estaciones.
2. Se calcula la media, $\overline{G}_d$, la desviación estándar, σ_{G_d}.
3. Se calcula la distancia estadística de cada estación al conjunto:

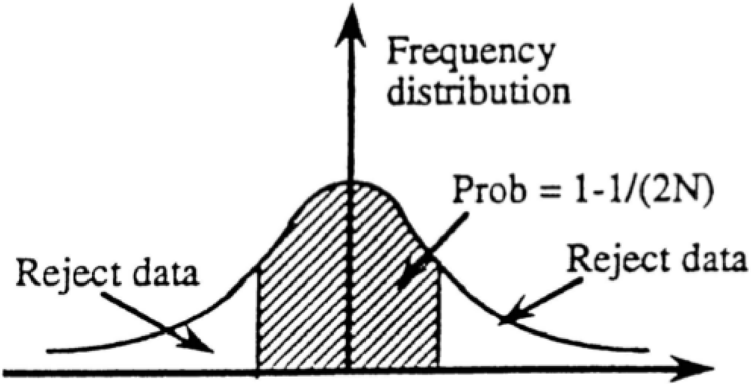
$$d_i = \frac{G_d(x_i) - \overline{G}_d}{\sigma_{G_d}}$$

4. En una distribución gaussiana se calcula la distancia estadística equivalente a la probabilidad límite, $1/2N$, teniendo en cuenta las dos colas.
    - Por ejemplo, para un conjunto de 10 estaciones cada cola es $1/40 = 0.025$, el límite es $|d_{max}| = 1.96$.
5. Aquellas observaciones que superan la distancia son marcadas como outliers.

Método de Chauvenet

$$d_i = \frac{G_d(x_i) - \overline{G}_d}{\sigma_{G_d}}$$

$$|d_i| > |d_{max}|$$



Desviación entre modelo y observación

- Sea O el conjunto de observaciones (medidas) de una variable aleatoria.

$$\mathbf{O} = \{o_1 \ldots o_n\}$$

- Sea M el conjunto de resultados de un modelo que aproxima el comportamiento de la variable medida.

$$\mathbf{M} = \{m_1 \ldots m_n\}$$

- La desviación entre modelo y observación es:

$$\mathbf{D} = \mathbf{M} - \mathbf{O} = \{(m_1 - o_1) \ldots (m_n - o_n)\} = \{d_1 \ldots d_n\}$$

Estimadores frecuentes: MBD y RMSD

- Mean Bias Difference (MBD), diferencia media (indica si el modelo sobreestima o subestima):

$$MBE = \overline{\mathbf{D}} = \overline{\mathbf{M}} - \overline{\mathbf{O}} = \frac{1}{n}\sum_{i=1}^{n}(m_i - o_i)$$

- Root Mean Square Difference (RMSD), diferencia cuadrático media:

$$RMSD = \left(\frac{1}{n}\sum_{i=1}^{n} d_i^2\right)^{1/2} = \left(\frac{1}{n}\sum_{i=1}^{n}(m_i - o_i)^2\right)^{1/2}$$

Estimadores frecuentes: MBE y RMSD

El RMSD agrega información del promedio y la varianza de la diferencia:

$$RMSD^2 = \overline{\mathbf{D}}^2 + \sigma_{\mathbf{D}}^2$$

donde la varianza de la diferencia (unbiased RMSD) se calcula:

$$\sigma_{\mathbf{D}}^2 = \frac{1}{n}\sum_{i=1}^{n}(d_i - \overline{\mathbf{D}})^2$$

Otros estimadores: MAD

- Mean Absolute Deviation (MAD):

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|d_i| = \frac{1}{n}\sum_{i=1}^{n}|m_i - o_i|$$

- El RMSD no es robusto (un error puntual puede distorsionar el estimador) y depende del número de muestras:

$$MAD \leq RMSD \leq n^{1/2}MAD$$

Otros estimadores: t y d

- t de Student (valores pequeños indican buen comportamiento del modelo)
  - Permite añadir intervalos de confianza a las diferencias entre modelo y observación

$$t = \left(\frac{(n-1)MBD^2}{RMSD^2 - MBD^2} \right)^{1/2}$$

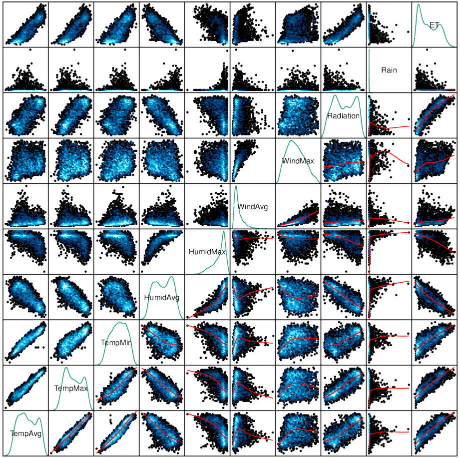
- d_1: Índice de concordancia de Willmott.
  - Limitado entre 0 (ausencia de concordancia) y 1 (concordancia total).
  - Robusto frente a *outliers*.

$$d_1 = 1 - \frac{\sum_{i=1}^{n} |m_i - o_i|}{\sum_{i=1}^{n} \left(\left| m_i - \overline{\mathbf{O}} \right| + \left| o_i - \overline{\mathbf{O}} \right| \right)}$$

Correlación

El coeficiente de correlación entre dos conjuntos de datos es una medida numérica de la relación **lineal** entre los dos conjuntos (si la relación no es lineal, este coeficiente no sirve):

$$r = \frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(\frac{o_i - \overline{\mathbf{O}}}{\sigma_{\mathbf{O}}} \right) \cdot \left(\frac{m_i - \overline{\mathbf{M}}}{\sigma_{\mathbf{M}}} \right)$$



Diagramas de Taylor

- Desarrollando $\sigma_{\mathbf{D}}^2$ y teniendo en cuenta la definición de r:

$$\sigma_{\mathbf{D}}^2 = \sigma_{\mathbf{O}}^2 + \sigma_{\mathbf{M}}^2 - 2 \cdot \sigma_{\mathbf{O}} \cdot \sigma_{\mathbf{M}} \cdot r$$

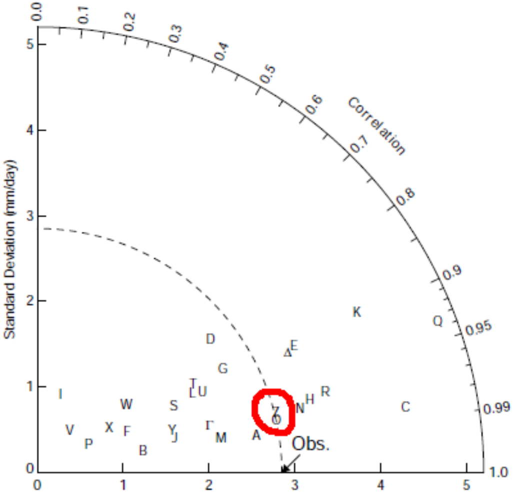
- Esta relación es semejante a la ley de los cosenos (c, a, b son lados de un triángulo y ϕ es el ángulo opuesto al lado c):

$$c^2 = a^2 + b^2 - 2 \cdot a \cdot b \cos \phi$$

Diagramas de Taylor

$$\sigma_{\mathbf{D}}^2 = \sigma_{\mathbf{O}}^2 + \sigma_{\mathbf{M}}^2 - 2 \cdot \sigma_{\mathbf{O}} \cdot \sigma_{\mathbf{M}} \cdot r$$

- $\sigma_{\mathbf{D}}^2$: Distancia al origen
- $\sigma_{\mathbf{O}}^2$: Eje horizontal
- $\sigma_{\mathbf{M}}^2$: Eje vertical
- r: acimut



Target Diagram

- Emplea la relación entre $RMSD$, $\sigma^2_{\mathbf{D}}$, y $\overline{\mathbf{D}}$, normalizadas con $\sigma_{\mathbf{O}}$:

$$RMSD' = RMSD/\sigma_{\mathbf{O}}$$

$$\sigma'_{\mathbf{D}} = \sigma_{\mathbf{D}}/\sigma_{\mathbf{O}}$$

$$\overline{\mathbf{D}}' = \overline{\mathbf{D}}/\sigma_{\mathbf{O}}$$

$$RMSD'^2 = \sigma'^2_{\mathbf{D}} + \overline{\mathbf{D}}'^2$$

$$sign_{\sigma} = sign(\sigma_{\mathbf{M}} - \sigma_{\mathbf{O}})$$

- Incorporan el signo de la diferencia entre desviaciones estándar de modelo y observación.

Target Diagram

- $\sigma'_{\mathbf{D}}$ (con signo): Eje horizontal
- $\overline{\mathbf{D}}'$: Eje vertical
- $RMSD'^2$: Distancia al origen