CLIMATE ANALYSIS FORECAST ENSEMBLE system

CAFE60v1: The CSIRO Climate re-Analysis Forecast Ensemble: 1960-present.

T. J. O'Kane, P. A. Sandery, V. Kitsios, P. Sakov, M.A. Chamberlain, R.J. Matear, D.T. Squire, M.A. Collier, C. Chapman S. Schroeter, I. Watterson, D. Harries, B. Sloyan, T. Moore, D.P. Monselesan, J.S. Risbey, B. Schroeter + DCFP

https://research.csiro.au/dfp/
www.csiro.au
http://nespclimate.com.au/decadal-prediction/

CSIRO

Motivation:

CAFE60 was developed to

- provide the first ensemble reanalysis of the climate over the past 60 years of sufficient size and temporal resolution that the evolving climate pdf might be accurately estimated
- generate self consistent and balanced initial conditions for O(100) forecasts each month from 1960 to present
- as part of the WMO Grand Challenge in Near Term Climate Prediction to develop an operational system capable of generating outputs sufficient to meet the conditions to become a WMO santioned global data producing centre (GDPC)
- push the boundaries of data assimilation - i.e. strongly coupled data assimilation

Reanalysis:

- 96 members
- Daily data: Atmosphere, surface fluxes, surface ocean, OBGC
- Monthly data: subsurface ocean, sea ice, land

Forecasts:

- 10 members 10 years every November
- ALCG: 96 members each month 2018-present

CAFE60v1 has completed all requirements for the WMO GDPC providing hindcast and forecasts.

Background studies:

- T.J. O'Kane, P.A. Sandery, D.P. Monselesan, P. Sakov, M.A. Chamberlain, R.J. Matear, M.A. Collier, D.T. Squire and L. Stevens (2019) "Coupled data assimilation and ensemble initialization with application to multi-year ENSO prediction", J. Climate, 32, 997—1024

- T.J. O'Kane, P.A. Sandery, V. Kitsios, R.J. Matear, T. Moore, J.S. Risbey, I. Watterson (2020) Enhanced ENSO prediction via augmentation of multi-model ensembles with initial thermocline perturbations. (J. Climate, 33, pp2281–2293)

- P.A. Sandery, T.J. O'Kane, V. Kitsios and P. Sakov (2020) State estimation of the climate system with the EnKF using variants of coupled data assimilation (EOR Mon. Wea. Rev.)

Model

- The atmospheric radiative forcing data used is based on CM2.1 CMIP5 historical forcings and was provided courtesy of GFDL

- This data was extended using RCP4.5 forcings for all the major radiative gases i.e. CO_2, CH_4, N_2O etc, and aerosols

- The only fields that change at different dates are volcanic sulphate aerosols, stratospheric O_3 and ocean CO_2 used to estimate ocean carbon

- For O_3 we used the spatially heterogeneous CMIP6 data which we tested this against the CMIP5 zonal mean O_3 over various assimilation periods finding better results i.e. lower RMS error.

- Volcanic emissions post 2000 were based on a “neutral” year

- Ocean CO_2 was based on data used in the ACCESS CMIP6

Observations

Table 1: Observation type and error estimates

Obs type	**Dataset**	**spatial distribution**	**error estimate**	R-Factor
Sea surface temperature SST	HadISST, OSISST, AVHRR, AMSR-E, AMSR-2, VIIRS, WindSat	point	0.5K, 0.5K, †, 0.5K, †, 0.25K, †	8
Sea level anomaly SLA	RADS	track	†	64
In situ ocean temperature TEM*	Argo, XBT, CTD, TAO, PIRATA	profiles	0.5K	8
In situ ocean salinity SAL*	Argo, CTD, TAO, PIRATA	profiles	0.075psu	8
Sea ice concentration SIC	HadISIC, OSISAF	gridded	0.1 [C], 0.1 [C]	8
Sea ice temperature SIT	HadISIT	gridded	0.1K	8
Zonal wind ARU	JRA55	gridded	$1ms^{-1}$	8
Meridional wind ARV	JRA55	gridded	$1ms^{-1}$	8
Air temperature ART	JRA55	gridded	1K	8
Specific humidity ARH	JRA55	gridded	0.05kg/kg	8

Table 2: *

* Low confidence CORA5.0 in situ data TEM2 & SAL2 have twice the error and four times the R-Factor to the high confidence data listed in the table.† error provided by vendor.

CORA5.0

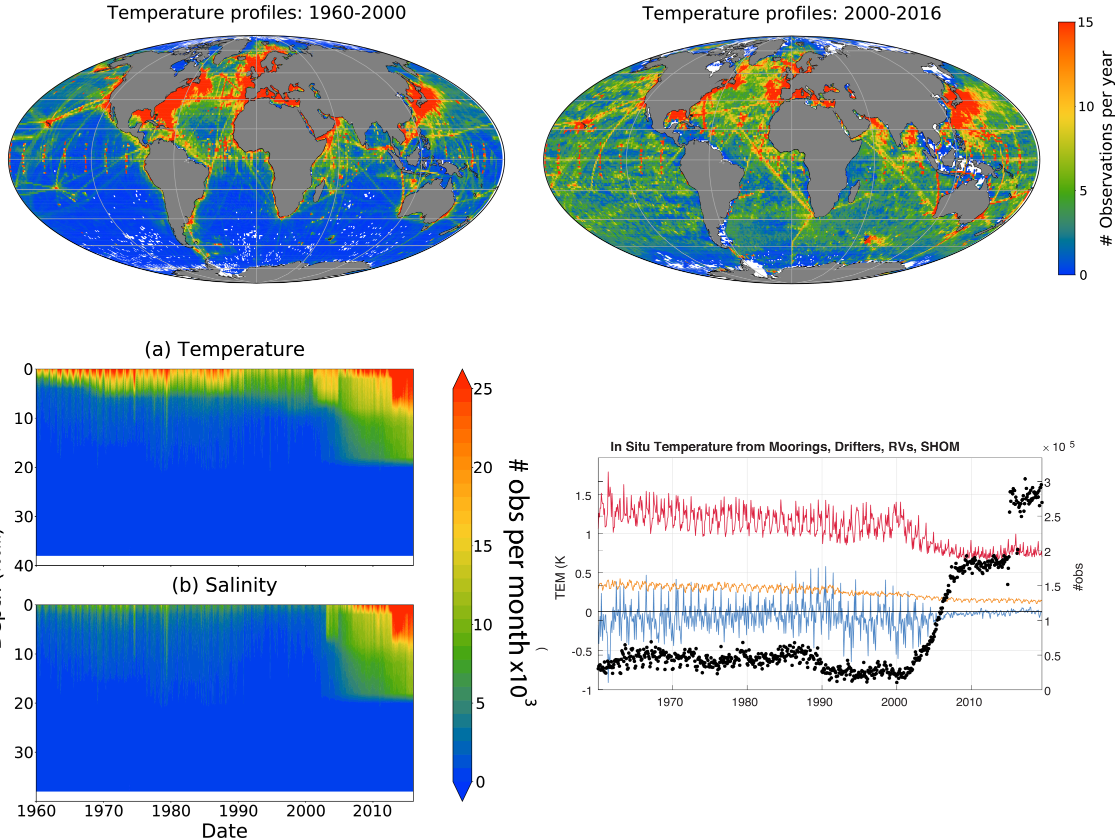


Figure 1:

EnKF scheme

The ETKF propagates first and second moments of $\mathbf{x}$ recursively and applies k-ensemble forecast and analysis anomalies $\mathbf{z}_i$ for $i = 1, 2, \ldots, k$ defined as

$$\mathbf{Z}^f = \frac{1}{\sqrt{k-1}}[\mathbf{z}_1^f, \mathbf{z}_2^f, \ldots, \mathbf{z}_k^f], \tag{1a}$$

$$\mathbf{Z}^a = \frac{1}{\sqrt{k-1}}[\mathbf{z}_1^a, \mathbf{z}_2^a, \ldots, \mathbf{z}_k^a] \tag{1b}$$

and where the state vectors are $\mathbf{z}_i^f = \mathbf{x}_i^f - \langle \mathbf{x}^f \rangle$ and $\mathbf{z}_i^a = \mathbf{x}_i^a - \langle \mathbf{x}^a \rangle$ which are n-dimensional in model space. The ETKF acts to choose appropriate initial forecast anomalies consistent with error covariance update equations within the vector subspace of ensemble anomalies formed as $\mathbf{P}^f = \mathbf{Z}^f \mathbf{Z}^{f^T}$ and $\mathbf{P}^a = \mathbf{Z}^a \mathbf{Z}^{a^T}$ and where the covariance update is given by

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f, \tag{2a}$$

$$\mathbf{K} = \mathbf{P}^f \mathbf{H}^T (\mathbf{H}\mathbf{P}^f \mathbf{H}^T + \mathbf{D})^{-1} \tag{2b}$$

where $\mathbf{K}$ is the $n \times p$ Kalman gain, $\mathbf{P}$ is the positive definite state covariance error matrix, $\mathbf{I}$ is the indentity matrix, and $\mathbf{H}$ is the $p \times n$ linearised observational operator mapping forecast grid point values onto observational points.

ETKF cont...

The Kalman gain $\mathbf{K}$ acts on the innovation $\left[\mathbf{d}_i - \mathbf{H}\langle\mathbf{x}^f\rangle\right]$ and specifically i.e.,

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}\left[\mathbf{d} - \mathbf{H}\mathbf{x}^f\right] \tag{3}$$

$$\mathbf{Z}^a = \mathbf{Z}^f\mathbf{T} \tag{4}$$

where the transform matrix $\mathbf{T}$ is now defined in terms of the $k \times (k-1)$ matrix of non-zero eigenvalues such that

$$\mathbf{T} = \mathbf{C}(\mathbf{\Gamma} + \mathbf{I})^{1/2}\mathbf{C}^T \tag{5}$$

where the $\mathbf{\Gamma}$ (non-zero eigenvalues) is $(k-1) \times (k-1)$ and $\mathbf{C}$ is $k \times (k-1)$, corresponds to the transform matrix in spherical simplex form.

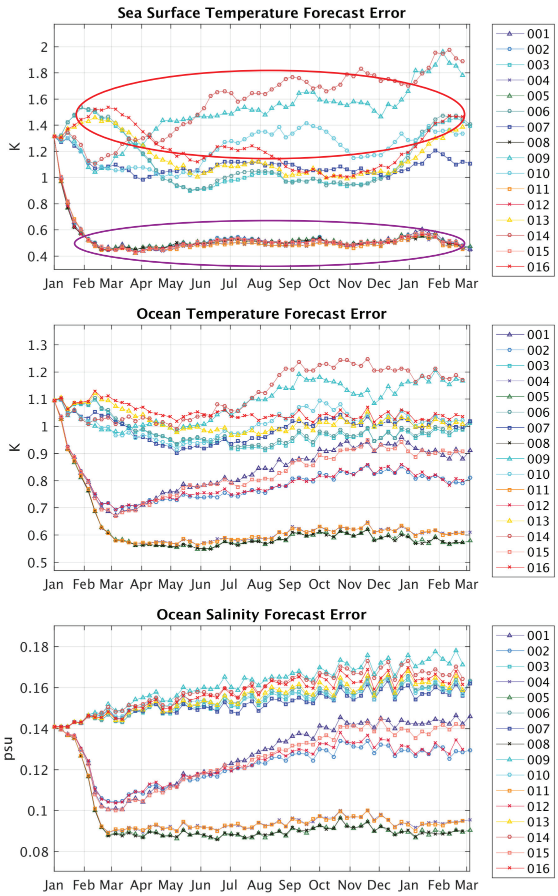
CAFE60 uses monthly mean observations and background (forecast) covariances to update the state estimates. For the atmosphere tis is radical as we seek to constrain only the large scale structures i.e.jet, Hadley and Ferrel cells etc.

Table 3: P.A. Sandery, T.J. O'Kane, V. Kitsios and P. Sakov (2020) State estimation of the climate system with the EnKF using variants of coupled data assimilation (EOR Mon. Wea. Rev.)

case	A-A	A-O	O-O	O-A	Atmospheric Increment	Ocean Increment	Type
1	1	1	1	1	atmospheric and ocean observations	atmospheric and ocean observations	strong
2	0	0	1	1		atmospheric and ocean observations	strong
3	1	1	0	0	atmospheric and ocean observations		strong
4	1	0	1	0	atmospheric observations	ocean observations	weak
5	0	0	1	0		ocean observations	weak
6	1	0	0	0	atmospheric observations		weak
7	0	1	0	1	ocean observations	atmospheric observations	strong
8	0	1	1	0	ocean observations	ocean observations	strong
9	1	0	0	1	atmospheric observations	atmospheric observations	strong
10	0	0	0	1		atmospheric observations	strong
11	1	1	1	0	atmospheric and ocean observations	ocean observations	strong
12	0	1	1	1	ocean observations	atmospheric and ocean observations	strong
13	0	1	0	0	ocean observations		strong
14	1	1	0	1	atmospheric and ocean observations	atmospheric observations	strong
15	1	0	1	1	atmospheric observations	atmospheric and ocean observations	strong
16	0	0	0	0			control

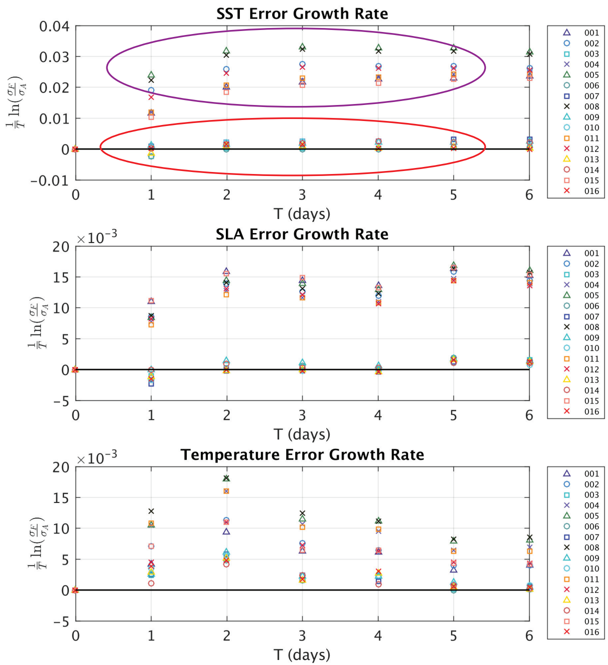
- A-A atmospheric covariances
- O-O ocean covariances
- A-O atmosphere-ocean cross-covariance: atmospheric increment due to ocean observations
- O-A ocean-atmosphere cross-covariance: ocean increment due to atmospheric observations

Ocean error growth: 7 day cycle



Global forecast innovation errors..

$$\sigma_G(T) = \frac{1}{T}\ln\left(\frac{\sigma_F(T)}{\sigma_A}\right)$$



Error growth rates for (a) sea surface temperature, (b) sea-level anomaly and (c) in-situ temperature.

Small innovation - error growth as skill decreases w.r.t. time

Large innovation - saturated errors

Table 4: Ocean domain

1st class	2nd class	3rd class	4th class
008: A-O, O-O	002: O-A, O-O	001: A-A, A-O, O-O, O-A	remaining variants saturated
005: O-O	012: A-O, O-O, O-A	011: A-A, A-O, O-O	
		015: A-A, O-O, O-A	

Table 5: Atmosphere domain

1st class	2nd class
001: A-A, A-O, O-O, O-A	remaining variants saturated
008: A-O, O-O	
011: A-A, A-O, O-O	
003: A-A, A-O	
009: A-A, O-A	

In the atmosphere domain and for a 7-day or longer cycle, an atmospheric increment based on ocean observations performs almost as well as directly assimilating atmospheric observations.

- A-A atmospheric covariances
- O-O ocean covariances
- A-O atmosphere-ocean cross-covariance: atmospheric increment due to ocean observations
- O-A ocean-atmosphere cross-covariance: ocean increment due to atmospheric observations

In addition CAFE60 has sea ice is strongly coupled to the ocean. OBGC is weakly coupled to the ocean via the cross domain covariance dependent on ocean observations. No assimilation of land based observations.

Atmosphere

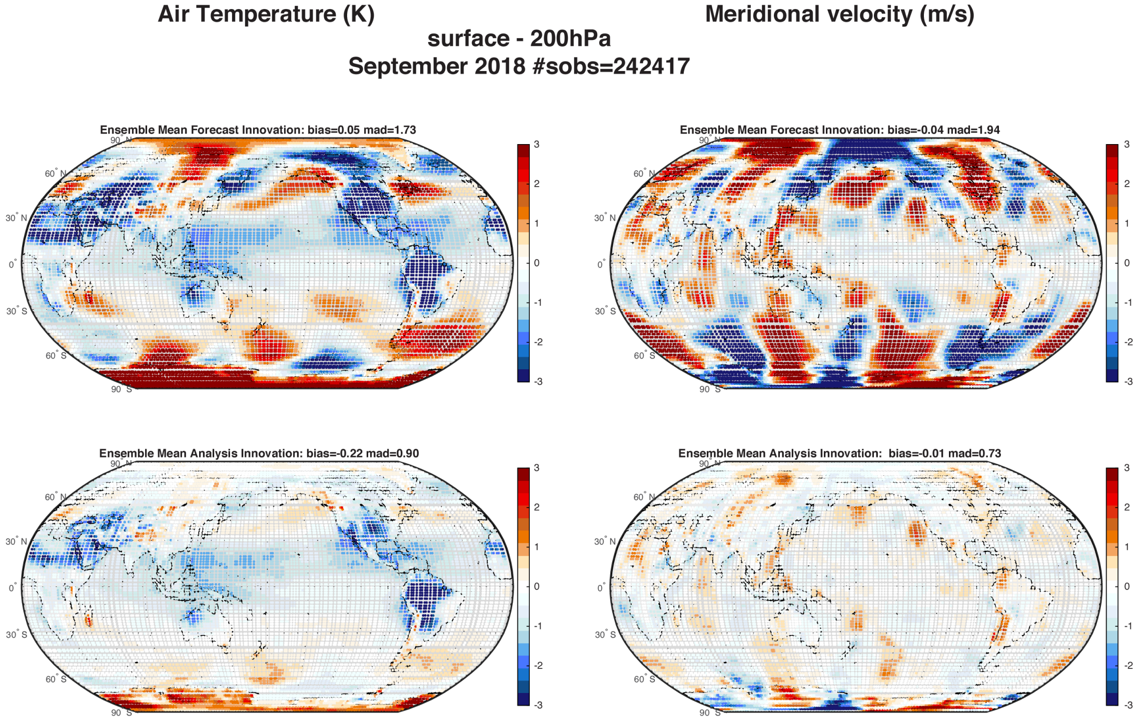


Figure 2:

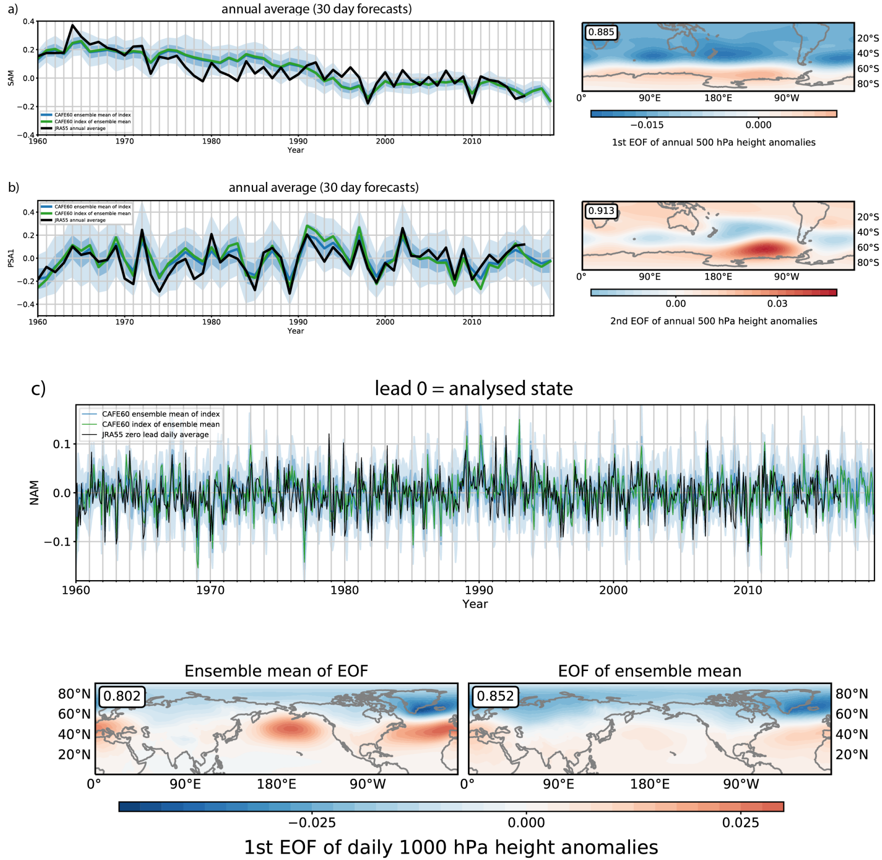


Figure 3: (a) Annual means of the CAFE60 SAM index calculated as the leading PC of daily 500hPa geopotential height data and the corresponding EOF1 spatial pattern and its correlation with JRA55. (b) Similar calculations for the annual mean PSA1 mode and (c) the Northern Annular mode / Arctic Oscillation EOF-based index for lead-zero daily PCs and corresponding EOF patterns. Here, both the ensemble mean of the EOF patterns and the EOF pattern of the ensemble mean are shown to illustrate model bias in the Pacific Ocean evident in the ensemble mean of member EOFs which are not evident whenconsidering the EOF of the ensemble mean.

IPO-TPI

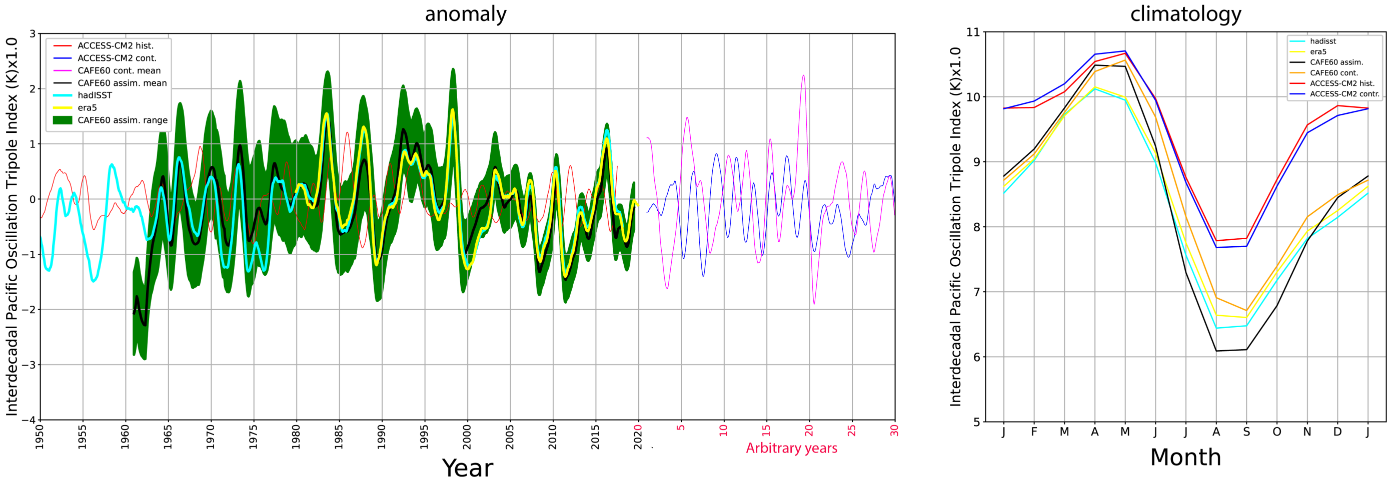


Figure 4: Interdecadal Pacific Oscillation tripole index.

MJO

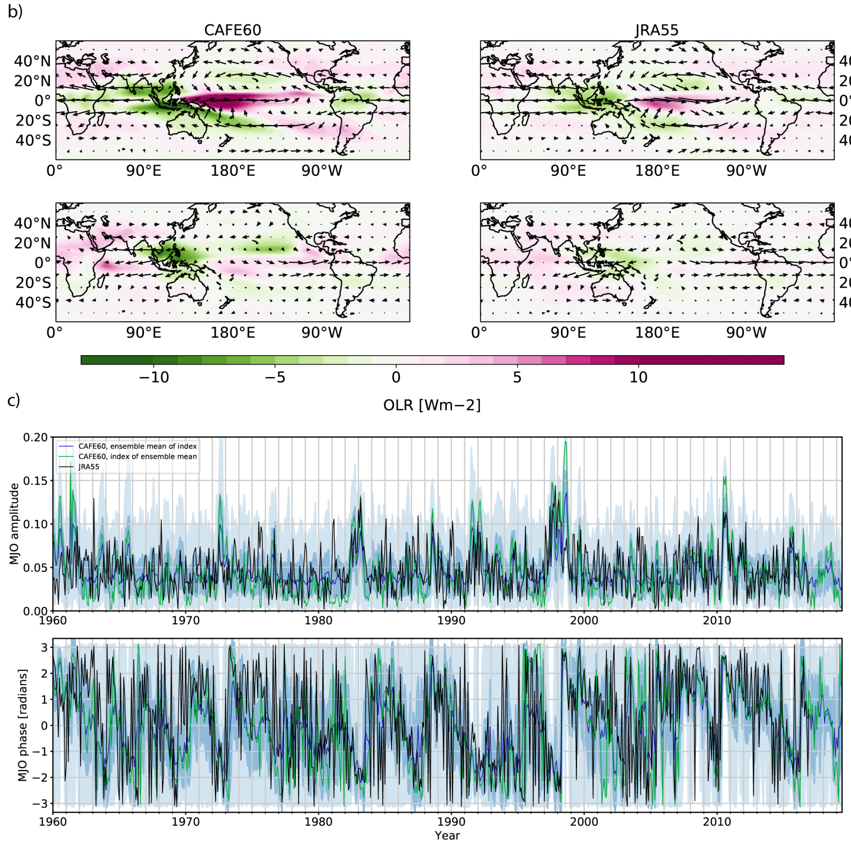


Figure 5: (b) The two leading principal components (PCs) of the daily 850 minus 150-hPa global velocity potential and their regression onto OLR and the ensemble mean wind direction pattern resulting from the regression of PCs onto u and v wind anomalies at 150 hPa. (c) Time series of amplitude and phase of the MJO in terms of daily 850 minus 150-hPa global velocity potential. Light blue shading indicates the CAFE60 ensemble spread, the darker blue shading indicates one standard deviation from the mean.

SAT and Precipitation

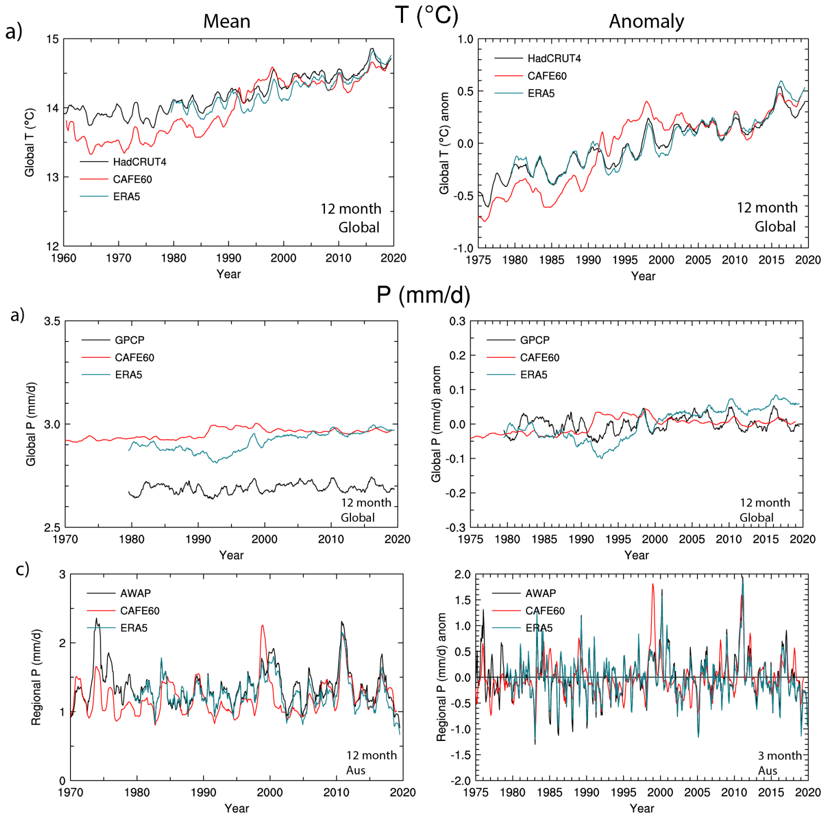


Figure 6: Comparison of CAFE60 ensemble mean precipitation and surface air temperature to HadCRUT4, ERA5, GPCP and AWAP.

Ertel PV

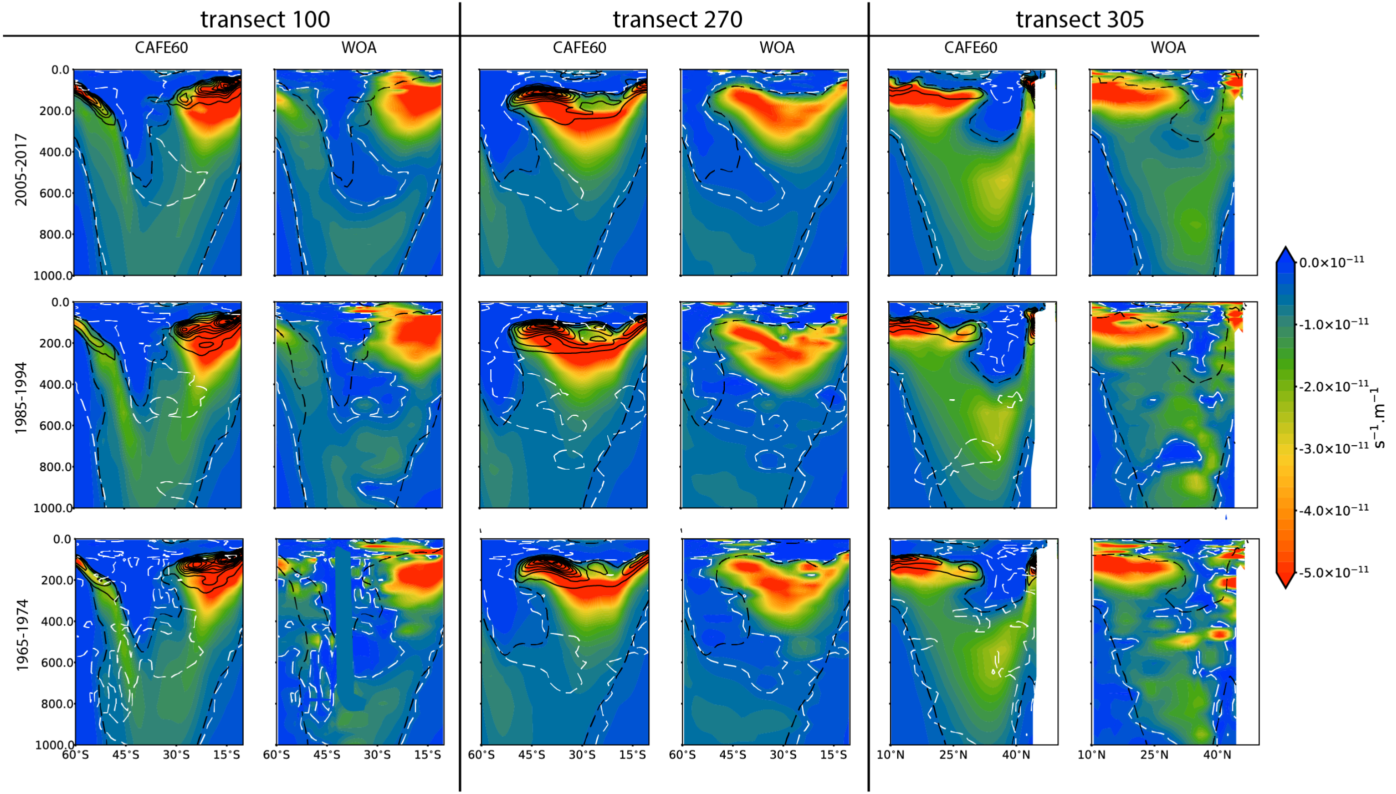


Figure 7:

Global MOC

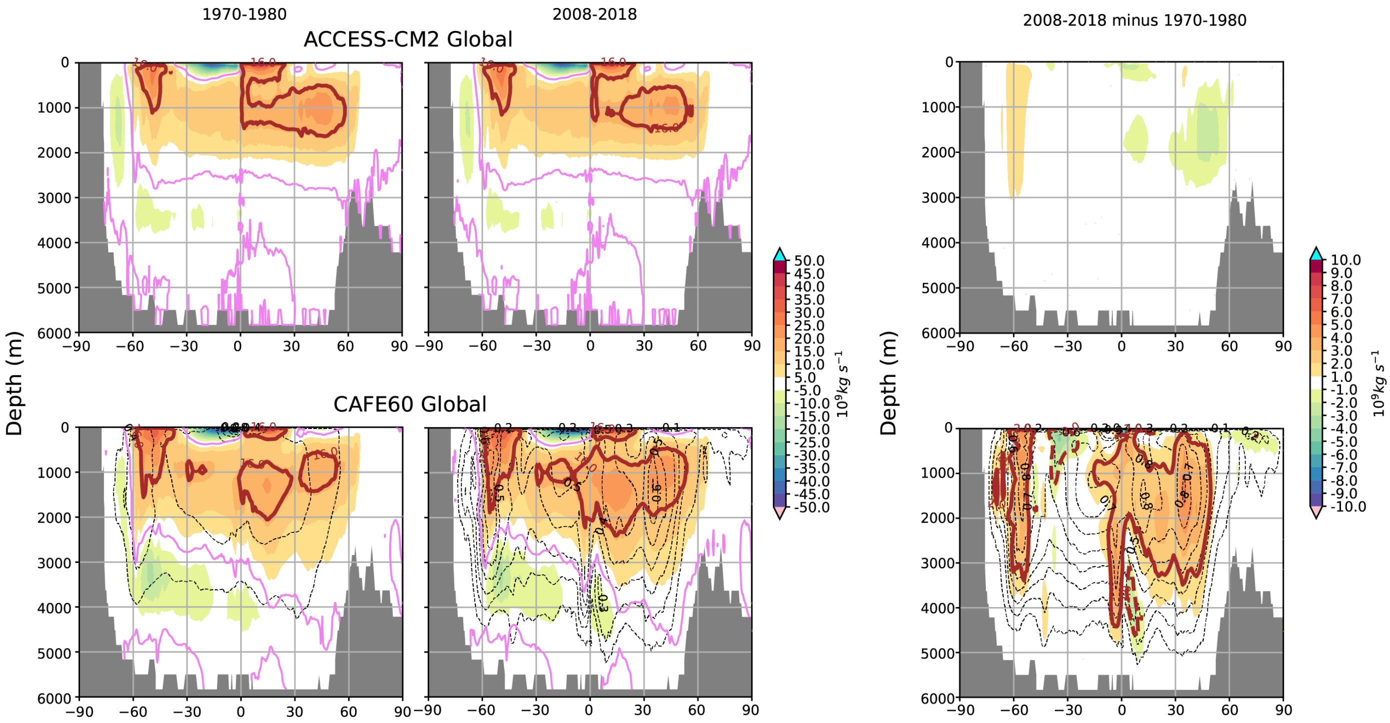


Figure 8: ACCESS-CM2 and CAFE60 meridional volume transport in depth coordinates defined in terms of the mass transport (kgs^{-1}) as a function of horizontal location and depth. Here the MOC are defined globally. Red contours indicate 16Sv and 2 Sv respectively. Black contours indicate stddvn.

Transports

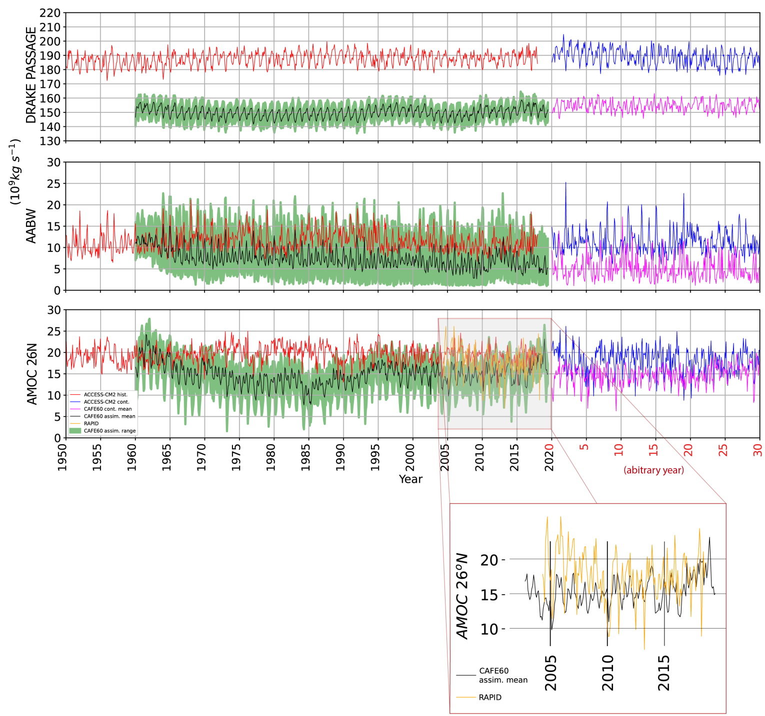


Figure 9: To obtain strength of overturning for a particular common index to compare to generally very limited observations, we simply identify the maximum value within a particular depth and latitude box, the range of these boxes are: $AMOC26^\circ N{-}25.5^\circ - 26.5^\circ, 0{-}6000m$ $AABW90^\circ S{-}60^\circ S, 0{-}3000m$ Note that $AMOC26N$ is identified at the two longitudes 25.5 and 26.5 that span the reqired $26^\circ N$.

Sea Ice Extent

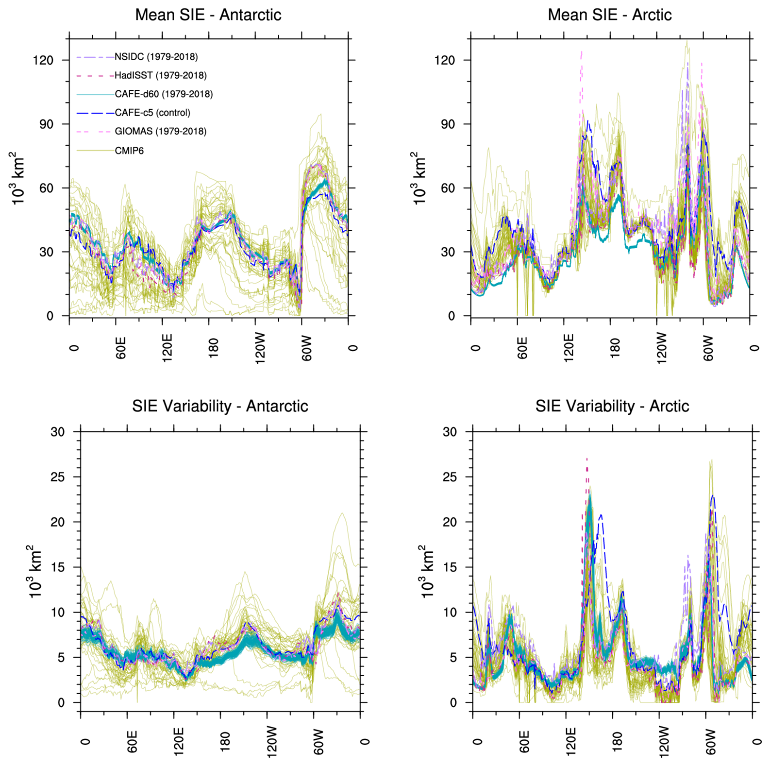


Figure 10: Arctic sea ice in JUL-SEP unrealistically low. Variability is good but trends are poor. Sea Ice is a huge challenge for long cycle length!

Summary

Challenges:

- Sea Ice
- blocking over Euro-Atlantic sector
- pre-Argo ocean

Future plans (just beginning):

- Australian Leadership Computing Grant forecasts.
- Improve sea ice!!
- Assimilation of land observations (leaf area index, soil moisture content), OBGC (Chlorophyl A).
- GFDL CM4.

CAFE60 offers a novel way to understand climate variability and predictability over the last 6 decades.

Thank You

CSIRO Oceans & Atmosphere

Terence O'Kane
t +61 3 6232 5066
e terence.okane@csiro.au
w http://people.csiro.au/O/T/Terence-OKane