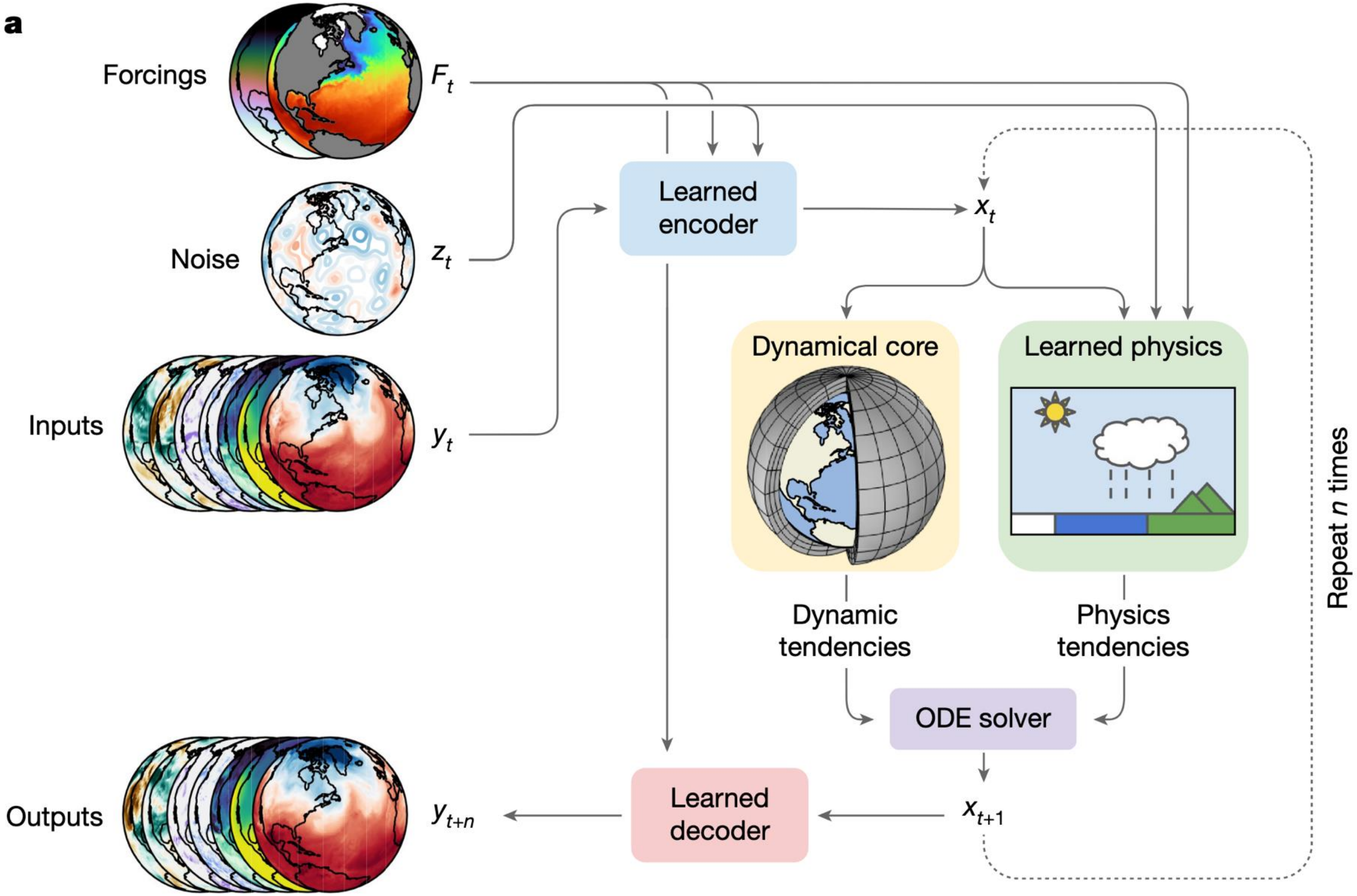
Neural general circulation models for modeling precipitation

30 January 2026 / Dominik Stiller / MLJC

Recap: NeuralGCM (Kochkov et al., 2024)

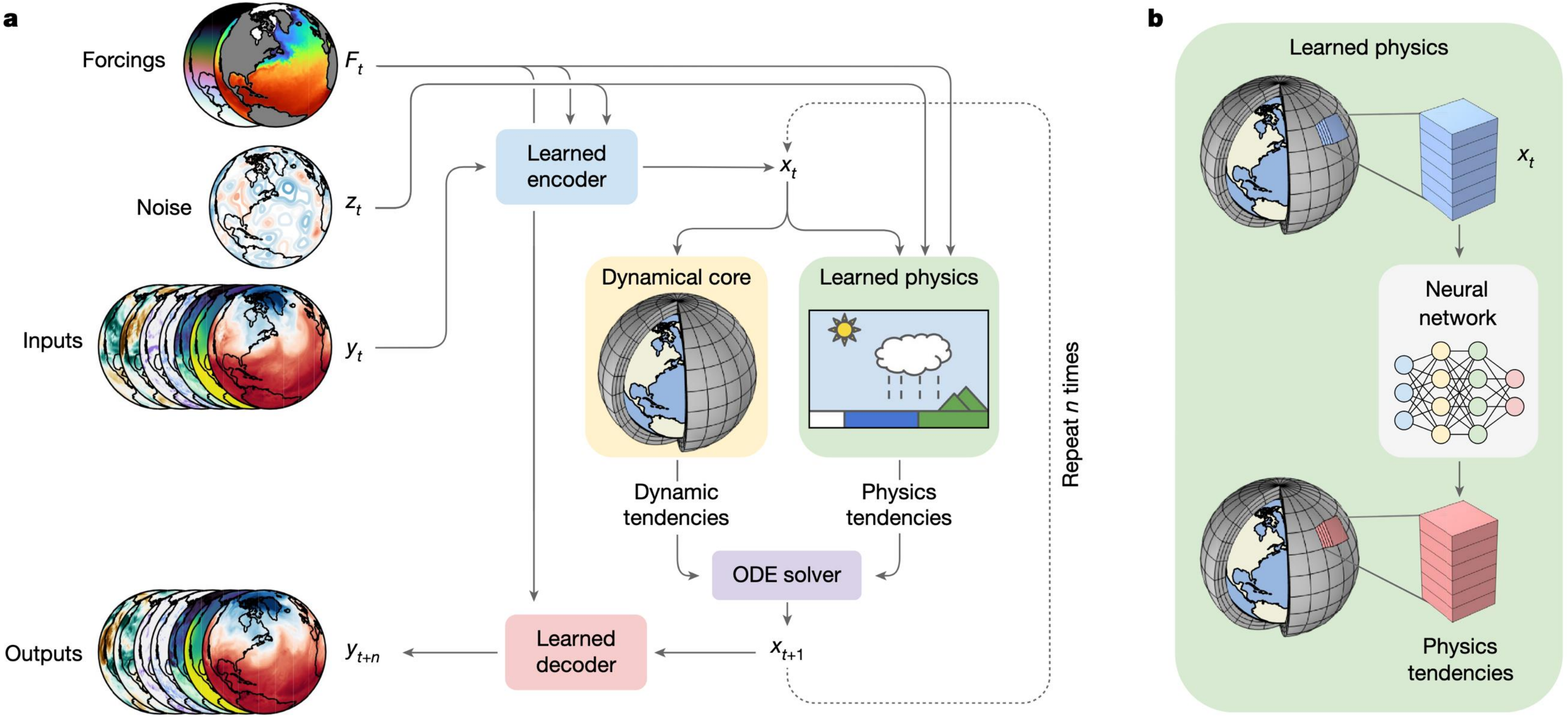
NeuralGCM is a hybrid model:

- Traditional dynamical core
- ML subgrid + physics parameterizations

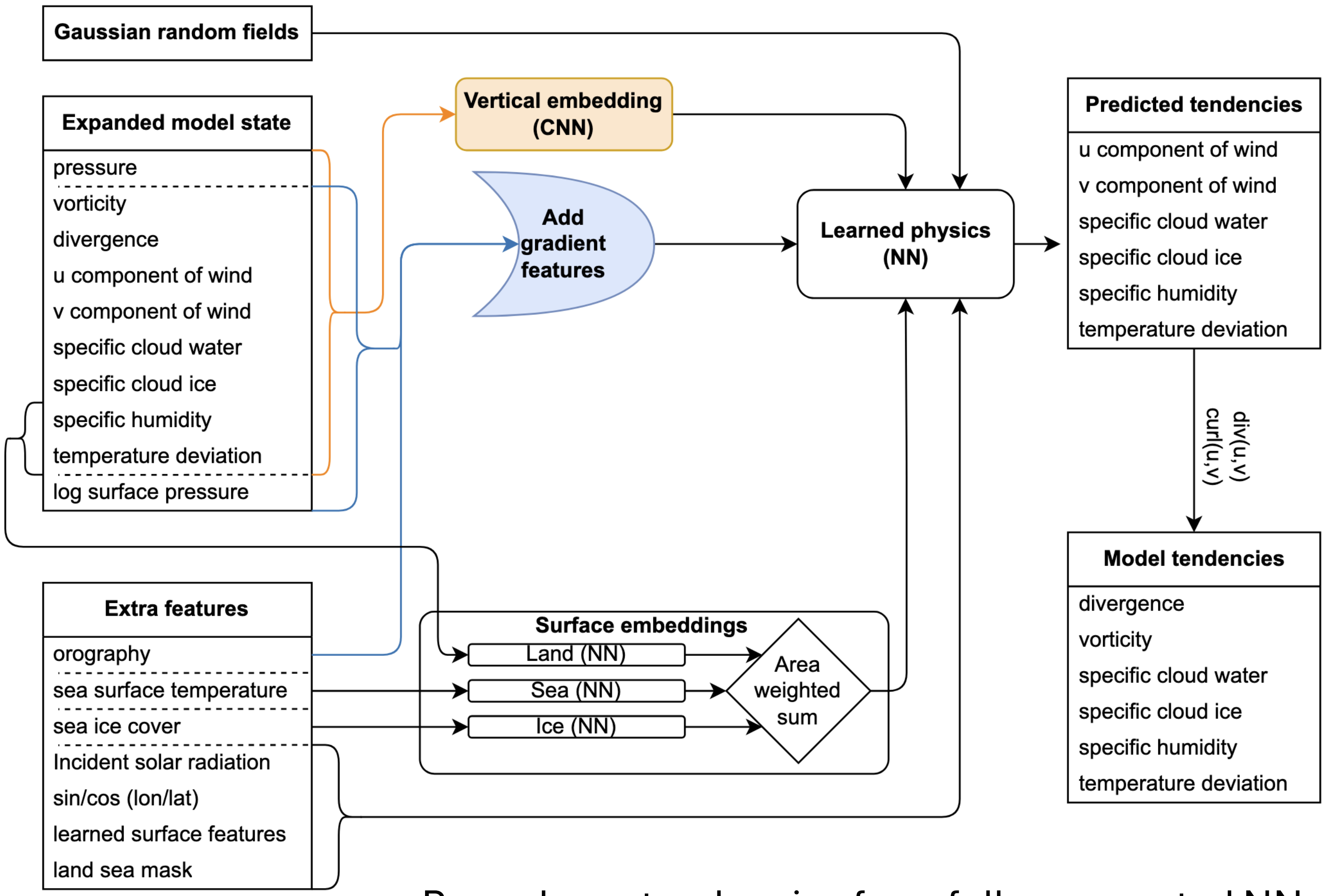


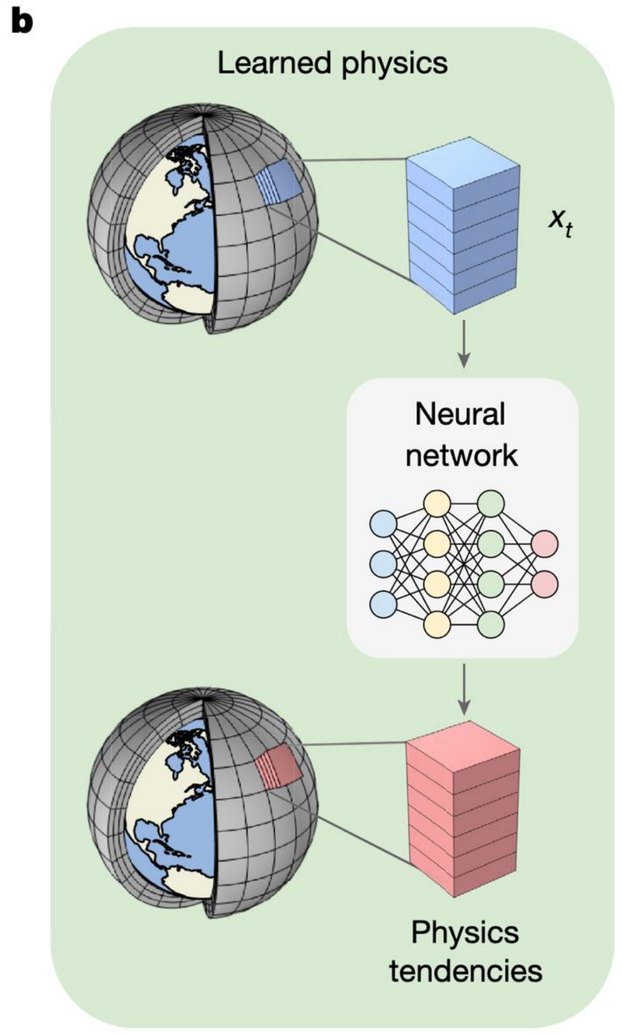
Recap: NeuralGCM (Kochkov et al., 2024)

NeuralGCM is a hybrid model:
- Traditional dynamical core
- ML subgrid + physics parameterizations



Recap: NeuralGCM (Kochkov et al., 2024)





Per-column tendencies from fully-connected NN

P−E in NeuralGCM (Kochkov et al., 2024)

- NeuralGCM does not diagnose precipitation, only P−E (precipitation minus evaporation)

$$P - E = \frac{1}{g}\int_0^1 \sum_{\mathrm{i}} \left(\frac{dq}{dt}\right)_{\mathrm{i}}^{\mathrm{NN}} p_{\mathrm{s}} d\sigma \qquad [\mathrm{kg\ m^{-2}\ s^{-1}}]$$

- dq/dt_i = column tendency of water species (vapor, liquid, ice)
- $p_s\, d\sigma = dp$ (integrating over all pressure levels)
- $\sigma = 1$ at surface (integral is from highest to lowest level)

"For weather forecasting, we expect that end-to-end learning with observational data will allow for better and more relevant predictions, including key variables such as precipitation."

SCIENCE ADVANCES | RESEARCH ARTICLE

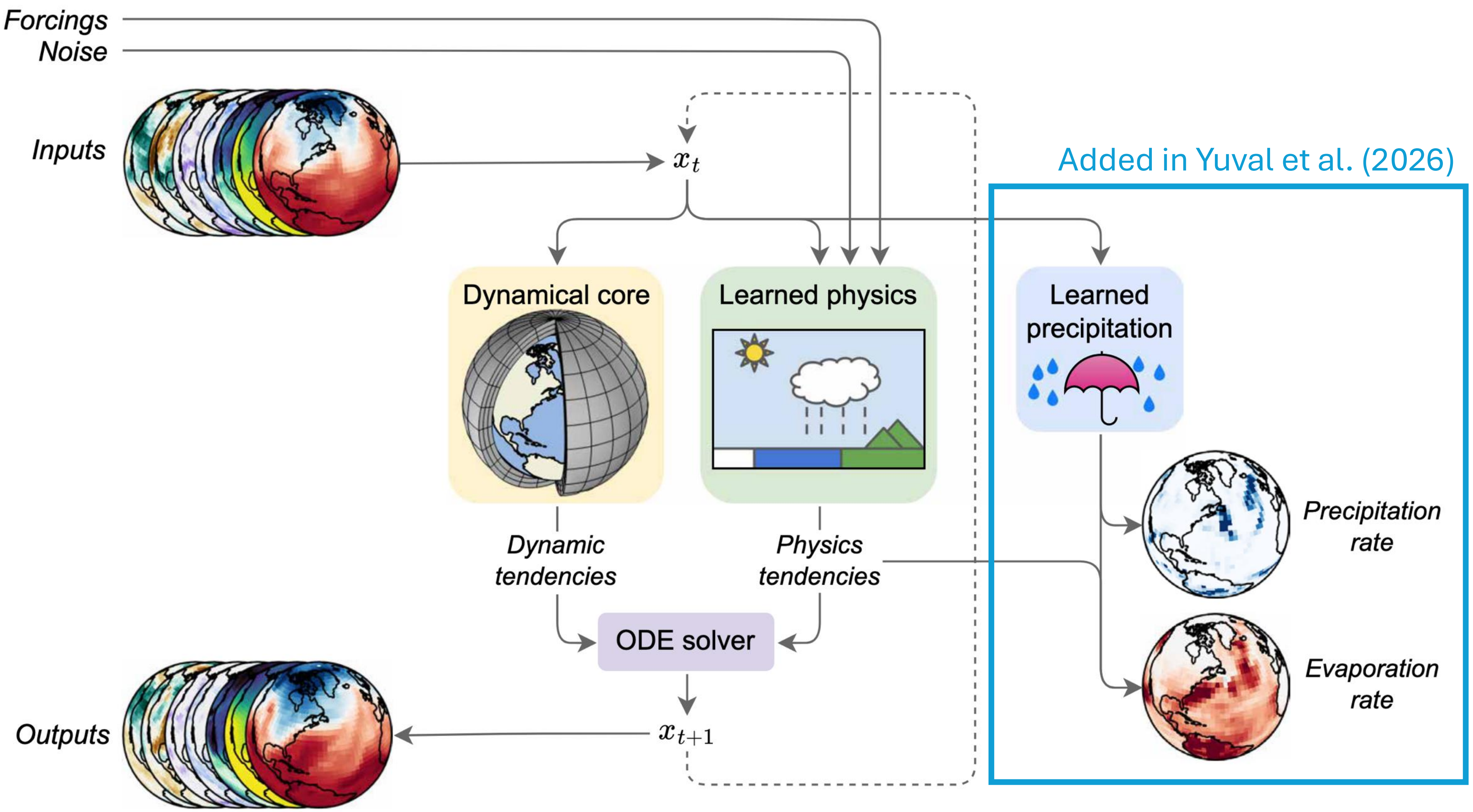
ATMOSPHERIC SCIENCE

Neural general circulation models for modeling precipitation

Janni Yuval*†, Ian Langmore†‡, Dmitrii Kochkov†, Stephan Hoyer†

- Trained to predict precipitation from satellite-based precipitation observations
- Consistent with column water budget (P−E)
- Better mid-range forecasts and climatology of precipitation than other models

Model structure



Two possible ways to constrain P and E, consistent with column water budget

Predict E, diagnose P

$$E = \mathrm{NN}_{\mathrm{evap}}(X),$$

$$P = \frac{1}{g}\int_0^1 \sum_{\mathrm{i}} \left(\frac{dq}{dt}\right)_{\mathrm{i}}^{\mathrm{NN}_{\mathrm{tend}}} p_{\mathrm{s}} d\sigma + \mathrm{NN}_{\mathrm{evap}}(X).$$

Pro: $\mathrm{NN}_{\mathrm{evap}}$ only requires surface values

Predict P, diagnose E

(this is what they did for the main text)

$$P = \mathrm{NN}_{\mathrm{precip}}(X)$$

$$E = \mathrm{NN}_{\mathrm{precip}}(X) - \frac{1}{g}\int_0^1 \sum_{\mathrm{i}} \left(\frac{dq}{dt}\right)_{\mathrm{i}}^{\mathrm{NN}_{\mathrm{tend}}} p_{\mathrm{s}} d\sigma.$$

Pro: $\mathrm{NN}_{\mathrm{precip}}$ can enforce non-negative P

For training:

- Evaporation from ERA5
- Precipitation from IMERG
- Relax the loss weight of specific humidity (to accommodate ERA5-IMERG inconsistency)
- P and E do not feed back into NeuralGCM (i.e., are not diagnostic), but are optimized together

Neural network to predict precipitation

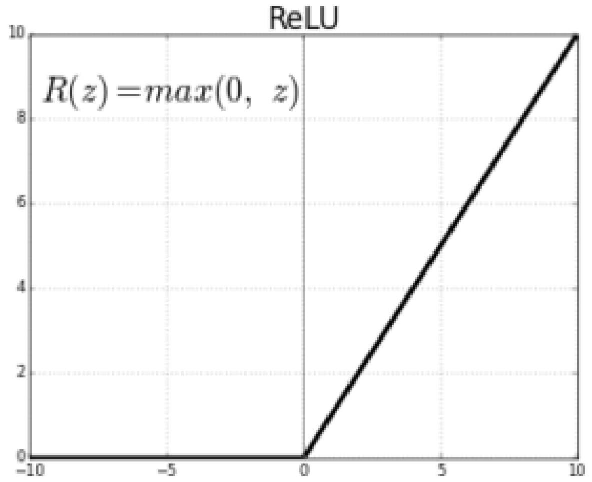
Inputs

- Zonal wind
- Meridional wind
- Temperature
- Specific humidity
- Specific cloud ice
- Specific liquid water
- Orography
- Land-sea mask
- Location embedding
- Surface embedding

$\mathbf{NN_{precip}}$

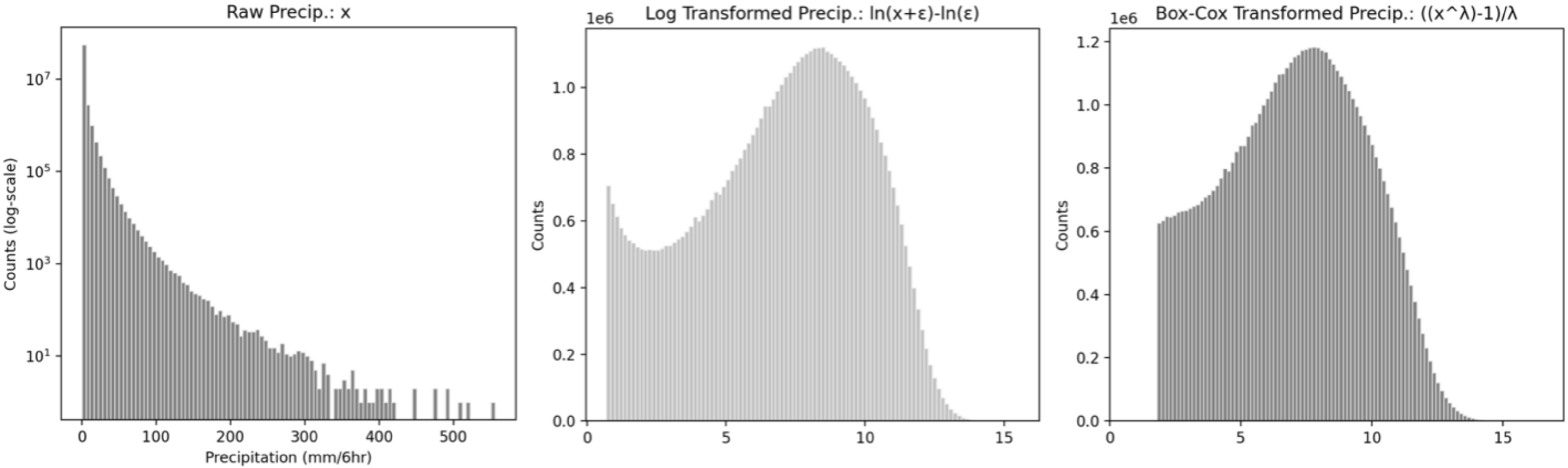
- Fully-connected NN with residual connections
- Six layers with hidden dimension = 64
- Encode–Process–Decode

Enforce non-negativity



Hourly precipitation rate (mm/hour)

P is heavily skewed



From Raul’s master’s thesis

Transformation of P encourages model to

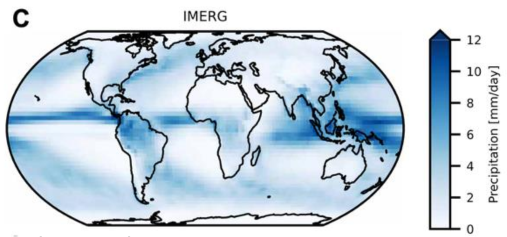
- Predict non-zero values
- Predict extreme values

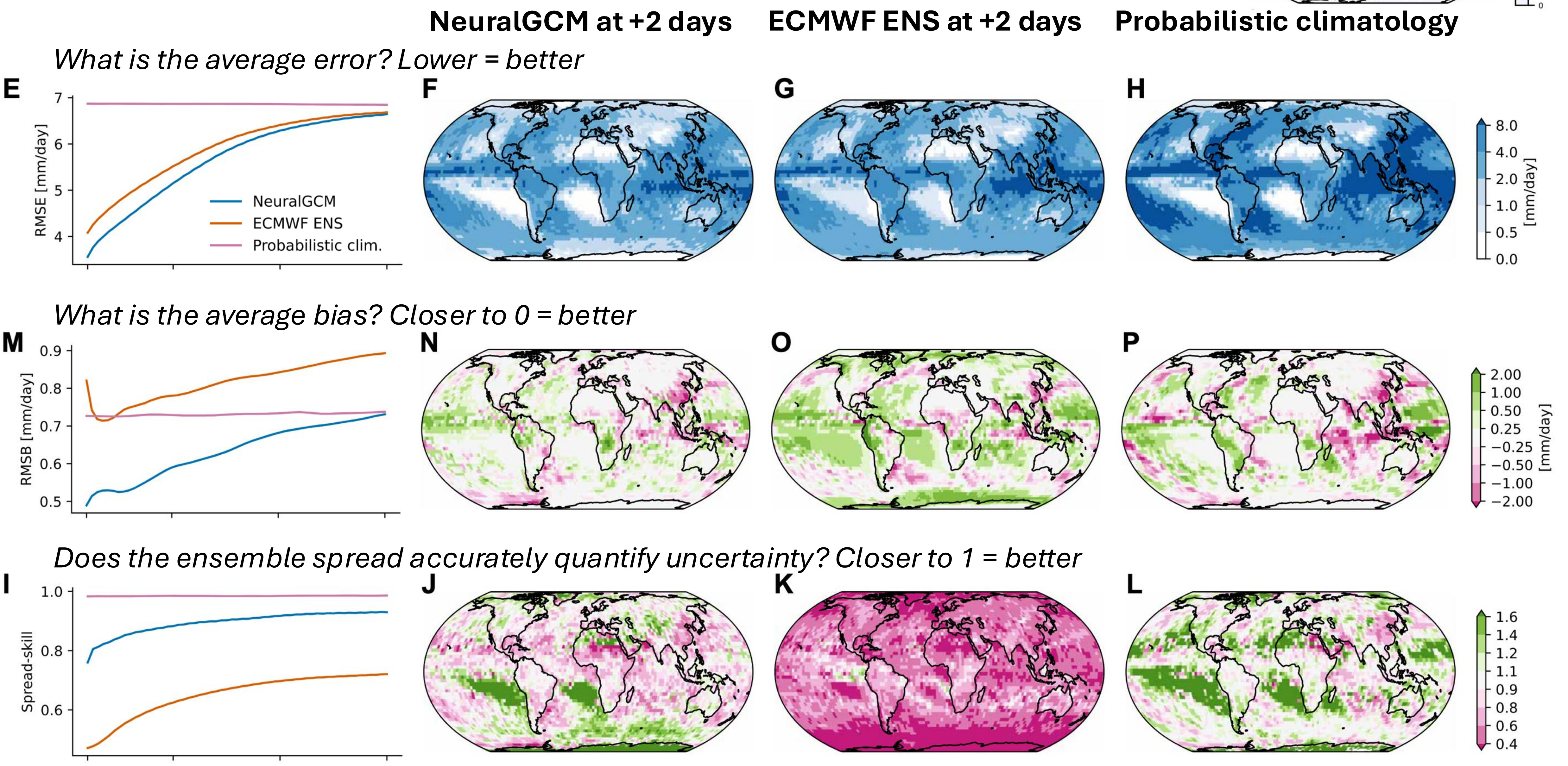
Discussion Q: Why can NeuralGCM predict P directly, while other models (DLESyM) need to transform P?

Results, finally

1. Medium-range forecasting performance (Fig. 2)
2. Climatological performance (Fig. 3-6)

Medium-range forecasting performance (against IMERG)



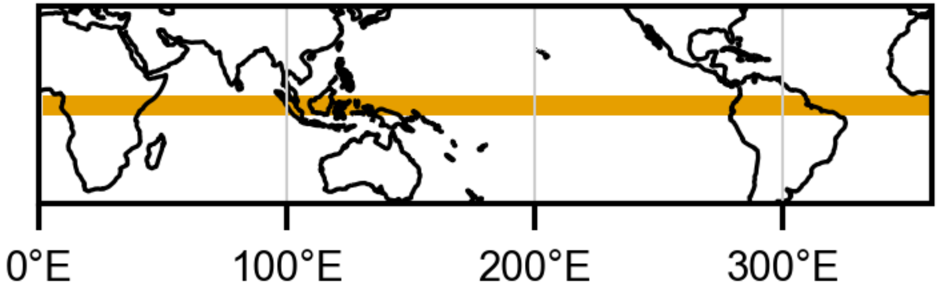


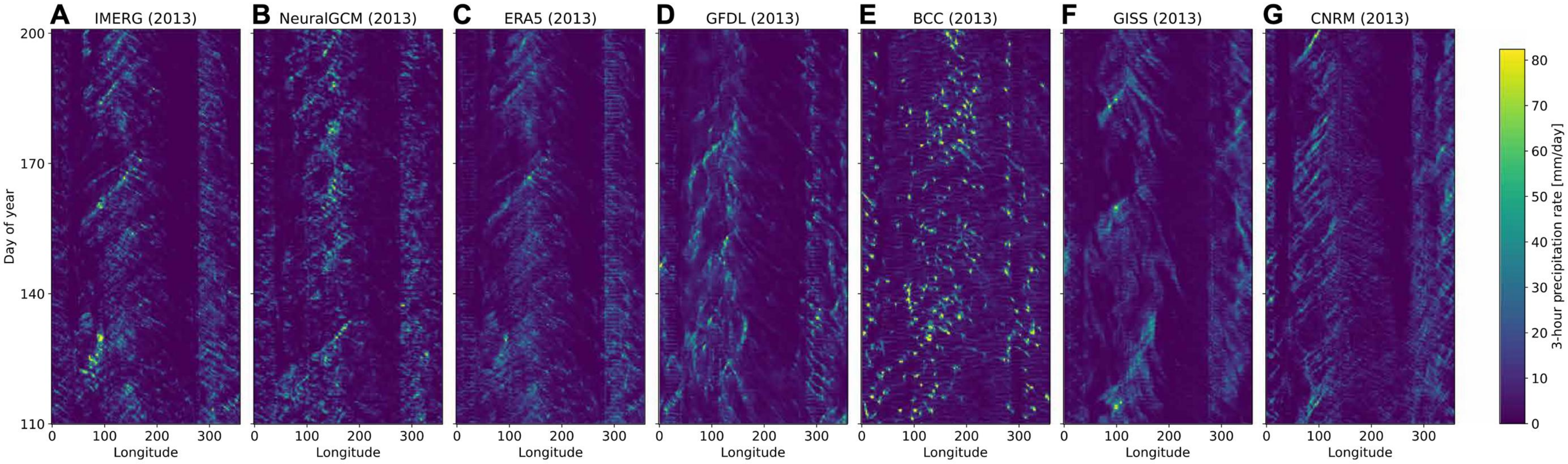
Climatological performance

Comparison of climatological precipitation over 2002-2014:

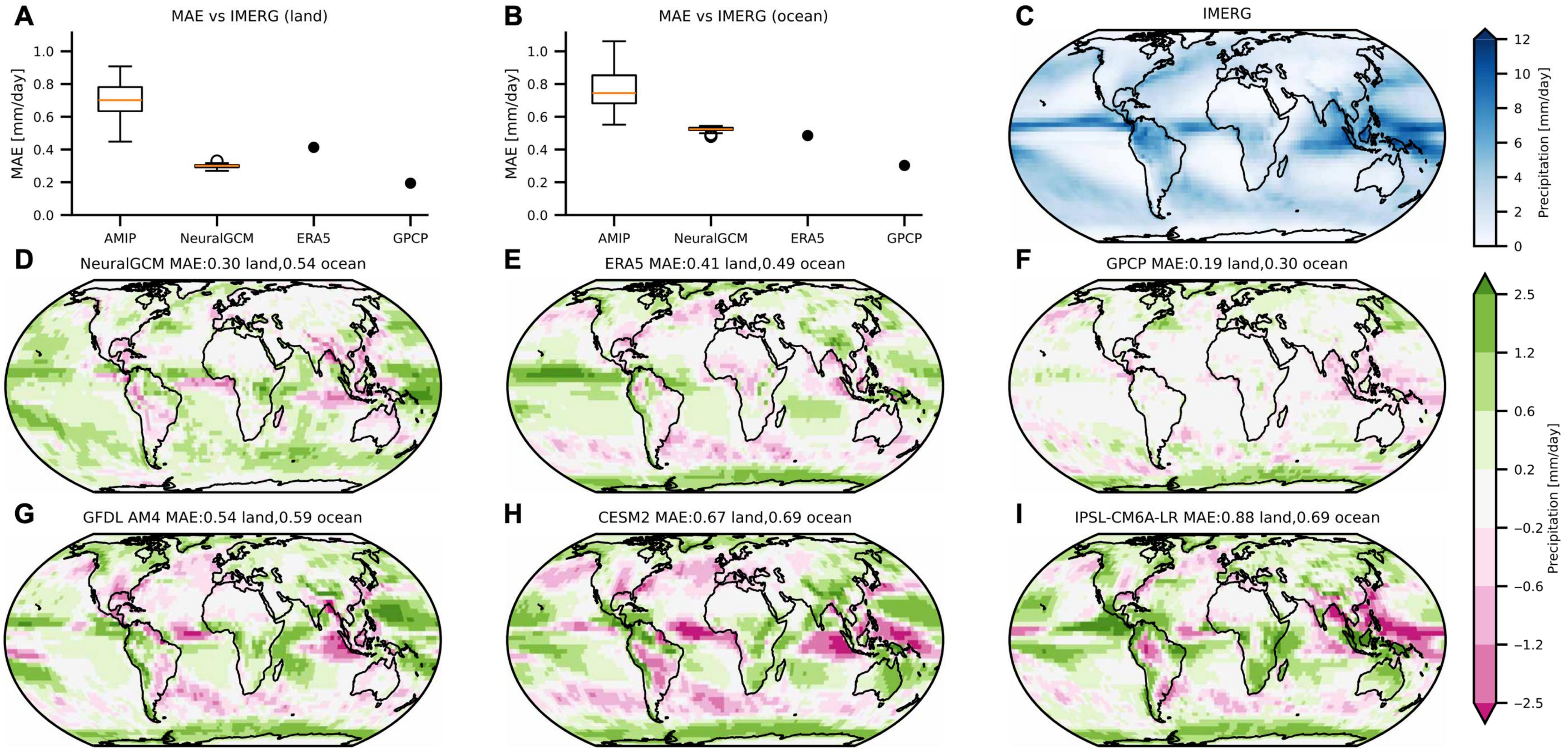
- NeuralGCM (37 ensemble members)
- ERA5 reanalysis
- GFDL X-SHiELD global cloud-resolving model
- CMIP6 AMIP (atmosphere-only) models (22 ensemble members/models)
- CMIP6 coupled models (8 ensemble members/models)

Qualitative comparison of tropical precipitation

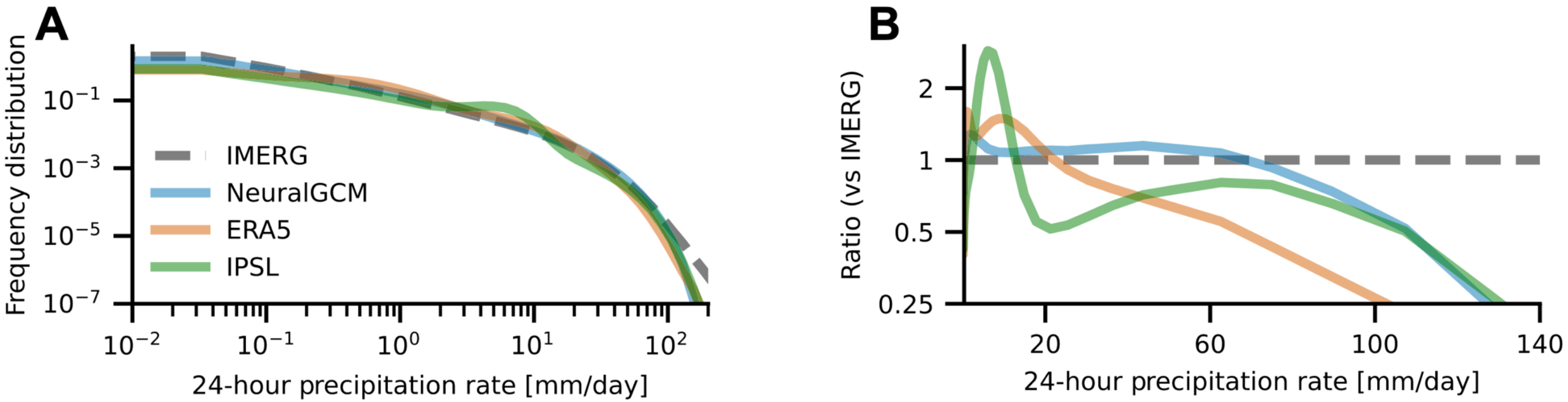




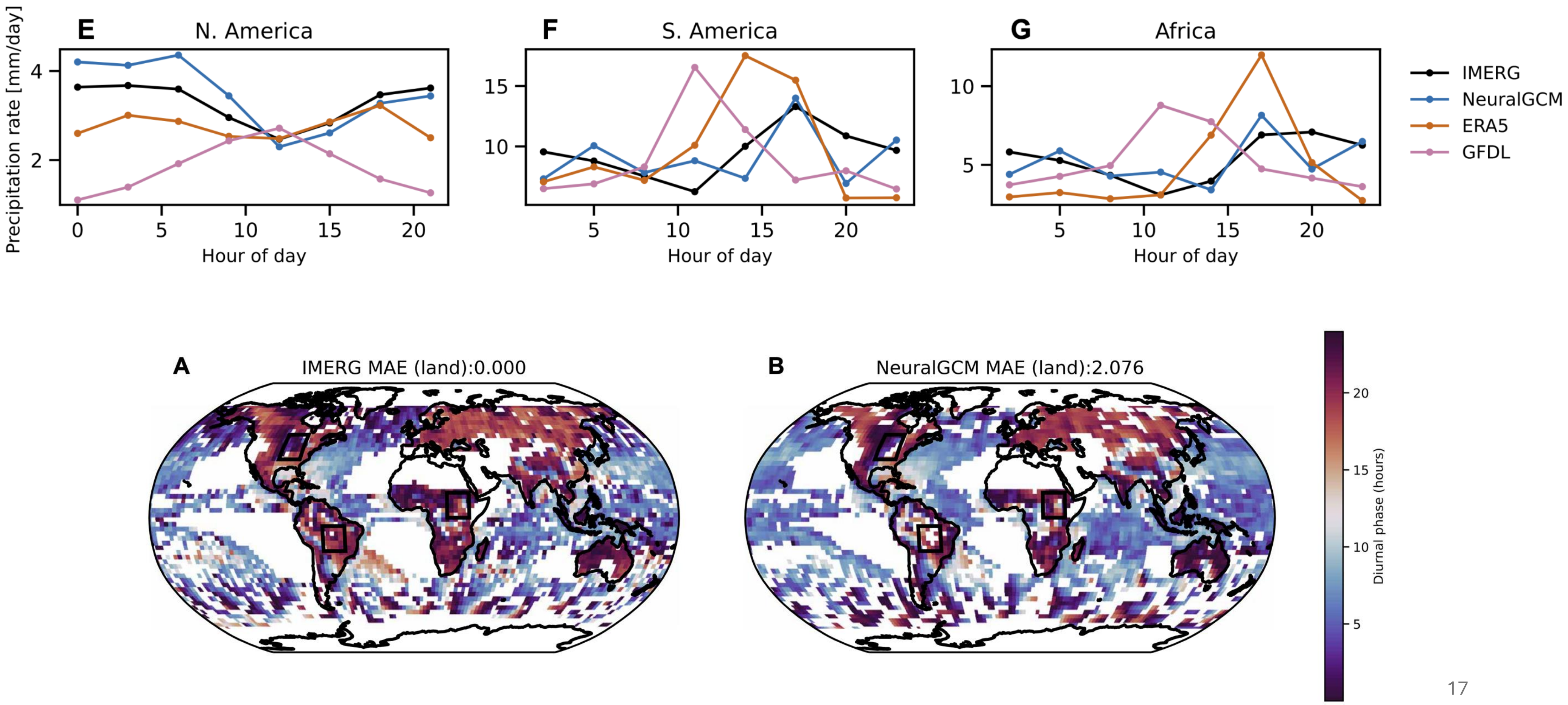
Bias in mean precipitation



Tropical precipitation distribution

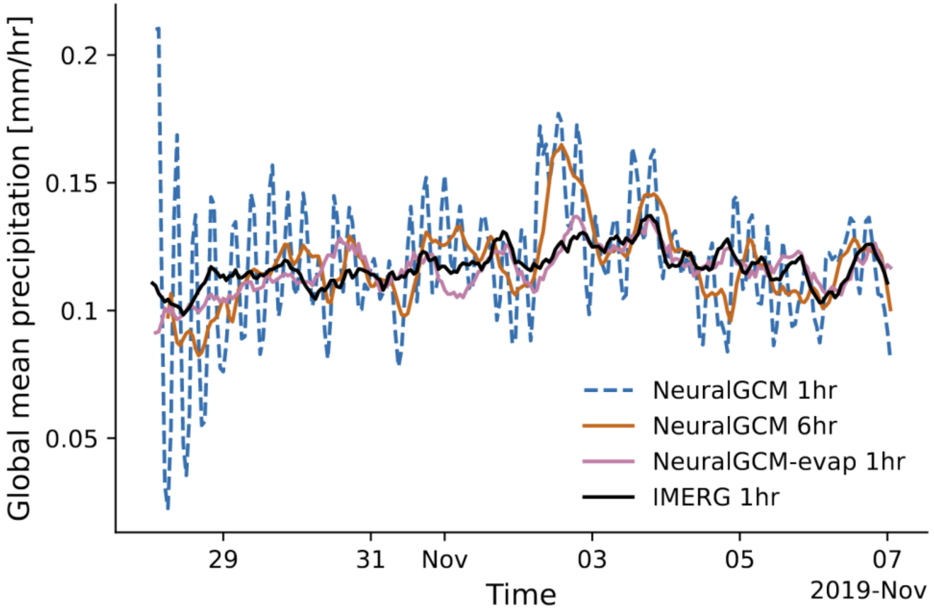


Diurnal cycle of summertime precipitation



Remaining issues with NeuralGCM

- “Sub–6-hour precipitation accumulations in NeuralGCM (but not NeuralGCM-evap) also show unrealistic oscillations in intensity, particularly during the first day of forecasting”
  - Time steps are 20 min for dynamical core and 60 min for physics NN



- “models trained with different random seeds (resulting in different initial model parameters or weights) exhibited notable variations in stability”
  - “One major limitation of our modeling framework is that obtaining models that are reliably stable over long rollouts requires training ~50 to 100 models”

Shoutout to JCM (JAX Circulation Model)

- Yuval et al.: “Although this study used an NN to parameterize all processes unresolved by the dynamical core, future work could explore coupling our differentiable dynamical core with a traditional parameterization suite and optimizing its free parameters”
- Preprint released on Monday: “JCM v1.0: A Differentiable, Intermediate-Complexity Atmospheric Mode”
  - Uses same dynamical core as NeuralGCM, but SPEEDY physics (simplified parameterization suite)
  - Built on JAX (same as NeuralGCM) → fully differentiable

https://doi.org/10.5194/egusphere-2025-6266
Preprint. Discussion started: 26 January 2026

EGUsphere
Preprint repository

JCM v1.0: A Differentiable, Intermediate-Complexity Atmospheric Model

Ellen H. Davenport[1,*], J. Varan Madan[1,*], Rebecca Gjini[1, 2], Jared Brzenski[1], Nick Ho[1], Tien-Yiao Hsu[1], Yueshan Liang[1], Zhixing Liu[1], Veeramakali Manivannan[1], Eric Pham[1], Rohith Vutukuru[1], Andrew I. L. Williams[1], Zhiqi Yang[1], Rose Yu[3], Nicholas J. Lutsko[1], Stephan Hoyer[5], and Duncan Watson-Parris[1,4]

Fig. 2. Precipitation forecasting accuracy scores for 24-hour accumulated precipitation, evaluated against IMERG

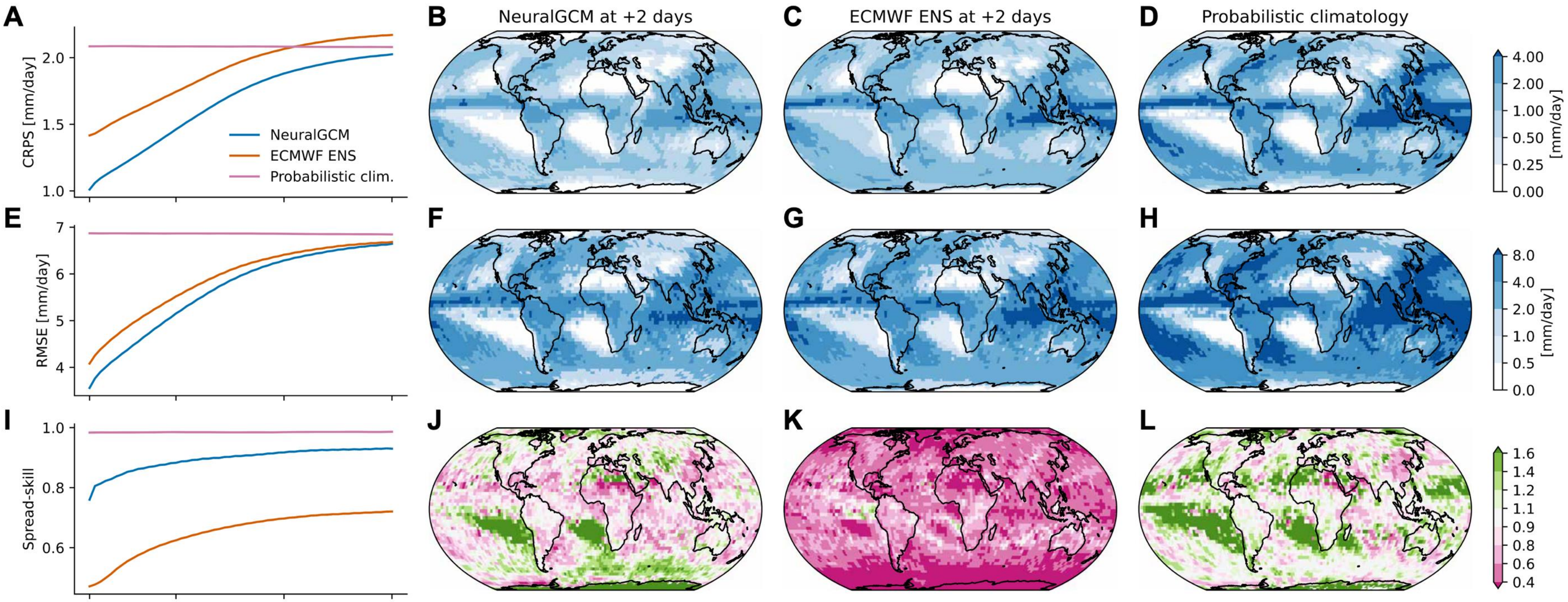


Fig. 2. Precipitation forecasting accuracy scores for 24-hour accumulated precipitation, evaluated against IMERG

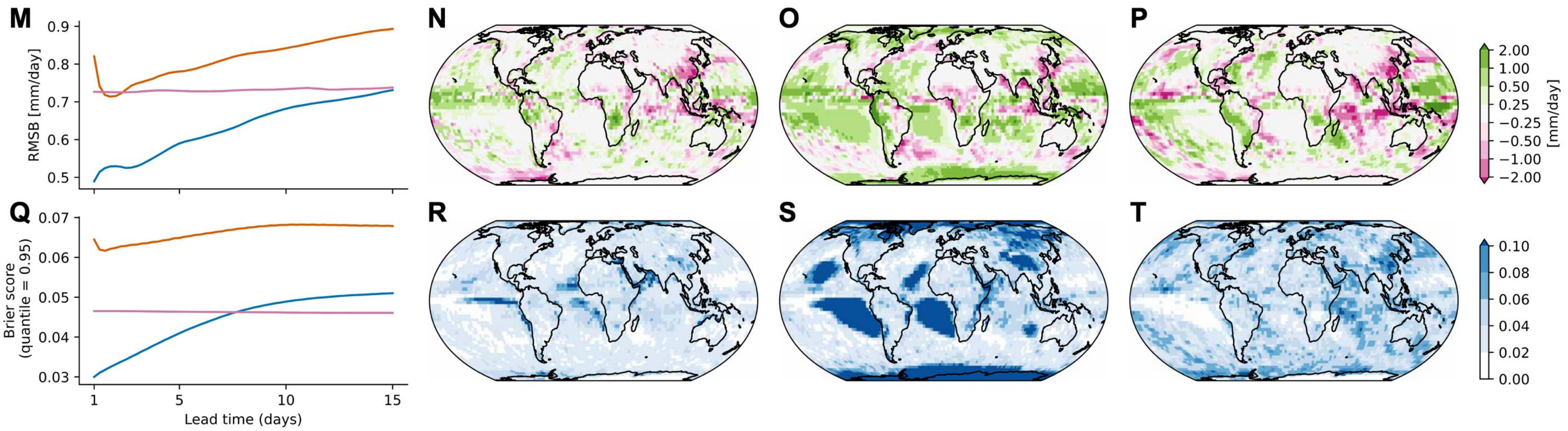


Fig. 3. Hovmöller tropical precipitation diagram for different models

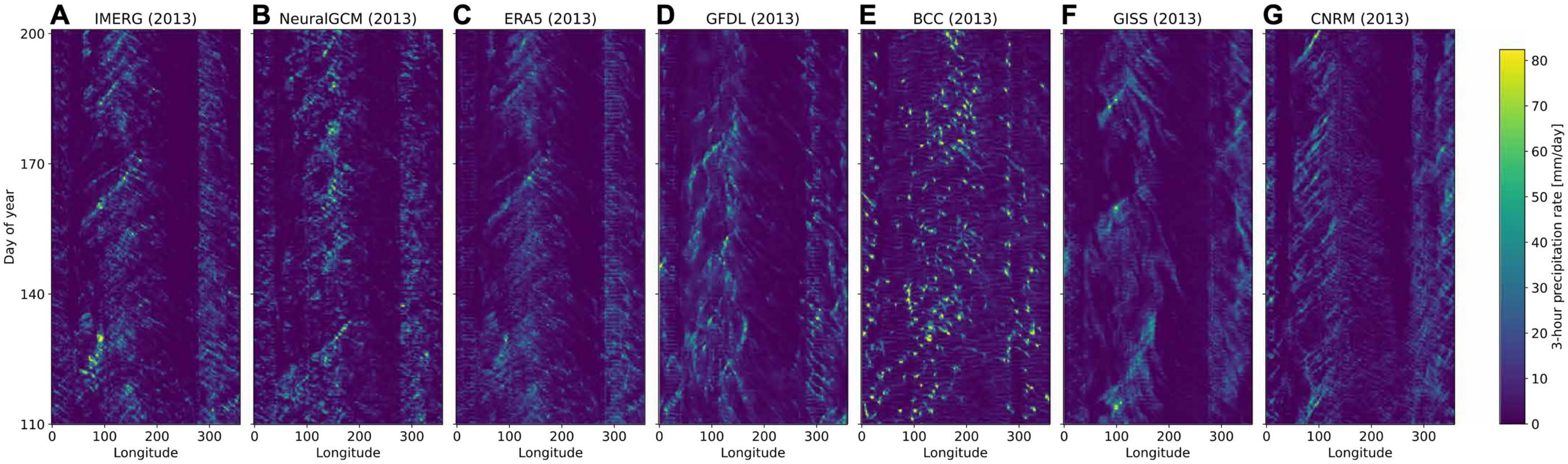


Fig. 4. Bias in mean precipitation averaged over 2002 to 2014

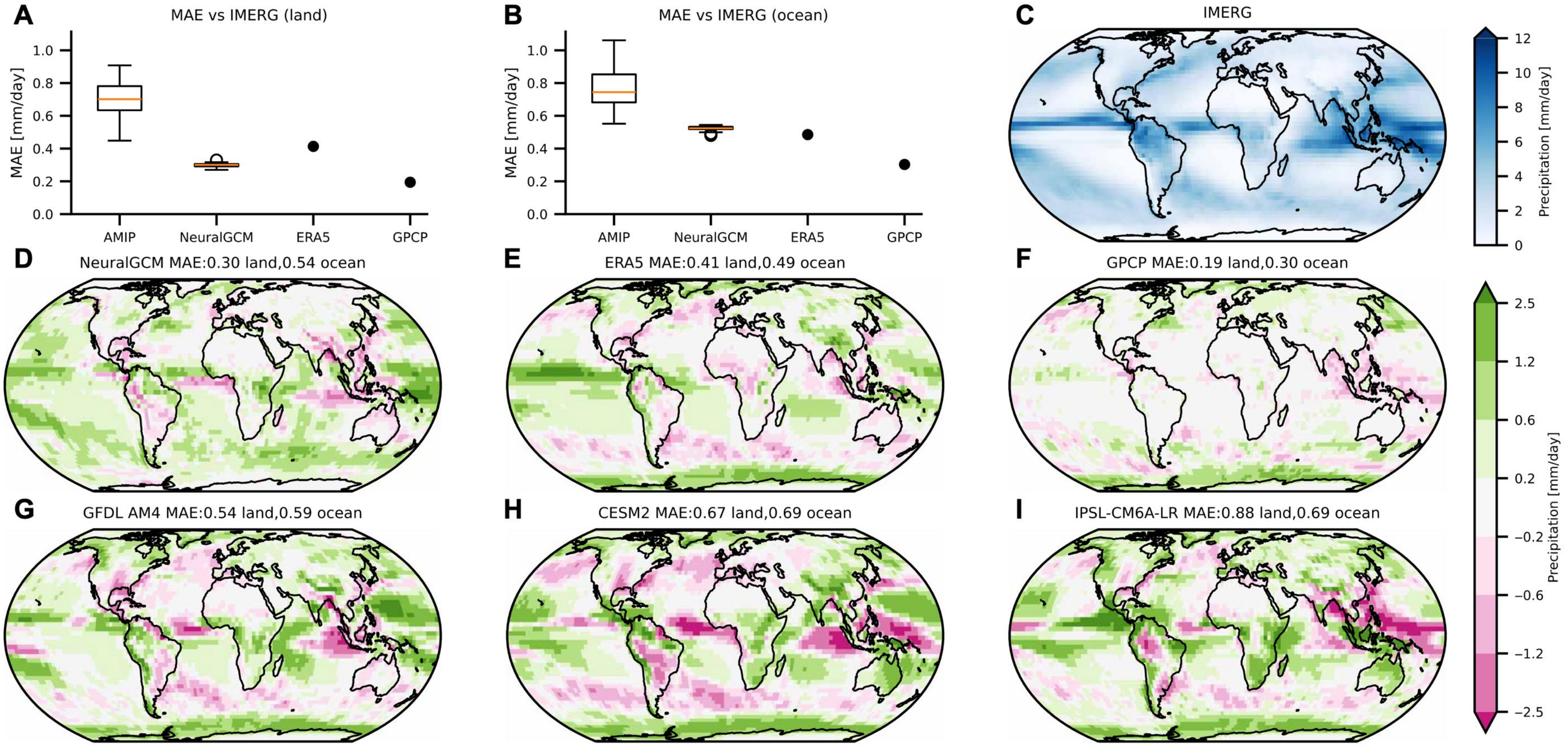


Fig. 5. Tropical precipitation rate distribution and annual maximum daily precipitation averaged over 2002 to 2014

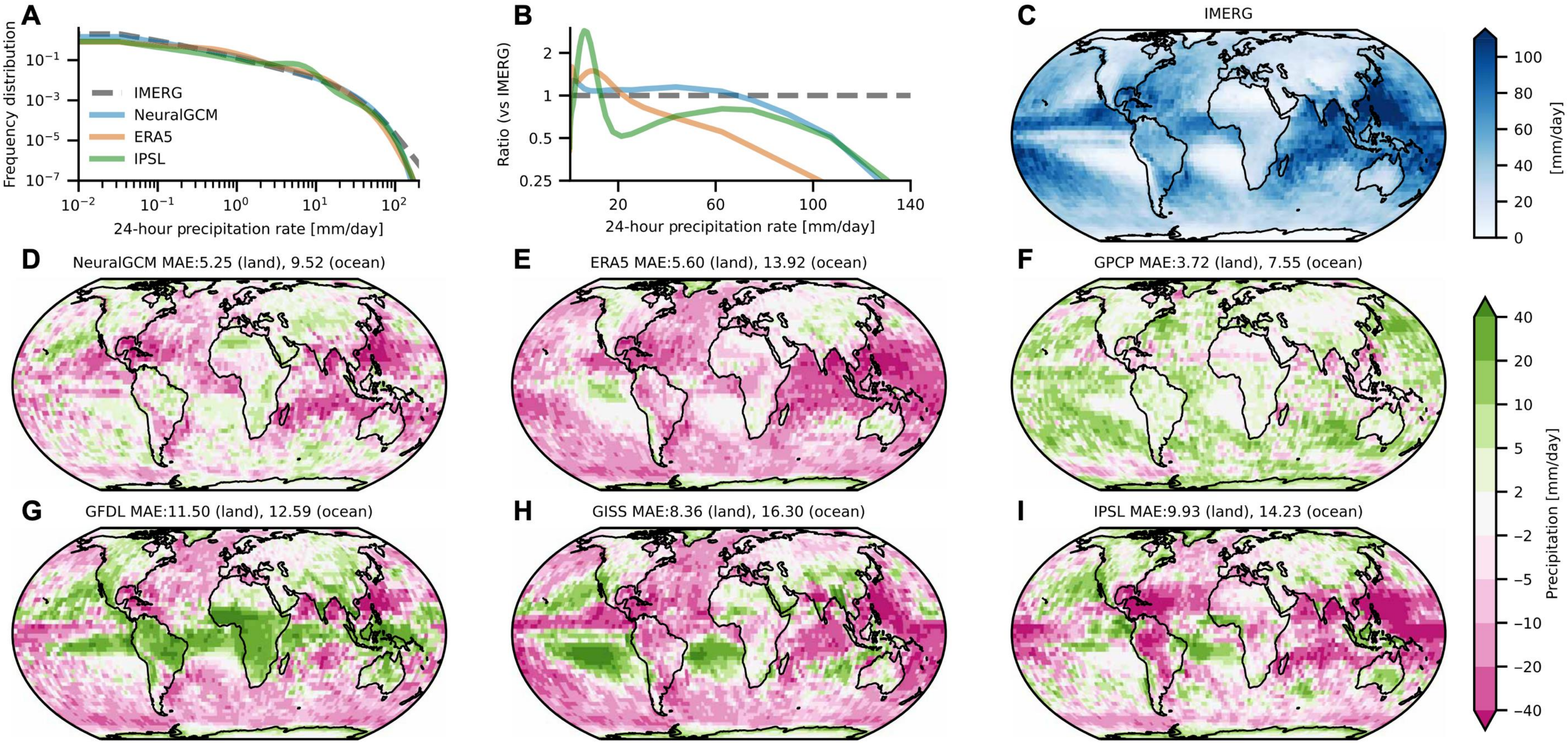


Fig. 6. Diurnal cycle of summertime precipitation (2002 to 2014)