GEOPHYSICAL RESEARCH LETTERS, VOL. 39, L22707, doi:10.1029/2012GL053326, 2012

Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology

W. A. Müller,[1] J. Baehr,[2] H. Haak,[1] J. H. Jungclaus,[1] J. Kröger,[1] D. Matei,[1] D. Notz,[1]
H. Pohlmann,[1] J. S. von Storch,[1] and J. Marotzke[1]

Received 31 July 2012; revised 19 October 2012; accepted 21 October 2012; published 30 November 2012.

[1] We examine the latest decadal predictions performed with the coupled model MPI-ESM as part of the Coupled Model Intercomparison Project Phase 5 (CMIP5). We use ensembles of uninitialized and yearly initialized experiments to estimate the forecast skill for surface air temperature. Like for its precursor, the initialization of MPI-ESM improves forecast skill for yearly and multi-yearly means, predominately over the North Atlantic for all lead times. Over the tropical Pacific, negative skill scores reflect a systematic error in the initialization. We also examine the forecast skill of multi-year seasonal means. Skill scores of winter means are predominantly positive over northern Europe. In contrast, summer to autumn means reveal positive skill scores over central and south-eastern Europe. The skill scores of summer means are attributable to an observed pressure-gradient response to the North Atlantic surface temperatures.
Citation: Müller, W. A., J. Baehr, H. Haak, J. H. Jungclaus, J. Kröger, D. Matei, D. Notz, H. Pohlmann, J. S. von Storch, and J. Marotzke (2012), Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, *Geophys. Res. Lett.*, *39*, L22707, doi:10.1029/2012GL053326.

1. Introduction

[2] Building on the merits of early research of decadal to multi-decadal predictability, but also because of the recent successes in initialising climate models for decadal climate predictions [e.g., *Smith et al.*, 2007], various modelling groups now perform near-term climate predictions. A suite of decadal hindcast experiments has recently been coordinated by the World Climate Research Program Coupled Model Intercomparison Project Phase 5 (CMIP5) [*Taylor et al.*, 2012] which will enter the fifth assessment report (AR5) of the Intergovernmental Panel on Climate Change (IPCC). Here we use the coupled Earth System model of the Max Planck Institute for Meteorology (MPI-ESM), which contributes to CMIP5, and investigate its prediction skill of surface temperature with a particular focus on the North Atlantic/European region.

[3] Previous studies have identified the North Atlantic as a key region for decadal climate predictions and have shown pronounced forecast skill for different parameters [e.g., *Pohlmann et al.*, 2009; *van Oldenborgh et al.*, 2012; *Matei et al.*, 2012a; *Gangstø et al.*, 2012]. Potential predictability is found for integrated quantities such as the Atlantic Meridional Overturning circulation (AMOC) [*Pohlmann et al.*, 2012] or the subpolar gyre (SPG) [*Matei et al.*, 2012a]. For the AMOC at 26.5°N, where observations are available from 2004 onward, recent model initialization reveals prediction skill for monthly mean AMOC strength up to 4 years in advance [*Matei et al.*, 2012b]. Further, North Atlantic surface temperatures and heat content have been shown to exhibit prediction skill for up to 10 years in advance [e.g. *Pohlmann et al.*, 2009; *Kröger et al.*, 2012] and are suggested to be linked with the heat transport of the overturning circulation [*Matei et al.*, 2012a]. Prediction skill is further assessed for climate impacts, such as for the multi-decadal variability of Atlantic tropical cyclones, which has been considered to origin from the SPG region [*Smith et al.*, 2010; *Dunstone et al.*, 2011].

[4] Typically, the verification of the predictions used for the mentioned studies is applied to yearly or multi-yearly means. However, the variability of parameters, such as surface air temperature (SAT), sea level pressure (SLP) and its underlying processes may change considerably in space and time as a function of season. An example is the local and remote response of atmospheric parameters to multi-decadal ocean-surface variability such as the Atlantic Multidecadal Oscillation (AMO) [e.g., *Sutton and Hodson*, 2005; *Hodson et al.*, 2010]. Here, observed SLP over large areas of North America co-varies with the different phases of the AMO, particularly during summer seasons. As shown by *Hodson et al.* [2010], multi-model experiments consistently confirm an observed low-pressure response across southern North America during summer to imposed North Atlantic warming and associated latent heat release over the Caribbean. In contrast, for winter seasons this co-variability is considerably reduced.

[5] Over the Eurasian continent, a consistent atmospheric response to the AMO is still controversial. A significant impact was found for observed summer SLP and near-surface temperatures [*Hodson et al.*, 2010]. The large-scale atmospheric structure of SLP is similar to the summer signature of the North Atlantic Oscillation (NAO) [*Bladé et al.*, 2011] and is associated with a large impact on SAT in the Mediterranean region [*Mariotti and Dell'Aquila*, 2012]. For winter means, the atmospheric response to multi-decadal variations of North Atlantic SST was suggested to be linked to the winter NAO (for review see *Czaja et al.* [2003]). Recent multi-model studies have addressed the influence of decadal variations of the AMOC on the atmosphere and find a negative phase of the NAO and associated processes such as decreased storm track activity following the positive phase

---

[1] Max Planck Institute for Meteorology, Hamburg, Germany.
[2] Institute of Oceanography, University of Hamburg, Hamburg, Germany.

Corresponding author: W. A. Müller, Max Planck Institute for Meteorology, Bundesstr. 53, DE-20146 Hamburg, Germany. (wolfgang.mueller@zmaw.de)

0094-8276/12/2012GL053326

of the AMOC after 1–10 years [*Gastineau and Frankignoul*, 2011].

[6] Here we examine the forecast skill of decadal predictions performed with MPI-ESM carried out for CMIP5. The skill scores are estimated for surface temperatures for yearly and multi-yearly averages as suggested by the CMIP5 protocol. In addition, given the pronounced seasonality of the observed atmospheric response to North Atlantic sea surface temperatures (SST), we consider multi-year seasonal means to account for the seasonal dependence of the skill. The variations of the skill scores are compared with an accompanying analysis of observed and modelled co-variability of North Atlantic SST and their remote response.

2. Data and Methods

[7] The coupled model MPI-ESM is used to perform historical runs (uninitialized) and decadal hindcast (initialized) experiments. The experiments are performed in a low-resolution configuration (MPI-ESM-LR) with the latest version of the ocean model MPIOM (J. H. Jungclaus et al., Characteristics of the ocean simulation in MPIOM, submitted to *Journal of Advances in Modeling Earth Systems*, 2012) and the atmospheric component ECHAM6 (B. Stevens et al., The atmospheric component of the MPI-M Earth-System-Model: ECHAM6, submitted to *Journal of Advances in Modeling Earth Systems*, 2012). The ocean model is run in a horizontal resolution of 1.5° on average and 40 vertical levels. For ECHAM6 the horizontal resolution is T63 with 47 vertical levels including the upper stratosphere up to 0.1 hPa.

[8] A total of 3 ensemble members each of uninitialized and initialized experiments are performed. The uninitialized experiments are started from a preindustrial control simulation and consider aerosol and greenhouse gas concentrations for the period 1850–2005 and the RCP4.5 scenario thereafter. The initial conditions for the initialized runs are taken from an assimilation experiment using the coupled model. In the assimilation experiments, the model state is nudged towards 3-dimensional daily ocean temperature and salinity anomalies, which are added to the model climatology. Temperature and salinity anomalies are taken from an MPIOM experiment forced with NCEP/NOAA reanalysis [*Kalnay et al.*, 1996] (for details see *Matei et al.* [2012a]). No assimilation of atmospheric parameters is applied. Three-dimensional atmospheric and ocean fields of the assimilation experiment are considered as initial conditions. The hindcasts are always started on January 1st and have a length of 10 years. In addition to the CMIP5-CORE experiments, which consider start dates every five years, we perform yearly initialized hindcasts for the period 1961–2010. The ensemble of the initialized runs is realized by 1-day-lagged initialisation of the assimilation experiment.

[9] For the analysis, anomalies are defined with respect to the 1961–2010 climatology. For the uninitialized runs, the climatology is calculated from the ensemble mean of the uninitialized runs. Anomalies of initialized runs are defined with respect to the climatology of the ensemble mean and lead time. Here a “leave-one-out” cross-validation is applied such that the considered prediction does not appear in the climatology [see also *Gangstø et al.*, 2012]. The predictions are verified with skill scores based on root-mean-squared-error (RMSE) and anomaly correlation coefficient. The RMSE skill score is defined as $SS_{RMSE} = 1 - (RMSE_{INI}/RMSE_{UNINI})$, where $RMSE_{INI}$ is the RMSE of the initialized experiments and $RMSE_{UNINI}$ is the RMSE of the uninitialized experiments. The RMSE are calculated taking into account the differences between the initialized/uninitialized runs and the observations. Further, a persistence forecast is performed, defined as $x(t + 1) = x(t)$, where x is any parameter and t is a discrete time step. To ensure a common period for all lead times, the RMSE and respective skill scores are calculated for the period 1969–2001.

[10] For verification of surface temperatures we use the HadISST [*Rayner et al.*, 2003] and GHCN-CAMS [*Fan and van den Dool*, 2008] datasets. Observed remote response (SLP, SAT) to North Atlantic SST backward in time is investigated with the 20th-century-reanalysis [*Compo et al.*, 2011]. We also examine correlations between an SST index defined as an average in the North Atlantic [40W-15W, 50N-60N] and the global distribution of surface temperatures and SLP. The significance of the correlation is assessed via bootstrap methods in which the SST index is resampled. The serial correlation of the time series is taken into account by resampling of blocks of the SST index, where the block length is defined by its e-folding time.

3. Forecast Skill of Surface Temperatures

[11] The distributions of the skill scores (SS_{RMSE}) of the ensemble mean surface temperature for yr2-5 and yr6-9 show two marked regions with opposite signs (Figures 1a and 1b). First, large areas in the tropical Pacific show skill scores less than −0.5 for yr2-5 (Figure 1a). For hindcasts with longer lead times (yr6-9) the values of the skill scores are larger than for yr2-5 and tend to zero. The negative values of the skill scores in the eastern tropical Pacific for the first years are also apparent in other CMIP5 models [*Kim et al.*, 2012] and indicate distinct model sensitivity to the initialization. An increase in RMSE of the initialised runs in the first years is also seen in quantities such as the global mean surface temperature (Figures 1c and 1e). Here, the RMSE of the ensemble mean and individual members of the initialized runs reveal small errors for the first year. For yr2-4 the errors are larger than the uninitialized runs but have similar magnitudes with longer lead times.

[12] The second region of noticeable – now, positive – skill scores is the North Atlantic. The eastern part of the North Atlantic shows an improvement of the initialized against uninitialized experiments of up to 50% for yr2-5 and yr6-9 (Figures 1a and 1b). The improvement of the ensemble mean of the initialized runs over the ensemble mean uninitialized runs is seen for all lead times and also for yearly means (Figures 1d and 1f). Here, the RMSE of almost all individual members of the initialized runs are smaller than those for the uninitialized runs. The RMSE of the persistence forecast has low values in the first four years, but becomes larger than the errors of the ensemble mean and single members of the initialized runs, for yearly and 4-yearly means after yr2 and yr2-5, respectively.

[13] As mentioned above, the atmospheric pattern co-varying with decadal-scale North Atlantic SST is modulated within different seasons. Hence, forecasts based on yearly to multi-yearly means may not reflect the seasonally varying remote response to North Atlantic SST anomalies. To examine this, the SS_{RMSE} for the different multi-year seasonal means is calculated (Figures 2a–2d). For winter means, the skill scores exhibit positive values over the North

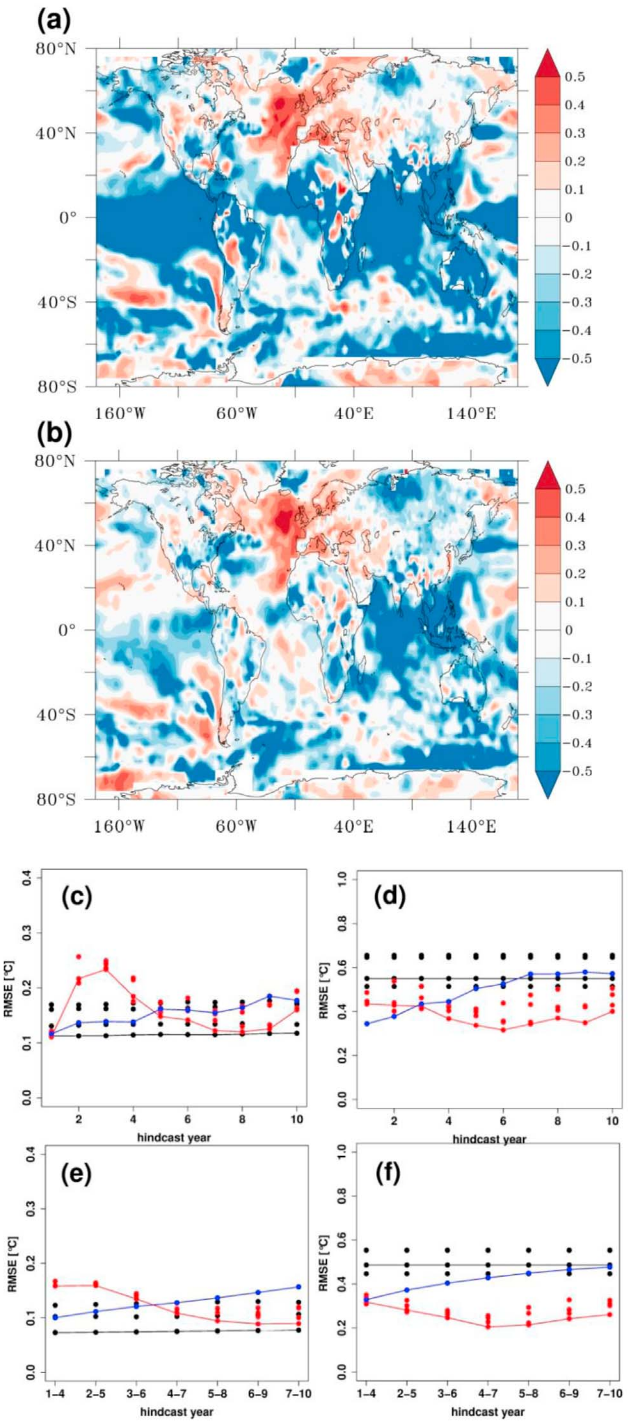


Figure 1. Root-mean-squared-error (RMSE) skill scores for surface temperature for (a) yr2-5 and (b) yr6-9. Reference forecast is the ensemble mean of the uninitialized runs. Also shown is the RMSE for yearly mean surface temperature for different lead times (c) globally averaged [180W-180E, 80S-80N] and (d) over the North Atlantic [40W-15W, 50N-60N]. Single dots show RMSE of individual members of (red) initialized, (black) uninitialized and (blue) persistence. Full dotted lines show ensemble mean. (e, f) Same as Figures 1c and 1d but for running 4 yr average.

Atlantic and northern Europe (Figure 2a). For northern Europe, the corresponding RMSE for the ensemble mean of the initialized runs are smaller than those of the ensemble mean of the uninitialized runs for all lead times (Figure 2e).

[14] From spring to autumn the skill scores are smaller in northern Europe and larger in central-to-eastern Europe. A similar transition is found for the correlation coefficients (not shown). For summer means, SS_{RMSE} in northern Europe is less than 0.1 (Figure 2c). For central Europe, the skill scores are in the range of 0.1–0.3 and appear in a larger region than in winter. The RMSE for central Europe indicates that the error of the ensemble mean of the initialized runs is smaller than that of the uninitialized runs for all lead times (Figure 2f). However, for yr3-6 and yr4-7 the difference is small. The RMSE of the individual ensemble members of the initialized runs are not clearly separated from the RMSE of the uninitialized runs. However, for yr1-4 and yr2-5 the ensemble of the initialized runs show a tendency towards smaller RMSE than the uninitialized runs. The RMSE of the persistence forecast has larger values than the RMSE of the initialized runs for yr2-5 and after yr4-7.

[15] For autumn means, skill scores larger than 0.3 are found in eastern Europe (Figure 2d). For this region, the RMSE of the ensemble means of the initialized runs are smaller than those of the uninitialized runs (Figure 2g). Individual members of the hindcasts have smaller RMSE than the uninitialized runs for all lead times. The RMSE of the persistence forecast has larger RMSE than the ensemble mean of the initialized runs after yr4-7.

4. Observed and Modelled Teleconnections

[16] The previous section has shown that the distribution of the forecast skill scores over European region varies with season, while the eastern North Atlantic provides a robust degree of forecast skill for all seasons (Figures 2a–2d). Because decadal variations of the North Atlantic SST can force the atmospheric circulation resulting in remote signals [e.g., *Czaja et al.*, 2003], we now investigate whether there is an atmospheric bridge linking the skill of the North Atlantic SST with the seasonally dependent skill of surface temperature over land. For brevity, we focus here on summer means, for which skill scores are positive in central Europe and a significant remote impact is found in observations [*Hodson et al.*, 2010]. Time series of a SST index defined as an average over the North Atlantic and a SAT index defined as an average over central Europe illustrate the close coherency of the two areas on multi-decadal timescales (auxiliary material, Figure S1c, correlation r = 0.58 for 1960–2010 and r = 0.43 for 1900–2010).[1] For winter and autumn means, the correlations between the SST index and SAT averaged in regions shown in Figures 2e and 2g are considerably lower.

[17] We examine the pattern of co-variances of observed summer surface temperatures and SLP with the SST index (Figures 3a, 3b, S1a, and S1b). For the period 1960–2010, surface temperatures exhibit a significant positive correlation over the subtropical Atlantic, from central Europe to Africa and over central Asia (Figure 3a). The correlation is significant with values larger than 0.6. Note that the correlations

---

[1] Auxiliary materials are available in the HTML. doi:10.1029/2012GL053326.

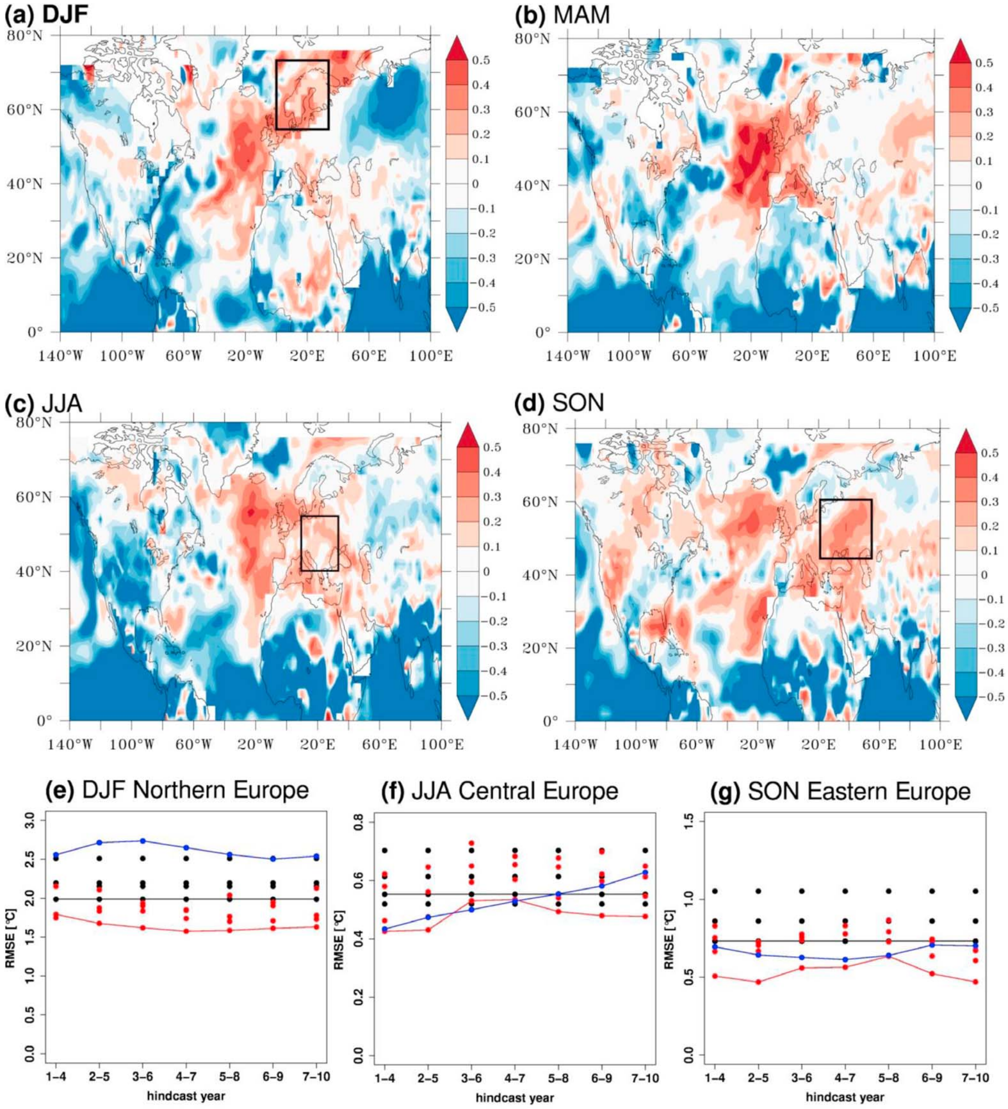
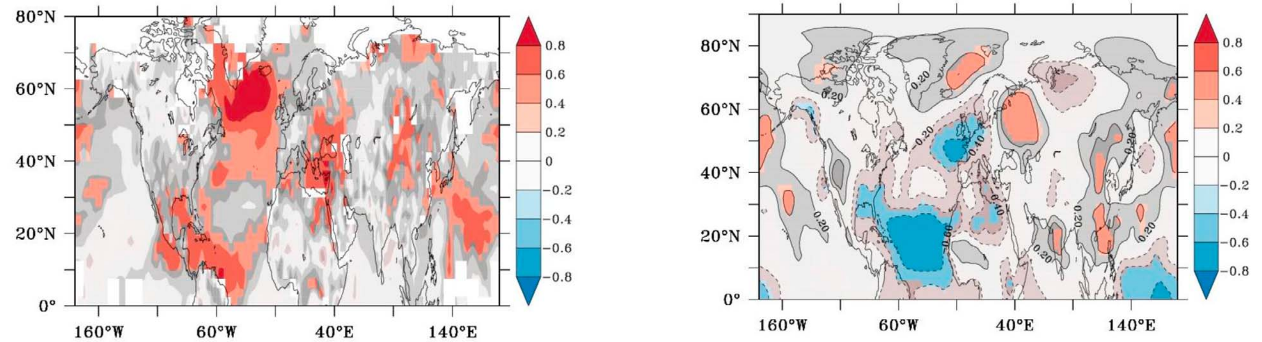


Figure 2. RMSE skill scores for surface temperature for yr2-5 for (a) DJF, (b) MAM, (c) JJA and (d) SON. Also shown are RMSE of running 4 yr-mean land only surface temperature for different seasons and selected regions: (red) hindcasts, (black) historical runs and (blue) persistence averaged over (e) DJF northern Europe [0-30E, 55N-75N], (f) JJA central Europe [10E-30E, 40N-55N] and (g) SON eastern Europe [20E-50E, 45N-60N]. Dots indicate RMSE of single members and full dotted lines show RMSE of ensemble mean.

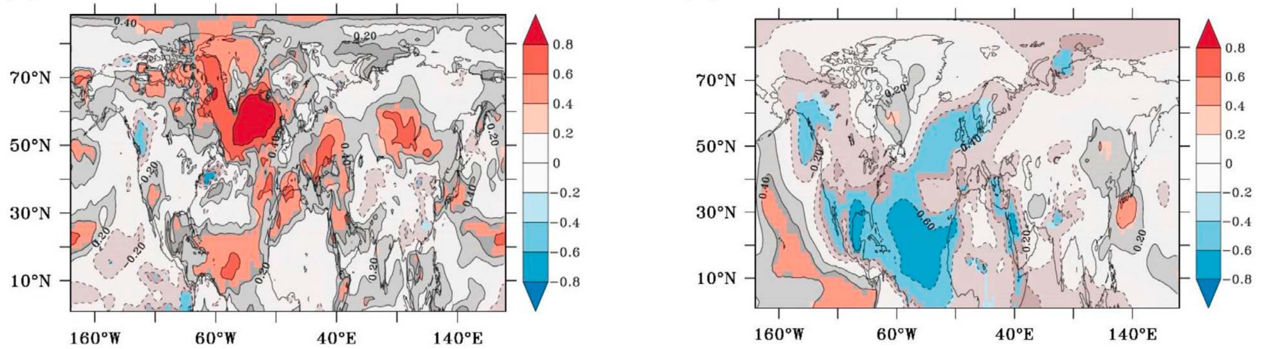
also remain significant considering a period from 1900–2007 (Figure S1a), which underlines that these correlations result from internal climate variability. The correlations of the SST index with the multi-year summer mean SLP (Figure 3b) reveal significant low-pressure anomalies over the subtropical Atlantic (similar to *Hodson et al.* [2010]). Moreover, negative values over the eastern North Atlantic vary in concert with positive anomalies north of Iceland and mark the characteristic pattern of the negative phase of the summer NAO [*Bladé et al.*, 2011]. A positive anomaly over eastern Europe, associated with negative values of the NAO, forms an east-west pressure-gradient. Such pressure-gradient geostrophically induces heat advection in central Europe and thus explains the temperature anomalies in this region. A similar correlation pattern is found for the period from 1900–2007 (Figure S1b). The time series of this pressure-gradient varies in phase with the SST index on decadal timescales (Figure S1d).

[18] The observational analysis suggests a significant covariability between the North Atlantic SST and summer SAT in central Europe, which is linked with a pressure-gradient across Europe. Though the observational analyses in Figures 3a and 3b give clear indications for central to eastern Europe being a potential beneficial region for prediction, the results in Figure 2f show only a small difference between the initialized and uninitialized experiments. To understand whether this small difference is due to the uninitialized experiments showing teleconnections similar to the observations, or whether the prediction system itself has deficiencies limiting the remote response to North Atlantic SST, a similar correlation analysis is performed for assimilation, historical and hindcast yr2-5.

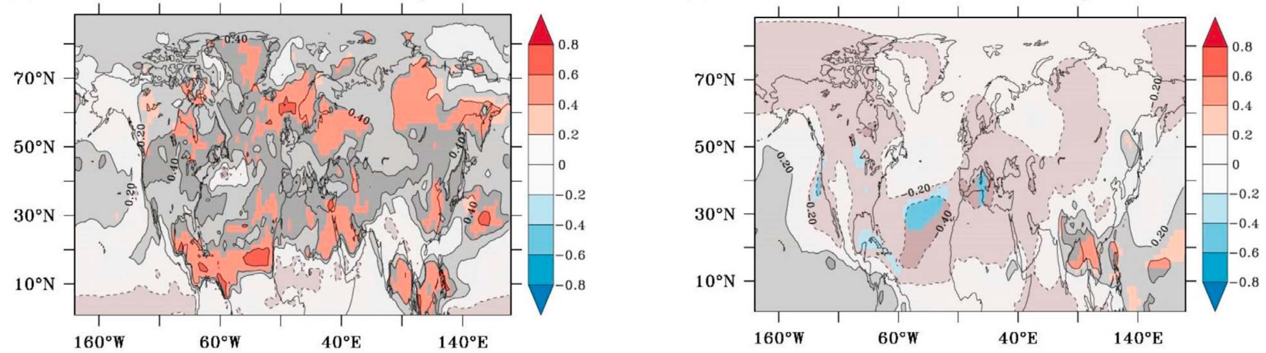
(a) HadISST/GHCN-CAMS **(b)** SLP 20th century reanalysis



(c) Tsurf Assimilation **(d)** SLP Assimilation



(e) Tsurf ensemble mean initialized yr2-5 **(f)** SLP ensemble mean initialized yr2-5



(g) Tsurf ensemble mean uninitialized **(h)** SLP ensemble mean uninitialized

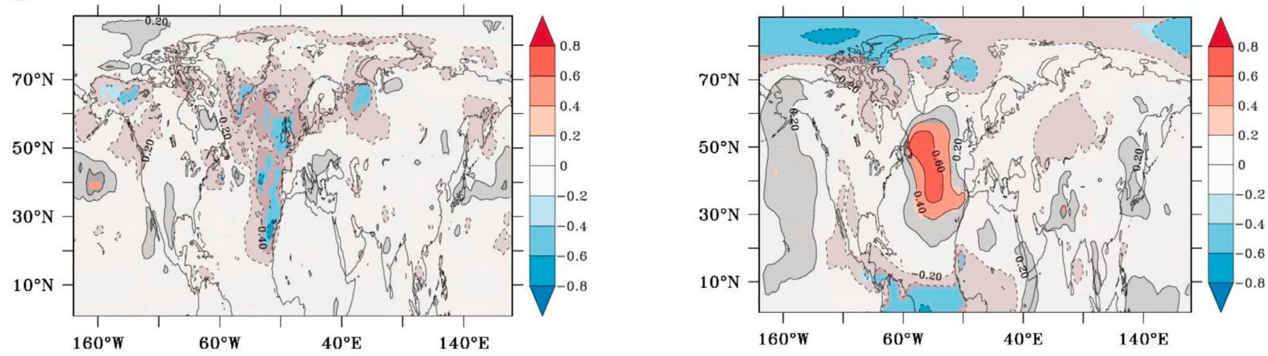


Figure 3. Correlation coefficients of (left) running 4 yr-summer-mean surface temperatures and (right) SLP with running 4 yr-mean HadISST surface temperature index as averaged over the North Atlantic [40W-15W, 50N-60N] for the period 1960–2010. Shown are (a, b) HadISST/GHCN-CAMS surface temperatures and NOAA 20th-century reanalysis SLP, (c, d) assimilation, (e, f) ensemble mean hindcast 2–5 yr and (g, h) ensemble mean uninitialized runs. Red and blue colors indicate significant values at a 90% confidence level. The confidence levels are calculated with 1000 times resampling of the temperature index. Grey shadings show non-significant values. Negative values are dashed.

[19] The correlations between the observed North Atlantic SST index and the assimilated surface temperatures (Figure 3c) show that key regions within the subtropical and North Atlantic are captured by the assimilation run. Though only ocean temperature and salinity are assimilated, the surface temperature is also similar to observations in central to south-eastern Europe and central Asia. Correlations with multi-year summer-mean SLP exhibit negative anomalies over the subtropical and North Atlantic; however, the positive anomalies over the Nordic Seas and Russia are not found. This clearly illustrates that the remote summer response to the North Atlantic SST is in principle transferred via initialisation to the coupled model. The figures, however, also show that this response is weaker than in observations, which might be caused by non-resolved components of the atmospheric response as illustrated in Figure 3d.

[20] For the initialized runs, the correlations are lower than for the assimilation run, although surface temperatures show significant correlations over the subtropical and North Atlantic. Still there are significant positive anomalies over central Europe and central Asia; however, their structures are less confined than in observations and assimilation. The corresponding SLP shows significant, albeit smaller, negative anomalies over the subtropical Atlantic. A significant east-west pressure-gradient, a prerequisite for the geostrophical advection of heat to central Europe, is not found. For the historical runs correlations for temperatures and SLP are different to observations (Figures 3g and 3h).

[21] A comparison of the skill scores (Figure 2c) with the correlation analyses applied to observations and the prediction system (Figure 3) illustrates that the forecast skill of the central European multi-year summer SAT stems from the initialisation of the coupled model. The assimilation experiment has structurally the same response in surface temperature as in observations. To some extent this response is maintained by the hindcasts, which explains the positive values of the skill scores (Figures 2c and 2f). The uninitialized runs show no similar co-variability with observed SST.

[22] However, Figure 3 reveals that the atmospheric response to a given SST is constrained by the model dynamics. The predicted SAT and SLP response over Europe is considerably reduced compared to the assimilation. A comparable analysis with the pre-industrial control simulation underlines these constraints, as no coherency is found between the modelled SST index and multi-year summer SLP and SAT response over central Europe (not shown). Moreover, Figure 3 shows that the correlation coefficients of predicted surface temperature over the eastern North Atlantic are significantly reduced, compared to the assimilation run. A less intense predicted temperature anomaly in the North Atlantic results in a weakening of the forcing of the overlaying atmosphere and hence further reduces or displaces the response in SLP and associated teleconnections.

5. Conclusions

[23] The MPI-ESM decadal predictions performed for CMIP5 are examined for skill scores of surface temperatures. For yearly and multi-year means, the predictions exhibit positive skill scores within the North Atlantic area and distinct regions over the European continent. Over the tropical Pacific, large negative skill scores reflect a systematic error in the initialisation.

[24] A multi-year seasonal decomposition shows pronounced seasonal variations of skill scores. For winter means, positive skill scores are predominantly located in northern Europe. For spring to autumn means, the positive skill scores appear in central and south-eastern Europe. An examination of summer means underlines that the skill scores stem from the initialisation of the coupled model, assimilating the remote response of central and eastern Europe surface temperatures to the North Atlantic SST. Given the different distribution of skill scores for the different seasons, we conclude that multi-year seasonal means would improve the description of skill for various regions, compared to multi-year means. For summer means, we have detected the central-to-eastern European region as a particular beneficial area for decadal climate predictions.

[25] The current model configuration shows a considerably different response of surface temperature and SLP to North Atlantic SST compared to observations (Figure 3). In addition, the assessment of a modelled atmospheric multi-year summer response to SST changes naturally depends on the modelled North Atlantic SST. The coupled model shows a large bias in SST in the North Atlantic, which is reflected in a zonal extension of the North Atlantic current. On the northern side, cold and fresh water masses intrude too far into the central North Atlantic (Jungclaus et al., submitted manuscript, 2012) and may alter simulated thermal wind and eddy growth rate and the atmospheric response.

[26] Given these model deficiencies, we expect in the light of our initialisation and observational analysis, that an improvement of the model components and their coupling will lead to enhanced prediction skill. For example for summer means, adapting the modelled to the observed SLP response such as shown in Figure 3 will also be reflected in an improved advection of heat and hence improved forecast skill of surface temperatures in the affected regions.

[27] **Acknowledgments.** We thank the DKRZ for running the CMIP5 experiments. This research was supported by the BMBF under the MiKlip and Nordatlantik II projects, the EU-FP7 projects THOR and COMBINE, and the DFG-funded project CliSAP.

[28] The Editor thanks Mark Liniger and an anonymous reviewer for their assistance in evaluating this paper.

References

Bladé, I., B. Liebmann, D. Fortuny, and G. J. van Oldenborgh (2011), Observed and simulated impacts of the summer NAO Europe: Implications for projected drying in the Mediterranean region, *Clim. Dyn.*, *39*, 709–727,doi:10.1007/s00382-011-1195-x.

Compo, G., et al. (2011), The twentieth century reanalysis project, *Q. J. R. Meteorol. Soc.*, *137*, 1–28, doi:10.1002/qj.776.

Czaja, A., A. Robertson, and T. Huck (2003), The role of coupled processes in producing NAO variability, in *An Overview of the North Atlantic Oscillation*, *Geophys. Monogr. Ser.*, vol. 134, edited by J. W. Hurrell et al., pp. 147–172, AGU, Washington, D. C.

Dunstone, N. J., D. M. Smith, and R. Eade (2011), Multi-year predictability of the tropical Atlantic atmosphere driven by the high latitude North Atlantic Ocean, *Geophys. Res. Lett.*, *38*, L14701, doi:10.1029/2011GL047949.

Fan, Y., and H. van den Dool (2008), A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.*, *113*, D01103, doi:10.1029/2007JD008470.

Gangstø, R., A. P. Weigel, M. A. Liniger, and C. Appenzeller (2012), Comments on the evaluation of decadal predictions, *Clim. Res.*, in press.

Gastineau, G., and C. Frankignoul (2011), Cold-season atmospheric response to the natural variability of the Atlantic meridional overturning circulation, *Clim. Dyn.*, *39*, 37–57, doi:10.1007/s00382-011-1109y.

Hodson, D. L. R., et al. (2010), Climate impacts of recent multidecadal changes in Atlantic Ocean sea surface temperature: A multimodel comparison, *Clim. Dyn.*, *34*, 1041–1058, doi:10.1007/s00382-009-0571-2.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kim, H.-M., P. J. Webster, and J. A. Curry (2012), Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts, *Geophys. Res. Lett.*, *39*, L10701, doi:10.1029/2012GL051644.

Kröger, J., W. A. Müller and J. S. von Storch (2012), Impact of different ocean reanalyses on decadal climate prediction, *Clim. Dyn.*, *39*, 795–810, doi:10.1007/s00382-012-1310-7.

Mariotti, A., and A. Dell'Aquila (2012), Decadal climate variability in the Mediterranean region: roles of large-scale forcings and regional processes, *Clim. Dyn.*, *38*, 1129–1145, doi:10.1007/s00382-011-1056-7.

Matei, D., H. Pohlmann, J. H. Jungclaus, W. A. Müller, H. Haak, and J. Marotzke (2012a), Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model, *J. Clim.*, doi:10.1175/JCLI-D-11-00633.1, in press.

Matei, D., J. Baehr, J. H. Jungclaus, H. Haak, W. A. Müller, and J. Marotzke (2012b), Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5°N, *Science*, *335*, 76–79, doi:10.1126/science.1210299.

Pohlmann, H., J. H. Jungclaus, A. Koehl, 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, doi:10.1175/2009JCLI2535.1.

Pohlmann, H., et al. (2012), Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system, *Clim. Dyn.*, in press.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res.*, *108*(D14), 4407, doi:10.1029/2002JD002670.

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, D., et al. (2010), Skillful multi-year predictions of Atlantic hurricane frequency, *Nat. Geosci.*, *3*(12), 846–849, doi:10.1038/ngeo1004.

Sutton, R. T., and D. L. R. Hodson (2005), Atlantic Ocean forcing of North American and European summer climate, *Science*, *309*, 115–118, doi:10.1126/science.1109496.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the Experiment Design, *Bull. Am. Meteorol. Soc.*, *93*, 485–498, doi:10.1175/BAMS-D-11-00094.1.

van Oldenborgh, G. J., F. J. Doblas-Reyes, B. Wouters, and W. Hazeleger (2012), Skill in the trend and internal variability in a multi-model decadal prediction ensemble, *Clim. Dyn.*, *38*(7), 1263–1280.