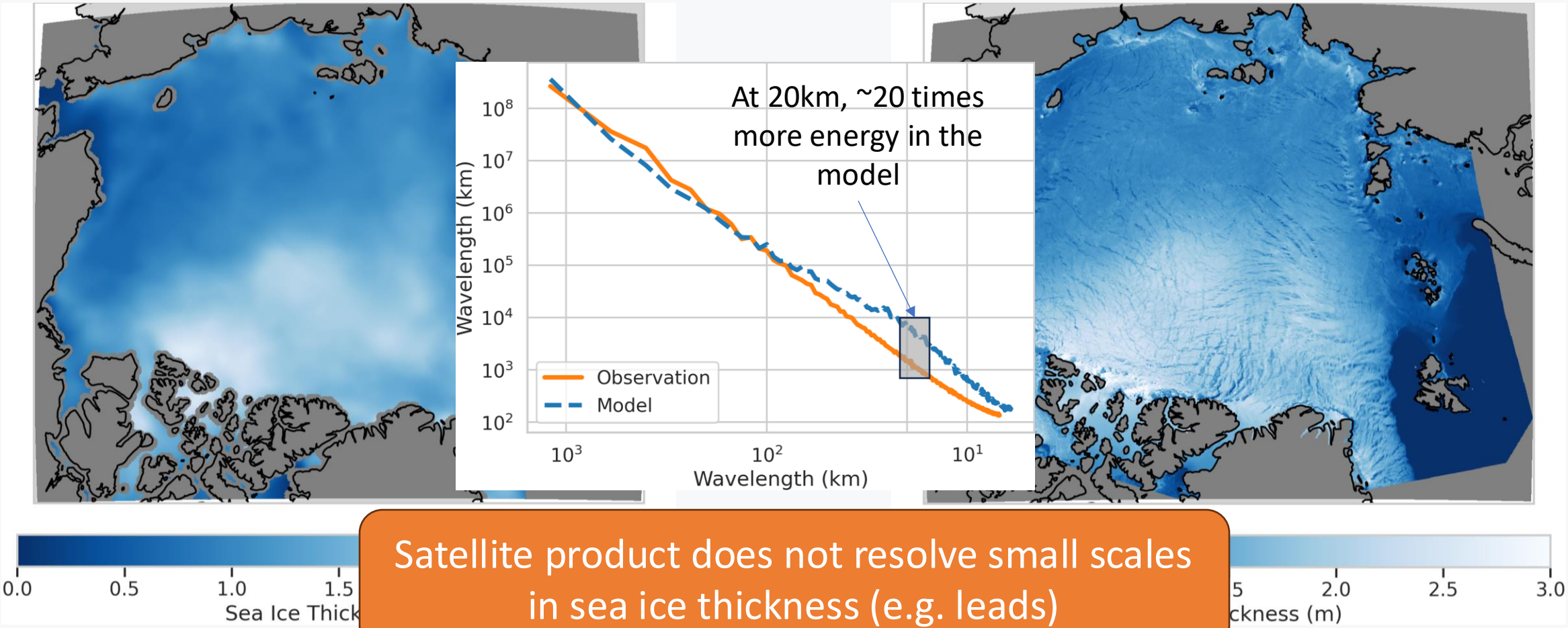
Super-resolution of satellite observations of sea ice thickness using diffusion models and physical modeling

Julien Brajard, Fabio Mangini, Anton Korosov, Yiguo Wang, Richard Davy

Motivation

Satellite observation product (CS2SMOS)

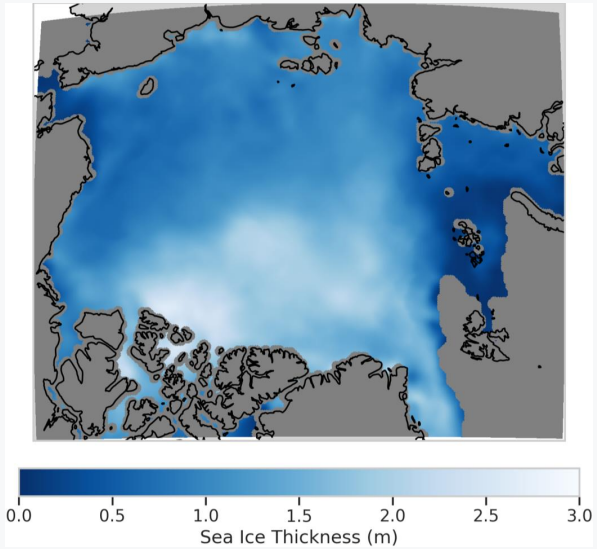
Physical model (NeXtSIM) forecast



Satellite product does not resolve small scales in sea ice thickness (e.g. leads)

Our Objective: downscaling

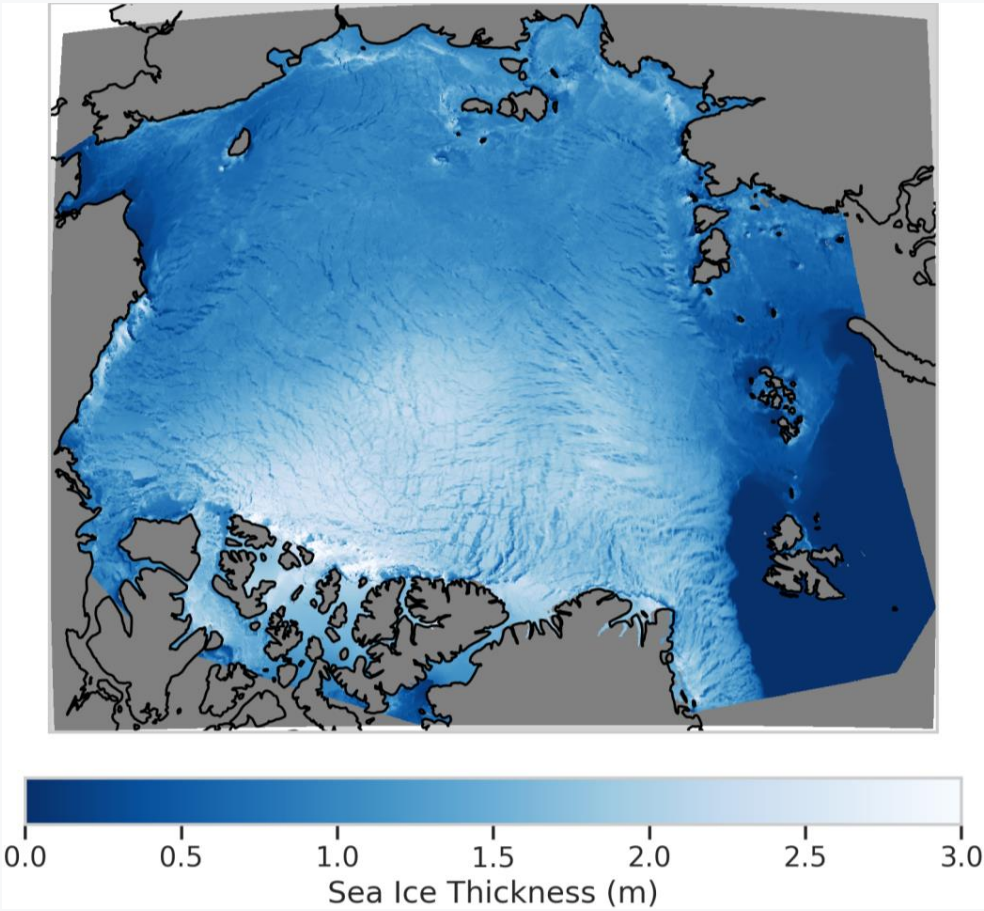
Low-resolution observation



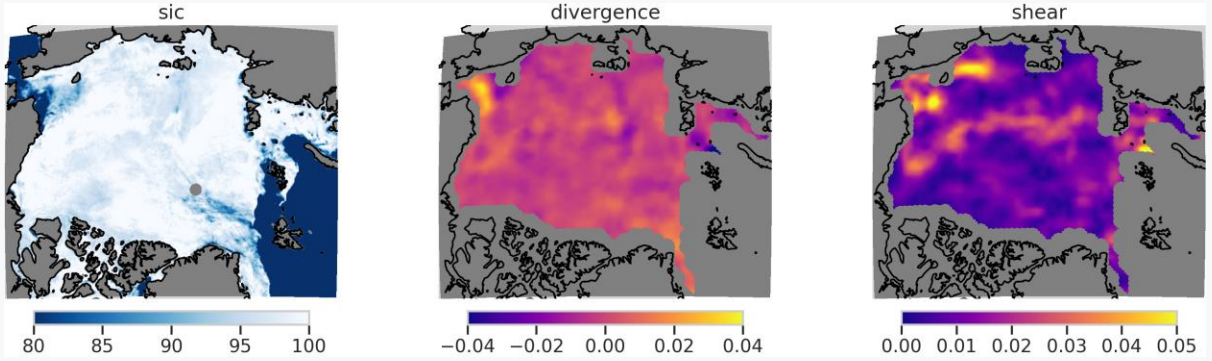
Downscaling Model

Neural network

High-resolution sea ice thickness



Other low-resolution observations



Our Objective: probabilistic downscaling

x

Low-resolution observations

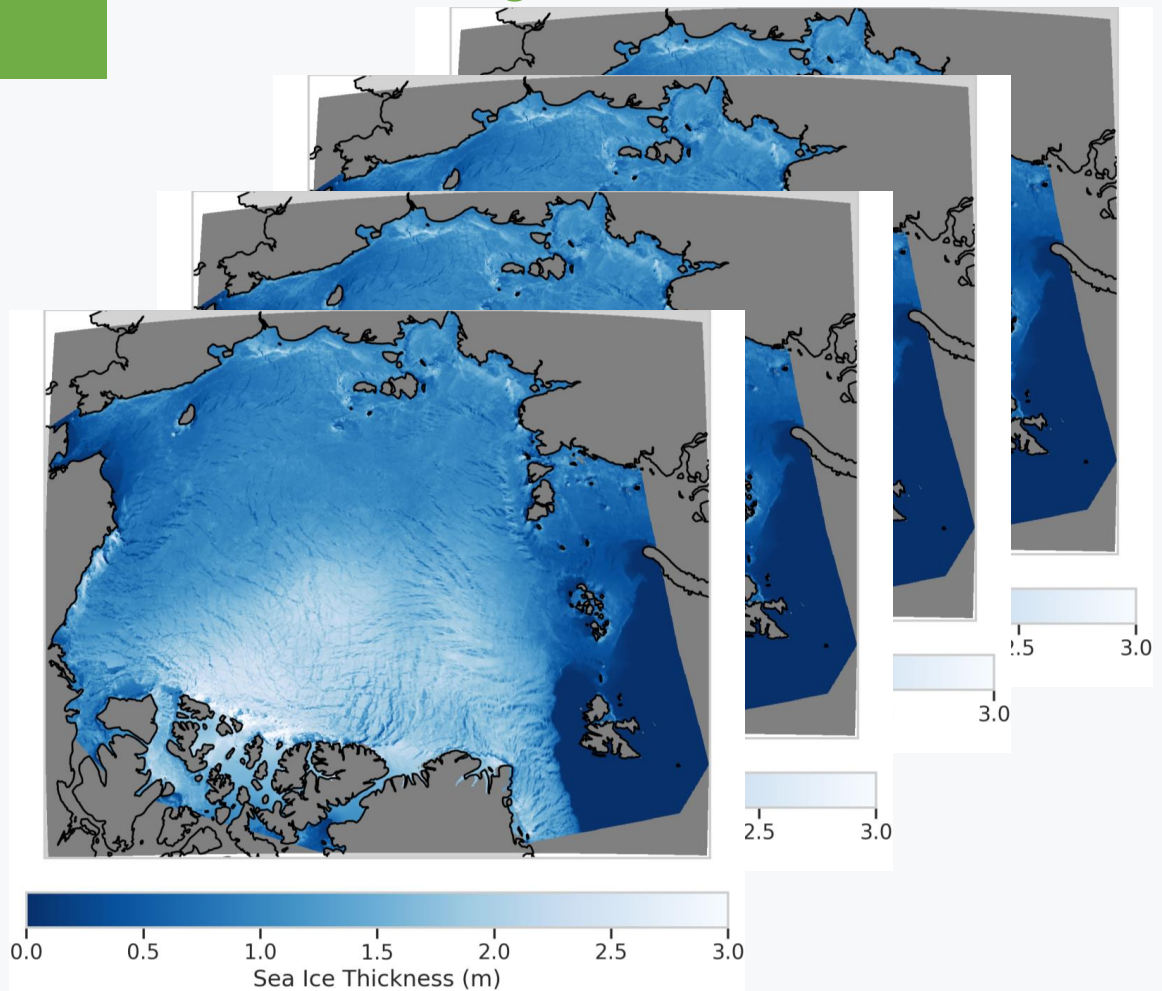
Downscaling Model

y

High-resolution sea ice thickness

Deterministic downscaling: $y = f(x)$

Probabilistic downscaling: $P(y|x)$



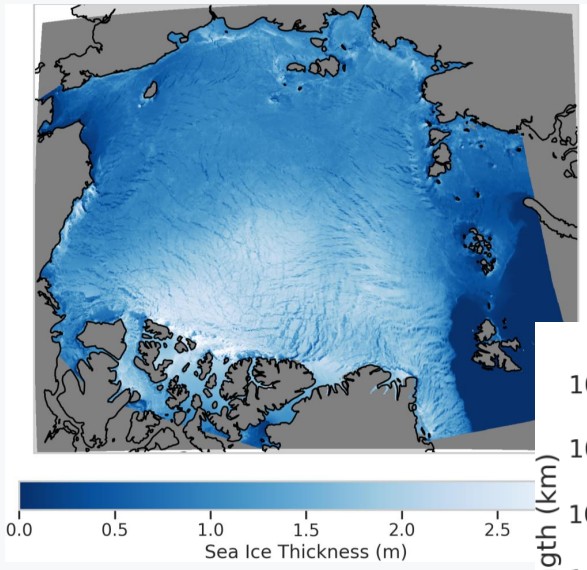
What do we need?

- ✓ A training set of matching pairs of low-resolution/high-resolution fields
- ✓ A probabilistic model
- ✓ Relevant metrics for validation
- ✓ Apply to observation

Dataset constitution

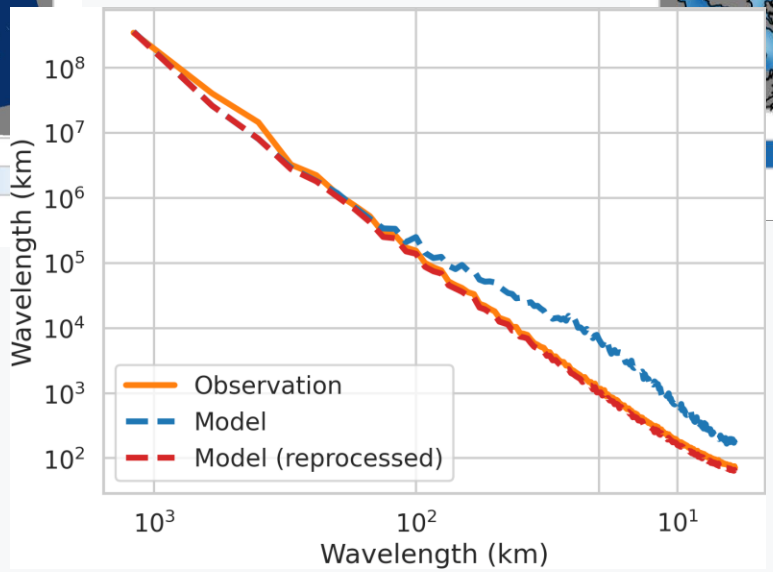
Principle: Using **high-resolution NeXtSIM simulations** [Ólason et al., 2022] and process them to match the resolution of **the observations**.

NeXtSIM sea ice thickness (res ~3km)

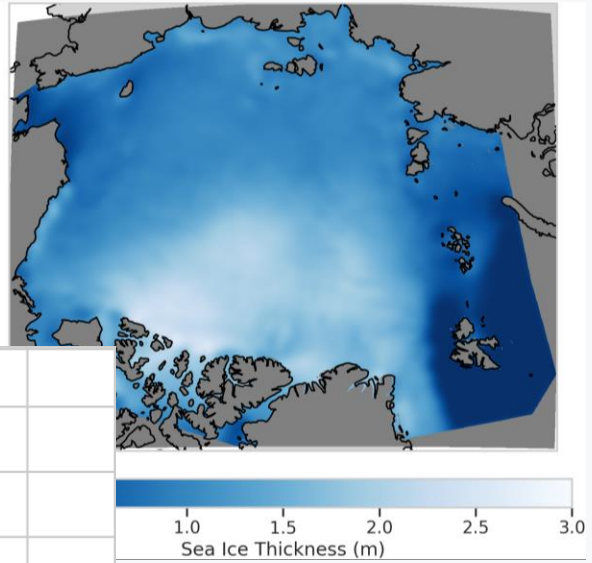


y

Smoothing with a Gaussian kernel
(size 33 km)

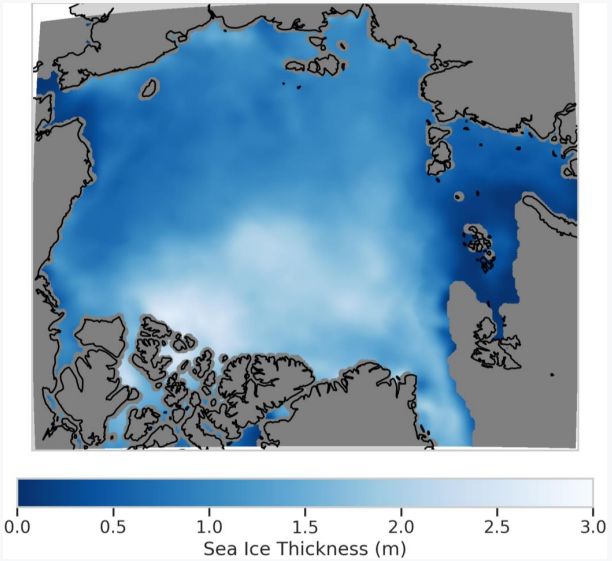


Reprocessed neXtSIM



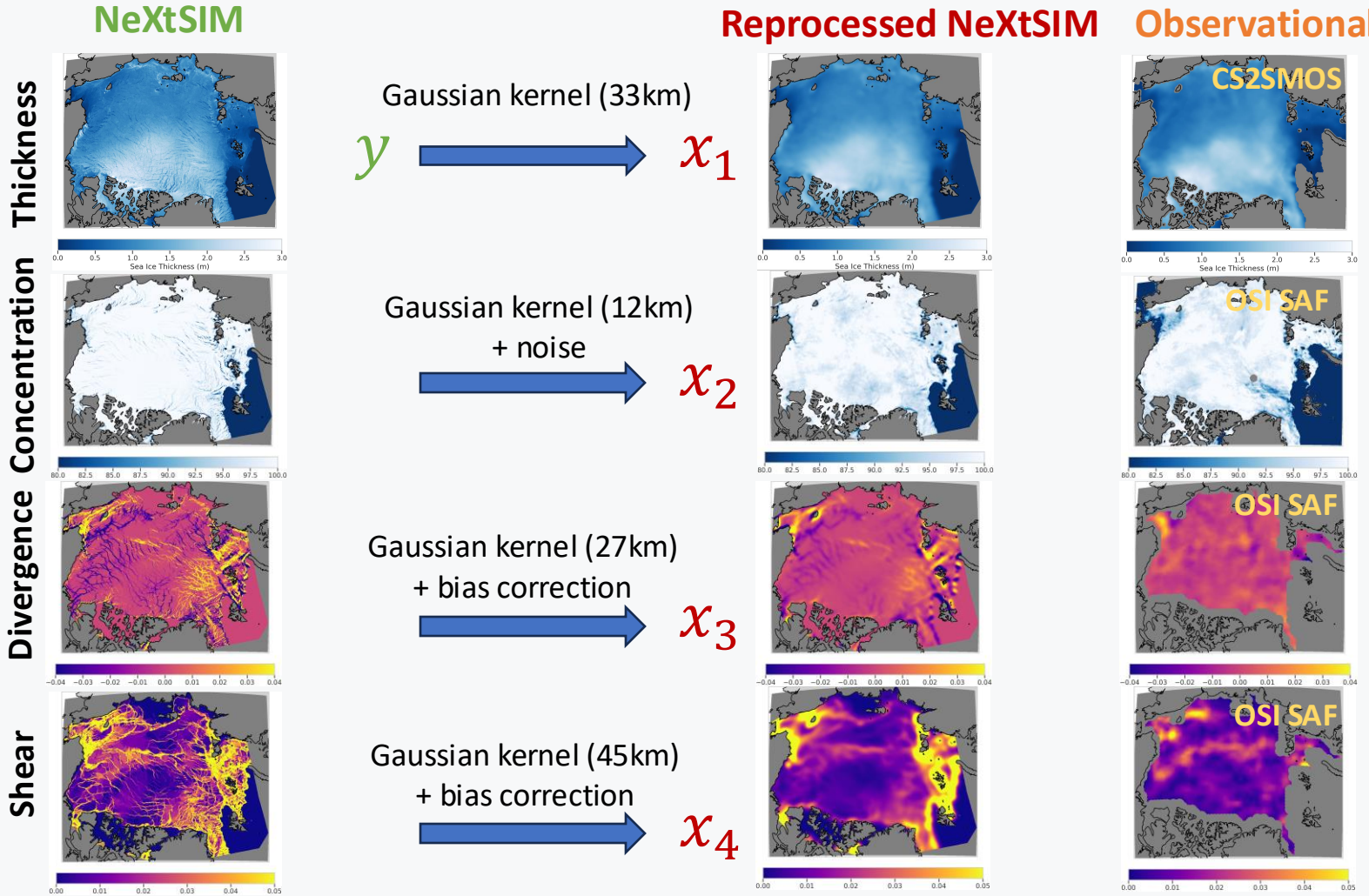
x

CS2SMOS (observational product)



Dataset constitution

Same procedure for Sea Ice concentration, divergence and shear (to be used as input feature)



Dataset: $([x_1, x_2, x_3, x_4], y)$

- ✓ Divergence and Shear are transformed into the total deformation
- ✓ A land mask is added
- ✓ Samples in freezing season:
  - ✓ Training: 2013-2020 (1157 samples)
  - ✓ Validation: 2020-2022 (360 samples)
  - ✓ Test: 2022-2023 (180 samples)

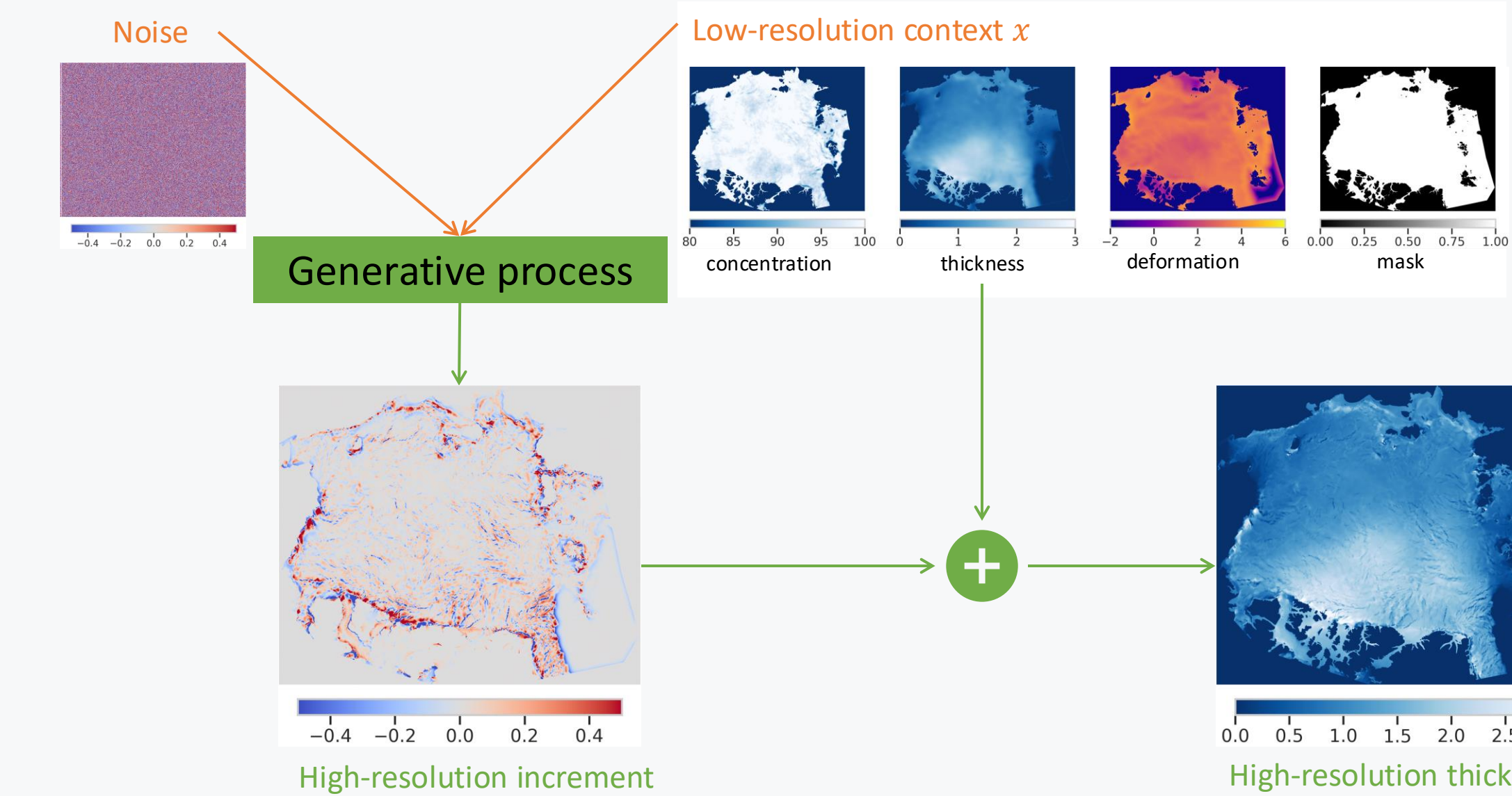
Download the dataset

https://archive.norstore.no/pages/public/datasetDetail.jsf?id=10.11582/2024.00126

What do we need?

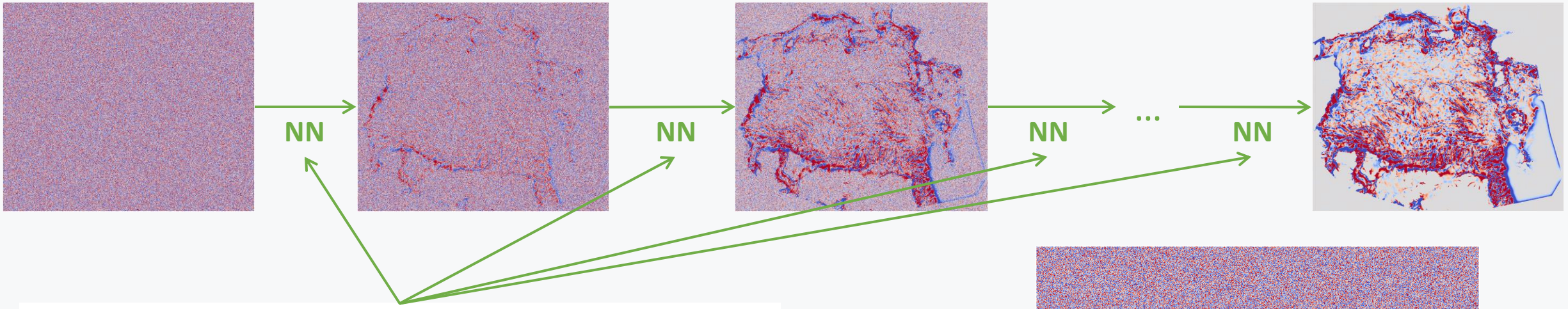
- ✓ A training set of matching pairs of low-resolution/high-resolution fields
- ✓ A probabilistic model
- ✓ Relevant metrics for validation
- ✓ Apply to observation

Applying the diffusion model to sea ice super-resolution

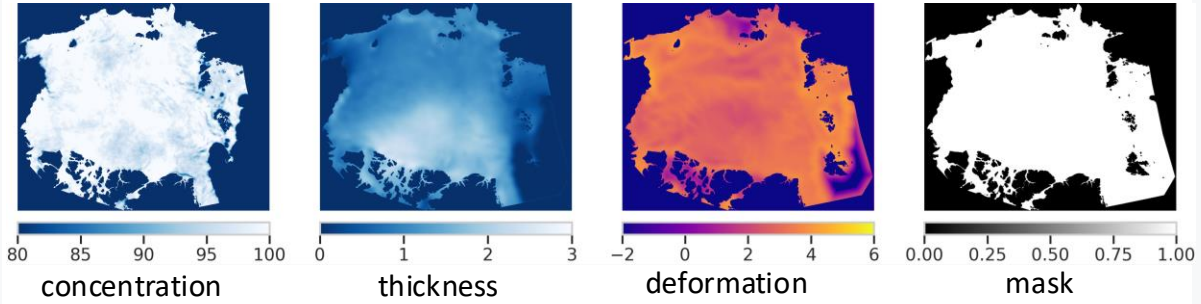


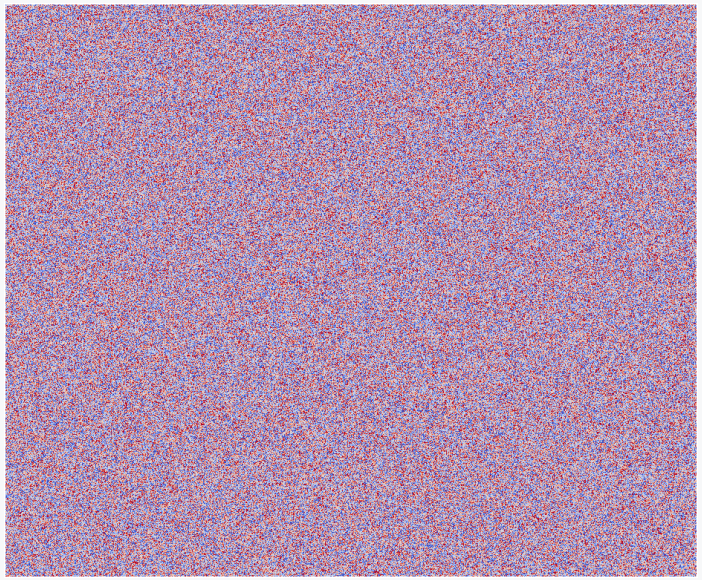
Diffusion models – how do they work?

A **neural network** as a recursive denoiser

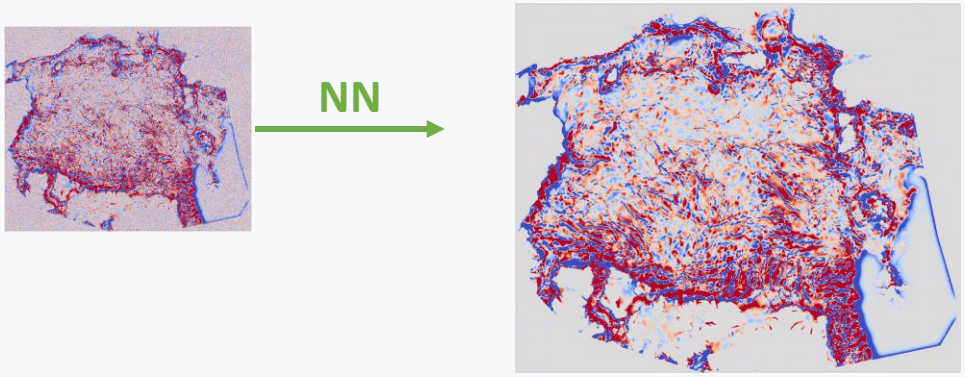


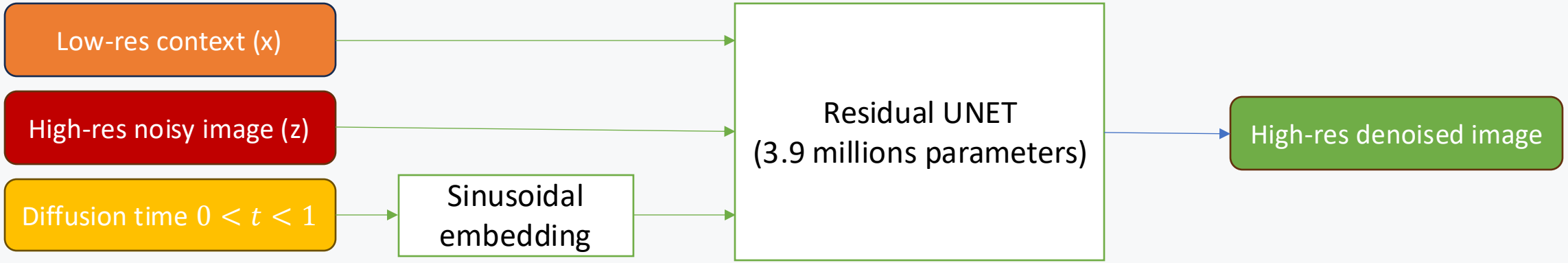
Low-resolution context x





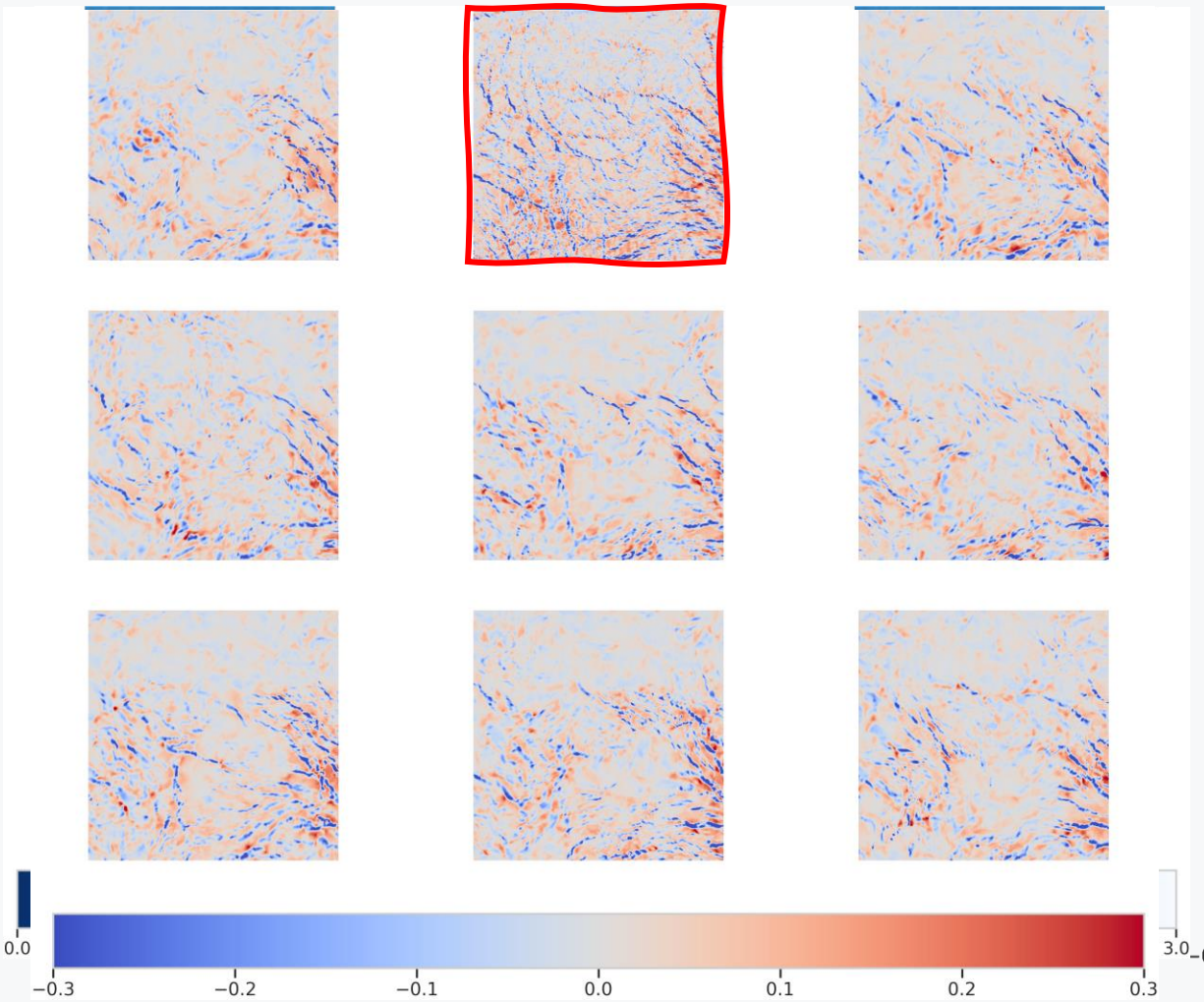
Implementation details

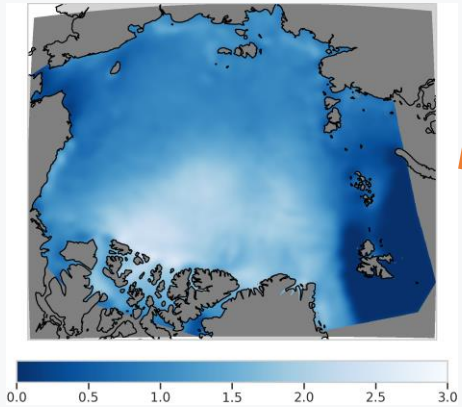




Generation January 1, 2021

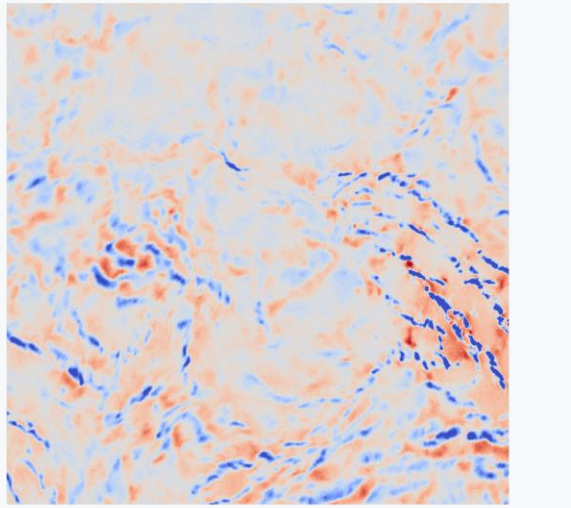
Generated ensemble of sea ice thickness

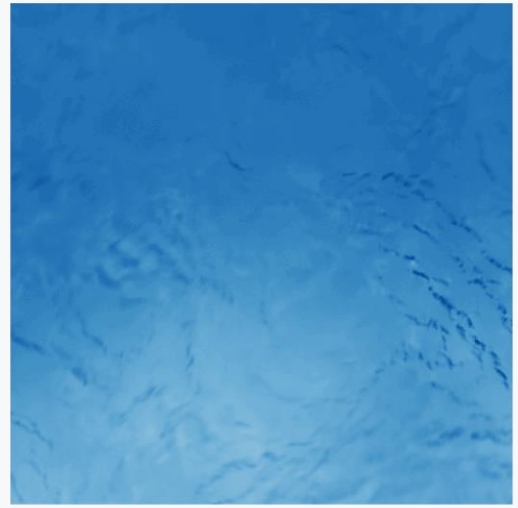




From the low-resolution tickness

SIT mem 0 - 20210101



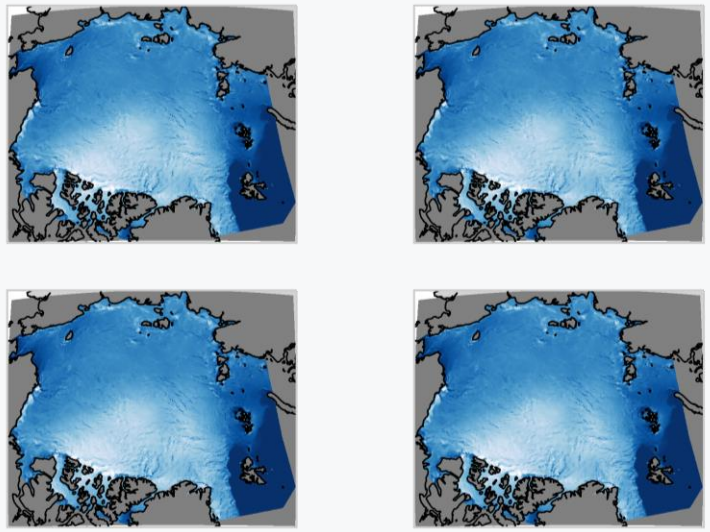


What do we need?

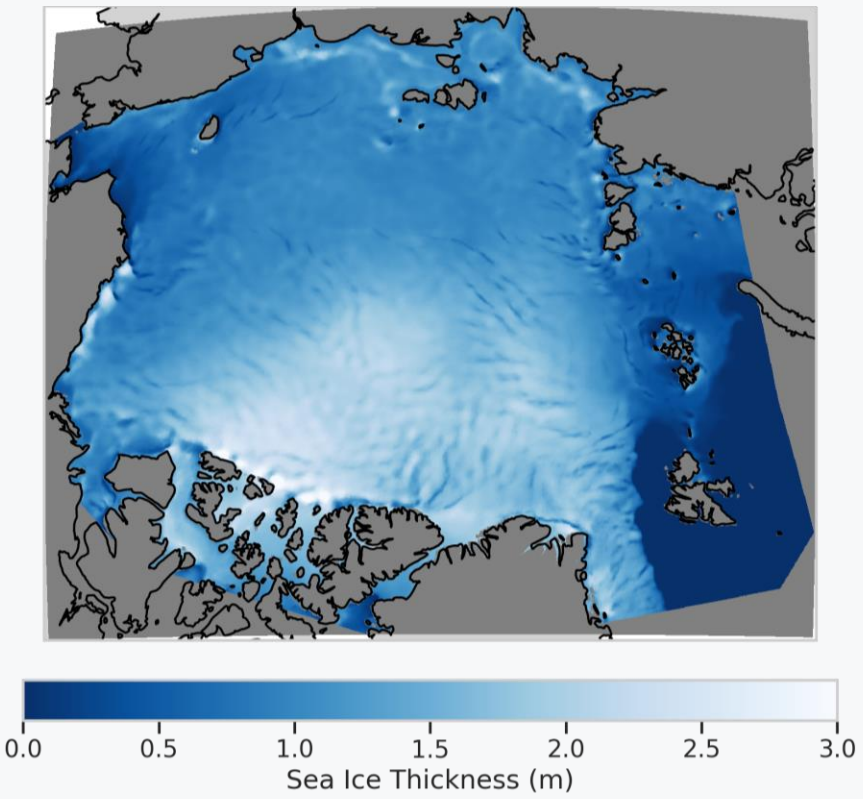
- ✓ A training set of matching pairs of low-resolution/high-resolution fields
- ✓ A probabilistic model
- ✓ Relevant metrics for validation
- ✓ Apply to observation

Different “products”

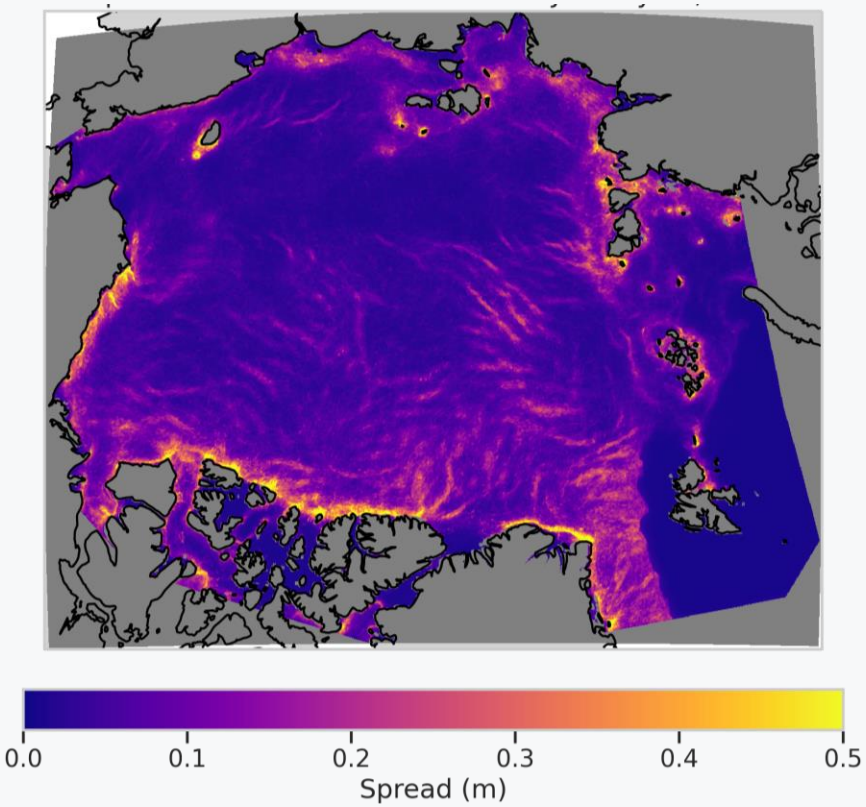
Individual members



Ensemble mean



Spread



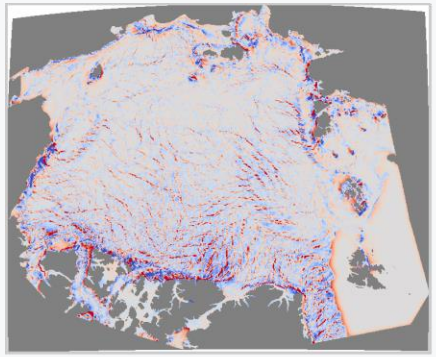
Used to assess
realism

Used to assess
accuracy

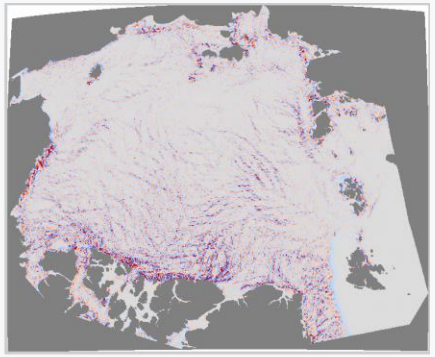
Used to assess
uncertainty

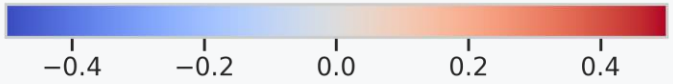
Accuracy of the super-resolution

Error low-resolution



Error AI ensemble mean



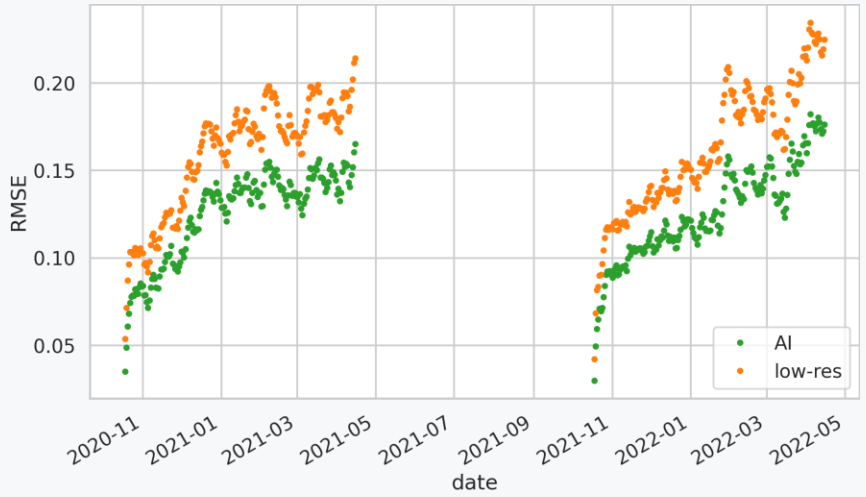


Root-mean square error (RMSE) of:

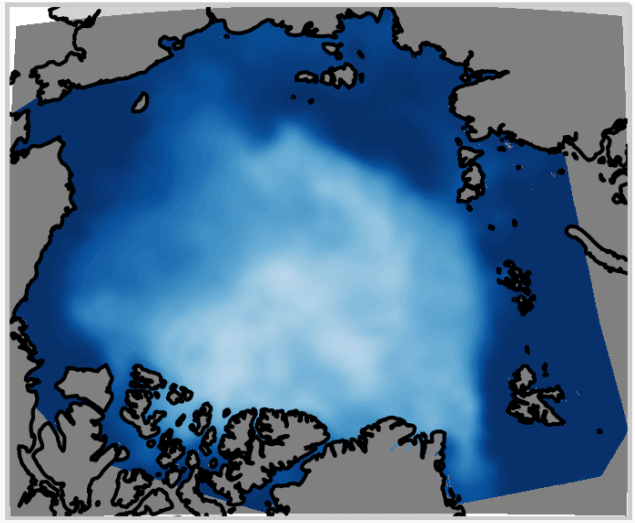
Low-resolution: 0.16 m

AI product: 0.13 m

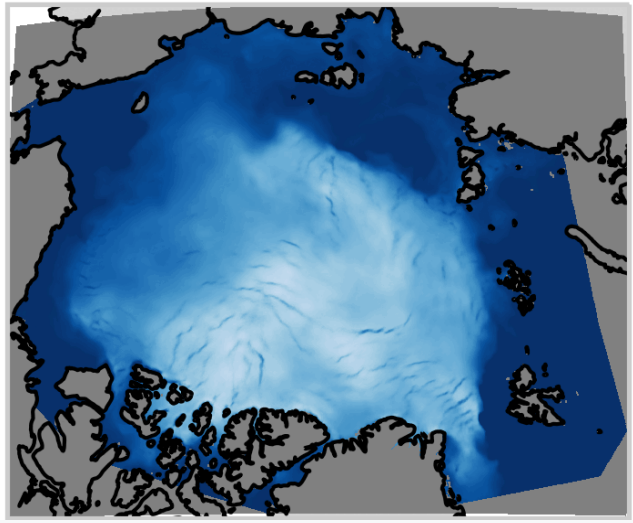
Improvement: **20%**



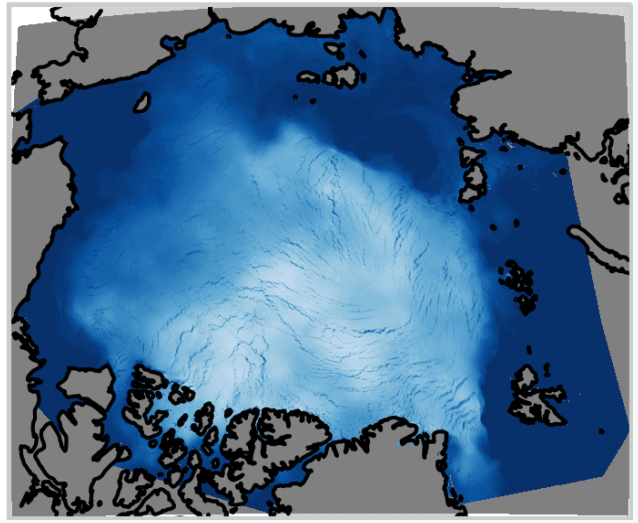
SIT low-res 20211022



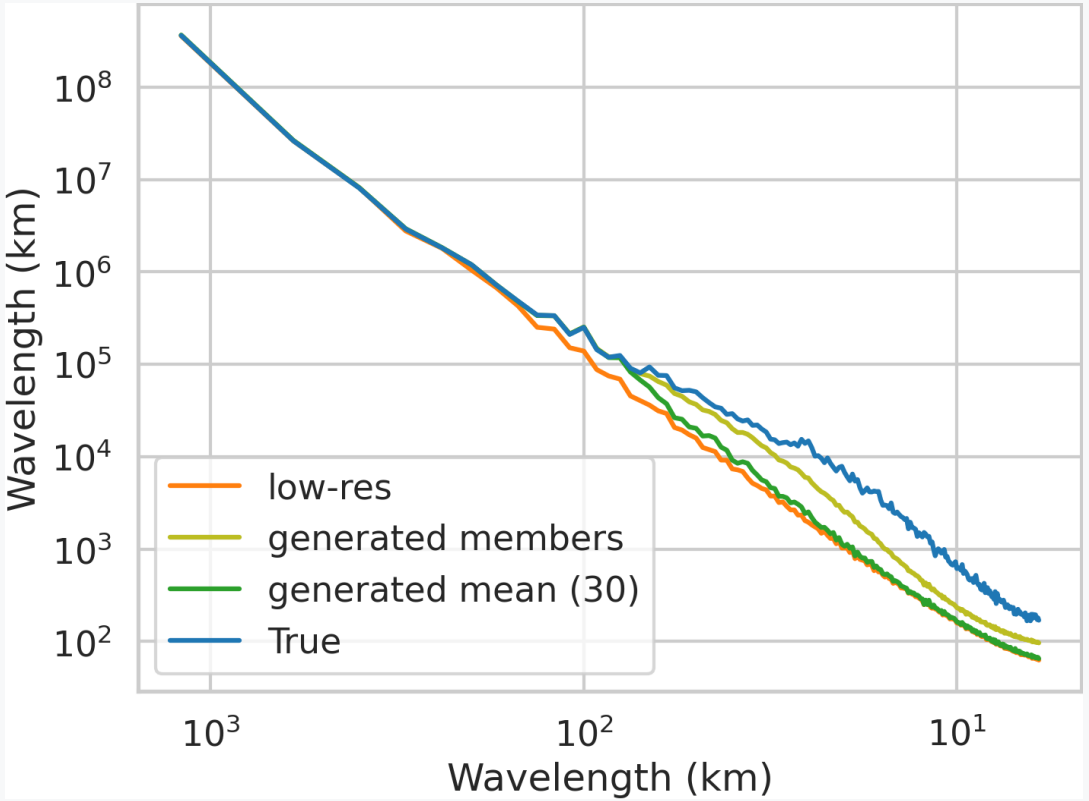
SIT AI 20211022

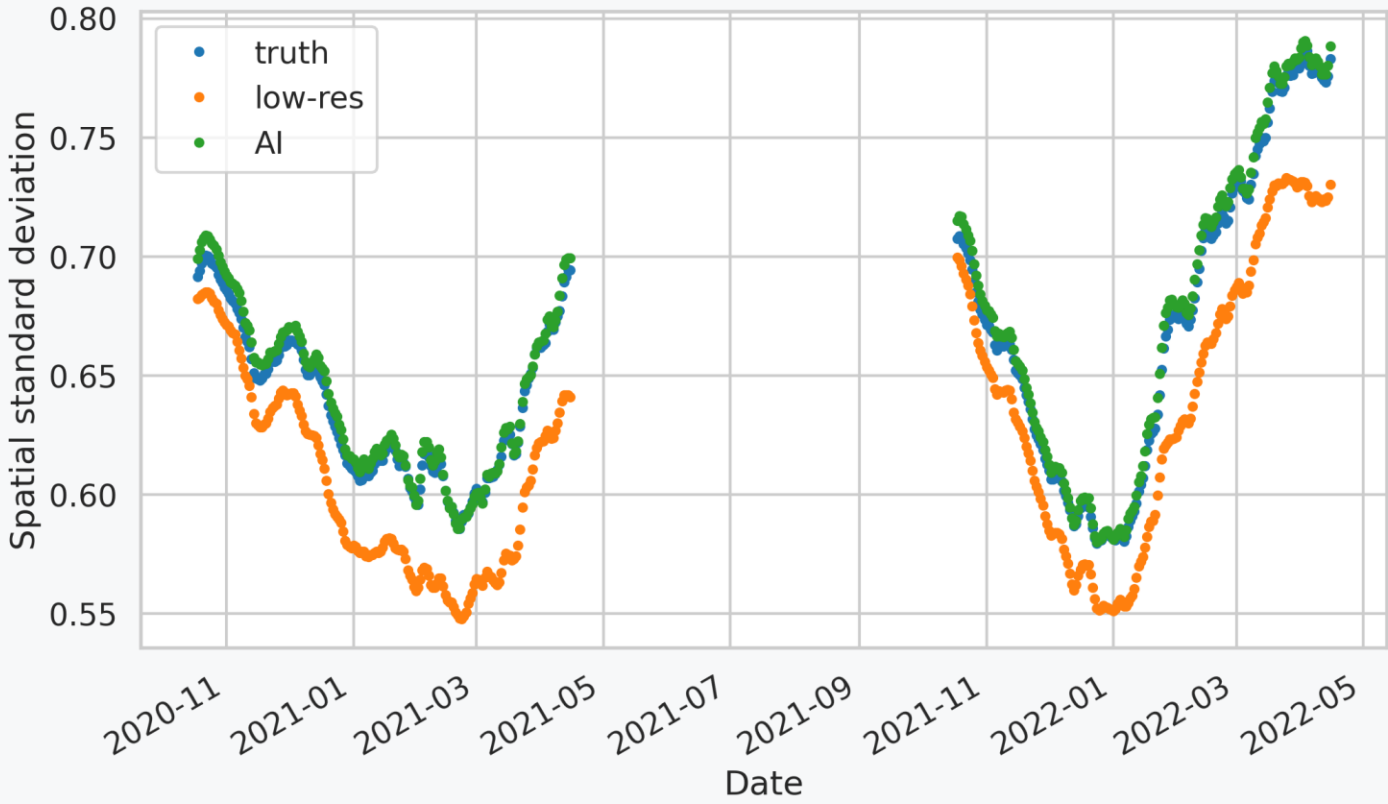


Ref 20211022



Realism


Wavelength (km)
10^8
10^7
10^6
10^5
10^4
10^3
10^2
low-res
generated members
generated mean (30)
True
10^3
10^2
10^1
Wavelength (km)


Spatial standard deviation
0.80
0.75
0.70
0.65
0.60
0.55
truth
low-res
AI
2020-11
2021-01
2021-03
2021-05
2021-07
2021-09
2021-11
2022-01
2022-03
2022-05
Date

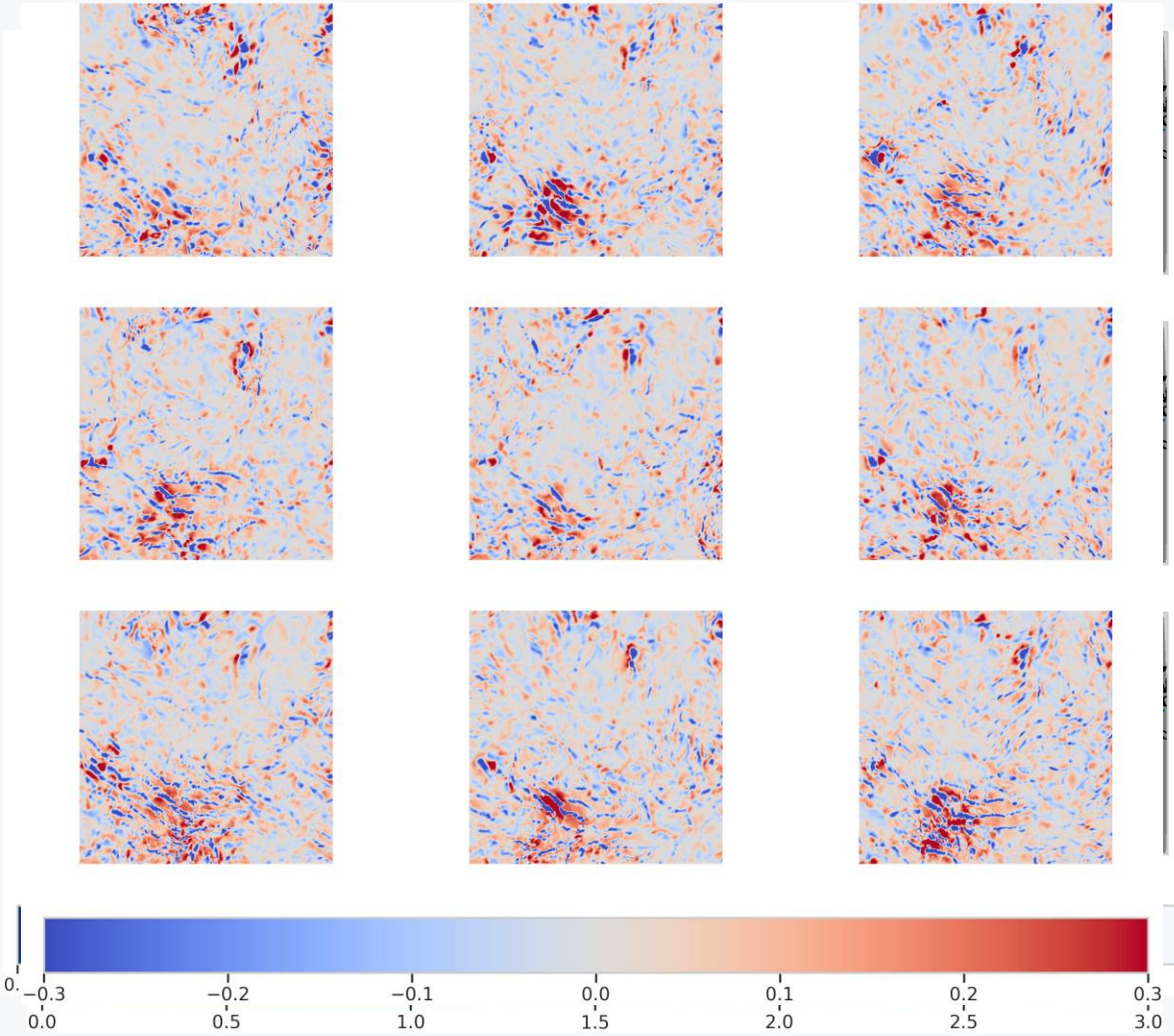
What do we need?

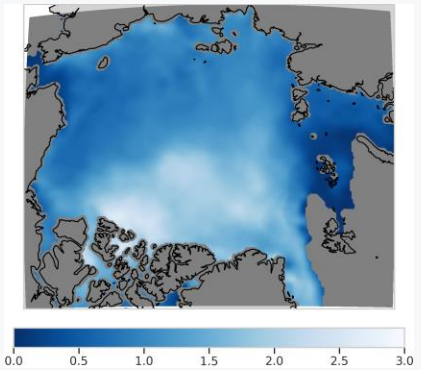
- ✓ A training set of matching pairs of low-resolution/high-resolution fields
- ✓ A probabilistic model
- ✓ Relevant metrics for validation
- ✓ Apply to observation

Generation from observations

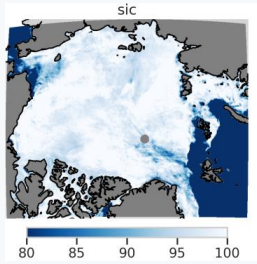
From the low-resolution thickness (CS2SMOS)

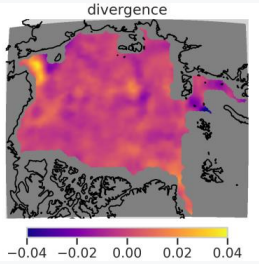
Generated ensemble of sea ice thickness

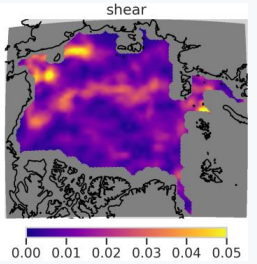




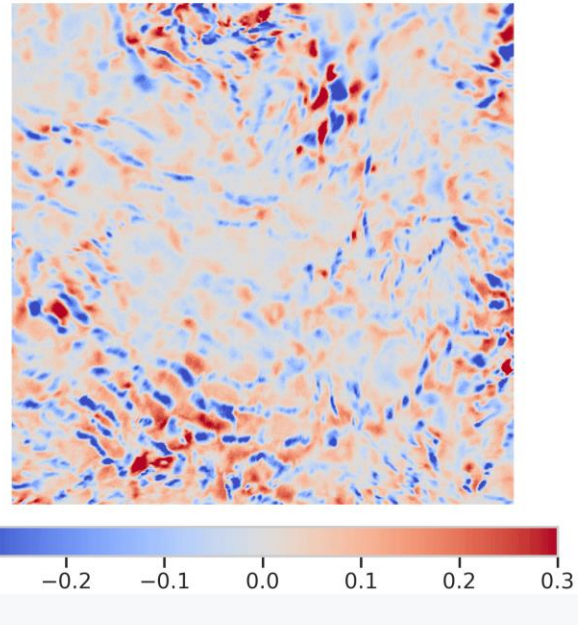
+ other observations

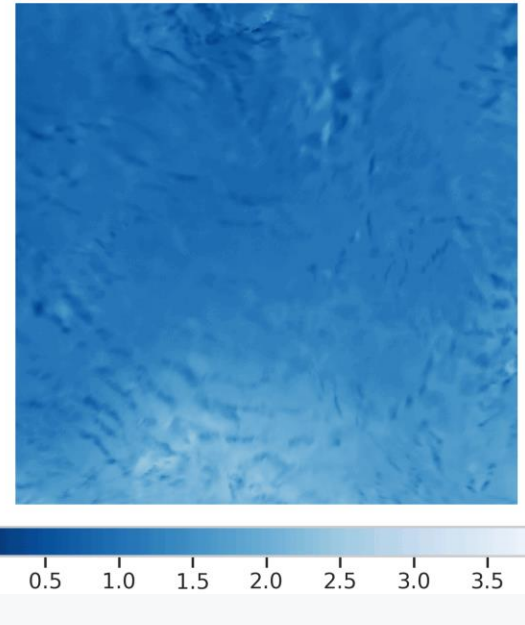






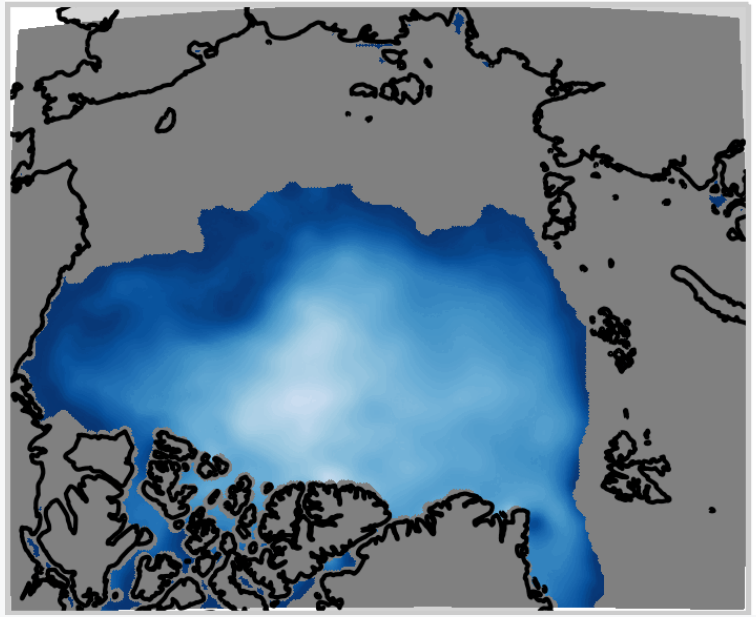
SIT mem 0 - 20210101



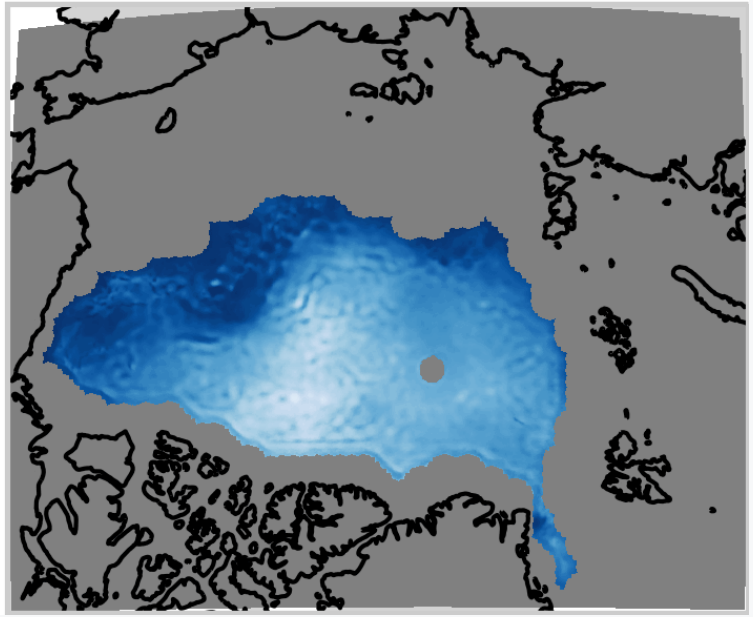


Observations 2020-2021

SIT low-res 20201022



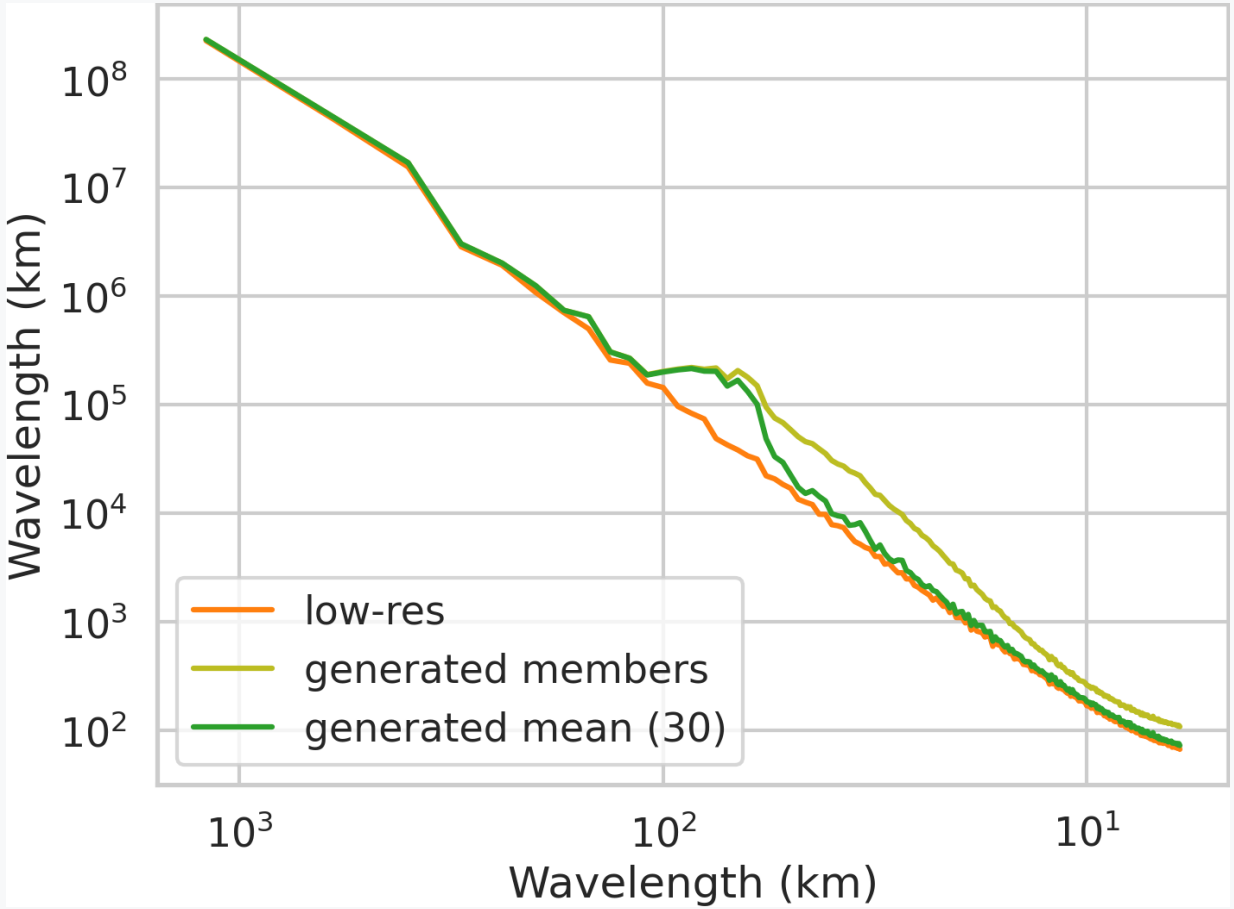
SIT AI 20201022



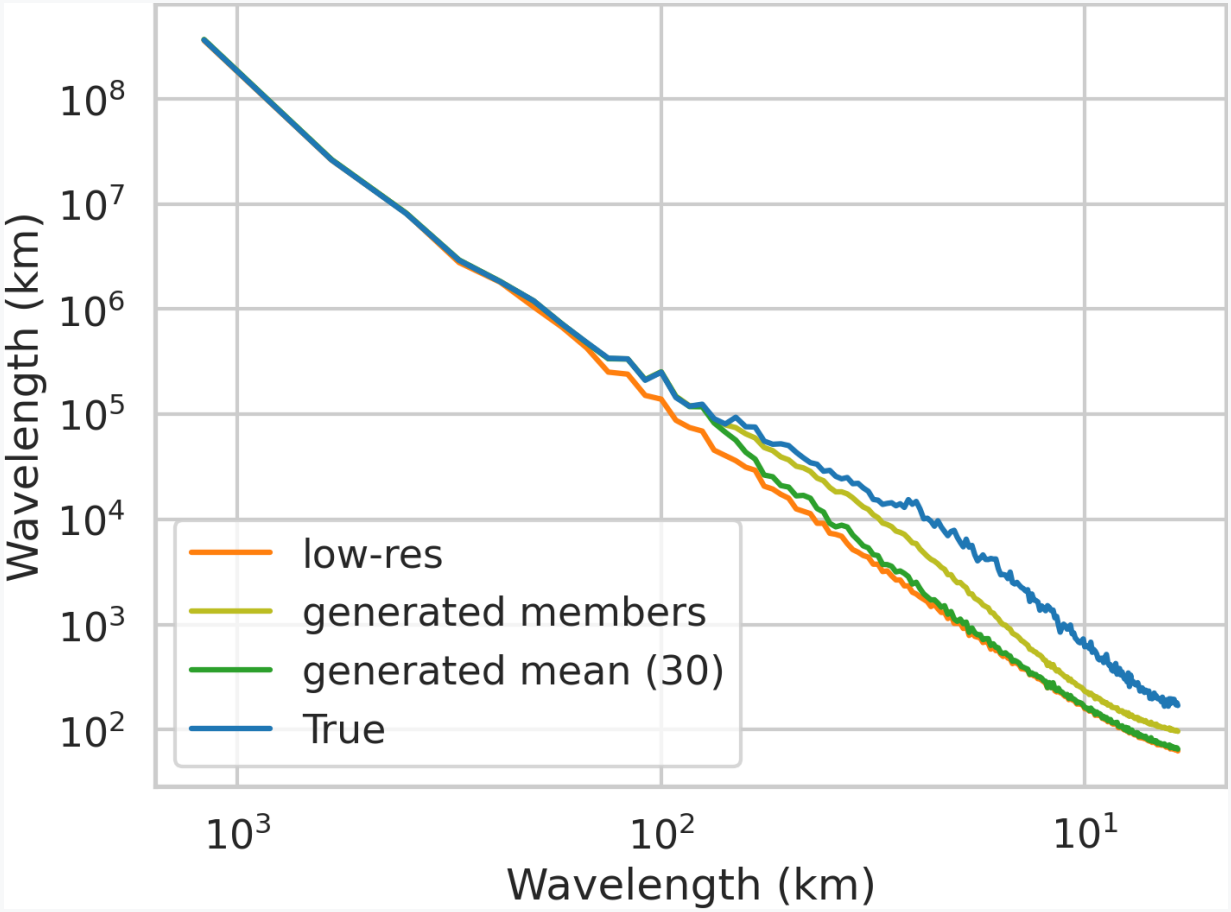
See the next presentations to evaluate the potential of this product

Observation spectrum

Spectrum of the observations reconstruction

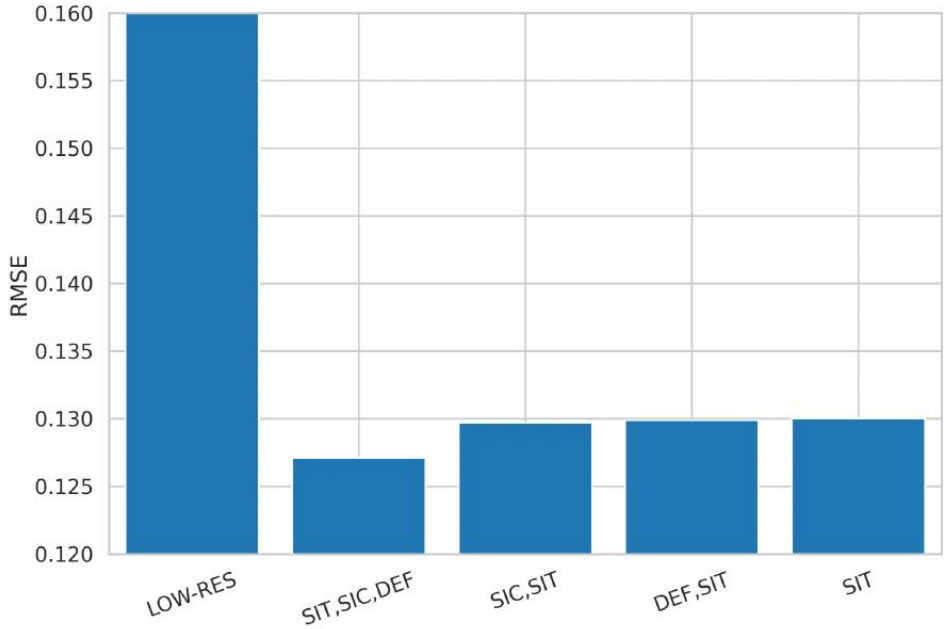


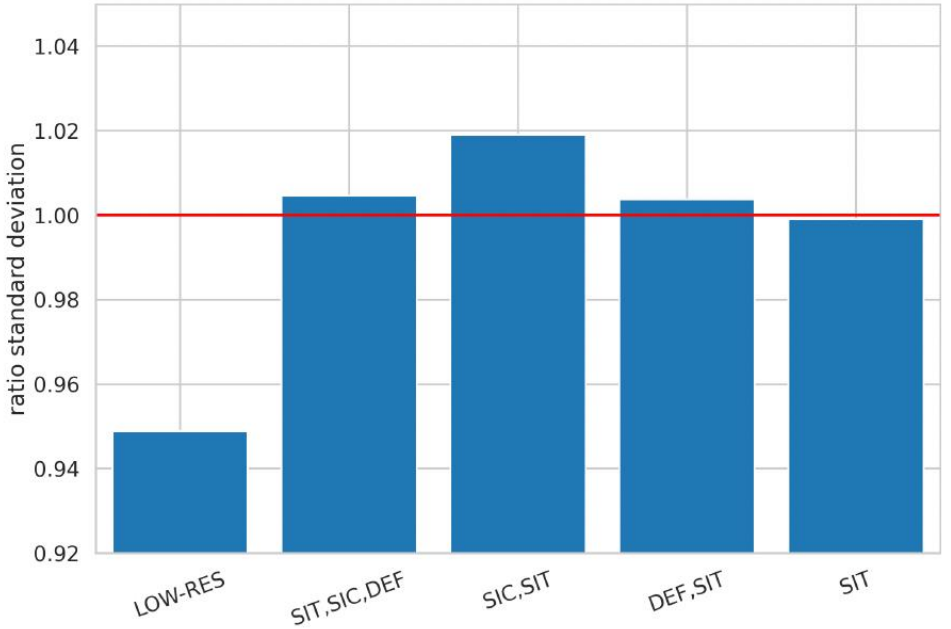
Spectrum of the NeXtSIM reconstruction



Input features

Trial	inputs
11	SIC, SIT, DEF
6	SIC, SIT
7	DEF, SIT
8	SIT





Take-home message

- **Diffusion models can be used to generate accurate and realistic high-resolution sea ice thickness fields**
  - Better accuracy and better realism compared with low-resolution field

- **A model trained on a realistic physical simulations can be applied, without retraining, on observations** (a few artifacts can appear)

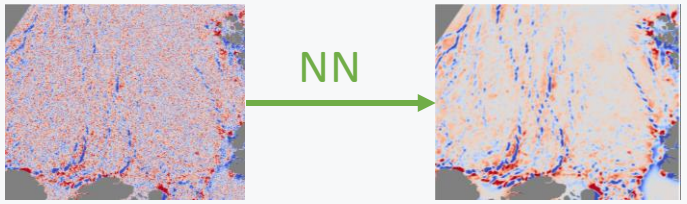
- **Super-resolution using diffusion models can be applied to other sea ice variables** (actually, any geophysical variable)

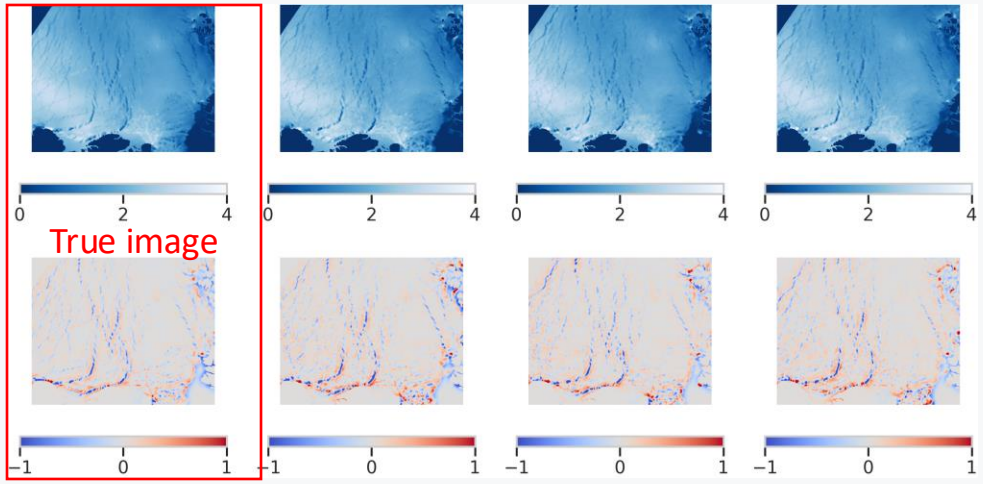
- The dataset (both physical simulations and AI generation) is **available** for download

Contact me! → Julien.brajard@nersc.no

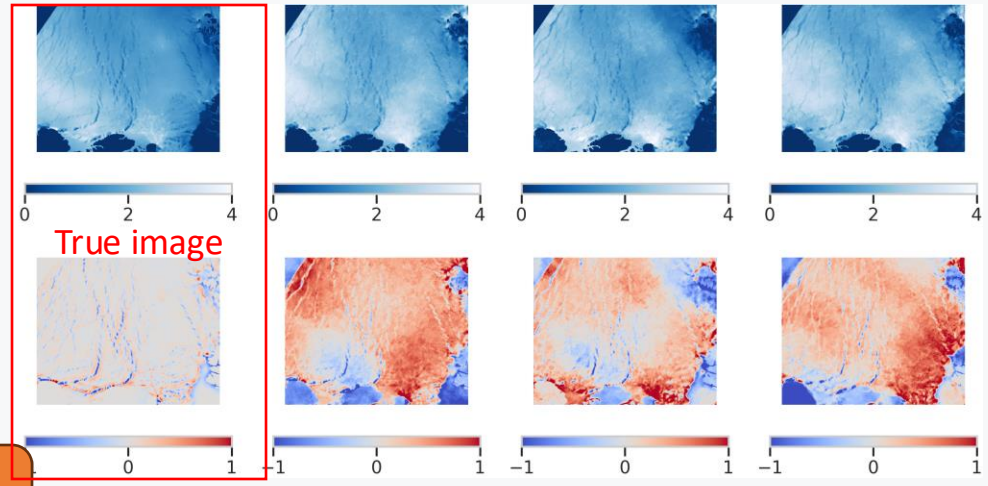
Anomaly Vs full field generation

Anomaly generation





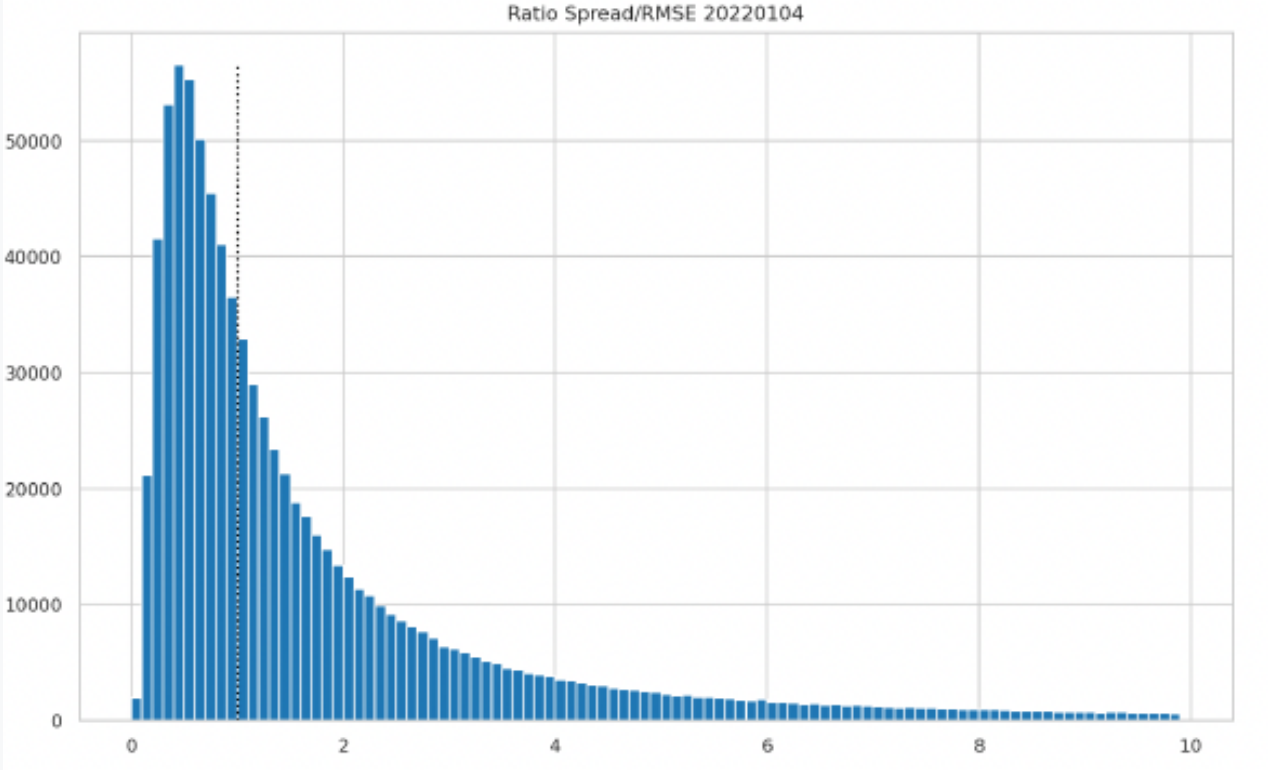
Full-field generation

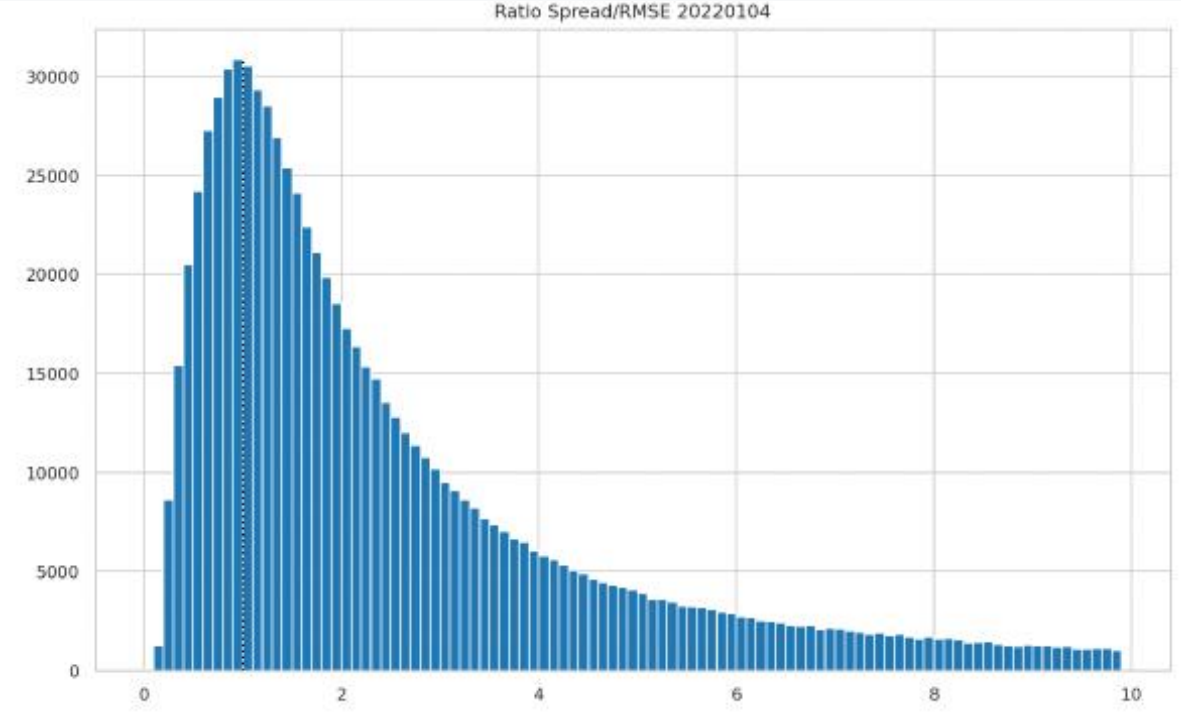


Full-field induces large-scale biases

Ensemble score

Ratio Spread / RMSE





Another training with only thickness and concentration in the context

Input features