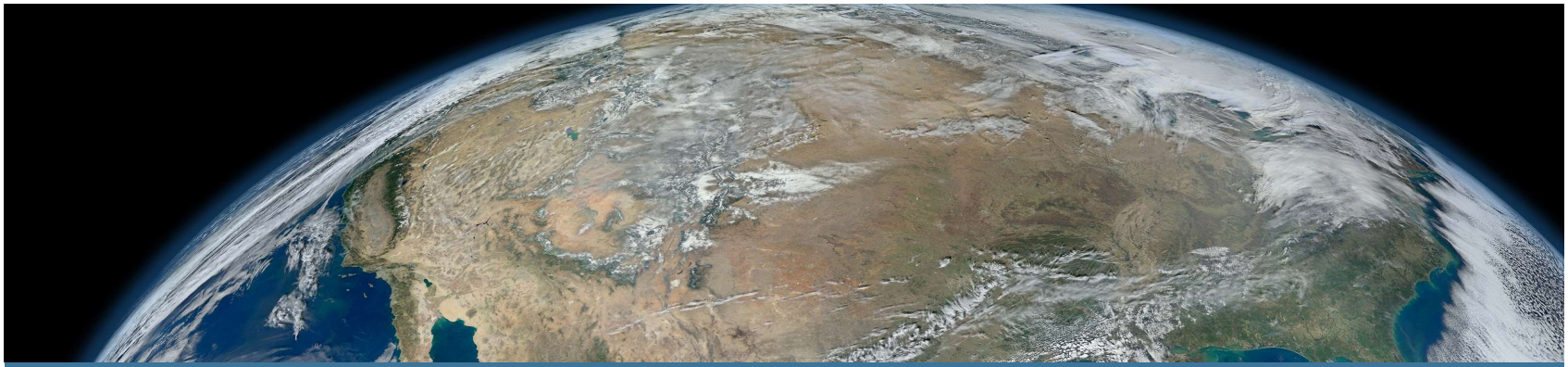


Importance of Uncertainties and Benchmark Metrics in the Diagnostics and Intercomparisons of Sea Surface Temperature Records

Boyin Huang, Xungang Yin, James Carton, Ligang Chen, Garrett Graham, Chunying Liu, Thomas Smith, Huai-Min Zhang*

NOAA National Centers for Environmental Information, Asheville, North Carolina, USA

* Presenter

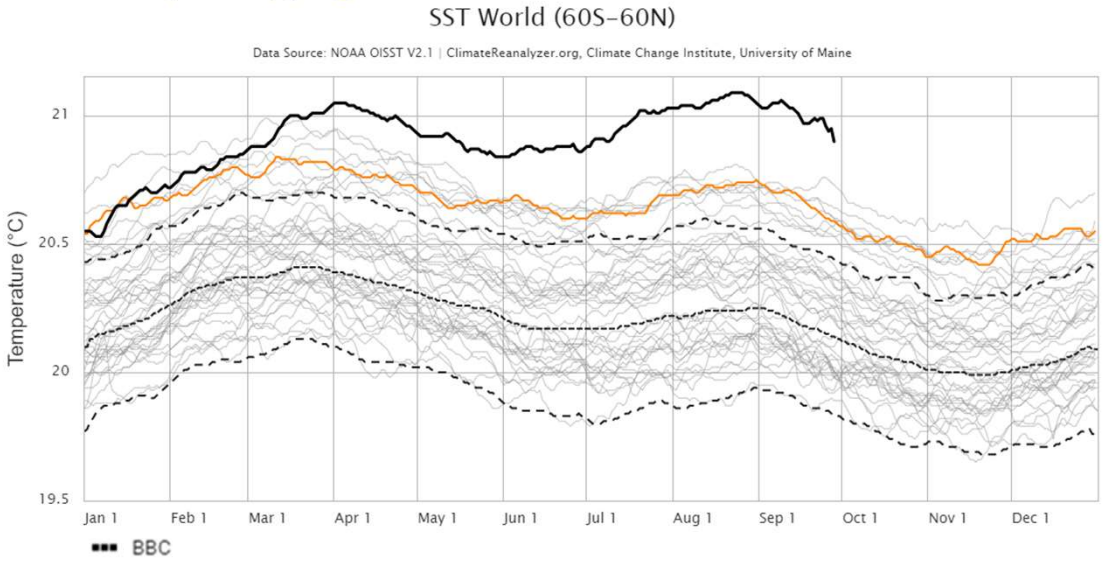
Ref: Huang et al. 2023, JTECH, DOI 10.1175/JTECH-D-22-0081.1.

Outline

1. Objective of this intercomparison study
2. Data sets
3. Metrics of comparisons
4. Comparisons against buoy SST (ICOADS, drifting + moored)
5. Comparisons against Argo SST
6. Comparisons against drifting buoy SST
7. Comparisons against iQuam SST and role of QCs
8. Summary

Hot in the News:
Rapid record ocean surface warming staring March 2023

https://climatereanalyzer.org/clim/sst_daily/
based on NOAA/NCEI OISST



BBC

Recent, rapid ocean warming ahead of El Niño alarms scientists

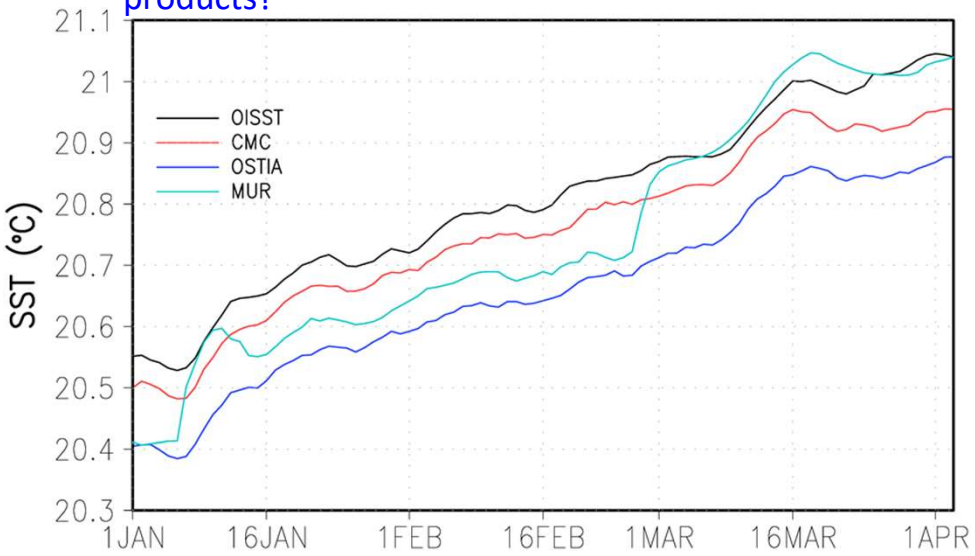
PBS

Scientists issue dire warnings as ocean temperatures spike

National Geographic

Why are our oceans getting warmer?

Media reports focused on record breaking even to the 2nd decimal points: 1) Is that scientifically/statistically significant? 2) How to evaluate among the different products?



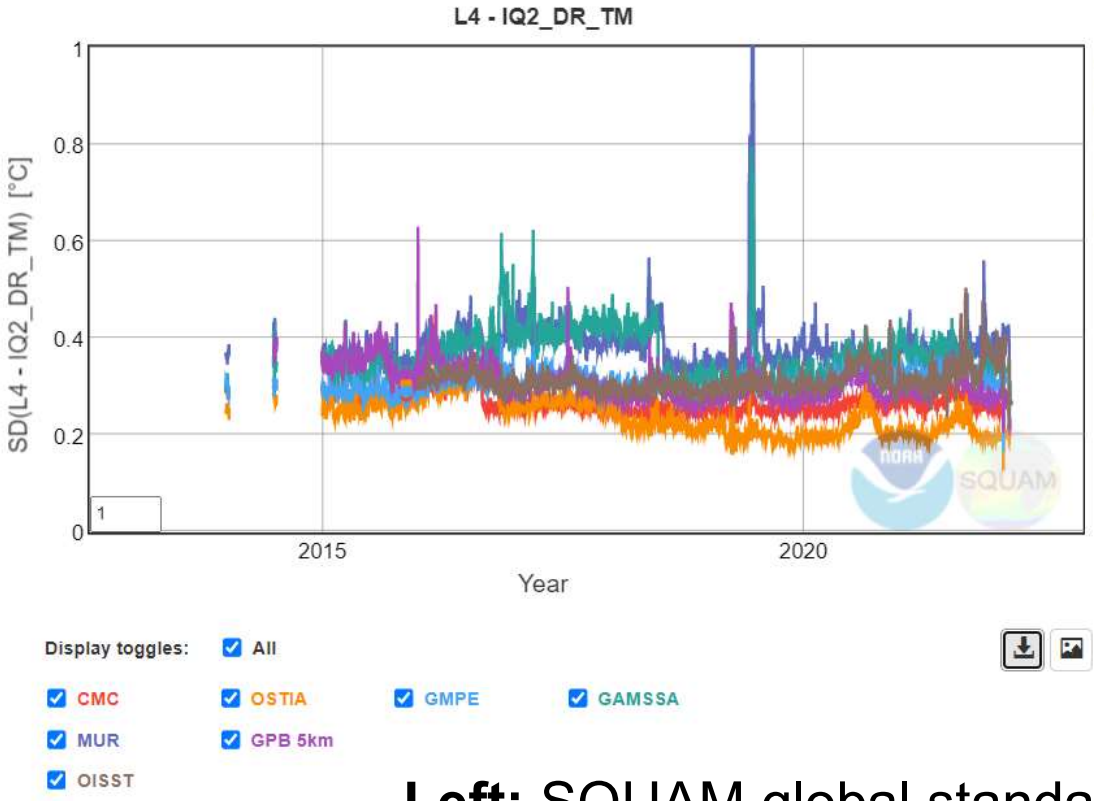
The Weather Channel

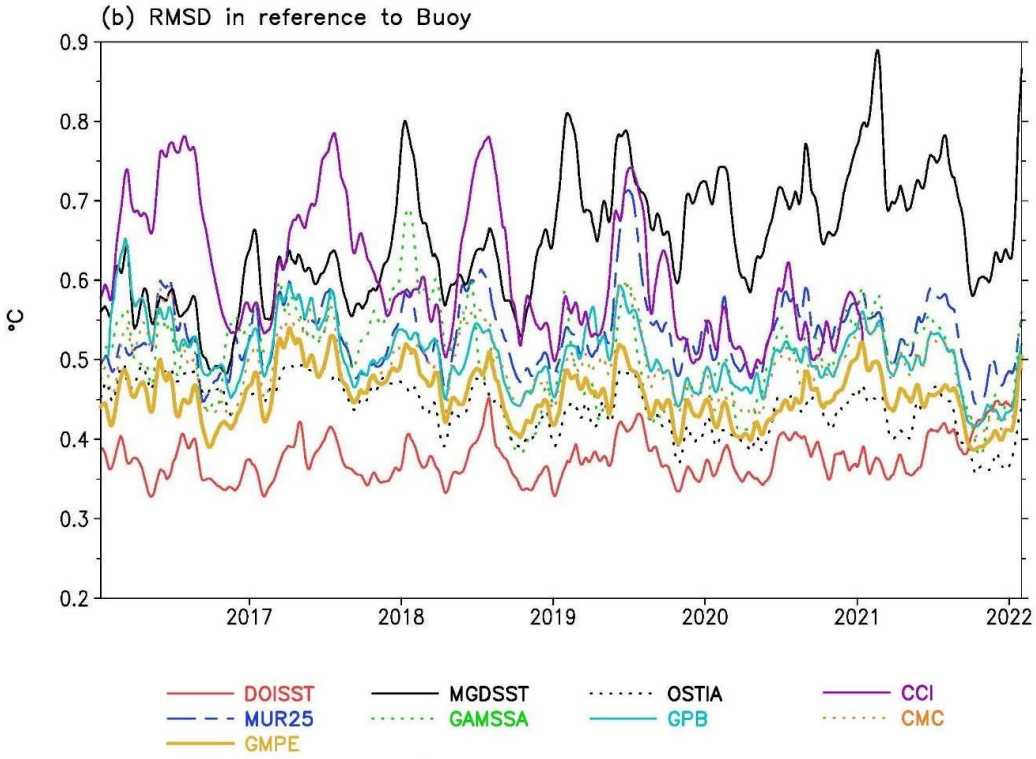
Record Sea Temperatures, El Niño Could Cause Extreme Impacts, Warn Scientists | Weather.com

Boston 25 News

New England coast heating up faster than other bodies of water, creating challenges for marine life

Motivation: To understand the differences in intercomparisons





Left: SQUAM global standard deviation (SD) relative to Buoy SST

Right: Huang et al. (2021a,b) RMSD relative to Buoy SST

The Discrepancy: Potential Orange versus Apple Comparisons

Huang et al. (2021a,b)

- SST products on 0.25-deg grid boxes
- Buoy SST under OISST first-guess (FG) QC on 0.25-deg grids
- Interpolation: None

SQUAM

- SST products on their original resolution (0.05-deg or 0.1-deg)
- Buoy SST from iQuam on in situ locations - Pointwise
- Interpolation to observation points: Dash et al. 2012.

"(a) averaging or interpolating all the L4 SSTs into a common grid (GMPE approach),
(b) interpolating the first term (L41 in ΔTS=L41−L42) to the resolution of the second term (L42), using various linear or cubic formulations or inverse distance-weighted methods, or,
(c) selecting the nearest neighbor (NN)."

Metrics for comparisons

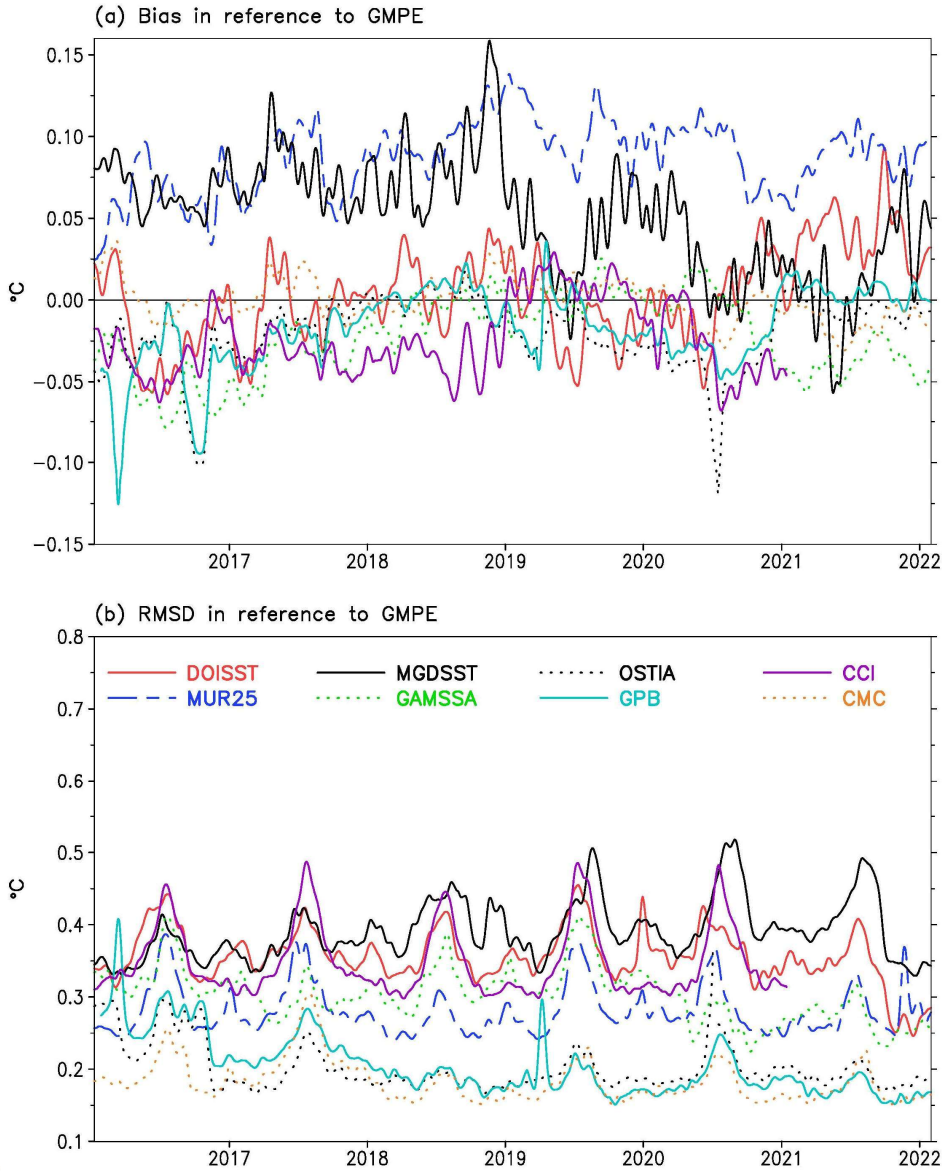
Bias: Difference b/w SST products and reference SSTs,
weighted by cosine(latitude) for area-averaged SSTs (0.25-deg boxes)

RMSD: Root-Mean Square Difference
weighted by cosine(latitude) for area-averaged SSTs (0.25-deg boxes)

DIFF: Pointwise Bias without cosine latitude weighting

SD: Standard Deviation

A common approach is compared with GMPE; BUT, GMPE itself has biases!



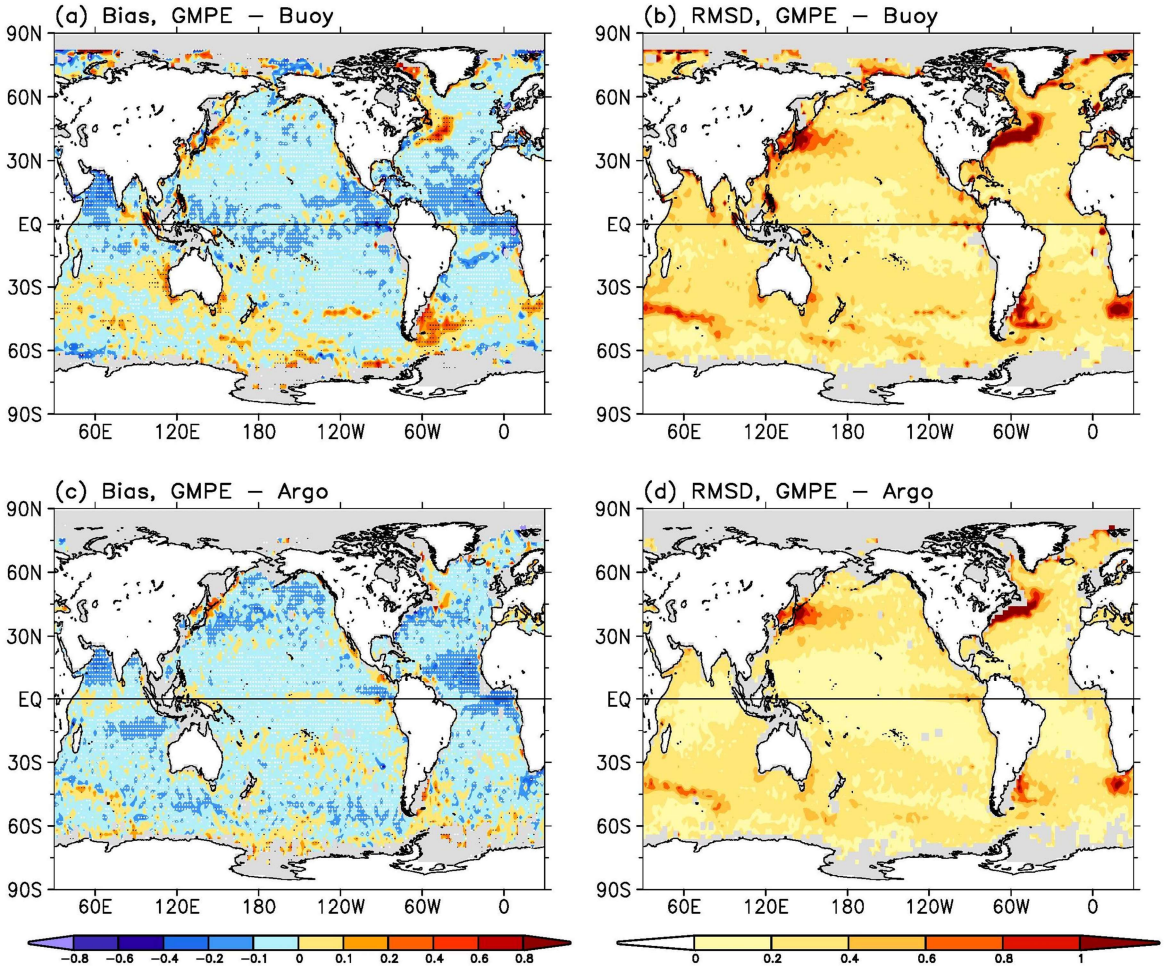
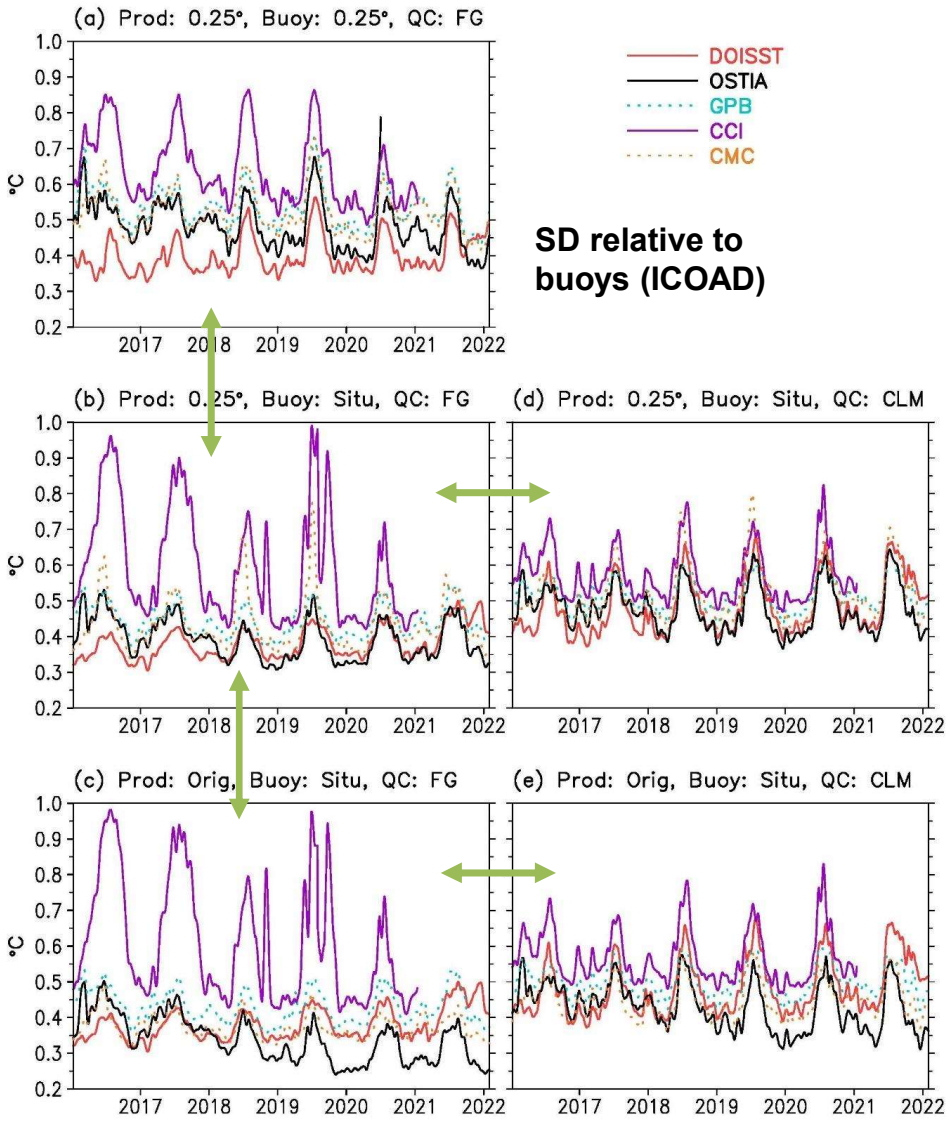


TABLE 2. Averaged biases and RMSDs (°C) in reference to GMPE, Buoy, and Argo SSTs on 0.25° × 0.25° grids from 1 Jan 2016 to 31 Jan 2022 in Figs. 1 and 3, and S1. The ± values represent the uncertainty at 95% confidence level that is determined by the lagged autocorrelation, effective sampling number, and the standard deviation (SD) (Huang et al. 2021b).

	GMPE reference		Buoy reference		Argo reference	
SST product	Bias	RMSD	Bias	RMSD	Bias	RMSD
DOISST v2.1	0.002 ± 0.017	0.357 ± 0.014	−0.018 ± 0.013	0.376 ± 0.009	−0.033 ± 0.007	0.346 ± 0.002
MUR25	0.087 ± 0.010	0.281 ± 0.011	0.038 ± 0.010	0.531 ± 0.015	0.036 ± 0.005	0.377 ± 0.004
MGDSST	0.051 ± 0.029	0.391 ± 0.016	0.028 ± 0.009	0.650 ± 0.047	0.006 ± 0.019	0.523 ± 0.011
GAMSSA	−0.024 ± 0.019	0.303 ± 0.022	−0.071 ± 0.011	0.505 ± 0.024	−0.088 ± 0.010	0.480 ± 0.005
OSTIA	−0.020 ± 0.011	0.200 ± 0.016	−0.045 ± 0.011	0.431 ± 0.012	−0.069 ± 0.011	0.370 ± 0.042
GPB	−0.016 ± 0.014	0.203 ± 0.041	−0.051 ± 0.015	0.505 ± 0.016	−0.066 ± 0.011	0.381 ± 0.020
CCI	−0.025 ± 0.013	0.349 ± 0.018	−0.054 ± 0.011	0.608 ± 0.036	−0.068 ± 0.008	0.429 ± 0.017
CMC	−0.001 ± 0.006	0.182 ± 0.012	−0.052 ± 0.009	0.492 ± 0.015	−0.056 ± 0.007	0.380 ± 0.002
GMPE	—	—	−0.036 ± 0.012	0.454 ± 0.012	−0.055 ± 0.007	0.363 ± 0.011

Top 3 scores



SD relative to buoys (ICOAD)

Effects of Aera-Average vs Pointwise and QCs

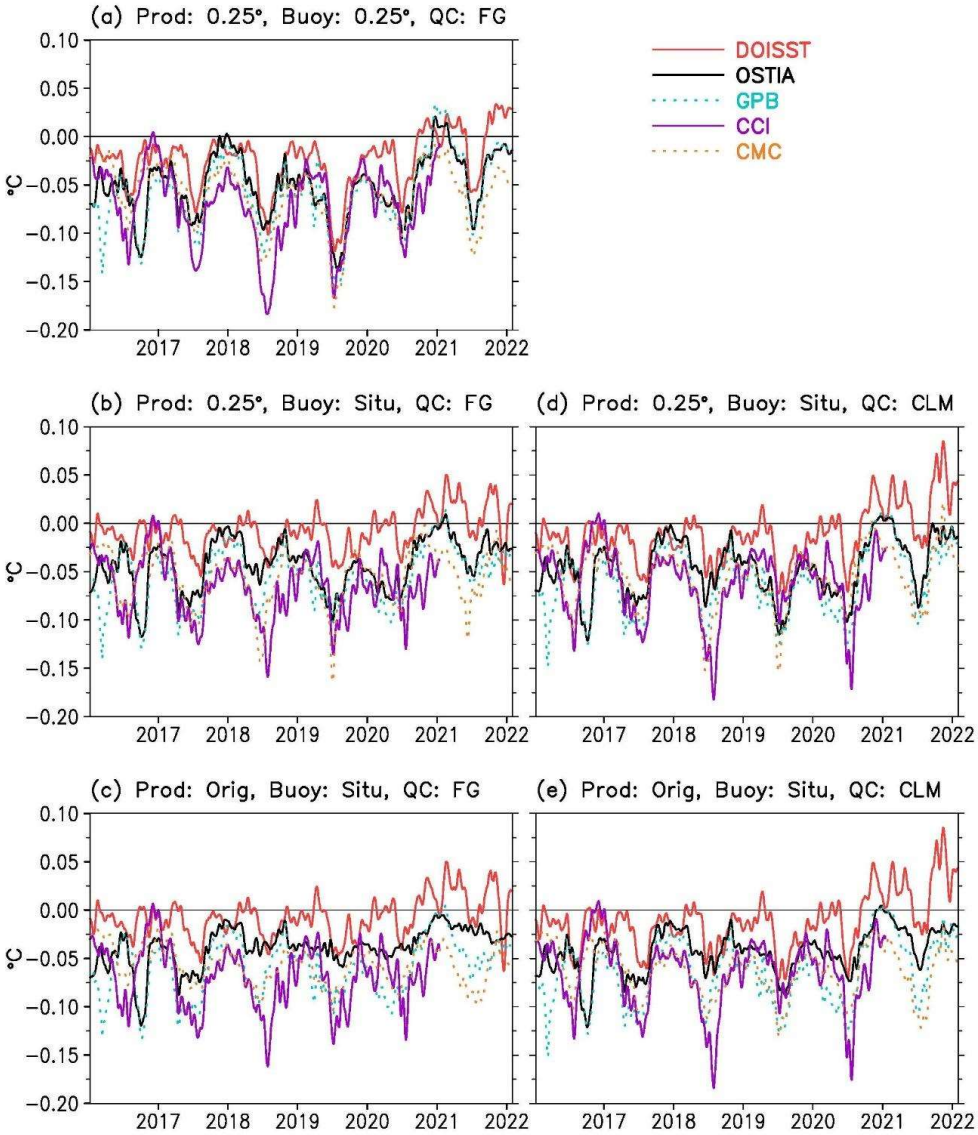
Panels	Product resolution	In-Situ/Buoy data resolution	Situ QC process
(a)	0.25-deg	0.25-deg	First-Guess
(b)	0.25-deg	In situ/pointwise	First-Guess
(c)	Original	In situ/pointwise	First-Guess
(d)	0.25-deg	In situ/pointwise	Climatology
(e)	Original	In situ/pointwise	Climatology

Features:

(1) SDs in DOISST is small at 0.25-deg in (a).

(2) SDs of OSTIA and CMC decrease in (b) and (c)
due to high resolution and
small scale (~25km) of matchups in bias-correction.

(3) SDs are generally larger in (d) and (e) than (a)-(c)
except for CCI due independence from in situ obs.



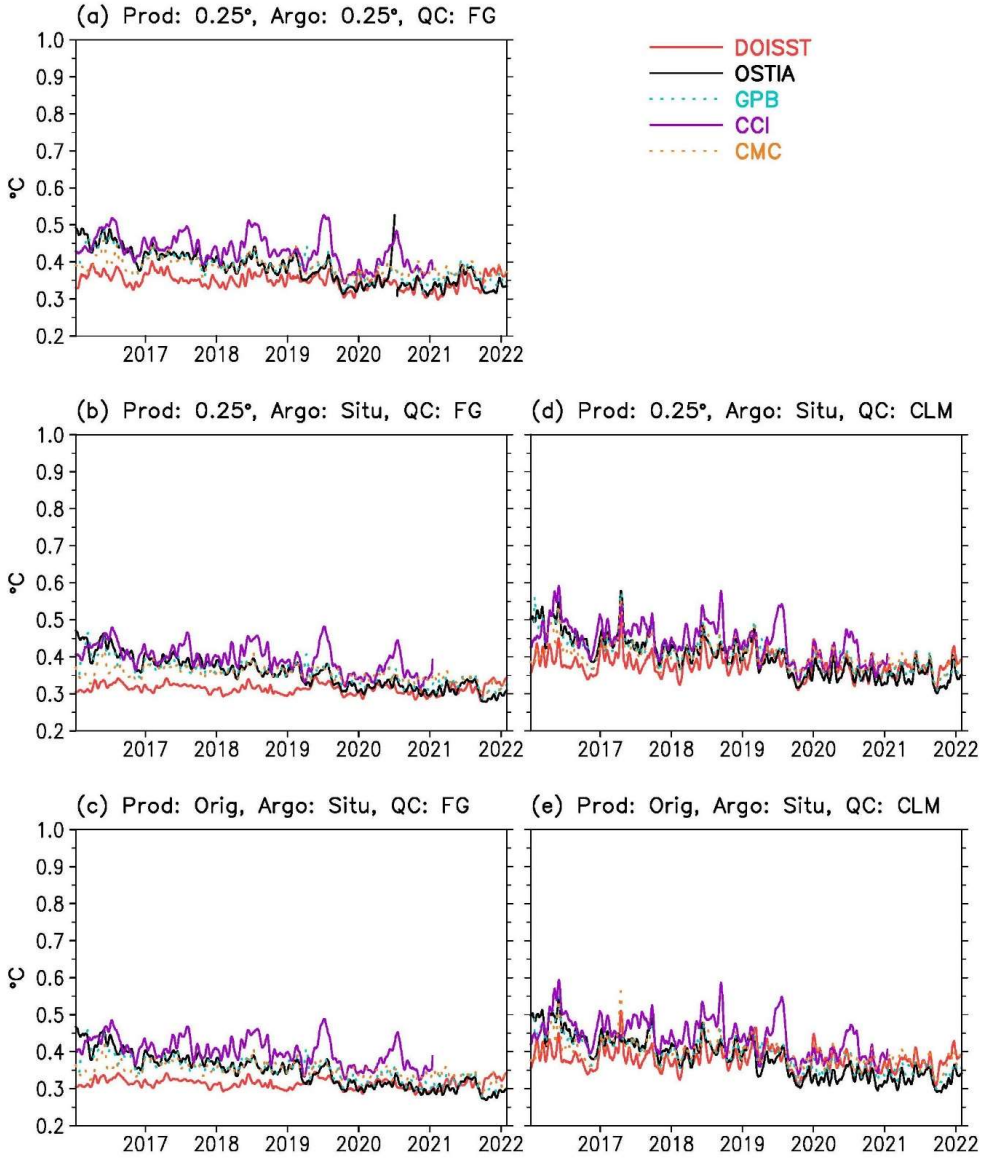
DIFFs relative to Buoy SST (ICOADS)

Panels	Products	Buoy	QC
(a)	0.25-deg	0.25-deg	First-Guess
(b)	0.25-deg	In situ	First-Guess
(c)	Original	In situ	First-Guess
(d)	0.25-deg	In situ	Climatology
(e)	Original	In situ	Climatology

Features:

(1) Relative performance of DIFFs
do not change much in (a)-(e)
due to error-cancellation.

(2) Low DIFF in DOISST may result from
its largescale (3000km x 500km) bias-correction
to satellite SSTs.



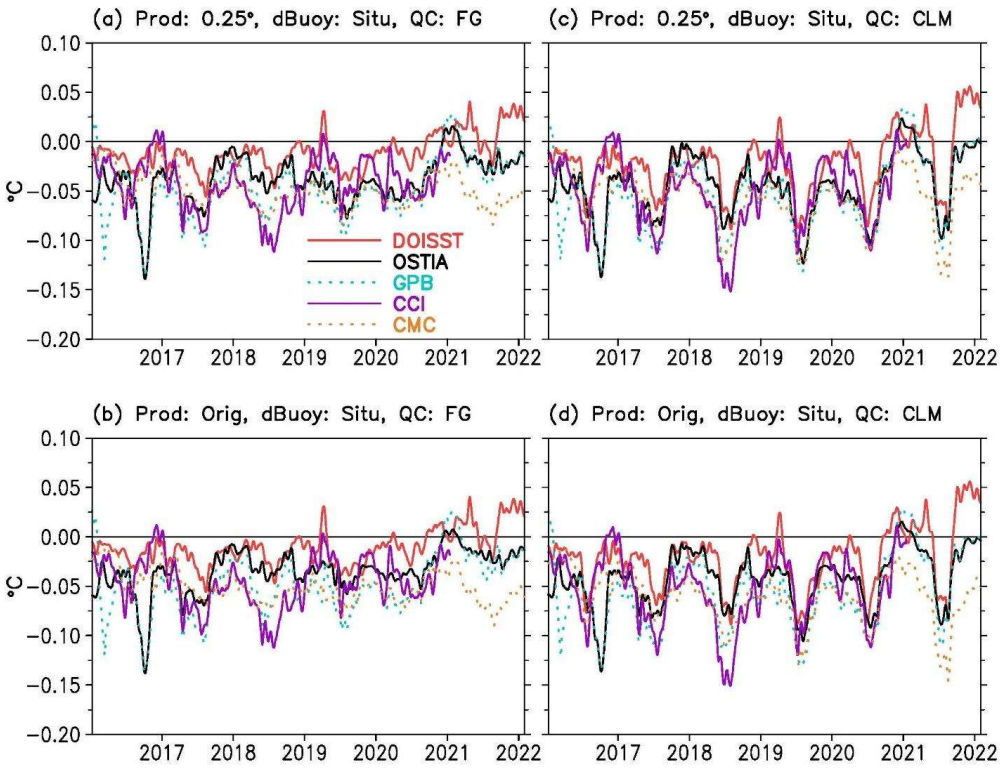
SD relative to Argo SST

Panels	Products	Buoy	QC
(a)	0.25-deg	0.25-deg	First-Guess
(b)	0.25-deg	In situ	First-Guess
(c)	Original	In situ	First-Guess
(d)	0.25-deg	In situ	Climatology
(e)	Original	In situ	Climatology

Features:

(1) Magnitude of SDs is smaller than that against Buoy SST due to longer observing interval of Argo (10 day) in contrast buoy (6 m to 1 h).

(2) Relative performance is similar to that against Buoy SST.



DIFF relative to Drifting-Buoy SST

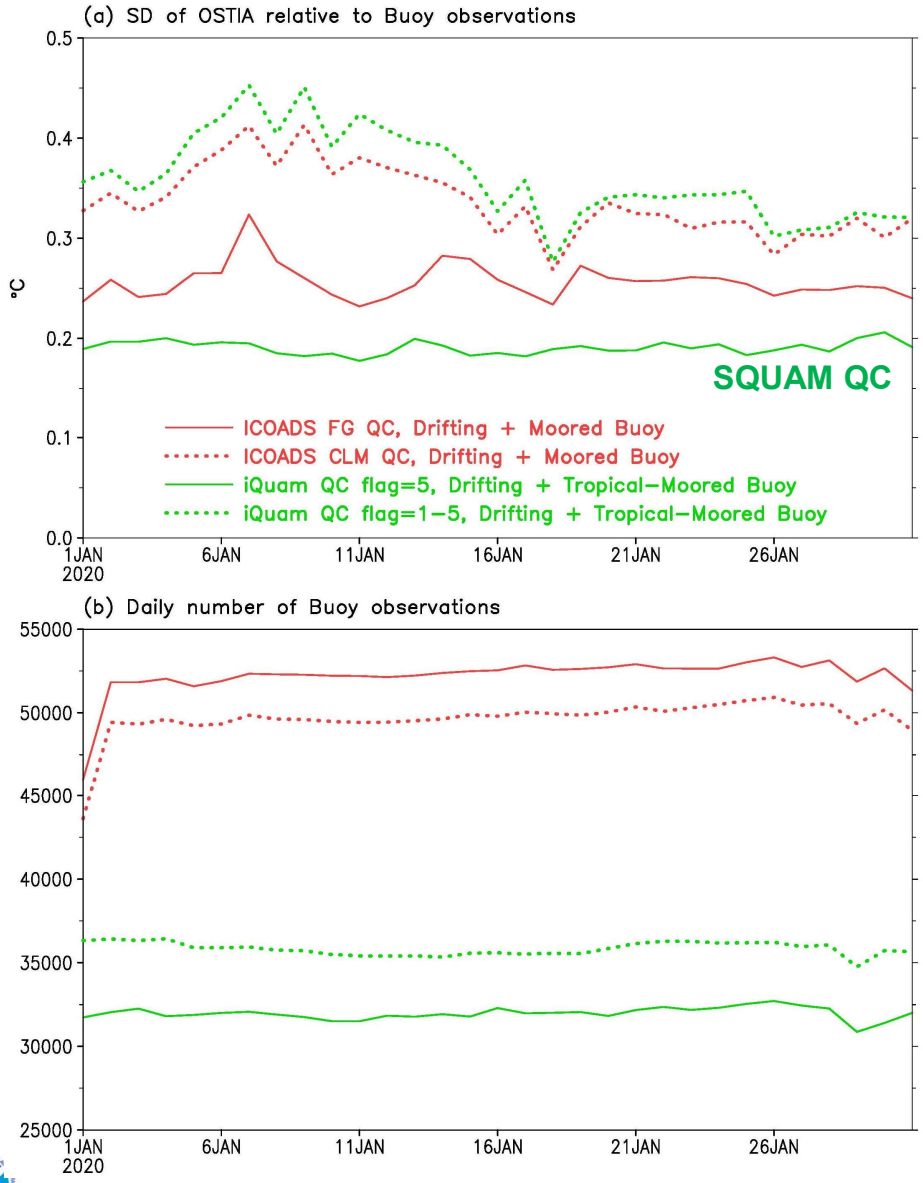
Panels	Products	Buoy	QC
(a)	0.25-deg	0.25-deg	First-Guess
(b)	Original	In situ	First-Guess
(c)	0.25-deg	In situ	Climatology
(d)	Original	In situ	Climatology

Features:

(1) Relative performance of DIFFs do not change much in (a)-(d).

(2) Relative performance is similar to that against Buoy SST.

(3) The good performance of DOISST is seen.



Another reason

for the differences in intercomparisons:

Using same reference Buoy data but different QC.

For example,

ICOADS buoy SST: FG or CLM QC

iQuam buoy SST: QC flag-5

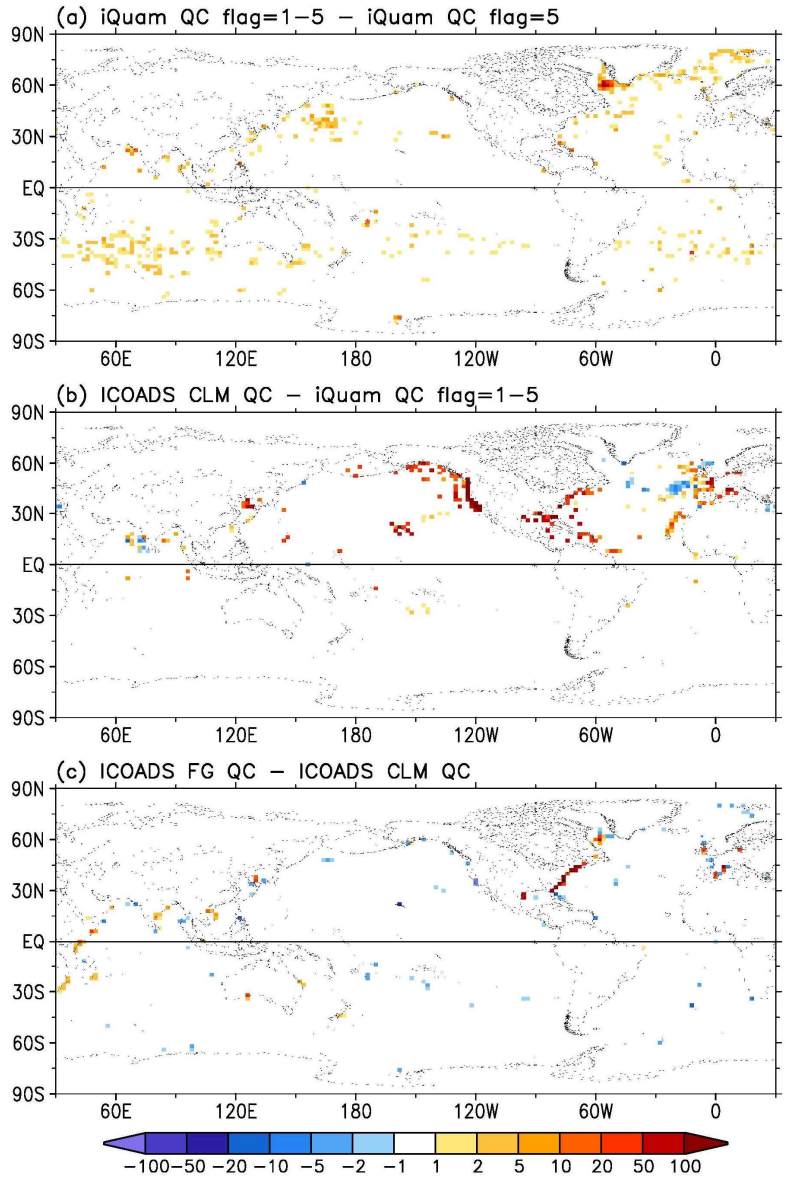
(a) OSTIA SD vs different SST references

(b) Daily number of buoy observations of different SST references

Features:

(1) SQUAM observation number is much lower than ICOADS.

(2) SDs vs SQUAM (iQuam QC flag-5) are lower than those vs ICOADS.



Observation difference and its spatial distribution

(a) iQuam data with QC Flag 1-4

(b) ICOADS CLM QC – iQuam QC 1-5

(c) ICOADS FG QC – ICOADS CLM QC

Summary

(1) “Discrepancies” of GHRSST L4 performance on different sites/studies are assessed: Results are different for aera-averaged metrics vs pointwise metrics, as well as the exact data QC used

(2) DOISST has a good performance in long term time series against in-situ data (lower SD and DIFF). OISS’s performance suffers in localized areas, due to the bias-correction at large scale – needs to be improved.

(3) SDs decrease to the lowest in OSTIA and CMC, when they are interpolated from their native grids and compared to in-situ point observations, which may result from their high resolutions and bias correction by matchups in a small scale.

(4) SDs relative to Argo and drifting buoys are smaller, which may result from the longer observing interval; Including high-resolution (6m – 1 h) moored buoy data increases the SDs.

(5) SDs in SQUAM (relative to iQuam QC flag-5) are smaller than those relative to ICOADS (FG or CLM QC) due to smaller numbers of observations near the ocean coasts.

(6) GMPE has its own biases and are not necessarily better than the input L4 products

- Backup slides

SST Data Sets (2016-2022)

DOISST v2.1, 0.25-deg
OSTIA, 0.25-deg
GPB, 0.25-deg
CCI, 0.25-deg
CMC, 0.25-deg

OSTIA, 0.05-deg
GPB, 0.05-deg
CCI, 0.05-deg
CMC, 0.10-deg

Reference SST Data Sets

GMPE 0.25-deg

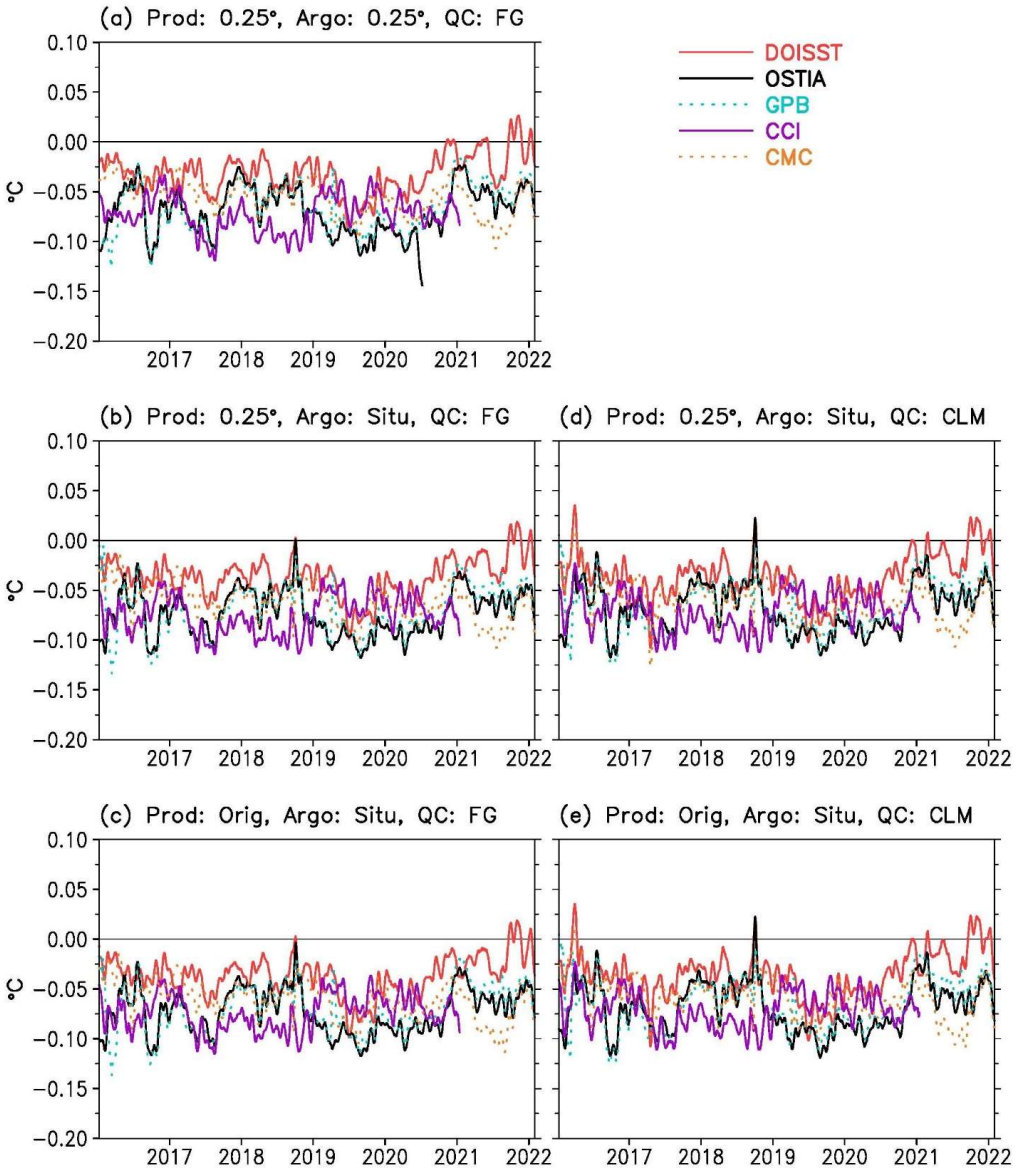
First-Guess QC-ed Buoy 0.25-deg
First-Guess QC-ed Argo 0.25-deg

First-Guess QC-ed Buoy in situ
First-Guess QC-ed Argo in situ

Climatology QC-ed Buoy 0.25-deg
Climatology QC-ed Argo 0.25-deg

Climatology QC-ed Buoy in situ
Climatology QC-ed Argo in situ

iQuam buoy SST
Drifting-buoy SST



DIFF relative to Argo SST

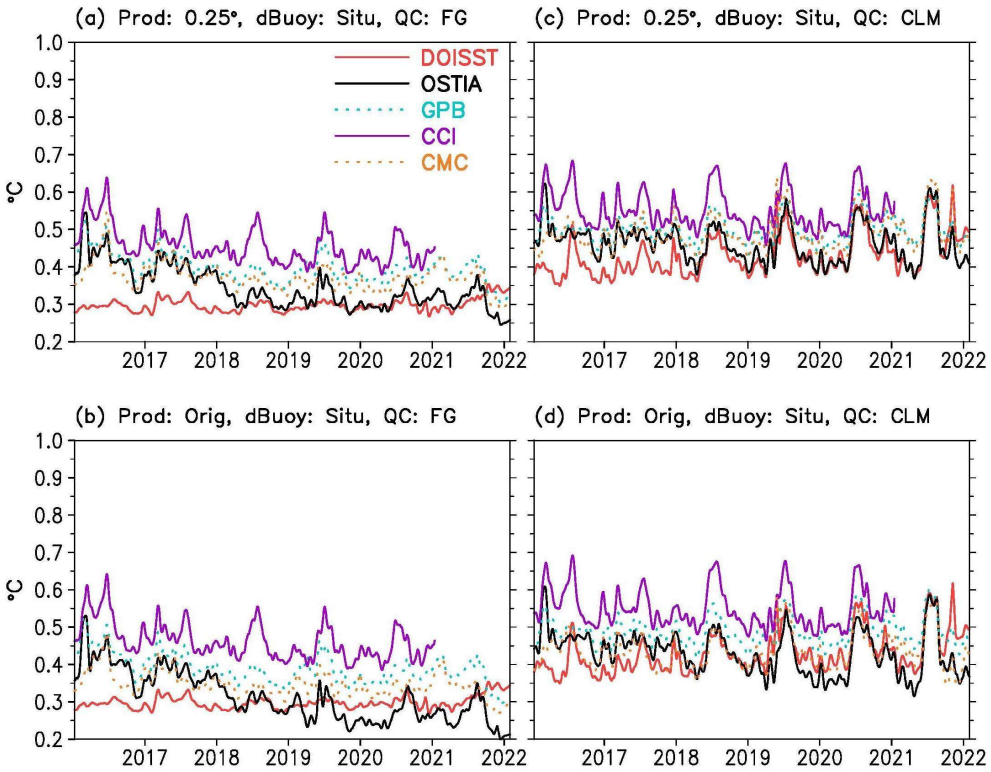
Panels	Products	Buoy	QC
(a)	0.25-deg	0.25-deg	First-Guess
(b)	0.25-deg	In situ	First-Guess
(c)	Original	In situ	First-Guess
(d)	0.25-deg	In situ	Climatology
(e)	Original	In situ	Climatology

Features:

(1) Relative performance of DIFFs do not change much in (a)-(e).

(2) Relative performance is similar to that against Buoy SST.

Why are the SDs relative to Buoy larger than those relative to Argo SSTs?



SD relative to Drifting-Buoy SST

Panels	Products	Buoy	QC
(a)	0.25-deg	In situ	First-Guess
(b)	Original	In situ	First-Guess
(c)	0.25-deg	In situ	Climatology
(d)	Original	In situ	Climatology

Features:

(1) Magnitude of SDs is smaller than that against buoy due to longer observing interval of drifting-buoy (1h) vs moored-buoy (6m—1h), which may be the reason for the small SDs against Argo.

(2) Relative performance are similar to that against Buoy.

(3) The good performance of DOISST is seen.