Ocean Science

High frequency fluctuations in the heat content of an ocean general circulation model

A. M. Huerta-Casas and D. J. Webb

National Oceanography Centre, Southampton SO14 3ZH, UK

Correspondence to: A. M. Huerta-Casas (adriana.huerta-casas@awi.de)

Abstract. The transport and storage of heat by the ocean is of crucial importance because of its effect on ocean dynamics and its impact on the atmosphere, climate and climate change. Unfortunately, limits to the amount of data that can be collected and stored mean that many experimental and modelling studies of the heat budget have to make use of mean datasets where the effects of short term fluctuations are lost.

In this paper we investigate the magnitude of the resulting errors by making use of data from OCCAM, a high resolution global ocean model. The model carries out a proper heat balance every time step so any imbalances that are found in the analysis must result from the use of mean fields.

The study concentrates on two areas of the ocean affecting the El Nino. The first is the region of tropical instability waves north of the Equator. The second is in the upwelling region along the Equator.

It is shown that in both cases, processes with a period of less than five days can have a significant impact on the heat budget. Thus, analyses using data averaged over five days or more are likely to have significant errors. It is also shown that if a series of instantaneous values is available, reasonable estimates can be made of the size of the errors. In model studies, such values are available in the form of the datasets used to restart the model. In experimental studies they may be in the form of individual unaveraged observations.

1 Introduction

Quantifying the heat budget of the ocean, including its ability to store and transport heat, is a key element in understanding the mechanisms that drive the ocean circulation and its role in climate and climate change. The ocean has the largest heat capacity of any single component of the climate system and for this reason it is essential that we have an accurate understanding of the global and regional oceanic heat budgets at all time scales.

Over the past 40 yr the ocean has been responsible for major changes in global heat content (Levitus et al., 2001). But despite the developments in global ocean observing system, the problems of quantifying heat budgets and their variability at regional and global scales remains difficult. Errors in such estimates can be large and usually go unreported in the literature (Willis et al., 2004).

A number of papers have focussed on generating accurate estimates of heat content. Willis et al. (2004) combined satellite altimetric height and historically available in-situ temperature data to produce global estimates of upper ocean heat content, thermosteric expansion, and temperature variability over a 10.5 yr period.

Dong et al. (2007a) studied the relative role of the air–sea heat flux and oceanic processes in the heat balance in the Southern Ocean. Their study identified an imbalance in the heat equation, which they attributed partly to the complex feedback processes of the coupling system in these regions. These are not well resolved by existing measurements. They also found that the air–sea heat flux datasets for the Southern Ocean showed large differences during the winter months, most probably due to the sparseness of the basic raw data.

Using in-situ and altimetry data for the western North Atlantic, Dong et al. (2007b) studied how the subsurface thermal structure changes correspond to changes in the upper ocean heat content. The residual term in their heat equation was ignored, as previous studies suggest that it is small in gyre regions. In order to balance the heat budget they used an

inverse method to adjust velocity estimates in each isothermal layer. They also estimate the residual by computing the RMS difference between the sea surface height and heat content derived from a 3-D model study.

Wells et al. (2009) analyzed the upper ocean heat budget of the North Atlantic using Argo profiling floats and NCEP/NCAR and NOC surface flux datasets. In their work they considered that high frequency contributions could make a contribution to the heat advection if there were associated high frequency variations in the averaged temperature between the depth of the upper ocean (300 m) and the sea surface. They believed those variations are only likely near frontal zones associated with mesoscale eddies and these were not resolved in their study. Therefore, the term was ignored.

As a final example of the approaches that have been used, Lee et al. (2004) pointed out the need for caution in identifying the mechanisms controlling the heat content of a region. They explored an alternative scheme for advecting temperature which should make it easier to do this.

Although studies such as these usually justify the use of mean datasets, there is always the possibility that the fluctuations are making important contributions that are missed when using the mean datasets.

The present study is therefore concerned with learning more about the errors that arise when mean datasets are used to estimate the advection and diffusion of heat within the ocean. We do this using the archived data from a run of the high resolution OCCAM ocean general circulation model to study three regions of the Eastern Tropical Pacific. The latter is a region of energetic variability mainly due to the presence of the tropical instability waves (Fig. 1).

The archive data include the 5-day average datasets and the instantaneous datasets, from the end of each 5-day period, which are used to restart the model. The model ensures a proper heat balance every model time step so any errors that arise in the analysis must result from the use of mean datasets.

Section 2 of the paper discusses the heat balance equation and its separation into the mean and fluctuating fields. Section 3 briefly describes the model and the equations it uses. The numerical details, the analysis for a single column and the estimate of the error, or high frequency term, are described in Sect. 4. Section 5 describes the analysis in two larger regions on the eastern Equatorial Pacific. The paper concludes with a brief summary and discussion of the results.

2   Mean field and fluctuations in time

Changes in the ocean's heat content result from the effects of advection and diffusion within the ocean. If Q is the heat content per unit volume of the ocean, then the conservation of heat equation has the form:

$$\frac{\partial Q}{\partial t} = \nabla \cdot \boldsymbol{A} + \nabla \cdot \boldsymbol{D} + S, \tag{1}$$

where the first term on the righthand side represents heat advected by the large scale current field, the second represents the diffusion of heat by small scale oceanic processes and S is the surface heat flux.

If the specific heat of sea water is constant, as is assumed in the ocean model, then this can be also written as:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \nabla \cdot \boldsymbol{u} T + \nabla \cdot \Lambda \nabla T + S, \tag{2}$$

where T is temperature, $\boldsymbol{u}$ is the three dimensional current vector (u, v, w) and Λ the diffusion tensor. In the OCCAM model, Λ is a diagonal tensor, such that

$$\Lambda \nabla T = A_{\mathrm{h}} \nabla_{\mathrm{h}} T + A_{\mathrm{z}} \partial T / \partial z. \tag{3}$$

∇_{h} is the horizontal gradient operator. A_{h} and A_{z} are the horizontal and vertical diffusion coefficients.

It is convenient to split the temperature and velocity fields into mean and fluctuating parts,

$$T(t) = \overline{T}(t) + T'(t), \tag{4}$$

$$\boldsymbol{u}(t) = \overline{\boldsymbol{u}}(t) + \boldsymbol{u}'(t), \tag{5}$$

where the terms with overbars are the mean during the time period τ, i.e. $\bar{T} = \frac{1}{\tau} \int_0^\tau T \mathrm{d}t$ and $\bar{\boldsymbol{u}} = \frac{1}{\tau} \int_0^\tau \boldsymbol{u} \mathrm{d}t$; the prime terms represent fluctuations during this period. The product $\boldsymbol{u} T$ in the advective term of Eq. (2) then becomes

$$\boldsymbol{u} T = (\bar{\boldsymbol{u}} + \boldsymbol{u}')(\bar{T} + T'). \tag{6}$$

Integrating over the period τ,

$$\int_0^\tau \boldsymbol{u} T \mathrm{d}t = \int_0^\tau \bar{\boldsymbol{u}} \bar{T} \mathrm{d}t + \int_0^\tau \bar{\boldsymbol{u}} T' \mathrm{d}t \tag{7}$$

$$+ \int_0^\tau \boldsymbol{u}' \bar{T} \mathrm{d}t + \int_0^\tau \boldsymbol{u}' T' \mathrm{d}t.$$

But $\bar{T}^t \int_0^\tau \boldsymbol{u}' \mathrm{d}t = 0$ and $\bar{\boldsymbol{u}}^t \int_0^\tau T' \mathrm{d}t = 0$. Thus,

$$\int_0^\tau \boldsymbol{u} T \mathrm{d}t = \int_0^\tau \bar{\boldsymbol{u}} \bar{T} \mathrm{d}t + \int_0^\tau \boldsymbol{u}' T' \mathrm{d}t, \tag{8}$$

or

$$\overline{\boldsymbol{u} T} = (\bar{\boldsymbol{u}} \bar{T}) + \overline{(\boldsymbol{u}' T')}. \tag{9}$$

As a result, the total advective heat flux can be split into a contribution only dependent on the mean values of temperature and velocity plus a contribution only dependent on the fluctuations.
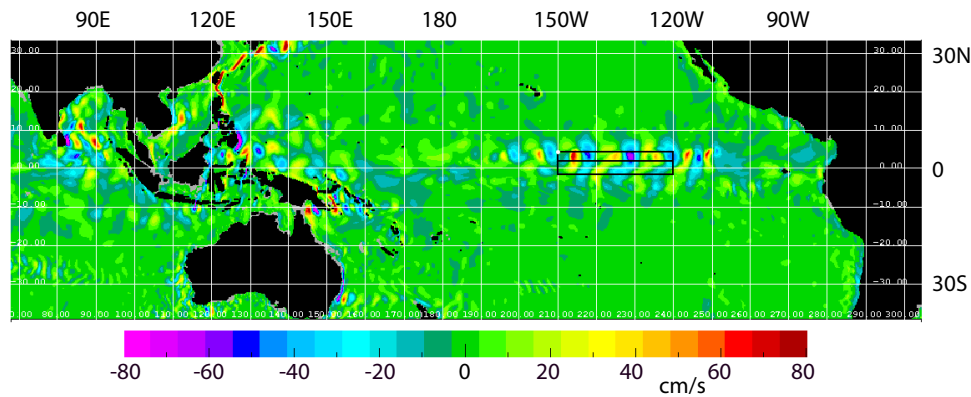
Fig. 1. Meridional velocity (cm s^{-1}) at 30 m depth in September 1993 (from an OCCAM restart dataset). The rectangular areas outline the regions used for the flux study described later in this paper.

If D represents the diffusive term of Eq. (2), then separating the mean and fluctuation parts in the same way,

$$D = \Lambda \nabla (\overline{T} + T^{'}). \tag{10}$$

If the diffusion tensor is constant, as is often assumed, then when the equation is integrated over a period τ, the fluctuation term vanishes so

$$D = \Lambda \nabla \overline{T}. \tag{11}$$

However, in regions where any of the terms of the diffusion tensor vary in time, there will be an extra diffusive contribution from the fluctuations.

3 The model

The Ocean Circulation and Climate Advanced Model (OC-CAM) is based on the Bryan-Cox-Semtner (Cox, 1984; Bryan, 1969; Semtner, 1974) ocean general circulation model. This is a "primitive equation" model (Bryan, 1969) which differs from other primitive equation models in the use of an Arakawa-B grid in the horizontal (Arakawa, 1966) and level surfaces in the vertical.

The version of OCCAM used here has a resolution of a quarter of a degree in both the east–west and north–south directions. It has thirty-six levels in the vertical, with thicknesses varying smoothly from 20 m in the surface layer to 255 m at a depth of 5500 m.

The model includes a number of important developments in both model physics and model numerics[1]. Physically it

[1]OCCAM uses asynchronous message passing, enabling it to be run efficiently on modern multi-processor computers. It also differs from other ocean models in that it is vectorised in the vertical. This arrangement optimises processor usage by limiting main memory transfers when the model is run either on short (~ 100) vector processors or on modern Intel type processors with very high speed memory caches.

replaces the rigid lid of the standard Bryan-Cox-Semtner scheme with a free surface and solves the resulting barotropic equations using a simple tidal model with a time-step of 18 s. It uses the Webb (1995) scheme for the vertical advection of momentum and the split-QUICK scheme (Webb et al., 1998) for the horizontal advection of tracers such as temperature and salinity. It uses the Pacanowski and Philander (1981) scheme for the vertical mixing of tracers.

At the start of the run, the potential temperature and salinity fields were initialized from the Levitus (1982) global annual average database. There was then a spin-up phase lasting 1440 model days (4 yr), during which the model was forced by the monthly averaged ECMWF wind stress climatology, calculated by Siefridt and Barnier (1993) for the years 1986 to 1988 inclusive. The data was corrected so that when linearly interpolated in time, the correct average stress was applied during each model month (Killworth et al., 1991). The surface fluxes of heat and fresh water were calculated so as to relax the surface layer of the model to the Levitus monthly average values (Levitus and Boyer, 1994; Levitus, 1982; Levitus et al., 1994).

Following the spin-up phase, the model was forced using six-hourly winds and surface flux data from the ECMWF analyses. The analysis reported here used archive data for the year 1994 and, as in the model tracer calculation, assumes a sea water density of 1g cm^{-3} and a specific heat of $1 \text{cal g}^{-1} \text{C}^{-1}$.

Model archive datasets were generated every five days. These contain the sea surface height field and the full three dimensional fields of temperature, salinity and horizontal velocity. There are two copies of each field, the first being the average value over the previous five days and the second containing instantaneous values at the end of each five day period. The latter were saved so that the model could be restarted from this point if required. Because of storage limitations the model did not archive any non-linear fields.

3.1 Analysis of the heat equation

The ocean model solves Eq. (1) by first partitioning the ocean into cells whose interfaces lie along lines of constant latitude, longitude and depth (see Fig. 2). Model temperatures are defined at the centre of each cell (or model grid box). The range of depths used in this analysis means that the regions are not immediately affected by surface fluxes but at the same time they may be typical of the upper part of the water column in the Pacific Ocean, therefore $S = 0$. Integrating Eq. (1) over each grid box gives the flux conserving equation,

$$V_n \frac{\partial T_n}{\partial t} = \sum_m (I_{nm} U_{nm} T_{nm} + I_{nm} D_{nm}), \tag{12}$$

where the sum is over the cells m surrounding cell n. V_n is the volume of the cell, T_n its average temperature, and I_{nm} the area of the interface between cells n and m. T_{nm} is the average temperature on the interface and U_{nm} is the average velocity through the interface. The first term on the righthand side is called the advective term, the second is the diffusive term. The latter represents all the small scale dynamical and molecular diffusive processes not represented by the advective term.

The above equation is stepped forward in time by approximating the time derivation term,

$$\frac{\partial T}{\partial t} = \frac{T(t + dt) - T(t - dt)}{2dt}, \tag{13}$$

where dt is the model time step. If the forcing on the right of Eq. (12) is represented by $F(t)$, then this gives the equation

$$T(t + dt) \approx T(t - dt) + 2dt\, F(t). \tag{14}$$

The approximation is correct to order dt^2, and if the time step is small enough, it is both accurate and stable when used for the advective term in Eq. (12). Unfortunately, diffusive terms make it unstable and so in the model it is modified to

$$T(t + dt) = T(t - dt) + 2dt\,(A(t) + D(t - dt)). \tag{15}$$

This equation is stable for both the advection ($A(t)$) and diffusion ($D(t - dt)$) terms.

In the Arakawa B-grid used by the model, shown in Fig. 2b, horizontal velocities are defined at the corners of the cells at the same depth as the central tracer (temperature) grid point. The average velocities through the horizontal interfaces needed for Eq. (12) are thus estimated in the model by averaging the two neighbouring corner values.

As the total flux of water in and out of each grid cell is zero, the ocean being assumed to be incompressible, the horizontal fluxes are then used to calculate the difference in fluxes through the bottom and top of each cell. These differences are summed from the bottom of the ocean, where the vertical flux is zero, to give the velocity flux through each of the levels.
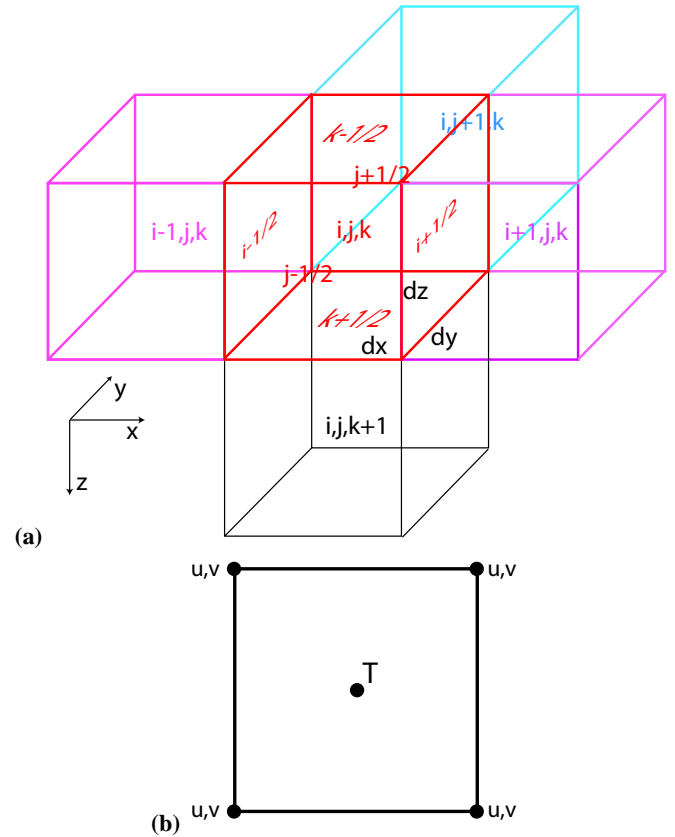
Fig. 2. (a) Model grid box around the point i, j, k; i increases with longitude, j with latitude and k with model depth. The length of each side of the box are given by dx, dy and dz, and interfaces between boxes are represented by $i + \frac{1}{2}$, $i - \frac{1}{2}$, etc. **(b)** Horizontal arrangement of tracer (temperature) and velocity components in the Arakawa B-grid.

The average tracer at the interface can be estimated in a similar way. Thus, replacing n in Eq. (12) by the triplet i, j, k, representing the longitude, latitude and depth indices, the mean temperature on the interface between the i, j, k and $i + 1, j, k$ cells can be written (see Fig. 2)

$$T_{i+\frac{1}{2},j,k} = \frac{(T_{i,j,k} + T_{i+1,j,k})}{2}. \tag{16}$$

This was the approximation used in the earliest ocean models, and like all the other approximations discussed here, is adequate for modelling features with spatial scales much larger than the grid spacing and time scales much longer than a model time step. Unfortunately, in ocean models there are often important features, such as ocean fronts and eddies, which are only a few grid points across. In such cases the approximation involved in Eq. (16) results in such short wavelength features being advected at the wrong speed or even in the wrong direction.

A large number of papers have been written about the problem and possible solutions, one of the most successful solutions being the QUICK scheme devised by Leonard

(1979). Farrow and Stevens (1995) adapted QUICK for use in an ocean model, but to ensure stability they replaced Eq. (15) by a time stepping method that was computationally inefficient. Later, Webb et al. (1998) showed that the QUICK advective operator included a diffusive-like operator and that if this was treated separately as an extra contribution to the main diffusion term, then Eq. (15) could again be used. The same analysis showed that slightly better accuracy could be obtained by changing the term multiplying the second term in the following equation from 1 / 16 to 1 / 12.

The resulting MSQ (modified split-QUICK) scheme is the one used by the OCCAM model for the run analysed here. Dropping the j and k indices in Eq. (16), the MSQ scheme is

$$T_{i+\frac{1}{2}} = \frac{T_{i+1} + T_i}{2} - \frac{T_{i+2} - T_{i+1} - T_i + T_{i-1}}{12} \quad (17)$$
$$+ \gamma \frac{|U|}{U} \frac{T_{i+2} - 3T_{i+1} + 3T_i - T_{i-1}}{16}.$$

The second term on the right improves the advection of features with wavelengths of four model grid points or more. The final term is the additional diffusion introduced by the MSQ scheme. Its effect is to damp out the very short waves which, in the model, can generate large errors.

Note that MSQ was not used in the vertical. This is because, in the presence of realistic surface wind forcing, most of the near-surface vertical velocities are oscillatory due to the internal waves generated. The oscillations tend to cancel out the errors due to the central difference scheme. At the same time, if MSQ is used with a rapidly oscillating current, it produces unwanted extra numerical diffusion.

4 Analysis

The main objective here is to quantify the errors that arise when calculating oceanic heat fluxes using averaged datasets. To do this we apply Eq. (12) to closed volumes of ocean and compare the heat flux, calculated using OCCAM 5-day mean datasets for model year 1994, with the rate of change of heat content calculated from the instantaneous datasets from the beginning and end of the year under study.

Results are presented for three regions in the Equatorial Pacific. The Equatorial Pacific was chosen because of the importance of heat transports there in the development of the El Nino (Halpern, 1987).

4.1 Numerical details

Equation (12) is implemented for each of the regions studied. In terms of the model variables, the advective term on an horizontal interface $i + \frac{1}{2}$ is

$$A_{i+\frac{1}{2},t_i} = \rho I_{i+\frac{1}{2}} C_p u_{i+\frac{1}{2},t_i} \left[\frac{(T_{i+1,t_i} + T_{i,t_i})}{2} \right. \quad (18)$$
$$\left. - \frac{(T_{i+2,t_i} - T_{i+1,t_i} - T_{i,t_i} + T_{i-1,t_i})}{12} \right].$$

The first term on the right is the contribution from the central difference approximation and the second is the extra advective component of the MSQ scheme discussed in Sect. 3.1. Vertical fluxes are calculated similarly but with just the central differences term.

The total heat advection into each region at time t_i is obtained by summing Eq. (19) over all the interfaces surrounding the region. Thus, if $\bar{A}_{\text{tot},t_i}$ is the total heat gain by advection in the region at time t_i,

$$\bar{A}_{\text{tot},t_i} = \sum_m A_{n,m,t_i}, \quad (19)$$

where the sum is over all the interfaces surrounding the region.

The heat gain by diffusion is obtained in a similar way. For each interface $i + \frac{1}{2}$ it is given by

$$D_{i+\frac{1}{2},t_i} = \Delta y \Delta z \left[A_{\text{h}} \left(\frac{T_{i+1,t_i} - T_{i,t_i}}{\Delta x} \right) \right. \quad (20)$$
$$\left. + \gamma \left(\frac{|U|}{U} \left(\frac{T_{i+2,t_i} - 3T_{i+1,t_i} + 3T_{i,t_i} - T_{i-1,t_i}}{16} \right) \right) \right],$$

where A_{h} represents the horizontal diffusion coefficient and γ is a coefficient set to one (Webb et al., 1998). The last term is the diffusive component of the MSQ scheme. The expression for interfaces above and below the study region is similar except that A_{h} is replaced by the vertical diffusion coefficient and the MSQ term is missing.

The total diffusive heat flux into the region is then

$$\bar{D}_{\text{tot},t_i} = \sum_m D_{n,m,t_i}, \quad (21)$$

where again the sum is over all the surrounding interfaces. For each analysis period, Eq. (12) can then be written as,

$$\frac{\partial Q}{\partial t} = \bar{A}_{\text{tot},t_i} + \bar{D}_{\text{tot},t_i}. \quad (22)$$

4.2 A single vertical column

We first analysed a single vertical column of model grid cells. The south-west corner of the column was at point 150° W, 4.5° N and it extended from a depth of 41.2 m to 146.79 m (see Fig. 1).

The heat content is given by

$$Q_{t_i} = \sum_n C_p \rho \, \mathrm{d}x_n \mathrm{d}y_n \mathrm{d}z_n T_{n,t_i}, \qquad (23)$$

where the sum is over the model cells n within the analysis region. If the averaging time is $\triangle t$, then the mean rate of change of heat content is

$$\frac{\triangle Q_{n,t_i}}{\triangle t_i} = C_p \rho \, \mathrm{d}x \mathrm{d}y \mathrm{d}z \left(\frac{T_{n,t_{i+1}} - T_{n,t_i}}{\triangle t_i} \right). \qquad (24)$$

Figure 3 shows the results for the single column for the year studied. The volume of the box is $8.1 \times 10^{10} \, \mathrm{m}^3$, so a flux of 1 TW is equivalent to a temperature change of approximately $0.25 \,^\circ\mathrm{C} \, \mathrm{day}^{-1}$. The results show that rates greater than 0.1 TW are common and that there is a large amount of short term variability. The differences between the actual heat flux and that calculated from the 5-day mean datasets are small, but they are systematic and may occasionally exceed 10 % of the total flux.

The second part of the figure shows the heat fluxes integrated over a year. In this case a change in heat content of 1 EJ is equivalent to an average temperature change of 2.9 °C.

The figure confirms that the errors are systematic, with the mean fields overestimating the total heat flux into the region during the year by approximately 1 EJ. The flux errors appear to be largest early in the year and from July onwards when tropical instability waves are most developed.

4.3 An independent estimate of the advective error

Figure 3 shows that flux calculations using the mean datasets produces errors, but it does not show whether they are primarily due to errors in calculating the advective or diffusive fluxes. However, it is possible to make an estimate of the advective error from the differences between the mean and instantaneous datasets.

Starting from Eq. (9), for a single interface,

$$\overline{\boldsymbol{u}T} = (\bar{u}\bar{T}) + \overline{(u'T')}; \qquad (25)$$

if $\bar{T}$ and $\bar{u}$ are values from the 5-day mean datasets, then the errors in the advective flux calculation correspond to the term $u'T'$.

If the fields $u(t)$ and $T(t)$ oscillate many times during each averaging interval, then – as discussed in Appendix A – $u'T'$ will have the same statistical properties as $F_e(m)$, where

$$F_e(m) = (u_e(m) - \bar{u}(m))(T_e(m) - \bar{T}(m)), \qquad (26)$$

and $u_e(m)$ and $T_e(m)$ are the instantaneous values at the end of each interval m available from the restart datasets. The mean value and variance of $F_e(m)$ over a period of time should then be a good approximation to the mean and variance of $u'T'$ during the same period.

In the other limit when the fields $u(t)$ and $T(t)$ are slowly varying, so that they can be approximated as linear functions
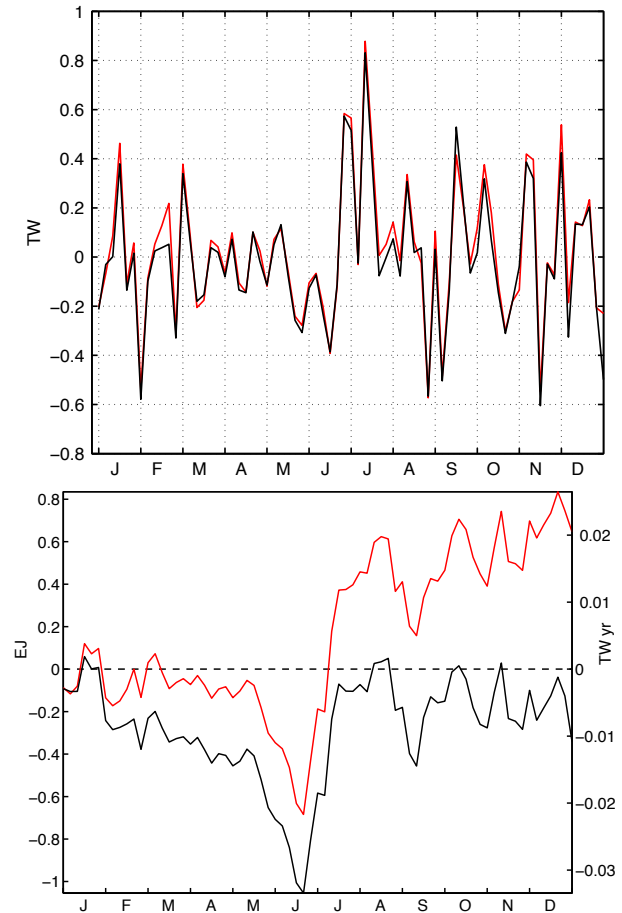
Fig. 3. Results for the single column of model cells at 150° W, 4.5° N during model year 1994. **(a)** Sum of the advection and diffusion terms (red – calculated from the 5-day mean datasets) and rate of change of heat content (black – calculated from the instantaneous datasets). Heat flux in TW (10^{12} W). **(b)** Cumulative integrals of both terms. Heat content in EJ (10^{18} J).

of time, u' and T' have their maximum amplitudes at the beginning and end of each 5-day period. As shown in Appendix A, Eq. (26) then overestimates $u'T'$ by a factor of three.

Using this independent estimate, the total error in using mean datasets to calculate the advective fluxes into a region of the model ocean is

$$F_{e,\mathrm{tot}} = \sum_{m-1}^{M} (A_m S_m (u_e(m) - \bar{u}(m))(T_e(m) - \bar{T}(m))), \quad (27)$$

where the sum is over the interfaces surrounding the region being considered. A_m is the area of the m'-th interface and S_m equals +1 if u represents inflow and -1 if it represents outflow. The estimate of the advective error will be referred to as the high frequency term.

Figure 4 compares this estimate of the advective error for the column considered earlier, with the actual error
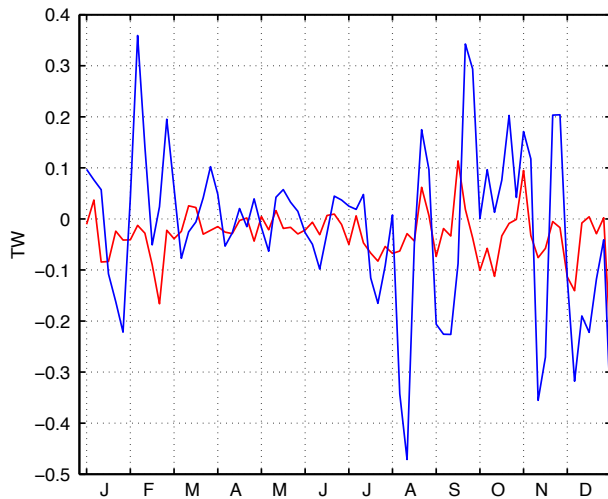
Fig. 4. Estimates of the heat flux error (in TW) due to neglecting short term fluctuations (blue – Eq. A6) compared with, in red, the difference between the actual change in model heat content and that estimated from the 5-day mean datasets.

calculated from the difference between the two fluxes in Fig. 3a. At most times the advective estimate is of the same order but slightly larger than the actual error. This confirms that the error estimate F_e, is comparable in magnitude with the error arising from using mean datasets. The average ratio between the two flux magnitudes appears to be larger than the figure of three given above, but overall the result shows that the instantaneous datasets can provide useful extra information on the transport of heat within the ocean.

5 The eastern Equatorial Pacific

We extend this analysis to two regions in the Equatorial Pacific. The first, the most northerly of the two, lies in the eastern Equatorial Pacific, between 150° W and 120° W and between 2.25° N and 4.75° N (see Fig. 1). It covers the same range of depths as the single column. This region is characterised by strong horizontal shears due to the North Equatorial Counter Current and the Equatorial Undercurrent. It is also an area of energetic variability due to the presence of tropical instability waves.

As variability is important in this analysis, the model surface currents on the northern and southern boundaries of this region were compared with satellite derived estimates (Bonjean and Lagerloef, 2002). Variances in this area of ocean depend primarily on the distance from the Equator as zonal gradients are small. On the northern boundary, at 4.75° N and 135° W, the root mean square difference between the surface velocity in the model and the mean velocity during 1994 was 7.7 cm s^{-1} and this compared with 5.9 cm s^{-1} estimated from the satellite observations. On the southern boundary, at

2.25° N and the same longitude, the corresponding values are 16.6 cm s^{-1} and 16.5 cm s^{-1}.

In the model, tropical instability waves grew rapidly each northern summer (i.e. in July) and following this the largest values of velocity variance occurred in the autumn. The satellite derived estimates differed in that the variance was largest in winter and the seasonal differences were smaller. However, overall the comparison indicates that the model results presented below should be reasonably representative of the real ocean.

5.1 The advection term

Figure 5a shows the rate of change of heat content within the northern region and the heat flux due to advection and diffusion, calculated from the averaged datasets. Both functions show a large amount of variability so Fig. 5b shows the values integrated over time[2].

In the first figure the two flux curves follow each other closely for most of the year. However, from July onward the integral calculated from the mean datasets produces a much larger increase in the total heat content than is actually observed. By the end of the year, the resulting error is as great as the annual variation in the heat content of the region.

Figure 6 compares the error in the heat flux calculation with the estimate of the advective error calculated using Eq. (27). Both terms have the same overall behaviour and it is noticeable that during July, when the errors are largest, they show a remarkable amount of agreement.

In the OCCAM model, this region north of the Equator is one where tropical instability waves are observed. These grow rapidly each June and have their greatest amplitudes between July and January. It seems likely that the waves are advecting heat out of the region and that this process is not being captured by the mean datasets.

The second region studied has the same longitude range but spans the Equator, running from 1° S to 2.25° N. Advection in the region is dominated by the westward flowing Equatorial Undercurrent, but there is also strong equatorial upwelling and meridional convergences and divergences.

Figure 7 shows the rate of change of heat content and the estimated heat flux due to advection and diffusion using the mean datasets. For this region a flux of 0.1 PW is equivalent to a temperature change of 0.016 °C day^{-1} or 83 W m^{-2} through its upper surface. The integrated values are shown in Fig. 7b, a difference of 1 ZJ being equivalent to a temperature change of 1.9 °C.

The heat flux calculated from the 5-day mean datasets is generally in good agreement with the full model calculation but this time the mean datasets tend to slightly underestimate the fluxes.

---

[2]In terms of other practical units, a flux of 0.1 PW is equivalent to a mean rate of temperature change within the region of 0.021 °C day^{-1} or a flux of 108 W m^{-2} though its upper surface. Similarly 1 ZJ is equivalent to a mean temperature change of 2.4 °C.
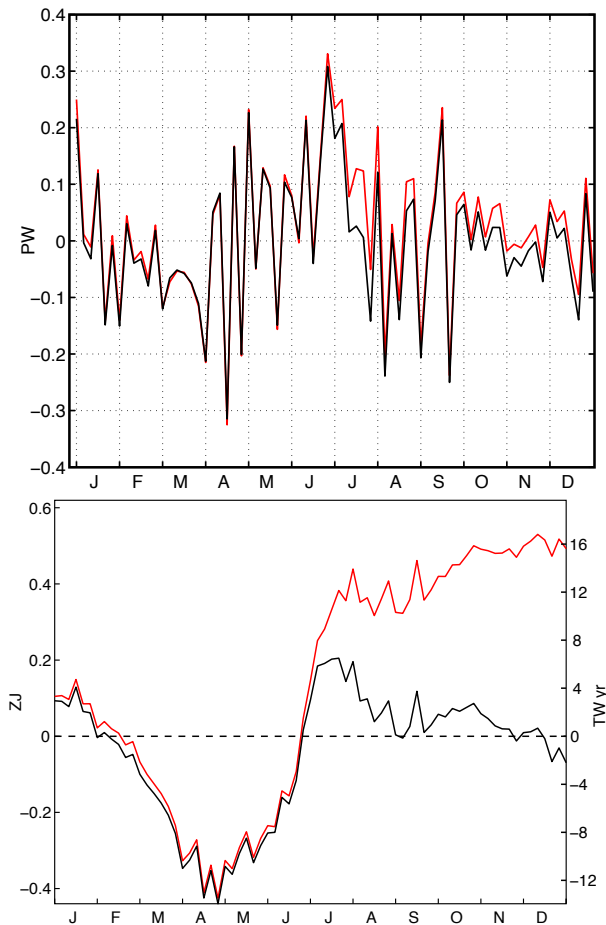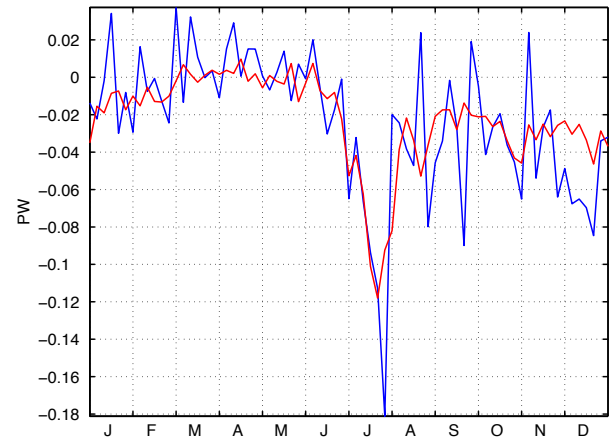
Fig. 5. Results for the northern region during model year 1994:
(a) rate of change of heat content (black – calculated from the in-
stantaneous datasets) and sum of the advection and diffusion terms
(red – calculated from the 5-day mean datasets). Heat flux in PW
(10^{15} W). **(b)** Integrals over time of both terms. Heat content in ZJ
(10^{21} J), also shown as TW years.

This differences shows up more clearly when integrated
over a year (Fig. 7b). The differences are less than they were
for the northern region, but whereas before the mean datasets
were overestimating the heat gain, here they overestimate the
heat loss. Thus the missing high frequency terms must be
warming the region.

Figure 8 shows that the warming may be due to the high
frequency advective term. However, especially from Septem-
ber onwards, the estimate of the missing advective contribu-
tion is much larger than the actual error. Given the earlier
discussion on the different sources of error, this implies that
in this region we also need to consider a possible high fre-
quency contribution to vertical mixing.

To summarise the results of this section, in the eastern
tropical Pacific we find that there are important heat transport
processes which are not captured by 5-day mean datasets. In
the region of the tropical instability waves we find that over

Fig. 6. Results for the northern region: Estimates of the heat flux er-
ror (in PW) due to neglecting short term fluctuations (blue – Eq. A6)
compared with, in red, the difference between the actual change in
model heat content and that estimated from the 5-day mean datasets.

the course of a year the cumulative error can be as large as
the seasonal signal. An estimate of the high frequency ad-
vective terms indicates that these may be responsible for the
discrepancy.

Near the Equator we find that the cumulative error from
using the mean datasets is smaller but that the high frequency
advective terms are likely to over correct the error. Thus, in
this region errors due to fluctuations in the diffusion terms
also need to be considered.

5.2 The diffusion terms

As discussed earlier, the scheme used to represent horizon-
tal diffusion in the ocean is linear and so is not affected by
high-frequency non-linear effects. In the vertical, however,
the model uses the Pacanowski and Philander (1981) scheme
and this is non-linear and so will be affected by the high-
frequency fluctuations.

The Pacanowski and Philander (1981) vertical mixing
scheme depends on the Richardson number, which in OC-
CAM is calculated as

$$Ri_k = g \frac{(z_{k+1} - z_k)}{\rho} \frac{(\rho_k - \rho_{k+1})}{((u_k - u_{k+1})^2 + (v_k - v_{k+1})^2)}, \quad (28)$$

where z_k is the layer depth, ρ is the density, u and v are the
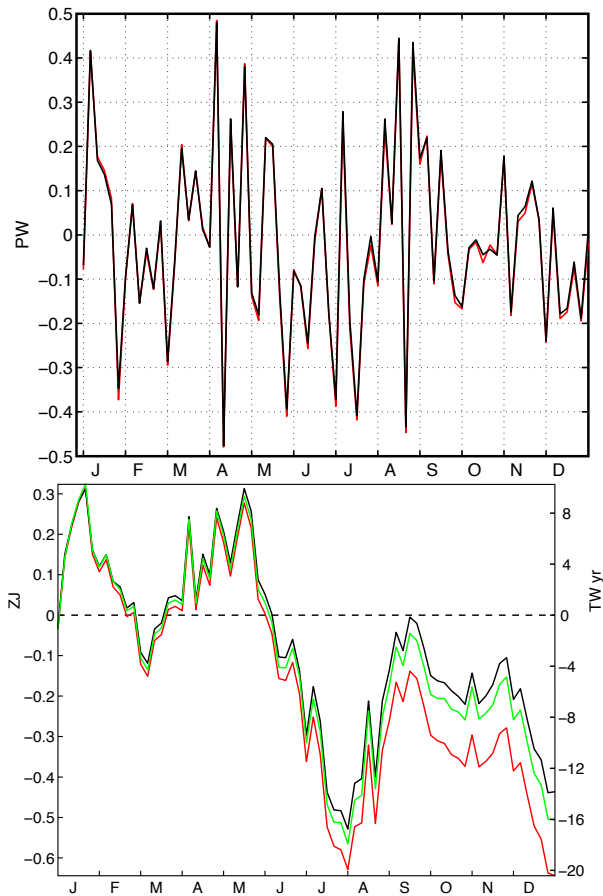zonal and meridional velocities, and g is gravity.

The vertical mixing coefficient is given by

$$A_z = \frac{\nu_0}{(1 + \alpha Ri)^n} + \nu_b, \quad (29)$$

where $\alpha = 5$, $n = 3$, $\nu_0 = 50$ and $\nu_b = 0.5 \, \text{cm}^2 \, \text{s}^{-1}$.

When the Richardson number is large ($Ri > 0.2$), the
flow is stable to shear instabilities and the vertical mixing
coefficient is small ($\sim 0.5 \, \text{cm}^2 \, \text{s}^{-1}$). Conversely, when the
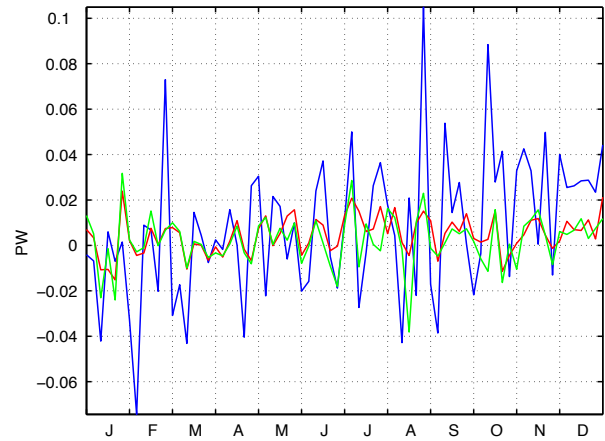
Fig. 7. Results for the equatorial region: **(a)** rate of change of heat content (black – calculated from the instantaneous datasets) and sum of the advection and diffusion terms (red – calculated from the 5-day mean datasets). Heat flux in PW. **(b)** Total heat content (black – from instantaneous datasets) plus cumulative integrals of heat fluxes calculated using mean datasets (red) and instantaneous datasets (green). Heat content in ZJ (10^{21} J), also shown as TW years.

Richardson number is small, shear instabilities develop and the mixing coefficient is large ($\sim 50\,\mathrm{cm^2\,s^{-1}}$).

Over most of the ocean the Richardson number is large, so shear instabilities have little effect. However there exist regions, such as the equatorial undercurrent, where vertical shear is large and the stratification small. The Richardson number is then small and there is strong vertical mixing.

We investigated this effect by calculating the Richardson number at the top of the cross equatorial region for each analysis time step and grid point, during two time periods – January to May and August to December. This was carried out using both the average and the instantaneous datasets. The region had 1920 grid boxes in the top layer and the calculation was repeated for 28 datasets for both periods, giving a total of 53 760 data points. Points with Richardson number bigger than 0.8 have not been plotted.

Fig. 8. Result for the equatorial region: estimate of the heat flux error (in blue) resulting from ignoring the high frequency fluctuations compared with (red), the difference between rate of change of the model heat content and the fluxes calculated using the 5-day mean datasets, and (green) the same but using the instantaneous datasets to calculate Richardson numbers.
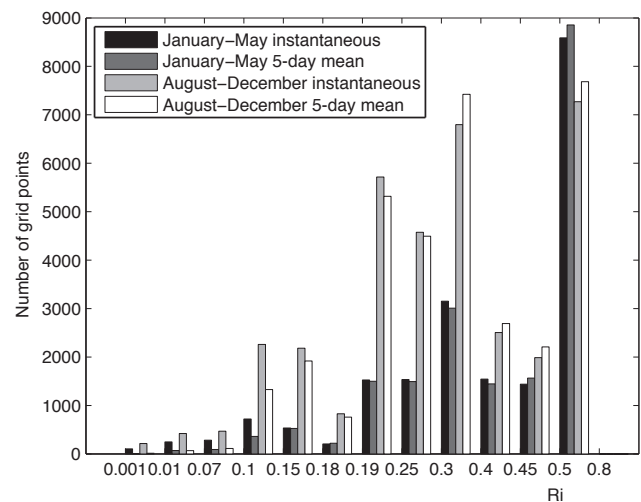
Fig. 9. Number of grid points at 41.20 m (model level three) with values of the Richardson number within the interval shown (see legend).

The results are shown in Fig. 9. At a Richardson number of 0.8, the vertical mixing coefficient has a value of $0.9\,\mathrm{cm^2\,s^{-1}}$. It increases to $1.6\,\mathrm{cm^2\,s^{-1}}$, at a Richardson number of 0.2, to $15\,\mathrm{cm^2\,s^{-1}}$ at 0.1 and $44\,\mathrm{cm^2\,s^{-1}}$ at 0.01.

The figure confirms that large Richardson numbers are the most common, but more importantly it also shows systematic differences between the values calculated from the instantaneous and 5-day mean datasets. Thus, during the first analysis period the number of grid points in the instantaneous datasets with Richardson number less than 0.1 is ten times larger than for the five day datasets. For the second analysis

period there is a similar order of magnitude difference for Richardson numbers of less than 0.07.

In both cases the results imply that the mean datasets are missing many of the strong vertical mixing events. For Richardson numbers larger than 0.15, the instantaneous and mean dataset results tend to have similar numbers of points within each of the plotted bands. Thus, although there is a strong seasonal signal most of it is captured by the 5-day mean datasets.

Figure 7b compares the integrals of the heat fluxes obtained when using either the mean or instantaneous datasets when calculating the Richardson number, with the rate of change of the total heat content of the region being studied. It shows that the instantaneous datasets give the better agreement; presumably, as discussed above, they are better at capturing the periods of strong mixing.

Figure 8 shows the difference between the rate of change of heat content and the total heat flux into the volume using 5-day mean and instantaneous datasets. The figure also shows the estimate of the error or high frequency term. It can be seen that while both curves follow each other most of the time, the use of instantaneous datasets gives stronger fluctuations; at times the latter is closer to the high frequency term than the error produced by the use of the 5-day mean datasets. This is best seen during January and between June and September.

5.3   Annual average fluxes in the north equatorial region

Table 1 shows the annual average heat fluxes through each of the interfaces for the region north of the Equator for the year under study. As before, the advection and diffusion terms are calculated from the 5-day average datasets and the estimate of the high frequency contributions to the advective terms are calculated from the instantaneous restart datasets.

The figures show that the region is warmed by the downward advection of heat and that this is largely balanced by a loss of heat due to meridional advection. A smaller amount of heat is lost to zonal advection. Thus, when using the 5-day mean fields, the net effect of advection is to warm the region at a rate of 29.8 TW[3].

The same mean fields produce a diffusive cooling at a rate of 14.2 TW. This is due primarily to a loss of heat to layers deeper in the water column. The combined effect of advection and diffusion, using the 5-day average values, thus produces a heating at a rate of 15.6 TW. However, over the course of the year the model region actually cooled at a rate of 2.2 TW, which means that the mean fields are generating a spurious heating at a rate of 17.8 TW.

The estimates of the contributions to the high frequency advective term show that the main contribution comes from

_____

[3]Other studies are often concerned with the heat flux through the surface of the ocean. To convert TW to the equivalent flux in $W\,m^{-2}$ through the top surface of each volume studied, multiply by 1.08 for the north equatorial region and 0.83 for the larger equatorial region.

Table 1. Heat fluxes averaged over the year at each interface of the region north of the Equator. The columns show the heat fluxes due to advection and diffusion, both calculated from the 5-day mean datasets, and the estimate F_e of the error in the advective flux due to ignoring high frequency fluctuations. The last column shows the actual mean rate of change of heat content during the period. Positive signs indicate flux into the region. Units are TW (10^{12} W).

interface	advection	diffusion	F_e	$\frac{\partial Q}{\partial t}$
top	417.61	7.66	−26.16	
bottom	83.62	−21.33	1.15	
east	340.28	0.01	0.59	
west	−398.69	−0.01	−0.17	
north	697.51	0.39	12.27	
south	−1110.50	−0.93	−10.84	
Total	29.83	−14.20	−23.15	−2.20

advection of heat through the upper surface of the region. There is also a significant southward heat flux but the two terms balance out so the net effect is small. Overall, the fluctuations in the advective term are estimated to generate a heat flux cooling of 23.2 TW – close to the value needed to explain the discrepancy when using the 5-day mean fields.

5.4   Annual averaged fluxes in the equatorial region

Table 2 shows the corresponding values for the region centred at the Equator. The advection terms, calculated using the 5-day mean datasets, generate a net cooling over the year of 32.54 TW. The meridional flows act to heat the region but this is balanced by large vertical and zonal heat losses.

As with the northern region, diffusion is dominated by the vertical terms, but this time they act to warm the region. The total diffusive flux is 12.14 TW, as a result of which the advection and diffusion using the 5-day datasets gives a heat flux loss of 20.4 TW. Over the year the model region actually cooled at a rate of 13.8 TW, meaning that the mean fields overcooled the region by 6.6 TW.

The estimates of the heat advected by the high frequency fluctuations indicates that they generate a southward flux of heat into the region which is partly compensated by a loss to the surface layers of the ocean. The net heat flux by these terms (8.0 TW) is certainly sufficient to explain the errors when using the 5-day mean datasets.

However, the equatorial region is also one where shear instabilities above the equatorial undercurrent are likely to be important. For this reason Table 2 also includes a column showing the estimate of the diffusion obtained when the instantaneous restart values are used to calculate the Richardson number. This shows an increase in the downward diffusion of heat, the net effect being an increase of 4.6 TW in the diffusive heat flux into the region.

If the high-frequency and instantaneous Richardson number contributions are taken together, their sum (12.2 TW) is

Table 2. As Table 1 but for the region centred at the Equator. In addition the table shows the vertical diffusive flux when using the restart datasets when calculating the Richardson number. Flux in TW.

| interface | advection | diffusion | | F_e | $\frac{\partial Q}{\partial t}$ |
		5-day	restart		
top	−2298.29	29.81	34.17	−2.92	
bottom	429.55	−18.68	−18.69	−0.25	
east	−1002.88	0.01	0.01	0.48	
west	911.50	0.02	0.02	0.08	
north	1110.50	0.93	0.93	10.84	
south	817.08	0.05	0.05	−0.25	
Total	−32.54	12.14	16.50	7.98	−13.84

almost double the amount needed to explain the discrepancy when using the mean datasets. However, if the high frequency term is being overestimated (see Appendix A), which is possible given the long term drift in heat content (see Fig. 7), then its actual value could be nearer 2.7 TW. Together with the Richardson number contribution, this would give a net cooling of 13.3 TW, close to the actual change of 13.8 TW.

6 Discussion and conclusions

The overall purpose of this study was to shed light in the limitations of using 5-day mean datasets in heat budget analyses in the ocean. The results show that processes with a period of less than five days can have a significant impact on the heat budget and that model restart datasets can be used to estimate the size of the resulting errors.

In the region studied north of the Equator, calculation of the advection and diffusion terms using the 5-day mean datasets showed a net heating of 15.6 TW during the year studied, whereas the model, which conserves heat every time step, actually cooled at a rate of 2.2 TW. The discrepancy of 17.8 TW is in reasonable agreement with the value of 23.2 TW, estimated as being the contribution of the advective processes with periods of less than 5 days.

In the equatorial region, the 5-day mean datasets produced a net cooling of 20.4 TW. Temperatures in the model did drop during the year but only at a rate of 13.8 TW, leaving a discrepancy of 6.6 TW. Estimates of the high-frequency contributions indicated that the discrepancy may result from both the extra advection and diffusion terms.

The work of Markus et al. (2007) suggests that daily averaged datasets provide useful estimates of tropical instability wave (TIW) temperature advection in the mixed layer. They also show that the zonal temperature advection due to the TIWs is as important as the meridional one. On the other hand, the study of Hansen and Paul (1984) in a similar re-

gion concluded that the zonal eddy heat flux is small. They used datasets which had been averaged over several days so this may explain why their analysis came to this conclusion.

Overall the results stress the importance of short period fluctuations in the transport of heat within the ocean. Such advective fluctuations are often called eddy heat fluxes and, as concluded by Bryden and Brady (1989), they make a substantial contribution to the heat balances in the Equatorial Pacific Ocean. The model results indicate that in the regions studied the "eddies" are primarily tropical instability waves and that datasets, averaged over periods of only 5 days, are insufficient to fully resolve their effect.

It is possible that such terms have their greatest effect in the tropical regions studied here but, until it is shown otherwise, they need to be treated with care in all regions of the ocean.

Appendix A

Estimating the contribution of short period fluctuations

Assume that the model is run for a period of time to estimate uT. Let the period be made up of M intervals, each involving N model time steps of length dt. Then if $u(i)$ and $T(i)$ are the values of u and T on an interface at time step i, the total heat flux, F_a, through the interfaces is

$$F_a = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{u(i)T(i)}{MN},$$ (A1)

where i equals $(m-1)N + n$.

At the end of the model run the data archive includes the instantaneous values $u_e(m)$ and $T_e(m)$ (equal to $u(Nm)$ and $T(Nm)$) from the last time step of each of the M intervals, plus $\bar{u}(m)$ and $\bar{T}(m)$ the mean values over each of the intervals.

Let $u'(i)$ and $T'(i)$ be defined as the differences between the actual value at the current time step i and the mean value during the current interval m. Then

$$u'(i) = u(i) - \bar{u}(m),$$ (A2)

$$T'(i) = T(i) - \bar{T}(m).$$ (A3)

Then following a similar argument to that used to derive Eq. (9),

$$F_a = \sum_{m=1}^{M} \frac{\bar{u}(m)\,\bar{T}(m)}{M} + \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{u'(i)T'(i)}{MN},$$ (A4)

where i is related to m and n as before.

If heat fluxes are calculated from just the mean values during each interval, then the first term is the resulting estimate of the heat flux. The second term, the contribution due to

short term fluctuations, becomes the error when such fluctuations are neglected. In the following, this term is represented by F_e.

$$F_e = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{u'(i)\, T'(i)}{MN}. \tag{A5}$$

If there are a large number of fluctuations of u and T within each averaging interval and the amplitude of the fluctuations are large compared with changes in the mean values, then the set of archived instantaneous values $u'_e(m)$ and $T'_e(m)$ (equal to $u'(mN)$ and $T'(mN)$), should have similar statistics as the full set of values, $u'(i)$ and $T'(i)$.

If this is so, then a good approximation to F_e is given by

$$E(F_e) = \frac{1}{M} \sum_{m=1}^{M} u'_e(m) T'_e(m), \tag{A6}$$

where

$$u'_e(m) = u_e(m) - \bar{u}(m), \tag{A7}$$

$$T'_e(m) = T_e(m) - \bar{T}(m). \tag{A8}$$

$E(Q_e)$ is easily calculated from the restart data.

Also, if $F_e(m)$ is the contribution from interval m, and ϵ is r.m.s. amplitude of such contributions over the period studied, then

$$\varepsilon = \sqrt{\frac{\sum (u'_e(m) T'_e(m))^2}{n-1}}, \tag{A9}$$

where n is the number of datasets during one year.

If instead u' and T' vary slowly with time, then the product $u'T'$ will tend to be largest near the beginning and end of each averaging interval. In this case Eq. (A6) will overestimate the contribution of fluctuations to the total heat flux.

The extreme limit occurs when u' and T' can be approximated as linear functions during each averaging interval. Then during such an interval,

$$u'(i) = (u_e(m) - \bar{u}(m)) \frac{i - (N+1)/2}{(N-1)/2}, \tag{A10}$$

$$T'(i) = (T_e(m) - \bar{T}(m)) \frac{i - (N+1)/2}{(N-1)/2}. \tag{A11}$$

The error in using u_m and T_m alone during the period is

$$F_e(m) = u'_e(m)\, T'_e(m) \frac{1}{N} \sum_{i=1}^{N} \left(\frac{i - (N+1)/2}{(N-1)/2} \right)^2. \tag{A12}$$

Using standard formula for the sum of integers and their squares, it is straightforward to show that

$$F_e(m) = u'_e(m) T'_e(m) \left(\frac{1}{3} \frac{N-1}{N+1} \right). \tag{A13}$$

When N is large,

$$F_e(m) = \frac{1}{3} u'_e(m) T'_e(m). \tag{A14}$$

Thus a good estimate of the error in the flux arising from the use of mean values is given by

$$E(F_e) = \frac{1}{3M} \sum_{m=1}^{M} u'_e(m) T'_e(m). \tag{A15}$$

This shows that, in the worst case, Eq. (A6) overestimates the error by a factor of three. In general we may expect the factor to lie between 1 and 1/3, although the error in any one of the averaging intervals may be greater or less than that estimated using the above equations.

The error in determining the net flux out of a volume of ocean is given by a sum of terms with the form of Eqs. (A6) and (A15). The magnitude will usually be proportional to the surface area of the region, but will be reduced for small volumes where the velocity and temperatures on different faces are correlated. The average divergence is given by the net flux divided by the volume and so, as would be expected, with similarly shaped regions of different scales the error will usually be smallest at the largest scale.

Edited by: M. Hecht

References

Arakawa, A.: Computational design for long-term numerical integration of the equations of fluid motion: Two dimensional incompressible flow, Part I, J. Comput. Phys., 1, 119–143, 1966.

Bonjean, F. and Lagerloef, G.: Diagnostic Model and Analysis of the Surface Currents in the Tropical Pacific Ocean, J. Phys. Oceanogr., 32, 2938–2954, 2002.

Bryan, K.: A numerical method for the study of the circulation of the world ocean, J. Comput. Phys., 4, 347–376, 1969.

Bryden, H. and Brady, E.: Eddy momentum and heat fluxes and their effects on the circulation of the equatorial Pacific-Ocean, J. Mar. Res., 47, 55–79, 1989.

Cox, M. D.: A primitive equation 3-dimensional model of the ocean, Tech. Rep. 1, GFDL Ocean Group Techincal Report, Princeton Univ., Princeton, NJ, USA, 1984.

Dong, S., Gille, S. T., and Sprintall, J.: An assessment of the Southern Ocean mixed layer heat budget, J. Climate, 20, 4425–4442, doi:10.1175/JCLI4259.1, 2007a.

Dong, S., Hautala, S. L., and Kelly, K. A.: Interannual variations in upper-ocean heat content and heat transport convergence in the western North Atlantic, J. Phys. Oceanogr., 37, 2682–2697, doi:10.1175/2007JPO3645.1, 2007b.

Farrow, D. E. and Stevens, D. P.: A New Tracer Advection Scheme for Bryan and Cox Type Ocean General Circulation Models, J. Phys. Oceanogr., 25, 1731–1741, 1995.

Halpern, D.: Observations of annual and El Nino thermal and flow variations at 0°, 110° W and 0°, 95° W during 1980–1985, J. Geophys. Res.-Oceans, 92, 8197–8212, 1987.

Hansen, D. and Paul, C.: Genesis and Effects of Long Waves in the Equatorial Pacific, J. Geophys. Res., 89, 10431–10440, 1984.

Killworth, P., Stainforth, D., Webb, D., and Paterson, S.: The development of a free surface Bryan-Cox-Semtner ocean model, J. Phys. Oceanogr., 21, 1333–1348, 1991.

Lee, T., Fukumori, I., and Tang, B.: Temperature advection: Internal versus external processes, J. Phys. Oceanogr., 34, 1936–1944, 2004.

Leonard, B.: A stable and accurate convection modelling procedure based on quadratic upstream interpolation, Comput. Method. Appl. M., 19, 59–98, 1979.

Levitus, S.: Climatological Atlas of the World Oceans, NOAA professional Paper, Washington DC, p. 173, 1982.

Levitus, S. and Boyer, T.: World Ocean Altas 1994, Temperature, NOAA Atlas, 4, p. 177, 1994.

Levitus, S., Burgett, R., and Boyer, T.: World Ocean Atlas 1994, Salinity, NOAA Atlas, 3, 99 pp., 1994.

Levitus, S., Antonov, J., Wang, J., Delworth, T., Dixon, K., and Broccoli, A.: Anthropogenic warming of Earth's climate system, Science, 292, 267–270, 2001.

Markus, J., Cronin, M., Kessler, W., and Shea, D.: Observed horizontal temperature advection by tropical instability waves, Geophys. Res. Lett., 34, L09604, doi:10.1029/2007GL029416, 2007.

Pacanowski, R. C. and Philander, S. G.: Parameterization of vertical mixing in numerical models of Tropical Oceans, J. Phys. Oceanogr., 11, 1443–1451, 1981.

Semtner, A.: A general circulation model for the World Ocean., Tech. Rep. 9, Department of Meteorology, University of California, Los Angeles, 1974.

Siefridt, L. and Barnier, B.: Banque de Donees AVISO Vent/flux: Climatologie des Analyses de Surface du CEPMMT, Report No. 91, 025, 43, 1993.

Webb, D. J.: The Vertical Advection of Momentum in Bryan-Cox-Semtner Ocean General Circulation Models, J. Phys. Oceanogr., 25, 3186–3195, 1995.

Webb, D. J., de Cuervas, B. A., and Richmond, C. S.: Improved advection schemes for ocean models, J. Atmos. Ocean. Tech., 1998.

Wells, N. C., Josey, S. A., and Hadfield, R. E.: Towards closure of regional heat budgets in the North Atlantic using Argo floats and surface flux datasets, Ocean Sci., 5, 59–72, doi:10.5194/os-5-59-2009, 2009.

Willis, J., Roemmich, D., and Cornuelle, B.: Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales, J. Geophys. Res.-Oceans, 109, doi:10.1029/2003JC002260, 2004.