Development of an Ensemble-Based Ocean Data Assimilation System at ECCC

Sergey Skachko, Mark Buehner, Gregory Smith, Andrew Peterson, Dorina Surcel-Colan

Environment and Climate Change Canada

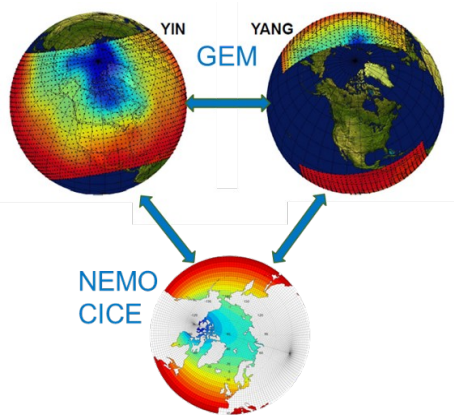


Environment and Climate Change Canada | Environnement et Changement climatique Canada

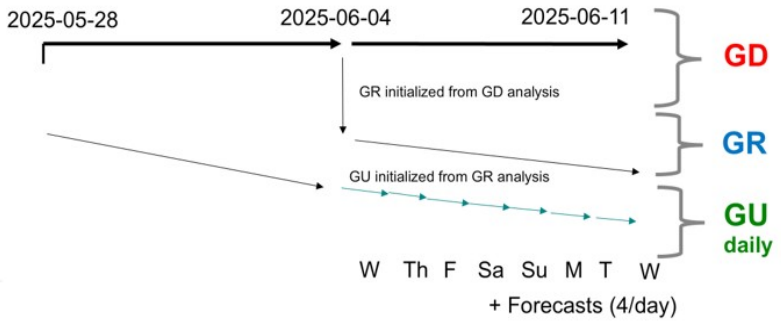
Canada

Sergey.Skachko@ec.gc.ca

Motivation and Objectives



- ECCC is developing a fully coupled **atmosphere-ocean-sea ice ensemble data assimilation system** for operational Earth system prediction.
- The current operational **ocean** analysis system is deterministic and consists of:
  - **Weekly analysis** component assimilating SST, T/S profiles, and altimetry.
  - **Daily (3D) analysis** component assimilating SST observations.
- Main objective:
  - Replace deterministic ocean analyses with an **ensemble-based framework** consistent with ECCC's atmospheric data assimilation systems.
- Expected benefits:
  - Flow-dependent background-error statistics.
  - Improved representation of uncertainty.
  - Stronger consistency across Earth system components.



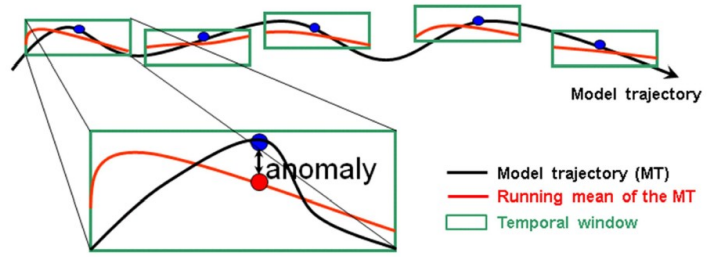


Sub-monthly SST anomaly

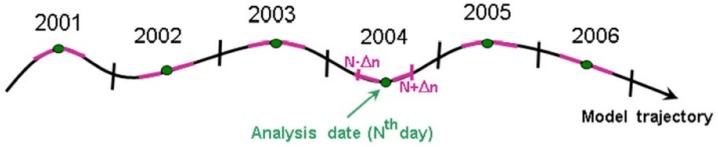
SAM2 ocean data assimilation system

$$K = S_n \left[I + (HS_n)^T R^{-1} (HS_n) \right]^{-1} (HS_n)^T R^{-1}$$

- Based on a SEEK approach
  - SAM2v2 version includes LETKF option
- Background error specified using Multivariate anomalies
  - Adaptivity scheme (α) to adjust background error variance
- Fortran/python loaders for all common ocean observations types/datasets
- Scale-dependent observation operator (Shapiro filter)
  - Including for large-scale low-frequency barotropic height (HBRST)
- Spatial representivity error (SLA, MDT, vertical profiles)
- Online quality control for vertical profile data
- Bogus (zero innovation) observations under ice, in tropics and at river mouths
- 2D Localization
- Heterogeneous reduced-grid approach
  - Error modes, assimilation grid, increment
  - GIOPS (2x), RIOPS (4x)
- Compression to eliminate land points (cpmx format)
- Configurable Incremental Analysis Updating (IAU) Scheme
- 3DVAR T/S profile bias correction scheme



(b)



MERCATOR OCEAN INTERNATIONAL

$$d^T [\alpha (HS)(HS)^T + R]\, d = n$$

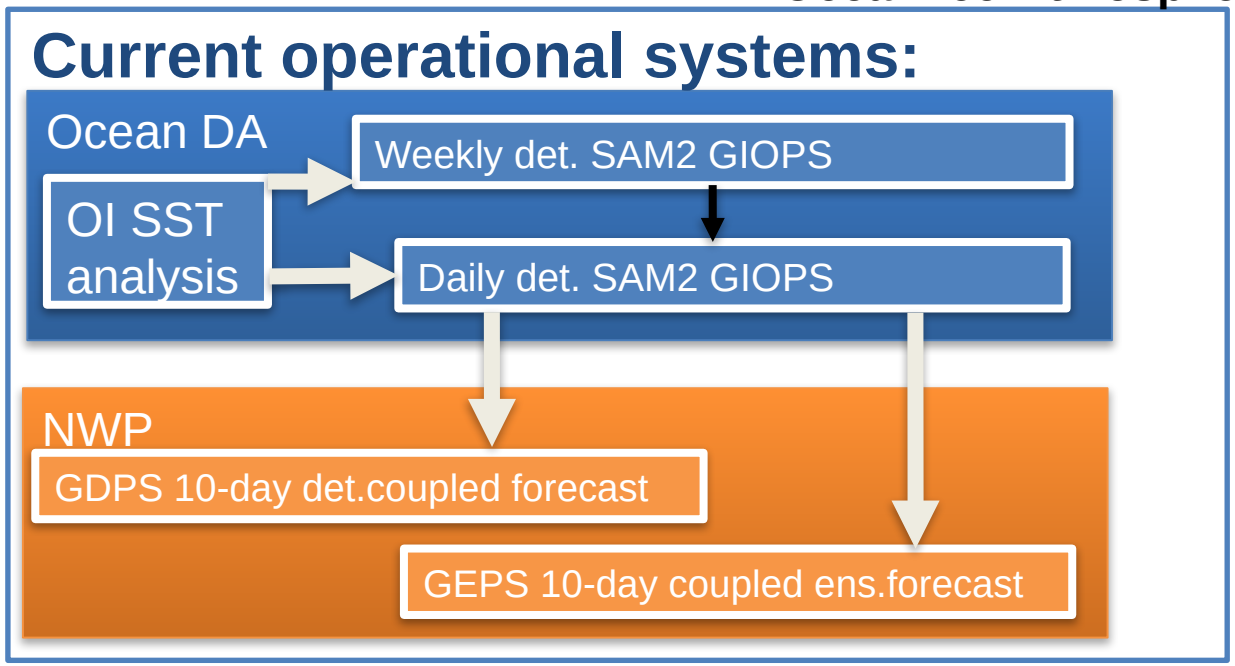
$$n = number\ of\ independent\ data$$

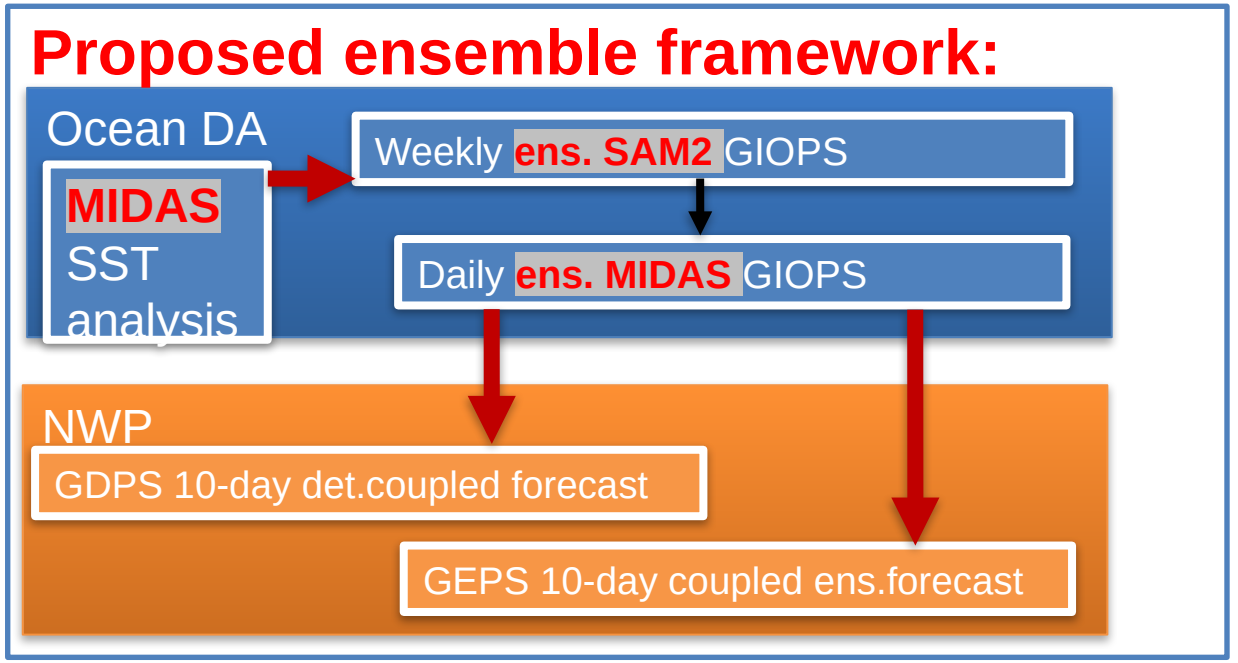
Lellouche et al. (2013; 2018); Smith et al. (2018); Benkiran et al. (2021)

MIDAS: Modular and Integrated Data Assimilation System

- MIDAS is a unified assimilation software framework for research and operational use at ECCC with the objective to avoid redundant work within the organization;
- Facilitate transition towards CDA of different Earth system components;
- Modular and generalized implementation of DA elements (i.e. **H**, **R**, **B**, **x**, **y**) enables use of DA algorithms for various applications (i.e. atmosphere, sea ice, ocean, land surface) using mostly the same code;
- MIDAS currently includes:
  - Assimilation algorithms: 2/3D-Var, 4D-EnVar, LETKF;
  - Bias correction, QC and thinning algorithms for all operational NWP obs types;
  - Observation impact estimation with hybrid-FSOI and ensemble-FSOI approaches;
  - Operational:
    - atmospheric DA systems, including atmospheric LETKF;
    - Sea ice and SST analysis using 2D-Var
  - Upcoming ocean LETKF DA;
  - Upcoming sea-ice LETKF DA (and Strongly Coupled ocean-sea-ice DA);
- MIDAS paper: Buehner et al. 2025, GMD

Towards a Consistent Ensemble Ocean-Ice-Atmosphere Prediction Framework

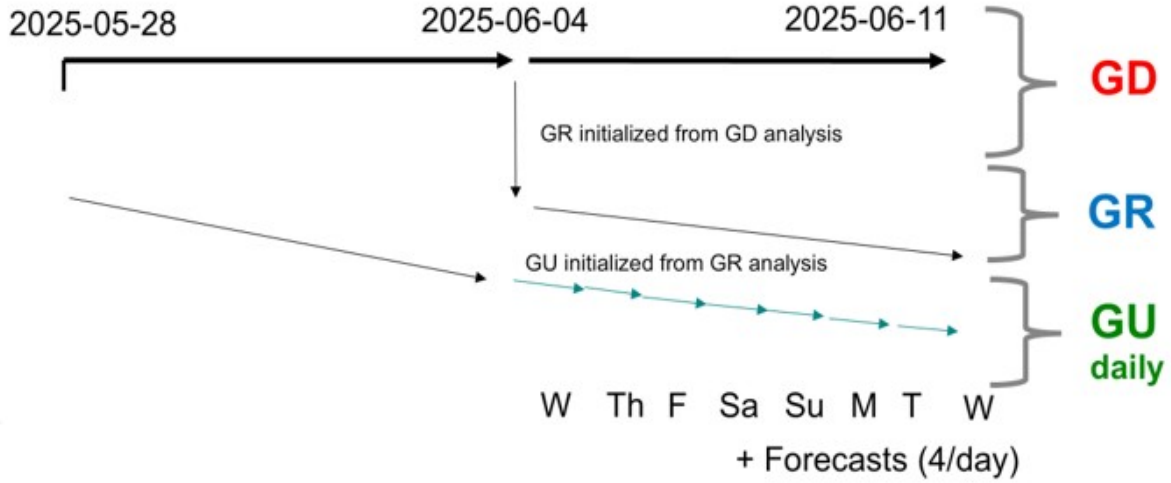




- **Consistent Ocean Initial Conditions**
  GDPS and GEPS inherit ocean states generated from the same ensemble framework.
- **Flow-Dependent Error Covariances**
  Deterministic analyses benefit from ensemble information.
- **Direct SST Assimilation**
  Removes dependence on a separate SST analysis system.
- **Improved Coupled System Consistency**
  Greater coherence between ocean, ice, and atmospheric prediction components.

Ensemble Data Assimilation Methodology

- Maintain the existing **SAM2 weekly ocean analysis** component to provide robust constraints on mesoscale circulation and water mass properties.
- Use the **MIDAS framework** for **daily** high-frequency ensemble-based ocean analyses and improved flow-dependent error representation.
- Leverage MIDAS-LETKF to better capture atmosphere-ocean-sea-ice interactions within a coupled data assimilation (CDA) environment.
- Conduct CDA experiments in which atmospheric, oceanic, and sea-ice observations are assimilated simultaneously within a unified ensemble framework.
- Assess the benefits of coupling for analysis consistency and forecast skill across Earth system components.



GD and GR are weekly delayed and real-time SAM2 cycles

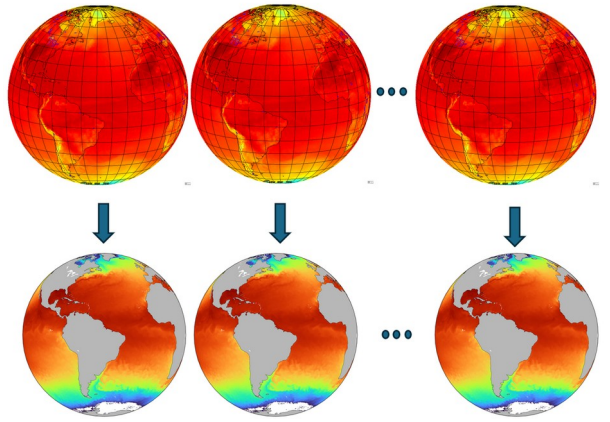
Daily GU cycle replaced with MIDAS LETKF

Transition from Deterministic to Ensemble-Based Ocean Analysis

New Ensemble-Based System: Weekly and Daily

- Ensemble of **NEMO-CICE** ocean-sea-ice analyses.
  - Peterson et al. (GRL; 2026)
- Background and IAU runs perturbed by:
  - **Global Ensemble Prediction System (GEPS)** atmospheric forcing
  - **STOPACK** (Storto & Andriopoulos 2021 QJRMS)
    - Stochastic parameter perturbations (SPP)
    - Stochastically Perturbed Parametrization Tendencies (SPPT)

GEPS: atmospheric ensemble DA

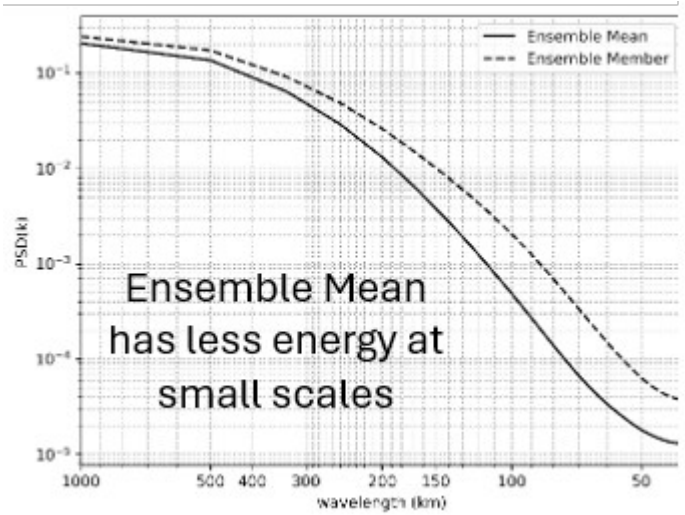


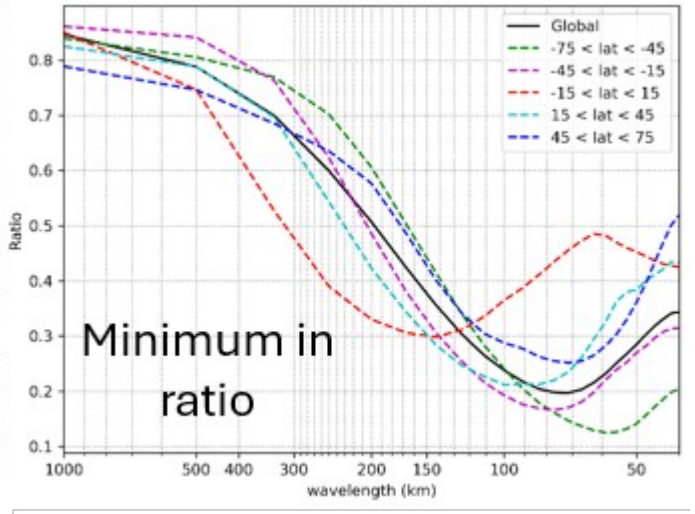
GIOPS: ocean ensemble DA

Weekly Ensemble Data Assimilation component

Extend deterministic weekly SAM2 to Ensemble Weekly Ocean Data Assimilation System

- Produce 20+1 ensemble members using GEPS ensemble forcing and STOPACK (SPP, SPPT perturbations)

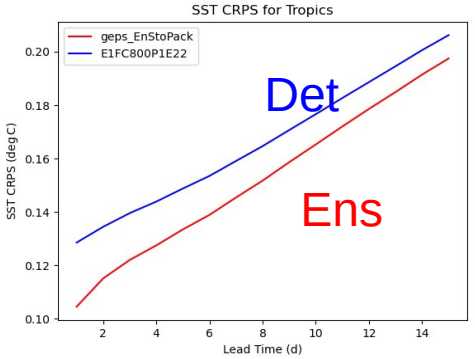




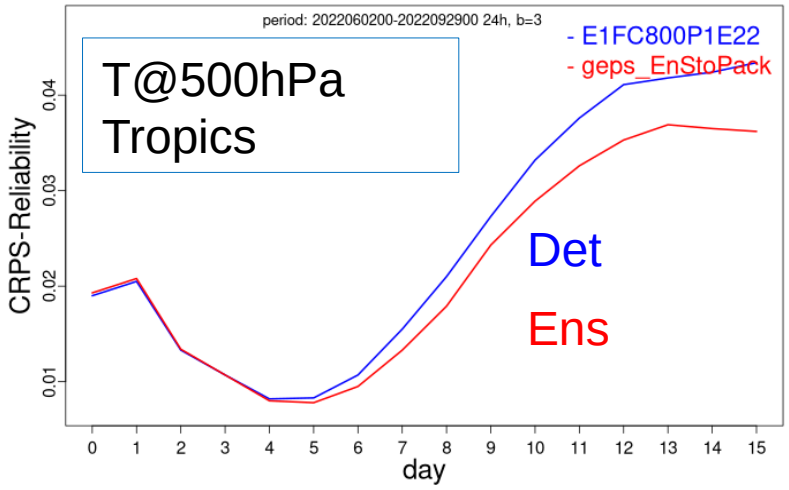
Advantages

- Probabilistic ocean products
- Ensemble mean can be used to remove unconstrained variability (Peterson et al., 2026)
- Provides ensemble initial conditions for Coupled GEPS forecasts
- Ensemble initial conditions for daily LETKF

CRPS SST Tropics

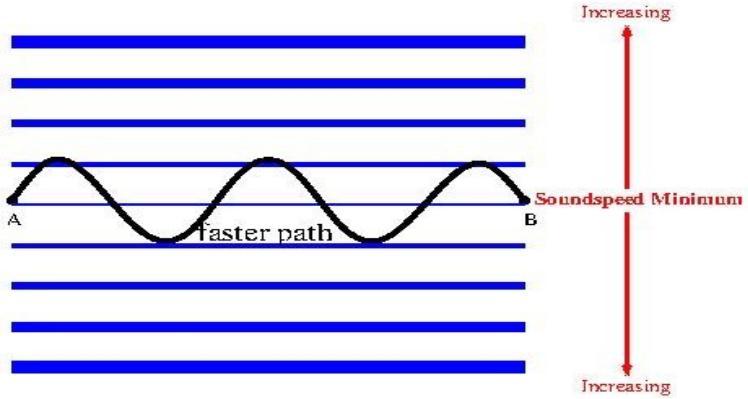


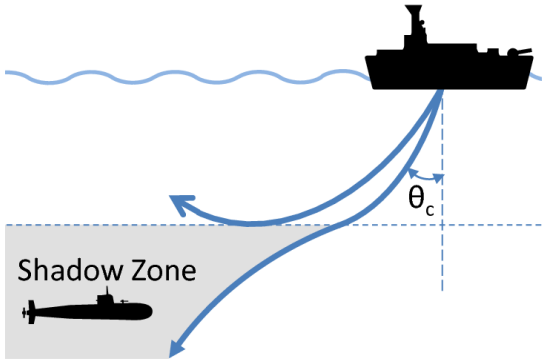
Impact of ensemble initial conditions on coupled GEPS forecasts



Ensemble of Sound Speed Profiles Near Surface Sound Ducts

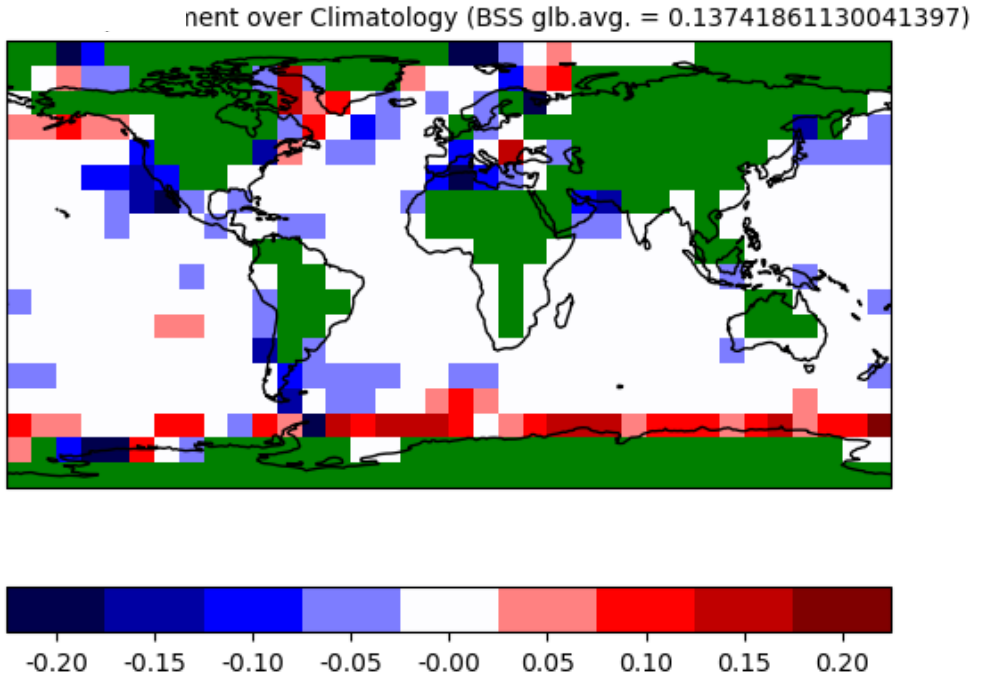
Near Surface Sound Channel



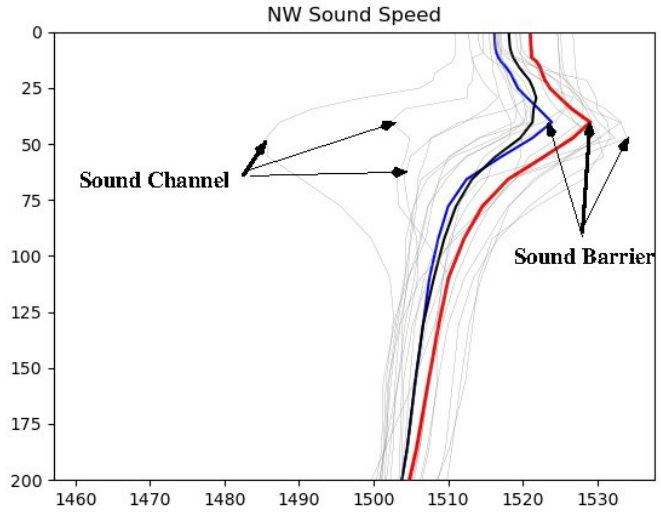


Shadow Zone underneath sound channel

Brier Score Improvement (Lowering/Blue) over Climatology



Ensemble of Sound Speed Profiles from North Wall of Gulf Stream



Daily Ensemble Data Assimilation component

MIDAS LETKF Daily Ocean Data Assimilation System

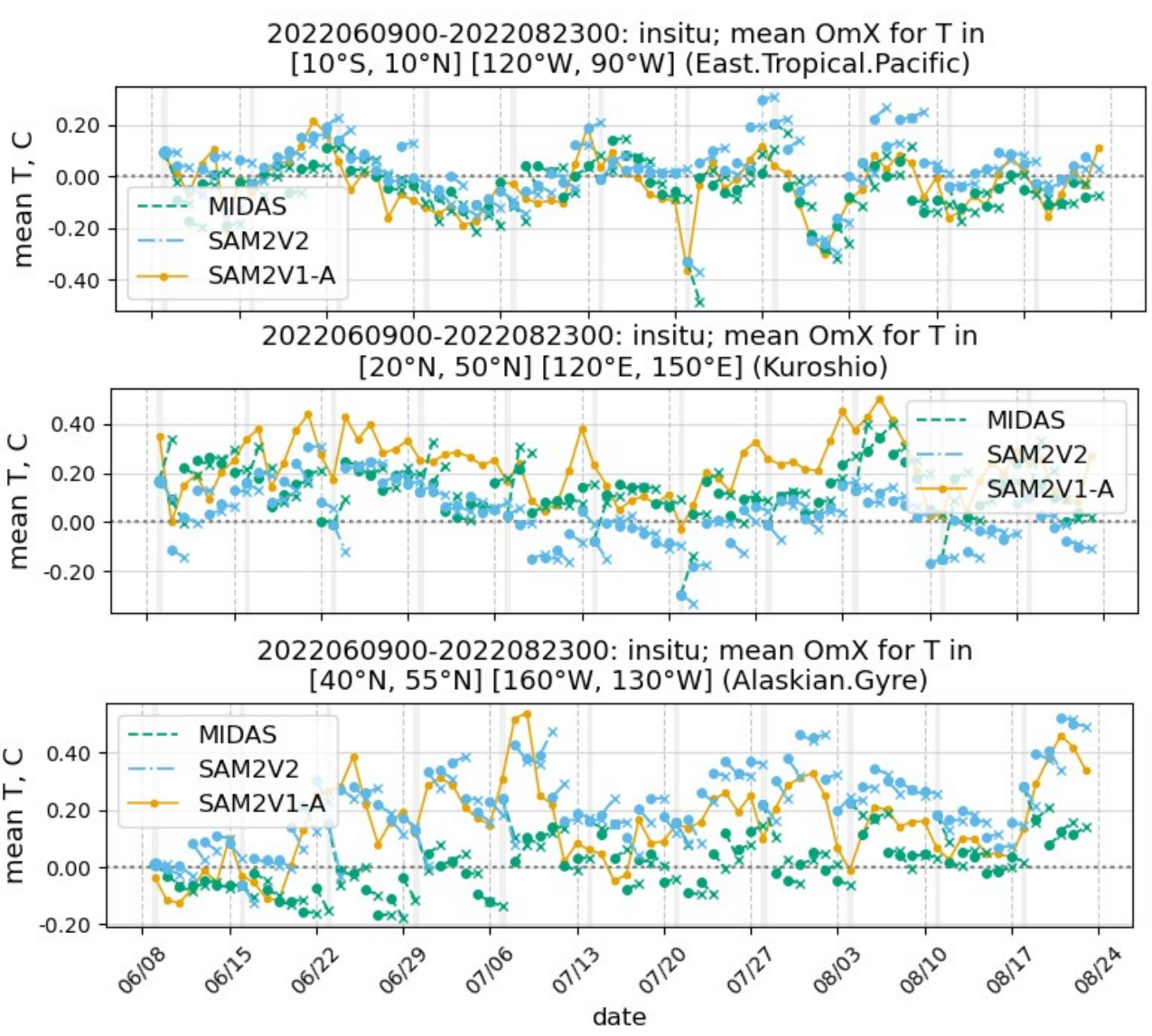
- Uses the **Local Ensemble Transform Kalman Filter (LETKF)**.
- Direct assimilation of:
  - Satellite-derived SST observations.
  - In-situ SST observations.
- Consistent methodology with ECCC's atmospheric ensemble data assimilation system.
- Daily and weekly ocean analysis components interact through a common ensemble ocean state and error representation.

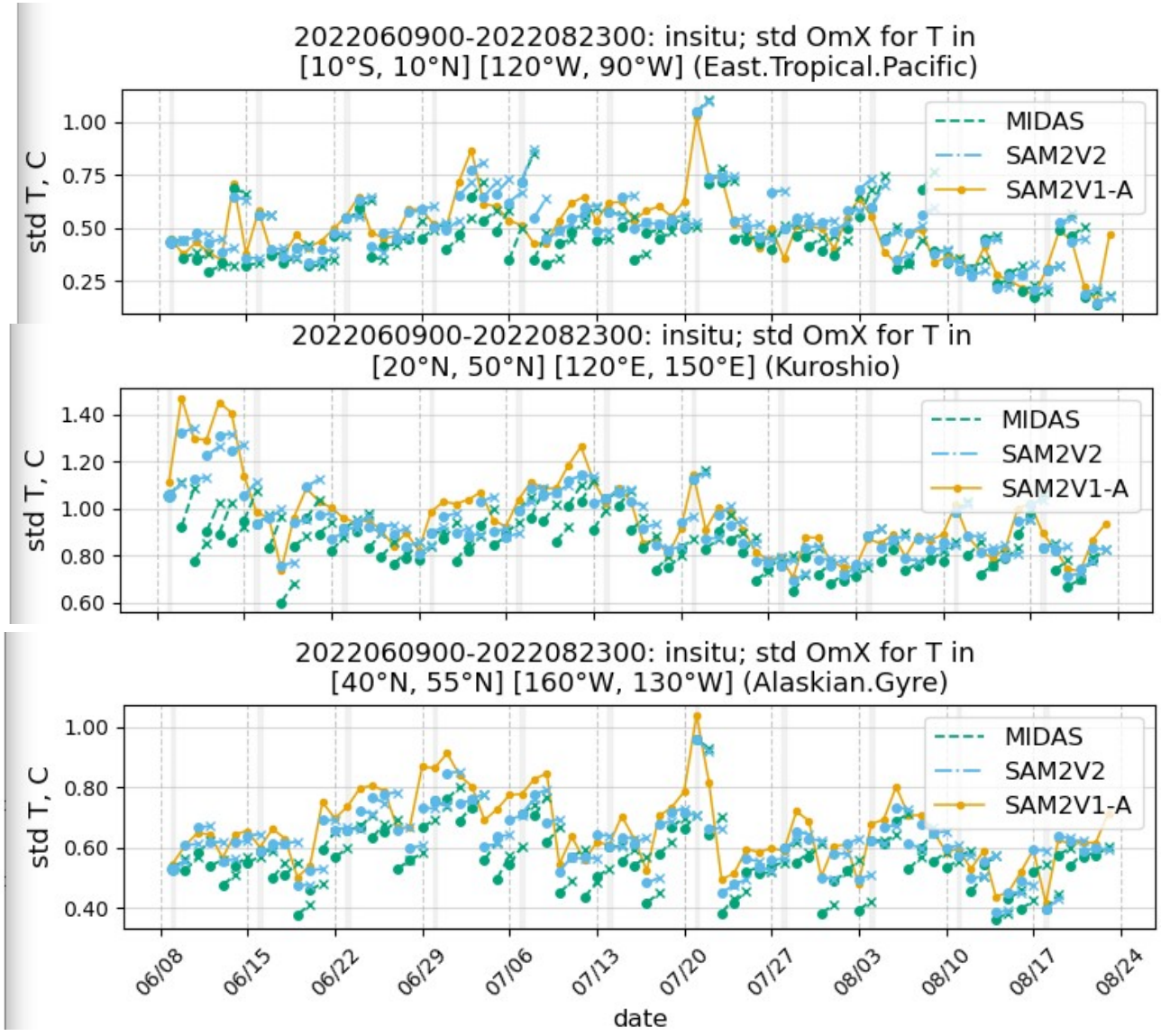
Advantages

- Dynamically evolving error estimates.
- Common framework for CDA

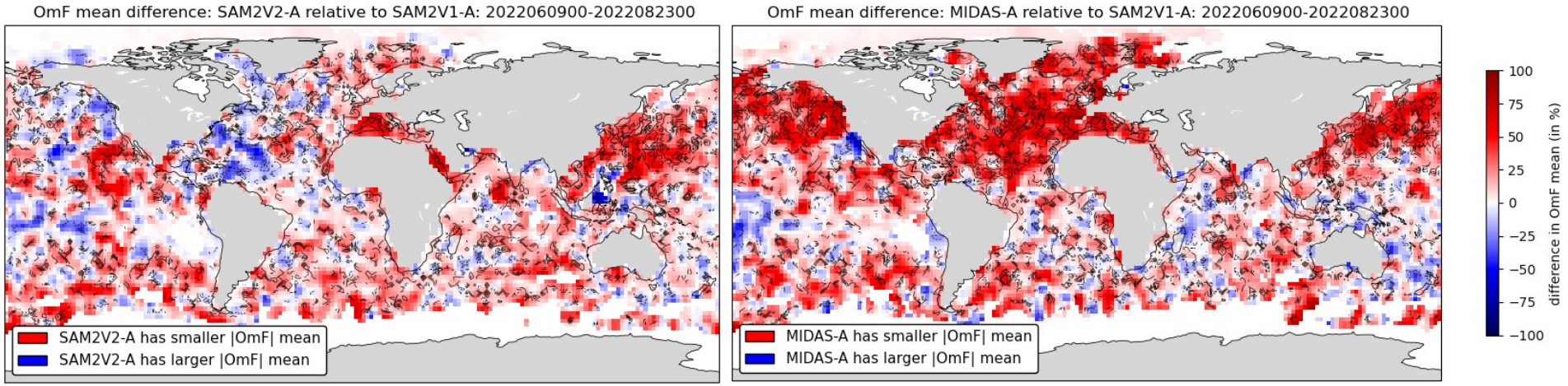
OmF error evolution for daily MIDAS, SAM2-ens and SAM2-det

Mean **STD**





Mean

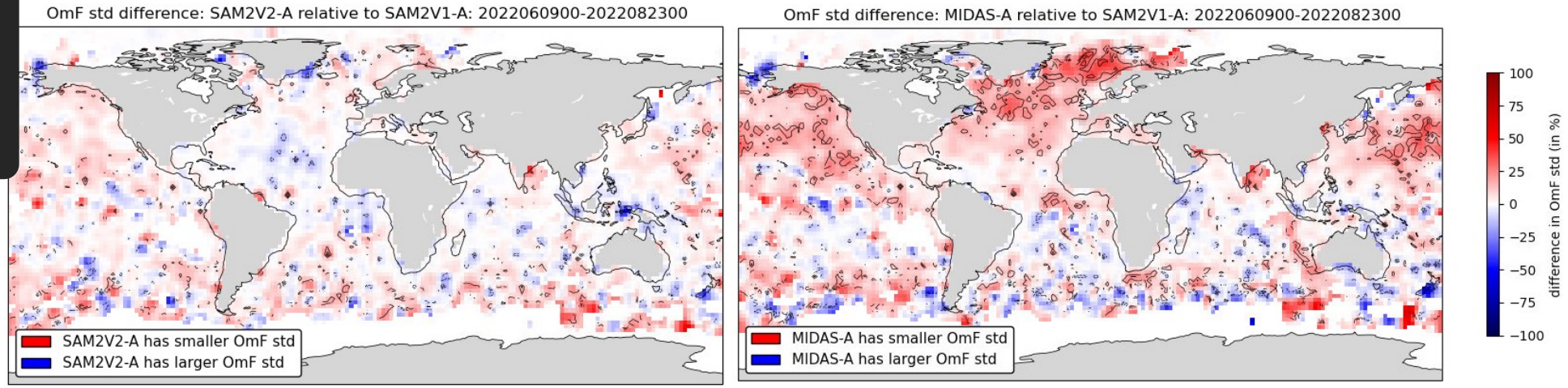


Comparing ensemble relative to determinist DA

SAM2v2 Ens vs SAM2v1 Det

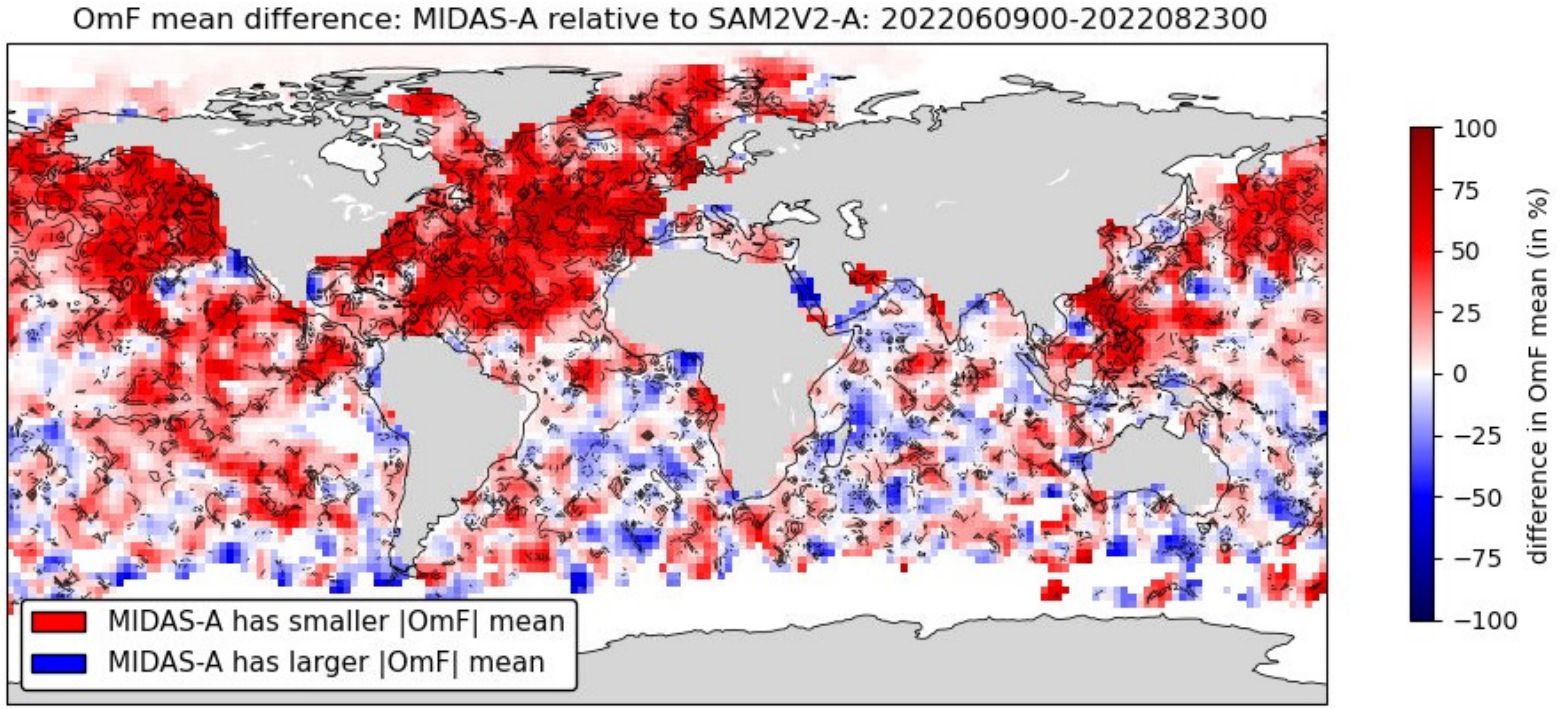
MIDAS vs SAM2v1 Det

STD

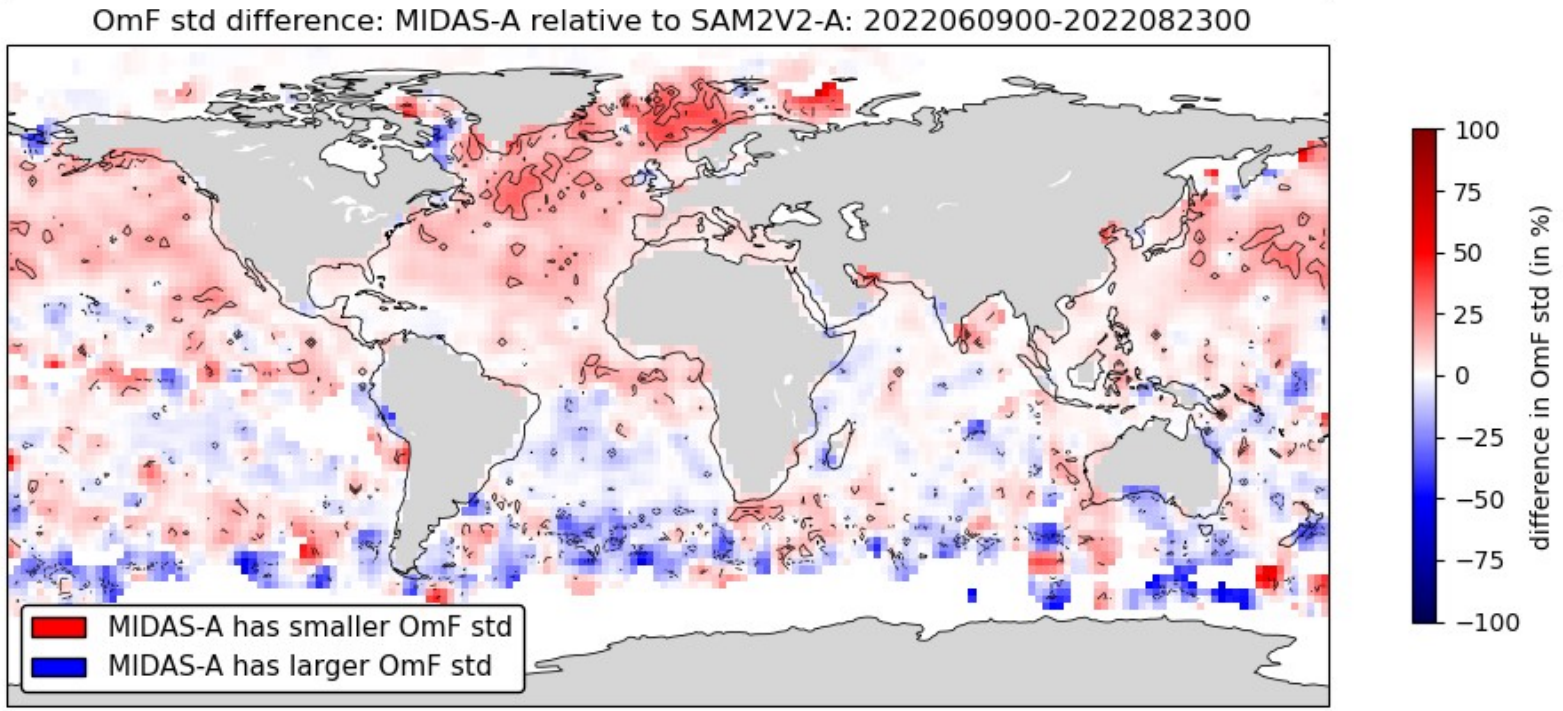


Comparing analyses: MIDAS vs SAM2V2

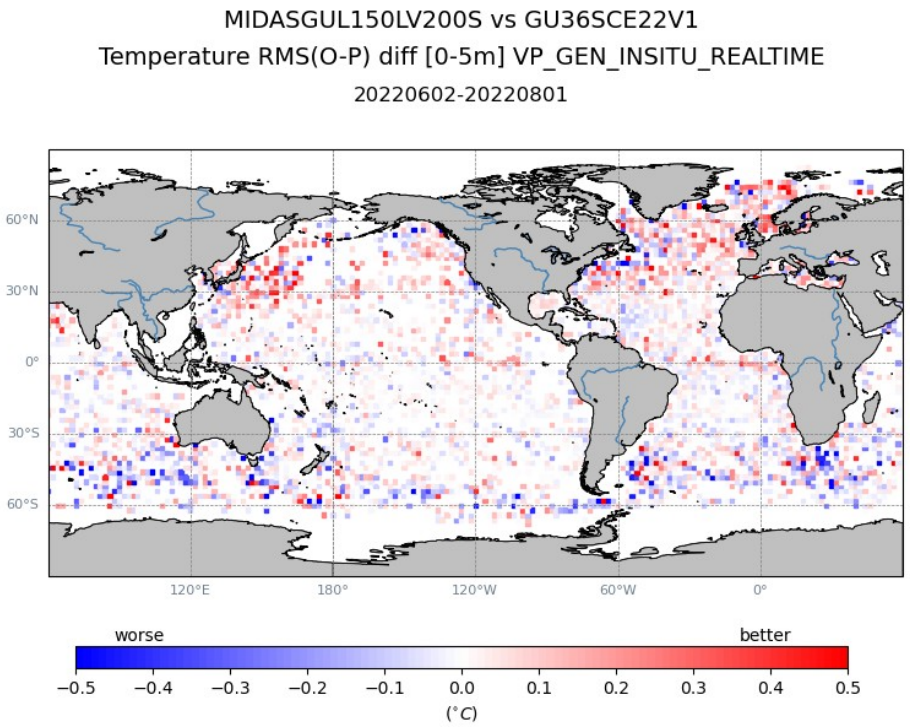
Mean



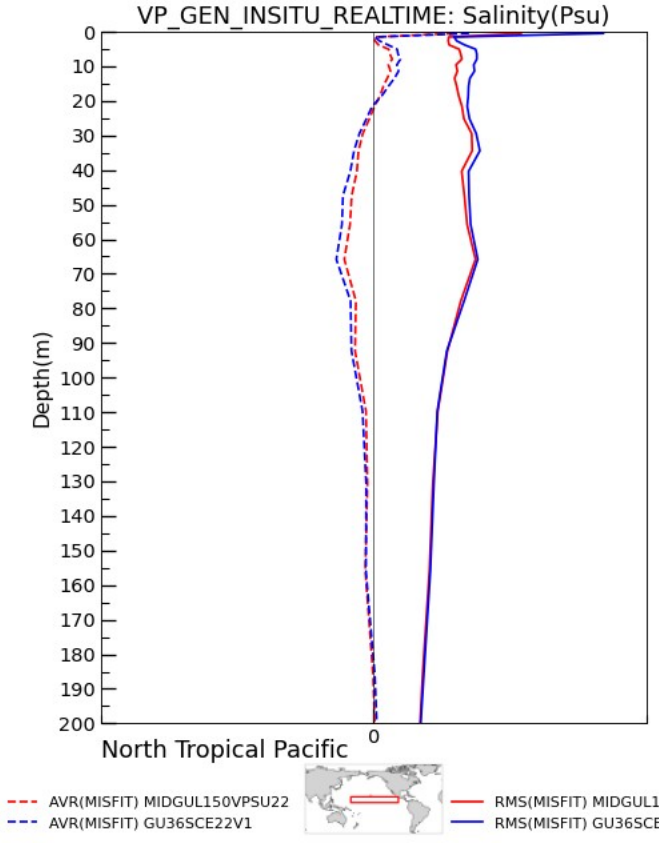
STD

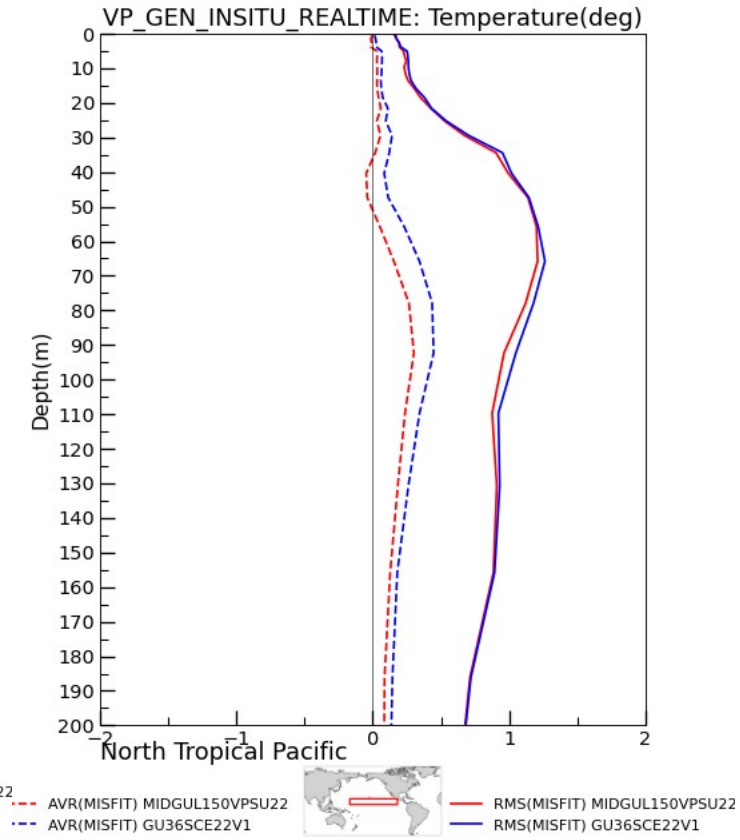


Daily MIDAS-based ensemble DA results: T/S profiles



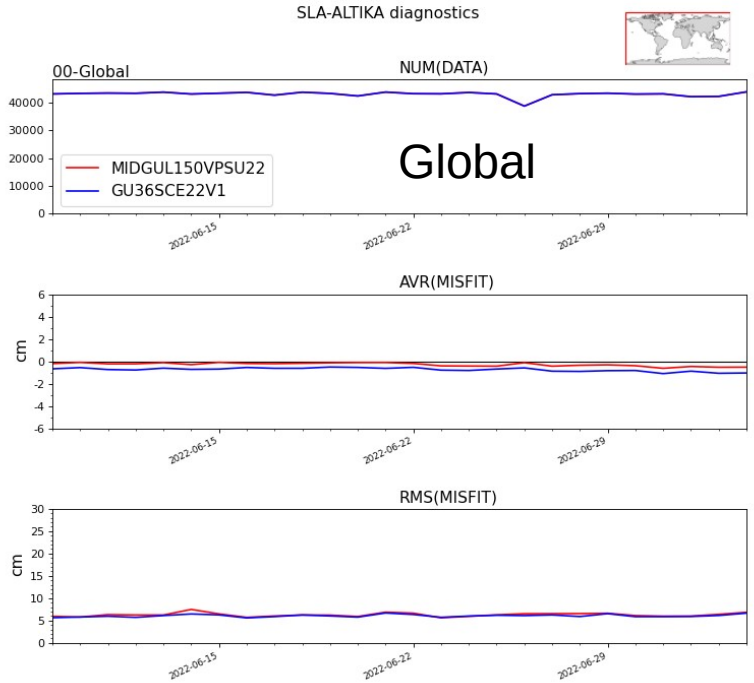
OmF RMS difference in innovations w.r.t in situ T data (0-5m)

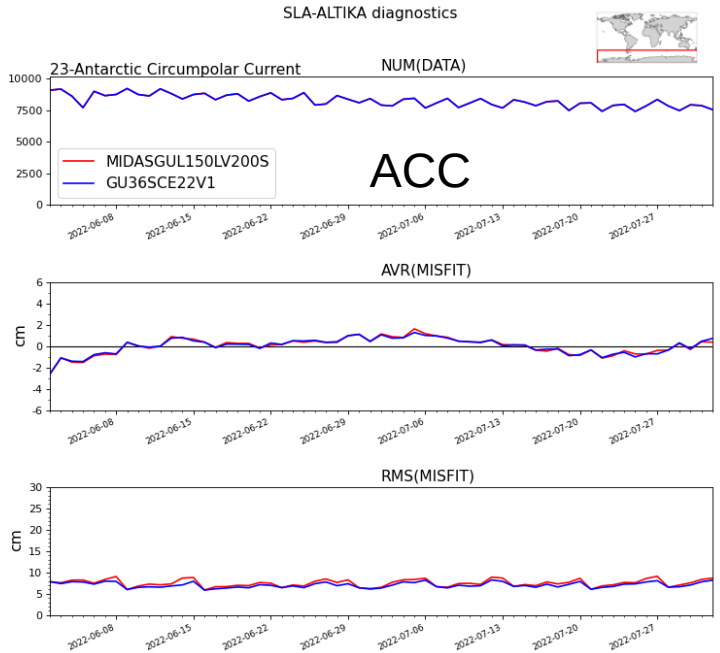




Using MIDAS daily ocean DA (comparing to ensemble SAM2 V2), we improve T and S bias and std as compared to independent Argo observations in many regions.

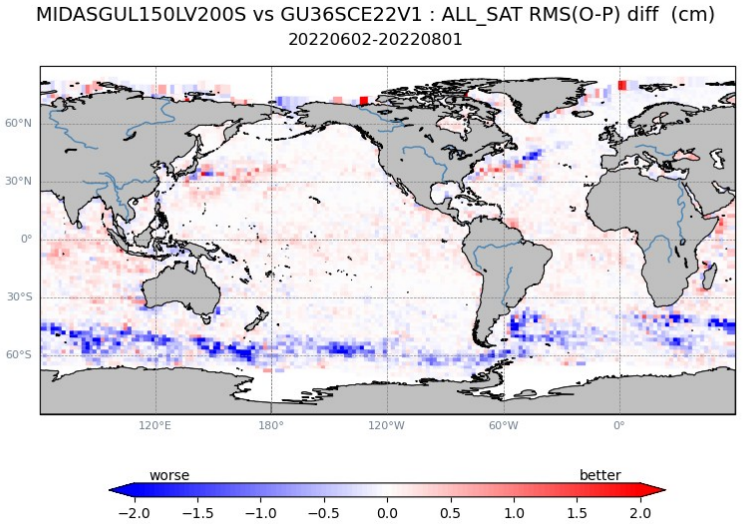
Daily MIDAS-based ensemble DA results: SLA





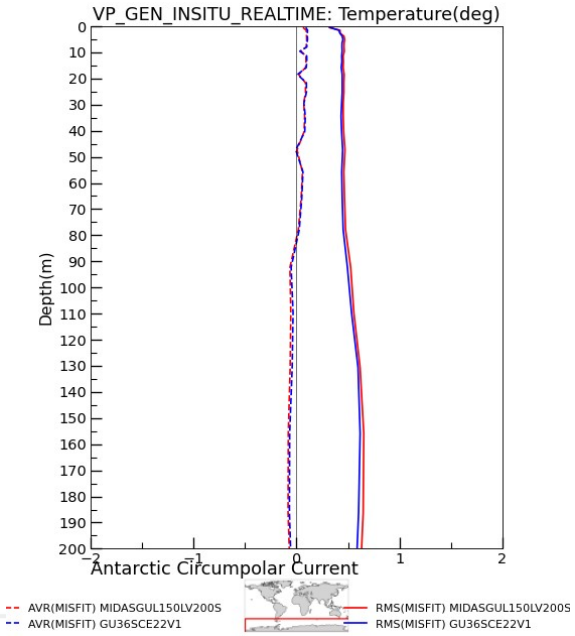
Temporal evolution of bias and std OmF errors for MIDAS (red) and SAM2 V2 (blue) with respect to ALTIKA observations

SLA innovations degraded in MIDAS over ACC.

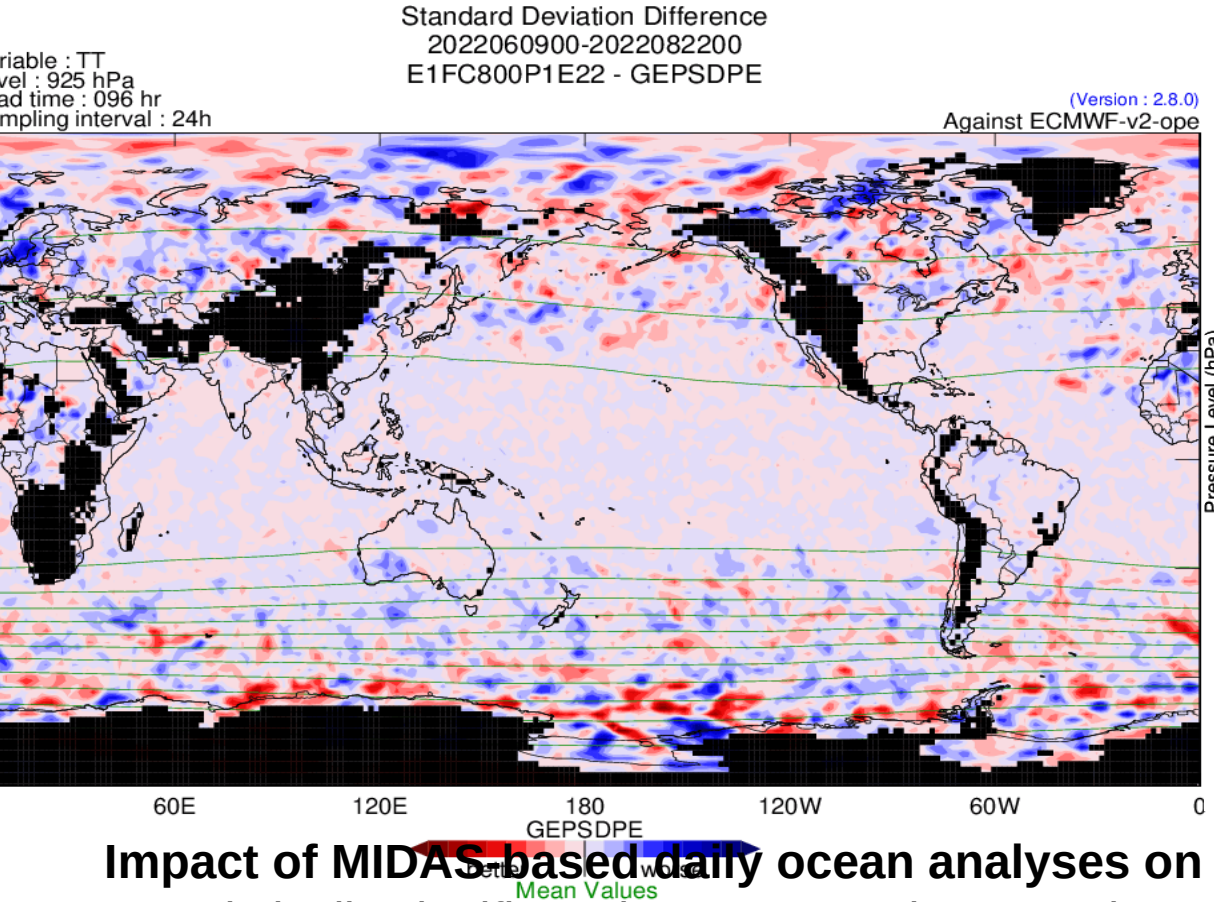


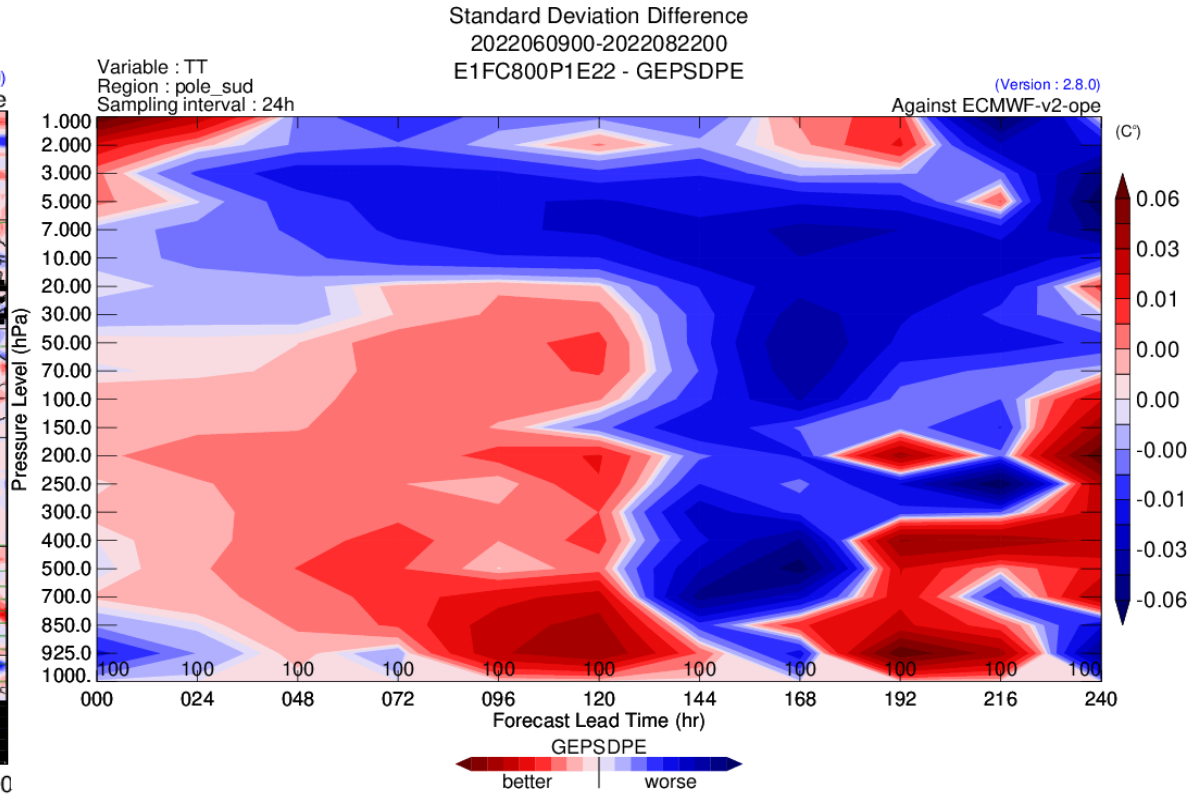
L_v = 200m

OmF errors (bias and std) for T profile in ACC



Daily MIDAS-based ensemble DA results: Coupled forecasts

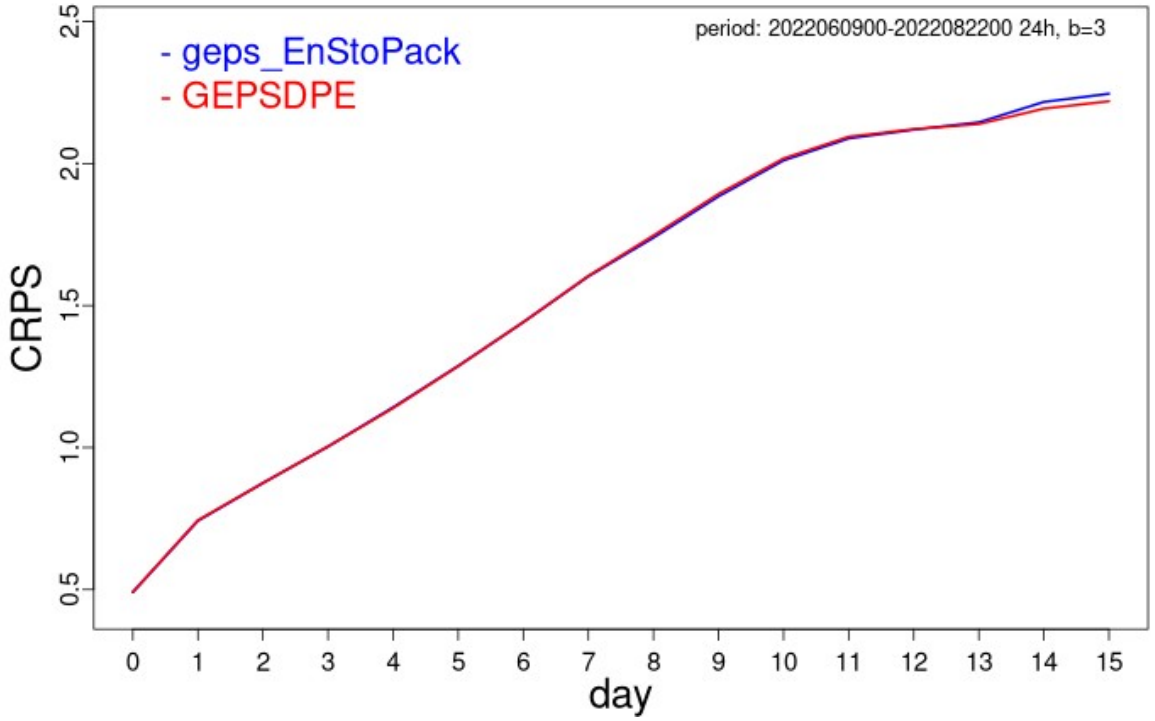


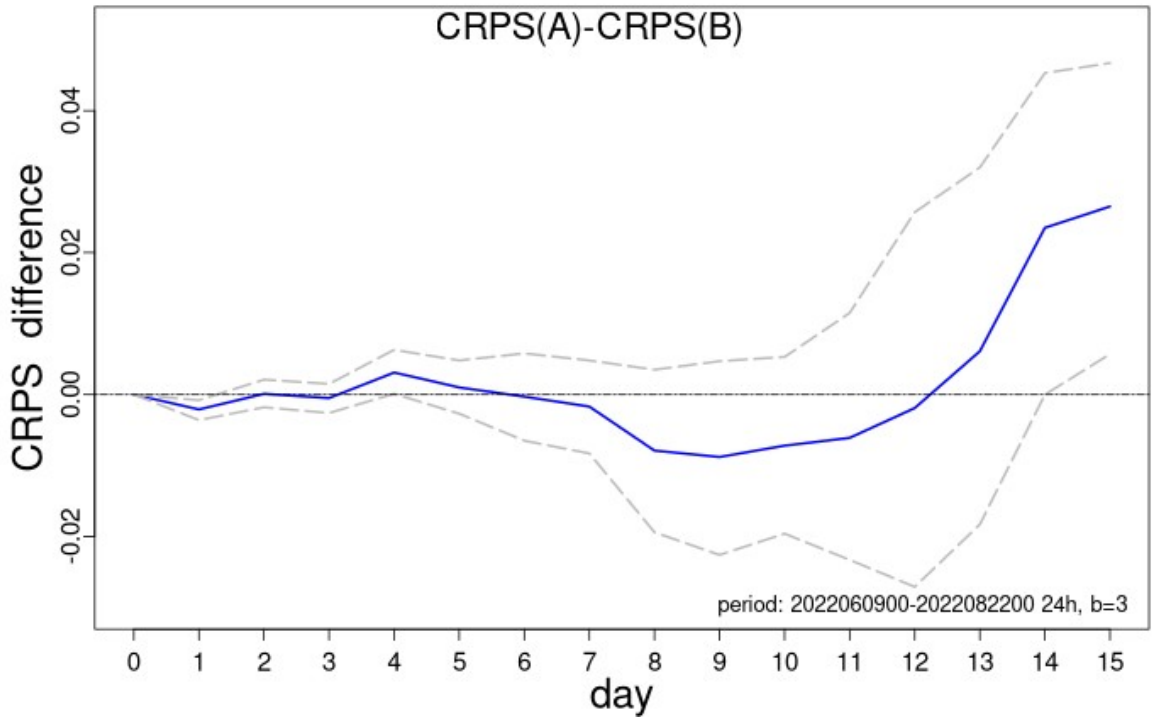


Impact of MIDAS-based daily ocean analyses on atmospheric prediction:

A statistically significant improvement in Antarctic near-surface air temperature forecasts is observed, despite degraded ocean forecast skill as measured by SLA OmF standard deviation.

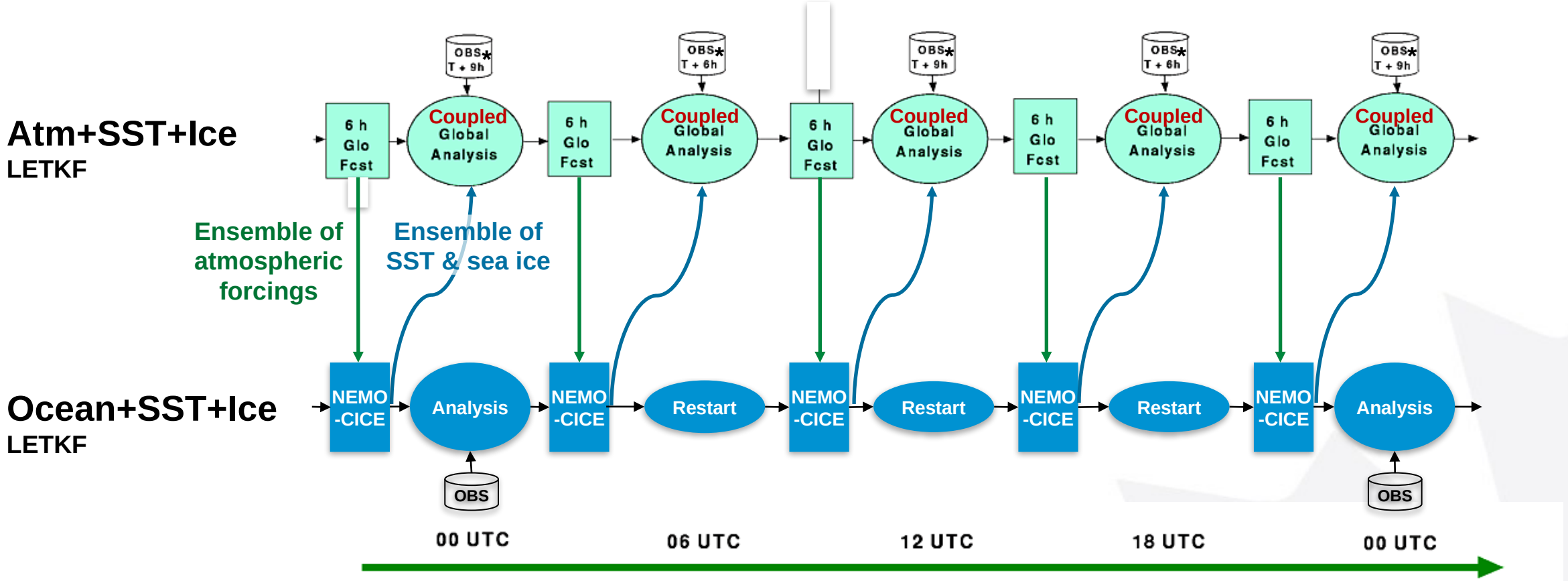
Daily MIDAS-based ensemble DA results: Coupled ensemble forecasts





An example GEPS score (Global) of using MIDAS-based (in red) daily ocean analyses on atmospheric forecasts vs SAM2 V2 (blue)

Towards coupled ensemble DA



Atmospheric assimilation 4 times per day, Ocean assimilation 1 time per day

SUMMARY

- We have created two global ensemble versions of GIOPS: SAM2v2 (weekly) and MIDAS (daily)
  - Moving towards operationalizing as part of Innovation Cycle 5 (July 2027)
- Use of ensemble can reduce uncertainty at length scales that are not constrained by observation (Peterson et al., 2025)
- The ensemble mean shows considerable improvements in fit to observations over deterministic version.
- But can further provide uncertainty ranges or probabilities to diagnostics
  - like sound speed profiles / likelihood of sound ducts
- Daily ocean MIDAS-based ensemble DA:
  - improved SST analyses and forecasts relative to both deterministic and ensemble SAM2 configurations;
  - generally neutral impact on GEPS and GDPS forecast skill, indicating successful integration of ocean ensemble DA without degrading NWP performance;
  - degradation of SLA in the ACC region (work in progress);
  - major step toward CDA by establishing an ocean DA compatible with other components of Earth system DA;
  - Future work: developing ensemble ocean-ice Strongly Coupled DA (SCDA);
  - Longer term objective: developing ocean-ice-atmosphere SCDA, initially through the interface-based coupling approach, toward improved Earth system analysis employing cross-component error covariance