iML-WKW Use Case TurboSWAN

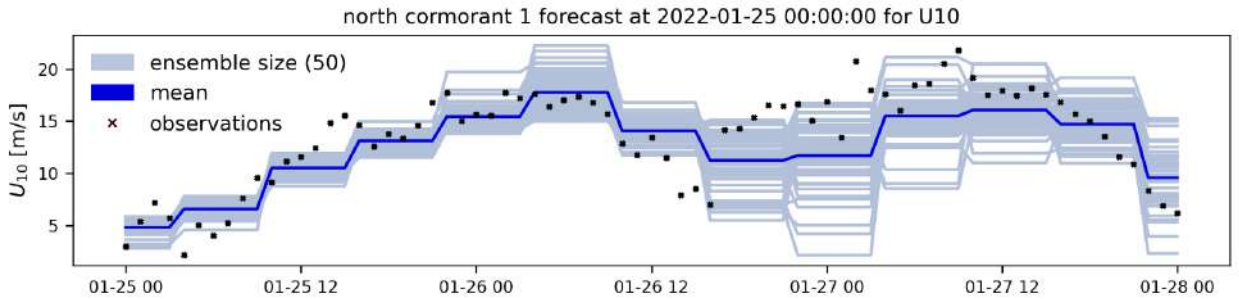
AIWAM meets iML-WKW

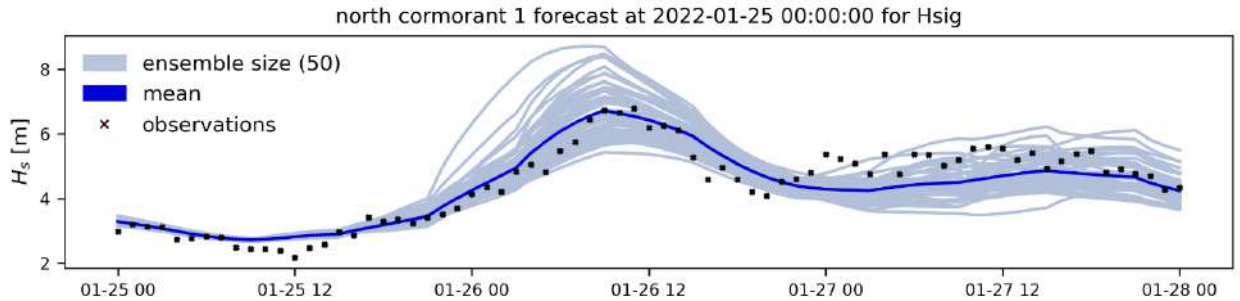
Caroline Gautier
Joost van Gijn
Madhav Mohan

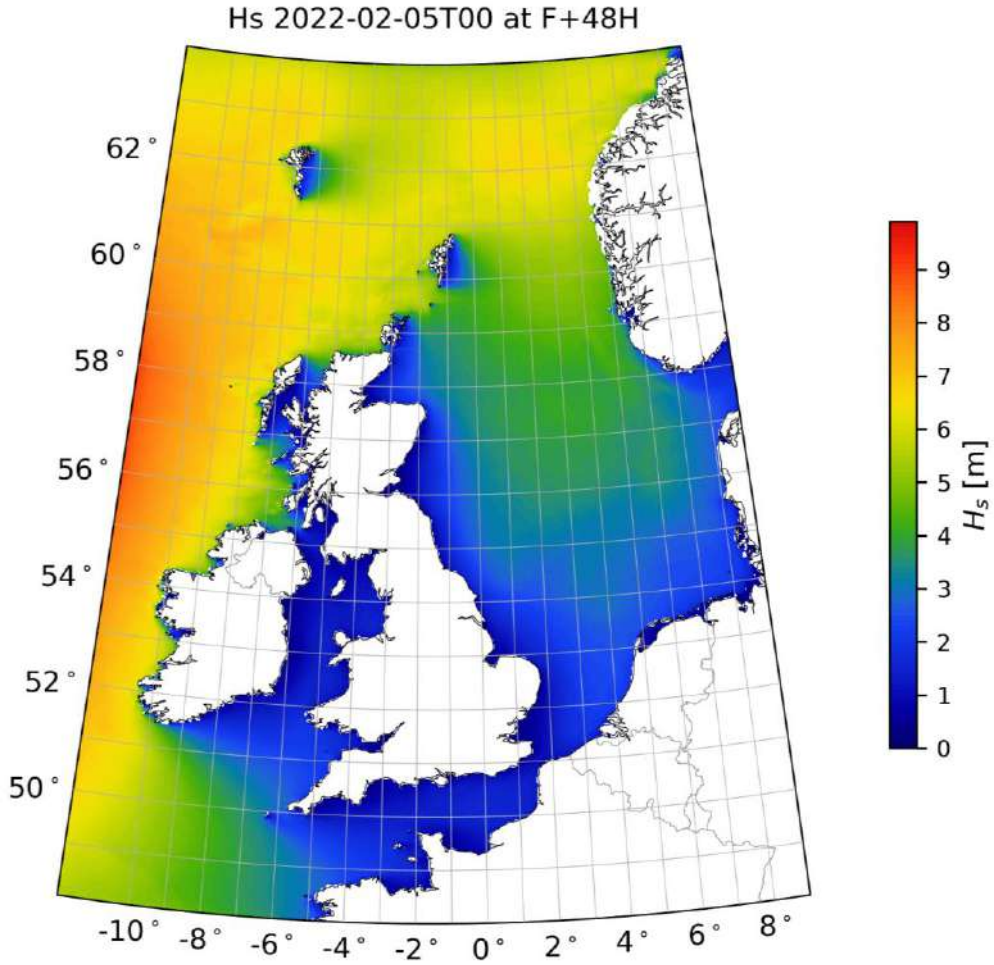


Use case: speed up training TurboSWAN

- SWAN-North Sea for operational wave forecasts







weer
klimaat
water

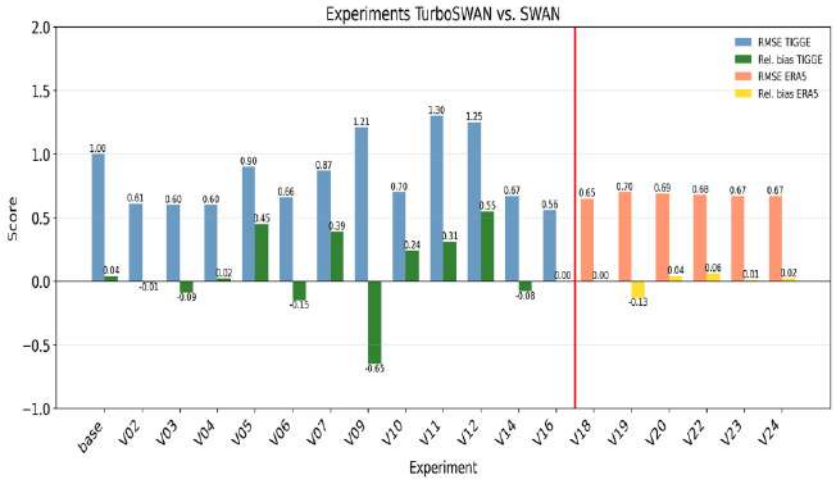
Use case: speed up training TurboSWAN

Trainingsdata: SWAN-DCSM 1979-2022 model results

incl ERA5 wind fields, water levels, boundary conditions.

Set up the fast surrogate model: TurboSWAN

- Time lag in wind?
- Waterdepth or water level?
- Loss function relative or absolute, …
- …



runid			input								
	dataset	Lag [hr]	Wave B.C.	Wind	Lev	Flow	Augm	Architecture	Loss	Mask	Output
2024	TIGGE*	0, -6	yes	U,V	WL	yes	-	0	sMSE	-	5 pars
Base	TIGGE	0, -6	yes	U,V	WL	yes	-	0	sMSE	-	5 pars
V02	TIGGE	0, -6	yes	U,V	dep	yes	-	1	sMSE	-	5 pars
V03	TIGGE	0, -6	yes	U,V	dep	yes	-	2	sMSE	-	5 pars
V04	TIGGE	0, -6	yes	U,V	dep	yes	-	3	sMSE	-	5 pars
V05	TIGGE	0, -6	yes	U,V	dep	yes	-	4	sMSE	-	5 pars
V06	TIGGE	0	yes	U,V	dep	yes	-	4	sMSE	-	5 pars
V07	TIGGE	0, -12	yes	U,V	dep	yes	-	4	sMSE	-	5 pars
V09	TIGGE	0, -6	yes	U,V	dep	yes	-	3	sMSE	yes	5 pars
V10	TIGGE	0	yes	U,V	dep	yes	-	3	sMSE	-	5 pars
V11	TIGGE	0, -6	-	τ_x, τ_y	dep	-	-	3	sMSE	-	Hs
V12	TIGGE	0, -6	-	τ_x, τ_y	dep	-	-	3	sMSE	yes	Hs
V14	TIGGE	0, -6	-	τ_x, τ_y	dep	-	yes	3	sMSE	yes	Hs
V16	TIGGE	0	-	U,V	dep	-	-	3	sMSE	-	Hs
V18	18yrERA5	0	yes	U,V	dep	-	-	3	sMSE	-	Hs
V19	18yrERA5	0	yes	U,V	dep	-	-	3	rMAE	-	Hs
V20	18yrERA5	0	yes	U,V	dep	-	-	3	uMSE	-	Hs
V22	18yrERA5	0, -3	yes	U,V	dep	-	-	3	uMSE	-	Hs
V23	18yrERA5	0, -6	yes	U,V	dep	-	-	3	uMSE	-	Hs
V24	18yrERA5	0, -3 -6 -9	yes	U,V	dep	-	-	3	uMSE	-	Hs

iML weer klimaat water

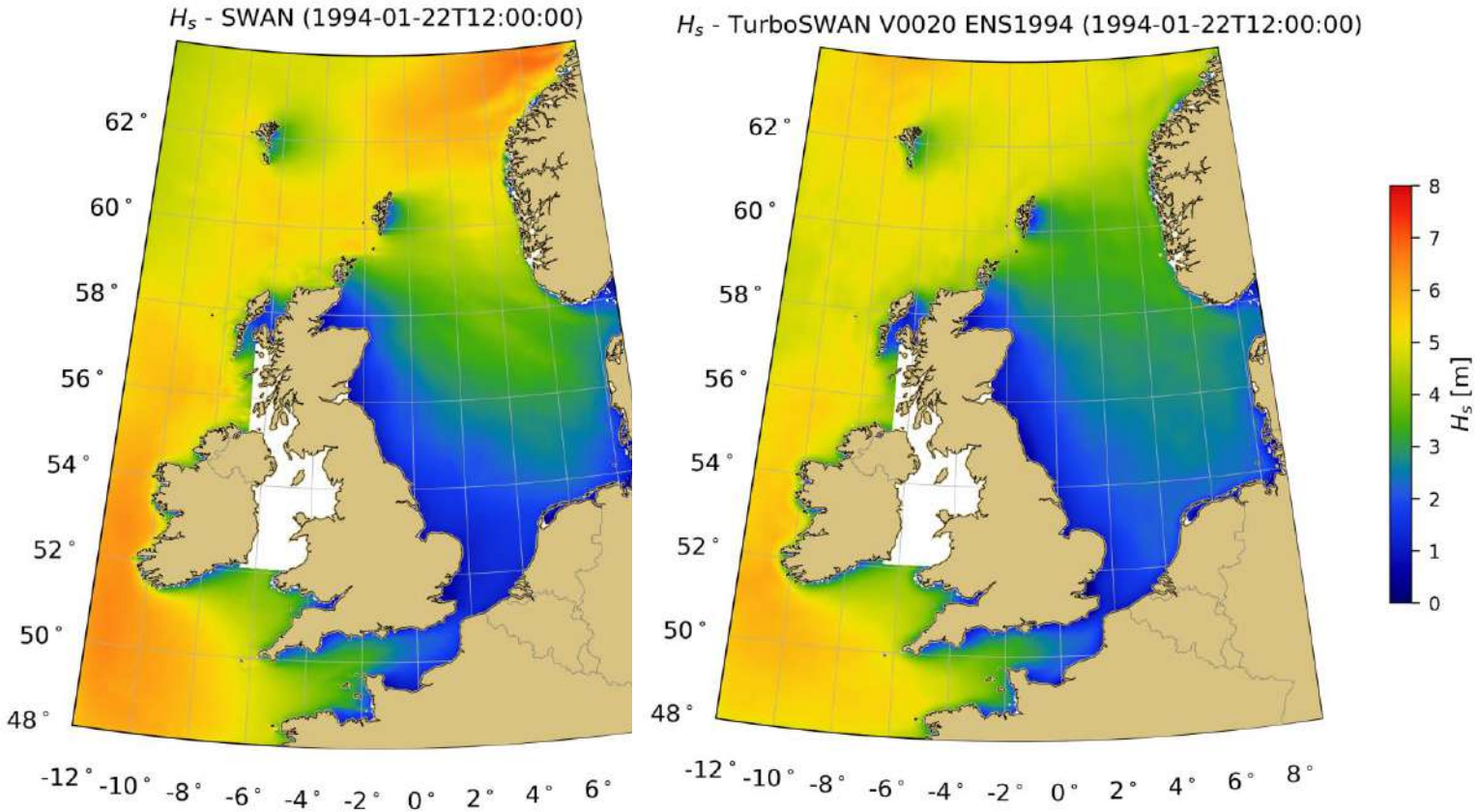
Use case: speed up training TurboSWAN

How can we optimise this process?

- Faster
- Better structured
- Automated

100 iterations…??

-> Let ‘s ask iML-WKW!



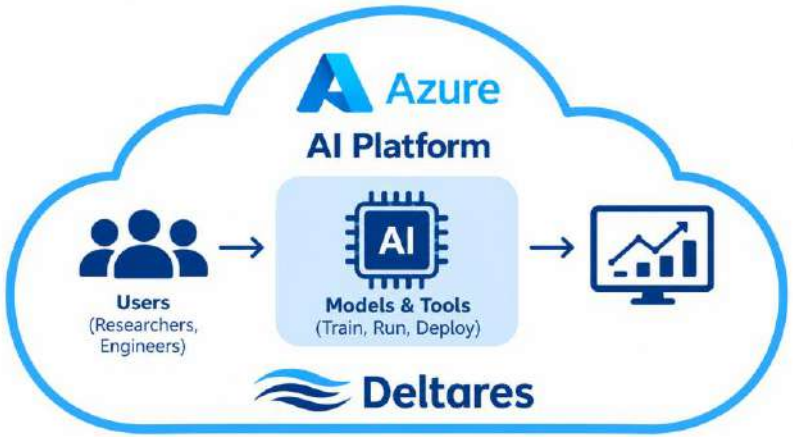
iML weer klimaat water

Perfect timing

- The MLX team had just expanded
- MLX needed a new use case at Deltares, to help onboard the new team-members
- Deltares was quite far with building an Proof of Concept AI platform, which we could use

MLX
MACHINE LEARNING EXCHANGE
MAKING ML WORK IN PRACTICE

USE case



iML weer klimaat water

How can we help?

To help the Turboswan use case we needed:

- Orchestration
- More compute
- Parallelisation
- Some more compute



The AI platform already had orchestration

Compute on Azure would be expensive (€ 100k+ for the estimated 70.000 GPUhours), so we had to use Surf (Snellius)

New features:

- Facilitate parallelisation
- Optimize runs
- Optimize GPU usage

- (Integrate the AI Platform with Snellius)

iMLweer klimaat water

Overview

AIWaM and iML-WKW

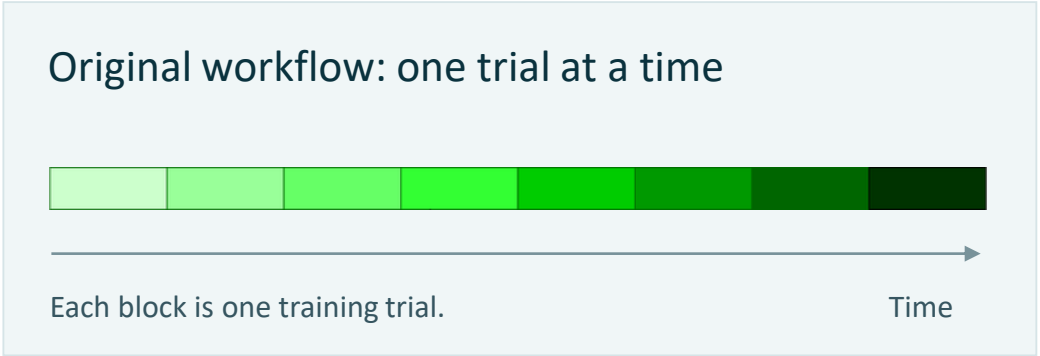
- Make experiments faster: run many trials at once instead of one at a time.
- Make results accessible: record data, settings and scores in one place.
- Make training cheaper: on Snellius, the optimised run process 78% more samples per second.



iML weer klimaat water

Make research faster

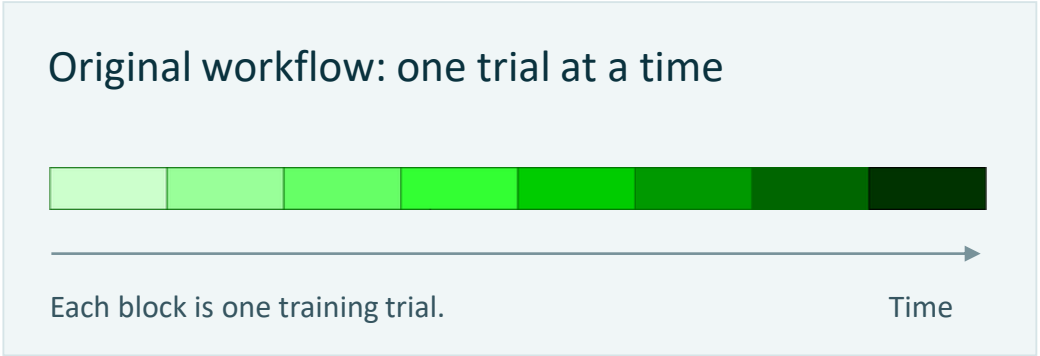
- Original: Test one set of training settings at a time.

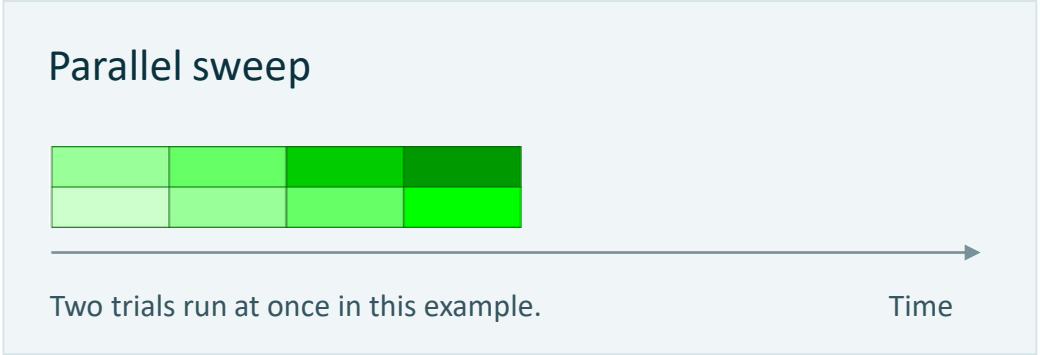


weer
klimaat
water

Make research faster

- Original: Test one set of training settings at a time.
- With a sweep, several trials can run at the same time.



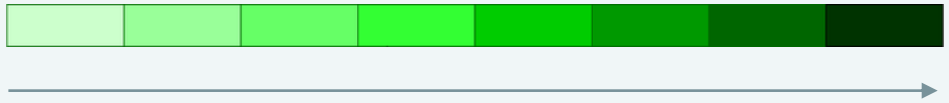


iML weer klimaat water

Make research faster

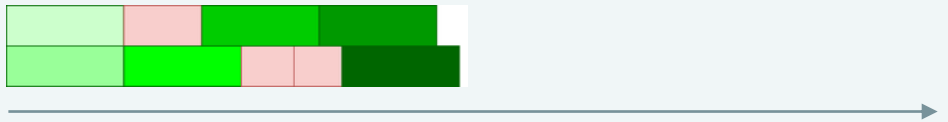
- Original: Test one set of training settings at a time.
- With a sweep, several trials can run at the same time.
- The results can guide the next trials, and trials that make little progress can stop early.

Original workflow: one trial at a time



Each block is one training trial. Time

Parallel sweep with early stopping



The pink blocks show trials that stopped early. Time

iML weer klimaat water

Make research accessible

The entire team can:

- See which data and settings each run used.

Inputs and settings in MLflow



Example training recipe

```
[training]
learning_rate = 1e-4
lr_decay_steps = 10000
lr_decay_rate = 0.9
n_epochs = 3
batch_size = 2
```

iML weer klimaat water

Make research accessible

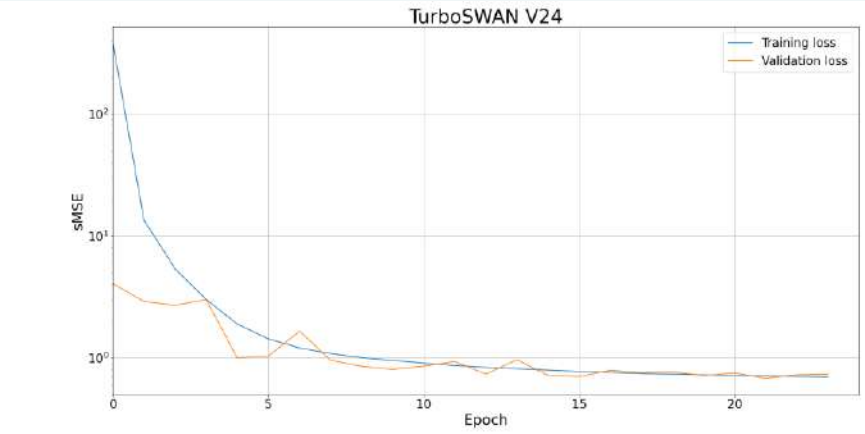
The entire team can:

- See which data and settings each run used.
- Compare training progress and scores.

Inputs and settings in MLflow



Training progress



iML weer klimaat water

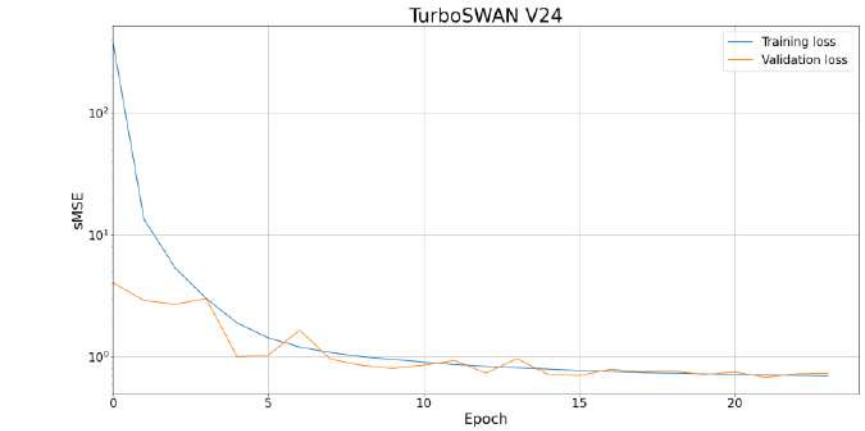
Make research accessible

The entire team can:

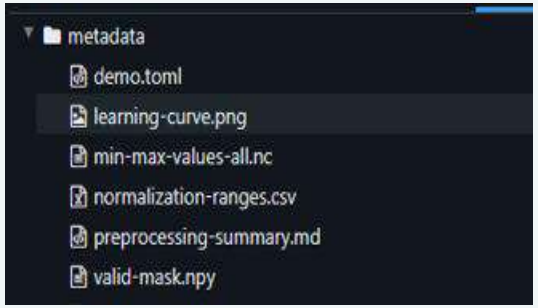
- See which data and settings each run used.
- Compare training progress and scores.

- Share a link to the trained model and its plots.

Training progress



Saved plots and files



iML weer klimaat water

Make research cheaper

- Parallel trials finish sooner but can use just as much GPU time.

Example: two one-hour trials

One after the other	Together
2 hours waiting	1 hour waiting

Both use two GPU-hours in total.

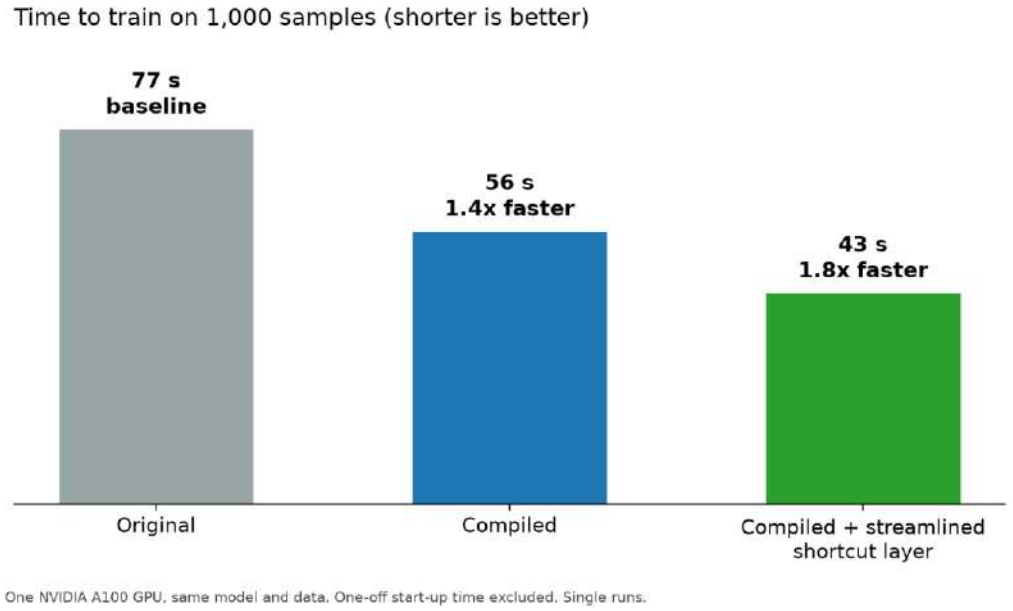
Make research cheaper

- Parallel trials finish sooner but can use just as much GPU time.
- We benchmarked TurboSWAN on Snellius GPUs.
- Grouped small GPU operations with compilation and streamlining.
- Now process 78% more training samples per second.

Example: two one-hour trials

One after the other	Together
2 hours waiting	1 hour waiting

Both use two GPU-hours in total.



iML weer klimaat water

What iML enables

- Allow you to run more experiments at the same time.
- Each optimised run takes less time and uses fewer computing resources.
- In short: enable easier, faster experimentation for researchers☺



Bring us a use case you would like help with!

iML weather climate water