



HAL
open science

Dependence of the climate prediction skill on spatiotemporal scales: Internal versus radiatively-forced contribution

D. Volpi, J. Doblas-Reyes, Javier García-Serrano, V. Guemas

► **To cite this version:**

D. Volpi, J. Doblas-Reyes, Javier García-Serrano, V. Guemas. Dependence of the climate prediction skill on spatiotemporal scales: Internal versus radiatively-forced contribution. *Geophysical Research Letters*, 2013, 40 (12), pp.3213 - 3219. 10.1002/grl.50557 . hal-01495328

HAL Id: hal-01495328

<https://hal.science/hal-01495328>

Submitted on 10 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Dependence of the climate prediction skill on spatiotemporal scales: Internal versus radiatively-forced contribution

D. Volpi,^{1,2} F. J. Doblas-Reyes,^{1,3} J. García-Serrano,^{1,4} and V. Guemas^{1,5}

Received 19 March 2013; revised 10 May 2013; accepted 10 May 2013; published 19 June 2013.

[1] This article aims at quantifying the improvement in climate prediction skill as a function of temporal (from monthly to decadal) and spatial scales (from grid point to global) when initializing a perturbed parameter ensemble of the Hadley Centre Climate Model. The focus is on near-surface temperature and precipitation in the Tropical band, the Northern and Southern hemispheres. For temperature, the forecast system reproduces the dominant impact of the external forcing at global spatial scale and at decadal time scales. There are significant improvements with initialization for the first 40 forecast months in the global and tropical domains. In the Northern (Southern) hemisphere, the initialization increases the skill in the first 12 (20) months on regional but not hemispheric scales. The initialization has a stronger impact in the model variants with a weaker global-mean temperature trend. For precipitation, the initialization corrects the negative correlation found at global and tropical scales. **Citation:** Volpi, D., F. J. Doblas-Reyes, J. García-Serrano, and V. Guemas (2013), Dependence of the climate prediction skill on spatiotemporal scales: Internal versus radiatively-forced contribution, *Geophys. Res. Lett.*, 40, 3213–3219, doi:10.1002/grl.50557.

1. Introduction

[2] Providing climate change information at interannual time scales that could be useful to the agriculture, energy, or health sectors is one of the main societal priorities. Decadal predictions aim at satisfying such urgent demand, focusing on time scales of several years to a few decades. While climate change projection focuses on reproducing the long-term trend in climate variables, decadal prediction also aims at modeling the low-frequency variability superimposed on any radiatively forced climate change [Meehl *et al.*, 2009]. The potential predictability can be quantified by the ratio of the temporal variability filtered over decadal timescales, over the total variability. Regions

exhibiting potential predictability indicate where there is a chance to find predictive skill at such timescales. Because of the large heat-storage capacity of the ocean and its slow release, most of such low-frequency variabilities come from the ocean and are driven by different mechanisms. Decadal potential predictability is found over the oceans at mid to high latitudes [Boer and Lambert, 2008]. Previous works have shown that the Atlantic meridional overturning circulation is potentially predictable a decade in advance, and it has also been shown (through a perfect model study of potential predictability) that in the North Atlantic, many variables such as sea surface temperature (SST), salinity, heat content, or meridional transport could be potentially predictable for many years [Griffies and Bryan, 1997a, 1997b]. In the Pacific, the model-based study from Branstator *et al.* [2012] shows that the decadal variability observed in the North Pacific sea surface temperature (SST) is pronounced but not necessarily predictable. Guemas *et al.* [2012] rather incriminate model deficiencies in the inability to predict the North Pacific SST.

[3] The study of potential predictability has led to the attempt of producing real predictions. Smith *et al.* [2007], Keenlyside *et al.* [2008], Pohlmann *et al.* [2009], and Mochizuki *et al.* [2010] obtained good results in the North Atlantic and North Pacific. In climate predictions, the following sources of uncertainty compromise the forecast quality [Robson, 2010]:

[4] 1. Internal variability: the natural (and not externally radiatively forced) variability of the climate system.

[5] 2. Model inadequacy: the parametrization of the physical processes is a source of uncertainty as the estimation of the parameters introduces errors into the model. Moreover, some processes are not even simulated because they are not known yet.

[3] 3. Scenario uncertainty: uncertainties due to the unforeseeable evolution of socioeconomic conditions, which influence the change in greenhouse gas emissions.

[7] The predictability of the internal variability is associated with information contained in the initial conditions. It is measured by determining for how long the predicted distribution of an ensemble of similar initial states is distinguishable from the climatological distribution [Teng and Branstator, 2010]. While weather and interannual climate predictions attempt to address this source of uncertainty, climate change projections do not. The relative importance of the initial conditions in climate prediction is supposed to vary with the time scale and has been assumed to be a continuous function that decreases with forecast time, becoming negligible after several decades [Hawkins and Sutton, 2009]. It has been shown that there is skill beyond the first forecast year and that the quality of the information about the initial state can improve the climate forecasts in different regions [Doblas-Reyes *et al.*, 2013].

Additional supporting information may be found in the online version of this article.

¹Institut Català de Ciències del Clima (IC3), Barcelona, Spain.

²School of Mathematical and Physical Sciences, University of Reading, Reading, UK.

³Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

⁴Laboratoire d'Océanographie et du Climat (LOCEAN-IPSL), Université Pierre et Marie Curie (UPMC), Paris, France.

⁵Centre National de Recherches Météorologiques, Groupe d'Etude de l'Atmosphère Météorologique Météo-France, Toulouse, France.

Corresponding author: D. Volpi, Institut Català de Ciències del Clima (IC3), Doctor Trueta 203, Barcelona, 08005, Spain. (danila.volpi@ic3.cat)

[8] This work aims at comparing the predictive skill for the near-surface temperature and precipitation of an experiment initialized with observed data and one with no initialization, both experiment accounting for the variable radiative forcing. The reader has to distinguish that this is not a predictability study in the sense of *Hawkins and Sutton* [2009], but instead is an analysis of the actual predictive skill. The following objectives have been addressed:

[9] 1. the spatial and time scales at which maximum skill is found, and

[10] 2. a quantification of the skill improvement provided by the initialization

[11] Section 2 presents the details of the data and the method implemented. Section 3 shows the results for near-surface temperature and precipitation, and section 4 aims at drawing the conclusions and suggests open issues that could lead to further work.

2. Data and Method

[12] The decadal hindcasts employed are from the perturbed-parameter ensemble of the MetOffice Decadal Prediction System (DePreSys_PP). DePreSys [Smith et al., 2007; Robson, 2010] is based on the global coupled ocean-atmosphere model HadCM3 [Gordon et al., 2000]. The perturbed-parameter ensemble of DePreSys_PP is composed of eight model variants with simultaneous perturbations to 29 atmosphere and sea-ice parameters [Murphy et al., 2004], plus the standard model version. The selection of the variants guarantees an approximately uniform sample of climate sensitivity and a wide range of different parameter settings to sample the model uncertainty.

[13] The atmospheric resolution is $2.5^\circ \times 3.75^\circ$ with 19 vertical levels, while the ocean component has a resolution of $1.25^\circ \times 1.25^\circ$ with 20 vertical levels.

[14] The decadal hindcasts consist in a set of 10 yearlong retrospective forecasts, starting every November from 1960 until 2005. Here two different experiments of the same forecast system have been compared: NoAssim decadal hindcasts are initialized from nine transient simulations with information about greenhouse gases, tropospheric and stratospheric ozone concentration, and sulfur emissions taken from observations. The volcanic aerosol load is damped with a 1 year e -folding time. The variability in solar radiation is represented by repeating the previous 11 year solar cycle. This gives confidence on the reliability of the operational decadal forecast system, unlike the Fifth Coupled Model Intercomparison Project (CMIP5, [Taylor et al., 2012]) hindcasts that prescribe observed volcanic aerosols and solar irradiance along the predictions. From the same transient runs with identical external forcings, the Assim experiment is initialized by assimilating atmosphere observations of horizontal winds, temperature and surface pressure, and ocean observations of temperature and salinity.

[15] The study has been carried out considering the following domains: the global (poles excluded, i.e., 60°N – 60°S), the Northern hemisphere (NH, poles excluded, i.e., 20°N – 60°N), the Southern hemisphere (SH, poles excluded, i.e., 20°S – 60°S), and the Tropical band (TRO, 20°S – 20°N). The skill measure used here is the correlation computed along the space and time dimensions. For the temperature, the reference data are GHCN [Fan and van den Dool, 2008] for land and ERSST for ocean [Smith et al., 2008],

both available until early 2010 at the time of the study. For precipitation, the reference is the CRU data [Brohan et al., 2006], which is available over land until the end of 2006.

[16] The anomalies at each grid point and for each one of the 120 time steps of the hindcasts have been computed using the per-pair method [García-Serrano and Doblas-Reyes, 2012] in which the computation of the lead-time-dependent climatology accounts only for the years in which both observational and model data are available. Depending on the reference data available and in order to guarantee the same validation sample at all forecast times (i.e., same amount of reference data available at all forecast time), the analysis employs the start dates included in the range 1960–1999 (hindcast starting in November 1999 and finishing in October 2009, when the observation were still available) for temperature and 1960–1996 (hindcast starting in November 1996 and finishing in October 2006, when the observation was still available) for precipitation.

[17] The forecast-time accumulation is performed by accumulating data for consecutive forecast months, up to accumulating the whole forecast period:

$$x_{1,i,j}; \dots; \frac{x_{1,i,j} + x_{2,i,j} + \dots + x_{t,i,j}}{t}; \frac{x_{1,i,j} + x_{2,i,j} + \dots + x_{120,i,j}}{120}$$

where $x_{t,i,j}$ is the ensemble mean anomaly at forecast time t [$1;120$], latitude $i \in [\text{min latitude of domain}; \text{max latitude of domain}]$ and longitude $j \in [\text{min longitude of domain}; \text{max longitude of domain}]$. The ensemble mean is the average of the anomalies obtained for each model version.

[18] In order to illustrate the skill dependence with the spatial scale, between grid point level and global average, the immediate neighbors at each grid point have been averaged along all the possible directions. The case of zero neighbors is the original grid, and the maximum amount of neighbors represents the area average of the domain defined by its latitudinal extension. Successively, the temporal variances and covariances between model and observed anomalies have been computed for each spatial averaging. Finally, the correlation has been computed from the spatial average of variance and covariance values:

$$\rho_g = \frac{(\overline{\text{Cov}_{xy}})_g}{\sqrt{(\overline{\text{Var}_x})_g \cdot (\overline{\text{Var}_y})_g}}$$

where ρ_g is the correlation for the degree of spatial averaging g , $(\overline{\text{Cov}_{xy}})_g$ is the temporal covariance between the model (x) and the reference data (y), averaged over the spatially averaged data g ; analogously $(\overline{\text{Var}_x})_g$ is the temporal variance of the model averaged with the spatial averaging g , and $(\overline{\text{Var}_y})_g$ is the temporal variance of the reference data averaged over the spatial averaging g .

[19] A Student's t -test has been applied, and the p -values 0.01, 0.05, and 0.1 have been calculated and plotted. As the number of degrees of freedom depends on the number of independent data in time and space, a time and space dependency has been arbitrarily chosen for our study. A time dependence between 10 consecutive start dates has been considered, which corresponds to 10 years (the Atlantic multidecadal oscillation, for example, can stay in the same phase for more than 10 years). An area dependency of $n_{pt} = 5 \times 5$ grid points that approximately corresponds to an area of $2000 \text{ km} \times 2700 \text{ km}$ has been considered. A smaller

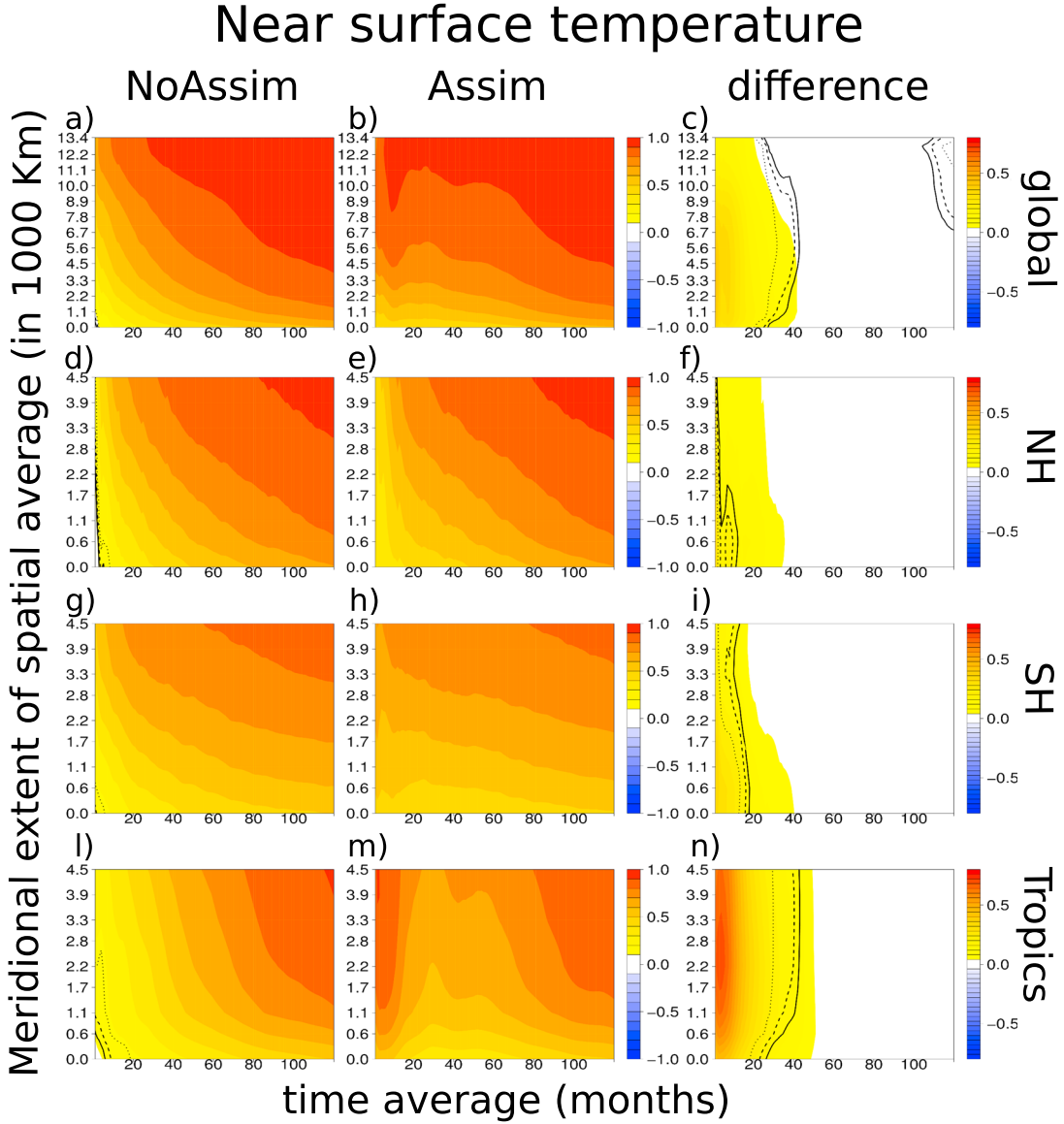


Figure 1. Correlation of Assim and NoAssim near-surface temperature with GHCN and ERSST data depending on the forecast time and the spatial averages. (a, b, and c) The global domain of, respectively, Assim and NoAssim experiments and their difference. Analogously, (d, e, and f) the Northern Hemisphere domain, (g, h, and i) the Southern Hemisphere, and (l, m, and n) the Tropical band. Note that the color bar for the third column is different from the others. The x axis indicates the number of consecutive months included in the average (from 1 up to the whole forecast period of 120 months). The y axis represents the meridional extent of the spatial average at which the variances and the covariances between model and observations are computed. From those values, the area average has been computed. The correlation has been calculated from the area average of the variances and covariances. The one-tailed Student's t -test has been computed, and the black dotted, dashed, and solid contours are, respectively, the p -values 0.01, 0.05, and 0.1. See details on the computation of the number of degrees of freedom in the text.

area dependence has been considered for the precipitation, and an $n_{pt}=3 \times 3$ grid points corresponding to approximately $1250 \text{ km} \times 1650 \text{ km}$ has been arbitrarily chosen. The degrees of freedom for each spatial averaging g are then calculated as follows:

$$\text{DOF}_g = \text{indep}_{sd} \cdot \min\left(\frac{nlat \cdot nlon}{n_{pt}}, nlat_g \cdot nlon_g\right)$$

where indep_{sd} is the number of independent start dates, $nlat$ and $nlon$ are the number of grid points in the original grid,

and $nlat_g$ and $nlon_g$ are the number of grid points for the degree of spatial averaging g . For precipitation, this quantity is multiplied by the proportion of land data, as there is precipitation observational data available only over land.

3. Results

[20] Figure 1 shows the 2 m temperature correlation between the decadal hindcasts and the reference data as a function of forecast time and spatial averaging. Each row

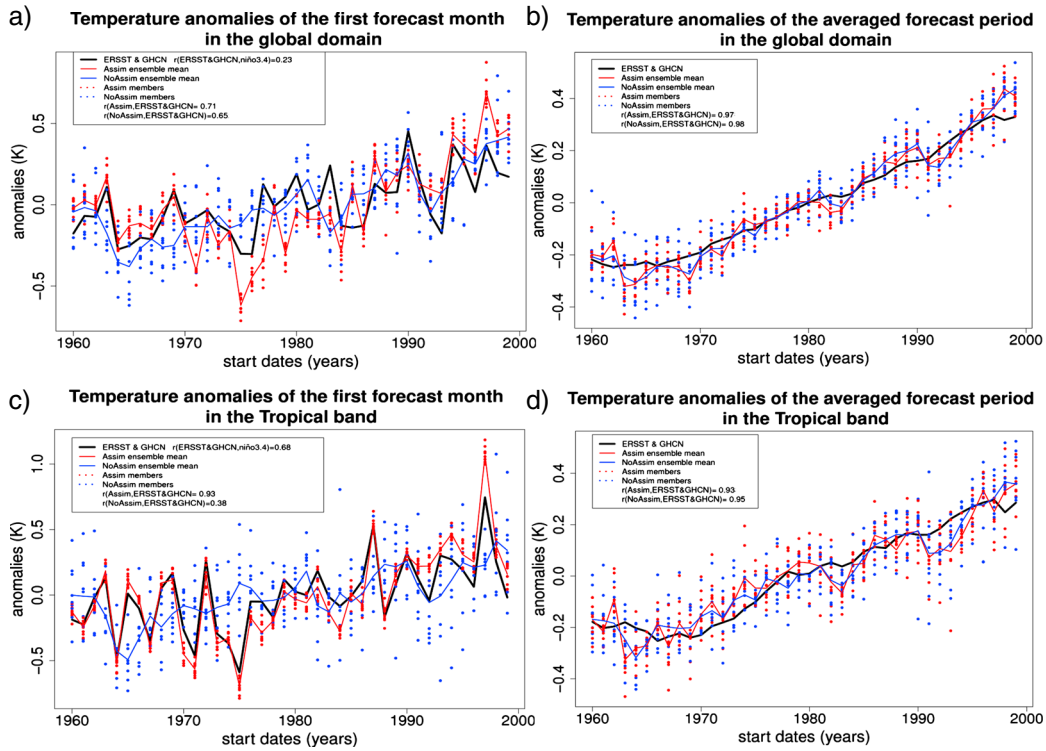


Figure 2. Anomaly time series of Assim and NoAssim near-surface temperature with GHCN and ERSST in the (a,b) global domain and (c,d) tropical band. Figure 2a shows the time series of the first forecast month with a spatial average over the whole global domain; Figure 2b is the anomalies time series of the 120 forecast months average over the whole global domain. In solid black are the reference data. Assim ensemble mean anomalies are in solid red, while the Assim ensemble members in dotted red. Analogously, the NoAssim ensemble mean anomalies are in solid blue, and the NoAssim ensemble members in dotted blue. Figures 2c and 2d are the same as Figures 2a and 2b, but in tropical band. The year in the abscissa corresponds to the start date of each individual hindcast. Note that the legend indicates the correlation of Assim with the reference data and the correlation of NoAssim with the reference data. In Figures 2a and 2c, the correlation between the observed Niño 3.4 SST index with the tropically averaged reference data is also shown.

represents a different domain, the global, the Northern hemisphere (NH), the Southern hemisphere (SH), and the Tropical band, respectively. The first column shows the skill of the NoAssim hindcasts, the second column the Assim skill, and the third one is the skill improvement of Assim over NoAssim. The first column of Figure 1 (NoAssim experiment near-surface temperature skill) shows that when the model is not initialized with observations, the skill grows nearly monotonically with spatial and time averaging. This leads to a maximum of skill regardless of the domain considered at the top right corner of the figure, where the spatial averaging and temporal accumulation are the largest (i.e., the whole domain is averaged, and the time accumulation includes all 120 forecast months), which suggests an increasing role of the varying forcing as the time series are smoothed. The smaller domains show that the NoAssim skill is larger in the extratropics, pointing toward a larger relative contribution to the climate signals of the external radiative forcings over the internal variability in the extratropics compared to the tropics, particularly for the 1 to 5 year time scale.

[21] Assim shows an additional maximum of skill in the global domain at the beginning of the forecast centered at the accumulated month 10 (Figure 1b). In the NH and SH of the initialized experiment (Figures 1e and 1h), there is also an increase in skill with space and time accumulation without the peak that appeared in the global domain at the beginning

of the forecast accumulation. The skill of Assim in the Tropical band (Figure 1m) shows in addition to the maximum at large spatial averaging, a peak for short accumulation forecast times. This peak seems to originate from the added value of the initialization associated with the El Niño Southern Oscillator (ENSO) as illustrated in Figure 2, which is explained in detail below. Figure 1c shows that the improvement given by the initialization in the global domain is significant for the accumulation of the first 40 forecast months. Moreover, the maximum improvement due to the initialization appears at intermediate spatial scales, which is an additional motivation to analyze the skill in smaller domains such as the NH, the SH, and the Tropical band.

[22] Comparing the skill of the NH and SH (Figures 1d and 1g) for the NoAssim experiment, SH shows less skill. Also, for the Assim experiment (Figures 1e and 1h), there are less steep contours in SH. When looking at the initialization improvement (Figures 1f and 1i), significant results are shown for the first 12 accumulated months in the NH and 18 accumulated months in SH. At the larger scale spatial averaging, which corresponds to the whole NH/SH domain (top of Figures 1f and 1i, respectively), the improvement is not significant at the 90% confidence level. Figures S1 and S2 in the supporting information show the results of the study using land-only and ocean-only data, respectively. The maximum skill in the GLO, NH, and SH is lower than that in

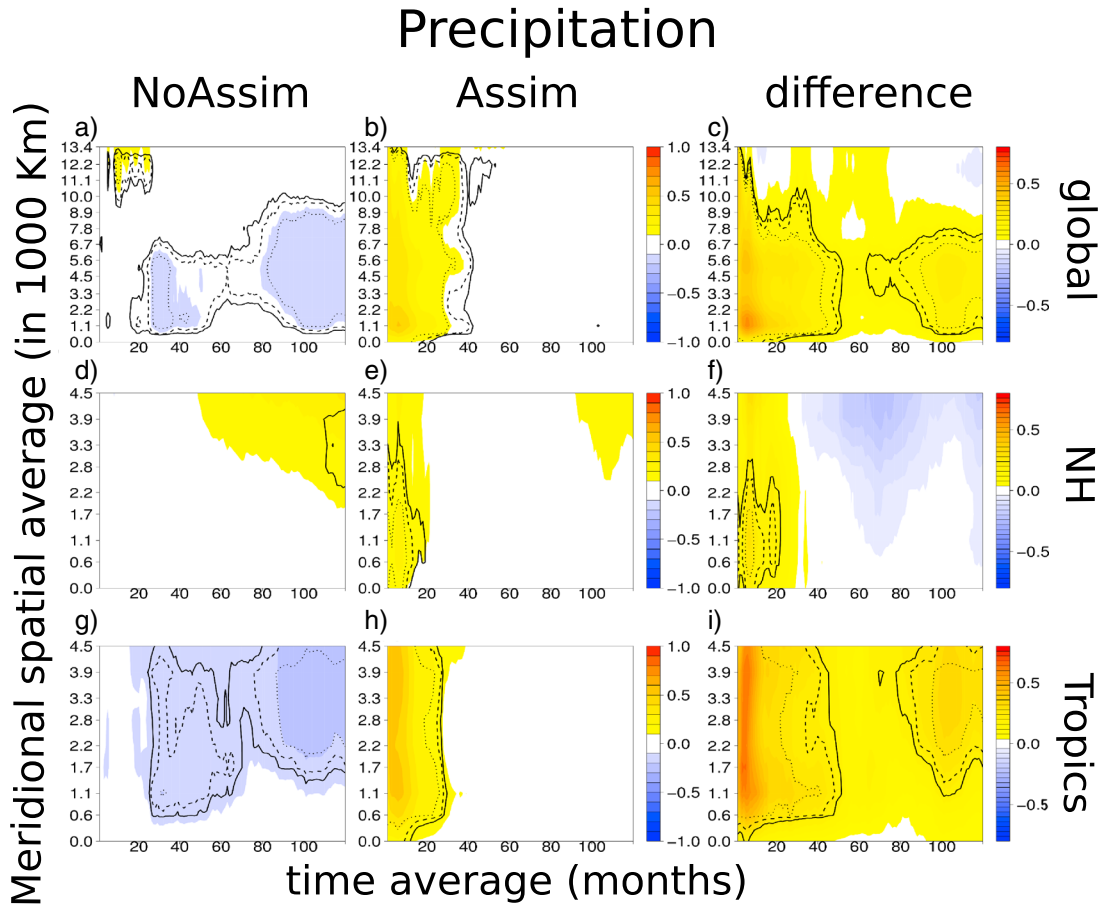


Figure 3. Correlation of Assim and NoAssim precipitation with CRU data depending on the forecast time and the spatial averages. (a, b, and c) The global domain of, respectively, Assim and NoAssim experiments, and their difference. Analogously, (d, e, and f) the Northern Hemisphere domain, and (g, h, and i) the Tropical band. Note that the color bar for the third column is different from the others. The x axis indicates the number of consecutive months included in the average (from 1 up to the whole forecast period of 120 months). The y axis represents the meridional extent of the spatial average at which the variances and the covariances between model and observations are computed. From those values, the area average has been computed. The correlation has been calculated from the area average of the variances and covariances. The one-tailed Student's t -test has been computed, and the black dotted, dashed, and solid contours are, respectively, the p -values 0.01, 0.05, and 0.1. Details on the computation of the number of degrees of freedom in the text.

Figure 1, especially over land, due to the fact that the effective averaged areas are smaller than the one when both ocean and land are considered together. The monotonic increase in skill with spatial and temporal averaging observed in Figure 1 is lost in Figures S1 and S2. When looking at the difference between Assim and NoAssim, the ocean-only figure (Figure S2) shows a statistically significant (confidence level of 90%) impact of the initialization for longer time scales when compared to the mixed ocean and land data. To better identify the different origins of the skill for the Assim and NoAssim, Figure 2 illustrates the temperature anomalies in different domains (respectively, the global domain in the first row and the tropical domain in the second row) and at different forecast time accumulations (first forecast month in the first column, and for the accumulation of all forecast months in the second column). All the grid points of each domain are averaged in all panels. In particular, Figure 2a corresponds to the temperature anomalies at the first forecast month for the 46 start dates used in this study. The anomaly correlation coefficients generated by those anomalies correspond to the top left points in Figures 1a (NoAssim) and 1b (Assim).

Similarly, Figure 2b illustrates the temperature anomalies of the averaged forecast period, and the corresponding anomaly correlation coefficient is represented by the top right corner points of Figures 1a and 1b. Analogously, the top left and top right anomaly correlation coefficients in Figures 1l and 1m are generated, respectively, by the anomalies illustrated in Figures 2c and 2d. Figure 2a shows that NoAssim ensemble has more spread than Assim (this is also shown in Figures 2c and 2d for the tropical domain discussed below). When integrating over time, the role of the radiative forcings gets dominant, and both Assim and NoAssim have a correlation with the reference data higher than 0.9 (Figure 2b).

[23] Compared to other domains, the Tropical band of NoAssim (Figure 1l) has the lowest skill. The skill figure of Assim in the Tropical band (Figure 1m) shows, in addition to the maximum at global average, a peak at the beginning of the forecast accumulation. The maximum improvement of the initialization is given during the accumulation of the first 4 forecast months (Figure 1n), while significant improvements are shown until 40 forecast months accumulation, and improvements last up to the first 50 forecast months

accumulation. Such improvement at the beginning of the forecast is also shown in Figure 2. Figure 2c shows the temperature anomalies of the first forecast month when averaging over the whole tropical band. Assim has a correlation of 0.94, while NoAssim has a correlation of 0.38. These correlations are significantly different with at least 99% confidence level. NoAssim reflects the continuous warming without showing any variability consistent in time with the reference data. This is mainly due to the fact that NoAssim does not have any information about the phase and amplitude of contemporaneous events, while Assim does. The correlation of the observed Niño 3.4 SST index with the tropically averaged reference data is of 0.69, which also suggests that the variability in the Tropical band is dominated by the variability of ENSO. Figure S3a of the supporting information shows the Niño 3.4 SST index anomalies for the first forecast month. Assim has a correlation with ERSST of 0.99, while NoAssim has a negative correlation.

[24] Similar to Figure 2b for the global domain, in Figure 2d for the Tropics, when the forecast time is accumulated over the 120 forecast months, both experiments have a correlation with the observations higher than 0.9 due to the large role of the external forcings. The results suggest that most of the skill improvement with initialization found in the global domain is associated with the tropical region. Figure S3b of the supporting information shows that with the accumulation of the whole forecast period, the Niño 3.4 SST index skill decreases with respect to that shown in Figure S3a. As a result, Assim and NoAssim have similar skill of around 0.55. Figure S3c shows how the Niño 3.4 correlation difference evolves as the forecast-time accumulation increases, with the situations in Figures S3a and S3b being the two extremes. The difference in correlation decreases with forecast-time accumulation. During the first forecast year, the correlation difference is greater than 1, which means that the initialization actually corrects the sign of the NoAssim correlation. Moreover, the correlation is significantly different at 90% level for accumulations of up to 40 forecast months, which explains the results found for the TRO in Figure 1.

[25] The results described above are consistent across the different model version used. Figure S4 shows the skill in the global domain for each individual version of the forecast system. When using the linear trend of the global-mean temperature as a proxy for the climate sensitivity, it was found that when the trend is stronger, the NoAssim experiment (first column Figure S4) has a larger skill at all spatial and temporal scales. Moreover, the Assim panels (second column figure S4) show that the impact of initialization is stronger when the trend is weaker. The model versions in Figure S4 are ranked following the climate sensitivity estimates shown in Figure S5, from the model version with higher climate sensitivity to the one with the lowest.

[26] The NoAssim precipitation (first column Figure 3) has negative correlation with the CRU data in the global domain (Figure 3a) and the Tropical band (Figure 3g). Some positive skill is found in the NH (Figure 3d) at large-scale averaging and full forecast-time accumulation. The correlation with CRU of the initialized experiment displays significantly positive skill (confidence level of 90%) in every domain (second column Figure 3) for an accumulation of up to 20 forecast months. The maximum skill is found in the Tropics (Figure 3h) for an accumulation of 12 forecast months. Improvement with the initialization is found in the global

domain (Figure 3c) for all forecast-time accumulations and spatial scales averaging, except for the averaging of the largest spatial scales. These improvements are the strongest over the TRO region (Figure 3i) and are mainly due to the correction of the NoAssim negative skill by the initialization.

4. Conclusions

[27] In this work, the improvements associated with the initialization of a decadal forecast system have been quantified. Moreover, it has been documented how the added value of the initial condition information varies with the temporal and spatial scales.

[28] 1. For near-surface temperature, it has been found that when increasing the spatial scales and temporal accumulation, the external forcing influence becomes more important. DePreSys_PP correctly reproduces the surface temperature response to the variation of external forcings. This leads to a maximum of skill at time scales of 10 accumulated years, which is the maximum time scale considered in this study, and for regional to large scales. This is not the case for land precipitation, for which the sign of the correlation with the observations is negative.

[29] 2. By introducing information of the state of the climate system through the initialization, a new peak of surface temperature skill appears from the beginning of the forecast to the first 40 forecast months in the global domain and the tropical band. This seems to be due to the correct prediction of ENSO, which is usually considered limited to 1 year. Globally, the skill improvements due to the initialization are mainly coming from the Tropics.

[30] 3. The NH is more skillful than the SH for near-surface temperature. The improvements brought by the initialization are statistically significant with a confidence level of 90%, respectively, in the NH for the accumulation of the first 12 forecast months and in the SH for the first 20 forecast months and from the grid point scale to the regional spatial scale. The reason for not getting statistically significant results when averaging over the whole hemisphere might be due to the small number of independent data available. A longer reforecast period would be necessary to get more robust results.

[31] 4. The skill results for near-surface temperature are consistent across the different model versions. When using the linear trend of the global-mean temperature as a proxy for the climate sensitivity, it was found that in the NoAssim experiment, the stronger the trend, the larger the skill, at all spatial and temporal scales. The Assim experiment shows that the impact of initialization is stronger when the trend is weaker.

[32] 5. NoAssim has no precipitation skill in the global domain and the Tropics at almost any spatial and time scale. The initialization corrects the negative sign of the NoAssim correlation and has a beneficial impact for all time scales and spatial averaging in both domains. Much is still needed to improve multiannual precipitation forecasts, especially considering that precipitation is a key variable with large socioeconomic consequences.

[33] **Acknowledgments.** This work was supported by the EU-funded QWeCI (FP7-ENV-2009-1- 243964), SPECS (FP7-ENV-2012-308378), NAELIM (FP7-ENV-2012-308299), the MICINN-funded RUCSS (CGL2010-20657) projects and the Catalan Government. The authors thankfully acknowledge the computer resources, technical expertise, and assistance provided by the Red Española de Supercomputación (RES).

[34] The Editor thanks one anonymous reviewer for his/her assistance in evaluating this paper.

References

- Boer, G., and S. Lambert (2008), Multi-model decadal potential predictability of precipitation and temperature, *Geophys. Res. Lett.*, *35*, L05706, doi:10.1029/2008GL033234.
- Branstator, G., H. Teng, G. A. Meehl, M. Kimoto, J. R. Knight, M. Latif, and A. Rosati (2012), Systematic estimates of initial-value decadal predictability for six AOGCMs, *J. Clim.*, *25*, 1827–1846.
- Brohan, P., J. Kennedy, I. Harris, S. Tett, and P. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850, *Geophys. Res. Lett.*, *11*, D12106, doi:10.1029/2005JD006548.
- Doblas-Reyes, F., I. Andreu-Burillo, Y. Chikamoto, J. García-Serrano, V. Guemas, M. Kimoto, T. Mochizuki, L. R. L. Rodrigues, and G. J. van Oldenborgh (2013), Initialized near-term regional climate change prediction, *Nat. Commun.*, *4*(1715), doi:10.1038/ncomms2704.
- Fan, Y., and H. van den Dool (2008), A global monthly land surface air temperature analysis for 1948–present, *Geophys. Res. Lett.*, *113*, D01103, doi:10.1029/2007JD008470.
- García-Serrano, J., and F. Doblas-Reyes (2012), On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast, *Clim. Dyn.*, doi:10.1007/s00382-012-1413-1.
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat and ocean heat transports in a version of the hadley centre coupled model without flux adjustments, *Clim. Dyn.*, *16*, 147–168.
- Griffies, S. M., and K. Bryan (1997a), Predictability of North Atlantic multidecadal climate variability, *Science*, *275*(5297), 181–184.
- Griffies, S. M., and K. Bryan (1997b), A predictability study of simulated North Atlantic multidecadal variability, *Clim. Dyn.*, *13*, 459–488.
- Guemas, V., F. Doblas-Reyes, F. Lienert, H. Du, and Y. Soufflet (2012), Identifying the causes for the low decadal climate forecast skill over the North Pacific, *J. Geophys. Res.*, *117*, D20111, doi:10.1029/2012JD018004.
- Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *Bull. Amer. Meteorol. Soc.*, *90*(8), 1095–1107.
- Keenlyside, N., M. Latif, J. Jungclauss, L. Kornblueh, and E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, *453*, 84–88, doi:10.1038/nature06921.
- Meehl, G., et al. (2009), Decadal prediction: Can it be skillful?, *Bull. Amer. Meteorol. Soc.*, *90*, 1467–1485.
- Mochizuki, T., et al. (2010), Pacific decadal oscillation hindcasts relevant to near-term climate prediction, *PNAS*, *107*, 1833–1837, doi:10.1073/pnas.0906531107.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, *430*, 768–772.
- Pohlmann, H., J. Jungclauss, A. Köhl, D. Stammer, and J. Marotzke (2009), Initializing decadal climate predictions with the gecco oceanic synthesis: Effects on the North Atlantic, *J. Clim.*, *22*, 3926–3938.
- Robson, J. I. (2010), Understanding the performance of a decadal prediction system. Ph.D. thesis, University of Reading, dept. of Meteorology, 233 pp.
- Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy (2007), Improved surface temperature prediction for the coming decade from a global climate model, *Science*, *317*, 796–799.
- Smith, T., R. Reynolds, T. Peterson, and J. Lawrimore (2008), Improvements to NOAA’s historical merged land-ocean surface temperature analysis (1880–2006), *J. Clim.*, *21*, 2283–2296.
- Taylor, K., R. Stouffer, and G. Meehl (2012), An overview of CMIP5 and the experimental design, *Bull. Amer. Meteorol. Soc.*, *93*, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Teng, H., and G. Branstator (2010), Initial-value predictability of prominent modes of North Pacific subsurface temperature in a CGCM, *Clim. Dyn.*, doi:10.1007/s00382-010-0749-7.