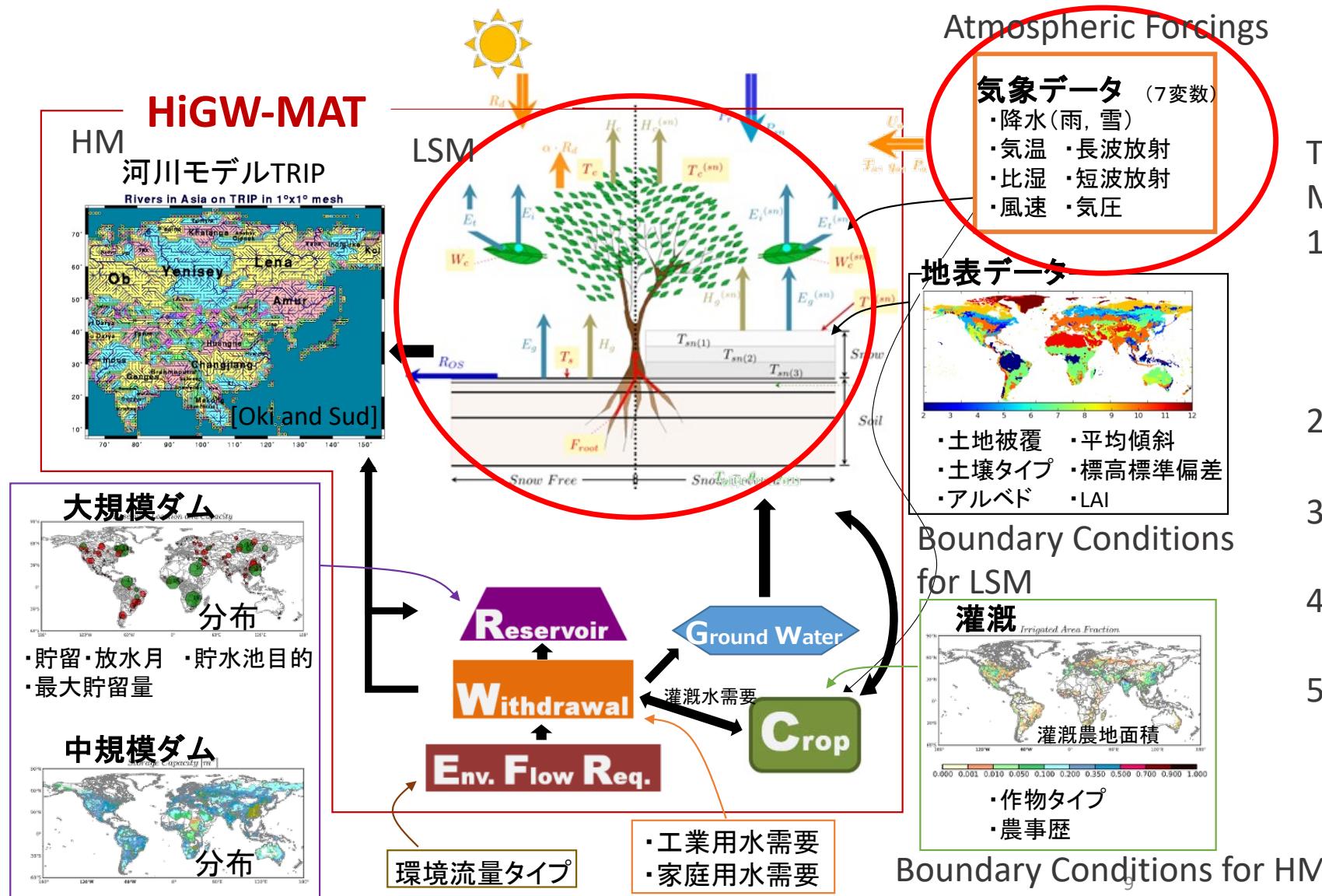


Machine Learnings for Terrestrial Simulations

Kei Yoshimura (The University of Tokyo)

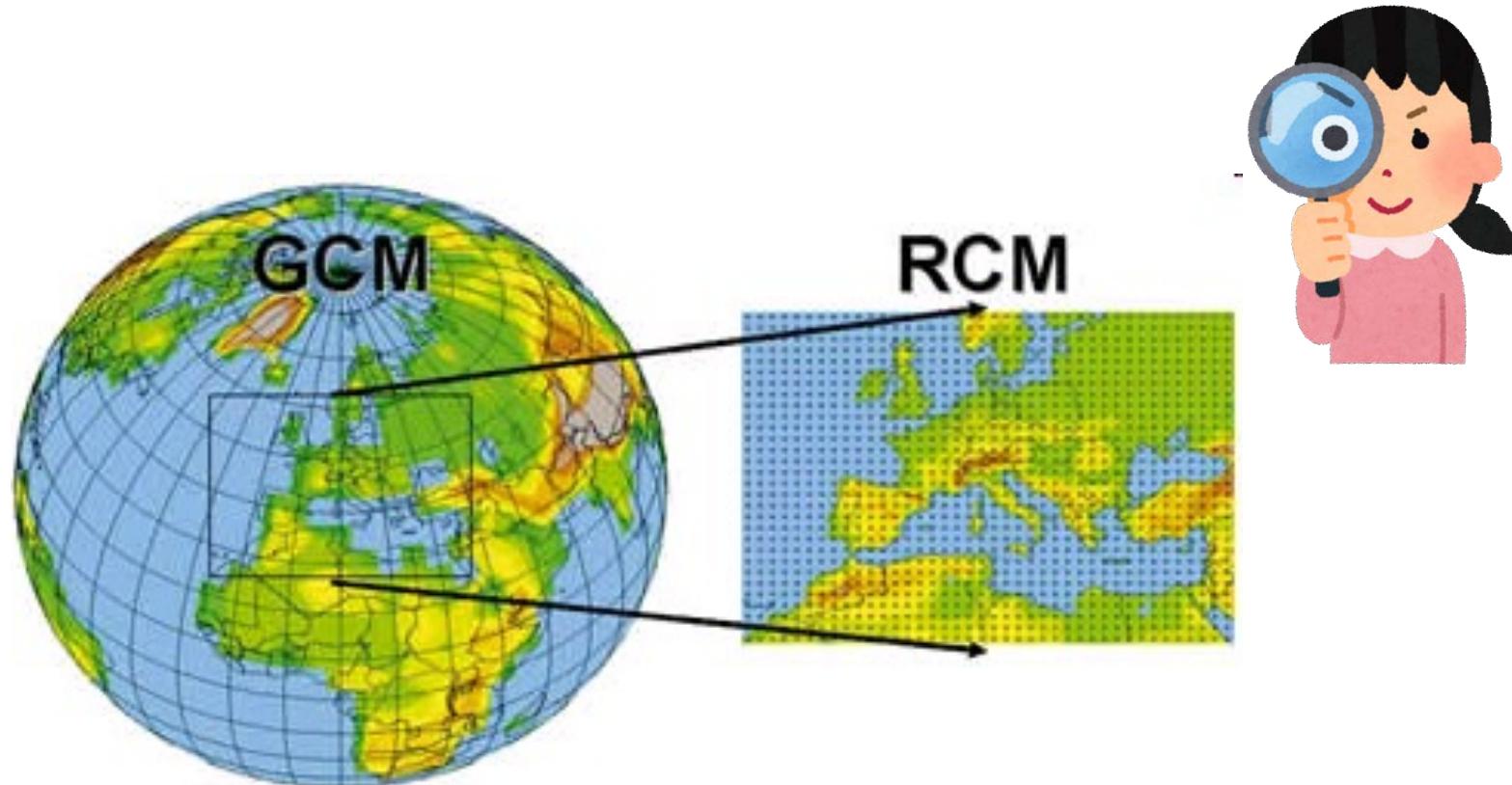
Many thanks to: Tomoko Nitta, Takao Yoshikane, Gaohong Yin, Roman Olson, Wenpeng Xie, Hongmei Li, Tamima Amin,

Typical Terrestrial Simulation: Offline Experiments with LSM/HM



- Today's Topics:
- Machine learnings for
1. Downscaling and bias correction of atmospheric forcing data (for both historical and future)
 2. Improvement of physical understanding
 3. Parameter calibration of LSM/HM
 4. Speed-up of calculation using Emulator
 5. Terrestrial Water Storage estimation

(Dynamical) Downscaling



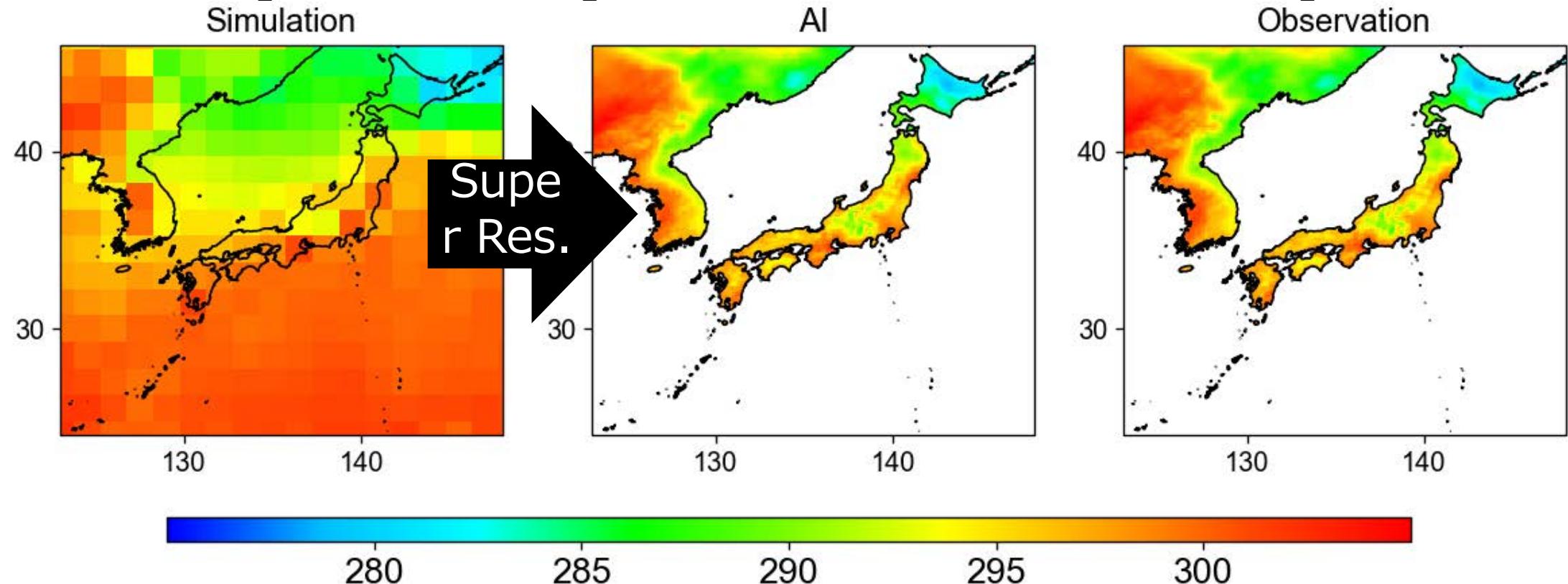
http://www.wmo.int/wcc3/bulletin/57_2_en/images/bull_57_2_9.jpg

- Popular method to make a high-resolution surface atmospheric conditions from low-resolution observation or model simulation.
- Necessary for impact assessment, risk analysis, etc.
- Computational (very) heavy.

Downscaling with AI (Super Resolution)

Machine Learning Super-Resolution Technology: Learning the Difference Between Low-Resolution and High-Resolution Images

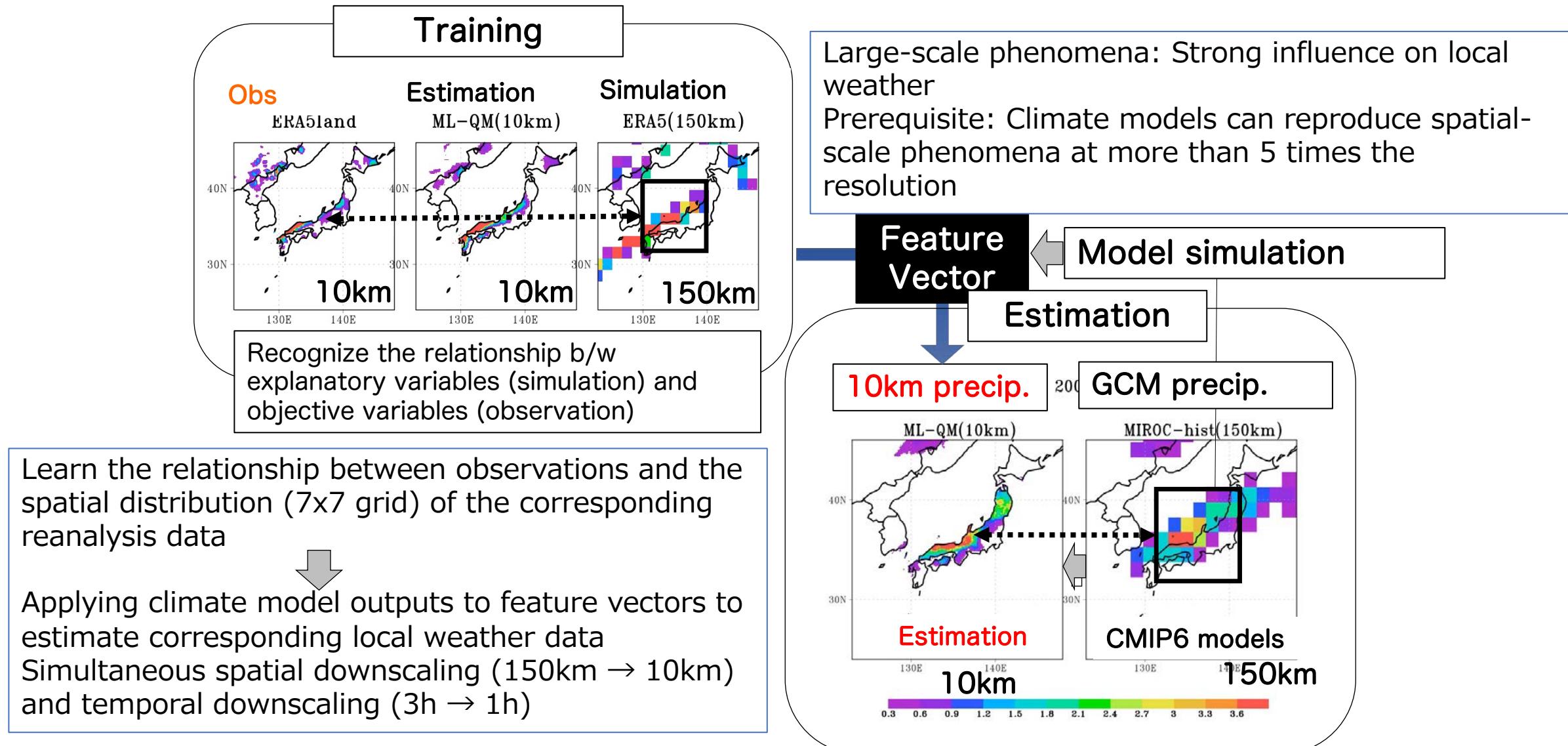
→ Generate high-resolution images from unlearned low-resolution images



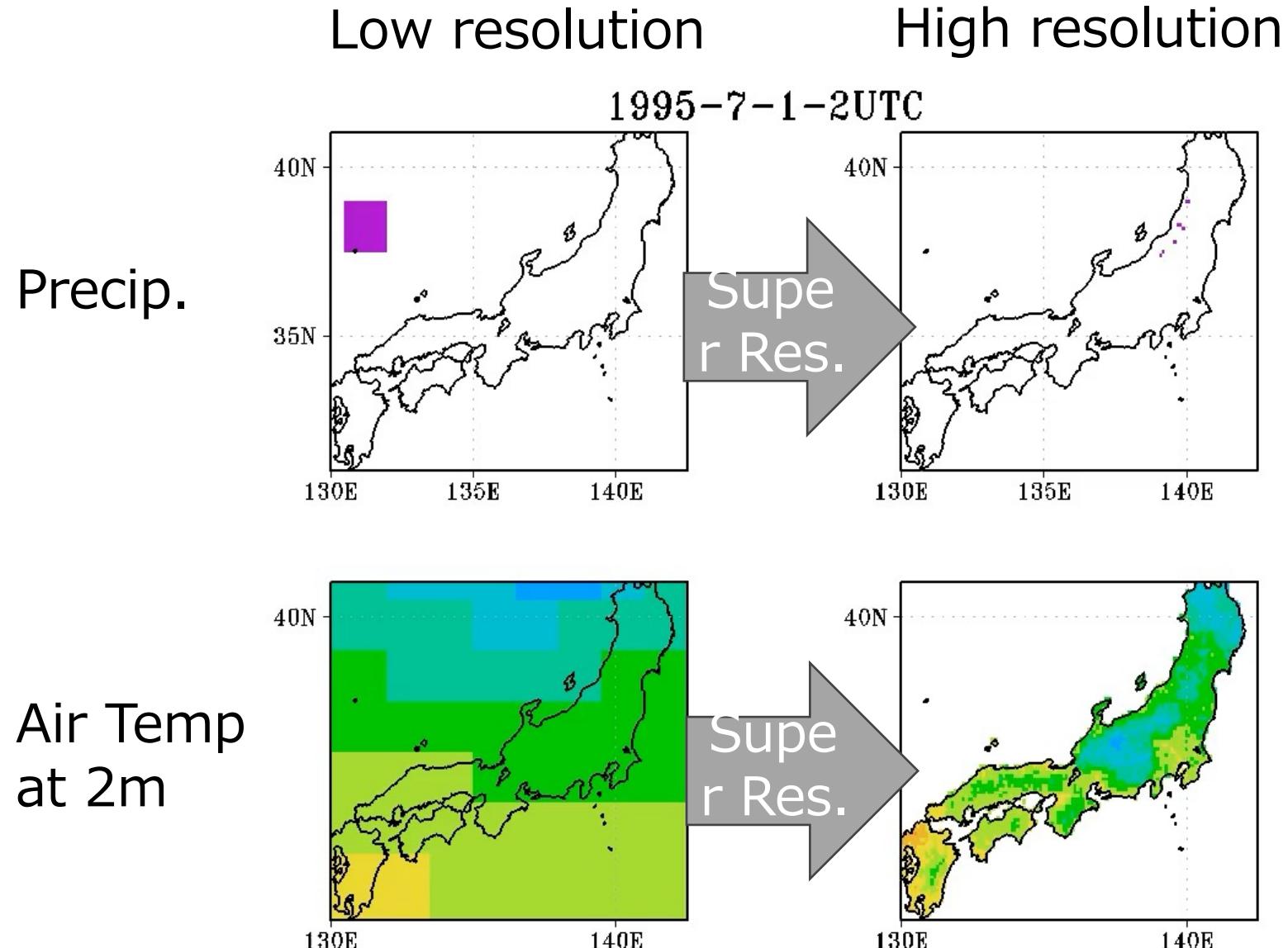
Our method: Using the characteristic that local weather is strongly influenced by broad-area weather.
→ Pattern Recognition of Relationships between “Time-varying Weather Characteristics at Observation Sites” and “Movement of the wide-area weather system reproduced by the climate model”.

Downscaling (super-resolution) using machine learning

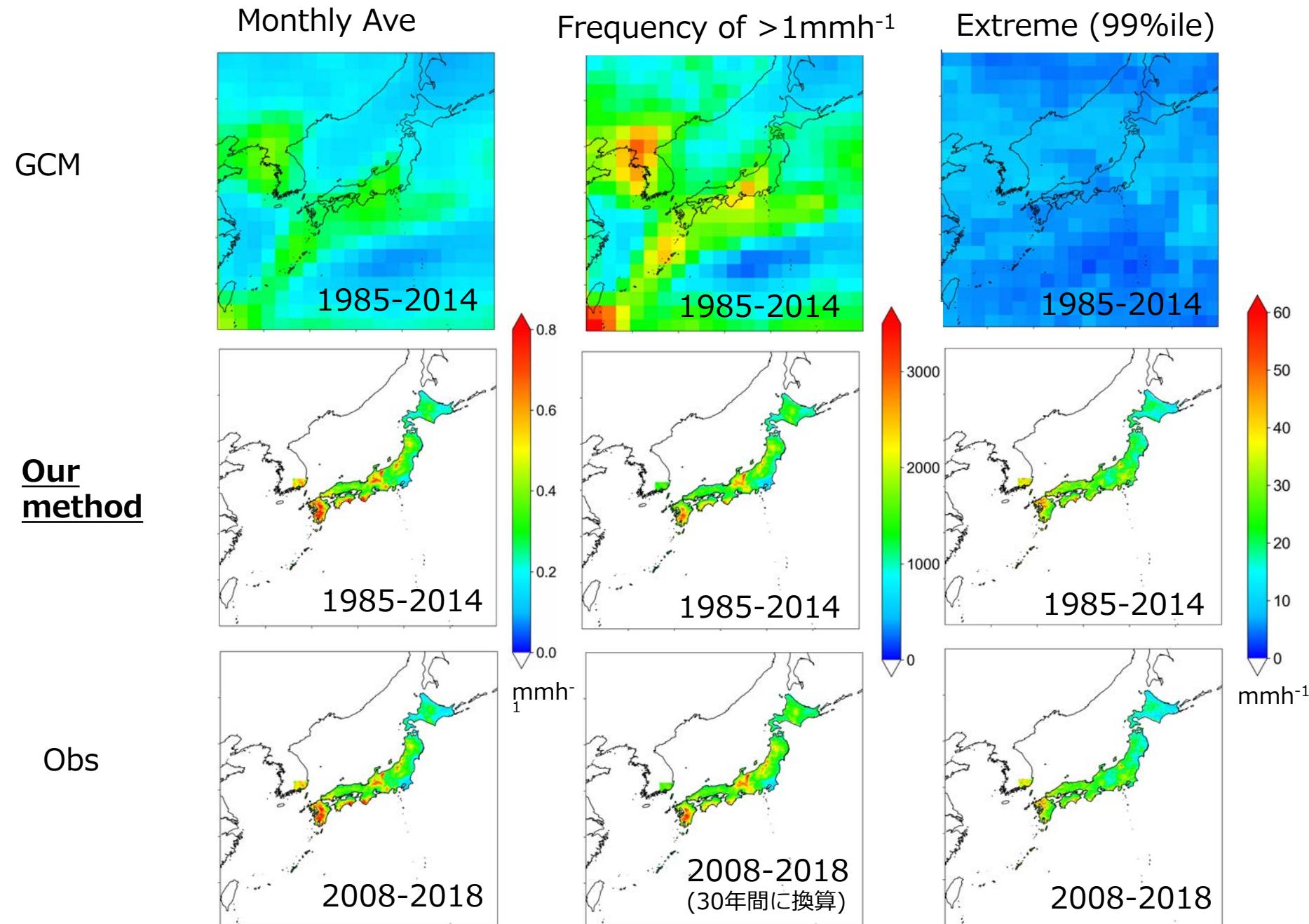
(from 150km to 10km)



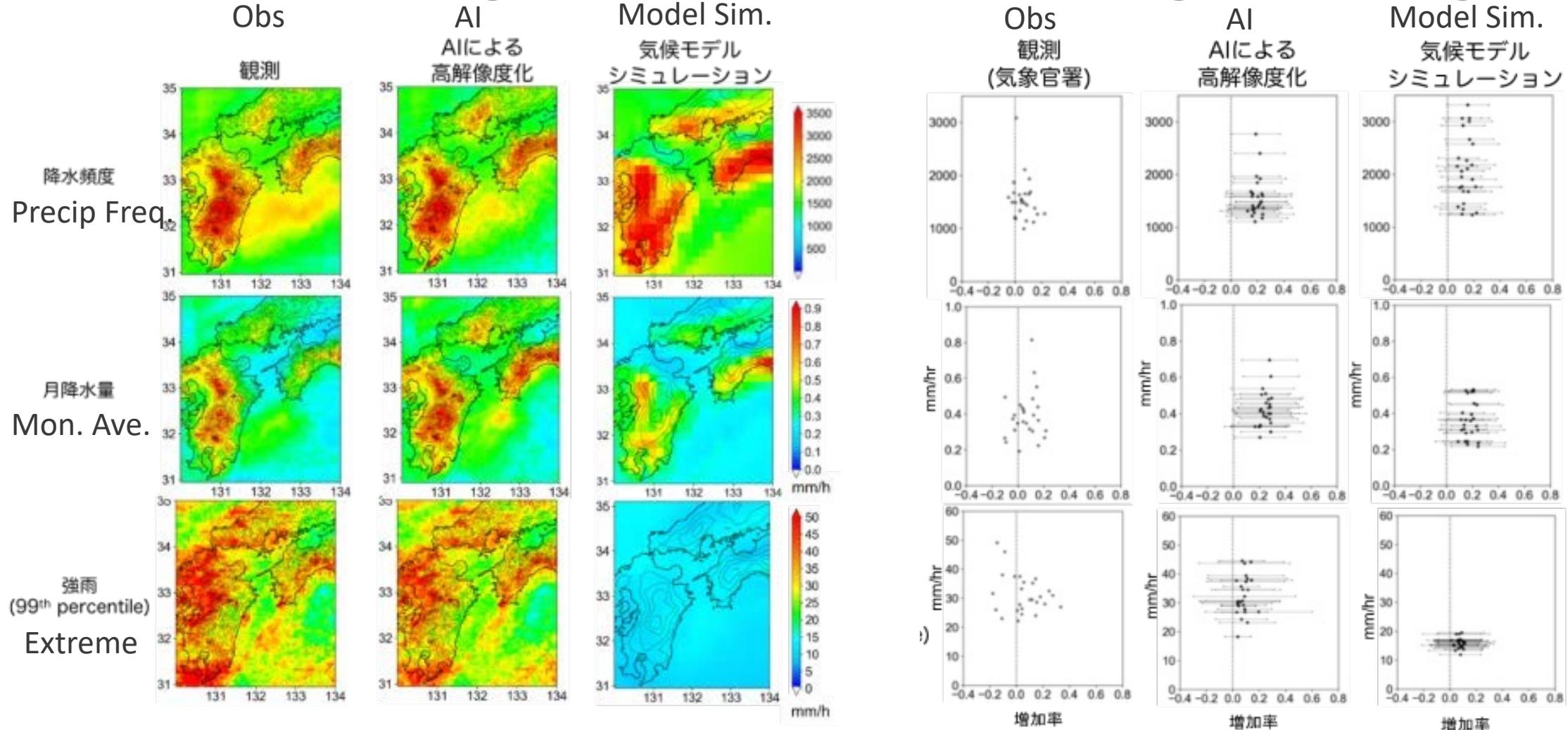
Results



Climatological Characteristics (July)

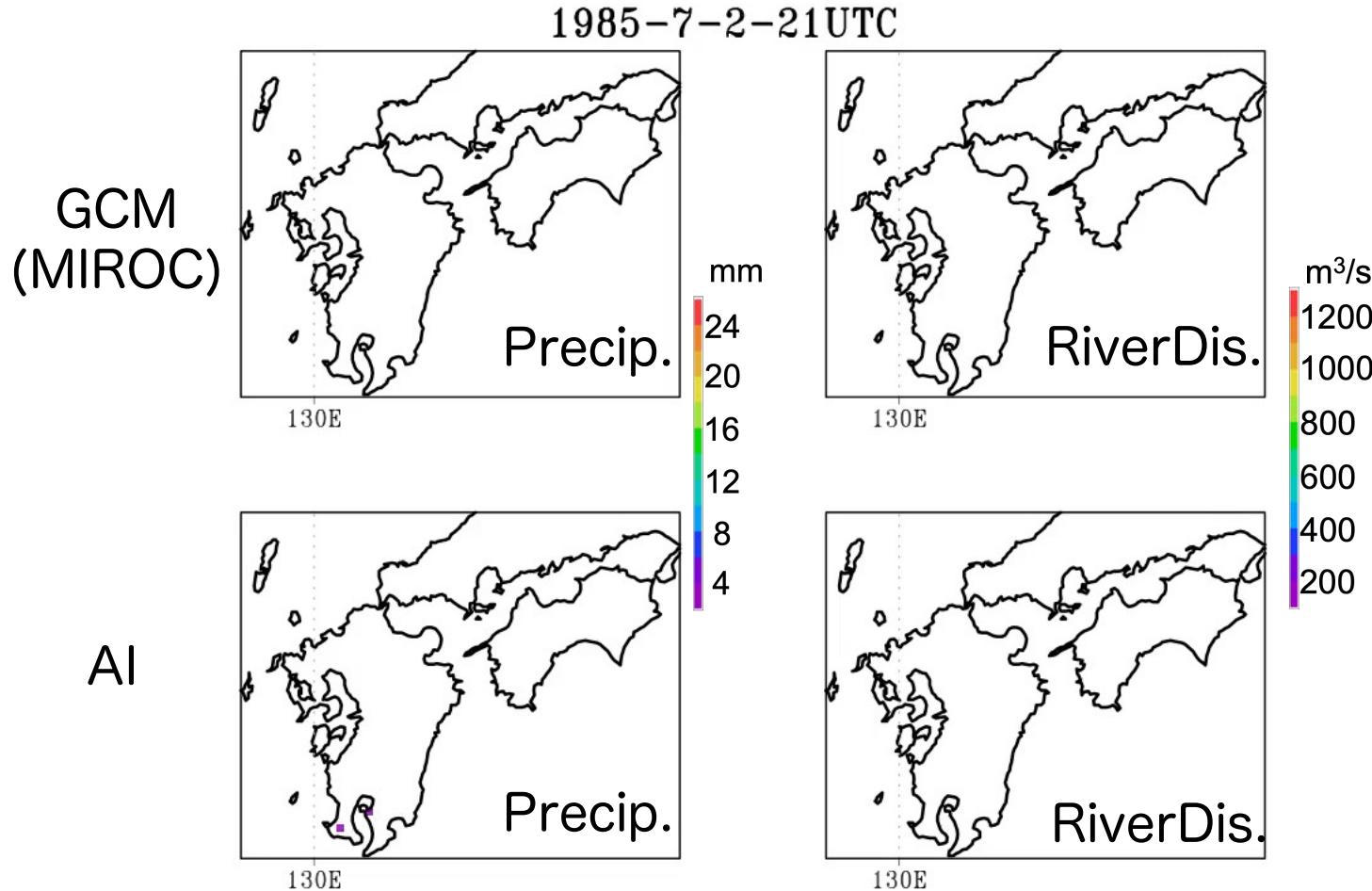


Downscaling of Simulated Climatological Changes



Recent 30yr (1982-2011) - Past 30yr (1952-81)

Snapshot simulation and hydrological impact

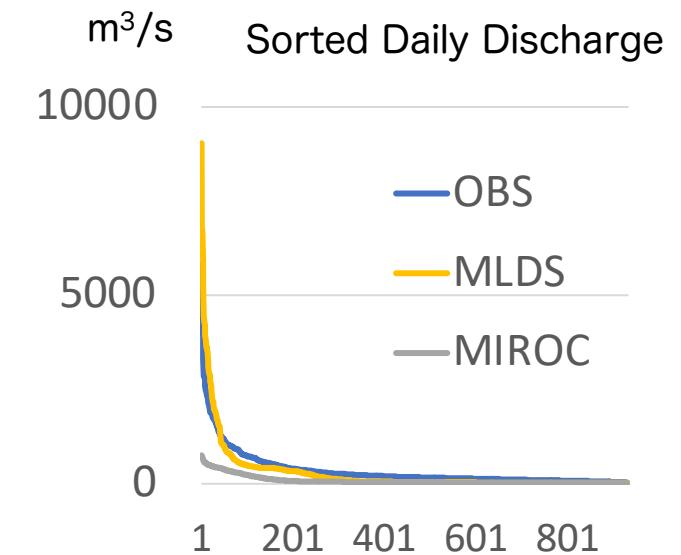
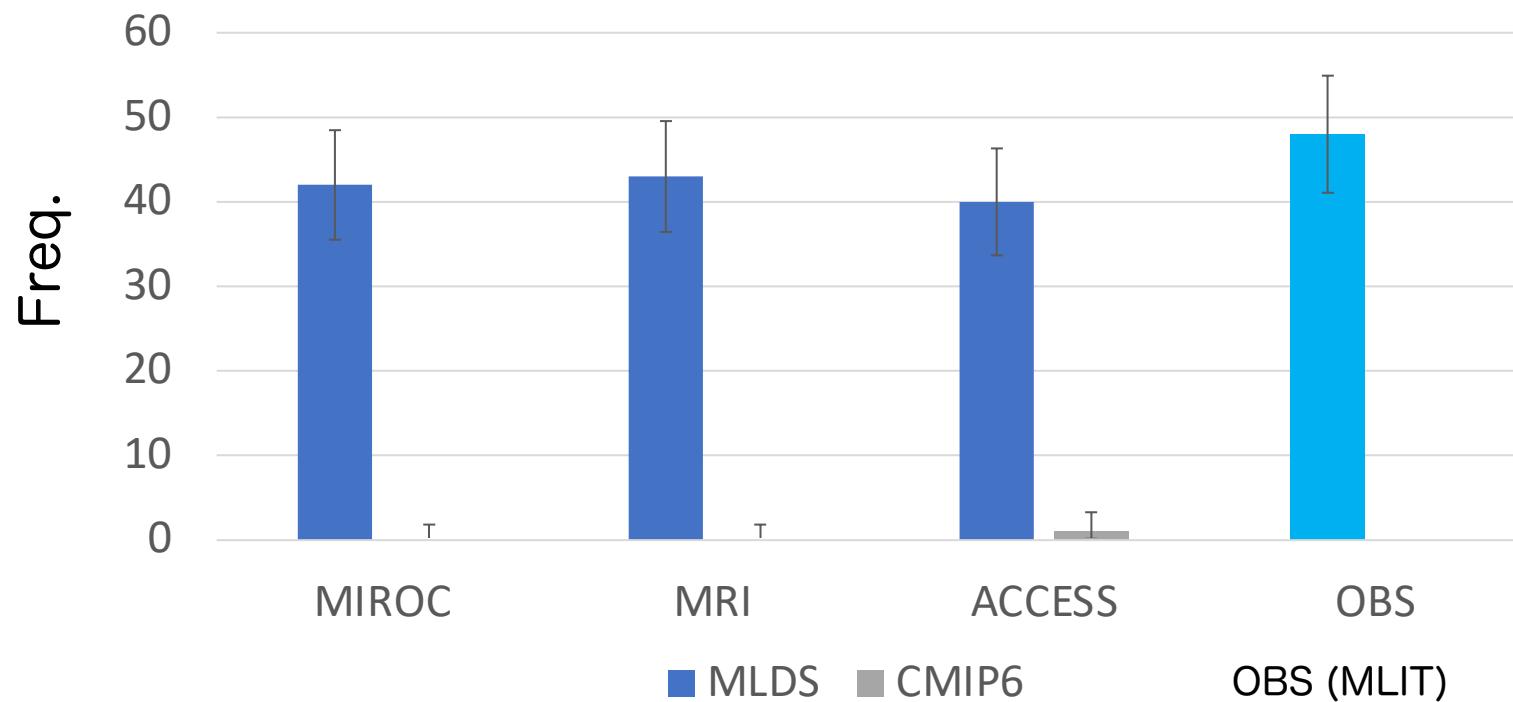


Hydrological Impact (Retension Curves)



Kuma river basin
(1,880km²)

Frequency of daily discharge in Kuma River >1200m³/s (95%ile of obs.)

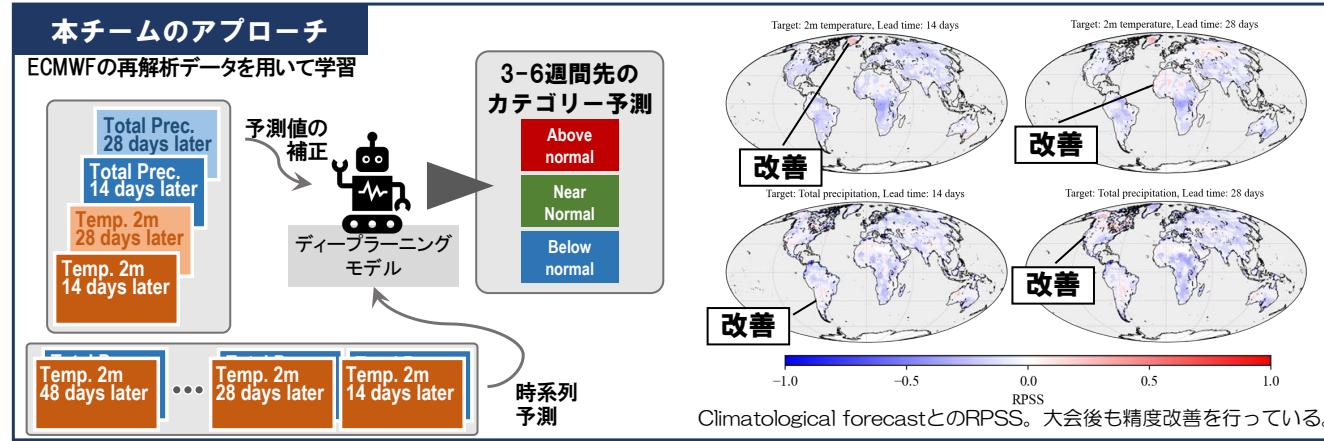


What is the potential application to intra-seasonal-seasonal forecasting (S2S)? S2S AI Challenge by WMO

S2S予測への取り組み

概要

- WMOがECMWFなどと共同で開催したChallenge to improve Sub-seasonal to Seasonal Predictions using Artificial Intelligenceに参加→**8位入賞**！
- 総降水量・2m気温のS2S予測を、機械学習(ML)・深層学習(DL)で行いその精度を競った。

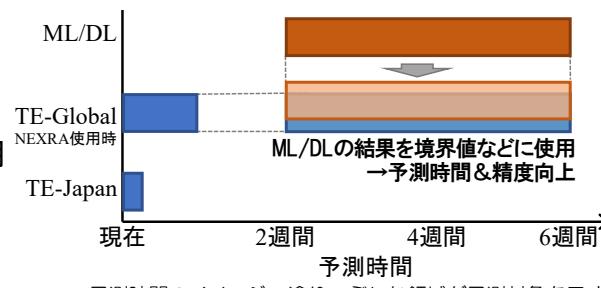


S2SデータのTE-Globalでの利用可能性検討

本取り組みは、TE-Globalの予測時間向上に用いられ、**洪水予測などに貢献できる可能性**がある。

- JMA/MRI-CPS2の季節予測など多様なデータを利用
- 前処理・後処理へのML/DLの適用
- 計算時間の短縮

※ML/DLの入出力をTE-Globalに合わせ柔軟に変更し貢献可能



CERTIFICATE OF APPRECIATION

THIS CERTIFICATE IS PRESENTED TO

Ryo Kaneko, Gaohong Yin, Wenchao Ma,
Kinya Toride, Gen Hayakawa, and Kei Yoshimura

IN RECOGNITION OF CONTRIBUTION TO CHALLENGE TO IMPROVE SUB-SEASONAL TO SEASONAL PREDICTIONS USING ARTIFICIAL INTELLIGENCE OF THE WORLD WEATHER RESEARCH PROGRAMME AND THE WORLD CLIMATE RESEARCH PROGRAMME OF THE WORLD METEOROLOGICAL ORGANIZATION SUB-SEASONAL TO SEASONAL PROJECT

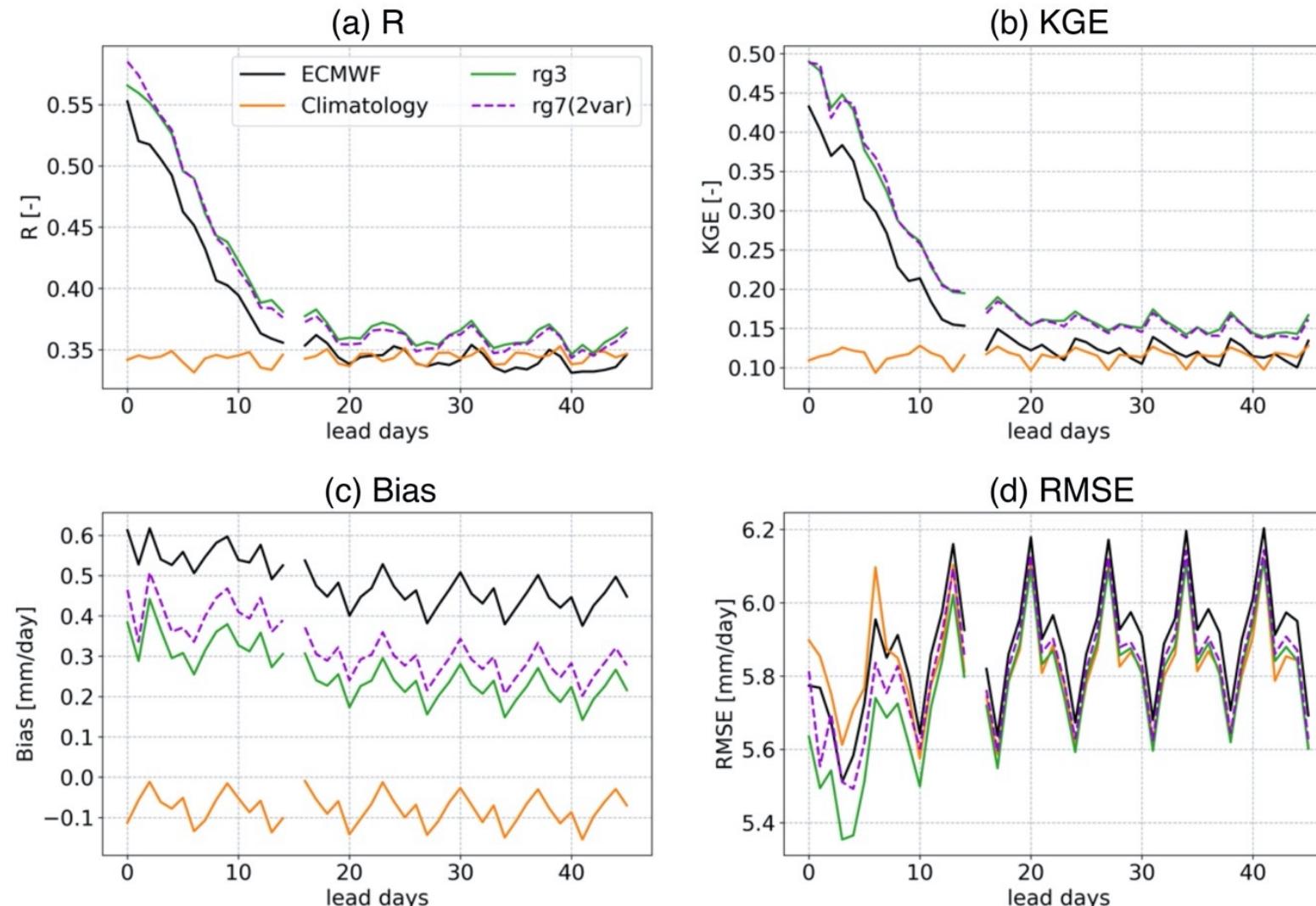
F. Vitart
A.W. Robertson

Frederic Vitart and
Andrew W. Robertson
S2S Project Leaders

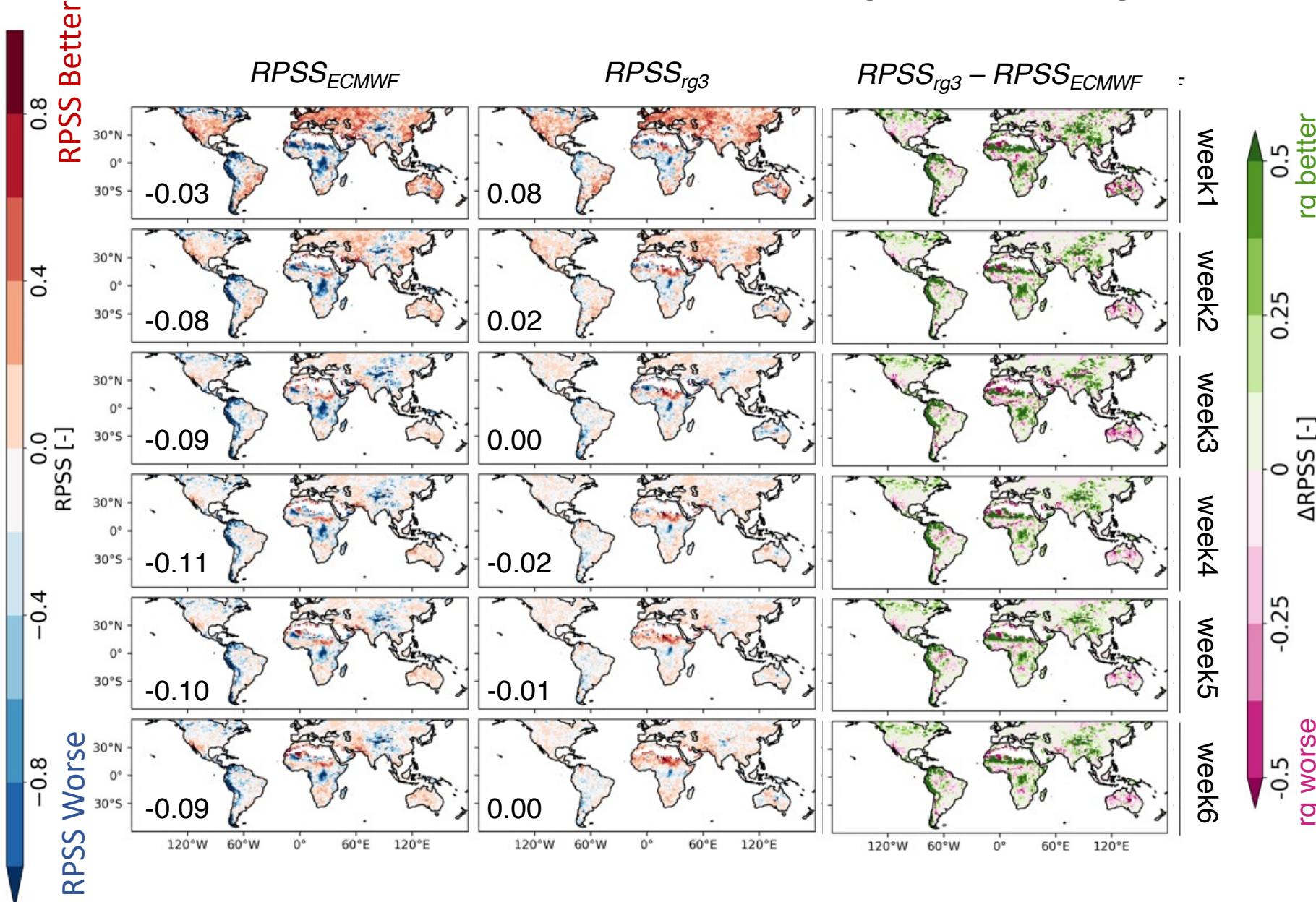
Estelle De Coning and
Michael Sparrow
Heads of the WWRP and
WCRP Secretariats



Application to intra-seasonal-seasonal forecasting (S2S)



BC/DS of S2S Precipitation by ECMWF

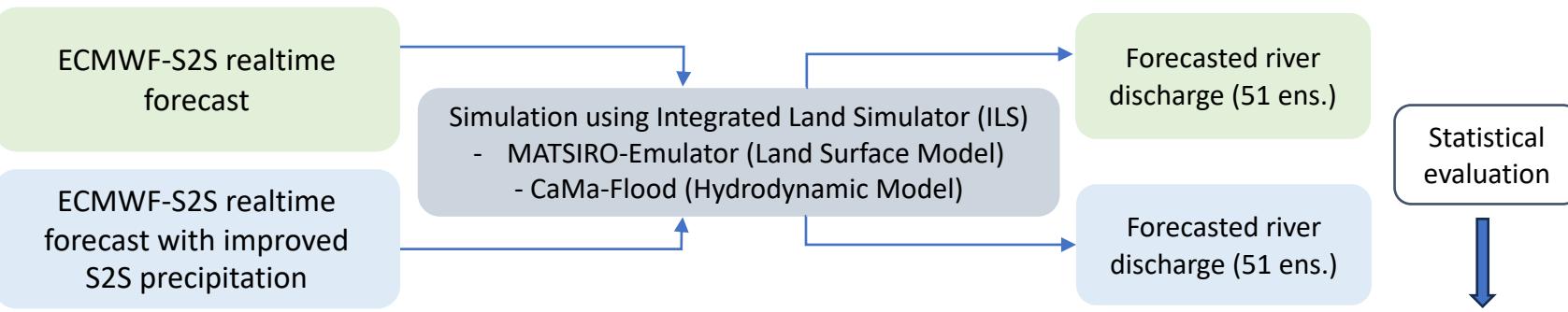


- Clearly, ranking performance improved.
- The improvement was most pronounced in complex terrain where precipitation forecasts were less accurate.
- Post-processed precipitation forecasts showed potential for use in seamless hydrologic forecasting.

The predictability of global river discharge forecast at sub-seasonal to seasonal (s2s) timescale

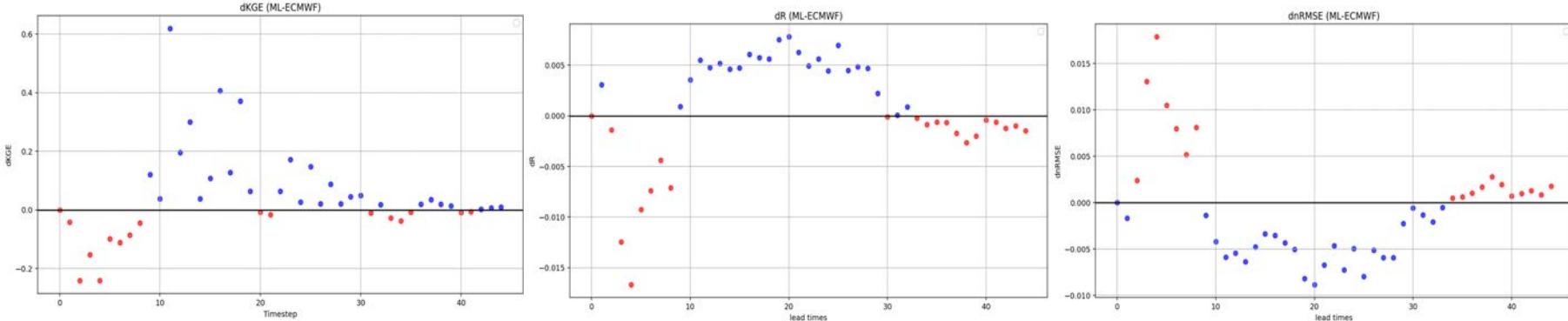
□ Study framework

53 forecasts initiated in 2020, each has 51 ens.



By incorporating ML-improved precipitation forecast, is it applicable to improve the skill of ILS-simulated river discharge forecast?

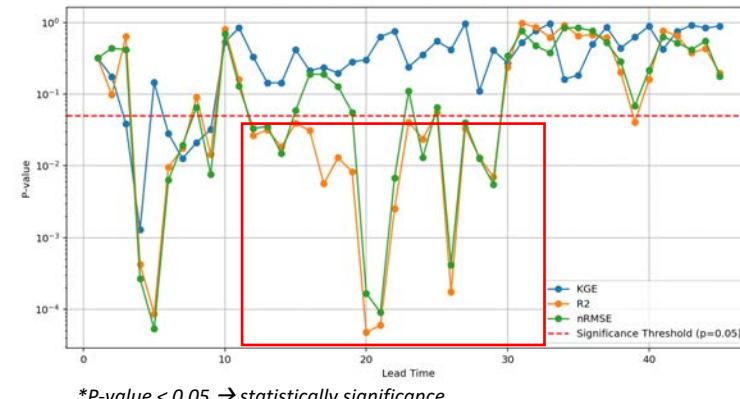
□ Results - Statistics of predictability based on lead times



- According to KGE and R² globally averaged value, skill enhancement becomes noticeable between 10 and 45 days of lead time.
- Statistical significance analysis of R² and nRMSE demonstrates that ML-forcings can enhance the skill of forecasted river discharge with p-value also indicated significant value (< 0.05) in 10-30 days lead times.

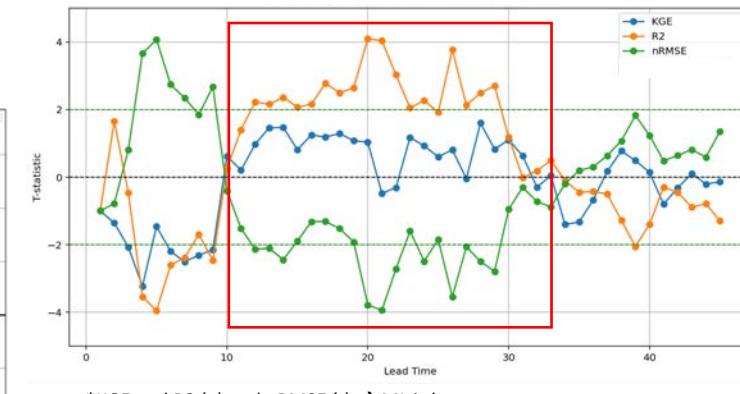
□ Statistical significance (587 sites)

P-value



*P-value < 0.05 → statistically significant

T-statistics

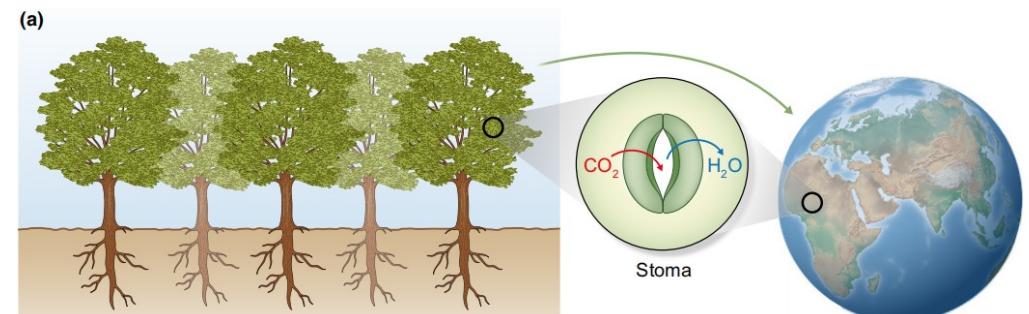


□ Key message

- Despite it is not significant improvement, the ML-forced precipitation are able to enhance the skill of forecasted river discharge in later lead times.

A Data-Driven Perspective on the Role of Leaf-to-Air Temperature Difference in Stomatal Regulation

- **The limited homeothermy hypothesis:** specific suites of leaf traits have evolved through natural selection to buffer environmental temperature fluctuations, thereby maintaining leaf temperatures within a narrower range around metabolic optima
- Plant thermoregulation, represented by the leaf-to-air temperature difference (ΔT), has been documented across various species.
- Stomatal opening has traditionally been attributed to the **diffusion of water vapor** along a concentration gradient.
- The role of thermodiffusion in stomatal regulation remains largely unexplored due to **the difficulty of isolating the independent effects of ΔT . Here, we used explainable machine learning to address this challenge.**



- Stomata control the diffusive exchange of CO₂ ([photosynthesis](#)) and water vapor ([transpiration](#)) between the ecosystem and atmosphere

Experiments design

- Traditional model (Medlyn et al., 2017)

$$gs = g_0 + 1.6 \times (1 + g_1 / \sqrt{VPD}) \times A_n / CO_2$$

- Machine learning models

Models	Experiment overview	Experiment name	Input variables
RF, SVM, CDNN	Feature importance ranking	PFI	VPD, An, CO ₂ , SWP, PAR, Tleaf, ΔT
RF, SVM, CDNN	Reference	Experiment1	VPD, An, CO ₂
RF, SVM, CDNN	Baseline	Experiment2	VPD, An, CO ₂ , SWP, PAR
RF, SVM, CDNN	Controlled experiments	Experiment3	VPD, An, CO ₂ , SWP, PAR, Tleaf
RF, SVM, CDNN		Experiment4	VPD, An, CO ₂ , SWP, PAR, ΔT
RF, SVM, CDNN		Experiment5	VPD, An, CO ₂ , SWP, PAR, Tleaf, ΔT
RF, SVM, CDNN		Experiment6	VPD, An, CO ₂ , SWP, PAR, Tleaf, Tair
CDNN	Mechanism analysis	ALE	VPD, An, CO ₂ , SWP, PAR, ΔT

RF: Random Forest

SVM: Support Vector Machine

CDNN: Convolutional Deep Neural Network

VPD: leaf-to-air vapour deficit

An: photosynthesis

CO₂: CO₂ concentration.

Tair: air temperature

- data information used in this work:
plant species: *Pinus Sylvestris*
Location: *Finland* (61.51°N, 24.17°E)
Time: June-August in 2006-2008
Scale: leaf

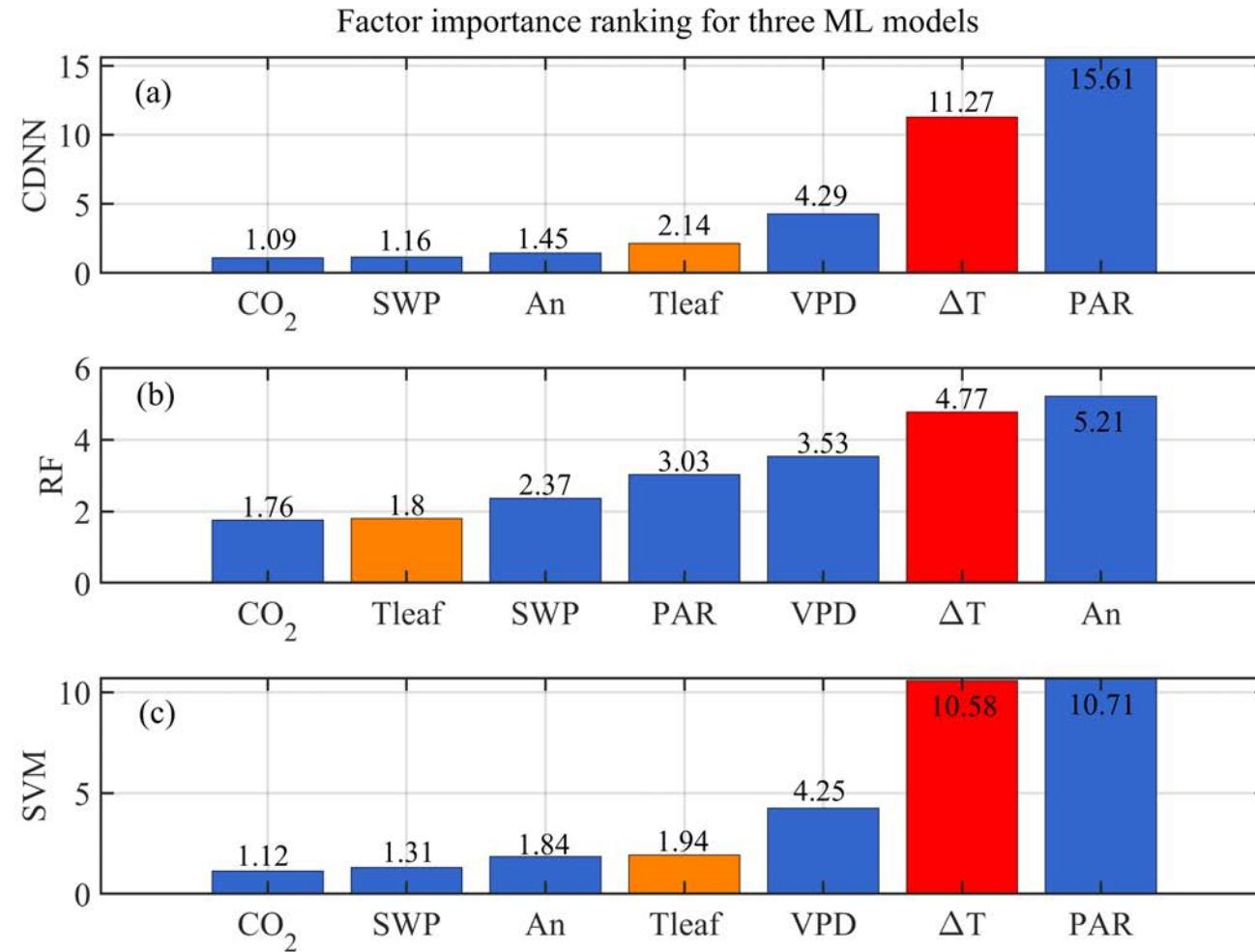
SWP: soil water potential

PAR: photosynthetic active radiation

Tleaf: leaf temperature

ΔT: leaf-to-air temperature difference

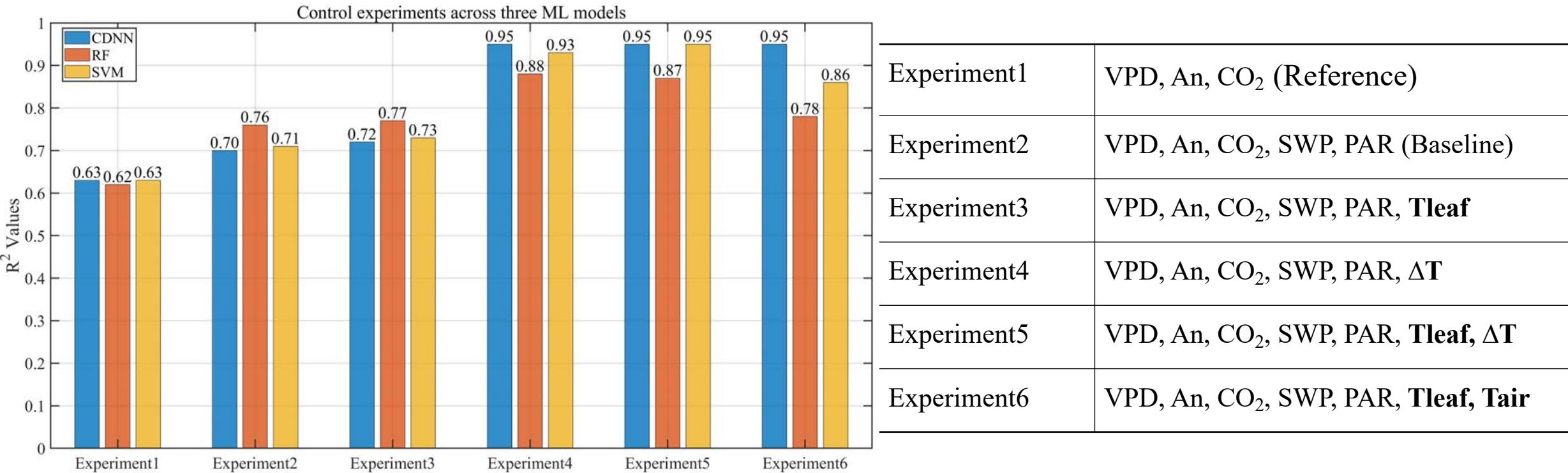
Feature Importance Ranking



The bar represents the mean importance of each feature based on RMSE loss over 50 permutations of the *Pinus Sylvestris* training datasets. The unit is %.

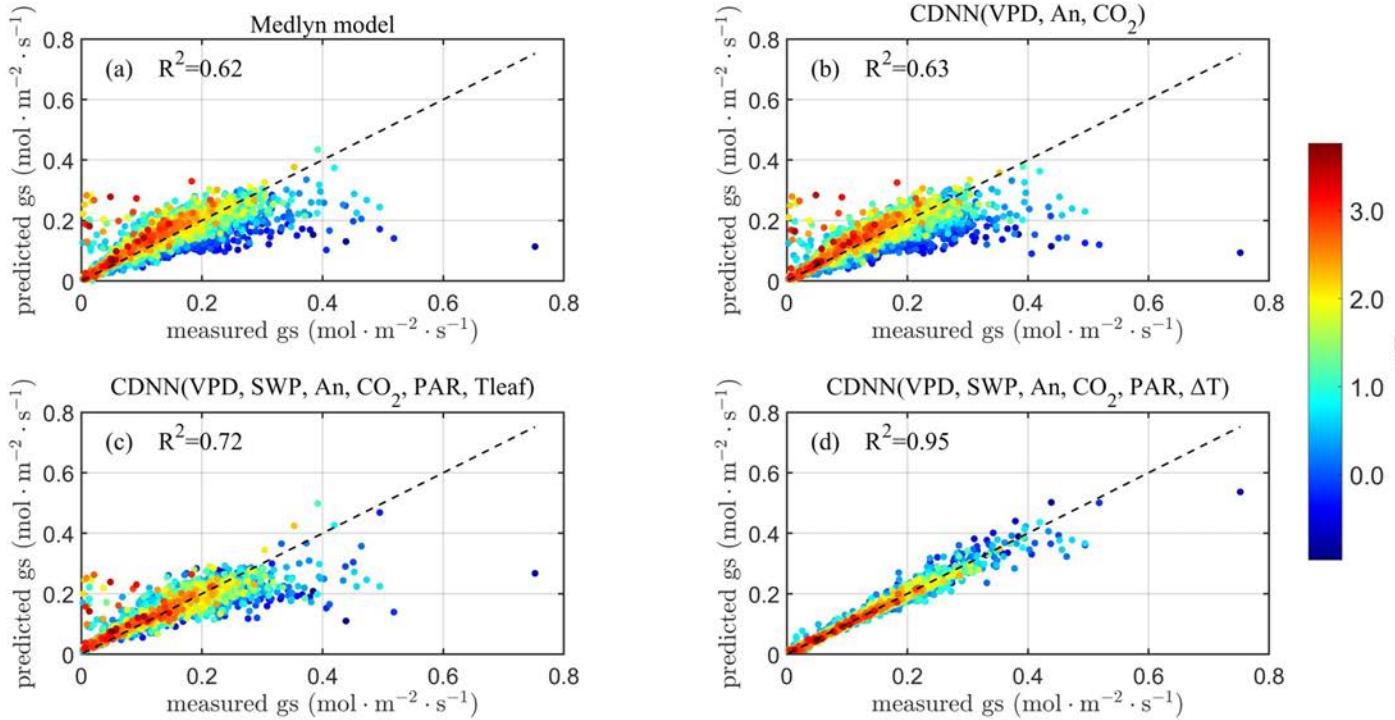
- ΔT ranks second in feature importance across all three machine learning models.
- Stomatal conductance is more sensitive to ΔT than to Tleaf.

Control Experiments Results



- The incorporating ΔT significantly improves all three ML model performance. The R^2 values of CDNN, RF, and SVM in Experiment 4 improved by 35.7%, 15.8%, and 31% compared to baseline experiment. The inclusion of Tleaf did not enhance model accuracy (Experiment3 compared to baseline).
- These results suggest that, from a data-driven perspective, Tleaf does not directly influence stomatal conductance. Instead, the direct effect is driven by ΔT .

Control Experiments Results

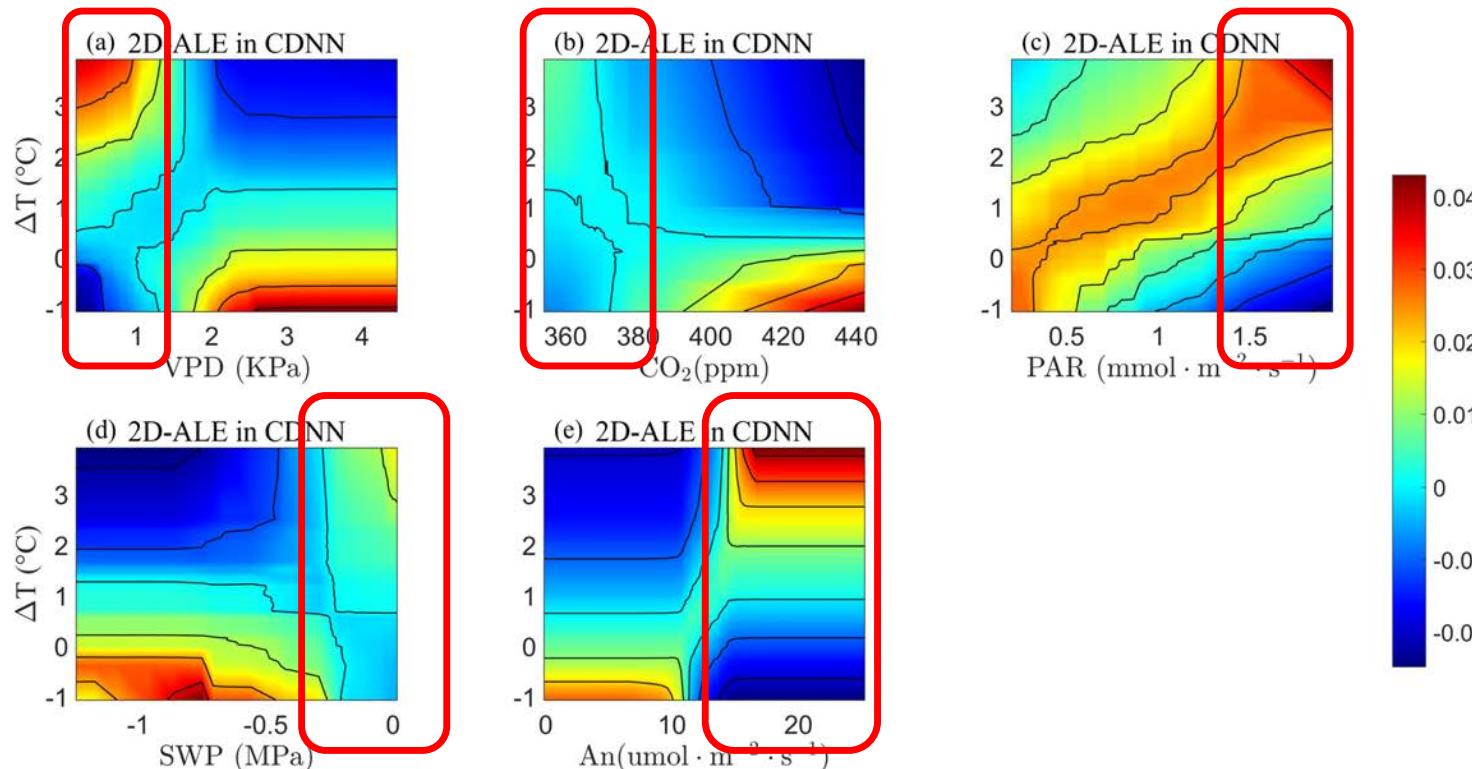
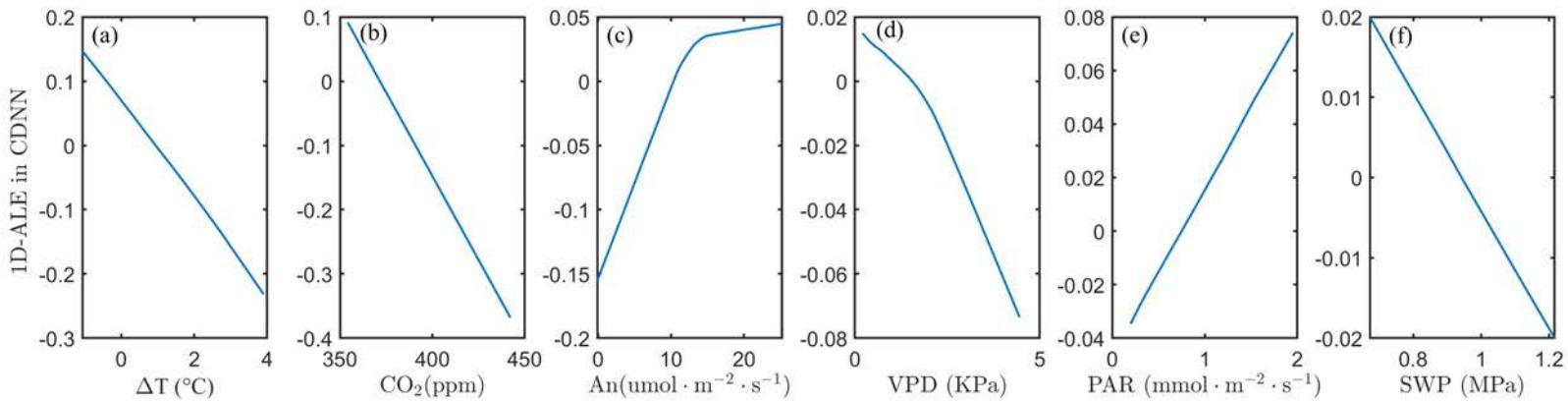


- When temperature information is missing, both the Medlyn model and CDNN perform poorly when T_{leaf} is lower than T_{air} .
- Incorporating ΔT , rather than using T_{leaf} , effectively addresses this issue.

RMSE in different ΔT ranges	$\Delta T < -1$	$-1 < \Delta T < -0.5$	$-0.5 < \Delta T < 0$	$0 < \Delta T < 0.5$	$0.5 < \Delta T < 1$	$1 < \Delta T < 1.5$	$\Delta T > 1.5$
Medlyn(VPD, An, CO ₂)	0.13	0.07	0.037	0.04	0.032	0.032	0.039
CDNN(VPD, An, CO ₂)	0.14	0.07	0.039	0.037	0.029	0.028	0.036
CDNN(VPD, SWP, An, CO ₂ , PAR, Tleaf)	0.1	0.051	0.034	0.036	0.027	0.026	0.03
CDNN(VPD, SWP, An, CO ₂ , PAR, ΔT)	0.035	0.018	0.014	0.014	0.009	0.008	0.006

Explainable ML results

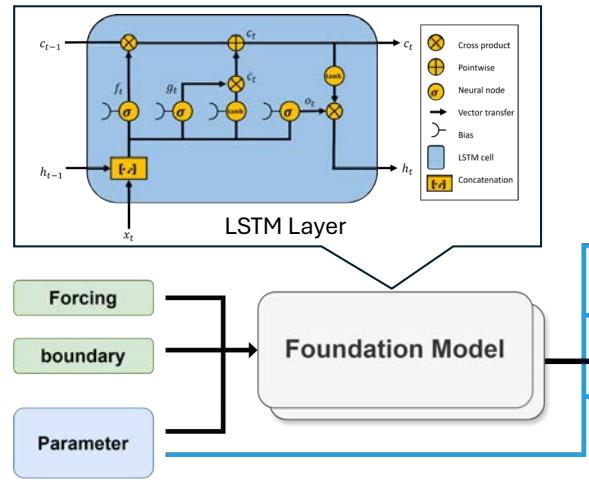
1D-ALE results: ΔT primarily produce a negative impacts on stomatal opening.



2D-ALE results:

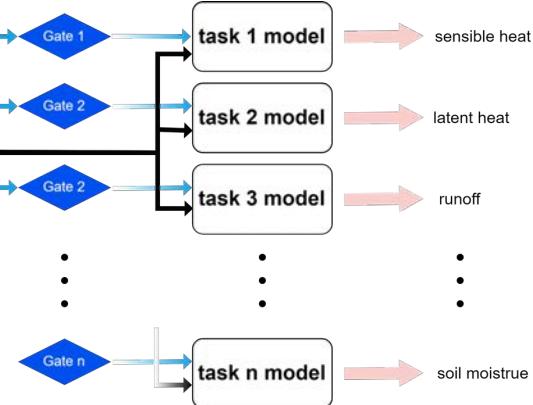
- stomatal conductance generally decreased with rising ΔT under unfavorable environmental conditions.
- However, a potential positive response was observed under conditions of sufficient water supply, low CO_2 concentration, low transpiration demand, and high photosynthetic demand.

Calibration of Model Parameters: Deep Learning Approach

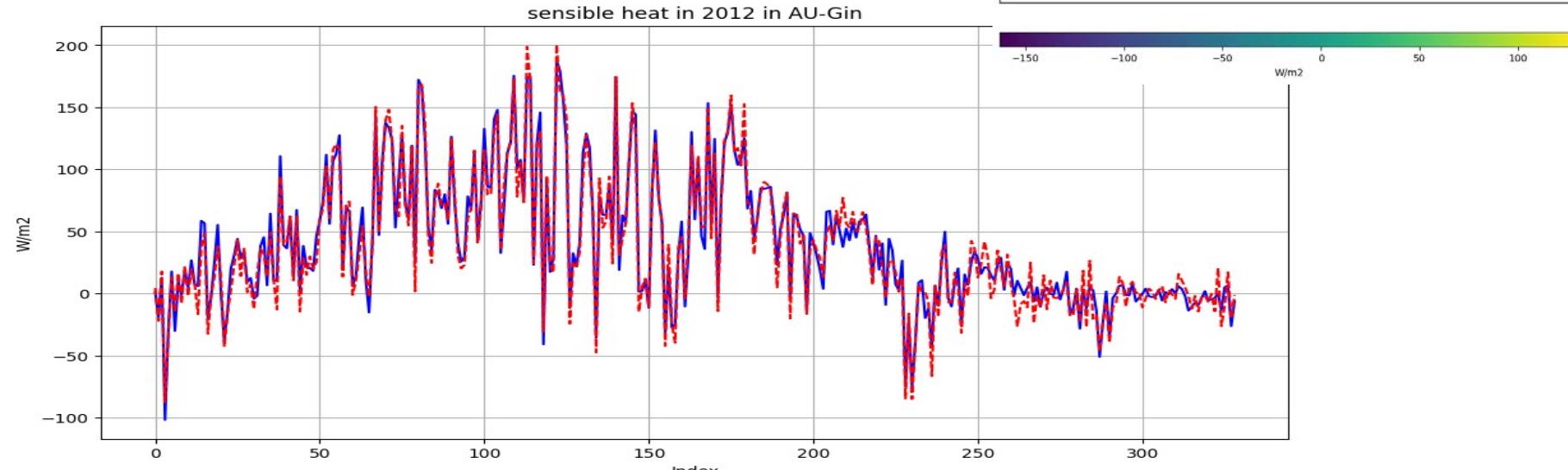
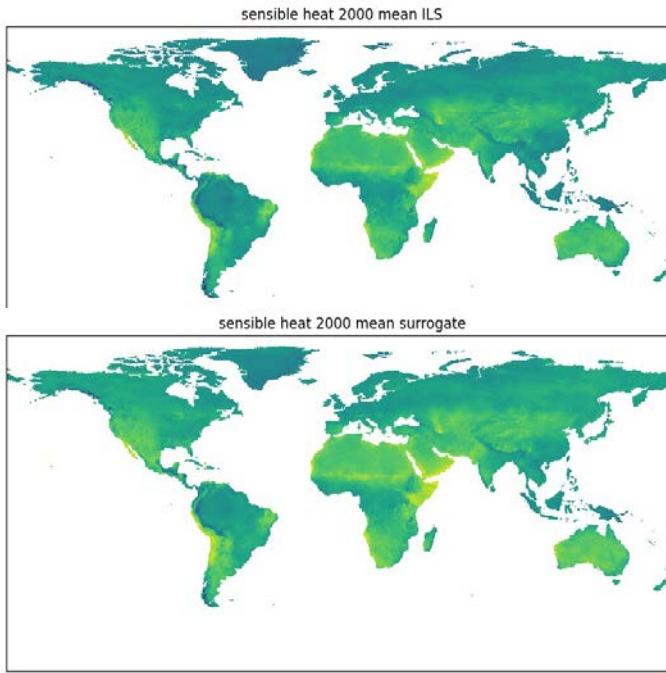


In the field of deep learning, "Foundation Models" refer to large, pre-trained neural network models that are trained on extensive datasets and can be fine-tuned for a variety of specific tasks.

Inspired by dPL and multi-task learning,
we created
Foundation Surrogate Model
for parameter calibration.



Rotation Training

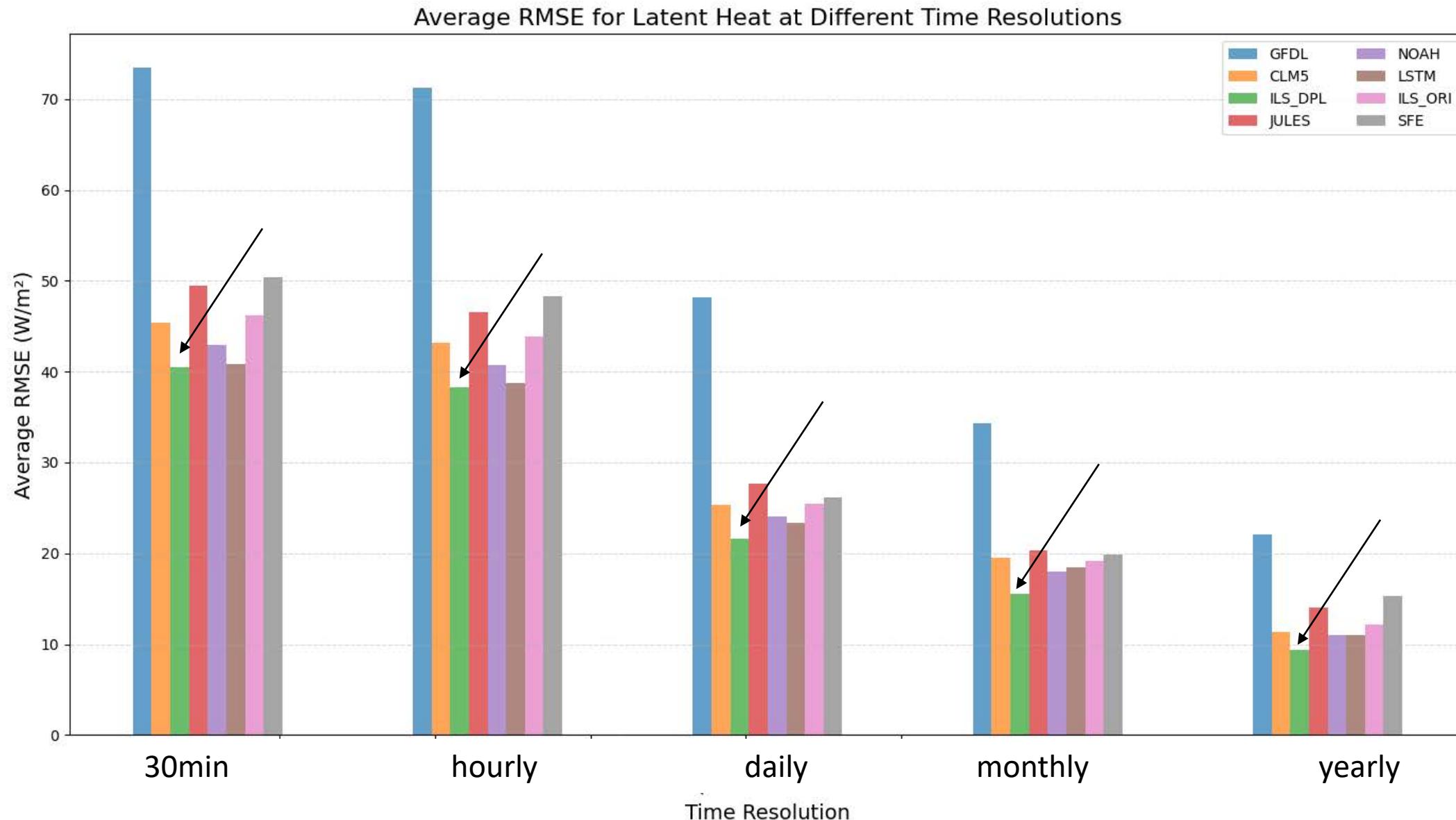


PLUMBER2 sites information

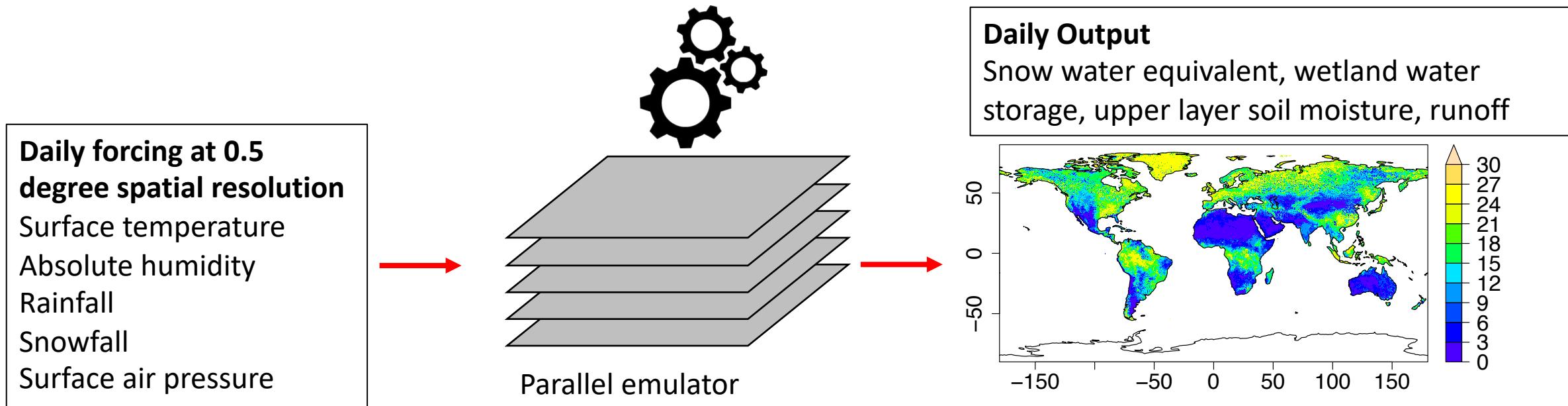
Vegetation Type	Site name	Location (lon, lat)	Observation period	Climate	Country
broadleaf evergreen forest	CN-Din	23.17, 112.54	2003 - 2005	Humid Subtropical	China
	ID-Pag	-2.32, 113.90	2002 - 2003	Tropical rain forest	Indonesia
	PT-Esp	38.64, -8.60	2002 - 2004	SubTropical	Portugal
	PT-Mil	38.54, -8.00	2005 - 2005	SubTropical	Portugal
mixed coniferous & broadleaf deciduous forest & woodland	AR-SLu	-33.46, -66.46	2010 - 2010	SubTropical	Argentina
	CN-Cha	42.40, 128.10	2003 - 2005	Temperate	China
	DE-Meh	51.28, 10.66	2004 - 2006	Temperate	Germany
	JP-SMF	35.26, 137.08	2003 - 2006	SubTropical	Japan
	US-Bar	44.06, -71.29	2005 - 2005	Temperate	USA
wooded & grassland	AU-Emr	-23.86, 148.47	2012 - 2013		Australia
	CN-Dan	30.85, 91.08	2004 - 2005	Arctic	China
	CN-Du2	42.05, 116.28	2007 - 2008	Temperate	China
	DK-Lva	55.68, 12.08	2005 - 2006	Temperate	Denmark
	IE-Dri	51.99, -8.75	2003 - 2005	Temperate	Ireland
	PL-wet	52.76 , 16.31	2004 - 2005	Temperate	Poland
cultivation	DE-Seh	50.87, 6.45	2008 - 2010	Temperate	Germany
	DK-Fou	56.48, 9.59	2005 - 2005	Temperate	Denmark
	IE-Ca1	52.86, -6.92	2004 - 2006	Temperate	Ireland
	IT-BCi	40.52, 14.96	2005 - 2010	SubTropical	Italy
	IT-CA2	42.38, 12.03	2012 - 2013	SubTropical	Italy

- **Parameters: plant function type (PFT) (12)**
- **Targeted output: Sensible heat and Latent heat**
- **Model: MATSIRO**

Evaluation of Calibrated ILS and Other LSMs with different resolution



Development of MATSIRO Emulator



- The emulator is written in coarray Fortran (parallel)
- The emulator is based on rudimentary physical equations that simplify the original MATSIRO equations
- Emulator parameters are optimized separately for each longitude, latitude, and month
- Relevant publication accepted in March 2024 in the Journal of Hydrology

Emulator Equations

$$\frac{dS_{snow}}{dt} = a(P_{snow} + A_{sn}P_{rain}) - bH(T_{air} - T_c)(T_{air} - T_c)^\chi$$

snow storage change snowfall rainfall snowmelt

$$P_{in} = cP_{melt} + d(1 - A_{sn})P_{rain} + \omega S_{wet}, A_{sn} = \min\left(\sqrt{\frac{S_{snow}}{100 \text{ kg m}^{-2}}}, 1\right)$$

water flux onto the soil surface

$$\frac{dy}{dt} = f(y, P_{in})P_{in} - gH(-S_{snow})(q_{air}^*(T_{air}, p_{air}) - q_{air}) - e(y - y_0)$$

top layer soil moisture percolation from above evaporation exchange with deep layer

$$\frac{dS_{wet}}{dt} = (1 - \alpha)(1 - f(y, P_{in}))P_{in} - \omega S_{wet}$$

wetland storage change surface runoff outflow

$$R_r = \alpha(1 - f(y, P_{in}))P_{in} + h$$

river runoff surface runoff baseflow

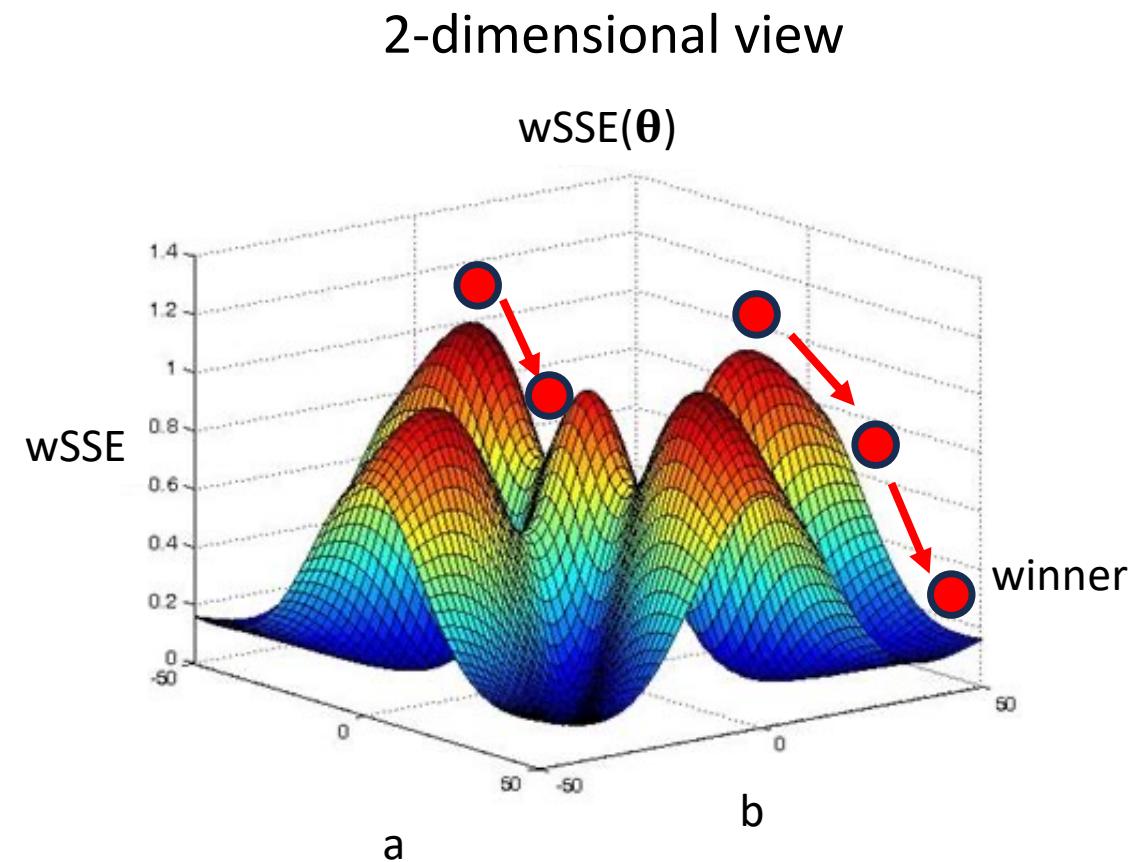
$$f(y, P_{in}) = \frac{1}{[1+e^{k_1(y-y_*)}][1+e^{k_2(P_{in}-P_{in}^*)}]} \times H(T_{air})$$

soil percolation fraction wetland outflow

$$\omega = \frac{1}{\beta}$$

Parameter Optimization

- The vector of parameters is $\theta = (a, b, T_c, \chi, c, d, \omega, g, e, y_o, h, k_1, y^*, k_2, P_{in}^*)$
- It is optimized for each latitude, longitude and month to minimize the objective function
- The objective function is the **mean weighted sum of squared errors (wSSE)** across all years between the emulator and MATSIRO state
- Different variables could be weighted differently, as per user settings
- The process can be repeated multiple times with initial starting values to pick the best result



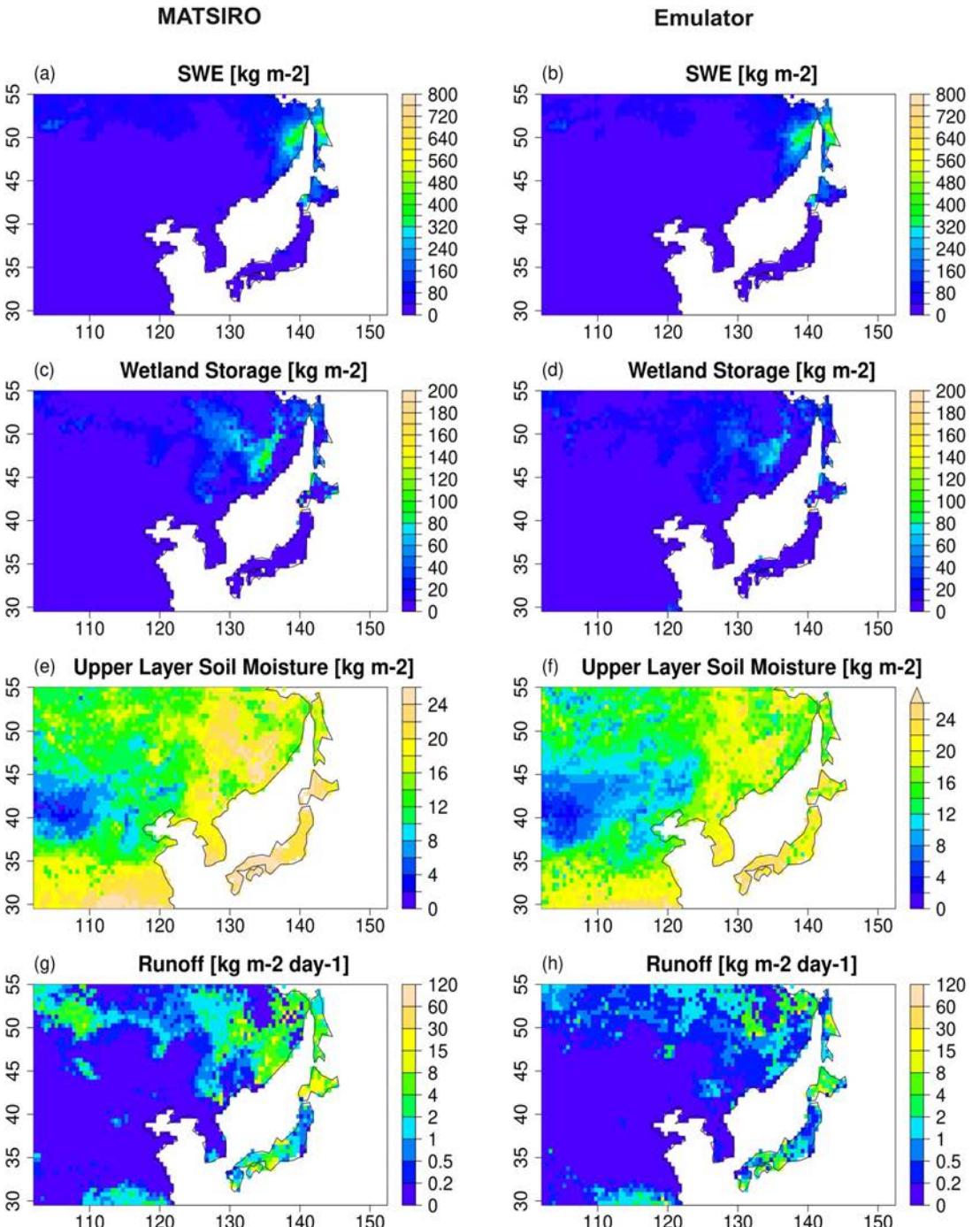
15 dimensions are used in reality*

Emulator Performance

- Training: 1911-2010
- Cross-validation: 1901-1910
- Single starting parameter value for optimization

Model	Time requirement per model year
MATSIRO	1448.19 s
Emulator (36 procs)	≈ 0.2 s

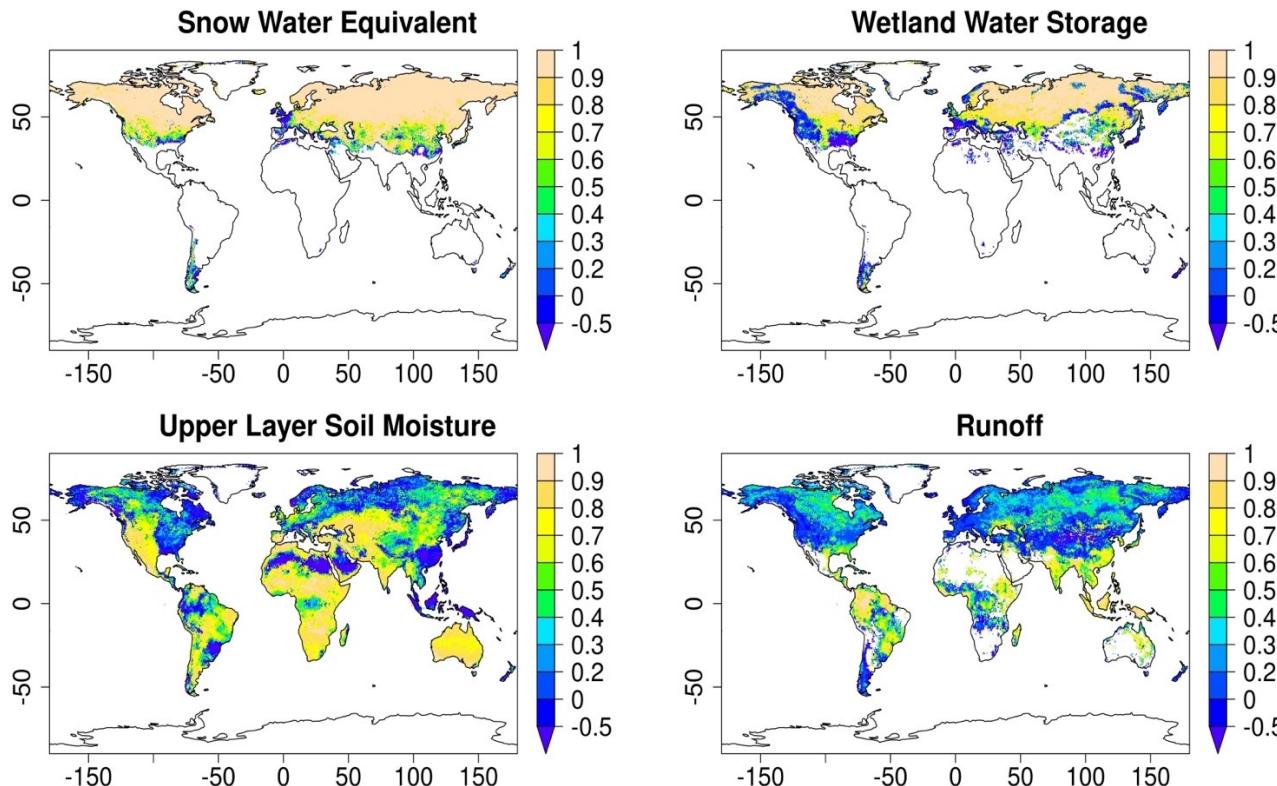
NOTE: I/O and parameter optimization are computationally expensive for the emulator



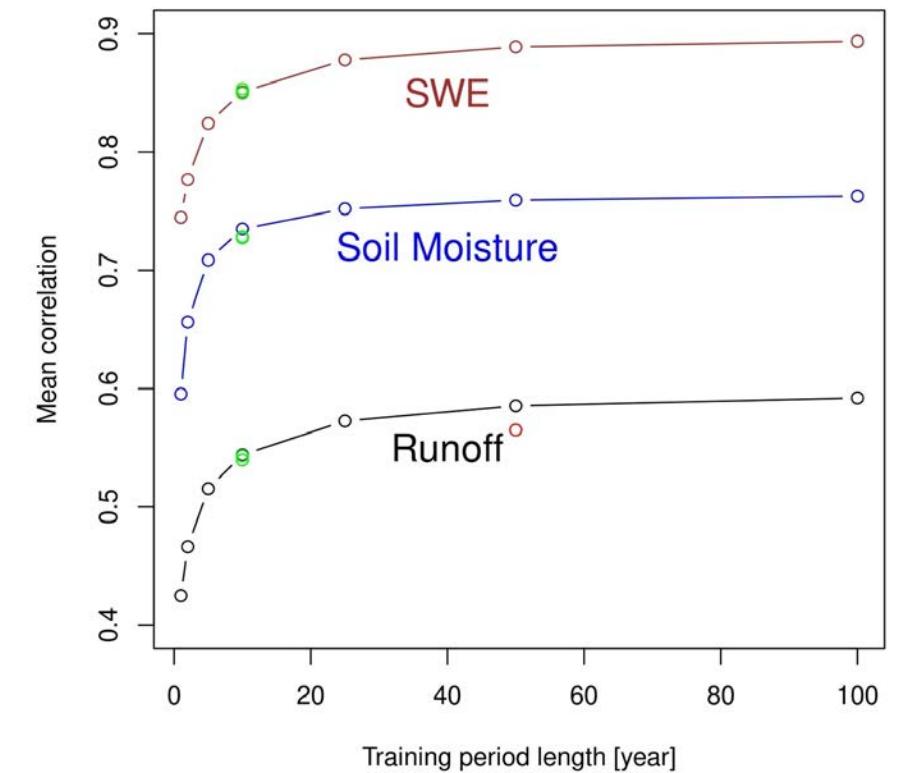
April 10, 1910

Effect of training period length on skill

Nash-Sutcliffe Efficiency for years 1901-1910



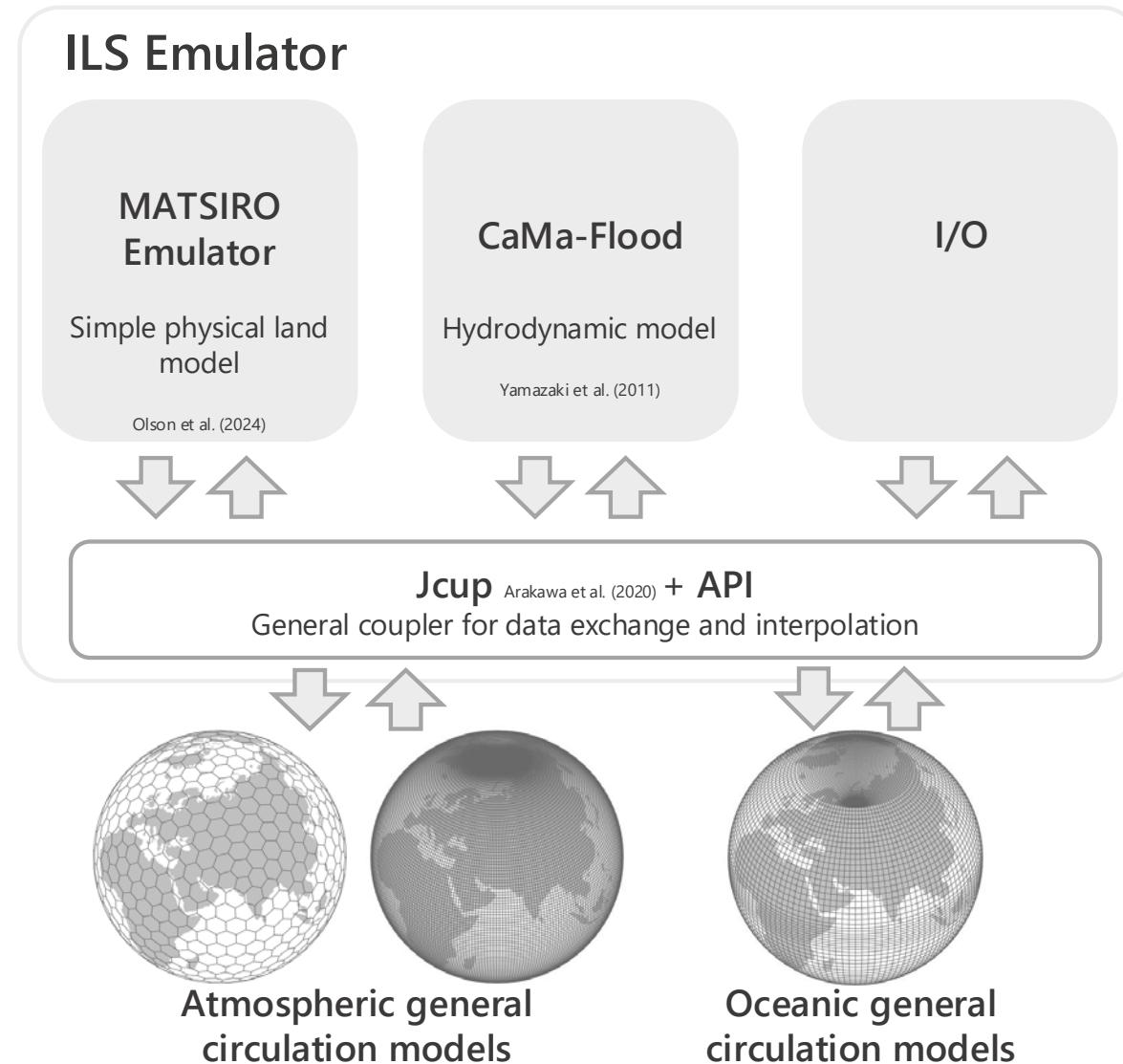
Values above 0.4 are acceptable for daily simulations (Moriasi et al., 2008)



- cross-validation, years 1901-1910
- cross-validation, years 2001-2010 (mild climate change)
- different training periods (1921-1930 and 1931-1940)

Implement Emulator to Integrated Land Simulator (ILS)

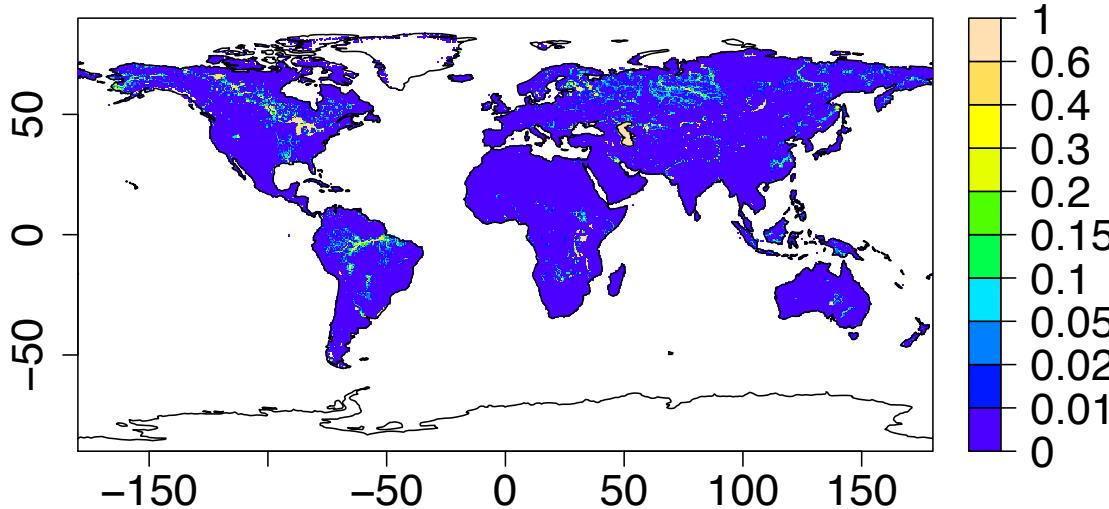
- The purpose of the emulator is to speed up the Integrated Land Simulator (ILS)
- Here we evaluate the skill of the emulator at representing the original ILS, as well as computational efficiency
- MATSIRO emulator improved through
 - Incorporation of baseflow
 - Implicit numerical scheme
 - Global parameter optimization using differential evolution



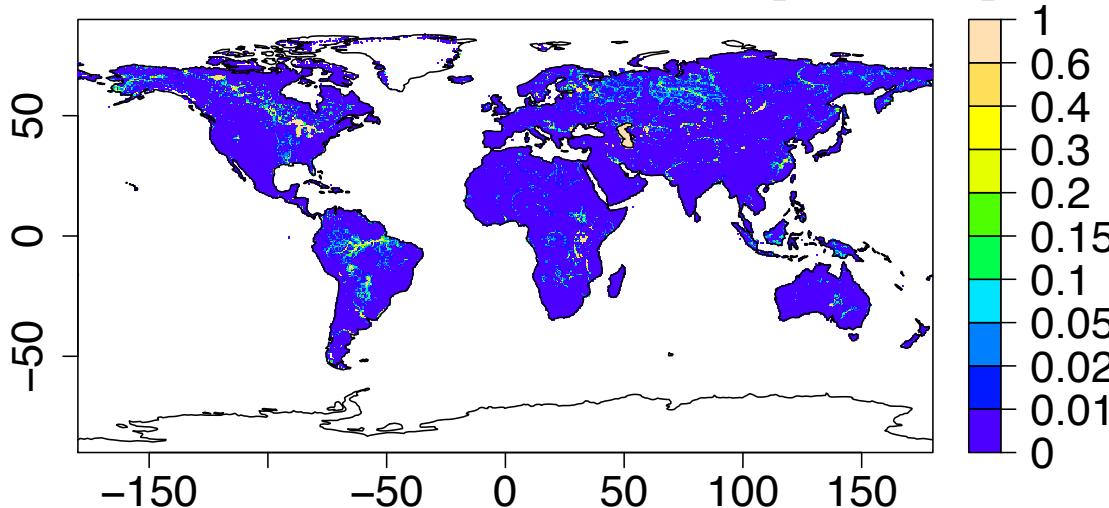
Implement Emulator to Integrated Land Simulator (ILS)

May 2009

ILS Flood Fraction [Unitless]

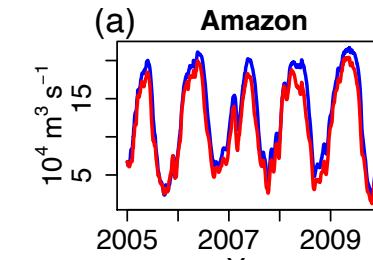


Emulated Flood Fraction [Unitless]

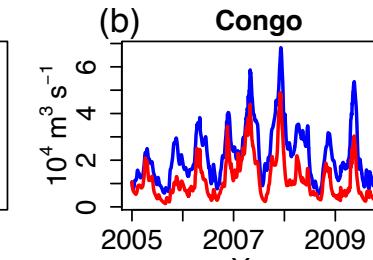


Discharge for major world rivers

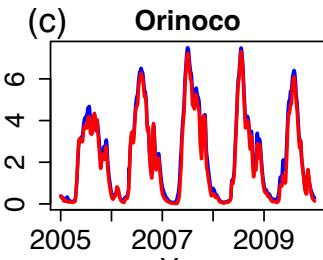
(a) Amazon



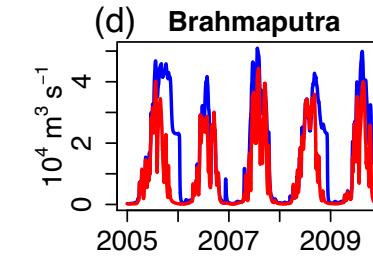
(b) Congo



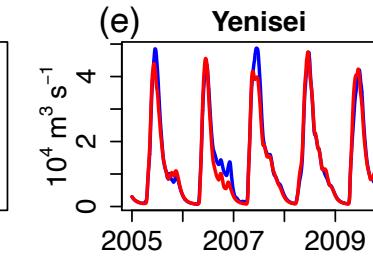
(c) Orinoco



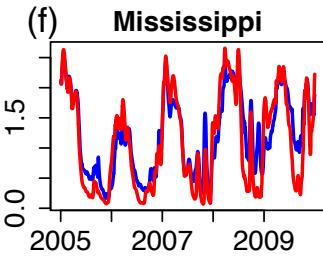
(d) Brahmaputra



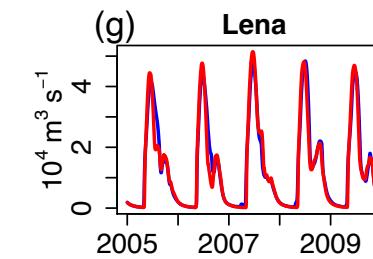
(e) Yenisei



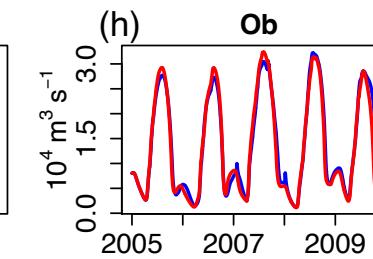
(f) Mississippi



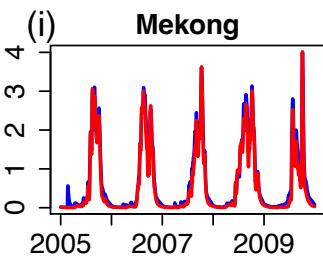
(g) Lena



(h) Ob



(i) Mekong



emulator

ILS

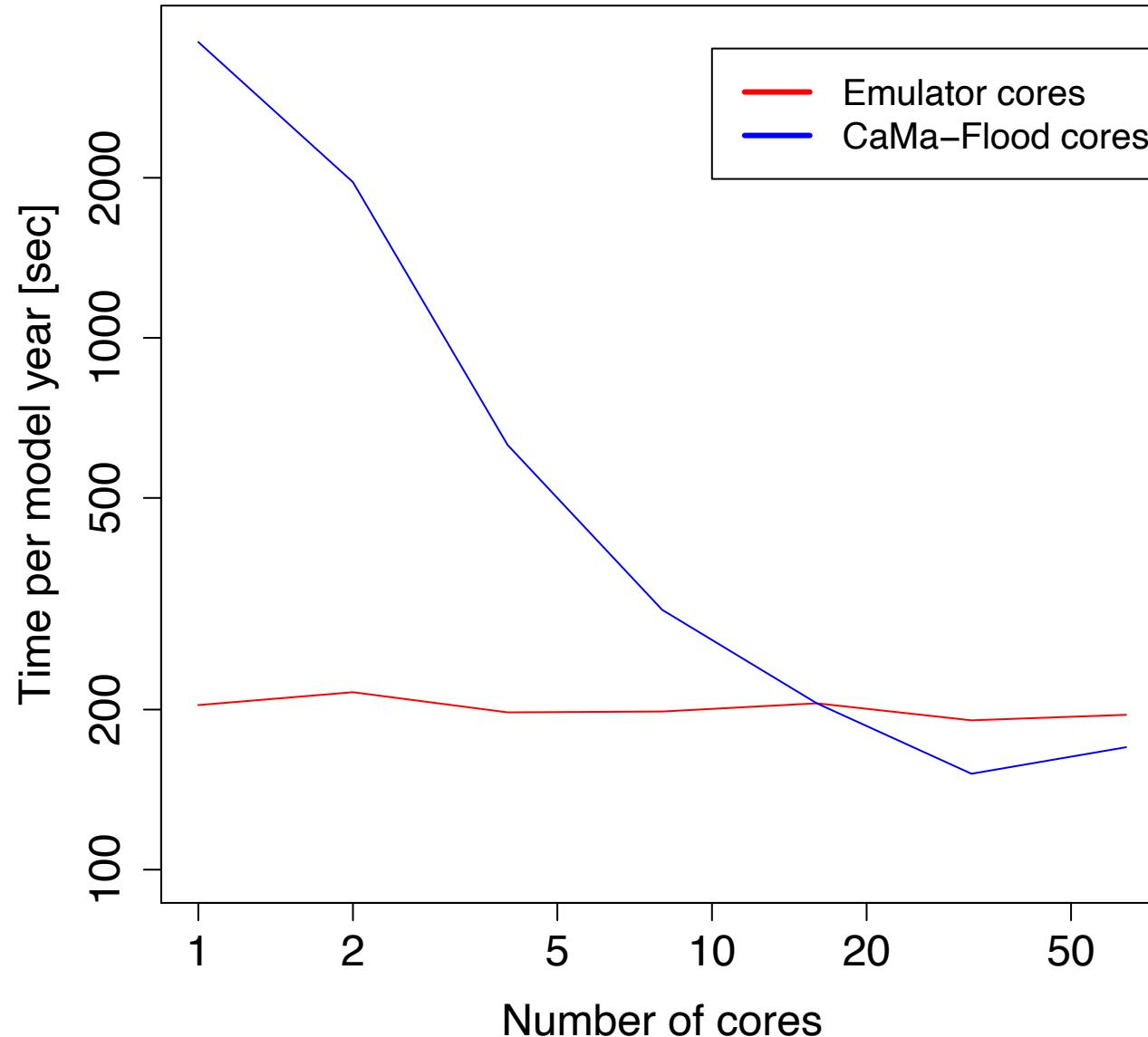
Implement Emulator to Integrated Land Simulator (ILS)

* I/O always uses one core

Corresponding time for original ILS was **1719 sec**

- Using 50 cores for MATSIRO, 7 for CaMa-Flood, and 7 for I/O.

In terms of the node-hours metric, the emulator is **~30 times faster** than the ILS



GRACE TWS Estimation by ML

1. Can GRACE TWS Anomalies (TWSA) be downscaled to a finer spatial resolution with high accuracy using a deep learning approach?
2. Can the downscaled TWSA better capture the water mass variation at a sub-regional to local scale for flood and drought monitoring?

Yin, G., Park, J., & Yoshimura, K. (2025). Spatial Downscaling of GRACE Terrestrial Water Storage Anomalies for Drought and Flood Potential Assessment. Journal of Hydrology. (Accepted)

Data Sets



GRACE/GRACE-FO

- **TWSA**
- Resolution: ~monthly, $3^\circ \times 3^\circ$ equal area caps (mascon)
- Jet Propulsion Laboratory (JPL) Release06 Version 2

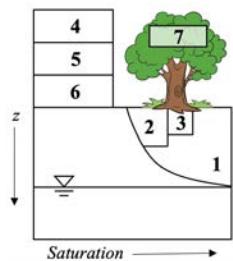
} Predictand
(real-world)



Today's Earth – Global (TE)

- **Precipitation, temperature, canopy interception water, snow water equivalent, soil moisture, latent heat, surface runoff**
- Resolution: daily, $0.5^\circ \times 0.5^\circ$
- MATSIRO + CaMa-Flood

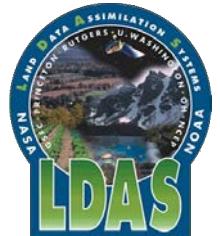
} Predictors



Catchment Land Surface Model (CLSM)

- **Groundwater**
- Resolution: 3 hourly, $1^\circ \times 1^\circ$
- GLDAS Catchment Land Surface Model Level 4 V2.1

} Predictand
(synthetic)

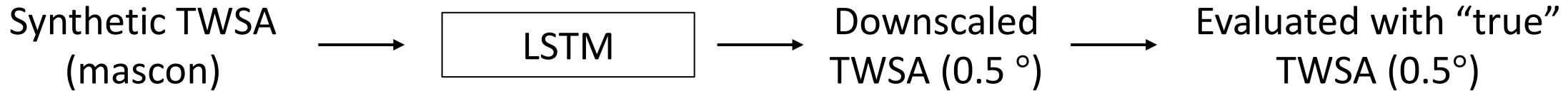


Noah Land Surface Model

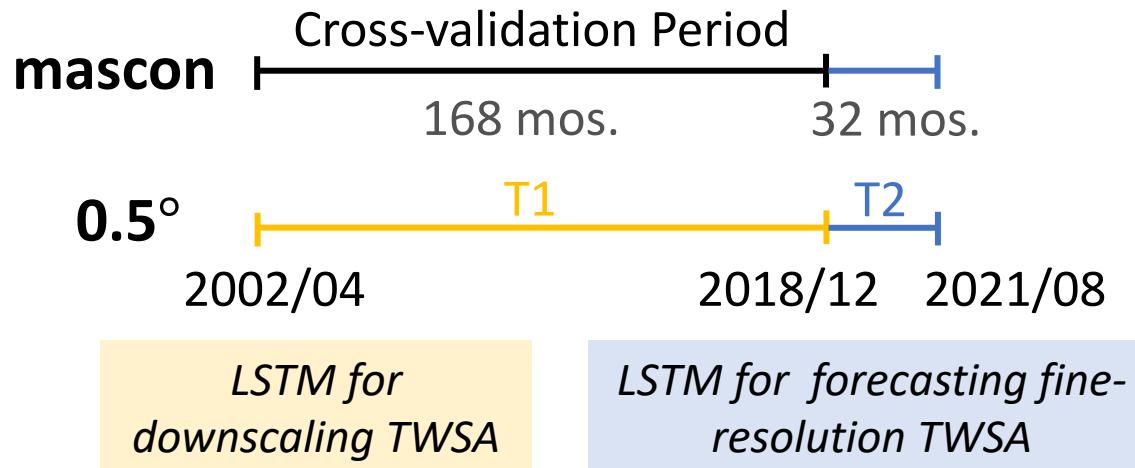
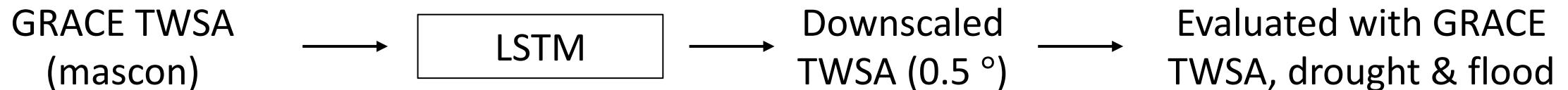
- **TWSA**
- Resolution: 3 hourly, $0.25^\circ \times 0.25^\circ$
- GLDAS Noah Land Surface Model Level 4 V2.1

Experiment Design

Experiment I (Synthetic)



Experiment II (Real)



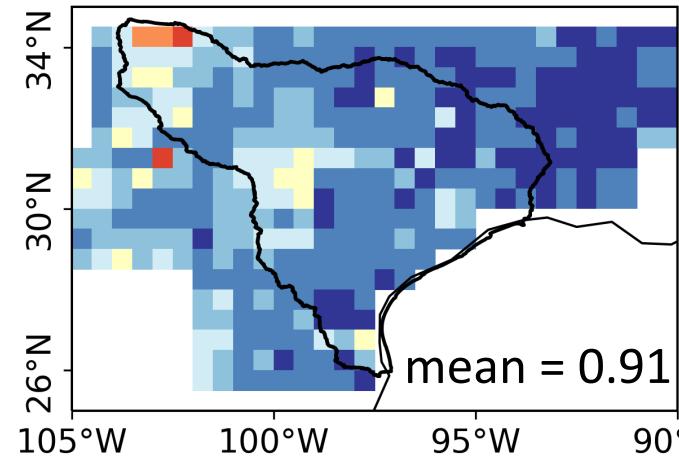
Notes:

- ✓ Cross-validation period: 2002/04 – 2018/12
- ✓ Downscaling period: 2002/08 – 2021/08
- ✓ Build a LSTM model for each mascon grid
- ✓ Lag time up to 2 months selected; 50 ensembles
- ✓ Sensitivity analysis to find the optimal combination of features

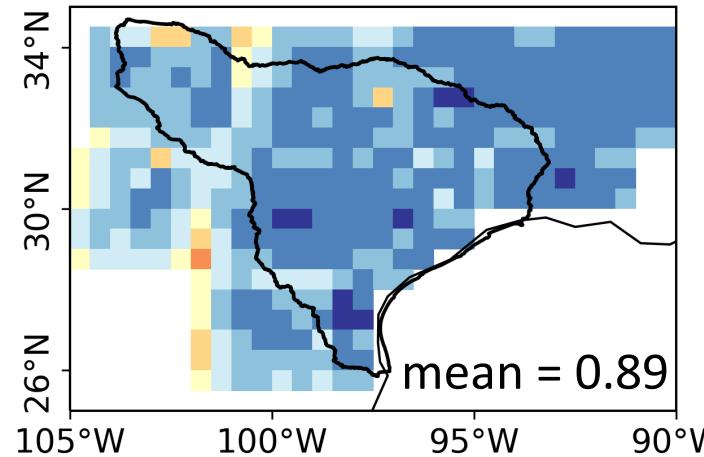
Synthetic Experiment Results

T1

LSTM

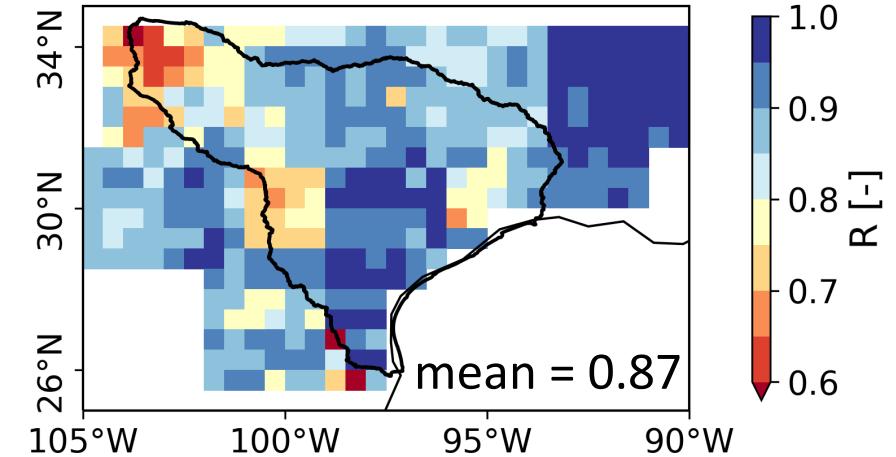


Interpolation

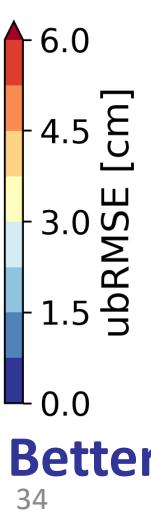
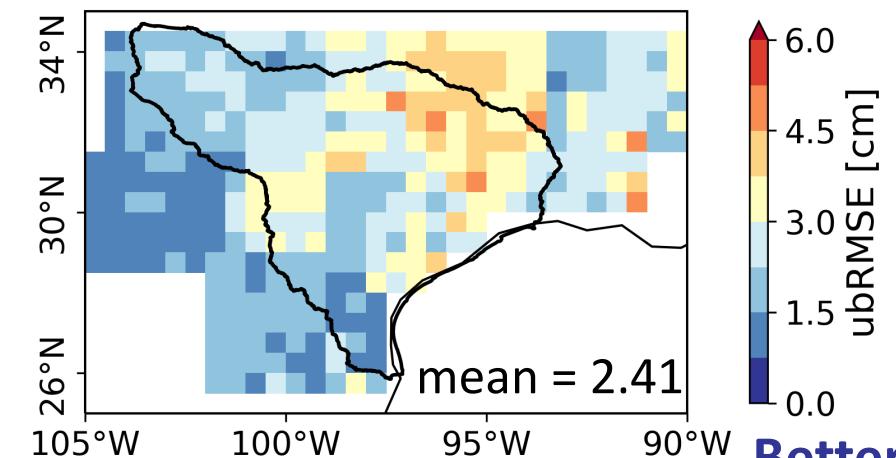
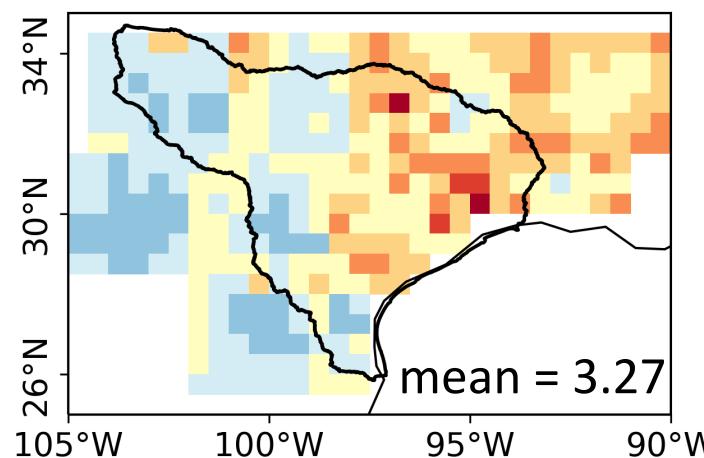
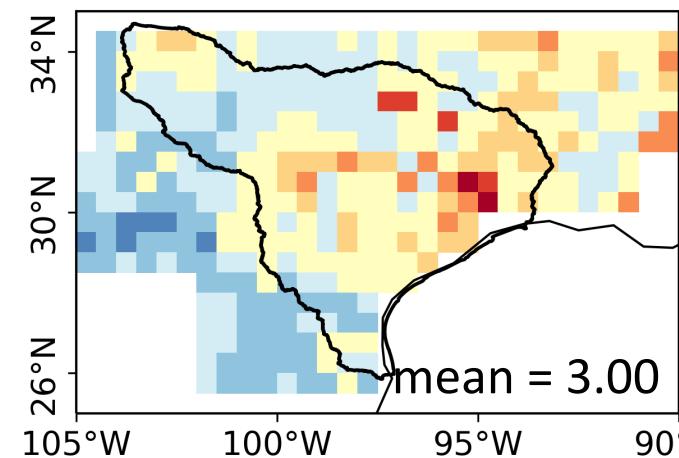
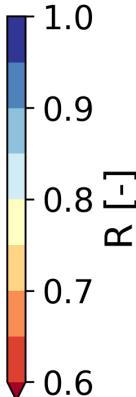


T2

LSTM



Better



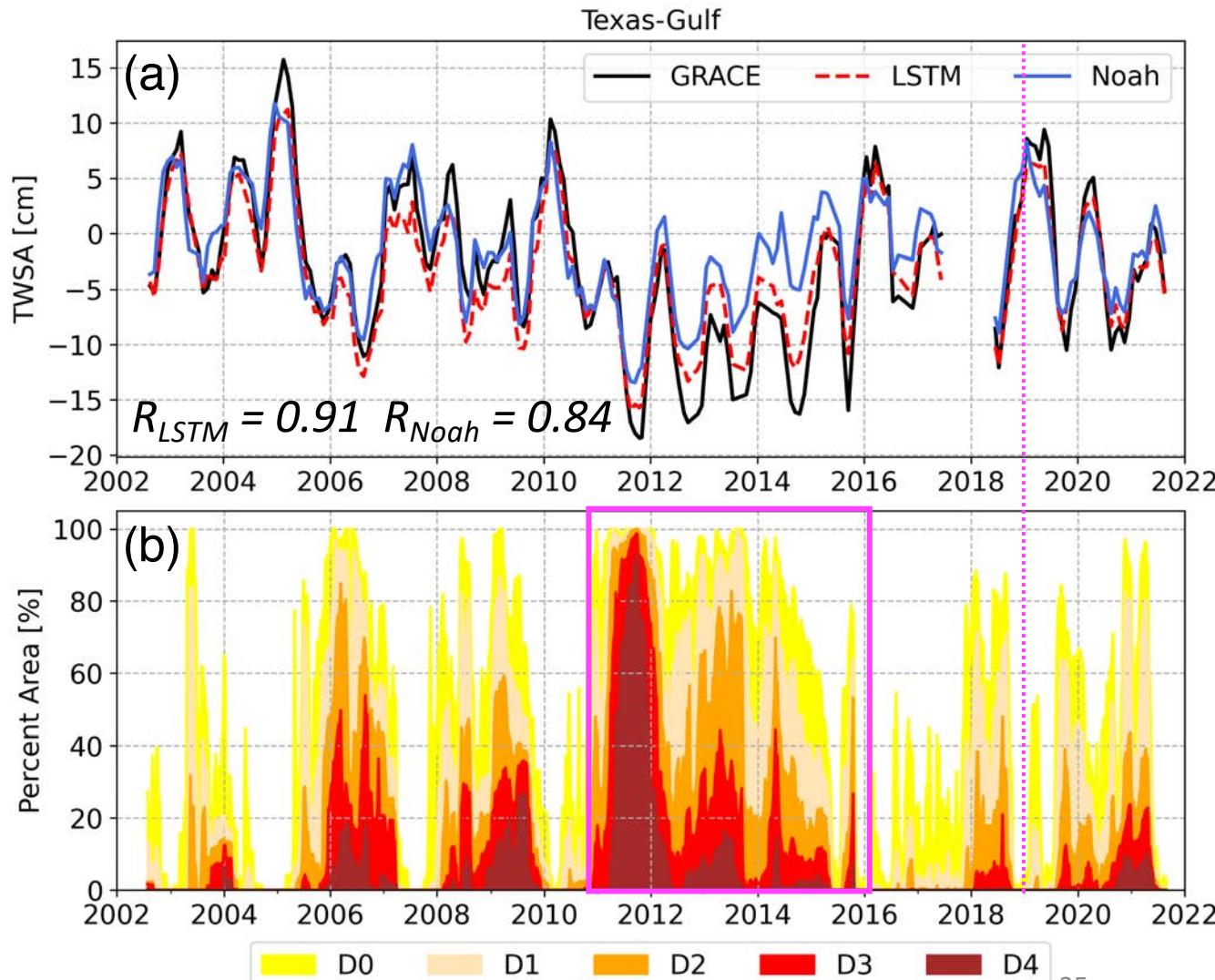
Real-World Experiment Results

U.S. Drought Monitor (USDM)

- Map released every Thursday showing parts of the U.S. that are in drought.
- Five level classifications (D0-D4)
- Relies on experts to synthesize the best available data from multiple sources.

Category	Description	Possible Impacts	Ranges				
			Palmer Drought Severity Index (PDSI)	CPC Soil Moisture Model (Percentiles)	USGS Weekly Streamflow (Percentiles)	Standardized Precipitation Index (SPI)	Objective Drought Indicator Blends (Percentiles)
D0	Abnormally Dry	Going into drought: <ul style="list-style-type: none"> • short-term dryness slowing planting, growth of crops or pastures Coming out of drought: <ul style="list-style-type: none"> • some lingering water deficits • pastures or crops not fully recovered 	-1.0 to -1.9	21 to 30	21 to 30	-0.5 to -0.7	21 to 30
D1	Moderate Drought	<ul style="list-style-type: none"> • Some damage to crops, pastures • Streams, reservoirs, or wells low, some water shortages developing or imminent • Voluntary water-use restrictions requested 	-2.0 to -2.9	11 to 20	11 to 20	-0.8 to -1.2	11 to 20
D2	Severe Drought	<ul style="list-style-type: none"> • Crop or pasture losses likely • Water shortages common • Water restrictions imposed 	-3.0 to -3.9	6 to 10	6 to 10	-1.3 to -1.5	6 to 10
D3	Extreme Drought	<ul style="list-style-type: none"> • Major crop/pasture losses • Widespread water shortages or restrictions 	-4.0 to -4.9	3 to 5	3 to 5	-1.6 to -1.9	3 to 5
D4	Exceptional Drought	<ul style="list-style-type: none"> • Exceptional and widespread crop/pasture losses • Shortages of water in reservoirs, streams, and wells creating water emergencies 	-5.0 or less	0 to 2	0 to 2	-2.0 or less	0 to 2

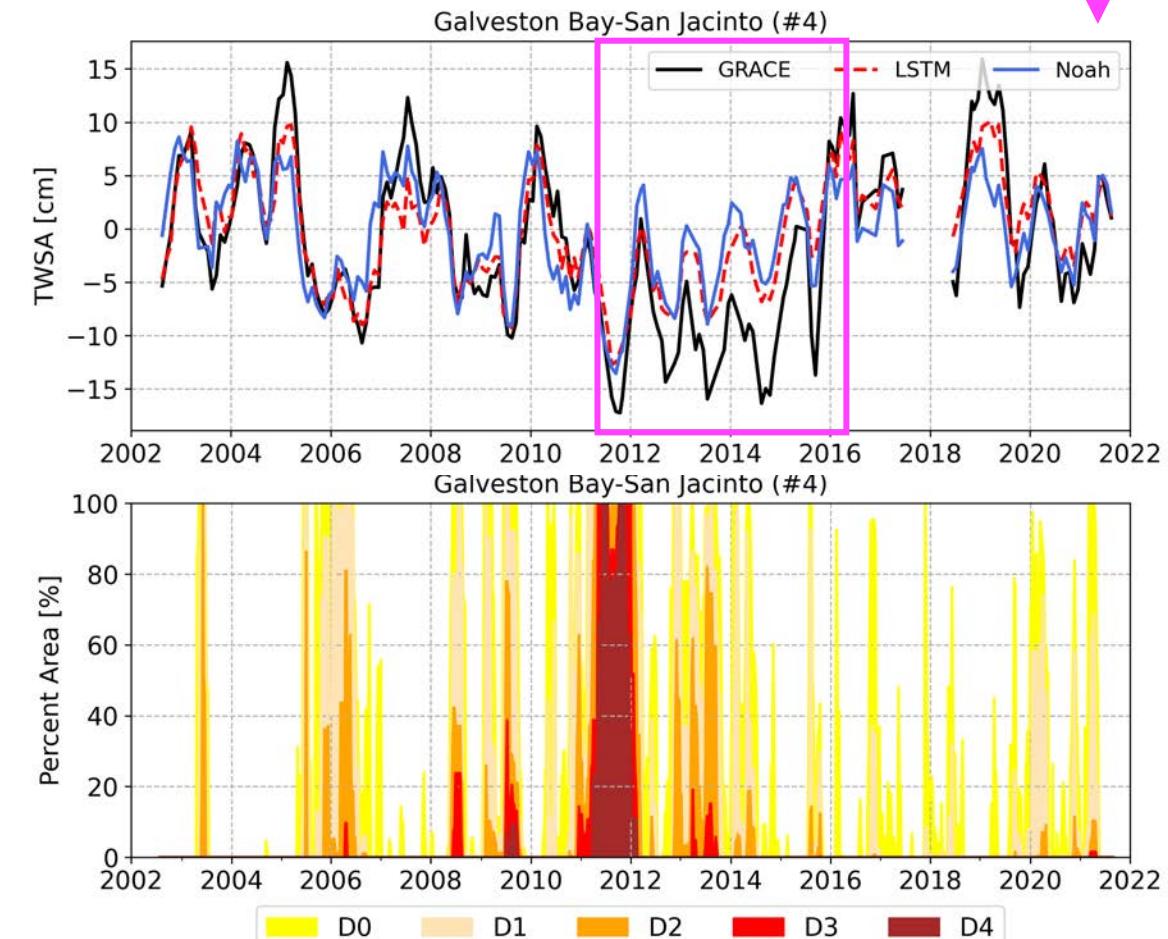
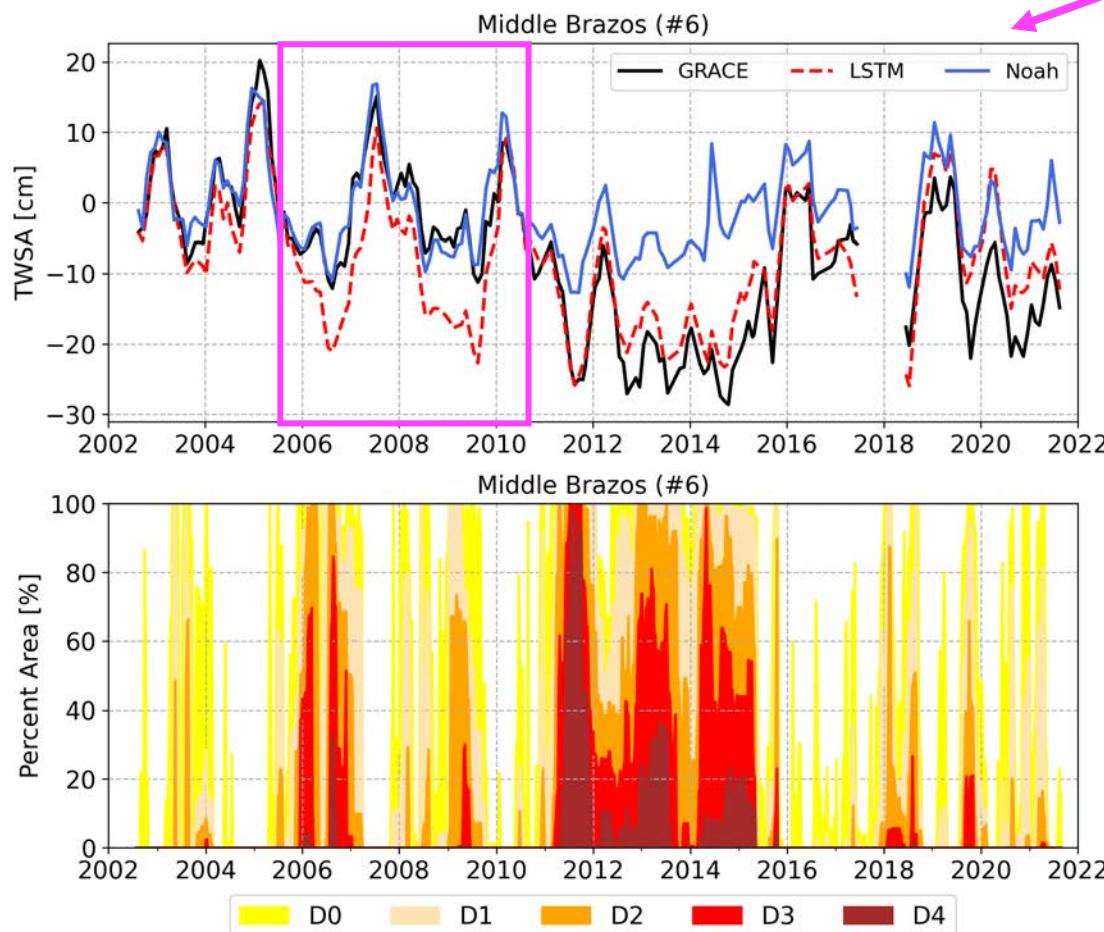
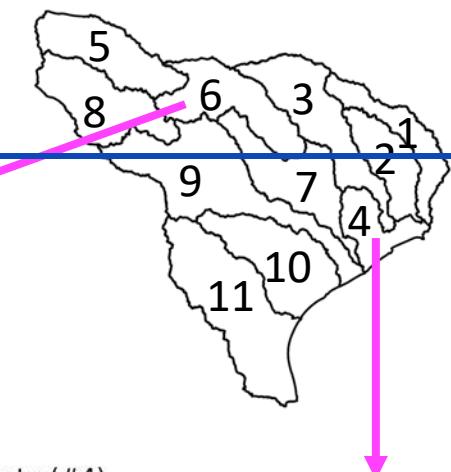
❖ TWSA versus USDM at basin scale



Real-World Experiment Results

❖ TWSA versus USDM at sub-basin scale

Downscaled TWSA better captured the variation of TWS at sub-region scale.



Real-World Experiment Results

❖ Flood Analysis

Storm Events Database

- ❑ National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI)
- ❑ Documents the occurrence of storms and other significant weather phenomena (e.g., flash flood, flood, hail) having sufficient intensity to cause lose.

Flood Potential Index (FPI)

- ❑ A quantitative, effective storage-based indicator of when a region is at risk of flooding (Reager & Famiglietti, 2009)
- ❑ Emphasize the information contained within the GRACE data pertinent

$$S_{DEF}(t) = S_{MAX} - S(t - 1)$$

$$F(t) = P_{MON}(t) - S_{DEF}(t)$$

$$FPI(t) = \frac{F(t)}{\max(F(t))}$$

S_{DEF} : storage deficit

S_{MAX} : historical maximum storage anomaly

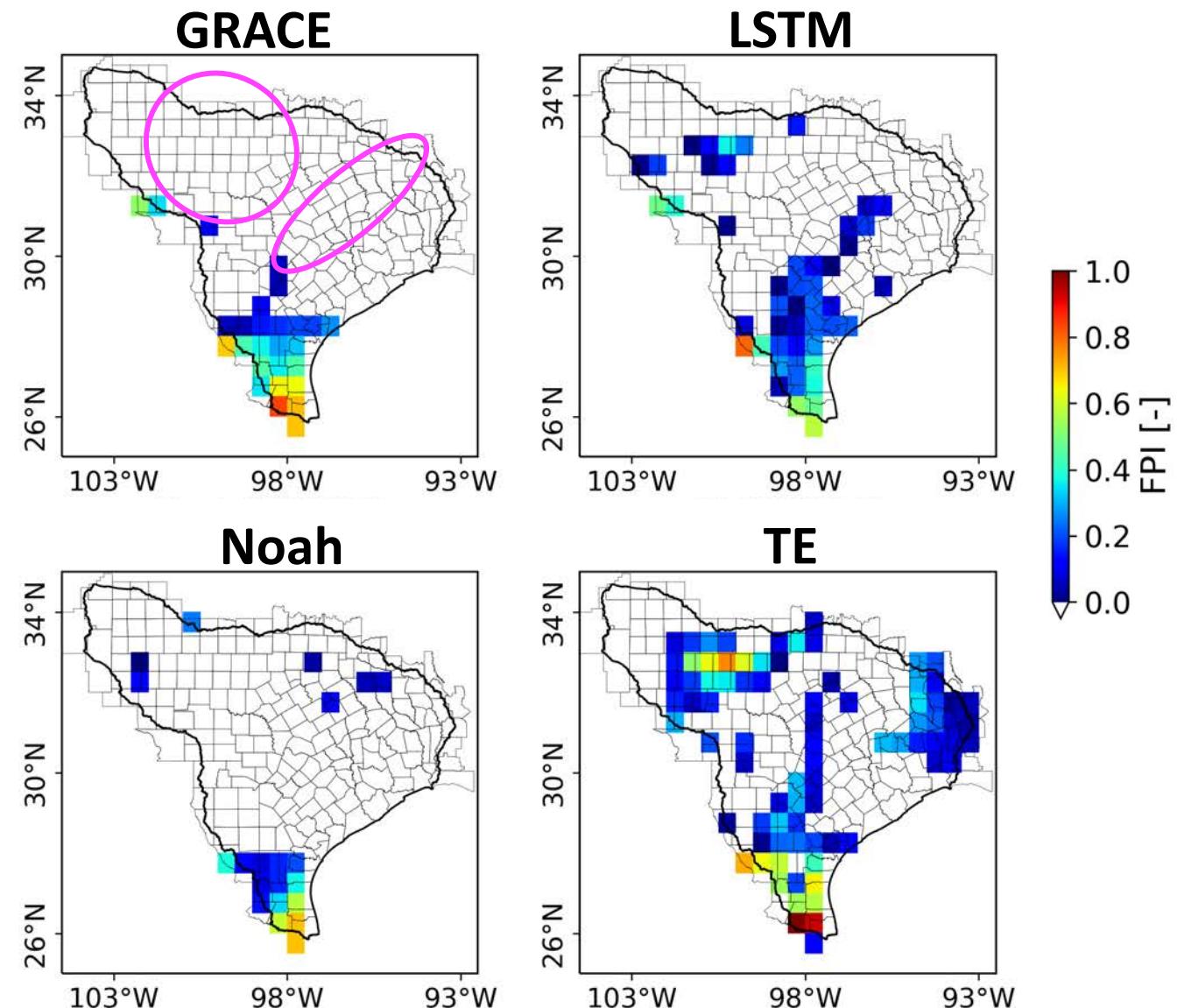
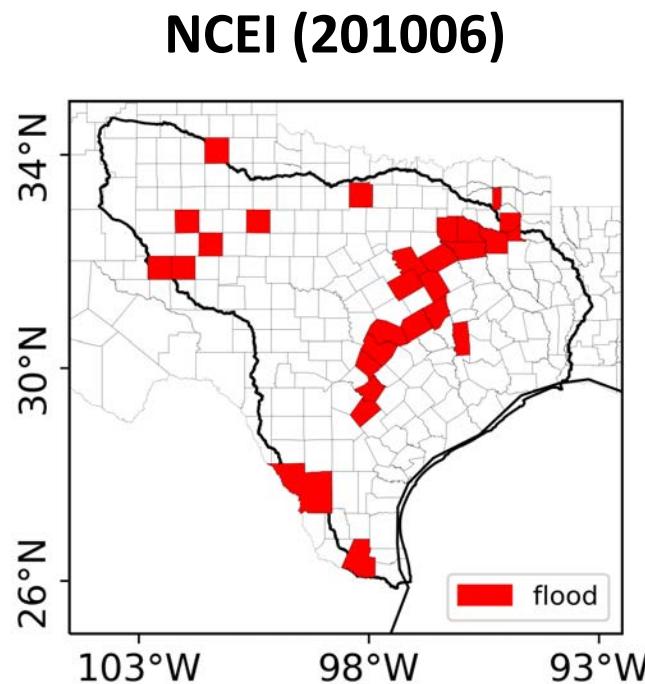
$S(t - 1)$: TWSA of the previous month

P_{MON} : monthly cumulative precipitation (from CPC in the study)

FPI → 1, high flood likelihood

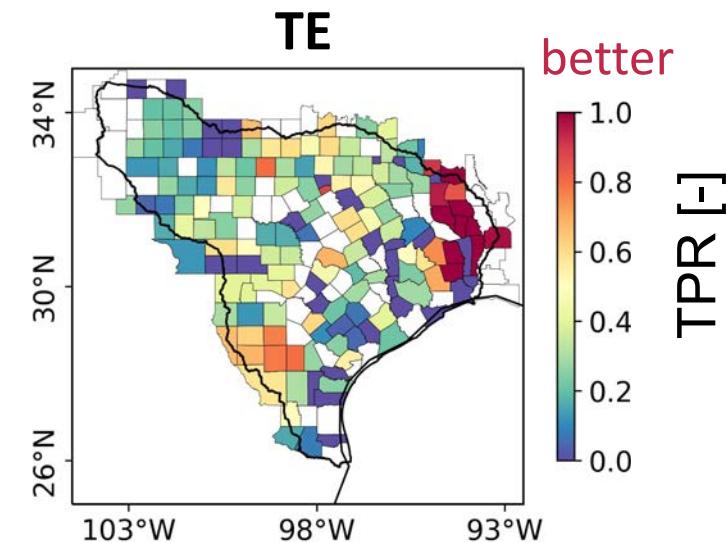
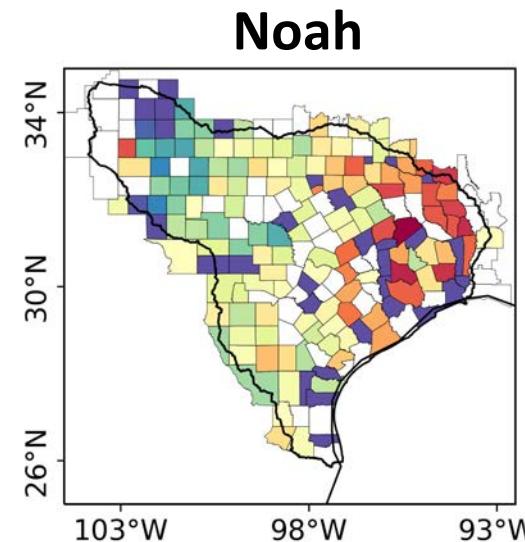
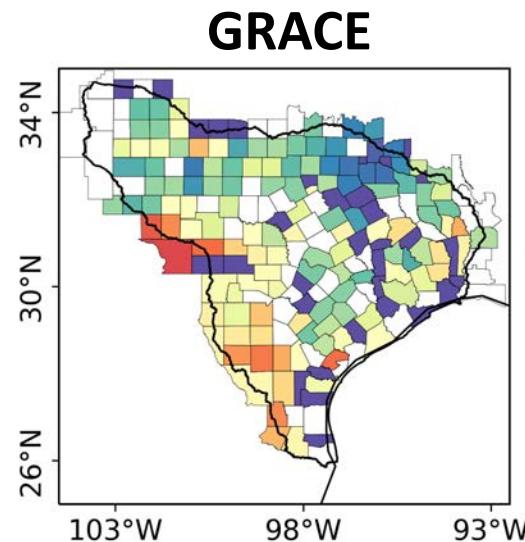
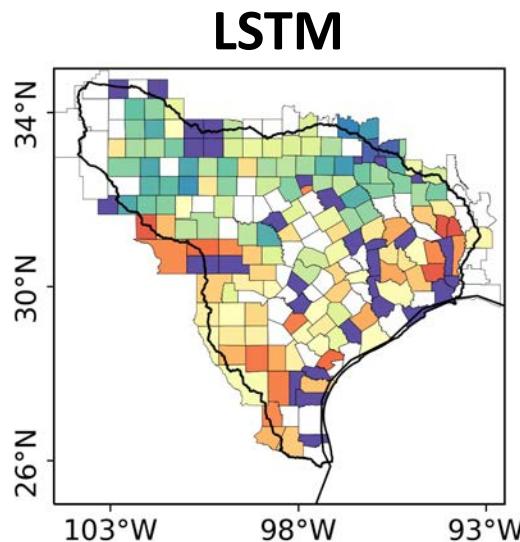
Real-World Experiment Results

❖ FPI versus NCEI



Real-World Experiment Results

❖ FPI versus NCEI

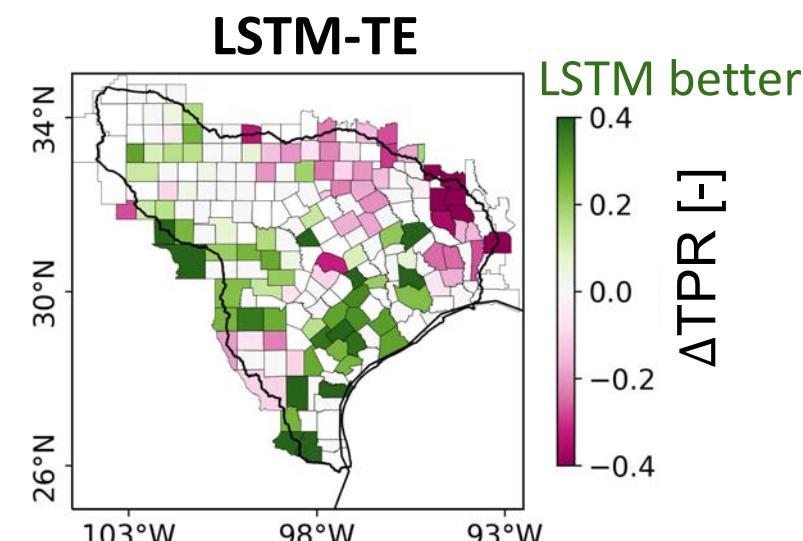
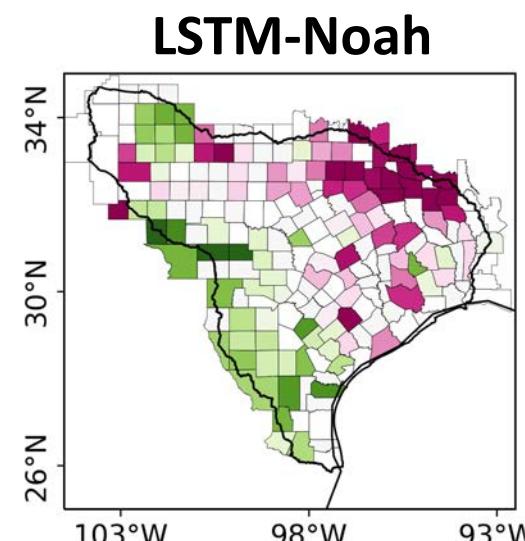
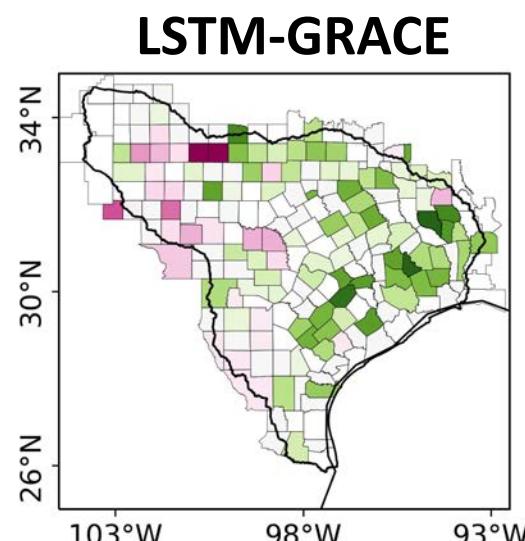


better

TPR [-]

True Positive Rate

$$TPR = \frac{TP}{TP+FN}$$



LSTM better

ΔTPR [-]

Summary and Publication List

I introduced our efforts to develop Machine learnings for Terrestrial Simulations.

1. Downscaling and bias correction of atmospheric forcing data (for both historical and future)
 - Yoshikane, T. and K. Yoshimura, PLOS water, 2022. <https://doi.org/10.1371/journal.pwat.0000016>
 - Yin, G., T. Yoshikane, K. Yoshimura, K. Yamamoto, and T. Kubota, J. Hydrol., 2022. <https://doi.org/10.1016/j.jhydrol.2022.128125>
 - Yoshikane, T. and K. Yoshimura, Sci. Rep. 13, 9412, 2023. <https://doi.org/10.1038/s41598-023-36489-3>
 - Yin, G., T. Yoshikane, R. Kaneko, and K. Yoshimura, JGR-Atmos, 2023. <https://doi.org/10.1029/2023JD038929>
2. Improvement of physical understanding
 - Li, H., G. Zhao, W. Xie, R. Olson, and K. Yoshimura, in revision.
3. dPL Parameter calibration of LSM/HM
 - Xie, W. et al., in prep.
4. Speed-up of LSM calculation using Emulator
 - Olson, R., T. Nitta, and K. Yoshimura, J. Hydrology, 2024. <https://doi.org/10.1016/j.jhydrol.2024.131093>
 - Olson, R., T. Nitta, T. Arakawa, and K. Yoshimura, submitted.
5. Terrestrial Water Storage estimation
 - Yin, G., J. Park, and K. Yoshimura, J. Hydrol., 2025. (Accepted)

Thank you!