

Big data, clima y meteorología

Big Data

Características del Big Data

**Volumen
y**

Variedad

Tamaño de los datos
y diversidad de tipos

Velocidad

Velocidad a la que se
generan/procesan los datos

Valor



Modelos.



Redes de observación.



3/15



Velocidad

Restricción temporal
para obtener los
resultados

Medio y largo plazo

Ejecutar modelos cada día mas complejos y a la vez aumentar el número de simulaciones.

Por ejemplo: PALEOSTRAT: proyecto para realizar un conjunto de simulaciones de 1.000 años

Corto plazo => Predicción meteorológica. Ejecutar el modelo de predicción en un intervalo de tiempo determinado.





Valor



Mejor comprensión del sistema climático que rige
nuestra atmósfera

Conseguimos información sobre como
está cambiando el clima

cómo cambió en el pasado y
cómo cambiará en el futuro

Entendemos mejor los eventos extremos
y sus impactos

Así como el papel de la meteorología
en la calidad del aire.

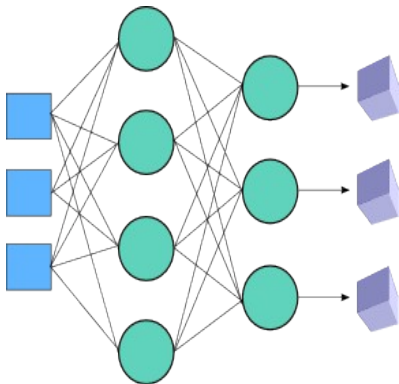


El desarrollo de técnicas de **Machine-Learning** nos ofrece nuevas y robustas herramientas de extracción de conocimiento.

Optimización evolutiva
(HS, CRO)

Clasificación automática
(MLP, SVM-C)

Clustering supervisado
y no supervisado



Técnicas de regresión
(ELM, GMDH, ϵ -SVM)



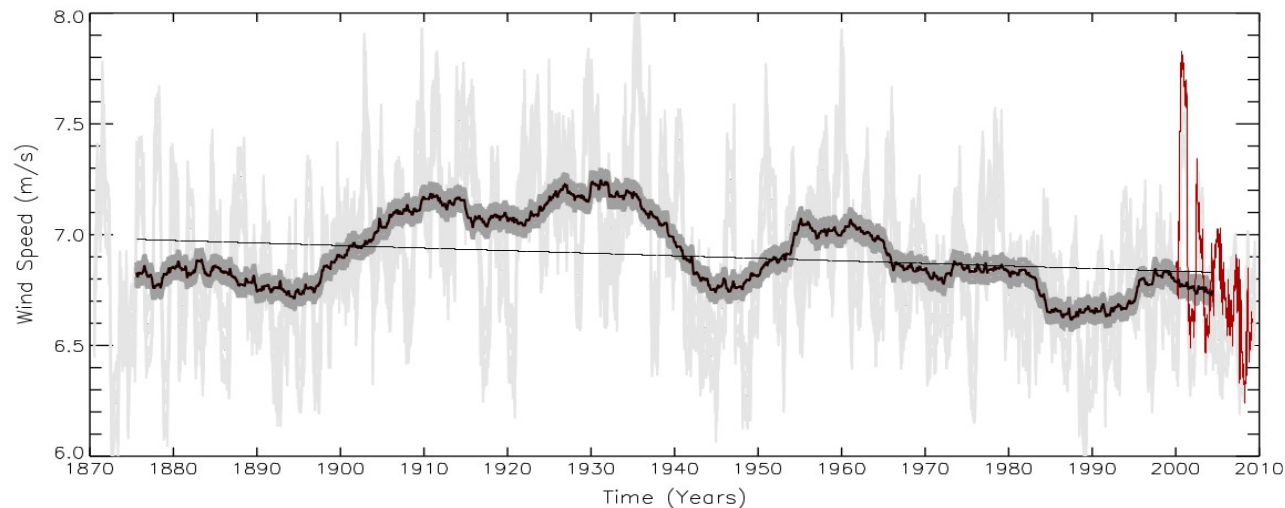
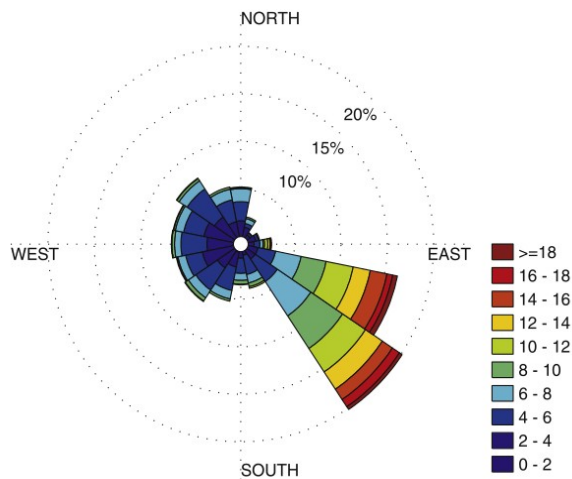


Valor (El recurso eólico)



Caracterización del recurso eólico a largo plazo.

N. Kirchner-Bossi et al. / Applied Energy 105 (2013) 30–46



Extraction of synoptic pressure patterns for long-term wind speed estimation in wind farms using evolutionary computing

L. Carro-Calvo et al.

Energy. 2011

Wind speed reconstruction from synoptic pressure patterns using an evolutionary algorithm

L. Carro-Calvo et al.

Applied Energy 2012

Multi-decadal variability in a centennial reconstruction of daily wind

N. Kirchner-Bossi, et al.

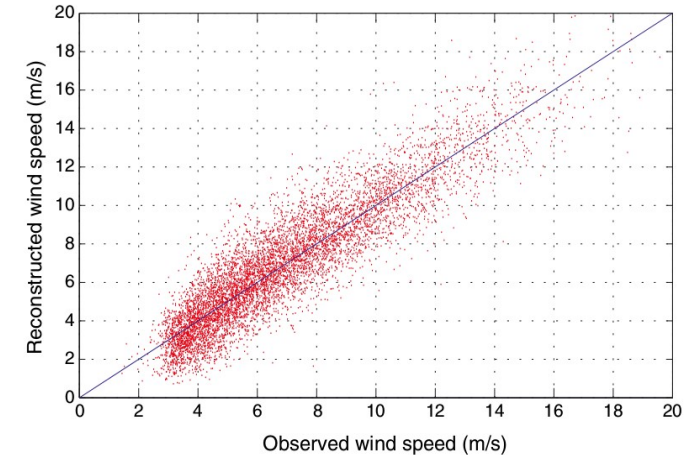
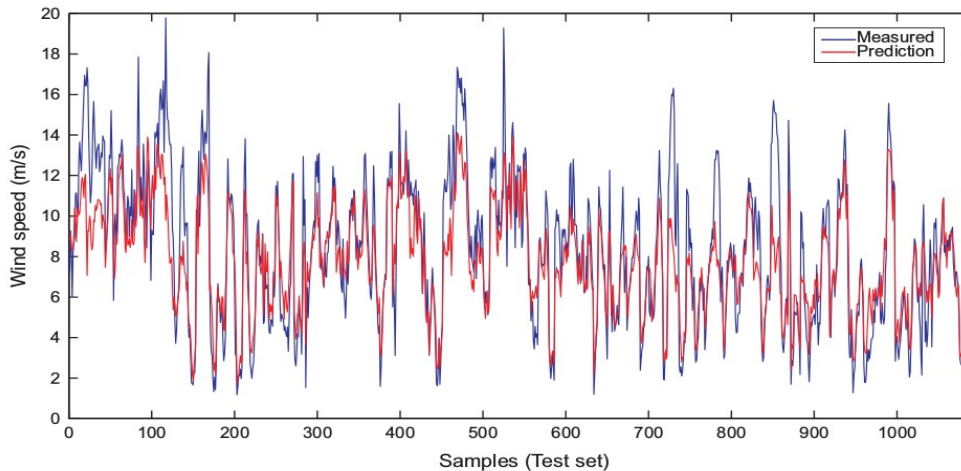
Applied Energy 2013



Valor (El recurso eólico)



Estimación del recurso eólico con alta resolución temporal.



Very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms.

Saavedra-Moreno, et al.

Journal of Wind Engineering and Industrial Aerodynamics 2013

Surface wind speed reconstruction from synoptic pressure fields: machine learning versus weather regimes classification techniques

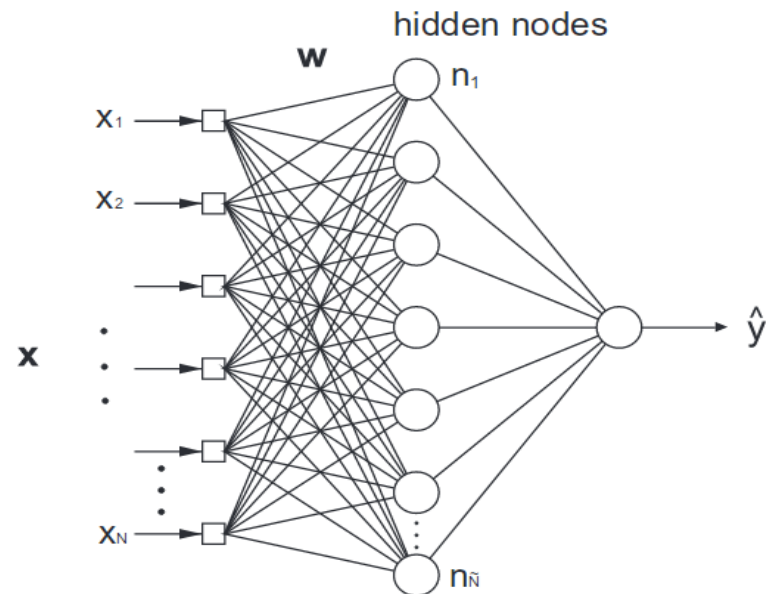
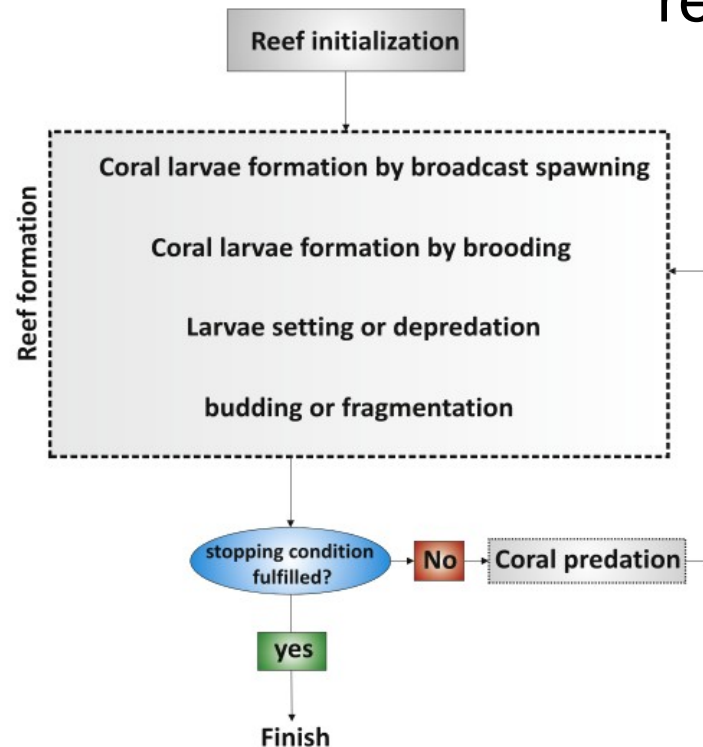
B. Saavedra-Moreno, et al.

Wind Energy 2014

Valor (El recurso eólico)



Estimación del recurso eólico con alta resolución temporal.



Outline of the extreme learning machine structure.

Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization – Extreme learning machine approach.

S. Salcedo-Sanz, et al.

Energy Conversion and Management 2014

Valor (PALEOSTRAT)

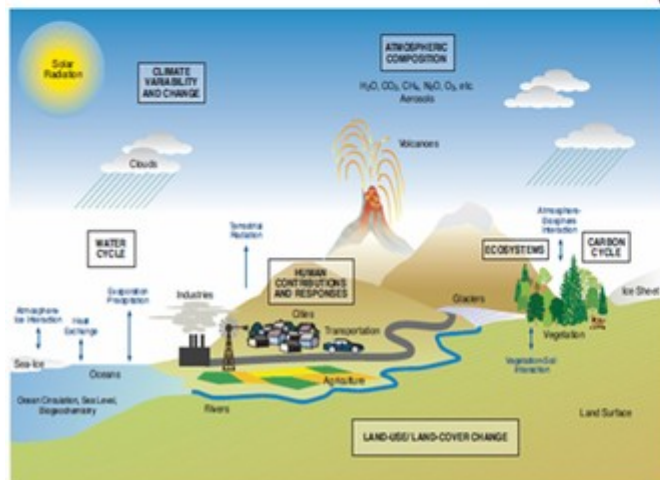
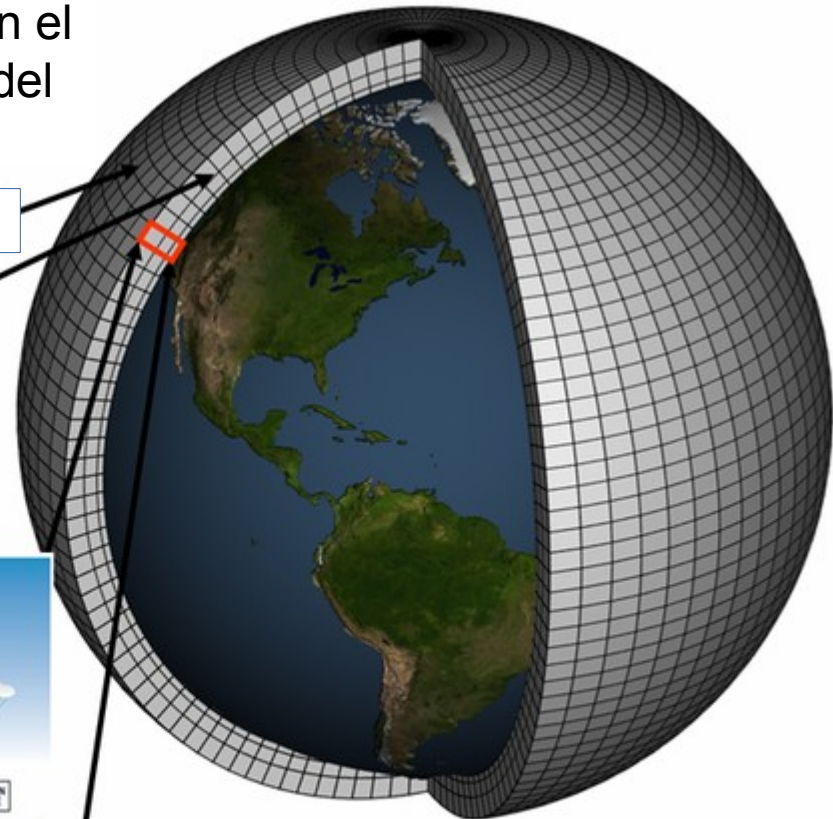


PALEOSTRAT investigará el impacto de la estratosfera en el clima usando simulaciones del último milenio.

Resolución espacial $1.9^{\circ} \times 2.5^{\circ}$

66 niveles verticales (0-140 km)

Pasos de simulación de 30 minutos a lo largo de 1000 años





Valor (PALEOSTRAT)



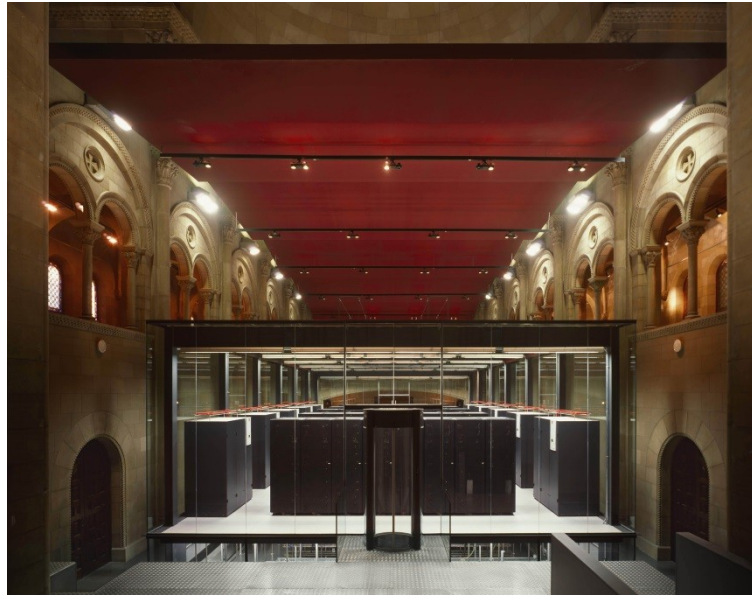
WACCM / CAM	1	2	3	4	5
1	5,5	7,1	8,6	10,1	11,6
2	9,6	11,1	12,6	14,1	15,7
3	13,6	15,1	16,6	18,2	19,7
4	17,6	19,1	20,6	22,2	23,7
5	21,6	23,1	24,7	26,2	27,7

WACCM / CAM	1	2	3	4	5
1	120	180	240	300	360
2	180	240	300	360	420
3	240	300	360	420	480
4	300	360	420	480	540
5	360	420	480	540	600

Meses de CPU usando 1024 núcleos y capacidad de almacenamiento necesario en función del número de simulaciones WACCM y CAM.



Valor (Paleostrat)



MareNostrum

Ranking Top500 List / Green List

Max. / Peak Performance

Architecture Computer

Processor

Nodes / Processors

Memory

Disk / Archive Storage

Interconnection

OS

BSC (Barcelona Supercomputing Center)

91 / 185

0.9 / 1.1 PFLOPS

Cluster IBM iDataPlex DX360M4

Intel SandyBridge Xeon E5-2670 8C 2.6GHz

3,056 / 48,896

100.8 TB

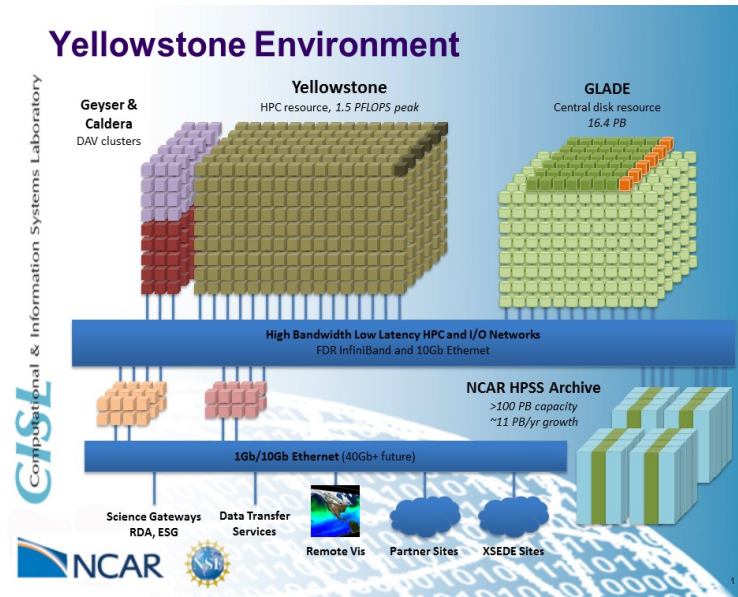
2,000 / 8,500 TB

Infiniband FDR10, GbE

SUSE Linux Es



Valor (Paleostrat)



Yellowstone

Ranking Top500 List / Green List

Max. / Peak Performance

Architecture Computer

Processor

Nodes / Processors

Memory

Disk / Archive Storage

Interconnection

OS

NCAR (National Center for Atmospheric Research)

57 / 198

1.3 / 1.5 PFLOPS

Cluster IBM iDataPlex DX360M4

Intel SandyBridge Xeon E5-2670 8C 2.6GHz

4,536 / 72,576

144.6 TB

16,000 TB

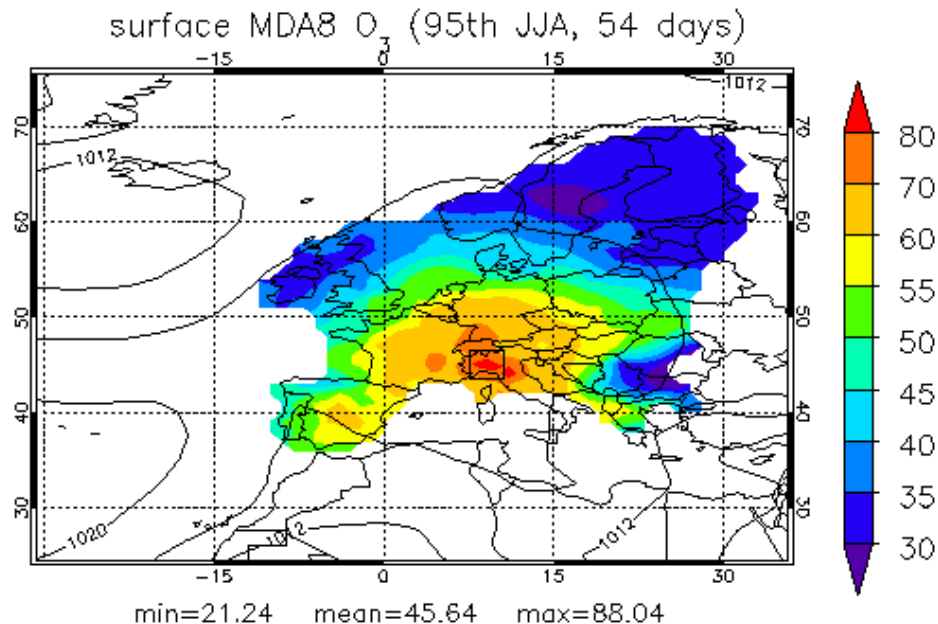
Infiniband FDR10, GbE

Red Hat Linux



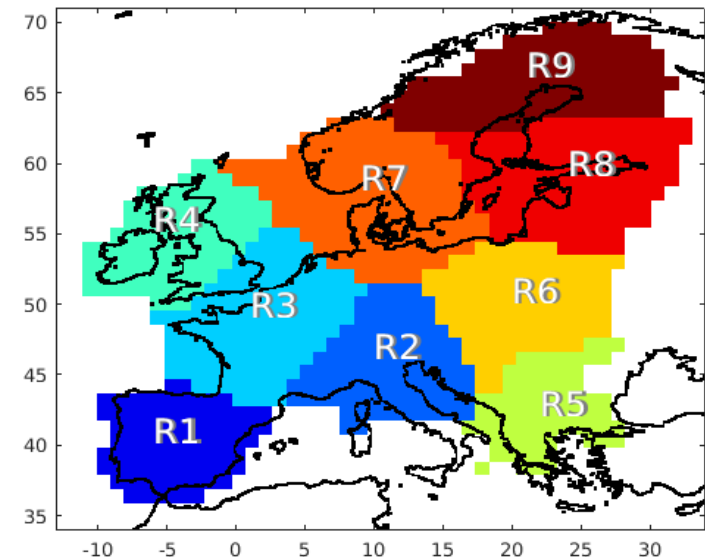
Valor (Calidad del aire)

Estudio de los factores meteorológicos en la calidad del aire



Encontramos que patrones climáticos tienen un efecto sobre las concentraciones de ozono en Europa.

Los algoritmos de clustering nos permiten obtener regionalizaciones definidas por los diferentes comportamiento de los contaminantes entre las regiones.





UNIVERSIDAD
COMPLUTENSE
MADRID



Stream-ucm.es

Leopoldo Carro Calvo Leocarro@ucm.es