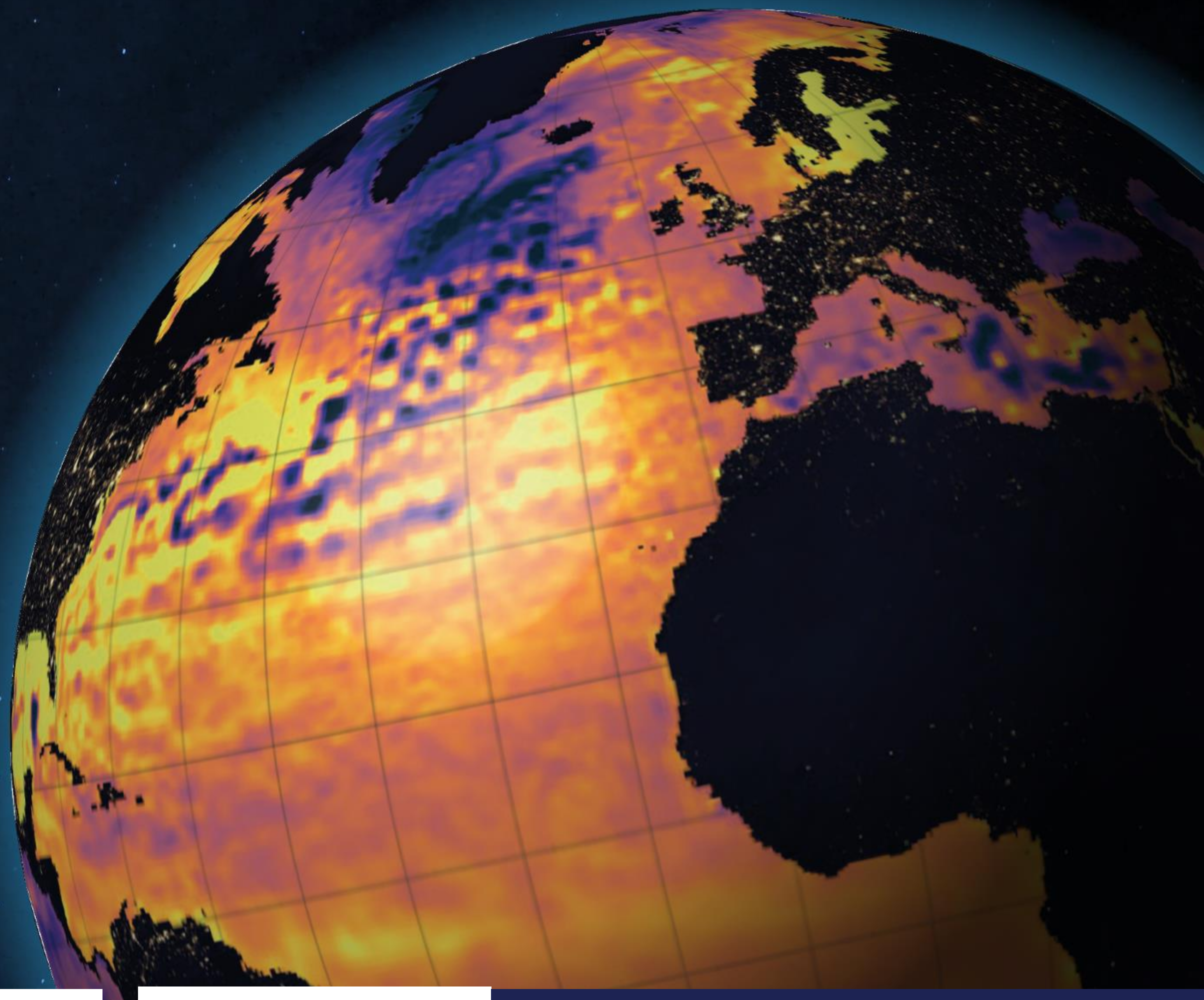
Ocean Predict

SYMPOSIUM OP'24

In partnership with unesco Intergovernmental Oceanographic Commission

2021 2030 United Nations Decade of Ocean Science for Sustainable Development



S. Barthélémy, F. Counillon, J. Brajard, L. Bertino

Universitas Bergensis

NERSC Stiftelsen Nansen Senter for Miljø og Fjernmåling

Bjerknes Climate Prediction Unit

Bjerknes Centre for Climate Research

Trond Mohn Foundation

IMPETUS 4CHANGE

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081555

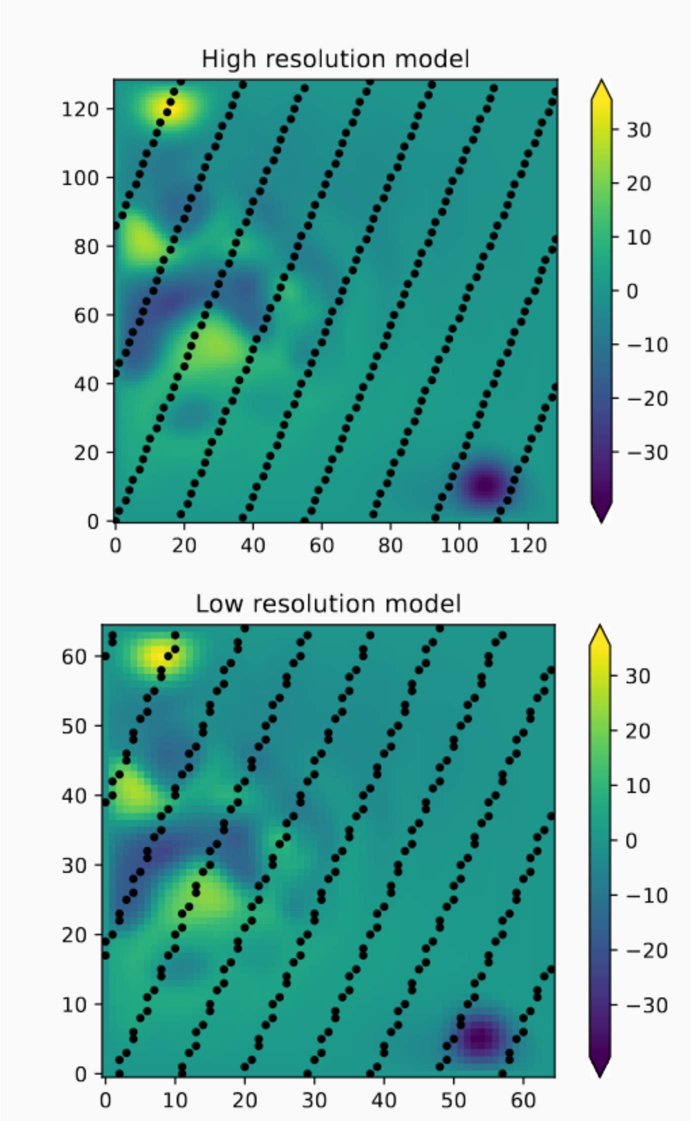
Introduction

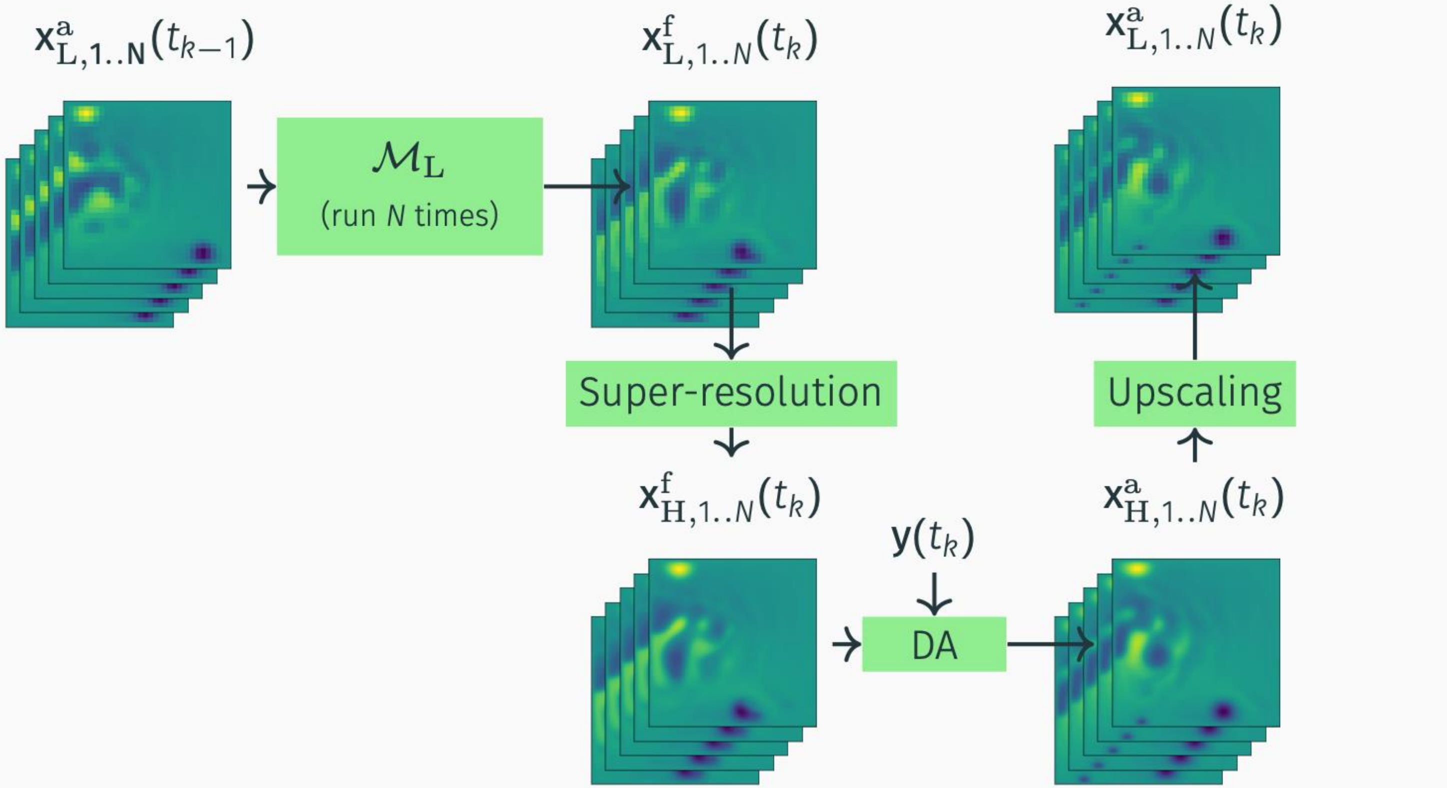
- Many evidences that increasing model resolution reduces model error
- However, the cost of increasing the resolution is ramping up (>*8) , often at the expenses of the complexity of data assimilation methods
- The observational data set in the ocean is very short and sparse making it challenging for ML to build emulator

Can we use ML to learn the resolution increase and improve our data assimilation system ?

Super-resolution data assimilation (SRDA)

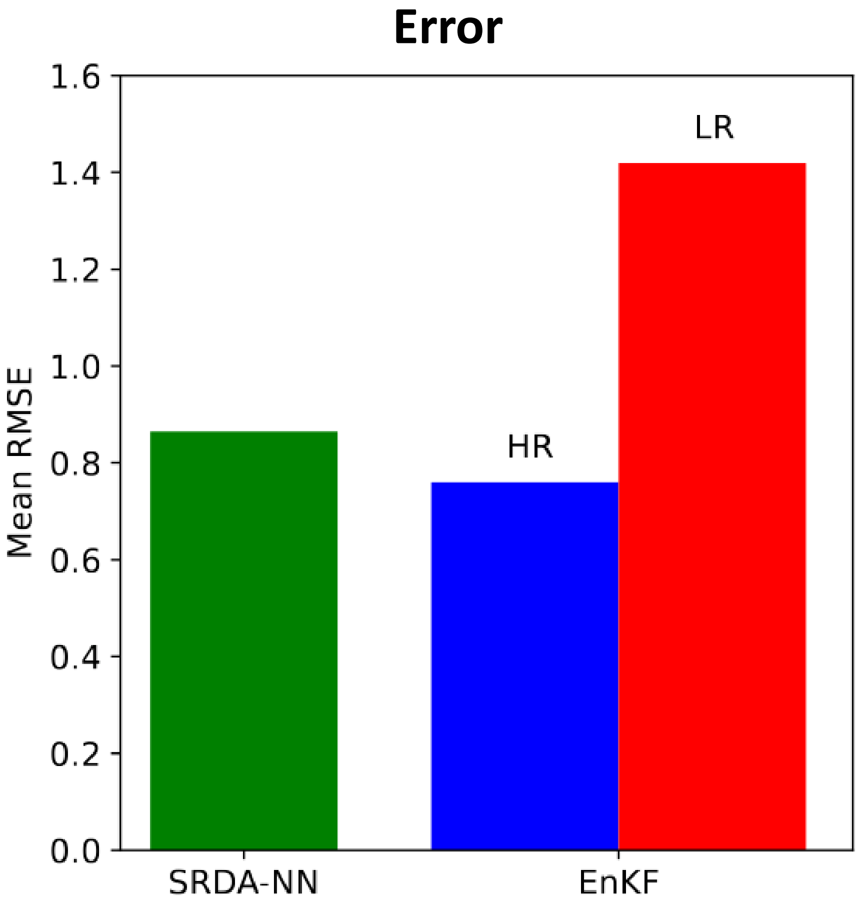
Robustly tested with a quasi-geostrophic model at 2 resolutions (HR,LR) in twin experiment

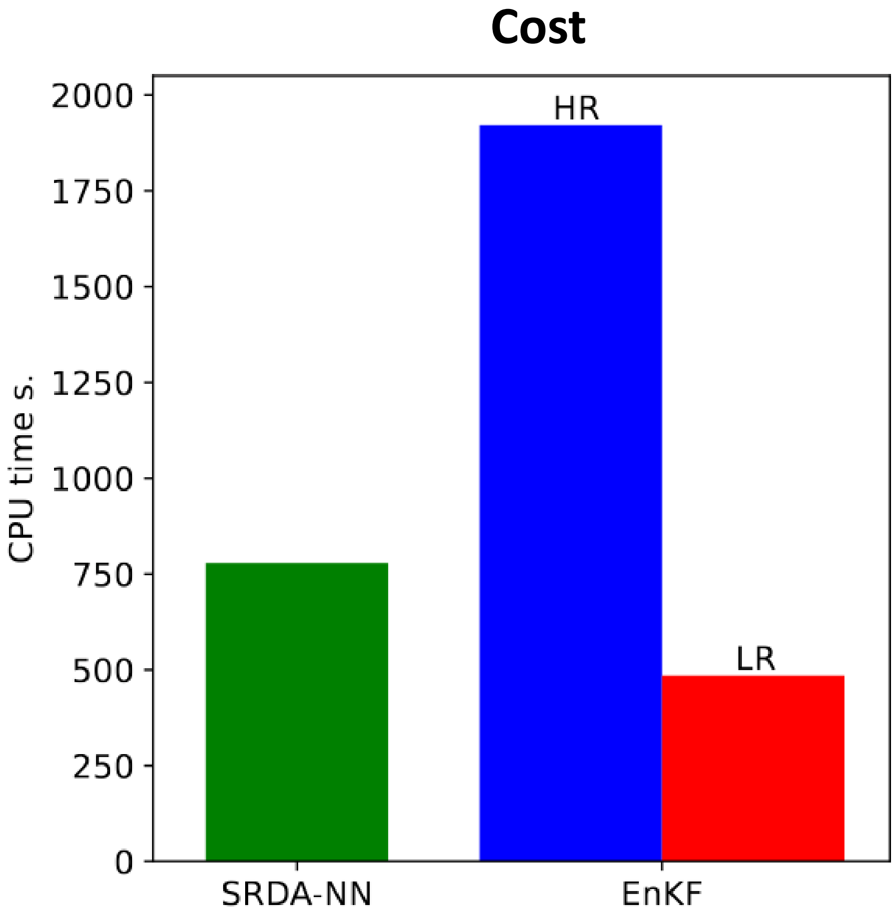




	EnKF-LR	EnKF-HR	SRDA
Observation error	High✔	Low✔	Low✔
High-resolution processes	Poorly resolved✔	Resolved✔	Emulated✔
Computational cost	Low✔	High, $\mathcal{O}\left(n^3\right)$✔	Low✔
Ensemble size	Large✔	Small✔	Large✔
Error to the true $\mathbf{P}^{f}$	Large✔	Small✔	Medium✔

How does SRDA compares to EnKF-LR & EnKF-HR

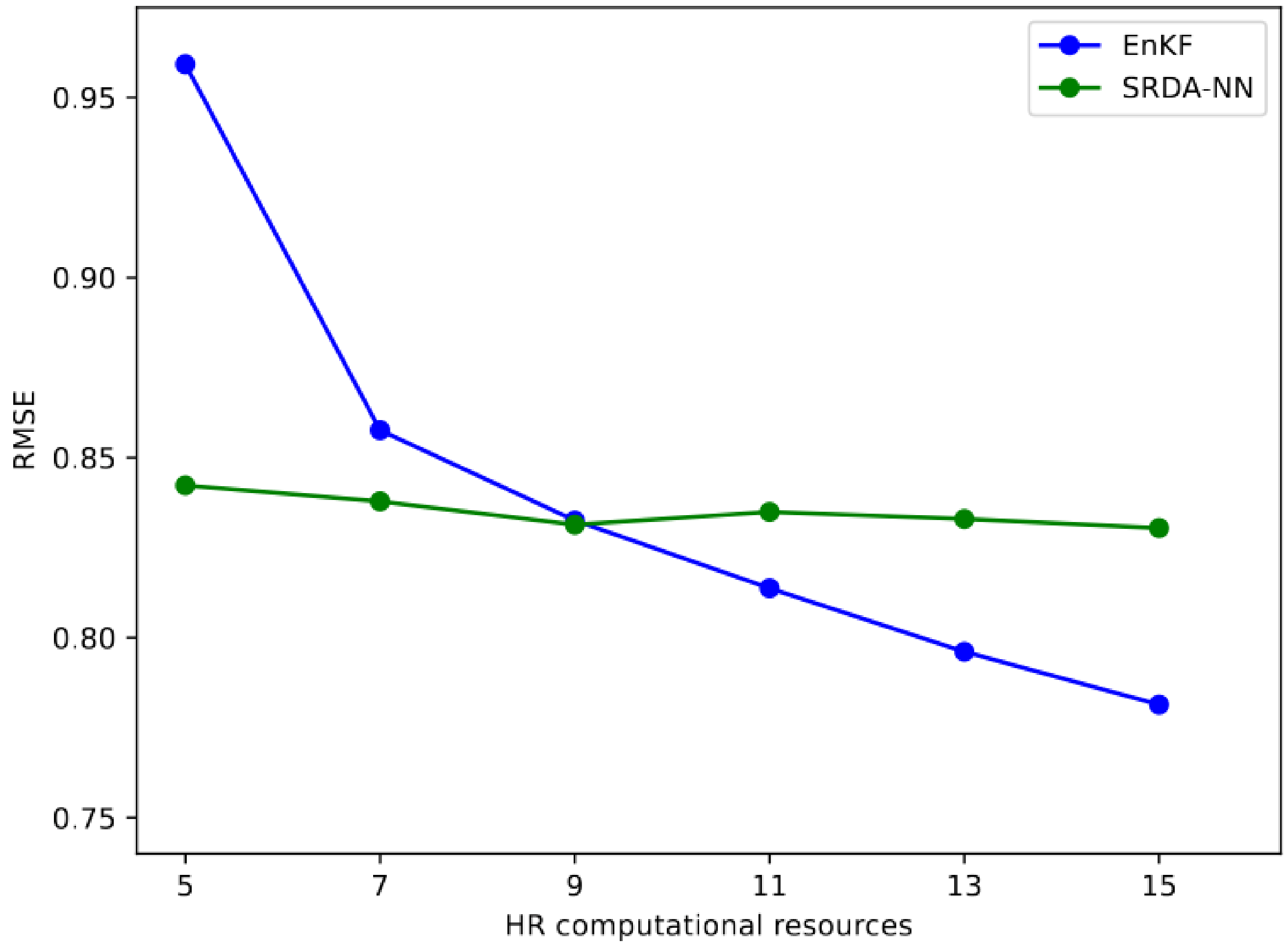




- Improved use of HR observations
- Corrects model bias

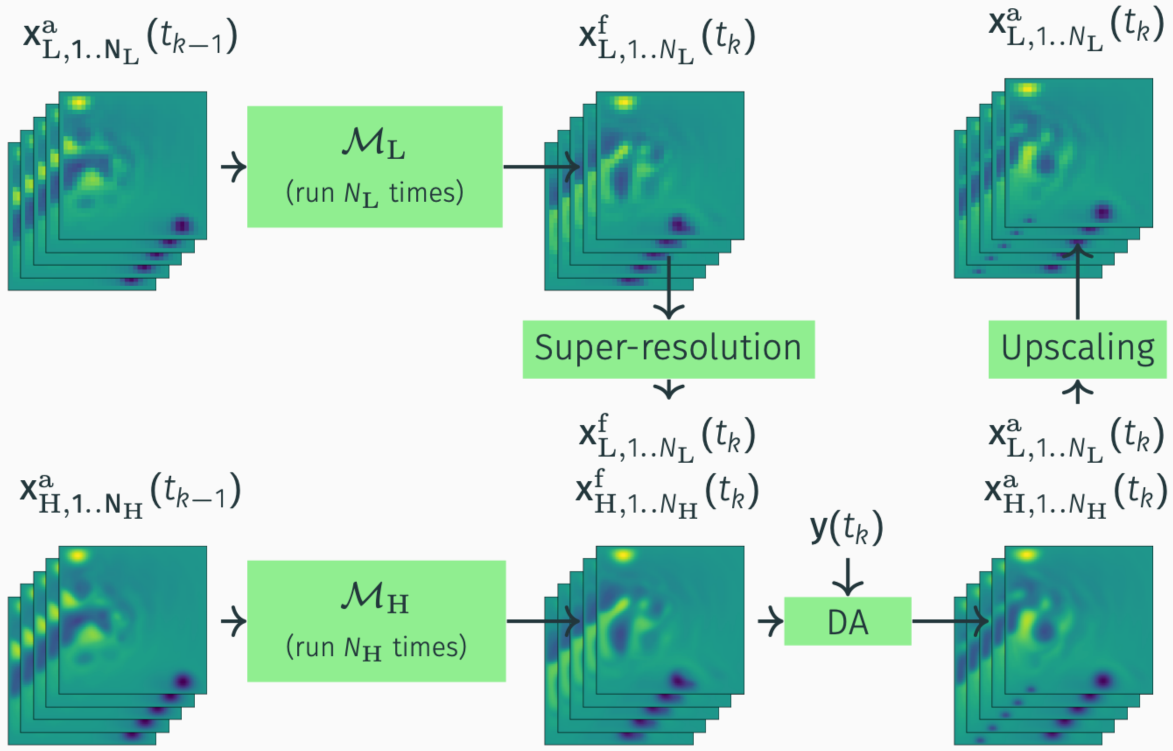
Barthelemy et al. 2022

But performance quickly saturate



This may relate to irreducible error in the NN emulator

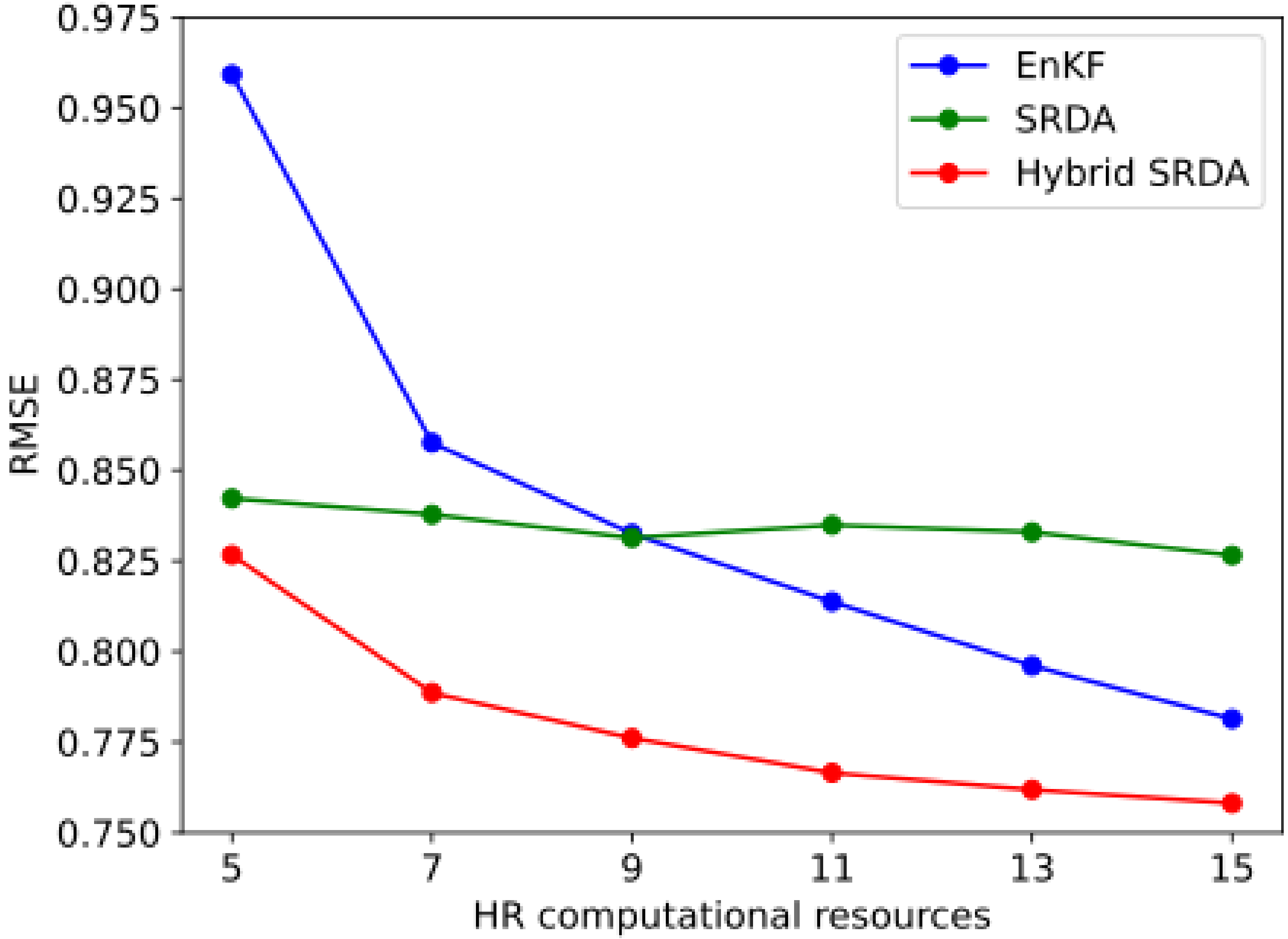
Hybrid Covariance SRDA



	EnKF-LR	EnKF-HR	SRDA	Hybrid SRDA
Observation error	High✔	Low✔	Low✔	Low✔
HR processes	Poorly resolved✔	Resolved✔	Emulated✔	Emulated (LR)✔/ resolved (HR)✔
Computational cost	Low✔	High, $\mathcal{O}\left(n^3\right)$✔	Low✔	Customizable(✔–✔)
Ensemble size	Big✔	Small✔	Big✔	Big✔
Error to the true P^f	Large✔	Small✔	Medium✔	Customizable(✔–✔)

1. We run in parallele few HR member & a large LR ensemble
2. Assimilate hybrid covariance(HR+NN(LR))
3. Upscale back the LR members

Hybrid covariance SRDA



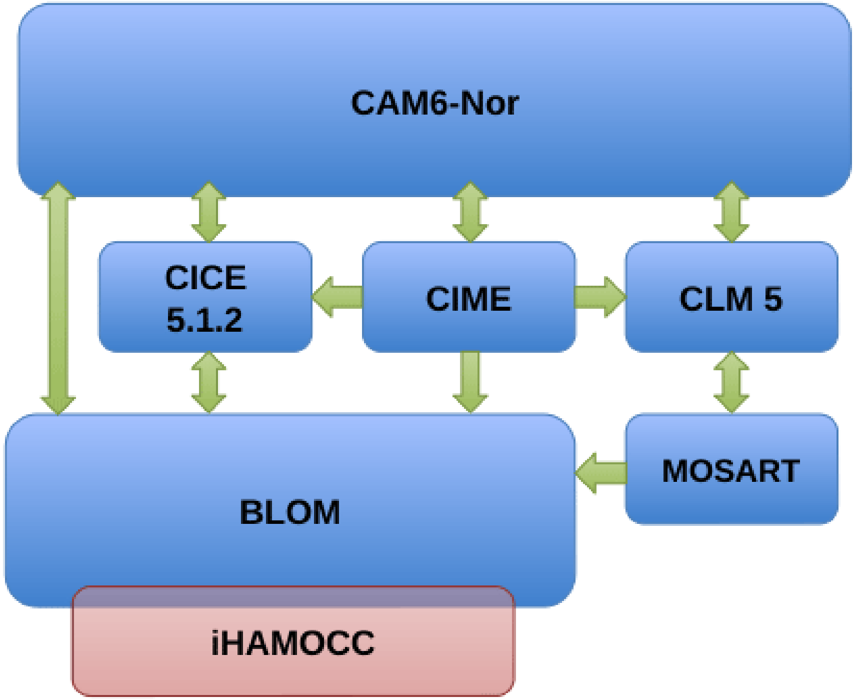
Using of a few HR can greatly enhance the performance

Barthelemy et al. 2024

Application of (hybrid) SRDA with NorESM

- Trainning based on a 400 year long pre-industrial run
- Start NorESM-LM from NorESM-MM every month:
  - Training (80%) year 1056—1375
  - Validation (10%) year 1376—1415
  - Test (10%) 1416—1455

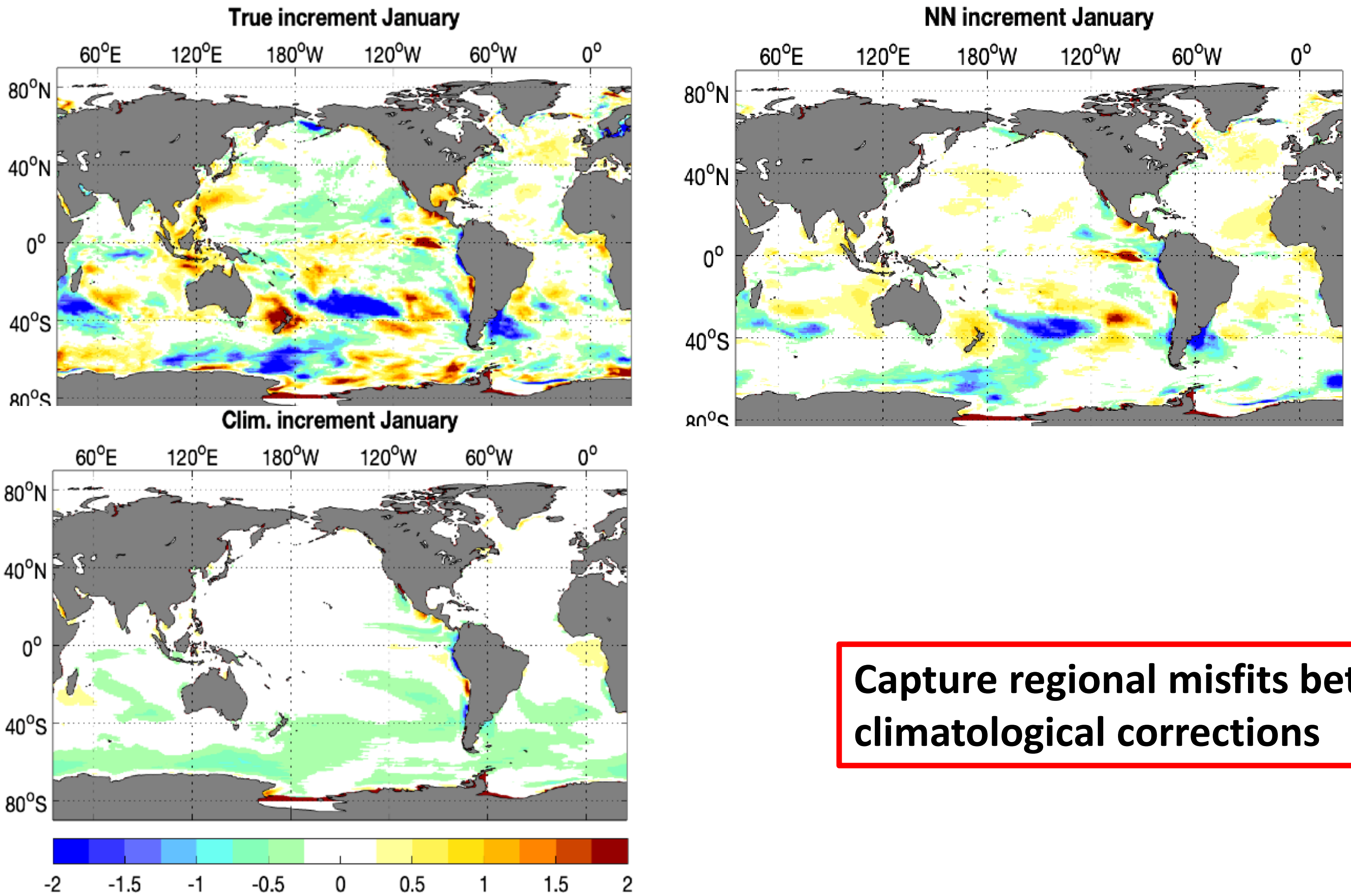
We learn the SST mistmatch after 1 month integration from the 3D ocean state



	Atmosphere / Land resolution	Ocean/sea ice resolution
NorESM2-LM	2°	tripolar 1°
NorESM2-MM	$0.9 \times 1.25^\circ$	tripolar 1°

Predicted mismatch between HR and LR

An example during the test period

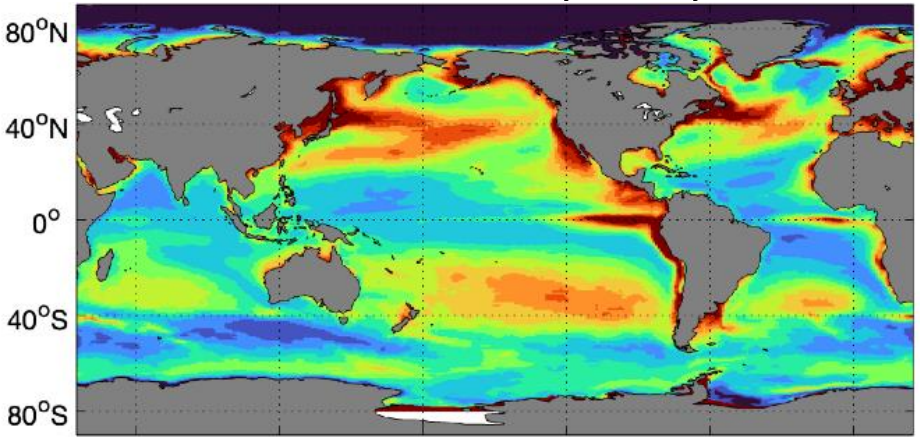


Capture regional misfits better than standard climatological corrections

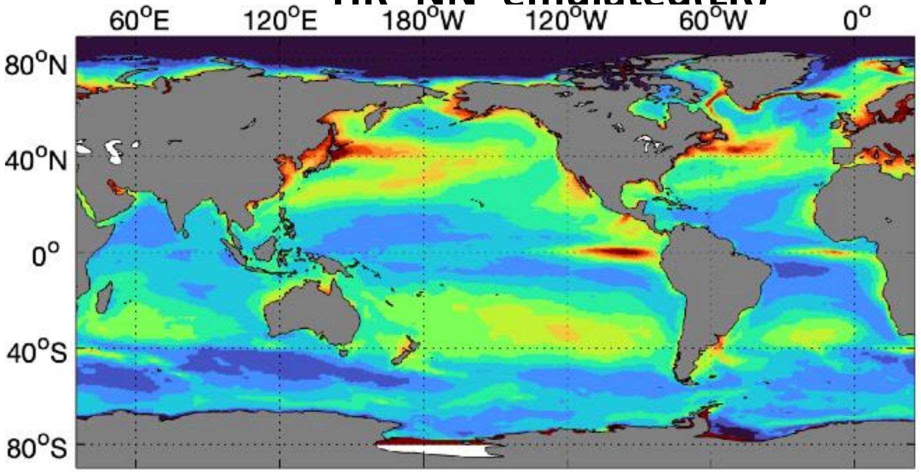
Predicted mismatch between HR and LR

Root mean squarre error (test period)

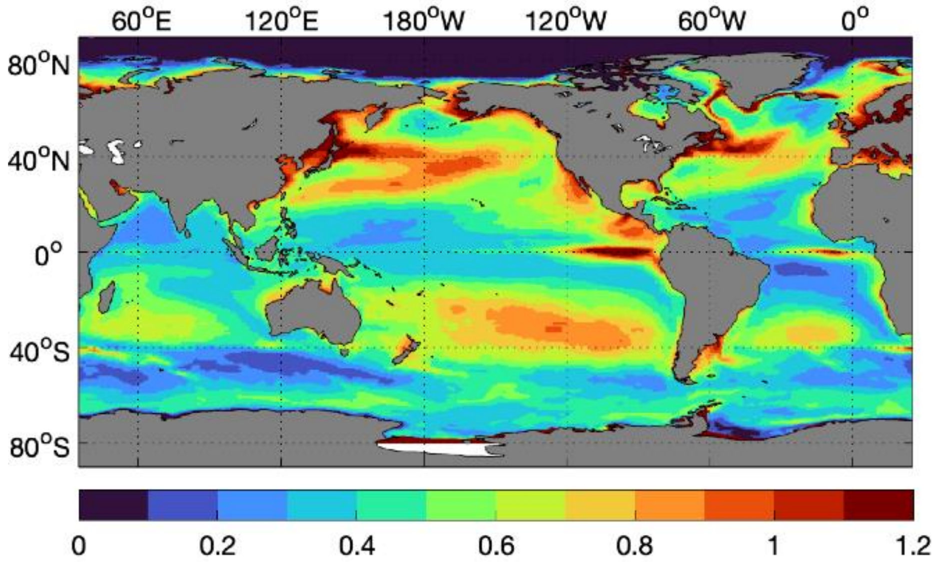
No correction (HR-LR)



HR- NN emulated(LR)



HR- seasonal_clim_corr(LR)



NN can strongly reduce error for SST over the test period

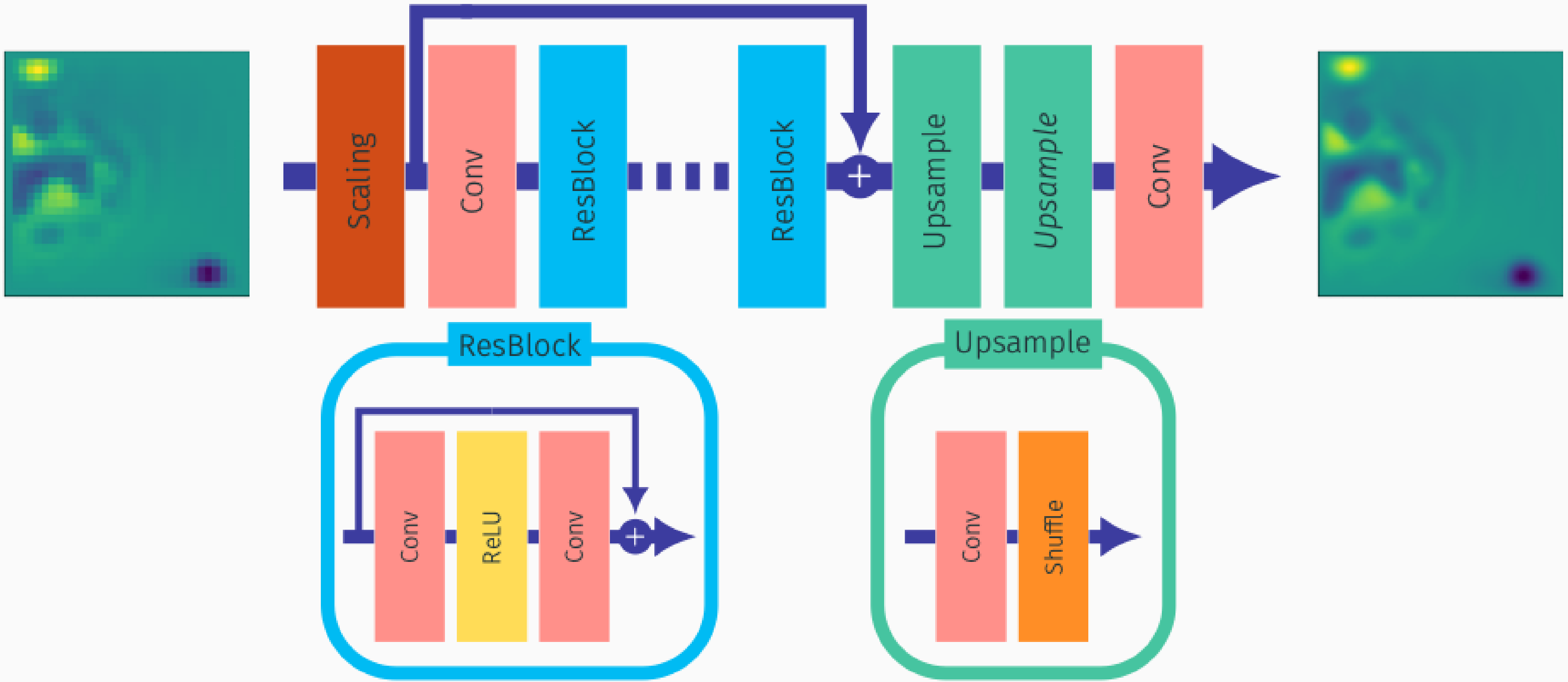
Conclusions

- We build an emulator for model resolution increase
  - Benefits:
    - We can train the emulator robustly on an extended multivariate data set
    - It is cheap if the HR simulation already exist
    - Less to learn since the model between LR and HR is the same
    - Can directly be used within the DA system
  - Limitation:
    - No guarantee that the HR model is best
- It can approaches the performance of the HR DA system at a fraction of the computational cost
- Hybridizing few dynamical members with a large ensemble of emulated LR member shows best performance
- We are testing the appoach for ESM in idealise twin experiment

Future perspectives

- We will test the added value of the SRDA for seasonal predictions:
  - Is it best to train the NN on full 3D corrections or only on SST and use DA to formulate multivariate corrections ?
  - Would the emulator train under pre-industrial configuration also work for the recent period ?
  - How far can we push the resolution increase ?

Machine Learning training



Architecture of the enhanced deep super-resolution network (EDSR) [Lim et al., 2017]