Quantifying the Scientific Skill of High-Resolution Models with *pyhanami*

Marta Alerany Solé[1] **(marta.alerany@bsc.es)**, Kai Keller[1], Chihiro Kodama[2], Masuo Nakano[2], Tomoe Nasuno[2], Daisuke Takasuka[3], Mario Acosta[1]

1: Barcelona Supercomputing Center, Spain; 2: Japan Agency for Marine-Earth Science and Technology, Japan; 3: Graduate school of science, Tohoku University, Japan

The Package: *pyhanami*

An open-source Python package for evaluating the **scientific skill** and **replicability** of Earth System Models (ESMs) through visualization tools, objective scalar-based metrics, and statistical testing.

pyhanami

Check our repo:

Easy input handling

SimulationData class

1. Load climate simulation ensembles from a **NetCDF file** or an **xarray.Dataset** object.
2. Automatically verify the integrity and consistency of the input datasets.

Intuitive syntax:

```
import pyhanami
sim_data = pyhanami.SimulationData('sim_data_path', name='sim name')
```

Model evaluation

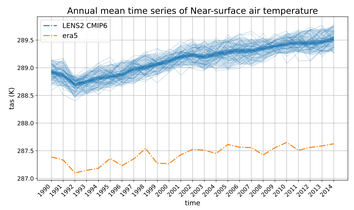
DataDiagnostics class

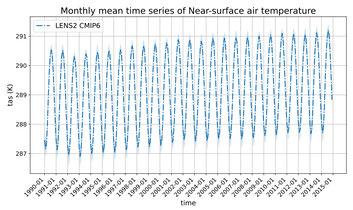
Diagnostics of **ensemble** simulations for a given variable:

1. **Time series**: multiple simulations + observations
2. **Spatial plots**: absolute difference + effect size (Cohen's d)

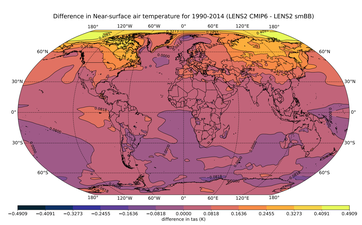
Usage examples employing CESM2 LENS2 data:

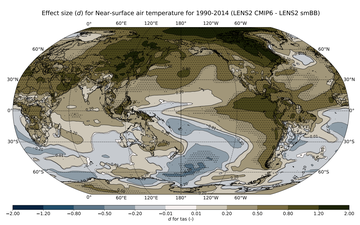
```
diags = pyhanami.DataDiagnostics(lens2_cmip6)
diags.time_series_plot('tas', start_year=1990, end_year=2014, obs=True, plot_ens=True)
diags.time_series_plot('tas', start_year=1990, end_year=2014, time_freq='monthly')
```






```
diags.add_dataset(lens2_smbb)
diags.abs_diff_plot('tas', start_year=1990, end_year=2014)
diags.eff_size_plot('tas', start_year=1990, end_year=2014, clon=180)
```





ScientificEvaluation class

Plots and **scalar scores** for evaluating scientific model skill. Examples using IFS-NEMO simulations at different resolutions:

```
sciskill = pyhanami.ScientificEvaluation([ifs_nemo_5km, ifs_nemo_10km, ifs_nemo_25km])
x_analysis = sciskill.compute_x_scores('IFS-NEMO 10km', start_year=1990, end_year=2014)
```

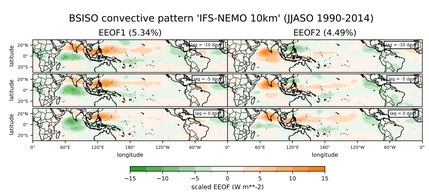
1. Tropical Intraseasonal Oscillation (ISO)

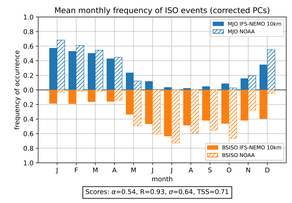
Predominant phenomenon in the tropics throughout the year:

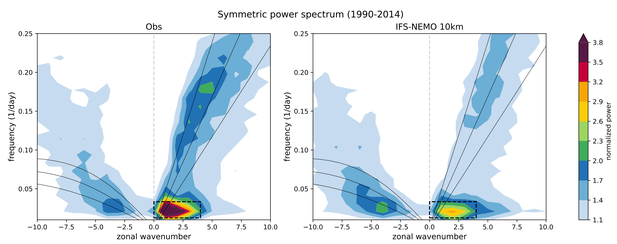
i. In boreal winter: **Madden-Julian Oscillation (MJO)**

ii. In boreal summer: **Boreal Summer ISO (BSISO)**

Evaluation: Empirical Orthogonal Function (EOF) analyses + Wavenumber-frequency power spectra







2. Tropical Cyclones (TCs)

Warm-core, rapidly rotating storm systems characterized by strong winds and heavy precipitation that begin over the tropical oceans. Evaluation:

Track TC trajectories ⇒ Compute TC metrics (duration, intensity, latitude) ⇒ Compute TC scores (bias, correlations)

Examples of usage employing several reanalyses and NICAM16-7S data (with $\sim$ 56 km horizontal resolution):

Global climatological mean bias (1990-2010)

5° x 5°	$\overline{b}_{clim,count}$ (#)	$\overline{b}_{clim,tcd}$ (days)	$\overline{b}_{clim,ace}$ (10^{-4} kn^2)	$\overline{b}_{clim,pace}$ (10^{-4} kn^2)	$\overline{b}_{clim,lmi}$ (° lat.)
IBTrACS	107.7	851.9	772.3	743.7	18.6
ERA5	-26.9	-329.0	-499.7	-347.1	5.0
JRA-55	-36.5	-377.3	-543.9	-458.8	4.9
NICAM16-7S	-32.3	-282.8	-408.1	-219.2	3.3

Negative bias — No bias — Positive bias

Extra and Future Features

- ✓ Spatial bias plots (simulations - observations)
- ✓ General scalar scores (bias, RMSE, correlation)
- ✓ Replicability test (following (K.Keller et al., 2025))
- ✓ Comprehensive documentation in ReadTheDocs
- ✍ Additional climate phenomena (e.g., precipitation)
- ✍ Intake data catalogue (standardized data interface)

Documentation:

This work received funding from the European High Performance Computing Joint Undertaking (EuroHPC JU) under the European Union's Horizon Europe framework program for research and innovation and Grant Agreement No. 101136269.

Barcelona Supercomputing Center — Centro Nacional de Supercomputación

HANAMI EU-Japan Alliance in HPC

Funded by the European Union