JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2021MS002888

Improved Surface Mass Balance Closure in Ocean Hindcast Simulations

Matthew Harrison[1], **Alistair Adcroft**[2], **Robert Hallberg**[1,2], **and Olga Sergienko**[2]

[1]Oceans and Cryosphere Group, NOAA/GFDL, Princeton, NJ, USA, [2]Princeton University, Princeton, NJ, USA

Special Section:
Dynamical cores of oceanic models across all scales and their evaluation

Key Points:

- Including a land-hydrology model component in a forced global ocean/sea-ice configuration allows for a more realistic closure of the freshwater cycle
- Improved freshwater closure can substantially reduce ocean temperature and salinity initialization drift - without the need for artificial salinity restoring fluxes
- A mid-latitude poleward moisture transport constraint derived from CMIP6 historical simulations is consistent with further increase in hindcast performance

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:
M. Harrison,
matthew.harrison@noaa.gov

Citation:
Harrison, M., Adcroft, A., Hallberg, R., & Sergienko, O. (2022). Improved surface mass balance closure in ocean hindcast simulations. *Journal of Advances in Modeling Earth Systems*, *14*, e2021MS002888. https://doi.org/10.1029/2021MS002888

Received 4 NOV 2021
Accepted 23 MAY 2022

Abstract Forced global ocean/sea-ice hindcast simulations are subject to persistent surface mass flux estimation biases, for example, configurations with an explicit-free surface may not take into account the seasonal storage of water on land when constraining sea level. We present a physically motivated surface mass flux closure, that results in: reduced watermass drift from initialization; improved Atlantic meridional overturning cirulation intensity; and more realistic rates of ocean heat uptake, in simulations using global ocean/sea-ice/land (MOM6/SIS2/LM3) model configurations, forced with atmospheric reanalysis data. In addition to accounting for the land storage, the area-integrated subpolar-to-polar (40°–90°N/S) surface mass fluxes are constrained, using a climatological estimate derived from the the CMIP6 historical ensemble, which helps to further improve hindcast performance. Simulations using MERRA-2 and JRA55-do forcing, subject to identical hydrologic constraints, exhibit similar reductions in drift.

Plain Language Summary Weather-reanalysis products combine atmospheric numerical models and observations in order to provide a historical record of the near-surface atmospheric state. This information is used to drive numerical ocean/sea-ice models, and the resulting retrospective simulations can be compared to independent ocean observations in order to evaluate model skill. Forced ocean general circulation models generally experience large initialization drift away from observations in response to reanalysis forcing. In this study, we evaluate the impact of a more realistic representation of the hydrologic cycle, or the flow of freshwater through the ocean-atmosphere-land system. We demonstrate considerable model skill enhancement, with reduced initialization drift, when using our new approach. This furthers the ability of ocean modelers to simulate and understand the physical evolution of the ocean during recent decades.

1. Introduction

Coupled general circulation models that include atmospheric, land, sea-ice and oceanic components should approximately conserve water, that is, evaporation, precipitation, moist transport and associated phase changes, occur through the climate system without spurious water gain or loss. In contrast, forced ocean hindcast models typically do not contain explicit atmospheric and land model components, and therefore require a means to close the global hydrologic cycle.

Forced ocean hindcasts specify precipitation, evaporation and runoff in some way, but this results in non-conservation errors due to residual imbalances between large terms. This may in turn lead to unrealistic trends in sea level. In addition to aggregate errors of this nature, there are uncertainties in local and regional mass-balance estimates, which together with associated errors in surface buoyancy fluxes, may degrade interior properties and ventilation rates. Thus, a common approach used in ocean hindcast modeling is to include an ad-hoc restorative damping of near surface salinity toward a prior climatology (Behrens et al., 2013; Danabasoglu et al., 2014; Griffies et al., 2016). While the need for restoring has diminished as forcing datasets and models have improved, it remains a common practice.

Global ocean hindcasts often use a natural boundary condition for freshwater, in conjunction with an evolving and explicit free surface formulation (Griffies et al., 2001); precipitation (P) is prescribed and evaporation (E) is computed with bulk aerodynamic formula, using boundary layer temperature, relative humidity and wind speeds from reanalysis products. Mass continuity over an arbitrary region of the globe can be written as,

$$\int_V Q_t dV = -\int_S \mathbf{T}_q \cdot \hat{n} dS + \int_A (E - P) dA, \quad (1)$$

Author Contributions:
Conceptualization: Matthew Harrison, Alistair Adcroft, Robert Hallberg
Data curation: Matthew Harrison
Formal analysis: Matthew Harrison
Funding acquisition: Robert Hallberg
Investigation: Matthew Harrison
Methodology: Matthew Harrison
Project Administration: Robert Hallberg
Software: Matthew Harrison
Supervision: Alistair Adcroft, Robert Hallberg
Validation: Matthew Harrison
Visualization: Matthew Harrison
Writing – original draft: Matthew Harrison
Writing – review & editing: Matthew Harrison, Alistair Adcroft, Robert Hallberg, Olga Sergienko

where the LHS term is the time-tendency of the integrated atmospheric water concentration, Q [kg m-3], $\mathbf{T}_q$ [kg m-2 s-1] is the lateral transport of moisture per unit area in the atmosphere and $(E-P)$ [kg m-2 s-1] is the net rate of surface evaporation, which are respectively integrated over a control volume, V, in the atmosphere bounded laterally by sides, S, with outward facing normal, $\hat{n}$, and at the surface by A. For suitable control volumes over sufficiently large time and space scales, net atmospheric moisture storage is small and Equation 1 reduces to,

$$\int_S \mathbf{T}_q \cdot \hat{n} dS \approx \int_A (E-P) dA. \tag{2}$$

Integrated globally, both terms become identically zero in this approximation. While atmospheric residence times for water range from hours to days, land residence time are seasonal and longer; in the high Arctic for example, peak Spring runoff occurs following melting of ice and snow stored on land during Winter. Seasonal land storage is often neglected in forced ocean hindcasts. For example, Adcroft et al. (2019) (hereafter, A19) used an alternative global mass closure to Equation 2, based on an ocean-sided mass conservation implementation,

$$\int_O (E-P-R) dA. = 0, \tag{3}$$

where combined land runoff and calving (R) at coastlines are specified and the integration occurs over the ocean only. Residual imbalances from Equation 3 were used to calculate a global offset for precipitation in order to retain global mass conservation (at each air-sea coupling timestep) during the simulation described in A19. While Equation 3 is reasonable when averaged over annual and longer timescales, for which land storage changes are relatively small, it is not applicable over seasonal timescales, for which land storage is not negligible.

The present study demonstrates, using the MOM6 ocean and SIS2 sea-ice configuration (termed OM4) from A19, the sensitivity to more realistic hydrologic closure using Equation 2. Two atmospheric reanalyzes are used to evaluate our approach: MERRA-2 (Rienecker et al., 2011) and JRA55-do (Tsujino et al., 2018). In the proposed configurations, we incorporate a comprehensive land model - with its associated network of groundwater, lakes and rivers. Global mass balance is maintained by diagnosing and adjusting the precipitation over the entire earth's surface, rather than solely the ocean, as in previous studies.

Additionally we describe simulations with a constraint on the implied mid-latitude poleward moisture transport (PMT). This constraint is a pre-computed climatology, based on a sampling of the subpolar-to-polar surface mass balance in the CMIP6 historical ensemble. This requires re-scaling precipitation independently within the northern and southern subpolar-to-polar domains according to the regional moisture transport convergence implied by the CMIP6 ensemble-average fluxes.

2. Methods

In this section, we provide descriptions of: the model configuration used in this study (2.1); the reanalysis data used to force the hindcast simulations (2.2); the surface fluxes (2.3); observational data used for model evaluation and initialization (2.4); and the implied poleward moisture constraint implementation (2.5).

2.1. Model Description and Initialization

The model is comprised of a layered ocean general circulation model, MOM6, and the SIS2 sea-ice model, on a common $\frac{1}{4}^\circ$ nominal grid (Adcroft et al., 2019). A land model, LM4.0 (Zhao et al., 2018), provides continental discharge and is discretized on a cubed-sphere grid at approximately 50 km resolution. These model components exchange energy and mass at the surface using prescribed atmosphere boundary conditions from the atmospheric reanalysis, which are interpolated to the cubed-sphere atmospheric grid, prior to the computation of the exchange fluxes (Balaji et al., 2006).

The ocean initial conditions, unless otherwise noted, are taken from the World Ocean Atlas version 2005. Additional experiments were performed with initial January 1988 ocean conditions from the SODA reanalysis (Carton et al., 2005). The simulations presented here were initialized without sea-ice. Additional simulations (not shown) were initialized using a sea-ice climatology. This greatly reduced an initial high latitude near-surface salinity shock, but otherwise did not significantly impact our results.

A 300 years land initialization, using a reanalysis-based forcing data set (Sheffield et al., 2006), with fixed land-use, was used in order to bring deep soil, groundwater and lake reservoirs into approximate balance. This was followed by an additional 20 years simulation (1982–2002) using MERRA-2 forcing data. Global RMS land storage tendency following the initialization procedure is less than 10^4 m^3 s^{-1} (0.01 Sv), and results presented here are robust to additional iterations of the land model spin-up procedure. It should be noted that glacial mass changes, and associated coastal runoff and calving are not accounted for in these simulations.

2.2. Forcing Data

Simulations were conducted using the MERRA-2 and JRA55-do reanalyzes. MERRA-2 data are provided hourly and JRA55-do data are at 3 hourly intervals, both on nominal 50 km horizontal grids. Boundary conditions are linearly interpolated to the model grid at each coupling timestep. MERRA-2 precipitation, prior to 2016 were replaced with an earlier version due to issues related to satellite assimilation (Cullather & Bosilovich, 2011) identified in MERRA-2 (additional simulations using MERRA-2 precipitation produced similar results).

It should be noted that JRA55-do data were subject to adjustments from the original reanalysis data, consistent with previous ocean forcing protocols (Large & Yeager, 2009; Tsujino et al., 2018), whereas no modifications were applied to MERRA-2 data used in this study, apart from precipitation adjustments consistent with surface mass balance constraints in Equation 2.

2.3. Surface Fluxes

Monin-Obukhov stability theory is used to calculate stress, heat and moisture exchange coefficients, as in fully coupled model configurations (Held et al., 2019), and the 10-m data supplied from the reanalyzes are assumed to lie within the constant turbulent flux layer. Ocean roughness depends on kinematic viscosity at low wind speeds and a Charnock relation for high wind conditions (Charnock, 1955).

Incident shortwave/longwave radiation and precipitation are specified using reanalysis data, and reflected shortwave radiation at the air-sea interface is parameterized by a zenith-dependent ocean/sea-ice albedo scheme. Stefan-Boltzmann law is used for longwave emissions.

Additional simulations were performed using neutral turbulent exchange coefficients (Large & Yeager, 2009), without significant impact on the results presented here.

2.4. Observational Datasets

For model evaluation, ocean temperature and salinity observations are taken from the EN4v3.4.2 data set, based on quality-controlled gridded ocean data (Good et al., 2013) for the 1982–2018 time period. Ocean observations are more scarce prior to the ARGO period, which starts roughly in 2005. For the latter time period, a gridded analysis based on ARGO-only temperature and salinity profiles to 2000 m are used for additional model comparisons (Roemmich et al., 2009).

An observational array-based timeseries of the Atlantic depth-space overturning circulation is available beginning around 2005 (Frajka-Williams et al., 2021). This is used to evaluate the Atlantic meridional overturning circulation (AMOC) strength during the last 14 years of the simulations. The estimated average AMOC intensity at 26°N for the 2005–2018 time period is 16.9 $\pm$ 0.9 Sv (1 Sv = 10^6 m^3 s^{-1}).

Ocean initial potential temperature and salinity are taken from the World Ocean Atlas, version 2005 (Antonov et al., 2006; Locarnini et al., 2018). An exception is Exp:M*-88, which uses initial values in 1988 from the SODAv3.4.2 analysis (Carton et al., 2005). The SODA reanalysis is based on an earlier version of MOM and assimilates historical temperature and salinity profiles into the model using a statistical interpolation procedure.

2.5. Implied Poleward Moisture Constraint

For all simulations, unless otherwise noted, the surface mass flux balance is consistent with Equation 2—in other words, land storage is accounted for when closing the mass budget, compared to previously reported simulations

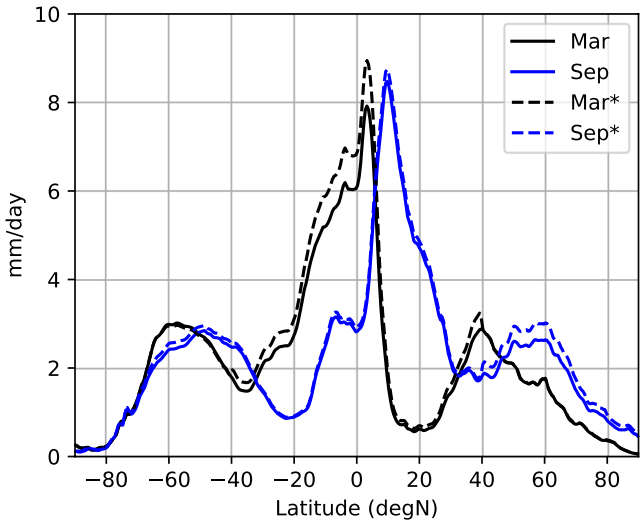

Figure 1. March (black) and September (blue) 2010 original MERRA zonal mean precipitation (solid, mm/day) and the adjusted version in Exp:M* (dashed).

(e.g., A19, using Equation 3), which do not account for seasonal water stored on land. We additionally present results which demonstrate model sensitivity to an implied poleward moisture transport constraint.

The mid-latitude PMT has important implications for the deep overturning in the ocean through the associated flux of buoyant freshwater into the subpolar and polar seas. While the advective fluxes of freshwater in the ocean dwarf the surface fluxes, the latter drives the thermohaline circulation and influences the salt budget of oceanic volumes (Ferreira et al., 2010; Stommel, 1961; Wijffels et al., 1992). When forcing ocean models using atmospheric reanalysis products, which may contain persistent regional data-assimilation related sources or sinks of moisture with resulting surface mass flux imbalances (Cullather & Bosilovich, 2011), steps are needed to compensate for these and other model-related errors.

In this study, we propose a mid-latitude PMT constraint which introduces two additional forcing parameters for the subpolar-to-polar mass balance in each hemisphere (40–90°N/S). We utilize the CMIP6 historical ensemble in order to provide an estimate of present-day PMT, by diagnosing the surface mass balance poleward of 40° latitude in each hemisphere. The selected latitudes corresponds roughly to the transition from the subtropics, where evaporation dominates, to the subpolar environment, where precipitation exceeds evaporation.

Monthly average binning of the surface fluxes was performed using data from historical simulation years 1982–2015. It should be noted that internal variability in CMIP6, for instance related to ENSO, is not in phase with the actual record. The ensemble-time averaging suppresses aliasing of internal variability as well as inter-model differences. Two caveats inherent in our approach pertain to the: (a) elimination of internal PMT variability, apart from a smoothly varying representation of the seasonal cycle; (b) systematic biases in the hydrologic representation for CMIP6 models, which could impact our results (e.g., Harrison et al. (2014)). Northern hemisphere CMIP6 PMT values range from 0.77 to 1.1 Sv, with an ensemble mean of 0.91 Sv. Southern transports are larger, 0.86–1.2 Sv, with an average of 1.1 Sv (see Supporting Information S1 for more details).

For the PMT-constrained simulations, the CMIP6 monthly climatological estimates are applied such that at each coupling time-step, when atmospheric thermodynamic fluxes are calculated, the time-interpolated target values of the respective transports, $T_n^*(t)$ and $T_s^*(t)$, are applied within the $40°S$–$40°N$ domain, such that

$$\bar{T}_n^* + \bar{T}_s^* = \int_{40°S}^{40°N} (E - P^*)\, dA, \tag{4}$$

$P^* = \alpha P$, is the adjusted precipitation within the central domain. Solving for the integrated adjusted precipitation, $\left(P_\Sigma^* = \int_{40°S}^{40°N} P^* dA\right)$,

$$P_\Sigma^* = \int_{40°S}^{40°N} E dA - \left(\bar{T}_n^* + \bar{T}_s^*\right) \tag{5}$$

and using the integrated original precipitation data (P_Σ), taking the ratio yields a re-scaling factor for the central domain,

$$\alpha = \frac{P_\Sigma^*}{P_\Sigma}, \tag{6}$$

which is applied to the precipitation at each gridpoint from 40°S to 40°N. These steps are repeated for the northern (40–90°N) and southern (40–90°S) regions, using their respective target transport divergence ($-\bar{T}_n^*$ and $-\bar{T}_s^*$, respectively in Equations 4 and 5). This results in a discontinuity across the mid-latitude boundaries in the re-scaled precipitation, an example of which is shown in Figure 1 during March and September 2010. Monthly

Table 1
Experiment Summary

Name	Description
Exp:M-ref	MERRA-2 forcing. Global freshwater forcing is balanced using Equation 3 as in A19, which neglects land storage
Exp:M	MERRA-2 forcing. Global freshwater forcing is balanced using Equation 2, which accounts for land storage.
Exp:M*	As in Exp:M, with the CMIP6-based mid-latitude PMT constraint
Exp:M*-88	As in Exp:M*, initialized in 1988 with SODA reanalysis temperature and salinity.
Exp:J	JRA55-do forcing. Global freshwater forcing is balanced using Equation 2, which accounts for land storage.
Exp:J*	As in Exp:J, with the CMIP6-based mid-latitude PMT constraint.

average precipitation re-scaling with MERRA-2 is generally positive and less than ∼15% in any domain. Discontinuities in the adjusted precipitation are detectable at the mid-latitude boundaries. These discontinuities could be partly alleviated by applying a time-filtered approach in Equation 4, which would additionally retain the internal variability in the fluxes - by constraining the slowly varying PMT, rather than the instantaneous values - but is this is considered beyond the scope of the present study.

3. Results

Tables 1 summarize the experiments presented in this study. A reference case (Exp:M-ref) does not take into account land storage when rescaling precipitation, as in A19. In PMT constrained hindcasts Exp:M* and Exp:J*, the moisture transports at 40°N/S are constrained to the monthly climatological CMIP6 estimate - the transports in Exp:M and Exp:J are a model result, in which global mass conservation is applied using Equation 2, without the PMT constraint. Initial condition sensitivity is partly evaluated in Exp:M*-88, which uses monthly average January 1988 SODA temperature and salinity, instead of WOA05, to initialize the model in 1988 (instead of 1982).

Table 2 shows freshwater (FW) forcing metrics: the implied moisture transports at 40°N/S (cols a,d); the Atlantic moisture convergence north of 40°N (col b); and from 34°S to 40°N (col c). The Atlantic overturning index along 26°N is shown (col e), along with ocean heat content tendencies in the upper (0–1500 m) and deep (1500–4500 m) ocean between pentads centered on 2007 and 2016, during the ARGO period (col f).

The MERRA-2 simulation using the original land closure (Exp:M-ref) significantly under-estimates northern hemisphere PMT and over-estimates southern transport relative to CMIP6, while MERRA-based PMT with global land closure (Exp:M) is within the CMIP6 ensemble range in both hemispheres. The difference in implied PMT between these experiments accounts for a significantly stronger rate of initial cooling in Exp:M-ref (Figure 2, panel a), a relatively strong AMOC intensity (Table 2, col e), and less ocean heat update during the analysis period (Table 2, col f).

Table 2
Experiment Metrics (2000–2018): (a) FW Forcing 40–90°N (Sv); (b) Atlantic/Arctic FW Forcing 40–90°N (Sv); (c) Atlantic FW Forcing 34°S–40°N (Sv); (d) FW Forcing 40–90°S (Sv); (e) 2005–2018 AMOC Index at 26°N (Sv), and the Linear Trend (Sv yr^{-1}); (f) Recent Pentadal Rate of Change of Ocean Heat Content (2014–2018) Minus (2005–2009), for the Top 1500 m and 1500–4500 m (W m^{-2} Over Ocean)

Name	(a) 40N+ (Sv)	(b) Atl40 N+ (Sv)	(c) Atl 34S-40 N (Sv)	(d) 40S- (Sv)	(e) AMOC (Sv)	(f) OHC (W m^{-2})
M-ref	0.7	0.4	−0.74	1.24	19.5 (0.2)	0.16/0.04
M	0.86	0.47	−0.89	1.1	15.9 (0)	0.64/−0.07
M*	0.87	0.48	−0.87	1.13	15.5 (−0.1)	0.48/−0.04
J	0.81	0.38	−0.78	0.94	22.9 (−0.01)	−0.04/0.27
J*	0.86	0.45	−0.88	1.13	22.3 (−0.05)	0.23/0.37

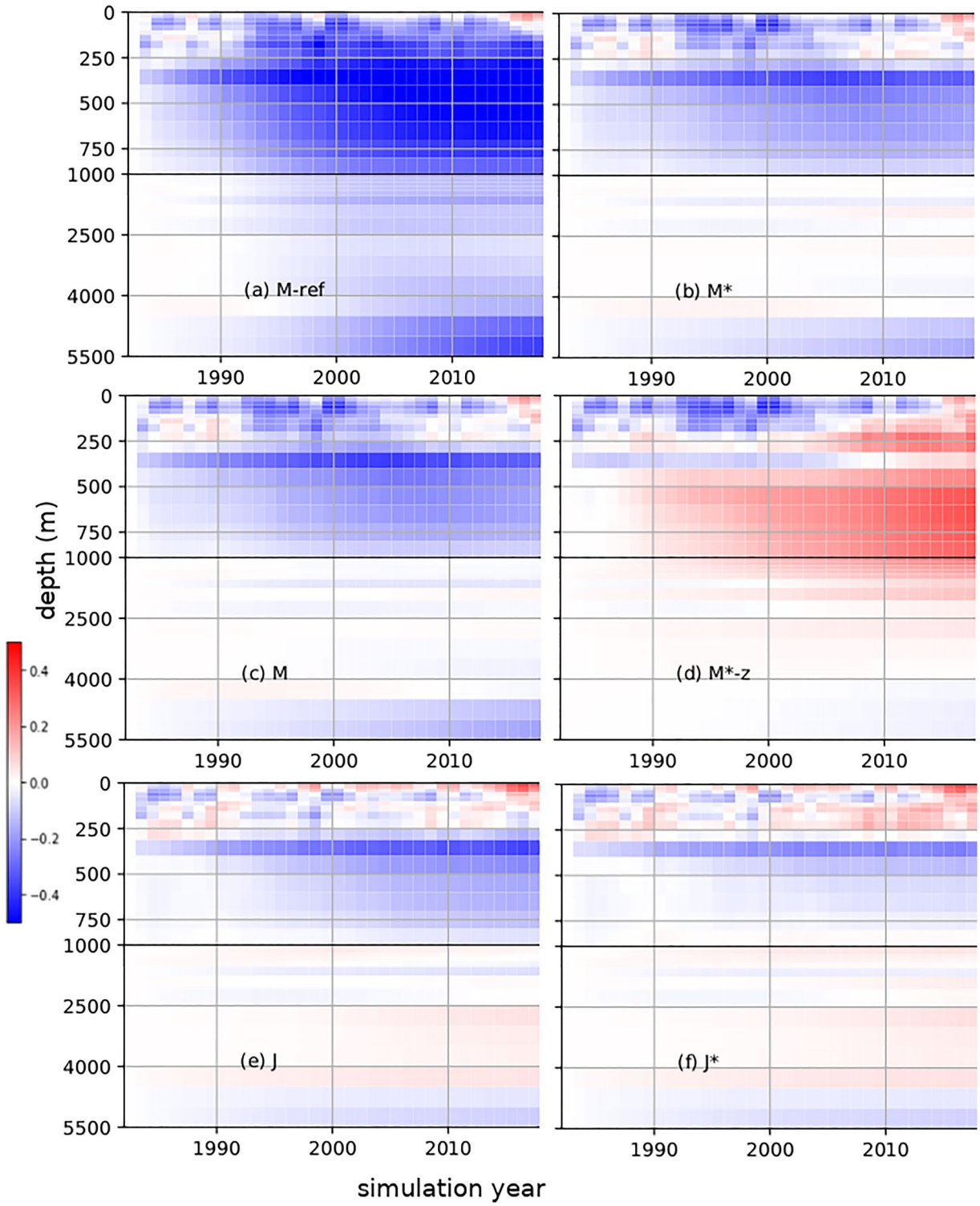

Figure 2. Global volume average temperature anomalies with respect to year one of the simulation (degC) for experiments forced with MERRA-2: Exp:M-ref (no land storage), Exp:M (land storage), and Exp:M* (land storage with poleward moisture transport [PMT] constraint). JRA55-do forced simulations: Exp:J (land storage) and Exp:J* (land storage and PMT constraint). An additional simulation is shown here (Exp:M*-z) which is using a geopotential-based (z) vertical coordinate representation in MOM6, instead of the hybrid z-potential density coordinate used in the simulations presented in this study.

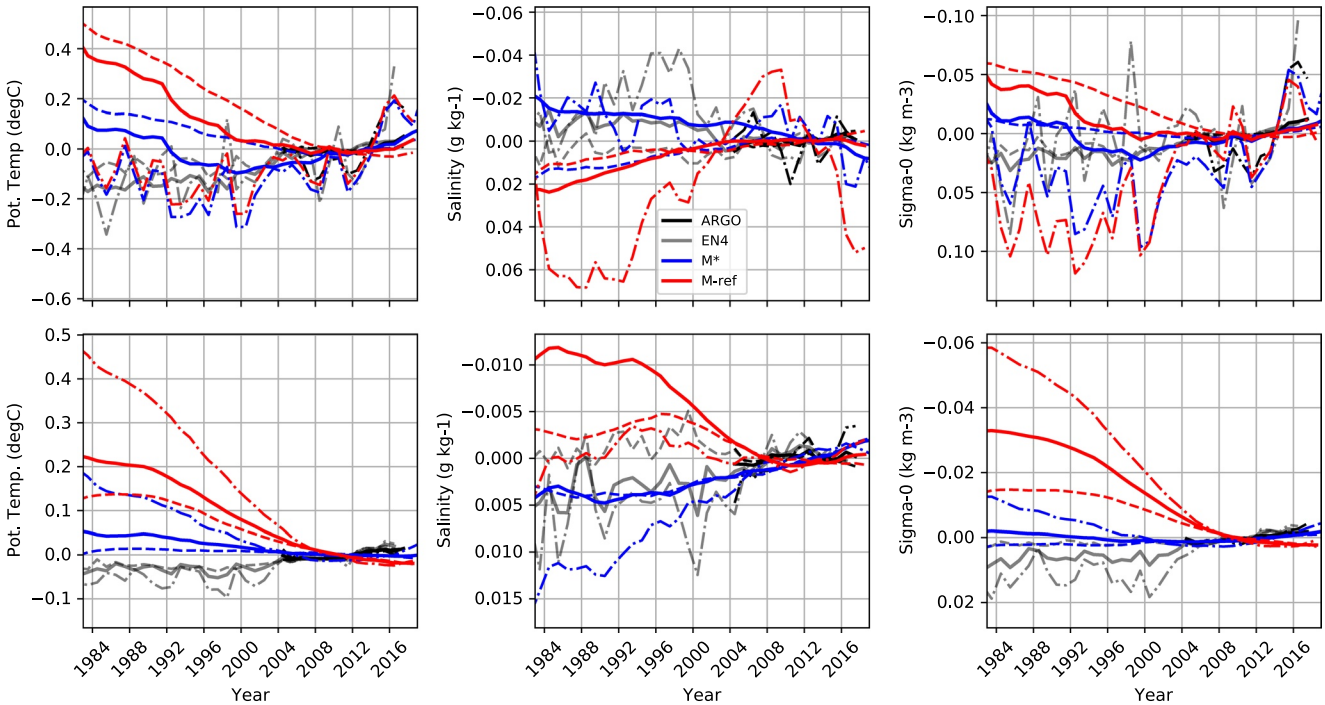

Figure 3. Annual average anomalies with respect to the 2005–2018 time period in the upper 500 m (top panels) and 500–1500 m (bottom panels). From left to right: potential temperature (degC), salinity (psu), and sigma-0 (kg m-3). Exp:M-ref and Exp:M* are indicated by red and blue lines, respectively. The area average at the upper (dot-dashed) and lower (dashed) interfaces are shown along with the volume average (solid). Observational estimates from EN4 and ARGO are shown in gray and black respectively.

While the implied moisture transports in Exp:M are within the CMIP6 ensemble range, the JRA55-do simulation (Exp:J) is near the lower range of CMIP6 in both hemispheres. The JRA55-do forced simulation responded similarly to the inclusion of land storage and the PMT constraint, with reduced thermal drift. For the analysis period, JRA55-do forced simulations produce reduced rates of upper ocean heat uptake, and stronger deep ocean heating, compared to the MERRA-2 simulations, both with and without the PMT constraint. The net freshwater input into the Atlantic (34°-40°N) basins are shown in Table 2, columns (b) and (c), respectively - surface mass budgets are similar in simulations using the PMT constraint, with both MERRA-2 and JRA55-do forcing. Likewise, Exp:M and Exp:M* exhibit consistent Atlantic/Arctic sector mass budgets, but Exp:J has a relatively weak Atlantic hydrologic intensity.

The MERRA-2 simulations, with and without the PMT constraint, have similar Atlantic overturning rates while Exp:M-ref has a significantly stronger overturning, which is consistent with reduced freshwater input into the North Atlantic. Exp:J*, produces a much stronger overturning circulation compared to MERRA-2, despite having a similar Atlantic mass budget, as indicated in columns (b, c). This appears to be due to locally stronger wintertime surface heat exchange in the northwest Atlantic using JRA55-do data (Supporting Information S1).

3.1. Sensitivity of Multi-Decadal Water-Mass Drift

As shown in Figure 2, global volume-average temperature drift from initialization is strongly impacted by the hydrologic closure. With the original scheme, the model rapidly cools to depth, whereas experiments using the updated closure with land storage and with the PMT constraint, exhibit substantially less initialization drift. Rapid initial cooling and stronger AMOC intensity without land storage is associated with an under-estimation of the subpolar-to-polar mass fluxes in both hemispheres.

In Figures 3–7, global and regional watermass tendencies for the entire hindcast period are evaluated in selected depth ranges. Area-average temperature, salinity and potential density anomalies relative to the 2005–2018 time period are shown at the top and bottom interfaces for each depth bin, in addition to the volume-average anomalies (solid). Plots are oriented with density increasing downward. In the upper 1500 m, timeseries derived from the

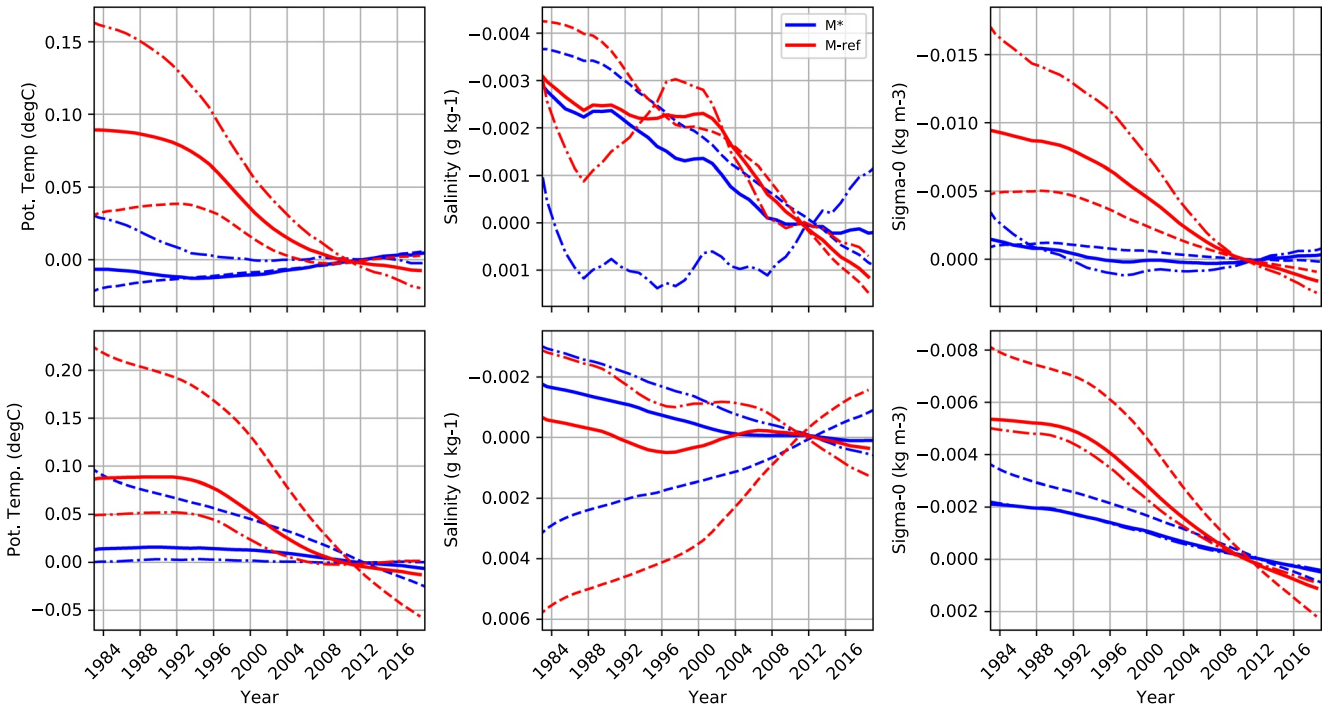

Figure 4. Global 1500–4500 m anomalies, as in Figure 3 (without observational estimates) for the (top) 1500–2500 m and (bottom) 2500–4500 m depth ranges.

EN4 and ARGO analyses are shown for comparison. Below 1500 m, sufficient long term global observations do not exist.

Figure 3 shows global watermass tendencies from 60°S to 60°N in the upper 1500 m starting from WOA05 initial conditions for Exp:M-ref and Exp:M*. This evaluates the proposed hydrologic closure, including both

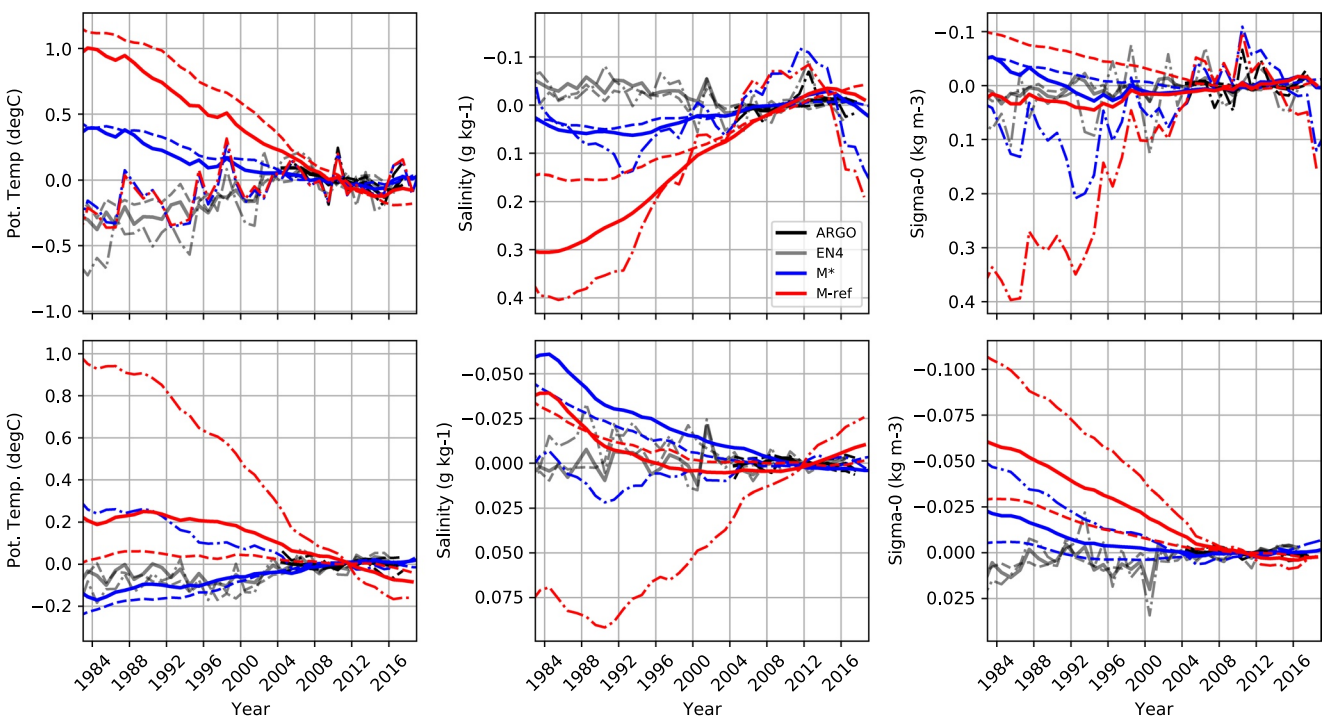

Figure 5. 0–1500 m anomalies, as in Figure 3 in the North Atlantic (0–65N).

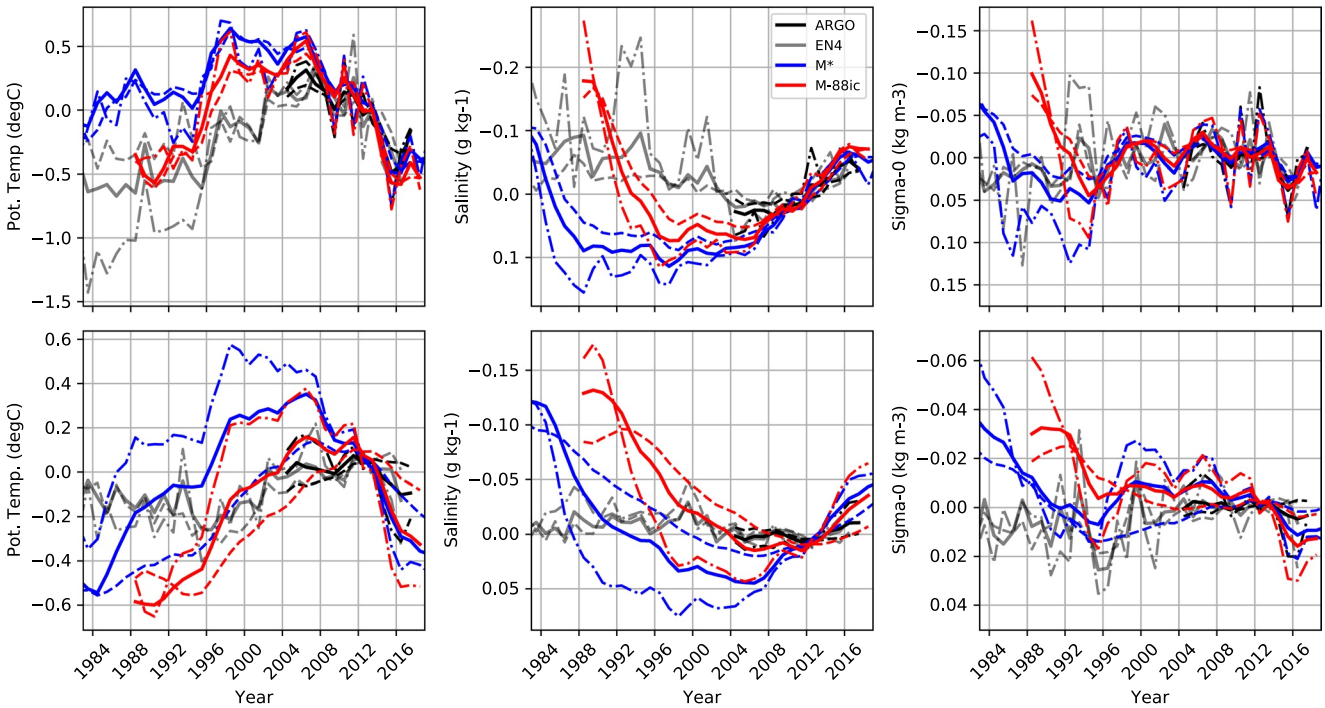

Figure 6. As in Figure 3 in the 0–1500 m subpolar North Atlantic (60–0W; 45–65N) for Exp:M* and Exp:M-88ic.

land storage and the PMT constraint, relative to the reference simulation, which neglects land storage. Focusing on column 1—both simulations produce nearly identical surface temperature evolution, which is reflective of the strong coupling to the reanalysis data at the surface. However, in the subsurface through 500 m, persistent cooling occurs during the first two decades in Exp:M-ref, with temperatures at 500 m averaging roughly 0.5

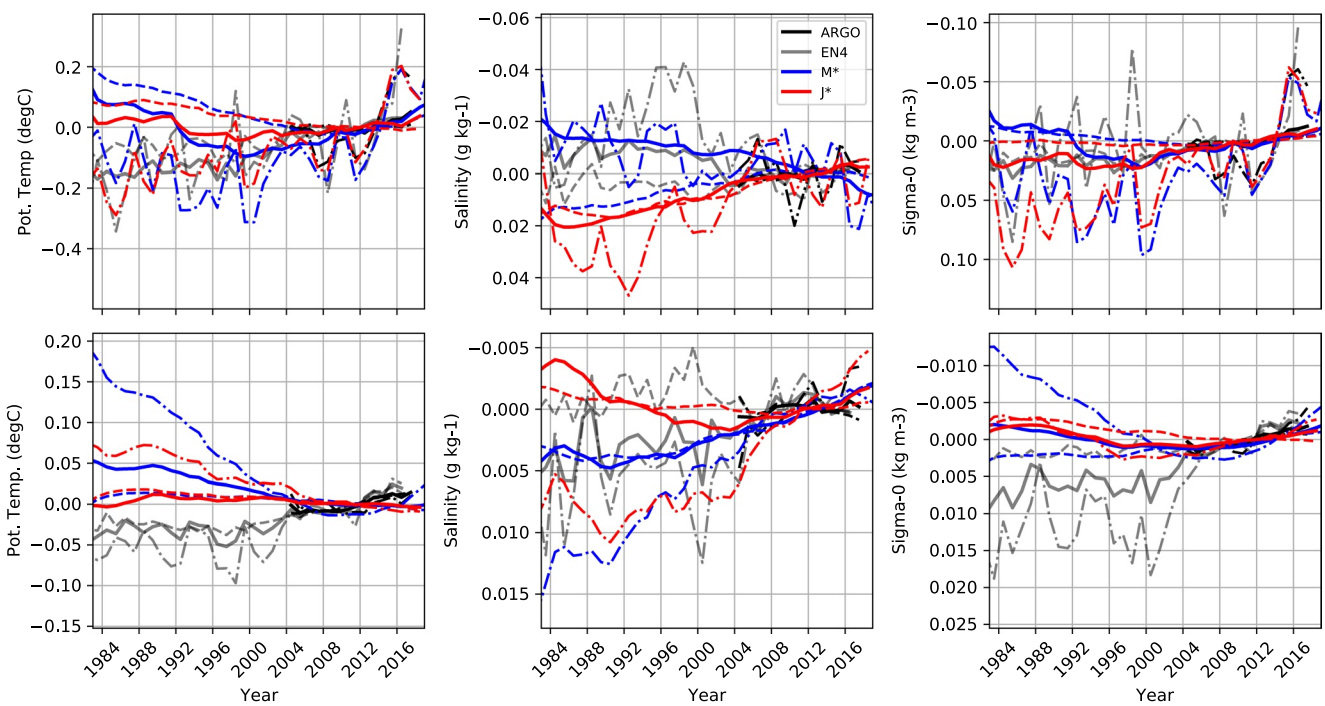

Figure 7. Global 0–1500 m anomalies, as in Figure 3 for Exp:M* and Exp:J*.

degC cooler after 20 years of simulation. This initial cooling is reduced by nearly a factor of four in Exp:M* on average in the upper 500 m, due solely to the modified freshwater closure. Below 1500 m, we observe similar reductions in global watermass drift (Figure 4), where an initial cooling exceeding 0.2 degC (1500–4500 m) in Exp:M-ref is substantially reduced in Exp:M*. Upper ocean salinity is well outside the observational range in Exp:M-ref (Figure 3, column 2). This is a ramification of the improper hydrologic closure, and the lack of surface salinity restoring in this simulation makes clear the degradation that results. In contrast, upper 500 m salinity in Exp:M* is broadly consistent with independent observational estimates, with an increasing trend over the simulation period, and freshening between 500 and 1500 m (Cheng et al., 2020; Durack et al., 2012). Below 1500 m (Figure 4, col. 2), both Exp:M* and Exp:M-ref produce similar rates of average salinity change over the course of the simulation. The vertical distribution is, however, impacted with Exp:M* having reduced drift near 1500 m and at depth. Global potential density drift is substantially reduced with the updated mass closure, both in the upper 500 m, where the improvements are dominated by both temperature and salinity, and below 1500 m, where drift reductions are dominated by temperature (Figure 4, col. c).

In addition to global drift reduction with the updated land closure, basin-scale model improvements occur as well. For example, North Atlantic (0°N–65°N) 0–1500 m tendencies are shown in Figure 5—upper ocean initial cooling (panel a) is reduced by more than a factor of two with the proposed closure. In the 500–1500 m depth range, Exp:M* is better correlated with observations of moderate warming since the mid-1990s (panel d), while in contrast, Exp:M-ref exhibits a dramatic long-term cooling trend, reflective of the trend above, through most of the integration period. The cooling trend in Exp:M-ref compensates strong surface intensified freshening in this region (panel b). Exp:M-ref exhibits strong upper-ocean freshening and density decrease through the first two decades of the simulation, whereas Exp:M* is more consistent with observations. Between 500 and 1500 m, Exp:M* is however somewhat *less* consistent with observations. Southern Ocean hydrologic sensitivity (see supplemental figures in Supporting Information S1) demonstrates notable reductions in salinity and potential density initialization drift as well. Overall, the new land closure significantly reduces longer term model drift, both globally and at basin scales in this model configuration, although considerable regional model biases remain.

An additional experiment (Exp:M*-z), demonstrates the high sensitivity of our model results to ocean numerical representation. In this case, Exp:M*-z (Figure 2, panel d) - which is identical to Exp:M*, but uses a fixed 75 level fixed geopotential coordinate system - substantial warming occurs through the upper kilometer of the global ocean, due to the more diffusive nature of this coordinate, relative to the more closely adiabatic hybrid coordinate used in the experiments presented in this study and in A19.

3.2. Sensitivity to Initial Conditions

We partially evaluate the sensitivity of our results to ocean initial conditions, using an additional experiment, starting in January 1988 and using ocean reanalysis data from SODA (Exp:M*-88), rather than WOA05.

Globally, in comparison to the control simulation (Exp:M*), this hindcast produced similar results approximately 10–15 years following initialization (see supplemental figure in Supporting Information S1). In the North Atlantic (60-0°W; 40–65°N, Figure 6)—initial upper 500 m temperatures are consistent with observations in Exp:M*-88 (due to the assimilation of profile data in SODA), however the model experiences excess initial warming in the upper 1500 m and subsequently diverges from observations toward the control solution. This is accompanied by an initial 0–1500 m increasing salinity and potential density drift which mimics initial trends in Exp:M*. These limited results suggest that the inability of the model to retain observed watermass characteristics in the SPNA. Additional factors, beyond the hydrologic closure presented here, should be considered. Further evaluation of model initial drift sensitivity - related to, for example, forcing uncertainty and ocean representation of eddies and diabatic processes, is beyond the scope of the present study.

3.3. Sensitivity to Reanalysis Forcing Data

The JRA55-do forced simulations (Exp:J* and Exp:J-ref) had a similar sensitivity to land closure, exhibiting broad reductions in drift from initialization using the update mass closure. Despite having identical hydrologic constraints, Exp:M* and Exp:J* differ significantly (Figure 7) - for example, Exp:M* exhibits stronger initial global cooling between 500 and 1500 m, relative to Exp:J*, which is more consistent with observations. Salinity

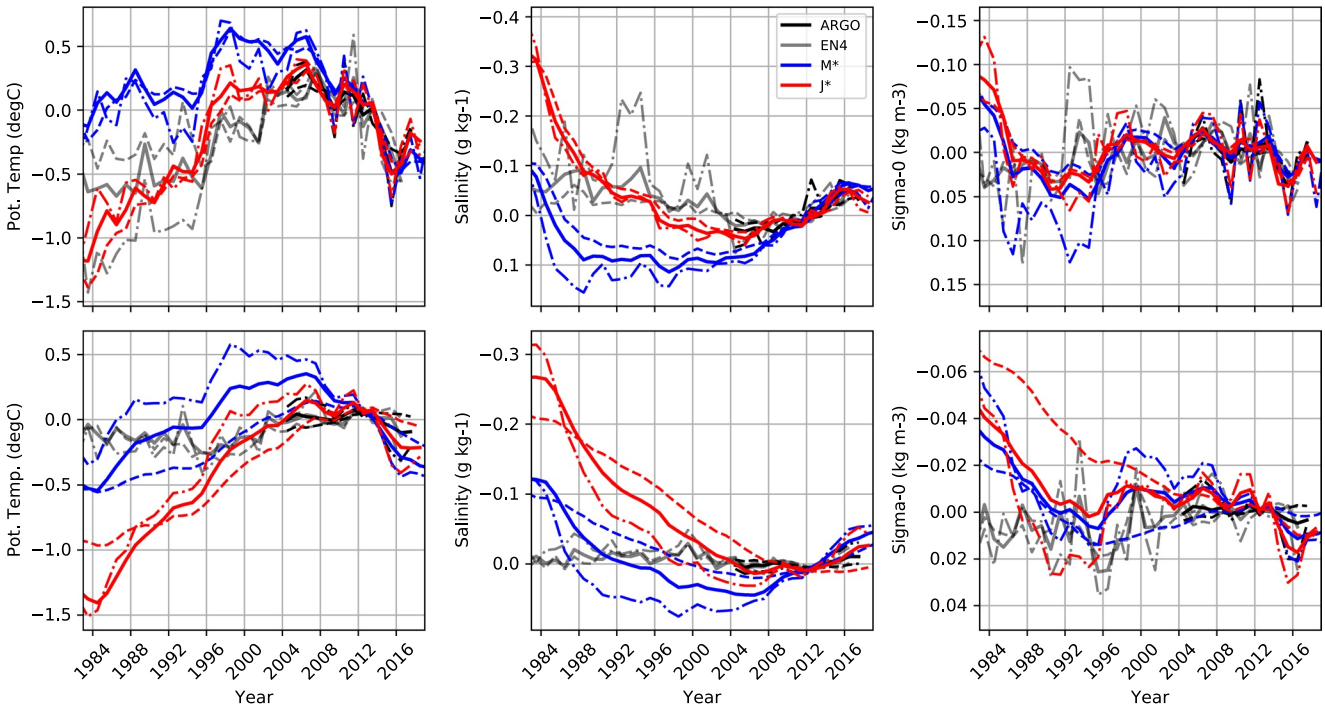

Figure 8. As in Figure 6 in the subpolar north Atlantic for Exp:M* and Exp:J*.

trends are broadly consistent with observations in Exp:M*, while Exp:J* exhibits opposing salinity tendencies in the upper 1500 m.

In the SPNA (Figure 8), both reanalysis simulations produce unrealistic initial freshwater decrease in the upper 1500 m, with somewhat larger drift in Exp:J*. This is consistent with stronger AMOC intensity in Exp:J* and more intense deep convection in this region. Further examination of the air-sea fluxes in SPNA reveal stronger wintertime cooling in this region due to colder/drier continental air-masses in JRA55-do, relative to MERRA-2 (see supplemental figures in Supporting Information S1). An additional simulation was performed using neutral turbulent exchange coefficients (Large & Yeager, 2009) instead of stability-dependent coefficients, with negligible impact on our results.

4. Discussion and Conclusions

We have demonstrated that a global ocean/sea-ice model configuration has as strong sensitivity to a more realistic hydrologic closure. Unlike in previous studies, in order to prevent spurious sea level drift in the model we account for seasonal land storage while re-scaling precipitation. Accounting for seasonal land storage using a land model, with its associated network of groundwater, lakes and rivers, significantly improved the model results. Neglecting land storage, for example, yields unrealistic rates of ocean heat uptake during the analysis period, and significantly larger drifts from initialization in this particular model configuration. Additionally, the state of AMOC shows improved model-observation agreement with the proposed closure, particularly using MERRA-2 data.

In addition to taking into account land storage, we investigated the sensitivity to a constraint on the implied mid-latitude atmospheric moisture transport. The underlying importance of the hydrologic cycle to ocean ventilation is clear, but has been largely ignored in previous studies, due perhaps to a lack of confidence in its representation in models and reanalyzes. For contextual reference, the original Stommel 3-box model of the thermohaline circulation (Stommel, 1961) relied on a negative prescribed salinity flux in the polar regions and a positive salinity flux in the tropics/sub-tropics. In the same manner, we use the climatological flux of moisture to the polar regions (or mid-latitude PMT) as a more physically plausible boundary condition for the model, rather than being reliant on surface salinity nudging.

Based on our analysis, CMIP6 historical simulations are in relatively good agreement with regard to present-day mid-latitude PMT, and our hindcast simulations indicate that the CMIP6-based PMT constraint is furthermore consistent with a significant reduction in long-term drift in this model configuration. The crude nature of the hydrologic closure, which introduces a discontinuity in precipitation at the mid-latitude boundary, needs to be refined; however, these sensitivity results suggest the feasibility of this approach. The relative agreement within CMIP6 models is perhaps indicative of its utility as a surface mass balance constraint, although there are likely systematic biases within the ensemble owing for example, to common physics and resolution. Fluxes from the high resolution model inter-comparison project (Haarsma et al., 2016) could be compared to CMIP6, for example, Additional sensitivity tests were performed in order to examine the dependence of our results on the PMT constraint. For example, increasing PMT by 10% (see supplemental figures in Supporting Information S1) showed a moderate impact on the results presented here. This suggests the need for further evaluation of the PMT constraints used in this study, based on additional independent model data.

Despite the fact that these hindcasts are not assimilating ocean observations, there is encouraging agreement using the proposed hydrology closure. For example, hindcasts Exp:M and Exp:M* have an implied top of atmosphere radiative imbalance of 0.4 and 0.3 W m^{-2} during the recent decade, respectively, with the associated warming occurring above 1500 m (see supplemental figures in Supporting Information S1). A recent estimate for the 1955–2017 period is similar, at approximately 0.3 W m^{-2} (Zanna et al., 2019), however some observational estimates for recent decades are higher (Balmaseda et al., 2013).

Initializing the model with a start date of 1988 using SODA reanalysis (instead of WOA05) gives similar results during the 2005–2018 analysis period. This suggests that our results are potentially less sensitive to initialization, however this requires further evaluation.

Longer term model drift is sensitive to other factors, in addition to hydrologic closure—for example, switching to a fixed-geopotential ocean vertical coordinate, which is inherently more diffusive (Ilıcak et al., 2012), causes the global temperature initialization drift to change sign from cooling to warming. Sensitivity to model physical and numerical representation should be considered in future studies, in conjuction with potential further refinements to hydrologic closure, in order to more fully evaluate the proposed methodology. Further model diagnosis of mode, intermediate and deep water masses and transports should addressed in future studies in order to better evaluate sources of model drift.

We applied the proposed hydrologic closure to two atmospheric reanalyzes: MERRA-2 and JRA55-do. The JRA55-do forced simulations are somewhat degraded relative to MERRA-2, with larger rates of deep ocean heat uptake and a stronger AMOC, which is attributed to elevated SPNA wintertime cooling in these simulations. Further investigation of these differences is warranted, and additional atmospheric reanalysis products could help to better constrain the air-sea fluxes above dense water formation regions.

Our results suggest that modifications to existing protocols (e.g., Griffies et al., 2016) should properly take into account land storage as well as PMT in order to reduce and potentially eliminate their reliance on artificial salinity nudging. An additional simulation was performed using the new methodology, and with additional salinity nudging toward a pre-computed monthly climatology (as in A19). Results were sensitive to nudging (see supplemental figure in Supporting Information S1), but we suggest that its usage of should be discarded in favor of more physically based approach. This of course comes at the cost of added model software complexity—with the introduction of a land model component, however, recent developments in community-based global coupled model infrastructure (e.g., Sandgathe et al. (2011)), should facilitate their utilization for this purpose. As an alternative to using an explicit land model component, seasonally varying land storage estimates within latitude bins could be introduced, in addition to the PMT constraint in order to similarly adjust ocean precipitation in a consistent manner.

Data Availability Statement

The run-time configuration of the numerical model results presented in this study, and scripts for generating output are archived with an associated https://doi.org/10.5281/zenodo.6342240.

Acknowledgments
The authors would like to thank Raphael Dussin and Charles Stock for their valuable feedback on an early version of this manuscript. Many discussions with Tom Delworth and Tony Rosati provide useful guidance. We would also like to thank Elena Shevliakova, Sergey Malyshev, Chris Milly and Krista Dunne for their assistance with the land model configuration. A. Adcroft and O. Sergienko were supported by award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, US Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the US Department of Commerce.

References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167–3211. https://doi.org/10.1029/2019MS001726

Antonov, J. I., Locarnini, R., Boyer, T., Mishonov, A., Garcia, H., & Levitus, S. (2006). *World Ocean atlas 2005 volume 2: Salinity*. Noaa atlas nesdis. (2).

Balaji, V., Anderson, J., Held, I., Winton, M., Durachta, J., Malyshev, S., & Stouffer, R. J. (2006). The exchange grid: A mechanism for data exchange between Earth system components on independent grids. *Parallel computational fluid dynamics*, 179–186.

Balmaseda, M. A., Trenberth, K. E., & Källén, E. (2013). Distinctive climate signals in reanalysis of global ocean heat content. *Geophysical Research Letters*, *40*(9), 1754–1759. https://doi.org/10.1002/grl.50382

Behrens, E., Biastoch, A., & Böning, C. W. (2013). Spurious AMOC trends in global ocean sea-ice models related to subarctic freshwater forcing. *Ocean Modelling*, *69*, 39–49. https://doi.org/10.1016/j.ocemod.2013.05.004

Carton, J. A., Giese, B. S., & Grodsky, S. A. (2005). Sea level rise and the warming of the oceans in the Simple Ocean Data Assimilation (SODA) ocean reanalysis. *Journal of Geophysical Research*, *110*(C9), C09006. https://doi.org/10.1029/2004jc002817

Charnock, H. (1955). Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, *81*(350), 639–640. https://doi.org/10.1002/qj.49708135027

Cheng, L., Trenberth, K. E., Gruber, N., Abraham, J. P., Fasullo, J. T., Li, G., et al. (2020). Improved estimates of changes in upper ocean salinity and the hydrological cycle. *Journal of Climate*, *33*(23), 10357–10381. https://doi.org/10.1175/jcli-d-20-0366.1

Cullather, R. I., & Bosilovich, M. G. (2011). The moisture budget of the polar atmosphere in MERRA. *Journal of Climate*, *24*(11), 2861–2879. https://doi.org/10.1175/2010jcli4090.1

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., Bi, D., et al. (2014). North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I: Mean states. *Ocean Modelling*, *73*, 76–107. https://doi.org/10.1016/j.ocemod.2013.10.005

Durack, P. J., Wijffels, S. E., & Matear, R. J. (2012). Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, *336*(6080), 455–458. https://doi.org/10.1126/science.1212222

Ferreira, D., Marshall, J., & Campin, J.-M. (2010). Localization of deep water formation: Role of atmospheric moisture transport and geometrical constraints on ocean circulation. *Journal of Climate*, *23*(6), 1456–1476. https://doi.org/10.1175/2009jcli3197.1

Frajka-Williams, E., Moat, B., Smeed, D., Rayner, D., Johns, W., Baringer, D., et al. (2021). *Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26n from 2004 to 2020 (v2020.1)*. British Oceanographic Data Centre—Natural Environment Research Council. https://doi.org/10.5285/cc1e34b3-3385-662b-e053-6c86abc03444

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*(12), 6704–6716. https://doi.org/10.1002/2013jc009067

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Boning, C. W., et al. (2016). OMIP contribution to CMIP6. *Geoscientific Model Development*, *9*(9), 3231–3296. https://doi.org/10.5194/gmd-9-3231-2016

Griffies, S. M., Pacanowski, R. C., Schmidt, M., & Balaji, V. (2001). Tracer conservation with an explicit free surface method for z-coordinate ocean models. *Monthly Weather Review*, *129*(5), 1081–1098. https://doi.org/10.1175/1520-0493(2001)129<1081:tcwaef>2.0.co;2

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High resolution model intercomparison project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development*, *9*(11), 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Harrison, M., Adcroft, A., & Hallberg, R. (2014). Atlantic watermass and circulation response to persistent freshwater forcing in two coupled general circulation models. *Climate Dynamics*, *42*(1), 59–68. https://doi.org/10.1007/s00382-013-1798-5

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., et al. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691–3727. https://doi.org/10.1029/2019MS001829

Ilıcak, M., Adcroft, A. J., Griffies, S. M., & Hallberg, R. W. (2012). Spurious dianeutral mixing and the role of momentum closure. *Ocean Modelling*, *45*, 37–58. https://doi.org/10.1016/j.ocemod.2011.10.003

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

Locarnini, M., Mishonov, A., Baranova, O., Boyer, T., Zweng, M., Garcia, H., et al. (2018). *World Ocean atlas 2018, volume 1: Temperature*. NOAA Technical Reports.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, *24*(14), 3624–3648. https://doi.org/10.1175/jcli-d-11-00015.1

Roemmich, D., Johnson, G. C., Riser, S., Davis, R., Gilson, J., Owens, W. B., et al. (2009). The Argo program: Observing the global ocean with profiling floats. *Oceanography*, *22*(2), 34–43. https://doi.org/10.5670/oceanog.2009.36

Sandgathe, S., O'Connor, W., Lett, N., McCarren, D., & Toepfer, F. (2011). National unified operational prediction capability initiative. *Bulletin of the American Meteorological Society*, *92*(10), 1347–1351. https://doi.org/10.1175/2011bams3212.1

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, *19*(13), 3088–3111. https://doi.org/10.1175/jcli3790.1

Stommel, H. (1961). Thermohaline convection with two stable regimes of flow. *Tellus*, *13*(2), 224–230. https://doi.org/10.3402/tellusb.v13i2.12985

Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., et al. (2018). JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean Modelling*, *130*, 79–139. https://doi.org/10.1016/j.ocemod.2018.07.002

Wijffels, S. E., Schmitt, R., Bryden, H. L., & Stigebrandt, A. (1992). Transport of freshwater by the oceans. *Journal of Physical Oceanography*, *22*(2), 155–162. https://doi.org/10.1175/1520-0485(1992)022<0155:tofbto>2.0.co;2

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., & Heimbach, P. (2019). Global reconstruction of historical ocean heat storage and transport. *Proceedings of the National Academy of Sciences*, *116*(4), 1126–1131. https://doi.org/10.1073/pnas.1808838115

Zhao, M., Golaz, J.-C., Held, I. M., Guo, H., Balaji, V., Benson, R., et al. (2018). The GFDL global atmosphere and land model AM4. 0/LM4. 0: 2. Model description, sensitivity studies, and tuning strategies. *Journal of Advances in Modeling Earth Systems*, *10*(3), 735–769. https://doi.org/10.1002/2017ms001209