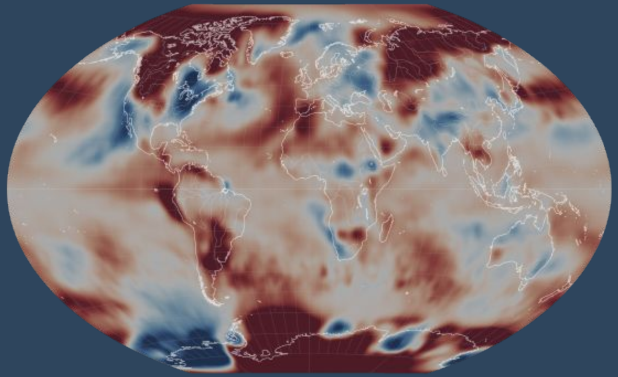
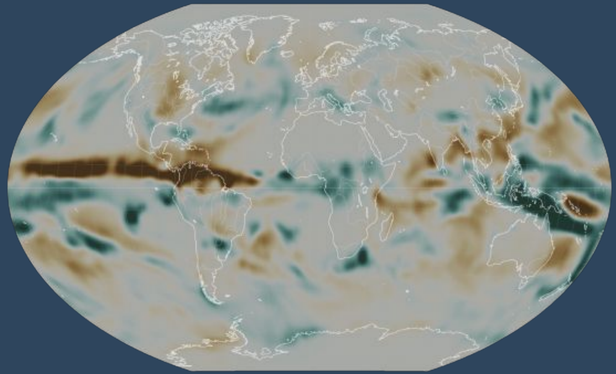
H2020 CLINT - Challenges in developing an AI-Enhanced operational demonstrator of sub-seasonal and seasonal forecasting of detection and attribution of heatwaves and warm nights



H. Loukos - the climate data factory



Using ECMWF Forecasts (UEF2023)
05-08 June 2023 - Reading

the climate data factory

Service & data provider

Climate data management & processing

H2020 CLINT
Climate Intelligence

AI/ML framework
Detection/Causation/Attribution
Extreme Events + impacts
S2S + Projections
15 partners
2021-2025

CLINT CLIMATE INTELLIGENCE

CLINT CLIMATE INTELLIGENCE

OUR MISSION

The **main objective** of CLINT is the development of an **Artificial Intelligence framework** composed of Machine Learning techniques and algorithms to process **big climate datasets** for improving **Climate Science** in the **detection, causation, and attribution of Extreme Events** (EEs), namely tropical cyclones, heatwaves and warm nights, droughts, and floods. The CLINT AI framework will also cover the quantification of the EE impacts on a variety of socio-economic sectors under historical, forecasted, and projected climate conditions, and across different spatial scales (from European to local), ultimately developing innovative and sectorial **AI-enhanced Climate Services**. Finally, these services will be operationalized into Web Processing Services, according to the most advanced open data and software standards by **Climate Services Information Systems,** and into a **Demonstrator** to facilitate the uptake of project results by public and private entities for research and Climate Services development.

Read more about the project

EXTREME EVENTS

HOTSPOTS

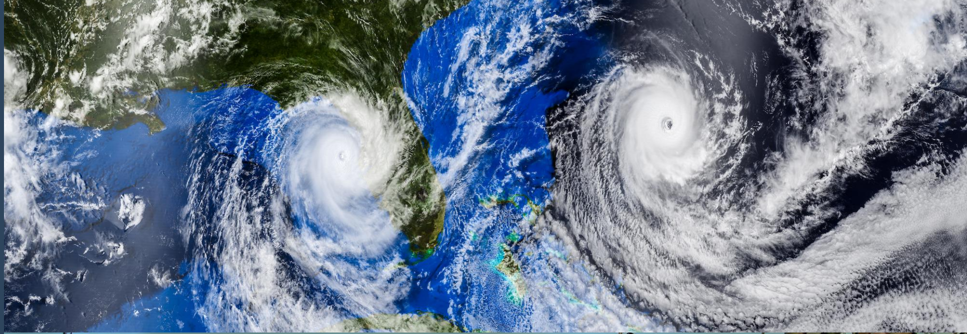
MACHINE LEARNING

EXTREME EVENTS

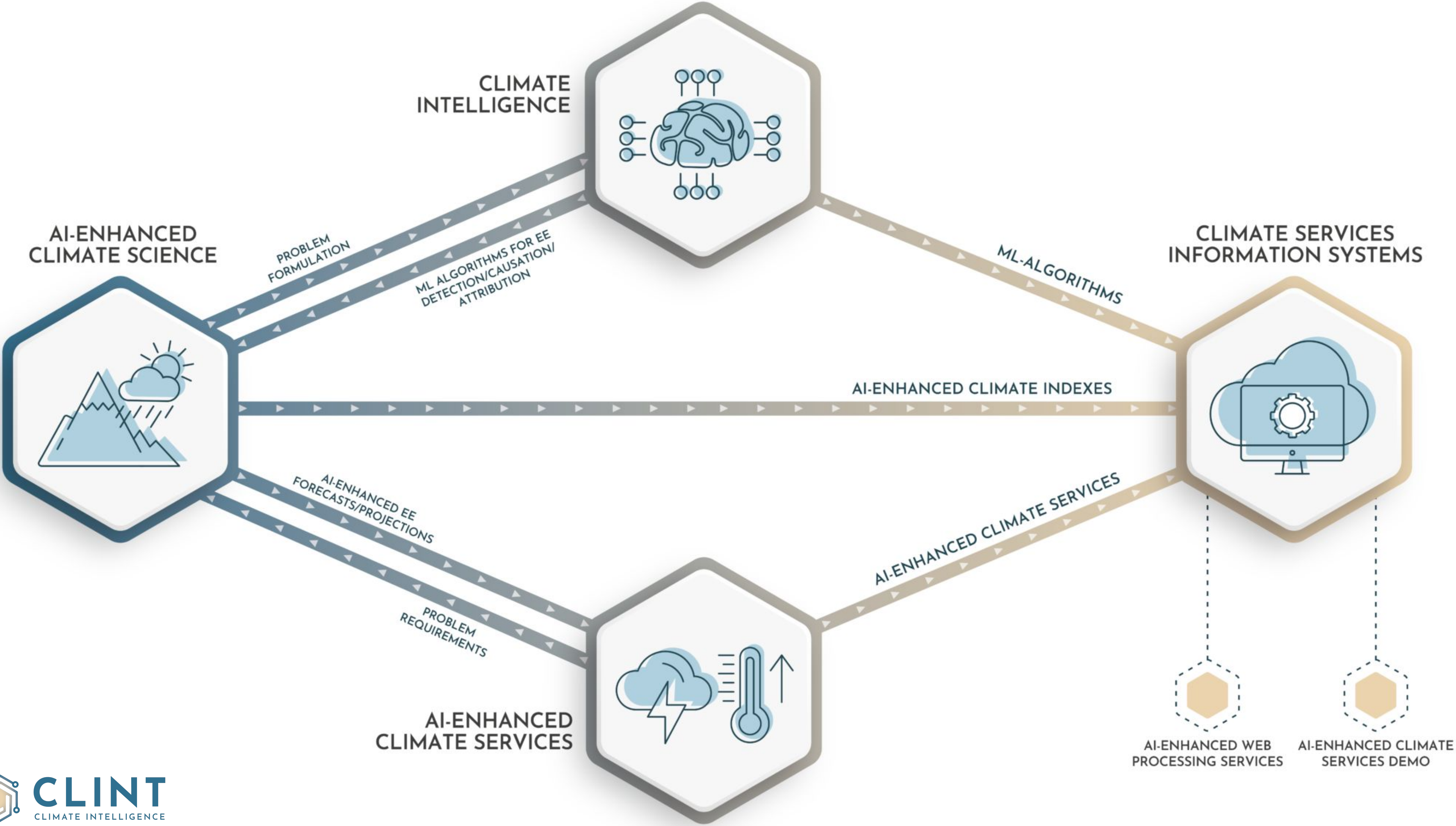
Tropical Cyclones
Heatwaves and Warm Nights
Extreme Droughts
Compound and Concurrent EE

AND THEIR IMPACTS

Water-Food-Energy Nexus



CLINT
CLIMATE INTELLIGENCE



CLINT
CLIMATE INTELLIGENCE


Semiarid Region
Douro basin
River deltas
Rhine delta
Snow-dependent basins
Lake Como basin
Semiarid Region
Zambezi watercourse
CLINT
CLIMATE INTELLIGENCE

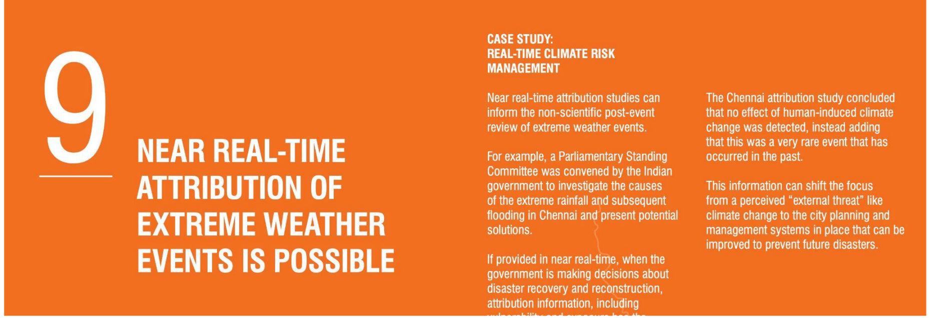
Near real-time attribution of EE events is possible

Raising Risk Awareness Project 2017

April 2023 Spain/Morocco HW

Event: April 26-28

Attribution: May 5



9

CASE STUDY:
REAL-TIME CLIMATE RISK MANAGEMENT

NEAR REAL-TIME ATTRIBUTION OF EXTREME WEATHER EVENTS IS POSSIBLE

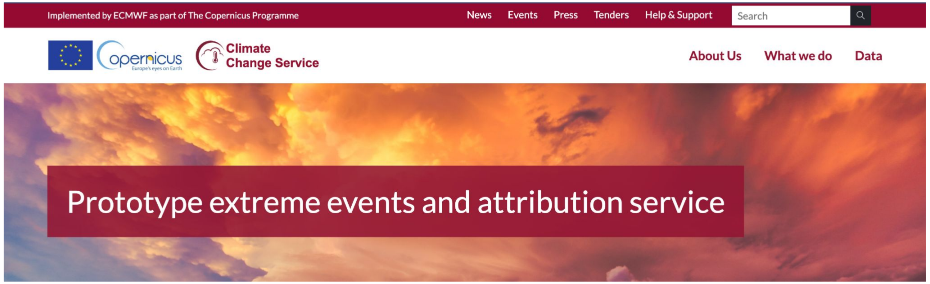
Near real-time attribution studies can inform the non-scientific post-event review of extreme weather events.

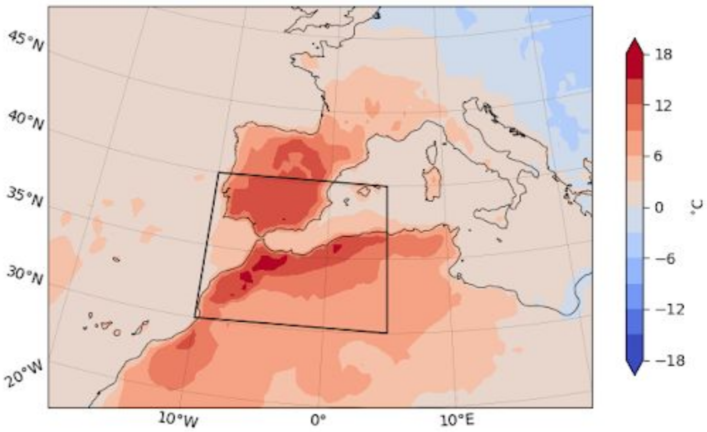
For example, a Parliamentary Standing Committee was convened by the Indian government to investigate the causes of the extreme rainfall and subsequent flooding in Chennai and present potential solutions.

If provided in near real-time, when the government is making decisions about disaster recovery and reconstruction, attribution information, including

The Chennai attribution study concluded that no effect of human-induced climate change was detected, instead adding that this was a very rare event that has occurred in the past.

This information can shift the focus from a perceived "external threat" like climate change to the city planning and management systems in place that can be improved to prevent future disasters.





World Weather Attribution

Challenges in developing an AI-Enhanced operational demonstrator of sub-seasonal and seasonal forecasting of detection and attribution of heatwaves and warm nights

Scientific - Technical - Computational - (Communicational)

Scientific

Improve Detection
Improve Forecast
Improve Attribution
of Heatwaves/Warm nights

- Define & Detect H/WN
  - Indices: many definitions of Heatwaves
  - Drivers (causation & detection)

- ML Forecast framework
  - Data driven/forecast postprocessing/hybride.
  - Seamless S2S ?

- Attribution
  - Existing Frameworks/Methods
    eg Oldenborgh et al. 2021, Leach et al. 2021.
  - Not many references using ML
  - Which ones are applicable to forecast?

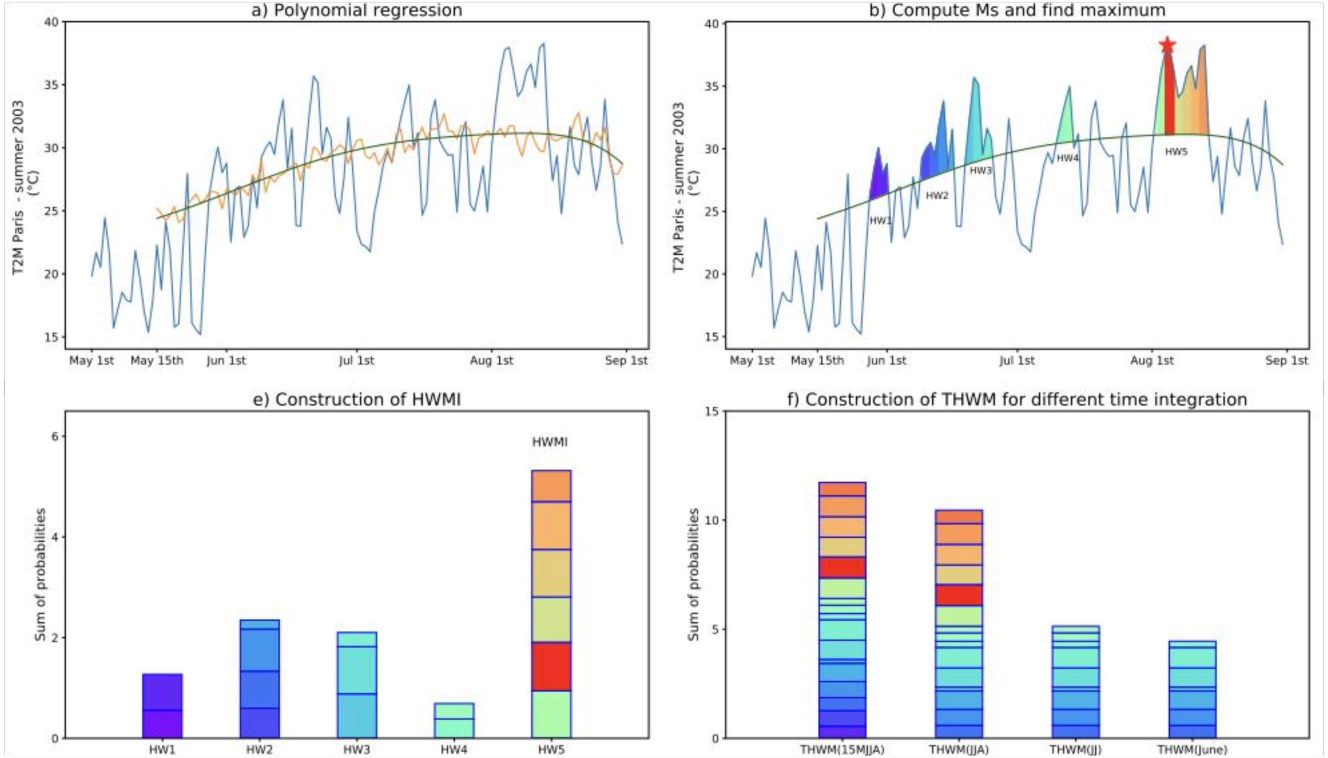
van Oldenborgh et al. 2021 Pathways and pitfalls in extreme event attribution.

Leach et al, 2021. Forecast-based attribution of a winter heat wave within the limit of predictability.

- Define, Detect and Forecast

Heat Wave Magnitude Index (HWMI)

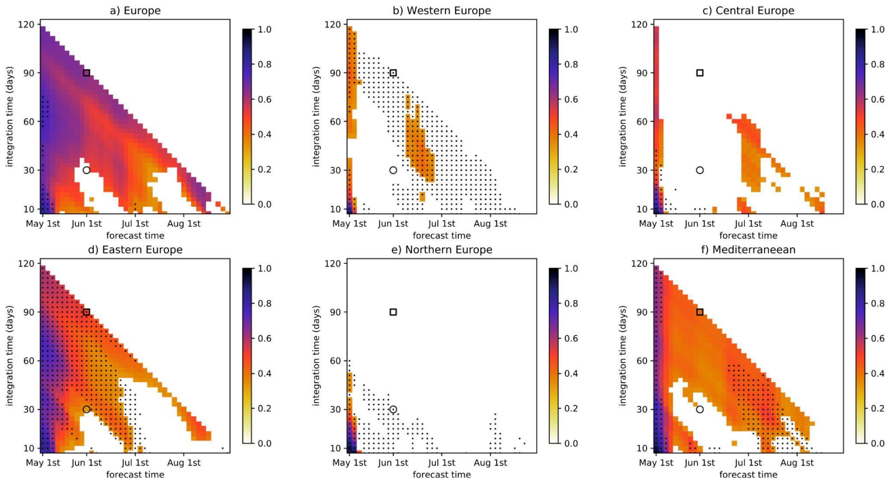
Summer heatwave **propensity** (THWMI)



Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., ... & García-Serrano, J. (2022). Seasonal prediction of European summer heatwaves. Climate Dynamics, 58(7), 2149-2166.

Forecast skill

"Useful information at the regional scale up to two months ahead for THWM integrated over one, two or three months for the whole European domain, the Mediterranean region and Eastern Europe."



Prodhomme, C., Materia, S., Ardilouze, C., White, R. H., Batté, L., Guemas, V., ... & García-Serrano, J. (2022). Seasonal prediction of European summer heatwaves. Climate Dynamics, 58(7), 2149-2166.

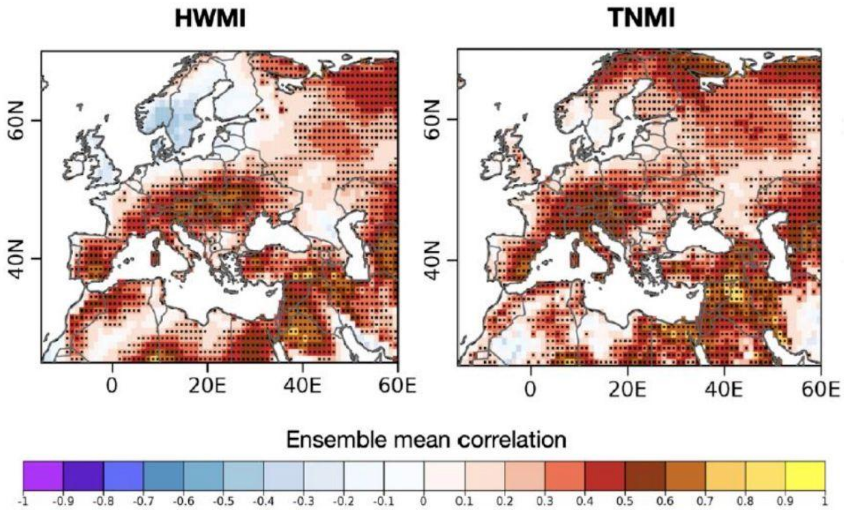
HW & TN DETECTION: INDICES & FORECAST SKILL

Key achievement of first phase -> analysis of multi-model seasonal forecast skill of TNs, to complement existing literature on HWs (e.g. Prodhomme et al, 2021):

- Torralba V., S. Materia, L. Cavicchia, C. Prodhomme, M. C. Álvarez-Castro, E. Scoccimarro, and S.Gualdi. *Seasonal forecast skill of European warm nights. In preparation.*

Providing a benchmark for ML-based techniques
Identify where ML-enhancements are most needed (e.g. Scandinavia, parts of central Mediterranean)

MS90 & D3.1 (2022)
Definition of indices, seasonal forecast skill and analysis of indivdual events (e.g. summer of 2022).



Skill (correlation score) of the CMCC SPS3.5 seasonal prediction system for Heat Wave Magnitude index (HWMI) and Tropical Night Magnitude index (TNMI).
Period: 1993-2016. Start date: 1st of May. Benchmark: ERA5

Torralba V. et al. Seasonal forecast skill of European warm nights. et al. EGU23 (2023)

Zaninelli P.G., Barriopedro D., Drouard M., Garrido-Pérez J.M., Pérez-Aracil J., Fister D., García-Herrera R., Salcedo-Sanz S., Alvarez-Castro M.C.: Deep learning techniques applied to an attribution study for heatwaves in the Iberian Peninsula. General Assembly of European Geosciences Union (EGU) 2021. 23-28 abril 2023, Viena, Austria.

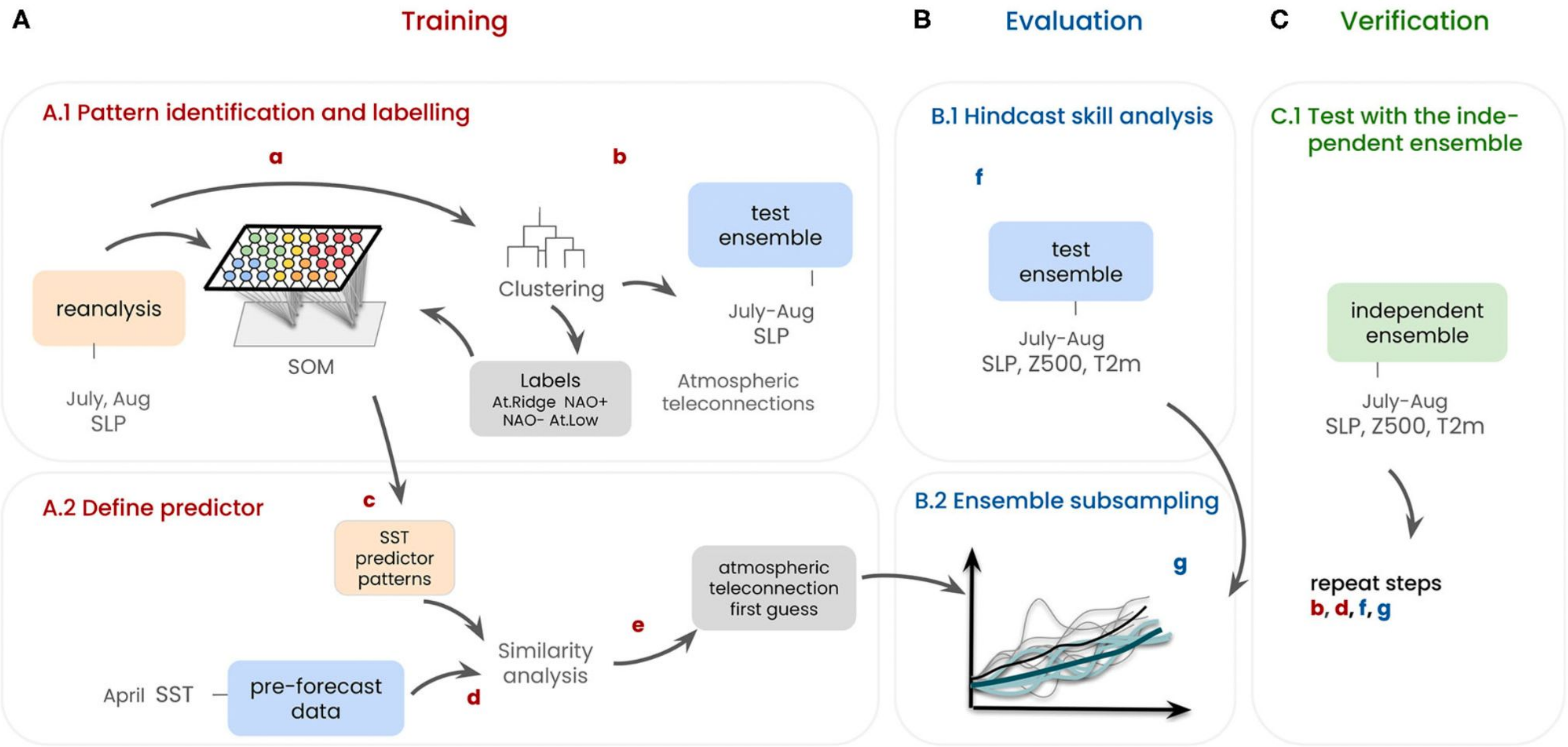
Drouard M., Pérez-Aracil J., Barriopedro D., Zaninelli P.G., Garrido-Pérez J.M., Fister D., Salcedo-Sanz S., García-Herrera R.: S2S prediction of summer heatwaves in the Iberian Peninsula using convolutional networks. General Assembly of European Geosciences Union (EGU) 2021. 23-28 abril 2023, Vienna, Austria

Alvarez-Castro M. C., Torralba V., L. Cavicchia, E. Scoccimarro, and S.Gualdi. Predictability of the 2022 extreme summer at seasonal scale **(In preparation)**

Torralba V., S. Materia, L. Cavicchia, C. Prodhomme, M. C. Álvarez-Castro, E. Scoccimarro, and S.Gualdi. *Seasonal forecast skill of European warm nights.* **(In preparation)**

CLINT
CLIMATE INTELLIGENCE

- Forecast postprocessing (ML member picking)



Set 1: Training & Evaluation (1902-2008) Set 2 Verification (1980-2016)

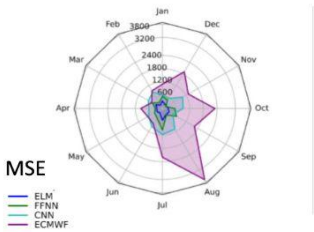
Carvalho-Oliveira, J., Borchert, L. F., Zorita, E., & Baehr, J. (2022). Self-organizing maps identify windows of opportunity for seasonal European summer predictions. Frontiers in Climate, 4.

● Data driven ML forecasting

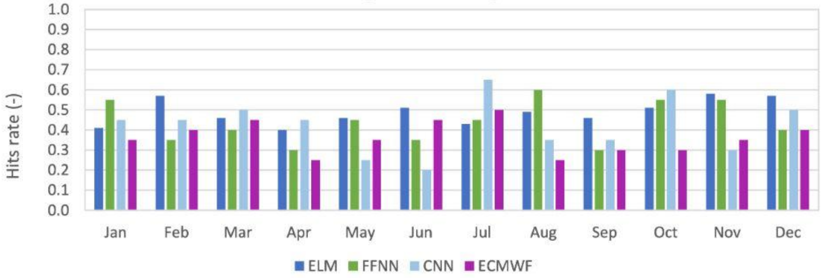
TASK 2.4: MACHINE LEARNING FOR EXTREME EVENTS FORECASTING

- Sub-seasonal drought forecasting via machine learning to leverage climate data at different spatial scales
  - Forecast total precipitation for the next 30-days and then compute SPI
  - Three frameworks developed in Rijnland: Niño Index Phase Analysis (NIPA) + Extreme Learning Machine (ELM), Feed Forward Neural Network (FFNN), Convolutional Neural Network (CNN)

Precipitation forecasting



Drought forecasting



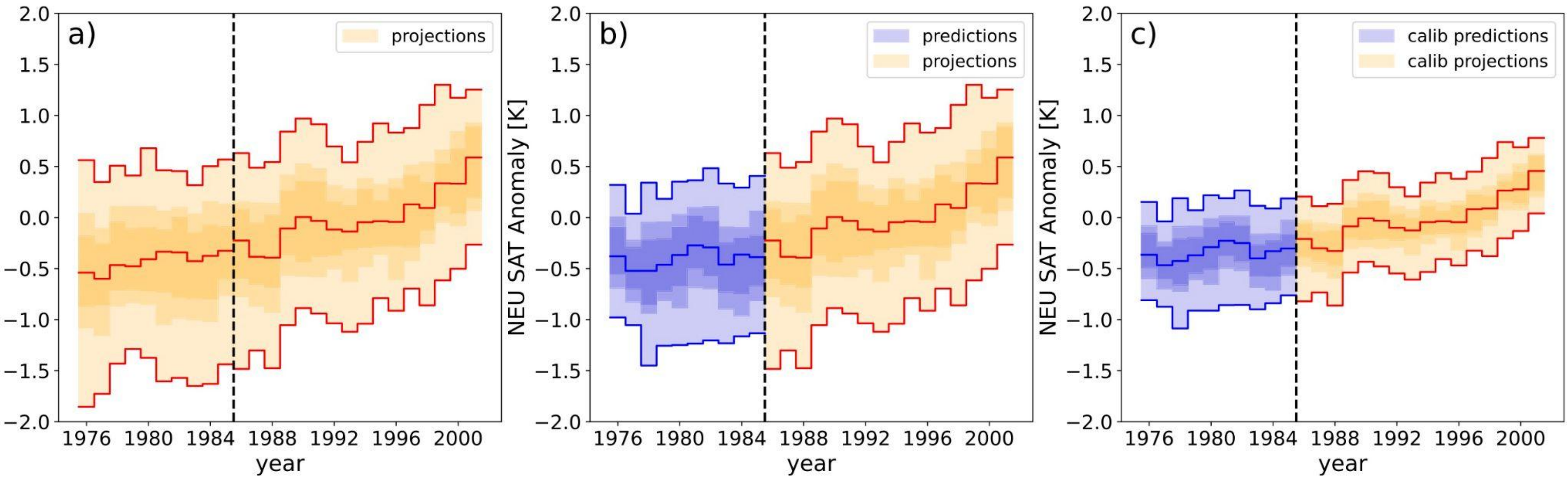
Results:

- ELM models forecast precipitation better than ECMWF benchmark for every month
- All the models outperform ECMWF forecasts in terms of SPI
- There is not an only ML model forecasting better SPI throughout all the year

- Other recently started works:
  - Forecasting tropical cyclones probability in the next 5 days with CNN
  - Forecasting daily rainfall for the next 30 days with LSTM in Rijnland

CLINT
CLIMATE INTELLIGENCE

- To be seamless or to to be ?

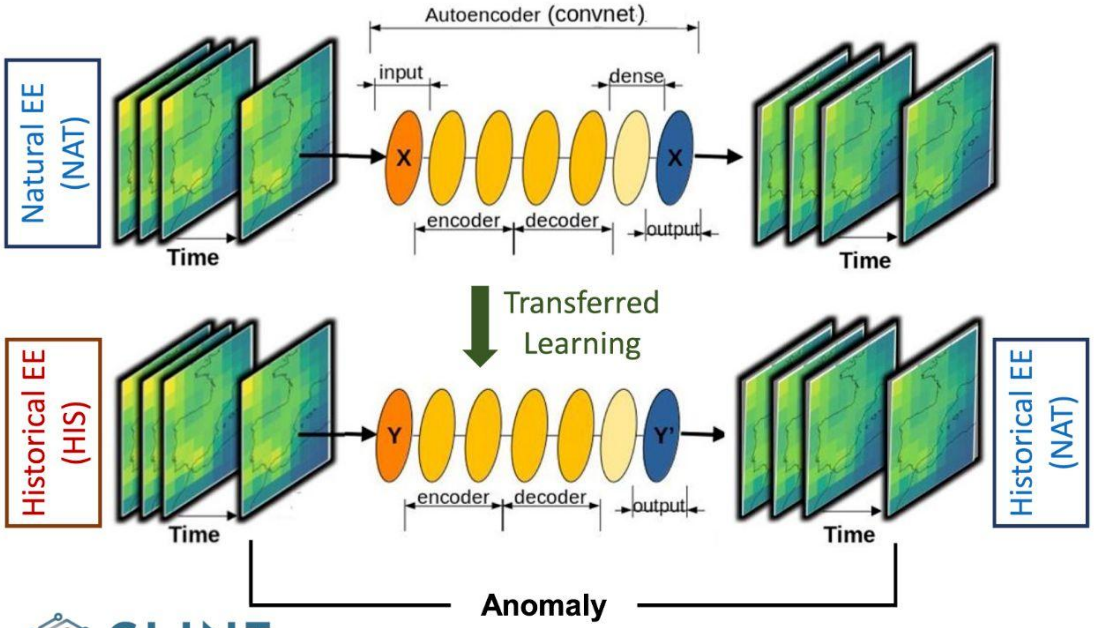


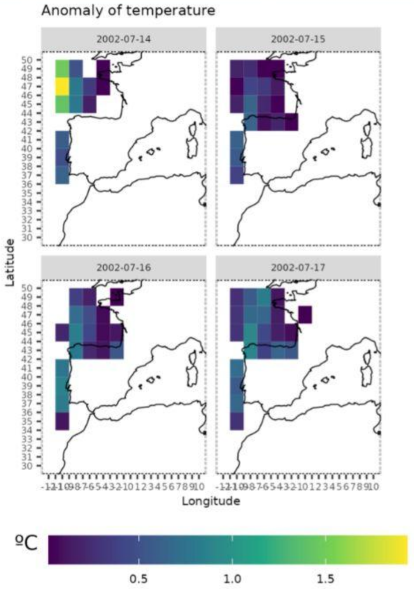
From Befort, D. J., Brunner, L., Borchert, L. F., O'reilly, C. H., Mignot, J., Ballinger, A. P., ... & Weisheimer, A. (2022). Combination of decadal predictions and climate projections in time: Challenges and potential solutions. Geophysical Research Letters, 49(15), e2022GL098568..

TASK 5.1: ATTRIBUTION OF EXTREME EVENTS

AI Developments

- ML algorithms trained in natural world and applied to present-day climates





Anomaly detection for a 4-day heatwave: temp. anomaly exceedance above a threshold that would have never been experienced in a natural world (ºC)

CLINT
CLIMATE INTELLIGENCE

Drouard et al. (2023). *EGU23-5742*
Fister et al. (2023). *EGU23-7457*
Zaninelli et al. (2023). *EGU23-6732*

Technical

Data availability
Data handling and processing
Coherency

- Data availability
  - Exploit large reanalysis datasets
    - ERA5 (1950-now, 0.25°, Hourly)
    - ERA20C (1900-2010, 1°, 3-Hourly)
    - 20CrV3 (1836-2015, 1°, 3-Hourly)
  - Exploit large S2S datasets
    - C3S (seasonal), S2S project
    - ML benchmarks datasets ?
    - But hindcast period is 20y (need more)
- Data handling and processing
  - Burst & Lagued forecast members (complicates processing & intercomparison)
  - Object storage: NCZarr (NetCDF-c+zarr)
  - CDO libraries optimisation/validation
- Coherency
  - Daily data calibration (Quantile Mapping)
  - Forecast calibration/Indicator definition
  - Between forecasted indicators (if ML training specific to each one)

Description of the C3S seasonal multi-system

This table shows the centres that provide data to this project together with the **latest** configuration of their systems. Follow the link of each Data Provider for specific model description.

Status on 13 Feb 2022	Time range (forecasts and hindcasts)	Resolution of model	Forecast initial conditions	Forecast ensemble size	Hindcast initial conditions	Hindcasts ensemble size	Hindcast period	Hindcast production schedule
ECMWF (ecmf)	215 days	T_{CO}319/L91 Dynamics:T_{CO}319 cubic octahedral grid Physics: O320 Gaussian grid (36 km) 91 levels in vertical, to 0.01hPa (80km)	1st of month 🙂	51 members	1st of month	25	1981-2016	fixed
UKMO (egrr)	215 days	N216/L85 0.83° x 0.56° (~ 60km in mid-latitudes) 85 levels in vertical, to 85km	each day of month 😐	2 members/day[4]	1st, 9th, 17th, 25th of month	7 members/start time	1993-2016	on-the-fly[1]
Météo-France[3] (lfpw)	7 calendar months	TL359/L137 (0.5°) 137 levels in vertical, to 0.01hPa	last and penultimate Thursday of previous month 1st of month 🙂	25 members each 1 member	last and penultimate Thursday of previous month 1st of month	12 members each 1 member	1993-2018	fixed
DWD (edzw)	6 calendar months	T127 (~100 km) 95 levels in vertical, to 0.01hPa	1st of month 🙂	50 members	1st of month	30 members	1993-2019	fixed
CMCC (cmcc)	6 calendar months	approx 0.5° lat-long 46 levels in vertical, to 0.2hPa	1st of month 🙂	50 members	1st of month	40 members	1993-2016	fixed
NCEP (kwbc)	215 days	T128/L64 (~ 1°) 64 levels in vertical, to 0.02hPa	each day of month 😐 members initialised every 6 [illegible]: 0h, 6h, 12h and 18h UTC)	4 members/day	every 5 days [5] members initialised every 6 hours (at 0h, 6h, 12h and 18h UTC)	4 members/start date	1993-2016	fixed
JMA (rjtd)	215 days	TL319 (approx. 55km) 100 levels in vertical, to 0.01hPa	every day of month 😐	5 members/day	2 start dates lagged by 15 days [6]	5 members/start date	1993-2016	fixed
ECCC (cwao) [7] CanCM4i (component of CanSIPSv2.1)	214 days	T63 (~2.8° lat-long) 35 levels in vertical, to 1hPa	1st of the month 🙂	10 members	1st of the month	10 members	1993-2020	fixed
ECCC (cwao) [7] GEM5-NEMO (component of CanSIPSv2.1)	214 days	~1.1° lat-long (~110 km) 85 levels in vertical, to 0.1hPa	1st of the month 🙂	10 members	1st of the month	10 members	1993-2020	fixed

BA with Quantile Mapping

Advantages: suitable to daily & global

CDFt : more adapted to "unseen" future distributions

Bias-adjustment

Global with ERA5 at 1°

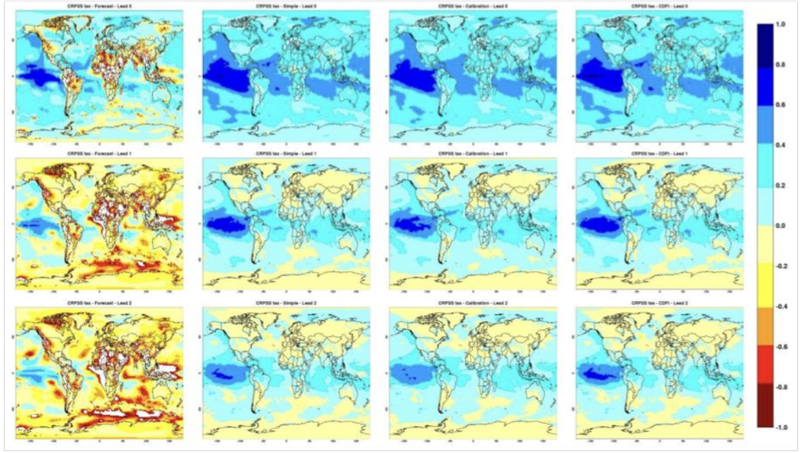
Downscaling

ERA5 (0.25°), ERA5-Land (0.1°)

Regional for now

"Our overall recommendation would be the use of versatile, easy to implement BA methods for those cases for which the use of MOS and PP methods cannot be carefully tested by experts."

Manzanas, R., et al. Statistical adjustment, calibration and downscaling of seasonal forecasts: a case-study for Southeast Asia. Clim Dyn 54, 2869–2882 (2020).

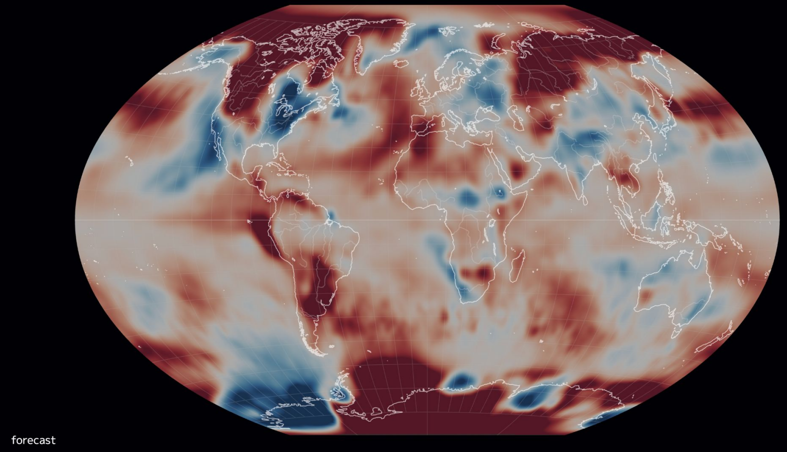


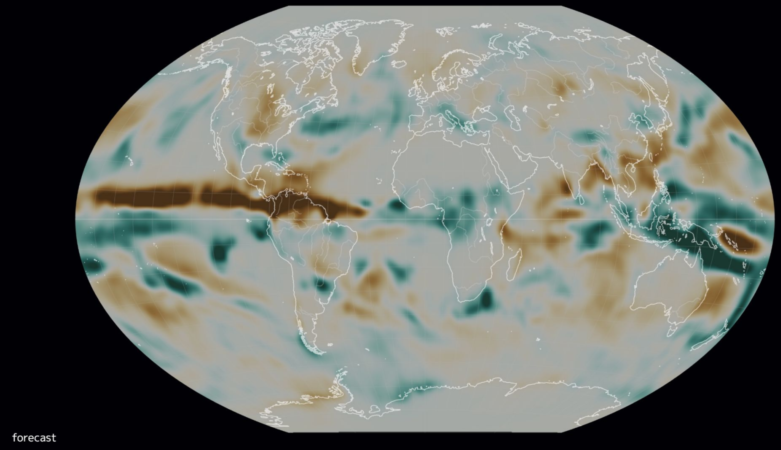
Skill assessment and comparison of methods for sub-seasonal and seasonal forecast systems for the energy sector. Deliverable D4.4, H2020 S2S4E (2017).

4-system Multimodel

Unit		Multiple
1 system		x 4
1 variable		x 4
	x 16	
2.2 Go		35 Go

7.5h processing time for a “Unit”
(1°x1°, 50 members, 180 days)
2.4 To Fsx + 96 CPU + 360 Go RAM





Computational (Cloud)

Performance/Cost
Budget

- Performance/Cost ratio
  - Acceptable performance and cost
  - Time of delivery vs original forecast
  - Costs vs skill (forecast's added value)

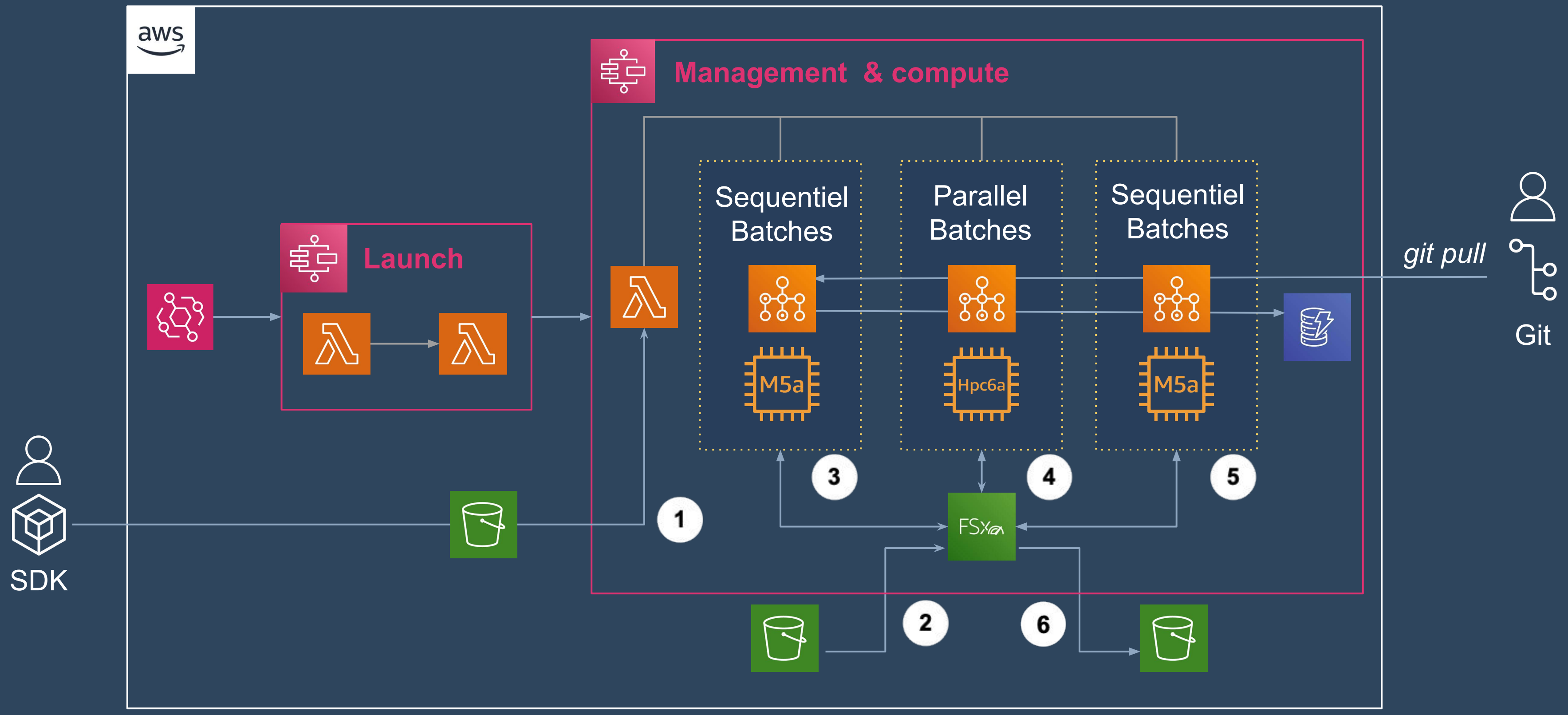
    Cost-function including HW/WN - CS tailoring

- Budget
  - Cloud benefits
    - Infrastructure-as-Code (IaC)
    - Serverless (compute environment)
    - Easy testing/scaling
  - Project necessary resources are unspecified
    - Project budget will shape the demo

Seasonal BA-Calibration batch chain (cloud): 1 model 1 variable 50 members Global 1°

Instances	C2	C2	M2	M2	M16 : M16	R8	M2	
Time (7H5)	90’	122’	44’	25’	135’	20’	7	SSD FSX200 (125-1000 MB/S)
Batch (Code)	T1	T2	T3	T4	T5 : T5 (6 Lead times)	T6	T7	

“HPC serverless”architecture (for one “Unit”)



"**ML** post-processing can act as a **bridge** between the **physical representation** of the atmosphere provided by numerical weather prediction and the **decision-making** requirements of end-users."

Haupt et al. 2021 Towards implementing artificial intelligence post-processing in weather and climate, Phil. Trans. R. Soc. A 379: 20200091.

Thanks

harilaos@theclimatedatafactory.com

the climate data factory