JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002126

The Evaluation of the North Atlantic Climate System in UKESM1 Historical Simulations for CMIP6

Jon Robson[1], Yevgeny Aksenov[2], Thomas J. Bracegirdle[3], Oscar Dimdore-Miles[4], Paul T. Griffiths[5], Daniel P. Grosvenor[6], Daniel L. R. Hodson[1], James Keeble[5], Claire MacIntosh[7], Alex Megann[2], Scott Osprey[8], Adam C. Povey[9], David Schröder[10], Mingxi Yang[11], Alexander T. Archibald[5], Ken S. Carslaw[12], Lesley Gray[8], Colin Jones[6], Brian Kerridge[13], Diane Knappett[13], Till Kuhlbrodt[1], Maria Russo[5], Alistair Sellar[14], Richard Siddans[13], Bablu Sinha[2], Rowan Sutton[1], Jeremy Walton[14], and Laura J. Wilcox[1]

[1]National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK, [2]National Oceanography Centre, Southampton, UK, [3]British Antarctic Survey, Cambridge, UK, [4]Department of Physics, University of Oxford, Oxford, UK, [5]National Centre for Atmospheric Science, Department of Chemistry, University of Cambridge, Cambridge, UK, [6]National Centre for Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK, [7]National Centre for Earth Observation, Department of Meteorology, University of Reading, Reading, UK, [8]National Centre for Atmospheric Science, Department of Physics, University of Oxford, UK, [9]National Centre for Earth Observation, Department of Physics, University of Oxford, UK, [10]Centre for Polar Observations and Modelling, Department of Meteorology, University of Reading, Reading, UK, [11]Plymouth Marine Laboratory, Plymouth, UK, [12]School of Earth and Environment, University of Leeds, Leeds, UK, [13]Remote Sensing Group, National Centre for Earth Observation, STFC Rutherford Appleton Laboratory, Chilton, UK, [14]Met Office Hadley Centre, Exeter, UK

Key Points:
- UKESM1 performs well across a range of variables and components of the North Atlantic climate system when compared with observations
- A range of shortcomings are still present in the simulations, however, with many shortcomings similar to those of the physical climate model
- Development of the physical model, alongside the Earth system components, is crucial for improved climate simulations

Supporting Information:
- Supporting Information S1

Correspondence to:
J. Robson,
j.i.robson@reading.ac.uk

Citation:
Robson, J., Aksenov, Y., Bracegirdle, T. J., Dimdore-Miles, O., Griffiths, P. T., Grosvenor, D. P., et al. (2020). The evaluation of the North Atlantic climate system in UKESM1 historical simulations for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002126. https://doi.org/10.1029/2020MS002126

Received 30 MAR 2020
Accepted 5 AUG 2020
Accepted article online 10 AUG 2020

©2020. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract Earth system models enable a broad range of climate interactions that physical climate models are unable to simulate. However, the extent to which adding Earth system components changes or improves the simulation of the physical climate is not well understood. Here we present a broad multivariate evaluation of the North Atlantic climate system in historical simulations of the UK Earth System Model (UKESM1) performed for CMIP6. In particular, we focus on the mean state and the decadal time scale evolution of important variables that span the North Atlantic climate system. In general, UKESM1 performs well and realistically simulates many aspects of the North Atlantic climate system. Like the physical version of the model, we find that changes in external forcing, and particularly aerosol forcing, are an important driver of multidecadal change in UKESM1, especially for Atlantic Multidecadal Variability and the Atlantic Meridional Overturning Circulation. However, many of the shortcomings identified are similar to common biases found in physical climate models, including the physical climate model that underpins UKESM1. For example, the summer jet is too weak and too far poleward; decadal variability in the winter jet is underestimated; intraseasonal stratospheric polar vortex variability is poorly represented; and Arctic sea ice is too thick. Forced shortwave changes may be also too strong in UKESM1, which, given the important role of historical aerosol forcing in shaping the evolution of the North Atlantic in UKESM1, motivates further investigation. Therefore, physical model development, alongside Earth system development, remains crucial in order to improve climate simulations.

Plain Language Summary The North Atlantic climate system plays an important role in regulating Earth's climate, and variability within the Atlantic can have important impacts on society. However, we do not understand all the linkages between different parts of the North Atlantic. Furthermore, climate simulations, which are an essential tool for improving our understanding, have shortcomings that can affect their utility. New developments in Earth System climate simulations could remedy these shortcomings. However, we do not understand the extent to which the addition of complex Earth system developments have changed or improved the simulation of the physical climate. Therefore, in this paper, we present a multidisciplinary evaluation of the North Atlantic climate in simulations made with the UK's Earth System Model, UKESM1. We find that simulations made with UKESM1 capture many aspects of the North Atlantic climate and that human activities have a large impact on the North Atlantic in UKESM1. Nevertheless, we also highlight some shortcomings of the simulations, many of which are like those seen in physical climate simulations. Therefore, further development of both the physical and Earth system components is needed to improve climate simulations in the future.

1. Introduction

1.1. Motivation

The North Atlantic region plays a key role in the Earth System. In the current climate the North Atlantic Ocean takes up considerable amounts of heat and carbon from the atmosphere. Hence it is a first-order regulator of global climate (Stocker et al., 2013). The subpolar North Atlantic ocean has a crucial role for the formation and maintenance of the Atlantic meridional overturning circulation (AMOC, Buckley & Marshall, 2016; Kuhlbrodt et al., 2007) and its subsequent large northward heat transport (Johns et al., 2011). This large heat transport results in the Atlantic being the only ocean basin with substantial cross-equator ocean transports (Marshall et al., 2014). Changes in atmospheric composition and chemistry across the North Atlantic, including tropospheric ozone, methane, and aerosols, can have a large impact on regional air quality (Monks et al., 2015) and climate through their radiative forcings (Booth et al., 2012; Wilcox et al., 2015). The North Atlantic climate system is also the location of a strong and variable atmospheric jet stream that can have significant impact over land (Woollings et al., 2014), and the North Atlantic can have significant impact on the Arctic. Variability in the atmospheric circulation, closely linked to variability over the Atlantic, plays an important role in driving the Arctic Ocean circulation (Inoue & Kikuchi, 2007), and the subpolar North Atlantic is the gateway through which the Atlantic and Arctic oceans interact (Holliday et al., 2018). The North Atlantic basin is also surrounded by large regions of high human population density in Europe, North Africa, as well as North and Central America. Therefore, changes in the North Atlantic climate system can have substantial societal impacts on a wide range of sectors and locations (Monerie et al., 2019; Scaife et al., 2008; Smith et al., 2010; Sutton & Hodson, 2005).

1.2. Variability and Coupled Interactions

The North Atlantic climate system has varied significantly across a range of variables, and on a range of time scales (Robson et al., 2018). For example, on intra-annual and interannual time scales there is substantial variability in the basin mean latitude and speed of the atmospheric jet, and the North Atlantic Oscillation (NAO) (Hurrell, 1995; Woollings et al., 2014, 2015). In winter, the jet/NAO variability is often linked to significant intra-annual and interannual variability in the stratospheric polar vortex (Baldwin & Dunkerton, 2001; Mitchell et al., 2013). The strength of the AMOC (Smeed et al., 2018), sea surface temperatures (Josey et al., 2018), and sea ice extent (Swart et al., 2015) have also been shown to vary significantly on interannual time scales.

There is also significant variability on decadal-to-multidecadal time scales across a range of variables. Multidecadal variability in North Atlantic sea surface temperatures (SSTs), where North Atlantic SSTs warmed or cooled relative to global mean changes, is particularly prominent over the observed period and has become known as Atlantic multidecadal variability (AMV, Sutton et al., 2018), or the Atlantic multidecadal oscillation (AMO Kerr, 2000; however, note that the term oscillation implies a preferred frequency, and hence we use the more generic AMV here). However, substantial decadal time scale variability has also been observed in other variables, including, but not limited to, the NAO and jet speed (Hurrell, 1995; Woollings et al., 2015), sea ice extent (Swart et al., 2015), and ocean transports (Smeed et al., 2018). Decadal time scale variability is especially clear within the subpolar North Atlantic, with large changes observed in both surface properties and upper ocean temperature and salinity anomalies (Reverdin, 2010; Robson et al., 2016, 2018; Ruiz-Barradas et al., 2018). Large shifts in the strength and location of surface currents have also been observed in the subpolar North Atlantic in recent decades (Chafik et al., 2019; Häkkinen & Rhines, 2004; Holliday et al., 2020; Sarafanov et al., 2008). Long-term trends are also present in many variables across the North Atlantic region, including surface temperatures in the atmosphere and ocean; sea ice; and in atmospheric composition, such as surface ozone (Stocker et al., 2013).

The variability and change across the different variables and components of the North Atlantic climate system does not occur in isolation; there are a myriad of linkages and interactions on different temporal and spatial scales. Changes in the NAO, the North Atlantic jet, and atmospheric blocking are well known to drive significant variability across the extratropical and subpolar North Atlantic ocean, including in SSTs, heat content, and sea ice extent and thickness through changes in surface fluxes and mechanical wind forcing (Häkkinen et al., 2011; Lozier et al., 2008; Moat et al., 2019; Robson et al., 2012; Visbeck et al., 1998).

A key hypothesis for the existence of the AMV is a lagged response of the AMOC, and its related heat transport, to low-frequency NAO forcing (O'Reilly et al., 2019; Sutton et al., 2018). However, it is widely recognized that AMV involves a complex array of local and regional interactions which are different between the subpolar and tropical regions. In the tropical North Atlantic, processes including changes in wind, evaporation, and clouds (Bellomo et al., 2016; Brown et al., 2016; Martin et al., 2014; Middlemas et al., 2019; Myers & Mechoso, 2020; Sutton et al., 2018), as well as local radiative forcing due to atmospheric composition changes (Booth et al., 2012; Yuan et al., 2016), are thought to be particularly important. In the subpolar region many studies have identified AMOC variability as a key driver of low-frequency SST and heat content anomalies (Delworth et al., 2017; Robson et al., 2012, 2016; Yeager et al., 2012) as well as long-term externally forced change (Caesar et al., 2018; Drijfhout et al., 2012). However, evidence also suggests AMOC variability can also play an important role in the tropical and subtropical North Atlantic (Alexander-Turner et al., 2018; Cunningham et al., 2013; Duchez et al., 2016). Moreover, a wide array of anomalies in tropical and extratropical North Atlantic SSTs, sea ice, and resulting surface flux anomalies, whether they are related to AMV or not, are also thought to exert a significant influence on the atmosphere. For example, SST and sea ice anomalies and have been linked to significant shifts in the NAO/jet on seasonal, decadal, and multidecadal time scales in both summer and winter (Deser et al., 2010; Dong et al., 2013; Omrani et al., 2016; Screen, 2017; Woollings et al., 2012). Stratospheric variability, including the polar vortex and its interactions with the quasi-biennial oscillation, can also modulate the tropospheric circulation variability in winter (Andrews et al., 2019; Baldwin & Dunkerton, 2001; Mitchell et al., 2013) and may play a crucial aspect in modulating the impact of AMV-related SSTs on the atmosphere (Omrani et al., 2014, 2016).

Changes in the atmospheric circulation are also known to affect atmospheric composition, for example, regional patterns of tropospheric ozone and dust (Pausata et al., 2012; Robson et al., 2018; Yuan et al., 2016). The correct simulation of clouds and aerosol changes are also key to simulate aerosol-cloud interactions and their impact on the North Atlantic (Allen et al., 2015; Booth et al., 2012; Chang et al., 2011). Furthermore, long-term changes in atmospheric composition, both gas-phase and aerosols, caused by human influences and natural forcing changes, such as volcanic and solar change, can drive a wide range of changes across the North Atlantic climate system (Caesar et al., 2018; Gray et al., 2013, 2016; Menary et al., 2013; Otterå et al., 2010; Stocker et al., 2013).

1.3. Climate Simulations and Earth System Models

In order to provide the best climate information possible it is important to understand climate variability and change across the North Atlantic. To be able to provide the best estimate of potential future climate risks, climate models also need to simulate the relevant physical and Earth System processes and the interaction between them. However, there remain significant questions on the relative importance of different linkages between components. For example, which anomalies (i.e., what variables, spatial locations, or time scales) are the most important for driving interactions, and how do forced changes and internal variability in different components of the climate system interact? Unfortunately, due to the short observational record, and the large range of interactions on different spatial and temporal scales, untangling the complex array of interactions is a challenging task. Therefore, models are an essential tool in unraveling this complexity and, particularly, for understanding the relative importance of different processes in controlling past change (i.e., attribution), to understand how the climate system will change in future (i.e., predictions and projections), and to understand whether the linkages between different components have or will change.

However, the North Atlantic remains a region where significant challenges exist in providing robust climate simulations. For example, the North Atlantic is a region that has been susceptible to long-standing and pervasive biases in coupled climate models, and this remains the case for modern models especially when using ocean configurations of order 100 km (Kuhlbrodt et al., 2018; Roberts et al., 2019). As the North Atlantic is also a highly coupled region, biases in one variable or component can have important implications elsewhere. For example, SST biases in the gulf stream region can have a large impact on errors in the atmospheric jet (Keeley et al., 2012). Large uncertainty remains in the simulation of the AMOC (Reintges et al., 2017) and AMV (Martin et al., 2014; Menary et al., 2018). Furthermore, recent decades have seen large swings in the number of sudden stratospheric warmings events (Cohen et al., 2009; Gillett et al., 2002), which are known to have impacts on North Atlantic seasonal weather (Baldwin et al., 2020). The distinction between stratospheric forcing and response with other modes of tropospheric and ocean

variability is currently an outstanding question (Omrani et al., 2014). There is also considerable debate on the importance of the forced response of the North Atlantic, especially to anthropogenic sulfate aerosols (Bellomo et al., 2018; Booth et al., 2012; Kim, Yeager & Danabasoglu 2018; R. Zhang et al., 2013).

In the past, due to the computational cost, many global climate models neglected first-order Earth system processes, for example, interactive atmospheric chemistry or ocean biogeochemistry. The current generation of Earth system models remedies this by including and coupling a greater number of Earth system processes (e.g., Séférian et al., 2019; Sellar et al., 2019). Although these processes may have the largest impact when exploring a range of future climate states, there is reason to expect that aspects of the current climate system may be different, or indeed improved, when these Earth system components are included. For example, interactive ozone can provide an important feedback within the polar vortex, leading to a strengthened polar vortex and more robust connections between vortex disturbances (e.g., sudden stratospheric warmings) and the midwinter North Atlantic circulation (Haase & Matthes, 2019; Oehrlein et al., 2020). Changes in tropospheric ozone and methane affect the oxidizing capacity of the atmosphere (Becker et al., 1999), and hence secondary aerosol formation (Rae et al., 2007). It follows that models which simulate chemical oxidants interactively could, therefore, have different changes in sulfate aerosols, cloud properties, and, hence, radiative forcing. Other changes, including interactive dimethyl sulfide emissions, biogenic emissions from vegetation, and interactive vegetation (Sellar et al., 2019) may also have large impacts on the North Atlantic through their impact on aerosol radiative forcing (Boucher et al., 2003; Yuan et al., 2016). Therefore, at present there is the opportunity and the need to evaluate the North Atlantic climate system in complex Earth system models that explicitly simulate a wide range of interactions between climate components.

1.4. Objectives

In this paper we undertake a multidisciplinary and multivariate evaluation of the North Atlantic climate system in the UK's Earth System Model, UKESM1 (Sellar et al., 2019). In part, we address the question of how the simulation of the North Atlantic climate system is affected by the inclusion of Earth System components. Specifically, we analyze key variables across a range of components of the North Atlantic climate system, including: atmospheric circulation; atmospheric composition and aerosols; clouds and radiation; ocean state and circulation; and Arctic-Atlantic exchange and sea ice. In order to compare to a diverse set of observations, we focus on the historical simulations performed for the sixth climate model inter-comparison project (CMIP6). We prioritize a broad holistic analysis of evaluating mean state and changes on the decadal time scale change across a wide range of Earth System variables.

Many aspects of the North Atlantic climate system have been evaluated extensively in the latest UK physical climate model, HadGEM3-GC3.1 (Andrews et al., 2020; Kuhlbrodt et al., 2018; Menary et al., 2018; Mulcahy et al., 2018; Williams et al., 2018). Therefore, here we focus our analysis exclusively on UKESM1 and compare with the published examples in those papers. However, we also take the opportunity to explore a wider basket of important variables including those that cannot be evaluated in the physical model in order to provide a more holistic overview. These variables include atmospheric composition, radiative properties, sea ice, and Arctic-Atlantic exchange that have not been evaluated in detail before. Where relevant, we will also compare with the wider literature. Detailed model intercomparisons, evaluation of other time scales, and detailed process-based evaluations of any one variable are deliberately left for future work.

The paper is organized as follows. Section 2 provides a brief description of the model and observational data used. Results are described in section 3, before a brief discussion and a summary of the key conclusions in sections 4 and 5, respectively.

2. Model and Data

2.1. UKESM1 Configuration and Simulations

The UK Earth System model (UKESM1) was jointly developed by the UK Met Office and the Natural Environment Research Council (NERC) as the successor to the HadGEM2-ES model (Collins et al., 2011). For a full description of UKESM1 and its development we refer the reader to Sellar et al. (2019), but for completeness we provide a brief description here.

UKESM1 is based on the low-resolution version of the UK Met Office's physical coupled climate model, HadGEM3-GC3.1 (known as HadGEM3-GC3.1-LL in CMIP6 nomenclature, but hereafter GC3.1,

Kuhlbrodt et al., 2018) with additional Earth system components included. GC3.1 is a coupled atmosphere-land-ocean-sea ice model. The atmospheric component is the Unified Model GA7.1 configuration at N96 horizontal resolution (~135 km in the extratropics) with 85 vertical levels up to a model lid at 85 km (35 levels are above 18 km) (Walters et al., 2019). The ocean component is based on the NEMO ocean model in the GO6.0 configuration at 1° resolution with 75 levels (Kuhlbrodt et al., 2018; Storkey et al., 2018). GC3.1 uses the CICE model for sea ice (GSI8.1, Ridley et al., 2018) and the JULES model for land surface processes (GL7.0, Walters et al., 2019). More details about the development of the physical core of UKESM1 (i.e., GC3.1) is given in Kuhlbrodt et al. (2018). In UKESM1, terrestrial biogeochemistry is added through JULES, with developments to plant physiology, land use changes, and the addition of nitrogen cycling (Harper et al., 2016, 2018); Ocean biochemistry is simulated through the addition of the MEDUSA model (Yool et al., 2013); and unified stratospheric-tropospheric chemistry is simulated with the addition of the UKCA model which represents the interaction between chemistry and aerosol processes (Archibald et al., 2019). In both GC3.1 and UKESM1 aerosol properties are simulated using the two-moment pseudo-modal aerosol scheme, GLOMAP-mode (Mann et al., 2010). However, oxidants used in secondary aerosol formation are interactively simulated in UKESM1 and coupled to GLOMAP-mode, whereas in GC3.1 they are prescribed as time invariant climatologies.

Here we focus on the evaluation of the CMIP6 historical simulations which cover the period 1850–2014 (Eyring et al., 2016). These simulations include the forcing due to anthropogenic changes in atmospheric composition (i.e., well-mixed greenhouse gases and aerosols) and land use changes, and through changes in natural forcing factors, such as changes in solar insolation and volcanic eruptions. We use the same nine-member ensemble of historical simulations as reported in Sellar et al. (2019). These nine historical simulations were all branched from different points of the preindustrial control simulation (henceforth, PiControl) in order to sample a range of internal variability states (see Sellar et al. (2019) for more details). We will focus on the analysis of the simulation of ensemble mean and ensemble spread of various variables in the historical simulations and compare with available observations described in section 2.2. We will also provide an estimate of the forced response by comparing with the variability simulated in the UKESM1 PiControl simulation (i.e., where there are no changes in external forcing factors).

2.2. Diagnostics and Observational Data

2.2.1. Atmospheric Data and Diagnostics of NAO, Jet Speed, and Jet Latitude

To evaluate atmospheric variables we compare UKESM1 to atmospheric reanalyses. Stratospheric variables are compared with ERA-Interim (Dee et al., 2011), which includes observationally informed assimilated variables in the upper atmosphere. However, a longer comparison can be made for lower atmospheric fields using two longer-term reanalysis data sets: NOAA/CIRES Twentieth Century Global Reanalysis Version 2c (20CRv2c) (Compo et al., 2011) and ECMWF ERA-20C (Poli et al., 2016). Both reanalyses use a restricted range of observational data types to help avoid issues relating to the sudden introduction of new types of data, such as is commonly found following the introduction of routine satellite observations (circa 1978). The 20CRv2c extends back to 1850, but the first decade is discarded due to model spinup effects, so 1861 is taken as the starting year in our study. ERA-20C spans the period 1900 to present day.

The NAO index used here is based on empirical orthogonal functions (EOF) following Hurrell (1995). The leading EOF of seasonal (winter and summer) mean PMSL anomalies is calculated over the NA region (20–80°N, 90°W to 40°E). This is then used to generate the principal component time series over each reanalysis or UKESM1 ensemble member. The NA tropospheric eddy driven jet is evaluated using indices of jet latitude and speed that closely follow the definitions of Woollings et al. (2015). They are based on zonal wind on the 850 hPa pressure level (UA850). At each latitude between 15°N and 75°N the longitudinal mean between 60°W and 0°W is calculated for each of the monthly mean fields. The maximum value across those latitudes is then defined as the jet speed and the latitude of this maximum is the jet latitude. Since monthly mean data were used here, rather than the more standard daily mean, the notation was chosen to reflect this. For example, the jet latitude index is denoted as "JLI_{mon}" rather than "JLI". Note that Bracegirdle et al. (2018) was also largely based on monthly data due to the unavailability of daily fields in CMIP5.

2.2.2. Atmospheric Composition and Aerosols

We use satellite observations to evaluate annual mean tropospheric column ozone and total column methane. Tropospheric column ozone observations is taken from (OMI/MLS, Ziemke et al., 2006). Total

column methane is taken from the atmospheric infrared sounder (AIRS) spectrometer, aboard the Earth observing system (EOS) platform Aqua (Xiong et al., 2008) and the infrared atmospheric sounding interferometer (IASI) aboard MetOp-A (Siddans et al., 2017), respectively. To facilitate the methane comparison, the model methane total column has been calculated by multiplying the methane mass mixing ratio by the mass of air and then dividing by surface area; IASI data have been converted from column average volume mixing ratio to column using an annual time series of average surface pressure for the North Atlantic domain; the AIRS data are already in column units.

Due to the uncertainty in observations of aerosol optical depths (AODs), we have evaluated the simulated AOD in UKESM1 against a range of satellite products. These include Dark Target, which evaluated the MODerate Imaging Spectroradiometer (MODIS) sensor on both the Terra and Aqua platforms (Levy et al., 2013); the ATSR-2 and AATSR sensors (along-track scanning radiometer and advanced ATSR, respectively), which were evaluated by ATSR-Dual View (Kolmonen et al., 2016); the Optimal Retrieval of Aerosol and Cloud (Thomas et al., 2009); and Swansea University's algorithm (Bevan et al., 2012). We have included all eight data sets in the evaluation, which can be downloaded from this site (https://search.earthdata.nasa.gov/search and https://cci.esa.int/data).

2.2.3. Clouds and Radiation

Cloud droplet concentrations (N_d) are compared with observations from MODIS (Salomonson et al., 1989). The MODIS data set is at $1° \times 1°$ resolution and is calculated from retrievals of cloud optical depth (τ_c) and effective radius (r_e) (Nakajima & King, 1990). Here we use the data set described in Grosvenor, Sourdeval, Zuidema, et al. (2018), which is based on the methods from Grosvenor, Sourdeval, and Wood (2018) and Grosvenor and Wood (2014). The 3.7 μm r_e is used for the N_d calculations, which has been suggested to be less prone to errors due to cloud heterogeneity (Grosvenor, Sourdeval, Zuidema, et al., 2018; Zhang et al., 2012, 2016). To avoid retrieval biases, the data set excludes $1° \times 1°$ data points with mean solar zenith angle greater than 65° (Grosvenor & Wood, 2014). Mean cloud top heights greater than 3.2 km are also excluded in order to restrict the retrievals to the cloud types that are most suited for this retrieval (see Grosvenor, Sourdeval, Zuidema, et al., 2018). Two-dimensional fields of in-cloud N_d are produced by the retrieval since it is assumed that N_d is constant throughout the depth of the cloud, which has been shown to be a good approximation by aircraft observations of stratocumulus (Painemal & Zuidema, 2011). N_d values (for the satellite and model) are here always defined as in-cloud values since N_d is undefined when there is no cloud. The model in-cloud N_d at a given model level is output as monthly means by weighting the data at each time step by the cloud fraction. Two-dimensional N_d data are obtained from the monthly mean 3-D model fields by calculating a vertical average that is weighted by the monthly mean cloud fraction on each level. This ensures that the vertical levels with the most cloud coverage contribute most to the vertical average. To match the sampling for MODIS the model data are also only sampled below 3.2 km in height.

Shortwave (SW) upwelling top-of-the-atmosphere (TOA) radiative fluxes are obtained from the monthly mean CERES-EBAF data product. This uses data from both the Terra and Aqua satellites as well as geostationary satellites to approximate averaging across the diurnal cycle, and the TOA net flux is constrained to the ocean heat storage (Kato et al., 2018; Loeb et al., 2018).

The cloud fraction data set used is based on the CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) satellite LIDAR instrument. Specifically, it is version 3 of the CALIPSO-GOCCP (GCM Oriented Cloud Calipso Product Chepfer et al., 2010). For comparisons of the model to this data set we use model output from the COSP (Cloud Feedback Model Intercomparison Project Observation Simulator Package Bodas-Salcedo et al., 2011) satellite simulator in order to get a fairer comparison between the model and satellite.

2.2.4. Ocean Data

We use ocean observations and reanalysis to evaluate the North Atlantic Ocean. We compare with SST from HadISST 1.1 (Rayner et al., 2003). Upper ocean (0–700 m) mean temperature and salinity, as well as surface salinity, is computed from EN4 (version 4.2.1 (Good et al., 2013), with XBT bias corrections applied (Gouretski & Reseghetti, 2010)). It should be noted that subsurface ocean observations were sparse before the year 2000 (i.e., before ARGO observations), which is reflected in the larger uncertainty before this time (see MacIntosh et al., 2017). AMOC is compared with an observed estimate at 26.5N from the RAPID array over the 2004–2014 period (Smeed et al., 2018), and ocean heat transport is compared with data from the

RAPID-MOCHA array (Johns et al., 2011). The standard observational mixed layer depth climatology is Montégut et al. (2004), but this climatology uses a diagnostic not available in UKESM1. Instead, March mixed layer depths were evaluated from March mean values of temperature and salinity from the EN4 data set using the mixed layer definition used in UKESM1; the depth (linearly interpolated between ocean levels) at which the density is 0.03 kg m^{-3} higher than that at the surface.

We also compute indices to represent Atlantic Multidecadal Variability (AMV). The basin mean AMV index (i.e., mean area-weighted SST anomaly over 0–60°N 7.5–75°W) is a common metric to quantify AMV (Sutton & Hodson, 2005). However, this simple metric has been criticized as it does not discriminate between Atlantic-only and global signals (Qasmi et al., 2017; Zhang et al., 2013). Hence, a fairer comparison between the model and observed AMV may be made if the global signal is removed (Zhang et al., 2013). Various methods have been proposed to do this, but we follow Moat et al. (2019) and Sutton and Dong (2012) to obtain AMVI—a smoothed (10 year running mean) and detrended AMV index:

$$AMVI = \overline{SST_{NAtl}} - \overline{SST_{ROTW}} \quad (1)$$

Where $\overline{SST_{ROTW}}$ is the SST averaged globally, except over the North Atlantic. These indices are then smoothed using a 10 year running mean and then standardized by dividing by their standard deviation before taking the difference.

2.2.5. Arctic-Atlantic Exchange and Sea Ice Data, and Forced Ocean-Sea Ice Simulations

For Arctic sea ice thickness we evaluate against Cryosat-2 satellite data (Laxon et al., 2013; Tilling et al., 2018) and PIOMAS (Zhang & Rothrock, 2003), and we use the SSM/I Bootstrap data to evaluate the sea ice edge (Comiso, 2017). Sea ice export through the Fram Strait is compared with existing observations (Kwok et al., 2004; Spreen et al., 2009; Vinje et al., 1998).

Estimates of Arctic-Atlantic exchanges of the volume and fresh water transports obtained from the mooring data vary substantially between different published sources. The key reasons for this are the large distance between the moorings, changes in instrumental layout from year-to-year, and absence of the observations on the Greenland continental shelf. The later requires the gap in the shelf transports to be filled with model estimates or extrapolate the data from nearest moorings (e.g., de Boer et al., 2018). The differences in the interpolation procedures can also introduce a large uncertainty in transport calculations. To address this uncertainty the volume and freshwater transports through Fram Strait are assessed using observations from an array of moorings presented by de Steur et al. (2018). We make use of both the reduced array, covering the outer Greenland shelf up to 6.5°W, and the full array, which includes the shelf up to 8°W.

We also include results from three forced ocean-sea ice simulations. All of them use the same ocean model NEMO GO6.0 (Storkey et al., 2018) and sea ice model CICE configuration GSI8.1 (Ridley et al., 2018) as UKESM1. This allows us to study the impact of the interaction with the atmosphere as well as the impact of changing sea ice physics and atmospheric forcing. CORE-NEMO-CICE-default is forced with CORE II atmospheric reanalysis data for the period 1950 to 2009 (Large & Yeager, 2009). The second simulation CORE-NEMO-CICE-CPOM uses sea ice settings suggested by Schröder et al. (2019) optimized based on a comparison of a stand-alone sea ice simulation with sea ice thickness estimates from Cryosat-2. The new settings include the elastic anisotropic plastic rheology (Tsamados et al., 2013), the bubbly conductivity formulation from Pringle et al. (2007), a simple scheme to account for the loss of drifting snow, increased longwave emissivity of sea ice from 0.95 to 0.976, and the maximum meltwater added to melt ponds, rfrac-max, is reduced from 100% to 50% (Schröder et al., 2019). The third simulation DFS-NEMO-CICE-CPOM uses the same sea ice setup as the second, but the atmospheric forcing data set DFS5.2 (Dussin et al., 2016) instead of CORE II.

3. Results

3.1. Atmospheric Circulation

3.1.1. NAO and Tropospheric Eddy-Driven Jet

Figure 1 shows the mean leading EOF of atmospheric pressure at mean sea level (PMSL) for the 20CRv2c and UKESM1 historical ensembles. During December-January-February (DJF) a broad correspondence is seen between the leading pattern of variability of both the reanalysis and UKESM1. This broad agreement

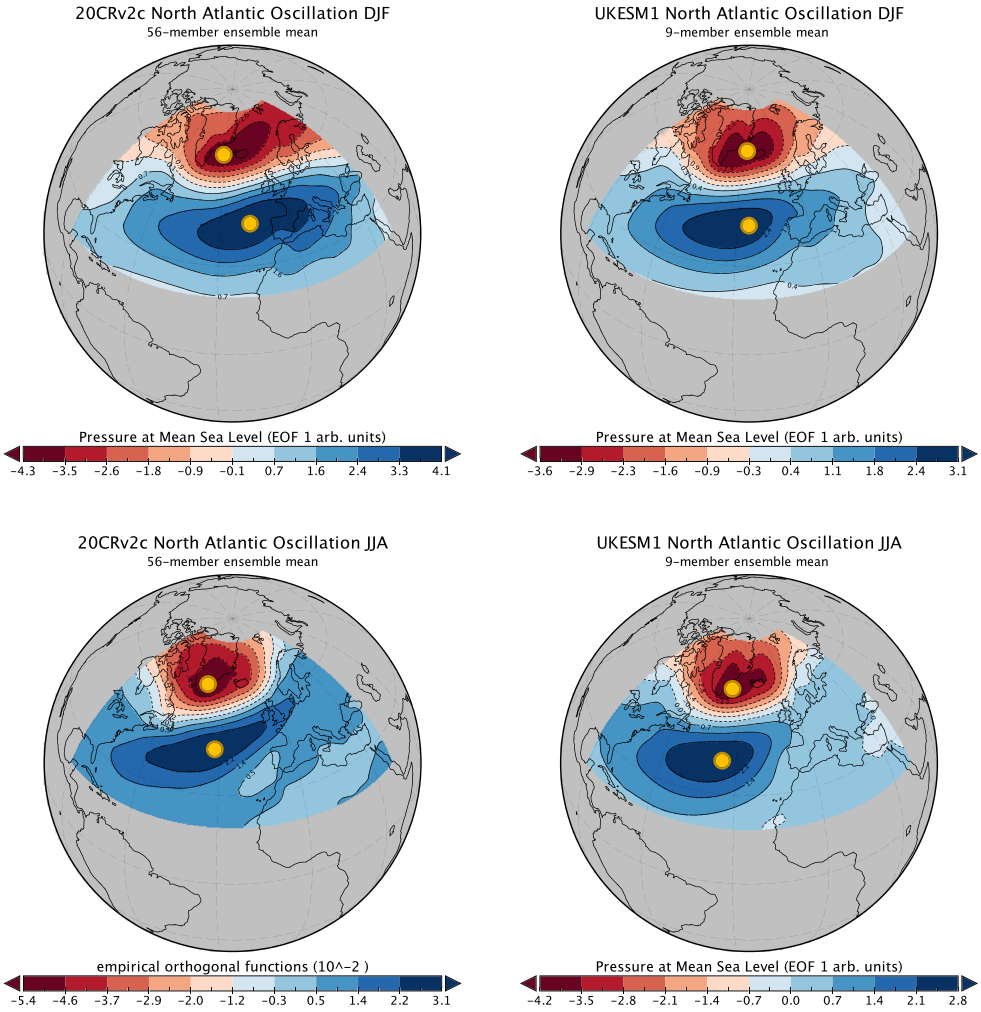


Figure 1. Leading empirical orthogonal function of December-February (top) and June-August (bottom) pressure at mean sea level (PMSL) for the 56-member ensemble mean 20CRv2c reanalysis (left) and nine-member ensemble mean historical UKESM1 simulations (1850–2014, right). Circles denote positions of the "centers of action" (e.g., the extrema of the EOF pattern). Note, the EOF patterns are not scaled by their variance so the units are arbitrary and magnitudes are not comparable between panels.

is consistent with Menary et al. (2018) who state that the NAO in GC3.1 is similar to that observed. However, differences are seen between reanalysis and UKESM1 and in the eastward extent of the pattern, which reaches western Europe in the reanalyses but not UKESM1. Furthermore, the contours are more zonally oriented within UKESM1; this would be consistent with a more zonal extension of the NA storm track. Similar characteristics are seen during June-July-August (JJA), with zonally flat contours and limited eastward extension into continental Europe within UKESM1. Another noticeable difference is the degree of pattern similarity between DJF and JJA in UKESM1 but not in the reanalyses. Note that the EOF patterns computed from each member are very similar, especially for DJF (see Figures S1 and S2 in the supporting information).

Figure 2 shows time series of the NAO and the Atlantic eddy-driven jet latitude and speed for 1850–2014. In winter (DJF) there is a general consistency between UKESM1 and reanalyses; the mean state of the jet latitude and speed is well represented, and the NAO index is largely within the ensemble spread (see Figures 2a–2c). The ensemble mean changes do not capture the observed variability, with small correlations between ensemble mean UKESM1 and ensemble mean 20CRv2c time series providing little evidence of a forced response. However, the previously documented spectral peak in winter low-frequency variability

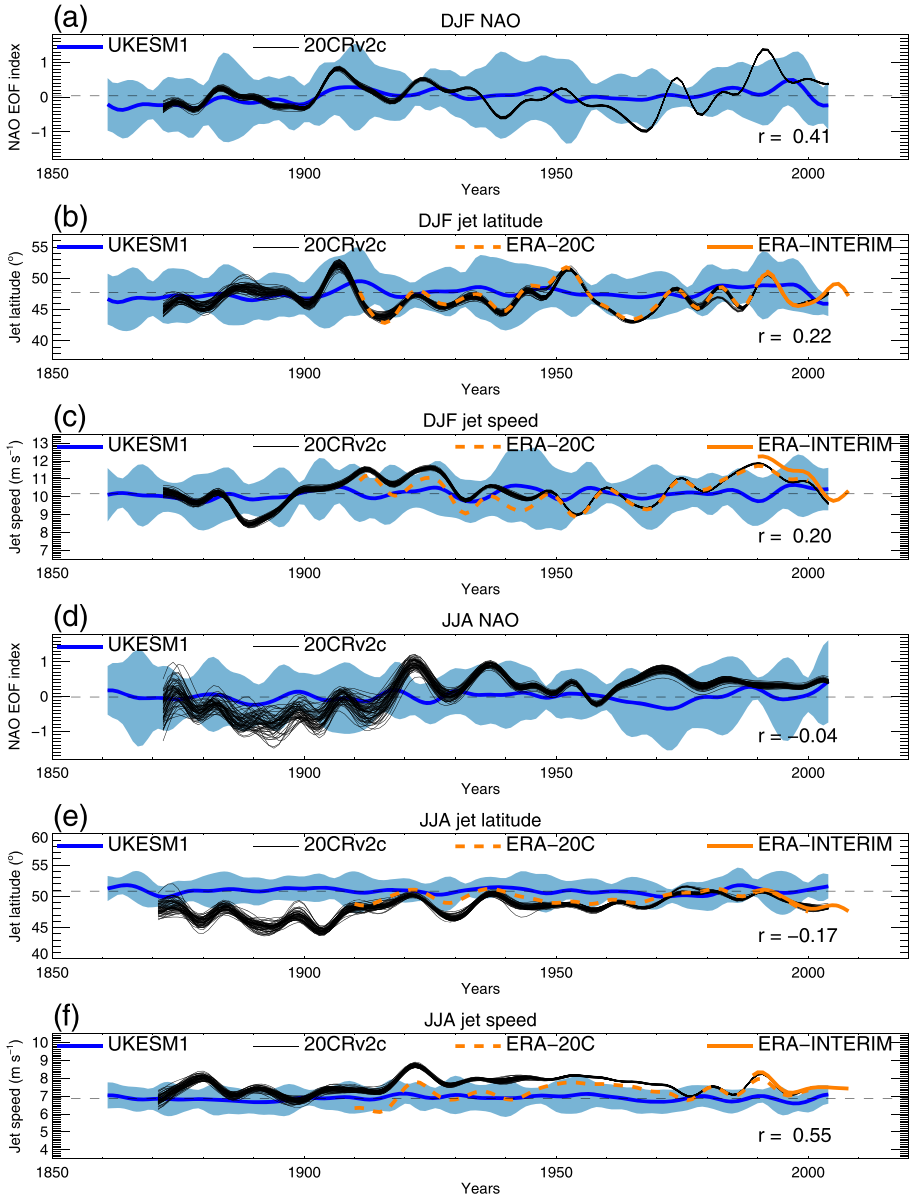


Figure 2. Time series of NAO EOF-based index (a, d), jet latitude (b, e), and speed (c, f) in winter (DJF, a–c) and summer (JJA, d–f). Each time series was filtered using a low-pass filter (greater than decadal time scales). UKESM1 ensemble means of historical simulations are shown by the thick blue line with the light blue shading showing 2 standard deviations of the ensemble spread. The thin black lines show the 20CRv2c reanalysis ensemble members. In panels (b), (c), (e), and (f) the thick orange lines show ERA-20C (dashed) and ERA-Interim (solid). Only 20CRv2c is used for the NAO time series reanalysis evaluation due to the potential for differences in EOF patterns extracted from shorter time periods to influence principal component time series. See section 2.2.1 for definition of indices used. The horizontal dashed lines show the time mean ensemble mean values for each diagnostic. In each panel, correlations (r) between the UKESM1 ensemble mean and the 20CRv2c ensemble mean time series are shown. Note that each time series is linearly detrended before calculating the correlations.

of jet speed in reanalyses at ~70 years (e.g., Woollings et al., 2014) is not reproduced by any of the nine UKESM1 ensemble members (see Figure 3e), and winter NAO variability also appears to be underestimated (see Figure 3e). This underestimation of multidecadal jet speed variability was also seen across the CMIP5 models (Bracegirdle et al., 2018), and the underestimation of NAO variability has been reported in other models of this resolution (Kim, Yeager, Chang & Danabasoglu, 2018), including GC3.1 (Menary et al., 2018). However, there remains some uncertainty over whether this difference is due to

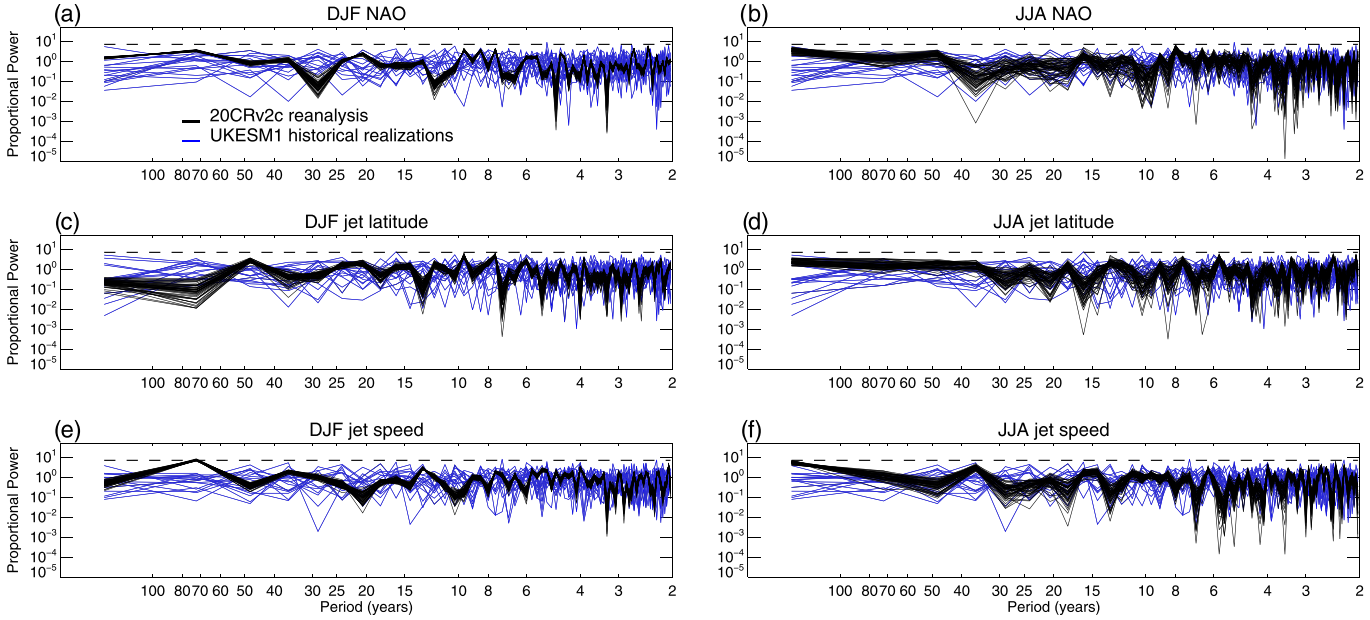


Figure 3. Proportional power spectral density (i.e., divided by the mean across all frequency components) of (a, b) NAO EOF index, (c, d) jet latitude, and (e, f) jet speed. The blue lines denote individual UKESM1 ensemble members, and the black lines show 20CRv2c ensemble members. Each time series was linearly detrended prior to the spectral analysis, and the time range was restricted to the period of 1862–2005 for comparison of the same frequency bins. The dashed line indicates the $p = 0.05$ threshold for the significance of a peak in any particular frequency bin using Fisher's g statistic.

reanalysis biases, model deficiencies in simulating the natural variability or forced response, or the occurrence of a low-probability natural variability event in nature (Bracegirdle et al., 2018).

Examining the DJF jet variability in more detail, the frequency distribution of JLI_{mon} over the historical period again shows consistency between UKESM1 and reanalyses (cf. Figures 4a and 4c). The model appears to simulate the general features of the well-documented trimodal structure in the distribution of observed jet position. However, UKESM1 generally has some difficulty in reproducing the frequency distribution minimum between 50°N and 55°N that is present in 20CRv2c. We note that this feature of the trimodal structure seems to occur more clearly in some of the UKESM1 ensemble members (e.g., r8i1p1f2 in Figure S3 of the supporting information), but not in others (e.g., r6i1p1f2), suggesting that it is sensitive to internal variability. A larger ensemble would help to evaluate potential role of internal variability in this difference. There is much more consistency in the reanalysis and model-simulated southernmost and middle-frequency distribution maxima. Note that UKESM1 is similar to many CMIP5 models of similar resolution, which generally produce jet latitude frequency distributions with clear differences from those observed (Iqbal et al., 2018).

For summer (JJA) there is less consistency between UKESM1 and reanalyses (Figures 2d–2f). During the 1979–2014 period the UKESM1 ensemble mean jet is predominantly weaker and more poleward than in reanalyses. However, the reanalyses sit mostly within the 2 standard deviation range of the ensemble spread. Previous to 1979 (i.e., before the introduction of widespread satellite remote sensing data) the general picture is similar, with the Atlantic jet weaker and more poleward, and a more negative NAO index. As for winter, there is no consistent evidence of a forced signal in UKESM1 decadal variability with generally only weak correlations with 20CRv2c. However, interreanalysis differences (and a larger ensemble spread across 20CRv2c ensemble members), particularly for jet speed diagnostics, suggest caution in making quantitative conclusions for this period. In terms of power spectra, there are no clear differences between UKESM1 and reanalysis ensemble members in summer (Figure 3). Frequency distributions of jet latitude (Figure 4) show that UKESM1 successfully captures the reanalysis-estimated range in latitude between about 38°N and 60°N, but the mode of the distribution is too equatorward in UKESM1.

3.1.2. Polar Vortex and Sudden Stratospheric Warmings

Variability in the NH winter stratospheric polar vortex has been found to significantly influence north Atlantic climate through the strength of the westerly polar vortex as well the abundance of vortex

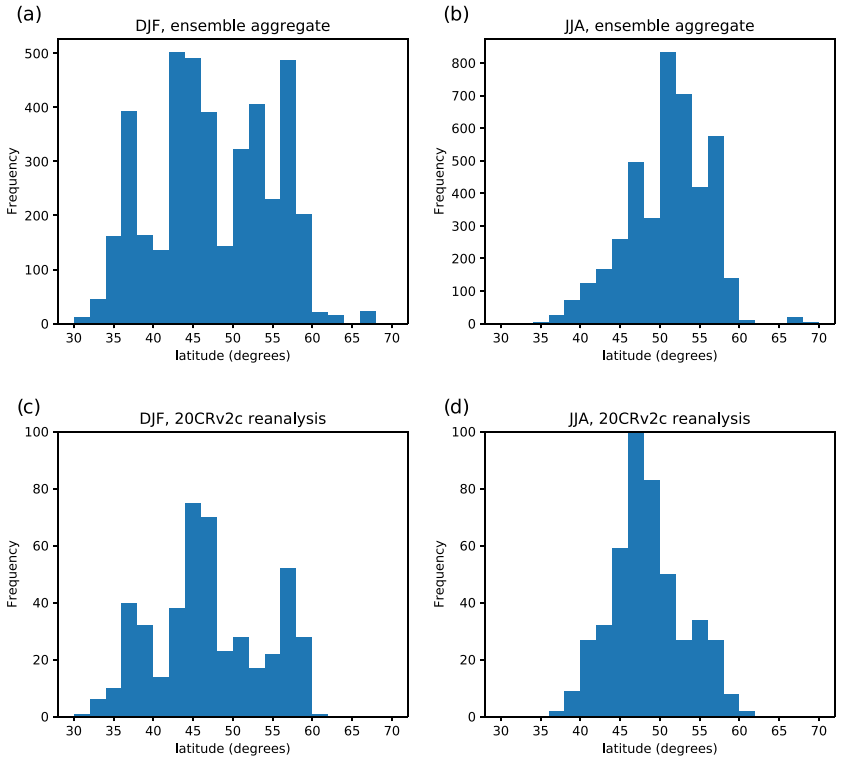


Figure 4. Frequency distribution of monthly JLI_{mon} for winter (DJF) (a, c) and summer (JJA) (b, d) from aggregates of all UKESM1 ensemble members (a, b) and 20CRv2c ensemble mean (c, d).

disruption events (sudden stratospheric warmings) on tropospheric winds (Baldwin & Dunkerton, 2001; Mitchell et al., 2013).

Figure 5 shows a comparison of the mean state and variability in stratospheric polar vortex strength in UKESM1 and ERA-Interim, as measured by the zonal mean zonal wind at 60°N on the 10 hPa level. The comparison indicates that the vortex forms early but is significantly weaker in early winter (November) in UKESM1 compared to ERA-Interim. Additionally, the UKESM1 ensemble mean exhibits a larger degree of interannual variability than ERA-Interim. These discrepancies are broadly similar to those seen in the physical climate model, GC3.1, especially in the early winter (Menary et al., 2018). The early formation could account for the weak early winter vortex as earlier westerlies will permit earlier vertical propagation of planetary waves into the stratosphere according to the Charney-Drazin Criterion (Charney & Drazin, 1961). These waves act to decelerate the westerly vortex. Figure 5b shows the mean state of wave activity

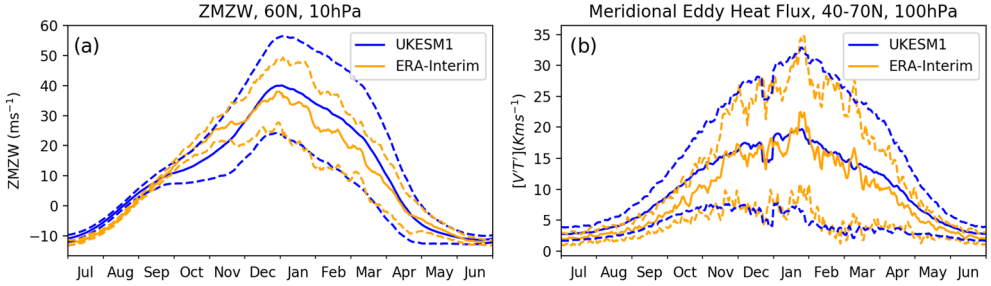


Figure 5. Panel (a) shows daily climatologies of zonal mean zonal wind at 60°N on the 10 hPa level over the 1979–2014 period for UKESM1 (blue) and ERA-Interim (orange). Solid lines denote mean values and dashed lines indicate the $\pm 1\sigma$ spread. Panel (b) shows the same as (a) but now for the eddy meridional heat flux averaged over 40–70°N on the 100 hPa level.

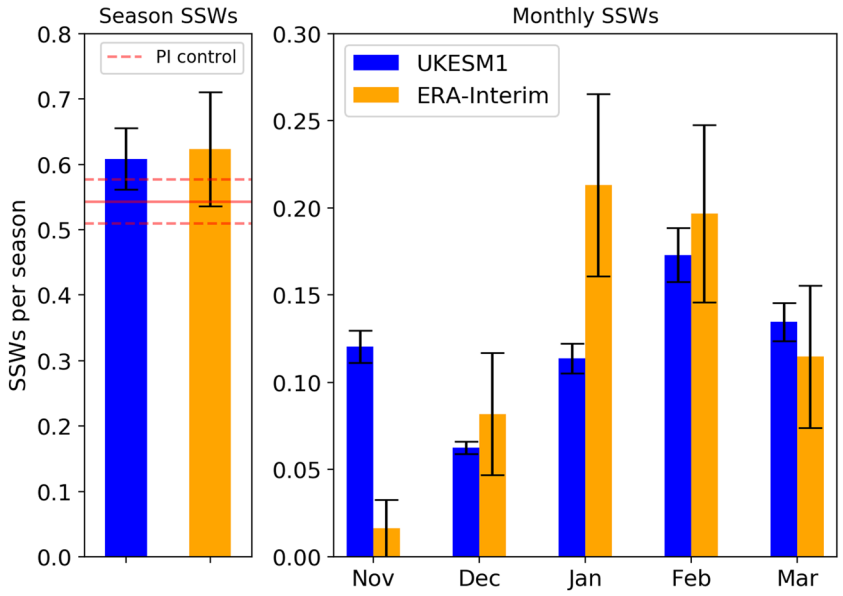


Figure 6. Histograms of mean SSWs per winter season (left) and per winter months (right) for UKESM1 historical ensembles (blue) and ERA-Interim (orange). Black error bars show ±1 standard error. Mean SSWs per season in the UKESM1 PiControl is also included as solid red horizontal line on the left plot, with dashed red lines showing ±1 standard error.

propagating upward from the troposphere and shows elevated meridional heat flux in UKESM1 during October and November compared to ERA-Interim. In this early winter period, wave driving in UKESM1 is ∼1 standard deviation above that of ERA-Interim. The cause of this overestimation in early winter tropospheric wave driving is not fully understood, but UKESM1's representation is generally reasonable over the remainder of the season.

Figure 6 shows the number of polar vortex disruption events (sudden stratospheric warmings or SSWs). UKESM1 and ERA-Interim broadly agree on the number of SSWs when considering the winter season as a whole (Figure 6, left). However, UKESM1 does exhibit discrepancies in the temporal distribution of SSW events within the winter season. For example, UKESM1 significantly overestimates SSW occurrence in early winter (November). However, the early winter overestimation of SSWs in UKESM1 is balanced by underrepresentation in January. Again, the early winter overestimation is consistent with the intraseason bias in GC3.1 (Menary et al., 2018), although UKESM1 appears to have fewer in January. The overestimated occurrence of SSW in November may be due to a combination of elevated wave driving in this interval as well as an early forming vortex allowing wave propagation (Figure 5). However, the causes for these discrepancies are as yet unexplained. Finally, UKESM1 historical runs also exhibit a marginally higher rate of SSWs than the preindustrial control simulation of the same model. However, the difference in SSW rate is not statistically significant. Interestingly, overall there appears to be less SSWs than in GC3.1, and so UKESM1 agrees better with reanalysis (not shown, but see Menary et al., 2018). This reduction in SSWs is consistent with previous work that showed that interactive chemistry reduces SSW (Haase & Matthes, 2019).

3.2. Atmospheric Composition

3.2.1. Tropospheric Ozone and Methane

Here we evaluate tropospheric column ozone and total column methane in UKESM1, which are two key composition metrics that are linked but have very different lifetimes. We assess their changes from 1850–2014 in the North Atlantic region. Both are simulated interactively in UKESM1, but are prescribed in GC3.1. Furthermore, changes in the oxidizing capacity (through changes in the hydroxyl radical and ozone) caused by (i) changes in emissions, (ii) changes in atmospheric circulation, and (iii) changes in humidity and temperature can lead to changes in the lifetimes and fate of aerosols and aerosol precursors.

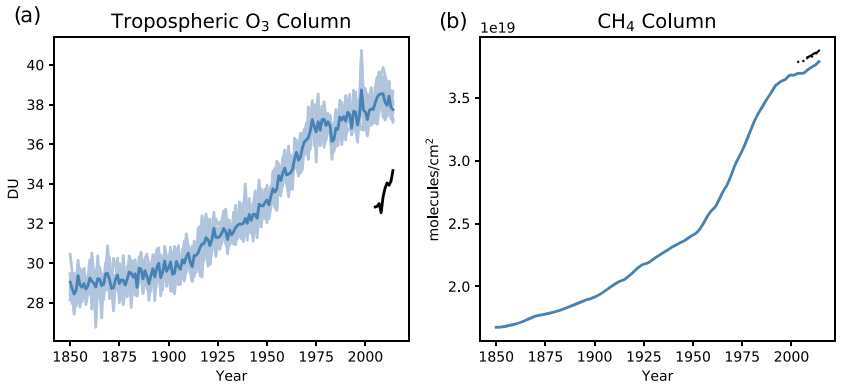


Figure 7. Panel (a) shows the annual mean tropospheric column ozone (in DU) averaged over the North Atlantic basin (60–10°W and 10–60°N) in UKESM1 for each ensemble member (light blue) and the ensemble mean (dark blue). Black shows satellite measurements from OMI/MLS. (b) Same as (a) but now for total column methane (in molecules cm^{-2}), with satellite observations from AIRS (dotted black line) and IASI (solid black line).

Tropospheric ozone is a greenhouse gas and a pollutant that is harmful toward human health, agriculture, and ecosystems. Formed by reactions of nitrogen oxides (largely anthropogenic) with volatile organic compounds (VOCs, including methane and carbon monoxide) in the presence of sunlight, tropospheric ozone has a lifetime of weeks (Monks et al., 2015). The relatively short lifetime and the temperature dependence in ozone formation mean that tropospheric ozone is very sensitive to changes in the emissions of ozone precursors as well as to changing weather and climate patterns. Furthermore, ozone acts as a precursor to the hydroxyl radical (OH), which itself governs the oxidizing capacity of the troposphere and determines the chemical lifetimes of the species consumed by OH. The OH radical also controls the gas phase conversion of sulfur dioxide to sulfuric acid (H_2SO_4) so is intimately related to changes in aerosols, clouds, and regional climate. OH is predicted interactively in UKESM1, but prescribed as a 3-D but time-invariant field in GC3.1.

Methane is an important greenhouse gas and it has a complex relationship with ozone. Methane acts as an important ozone precursor, but the methane abundance is regulated by the amount of OH. The main sources of methane are both natural (e.g., wetlands) and anthropogenic (e.g., fossil fuel extraction, agriculture; Nisbet et al., 2016; Rice et al., 2016). Methane is slowly removed from the atmosphere via the reaction with OH, resulting in a methane lifetime of approximately a decade. Changes in the Earth's surface types as a result of increasing temperature (e.g., melting of permafrost) and changes in the OH concentration could affect the long-term trends in methane.

Figure 7a shows annual mean tropospheric column ozone averaged over the North Atlantic Basin (60–10°W and 10–60°N). Tropospheric column ozone in the model has increased from 29 to 38 DU during the period of 1850–2014 (see Figure 7a). UKESM1 simulates a relatively slow increase in tropospheric ozone (~0.2 DU decade^{-1}) from 1850 until the early 1900s, after which there is a relatively rapid increase until the 1970s (~1.1 DU decade^{-1}). UKEMS1 continues to simulate an increase in tropospheric ozone column over the last few decades, but the rate of increase is slower (~0.4 DU decade^{-1}) and is closer to that seen at the early part of the time series. This trend in tropospheric ozone closely mimics the changes to NO_x emissions from 1850–2014 (see Hoesly et al., 2018) and the change in methane (Figure 7b), suggesting that increases in ozone precursors have driven the increase in the tropospheric column ozone changes in the North Atlantic region. However, other processes such as emissions of other ozone precursors (i.e., VOCs), transport of anthropogenic emissions into the basin from the surrounding continents, climate change, stratosphere-troposphere transport, and in situ chemical processing could also affect the tropospheric column ozone trend over the North Atlantic.

Figure 7a also includes a comparison with OMI/MLS observations (see section 2.2.2 for details). The satellite products used here span a relatively short part of the modeled time series but provide a more holistic representation of the North Atlantic than individual surface measurement sites. Averaged over the 2005–2014 time period, for which there are modeled and satellite measurements, the data show that tropospheric

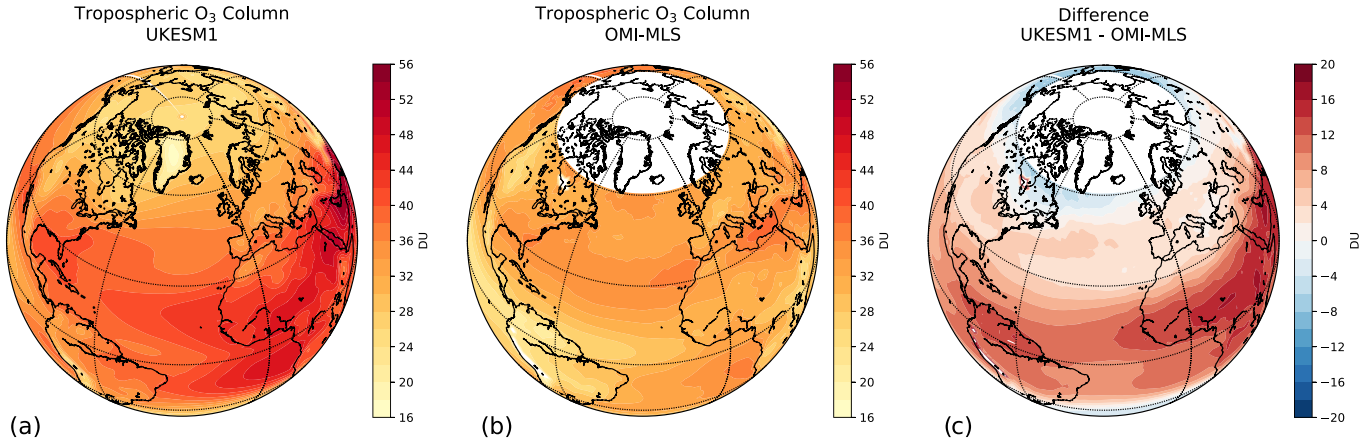


Figure 8. The spatial patterns of tropospheric ozone column (in DU), averaged from 2000–2014 in (a) UKESM1 historical simulations and (b) for satellite observations (OMI-MLS). Panel (c) shows the difference between model and satellite estimates (in DU).

column ozone values in UKESM1 are on average 4.7 DU higher than those from OMI/MLS. The satellite ozone column also increases at a faster rate than the model estimate over this time period.

Spatially, the modeled and observed tropospheric ozone demonstrate significant differences. Figure 8 shows the ensemble mean tropospheric column ozone, averaged from 2000–2014 from the model compared to satellite observations. Tropospheric column ozone values are highly variable across the North Atlantic basin in both the model and observations, as may be expected as a result of its relatively short lifetime (weeks) and the spatial heterogeneity of ozone precursor emissions. We find that the overestimation of the basin mean column ozone is largely explained by UKESM1 overestimating tropospheric column ozone over lower latitudes (see Figure 8c). Conversely, UKESM1 slightly underestimates tropospheric column ozone near the Arctic Circle. The potential causes for the discrepancy between the model and satellite estimates are numerous and could include different definitions of tropopause heights, inadequate model representations of tropospheric ozone chemistry, lightning NO_x, transport of ozone precursors from the tropics to higher latitudes, or oceanic deposition. These model-observation differences are topics of ongoing research, as are the drivers of the ozone trend over the North Atlantic. Thus, we will not discuss them in further detail here. However, we note that these biases in tropospheric ozone are likely to lead to an overestimate of the oxidizing capacity of the North Atlantic Basin. This overestimated oxidizing capacity will likely mean that the lifetime of gases like sulfur dioxide are too short in this region, assuming the other processes are well represented, and could have implications for cloud and radiative properties.

In the case of methane, measurements from surface stations (1984–2013) and glacial sites (pre-1984), globally averaged, were used to force the surface concentrations in the UKESM1 model (Meinshausen et al., 2017; Sellar et al., 2019). UKESM1 simulates an increase in basin mean column methane over the North Atlantic throughout the period of 1950–2014, from 1.7E19 molecules cm^{-2} to 3.8E19 molecules cm^{-2} (see Figure 7b). As with tropospheric column ozone, the rate of column methane increase has not been constant, with the fastest increase between 1950 and 1990 (~3.1E18 molecules cm^{-2} $decade^{-1}$) and a slower rate of increase over the last few decades (~7.7E17 molecules cm^{-2} $decade^{-1}$). These changes correspond closely to changes in methane emissions from 1850–2014 (e.g., Hoesly et al., 2018), but may also reflect changes in atmospheric oxidants including the OH radical (e.g., Rigby et al., 2017; Turner et al., 2017).

The simulated methane over the North Atlantic is slightly lower than both the AIRS and the more recent IASI satellite observations by about 0.1E19 molecules cm^{-2}, a difference of ~2.7%. This underestimation could in part be caused by the fact that globally averaged methane concentration was used as the lower boundary condition for the model, but in reality methane over Northern Hemisphere is several percent higher than in the Southern Hemisphere. Nevertheless, the rate of increase in methane column over the North Atlantic is comparable between the model and the satellite observations, suggesting that the methane sources in the model are largely reasonable.

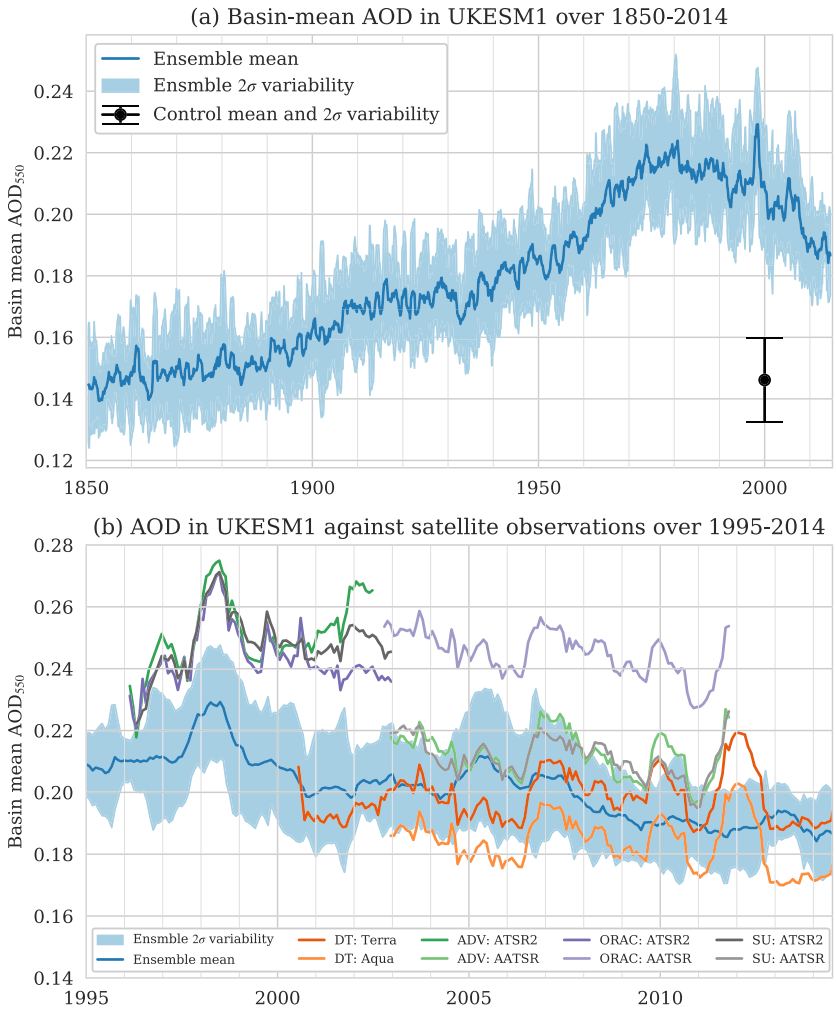


Figure 9. Variation in aerosol optical depth simulated by UKESM1 over the north Atlantic and its evaluation against observations. Panel (a) shows the evolution of AOD at 550 nm averaged over the North Atlantic basin (7.5–75°W and 0–65°N) over the period 1850–2014. Historical simulations are shown in blue, with 2σ ensemble spread shown as shading. The mean and the variability of the preindustrial control is shown in black. Panel (b) shows the same as (a) but now for the period when observations are available (e.g., 1995–2014). UKESM1 (blue) is compared to a selection of satellite data records (see legend). Observed data are described in section 2.2.2. All data areshown as a 12-month running mean of monthly fields.

3.2.2. AOD

Here we provide a broad evaluation of aerosol properties over the North Atlantic by focusing on the AOD. As mentioned in section 2.1, aerosols are simulated in UKESM1 using GLOMAP-Mode; the same scheme as used in GC3.1. However, UKESM1 uses interactive oxidants and has interactive vegetation, which are both prescribed in GC3.1.

Figure 9a illustrates the general behavior of basin mean AOD in the UKESM1 historical simulations. AOD increased significantly over the North Atlantic until the 1970s–1980s, when clean-air legislation was passed, and a steady reduction has occurred since then, though levels still remain much greater than those observed in the PiControl. There is significant seasonal variability (not shown), but variations across the ensemble are relatively small ($\simeq$0.01).

Figure 9b provides a qualitative comparison of Atlantic basin mean AOD in UKESM1 compared to a sample of the satellite data record. The UKESM1 time series falls roughly in the middle of the range of values reported by the satellites, with AODs of 0.21 in the late 1990s and 0.20 up to 2007, followed by a slow decrease to 0.18. Although observational products report a large spread in mean state, similar variability

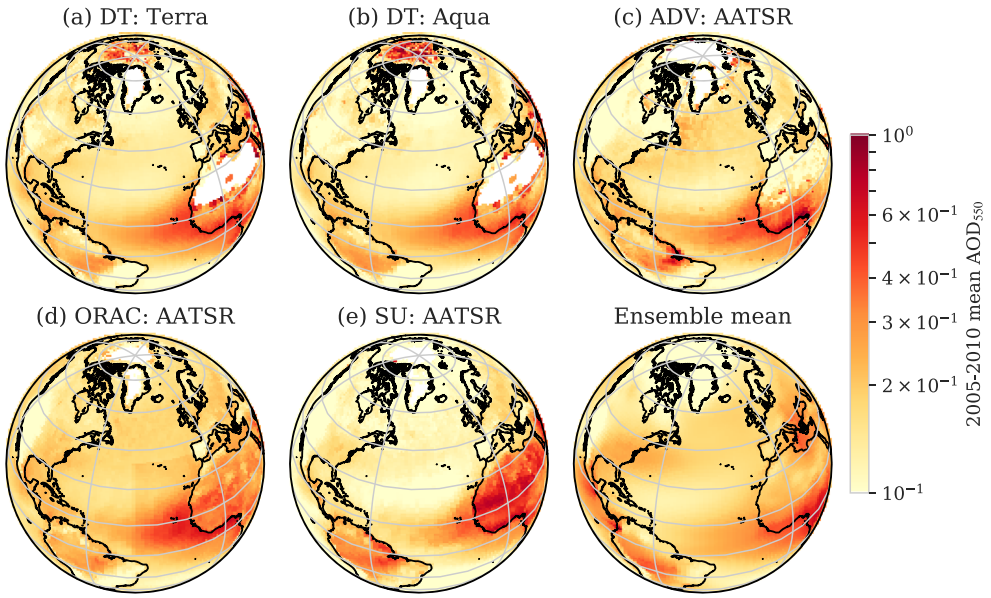


Figure 10. Distribution of annual mean aerosol optical depth (AOD) over 2005–2010 for observations, including Dark Target for (a) MODIS Terra and (b) Aqua, (c) ATSR Dual-View for AATSR, (d) the Optimal Retrieval of Aerosol and Cloud for AATSR, and (e) Swansea University for AATSR. The bottom right shows the ensemble mean of the UKESM1 historical simulations. Note the logarithmic color scale. See section 2.2.2 for details on observations.

can be qualitatively identified. However, it should be noted that the magnitude of these trends over 1995–2014 (~0.03) is less than the typical uncertainty reported by these satellite products (~0.05). An enhancement in AOD during the El Niño of 1998 is satisfyingly reported by observations and UKESM1. However, UKESM1 does not reproduce a sharp increase in AOD during 2011 seen by all of the satellites.

Figure 10 shows the spatial distribution of AOD during the period 2005–2010, when several satellites were in operation. UKESM1 captures the general pattern of AODs, and is qualitatively similar to the pattern of AOD in GA7.1 (i.e., the physical version of the model Mulcahy et al., 2018). High values are present in the Saharan outflow, reducing east of the Caribbean and increasing in the storm track region. However, UKESM1's AOD over the extratropical North Atlantic ocean, including the region downwind of the eastern seaboard of the United States, is generally on the high end of the observed range. In contrast, AODs over the Saharan outflow are not as large as those reported by satellites. UKESM1 also places the Saharan plume a few degrees further south than observed. This is partially an effect of the satellites-only sampling clear sky at particular times of day but also depends on the model's positioning of dust source regions.

Note that these comparisons are intended to be purely qualitative, and no effort has been made to address the differences in sampling between the model (which reports an all-time average) and the satellites (which report a clear-sky average at their overpass time). These differences have been shown to be significant (Schutgens et al., 2016). An analysis using hourly model output will be reported in a later study. There are also plans to adjust the model's aggregation to report more detailed statistics of AOD (Povey & Grainger, 2019).

3.3. Clouds and Radiation

3.3.1. Cloud Droplet Concentrations

Here cloud droplet concentrations (N_d) from the UKESM1 model are evaluated against observations from the MODIS satellite (see section 2.2.3 for observation details). N_d is an important quantity because, it is the first step in the chain of processes by which aerosols affect the microphysical properties of clouds. N_d gives some indication of the number of CCN (cloud condensation nuclei) that were available to produce cloud droplets in cloudy situations where aerosol retrievals (e.g., AOD) are not currently possible.

Figure 11 shows maps of the time mean N_d from both the UKESM1 and MODIS. In the North Atlantic region both show a similar spatial pattern; larger values are seen over landmasses and off the coasts of US/Canada

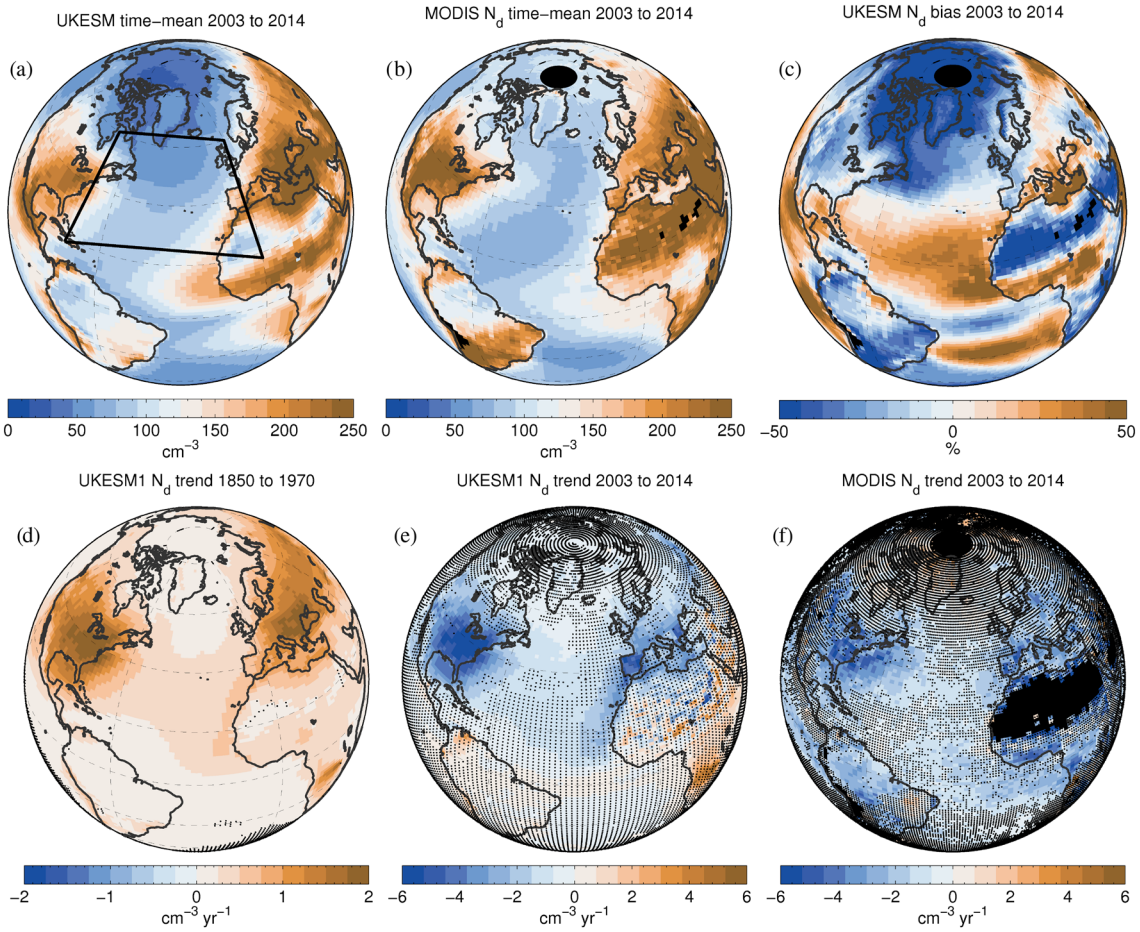


Figure 11. Time means and trends of cloud droplet concentration (N_d). Panels (a) and (b) show maps of the time mean N_d between 2003 and 2014 for UKESM1 and MODIS, respectively. Panel (c) shows the percentage model bias. Panel (d) shows a map of the linear trend in N_d in UKESM1 between 1850 and 1970. Panel (e) shows the same as 9d) but now for a trend between 2003 and 2014. Panel (f) shows the linear trend in N_d from the MODIS satellite between 2003 and 2014. Stippling denotes trends that are not significant at the 5% level using a two-tailed *Student's t*-test.

and western Africa, whist lower values are located in the middle of the Atlantic. However, somewhat different spatial patterns are seen at high latitudes. UKESM1 generally shows N_d values that decrease with latitude, whereas the satellite indicates higher values near the landmasses and lower values in the mid-Atlantic up to latitudes as high as Iceland. The satellite values are also generally higher than UKESM1 at high latitudes, although caution is required in this region due to the potential presence of sea ice that would make the retrievals less reliable (King et al., 2004; Platnick et al., 2001). At lower latitudes the model shows positive biases in N_d, particularly off the west coast of Africa. This suggests potential issues with aerosols sources from the landmass in this region, but also there is potential for dust and biomass burning aerosol to adversely affect the N_d retrievals (Haywood et al., 2004). Indeed, Figure 10 shows that satellite AOD is high in this region suggesting a large dust concentration there. See Grosvenor and Carslaw (2020) for additional evaluation of N_d in the North Atlantic region.

Maps of the linear trend in N_d are also shown in Figure 11 for both UKESM1 and the MODIS observations. The 1850–1970 trend (chosen because 1970 represents the peak of the N_d time series, see Figure 12) shows that UKESM1 N_d has increased most strongly over the American and European continents. Over the Atlantic Ocean there are larger increases downwind of the American continent (where the prevailing wind is westerly) and downwind of Europe (prevailing wind is northerly). In the middle of the Atlantic trends are smaller, likely indicating the fact that CCN from America are scavenged as air travels eastward, leading to fewer cloud droplets from anthropogenic aerosols.

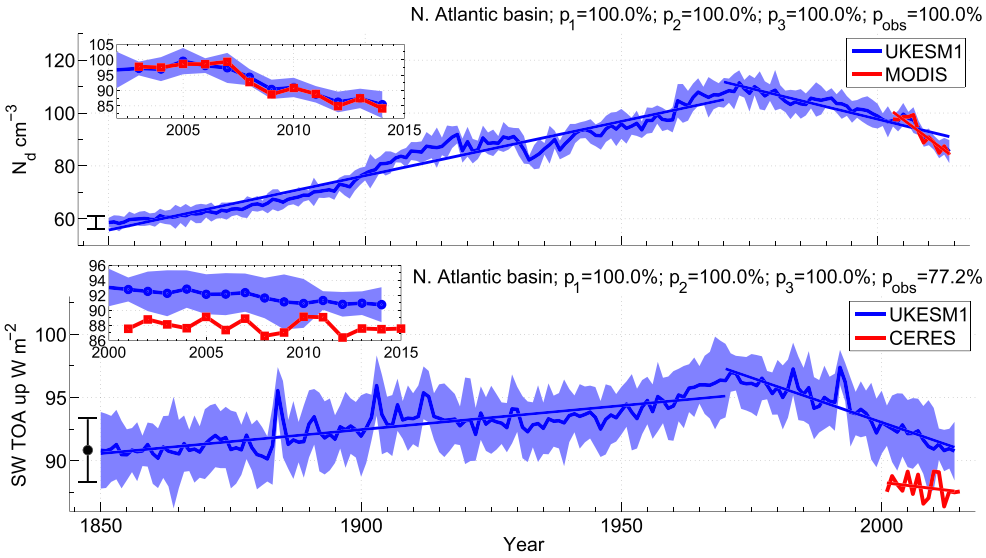


Figure 12. Time series of the annual mean cloud droplet number concentration (N_d, top) and SW TOA flux (bottom) averaged over the North Atlantic basin region marked on Figure 11a. Lines are shown for the UKESM1 ensemble mean and the MODIS and CERES satellite instruments. The blue shading denotes ±2 times the intermodel standard deviation across the ensemble. The black error bar plotted at year 1847 shows the time mean (black dot) and a range of $\pm 2\sigma$ of the annual averages of the PiControl run for the same region (for clarity, the time mean is not shown in the N_d plot because the range is small, but would lie in the middle of the error bar). The p_1, p_2, and p_3 values in the titles of each panel are the significance levels (using a two-tailed *Student's t*-test) of the computed model trends for 1850–1970, 1970–2014, and 2001–2014, respectively. The p_{obs} is the same for the observed trend over the period of the observations (2003–2014 for MODIS and 2001–2015 for CERES). The inset figure shows a close-up of the time period with observations using the same axes.

In contrast, the UKESM1 trend in N_d over the period of the satellite observations (2003–2014) is negative across all regions of the North Atlantic. This decrease likely reflects the impact of clean air legislation. The pattern and magnitude of the observed and modeled trend for 2003–2014 are generally similar with both showing large negative trends downwind of America and Europe and smaller trends in the mid-Atlantic. However, the region of strong negative trends downwind of Europe extends further south down the west coast of Africa in the model. This could indicate insufficient aerosol scavenging of European aerosols by UKESM1, or that changes in aerosol emissions from Europe or west Africa have changed too quickly. We also note that UKESM1 does not appear to capture the observed trends in N_d across the Tropical Atlantic. This difference could be related to erroneous positive N_d trend simulated over southern Africa, but further analysis of aerosol and cloud properties is needed here.

Figure 12 shows time series of the annual mean N_d for the North Atlantic basin region as marked in Figure 11. It shows an increase in N_d from 1850 onward up to a peak at around 1970, which is again consistent with the timing of clean air legislation. The temporal pattern on these long time scales is remarkably similar to that of AOD (see Figure 9). The UKESM1 values match the observed values for 2003–2014 very well with the time mean agreeing within 0.3%. However, as indicated by Figure 11 the good match is likely the result of some cancelation between negative biases in the north and positive ones in the south. The UKESM1 trend is also very similar to the observed one showing a general decrease over this period. The variability over time of the PiControl is quite small, hence the evolution of N_d is largely forced.

There are some caveats to the comparison with MODIS. Particularly, MODIS N_d observations have some uncertainty, possibly up to 50% according to an error analysis (Grosvenor, Sourdeval, Zuidema, et al., 2018). Therefore comparing absolute magnitudes may require caution. On the other hand, comparisons of MODIS N_d with in-situ aircraft actually suggest very good agreement (Ahn et al., 2018; Painemal & Zuidema, 2011). Nevertheless, we assume that the evaluation of the trends using MODIS is likely to be more robust than the absolute values. Hence, the good agreement in trends suggests a good a representation of aerosol emissions and their changes over recent times and suggests a good representation of aerosol activation, scavenging,

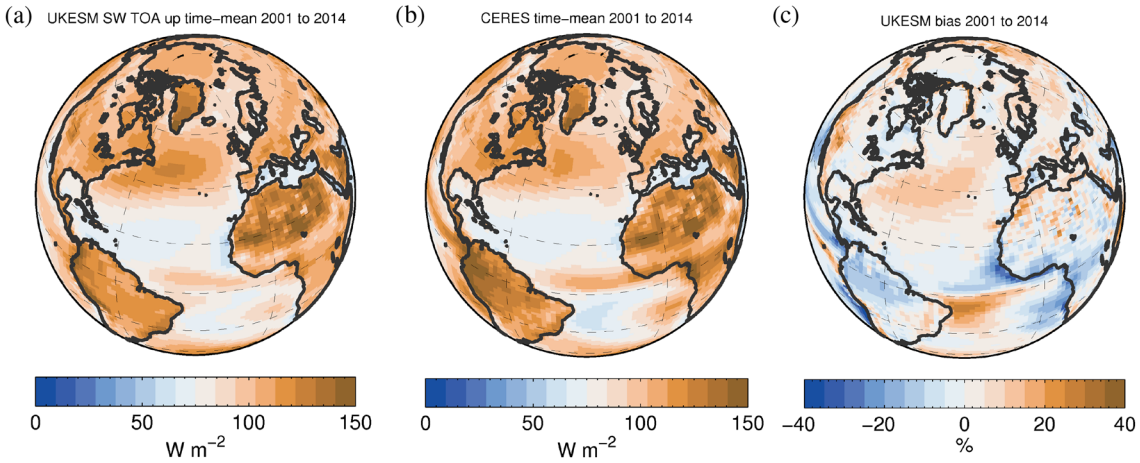


Figure 13. Maps of the time mean shortwave top of the atmosphere (SWTOA) fluxes for 2001 to 2014. (a) The UKESM1. (b) The CERES-EBAF satellite product. (c) The UKESM1 percentage bias.

and cloud processes. However, it is possible that individual biases in these processes compensate to give the correct end result; evaluation of the individual processes would be required to test this. We note that the pattern of the 2003–2014 trend is very similar to that between 1851–1970. Hence, the good performance in recent times could be indicative of long-term performance. However, more detailed further work will be needed to test this.

3.3.2. SW Fluxes at the TOA

Figure 13 shows the evaluation of time mean SW TOA radiative fluxes from UKESM1 using monthly CERES satellite data. The broad spatial pattern for UKESM1 is very close to that from the observations. The largest fluxes occur in the northwestern part of the North Atlantic off the coast of Newfoundland due to the large areal cloud cover in that region as shown in Figure 14. Just south of that region there is a region of positive SWTOA bias of around 15%, which represents absolute biases of around 15 W m^{-2} (see Figure 13c). The region of positive bias coincides with a region of positive total cloud fraction (see Figure 14c). Over the whole North Atlantic basin region there is a time mean model bias of 3.8 W m^{-2} (see Figure 12), with the observed value being outside the 2-standard deviation spread of the model ensemble. Radiative calculations in Grosvenor and Carslaw (2020) show that the model SWTOA biases in this region are mostly caused by the positive UKESM1 cloud fraction biases. The latter paper also

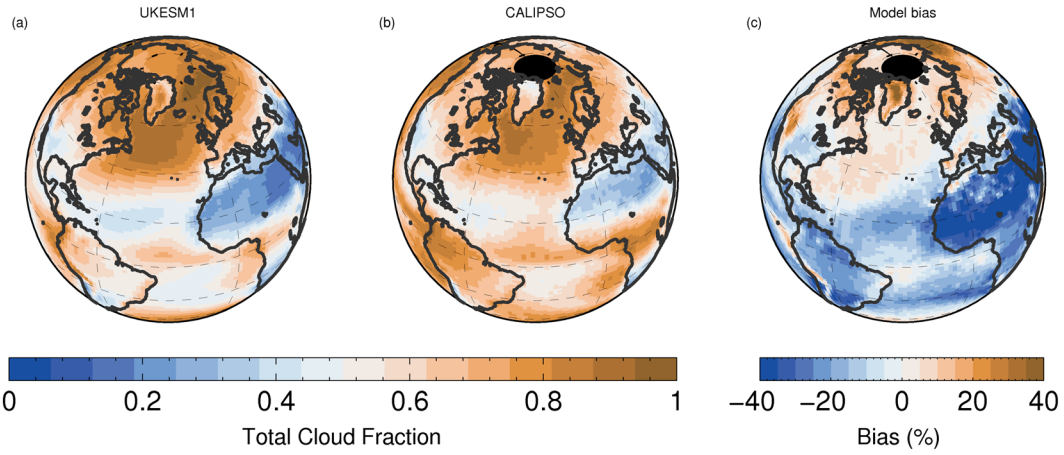


Figure 14. As in Figure 13 except for the total cloud fraction and for the period 2007–2014. CALIPSO is used for the observations and the COSP CALIPSO satellite simulator is used for the model.

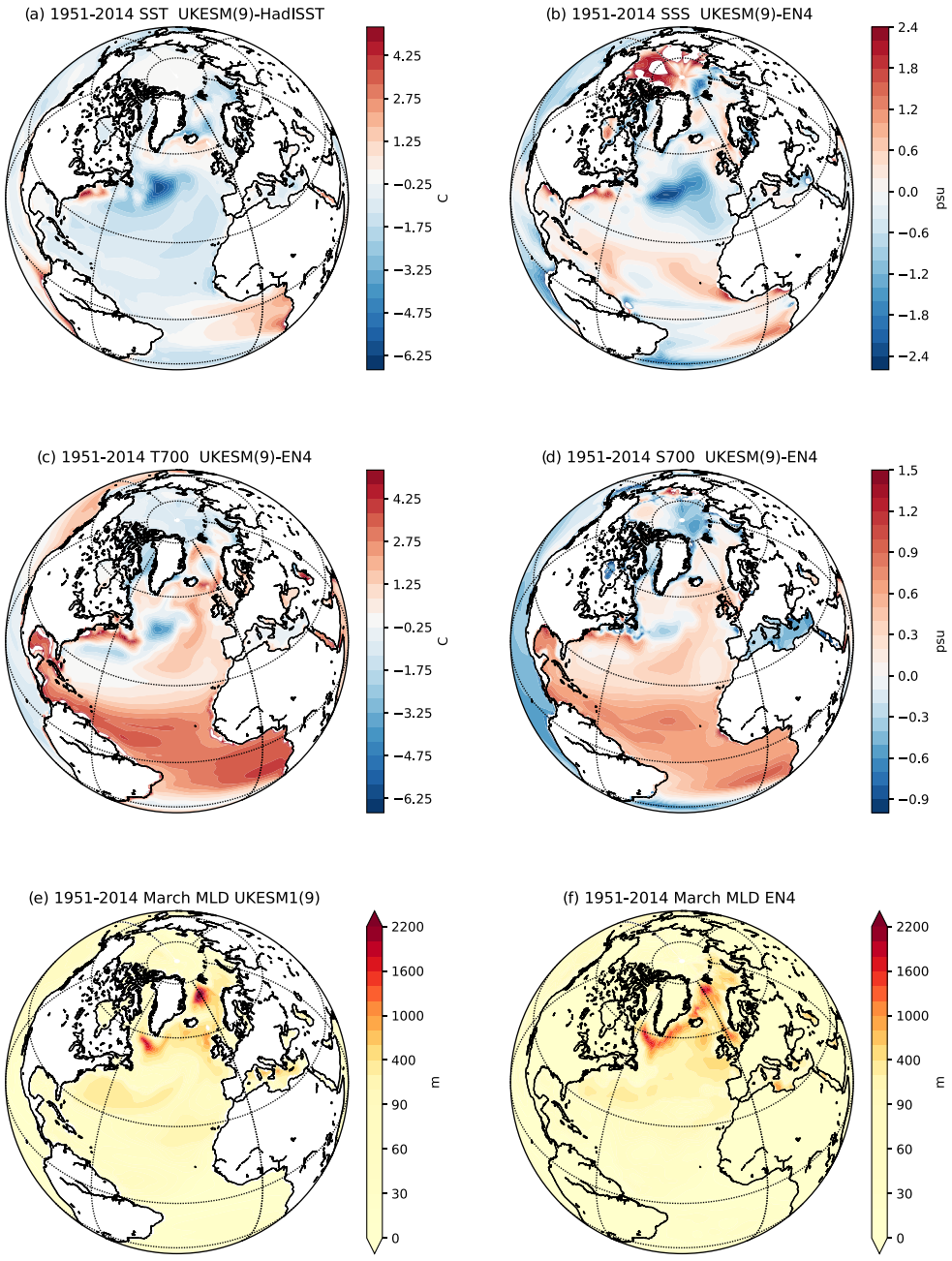


Figure 15. UKESM1 model biases (UKESM1—observations) in annual mean. Panel (a) shows sea surface temperature (SST, °C). Panel (b) shows sea surface salinity (PSU). Panels (c) and (d) show upper ocean (0–700 m) mean temperature (T700, °C) and salinity (S700, PSU), respectively. Panels (e) and (f) show ensemble mean UKESM1 and observed absolute March mixed layer depth (m), respectively. In panels (a) to (d) differences that are smaller than 1 standard deviation of the ensemble spread are shaded white. All differences shown are for the 1951–2014 mean. See section 2.2.4 for details of observational data sets used and for the details of the mixed layer depth definition.

gives further in-depth evaluation of cloud and radiative properties in the UKESM1 model over the North Atlantic along with analysis of its aerosol forcing.

As for N_d, the time series of the regional average ensemble mean SWTOA flux (Figure 12b) shows substantial multidecadal change in UKESM1. Over 1850–1970 there is a positive trend of $0.038 \pm 0.008\,\mathrm{W\,m^{-2}\,yr^{-1}}$, which is significant at the >99.9% level using a two-tailed *t*-test. Between 1970 and 2014 there is a

negative ensemble mean trend of $-0.14 \pm 0.03\,\mathrm{W\,m^{-2}\,yr^{-1}}$ (significant at the >99.9% level) in agreement with the modeled N_d trend. Furthermore, the trends for the individual ensemble members are all negative ranging from -0.18 to $-0.12\,\mathrm{W\,m^{-2}\,yr^{-1}}$.

Over the observed period (from 2001–2014) there appears to be a disagreement between the strength of the simulated and observed trends in SWTOA. UKESM1 continues to decline with a negative trend of $-0.17 \pm 0.04\,\mathrm{W\,m^{-2}\,yr^{-1}}$ (significant at the >99.9% level), but the observed trend is much flatter ($-0.05 \pm 0.13\,\mathrm{W\,m^{-2}\,yr^{-1}}$), and it is not significantly different to zero. Note that these values are not overly sensitive to the exact boundaries of the North Atlantic region used. All else being equal, the negative N_d trend would be expected to cause a negative SWTOA trend. Therefore, the lack of an observed trend indicates that trends in other factors are opposing the expected N_d effect. Likely candidates are trends in cloud fraction or cloud liquid water path as albedo has a stronger dependence on these than on N_d. Cloud fraction and cloud liquid water path are also quite variable from year to year (Malavelle et al., 2017), which may mask an observed trend in SWTOA. The range of simulated trends across the ensemble for this period is -0.35 to $-0.05\,\mathrm{W\,m^{-2}\,yr^{-1}}$ with the observed trend being 1.4σ above the ensemble mean trend, where σ is the interensemble standard deviation. This indicates that there is a high degree of internal variability within the system that allows a large range of SWTOA trend values to be feasible. Therefore, the lack of an observed trend may be consistent with natural variability; however, further detailed investigation, ideally over a longer time period, is warranted to understand if the simulated and observed trends are different.

3.4. Ocean State and Circulation

We now turn our attention to the behavior of the Atlantic Ocean in UKESM1. First, we will consider the biases in the mean state, then we will discuss the time evolution of key ocean metrics (AMV and AMOC).

3.4.1. Mean State 1950–Present

Figure 15 summarizes the model annual mean climatological biases with respect to observations in the North Atlantic Ocean. Subsurface ocean data are relatively sparse before 1950, therefore we restrict the comparison of UKESM1 with observations to 1951–2014.

There are significant biases present in the surface fields in UKESM1. Figure 15a shows that SSTs in UKESM1 are generally colder than the observed record with a notable maximum bias of ∼−4°C within the southwestern subpolar North Atlantic (SPNA, 45–67°N, 7.5–75°W). This widespread cold bias is consistent with the cold bias reported by Williams et al. (2018) in earlier versions of GC3.1 run under present-day forcings. That paper partly attributed this cold bias to aerosol-cloud interactions. Warm regional biases also occur, most notably in the ocean upwelling region in the eastern tropical south Atlantic (∼1°C). However, the dipole in the bias, with warm anomalies along the eastern U.S. coast and cold anomalies off Newfoundland, is consistent with a systematic bias in the position of the Gulf Stream and North Atlantic Current in models of this resolution, including GC3.1 (Kuhlbrodt et al., 2018). A similar bias in ocean surface currents is seen in UKESM1, with a Gulf Stream separation that occurs too far north, and a too zonal North Atlantic Current (see figure S5). These patterns of SST bias are generally replicated in sea surface salinity (SSS, Figure 15b), with a notable fresh bias in the SPNA. The SSS bias is especially large at the southern gyre boundary ∼45°N and is also consistent with the too zonal North Atlantic Current and the related southward displacement of fresh subpolar waters (see Figure S5). Further south, in the tropical North Atlantic, there is a widespread positive SSS bias (0–30°N).

Over the depth of the upper Atlantic Ocean (0–700 m) many biases are opposite to those at the surface; with widespread warmer and more saline water peaking in the tropics (see Figures 15c and 15d, respectively). Cold and fresh biases only persist throughout the upper ocean in the Gulf Stream Extension and the southern SPNA. Within the SPNA the spatial pattern of the biases in temperature and salinity are similar to those in the GC3.1 PiControl (Menary et al., 2018). This suggests that the SPNA biases are due to mean state model biases, rather than biases in the UKESM1's response to the twentieth century changes in radiative forcings. Over the wider Atlantic, the contrast between the colder and fresher surface biases and the warmer and more saline biases below, suggests that ocean stratification is weaker in UKESM1 than in observations. This weaker stratification in UKESM1 is confirmed through analysis of temperature and salinity profiles (see Figure S6).

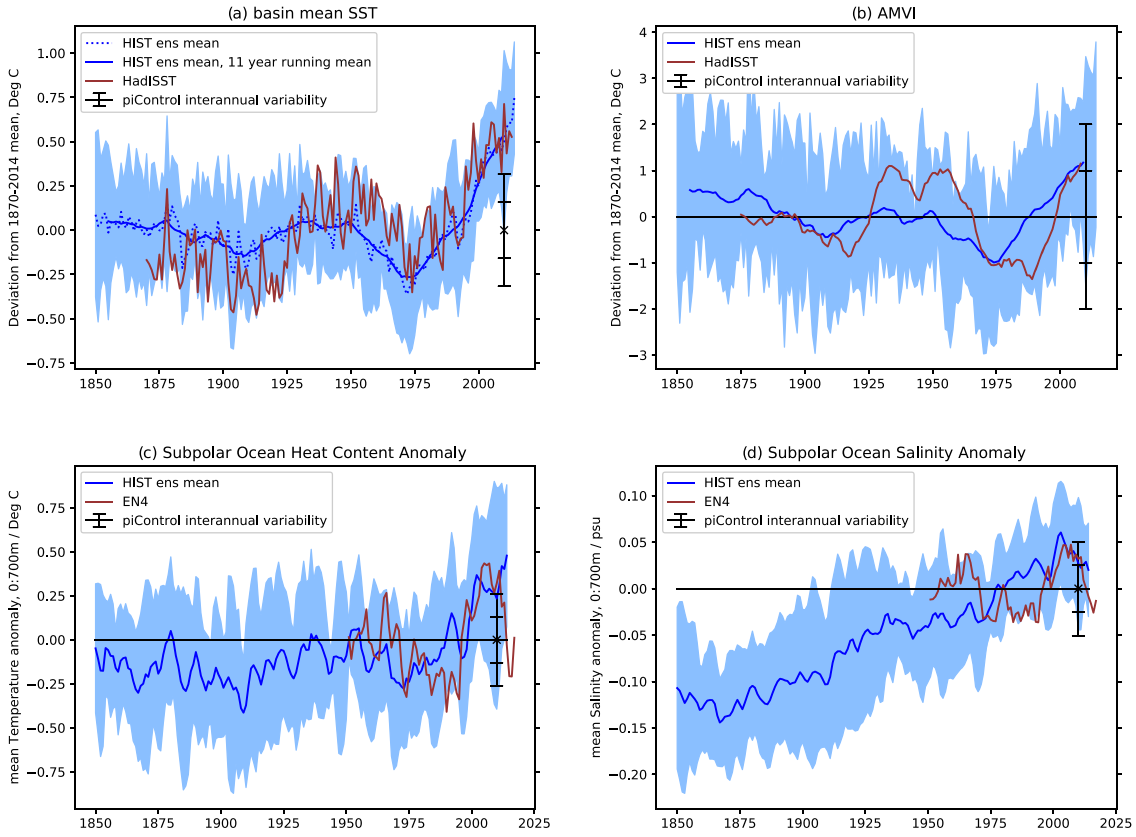


Figure 16. Atlantic basin ocean indices. (a) Atlantic multidecadal variability (AMV: represented by basin mean SST over 0–60°N, 7.5–75°W). Dotted blue shows UKESM1 ensemble mean and solid blue shows its 11-year running mean. Red shows observed AMV. Both are anomalies with respect to the observational period (1860–2014). Panel (b) shows the the same as (a) but now for the AMVI index computed as AMV, but then removing the global mean SST (excluding AMV box) by linear regression—see section 2.2.4 for details. Panels (c) and (d) show upper (0–700 m) ocean mean temperature and salinity anomaly in the SPNA (45–67°N, 7.5–75°W), respectively. In (c) and (d) red shows the same from EN4 observations, and anomalies are with respect to the observation period (1951–2017). In all plots, the shaded envelope shows 2σ ensemble spread for UKESM1. Black bar on the right of each plot shows the time mean and 1 and 2σ computed from the UKESM1 PiControl.

Figure 15e shows March mixed layer depths in UKESM1. Previous modeling studies (e.g., Chanut et al., 2008; Marzocchi et al., 2015) have shown the importance of mesoscale eddies in the formation of Labrador Sea Deep Water, particularly in their role in transporting heat from the boundary currents into the center of the Labrador Sea, where doming of the isopycnals preconditions the water columns for deep convection in intense heat loss events. For this reason, we would not expect UKESM1, whose resolution does not permit mesoscale variability, to simulate deep water formation in the Labrador Sea region particularly realistically. Nevertheless, Figure 15f confirms that the winter mixing in the Labrador Sea in UKESM1 is reasonably consistent in both its spatial distribution and depth with estimates of the observed climatology. The deep winter mixing volume in the Nordic Seas is also generally comparable between model and observations. The winter mixing is relatively evenly distributed across the Nordic Seas in observations, with mixing down to between 400 and 1,400 m. However, mixing is more concentrated in the region immediately south of Fram Strait in UKESM1, with the MLD extending to below 2,000 m there. In contrast, UKESM1 simulates shallower MLDs in the Barents Sea and the Norwegian Sea compared to observations.

3.4.2. Time Evolution of Key Variables

We now turn our attention to key ocean metrics to assess how well UKESM1 reproduces the observed evolution of the Atlantic Ocean over the observed period.

It is well known that the observed Atlantic SST record displays pronounced multidecadal variability (Sutton et al., 2018). The basin mean AMV index is shown in Figure 16a. The ensemble mean index captures some of the multidecadal variability exhibited in observations and the two are nearly in phase; although the early

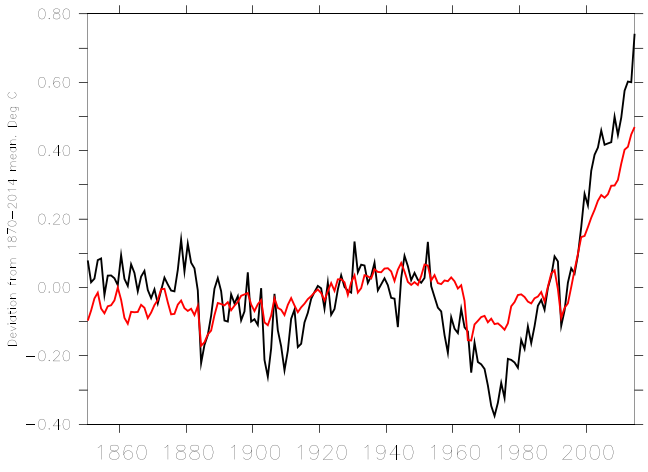


Figure 17. Comparison of SST changes inside and outside the Atlantic. UKESM1 AMV index (black) computed as SST mean over AMV region (0–60°N, 7.5–75°W). Mean of SSTs (red) between 60°S and 60°N excluding AMV region.

twentieth century warming begins somewhat earlier and the cooling post 1960 begins earlier in UKESM1. This phasing of the AMV index in UKESM1 is consistent with the evolution of AMV in GC3.1 (Andrews et al., 2020) and suggests that external forcings have contributed to the multidecadal signal in the AMV. The difference in phasing between UKESM1 and observations is likely to be consistent with a role for internal variability in the observed changes, but could represent a poorly simulated forced response. Nevertheless, while the ensemble mean basin mean AMV index only captures some of the observed variance, the observations do sit within the 2σ ensemble spread.

While the basin mean AMV index shows a clear multidecadal signal, this may be superimposed on the globally forced signal. We estimate and then remove a globally forced signal to produce an AMVI index (see section 2.2.4). Figure 16b suggests that the ensemble mean AMVI in UKESM1 has less multidecadal variability than observed; that is, forcings do not explain all of the observed variability. However, the AMVI does capture some aspects of the cooling in the 1960s. However, in using the global mean SST as an estimate of the global externally forced signal to isolate AMV, we make a prior assumption that such multidecadal variability is only confined to the Atlantic. However, in UKESM1 the global mean SST index does show pronounced multidecadal variability that is in phase with the AMV index (See Figure 17 or Sellar et al., 2019). Furthermore, the spatial pattern associated with the AMVI index also appears to be largely interhemispheric, with large anomalies that explain substantial amounts of variance across much of the Northern Hemisphere oceans including the northwest Pacific, the Mediterranean, and the subarctic (see Figure 18). This spatial pattern is similar to that seen in GC3.1 (Andrews et al., 2020), but is different to that found in HadISST observations where the spatial pattern associated with AMVI is primarily located in the Atlantic. The spatial pattern within the Atlantic is also different in the extratropics; UKESM1s strongest positive anomalies appear in the Gulf Stream extension, but has a cooling of the western SPNA (again similar to GC3.1, Andrews et al., 2020). In observations, the positive anomalies are spread broadly across the whole SPNA. Note that the comparison of AMV spatial pattern between UKESM1 and observations is also not sensitive to the exact definition of the AMV index used (see Figure S7). Therefore, the extent to which the AMV in UKESM1 is part of a global signal or part of an Atlantic-only signal is difficult to untangle.

The SPNA is also a key region of multidecadal variability in both the observed record (Reverdin, 2010; Robson et al., 2016) and in climate models (Menary et al., 2015; Visbeck et al., 1998). Figure 16c shows the mean upper (0–700 m) ocean temperature (T700) anomalies in this region. In UKESM1, the SPNA T700 echoes the model AMV index, showing some multidecadal variability, but no overall trend until the mid 1970s, followed by a rapid rise. Observations for the upper ocean are more limited in time than those for SSTs; hence, reliable records only exist in the SPNA back to the 1950s. Despite this, UKESM1 appears to capture some of the 1960s cooling of the SPG and the subsequent post-1975 warming seen in the observed record. However, the post-1960s cold minimum occurs too early, and the post 1970s' warming is slower than observed. Additionally, UKESM1 does not capture the recent (2005–present) observed cooling. Furthermore, although UKESM1 appears to capture some of the multidecadal variability of upper ocean heat content in the SPNA, we note that the T700 trends across the North Atlantic more broadly are generally larger in magnitude than that observed, particularly for the 1960s cooling period (1951–1980; see Figure S8). Therefore, further investigation of the upper ocean heat content changes are needed.

The SPNA has also undergone pronounced salinity changes over the observed period (Belkin et al., 1998; Dickson et al., 1988). Figure 16d compares the upper ocean (0–700 m) mean salinity anomalies in the SPNA. It is notable that the ensemble mean signal in UKESM1 does not capture the observed multidecadal variability. Although the observed variation is just within the 2σ UKESM1 ensemble spread over this period. However, UKESM1 displays a near-monotonic positive trend in salinity throughout the simulation. This trend in the ensemble mean (0.1 PSU/century) is large compared to the ensemble spread of the historical runs and is greater than both the trends seen in the EN4 observations (0.02 PSU/century) and the

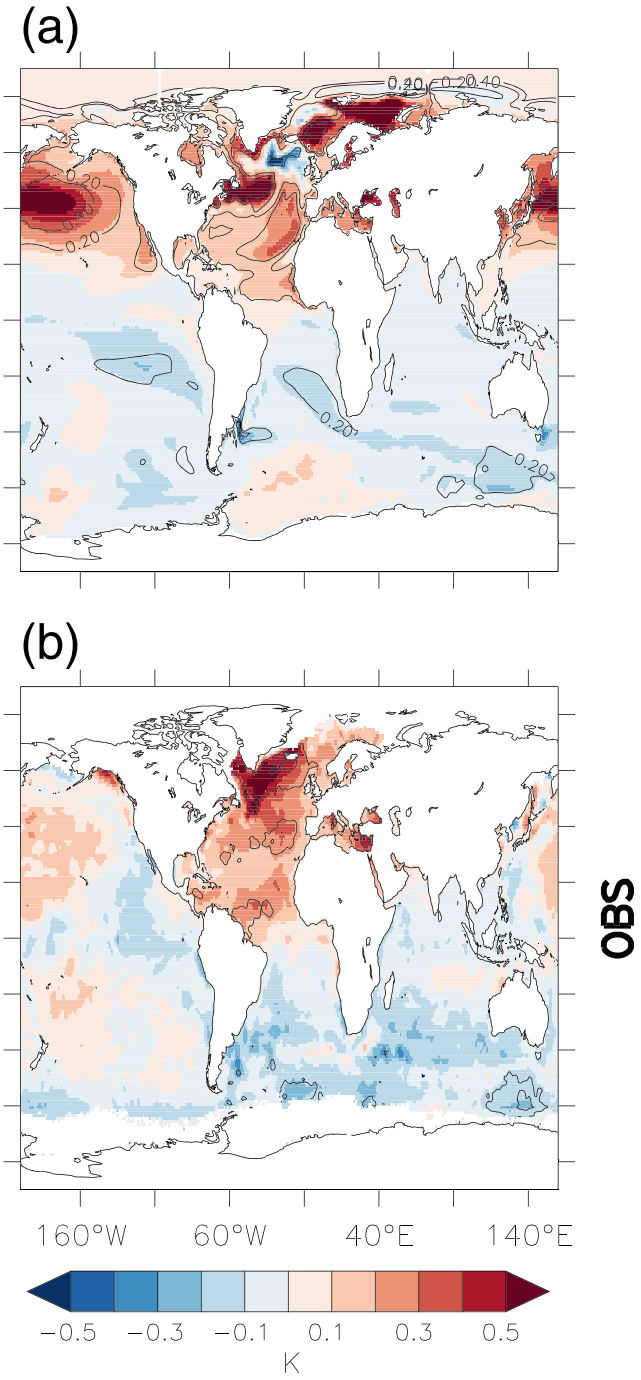


Figure 18. Spatial pattern of Atlantic multidecadal variability (AMV) in UKESM1 and observations. Panel (a) shows the regression of ensemble mean SST anomalies onto the ensemble mean AMVI index in UKESM1 in colors. SST anomalies have had the global mean temperature anomalies removed by regression on global mean temperatures. Black contours show the variance explained computed as the square of the correlation coefficient. Panel (b) shows the same as (a) but for HadISST observations.

PiControl (−0.02 PSU/century). By 2014, the salinity in the SPNA in UKESM1 exceeds the observed values (UKESM1 2014: 35.14 PSU, Observed 1950–2014 mean: 34.919 PSU, see Figure 15). We note that the SPNA S700 simulated in PiControl does display considerable centennial variability and exhibits some trends comparable to those seen in the historical ensemble mean. However, the ensemble mean historical trend is larger than 80% of all possible 164 year trends in the PiControl and each member in the historical

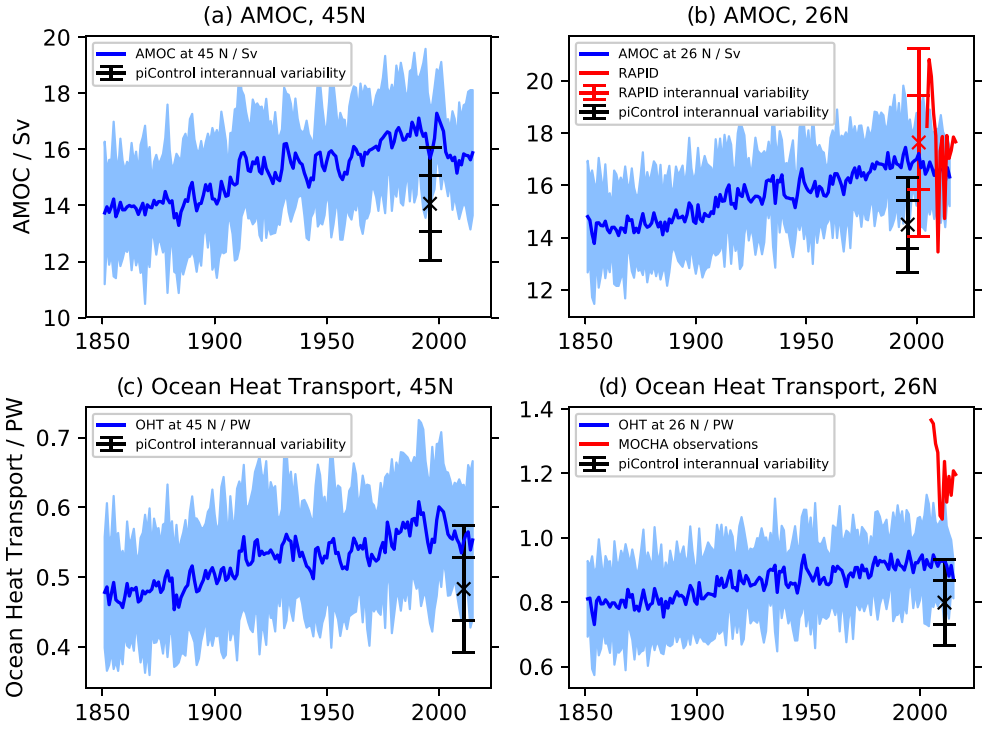


Figure 19. Atlantic meridional overturning circulation (AMOC) and ocean heat transport. (a) Blue shows ensemble mean AMOC at 45°N (maximum overturning stream function between 43°N and 46°N) in UKESM1 historical simulations in Sverdrups (shading shows 2σ spread across the ensemble). Mean and 1 and 2σ from PiControl are shown in black to the right. (b) As in (a) but for 26°N (maximum overturning stream function between 24°N and 27°N). (red) Annual mean RAPID AMOC observational data. Red bar shows 1σ RAPID annual mean range. (c) and (d) Same as (a) and (b) but now for northward ocean heat transport in petawatts. The MOCHA observed estimate of ocean heat transport is used here.

ensemble was initialized from a different point in the PiControl to average out centennial variability. Therefore, the trend in S700 is a forced response in UKESM1. Interestingly, the changes in subpolar SSS in UKESM1 historical simulations are very similar to those in S700, with a near-monotonic increase of ~0.2 PSU/century (see Figure S9). However, we note that observations suggest that subpolar SSS likely decreased from 1900 (Friedman et al., 2017), which could indicate that the forced response in UKESM1 is not correct.

3.4.3. Ocean Transports

Figure 19 shows the AMOC and the OHT for two key latitudes within the North Atlantic basin (26°N and 45°N). The AMOC at both latitudes (Figure 19 upper) shows a positive trend throughout the twentieth century, until the 1990s, where it begins to decrease. The twentieth century rise is outside the 2 standard deviation range of the PiControl. Taken together with the small spread across the historical ensemble, this suggests that these variations arise due to external forcings. It is also interesting to note that these low-frequency changes in the AMOC are similar to those seen in AOD (Figure 9) and N_d (Figure 12), albeit with a lag. Therefore, the changes in AMOC are consistent with the aerosol forcing (Menary et al., 2013, 2020) and are also consistent with the forced AMOC changes in GC3.1 (Andrews et al., 2020). However, we note that the rising AMOC is also accompanied by some multidecadal variability about the long-term AMOC trend with similar time scales to AMV (Figure 16a). For example, the AMOC peaks above the long-term trend around 1920 and 1980 and falls below the trend around 1940 and after 2000. Similar variations (of the opposite sign) are seen in the variations in the AMV around the long-term trend.

By the 21st century, the mean AMOC at 26°N in UKESM1 is comparable to the contemporary AMOC RAPID observations (Smeed et al., 2018), and GC3.1 (albeit slightly weaker). However, it appears that UKESM1 underestimates AMOC variability, consistent with other physical climate models (Roberts

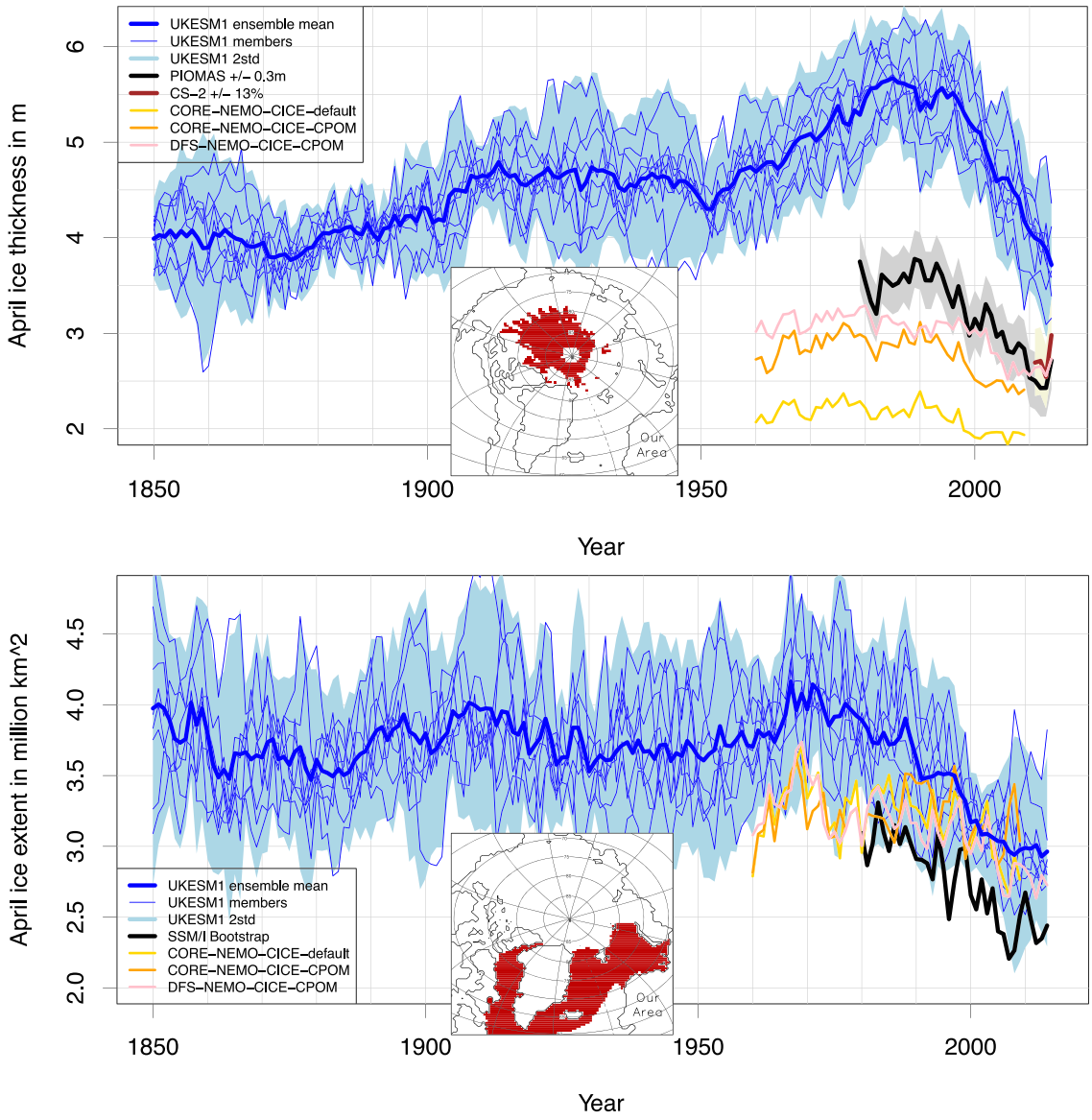


Figure 20. Evolution of sea ice thickness and sea ice extent in UKESM1. (top) April ice thickness averaged over the Central Arctic (red region in inset). UKESM1 (blue) is compared with forced ocean-sea ice simulations and estimates from Cyrosat-2 (CS-2) and PIOMAS (see legend). (bottom) Combined sea ice extent in the three Arctic Seas connected to the North Atlantic: Barents Sea, East Greenland Sea, and Buffin Bay (see red region in inset). UKESM results compared with forced ocean-sea ice simulations and data derived by the SSM/I Bootstrap algorithm (see legend).

et al., 2014). The AMOC decline post-2000 in UKESM1 ($-0.025 \pm 0.012\,\mathrm{Sv\,yr^{-1}}$) is the same sign as, but weaker than, the observed decline ($-0.5\,\mathrm{Sv\,yr^{-1}}$ reported over the 2004–2014 period Smeed et al., 2018). This AMOC decline is also similar to that seen in GC3.1 (Andrews et al., 2020).

The northward ocean heat transport across both 26°N and 45°N in UKESM1 also shows a pronounced upward trend (see Figures 19c and 19d). Again, the ensemble spread is small compared to this trend, suggesting this trend is a forced response. There is some evidence of a decline in heat transport after 1990. Comparing these heat transport trends with those of the AMOC, we conclude that the increased heat transport is likely to be driven by the changes in volume transport, rather than changes in temperature (i.e., $(VT)' \sim V'\overline{T}$). However, it is important to note that the ocean heat transport at 26°N in UKESM1 is significantly weaker than observed at the RAPID-MOCHA array (Johns et al., 2011). This bias in OHT at 26°N is again qualitatively similar to the weak OHT in GC3.1 (Menary et al., 2018).

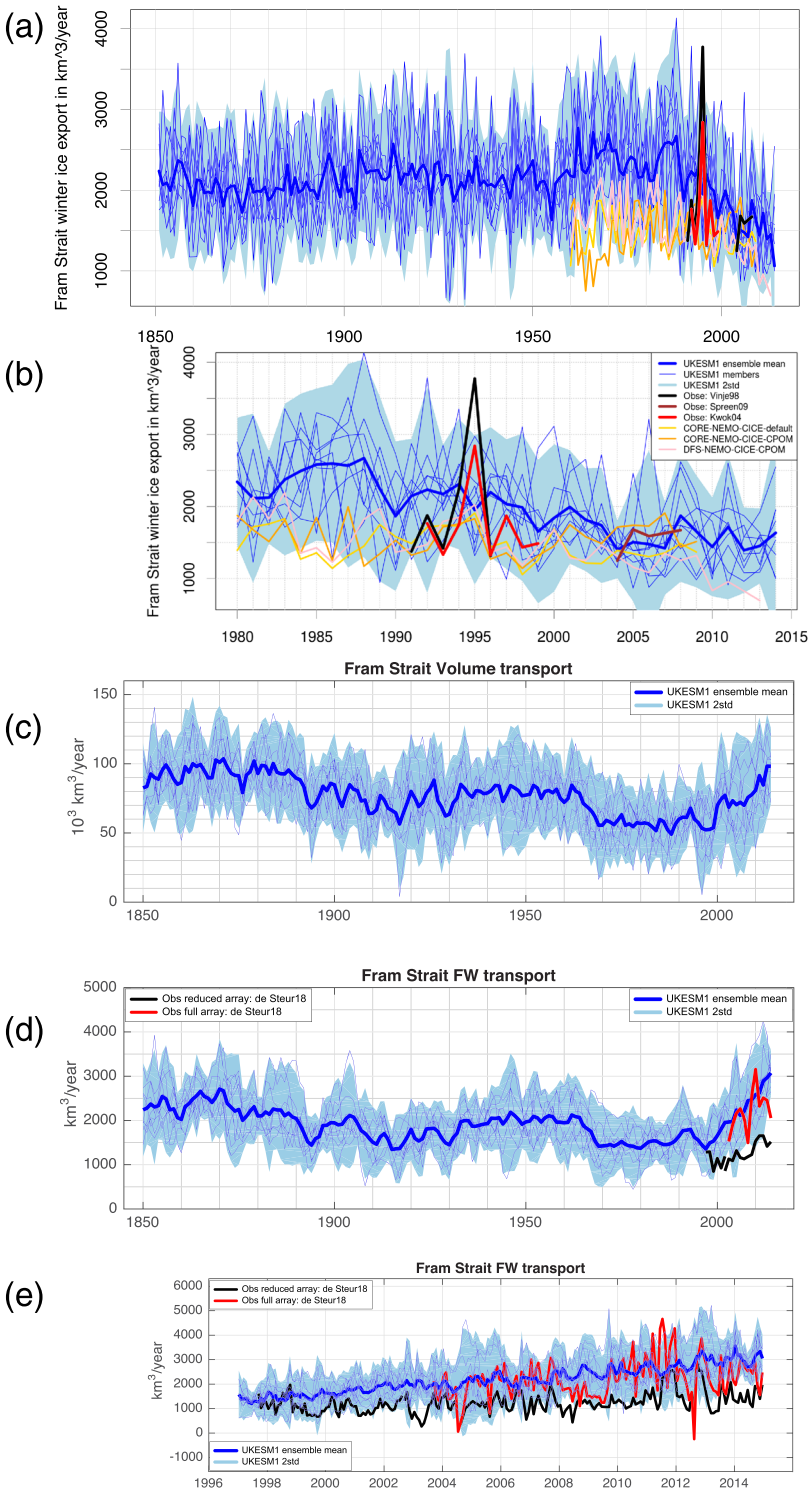


Figure 21. Transport through the Fram Strait at 81°N: (a) Sea ice transport (1850 to 2014), (b) sea ice transport (1980 to 2014), (c) volume transport (1850 to 2014), (d) liquid fresh water transport FW (1850 to 2014), and (e) liquid fresh water transport (1997 to 2014). Export from the Arctic to the Atlantic is positive. UKESM1 results are compared with forced ocean-sea ice simulations and observations (see legends, and section 2.2.5 for more details). Winter sums (October to April) are shown for sea ice transport in (a) and (b) to enable comparison with observations. Annual sums for the volume and liquid freshwater transport in (c) and (e). And monthly transports for liquid freshwater transports in (e). Simulated liquid freshwater transports are referenced to 34.90. Individual model ensemble members are shown in thin blue lines; mean and 2-standard deviation (STD) envelope are shown in solid thick blue and shading.

3.5. Atlantic-Arctic Exchanges and Sea Ice

3.5.1. Sea Ice

Figure 20 shows time series of the mean April sea ice thickness from 1850 to 2014 for the Central Arctic. The UKESM1 ensemble mean thickness increases from about 4 m in 1850 to 5.6 m in 1985. In 1995 a strong decrease starts with a mean value of 3.7 m in 2014. This is an overestimation in Central Arctic sea ice thickness of between 1.5 and 2.5 m with respect to Cryosat-2 and PIOMAS (Zhang & Rothrock, 2003). The difference is much larger than the estimated 0.3 m uncertainty of PIOMAS for monthly mean data averaged over such a large area and given the 2σ ensemble range of the individual UKESM1 members of ~0.6 m. The bias in ice thickness is present the whole year-round and is in accordance with the cold bias in the Northern Hemisphere near-surface air temperature (see Sellar et al., 2019). The overestimation of Arctic sea ice, and the forced variability, is also seen in GC3.1 (Andrews et al., 2020).

Interestingly, the forced NEMO-CICE simulation with the same ocean and sea ice setup as in UKESM1 CORE-NEMO-CICE-default shows a thickness bias of the opposite sign with mean values ranging from ~1.8 to ~2.8 m. This comparison indicates that interaction processes between atmosphere and sea ice play an important role regarding the thickness bias in UKESM1. However, we find that changing the sea ice physics leads to a strong improvement of the simulated ice thickness in forced simulations; DFS-NEMO-CICE-CPOM, which has improved sea ice physics, agrees to Cryosat-2 within range of uncertainty. However, in contrast to UKESM1, the forced ocean-sea ice models appear to underestimate the ice thickness trend in PIOMAS.

In spite of the strong overestimation of ice thickness in the Central Arctic, the winter ice edge in the North Atlantic is only slightly overestimated in UKESM1 (see bottom panel in Figure 20, which shows the mean April sea ice extent for the three adjacent seas: the Barents Sea, East Greenland Sea, and Baffin Bay). The ensemble mean sea ice extent in the North Atlantic region has little trend up to ~1970, but there is considerable simulated internal variability. However, following ~1970 there is a consistent simulated decline in extent that appears similar to the observed decline in the SSM/I Bootstrap data from 1980–2014.

3.5.2. Arctic Freshwater Export

Figure 21a shows time series of winter (October to April) ice export through the Fram Strait in UKESM1 (the same for the period 1990–2014 is shown in Figure 21b). The ensemble mean ice export increases slightly to 1850–~1980. However, it is fairly constant from 1850 to 1950 with export of around 2000 km^3/yr. From 1950 ice export increases up to a maximum of ~2,500 km^3/yr in the 1980s and subsequently decreases to a minimum value in 2012 of ~1,400 km^3/yr. Examining the individual members shows that there is considerable internal variability. The 2σ ensemble spread of order ~1,500 km^3/yr from 1850–1990, but the spread appears to reduce following 1990. All observations fall within the range of the ensemble members, and the large flux observed in 1995 also appears consistent with the ensemble spread. In comparison, the forced NEMO-CICE simulations show a weaker trend with a weaker flux during 1960 to 2000, but similar values after 2000. 70% of the ice export occurs during the period October to April. Thus, the annual UKESM1 ice export decreased from ~3,500 km^3/yr in the 1980s to a minimum value in 2012 of ~2,000 km^3/yr.

Figure 21d also shows time series of volume and liquid freshwater transports through Fram Strait. The UKESM1 ensemble mean annual liquid freshwater transport varies between ~1,500 km^3/yr and ~2,500 km^3/yr on multidecadal time scales. In contrast to the sea ice export, liquid water transport appears to show a slight decrease from 1850–~1990. Furthermore, the consistency of the phasing of the multidecadal variability across the ensemble suggests that external forcings are an important driving factor. After 1995, there is almost a twofold increase in the Fram Strait fresh water transport to ~3,000 km^3/yr, with ensemble mean freshwater export returning to values simulated in ~1850. The post-1995 increase is also consistent with the latest observations (de Steur et al., 2015). It is noteworthy that the total freshwater transport (liquid and frozen) is remarkably constant over time. The reduction in ice export is balanced by an increase of liquid freshwater transport.

In contrast to Fram Strait, freshwater transport through the Davis Strait in UKESM1 does not show any clear multidecadal variability or trend (see Figure S10). Taken together with the results in Figure 21, there is no obvious centennial trend in the exports through both straits. However, there do appear to be opposite trends in transports in the Fram Strait (increasing) and Davis Strait (decreasing) after 1995. This anticorrelation

between the Arctic oceanic outflows east and west of Greenland is also present in the observations (Curry et al., 2014; de Steur et al., 2009; Wang et al., 2016) and has been also noticed in higher-resolution model results (e.g., Aksenov et al., 2010, 2016; de Boer et al., 2018). In both straits the volume transport dominates freshwater transport variations (e.g., consistent with Jahn et al., 2012). It should be noted that the freshwater flux divergence due to run-off from the Greenland shelf, precipitation-evaporation and sea ice and snow melt can alter the freshwater exports through the Nordic Seas (Chafik & Rossby, 2019). However, given the short observational time series, the evaluation of the Nordic Seas freshwater budget is beyond the scope of the present study.

Generally we find Arctic to North Atlantic oceanic volume and freshwater transports simulated with UKESM1 compare well with the available observational estimates and are accurate enough for the model to simulate the historical state of the Arctic-North Atlantic exchanges. The mean volume transport through the Fram Strait in UKESM1 simulated over 1997–2014 of 2.4 ± 0.98 Sv is well within the observational range of 2.2 ± 2.1 Sv (de Boer et al., 2018). Observational estimates of the Fram Strait fresh water transport depend on the reference salinity, whether shelf or Atlantic water component are accounted for, on the layout of the moored instruments, and on the data interpolation strategy (as highlighted by the observed uncertainty in Figures 21d and 21e). Estimates vary between $1{,}753 \pm 713\ km^3\ yr^{-1}$ (Marnela et al., 2013) and $2{,}217 \pm 682\ km^3\ yr^{-1}$ (de Steur et al., 2009). The mean fresh water transport in UKESM1 is $2{,}198 \pm 663\ km^3\ yr^{-1}$ and within the range of that observed. In the Davis Strait the simulated annual mean volume transport for 2004–2010 is is 1.8 ± 0.6 Sv and is close to the observed values of 1.6 ± 0.5 Sv. Corresponding fresh water transports from UKESM1 and observations for 2004–2010 are $2{,}889 \pm 867$ and $2{,}933 \pm 189\ km^3\ yr^{-1}$ (Curry et al., 2014), respectively. It is important to note that Kuhlbrodt et al. (2018) found that northward heat transport through Fram Strait in GC3.1, the physical version of the model, was 1 order of magnitude too small, and we expect a similar bias in UKESM1.

4. Discussion

The analysis presented here is intentionally broad in order to cover a large range of variables across different components of the North Atlantic climate system. Therefore, we have focused on a relatively simple analysis of mean state, and decadal time scale variability and change. Not surprisingly, there are a number of questions that require further investigation. For example, how do changes in different components of the North Atlantic climate system affect each other, and do these changes happen at different time scales? For example, it is well known that SST biases in the Gulf Stream region impact on the wider atmospheric circulation (Keeley et al., 2012; Lee et al., 2018). As a consequence, to what extent do these SST biases affect the errors in jet latitude and speed in UKESM1, especially in summer? What is the impact of ozone overestimation on the distribution of sulfate aerosols and, hence, AOD over the North Atlantic; in particular, does the overestimation in ozone in the tropics contribute to the relatively low AOD there, or is this bias related to dust? Why are Arctic-Atlantic ice and freshwater transports through Fram Strait compensating, and how are they related to the changes in the Atlantic Ocean further south? What controls the bias in the intraseasonal distribution of SSWs, and is this related in any way to the underestimation of variability in the winter (i.e., DJF) NAO?

We also have not explored the relative role of different forcing factors in controlling changes across the North Atlantic climate system in UKESM1, including external forcings (greenhouse gases, aerosols, etc.) and changes outside the North Atlantic region (e.g., the Pacific). However, the similarity between the temporal evolution in basin mean AOD and many variables does suggest that anthropogenic aerosol emissions play an important role in decadal-to-multidecadal time scale change across the North Atlantic climate system in UKESM1. This role of anthropogenic aerosols is consistent with previous modeling studies (Bellomo et al., 2018; Booth et al., 2012; Menary et al., 2013, 2020; Watanabe & Tatebe, 2019). However, it is important to recognize that there continues to be debate on the relative role of aerosols in driving recent changes in the North Atlantic, and in AMV (Kim, Yeager & Danabasoglu, 2018; Robson et al., 2016; Yan et al., 2019; Zhang et al., 2013). There are also clearly issues in the simulation of multidecadal variability and long-term change in UKESM1; in particular, the simulation of subpolar surface salinity does not appear to agree with observations (i.e., Friedman et al., 2017) and needs further investigation. Furthermore, some of the mean-state deficiencies highlighted here will impact on the simulation of AMV. For example, cloud

biases in the tropical North Atlantic will likely affect cloud feedbacks to natural and forced variability (e.g., Brown et al., 2016) and the strength of the forced response (Wilcox et al., 2015). Surface temperatures in the Northern Hemisphere in UKESM1 historical simulations are also known to cool excessively as anthropogenic aerosol forcing increased through the 1950s–1980s (Sellar et al., 2019), which undoubtedly projects onto AMV. In this context, the analysis showing that UKESM1 appears to overestimate recent Atlantic basin mean changes in SWTOA (i.e., Figure 12) is particularly interesting and warrants further study. Specifically, is the difference in recent SWTOA trends a model bias (and if so what processes need improving), or does it represent a large role for internal variability that is not captured by a limited ensemble? Either way, further in depth process-based analysis is key to making progress.

It is important to recognize that many of the deficiencies in UKESM1 identified here are consistent with physical climate model biases, especially seen for models of similar resolution. Indeed, many of the biases identified in UKESM1 are also seen in GC3.1, the physical version of the model which forms the basis of UKESM1. For example, SST biases in the Gulf Stream and Grand Banks region (Kuhlbrodt et al., 2018), the overestimation of SSW in early winter (Menary et al., 2018), and too thick Arctic sea ice (Andrews et al., 2020), are all seen in the physical climate model. The externally forced changes in many variables across the Atlantic is also similar between GC3.1 and UKESM1, and the forced signal also exhibits similar deficiencies, including a forced AMV signal that is too "global" (Andrews et al., 2020). It is, therefore, important to emphasize that continued development of the physical core model, as well as of the Earth system component models, remains crucial for delivering reliable projections of future climate and Earth system change over the North Atlantic. One potential route for improving some systematic biases may be to move to higher resolution, where improvements in some key biases are seen, including in low-frequency NAO variability (Menary et al., 2018), SSTs and the dynamics of the Gulf Stream (Roberts et al., 2019). However, higher-resolution models do not improve all aspects of the simulation of the current climate (Menary et al., 2018; Roberts et al., 2019), and continued development of Earth system components is also required, both to explore a broader range of future system responses and to investigate the efficacy of future climate mitigation policies.

5. Conclusions

In this paper we have undertaken a broad multivariate evaluation of the North Atlantic climate system in the UK's Earth System Model (UKESM1). We have focused on the comparison of UKESM1's CMIP6 historical experiments with observations across different components of the North Atlantic climate system, including atmosphere dynamics; atmospheric composition; clouds and radiation; ocean state and dynamics; Arctic-Atlantic exchanges; and sea ice. Where possible, we have also compared with published results from the physical model on which UKESM1 is based, HadGEM3-GC3.1-LL (GC3.1). The main results are as follows:

- The general spatial structure of the NAO compares well between UKESM1 and reanalyses (20CRv2c) in both winter (DJF) and summer (JJA). However, in both winter and summer the spatial pattern of the NAO tends to be too zonal, consistent with the North Atlantic Current also being too zonal (Menary et al., 2018). Furthermore, the NAO centers of action are largely confined to the North Atlantic Ocean, rather than extending over Northern Europe. Generally, the observed NAO indices fall within the ensemble spread, but there is little evidence for forced variability in the NAO.
- The mean state of the North Atlantic eddy-driven jet is generally well represented in UKESM1 in winter (DJF) in terms of jet latitude and speed. UKESM1 also generally captures the trimodal structure of jet latitude variance, although it appears to overestimate the variability between 50°N and 55°N.
- There is less consistency for the North Atlantic eddy-driven jet in summer (JJA). During the 1979–2014 period the jet is weaker and more poleward than in reanalysis. The reanalysis does fall within the ensemble spread, however.
- The number of SSWs in UKESM1 compares well with ERA-Interim over the extended winter season (November–March). However, there is significant bias within the winter season, with too many SSWs in early winter (i.e., November), and too few in midwinter (i.e., January). The overestimation in November SSWs is also seen in the physical base model, GC3.1 (Menary et al., 2018), and appears related to an early formation of the polar vortex and an overestimation of vertical propagation of planetry waves in August to October.

- Tropospheric column ozone over the North Atlantic is overestimated in UKESM1, particularly over the tropical North Atlantic. Basin mean tropospheric column ozone increases from ~29 to ~38 DU over the period 1850–2014. However, basin mean values in UKESM1 are ~4.7 DU larger than satellite-derived observations. These biases are likely to lead to an overestimate of the oxidizing capacity of the atmosphere that will likely mean that the lifetime of gases like SO_2 are too short over the North Atlantic Basin (assuming all other relevant processes are well represented). These tropospheric ozone changes will not be simulated in GC3.1 due to its use of prescribed oxidants.
- AOD averaged over the North Atlantic basin peaks around 1980 in UKESM1 and decreases thereafter (neglecting the influence of the Pinatubo eruption). Over the period 1995–2014, UKESM1 produces qualitatively similar behavior to an ensemble of satellite observations. Spatially, the model produces a realistic pattern of AOD compared to observations, and the pattern is also similar to the physical version of the model (i.e., Mulcahy et al., 2018). However, UKESM1 may be dispersing dust more widely over the basin than observed.
- UKESM1 appears to simulate the observed cloud droplet concentration (N_d) over the North Atlantic well, including the mean and the trend over 2003–2014. However, the basin mean N_d masks significant regional biases in the mean state, particularly in the tropical North Atlantic which is overestimated in the mean and has too weak a trend. Changes in basin mean N_d over 1850–2014 generally follows the changes in AOD.
- The basin mean outgoing shortwave flux at the top of the atmosphere (SWTOA) is overestimated in comparison to observations over 2001–2014 by ~4 W m^{-2}. The basin mean bias is dominated by a too bright North Atlantic storm track region. This is a region with a positive total cloud cover bias, which is suggested to be the cause in Grosvenor and Carslaw (2020). However, the tropical North Atlantic has a negative total cloud cover bias.
- Over 2001–2014 UKESM1 simulates a declining trend in Atlantic basin mean SWTOA that appear to follow changes in AOD and N_d. However, the ensemble mean trend is significantly larger than that observed, and the observed trend is marginally outside the range of the nine-member ensemble.
- The UKESM1 historical simulations display some significant ocean biases. In the North Atlantic UKESM1 is generally too cold and saline at the surface, except for the subpolar North Atlantic (SPNA), where it is both cold and fresh, (consistent with long running biases in low-resolution climate models such as GC3.1, cf. Kuhlbrodt et al., 2018). However, the subsurface ocean, that is, the entire layer 0–700 m, is generally too warm and saline compared to observations. Therefore, these biases suggest that the ocean stratification may be too weak.
- Over the twentieth century, UKESM1 simulates significant multidecadal variability in basin mean SST, which is similar to the observed AMV (Figure 16), and the variability simulated in GC3.1 (Andrews et al., 2020). This variability is consistent with a forced response due to changes in anthropogenic aerosols (Booth et al., 2012). However, the ensemble mean AMV in UKESM1 has a lower amplitude and a slight shift in phase compared to the observed AMV. Furthermore, in UKESM1 the spatial pattern of AMV is different to observations; particularly, it does not capture the pattern in the subpolar North Atlantic and the pattern is related to changes across the entire Northern Hemisphere, as similar to GC3.1 (Andrews et al., 2020).
- There is also a forced response in the SPNA upper ocean (0–700 m) heat content and salinity. SPNA-mean upper ocean heat content changes are in phase with the simulated AMV after 1951, but weaker than those observed. However, over the wider-Atlantic basin heat content trends, particularly over 1951–1980, are substantially larger than observed. Upper ocean and surface salinity increases almost monotonically from 1850–2014 in UKESM1 historical simulations. However, the long-term changes in SPNA surface salinity appear to be opposite to long-term trends suggested by observations (i.e., Friedman et al., 2017)
- The AMOC and northward heat transport at 26°N shows a forced upward trend for much of the twentieth century in UKESM1, before starting to decline around 1990. The increase in AMOC is also consistent with a forced response due to changes in anthropogenic aerosols (Menary et al., 2013, 2020). By the end of the twentieth century the AMOC is comparable to the observed AMOC strength (as measured by RAPID), but the northward heat transport is too low. The decline in AMOC in UKESM1 after 1990 is comparable in sign to, but much weaker than, the observed decline after 2004. However, significant observational uncertainties (Sinha et al., 2018) and the fact that the UKESM1 underestimates the observed AMOC variability, make direct comparisons difficult. The simulation of AMOC (including mean state and forced trend) is

generally similar to that found in GC3.1, although the mean AMOC and mean northward heat transport at 26°N is stronger in GC3.1 (Andrews et al., 2020; Menary et al., 2018).

- Sea ice thickness in the central Arctic is overestimated significantly in UKESM1 and is between 1 and 2 m thicker than in observations (CryoSat-2) or observationally constrained data sets (PIOMAS). There is also a large forced response in thickness, which increases up to the 1980s, and decreases rapidly thereafter. However, sea ice extent in the North Atlantic region compares more favorably with observations, although the ensemble mean is generally larger than observations. The general overestimation of sea ice volume, and the forced response, is largely consistent with GC3.1 (Andrews et al., 2020).
- Winter sea ice flux through Fram Strait and annual mean volume and freshwater fluxes through both Fram and Davis Strait compare well with observed estimates. There is considerable multidecadal forced variability in Fram Strait freshwater and sea ice fluxes, but the variability appears to be compensating. Fram Strait freshwater transport shows a slight decrease up until the ∼1990s, but increases rapidly thereafter. Fram strait winter sea ice flux increases to the∼1990s before decreasing. The latter period, post 1990 is consistent with observations.

Overall UKESM1 appears to simulate many general aspects of the North Atlantic climate system well across a range of different components. This includes, but is not limited to, the location and strength of the atmospheric jet in winter; the spatial pattern and basin mean AOD; the basin mean and spatial changes in cloud droplet concentration; multidecadal SST variability; and Atlantic sector sea ice extent and sea ice export through the Fram Strait. Therefore, we believe UKESM1 is a useful tool for understanding linkages between different components of the North Atlantic climate system, both across domains of the coupled physical climate and between physical climate and Earth system phenomena. In addition to understanding such interactions, UKESM1 will also be valuable for exploring past and potential future changes.

Nevertheless, UKESM1 does suffer from significant deficiencies which need to be considered. For example, Arctic sea ice is too thick; the summer jet is too weak and positioned too far poleward; and tropospheric ozone column is significantly overestimated. The implications of all these deficiencies for model predictions across the wider North Atlantic climate system, or for the attribution of past changes in the North Atlantic, is still not currently known. Therefore, there is still a need to understand the causes of the deficiencies and to understand the implications of these shortcomings. Identification of such systematic biases, along with the underpinning cause of these biases, provides important pointers for future development of UKESM models including their physical core, the HadGEM family of models.

Finally, although we highlight some deficiencies in UKESM1, it is important to emphasize that the representation of the North Atlantic climate system is largely comparable to GC3.1, the physical climate model that forms the basis of UKESM1. Given UKESM1 was specifically developed to include as many prognostic cross-domain and cross-phenomena couplings as possible (e.g., Archibald et al., 2019; Mulcahy et al., 2018; O'Connor et al., 2020; Sellar et al., 2019), and thereby increasing the model degrees of freedom, this similarity of performance is encouraging. A priority for future development of UKESM1 is to better constrain such cross-domain and phenomena coupling, using observations and detailed process models. Earth system models, such as UKESM1, are important tools for understanding the complex array of interactions within the North Atlantic climate system, and will be critical for developing robust scenarios of how the coupled North Atlantic may change in the future.

Data Availability Statements

The UKESM1 climate model simulations used in this study are available on the CMIP6 archive. EN4 data are available at this site (https://www.metoffice.gov.uk/hadobs/en4/). HadISST data are available at this site (https://www.metoffice.gov.uk/hadobs/hadisst/). RAPID-MOCHA data are available here (http://www.rapid.ac.uk/rapidmoc/). The Twentieth Century Reanalysis data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site (https://www.esrl.noaa.gov/psd/). Satellite-based data sets for aerosol optical depth described in section 2.2.2 can be download from these sites (https://search.earthdata.nasa.gov/search and https://cci.esa.int/data). The cloud droplet concentration data set is available from this site (https://catalogue.ceda.ac.uk/uuid/cf97ccc802d348ec8a3b6f2995dfbbff). CERES-EBAF data product is available at this site (https://ceres.larc.nasa.gov/order_data.php). The CALIPSO-GOCCP data is available at this site (https://climserv.ipsl.polytechnique.fr/cfmip-obs/index.html).

Acknowledgments

We thank Léon Chafik and an anonymous reviewer for their comments that helped improve the clarity of the paper. This article was written with support from National Environmental Research Council (NERC) national capability grant for the North Atlantic Climate System: Integrated study (ACSIS) program (Grants NE/N018001/1, NE/N018044/1, NE/N018028/1, and NE/N018052/1). C. J., T. K., and A. P. were funded by the NERC national capability grant for the UK Earth System Modelling project (Grant NE/N017951/1). C. J. and T. K. were also supported by the EU Horizon 2020 Research Programme CRESCENDO project, Grant Agreement 641816. A. P. was funded as part of NERC's support of the National Centre for Earth Observation, contract PR140015. Y. A. was additionally supported by the NERC APEAR project (Grant NE/R012865/1) as part of the Changing Arctic Ocean program, the NERC UK Earth System Modelling project (Grant NE/N018036/1), and the EU Horizon 2020 Reserach Program COMFORT project (Grant Agreement 820989). A. S. and J. W. were supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra. The work reflects only the authors' view; the European Commission and their executive agency are not responsible for any use that may be made of the information the work contains. Support for the Twentieth Century Reanalysis Project data set is provided by the U. S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office.

References

Ahn, E., Huang, Y., Siems, S. T., & Manton, M. J. (2018). A comparison of cloud microphysical properties derived from MODIS and CALIPSO with in situ measurements over the wintertime Southern Ocean. *Journal of Geophysical Research: Atmospheres*, *123*, 11,120–11,140. https://doi.org/10.1029/2018JD028535

Aksenov, Y., Bacon, S., Coward, A. C., & Holliday, N. P. (2010). Polar outflow from the Arctic Ocean: A high resolution model study. *Journal of Marine Systems*, *83*(1), 14–37. https://doi.org/10.1016/j.jmarsys.2010.06.007

Aksenov, Y., Karcher, M., Proshutinsky, A., Gerdes, R., de Cuevas, B., Golubeva, E., et al. (2016). Arctic pathways of Pacific Water: Arctic Ocean model intercomparison experiments. *Journal of Geophysical Research: Oceans*, *121*, 27–59. https://doi.org/10.1002/2015JC011299

Alexander-Turner, R., Ortega, P., & Robson, J. I. (2018). How robust are the surface temperature fingerprints of the Atlantic overturning meridional circulation on monthly time scales? *Geophysical Research Letters*, *45*, 3559–3567. https://doi.org/10.1029/2017GL076759

Allen, R. J., Evan, A. T., & Booth, B. B. B. (2015). Interhemispheric aerosol radiative forcing and tropical precipitation shifts during the late twentieth century. *Journal of Climate*, *28*(20), 8219–8246. https://doi.org/10.1175/JCLI-D-15-0148.1

Andrews, M. B., Knight, J. R., Scaife, A. A., Lu, Y., Wu, T., Gray, L. J., & Schenzinger, V. (2019). Observed and simulated teleconnections between the stratospheric quasi-biennial oscillation and Northern Hemisphere winter atmospheric circulation. *Journal of Geophysical Research: Atmospheres*, *124*, 1219–1232. https://doi.org/10.1029/2018JD029368

Andrews, M., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., et al. (2020). Historical simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001995. https://doi.org/10.1029/2019MS001995

Archibald, A. T., O'Connor, F. M., Abraham, N. L., Archer-Nicholls, S., Chipperfield, M. P., Dalvi, M., et al. (2019). Description and evaluation of the UKCA stratosphere-troposphere chemistry scheme (StratTrop vn 1.0) implemented in UKESM1. *Geoscientific Model Development Discussions*, *2019*, 1–82. https://doi.org/10.5194/gmd-2019-246

Baldwin, M. P., Ayarzagüena, B., Birner, T., Butchart, N., Charlton-Perez, A. J., Butler, A. H., et al. (2020). Sudden stratospheric warmings. *Earth and Space Science Open Archive*, 49. https://doi.org/10.1002/essoar.10502884.1

Baldwin, M. P., & Dunkerton, T. J. (2001). Stratospheric harbingers of anomalous weather regimes. *Science*, *294*(5542), 581–584.

Becker, K.-H., Barnes, I., Ruppert, L., & Wiesen, P. (1999). Is the oxidising capacity of the troposphere changing? In D. Möller (Ed.), *Atmospheric Environmental Research* (pp. 105–132). Berlin, Heidelberg: Springer.

Belkin, I. M., Levitus, S., Antonov, J., & Malmberg, S.-A. (1998). "Great salinity anomalies" in the North Atlantic. *Progress in Oceanography*, *41*(1), 1–68 en. https://doi.org/10.1016/S0079-6611(98)00015-9

Bellomo, K., Clement, A. C., Murphy, L. N., Polvani, L. M., & Cane, M. A. (2016). New observational evidence for a positive cloud feedback that amplifies the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, *43*, 9852–9859. https://doi.org/10.1002/2016GL069961

Bellomo, K., Murphy, L. N., Cane, M. A., Clement, A. C., & Polvani, L. M. (2018). Historical forcings as main drivers of the Atlantic multidecadal variability in the CESM large ensemble. *Climate Dynamics*, *50*(9-10), 3687–3698. https://doi.org/10.1007/s00382-017-3834-3

Bevan, S. L., North, P. R. J., Los, S. O., & Grey, W. M. F. (2012). A global dataset of atmospheric aerosol optical depth and surface reflectance from AATSR. *Remote Sensing of Environment*, *116*, 199–210. Advanced Along Track Scanning Radiometer(AATSR) Special Issue) https://doi.org/10.1016/j.rse.2011.05.024

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., et al. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., & Bellouin, N. (2012). Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, *484*(7393), 228.

Boucher, O., Moulin, C., Belviso, S., Aumont, O., Bopp, L., Cosme, E., et al. (2003). DMS atmospheric concentrations and sulphate aerosol indirect radiative forcing: A sensitivity study to the DMS source representation and oxidation. *Atmospheric Chemistry and Physics*, *3*(1), 49–65. https://doi.org/10.5194/acp-3-49-2003

Bracegirdle, T. J., Lu, H., Eade, R., & Woollings, T. (2018). Do CMIP5 models reproduce observed low-frequency North Atlantic jet variability? *Geophysical Research Letters*, *45*, 7204–7212. https://doi.org/10.1029/2018GL078965

Brown, P. T., Lozier, M. S., Zhang, R., & Li, W. (2016). The necessity of cloud feedback for a basin-scale Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, *43*, 3955–3963. https://doi.org/10.1002/2016GL068303

Buckley, M. W., & Marshall, J. (2016). Observations, inferences, and mechanisms of the Atlantic meridional overturning circulation: A review. *Reviews of Geophysics*, *54*, 5–63. https://doi.org/10.1002/2015RG000493

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., & Saba, V. (2018). Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature*, *556*(7700), 191.

Chafik, L., Nilsen, J. E. O., Dangendorf, S., Reverdin, G., & Frederikse, T. (2019). North Atlantic Ocean circulation and decadal sea level change during the altimetry era. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-018-37603-6

Chafik, L., & Rossby, T. (2019). Volume, heat, and freshwater divergences in the subpolar North Atlantic suggest the Nordic Seas as key to the state of the meridional overturning circulation. *Geophysical Research Letters*, *46*, 4799–4808. https://doi.org/10.1029/2019GL082110

Chang, C.-Y., Chiang, J. C. H., Wehner, M. F., Friedman, A. R., & Ruedy, R. (2011). Sulfate aerosol control of tropical Atlantic climate over the twentieth century. *Journal of Climate*, *24*(10), 2540–2555. https://doi.org/10.1175/2010JCLI4065.1

Chanut, J., Barnier, B., Large, W., Debreu, L., Penduff, T., Molines, J. M., & Mathiot, P. (2008). Mesoscale eddies in the Labrador Sea and their contribution to convection and restratification. *Journal of Physical Oceanography*, *38*(8), 1617–1643. https://doi.org/10.1175/2008JPO3485.1

Charney, J. G., & Drazin, P. G. (1961). Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, *66*(1), 83–109. https://doi.org/10.1029/JZ066i001p00083

Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., et al. (2010). The GCM-oriented CALIPSO cloud product (CALIPSO-GOCCP). *Journal of Geophysical Research*, *115*, D00H16. https://doi.org/10.1029/2009JD012251

Cohen, J., Barlow, M., & Saito, K. (2009). Decadal fluctuations in planetary wave forcing modulate global warming in late boreal winter. *Journal of Climate*, *22*(16), 4418–4426. https://doi.org/10.1175/2009JCLI2931.1

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., et al. (2011). Development and evaluation of an Earth-system model HadGEM2. *Geoscientific Model Development*, *4*(4), 1051–1075. https://doi.org/10.5194/gmd-4-1051-2011

Comiso, J. (2017). *Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, version* (Vol. 3). Boulder, CO: NASA National Snow and Ice Data Center Distributed Active Archive Center.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Cunningham, S. A., Roberts, C. D., Frajka-Williams, E., Johns, W. E., Hobbs, W., Palmer, M. D., et al. (2013). Atlantic meridional overturning circulation slowdown cooled the subtropical ocean. *Geophysical Research Letters*, *40*, 6202–6207. https://doi.org/10.1002/2013GL058464

Curry, B., Lee, C. M., Petrie, B., Moritz, R. E., & Kwok, R. (2014). Multiyear volume, liquid freshwater, and sea ice transports through Davis Strait, 2004-10. *Journal of Physical Oceanography*, *44*(4), 1244–1266. https://doi.org/10.1175/JPO-D-13-0177.1

de Boer, A. M., Gavilan Pascual-Ahuir, E., Stevens, D. P., Chafik, L., Hutchinson, D. K., Zhang, Q., et al. (2018). Interconnectivity between volume transports through Arctic Straits. *Journal of Geophysical Research: Oceans*, *123*, 8714–8729. https://doi.org/10.1029/2018JC014320

de Steur, L., Hansen, E., Gerdes, R., Karcher, M., Fahrbach, E., & Holfort, J. (2009). Freshwater fluxes in the East Greenland current: A decade of observations. *Geophysical Research Letters*, *36*, L23611. https://doi.org/10.1029/2009GL041278

de Steur, L., Peralta-Ferriz, C., & Pavlova, O. (2018). Freshwater export in the East Greenland current freshens the North Atlantic. *Geophysical Research Letters*, *45*, 13,359–13,366. https://doi.org/10.1029/2018GL080207

de Steur, L., Pickart, R. S., Torres, D. J., & Valdimarsson, H. (2015). Recent changes in the freshwater composition east of Greenland. *Geophysical Research Letters*, *42*, 2326–2332. https://doi.org/10.1002/2014GL062759

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Delworth, T. L., Zeng, F., Zhang, L., Zhang, R., Vecchi, G. A., & Yang, X. (2017). The central role of ocean dynamics in connecting the North Atlantic oscillation to the extratropical component of the Atlantic multidecadal oscillation. *Journal of Climate*, *30*(10), 3789–3805. https://doi.org/10.1175/JCLI-D-16-0358.1

Deser, C., Tomas, R., Alexander, M., & Lawrence, D. (2010). The seasonal atmospheric response to projected Arctic sea ice loss in the late twenty-first century. *Journal of Climate*, *23*(2), 333–351. https://doi.org/10.1175/2009JCLI3053.1

Dickson, R. R., Meincke, J., Malmberg, S.-A., & Lee, A. J. (1988). The "great salinity anomaly" in the Northern North Atlantic 1968–1982. *Progress in Oceanography*, *20*(2), 103–151. https://doi.org/10.1016/0079-6611(88)90049-3

Dong, B., Sutton, R. T., Woollings, T., & Hodges, K. (2013). Variability of the NorthAtlantic summer storm track: Mechanisms and impacts on European climate. *Environmental Research Letters*, *8*(3), 34037. https://doi.org/10.1088/1748-9326/8/3/034037

Drijfhout, S., van Oldenborgh, G. J., & Cimatoribus, A. (2012). Is a decline of AMOC causing the warming hole above the North Atlantic in observed and modeled warming patterns? *Journal of Climate*, *25*(24), 8373–8379. https://doi.org/10.1175/JCLI-D-12-00490.1

Duchez, A., Courtois, P., Harris, E., Josey, S. A., Kanzow, T., Marsh, R., et al. (2016). Potential for seasonal prediction of Atlantic sea surface temperatures using the RAPID array at 26°N. *Climate Dynamics*, *46*(9–10), 3351–3370. https://doi.org/10.1007/s00382-015-2918-1

Dussin, R., Barnier, B., Brodeau, L., & Molines, J. M. (2016). The making of the DRAKKAR forcing set DFS5. DRAKKAR (*MyOcean report 01-04-16*): LGGE, Grenoble, France.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Friedman, A. R., Reverdin, G., Khodri, M., & Gastineau, G. (2017). A new record of Atlantic sea surface salinity from 1896 to 2013 reveals the signatures of climate variability and long-term trends. *Geophysical Research Letters*, *44*, 1866–1876. https://doi.org/10.1002/2017GL072582

Gillett, N. P., Allen, M. R., McDonald, R. E., Senior, C. A., Shindell, D. T., & Schmidt, G. A. (2002). How linear is the Arctic Oscillation response to greenhouse gases? *Journal of Geophysical Research*, *107*(D3), ACL–1. https://doi.org/10.1029/2001JD000589

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*, 6704–6716. https://doi.org/10.1002/2013JC009067

Gouretski, V., & Reseghetti, F. (2010). On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep Sea Research Part I: Oceanographic Research Papers*, *57*(6), 812–833. https://doi.org/10.1016/j.dsr.2010.03.011

Gray, L. J., Scaife, A. A., Mitchell, D. M., Osprey, S., Ineson, S., Hardiman, S., et al. (2013). A lagged response to the 11 year solar cycle in observed winter Atlantic/European weather patterns. *Journal of Geophysical Research: Atmospheres*, *118*, 13,405–13,420. https://doi.org/10.1002/2013JD020062

Gray, L. J., Woollings, T. J., Andrews, M., & Knight, J. (2016). Eleven-year solar cycle signal in the NAO and Atlantic/European blocking. *Quarterly Journal of the Royal Meteorological Society*, *142*(698), 1890–1903. https://doi.org/10.1002/qj.2782

Grosvenor, D. P., & Carslaw, K. S. (2020). How does the UKESM1 climate model produce its cloud-aerosol forcing in the North Atlantic? *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–57. https://doi.org/10.5194/acp-2020-502

Grosvenor, D. P., Sourdeval, O., & Wood, R. (2018). Parameterizing cloud top effective radii from satellite retrieved values, accounting for vertical photon transport: Quantification and correction of the resulting bias in droplet concentration and liquid water path retrievals. *Atmospheric Measurement Techniques*, *11*(7), 4273–4289. https://doi.org/10.5194/amt-11-4273-2018

Grosvenor, D. P., Sourdeval, O., Zuidema, P., Ackerman, A., Alexandrov, M. D., Bennartz, R., et al. (2018). Remote sensing of droplet number concentration in warm clouds: A review of the current state of knowledge and perspectives. *Reviews of Geophysics*, *56*, 409–453.

Grosvenor, D. P., & Wood, R. (2014). The effect of solar zenith angle on MODIS cloud optical and microphysical retrievals within marine liquid water clouds. *Atmospheric Chemistry and Physics*, *14*(14), 7291–7321. https://doi.org/10.5194/acp-14-7291-2014

Häkkinen, S., Rhines, P. B., & Worthen, D. L. (2011). Atmospheric Blocking and Atlantic multidecadal ocean variability. *Science*, *334*(6056), 655–659. https://doi.org/10.1126/science.1205683

Haase, S., & Matthes, K. (2019). The importance of interactive chemistry for stratosphere–troposphere coupling. *Atmospheric Chemistry and Physics*, *19*(5), 3417–3432. https://doi.org/10.5194/acp-19-3417-2019

Häkkinen, S., & Rhines, P. B. (2004). Decline of subpolar North Atlantic circulation during the 1990s. *Science*, *304*(5670), 555–559. https://doi.org/10.1126/science.1094917

Harper, A. B., Cox, P. M., Friedlingstein, P., Wiltshire, A. J., Jones, C. D., Sitch, S., et al. (2016). Improved representation of plant functional types and physiology in the joint UK land environment simulator (JULES v4.2) using plant trait information. *Geoscientific Model Development*, *9*(7), 2415–2440. https://doi.org/10.5194/gmd-9-2415-2016

Harper, A. B., Wiltshire, A. J., Cox, P. M., Friedlingstein, P., Jones, C. D., Mercado, L. M., et al. (2018). Vegetation distribution and terrestrial carbon cycle in a carbon cycle configuration of JULES4.6 with new plant functional types. *Geoscientific Model Development*, *11*(7), 2857–2873. https://doi.org/10.5194/gmd-11-2857-2018

Haywood, J. M., Osborne, S. R., & Abel, S. J. (2004). The effect of overlying absorbing aerosol layers on remote sensing retrievals of cloud effective radius and cloud optical depth. *Quarterly Journal of the Royal Meteorological Society*, *130*(598), 779–800. https://doi.org/10.1256/qj.03.100

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., et al. (2018). Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS). *Geoscientific Model Development*, *11*(1), 369–408. https://doi.org/10.5194/gmd-11-369-2018

Holliday, N. P., Bacon, S., Cunningham, S. A., Gary, S. F., Karstensen, J., King, B. A., et al. (2018). Subpolar North Atlantic overturning and gyre-scale circulation in the summers of 2014 and 2016. *Journal of Geophysical Research: Oceans*, *123*, 4538–4559. https://doi.org/10.1029/2018JC013841

Holliday, N. P., Bersch, M., Berx, B., Chafik, L., Cunningham, S., Florindo-López, C., et al. (2020). Ocean circulation causes the largest freshening event for 120 years in eastern subpolar North Atlantic. *Nature Communications*, *11*(1), 1–15. https://doi.org/10.1038/s41467-020-14474-y

Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, *269*(5224), 676–679. https://doi.org/10.1126/science.269.5224.676

Inoue, J., & Kikuchi, T. (2007). Outflow of summertime Arctic sea ice observed by ice drifting buoys and its linkage with ice reduction and atmospheric circulation patterns. *Journal of the Meteorological Society of Japan. Ser. II*, *85*(6), 881–887. https://doi.org/10.2151/jmsj.85.881

Iqbal, W., Leung, W.-N., & Hannachi, A. (2018). Analysis of the variability of the North Atlantic eddy-driven jet stream in CMIP5. *Climate Dynamics*, *51*(1), 235–247. https://doi.org/10.1007/s00382-017-3917-1

Jahn, A., Aksenov, Y., de Cuevas, B. A., de Steur, L., Häkkinen, S., Hansen, E., et al. (2012). Arctic Ocean freshwater: How robust are model simulations? *Journal of Geophysical Research*, *117*, C00D16. https://doi.org/10.1029/2012JC007907

Johns, W. E., Baringer, M. O., Beal, L. M., Cunningham, S. A., Kanzow, T., Bryden, H. L., et al. (2011). Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N. *Journal of Climate*, *24*(10), 2429–2449. https://doi.org/10.1175/2010JCLI3997.1

Josey, S. A., Hirschi, J. J.-M., Sinha, B., Duchez, A., Grist, J. P., & Marsh, R. (2018). The recent Atlantic cold anomaly: Causes, consequences, and related phenomena. *Annual Review of Marine Science*, *10*(1), 475–501. https://doi.org/10.1146/annurev-marine-121916-063102

Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. J., Loeb, N. G., Doelling, D. R., et al. (2018). Surface irradiances of edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product. *Journal of Climate*, *31*(11), 4501–4527. https://doi.org/10.1175/JCLI-D-17-0523.1

Keeley, S. P. E., Sutton, R. T., & Shaffrey, L. C. (2012). The impact of North Atlantic sea surface temperature errors on the simulation of North Atlantic European region climate. *Quarterly Journal of the Royal Meteorological Society*, *138*(668), 1774–1783. https://doi.org/10.1002/qj.1912

Kerr, R. A. (2000). A North Atlantic climate pacemaker for the centuries. *Science*, *288*(5473), 1984–1985. https://doi.org/10.1126/science.288.5473.1984

Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2018). Low-frequency North Atlantic climate variability in the community earth system model large ensemble. *Journal of Climate*, *31*(2), 787–813. https://doi.org/10.1175/JCLI-D-17-0193.1

Kim, W. M., Yeager, S. G., & Danabasoglu, G. (2018). Key role of internal ocean dynamics in Atlantic multidecadal variability during the last half century. *Geophysical Research Letters*, *45*, 13,449–13,457. https://doi.org/10.1029/2018GL080474

King, M. D., Platnick, S., Yang, P., Arnold, G. T., Gray, M. A., Riedi, J. C., et al. (2004). Remote sensing of liquid water and ice cloud optical thickness and effective radius in the Arctic: Application of airborne multispectral MAS data. *Journal of Atmospheric and Oceanic Technology*, *21*(6), 857–875. https://doi.org/10.1175/1520-0426(2004)021<0857:RSOLWA>2.0.CO;2

Kolmonen, P., Sogacheva, L., Virtanen, T. H., de Leeuw, G., & Kulmala, M. (2016). The ADV/ASV AATSR aerosol retrieval algorithm: Current status and presentation of a full-mission AOD dataset. *International Journal of Digital Earth*, *9*(6), 545–561. https://doi.org/10.1080/17538947.2015.1111450

Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., & Rahmstorf, S. (2007). On the driving processes of the Atlantic meridional overturning circulation. *Reviews of Geophysics*, *45*, RG2001. https://doi.org/10.1029/2004RG000166

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., et al. (2018). The low-resolution version of HadGEM3 GC3.1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, *10*, 2865–2888. https://doi.org/10.1029/2018MS001370

Kwok, R., Cunningham, G. F., & Pang, S. S. (2004). Fram Strait sea ice outflow. *Journal of Geophysical Research*, *109*, C01009. https://doi.org/10.1029/2003JC001785

Large, W. G., & Yeager, S. G. (2009). The global climatology of an interannually varying air–sea flux data set. *Climate Dynamics*, *33*(2), 341–364. https://doi.org/10.1007/s00382-008-0441-3

Laxon, S. W., Giles, K. A., Ridout, A. L., Wingham, D. J., Willatt, R., Cullen, R., et al. (2013). CryoSat-2 estimates of Arctic sea ice thickness and volume. *Geophysical Research Letters*, *40*, 732–737. https://doi.org/10.1002/grl.50193

Lee, R. W., Woollings, T. J., Hoskins, B. J., Williams, K. D., O'Reilly, C. H., & Masato, G.(2018). Impact of gulf stream SST biases on the global atmospheric circulation. *Climate Dynamics*, *51*(9-10), 3369–3387. https://doi.org/10.1007/s00382-018-4083-9

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C.(2013). The collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, *6*(11), 2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018). Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *Journal of Climate*, *31*(2), 895–918. https://doi.org/10.1175/JCLI-D-17-0208.1

Lozier, M. S., Leadbetter, S., Williams, R. G., Roussenov, V., Reed, M. S. C., & Moore, N. J. (2008). The spatial pattern and mechanisms of heat-content change in the North Atlantic. *Science*, *319*(5864), 800–803. https://doi.org/10.1126/science.1146436

MacIntosh, C. R., Merchant, C. J., & von Schuckmann, K. (2017). Uncertainties in steric sea level change estimation during the satellite altimeter era: Concepts and practices. *Surveys in Geophysics*, *38*(1), 59–87. https://doi.org/10.1007/s10712-016-9387-x

Malavelle, F. F., Haywood, J. M., Jones, A., Gettelman, A., Clarisse, L., Bauduin, S., et al. (2017). Strong constraints on aerosol–cloud interactions from volcanic eruptions. *Nature*, *546*(7659), 485.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., et al. (2010). Description and evaluation of GLOMAP-mode: A modal global aerosol microphysics model for the UKCA composition-climate model. *Geoscientific Model Development*, *3*(2), 519–551. https://doi.org/10.5194/gmd-3-519-2010

Marnela, M., Rudels, B., Houssais, M.-N., Beszczynska-Möller, A., & Eriksson, P. B. (2013). Recirculation in the Fram Strait and transports of water in and north of the Fram Strait derived from CTD data. *Ocean Science*, *9*(3), 499–519. https://doi.org/10.5194/os-9-499-2013

Marshall, J., Donohoe, A., Ferreira, D., & McGee, D. (2014). The ocean's role in setting the mean position of the Inter-Tropical Convergence Zone. *Climate Dynamics*, *42*(7), 1967–1979. https://doi.org/10.1007/s00382-013-1767-z

Martin, E. R., Thorncroft, C., & Booth, B. B. B. (2014). The multidecadal Atlantic SST-Sahel rainfall teleconnection in CMIP5 simulations. *Journal of Climate*, *27*(2), 784–806. https://doi.org/10.1175/JCLI-D-13-00242.1

Marzocchi, A., Hirschi, J. J. M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., & Coward, A. C.(2015). The North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, *142*, 126–143. https://doi.org/10.1016/j.jmarsys.2014.10.007

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., et al. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, *10*(5), 2057–2116. https://doi.org/10.5194/gmd-10-2057-2017

Menary, M. B., Hodson, D. L. R., Robson, J. I., Sutton, R. T., & Wood, R. A. (2015). A Mechanism of internal decadal Atlantic Ocean variability in a high-resolution coupled climate model. *Journal of Climate*, *28*(19), 7764–7785. https://doi.org/10.1175/JCLI-D-15-0106.1

Menary, M. B., Kuhlbrodt, T., Ridley, J., Andrews, M. B., Dimdore-Miles, O. B., Deshayes, J., et al. (2018). Preindustrial control simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, *10*, 3049–3075. https://doi.org/10.1029/2018MS001495

Menary, M. B., Roberts, C. D., Palmer, M. D., Halloran, P. R., Jackson, L., Wood, R. A., et al. (2013). Mechanisms of aerosol-forced AMOC variability in a state of the art climate model. *Journal of Geophysical Research: Oceans*, *118*, 2087–2096. https://doi.org/10.1002/jgrc.20178

Menary, M. B., Robson, J., Allan, R., Booth, B. B. B., Cassou, C., Gastineau, G., et al. (2020). Aerosol-forced AMOC changes in CMIP6 historical simulations. *Geophysical Research Letters*, *47*, e2020GL088166. https://doi.org/10.1029/2020GL088166

Middlemas, E., Clement, A., & Medeiros, B. (2019). Contributions of atmospheric and oceanic feedbacks to subtropical northeastern sea surface temperature variability. *Climate Dynamics*, *53*(11), 6877–6890. https://doi.org/10.1007/s00382-019-04964-1

Mitchell, D. M., Gray, L. J., Anstey, J., Baldwin, M. P., & Charlton-Perez, A. J. (2013). The influence of stratospheric vortex displacements and splits on surface climate. *Journal of Climate*, *26*(8), 2668–2682. https://doi.org/10.1175/JCLI-D-12-00030.1

Moat, B. I., Sinha, B., Josey, S. A., Robson, J., Ortega, P., Sévellec, F., et al. (2019). Insights into decadal North Atlantic sea surface temperature and ocean heat content variability from an eddy-permitting coupled climate model. *Journal of Climate*, *32*(18), 6137–6161. https://doi.org/10.1175/JCLI-D-18-0709.1

Monerie, P.-A., Robson, J., Dong, B., Hodson, D. L. R., & Klingaman, N. P. (2019). Effect of the Atlantic multidecadal variability on the global monsoon. *Geophysical Research Letters*, *46*, 1765–1775. https://doi.org/10.1029/2018GL080903

Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., et al. (2015). Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. *Atmospheric Chemistry and Physics*, *15*, 8889–8973.

Montégut, C. B., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, *109*, C12003. https://doi.org/10.1029/2004JC002378

Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., et al. (2018). Improved aerosol processes and effective radiative forcing in HadGEM3 and UKESM1. *Journal of Advances in Modeling Earth Systems*, *10*, 2786–2805. https://doi.org/10.1029/2018MS001464

Myers, T. A., & Mechoso, C. R. (2020). Relative contributions of atmospheric, oceanic, and coupled processes to North Pacific and North Atlantic variability. *Geophysical Research Letters*, *47*, e2019GL086321. https://doi.org/10.1029/2019GL086321

Nakajima, T., & King, M. D. (1990). Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory. *Journal of the Atmospheric Sciences*, *47*(15), 1878–1893. https://doi.org/10.1175/1520-0469(1990)047<1878:DOTOTA>2.0.CO;2

Nisbet, E. G., Dlugokencky, E. J., Manning, M. R., Lowry, D., Fisher, R. E., France, J. L., et al. (2016). Rising atmospheric methane: 2007–2014 growth and isotopic shift. *Global Biogeochemical Cycles*, *30*, 1356–1370. https://doi.org/10.1002/2016GB005406

O'Connor, F. M., Abraham, N. L., Dalvi, M., Folberth, G., Griffiths, P., Hardacre, C., et al. (2020). Assessment of pre-industrial to present-day anthropogenic climate forcing in UKESM1. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–49. https://doi.org/10.5194/acp-2019-1152

O'Reilly, C. H., Zanna, L., & Woollings, T. (2019). Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices. *Journal of Climate*, *32*(22), 7727–7745. https://doi.org/10.1175/JCLI-D-19-0177.1

Oehrlein, J., Chiodo, G., & Polvani, L. M. (2020). The effect of interactive ozone chemistry on weak and strong stratospheric polar vortex events. *Atmospheric Chemistry and Physics Discussions*, *2020*, 1–29. https://doi.org/10.5194/acp-2020-187

Omrani, N.-E., Bader, J., Keenlyside, N. S., & Manzini, E. (2016). Troposphere–stratosphere response to large-scale North Atlantic ocean variability in an atmosphere/ocean coupled model. *Climate Dynamics*, *46*(5), 1397–1415. https://doi.org/10.1007/s00382-015-2654-6

Omrani, N.-E., Keenlyside, N. S., Bader, J., & Manzini, E. (2014). Stratosphere key for wintertime atmospheric response to warm Atlantic decadal conditions. *Climate Dynamics*, *42*(3-4), 649–663. https://doi.org/10.1007/s00382-013-1860-3

Otterå, O. H., Bentsen, M., Drange, H., & Suo, L. (2010). External forcing as a metronome for Atlantic multidecadal variability. *Nature Geoscience*, *3*(10), 688.

Painemal, D., & Zuidema, P. (2011). Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements. *Journal of Geophysical Research*, *116*, D24206. https://doi.org/10.1029/2011JD016155

Pausata, F. S. R., Pozzoli, L., Vignati, E., & Dentener, F. J. (2012). North Atlantic Oscillation and tropospheric ozone variability in Europe: model analysis and measurements intercomparison. *Atmospheric Chemistry and Physics*, *12*(14), 6357–6376. https://doi.org/10.5194/acp-12-6357-2012

Platnick, S., Li, J. Y., King, M. D., Gerber, H., & Hobbs, P. V. (2001). A solar reflectance method for retrieving the optical thickness and droplet size of liquid water clouds over snow and ice surfaces. *Journal of Geophysical Research*, *106*(D14), 15,185–15,199. https://doi.org/10.1029/2000JD900441

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., et al. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*, 4083–4097. https://doi.org/10.1175/JCLI-D-15-0556.1

Povey, A. C., & Grainger, R. G. (2019). Toward more representative gridded satellite products. *IEEE Geoscience and Remote Sensing Letters*, *16*(5), 672–676. https://doi.org/10.1109/LGRS.2018.2881762

Pringle, D. J., Eicken, H., Trodahl, H. J., & Backstrom, L. G. E. (2007). Thermal conductivity of landfast Antarctic and Arctic sea ice. *Journal of Geophysical Research*, *112*, C04017. https://doi.org/10.1029/2006JC003641

Qasmi, S., Cassou, C., & Boé, J. (2017). Teleconnection between Atlantic multidecadal variability and European temperature: Diversity and evaluation of the Coupled Model Intercomparison Project Phase 5 models. *Geophysical Research Letters*, *44*, 11,140–11,149. https://doi.org/10.1002/2017GL074886

Rae, J. G. L., Johnson, C. E., Bellouin, N., Boucher, O., Haywood, J. M., & Jones, A. (2007). Sensitivity of global sulphate aerosol production to changes in oxidant concentrations and climate. *Journal of Geophysical Research*, *112*, D10312. https://doi.org/10.1029/2006JD007826

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14), 4407. https://doi.org/10.1029/2002JD002670

Reintges, A., Martin, T., Latif, M., & Keenlyside, N. S. (2017). Uncertainty in twenty-first century projections of the Atlantic Meridional Overturning Circulation in CMIP3 and CMIP5 models. *Climate Dynamics*, *49*(5-6), 1495–1511. https://doi.org/10.1007/s00382-016-3180-x

Reverdin, G. (2010). North Atlantic subpolar gyre surface variability (1895–2009). *Journal of Climate*, *23*(17), 4571–4584. https://doi.org/10.1175/2010JCLI3493.1

Rice, A. L., Butenhoff, C. L., Teama, D. G., Röger, F. H., Khalil, M. A. K., & Rasmussen, R. A. (2016). Atmospheric methane isotopic record favors fossil sources flat in 1980s and 1990s with recent increase. *Proceedings of the National Academy of Sciences*, *113*(39), 10,791–10,796. https://doi.org/10.1073/pnas.1522923113

Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., & Schroeder, D.(2018). The sea ice model component of HadGEM3-GC3.1. *Geoscientific Model Development*, *11*(2), 713–723. https://doi.org/10.5194/gmd-11-713-2018

Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., et al. (2017). Role of atmospheric oxidation in recent methane growth. *Proceedings of the National Academy of Sciences*, *114*(21), 5373–5377. https://doi.org/10.1073/pnas.1616426114

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., et al. (2019). Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments. *Geoscientific Model Development*, *12*(12), 4999–5028. https://doi.org/10.5194/gmd-12-4999-2019

Roberts, C. D., Jackson, L., & McNeall, D. (2014). Is the 2004–2012 reduction of the Atlantic meridional overturning circulation significant?. *Geophysical Research Letters*, *41*, 3204–3210. https://doi.org/10.1002/2014GL059473

Robson, J., Ortega, P., & Sutton, R. (2016). A reversal of climatic trends in the North Atlantic since 2005. *Nature Geoscience*, *9*(7), 513. https://doi.org/10.1038/ngeo2727

Robson, J., Sutton, R. T., Archibald, A., Cooper, F., Christensen, M., Gray, L. J., et al. (2018). Recent multivariate changes in the North Atlantic climate system, with a focus on 2005–2016. *International Journal of Climatology*, *38*(14), 5050–5076. https://doi.org/10.1002/joc.5815

Robson, J., Sutton, R., Lohmann, K., Smith, D., & Palmer, M. D. (2012). Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. *Journal of Climate*, *25*(12), 4116–4134. https://doi.org/10.1175/JCLI-D-11-00443.1

Ruiz-Barradas, A., Chafik, L., Nigam, S., & Häkkinen, S. (2018). Recent subsurface North Atlantic cooling trend in context of atlantic decadal-to-multidecadal variability. *Tellus A: Dynamic Meteorology and Oceanography*, *70*(1), 1–19. https://doi.org/10.1080/16000870.2018.1481688

Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E., & Ostrow, H. (1989). MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Transactions on Geoscience and Remote Sensing*, *27*(2), 145–153. https://doi.org/10.1109/36.20292

Sarafanov, A., Falina, A., Sokov, A., & Demidov, A. (2008). Intense warming and salinification of intermediate waters of southern origin in the eastern subpolar North Atlantic in the 1990s to mid-2000s. *Journal of Geophysical Research*, *113*, C12022. https://doi.org/10.1029/2008JC004975

Scaife, A. A., Folland, C. K., Alexander, L. V., Moberg, A., & Knight, J. R. (2008). European climate extremes and the North Atlantic Oscillation. *Journal of Climate*, *21*(1), 72–83. https://doi.org/10.1175/2007JCLI1631.1

Schröder, D., Feltham, D. L., Tsamados, M., Ridout, A., & Tilling, R. (2019). New insight from CryoSat-2 sea ice thickness for sea ice modelling. *The Cryosphere*, *13*(1), 125–139. https://doi.org/10.5194/tc-13-125-2019

Schutgens, N. A. J., Gryspeerdt, E., Weigum, N., Tsyro, S., Goto, D., Schulz, M., & Stier, P. (2016). Will a perfect model agree with perfect observations? The impact of spatial sampling. *Atmospheric Chemistry and Physics*, *16*(10), 6335–6353. https://www.atmos-chem-phys.net/16/6335/2016/, https://doi.org/10.5194/acp-16-6335-2016

Screen, J. A. (2017). Simulated atmospheric response to regional and Pan-Arctic Sea Ice Loss. *Journal of Climate*, *30*(11), 3945–3962. https://doi.org/10.1175/JCLI-D-16-0197.1

Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., et al. (2019). Evaluation of CNRM Earth system model, CNRM-ESM2-1: Role of Earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, *11*, 4182–4227. https://doi.org/10.1029/2019MS001791

Sellar, A. A., Jones, C. G., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., et al. (2019). UKESM1: Description and evaluation of the UK Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*, 4513–4558. https://doi.org/10.1029/2019MS001739

Siddans, R., Knappett, D., Kerridge, B., Waterfall, A., Hurley, J., Latter, B., et al. (2017). Global height-resolved methane retrievals from the Infrared Atmospheric sounding interferometer (IASI) on MetOp. *Atmospheric Measurement Techniques*, *10*(11), 4135–4164. https://doi.org/10.5194/amt-10-4135-2017

Sinha, B., Smeed, D. A., McCarthy, G., Moat, B. I., Josey, S. A., Hirschi, J. J.-M., et al. (2018). The accuracy of estimates of the overturning circulation from basin-wide mooring arrays. *Progress in Oceanography*, *160*, 101–123. https://doi.org/10.1016/j.pocean.2017.12.001

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., et al. (2018). The North Atlantic Ocean is in a State of reduced overturning. *Geophysical Research Letters*, *45*, 1527–1533. https://doi.org/10.1002/2017GL076350

Smith, D. M., Eade, R., Dunstone, N. J., Fereday, D., Murphy, J. M., Pohlmann, H., & Scaife, A. A. (2010). Skilful multi-year predictions of Atlantic hurricane frequency. *Nature Geoscience*, *3*(12), 846. https://doi.org/10.1038/ngeo1004

Spreen, G., Kern, S., Stammer, D., & Hansen, E. (2009). Fram Strait sea ice volume export estimated between 2003 and 2008 from satellite data. *Geophysical Research Letters*, *36*, L19502. https://doi.org/10.1029/2009GL039591

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., et al. (2013). *Climate change 2013: The physical science basis*. Cambridge: Cambridge University Press.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., et al. (2018). UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. *Geoscientific Model Development*, *11*(8), 3187–3213. https://doi.org/10.5194/gmd-11-3187-2018

Sutton, R. T., & Dong, B. (2012). Atlantic Ocean influence on a shift in European climate in the 1990s. *Nature Geoscience*, *5*(11), 788–792. https://doi.org/10.1038/ngeo1595

Sutton, R. T., & Hodson, D. L. R. (2005). Atlantic Ocean forcing of North American and European summer climate. *Science*, *309*(5731), 115–118. https://doi.org/10.1126/science.1109496

Sutton, R. T., McCarthy, G. D., Robson, J., Sinha, B., Archibald, A. T., & Gray, L. J. (2018). Atlantic multidecadal variability and the U.K. ACSIS Program. *Bulletin of the American Meteorological Society*, *99*(2), 415–425. https://doi.org/10.1175/BAMS-D-16-0266.1

Swart, N. C., Fyfe, J. C., Hawkins, E., Kay, J. E., & Jahn, A. (2015). Influence of internal variability on Arctic sea-ice trends. *Nature Climate Change*, *5*(2), 86.

Thomas, G. E., Poulsen, C. A., Sayer, A. M., Marsh, S. H., Dean, S. M., Carboni, E., et al. (2009). The GRAPE aerosol retrieval algorithm. *Atmospheric Measurement Techniques*, *2*(2), 679–701. https://doi.org/10.5194/amt-2-679-2009

Tilling, R. L., Ridout, A., & Shepherd, A. (2018). Estimating Arctic sea ice thickness and volume using CryoSat-2 radar altimeter data. *Advances in Space Research*, *62*(6), 1203–1225. The CryoSat Satellite Altimetry Mission: Eight Years of Scientific Exploitation https://doi.org/10.1016/j.asr.2017.10.051

Tsamados, M., Feltham, D. L., & Wilchinsky, A. V. (2013). Impact of a new anisotropic rheology on simulations of arctic sea ice. *Journal of Geophysical Research: Oceans*, *118*, 91–107. https://doi.org/10.1029/2012JC007990

Turner, A. J., Frankenberg, C., Wennberg, P. O., & Jacob, D. J. (2017). Ambiguity in the causes for decadal trends in atmospheric methane and hydroxyl. *Proceedings of the National Academy of Sciences*, *114*(21), 5367–5372. https://doi.org/10.1073/pnas.1616020114

Vinje, T., Nordlund, N., & Kvambekk, A. A. (1998). Monitoring ice thickness in Fram Strait. *Journal of Geophysical Research*, *103*(C5), 10,437–10,449. https://doi.org/10.1029/97JC03360

Visbeck, M., Cullen, H., Krahmann, G., & Naik, N. (1998). An ocean model's response to North Atlantic Oscillation-like wind forcing. *Geophysical Research Letters*, *25*(24), 4521–4524. https://doi.org/10.1029/1998GL900162

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., et al. (2019). The Met Office unified model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. *Geoscientific Model Development*, *12*(5), 1909–1963. https://doi.org/10.5194/gmd-12-1909-2019

Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., et al. (2016). An assessment of the Arctic Ocean in a suite of interannual CORE-I I simulations. Part II: Liquid freshwater. *Ocean Modelling*, *99*, 86–109. https://doi.org/10.1016/j.ocemod.2015.12.009

Watanabe, M., & Tatebe, H. (2019). Reconciling roles of sulphate aerosol forcing and internal variability in atlantic multidecadal climate changes. *Climate Dynamics*, *53*(7-8), 4651–4665. https://doi.org/10.1007/s00382-019-04811-3

Wilcox, L. J., Highwood, E. J., Booth, B. B. B., & Carslaw, K. S. (2015). Quantifying sources of inter-model diversity in the cloud albedo effect. *Geophysical Research Letters*, *42*, 1568–1575. https://doi.org/10.1002/2015GL063301

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., et al. (2018). The Met Office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*, 357–380. https://doi.org/10.1002/2017MS001115

Woollings, T., Czuchnicki, C., & Franzke, C. (2014). Twentieth century North Atlantic jet variability. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 783–791. https://doi.org/10.1002/qj.2197

Woollings, T., Franzke, C., Hodson, D. L. R., Dong, B., Barnes, E. A., Raible, C. C., & Pinto, J. G. (2015). Contrasting interannual and multidecadal NAO variability. *Climate Dynamics*, *45*(1), 539–556. https://doi.org/10.1007/s00382-014-2237-y

Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M., & Brayshaw, D. J. (2012). Response of the North Atlantic storm track to climate change shaped by ocean–atmosphere coupling. *Nature Geoscience*, *5*(5), 313.

Xiong, X., Barnet, C., Maddy, E., Sweeney, C., Liu, X., Zhou, L., & Goldberg, M. (2008). Characterization and validation of methane products from the Atmospheric Infrared Sounder (AIRS). *Journal of Geophysical Research*, *113*, G00A01. https://doi.org/10.1029/2007JG000500

Yan, X., Zhang, R., & Knutson, T. R. (2019). A Multivariate AMV Index and Associated Discrepancies Between Observed and CMIP5 Externally Forced AMV. *Geophysical Research Letters*, *46*, 4421–4431. https://doi.org/10.1029/2019GL082787

Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., & Teng, H. (2012). A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content. *Journal of Climate*, *25*(15), 5173–5189. https://doi.org/10.1175/JCLI-D-11-00595.1

Yool, A., Popova, E. E., & Anderson, T. R. (2013). MEDUSA-2.0: An intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies. *Geoscientific Model Development*, *6*(5), 1767–1811. https://doi.org/10.5194/gmd-6-1767-2013

Yuan, T., Oreopoulos, L., Zelinka, M., Yu, H., Norris, J. R., Chin, M., et al. (2016). Positive low cloud and dust feedbacks amplify tropical North Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, *43*, 1349–1356. https://doi.org/10.1002/2016GL067679

Zhang, Z., Ackerman, A. S., Feingold, G., Platnick, S., Pincus, R., & Xue, H. (2012). Effects of cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet effective radius: Case studies based on large-eddy simulations. *Journal of Geophysical Research*, *117*, D19208. https://doi.org/10.1029/2012JD017655

Zhang, R., Delworth, T. L., Sutton, R., Hodson, D. L. R., Dixon, K. W., Held, I. M., et al. (2013). Have aerosols caused the observed Atlantic Multidecadal variability? *Journal of the Atmospheric Sciences*, *70*(4), 1135–1144. https://doi.org/10.1175/JAS-D-12-0331.1

Zhang, J., & Rothrock, D. A. (2003). Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. *Monthly Weather Review*, *131*(5), 845–861. https://doi.org/10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2

Zhang, Z., Werner, F., Cho, H.-M., Wind, G., Platnick, S., Ackerman, A. S., et al. (2016). A framework based on 2-D Taylor expansion for quantifying the impacts of subpixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the bispectral method. *Journal of Geophysical Research: Atmospheres*, *121*, 7007–7025. https://doi.org/10.1002/2016JD024837

Ziemke, J. R., Chandra, S., Duncan, B. N., Froidevaux, L., Bhartia, P. K., Levelt, P. F., & Waters, J. W. (2006). Tropospheric ozone determined from Aura OMI and MLS: Evaluation of measurements and comparison with the global modeling initiative's chemical transport model. *Journal of Geophysical Research*, *111*, D19303. https://doi.org/10.1029/2006JD007089