

Towards machine learning-based Earth system models

Christian Lessig, European Centre for Medium-Range Weather Forecasts
christian.lessig@ecmwf.int



Machine learning-based Earth system models?

ERA5

Weather forecasting

Machine learning-based Earth system models?

ERA5

Weather forecasting

Post-processing

Climate projections

Scenario generation

CO₂ monitoring

• • •

Machine learning-based Earth system models?

ERA5

Weather forecasting

Post-processing

Climate projections

Scenario generation

CO₂ monitoring

• • •

Machine learning-based Earth system models?

ERA5
CERRA
COSMO-REA6
...

Weather forecasting

Post-processing

Climate projections

Scenario generation

CO₂ monitoring

...

Machine learning-based Earth system models?

ERA5
CERRA
COSMO-REA6
...
IFS
CMIP6
ICON

Weather forecasting

Post-processing

Climate projections

Scenario generation

CO₂ monitoring

...

Machine learning-based Earth system models?

ERA5
CERRA
...
IFS
CMIP6
ICON
...
AMSU-A, METOP, ...
Landsat, ...
...
Local GDP, tax revenue, ...

Weather forecasting

Post-processing

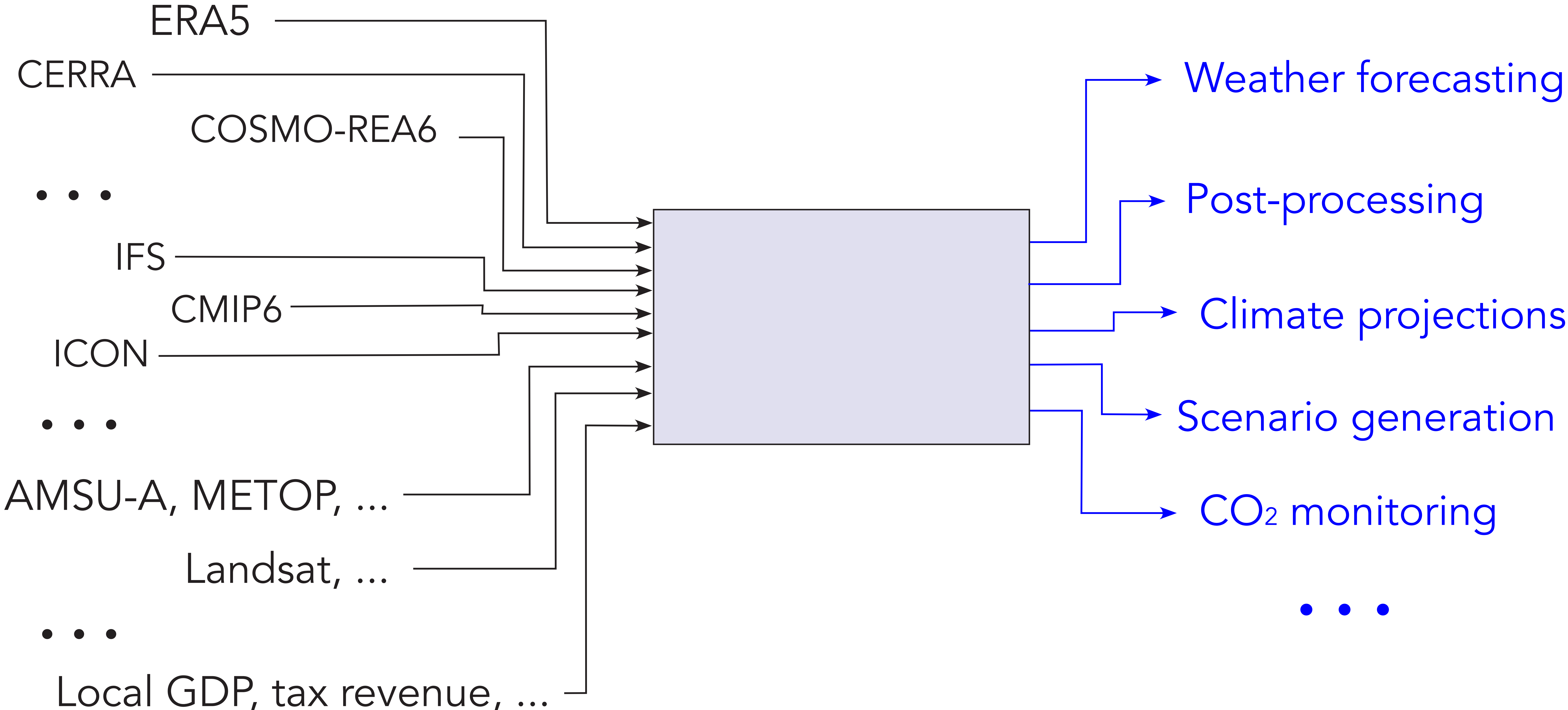
Climate projections

Scenario generation

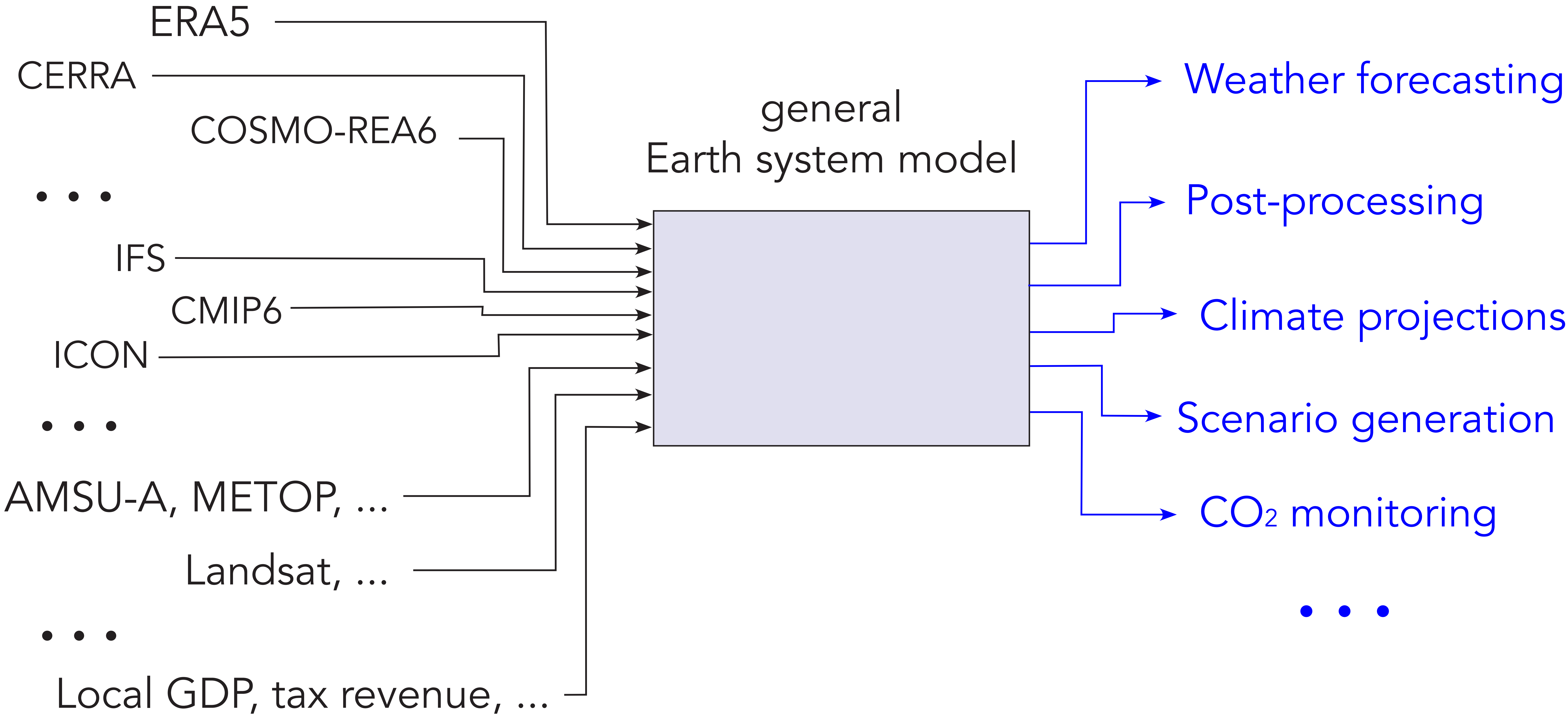
CO₂ monitoring

...

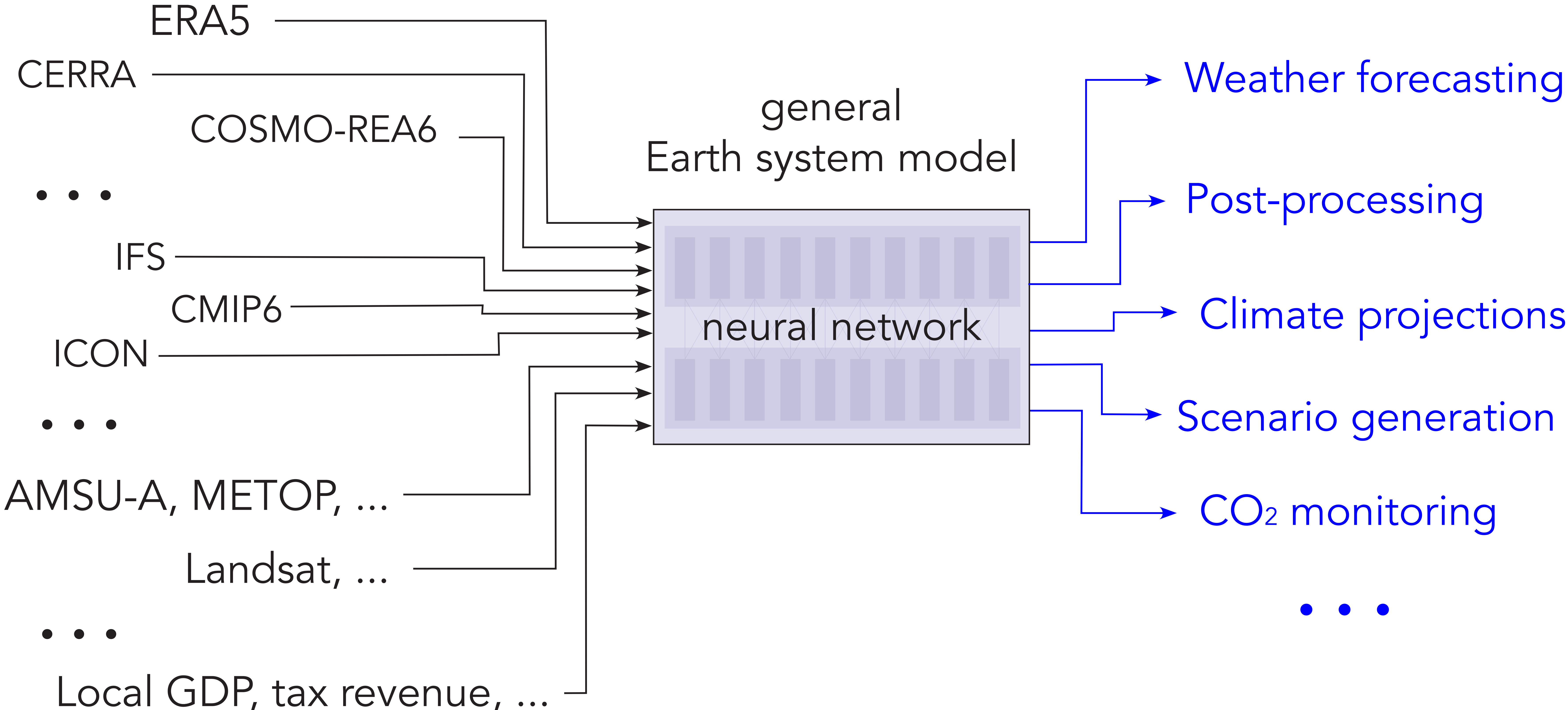
Machine learning-based Earth system models?



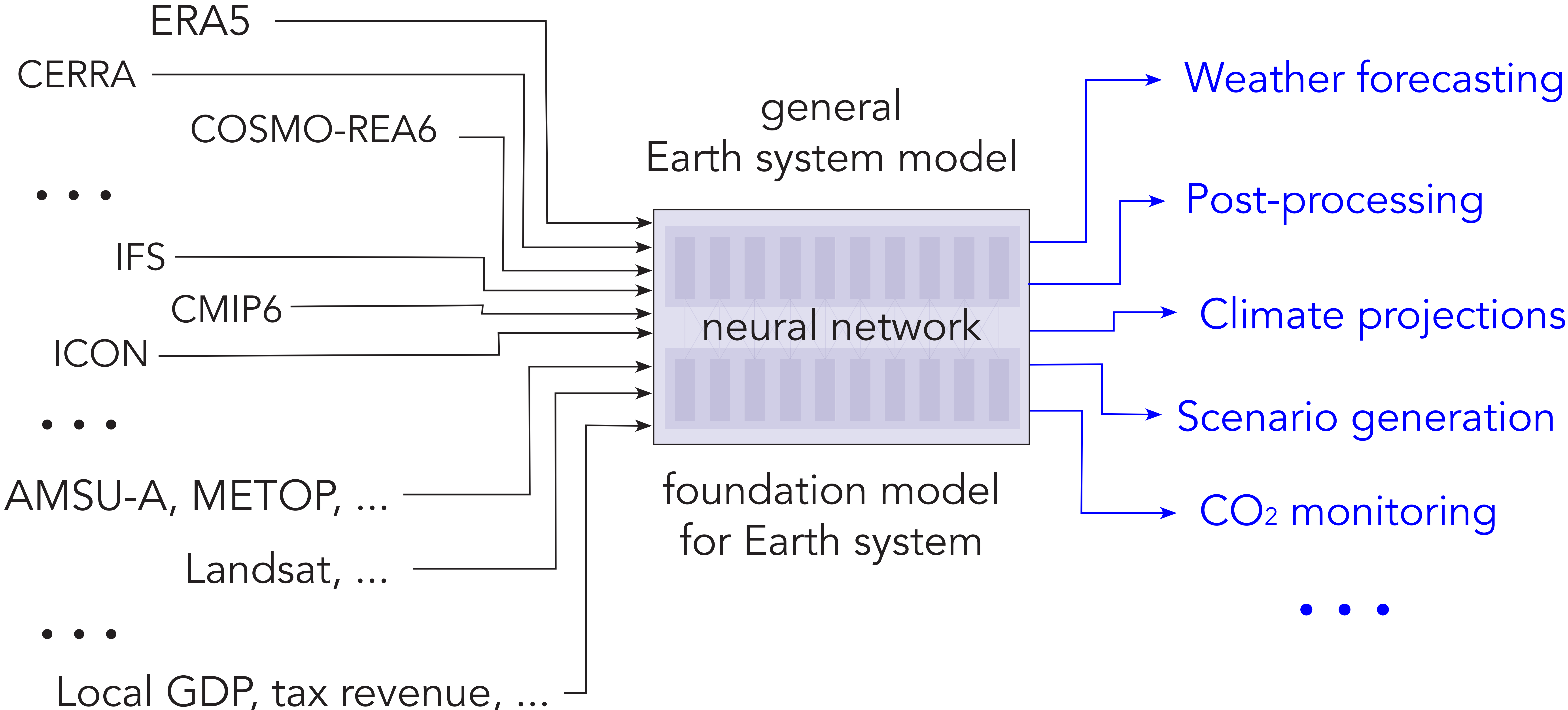
Machine learning-based Earth system models?



Machine learning-based Earth system models?

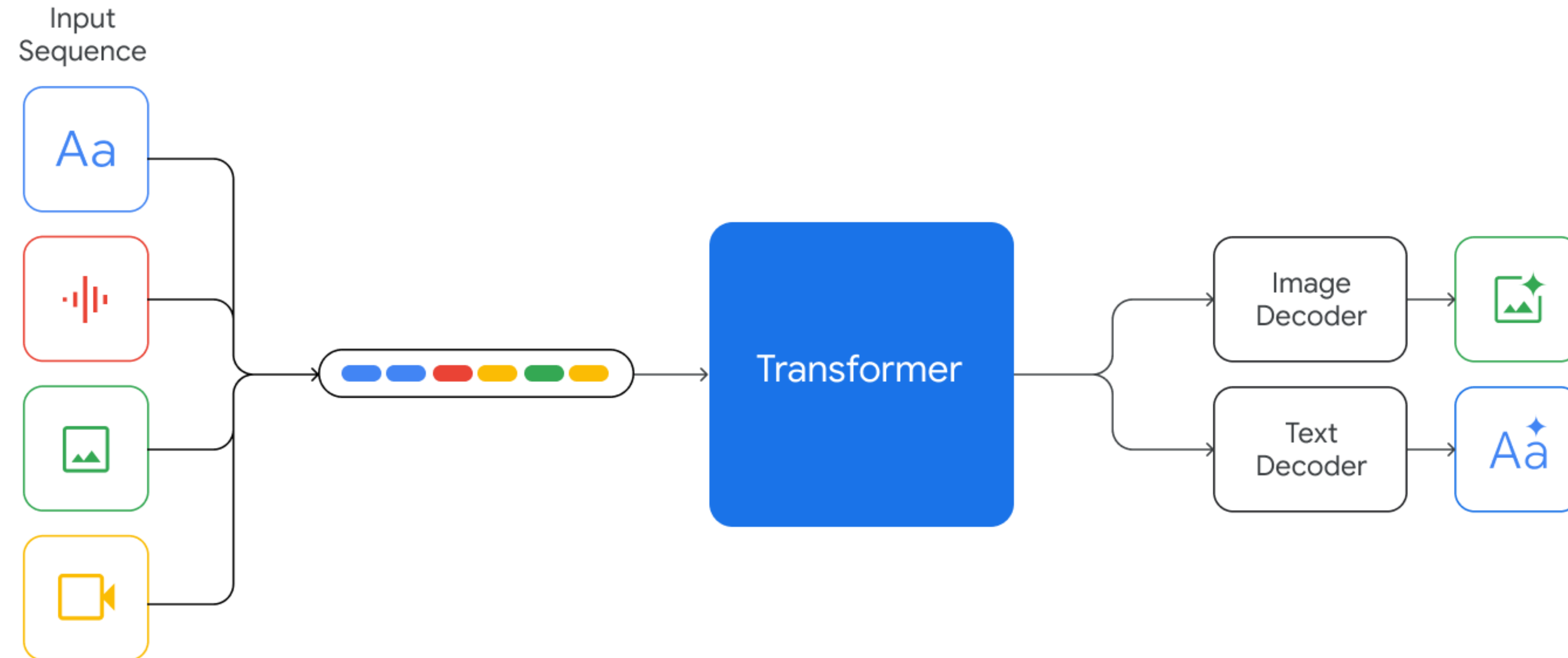


Machine learning-based Earth system models?



An example

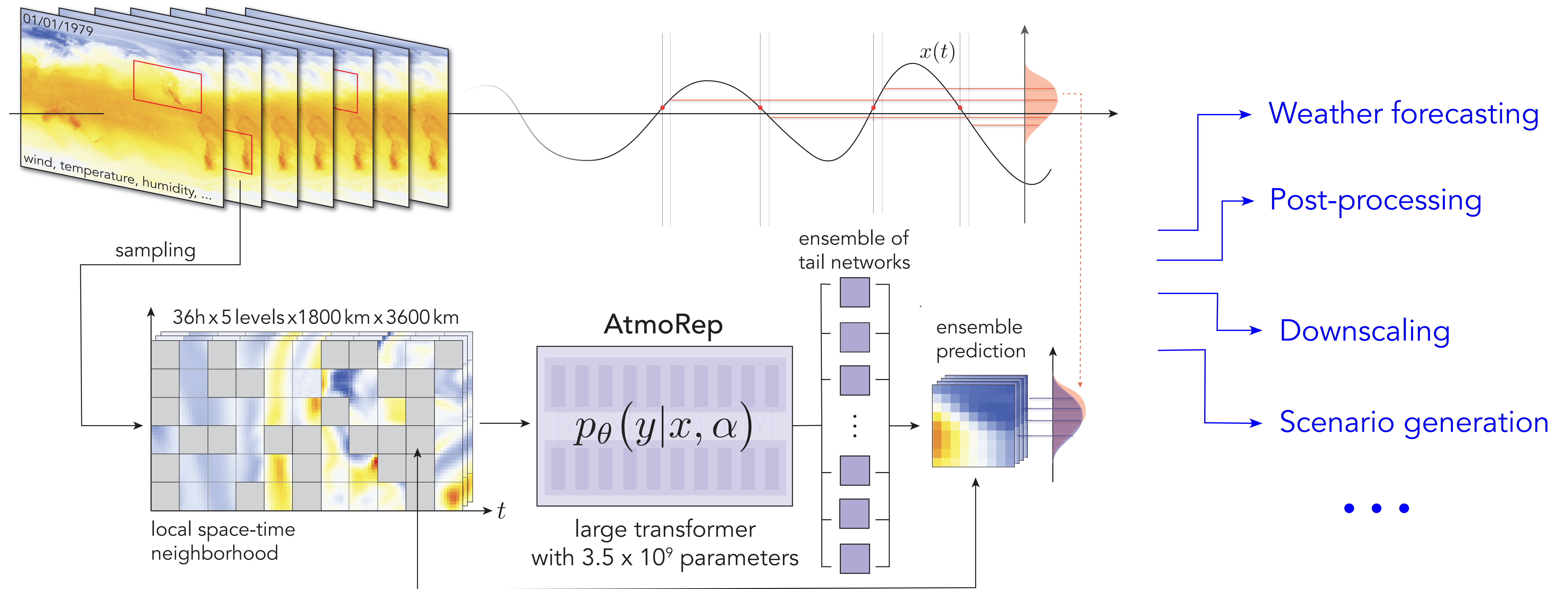
- Google's Gemini model:¹



¹ Google Team. Gemini: A family of highly capable multimodal models, 2023.

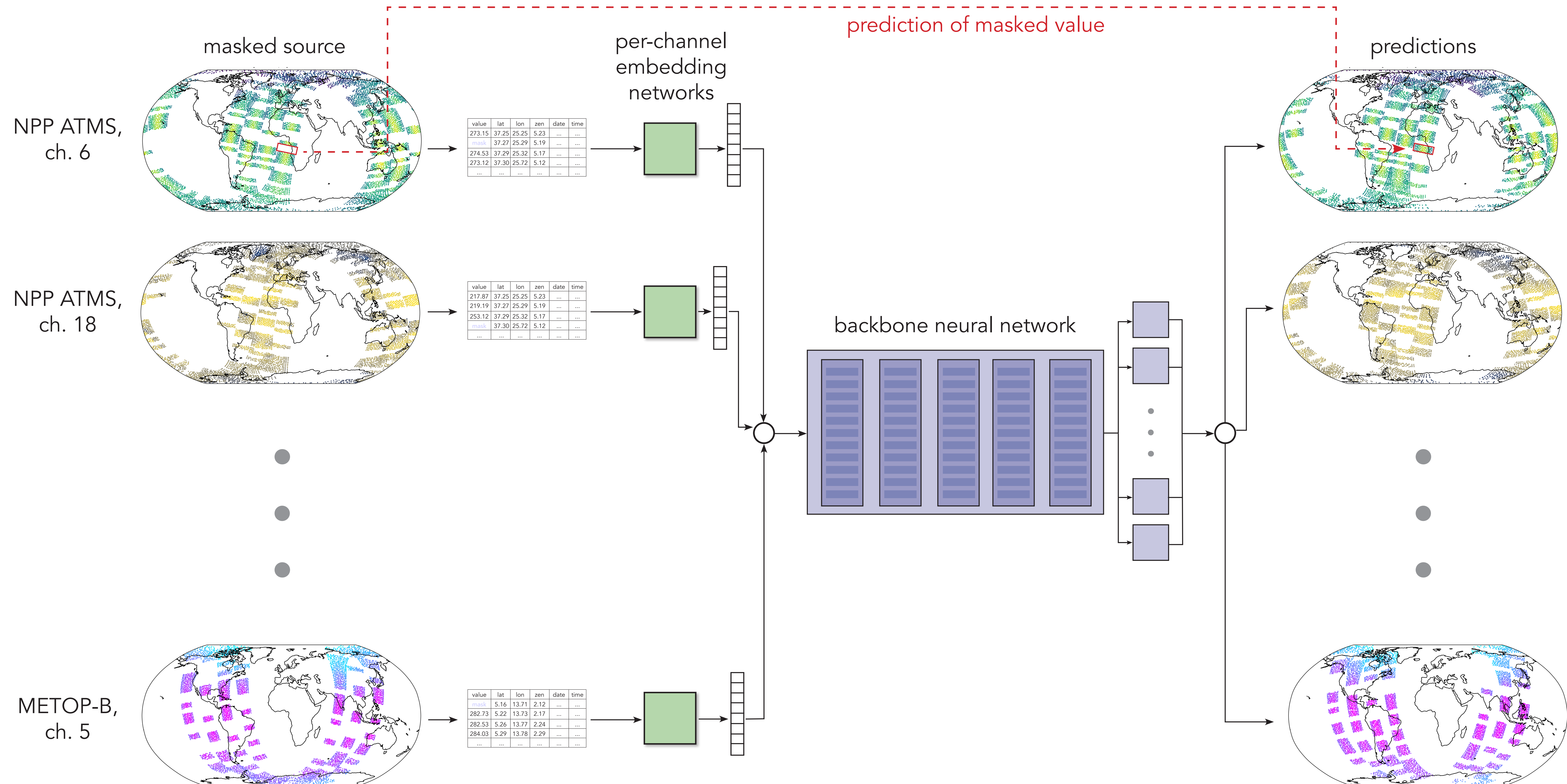
First steps: AtmoRep

pre-processed historical observational record $x(t)$ (ERA5 reanalysis)



² C. Lessig, I. Luise, B. Gong, M. Langguth, S. Stadler, and M. Schultz. Atmorep: A stochastic model of atmosphere dynamics using large scale representation learning, 2023; <https://arxiv.org/abs/2308.13280>

First steps: Learning from observations



Summary

- Machine learning-based Earth system model is plausible
 - › Extension of multimodal model (with many challenges)
- Integration of many different data streams in the network
 - › Correlations between them learned from the data
 - › Include Earth observations might allow to obtain models with better skill than conventional ones
- Would (also) allow bottom-up and top-down approaches for CO₂ monitoring