Thermodynamic neutral density: a new physically-based materially conserved neutral density variable for studying mixing and tracking water masses in the ocean

Rémi Tailleux *

University of Reading, Reading, United Kingdom

* *Corresponding author address:* Rémi Tailleux, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, RG6 6BB, United Kingdom.

E-mail: R.G.J.Tailleux@reading.ac.uk

ABSTRACT

Thermodynamic neutral density — denoted γ^T — is proposed as a new quasi-neutral and quasi-material density variable superior to Jackett and McDougall (1997) empirical γ^n variable. γ^T is the difference between the potential density of the fluid parcel referenced to the pressure it would have in the reference state of minimum potential energy entering Lorenz theory of available potential energy and a correction for pressure that is empirically chosen here to minimise differences between γ^T and γ^n for the WOCE dataset, but which could be physically-based if desired.

Thermodynamic neutral density possesses the following advantages over empirical neutral density: 1) it is fully justified from first principles and has a precise and rigorous mathematical definition; 2) its physical basis is the same as that used to rigorously quantify turbulent diapycnal mixing by the turbulent mixing community over the past 20 years; 3) it is materially conserved, and therefore more suited to quantifying ocean mixing; 4) it can be computed accurately and efficiently, making it possible in principle to compute it on the fly in numerical ocean models; 5) density inversions are very rare and confined to very weakly stratified regions, making it suitable as a vertical coordinate for use in isopycnal models.

1. Introduction

The problem of how best to construct a density variable suitably corrected for pressure is a longstanding fundamental issue in oceanography whose answer is vital for many key applications ranging from the study of mixing to ocean climate studies. These include but are not limited to: the separation of mixing into "isopycnal" and "diapycnal" components necessary for the construction of rotated diffusion tensors in numerical ocean models (Redi 1982; Griffies 2004), the construction of climatological datasets for temperature and salinity devoid of spurious water masses (Lozier et al. 1994), the construction of inverse models of the ocean circulation (Wunsch 1996), the tracking and analysis of water masses (Montgomery 1938; Walin 1982), the construction of isopycnal models of the ocean based on generalised coordinate system (Griffies et al. 2000; de Szoeke 2000), the study of the residual circulation (Wolfe 2014), and the parameterisation of meso-scale eddy induced mass fluxes (Gent et al. 1995).

Physically, it is generally agreed that a suitable density variable γ should possess the desirable dual thermodynamic and dynamic attributes of defining adiabatic surfaces along which fluid parcels experience no net buoyancy force, e.g., McDougall (1987); de Szoeke and Springer (2000); Huang (2014). The first attribute, which is equivalent to material conservation, poses no difficulty as it can always be enforced by requiring γ to be a function of potential temperature θ and salinity S only. The second attribute — usually referred to as the *neutral property* — is problematic, however, as it can only be satisfied in special circumstances not usually encountered in the ocean. To satisfy exact neutrality, $\nabla\gamma$ would need to be parallel at every point to the local neutral vector $\mathbf{d} = g[\alpha\nabla\theta - \beta\nabla S] = -(g/\rho)[\nabla\rho - c_s^2\nabla p]$, where α and β are the thermal expansion and haline contraction coefficients, g is the acceleration of gravity, c_s^2 is the squared speed of sound, ρ is in-situ density and p is pressure. To understand why the latter property cannot be satisfied in general, it is useful to decompose $\nabla\gamma$ into components parallel and orthogonal to $\mathbf{d}$ as follows:

$$\nabla\gamma = b\left(\nabla\rho - \frac{1}{c_s^2}\nabla p\right) + \mathcal{R} = b\rho(\beta\nabla S - \alpha\nabla\theta) + \mathcal{R} = -\frac{\rho b\mathbf{d}}{g} + \mathcal{R}, \tag{1}$$

where b is an integrating factor, and $\mathcal{R}$ a residual term perpendicular to $\mathbf{d}$. Taking the curl of (1) and multiplying the result by $\mathbf{d}$ gets rid of $\nabla\gamma$ and yields an equation for the residual $\mathcal{R}$, viz.,

$$-\frac{bH}{g} + \rho\mathbf{d}\cdot[\nabla\times\mathcal{R}] = 0, \tag{2}$$

where the term $H = \mathbf{d}\cdot(\nabla\times\mathbf{d})$ is the *helicity* of the neutral vector $\mathbf{d}$, which shows that exact neutrality can only be achieved when $H = 0$, a well known result (McDougall 1987; de Szoeke and Springer 2000; Huang 2014), with Eden and Willebrand (1999) discussing some of the conditions necessary for the helicity to vanish. In practice, achieving $H = 0$ in the ocean would either require the ocean to be at rest — as ρ would then be a function of pressure p alone — or in absence of density-compensated temperature/salinity variations along surfaces $\gamma = \text{const}$, which is equivalent to say that the ocean would then have a well defined temperature/salinity relationship of the form $\theta = \theta(\gamma)$ and $S = S(\gamma)$. In the ocean, however, the existence of density-compensated θ/S anomalies conspire with thermobaricity (the pressure dependence of the thermal expansion coefficient) to make H non-zero and hence forbid the construction of exactly neutral density variables.

If so, what then are the physical principles determining the degree of non-neutrality that γ should have? In particular, should material conservation be retained, or sacrificed to improve neutrality? For lack of clear physical basis about how to address the above questions, most attempts at constructing density variables so far seem to originate in the concept of potential density. Most likely, this is because in absence of salinity or thermobaricity, potential density referenced to any arbitrary fixed reference pressure p_r (referred to as σ_r) would be both materially conserved and neutral. In the ocean, however, the neutrality of potential density σ_r deteriorates proportionally to the the pressure difference $p - p_r$ times the thermobaric parameter (McDougall 1987) as one moves away from p_r. In order to achieve better neutral properties, Lynn and Reid (1968) introduced the concept of patched potential density (PPD), that is, potential density referenced to a piecewise constant reference pressure p_r depending on the depth range considered, with some studies using up to 10 different reference pressures, e.g., Reid (1994). However, while this makes PPD more neutral than σ_r, this is done at the expenses of continuous behaviour and material conservation, which both break down near the depthd at which the reference pressured change discontinuously.

The unsatisfactory discontinuous character of PPD prompted Jackett and McDougall (1997) (JMD97 thereafter) and de Szoeke and Springer (2000) (SS00 thereafter) to propose empirical neutral density γ^n and orthobaric density respectively as continuous analogues of PPD. The way in which each variable can be regarded as an extension of PPD is somewhat subtle, however. With regard to γ^n, its connection to PPD appears to rely on the assumption that any nonzero angle between ∇(PPD) and the local neutral vector $\mathbf{d}$ would ultimately vanish in the asymptotic limit of an infinite number of reference pressures [1], thus prompting JMD97 to define γ^n as the density variable minimising the residual $\mathcal{R}$ in (1) in some sense [2]. In contrast, orthobaric density's connection to PPD appears to stem from the possibility to integrate (1) exactly whenever a well defined θ/S relationshp exists, thus motivating SS00 to define orthobaric density as as the function of in-situ density ρ and pressure p solving (1) for a θ/S relationship that best approximates the present day θ/S properties of the ocean.

[1] It is important to note that the mathematical validity of the procedure is questionable, since increasing the number of reference pressures towards infinity seems to result in the reference pressure converging towards $p_r = p_r(z)$ a function of z only. This in turn seems to result in PPD converging towards a form of Boussinesq in-situ density, which is physically unacceptable. This mathematical difficulty seems to have been overlooked so far, yet it is central for making sense of the concept of 'locally referenced potential density', as further discussed in the text.

[2] In JMD97, the minimisation of $\mathcal{R}$ is not formally defined and appears to be done subjectively through a trial and error procedure, but it could presumably be made more rigorous by defining a minimising cost function

$$\text{Cost Function} = \int_V W(\mathbf{x}) \|\mathcal{R}\|^2 \mathrm{d}V, \tag{3}$$

for some weighting function $W(\mathbf{x})$, similarly as in Eden and Willebrand (1999).

Although orthobaric density appears to be somewhat less neutral than γ^n, e.g., McDougall and Jackett (2005), orthobaric density possesses nevertheless several attractive advantages over γ^n, such as an exact geostrophic streamfunction and well defined formal properties, making it more suited to theoretical studies or as a generalised vertical coordinate, e.g., de Szoeke (2000). Neither JMD97 nor SS00 advocates material conservation as essential, yet both McDougall and Jackett (2005) and de Szoeke and Springer (2009) seem to agree that non-material conservation is undesirable; meanwhile, Eden and Willebrand (1999) have advocated that γ should minimise non-neutrality while retaining material conservation, an approach that they illustrated only for the Atlantic ocean case.

From a fundamental viewpoint, none of the above approaches is really satisfactory, however, for they all rely to varying degrees on ad-hoc assumptions having no clear physical justification. In this paper, we introduce a new quasi-neutral pressure-corrected density variable — called thermodynamic neutral density γ^T — which in contrast to previous approaches can entirely be constructed from first physical principles. Moreover, its physical basis is remarkably simple: in order for two fluid parcels to have the same γ^T label (in JMD97 speak), they need to belong to the same density surface in Lorenz reference state, that is, the notional state of rest that can in principle be obtained by means of an adiabatic and isohaline re-arrangement of the actual state, first defined in the theory of available potential energy (Lorenz 1955; Tailleux 2013a). Contrary to what is often assumed, e.g., Roquet (2013), Lorenz reference state is well defined even for an ocean with a realistic nonlinear equation of state (Saenz et al. 2015; Hieronymus and Nycander 2015).

Specifically, we define γ^T as the difference between Lorenz reference density and an empirical pressure correction term. Physically, Lorenz reference density $\rho(S, \theta, p_r)$ is potential density referenced to Lorenz reference pressure. If the time-dependence of Lorenz reference state is neglected — which is sufficient for the present purposes — Lorenz reference pressure $p_r = p_r(S, \theta)$ is then materially conserved, e.g., Tailleux (2013b), and so is γ^T. The empirical pressure correction is a function of p_r only, and can either be constructed from first phyiscal principles or calibrated to make γ^T traceable to γ^n, which is the approach explored here. Traceability, as defined in Huber et al. (2015), aims to make one quantity behave as much as a given target quantity via some calibration process, in order to facilitate the interpretation of the differences between the two quantities directly in terms of differences in methodologies, rather than due to some of the arbitrary choices usually entering the construction of such quantities. The motivation here is that if the procedure succeeds in making γ^T and γ^n virtually indistinguishable from each other in most of θ/S space or on oceanographic sections, which appears to be the case, one will be able to argue that γ^n might actually represent a previously unrecognised attempt at approximating γ^T. The present approach is very different to that previously pursued by SS00 for instance, who did not attempt to make orthobaric density traceable to γ^n (for instance, by constructing it based on a T/S relationship that would minimise its differences with γ^n, rather that by minimising some ad-hoc cost function). Although both group of investigators insist that orthobaric density and neutral density should be regarded as distinct concepts, we argue that this can only unambiguously established by comparing γ^n with a traceable form of orthobaric density, which remains to be done. Indeed, since there is no unique way to construct orthobaric density, the differences discussed by McDougall and Jackett (2005) or de Szoeke and Springer (2009) lack fundamental significance.

This paper's original aim was to test JMD97's claim that neutral density is best interpreted as a form of 'locally referenced potential density' (LRPD). Indeed, it would seem that a key underlying assumption of JMD97 is that if γ^n is initialised at some point A to behave as potential density referenced to the local pressure p_A (as done at JMD97's Pacific reference cast), then γ^n will also behave as potential density referenced to the local pressure p_B at some distant point B if B is linked to A via a succession of neutral paths, regardless of the distance separating B from A. The refutation of this idea is given in Section 3, and is based on the empirical finding that γ^n appears to behave much more like Lorenz reference density than a LRPD, thus motivating the construction of thermodynamic neutral density γ^T presented here. The paper first reviews some theoretical background on quasi-neutral density variables in Section 2, then proceeds on constructing a general form of Patched Potential Density as a preliminary step to the construction of γ^T in Section 3, and concludes with a discussion of some of the implications of the present results in Section 4.

2. Theoretical background

a. What physical basis for quasi-neutral pressure-corrected density variables?

The various density variables discussed above tend to rely on distinct physical principles, which it is hence important to review in order to identify which one(s) should be regarded as the most rigorous and likely to provide the most systematic construction. Perhaps the most widely used framework (especially in papers by McDougall and co-authors) is to pose the problem in physical space via Eq. (1). How to obtain the latter from a systematic analysis of the primitive equations is unclear, however, since (1) is defined in terms of *mean* variables, thus suggesting that it is to be obtained from some averaging process, yet possesses no eddy-correlation terms. The main alternative, which underlies SS00's construction of orthobaric density, takes as its starting point the evolution equation for density written in the form

$$\frac{D\rho}{Dt} - \frac{1}{c_s^2(S,\theta,p)}\frac{Dp}{Dt} = q, \tag{4}$$

where c_s^2 is the squared speed of sound, while q represents diabatic effects due to irreversible molecular diffusive processes. The left-hand side of (4) defines the differential form $\delta\varpi = \mathrm{d}\rho - c_s^{-2}\mathrm{d}p$, which in general is not perfect and hence not integrable because of the non-zero helicity of the neutral vector; otherwise, it is well accepted that ϖ would define the most natural choice of quasi-neutral pressure-corrected density variable. Mathematically, this is equivalent to state that the total differential $\mathrm{d}\gamma$ of any mathematically well-defined quasi-neutral density variable γ can at best be written in the form

$$\mathrm{d}\gamma = \underbrace{b\left(\mathrm{d}\rho - \frac{1}{c_s^2}\mathrm{d}p\right)}_{\delta\varpi} + \delta w = -\rho b\left[\alpha \mathrm{d}\theta - \beta \mathrm{d}S\right] + \delta w, \qquad (5)$$

and involves a non-vanishing residual imperfect differential form δw, with b an integrating factor. Eq. (5) seems to be the basis for (1), as it is easily seen that the latter can be obtained from the former upon making the following substitutions $\mathrm{d}\gamma \rightarrow \nabla\gamma$, $\mathrm{d}\theta \rightarrow \nabla\theta$, $\mathrm{d}S \rightarrow \nabla S$, $\delta w \rightarrow \mathcal{R}/\pi$, as well as by interpreting S and θ as their climatological values rather than their instantaneous ones. The main advantage of (1) is that it defines the problem in standard Euclidean vector space, which makes it easy to define the 'smallness' of the residual $\mathcal{R}$ or its orthogonality with the neutral vector $\mathbf{d}$. In the space of differential forms, however, there is no natural way to define the distance or orthogonality between two differential forms. The mathematical analysis of (5) must therefore proceed differently, and rely on identifying the precise conditions that would make ϖ an exact differential and hence δw vanish, such as the existence of a well-defined θ/S relationship of the form $\theta = \theta(\gamma)$ and $S = S(\gamma)$. Such a discussion, however, is not needed for what follows, and hence beyond the scope of this paper.

b. Potential density and its generalisation(s)

SS00's orthobaric density represents one possible way to construct a density variable based on (4) or (5), but this is by no means the only possible approach. A different approach, which leads to the concept of potential density as a particular case, consists in integrating (5) by parts as follows:

$$\frac{D\rho}{Dt} - \frac{1}{c_s^2}\frac{Dp}{Dt} = \frac{D}{Dt}\left[\rho - \int_{p_r}^{p} \frac{\mathrm{d}p'}{c_s^2(S,\theta,p')}\right]$$
$$-\frac{1}{c_s^2(S,\theta,p_r)}\frac{Dp_r}{Dt} + \int_{p_r}^{p} \rho_{pS} dp' \frac{DS}{Dt} + \int_{p_r}^{p} \rho_{p\theta} dp' \frac{D\theta}{Dt}. \qquad (6)$$

The quantity within square brackets can be recognised as $\rho(S,\theta,p_r)$, that is, the potential density referenced to the reference pressure p_r. So far, only the cases of a constant or piecewise constant reference pressure p_r appear to have been discussed in the literature, so the novelty here is in extending the discussion to the case where $p_r = p_r(\mathbf{x})$ is a continuous function of space (time dependence can also be included if desired but discarded here for simplicity). Doing so, however, introduces the additional term $c_s^{-2}(S,\theta,p_r)\nabla p_r$ in the gradient of $\rho(S,\theta,p_r)$, as well as the term proportional to Dp_r/Dt in (6). Unless p_r can be defined so that ∇p_r is aligned or closely aligned with the neutral vector $\mathbf{d}$, the potential density thus defined is likely to suffer from the same undesirable compressibility dependence as in-situ density. To make progress, there seems only be two choices: either giving up on the concept of potential density referenced to a continuous reference pressure field altogether, or find a way to correct for the compressibility effects introduced by retaining the spatial variations of p_r. We explore the second of these choices by subtracting from $\rho(S,\theta,p_r)$ a density offset of the form $\sigma_r = \sigma_r(\mathbf{x})$, which leads to the following density variable

$$\gamma = \sigma(S,\theta,p_r(\mathbf{x})) - \sigma_r(\mathbf{x}), \tag{7}$$

where $\sigma(S,\theta,p) = \rho(S,\theta,p) - 1000\,\mathrm{kg.m^{-3}}$, while $p_r(\mathbf{x})$ and $\sigma_r(\mathbf{x})$ are *a priori* spatially variable reference pressure and density offset fields whose specification is the key focus of this paper. Physically, (7) can be interpreted as a form of potential density referenced to a continuously varying reference pressure field $p_r(\mathbf{x})$ (the first term), empirically corrected for pressure (the second term).

It is easily seen that (7) includes PPD as a special case for the particular choices $\sigma_r = 0$ and piecewise constant p_r, and hence that it represents a continuous analog of PPD. Since JMD97 have made the same claim for γ^n, does that mean that (7) could be a suitable mathematical descriptor of γ^n? If so, this would be of considerable interest, for one of the main drawback of γ^n is its lack of explicit mathematical expression, which has so far prevented the systematic analysis of its formal properties. As it turns out, the answer is positive, but this is not in fact surprising, as the form of (7) is in fact so general that it is always possible to find a continuous $p_r(\mathbf{x})$ and $\sigma_r(\mathbf{x})$ so that $\gamma(\mathbf{x}) = \gamma^n(\mathbf{x})$ at every point in the ocean [3]. In other words, the possibility of writing γ^n in the form (7) does not in itself shed any light on the problem, unless one is able to further constrain the form of the density offset $\sigma_r(\mathbf{x})$. That this is in fact possible is established in next section, which suggests that the density offset can in fact be constrained to be of the form $\sigma_r(\mathbf{x}) = \sigma_{r1d}(p_r(\mathbf{x}))$, that is, as a function of p_r alone, allowing one to rewrite γ^n as

$$\gamma^n = \sigma(S, \theta, p_r(\mathbf{x})) - \sigma_{r1d}(p_r(\mathbf{x})) = \gamma^n(S, \theta, p_r). \tag{8}$$

If $\partial\gamma^n/\partial p_r \neq 0$, the problem of computing γ^n is then equivalent to that of computing p_r (assuming that $\sigma_{r1d}(p)$ has been determined in some way, which is discussed below), as it is then possible to compute p_r from the knowledge of γ^n and conversely. In fact, we hypothesise that all quasi-neutral density variables can be written in the form (8), with differences between different density variables arising from differences in the continuous reference pressure field p_r they implicitly rely on.

It is useful to note that (8) can also be interpreted as a classical density anomaly

$$\gamma^n = \sigma(S, \theta, p) - \sigma_{r3d}(\mathbf{x}), \tag{9}$$

defined relative to some background density field σ_{r3d} related to σ_{r1d} via

$$\sigma_{r3d}(\mathbf{x}) = \sigma_{r1d}(p_r(\mathbf{x})) + \int_{p_r}^{p} \frac{\mathrm{d}p'}{c_s^2(S, \theta, p')}, \tag{10}$$

as it is under the form (9) that JMD97 initialises the vertical profile of γ^n at the central Pacific cast located at $\mathbf{x}_{pc} = (16°S, 188°E)$ that is the starting point for their method. At this cast, JMD97 set $b = 1$ and constrains γ^n to satisfy the following equation exactly

$$\gamma^n_{\mathrm{ref}}(z) = \sigma_0(0) + \int_z^0 \frac{\rho N^2}{g} \mathrm{d}z., \tag{11}$$

which is easily integrated, using the definition of $N^2 = -(g/\rho)[\partial\sigma/\partial z + \rho g c_s^{-2}]$, as follows

$$\gamma^n_{\mathrm{ref}}(z) = \sigma(S, \theta, p) - \int_z^0 \frac{\rho g}{c_s^2} \mathrm{d}z'. \tag{12}$$

[3] For instance, take $p_r = p_0 =$ constant, and define $\sigma_r(\mathbf{x}) = \sigma(S(\mathbf{x}), \theta(\mathbf{x}), p_0) - \gamma^n(\mathbf{x})$.

This implies that at $\mathbf{x} = \mathbf{x}_{pc}$,

$$\sigma_{r3d}(\mathbf{x}_{pc}, z) = \int_z^0 \frac{\rho g}{c_s^2} \mathrm{d}z' > 0, \tag{13}$$

which in turn imposes a constraint on σ_{r1d} through (10). The question, of course, is how to formulate mathematical equations for determining the reference fields $p_r(\mathbf{x})$ and density offset $\sigma_{r3d}(\mathbf{x})$ (or $\sigma_{r1d}(p)$) in practice? This issue is addressed next.

3. Physical basis for $p_r(\mathbf{x})$ and connection with Lorenz theory of available potential energy

a. Neutral density and generalised patched potential density

In order to test our hypothesis that (7) is a useful mathematical descriptor of γ^n, we need to understand the physical principles governing the continuous reference pressure p_r and density offset σ_r that enter it. Because as mentioned earlier, the problem appears to be under-determined, we seek insights into the issue by first considering a simpler problem aiming to restrict the range of possible p_r and σ_r by introducing the following discrete version of (7), called generalised patched potential density (GPPD),

$$\gamma^{GPPD} = \sigma(S, \theta, p_{ijk}) - \sigma_{ijk} \tag{14}$$

where both p_{ijk} and σ_{ijk} are piecewise constant fields, based on a partition $V = \bigcup V_{ijk}$ of the total ocean volume, which in principle can be taken to vary in all three spatial directions, even if in what follows a two-dimensional latitude/depth partition V_{jk} is used for simplicity.

Using the γ^n field supplied as part of Gouretski and Koltermann (2004) WOCE dataset, we then seek to compute the 2D fields p_{jk} and σ_{jk} that minimise the misfit between γ^n and γ^{GPPD} for a given partition V_{jk} of the ocean. This is done here in the particular case of the two-dimensional partition of the ocean volume depicted in the top panel of Fig. 2, with $\Delta z = 500\,\mathrm{m}$ and $\Delta y \approx 40°$, using a least-square approach to find the optimal values of p_{jk}

and σ_{jk} in each subdomain. The results are illustrated in the top panels of Figs. 2 and 3 respectively; interestingly, they strongly suggest that σ_{jk} is a function of p_{jk} only, which is confirmed by a standard regression analysis, and which we use as the basis for constraining σ_r in (7) be a function of p_r alone in the rest of the paper. The associated plot for $\gamma^{\rm GPPD}$ is given in the top panel of Fig. 1 for the $30°W$ latitude/depth section in the Atlantic ocean, which can be compared with the corresponding section for γ^n in the middle panel. The strong similarity between the two figures is striking, given that the ability of $\gamma^{\rm GPPD}$ to reproduce the main features of γ^n is achieved with only $7 \times 11 = 77$ discrete reference pressures p_{jk}; the visual agreement is further confirmed by the scatter plot of γ^n against $\gamma^{\rm GPPD}$ depicted in the top panel of Fig. 4, which shows a near perfect correlation between the two quantities (the outliers seemingly originating from somewhat strange values of WOCE γ^n in enclosed seas). An histogram of the differences $\gamma^{\rm GPPD} - \gamma^n$ (blue bars in bottom panel of Fig. 4) shows that $\gamma^{\rm GPPD}$ approximates γ^n to better than $0.01\,{\rm kg.m^{-3}}$ in most of the ocean, which is remarkable.

In addition to provide a crucial constraint on the form of the density offset, the above procedure is also useful for suggesting that the reference pressure 'seen' by γ^n is a function of both depth *and* latitude. This result is important, because it conflicts with JMD97's claim that γ^n can be interpreted as a continuous analog of PPD (recall that the piecewise constant pressures used by PPD vary only with depth), or that γ^n can be interpreted as a form of 'locally referenced potential density (LRPD)', since some of the p_{jk}'s appear to occasionally depart significantly from the local box mean pressure. As it turns out, the structure of the p_{jk}'s is in fact much more reminiscent of that of the reference pressures that fluid parcels would have in their reference state of minimum potential energy that have been recently described in some recent advances in APE theory by Tailleux (2013b) and Saenz et al. (2015). The possibility to use APE theory to provide a physical basis for p_r is discussed next.

b. *APE theory as a physical basis for $p_r(\mathbf{x})$*

Motivated by the results of the previous section, we introduce a new quasi-neutral density variable, called thermodynamic neutral density γ^T, defined as

$$\gamma^T(S,\theta) = \sigma(S,\theta,p_r^{LZ}(S,\theta)) - \sigma_{r1d}(p_r^{LZ}(S,\theta)), \tag{15}$$

where $p_r^{LZ}(S,\theta)$ is the reference pressure that a parcel would have if brought in a notional reference state of rest obtained by means of an adiabatic and isohaline re-arrangement of the actual state. As shown recently by Tailleux (2013b) and Saenz et al. (2015), the reference pressure p_r^{LZ} that fluid parcels would have in Lorenz reference state of minimum potential energy $\rho_r^{LZ}(z)$ is the solution of the level of neutral buoyancy (LNB) equation

$$\rho(S,\theta,p_r^{LZ}(z_r)) = \rho_r^{LZ}(z_r), \tag{16}$$

where the possible time-dependence of the reference state, e.g., Tailleux (2013a), is neglected for simplicity. Importantly, the LNB equation (16) implies that the reference depth of fluid parcels $z_r = z_r(S,\theta)$ is a materially conserved quantity; solving (16) at all points in the ocean provides the following explicit construction for the continuous reference pressure field $p_r(\mathbf{x})$, namely

$$p_r(\mathbf{x}) = p_r^{LZ}(z_r(S(\mathbf{x}),\theta(\mathbf{x}))). \tag{17}$$

The reference density profile $\rho_r^{LZ}(z)$ was estimated for the WOCE dataset following the methodology detailed in Saenz et al. (2015), with an example of the resulting $p_r(\mathbf{x})$ field at $30°W$ in the Atlantic ocean being illustrated in the bottom panel of Fig. 2).

c. *Comparison between γ^T and γ^n*

In order to compare γ^T with γ^n, one first needs to find a way to define the pressure-dependent density offset $\sigma_{r1d}(p)$. This was done here by means of a joint pdf analysis of the respective distributions of $\rho(S,\theta,p_r^{LZ})$ and γ^n, with $\sigma_{r1d}(p)$ constructed so as to minimise the misfit between γ^T and γ^n. The distribution for γ^T obtained from such a procedure is depicted in the bottom of Fig. 1 for the same Atlantic ocean section at $30°W$ previously used. Clearly, γ^T appears to capture all the main features of γ^n except in the upper region of the ACC where γ^T displays some inversions not seen in γ^n. At the same time, it seems important to point out that as explained in JMD97, the values of γ^n in the Southern Ocean are not obtained from the actual Levitus data, but from modified ones, an approach that is avoided here. Indeed, JMD97 found it necessary to modify the Levitus data owing to the difficulty of neutrally connecting southern ocean values with values further north with the original Levitus data. It seems plausible, therefore, that this is the main reason for the observed differences between γ^T and γ^n in this region. Apart from this issue, Fig. 4) (top panel) shows that γ^T and γ^n are otherwise extremely well correlated. The bottom panel shows an histogram of the differences between the two variables, which reveal that γ^T does in general better than γ^{GPPD} at approximating γ^n, although it also reveals a few instances of rather large differences between γ^T and γ^n that do not exist for γ^{GPPD}.

Another way to compare γ^T and γ^S is directly in (θ, S) space. Although γ^n is not materially conserved, it is nevertheless possible to write it as a sum of a materially conserved part $\gamma^n_{material}(S,\theta)$ plus some residual $\delta\gamma$. For the present purposes, we estimated $\gamma^n_{material}(S,\theta)$ as the bin-average of γ^n in (θ, S) space, using $\Delta S = 0.1$ psu and $\Delta\theta = 0.1°C$ for the binning, which is equivalent to defining $\gamma^n_{material}$ as the materially conserved function of θ and S that best approximates γ^n in a least-square sense (see also McDougall and Jackett (2005) for an alternative take on the same issue). Fig. 5 top, middle and bottom panels show $\gamma^n_{material}$, γ^T and their residual respectively. Remarkably, γ^T and $\gamma^n_{material}$ appear to exhibit the same functional dependence on S and θ for most of the ocean water masses, suggesting that the non-materiality of γ^n might be the primary cause for the observed differences between γ^T and γ^n, even though the residual $\gamma^T - \gamma^n$ appears to have a rather complex structure. Since the estimation of the non-materiality of γ^n has proven so far technically complex and controversial (see de Szoeke and Springer (2009) versus McDougall and Jackett (2005)), the present results are interesting as they might point to a potentially much simpler way to quantify the non-materiality of γ^n, which is beyond the scope of this study.

To conclude this paragraph, it is important to point out that the structure of the differences between γ^T and γ^n is somewhat sensitive to the way — by no means unique — that the function $\sigma_{r1d}(p_r)$ is constructed, and hence that these differences should not be interpreted literally or as being definitive, as there might be alternative ways to construct $\sigma_{r1d}(p_r)$ that would result in an even better agreement between γ^n and γ^T. On the other hand, it is also important to recognise that rather than constructing $\sigma_{r1d}(p_r)$ to minimise the differences between γ^T and γ^n, one might prefer to define it based in physical arguments. The most natural approach would be in terms of a globally-defined θ/S relationship parameterised in terms of p_r, that is of the form $S_r(p_r)$ and $\theta_r(p_r)$, which would yield

$$\sigma_{r1d}(p_r) = \frac{1}{c_s^2(S_r(p_r), \theta_r(p_r), p_r)}. \tag{18}$$

This approach, however, is beyond the scope of the present paper, and will be discussed in a subsequent study.

d. A posteriori rationalisation of the relevance of Lorenz reference state to the theory of quasi-neutral density variables

The strong agreement found between γ^n and γ^T suggests that Lorenz APE theory is key to the theoretical understanding of neutral density variables; can this be rationalised *a posteriori*? To see that this is indeed the case, it seems sufficient to remark that the construction of neutral density as proposed by JMD97 would be trivial for a resting ocean, as neutral surfaces would then coincide with Lorenz reference density surfaces. Now, if neutral density γ^n had been constrained to be materially conserved, neutral surfaces would still have to coincide with Lorenz reference density surfaces in the actual state, since Lorenz reference density is a materially conserved variable by construction. It follows that differences between γ^n surfaces and γ^T surfaces can only come from the non-material conservation of γ^n, which is generally considered to be small, establishing *a posteriori* that γ^n and γ^T should indeed be expected to be strongly correlated. This also establishes *a posteriori* that γ^T should be regarded as the most natural definition of quasi-neutral density if material conservation is retained, which seems essential for studying ocean mixing. QED.

4. Discussion

In this paper, we have revisited the theory of neutral density by establishing that contrary to what is commonly assumed, JMD97's empirical neutral density γ^n does not behave as a locally referenced potential density (LRPD), but as the potential density referenced to its Lorenz reference state pressure. This was established by introducing a new materially-conserved neutral density variable — called thermodynamic neutral density γ^T — a function of Lorenz reference density only, which was calibrated to minimise its misfit with γ^n. The close agreement between the two variables, and the fact that they are often virtually indistinguishable from each other when plotted on oceanographic sections or in (θ/S) space, suggest that JMD97 neutral surfaces actually represent a previously unrecognised attempt at recovering Lorenz reference density surfaces.

This is an important result, for it makes it possible to reconcile the theory of neutral density — which so far has been the basis for thinking about how to define isopycnal and diapycnal directions — with at least two important developments over the past 20 years or so, all pointing to the key role of Lorenz reference state for studying turbulent diapycnal mixing and meso-scale ocean eddies, namely: Winters et al. (1995)'s proposal to use Lorenz reference state to rigorously quantify turbulent diapycnal mixing and Gent et al. (1995)'s proposal to parameterise the effect of meso-scale ocean eddies as net sinks of available potential energy. Moving towards defining isopycnal and diapycnal mixing based on γ^T rather than in terms of γ^n or local neutral tangent planes would help putting the study of ocean mixing on a more rigorous footing that has been the case so far, as APE theory already benefits from a considerable body of literature while currently undergoing rapid and exciting developments. Moreover, it would also allows the study of ocean mixing in terms of a strictly materially conserved quantity, also defining it in terms of a strictly materially conserved quantity, and therefore make it more easily applied to Walin (1982)-type water mass analysis for instance.

From a practical viewpoint, there are also considerable advantages in using γ^T in place of γ^n. Indeed, investigators currently willing to use γ^n have no choice but to use the computational software made available to the community by JMD97. This software — which most investigators use as a black box — is known to be computationally expensive and restricted to the analysis of present-day climatologies of temperature and salinity (and pre-TEOS10) (which does not stop its use for different kinds of climatology). In contrast, the construction of Lorenz reference state proposed by Saenz et al. (2015) is straightforward and computationally cheap, and physically amounts to map water masses volume in thermohaline (θ/S) space onto physical space, a much simpler and cleaner approach than sorting previously used by Huang (2005). It therefore does not require integration along characterestics, which is arguably mathematically ill-suited to the construction of a density variable in the ocean owing to the presence of orographic features and surface boundaries, e.g., Klocker et al. (2009).

In conclusion, we believe that the above arguments represent overwhelming evidence that γ^T represents the elusive quasi-neutral pressure-corrected density variable that oceanographers have been seeking since Montgomery (1938) pioneering analysis, which we hope to demonstrate further in subsequent studies.

Acknowledgments.

This work was supported by NERC Grant NE/K016083/1, "Improving simple climate models through a traceable and process-based analysis of ocean heat uptake (INSPECT)". Comments by David Marshall, Geoff Stanley, Juan Saenz and Magnus Hieronymus significantly improved clarity and presentation, for which the author is grateful. Data and software needed to reproduce the results presented here are made freely available on the author's software pages at: www.met.reading.ac.uk/ sws04rgt/software. The study also benefited from stimulating discussions with Trevor McDougall, and current and former members of his group, Paul Barker, Andreas Klocker and Stefan Riha.

REFERENCES

de Szoeke, R. A., 2000: Equations of motion using thermodynamic coordinates. *J. Phys. Oceanogr.*, **30**, 2184–2829.

de Szoeke, R. A. and S. R. Springer, 2000: Orthobaric density: A thermodynamic variable for ocean circulation studies. *J. Phys. Oceanogr.*, **30**, 2830–2852.

de Szoeke, R. A. and S. R. Springer, 2009: The materiality and neutrality of neutral density and orthobaric density. *J. Phys. Oceanogr.*, **39**, 1779–1799.

Eden, C. and J. Willebrand, 1999: Neutral density revisited. *Deep Sea Res. Part II*, **46**, 33–54.

Gent, P. R., J. Willebrand, T. J. McWilliams, and J. C. McWilliams, 1995: Parameterizing eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.*, **25**, 463–474.

Gouretski, V. V. and K. P. Koltermann, 2004: WOCE global hydrographic climatology. *Berichte des Bundesamtes für Seeshifffahrt und Hydrographie Tech. Rep. 35/2004*, 49pp.

Griffies, S. M., 2004: *Fundamentals of ocean models*. Princeton University Press, 424pp pp.

Griffies, S. M., et al., 2000: Developments in ocean climate modeling. *Ocean Modelling*, **2**, 123–192.

Hieronymus, M. and J. Nycander, 2015: Finding the minimum potential energy state by adiabatic parcel rearrangements with a nonlinear equation of state: an exact solution in polynomial time. *J. Phys. Oceanogr.*, **45**, 1843–1857.

Huang, R. X., 2005: Available potential energy in the world's oceans. *J. Mar. Res.*, **63**, 141–158.

Huang, R. X., 2014: Adiabatic density surface, neutral density surface, potential density surface, and mixing parth. *J. Trop. Oceanogr.*, **33**, 1–19.

Huber, M., R. Tailleux, D. Ferrerira, T. Kuhlbrodt, and J. M. Gregory, 2015: A traceable physical calibration of the vertical advection-diffusion equation for modeling ocean heat uptake. *Geophis. Res. Lett.*, **42**, 2333–2342, doi:10.1002/2015GL063383.

Hughes, G. O., A. M. Hogg, and R. W. Griffiths, 2009: Available potential energy and irreversible mixing in the meridional overturning circulation. *J. Phys. Oceanogr.*, **39**, 3130–3146.

Iudicone, D., G. Madec, and T. J. McDougall, 2008: Water-mass transformations in a neutral density framework and the key role of light penetration. *J. Phys. Oceanogr.*, **38**, 1357–1376.

Jackett, D. R. and T. J. McDougall, 1997: A neutral density variable for the world's oceans. *J. Phys. Oceanogr.*, **27**, 237–263.

Klocker, A., T. J. McDougall, and D. R. Jackett, 2009: A new method for forming approximately neutral surfaces. *Ocean Sci.*, **5**, 155–172.

Lorenz, E. N., 1955: Available potential energy and the maintenance of the general circulation. *Tellus*, **7**, 138–157.

Lozier, M. S., M. S. McCartney, and W. B. Owens, 1994: Anomalous anomalies in averaged hydorgraphic data. *J. Phys. Oceanogr.*, **24**, 2624–2638.

Lynn, R. J. and J. L. Reid, 1968: Characteristics and circulation of deep and abyssal waters. *Deep-Sea Res.*, **15**, 577–598.

McDougall, T. J., 1987: Neutral surfaces. *J. Phys. Oceanogr.*, **17**, 1950–1964.

McDougall, T. J. and D. R. Jackett, 2005: The material derivative of neutral density. *J. Mar. Res.*, **63**, 159–185.

Montgomery, R. B., 1938: Circulation in the upper layers of the southern north atlantic, deduced with the use of isentropic analysis. *Pap. Phys. Oceanogr. Meteor*, **6(2)**, 55 pp.

Nycander, J., 2011: Energy conversion, mixing energy, and neutral surfaces with a nonlinear equation of state. *J. Phys. Oceanogr.*, **41**, 28–41.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158.

Reid, J. L., 1994: On the total geostrophic circulation of north atlantic ocean: flow patterns, tracers and transports. *Prog. Oceanogr.*, **33**, 1–92.

Roquet, F., 2013: Dynamical potential energy: A new approach to ocean energetics. *J. Phys. Oceanogr.*, **43**, 457–476.

Saenz, J. A., R. Tailleux, E. D. Butler, G. O. Hughes, and K. I. C. Oliver, 2015: Estimating lorenz's reference state in an ocean with a nonlinear equation of state for seawater. *J. Phys. Oceanogr.*, **45**, 1242–1257.

Tailleux, R., 2009: On the energetics of turbulent stratified mixing, irreversible thermodynamics, boussinesq models, and the ocean heat engine controversy. *J. Fluid Mech.*, **638**, 339–382.

Tailleux, R., 2013a: Available potential energy and exergy in stratified fluids. *Ann. Rev. Fluid Mech.*, **45**, 35–58.

Tailleux, R., 2013b: Available potential energy density for a multicomponent boussinesq fluid with a nonlinear equation of state. *J. Fluid Mech.*, **735**, 499–518.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34**, 187–195.

Winters, K. B., P. N. Lombard, J. J. Riley, and E. A. d'Asaro, 1995: Available potential energy and mixing in density stratified fluids. *J. Fluid Mech.*, **289**, 115–128.

Wolfe, C. L., 2014: Approximations to the ocean's residual circulation in arbitrary tracer coordinates. *Ocean Modelling*, **75**, 20–35.

Wunsch, C., 1996: *The ocean circulation inverse problem*. Cambridge University Press.

List of Figures

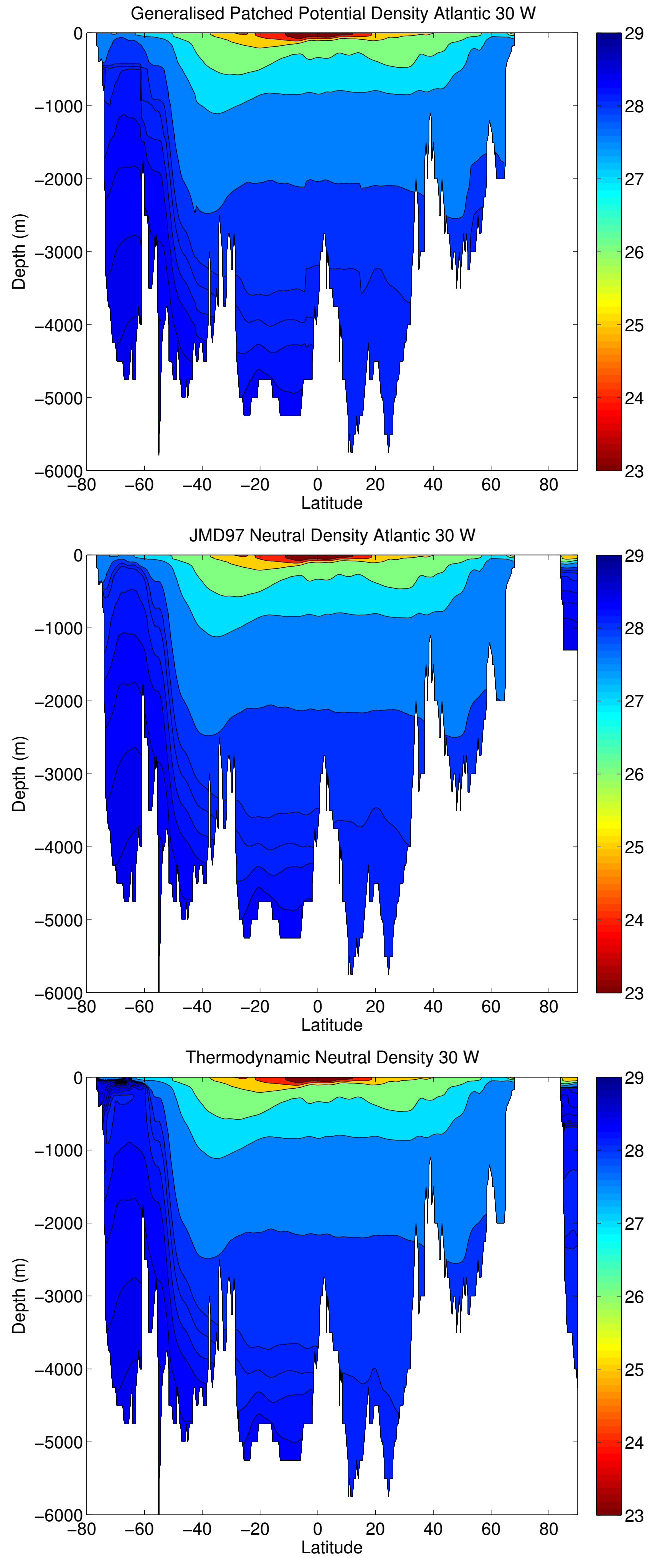

FIG. 1. (Top) Generalised patched potential density (GPPD) based on the GPPD reference pressure depicted in the top of Fig. 2 and GPPD density offset depicted in Fig. 3 at $30°W$ in the Atlantic ocean. (Middle) Neutral density γ^n at the same longitude. (Bottom) Lorenz neutral density based on the reference pressure depicted in the bottom of Fig. 2.

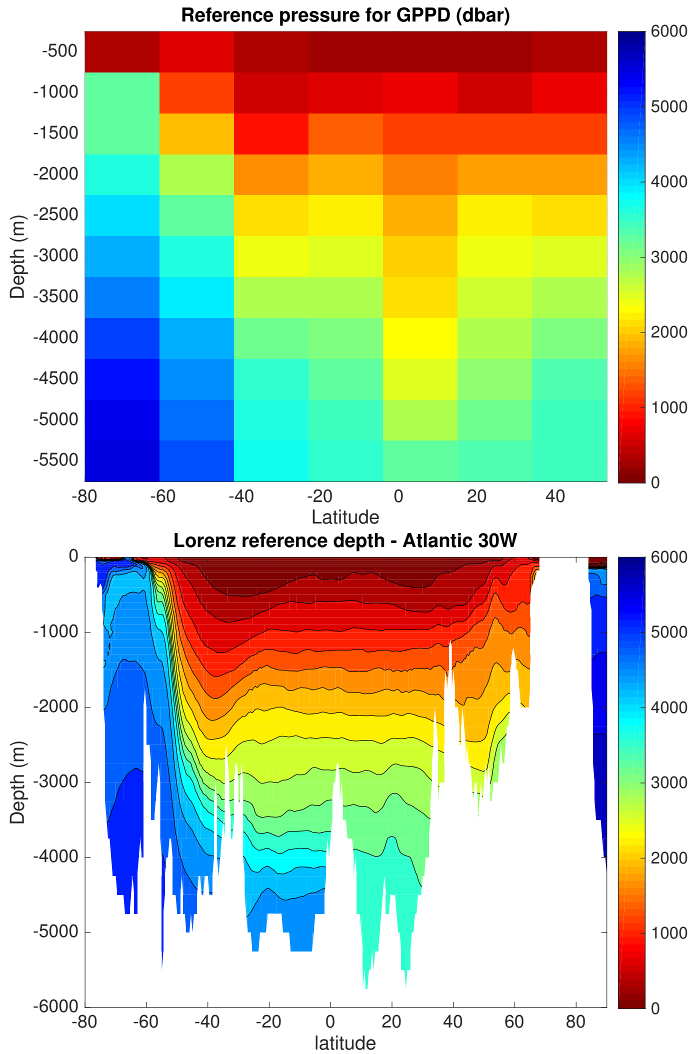

Fig. 2. (Top) The latitude-depth dependent reference pressure seen by the Generalised Patched Potential Density depicted in the top panel of Fig. 1, as obtained through regression against neutral density in the discrete domains indicated by the grid. (Bottom) The reference pressure associated with Lorenz reference state underlying the Lorenz neutral density depicted in the bottom of Fig. 1.

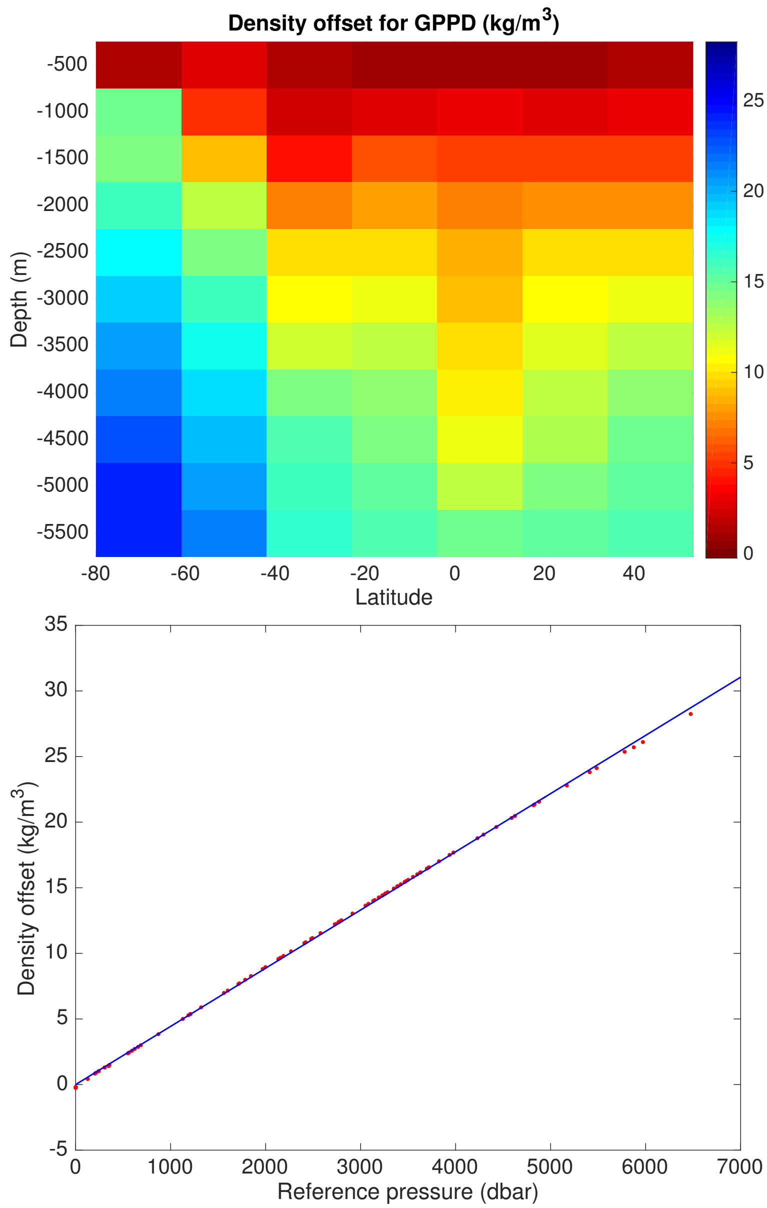

Fig. 3. (Top) The density offset used in the construction of the GPPD depicted in the top Fig. 1. (Bottom) Scatter plot of the GPPD density offset against the GPPD reference pressure, with the straight line being the best fit linear regression.

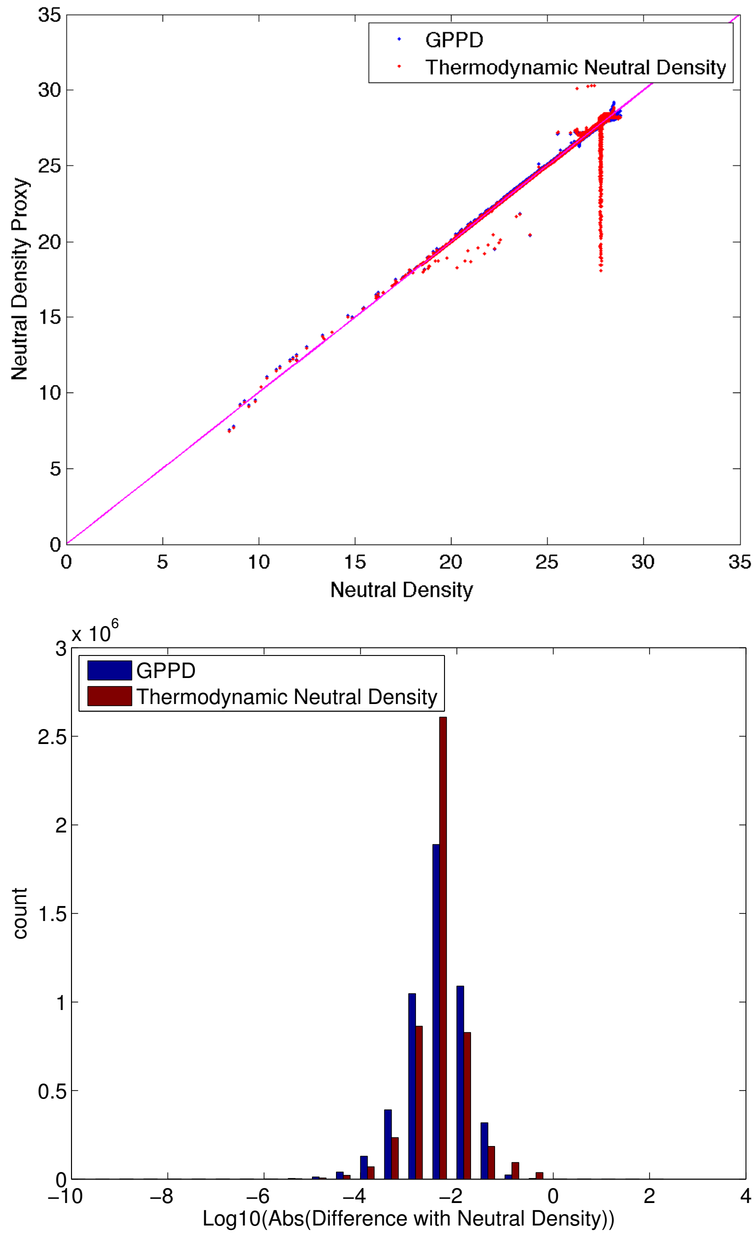

FIG. 4. (Top) Scatter plot of Neutral density against GPPD (blue dots) and of neutral density against Lorenz neutral density (red dots), with the magenta line being the straight line of equation $y = x$. (Bottom) Histogram of the decimal logarithm of the absolute value of the differences between neutral density and GPPD (blue bars) and between neutral density and Lorenz neutral density (red bars).

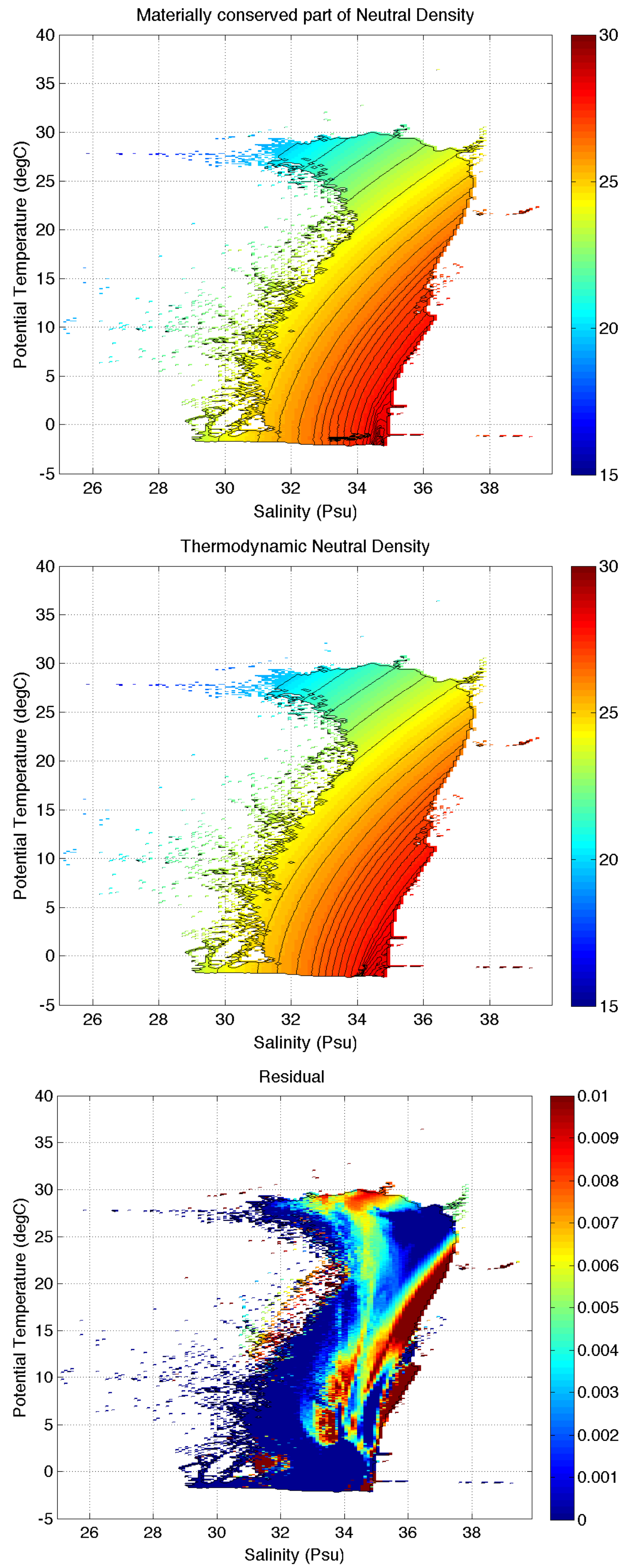

FIG. 5. (Top) Materially-conserved part of γ^n obtained by bin-averaging γ^n in (θ, S) space. (Middle) The quasi-material Lorenz neutral density bin-averaged in the same way as γ^n. (Bottom) difference between the top and middle panels.