Climate Dynamics manuscript No.
(will be inserted by the editor)

Global heat balance and heat uptake in potential temperature coordinates

Antoine Hochet · Rémi Tailleux · Till Kuhlbrodt · David Ferreira

Received: date / Accepted: date

Abstract The representation of ocean heat uptake in Simple Climate Models used for policy advice on climate change mitigation strategies is often based on variants of the one-dimensional Vertical Advection/Diffusion equation (VAD) for some averaged form of potential temperature. In such models, the effective advection and turbulent diffusion are usually tuned to emulate the behaviour of a given target climate model. However, because the statistical nature of such a "behavioural" calibration usually obscures the exact dependence of the effective diffusion and advection on the actual physical processes responsible for ocean heat uptake, it is difficult to understand its limitations and how to go about improving VADs. This paper proposes a physical calibration of the VAD that aims to provide explicit traceability of effective diffusion and advection to the processes responsible for ocean heat uptake. This construction relies on the coarse-graining of the full three-dimensional advection diffusion for potential temperature using potential temperature coordinates. The main advantage of this formulation is that the temporal evolution of the reference temperature profile is entirely due to the competition between effective diffusivity that is always positive definite, and the water mass transformation taking place at the surface, as in classical water mass analyses literature. These quantities are evaluated in numerical simulations of present day climate and global warming experiments. In this framework, the heat uptake in the global warming experiment is attributed to the increase of surface heat flux at low latitudes, its decrease at high latitudes and to the redistribution of heat toward cold temperatures made by diffusive flux.

Keywords Heat uptake · Simple Climate Model

A. Hochet
Univ of Brest, CNRS, Ifremer, IRD, Laboratoire d'Océanographie Physique et Spatiale (LOPS, UMR 6523), IUEM, Brest, France.
E-mail: antoine.hochet@univ-brest.fr

R. Tailleux, T. Kuhlbrodt and D. Ferreira
Department of Meteorology, University of Reading, UK

1 Introduction

Ocean heat uptake is of great importance in climate change predictions: 90 % of the anthropogenic increase in heat stored in the climate system ends up in the oceans ([17]), thus contributing to sea level rise via thermal expansion. The main effects controlling the heat balance include the upwelling of deep water driven by the Southern Ocean winds, cooling by deep water formation, as well as isopycnal and diapycnal mixing, most of which require to be parameterized in current AOGCMs (see for instance [18] and references therein). The ocean heat uptake efficiency, defined as the ratio of net heat flux into the climate system over the global mean surface air temperature change ([7,22]), has been found to vary by a factor of 2 across CMIP5 models ([15]) outlining its high sensitivity to the parametrization choices made by the various modelling groups. A better understanding of the heat balance processes will thus also help constrain mixing parameters in these models. One of the most common method to rationalize the heat balance in the ocean consist in studying its vertical structure from its horizontally-averaged properties. This method is justified by its simplicity but also by the interest in the vertical structure of the temperature which is linked with the idea of ocean heat storage.

The vertical heat transport described by the horizontally-averaged heat balance is often compared with the theory of early models of the deep circulation such as [28] where dense water downwells at high latitude due to convection in very localised regions and upwells uniformly at mid and low latitudes. This leads to the classical view where the upwelling of cold/dense waters is balanced by downward diffusion of heat. This model is frequently referred to as the one-dimensional Vertical Advection/Diffusion (VAD hereafter) model in the literature. So far, however, it has proved difficult to reconcile the classical view of heat balance offered by the VAD model with that resulting from numerous studies of the horizontally-averaged heat balance such as [6,27,16,3]. Indeed, in such studies the horizontally averaged advective heat fluxes are often found to be downward and the horizontally averaged diffusive or eddy-resolved heat fluxes (thus the average of a combination of iso and diapycnal diffusive fluxes) upward, which is seemingly the opposite of what the standard VAD model predicts ([6]).

Yet, the VAD model appears nevertheless successful at emulating the temperature variations of complex AOGCM ([21]). As a result, the VAD model has formed the basis for the one-dimensional representation of ocean heat uptake in Simple Climate Models (SCMs) such as MAGICC ([20]). SCMs are used for instance to evaluate the amount of CO_2 that can be released in the atmosphere before reaching the 2 °C limit ([19]) and play an important role in policy making decisions about global warming mitigation strategies.

To reconcile these two approaches, [14] proposed to calibrate the VAD equation (i.e. the set-up of vertical velocity w and diffusive coefficient K) using a physical approach rather than the behavorial approach used in previous studies such as [21]. The two approaches differ in that the behavioural approach calibrates the VAD model parameters to mimic the temperature variations of complex AOGCMs using statistical techniques, whereas the physical approach seeks to calibrate such parameters by linking them to the processes that control them.

However, when horizontal averaging is used as the underlying basis for the physical calibration, the diffusion coefficient can occasionally be negative owing to the heat diffusion being occasionally upward in parts of the ocean. Moreover, the

time variation of K and w were found crucial in emulating correctly the temperature of AOGCM thus complicating the practical implementation of the method. We have thus identified the two following points: 1) The possibility to justify the VAD model from horizontally-averaging the three-dimensional advection/diffusion equation for heat is far from obvious; 2) the occasional up-gradient nature of the horizontally-averaged heat flux complicates the construction of a one-dimensional VAD model because it does not act to reduce the vertical temperature gradient as is expected physically. To circumvent this difficulty, we adopt a different approach: instead of averaging on constant depth surfaces we average on constant potential temperature (θ hereafter) surfaces, following an approach similar to that recently developed by [11]. The averaged diapycnal diffusion is then automatically downgradient and we will further show that the advection through θ surfaces is in theory zero, leading to a much simpler equation than that obtained with constant depth surfaces. [4] has used a similar approach to study the ocean heat transport in order to filter out any recirculation of waters at constant temperature. [12] also used a similar approach to study the diathermal heat transport in a global ocean sea ice model.

The heat balance averaged in temperature coordinates can be expected to be quite different from the well studied horizontally averaged heat balance in depth coordinates. Indeed, because nearly all isotherms outcrop at the ocean surface, heat fluxes through the coldest temperature classes may either reflect processes at great depth or at high latitudes. In the standard VAD model heat fluxes through the coldest horizontally averaged temperature only pertain to processes at great depth. It might be useful to keep in mind the results from horizontal averages of AOGCM outputs in Control Run (CR hereafter) with constant present day CO_2 concentration and warming climates (see for instance [6, 13, 1, 16]). In CR, the strongest downward heat transport comes from the mean advection while the largest upward heat transport comes from eddy induced advection (resolved or parametrized). In warming climates, the heat uptake takes place mostly in the Southern ocean and is due to the reduction of along-isopycnal mixing and of deep convection. We analyse the outputs of the ocean component of the HiGEM1.2 coupled atmosphere ocean general circulation model (AOGCM), which include a detailed set of temperature tendency diagnostics. HiGEM1.2 is a CMIP5-type model and this study thus contributes to the understanding of heat uptake in this class of models. To analyse the processes controlling ocean heat uptake, we study the heat balance in temperature coordinates first in a control run of the HiGEM model that we then compare to a warming climate run where the pre-industrial CO_2 has been doubled.

The article is organized as follows: in section 2, we derive an alternative one dimensional equation of heat uptake using potential temperature coordinates and show that it allows to remove the effect of advection and to obtain a downgradient diffusion. In section 3, we apply this new method to the study heat uptake first in the CR of HiGEM, then on a simulation where CO_2 concentration is doubled. The last section concludes and discusses the results.

2 Method

Because AOGCM outputs are generally averaged over a period of time (1 month here) all terms of the temperature budget are decomposed into time mean and anomalies:

$$X = \overline{X} + X' \tag{1}$$

where X represents any term of the heat budget, $\overline{(.)}$ the monthly average and $(.)'$ the deviation from this monthly average so that $\overline{X'} = 0$. The time mean potential temperature $\overline{\theta}$ conservation can be written as:

$$\frac{\partial \overline{\theta}}{\partial t} + \overline{\mathbf{v}} \cdot \nabla \overline{\theta} + \overline{\nabla \cdot \mathbf{v}'\theta'} = \nabla \cdot \left(\mathbf{K} \nabla \overline{\theta} \right) + \overline{\mathrm{VM}} + \overline{\mathrm{Q_{net}}} \tag{2}$$

For clarity we will drop the overline notation in what follows and keep it only when it involves anomalies. $\mathbf{v}$ is the 3D velocity vector, $\overline{\nabla \cdot \mathbf{v}'\theta'}$ is a term representing the effect of sub-monthly advection, $\mathbf{K}$ a diffusion tensor representing the effect of unresolved advection and small-scale irreversible mixing, $\mathbf{K}\nabla\theta \cdot \mathbf{k}$ with $\mathbf{k}$ the upward unit vector is thus zero at the surface. $\mathbf{K}$ thus contains the parameterization of both the isopycnal and diapycnal mixing terms. VM is a term representing all parameterized non-diffusive terms like convection and $\mathrm{Q_{net}}$ the net heat flux through the surface. For comparison, the equations of the physical calibration of the VAD using the horizontal average of equation (2) are derived in appendix A. Building on [25]'s work, we first define a reference level z_r of the temperature θ. The use of a reference level will be useful to obtain an 1D evolution equation for the temperature along surfaces of constant reference depth as will become clear below. z_r is the depth of isotherm θ in the reference state which is obtained after an adiabatic rearrangement of each fluid parcel so that isotherms are horizontal and in ascending order. Note that unlike the reference state described in [25], this reference state is not a state of rest because the density is here also a (non-linear) function of salinity and pressure. Such a rearrangement being volume conserving, the reference depth z_r can thus be computed using the fact that the volume of water with temperature larger than θ is the same after and before the adiabatic rearrangement i.e.:

$$\int_{z_r}^{0} A(z)\mathrm{d}z = \int_{V(\theta,t)} \mathrm{d}V, \tag{3}$$

where $V(\theta, t)$ is the volume of ocean with temperature θ_l satisfying $\theta < \theta_l < \theta_{max}$ with θ_{max} the maximum temperature in the ocean and $A(z)$ is the ocean area at depth z. The definition (3) of the reference depth makes it possible to rewrite the temperature $\theta(x, y, z, t)$ as a function of z_r: $\theta_r(z_r, t) = \theta(x, y, z, t)$. θ_r can be inverted to yield $z_r = z_r(\theta, t)$ or $z_r = z_r(x, y, z, t)$. Note that Eq. (3) shows that the volume $V(\theta, t) = V(z_r)$ of water of temperatures greater than θ is a function of z_r alone and hence that it can be treated as a constant independent of time at fixed z_r. An alternative definition of z_r, that can be found for instance in [25], is:

$$z_r(x, y, z, t) = \int_V \mathcal{H}\left[\theta(\tilde{x}, \tilde{y}, \tilde{z}, t) - \theta(x, y, z, t)\right] \mathrm{d}\tilde{V} \tag{4}$$

where $\mathcal{H}$ is the Heaviside step function and V represents the ocean volume. The schematic shown on figure 1 summarizes the calculation of the reference depth as explained above.

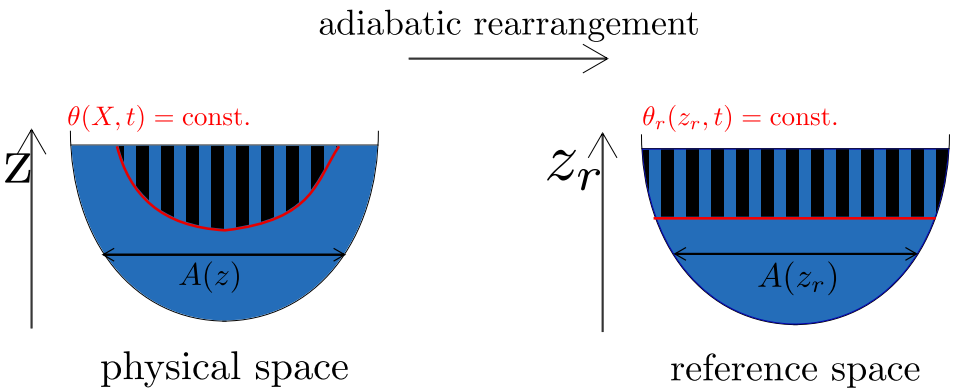

Fig. 1 Schematic showing how the reference depth z_r associated with temperature θ is obtained. On the left is the physical space, on the right the reference space which is obtained through an adiabatic rearrangement of all fluid parcel in the physical space. Isotherms in the reference space are horizontal and only depend on z_r, temperature in the reference space is described by the function $\theta_r(z_r, t)$. The volume for all water parcels warmer than $\theta(X, t) = \text{const.}$ is shown by black stripes in both physical and reference space. This volume is the same in both spaces, and this property is used in formula (3) with $A(z)$ the ocean area at depth z to compute the reference depth associated with the temperature.

We now seek an evolution equation for $\theta_r(z, t)$ by integrating (2) over the volume $V(z_r)$, which after some manipulation yields:

$$\int_{V(z_r)} \frac{\partial \theta_r}{\partial t} \mathrm{d}V + \theta_r(z_r, t) \int_{z_r=\text{const.}} \mathbf{v} \cdot \mathbf{n} \mathrm{d}S + \int_{V(z_r)} \nabla \cdot \overline{\mathbf{v}'\theta'} \mathrm{d}V = \int_{z_r=\text{const.}} \mathbf{K}\nabla\theta \cdot \mathbf{n} \mathrm{d}S + \int_{V(z_r)} \mathrm{VM} \mathrm{d}V + \int_{V(z_r)} \mathrm{Q_{net}} \mathrm{d}V, \quad (5)$$

where $\mathbf{n} = -\nabla\theta/|\nabla\theta| = -\nabla z_r/|\nabla z_r|$ is the outward unit normal vector to the isothermal surface $\theta = \text{constant}$, which at fixed time coincides with the surface $z_r(x, y, z, t) = \text{constant}$. The diffusion term can be written as:

$$\int_{z_r=\text{const.}} \mathbf{K}\nabla\theta \cdot \mathbf{n} \mathrm{d}S = -K_{\text{eff}}(z_r, t) A(z_r) \frac{\partial \theta_r}{\partial z_r}(z_r, t) \quad (6)$$

with $K_{\text{eff}} = \frac{1}{A(z)} \int_S K^{\text{loc}}_{\text{eff}} \mathrm{d}S'$ and $K^{\text{loc}}_{\text{eff}}$ a positive quantity independent of $\partial_{z_r}\theta_r$, indeed:

$$\mathbf{K}\nabla\theta \cdot \mathbf{n} = -\frac{\partial \theta_r}{\partial z_r} \left(K_i \left(\nabla z_r - (\nabla z_r \cdot \mathbf{d})\mathbf{d}\right) + K_d (\nabla z_r \cdot \mathbf{d})\mathbf{d}\right) \cdot \frac{\nabla z_r}{|\nabla z_r|} = -\frac{\partial \theta_r}{\partial z_r} \underbrace{\left(K_i \sin^2(\nabla z_r, \mathbf{d}) + K_d \cos^2(\nabla z_r, \mathbf{d})\right) |\nabla z_r|}_{K^{\text{loc}}_{\text{eff}}} \quad (7)$$

where K_i and K_d are the isoneutral and dianeutral turbulent diffusivities respectively. Using the non-divergence of the velocity field and neglecting the contribution of the freshwater fluxes (whose expression is derived in Appendix B) so that

$w = 0$ at the surface, we have:

$$\int_S \mathbf{v}\cdot\mathbf{n}\mathrm{d}S = \int_{V(z_r)} \nabla\cdot\mathbf{v}\mathrm{d}V = 0. \tag{8}$$

This equation holds even under a non-steady state and means that a closed volume cannot increase or decrease due to advection by a non-divergent velocity through its boundaries. The more general case for which $w(z=0) = E - P + R$ with E, P and R respectively the evaporation precipitation and river run-off is discussed in details in [10] and described briefly in appendix B. Using Eqs. (6) and (8) in Eq. (5) gives:

$$\begin{aligned}\int_{V(z_r)} \frac{\partial\theta_r}{\partial t}\mathrm{d}V = &-K_{\mathrm{eff}}(z_r,t)A(z_r)\frac{\partial\theta_r}{\partial z_r}(z_r,t)\\ &-\int_{V(z_r)} \nabla\cdot\overline{\mathbf{v}'\theta'}\mathrm{d}V + \int_{V(z_r)} VM\mathrm{d}V + \int_{V(z_r)} Q_{net}\mathrm{d}V.\end{aligned} \tag{9}$$

This equation links the volume integral on $V(z_r)$ of the time derivative of the temperature to the diffusive flux, the sub-monthly advection, the vertical mixing and the surface heat flux. Calculating the derivative of eq. (9) with respect to z_r and dividing by $A(z_r)$ gives an evolution equation for θ_r:

$$\begin{aligned}\frac{\partial\theta_r}{\partial t} = &\frac{1}{A(z_r)}\frac{\partial}{\partial z_r}\left(K_{\mathrm{eff}}(z_r,t)A(z_r)\frac{\partial\theta_r}{\partial z_r}\right)\\ -&\frac{1}{A(z_r)}\frac{\partial}{\partial z_r}\int_{V(z_r)} \nabla\cdot\overline{\mathbf{v}'\theta'}\mathrm{d}V - \frac{1}{A(z_r)}\frac{\partial}{\partial z_r}\int_{V(z_r)} VM\mathrm{d}V - \frac{1}{A(z_r)}\frac{\partial}{\partial z_r}\int_{V(z_r)} Q_{net}\mathrm{d}V,\end{aligned} \tag{10}$$

where we have used:

$$\frac{\partial}{\partial z_r}\int_{V(z_r)} \frac{\partial\theta_r}{\partial t}\mathrm{d}V = \frac{\partial}{\partial z_r}\int_{z_r}^{0} A(z)\frac{\partial\theta_r(z,t)}{\partial t}\mathrm{d}z = -A(z_r)\frac{\partial\theta_r(z_r,t)}{\partial t} \tag{11}$$

The possibility to obtain a 1D equation for $\theta_r(z_r,t)$ as given by Eq. (10) is one of the main advantage of the use of a reference level. Note that equation (10) is similar to equation (17) in [26] with vertical mixing, sub-monthly advection and heating terms added. In agreement with [9], Eq. (9) and (10) establish that the time evolution of the reference potential temperature is only a function of the effective diffusion, of the sub-monthly advection, of the forcing and of the vertical mixing. The (resolved) monthly advection does not play any role in the evolution of θ_r and the diffusive part is only due to the divergence of the downgradient diffusive flux. In the remaining of this paper we use Eq. (9) and (10) to study heat uptake in the Control Run and 2x CO_2 run of a climate model.

3 Results

3.1 Model

HiGEM1.2 is an AOGCM pertaining to the CMIP5-type models. It is based on the UK MetOffice coupled AOGCM HadGEM1, but has a higher spatial resolution, of 0.83° lat. × 1.25° lon. (N144) in the atmosphere and 1/3°×1/3° with 40 levels in the ocean. An implicit linear free surface scheme based on [2] with explicit fresh water fluxes is used. Lateral mixing of tracers uses the isopycnal formulation of [8], and the [5] (GM) adiabatic mixing scheme is not used. A detailed description of this model can be found in [24]. We use two different runs of HiGEM1.2: 1) a Control Run (CT hereafter) where present-day boundary conditions are used, in particular, the atmospheric CO_2 concentration is set to 345 ppm, reflecting conditions in the 1980s, and 2) a perturbed run where atmospheric CO_2 concentration is doubled (2× CO_2). The control run length used in this article is 50 years and the 2× CO_2 perturbation run length is 70 years.

The HiGEM diagnostics used here consist in monthly means of the potential temperature tendencies i.e. all terms at each grid point contributing to local changes in potential temperature. These terms comprise potential temperature change due to advection, diffusion (separately in the x, y and z directions), convection, mixed layer physics, ice physics, penetrating solar radiation and other surface fluxes. Note that there is no GM parameterisation, the advection diagnostic thus contains both the mean and resolved eddy-induced advection. We regroup in what follows convection and mixed layer dynamics into a vertical mixing (VM) term and penetrative solar, surface fluxes, ice physics into a forcing term. We are thus left with four terms: diffusion, advection, vertical mixing and forcing.

As seen from equation (5) the integral is performed on volumes defined by surfaces of constant z_r. For each time t, θ = const. surfaces are exactly the same as the z_r = const. surfaces. However, the fact that $\theta_r(z_r, t)$ is also a function of time implies that the temperature associated with a given reference level is time dependent. Practically it means that we need to calculate the reference level for every monthly mean outputs and then perform the volume integral of the tendencies. The method used to calculate the volume integral of the heat tendencies is described in appendix C. The reference levels and volume integral of heat tendencies are calculated for monthly means for both the Control Run and the 2×CO_2 run. They are then averaged over a 50 years period for the CR and on the 70 years of the 2×CO_2 run.

3.2 Control Run

3.2.1 Reference level

The 50 years mean reference level is shown on the left panel of figure 2. As expected, it is a monotonic function of temperature, deepest (shallowest) z_r correspond to coldest (warmest) temperatures. Because most of the volume of the ocean has small temperatures below 5 °C, the range of temperature between −1000 m and 0 m is much larger (∼ 25 °C) than at deeper depth: ∼ 7 °C between −5500 m and −1000 m. The reference temperature gradient will therefore be much larger at

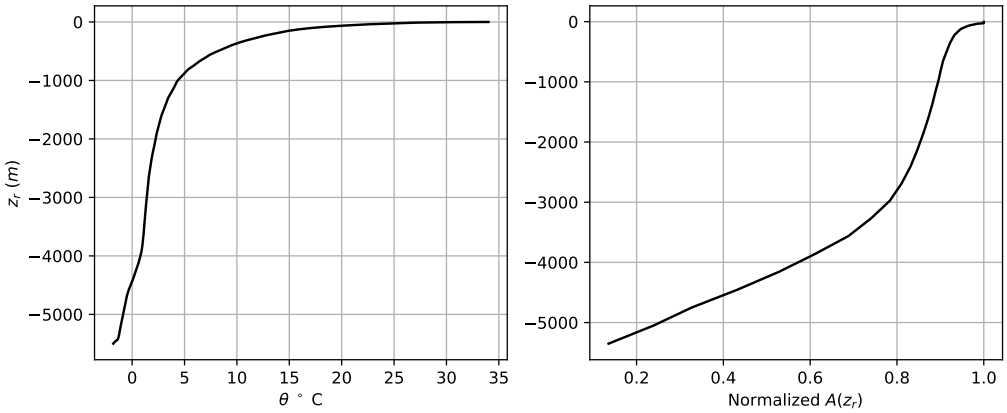

Fig. 2 Left panel: time mean of the reference temperature (in °C) in the CR as a function of the reference depth in meters. Right panel: surface of the ocean $A(z_r)$ normalized by its maximum value as a function of the reference depth.

shallow reference depth than at deep reference depth. The ocean area as a function of depth $A(z)$ calculated for the the HiGEM grid and used in the reference depth calculation (see formula 3) is shown on the right panel of figure 2.

3.2.2 Time mean of the volume integral of the heat tendencies as a function of the reference depth

At each grid cell, heat tendencies are decomposed using the following equation:

$$\frac{\partial \theta}{\partial t} = \text{advection} + \text{diffusion} + \text{VM} + \text{forcing} \tag{12}$$

where "advection", "diffusion", "VM", "forcing" are respectively the three dimensional heat tendencies due to advection, diffusion, VM, and forcing described in the last section. Equation (12) is then integrated on volume $V(z_r)$ described in section 2:

$$\int_{V(z_r)} \frac{\partial \theta}{\partial t} \mathrm{d}V' = \int_{V(z_r)} \text{advection } \mathrm{d}V' + \int_{V(z_r)} \text{diffusion } \mathrm{d}V' + \int_{V(z_r)} \text{VM } \mathrm{d}V' + \int_{V(z_r)} \text{forcing } \mathrm{d}V' \tag{13}$$

Figure 3 shows the time mean volume integral of the heat tendencies as a function of the reference level for the CR.

The time mean of the integral of the tendencies is negative at all the reference depth for the diffusion, advection and vertical mixing and always positive for the forcing. This shows that for all z_r, diffusion, advection and vertical mixing act together to reduce the temperature of the volume of water parcels with z_r' larger

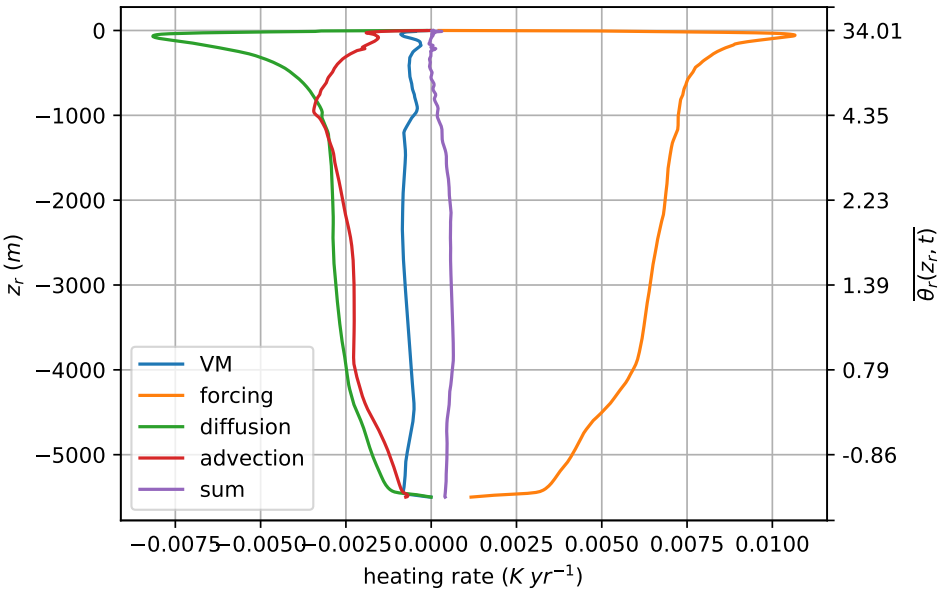

Fig. 3 Volume integral of heat tendencies ($\mathrm{K\,yr^{-1}}$) on $V(z_r)$ associated with vertical mixing (convection+mixed layer dynamcis) in blue, forcing (surface fluxes+ ice) in orange, advection in red, diffusion in green as a function of the reference depth z_r in meters. The time mean of the temperature as a function of z_r is shown on the right. The sum of all terms is shown in purple. The time mean of the temperature (over the CR) as a function of z_r is shown on the right.

than z_r while the forcing acts to increase it. Figure 3 is similar to figure 3 of [12] where the budget for the internal heat content of a global ocean sea ice model is expressed in terms of surface forcing, vertical mixing and "numerical" mixing (which is calculated as a residual and thus contains the isopycnal mixing). Our forcing term looks similar to theirs, the detailed comparison for the two other terms is less straigthforward because they do not represent the same processes as ours but overall the sum of our VM, diffusion and advection terms act as the sum of their "numerical" and vertical mixing i.e. in opposition to the forcing.

The effect of a given tendency term over the entire volume of the ocean is given by its value at the deepest reference depth i.e. $-5500\,\mathrm{m}$. At this depth, the diffusion and vertical mixing are both zero, while the forcing is positive and the effect of advection is negative. The volume integral of the advection is negative because of the imperfect way the free surface boundary condition is formulated in the model as explained in [16]. As explained in section 2, the advection made by the monthly mean velocity on the monthly mean temperature is zero when volume integrated on $V(z_r)$ and is therefore not part of the advection term in equation (13). The volume integral of the forcing on the entire volume of the ocean is positive because of the small control run drift.

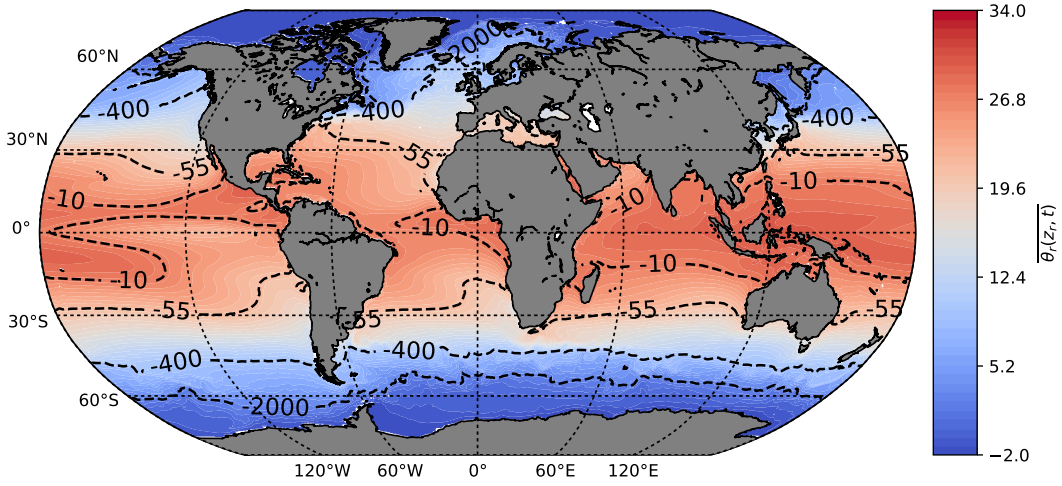

Fig. 4 Time average of the temperature (in °C) at the surface in the CR. $-2000, -400, -55$ and -10 m contours of the reference depth are shown with dashed black contours.

All of the four terms have a large slope change at very shallow reference depth, around -55 m. It is explained by the fact that low and mid latitudes have shallow reference depths because their surface temperature is mostly contained between approximately 10 and 30 °C whereas the deeper reference depths are confined to high latitudes regions (figure 4). The heating thus only occurs for reference depths shallower than -55 m, while the cooling occurs on a much larger range of reference depths: between -5500 m and -55 m.

The negative sign of the volume integrated tendency due to diffusive processes (see figure 3) is consistent with the downgradient nature of heat diffusion. Indeed, writing the diffusion term as the divergence of a downgradient heat flux $\mathbf{F}_{\rm diff}$ as $-\nabla\cdot\mathbf{F}_{\rm diff}$, with $\mathbf{F}_{\rm diff}\cdot\nabla\theta<0$, shows that:

$$\int_{V(z_r)} \text{diffusion}\,\mathrm{d}V = -\int_{z_r=\text{const.}} \mathbf{F}_{\rm diff}\cdot\mathbf{n}\,\mathrm{d}S = \int_{z_r=\text{const.}} \frac{\mathbf{F}_{\rm diff}\cdot\nabla\theta}{|\nabla\theta|}\,\mathrm{d}S < 0 \quad (14)$$

where $\mathbf{F}_{\rm diff}$ is the diffusive flux.

Finally the sum of the advective and diffusive terms almost completely balance the forcing term because the VM is small compared to the three other terms. This is in contrast with the horizontally-averaged heat balance, for which the mean diffusive flux may occasionally be upward and balanced by a mean downward advection (see for instance [16]. Here the main balance is between a downward (toward deeper z_r) diffusion (and advection) and an upward (toward shallower z_r) forcing flux where forcing flux can be defined as follows:

$$\frac{\partial\ \text{forcing flux}}{\partial z_r} = \frac{\partial}{\partial z_r}\left(\int_{V(z_r)} \text{forcing}\mathrm{d}V\right) \quad (15)$$

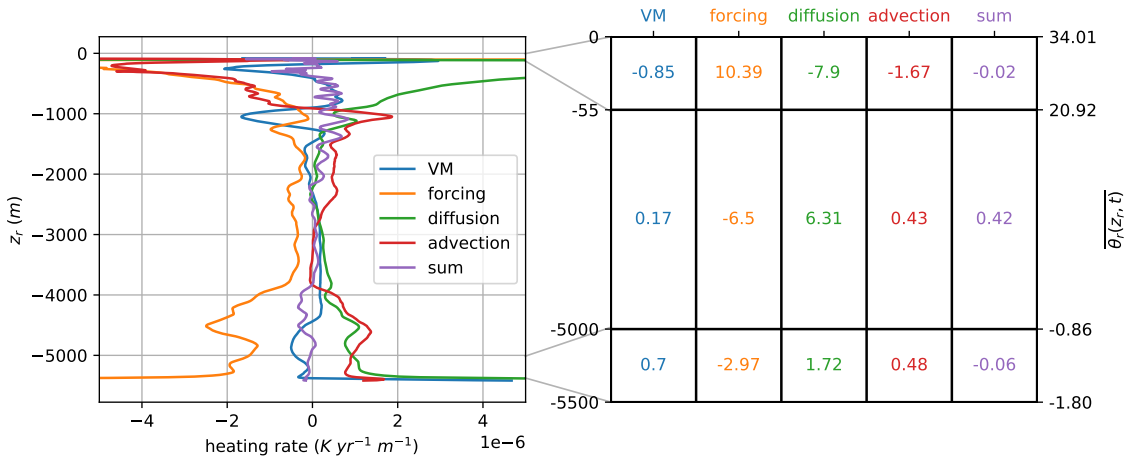

Fig. 5 Left panel: divergence of the volume integrated tendency terms as a function of reference depth (z_r) (derivative with respect to z_r of terms shown on figure 3). Right panel: vertical integral of the left panel quantities over 3 different ranges of z_r. The 3 ranges are: $[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$ (third row), $[-5000\,\mathrm{m}, -55\,\mathrm{m}]$ (second row) and $[-55\,\mathrm{m}, 0\,\mathrm{m}]$ (first row). Unit is 10^{-3} K yr^{-1}. The corresponding time mean temperature (over the CR) is shown for each reference depth range on the right.

Figure 5 shows $\frac{\partial}{\partial z_r}\left(\int_{V(z_r)} \text{term}\right)$ with "term" replaced by either forcing, advection, diffusion or VM. Positive values act to increase the temperature while negative values decreases the temperature. To facilitate the interpretation of these noisy terms we also show their integration over three ranges of reference depth: $[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$, $[-5000\,\mathrm{m}, -55\,\mathrm{m}]$ and $[-55\,\mathrm{m}, 0\,\mathrm{m}]$. In $[-55\,\mathrm{m}, 0\,\mathrm{m}]$, advection, diffusion and VM all act to decrease the temperature and are balanced by the forcing. In $[-5000\,\mathrm{m}, -55\,\mathrm{m}]$ the forcing decreases the temperature and is almost entirely balanced by the diffusion. The sum of all terms in this range of z_r is slightly positive because of the CR drift. In $[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$ the forcing is negative and balances the sum of the remaining terms. The magnitude of the diffusion and forcing values in the shallowest and deepest ranges are respectively about two times and one third of that found in the intermediate range although both correspond to a much smaller volume (50 m and 500 m of reference depth vs almost 5000 m). This emphasize the importance of these two ranges of reference depth for the ocean heat budget.

3.2.3 Advective term

In this section we show that the non-zero advection appearing in the above budget (Eq. 12) is approximately balanced by sub-monthly diffusion. To understand what term balance this sub-monthly advection term, we have run the control run of HiGEM on a year with daily means outputs and repeated the calculation that led to figure 3. The comparison between results from monthly means and daily means outputs for the same year is on figure 6.

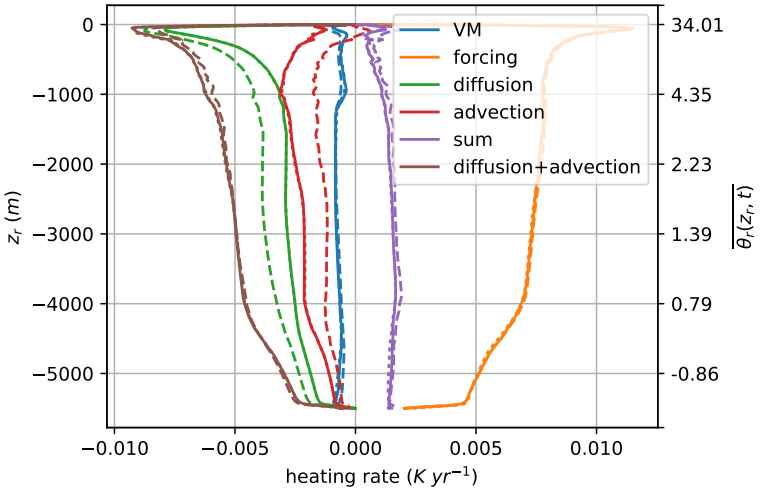

Fig. 6 Comparison of the monthly outputs (plain line) average vs daily outputs (dash line) average of the volume integral of the heat tendencies (in $\mathrm{K\,yr^{-1}}$) respectively as a fonction of $\overline{z_r}$ (calculated on monthly means θ) and z_r (calculated on daily mean of θ). VM is in blue, forcing in orange, diffusion in green, advection in red, sum of all terms in purple and the sum of diffusion and advection is in brown. Also shown (dotted, indistinguishable from plain) are the monthly means of the heat tendencies integrated on the volumes defined by z_r instead of $\overline{z_r}$, but both are very close and thus indistinguishable on the figure. Note that the difference with figure 3 for the monthly outputs is due to the time mean performed on only one year compared to the 50 years of figure 3. The time mean of the temperature as a function of z_r is shown on the right.

We first notice that the differences between the two time resolution for the forcing and the VM are very small. Secondly, as expected, the advection term is closer to zero when daily means are used rather than monthly means. Recall that it cannot be zero because of the problem in HiGEM with the free surface boundary condition. The diffusive flux is larger with daily means outputs than with monthly means so that the sum of the diffusion and of the advection remains approximately constant between the two outputs frequency (Fig. 6). To understand this, we first write Eq. (2) using monthly mean and anomalies:

$$\frac{\partial\overline{\theta}}{\partial t}+\frac{\partial\theta'}{\partial t}+\mathbf{v}\cdot\nabla\theta+\mathrm{Ae}=\overline{D}+D'+\overline{VM}+VM'+\overline{F}+F', \tag{16}$$

where Ae represents the effect of the imperfect formulation of the free surface boundary condition in HiGEM ([16]), and D the diffusion. We integrate it over the volume, $V(z_r)$, calculated from the daily outputs and average the result:

$$\underbrace{\overline{\overline{\int_{V(z_r)} \frac{\partial \overline{\theta}}{\partial t} \mathrm{d}V}}}_{\approx \overline{\overline{\int_{V(\overline{z_r})} \frac{\partial \overline{\theta_r}}{\partial t} \mathrm{d}V}}} + \underbrace{\overline{\overline{\int_{V(z_r)} \frac{\partial \theta'}{\partial t} \mathrm{d}V}}}_{\approx 0} + \underbrace{\overline{\overline{\int_{V(z_r)} \mathrm{Ae} \mathrm{d}V}}}_{\overline{\overline{\int_{V(\overline{z_r})} \mathrm{Ae} \mathrm{d}V}}}$$
$$= \underbrace{\overline{\int_{V(z_r)} \overline{\mathrm{D}} \mathrm{d}V}}_{\approx \overline{\overline{\int_{V(\overline{z_r})} \overline{D} \mathrm{d}V}}} + \overline{\overline{\int_{V(z_r)} \mathrm{D}' \mathrm{d}V}} + \underbrace{\overline{\overline{\int_{V(z_r)} \overline{\mathrm{V}}\overline{\mathrm{M}} \mathrm{d}V}}}_{\overline{\overline{\int_{V(\overline{z_r})} \overline{VM} \mathrm{d}V}}} + \underbrace{\overline{\overline{\int_{V(z_r)} \mathrm{V}\mathrm{M}' \mathrm{d}V}}}_{\approx 0} + \underbrace{\overline{\overline{\int_{V(z_r)} \overline{\mathrm{F}} \mathrm{d}V}}}_{\approx \overline{\overline{\int_{V(\overline{z_r})} \overline{F} \mathrm{d}V}}} + \underbrace{\overline{\overline{\int_{V(z_r)} F' \mathrm{d}V}}}_{\approx 0}. \tag{17}$$

where $\overline{\overline{(.)}}$ is used to indicate a time average over the 50 years of the CR. Figure 6 shows that the time mean of the volume integral of a monthly mean term is very similar when calculated on daily outputs $V(z_r)$ or monthly outputs $V(\overline{z_r})$, and that the volume integral of $\frac{\partial \theta'}{\partial t}$, VM' and F' are negligible, giving all equalities added in equation (17). Comparing Eq. (9) and (17) we deduce that:

$$\overline{\overline{\int_{\overline{z_r}=\text{const.}} \overline{\theta' \mathbf{v}'} \cdot \overline{\mathbf{n}} \mathrm{d}S}} \approx \overline{\overline{\int_{V(z_r)} \mathrm{D}' \mathrm{d}V}} \tag{18}$$

which shows that the residual advection appearing when volume integrating with monthly means is approximately equal to the higher frequency diffusion of temperature.

To sum up, we showed in this section that part (the other part is associated with Ae) of the volume integrated advection is associated with the sub-monthly diffusion.

3.2.4 Effective diffusivity

In what follows, we return to the analysis of the monthly means. The above results motivates us to include the non-vanishing advection term as part of our definition of effective diffusivity. The effective diffusivities associated with sub-monthly diffusion via the advection and associated to the monthly mean diffusion are calculated using the two following formulas:

$$K_{\text{eff}}(z_r) = K_{\text{eff}}^{diff}(z_r) + K_{\text{eff}}^{adv}(z_r) \tag{19}$$

with:

$$K_{\text{eff}}^{diff}(z_r) = -\overline{\overline{\frac{1}{A(z_r)\frac{\partial \theta_r}{\partial z_r}} \int_{V(z_r)} \text{diffusion} \mathrm{d}V}}, \tag{20}$$

$$K_{\text{eff}}^{adv}(z_r) = -\overline{\overline{\frac{1}{A(z_r)\frac{\partial \theta_r}{\partial z_r}} \int_{V(z_r)} \text{advection} \mathrm{d}V}}, \tag{21}$$

where the double overline denotes here the time mean over the 50 years of the CR. K_{eff} is shown on the left panel of figure 7 and is seen to increase with depth

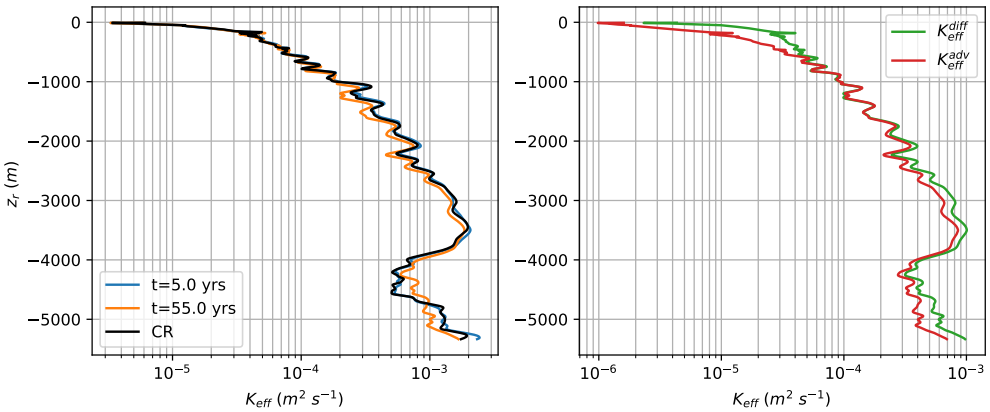

Fig. 7 Left panel: 10-years time average of the effective diffusivity centered at year 5 (orange) and 55 (blue) of the 2×CO_2 run. Right panel: comparison of the 10 years averaged $K_{\rm eff}^{diff}$ (green) and $K_{\rm eff}^{adv}$ (red) centered at year 55 of the 2×CO_2 run. $K_{\rm eff}^{diff}$ is associated with monthly mean diffusion and $K_{\rm eff}^{adv}$ with sub-monthly diffusion because sub-monthly advection is approximately balanced by sub-monthly diffusion, as explained in section 3.2.3.

from values around $1\times10^{-6}\,\mathrm{m^2\,s^{-1}}$ for $z_r=0$ to $2\times10^{-3}\,\mathrm{m^2\,s^{-1}}$ at $-3500\,$m. It then decreases to $5\times10^{-4}\,\mathrm{m^2\,s^{-1}}$ at approximately $-4500\,$m and increases again to $2\times10^{-3}\,\mathrm{m^2\,s^{-1}}$ for the deepest z_r. Note that these values of the diathermal diffusive coefficient are at least one order of magnitude larger than the values $0(10^{-5}\mathrm{m^2.s^{-1}})$ commonly observed in the thermocline. This is mainly because the temperature gradient is generally not parallel to the neutral direction so that part of the large isoneutral mixing occurs in the diathermal direction. Warm waters associated with very shallow reference depth ($>-55\,$m) have an effective diffusivity smaller than $10^{-5}\,\mathrm{m^2\,s^{-1}}$ down to $10^{-6}\,\mathrm{m^2\,s^{-1}}$. This is partly explained by the large temperature gradient (i.e. $\frac{\partial\theta_r}{\partial z_r}$) found at these reference depths as can be seen on figure 2.

In the following section, we study the heat balance under a warming climate using a HiGEM run where the atmospheric concentration of CO_2 is doubled.

3.3 2×CO_2 run

3.3.1 Time evolution of the reference temperature

The left panel of figure 8 shows the time evolution of the isotherms' reference depth in the 70 years of the 2×CO_2 run. All isotherms are seen to progressively deepen with time as the ocean is getting warmer. The net warming at the end of the 70 year period, defined as the difference between the temperatures at the end and beginning of the period, is depicted in the right panel. This shows that the largest increase ($\approx 2.5\,^\circ$C) occurs at shallow reference depth i.e. at high temperature. The temperatures between reference depths of $-4000\,$m and $-2000\,$m remains almost

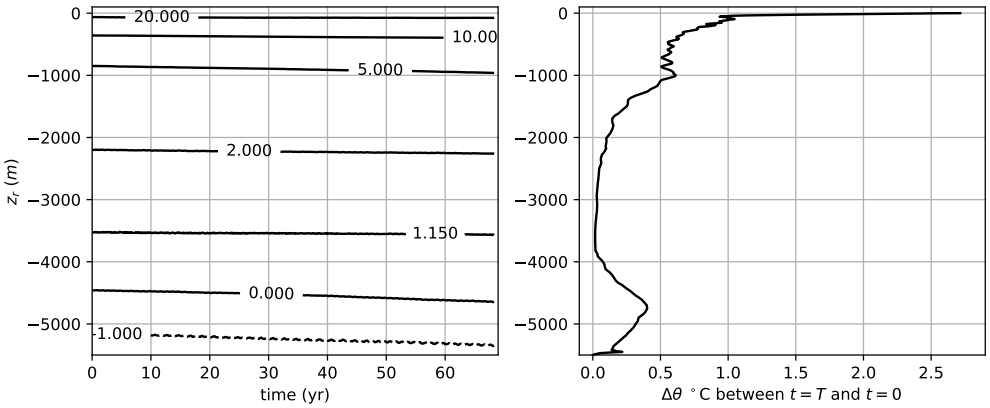

Fig. 8 Left: contours of $\theta_r(z_r, t)$ as a function of time (years) and reference depth (z_r) in the $2\times CO_2$ run. Right: temperature difference (in °C) between the end and the beginning of the $2\times CO_2$ run as a function of z_r.

constant while the temperature between between -5500 m and -4000 m i.e. the coldest waters, increases significantly (≈ 0.4 °C).

3.3.2 Effective diffusivity

The time mean between years 50 and 60 of the $2\times CO_2$ run of the effective diffusivities associated with diffusion (Eq. (20)) and advection (Eq. (21)) are shown on the right panel of figure 7. Despite the differences between the volume integral of the temperature tendencies due to advection and diffusion (see figure 3), K_{eff}^{diff} and K_{eff}^{adv} have similar variation because their reference depth dependence is mainly controlled by the variation of reference temperature gradient $(\overline{\overline{\frac{\partial\theta_r}{\partial z_r}}})^{-1}$ (not shown). K_{eff}^{diff} and K_{eff}^{adv} have very similar magnitude except between -500 m and 0 m, where K_{eff}^{adv} is almost one order of magnitude smaller than K_{eff}^{diff}. The total effective diffusivity during the $2\times CO_2$ run remains nearly constant in the upper 1000 m, increases in the range 3500 m to 4750 m, and decreases everywhere else.

3.3.3 Volume average of the tendencies and heat flux convergence

Figure 9 shows the difference between the 70 years time mean of the volume average of the temperature tendencies in the $2\times CO_2$ and in the CR as a function of reference depth. The temperature increase found at all reference depths, as shown on figure (8), can mainly be attributed to the increase in forcing found in $2\times CO_2$. The volume integral of the forcing in $2\times CO_2$ is indeed much larger than that of the CR, with a difference close to $4 \times 10^{-3}\,\mathrm{K\,yr^{-1}}$.

The heat flux convergences are studied below to understand the time evolution of the temperature at each reference depth. Following equation (10) the heat flux convergence at reference depth z_r are obtained by calculating the z_r derivative of

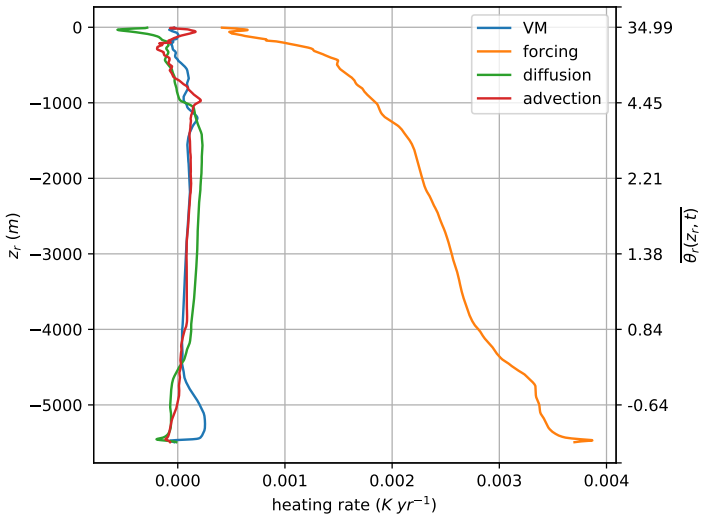

Fig. 9 Volume average of the difference between the $2\times CO_2$ and Control Run of the temperature tendencies due to VM (blue), forcing (orange), diffusion (green) and advection (red) as a function of reference depth (z_r). The time mean of the temperature in the reference space over the $2\times CO_2$ run is shown on the right.

the volume integral of the tendencies. Then, we substract the CR heat flux convergence from the $2\times CO_2$ heat flux convergence to understand what term drives the increase in temperature as shown on figure 8. As for figure 5, the convergence terms are noisy and we thus integrate the results on 5 different ranges of reference level to facilitates the interpretation. This ranges are: $[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$, $[-5000\,\mathrm{m}, -4000\,\mathrm{m}]$, $[-4000\,\mathrm{m}, -2000\,\mathrm{m}]$, $[-2000\,\mathrm{m}, -55\,\mathrm{m}]$ and $[-55\,\mathrm{m}, 0\,\mathrm{m}]$ and are chosen to represent the vertical variation of the convergence terms. Right panel of figure 10 shows that, as expected, the sum of the four processes (advection, diffusion, VM and forcing) is always positive. The difference between the forcing of the $2\times$ CO_2 and the CR is positive for all ranges of reference depth while the diffusion is negative everywhere except in $[-2000\,\mathrm{m}, -55\,\mathrm{m}]$. The heat flux from the atmosphere to the ocean thus increases at shallow reference depths and low latitudes (see figure 4) whereas the ocean loss of heat to the atmosphere that occurs at deeper reference depth and at mid and high latiutdes is reduced. The diffusion intensity increases for low reference depths in the $2\times CO_2$ run: a larger amount of heat is diffused toward low temperatures than in the CR resulting in a cooling of the ocean for $z_r > -55\,\mathrm{m}$ and in a warming for z_r between $-2000\,\mathrm{m}$ and $-55\,\mathrm{m}$. As shown on figure 7, K_{eff} remains approximately constant in the $2\times$ CO_2 run while the gradient of theta ($\frac{\partial \theta_r}{\partial z_r}$) increases at shallow reference depth (see figure 8). The increase in diffusive flux toward low reference depth is thus explained by the increase in the temperature gradient at low reference depth. The largest value of the sum of all terms is found at shallow reference depth in the $[-2000, -55]$ range ($25.77 \times 10^{-3}\,\mathrm{K\,yr^{-1}}$) where diffusion ($7.12 \times 10^{-3}\,\mathrm{K\,yr^{-1}}$) and forcing ($17.34 \times 10^{-3}\,\mathrm{K\,yr^{-1}}$) act together to increase the temperature. This range represents 72% of the difference between $2\times CO_2$ and CR total heating rate.

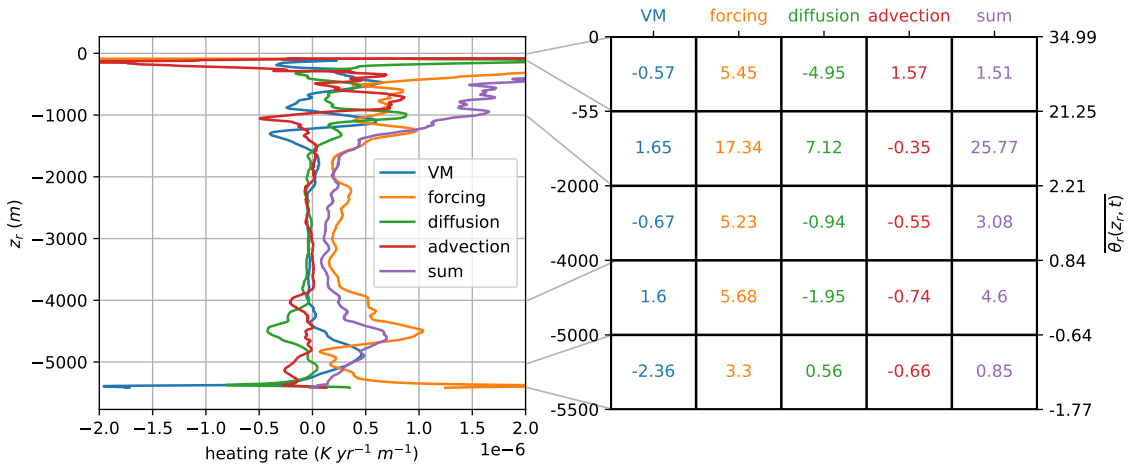

Fig. 10 Left panel: difference between the $2\times CO_2$ and Control runs of the temperature tendencies terms divergence (positive values mean that temperature in the $2\times CO_2$ increases compared to CR) as a function of reference depth (z_r). VM is in blue, forcing in orange, diffusion in green, advection in red and the sum of all terms in purple. Right panel: vertical integral of the left panel terms over 5 ranges of reference level. The 5 ranges are: $[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$ (fifth row), $[-5000\,\mathrm{m}, -4000\,\mathrm{m}]$ (fourth row), $[-4000\,\mathrm{m}, -2000\,\mathrm{m}]$ (third row), $[-2000\,\mathrm{m}, -55\,\mathrm{m}]$ (second row) and $[-55\,\mathrm{m}, 0\,\mathrm{m}]$ (first row). Unit is $10^{-3}\mathrm{K\,yr^{-1}}$. The corresponding time mean temperature (over the $2\times CO_2$) is shown for each reference depth range on the right.

At the surface the $[-2000\,\mathrm{m}, -55\,\mathrm{m}]$ range is approximately located between 60°S and 30°S in the Southern hemisphere, and between 30°N and 60°N in the Northern hemisphere (see figure 4). In the deepest range ($[-5500\,\mathrm{m}, -5000\,\mathrm{m}]$) the warming effect of the forcing (due to a weaker heat transfer to the atmosphere) is almost entirely balanced by the reduced VM found in the $2\times$ CO_2 compared to the CR.

To sum up, the increase of temperature in the $2\times$ CO_2 is mainly attributed to the increase of heat flux from the atmosphere to the ocean at low reference depth and to the decrease of the heat flux from the ocean to the atmosphere at deeper reference depth, particularly between $[-2000\,\mathrm{m}, -55\,\mathrm{m}]$. Diffusion acts to decrease the temperature at shallow reference depth ($[-55\,\mathrm{m}, 0\,\mathrm{m}]$) and to increase the temperature in the range below ($[-2000\,\mathrm{m}, -55\,\mathrm{m}]$). This is explained by the intensification of the diffusive flux in the upper reference depths of the $2\times$ CO_2, associated with the increased temperature gradient, that results in a higher transfer of heat from high to low temperatures.

4 Conclusions

Following up on [14], this paper explores an alternative way to develop a physical calibration of the classical VAD for the purposes of representing the ocean heat balance and ocean heat uptake in SCMs. VADs based on an Eulerian horizontal average — which represent the majority of existing SCM VAD — are not well suited to the development of physical calibrations, because one of the key terms controlling their time evolution involves the correlation between w' and θ', defined

as departures from a horizontal mean (see appendix A). Physically, we know that θ' and w' must be controlled both by surface buoyancy fluxes, wind forcing and interior mixing processes, meaning that we should expect their correlation to be partly advective, partly diffusive. How to perform such a separation in practice is not understood. However, interpreting such a term as purely diffusive reveals that it generally tends to act anti-diffusively, which in [14] was found to be responsible for occasionally making the effective diffusivity negative, thus explaining the behaviour seen in studies such as [16]. This effect is in theory suppressed when the average is performed along constant θ surfaces. Indeed, by definition of this average, deviation from isotherms are zero (i.e. $\theta' = 0$) and thus cannot influence the evolution equation. The temperature time evolution is then only due to diathermal diffusion (toward low temperature), to surface heat fluxes and to parameterized convection/mixed layer dynamics, while the temperature advection plays no role.

Using this new framework, we studied the heat balance and heat uptake in two HiGEM runs, one where the CO_2 concentration is set to 345 ppmv reflecting conditions in the 1980s (the control run), and one where the CO_2 concentration is doubled. In the CR, the balance is mainly between the downward (i.e. toward colder temperatures) sum of advection and diffusion and the upward forcing. Heat flux convergences (i.e. the reference depth derivative of the total heat fluxes) show that above a reference depth of approximately $-55\,\mathrm{m}$, the diffusion, advection, VM cool the ocean while the forcing heats the ocean and compensate almost completely this cooling. Below this reference depth (i.e. for most of the ocean volume), the main equilibrium is between the sum of advection and diffusion that heats the ocean and the forcing that cools the ocean. We showed that the advection term is in theory zero in this framework but that in practice it is true only when the outputs frequency is large enough (smaller than a month here). However, we showed that the advection term that appears when the monthly means are used can conveniently be linked to the higher frequency diffusion and that the total diffusion that would have been obtained with high frequency outputs is very close to the sum of the diffusion and advection obtained from monthly average. Further work need to be done to understand if this result can be generalized to other models or to the ocean. In HiGEM, sub-monthly forcing is negligible compared to sub-monthly advection and sub-monthly diffusion, however the question whether this is true in models with a more realistic representation of sub-monthly forcing remains to be addressed.

The effective diffusivity coefficient (K_{eff}) of the diathermal diffusion has then be calculated using the sum of tendencies from advection and diffusion. K_{eff} is around $1 \times 10^{-3}\,\mathrm{m\,s^{-2}}$ between $-5500\,\mathrm{m}$ to $-2000\,\mathrm{m}$ and increases from $1 \times 10^{-6}\,\mathrm{m\,s^{-2}}$ to $1 \times 10^{-3}\,\mathrm{m\,s^{-2}}$ between 0 to $-2000\,\mathrm{m}$. The fact that these values are at least one order of magnitude larger than the prescribed vertical diffusion in HiGEM indicates that isopycnal mixing for reference depth below $-2000\,\mathrm{m}$ plays an important role in the heat budget. In the $2\times CO_2$ run temperature increases at every reference depth, particularly at shallow reference depth. This temperature increase is attributed to the increase in forcing at all reference depths: the heat flux from the atmosphere to the ocean increases (low reference depths, high temperatures) while the heat flux from the ocean to the atmosphere (mid and deep reference depths, low temperatures) decreases. The diffusive flux increases for reference depths between $-2000\,\mathrm{m}$ and $0\,\mathrm{m}$ which results in a cooling above $-55\,\mathrm{m}$ and a warming between $-2000\,\mathrm{m}$ and $-55\,\mathrm{m}$. It contrasts with the results obtained with the horizontal

average as in [16] where the warming of the horizontally averaged temperature is attributed to the vertical mixing in the top 1000 m and from increased downwelling below −1000 m. Vertical mixing plays no significant role in heat uptake using θ-coordinates except for the deepest reference depths (between −5500 m and −5000 m) where it almost balances the warming due to the reduction of heat transfer to the atmosphere. Similarly, downwelling (i.e. advection) is in theory zero as explained above and can in practise be linked to the diffusion so that it does not play any significant role in the θ-averaged model.

Even if the evolution equation for the θ-averaged model is simpler than for the horizontal average model, some work remains to be done before the θ averaged model can be used to calibrate SCM. The two main point that we identify are the time evolution of the effective diffusivity coefficient K_{eff} and of the reference level at the surface (i.e. $z_r(x, y, z = 0, t)$) under a warming climate. Indeed, once the time evolution of the reference level z_r at the surface is known, the surface temperature can be deduced from the $\theta(z_r, t)$. The heat exchanges between the atmosphere and the ocean could then be deduced from the knowledge of this temperature. Understanding this two points would help to predict the evolution of the diffusivity and of the forcing in different θ classes.

A Revisiting the horizontally-averaged interpretation of the VAD in the light of the θ-based framework

In this appendix, we propose an alternative construction of the horizontally-averaged temperature previously considered by [14] aimed at making it more easily comparable to the θ-based framework considered in this paper.

To obtain an equation for the horizontal average of the temperature, θ is first decomposed into its horizontal average plus departure from it:

$$\theta = \langle\theta\rangle\,(z,t) + \theta'(x,y,z,t). \tag{22}$$

where $< . >$ is the horizontal average and (x, y, z, t) are respectively the longitude, latitude, depth and time coordinates. The temperature departure from the horizontal average θ' is obtained using $\theta' = \theta - \langle\theta\rangle$ and satisfies: $< \theta' >= 0$. Substituting θ in Eq. (2) by (22) and volume integrating between the ocean surface and some depth z yields:

$$\int_{V(z)} \frac{\partial\theta'}{\partial t}\mathrm{d}V + \frac{\partial}{\partial t}\int_{V(z)} \langle\theta\rangle\,\mathrm{d}V + \langle\theta\rangle \int_{x,y} w\mathrm{d}x\mathrm{d}y + \int_{x,y} w\theta'\mathrm{d}x\mathrm{d}y = \int_{x,y} \mathbf{K}\nabla\,\langle\theta\rangle\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y + \int_{x,y} \mathbf{K}\nabla\theta'\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y + \int_{V(z)} \mathrm{VM}\mathrm{d}V + \int_{V(z)} \mathrm{Q_{net}}\mathrm{d}V \tag{23}$$

where $V(z)$ is the volume between the surface and depth z and $\int_{x,y}\mathrm{d}x\mathrm{d}y$ the horizontal integral on the whole ocean surface. The non-divergence of the velocity field has been used to transform the volume integral of the advection into a surface integral at depth z. The term involving the surface heating is zero in a statistical steady-state but positive in global warming experiments.

By definition of θ', the first term on the left hand side is zero and the third term of the lhs is also zero because of volume conservation, Eq. (23) then becomes:

$$\frac{\partial}{\partial t}\int_{V(z)}\langle\theta\rangle\,\mathrm{d}V+\int_{x,y}w\theta'\mathrm{d}x\mathrm{d}y= \int_{x,y}\mathbf{K}\nabla\langle\theta\rangle\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y+\int_{x,y}\mathbf{K}\nabla\theta'\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y+\int_{V(z)}\mathrm{VM}\mathrm{d}V+\int_{V(z)}\mathrm{Q_{net}}\mathrm{d}V. \quad (24)$$

To make the calculation concrete, we assume that $\mathbf{K}=K_i(\mathbf{I}-\mathbf{dd}^T)+K_d\mathbf{dd}^T$ is a rotated diffusion tensor ([23]), with $\mathbf{I}$ the identity tensor, $\mathbf{d}$ the unit normal vector pointing in the dianeutral direction, K_i and K_d the turbulent isoneutral and dianeutral mixing coefficients respectively. As a result, the projection of the diffusive flux of the horizontally average θ i.e.: $\mathbf{K}\nabla\langle\theta\rangle\cdot\mathbf{k}$ may be written as:

$$\mathbf{K}\nabla\langle\theta\rangle\cdot\mathbf{k}=\frac{\partial\langle\theta\rangle}{\partial z}\mathbf{K}\mathbf{k}\cdot\mathbf{k}=\frac{\partial\langle\theta\rangle}{\partial z}\left(K_i\left(\mathbf{k}-(\mathbf{k}\cdot\mathbf{d})\mathbf{d}\right)+K_d(\mathbf{k}\cdot\mathbf{d})\mathbf{d}\right)\cdot\mathbf{k} = \frac{\partial\langle\theta\rangle}{\partial z}\underbrace{\left(K_i\sin^2(\mathbf{k},\mathbf{d})+K_d\cos^2(\mathbf{k},\mathbf{d})\right)}_{=K_{\mathrm{eff}}^{\mathrm{loc}}} \quad (25)$$

$K_{\mathrm{eff}}^{\mathrm{loc}}$ is the local effective mixing coefficient and is always a positive quantity. Let us now take the derivative of (24) with respect to z and divide the result by $A(z)$, the depth-dependent ocean area at depth z, which yields the following equation for $\langle\theta\rangle$:

$$\frac{\partial\langle\theta\rangle}{\partial t}+\frac{1}{A(z)}\frac{\partial}{\partial z}\int_{x,y}w\theta'\mathrm{d}x\mathrm{d}y=\frac{1}{A(z)}\frac{\partial}{\partial z}\left(K_{\mathrm{eff}}A(z)\frac{\partial\langle\theta\rangle}{\partial z}\right)+\frac{1}{A(z)}\frac{\partial}{\partial z}\left(\int_{x,y}\mathbf{K}\nabla\theta'\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y\right) + \frac{1}{A(z)}\frac{\partial}{\partial z}\left(\int_{V(z)}\mathrm{VM}\mathrm{d}V\right)+\frac{1}{A(z)}\frac{\partial}{\partial z}\left(\int_{V(z)}\mathrm{Q_{net}}\mathrm{d}V\right) \quad (26)$$

where

$$K_{\mathrm{eff}}=\frac{1}{A(z)}\int_{x,y}K_{\mathrm{eff}}^{\mathrm{loc}}\,\mathrm{d}x\mathrm{d}y \quad (27)$$

is the horizontally-averaged $K_{\mathrm{eff}}^{\mathrm{loc}}$, and where we have used:

$$\frac{\partial}{\partial z}\frac{\partial}{\partial t}\int_{V(z)}\langle\theta\rangle\,\mathrm{d}V=\frac{\partial}{\partial z}\int_z^0 A(z^v)\frac{\partial\langle\theta\rangle}{\partial t}(z^v)\mathrm{d}z^v=-A(z)\frac{\partial\langle\theta\rangle}{\partial t} \quad (28)$$

Note that the derivation of (26) could have been done using directly an horizontal average on (2) instead of a volume integral followed by a vertical derivation. This method is preferred here to emphasize the similarity with the average along θ surfaces as will become clear in the following section.

In [14], the VAD equation is written in the form:

$$\frac{\partial\langle\theta\rangle}{\partial t}+w^*\frac{\partial\langle\theta\rangle}{\partial z}=\frac{\partial}{\partial z}\left(k^*\frac{\partial\langle\theta\rangle}{\partial z}\right)+Q, \quad (29)$$

where k^* is an effective vertical diffusive coefficient, w^* an effective vertical velocity and Q a source term. Comparing this equation with Eq. (26) and identifying like for like terms suggests the following associations:

$$\frac{1}{A(z)}\frac{\partial}{\partial z}\left(K_{\mathrm{eff}}A(z)\frac{\partial\langle\theta\rangle}{\partial z}\right)+\frac{1}{A(z)}\frac{\partial}{\partial z}\left(\int_{x,y}\mathbf{K}\nabla\theta'\cdot\mathbf{k}\mathrm{d}x\mathrm{d}y\right)\quad\leftrightarrow\quad\frac{\partial}{\partial z}\left(k^*\frac{\partial\langle\theta\rangle}{\partial z}\right) \quad (30)$$

$$\frac{1}{A(z)}\left(\frac{\partial}{\partial z}\left(\int_{V(z)}\mathrm{VM}\mathrm{d}V\right)+\frac{\partial}{\partial z}\left(\int_{x,y}w\theta'\mathrm{d}x\mathrm{d}y\right)\right)\quad\leftrightarrow\quad w^*\frac{\partial\langle\theta\rangle}{\partial z} \quad (31)$$

$$\frac{1}{A(z)}\frac{\partial}{\partial z}\left(\int_{V(z)} \mathrm{Q_{net}} \mathrm{d}V\right) \quad \leftrightarrow \quad Q \tag{32}$$

The diffusive terms in Eq. (26) are identified with the diffusive part of Eq. (29), the advective and vertical mixing terms with the advective part of (29) and the forcing with the forcing term of (29). This identification can be used to obtain a physical calibration of the VAD equation w^* and k^* as in [14]. As noted by [14], this choice is not unique and the VM term could also be attributed to the diffusive part of the VAD for instance. The main novelty here is that the divergence of the diffusive flux is explicitly split into a part involving the vertical gradient of $<\theta>$ which is always downgradient and a part involving the departure of θ from the horizontal average which can a priori be negative or positive. The fact that the horizontal mean diffusive heat flux can occasionally transport heat upwards, as first showed by [6], means that the latter term may occasionally counteract the effect of the former term.

Due to volume conservation, the horizontally-averaged vertical velocity must vanish at all depths (i.e. $<w>=0$), and therefore cannot contribute to the effective advection w^* of $<\theta>$. As shown by Eq. (31) w^* is rather associated with the advection of θ' through horizontal surfaces and to VM. Negative values around $w^* \approx -0.5 \times 10^{-7}\,\mathrm{m\,s^{-1}}$ are found in [14] for the resolved and eddy parametrized advection: advection of θ' transports heat downward.
To sum up, averaging the heat budget along depth levels introduces two new non-negligible terms involving horizontal temperature anomalies θ', which complicates the description of the horizontal mean temperature $<\theta>$. This motivates us to seek an approach that avoid the introduction of such anomalies.

From equation (26), the only way to remove the terms involving θ' in (31) and (30) is to perform the average along temperature surfaces instead of horizontally. Indeed, θ' is then zero by definition.

B General framework accounting for freshwater fluxes

In this appendix, we consider the more general case where the effect of the free surface and of freshwater fluxes are not neglected in Eq. (5). Note that the full derivation of this equation is given in more details in [10], for the sake of conciseness, we only give the main result here.

At the top, the ocean is bounded by a free surface of equation $z = \eta(x, y, t)$. With the effect of freshwater fluxes and of the free surface included, equation (5) of section 2 becomes:

$$\int_{V(z_r)} \frac{\partial \theta_r}{\partial t}\mathrm{d}V + \int_{S(z_r)} (\theta_s - \theta_r)(E - P + R)dS + \int_{V(z_r)} \nabla \cdot \overline{\mathbf{v}'\theta'}\mathrm{d}V = \int_{z_r=\mathrm{const.}} \mathbf{K}\nabla\theta \cdot \mathbf{n}\mathrm{d}S + \int_{V(z_r)} \mathrm{VM}\mathrm{d}V + \int_{V(z_r)} \mathrm{Q_{net}}\mathrm{d}V, \tag{33}$$

where P, E and R represents respectively precipitation, evaporation and river runoff, θ_s is the temperature at the surface, $S(z_r)$ is the outcropping surface corresponding to the surface $z_r = \mathrm{const.}$ and where we have use the fact that $\nabla \cdot \mathbf{v} = 0$ imposes at each time:

$$\int_{z_r=\mathrm{const.}} \mathbf{v} \cdot \mathbf{n}\,\mathrm{d}S = \int_{S(z_r)} (E - P + R)\mathrm{d}S. \tag{34}$$

C Volume integral calculation

In this appendix we describe how the volume integral of the different terms in the temperature tendencies budget (i.e. Eq. 13) is calculated.
For each monthly mean output, each grid cell is vertically divided into 10 smaller volumes at the center of which θ is linearly interpolated. The value of the term we want to integrate is divided by 10 and attributed to each of the 10 sub-volumes corresponding to each grid cell. This procedure allows one to have a better resolution and to conserve the volume integral of the term. We experimentally found that using no subdivision leads to a larger amount of noise when calculating the z_r derivative of the integrated term and that a larger number of vertical subdivision (> 10) has no significant effect on the results. Then, for each time step t, the minimum θ_{min} and maximum θ_{max} of θ are obtained. An array $\theta_{vec} = [\theta_{min}, \theta_{min} + \Delta\theta, \theta_{min} + 2\Delta\theta, ..., \theta_{max}]$ is then constructed with $\Delta\theta = \frac{\theta_{max}-\theta_{min}}{N}$ and $N = 1000$. For each temperature θ^i_{vec} (with $i \in 1, ..., 1000$) in this array we sum the volume times the integrated term of all parcels with θ satisfying $\theta > \theta^i_{vec}$:

$$\sum_{j=1}^{N_i} \text{term}_j \times \Delta \text{V}_\text{j} \tag{35}$$

where ΔV_j is the parcel's volume and N_i the number of parcels with θ satisfying $\theta > \theta^i_{vec}$. Using continuous notation for clarity we now have:

$$\int_{V(\theta,t)} \text{term dV} \tag{36}$$

And we make use of the previously calculated $z_r(\theta, t)$ to obtain:

$$\int_{V(z_r)} \text{term dV} \tag{37}$$

Acknowledgements This work was supported by the OUTCROP NERC grant NE/R010536/1. A. Hochet is supported by EU Marie Curie IF grant number 749924. We thank the two anonymous reviewers who helped improved a first version of this manuscript.

References

1. Brierley, C.M., Collins, M., Thorpe, A.J.: The impact of perturbations to ocean-model parameters on climate and climate change in a coupled model. Climate dynamics **34**(2-3), 325–343 (2010)
2. Dukowicz, J.K., Smith, R.D.: Implicit free-surface method for the bryan-cox-semtner ocean model. Journal of Geophysical Research: Oceans **99**(C4), 7991–8014 (1994)
3. Exarchou, E., Kuhlbrodt, T., Gregory, J.M., Smith, R.S.: Ocean heat uptake processes: A model intercomparison. Journal of Climate **28**(2), 887–908 (2015)
4. Ferrari, R., Ferreira, D.: What processes drive the ocean heat transport? Ocean Modelling **38**(3), 171–186 (2011)
5. Gent, P.R., Mcwilliams, J.C.: Isopycnal mixing in ocean circulation models. Journal of Physical Oceanography **20**(1), 150–155 (1990)
6. Gregory, J.M.: Vertical heat transports in the ocean and their effect on time-dependent climate change. Climate Dynamics **16**(7), 501–515 (2000)
7. Gregory, J.M., Mitchell, J.F.: The climate response to co2 of the hadley centre coupled aogcm with and without flux adjustment. Geophysical Research Letters **24**(15), 1943–1946 (1997)
8. Griffies, S.M.: The Gent–McWilliams Skew Flux. JOURNAL OF PHYSICAL OCEANOGRAPHY **28**, 11 (1998)
9. Hieronymus, M., Nilsson, J., Nycander, J.: Water mass transformation in salinity temperature space. Journal of Physical Oceanography **44**(9), 2547–2568 (2014). DOI 10.1175/jpo-d-13-0257.1. URL ¡Go to ISI¿://WOS:000341334300016. 3
10. Hochet, A., Tailleux, R.: Comments on "diathermal heat transport in a global ocean model". Journal of Physical Oceanography **49**(8), 2189–2193 (2019)

11. Hochet, A., Tailleux, R., Ferreira, D., Kuhlbrodt, T.: Isoneutral control of effective diapycnal mixing in numerical ocean models with neutral rotated diffusion tensors. Ocean Science **15**(1), 21–32 (2019)
12. Holmes, R.M., Zika, J.D., England, M.H.: Diathermal Heat Transport in a Global Ocean Model. Journal of Physical Oceanography **49**(1), 141–161 (2018). DOI 10.1175/JPO-D-18-0098.1. URL https://doi.org/10.1175/JPO-D-18-0098.1
13. Huang, B., Stone, P.H., Sokolov, A.P., Kamenkovich, I.V.: The deep-ocean heat uptake in transient climate change. Journal of climate **16**(9), 1352–1363 (2003)
14. Huber, M., Tailleux, R., Ferreira, D., Kuhlbrodt, T., Gregory, J.: A traceable physical calibration of the vertical advection-diffusion equation for modeling ocean heat uptake. Geophysical Research Letters **42**(7), 2333–2341 (2015). DOI 10.1002/2015GL063383
15. Kuhlbrodt, T., Gregory, J.: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. Geophysical Research Letters **39**(18) (2012)
16. Kuhlbrodt, T., Gregory, J., Shaffrey, L.: A process-based analysis of ocean heat uptake in an aogcm with an eddy-permitting ocean component. Climate Dynamics **45**(11-12), 3205–3226 (2015)
17. Levitus, S., Antonov, J.I., Boyer, T.P., Baranova, O.K., Garcia, H.E., Locarnini, R.A., Mishonov, A.V., Reagan, J., Seidov, D., Yarosh, E.S., et al.: World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. Geophysical Research Letters **39**(10) (2012)
18. Marshall, D.P., Zanna, L.: A conceptual model of ocean heat uptake under climate change. Journal of Climate **27**(22), 8444–8465 (2014). DOI 10.1175/jcli-d-13-00344.1
19. Meinshausen, M., Meinshausen, N., Hare, W., Raper, S.C., Frieler, K., Knutti, R., Frame, D.J., Allen, M.R.: Greenhouse-gas emission targets for limiting global warming to 2 c. Nature **458**(7242), 1158–1162 (2009)
20. Meinshausen, M., Raper, S., Wigley, T.: Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, magicc6–part 1: Model description and calibration. Atmospheric Chemistry and Physics **11**(4), 1417–1456 (2011)
21. Raper, S., Gregory, J.M., Osborn, T.: Use of an upwelling-diffusion energy balance climate model to simulate and diagnose a/ogcm results. Climate Dynamics **17**(8), 601–613 (2001)
22. Raper, S.C., Gregory, J.M., Stouffer, R.J.: The role of climate sensitivity and ocean heat uptake on aogcm transient temperature response. Journal of Climate **15**(1), 124–130 (2002)
23. Redi, M.H.: Oceanic isopycnal mixing by coordinate rotation. Journal of Physical Oceanography **12**(10), 1154–1158 (1982)
24. Shaffrey, L.C., Stevens, I., Norton, W.A., Roberts, M.J., Vidale, P.L., Harle, J.D., Jrrar, A., Stevens, D.P., Woodage, M.J., Demory, M.E., Donners, J., Clark, D.B., Clayton, A., Cole, J.W., Wilson, S.S., Connolley, W.M., Davies, T.M., Iwi, A.M., Johns, T.C., King, J.C., New, A.L., Slingo, J.M., Slingo, A., Steenman-Clark, L., Martin, G.M.: U.K. HiGEM: The New U.K. High-Resolution Global Environment Model: Model Description and Basic Evaluation. Journal of Climate **22**(8), 1861–1896 (2009). DOI 10.1175/2008JCLI2508.1
25. Winters, K.B., D'Asaro, E.A.: Diascalar flux and the rate of fluid mixing. Journal of Fluid Mechanics **317**, 179–193 (1996)
26. Winters, K.B., Lombard, P.N., Riley, J.J., D'Asaro, E.A.: Available potential energy and mixing in density-stratified fluids. Journal of Fluid Mechanics **289**, 115–128 (1995). DOI 10.1017/S002211209500125X
27. Wolfe, C., Cessi, P., McClean, J., Maltrud, M.: Vertical heat transport in eddying ocean models. Geophysical Research Letters **35**(23) (2008)
28. Wyrtki, K.: The thermohaline circulation in relation to the general circulation in the oceans. Deep Sea Research (1953) **8**(1), 39–64 (1961)