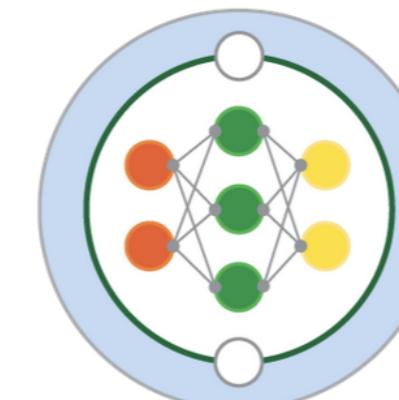


Online deployment of pre-trained machine learning components within geoscientific models through OASIS

Alexis Barge — Julien Le Sommer

22th July 2024



Motivation: interface for hybrid physics / AI modeling

Typical use cases of hybrid modeling

Parameterization from hi-fidelity models (LES, km-scale models)

[Sane et al. 2023](#) [Frezat et al. 2022](#)

Correction of model error from reanalysis or DA increments

[Gregory et al. 2024](#) [Chapman and Berner. 2023](#)

Acceleration of code components with neural emulators

[Hogan and Bozzo. 2018](#) [Chantry et al. 2021](#)

Needs for interfacing

Existing solutions

Implement or convert NN in Fortran

[neural-fortran](#) [Fortran-Keras Bridge](#) [FNN](#) [inference-engine](#)

Call Python scripts from Fortran with Python bindings

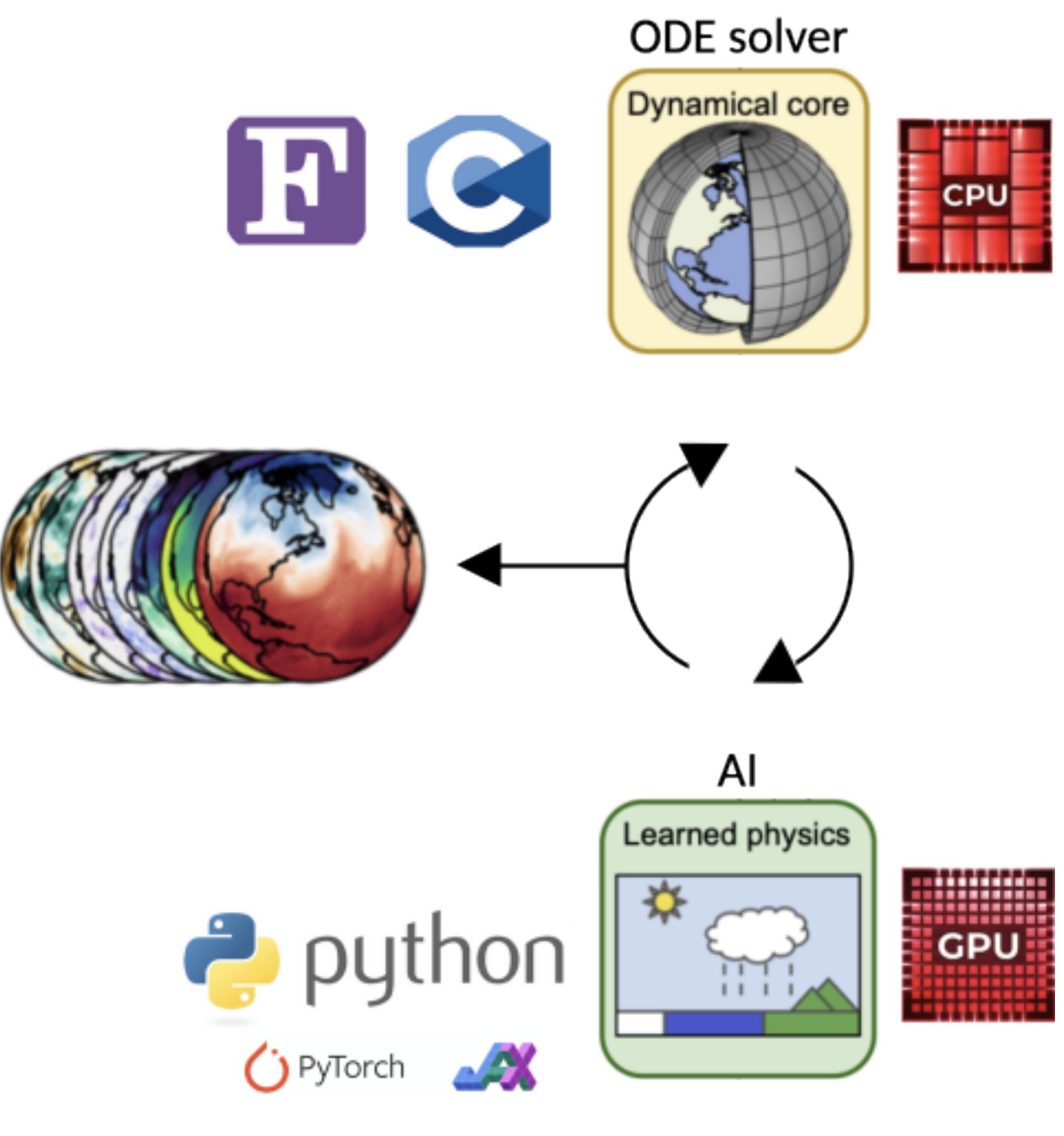
[call_py_fort](#) [Pyfort](#)

Leverage the existing C/C++ bindings of specific ML libraries

[infero](#) [Ftorch](#) [TF-lib](#) [TorchClim](#)

Leverage a more generic interface between Fortran and Python

[smartsim](#) [PhyDLL](#)



OASIS coupler for hybrid modeling

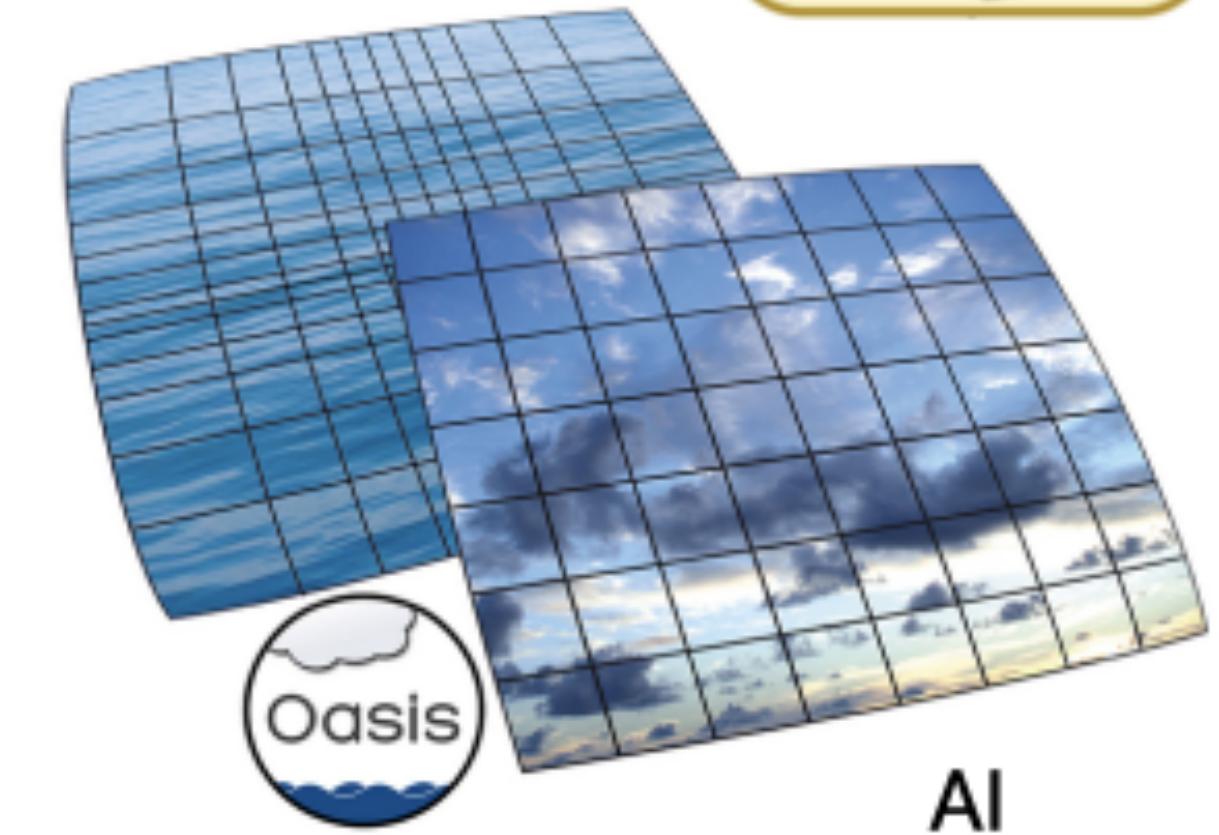
OASIS3-MCT <https://oasis.cerfacs.fr/en>

Coupling library between different codes
Interpolate and exchange 2D/3D fields



Widely deployed

Used by 67 climate modeling groups
Used in 5 of the 7 European ESMs used for CMIP6

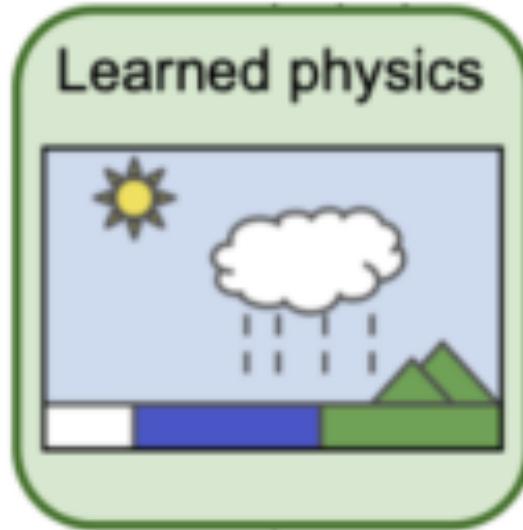


New functionalities since OASIS3-MCT_5.0

Python and C/C++ API



Can we leverage OASIS for hybrid modeling ?



OASIS coupler for hybrid modeling: drawing

OASIS3-MCT <https://oasis.cerfacs.fr/en>

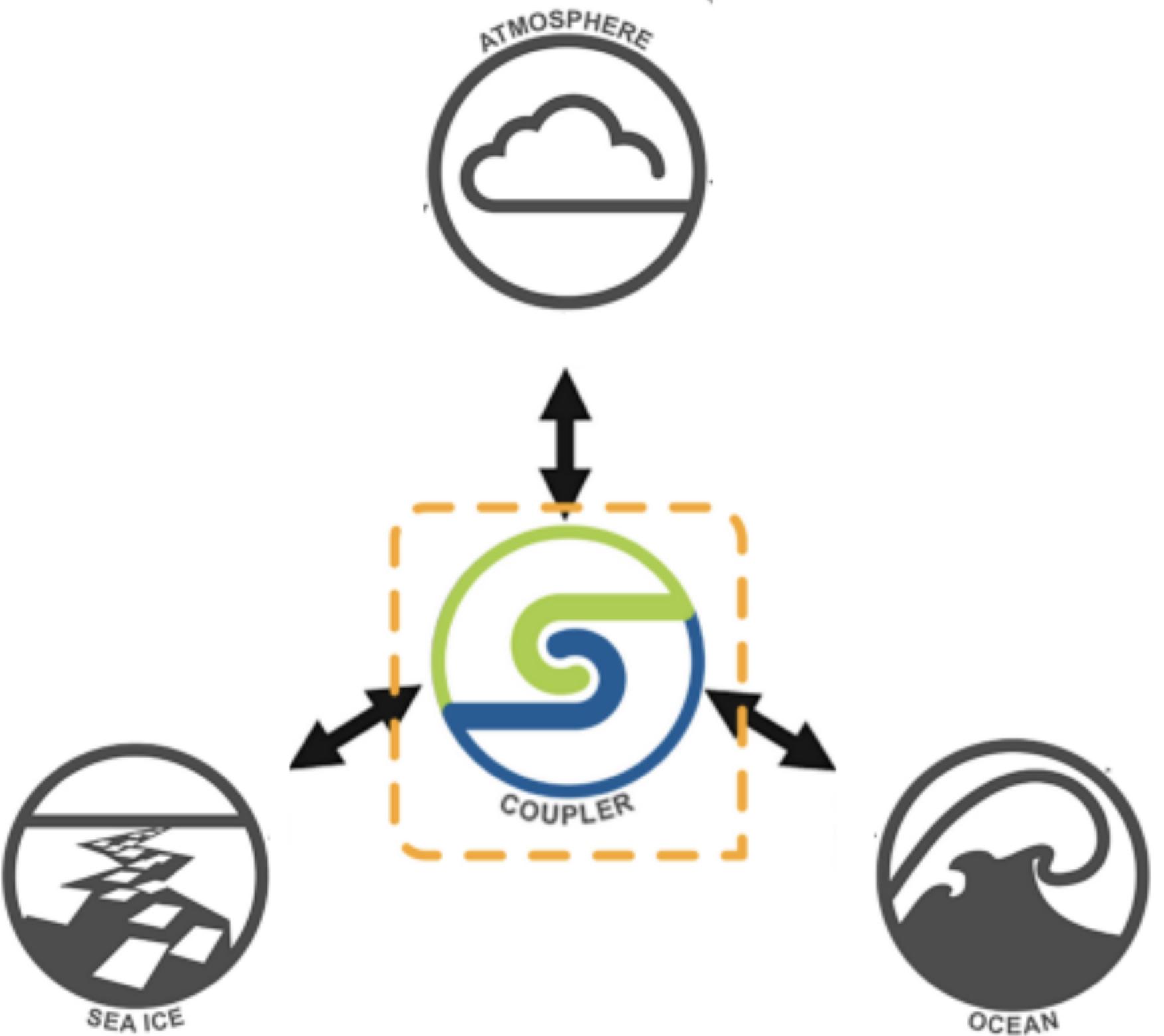
Coupling library between different codes
Interpolate and exchange 2D/3D fields

Widely deployed

Used by 67 climate modeling groups
Used in 5 of the 7 European ESMs used for CMIP6

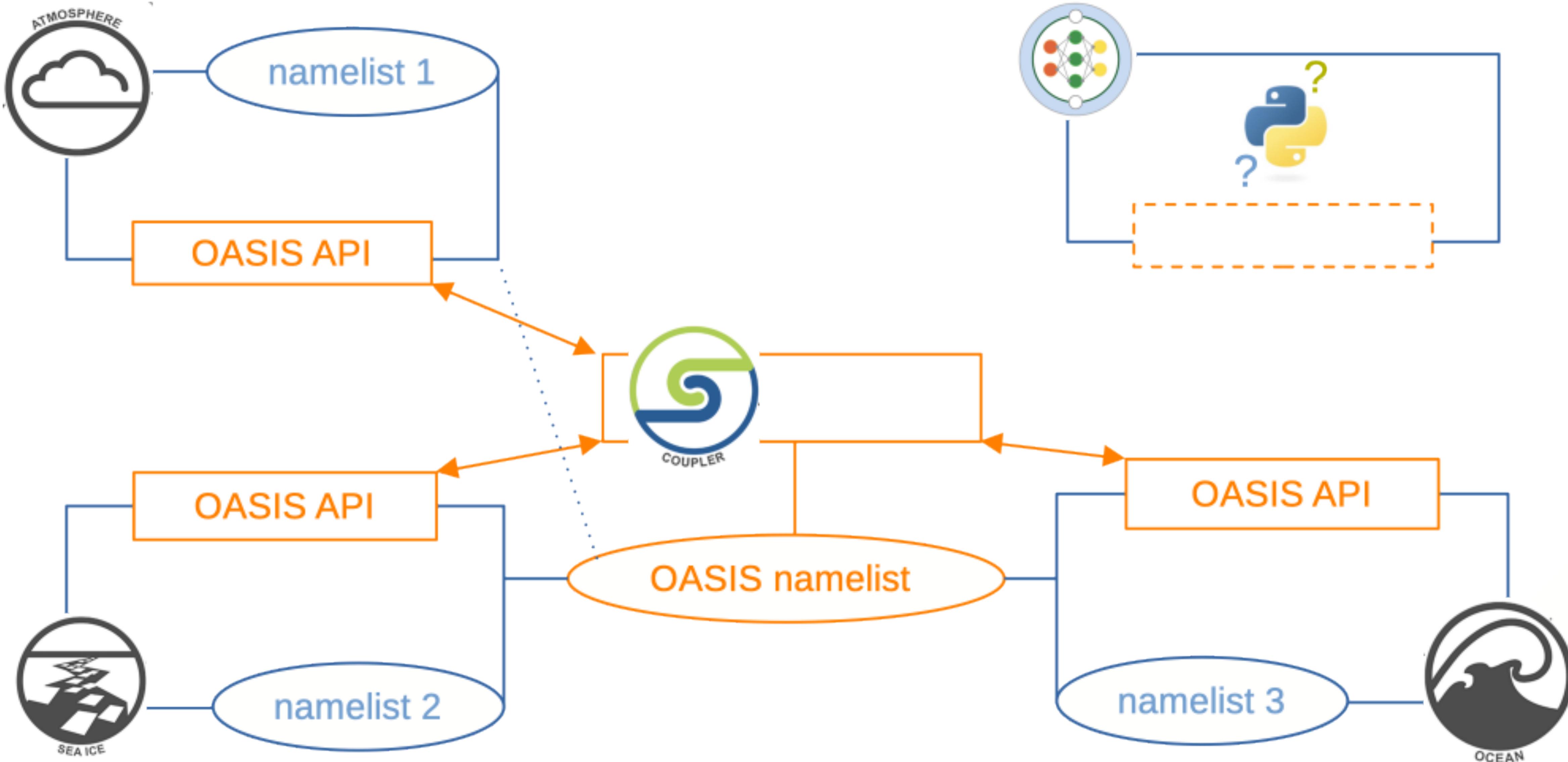
New functionalities since OASIS3-MCT_5.0

Python and C/C++ API



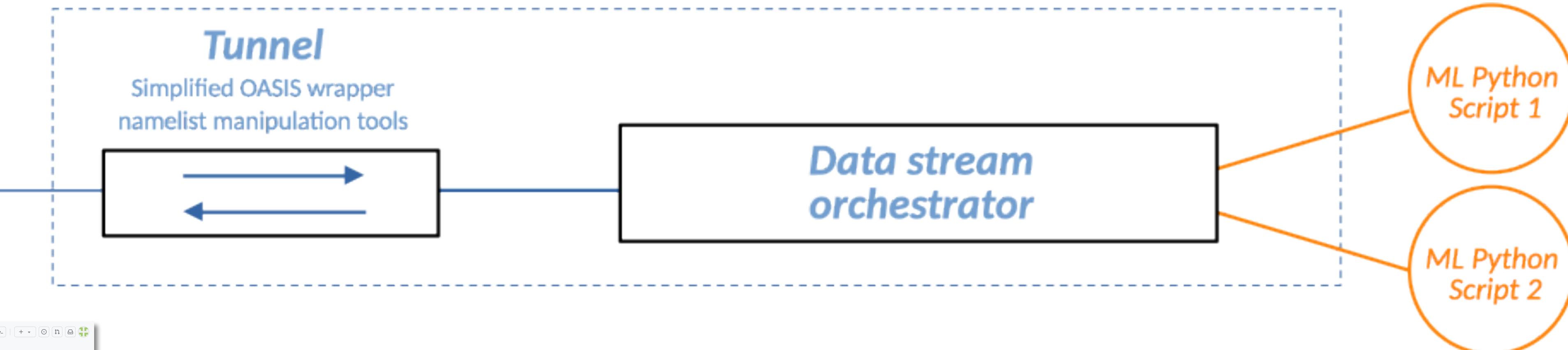
Can we leverage OASIS for hybrid modeling ?

OASIS coupler for hybrid modeling: reality



EOPHIS: a library for deploying ML models through OASIS

OASIS usage	Coupling setup Interpolation / restart definition Namelist
Aimed usage	No interpolation / restart Back and forth exchange only Use by non-experts



The screenshot shows the GitHub repository page for 'meom-group/eophis'. It displays the repository's structure, including files like .readthedocs.yaml, .github/workflows, conf.py, and README.md. The repository has 2 branches, 2 tags, and 71 commits. The 'About' section provides a brief description of the project: 'Coupling inference models with Earth-System simulation codes via OASIS'. The 'Packaging' section indicates no packages are published yet.

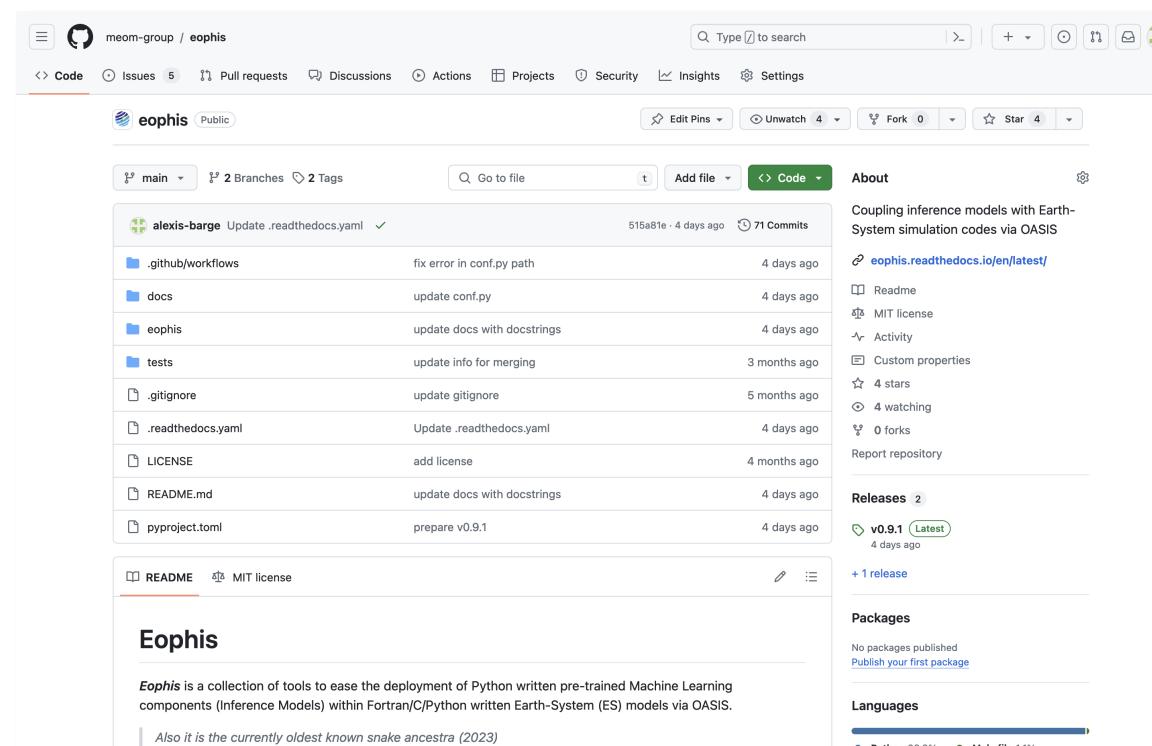
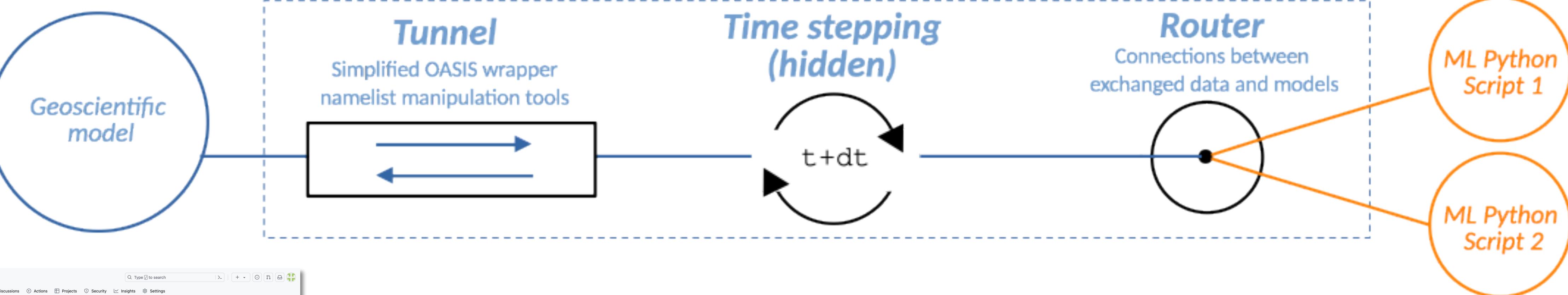
<https://github.com/meom-group/eophis>
MIT Licensed
Open source

Key features

- ML-framework agnostic solution (TF, PyTorch, Jax...)
- ML-models as pluggable scripts with minimal bindings
- Easy sharing of ML-models across geoscientific codes

EOPHIS: a library for deploying ML models through OASIS

OASIS usage	Coupling setup Interpolation / restart definition Namelist	Time advancement for exchange synchronicity
Aimed usage	No interpolation / restart Back and forth exchange only Use by non-experts	Time not required



<https://github.com/meom-group/eophis>
MIT Licensed
Open source

Key features

- ML-framework agnostic solution (TF, PyTorch, Jax...)
- ML-models as pluggable scripts with minimal bindings
- Easy sharing of ML-models across geoscientific codes

NEMO: Current OASIS implementation

OASIS typical steps

MPI environment



cpl_init()

Grid partitioning



cpl_define()

Fields to exchange



cpl_rcv()

Setup coupling



cpl_snd()

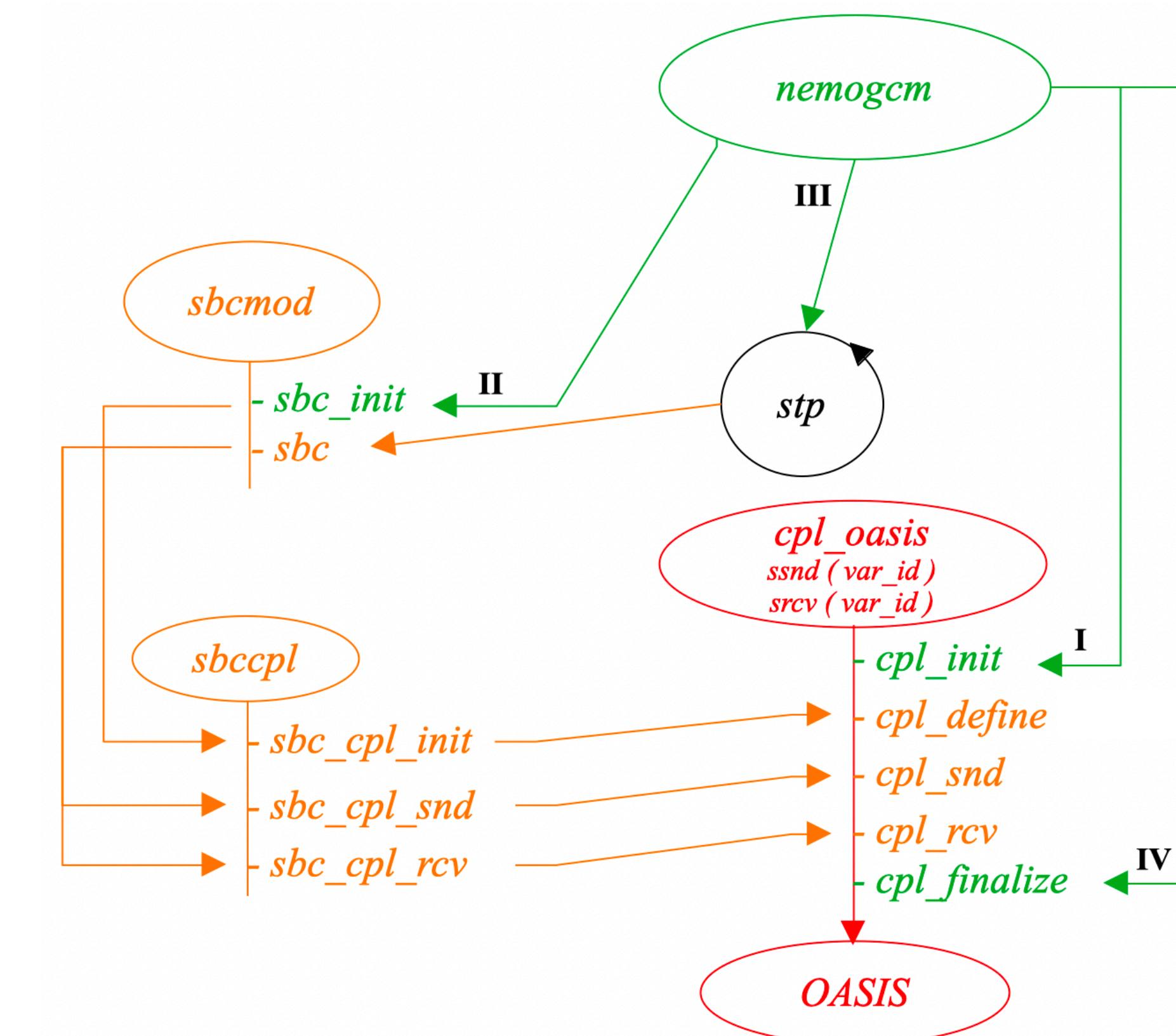
Receive a field



Send a field



NEMO routines



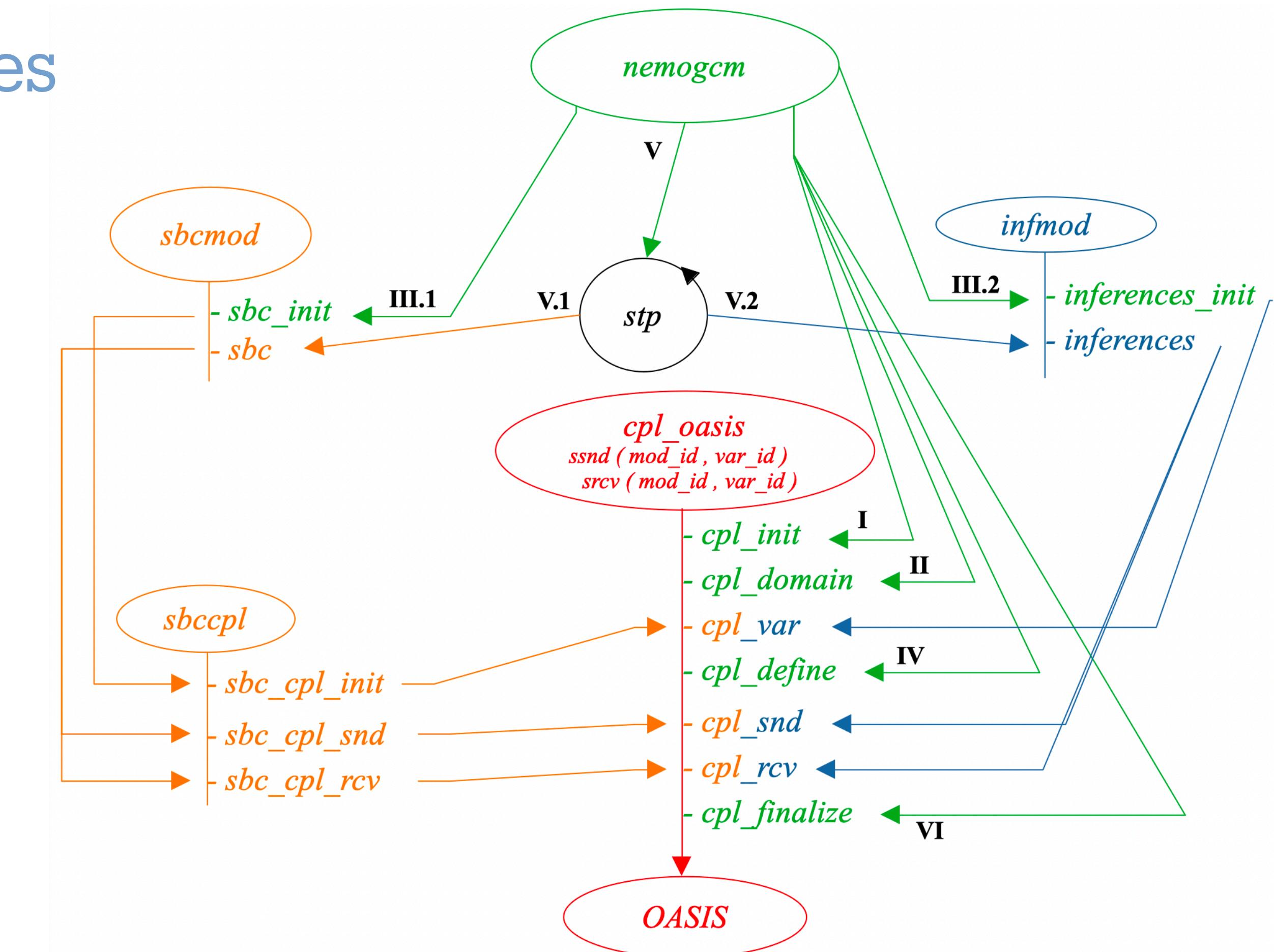
Coupling definition exclusive to SBC

NEMO: Independent OASIS module

OASIS typical steps



Splited routines



Other modules may participate to coupling definition

NEMO: 3D coupling scattered among modules

Coupling fields properties

```
TYPE, PUBLIC :: FLD_CPL          !: Type for coupled field informations
  LOGICAL           :: laction   ! To be coupled or not
  CHARACTER(len = 8) :: cname    ! Field alias used by OASIS
  CHARACTER(len = 1) :: clgrid   ! Grid type
  REAL(wp)          :: nsgn     ! Control of the sign change
  INTEGER, DIMENSION(nmaxcat,nmaxcpl) :: nid   ! Id of the field (no more than 9 categories)
  INTEGER            :: nct      ! Number of categories in field
  INTEGER            :: ncplmodel ! Maximum number of models to/from which this variable
  INTEGER            :: nlvl     ! Number of grid level to exchange, set 1 for 2D fields
END TYPE FLD_CPL
```

Arrays of coupling properties **ssnd / srcv**
Filled during initialization

New attribute for 3D

Sort with modules

ssnd / srcv : now batches of coupling properties arrays

Two-level dimensions with modules and fields IDs

Different size allocation for memory optimization

```
ssnd( mod_ID_1 )%fld( field_ID_1 )%field_property_1
srcv( mod_ID_1 )%fld( field_ID_2 )%field_property_2
ssnd( mod_ID_2 )%fld( field_ID_3 )%field_property_3
```

Module ID passed through OASIS interface

Deployment of ML models in idealized NEMO

GZ21 Model

Stochastic subgrid forcing

Infer stochastic elements from 2D velocities

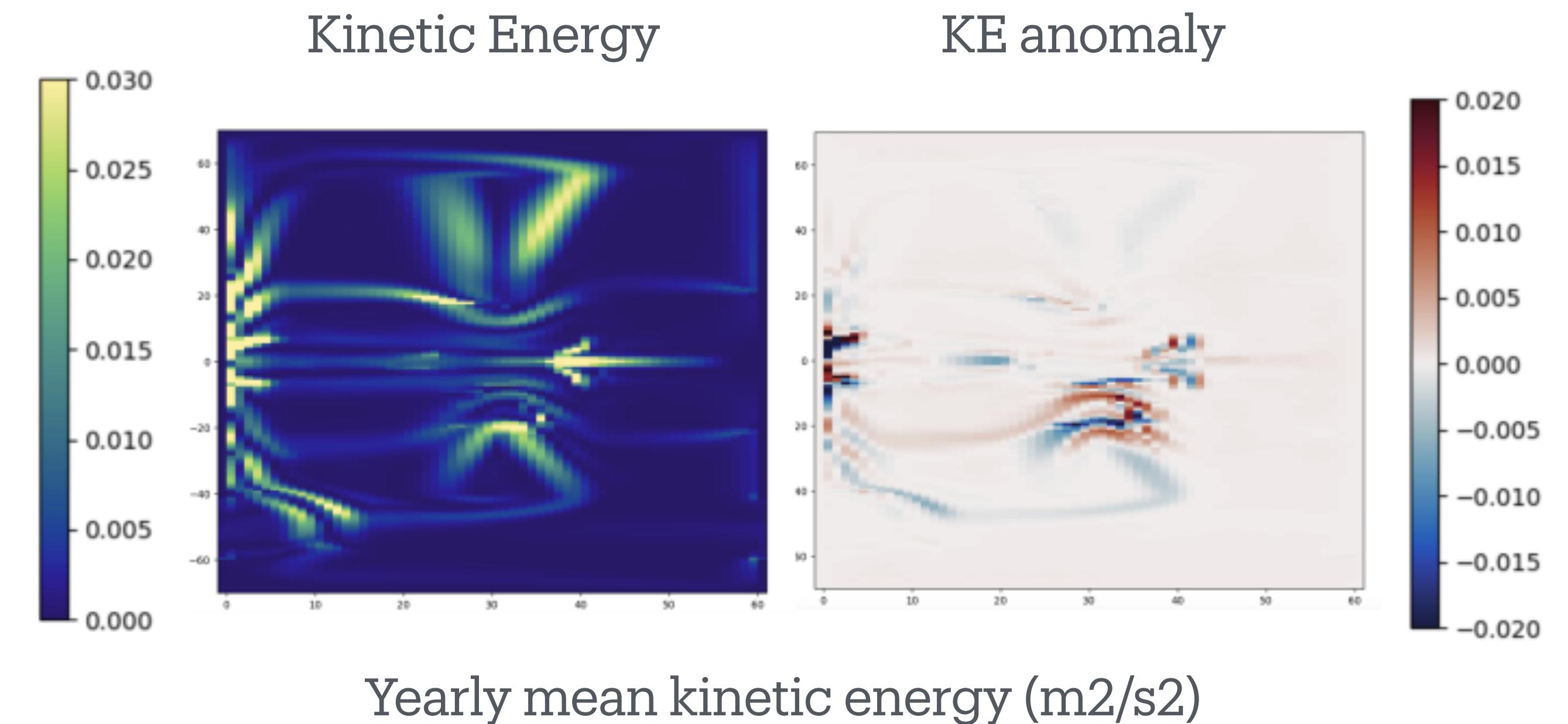
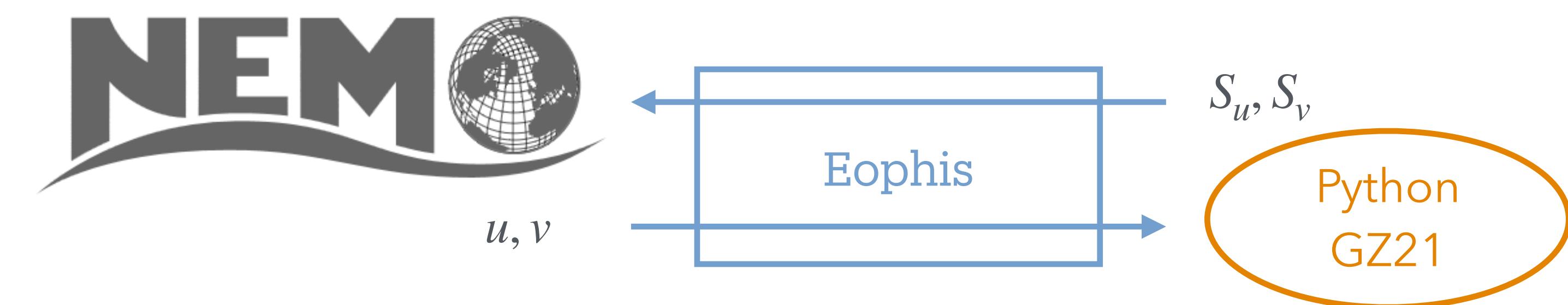
[Guillaumin and Zanna, 2021](#)

DINO.GZ21

5 years low-res *DIabatic NeverWOrld2*

Implementation of retroactive GZ21

Collaboration with E.Meunier and D.Kamm



Deployment of ML models in realistic NEMO

Restratification due to submesoscale eddies

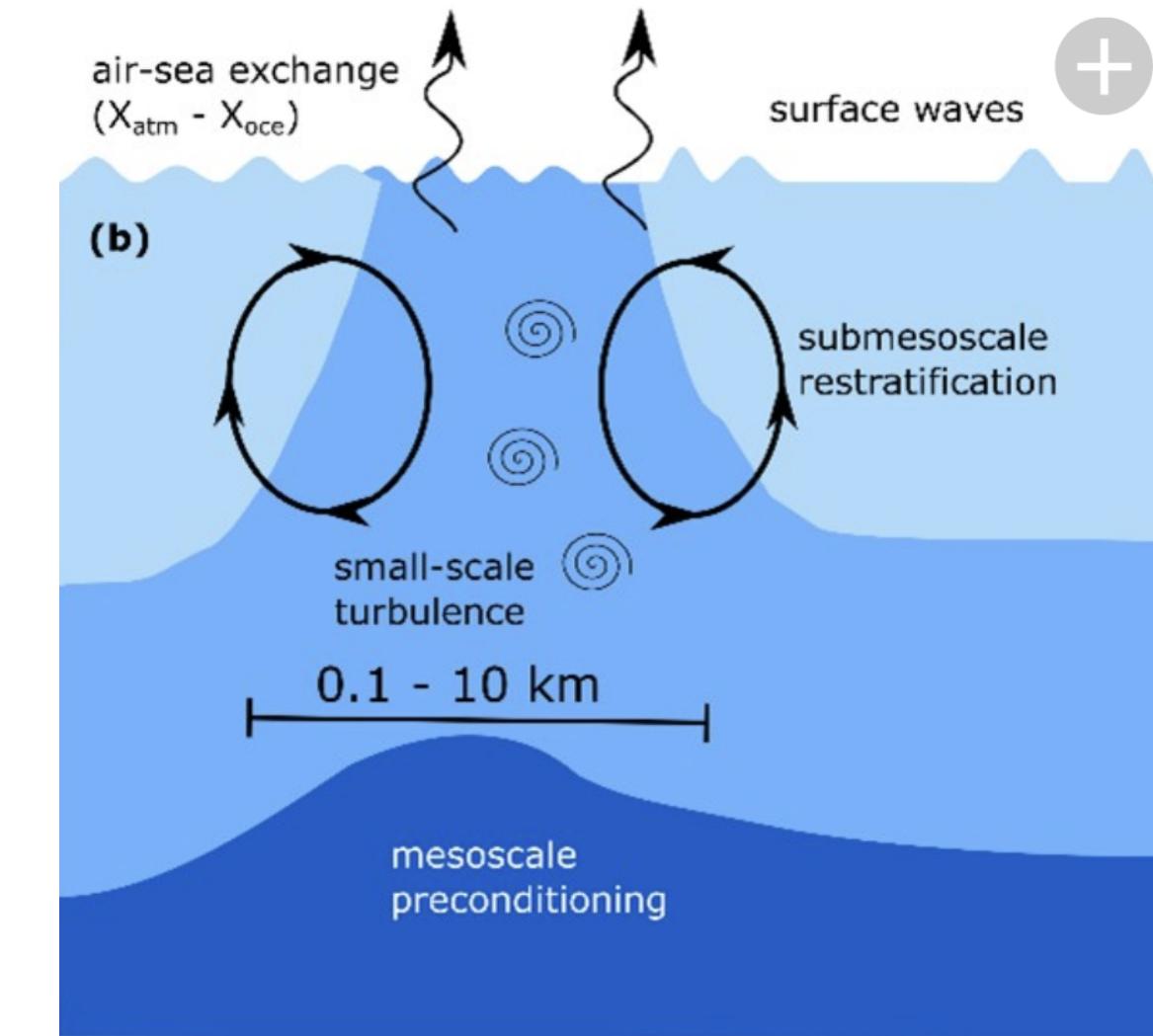
Unresolved vertical transports in ocean near-surface mixed layer

Modeled by Mixed Layer Eddies (MLE) parameterizations

Use of ML parameterizations currently an active research field

[Bodner et al. submitted to JAMES 2024](#)

[Zhou et al., submitted to JAMES 2024](#)



Existing closure framework

[Fox-Kemper et al. 2008b](#)

[Calvert et al. 2020](#)

Submesoscale vertical buoyancy fluxes

$\overline{w'b'_\psi} = \underline{\psi} \times \nabla_H \overline{b^z} \rightarrow$ streamfunction-induced transport

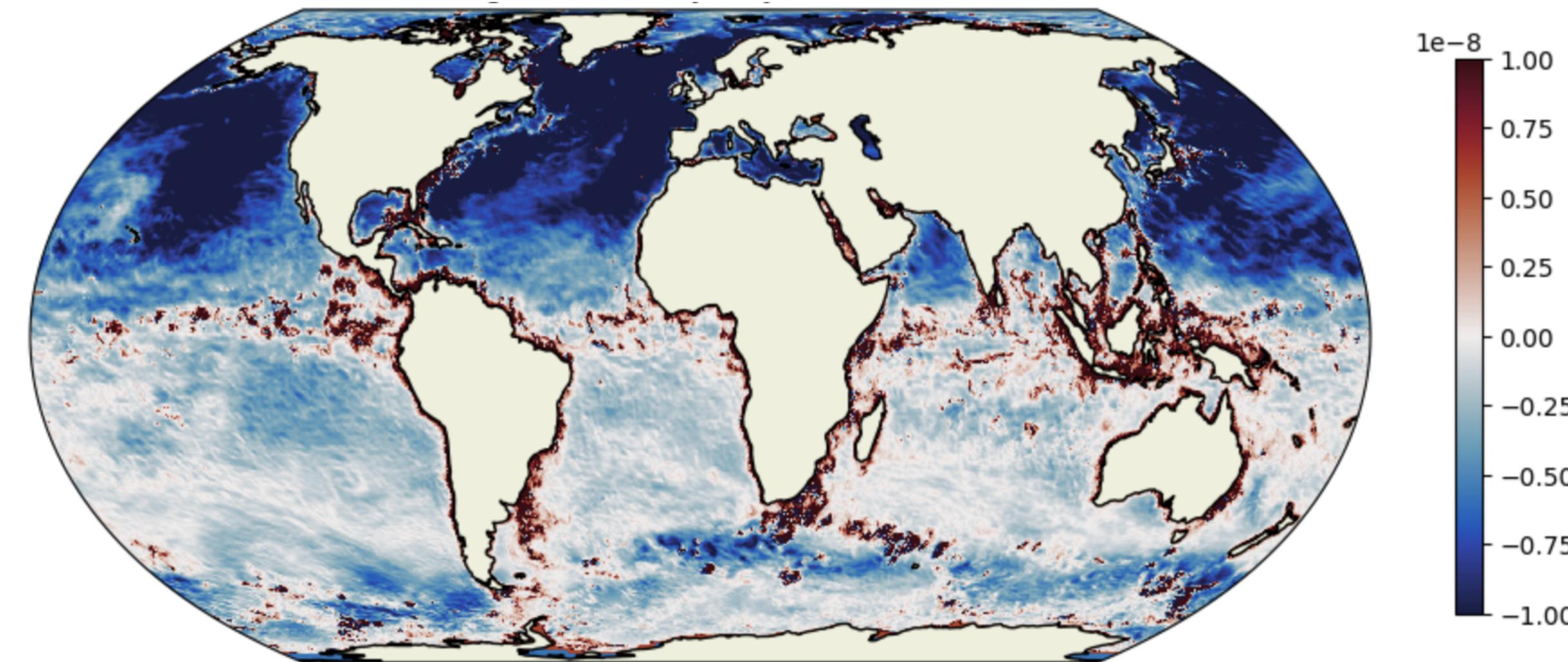
$$\underline{\psi} = (|\overline{\nabla b^z}|, \bar{f}, \overline{H}, \overline{Q^*}, \text{div}(u), \text{rot}(u), |\tau|, \overline{\sigma^z})$$

eORCA025-MLE.BBZ24

10 years forced global 1/4deg simulation

Infer 2D vertical buoyancy fluxes with Bodner et al. model

Collaboration with A.Bodner and D.Balwada



Snapshot of outsourced vertical buoyancy flux (W/m²)

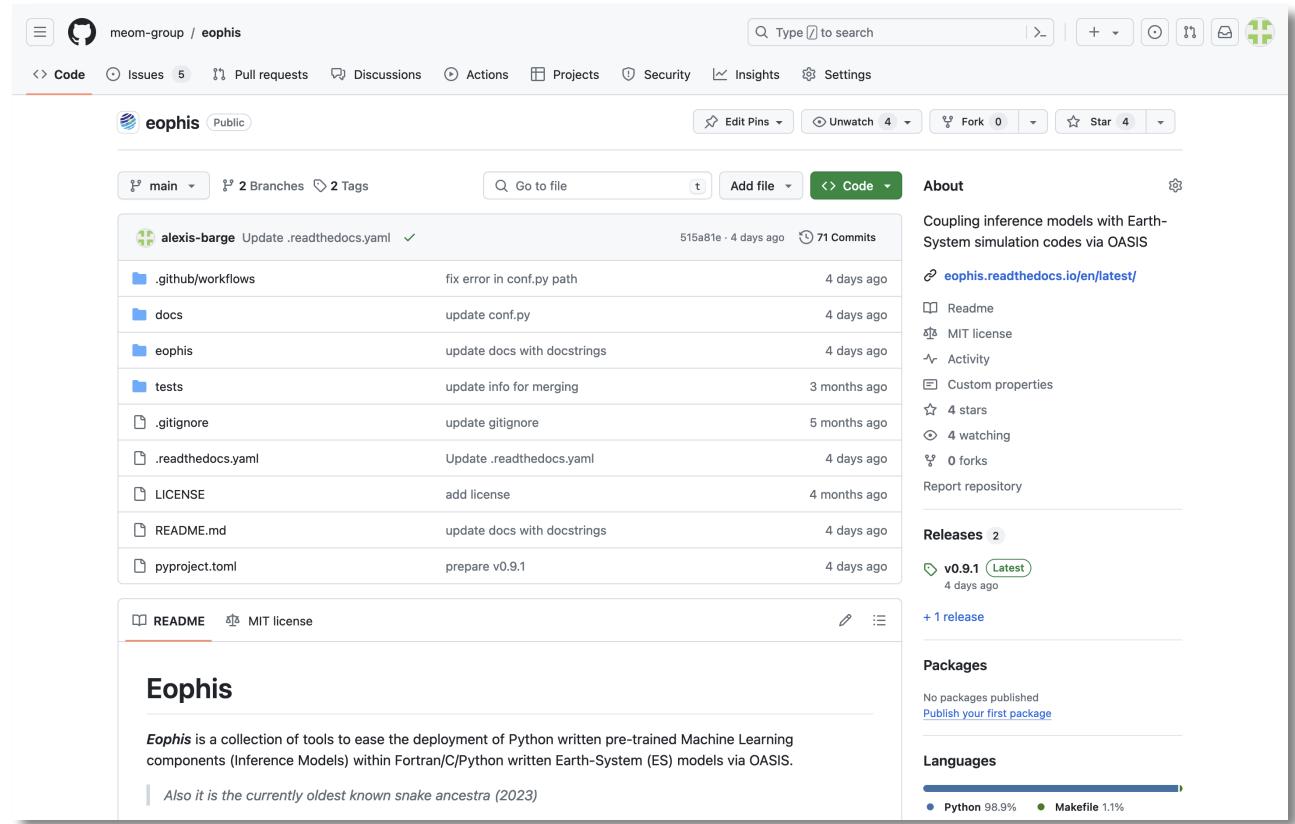
Wrap-up

Eophis: a framework agnostic interface for ML models

Future developments :

- parallel execution with data exchanges across multiple processors

<https://github.com/meom-group/eophis>
MIT Licensed , Open Source



A Github org for sharing ML-based ocean closures (w/ Eophis)

MORAYS-community: examples deployment and use cases

Ocean model agnostic (NEMO, CROCO) with commonly agreed templates

Toward reproducible hybrid ocean modeling

<https://github.com/morays-community>

A screenshot of the GitHub organization page for 'Morays-community'. It shows 6 repositories, 4 teams, and 4 people. The 'Morays' repository is highlighted, showing its README.md file which describes 'Morays' as a project that illustrates how to deploy ML-based components in ocean models with Eophis. It also serves as a platform for collaborative collections of ocean-models experiments conducted within this framework. The repository has 1 star and was last updated 1 week ago.

Getting Started

https://morays-doc.readthedocs.io/en/latest/getting_started.html

A screenshot of the Morays documentation page for 'Running the experiment'. It provides instructions for executing a hybrid experiment. It includes a command-line example:

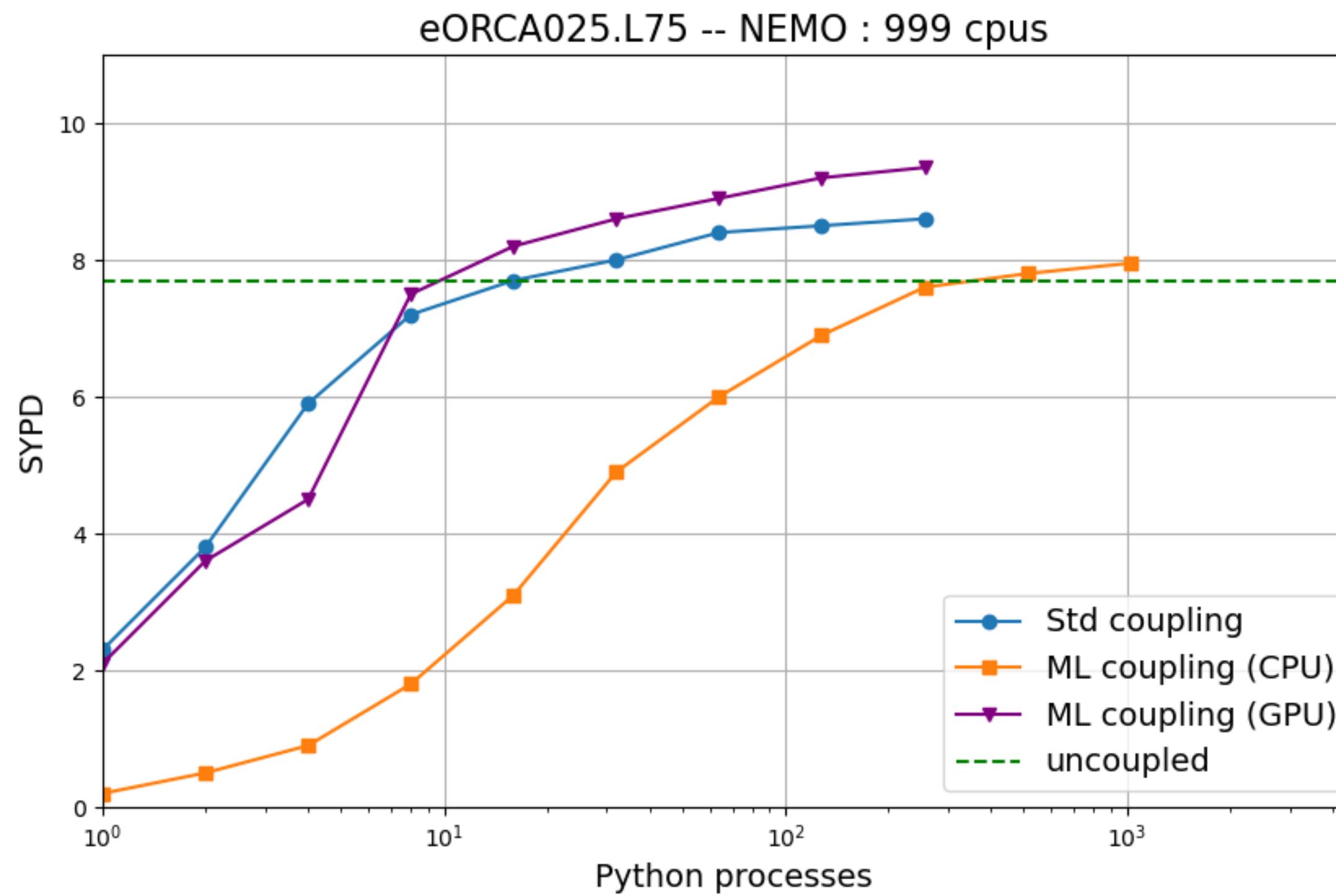
```
cd ~/nemo_v4.2.1/tests/MY_DINO_GZ21/EXP00/
# clean working directory
touch namcouple
rm namcouple*
# execute eophis in preproduction mode to generate namcouple
python3 ./main.py --exec preprod
# save eophis preproduction logs
mv eophis.out eophis_preprod.out
mv eophis.err eophis_preprod.err
# run coupled NEMO-Python
mpirun -np 1 ./nemo : -np 1 python3 ./main.py --exec prod
```

Appendix: Performances

Simulation speed

Realistic global eORCA025 cases

Pure CPU vs CPU/GPU



Time costs

Percentage of iteration time taken by
Total Coupling, Comms and Preds

