

SUMMARY

The limits for the predictability of the ENSO have been long discussed. Even when ENSO prediction skill is expected to be limited, questions remain as to which are the limiting factors. The role of atmospheric noise for ENSO initiation, the growth of initial errors and inadequate models have been identified as key factors. Past studies have used retrospective forecasts from seasonal prediction systems to evaluate the predictability of ENSO for up to 24 months. The present study makes use of the recently released CMIP5 decadal hindcasts to explore ENSO predictability beyond that threshold. A set of retrospective forecasts from 9 different modeling systems that were initialized every year starting in 1961 and run for 120 months were considered to explore long-lead ENSO predictability and suggest that some skill exists for leads longer than 24 months. In addition, the performance of the multi-model ensemble mean is explored and compared to the multi-model mean based solely on the most skillful systems; the latter is found to yield better results for the deterministic metrics. Finally, an analysis of the near-surface temperature and precipitation teleconnections reveals that the ability of the systems to detect ENSO events far in advance could translate into predictive skill over land for several lead years, though with reduced amplitudes compared to observations.

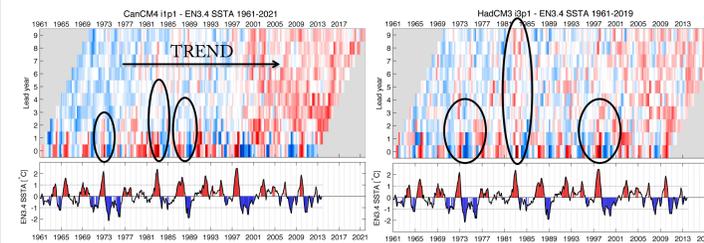
DATA

The activity of El Niño – Southern Oscillation (ENSO) is described using the sea surface temperature (SST) averaged over the EN3.4 region.

Decadal hindcasts are used for the following 9 modeling systems, which were initialized every year for the period indicated in the table. EN3.4 from those simulations was compared with the observed EN3.4 calculated from the ERSST version 3b dataset.

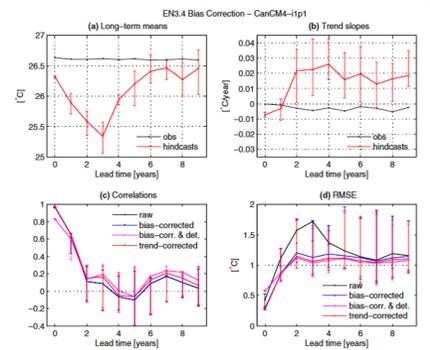
Model	Center	Members	Initialization	Starts	Atmospheric resolution
BCC-CSM1.1	BCC, China	3	Full field	1961:1:2007	2.81° lat X 2.81° lon
CanCM4	CCCMA, Canada	10	Full field	1961:1:2012	2.8° lat X 2.8° lon
EC-Earth i1	Consortium, Europe	5	Full field	1961:1:2006	1.125° lat X 1.125° lon
EC-Earth i3	Consortium, Europe	8	Anomaly	1961:1:2006	1.125° lat X 1.125° lon
GFDL-CM2.1	GFDL, USA	10	Full field	1961:1:2012	2° lat X 2.5° lon
HadCM3 i2	Hadley Center, UK	10	Anomaly	1961:1:2010	2.75° lat X 3.75° lon
HadCM3 i3	Hadley Center, UK	10	Full field	1961:1:2010	2.75° lat X 3.75° lon
MIROC5	MIROC, Japan	6	Anomaly	1961:1:2011	1° lat X 1° lon
MPI-ESM-LR	MPI-M, Germany	3	Anomaly	1961:1:2011	1.9° lat X 1.9° lon

MOTIVATIONS AND PREPROCESSING



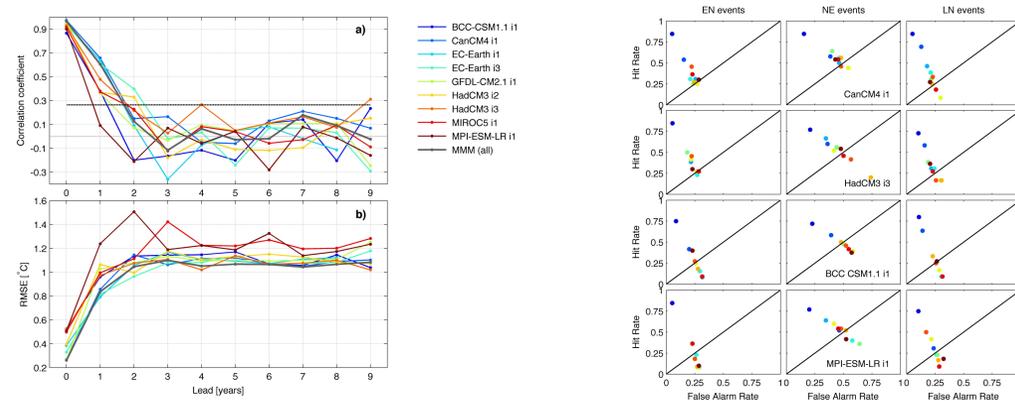
When exploring EN3.4 SST monthly anomalies in some of the CMIP5 decadal hindcasts we noticed that some of the modeling systems seemed to be able to simulate the strong ENSO events (both La Niña and El Niño) for long leads.

Some of the hindcasts exhibit biases, and we therefore worked with **cross-validated** and **bias-corrected anomalies**. Additionally, the anomalies were **trend-corrected** by fitting linear trend coefficients for each lead time.



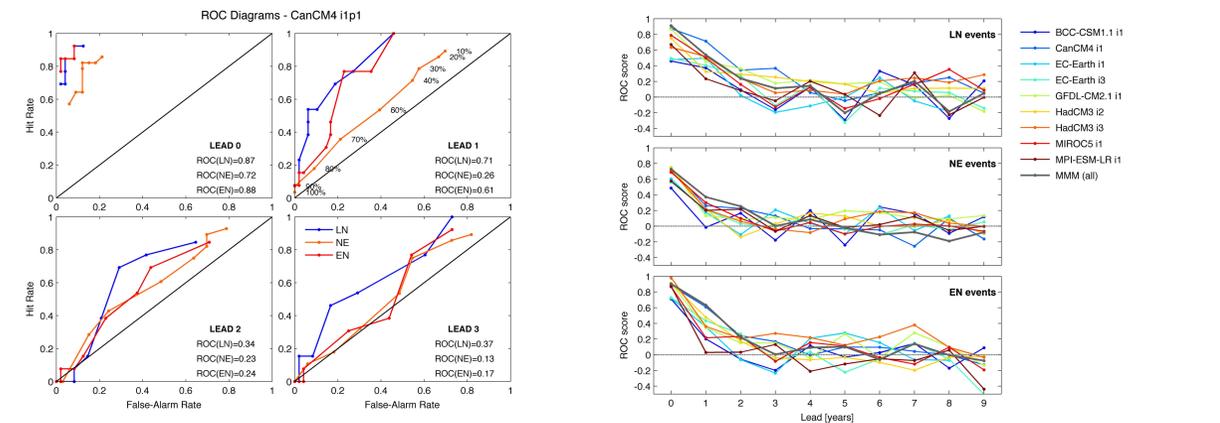
DETERMINISTIC SKILL

The deterministic skill of the systems to represent the EN3.4 SST variability, as well as their ability to capture the observed El Niño (EN), La Niña (LN) and Neutral (NE) events was assessed. Most models have significant ability for lead years 0 and 1. Some of them, such as the HadCM3 and EC-Earth i3 systems, are also skillful at lead year 2 (24–26 months). These results are comparable to those found for seasonal prediction systems for up to 12 months.



PROBABILISTIC SKILL

The probabilistic assessments of the forecast systems use the ROC diagrams and ROC scores, calculated for the EN, LN and NE events. Skill is evident in single models up to 2 years out, and for the multi-model mean for up to 3-4 years ahead. The probabilistic skill for event detection is larger for LN events than for EN events. NE years are the hardest to predict in every case.

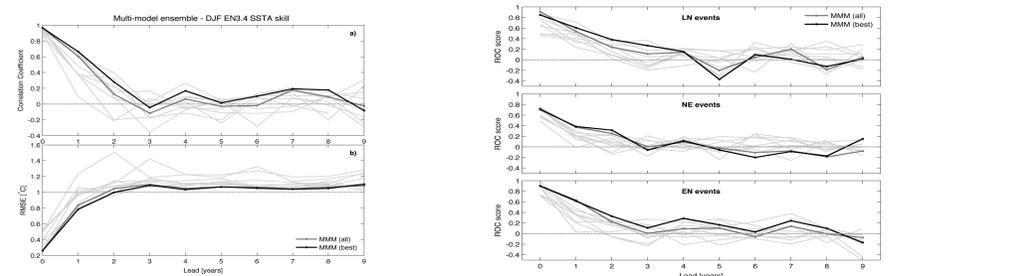


BEST MODELS ENSEMBLE MEAN

The decadal prediction systems are ranked based on their representation of the seasonal cycle of EN3.4 SST, as well as their skill metrics. The 5 highest-ranking models are considered in the “best” models ensemble.

This new “best” ensemble mean shows improved deterministic skill, compared to the full ensemble, for lead years up to 3.

MODEL	BCC-CSM1.1	CanCM4	EC-Earth i1	EC-Earth i3	GFDL-CM2.1	HadCM3 i2	HadCM3 i3	MIROC5	MPI-ESM-LR
Properties of seasonal cycle									
lowest RMSE of cycle				2nd		1st		3rd	
lowest RMSE for amplitude			3rd		3rd		1st		2nd
lowest RMSE of var									
SEASONAL CYCLE POINTS	0	1	0	3	0	6	2	4	2
Deterministic EN3.4 index skill									
highest ACC		1st		2nd				3rd	
lowest RMSE		2nd		3rd		1st			
highest HR		1st		2nd				3rd	
lowest FAR		1st		2nd				3rd	
Deterministic ENSO skill									
highest HR		2nd		3rd			1st		
lowest FAR		2nd		3rd			1st		
highest HR		1st		3rd		2nd			
lowest FAR		1st		3rd		2nd			
highest ROC score		1st		2nd			2nd		3rd
lowest ROC score		1st		2nd			2nd		3rd
Probabilistic ENSO skill									
highest ROC score		1st		2nd			3rd		
lowest ROC score		1st		2nd			3rd		
ENSO SKILL POINTS	0	30	3	13	9	7	12	7	2
TOTAL POINTS	0	31	3	13	9	7	12	7	2



TEMPERATURE AND PRECIPITATION COMPOSITES

The EN and LN teleconnections are computed for near-surface temperature and precipitation, to assess if the ability of the systems to capture the ENSO variability translates into potential predictive skill over land (shown here for EN events). These composites show some potential skill, though the amplitudes become smaller than the observed ones at longer lead-times. Additionally, the “best” model ensemble has better skill than the best single model (CanCM4) in some cases.

