cmcc
Centro Euro-Mediterraneo sui Cambiamenti Climatici
20 Years

UNIBO **GUEST LECTURES**

Expanding CMCC Seasonal Prediction System v3.5 applications to the local scale through STATISTICAL DOWNSCALING TECHNIQUES

Leonardo Aragão[1], Andrea Borrelli[2] & Silvio Gualdi[1]

[1]Earth SYstem modeling and Data Assimilation (ESYDA)
[2]CLImate VAriability and Predictions (CLIVAP)

www.cmcc.it


Seasonal Prediction


Seasonal Prediction
provide a long-range outlook of changes in the Earth system over periods of a few weeks or months, as a result of predictable changes in some of the slow-varying components of the system.

Seasonal Prediction

provide a **long-range outlook** of changes in the **Earth system** over periods of a few weeks or months, as a result of **predict**able changes in some of the **slow-varying** components of the system.

to forecast X to predict

weather climate

Seasonal Prediction

provide a **long-range outlook** of changes in the **Earth system** over periods of a few weeks or months, as a result of predictable changes in some of the **slow-varying** components of the system.

to forecast ✗ to predict

to forecast	to predict
weather	climate
product	process
time-stamped	outlook

Seasonal Prediction

provide a **long-range** outlook of changes in the **Earth system** over periods of a few weeks or months, as a result of predictable changes in some of the **slow-varying** components of the system.

to forecast	to predict
weather	climate
product	process
time-stamped	outlook
short-range	long-range
event	slow-varying
deterministic model	Earth system

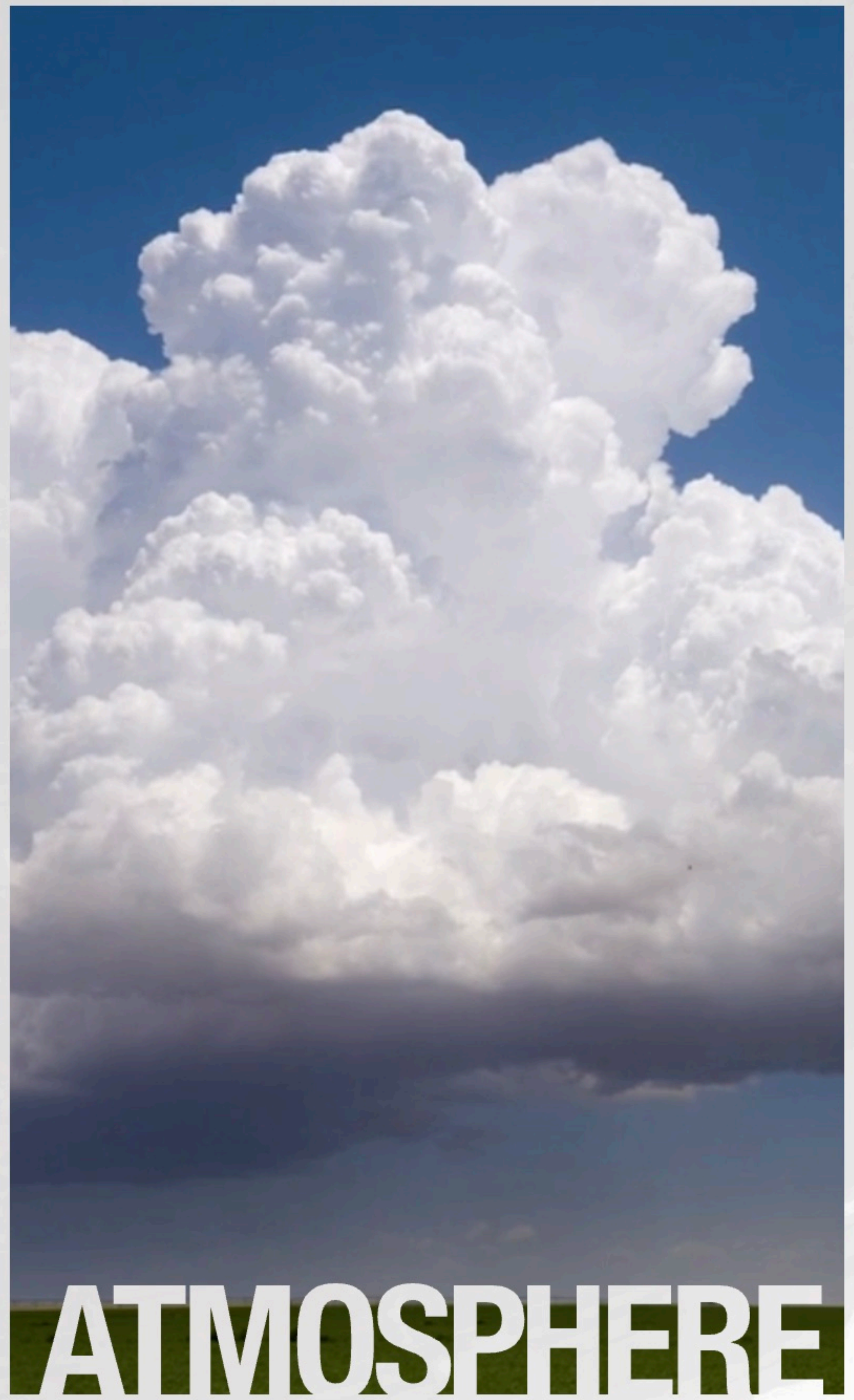
Seasonal Prediction

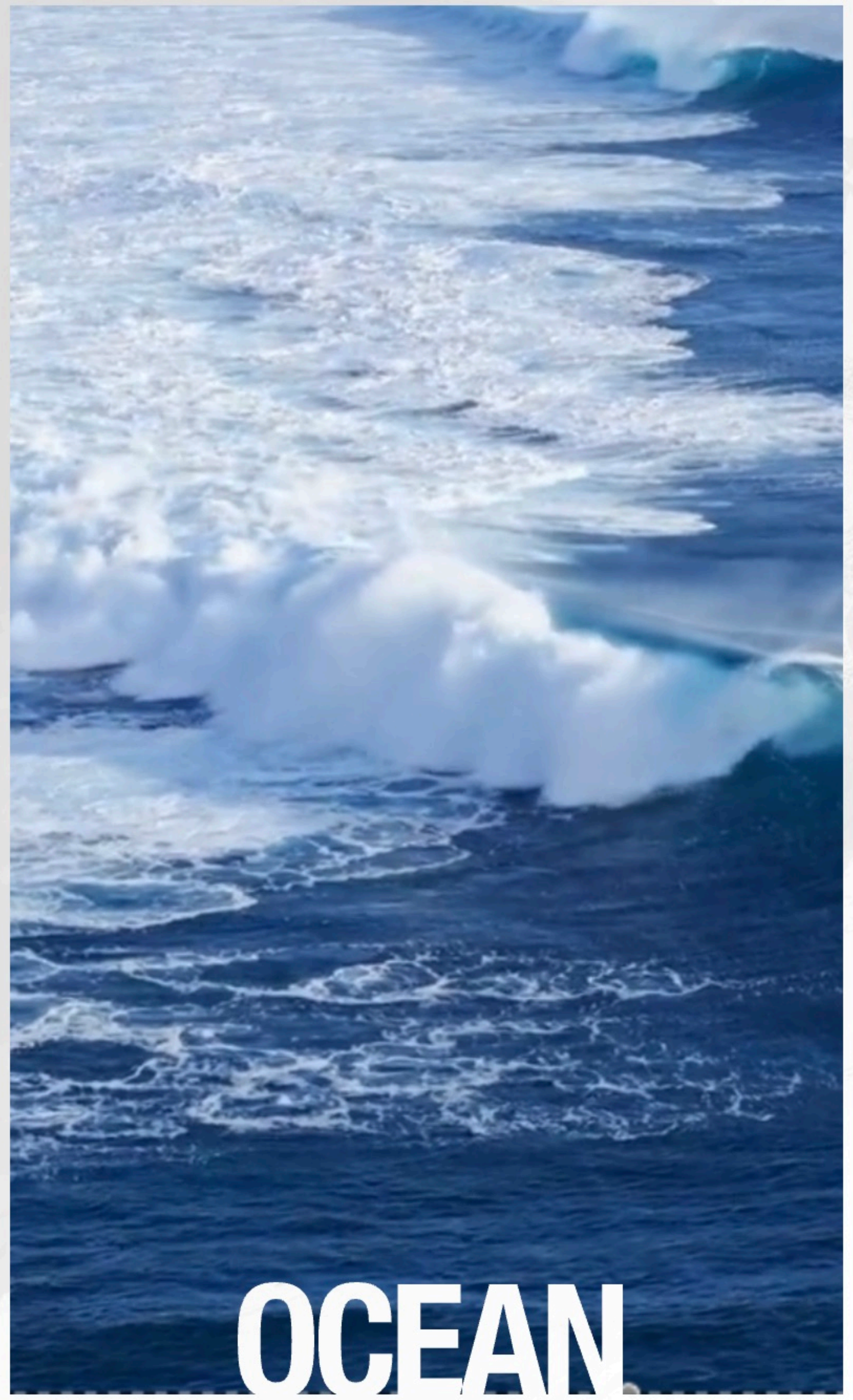
provide a long-range outlook of changes in the Earth system over periods of a few weeks or months, as a result of predictable changes in some of the slow-varying components of the system.

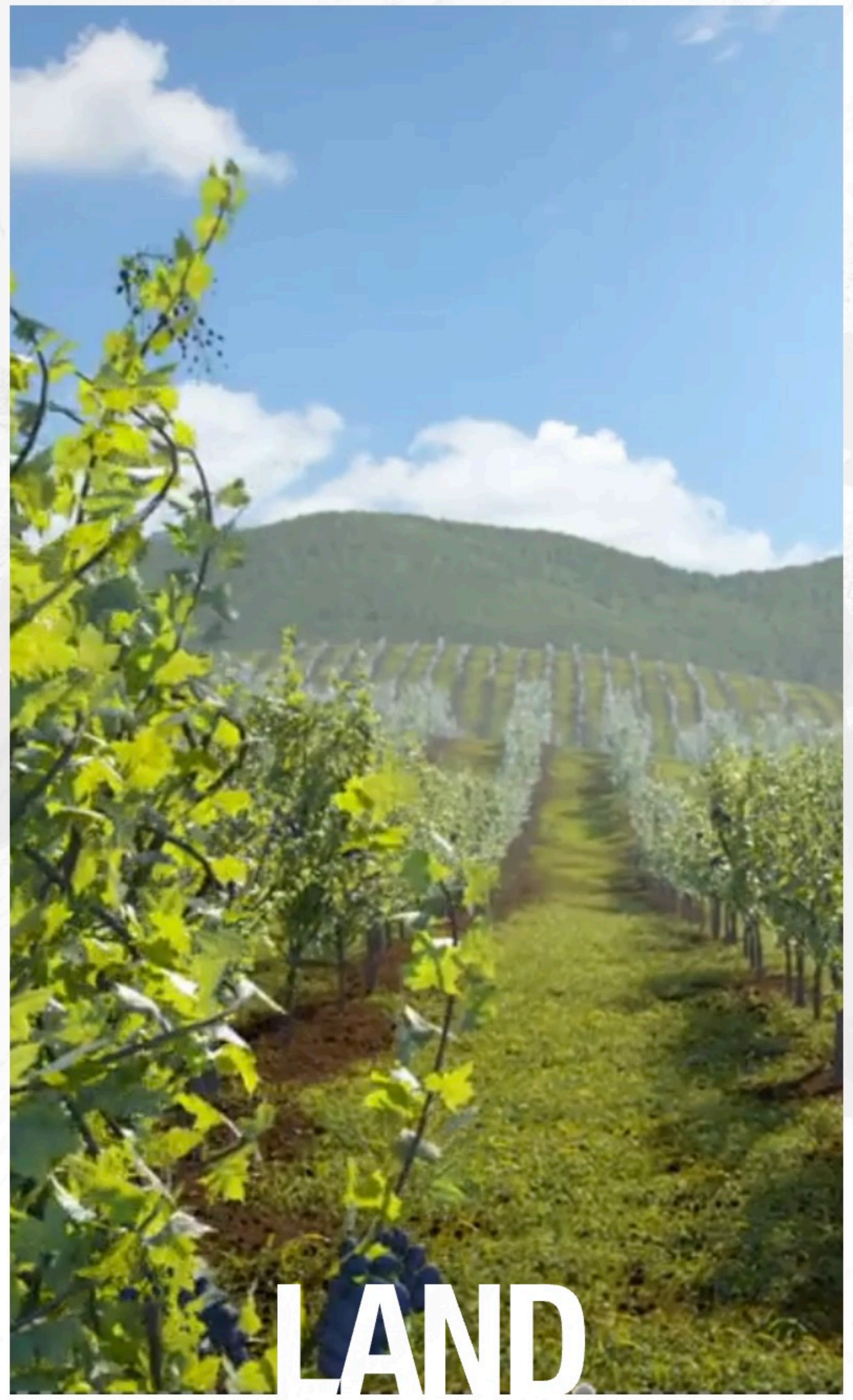
EARTH SYSTEM

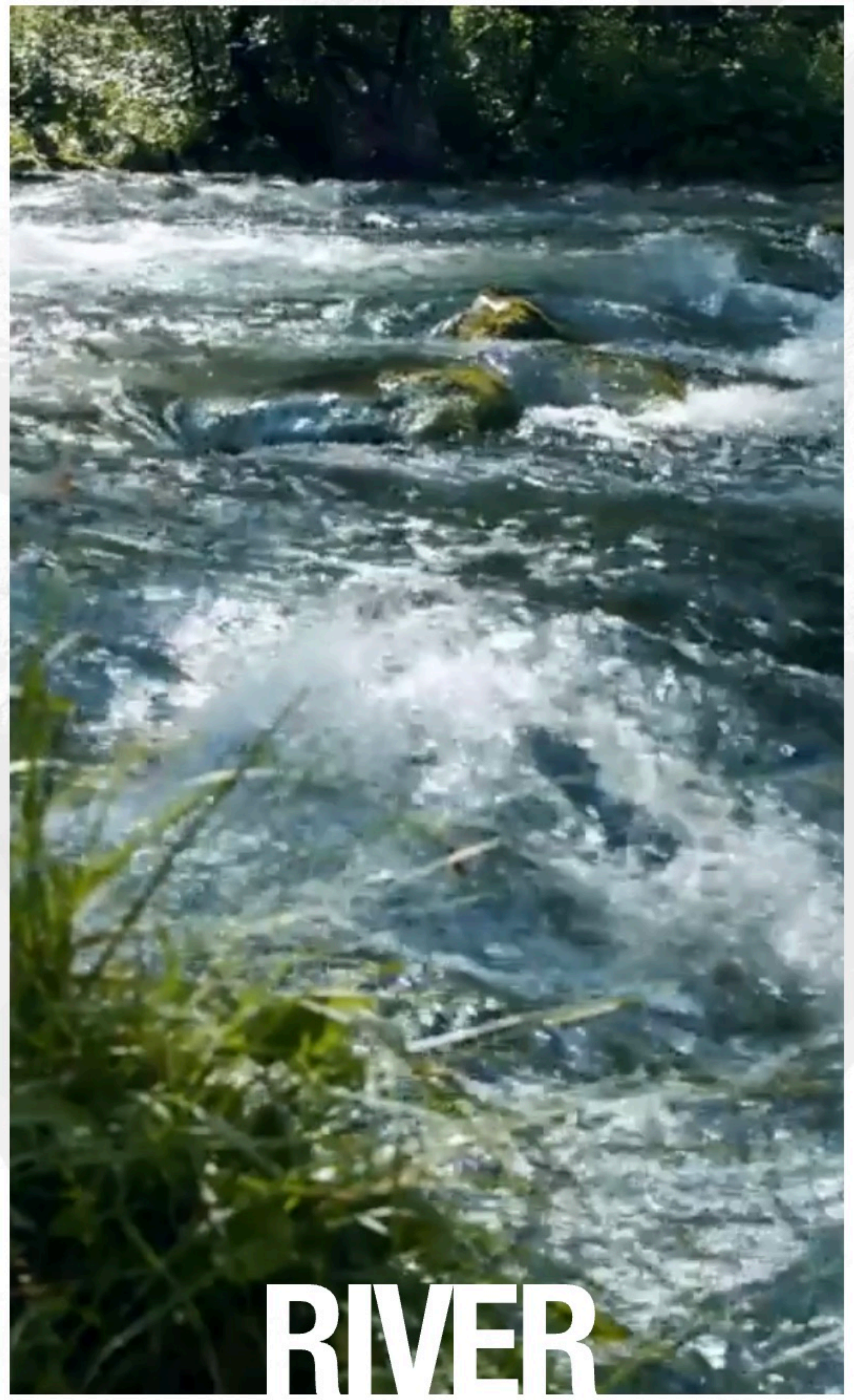
EARTH SYSTEM MODELS

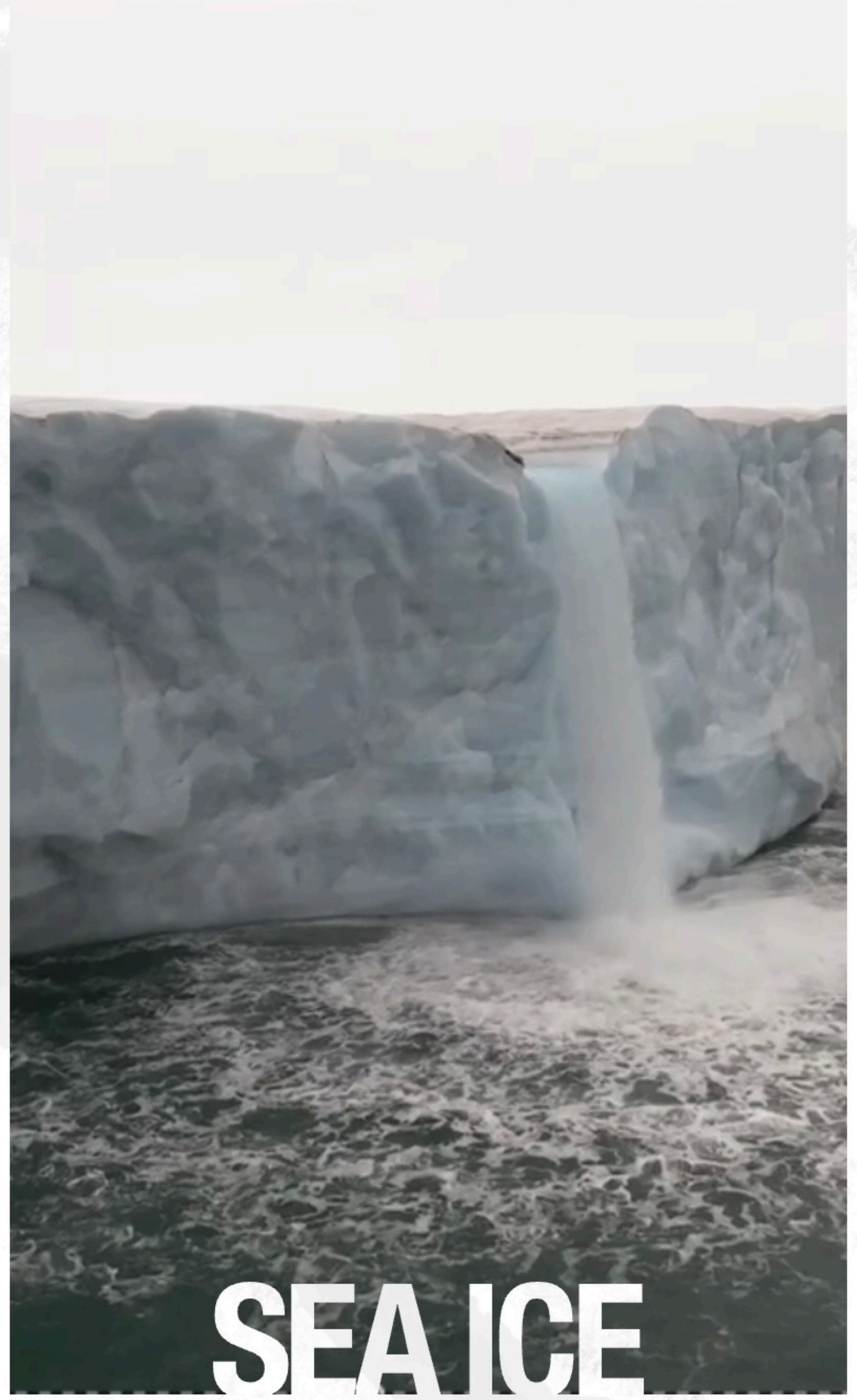
CMCC Seasonal Prediction System v4





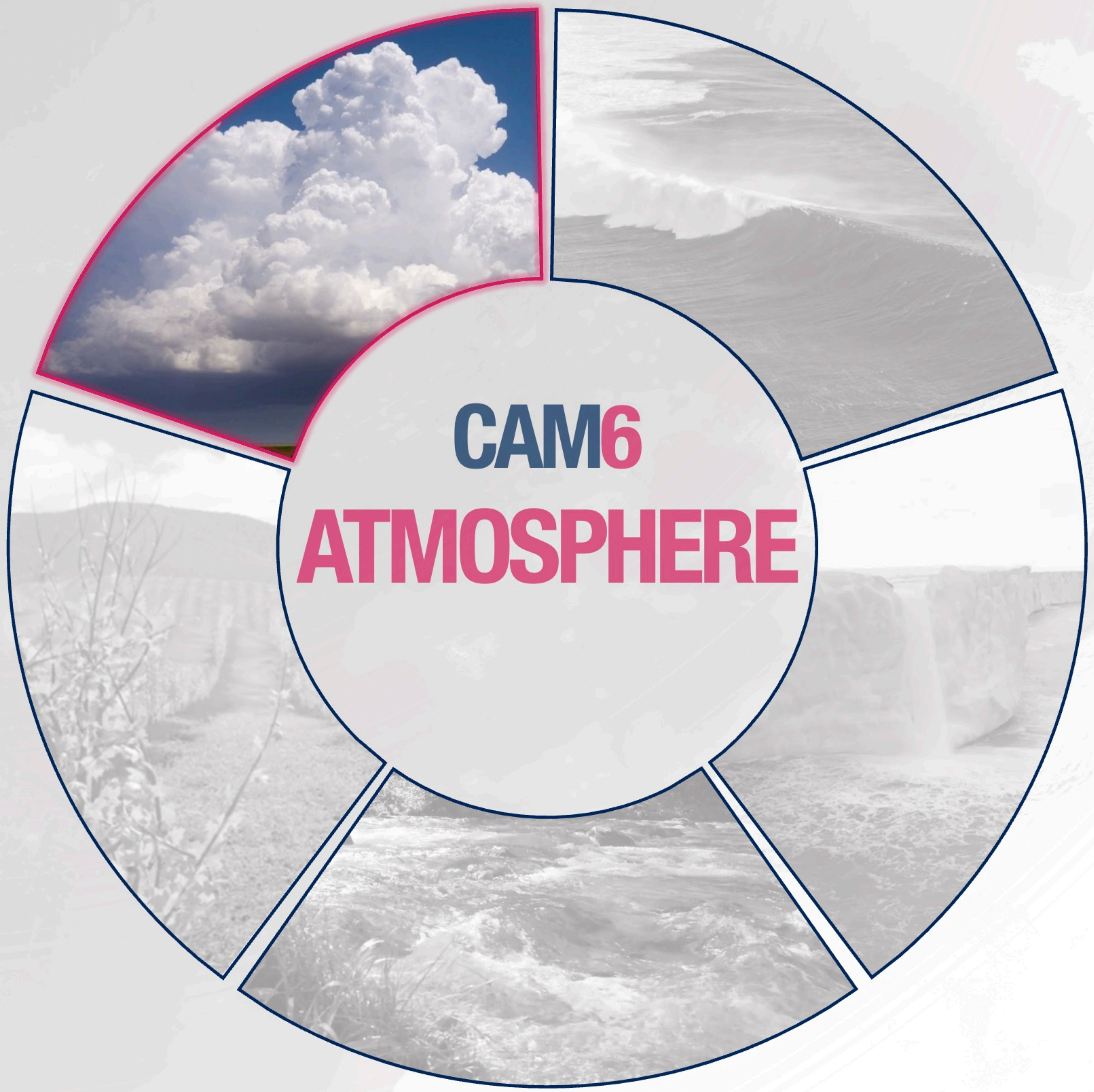






EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



Community Atmosphere Model version 6 (5.3)

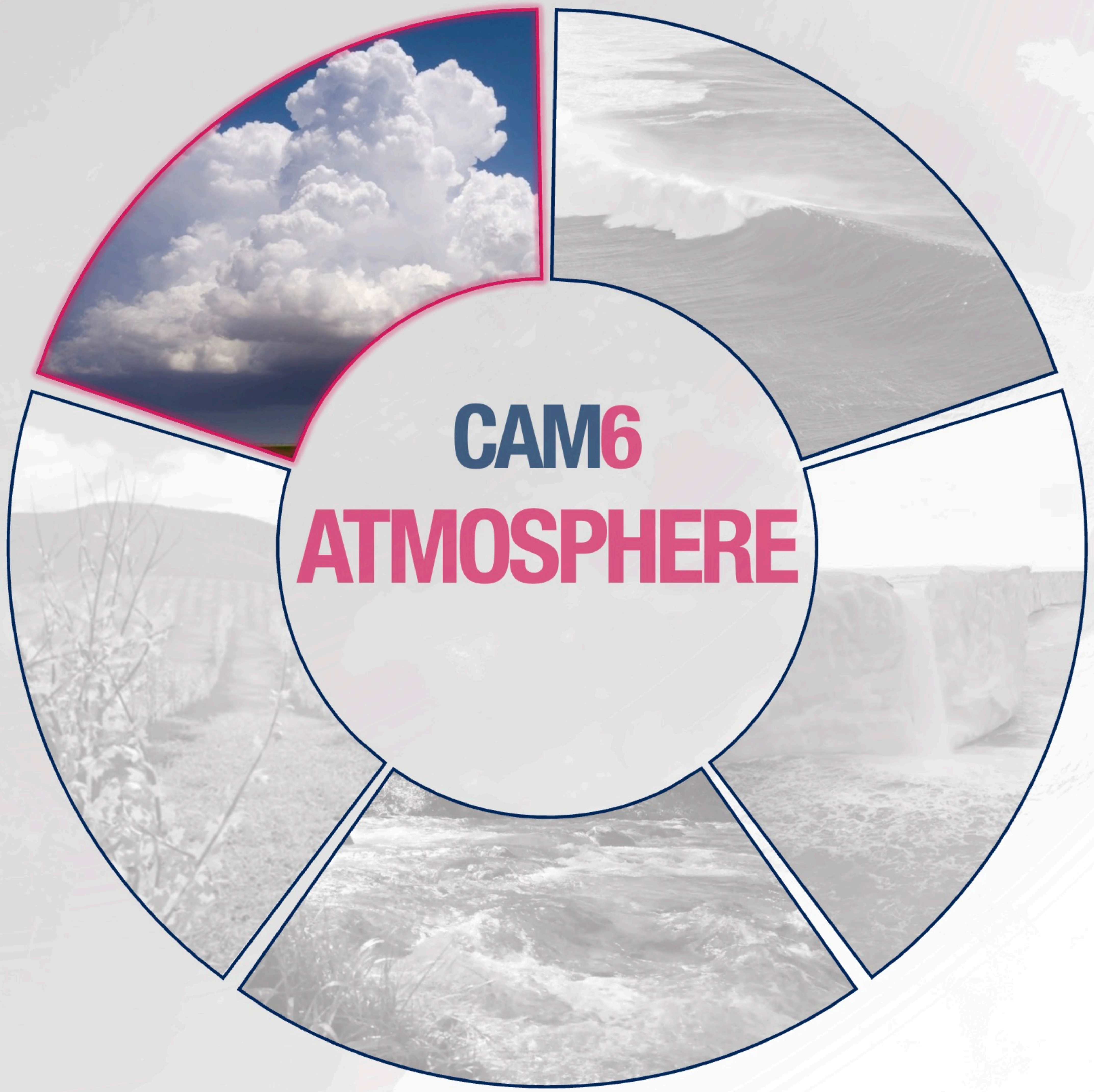
Craig *et al.* (2021)

- Atmospheric component of the NCAR's Community Earth System Model (CESM)
- Regular grid of 0.47° (latitude) × 0.63° (longitude)
- 83 vertical hybrid levels
- Troposphere and Stratosphere up to 0.01 hPa

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)

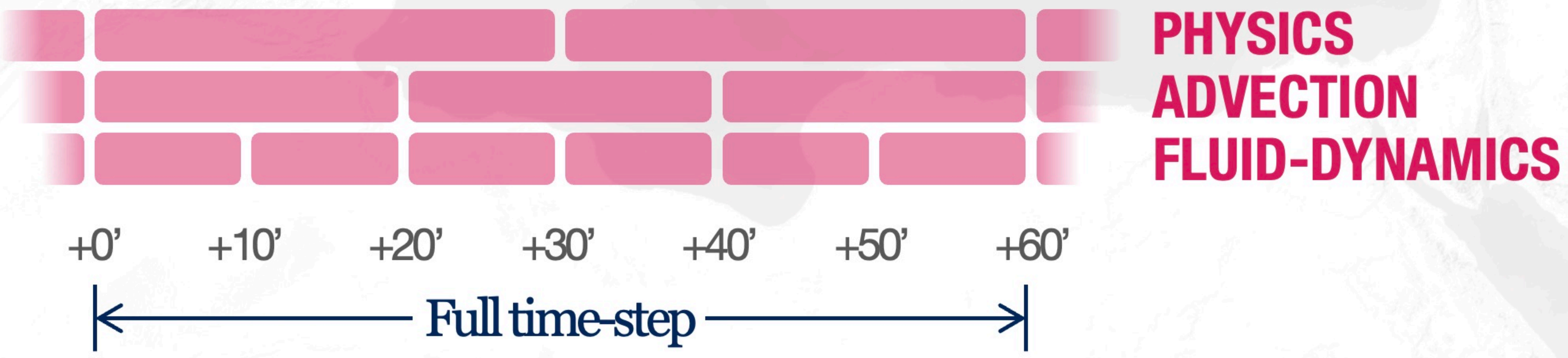


Community Atmosphere Model version 6 (5.3)

Craig *et al.* (2021)

- Atmospheric component of the NCAR's Community Earth System Model (CESM)
- Regular grid of 0.47° (latitude) × 0.63° (longitude)
- 83 vertical hybrid levels
- Troposphere and Stratosphere up to 0.01 hPa

Time-steps:

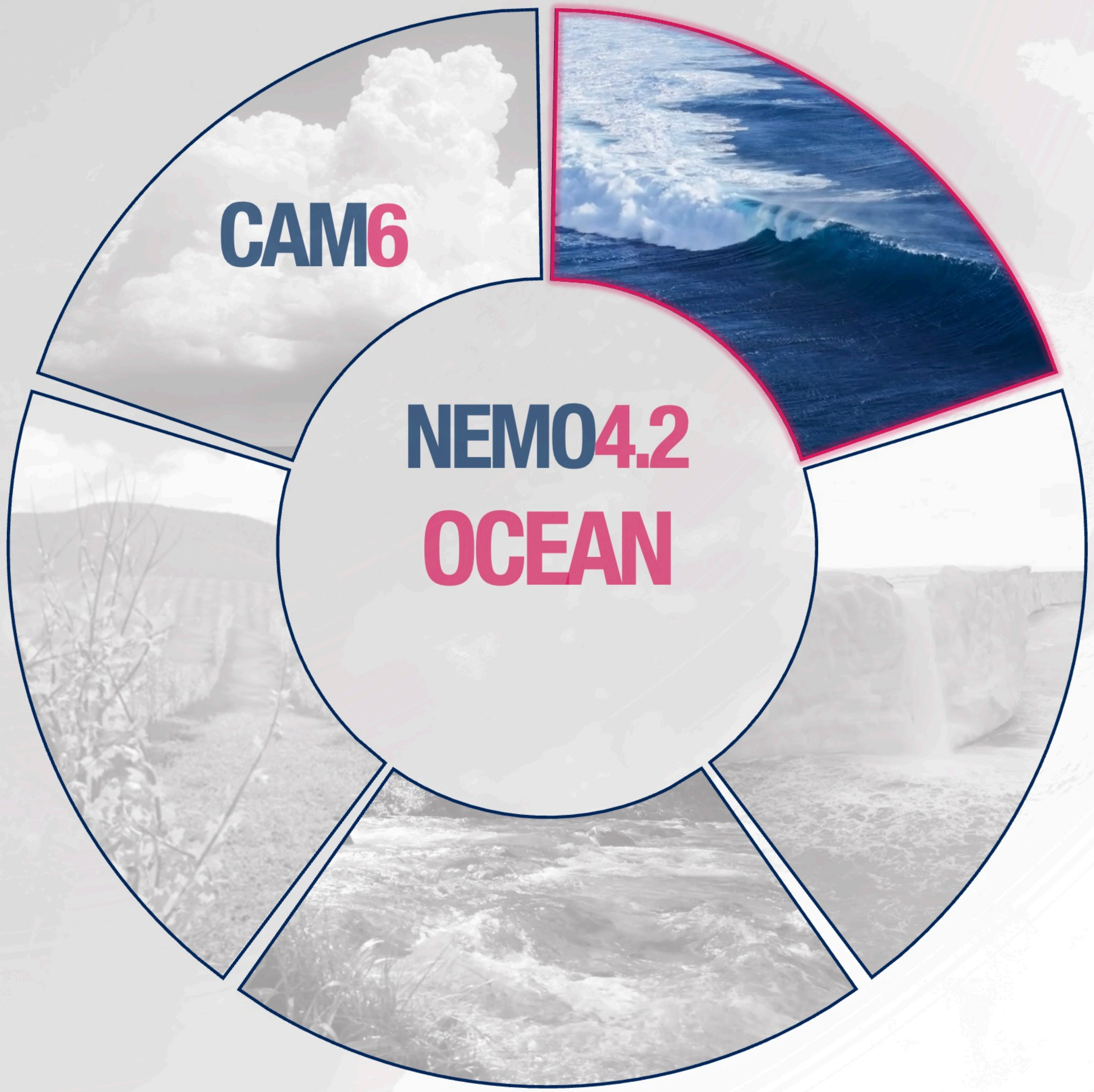


More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



Nucleus for European Modelling of the Ocean model version 4.2 (3.4)

Madec *et al.* (2022)

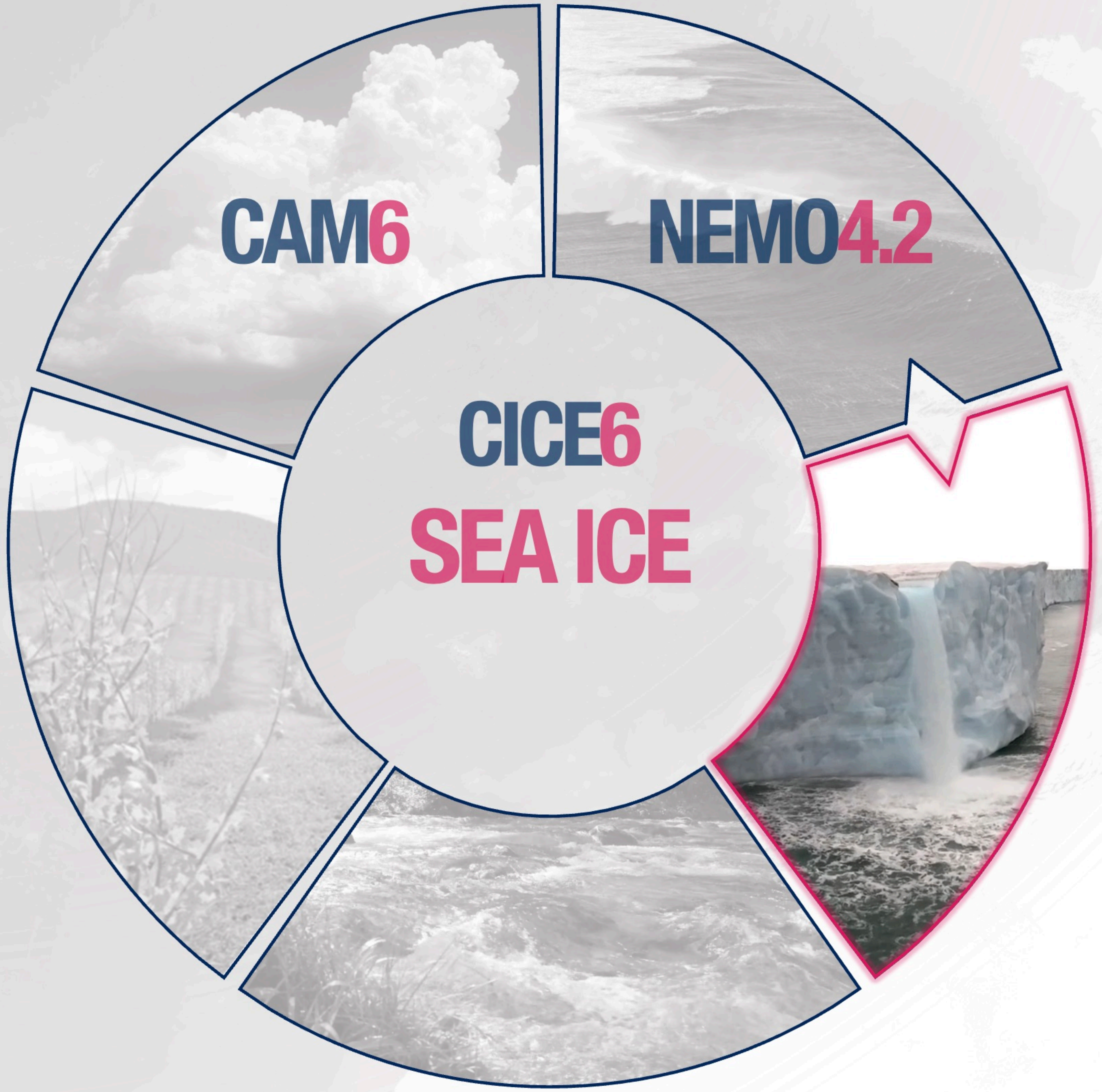
- Tripolar ORCA grid of 0.25 °
- 75 vertical levels
- 20-minute time-step
- Prognostic variables: three velocity components,
  the sea surface height,
  the potential temperature, and
  salinity

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici | 20 Years

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



Los Alamos sea ICE model version 6 (4)

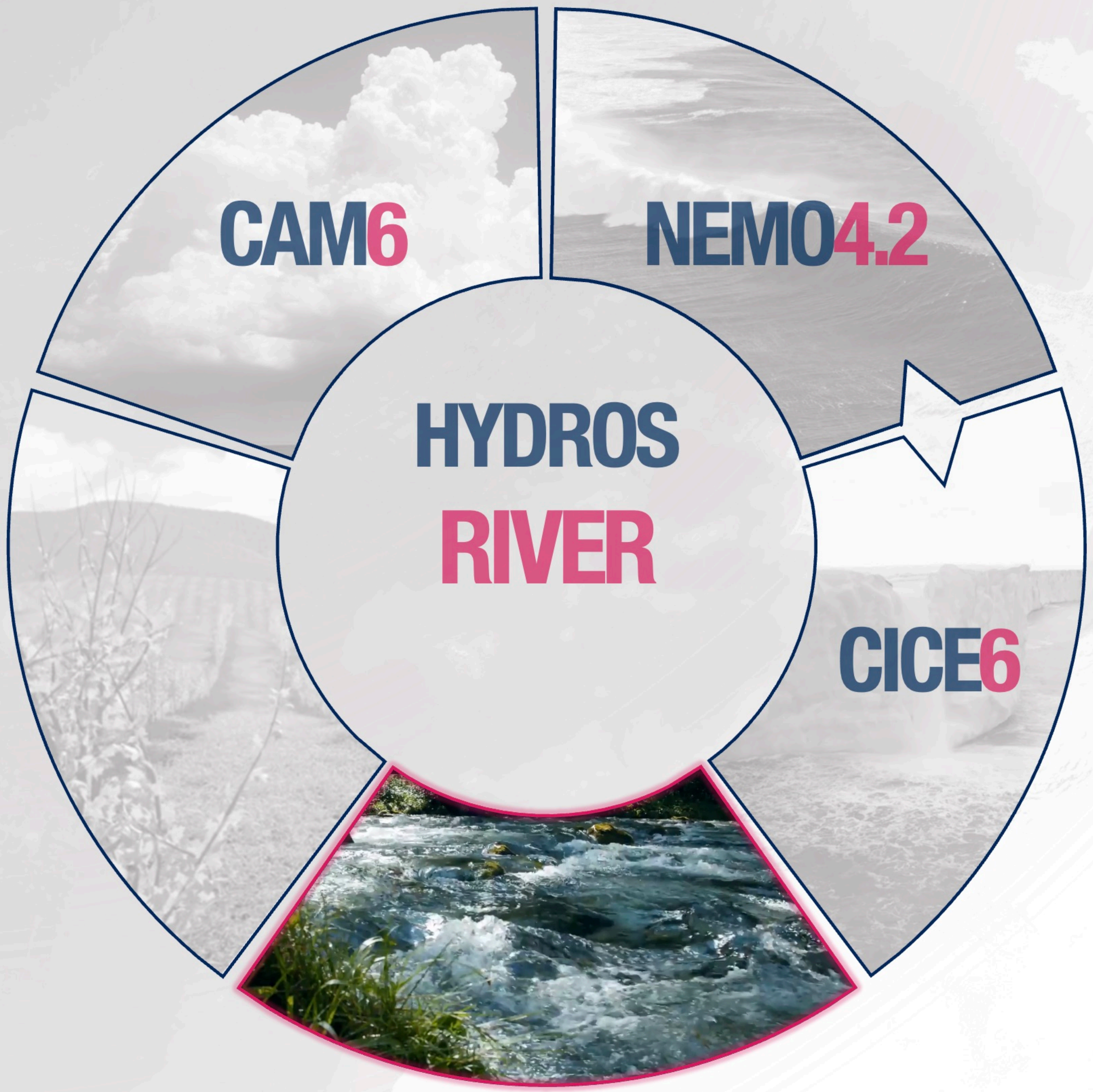
Hunke *et al.* (2018)

- Consortium ICE
- Integrated with NEMO (coupled): Tripolar ORCA grid of 0.25°
  20-minute time-step
- 8 ice layers
- 3 snow layers
- 5 ice categories (single)
- No wave model

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



HYdro-Dynamic ROuting Scheme (RTM)

Materia *et al.* (2020)

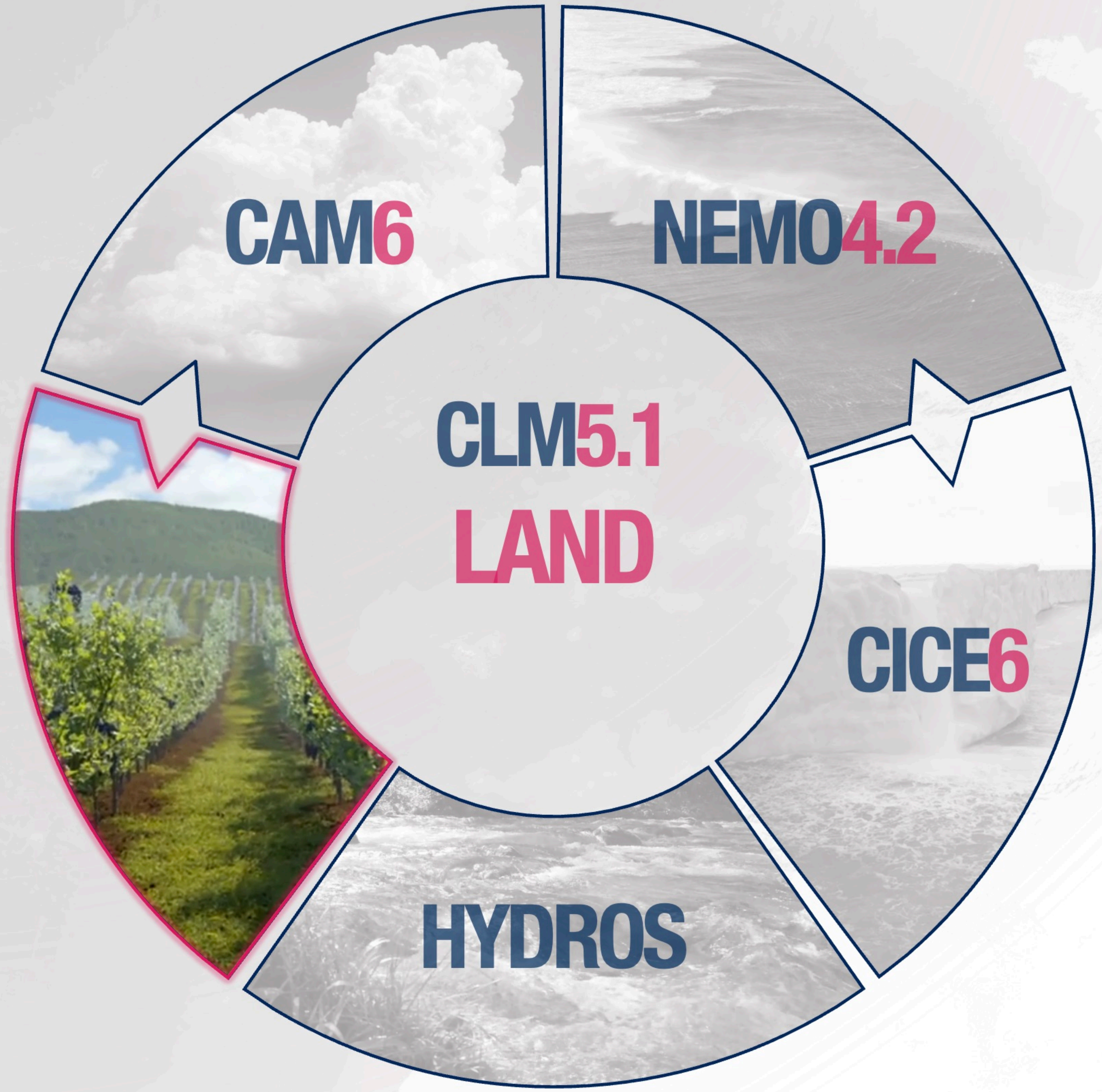
- Replaces the River Transport Model (RTM, Graham *et al.* 1999)
- A more complex scheme for computing the flow velocity
- Generates time-dependent flow velocity influenced by orography and water availability
- Regular grid of 0.5°
- 3-hour time step
- Receive input data from the land surface model

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



Community Land Model version 5.1 (4.5)

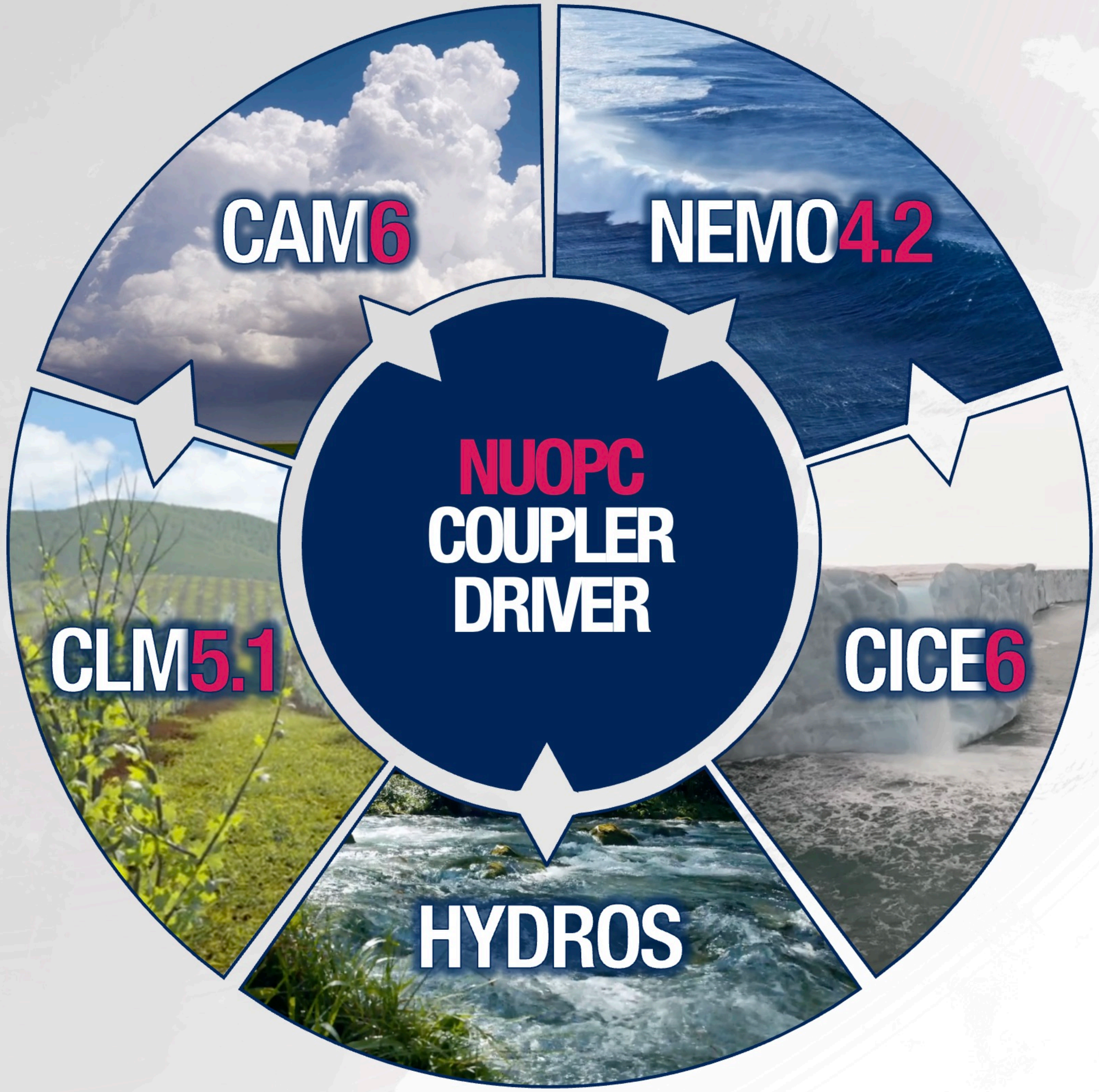
Lawrence *et al.* (2019)

- Land/Surface component of the NCAR's Community Earth System Model (CESM)
- Integrated with CAM6 (coupled): Regular grid of 0.47° (latitude) × 0.63° (longitude)
30-minute time step
- 20 soil layers (down to 20 m)
- 5 bedrock layers (down to 42 m)
- Snow density parameterisation accounting for a tuning factor to adjust the snow ageing
- Biogeochemical cycles included

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)



National Unified Operational Prediction Capability

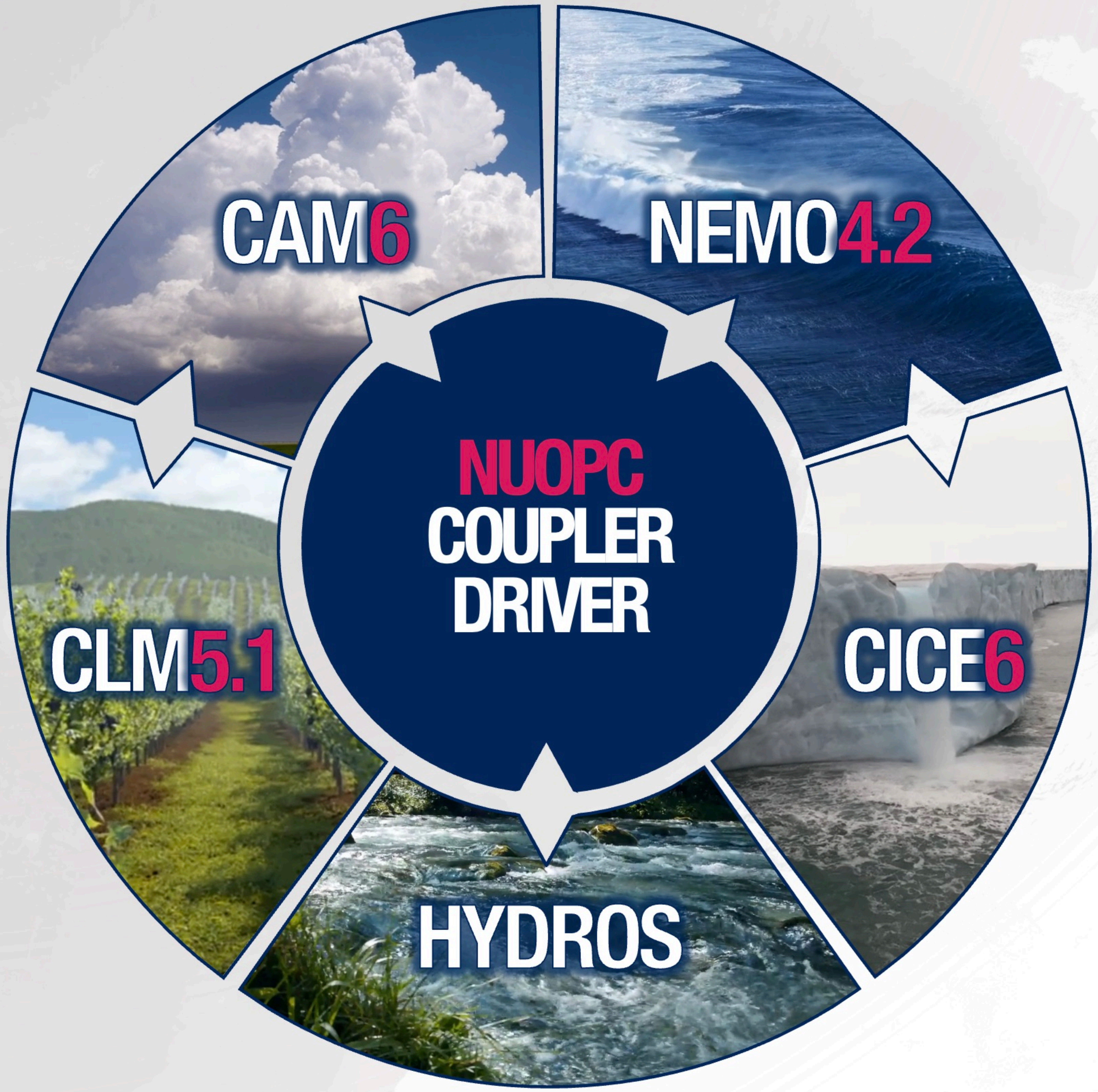
Earth System Modelling Framework

- Coupler component of the NCAR's Community Earth System Model (CESM)
- Atmosphere-Ocean 60 minutes (full time-step of the atmospheric model)
- Atmosphere-Land 30 minutes (internally in CAM6)
- Ocean-Sea Ice 20 minutes (internally in NEMO4.2)
- Land-River 180 minutes (3x full time-step of the atmospheric model)

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS
CMCC Seasonal Prediction System v4 (v3.5)



INITIALISATIONS & ENSEMBLE
Copernicus Climate Change Services (C3s) framework

Perturbed Initial Conditions

10	x	3	x	9	=	270
Atmosphere		Land		Ocean		Possibilities

More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)

INITIALISATIONS & ENSEMBLE

Copernicus Climate Change Services (C3s) framework

270 Possibilities → Random Selection → **50** Ensemble Members*

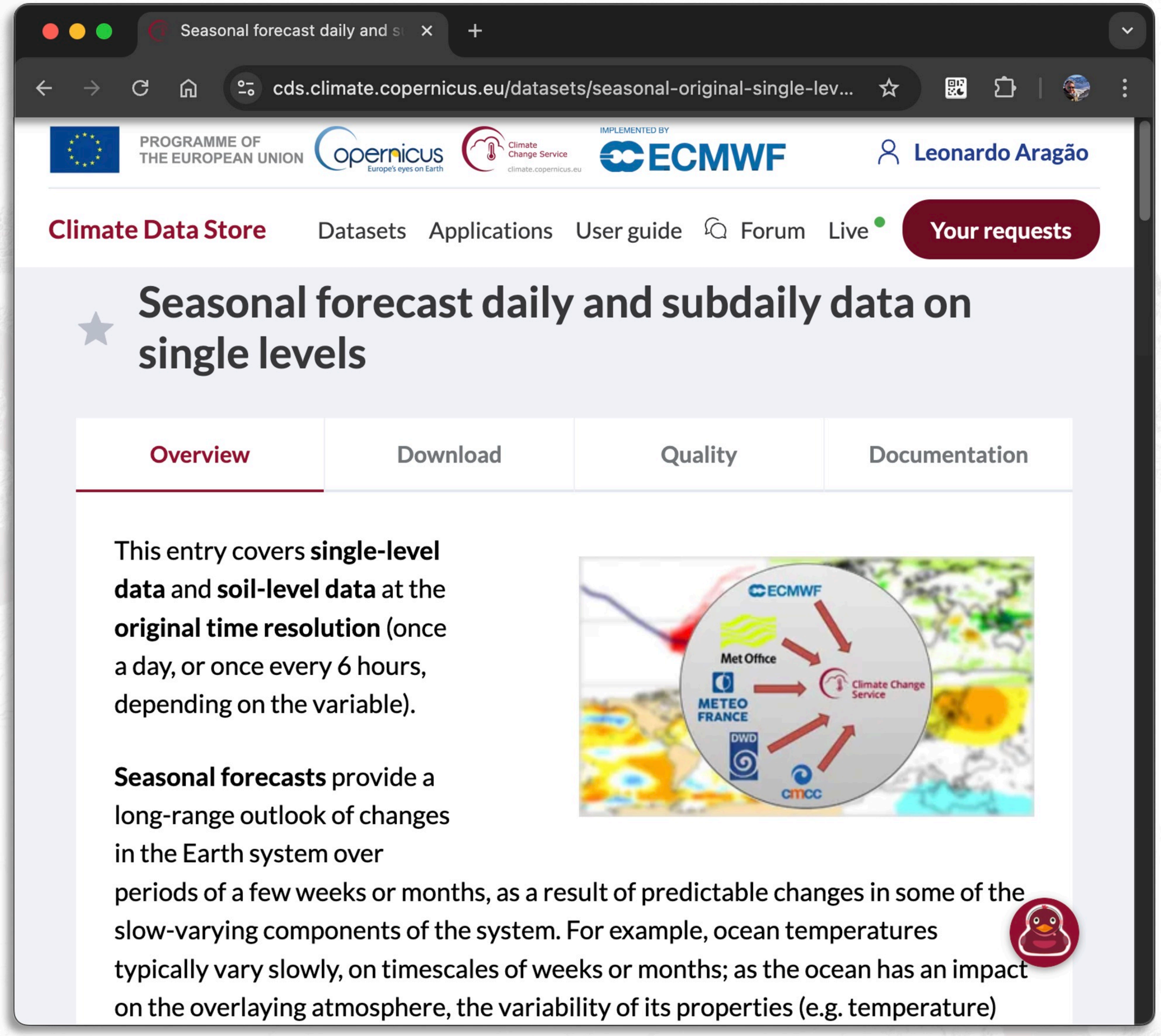
*Hindcast mode (30 in forecast mode).
More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS
CMCC Seasonal Prediction System v4 (v3.5)

INITIALISATIONS & ENSEMBLE
Copernicus Climate Change Services (C3s) framework

270 Possibilities → Random Selection → **50** Ensemble Members*



*Hindcast mode (30 in forecast mode).
More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

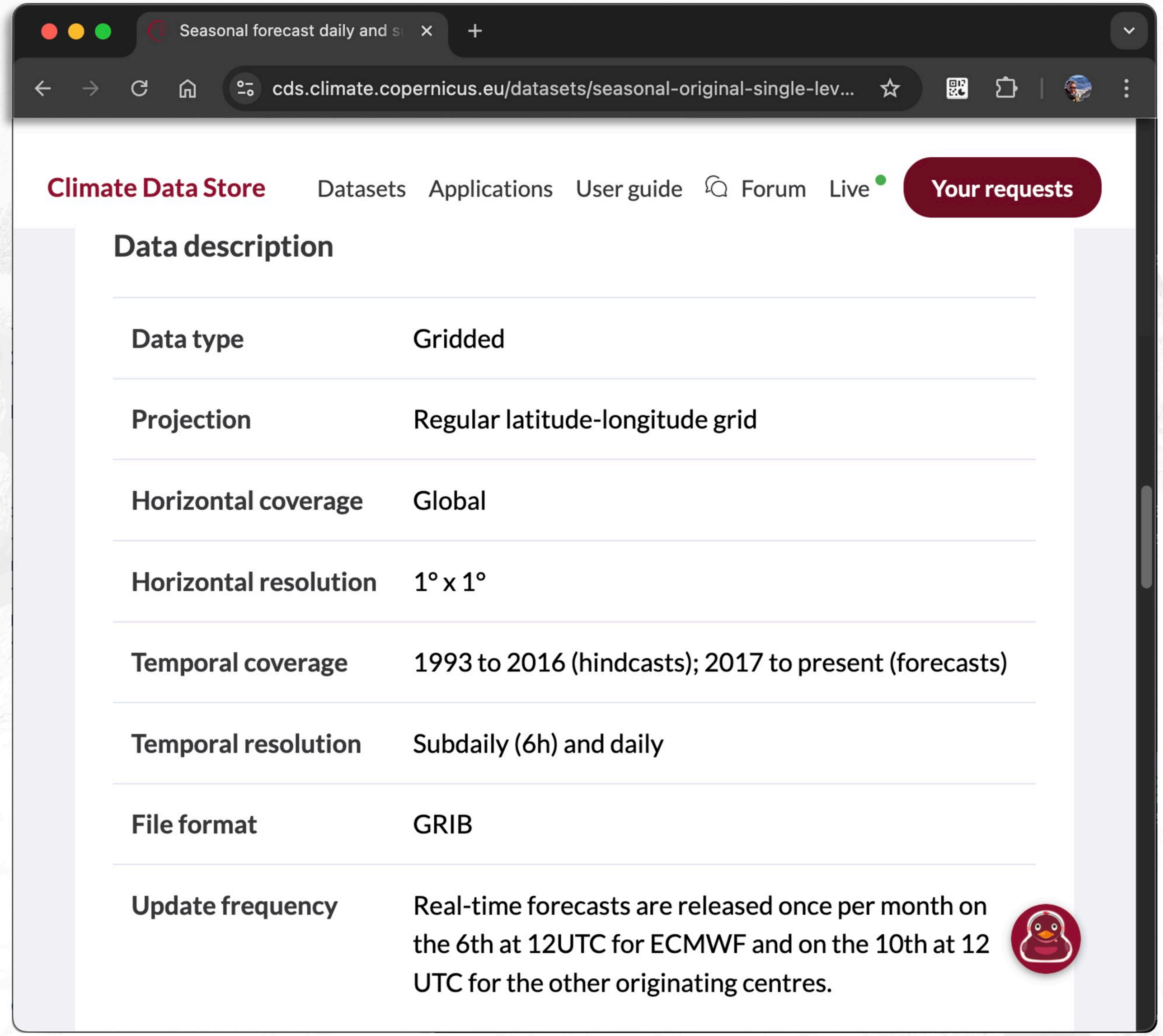
EARTH SYSTEM MODELS

CMCC Seasonal Prediction System v4 (v3.5)

INITIALISATIONS & ENSEMBLE

Copernicus Climate Change Services (C3s) framework

270 Possibilities → Random Selection → **50** Ensemble Members*


Seasonal forecast daily and s
cds.climate.copernicus.eu/datasets/seasonal-original-single-lev...
Climate Data Store | Datasets | Applications | User guide | Forum | Live | Your requests

Data description

Data type	Gridded
Projection	Regular latitude-longitude grid
Horizontal coverage	Global
Horizontal resolution	1° x 1°
Temporal coverage	1993 to 2016 (hindcasts); 2017 to present (forecasts)
Temporal resolution	Subdaily (6h) and daily
File format	GRIB
Update frequency	Real-time forecasts are released once per month on the 6th at 12UTC for ECMWF and on the 10th at 12 UTC for the other originating centres.

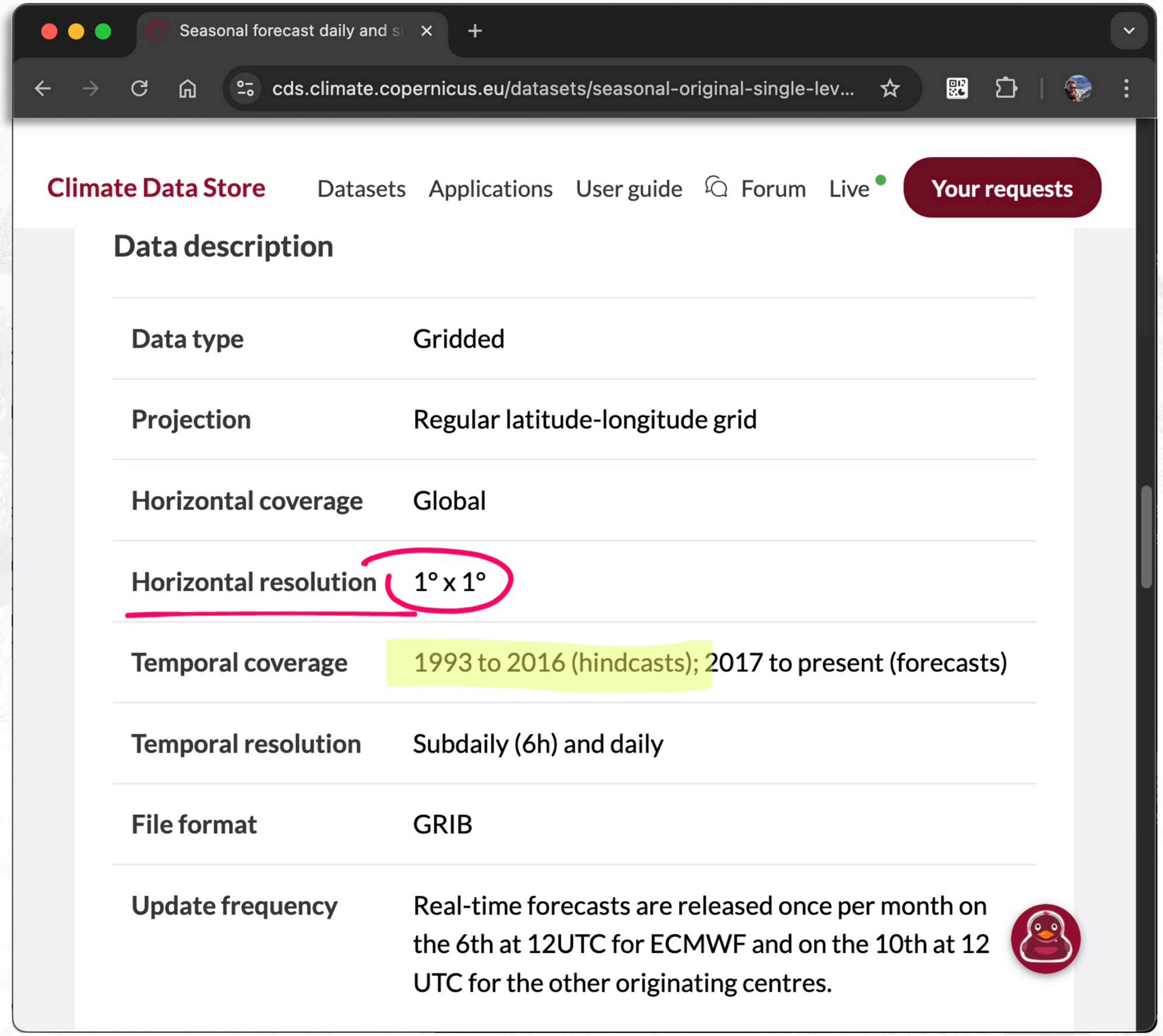
*Hindcast mode (30 in forecast mode).
More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

EARTH SYSTEM MODELS
CMCC Seasonal Prediction System v4 (v3.5)

INITIALISATIONS & ENSEMBLE
Copernicus Climate Change Services (C3s) framework

270 Possibilities → Random Selection → **50** Ensemble Members*


Seasonal forecast daily and s
cds.climate.copernicus.eu/datasets/seasonal-original-single-lev...

Climate Data Store | Datasets | Applications | User guide | Forum | Live | Your requests

Data description

Data type	Gridded
Projection	Regular latitude-longitude grid
Horizontal coverage	Global
Horizontal resolution	1° x 1°
Temporal coverage	1993 to 2016 (hindcasts); 2017 to present (forecasts)
Temporal resolution	Subdaily (6h) and daily
File format	GRIB
Update frequency	Real-time forecasts are released once per month on the 6th at 12UTC for ECMWF and on the 10th at 12 UTC for the other originating centres.

*Hindcast mode (30 in forecast mode).
More information in: Sanna *et al.* (2025) // DOI 10.25424/cmcc-dkcv-fs25

STATISTICAL DOWNSCALING

OUR SCOPE

"The general objective is to develop a computational algorithm in Python language that allows downscaling low-resolution meteorological data (historical or forecast) using machine learning techniques."

OUR SCOPE

The general objective is to develop a computational algorithm in Python language that allows downscaling low-resolution meteorological data (historical or forecast) using machine learning techniques.

EXPERIMENT

> *"To evaluate two different machine learning methods, Empirical Quantile Mapping (EQM) and k-Nearest Neighbours (kNN), in downscaling seven meteorological fields over the Italian Peninsula based on the 24-year hindcasts (1993-2016) of the CMCC Seasonal Prediction System v3.5 (SPS)."*

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

REG	Global
DOM	180×360
RES	1.00° Daily
PER	1993-2016

SPS
T2m T2N T2X SST
Prec U10 V10

ERA5
T2m T2N T2X
SST U10 V10

CHIRPS
Prec

***k*NN**
T2m T2N T2X SST
Prec U10 V10

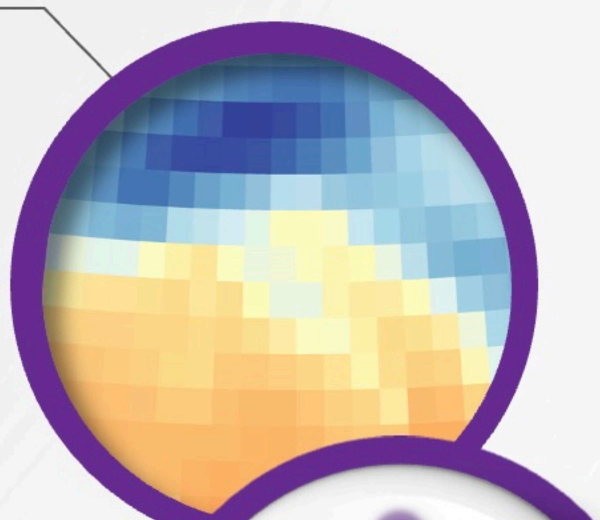
EQM
T2m T2N T2X SST
Prec U10 V10

T2m	Daily average 2m Temperature
T2N	Daily minimum 2m Temperature
T2X	Daily maximum 2m Temperature
SST	Daily average Sea Surface Temperature
Prec	Daily Precipitation (accumulated)
U10	Daily average zonal wind component
V10	Daily average meriodional wind component

DATA & METHODS

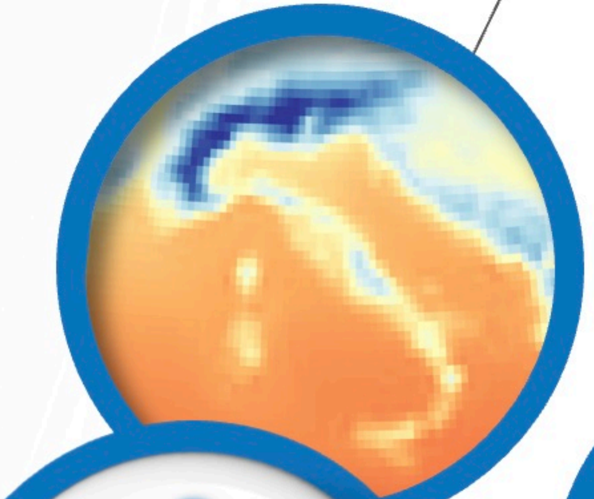
SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

REG Global
DOM 180×360
RES 1.00° Daily
PER 1993-2016



SPS
T2m T2N T2X SST
Prec U10 V10

REG Global
DOM 720×1440
RES 0.25° Hourly
PER 1940-Now



ERA5
T2m T2N T2X
SST U10 V10

REG 50°N – 50°S
DOM 2000×7200
RES 0.05° Daily
PER 1981-Now

CHIRPS
Prec

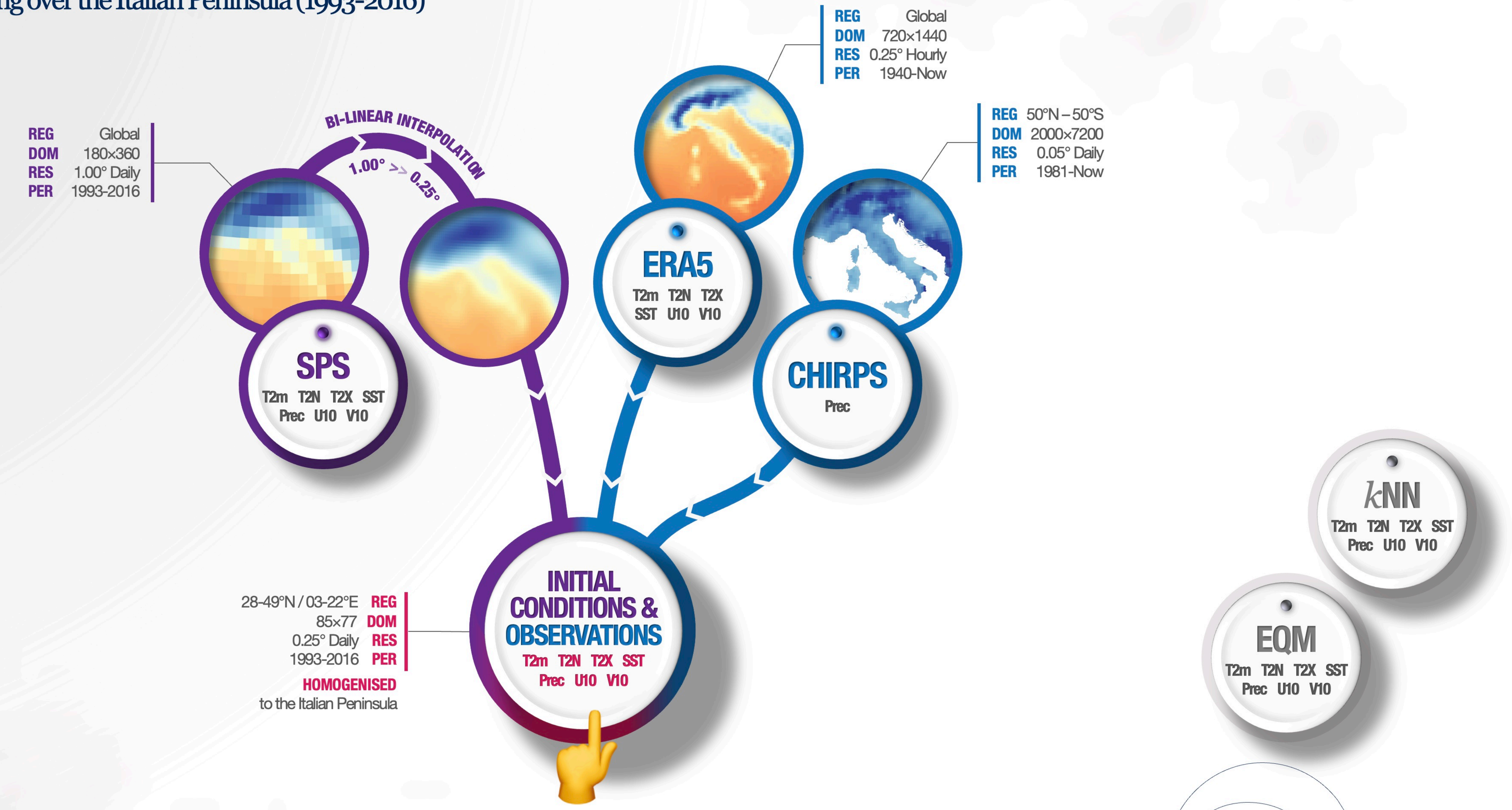
***k*NN**
T2m T2N T2X SST
Prec U10 V10

EQM
T2m T2N T2X SST
Prec U10 V10

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)



DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

$$\mathrm{M}(x_{SPS}) = F_{Obs}^{-1}\left(F_{SPS}(x_{SPS})\right)$$

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici
20 Years

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

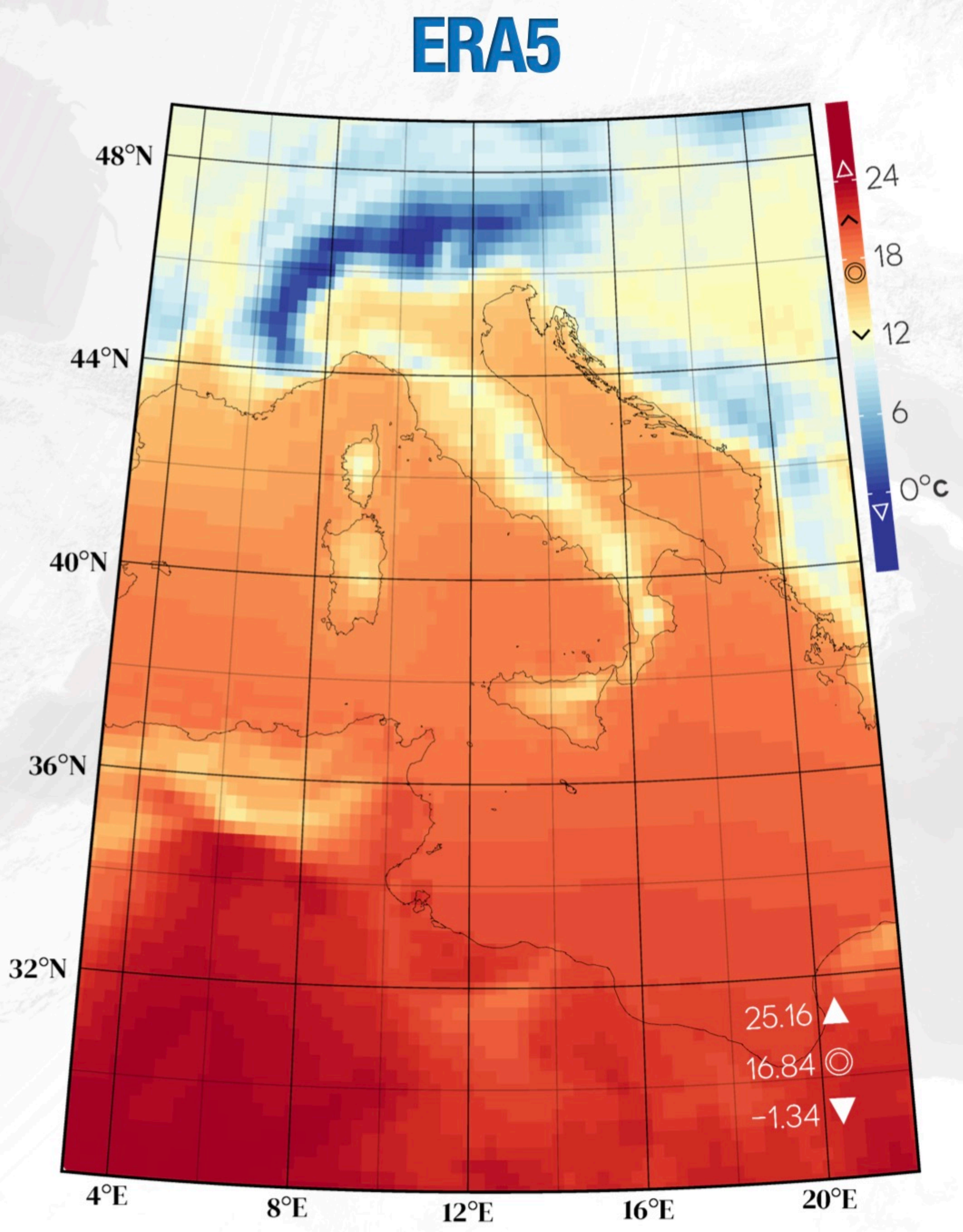
DATA & METHODS

SPS3.5 statistical downscaling over the Italian Peninsula (1993-2016)

SUMMARISED RESULTS

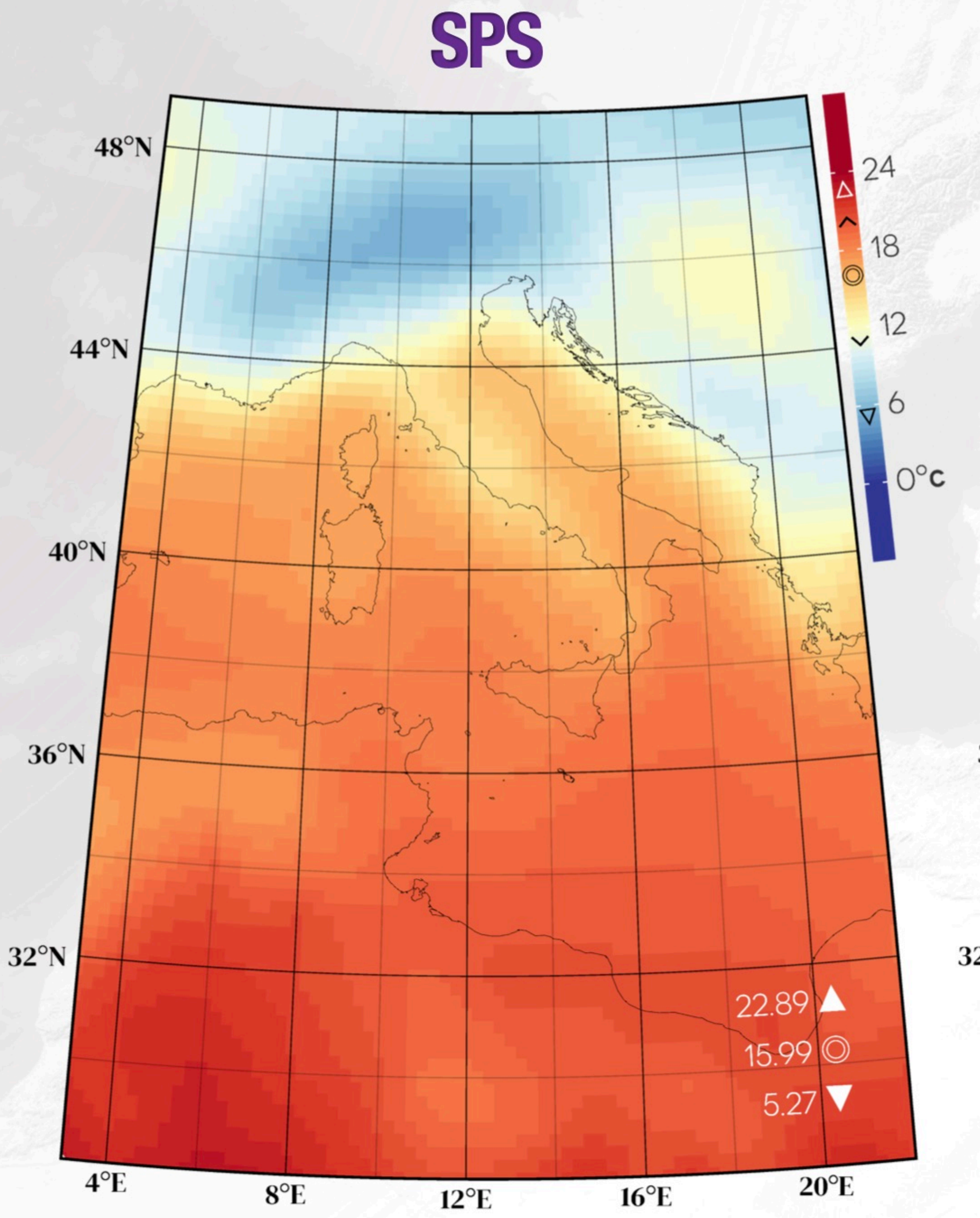
SUMMARISED RESULTS

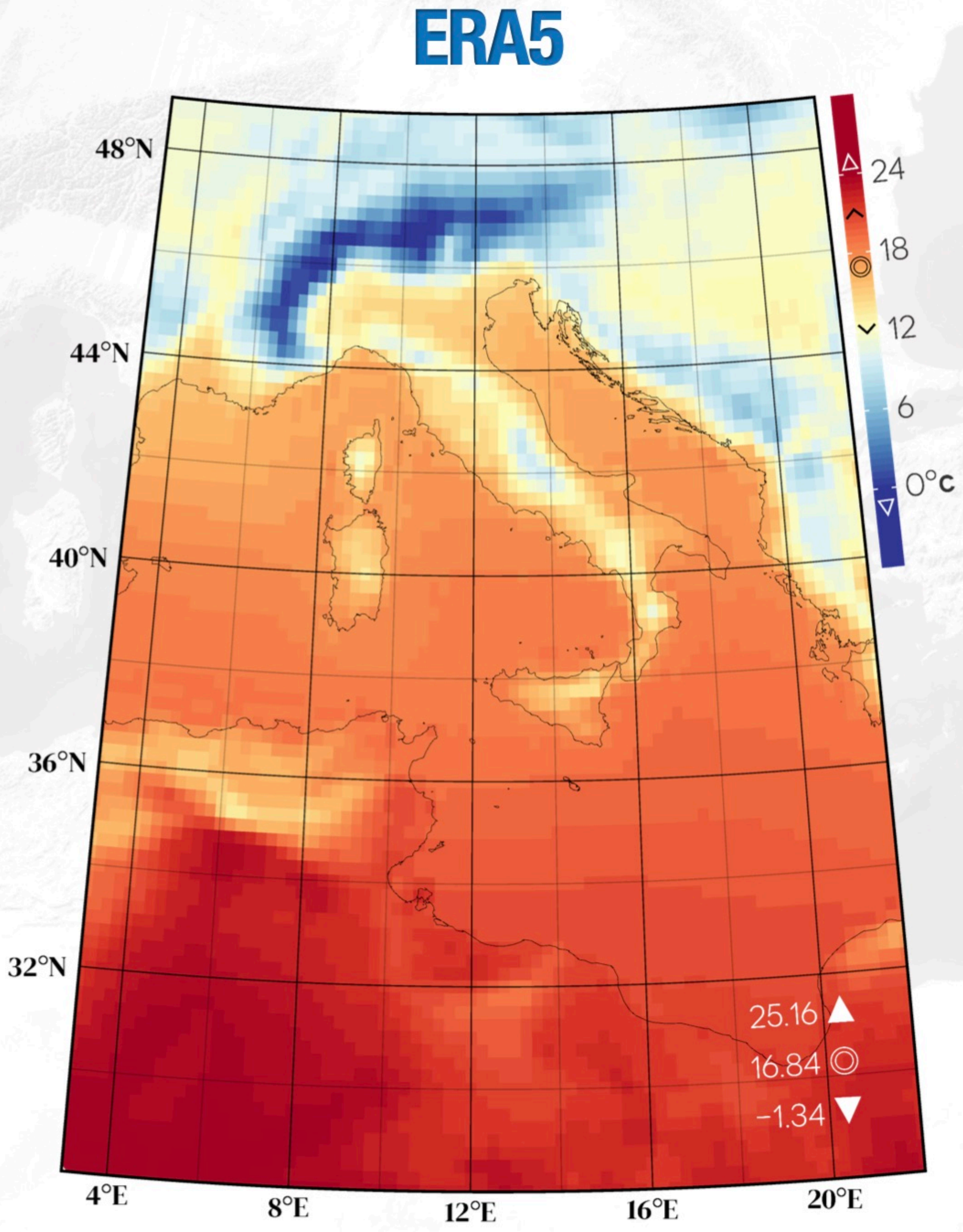
Ensemble annual mean (lead month 1) for the daily average 2m temperature (1993-2016)



SUMMARISED RESULTS

Ensemble annual mean (lead month 1) for the daily average 2m temperature (1993-2016)





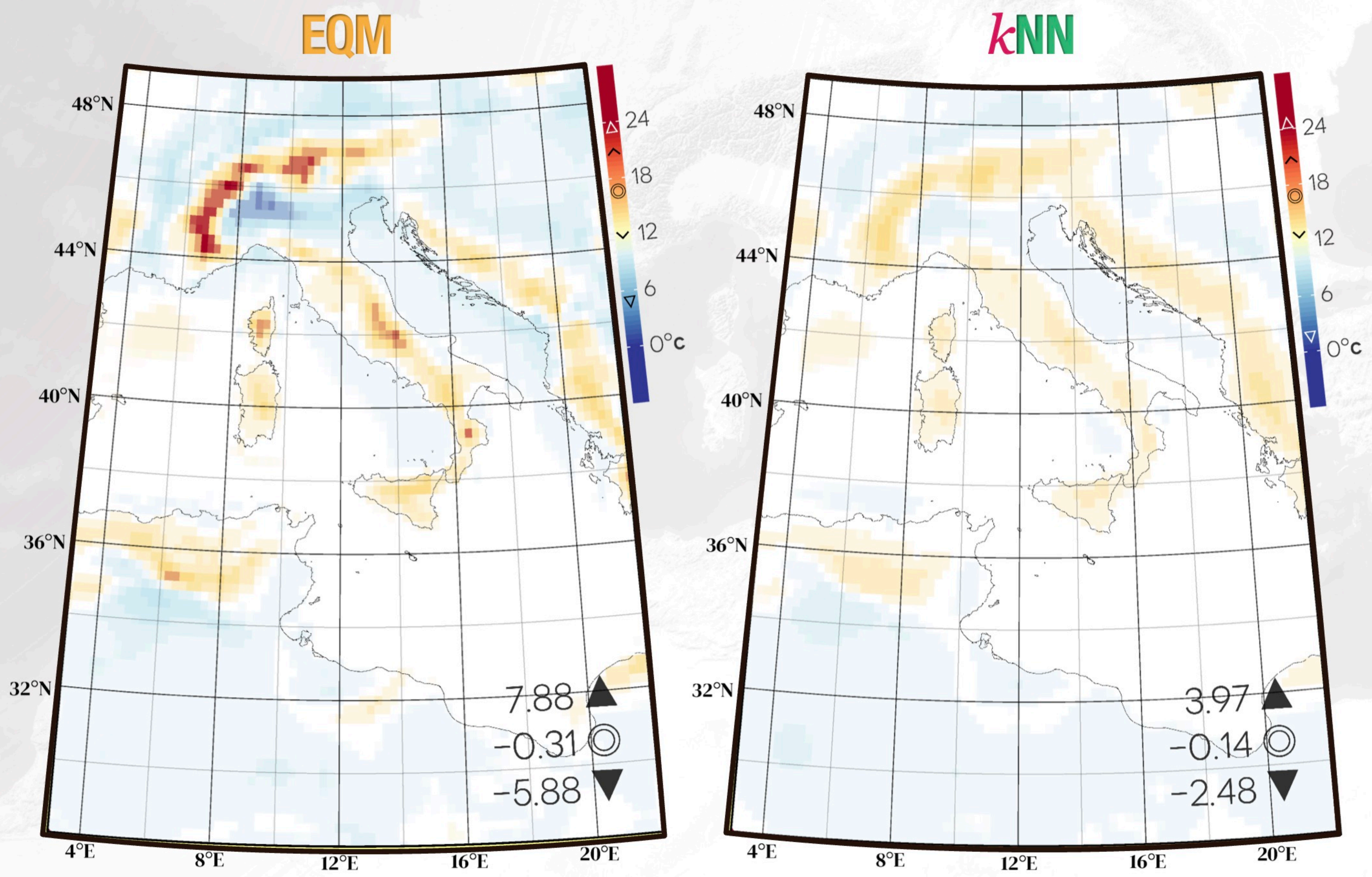
cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici 20 Years

SUMMARISED RESULTS

Ensemble annual mean (lead month 1) for the daily average 2m temperature (1993-2016)

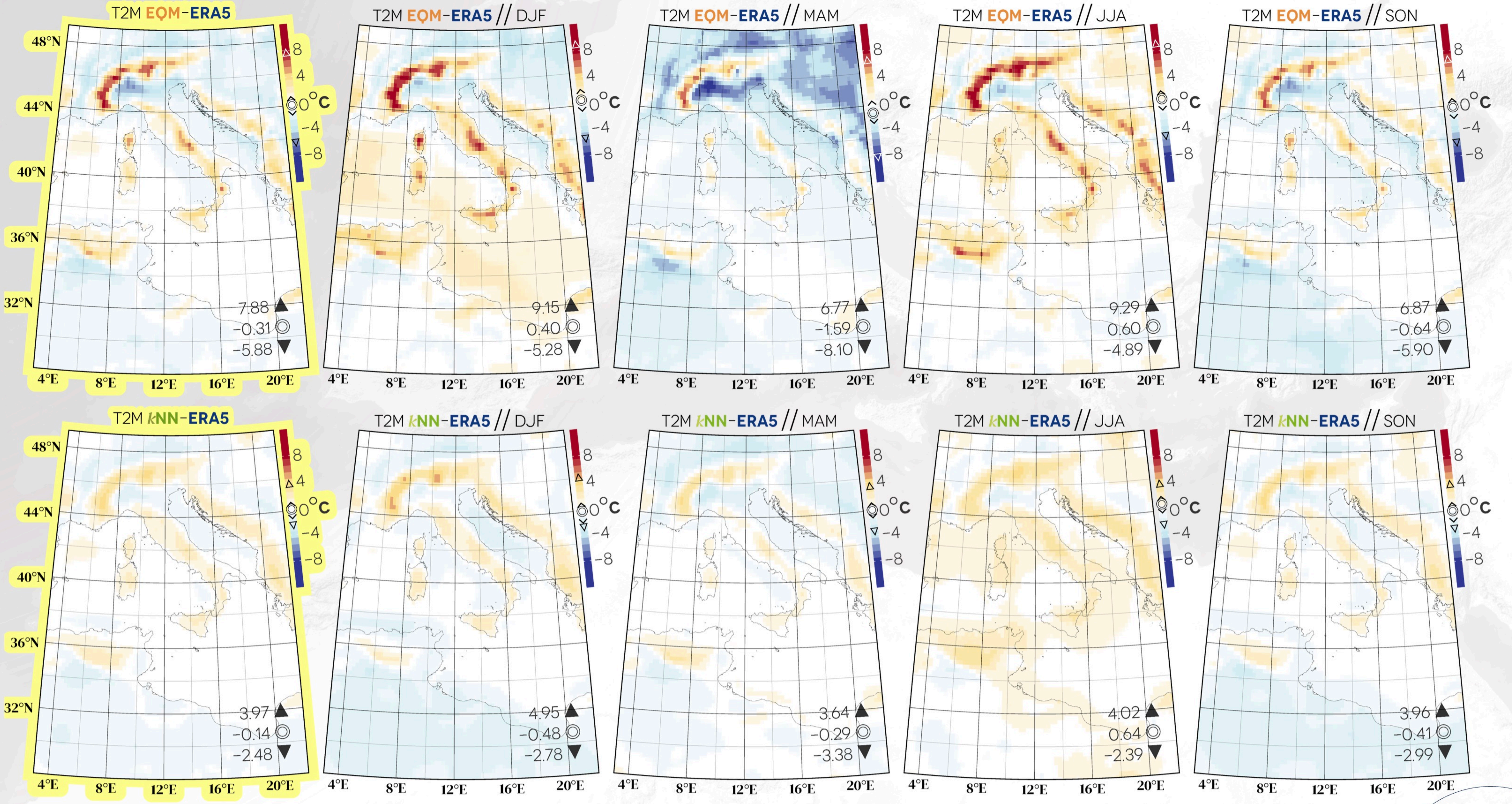
SUMMARISED RESULTS

Ensemble annual mean (lead month 1) for the daily average 2m temperature Bias (1993-2016)



SUMMARISED RESULTS

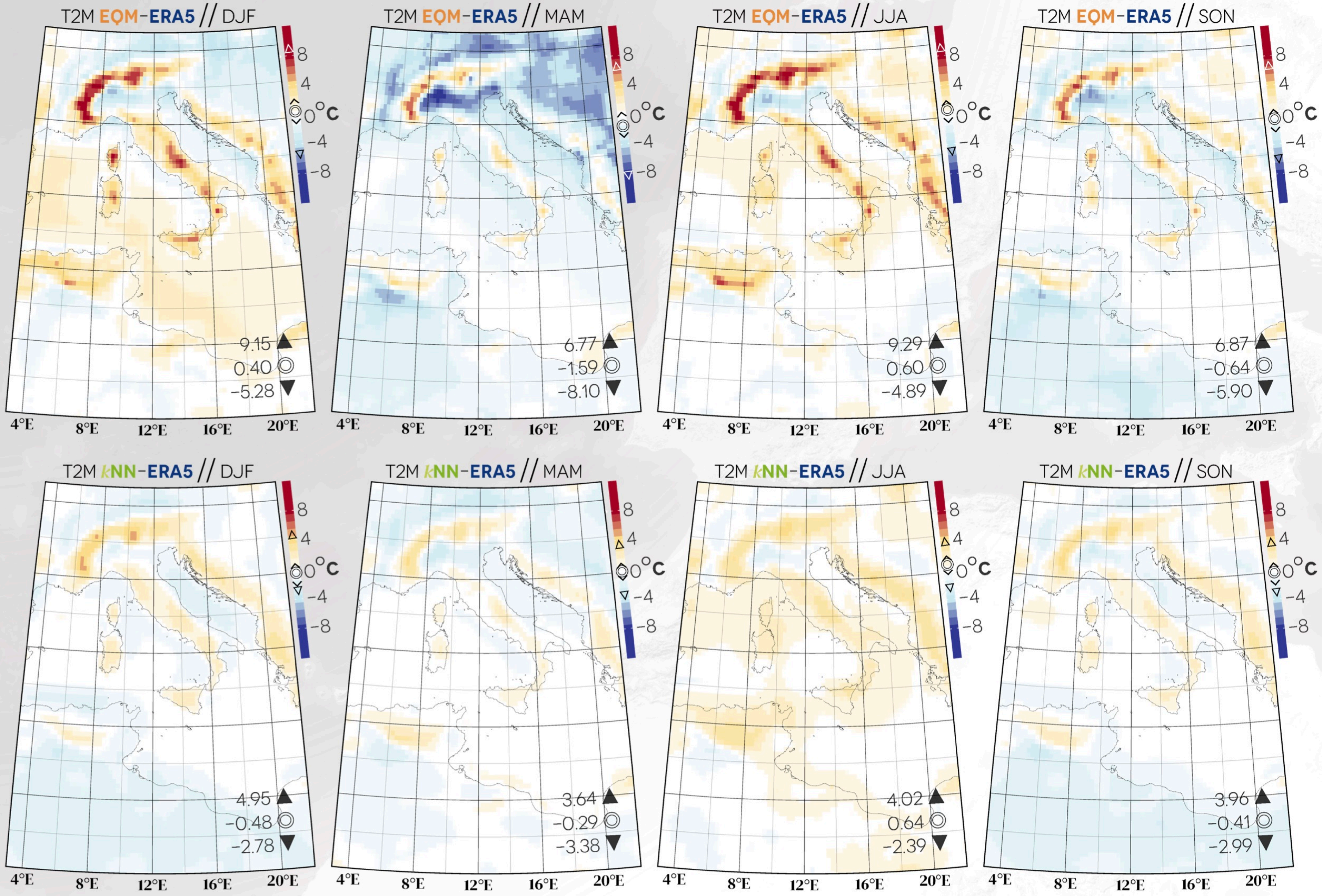
Ensemble Seasonal mean (lead month 1) for the daily average 2m temperature Bias (1993-2016)

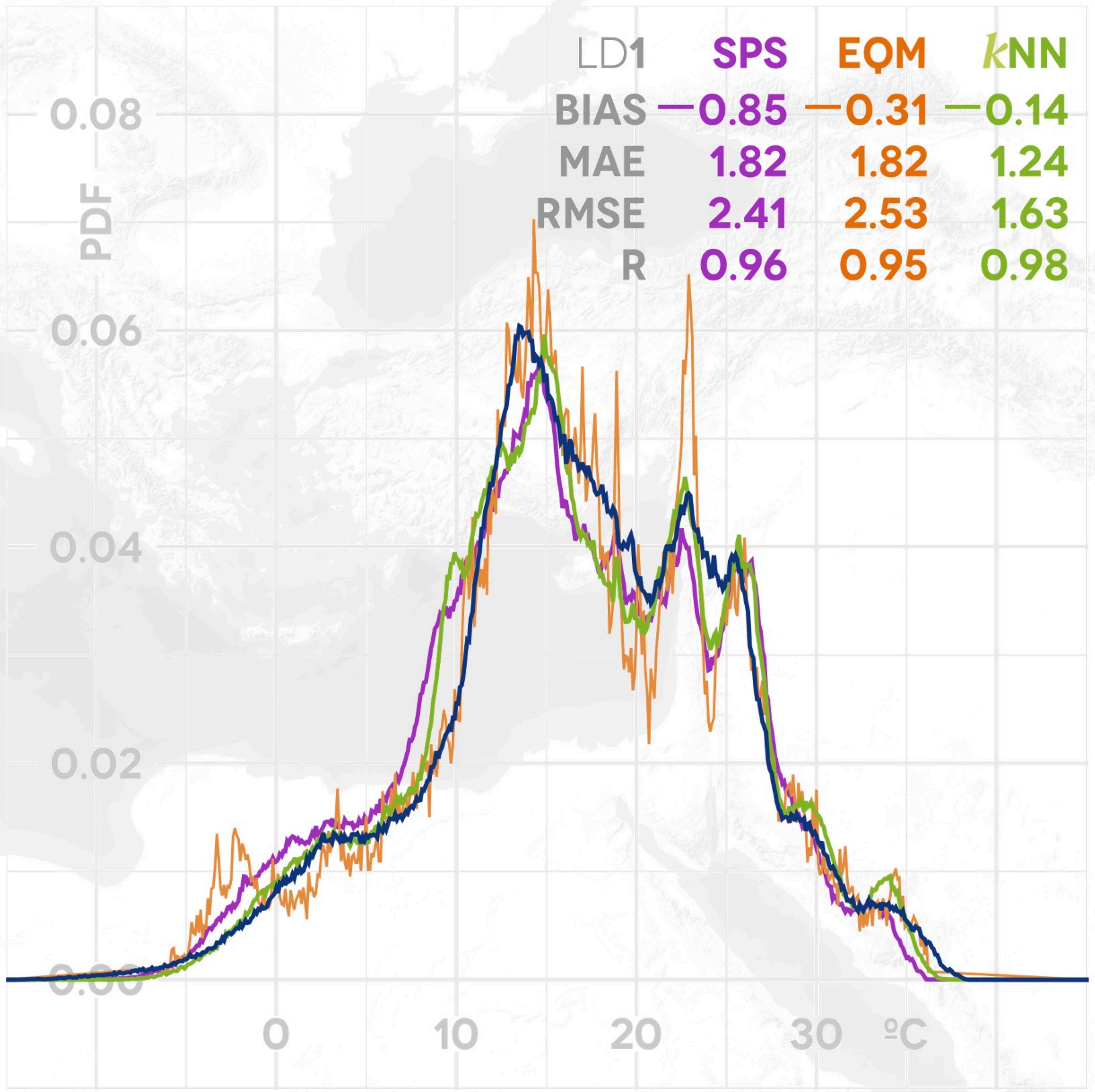


cmcc

SUMMARISED RESULTS

Ensemble Seasonal mean (lead month 1) for the daily average 2m temperature Bias (1993-2016)



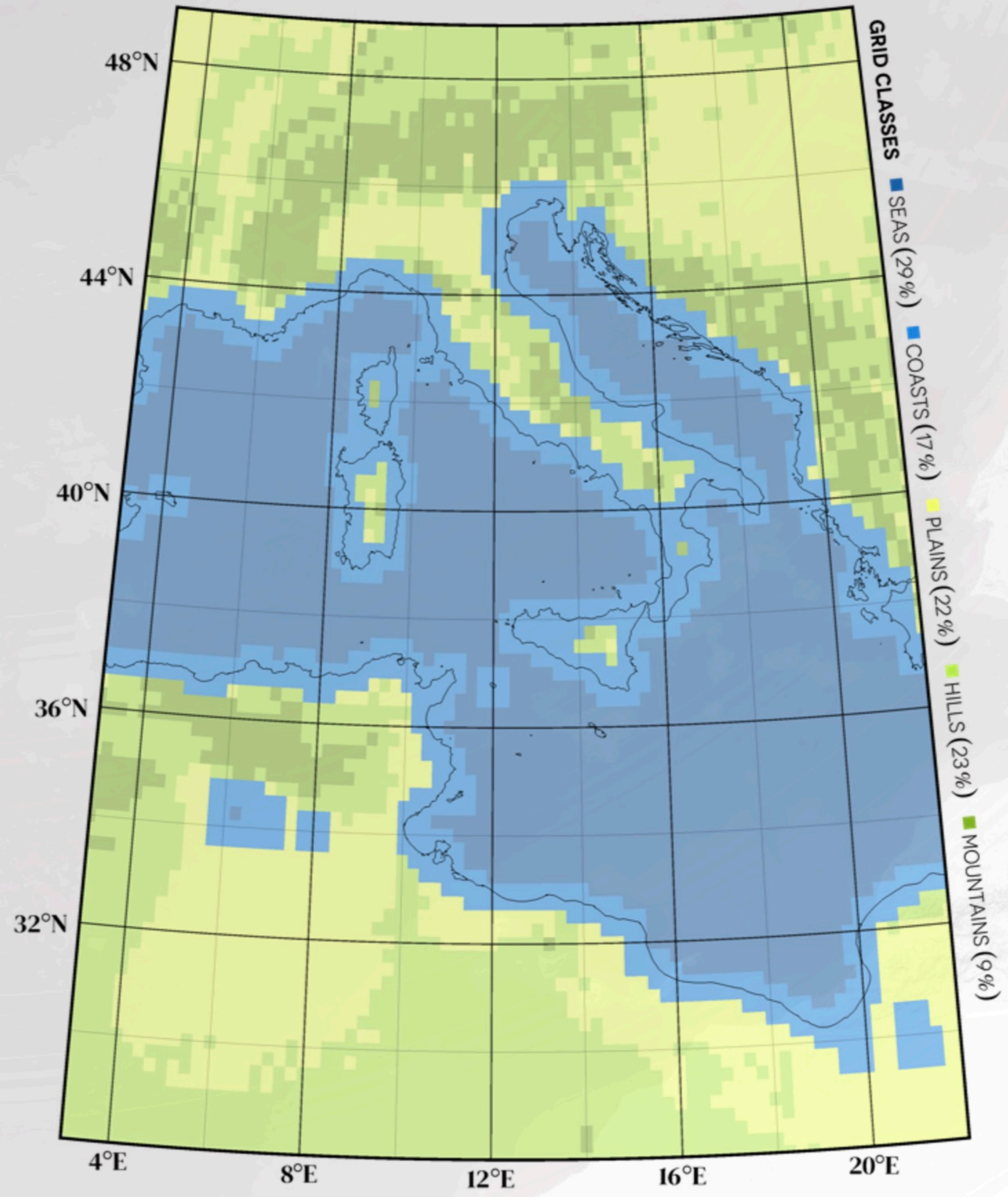


SUMMARISED RESULTS

Statistical indices by lead month of the daily average 2m temperature (1993-2016)

SUMMARISED RESULTS

Statistical indices by environment (lead month 1) of the daily average 2m temperature (1993-2016)



		DOMAIN	SEAS	COASTS	PLAINS	HILLS	MOUNTAINS
Bias	SPS	−0.85	−0.7	−1.12	−1.65	−1.03	1.66
	EQM	−0.31	−0.07	−0.25	−1.25	−0.62	1.96
	*k*NN	−0.14	−0.08	0.04	−0.72	−0.41	1.48
MAE	SPS	1.82	1.15	2.06	2.12	1.94	2.49
	EQM	1.82	1.03	1.67	2.22	2.11	3.05
	*k*NN	1.24	0.69	1.1	1.49	1.52	1.98
RMSE	SPS	2.41	1.67	2.74	2.66	2.38	3.15
	EQM	2.53	1.36	2.16	3.06	2.75	3.79
	*k*NN	1.63	0.9	1.39	1.84	1.89	2.44
R	SPS	0.96	0.97	0.94	0.98	0.98	0.95
	EQM	0.95	0.96	0.94	0.96	0.96	0.93
	*k*NN	0.98	0.98	0.98	0.98	0.98	0.97

SEAS
Gridpoints at the sea level (Altitude h = 0m)

COASTS
Gridpoints at the sea level (h = 0m) surrounded by at least one gridpoint with h > 0m, plus those surrounded by gridpoints at the sea level.

PLAINS
Gridpoins with 0m < h ≤ 300m

HILLS
Gridpoins with 300m < h ≤ 800m

MOUNTAINS
Gridpoins with h > 800m

SUMMARISED RESULTS

Cross-validation of the ensemble mean within the SPS3.5 hindcast period over the Italian domain (1993-2016)

		T2m	Prec	SST	U10	V10	T2N	T2X
Bias	SPS	-0.85	-4.39	-0.44	0.15	-0.04		
	EQM	-0.31	8.76	0.45	-0.06	0.13	-0.44	-0.20
	*k*NN	-0.14	-20.48	0.06	0.01	-0.06	-0.24	-0.08
MAE	SPS	1.82	19.89	0.78	0.81	0.72		
	EQM	1.82	27.35	0.89	0.86	0.78	1.92	1.95
	*k*NN	1.24	24.72	0.49	0.73	0.63	1.23	1.33
RMSE	SPS	2.41	33.81	1.15	1.09	0.95		
	EQM	2.53	42.91	1.12	1.17	1.05	2.62	2.74
	*k*NN	1.63	41.45	0.65	1.01	0.86	1.63	1.79
R	SPS	0.96	0.76	0.97	0.74	0.61		
	EQM	0.95	0.68	0.97	0.69	0.52	0.95	0.95
	*k*NN	0.98	0.73	0.99	0.77	0.69	0.98	0.98

CONCLUSIONS & NEXT STEPS

CONCLUSION

The analysis of spatial variability shows that k**NN** *produces more detailed and realistic* **T2m** *fields across Italy than* **EQM**. *Unlike* **EQM**'*s smooth and generalised fields,* k**NN** *captures key geographic features like the Po Valley and the Alps, aligning closely with* **ERA5**. *Statistically,* k**NN** *achieves better performance with a 0.98 correlation and 1.63°C RMSE, outperforming EQM.*

T2m *seasonal bias assessment reveals that* **EQM** *has higher winter biases (up to 8°C in the Alps), while* k**NN** *shows smoother and more accurate seasonal cycles.* k**NN** *better reflects* **ERA5** *data across all seasons, especially in marine regions.* **EQM**'*s apparent low marine bias is misleading, caused by seasonal over- and underestimations balancing each other.*

Environmental analysis across Italy shows that k**NN** *outperforms* **EQM** *in all regions, especially in coastal and mountainous areas where* **EQM** *fails to capture contrasts. Both methods perform better over marine areas, but* k**NN** *shows lower errors inland. The persistent bias in mountainous zones indicates challenges in representing cold extremes.*

In seasonal forecasting (up to 6 months), k**NN** *shows lower bias and RMSE than* **EQM** *and* **SPS**, *preserving initial signals while improving local-scale detail.* **EQM**'*s bias correction introduces uncertainty due to inter-member spread.* k**NN**'*s ensemble is more consistent, offering better reliability for climate-sensitive applications.*

Expanding these analyses to other variables yields a similar scenario. While k**NN** *tries to align with* **ERA5**, *preserving the initial conditions from* **SPS**, **EQM** *shows greater spread and less physical coherence, abruptly forcing* **SPS** *towards the observations. Thus,* k**NN** *better replicates observed variability in spatially structured environments.*

Precipitation downscaling remains difficult for statistical methods. **EQM** *overestimates light rain and misses seasonal variability, especially in cold months and orographic zones.* k**NN** *better captures spatial patterns in spring and autumn but underrepresents convective summer extremes. These challenges demand more sophisticated inputs and frameworks.*

Due to analogs averaging, k**NN** *struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions. Minor errors in analogs selection lead to significant rainfall underestimations. Hybrid methods and convective indices could enhance* k**NN**'*s ability to capture such extremes.*

When used as a weather generator, k**NN** *accurately reproduces daily temperature extremes (***T2N** *and* **T2X***), aligning closely with* **ERA5** *in terms of spread and skewness, unlike* **EQM**. *These results confirm the value of* k**NN**'*s multivariate coherence and spatial sensitivity. However, improving its treatment of extremes remains key for broader applicability.*

CONCLUSION

> *Unlike EQM's smooth and generalised fields, kNN captures key geographic features like the Po Valley and the Alps, aligning closely with ERA5.*

CONCLUSION

The analysis of spatial variability shows that **kNN** *produces more detailed and realistic* **T2m** *fields across Italy than* **EQM**. *Unlike* **EQM**'*s smooth and generalised fields,* **kNN** *captures key geographic features like the Po Valley and the Alps, aligning closely with* **ERA5**. *Statistically,* **kNN** *achieves better performance with a 0.98 correlation and 1.63°C RMSE, outperforming EQM.*

T2m *seasonal bias assessment reveals that* **EQM** *has higher winter biases (up to 8°C in the Alps), while* **kNN** *shows smoother and more accurate seasonal cycles.* **kNN** *better reflects* **ERA5** *data across all seasons, especially in marine regions.* **EQM**'*s apparent low marine bias is misleading, caused by seasonal over- and underestimations balancing each other.*

Environmental analysis across Italy shows that **kNN** *outperforms* **EQM** *in all regions, especially in coastal and mountainous areas where* **EQM** *fails to capture contrasts. Both methods perform better over marine areas, but* **kNN** *shows lower errors inland. The persistent bias in mountainous zones indicates challenges in representing cold extremes.*

In seasonal forecasting (up to 6 months), **kNN** *shows lower bias and RMSE than* **EQM** *and* **SPS**, *preserving initial signals while improving local-scale detail.* **EQM**'*s bias correction introduces uncertainty due to inter-member spread.* **kNN**'*s ensemble is more consistent, offering better reliability for climate-sensitive applications.*

Expanding these analyses to other variables yields a similar scenario. While **kNN** *tries to align with* **ERA5**, *preserving the initial conditions from* **SPS**, **EQM** *shows greater spread and less physical coherence, abruptly forcing* **SPS** *towards the observations. Thus,* **kNN** *better replicates observed variability in spatially structured environments.*

Precipitation downscaling remains difficult for statistical methods. **EQM** *overestimates light rain and misses seasonal variability, especially in cold months and orographic zones.* **kNN** *better captures spatial patterns in spring and autumn but underrepresents convective summer extremes. These challenges demand more sophisticated inputs and frameworks.*

Due to analogs averaging, **kNN** *struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions. Minor errors in analogs selection lead to significant rainfall underestimations. Hybrid methods and convective indices could enhance* **kNN**'*s ability to capture such extremes.*

When used as a weather generator, **kNN** *accurately reproduces daily temperature extremes (***T2N** *and* **T2X***), aligning closely with* **ERA5** *in terms of spread and skewness, unlike* **EQM**. *These results confirm the value of* **kNN**'*s multivariate coherence and spatial sensitivity. However, improving its treatment of extremes remains key for broader applicability.*

CONCLUSION

> *While kNN tries to align with ERA5, preserving the initial conditions from SPS, EQM shows greater spread and less physical coherence, abruptly forcing SPS towards the observations.*

CONCLUSION

The analysis of spatial variability shows that kNN *produces more detailed and realistic* T2m *fields across Italy than* EQM. *Unlike* EQM*'s smooth and generalised fields,* kNN *captures key geographic features like the Po Valley and the Alps, aligning closely with* ERA5. *Statistically,* kNN *achieves better performance with a 0.98 correlation and 1.63°C RMSE, outperforming EQM.*

T2m *seasonal bias assessment reveals that* EQM *has higher winter biases (up to 8°C in the Alps), while* kNN *shows smoother and more accurate seasonal cycles.* kNN *better reflects* ERA5 *data across all seasons, especially in marine regions.* EQM*'s apparent low marine bias is misleading, caused by seasonal over- and underestimations balancing each other.*

Environmental analysis across Italy shows that kNN *outperforms* EQM *in all regions, especially in coastal and mountainous areas where* EQM *fails to capture contrasts. Both methods perform better over marine areas, but* kNN *shows lower errors inland. The persistent bias in mountainous zones indicates challenges in representing cold extremes.*

In seasonal forecasting (up to 6 months), kNN *shows lower bias and RMSE than* EQM *and* SPS, *preserving initial signals while improving local-scale detail.* EQM*'s bias correction introduces uncertainty due to inter-member spread.* kNN*'s ensemble is more consistent, offering better reliability for climate-sensitive applications.*

Expanding these analyses to other variables yields a similar scenario. While kNN *tries to align with* ERA5, *preserving the initial conditions from* SPS, EQM *shows greater spread and less physical coherence, abruptly forcing* SPS *towards the observations. Thus,* kNN *better replicates observed variability in spatially structured environments.*

Precipitation downscaling remains difficult for statistical methods. EQM *overestimates light rain and misses seasonal variability, especially in cold months and orographic zones.* kNN *better captures spatial patterns in spring and autumn but underrepresents convective summer extremes. These challenges demand more sophisticated inputs and frameworks.*

Due to analogs averaging, kNN *struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions. Minor errors in analogs selection lead to significant rainfall underestimations. Hybrid methods and convective indices could enhance* kNN*'s ability to capture such extremes.*

When used as a weather generator, kNN *accurately reproduces daily temperature extremes (*T2N *and* T2X*), aligning closely with* ERA5 *in terms of spread and skewness, unlike* EQM. *These results confirm the value of* kNN*'s multivariate coherence and spatial sensitivity. However, improving its treatment of extremes remains key for broader applicability.*

CONCLUSION

"EQM overestimates light rain and misses seasonal variability, especially in cold months and orographic zones. *k*NN better captures spatial patterns in spring and autumn but underrepresents convective summer extremes."

CONCLUSION

The analysis of spatial variability shows that **kNN** *produces more detailed and realistic* **T2m** *fields across Italy than* **EQM**. *Unlike* **EQM***'s smooth and generalised fields,* **kNN** *captures key geographic features like the Po Valley and the Alps, aligning closely with* **ERA5**. *Statistically,* **kNN** *achieves better performance with a 0.98 correlation and 1.63°C RMSE, outperforming EQM.*

T2m *seasonal bias assessment reveals that* **EQM** *has higher winter biases (up to 8°C in the Alps), while* **kNN** *shows smoother and more accurate seasonal cycles.* **kNN** *better reflects* **ERA5** *data across all seasons, especially in marine regions.* **EQM***'s apparent low marine bias is misleading, caused by seasonal over- and underestimations balancing each other.*

Environmental analysis across Italy shows that **kNN** *outperforms* **EQM** *in all regions, especially in coastal and mountainous areas where* **EQM** *fails to capture contrasts. Both methods perform better over marine areas, but* **kNN** *shows lower errors inland. The persistent bias in mountainous zones indicates challenges in representing cold extremes.*

In seasonal forecasting (up to 6 months), **kNN** *shows lower bias and RMSE than* **EQM** *and* **SPS**, *preserving initial signals while improving local-scale detail.* **EQM***'s bias correction introduces uncertainty due to inter-member spread.* **kNN***'s ensemble is more consistent, offering better reliability for climate-sensitive applications.*

Expanding these analyses to other variables yields a similar scenario. While **kNN** *tries to align with* **ERA5**, *preserving the initial conditions from* **SPS**, **EQM** *shows greater spread and less physical coherence, abruptly forcing* **SPS** *towards the observations. Thus,* **kNN** *better replicates observed variability in spatially structured environments.*

Precipitation downscaling remains difficult for statistical methods. **EQM** *overestimates light rain and misses seasonal variability, especially in cold months and orographic zones.* **kNN** *better captures spatial patterns in spring and autumn but underrepresents convective summer extremes. These challenges demand more sophisticated inputs and frameworks.*

Due to analogs averaging, **kNN** *struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions. Minor errors in analogs selection lead to significant rainfall underestimations. Hybrid methods and convective indices could enhance* **kNN***'s ability to capture such extremes.*

When used as a weather generator, **kNN** *accurately reproduces daily temperature extremes (***T2N** *and* **T2X***), aligning closely with* **ERA5** *in terms of spread and skewness, unlike* **EQM**. *These results confirm the value of* **kNN***'s multivariate coherence and spatial sensitivity. However, improving its treatment of extremes remains key for broader applicability.*

CONCLUSION

> *Due to analogs averaging, kNN struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions.*

CONCLUSION

The analysis of spatial variability shows that kNN *produces more detailed and realistic* T2m *fields across Italy than* EQM. *Unlike* EQM*'s smooth and generalised fields,* kNN *captures key geographic features like the Po Valley and the Alps, aligning closely with* ERA5. *Statistically,* kNN *achieves better performance with a 0.98 correlation and 1.63°C RMSE, outperforming EQM.*

T2m *seasonal bias assessment reveals that* EQM *has higher winter biases (up to 8°C in the Alps), while* kNN *shows smoother and more accurate seasonal cycles.* kNN *better reflects* ERA5 *data across all seasons, especially in marine regions.* EQM*'s apparent low marine bias is misleading, caused by seasonal over- and underestimations balancing each other.*

Environmental analysis across Italy shows that kNN *outperforms* EQM *in all regions, especially in coastal and mountainous areas where* EQM *fails to capture contrasts. Both methods perform better over marine areas, but* kNN *shows lower errors inland. The persistent bias in mountainous zones indicates challenges in representing cold extremes.*

In seasonal forecasting (up to 6 months), kNN *shows lower bias and RMSE than* EQM *and* SPS, *preserving initial signals while improving local-scale detail.* EQM*'s bias correction introduces uncertainty due to inter-member spread.* kNN*'s ensemble is more consistent, offering better reliability for climate-sensitive applications.*

Expanding these analyses to other variables yields a similar scenario. While kNN *tries to align with* ERA5, *preserving the initial conditions from* SPS, EQM *shows greater spread and less physical coherence, abruptly forcing* SPS *towards the observations. Thus,* kNN *better replicates observed variability in spatially structured environments.*

Precipitation downscaling remains difficult for statistical methods. EQM *overestimates light rain and misses seasonal variability, especially in cold months and orographic zones.* kNN *better captures spatial patterns in spring and autumn but underrepresents convective summer extremes. These challenges demand more sophisticated inputs and frameworks.*

Due to analogs averaging, kNN *struggles with extremes, often underestimating cold days and intense summer precipitation, especially in mountainous regions. Minor errors in analogs selection lead to significant rainfall underestimations. Hybrid methods and convective indices could enhance* kNN*'s ability to capture such extremes.*

When used as a weather generator, kNN *accurately reproduces daily temperature extremes (*T2N *and* T2X*), aligning closely with* ERA5 *in terms of spread and skewness, unlike* EQM. *These results confirm the value of* kNN*'s multivariate coherence and spatial sensitivity. However, improving its treatment of extremes remains key for broader applicability.*

CONCLUSION

When used as a weather generator, kNN accurately reproduces daily temperature extremes (T2N and T2X), aligning closely with ERA5 in terms of spread and skewness, unlike EQM.

NEXT STEPS

SCIENTIFIC PUBLICATION
The full version of this experiment considering the whole set of variables, analysis and discussions

OPERATIONAL VERSION
Integrate the ***k*****NN** *tool to* **SPS4** *to provide statistical downscaled predictions monthly*

MedCORDEX DOMAIN
6 times faster is not enough

TOWARDS EXTREMES
Review the list of predictors to include large-scale variables and enhance the ***k*****NN** *skill to a daily level*

ARAGÃO, Borrelli & Gualdi (2025)
CMCC Seasonal Prediction System v3.5 hindcasts downscaled from 1-degree to 0.25-degree using Empirical Quantile Mapping and k-Nearest Neighbours.
CMCC Data Delivery System, DOI 10.25424/cmcc-4brr-ym18

ARAGÃO, Borrelli & Gualdi (2025) ▫ **Under Review**
Expanding CMCC Seasonal Prediction System v3.5 applications to the local scale through statistical downscaling techniques.
Earth System Science Data, DOI 10.5194/essd-2025-371

cmcc | 20 Years

Centro Euro-Mediterraneo sui Cambiamenti Climatici

Expanding CMCC Seasonal Prediction System v3.5 applications to the local scale through

STATISTICAL DOWNSCALING TECHNIQUES

Thank you!

Leonardo Aragão, Andrea Borrelli & Silvio Gualdi

leonardo.aragao@cmcc.it

www.cmcc.it