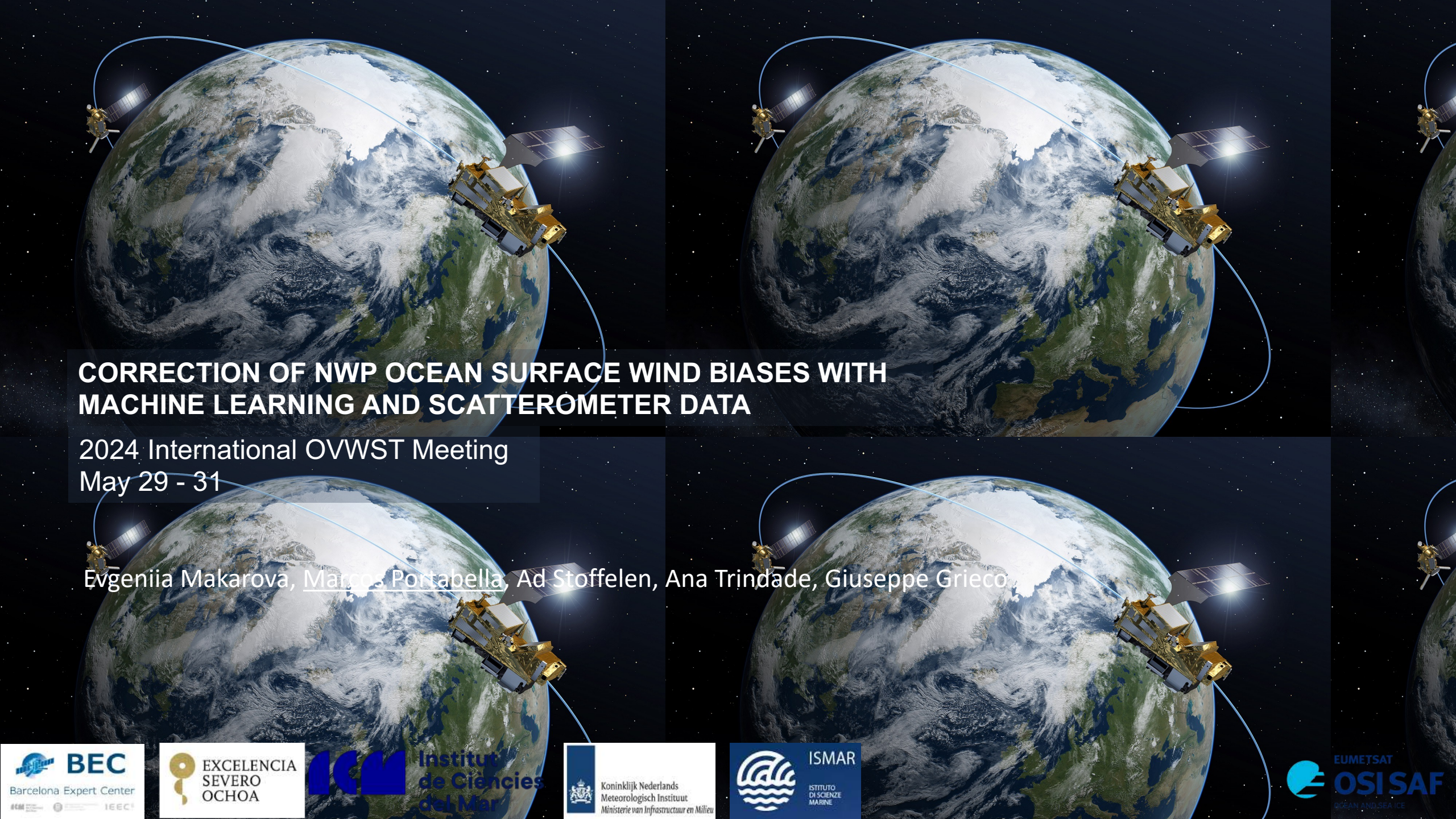
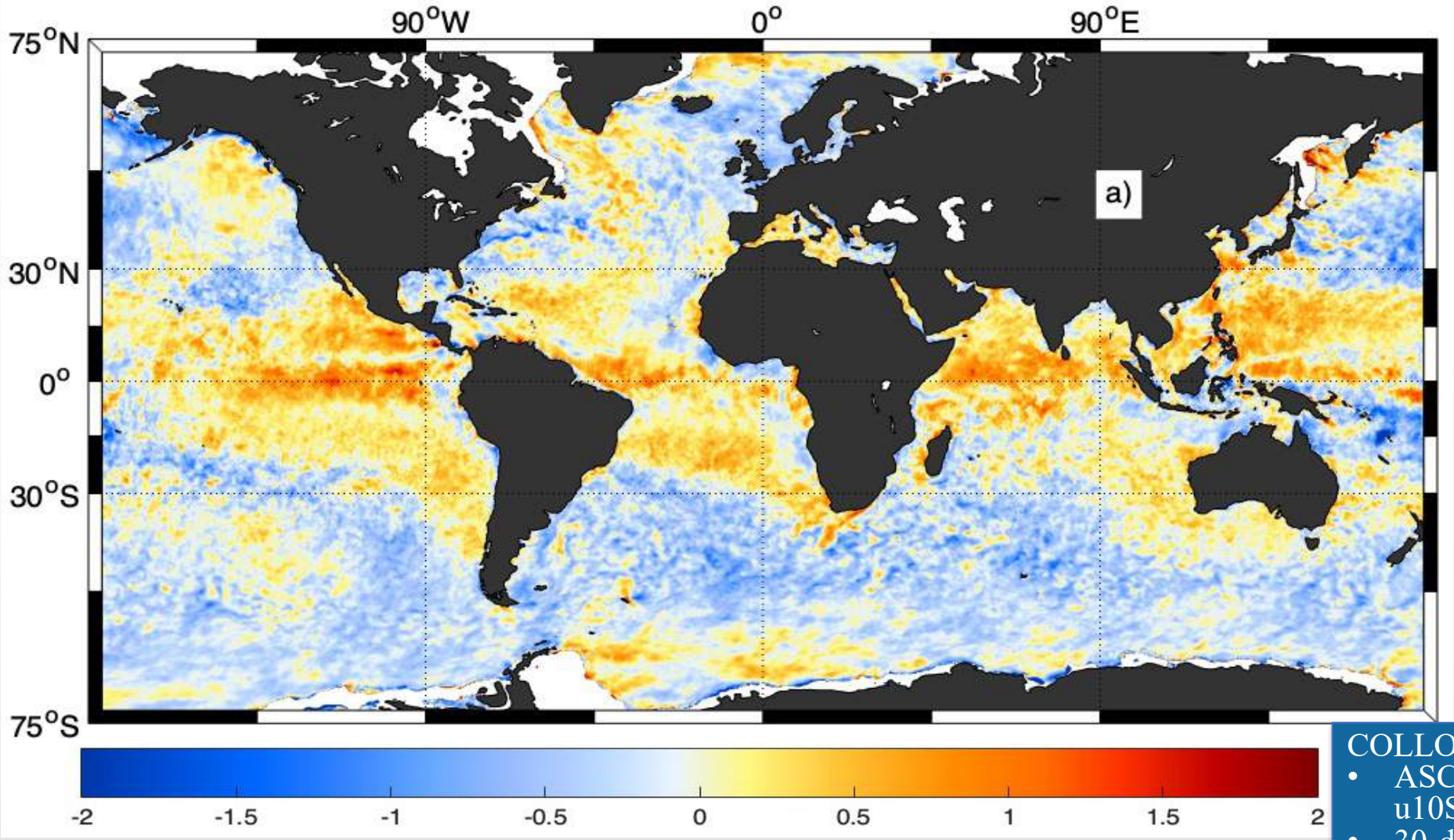


CORRECTION OF NWP OCEAN SURFACE WIND BIASES WITH MACHINE LEARNING AND SCATTEROMETER DATA

2024 International OVWST Meeting
May 29 - 31

Evgeniia Makarova, Marcos Portabella, Ad Stoffelen, Ana Trindade, Giuseppe Grieco

Systematic differences between NWP and scatterometer



COLLOCATIONS
- ASCAT-A/B/C – ERA5 u10S
- 30-d Temporal Window

Zonal component (v10S)

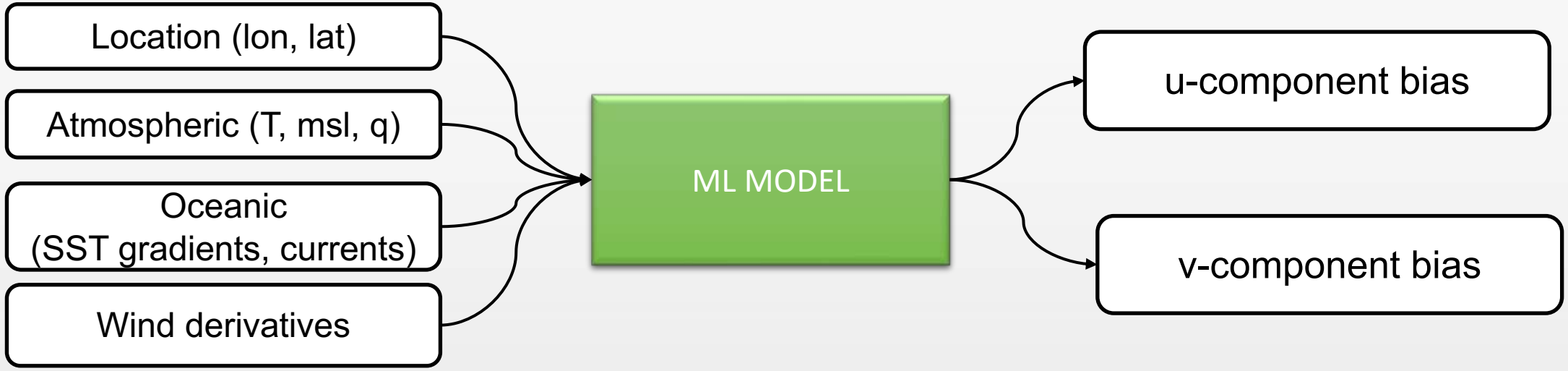
Previous approach: ERA*

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)

Objectives

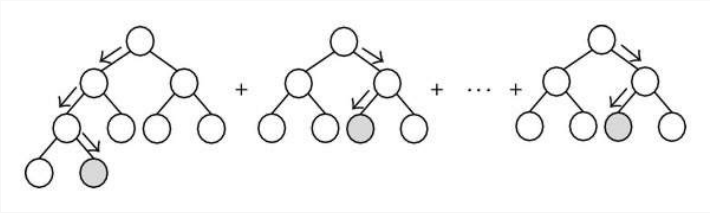
- Create a preliminary ML model to predict NWP stress-equivalent wind (U10S) biases
- Fit a regression that finds functional relationship between several NWP parameters and U10S biases



- Check the viability of the approach and compare performance of models based on different ML algorithms and libraries
- Train and validate models over a relatively small subset of data

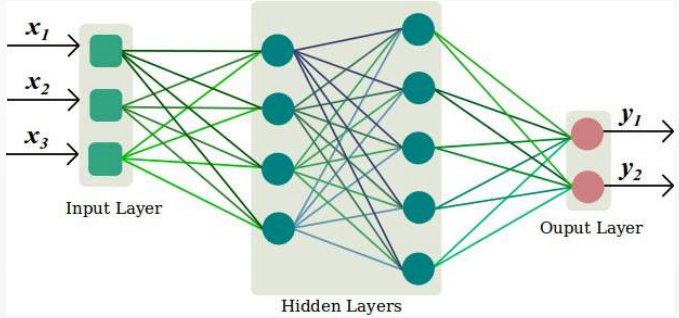
ML Models

Gradient-boosted decision trees (GBDT)



XGBoost

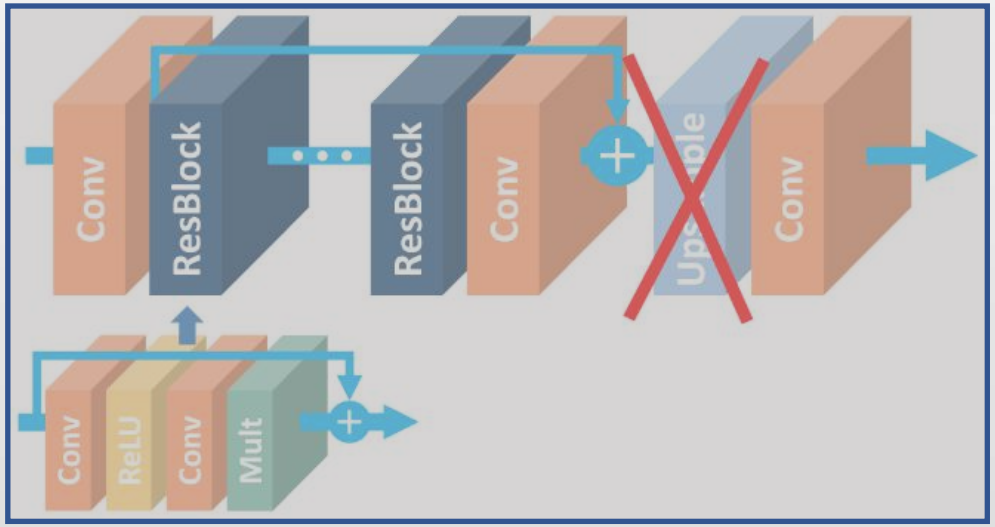
Fully-connected feed forward neural networks



PyTorch (currently)

2nd stage (future development):

- Convolutional neural networks for super-resolution adapted to regression/generative networks



Modified EDSR network (arXiv:1707.02921)

Methodology: Validation

Metrics:

Vector Root Mean Square Difference

$$VRMS = \sqrt{\frac{1}{N}\sum_i \left(u_i^{scat} - u_i^{model}\right)^2 + \left(v_i^{scat} - v_i^{model}\right)^2}$$

$$relativeerrorvariancereduction(\%) = \frac{VRMS_{ERA5}^2 - VRMS_{ML}^2}{VRMS_{ERA5}^2} * 100$$

(Train dataset: 02/01 - 09/03/2020)

Test dataset: 10/03 - 01/05/2020

vs ASCAT-A:

- Complete ASCAT-A swaths, no reduction
- Same ground truth instrument as in training
- Validation at same local times

vs HSCAT-B:

- Predictions are generated for complete ERA5 forecasts for this period
- Corrected forecasts are collocated with HSCAT-B
- Local times 3.5 hours apart from training times

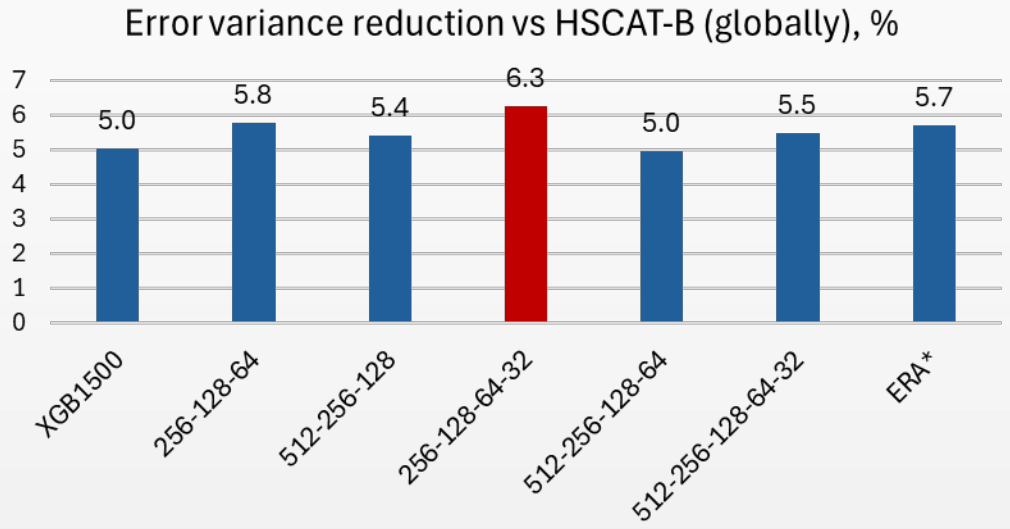
Additional validation metrics for 01/02/2019 – 30/04/2019 vs HSCAT-B & ASCAT-A

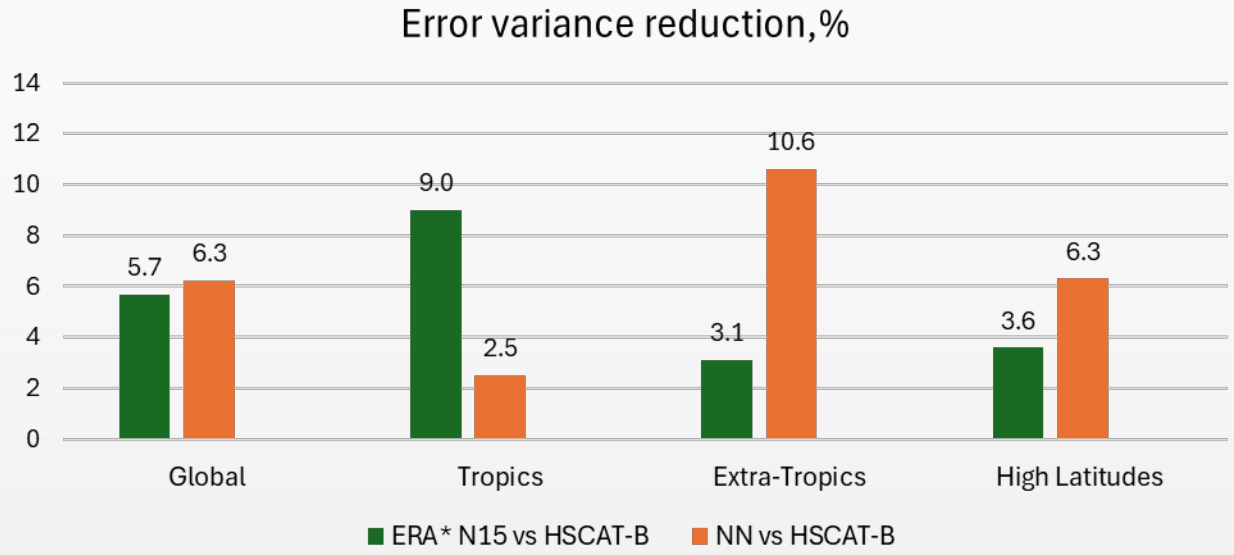
- Covers train* and test months but the period is independent

* January not included as HSCAT-B data is only available from 16/01/2019

2019 test period

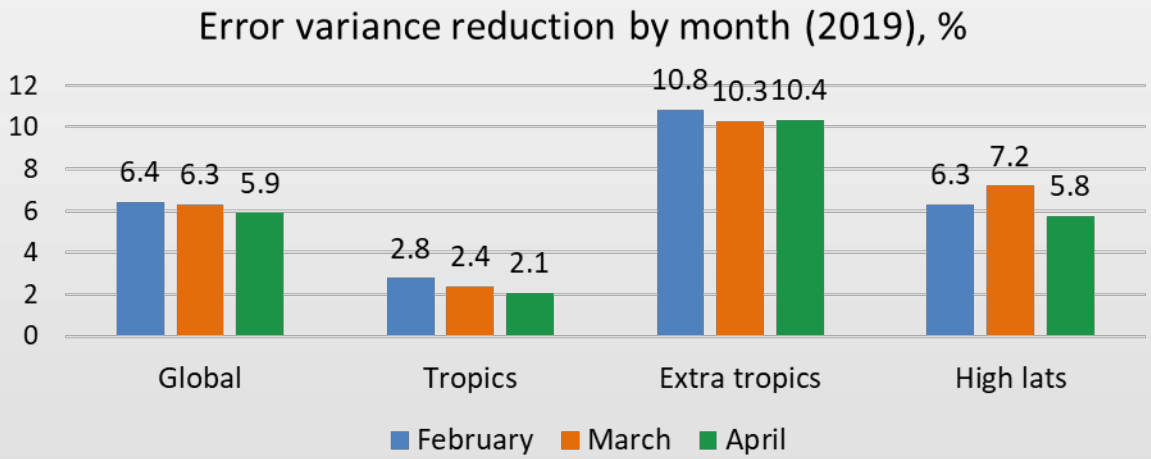
01/02 – 30/04 2019





ERA5 VRMS

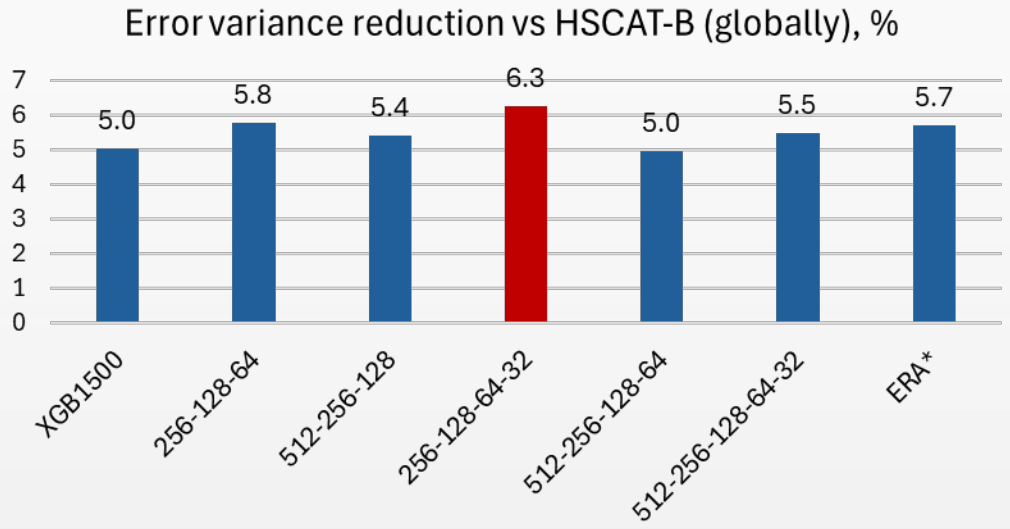
ERA5	ASCAT-A	HSCAT-B
Global	2.082	1.647
Tropics	2.056	1.602
Extra-Tropics	2.047	1.622
High Lats	2.2	1.794

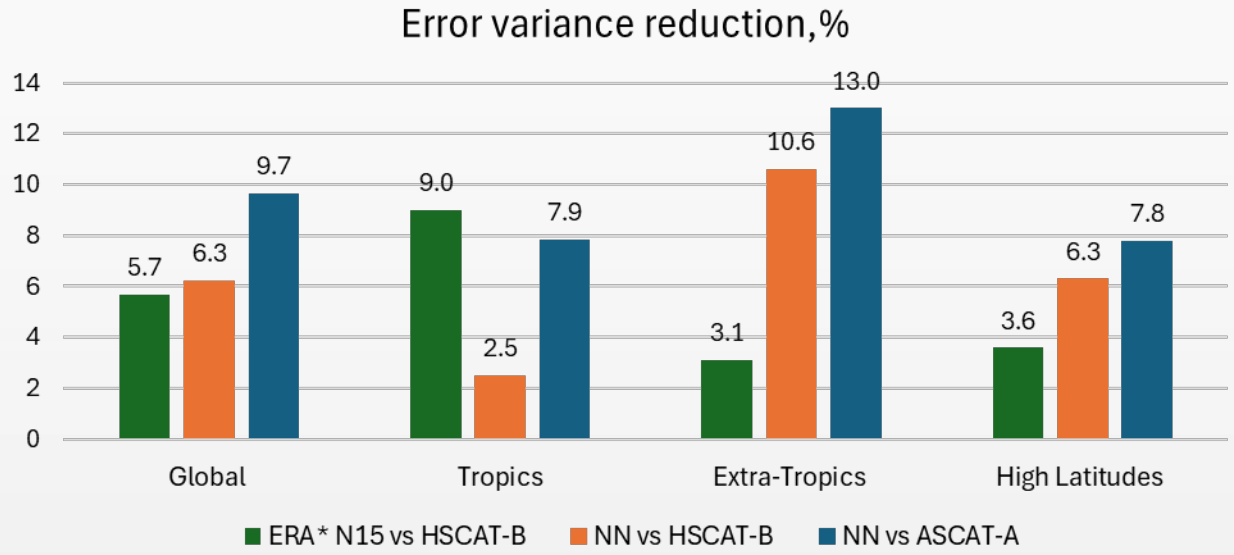


Better metrics for months seen in training but applied to another year -> seasonal models?

2019 test period

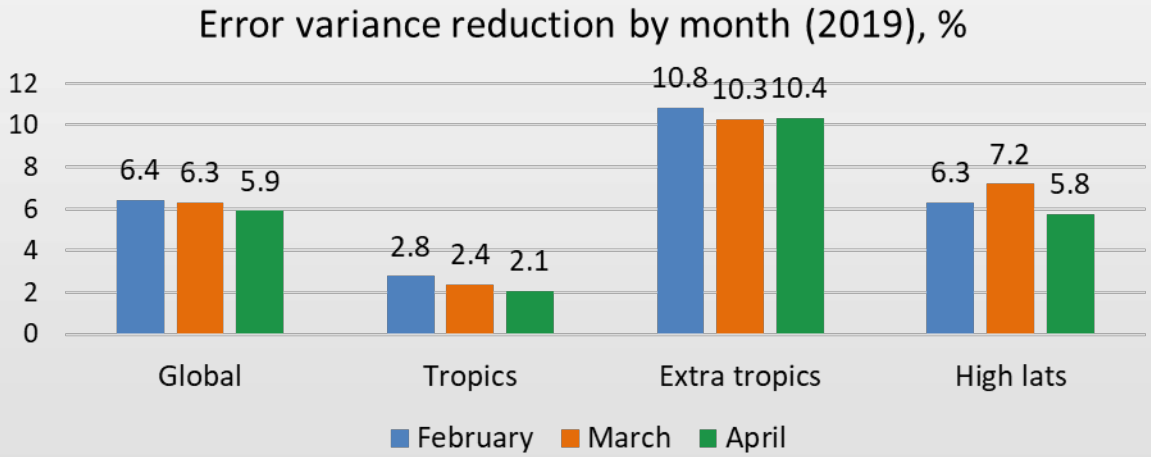
01/02 – 30/04 2019





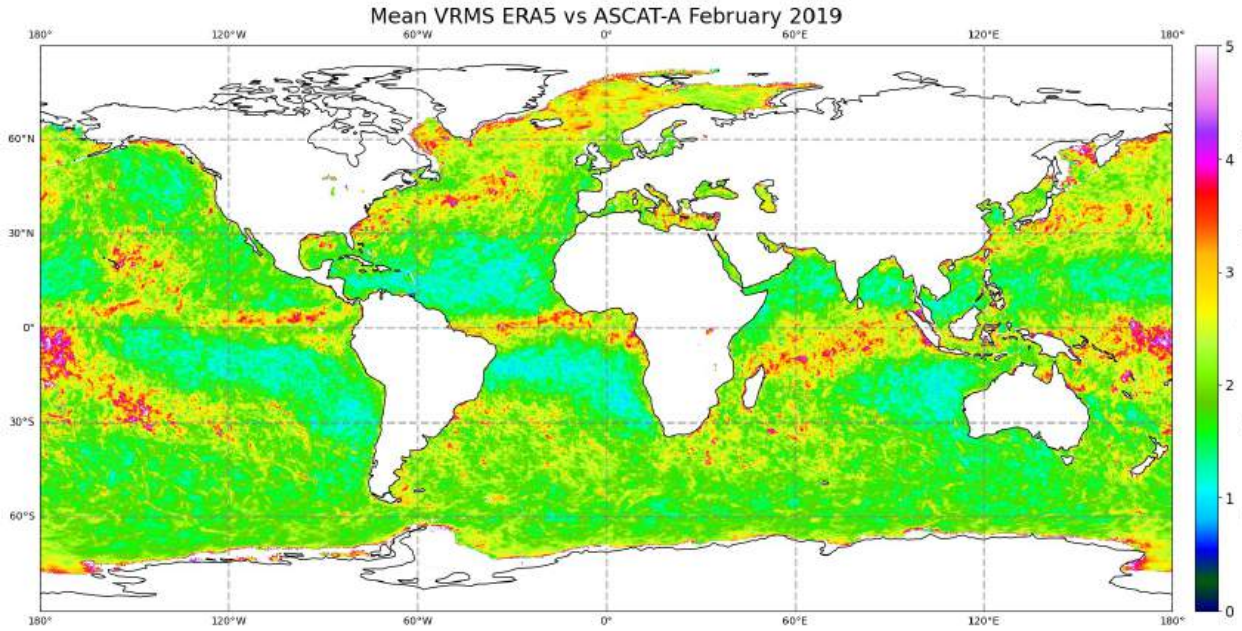
ERA5 VRMS

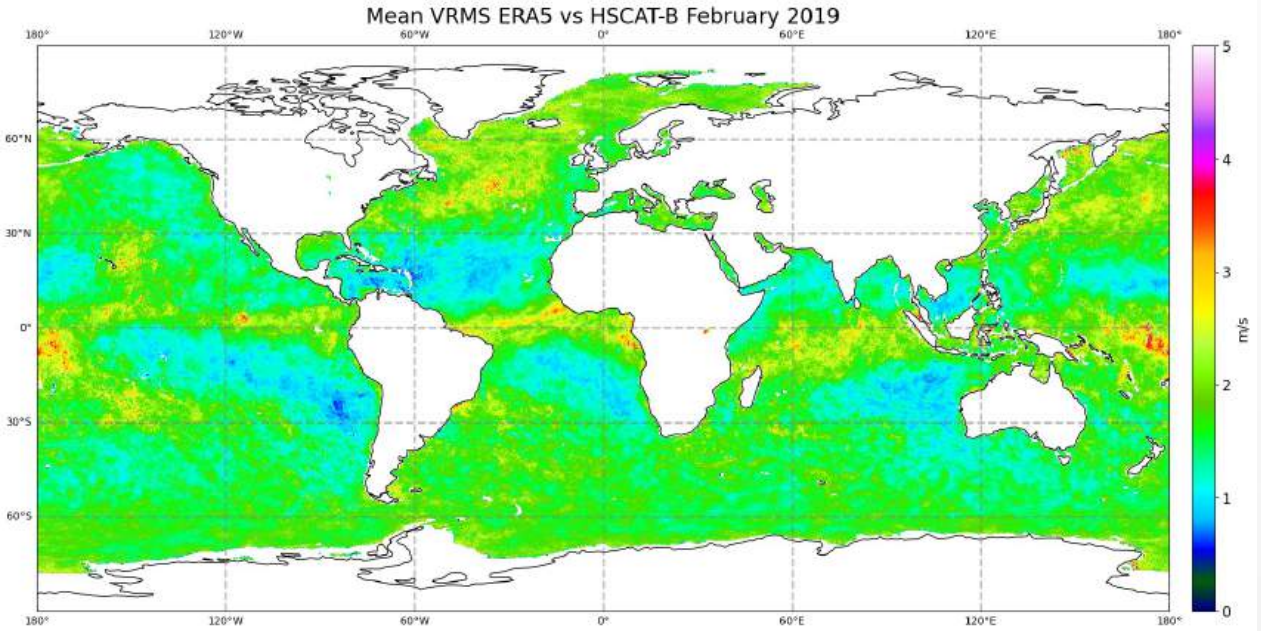
ERA5	ASCAT-A	HSCAT-B
Global	2.082	1.647
Tropics	2.056	1.602
Extra-Tropics	2.047	1.622
High Lats	2.2	1.794

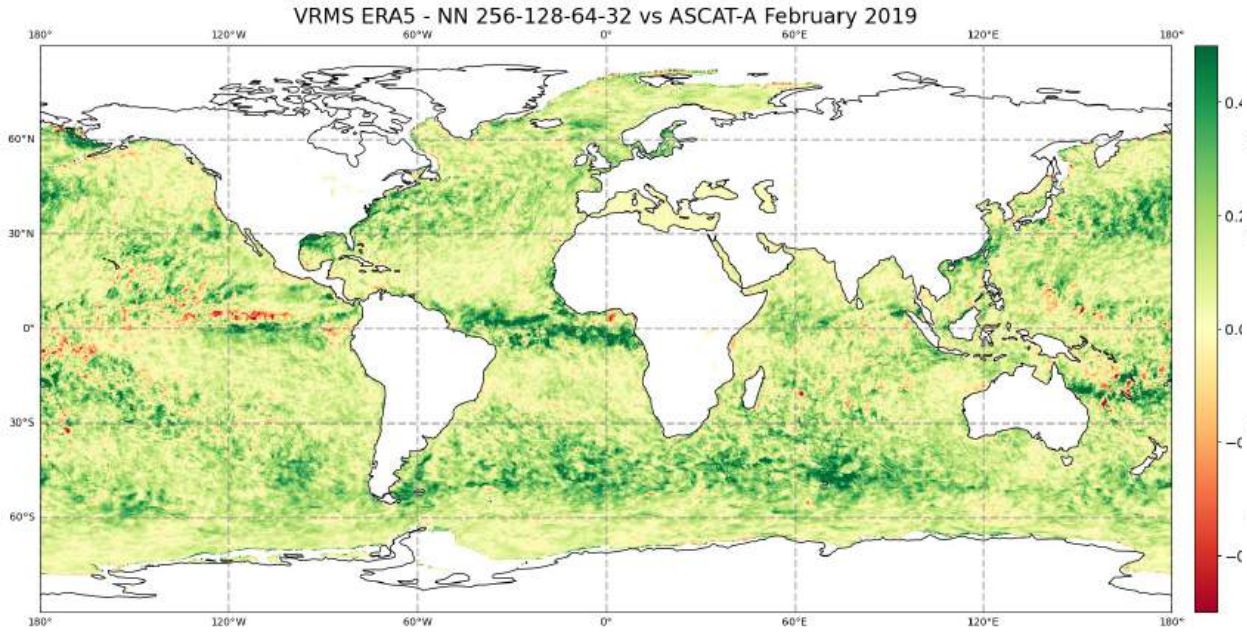


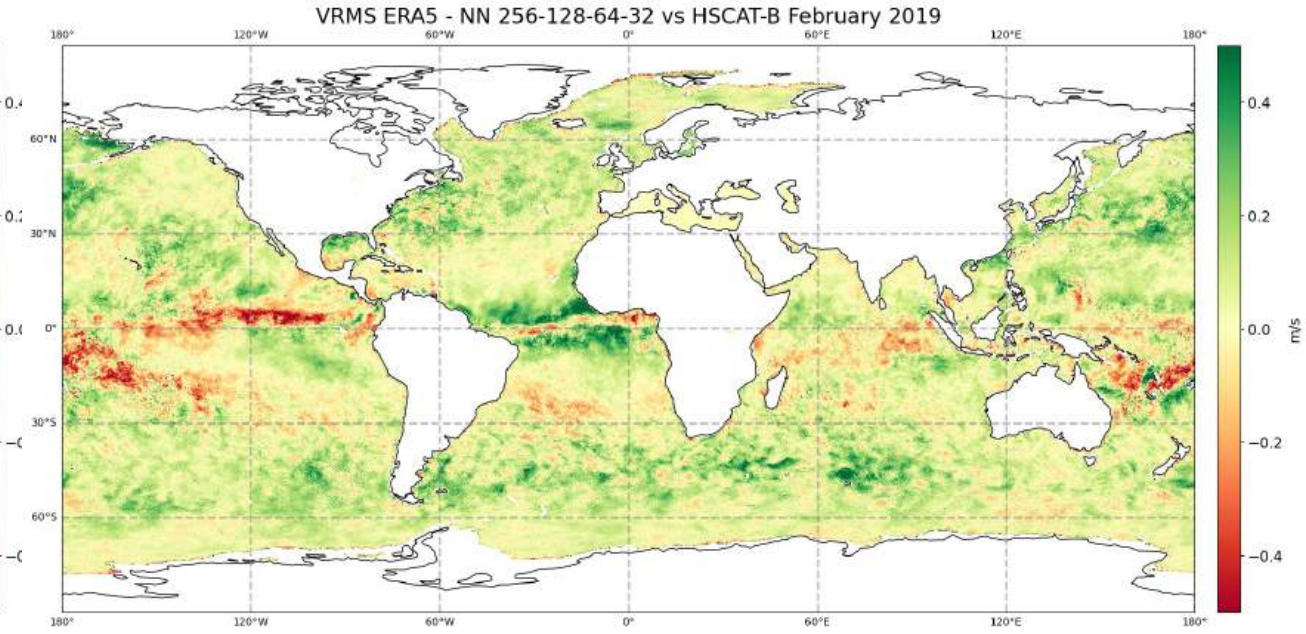
Better metrics for months seen in training but applied to another year -> seasonal models?

Spatial distribution of VRMS vs ASCAT-A and HSCAT-B

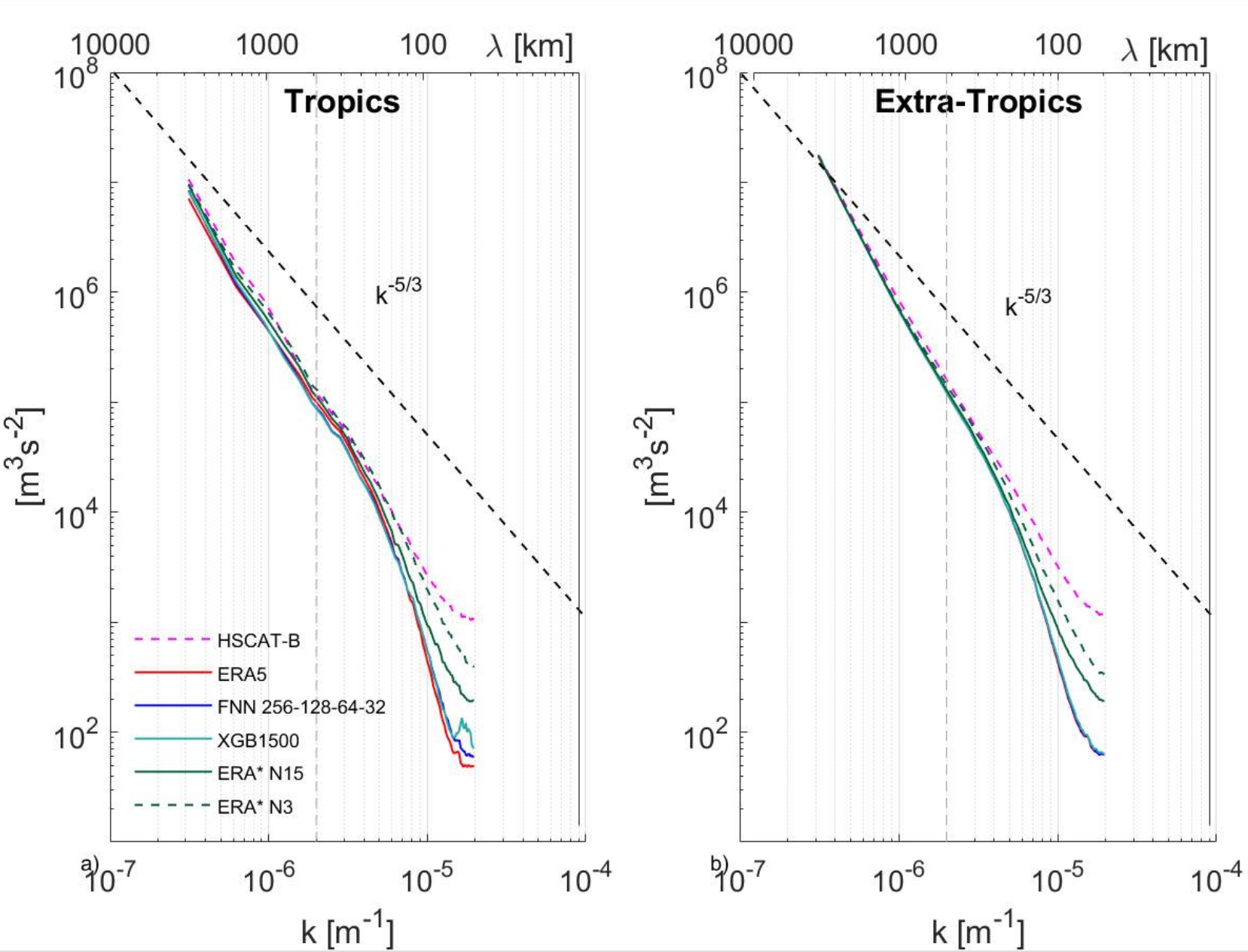








Spectral density



v-component
February 2019
Collocations vs HSCAT-B

- Lower variance at 500-km scale in tropics compared to ERA5
- Corrections are smoothing the abrupt changes in the forecast fields
- Possibly due to the RMSE target loss function

Future Work

- Create seasonal models
- Generate and validate the corrections for several years, incl. validation against inclined orbit HY-2C & -2D to assess daily cycle effects.
- Interpretation of the resulting model to assess in which conditions ERA5 is prone to errors
- Try more complex algorithms including convolutional neural networks (CNNs) and generative AI (generative adversarial networks, diffusion models)

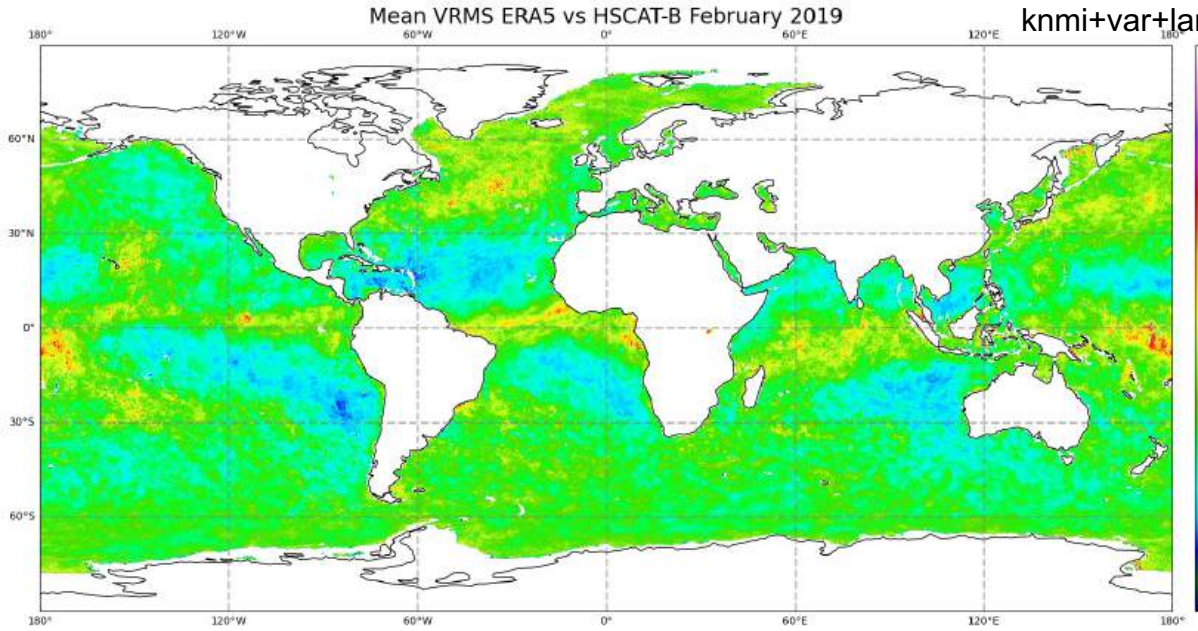
Acknowledgement: This work has been funded by the EUMETSAT OSI SAF through the OSI_VSA22_01 and OSI_VSA24_01 studies

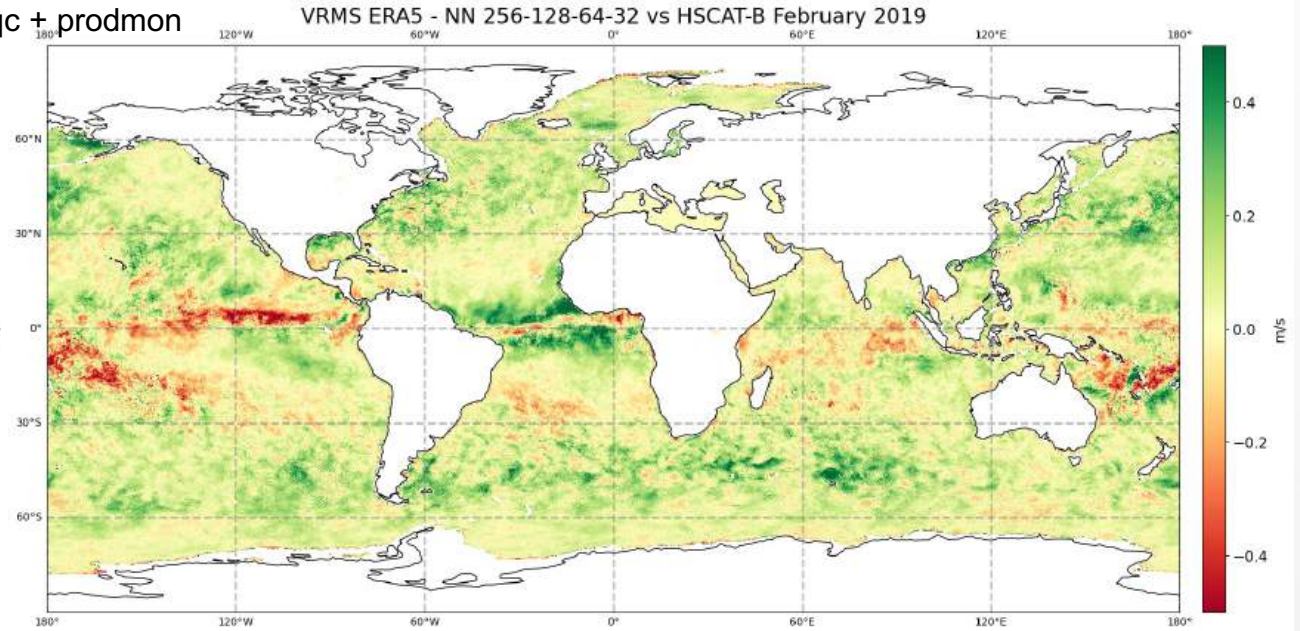
EUMETSAT OSI SAF OCEAN AND SEA ICE

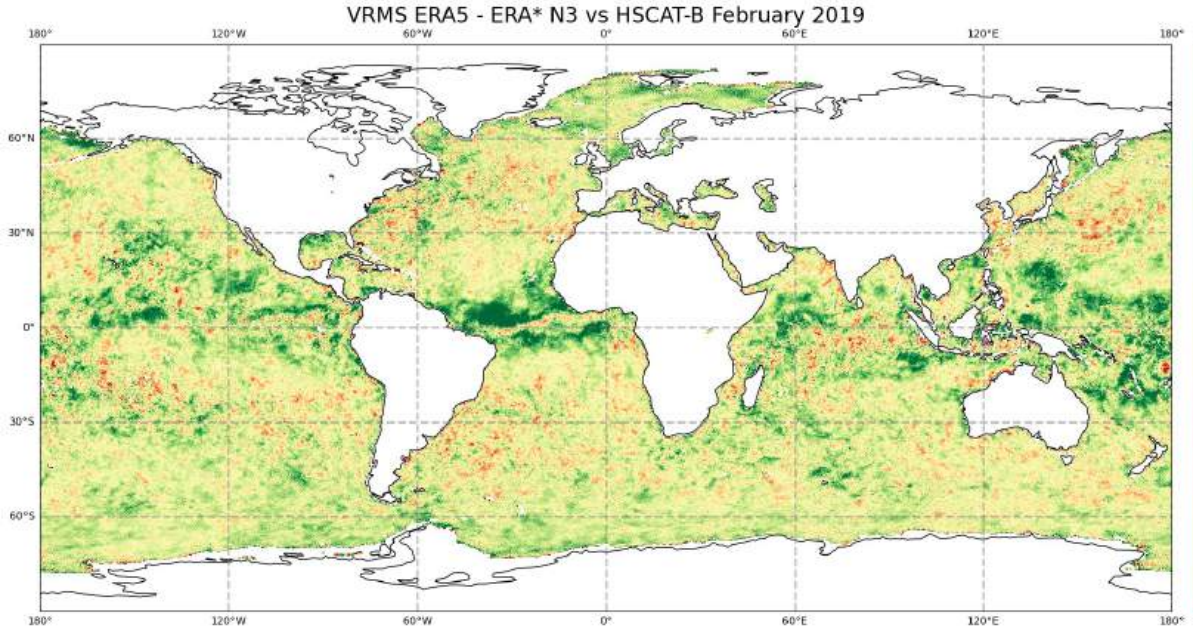
Additional slides

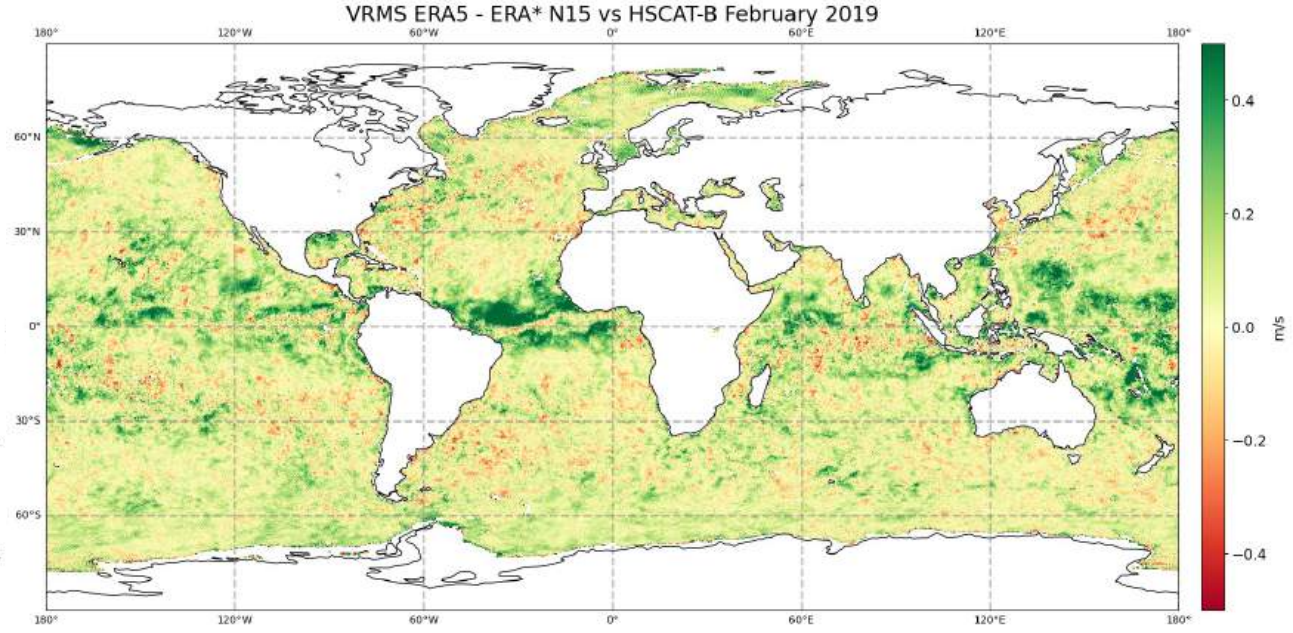
Spatial distribution of VRMS vs HSCAT-B (2019)

knmi+var+land qc + prodmon

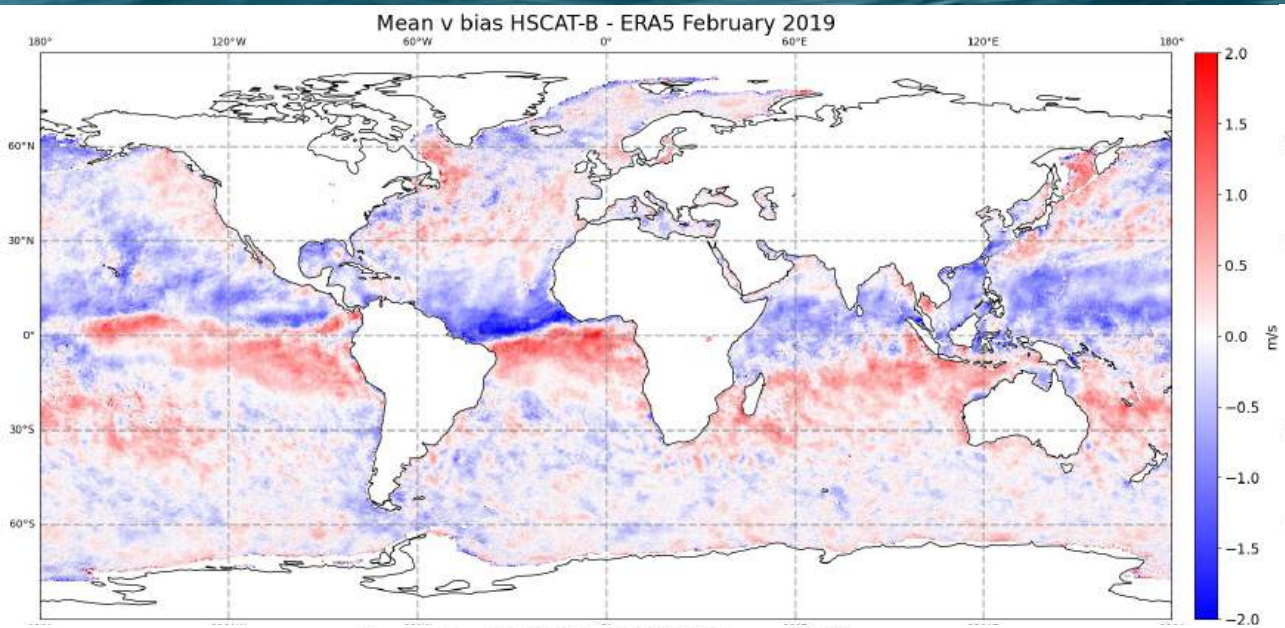


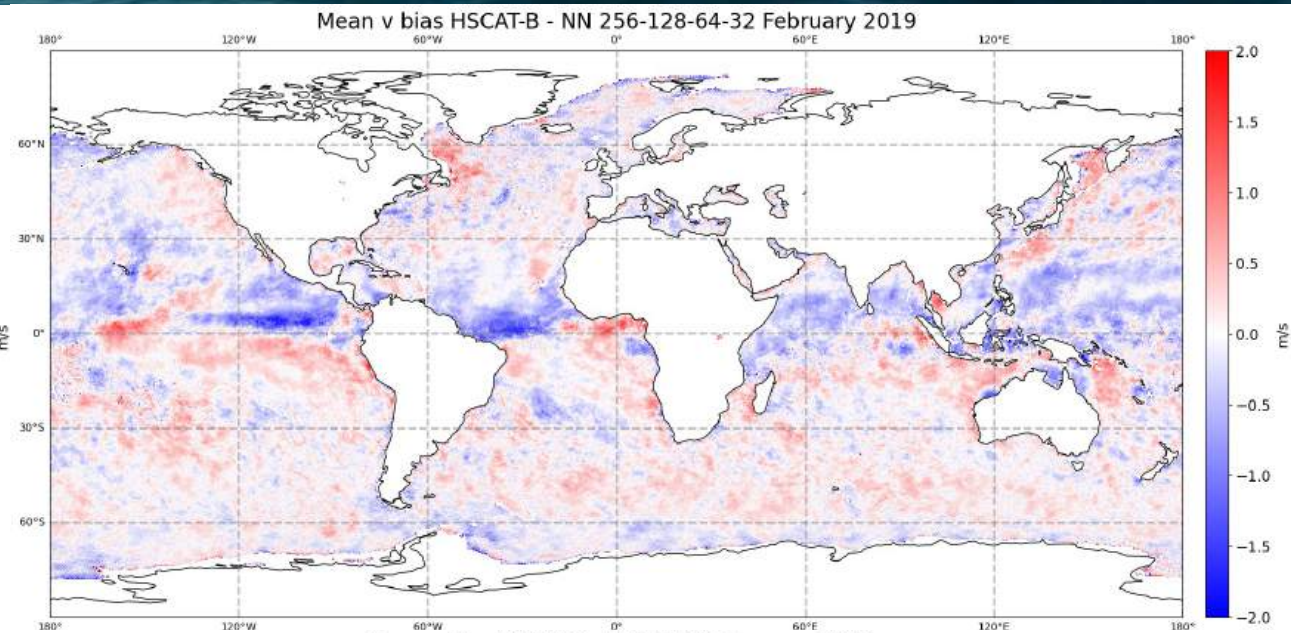


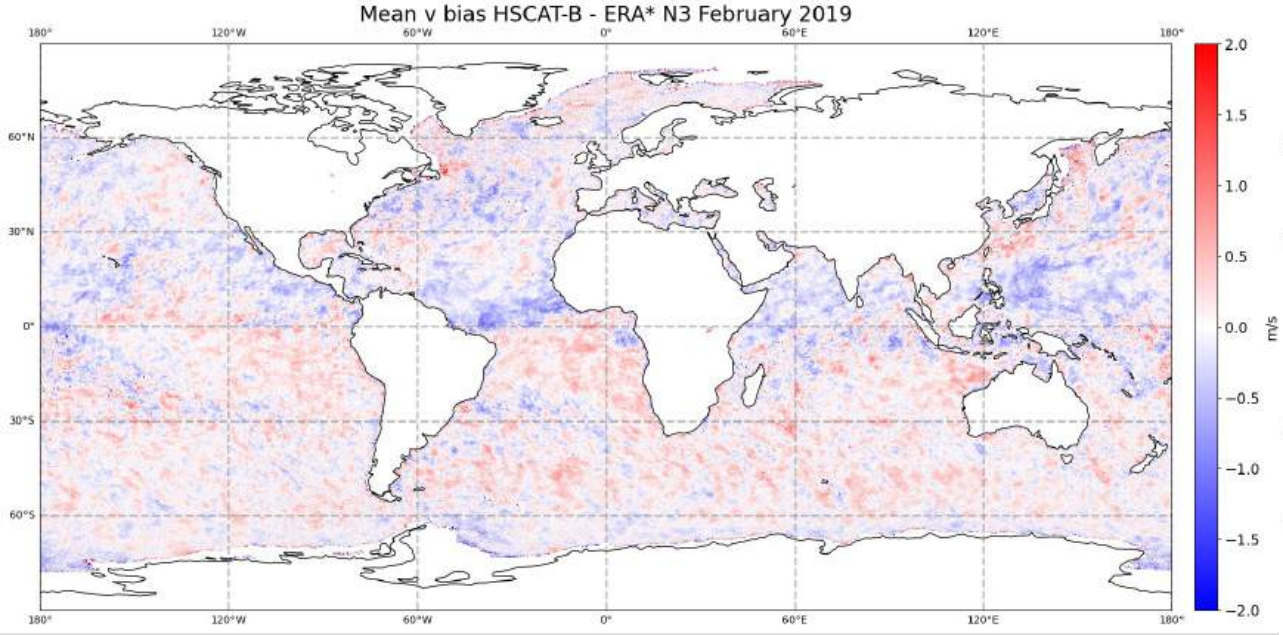


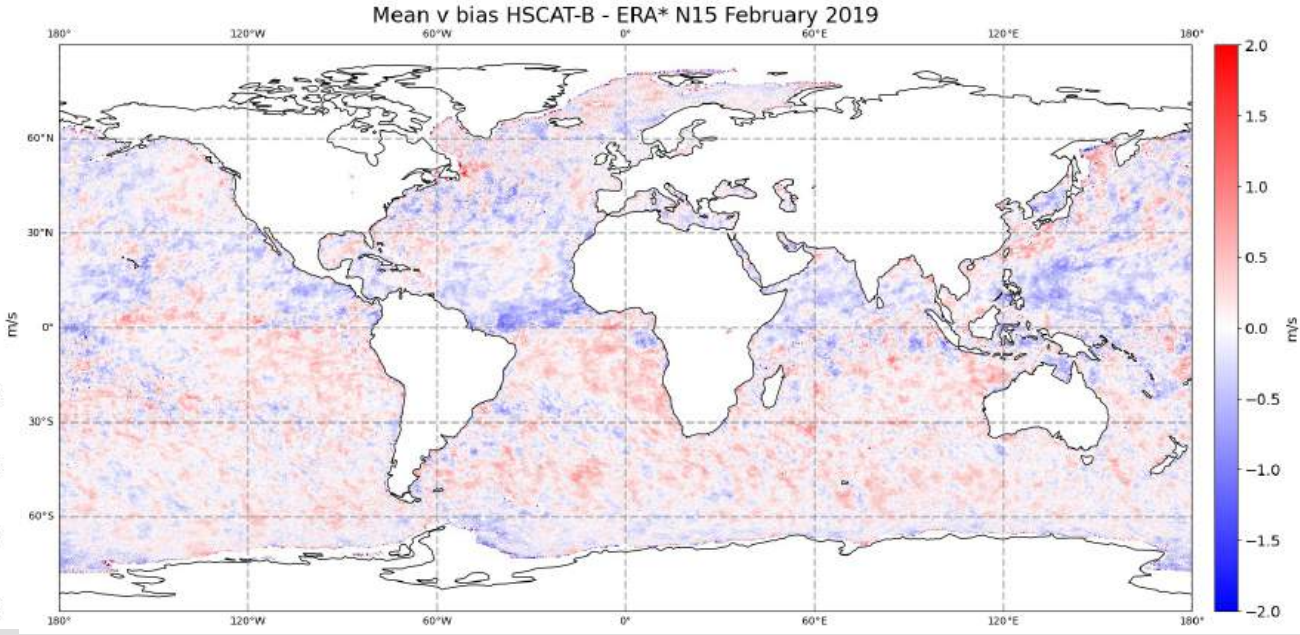


v bias vs HSCAT-B

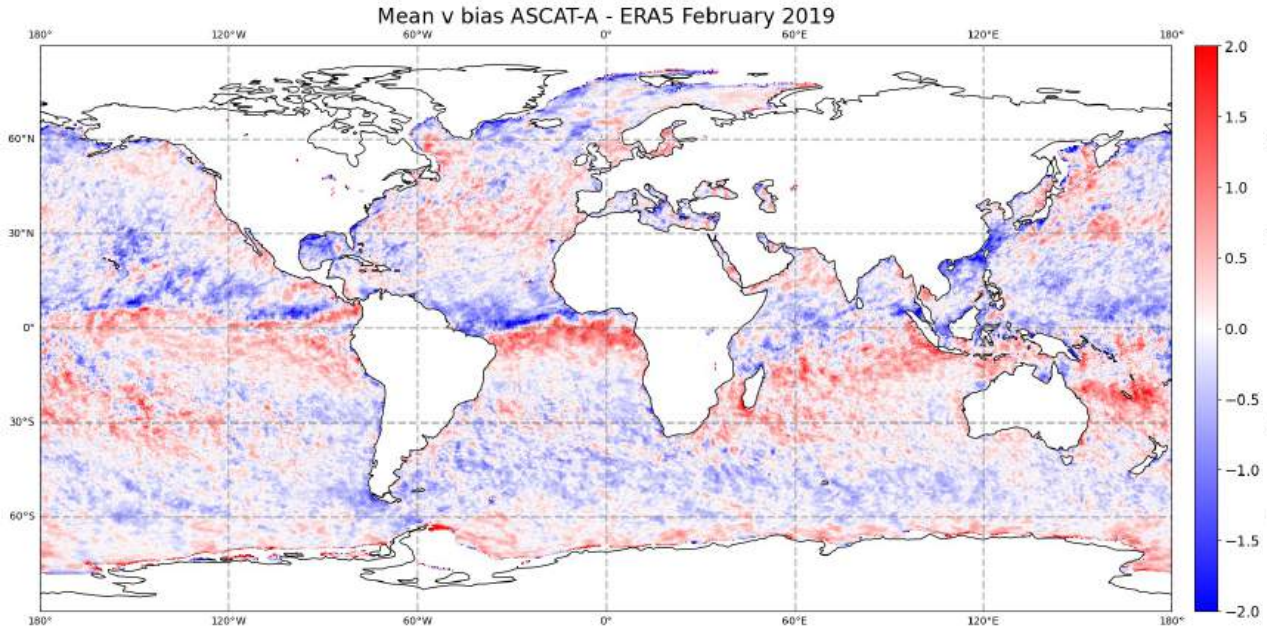


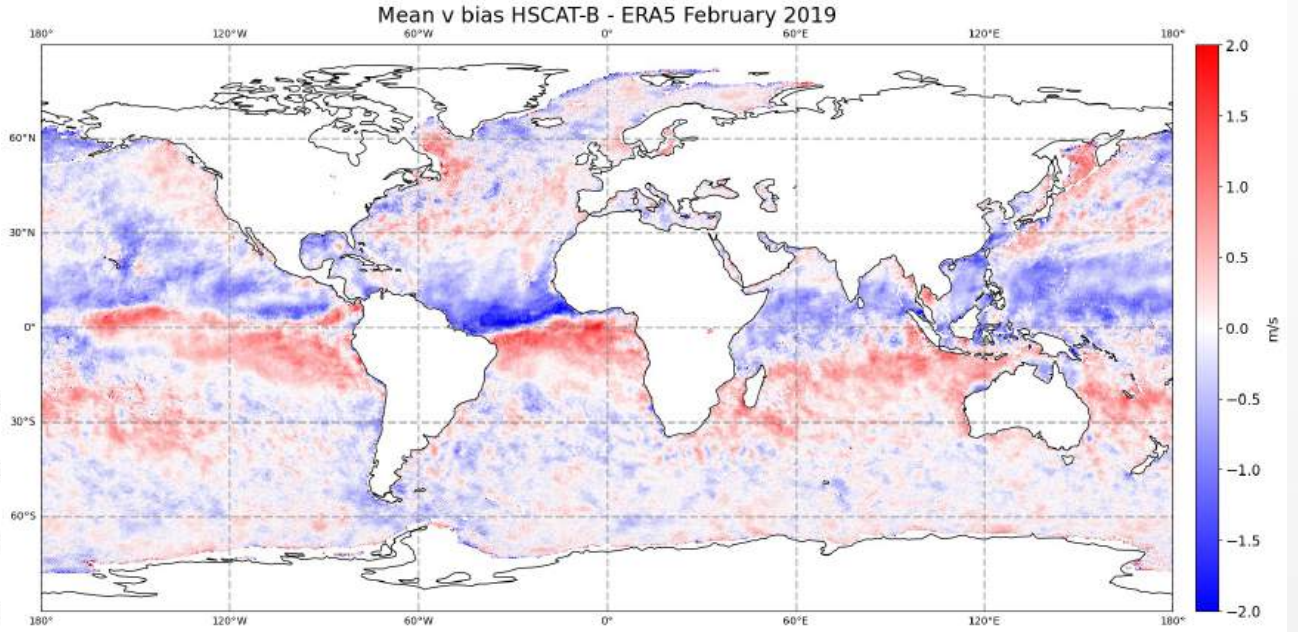


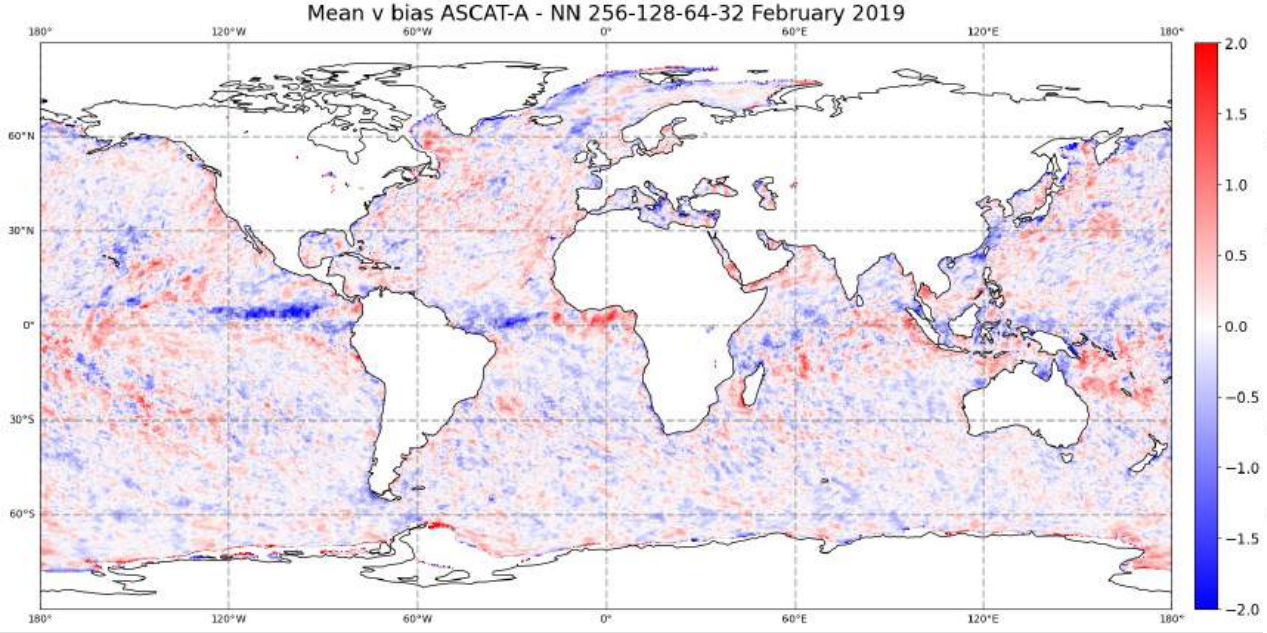


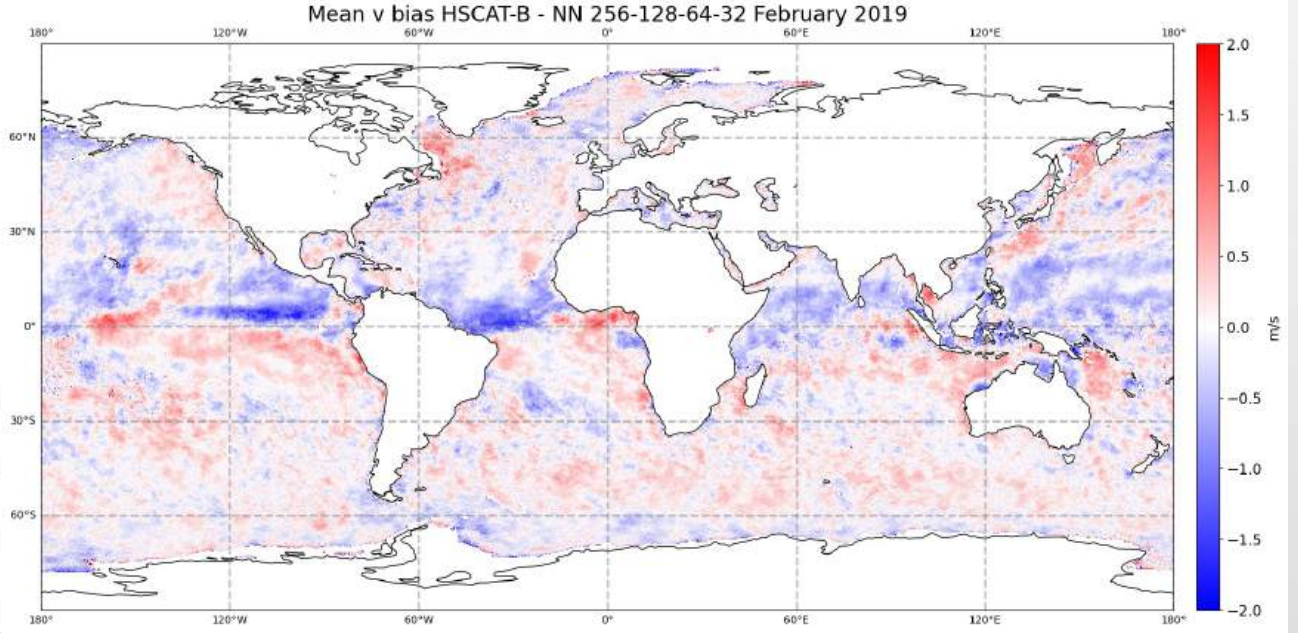


v bias vs ASCAT-A and HSCAT-B

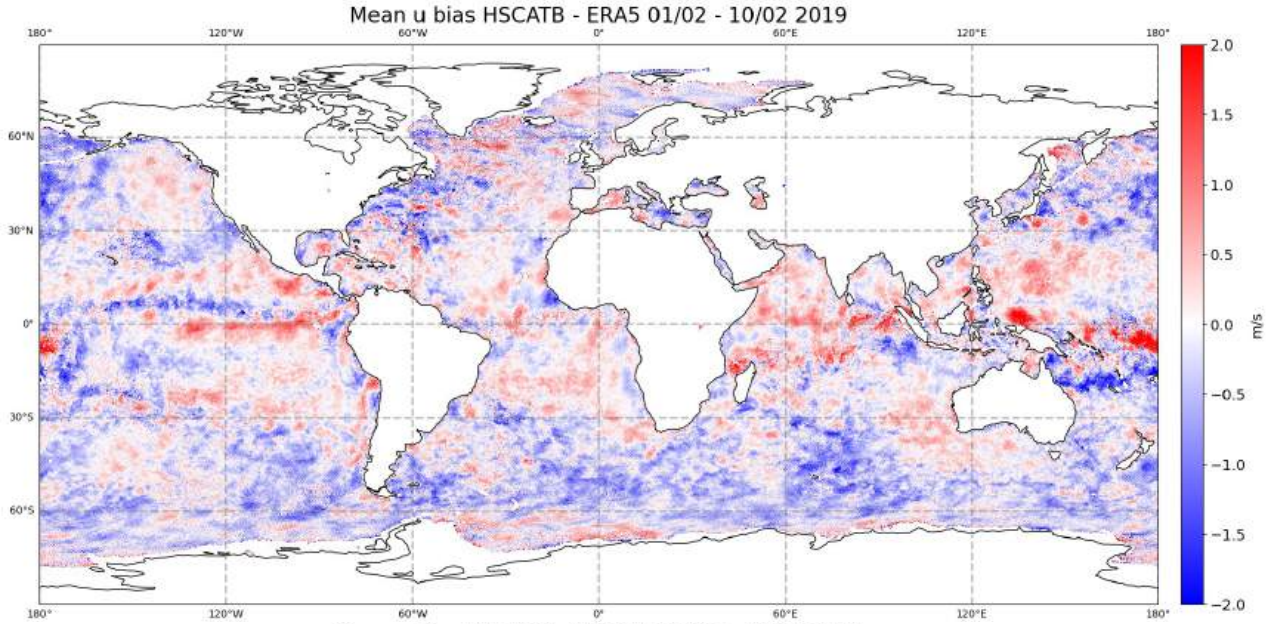


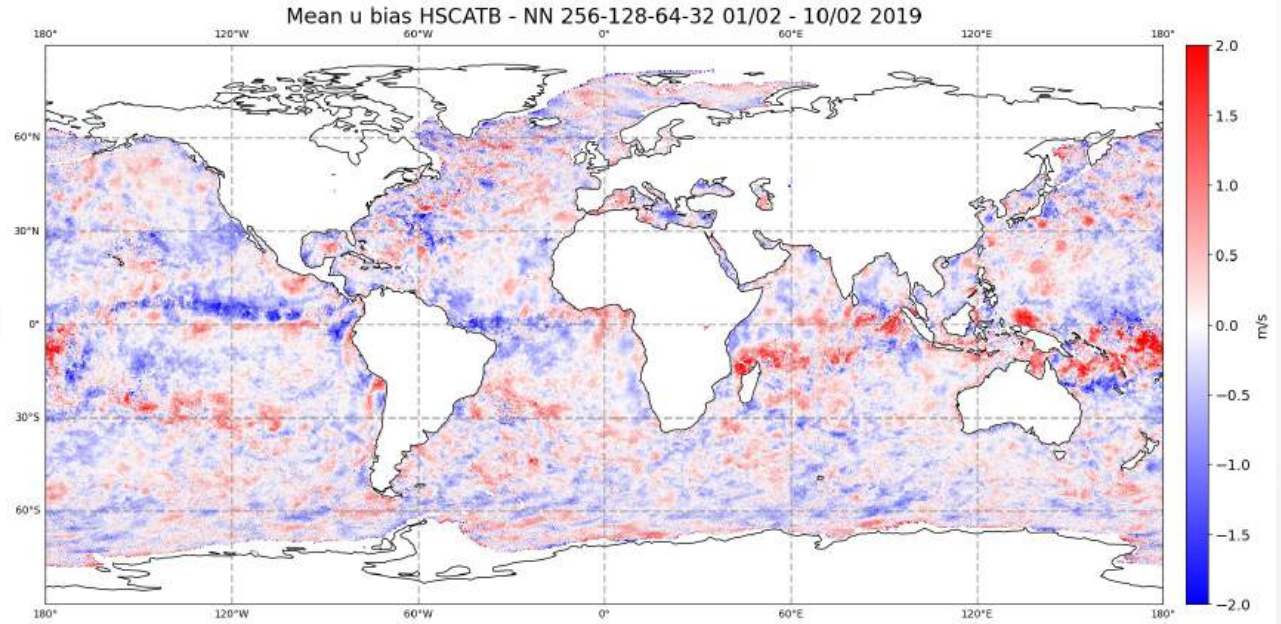


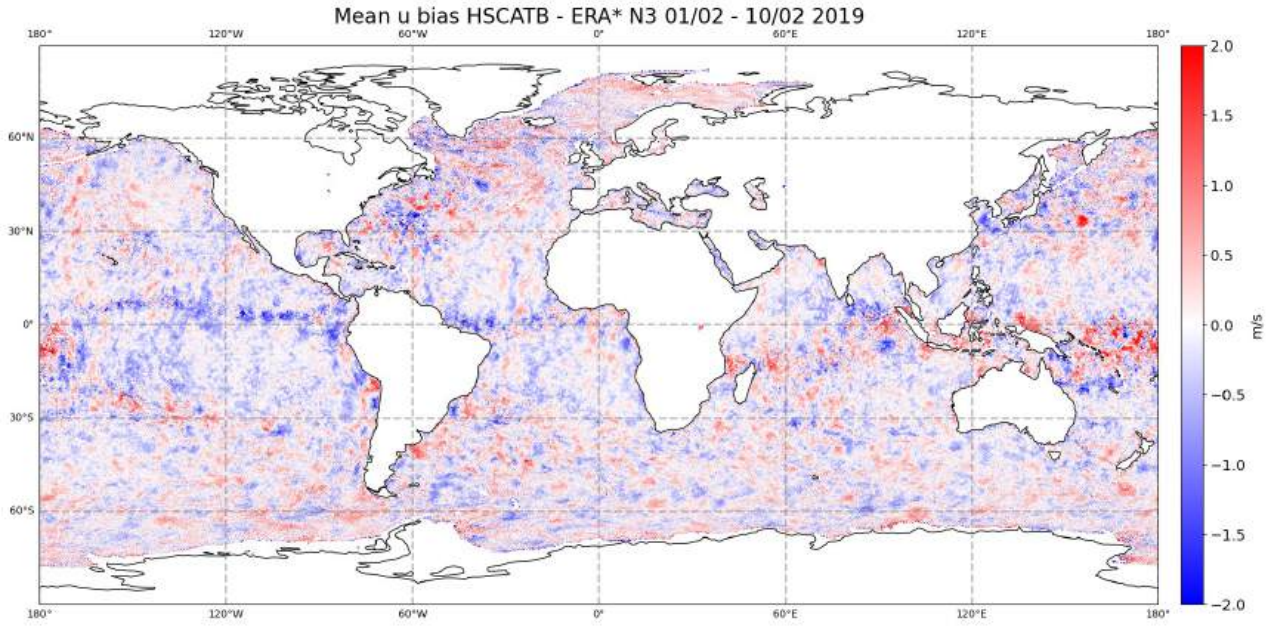


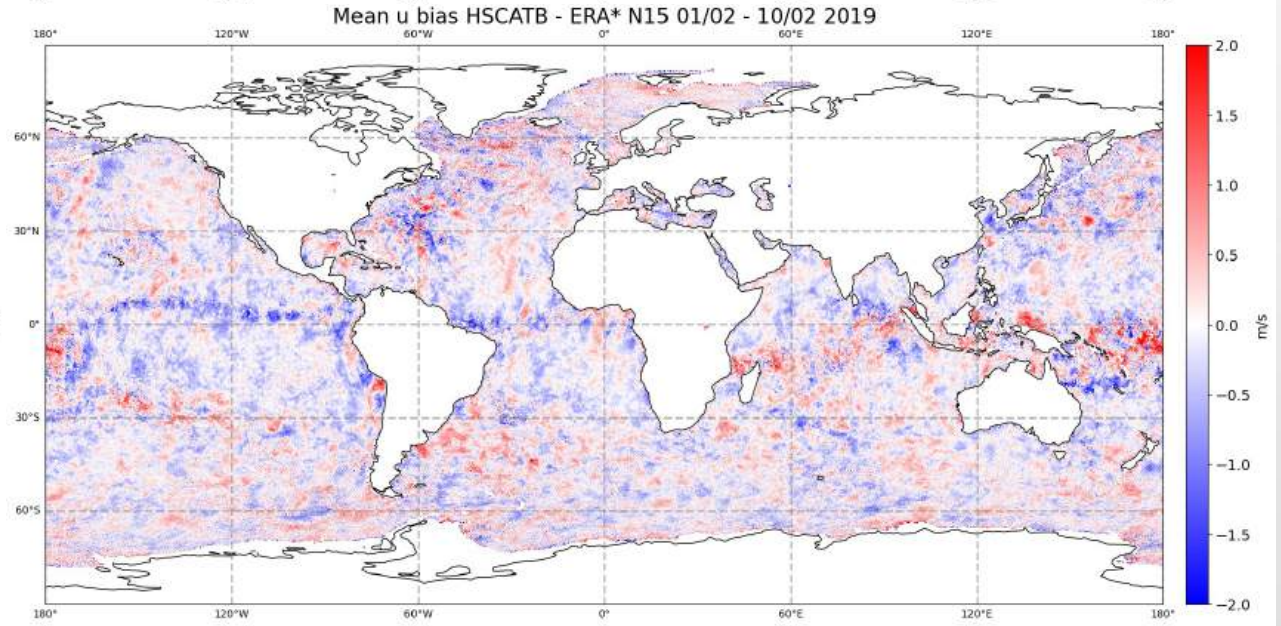


u bias vs HSCAT-B

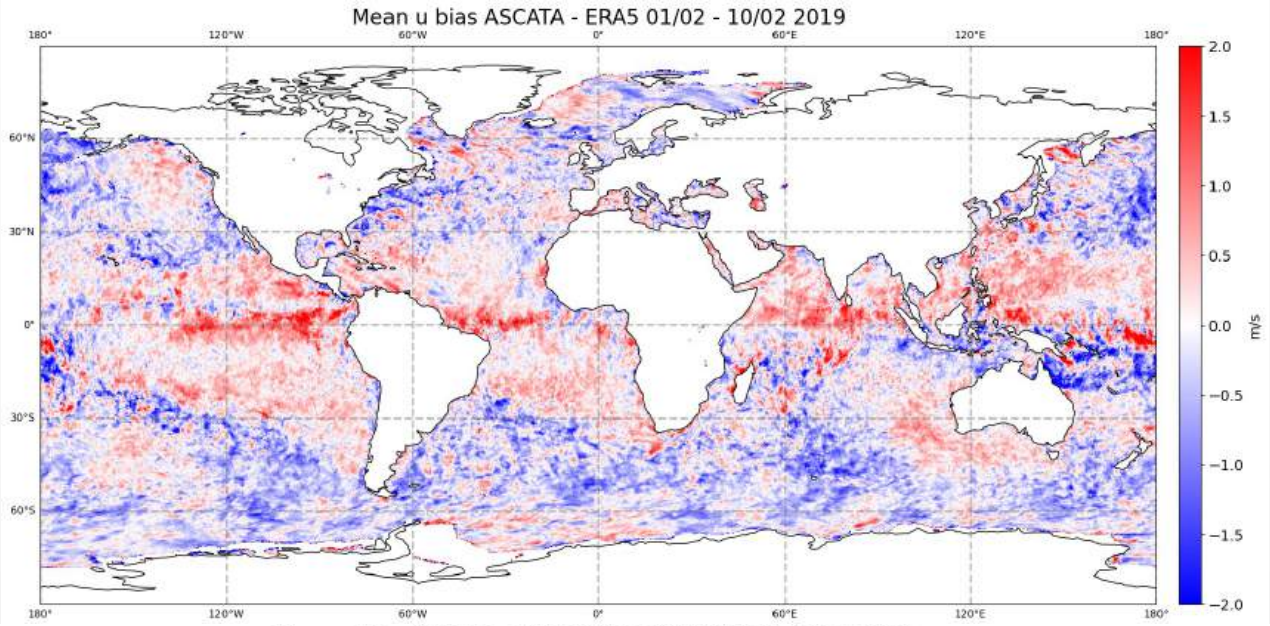


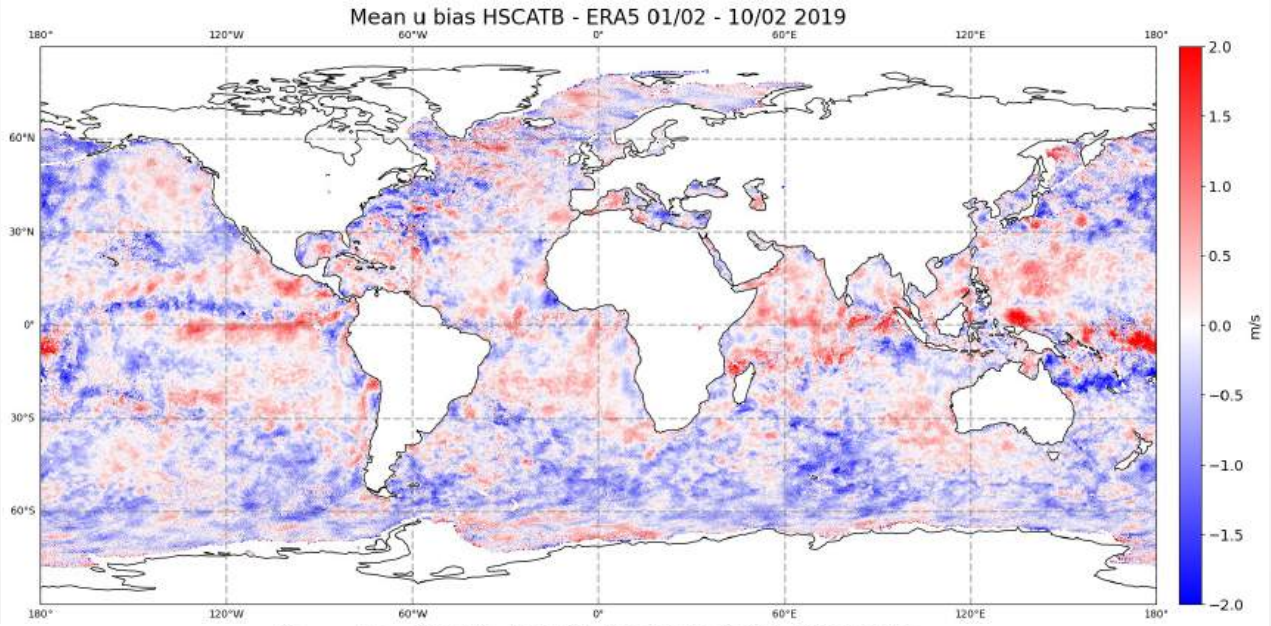


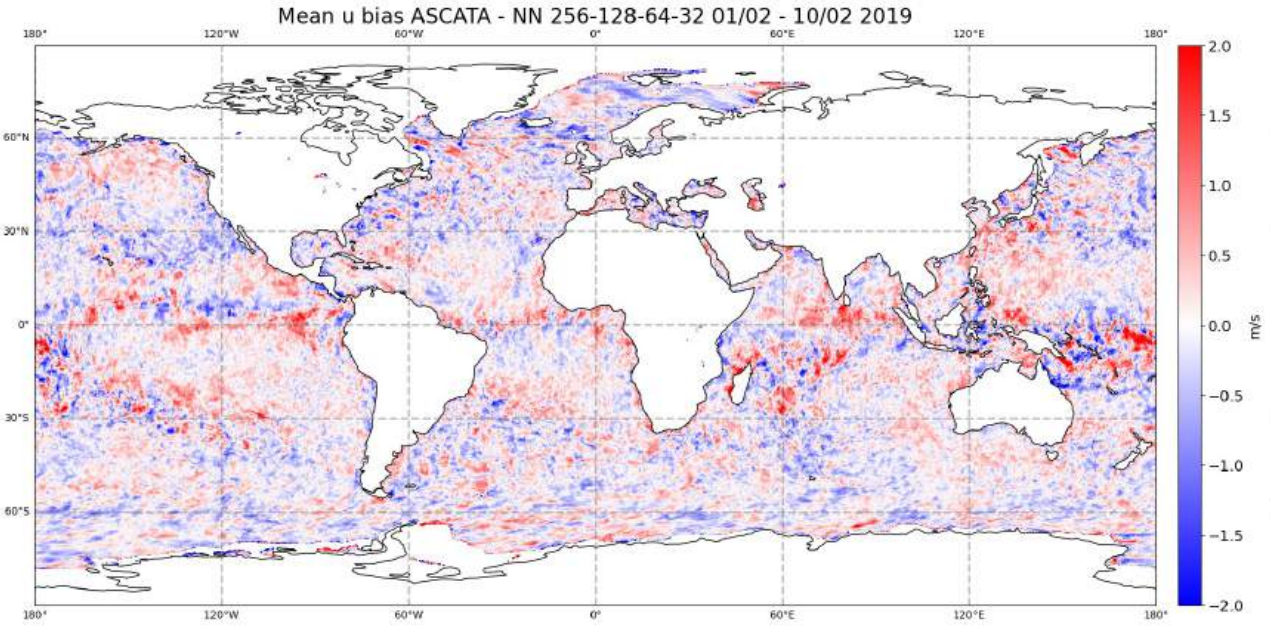


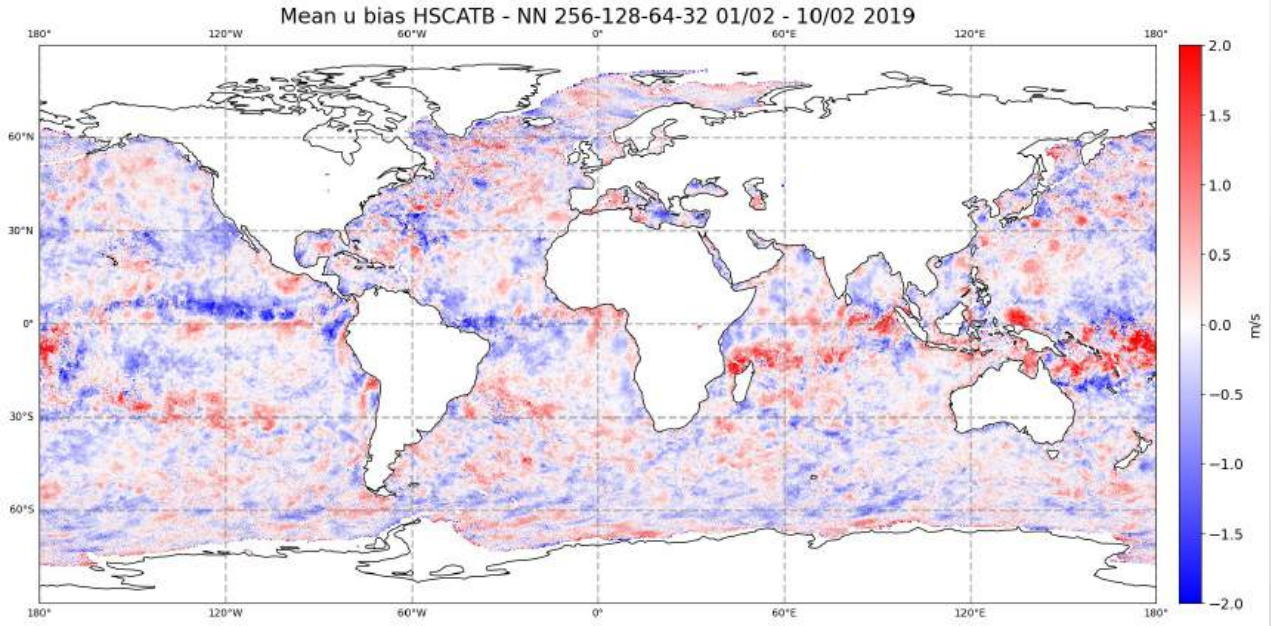


u bias vs ASCAT-A and HSCAT-B



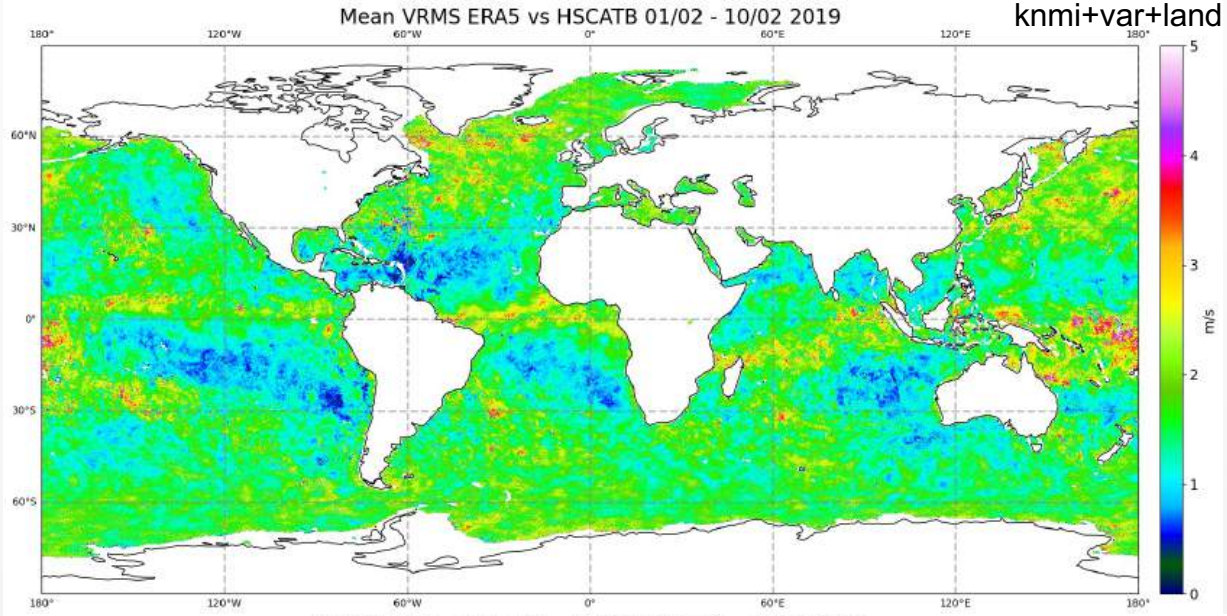


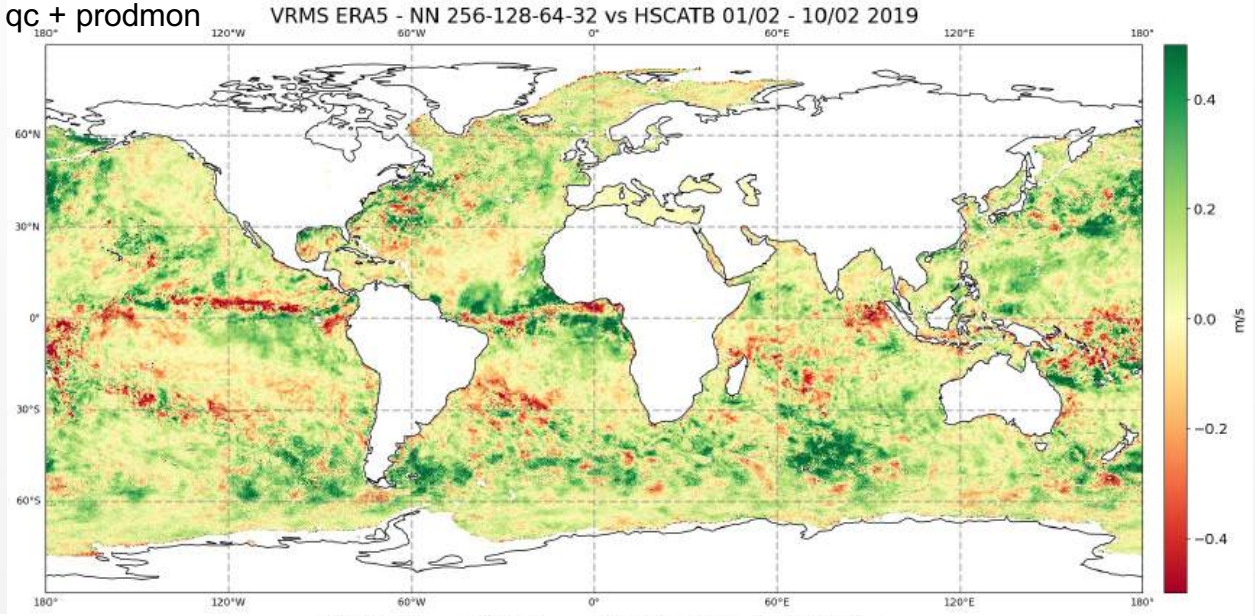


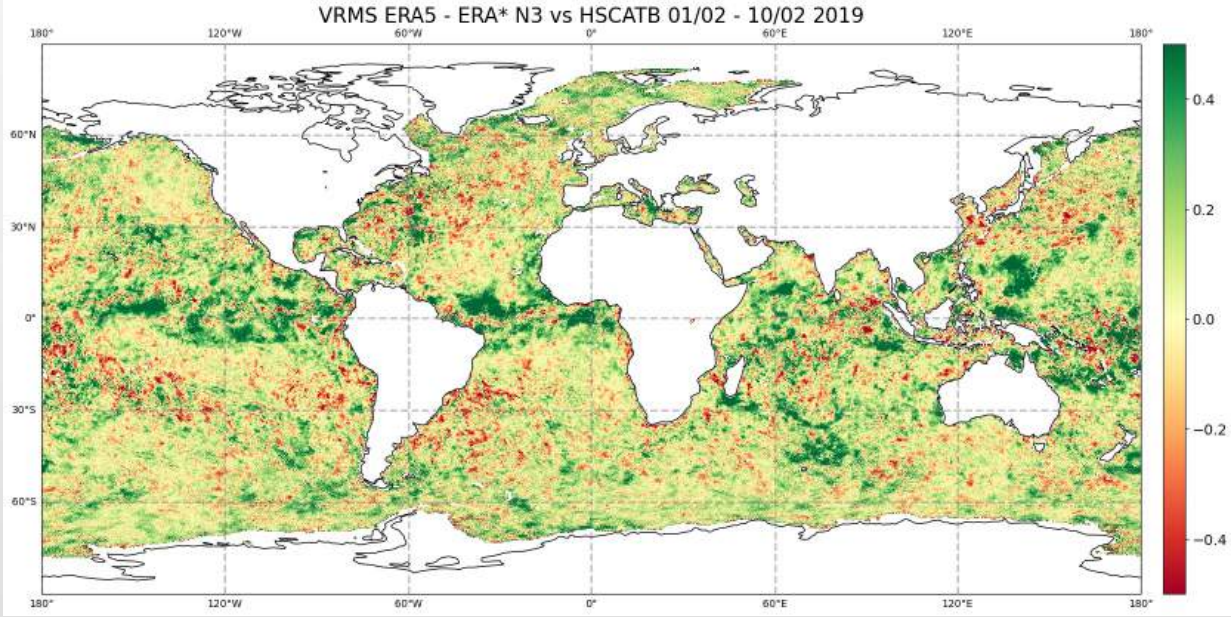


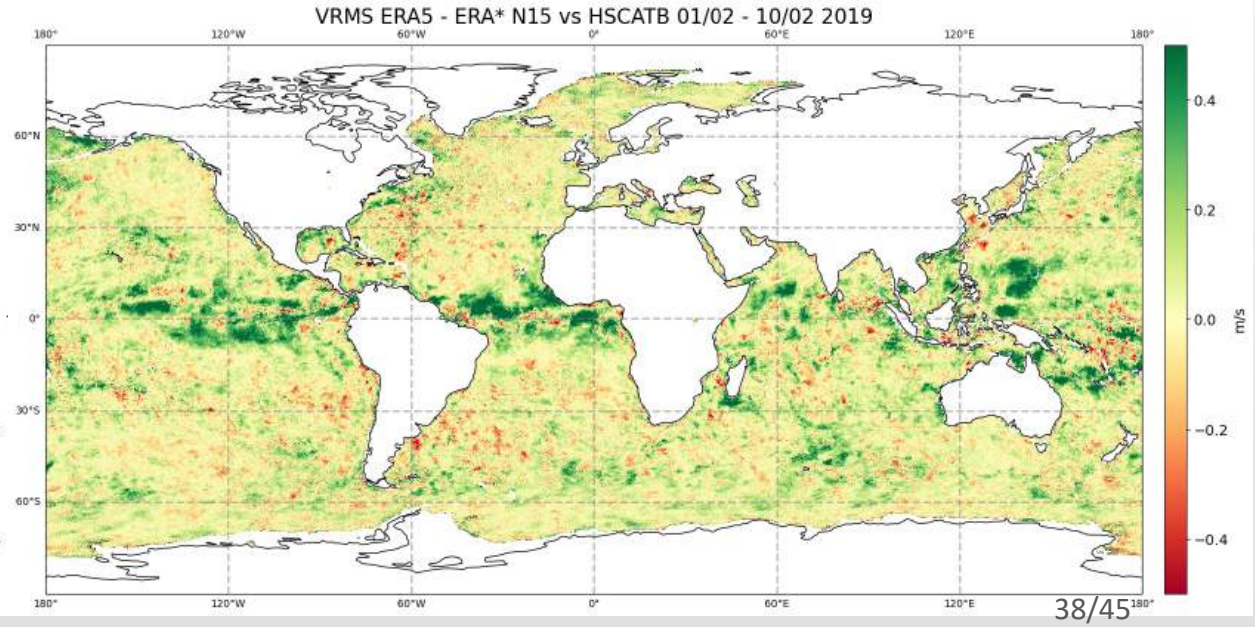
Spatial distribution of VRMS vs HSCAT-B (2019)

knmi+var+land qc + prodmon









Spatial distribution of VRMS vs ASCAT-A and HSCAT-B