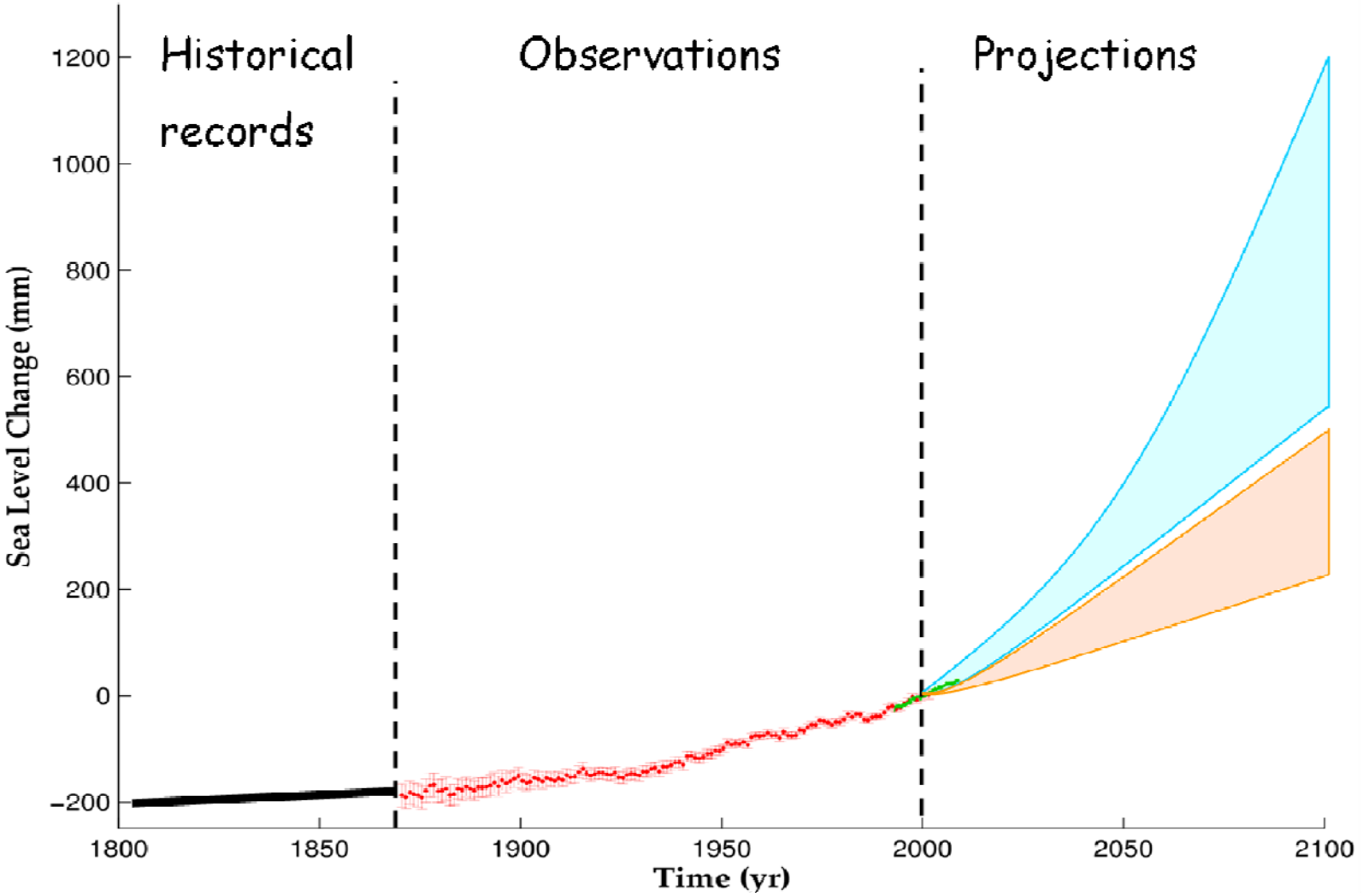
Reliability estimates of decadal sea level trend hindcasts

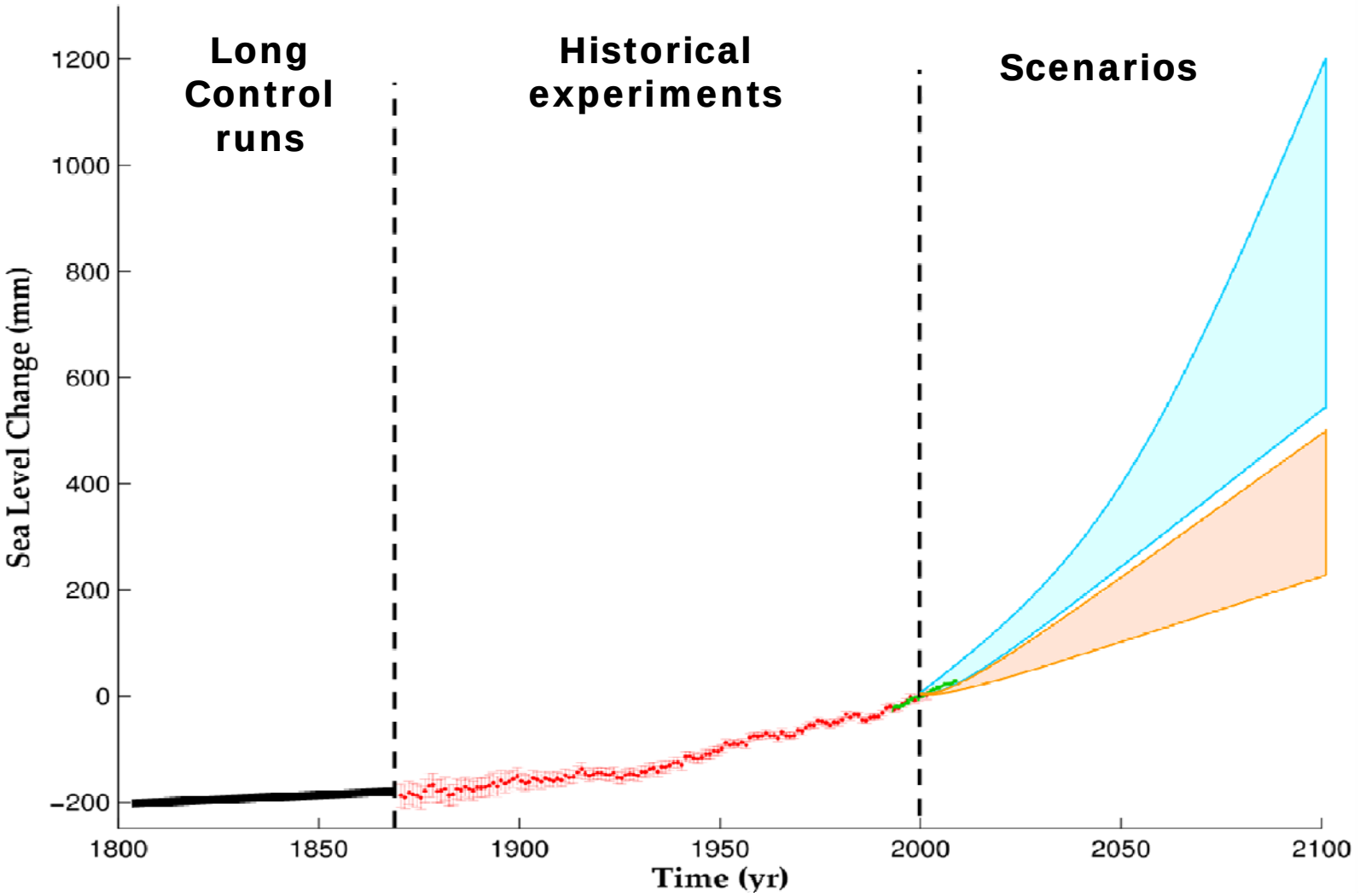
Philippe Rogel (CERFACS)
David Salas Y Melia (CNRM Météo-France)
Emilia Sanchez, Elodie Fernandez, Christophe Cassou (CERFACS)
Mélanie Becker, Benoit Meyssignac (LEGOS)

- Motivations
- Probabilistic projections
- Reliability estimates and recalibrated probabilities
- Conclusions & perspectives


Historical records
Observations
Projections
Sea Level Change (mm)
1200
1000
800
600
400
200
0
−200
1800
1850
1900
1950
2000
2050
2100
Time (yr)

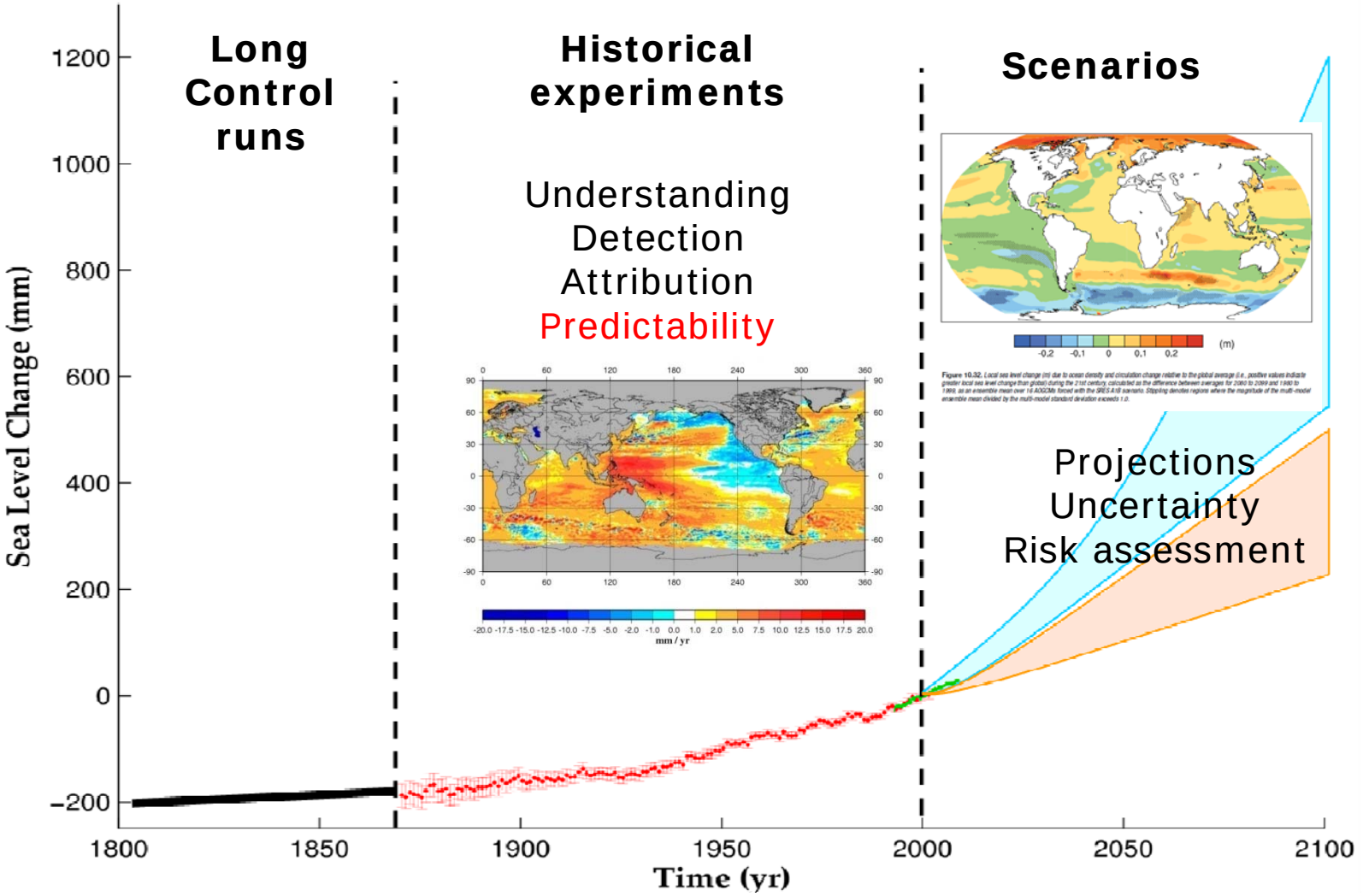
?

In the world of models



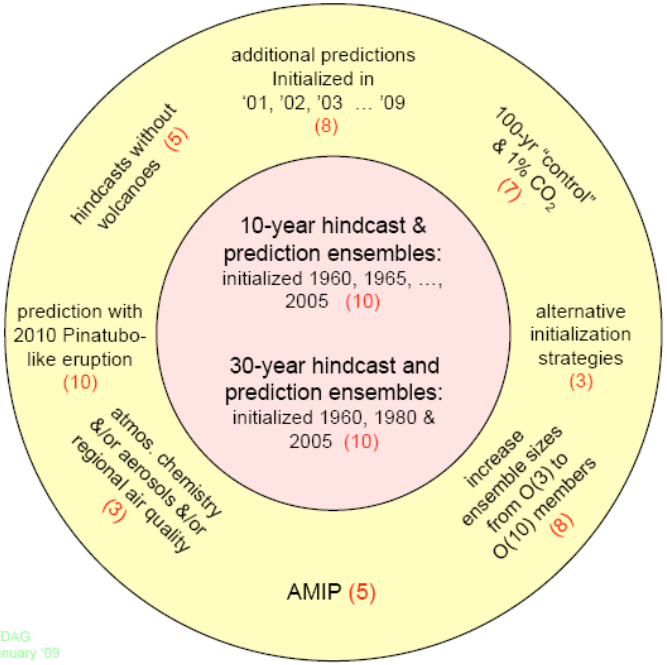
?

In the world of models



The CMIP5 Experiments

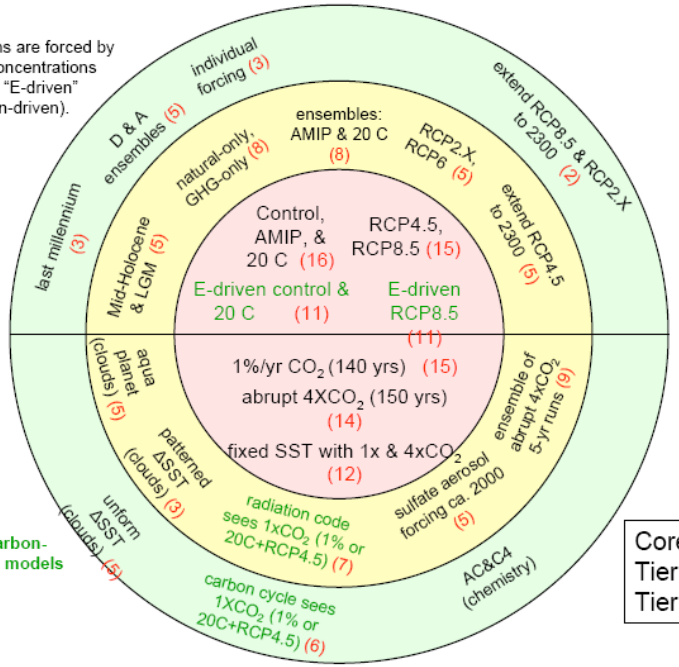
CMIP5 Decadal Prediction Experiments



IDAG
23 January '09

CMIP5 Long-term Experiments

All simulations are forced by prescribed concentrations except those "E-driven" (i.e., emission-driven).



Coupled carbon-cycle climate models only

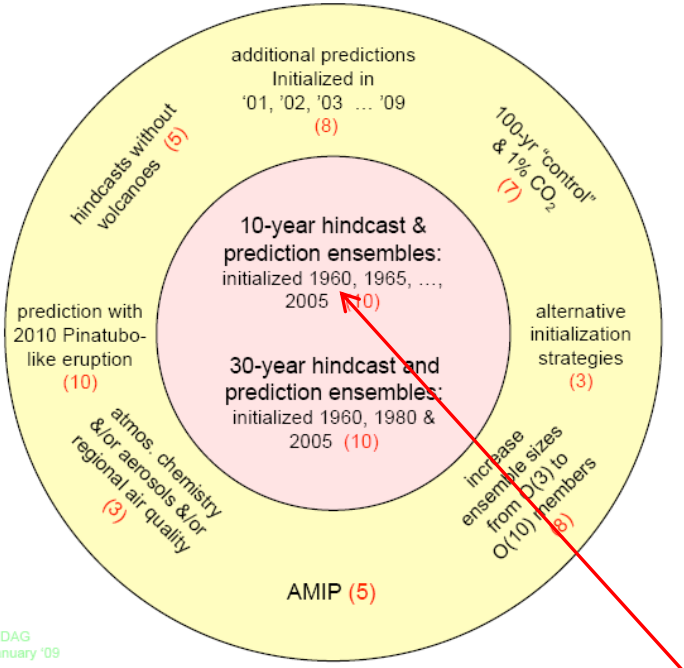
Core:	≥1718 yrs
Tier 1:	≥1727 yrs
Tier 2:	≥2038 yrs

IDAG
23 January '09

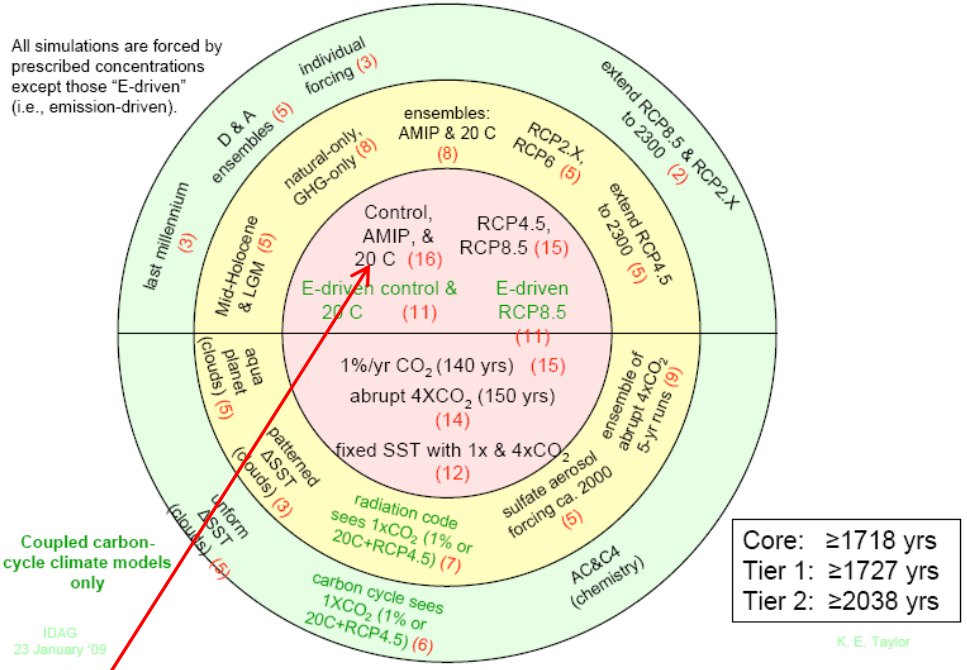
K. E. Taylor

The CMIP5 Experiments

CMIP5 Decadal Prediction Experiments



CMIP5 Long-term Experiments



Core: ≥1718 yrs
Tier 1: ≥1727 yrs
Tier 2: ≥2038 yrs

K. E. Taylor

Experiments considered here

Experiments considered here:

•NEMOVAR Reanalysis of Temperature and Salinity profiles (1960-present) produced by ECMWF (Balmaseda, 2011)

•Decadal hindcast ensembles performed with the CNRM-CM5 (ARPEGE-NEMO) coupled model developed for CMIP5 (« Decadal Hindcasts »):

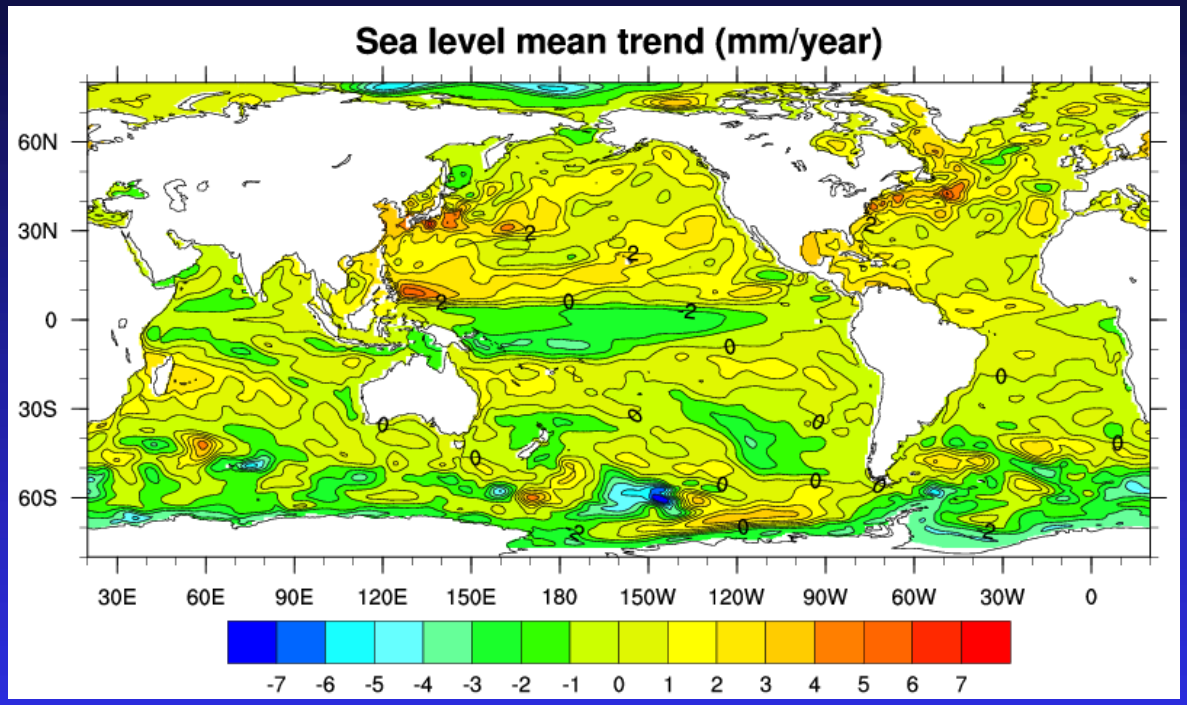
- One 10-year-ahead hindcast ensemble initialised in 1960, 1965…. 2005, from NEMOVAR Reanalysis
- 10-member ensembles obtained through atmospheric perturbations

• 10 all-forcing (natural + anthropogenic) 20th Century « Historical » simulations

• Exactly the same ocean model for all experiments (including Reanalysis)

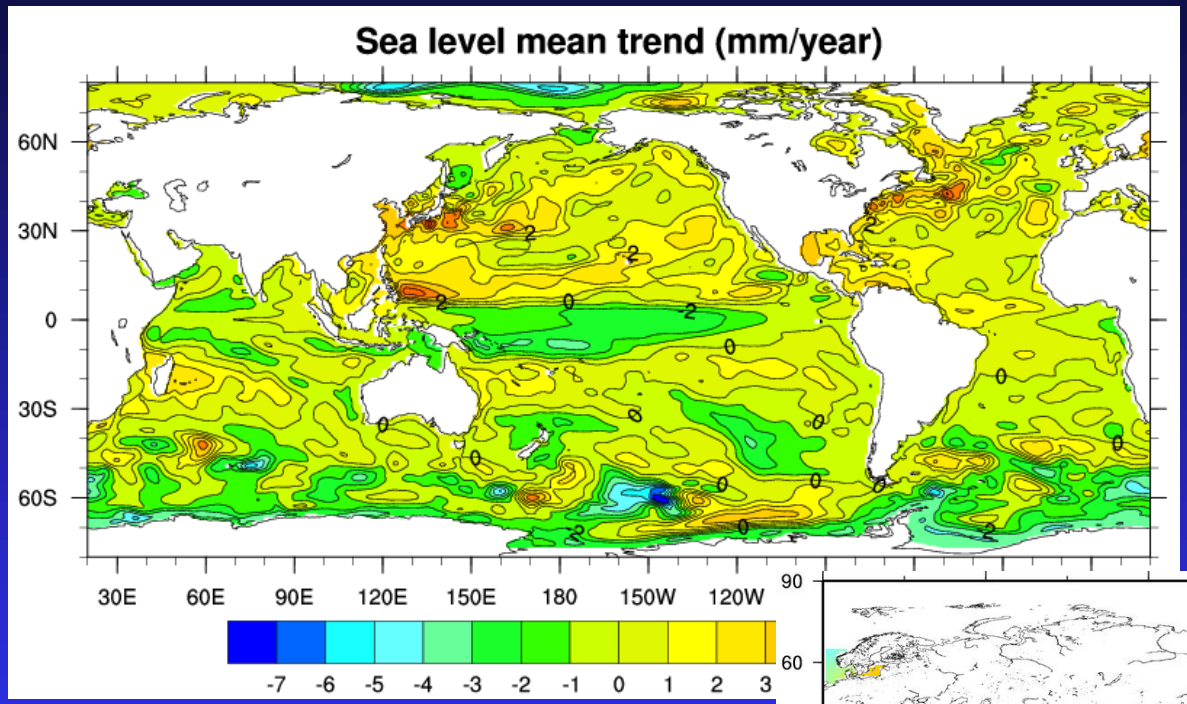
• Sea level trend maps extracted for leadtimes 0-2 years to 0-10 years

Mean Sea level regional trends 1960-2005



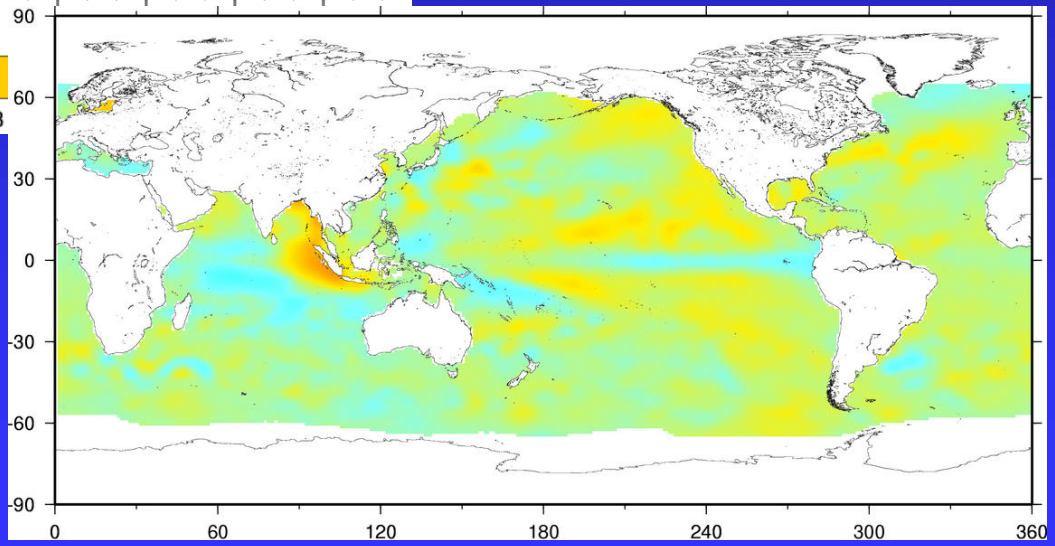
← Reanalysis

Mean Sea level regional trends 1960-2005

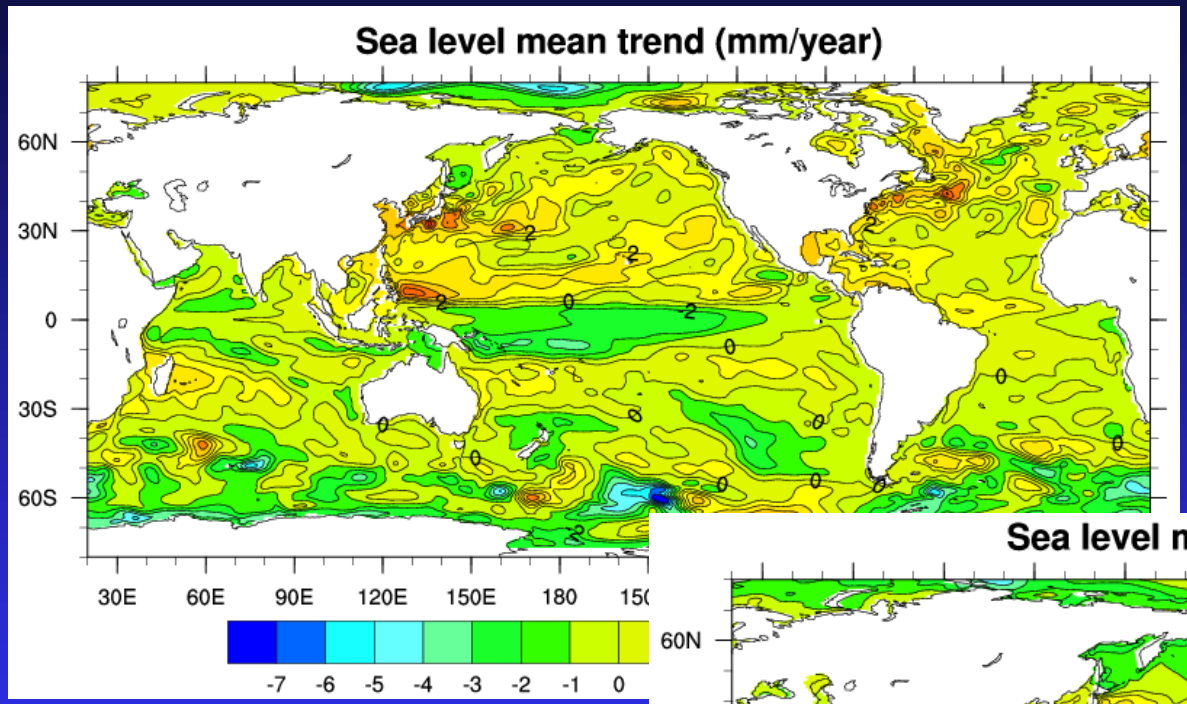


← Reanalysis

Church 2004 →

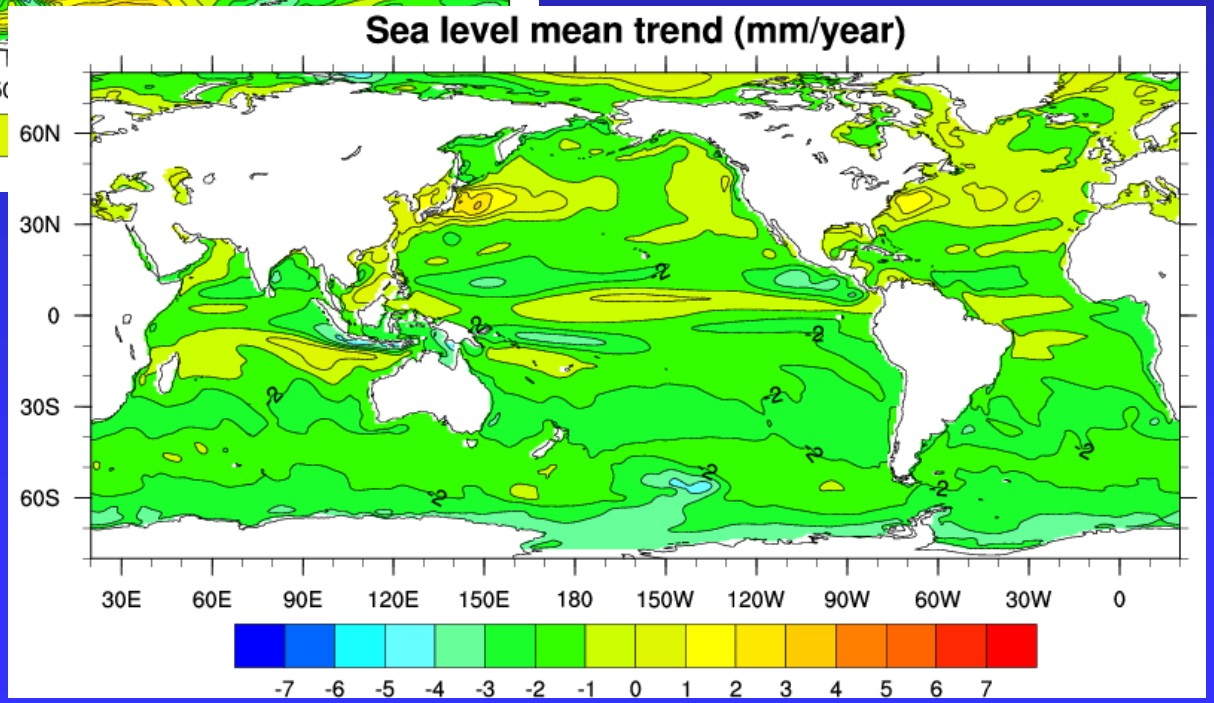


Mean Sea level regional trends 1960-2005

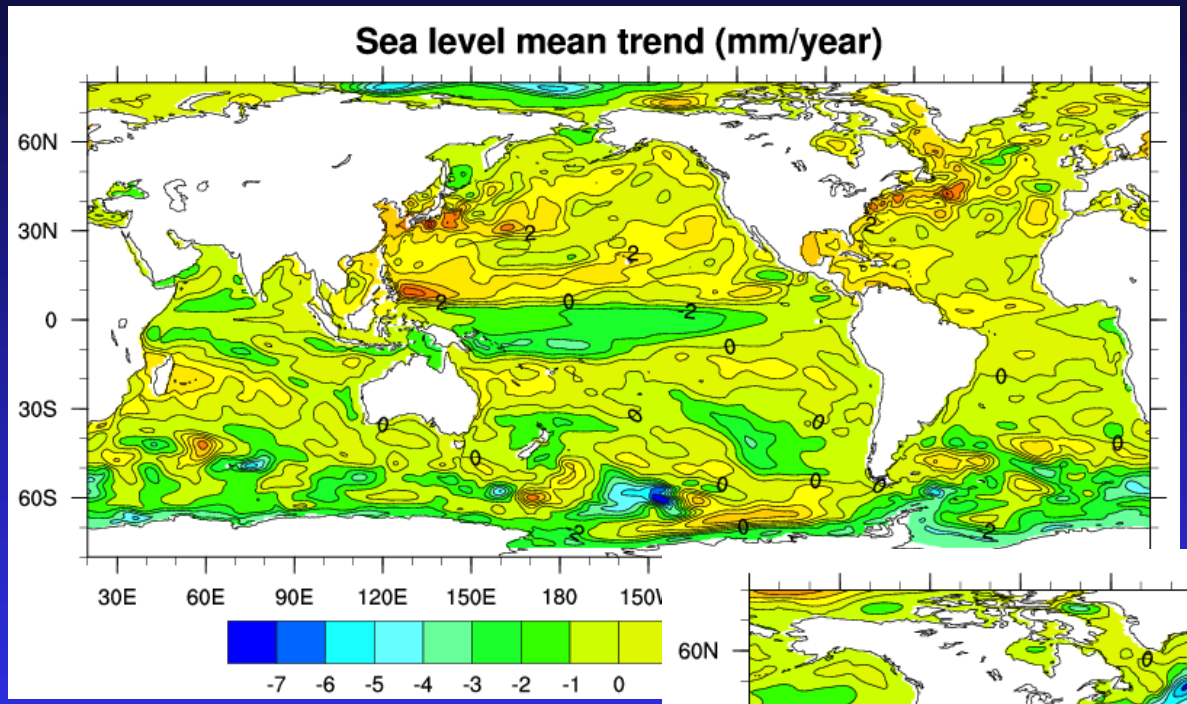


← Reanalysis

Historicals →

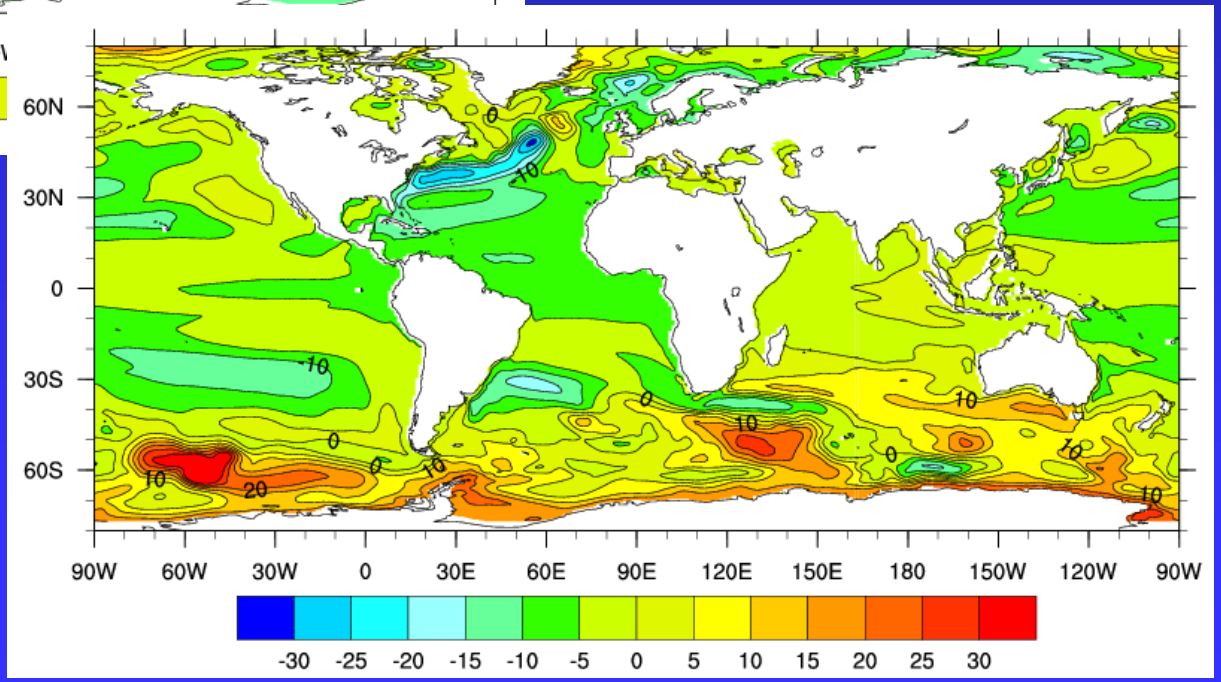


Mean Sea level regional trends 1960-2005

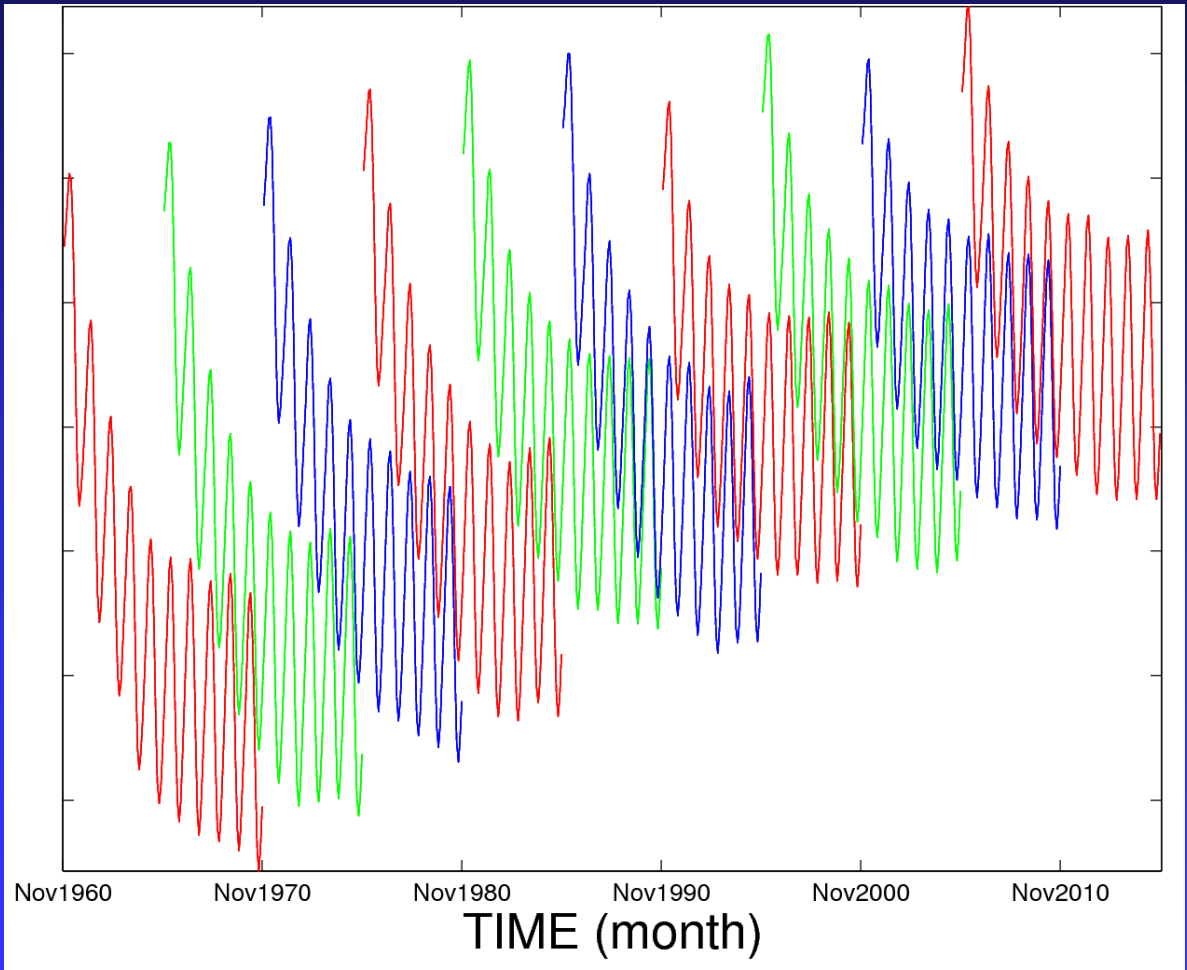


← Reanalysis

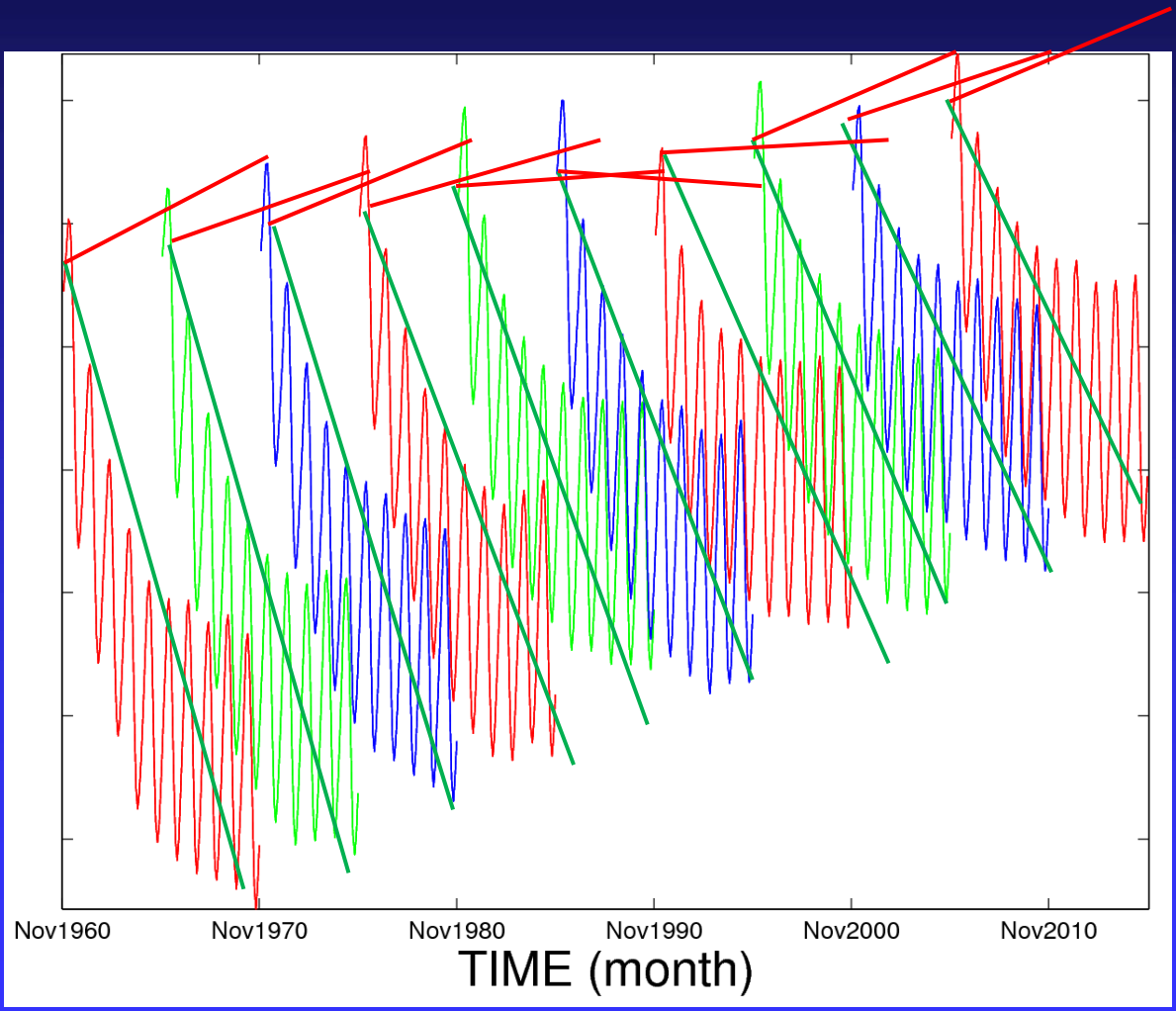
Decadal Hindcasts →



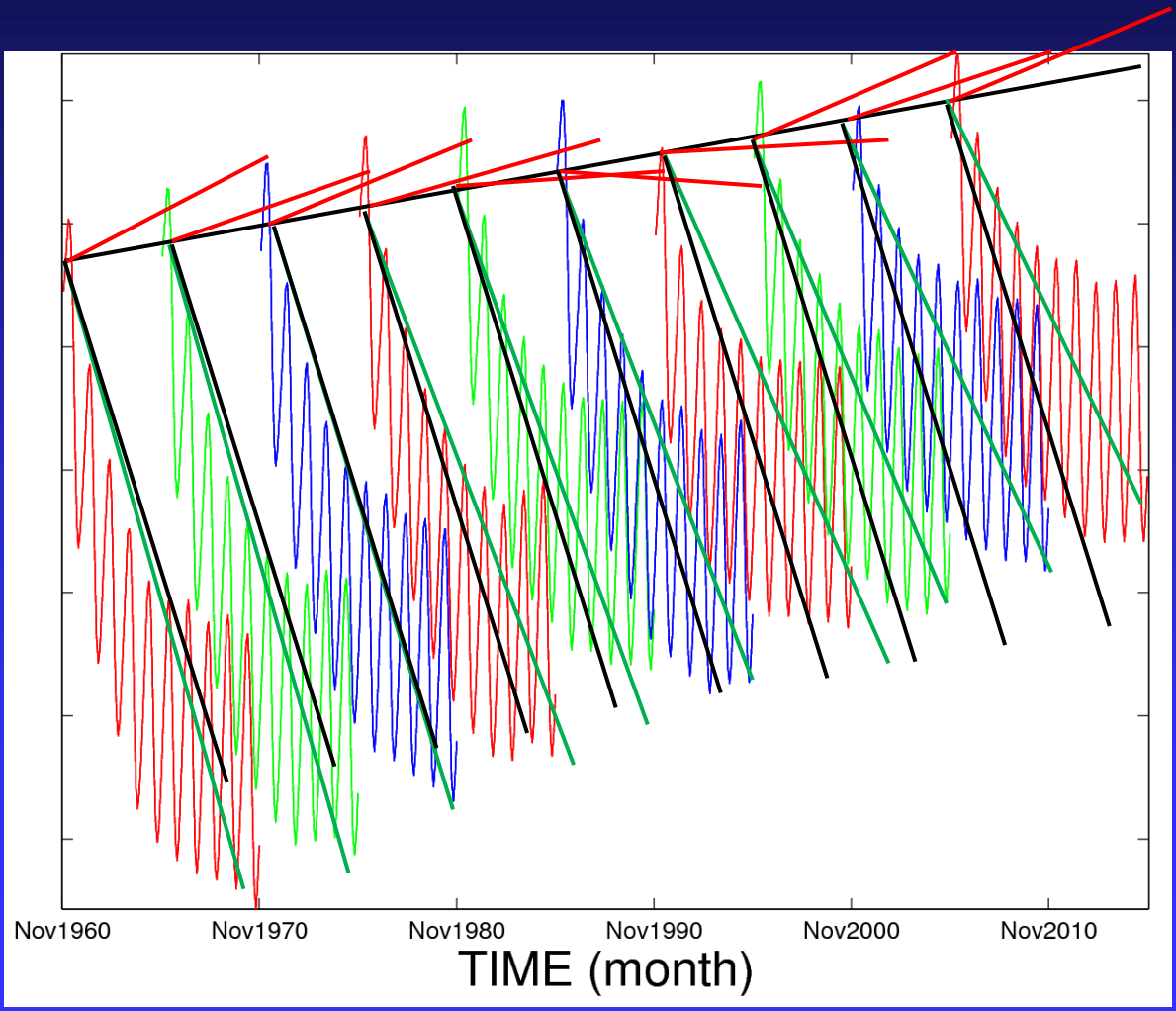
Hindcasts of Sea level trends



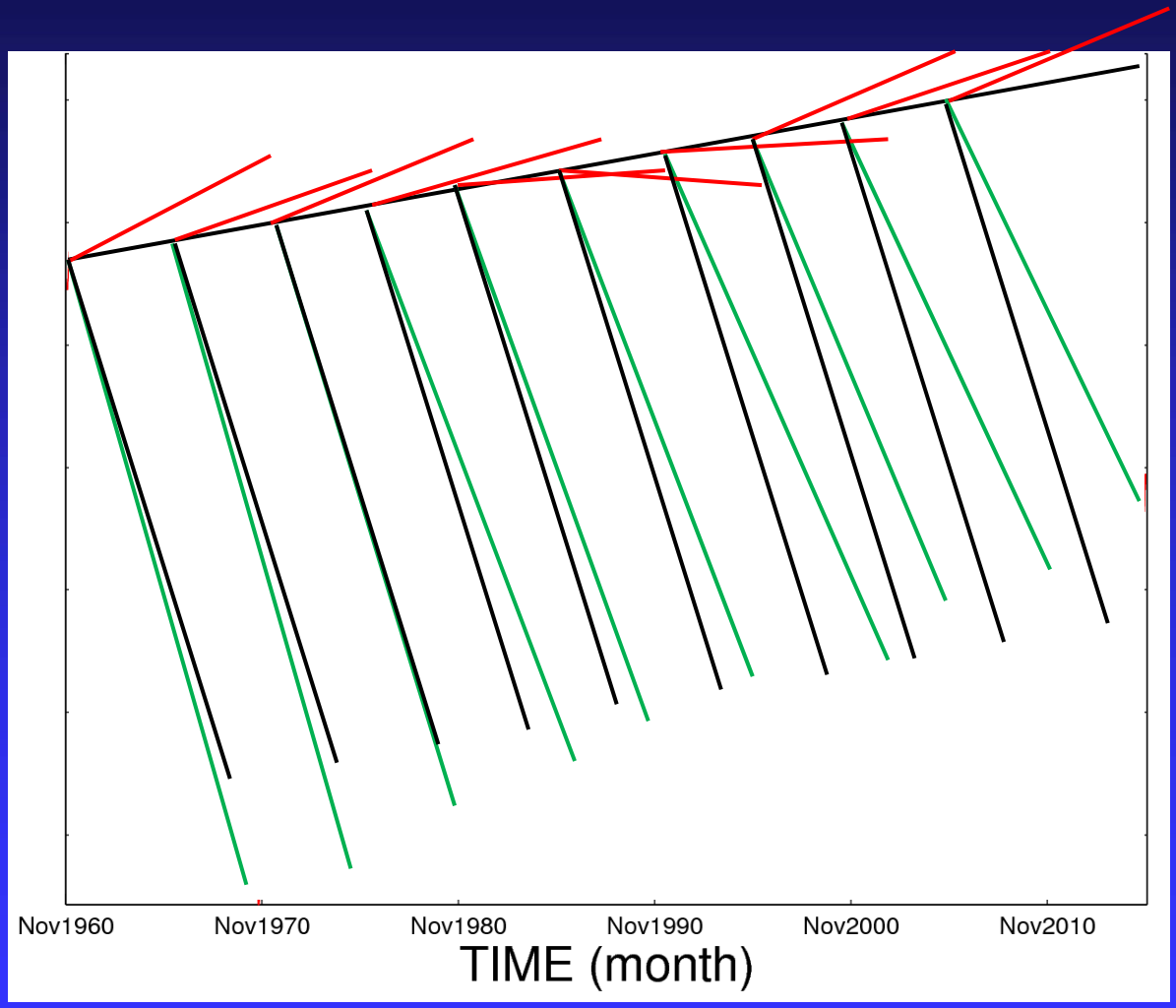
Hindcasts of Sea level trends



Hindcasts of Sea level trends



Hindcasts of Sea level trends



Probabilistic prediction of sea level trends

In the following, we want to assess the predictability of the event E:

« The regional sea level trend will be higher than the long term average global mean sea level trend » (i.e. 1.8 mm/year in observations)

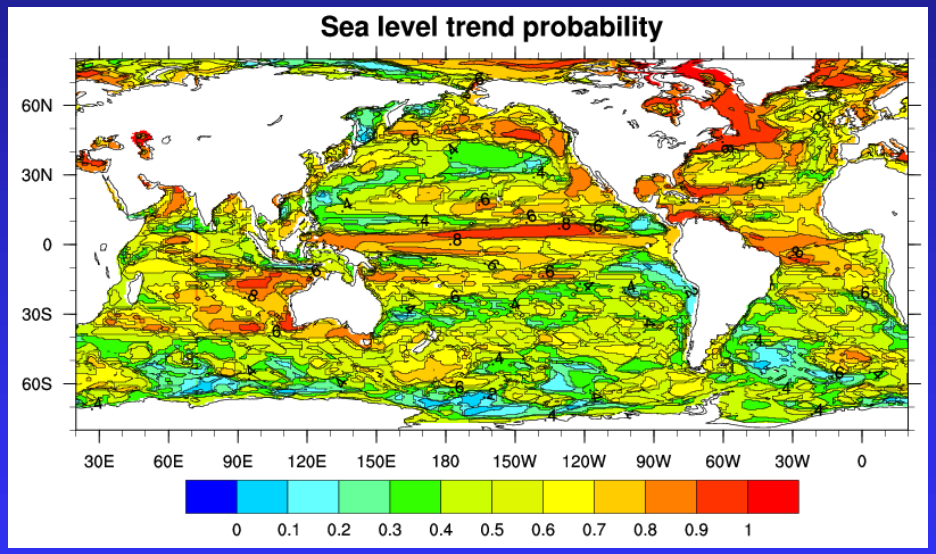
The forecasted probability that E occurs is

$$p = j/P$$

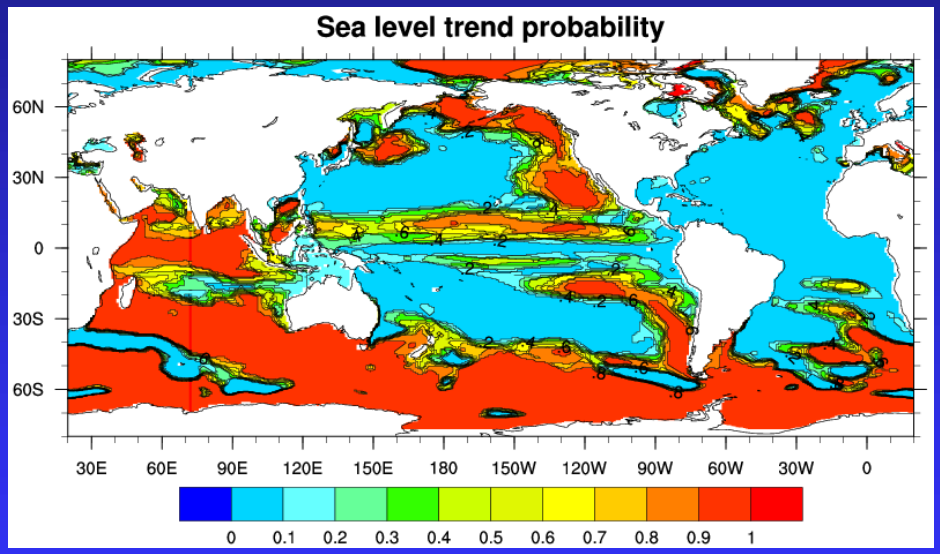
where P is the ensemble size, j is the number of individual members where E occurs

Probabilistic prediction of sea level trends

Probability that the 2005-2015 sea level trend is higher than the mean 1960-2005 global mean trend



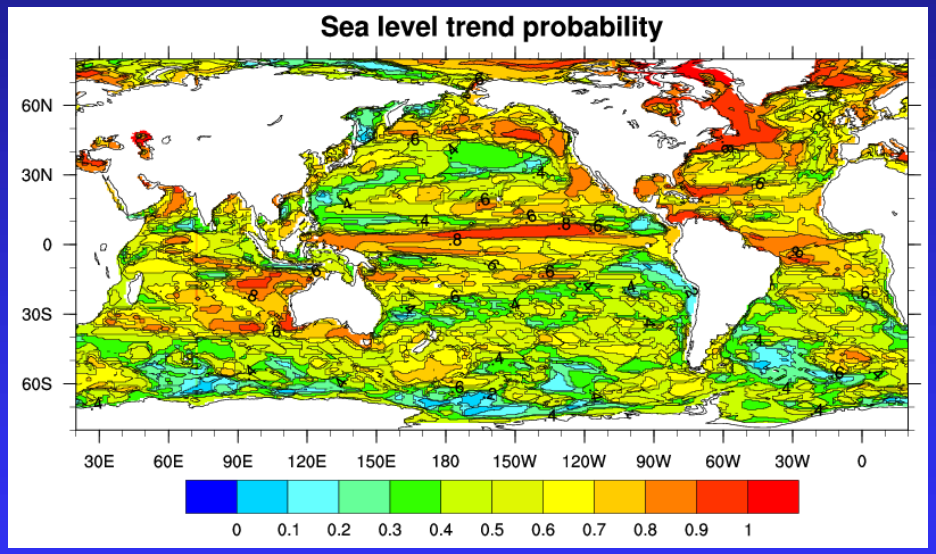
Historical

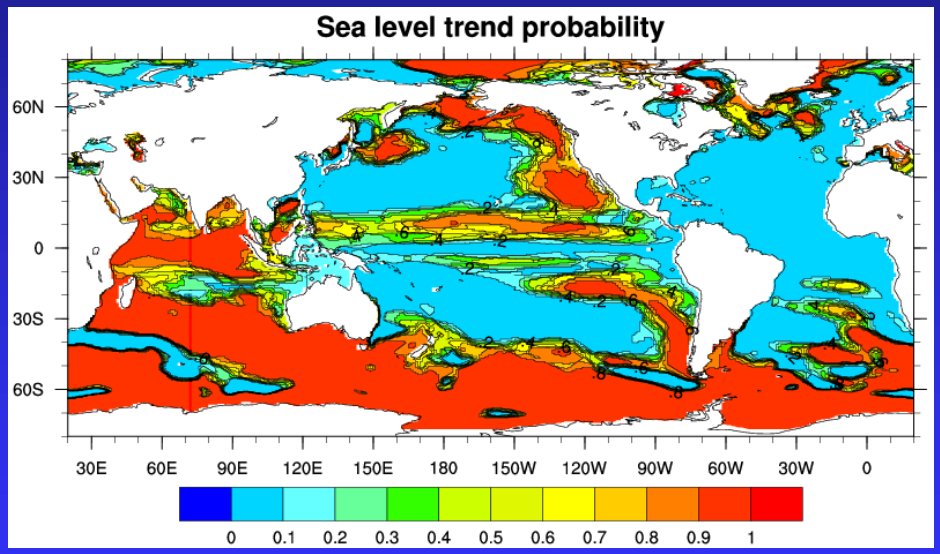


Decadal hindcasts

Probabilistic prediction of sea level trends

Probability that the 2005-2015 sea level trend is higher than the mean 1960-2005 global mean trend





Historical

Decadal hindcasts

Q: Is there some reason to believe these maps ? Is one more reliable than the other ?

Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

Observation : o_i = 0 or 1

Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

Observation : o_i = 0 or 1

Verification dataset needed !

Here: Reanalysis

Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

$$= \int_0^1 [p - o(p)]^2 g(p) dp - \int_0^1 [\bar{o} - o(p)]^2 g(p) dp + \bar{o}[1 - \bar{o}]$$

$$BSS = 1 - \frac{BS}{\boldsymbol{BSc}}$$

$$BSS = BSS_{res} + BSS_{rel}$$

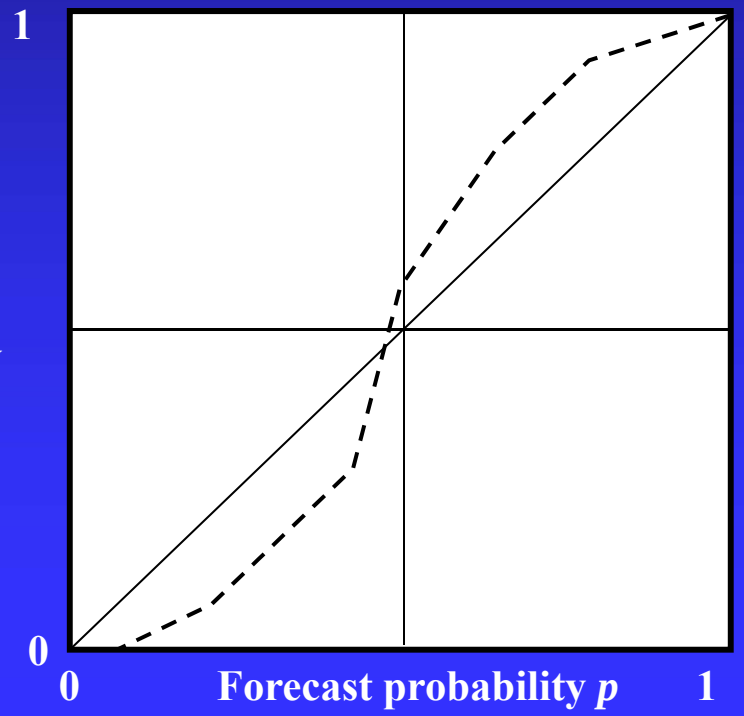
Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

$$= \int_0^1 [p - o(p)]^2 g(p) dp - \int_0^1 [\bar{o} - o(p)]^2 g(p) dp + \bar{o}[1 - \bar{o}]$$

$$BSS = 1 - \frac{BS}{\boldsymbol{BSc}}$$

$$BSS = BSS_{res} + BSS_{rel}$$



e.g. Palmer, 2002

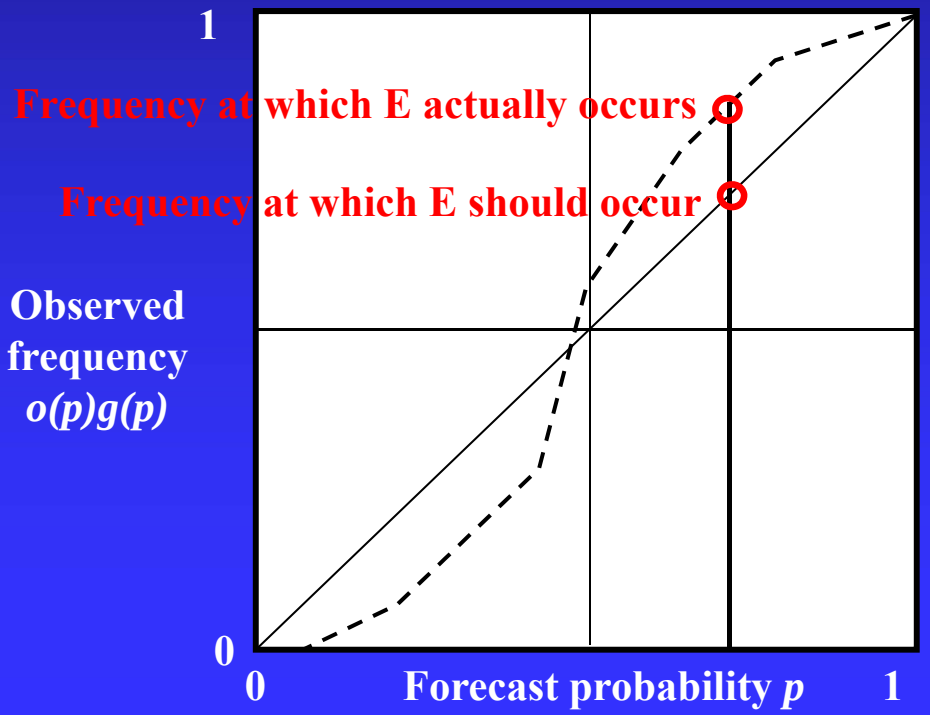
Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

$$= \int_0^1 [p - o(p)]^2 g(p) dp - \int_0^1 [\bar{o} - o(p)]^2 g(p) dp + \bar{o}[1 - \bar{o}]$$

$$BSS = 1 - \frac{BS}{BSc}$$

$$BSS = BSS_{res} + BSS_{rel}$$



e.g. Palmer, 2002

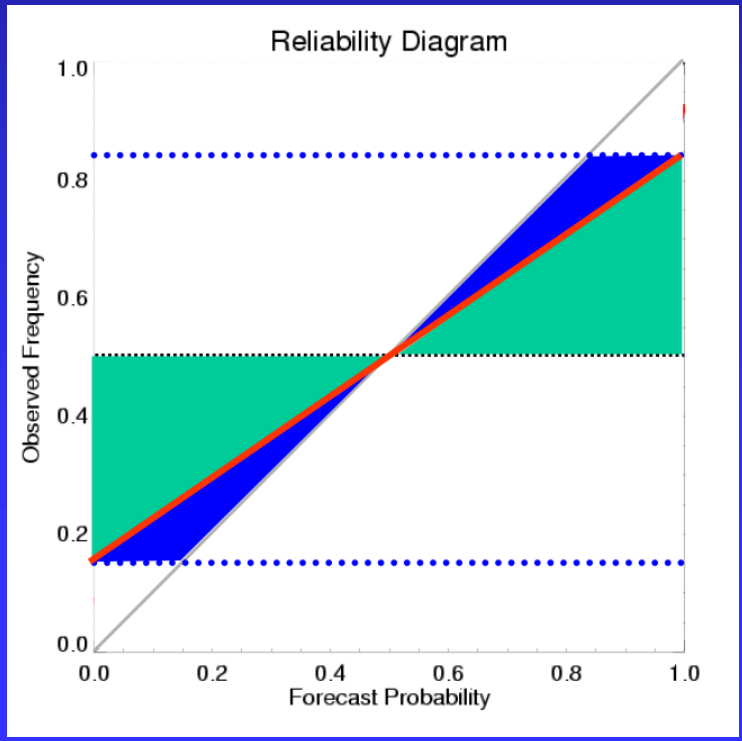
Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

$$= \int_0^1 [p - o(p)]^2 g(p) dp - \int_0^1 [\bar{o} - o(p)]^2 g(p) dp + \bar{o}[1 - \bar{o}]$$

$$BSS = 1 - \frac{BS}{\boldsymbol{BSc}}$$

$$BSS = BSS_{res} + BSS_{rel}$$



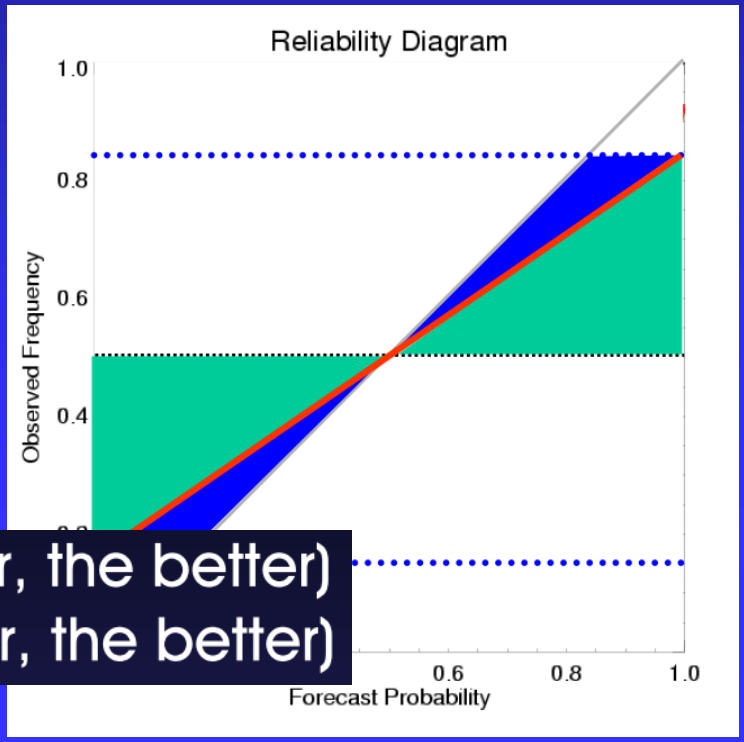
Reliability Estimates (Brier score)

$$BS = \sum_{i=1}^{N} (p_i - o_i)^2$$

$$= \int_0^1 [p - o(p)]^2 g(p) dp - \int_0^1 [\bar{o} - o(p)]^2 g(p) dp + \bar{o}[1 - \bar{o}]$$

$$BSS = 1 - \frac{BS}{\boldsymbol{BSc}}$$

$$BSS = BSS_{res} + BSS_{rel}$$

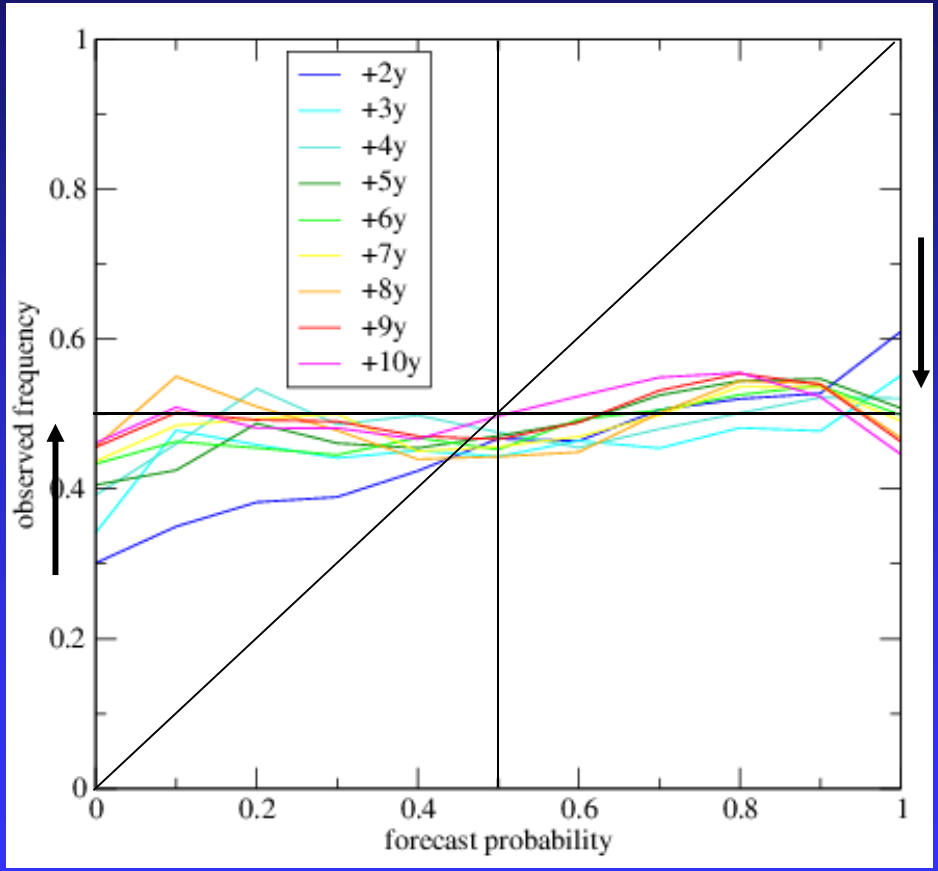


Reliability score (the smaller, the better)

Resolution score (the bigger, the better)

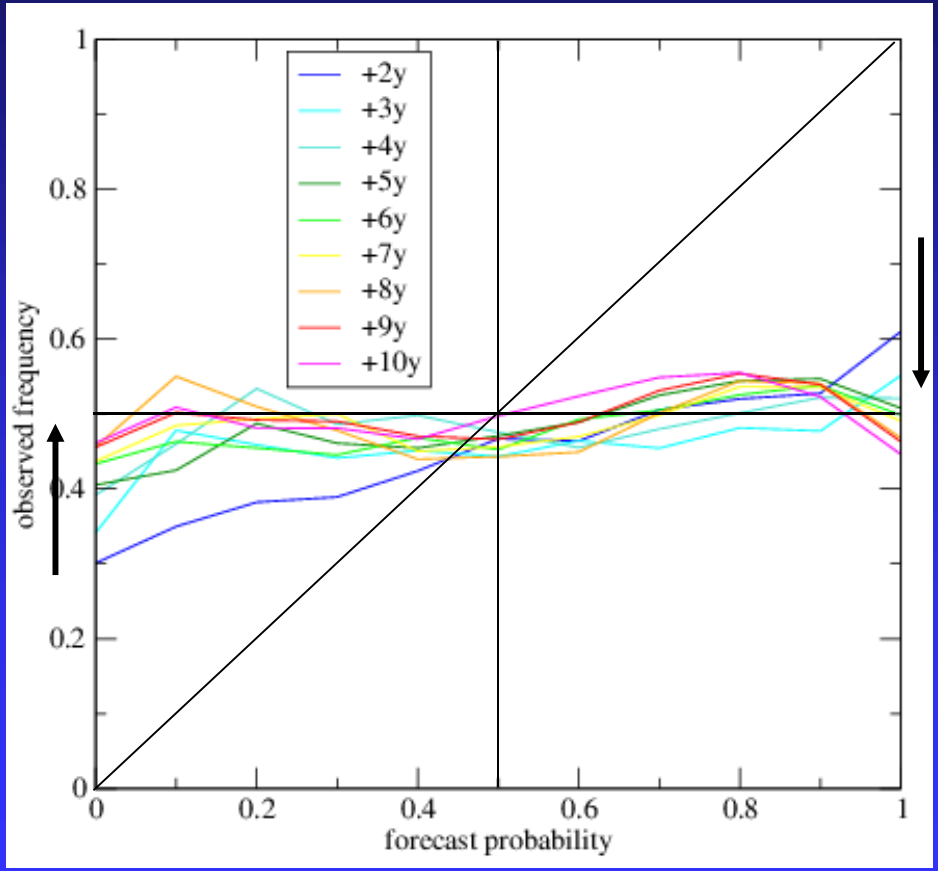
Reliability estimates of probabilistic prediction that the sea level trend is higher than the mean 1960-2005 global mean trend

Decadal hindcasts



Reliability estimates of probabilistic prediction that the sea level trend is higher than the mean 1960-2005 global mean trend

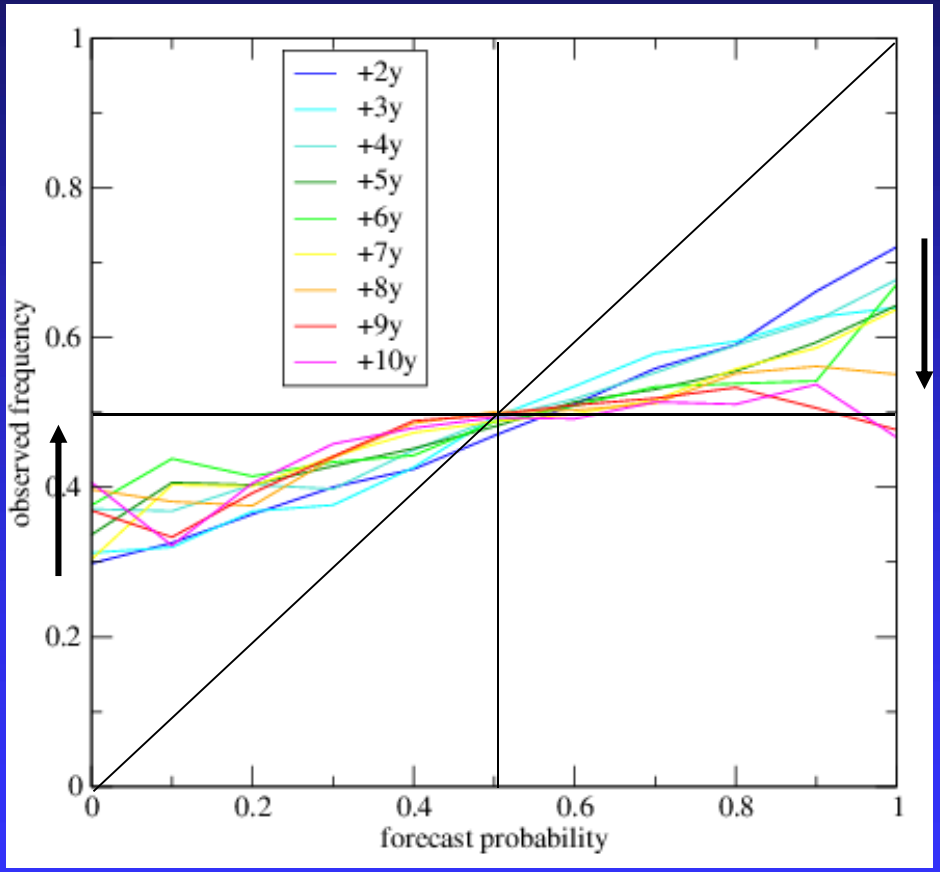
Decadal hindcasts



Reliability diagrams pathological of an « over-confident » prediction system, the ensemble spread is too weak.

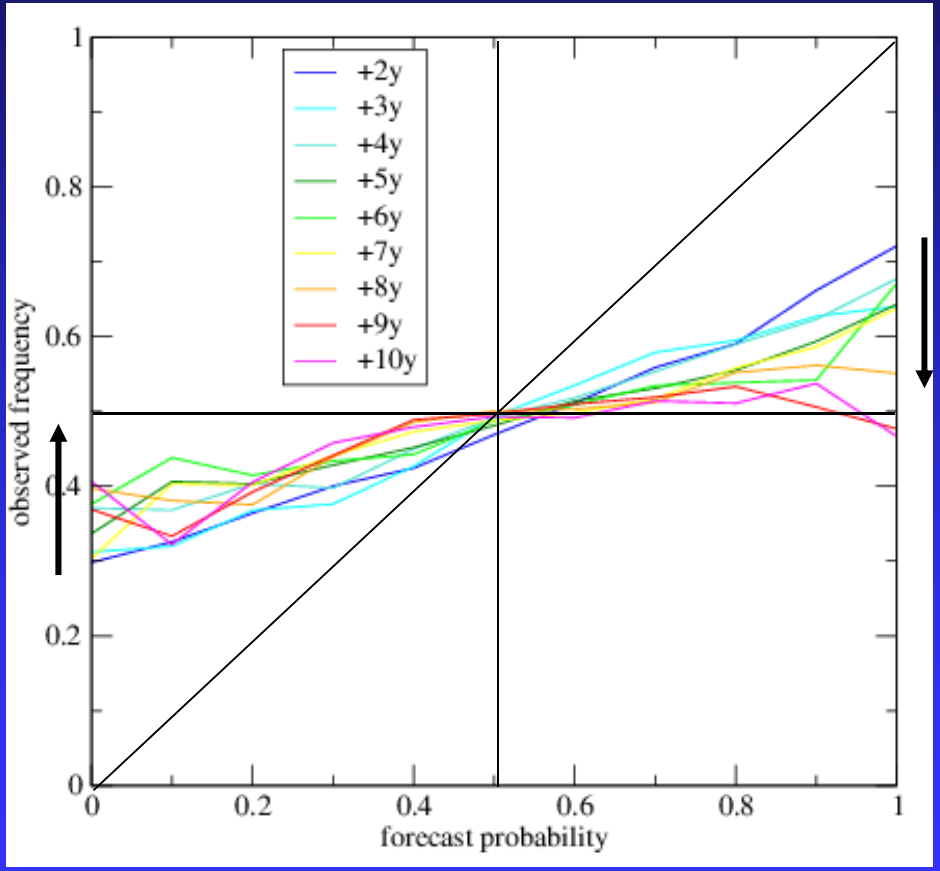
Reliability estimates of probabilistic prediction that the sea level trend is higher than the mean 1960-2005 global mean trend

Historicals



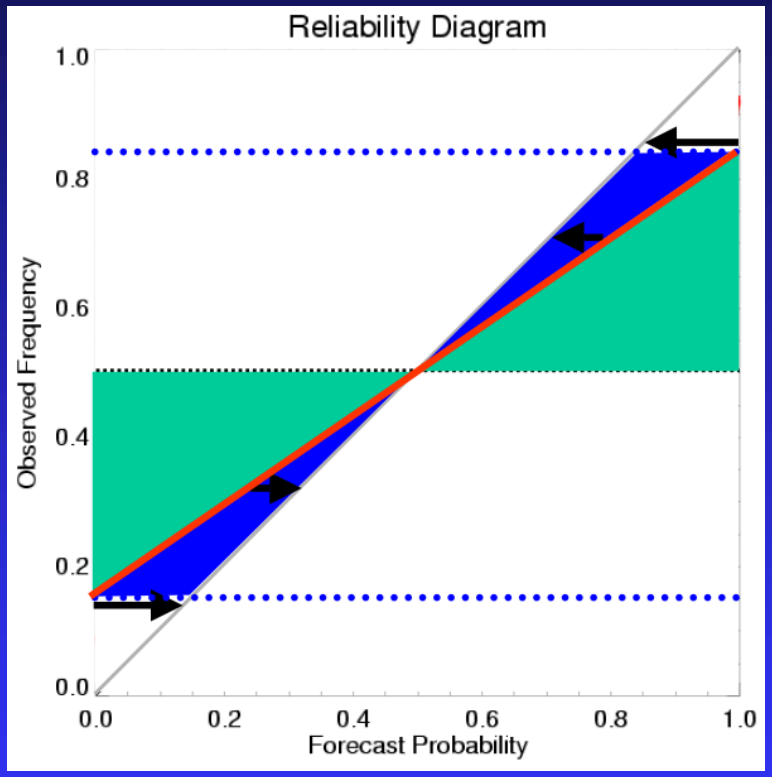
Reliability estimates of probabilistic prediction that the sea level trend is higher than the mean 1960-2005 global mean trend

Historicals



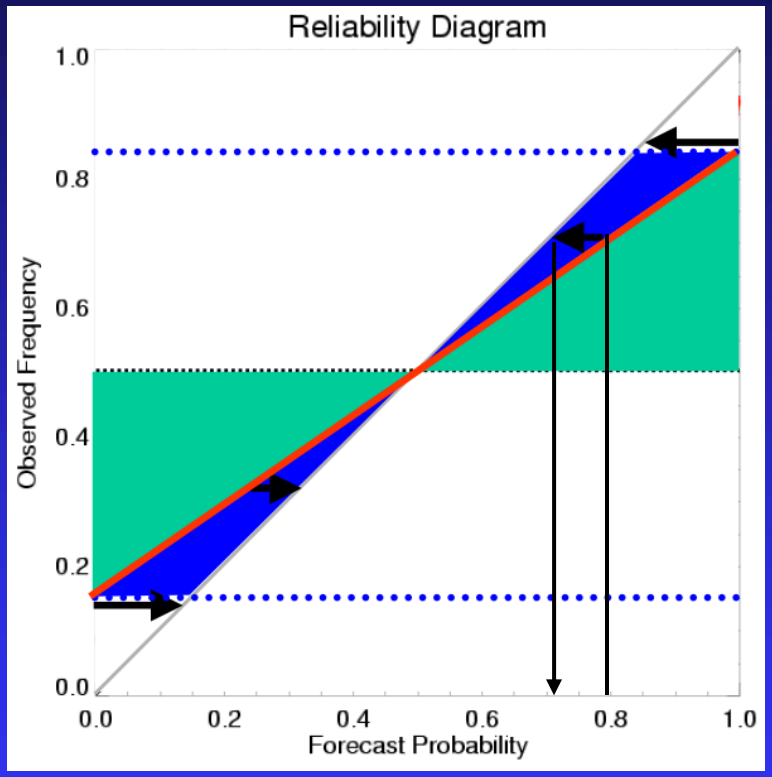
These diagrams lead to negative scores, meaning that no additional information beyond climatology is brought

Reliability diagram or calibration curves



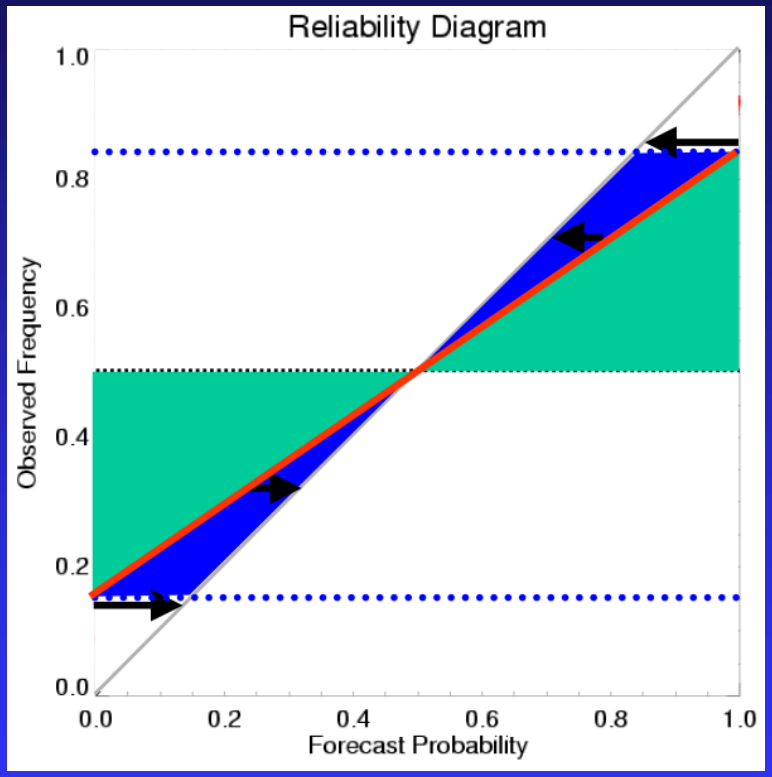
Knowing past (independent) performances of the system allows user to recompute probabilities based on these diagrams.

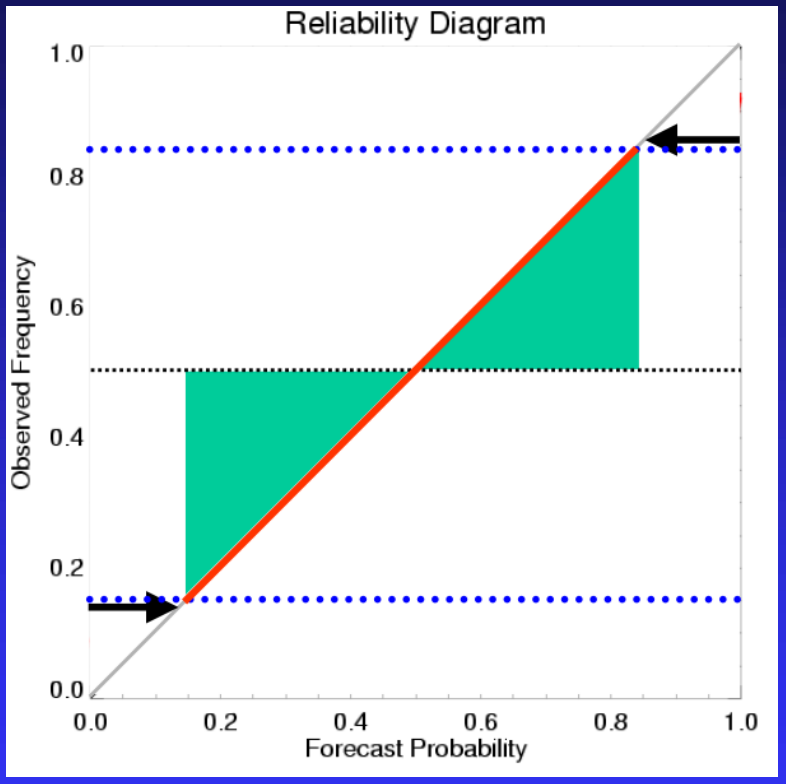
Reliability diagram or calibration curves



Knowing past (independent) performances of the system allows user to recompute probabilities based on these diagrams.

Reliability diagram or calibration curves





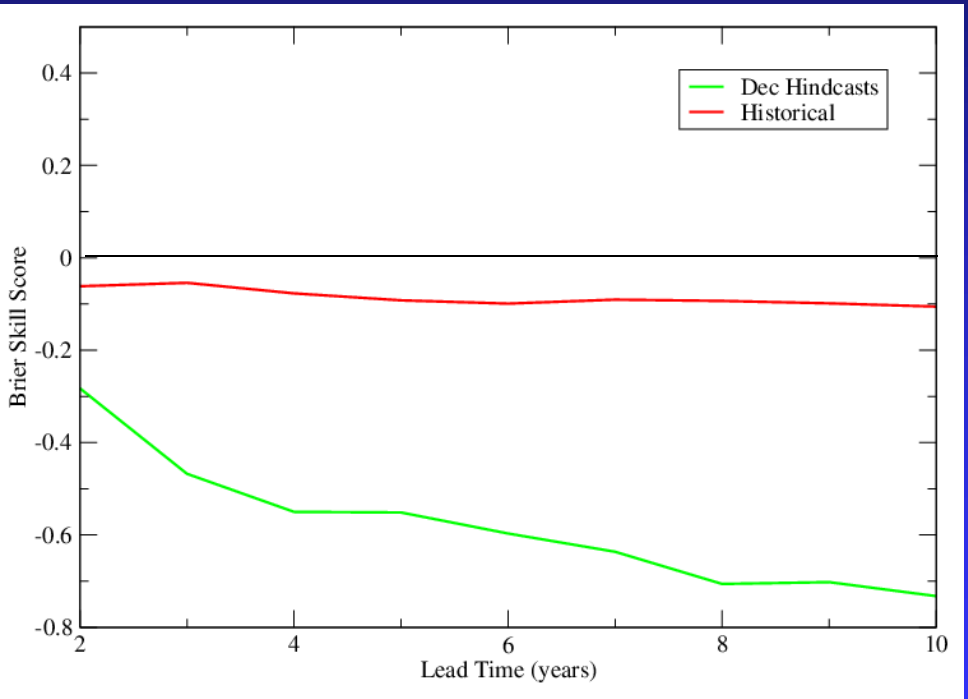
Courtesy of FJ Doblas-Reyes

It leads to more reliable hindcasts, at the cost of a loss of resolution.

It must be made in cross calibration mode, excluding the forecasted year from the calibration phase.

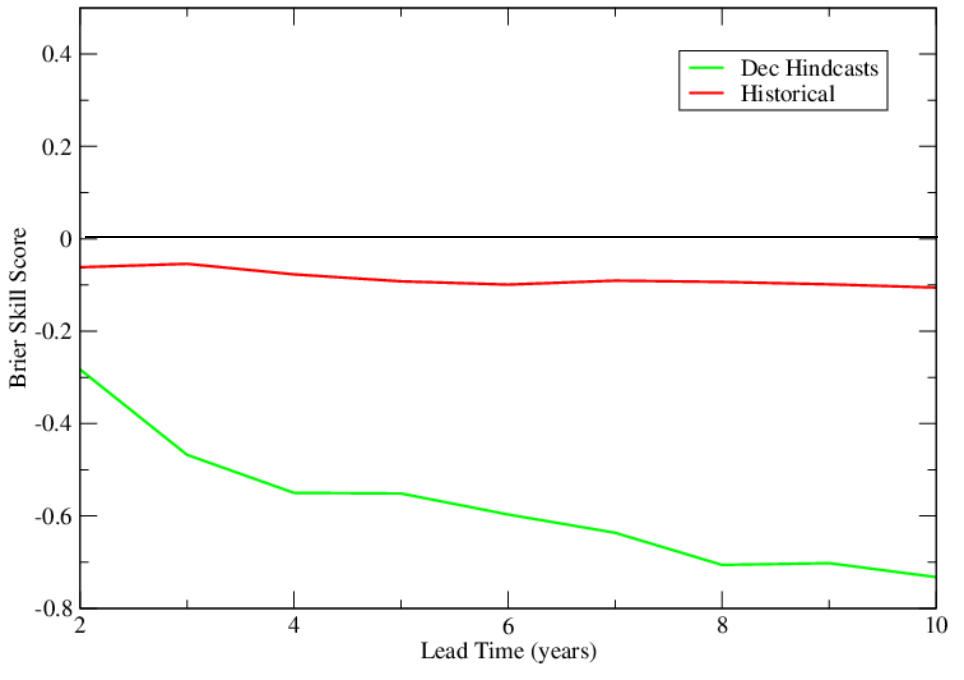
Recalibrated hindcast performances

Before Calibration

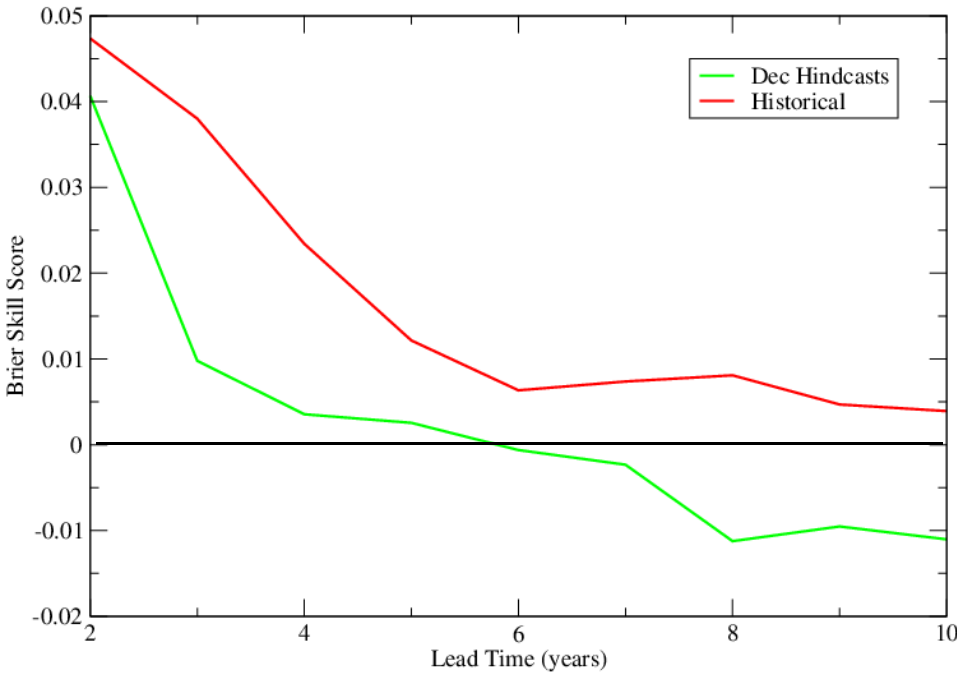


Recalibrated hindcast performances

Before Calibration



After Calibration



Conclusions

• Ensembles of sea level decadal prediction are not reliable if they are not calibrated using the whole hindcast history

• Very moderate skill seems to be found after recalibration; but decadal hindcasts still seem less skilfull than historical experiments; need to revisit initialisation schemes and ensemble generation methods

•Verification dataset issue: we are close to a predictability study; need for further tests, especially against reconstructions

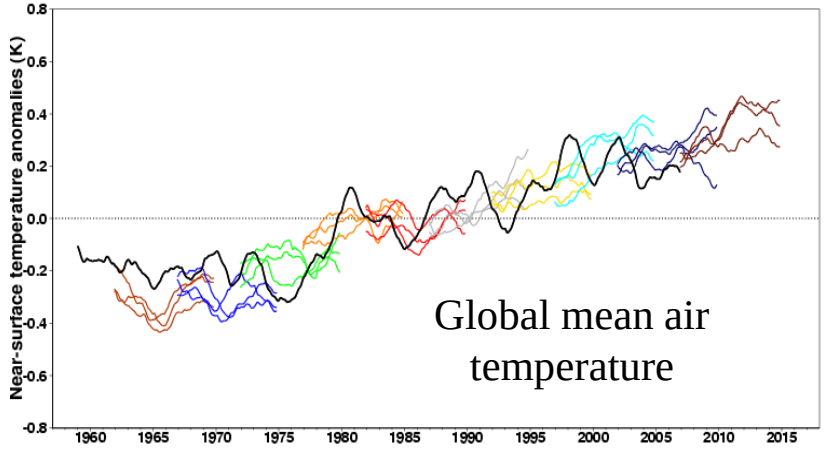
• Ten-member ensembles seem a minimum required to hope to beat climatology; clear hope that multi-model approaches will do better within CMIP5.

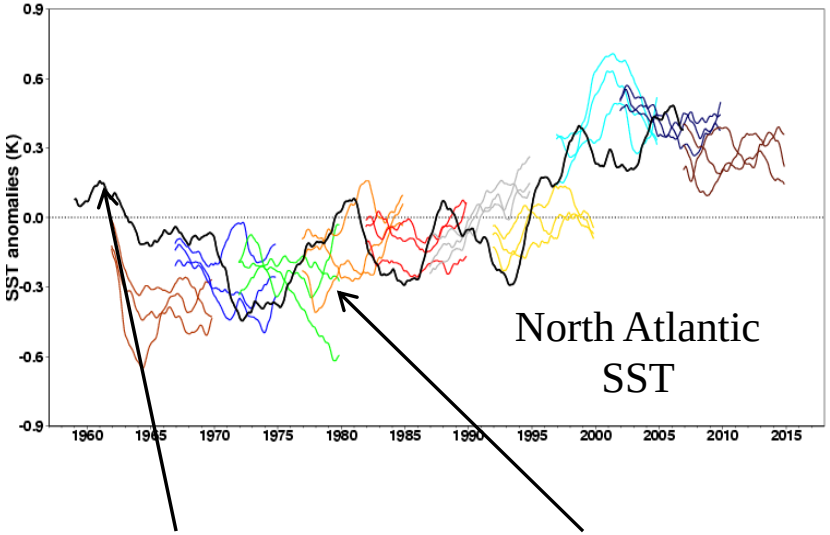
Thank you

http://www.anr-cecile.fr/

Decadal hindcasts, preliminary experiments from the ENSEMBLES project (examples of temperature predictions)

CERFACS (3 members)





Observations

"Forecasts"

Multi model (ECMWF, CERFACS, IfM Kiel, Hadley)