Publisher's PDF, also known as Version of record

Link to published version (if available):
[10.5194/bg-10-1799-2013](10.5194/bg-10-1799-2013)

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research
General rights

Biogeosciences

Detecting an external influence on recent changes in oceanic oxygen using an optimal fingerprinting method

O. D. Andrews[1], **N. L. Bindoff**[2,3,4], **P. R. Halloran**[5], **T. Ilyina**[6], **and C. Le Quéré**[1]

[1]Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK
[2]IMAS, University of Tasmania, Hobart, Australia
[3]CSIRO Marine and Atmospheric Research, Hobart, Australia
[4]ACE CRC, University of Tasmania, Hobart, Australia
[5]Met Office Hadley Centre, FitzRoy Road, Exeter, UK
[6]Max Planck Institute for Meteorology, Hamburg 20146, Germany

Correspondence to: O. D. Andrews (o.andrews@uea.ac.uk)

Abstract. Ocean deoxygenation has been observed in all major ocean basins over the past 50 yr. Although this signal is largely consistent with oxygen changes expected from anthropogenic climate change, the contribution of external forcing to recent deoxygenation trends relative to natural internal variability is yet to be established. Here we conduct a formal optimal fingerprinting analysis to investigate if external forcing has had a detectable influence on observed dissolved oxygen concentration ($[O_2]$) changes between ~ 1970 and ~ 1992 using simulations from two Earth System Models (MPI-ESM-LR and HadGEM2-ES). We detect a response to external forcing at a 90 % confidence level and find that observed $[O_2]$ changes are inconsistent with internal variability as simulated by models. This result is robust in the global ocean for depth-averaged (1-D) zonal mean patterns of $[O_2]$ change in both models. Further analysis with the MPI-ESM-LR model shows similar positive detection results for depth-resolved (2-D) zonal mean $[O_2]$ changes globally and for the Pacific Ocean individually. Observed oxygen changes in the Atlantic Ocean are indistinguishable from natural internal variability. Simulations from both models consistently underestimate the amplitude of historical $[O_2]$ changes in response to external forcing, suggesting that model projections for future ocean deoxygenation may also be underestimated.

1 Introduction

The oceanic oxygen inventory is coupled to the climate system via a number of physical and biogeochemical processes making oxygen a useful tracer for detecting changes in the state of the earth system (Joos et al., 2003; Brennan et al., 2008). In particular, changes in temperature, ocean circulation, biological production and respiration, expected in response to global climate change, all exert a major control on dissolved oxygen concentration ($[O_2]$). Open ocean deoxygenation has been recorded in nearly all ocean basins during the second half of the 20th century (Helm et al., 2011) with more acute $[O_2]$ decreases in the coastal ocean (Gilbert et al., 2010). Oxygen time series measurements also show historical expansion and intensification of established oxygen minimum zones (OMZs) in the eastern tropical Atlantic and equatorial Pacific since 1960 (Stramma et al., 2008) along with long-term $[O_2]$ decreases in the subarctic North Pacific between 1956 and 2006 (Whitney et al., 2007).

Historical deoxygenation has been associated with global climate change, chiefly due to enhanced ocean stratification in a warming climate which increases the ventilation age of downwelling water masses, and augmented by the reduced solubility of dissolved oxygen at higher temperatures (Keeling et al., 2010). These findings are supported by prognostic simulations from a suite of Atmosphere-Ocean General Circulation Models (AOGCMs), which show ventilation-driven reductions in global mean $[O_2]$ between 3 and 12 μmol kg^{-1}

by 2100 (Sarmiento et al., 1998; Bopp et al., 2002; Matear and Hirst, 2003; Frölicher et al., 2009). Global $[O_2]$ decreases have also been simulated over longer timescales in response to protracted ocean warming. For example, model simulations using Intermediate Complexity Earth System Models (EMICs) produce a tripling in the volume of suboxic ($[O_2] < 5\,\mu mol\,kg^{-1}$) waters by 2500 (Schmittner et al., 2008). Ocean deoxygenation and expansion of the OMZs has also been projected to persist on millennial timescales for EMIC simulations with high greenhouse gas emissions or high climate sensitivity (Shaffer et al., 2009).

All global models simulate that the most significant $[O_2]$ decreases occur at mid- to high latitudes, consistent with observations and caused by decreasing ventilation of the oceanic interior in response to enhanced upper ocean stratification. However, models generally do not reproduce the deoxygenation observed at low latitudes, which leads to considerable uncertainties for future projections. The predictive skill of coarse resolution ocean models at low latitudes is limited by the need to resolve zonal currents thought to be important in supplying oxygen into the tropical OMZs (Keeling et al., 2010; Brandt et al., 2010; Stramma et al., 2010). Climate change integrations of coupled AOGCMs often simulate $[O_2]$ increases in the tropical thermocline associated with reductions in the volume of O_2-depleted waters at low latitudes, contrary to recent trends (Matear and Hirst, 2003; Gnanadesikan et al., 2012). Gnanadesikan et al. (2007) propose that this oxygenation is driven by reduced exchange between poorly ventilated deep water and the tropical thermocline in a more stratified ocean. However, this process has been posited to be a numerical artifact tied to unrealistically high rates of diapycnal mixing in coarse resolution ocean models (Keeling et al., 2010). A recent sensitivity study by Duteil and Oschlies (2011) confirms the importance of vertical diffusivity and its parameterisation on the evolution of suboxic areas in global models. Simulated oxygenation of suboxic zones has also been attributed to an elevated supply of oxygen from isopycnal diffusion in one study using the GFDL ESM2.1 (Gnanadesikan et al., 2012). Up to now, biogeochemical models can reproduce observed low-latitude deoxygenation and expansion of hypoxic and suboxic waters in the tropical OMZs only when they take into account poorly constrained pCO$_2$-driven changes to the C:N ratio in primary production (Oschlies et al., 2008; Tagliabue et al., 2011) or the alteration to $CaCO_3$ mineral ballasting in response to ocean acidification (Hofmann and Schellenhuber, 2009).

The implications of a climate change signal in $[O_2]$ changes are profound, owing to the deleterious impact of reduced oxygen concentrations and expanding hypoxic zones on marine ecosystems and fisheries (e.g., Doney et al., 2012; Stramma et al., 2012b), and the sensitivity of biolimiting marine nutrient cycles such as nitrogen and phosphorous to oceanic redox conditions (Keeling et al., 2010). However, attribution of secular $[O_2]$ changes to global climate change is confounded by natural interannual to decadal variability

driven by the leading climate modes which act to mask any climate-driven signal (e.g., the North Atlantic Oscillation, Frölicher et al., 2009). To this end, modelling studies have demonstrated that internal variability associated with the major climate modes in the North Pacific drive propagating O_2 anomalies that are large enough to frequently preclude unequivocal detection of anthropogenic trends in dissolved oxygen on decadal timescales and regional scales (Deutsch et al., 2006; Ito and Deutsch, 2010). However, Deutsch et al. (2005, 2006) highlight that in the North Pacific strong historical $[O_2]$ decreases within the lower ventilated thermocline are reproduced by models and are likely to be driven by long-term changes in ocean ventilation and circulation, but a formal detection study is required to separate external causes from internal variability.

We use an optimal fingerprinting methodology to objectively quantify the detectability of externally forced changes in oceanic $[O_2]$ relative to natural internal periodicity. Numerous detection studies have been conducted investigating the impact of external forcing on oceanic variables such as ocean temperature (e.g., Barnett et al., 2005; Pierce et al., 2006; Gleckler et al., 2012) and salinity (Stott et al., 2008; Terray et al., 2012), but these techniques have not yet been used for biogeochemical tracers like oxygen. $[O_2]$ changes, therefore, provide a novel means for detecting and attributing anthropogenic climate change in the oceans and for the evaluation of model performance (Hegerl et al., 2006). As such, this study aims to conduct a formal optimal fingerprinting analysis to investigate the "signal in noise" problem of historical changes in oceanic oxygen using $[O_2]$ data and simulations from two Earth System Models (ESMs) participating in Phase 5 of the Coupled Model Intercomparison Project (CMIP5).

2   Methods

2.1   Observations and CMIP5 models

We use a recent collation of global World Ocean Circulation Experiment (WOCE) O_2 data compared with earlier ocean oxygen profiles (Helm et al., 2011) to investigate the cause of the observed historical changes in global marine oxygen distribution. This compilation maps all historical $[O_2]$ data (1940–1988; mean year ~ 1970) onto the locations of WOCE profiles to facilitate comparison with WOCE measurements (~ 1989–2000; mean year ~ 1992) using an optimal-interpolation technique that adapts the length scales to suit the density and distribution of oxygen data (Helm et al., 2010). The $[O_2]$ data is binned into regular $1° \times 1°$ grid cells with 101 fixed pressure levels, excluding shallow (< 1000 m) profiles and observations above 100 m where seasonal variability dominates the oxygen signal (Garcia et al., 2005). Noise variance calculated using the difference between neighbouring data points is used to provide an a priori

estimate of natural variability in [O_2] data (Helm et al., 2010, 2011; Bindoff and Wunsch, 1992). This technique accounts for mesoscale processes and to some extent longer period internal variability such as the dominant climate modes.

This dataset is used in an optimal detection analysis along with the biogeochemical output from two state-of-the-art ESMs (MPI-ESM-LR and HadGEM2-ES) participating in the CMIP5 experiments (Taylor et al., 2009, 2012). [O_2] fields from MPI-ESM-LR and HadGEM2-ES have been selected here because they display a higher level of realism in simulating both climatological [O_2] distribution and historical oxygen changes compared to the other available CMIP5 ESMs. Sufficient qualitative agreement between modelled and observed [O_2] distribution allows for direct quantitative comparison using an optimal fingerprinting algorithm without any requirement for a prior transformation of the model's variables (e.g., Banks and Bindoff, 2003).

The Max Plank Institute for Meteorology Earth System Model used here (MPI-ESM-LR; Giorgetta et al., 2012) is a low resolution model version that includes the ECHAM6 atmospheric GCM with T63 horizontal resolution (1.875°) and 47 vertical layers (Stevens et al., 2013; developed from ECHAM5 (Roeckner et al., 2006) with modifications in the shortwave radiative transfer and representation of the middle atmosphere). It is coupled to the MPIOM physical ocean model (Jungclaus et al., 2013; Marsland et al., 2003) which includes a thermodynamic-dynamic sea ice component (Notz et al., 2013) and ocean biogeochemistry model (HAMOCC5; Ilyina et al., 2013; Maier-Reimer, 1993; Maier-Reimer et al., 2005) implemented on a curvilinear bipolar orthogonal grid with a nominal horizontal resolution of 1.5° and 40 z-levels which increase in thickness with depth. The HAMOCC5 ocean biogeochemistry component includes an extended Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) type ecosystem model (Six and Maier-Reimer, 1996) along with over 30 prognostic variables in the water column and the upper sediments including co-limiting nutrients nitrate, phosphate and iron, as well as DIC and oxygen. The ocean and atmosphere components are coupled daily without flux corrections.

The Hadley Centre Global Environment Model Version 2 Earth System Model (HadGEM2-ES, Collins et al., 2011; HadGEM2 Development Team, 2011) is also a coupled AOGCM and has an atmospheric resolution of N96 (1.875° × 1.25°) with 38 vertical levels. The ocean model has a 1° × 1° horizontal resolution (increasing smoothly to 1/3° at the equator) with 40 unevenly spaced depth levels. The ocean biogeochemistry component of this model is Diat-HadOCC, which is an NPZD model with two phytoplankton functional types (developed from the HadOCC model; Palmer and Totterdell, 2001), and includes the nutrient cycles of nitrogen, silica and iron along with prognostic tracers such as dissolved oxygen.

We use output from the ∼ 1000 yr control (piControl) experiments of MPI-ESM-LR and HadGEM2-ES, which pre-

Table 1. Summary of CMIP5 model output used in the optimal detection, comprising MPI-ESM-LR and HadGEM2-ES historical and piControl integrations. ALL = natural + anthropogenic external forcing. * control length after drift removal.

Model	Experiment (name)	Forcing	Ensemble size (n)	Control length (years)
MPI-ESM-LR	Historical (r1:ni1p1)	ALL	3	
	piControl (r1i1p1)		1	1000
HadGEM2-ES	Historical (r1:ni1p1)	ALL	4	
	piControl (r1i1p1)		1	965*

scribe non-evolving pre-industrial forcings, and an ensemble of historical experiments for the period ∼ 1850 to 2005 (Table 1). The piControl runs were initialised from a pre-industrial spin-up to pseudo-equilibrium. Boundary conditions for the historical experiments are prescribed from observations as an evolving record of climate forcing factors. These include external forcings from: historical greenhouse gas concentrations, tropospheric and stratospheric ozone changes, surface emissions of tropospheric aerosols and land use changes as well as the natural forcings from changes in solar irradiance and volcanic aerosols. Each member of the model ensemble is initialised from a different point in the control simulation in order to create a spectrum of equally plausible historical simulations each starting at a different phase of internal variability. The MPI-ESM-LR and HadGEM2-ES CMIP5 experimental design and spin-up are described in Mauritsen et al. (2012) and Jones et al. (2011), respectively. A persistent climate drift in HadGEM2-ES oxygen fields has been isolated from a 20 yr low-pass filtered version of the control run and subtracted at each grid point of the historical [O_2] data using the corresponding segment of the piControl integration. Since residual drift can also bias noise estimates required for the optimal fingerprinting procedure (as estimated from long control experiments) an equivalent point-by-point drift is also removed from the HadGEM2-ES piControl integration. In this case removal of a linear trend diagnosed from a low pass filter provides an adequate representation of climate drift behaviour, as distinct from natural internal periodicity.

2.2 Optimal fingerprinting method

We test the null hypothesis that historical changes in [O_2] are indistinguishable from natural internal variability ($\beta = 0$) using an optimal detection algorithm (Hasselmann, 1997; Allen and Tett, 1999). This statistical technique is widely used in the detection and attribution of climate change (e.g., IDAG, 2005; Hegerl et al., 2010) and also provides a powerful test of ESM performance which includes the effect of natural internal variability. We use this technique to regress model simulated patterns of [O_2] change (x_i) against corresponding observed patterns (y) using a Total Least Squares (TLS)
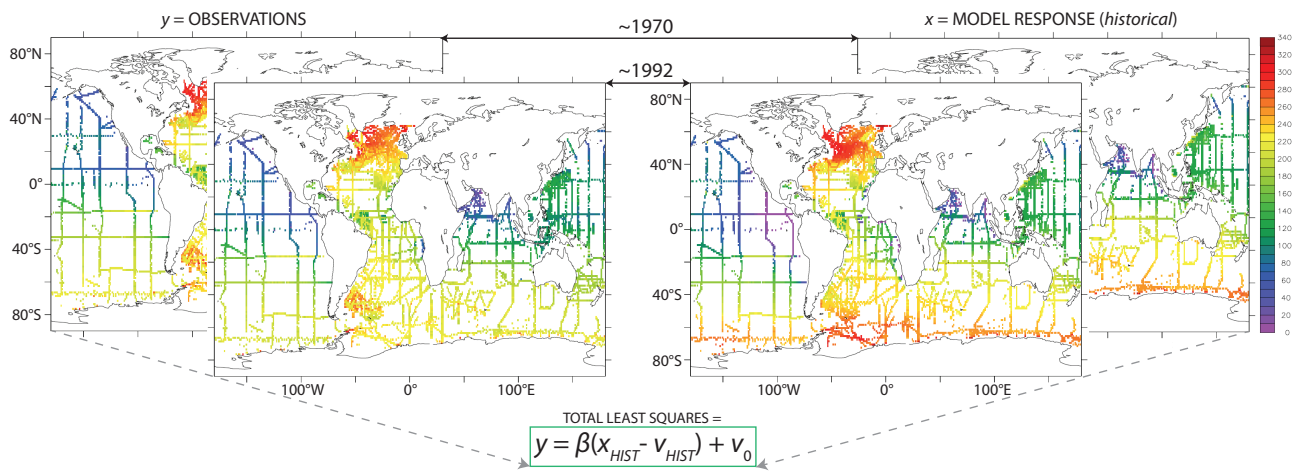
Fig. 1. Schematic summary of the optimal detection method used here. The left panels show the observed distribution of dissolved oxygen data (y; μmol kg^{-1}) for ~ 1970 and ~ 1992, averaged between 100–3000 m. The right panels show the masked and re-mapped ensemble mean model output (x_{HIST}) from historical simulations of ESMs (MPI-ESM-LR or HadGEM2-ES) averaged between 100–3000 m. y and x_{HIST} are then provided to the Total Least Squares regression as zonal mean [O$_2$] differences between the two time periods, where v_{HIST} and v_0 are estimates of the noise in the model response and observations, respectively.

method (Allen and Stott, 2003) which estimates the scaling factors (β_i) required to match simulated and observed changes following Eq. (1):

$$y = \sum_{i=1}^{1} \beta_i (x_i - v_i) + v_0 \qquad (1)$$

where in the single signal case, for the i-th forcing, v_i is the error in the model response x_i and v_0 is the climate noise in the observations. If the confidence interval that contains β exceeds zero a signal is detected in response to an imposed forcing i. If this scaling factor is consistent with one the simulated and observed responses are said to be similar in magnitude. In this study, the simulation consists of the ESM, all of its parameters and all of the temporal and spatial variations of the external forcing i transformed through the physical and chemical processes represented in the model. Thus, a value of β consistent with one demonstrates that forcing i and the theory are consistent with the observations, and a β estimate which is significantly greater than 0 demonstrates that the null hypothesis of no contribution from forcing i can be rejected. TLS is likely to yield a more robust β estimate than an Ordinary Least Squares (OLS) approach because it also includes a signal error term (v_i) arising from averaging across a finite model ensemble to obtain the simulated response pattern. Signal error (v_i) is inversely proportional to the model ensemble size (n) and can negatively bias scaling factors (Allen and Tett, 1999), particularly for variables where the forced response is small relative to internal variability. Thus, TLS is widely used in optimal detection studies as a more conservative approach which explicitly accounts for the effect of noise in simulated response patterns relative

to that which is found in the observations (e.g., Stott et al., 2003, 2008; Terray et al., 2012).

We focus our single fingerprint analysis on the "all forcings" historical scenario of MPI-ESM-LR and HadGEM2-ES, with model output being bi-linearly interpolated onto a $1° \times 1°$ grid and masked to emulate the pattern of missing values found in the observations of Helm et al. (2011) (Fig. 1). [O$_2$] changes between ~ 1970 and ~ 1992 are then calculated for these experiments and provided as model response patterns (x_{HIST}) in the TLS regression against corresponding observed [O$_2$] changes (y) in order to estimate scaling factors (β). The instrumental record of dissolved O$_2$ measurements is not sufficiently long to get a reliable approximation of internal climate variability (v_0), and also includes perturbations driven by external forcing. In order to characterise unforced climate variability in [O$_2$] on a global scale, we estimate v_0 by sampling non-overlapping 22 yr slabs of [O$_2$] fields taken from the long (~ 1000 yr) control integrations of MPI-ESM-LR and HadGEM2-ES. The non-overlapping model [O$_2$] fields are also masked and re-gridded as "pseudo-observations". Subsampled model piControl output is then used to (1) estimate internal variability in [O$_2$] data and (2) place 5–95 % uncertainty limits on calculated scaling factors.

The TLS regression is carried out in a reduced dimension space where model and observed data are projected onto k leading Empirical Orthogonal Functions (EOFs) of simulated internal variability. Signal-to-noise ratios are optimised in a standard way via normalisation of observations and model response patterns by internal climate variability (e.g., IDAG, 2005), a transformation which down-weights patterns of [O$_2$] change with high internal variability and vice versa. Dimensionality of the detection space in this study is further reduced by averaging across multiple en-

semble members in each historical CMIP5 experiment (see Table 1), and by calculating zonal means for observed and modelled [O_2] changes. The analysis of Helm et al. (2011) shows pronounced global [O_2] decreases between ~ 1970 and ~ 1992 in all ocean basins, accommodating a zonally averaged spatial domain. This removes small-scale variability in [O_2] and allows the detection analysis to focus on global and basin-scale patterns of oxygen change. Gridded dissolved oxygen data are also smoothed vertically and horizontally. Modelled [O_2] changes are calculated as temporal averages of annual data between 1960–1980 and 1990–2000 in order to be consistent with the historical spread of observations, with sensitivity tests demonstrating simulated [O_2] changes to be relatively invariant across a range of small changes in averaging period.

We prepare model fingerprints and observations for the optimal detection analysis using two different zonal averaging schemes: (1) 1-D analysis: zonal mean of depth-averaged [O_2] changes between 100 and 3000 m; (2) 2-D analysis: zonal mean [O_2] changes explicitly resolve depth (between 100–3000 m) globally and separately for the Atlantic and Pacific basins. Both MPI-ESM-LR and HadGEM2-ES models are used in the 1-D analysis. The MPI-ESM-LR model only is used to extend the work and provide a more detailed 2-D analysis. For the 2-D optimal detection, masked MPI-ESM-LR output and observations are remapped onto a $\sim 5°$ latitude by $\sim 10°$ longitude grid with 40 unevenly spaced z-levels to mediate the effects of internal noise on the signal whilst still retaining the depth structure of meridional [O_2] change between ~ 1970 and ~ 1992. The use of several models and several spatial averaging schemes provides multiple model fingerprints that are used to quantify possible errors in model response patterns ("structural uncertainty") driven by inadequate representation of physical and biogeochemical process in ESMs (e.g., Hegerl and Zwiers, 2011).

In order to avoid spurious detection it is a necessary prerequisite that the internal variability estimated from control simulations (v_0) provides a realistic estimate of observed climate noise in [O_2]. As such, the number of EOFs retained in the optimal detection analysis is guided by checking the fidelity of model simulated internal variability against the residual observed variance at k truncations using a standard residual consistency F test (Allen and Tett, 1999). This check is used to test the null hypothesis that internal variability as simulated by models is consistent with observed variability on the scales retained in the analysis. Failure of the residual consistency test could also indicate that the timing or pattern of ESM response is incorrect. We choose to truncate the 1-D MPI-ESM-LR analysis at 40 EOFs ($k = 40$) since this is the first truncation for which the residual test passes the consistency check, reducing the likelihood of keeping poorly sampled higher order EOFs in the optimal detection. Fewer truncations ($k = 31$) are needed to pass the residual consistency check in the 1-D HadGEM2-ES analysis, however, scaling factors are not substantially different at higher k

Table 2. Comparison of decadal standard deviations of [O_2] (μmol kg^{-1}) between observations and CMIP5 piControl experiments at Ocean Station Papa in the Eastern North Pacific (50° N, 145° W, $\sigma = 26.9$, ~ 250–350 m) and the Oyashio Current region in the Western North Pacific (39–42° N, 143–145° E, $\sigma = 26.9$, ~ 400–450 m). piControl output from MPI-ESM-LR and HadGEM2-ES is sampled to calculate multiple estimates of decadal standard deviations for each time series. The mean and (10th–90th) percentile ranges of piControl estimates are shown. Observations are detrended using a linear fit to extract variability. Drift has been removed from the control integration of HadGEM2-ES as described in Sect. 2.1.

	Observations	MPI-ESM-LR	HadGEM2-ES
Ocean Station Papa	11.1	9.1 (6.2–12.0)	5.3 (3.9–6.7)
Oyashio Current region	8.1	9.9 (4.9 – 15.1)	1.4 (0.4–2.4)

values (see Sect. 3.2). The 2-D analysis both globally and between ocean basins required the maximum number of EOFs ($k = 49$) to be retained in the regression to pass the residual consistency check (apart from for the Atlantic basin where the F test p value falls marginally outside the 5–95 % range). A higher number of retained EOFs is consistent with the greater amount of information needed to describe spatial patterns of both depth and latitude variations, with more modes being necessary to explain variability in the depth-resolved signal.

In addition to the residual consistency test, we assess the reliability of model simulated climate variability by comparing piControl output with detrended subsurface [O_2] measurements from two long-term time series: Ocean Station Papa (1956–2007, Whitney et al., 2007) and the Oyashio Current region (1968–1998, Ono et al., 2001). Observed decadal standard deviations calculated for both time series fall within the 10–90 % ranges of MPI-ESM-LR control simulation estimates, demonstrating that this model provides a robust estimate of internal variability in [O_2] on decadal timescales (Table 2). The HadGEM2-ES control simulation significantly underestimates decadal variability in [O_2] when compared to time series data and is, thus, less reliable than simulations from MPI-ESM-LR in the context of our analysis.

3 Results

3.1 Model-data comparison

We evaluate model performance by comparing simulated and observed zonally averaged [O_2] distributions for the ~ 1992 time period. Historical integrations of MPI-ESM-LR and HadGEM2-ES are able to capture the general latitude-depth pattern of [O_2] distribution present in the observations (Fig. 2). The [O_2] minimum between 20° S and 20° N is reproduced at mid-depths by both models, with increasing [O_2] towards higher latitudes in both hemispheres. However,

HadGEM2-ES simulates higher than observed oxygen concentrations in the [O_2] minimum. MPI-ESM-LR overestimates the spatial extent of low-latitude [O_2] minimum waters and produces erroneously high [O_2] south of 60° S. Ilyina et al. (2013) present a detailed comparison between biogeochemical tracers in MPI-ESM-LR CMIP5 historical simulations and observations using a range of statistical metrics to assess model capability.

A marked meridional structure also exists in observed depth-averaged zonal mean [O_2] changes, with deoxygenation increasing with latitude poleward of 40° (up to 12 µmol kg^{-1}) in both hemispheres (Fig. 3). Pronounced deoxygenation in the mid- to high-latitude ocean is opposed by no change or a small zonal mean [O_2] increase of $\sim$ 1–2 µmol kg^{-1} between 20° S and 20° N. Both MPI-ESM-LR and HadGEM2-ES historical integrations show predominant global decreases in [O_2] at high latitudes consistent with but smaller ($<$ 5 µmol kg^{-1}) than the observations, with no net oxygen change at low latitudes. Observed and modelled global mean [O_2] decreases both reach their maxima at $\sim$ 60° N. In some regions MPI-ESM-LR and HadGEM2-ES historical experiments show a much reduced or inverse [O_2] change signal, particularly in areas of the Southern Ocean (65° S) and at $\sim$ 50° N where significant observed zonal mean [O_2] decreases (3–6 µmol kg^{-1}) contrast with small oxygen increases in both models.

The consistency between models and observations is further examined in zonal mean sections as a function of depth (Fig. 4). [O_2] data show acute deoxygenation of the ventilated thermocline (100–1000 m depth) at all latitudes, with large oxygen decreases ($>$ 10 µmol kg^{-1}) extending throughout the water column poleward of 40° in both hemispheres (Fig. 4a). These regions of deoxygenation are countered by areas of increasing [O_2] (5–10 µmol kg^{-1}) located between 30° S and 30° N below 1000 m depth and within the shallow subtropical gyres (15–30°). Both models generally reproduce the key zonal mean features of observed [O_2] change, simulating [O_2] decreases at high latitudes and within the thermocline, with an area of increasing [O_2] beneath $\sim$ 1500 m depth at mid- to low latitudes. In agreement with data, MPI-ESM-LR simulates strong [O_2] depletion extending throughout the water column north of 40° N and shows a distinct region of increased [O_2] beneath $\sim$ 1500 m between 30° S and 30° N (Fig. 4b). MPI-ESM-LR also displays significant deoxygenation between 40° S and 60° S, but this is opposed by a limb of [O_2] increase within the interior of the Southern Ocean (a feature which is entrained into the depth-averaged trend of Fig. 3). HadGEM2-ES exhibits deoxygenation in the upper ocean at mid- to high latitudes in both hemispheres, with [O_2] depletion extending down to $\sim$ 3000 m depth south of 65° S and at $\sim$ 60° N (Fig. 4c). Near ubiquitous [O_2] decreases are also simulated in the ventilated thermocline, apart from within the subtropical gyres (15–30°) where HadGEM2-ES reproduces, in part, the observed [O_2] increases. However, the oceanic interior between 60° S and

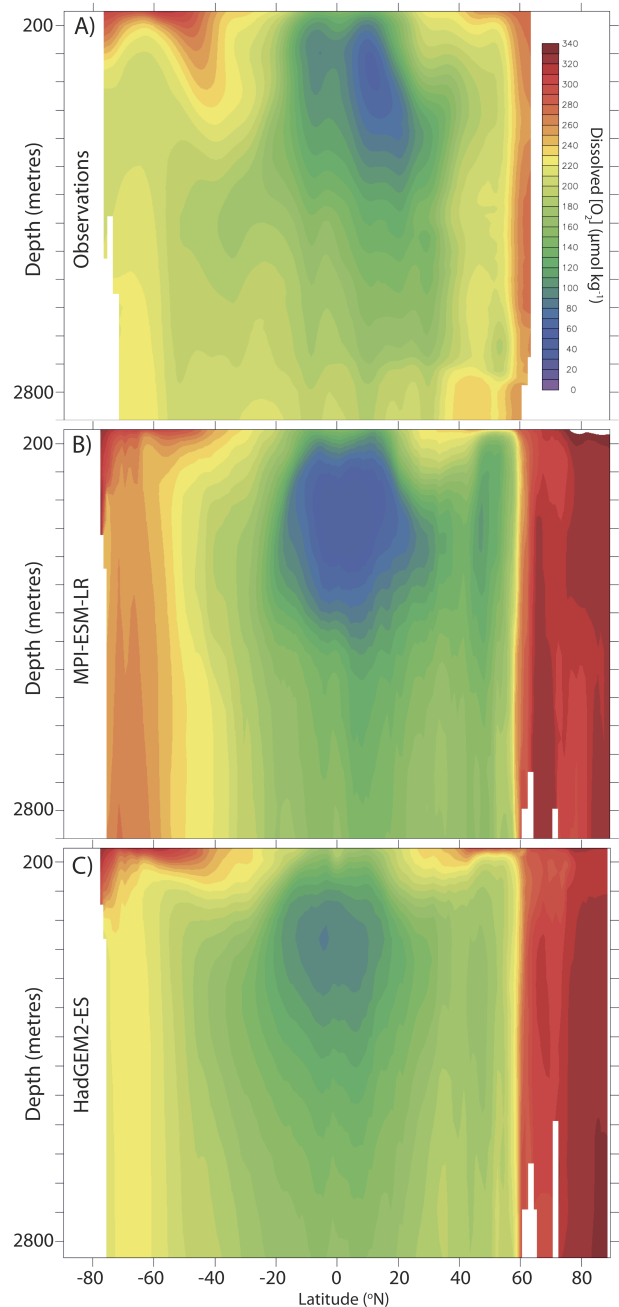**Fig. 2.** Zonal mean [O_2] distribution (µmol kg^{-1}) for $\sim$ 1992 as a function of latitude and depth from **(A)** observations (Helm et al., 2011), and historical integrations of **(B)** MPI-ESM-LR and **(C)** HadGEM2-ES.

40° N becomes increasingly oxygenated over the modelled analysis period, with an anomalous region of positive [O_2] change below $\sim$ 1500 m in the polar north and protruding into the upper ocean at $\sim$ 65° S. Both models fail to emulate the pattern of [O_2] loss recorded by observations in the low-latitude OMZs ($<$ 1000 m), displaying small [O_2] increases within the tropical thermocline between 20° S and 20° N.
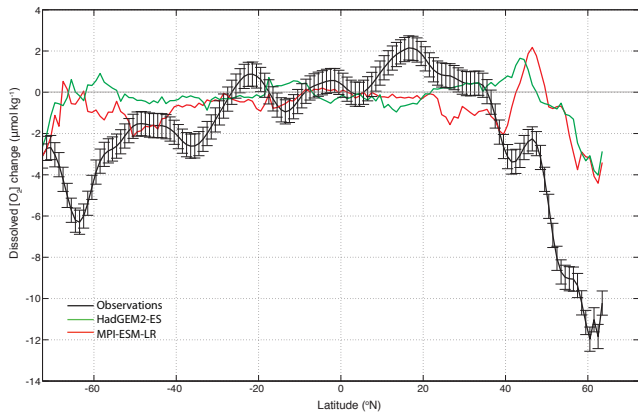
Fig. 3. Zonal mean change ($\sim$1992 minus $\sim$1970) in [O_2] (μmol kg^{-1}) as a function of latitude averaged between 100–3000 m depth for observations (black; Helm et al., 2011), and HadGEM2-ES (green) and MPI-ESM-LR (red) historical integrations. Error bars in observed [O_2] are given at the 95 % confidence level and are associated with instrumental uncertainty, un-resolved ocean processes and methodological uncertainty in forming the zonal averages (derived from an a priori estimate of noise using the method of Bindoff and Wunsch, 1992).

This failure of the models to replicate the low-latitude OMZs is discussed in Sect. 1 and is well known. Overall, simulated [O_2] changes for both models tend to be similar but smaller in magnitude than recorded by observations, with MPI-ESM-LR exhibiting generally more skill at reproducing patterns of observed [O_2] change as a function of depth.

On a basin scale (Fig. 5) there is a higher level of agreement between modelled and observed water mass changes over the analysis period in the Pacific compared to the Atlantic. MPI-ESM-LR reproduces observed deoxygenation of the high-latitude North Pacific down to $\sim$3000 m depth, and shows a general trend towards decreasing [O_2] levels within the ventilated thermocline. In comparison, the structure of modelled oxygen changes within the Atlantic Ocean is largely inconsistent with observations. MPI-ESM-LR simulates extensive deoxygenation (up to 4 μmol kg^{-1}) throughout much of the Atlantic between 1000 and 3000 m depth, and does not resolve major regions of observed [O_2] increase in the interior of the Atlantic between 30° S and 30° N. In both the Atlantic and Pacific basins, following the global trend, MPI-ESM-LR does not reproduce the pronounced deoxygenation signal recorded within the low-latitude OMZs. This model-data mismatch is most apparent in the thermocline of the tropical Atlantic where MPI-ESM-LR shows strong [O_2] increases of > 5 μmol kg^{-1} in opposition to observed [O_2] depletion.

Changes in upper ocean stratification are associated with changes in the ventilation (and oxygenation) of subsurface water masses. Figure 6 shows historical changes in zonal mean upper ocean density stratification as a function of lat-
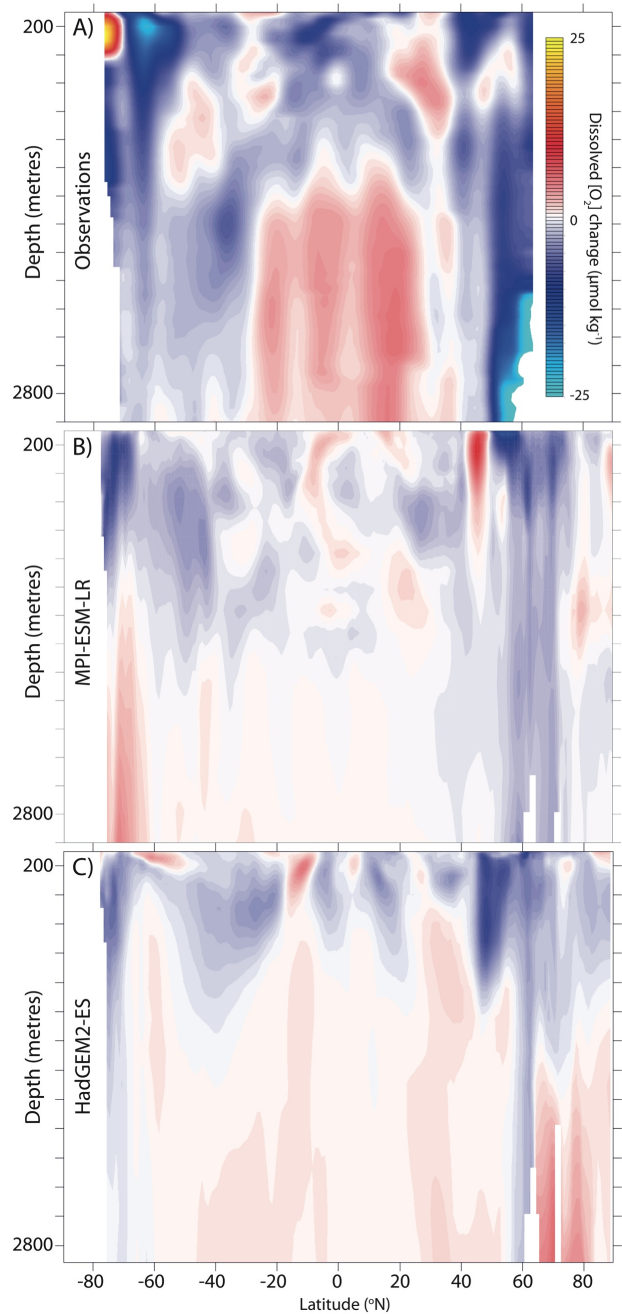

Fig. 4. Zonal mean change ($\sim$1992 minus $\sim$1970) in [O_2] (μmol kg^{-1}) as a function of latitude and depth from **(A)** observations (Helm et al., 2011), and historical integrations of **(B)** MPI-ESM-LR and **(C)** HadGEM2-ES.

itude (between 200 m and 1000 m) from observations and for MPI-ESM-LR and HadGEM2-ES models. Observed and model derived zonal mean density stratification generally increases at mid- to high latitudes between $\sim$1970 and $\sim$1992, with decreases in global stratification within the subtropical and equatorial ocean. However, at high latitudes in the
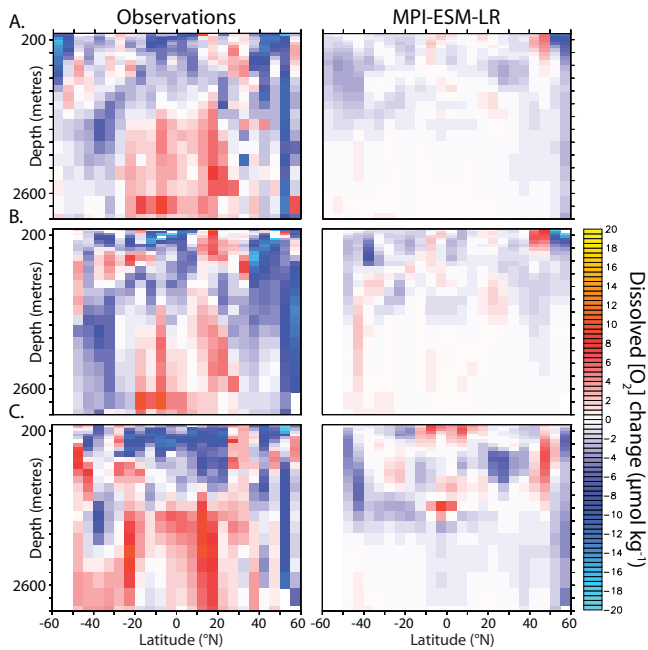
Fig. 5. Observed (left) and simulated (right) patterns provided to the 2-D optimal detection as zonal mean [O$_2$] changes (µmol kg^{-1}) as a function of latitude and depth for the global ocean (60° S–60° N; **A**), and for the Pacific (**B**) and Atlantic (**C**) ocean basins. Masked MPI-ESM-LR [O$_2$] fields and observations are interpolated onto a ∼ 5° latitude by ∼ 10° longitude grid with 40 unevenly spaced depth levels and zonally averaged.

Fig. 6. Zonal mean change (∼ 1992 minus ∼ 1970) in upper ocean stratification as a function of latitude from observations (black; Helm, 2008), and HadGEM2-ES (green) and MPI-ESM-LR (red) historical integrations. The Stratification Index is calculated as the vertical density gradient ($\delta\rho/\delta z$) between 200–1000 m depth. Model stratification is estimated by calculating density (kg m^{-3}) from CMIP5 historical temperature and salinity fields at 200 m and 1000 m depth levels. Positive values indicate an increase in upper ocean stratification over time. These stratification changes differ from those presented by Helm et al. (2011) because we exclude surface data (< 100 m) from the Hadley SST climatology. Error bars for observed stratification changes are given at the 95 % confidence level as in Fig. 3.

Southern Ocean both MPI-ESM-LR and HadGEM2-ES underestimate the observed stratification change signal, with major increases in stratification between 50° S and 60° S not reproduced by either model. Analysis of the density structure of these models show that in this region the density profile is more homogeneous than observed with more vigorous vertical mixing.

3.2   Optimal detection analysis

We apply the optimal detection approach outlined in Sect. 2.2 to quantitatively investigate the consistency of modelled and observed changes in [O$_2$], and to assess the detectability of observed changes in response to external forcing. Figure 7 shows observations and model response patterns provided to the 1-D TLS regression using MPI-ESM-LR and HadGEM2-ES oxygen fields, plotted against the model-estimated range of internal variability. The envelope of estimated internal variability in [O$_2$] scales with latitude in both models such that sub-polar and polar regions where a strong zonal mean deoxygenation signal is observed coincide with areas of largest internal variability (± 2–4 µmol kg^{-1}). Nevertheless, the magnitude of observed zonal mean [O$_2$] change exceeds the spread of simulated internal variability in the mid- to high-latitude ocean, with a deoxygenation signal emerging
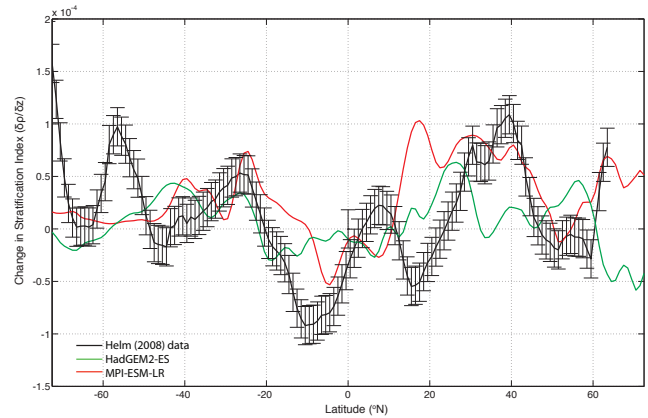
above natural internal variability poleward of 30–40° in both hemispheres. Simulated internal variability is smaller in the HadGEM2-ES piControl experiment such that observed [O$_2$] decreases become distinct from internal variability across more of the mid-latitude ocean. Both MPI-ESM-LR and HadGEM2-ES historical integrations simulate more muted deoxygenation signals compared to the observations which are largely within the range of internal variability at all latitudes. Structural errors in the pattern of [O$_2$] change simulated by both models also generally fall within the spread of internal variability. Observations and model response patterns have also been provided to the TLS regression from depth-resolved 2-D [O$_2$] changes globally and for the Pacific and Atlantic basins between 100–3000 m depth (Fig. 5).

Scaling factors (β) resulting from the single fingerprint analysis confirm that a statistically significant change in observed [O$_2$] in response to external forcing is detected and inconsistent with simulated internal variability at a 90 % confidence level (Fig. 8). Projection of observations onto the simulated "all forcings" response of MPI-ESM-LR and HadGEM2-ES shows a detectable change in depth-averaged zonal mean [O$_2$] between ∼ 1970 and ∼ 1992, since the best estimates of the scaling factors and their confidence intervals at the global scale are significantly different from 0. Scaling factors calculated for the MPI-ESM-LR and HadGEM2-ES 1-D analyses are 2.82 and 3.59, respectively, indicating that the effect of external forcing is detectable and that the
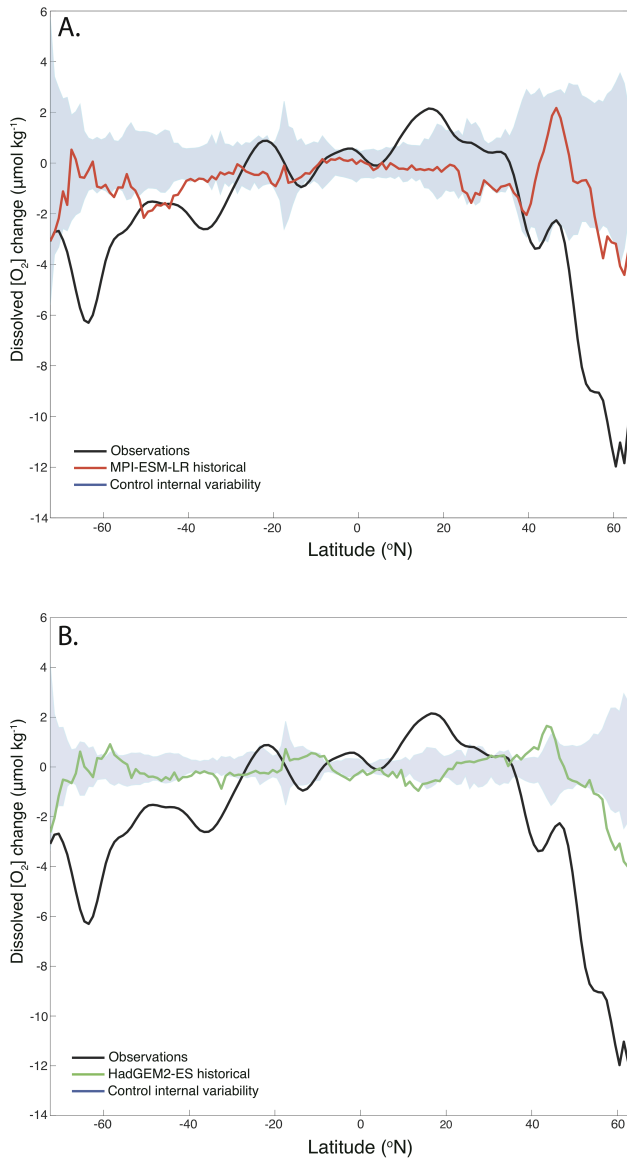
Fig. 7. Observed (black) and simulated (red or green) patterns provided to the 1-D optimal detection as zonal mean [O_2] changes (µmol kg^{-1}) averaged between 100–3000 m depth using MPI-ESM-LR (**A**) and HadGEM2-ES (**B**) historical experiments. The spread of internal variability (shown by the shaded area) is estimated by sampling model piControl simulations, as described in Sect. 2.2.

Fig. 8. Best estimates and 5–95 % uncertainty ranges of regression coefficients (β) from single fingerprint optimal detection analysis. Two types of model response pattern are projected onto dissolved oxygen observations. Left: a 1-D global depth-averaged (100–3000 m) zonal mean [O_2] change using MPI-ESM-LR ($k = 40$) and HadGEM2-ES ($k = 31$). Right: a 2-D depth-resolving zonal mean [O_2] change globally and for the Atlantic and Pacific Ocean basins ($k = 49$) using MPI-ESM-LR only (* indicates a failed residual consistency test at all EOF truncations).

models and observations have significant correlations in the latitudinal pattern of variation. However, since these β values are greater than one we can infer that the ESM simulated [O_2] responses to external forcing are significantly underestimated and need to be amplified (by a factor of between ~ 2 and ~ 4) to be consistent with observed changes. The residual consistency test passes for both model experiments indicating no inconsistency between residual observed variance and model simulated internal variability, and suggesting that
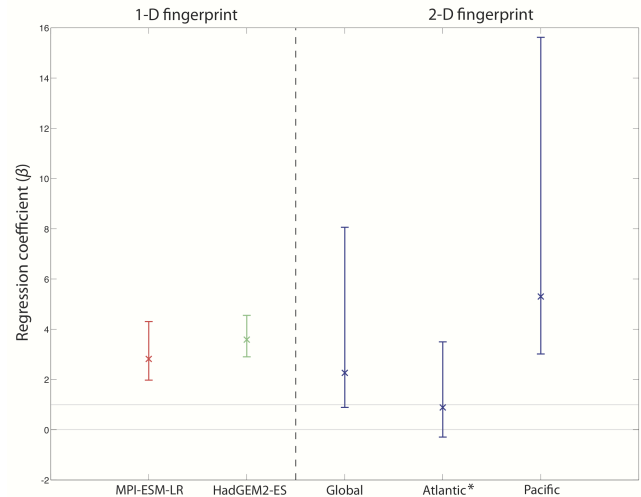
both ESMs simulate the externally forced signal adequately to explain observed [O_2] changes. For both MPI-ESM-LR and HadGEM2-ES β estimates for the 1-D analysis are robust across a range of truncations proximal to the chosen values of k, adding confidence to the presented optimal detection results (Fig. 9).

In agreement with the 1-D result, the 2-D detection analysis using MPI-ESM-LR yields a best estimate scaling factor of greater than 1 ($\beta = 2.26$) for the global ocean, although in this case the 5–95 % uncertainty bounds on this value are consistent with 1. This range of estimated β values demonstrates that the simulated response of the MPI-ESM-LR model provides a good fit to observed zonal mean [O_2] changes between ~ 1970 and ~ 1992 in response to external forcing. A positive detection result ($\beta = 5.30$) is also found for observed [O_2] changes in the Pacific Ocean using a 2-D pattern with MPI-ESM-LR. However, this scaling factor is much larger than 1 indicating poor model-data agreement in amplitude. In contrast, the null hypothesis that observed changes in marine oxygen are caused by natural internal variability cannot be rejected for the Atlantic Ocean, where the 5–95 % range of β estimates are indistinguishable from zero, with a best estimate scaling factor of less than 1 ($\beta = 0.89$). The residual consistency test fails for this region only, suggesting that there are elements of the observed [O_2] variability which are not well captured by the model in the Atlantic Ocean. Overall, an external influence on historical changes in observed [O_2] is detected for the zonally averaged global
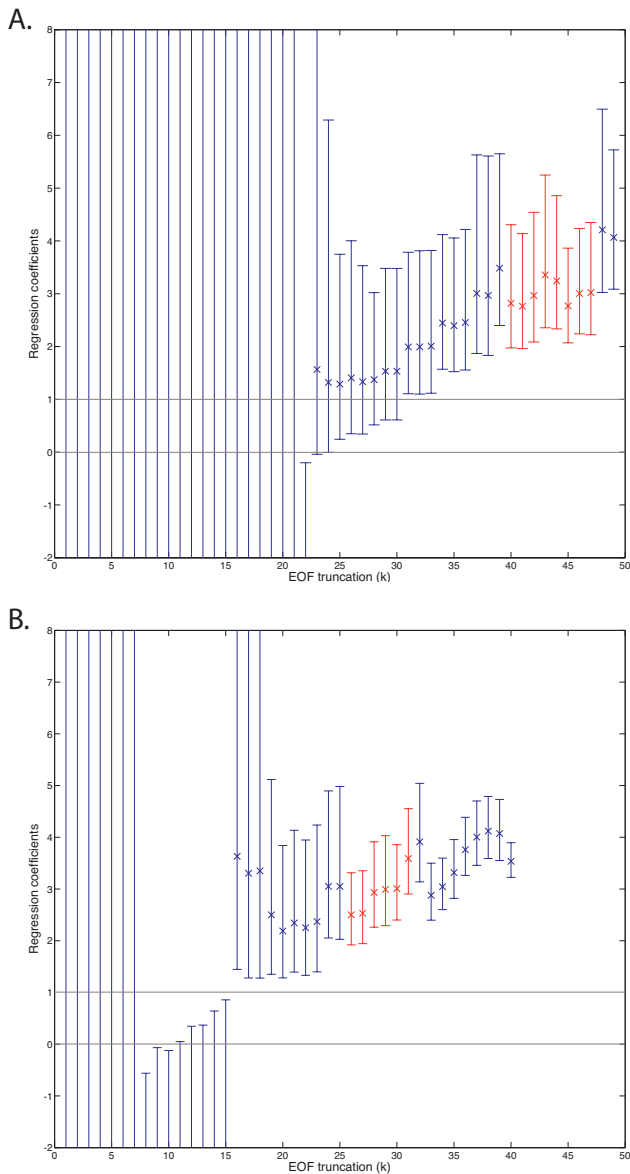
Fig. 9. Regression coefficients (β) and their 90 % confidence intervals plotted as a function of EOF truncation (k) for 1-D MPI-ESM-LR (**A**) and HadGEM2-ES (**B**) optimal detection analyses. β estimates are relatively invariant where the residual consistency check passes (at k values shown in red).

ocean and Pacific basin, although the observed response is larger than that simulated by ESMs at all scales.

4   Discussion

This study presents, for the first time, an optimal fingerprinting analysis detecting statistically significant global decreases in observed oceanic [O$_2$] in response to external forcing, as distinct from internal variability driven by the leading climate modes. The primary natural external forcing imposed

on [O$_2$] is explosive volcanism, a perturbation on the climate system that is generally limited to the ocean's upper ~ 500 m on interannual timescales (Frölicher et al., 2009). Because the observed and modelled deoxygenation occurs throughout the water column and on interdecadal timescales, we can infer that rising greenhouse gas concentration is the main driver of a detectable external forcing on historical [O$_2$] changes between ~ 1970 and ~ 1992. In the 1-D analysis we find the most detectable [O$_2$] change signal relative to internal variability at mid- to high latitudes in regions of water mass renewal where the observed subsurface zonal mean oxygen decreases are largest. Helm et al. (2011) attribute ~ 85 % of global ocean deoxygenation in polar regions to elevated upper ocean stratification which increases the ventilation age of downwelling water parcels allowing for more biological oxygen consumption to occur, consistent with a range of prognostic modelling studies (Sarmiento et al., 1998; Bopp et al., 2002; Matear and Hirst, 2003). We also find significant increases in upper ocean stratification between ~ 1970 and ~ 1992 at high latitudes for MPI-ESM-LR and HadGEM2-ES historical integrations concurrent with enhanced subsurface deoxygenation. Moreover, the inability of either model to reproduce observed increases in zonal mean stratification in the Southern Ocean could provide a mechanistic explanation for why modelled [O$_2$] decreases are considerably smaller than observed in this region. Jungclaus et al. (2013) also show that the ocean model within MPI-ESM-LR overestimates ventilation in the Antarctic Circumpolar Current system, producing enhanced oxygenation of the oceanic interior in this region. Large uncertainties are entrained into model estimates of physical circulation changes owing to incomplete assessments of the competing roles of stratification and wind forcing (e.g., Le Quéré et al., 2007; Böning et al., 2008; Downes et al., 2011) in controlling ventilation processes in a warmer world.

Although we find a clearly identifiable influence of external forcing on recorded patterns of [O$_2$] change, best estimates of β for both 1-D and 2-D global analyses are consistently greater than one suggesting that model responses need to be scaled up to match the observations. There are a number of potential explanations for this result. One possibility, endemic to climate sensitive variables that are locally heterogeneous and difficult to simulate (Masson and Knutti, 2011), is that [O$_2$] data could exhibit too much small-scale variability to be directly comparable with smoother model fields. This potential discrepancy could yield observed change signals with artificially inflated amplitudes relative to model response patterns, although we try to account for this effect by spatially smoothing the data prior to model evaluation. Detection results are robust across a range of EOF truncations and temporal averaging schemes with no evidence of under or over fitting within the TLS regression, such that methodological choices about data processing are unlikely to have significantly biased scaling factors. Model errors could also contribute to elevated β values if imposed forcings in

CMIP5 historical experiments are too weak or ESMs are less responsive to external forcing than in reality. Since CMIP5 historical forcings are prescribed from observations (Taylor et al., 2009, 2012) the former seems unlikely; however, there is evidence to support the proposition that the current generation of ESMs systematically underestimate observed variability in [O_2]. For example, in good agreement with our 1-D scaling factors, hindcast simulations initialised globally (Rödenbeck et al., 2008) and for the North Pacific (Deutsch et al., 2005, 2006) underestimate observed interannual to decadal variability in ocean oxygen by a factor of between ~ 2 and ~ 3. It has been suggested that coarse resolution models might underestimate variability in simulated passive tracers since they do not resolve mesoscale ocean dynamics (Hirst et al., 2000; Jochum et al., 2007). Moreover, coupling frequency between the ocean and atmosphere (once a day in MPI-ESM-LR and HadGEM2-ES simulations) might also be insufficient to fully capture climate variability in state-of-the-art ESMs. This could support the assertion that the response of models to secular climate change may also be underestimated, since a more naturally variable ocean-climate system will likely be more sensitive to imposed external forcings (von Storch and Zwiers, 1999; Swanson et al., 2009). However, the piControl derived estimates of natural internal variability used in this analysis are shown to be consistent with observed variance using a standard residual consistency test and comparison of simulated noise with two observational estimates (Sect. 2.2 and Table 2).

Qualitative model-data comparison (Sect. 3.1) suggests that regional differences between modelled and observed patterns of [O_2] change could also contribute to the weaker simulated zonal mean signal. However, consistent with the assumption of the TLS detection model that "structural uncertainty" has the same structure as internal variability (e.g., Terray et al., 2012), model errors generally fall within the range of internal variability (Fig. 7). Regional differences probably result from the omission of key biogeochemical and physical processes that control marine oxygen distribution at a local scale. Particularly, 2-D model response patterns are unable to reproduce a net deoxygenation signal in the tropical thermocline, consistent with deficiencies in modelling studies (Bopp et al., 2002; Matear and Hirst, 2003; Stramma et al., 2012a). Stramma et al. (2012a) report that the spurious tropical [O_2] increases at 300 dbar in the UVic ESM may be primarily related to the inability of coarse resolution ocean models to resolve the fine scale equatorial currents which control the oxygen budget of the OMZs (Brandt et al., 2010; Stramma et al., 2010). Their sensitivity study suggests that alternative proposed model deficiencies such as artificially high rates of diapycnal mixing within the tropical thermocline (Gnanadesikan et al., 2007) and the omission of a pCO$_2$ sensitive C:N stoichiometry in primary production (Oschlies et al., 2008; Tagliabue et al., 2011) do not resolve the erroneous tropical [O_2] increases in their model. In the case of MPI-ESM, a comparison between the low resolution

(analysed in this study) and eddy-permitting model configurations (MPI-ESM-MR with a nominal resolution of $\sim 0.4°$) shows only a slight improvement in the representation of [O_2] in the thermocline (Ilyina et al., 2013). This indicates that eddy-permitting spatial resolution of the ocean model alone is insufficient to solve the coupled models deficiencies with respect to [O_2] dynamics. Other model-data discrepancies, particularly simulated oxygen increases within the ocean interior at high latitudes, are likely to be related to persistent errors in model physical mixing and deep ocean ventilation, as reported for historical CMIP5 experiments of IPSL-CM5A and CNRM-CM5.1 (Séférian et al., 2012). In addition, the inability of models to capture [O_2] dynamics in ice-covered high-latitude areas can be attributed to uncertainties in the underlying sea-ice models related to the growth and melting of seasonal ice (e.g., Notz et al., 2013).

Counter to the global and Pacific Ocean detection results, our analysis finds observed patterns of [O_2] change in the Atlantic basin to be indistinguishable from natural internal variability. The high internal variability of Atlantic oxygen has been reported by others (e.g., Johnson and Gruber, 2007) and in an analysis of the essential ocean indices the Atlantic scored poorly from the perspective of the strength of signal-to-noise ratio (Banks and Wood, 2002). These findings are supported by modelling studies which show secular increases in Apparent Oxygen Utilisation (AOU) of the subsurface North Atlantic to be rendered statistically insignificant by internal variability associated with the North Atlantic Oscillation (Frölicher et al., 2009). Thus, in the Atlantic, high levels of internal variability coupled to a smaller observed mid- to high-latitude deoxygenation trend (relative to the Pacific and Southern Oceans; Helm et al., 2011) reduce the signal-to-noise ratio, contributing to the null detection result for this basin. The same conclusion was reached for sea surface salinity where the single fingerprint analysis of Terray et al. (2012) shows robust detection of observed salinity changes in response to anthropogenic forcing for the Pacific basin but with scaling factors consistent with 0 for the Atlantic Ocean.

5 Summary and conclusions

Our analysis shows that the global consistency and pronounced meridional structure of [O_2] changes between ~ 1970 and ~ 1992 recorded by Helm et al. (2011) provide a distinct fingerprint of climate change in the oceans, which is robustly detected against internal variability as simulated by models. The unprecedented spatiotemporal coverage of data and sensitivity of [O_2] to climate perturbation provides a unique opportunity to validate the response of ESMs to conditions of global change on multidecadal timescales. Rigorous assessment of ESMs against observations using detection and attribution methods supports the need to improve model physics and biogeochemistry, but also reveals that

these models already have significant capacity to simulate many aspects of the ocean system, particularly the deoxygenation at mid- to high latitudes. However, both MPI-ESM-LR and HadGEM2-ES underestimate the observed [O_2] response to external forcing, suggesting that model projections for future ocean deoxygenation in response to climate change may be too conservative. Furthermore, both models perform poorly at low latitudes, indicating that model projections in that region may not be reliable.

Detectable anthropogenic contributions to recent trends in ocean temperature (Barnett et al., 2005; Pierce et al., 2006; Gleckler et al., 2012) and salinity (Stott et al., 2008; Terray et al., 2012) have already been identified. Taken together, these marked changes to global heat and freshwater fluxes implicitly support our finding that stratification-driven global deoxygenation has started to emerge from the envelope of internal variability. We find the most detectable changes in [O_2] relative to internal variability to occur at high latitudes where the independent and synergistic effects of secular ocean warming, acidification and deoxygenation could have a major impact on polar ecosystems and biogeochemical cycles (Gruber, 2011). Subsurface [O_2] changes at high latitudes (particularly in the Pacific and Southern Oceans) should, therefore, be monitored closely alongside more widely documented [O_2] decreases within the OMZs (e.g., Stramma et al., 2008) to better constrain the anthropogenic fingerprint of climate change in the oceans. Models predict that ocean deoxygenation at mid- to high latitudes will continue to intensify under global warming conditions, such that climate-driven perturbation to oceanic oxygen will become more distinct with time. Therefore, on-going observational efforts from time-series, repeat hydrographic sections and global in-situ profiling floats (e.g., Argo; Gruber et al., 2010) are crucial to better understanding natural variability in marine O_2 on multidecadal timescales and improving the detectability of emergent anthropogenic trends.

Edited by: B. Dewitte

References

Allen, M. R. and Stott, P. A.: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory, Clim. Dynam., 21, 477–491, 2003.

Allen, M. R. and Tett, S. F. B.: Checking for model consistency in optimal fingerprinting, Clim. Dynam., 15, 419–434, 1999.

Banks, H. T. and Bindoff, N. L.: Comparison of Observed Temperature and Salinity Changes in the Indo-Pacific with Results from the Coupled Climate Model HadCM3: Processes and Mechanisms, J. Clim., 16, 156–166, doi10.1175/1520-0442(2003)016¡0156:COOTAS¿2.0.CO;2, 2003.

Banks, H. T. and Wood, R. A.: Where to look for anthropogenic climate change in the ocean, J. Clim., 15, 879–891, doi:10.1175/1520-0442(2002)015¡0879:WTLFAC¿2.0.CO;2, 2002.

Barnett, T. P., Pierce, D. W., AchutaRao, K. M., Gleckler, P. J., Santer, B. D., Gregory, J. M., and Washington, W. M.: Penetration of Human-Induced Warming into the World's Oceans, Science, 309, 284–287, doi:10.1126/science.1112418, 2005.

Bindoff, N. L. and Wunsch, C.: Comparison of Synoptic and Climatologically Mapped Sections in the South Pacific Ocean, J. Clim., 5, 631–645, 1992.

Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R., and Schwarzkopf, F. U.: The response of the Antarctic Circumpolar Current to recent climate change, Nat. Geosci., 1, 864–869, doi:10.1038/ngeo362, 2008.

Bopp, L., Le Quéré, C., Heimann, M., Manning, A. C., and Monfray, P.: Climate-induced oceanic oxygen fluxes: Implications for the contemporary budget, Glob. Biogeochem. Cy., 16, 1022, doi:10.1029/2001GB001445, 2002.

Brandt, P., Hormann, V., Körtzinger, A., Visbeck, M., Krahmann, G., and Stramma, L.: Changes in the Ventilation of the Oxygen Minimum Zone of the Tropical North Atlantic, J. Phys. Oceanogr., 40, 1784–1801, doi:10.1175/2010JPO4301.1, 2010.

Brennan, C. E., Matear, R. J., and Keller, K.: Measuring oxygen concentrations to improve the detection capabilities of an ocean circulation observation array, J. Geophys. Res., 113, C01019, doi:10.1029/2007JC004113, 2008.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-System model – HadGEM2, Geosci. Model Dev., 4, 1051–1075, doi:10.5194/gmd-4-1051-2011, 2011.

Deutsch, C., Emerson, S., and Thompson, L.: Fingerprints of climate change in North Pacific oxygen, Geophys. Res. Lett., 32, L16604, doi:10.1029/2005GL023190, 2005.

Deutsch, C., Emerson, S., and Thompson, L.: Physical-biological interactions in North Pacific oxygen variability, J. Geophys. Res., 111, C09S90, doi:10.1029/2005JC003179, 2006.

Doney, S. C., Ruckelshaus, M., Duffy, J. E., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J. M., Hollowed, A. B., Knowlton, N., Polovina. J., Rabalais, N. N., Sydeman, W. J., and Talley, L. D.: Climate change impacts on marine ecosystems, Ann. Rev. Mar. Sci., 4, 4.1–4.27, doi:10.1146/annurev-marine-041911-111611, 2012.

Downes, S. M., Bindoff, N. L., and Rintoul, S. R.: Changes in the Subduction of Southern Ocean Water Masses at the End of the Twenty-First Century in Eight IPCC Models, J. Clim., 23, 6526–6541, doi:10.1175/2010JCLI3620.1, 2010.

Duteil, O. and Oschlies, A.: Sensitivity of simulated extent and future evolution of marine suboxia to mixing intensity, Geophys. Res. Lett., 38, L06607, doi:10.1029/2011GL046877, 2011.

Frölicher, T. L., Joos, F., Plattner, G.-K., Steinacher, M., and Doney, S. C.: Natural variability and anthropogenic trends in oceanic oxygen in a coupled carbon-climate model ensemble, Glob. Biogeochem. Cy., 23, GB1003, doi:10.1029/2008GB003316, 2009.

Garcia, H. E., Boyer, T. P., Levitus, S., Locarnini, R. A., and Antonov, J.: On the variability of dissolved oxygen and apparent oxygen utilization content for the upper world ocean: 1955 to 1998, Geophys. Res. Lett., 32, L09604, doi:10.1029/2004GL022286, 2005.

Gilbert, D., Rabalais, N. N., Díaz, R. J., and Zhang, J.: Evidence for greater oxygen decline rates in the coastal ocean than in the open ocean, Biogeosciences, 7, 2283–2296, doi:10.5194/bg-7-2283-2010, 2010.

Giorgetta, M. A., Jungclaus, J. H., Reick, C. H., Legutke, S., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W. A., Notz, D., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K., Stockhause, M., Wegner, J., Widmann, H., Wieners, K.-H., Claussen, M., Marotzke, J., and Stevens, B.: Climate change from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project 5, J. Adv. Model. Earth Syst., submitted, 2012.

Gleckler, P. J., Santer, B. D., Domingues, C. M., Pierce, D. W., Barnett, T. P., Church, J. A., Taylor, K. E., AchutaRao, K. M., Boyer, T. P., Ishii, M., and Caldwell, P. M.: Human-induced global ocean warming on multidecadal timescales, Nature Clim. Change, 2, 524–529, doi:10.1038/nclimate1553, 2012.

Gnanadesikan, A., Russell, J. L., and Zeng, F.: How does ocean ventilation change under global warming?, Ocean Sci., 3, 43–53, doi:10.5194/os-3-43-2007, 2007.

Gnanadesikan, A., Dunne, J. P., and John, J.: Understanding why the volume of suboxic waters does not increase over centuries of global warming in an Earth System Model, Biogeosciences, 9, 1159–1172, doi:10.5194/bg-9-1159-2012, 2012.

Gruber, N.: Warming up, turning sour, losing breath: ocean biogeochemistry under global change, Phil. Trans. R. Soc. A, 369, 1980–1996, doi:10.1098/rsta.2011.0003, 2011.

Gruber, N., Doney, S. C., Emerson, S. R., Gilbert, D., Kobayashi, T., Körtzinger, A., Johnson, G. C., Johnson, K. S., Riser, S. C., and Ulloa, O.: Adding oxygen to Argo: Developing a global in-situ observatory for ocean deoxygenation and biogeochemistry, in: Proceedings of the "OceanObs09: Sustained Ocean Observations and Information for Society" Conference (Vol. 2), Venice, Italy, 21–25 September 2009, 1–10, 2010.

Hasselmann, K.: Multi-pattern fingerprint method for detection and attribution of climate change, Clim. Dynam., 13, 601–612, 1997.

Hegerl, G. C. and Zwiers, F.: Use of models in detection and attribution of climate change, WIREs Clim. Change, 2, 570–591, doi:10.1002/wcc.121, 2011.

Hegerl, G. C., Karl, T. R., Allen, M., Bindoff, N. L., Gillett, N., Karoly, D., Zhang, X. B., and Zwiers, F.: Climate change detection and attribution: Beyond mean temperature signals, J. Clim., 19, 5058–5077, doi:10.1175/JCLI3900.1, 2006.

Hegerl, G. C., Hoegh-Guldberg, O., Casassa, G., Hoerling, M. P., Kovats, R. S., Parmesan, C., Pierce, D. W., and Stott, P. A.: Good practice guidance paper on detection and attribution related to anthropogenic climate change, in: Meeting report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution Related to Anthropogenic Climate Change, 14–16 September 2009, IPCC Working Group I Technical Support Unit, University of Bern, Bern, Switzerland, 1–8, 2010.

Helm, K. P.: Decadal ocean water-mass changes: Global observations and interpretation, Ph.D. thesis, University of Tasmania, Australia, 2008.

Helm, K. P., Bindoff, N. L., and Church, J. A.: Changes in the global hydrological-cycle inferred from ocean salinity, Geophys. Res. Lett., 37, L18701, doi:10.1029/2010GL044222, 2010.

Helm, K. P., Bindoff, N. L., and Church, J. A.: Observed decreases in oxygen content of the global ocean, Geophys. Res. Lett., 38, L23602, doi:10.1029/2011GL049513, 2011.

Hirst, A. C., O'Farrell, S. P., and Gordon, H. B.: Comparison of a Coupled Ocean-Atmosphere Model with and without Oceanic Eddy-Induced Advection, Part I: Ocean Spinup and Control Integrations, J. Clim., 13, 139–163, doi:10.1175/1520-0442(2000)013¡0139:COACOA¿2.0.CO;2, 2000.

Hofmann, M. and Schellnhuber, H. J.: Ocean acidification affects marine carbon and triggers extended marine oxygen holes, Proc. Natl. Acad. Sci., 106, 3017–3022, doi:10.1073/pnas.0813384106, 2009.

IDAG (The International Ad Hoc Detection and Attribution Group): Detecting and Attributing External Influences on the Climate System: A Review of Recent Advances, J. Clim., 18, 1291–1314, 2005.

Ilyina T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez-Riboni, I.: The global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth System Model in different CMIP5 experimental realizations, J. Adv. Model. Earth Syst., accepted, 2013.

Ito, T. and Deutsch, C.: A conceptual model for the temporal spectrum of oceanic oxygen variability, Geophys. Res. Lett., 37, L03601, doi:10.1029/2009GL041595, 2010.

Jochum, M., Deser, C., and Philips, A.: Tropical Atmospheric Variability Forced by Oceanic Internal Variability, J. Clim., 20, 765–771, doi:10.1175/JCLI4044.1, 2007.

Johnson, G. C. and Gruber, N.: Decadal water mass variations along 20° W in the Northeastern Atlantic Ocean, Prog. Oceanogr., 73, 277–295, 2007.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S.,

Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geosci. Model Dev., 4, 543–570, doi:10.5194/gmd-4-543-2011, 2011.

Joos, F., Plattner, G.-K., Stocker, T. F., Körtzinger, A., and Wallace, D. W. R.: 2003. Trends in marine dissolved oxygen: implications for ocean circulation changes and carbon budget, Eos Trans. AGU, 84, 197–204, doi:10.1029/2003EO210001, 2003.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J.-S.: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth System Model, J. Adv. Model. Earth Syst., accepted, 2013.

Keeling, R. F., Körtzinger, A., and Gruber, N.: Ocean deoxygenation in a warming world, Ann. Rev. Mar. Sci., 2, 199–229, doi:10.1146/annurev.marine.010908.163855, 2010.

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., and Heimann, M.: Saturation of the Southern Ocean CO_2 sink due to recent climate change, Science, 316, 1735–1738, 2007.

Maier-Reimer, E.: Geochemical cycles in an ocean general circulation model: preindustrial tracer distributions, Glob. Biogeochem. Cy., 7, 645–677, 1993.

Maier-Reimer, E., Kriest, I., Segschneider, J., and Wetzel, P.: The HAMburg Ocean Carbon Cycle model HAMOCC5.1 – Technical description., Berichte zur Erdsystemforschung 14/2005, 14, Max-Planck-Institut für Meteorologie, Hamburg, Germany, available at: http://www.mpimet.mpg.de (last access: 12 May 2012), 2005.

Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Röske, F. R.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Model., 5, 91–127, 2003.

Masson, D. and Knutti, R.: Spatial-scale dependence of climate model performance in the CMIP3 ensemble, J. Clim., 24, 2680–2692, 2011.

Matear, R. J. and Hirst, A. C.: Long-term changes in dissolved oxygen concentrations in the ocean caused by protracted global warming, Glob. Biogeochem. Cy., 17, 1125, doi:10.1029/2002GB001997, 2003.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M. A., Haak, H., Jungclaus, J. H., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Syst., 4, M00A01, doi:10.1029/2012MS000154, 2012.

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., and Marotzke, J.: Arctic sea-ice evolution as modeled by MPI-ESM, J. Adv. Model. Earth Syst., accepted, 2013.

Oschlies, A., Schultz, K. G., Riebesell, U., and Schmittner, A.: Simulated 21st century's increase in oceanic suboxia by CO_2-enhanced biotic carbon export, Glob. Biogeochem. Cy., 22, GB4008, doi:10.1029/2007GB003147, 2008.

Ono, T., Midorikawa, T., Watanabe, Y. W., Tadokoro, K., and Saino, T.: Temporal increases of phosphate and apparent oxygen utilization in the subsurface waters of western subarctic Pacific from 1968 to 1998, Geophys. Res. Lett., 28, 3285–3288, 2001.

Palmer, J. R. and Totterdell, I. J.: Production and export in a global ocean ecosystem model, Deep-Sea Res. Pt. I, 48, 1169–1198, 2001.

Pierce, D. W., Barnett, T. P., AchutaRao, K. M., Gleckler, P. J., Gregory, J. M., and Washington, W. M.: Anthropogenic warming of the oceans: observations and model results, J. Clim., 19, 1873–1900, doi:10.1175/JCLI3723.1, 2006.

Rödenbeck, C., Le Quéré, C., Heimann, M., and Keeling, R. F.: Interannual variability in oceanic biogeochemical processes inferred by inversion of atmospheric O_2/N_2 and CO_2 data, Tellus B, 60, 685–705, 2008.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, J. Clim., 19, 3771–3791, 2006.

Sarmiento, J. L., Hughes, T. M. C., Stouffer, R. J., and Manabe, S.: Simulated response of the ocean carbon cycle to anthropogenic climate warming, Nature, 393, 245–249, 1998.

Schmittner, A., Oschlies, A., Matthews, H. D., and Galbraith, E. D.: Future changes in climate, ocean circulation, ecosystems, and biogeochemical cycling simulated for a business as-usual CO2 emission scenario until year 4000 AD, Glob. Biogeochem. Cy., GB1013, doi:10.1029/2007GB002953, 2008.

Séférian, R., Bopp, L., Gehlen, M., Orr, J., Ethé, C., Cadule, P., Aumont, O., Salas-y-Mélia, D., Voldoire, A., and Madec, G.: Skill assessment of three Earth system models with common marine biogeochemistry, Clim. Dynam., in press, doi:10.1007/s00382-012-1362-8, 2012.

Shaffer, G., Olsen, S. M., and Pedersen, J. O. P.: Long-term ocean oxygen depletion in response to carbon dioxide emissions from fossil fuels, Nat. Geosci., 2, 105–109, 2009.

Six, K. D. and Maier-Reimer, E.: Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model, Glob. Biogeochem. Cy., 10, 559–583, 1996.

Stevens, B., Giorgetta, M. A., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E.: The atmospheric component of the MPI-M Earth System Model: ECHAM6, J. Adv. Model. Earth. Syst., accepted, 2013.

Stott, P. A., Allen, M. R., and Jones, G. S.: Estimating signal amplitudes in optimal fingerprinting. Part II: application to general circulation models, Clim. Dynam., 21, 493–500, doi:10.1007/s00382-003-0314-8, 2003.

Stott, P. A., Sutton, R. T., and Smith, D. M.: Detection and attribution of Atlantic salinity changes, Geophys. Res. Lett., 35, L21702, doi:10.1029/2008GL035874, 2008.

Stramma, L., Johnson, G. C., Sprintall, J., and Mohrholz, V.: Expanding oxygen-minimum zones in the tropical oceans, Science, 320, 655–658, 2008.

Stramma, L., Johnson, G. C., Firing, E., and Schmidtko, S.: Eastern Pacific oxygen minimum zones: Supply paths and multidecadal changes, J. Geophys. Res., 115, C09011, doi:10.1029/2009JC005976, 2010.

Stramma, L., Oschlies, A., and Schmidtko, S.: Anticorrelated observed and modelled trends in dissolved oceanic oxygen over the last 50 years, Biogeosciences Discuss., 9, 4595–4626, doi:10.5194/bgd-9-4595-2012, 2012a.

Stramma, L., Prince, E. D., Schmidtko, S., Luo, J., Hoolihan, J. P., Visbeck, M., Wallace, D. W. R., Brandt, P., and Körtzinger, A.: Expansion of oxygen minimum zones may reduce available habitat for tropical pelagic fishes, Nature Clim. Change, 2, 33–37, doi:10.1038/nclimate1304, 2012b.

Swanson, K. L., Sugihara, G., and Tsonis, A. A.: Long-term natural variability and 20th century climate change, Proc. Natl. Acad. Sci., 106, 16111–16123, doi:10.1073/pnas.0908699106, 2009.

Tagliabue, A., Bopp, L., and Gehlen, M.: The response of marine carbon and nutrient cycles to ocean acidification: Large uncertainties related to phytoplankton physiological assumptions, Glob. Biogeochem. Cy., 25, GB3017, doi:10.1029/2010GB003929, 2011.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: L A Summary of the CMIP5 Experiment Design, available at: http://www-pcmdi.llnl.gov/ (last access: 12 May 2012), 2009.

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

Terray, L., Corre, L., Cravatte, S., Delcroix, T., Reverdin, G., and Ribes, A.: Near-surface salinity as Nature's rain gauge to detect human influence on the tropical water cycle, J. Climate, 25, 958–977, doi:10.1175/JCLI-D-10-05025.1, 2012.

The HadGEM2 Development Team: Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model Climate configurations, Geosci. Model Dev. Discuss., 4, 765–841, doi:10.5194/gmdd-4-765-2011, 2011.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, Cambridge, 1999.

Whitney, F. A, Freeland, H. J., and Robert, M.: Persistently declining oxygen levels in the interior waters of the eastern subarctic Pacific, Prog. Oceanogr., 75, 179–199, 2007.