16 July 2025

EPESC – LEADER Science Meeting (15-18 July 2025)

APEC Climate Center, Busan, Republic of Korea

WCRP World Climate Research Programme

Exploring Multi-year Predictability of Terrestrial Heatwaves in Global Hotspot Regions

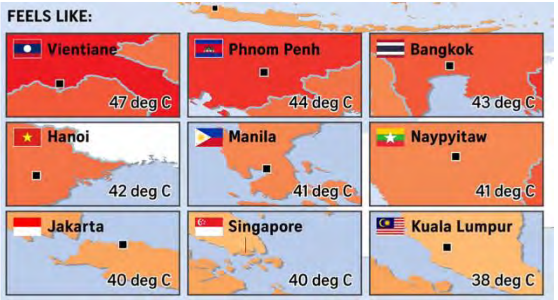
Alexia Karwat[1], June-Yi Lee[1,2], Yong-Yub Kim[2], Jeong-Eun Yun[1], and Sun-Seon Lee[2]

[1] Research Center for Climate Sciences, Pusan National University, Busan, Republic of Korea
[2] IBS Center for Climate Physics, Busan, Republic of Korea

Greenland

Central America

Arabian Peninsula



Southeast Asia

부산대학교 PUSAN NATIONAL UNIVERSITY

Research Center for Climate Sciences (RCCS)

contact:
alexia.karwat@pusan.ac.kr

Motivation

- Terrestrial heatwaves pose significant risks to ecosystems, human health, and socio-economies, **with limited understanding of their multi-year predictability in relation to energy, electricity and cooling demand** *(IPCC AR6 report; Zhang et al. 2024)*

- Prediction of statistics (e.g., frequency) over **multi-year time scales** remains challenging due to the complex interactions between **internal variability, large-scale drivers, and local processes** *(Hamilton et al. 2011; Luo et al. 2020; Kim et al., 2025 preprint)*

- **Sources** of multi-year predictability are not well understood for regional hotspots *(e.g., Qasmi et al. 2020; Pyrina and Domeisen 2022; Boisseson and Balmaseda 2023)*

→ Improving our understanding of the underlying mechanisms can enhance long-term heatwave forecasts, guide climate adaptation strategies, and inform proactive measures to mitigate risks in vulnerable regions

CESM2 Multi-year Climate Prediction System (CESM2-MP)

- **it consists of ocean assimilations, 5-year hindcasts and uninitialized large-ensemble historical simulations (CESM2-LE)** *(Kim et al. 2025, preprint)*

- atmospheric component: Community Atmosphere Model version 6 (CAM6) *(Danabasoglu et al. 2020)*

- ocean and sea ice models: Parallel Ocean Program version 2 (POP2) *(Smith et al. 2010)* and CICE version 5.1.2 (CICE5) *(Bailey et al. 2020)*

- **external forcing: historical and SSP3-7.0 warming scenario based on CMIP6** *(van Marle et al. 2017, Rodgers et al. 2021)*

- *here:* 50-member CESM2 Large Ensemble

Overview of Data

- 2m air temperature (TREFHT → model / t2m in observation)
- 2m relative humidity data (RH2M in model and observation; assimilated over land)

data set	original resolution		common resolution
hindcast (HIND)	0.94	1.25	1.25° x 1.25°
CESM2-LE (UNIN)	0.94	1.25	
ERA5	0.25	0.25	
AgERA5	0.1	0.1	

- we consider the boreal spring and summer season during 1981-2020 (daily resolution)

How can we define “Predictive Skill“?

1) Total Skill: HIND (LY1-5)$_{anomaly}$ vs. OBS$_{anomaly}$ → *compute correlation*

2) Skill from Forcing: UNIN vs. OBS

$$\left.\boxed{\begin{array}{l} \text{UNIN}_{\,anomaly}\,(t) = \text{UNIN}\,(t) - \text{UNIN}\,(t)_{seasonal} \\ \text{OBS}_{\,anomaly}\,(t) = \text{OBS}\,(t) - \text{OBS}\,(t)_{seasonal} \end{array}}\ \right\} \textit{as above}$$

3) Skill from Internal Variability: HIND (LY1-5) – LE vs. OBS – LE

$$\left.\boxed{\begin{array}{l} \text{HIND}_{\,int.\,var.}\,(t) = \text{HIND}_{\,anomaly}\,(t) - \text{UNIN}\,(t)_{seasonal} \\ \text{OBS}_{\,int.\,var.}\,(t) = \text{OBS}_{\,anomaly}\,(t) - \text{UNIN}\,(t)_{seasonal} \end{array}}\ \right\} \textit{as above}$$

How can we define “Predictive Skill“?

1) Total Skill: HIND (LY1-5)$_{anomaly}$ vs. OBS$_{anomaly}$ → *compute correlation*

2) Skill from Forcing: UNIN vs. OBS

$$\left.\begin{aligned} \text{UNIN}_{\,anomaly}(t) &= \text{UNIN}(t) - \text{UNIN}(t)_{seasonal} \\ \text{OBS}_{\,anomaly}(t) &= \text{OBS}(t) - \text{OBS}(t)_{seasonal} \end{aligned}\right\} \textit{as above}$$

3) Skill from Internal Variability: HIND (LY1-5) – LE vs. OBS – LE

$$\left.\begin{aligned} \text{HIND}_{\,int.\,var.}(t) &= \text{HIND}_{\,anomaly}(t) - \text{UNIN}(t)_{seasonal} \\ \text{OBS}_{\,int.\,var.}(t) &= \text{OBS}_{\,anomaly}(t) - \text{UNIN}(t)_{seasonal} \end{aligned}\right\} \textit{as above}$$

Classifying Terrestrial Heatwaves: Thermal, Dry, and Wet

Thermal Heatwave (THW) = 3 consecutive days ≥ 90th percentile of 2mT

(e.g., Perkins 2015; Domeisen et al. 2022)

Stricter criteria:

Dry Heatwave (DHW) = 3 consecutive days ≥ 90th pct. of 2mT & ≤ 33% of RH2M

Wet Heatwave (WHW) = 3 consecutive days ≥ 90th pct. of 2mT & ≥ 66% of RH2M

apply skill concept

(e.g., Ha, Seo et al. 2022)

Classifying Terrestrial Heatwaves: Thermal, Dry, and Wet

Thermal Heatwave (THW) = 3 consecutive days ≥ 90th percentile of 2mT

(e.g., Perkins 2015; Domeisen et al. 2022)

Stricter criteria:

Dry Heatwave (DHW) = 3 consecutive days ≥ 90th pct. of 2mT & ≤ 33% of RH2M

Wet Heatwave (WHW) = 3 consecutive days ≥ 90th pct. of 2mT & ≥ 66% of RH2M

apply skill concept

(e.g., Ha, Seo et al. 2022)

Hypothesis: Thermal and humidity-driven heatwave types **enable predictable cooling demand extremes** across Northern Hemisphere hotspots

Cooling Degree Days and Cooling Demand

Dry Cooling Degree Days = The sum of degrees by which **daily temperatures** exceed 22°C, reflecting cooling demand.

$CDD_{dry} = T_{2m,mean} - T_{2m,base}$

(e.g., Ember Energy UK)

Wet Cooling Degree Days = The sum of degrees by which **daily wet-bulb temperatures** exceed 24°C, reflecting cooling demand.

$CDD_{wet} = T_{wb,mean} - T_{wb,base}$

Cooling Demand Index = α x heatwave frequency$_{type}$ + β x CDD$_{type}$

→ apply skill concept

Cooling Degree Days and Cooling Demand

Dry Cooling Degree Days = The sum of degrees by which **daily temperatures** exceed 22°C, reflecting cooling demand.

$CDD_{dry} = T_{2m,mean} - T_{2m,base}$

(e.g., Ember Energy UK)

Wet Cooling Degree Days = The sum of degrees by which **daily wet-bulb temperatures** exceed 24°C, reflecting cooling demand.

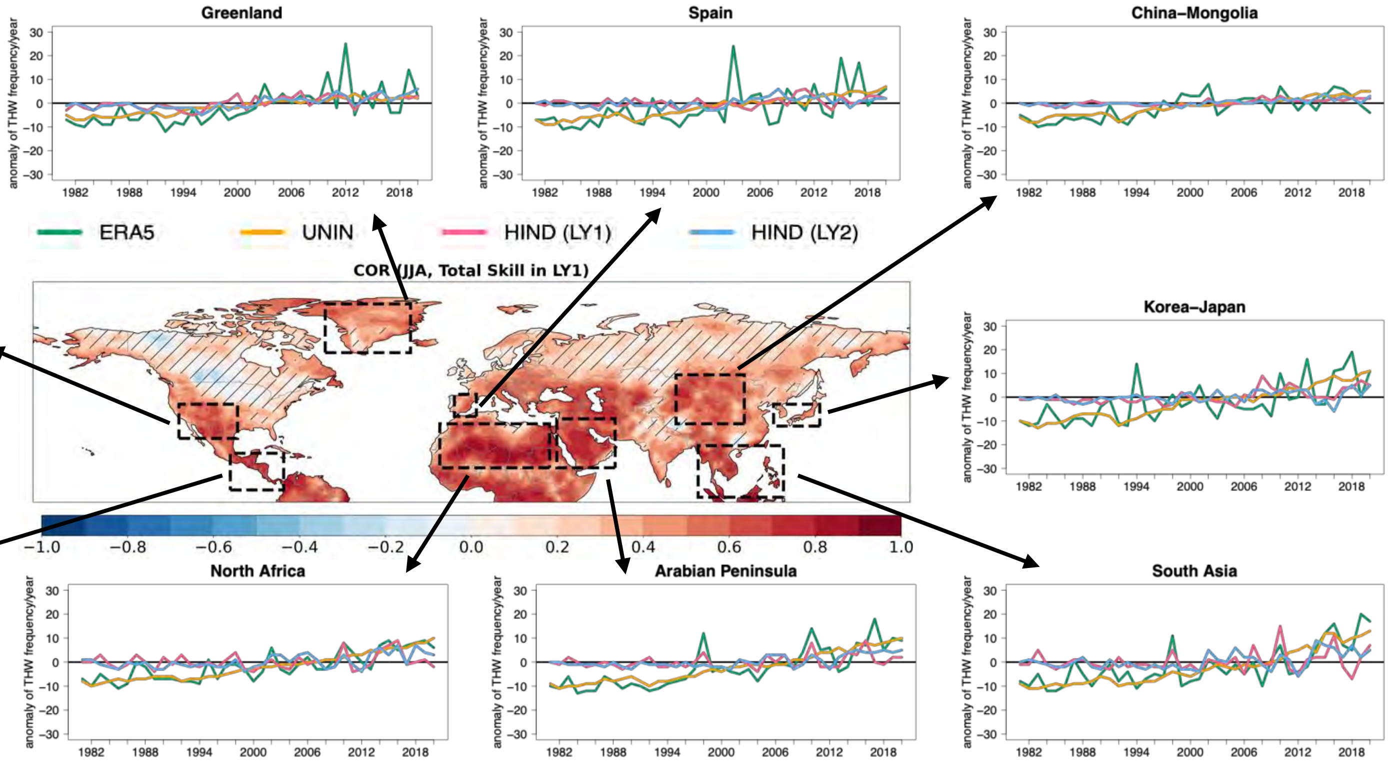
$CDD_{wet} = T_{wb,mean} - T_{wb,base}$

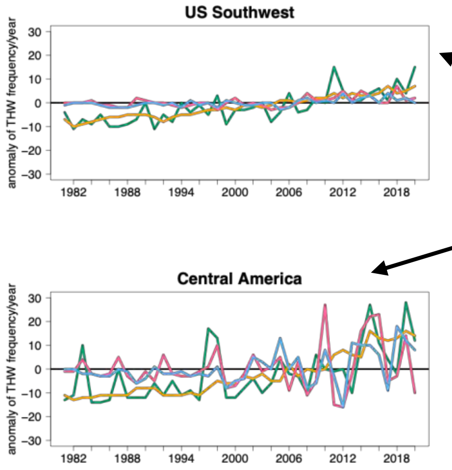
Cooling Demand Index = α x heatwave frequency$_{thermal}$ + β x CDD$_{dry}$

→ apply skill concept

Anomaly of Thermal Heatwave Frequency (1981-2020, JJA)

- Increase in THW frequency
- UNIN captures overall trend
- HIND LY1-2 occasionally better predicts peaks (e.g., 1998 ENSO)

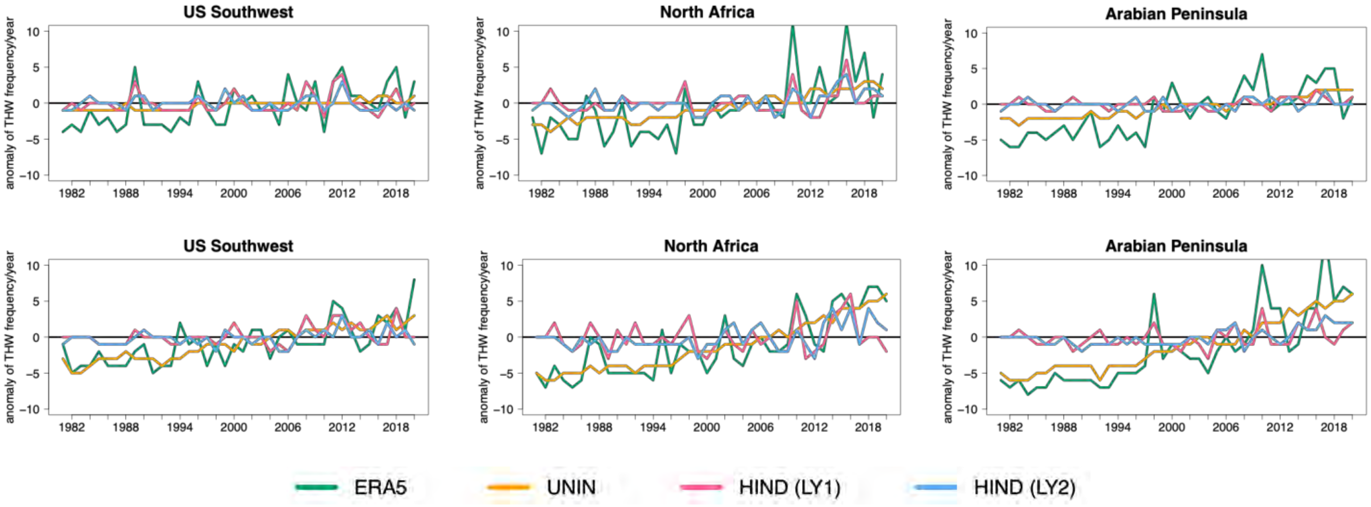




Anomaly of Dry Heatwave Frequency (1981-2020)

MAM

JJA



Strong increasing trends in all 3 regions and both seasons, **mostly evident after 2000**

Good model agreement on the general trend, but the magnitude varies

MAM: A clear increase in dry heatwave frequency since around 2000

JJA: Less strong upward trend but still a slight increase, model agreement seems generally better

Both seasons show a gradual increase in dry heatwave frequency anomalies, with more variability and stronger upward trend after 2000

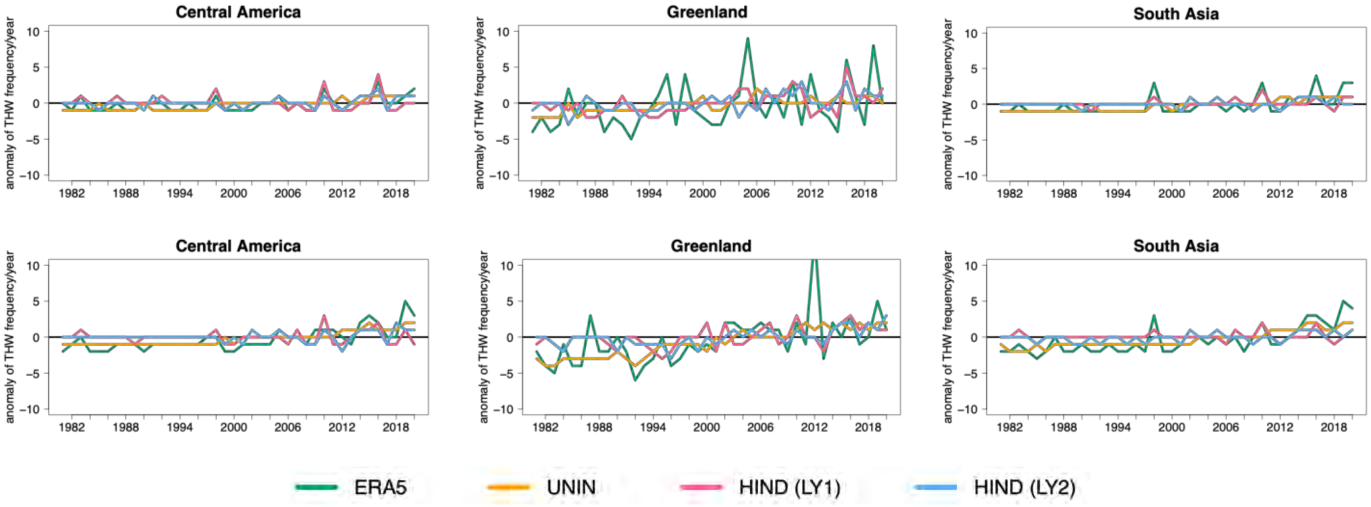
MAM: Large interannual variability with a strong increase after 2000

JJA: Similar trend but smoother and more consistent rise

Anomaly of Wet Heatwave Frequency (1981-2020)

MAM

JJA



Less clear, weaker long-term trends compared to dry heatwaves

Greenland shows episodic increases, while **Central America and South Asia show more recent extremes**

MAM: Generally weak trends in anomalies up to about 2010, followed by a slight increase afterward

JJA: Mostly flat until 2010, then slight upward trend, but less pronounced than for dry heatwaves

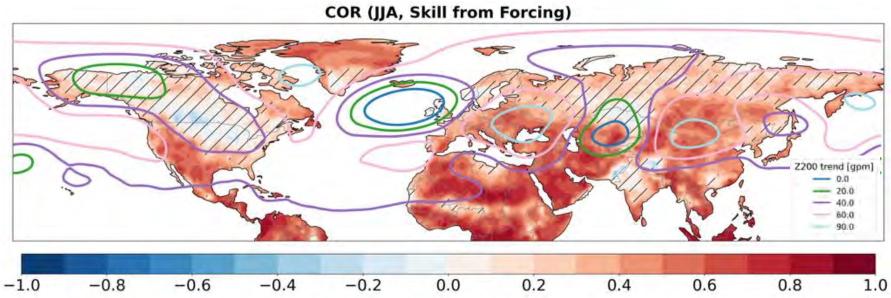
MAM: Large interannual variability

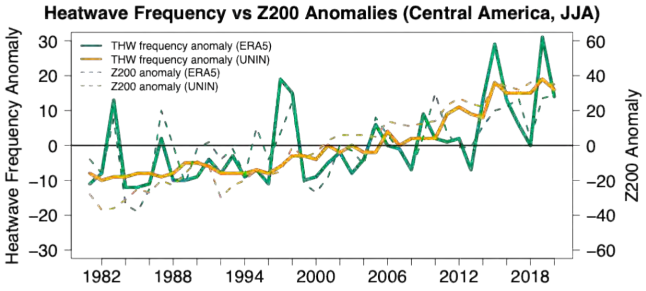
JJA: Noticeable variability and an upward trend in recent years

MAM: Nearly no change before 2000, then slight increase in anomalies

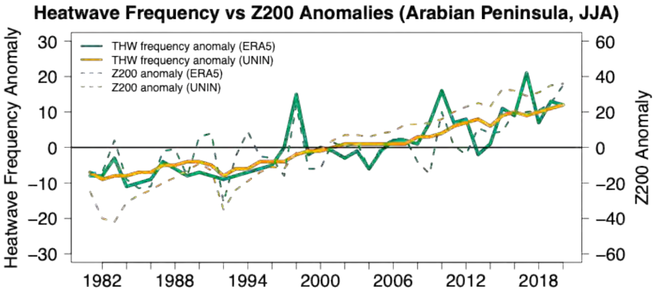
JJA: Similar to MAM

Skill from External Forcing: Frequency of Thermal Heatwaves & Trend in Z200 (JJA)





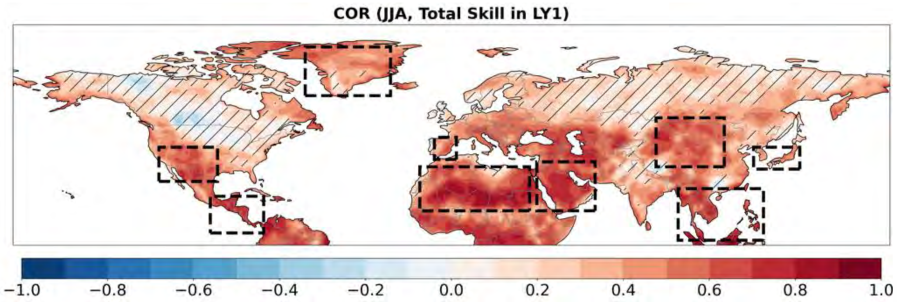
Anomaly Correlation Coefficient, 1981-2020, UNIN and ERA5



- Heatwave frequency anomalies show a strong correlation with atmospheric circulation anomalies (Z200)
- Higher correlation for hotspots in tropics and subtropics (e.g., Central America, Arabian Peninsula, US Southwest etc.)
- Anthropogenic influence rather dominant across many regions

Heatwaves, Cooling Degree Days & Cooling Demand (JJA)

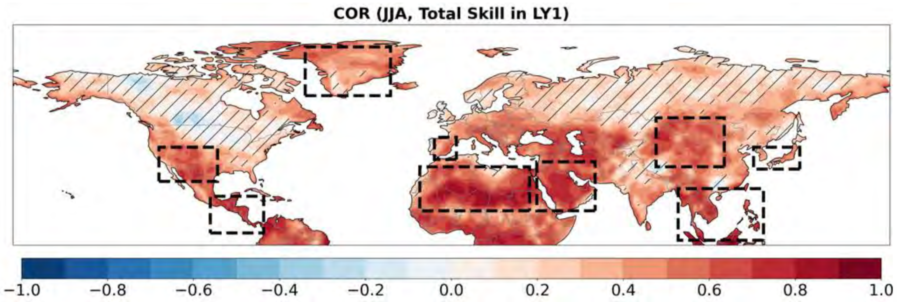
Can we predict **Cooling Demand** based on heatwave frequency and Cooling Degree Days (CDDs) information?



Anomaly Correlation Coefficient of Thermal Heatwave Frequency 1981-2020 HIND and ERA5

Heatwaves, Cooling Degree Days & Cooling Demand (JJA)

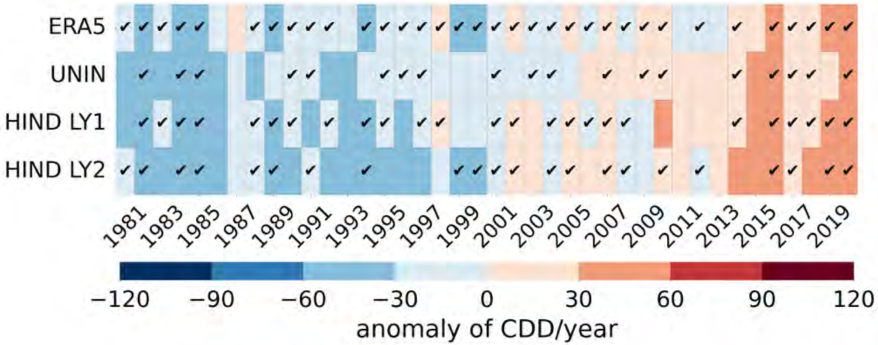
Can we predict **Cooling Demand** based on heatwave frequency and Cooling Degree Days (CDDs) information?



Anomaly Correlation Coefficient of Thermal Heatwave Frequency 1981-2020 HIND and ERA5

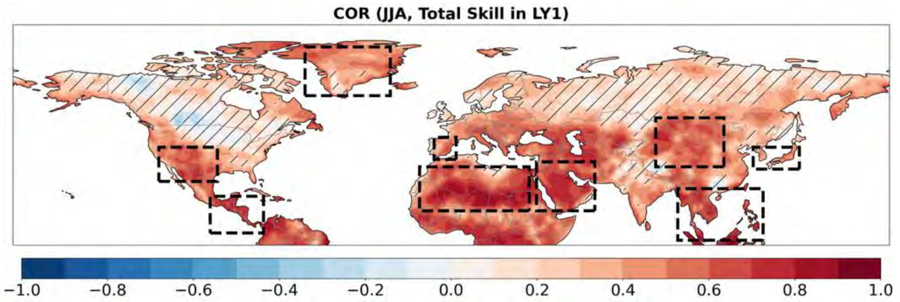
Heatmap of the Anomaly of CDD_{dry}, South Asia

(34/40 seasons -> 85% agreement)



Heatwaves, Cooling Degree Days & Cooling Demand (JJA)

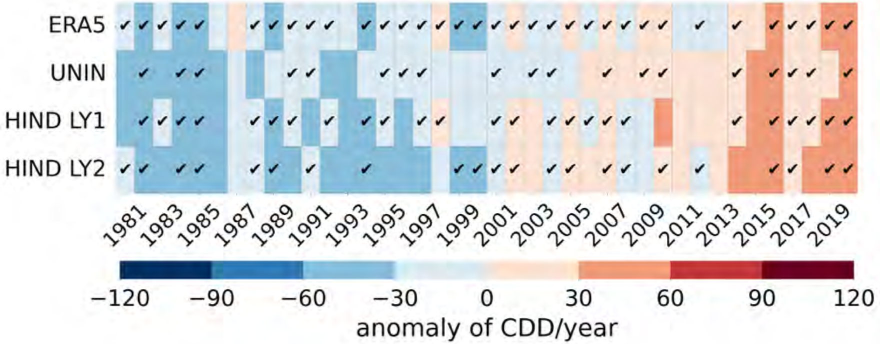
Can we predict **Cooling Demand** based on heatwave frequency and Cooling Degree Days (CDDs) information?



Anomaly Correlation Coefficient of Thermal Heatwave Frequency 1981-2020 HIND and ERA5

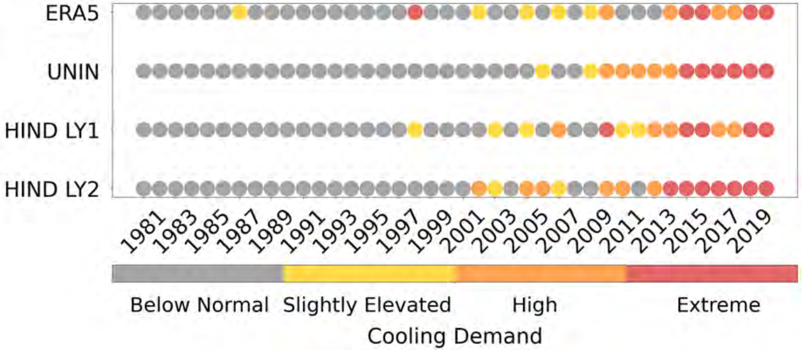
Heatmap of the Anomaly of CDD_{dry}, South Asia

(34/40 seasons -> 85% agreement)



Cooling Demand, South Asia

(35/40 seasons -> 87.5% agreement)



How skillful are the CDD_{dry} predictions?

Hotspot	MAM	JJA	Difference
US Southwest	87.5%	47.5%	Much higher in MAM (+40%)
Central America	72.5%	72.5%	Same in both seasons
Greenland	-	-	No data
Spain	82.5%	62.5%	Higher in MAM (+20%)
North Africa	62.5%	60%	Slightly higher in MAM (+2.5%)
Arabian Peninsula	57.5%	75%	**Higher in JJA (+17.5%)**
China-Mongolia	70%	67.5%	Slightly higher in MAM (+2.5%)
South Asia	72.5%	85%	**Higher in JJA (+12.5%)**
Korea-Japan	85%	42.5%	Much higher in MAM (+42.5%)

Overall, most hotspots show higher skill in MAM than in JJA, except for the Arabian Peninsula and South Asia, where skill is higher in JJA.

How skillful are the Cooling Demand predictions?

Hotspot	MAM	JJA	Difference
US Southwest	82.5%	80%	Slightly higher in MAM (+2.5%)
Central America	85%	80%	Slightly higher in MAM (+5%)
Greenland	82.5%	82.5%	Same in both seasons
Spain	70%	67.5%	Slightly higher in MAM (+2.5%)
North Africa	65%	62.5%	Slightly higher in MAM (+2.5%)
Arabian Peninsula	70%	90%	**Much higher in JJA (+20%)**
China-Mongolia	75%	57.5%	**Higher in MAM (+17.5%)**
South Asia	77.5%	87.5%	**Higher in JJA (+10%)**
Korea-Japan	87.5%	62.5%	**Much higher in MAM (+25%)**

Overall, most hotspots have slightly higher skill in MAM than in JJA, except for the Arabian Peninsula and South Asia, where skill is higher in JJA. Korea-Japan and China-Mongolia show particularly strong MAM advantages.

Summary & Conclusions

Thermal Heatwaves and dry CDDs are predictable on multiyear timescales:

→ **External forcing is a strong driver of multi-year predictability** of heatwaves in many hotspot regions.

→ **Dry heatwaves are more predictable** since trends are stronger, more widespread, and better captured in the model:

e.g., the US Southwest, which has been affected by severe droughts and wildfires in recent years.

→ **Wet heatwaves are less predictable**, since trends are weaker and less consistent, with regional differences and greater uncertainty. Particularly relevant for heat stress predictability across South(east) Asia and Central America.

→ **Internal variability is limited to 1-2 years** and constraint by ENSO predictability (12-14 months in CESM2-MP).

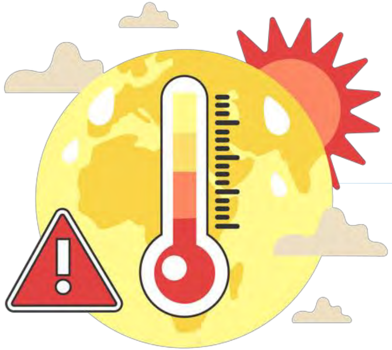
→ **Skilful prediction of the cooling demand during THWs** helps prevent power outages and improve energy management.

Outlook

- Examine prediction skills of **CDD_{wet} + cooling demand from strictly dry/wet heatwaves**
- Explore concept of **sudden day2day temperature changes**

- Correlation with (NH) blocking
- Linkage with SST anomalies / Marine Heatwaves *(Karwat et al. 2025, under review)*
- Urban Heat Island (UHI) effect

- Consider the predictability of other extremes in the CESM2-MP and hybrid/AI model approaches, e.g., for drought, malaria, heavy precipitation and storm prediction.



Thank you very much for your attention!

감사합니다

<u>Contact:</u>

alexia.karwat@pusan.ac.kr

Appendix

Spherical Convolutional Wasserstein Distance (SCWD)

- SCWD is **a similarity measure to validate climate models** by comparing the distance between a climate model and the observed data *(Garrett et al. 2024)*

- concept based on the global mean **Wasserstein distance** *(Vissio et al. 2020)*

- **SCWD evaluates the distributions of spatial fields while taking into account localized extreme events:**

  a convolution slicer takes a weighted mean of data around each location to calculate local distances that are then incorporated in the computation of the SCWD *(Garrett et al. 2024)*

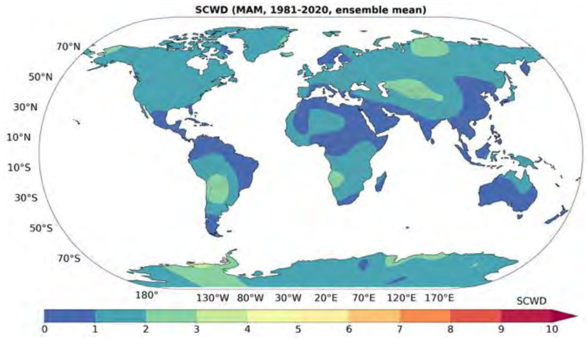
- **low values indicate that the climate model is similar to the observation**

→ we define a critical SCWD = 3 as the distance where the climate model is considerably different from the observation

Climate Model Validation: 2m Temperature

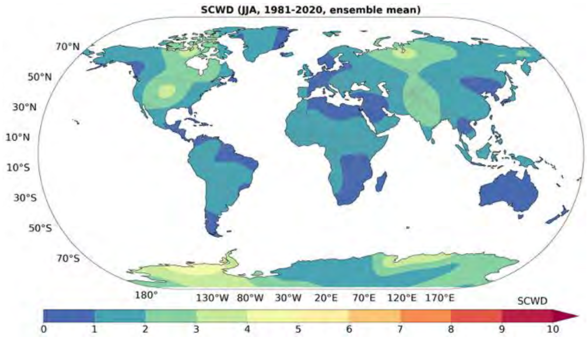
UNIN *TREFHT* vs ERA5 Reanalysis *t2m*

MAM



→ high global similarity up to 3

JJA



→ more differences over, e.g.,
central North America,
West Antarctica, and
Northwest-central Siberia (up to 5)

Spherical Convolutional Wasserstein Distance (SCWD)

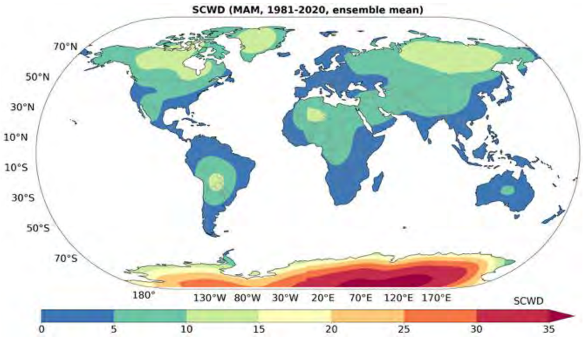
1981-2020

UNIN and ERA5

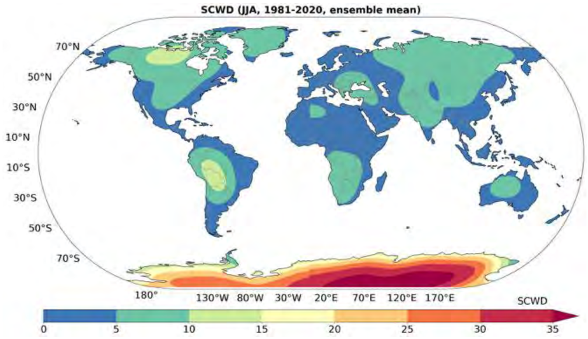
Climate Model Validation: 2m Relative Humidity

UNIN *RH2M* vs AgERA5 Reanalysis *RH2M*

MAM



JJA



Spherical Convolutional Wasserstein Distance (SCWD)

1981-2020

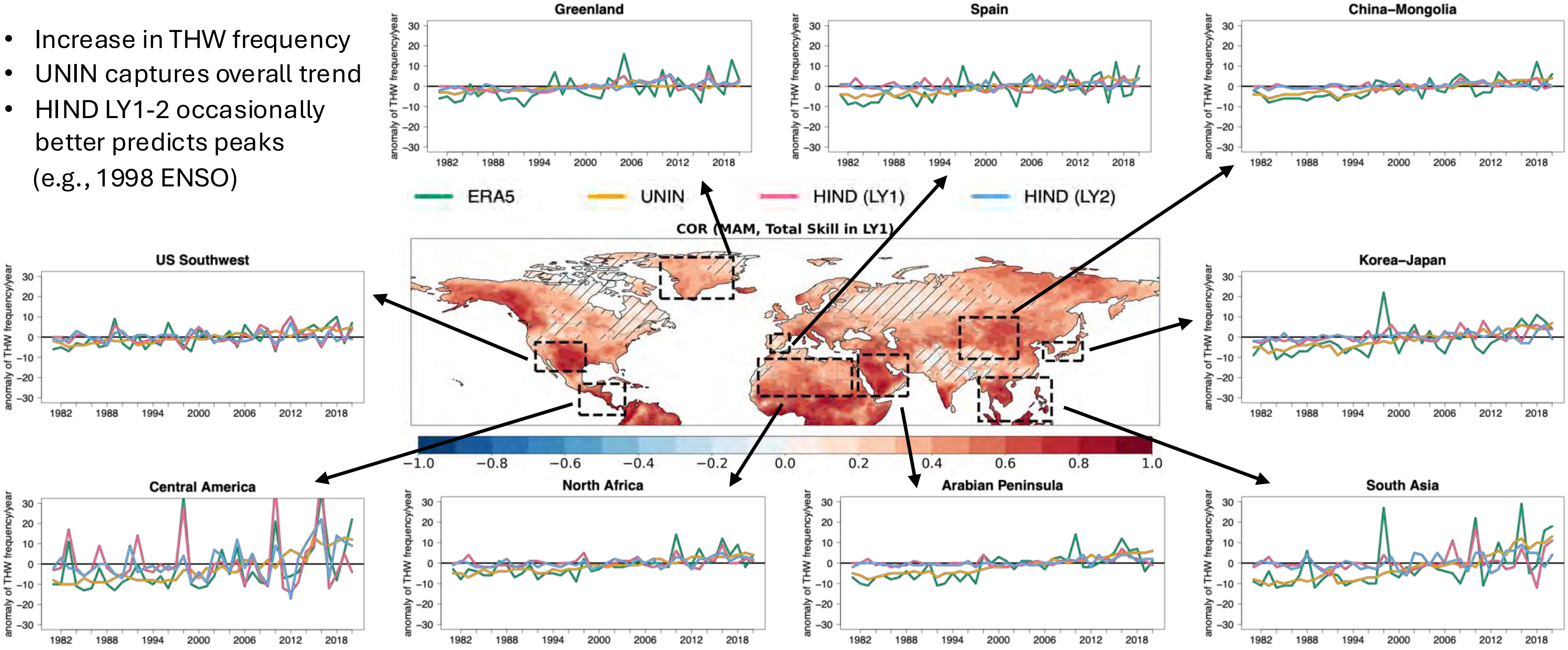
UNIN and AgERA5

→ high similarity over most regions except for Antarctica

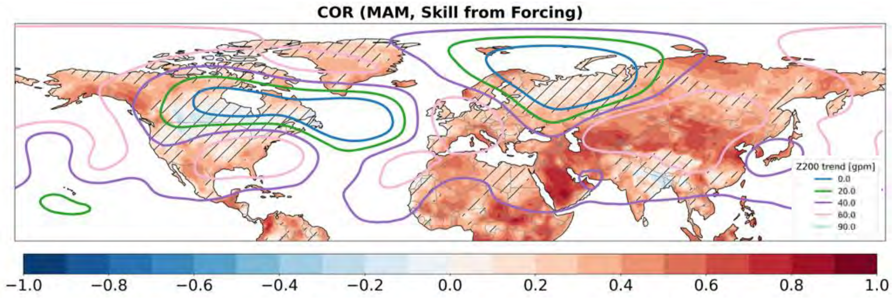
→ similar to MAM, with slightly higher similarity over North Africa and Siberia

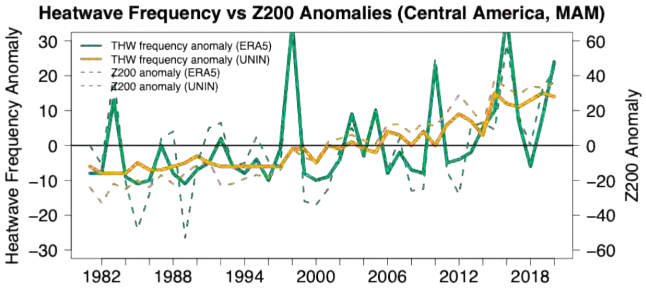
Anomaly of Thermal Heatwave Frequency (1981-2020, MAM)

- Increase in THW frequency
- UNIN captures overall trend
- HIND LY1-2 occasionally better predicts peaks (e.g., 1998 ENSO)

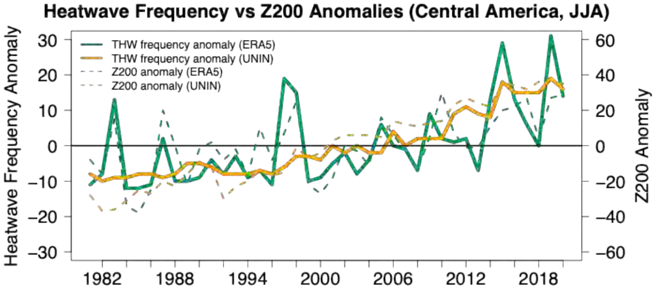


Skill from External Forcing: Frequency of Thermal Heatwaves & Trend in Z200 (MAM)





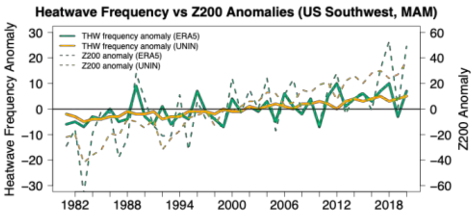
Anomaly Correlation Coefficient, 1981-2020, UNIN and ERA5



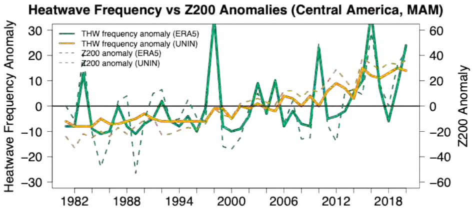
- Heatwave frequency anomalies show a strong correlation with atmospheric circulation anomalies (Z200)
- Higher correlation for hotspots in tropics and subtropics (e.g., Central America, Arabian Peninsula, US Southwest etc.)
- Anthropogenic influence rather dominant across many regions

Heatwave Frequency vs Z200 Anomalies (MAM)

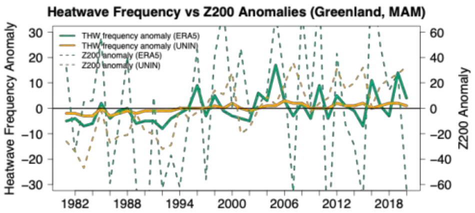
US Southwest



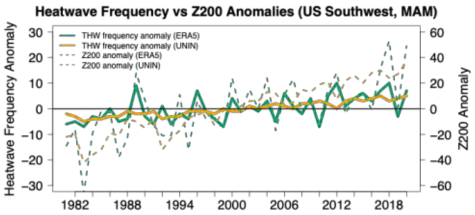
Central America



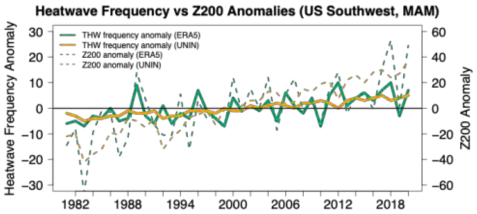
Greenland



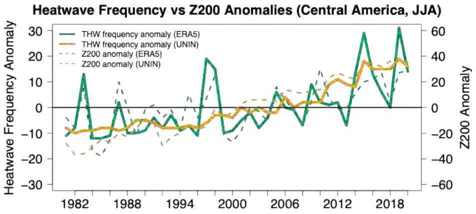
Spain



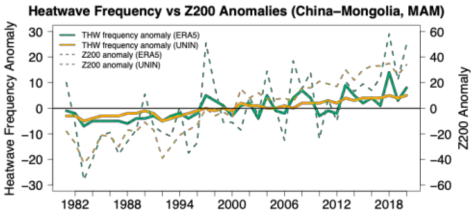
North Africa



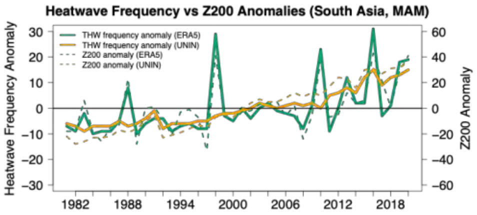
Arabian Peninsula



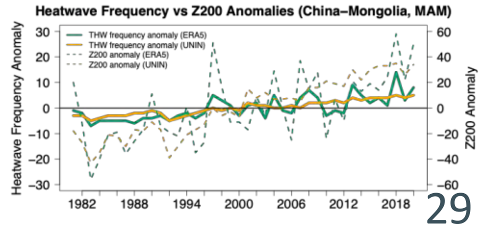
China-Mongolia



South Asia

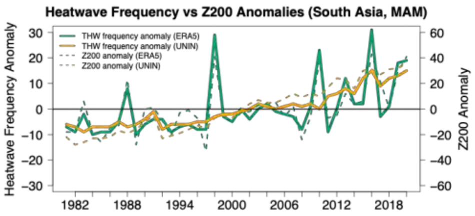


Korea-Japan

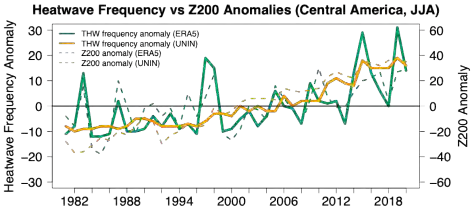


Heatwave Frequency vs Z200 Anomalies (JJA)

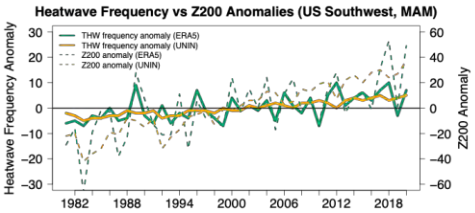
US Southwest



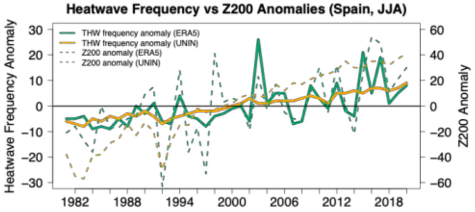
Central America



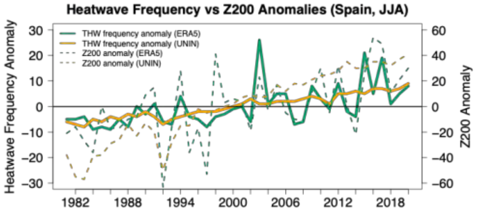
Greenland



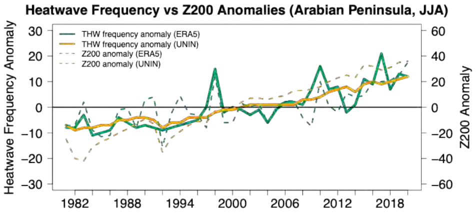
Spain



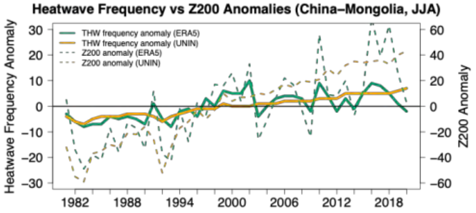
North Africa



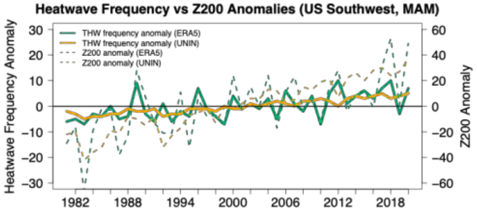
Arabian Peninsula



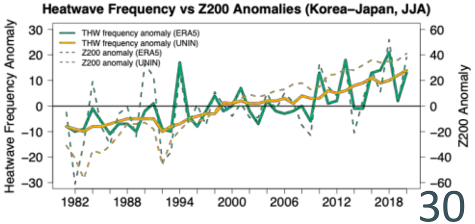
China-Mongolia



South Asia

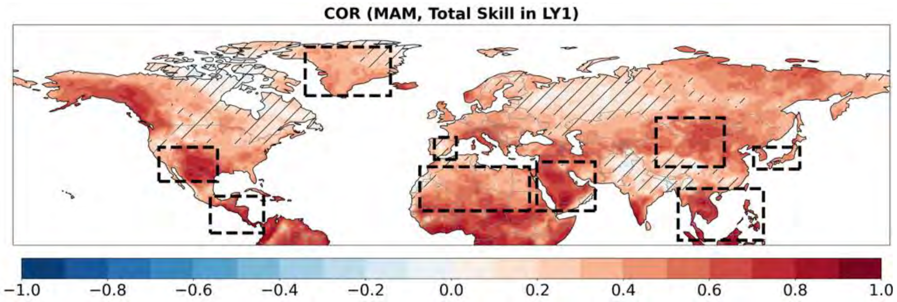


Korea-Japan



Heatwaves, Cooling Degree Days & Cooling Demand (MAM)

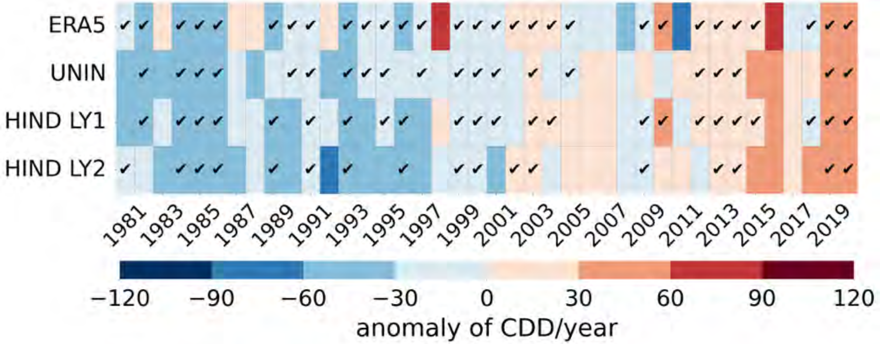
Can we predict **Cooling Demand** based on heatwave frequency and Cooling Degree Days (CDDs) information?



Anomaly Correlation Coefficient of Thermal Heatwave Frequency 1981-2020 HIND and ERA5

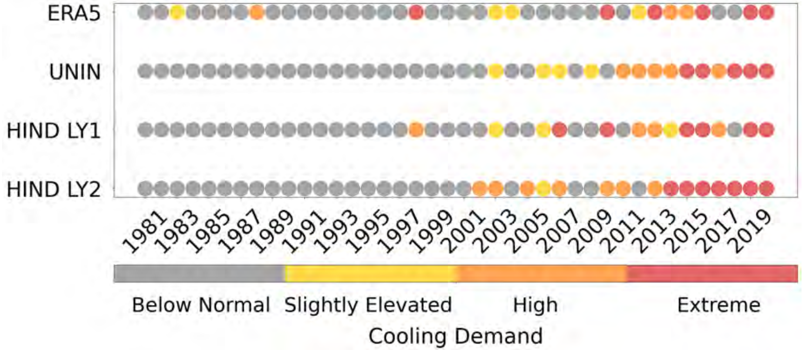
Heatmap of the Anomaly of CDD_{dry}, South Asia

(29/40 seasons -> 72.5% agreement)



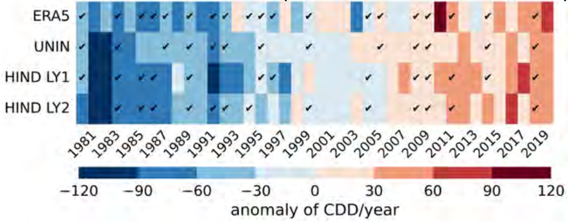
Cooling Demand, South Asia

(31/40 seasons -> 77.5% agreement)

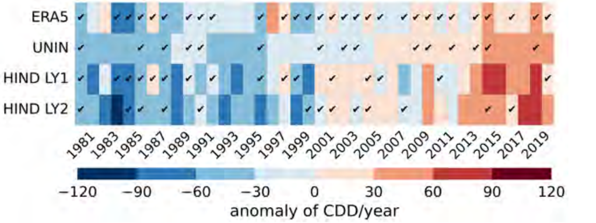


Heatmaps of the Anomaly of CDD_{dry} (JJA)

US Southwest (19/40 seasons -> 47.5%)



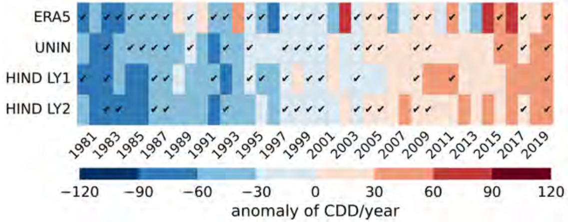
Central America (29/40 seasons -> 72.5%)



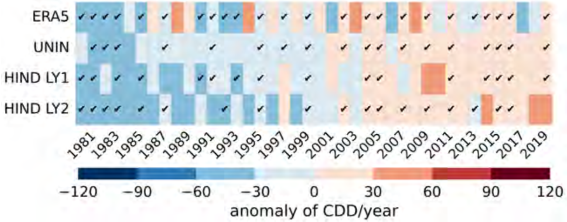
Greenland

No cooling demand since the average temperature (often below 10°C) is lower than the base threshold.

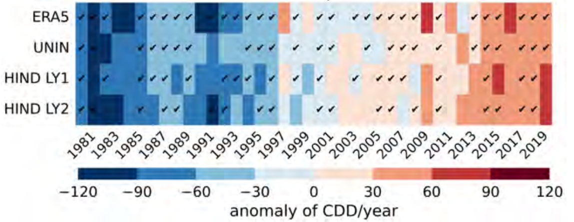
Spain (25/40 seasons -> 62.5%)



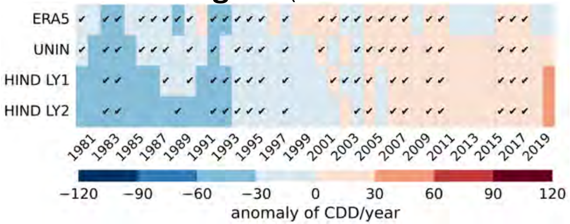
North Africa (24/40 seasons -> 60%)



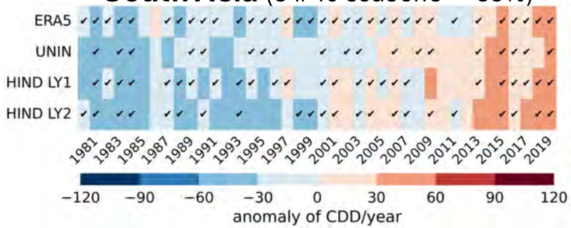
Arabian Peninsula (30/40 seasons -> 75%)



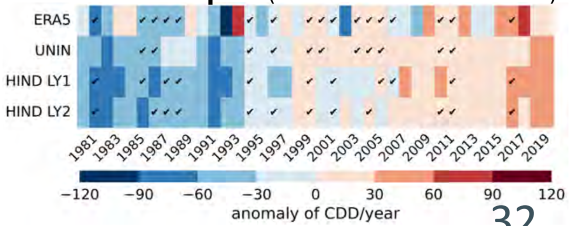
China-Mongolia (27/40 seasons -> 67.5%)



South Asia (34/40 seasons -> 85%)

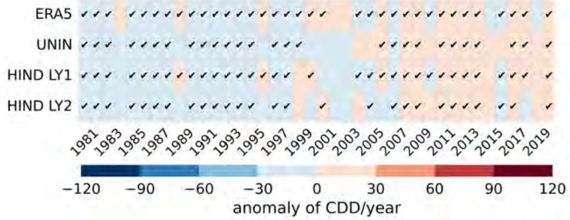


Korea-Japan (17/40 seasons -> 42.5%)

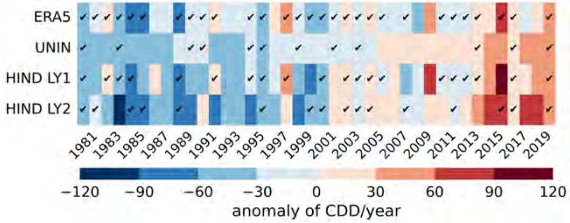


Heatmaps of the Anomaly of CDD_{dry} (MAM)

US Southwest (35/40 seasons -> 87.5%)



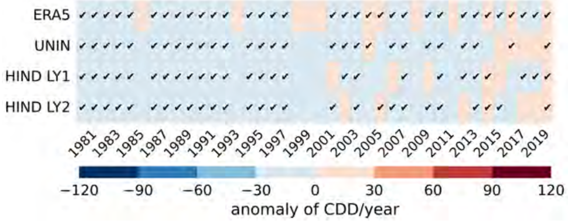
Central America (29/40 seasons -> 72.5%)



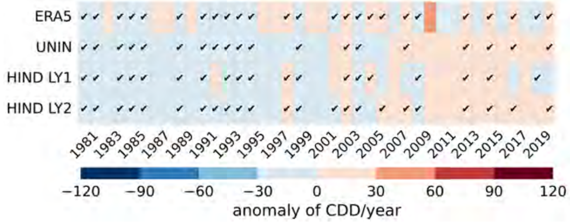
Greenland

No cooling demand since the average temperature (often below 10°C) is lower than the base threshold.

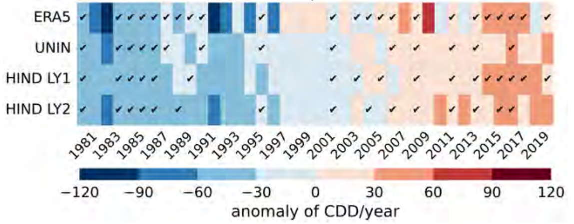
Spain (33/40 seasons -> 82.5%)



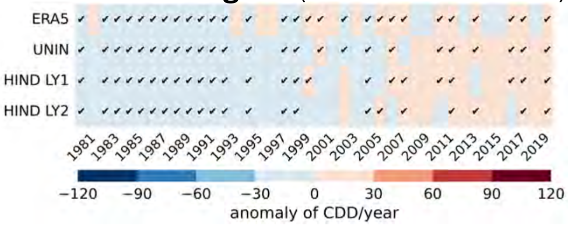
North Africa (25/40 seasons -> 62.5%)



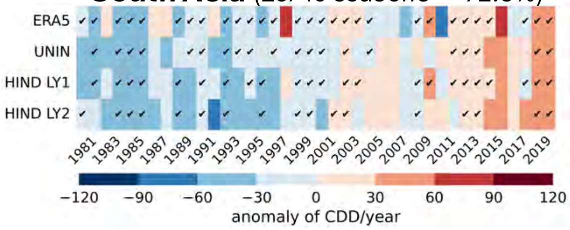
Arabian Peninsula (23/40 seasons -> 57.5%)



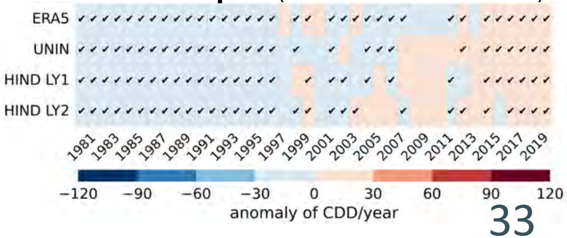
China-Mongolia (28/40 seasons -> 70%)



South Asia (29/40 seasons -> 72.5%)



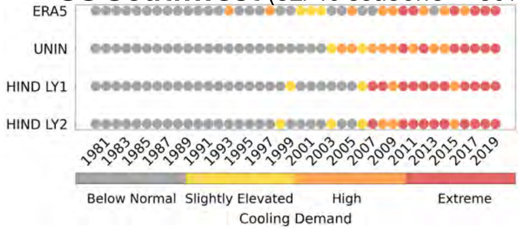
Korea-Japan (34/40 seasons -> 85%)



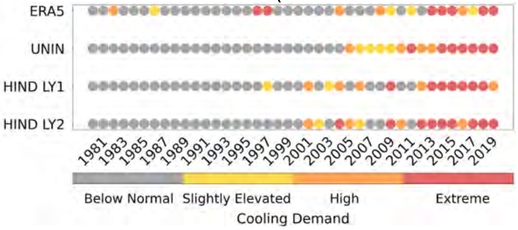
Can we predict **Cooling Demand** during JJA?

Determined only by THW frequency

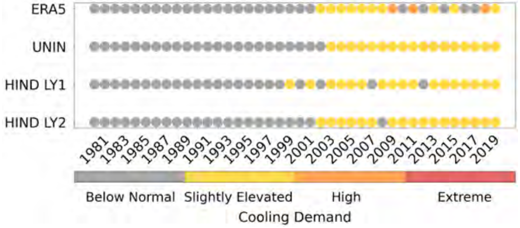
US Southwest (32/40 seasons -> 80%)



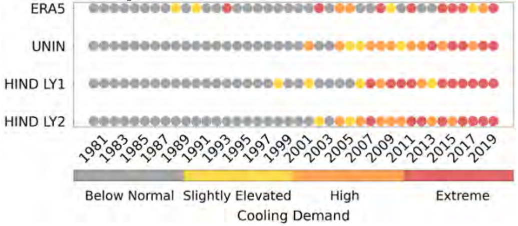
Central America (32/40 seasons -> 80%)



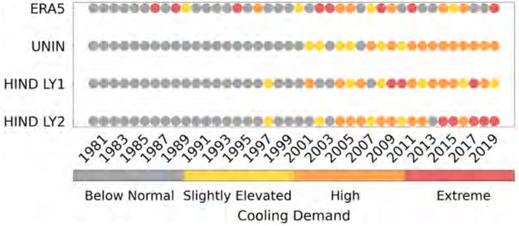
Greenland (33/40 seasons -> 82.5%)



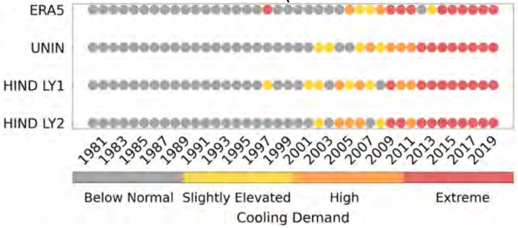
Spain (27/40 seasons -> 67.5%)



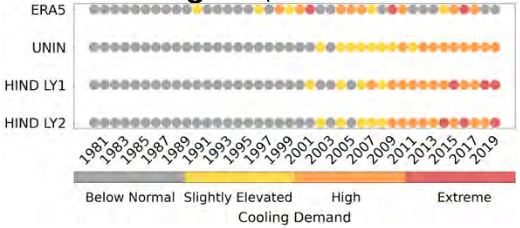
North Africa (25/40 seasons -> 62.5%)



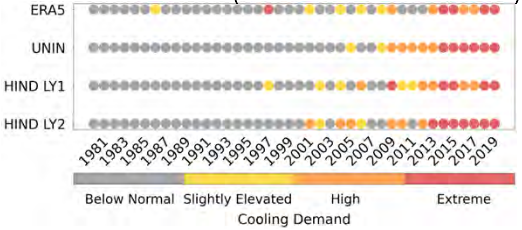
Arabian Peninsula (36/40 seasons -> 90%)



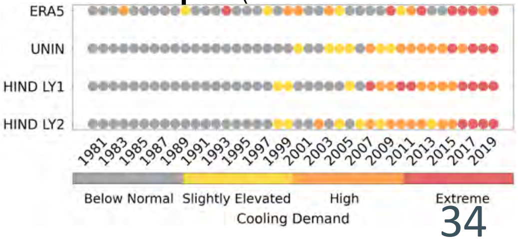
China-Mongolia (23/40 seasons -> 57.5%)



South Asia (35/40 seasons -> 87.5%)



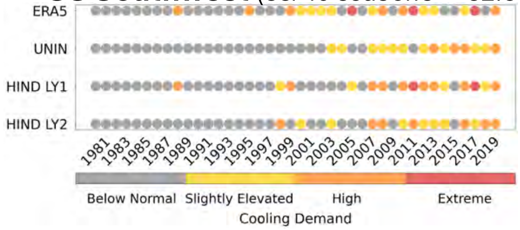
Korea-Japan (25/40 seasons -> 62.5%)



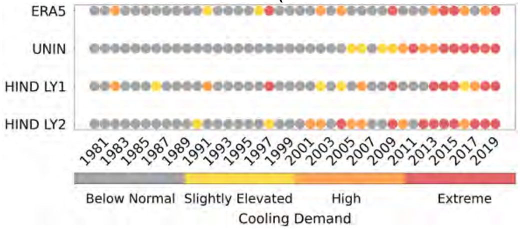
Can we predict **Cooling Demand** in MAM?

Determined only by THW frequency

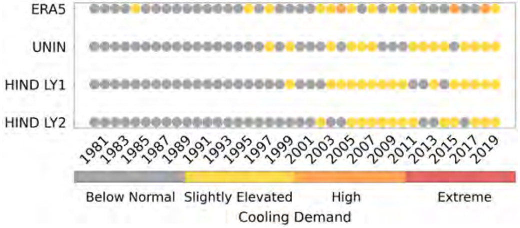
US Southwest (33/40 seasons -> 82.5%)



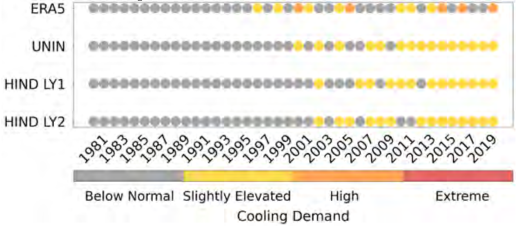
Central America (34/40 seasons -> 85%)



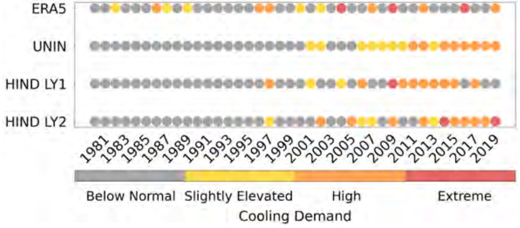
Greenland (33/40 seasons -> 82.5%)



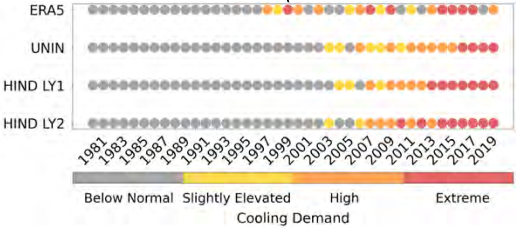
Spain (28/40 seasons -> 70%)



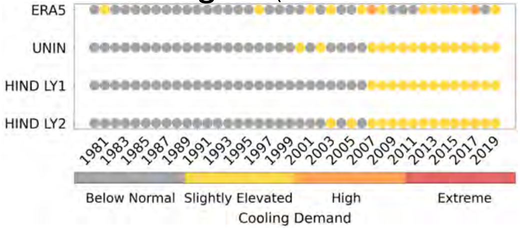
North Africa (26/40 seasons -> 65%)



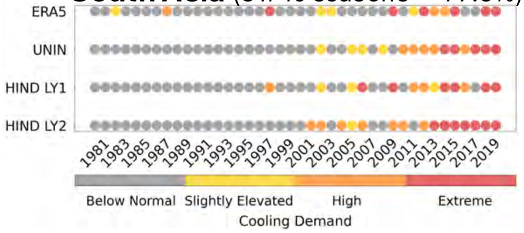
Arabian Peninsula (28/40 seasons -> 70%)



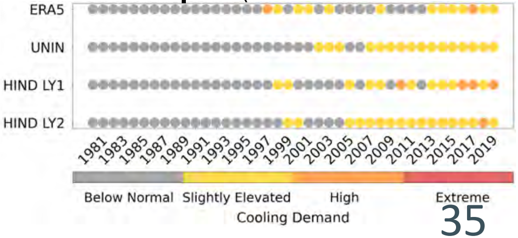
China-Mongolia (30/40 seasons -> 75%)



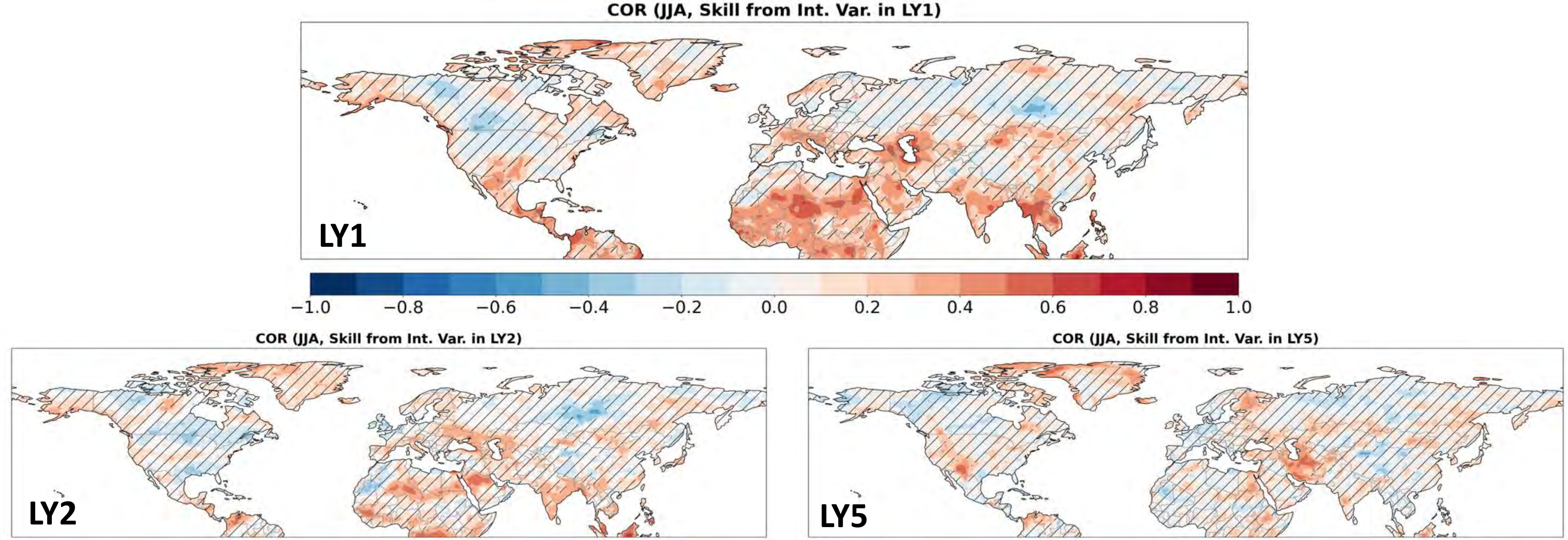
South Asia (31/40 seasons -> 77.5%)



Korea-Japan (35/40 seasons -> 87.5%)



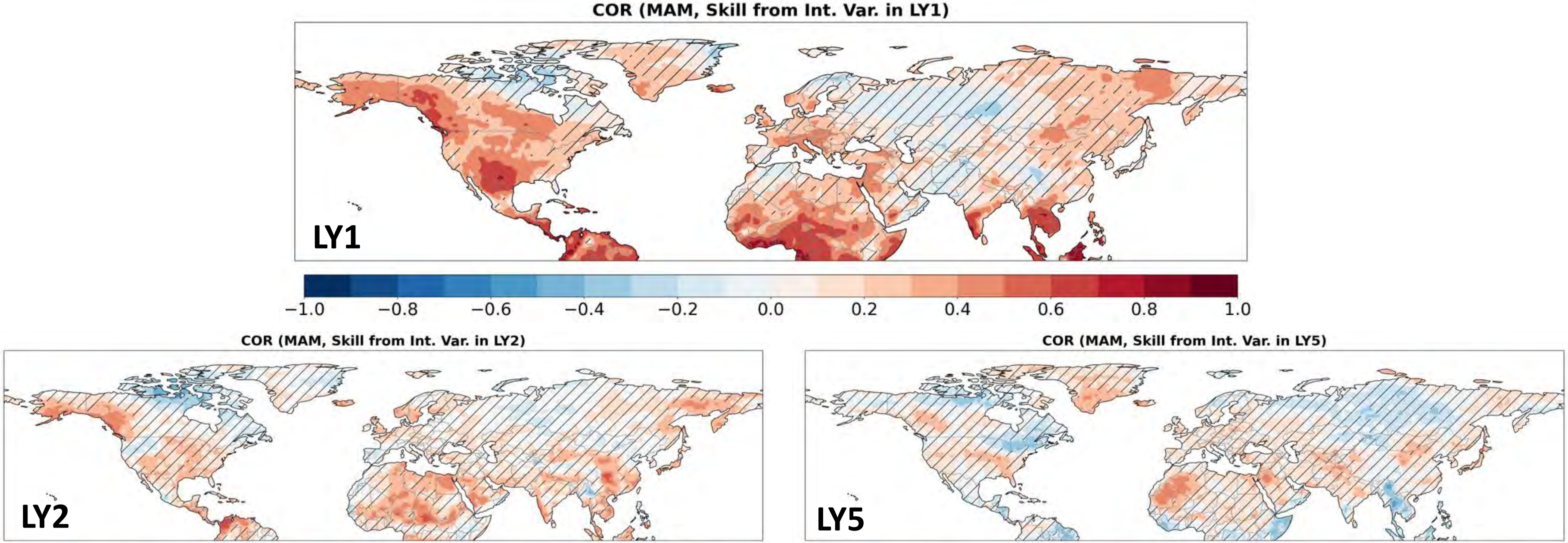
Skill from Internal Variability: Frequency of Thermal Heatwaves (JJA)



Anomaly Correlation Coefficient / JJA 1981-2020 / HIND-LE and ERA5-LE

5-, 17-, 53-Month Lead

Skill from Internal Variability: Frequency of Thermal Heatwaves (MAM)



Anomaly Correlation Coefficient / MAM 1981-2020 / HIND-LE and ERA5-LE

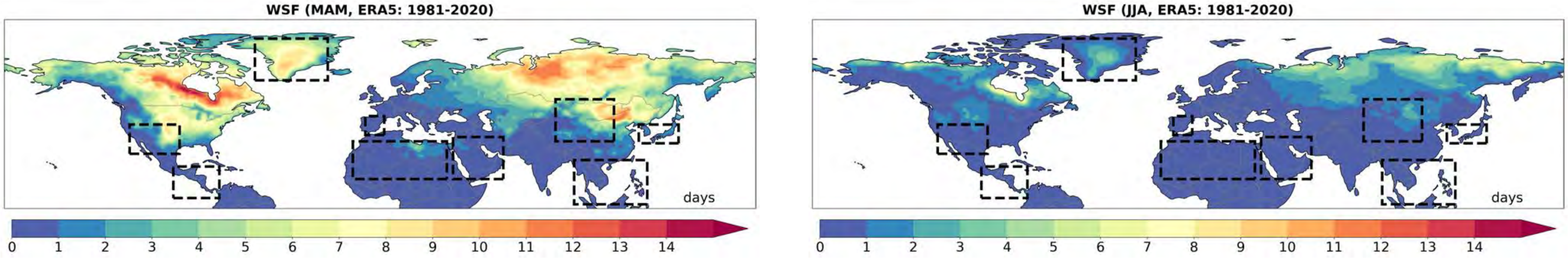
2-, 14-, 50-Month Lead

Sudden Day2Day Temperature Spikes:
Warm Shock Frequency (ERA5, 1981-2020)

- previous statistics (THW freq., CDDs) focus on sustained heat and cumulative thermal load, however, “rapid temperature flips“ *(Wu et al. 2025)* may also pose risks to power grid reliability

- short-term temperature volatility, if predictable, may signal the start of a heatwave

- we choose a threshold of Δt2m ≥ 5°C as most suitable for use in early warning systems



→ very few days of “rapid temperature flips” in the NH → high predictability in CESM2-MP?