NCAR
OPERATED BY UCAR
CGD Laboratory

MESACLIP
TAMU – NSF NCAR

CVP
Climate Variability and Predictability

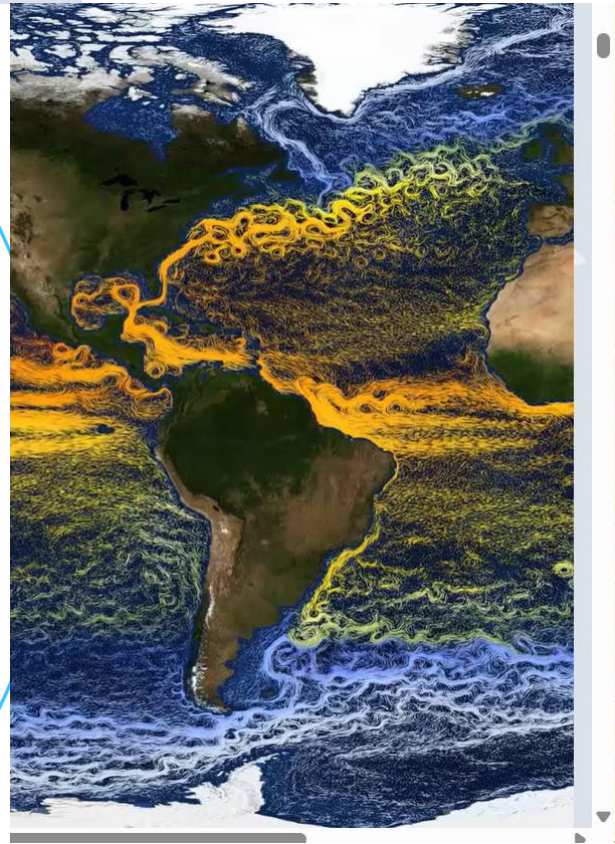
NOAA grant
NA22OAR4310615-T1-01

OMWG Meet, February 6, 2026

Predictability of ocean eddies and precipitation:

An S2S application of HRDP

Justin Small (NSF-NCAR),
Steve Yeager, Fred Castruccio, Gokhan Danabasoglu, NSF-NCAR,
Ping Chang (TAMU), Ian Grooms, Ben Lombardi (CU),
Niklas Schneider (UH), Lucas Laurindo (UM), Rhys Parfitt (FSU)



This material is based upon work supported by the NSF National Center for Atmospheric Research, a major facility sponsored by the U.S. National Science Foundation and managed by the University Corporation for Atmospheric Research. Any opinions, findings and conclusions or recommendations expressed in this material do not necessarily reflect the views of NSF.

Introduction

- Importance of ocean mesoscale eddies
  - ocean ecosystems, mixing, transport of heat, freshwater, carbon, and underwater acoustics
  - Net effect on ocean circulation
  - Air-sea interaction
- Can we predict eddies?
  - Previous literature suggests predictability horizons of between 3 to 8 weeks
  - Previous literature is not comprehensive (regional, uses a variety of methods)
- Can we predict small-scale precipitation?
  - Small-scale convection is hard to predict
  - Is there a role of small-scale SST?
- Here we use a High-Resolution Decadal Prediction system
  - Developed for decadal studies but applied here to S2S timescales
- Focus on North-west Atlantic
  - Findings can be applied to other regions of the Globe
- Standard (community) predictability metrics are used
- Approach to investigate mesoscale eddies
  - Spatial filtering to isolate spatial scales less than ~ 500km
  - Eddy tracking is also applied

Prediction experiments

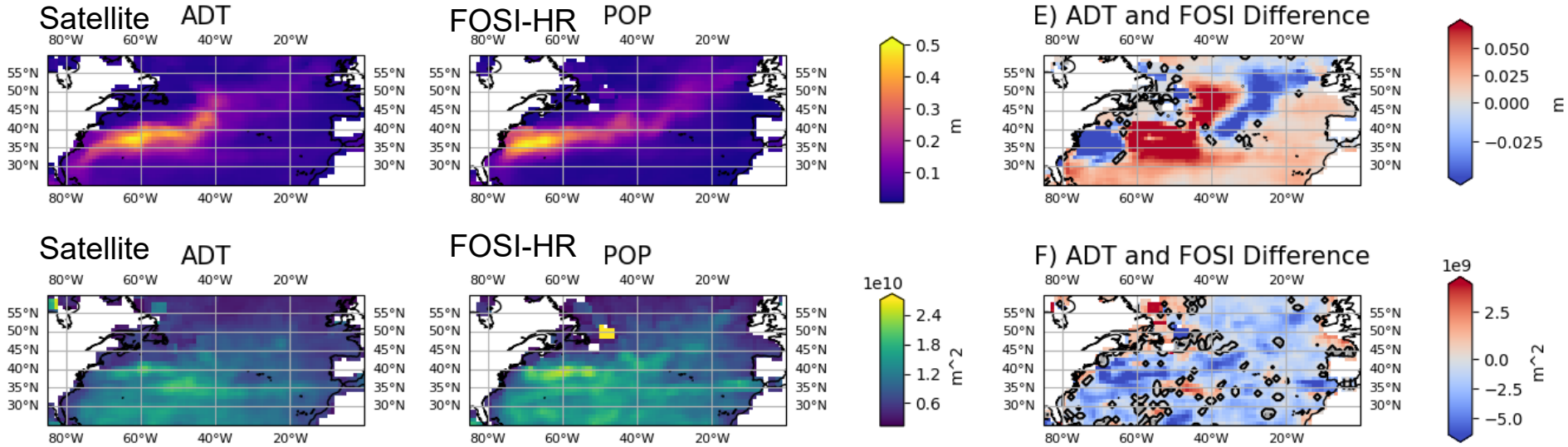
- High-Resolution Decadal Prediction experiment (HRDP)
  - **Focus on S2S timescales.**
  - CESM High-Res, fully coupled (Chang et al. 2020)
  - 1/10th deg. ocean model, 1/4deg. Atmosphere model.
  - 21 sets of November 1st start dates, every other year
  - 10 ensemble members.
  - **Perfect model experiments**
  - Yeager et al. (2023), Who Kim et al. (2023), Q. Zhang et al. (2024a,b)
  - **Initial atmosphere: JRA55 regridded.**

Yeager et al. 2023

- **Initial ocean conditions and perfect model: Forced-Ocean-Sea-Ice High-Resolution (FOSI-HR)**
  - Forced by JRA55-do (Tsujino et al. 2018)
  - **No data assimilation: eddies are present but not in correct location or amplitude – chaotic system**
  - Chassignet et al. (2020), Yeager et al. (2023), Small et al. (2024), Little et al. (2024)

A perfect model? (eddy SSH properties)

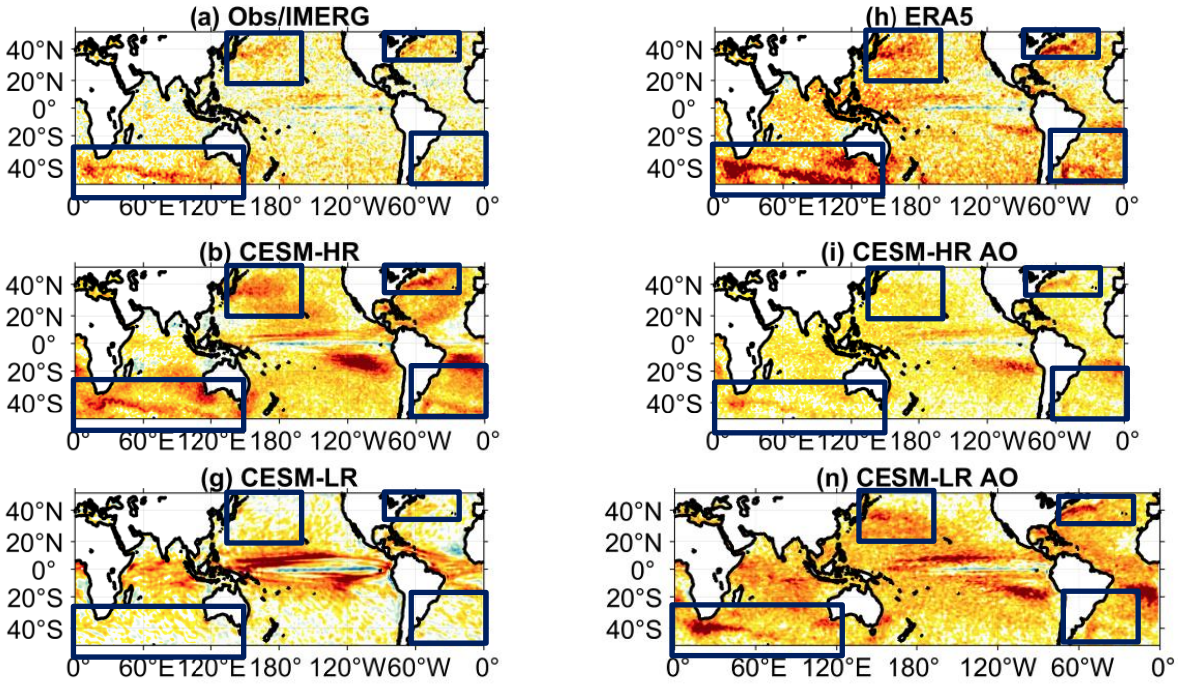
Amplitude (first row) and area (second row)

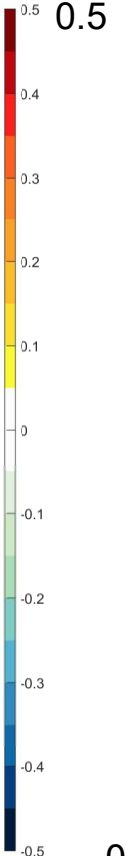


C. Pegliasco, A. Delepoulle, E. Mason, R. Morrow, Y. Faug´ere, and G. Dibarboure. Meta3.1exp: a new global mesoscale eddy trajectory atlas derived from altimetry. Earth System Science Data, 14(3):1087–1107, 2022a. doi:https://doi.org/10.5194/essd-14-1087-2022.

Benjamin Lombardi, Ian Grooms, and William Kleiber. Demographics of Mesoscale Eddies in an Eddy-Permitting Ocean Model and Reanalysis. Manuscript submitted to Progress in Oceanography.

A perfect model? (air-sea interaction)





Boxes denote strongly eddying regions of the Globe.

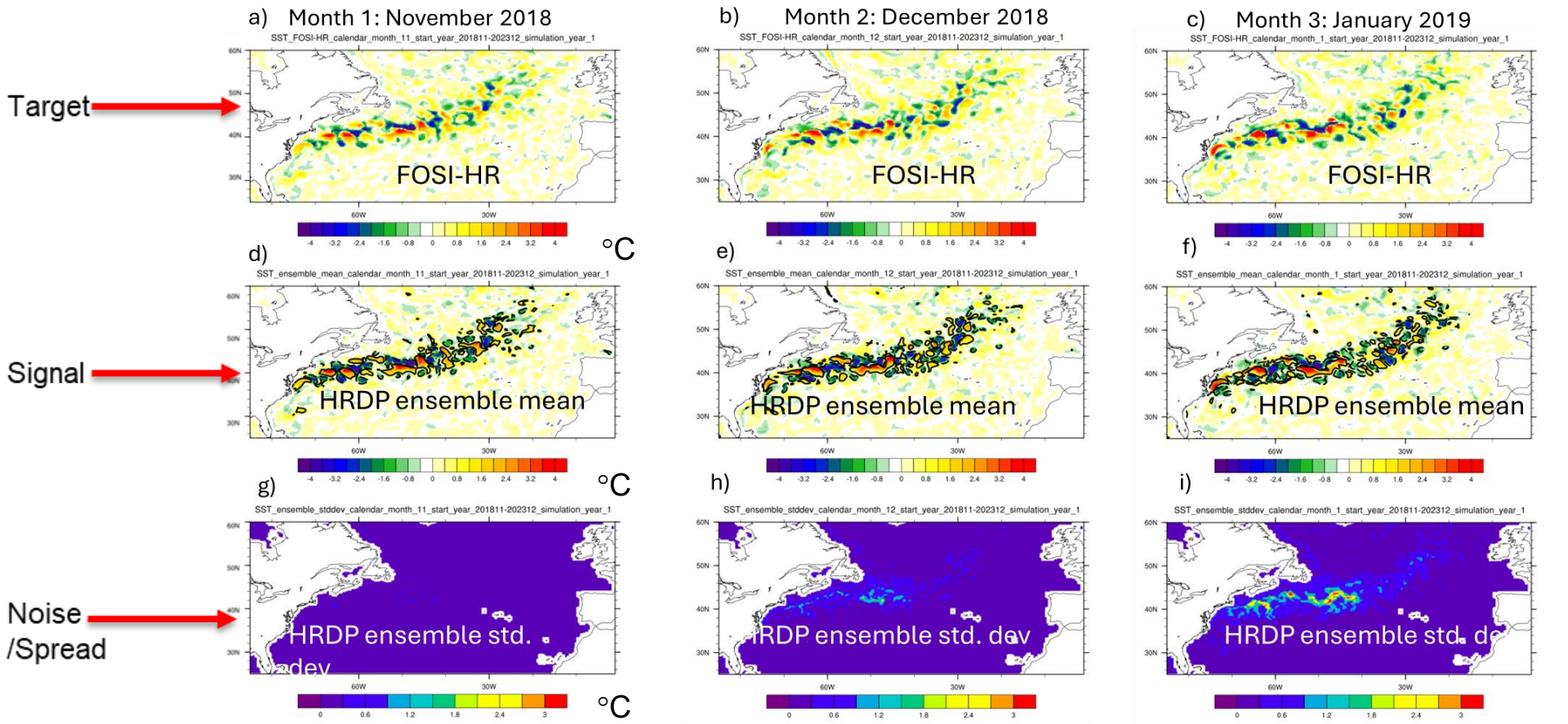
Left column are fully coupled systems. Right column is atmosphere-land-only (inc. ERA5).

Simultaneous correlation coefficients between monthly highpass-filtered SST and precipitation anomalies.

Wang, X., P. Chang, R. J. Small, D. Fu and M. Lyu, 2026. Precipitation response to mesoscale SST variability: insights from observations and multi-resolution models. In press, GRL.

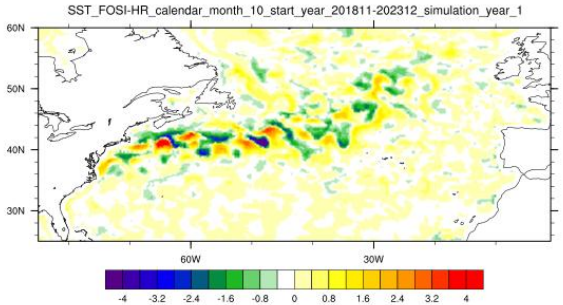
Case study. Predictions of small-scale SST.

Data high-pass filtered in space and climatology removed.

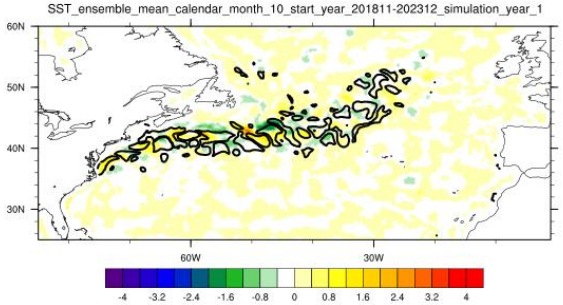


12th month

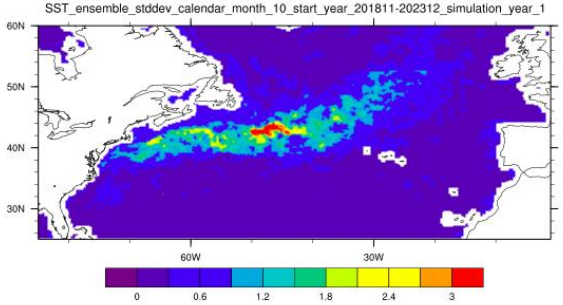
Target



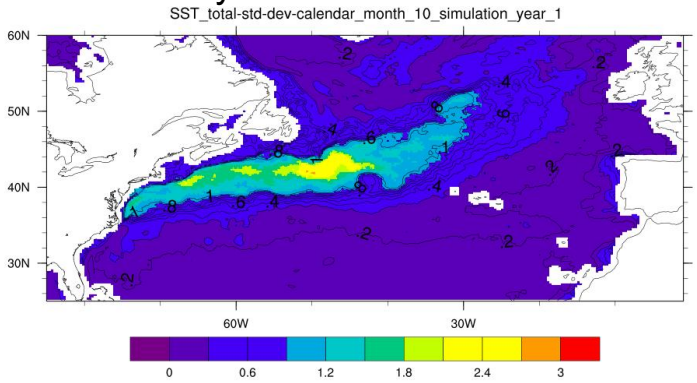
Signal



Noise /Spread



Standard deviation of natural variability



Some predictability metrics (Yeager et al. 2023 and references therein)

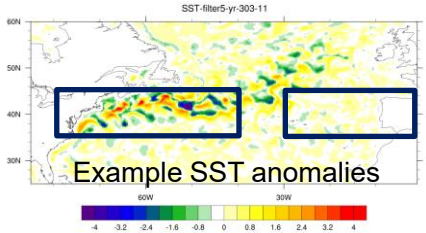
- *Anomalies are relative to appropriate climatologies*
  - *For HRDP, this is lead-time-dependent due to model drift*

- *Anomaly Correlation Coefficient (ACC) correlates HRDP signal and FOSI across start dates*
  - *Here signal=ensemble mean.*
  - *mainly phase information*

- *Mean square skill score (MSSS) relates the amplitude and phase of HRDP signal and FOSI across start dates*

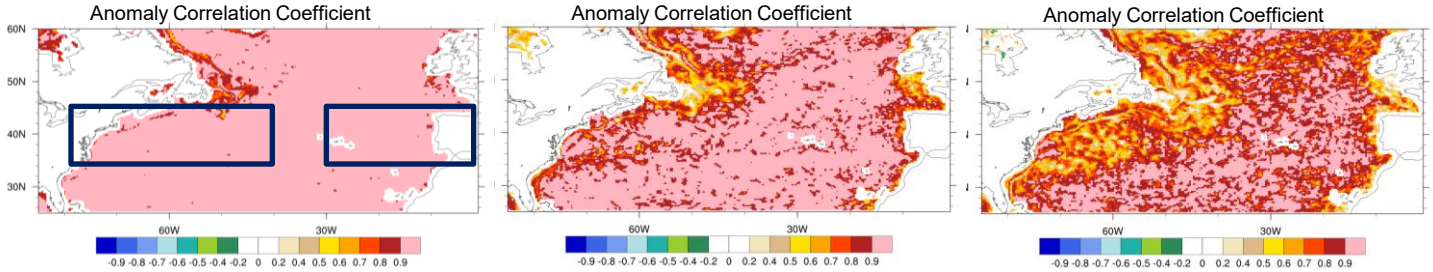
- *Signal-to-Total (S2T)= ratio of standard deviations of the signal and of the total variability of HRDP*
  - *"How well does the model predict itself?"*
  - *Equivalent to correlations between the signal and individual ensemble members*

Summary Skill scores (monthly)

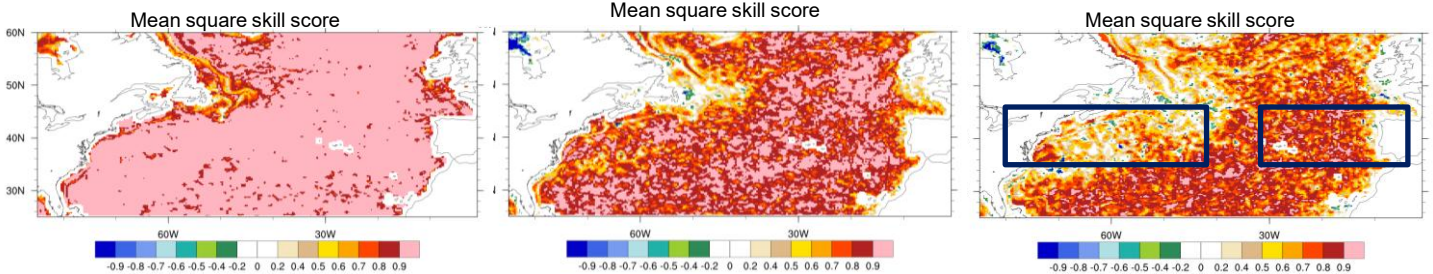


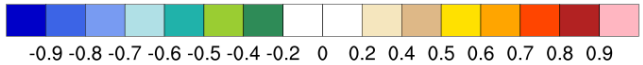
First month November | Second month December | Third month January

Anomaly Correlation Coefficient (ACC) correlates HRDP and FOSI across start dates



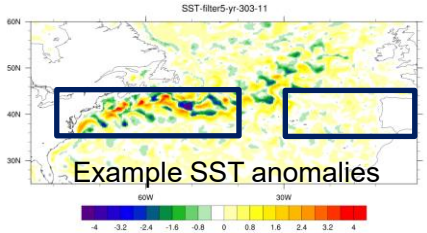
Mean square skill score (MSSS) relates the amplitude and phase of HRDP and FOSI across start dates





For first 3 months, Gulf Stream eddies are reasonably predicted relative to the perfect model (FOSI).

Summary Skill scores (monthly)

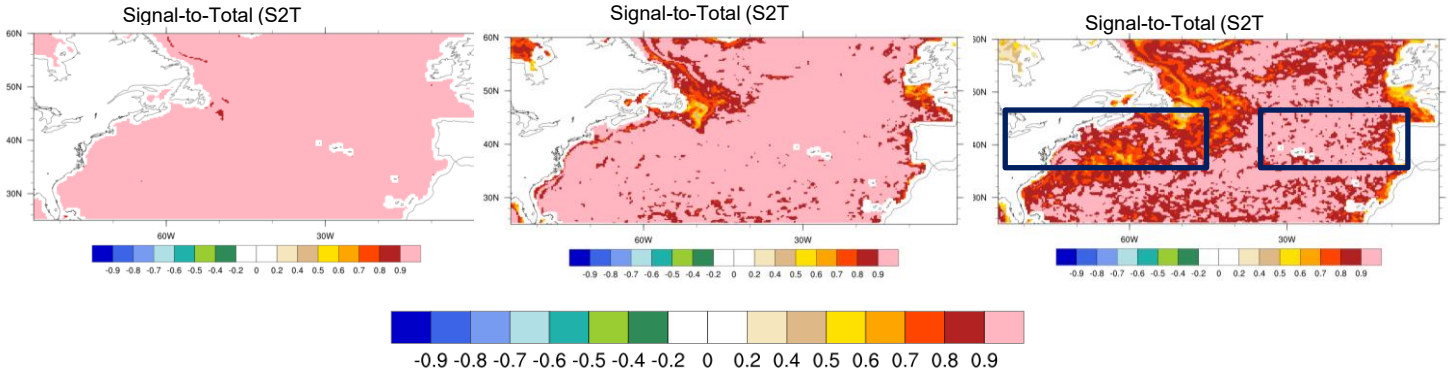


Signal-to-Total (S2T)= ratio of standard deviations of the signal and the total variability.
Here signal=ensemble mean.

First month November

Second month December

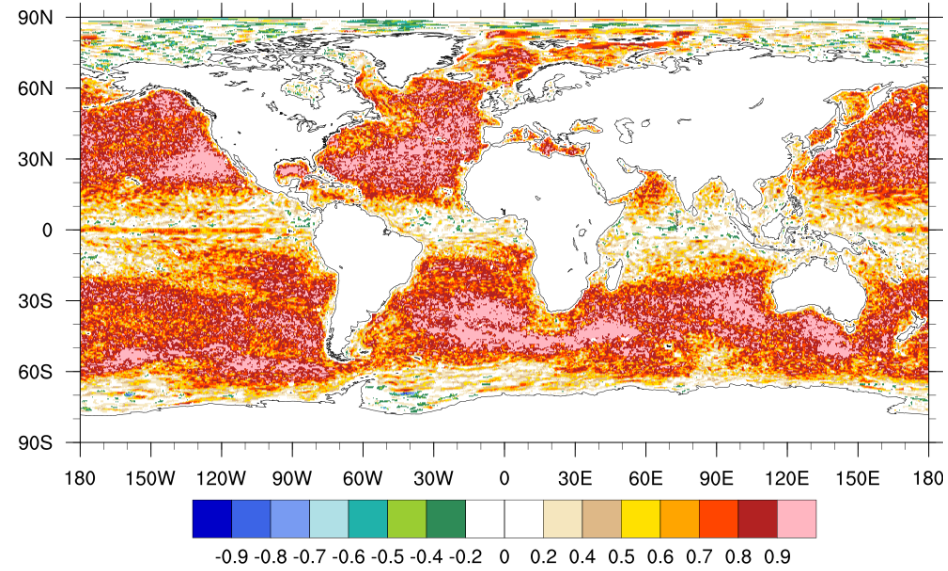
Third month January



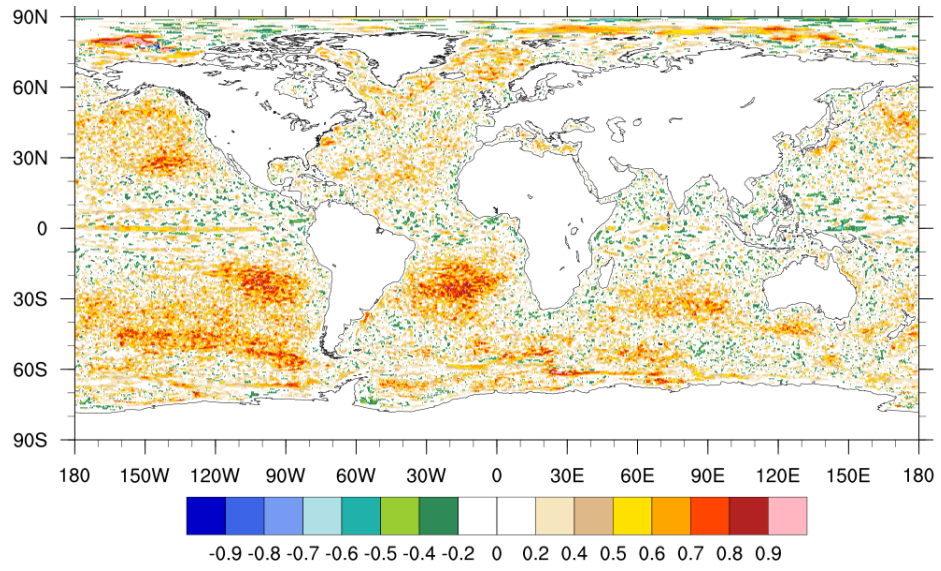
For first 3 months, Gulf Stream eddies are reasonably predicted relative to the perfect model (FOSI).

Summary Skill scores (Global)

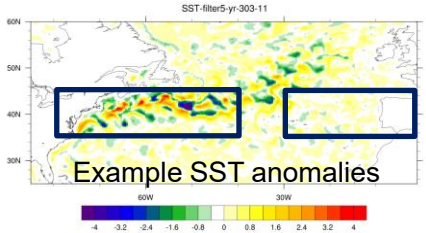
ACC month 3

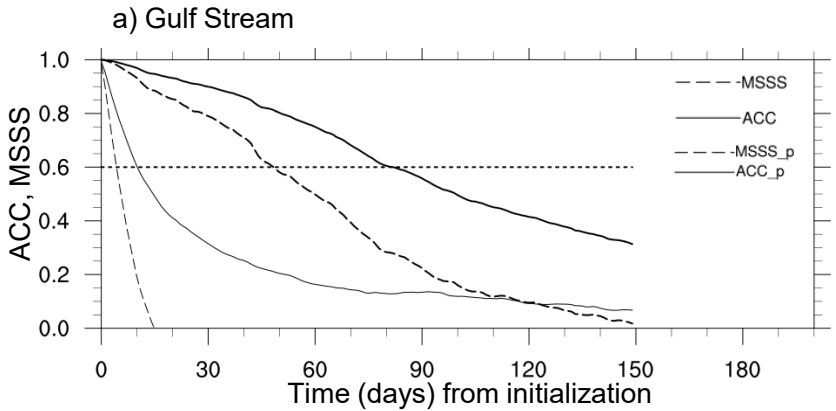


ACC month 12



Summary Skill scores (daily)





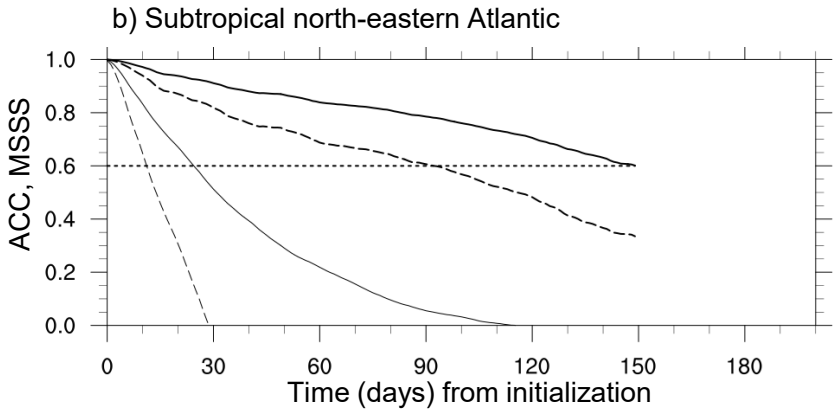
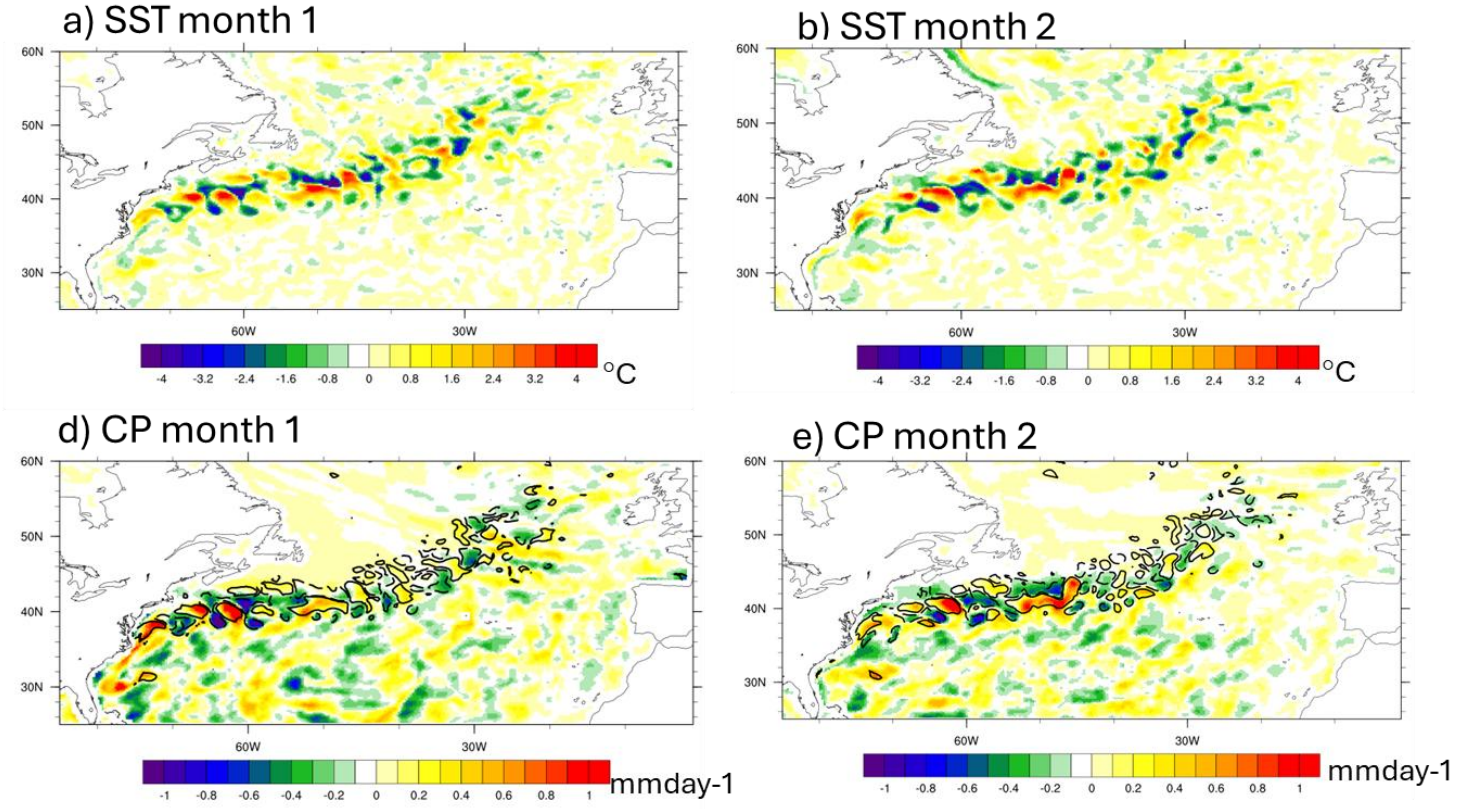


Fig. 5. Skill scores area- averaged over particular regions as a function of time. The scores are MSSS and ACC, see legend in a). The thin dotted line marks a score of 0.6, which may be used as a reference. a) Gulf Stream region, 35° N to 45° N, 80° W to 40° W, b) subtropical north-eastern Atlantic, 35° N to 45° N, 30° W to 1° W, *Thin lines denote the equivalent scores for persistence (labelled "_p").*

Case study. SST and convective precip (monthly)



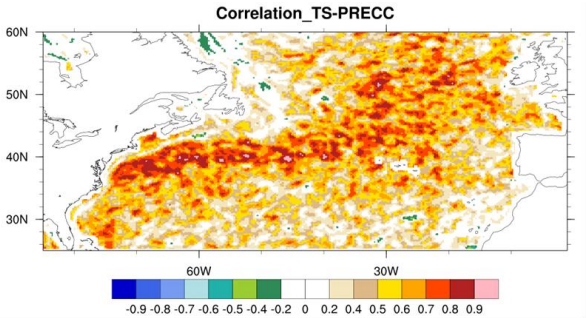
Ensemble means of SST and of convective precipitation (CP), for a start date of Nov. 1st 2018. a-c) SST anomalies, month 1-3, d-f) corresponding CP anomalies. CP=convective precip.

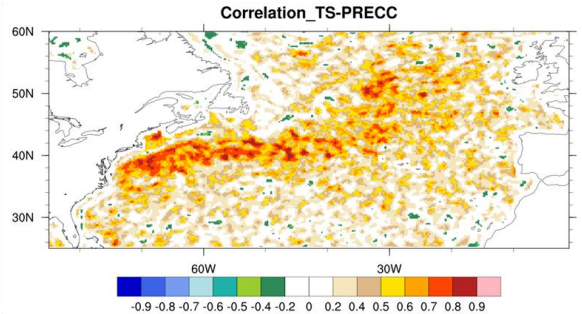
Prediction experiments

Correlations of ensemble-mean SST and ensemble-mean precip

10-members

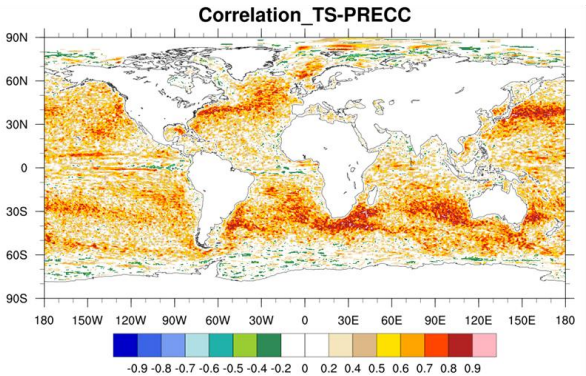
1-member

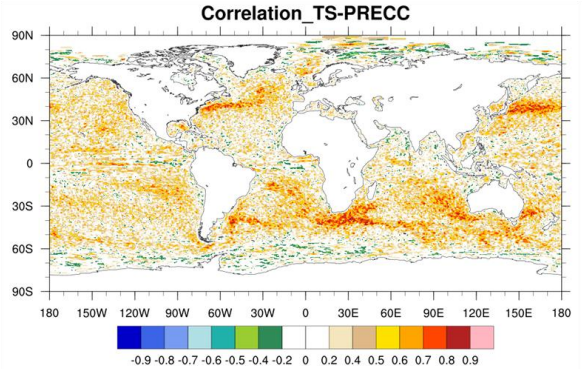




November

See also Kirtman et al. 2017, CHAOS, Effect of ensemble averaging

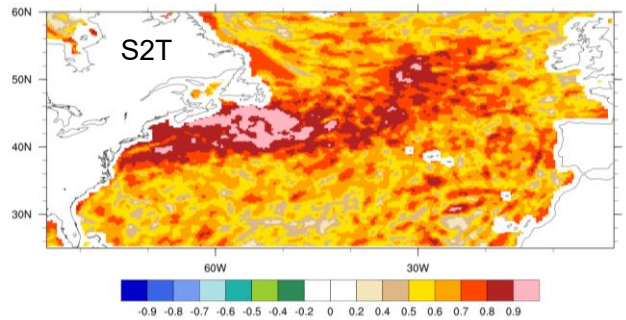




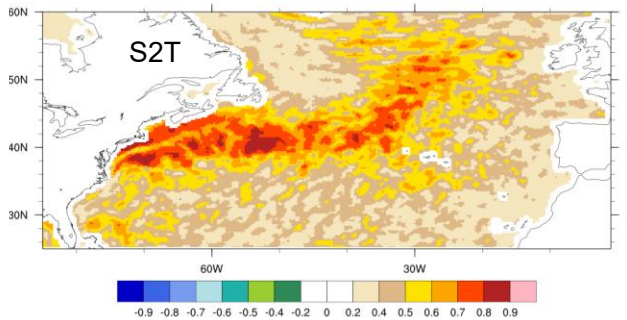
This is not a skill score – next slide shows a skill score.

Signal-to-Total for convective precip (HRDP, filtered)

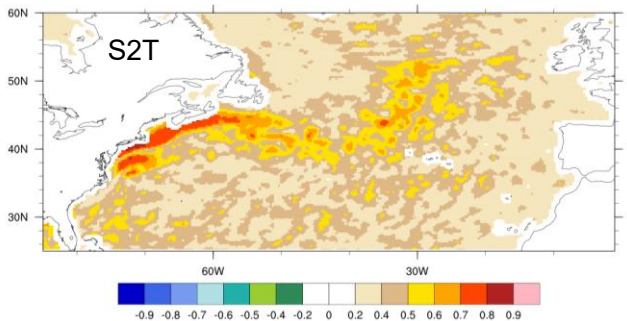
a) First month



b) Third month



c) Fifth month



Signal-to-Total for convective precip (HRDP, filtered and unfiltered, and DPLE)

DPLE= standard resolution decadal prediction, Yeager et al. 2018 BAMS

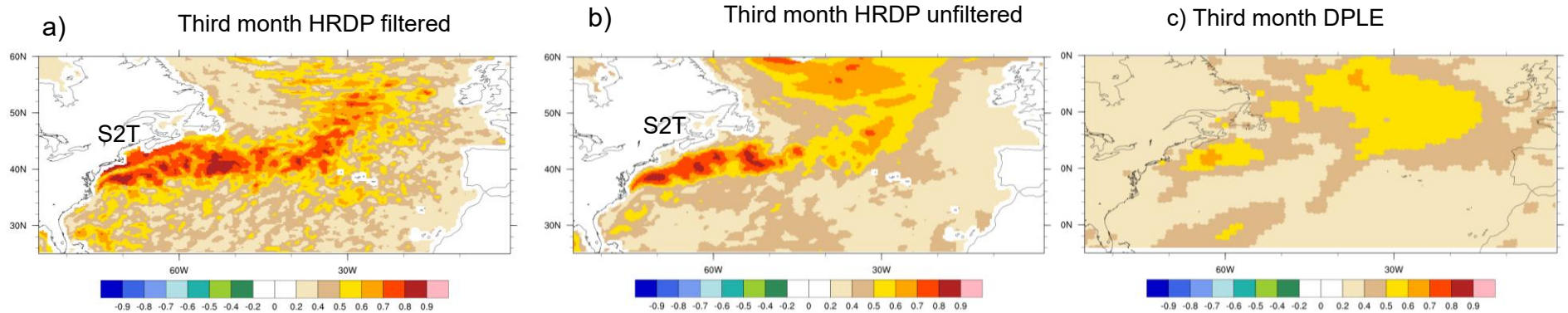


Fig. .Left: filtered, Middle) *unfiltered data. Right: DPLE*

Conclusions/Limitations/way-ahead

- Conclusions
  - The predictability of ocean eddies, and associated air-sea interactions, have been quantified using a high-resolution coupled prediction system. Applying a perfect model framework, some of the key results are:
  - For spatial scales less than around 500km, the predictability timescale for mesoscale SST in strongly eddying regions is between 50 and 70 days and it is much longer (e.g. 100 to 150 days) for weaker eddying regions.
  - The small-scale convective precipitation signal may also have some predictive skill, as seen within HRDP (S2T metric).
- Limitations:
  - Convective precip is only ~ one half of total precip in this region in CESM-HR
  - The targeted precipitation is over the ocean!
  - Perfect model?
- Way ahead:
  - Analysis of different timescales – down to daily (important to precip)
  - Non-local response of precip (storm track) to eddy field

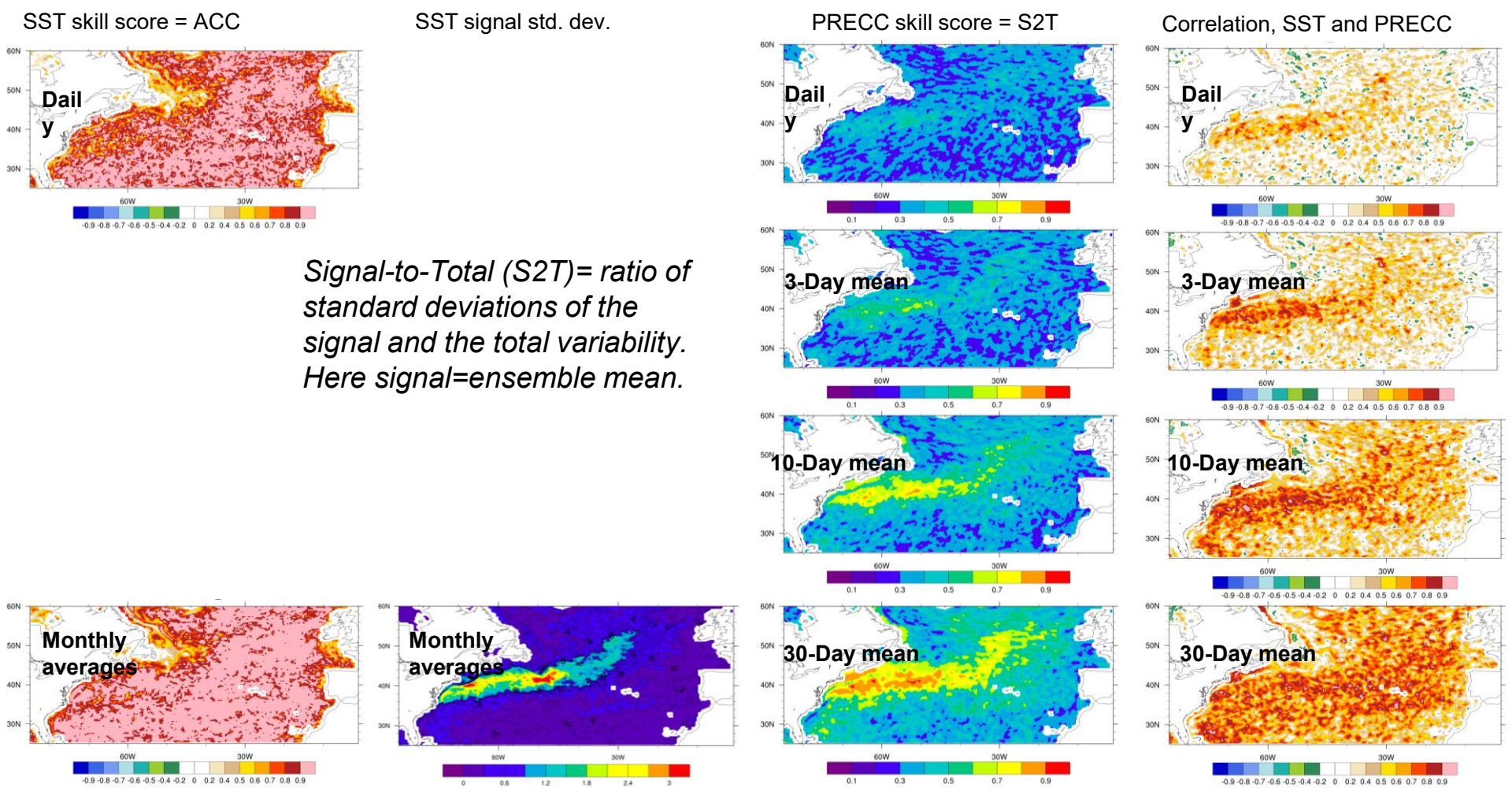
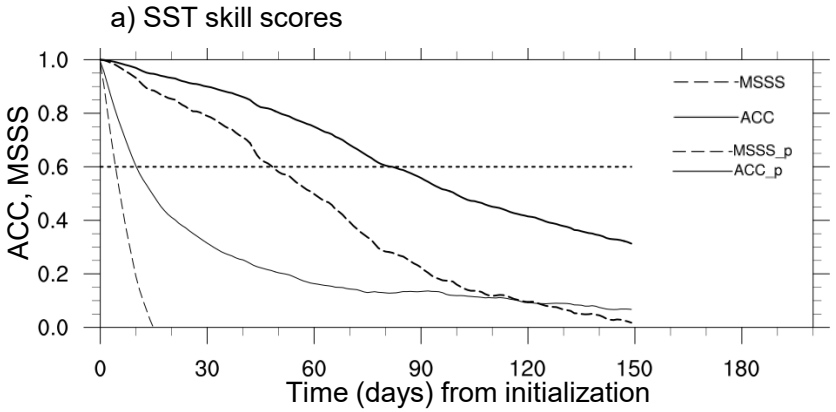
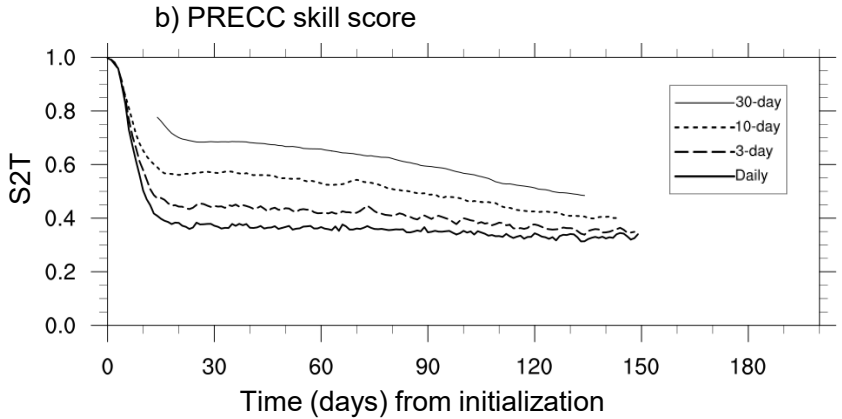
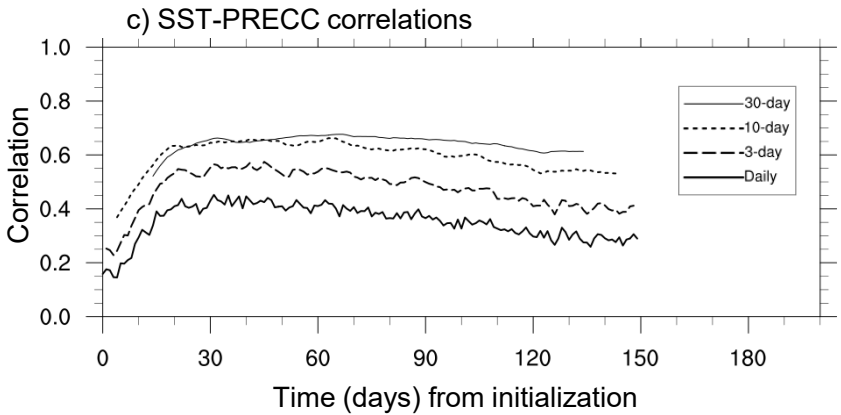


Fig. B. Skill scores and associated metrics for day 45 from daily data (or month 2 for monthly data). Left to right: SST skill score (ACC), SST signal standard deviation (signal=ensemble-means), PRECC skill score (S2T), and correlation between SST and PRECC. Top to bottom: timescales increase from daily to 30-day. Note that the SST skill score and std. dev. are not yet computed for all time-scales. All data has a spatial high-pass filter (5deg.) applied. See next slide for interpretation.







Possible interpretation

SST skill has a relatively slow decay from the initial time: it is the timescale for eddy predictability. b) Precipitation skill has two time scales: the first ~ 10 day phase in the initial condition/weather regime, with high skill in the first week then rapid decay. There is a slower reduction of skill with time after that which should relate to the correlations with SST shown in panel c). c) At short time-scales < 20 days the initial condition has an effect and the ICs for the atmosphere and ocean on small-scales are not matched. After that period the correlations reach a maximum at around 40-60 days, then slowly decay thereafter.

I don't fully understand the "asymptoting behavior" of PRECC skill scores at longer timescales (panel b)– why does it not just tend to zero with time? Maybe the timescale for decay to zero is extremely long?

Fig. A. SST and PRECC metrics area- averaged over particular regions as a function of time, based on daily data. Gulf Stream region, 35° N to 45° N, 80° W to 40° W, a) Skill scores ACC and MSSS for SST (see legend). *Thin lines denote the equivalent scores for persistence (labelled "_p").* b) Skill score S2T for PRECC. The thick solid line is for daily averages of PRECC, dashed and thin lines have a running average applied (number of days shown in the legend). c) The correlation of ensemble-mean SST and ensemble-mean PRECC based on daily data. The thick solid line is for daily averages of SST and PRECC, dashed and thin lines have a running average applied to PRECC. All data has a spatial high-pass filter (5deg.) applied.