Clim Dyn (2011) 36:19–39
DOI 10.1007/s00382-010-0781-7

A statistical–dynamical scheme for reconstructing ocean forcing in the Atlantic. Part I: weather regimes as predictors for ocean surface variables

Christophe Cassou · Marie Minvielle · Laurent Terray · Claire Périgaud

Received: 26 May 2009 / Accepted: 24 February 2010 / Published online: 7 April 2010

Abstract The links between the observed variability of the surface ocean variables estimated from reanalysis and the overlying atmosphere decomposed in classes of large-scale atmospheric circulation via clustering are investigated over the Atlantic from 1958 to 2002. Daily 500 hPa geopotential height and 1,000 hPa wind anomaly maps are classified following a weather-typing approach to describe the North Atlantic and tropical Atlantic atmospheric dynamics, respectively. The algorithm yields patterns that correspond in the extratropics to the well-known North Atlantic-Europe weather regimes (NAE-WR) accounting for the barotropic dynamics, and in the tropics to wind classes (T-WC) representing the alteration of the trades. 10-m wind and 2-m temperature (T2) anomaly composites derived from regime/wind class occurrence are indicative of strong relationships between daily large-scale atmospheric circulation and ocean surface over the entire Atlantic basin. High temporal correlation values are obtained basin-wide at low frequency between the observed fields and their reconstruction by multiple linear regressions with the frequencies of occurrence of both NAE-WR and T-WC used as sole predictors. Additional multiple linear regressions also emphasize the importance of accounting for the strength of the daily anomalous atmospheric circulation estimated by the combined distances to all regimes centroids in order to reproduce the daily to interannual variability of the Atlantic ocean. We show that for most of the North Atlantic basin the occurrence of NAE-WR generally sets the sign of the ocean surface anomaly for a given day, and that the inter-regime distances are valuable predictors for the magnitude of that anomaly. Finally, we provide evidence that a large fraction of the low-frequency trends in the Atlantic observed at the surface over the last 50 years can be traced back, except for T2, to changes in occurrence of tropical and extratropical weather classes. All together, our findings are encouraging for the prospects of basin-scale ocean dynamical downscaling using a weather-typing approach to reconstruct forcing fields for high resolution ocean models (Part II) from coarse resolution climate models.

Keywords Weather regimes · Climate variability · Atlantic Ocean · Oceanic forcing variables

C. Cassou (✉) · M. Minvielle · L. Terray
CERFACS/CNRS, Climate Modelling and Global Change Team, 42 avenue Gaspard Coriolis, 31057 Toulouse, France
e-mail: cassou@cerfacs.fr

C. Périgaud
JPL-NASA, Ocean Science Element, 4800 Oak Grove Drive, Pasadena, CA 91001, USA

1 Introduction

Ocean variability occurs over a broad range of spatio-temporal scales, from seasonal (associated with surface mixed layer processes) to decadal (e.g. gyres circulation) to centennial and longer (e.g. the meridional overturning circulation, MOC). Internal dynamics through mesoscale transient eddies and their interaction with the mean flow contributes to the evolution of the three-dimensional density distribution of the ocean. Sources for most of the low frequency variability can be found though at the ocean–atmosphere interface through changes in wind patterns and buoyancy (heat and freshwater) fluxes acting as a forcing. The wind-driven ocean circulation dominates the strong current system in the upper ocean, such as the subtropical and subpolar gyres, and interacts nonlinearly with the buoyancy-driven flow. This is especially true in the

Atlantic where deep water formation associated with deep convection and strong vertical mixing occurs in the Nordic Seas and in the Labrador Sea and feeds the lower limb of the MOC.

Observational estimates for heat content show an overall trend from 1955 to 2005 (Levitus et al. 2005). The Atlantic Ocean accounts for approximately half of the global linear trend over that period. Strong warming extending down to 1,000 m occurs along the Gulf Stream and North Atlantic current, as part of a broader warming of the entire subtropical gyre. Concurrently, the subpolar gyre intensifies, cools down and gets fresher. The North Atlantic Oscillation (hereafter NAO) and associated perturbation of surface wind patterns and buoyancy fluxes (Hurrell et al. 2003) is the main driver of the oceanic variations in the Atlantic gyres. Sea surface temperature (SST) and salinity (SSS) anomalies, mixed layer depth, are highly correlated with the NAO as well as the circulation (strength and position of the main currents) with some lags (see Bindoff et al. 2007 for a review). Evidence has been also provided that the NAO affects the production of Labrador Sea Water (LSW) that is a major contributor to the lower limb of the MOC (Yashayaev and Loder 2009). This example suggests that understanding the 3-dimensional variability of the Atlantic Ocean relies on the ability to correctly account for the variability of the surface forcings associated with the large-scale atmospheric modes. The ocean modelling community even goes further and states that the surface fluxes play a large part in determining the fidelity of the oceanic simulation.

Despite considerable improvements, the 4th Intergovernmental Panel on Climate Change (IPCC) report shows that Coupled General Circulation Models (CGCMs) traditionally used to study the climate evolution for the last century and to provide climate projections for the next, still suffer from strong biases (e.g. Randall et al. 2007). Some are directly attributed to surface fluxes and interface processes. Displaced wind patterns or erroneous mean strength, misrepresentation of cloud radiative forcing and consequently surface radiative exchanges, wrong precipitation–evaporation fluxes (Dai 2006), etc. are associated with altered ocean circulation and buoyancy budget leading to major errors of the climate mean state. In addition, multi-model comparisons show (Frankignoul et al. 2004) that coupled models in general are unable to simulate the correct feedbacks between SST and both surface turbulent and radiative fluxes at any time scale. Some biases are also attributed to the representation of oceanic processes themselves. For instance, mesoscale dynamics is not resolved in current IPCC experiments because of coarse resolution (about 2° on average for the ocean). Recent publications (e.g. Cunningham et al. 2007 from observational derivations of meridional heat transport changes, Roberts et al. 2004 and Biastoch et al. 2008 among others from ocean simulations) strongly suggest though that mesoscale physics is crucial to capture the interannual-to-decadal variability of the gyre circulation and the MOC, and is likely to control part of the climate sensitivity to anthropogenic forcings. All these biases taking place in a poor observational context somehow lower the strength of the IPCC conclusions about ocean change, and are thought to partly explain the large uncertainties in both climate projections and the modeling of the twentieth century variability of the 3D ocean (e.g. conflicting results for MOC evolution over the last decades, Knight et al. 2005, Bryden et al. 2005, Baehr et al. 2007).

An important challenge for the next IPCC exercise is to reduce the large discrepancies in both present climate simulations and future projections of the circulation and 3D thermohaline structure of the ocean. The present study participates to this difficult task by adapting to ocean purposes the downscaling and bias-correction approaches that are traditionally used in hydrology, wind-power estimation, etc. (e.g. Boé et al. 2006; Najac et al. 2008). Several attempts applied to relatively small oceanic basins can be found in literature (e.g. the Baltic Sea, Heyen et al. 1996, Heyen and Dippner 1998). Here we take up the challenge to set a scheme to reconstruct at best the surface ocean forcing variables for the entire Atlantic Ocean aiming at a better representation and understanding of the ocean variability for both present and future climates. The proposed method is divided into two steps:

- a statistical step to produce an unbiased forcing dataset for the ocean from CGCMs outputs. It is commonly admitted by the climate community that CGCMs have some skill in reproducing large scale atmospheric features while their performance is rather poor for local and surface processes. The idea is thus to use modelled large-scale atmospheric circulation (also referred to as predictors) to reconstruct the ocean forcing fields (or predictants) through a transfer function built beforehand from the observations, or their best estimates. Change in the predictants for future climate is obtained through this transfer function via change in the statistical properties of the predictors in input (see e.g. Wilby et al. 2004 for a review).
- a dynamical step also referred to as dynamical downscaling. This step is based on the use of a high resolution ocean global circulation model (OGCM) that is expected to account for key mesoscale processes (e.g. oceanic deep convection, vertical mixing, etc.) known to be crucial for low frequency ocean changes that we want to reproduce for present climate or to predict in future scenarios (circulation, MOC, etc.)

Within this framework, ocean processes and nonlinear effects are better resolved through the dynamical step, while the mean and variability properties of the ocean forcing are constrained through the statistics of the predictors built from observed data though the statistical step.

A stratified strategy of validation for the different stages of the combined statistico-dynamical method is presented in two companion papers. The design of the transfer function from predictors to predictants, and more specifically the choice for predictors, are clearly at the core of the approach. The present paper (or Part I) is devoted to the assumption that weather typing techniques leading to representative large-scale atmospheric circulation patterns can be relevant to extract predictors to reconstruct the full Atlantic Ocean surface variables used ultimately as forcing for OGCMs. In other words, one has to check first as a prerequisite that the observed weather regimes in both the tropics and midlatitudes are discriminatory for ocean forcings that we further want to reproduce. One must verify in particular that a large fraction of the observed interannual to decadal variability of the ocean surface variables can be diagnosed over the entire Atlantic basin through the observed frequency distribution and intrinsic properties of the weather regimes. Part II will be devoted to the description and validation of the transfer function using the weather regimes as predictors to reconstruct the ocean variables (Minvielle et al. 2010). Part II will also present the performance of the statistical method when the reconstructed datasets are used to force a high resolution OGCM within a dynamical downscaling framework. Before applying these methods to future climate projections, their evaluation is necessary for present-day climate; both Part I and Part II are accordingly devoted to this task.

Part I is organized as follows: concepts and definitions for weather regimes, data and methods, are presented in Sect. 2. Wintertime observed weather regimes and their relationship with sea surface variables are illustrated in Sect. 3 for both extratropical and tropical domains. Their relevance as predictors to reconstruct wintertime sea surface variables at interannual time scale as well as their trend is presented in Sect. 4. Emphasis is laid in Sect. 5 on the importance of both the occurrence and the intrinsic properties of the regimes to capture the observed daily to interannual variability. Section 6 is devoted to summertime and the results are summarized and further discussed in Sect. 7.

2 Weather regimes: concepts and definition

The weather regime paradigm has been extensively used to describe the midlatitude atmospheric variability. Travelling synoptic pressure systems or storms contribute to a significant fraction of the daily to interannual variability of the extratropical climate. Those are linked to the unstable nature of the upper-level westerly jet stream and interact with circulation patterns of larger scale, or weather regimes, in which they are embedded. Weather regimes could be understood as envelopes for daily atmospheric variability; they have a typical 6–10 day nominal persistence, are spatially well defined and limited in number. They can be viewed, in midlatitudes, as the preferred and/or recurrent quasi-stationary atmospheric circulation patterns produced by the interaction between planetary-scale and synoptic-scale atmospheric waves (e.g. Ghil and Roberston 2002). The day-to-day meteorological fluctuations can be described in terms of temporal transition between regimes. The year-to-year (or longer timescale) climate fluctuations can be interpreted as changes in their frequency of occurrence provided the hypothesis of long-term quasi-stationary climate (Yiou et al. 2007). This climate-oriented interpretation for weather regimes is shared with the so-called continuum paradigm elaborated, for instance, in Franzke and Feldstein (2005) to better understand the low-frequency fluctuations of the Northern Hemisphere atmospheric teleconnection patterns.

The weather regime entity is expected to be a promising candidate as predictor for surface ocean variability from intraseasonal to decadal timescale. Viewed as an efficient spatio-temporal filter of the mostly chaotic atmospheric flow at midlatitudes, it appears to be compatible with the time-integrator properties of the ocean and might thus be relevant to explain part of its low-frequency changes. Associated surface winds and air temperature anomalies, displaced cyclone activity within a given regime, result in recurrent anomalous fluxes of sensible and latent heat which could create ocean temperature anomalies extending down to the base of the mixed layer. Statistically or technically speaking, weather regimes are classes of atmospheric circulation patterns gathered together from a similarity criterion. Those classes are defined by their mean conditions, or centroids, by their variance and by their frequency of occurrence. They are traditionally obtained from clustering techniques: the most common are hierarchical algorithms (e.g. the Ward classification, Cheng and Wallace 1993 among others) and the *k*-means approach (e.g. Michelangeli et al. 1995). More complex approaches, e.g. the Self-Organizing Map (SOM) method arising from the field of artificial neural network (Johnson et al. 2008, among others) have been also recently proposed. As reported in all studies, there is always a part of subjectivity associated with the spatial domain retained for computation of weather representative patterns as well as their number. Their existence is still even controversial (Christiansen 2007). In statistical reconstruction techniques, choices are usually adapted to the predictants to be reconstructed and to the time-scale characteristics of the underlying physical

mechanisms. The domain considered here for predictants is the full Atlantic ocean and our goal is twofold: (a) to reconstruct at best all the forcing variables over the entire basin for present day climate variability, (b) to extract at best the climate change information from scenario experiments over the entire basin that is linked to the perturbed large-scale atmospheric changes that we trust the most in IPCC-class CGCMs.

Based on the truly different nature of the atmospheric dynamics between the tropics and the extratropics, the Atlantic is divided into a mid-to-high latitude domain, hereafter North Atlantic–Europe domain (NAE, 20°N–80°N/90°W–30°E), and a tropical band between 20°N and 20°S. Two large scale atmospheric circulation fields from the European Centre for the Medium-Range Weather Forecasts (ECMWF) ERA40 reanalysis (Uppala et al. 2005) have been classified over 1958–2002 using the *k*-means approach leading to NAE weather regimes (henceforth NAE-WR) and tropical weather regimes, their combination serving as predictors for sea surface variables over the Atlantic Ocean as a whole. In the extratropics, we use daily averaged 500 hPa geopotentiel height (Z500) interpolated on a 2.5°×2.5° grid; in the tropical band, we use both the meridional and zonal components of the wind at 1,000 hPa (UV1000). Z500 and UV1000 data are weighted by the square root of cosine of latitude and the decomposition is done for computational efficiency in the Empirical Orthogonal Functions (EOFs) phase space span by the first 20 EOFs and principal components retaining 90% and 85% of the total Z500 and UV1000 daily variance, respectively. Even if the seasonal cycle has been removed at each grid point (we refer to the resulting quantities as anomalies in the following), two seasons are considered to account for the seasonal latitudinal migration and strength of the Atlantic pressure centers of action.

The large size of the domains considered here contrasts with traditional downscaling or reconstruction techniques usually devoted to regional application (e.g. Boé and Terray 2008). NAE and tropical regimes appear to be a good compromise in our case because, as detailed in the course of the paper, they both capture the coherent variability among surface variables over the entire Atlantic, a property that must be respected in the reconstructed forcing when passed to the ocean model (see Part II).

3 Atlantic wintertime atmospheric circulation classes and associated surface ocean variables

3.1 North Atlantic-Europe Z500 weather regimes

Wintertime (December-March) is first documented in this section. Consistently with previous literature (e.g. Vautard 1990), four NAE-WR are obtained (Fig. 1a–d). The number of clusters ($k = 4$) has been chosen following statistical tests (not shown) based on reproducibility (Michelangeli et al. 1995), sampling dependence and variance ratio consideration like in Straus et al. (2007). The possibility of perhaps two regimes as documented in recent papers (Woollings et al. 2010, Hannachi 2010) was envisioned because the $k = 2$ partition also emerges from the statistical tests, but this option was abandoned because subsequent results (forecast skill like in Fig. 5) are considerably degraded in that case. We here use raw data for clustering as opposed to Fereday et al. (2008) in which filtering is applied before classification; we tested that results are indeed very similar for both options. Once the representative patterns determined, we classify each daily Z500 field in the 44-year dataset to one of the four regimes using the criterion of minimum Euclidian distance. The corresponding temporal evolution of the total number of days attributed to a given regime over the complete winter season is given in Fig. 1e, h. The two-first Z500 regimes can be viewed as the negative and positive phase of the North Atlantic Oscillation (NAO− and NAO+, respectively). Their respective occurrence is indeed strongly correlated (−0.89 and +0.73) to the traditional wintertime NAO index estimated from EOF (Fig. 1e, f). Note that spatial asymmetries between the two NAO phases are clearly evidenced here by clustering techniques that do not make any assumption for linearity. The year-to-year occurrence of the NAO regimes captures the strong interannual-to-interdecadal variability of the oscillation that has been extensively documented in literature (see e.g. Hurrell et al. 2003 for a review). The third regime is named Atlantic ridge (AR) and is reminiscent of the so-called East Atlantic pattern (negative phase, Barnston and Livezey 1987) viewed as a Euro-Atlantic wave train. The forth regime is often referred to as Scandinavian blocking (S-BL) and is characterized by a strong anomalous height anomaly over Northern Europe (Tyrlis and Hoskins 2008) while a mild deeper trough extends southeastward from the Labrador Sea to the Iberian Peninsula. These two regimes all together are excited about 50% of time on average in winter and exhibit considerable interannual variability (Fig. 1g, h, less loading in the decadal frequency band compared to NAO regimes).

Links between flow regimes and anomalous climate conditions over continental Europe have been extensively documented in literature from daily (e.g. Slonosky and Yiou 2001; Philipp et al. 2007) to decadal timescales (Hurrell 1995). ERA40 anomalous 10-m wind (hereafter UV10) and 2-m temperature (hereafter T2) associated with NAE-WR are shown in Fig. 2 over the ocean based on composites. The NAO regimes can be essentially considered as a measure of the strength of the prevailing

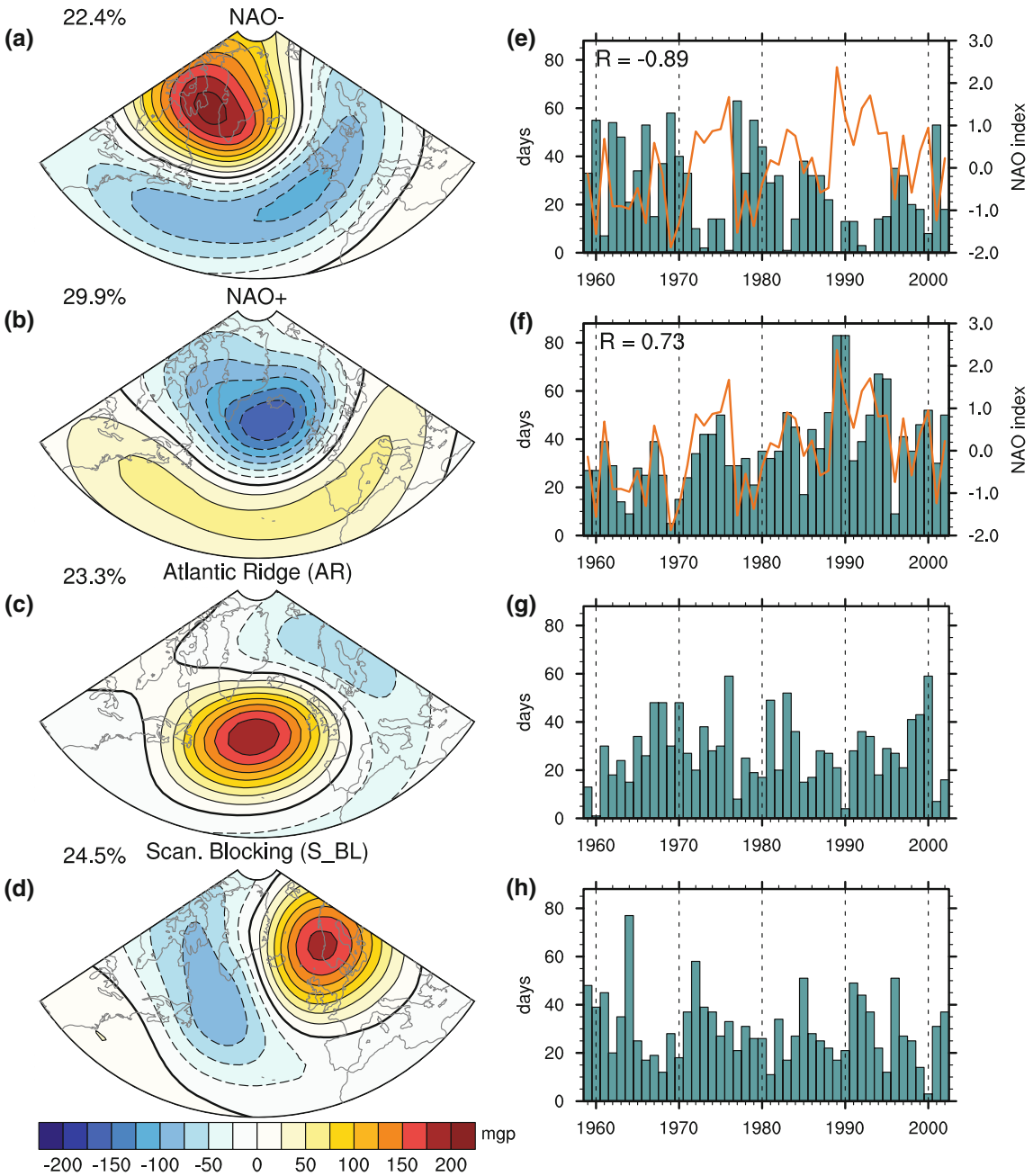

Fig. 1 **a–d** Centroids of the four wintertime NAE Z500 weather regimes (m). Each percentage represents the mean frequency occurrence of the regime computed over 1958–2002 from 1 December to 31 March. Contour intervals are 25 m. **e–h** Number of days of occurrence of each regime per winter from 1959 to 2002. The NAO index (*orange curve*) defined here by the normalized principal component of the leading EOF of averaged DJFM Z500 is superimposed on the upper two panels corresponding to the NAO regimes. Correlation (R) between the NAO index and the frequency of occurrence of the NAO regimes is provided

westerlies flow over the North Atlantic as well as the trade winds in the subtropics. Spatial asymmetry is found though in wind patterns between the two NAO regimes in agreement with the longitudinal shift of the anomalous centers of action and associated storm track that is intrinsically linked to two NAO phases (Cassou et al. 2004). Such an asymmetry is expected to have an important impact on the ocean/sea ice dynamics because of the nonlinearity properties of the processes governing the gyre circulations and the ocean mixed layer variability (vertical mixing, turbulent sensible and latent heat fluxes controlling a large part of the buoyancy, etc.). For instance it is clearly visible here

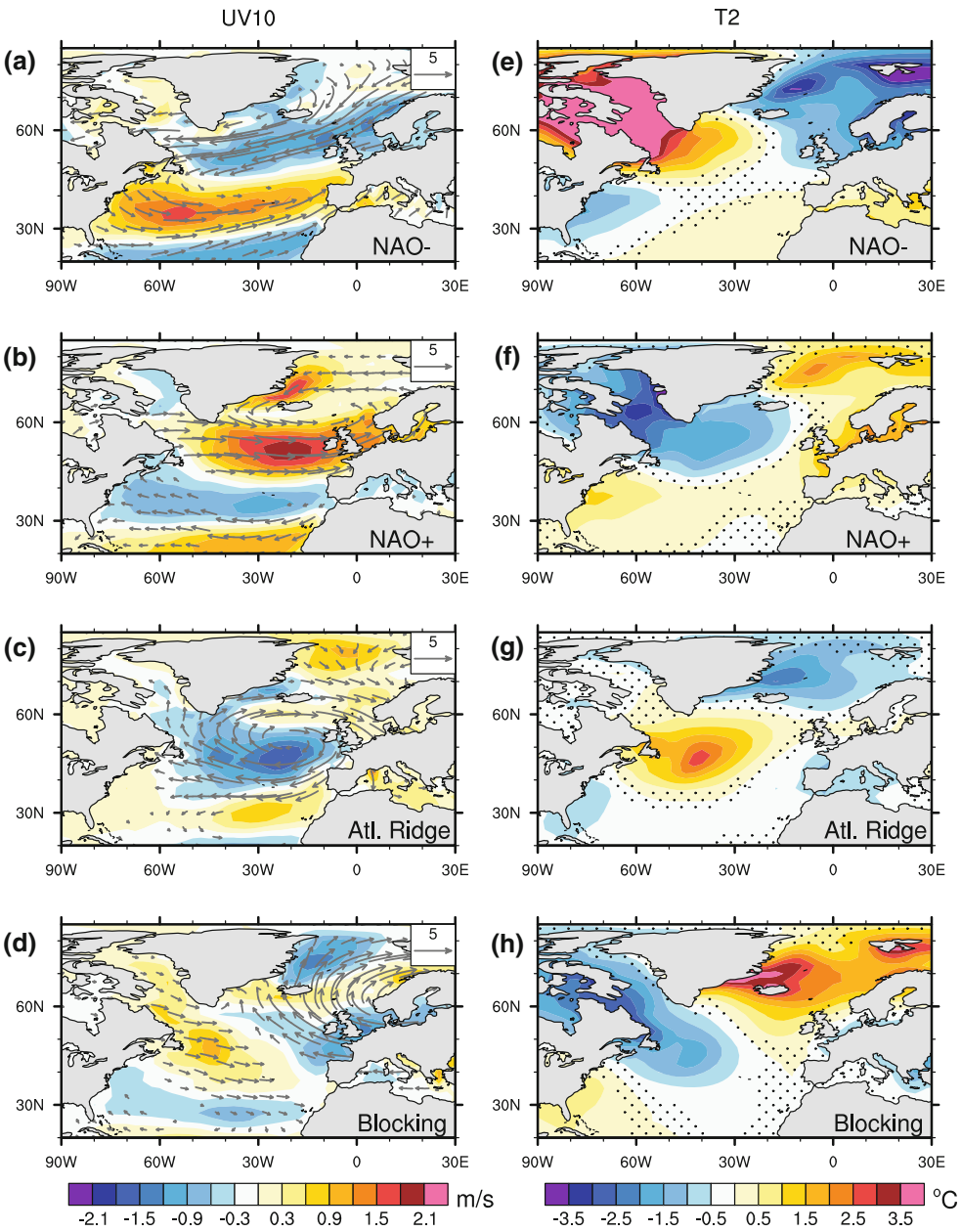

Fig. 2 Daily derived-regime anomaly composites for (**a**–**d**) 10-m wind (*arrows*, UV10) and its module, and for (**e**–**h**) 2-m temperature (T2). Contour intervals are 0.15 m s^{-1} and 0.25°C respectively. *Dotted shading* stands for non-significant area based on *t* statistics at the 95% level of confidence. Non significant wind anomaly *arrows* are omitted

that the occurrence of NAO− regimes will not be discriminatory to explain anomalous sea-ice export through the Fram Strait while NAO+ clearly will (Hilmer and Jung 2000). AR is associated with a strong surface anomalous cyclonic circulation located downstream the maximum Z500 high. It corresponds to diminished (reinforced) westerlies off Western Europe (along 60°N) while reinforced northerlies (southerlies) conditions dominate from 10°W eastward (in the Labrador Sea). The long-lasting high pressure center over Europe in S-BL is powerful enough to completely disrupt the prevailing westerly flow penetrating inland. Easterlies dominate off-Europe and the recirculation branch associated climatologically with the Icelandic Low is stopped in the Nordic Seas.

Interestingly, T2 anomalies linked to NAE-WR are rarely collocated with maximum wind changes (Fig. 2e, f). A strong cooling (warming) occurs in the Labrador Sea for NAO+ (NAO−) while the Nordic Seas as well as the midlatitudes are warmer (colder). It is clear that in addition to the direct impact of wind intensity changes, the mean advection of anomalous temperature responsible for altered local heat exchanges through turbulent fluxes over the ocean (Cayan 1992; Deser and Timlin 1997) is crucial to understand the NAO regime-derived T2 composite. This

inference holds for AR and S-BL where, respectively, maximum T2 anomalies are found off Newfoundland, i.e. upstream of the anomalous high core and maximum wind speed disturbance, and in the Nordic Seas.

3.2 Tropical Atlantic UV1000 wind classes

Similarly to midlatitudes, clustering is applied to tropical dynamics (20°N–20°S) with some adaptations. First UV1000 is preferred as predictor for ocean surface fields instead of Z500. The latter is not relevant anymore because of the baroclinic nature of the tropical variability while extratropical fluctuations are mostly barotropic due to strong eddy-mean flow interaction. Second, it is essential to capture the true coupled nature of the tropical variability between the surface ocean and the surface atmosphere. Because of this coupling and the intrinsic thermal inertia of the ocean, daily variability in the tropics is consequently weaker than at midlatitudes and leads to enhanced energy in power spectrum from intra-seasonal timescale onward. To preserve the coupled properties of the tropical Atlantic variability, we applied a 30-day low pass filter to UV1000 daily maps prior to classification based on Peña et al. (2004) who suggests that uncoupled atmospheric noise is concentrated in the 1–30 day frequency band. We verified though that the main results are insensitive to filtering. Because the core of the episodic paradigm associated with extratropical regimes is somehow violated in the tropics due to the nature itself of the tropical dynamics, and thus in order to avoid any possible misinterpretations of our conclusions, we will prefer the term "wind classes" to the term "regimes" to refer to the tropical UV1000 centroids (henceforth T-WC) in the following,

Four wintertime T-WC are obtained from *k*-means clustering and related statistics, and are displayed in Fig 3. Strengthened (relaxed) trade winds in both hemispheres are captured in T-WC1 (T-WC2) with maximum loading in the northern basin though. Significant changes are mostly found in the southern hemisphere for T-WC3 and T-WC4 encroaching on the equator. When Southeast (SE) trades are slackened (reinforced) in T-WC3 (T-WC4), NE trades are concurrently slightly intensified (diminished), the node of the anomalies being collocated around 7°N on the northward side of the climatological position of the Inter Tropical Convergence Zone (ITCZ). As opposed to NAE-WR, note that T-WC centroids have a fair amount of linearity in terms of spatial structures.

T-WC occurrences are dominated by interdecadal variability over 1958–2002, except for T-WC3 characterized by some interannual pulses superimposed on a significant trend. T-WC1 is almost totally absent before 1970 while T-WC2 virtually disappears post to 1970 (Fig 3e, f); such a shift can be simply interpreted as an intensification of the mean trades from the mid-1970's onward. The discriminatory skill of the UV1000 classification to capture such a decadal fluctuation is further confirmed in Fig 4 based on probability density functions (pdf) for raw daily 10-m winds (UV10) averaged over two selected domains. The pdf diagnosed from the whole dataset (black dashed line) is compared to the four-ones computed for each wind class taken separately (color curves). UV10 pdfs associated with T-WC1 and T-WC2 are clearly distinct (Fig. 4a) in the northern basin (10°–20°N/50°–30°W): upper quantiles in the full dataset clearly fall in T-WC1 while weak-to-moderate surface wind days belong to T-WC2. Note that the mean statistics of the direction of the surface wind assessed from the wind rose added in the upper-left corner of the graph is the same in both T-WC1 and T-WC2. The late 1960s–early 1970s decadal rupture detected here from wind clustering has been documented in many papers from other fields such as convection over adjacent continents (Sahel and Amazonian basin); in fact, it is part of world wide coordinated changes associated with a rapid global climate shift whose signatures are particularly pronounced in the Atlantic sector (see Baines and Folland 2007 for a detailed description). The origin of this shift in the tropical Atlantic atmosphere has been linked with more and more confidence to interdecadal changes in either local (Atlantic Multidecadal Oscillation—AMO, Knight et al. 2006) or remote tropical SSTs (e.g. Giannini et al. 2003).

A similar pdf analysis has been carried for a region straddling the equator in the western part of the tropical Atlantic basin (2°N–2°S/35°–20°W, Fig. 4b). Results suggest that the four wind classes are not discriminatory there for surface wind speed: the four pdfs, except maybe for T-WC4, are almost superimposed to each other. The wind roses for T-WC3 and T-WC4 show however that the latter classes are very skilful in capturing the direction properties of the surface wind along the equator. The proportion of northeasterly days is equal to 7% in T-WC4 whereas they represent about 40% for T-WC3. It appears rather complex to link the occurrence of T-WC3 and T-WC4 (Fig 3gh) to any specific SST pattern because it is very likely that several processes, operating on a range of timescales, play a role. Their occurrences are mostly correlated to the so-called interhemispheric gradient SST mode (Ruiz-Barradas et al. 2000) but are modulated by Atlantic Niño-type anomalies along the equator (Zebiak 1993; Carton and Huang 1994 among others) and by El Niño Southern Oscillation (ENSO) whose remote influence is maximum in DJF (Sutton et al. 2000). Note that all the T-WCs are tightly linked to both the position and strength of the Atlantic ITCZ. Finally, it is also evident that the equatorward penetration of the NAO and associated decadal variability blur the entire picture in winter as detailed in the next section.

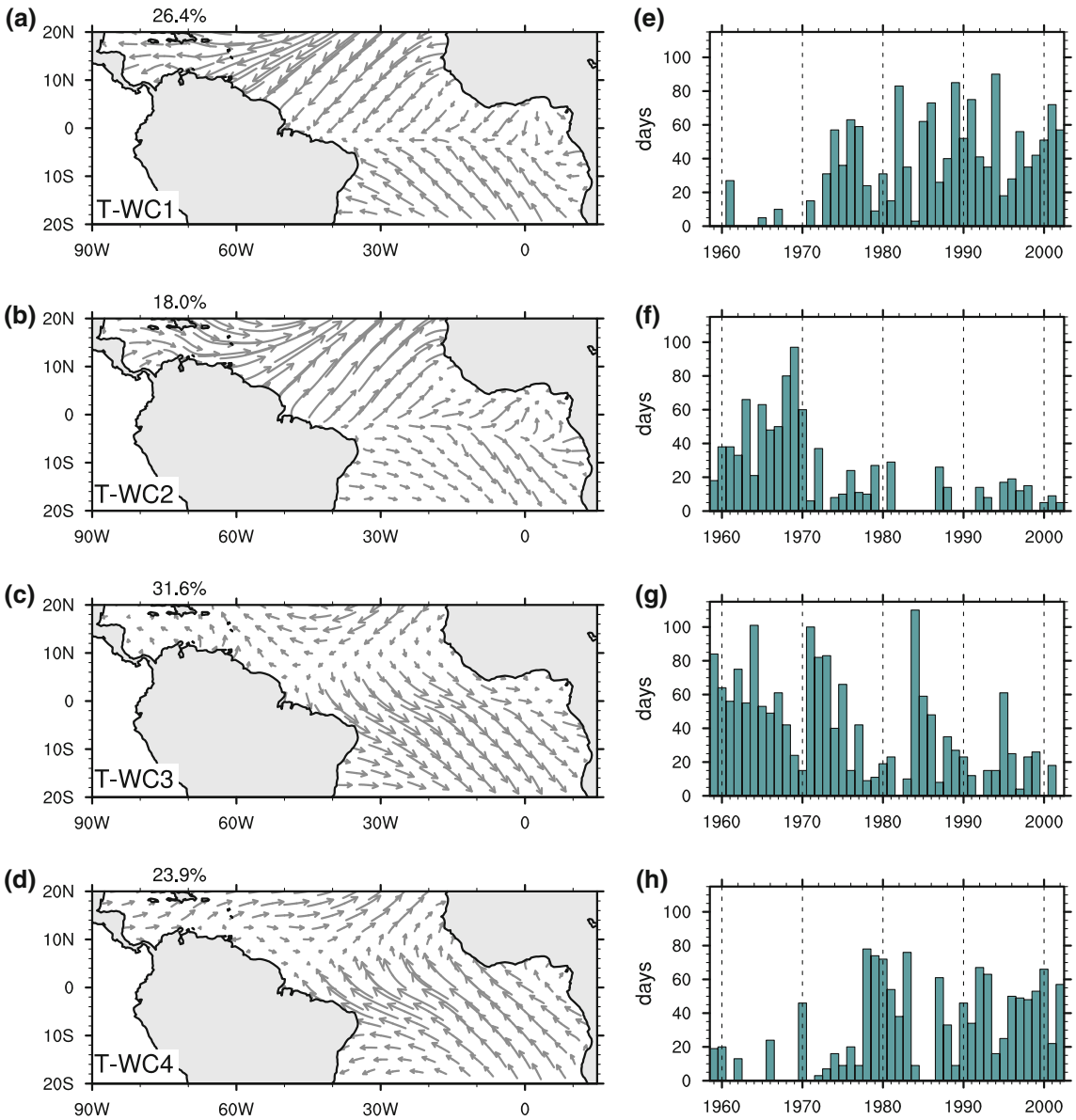

Fig. 3 **a–d** Centroids of the four wintertime tropical UV1000 classes. Each percentage represents the mean frequency occurrence of the class computed over 1958–2002 from 1 December to 31 March. **e–h** Number of days of occurrence of each wind class per winter from 1959 to 2002

4 Link between occurrence of atmospheric circulation and ocean low-frequency variability

4.1 Interannual variability of surface ocean variables

The ability of NAE-WR and/or T-WC to capture the observed low-frequency variability of the surface ocean variables over the Atlantic is now examined using least-square multiple linear regressions. The December–March frequencies of occurrence of the circulation patterns are used as predictors while U10, V10 and T2 are the predictants of the regression models. The interannual band is first addressed: time series of sea surface fields and frequencies of occurrence of regimes/classes are high pass filtered using a Lanczos filter to retain periods shorter than 7 years (classical frequency cutoff to separate interannual from decadal fluctuations in the Atlantic, e.g. Giannini et al. 2003). Correlation maps between observed and reconstructed fields are shown in Fig 5 from three different multiple regression models. Predictors are alternatively limited to NAE-WR or to T-WC in order to better isolate the respective role of the extratropical and tropical atmospheric dynamics on the interannual fluctuations of the surface ocean fields; their combination is additionally illustrated. We verify that multicolinearity between the predictors (correlation between two or more explanatory

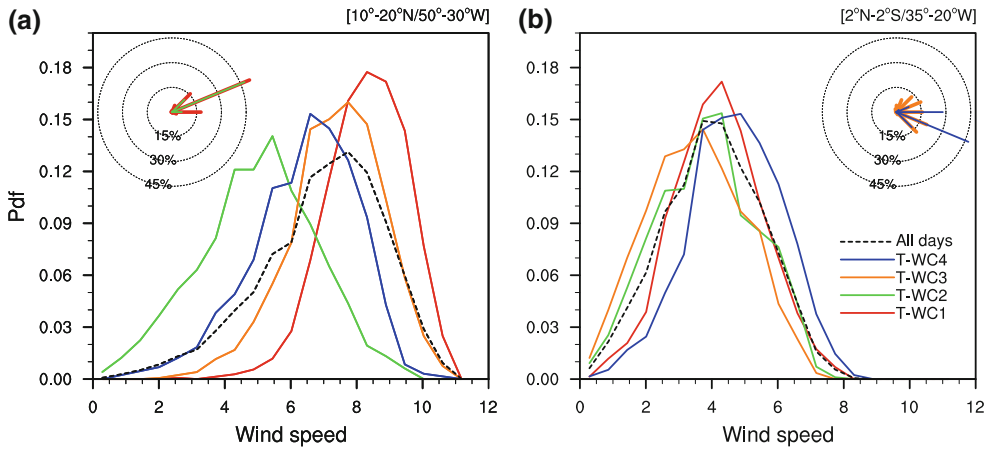

Fig. 4 10-meter wind speed (m s^{-1}) distribution within each tropical wind classes (*colored lines*) and for the whole dataset (*dashed black line*) for (**a**) the north tropical Atlantic basin (10°–20°N/50°–30°W), (**b**) along the equator (2°N–2°S/35°–20°W). Wind roses are shown for T-WC1 (*red*) and T-WC2 (*green*), and for T-WC3 (*orange*) and T-WC4 (*blue*) in the *upper-left* and *upper-right* corner, respectively. Each *circle* represents the percentage of time the wind blows from a particular direction. Contour interval is every 15% and the wind dial is divided into 16 directions represented by a color line depending of the wind class

variables, von Storch and Zwiers 1999) does not actually bias results. An alternative way to reconstruct surface fields would have been to multiply the frequency of occurrence of the circulation classes by the derived anomaly composite maps displayed in Fig 2, following Johnson et al. (2008) formulation applied to SOM-derived patterns. We verify that the latter approach also produces very similar results.

U10 correlation patterns for the NAE-WR regression model (Fig. 5a) range from 0.4 to 0.8 over the entire extratropical domain. Maximum values extend zonally along two latitudinal bands projecting on the actual NAO anomaly composite pattern for surface wind (Fig. 2a, b). Weaker, though significant, correlation penetrates southward in the tropical basin while skill is lost south of the ITCZ. As expected, most of the interannual variability for U10 in the tropics is captured in the T-WC regression model, especially in the northern basin where correlation values locally reach 0.7 (Fig 5b). Note that a narrow band of significant correlation is found off Newfoundland in the latter model and is indicative of the existence of a connection between tropics and extratropics as further detailed below. Using NAE-WR + T-WC as predictors is virtually additive for U10 interannual variability over the entire basin (Fig. 5c). This conclusion holds for V10 (Fig. 5f). The best reconstruction is found over a large Northeastern Atlantic basin and can be attributed to S-BL and AR signals (Fig. 5d). Fluctuations in trade winds and meridional wind changes along the equator associated with the anomalous position of the ITCZ are well captured in the T-WC regression model (Fig. 5e) consistently with T-WC3 and T-WC4 distribution properties discussed in the previous section. Combining NAE-WR and T-WC allows to reproduce V10 interannual variability relatively well over the entire basin, except at midlatitudes from Florida to the Azores archipelago. As to T2, much of the interannual variability is well reproduced at midlatitudes (north of 40°N) based on the sole NAE-WR frequencies of occurrence. The highest correlations surround Europe and are found in the Labrador and Irminger Seas (Fig. 5g). By construction, skill for T2 is poorer in the T-WC regression model compared to U10 and V10 because surface wind components are tightly linked to 1,000 hPa wind used for clustering while T2 is not (Sutton et al. 2000). The correlation pattern for T2 and T-WC regression model (Fig. 5h) is reminiscent of the North Atlantic tripole (Deser and Timlin 1997) and suggests again that T-WC occurrences are associated with anomalous extratropical atmospheric circulation through teleconnection.

The table of contingency between T-WC and NAE-WR daily occurrence confirms that extratropical and tropical dynamics are clearly not independent (Table 1) on a daily basis. Table 1 should be read in column. When T-WC1 is excited, NAO+ is overly prevalent (about 50% of the days) while NAO− and AR occurrences are clearly reduced. The opposite is found for T-WC2 and a similar balance between the four NAE-WR (NAO−/AR vs. NAO+/S-BL) is also obtained, although less pronounced, for T-WC3 and T-WC4. The tropical-extratropical relationship explored here through clustering is consistent with previous pictures described in literature. Evidence has been provided based on modeling (Sutton et al. 2001; Cassou et al. 2004 among others) that altered NE trades lead to changes in the probability of occurrence of NAE regimes via the modification of the local Hadley cells (Terray and Cassou 2002). The NAO regimes are the most sensitive to the tropical forcing which likely explains the tripole structure found in

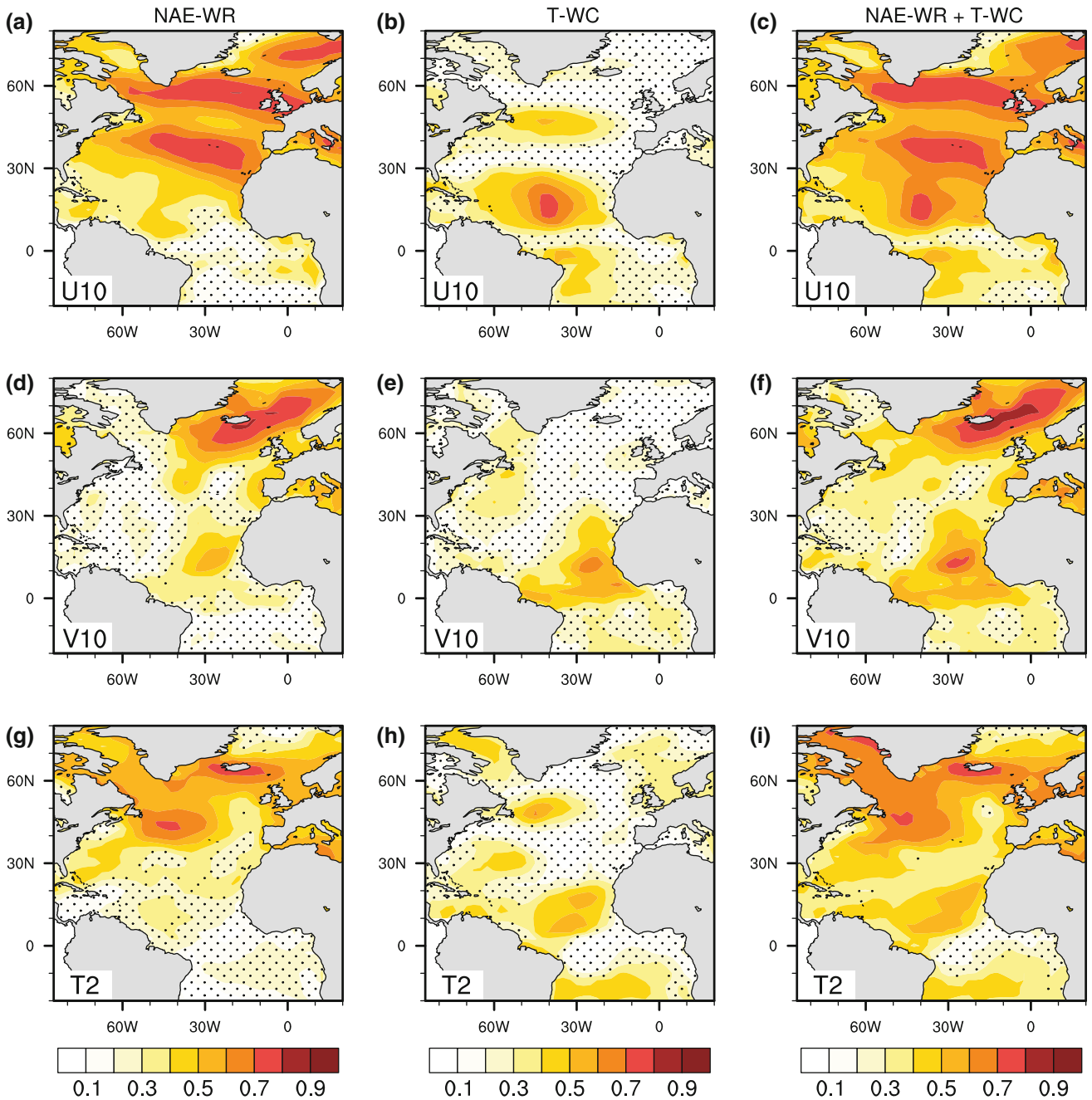

Fig. 5 Interannual linear correlation between wintertime ERA40 10-m zonal wind (U10, **a–c**), 10-m meridional wind (V10, **d–f**) and 2-m temperature (T2, **g–i**) and reconstructed fields via multiple linear regression when NAE-WR (*left panels*), T-WC (*middle panels*) frequencies of occurrence and their combination (*right panels*) are used as predictors. Contour interval is 0.1. Shading stands for non-significant correlation based on R-statistics at the 95% confidence level

the correlation map for T2 based on the occurrence of the sole T-WCs used as predictors (Fig. 5h). Conversely, in agreement with Sutton et al. (2000), the equatorward extension of the NAE-WR imprint highlighted in Fig. 5a (Fig. 5d) for U10 (V10) also supports the conclusion that a significant part of the northern tropical Atlantic atmospheric variability is controlled by the southward penetration of the extratropical modes. The NAO appears to be the most important such modes (Fig. 2).

4.2 Low-frequency trends of surface ocean variables

We use linear trend analysis as a simple way of characterizing changes in atmospheric circulation that have

Table 1 Table of contingency (in %) between T-WC and NAE-WR in winter

	T-WC1	T-WC2	T-WC3	T-WC4
NAO−	11.3	32.4	20.8	29.2
NAO+	48.6	12.5	32.2	19.1
AR	12.9	41.8	15.7	30.8
BL	27.5	13.2	31.2	20.9

The sum of percentages for a given T-WC (column) is equal to 100%. The sum of percentages for a given NAE-WR is not by construction because regimes and classes are not equally excited

occurred over 1958–2002. Seasonal anomalies are formed by subtracting the long-term seasonal mean based on the full period and the DJFM trends are computed using least squares linear regression. Observed trends are shown in Fig. 6 for U10, V10 and T2 (upper panels). Large-scale atmospheric circulation changes are characterized in the Atlantic by an intensification of the midlatitude westerlies and a concomitant strengthening of the NE trades (Fig. 6a) and easterlies along the equator. Northward displaced storm track is associated with prevalent southerlies north of 40°N whereas equatorward convergence associated with a southward retreat the ITCZ is enhanced (Fig. 6b) due to enhanced trade winds in both hemispheres. The node for the wind trend corresponds to the mean position of the Azores High and is consistent with the positive trend in the NAO (Hurrell 1995). Consistently, the largest trends for T2 occur in the western part of the North Atlantic and strongly project on the North Atlantic tripole (Fig. 6c). The causes of observed atmospheric circulation trends over the last half century are a subject of ongoing investigation. Latest studies suggest that direct atmospheric radiative forcing due to greenhouse gases (GHG) concentration and indirect SST forcing drive circulation responses that contribute about equally to the global pattern of circulation trends (Deser and Phillips 2009).

The pattern of trends reconstructed from multiple regression using raw NAE-WR + T-WC frequencies of occurrence as predictors compares very well with observations (Fig. 6, lower panels). The pattern correlations are 0.98, 0.93, 0.98 for U10, V10 and T2, respectively. The magnitudes of the reconstructed trends are slightly stronger than observed north of 40°N, while they are slightly

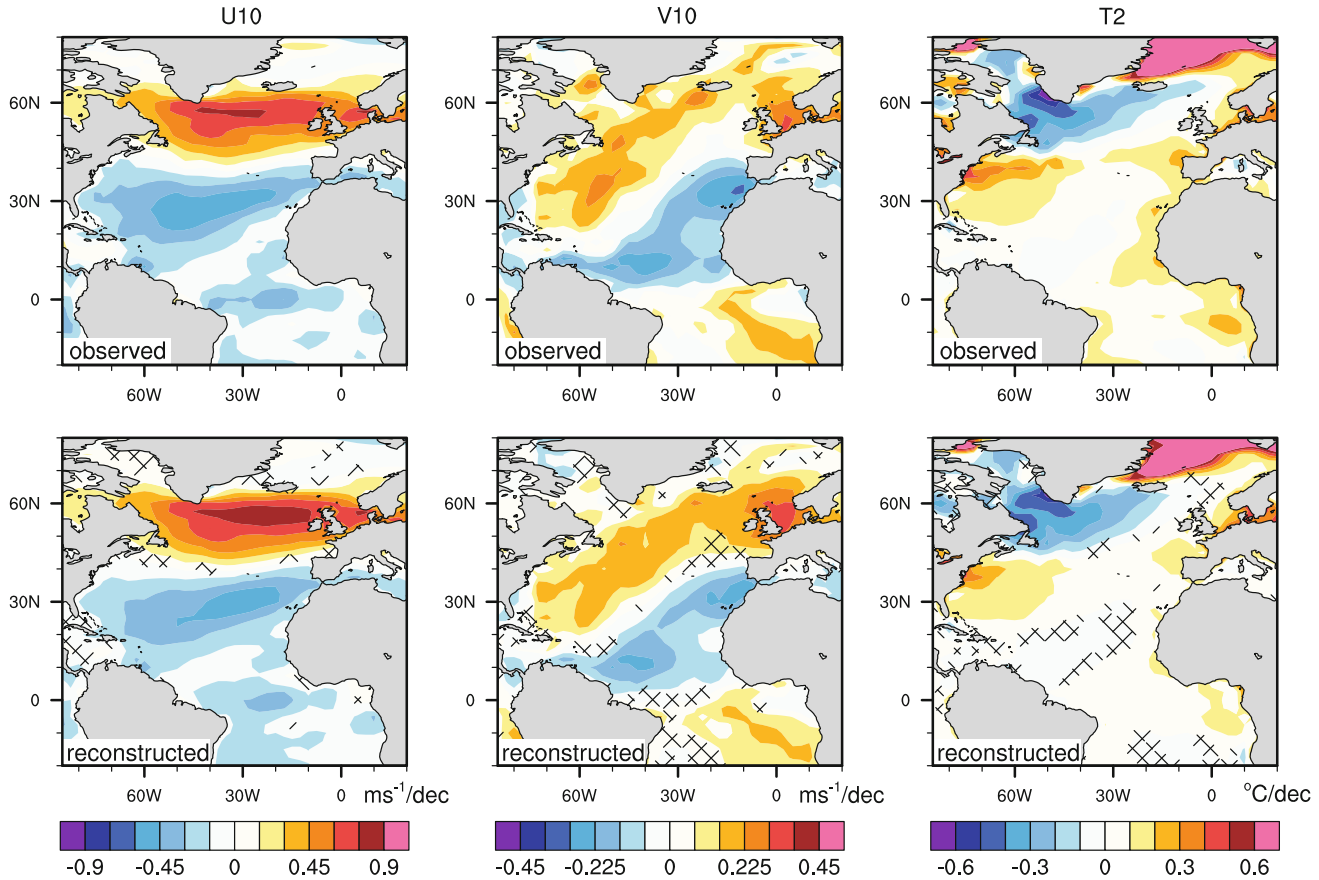

Fig. 6 Linear trends computed over 1959–2002 for wintertime (**a–c**) ERA40 U10, V10 and T2 and for (**d–f**) reconstructed fields from multiple linear regression using NAE-WR + T-WC occurrences as predictors. Hashing stands for sign disagreement between observed and reconstructed trends. Contour intervals are 0.15 m s^{-1}/decade, 0.075 m s^{-1}/decade for UV10 and V10, respectively, and 0.1°C for T2

underestimated over a large tropical band. Maximum discrepancies are found for T2 where the observed warming along the African coast is barely captured whereas the weak cooling over the northwest tropical basin is too broad in the reconstruction compared to observed estimates. Based on our findings, we conclude that observed trends in surface ocean fields can be viewed to a large extent as the temporal integration of the anomalous frequencies of occurrence of both NAE-WR and T-WC over the last 50 years or so. Similar conclusions have been put forward for trends over land (e.g. Hurrell 1995; Boé and Terray 2008) and appear to be fairly robust at midlatitudes. By contrast, underestimated warming south of 40°N likely indicates that the T2 tendency is additionally determined by influences other than the sole atmospheric dynamics; direct radiative contribution from increased GHG concentration appears to be a reasonable candidate, as well as the AMO.

4.3 Subsurface ocean

We now proceed to examine if the subsurface ocean variability in addition to the surface is also linked to the temporal excitation/recurrence of the regimes. We concentrate on the subsurface variability in the Labrador Sea where deep convection during winter forms and transforms dense water masses known to play a role in the 3D circulation of the ocean. Systematic physical observations have been collected over the past two decades along the so-called AR7W line extending from the Labrador Shelf to the Greenland Shelf straddling the main deep convection core (Våge et al. 2009). Since 2002, observations from the Argo floats nicely complement the AR7W survey but the period of the two combined ocean datasets is too short and the spatial coverage too sparse to assess a robust glimpse of the interannual variability of the subsurface. To circumvent this limitation, we make use of three ocean reanalyses products carried out within the ENSEMBLES EU-project (http://www.ecmwf.int/research/EU_projects/ENSEMBLES/index.html) over 1960–2005, hereafter referred to as CERFACS (Daget et al. 2009), ECMWF (Balmaseda et al. 2008) and UKMO (Martin et al. 2007). All are state-of-the-art model-data assimilation systems and are considered as the best estimates available today for a 3D coverage of the ocean variability. They differ in their assimilation scheme as well as in the OGCM used, but share the same atmospheric forcing derived from ERA40 or operational ECMWF analyses post to 2002. It is beyond the scope of this paper to document properties and performances of each product; we simply combine them in the present study to evaluate the uncertainties of our results.

Annual vertical profiles for ocean potential temperature and salinity are computed over a box straddling the ARW7 line and encompassing the deep convection area in the three reanalysis datasets. Absolute values of correlation as a function of depth between temperature and NAE-WR frequencies of occurrence are maxima between 500 and 1200 meter depth for NAO+ and NAO− where LSW are produced (Yashayaev and Loder 2009), and are close to zero through the entire ocean column for S-BL and AR (Fig. 7a). Significant correlation extends deeper for NAO+ regimes in all reanalysis products and is indicative of more active (shallow) convection during positive (negative) NAO winters as extensively documented in literature from observations (e.g. Dickson et al. 1996). The density properties of the reanalysed LSW as a function of the wintertime integrated occurrence are further evaluated for NAO− only from a T–S diagram at 750-m depth (Fig. 7b). Repeated (rare) appearance of NAO− days in winter tends to form warmer (colder) LSW while salinity is somehow less affected. Consequently, lighter waters are produced inhibiting the convection and the ventilation of the deep ocean during NAO− winter. The opposite is found for NAO+ (not shown). Even if there is not a one-to-one correspondence between the NAO and the production and the density properties of LSW since preconditioning and/or large-scale horizontal advection play a role in regulating the depth of the convection (Yashayaev 2007), we suggest that the regime approach adopted here seems to be relevant to capture part of the 3D ocean variability.

5 Link between strength of anomalous atmospheric circulations and surface ocean variables

The results presented so far establish that a large fraction of air-sea low-frequency variability can be traced back to the anomalous frequencies of occurrence of atmospheric circulation classes. To address more thoroughly the relevance of the regime paradigm in capturing surface ocean changes, we now evaluate how the strength of the atmospheric circulation within a given regime and/or with respect to the others can provide an additional insight. We include the Euclidean distance as a measure of similarity between the centroids of the regimes and a given day's anomalous circulation. Let D be that given day. We define the "intra-regime distance" as the distance between the anomalous circulation of D and the centroid of the regime whose D belongs to. We term "inter-regime distances" the distances between the anomalous circulation of D and the centroids of the other regimes. Because spatial resemblances are strong by construction within the pool of days attributed to the same regime, intra-regime distances mostly provide information on the differences in the anomaly magnitudes between D and its associated regime centroid; it can be viewed as the strength of the anomalous circulation.

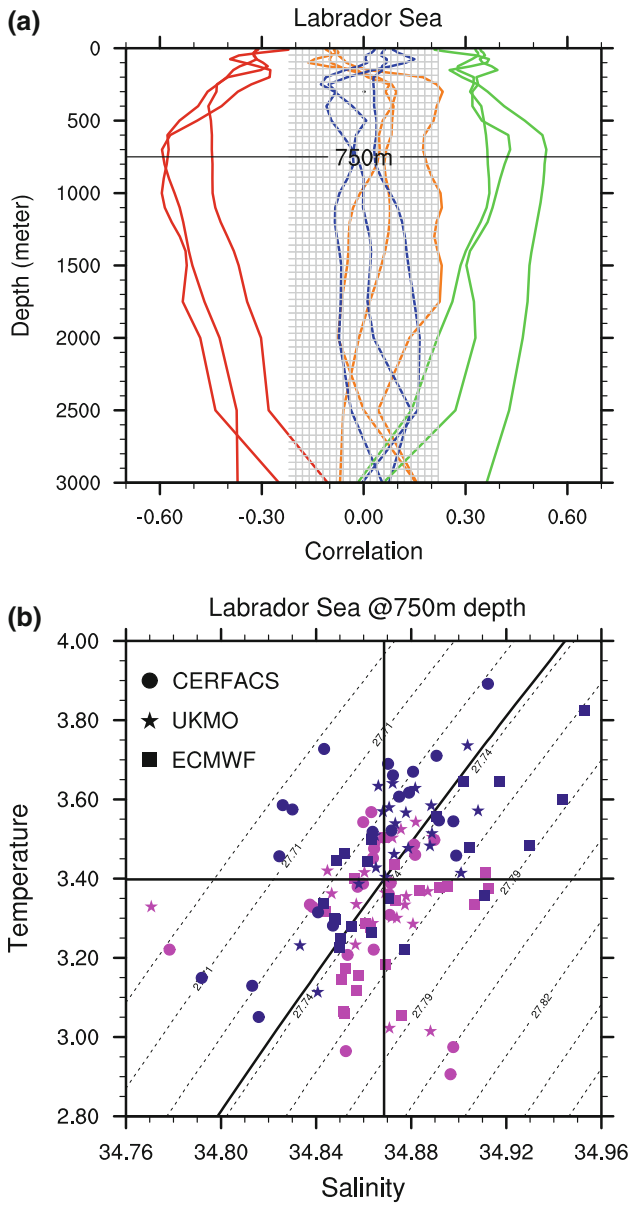

Fig. 7 **a** Correlation as a function of depth between the wintertime frequency of occurrence of the four NAE-WR regimes (*red* NAO+, *blue* S-BL, *orange* AR, *green* NAO−) and the corresponding annual ocean temperature box-averaged over 55°N–58°N/55°W–47°W across ARW7 for the three reanalysis products. Hatching stands for significance estimated from *t* statistics (**b**) T-S diagram at 750-m depth for the three reanalysis products; one point represents 1 year, the *blue color* (*magenta*) stands for above-average (below-average) occurrence of winter NAO− regime. Black thick lines correspond to the temperature, salinity and density means of the total sample

Figure 8 illustrates the relative importance of intra- versus inter-regime distances to interpret the strength of the midlatitude westerlies (upper panels). A daily index of anomalous U10 averaged over a box corresponding to the climatological position of the North Atlantic westerlies is computed. Intra-regime is investigated first (Fig. 8a). For NAO+, the amplitude of the anomalous westerlies is clearly independent on the intra-regime distance decomposed into deciles. In others words, it is suggestive that the more or less reinforcement of the westerlies when NAO+ is excited is not controlled by the strength of the NAO+ circulation itself. This conclusion is also valid for the other three regimes but to a lower extent though. NAO− slackened westerlies are all the more diminished when NAO− intra-regime distances are greater likely suggesting that the strength of the anomalous NAO− circulation (especially for extreme NAO− days) slightly matters for the amplitude of wind anomalies. Overall, Fig. 8a shows that the sign of the basin-wide westerlies anomalies (scale relevant for ocean low frequency variability) is primarily set by the excitation of the regime itself and that intra-dynamical characteristic within the excited regime plays a secondary role. Figure 8b suggests however that inter-regime distances are crucial to modulate the westerlies anomalies for a given regime, here NAO+. The three inter-regime distances between NAO + days and the other regimes are computed and classified into deciles similarly to Fig. 8a. Results indicate that the amplitude of the reinforced westerlies during NAO+ is clearly controlled by the distance to NAO− and/or AR centroids. The closer to NAO− and/or AR the circulation of the NAO+ day is, the weaker the westerlies intensification is following a quasi-linear relationship.

A similar analysis has been conducted for T2 over the Labrador Sea (Fig. 8c, d). Like for U10, we find that intra-regime distances do not play a primary role in modulating the T2 anomalies, except for the upper-decile of the NAO+ and NAO− regimes; cooling (warming) is more pronounced when NAO+ (NAO−) circulations are marked. The sign of the anomalies is clearly set by the regime and is consistent with the anomalous composite patterns displayed in Fig. 2. The inter-regime distances act however as a strong modulator of the amplitude of the T2 anomalies. "Very far" NAO+ days from the NAO− centroid are associated with very cold T2 anomalies while "very close" NAO+ days to NAO− centroid lead to normal T2 conditions. Both AR and NAO− directions in the regime phase space are found to be discriminatory for the magnitude of T2 anomalies during NAO+ in the Labrador Sea while the S-BL direction appears to be less of importance. The importance of inter-regime distances is not specific to the NAO+ regime but is also valid for the three others (not shown).

To go further, linear correlations between on one hand the Labrador T2 daily values, and on the other hand the daily distance to the NAE-WR centroids, are computed considering the four regimes separately (Table 2). This diagnostic is complementary to Fig. 8 and helps to determine the main direction or axis of the intra-regime

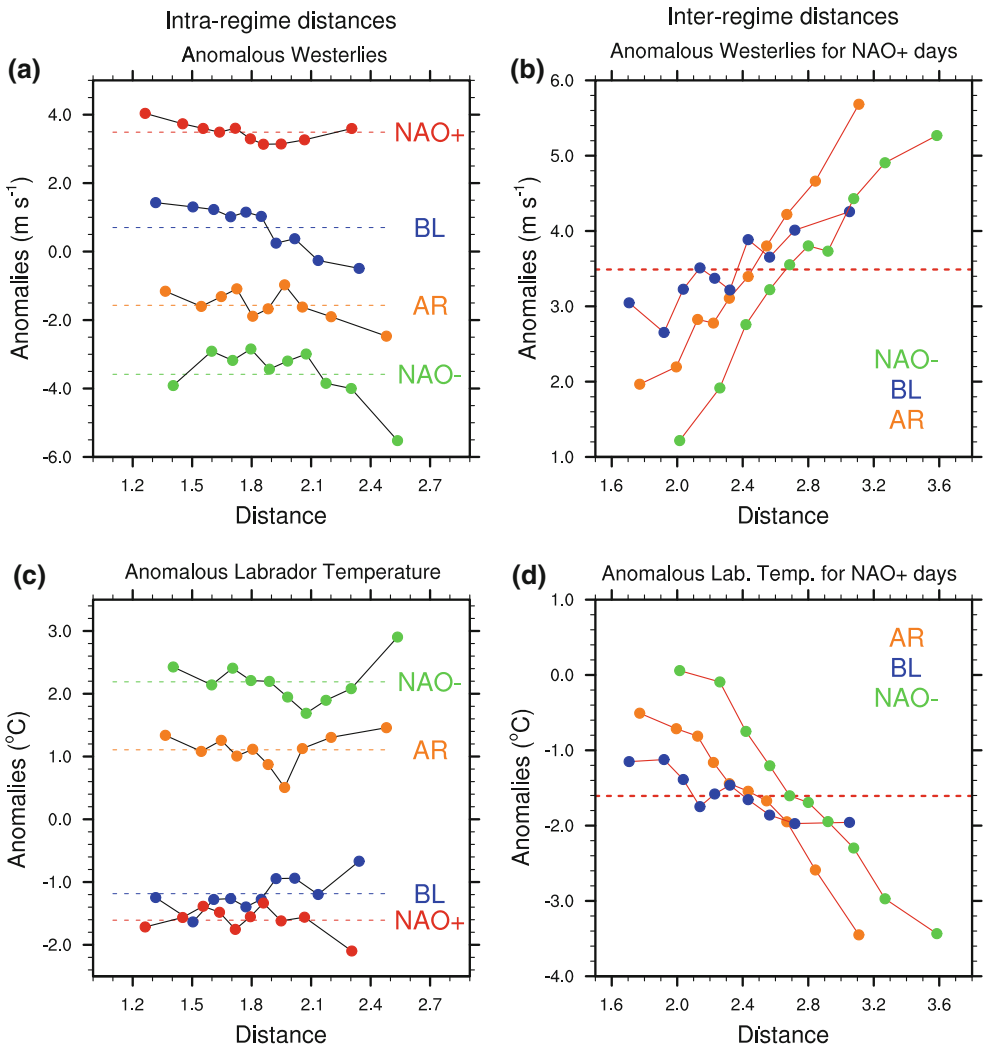

Fig. 8 **a** 10-m zonal wind box-averaged over 40°N–60°N/50°W–10°W and (**c**) 2-m temperature anomalies over the Labrador Sea (47°N–62°N/60°W–47°W) as a function of intra-regime distances binned in deciles for the four NAE-WR regimes (*colored dots*). *Dashed lines* stands for the mean anomalies as a function of the four NAE-WR (**b**) Same for inter-regime distances for NAO+ days only

variability. The table should be read by column. For instance, for NAO+ days, the Labrador T2 daily values are anti-correlated to NAO− and AR distances at −0.53 and −0.43, respectively, while no significant correlation is found between NAO+ and BL daily distances. Similar (opposite) results are found for S-BL (NAO−) while AR T2 anomalies are linked to all distances, except to its own ones. The last line of Table 2 indicates the correlation coefficient for a given regime between observed T2 daily values and T2 estimated by multiple linear regression using the four distances as predictors. Correlation values are higher in the latter case. This provides further insight that all the distances, as intrinsic parameters of decomposition in classes, are important factors to be used for the

Table 2 Linear correlations for a given NAE-WR regime between the daily distances to the four NAE-WR centroids taken separately and ERA40 2-m temperature daily anomalies over the Labrador Sea (47°N–62°N/60°W–47°W)

	NAO−	NAO+	AR	S-BL
NAO−	*0.02*	−0.53	−0.24	−0.48
NAO+	−0.48	*0.03*	0.34	0.10
AR	0.10	−0.43	*0.02*	−0.27
S-BL	0.44	−0.15	0.33	*0.02*
ALL	**0.64**	**0.62**	**0.60**	**0.58**

Intra-regime distance stands in italics. Last line (bold): correlation for a given regime between the ERA40 2-m temperature anomalous values and their reconstruction via multiple regression when the four distances are used as predictors

reconstruction of the surface ocean variables following a weather typing approach, in addition to the occurrence frequencies of the classes.

6 Atlantic summertime atmospheric circulation classes

The summer season (June–September) is now examined following the same line of reasoning as the one adopted so far for winter. Five summertime NAE-WR are obtained from *k*-means clustering (Fig. 9a–e). The first two NAE-WRs bear a strong resemblance to the negative and positive phases of the summertime NAO (Folland et al. 2009), respectively. Their respective occurrence integrated over the entire season is indeed strongly correlated (0.88 and 0.83) to the traditional summertime NAO index estimated from JJAS EOF. The teleconnection pattern is northward shifted compared to its winter counterpart (Fig. 1a, b) following the seasonal migration of the North Atlantic centers of action (Hurrell et al. 2003). Like in winter, a fair amount of asymmetry is found between the two summertime NAO weather regimes especially off Newfoundland; this yields significant impacts in terms of surface wind and T2 associated anomalies (not shown). The third regime is referred to as Atlantic Ridge because of its strong similarity to the winter pattern (Fig. 1c); the associated surface ocean anomalies are very close to the ones depicted in Fig. 2c, g, but with weaker amplitude. The forth regime is named Icelandic Low (IL). Interestingly, IL strongly projects on the wintertime NAO+ spatial structure but is not correlated with the summertime NAO index (the correlation value is equal to −0.31). The last regime is termed WES standing for Scandinavia–Western Europe. This regime is associated with bursts of thunderstorms over Western Europe because of southerly advection of warm air masses from Maghreb (van Delden 2001). Similarly to wintertime, summer NAO frequencies of occurrence exhibit decadal fluctuations (not shown) superimposed on considerable interannual variability as also shown in Folland et al. (2009).

Five NAE-WR are kept in the present manuscript as opposed to Cassou et al. (2005) or Guemas et al. (2010) papers using four summertime patterns defined from June to August days. We found that including September in the clustering step alters the robustness of the partition into four regimes while $k = 5$ appears to be more stable (i.e. not dependent on the sampling size/period as tested in Straus et al. 2007). Note though that $k = 5$ fails the Michelangeli et al. (1995) red noise test from which none of *k*-values are significant. Five summertime tropical wind classes are also chosen for the same reasons (Fig. 9f–j). T-WC1 is characterized by reinforced trade winds in both hemispheres. T-WC2 exhibits a strong reduction of the NE trades while signs for anomalous anticyclonic circulation in the southern hemisphere are visible. Both regimes are overly dominated by decadal variability that is even more pronounced in summer than in winter (not shown) and is clearly associated with the AMO. T-WC3 and T-WC4 exhibits a fair amount of linearity; T-WC3 (T-WC4) is characterized by enhanced (reduced) NE trades while SE trades are diminished (reinforced). Their occurrence is correlated with the cross-equator SST gradient mode. Maximum occurrence for T-WC3 (T-WC4) is found in the 1970s and 1980s (1960s and 1990s) with strong interannual pulses over the entire period. T-WC5 is characterized by a pronounced equatorial surface wind divergence and its appearance is dominated by an upward trend; T-WC5 is mostly active from the mid-1980s onward.

To what extent is decomposition in weather classes relevant for reconstructing the low-frequency changes of the surface ocean variables in summer? Following the exact same approach as for the winter season (Sect. 4), the frequencies of occurrence of summer NAE-WR or/and T-WC are entered as predictors in least-square multiple linear regressions. We only show correlation skills (Fig. 10) when both NAE-WR and T-WC are used in the regression equation, their relative contribution over the entire Atlantic basin being even more regionally additive than in winter. Indeed, there is no sign of significant tropical-extratropical connection in summertime as also confirmed by the table of contingency between summer NAE-WR and T-WC (same as Table 1, not shown). This result is consistent with the accepted notion that tropical-extratropical connections are mostly active during wintertime because of their dependence upon the climatological background of the atmospheric dynamics. Figure 10a shows that the U10 reconstruction based on summertime occurrence frequencies is as skillful as in winter with values ranging from 0.5 to 0.9. Marginal low correlation areas are found south of Newfoundland and around the Canaries Archipelago while performances are improved within the tropical band. Good correlation values are also obtained for summertime V10 except over a broad Northwest Atlantic basin (Fig. 10b). Reasonable skill is found for T2 although slightly lower north of 40°N compared to wintertime. Because no significant value for *k* is suggested by statistical tests, we verified that the reconstruction skill is not changed from *k* values ranging from 4 to 7 (not shown), the maximum reconstruction skill being obtained for $k = 5$.

Summertime trends are weak in U10, V10 and T2 in the Atlantic. Their reconstruction is however quite correct like in winter, except for T2 for which the warming is underestimated by 60–70% over the entire basin (not shown). This suggests again that part of the thermal trends of the ocean surface is not directly linked to the overlying atmospheric dynamics and is controlled by other factors.

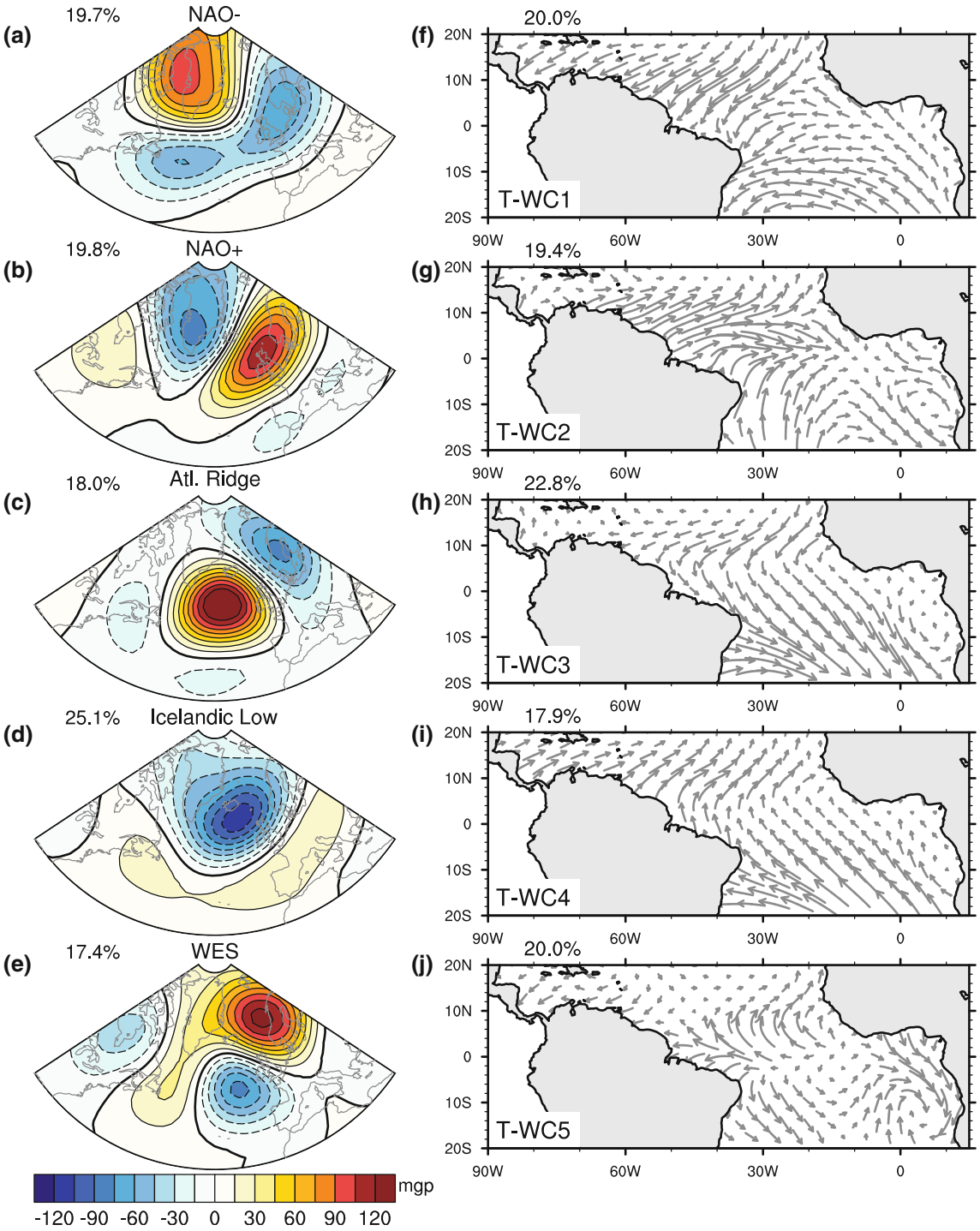

Fig. 9 Same as Figs. 1 and 2 but for summertime (over 1958–2002 from 1 June to 30 September)

Because of the shallowness of the summertime ocean mixed layer, radiative forcing from GHG increased concentration is expected to be even more efficient to produce a superficial warming, being subsequently entrained at depth in winter when the seasonal deepening of the mixed layer occurs.

7 Summary and discussion

In this study, we have investigated the link between the observed variability of the surface ocean variables over the Atlantic and the overlying atmosphere decomposed in classes of daily large-scale atmospheric circulation. The

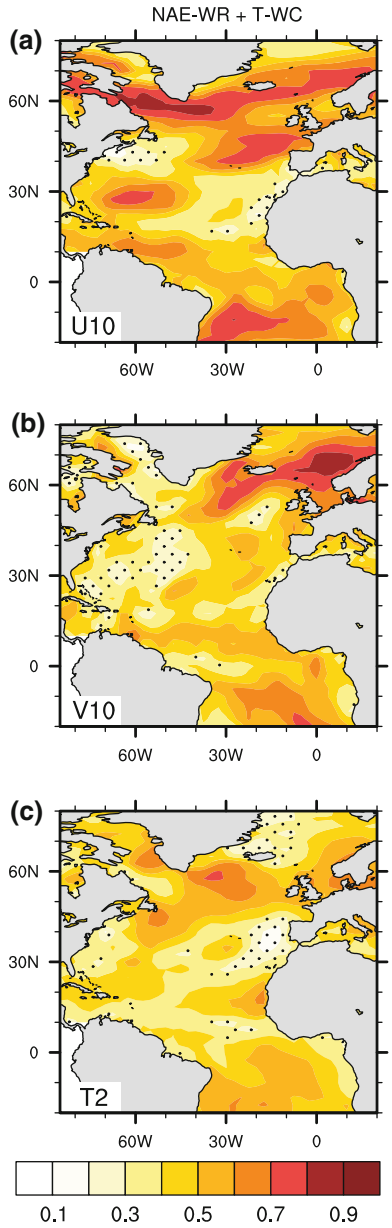

Fig. 10 Same as Fig 5c, f, i but for summertime

k-means partition method has been applied first to daily Z500 anomaly maps from ERA40 over a broad NAE domain from 1958 to 2002. The clustering analysis provides a set or representative patterns that correspond in the extratropics to well-known NAE weather regimes accounting for the barotropic dynamics. The *k*-means algorithm has been then applied to daily 30-day low pass filtered UV1000 anomaly maps in the tropical band; it yields to wind classes representing the alteration of the trades. Even if they are temporally connected, tropics and extratropics are treated separately because of the intrinsically different nature of their atmospheric dynamics and associated surface ocean changes. The primary aim of Part I of our study was to evaluate how relevant the decomposition of the atmospheric variability into daily midlatitude weather regimes and/or tropical classes is to capture the low-frequency changes of the surface ocean variables. This step appears to be a prerequisite before thinking about applying the weather-typing perspective to ocean dynamical downscaling as further described in Part II of this work (Minvielle et al. 2010).

Winter and summer seasons are analyzed individually. Consistent with many studies, four wintertime NAE-WR are detected: the NAO− and NAO+ regimes, the AR regime and the S-BL regime. Derived anomaly composites for UV10 and T2 exhibit significant ocean basin wide structures for each regime and are indicative of strong links between anomalous large-scale atmospheric circulations and surface ocean fields in the North Atlantic. Because there is no assumption for linearity in clustering, spatial asymmetries are found in surface fields, especially for UV10, between the two NAO regimes. We argue at that point that such asymmetries could be of great importance when surface fields associated with the NAO are used as forcing in OGCMs, and can therefore be considered as an added-value of the weather-typing approach. We indeed expect the ocean to be very sensitive to asymmetries because of nonlinear processes present in key regions for its dynamics and variability (deep convection, vertical entrainment in the mixed layer, etc.).

In the tropical Atlantic, four wintertime wind classes are obtained corresponding to either basin wide alteration of the trade wind intensity in both hemispheres or to changes in the cross-equatorial flow. Their occurrence is largely dominated by decadal changes. Clustering is able to capture the late 1960s-early 1970s climate shift that underwent a number of significant variables of the atmospheric circulation. We find that the prevalence of this shift renders difficult to firmly link the individual T-WC occurrence to any specific SST patterns at interannual to quasi-decadal timescale. Besides, there is clear evidence of tropical-extratropical connection as diagnosed by the table of contingency between wintertime NAE-WR and T-WC mean occurrences. Additional investigations are thus considerably needed to clarify the relative roles of the different actors in the tropical Atlantic climate. Those investigations might likely imply some refinement at the very first stage of the UV1000 clustering and our study should be treated as a first attempt.

Five NAE-WR and T-WC are obtained for summer. Similar conclusions hold for that season: the importance of the spatial asymmetry found in derived surface variables anomalies from NAE regimes, and the dominant signature of the late-1960 s climate shift in the tropics, although the relationship between T-WC summertime occurrence and local summertime SST appears to be less blurred in

agreement with the seasonal estimate of the trade winds potential predictability given in Sutton et al. (2000). By contrast to winter, there is no evidence of tropical-extra-tropical connection in summer. Note that such an assertion is drawn from mean statistics that does not preclude some possible connections for some specific events or years as documented in literature (Cassou et al. 2005).

The relevance of the decomposition in NAE-WR and/or T-WC to derive basin-wide low-frequency changes of surface ocean variables has been demonstrated. Our results highlight that both the type of regime, i.e. its occurrence, and its relative strength with respect to the others regimes, i.e. its inter-regime distances, are valuable predictors for surface ocean variables. Best results are obtained when NAE and tropical clusters are combined because of closely interwoven dynamics. Figure 11 provides a schematic means of visualizing and summarizing the latter findings. Let us consider two regimes A and B whose centroid position is represented by a thick black cross in a reduced phase space span by two arbitrary dimensions. For the sake of simplicity, let us consider T2 as the surface variable to be reconstructed from regime-derived information. Each individual dot represents a day and its color stands for the sign and the magnitude of the T2 anomalies of that day; reddish (bluish) color stands for warm (cold) T2 anomalies and the redder (bluer), the warmer (colder). For a given region, we found that:

- the attribution to a given regime controls to a great extent the sign of the surface anomalies. Schematically warm (cold) days mostly belong to regime B (regime A).
- the strength of the regime with respect to its own centroid does not explain the magnitude of the surface anomalies. Schematically, let us take two warm days from Regime B (colored stars) located at the same distance to Regime B centroid, i.e. along the same solid circle. Intra-regime distances (pink arrows) are thus identical and cannot explain the fact than 1 day is warmer than the other.
- the strength of the regime with respect to the centroids of the other regimes do explain the magnitude of the surface anomalies. Difference in inter-regime distances (blue arrows) explains why one day is warmer than the other despite the fact that both belong to the same regime.

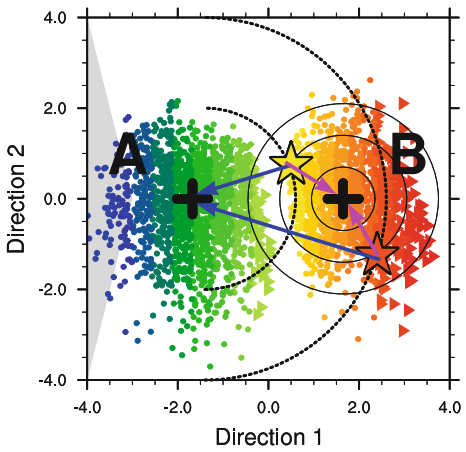

Fig. 11 Schematic summary diagram of inter- and intra-distance influence on surface variable anomalies

The latter conclusions are valid to a large extent over most of the NAE domain and have been validated for two key regions for ocean dynamics: the Labrador Sea (deep water formation, etc.) and a latitudinal band along the mean westerlies whose strength controls a large part of the variability of the ocean gyre dynamics.

We have verified that inter-regime distances used as predictors in least-square multiple linear regressions to reconstruct surface ocean variables yields significant correlations between observed and reconstructed daily values. We have shown that intra-regime distances do not play a significant role, a conclusion that is not straightforward. For instance, it is definitely not the case for precipitation or temperature over land and we believe that it might be related to the large size of the domain considered in our case. Note that this assertion may not hold for climate change studies. We have verified that using the combined frequencies of occurrence of NAE-WR + T-WC as predictors enables us to reproduce very reasonably the inter-annual variability of U10, V10 and T2 over the entire Atlantic basin We also have suggested that part of sub-surface water properties in the Labrador Sea where deep convection occurs can be linked to NAE-WR excitation providing further assurance in the relevance of our weather-typing approach to investigate the 3-dimensional ocean dynamics. Finally, we have provided some evidence that a large fraction of the low-frequency trends in the Atlantic observed at the surface over the last 50 years can be traced back to changes in occurrence of tropical and extratropical weather classes and can be interpreted as the time integration of higher frequency fluctuations in the atmosphere. One direct interpretation of these findings is that the weather-typing statistical scheme to be proposed in Part II of this study to reconstruct the sea surface variables should include both the frequency of occurrence and the distance as input parameters.

The last element to be considered is the tropical-extra-tropical connection that must be respected in the future reconstruction scheme. Complementary analyses (not shown) suggest that, in addition to the fact that T-WC and NAE-WR occurrences are not independent, the NAE-WR strength is modulated by the tropical states. This is

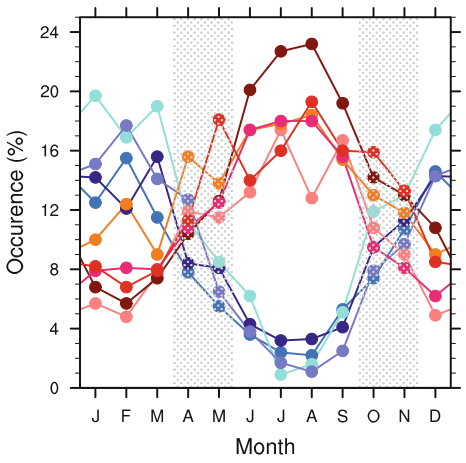

Fig. 12 Percentage of days attributed to each regime among the nine winter + summer NAE centroids as function of calendar month. *Bluish* (*reddish*) color stands for the winter (summer) NAE-WR

particularly true for the NAO regimes and to a lesser extent for S-BL and AR. This arises through atmospheric teleconnections sustained by a Rossby wave-like disturbances propagating northward from the western part of the tropical Atlantic (Terray and Cassou 2002). The triangles in Fig 11 schematically symbolize the modulator role of the tropics in the spatial distribution of the NAE regime. The gray large triangle on the left side represents the direction of the tropical forcing in the reduced phase space. Let regime B (A) be the NAO− (NAO+) regime and T2 over the Labrador Sea. Individual colored triangles stand for days where T-WC2 is excited; those are sparse in regime A and directed towards regime B (light green) whereas they are numerous in regime B but located on the opposite side "running away" from A and leading to stronger surface anomalies (dark red). Recall that Fig 11 is a simplified or noise-free view for heuristic purposes; the true nature of the dynamics is much more blurred.

We concentrated our analyses on winter and summer seasons. How about spring and fall? Because we are interested in low-frequency changes of the ocean (interannual variability and trends), a continuous full-year daily forcing must be provided to OGCMs. When clustering is applied to the two intermediate seasons, obtained classifications are not stable at all. This could be consistent with the idea that intermediate seasons may not be characterized by specific atmospheric circulation patterns; those should rather be treated as transition periods between summertime and wintertime distinct dynamics, the timing of shift between the two varying with years leading to unstable statistics. This interpretation is reinforced by Fig. 12 for NAE-WR. All anomalous Z500 calendar days are attributed to one of the 9 NAE-WRs formed by the 4 + 5 winter + summer weather regimes obtained from seasonal clustering as described above. We count the number of days attributed to each regime considering the 12 calendar months separately. As anticipated by construction, D/J/F/M days mostly fit in one of the four winter centroids while J/J/A/S days rarely picked them. Interestingly, occurrences are less segregated in winter than in summer. We found that weak Z500 anomalies days in winter are artificially closer to summertime regimes because the Euclidian distance mostly accounts for the amplitude of the anomalies. Such a behavior disappears if the pattern correlation distance is used instead or if non-persistent regimes (episodes lasting less than four consecutive days and interpreted as transition days) are removed from the statistics. Winter and summer regimes are about equally excited in spring and autumn months, except for May where summer regimes slightly dominate. This suggests that both summertime and wintertime dynamics statistically occur during the intermediate seasons making difficult the extraction of specific representative patterns through clustering for April, October, November and May to a lesser extent. This result highlights the need for considerable further work to understand the nature and physics of the intermediate season atmospheric dynamics.

All together, this paper, or Part I, tends to show that a weather-typing approach could be promising for ocean dynamical downscaling application. As a prerequisite, we provided some evidence that weather regimes/classes occurrence and strength are indeed discriminatory for the variability of the surface ocean fields used as forcing in OGCMs. Consequently, we have elaborated a statistical method based on the conclusions described here to estimate the ocean forcing variables from observed estimates, and have used a high resolution ocean model to test its relevance. The reader is invited to report to Minvielle et al. (2009) or Part II of the present study, for a complete description of the statistico-dynamical algorithm and its performance.

Acknowledgments We thank Julien Najac and Julien Boé for stimulating discussions. We are grateful to Soline Bielli and Didier Swingedouw for their help in improving the manuscript The figures were produced with the NCL software developed at NCAR. This work was supported by CERFACS, CNRS, Mercator-Ocean via the DESAGO project and by the European Community via the sixth framework ENSEMBLES project under Contract GOCE-CT-2003-505539.

References

Bindoff NL et al (2007) Observations: Oceanic Climate Change and Sea Level. In: Salomon S et al (eds) Climate Change 2007: the Physical Science Basis, vol 385. Cambridge University Press, Cambridge, p 432

Baehr J, Haak H, Alderson S, Cunningham SA, Jungclaus JH, Marotzke J (2007) Timely detection of changes in the meridional overturning circulation at 26°N in the Atlantic. J Clim 20:5827–5841

Baines PG, Folland CK (2007) Evidence for a Rapid Global Climate Shift across the Late 1960s. J Clim 20:2721–2744. doi:10.1175/JCL14177.1

Balmaseda MA, Vidard A, Anderson DLT (2008) The ECMWF ocean analysis system: ORA-S3. Mon Wea Rev 136:3018–3034

Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. Mon Wea Rev 115:1083–1126

Biastoch A, Böning CW, Getzlaff J, Molines JM, Madec G (2008) Causes of interannual-decadal variability in the meridional overturning circulation of the midlatitude North Atlantic Ocean. J Clim 21:6599–6615

Boé J, Terray L (2008) A weather type approach to analyzing winter precipitation in France: twentieth century trends and the role of anthropogenic forcing. J Clim 21:3118–3133

Boé J, Terray L, Habets F, Martin E (2006) A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling. J Geophys Res 111:D23106

Bryden HL, Longworth HR, Cunningham SA (2005) Slowing of the Atlantic meridional overturning circulation at 25°N. Nature 438:655–657. doi:10.1038/nature04385

Carton JA, Huang B (1994) Warm events in the tropical Atlantic. J Phys Oceanogr 24:888–903

Cassou C, Terray L, Hurrell JW, Deser C (2004) North Atlantic winter climate regimes: spatial asymmetry, stationarity with time and oceanic forcing. J Clim 17:1055–1068

Cassou C, Terray L, Phillips AS (2005) Tropical Atlantic influence on European heat waves. J Clim 18:2805–2811

Cayan DR (1992) Latent and sensible heat flux anomalies over the northern oceans: driving the sea surface temperature. J Phys Oceanogr 22:859–881

Cheng XH, Wallace JM (1993) Cluster analysis of the northern hemisphere winter 500-hPa height field: spatial patterns. J Atmos Sci 50:2674–2696

Christiansen B (2007) Atmospheric circulation regimes: can cluster analysis provide the number? J Clim 20:2229–2250

Cunningham SA, Coauthors (2007) Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. Science 317:935–938

Daget N, Weaver AT, Balmaseda MA (2010) Ensemble estimation of background-error variances in a three dimensional variational data assimilation system for the global ocean. Q J R Meteorol Soc (in press)

Dai AG (2006) Precipitation characteristics in eighteen coupled climate models. J Clim 19:4605–4630

Delden (2001) The synoptic setting for thunderstorms in western Europe. Atmos Res 56:89–110

Deser C, Phillips AS (2009) Atmospheric circulation trends, 1950–2000: the relative role of sea surface temperature forcing and direct atmospheric radiative forcing. J Clim 22:396–413

Deser C, Timlin MS (1997) Atmosphere–ocean interaction on weekly timescales in the North Atlantic and Pacific. J Climate 10:393–408

Dickson R, Lazier J, Meincke J, Rhines P, Swift J (1996) Long-term coordinated changes in the convective activity of the North Atlantic. Prog Oceanogr 38:241–295

Fereday DR, Knight JR, Scaife AA, Folland CK (2008) Cluster analysis of the North Atlantic-European circulation types and links with tropical Pacific Sea Surface Temperatures. J Clim 21:3687–3703

Folland CK, Knight J, Linderholm HW, Fereday D, Ineson S, Hurrell JW (2009) The Summer North Atlantic Oscillation: Past, Present, and Future. J Clim 22:1082

Frankignoul C, Kestenare E, Botzet M, Carril AF, Drange H, Pardaens A, Terray L, Sutton R (2004) An intercomparison between the surface heat flux feedback in five coupled models, COADS and the NCEP reanalysis. Clim Dyn 22:373–388

Franzke C, Feldstein SB (2005) The continuum and dynamics of Northern hemisphere teleconnection patterns. J Atmos Sci 62:3250–3267

Ghil M, Roberston AW (2002) "Waves" vs "Particles" in the atmospheric phase space: a pathway to long-range forecasting? Proc Natl Acad Sci USA 99:2493–2500

Giannini A, Saravanan R, Chang P (2003) Oceanic forcing of Sahel rainfall on interannual to interdecadal time scales. Science 302:1027–1030

Guemas V, Salas-Mélia D, Kageyama M, Giordani H, Voldoire A, Sanchez-Gomez E (2010) Summer interactions between weather regimes and surface ocean in the North Atlantic region. Clim Dyn. 34:527–546. doi:10.1007/s00382-008-0491-6

Hannachi A (2010) On the origin of planetary scale extratropical winter circulation regimes. J Atmos Sci (in press)

Heyen H, Dippner JW (1998) Salinity variability in the German Bight in relation to climate variability. Tellus 50A:545–556

Heyen H, Zorita E, von Storch H (1996) Statistical downscaling of monthly mean North Atlantic air-pressure to sea level anomalies in the Baltic Sea. Tellus 48A:312–323

Hilmer M, Jung T (2000) Evidence for a recent change in the link between the North Atlantic Oscillation and Arctic sea ice export. Geophys Res Lett 27:889–992

Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperature and precipitation. Science 26:676–679

Hurrell JW, Kushnir Y, Otterson G, Visbeck M (2003) An overview of the North Atlantic Oscillation. The North Atlantic Oscillation: Climate significance and environmental impact. Geophys Monogr 134(1):35

Johnson NC, Feldstein SB, Tremblay B (2008) The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organized maps. J Clim 21:6354–6371

Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycles in observed climate. Geophys Res Lett 32:L20708. doi:10.1029/2005GL024233

Knight JR, Folland CK, Scaife AA (2006) Climate impacts of the Atlantic multidecadal oscillation. Geophys Res Lett 33:L17706. doi:10.1029/2006GL026242

Levitus S, Antonov J, Boyer T (2005) Warming of the world ocean, 1955–2003. Geophys Res Lett 32:L02604. doi:10.1029/2004GL021592

Martin MJ, Hines A, Bell MJ (2007) Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. Q J R Meteorol Soc 133:981–995

Michelangeli P, Vautard R, Legras B (1995) Weather regimes: recurrence and quasi stationarity. J Atmos Sci 52:1237–1256

Minvielle M, Cassou C, Terray L, Bourdallé-Badie R, Najac J (2010) A statistical-dynamical scheme for reconstructing ocean forcing in the Atlantic. Part II: Methodology, validation and application to high resolution ocean models. Clim Dyn. doi: 10.1007/s00382-010-0761-y

Najac J, Boé J, Terray L (2008) A multi-model ensemble approach for assessment of climate change impact on surface winds in France. Clim Dyn. doi:10.1007/s00382-008-0440-4

Peña M, Cai M, Kalnay E (2004) Life span of subseasonal coupled anomalies. J Clim 17:1587–1603

Philipp A et al (2007) Long term variability of daily North Atlantic-Europe pressure patterns since 1850 classified by simulated annealing clustering. J Clim 20:4065–4095

Randall DA et al (2007) Climate models and their evaluation. In: Salomon S et al (eds) Climate Change 2007: the Physical

Science Basis. Cambridge University Press, Cambridge, pp 589–662

Roberts MJ, Banks H, Gedney N, Gregory J, Hill R, Mullerworth S, Pardaens A, Rickard G, Thorpe R, Wood R (2004) Impacts of an eddy permitting ocean resolution on control and climate change simulations with a global coupled GCM. J Clim 17:3–20

Ruiz-Barradas A, Carton JA, Nigam S (2000) Structure of interannual-to-decadal climate variability in the tropical Atlantic sector. J Clim 13:3285–3297

Slonosky VC, Yiou P (2001) The North Atlantic Oscillation and its relationship with near surface temperature. Geophys Res Lett 28:807–810

Straus DM, Corti S, Molteni F (2007) Circulation regimes: chaotic variability versus SST-Forced predictability. J Clim 20:2251–2272

Sutton RT, Jewson SP, Rowell DP (2000) The elements of climate variability in the tropical Atlantic region. J Clim 13:3261–3284

Sutton RT, Norton WA, Jewson SP (2001) The North Atlantic Oscillation—What role for the ocean? Atmos Sci Lett 1:89–100

Terray L, Cassou C (2002) Tropical Atlantic sea surface temperature forcing the quasi-decadal climate variability over the North Atlantic-Europe region. J Clim 15:3170–3187

Tyrlis E, Hoskins BJ (2008) Aspects of a Northern Hemisphere atmospheric blocking climatology. J Atmos Sci 65:1638–1652

Uppala SM, Coauthors (2005) The ERA40 re-analysis. Q J R Meteor Soc 131:2961–3012

Våge K, Pickart RS, Thierry V, Reverdin G, Lee CM, Petrie B, Agnew TA, Wong A, Ribergaard MH (2009) Surprising return of deep convection to the subpolar North Atlantic in winter 2007-08. Nature Geosci., in press

Vautard R (1990) Multiple weather regimes over the North Atlantic: analysis of precursors and processors. Mon Wea Rev 118:2056–2081

von Storch H, Zwiers FW (1999) Statistical Analysis in climate research. Cambridge University Press, Cambridge, 484 p

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. Data Distribution Centre of the Intergovernmental Panel on Climate Change. Available at http://ipcc-ddc.cru.uea.ac.uk/guidelines

Woollings T, Hannachi A, Hoskins BJ (2010) A regime view of the North Atlantic Oscillation and its response to anthropogenic forcing. J Clim (in press)

Yashayaev I (2007) Hydrographic changes in the Labrador Sea, 1960–2005. Prog Oceanogr 73:242–276. doi:10.1016/j.pocean.2007.04.015

Yashayaev I, Loder JW (2009) Enhanced production of Labrador Sea Water in 2008 Geophys. Res Lett 36:L010606. doi:10.1029/2008GL036162

Yiou P, Vautard R, Naveau P, Cassou C (2007) Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe. Geophys Res Lett 34:L21808. doi:10.1029/2007GL031981

Zebiak SE (1993) Air-sea interaction in the equatorial Atlantic region. J Clim 6:1567–1586