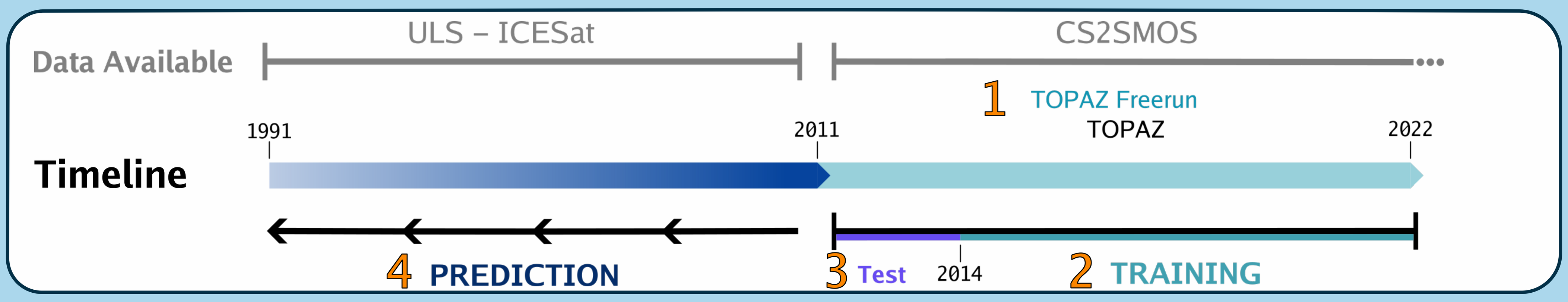
Reconstruction of Arctic sea ice thickness (1991–2010) based on a hybrid machine learning and data assimilation approach

Léo Edel[1], Jiping Xie[1], Calliopé Danton Laloy[1], Julien Brajard[1], Laurent Bertino[1]

[1] NERSC, Bergen, Norway

TARDIS



Baseline: simple correction used to evaluate our model

Monthly bias: TOPAZ – TOPAZ Freerun

The SIT from TOPAZ freerun is corrected daily by adding the monthly bias computed on 2014–2022.

1. Objectives and bias of SIT

Correct the sea ice thickness (SIT) by learning the bias of SIT between TOPAZ assimilating CS2SMOS and TOPAZ freerun between 2011–2022, assisted by additional information from environmental variables.
Prediction backward in time prior to 2011.

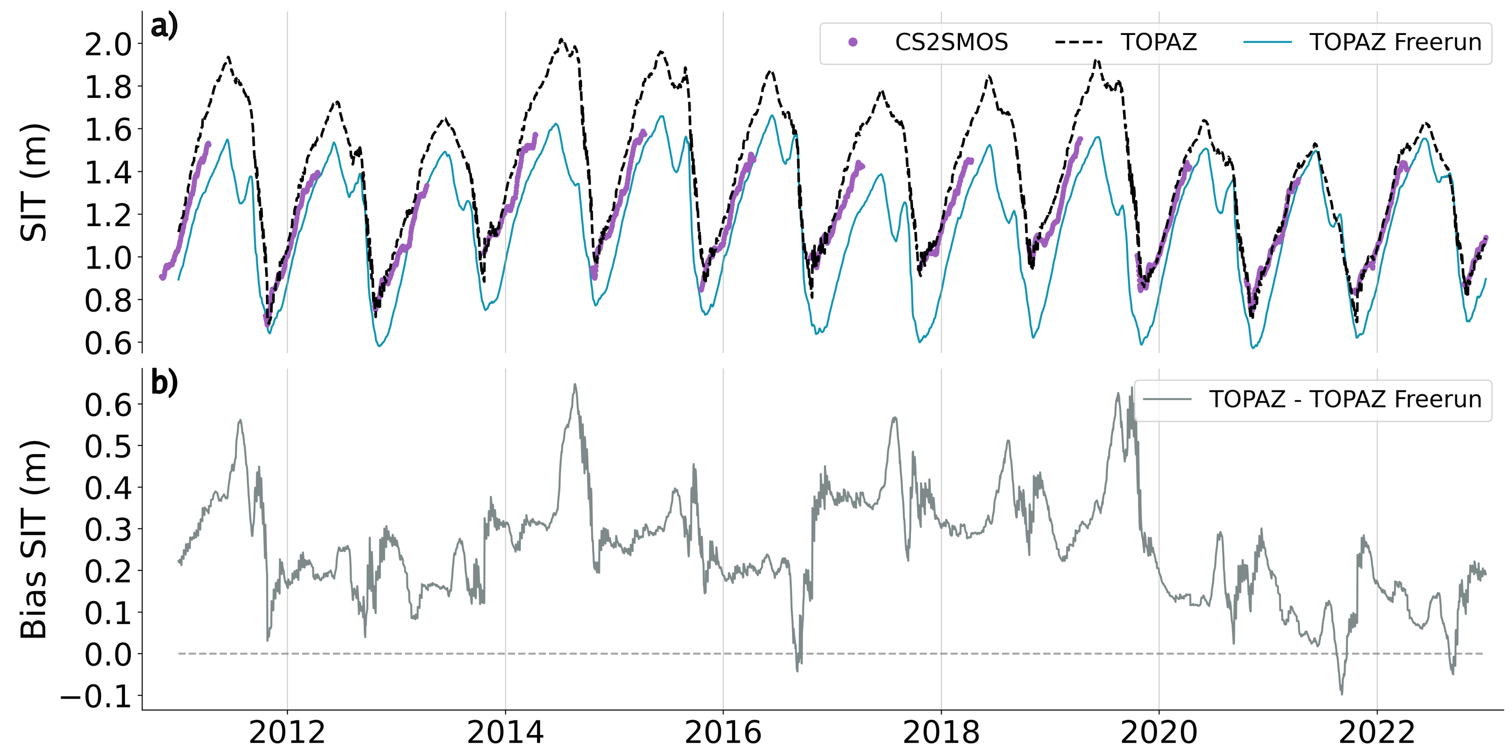
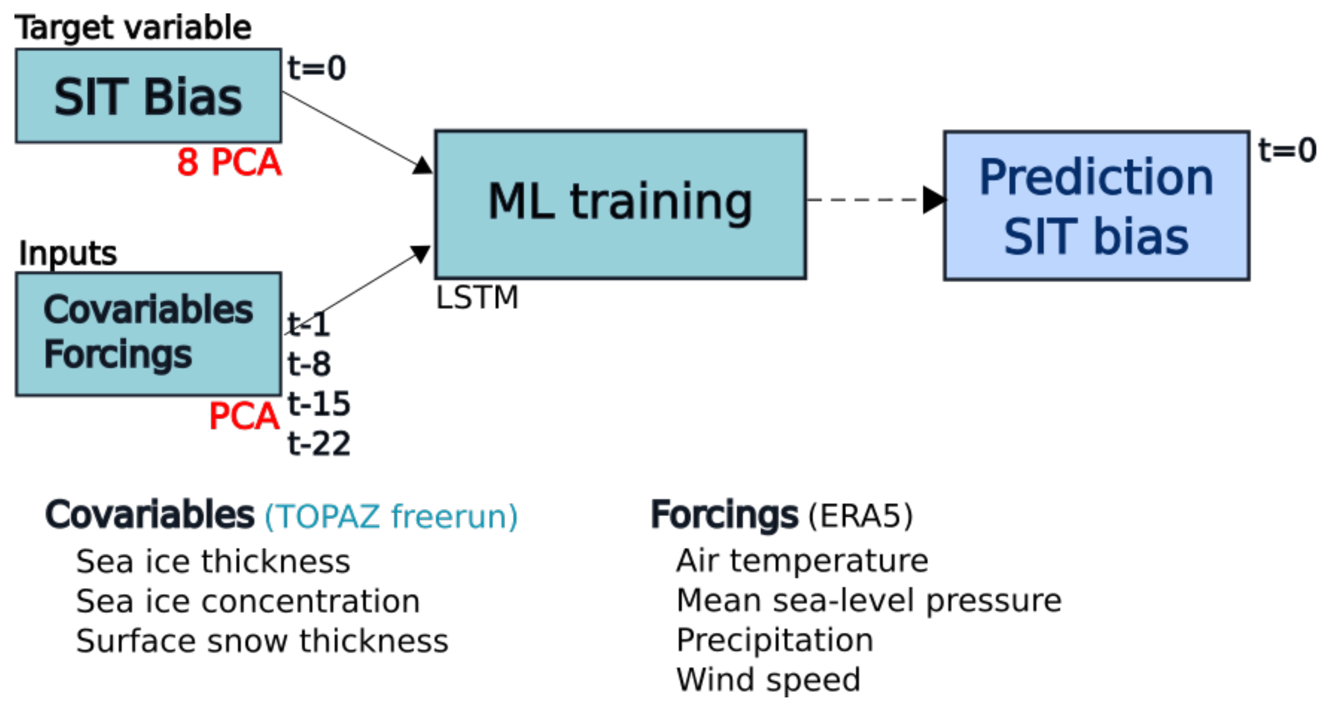


Fig. 1. a) Daily sea ice thickness (m) averaged over the Arctic for sea ice concentration (SIC) > 15%.
b) Bias of sea ice thickness (m) computed as: TOPAZ–TOPAZ freerun.

2. Method and training

Principal Component Analysis (**PCA**) reduces dimensionality.
Long Short-Term Memory Network (**LSTM**) trained to predict the bias of SIT **backward** in time.



3. Performances over test period

The **RMSE** (Fig. 2) attests that our ML algorithm is performing better than the baseline, with close performance to the optimal EOF capability.

A specific emphasis on the spatial distribution of the error must be acknowledged, as we can distinctly see that the baseline performance is not satisfactory, even when the average appears good (Fig. 4b).

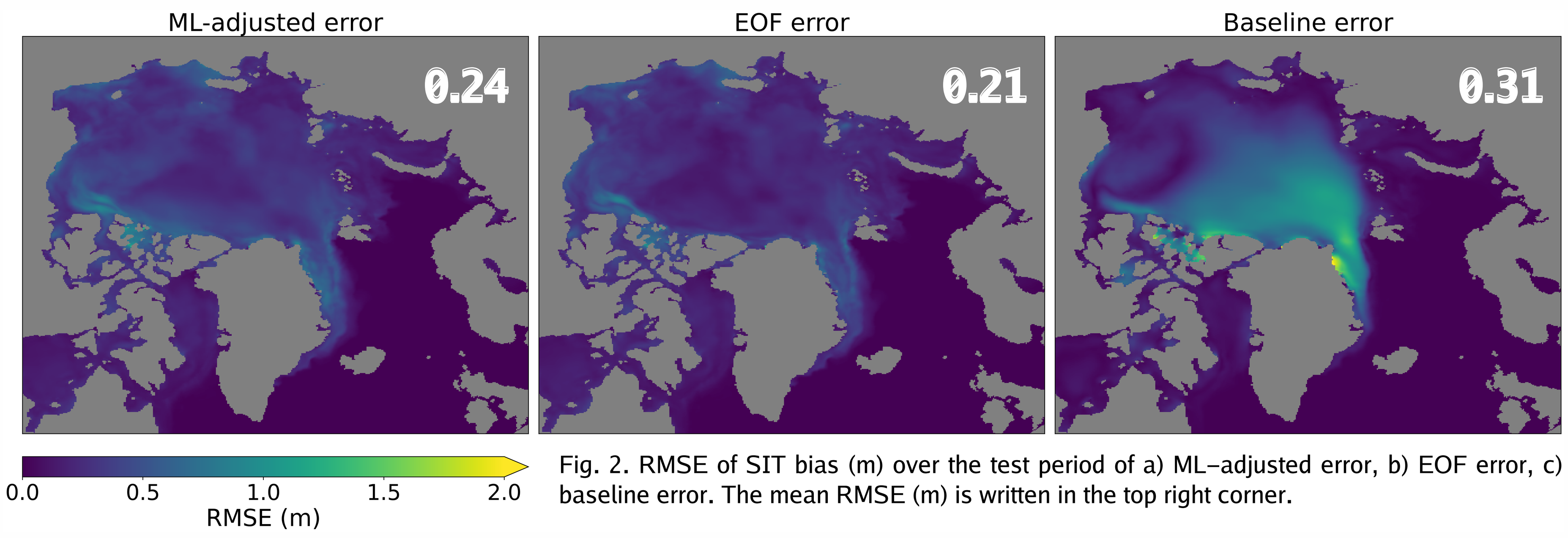


Fig. 2. RMSE of SIT bias (m) over the test period of a) ML-adjusted error, b) EOF error, c) baseline error. The mean RMSE (m) is written in the top right corner.

4. Application over 1991–2010

Assessement of SIT over mooring data

Results show clear improvement compared to TOPAZ freerun, better agreement with ML-adjusted than baseline, especially in Summer and Autumn (Fig. 3 and Table 1).
Good accordance with observations depends strongly on location as well as temporal and spatial scales.

Fig. 3. Daily SIT (m) for observation (mooring A), TOPAZ, TOPAZ freerun and ML-adjusted. The standard deviation of SIT for ULS A is displayed in grey.

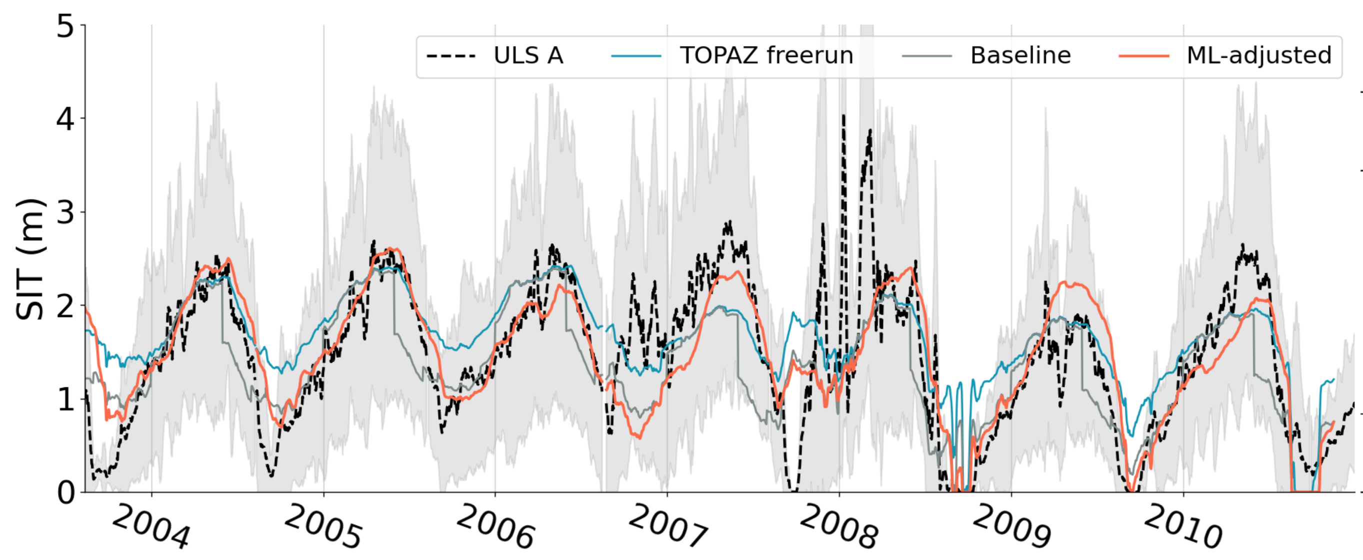


Table. 1. Bias (m), RMSE (m) and correlation between ML-adjusted SIT (baseline) and observed SIT for each BGEP buoy. Average is weighted by the number of days available.

Location	Bias (m)	RMSE (m)	Correlation	Duration (years)
BGEP A	0.021 (-0.105)	0.496 (0.546)	0.763 (0.715)	7
BGEP B	-0.113 (-0.034)	0.405 (0.440)	0.875 (0.849)	5
BGEP C	0.007 (-0.165)	0.513 (0.488)	0.7 (0.8)	5
BGEP D	0.13 (-0.029)	0.778 (0.717)	0.355 (0.364)	3
All BGEP	0.001 (-0.064)	0.368 (0.376)	0.505 (0.506)	-

Change of SIT distribution and mean SIT over time

We observe a decrease in the mean SIT from 2002 to 2012, surrounded by two periods without distinct trends (Fig. 4b). The manifestation of this trend illustrates the capacity of the ML algorithm for extrapolation.
The SIT distribution (Fig. 4a) shows a transition from a bimodal distribution (before 2007) during the growth period, to an unimodal distribution (after 2007). This corresponds to the shift from older sea ice towards younger sea ice in the Arctic.

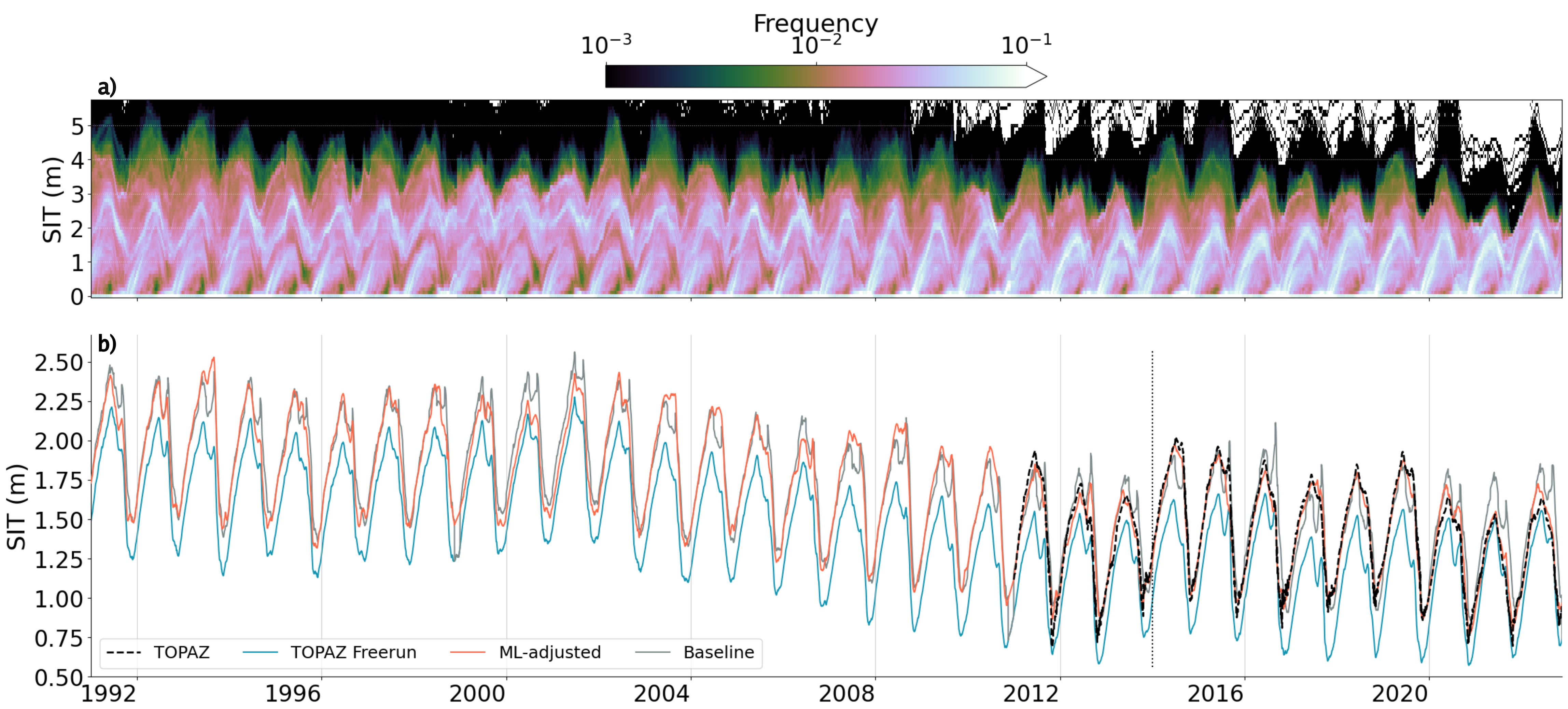


Fig. 4. a) Distribution of daily SIT (m) from 1991 to 2022. Bins of 0.1m are used and the color bar is a log scale. b) Daily SIT (m) averaged over the Arctic for SIC>15% for the same period. The ML algorithm is trained from 2014 to 2022, as indicated by the vertical line in 2014.

Conclusions

- innovative approach combining data assimilation and machine learning
- great potential to reconstruct the Arctic sea ice thickness in the past: time series expansion of 20 years
- uncertainty of the prediction of the SIT can be retrieved: possibility of a second iteration of assimilation
- in-situ datasets (underwater moorings, submarine measurements, or remote sensing) employed for validation