Multi-year prediction of the Atlantic Niño: A first approach from ENSEMBLES

JAVIER GARCÍA-SERRANO[1,2], IRENE POLO[3],
FRANCISCO J. DOBLAS-REYES[1,4] & REINDERT J. HAARSMA[5]

[1]Institut Català de Ciències del Clima (IC3), Barcelona, Spain
[2]LOCEAN-IPSL, Université Pierre et Marie Curie (UPMC), Paris, France
[3]NCAS-Climate, University of Reading, Reading, UK
[4]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[5]Koninklijk Nederlands Meteorologisch Instituut (KNMI), De Bilt, The Netherlands
jgarcia@ic3.cat

Received: 29/04/2013
Accepted: 12/11/2013

Abstract
This work presents for the first time a comprehensive exploration of reliability for decadal predictions upon the Atlantic Niño. Initialized and uninitialized decadal re-forecasts are used to assess multi-year prediction skill of the ATL3 SST index and to evaluate the oceanic regions, over the tropical Atlantic, providing that skill. The analysed decadal re-forecasts were carried out with the UK Met Office Decadal Prediction System (DePreSys) as part of the ENSEMBLES perturbed-parameter ensemble. Re-forecasts started once every year over 1960-2005 are employed; the verification period is 1961-2012, and the summer season considered June throughout September. Model (i.e. HadCM3) SST systematic errors and the impact of the initialization on local SST variability are described. The forecast quality assessment shows that initialized predictions of the ATL3 index yield higher correlation skill and lower root mean square error than the uninitialized counterparts until 2-5 years ahead. The results also indicate that initialization outperforms empirical predictions based on persistence over this forecast range, suggesting that ocean dynamics play a role in this near-term predictability of the Atlantic Niño.
Key words: Atlantic Niño, decadal prediction, ENSEMBLES.

Predicción multianual del Niño Atlántico: Una primera aproximación desde ENSEMBLES

Resumen
Este trabajo explora, por primera vez, la aplicabilidad de la predicción decadal sobre el Niño Atlántico. Predicciones retrospectivas inicializadas y no-inicializadas han sido usadas para evaluar el carácter predictivo del índice ATL3 a escalas multi-anuales, así como para identificar las regiones oceánicas del Atlántico tropical que contribuyen a esa predictabilidad. Las predicciones retrospectivas usadas en este estudio fueron generadas con el Sistema de Predicción Decadal (DePreSys) del UK Met Office, como parte de la contribución de parámetros-perturbados del proyecto ENSEMBLES. Estas predicciones retrospectivas fueron realizadas cada año durante el período 1960-2005; el período de verificación se extiende desde 1961 a 2012, y la media estacional empleada corresponde a Junio-Septiembre. Primeramente se describen los errores sistemáticos del modelo climático (HadCM3), así como el impacto de la inicialización en la variabilidad local de la temperatura del mar. Posteriormente, otros resultados muestran que las predicciones inicializadas poseen una mejor habilidad para reproducir la evolución observada del índice ATL3, que las no-inicializadas, hasta el rango 2-5 años, con una mayor correlación y un menor error cuadrático medio.

ISSN: 0214-4557
http://dx.doi.org/10.5209/rev_FITE.2013.v25.43435

Los resultados también sugieren que la inicialización mejora las predicciones estadísticas basadas en persistencia, indicando que la dinámica oceánica juega un papel importante en la predictabilidad del Niño Atlántico.
Palabras clave: Niño Atlántico, predicción decadal, ENSEMBLES.

Summary: 1. Introduction. 2. Data and methods. 3. Results, 3.1. Model bias and tropical Atlantic variability, 3.2. Forecast quality assessment. 4. Summary and conclusions. 5. Acknowledgements. 6. References.

Normalized reference
García-Serrano, J., Polo, I., Doblas-Reyes, F.J., Haarsma R.J. (2013). Multi-year prediction of the Atlantic Niño: a first approach from ENSEMBLES. *Física de la Tierra*, Vol. 25, 57-71.

1. Introduction

The Atlantic Niño is one of the phases of the leading mode of the tropical Atlantic variability and it consists in an anomalous warming over the eastern equatorial Atlantic which develops through Bjerknes feedback (Zebiak 1993; Carton et al. 1996; Ruiz-Barradas et al. 2000; Keenlyside and Latif 2007; Polo et al. 2008). The equatorial warming usually comes together with positive sea surface temperature (SST) anomalies over the Angola/Benguela coast shaping a characteristic "cold tongue" configuration (Polo et al. 2008). The opposite is true for the Atlantic Niña or cold equatorial Atlantic events. The Atlantic Niño can be characterised with the ATL3 index, which is the result of averaging SST over the region 20ºW-0ºE / 3ºN-3ºS (Zebiak 1993).

State-of-the-art coupled models fail in reproducing the Atlantic Niño mainly because they underestimate the annual cycle under an already too warm mean state (warm SST bias; see also Fig. 1). Much research has focused in the past on understanding the strong mean bias over the equatorial Atlantic, pointing out at key processes such as a complex land-sea interaction in the Atlantic basin or the oceanic coarse resolution for reproducing no enough upwelling. The overall conclusion is that the air-sea coupling results in a positive feedback that leads to the development of the SST error: models fail in especial areas where poor simulation of certain atmospheric phenomena such as deep convection over Brazil and the stratocumulus over Angola coast are driving wrong surface fluxes and winds, which in turn yield a too deep mean equatorial thermocline and a warm SST bias (Large and Danabasoglu 2006; Chang et al. 2007; Ritcher and Xie 2008; Ritcher et al. 2011; Wahl et al. 2011). In addition, some authors have investigated the processes related to the growth of the SST error in several CMIP5 decadal hindcasts over the south tropical Atlantic, concluding that, in particular, for HadCM3 the ocean integrates the anomalous deep thermocline without advecting the excess of heat content out (Toniazzo and Woolnough 2013).

However, despite the mean bias over the equatorial Atlantic, most models are able to simulate some kind of variability over the ATL3 region, although with smaller amplitude than in the observations (Collins et al. 2001; Breugem et al. 2006; see also Tokinaga and Xie 2011 additional material; Richter et al. 2012) or with the wrong seasonality. For instance, the interannual variability peak for the ATL3 index in HadCM3 occurs in boreal autumn instead of boreal summer as in the observations (Polo et al. 2013). Capturing the right seasonality of the events is important in terms of the teleconnections associated with the Atlantic Niño. The Atlantic Niño has an impact on precipitation over the Nordeste Brazil region and onto the West African monsoon (García-Serrano et al. 2008; Losada et al. 2010), which are systems phase-locked with the seasonal cycle.

Since the Atlantic Niño exerts an influence on the rainfall over the Tropics, where rain-fed agriculture is an important part of the economy, being able to predict its variability on seasonal to decadal timescales is key for decision-makers (Rodríguez-Fonseca et al. 2011).

Decadal prediction systems have been developed and many works have been devoted to evaluate SST predictions over the North Atlantic and tropical Pacific and its associated impacts (among others, García-Serrano et al. 2012; Robson et al. 2012).

The reliability of SST multi-year ensemble predictions has been evaluated globally (Doblas-Reyes et al. 2013; Ho et al. 2013) but it has never been specifically assessed over the equatorial Atlantic. For instance, Ho et al (2013) have shown that in general terms, the initialization significantly reduces forecast spread, in agreement with previous studies (Corti et al. 2012). However, overdispersion within the ensemble, related to an excess of interannual variability of the model (i.e. HadCM3), make predictions to be not reliable (Ho et al. 2013). Interestingly, the former authors have shown that over the equatorial Atlantic the spread-error ratio for short lead times in DePreSys is close to 1 (Figure 1 in Ho et al. 2013), which suggests that ATL3 region could be promising for future reliable predictions. Recently some authors have also evoked some skill for ATL3 as well as for the associated rainfall over the Gulf of Guinea (García-Serrano et al. 2013).

This work is the first attempt to comprehensively explore the reliability of decadal predictions upon the Atlantic Niño; and, to this aim decadal re-forecasts with yearly start dates have been employed. This work moves forward from García-Serrano et al. (2013) who analysed the ENSEMBLES decadal re-forecasts with a five-year interval between start dates. Increasing the start dates frequency has been recently shown to be of major relevance rather than increasing the ensemble size (García-Serrano and Doblas-Reyes 2012; Meehl et al. 2013).

2. Data and methods

This study uses the decadal re-forecasts of the perturbed-parameter ensemble predictions produced with the Met Office Decadal Climate Prediction System (DePreSys; Smith et al. 2007, 2010) as part of the EU project ENSEMBLES (Doblas-Reyes et al. 2010). Ten-year long re-forecasts were started on the first of November once every year over the period 1960-2005 using a nine-member ensemble of HadCM3 model variants. The set of initialized decadal re-forecasts used in this study was run explicitly prescribing the contemporaneous state of the climate system at the start date; these will be referred, simply, to as ASSIM. In order to assess the impact of the initialization an additional set of uninitialized re-forecasts (referred to as NOASSIM) with the same nine model versions was run.

What is unique in these two sets, in comparison with the Coupled Model Inter-comparison Project phase 5 (CMIP5) protocol (Taylor et al. 2012), is that they use information available in a real-time forecast context. For that reason we refer to these decadal integrations as re-forecasts, not hindcasts. The ENSEMBLES decadal re-forecasts do not include observed time-dependent variations in solar activity or volcanic aerosols, but they take into account changes in external forcing such as greenhouse gases, sulphate aerosols, and anthropogenic emissions. Further description of the experimental design is provided by Doblas-Reyes et al. (2010).

The study focuses on time-averaged predictions over an extended summer season, ranging from June to September (hereafter, JJAS). The ASSIM and NOASSIM model climate is estimated by averaging raw forecasts along the actual time (the start date dimension) according to availability of observations. The period for verification covers 1961-2012. All re-forecast anomaly time-series have been computed, for each model version separately, by removing a model climate estimate at each forecast time. A four-year forecast average is performed afterwards upon these drift-corrected anomaly time-series, to retain interannual-to-decadal predictability (Goddard et al. 2013). Four-year averages are considered as it represents a compromise between the capability of partially removing the unpredictable interannual variability in near-term dynamical forecasting (e.g. the link to ENSO) and the ability to partially represent skill evolution along the forecast time (García-Serrano and Doblas-Reyes 2012). The NOAA extended reconstructed SST v3b (ERSST; Smith et al. 2008) has been used as the reference dataset for verification. The subset of observational data used to estimate the respective climatology is selected taking into account only years when both observations and model data are available. In this case, a separate observational climatology is estimated for each four-year forecast average.

The purpose of this research is to evaluate multi-year forecast quality upon the Atlantic Niño. Firstly, the skill of the DePreSys decadal predictions of the Atlantic-3 SST (ATL3) index is assessed. The ATL3 index adopted in this study was defined in García-Serrano et al. (2013), which corresponds to a revised definition of the one by Zebiak (1993) with the aim of removing long-term trends: SST anomalies averaged over 20ºW-0ºE / 3ºS-3ºN minus global SST anomalies averaged over 60ºS-60ºN. This definition mimics the approach by Trenberth and Shea (2006) to revise the AMO index. Two different measures of forecast quality have been employed to analyse the relative merits of the forecast systems. The skill scores include the anomaly correlation coefficient (ACC) and root mean square error (RMSE) of the ensemble mean (Joliffe and Stephenson 2003; García-Serrano and Doblas-Reyes 2012). The DePreSys decadal predictions are compared against two empirical prediction methods. These methods are based on 1-year and 4-year persistence; the second one is intended to represent low-frequency persistence relevant for the prediction skill assessment (four-year forecast averages).

The second objective of this study is the identification of regions in the tropical Atlantic that potentially contribute to improved multi-year ATL3 prediction skill in ASSIM when compared to NOASSIM. To this aim, grid-point correlation maps between ensemble-mean detrended re-forecast SST anomalies and the observed (ERSST) ATL3 index are performed. Detrended anomalies are computed by subtracting global SST anomalies averaged over 60ºS-60ºN, for consistency with the ATL3 index employed. It is found that linear detrending or regressing out the global SST average yield very similar results (not shown). This approach, already applied to the AMO index (García-Serrano et al. 2012), provides a simple, systematic procedure for quantitatively assessing performance in different forecast systems and complements other initiatives in establishing a common verification framework for decadal prediction (Goddard et al. 2013; Meehl et al. 2013).

To avoid getting confidence levels that are too liberal, the computation of an effective sample size is used to estimate the statistical thresholds. This is tackled by computing the effective degrees of freedom as described in von Storch and Zwiers (2001) or Zieba (2010).

3. Results

3.1. Model bias and tropical Atlantic variability

Once initialized using observations, climate prediction models tend to develop, along the forecast time, systematic errors leading to a drift towards their preferred climate. This is specially marked when the observed state is prescribed in a *full-field initialization*, as in the ENSEMBLES multi-model (e.g. García-Serrano and Doblas-Reyes 2012). A way to reduce this model drift is prescribing observed anomalies into the imperfect model climate, which is known as *anomaly initialization* (Doblas-Reyes et al. 2010; Hazeleger et al. 2013; Smith et al. 2013). DePreSys, and then ASSIM, uses anomaly initialization (Smith et al. 2007, 2010). However, this initialization strategy does not prevent from having model biases in the course of the forecast time, since they are inherent of the climate model (i.e. the forecast system).

Figure 1 illustrates the SST systematic error over the tropical Atlantic in NOASSIM (a) and ASSIM (b); shown is the ensemble-mean of the nine model variants of HadCM3 from both ensembles for the forecast average 2-5 years. As can be seen, the initialization does not play a significant role in the model SST biases, as the two ensembles show identical patterns. A strong warm bias is evident over the Angola/Benguela upwelling region, which reaches around +2.5ºC compared to observations. This is accompanied by a weaker warm bias at equatorial latitudes over the cold tongue region. In the central tropical Atlantic, offshore Northeastern Brazil, a cold bias of around -1ºC is apparent. Likewise, weaker cold systematic errors are shown at subtropical latitudes of both hemispheres. This pattern compares well with the one described by Smith et al. (2013) from a ten-member historical simulation with HadCM3.

Figure 1 also depicts an estimate of the SST model drift in both systems, NOASSIM (c) and ASSIM (d), computed as the difference in the model climate between the forecast averages 7-10 and 1-4 years. As expected from the DePreSys set-up, with re-forecasts released from its model climate, the drift is quite small, reaching only 0.1-0.2ºC. The homogeneous pattern of warming drift in both systems suggests that the prescribed, varying external forcing is key in phasing the model response.

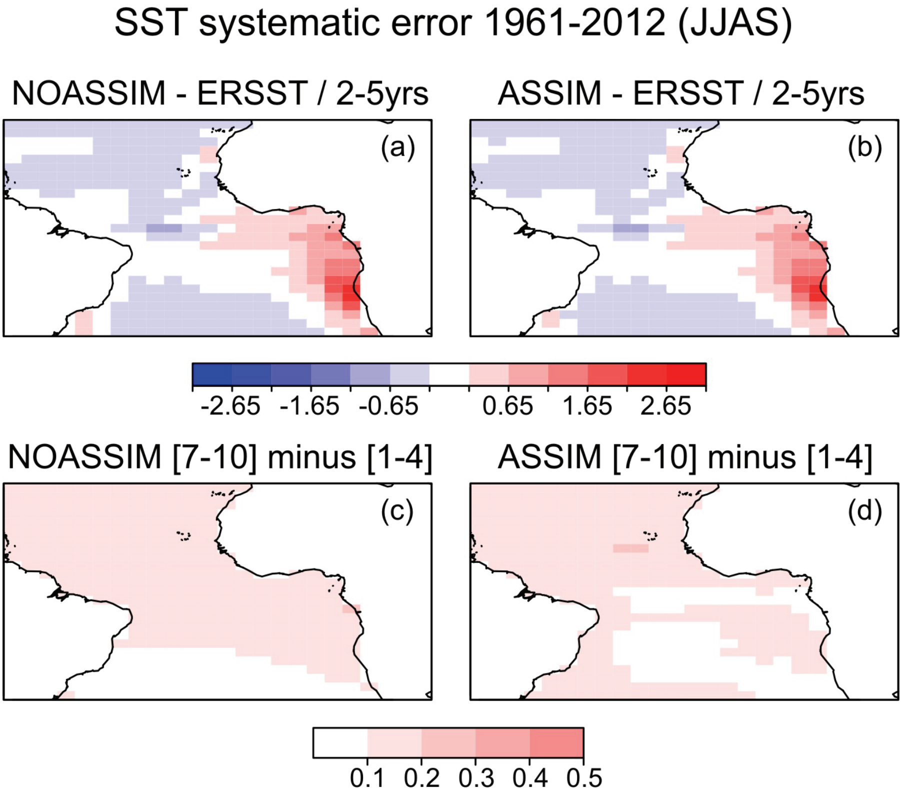

Fig. 1. (top) Ensemble-mean SST systematic error (ºC), with respect to ERSST, in the uninitialized (NOASSIM; a) and initialized (ASSIM; b) decadal re-forecasts for the forecast period 2-5 years. (bottom) Estimate of the SST model drift (ºC) in both systems (NOASSIM, c; ASSIM, d), computed as the difference in model mean-climate between the forecast periods 7-10 and 1-4 years. Both analyses are performed over the period 1961-2012 and after applying a four-year forecast average.

This attempt to point out some systematic errors of the forecast systems is intended to improve model performance and initialization procedures, which could lead to get more skilful predictions over the tropical Atlantic. In the following, an illustration of the benefits in local SST variability from initialization is presented.

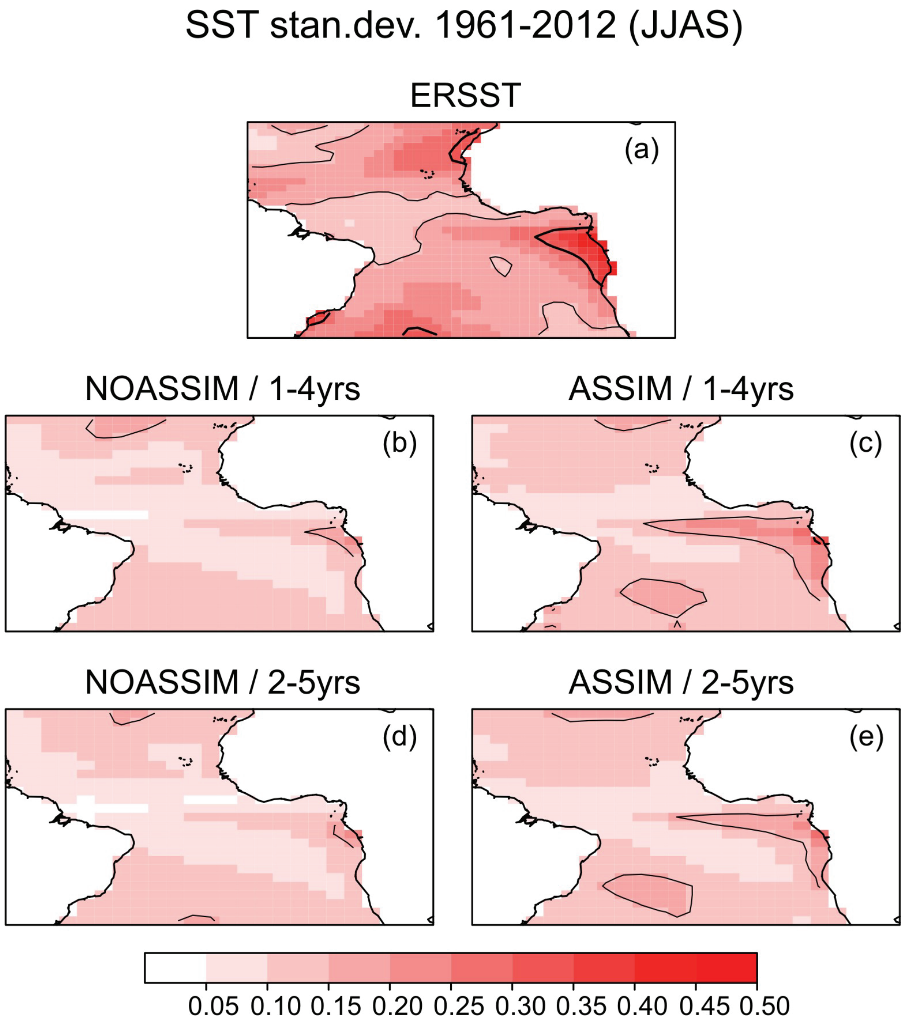

Fig. 2. (a) Observed and (b-e) ensemble-mean re-forecast SST standard deviation (ºC) computed from detrended anomalies over the period 1961-2012; shown for uninitialized (NOASSIM; b, d) and initialized (ASSIM; c, e) decadal re-forecasts are the forecast periods 1-4 (middle) and 2-5 years (bottom). Detrended anomalies are computed by subtracting global SST anomalies averaged over 60ºS-60ºN. Contours are drawn for 0.15 (thin black) and 0.30 (thick black) ºC.

Figure 2 shows the SST standard deviation at each grid-point from ERSST (a), over the verification period 1961-2012, and from NOASSIM (b,d) and ASSIM (c,e) ensembles for the first two forecast averages in which the impact of the initialization is expected to be maximum, namely 1-4 (middle row) and 2-5 (bottom) years. In the observations, and at tropical latitudes, the SST standard deviation peaks in the easternmost side, over the Angola and Guinea basins, and extends along the cold tongue region up to the central equatorial Atlantic. In the forecast systems an overall underestimation of the observed variability is noticeable. Even so, a clear enhancement of the SST standard deviation is apparent in ASSIM at equatorial latitudes, from the African coastline towards the central tropical Atlantic, for both forecast periods. This improvement is linked to the initialization. The question emerges whether this improvement translates into a better skill in re-forecasting the Atlantic Niño. This is addressed in the next section.

3.2. Forecast quality assessment

Figure 3 shows the ATL3 forecast skill assessment by means of the ACC (a) and RMSE (b) scores from the dynamical prediction systems, NOASSIM (pink) and ASSIM (purple), as well as two empirical models based on persistence, i.e. 1- (black solid) and 4-year (black dashed) persistence. The latter is considered to represent interannual-to-decadal predictability coming from low-frequency, simple damping processes, whereas the former indicates memory of the climate system due to the initial state. Note that in general 1-year persistence predictions are even less skilful than the 4-year persistence model, which points that the Atlantic Niño effectively has variability at decadal time-scale, as suggested by García-Serrano et al. (2013) from a power spectrum analysis of the ATL3 index.

Concerning the dynamical forecast systems, there appears that no statistically significant multi-annual prediction skill is found for the ATL3 index. Nonetheless, a systematic improvement in the skill scores from initialization can be noticed early in the forecast range. ASSIM shows correlation coefficients of 0.5 and seems to outperform predictions based on climatology in terms of RMSE at the forecast periods 1-4 and 2-5 years. For those forecast times, the initialized predictions (ASSIM) are more skilful than the uninitialized ones (NOASSIM). This result, using a larger sampling (yearly re-forecasts), reinforces the evidence shown by García-Serrano et al. (2013), from DePreSys re-forecasts with the standard 5-year interval between start dates, that initialization may bring skilful multi-year predictions in the tropical Atlantic Ocean.

To further explore the better performance of ASSIM, compared to NOASSIM, in re-forecasting the SST variability associated with the Atlantic Niño, correlation maps are computed using the observed ATL3 index (see Section 2). Figure 4a shows the SST correlation map of the ATL3 index in the observations for the first forecast average, 1-4 years, which represents the target pattern to be captured by the forecast systems. Note that all forecast averages depict a similar spatial pattern in the observations (not shown), indicating that the SST variability associated with the Atlantic Niño is robust over the verification period. In the climate prediction models, in both the NOASSIM (Fig. 4b) and ASSIM (Fig. 4c) systems, the

Atlantic Niño overestimates the observed meridional SST gradient with the northern tropical-subtropical latitudes while the observed second relative maximum in the southern hemisphere is absent. Besides these systematic errors, both forecast systems have a well-defined warm tongue extending over the equatorial region that resembles the observed Atlantic Niño. All forecast averages in the two ensembles show a similar pattern (not shown), which implies that the same dynamics, responsible of the model Atlantic Niño, is at play independently of the initialization.

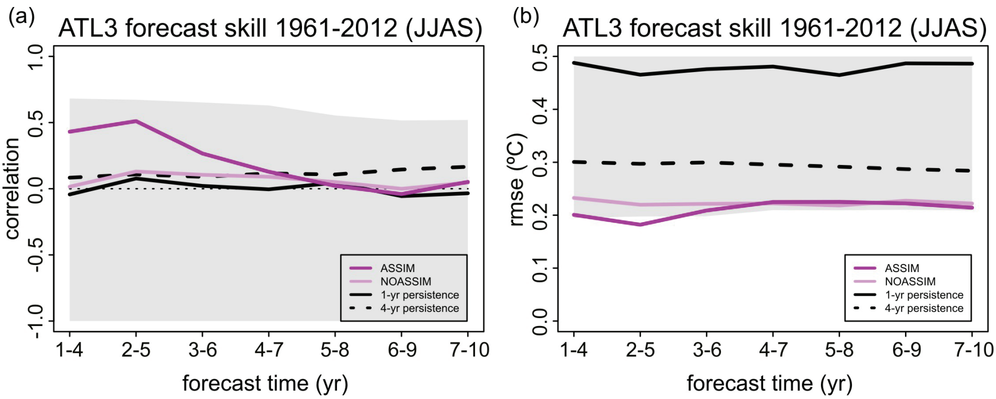

Fig. 3. Ensemble-mean anomaly correlation coefficient (a) and root mean square error (rmse, ºC; b) between the uninitialized (NOASSIM, pink) and initialized (ASSIM, purple) forecast systems, 1- (solid black) and 4-year (dashed black) persistence models, and the observational reference dataset (ERSST) for the ATL3 index, based on JJAS seasonal mean and computed separately for each four-year forecast average over the period 1961-2012. The statistical significance threshold for correlation skill in panel (a), estimated from a one-tailed *t*-test at 95% confidence level taking into account the autocorrelation of the ERSST ATL3 time-series, is drawn in grey shading. In panel (b), grey shading stands for values above the RMSE of predictions based on climatology, computed as the standard deviation of the ATL3 index at each four-year forecast average.

Correlating the observed ATL3 index with re-forecast SST anomalies provides prediction skill information, since evidence of grid-point ability in recapturing the observed low-frequency variability is shown. NOASSIM is unable to re-forecast the observed Atlantic Niño signature at equatorial latitudes during the 1-4 (Fig. 4d) and 2-5 (Fig. 4f) forecasting years. This result goes along with the lack of correlation skill for the ATL3 index in this system (Fig. 3a; pink).

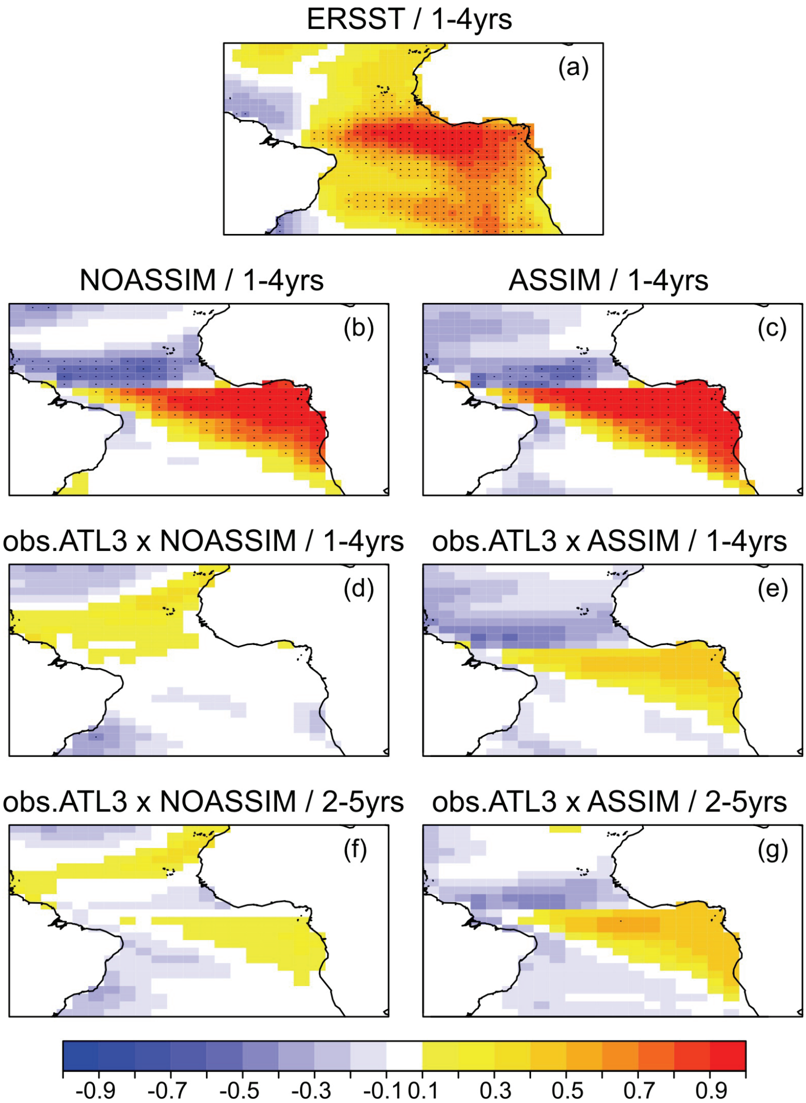

Fig. 4. (a) Correlation map of detrended JJAS ERSST anomalies onto the observed ATL3 (obs.ATL3; ERSST) index. (second row) Ensemble-mean correlation map of detrended JJAS SST anomalies onto the ATL3 index in the uninitialized (NOASSIM; b) and initialized (ASSIM; c) forecast systems. (third-fourth rows) Correlation at each grid point between ensemble-mean detrended JJAS SST anomalies from the uninitialized (NOASSIM; d, f) and initialized (ASSIM; e, g) decadal re-forecasts and the obs.ATL3 index (ERSST); shown are the forecast periods 1-4 (d-e) and 2-5 years (f-g). Detrended anomalies are computed by subtracting global SST anomalies over 60ºS-60ºN. Statistically significant areas, evaluated with a two-tailed *t*-test at 95% confidence level taking into account the autocorrelation of the ERSST ATL3 time-series, are indicated with dots.

By contrast the initialized system (ASSIM) shows skill, albeit not statistically significant, in predicting the observed Atlantic Niño warm tongue at these two early forecast averages (Figs. 4e,g), which is consistent with its ATL3 positive correlations (Fig. 3a; purple). According to the ATL3 skill decrease, with forecast time, in this system no other forecast average shows large, positive grid-point correlations with the observed ATL3 index over the tropical Atlantic (not shown).

All these results suggest that initializing climate models with the contemporaneous state of the system increases the ATL3 forecast skill and improves the capability of predicting SST variability associated with the Atlantic Niño up to 2-5 years ahead. This represents encouraging prospects for future efforts in the decadal prediction framework over the tropical Atlantic.

4. Summary and conclusions

In this study, initialized and uninitialized decadal re-forecasts have been used to assess multi-year forecast skill of the ATL3 SST index and to evaluate the oceanic regions, over the tropical Atlantic, providing that skill. Those decadal prediction experiments were carried out using DePreSys as part of the ENSEMBLES perturbed-parameter ensemble. This work is the first attempt to comprehensively explore the reliability of decadal predictions upon the Atlantic Niño; and, to this aim yearly decadal re-forecasts have been employed. This work takes over from García-Serrano et al. (2013) who analysed the ENSEMBLES decadal re-forecasts with a five-year interval between start dates. Increasing the start-date frequency has been recently shown to be of major relevance rather than increasing the ensemble size (Meehl et al. 2013). Despite limitations on statistical significance of the results, coming from the autocorrelation of the ATL3 index, this first approach has led to some relevant results that are summarized below:

- The impact of the initialization on predictions of the ATL3 index is noticeable early in the forecast range, up to 2-5 years ahead, whereby the initialized decadal prediction system (ASSIM) yields higher correlation skill and lower root mean square error than the uninitialized counterpart (NOASSIM).
- Initialization also outperforms empirical predictions based on one- and four-year persistence damping.
- The improvement in re-forecasting the ATL3 index and ATL3-related SST anomalies in the initialized decadal predictions is concomitant with enhanced local SST variability over the central-eastern equatorial Atlantic.

Our results suggest that ocean dynamics play a role in reproducing the Atlantic Niño evolution beyond the thermal inertia of the ocean. This is in agreement with the idea of changes in the Atlantic meridional overturning circulation impacting the mean thermocline and thus Atlantic Niño mode (Chang et al. 2008; Haarsma et al. 2008; Polo et al. 2013). It also implies that ATL3-related climate phenomena such as rainfall variability in Northeastern Brazil or the WAM may have some predictability at 2-5 years lead time, and could be skilfully predicted on those time scales if the models can reproduce the relevant climate processes. Likewise, this study illustrates that this internally-generated low-frequency variability is correctly initialized in the decadal prediction system (ASSIM), and envisages encouraging prospects for further development towards its skilful, near-term forecasts. As first step forward, a further assessment of our findings will be provided by the CMIP5 multi-model decadal hindcasts.

5. Acknowledgements

This work has been possible thanks to the Spanish projects TRACS (CGL2009-10285) and MULCLIVAR (CGL2012-38923-C02-01), and the EU-funded FP7 QWeCI project (ENV-2009-1-243964). I. Polo has been supported by a postdoctoral fellowship funded by the Spanish Government. J. García-Serrano and F. J. Doblas-Reyes have been supported by the EU-funded FP7 NACLIM (ENV-308299) and SPECS (ENV-308378) projects, respectively.

6. References

BREUGEM W.-P., W. HAZELEGER & R. J. HAARSMA (2006). Multimodel study of tropical Atlantic variability and change, *Geophys. Res. Lett.*, 33, L23706, doi: 10.1029/2006GL027831.

CARTON, J. A., X. CAO, B. S. GIESE & A. M. Da SILVA (1996). Decadal and interannual SST variability in the tropical Atlantic Ocean. *J. Phys. Oceanogr.*, 26, 1165-1175.

CHANG P. R. ZHANG, W. HAZELEGER, C. WEN, X WAN, K JI, R. HAARSMA, W. P. BREUGEM & H. SEIDEL (2008). Oceanic link between abrupt changes in the North Atlantic ocean and the African monsoon. *Nature Geosci.*, 1, 444-448.

CHANG C.Y., J.A. CARTON, S.A. GRODSKY & S. NIGAM (2007). Seasonal climate of the tropical Atlantic sector in the NCAR Community Climate System Model 3: error structure and probable causes of errors. *J. Climate*, 21, 1053-1070.

COLLINS M., S. F. B. TETT & C. COOPER (2001). The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments. *Clim. Dyn.*, 17, 61-81.

CORTI S., A. WEISHEIMER, T. N. PALMER, F. J. DOBLAS-REYES & L. MAGNUSSON (2012). Reliability of decadal predictions. *Geophys. Res. Lett.*, 39, doi: 10.1029/2012GL053354.

DOBLAS-REYES, F.J., A. WEISHEIMER, T.N. PALMER, J.M. MURPHY & D.M. SMITH (2010). *Forecast quality assessment of the ENSEMBLES seasonal-to-decadal Stream 2 hindcasts*. European Centre for Medium-Range Weather Forecasts, Reading, U.K., Tech. Memo. 621, 45pp.

DOBLAS-REYES, F.J., I. ANDREU-BURILLO, Y. CHIKAMOTO, J. GARCÍA-SERRANO, V. GUEMAS, M. KIMOTO, T. MOCHIZUKI, L.R.L. RODRIGUES & G.J. VAN OLDENBORGH (2013). Initialized near-term regional climate change prediction. *Nature Comms.*, doi: 10.1038/ncomms2704.

GARCÍA-SERRANO, J., T. LOSADA, B. RODRÍGUEZ-FONSECA & I. POLO (2008). Tropical Atlantic variability modes (1979-2002). Part II: time-evolving atmospheric circulation related to SST-forced tropical convection. *J. Clim.*, 24, 6467-6497.

GARCÍA-SERRANO, J. & F.J. DOBLAS-REYES (2012). On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Clim. Dyn.*, 39, 2025-2040.

GARCÍA-SERRANO, J., F.J. DOBLAS-REYES & C.A.S. COELHO (2012). Understanding Atlantic multi-decadal variability prediction skill. *Geophys. Res. Lett.*, 39, L18708.

GARCÍA-SERRANO, J., F.J. DOBLAS-REYES, R.J. HAARSMA & I. POLO (2013). Decadal prediction of the dominant West African monsoon rainfall modes. *J. Geophys. Res.*, 118, 5260-5279.

GODDARD, L., A. KUMAR, A. SALOMON, D. M. SMITH, G. BOER, P. GONZALEZ, V. KHARIN, W. MERRYFIELD, C. DESER & S.J. MASON (2013). A verification framework for interannual-to-decadal prediction experiments. *Clim. Dyn.*, 40, 245-272.

HAARSMA RJ, E. CAMPOS, W HAZELEGER & C SEVERIJNS (2008). Influence of the meridional overturning circulation on tropical Atlantic climate and variability. *J. Climate*, 21, 1403-1416.

HAZELEGER, W., V. GUEMAS, B. WOUTERS, S. CORTI, I. ANDREU-BURILLO, F.J. DOBLAS-REYES, K. WYSER & M. CAIAN (2013). Multi-year climate prediction using two initialisation strategies. *Geophys. Res. Lett.*, doi:10.1002/grl.50355.

HO C. K., E. HAWKINS, L. SHAFFREY, J. BROCKER, L. HERMANSON, J. M. MURPHY & D. M. SMITH (2013). Variation in the reliability of ensemble predictions of sea surface temperaturas from seasonal to decadal timescales. *Geophys. Res. Lett*. (in press)

JOLIFFE, I.T. & D.B. STEPHENSON (2003). *Forecast verification. A practitioner's guide in atmospheric science*. Wiley, New York.

KEENLYSIDE N. & M. LATIF (2007). Understanding equatorial Atlantic interanual variability. *J. Climate*, 20, 131-142.

LARGE W.G. & G. DANABASOGLU (2006). Attribution and impacts of upper ocean biases in CCSM3. *J. Climate*, 19, 2325-2346.

LOSADA, T., B. RODRÍGUEZ-FONSECA, S. JANICOT, S. GERVOIS, F. CHAUVIN & P. RUTI (2010). A multi-model approach to the Atlantic equatorial mode: impact on the West African monsoon. *Clim. Dyn.*, 35, 29-43.

MEEHL, G.A. & co-authors (2013). Decadal climate prediction: an update from the trenches. *Bull. Amer. Meteorol. Soc.*, doi: 10.1175/BAMS-D-12-00241.1.

POLO, I., B. RODRIGUEZ-FONSECA, T. LOSADA & J. GARCÍA-SERRANO (2008). Tropical Atlantic variability modes (1979-2002). Part I: Time-evolving SST modes related to West African rainfall. *J. Climate*, 21, 6457-6475, doi: 10.1175/2008JCLI2607.1.

POLO, I., B.W. DONG & R.T. SUTTON (2013). Changes in Tropical Atlantic interannual variability from a substantial weakening of the meridional overturning circulation. *Clim. Dyn*., doi: 10.1007/s00382-013-1716-x.

RICHTER, I., S.P. XIE, S K. BEHERA, T. DOI, Y. MASUMOTO (2012). Equatorial Atlantic variability and its relation to mean state biases in CMIP5. *Clim. Dyn.* , doi: 10.1007/s00382-012-1624-5

RICTHTER, I., S.P. XIE, A.T. WITTENGERG & Y. MASUMOTO (2011). Tropical Atlantic biases and their relation to surface wind stress and terrestrial precipitation, *Clim. Dyn*., doi:10.1007/ s00382-011-1038-9.

RICHTER, I., & S. P. XIE (2008). On the origin of equatorial Atlantic biases in coupled general circulation models. *Clim. Dyn*., 31, 587-598, doi: 10.1007/s00382-008-0364-z.

ROBSON J. I., R. T. SUTTON & D. M. SMITH (2012). Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. *Geophys. Res. Lett*., 39, doi: 10.1029/2012GL053370

RODRIGUEZ-FONSECA B., S. JANICOT, E. MOHINO, T. LOSADA, J. BADER, C. CAMINADE, B. FONTAINE, J. GARCÍA-SERRANO, S. GERVOIS, M. JOLY, I. POLO, P. RUTI, P. ROUCOU & A. VOLDOIRE (2011). Interannual and decadal SST-forced responses of the West African monsoon, *Atm Sci. Lett*., 12, 67-74, doi: 10.1002/asl.308

RUIZ-BARRADAS, A., J. A. CARTON & S. NIGAM (2000). Structure of interannual-to-decadal climate variability in the tropical Atlantic sector, *J. Climate*., 13, 3285 - 3297.

SMITH, D.M., S. CUSACK, A.W. COLMAN, C.K. FOLLAND, G.R. HARRIS & J.M. MURPHY (2007). Improved surface temperature prediction for the coming decade from a global climate model. *Science*, 317, 796-799.

SMITH, D.M., R. EADE, N.J. DUNSTONE, D. FEREDAY, J.M. MURPHY, H. POHLMANN & A.A. SCAIFE (2010). Skilful multi-year predictions of Atlantic hurricane frequency. *Nature Geosci.*, 3, 846-849.

SMITH, D.M., R. EADE & H. POHLMANN (2013). A comparison of full-field and anomaly initialization for seasonal to decadal prediction. *Clim. Dyn.*, doi: 10.1007/s00382-013-1683-2.

SMITH, T.M., R.W. REYNOLDS, T.C. PETERSON & J. LAWRIMORE (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *J. Clim.*, 21, 2283-2296.

TAYLOR, K.E., R.J. STOUFFER & G.A. MEEHL (2012). An overview of CMIP5 and the experimental design. *Bull. Amer. Meteorol. Soc.*, 93, 485-498.

TOKINAGA, J., S.P. XI (2011). Weakening of the equatorial Atlantic cold tongue over the past six decades. *Nature Geosci.*, 4, 222-226.

TONIAZZO T., & S. WOOLNOUGH (2013). Development of warm SST errors in the southern tropical Atlantic in CMIP5 decadal hindcasts. *Clim. Dyn*. (in press)

TRENBERTH, K.E. & D.J. SHEA (2006). Atlantic hurricanes and natural variability in 2005. *Geohys. Res. Lett.*, 33, L12704.

VON STORCH, H. & F.W. ZWIERS (2001). *Statistical analysis in climate research*. Cambridge University Press, UK.

WAHL S., M. LATIF, W. PARK & N. KEENLYSIDE (2011). On the Tropical Atlantic SST warm bias in the Kiel Climate Model. *Clim Dyn*. 36, 891-906, doi: 10.1007/s00382-009-0690-9.

ZEBIAK, S.E. (1993). Air-sea interaction in the equatorial Atlantic region. *J. Clim.*, 6, 1567-1586.

ZIEBA, A. (2010). Effective number of observations and unbiased estimators of variance for autocorrelated data - an overview. *Metrol. Meas. Syst.*, 1, 3-16.