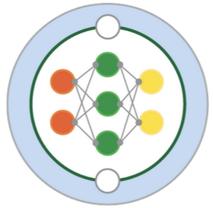
Online deployment of pre-trained machine learning components within geoscientific models through OASIS

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)

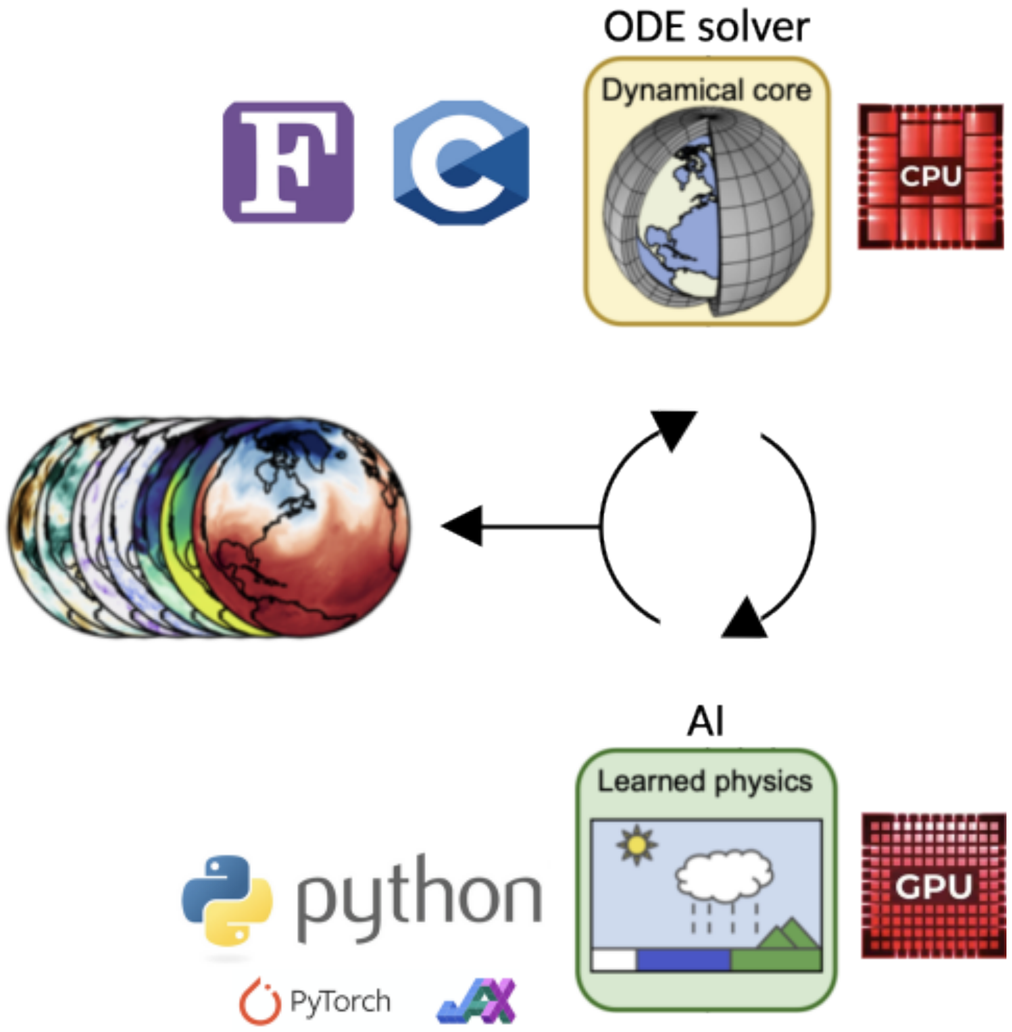
Sane et al. 2023 *Frezat et al. 2022*

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

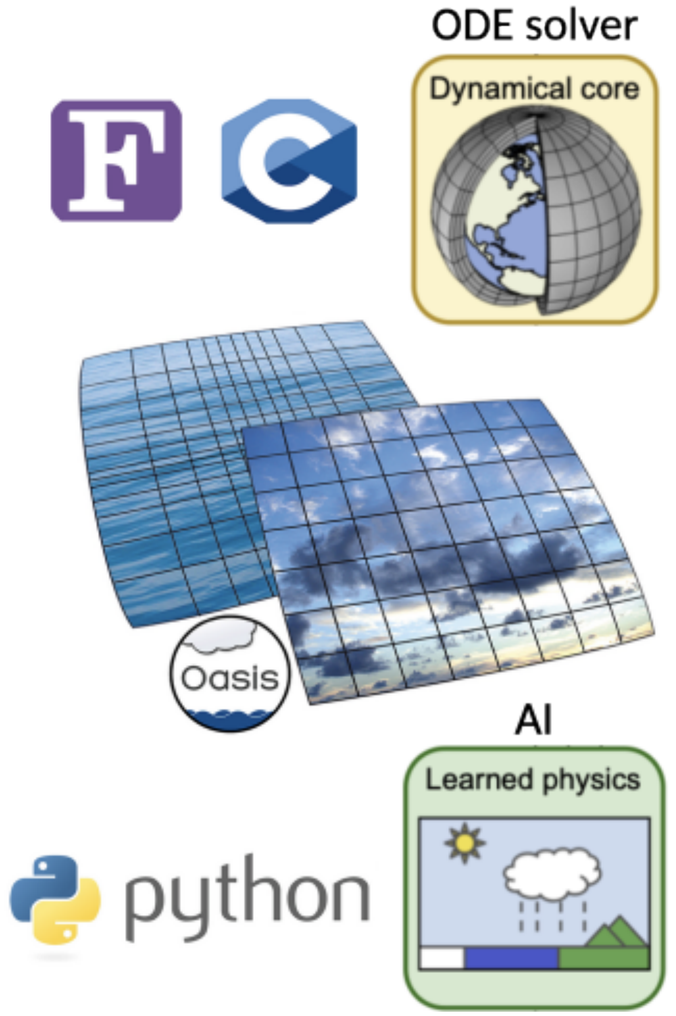
Coupling library between different codes
Interpolate and exchange 2D/3D fields

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

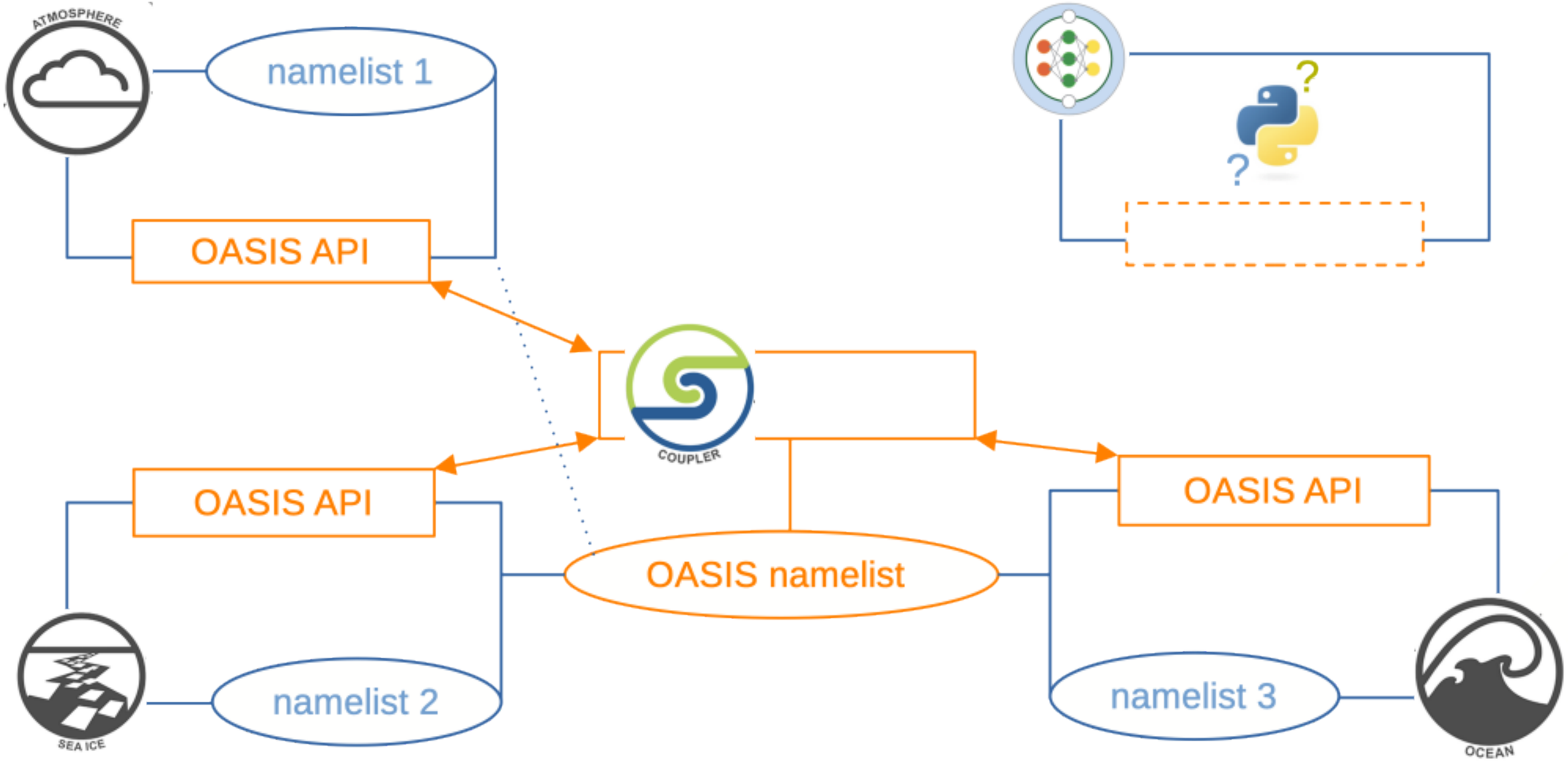
Used by 67 climate modeling groups
Used in 5 of the 7 European ESMs used for CMIP6

New functionalities since OASIS3-MCT_5.0

Python and C/C++ API

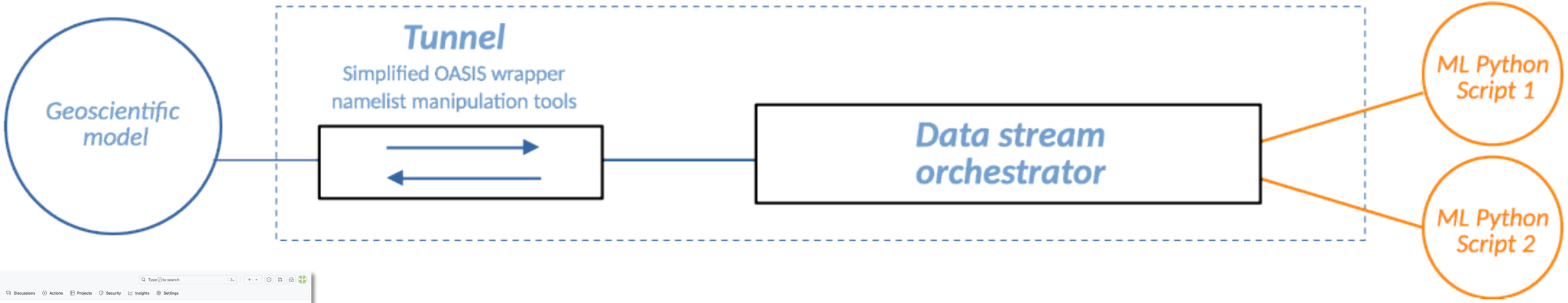
Can we leverage OASIS for hybrid modeling ?

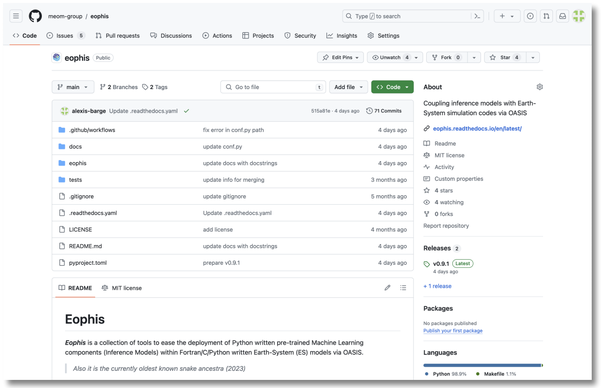
OASIS coupler for hybrid modeling: reality



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





https://github.com/meom-group/eophis

MIT Licensed

Open source

Key features

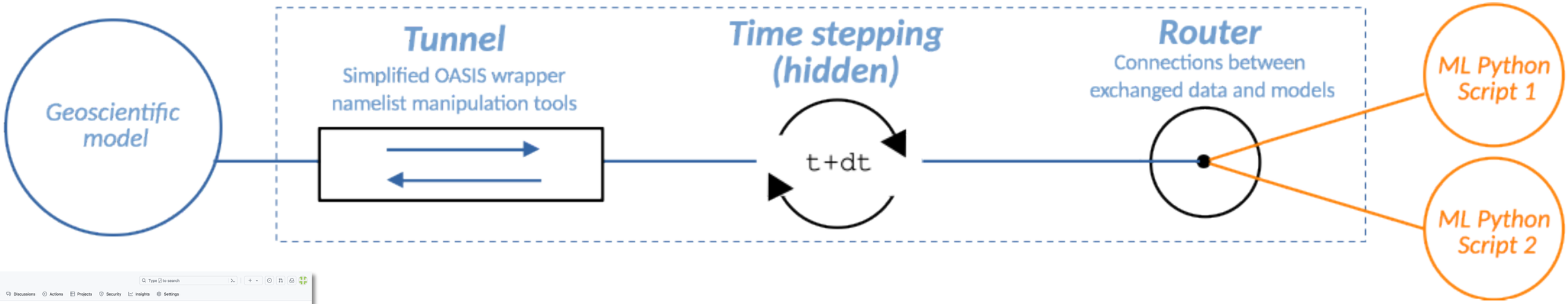
ML-framework agnostic solution (TF, PyTorch, Jax…)

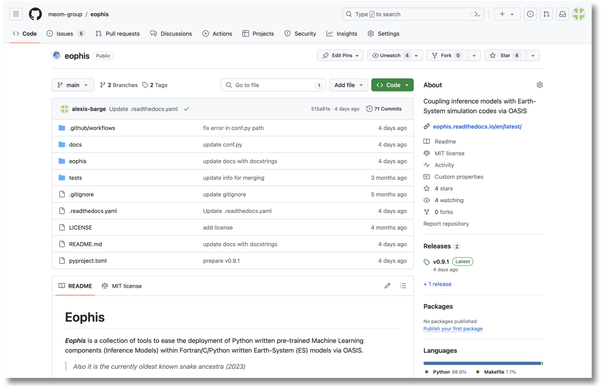
ML-models as pluggable scripts with minimal bindings

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

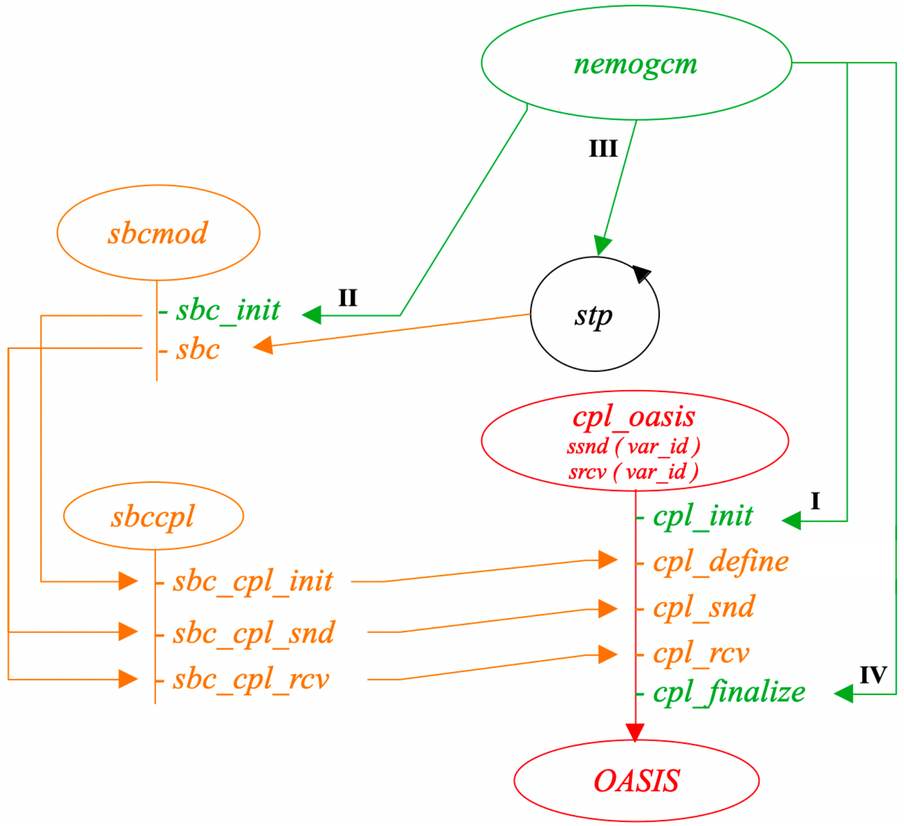
ML-framework agnostic solution (TF, PyTorch, Jax…)

ML-models as pluggable scripts with minimal bindings

Easy sharing of ML-models across geoscientific codes

NEMO: Current OASIS implementation

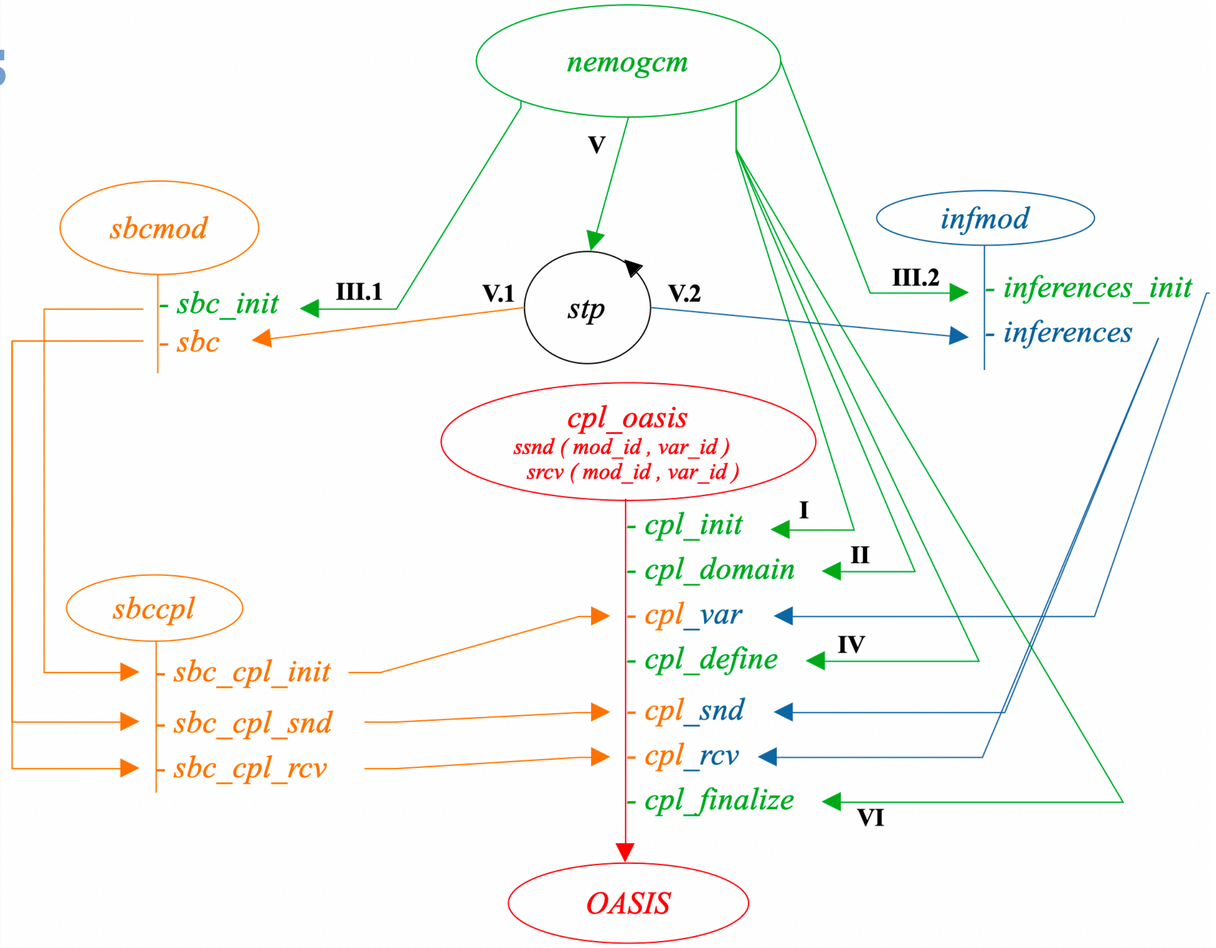
OASIS typical steps		NEMO routines
MPI environment	→	cpl_init()
Grid partitioning Fields to exchange Setup coupling	→	cpl_define()
Receive a field	→	cpl_rcv()
Send a field	→	cpl_snd()



Coupling definition exclusive to SBC

NEMO: Independent OASIS module

OASIS typical steps	Splited routines
MPI environment	cpl_init()
Grid partitioning	cpl_domain()
Fields to exchange	cpl_var()
Setup coupling	cpl_define()
Receive a field	cpl_rcv()
Send a field	cpl_snd()



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)      ::   clname    ! 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
```

New attribute for 3D

Arrays of coupling properties ssnd / srcv

Filled during initialization

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
```

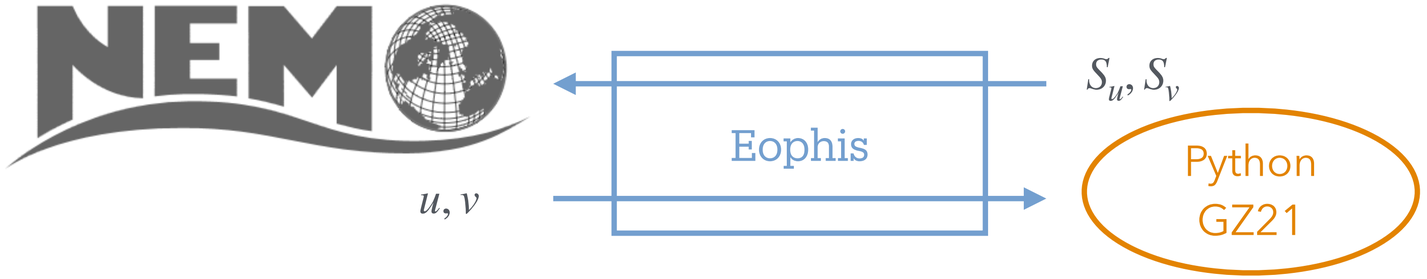
Module ID passed through OASIS interface

https://morays-doc.readthedocs.io/en/latest/nemo.api_5.html

Deployment of ML models in idealized NEMO

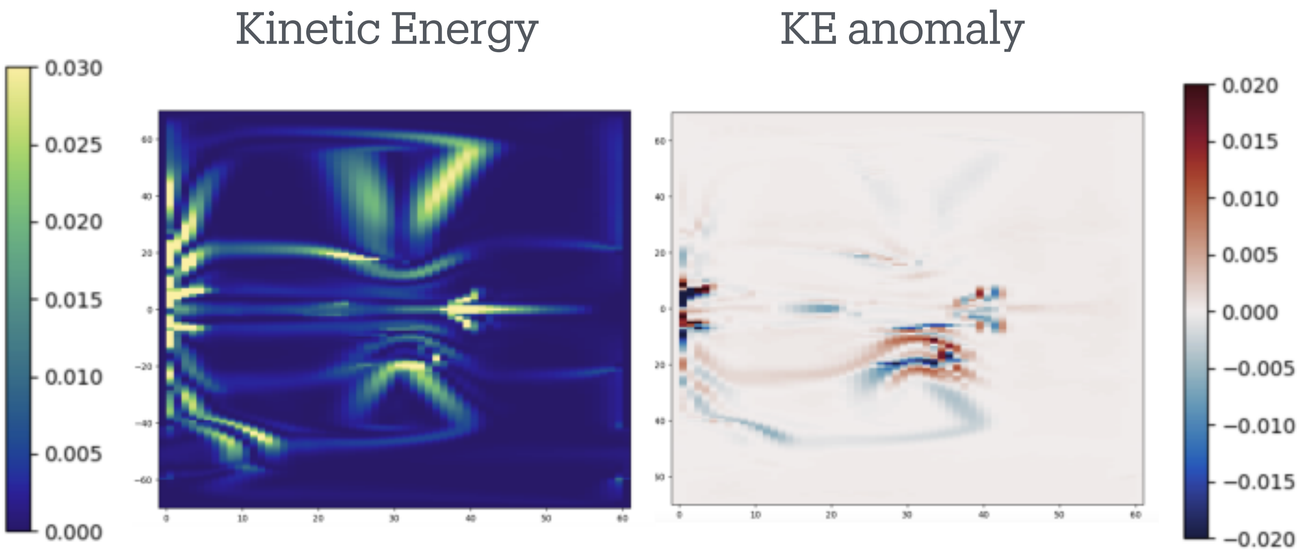
GZ21 Model

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



Yearly mean kinetic energy (m2/s2)

Deployment of ML models in realistic NEMO

Restratification due to submesoscale eddies

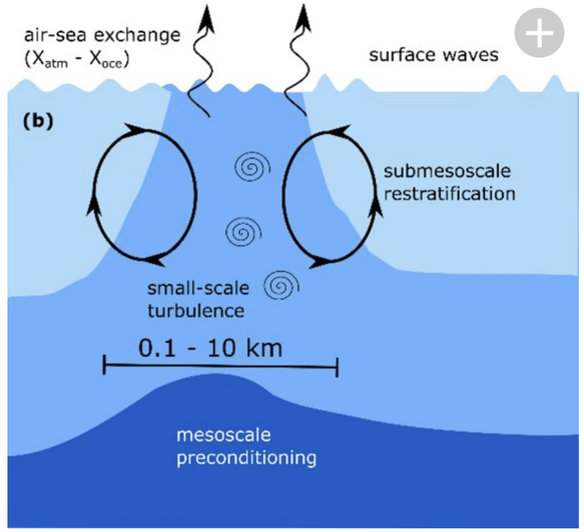
Unresolved vertical transports in ocean near-surface mixed layer

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

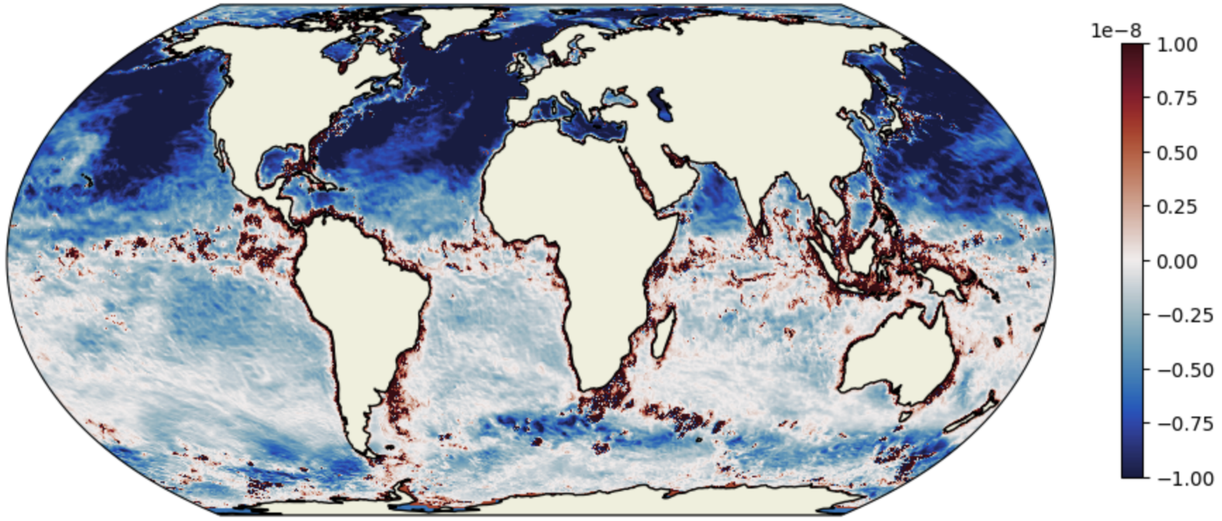
$\underline{\psi} = \left(|\overline{\nabla b^z}|, \bar{f}, \overline{H}, \overline{Q^*}, div(u), rot(u), |\tau|, \overline{\sigma^z} \right)$

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



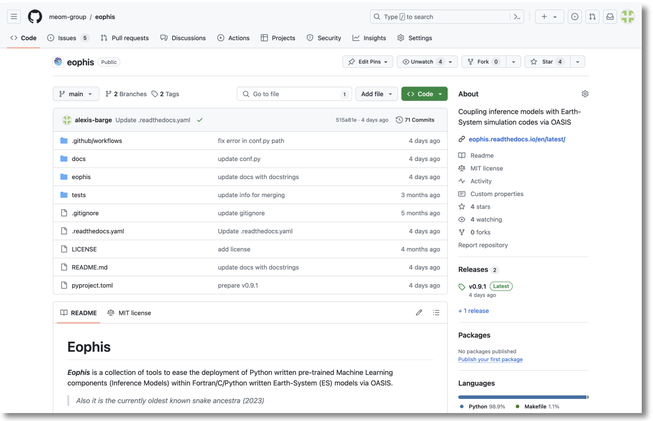
Snapshot of outsourced vertical buoyancy flux (W/m2)

Wrap-up

Eophis: a framework agnostic interface for ML models

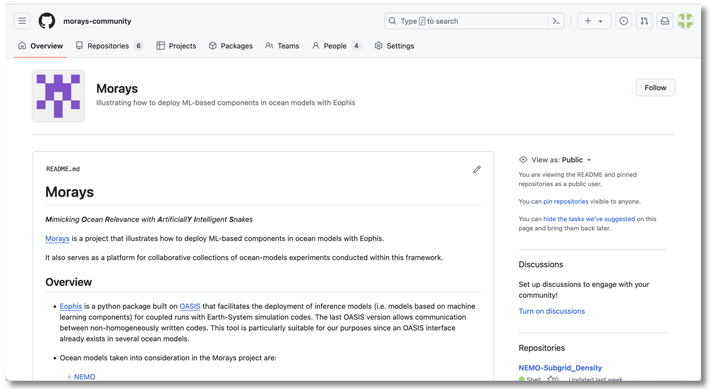
Future developments :

- 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

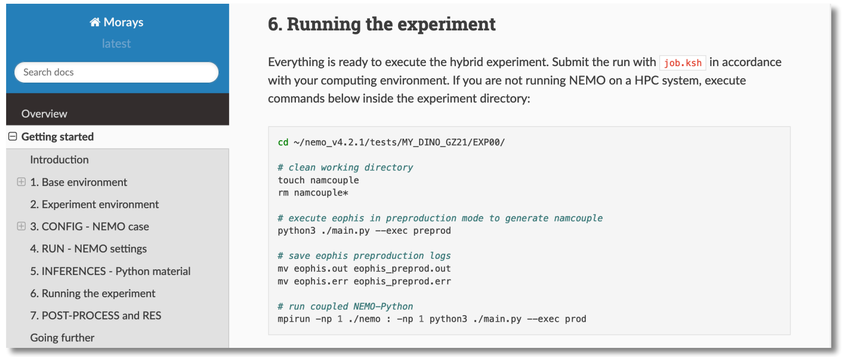
Ocean model agnostic (NEMO, CROCO) with commonly agreed templates

Toward reproducible hybrid ocean modeling

https://github.com/morays-community

Getting Started

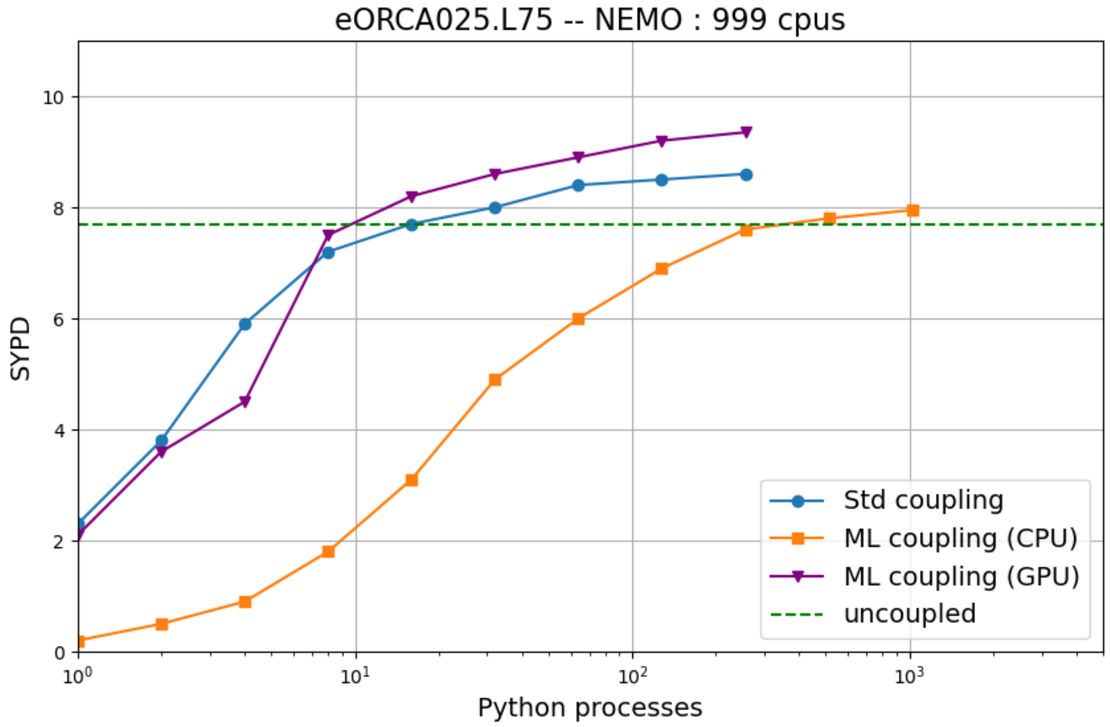
https://morays-doc.readthedocs.io/en/latest/getting_started.html



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