Correction of NWP ocean surface wind biases with machine learning and scatterometer data

E. Makarova (BEC, ICM–CSIC) makarova@icm.csic.es,
M. Portabella (BEC, ICM–CSIC), A. Stoffelen (KNMI),
V. Vilaplana (UPC), M. García (NOW), M. Sotillo (NOW)

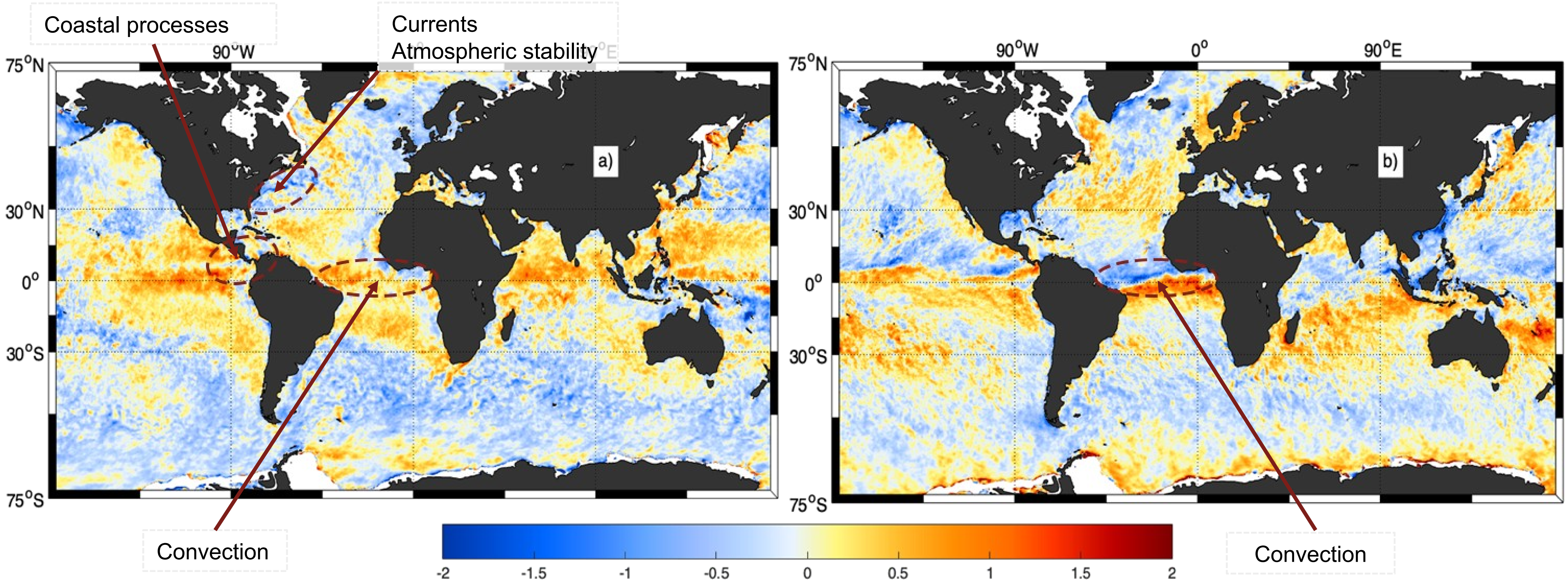
International Ocean Vector Winds Science Team Meeting
Darmstadt, Germany

8 May 2025

BEC
Barcelona Expert Center

Differences between NWP and scatterometer

Reported persistent **ERA5 U10S biases**:
Excessive mean model westerlies in the middle latitudes
Insufficient mean model poleward flow between 30° and 60°
Wind direction biased clockwise in NH and anticlockwise in SH



Portabella et al., 2022

Available corrected datasets

Current approach to reduce the biases is based on the accumulation of the differences between scatterometers and model stress-equivalent winds over a certain time window.

Dataset	**World Ocean Circulation ERA***	**Copernicus Global Ocean Hourly Reprocessed Sea Surface Wind and Stress from Scatterometer and Model**	**Copernicus Global Ocean Hourly Sea Surface Wind and Stress from Scatterometer and Model**
Time span	2010 - 2020	1 Jun 1994 to 22 Dec 2024	27 Apr 2023 to now
Spatial Resolution	0.125 degrees	0.125 and 0.25 degrees	0.125 degrees
Temporal resolution	1 hour	1 hour	1 hour
Time window	15 and 3 days	20 and 90 days	20 days
Corrected dataset	ERA5	ERA5	ECMWF operational forecasts
Scatterometers	Metop-A, Metop-B, Metop-C ASCAT, OceanSat-2 OSCAT, ScatSat-1 OSCAT	Metop-A, Metop-B, Metop-C ASCAT, QuikSCAT SeaWinds, ERS-1 and ERS-2 SCAT	Metop-B and Metop-C ASCAT
DOI	https://doi.org/10.20350/digitalCSIC/15436	https://doi.org/10.48670/moi-00185	https://doi.org/10.48670/moi-00305

ERA*

Current work on ERA* derived from WOC project:

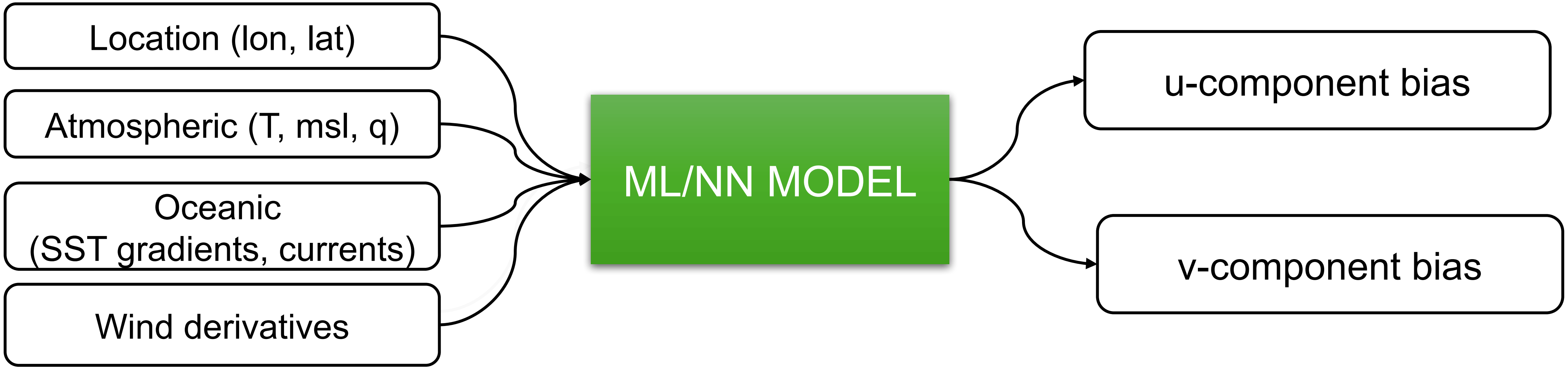
- New version of code solves the issue of gradient artifacts due to the interpolation
- Latest code will be used to reprocess the 2010 – 2020
- Generate outputs for 2021 - 2024

Limitations:

- Only corrects local biases persistent over several days
- Very sensitive to scatterometer sampling, especially over shorter time windows
- Doesn’t directly show NWP error dependence on both atmospheric and ocean state conditions
- Has limitations in operational use: computationally expensive and need to shift temporal window (which in turn degrades performance)

ERA* ML approach

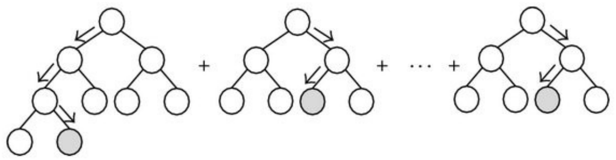
Regression ML model to predict NWP stress-equivalent wind (U10S) biases



ML Models

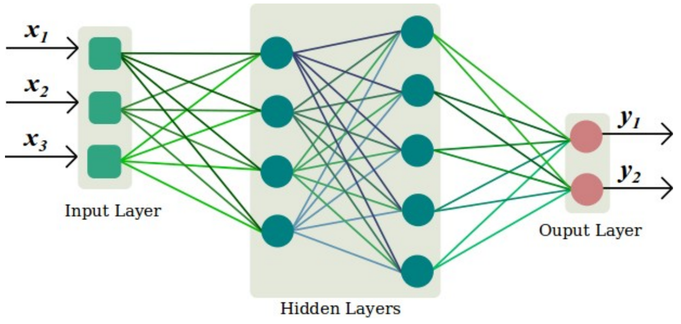
Gradient-boosted decision trees (GBDT) – fast training

XGBoost



Fully-connected feed forward neural networks

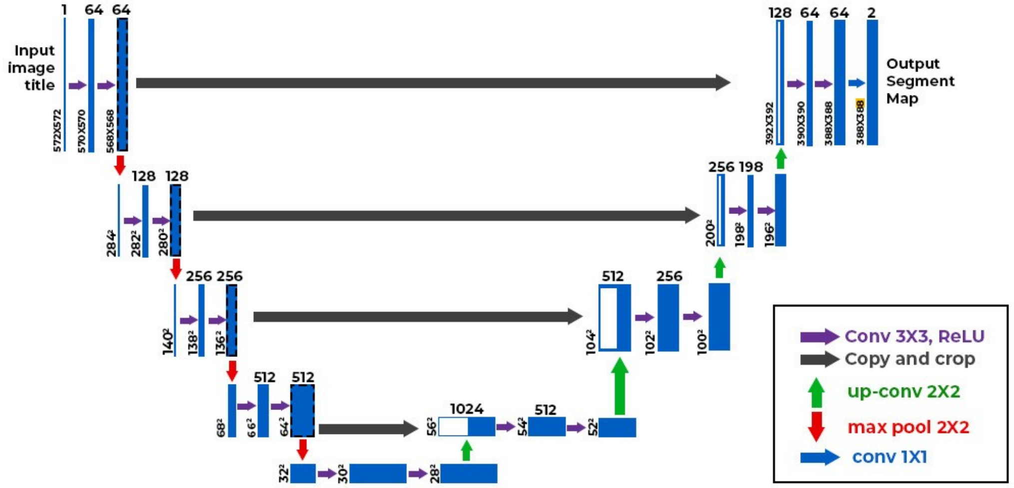
PyTorch



Spatial convolutions:

U-Net with padding

Inputs and outputs on 0.125° regular grid



ML Models

ML model	Gradient-boosted decision trees (GBDT)	Fully-connected feed forward neural networks	Convolutional NN
Implementation	XGBoost	PyTorch	U-Net (PyTorch)
Convolutions	No	No	Spatial
Data inputs grid	L2	L2	L3
Down-sampling	10%	10%	No
Regularization	Several parameters	Dropout	Weight decay
Advantages	Very fast training compared to NNs, built-in feature importance	Less overfit compared to XGBoost	Best VRMS metrics vs ASCAT
Disadvantages	More prone to overfitting	Slow training, manual calculation of spatial derivatives	Drop of spatial variance

Datasets

Inputs:

ERA5 reanalysis:

- U10S components, wind speed and direction
- Mean sea level pressure
- Air temperature
- Specific humidity
- SST, SST gradients

Currents: Global Total (COPERNICUS-GLOBCURRENT), Ekman and Geostrophic currents at the Surface and 15m

Daily mean surface velocities components (uo, vo)

Derivatives: ERA5 wind curl, divergence

Targets:

Differences between ASCAT-A and ERA5 U10S:

OSI SAF ASCAT-A 12.5 km U10S data

Periods:

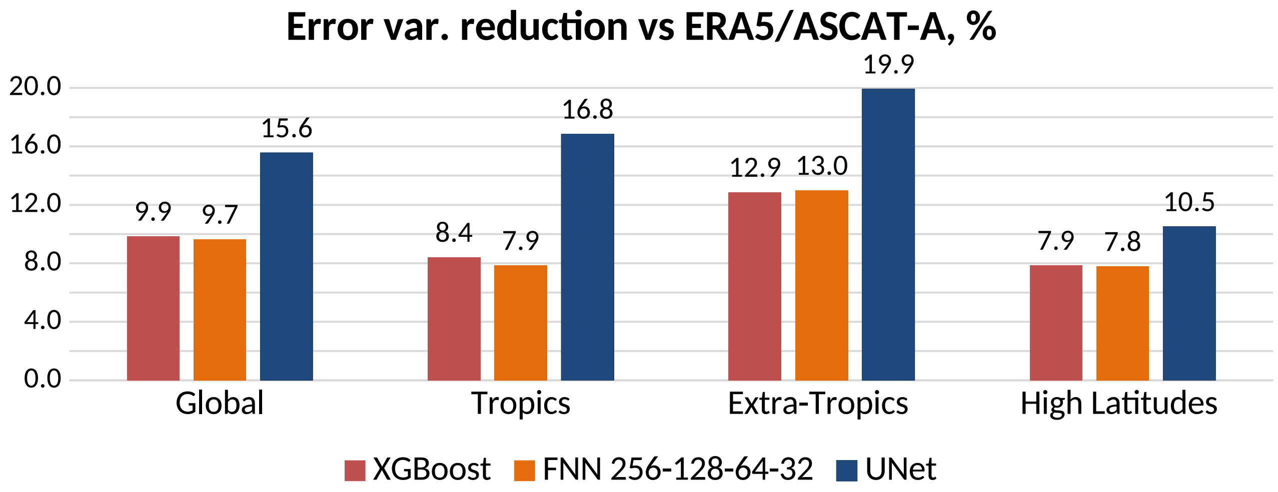
- 02/01/2020 – 01/05/2020

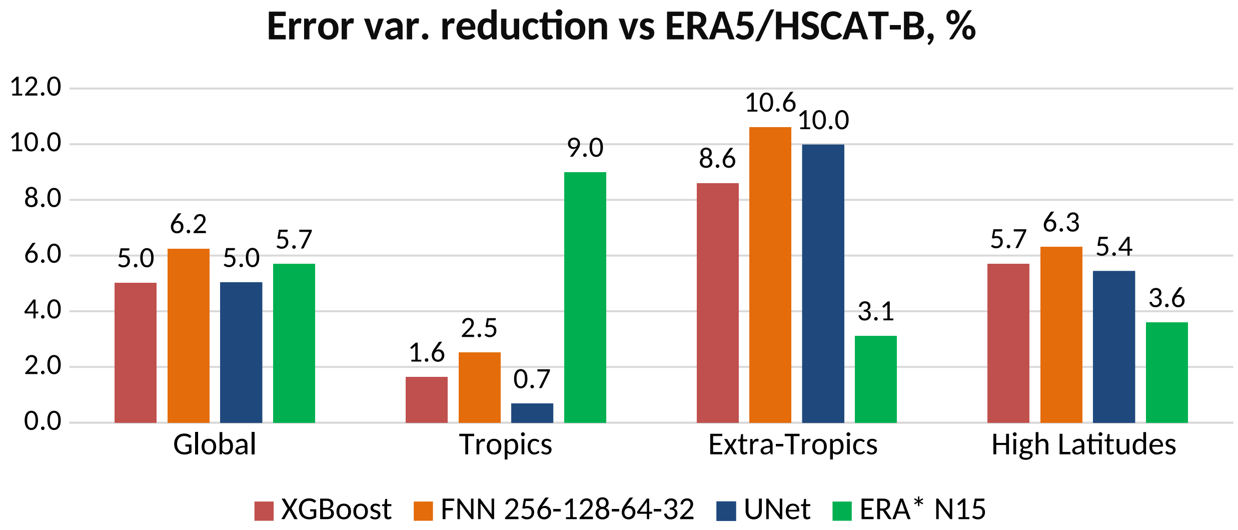
(split into train/validation/test)

- 01/02/2019 – 30/04/2019 (test)

Validation: 2019

01/02 – 30/04 2019



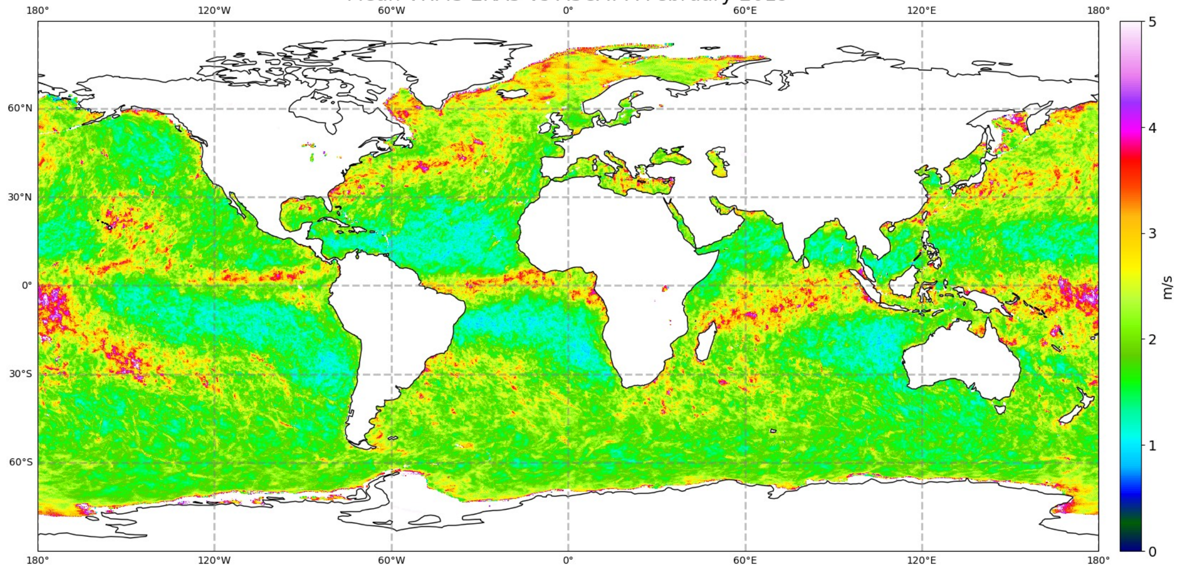


VRMS ERA5	ASCAT-A	HSCAT-B
Global	2.082	1.647
Tropics	2.056	1.602
Extra-Tropics	2.047	1.622
High Lats	2.2	1.794

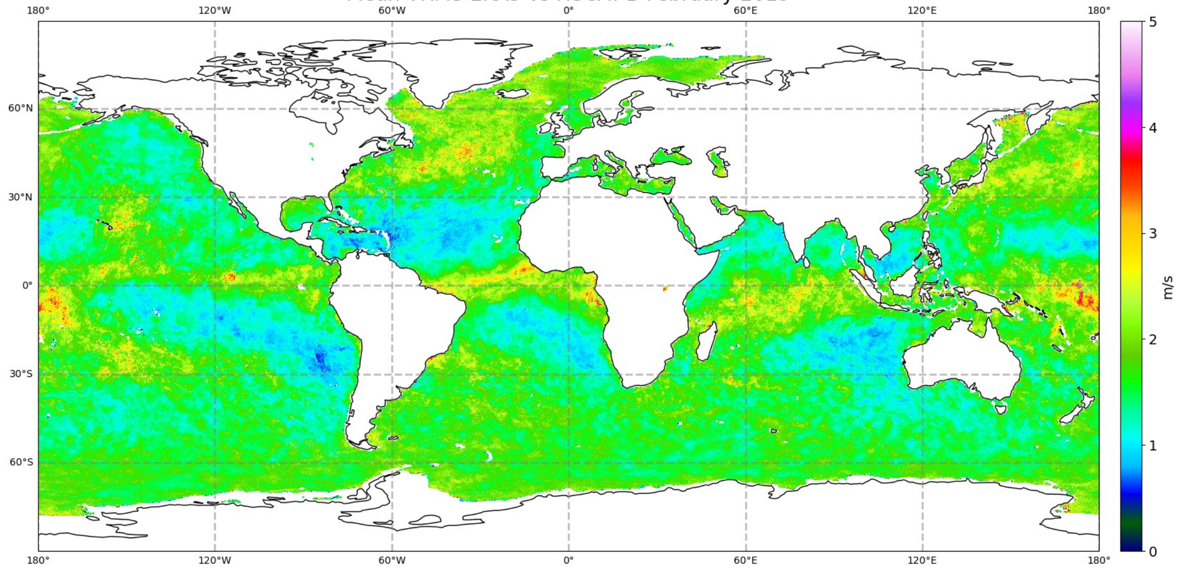
VRMS vs ASCAT-A & HSCAT-B

BEC
Barcelona Expert Center

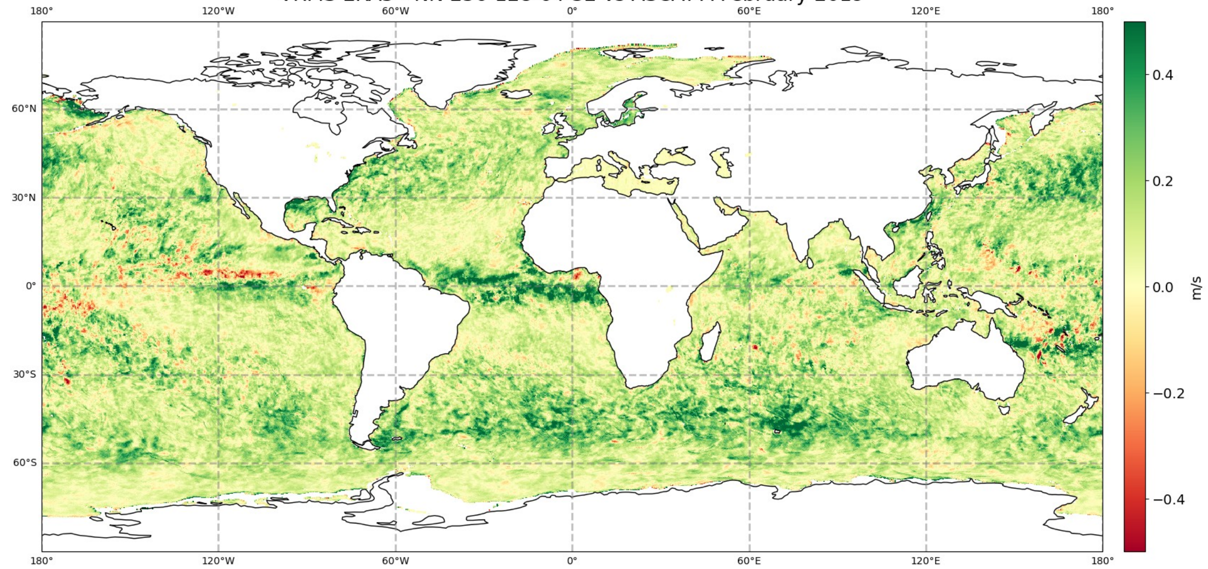
Mean VRMS ERA5 vs ASCAT-A February 2019



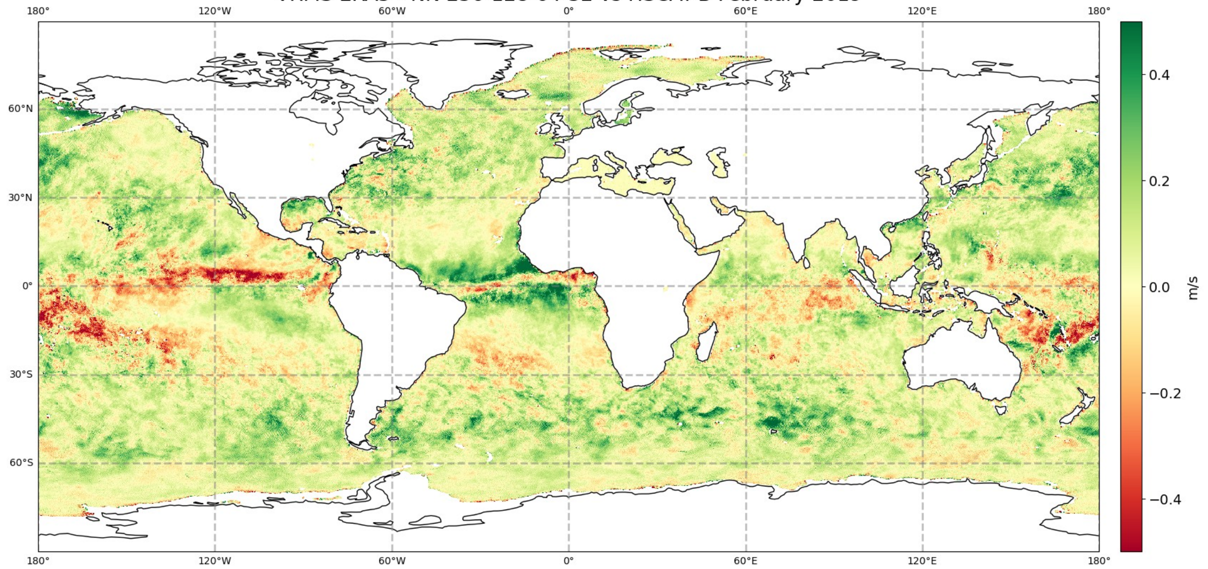
Mean VRMS ERA5 vs HSCAT-B February 2019



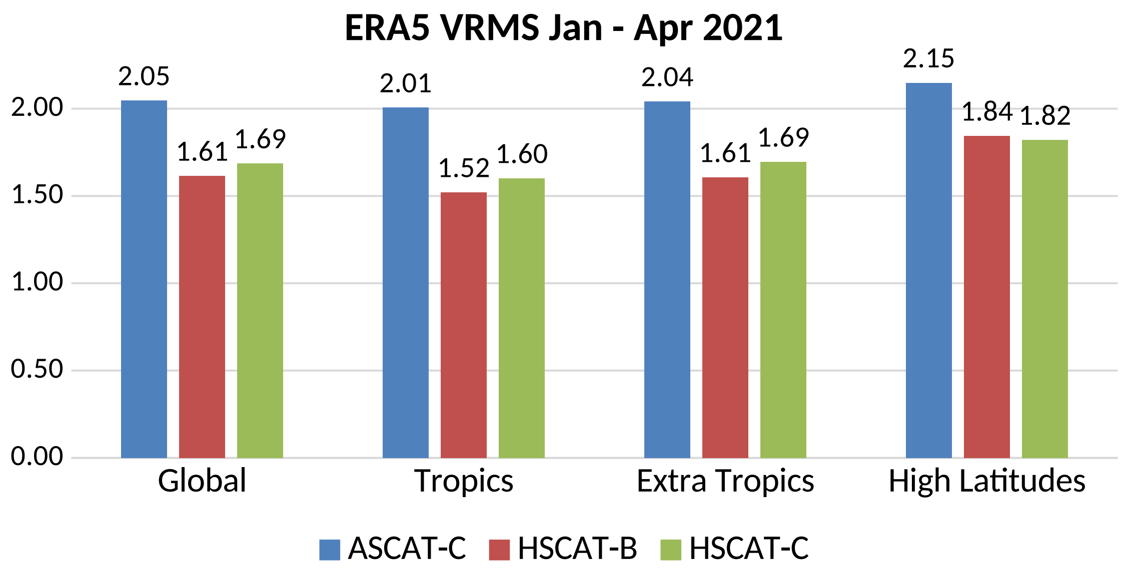
VRMS ERA5 - NN 256-128-64-32 vs ASCAT-A February 2019

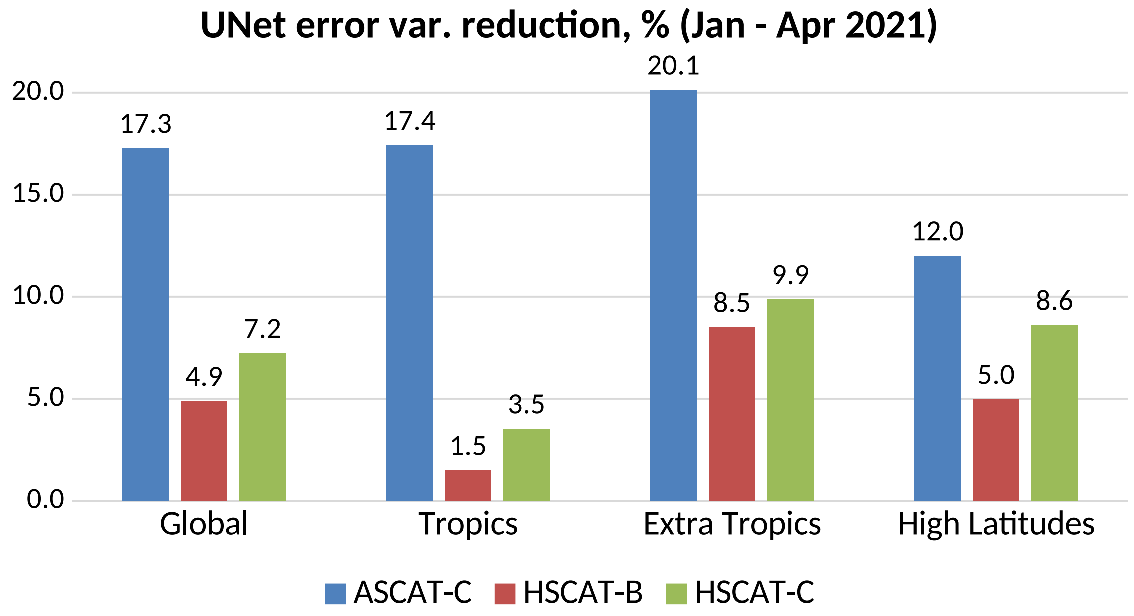


VRMS ERA5 - NN 256-128-64-32 vs HSCAT-B February 2019

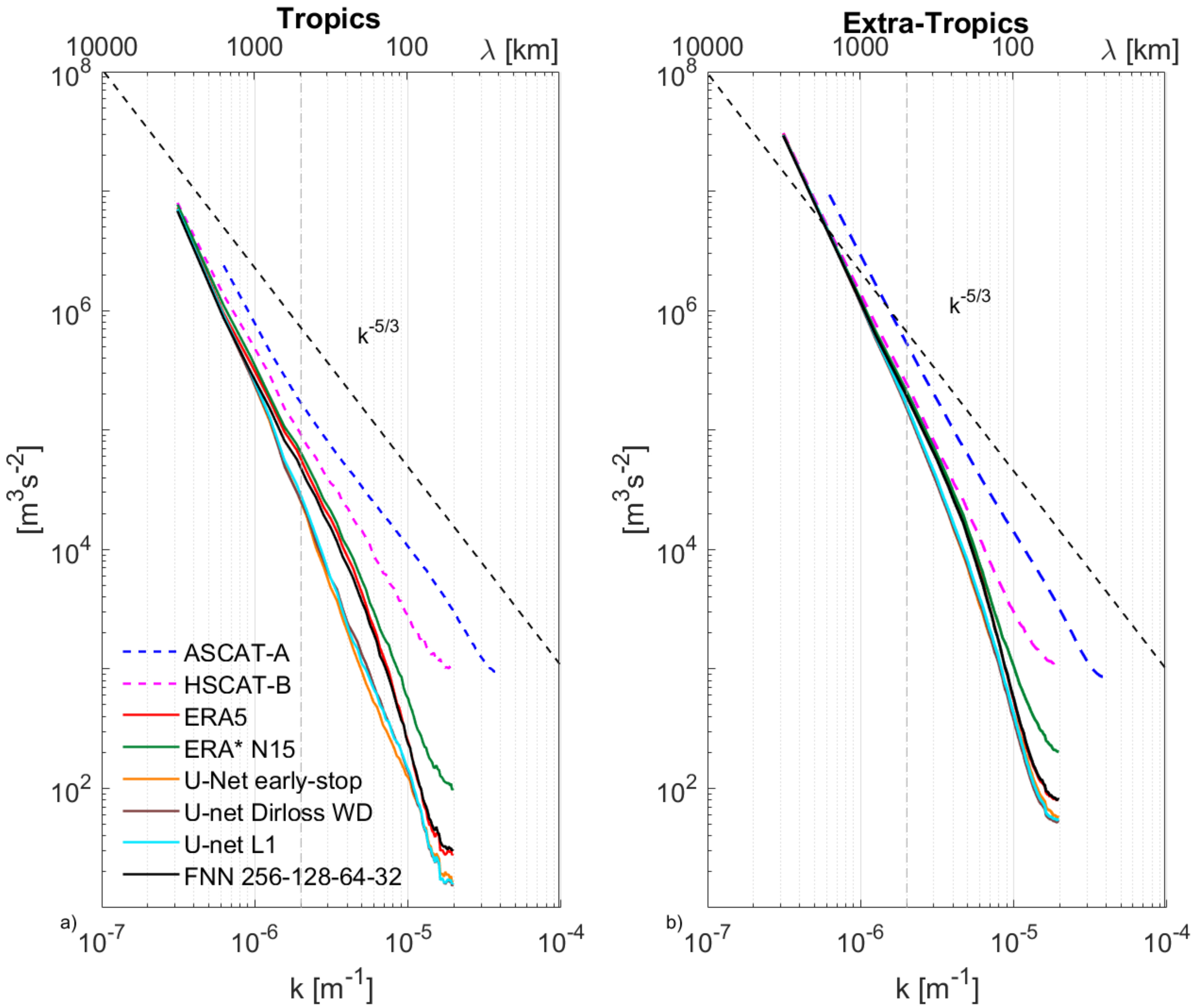


U-Net Validation ASCAT-C/HSCAT-B/HSCAT-C





Spectral power density



v-component
February 2019
Collocations vs HSCAT-B

FNN & XGBoost:

- Lower variance at 500-km scale in tropics compared to ERA5

U-Net:

- Important drop of spatial variance at scales < 1000 km

Loss function?

Conclusions

- Preliminary ML models show that it is possible to predict ERA5 NWP biases using other NWP variables as input.
- With FNN models 6.3% error variance reduction achieved globally and 10.6% in the extra-tropics (vs HSCAT-B, February – April 2019).
- Reduction is larger when validated against ASCAT (up to 9.7% globally and 13% in extra-tropics for FNNs) than against HSCAT. Diurnal effects? QC effects?
- Issues with the loss of spatial variance, especially for CNNs.
- Uses:
    - ➢ Improvement of the reanalysis datasets for the periods with no scatterometer observations
    - ➢ Improvement of the operational forecasts if trained with IFS

Future Work

- Train FNN models on 5 years of data (OSI_VSA24_01 ERAS_ML project).
- Try other loss functions for CNN approach to reduce loss of spatial variance.
- Explore architectures based on GANs and transformers.
- Analyse diurnal & QC effects by ASCAT-HSCATs collocation analysis
- CERAINE Copernicus Marine service evolution project:
  - Improve the North-East Atlantic wave and physics model solutions (IBI-MFC short-term forecast services) by correcting their operational forcing data with ANNs
- Improved near-surface temperature representation (ESA COMET)
- Improved SMOS-derived salinity retrievals (EO4TIP)

EUMETSAT
OSI SAF
OCEAN AND SEA ICE

COPERNICUS MARINE SERVICE
SERVICE EVOLUTION
25 months until
September 2026
Contract 23138-COP-INNO SCI

esa

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES

AGENCIA
ESTATAL DE
INVESTIGACIÓN

Contacts:

Eugenia Makarova
makarova@icm.csic.es

Marcos Portabella
portabella@icm.csic.es

International Ocean Vector Winds Science Team Meeting
Darmstadt, Germany

5 May 2025

BEC
Barcelona Expert Center