



**UNIVERSIDAD
DE LA RIOJA**

Seminario Mirian Andrés:

Usos de la **Inteligencia Artificial** en
meteorología y climatología

Daniel San Martín daniel@predictia.es

19 de enero de 2024

Índice

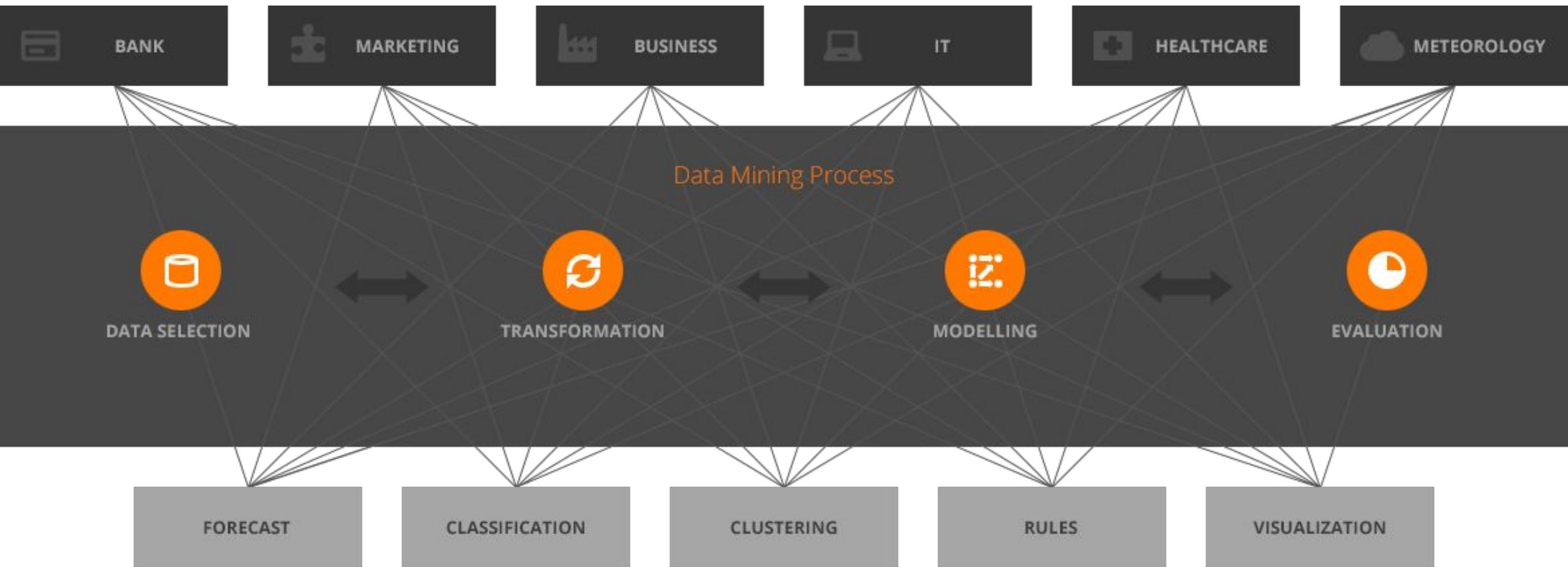
- **Predictia**
- **Modelización numérica** de la atmósfera (NWP)
- IA: aplicaciones en **meteo**
- IA: aplicaciones en **clima**



Origen

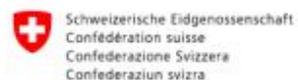
- Nace en 2008 como una **spin-off** de la **Universidad de Cantabria**
- Desarrollo de **software** a medida para **modelización** de datos
- Foco en **meteorología y clima**
- Equipo de 16 **técnicos y científicos** de **datos**





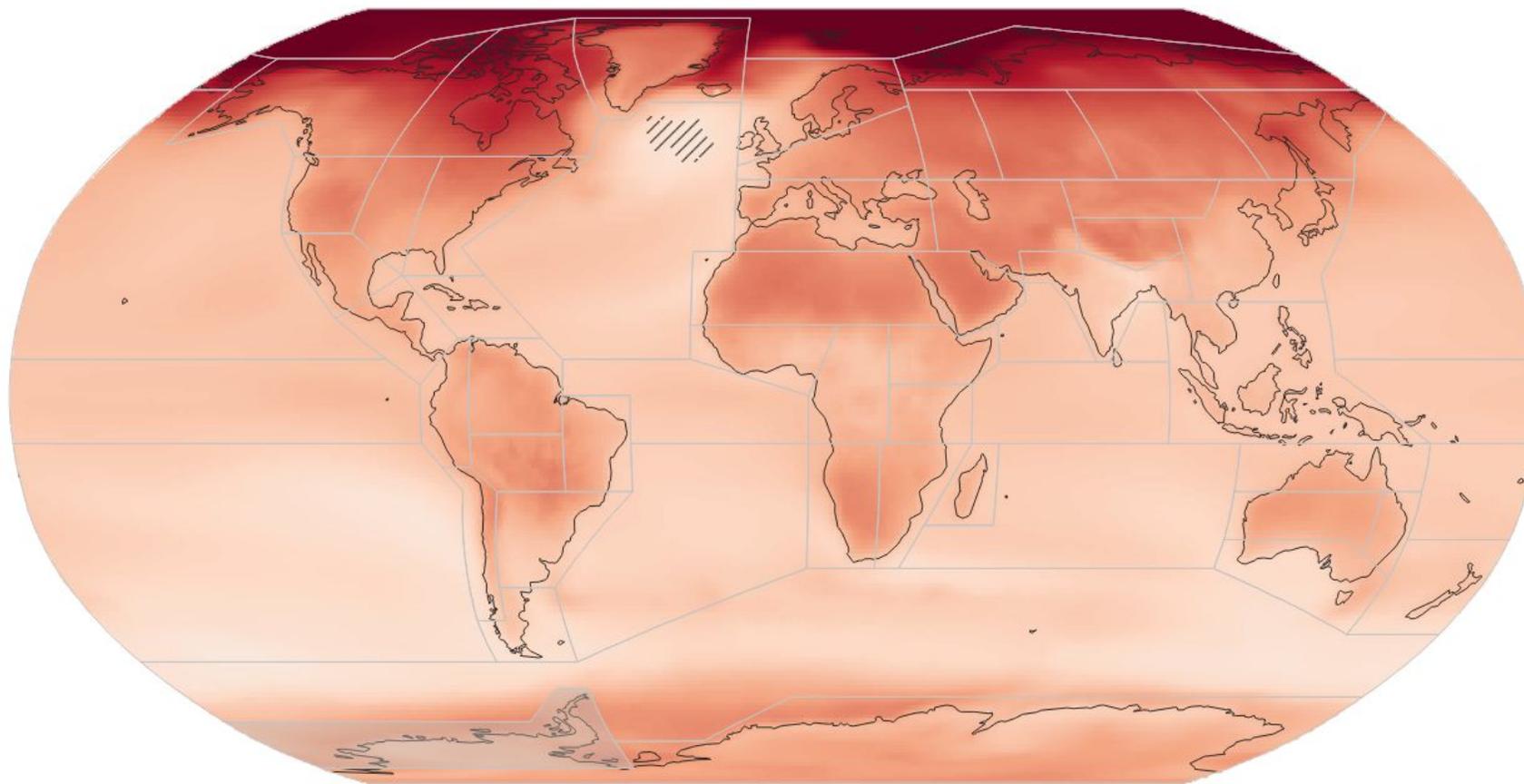
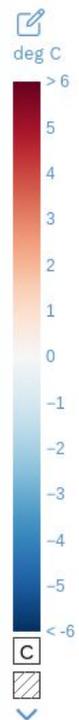


B/S/H/



VARIABLE QUANTITY & SCENARIO SEASON

Region Set:
WGI reference-re...
Uncertainty:
Simple



CMIP6 - Mean temperature (T) Change deg C - Warming 2°C SSP5-8.5 (rel. to 1850-1900) - Annual (34 models)

Copernicus Climate Data Store



Registered users

201,580

Running users

312

Queued users

88

Running requests

355

Queued requests

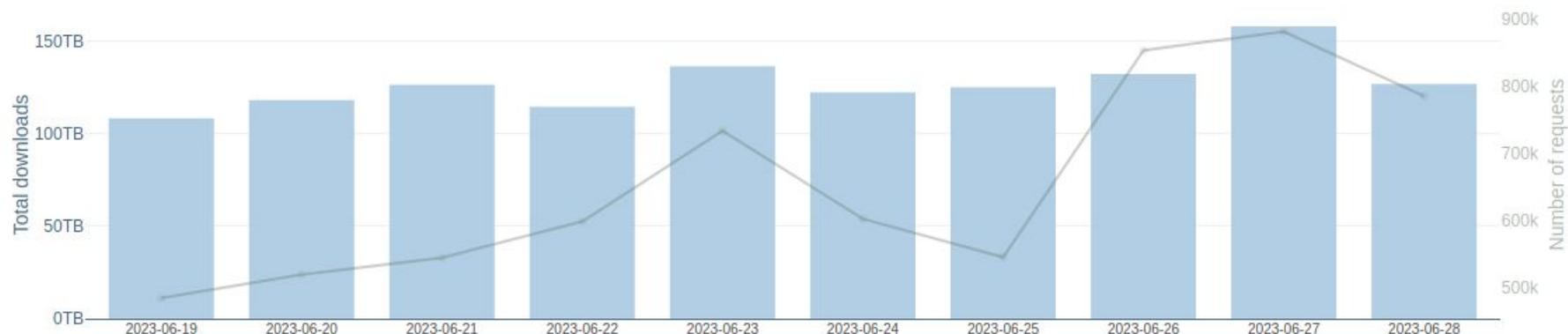
7,649

At

14:43

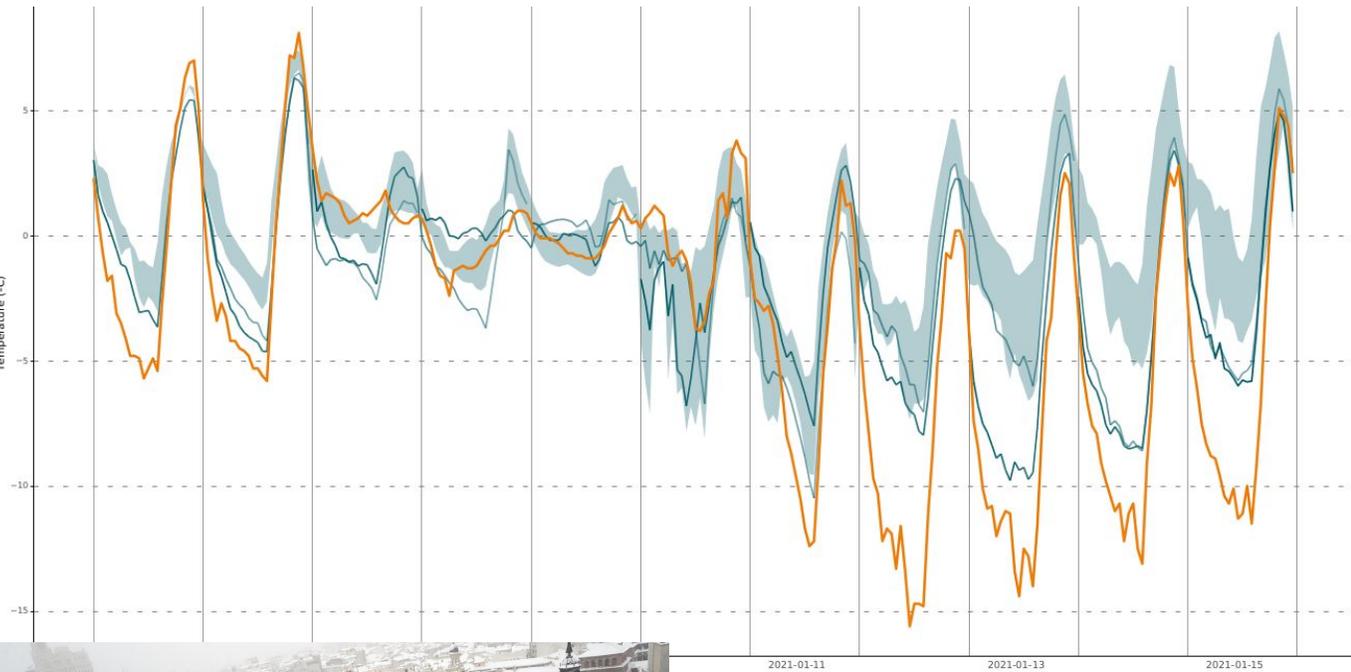
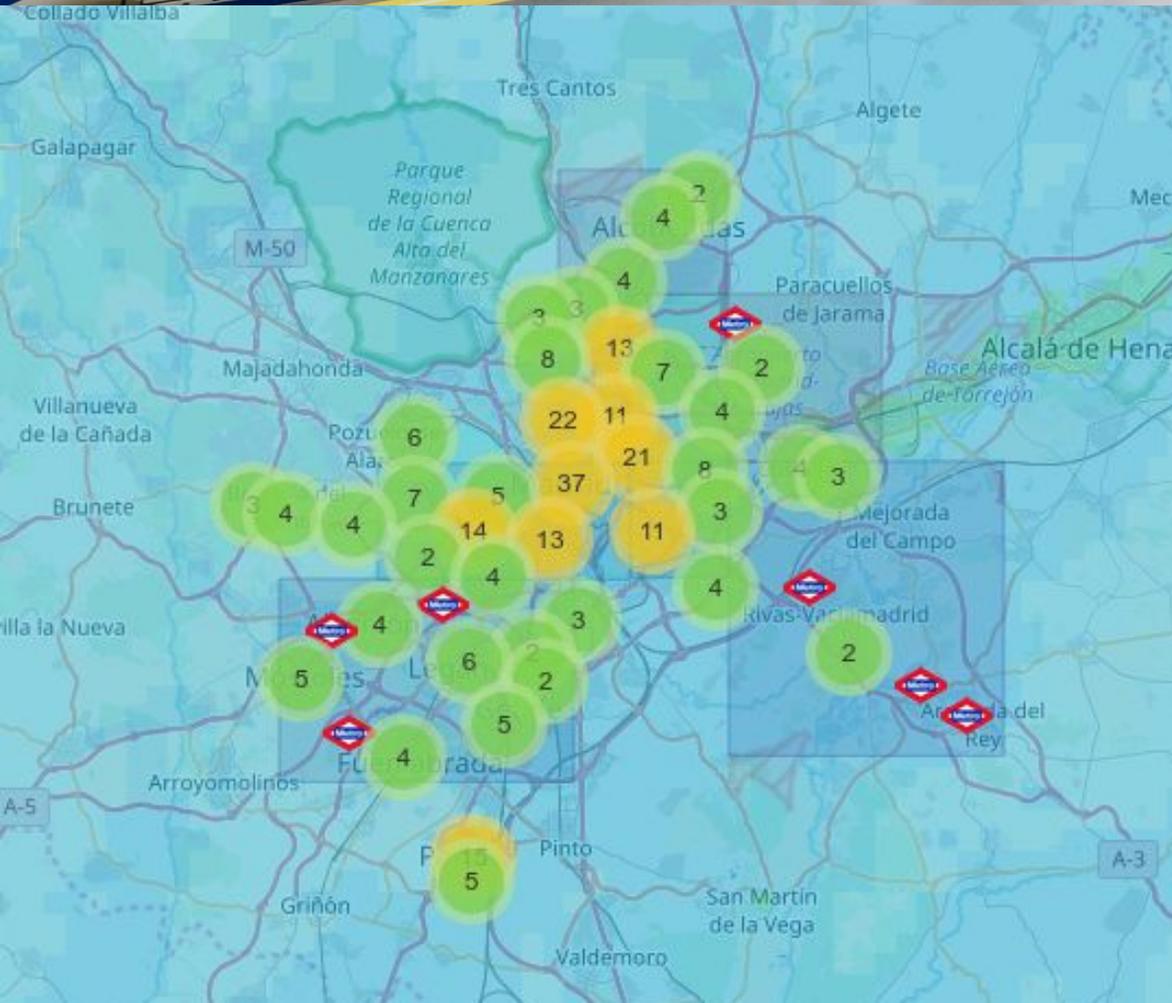
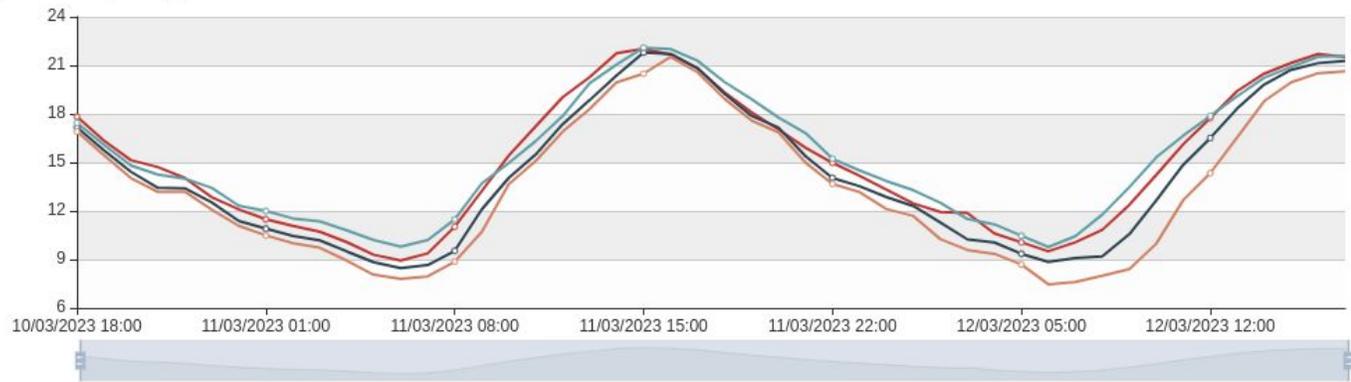
UTC | 29/06/2023

Daily downloads in Terabytes



Requests completed

Last 1h	14,779	5.6TB
Last 2h	31,559	10.2TB
Last 3h	66,583	16.2TB
Last 6h	174,742	34.8TB
Last 12h	271,468	65.3TB
Last 24h	474,722	120.4TB



METEOROLOGÍA >

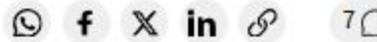
La inteligencia artificial de Google predice el tiempo que hará en el planeta en solo un minuto

El programa supera al principal sistema de predicción en más del 90% de las variables meteorológicas



MIGUEL ÁNGEL CRIADO

14 NOV 2023 - 16:00 CET



METEOROLOGÍA >

Así es la carrera por conseguir predicciones meteorológicas fiables a largo plazo

La comunidad científica se vuelca en conocer el tiempo para prevenir fenómenos adversos que cuestan cada año más de 131.000 millones de euros



<https://elpais.com/ciencia/2024-01-17/asi-es-la-carrera-por-conseguir-predicciones-meteorologicas-fiables-a-largo-plazo.html>



Predicción numérica

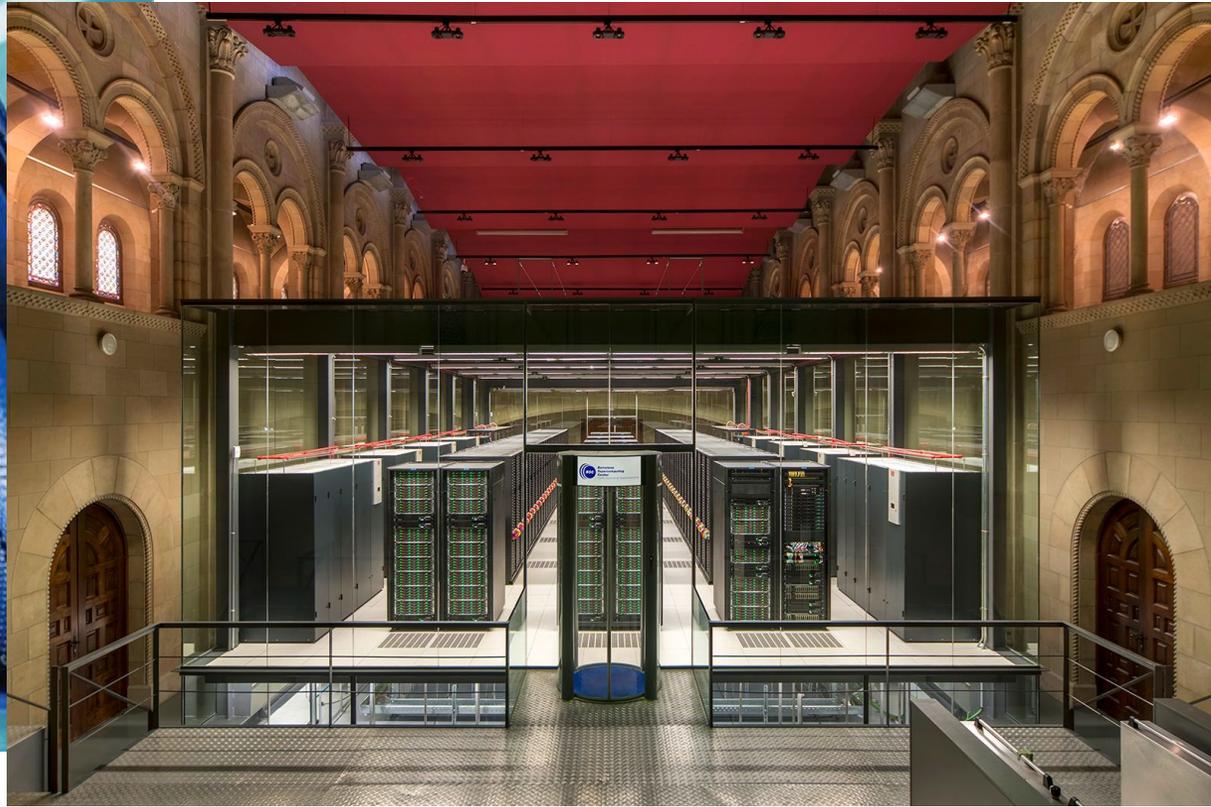


ENIAC 1950s





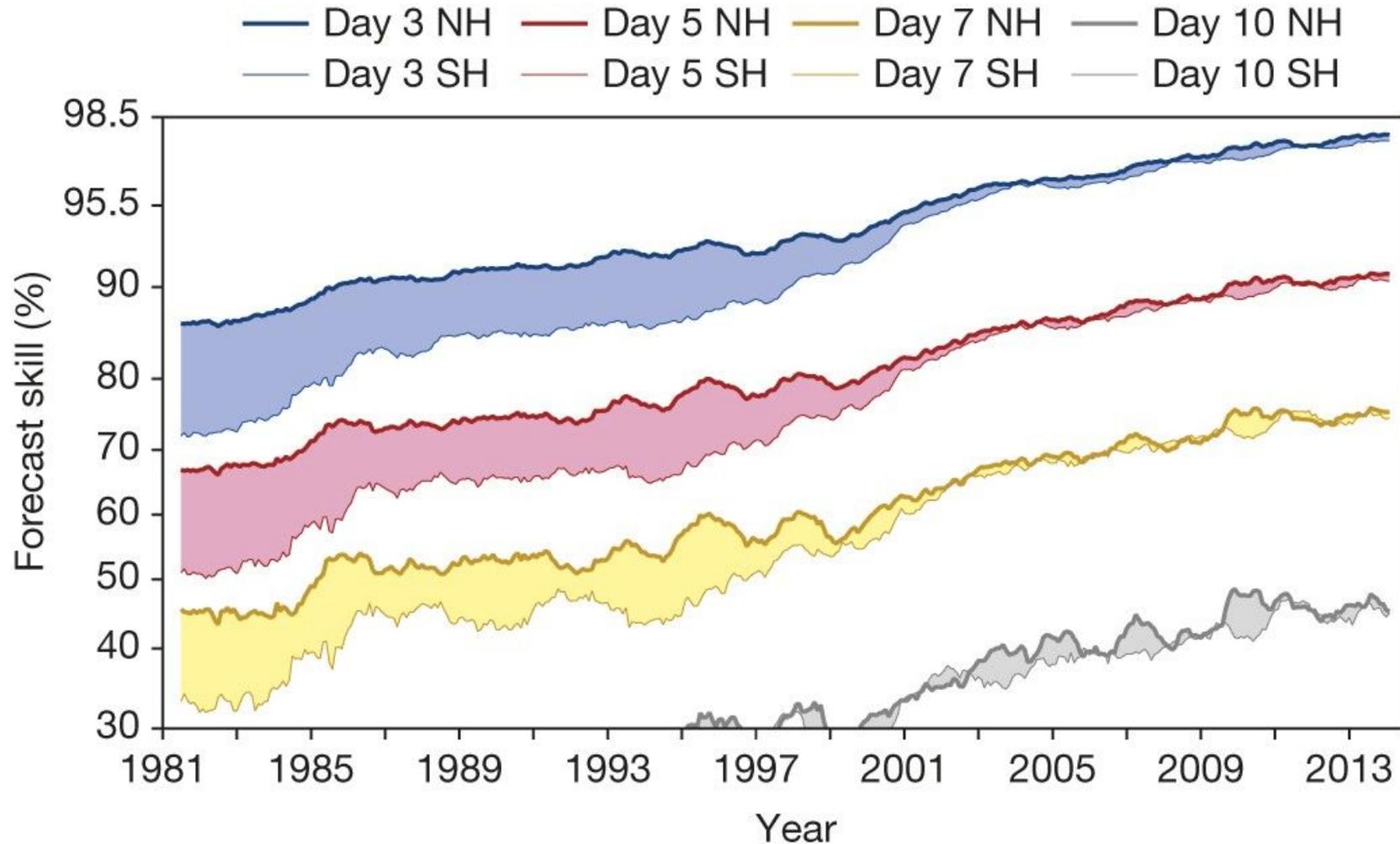
2017 - Cray XC40 supercomputer. MetOffice
460,000 cores , pico de 16 petaflops
2 petabytes de memoria



MareNostrum 5. BSC. (top 20)
Pico de 314 petaflops



The quiet revolution of NWP



Bauer, P., Thorpe, A., Brunet, G.,
2015. The quiet revolution of
numerical weather prediction.
Nature 525, 47–55.
doi:10.1038/nature14956



Dinámica de la atmósfera

El aire es un **fluido** compresible, por lo tanto puede modelarse con las ecuaciones de la mecánica de fluidos:

- Conservación de momento lineal (3 ecuaciones)
- Conservación de la energía (+1)
- Conservación de masa (+1)
- Estado de los gases (+1)
- Conservación del vapor de agua (+1)

$$\begin{aligned}\frac{d\vec{v}}{dt} &= -\frac{1}{\rho}\vec{\nabla}p - \vec{\nabla}\phi + \vec{F} - 2\vec{\Omega} \times \vec{v} \\ \frac{\partial\rho}{\partial t} &= -\vec{\nabla}\cdot(\rho\vec{v}) \\ p &= \rho RT \\ Q &= C_p \frac{dT}{dt} - \frac{1}{\rho} \frac{dp}{dt} \\ \frac{\partial\rho q}{\partial t} &= -\vec{\nabla}\cdot(\rho q\vec{v}) + \rho(E - C)\end{aligned}$$

7 ecuaciones y 7 incógnitas: (u, v, w, q, ρ, P, T)

Sistema de **ecuaciones** en **derivadas parciales** no lineal. Sin solución exacta.



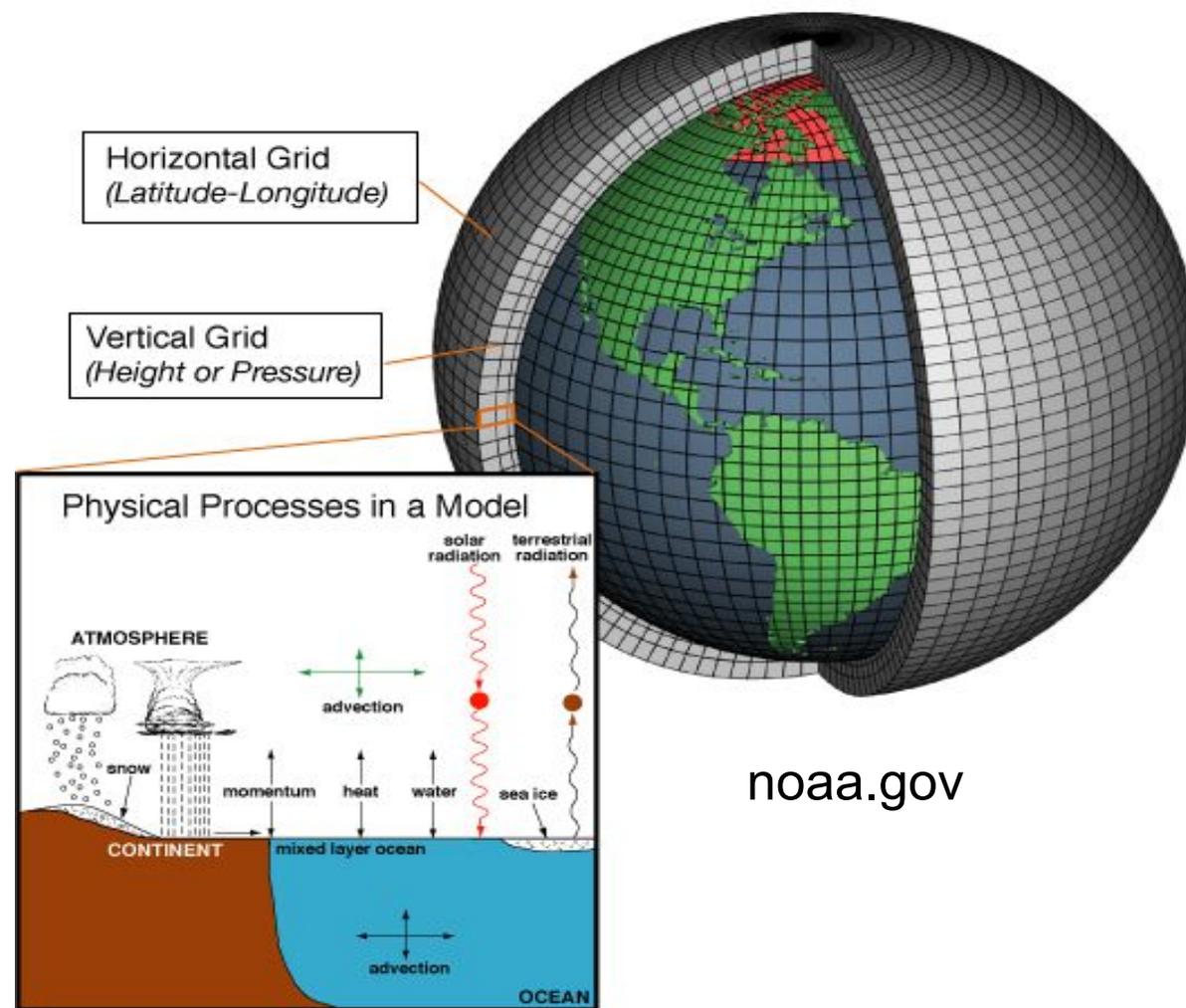
Núcleo dinámico

La parte del modelo que integra las ecuaciones de la dinámica atmosférica se denomina **núcleo dinámico**.

Rejilla hiper-dimensional.

Hay fenómenos importantes que no son resueltos. Tienen representados en el modelo por módulos llamados **parametrizaciones**.

Están basados también en las leyes de la física, pero con **aproximaciones** y **simplificaciones**, algunas empíricas.



noaa.gov

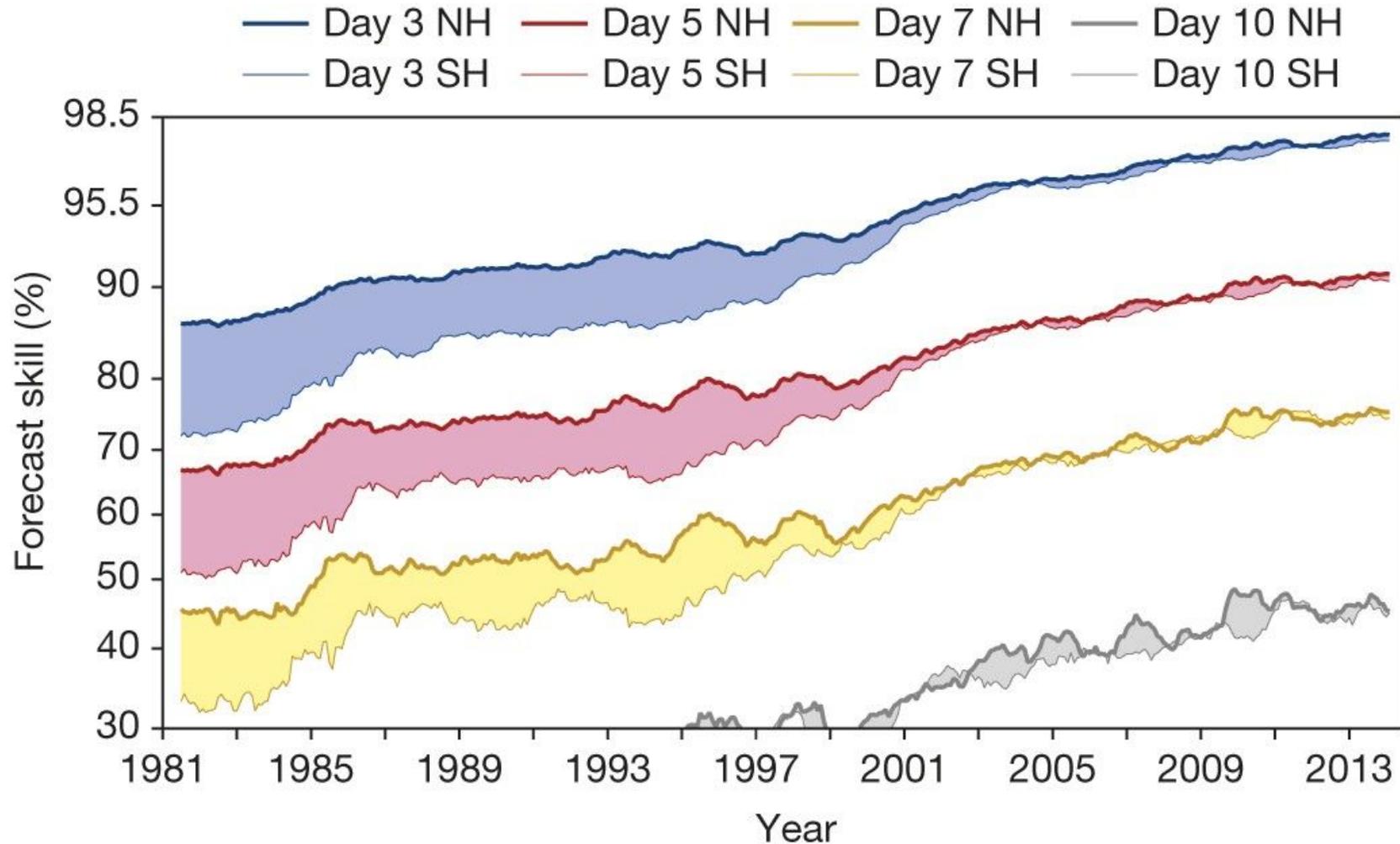


Fuentes de error

- Errores en las **condiciones iniciales**.
- Errores en las **parametrizaciones** (aproximaciones no exactas).
- Insuficiente **resolución** para representar correctamente la topografía y usos del suelo de la zona, y los fenómenos mesoescalares.
- Problemas en las **observaciones** de referencia usadas en la validación



The quiet revolution of NWP

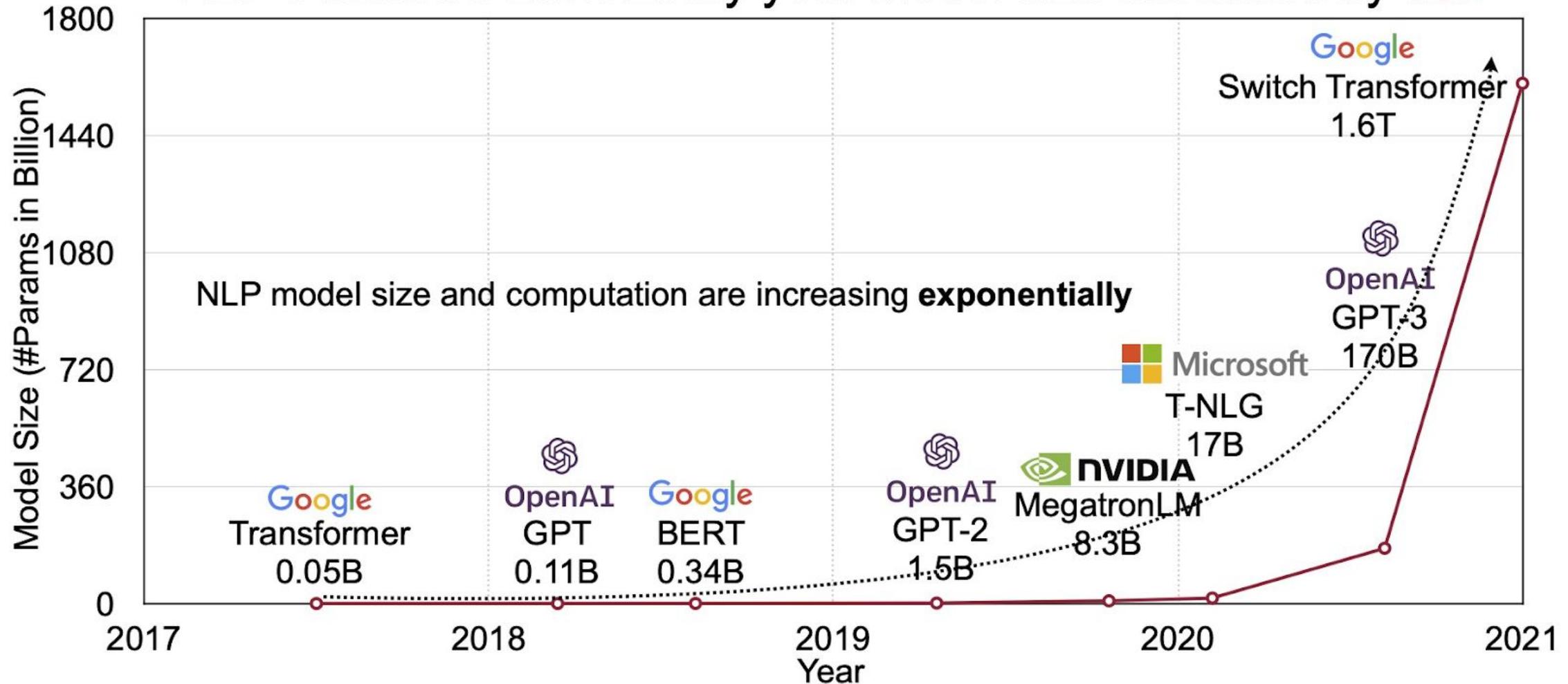


Bauer, P., Thorpe, A., Brunet, G.,
2015. The quiet revolution of
numerical weather prediction.
Nature 525, 47–55.
doi:10.1038/nature14956



and a non-quiet revolution: ML

NLP's Moore's Law: Every year model size increases by **10x**



Escalas de variabilidad

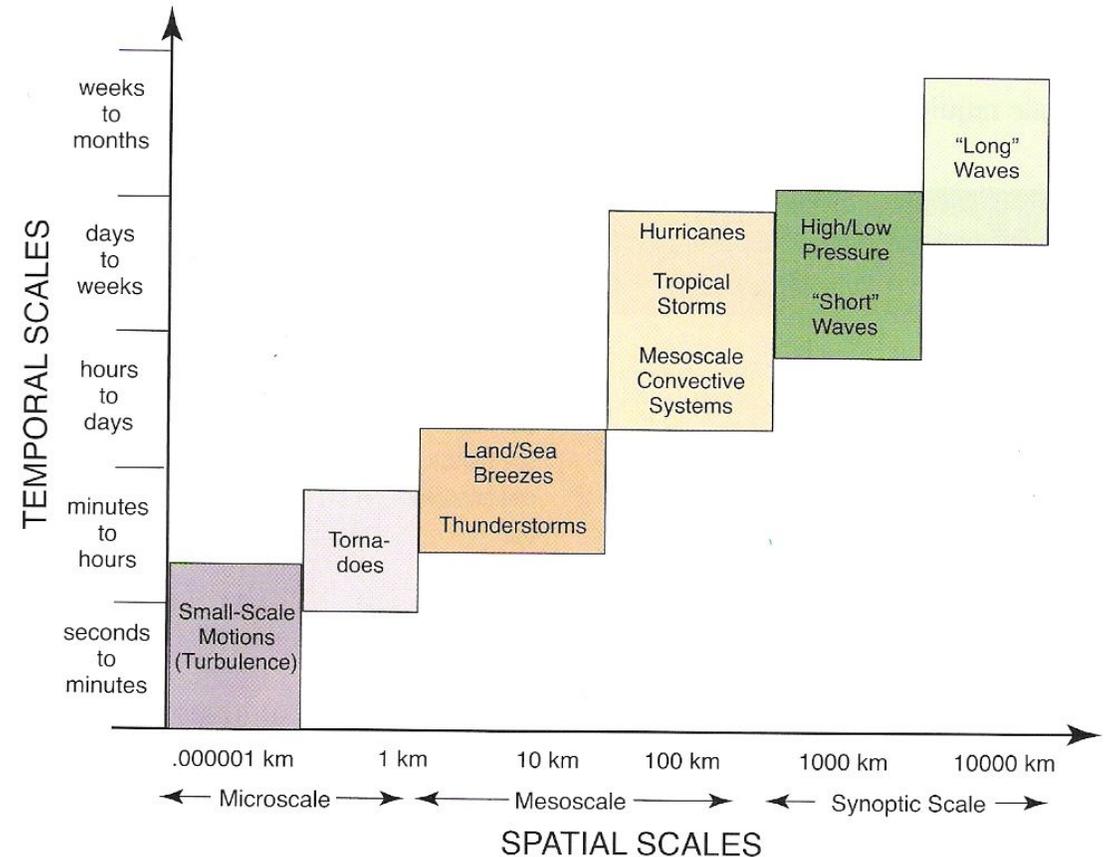
Cada fenómeno atmosférico tiene una escala temporal y espacial asociada.

Escala sinóptica: Borrascas, frentes

Mesoescala: Brisa marina, tormentas

Microescala: Térmicas, tornados, turbulencia

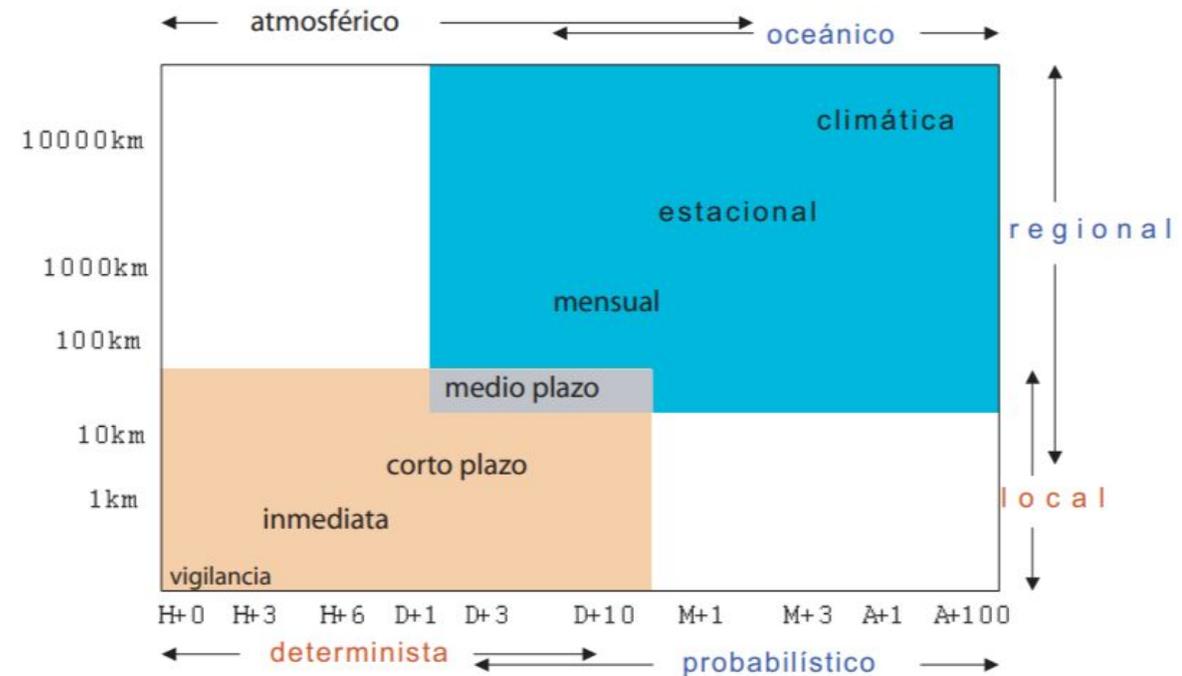
Es un **sistema caótico** y la predictibilidad depende de los horizontes de predicción



Problemas que resolvemos

- Predicción **meteorológica (corto y muy corto)**
- Predicción **sub-estacional y estacional**
- Proyecciones **climáticas**
- Pasado: reanálisis:

<https://www.youtube.com/watch?v=FAGobvUGl24>



jue. 18			vie. 19				sáb. 20		dom. 21		lun. 22	mar. 23	mié. 24
06-12 h	12-18 h	18-24 h	00-06 h	06-12 h	12-18 h	18-24 h	00-12 h	12-24 h	00-12 h	12-24 h			
14°C	9°C	7°C	5°C	4°C	2°C	-1°C							

Probabilidad de precipitación

0%	85%	75%	90%	85%	65%	15%	5%	0%	0%	0%	0%	0%	0%
----	-----	-----	-----	-----	-----	-----	----	----	----	----	----	----	----

Cota de nieve a nivel de provincia (m)

1500	1200	900	500	200	200								
------	------	-----	-----	-----	-----	--	--	--	--	--	--	--	--

Temperatura mínima y máxima (°C)

6 / 15	-3 / 7				-4 / 8		-4 / 7		-3 / 9	1 / 12	2 / 14
--------	--------	--	--	--	--------	--	--------	--	--------	--------	--------

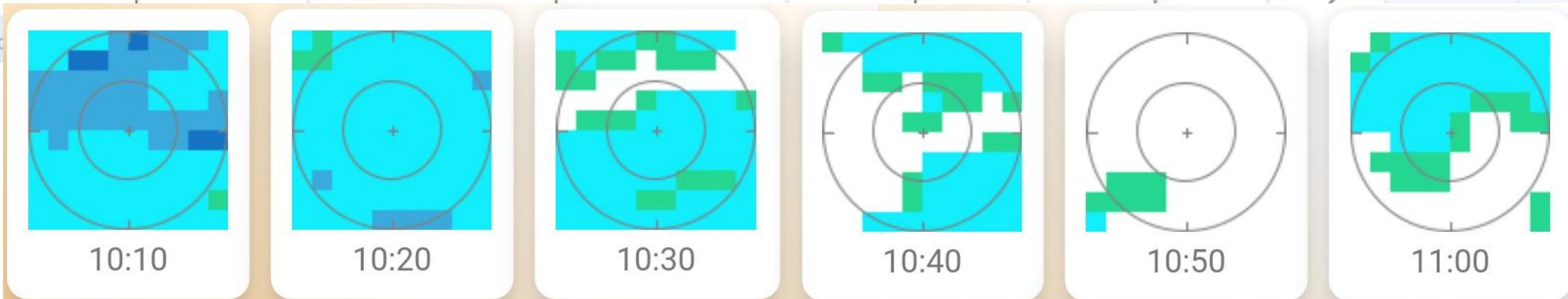
Dirección y velocidad del viento (km/h)

O	O	SO	NE	O	O	O	O	E	E	E	C	C	C
35	15	5	5	15	15	20	20	10	10	10	0	0	0

Índice ultravioleta máximo

1	1				1	1	1	2
---	---	--	--	--	---	---	---	---

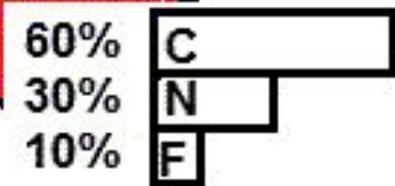
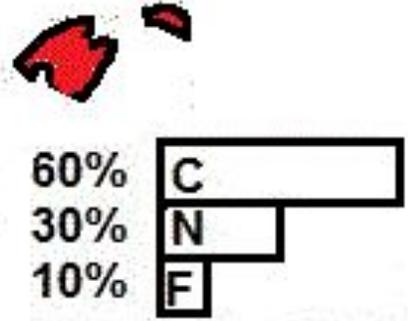
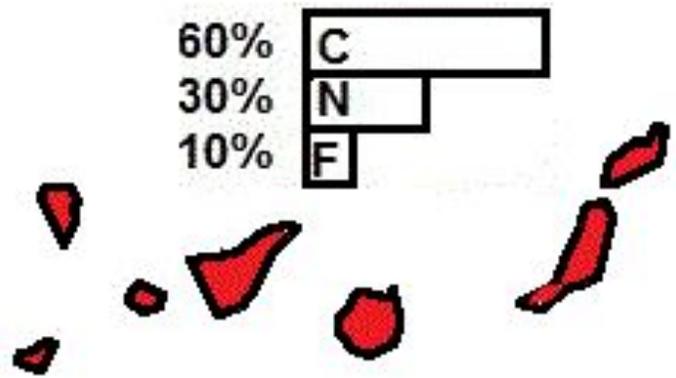
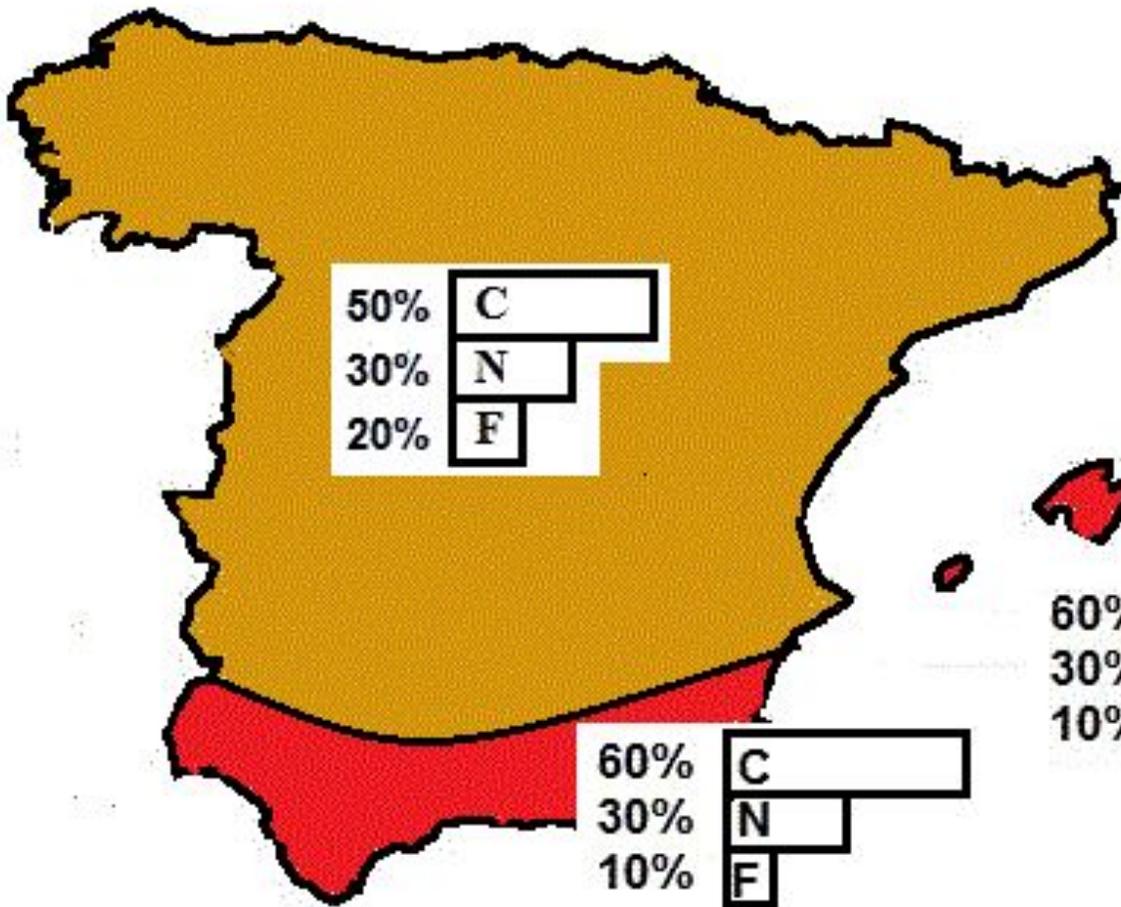
Aviso



PROBABILIDAD DE LA CATEGORÍA MÁS PROBABLE DE TEMPERATURA ENERO - FEBRERO - MARZO 2024

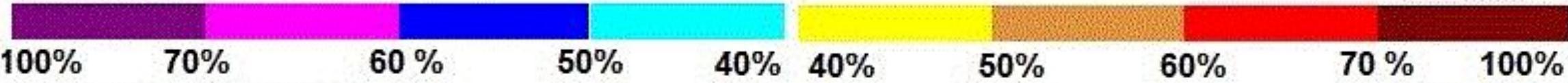
- C** Probabilidad tercil superior
- N** Probabilidad tercil central
- F** Probabilidad tercil inferior

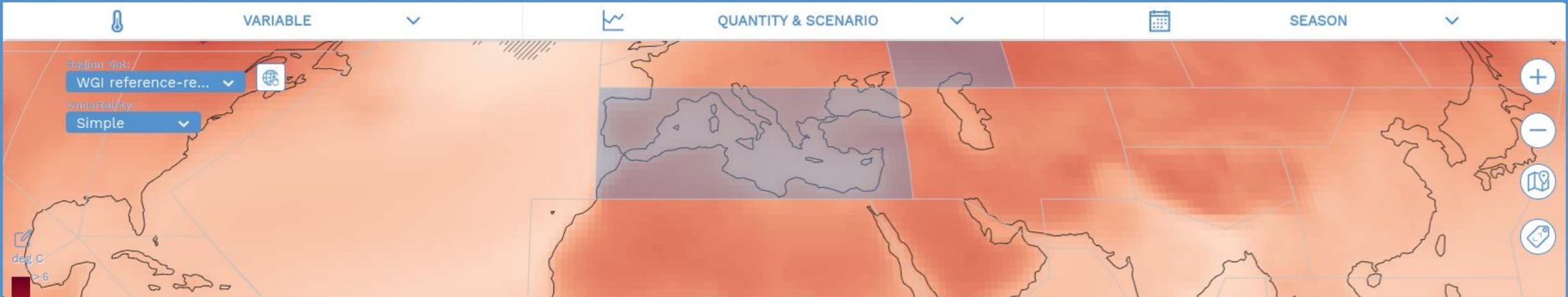
Los colores muestran la probabilidad de la categoría más probable.
El color blanco indica la climatología



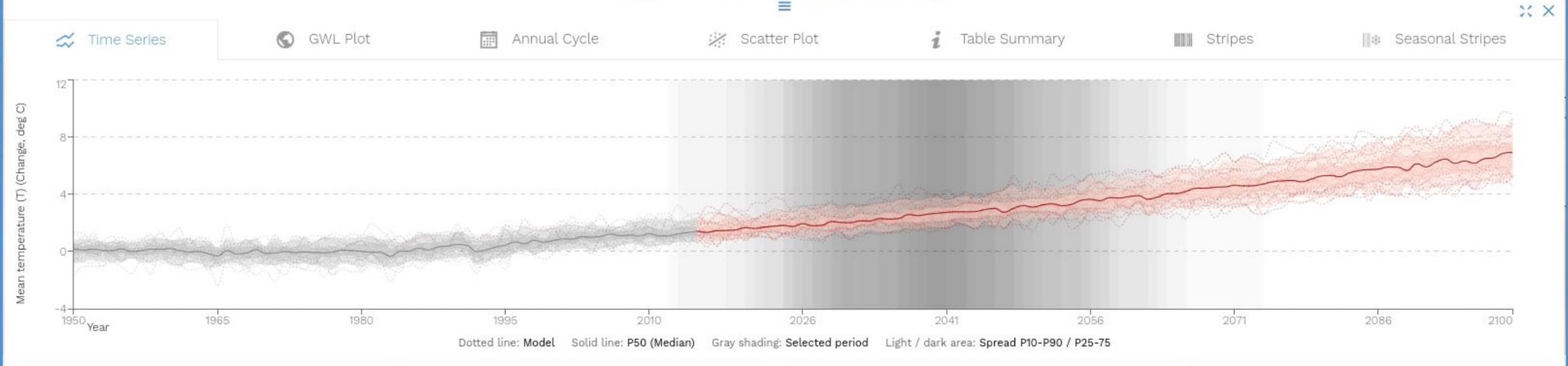
Probabilidad categoría inferior

Probabilidad categoría superior





CMIP6 - Mean temperature (T) Change deg C - Warming 2°C SSP5-8.5 (rel. to 1850-1900) - Annual (34 models)
Regions: Mediterranean, Eastern Europe



Export PDF



Export PNG



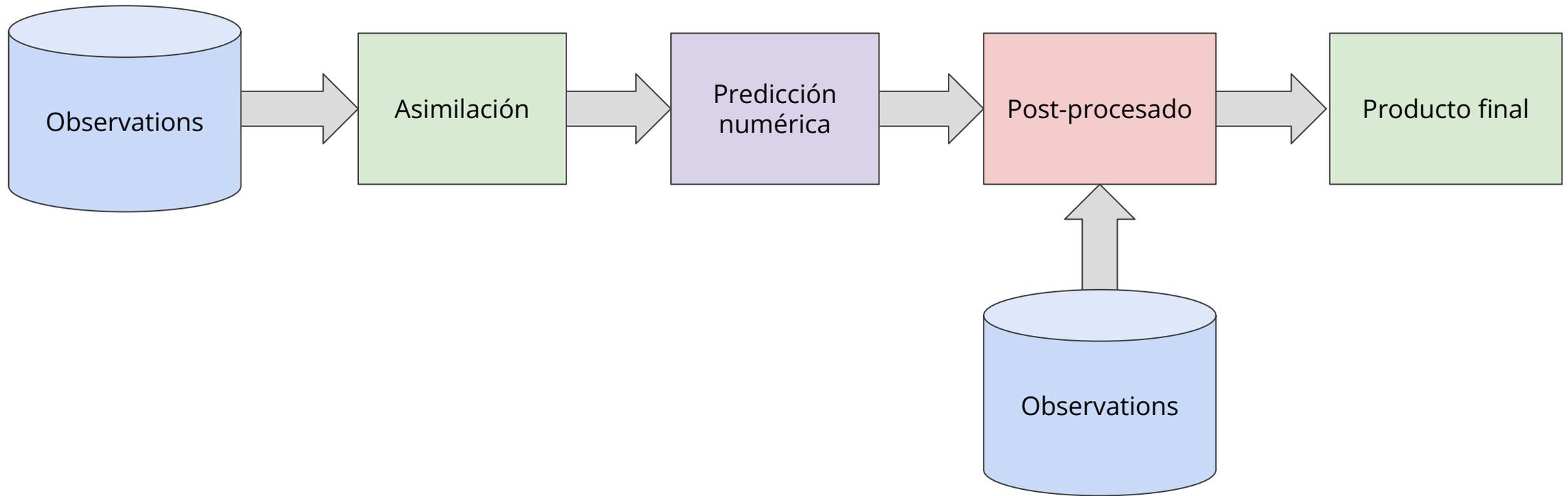
Mask: None



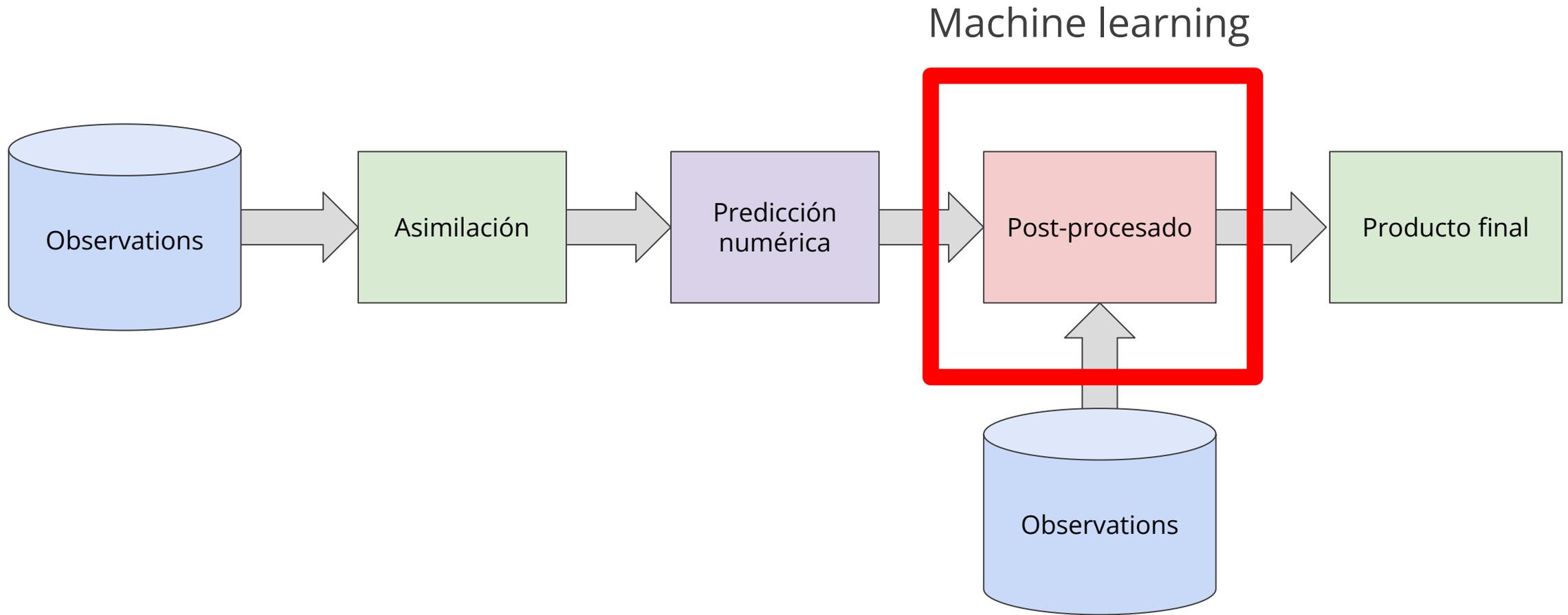
Models



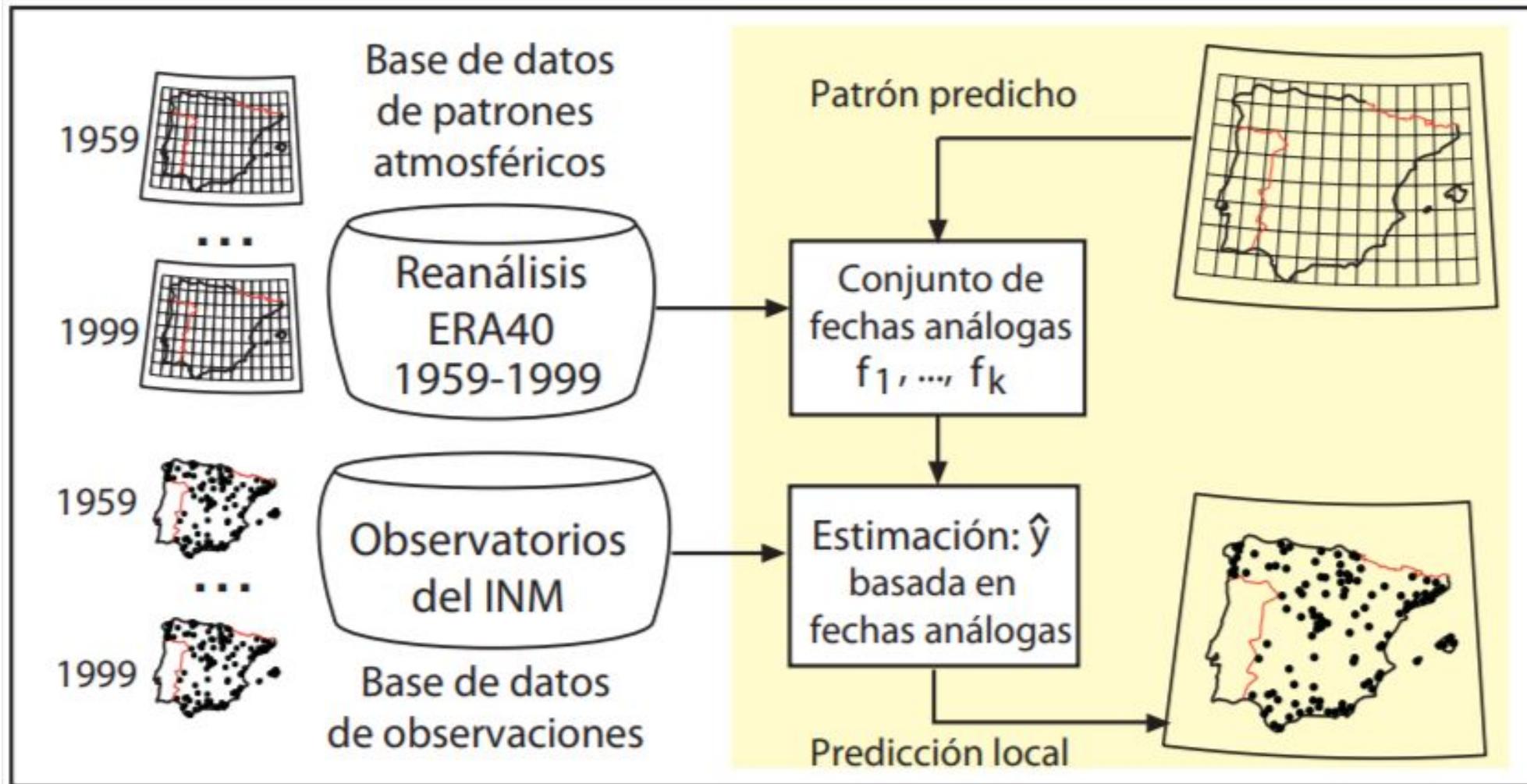
Predicción meteorológica



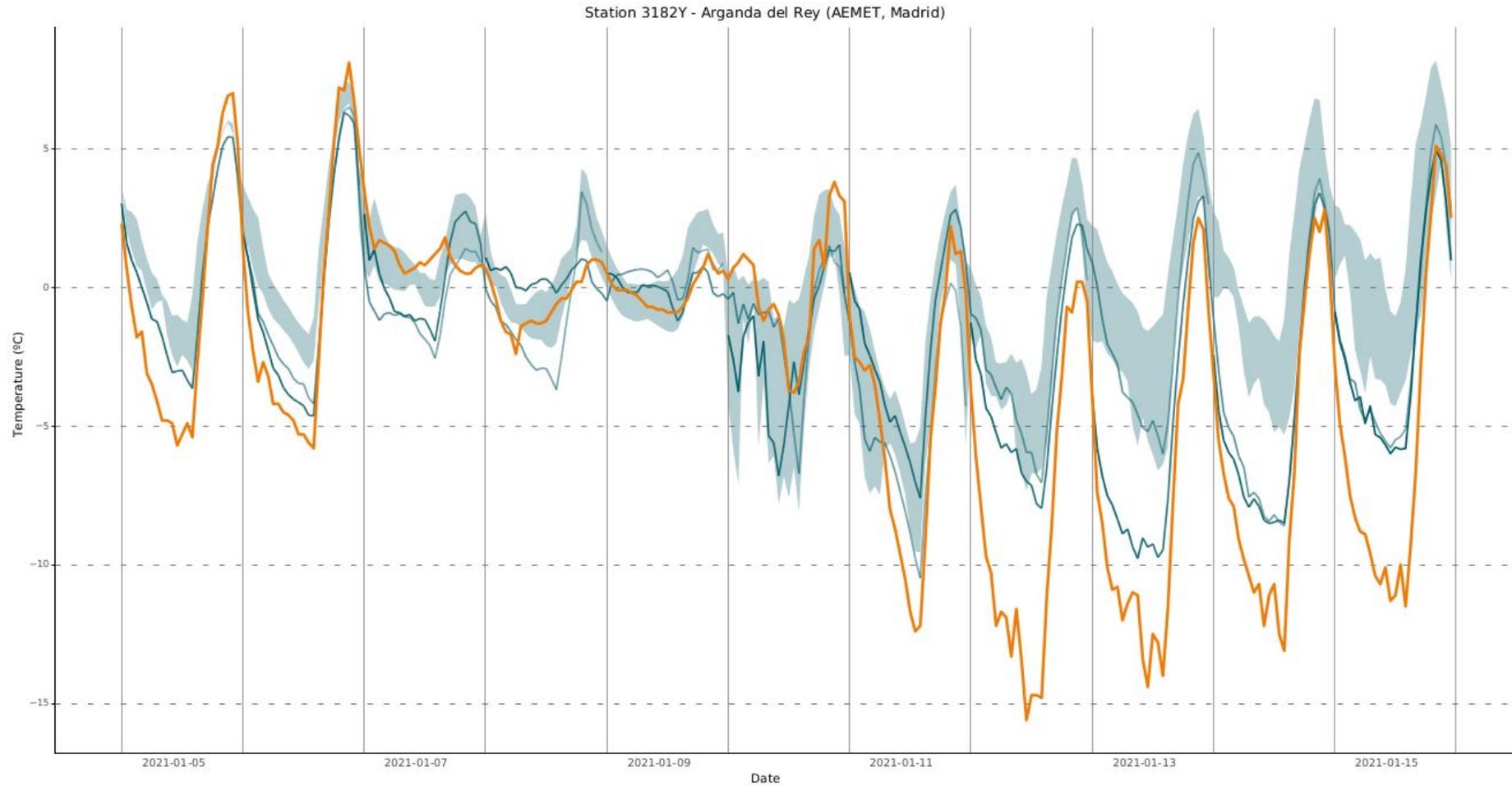
Predicción meteorológica



Análogos ~ KNN

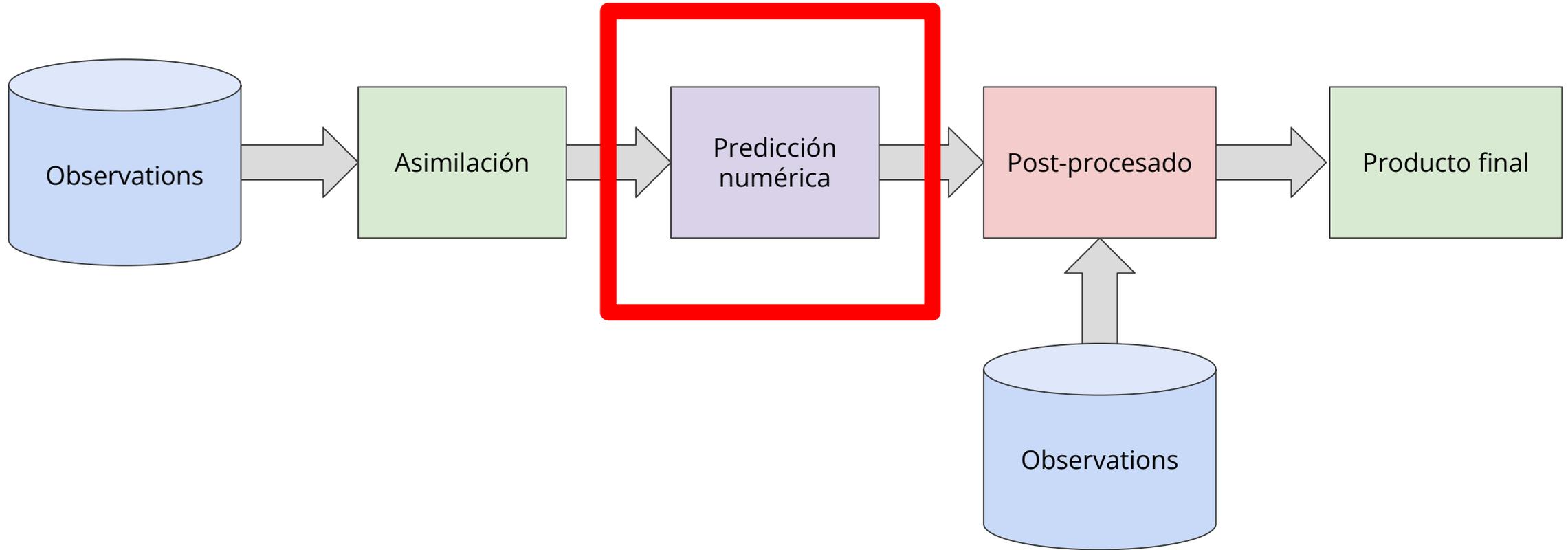


Ajustes dinámicos: KF, regresiones...



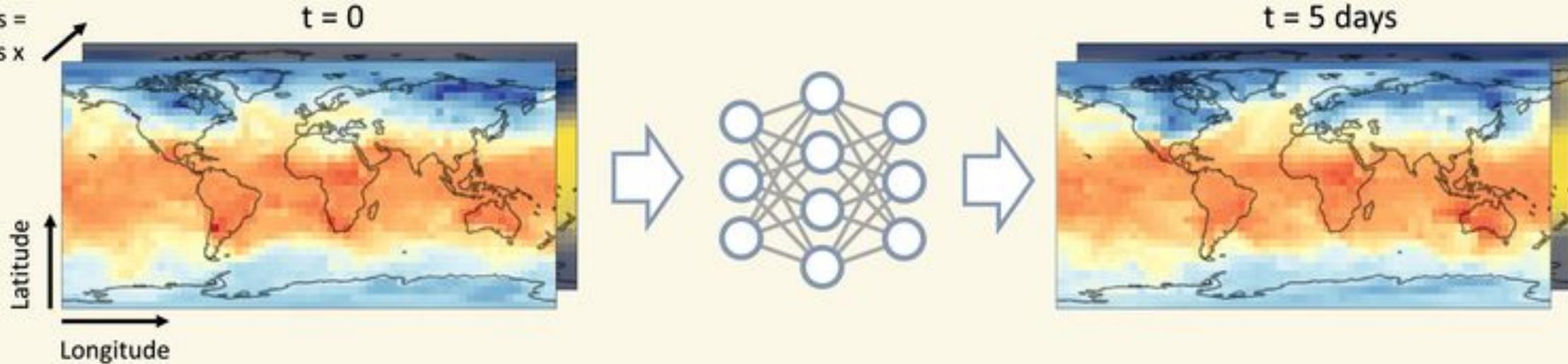
Predicción meteorológica

Machine learning

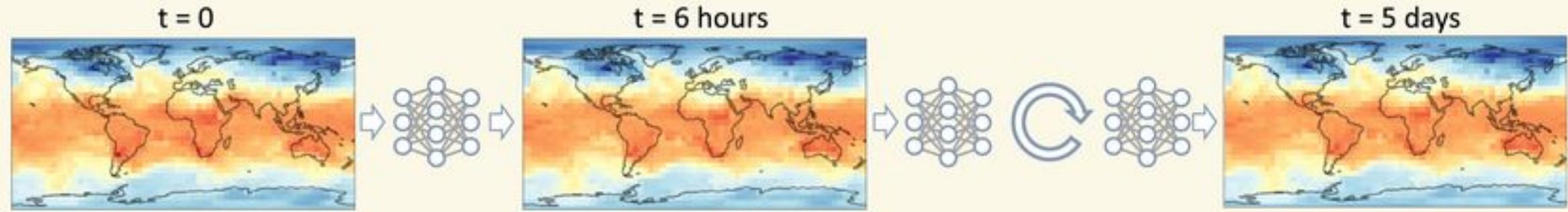


(a) Direct prediction

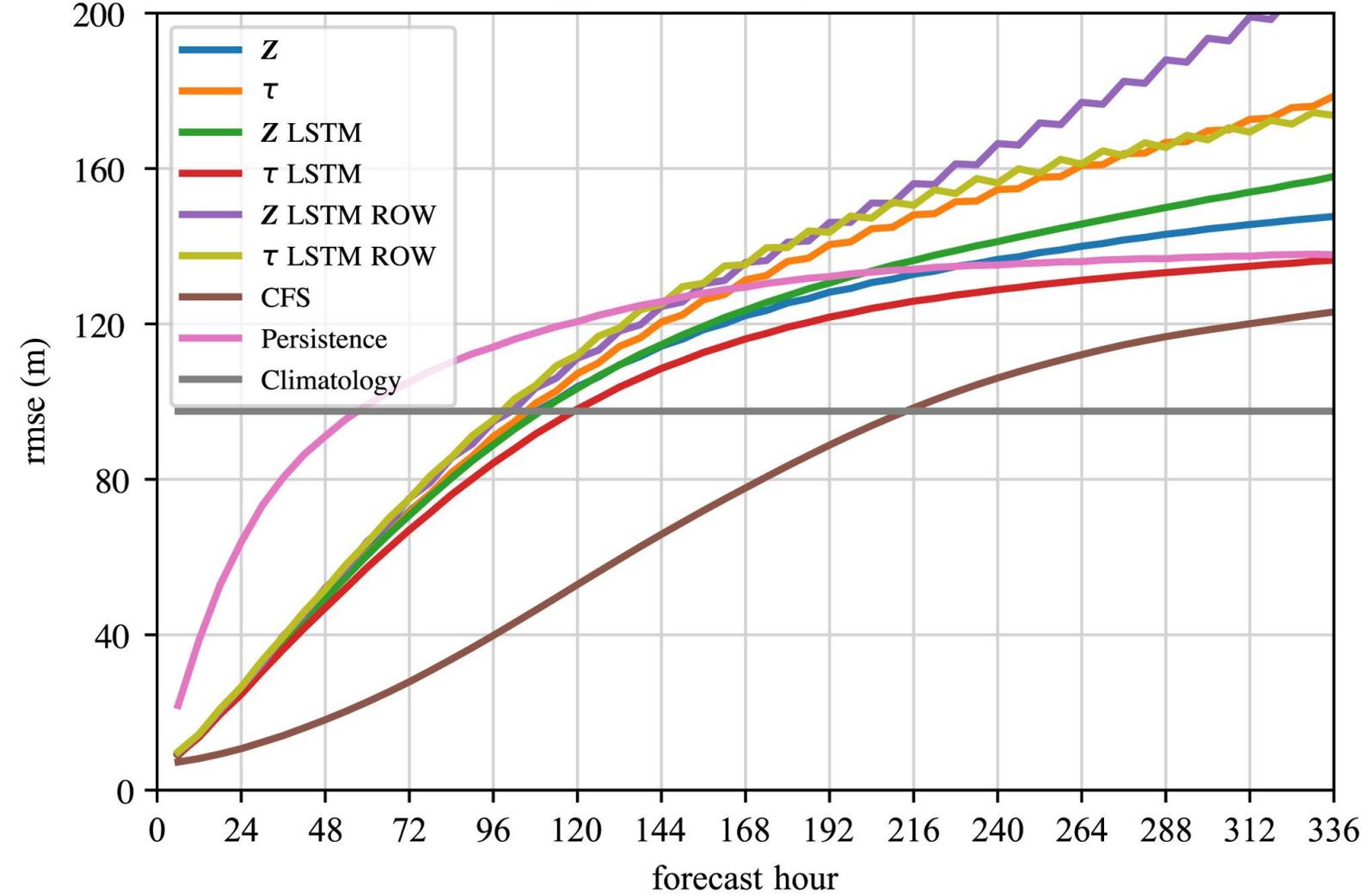
Channels =
Variables x
Levels



(b) Iterative prediction



Z_{500} ; 2007-2009; 20-70° N

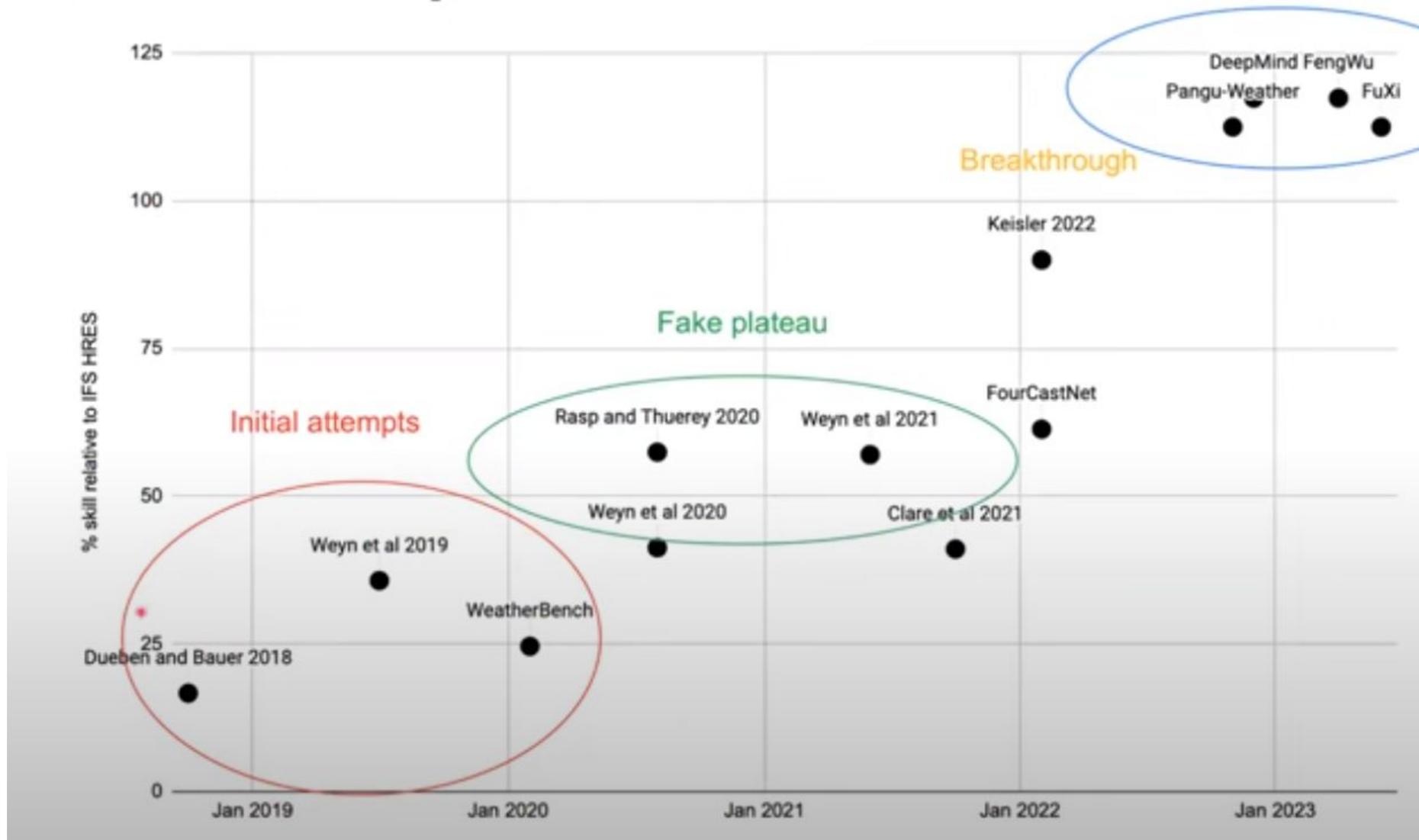


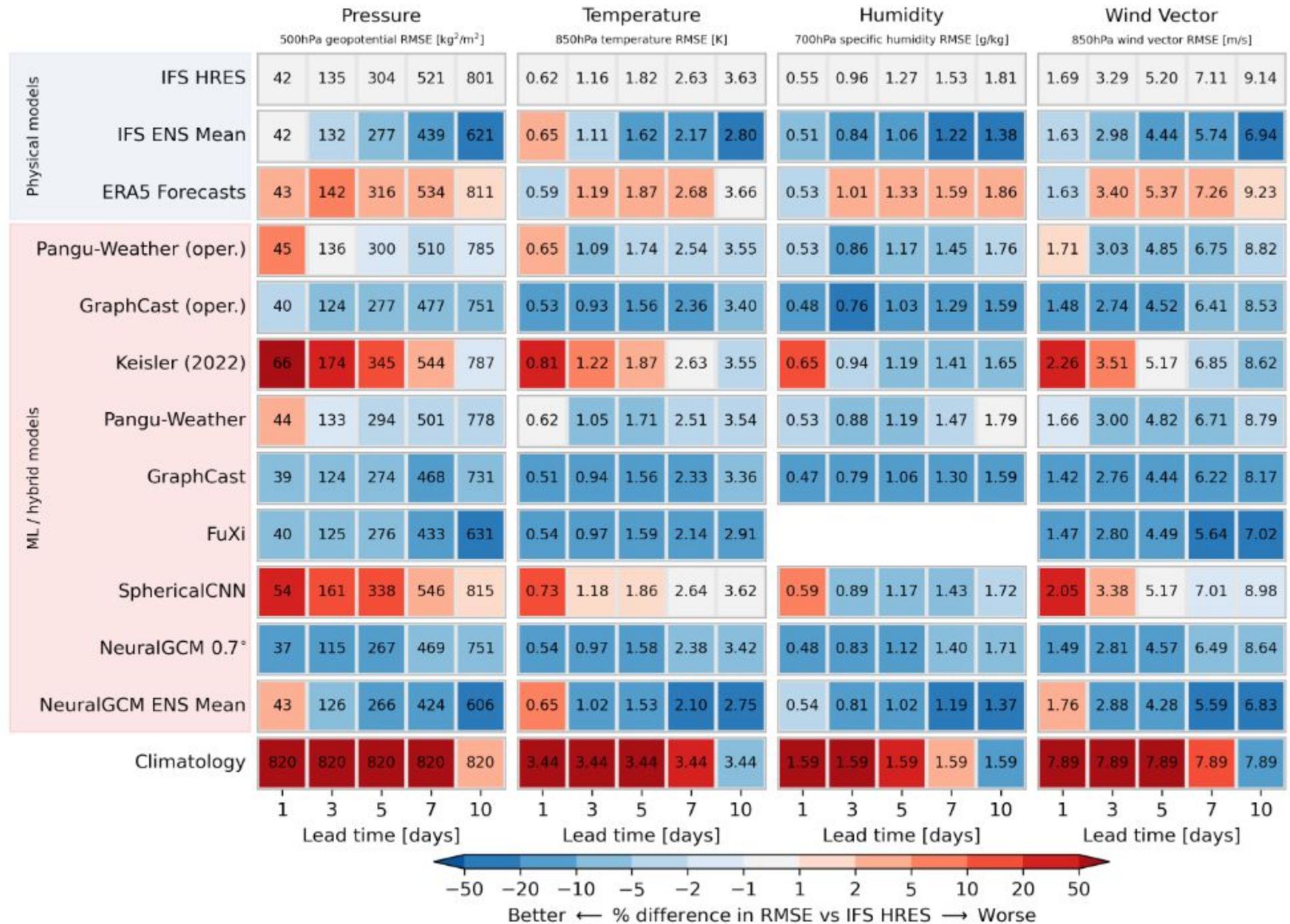
Can Machines Learn to Predict Weather? Using Deep Learning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data

2019



Second revolution of NWP

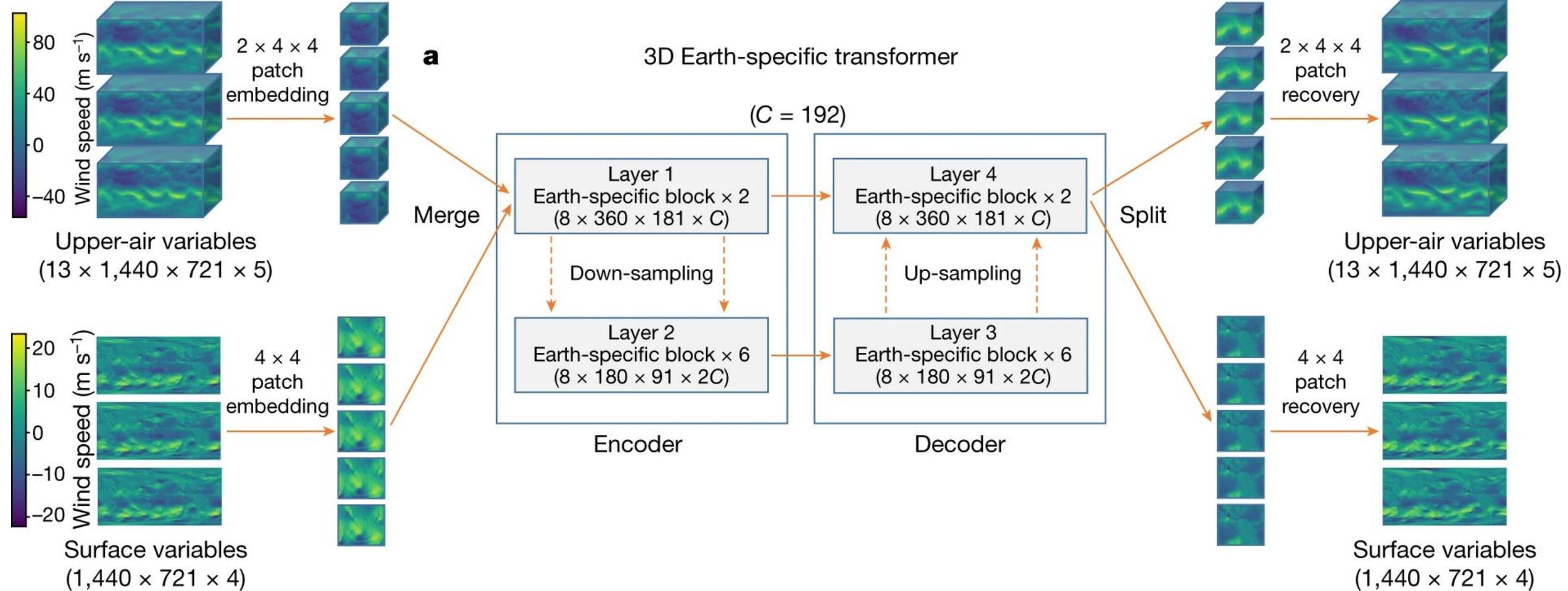




Arquitecturas

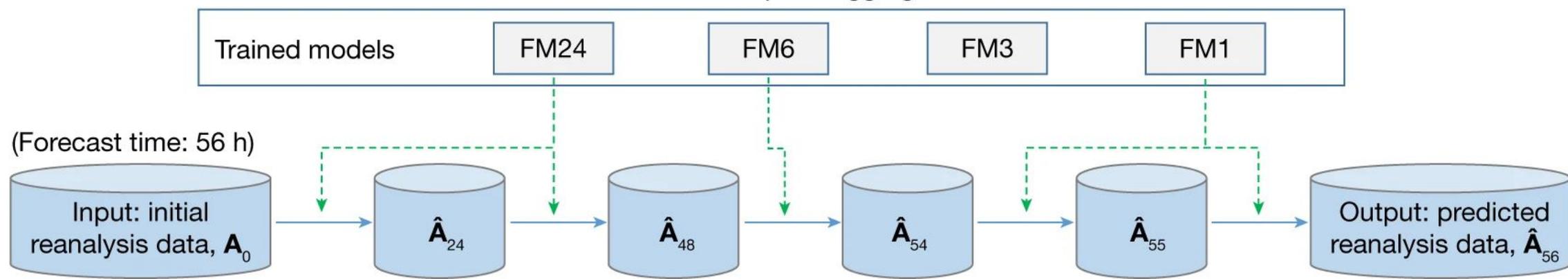
- La mayoría de ellos entrenados con **reanálisis**
- Usando:
 - CNN
 - Transformers
 - Redes neuronales basadas en grafos
 - Modelos de difusión
 - ...
- Muchos de ellos no son modelos “completos”

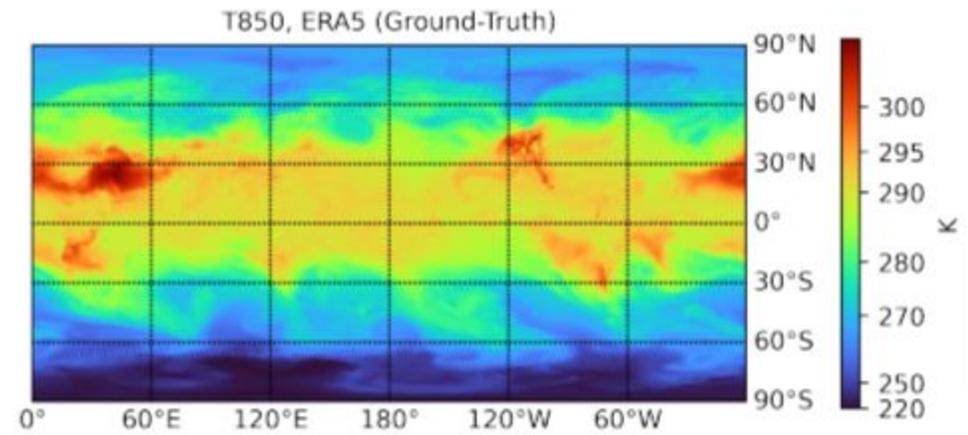
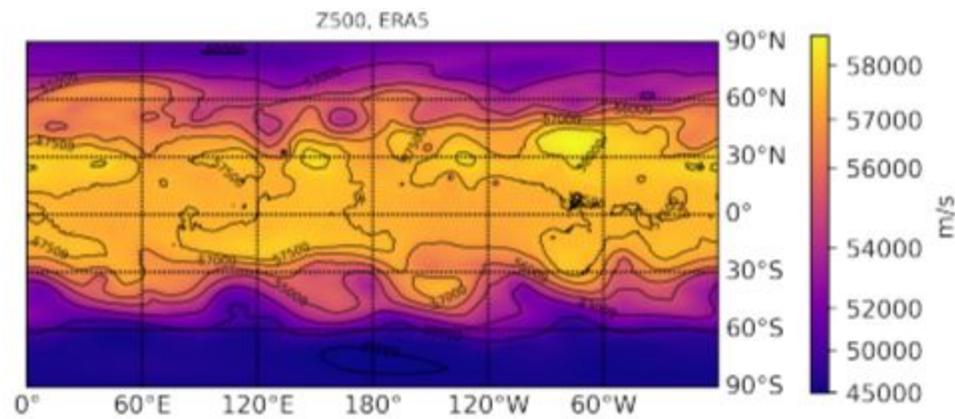
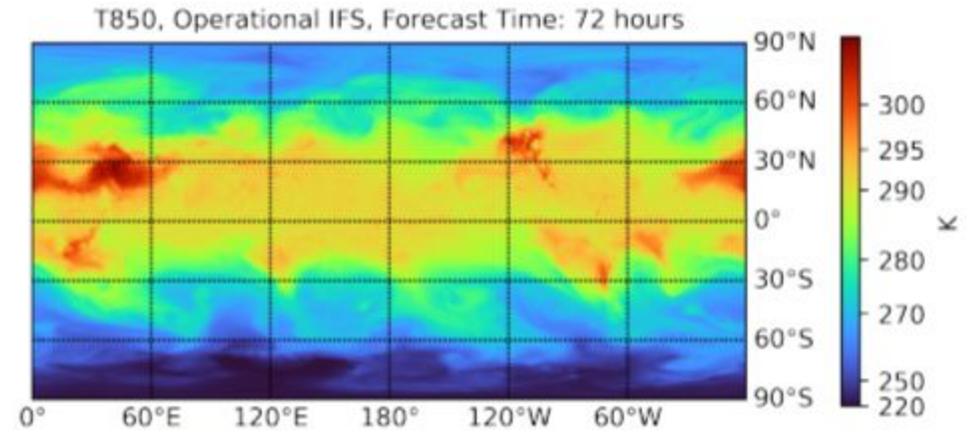
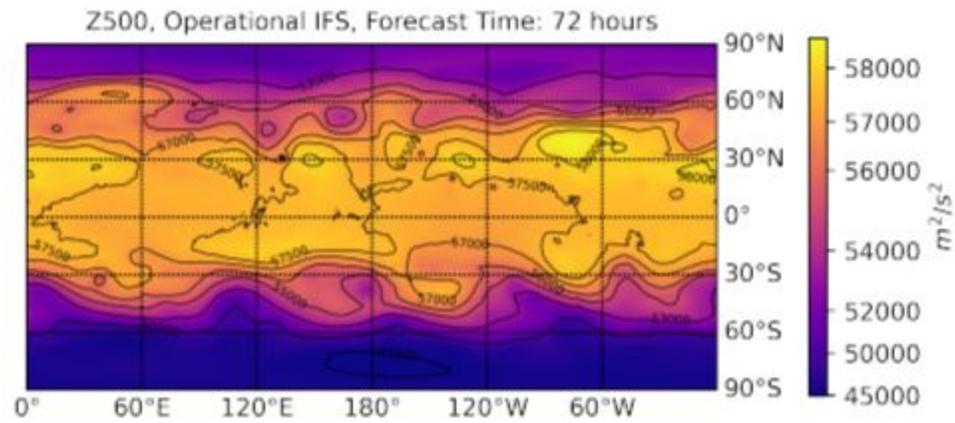
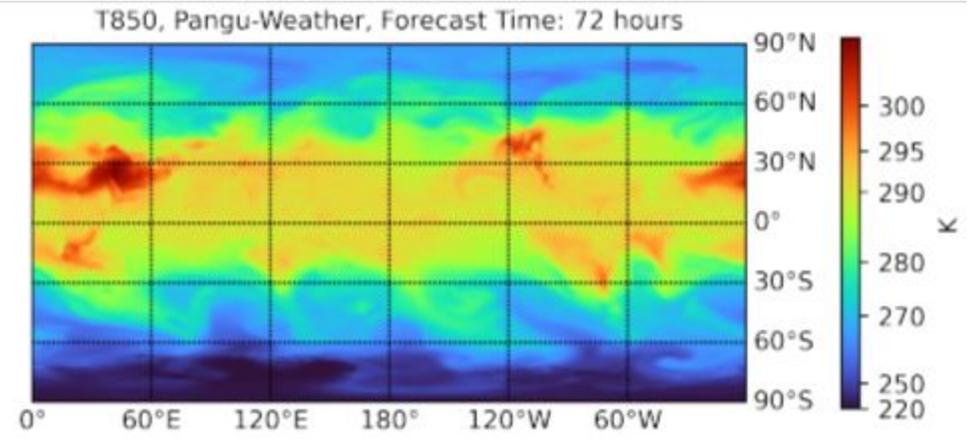
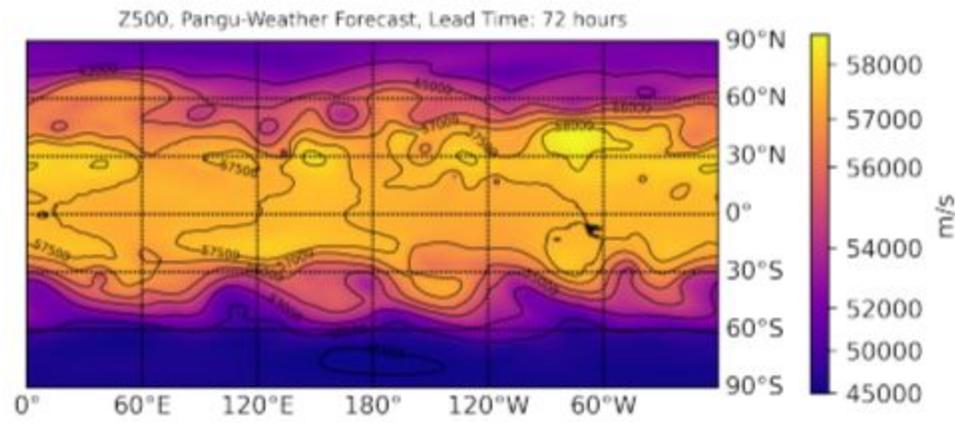




b

Hierarchical temporal aggregation





Estado actual

Ventajas:

- Rapidez
- Funcionan tan bien (o mejor) como los modelos físicos

Trabajo pendiente:

- Resolución
- Variables como precipitación
- Ensembles
- Validación



What the ML forecasts are showing: potential gain in time and energy

ERA5:
15 billion (one off)
(\$7.4Mio (compute only))

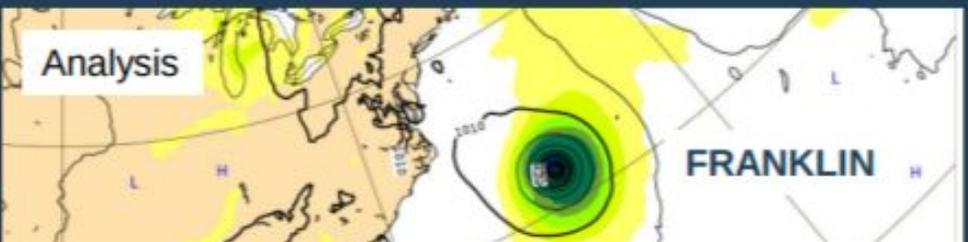
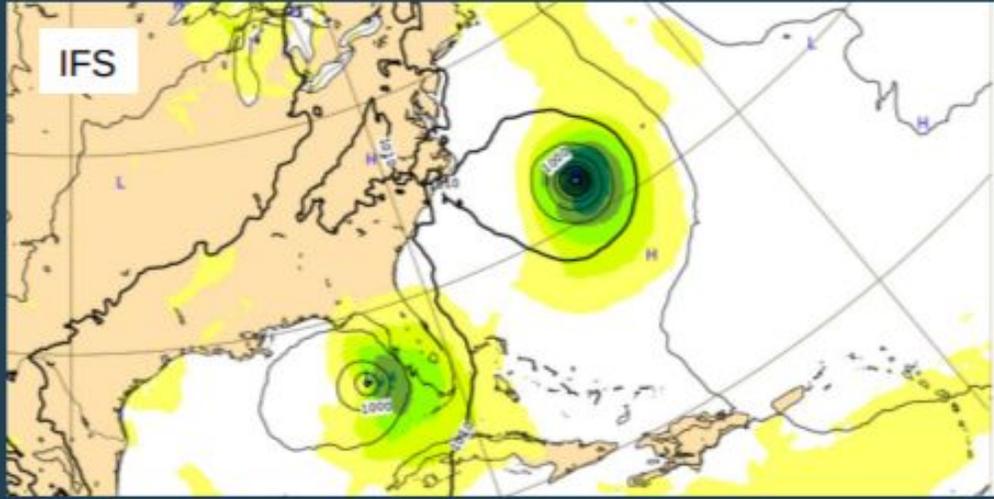
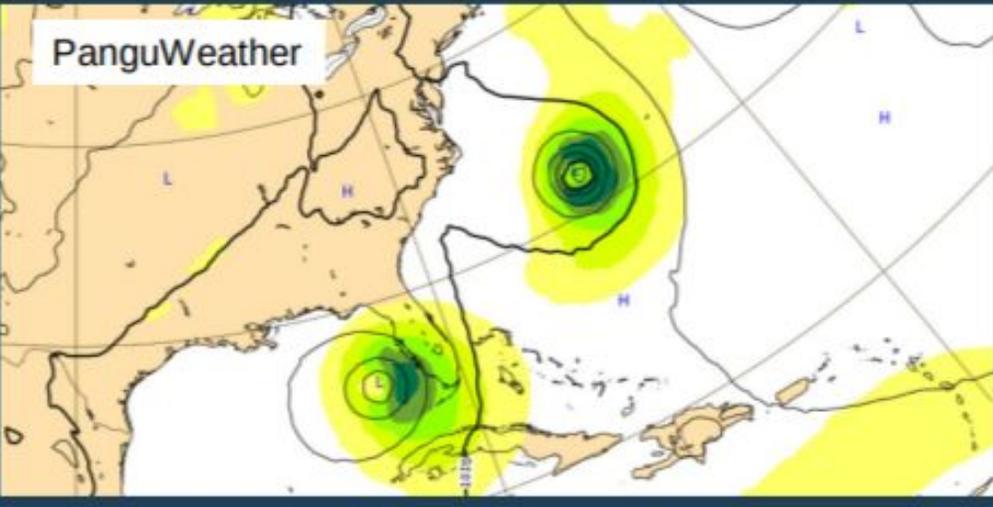
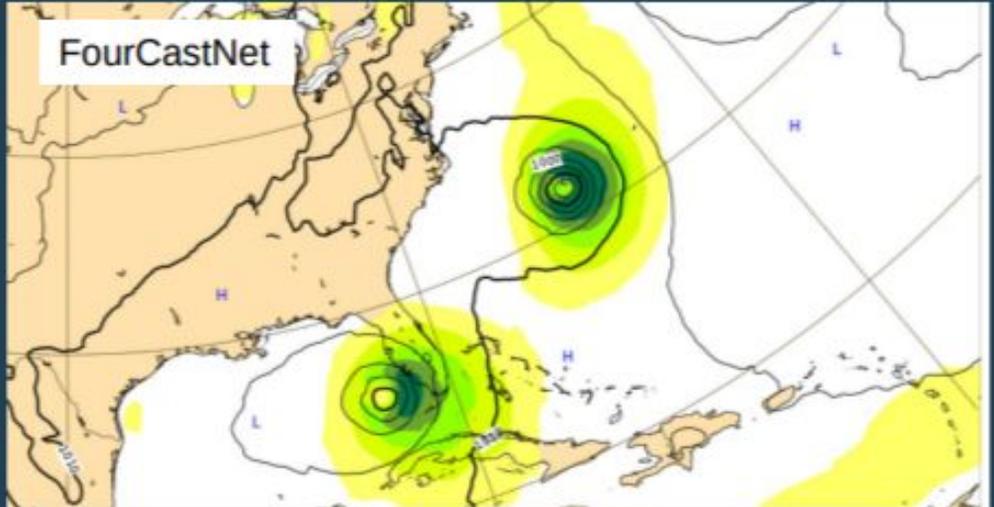


ECMWF HRES:
180 000 (\$90)
per forecast

Pangu:
0.3 (<€1)
per forecast

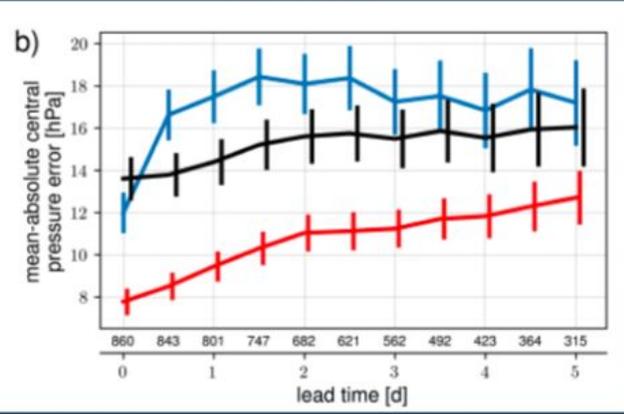
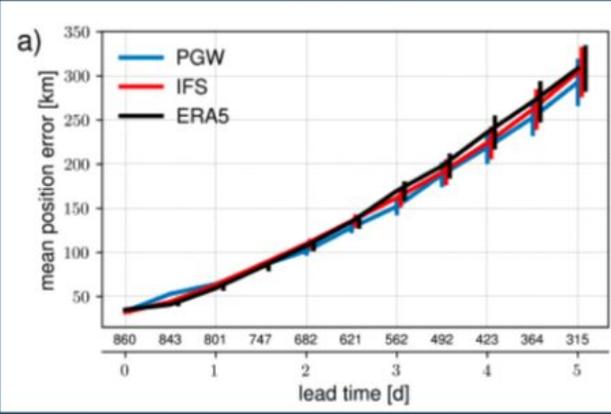


Tropical cyclones Idalia and Franklin (day 2 forecasts, valid on 30 Aug 2023 00UTC)



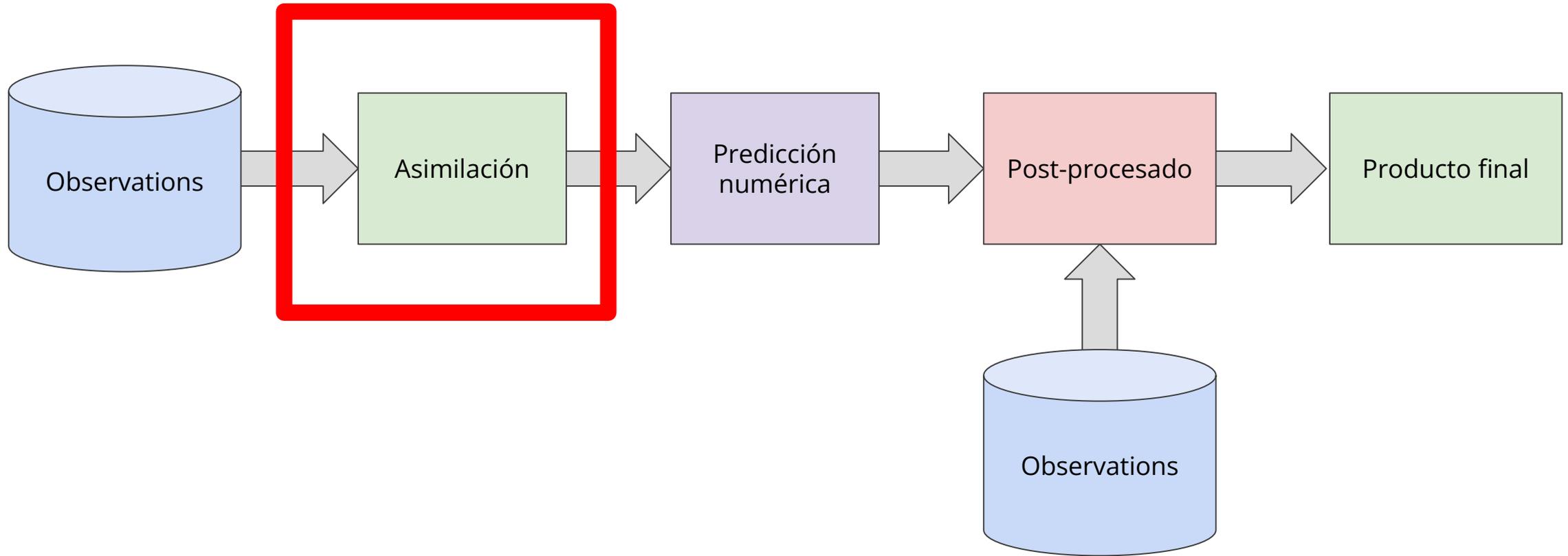
Position error

Intensity bias



Predicción meteorológica

Machine learning



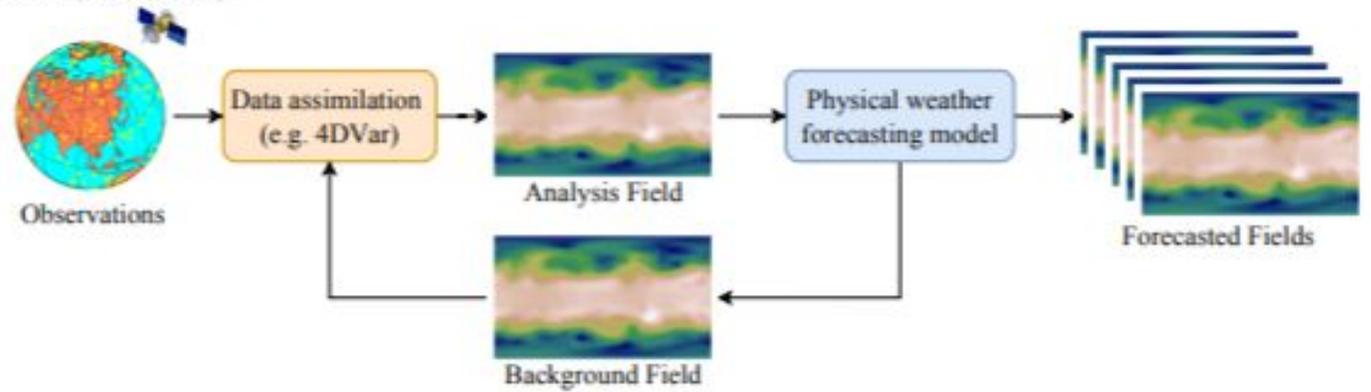
Asimilación

FengWu-4DVar: Coupling the Data-driven Weather Forecasting Model with 4D Variational Assimilation

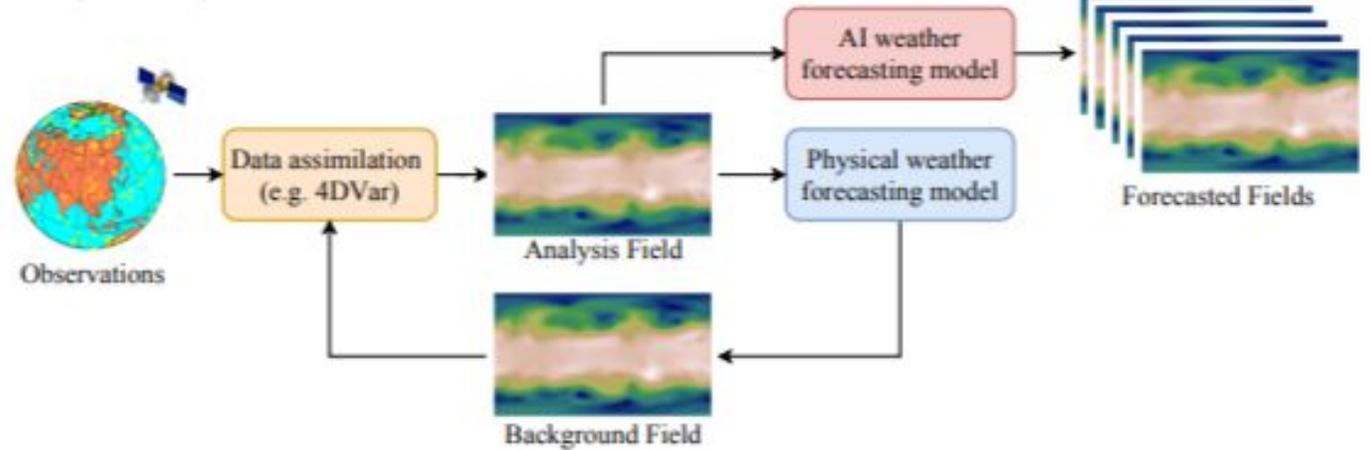
<https://arxiv.org/pdf/2312.12455.pdf>

Diciembre 2023

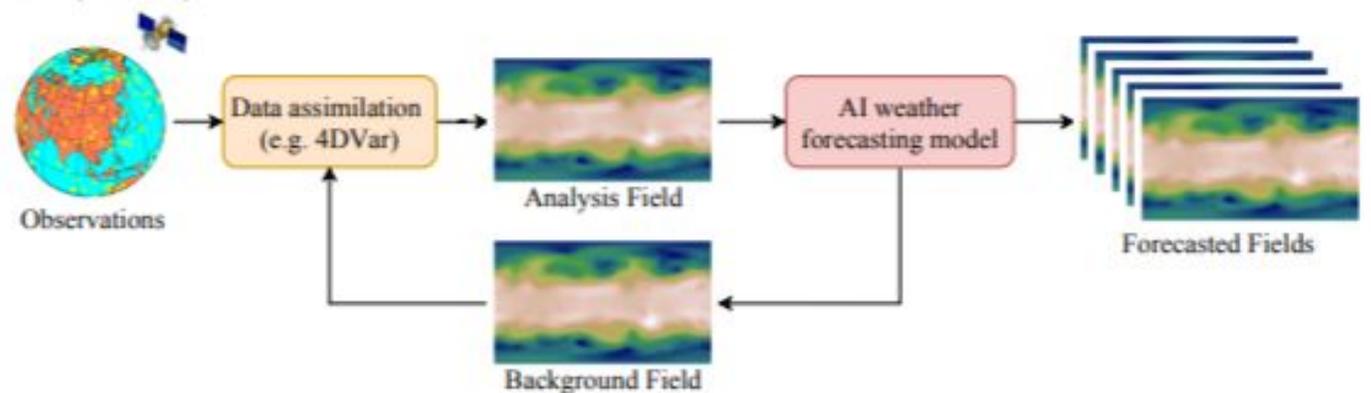
Phase 1 (Before 2022)



Phase 2 (2022-2023)



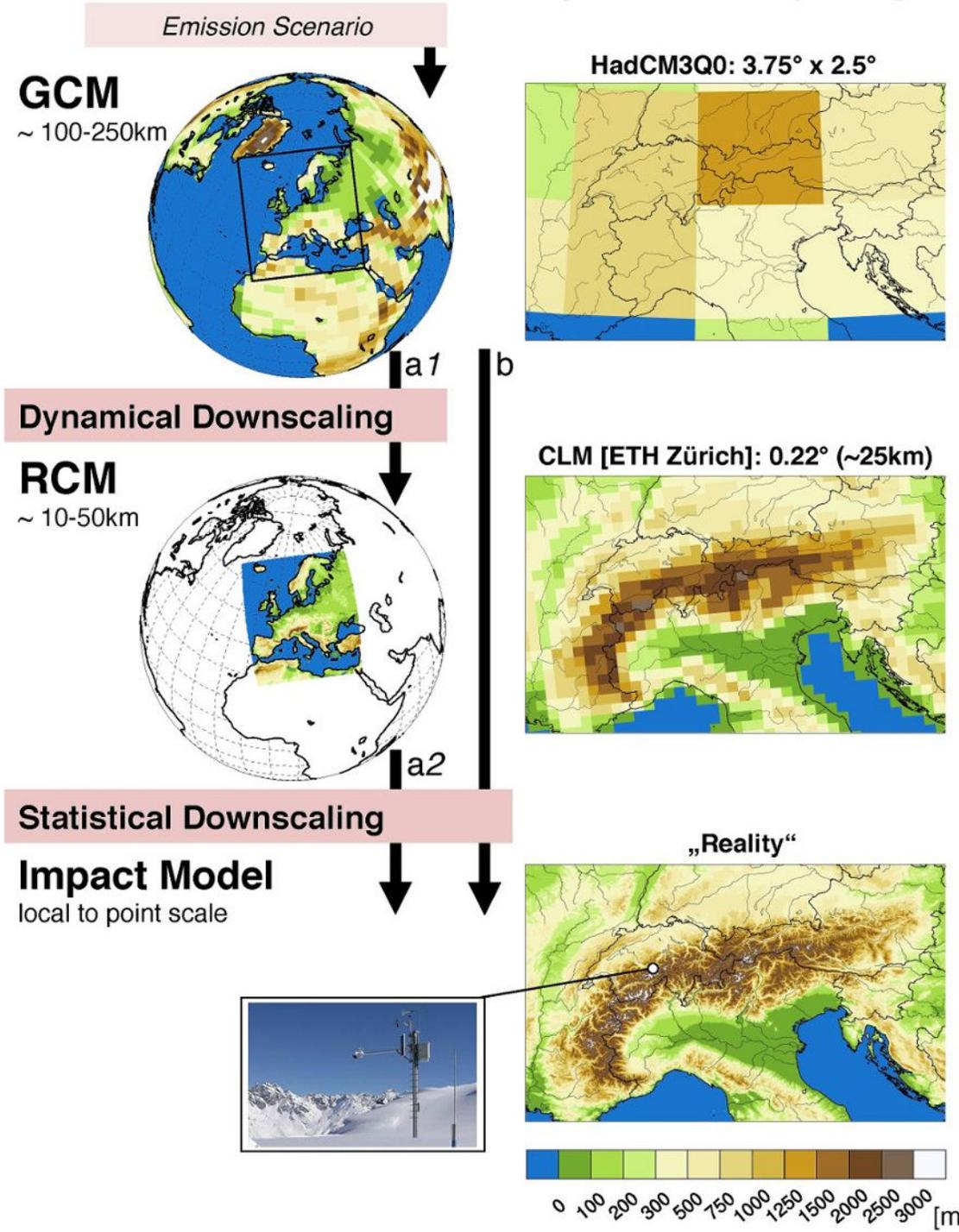
Phase 3 (our work)



Aplicaciones en modelización climática

Downscaling:

- Dinámico
- Estadístico





DATOS

Datos en rejilla ajustados (media)



VARIABLE

Precipitación



ESCENARIO

RCP 8.5

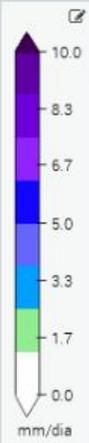
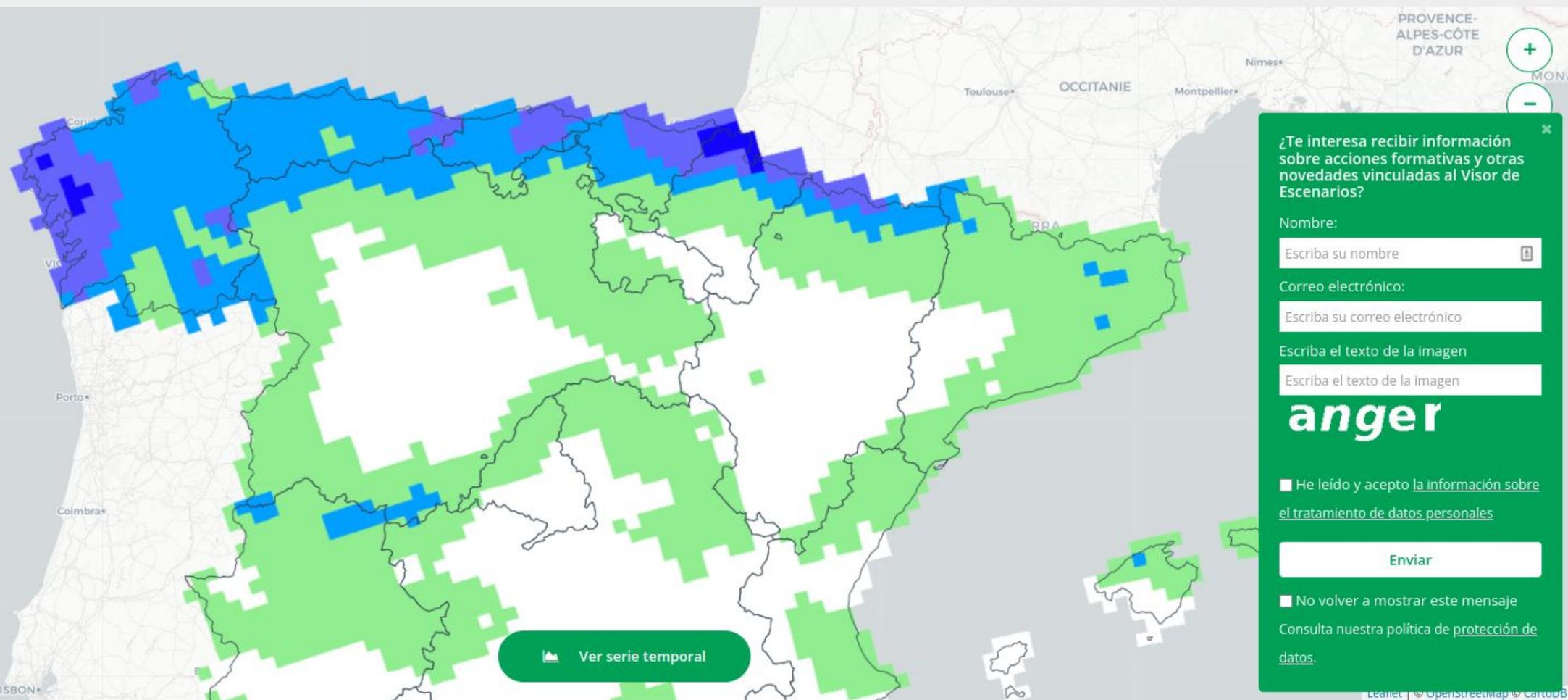


ESTACIÓN / MES

Año completo

Comunidades Autónomas

Área analizada (introduzca el nombre)



Periodo representado:

- Histórico ⓘ
- Futuro cercano ⓘ
- Futuro medio ⓘ
- Futuro lejano ⓘ

Ver serie temporal

¿Te interesa recibir información sobre acciones formativas y otras novedades vinculadas al Visor de Escenarios?

Nombre:

Correo electrónico:

Escriba el texto de la imagen

anger

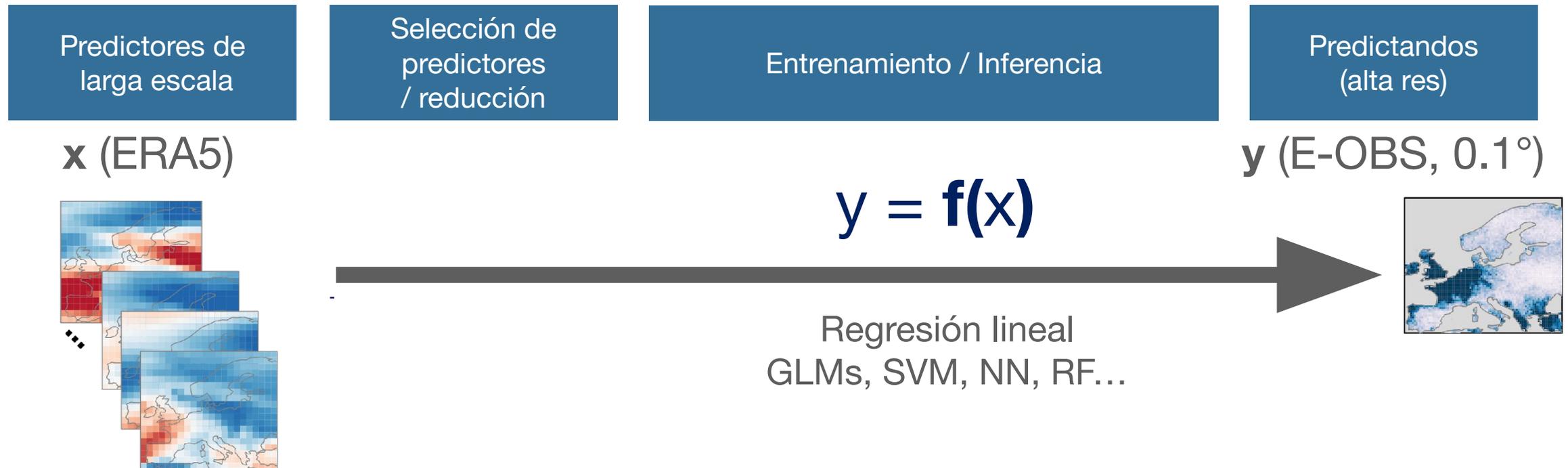
He leído y acepto [la información sobre el tratamiento de datos personales](#)

No volver a mostrar este mensaje

Consulta nuestra política de [protección de datos](#).

Statistical downscaling

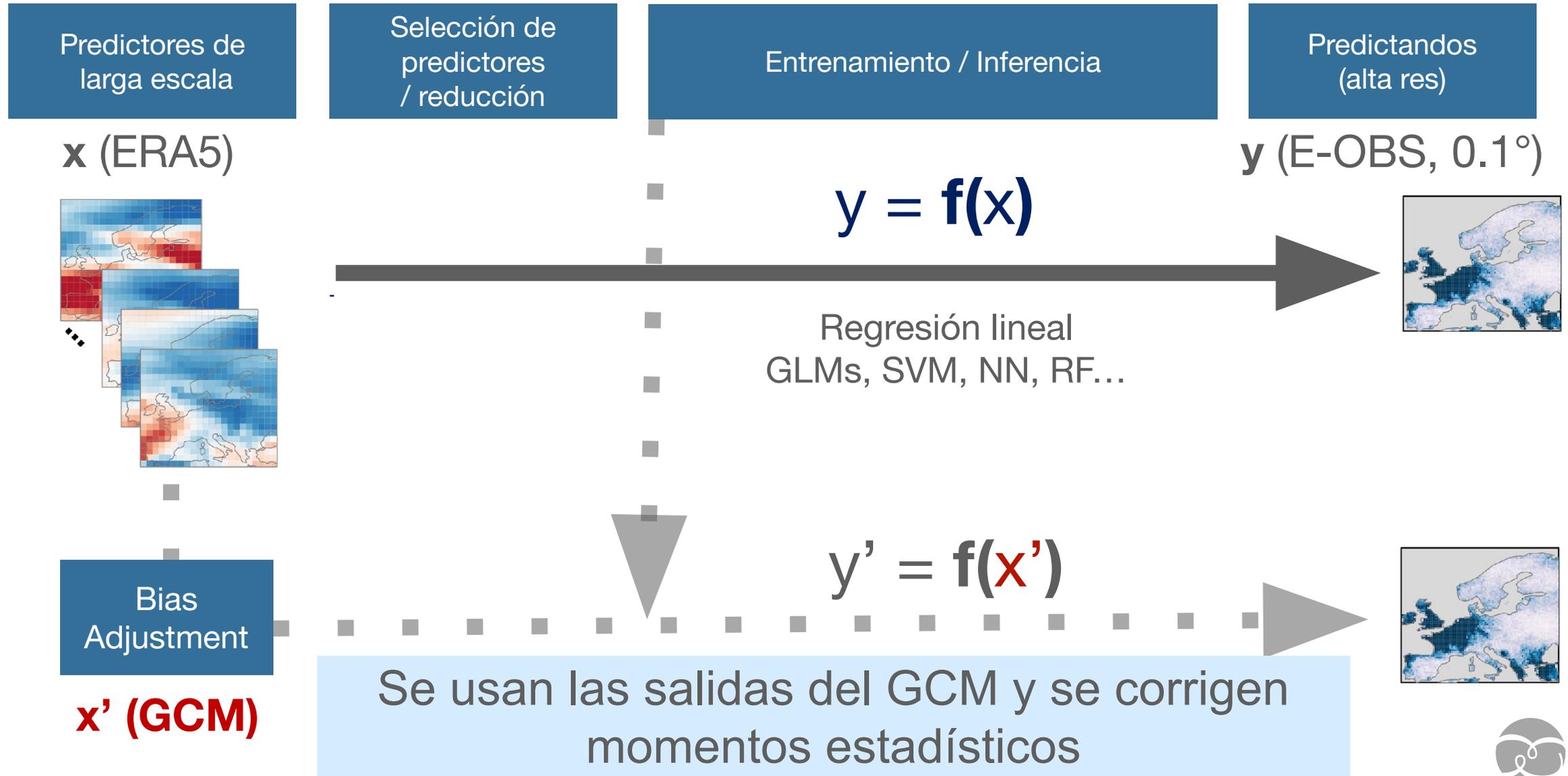
Gutiérrez et al. (2019)
<https://doi.org/10.1002/joc.5462>



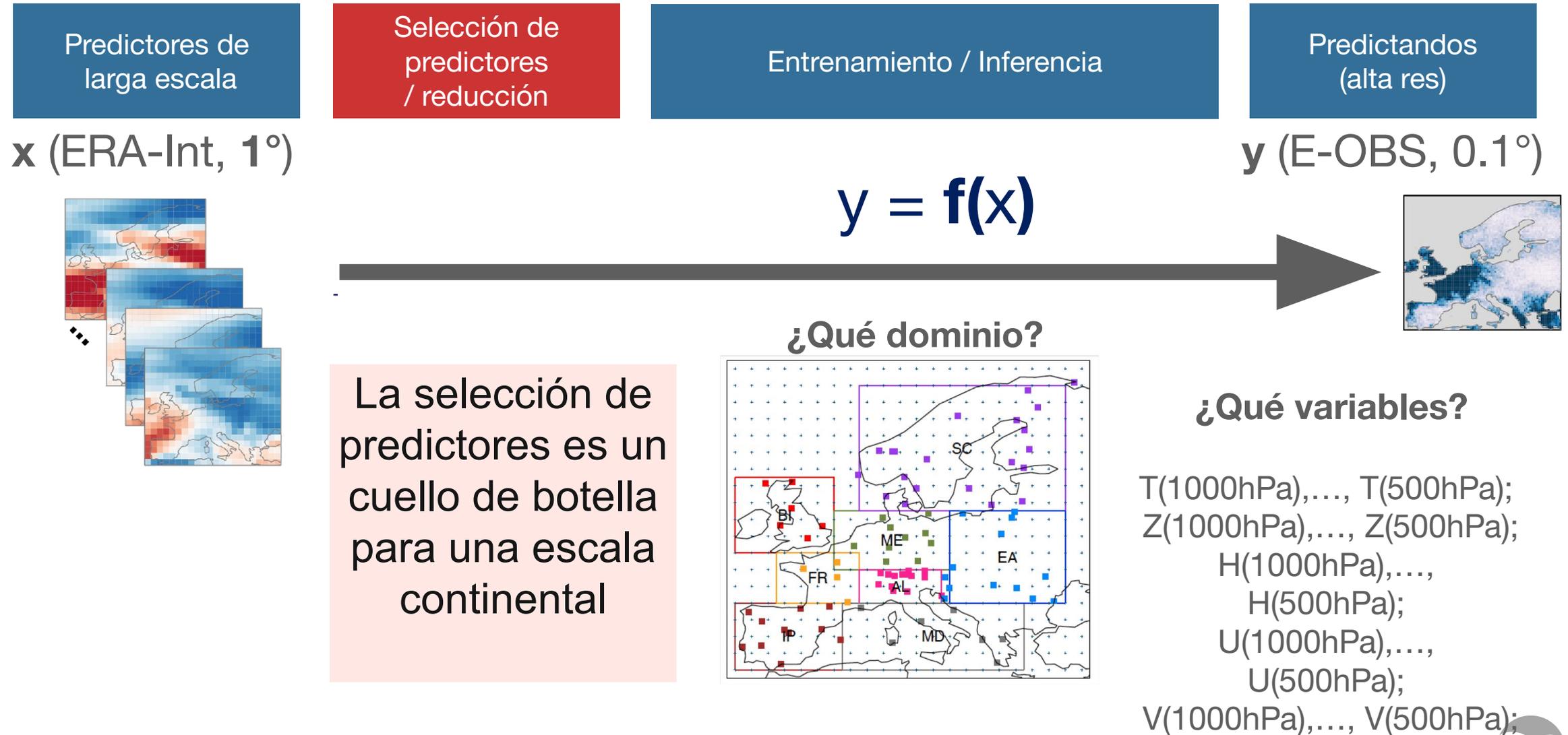
Perfect Prognosis (PP) aprende relaciones empíricas entre un conjunto de predictores atmosféricos de larga escala (reanálisis) y unos predictandos (target) de observados de alta resolución



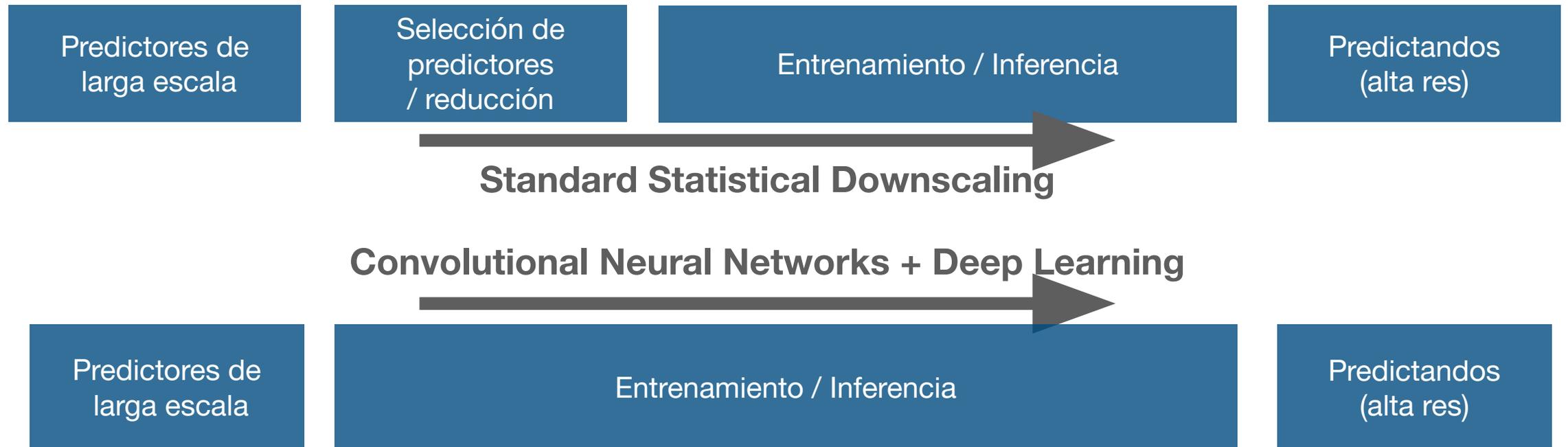
Statistical downscaling



Statistical downscaling: Problem



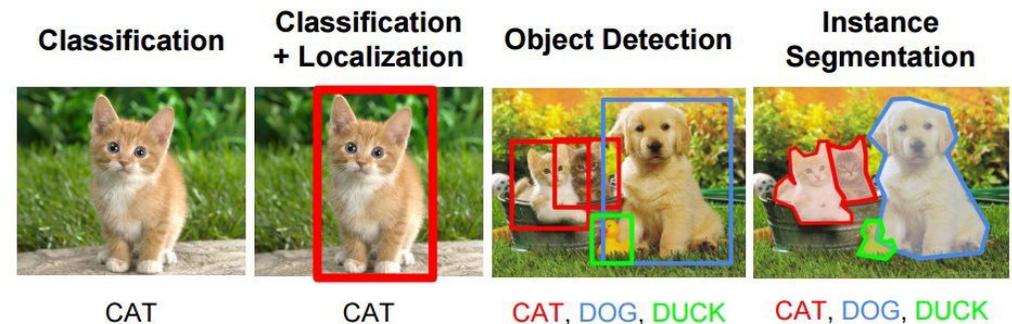
Statistical downscaling



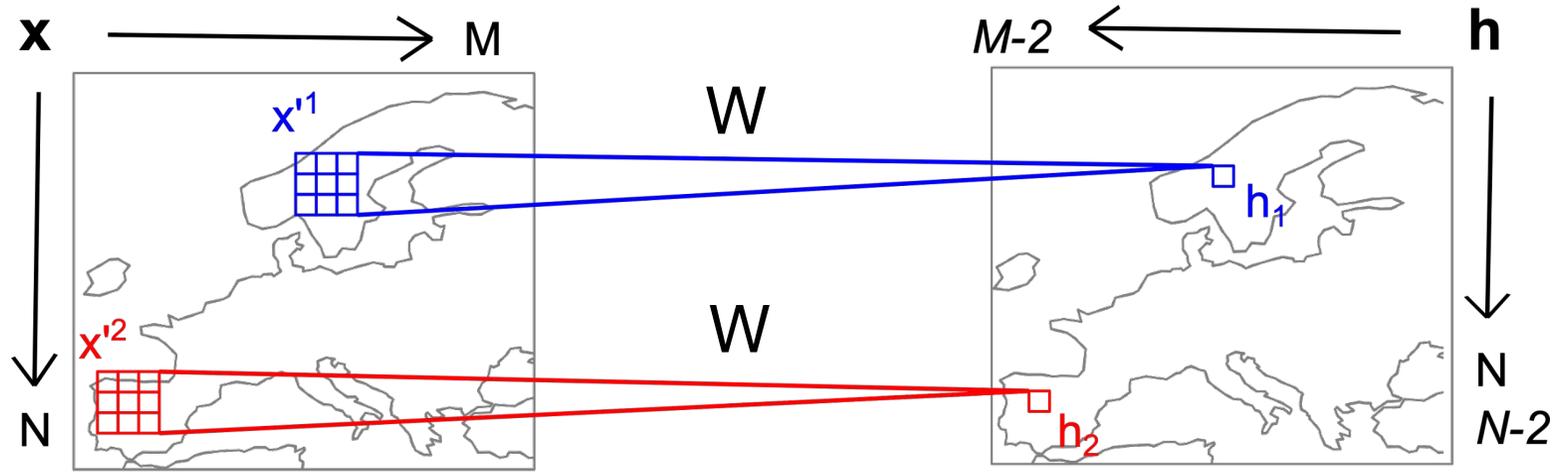
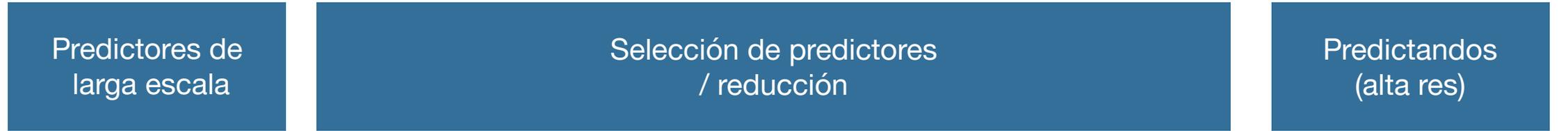
La selección de predictores se realiza durante el entrenamiento

...similar a su uso en otras disciplinas

Computer Vision Tasks



Convolutional Neural Networks (CNN)



Convolutional Layers:

Menos parámetros que en capas densas

Cada kernel representa un patrón espacial

$$h_1 = f(x'^1_1 w_1 + x'^1_2 w_2 + \dots + x'^1_9 w_9)$$

$$h_2 = f(x'^2_1 w_1 + x'^2_2 w_2 + \dots + x'^2_9 w_9)$$

Funciones de loss/error:

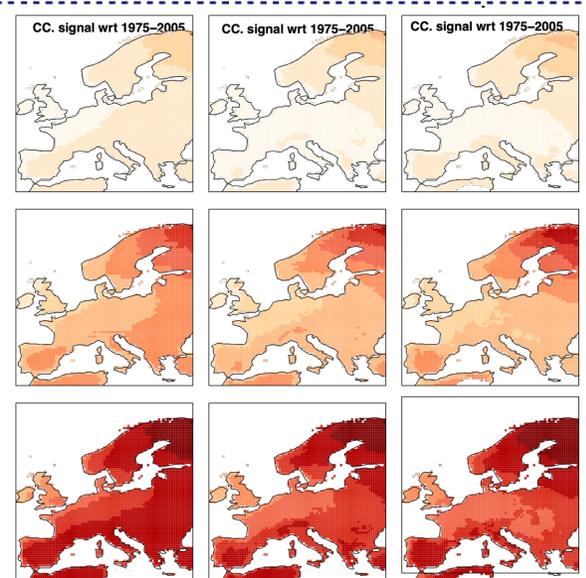
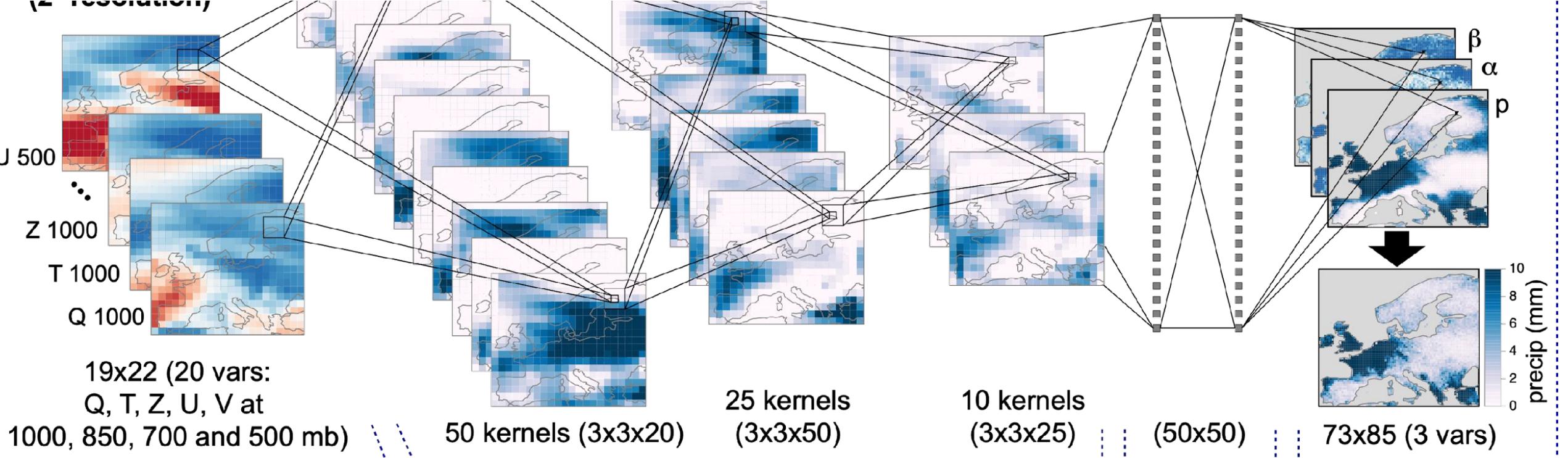
Loss: MSE.

Loss: Negative log-likelihood de una distribución paramétrica:

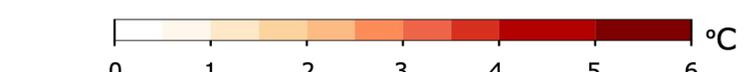
T = Gaussian

P = Bernoulli-Gamma





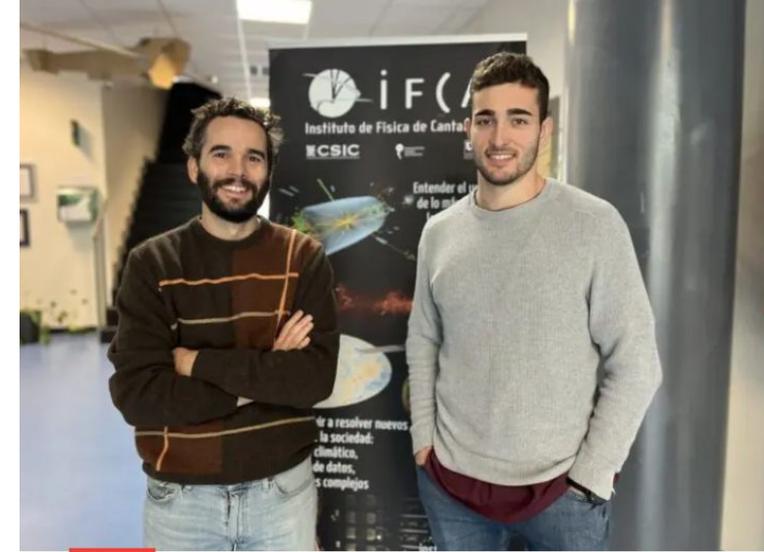
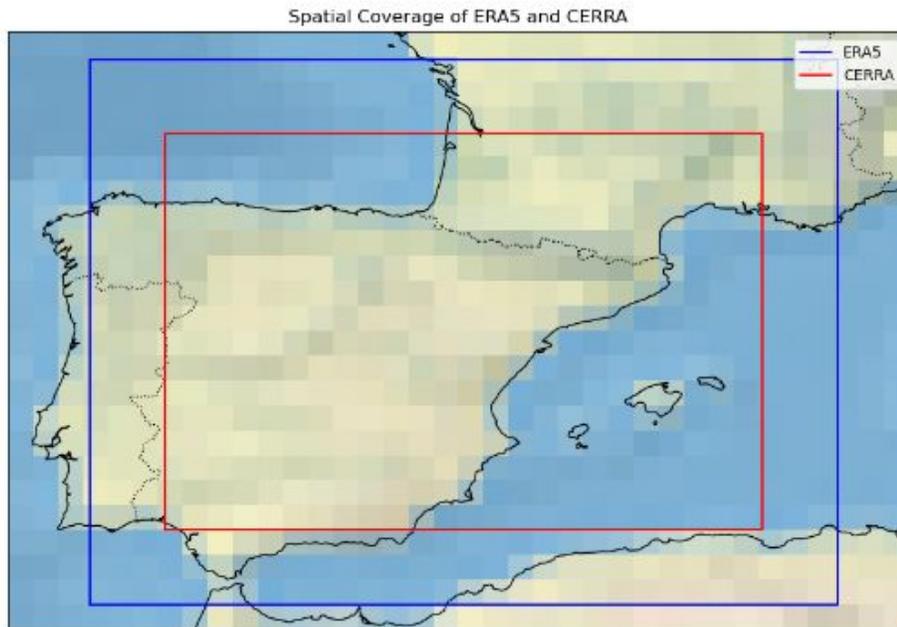
Baño-Medina et al. (2020)
<https://doi.org/10.5194/gmd-13-2109-2020>



Super resolución

CONVSWIN2SR: A 12-million-parameter vision model
<https://predictia.es/en/deep-12-million-vision-model>

Objetivo: de ERA5 (0.25°) a CERRA (0.05°)



INICIATIVA

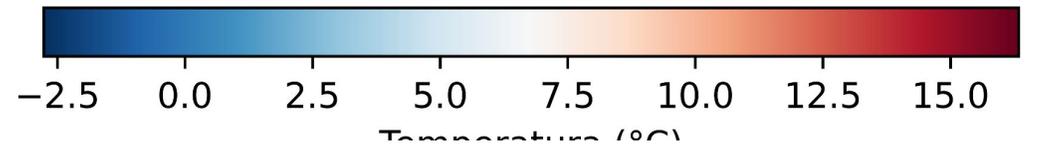
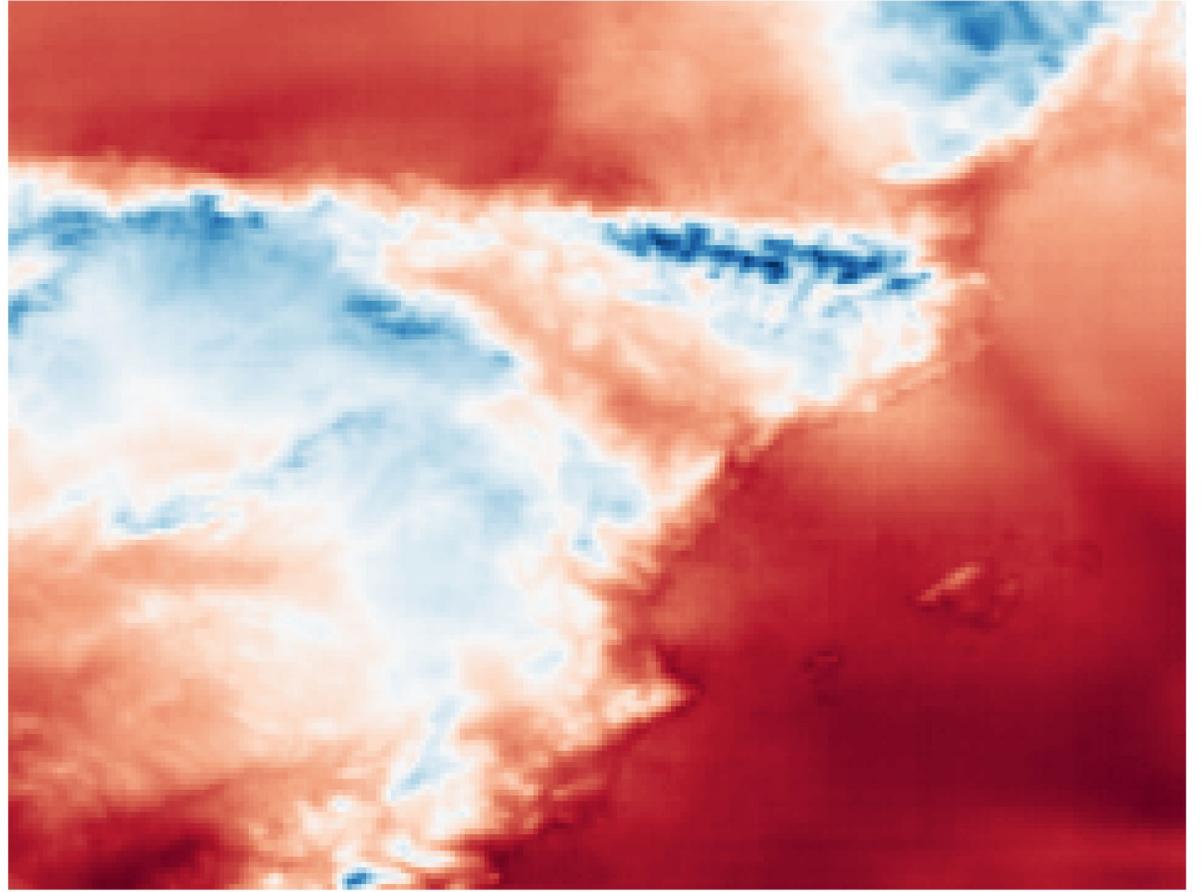
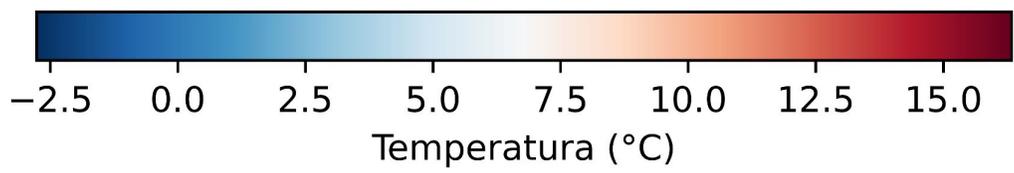
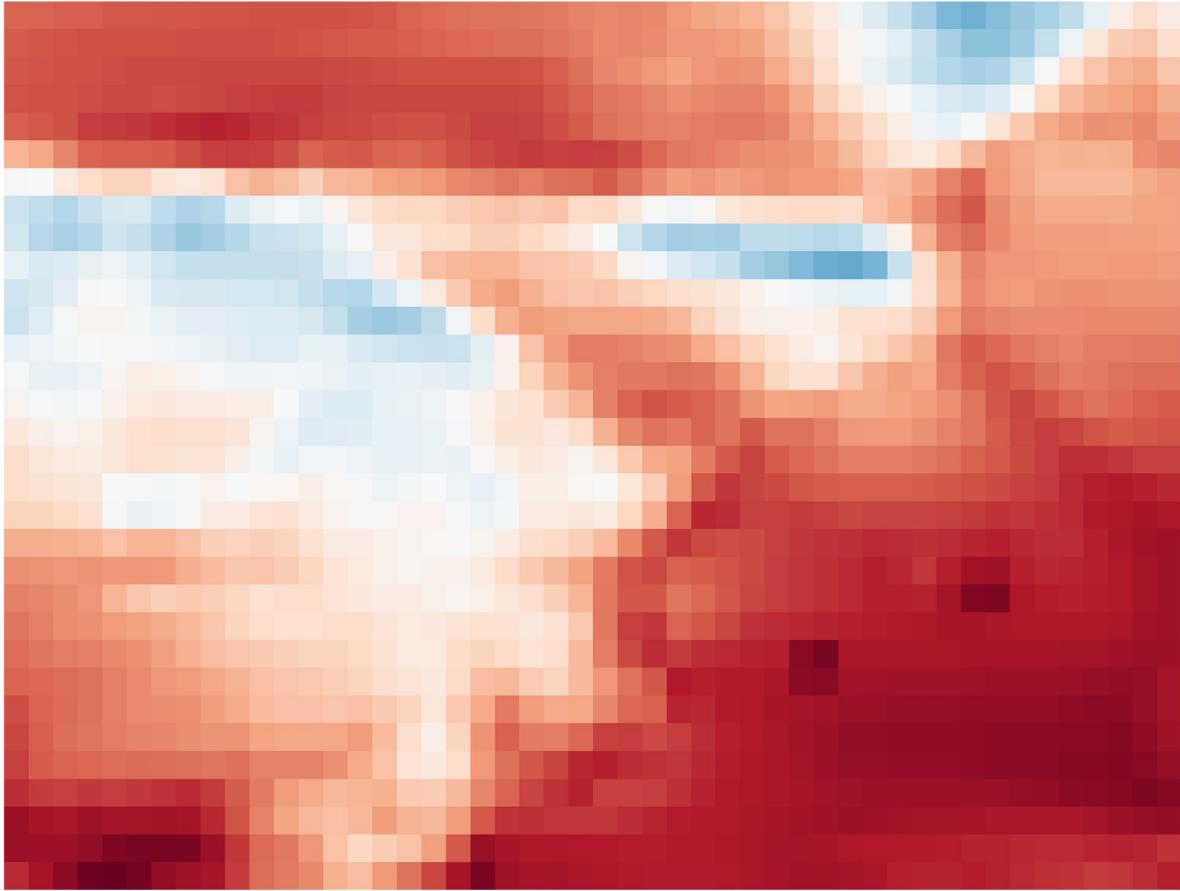
Investigadores cántabros desarrollan un modelo climático con inteligencia artificial

18/01/2024 - 11:08 am REDACCIÓN

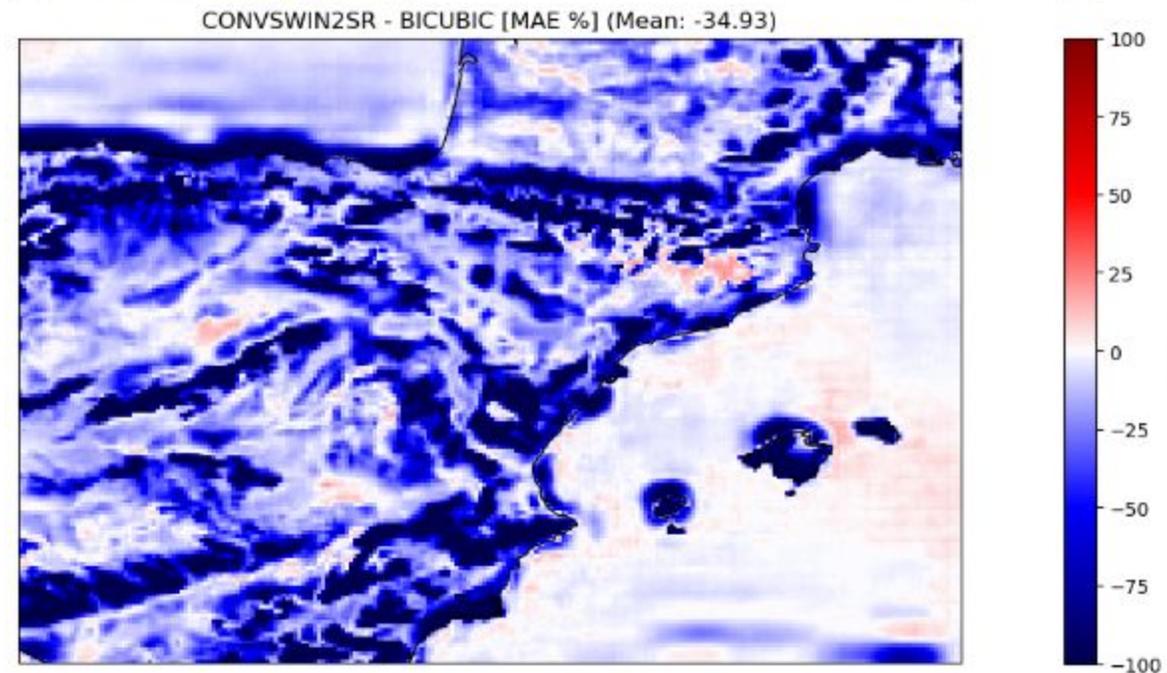
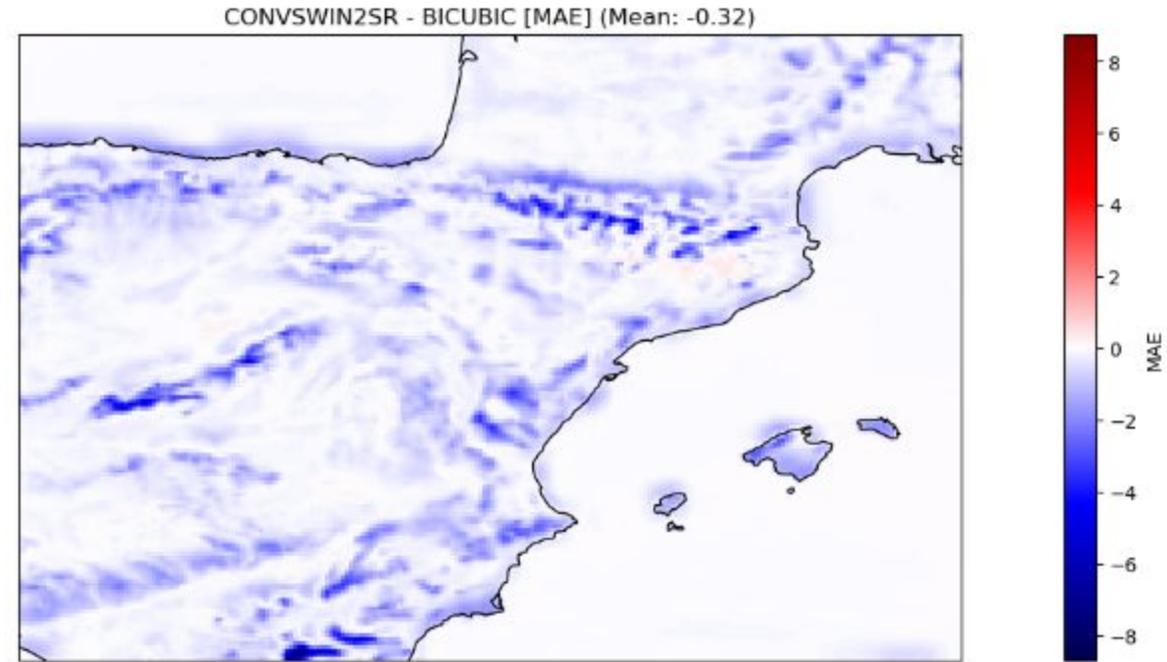
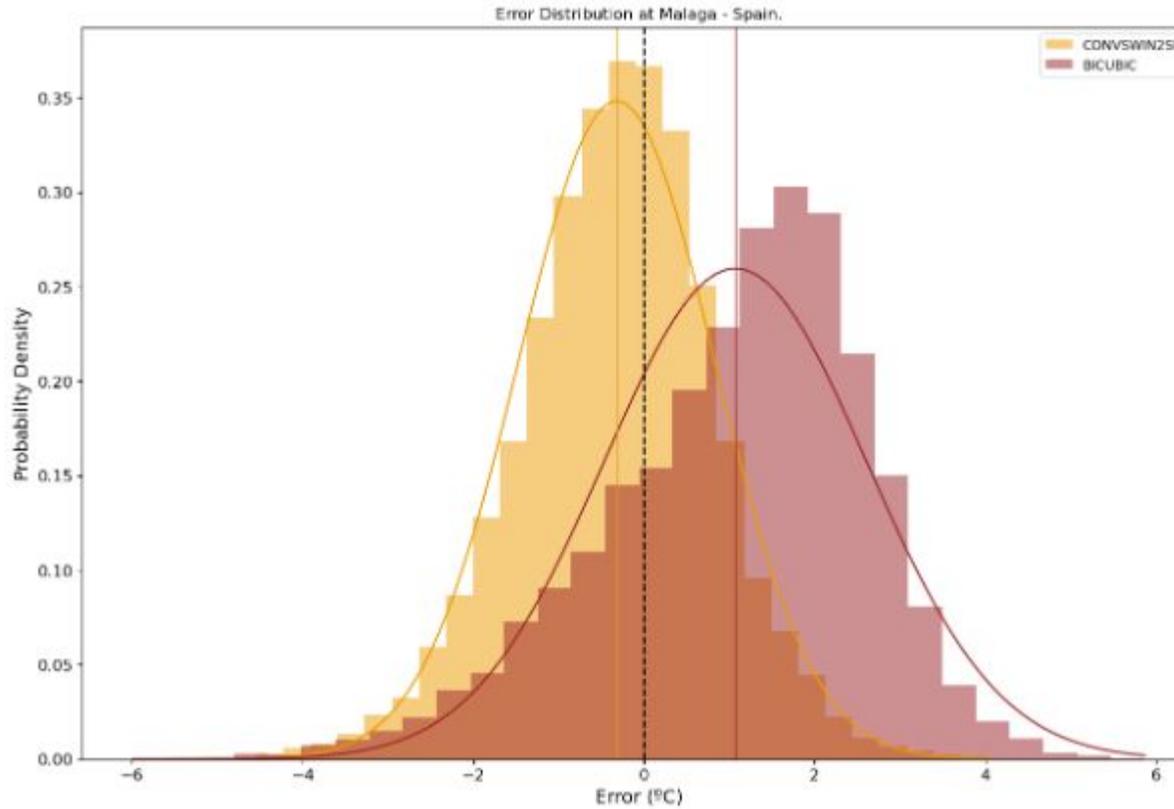
Tiempo de lectura: 4 min

Investigadores cántabros desarrollan un modelo climático con inteligencia artificial





Super resolución



Input

DeiT-III-B

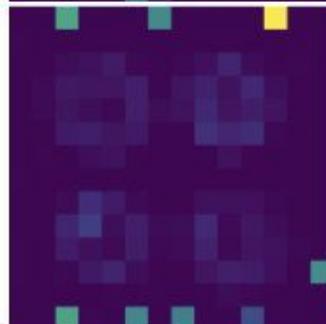
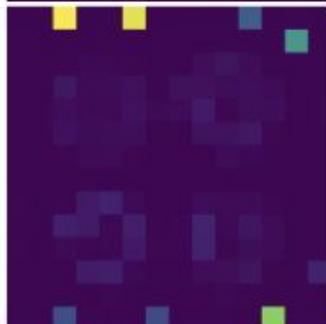
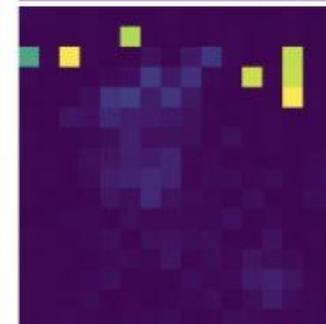
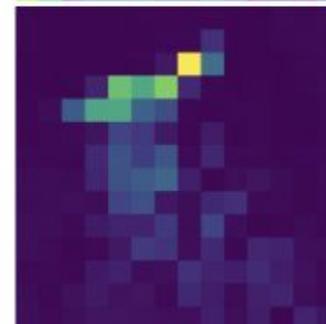
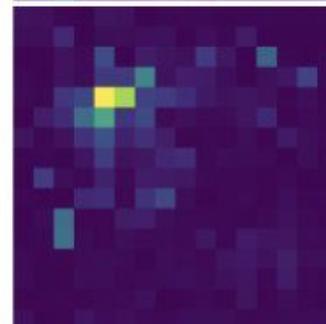
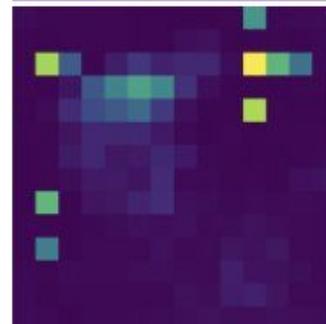
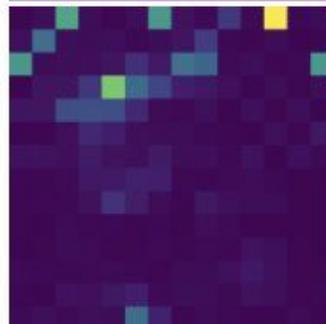
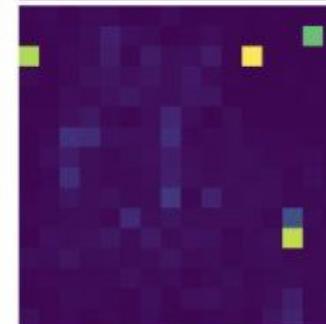
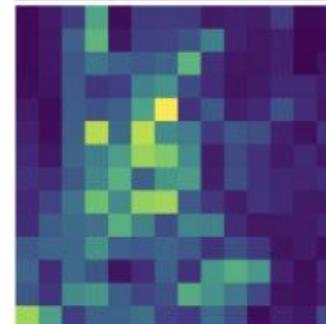
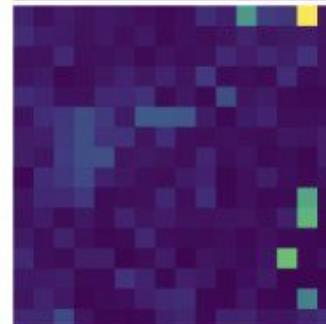
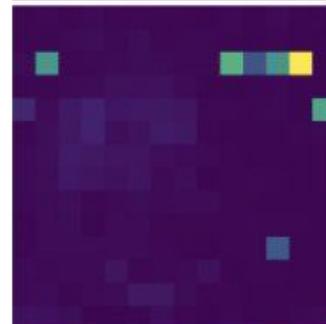
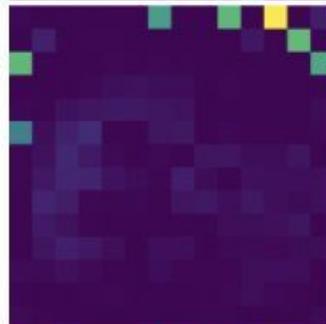
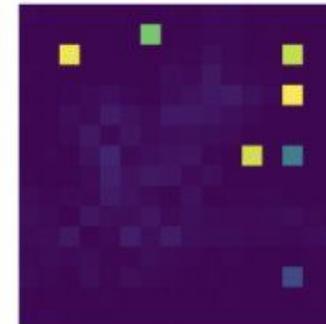
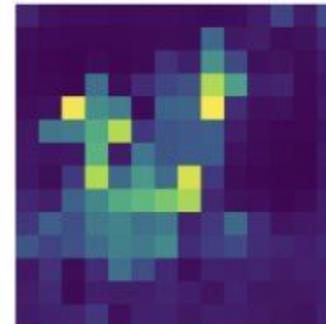
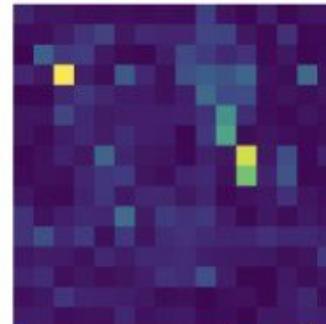
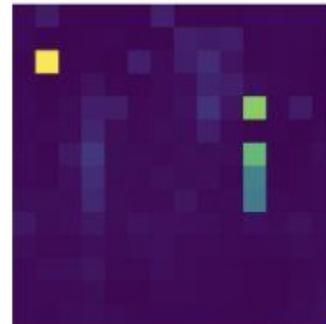
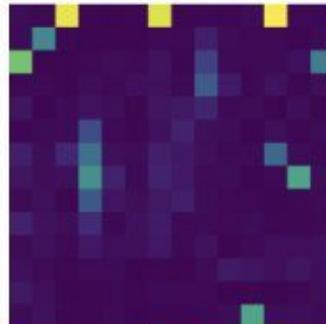
DeiT-III-L

OpenCLIP-B

OpenCLIP-L

DINO-B

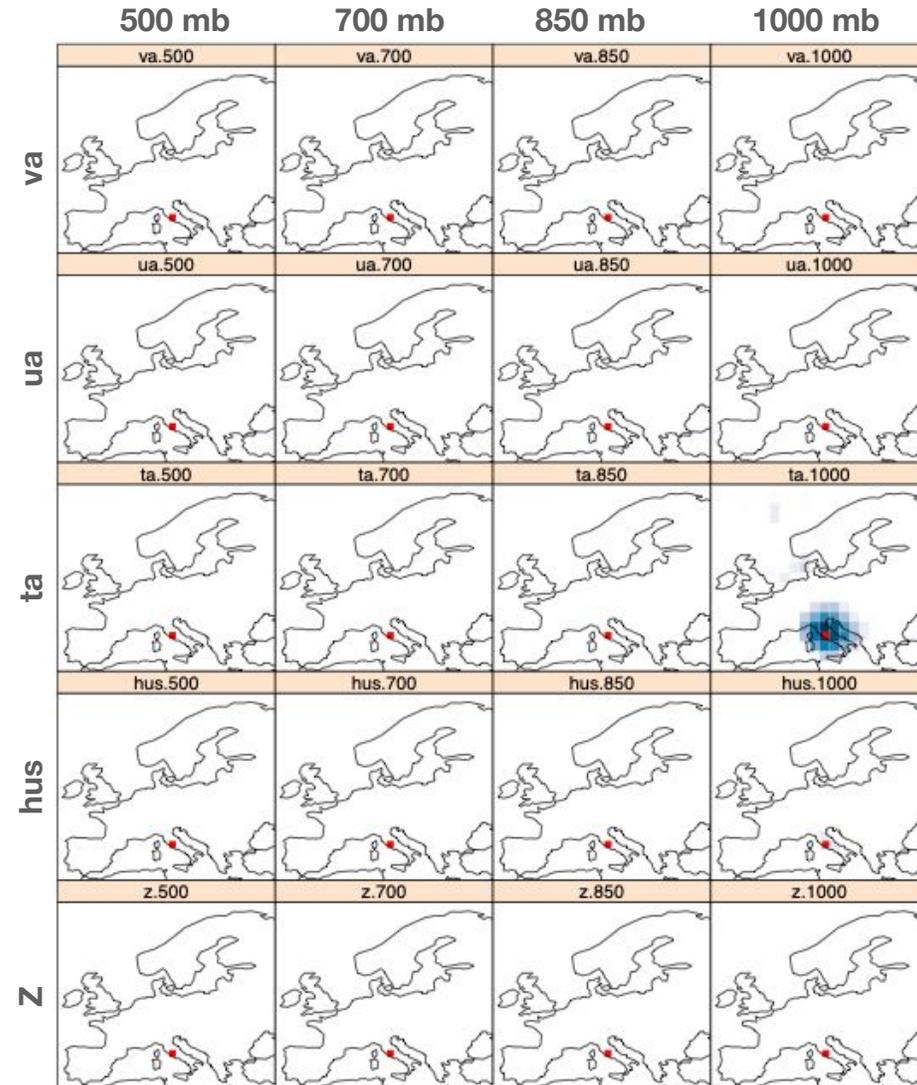
DINOv2-g



DINOv2: A Self-supervised Vision Transformer Model

Explainable Artificial Intelligence (XAI)

- Predictores
- Contexto espacial
- Saliency maps: spatial representation of the influence of the inputs in the model output
- Diagnostic y explicabilidad



TEMPERATURE

- Air temperature at 1000 hPa is the most relevant variable

- This predominance of T° at 1000 hPa is also observed for the other locations studied (see *the manuscript*).

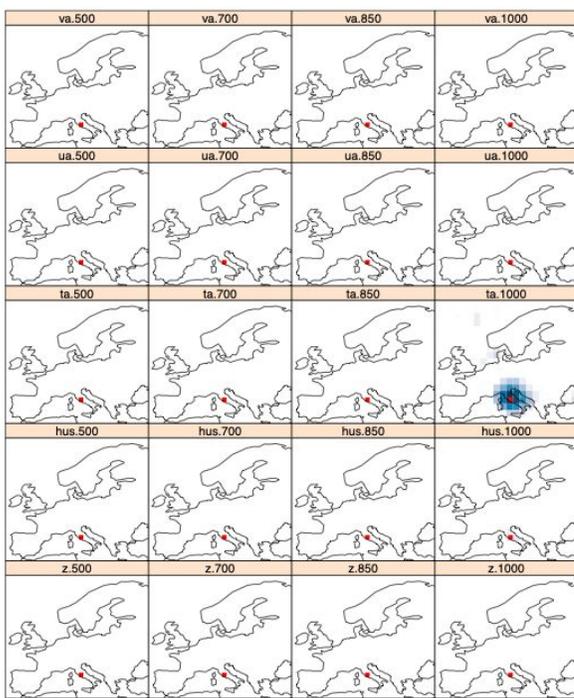
- The relevant spatial domain is located around the location of interest.

PRECIPITATION

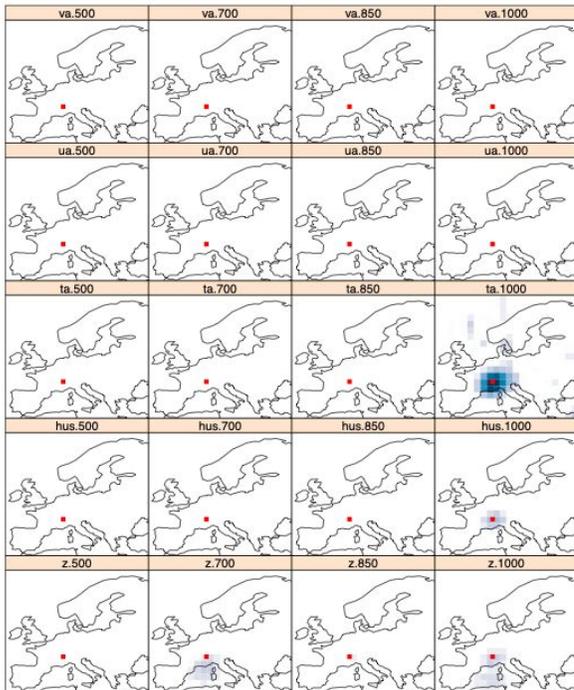
- Wind velocity, especially the meridional component at 700 and 850 hPa, together with the specific humidity and the geopotential height at 1000 hPa seem the most informative to downscale over Rome.

- For the Alps, zonal velocity is also relevant.

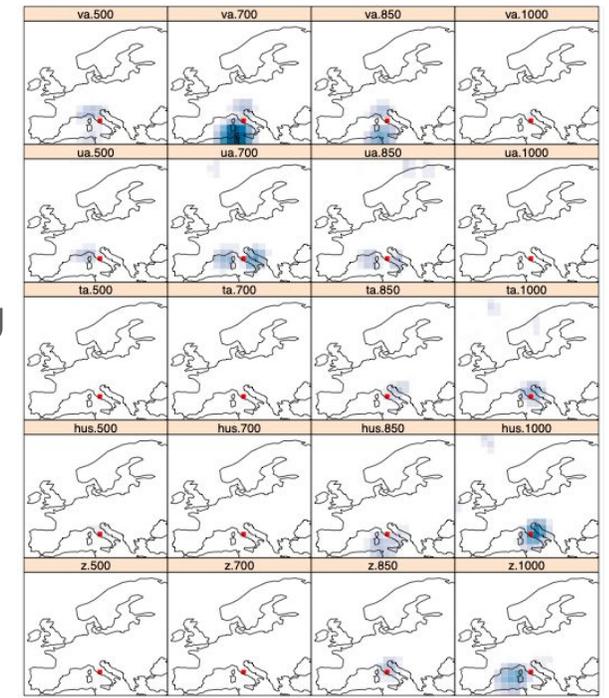
Rome



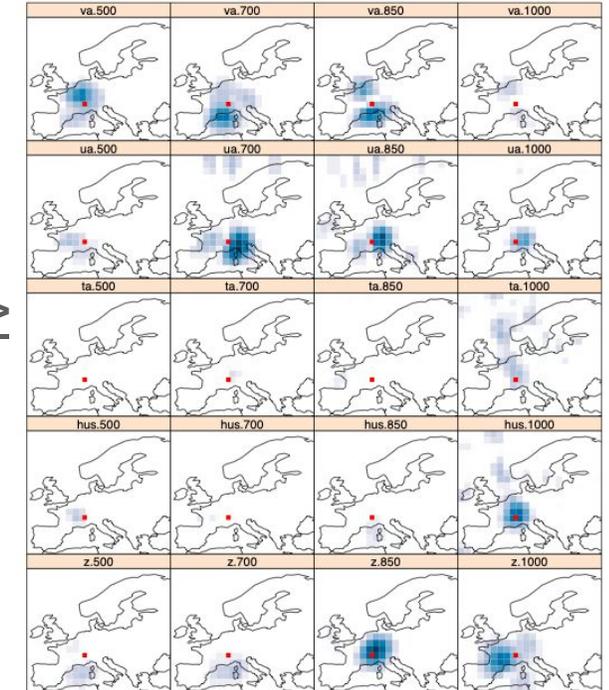
Alps



Rome



Alps



Explainable Artificial Intelligence (XAI)

Muchas **técnicas diferentes**: Gradient/Saliency, InputGradient, Integrated Gradients, Layerwise Relevance Propagation, SmoothGrad, NoiseGrad, FusionGrad...

Con diferentes medidas de **calidad**: **robustness**, **faithfulness**, **randomisation**, **localisation**, **complexity**...

Finding the right XAI method — A Guide for the Evaluation and Ranking of Explainable AI Methods in Climate Science. 2023. Bommer, P, et al.

<https://arxiv.org/pdf/2303.00652.pdf>

