# Decadal prediction skill using a high-resolution climate model

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work.  See Guidance on citing.

www.reading.ac.uk/centaur

# Decadal prediction skill using a high-resolution climate model.

Paul-Arthur Monerie[1], Laure Coquart[1], Éric Maisonnave[1], Marie-Pierre Moine[1], Laurent Terray[1],

Sophie Valcke[1].

[1]CECI UMR 5318 CNRS/CERFACS, Toulouse, France

Abstract

The ability of a high-resolution coupled atmosphere-ocean general circulation model (with a horizontal resolution of a quarter of a degree in the ocean and of about 0.5° in the atmosphere) to predict the annual means of temperature, precipitation, sea-ice volume and extent is assessed based on initialized hindcasts over the 1993-2009 period. Significant skill in predicting sea surface temperatures is obtained, especially over the North Atlantic, the tropical Atlantic and the Indian Ocean. The Sea Ice Extent and volume are also reasonably predicted in winter (March) and summer (September). The model skill is mainly due to the external forcing associated with well-mixed greenhouse gases. A decrease in the global warming rate associated with a negative phase of the Pacific Decadal Oscillation is simulated by the model over a suite of ten-year periods when initialized from starting dates between 1999 and 2003. The model ability to predict regional change is investigated by focusing on the mid-90's Atlantic Ocean subpolar gyre warming. The model simulates the North Atlantic warming associated with a meridional heat transport increase, a strengthening of the North Atlantic current and a deepening of the mixed layer over the Labrador Sea. The atmosphere plays a role in the warming through a modulation of the North Atlantic Oscillation: a negative sea level pressure anomaly, located south of the subpolar gyre is associated with a wind speed decrease over the subpolar gyre. This leads to a reduced oceanic heat-loss and favors a northward displacement of anomalously warm and salty subtropical water that both concur to the subpolar gyre warming. We finally conclude that the subpolar

26    gyre warming is mainly triggered by ocean dynamics with a possible contribution of atmospheric

27    circulation favoring its persistence.

28

29

## 1. Introduction

Climate prediction at interannual-to-decadal timescales can help decision makers to calibrate plans and actions related to climatic events due to climate variability and change (Hibbard et al., 2007; Cox and Stephenson, 2007). Model error and forcing uncertainties are however a source of near-term global prediction biases, as underlined by the global temperature fluctuation of the 2000's, i.e. the so-called "hiatus", that has not been predicted by the CMIP5 climate models (Kosaka and Xie., 2013; Watanabe et al., 2013, Meehl et al., 2014; among others). Decadal prediction experiments are an opportunity to analyze the physical mechanisms associated with climate predictability and variability in a near-future (2010-2035). They can thereby allow model physics and projection improvements. Meehl et al. (2011, 2014), Watanabe et al., (2013), Guemas et al. (2013b) and England et al. (2014) have thus used CMIP5 models to evaluate this 2000's temperature hiatus in climate models. This is a relevant question and is also one goal of this study. In this study we explore the ability of a newly developed high resolution climate model to predict the decadal evolution of temperature, sea-ice extent and volume. In addition, we assess conditional skill by focusing on two specific events, the 2000's temperature hiatus and the 1990's warming of the subpolar gyre.

Motivation for initialized decadal predictions comes from various studies that have successfully predicted regional climate on decadal time scales (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009; Mochizuki et al., 2012; Chikamoto et al., 2013; Bellucci et al., 2013, 2014; Doblas-Reyes et al., 2013a; Garcìa-Serrano et al., 2015; Karspeck et al., 2015). The strongest skill is found in retrospectively predicting surface air temperature over large areas. This is primarily determined by the response to well-mixed greenhouse gases (GHGs), especially for lead-time greater than 2 years (Oldenborgh et al., 2012). Additional skill is given by the major volcanic events that induce an abrupt cooling through the ejected volcanic aerosols (Guemas et al., 2013a; Mehta et al., 2013; Timmreck et al., 2016). Another source of predictability is the initialization of the ocean that can lead to additional skill for the first years of the hindcasts, improving the near-term predictability

57    (Yeager et al., 2012, Robson et al., 2012b; Matei et al., 2012; Chikamoto et al., 2013; Doblas-Reyes et

58    al., 2013a). This is especially relevant for the North Atlantic Ocean where an accurate initialization of

59    the Atlantic meridional overturning circulation (AMOC) allows Atlantic multidecadal variability to be

60    predicted a decade or more in advance (Griffies and Bryan, 1997; Boer, 2000; Collins et al., 2006;

61    Pohlmann et al., 2013). Model studies have then shown that it is possible to produce an accurate

62    simulation of the mid 90's subpolar gyre (SPG) warming thanks to a correct initialization of the ocean

63    mean state (Robson et al., 2012ab; Yeager et al., 2012; Msadek et al., 2014a). The SPG warming is

64    therefore an excellent case study to test the ability of a prediction system, due to its large magnitude

65    and potential predictability, and is thereby one of the main goal of this study.

66

67    The impact of the land surface initialization has also been assessed and is associated with a more

68    limited but significant improvement of the precipitation and temperature predictions (Koster and

69    Suarez, 2003; Paolino et al., 2012). Over the Arctic the sea ice thickness initialization plays a role in

70    the seasonal forecast skill of sea ice extent, has shown in Day et al. (2014) and suggested in Msadek et

71    al., (2014b). Important efforts are thus currently being undertaken to improve sea-ice initialization

72    (Lindsay et al., 2012). The role of other drivers has been summarized in the review of Bellucci et al.

73    (2015).

74

75    It has been suggested that models with larger systematic errors tend to produce lower prediction

76    performance (DelSole and Shukla, 2010). The improvement of the decadal prediction systems by

77    increasing spatial resolution and improving physical parametrization is one of the main challenges

78    proposed in Doblas-Reyes et al. (2013b). There is indeed some evidence that increasing model

79    resolution can improve some aspects of the mean climate. It has been shown that high resolution

80    ocean-atmosphere coupled models produce smaller errors than low resolution models in sea surface

81    temperature (SST) over coastal upwelling regions (Shaffrey et al., 2009; Gent et al., 2010; McClean et

82    al., 2011; Delworth et al., 2012; Sakamoto et al., 2012; Small et al., 2014), due to a better simulation

83    of the wind and its effect on the ocean (Gent et al., 2010; Sakamoto et al., 2012). It reduces the double

84    ITCZ in the tropics (Delworth et al., 2012), El-Niño Southern Oscillation (ENSO) (Shaffrey et al.,

85    2009; Sakamoto et al., 2012; Small et al., 2014) and the north Atlantic SSTs (Gent et al., 2010) are

86    better simulated. Gent et al. (2010) with an ocean-atmosphere coupled model and Jung et al (2012)

87    with a atmosphere model have shown a more accurate atmospheric circulation at higher resolution: in

88    Gent et al. (2010) the maximum sea-surface temperature biases in the major upwelling regions are

89    reduced by more than 60%, the precipitation patterns are improved in the summer Asian monsoon, and

90    the atmospheric circulation in the Arctic is also improved. In Jung et al. (2012) the tropical

91    precipitation and atmospheric circulation, the occurrence of Euro-Atlantic blocking and the

92    representation of extratropical cyclones are better represented in increasing horizontal resolution. Over

93    land, the improvement of the precipitation is partly due to a better resolved orography (Gent et al.,

94    2010; Delworth et al., 2012; Sakamoto et al., 2012) as shown with regional downscaling.

95

96    With the high resolution Community Atmosphere Model (CAM) Bacmeister et al. (2014) have found

97    however that their simulations are not dramatically better than low resolution simulations.

98    Furthermore some problems remain or are exacerbated at these resolutions: some simulations exhibit

99    too much warming and cooling, especially over polar regions where sea ice are not enough accurately

100    represented (McClean et al., 2011; Kirtman et al., 2012), and Bacmeister et al. (2014) have found an

101    exacerbated Pacific double ITCZ bias. The improvement due to a higher resolution is therefore model-

102    dependent. Increasing the resolution does not necessary lead to more accurate simulations.

103

104    Increasing the resolution can also lead to improved seasonal predictions: with the high-resolution

105    GFDL climate model Jia et al. (2015) have shown an improvement of the seasonal prediction of 2-m

106    air temperature and precipitation over land and of the Nino-3.4 index.

107    The question of the skill improvement with a high-resolution climate model remains a wide-opened

108    question. We address it here by using a set of decadal hindcasts performed with the CERFACS-HR

109    high-resolution model. Unfortunately, we cannot directly assess the potential improvements due to

110    resolution only as the model physics version is slightly different from our low resolution version that

111    was used for CMIP5 (Sanchez-Gomez et al., 2015). Furthermore, the initialization dates are different

112 making impossible a fair comparison of the two prediction systems. We thereby simply assess the skill

113 of our new modeling system and qualitatively compare it to the range obtained with low-resolution

114 models. Note also that we do not have historical simulations (uninitialized hindcasts) that would allow

115 us to clearly assess the additional influence of ocean initialization in addition to that of the external

116 forcing.

117

118 The aim of this paper is thus to address three main questions:

119    - What is the skill of our high-resolution coupled climate model in predicting global and regional

120 surface quantities (temperature and sea-ice)?

121    - Is the model able to represent the 2000's global mean temperature fluctuation when initialized

122 with observed ocean initial conditions?

123    - Is the model able to predict the mid-90's warming of the Atlantic subpolar gyre? What are the

124 dominant mechanisms?

125 The paper is structured as follows. The model, data and methodology are described in section 2.

126 Section 3 focuses on the model ability to predict the annual means of SST, air-surface temperature,

127 sea-ice volume and extent. We analyze the simulated 2000's global mean temperature (GMT) changes

128 in section 4. The case-study of the mid-90's north Atlantic warming is investigated in section 5. A

129 summary and discussion are given in section 6.

130

## 2. Data and Methodology

### 2.1 The CERFACS-HR model

CERFACS-HR is a high resolution Atmosphere-Ocean General Circulation Model (AOGCM) developed at Centre Européen de Recherche Avancée en Calcul Scientifique (CERFACS). The atmospheric model is the ARPEGE-climate v5.3 with a horizontal resolution of about 50 kilometers and 31 vertical levels (T359L31), developed at Météo-France/CNRM (Centre National de Recherches Météorologiques). This version of ARPEGE shares the same physic as ARPEGE v2, used in the CMIP5 exercice (CNRM-CM5.1, Voldoire et al., 2013), but runs at a higher resolution (truncature T359 instead of T127) and does not use the SURFEX (SURFface EXternalisé) modeling system. The land-surface Interaction between Soil Biosphere and Atmosphere (ISBA) model (Noilhan and Planton 1989; Noilhan and Mahfouf 1996) is used in order to represent continental surfaces (soil, vegetation, snow).

The ocean model is the Nucleus for European Models of the Ocean (NEMO) v3.4 (Madec, 2008) from the Institut Pierre-Simon Laplace (IPSL). NEMO is discretized on a ORCA025L75 grid (horizontal resolution of about 0.25° and on 75 vertical levels) at global scale decreasing poleward (an isotropic Mercator grid in the Southern Hemisphere, a quasi-isotropic bipolar grid in the Northern Hemisphere with poles over land at 107°W and 73°E). Vertical grid spacing is finer near the surface and increases with depth. Further information on the DRAKKAR ORCA-R025 grid is given in Barnier et al. (2006). The Louvain-la-Neuve Sea Ice model LIM2 3.3 (Vancoppenolle et al., 2009a, 2009b) is embedded into the ocean modeling system NEMO, on a C-grid dynamic-thermodynamic model and includes the representation of the subgrid-scale distributions of ice thickness, enthalpy, salinity and age. The atmospheric and oceanic components are coupled with OASIS3-MCT2.0 (Valcke et al., 2013).

### 2.2 Initialization and external forcing

The hindcasts are full-field initialized every year from 1993 to 2009 (initialized November 1$^{st}$ every year). Ocean and Sea-ice initial conditions are given by the GLORYS2V1 ocean reanalysis product

156 (Ferry et al., 2012). This reanalysis is based on an ocean and sea-ice general circulation model at 1/4°

157 horizontal resolution assimilating in situ profiles of temperature, salinity and sea surface temperature.

158 The atmosphere is initialized from a SST-forced ARPEGE simulation over the same period. The

159 different members are obtained by lagged perturbation of the atmospheric initial conditions (initialized

160 from a different day of the same month). Finally, ten ensembles (1993-2002) of five members

161 integrated over a ten-year period and seven additional ensembles (2003-2009) of five members

162 integrated over a five-year period are available. Each member shares the same radiative forcing and

163 differs due to a different atmospheric initialization and to the model internal variability.

164 The model is constrained by the observed external forcings such as past and current greenhouse gases

165 concentration, solar irradiance and various types of aerosols (black carbon, particulate organic matter,

166 dust, sea salt and sulfate). The optical depths of the tropospheric aerosols are similar as that of the

167 CMIP5 prescribed emissions (Szopa et al., 2013).

168 **2.3 Model evaluation**

169 The magnitude of the bias and drift of CERFACS-HR is assessed by comparing the raw hindcast data

170 with the ocean reanalysis. CERFACS-HR exhibits positive biases of SST over the Austral Ocean (up

171 to 6°C) and over the eastern equatorial Pacific and Atlantic Oceans (up to 4°C) (Fig. 1a). These strong

172 biases, in particular over the equatorial and subtropical Oceans, are commonly obtained with climate

173 model simulations. The bias in monthly GMST increases during the first two years, reaching 1.2°C,

174 and decreases then after 10 years (Fig. 1b).

175 The Arctic experiences a cold bias of 5°C, at the east of Greenland (Fig. 1a). The sea ice extent is

176 therefore anomalously large ($+0.5 \ 10^6 \ km^2$) during the first months of the simulations. The bias

177 increases with time and can reach up to $2.5 \ 10^6 \ km^2$ after 9 years (Fig. 1c).

178 The AMOC spatial pattern is reasonably simulated (Fig. 1d, e) despite a negative bias at the location

179 of the core of North Atlantic deep water (the AMOC is weaker in CERFAC-HR than in GLORYS2V3

180    north of 30°N). Note that the ocean reanalysis AMOC data may also suffer from large biases and does

181    not necessarily reflect the observed AMOC.

182

### 183 2.4 Observed data sets

184    The new Mercator product (GLORYS2V3; Lellouche et al., 2013) covers a longer period (1993-2012)

185    than GLORYS2V1 (1993-2009) and allows us to take into account all the hindcasts integrated over a

186    ten-year period. A comparison between the GLORYS2V1 and GLORYS2V3 products shows a similar

187    mean state of the SST (not shown). Furthermore these products were produced with the same ocean

188    NEMO global configuration, i.e. the ORCA025 grid (Lellouche et al., 2013), than CERFACS-HR.

189    The model ability to predict SSTs is investigated by comparing with the ERSST3b analysis. This

190    product covers the 1854-present period on a monthly frequency and on a 2° x 2° horizontal resolution.

191    The ERSST3b analysis includes the Extended Reconstruction SST version 2 (ERSST) improvements

192    due to analysis methods tuning and to the inclusion of bias-adjusted satellite data (Smith et al., 2008).

193    The observed sea level pressure (SLP) and surface fluxes are given by the ERA-interim reanalysis

194    (Simmons et al., 2007; Uppala et al., 2008; Dee et al., 2011) over a global domain and available from

195    1979 to present. ERA-interim corrects some of the ERA-40 errors: an improved representation of the

196    hydrological cycle, a more realistic stratospheric circulation and a better temporal consistency of the

197    reanalysis fields (Dee et al., 2011).

198    For the surface-air temperature, three sets of observations are used. The Hadley Centre/Climate

199    Research Unit Temperature version 4 (HadCRUT4) provides near-surface air temperature from 1850

200    to present on a 5° x 5° horizontal resolution (Morice et al., 2012). HadCRUT4 covers about 84% of

201    the globe (with a weaker concentration of observations at the poles and over Africa). The coverage is

202    varying in time and missing values are associated to grid points receiving a weak coverage of

203    observations. The incomplete global coverage is a potential source of bias in global temperature

204    reconstructions. HadCRUT4 has been improved in Cowtan and Way (2014) with a better infilling of

205    poorly sampled regions. This additional dataset is hereafter named COWT.

206    The NASA's Goddard Institute for Space Studies Surface Temperature Analysis (GISTEMP) consists

207    of an update of the Goddard Institute for Space Studies (GISS) analysis of surface temperature. The

208    GISTEMP record attempts to address the coverage issue by extrapolating temperatures into

209    unmeasured regions by means of kernel smoothing using a canonical kernel of radius 250 and 1200

210    km (Hansen et al., 2010). We used the data smoothed at 250 km. GISTEMP provides temperature

211    anomalies (the base period is 1960-1990) over land and ocean from 1880 to present on a 2° x 2°

212    horizontal resolution.

213    We also used the NOAA Merged Land-Ocean Surface Temperature (MLOST) version 3b of the

214    National Climatic Data Center (NCDC). MLOST is a merged land air and sea surface temperature

215    anomaly analysis, based on data from the Global Historical Climatology Network (GHCN) of land

216    temperatures and the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) of Sea

217    Surface Temperature (SST) data. The temperature anomalies, with respect to 1961-1990 are analyzed

218    separately and then merged to form the global analysis. More details are given in Smith et al. (2008)

219    and Vose et al. (2012). MLOST is a spatially gridded (5° x 5°) global surface temperature dataset, with

220    monthly resolution from January 1880 to present.

221    The Sea-Ice Extent (SIE) is given by the National Snow Ice Data Center (NSIDC) (Fetterer et al.,

222    2009). As for the NSIDC sea ice index, the CERFACS-HR SIE is computed as the surface of the grid

223    cells covered by ice when the ice concentration is above 0.15. The Sea-Ice Volume (SIV) time series is

224    reconstructed from the Pan-Arctic-Ice-Ocean-Modeling-System (PIOMAS) (Zhang et al., 2008;

225    Schweiger et al., 2011).

226    CERFACS-HR results were interpolated on different grids so to compare them with different as the

227    reanalysis, i.e. a 1.5° resolution for SLP and surface fluxes, a 2.5° resolution for precipitation and a 5°

228    resolution for surface air temperature. SSTs have been interpolated at the ERSST3b 2° resolution. The

229    oceanic variables used in the section 4 (temperature, salt, zonal and meridional velocities) are

230    produced on the same ORCA025 grid as GLORYS2V1 and GLORYS2V3, and have therefore not

231    been interpolated.

232    In section 4 the observed heat content is computed with the Hadley Center reanalysis (EN4 data; Good

233    et al., 2013) that has used all the available data sets including Argo floats of recent years. It is a global

234    quality controlled ocean temperature data set, available from 1900 to present at a monthly frequency at

235    1° resolution with 42 vertical levels.

236    **2.5 Methodology**

237     The hindcasts cover a short period starting every year over a seventeen-year period (1993-2009).

238    Hindcasts starting in 1993-2002 (2003-2009) were run for ten (five) years. Such a short period can be

239    strongly impacted by decadal variability and the removal of a parametric trend associated with the

240    increase in well-mixed GHGs may lead to spurious results. Removing the trend (linear or quadratic)

241    could therefore lead to remove part of the decadal signal. We thus chose to not detrend the temporal

242    data fields.

243    **2.5.1 Bias adjustment**

244    Hindcasts start from the initial conditions and drift away as time progresses. This drift is due to model

245    errors and reflects the model adjustment from the initial observed state back to its equilibrium state.

246    We used a standard procedure following the World Climate Research Program recommendations

247    (ICPO report on mean bias adjustment, 2011) to remove the drift, a posteriori and in a linear way.

248    Let $Y_j^i(\tau)$ the hindcasts and $X_j(\tau)$ the corresponding verifying observations for each member $i$ of the $m$

249    simulations; where j=1 to $n$ indicates the initial year of the prediction (start-date), and $\tau$ is the lead

250    time. For each lead-time $\tau$ we can thus define the mean model predictions $\overline{Y}(\tau)$ and observations $\overline{X}(\tau)$

251    as

252
$$\overline{Y}(\tau) = \frac{1}{nm} \sum_{j=1}^{n} \sum_{i=1}^{m} Y_j^i(\tau)$$

253
$$\overline{X}(\tau) = \frac{1}{n}\sum_{j=1}^{n} X_j(\tau)$$

254    The drift $dr(\tau)$ is thus estimated for each lead-time as the difference between the mean model

255    prediction $\overline{Y}(\tau)$ and the observation $\overline{X}(\tau)$, as

256
$$dr(\tau) = \overline{Y}(\tau) - \overline{X}(\tau)$$

257    The corrected model prediction, for each member $i$, initial year $j$ and lead-time $\tau$ is then defined as

258
$$\hat{Y}_j^i(\tau) = Y_j^i(\tau) - dr(\tau)$$

259    This method assumes that the drift is independent of the initial conditions. The drift is however not

260    stationary and the results may be sensitive to this methodology (in the case where there is strong non-

261    stationary drifts it could be preferable to use the method developed in Kruschke et al., 2015). Here we

262    assume that the ICPO method is a reliable bias adjustment method.

263    In section 3 we choose to remove the drift of SST based on the observed datasets used to assess skill

264    (ERSST3b). The drift in air-surface temperature is removed with ERA-interim since HadCRUT4 is

265    given in anomaly. We have test the sensitivity to the selected reanalysis used to remove the drift: we

266    have found weak differences among the used reanalysis (not shown). In section 4 we analyzed the

267    subpolar gyre warming and used the GLORYS2V3 reanalysis (that are close to the GLORYS2V1

268    reanalysis used to initialize the ocean model while spanning a longer time period) to remove the drift

269    on the original ORCA grid of the model. We used ERA-interim for heat fluxes, surface pressure and

270    surface winds.

271    **2.5.2 Evaluation of hindcast skill**

272    The skill is computed with an Anomaly Correlation Coefficient (ACC) $r$, given by

273
$$r(\tau) = \frac{\sum_{j=1}^{n}[X_j(\tau)Y_j(\tau)]}{\sqrt{\sum_{j=1}^{n}[X_j(\tau)]^2 \sum_{j=1}^{n}[Y_j(\tau)]^2}}$$

274    Where $Y_j$ is the ensemble mean anomaly for the $j^{\text{th}}$ hindcast starting in year $j$; $X_j$ is the observation

275    anomaly for the corresponding starting date $j$. $\tau$ is the lead-time and $n$ the number of starting date.

276    Anomalies for $X_j$ and $Y_j$ are calculated independently, both having a zero mean over the hindcast

277    period.

278    We estimate the significance of the ACC through re-sampling and 1000 permutations in a Monte-

279    Carlo framework. The values are judged significant at the 5% level if the correlation values are

280    stronger than 97.5% of the randomly obtained correlation values. We also use the root mean squared

281    error (RMSE). RMSE measures the magnitude of the error between the hindcasts and the observations

282    while ACC measures only the phase difference between observations and hindcasts. RMSE and ACC

283    thus provide complementary information on the prediction skill of CERFACS-HR.

284    The 4-year persistence is computed for the SST, SIE and SIV based on the observed values of SST,

285    SIE and SIV in the 4 years prior to the start date. The skill of this empirical model to predict the

286    climate depends on the initial condition given the climate system memory.

287

## 3. Model skill

### 3.1 Sea Surface Temperature

The ACC is computed for the first year of the hindcasts (lead-time 1) and averaged over two different lead-times: 2-5 and 6-9 years (based on the hindcasts for which the time-period is long enough). The skill of the first year is strongly dependent on the initial conditions and represents the seasonal-to-interannual predictions; the 2-5 years lead-time represents the interannual timescale (Goddard et al., 2013). The comparison of the 2-5 and 6-9 year means allows analyzing the dependence of the skill on lead time. The 4-year average reduces higher frequency noise. Another reason of the choice of the lead-times 2-5 and 6-9 is that they are frequently used (Bellucci et al., 2013, 2014; Goddard et al., 2013; Karspeck et al., 2015, among others) and results are thus comparable with other modeling studies.

The skill in predicting annual mean sea-surface temperature (SSTs) is given by the ACC between the hindcasts and the ERSST3b analysis. The first year exhibits significant correlations over the Atlantic Ocean (mostly over the tropical Atlantic and the Caribbean), the Indian Ocean and the Pacific Ocean, west of 150°W (Fig. 2a). The 2-5 year lead-time still exhibits skill over tropical and North Atlantic (but with a lesser extent), Indian and West of the Pacific Ocean (Fig. 2b). The 6-9 year lead-time exhibits negative correlations over the Pacific Ocean and the skill over the Caribbean vanishes, underlying the loss of skill in some areas (Fig. 2c).

It has been shown in Sanchez-Gomez et al. (2015) that the CNRM-CM5 model drift is due to quasi-systematic excitation of an ENSO event at the first year of the hindcasts. Since the drift is not totally removed by the ICPO adjustment method a similar mechanism could occur in CERFACS-HR (which also used the ARPEGE and NEMO model but at higher resolution) and can partly explain the weak skill obtained in the Pacific Ocean.

For the 6-9 year lead-time, the most prominent result is the skill obtained over the tropical Atlantic that is unexpected since biases, due to the poor representation of the mean intensity and dynamic of

14

313  the cold tongue and seasonal cycle of the SSTs (Okumura and Xie, 2004; Richter and Xie., 2008)

314  generally lead to low skill in these regions (Stockdale et al., 2006). Fig. 2c shows annual mean

315  correlation but significant correlations remain in March-April-May over the tropical Atlantic (not

316  shown), the period of the cold-tongue set-up and the season where the bias is at its maximum

317  (Zermeño-Diaz and Zhang, 2013).

318  The RMSE is strong but no significant over the equatorial Pacific and the subtropical north Atlantic,

319  for the first year (Fig. 3a). The 2-5 and 6-9 year lead-time areas with large and significant RMSE value

320  correspond to the areas of large and negative ACC value, i.e, the Pacific Ocean for 2-5 and 6-9 year

321  lead-time and the Austral Ocean and subtropical North Atlantic for the 6-9 year lead-time (Fig. 3b, c).

322  The loss of skill is thus due to inabilities of CERFACS-HR to reproduce both the magnitude and phase

323  of the SST.

324  Different boxes are defined in order to show the ability of the model to predict regional indexes: the

325  Atlantic multidecadal oscillation (AMO) index ([0°N-60°N; 80°W-0°W]), the nino3.4 index ([5°S-

326  5°N; 170°W-120°W]), the tropical Atlantic ([3°S-3°N; 20°W-10°E]) and Indian SSTs ([30°S-10°S;

327  60°E-100°E])(see Fig. 2a). The global mean SST domain is defined as the average of all the observed

328  grid points containing non-missing values.

329  Fig. 4a shows the correlation score for these five indexes and for 4-year averaged periods. Gray

330  shading represents the non-significant correlations (according to the Monte-Carlo test) and the 4-year

331  persistence is represented within the purple line.

332  When we consider the global domain, results are significant for the first lead-time, decrease as time

333  progresses, and become non-significant for the lead-time 3-6 year (Fig. 4a). It is almost zero at a lead-

334  time 6-9 year. The 4-year mean persistence is not beaten by CERFACS-HR leading to the conclusion

335  that the model ability to predict the mean global SSTs is mostly due to the external forcing. The

336  decrease of the skill is partly due to the negative correlation over the Pacific Ocean from the lead-time

337  2-5 to 6-9 year (Fig. 4e). This is consistent with the Fig. 2c which exhibits negative correlations over

338  the Pacific Ocean.

339    Correlations are significant for the AMO, Tropical Atlantic and Indian Ocean. The skill of the AMO is

340    positive and significant for the lead-times 1-4 year to 4-7 years and becomes non-significant for the

341    lead-times 5-8 and 6-9 year. A positive and significant skill for the AMO was obtained by numerous

342    authors (Kim et al., 2012; Chikamoto et al., 2013; Bellucci et al., 2014; Garcìa-Serrano et al. 2015).

343    This high predictability was attributed to multidecadal variability of the ocean dynamics (Pohlmann et

344    al., 2004) and to the initialization of the ocean (Garcìa-Serrano et al., 2015).

345    Over the tropical Atlantic, correlations using CERFACS-HR are significant and stronger than

346    persistence for lead-times up to 3-6 years. The skill in predicting tropical Atlantic SSTs is associated

347    with variables linked to the tropical Atlantic SSTs variability: the westerly winds (Richter et al., 2014a,

348    2014b), precipitation over the adjacent areas (Chang et al., 2008; Richter and Xie, 2008; Zermeño-

349    Diaz and Zhang, 2013) and deep convection (Richter et al., 2014b). Since no skill is found on the

350    pressure field (involving the subtropical gyre), the heat content (involving the memory of the ocean)

351    and the meridional wind and precipitations (Supplementary material S1) we hypothesize that this

352    result could be due to the sampling (we only used ten start-dates). The question of the relationship

353    between these variables and the tropical Atlantic SSTs in CERFACS-HR is however not under the

354    scope of this study.

355    The skill over the Indian Ocean is positive and marginally significant for several lead times (1-4 years,

356    2-5 years and 6-9 years). Guemas et al. (2013a) used a set of decadal predictions and found that the

357    Indian Ocean stands out as the region where the predictions of SST perform the best worldwide. The

358    models ability to predict the SSTs is mainly due to the long-term warming trend (as in Guemas et al.,

359    2013a).

360    The CERFACS-HR model is thus able to predict the decadal trends of several indexes of SST (AMO,

361    Tropical Atlantic and Indian) and with at least similar skill as that shown in studies based on low-

362    resolution models (Bellucci et al., 2013; Chikamoto et al., 2013; Guemas et al., 2013a).

363    **3.2 Sea ice extent and volume**

364 The observed shrinking of summer Arctic sea ice extent (SIE) (Serreze et al., 2007; Stroeve et al.,

365 2012) has increased accessibility to marine waters and has raised the possibility of an ice-free Arctic in

366 the near-future (Wang and Overland, 2009). This rapid decline also contributes to the polar

367 temperature amplification (Screen and Simmonds, 2010) and is thereby an indicator of climate change.

368 Accurate prediction of SIE and the sea ice volume (SIV) can thus lead to a better understanding of the

369 incoming climate change. There are plenty of prediction studies at a monthly to seasonal timescale

370 (Sigmond et al., 2013; Day et al., 2014; Msadek et al., 2014b; among others) about sea-ice but only a

371 few studies at decadal time-scales with for instance Blanchard-Wrigglesworth et al., (2011) and Germe

372 et al., (2014).

373 The skill in simulating the SIE and SIV 10-year trends is investigated with the Fig. 5 for ten start dates

374 (from the 1994-2003 to 2003-2012 trend). The observed SIE and SIV trends are included in the

375 CERFACS-HR spread for a majority of start dates, highlighting the importance of both the external

376 forcing and the initial conditions in predicting these quantities. The mean simulated and observed

377 trends are negative (except observed SIE in winter), indicating a continuous shrinking of SIE and SIV.

378 CERFACS-HR well simulates the sign of the long-mean trend with a better result in summer (Fig. 5bd)

379 than in winter (Fig. 5ac) (the SIE and SIV shrinking is not apparent in the observations in winter and is

380 over-estimated with CERFACS-HR). The internal variability of SIE and SIV is strong (the ensemble-

381 spread of summer SIE is of 2 million $km^2$ for the 2002-2011).

382

383

384

385

## 4. The global mean temperature fluctuation of the 2000's

386

387 Despite a continuous increase in atmospheric GHGs, the GMT has shown a quasi-stabilization since

388 1998. This recent temperature evolution has been named "the hiatus" and has been associated with

389 several possible factors: a negative phase of the Pacific decadal oscillation (PDO) with anomalously

390 cold SSTs over the eastern Pacific ocean and more heat uptake (Meehl et al., 2011; Kosaka and Xie,

391 2013; Watanabe et al., 2013; Trenberth and Fasullo, 2013; Guemas et al., 2013b; England et al., 2014;

392 Meehl et al., 2014; Douville et al., 2015), the recent stratospheric volcanic aerosol trend (Fyfe et al.,

393 2013; Santer et al. 2014; 2015; Haywood et al., 2014; Ridley et al., 2014; Schmidt et al., 2014; Brühl

394 et al., 2015; Mills et al., 2016), the decrease in tropospheric water vapor concentration (Solomon et al.,

395 2010) and the solar minimum around 2009 (Kopp and Lean., 2011).

396 We analyzed the ability of CERFACS-HR to simulate the so-called hiatus over the recent decades.

397 Figure 6 represents the decadal linear trends for the hindcasts initialized from November 1997 to

398 November 2002. The hindcast surface air temperatures are interpolated on the same horizontal grids

399 and with the same missing values as in COWT, MLOST and in GISTEMP and the respective trends

400 are calculated over the ten-year period starting in January following the initialization (e.g. over the

401 1998-2007 period for the hindcasts starting in November 1997).

402 The uncertainty of the annual-mean GMT evolution is addressed using these three observed datasets,

403 which trends are shown on Figure 6, i.e. black crosses for COWT, blue squares for MLOST and

404 orange triangles for GISTEMP. The trends computed over the last decades are weaker as time

405 progresses and highlight the decrease of the global warming rate. The observations are globally

406 consistent but with COWT exhibiting values about .1°C per 10 years higher. This is mainly due to the

407 differences in spatial coverage among the various observations and shows the impact of high-latitude

408 regions since regions north of 60°N present a trend of about 0.5°C per 10 years higher (Fig. 7b).

409 Indeed Cowtan and Way (2014) have shown that the coverage bias causes a cool bias in recent

410 temperature trend relative to the late 1990s, that increases from around 1998 to present. Hansen et al.

411 (2010) have also shown that the global temperature change is sensitive to estimated temperature in

412 polar regions where observations are limited.

413 The large spread among the hindcasts stands out when different grid points are replaced by missing

414 values as in COWT (black circles), MLOST (blue circles) and in GISTEMP (orange circle). This is

415 due to strong positive trends over the North Pole in the model (Fig. 7a), which are more or less

416 masked depending on the various observed masks. For each start date, the gray area represent the

417 spread (+/- one standard deviation) computed with all the hindcasts, and the white area is the average

418 over all the 3 interpolated hindcasts, for each start-date. The spread is strong with mean trends ranging

419 from -0.08 to +0.19 °C per ten years. From 1999 and on, the warming rate decreases with time in the

420 hindcasts as in the observations with the weakest trends occurring over the last decade (2003-2012).

421 The model is thus able to simulate the hiatus of the 2000's. Interestingly, the hindcasts considered with

422 the spatial coverage of MLOST (CERFACS-HR/MLOST, blue circles) exhibit 10-year trends closer to

423 the raw MLOST observations (blue squares) than CERFACS-HR/GISTEMP or CERFACS-

424 HR/COWT with their respective observations.

425 Note that for the hindcasts starting in November 1997, the model is initialized with an anomalously

426 warm SSTs over the eastern Pacific ocean (the 1998 El-Nino) and the 1998-2007 trends are largely

427 weaker than in the observations (this is even more evident when we only consider the tropical latitudes,

428 not shown). With the standard bias-correction method, it is assumed that the drift is stationary and

429 independent of the initial conditions. The strong discrepancy between the hindcasts and the

430 observations for these strong events suggests that the bias-adjustment method may not be able to

431 correctly remove the drift.

432 The last decade (2003-2012) exhibits the weakest 10-year trend and is reasonably simulated since the

433 difference between the observations and the model does not exceed the model spread (Fig. 6). Figure 7

434 shows the 10-year trend in air surface temperature for CERFACS-HR and COWT (Fig. 7a, b), and

435 ocean heat content [0-2000m] for CERFACS-HR and EN4 (Fig. 7c, d). We used the 0-2000m depth to

436 take into account the upper and mid ocean layers, which have undergone an increase in heat uptake

437  during the hiatus period (Meehl et al., 2011). The horizontal resolution of EN4 is weaker than for

438  CERFACS-HR and each grid cell point contains thus, by construction, more heat. The heat content

439  comparison was made available by scaling the observations and CERFACS-HR by the surface of each

440  grid cell. The heat content values represent thus the heat contained in a 1m$^2$ area down to a depth of

441  2000 meters. The trend in air surface temperature is positive over the Arctic and the Indian Ocean (Fig.

442  7a). The positive trend over the western Pacific and the negative trend over the eastern Pacific Ocean

443  indicate that CERFACS-HR simulates a negative phase of the PDO. This is consistent with the

444  observed trend (Fig. 7b). Over the Pacific Ocean stippling indicates that the trend is more negative in

445  the observations than in CERFACS-HR. This pattern is however consistent with the literature and is

446  associated with an increase (decrease) in the ocean heat content over the Western Pacific and Indian

447  (eastern Pacific) Ocean (Fig. 7c, d). The positive and significant heat content trend over the Indian

448  Ocean is consistent with Lee et al., (2015) who show that the heat has been transferred from the

449  Pacific Ocean to the Indian Ocean through the Indonesian straits.

450  During this decade the volcanic activity has led to an increase in stratospheric aerosol optical depth

451  that is able to explain a part of the hiatus (Fyfe et al., 2013; Santer et al. 2014; Haywood et al., 2014).

452  In CERFACS-HR the stratospheric aerosol optical depth is prescribed based on the data of Vernier et

453  al. (2011). Unfortunately a technical problem led to an inaccurate representation of these volcanic

454  aerosols in the model leading to a very small effect equivalent to that of a small and constant volcanic

455  aerosol background. Hence, the ability of CERFACS-HR to simulate the hiatus cannot be attributed to

456  the impact of the stratospheric aerosols and can therefore mainly be explained by the ocean

457  initialization and heat uptake. Additional simulations are needed to conclude on a possible impact of

458  the stratospheric aerosols on the modeled hiatus. This is the main topic of a current work with

459  CERFACS-HR.

460

461  The ability of CERFACS-HR to simulate the recent 2000's temperature fluctuation could

462  seem contradictory to inability of CERFACS-HR to simulate the GMST at decadal scale (6-9

463   years). These two results are however not directly linked since the former is associated with

464   the hindcasts of only a few start-dates and the latter is based on all the hindcasts.

465

## 466   5. North Atlantic warming of the mid-90's

467   In this section we address the question of the model ability to predict the North Atlantic warming that

468   occurred during the mid-90's, when the subpolar gyre warmed significantly by 1°C in just 2 years

469   (1995-1996) (Robson et al., 2012a). At the same time the North Atlantic Oscillation (NAO) went from

470   a long positive phase (1989-1995) to a negative phase (1996). Different hypotheses, linking the

471   subpolar gyre and the North Atlantic Ocean heat content (OHC) with the phase of the NAO, have been

472   proposed to explain this sudden warming (Hurrell, 1995; Bersch, 2002; Hurrell et al., 2003; Bersch et

473   al., 2007; Sarafanov et al., 2008; Ortega et al., 2011, 2012; Barrier et al., 2015). The first hypothesis is

474   based on the long positive phase of the NAO (NAO+) and the second on the short and strong negative

475   NAO (NAO-) phase that occurs in 1996 (Lohmann et al., 2009):

476   -   A NAO+ phase is associated with heat loss over the subpolar North Atlantic (Barrier et al.,

477       2014) through increased westerlies across the subpolar gyre (Hurrell et al., 2003; Visbeck et

478       al., 2003; Yashayaev, 2007; Barrier et al., 2014). Temperatures cool down, leading to an

479       increase in deep-water formation and to a strengthening of the AMOC (Lohmann et al., 2009,

480       Ortega et al., 2012). This favors an increase in northward oceanic heat transport leading to the

481       warming of the subpolar North Atlantic Ocean several years later (Johns et al, 2011; Msadek

482       et al, 2013).

483   -   The NAO- phase has been associated with an increase in OHC over the subpolar gyre through

484       a change of the oceanic circulation and air-surface heat fluxes (Bersch, 2002; Bersch et al.,

485       2007; Sarafanov et al., 2008; Lohmann et al., 2009; Barrier et al., 2015). The NAO- phase can

486       favor a shrinking and a weakening of the subpolar gyre and thus a northward motion of warm

487       subtropical waters into the subpolar gyre (Hatun et al., 2005; Sarafanov et al., 2008).The north

488       Atlantic warming could thus be linked to the abrupt change from a positive to a negative phase

21

489    of the NAO (as shown by Lohmann et al., 2009 maintaining NAO- conditions in numerical

490    experiments).

491

492    The first hypothesis involves the NAO- phase and therefore implies low-predictability (Collins, 2002;

493    Muller et al., 2005) while the second one implies higher predictability through a lagged AMOC

494    response (Griffied and Bryan, 1997; Pohlmann et al., 2004; Collins et al., 2006, Pohlmann et al.

495    (2013).

496    5.1 Oceanic changes

497    The subpolar gyre simulated in our hindcasts is represented with the annual-mean barotropic stream

498    function (BSF) in Fig. 8. Negative values indicate a cyclonic gyre that extents from the Labrador Sea

499    to the Iceland Basin with a maximum south-east of Greenland. The subpolar gyre (hereafter noted

500    SPG) area is defined from 60°W to 10°W and from 50°N to 65°N as in Robson et al. (2012ab) and

501    Msadek et al. (2014a) (see the box in Fig. 8).

502    Fig. 9 exhibits the OHC anomalies integrated over the SPG area and from the surface to a 500-meter

503    depth for CERFACS-HR and GLORYS2V3. We used the 0-500-meter integrated anomalies but

504    checked that the warming is consistent over the 0-1000-meter depth (Supplementary material S2).

505    Since the shift occurs in 1995-1996, we focus on hindcasts initialized few years before (i.e. in 1993

506    and 1994; hereafter DH93 and DH94) as well as those initialized in 1995 and 1996 (hereafter DH95

507    and DH96). GLORYS2V3 exhibits a strong warming from 1993 to 1998 (Fig. 9a; black line). DH93

508    and DH94 do not simulate a strong and abrupt warming (with a positive anomaly in heat content) over

509    the SPG (Fig. 9a, b). However the simulations initialized in 1995 and 1996 exhibit a shift (Fig. 9c, d),

510    as shown by the large and sudden increase of OHC during the first four years. DH95 simulates a

511    strong warming (year 1 to 6), with the SPG remaining anomalously warm during 5 years (year 4 to 8),

512    and finally cools down from year 6 to 10 (Fig. 9c). DH96 exhibits a warming that maintains itself

513    during 8 years (Fig. 9d) and thus simulates the observed persistent SPG warming. The following

514    analysis focuses on DH95, which is the first hindcast of the series to simulate the SPG warming.

515   A comparison of the vertically integrated temperature [0-300m] between the GLORYS2V1,

516   GLORYS2V3, ORA-S4 (Balmaseda et al., 2013), and Nemovar1 (Balmaseda et al., 2010) shows that

517   GLORYS2V1 and GLORYS2V3 are close to but warmer than ORA-S4 and Nemovar1 of about 0.5

518   degrees Celsius over the SPG area (Supplementary material S3). ORA-S4 and Nemovar1 cover a

519   longer period than GLORYS2V1 and GLORYS2V3; however, their lower temperature may lead to an

520   overestimation of the warming and therefore these reanalysis were not used throughout this study.

521   The evolution of the volume weighted annual-mean temperature anomalies [0-500m] are shown on

522   Fig. 10 over the 1996-2005 period. The first two years exhibit a positive temperature anomaly

523   centered at 45°N and 40°W (Fig. 10a, b). This anomaly then propagates northward following the

524   eastern side of the SPG (Fig. 10c, d) and reaches the northern SPG boundary during the fifth year (the

525   maximum of warming is located at the south-east of Greenland on Fig. 10e). As in Fig. 9c the SPG

526   then starts to cool down and the positive temperature anomaly slowly vanishes (Fig. 10h-j). The

527   timing of DH95 is not exactly synchronized with the reanalysis since the maximum warming occurs in

528   2001 (instead of 1998 in GLORYS2V3) but its amplitude is well reproduced (about 1°C).

529   Some studies have suggested that the SPG warming is due to the northward migration of subtropical

530   anomalously warm and salty waters. The salt anomalies [0-500m] simulated by DH95 are shown

531   within the figure 11. A positive salt anomaly propagates from the subtropical gyre (Fig. 11a) to the

532   SPG (Fig. 11e-h). At the same time the subtropical gyre exhibits a negative salt anomaly suggesting a

533   transfer from the subtropical to the SPG. As for temperature the maximum is located at the south-

534   eastern tip of Greenland. The positive anomaly located on the eastern side of the subtropical gyre

535   highlight the importance of the gyre as a source of salt and temperature for the SPG warming. This

536   anomaly is also found in Msadek et al. (2014a) in the simulation that reproduces the SPG warming. It

537   also suggests a possible role of the Mediterranean Sea in exporting anomalously warm and salty

538   waters into the subtropical gyre.

539   The subtropical gyre strengthens along with the North Atlantic current during the first five years of

540   DH95 (Fig. 12a-e). In 1998 a robust and positive anomaly extends from the north-west (55°W-40°N)

541    of the subtropical gyre to the east (40°W-55°N) of the SPG (Fig. 12c). It reinforces in 1999 (Fig. 12d)

542    denoting a northward shift of the Gulf Stream (for more clarity the mean climate is not shown but the

543    reader can refers to the Fig. 8). The BSF weakens from 1998 to 2000 over the eastern bouday of the

544    SPG. The SPG continue then to weaken (Fig. 12d-f), especially in 2003 (Fig. 12f).


545    In summary the strengthening of the north Atlantic current favors the northward move of a salty and

546    warm anomaly that reaches the Iceland Basin, south of Greenland and the Labrador Sea several years

547    later (see Fig 11e for the salt and Fig 10e for the temperature).


548    The evolution over the 1996-2002 period of the Meridional Heat Transport (MHT) integrated over the

549    whole oceanic column is analyzed in Fig. 13. MHT increases from 1996 to 2000 (with significant

550    anomalies for the first 2 years) north of 40°N with a maximum at 55°N (Fig. 13a). MHT decreases

551    after 2000. The MHT change is thus consistent with the temperature evolution over the SPG, from

552    40°N to 60°N (Fig. 10).


553    The MHT anomaly is decomposed in the anomalous advection of mean temperature ($\overline{T}v^{'}$), the mean

554    advection of anomalous temperature ($\overline{v}T^{'}$) and the cross-product of anomalies ($T^{'}v^{'}$) (Fig. 13).

555    This decomposition is computed at a monthly time-scale, bars denoting monthly means and prime the

556    departure from the monthly means. These products, averaged in annual means are shown in figure 13.

557    The anomalous advection $\overline{T}v^{'}$, denoting an enhanced northward heat transport over the SPG

558    associated with a strengthening of the oceanic circulation, dominates the MHT change (the anomalies

559    of $\overline{T}v^{'}$ are significant from 1996 to 2002) between 50°N and 55°N (Fig. 13b). The anomaly of MHT

560    due to the mean advection of anomalous temperature ($\overline{v}T^{'}$) is positive and significant north of 55°N

561    during the first five years of DH95 (Fig. 13c). This ($\overline{v}T^{'}$) anomaly is located at 55°N during the first

562    3 years, then moves northward and reaches 60°N from 1998 to 2001; it is generally weaker than the

563    $\overline{T}v^{'}$ anomaly but dominates the MHT anomalies at 60°N during the 1998-2000 period. This mean

564    advection of temperature anomalies is thus sustaining the warming first created by the anomalous

565    advection of mean temperature. The ($T^{'}v^{'}$) component is weak and can be neglected over the SPG

566 (Fig. 13d). During the last two years of DH95, the decrease of temperature is associated with a weaker

567 MHT linked with a decrease of both $\overline{v}T'$ and $\overline{T}v'$.

568 Fig. 14 shows the mixed-layer depth anomalies averaged over the years 1996-2000. The mixed layer is

569 deeper, consistent with more deep water formation over the Labrador Sea. It denotes an increase of

570 convective activity (Ortega at al., 2011), leading to an AMOC strengthening and an increase of

571 meridional heat transport (Johns et al, 2011; Msadek et al, 2013). The deepening of the mixed layer is

572 also consistent with the $\overline{T}v'$ increase and confirms (as in Msadek et al., 2014a) the prominent role of

573 the ocean dynamics on the SPG warming. The AMOC is particularly strong in November 1995 (and

574 stronger than for the other start date) in GLORYSV21 (not shown). The DH95 dynamical shift is

575 probably, at least partly, built-in in the initial condition. The mixed-layed depth anomaly weakens

576 after 2002 and vanishes during the second half of the simulation (2001-2005).

577 5.2 Atmospheric changes

578 Some studies have also suggested a possible role of heat fluxes in relation with the atmospheric

579 variability (Bersch, 2002; Bersch et al., 2007; Sarafanov et al., 2008; Lohmann et al., 2009). The

580 surfaces fluxes are shown in the Fig. 15. By convention positive (negative) values denote downward

581 (upward) anomalies. The net surface heat flux (the sum of net shortwave, net longwave, sensible and

582 latent heat fluxes) averaged over the SPG exhibits positive values over the 1998-2003 period (Fig.

583 15a), which means that the atmosphere is warming the ocean in DH95. The surface flux anomalies are

584 stronger over the Labrador Sea, the Irminger Sea and the Iceland Basin (Fig. 15b). However, the

585 strongest positive anomaly is located south of 50°N (southern boundary of the SPG). The net

586 shortwave and longwave fluxes show weak and non-significant anomalies (Fig. 15c, d). Sensible and

587 latent heat fluxes exhibit stronger positive anomalies (Fig. 15e, f), denoting a decrease in oceanic heat

588 loss.

589 We now discuss changes in the surface wind speed and sea-level pressure (SLP) average over the

590 1998-2003 period (Fig. 16). The SLP denote a negative phase of the NAO with positive (negative)

591   anomalies of SLP over Iceland (the subtropical latitudes). The SLP exhibits a cyclonic anomaly

592   between 30°N and 50°N and an anticyclonic anomaly north of 50°N. The wind decreases over the

593   south-eastern boundary of the SPG and acts to weaken the SPG, favoring the northward move of a

594   salty and warm anomaly from the subtropical to the SPG, as described in Hatùn et al. (2005). The

595   decrease of the wind speed over the SPG, associated with the pressure anomalies, is able to weaken

596   the oceanic heat loss through the weakening of the latent and sensible heat fluxes (Fig. 15e, f). These

597   changes are strong in winter (DJF) but not consistent in summer (JJA) (not shown). CERFACS-HR is

598   able to reproduce the negative phase of the NAO. However this does not indicate a skill in predicting

599   the NAO since only concerning one start-date: the NAO is not skillful on interannual time-scales in

600   CERFACS-HR (not shown). We have computed the fraction of heat increase due to the net oceanic

601   heat transport (difference between the heat entering the southern boundary of the SPG and the heat

602   leaving its northern boundary) and the part coming from the atmosphere (the net heat surface fluxes).

603   One third of the OHC change is due to atmospheric changes over the 1998-2003 period. The

604   anomalies of net surface heat fluxes are thus not negligible.

605   The pressure anomaly may be due to a remote effect of the Atlantic SST (Czaja and Frankignoul, 2002)

606   and/or East Pacific SST (Huang et al., 2002; Brönnimann, 2007) on the NAO. With only five members

607   we cannot assess the robustness of such a relationship. This will need to be confirmed when larger

608   ensemble sizes will be routinely used with high-resolution models.

609

## 6. Conclusion

CERFACS-HR was first evaluated in term of its ability to predict temperature and precipitation. CERFACS-HR shows some skill at predicting the AMO, Indian and tropical Atlantic SST as shown by related indexes that exhibit reasonable correlations for 1-4 to 3-6 year lead-time. However CERFACS-HR presents no skill in predicting the eastern pacific SSTs. Skill in predicting SSTs was found in other studies (Chikamoto et al., 2013; Bellucci et al., 2013, 2014; Garcìa-Serrano et al., 2015; Karspeck et al., 2015; among others) and are mainly due to the warming trend. The skill improvement due to a spatial resolution increase is not assessed in this study. We nevertheless observe that the CERFACS-HR skill in predicting temperature is slightly weaker than in the literature. This can be explained by several factors: (i) the model tuning aims to improve the ability of a climate model to simulate the mean climate. High resolution coupled climate models are still computationally very expensive. Less sensitivity tests are thus performed with high-resolution models meaning that they are usually less optimized than low-resolution AOGCMs. The bias reduction associated with the time devoted to preliminary tests is important as it is related with prediction performance (DelSole and Shukla, 2010). (ii) We only ran a few start dates covering a period with no major volcanic events. In contrast, the studies based on the CMIP5 models used longer time periods (1960-2009), including major volcanic events (El-Chichón, Pinatubo) that improve model skill (Guemas et al., 2013a). (iii) Another difference comes from the start-date frequency, every 5 years in the CMIP5 protocol and every year in the present study. Garcìa-Serrano et al. (2015) have noticed different results using 5-year rather than 1-year date frequency.

Interestingly, the model is able to simulate the observed slow-down of the global warming rate over the recent period. This hiatus is associated with a negative phase of the PDO and an increase of the ocean heat uptake in both CERFACS-HR and observations. This is consistent with Meehl et al. (2011), Kosaka and Xie (2013), and Watanabe et al. (2013) among others. Since the observed PDO has experienced a negative phase during the 2000's (Trenberth and Fasullo, 2013), the question arises as to the main driver of the PDO variations, external forcing (e.g. the volcanic forcing due to the small to

27

636 medium size eruptions after 2005) or internal variability. In order to confirm the role of ocean

637 initialization versus the external forcing, we have performed two other sets (6 members each) of 10-

638 year simulations starting from November 2002 and differing by the applied volcanic forcing (a

639 constant weak background is used in one set while the observed forcing is used in the second set). The

640 two sets simulate a negative phase of the PDO (Monerie et al, to be submitted) with the main

641 contribution coming from the ocean initialization and an additional cooling due to volcanic forcing

642 (Fyfe et al. 2013; Santer et al. 2014; Haywood et al. 2014; Ridley et al., 2014; Schmidt et al., 2014;

643 Brühl et al., 2015; Mills et al., 2016).

644 The last part of the paper focuses on the case study of the North Atlantic warming of the mid 1990's.

645 The ability of the model to reproduce such an event is not directly related to the overall skill over the

646 North Atlantic. The simulation initialized in November 1995 simulates a quick warming, of about 1°C,

647 in just five years. The analysis of temperature, salt and circulation anomalies reveals that this warming

648 is associated with a strengthening of ocean dynamics. The subtropical gyre strengthens and its north-

649 western boundary transports the anomalously warm water to the SPG. The SPG then warms and

650 weakens. The increase of the North Atlantic temperature through a strengthening of the oceanic

651 circulation is consistent with Robson et al. (2012ab), Yeager et al. (2012) and Msadek et al. (2014a).

652 DH95 exhibits a robust strengthening of the heat transport by the anomalous advection of mean

653 temperature ($\overline{T}v^{'}$). The SPG warming predictability comes mainly from the initialization of the ocean.

654 Note that DH96 simulates an increase of the mean SPG temperature with a MHT change largely

655 dominated by the mean advection of temperature anomalies ($\overline{v}T^{'}$) north of 50°N (Supplementary

656 material S4). In DH96 the shift has already occurred and the temperature anomaly is incorporated in

657 the initial conditions. The warm anomaly is located inside the SPG at the beginning of DH96 and stay

658 inside the SPG during the simulation period (Fig. 9d). The reproducibility of the shift mechanisms is

659 thereby dependent on the initial conditions. Interestingly analysis of DH96 confirms that the SPG

660 warming may be sustained by the $\overline{v}T^{'}$ anomaly.

661    There is also a contribution from the net surface heat flux indicating a warming of the ocean by the

662    atmosphere over the 1998-2003 period. A negative phase of the NAO is simulated during this period

663    and favors a decrease of the wind speed over most of the SPG. This has multiple consequences: (i) the

664    latent and sensible heat fluxes weaken and reduce heat loss from the ocean (ii) the decrease of the

665    wind speed over the south-eastern boundary of the SPG acts on the SPG circulation, favoring the

666    northward move of the warm anomaly along the gyre eastern boundary. The warming of the SPG in

667    CERFACS-HR is thus due to both an increase of the oceanic heat transport and of the local change

668    due to atmospheric forcing, which sustains the positive anomaly of temperature. This is consistent

669    with the relation found in Bersh (2002), Bersh et al. (2007), Sarafanov et al. (2008), Lohmann et al.

670    (2009) and Barrier et al. (2014) between a NAO- phase and anomalously warm SSTs over the North

671    Atlantic Ocean. We hypothesize that this negative phase of the NAO acts to reduce the meridional heat

672    transport and the anomaly of mixed-layer over the Labrador Sea, in consistency with Ortega et al.

673    (2012).

674    In this study we show that the north Atlantic warming of the mid-1990's is predictable due to the

675    ocean initialization (ocean memory): DH95 is initialized with a strong AMOC and thus with an

676    anomalously strong oceanic heat convergence.

682

**Captions**
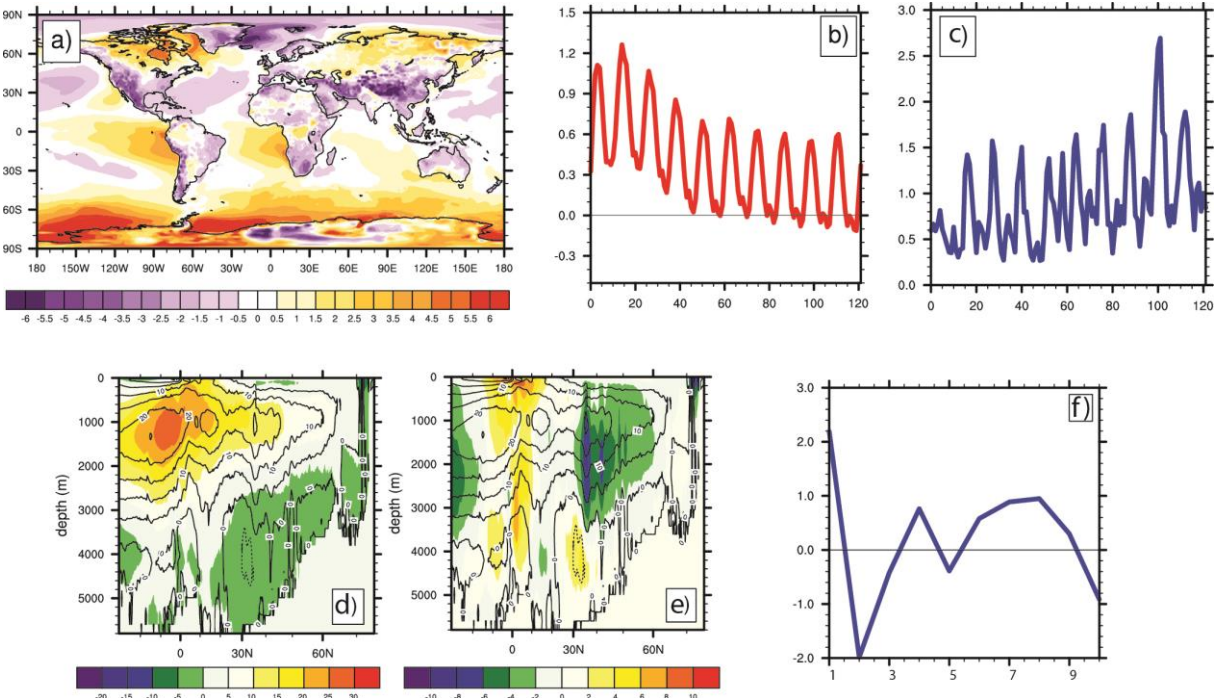
Figure 1: (a) CERFACS-HR minus ERAI difference in surface air-temperature (°C), averaged over all the start-dates and lead-time, (b) drift in global mean surface air temperature (°C), over the 122 months lead-time, (c) drift in Arctic sea-ice extent ($10^6$ km2) over the 122 months lead-time, (d) mean value of the Atlantic meridional overturning circulation (Sverdrup) computed with GLORYS2V3 (contour) and CERFACS-HR (color), and its (e) CERFACS-HR minus GLORYS2V3 difference (color, in Sverdrup). (f) Annual mean of the drift of the Atlantic meridional overturning circulation, i.e an index computed at 40°N and at a depth of 2000m, for the 10 years lead time.
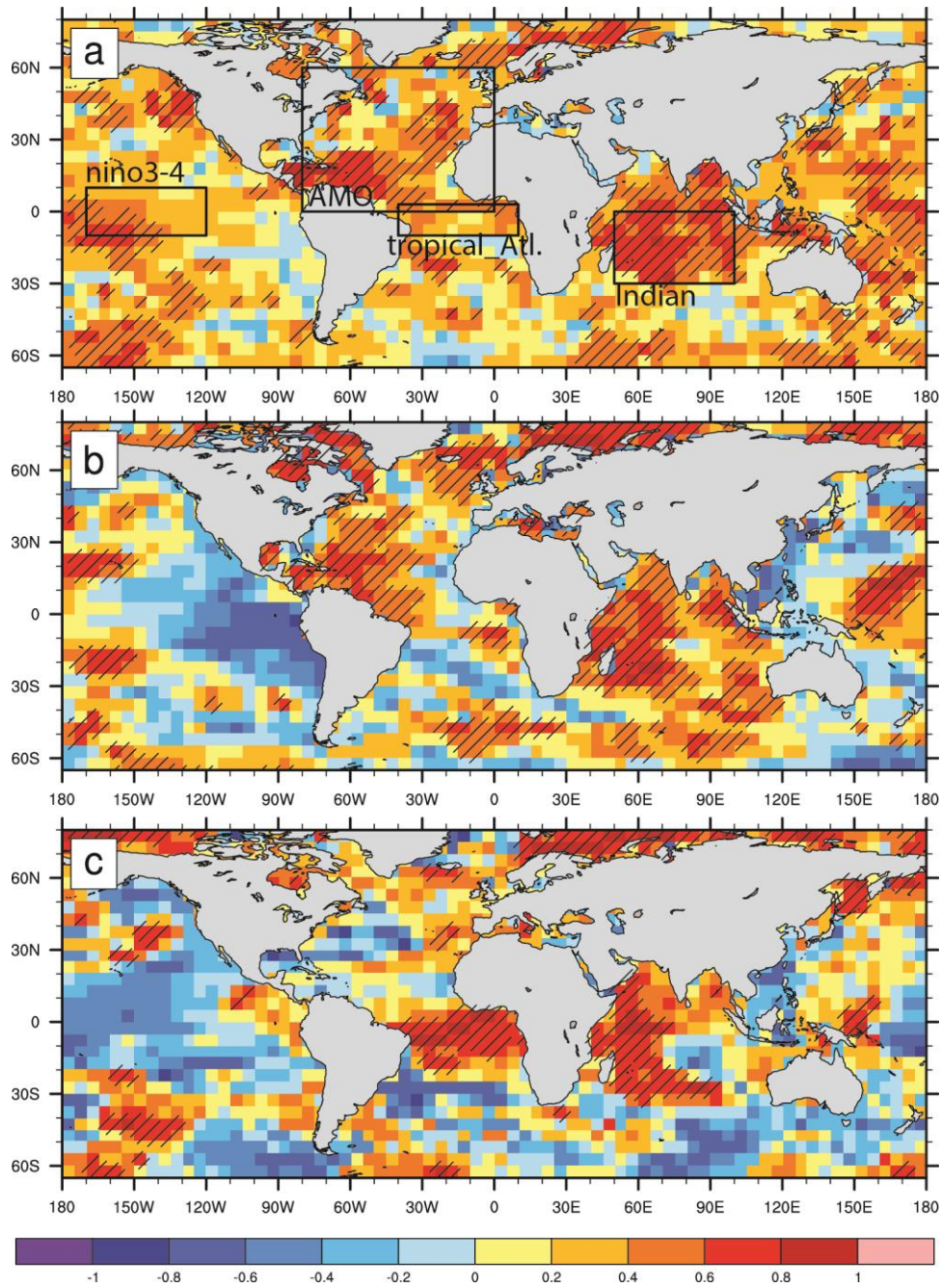
Figure 2: Anomaly coefficient correlation of the SST hindcasts (with respect to ERSST3b) for a) year 1, b) years 2-5 and c) 6-9. The trend is not removed. The hindcasts considered are the 17 5-year ones initialized every year from 1993 to 2009 for lead-times 1 year and 2-5 year, and the 10 10-year ones initialized every year from 1993 to 2002 for the lead-time 6-9 year. Hatching indicates that the ACC is positive and significant at the 95% confidence level according to a Monte-Carlo procedure.
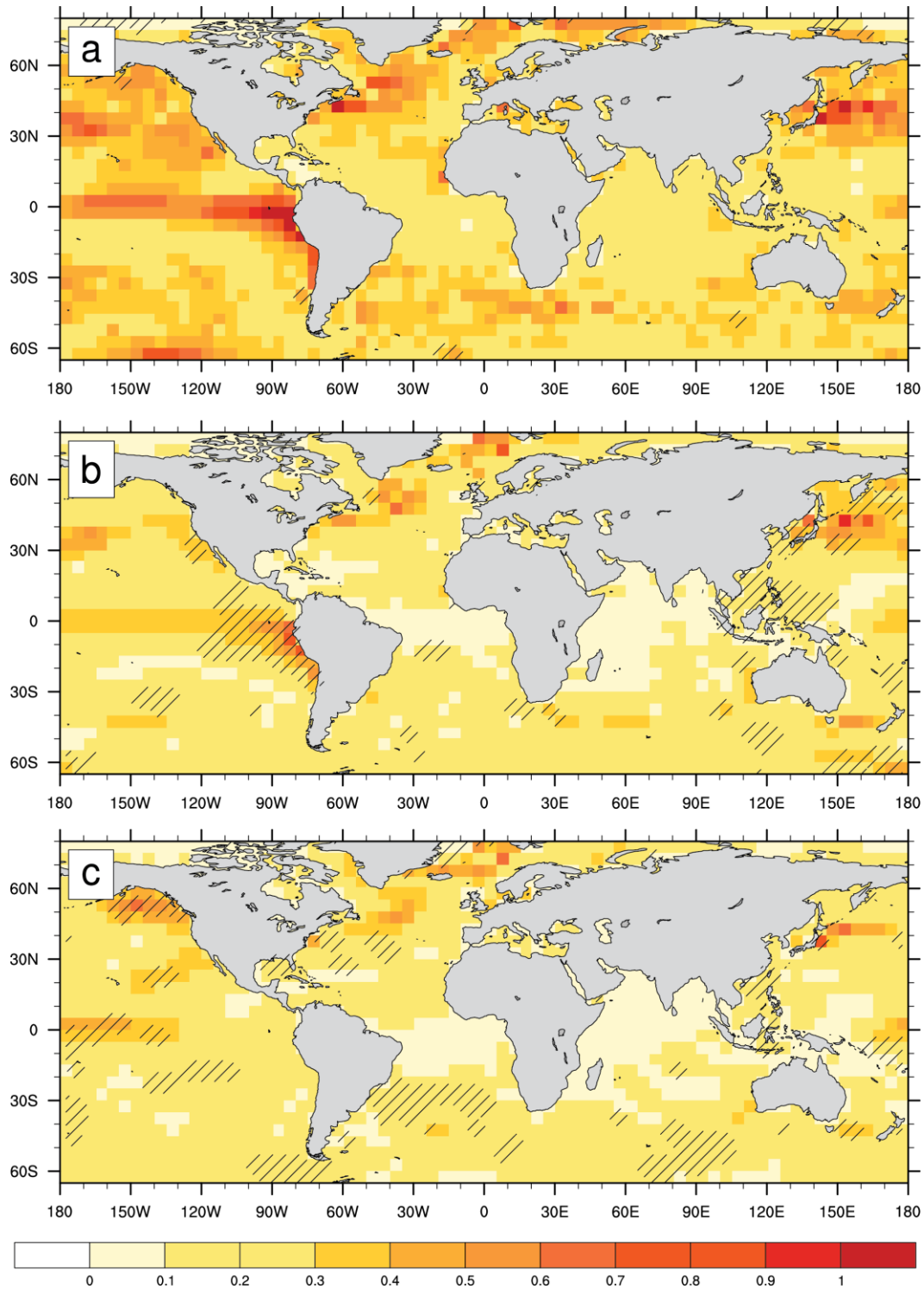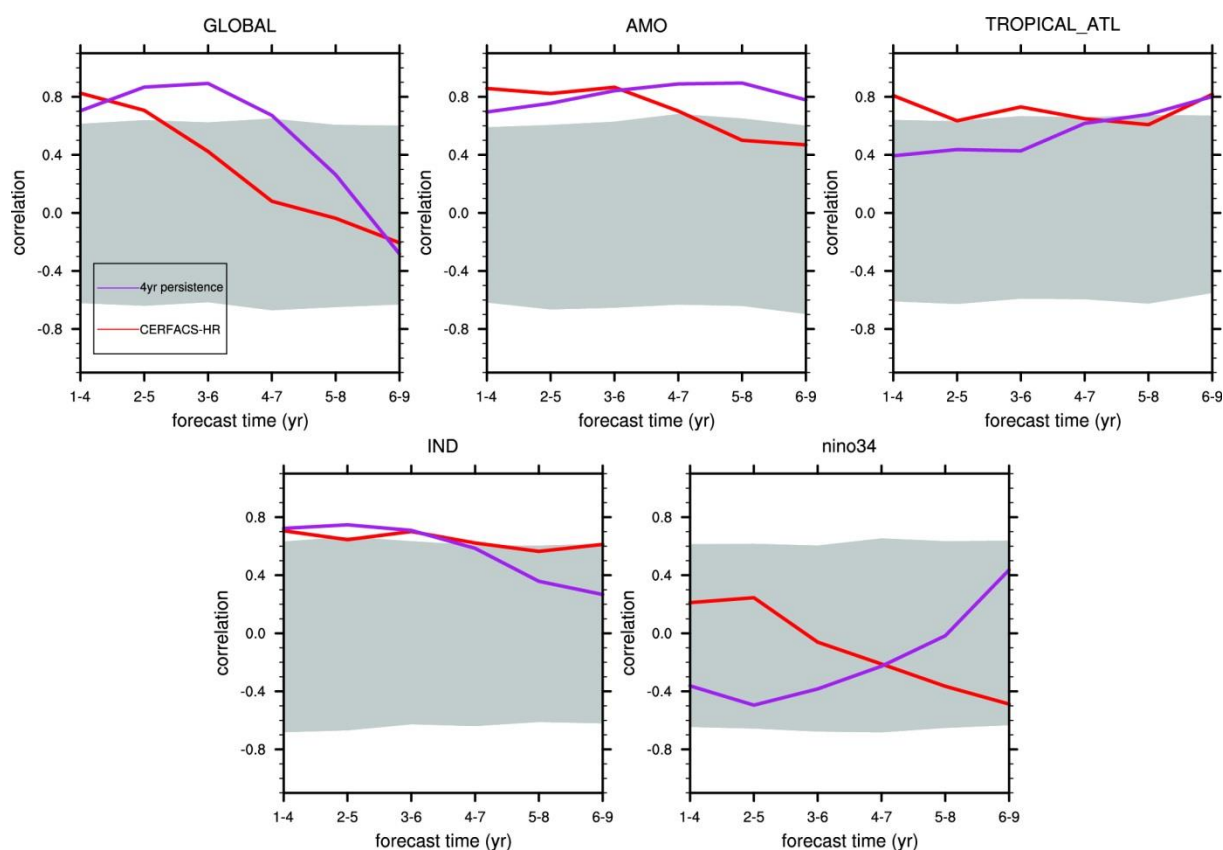
Figure 3: Root mean squared error, computed with the hindcasts and ERSST3b for the sea-surface temperature, and for a) year 1, b) years 2-5 and c) 6-9. The trend is not removed. The hindcasts considered are the 17 5-year ones initialized every year from 1993 to 2009 for lead-times 1 year and 2-5 year, and the 10 10-year ones initialized every year from 1993 to 2002 for the lead-time 6-9 year. Hatching indicates that the RMSE is not due to the sampling (a 95% confidence level according to a Monte-Carlo procedure

Figure 4: Anomaly coefficient correlation of the SST hindcasts (with respect to ERSST3b) for a) the global ocean, b) the AMO, c) the tropical Atlantic, d) the Indian Ocean and e) the nino3-4 index. The areas used to compute the indices are defined within the figure 1a. The red (purple) line denotes the hindacst (4-year persistence) skill (trends included). The hindcasts are initialized every year from 1993 to 2002 (i.e. 10 hindcasts of 10 years). The gray shading indicates that the ACC is not significant at the 95% confidence level according to a Monte-Carlo procedure.
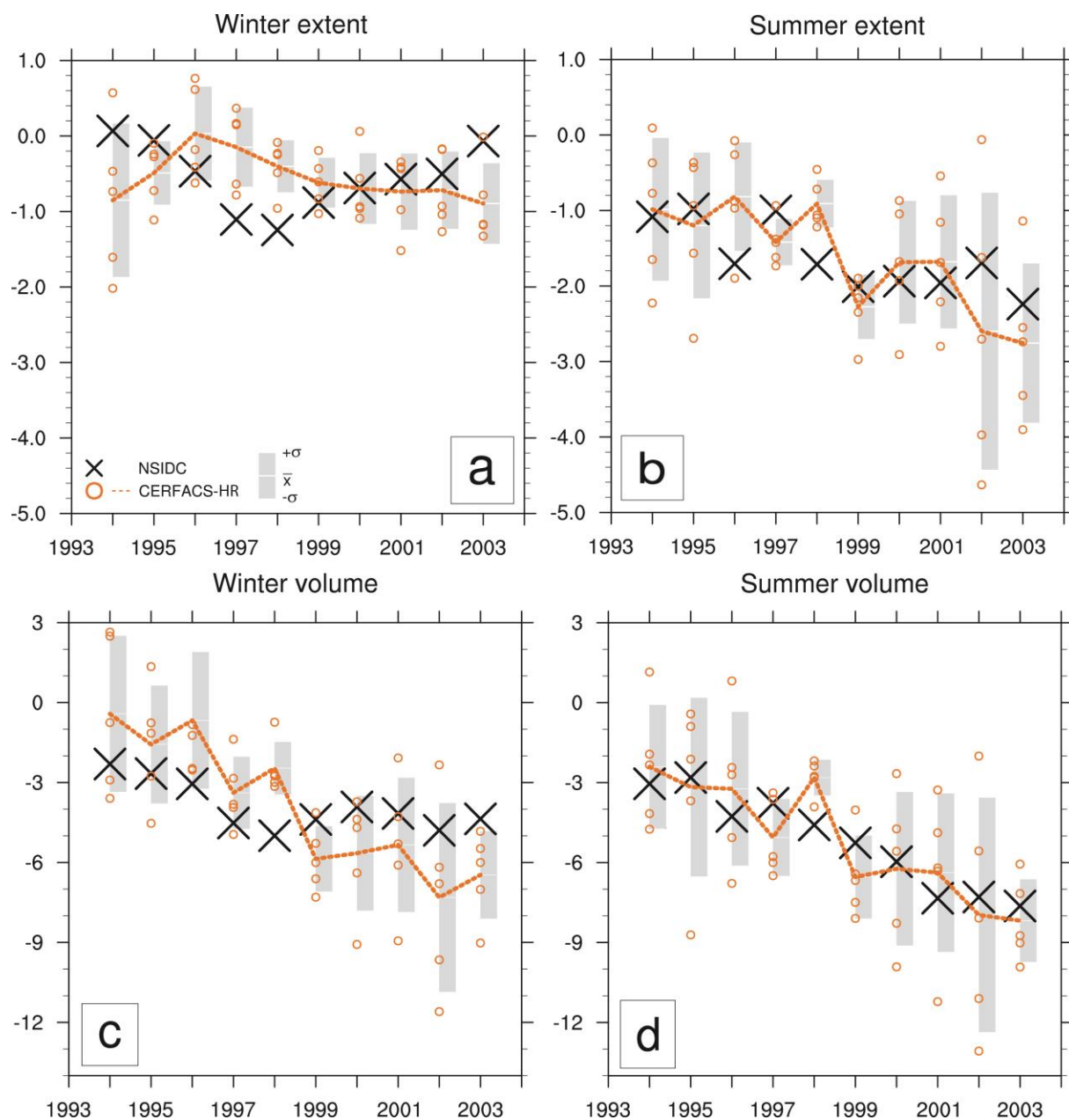
Figure 5: Ten-year linear trends for 1994-2003 to the 2003-2012 periods for a) winter SIE, b) summer SIE, c) winter SIV and d) winter SIV. The SIE trend is expressed in $10^6$ km$^2$ and the SIV trend in $10^3$ km$^3$. The observed SIE (from NSIDC) and SIV (from PIOMAS) are represented with a black cross. Each hindcast is represented by an orange circle and the ensemble mean by an orange dashed line. The gray areas represent the spread (more or less one standard deviation of the ensemble) computed from all the hindcasts of each start dates, and the white area is the average over all the hindcasts. The hindcasts are initialized every year from 1993 to 2002 (i.e. 10 hindcasts of 10 years with 5 members).

Figure 6: Ten-years linear trends (°C by ten years) for 1998-2009 to the 2003-2012 time-period for COWT (black cross), MLOST (blue square), GISTEMP (orange triangle) and for the hindcasts interpolated to the same horizontal resolution and

713  with the same spatial coverage of the observations. Each hindcast is represented by a circle and the ensemble mean by a

714  dashed line. The gray areas represent the spread (more or less one standard deviation) computed with the average of each

715  hindcast ensemble mean and the white area is the average over all the hindcasts. The hindcasts are initialized every year from

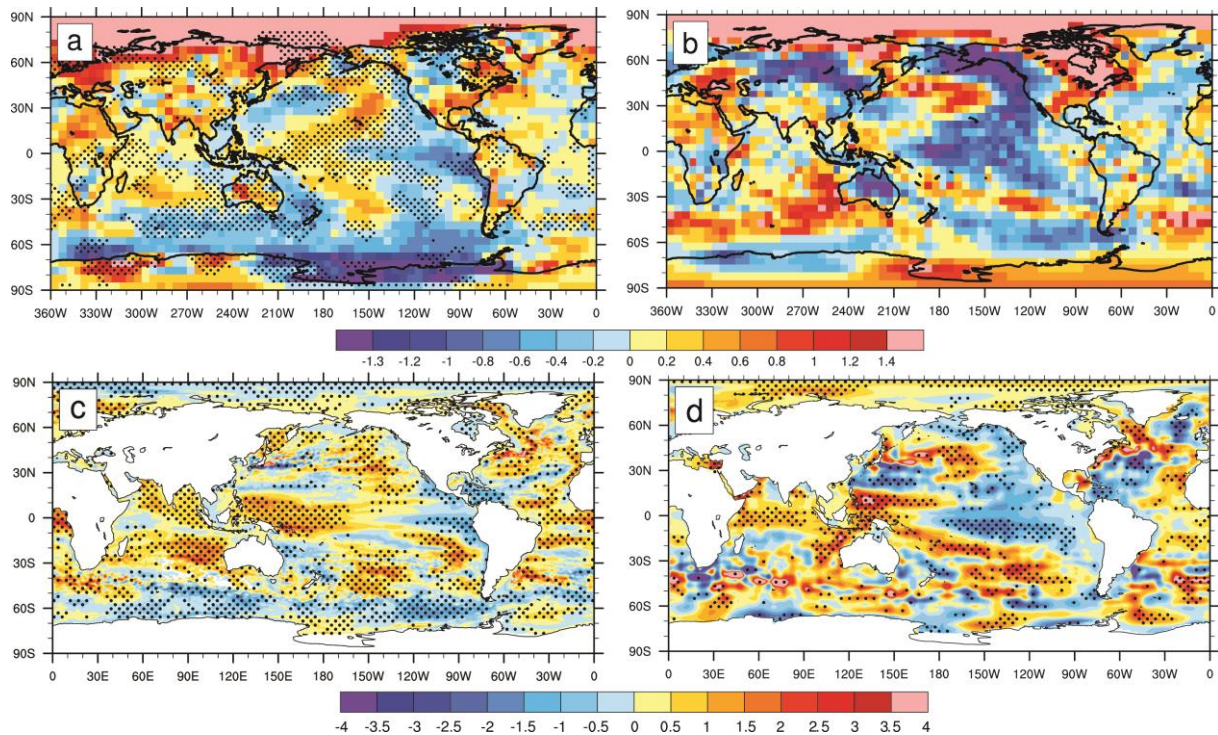716  1997 to 2002 (i.e. 6 hindcasts of 10 years with 5 members).

717  Figure 7: Ten-year linear trend (°C by decade) for the 2003-2012 period for the annual mean air-surface temperature for a)

718  CERFACS-HR (hindcast initialized in 2002) and b) COWT; stippling indicates zones where the trends are significantly

719  different (i.e. the observed trend is not included in the modeled mean trend +/- 1.64 standard deviation). Ten-year linear trend

720  ($10^7$ J.m$^{-2}$ by decade) for the heat content [0-2000m] with c) CERFACS-HR and d) EN4; stippling indicates that the trends

721  are significant at the 95% confidence level according to a Spearman's rank correlation test.
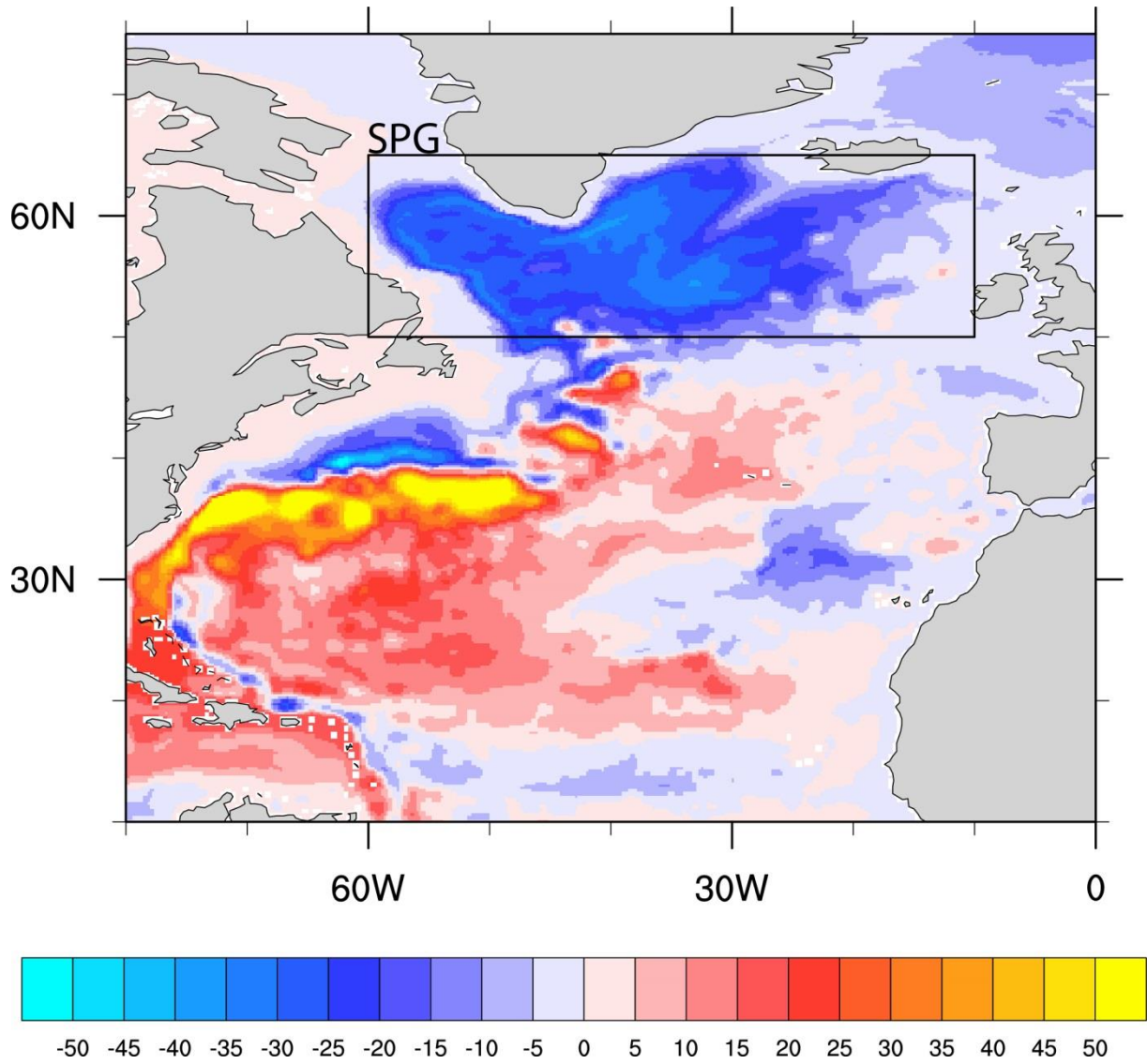
722  Figure 8: Annual mean barotropic streamfunction (Sv) averaged over all the hindcasts. The hindcasts are initialized every

723  year from 1993 to 2002 (i.e. 10 hindcasts of 10 years with 5 members). Negative (positive) values denote counterclockwise

724  (clockwise) circulation. The subpolar gyre area (SPG) is defined as the box represented in black: [60°W – 10°W; 50°N –
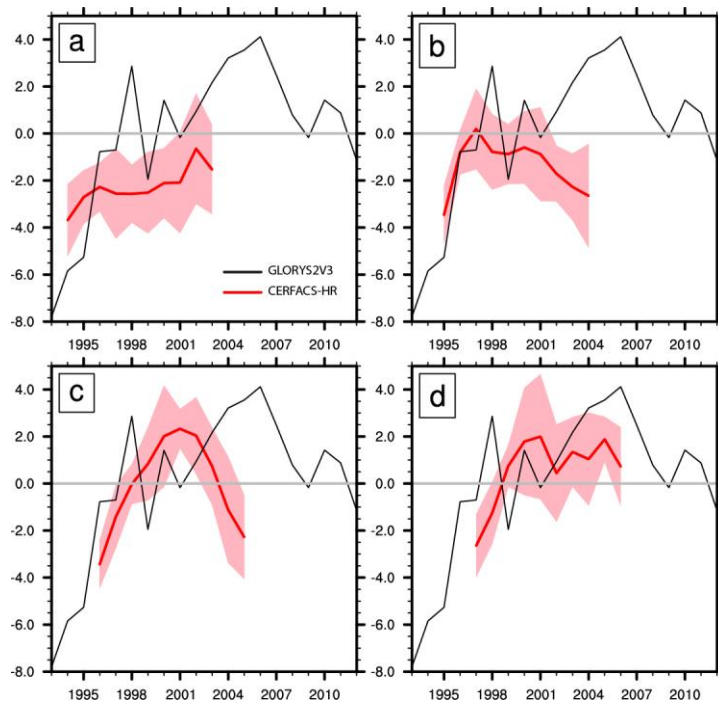
725  65°N].

726  Figure 9: Oceanic Heat Content ($10^{21}$ joules) [0-500m] over the SPG for GLORYS2V3 (black line) and a) DH93, b) DH94, c)

727  DH95 and d) DH96 (red line). The anomalies are computed with respect to the time average over the 10-years hindcasts. The

728  spread is computed on the 5 members as more or less 1 standard deviation (red shading). The zero line is in gray.

729  Figure 10: Temperature anomalies (°C) [0-500m] for DH95 and from 1996 to 2005. The anomalies are computed with

730  respect to the average over all the 10-years hindcasts. Dots indicate that anomalies are significant at the 95% confidence level

731  according to a Student t-test.

732  Figure 11: Same as Fig. 10 but for salt anomalies (psu).

733  Figure 12: Same as Fig. 10 but for barotropic streamfunction anomalies (Sv).

734  Figure 13: a) Meridional heat transport anomalies $tv$ (PW) and its decomposition into b) $\overline{T}v'$, c) $\overline{v}T'$ and d) $v'T'$ for

735  DH95. The anomalies are computed in regard to the average over all the 10-years hindcasts. Bars denote the monthly means

736  and prime the departure from the monthly means. Dots indicate that anomalies are significant at the 95% confidence level

737  according to a Student t-test.

738  Figure 14: Mixed layer depth anomalies (m) for the years 1-5 of DH95. The anomalies are computed with respect to the

739  average over all the 10-years hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to

740  a Student t-test.

741     Figure 15: a) Surface net fluxes (net shortwave + net longwave + sensible + latent) anomalies (W m-2) averaged over the

742     SPG, b) 3-8 year average of net surface flux anomalies and its decomposition into the c) net shortwave anomaly, d) net

743     longwave anomaly, e) sensible heat flux anomaly and f) latent heat flux anomaly. By convention positive (negative) values

744     denote downward (upward) anomalies. The anomalies are computed with respect to the average over all the 10-years

745     hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test.


746     Figure 16: 3-8 year average of sea-level pressure (hPa, color) and surface wind speed anomaly (m s-1, vector). The anomalies

747     are computed in regard to the average over all the 10-years hindcasts. Dots indicate that anomalies are significant at the 95%

748     confidence level according to a Student t-test. Only the significant wind anomalies are represented, in red (black) when

749     significant at the 95% (90%) confidence level according to a Student t-test.

750

**Figures**



Figure 1: (a) CERFACS-HR minus ERAI difference in surface air-temperature (℃), averaged over all the start-dates and lead-time, (b) drift in global mean surface air temperature (℃), over the 122 months lead-time, (c) drift in Arctic sea-ice extent ($10^6$ km2) over the 122 months lead-time, (d) mean value of the Atlantic meridional overturning circulation (Sverdrup) computed with GLORYS2V3 (contour) and CERFACS-HR (color), and its (e) CERFACS-HR minus GLORYS2V3 difference (color, in Sverdrup). (f) Annual mean of the drift of the Atlantic meridional overturning circulation, i.e an index computed at 40°N and at a depth of 2000m, for the 10 years lead time.

Figure 2: Anomaly coefficient correlation of the SST hindcasts (with respect to ERSST3b) for a) year 1, b) years 2-5 and c) 6-9. The trend is not removed. The hindcasts considered are the 17 5-year ones initialized every year from 1993 to 2009 for lead-times 1 year and 2-5 year, and the 10 10-year ones initialized every year from 1993 to 2002 for the lead-time 6-9 year. Hatching indicates that the ACC is positive and significant at the 95% confidence level according to a Monte-Carlo procedure.

Figure 3: Root mean squared error, computed with the hindcasts and ERSST3b for the sea-surface temperature, and for a) year 1, b) years 2-5 and c) 6-9. The trend is not removed. The hindcasts considered are the 17 5-year ones initialized every year from 1993 to 2009 for lead-times 1 year and 2-5 year, and the 10 10-year ones initialized every year from 1993 to 2002 for the lead-time 6-9 year. Hatching indicates that the RMSE is not due to the sampling (a 95% confidence level according to a Monte-Carlo procedure).

Figure 4: Anomaly coefficient correlation of the SST hindcasts (with respect to ERSST3b) for a) the global ocean, b) the AMO, c) the tropical Atlantic, d) the Indian Ocean and e) the nino3-4 index. The areas used to compute the indices are defined within the figure 1a. The red (purple) line denotes the hindcast (4-year persistence) skill (trends included). The hindcasts are initialized every year from 1993 to 2002 (i.e. 10 hindcasts of 10 years). The gray shading indicates that the ACC is not significant at the 95% confidence level according to a Monte-Carlo procedure.

788

789



Figure 5: Ten-year linear trends for 1994-2003 to the 2003-2012 periods for a) winter SIE, b) summer SIE, c) winter SIV and d) winter SIV. The SIE trend is expressed in $10^6$ km$^2$ and the SIV trend in $10^3$ km$^3$. The observed SIE (from NSIDC) and SIV (from PIOMAS) are represented with a black cross. Each hindcast is represented by an orange circle and the ensemble mean by an orange dashed line. The gray areas represent the spread (more or less one standard deviation of the ensemble) computed from all the hindcasts of each start date, and the white area is the average over all the hindcasts. The hindcasts are initialized every year from 1993 to 2002 (i.e. 10 hindcasts of 10 years with 5 members).
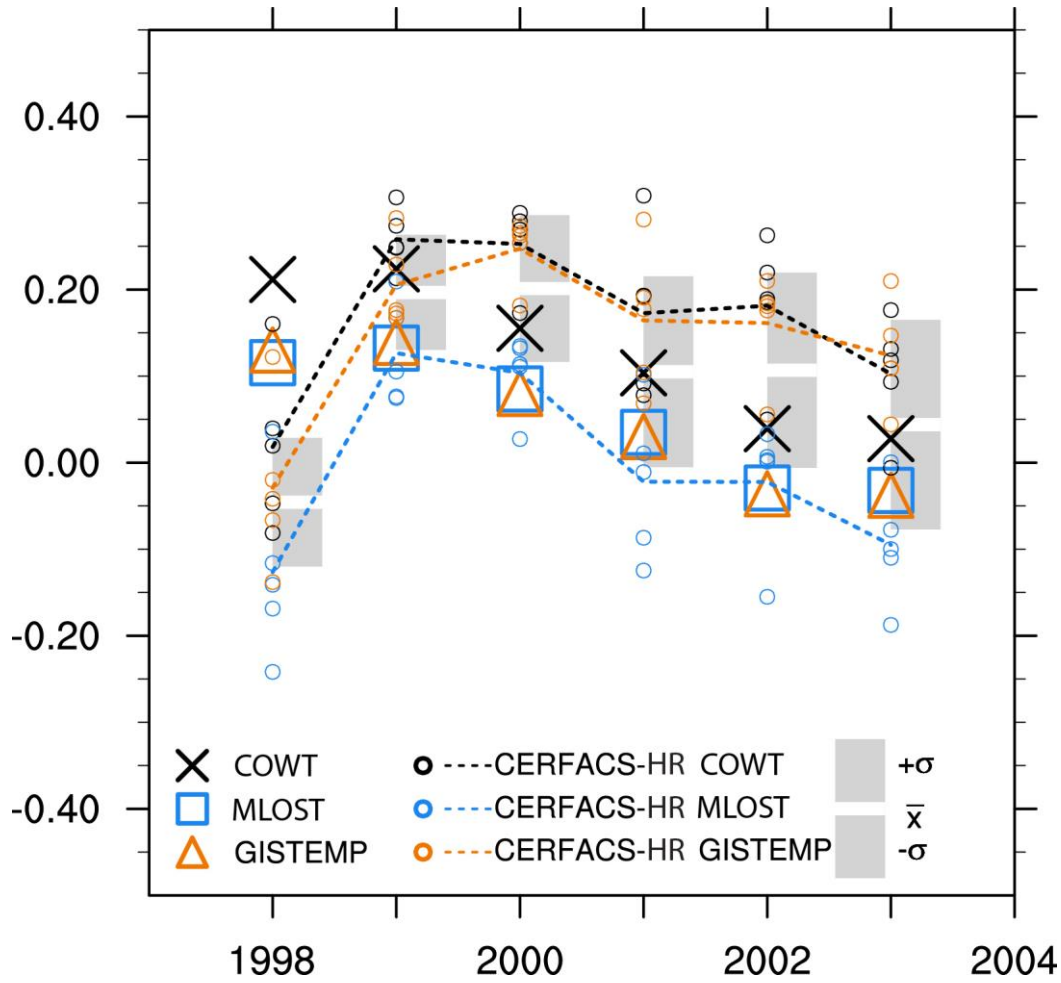
Figure 6: Ten-year linear trends (expressed in °C by ten years) for 1998-2007 up to the 2003-2012 time-period for COWT (black cross), MLOST (blue square), GISTEMP (orange triangle) and for the hindcasts interpolated to the same horizontal resolution and with the same spatial coverage the three sets of the observations. Each hindcast is represented by a circle and the ensemble mean by a dashed line. The gray areas represent the spread (more or less one standard deviation) computed from all the hindcasts (15 here) and the white area is the average over all the hindcasts. The spread and average are computed for each start dates. The hindcasts are initialized every year from 1997 to 2002 (i.e. 6 hindcasts of 10 years with 5 members).

Figure 7: Ten-year linear trend (°C by decade) for the 2003-2012 period for the annual mean air-surface temperature for a) CERFACS-HR (hindcast initialized in 2002) and b) COWT; stippling indicates zones where the trends are significantly different (i.e. the observed trend is not included in the modeled mean trend +/- 1.64 standard deviation). Ten-year linear trend ($10^7$ J.m$^{-2}$ by decade) for the heat content [0-2000m] with c) CERFACS-HR and d) EN4; stippling indicates that the trends are significant at the 95% confidence level according to a Spearman's rank correlation test.

Figure 8: Annual mean barotropic streamfunction (Sv) averaged over all the hindcasts. The hindcasts are initialized every year from 1993 to 2002 (i.e. 10 hindcasts of 10 years with 5 members). Negative (positive) values denote counterclockwise (clockwise) circulation. The subpolar gyre area (SPG) is defined as the box represented in black: [60°W – 10°W; 50°N – 65°N].

827

Figure 9: Oceanic Heat Content ($10^{21}$ joules) [0-500m] over the SPG for GLORYS2V3 (black line) and a) DH93, b) DH94, c)

DH95 and d) DH96 (red line). The anomalies are computed with respect to the time average over the 10-year hindcasts. The

spread is computed on the 5 members as more or less 1 standard deviation (red shading). The zero line is in gray.

842
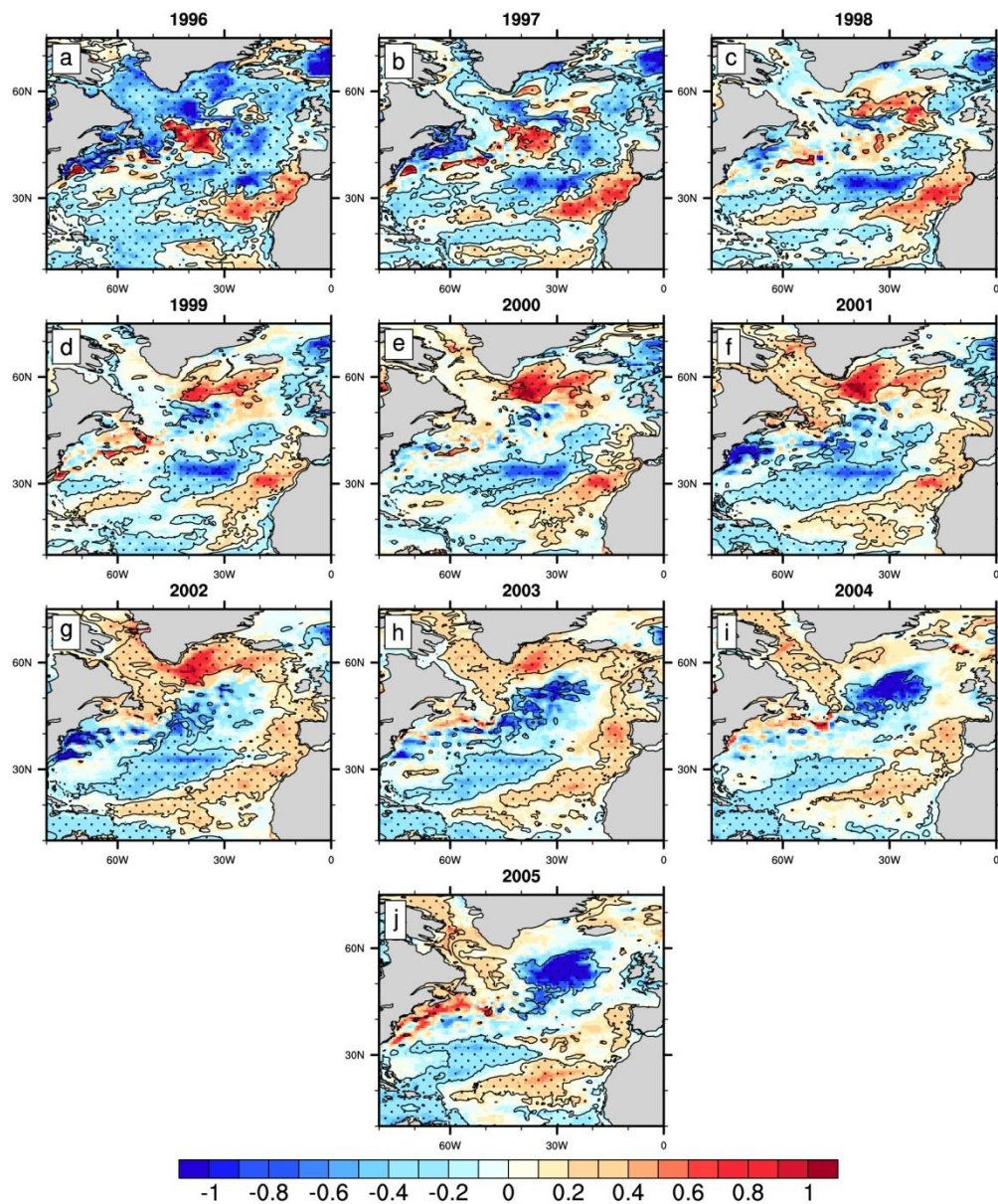


Figure 10: Temperature anomalies (°C) [0-500m] for DH95 from 1996 to 2005. The anomalies are computed with respect to the time average over the 10-year hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test.

843
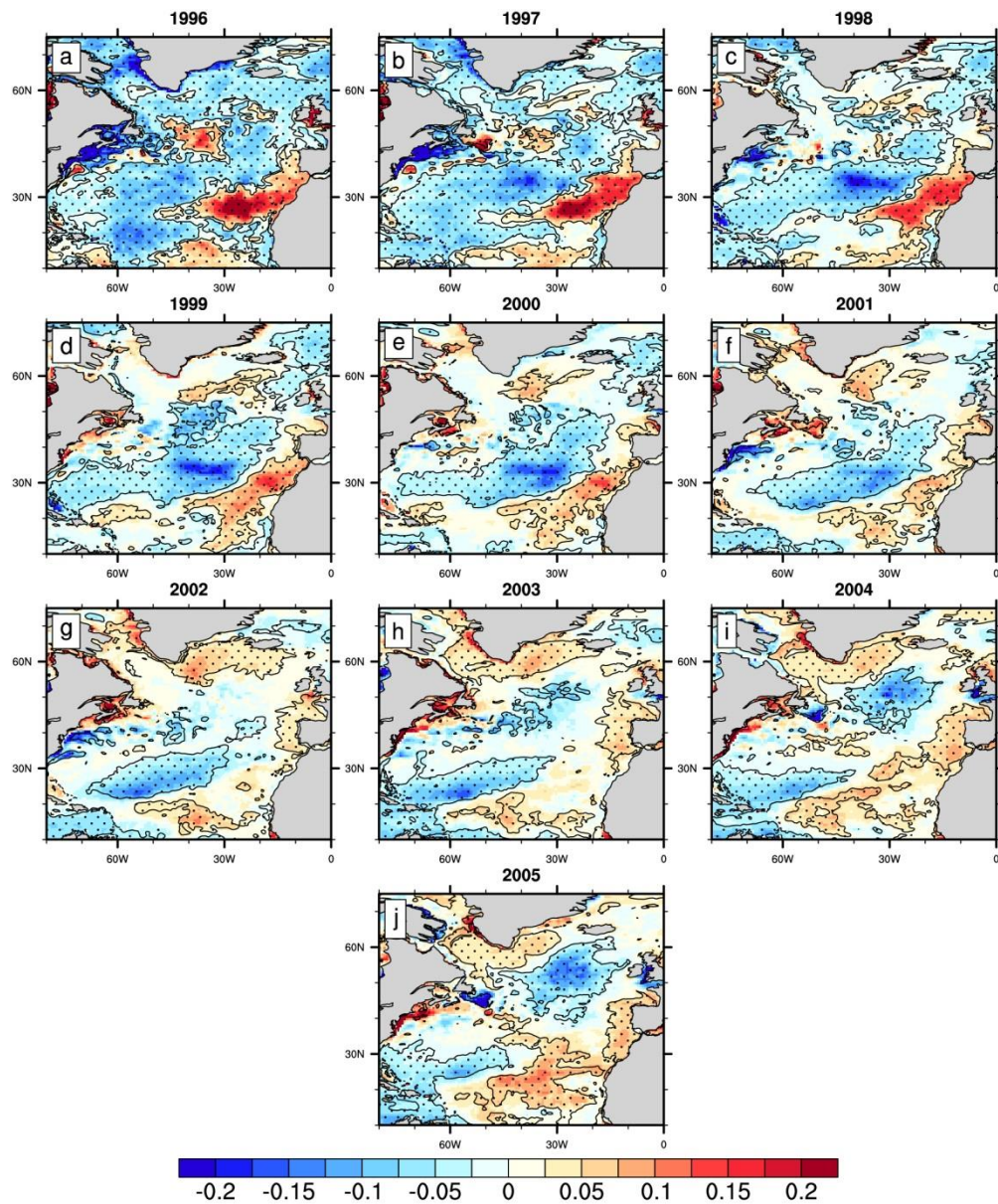
844

845

846

847

848

849

850

Figure 11: Same as Fig. 10 but for salt anomalies (psu).

851

852

853

854

855

856

857

858
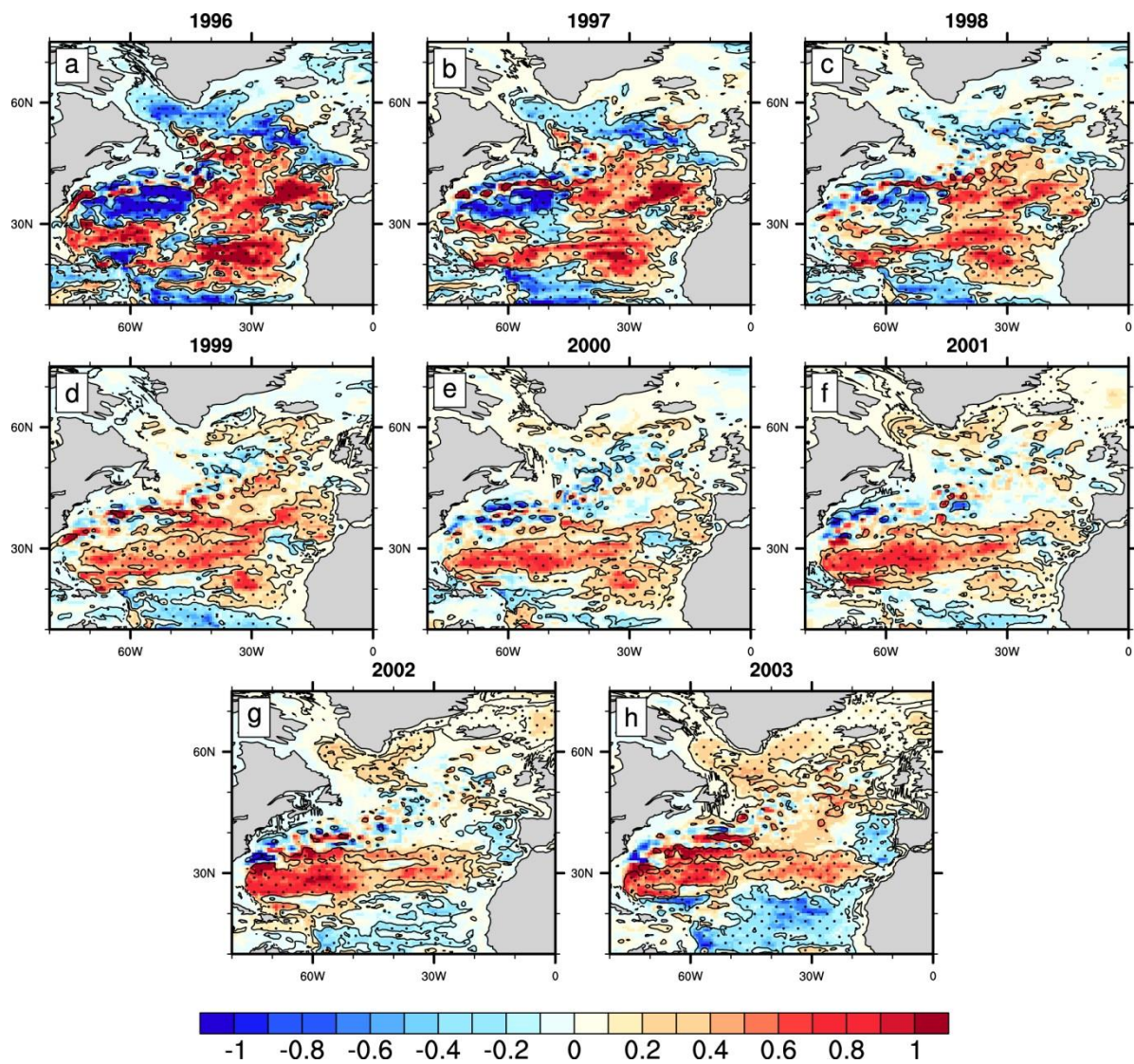
Figure 12: Same as Fig. 10 but for barotropic streamfunction anomalies (Sv).
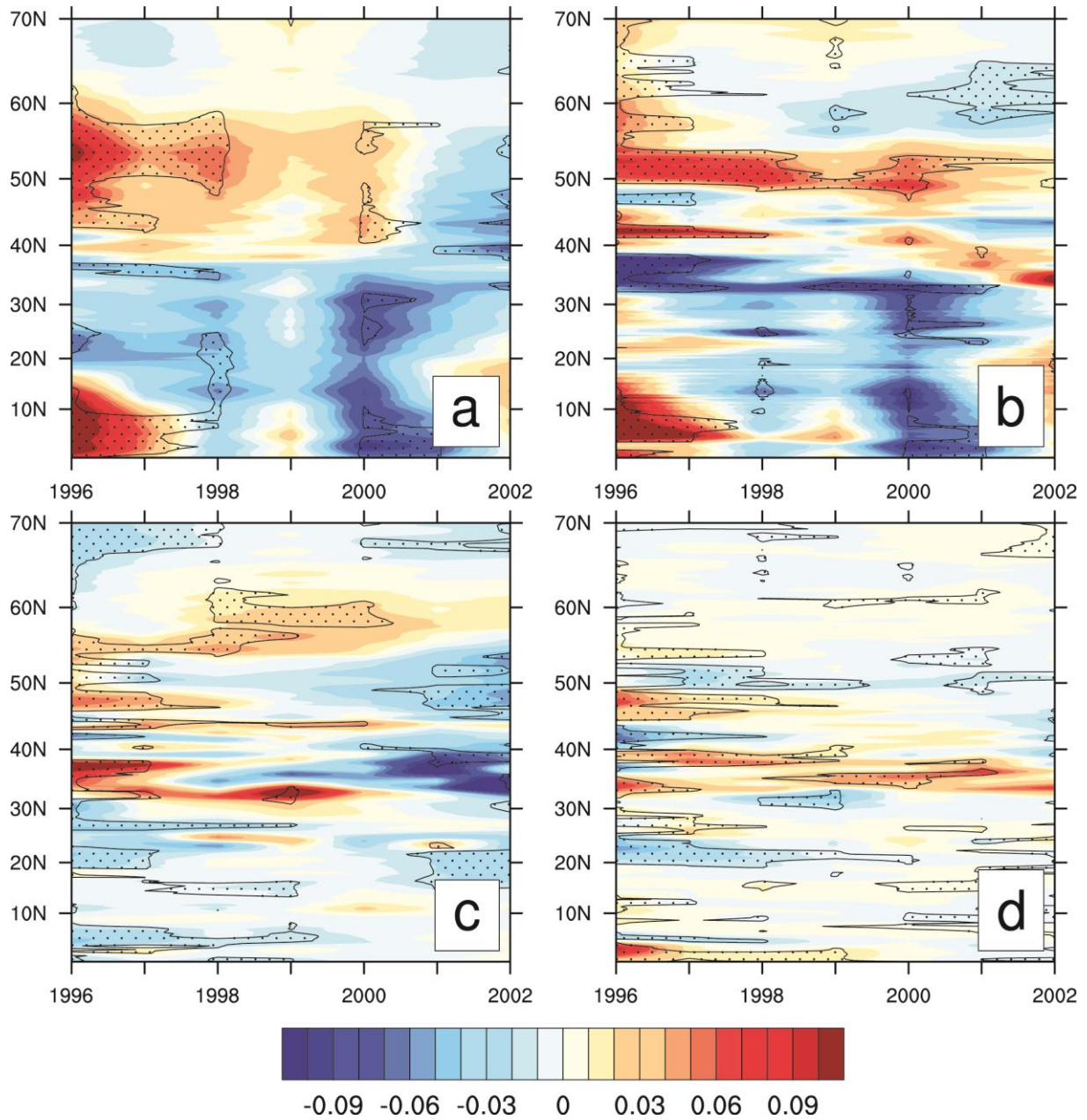
Figure 13: Evolution of vertically integrated (in PW) a) meridional heat transport anomalies $vT$ and its decomposition into b) $\overline{T}v'$, c) $\overline{v}T'$ and d) $v'T'$ for DH95. The anomalies are computed with respect to the average over all the 10-year hindcasts. Bars denote the monthly means and prime the departure from the monthly means. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test.
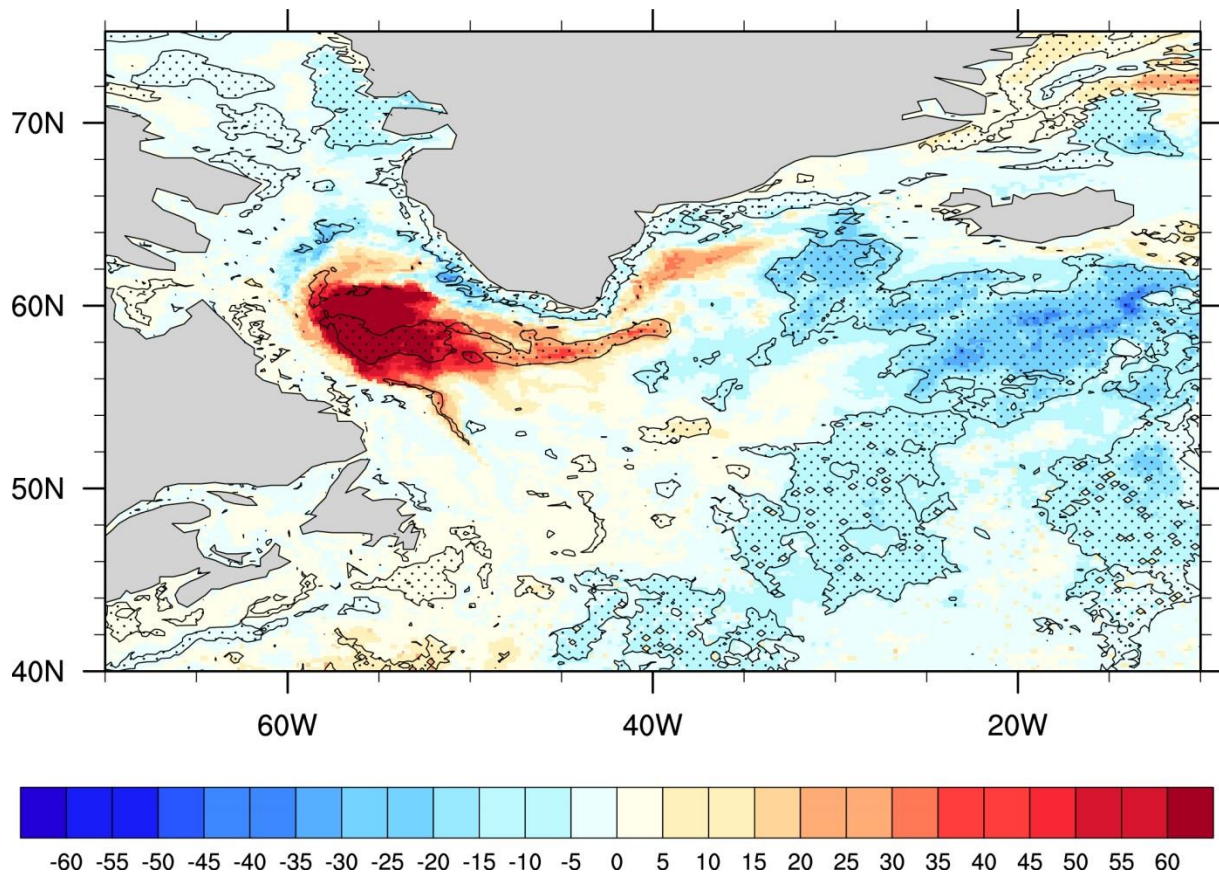
Figure 14: Mixed layer depth anomalies (m) for the years 1996-2000 of DH95. The anomalies are computed with respect to the average over all the 10-year hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test.
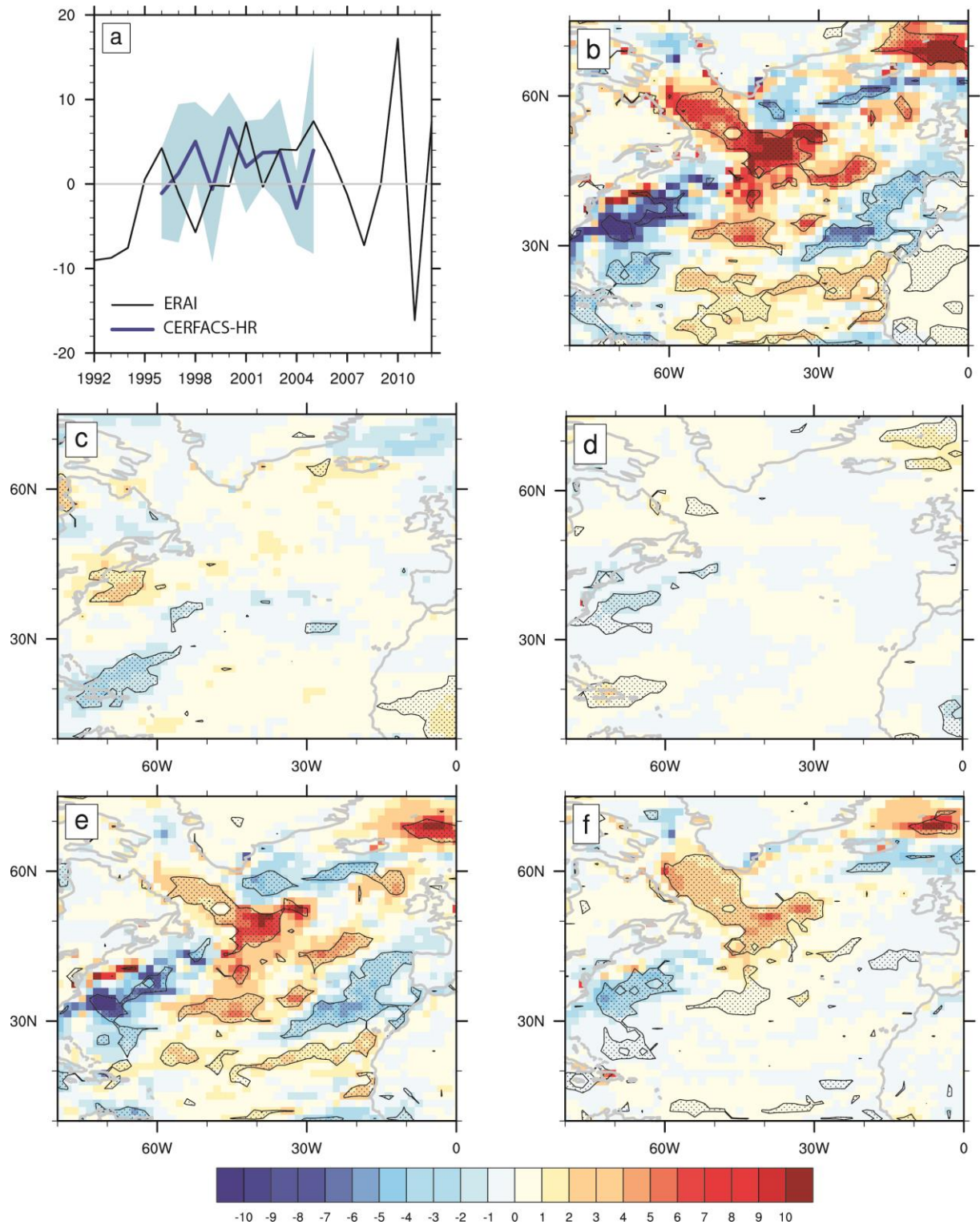
Figure 15: a) Surface net fluxes (net shortwave + net longwave + sensible + latent) anomalies (W.m$^{-2}$) averaged over the SPG, b) 3-8 year average of net surface flux anomalies and its decomposition into the c) net shortwave, d) net longwave, e) sensible heat flux and f) latent heat flux anomalies. By convention positive (negative) values denote downward (upward) anomalies. The anomalies are computed with respect to the average over all the 10-year hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test.
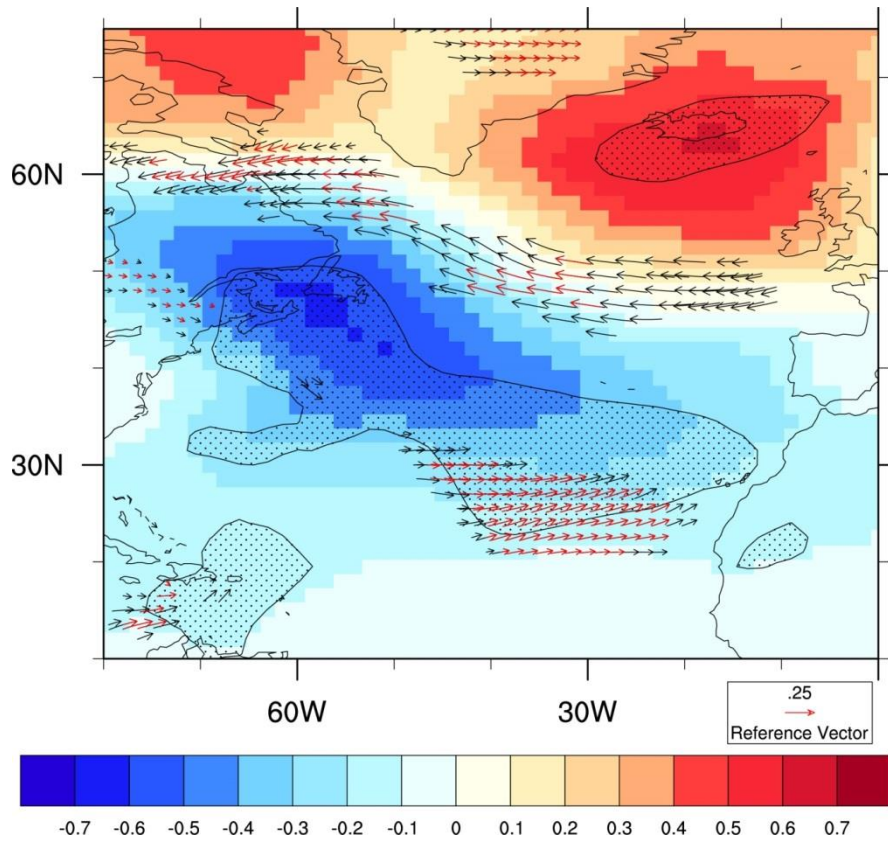
Figure 16: 1998-2003 (3-8 year lead time) average of sea-level pressure (hPa, color) and surface wind speed anomaly (m s-1, vector). The anomalies are computed with respect to the average over all the 10-years hindcasts. Dots indicate that anomalies are significant at the 95% confidence level according to a Student t-test. Only the significant wind anomalies are represented, in red (black) when significant at the 95% (90%) confidence level according to a Student t-test.

**References**

Bacmeister JT et al. (2014) Exploratory high-resolution climate simulations using the Community Atmosphere Model (CAM). Journal of Climate 27:3073-3099

Balmaseda M, Mogensen K, Molteni F, Weaver A (2010) The NEMOVAR-COMBINE ocean re-analysis. COMBINE Technical Report No 1

Balmaseda MA, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. Quarterly Journal of the Royal Meteorological Society 139:1132-1161

Barnier B et al. (2006) Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. Ocean dynamics 56:543-567

Barrier N, Cassou C, Deshayes J, Treguier A-M (2014) Response of North Atlantic Ocean circulation to atmospheric weather regimes. Journal of Physical Oceanography 44:179-201

Barrier N, Deshayes J, Treguier A-M, Cassou C (2015) Heat budget in the North Atlantic subpolar gyre: Impacts of atmospheric weather regimes on the 1995 warming event. Progress in Oceanography 130:75-90

Bellucci A et al. (2013) Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses. Climate Dynamics 40:1483-1497

Bellucci A et al. (2014) An assessment of a multi-model ensemble of decadal climate predictions. Climate Dynamics 44:2787-2806

Bellucci A et al. (2015) Advancements in decadal climate predictability: The role of nonoceanic drivers. Reviews of Geophysics 53: 165-202. Doi:10.1002/2014RG000473

Bersch M (2002) North Atlantic Oscillation–induced changes of the upper layer circulation in the northern North Atlantic Ocean. Journal of Geophysical Research: Oceans (1978–2012) 107:20-21-20-11

Bersch M, Yashayaev I, Koltermann KP (2007) Recent changes of the thermohaline circulation in the subpolar North Atlantic. Ocean dynamics 57:223-235

Blanchard‑Wrigglesworth E, Bitz C, Holland M (2011) Influence of initial conditions and climate forcing on predicting Arctic sea ice. Geophysical Research Letters 38

Boer G (2000) A study of atmosphere-ocean predictability on long time scales. Climate Dynamics 16:469-477

Brönnimann S (2007) Impact of El Niño–Southern Oscillation on European climate. Reviews of Geophysics 45

Brühl C, Lelieveld J, Tost H, Höpfner M, Glatthor N (2015) Stratospheric sulfur and its implications for radiative forcing simulated by the chemistry climate model EMAC. J. Geophys. Res. Atmos. 120: 2103–2118. doi: 10.1002/2014JD022430.

Chang C-Y, Nigam S, Carton JA (2008) Origin of the springtime westerly bias in equatorial Atlantic surface winds in the community atmosphere model version 3 (CAM3) simulation. Journal of Climate 21:4766-4778

Chikamoto Y et al. (2013) An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC. Climate Dynamics 40:1201-1222

Collins M (2002) Climate predictability on interannual to decadal time scales: the initial value problem. Climate Dynamics 19:671-692

Collins M et al. (2006) Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study. Journal of Climate 19:1195-1203

Cowtan K, Way RG (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. Quarterly Journal of the Royal Meteorological Society 140:1935-1944

Cox P, Stephenson D (2007) A Changing Climate for Prediction. Science 317:207-208

Czaja A, Frankignoul C (2002) Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation. Journal of Climate 15:606-623

Day J, Hawkins E, Tietsche S (2014) Will Arctic sea ice thickness initialization improve seasonal forecast skill? Geophysical Research Letters 41:7566-7575

953 Dee D et al. (2011) The ERA‑Interim reanalysis: Configuration and performance of the data
954     assimilation system. Quarterly Journal of the Royal Meteorological Society 137:553-597
955 DelSole T, Shukla J (2010) Model fidelity versus skill in seasonal forecasting. Journal of Climate
956     23:4794-4806
957 Delworth TL et al. (2012) Simulated climate and climate change in the GFDL CM2. 5 high-resolution
958     coupled climate model. Journal of Climate 25:2755-2781
959 Doblas-Reyes F et al. (2013a) Initialized near-term regional climate change prediction. Nature
960     communications 4:1715
961 Doblas‑Reyes FJ, García‑Serrano J, Lienert F, Biescas AP, Rodrigues LR (2013b) Seasonal climate
962     predictability and forecasting: status and prospects. Wiley Interdisciplinary Reviews: Climate
963     Change 4:245-268
964 Douville H, Voldoire A, Geoffroy O (2015) The recent global warming hiatus: What is the role of
965     Pacific variability? Geophysical Research Letters 42:880-888
966 England MH et al. (2014) Recent intensification of wind-driven circulation in the Pacific and the
967     ongoing warming hiatus. Nature Climate Change 4:222-227
968 Ferry N et al. (2012) GLORYS2V1 global ocean reanalysis of the altimetric era (1992–2009) at meso
969     scale. Mercator Ocean–Quaterly Newsletter 44
970 Fetterer F, Knowles K, Meier W, Savoie M (2009) Sea ice index. National Snow and Ice Data Center,
971     Boulder, CO, USA. Digital Media, updated
972 Fyfe J, Salzen K, Cole J, Gillett N, Vernier JP (2013) Surface response to stratospheric aerosol
973     changes in a coupled atmosphere–ocean model. Geophysical Research Letters 40:584-588
974 García-Serrano J, Guemas V, Doblas-Reyes F (2015) Added-value from initialization in predictions of
975     Atlantic multi-decadal variability. Climate Dynamics 44:2539-2555
976 Gent PR, Yeager SG, Neale RB, Levis S, Bailey DA (2010) Improvements in a half degree
977     atmosphere/land version of the CCSM. Climate Dynamics 34:819-833
978 Germe A, Chevallier M, y Mélia DS, Sanchez-Gomez E, Cassou C (2014) Interannual predictability of
979     Arctic sea ice in a global climate model: regional contrasts and temporal evolution. Climate
980     Dynamics 43:2519-2538
981 Goddard L et al. (2013) A verification framework for interannual-to-decadal predictions experiments.
982     Climate Dynamics 40:245-272
983 Good SA, Martin MJ, Rayner NA (2013) EN4: quality controlled ocean temperature and salinity
984     profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical
985     Research: Oceans 118:6704-6716
986 Griffies SM, Bryan K (1997) Predictability of North Atlantic multidecadal climate variability. Science
987     275:181-184
988 Guemas V, Corti S, García-Serrano J, Doblas-Reyes F, Balmaseda M, Magnusson L (2013a) The
989     Indian Ocean: the region of highest skill worldwide in decadal climate prediction*. Journal of
990     Climate 26:726-739
991 Guemas V, Doblas-Reyes FJ, Andreu-Burillo I, Asif M (2013b) Retrospective prediction of the global
992     warming slowdown in the past decade. Nature Climate Change 3:649-653
993 Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. Reviews of Geophysics
994     48
995 Hátún H, Sandø AB, Drange H, Hansen B, Valdimarsson H (2005) Influence of the Atlantic subpolar
996     gyre on the thermohaline circulation. Science 309:1841-1844
997 Haywood JM, Jones A, Jones GS (2014) The impact of volcanic eruptions in the period 2000–2013 on
998     global mean temperature trends evaluated in the HadGEM2‑ES climate model. Atmospheric
999     Science Letters 15:92-96
1000 Hibbard KA, Meehl GA, Cox PM, Friedlingstein P (2007) A strategy for climate change stabilization
1001     experiments. Eos, Transactions American Geophysical Union 88:217-221
1002 Huang B, Schopf PS, Pan Z (2002) The ENSO effect on the tropical Atlantic variability: A regionally
1003     coupled model study. Geophysical Research Letters 29:35-31-35-34
1004 Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and
1005     precipitation. Science 269:676-679

1006 Hurrell JW, Kushnir Y, Ottersen G, Visbeck M (2003) An overview of the North Atlantic oscillation.
1007         Geophysical Monograph-American Geophysical Union 134:1-36
1008 ICPO (2011) Data and bias correction for decadal climate predictions. International CLIVAR Project
1009         Office Publication Series 150:5 pp. [Available online at http://www.wcrp-climate.
1010 org/decadal/references/DCPP_Bias_Correction.pdf.]
1011 Jia L et al. (2015) Improved seasonal prediction of temperature and precipitation over land in a high-
1012         resolution GFDL climate model. Journal of Climate 28:2044-2062
1013 Johns WE et al. (2011) Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5 N.
1014         Journal of Climate 24:2429-2449
1015 Jung T et al. (2012) High-resolution global climate simulations with the ECMWF model in Project
1016         Athena: Experimental design, model climate, and seasonal forecast skill. Journal of Climate
1017         25:3155-3172
1018 Karspeck A, Yeager S, Danabasoglu G, Teng H (2015) An evaluation of experimental decadal
1019         predictions using CCSM4. Climate Dynamics 44:907-923
1020 Keenlyside N, Latif M, Jungclaus J, Kornblueh L, Roeckner E (2008) Advancing decadal-scale
1021         climate prediction in the North Atlantic sector. Nature 453:84-88
1022 Kim H-M, Webster PJ, Curry JA (2012) Evaluation of short-term climate change prediction in multi-
1023         model CMIP5 decadal hindcasts. Geophysical Research Letters 39:n/a-n/a.
1024         doi:10.1029/2012gl051644
1025 Kirtman BP et al. (2012) Impact of ocean model resolution on CCSM climate simulations. Climate
1026         Dynamics 39:1303-1328
1027 Kopp G, Lean J-L (2011) A new, lower value of total solar irradiance: Evidence and climate
1028 significance. Geophysical Research Letters 38. L01706. doi:10.1029/2010GL045777.
1029 Kosaka Y, Xie S-P (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling.
1030         Nature 501:403-407
1031 Koster RD, Suarez MJ (2003) Impact of land surface initialization on seasonal precipitation and
1032         temperature prediction. Journal of hydrometeorology 4:408-423
1033 Kruschke T, Rust HW, Kadow C, Müller WA, Pohlmann H, Leckebusch GC, Ulbrich U (2015)
1034         Probabilistic evaluation of decadal predictions of Northern Hemisphere winter storms.
1035         Meteorol. Z. doi:10.1127/metz/2015/0641
1036 Lee S-K, Park W, Baringer MO, Gordon AL, Huber B, Liu Y (2015) Pacific origin of the abrupt
1037         increase in Indian Ocean heat content during the warming hiatus. Nature Geoscience
1038 Lellouche J-M et al. (2013) Evaluation of global monitoring and forecasting systems at Mercator
1039         Océan. Ocean Science 9:57
1040 Lindsay R et al. (2012) Seasonal forecasts of Arctic sea ice initialized with observations of ice
1041         thickness. Geophysical Research Letters 39
1042 Lohmann K, Drange H, Bentsen M (2009) Response of the North Atlantic subpolar gyre to persistent
1043         North Atlantic oscillation like forcing. Climate Dynamics 32:273-285
1044 Madec G (2008) NEMO ocean engine. Note du Pole de mode´lisation, Institut Pierre-Simon Laplace
1045         (IPSL), France No 27 ISSN:No 1288–1619
1046 Matei D, Pohlmann H, Jungclaus J, Müller W, Haak H, Marotzke J (2012) Two tales of initializing
1047         decadal climate prediction experiments with the ECHAM5/MPI-OM model. Journal of
1048         Climate 25:8502-8523
1049 McClean JL et al. (2011) A prototype two-decade fully-coupled fine-resolution CCSM simulation.
1050         Ocean Modelling 39:10-30
1051 Meehl GA, Arblaster JM, Fasullo JT, Hu A, Trenberth KE (2011) Model-based evidence of deep-
1052         ocean heat uptake during surface-temperature hiatus periods. Nature Climate Change 1:360-
1053         364
1054 Meehl GA, Teng H, Arblaster JM (2014) Climate model simulations of the observed early-2000s
1055         hiatus of global warming. Nature Climate Change
1056 Mehta VM, Wang H, Mendoza K (2013) Decadal predictability of tropical basin average and global
1057         average sea surface temperatures in CMIP5 experiments with the HadCM3, GFDL‐CM2. 1,
1058         NCAR‐CCSM4, and MIROC5 global Earth System Models. Geophysical Research Letters
1059         40:2807-2812

Mills MJ et al (2016) Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1(WACCM). J. Geophys. Res. Atmos. 121:2332–2348. doi:10.1002/2015JD024290.

Mochizuki T et al. (2012) Decadal prediction using a recent series of MIROC global climate models. Journal of the Meteorological Society of Japan 90A:373-383

Monerie P-A, Moine M-P, Terray L, Valcke S. Quantifying the impact of early 21$^{st}$ century volcanic eruptions on global-mean surface temperature? To be submitted in Environmental Research Letters.

Morice CP, Kennedy JJ, Rayner NA, Jones PD (2012) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. Journal of Geophysical Research: Atmospheres (1984–2012) 117

Msadek R, Johns WE, Yeager SG, Danabasoglu G, Delworth TL, Rosati A (2013) The Atlantic meridional heat transport at 26.5 N and its relationship with the MOC in the RAPID array and the GFDL and NCAR coupled models. Journal of Climate 26:4335-4356

Msadek R et al. (2014a) Predicting a decadal shift in North Atlantic climate variability using the GFDL forecast system. Journal of Climate 27:6472-6496

Msadek R, Vecchi G, Winton M, Gudgel R (2014b) Importance of initial conditions in seasonal predictions of Arctic sea ice extent. Geophysical Research Letters 41:5208-5215

Müller W, Appenzeller C, Schär C (2005) Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature. Climate Dynamics 24:213-226

Noilhan J, Mahfouf J-F (1996) The ISBA land surface parameterisation scheme. Global and planetary Change 13:145-159

Noilhan J, Planton S (1989) A simple parameterization of land surface processes for meteorological models. Monthly Weather Review 117:536-549

Okumura Y, Xie S-P (2004) Interaction of the Atlantic Equatorial Cold Tongue and the African Monsoon*. Journal of Climate 17:3589-3602

Ortega P, Hawkins E, Sutton R (2011) Processes governing the predictability of the Atlantic meridional overturning circulation in a coupled GCM. Climate Dynamics 37:1771-1782

Ortega P, Montoya M, González-Rouco F, Mignot J, Legutke S (2012) Variability of the Atlantic meridional overturning circulation in the last millennium and two IPCC scenarios. Climate Dynamics 38:1925-1947

Paolino DA, Kinter III JL, Kirtman BP, Min D, Straus DM (2012) The impact of land surface and atmospheric initialization on seasonal forecasts with CCSM. Journal of Climate 25:1007-1021

Pohlmann H, Botzet M, Latif M, Roesch A, Wild M, Tschuck P (2004) Estimating the decadal predictability of a coupled AOGCM. Journal of Climate 17:4463-4472

Pohlmann H, Jungclaus JH, Köhl A, Stammer D, Marotzke J (2009) Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. Journal of Climate 22:3926-3938

Pohlmann H et al. (2013) Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. Climate Dynamics 41:775-785

Richter I, Xie S-P (2008) On the origin of equatorial Atlantic biases in coupled general circulation models. Climate Dynamics 31:587-598

Richter I, Behera SK, Doi T, Taguchi B, Masumoto Y, Xie S-P (2014a) What controls equatorial Atlantic winds in boreal spring? Climate Dynamics 43:3091-3104

Richter I, Xie S-P, Behera SK, Doi T, Masumoto Y (2014b) Equatorial Atlantic variability and its relation to mean state biases in CMIP5. Climate Dynamics 42:171-188

Ridley DA et al. (2014) Total volcanic stratospheric aerosol optical depths and implications for global climate change. Geophys. Res. Lett 41 :7763–7769. doi:10.1002/2014GL061541.

Robson J, Sutton R, Lohmann K, Smith D, Palmer MD (2012a) Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. Journal of Climate 25:4116-4134

Robson J, Sutton R, Smith D (2012b) Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. Geophysical Research Letters 39

Sakamoto T et al. (2012) MIROC4h—a new high-resolution atmosphere-ocean coupled general circulation model. Journal of the Meteorological Society of Japan 90:325-359

Sanchez-Gomez E, Cassou C, Ruprich-Robert Y, Fernandez E, Terray L (2015) Drift dynamics in a coupled model initialized for decadal forecasts. Climate Dynamics:1-22

1115   Santer BD et al. (2014) Volcanic contribution to decadal changes in tropospheric temperature. Nature
1116         Geoscience 7:185-189
1117   Santer BD et al. (2015) Observed multivariable signals of late 20th and early 21st century volcanic
1118         activity. Geophys. Res. Lett. 42 :500–509. doi:10.1002/2014GL062366.
1119   Sarafanov A, Falina A, Sokov A, Demidov A (2008) Intense warming and salinification of
1120         intermediate waters of southern origin in the eastern subpolar North Atlantic in the 1990s to
1121         mid‐2000s. Journal of Geophysical Research: Oceans (1978–2012) 113
1122   Schmidt GA, Shindell DT, Tsigaridis K (2014) Reconciling warming trends. Nat. Geosci. 7 :158-160.
1123         doi:10.1038/ngeo2105
1124   Schweiger A, Lindsay R, Zhang J, Steele M, Stern H, Kwok R (2011) Uncertainty in modeled Arctic
1125         sea ice volume. Journal of Geophysical Research: Oceans (1978–2012) 116
1126   Screen JA, Simmonds I (2010) The central role of diminishing sea ice in recent Arctic temperature
1127         amplification. Nature 464:1334-1337
1128   Serreze MC, Holland MM, Stroeve J (2007) Perspectives on the Arctic's shrinking sea-ice cover.
1129         Science 315:1533-1536
1130   Shaffrey LC et al. (2009) UK HiGEM: The new UK high-resolution global environment model-model
1131         description and basic evaluation. Journal of Climate 22:1861-1896
1132   Sigmond M, Fyfe J, Flato G, Kharin V, Merryfield W (2013) Seasonal forecast skill of Arctic sea ice
1133         area in a dynamical forecast system. Geophysical Research Letters 40:529-534
1134   Simmons A, Uppala S, Dee D, Kobayashi S (2007) ERA-Interim: New ECMWF reanalysis products
1135         from 1989 onwards. ECMWF newsletter 110:25-35
1136   Small RJ et al. (2014) A new synoptic scale resolving global climate simulation using the Community
1137         Earth System Model. Journal of Advances in Modeling Earth Systems 6:1065-1094
1138   Smith DM, Cusack S, Colman AW, Folland CK, Harris GR, Murphy JM (2007) Improved surface
1139         temperature prediction for the coming decade from a global climate model. Science 317:796-
1140         799
1141   Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements to NOAA's historical
1142         merged land-ocean surface temperature analysis (1880-2006). Journal of Climate 21:2283-
1143         2296
1144   Solomon S, Rosenlof KH, Portmann RW, Daniel JS, Davis SM, Sanford TJ, Plattner G-K (2010)
1145         Contributions of stratospheric water vapor to decadal changes in the rate of global warming.
1146         Science 327:1219-1223
1147   Stockdale TN, Balmaseda MA, Vidard A (2006) Tropical Atlantic SST prediction with coupled ocean-
1148         atmosphere GCMs. Journal of Climate 19:6047-6061
1149   Stroeve JC, Serreze MC, Holland MM, Kay JE, Malanik J, Barrett AP (2012) The Arctic's rapidly
1150         shrinking sea ice cover: a research synthesis. Climatic Change 110:1005-1027
1151   Szopa S et al. (2013) Aerosol and ozone changes as forcing for climate evolution between 1850 and
1152         2100. Climate Dynamics 40:2223-2250
1153   Timmreck C, Pohlmann H, Illing S, Kadow C (2016) The impact of stratospheric volcanic aerosol on
1154         decadal-scale climate predictions. Geophysical Research Letters 43:834-842.
1155         doi:10.1002/2015gl067431
1156   Trenberth KE, Fasullo JT (2013) An apparent hiatus in global warming? Earth's Future 1:19-32
1157   Uppala S, Dee D, Kobayashi S, Berrisford P, Simmons A (2008) Towards a climate data assimilation
1158         system: status update of ERA-Interim. ECMWF newsletter 115:12-18
1159   Valcke S, Craig T, Coquart L (2013) OASIS3-MCT User Guide OASIS3-MCT 2.0. CERFACS/CNRS
1160         SUC URA
1161   van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Decadal prediction skill in a
1162         multi-model ensemble. Climate Dynamics 38:1263-1280
1163   Vancoppenolle M, Fichefet T, Goosse H (2009a) Simulating the mass balance and salinity of Arctic
1164         and Antarctic sea ice. 2. Importance of sea ice salinity variations. Ocean Modelling 27:54-69
1165   Vancoppenolle M, Fichefet T, Goosse H, Bouillon S, Madec G, Maqueda MAM (2009b) Simulating
1166         the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and
1167         validation. Ocean Modelling 27:33-53
1168   Vernier JP et al. (2011) Major influence of tropical volcanic eruptions on the stratospheric aerosol
1169         layer during the last decade. Geophysical Research Letters 38

1170  Visbeck M, Chassignet EP, Curry RG, Delworth TL, Dickson RR, Krahmann G (2003) The ocean's
1171      response to North Atlantic Oscillation variability. Geophysical Monograph-American
1172      Geophysical Union 134:113-146
1173  Voldoire A et al. (2013) The CNRM-CM5. 1 global climate model: description and basic evaluation.
1174      Climate Dynamics 40:2091-2121
1175  Vose RS et al. (2012) NOAA's merged land-ocean surface temperature analysis. Bulletin of the
1176      American Meteorological Society 93:1677-1685
1177  Wang M, Overland JE (2009) A sea ice free summer Arctic within 30 years? Geophysical Research
1178      Letters 36
1179  Watanabe M et al. (2013) Strengthening of ocean heat uptake efficiency associated with the recent
1180      climate hiatus. Geophysical Research Letters 40:3175-3179
1181  Yashayaev I (2007) Hydrographic changes in the Labrador Sea, 1960–2005. Progress in
1182      Oceanography 73:242-276
1183  Yeager S, Karspeck A, Danabasoglu G, Tribbia J, Teng H (2012) A decadal prediction case study:
1184      Late twentieth-century North Atlantic Ocean heat content. Journal of Climate 25:5173-5189
1185  Zermeño-Diaz DM, Zhang C (2013) Possible root causes of surface westerly biases over the equatorial
1186      Atlantic in global climate models. Journal of Climate 26:8154-8168
1187  Zhang J, Steele M, Lindsay R, Schweiger A, Morison J (2008) Ensemble 1‑Year predictions of
1188      Arctic sea ice for the spring and summer of 2008. Geophysical Research Letters 35

1189

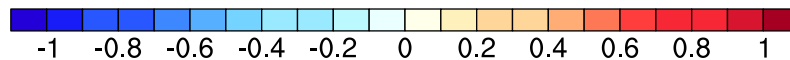Figure1
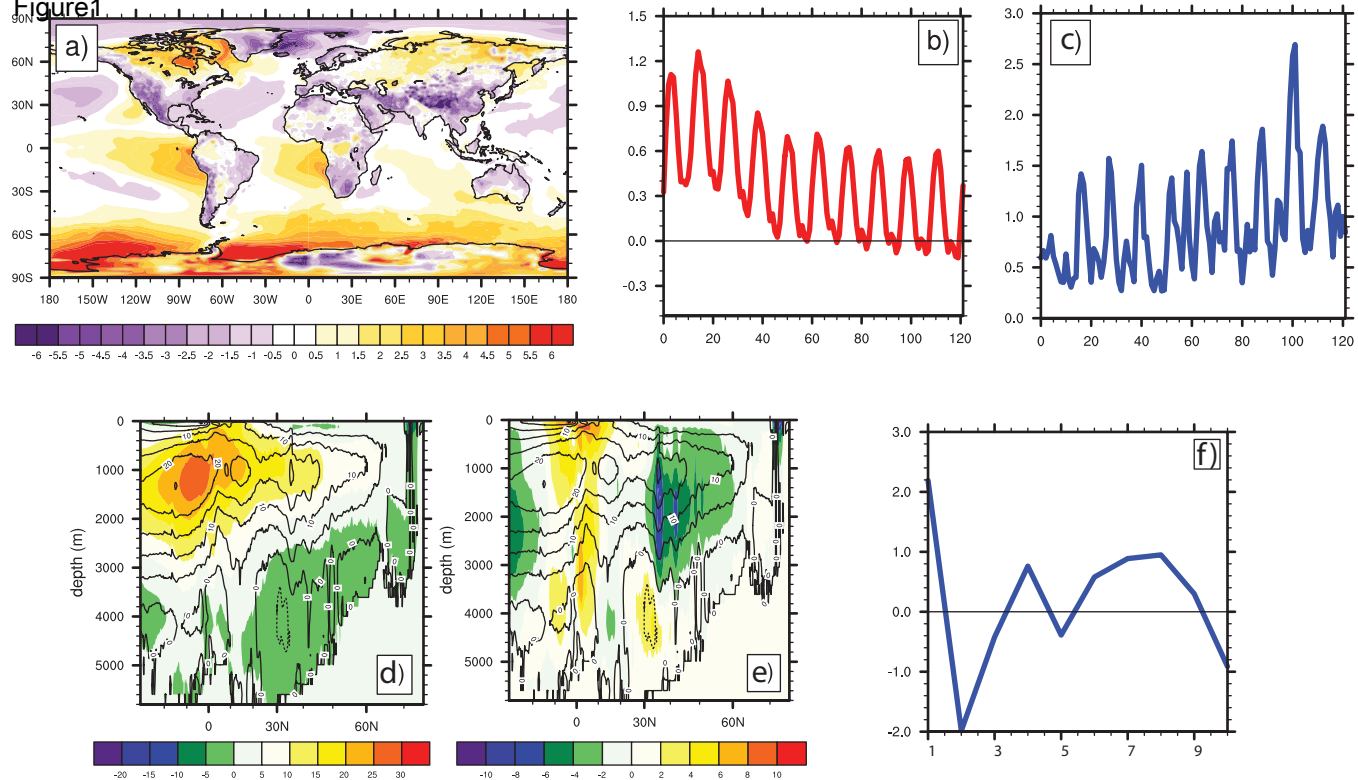
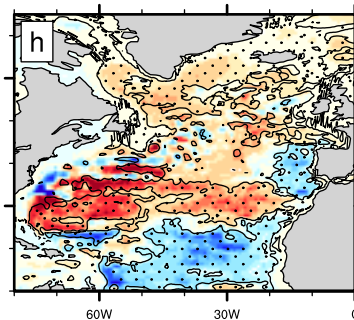Figure2
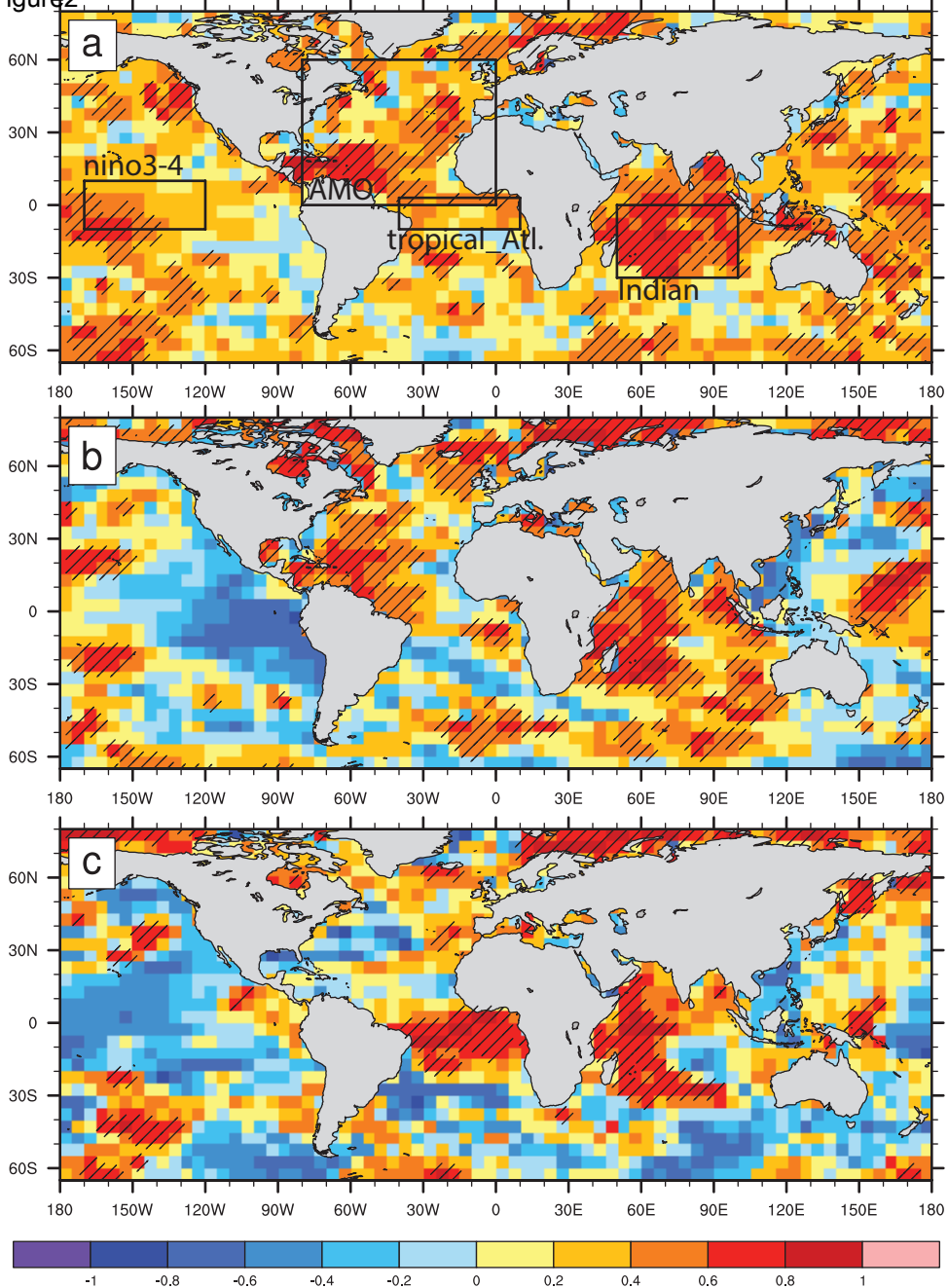
Figure3
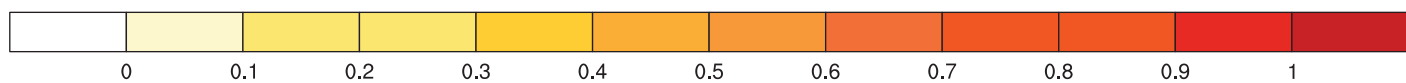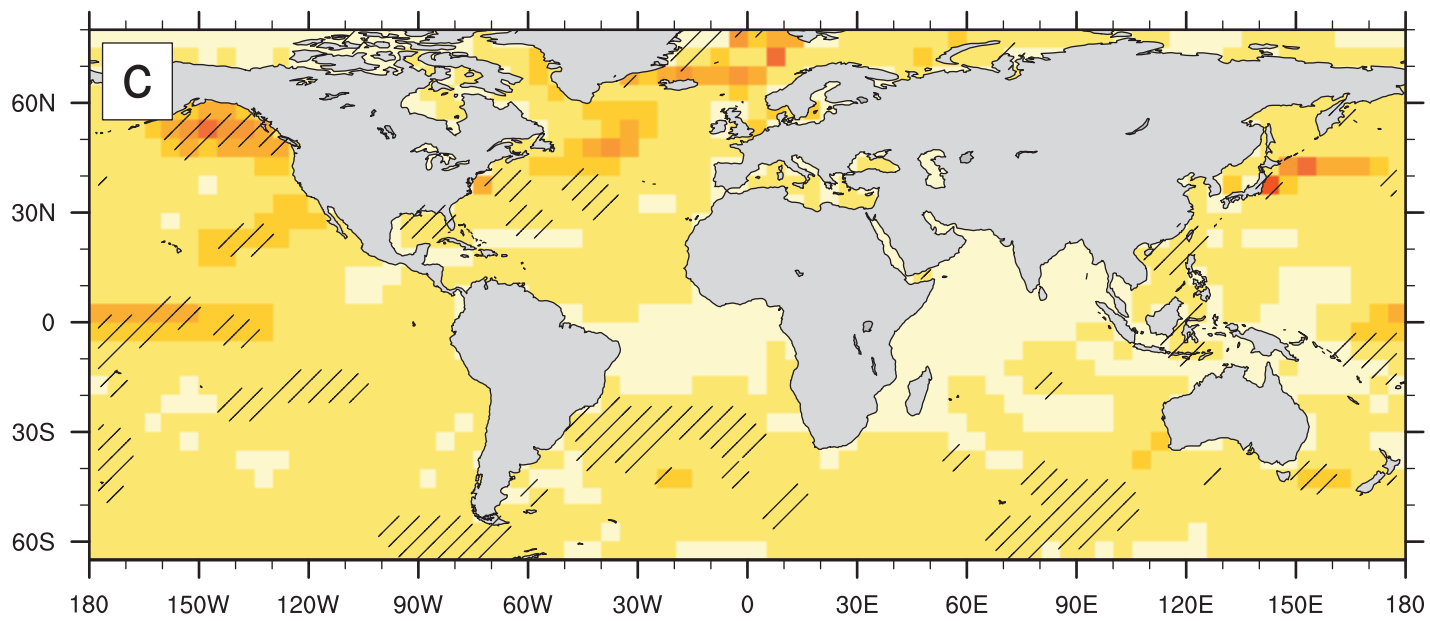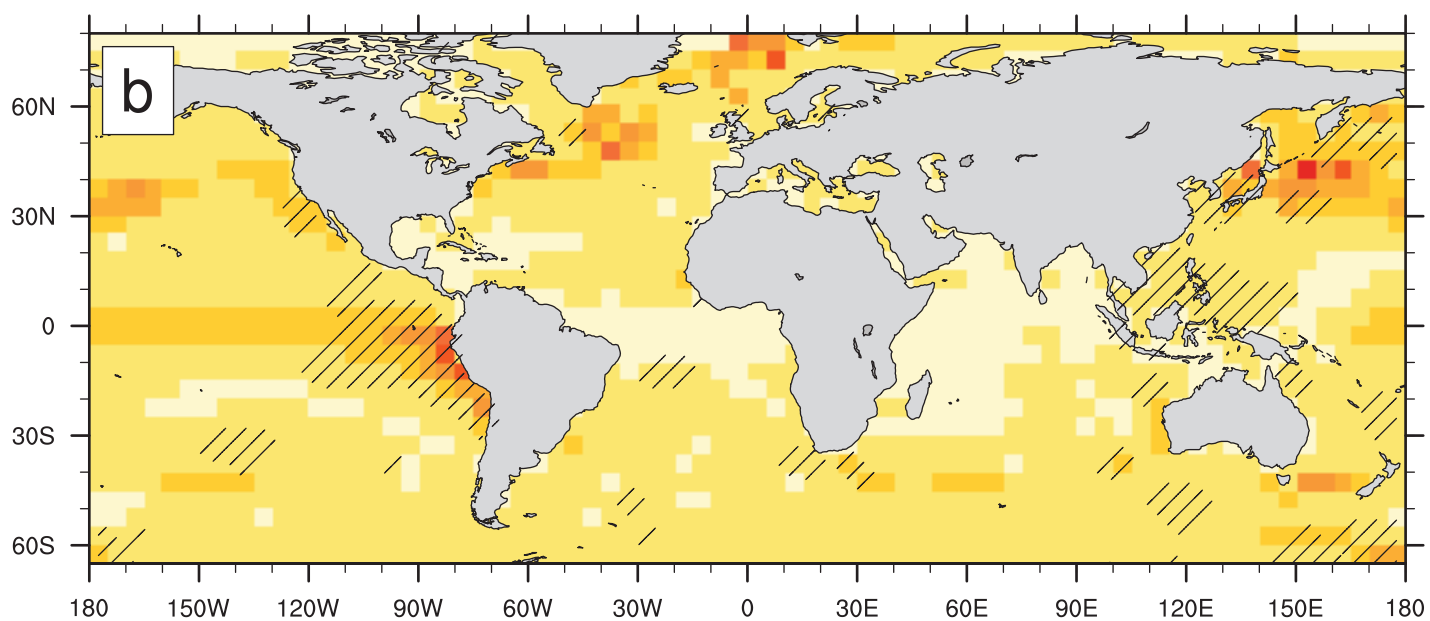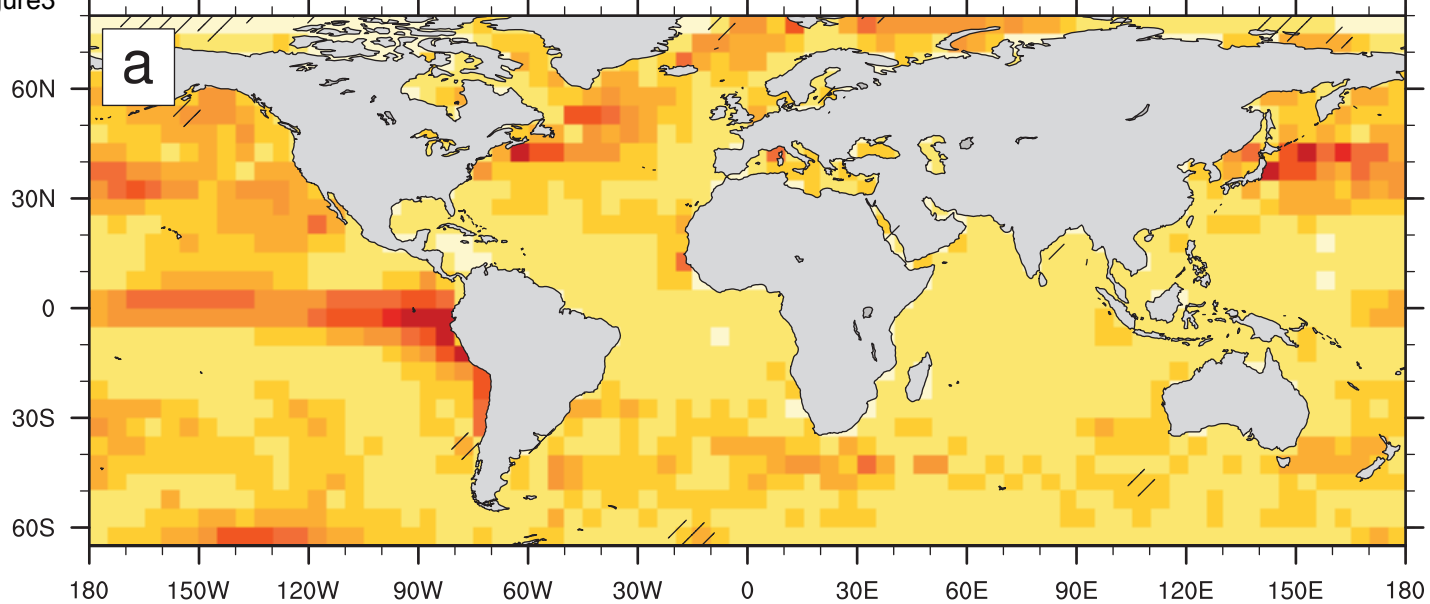
Figure4
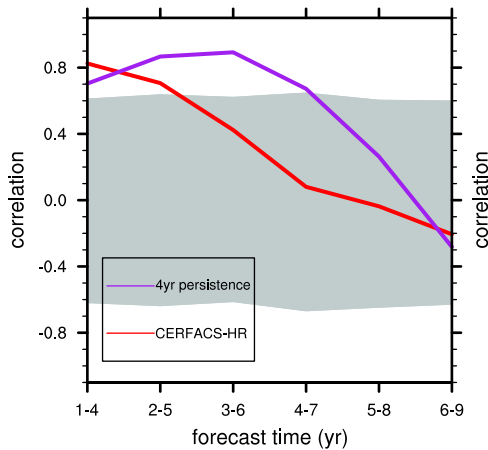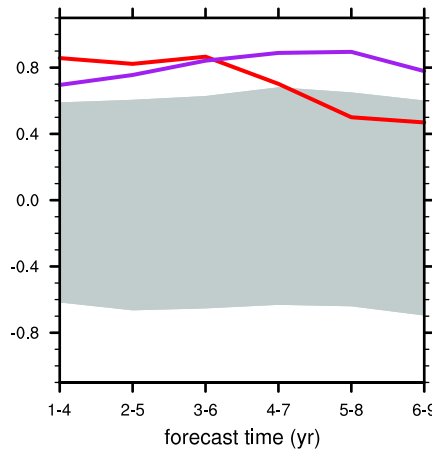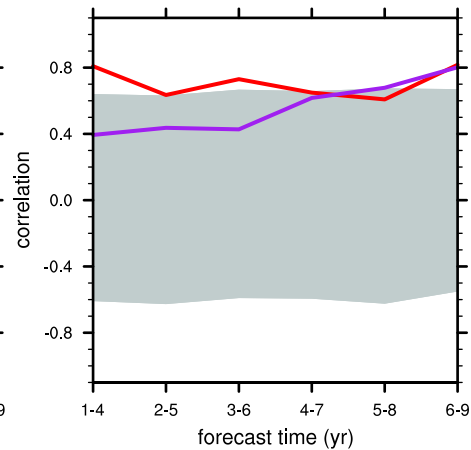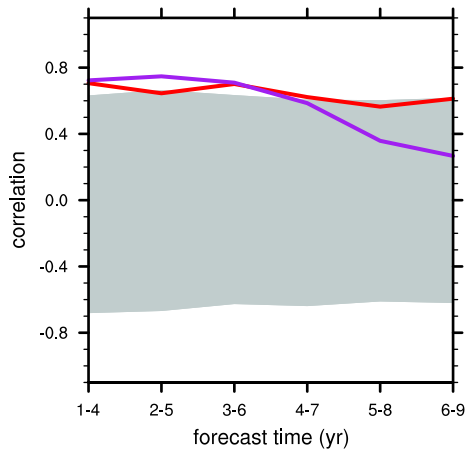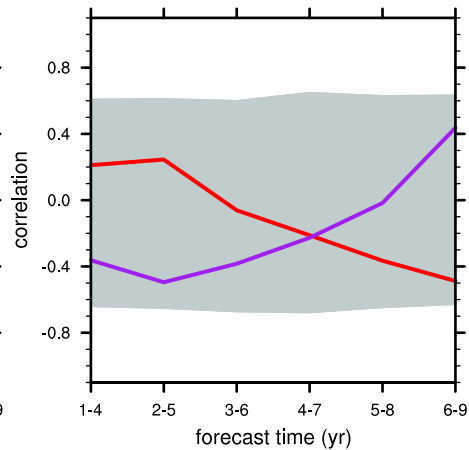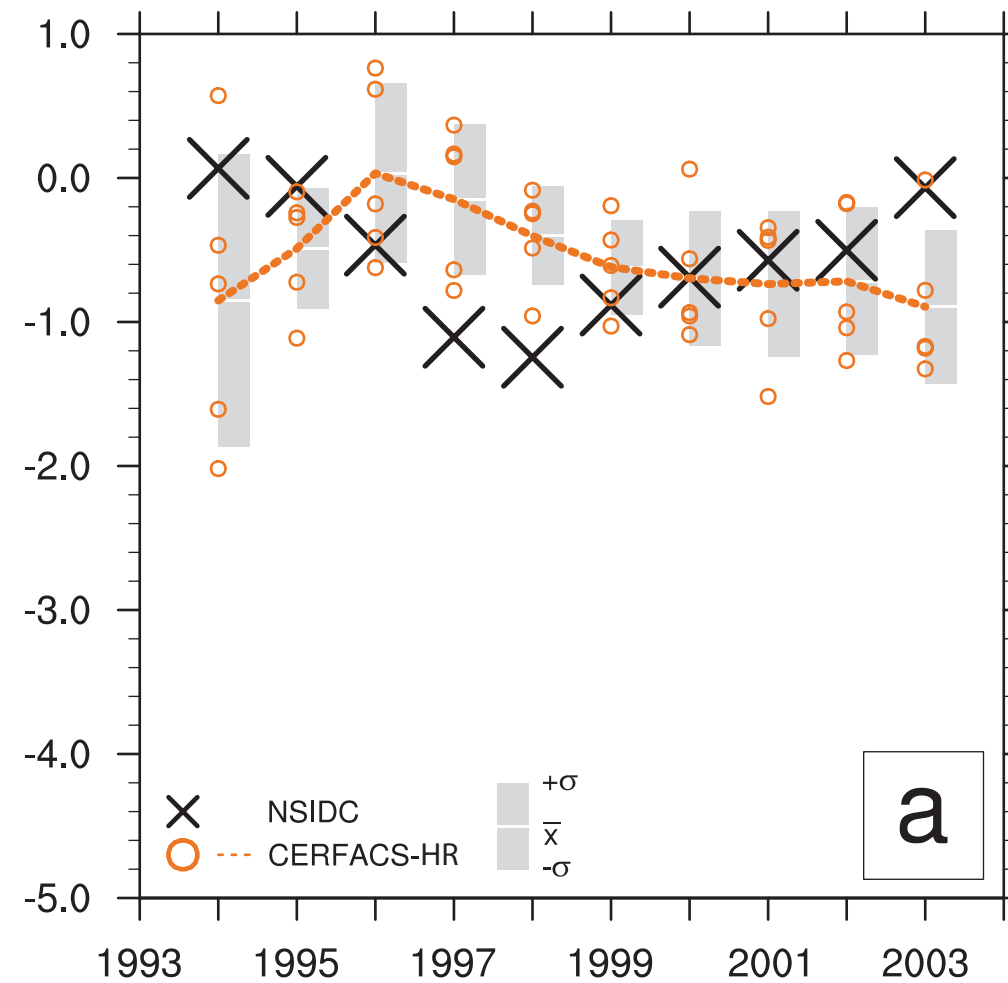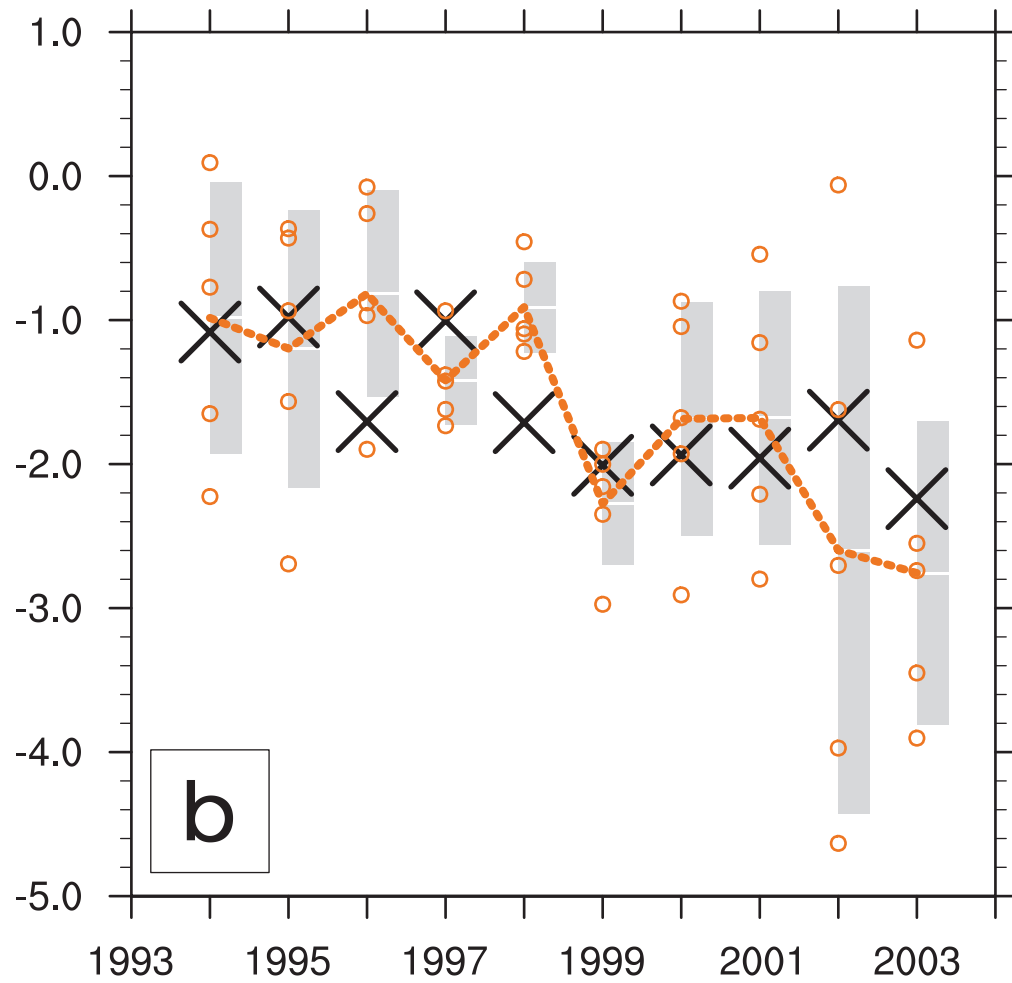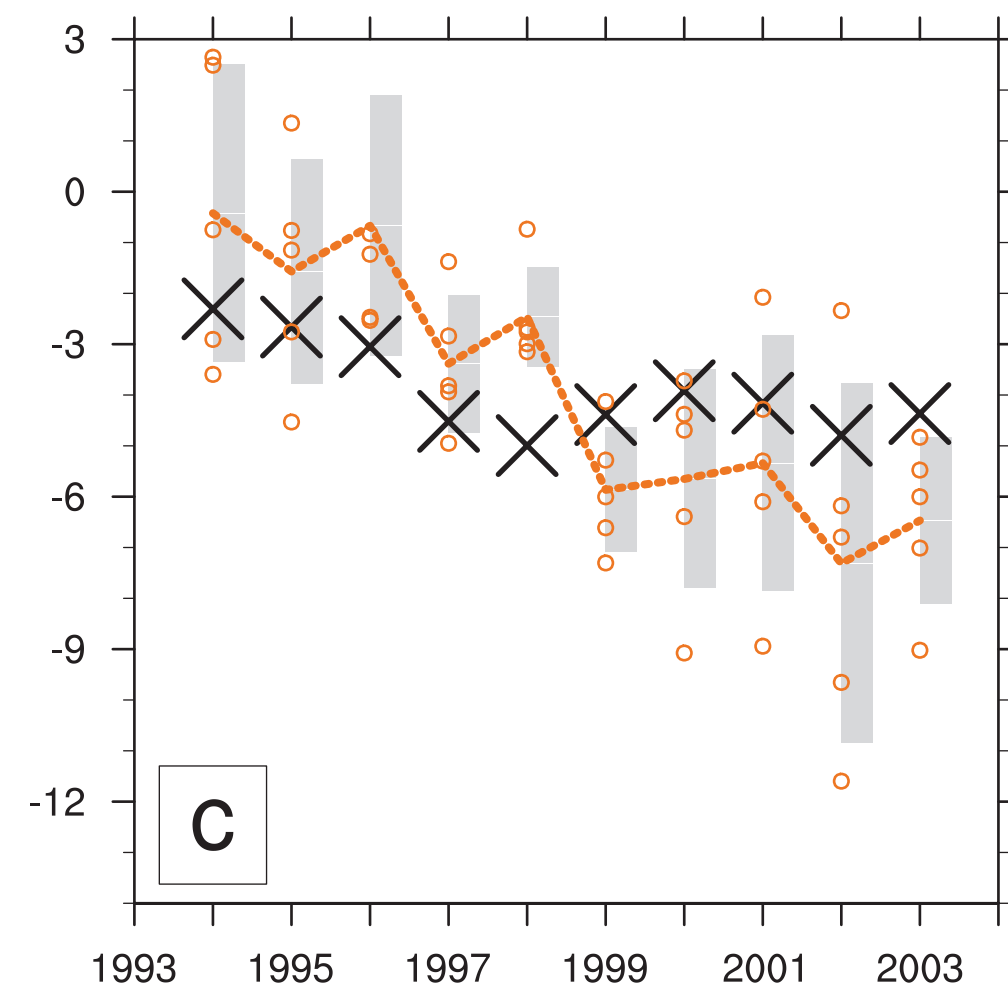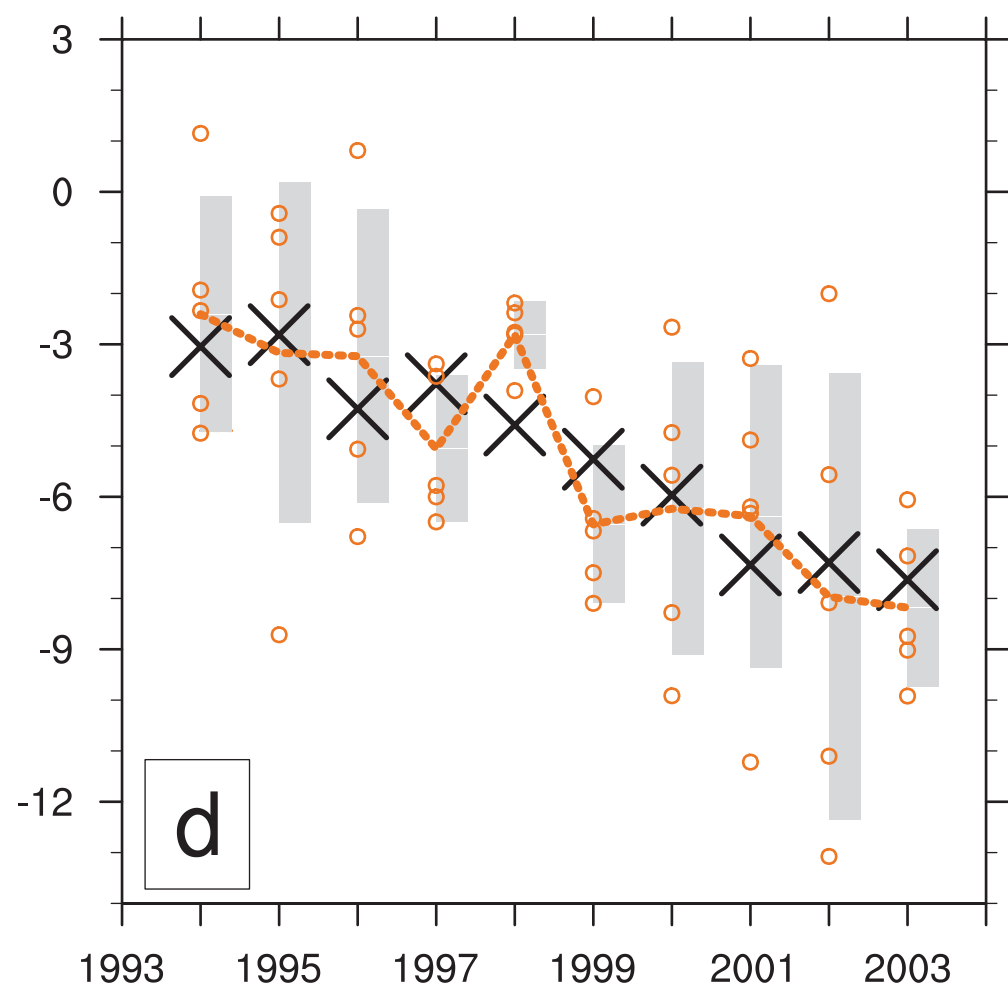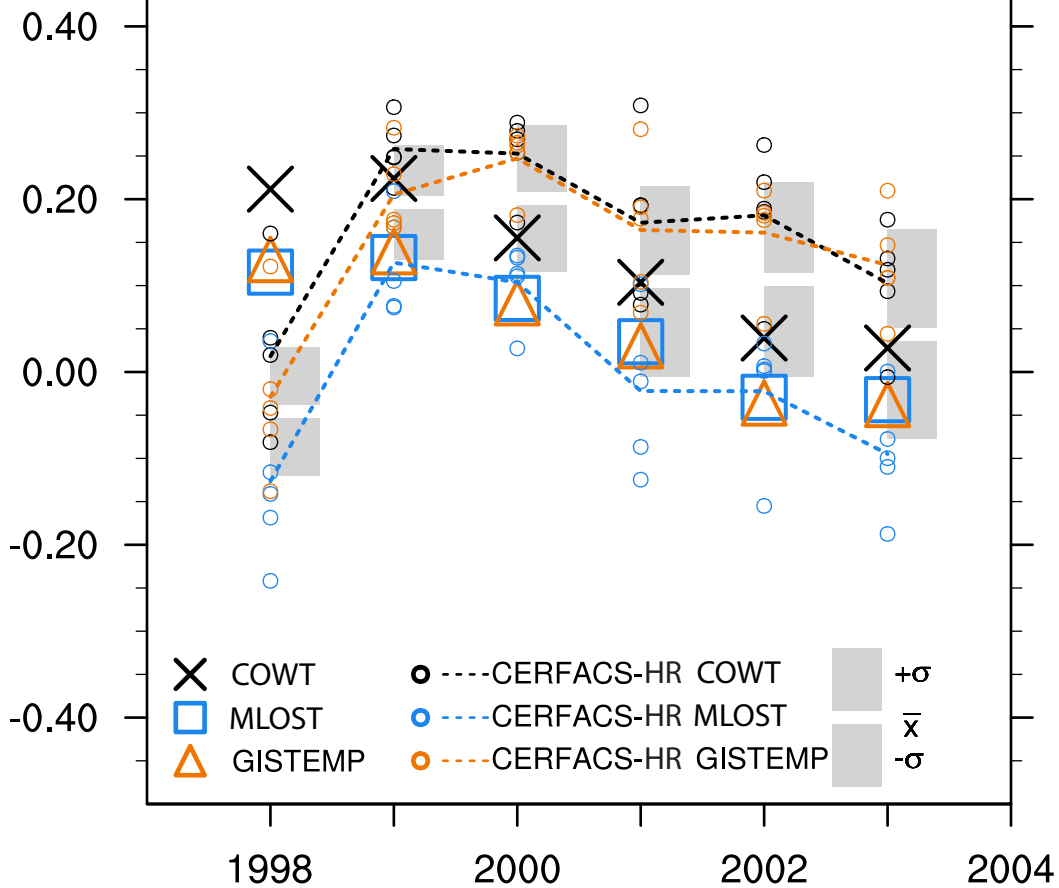
Figure5
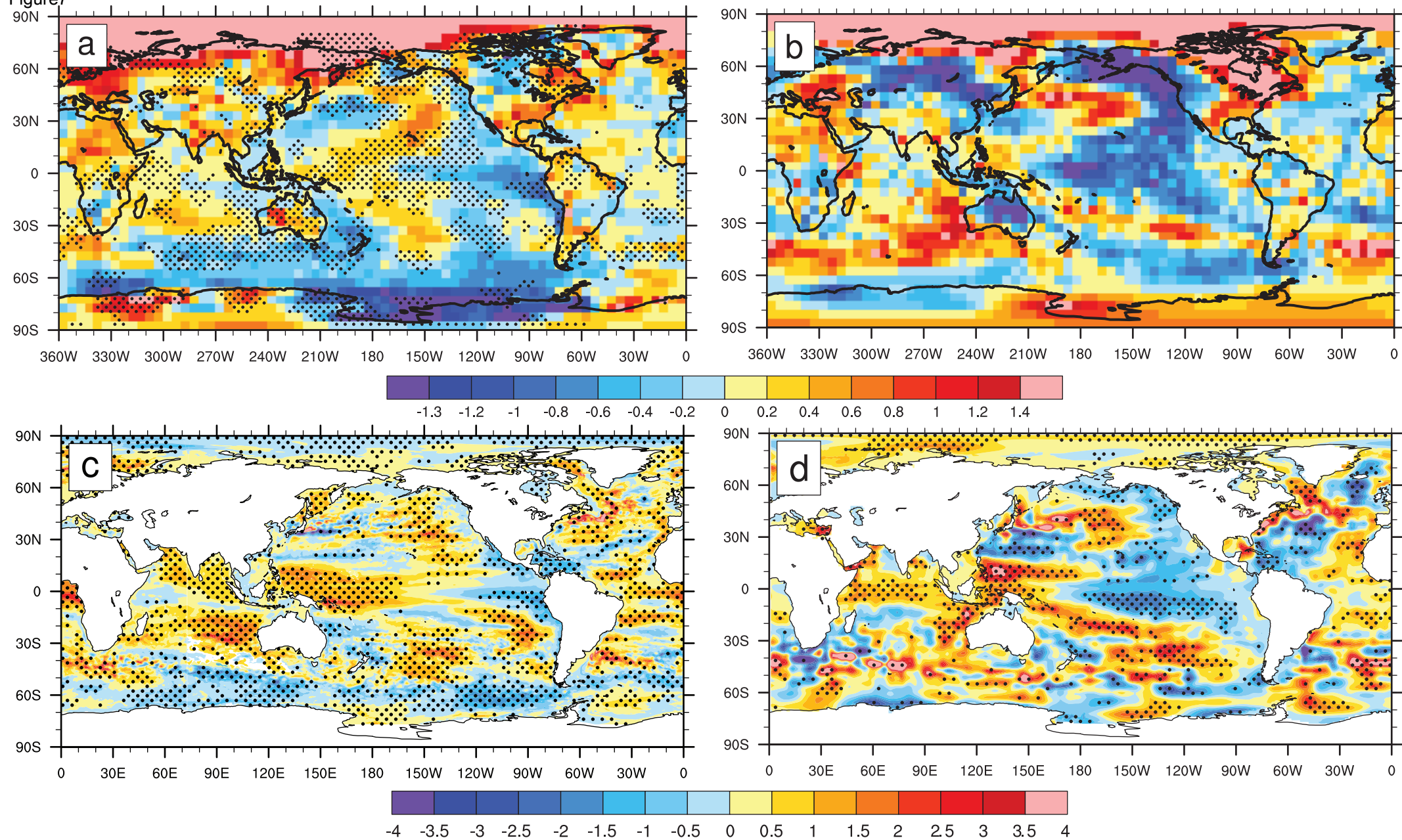
Figure6

Figure7

Figure8
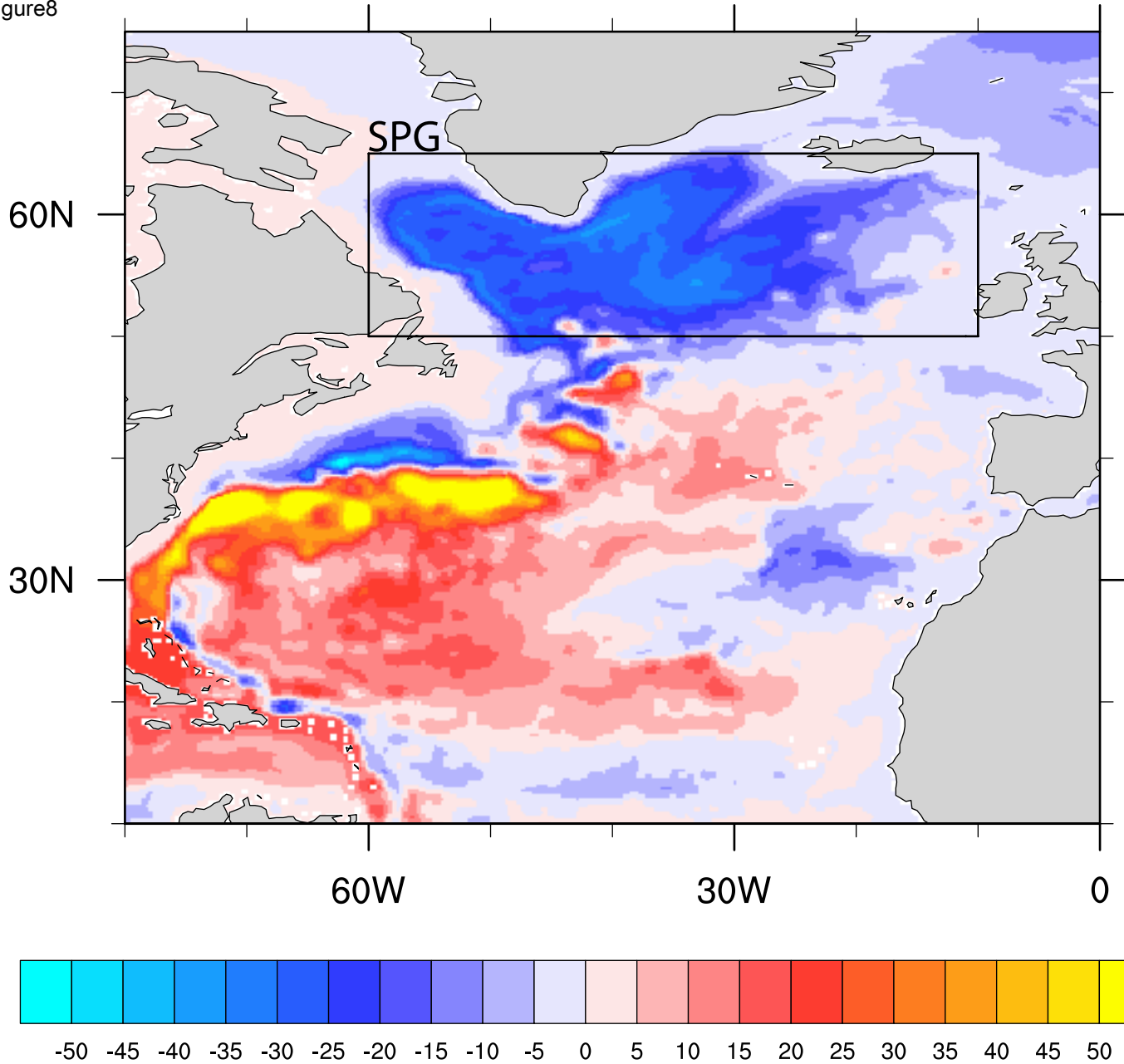
Figure9
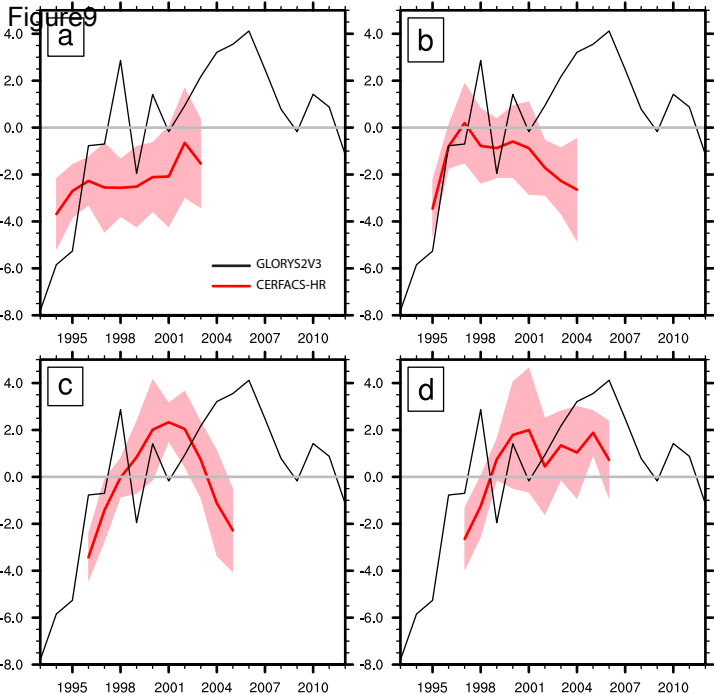
Figure10
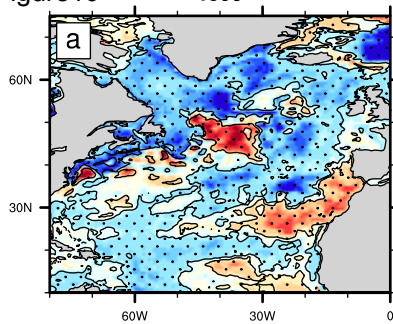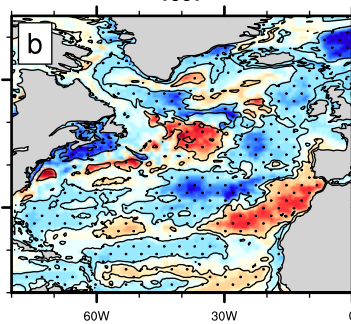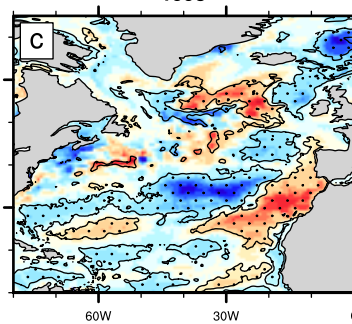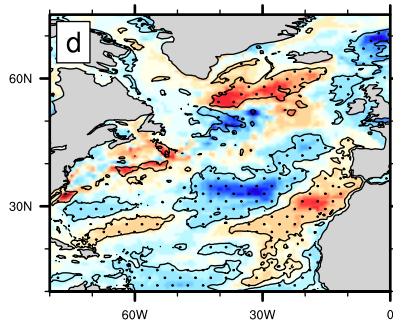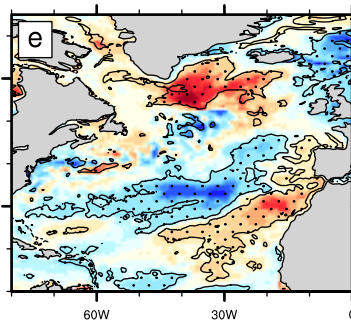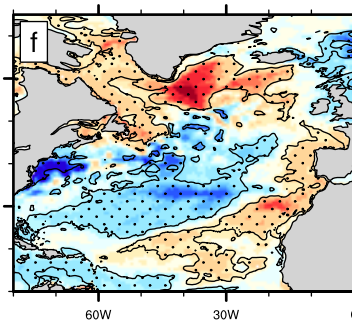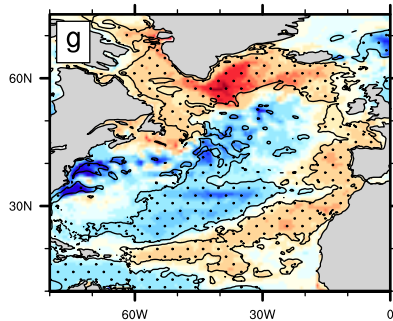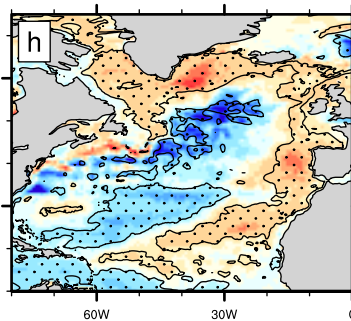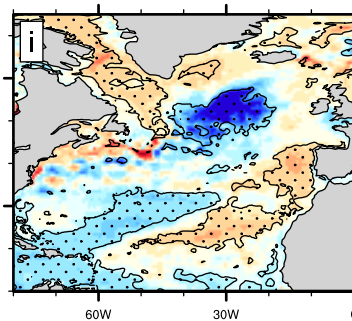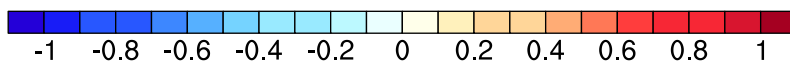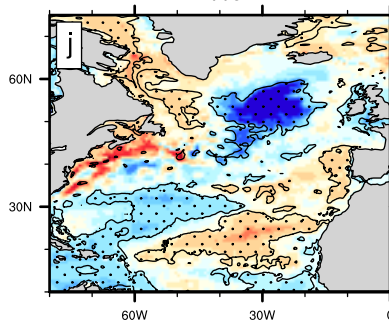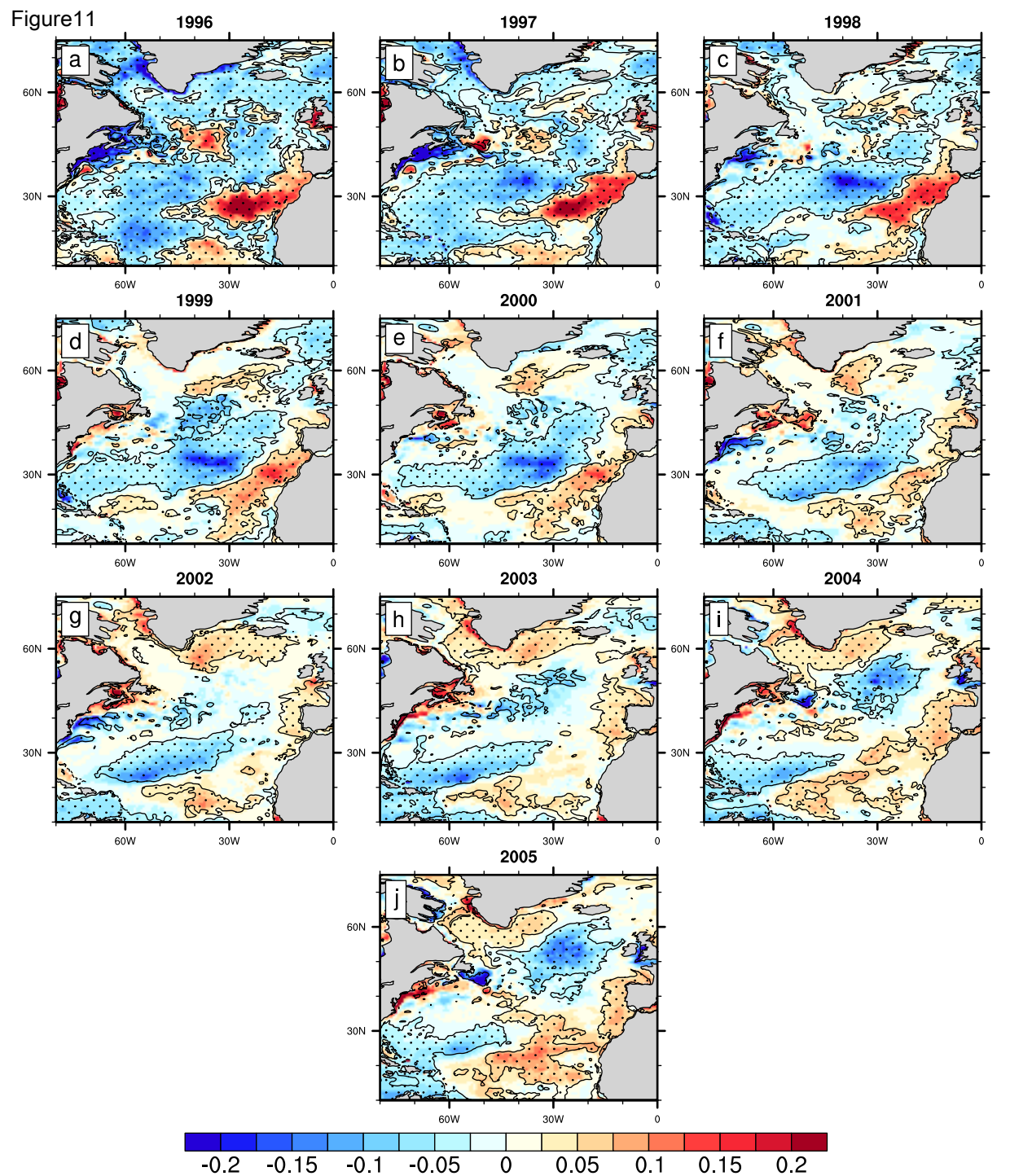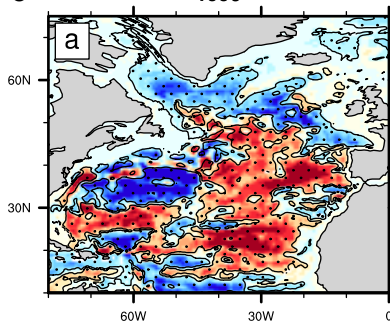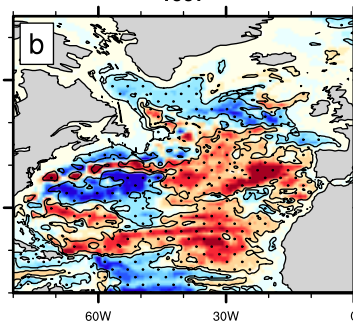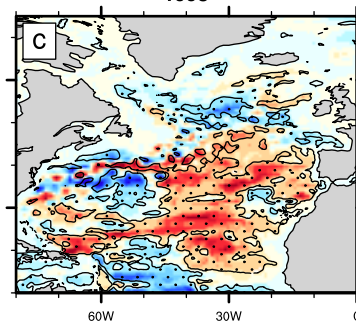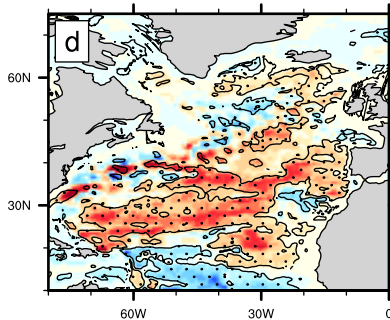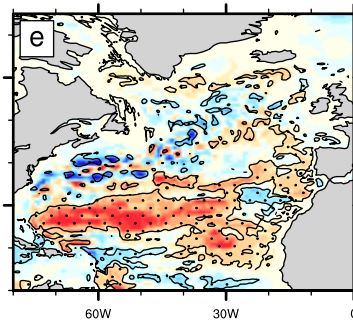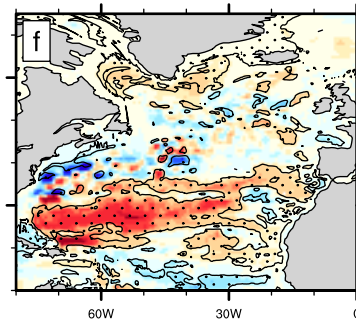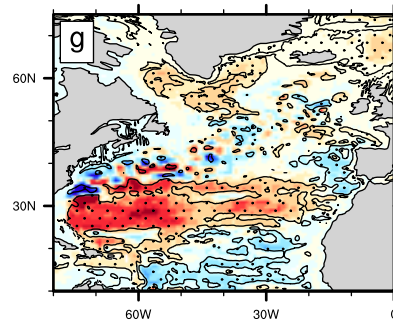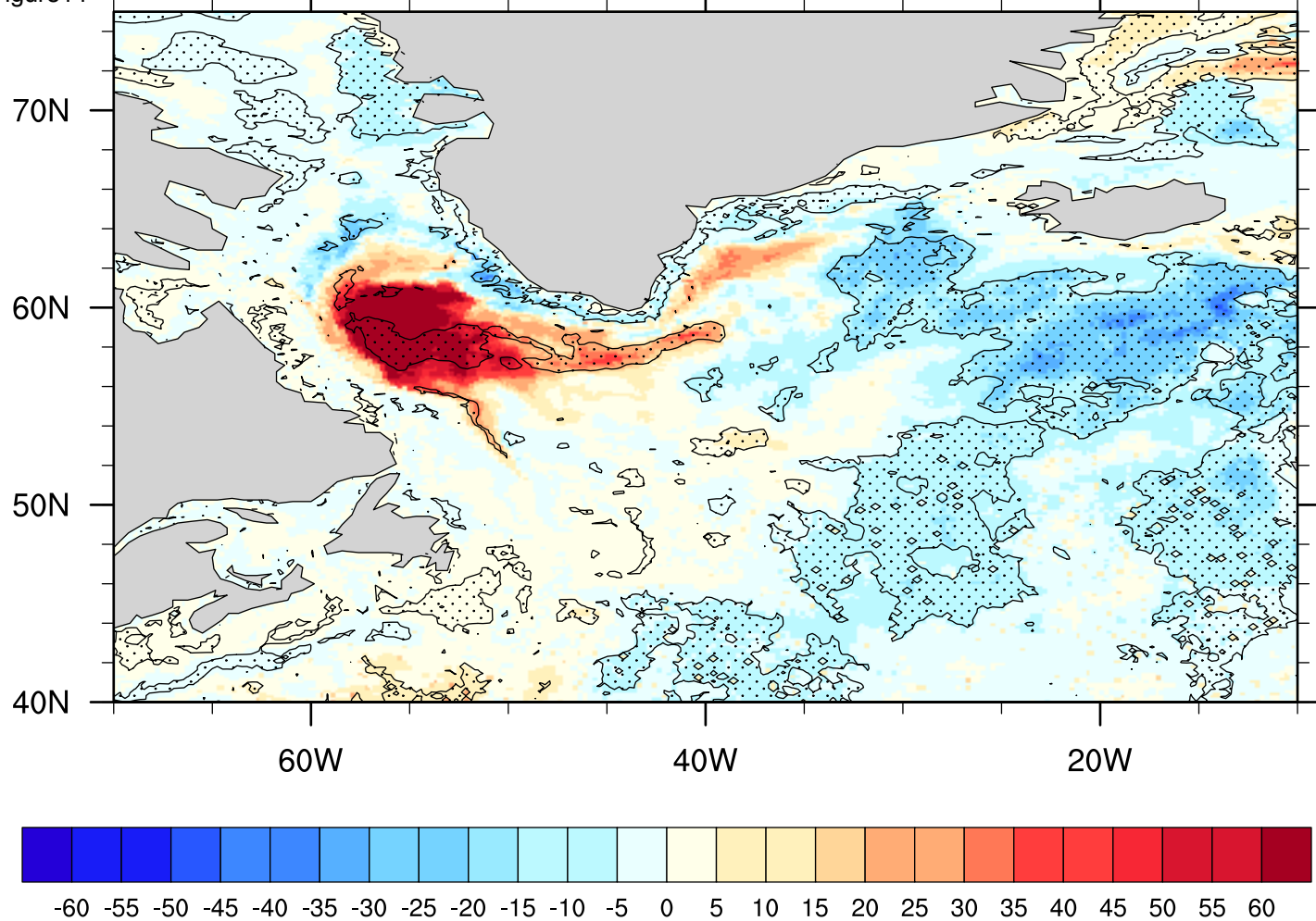
Figure11

Figure12

Figure13
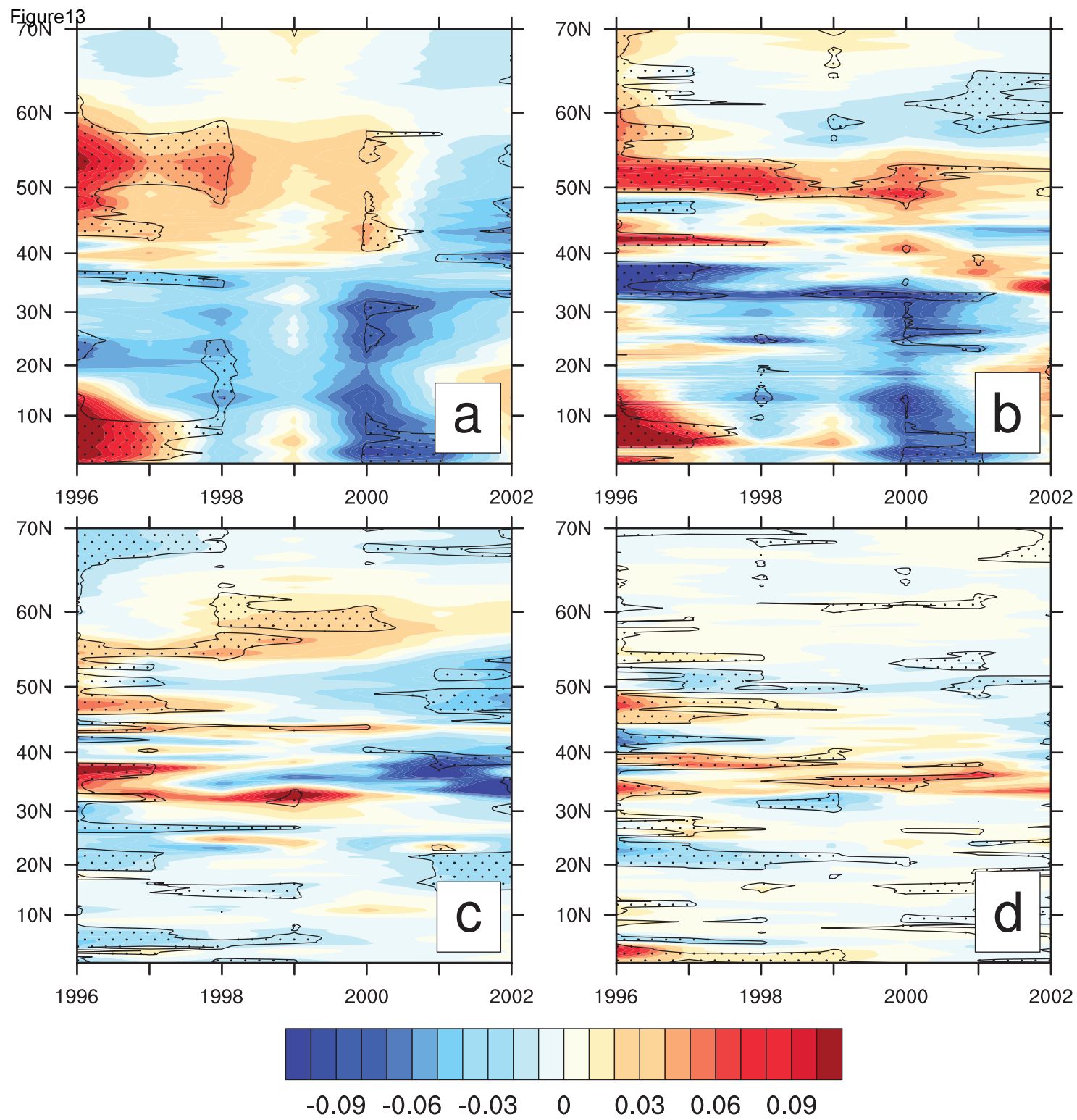
Figure14

Figure15
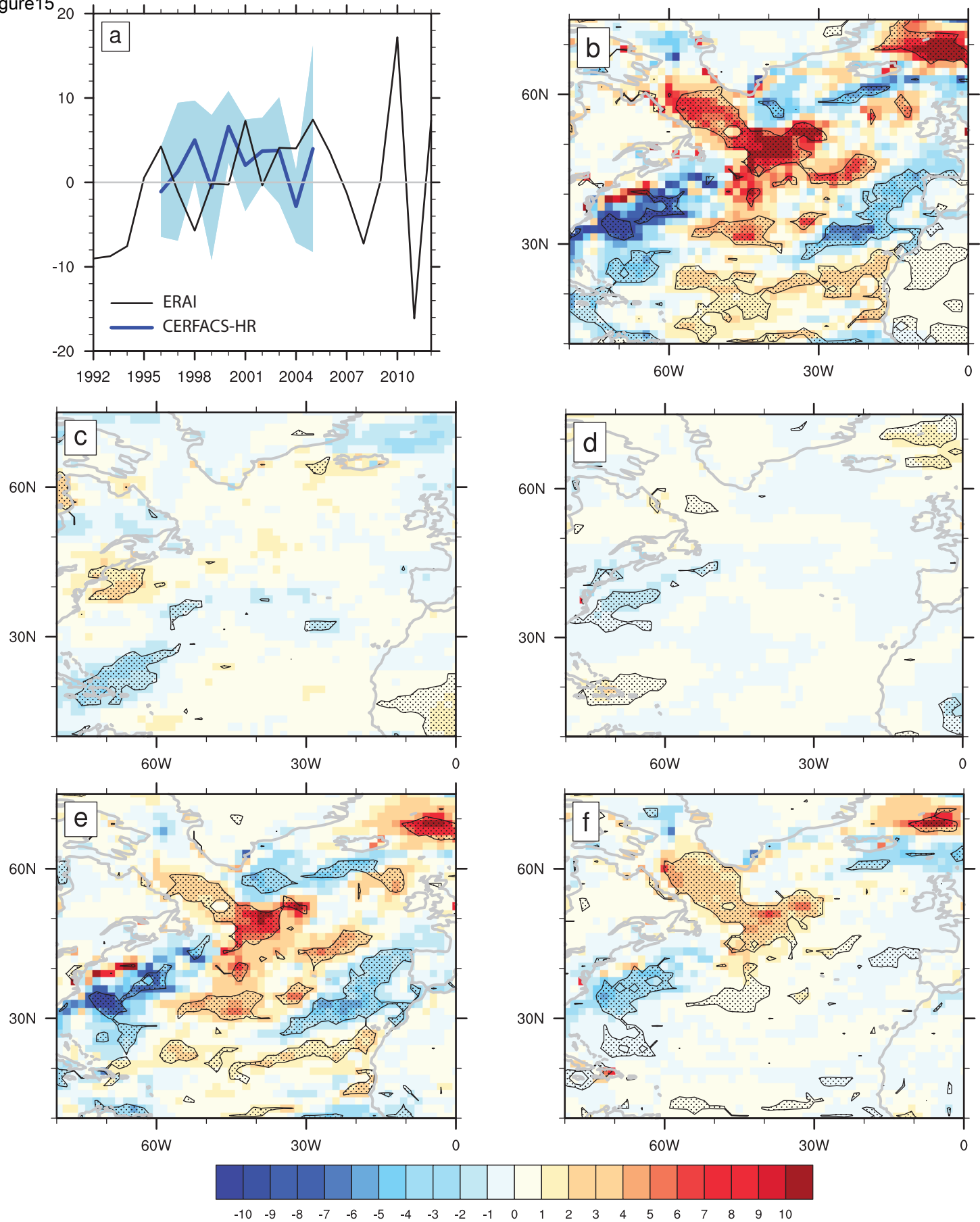
Figure16
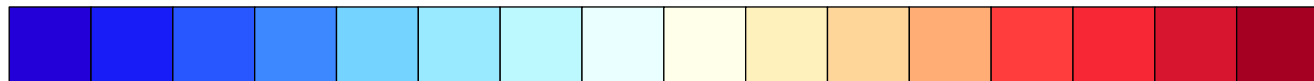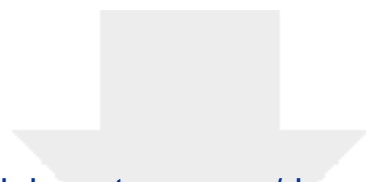
Click here to access/download
**Electronic Supplementary Material**
Auxiliary material for Paper.odt