

# Predicción de radiación solar por medio de la combinación de modelos meteorológicos y técnicas de postproceso estadístico

Jose Luis Casado ([jcasador@aemet.es](mailto:jcasador@aemet.es))

Área de Aplicaciones, AEMET

## Marco general



Cobertura de la demanda eléctrica peninsular. Año 2018	
Nuclear	20.6%
Viento	19.0%
Carbón	13.5%
Hidroeléctrica	13.2%
Cogeneración	11.2%
Ciclo combinado	10.2%
Solar fotovoltaica	2.9%
Solar térmica	1.7%

Fuente: REE - Avance del Informe del sistema eléctrico español, 2018

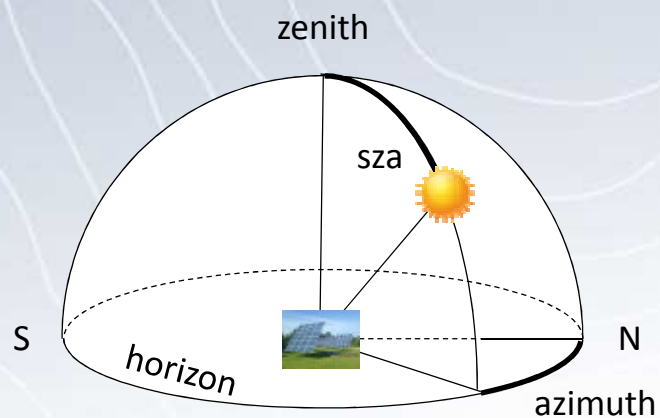
## Marco general



Parámetro principal:  
radiación global (GHI)



Parámetro principal:  
radiación directa normal (DNI)



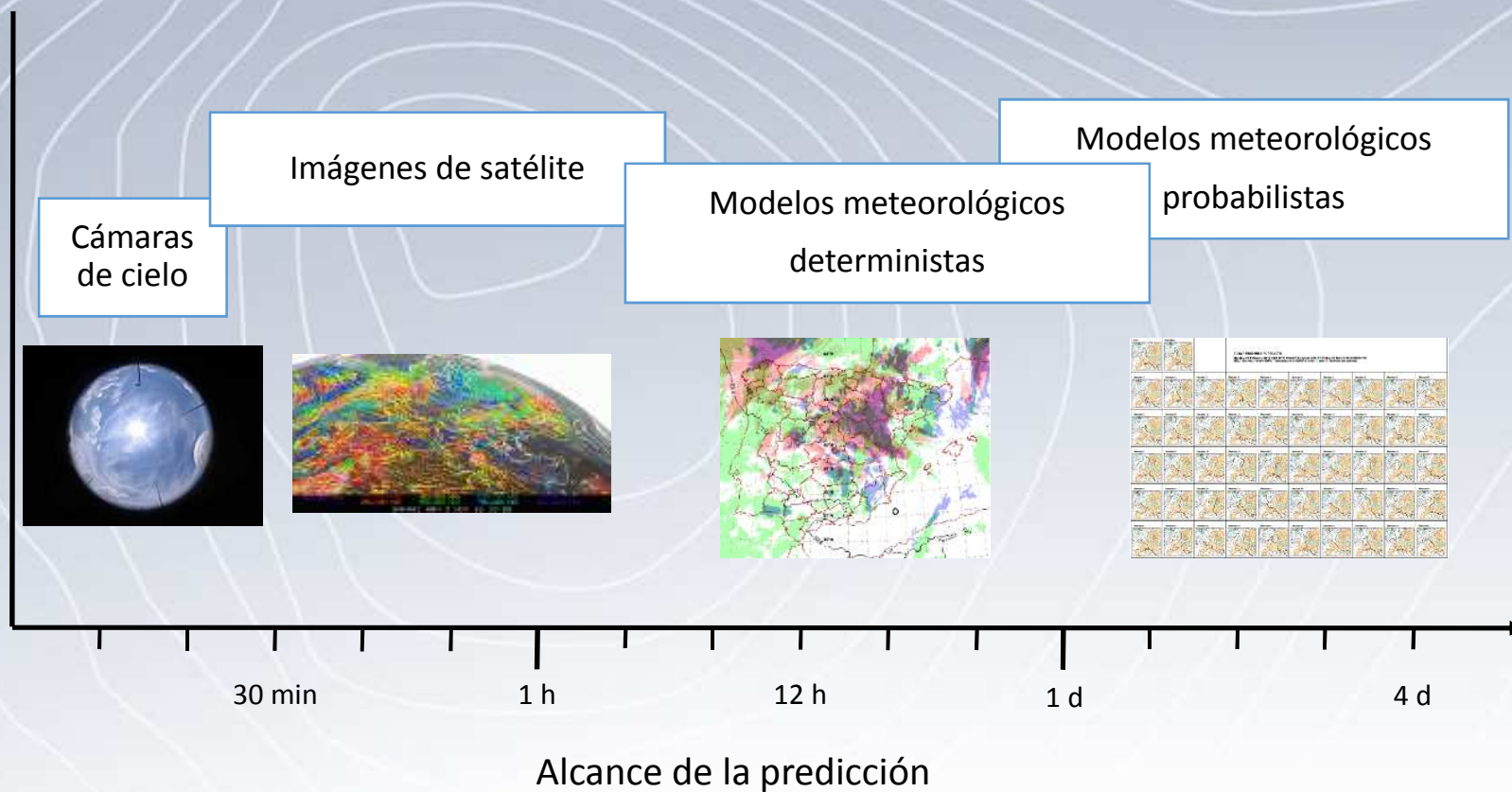
$$GHI = DNI \cos(sza) + Diffuse$$

sza: ángulo cenital

## Predicción con modelos meteorológicos



## Predicción con modelos meteorológicos



## Predicción con modelos meteorológicos



## Modelos meteorológicos utilizados en AEMET:

### ECMWF High Res

- Global, determinista
- 4 ejecuciones / día
- Resolución ~9 km
- Horario
- Con radiación directa desde 2011



### ECMWF EPS

- Global, por conjuntos (50 miembros)
- 4 ejecuciones / día
- Resolución ~18 km
- Horario
- Con radiación directa desde 2011



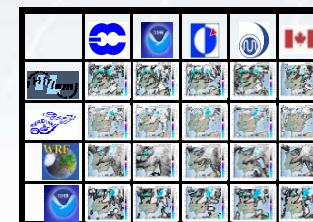
### Harmonie-Arome

- Área limitada, determinista
- 8 ejecuciones / día
- Resolución 2.5 km
- Cada 15 minutos
- Con radiación directa desde 2015



### gSREPS

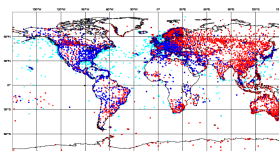
- Área limitada, por conjuntos (20 miembros)
- 2 ejecuciones / día
- Resolución 2.5 km
- Cada 3 horas
- Con radiación directa desde 2017



## Predicción con modelos meteorológicos

### Modelos meteorológicos

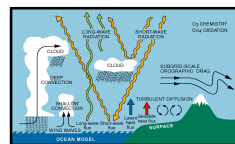
#### Observaciones



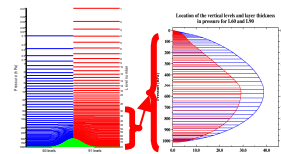
#### Asimilación



#### Parametrizaciones (incluyendo radiación)



#### Resolución numérica



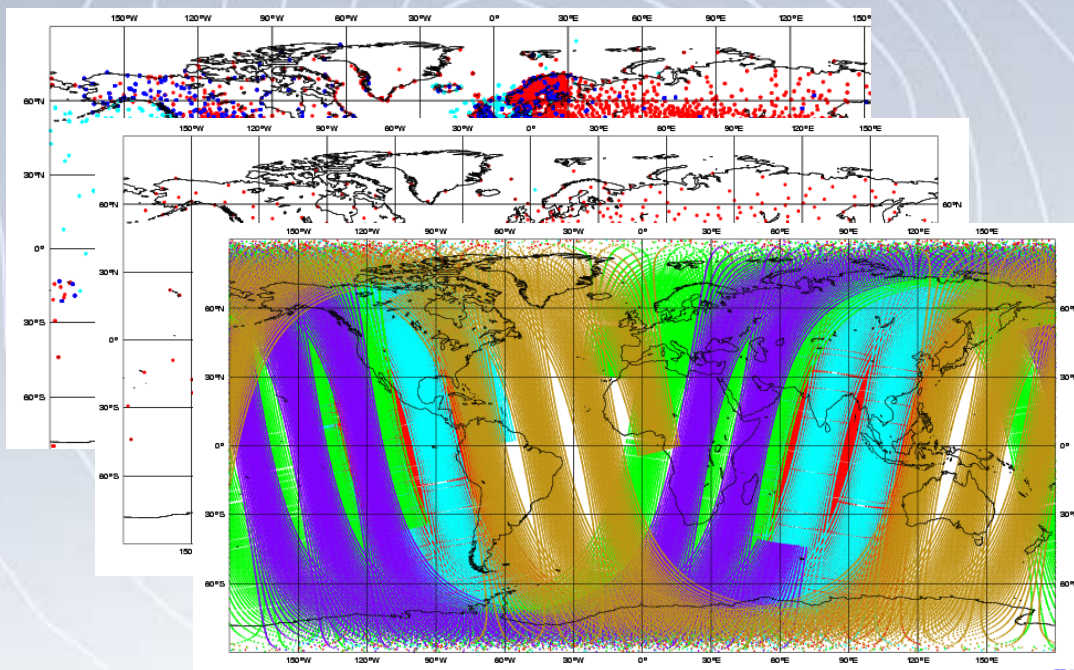
### ¿Cuáles son los problemas en la predicción meteorológica?

- Incertidumbre en las condiciones iniciales. Carencia y errores en las observaciones.
- Los errores de predicción dependen del tipo de flujo atmosférico.
- No linealidad.
- Multiplicidad de escalas espaciales y temporales.
- Conocimiento insuficiente de las leyes de la física. Necesidad de las parametrizaciones.
- Resolución limitada.



## Predicción con modelos meteorológicos

Incertidumbre en las condiciones iniciales. Carencia y errores en las observaciones



© 2019 European Centre for Medium-Range Weather Forecasts (ECMWF)

Millones de observaciones utilizadas.

La inmensa mayoría provienen de los satélites.

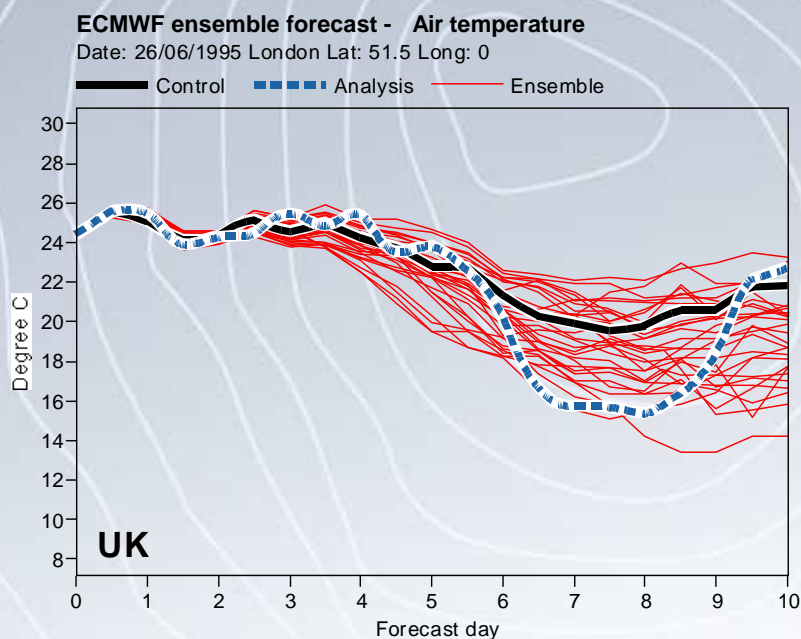


## Predicción con modelos meteorológicos

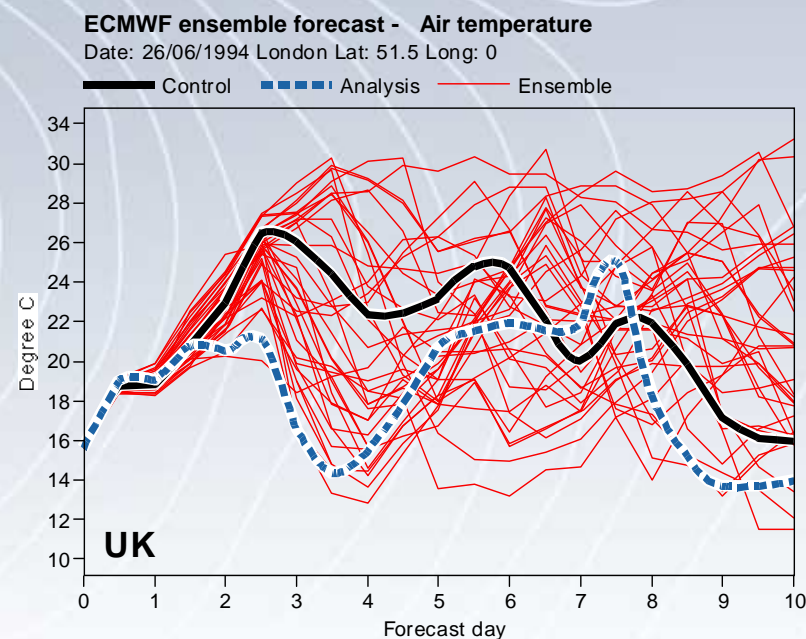


Los errores de predicción dependen del tipo de flujo atmosférico:

26 Junio 1995



26 Junio 1994

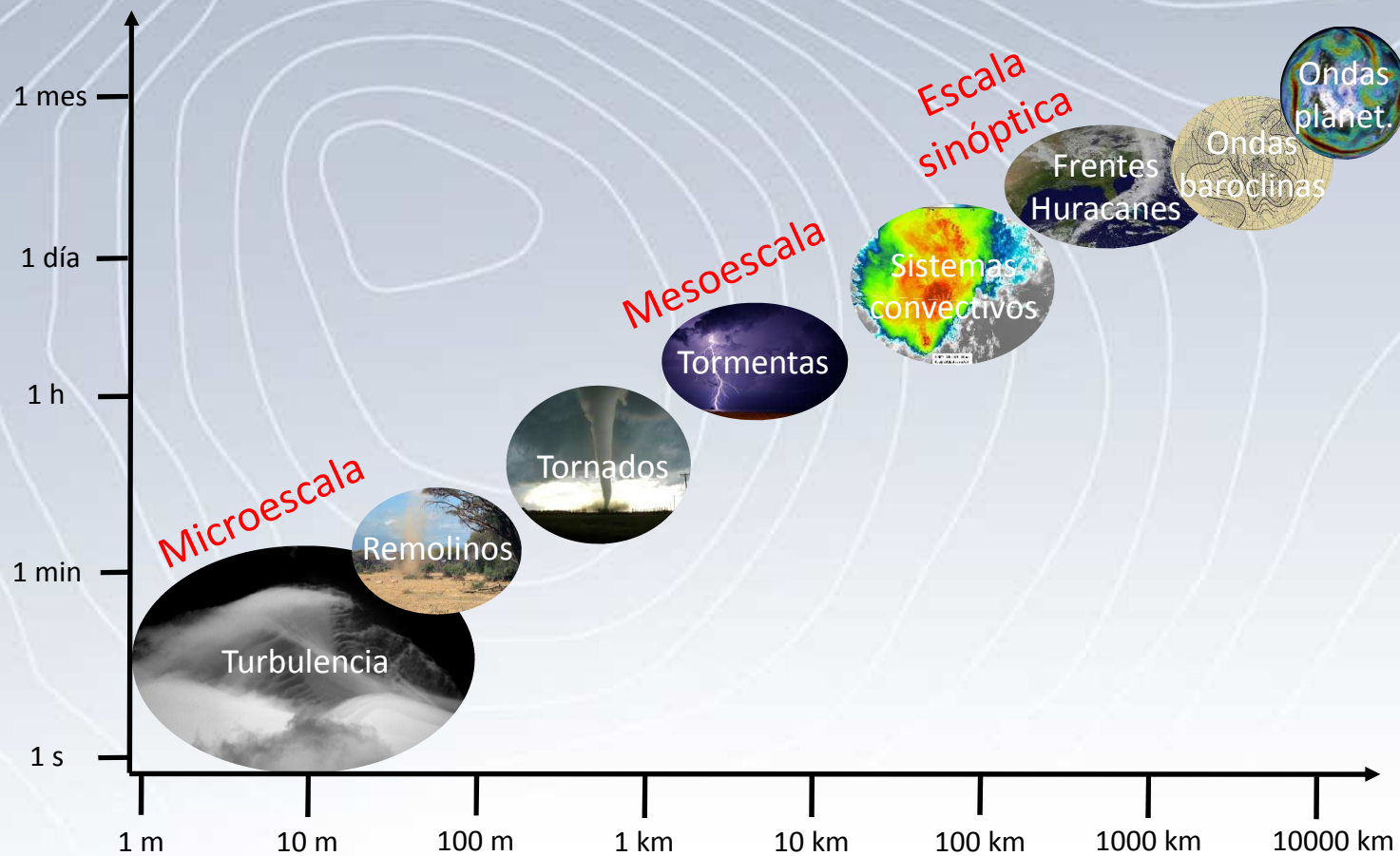


© 2019 European Centre for Medium-Range Weather Forecasts (ECMWF)

Si las predicciones son coherentes (pequeña dispersión), la atmósfera está en un estado más predecible que si las predicciones divergen (gran dispersión).

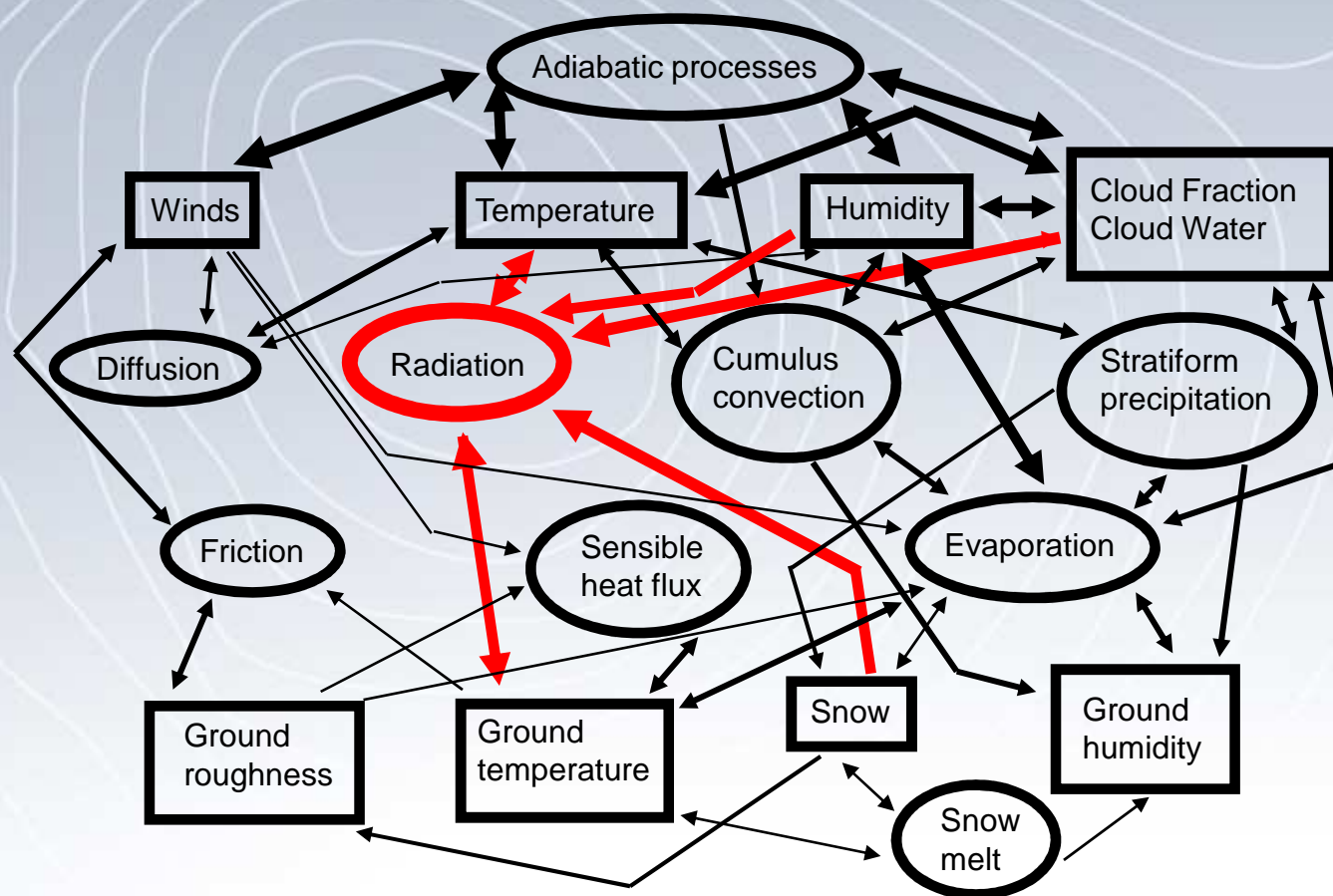
## Predicción con modelos meteorológicos

## Multiplicidad de escalas espaciales y temporales:



## Predicción con modelos meteorológicos

Conocimiento insuficiente de las leyes de la física. Parametrizaciones físicas.





## Predicción con modelos meteorológicos



Resolución limitada:

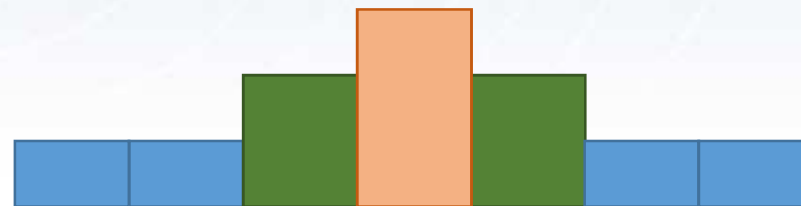
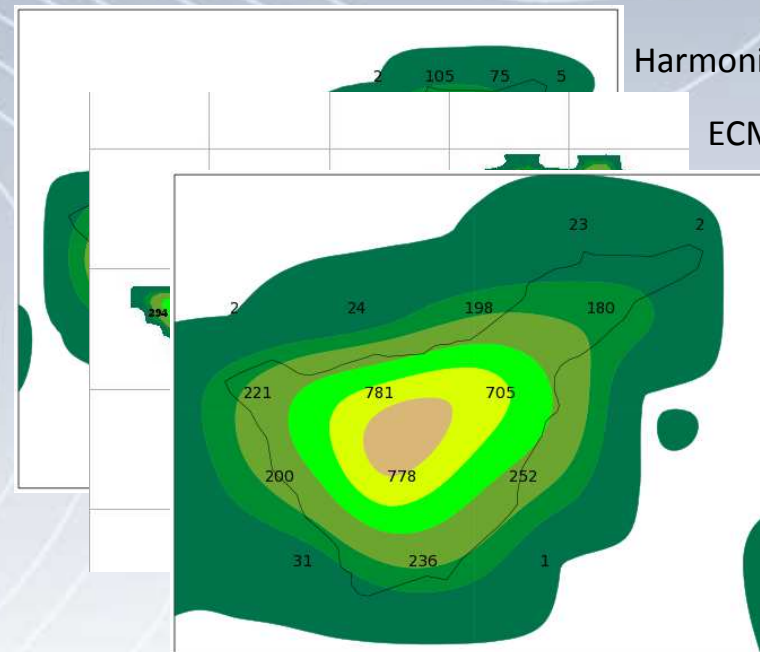
Realidad



ECMWF High Res

Harmonie-Arome

ECMWF EPS





Utilización de métodos estadísticos.  
Consideraciones específicas para la radiación

## Utilización de métodos estadísticos. Consideraciones



### Tamaño del problema:

900.000.000 puntos de grid X 10 variables  
7.500.000 puntos de grid X 200 variables

365 días al año

2 pasadas al día

120 alcances por pasada

20 años de predicciones

50 miembros en un ensemble

$10^{18}$  datos = 1000 petadatos

Tamaño real MARS (base de datos del ECMWF): docenas de petabytes

## Utilización de métodos estadísticos. Consideraciones



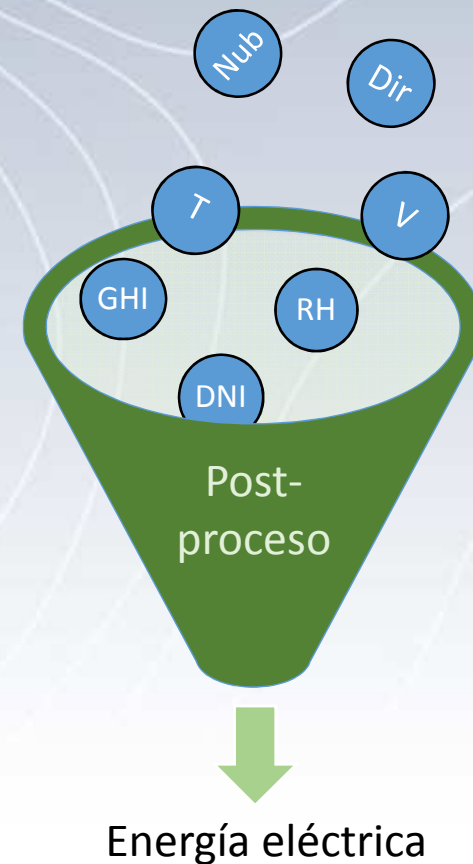
Queremos predecir el comportamiento futuro de un parámetro (por ejemplo, la energía eléctrica) en función de varias variables de entrada o predictandos que dependen de la meteorología.

Como primera aproximación, podemos suponer que la dependencia que tienen es lineal:

$$\text{Energía eléctrica} = a_0 + a_1 DNI + a_2 GHI + a_3 T + a_4 \text{Nub} + a_5 V + \dots$$

Algunos parámetros de entrada son obvios (radiación solar, nubosidad), pero podemos incluir también otros que quizás tengan una influencia indirecta.

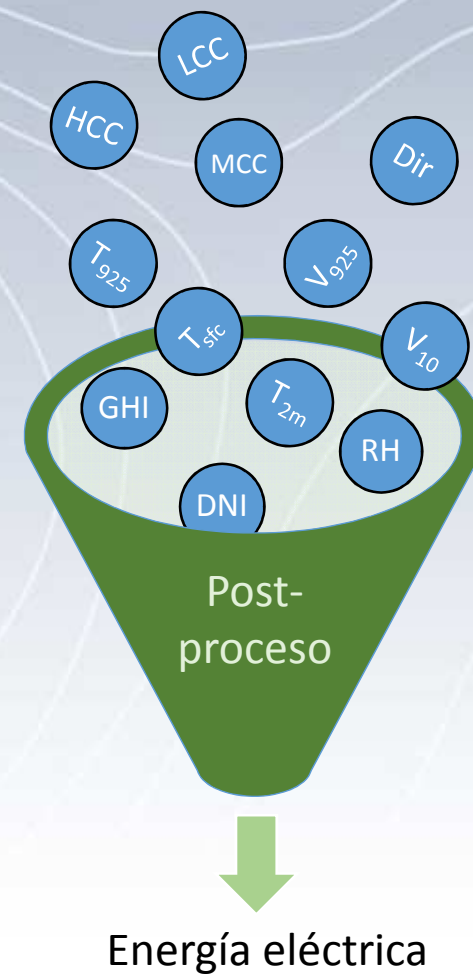
Tendremos que calcular los coeficientes  $a_i$  con algún método de tipo estadístico, por ejemplo por medio de una regresión.



## Utilización de métodos estadísticos. Consideraciones

Pero hay ciertos parámetros que pueden tomar distintos valores según la altura.

Por ejemplo la temperatura ( $T_{sfc}$ ,  $T_{2m}$ ,  $T_{925mb}$ ,  $T_{850mb}$ ), o la nubosidad (baja – LCC, media – MCC, alta – HCC).





## Utilización de métodos estadísticos. Consideraciones

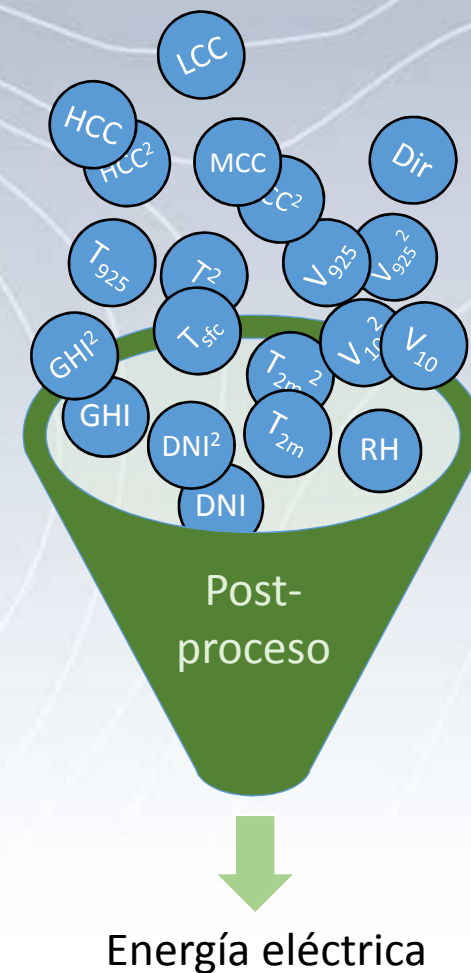
En realidad la dependencia respecto a otras variables no tiene por qué ser lineal. Será más genérica:

$$\text{Energía eléctrica} = f(\text{DNI}, \text{GHI}, T_{2m}, \dots)$$

Podemos simular la no linealidad, por ejemplo, por medio de un desarrollo de Taylor:

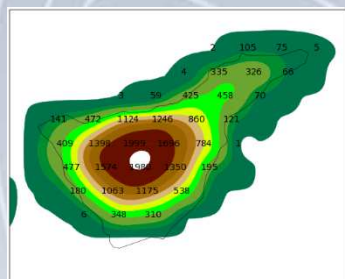
$$f(x) = f(0) + f(0)'x + \frac{f(0)''}{2!}x^2 + \dots$$

Introduciremos términos de orden superior (como  $\text{DNI}^2$ ) y términos cruzados (como  $T_{2m}\text{DNI}$ ).

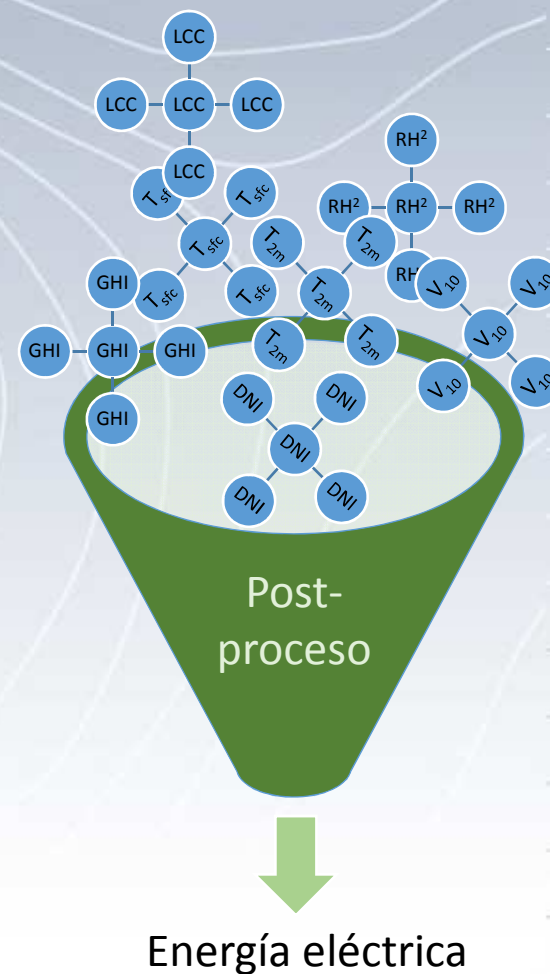
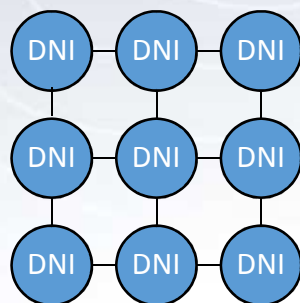


## Utilización de métodos estadísticos. Consideraciones

Hemos de tener en cuenta también que el punto de la malla más cercano no tiene por qué dar el mejor valor. Nuestra orografía es imperfecta, y es posible que algún otro punto de una aproximación más cercana a la realidad:



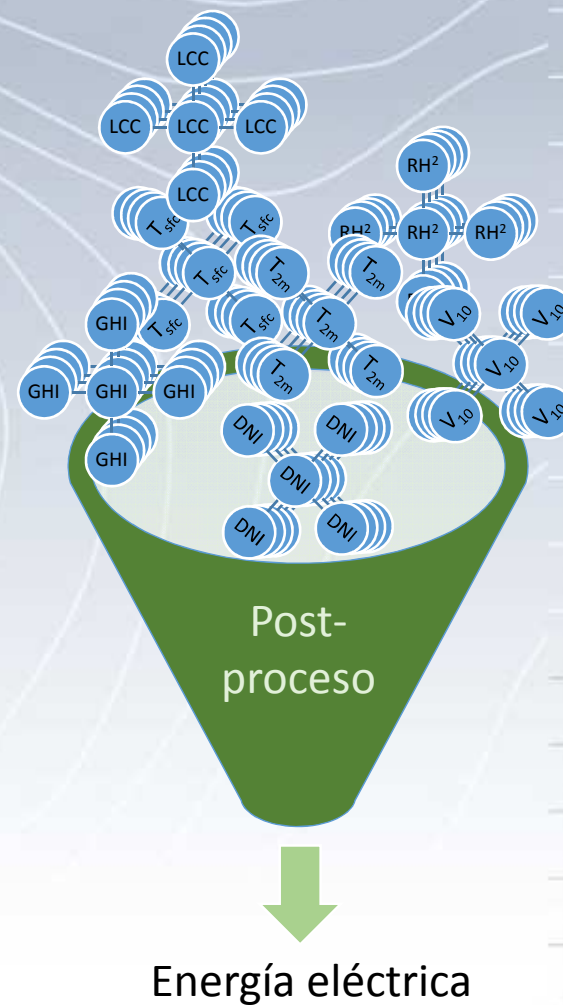
Es posible por ello que debamos incluir también como variables de entrada los 8 puntos que rodean al más cercano (o incluso más):



## Utilización de métodos estadísticos. Consideraciones

Si se dispone de un modelo de predicción por conjuntos, esto añade aún más variables de entrada.

De todas formas no es buena idea utilizar los modelos probabilistas de esta forma (y de hecho en modelos con miembros indistinguibles directamente no se puede).



### Situación:

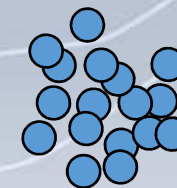
- Tenemos una gran cantidad de datos con los que podemos entrenar, y una serie de predictandos que pueden ayudar a predecir nuestra producción.
- Sería útil conocer qué predictandos son útiles, y las relaciones entre ellos.
- Dado el volumen de información, no podemos hacer esto de forma manual.
- Los métodos estadísticos modernos de aprendizaje automático pueden ser la solución en este caso.



## Utilización de métodos estadísticos. Consideraciones



Un método tradicional utilizado en meteorología es el Model Output Statistics (MOS):



$$\text{Energía eléctrica}_{D+1} = a_0 + a_1 \text{DNI}_{D+1} + a_2 \text{GHI}_{D+1} + \dots$$

$$\text{Energía eléctrica}_{D+2} = a'_0 + a'_1 \text{DNI}_{D+2} + a'_2 \text{GHI}_{D+2} + \dots$$

Se establecen relaciones entre las predicciones a futuro de los parámetros de entrada y el valor observado de la variable de salida.

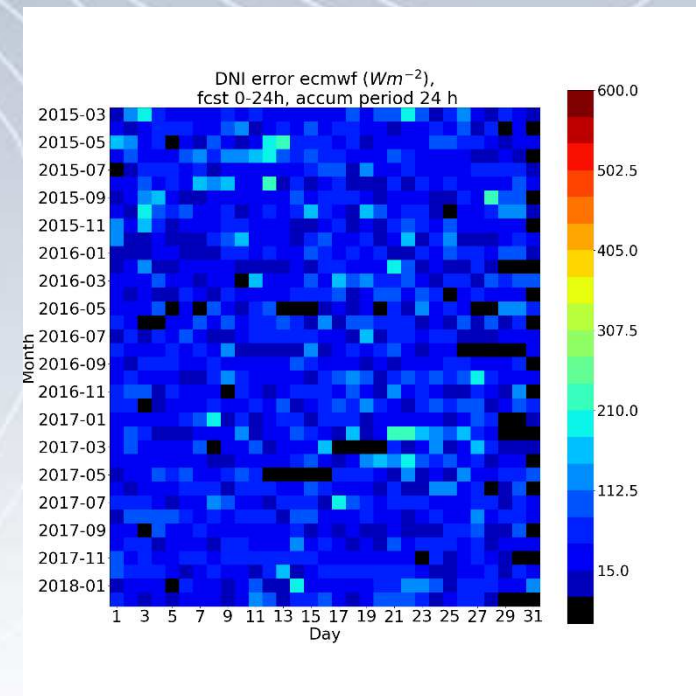
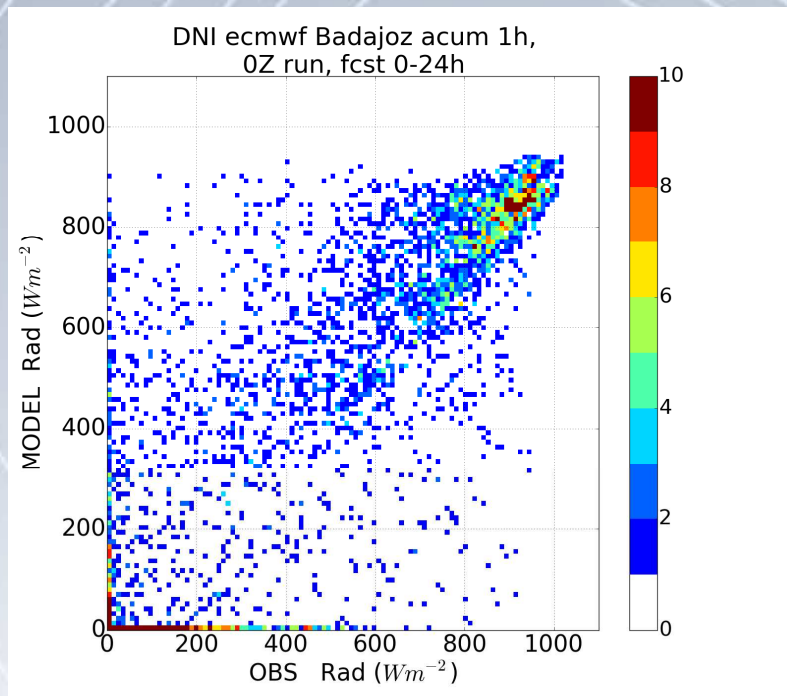
Hay que usar ecuaciones diferentes según el alcance de la predicción, ya que las características del error del modelo pueden cambiar dependiendo del alcance.



## Utilización de métodos estadísticos. Consideraciones



Advertencia 1: Conoce a tu enemigo: usa métodos de exploración



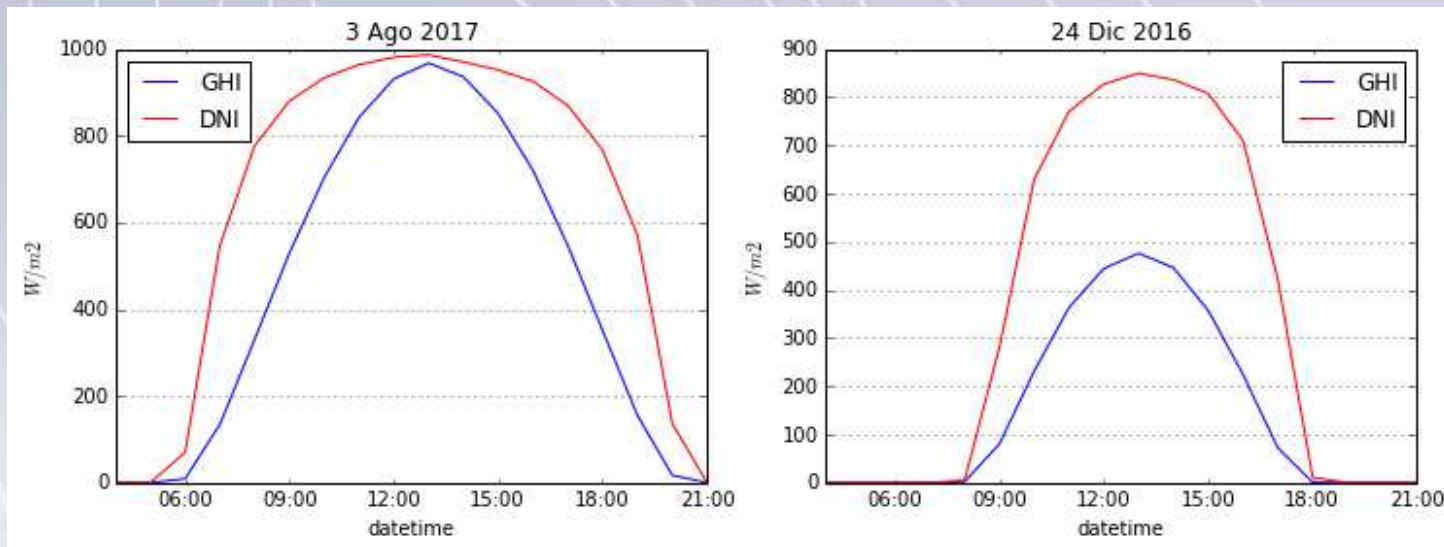
Al ser un problema de dimensionalidad alta, hay que elegir bien los parámetros cuya dependencia estudiamos.

## Utilización de métodos estadísticos. Consideraciones



Advertencia 2: Las propiedades físicas de los parámetros de entrada pueden cambiar con el tiempo

La radiación tiene un ciclo diurno y anual. ¿Incluimos la noche en la verificación?



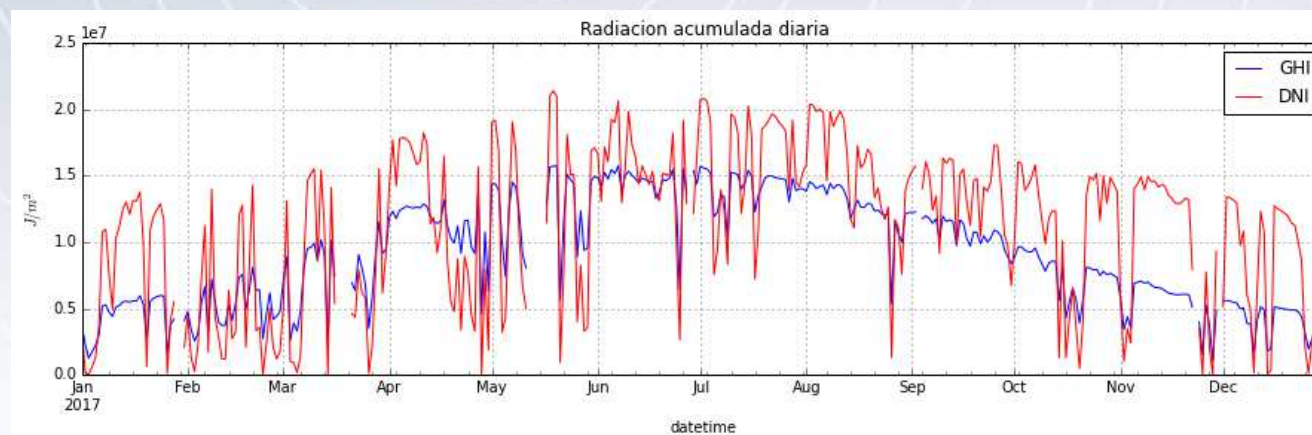
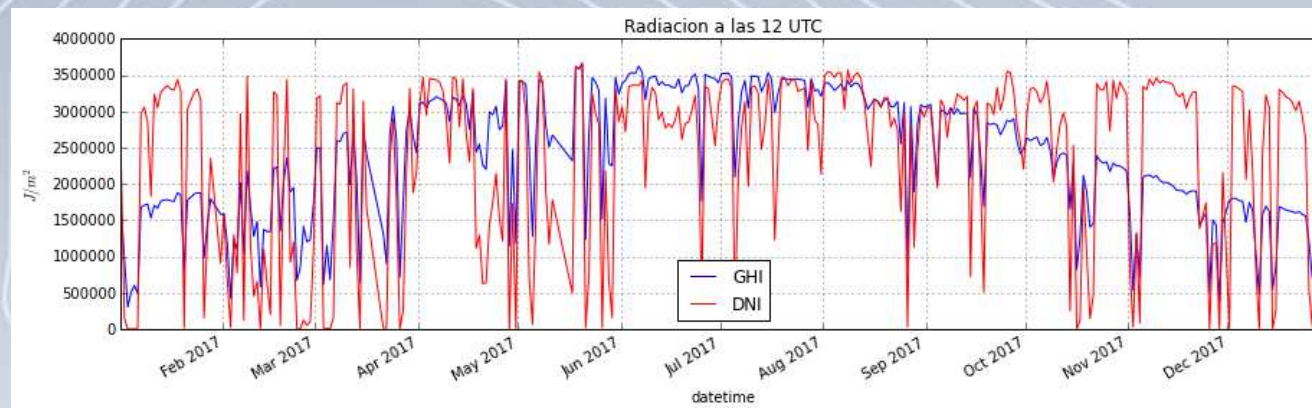
Si mezclamos datos no homogéneos podemos incurrir en errores.



## Utilización de métodos estadísticos. Consideraciones



Advertencia 2: Las propiedades físicas de los parámetros de entrada pueden cambiar con el tiempo



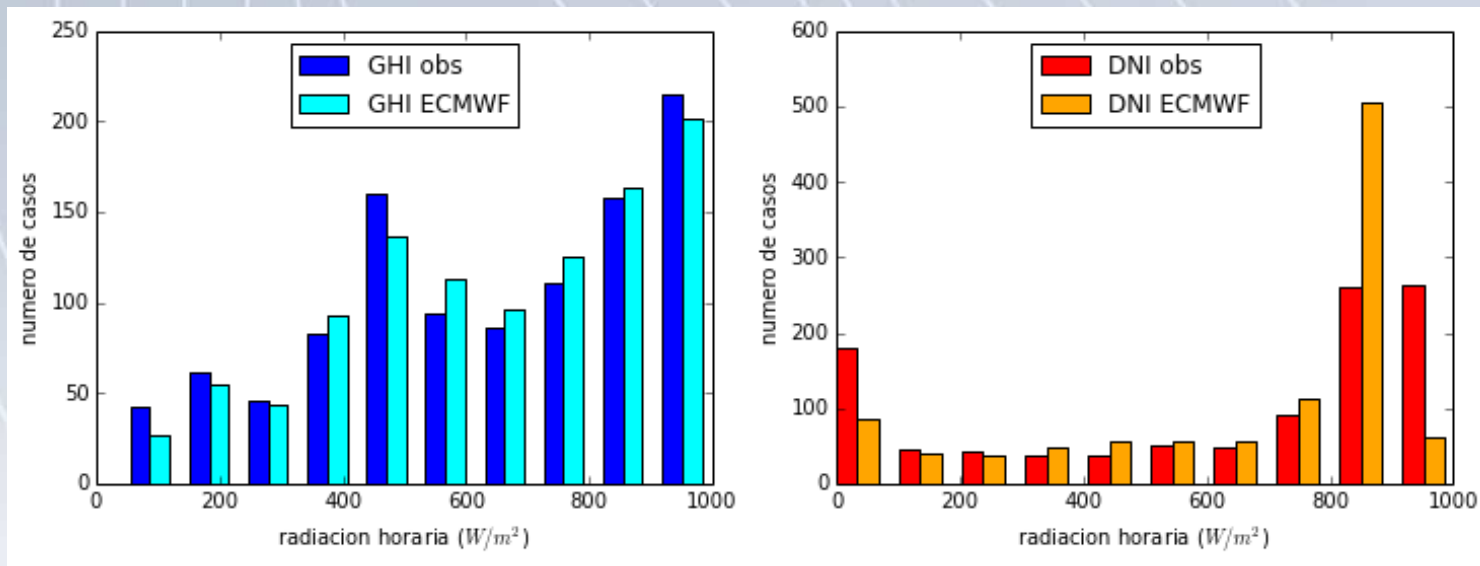


## Utilización de métodos estadísticos. Consideraciones



### Advertencia 3: La radiación tiene características particulares

Tiene un límite inferior y superior, y no sigue la distribución normal:



## Utilización de métodos estadísticos. Consideraciones



### Advertencia 3: La radiación tiene características particulares

No todos los parámetros relevantes se definen igual en los modelos meteorológicos:

Radiación: parámetro acumulado o instantáneo (normalmente acumulado).

Nubosidad: parámetro instantáneo y normalizado.

Viento: definido como el viento medio en los 10 minutos anteriores a la hora del modelo.

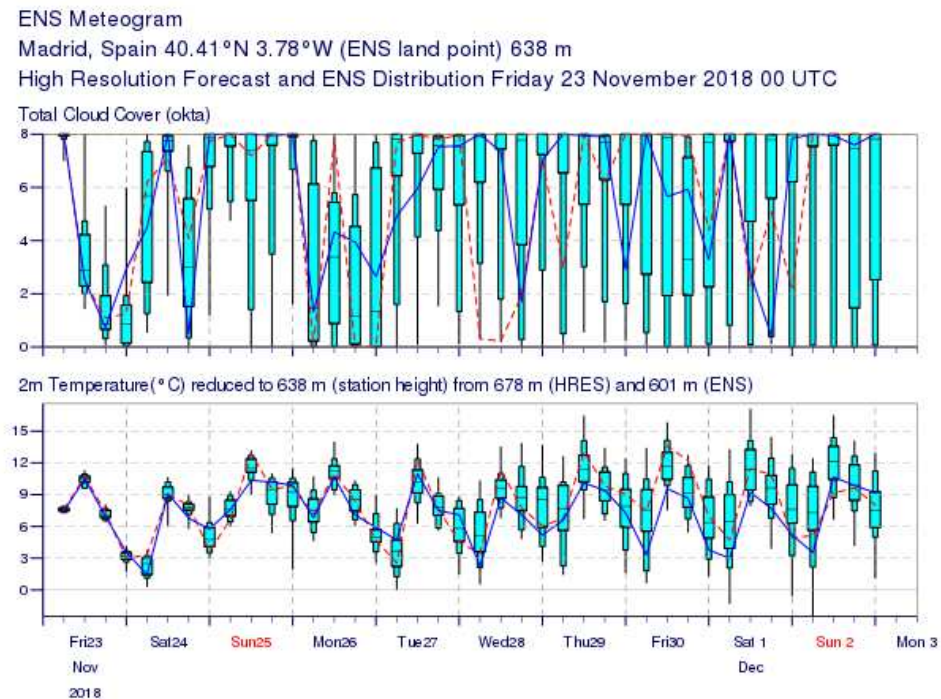
Por otra parte, no hay que olvidar que la radiación depende de la latitud.

## Utilización de métodos estadísticos. Consideraciones



### Advertencia 3: La radiación tiene características particulares

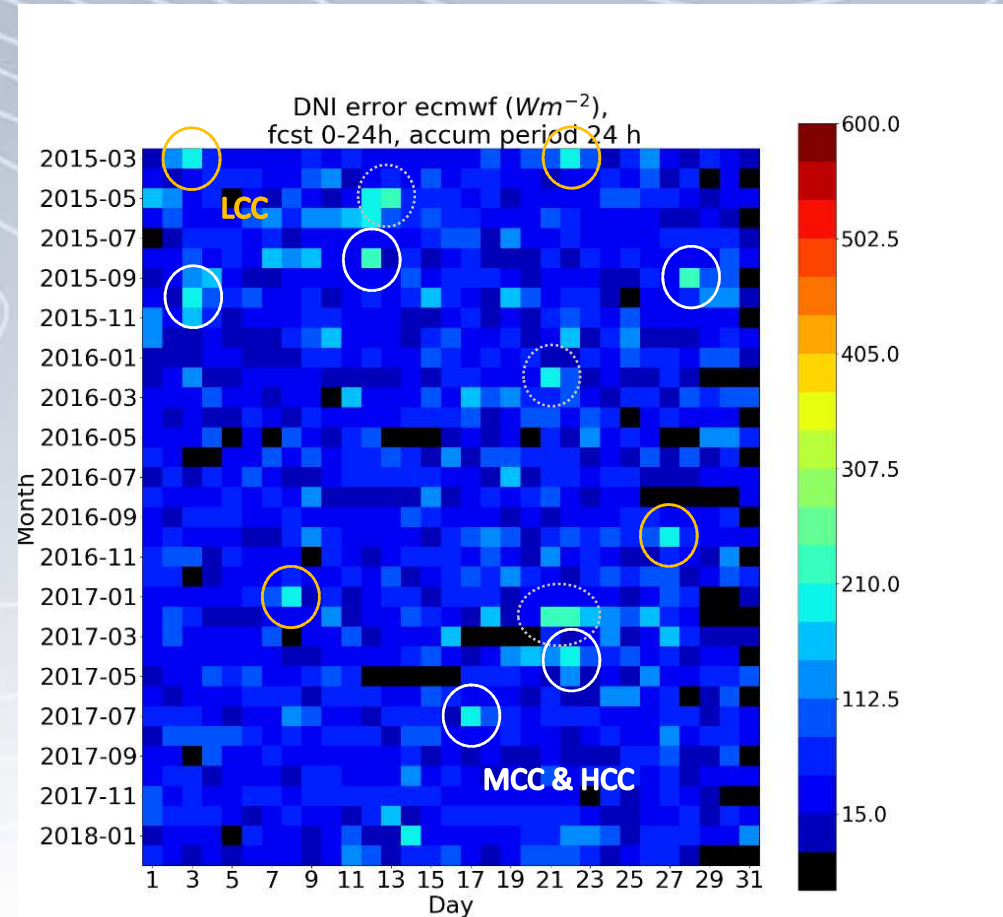
La incertidumbre al predecir la nubosidad repercute en la radiación:



## Utilización de métodos estadísticos. Consideraciones



Advertencia 4: “De donde no hay no se puede sacar”

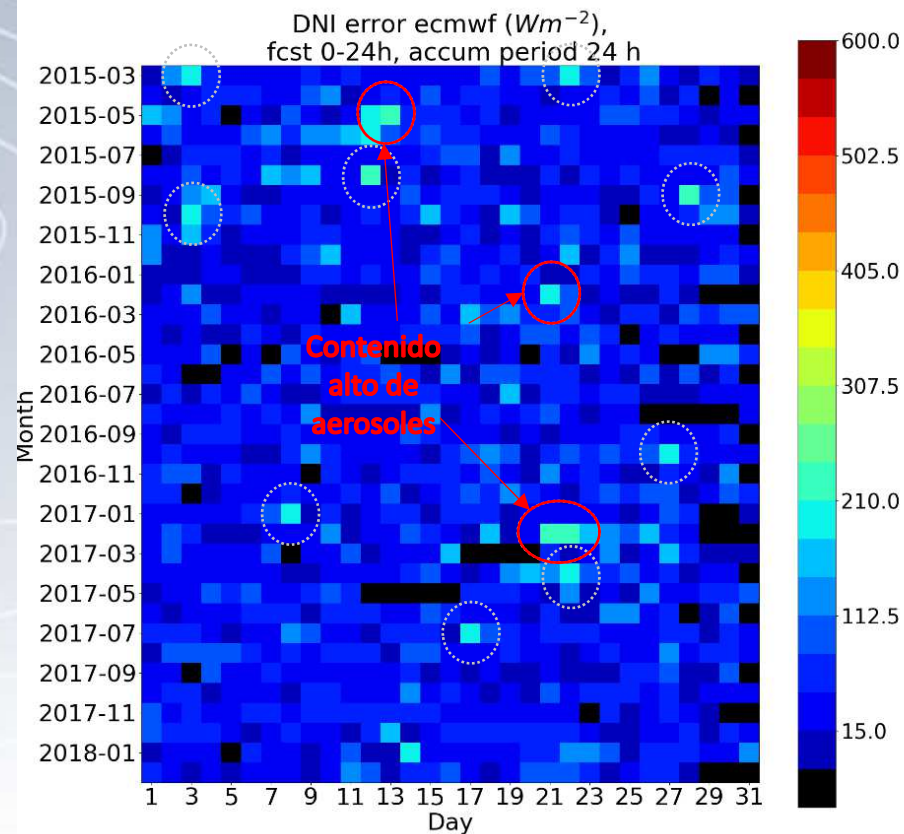
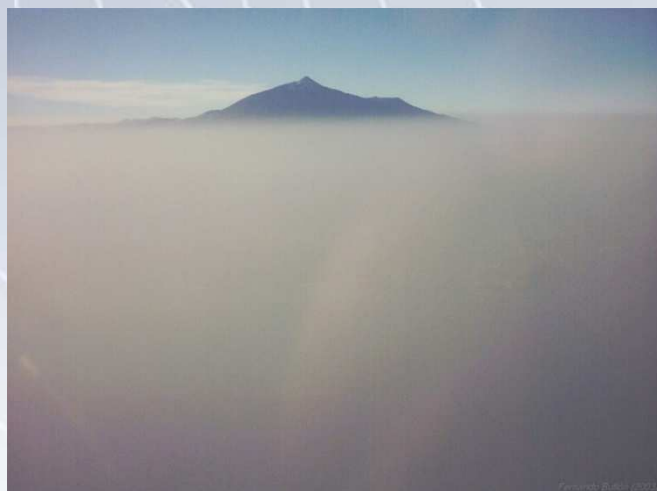




## Utilización de métodos estadísticos. Consideraciones



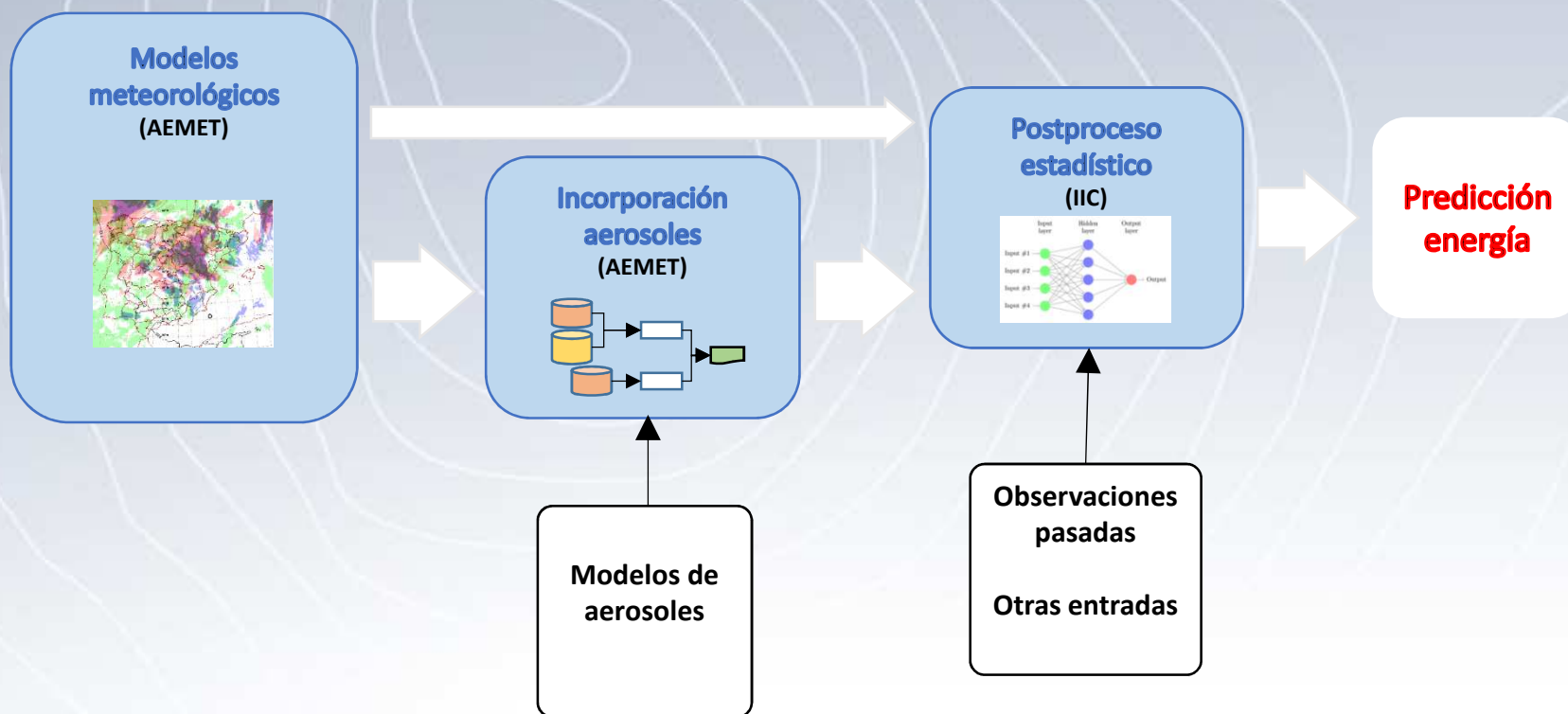
Advertencia 4: “De donde no hay no se puede sacar”



## Utilización de métodos estadísticos. Consideraciones



Advertencia 4: “De donde no hay no se puede sacar”



## Utilización de métodos estadísticos. Consideraciones



Advertencia 4: “De donde no hay no se puede sacar”

Puedo usar métodos estadísticos para tener en cuenta problemas no previstos, pero...  
¿dispondré de un periodo de entrenamiento suficientemente largo?



Cuando arreglo un problema, me aparece otro...



Radiación  
circumsolar



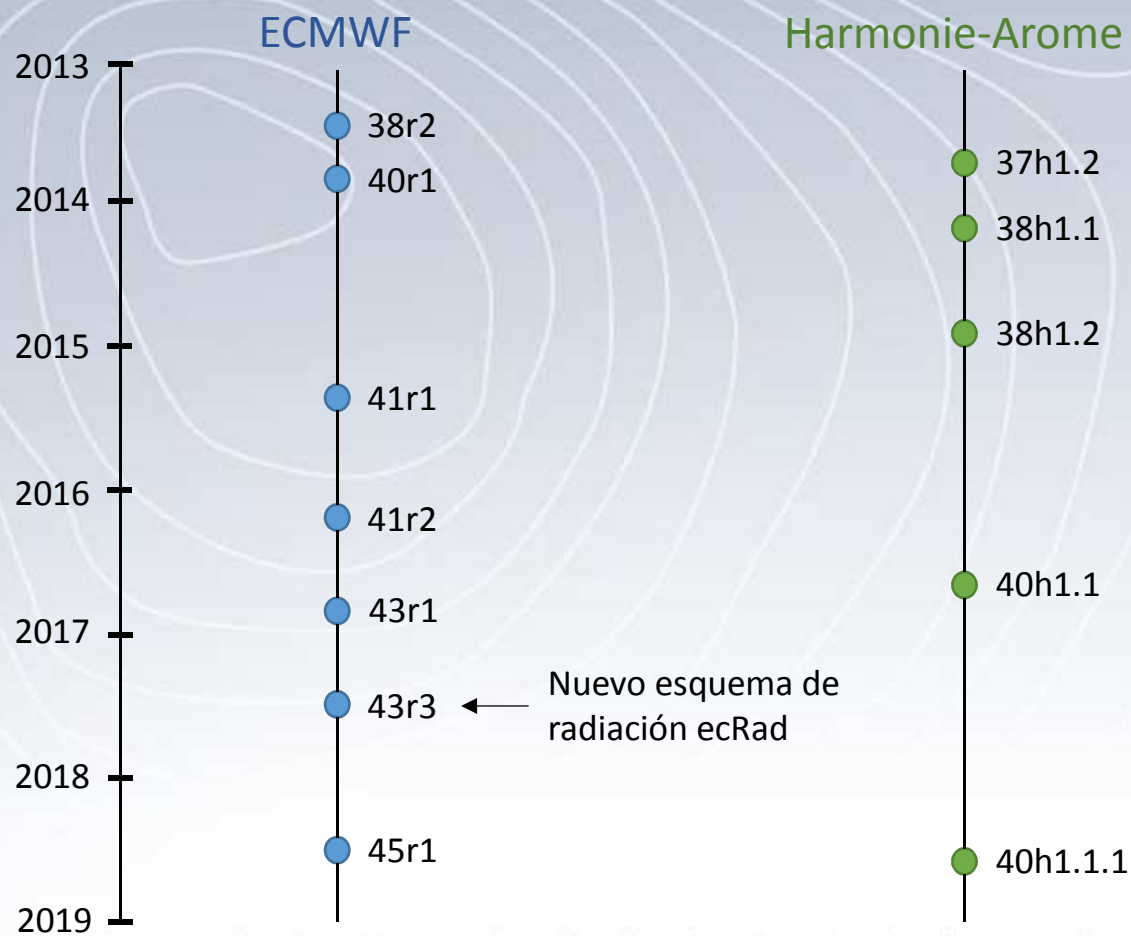
Suciedad



## Utilización de métodos estadísticos. Consideraciones



Advertencia 5: la versión de los modelos cambia cada cierto tiempo







### Advertencia 6: estratificación vs. agrupación (pooling)

#### Estratificación:

- + Permite ver el comportamiento de las predicciones para distintos regímenes.
- Disminuye la muestra, por lo que la confianza en los resultados es menor.

#### Agrupación:

- + Aumenta el tamaño de la muestra, con lo que las predicciones son más fiables.
- Si la muestra no es climatológicamente homogénea, puede enmascarar variaciones en la habilidad (skill) de las predicciones. Puede sobreestimar el skill en ocasiones (Hamill and Juras, doi: [10.1256/qj.06.25](https://doi.org/10.1256/qj.06.25))

## Utilización de métodos estadísticos. Consideraciones



### Advertencia 6: estratificación vs. agrupación (pooling)

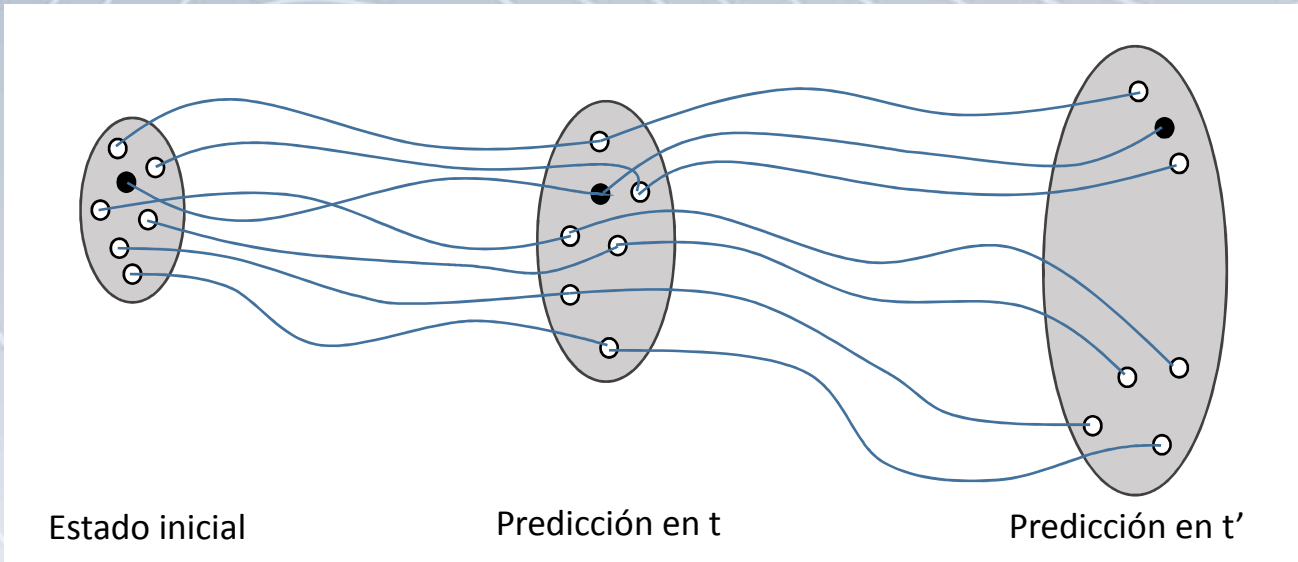
Para estudiar mejor los resultados, podemos estratificar según la hora del día, el día del año, y/o el ángulo cenital.

También podemos utilizar el clearness index, el cociente entre la radiación en el tope de la atmósfera y la que llega a la superficie:

$$KT_{f,o} = \frac{G_{f,o}}{E}$$

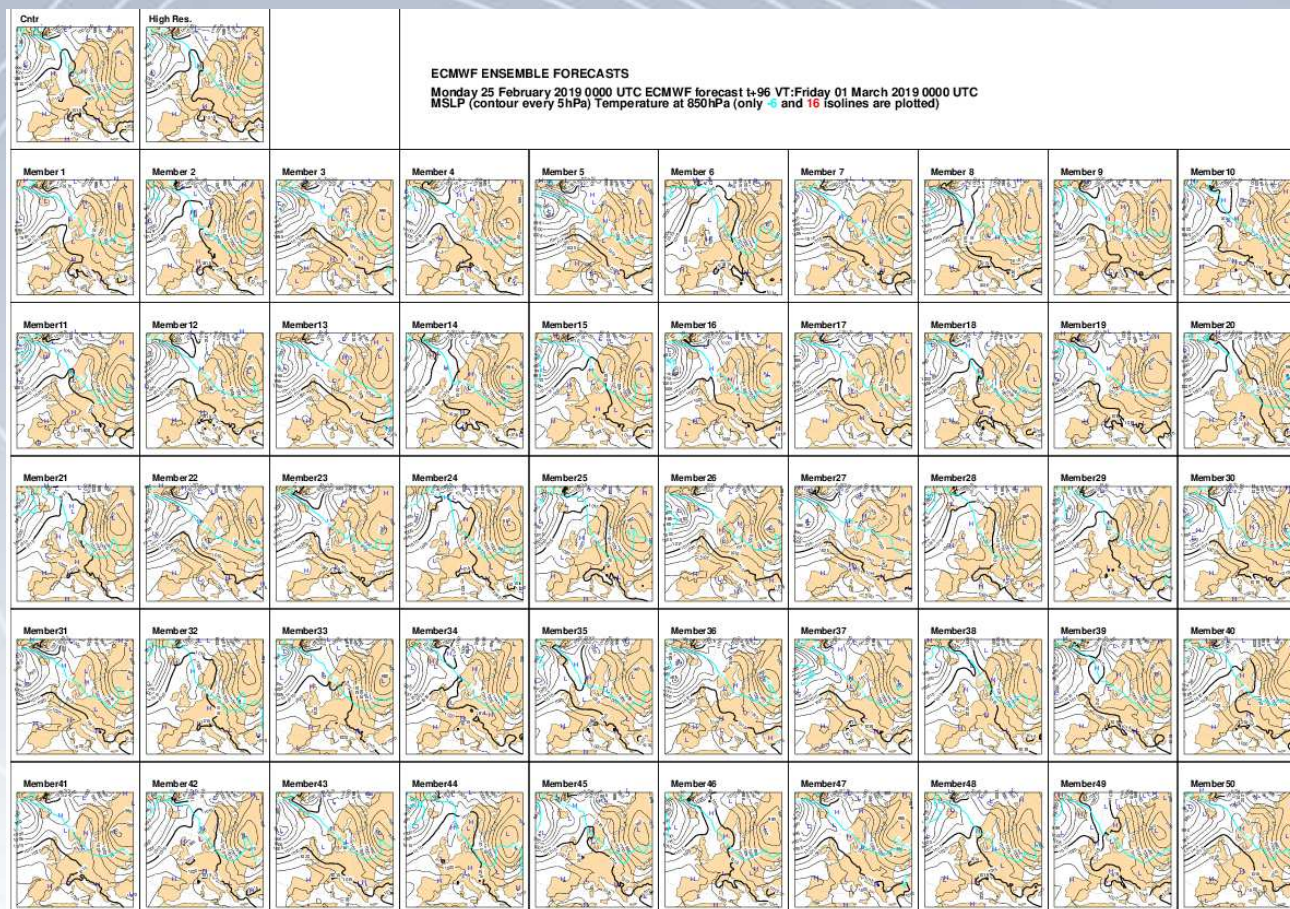
## Modelos de predicción por conjuntos (probabilistas)

# Modelos de predicción por conjuntos





## Modelos de predicción por conjuntos



## Modelos de predicción por conjuntos



“Si pronostico que hay un 50% de posibilidades de que llueva mañana, no me estoy arriesgando nada: voy a acertar parcialmente pase lo que pase”.

Supongamos que se prevé un 20% de posibilidades de que llueva mañana. ¿Cómo puedo tomar una decisión con esta incertidumbre?



## Modelos de predicción por conjuntos

### Modelos de coste-pérdida:

Supongamos que se preve un 20% de posibilidades de que llueva mañana. ¿Cómo podemos tomar una decisión con esta incertidumbre?

	El suceso ocurre	El suceso no ocurre
Se toman medidas	C	C
No se toman medidas	L	0

Si solo salimos para ir al cine



Coste (C) ↑  
Pérdida (L) ↓↓

C/L ↑

20% de lluvia es poco

Si vamos a organizar una fiesta al aire libre



Coste (C) ↔  
Pérdida (L) ↑↑

C/L ↓

20% de lluvia es mucho

## Modelos de predicción por conjuntos



$$\text{Valor relativo} = \frac{\min\left(\frac{C}{L}, \bar{o}\right) - F\frac{C}{L}(1 - \bar{o}) + H\bar{o}(1 - \frac{C}{L}) - \bar{o}}{\min\left(\frac{C}{L}, \bar{o}\right) - \bar{o}\frac{C}{L}}$$

siendo:

$\bar{o}$ : probabilidad del suceso (*depende del suceso*)

H, F: hit rate, false alarm rate (*depende del modelo*)

C/L: cociente coste/pérdida (*depende del usuario*)

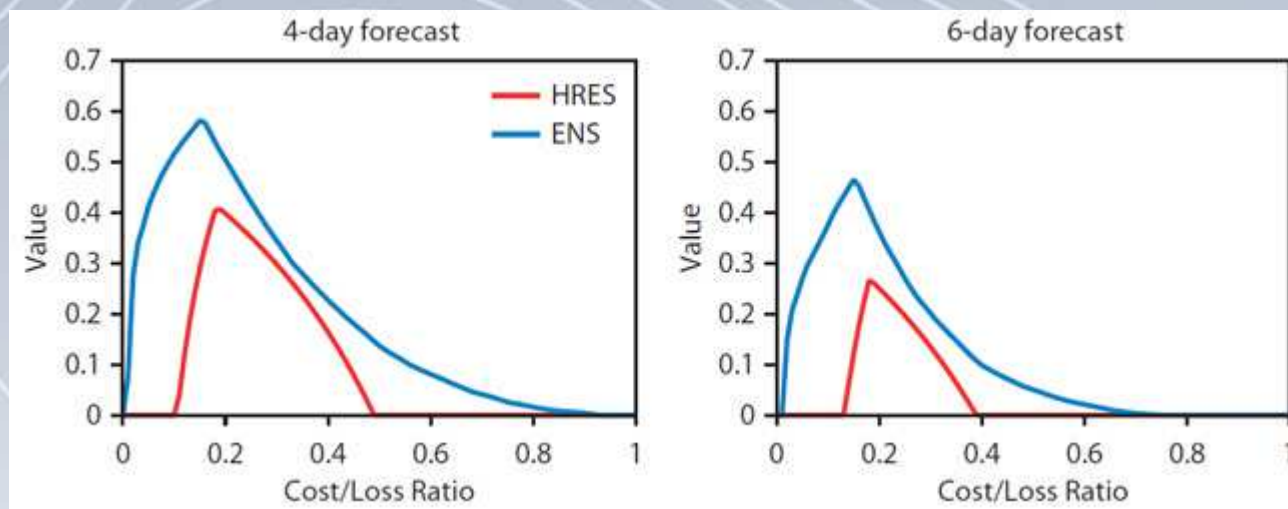
Un modelo determinista “determina” cuándo un suceso cree que va a ocurrir. Pero para un modelo probabilista nosotros podemos decidir a partir de qué probabilidad  $p$  consideramos que el suceso va a ocurrir, y calcular de ahí la H y F del modelo.

Podemos escoger la  $p$  que nos de el mayor valor de V para cada cociente C/L. En concreto se puede deducir que esto ocurre si  $p = C/L$  (para modelos **fiabiles**).



## Modelos de predicción por conjuntos

## Modelos de coste-pérdida



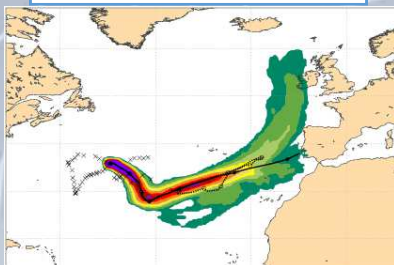
(extraído de la ECMWF Newsletter No. 141)  
© 2019 European Centre for Medium-Range Weather Forecasts (ECMWF)

Los modelos de predicción por conjuntos tienen un mayor valor económico, cuando se estima una predicción utilizando modelos coste-pérdida.

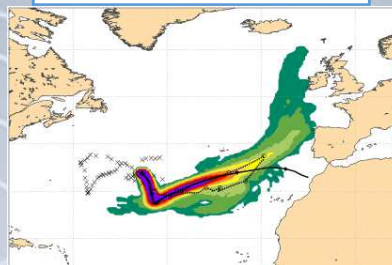
## Modelos de predicción por conjuntos

### Caso del huracán Leslie

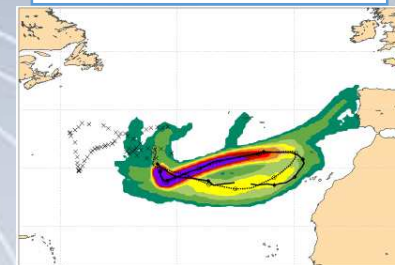
Lunes 8, 00 UTC



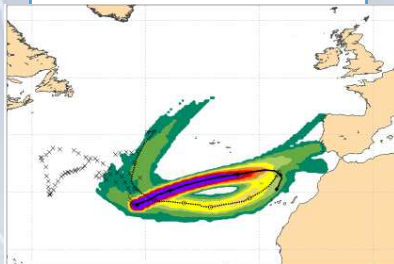
Martes 9, 00 UTC



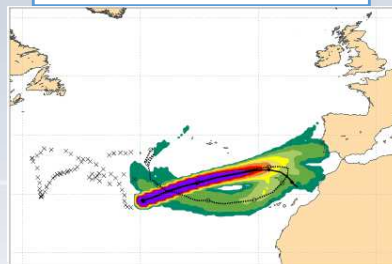
Miércoles 10, 00 UTC



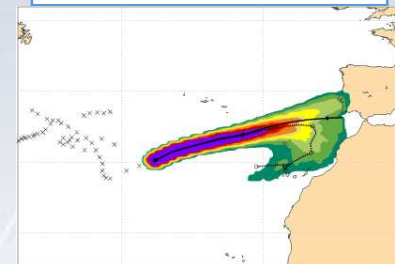
Jueves 11, 00 UTC



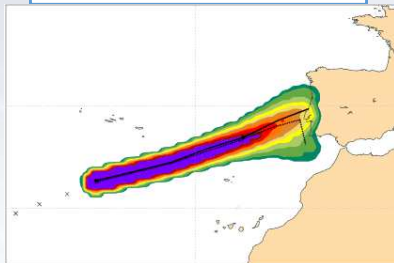
Jueves 12, 12 UTC



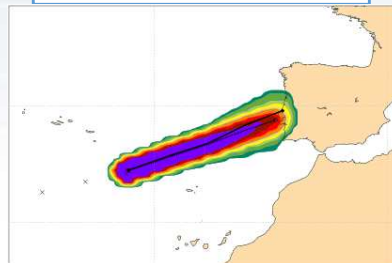
Viernes 12, 00 UTC



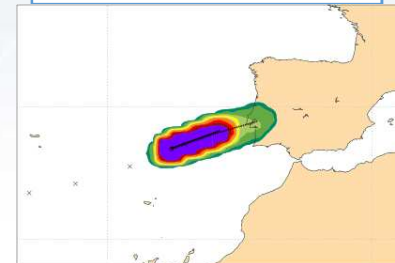
Viernes 12, 12 UTC



Sábado 13, 00 UTC



Sábado 13, 12 UTC



## Modelos de predicción por conjuntos



“Prefiero no complicarme la vida. Tomo la media de los miembros y ya está”

=> Efectivamente tomar la media de las predicciones probabilísticas es útil: da mejores resultados que usar un miembro determinado.

=> Pero así no utilizamos la dispersión del modelo, y además la media suele dar problemas en los cambios de régimen.

“No sé para qué se hacen ciertas predicciones en las que se falla casi siempre.”

=> Calidad vs. Valor:

- Calidad: las predicciones se ajustan bien a las observaciones de acuerdo a un criterio (objetivo o subjetivo).
- Valor: las predicciones ayudan al usuario a tomar mejores decisiones.

Predicciones con poca calidad pueden sin embargo tener bastante valor.

## Modelos de predicción por conjuntos



Dos propiedades importantes en un modelo de predicción por conjuntos:

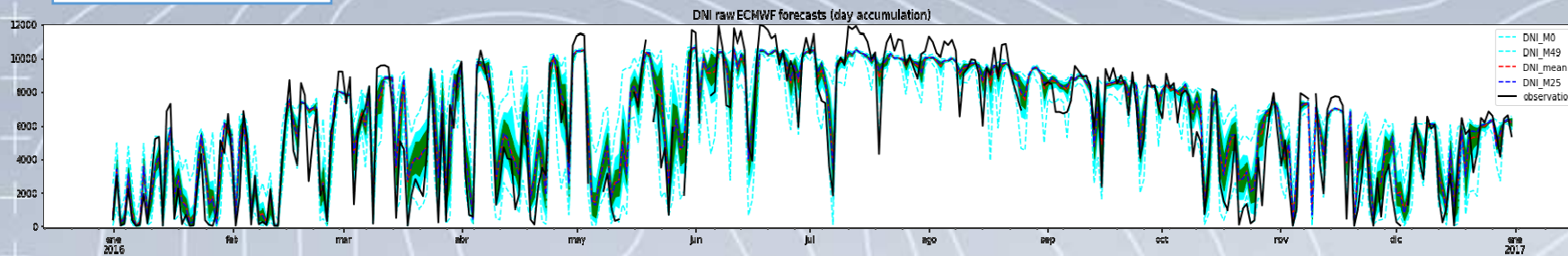
- Resolución: capacidad de la predicción de dividir los sucesos en subconjuntos con diferentes frecuencias.
- Fiabilidad (reliability): hay correspondencia entre la probabilidad prevista y la frecuencia de observación condicionada.



## Modelos de predicción por conjuntos



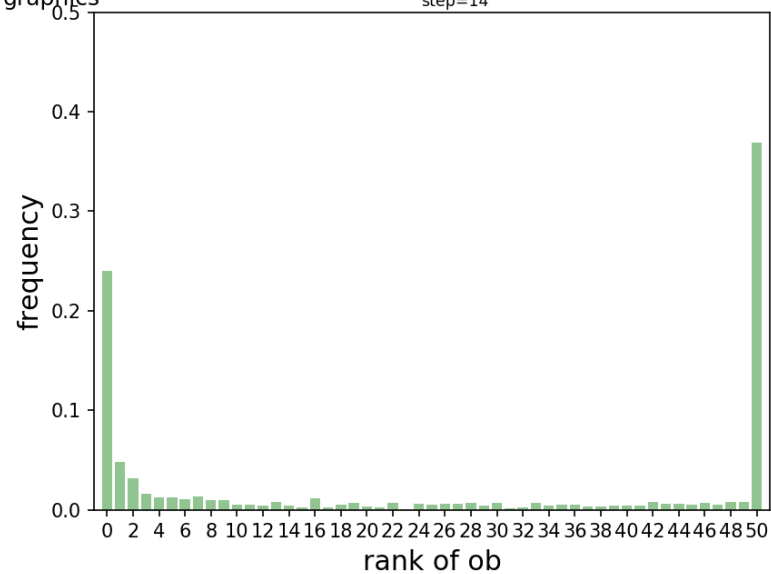
## Datos brutos



### Talagrand

Performance measures  
graphics

ens\_label=ECMWF  
parameter=DNI  
vt\_date=None  
step=14



## Modelos de predicción por conjuntos



### Regresión por cuantiles

Cada cuantil  $\tau$  se ajusta linealmente como:

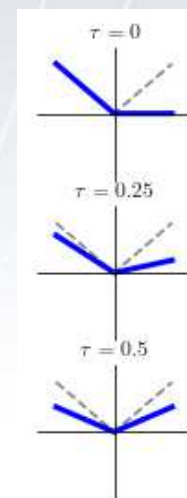
$$q_{\tau}(y|x) = \beta_0 + \beta x$$

Los coeficientes  $\beta$  (distintos para cada cuantil  $\tau$ ) se calculan de forma que se minimice:

$$\operatorname{argmin}_{\beta_0, \beta} \sum_{i=1}^n \rho_{\tau}(y_i - \beta_0 - \beta x_i)$$

siendo  $\rho$  la check function:

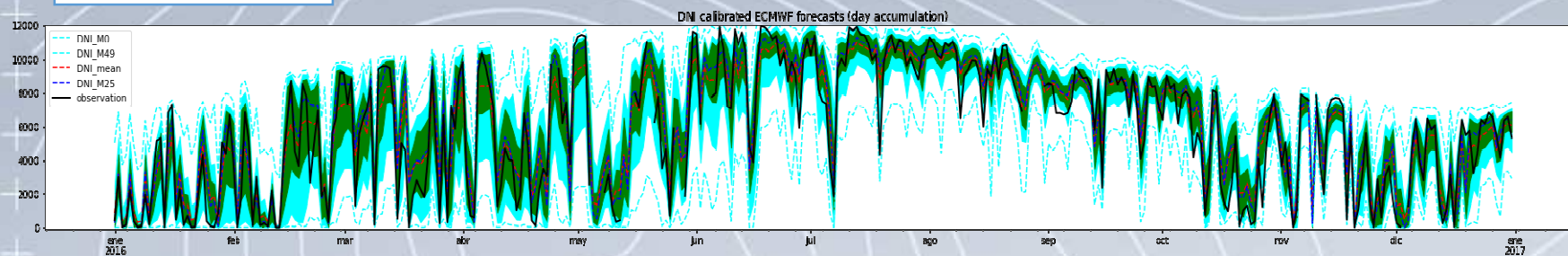
$$\rho_{\tau}(u) = u[\tau - I(u < 0)] = \begin{cases} \tau u & \text{si } u \geq 0 \\ (\tau - 1)u & \text{si } u < 0 \end{cases}$$



## Modelos de predicción por conjuntos



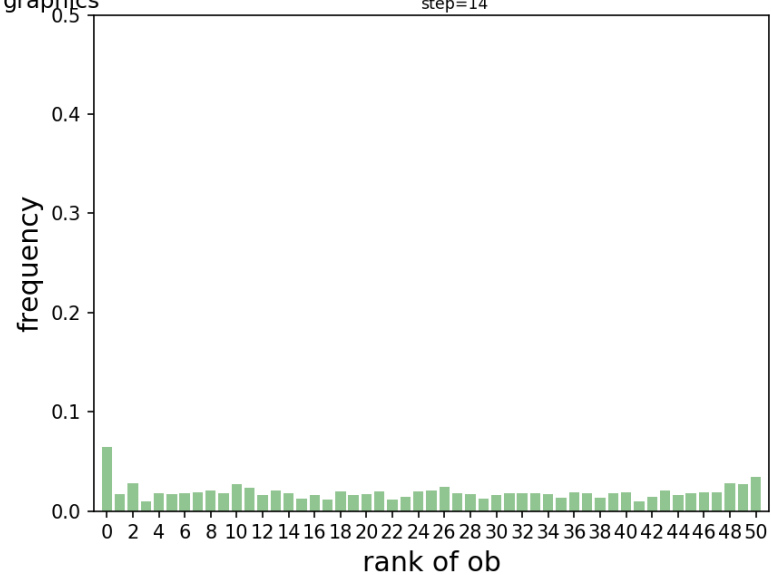
### Datos calibrados



### Talagrand

Performance measures  
graphics

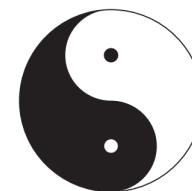
ens\_label=ECMWF  
parameter=DNI  
vt\_date=None  
step=14



## Conclusiones



- Los modelos meteorológicos son una herramienta básica para predecir la radiación solar (y por ello la energía que pueden generar las centrales solares) en el corto y medio plazo.
- Los métodos estadísticos (especialmente los más modernos) permiten corregir muchos de los errores cometidos por los modelos. Es conveniente tener en cuenta el problema físico real para evitar trampas de origen estadístico.
- Los modelos probabilísticos dan mejores resultados por su propia naturaleza, aunque son más complicados de utilizar. Esto también puede convertirse en una oportunidad.
- El yin y el yang de la meteorología y la computación:  
“A statistically literate person with insight into the physical problem may be more successful than a statistician at devising a forecast equation” (Wilks)





Más información



- Santos, C.: Física del caos en la predicción meteorológica (doi: [10.31978/014-18-009-X](https://doi.org/10.31978/014-18-009-X))
- ECMWF Forecast User Guide (doi: [10.21957/m1cs7h](https://doi.org/10.21957/m1cs7h))
- Richardson, D.S.: Skill and relative economic value of the ECMWF ensemble prediction system (doi: [10.1002/qj.49712656313](https://doi.org/10.1002/qj.49712656313))
- Wilks, D.: Statistical Methods in the Atmospheric Sciences (doi: [10.1016/B978-0-12-385022-5.00007-5](https://doi.org/10.1016/B978-0-12-385022-5.00007-5))