Machine Learning models to emulate Gravity Wave Drag by Atos Center of Excellence

Alexis Giorkallos, Christophe Bovalo – *Atos BDS R&D AI4Sim*
Matthew Chantry, Peter Düben - *ECMWF*

ECMWF 19th workshop on HPC in Meteorology
24/09/2021

Atos ThinkAI

Strategize your AI journey, for any industry now!

ThinkAI

Advise — Center of Excellence in Advanced Computing

Industry-contextualized & sustainable high-performance AI

Architect	Best-of-breed hardware & software	AI-driven optimization toolbox	Digital security
Accomplish	E2E solution orchestration	Operationalization & industrialization	Scalability

Visit: Atos-Thinkai

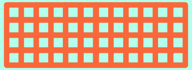
01. AI4Sim at a glance

Atos

Projection in a subspace where it’s easy to work. Usually, domain reduced with symmetries, etc…

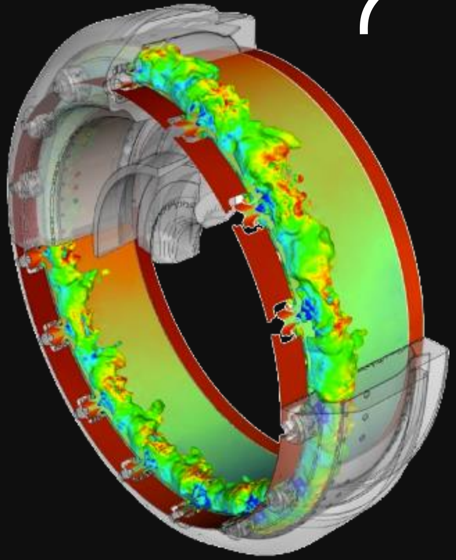
flattened domain

Unstructured mesh.
Raw data



Sub-domain where the phenomenon is happening. Depending on the model, possibly cartesian grid

Replace computation of terms of the equation traditionally done by the solver by a pre-trained neural model, coupled with the solver



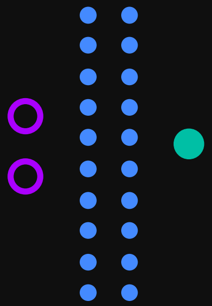
Real simulation within combustion chamber

Solving the equations is an intractable computational nightmare.

So the idea is to capture the essence of the Physics with a data-driven approach.

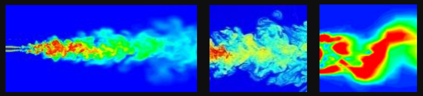
+ mesh-free, @ every energy scale

- **RAISE.** H2 combustion chamber
- **ECMWF.** Weather Forecast
- **AIRSEA.** Oceanic-atmospheric coupling



- MLPs
- CNNs
- Unets
- GNNs
- PINNs

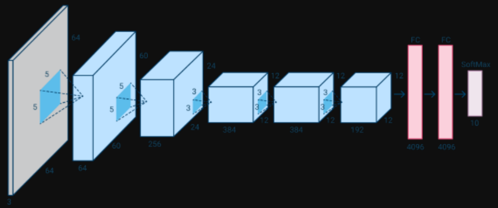
NN trained on simpler simulations that isolate the Physics to capture. Look for good generalization properties



AI4Sim at a glance

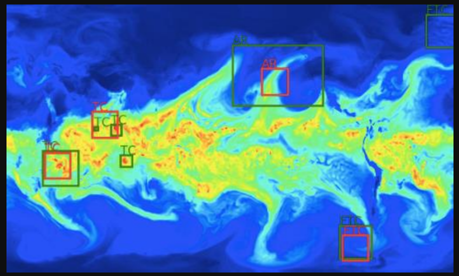
AI to back simulations

Atos



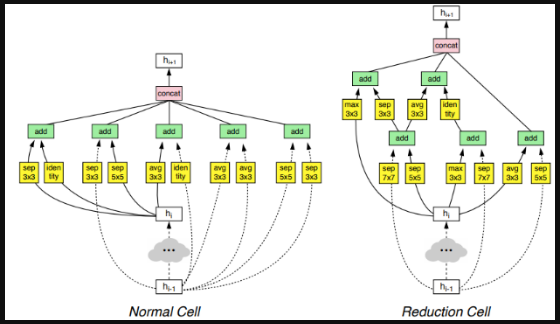
Model Architecture

- Exploring DL to surrogate Physical Models (MLP, CNN)
- Working with unstructured grids, mesh-free approach (GNN), Physics-Informed modeling (PINN, HNN)



Coupling AI + Sim

- Advanced data coupling between ML inference engine & numerical solvers
- AI/HPC workflow orchestration for continuous improvement



Meta-learning

- HPO and topologies Optimization (Bayesian, NAS)
- Automatic data refinement for surrogate modeling and simulation efficiency

AI4Sim at a glance

General approach to fluid simulation problems

02. Center of Excellence in Weather & Climate modeling

Atos

Center of Excellence in Weather & Climate modelling

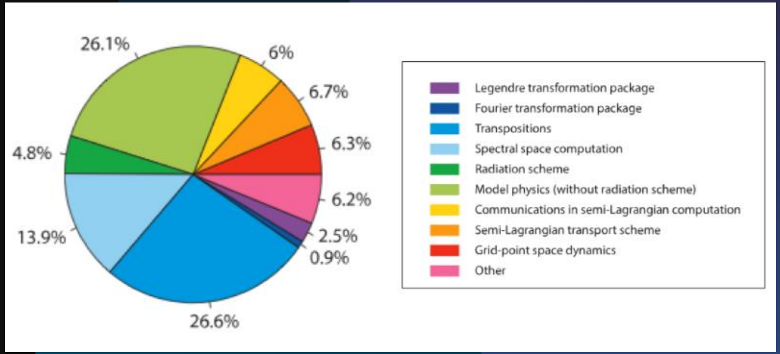
AI for Weather Forecast objectives

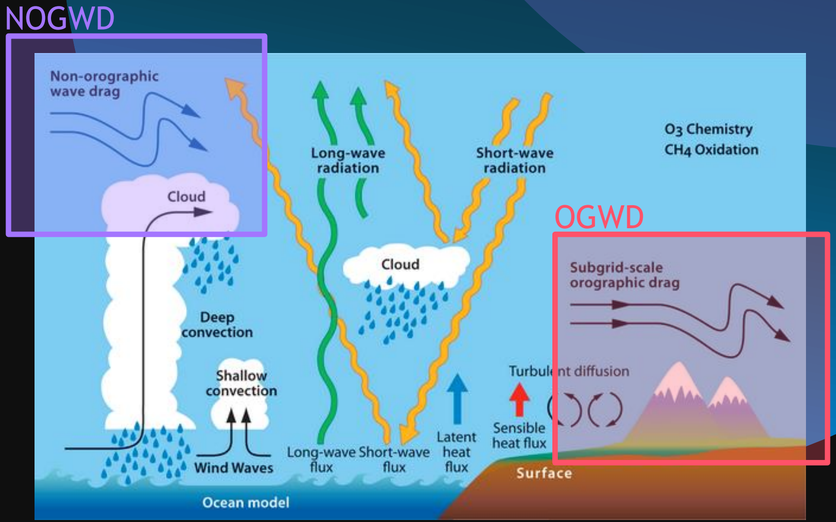
- ✓ Develop vanilla solutions (DL Models) for the emulation of physical parametrization schemes
- ✓ Develop vanilla solutions for ML applications that take the 3D state of the atmosphere on unstructured grids as well as scale interactions in both space and time into account.
- ✓ Develop vanilla solutions for feature detection in 3D IFS model output.
- ✓ Deliver a fully functional workflow ML-IFS (coupling IFS with ML libraries, enhance data workflow to enhance ML applications that use IFS outputs, integrate ML in the product-generation workflow)

03. Emulating the Orographic Gravity Wave Drag parameterization

General Context
Physical parameterizations

- Within IFS, not all physical processes are explicitly resolved by the model **BUT** they need to be represented. They account for 30% of the computational cost
- **Parameterization** (subgrid-scale scheme) = simplified representation of physical processes that are too complex or too small-scale (< grid resolution)
- Within IFS, one parameterization has been successfully emulated (non-orographic gravity wave drag by Chantry et al., 2021)
- The “real” evaluation is done by injecting the trained NN into IFS



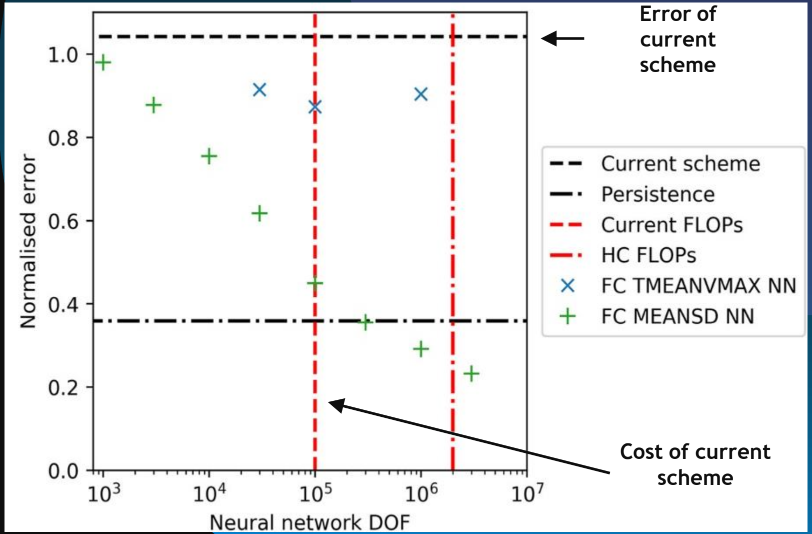


Bauer et al., 2015

Chantry et al.'s study

Emulation of the Non-Orographic Gravity Wave drag (NOGWD) parameterization

- Training set comes from a higher complexity version of the operational NOGWD parameterization
- The accuracy linearly increases with the NN DoF
- The NN version of NOGWD outperforms the operational version within the same time constraint
- Dramatic performance gain when inference is done on GPU
- More complicated to export to the OGWD scheme

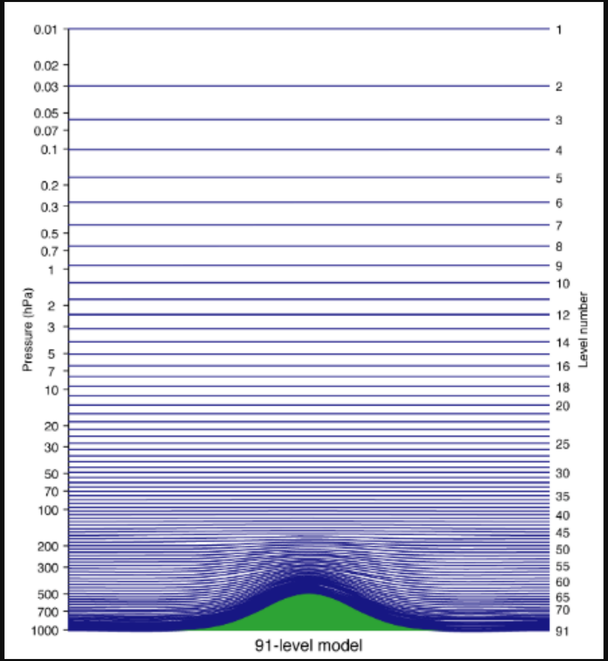


Chantry et al., 2021

Chantry, M., Hatfield, S., Dueben, P., Polichtchouk, I., & Palmer, T. (2021). Machine learning emulation of gravity wave drag in numerical weather forecasting. *Journal of Advances in Modeling Earth Systems*, 13, e2021MS002477. https://doi.org/10.1029/2021MS002477

Emulation of the OGWD parameterization

Problem and Inputs



IFS L91 model levels

Supervised regression problem

Raw dataset (~60 Gb) from **IFS simulations**

N ~ 28M

$\boldsymbol{X}$ (5 x 91 + 4 = 459 features)

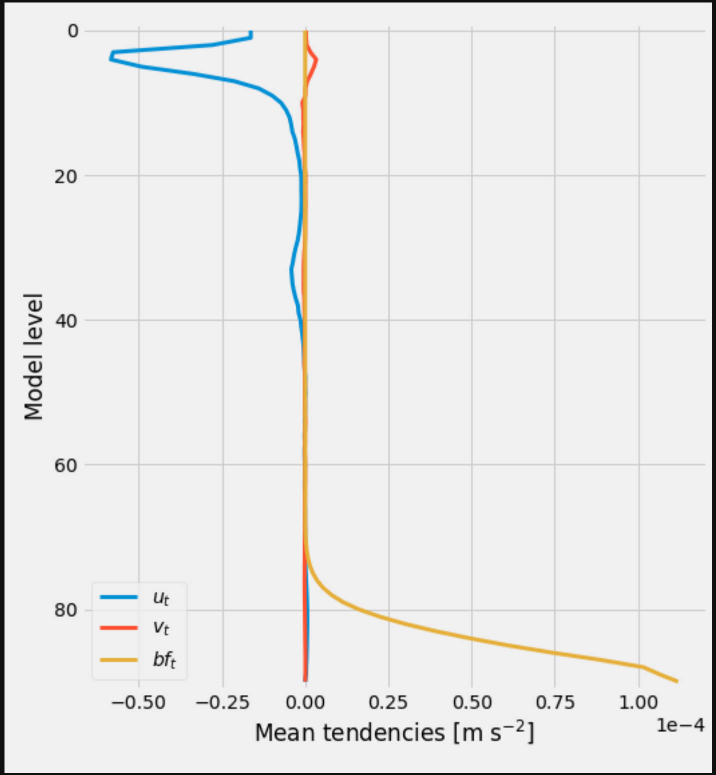
- *on each model level*
  - *u, v horizontal wind velocities*
  - *T temperature*
  - *p pressure*
  - *g geopotential*
- *+ 4 surface params* defining the subscale orography

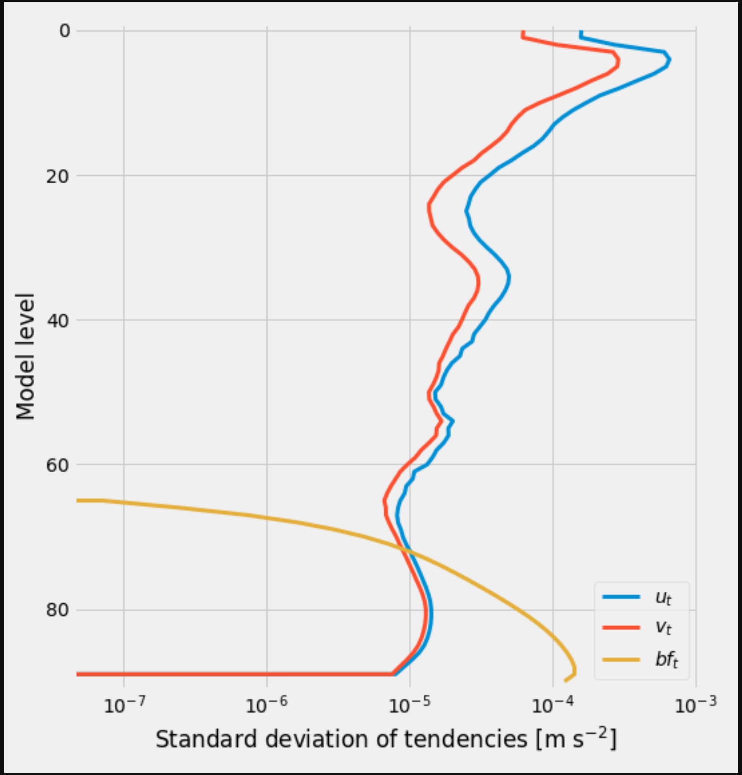
$\boldsymbol{Y}$ ground truth (3 x 91 = 273 targets)

- *u_t v_t gravity-wave drag tendencies*
- *bf_t blocking drag tendencies*

Description of the targets

Mean and Variance of targets



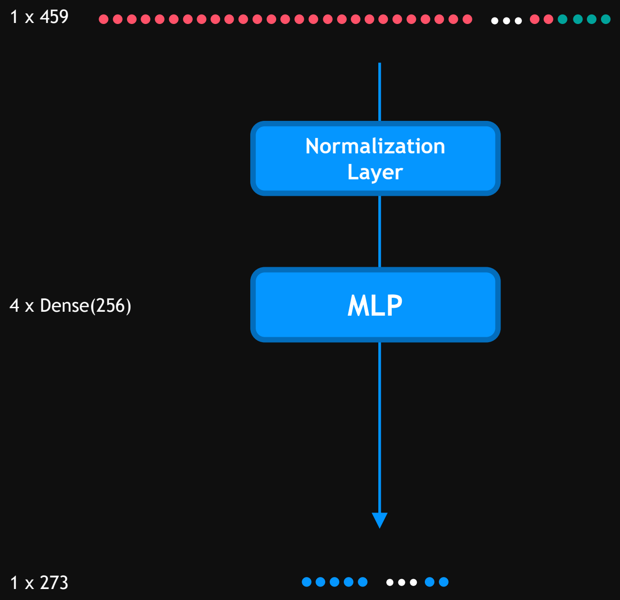


Models

Currently, two approaches

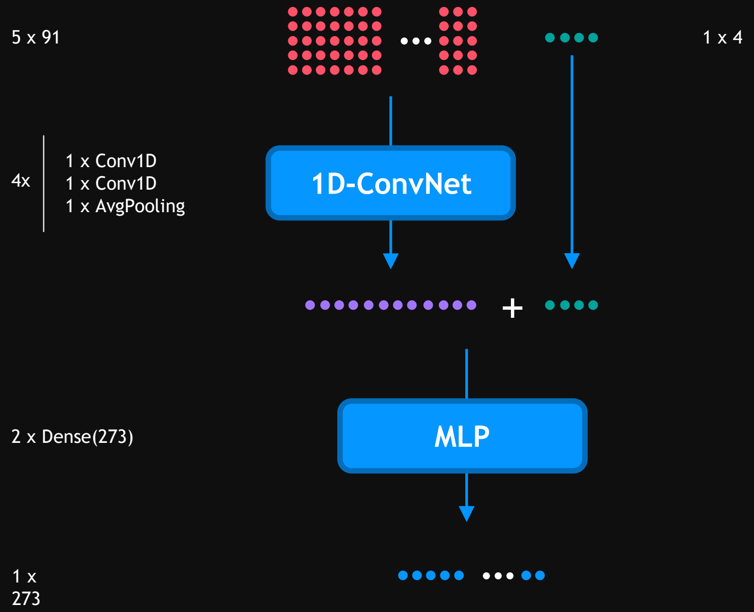
Multi-Layer Perceptron (MLP)

Input = raw X



1D Convolutional Neural Network (CNN)

Input = 2D map built from X, then normalized



Emulation of OGWD
Preliminary results

- Current best fit (Sazna) models on splitted labels.
  - 1x MLP for u_t and v_t
  - 1x MLP for bf_t
- The emulation of physical parameterizations is a challenging and ongoing task
- We are still working side-by-side with Matthew and Peter to achieve the emulation of OGWD scheme

Conclusions and Perspectives

- Active collaboration between Atos and ECMWF through the CoE in Weather & Climate modeling
- Partner’s results and approach reproduced
- OGWD, a challenging problem !
- Current implementations are limited by the triggering method of the scheme
- Still exploring modeling options to improve the model quality
- A good model quality in train/test process does not necessarily imply a proper behavior within IFS

1. Improving the input data 'quality'
   - Better understanding the data distribution
   - Filter data when scheme is not active
2. Atos AI-Simulation coupling solutions
   - **Weak coupling** : inter-nodes inference engine containerized in Docker/Singularity images with MPI communication
   - **Strong coupling** : intra-nodes C++ Tensorflow/PyTorch (binding C++ and Fortran) with zero copy data exchange
3. Modeling
   - Train a NN to predict the blocking drag tendencies
   - Ongoing experiments with GNNs

Thank you!

For more information please contact:

Product Owner
gael.goret@atos.net

Product Managers
cedric.bourrasset@atos.net
arturo.mondragon@atos.net

Developers
christophe.bovalo@atos.net
alexis.giorkallos@atos.net