Referee report on: The interpretation of temperature and salinity variables in numerical ocean model output, and the calculation of heat fluxes and heat content
by McDougall, Barker, Holmes, Pawlowicz, Griffies and Durack

General comments The main aim of this paper is to argue that the potential temperature θ calculated by EOS-80 based numerical ocean models should in fact be interpreted as Conservative Temperature, thus allowing it to be directly compared to the CT calculated by TEOS-10 based models. In this respect, the authors strongly dissent from Griffies et al. (2016), who recommended that TEOS-10 based models diagnose θ from CT as the temperature variable to be compared with that of EOS-80 based models. The authors also recommend that the salinity variable of models should be interpreted as preformed salinity, in contrast to Griffies et al. (2016), who recommended to retain the same interpretation of salinity as before TEOS-10.

As a practicing theoretician with over 30 years experience in analysing and constructing theoretical models of the ocean and atmosphere, there is absolutely no doubt in my mind that the new recommendations proposed are incorrect and based on flawed logic, and that ocean modellers should stick with Griffies et al. (2016)'s recommendations for the time being. From a theoretical viewpoint, the identity and nature of the temperature and salinity variables used by a numerical model must be consistent with the choice of equation of state. As regards to salinity, the fact that all the thermodynamic properties that can be derived from the TEOS-10 Gibbs function are only well defined for Reference Composition Salinity S_R means that there is no choice but to interpret salinity in TEOS-10 models as S_R. It may be true that such an interpretation results in errors in the estimation of density gradients, but the fact is that accounting for variations in composition in the binary fluid idealisation underlying current theories of the ocean circulation and ocean models is fundamentally a mathematically ill-posed problem, which implies that more sophisticated interpretations of salinity are irrelevant for current modelling practices, as correctly recognised by Griffies et al. (2016). In other words, any ocean model that would attempt to use Absolute Salinity instead of S_R should be regarded as fundamentally physically suspicious. Although the authors are correct to point out that using density salinity or absolute salinity may be useful for improving the accuracy of the thermal wind and density gradients, studies such as Feistel et al. (2010) suggest that the approach may however introduce errors in other key quantities such as the speed of sound, which in turn would introduce errors in the estimation of potential densities, buoyancy frequency, enthalpy and total energy for instance. Encouraging ocean modellers to test the impact of using Absolute Salinity without telling them about its potential detrimental effects seems a bit misleading and likely to generate a lot of pseudo-scientific papers of dubious physical soundness. Is this what we really want? As regards to EOS-80, the fact that is has been calibrated using practical salinity S_p similarly implies that salinity in EOS-80 based models must be interpreted as S_p. As to the evolution equations used to forward the prognostic variables forward in time, they are not supposed to have any bearing on how such variables should be interpreted, because it is always possible to interpret such equations as approximations to the 'true' equations. As a result, the 'normal' and expected behaviour of anybody who discovers that the some of the equations used may turn out to be less accurate than previously thought is to propose a way to correct the inaccurate equations, and to run comparison tests demonstrating the superiority of the new approach over the old one. For over 17 years now, Prof. McDougall have tried to construct a 'perfect' and 'irrefutable' argument based on logic and reasoning, of which the present paper is the latest incarnation, that would convince the ocean modelling community that treating potential temperature as conservative is very bad and harmful. His proposed recommendation, namely that models should switch to CT is interesting and physically sound as far as I can judge; however, I am not at all convinced that it would be as beneficial as Prof. McDougall thinks it is. The problem is that switching to using CT does nothing in itself to confirm or refute that treating potential temperature as conservative in EOS-80 based models is as inaccurate and defective as he claims. Yet, the issue could have long been settled once and for all if Prof. McDougall had put money where his mouth is. Indeed, one important outcome of McDougall (2003), Graham and McDougall (2013), IOC et al. (2010) and my own work is that we now understand how to correct the equation for potential temperature in EOS-80 based models to conform to Prof. McDougall's requirements. All is needed is to replace the boundary condition Q/c_p by replacing the currently constant heat capacity by $c_p(\theta, S, p_a)$, and by adding the missing non-conservation terms by diagnosing these in the manner proposed by Graham and McDougall (2013) or Tailleux (2015).

As far as I can judge, running an EOS-80 based model with the current and corrected equation is the only way to ascertain that the current way of treating potential temperature as conservative is as based as Prof. McDougall claims. It is therefore imperative if we are to settle this issue once and for all that somebody, preferably Prof. McDougall and his team, perform such an experiment, which is the only way I can think of to establish a sound and rigorous physical basis for switching to CT. Failing this, whether or not treating θ as conservative is as bad as Prof. McDougall claims will remain speculative and purely based on indirect evidence. In any case, I don't think that the authors' ideas and recommendations should be published until their scientific merits has been established by running an EOS-80 based model with the 'correct' potential temperature equation.

In the following, I list the major points of contention and disagreement I have with the authors, as well as specific comments on the text, which in my opinion refute the validity of their recommendations. I hope that my comments will convince the authors that they have overlooked and misunderstood too many key issues, and hence that their best course of action is to withdraw their manuscript and retract their recommendations.

Major points of contention and disagreement

1. **'Interpreting θ as CT' is equivalent to 'interpreting an orange as an apple'** The authors' recommendation that θ in EOS-80 models should be interpreted as CT presupposes that the two quantities are of the same nature, but we all know that this is not the case. Indeed, while θ is truly a temperature that can be experimentally measured, CT is truly a non-measurable re-scaled energy quantity disguising as temperature. Moreover, since enthalpy and potential enthalpy are defined up to a linear function of salinity, it follows that the construction of CT involves the specification of three arbitrary parameters, two associated with the said linear function of salinity, one associated with the least-square determination of c_p^0. Now, while Prof. McDougall knows about the values of the three arbitrary parameters determining his construction of CT, the potential temperature of an EOS-80 model obviously does not. Can the authors explain how it is somehow possible for θ to morph into CT without θ having any knowledge of the particular determination of CT it is supposed to morph into? How is it possible for θ to somehow morph into CT if it does not know which determination of CT it is supposed to morph into?

2. **$\theta - \Theta$ is a physically-meaningless object that is completely devoid of physical meaning** Because standard physics teaches us that two quantities can only be compared if they are of the same nature, it follows that one should regard the quantity $\theta - \Theta$ as completely devoid of physical meaning, since it is the difference between a temperature and an energy, whose value depends on the specification of three arbitrary parameters. Yet, the authors seem to suggest that the values of $\theta - \Theta$ — a physically meaningless object — should be regarded as somehow representative of the errors arising from treating θ as conservative in EOS-80 models. How is that possible? Have the authors somehow being taught differently in their physics classes?

3. **One of the premises of the syllogism used by the authors to prove their argument is flawed** As far as I can judge, the authors arrive at their conclusions that θ should be interpreted as CT in EOS-80 based models by using the following syllogism:

   - *Numerical models assume potential temperature to be conservative.*
   - *We know that potential temperature is not conservative.*
   - *Therefore, potential temperature in EOS80-based ocean models cannot truly be potential temperature and hence should be interpreted as Conservative Temperature.*

   Although I agree that the use of syllogisms represents a valid tool in logic to derive a conclusion deductively, it is also well understood that the validity of doing so crucially depends on the validity of the premises. While the validity of 'We know that potential temperature is not conservative' is indubitable thanks to McDougall

(2003), this is not so of 'Numerical models assume potential temperature to be conservative', which is arguably quite a misleading way to characterise pre-TEOS10 ocean modelling practice. Indeed, a much fairer characterisation closer to the truth is 'EOS80 based models assume that the errors made in using a constant heat capacity to compute surface fluxes of potential temperature and in neglecting interior non-conservation terms are sufficiently small that they are irrelevant in practice'. If this characterisation is used instead of the authors' premise, their syllogism no longer makes sense. As argued above, the authors would have a much stronger case if they could demonstrate the impact of correcting the potential temperature equation in an EOS-80 based model, which would be much more easily understandable by numerical ocean modellers.

4. **CT is only 2 or 3 times more conservative than θ. The conservativeness of CT has been greatly exaggerated so far** I admit to being quite confused about the authors' explanation of the origin for the non-conservative production/destruction of θ and CT, which conflicts with the exact results of Tailleux (2010). Indeed, according to the latter, the non-conservative terms $\dot{\theta}_{irr} = \dot{\theta}_{irr}^{diff} + \dot{\theta}_{irr}^{visc}$ and $\dot{\Theta}_{irr} = \dot{\Theta}_{irr}^{diff} + \dot{\Theta}_{irr}^{visc}$ are the sum of a diffusive and viscous contribution, with $\dot{\theta}_{irr}^{visc} \approx \dot{\Theta}_{irr}^{visc} \approx \varepsilon_K$, where ε_K is the local viscous dissipation rate. Yet, the authors's comparison of the relative non-conservativeness of θ and CT appears to be based solely on the diffusive part. Assuming that viscous dissipation balances about 3 TW of mechanical energy input by the wind, tides and surface buoyancy fluxes (which is most likely a significant underestimate) yields an equivalent surface flux of 10 mK.m^{-2}. Based on Graham and McDougall (2013)'s estimates, a summary of the diffusive and viscous contribution to the non-conservation of each variable is therefore:

Non-conservation	diffusive	viscous	sum
$\dot{\theta}_{irr}$	$16mK.m^{-2}$	$10mK.m^{-2}$	$26mK.m^{-2}$
$\dot{\Theta}_{irr}$	$0.3mK.m^{-2}$	$10mK.m^{-2}$	$10.3mK.m^{-2}$

While these results show that Θ is about 50 times more conservative under the action of diffusive mixing — admittedly an important result in itself and worth mentioning — its total degree of non-conservativeness accounting for viscous dissipation is only a factor 2-3 better than that of θ overall. A priori, it is the total degree of non-conservativeness that should be compared, not just that due to diffusion effects, in order to establish whether it is justified to treat CT as exactly conservative in a numerical ocean model. Moreover, as shown by Tailleux (2015), the non-conservation arising from using a constant c_p in the estimation of the surface fluxes for θ tends to be balanced by the non-conservation due to diffusive effects, at least globally, which means that one should expect to compensate, at least to some extent. As a result, it is by no means obvious that Prof. McDougall is correct. Again, this could easily be tested by running an EOS-80 based model with a corrected evolution equation for θ.

Specific comments

1. Line 63: *namely heat content ("temperature")* Given that standard thermodynamics teaches us that the concepts of 'heat' and 'temperature' should not be confused, starting the paper by confusing the two concepts does not bode well for the following, especially coming from the previous chair of WG127 and current chair of JSC who are supposed to teach us the right way of doing thermodynamics.

2. Lines 64-65. *For computational reasons, it is useful for numerical schemes involved to be conservative [...]* The authors seem to make this a central tenet of their argumentation, even though adding non-conservative terms in a conservative equation does not pose any particular challenge from a numerical viewpoint. Why do the authors consider it would problematic to add the missing nonconservative terms in the potential temperature equation and use the correct boundary condition for the surface fluxes of θ?

3. Lines 70-72. The property of any 'heat' variable to be non-conservative is a generic property of any fluid, which is not limited to seawater, and which can be defined independently of the development of any thermodynamic standard for seawater.

4. Lines 75-77. I agree that it is now widely recognised that potential temperature is not truly conservative. However, there is nothing in thermodynamics that says that the appropriate measure of heat should be conservative or approximately conservative, which it seems important to point out. The idea that 'heat' should be a conservative is idiosyncratic to Prof. McDougall and has absolutely no root in classical thermodynamics or anywhere else in the development of the subject. The only two conservative quantities for seawater idealised as a binary fluid is salinity and total energy. As shown by Tailleux (2010), it is not possible for total energy and any 'heat' variable to be simultaneously conservative. Assuming 'heat' to be conservative is strictly equivalent to assuming that total energy is not conservative, which the authors appear to have overlooked. Given that recent developments seem to focus on the construction of energetically consistent models, e.g., Eden et al. (2014), one should anticipate that ocean modellers will seek to understand how to add the missing non-conservation of 'heat' in their models in order to achieve total energy conservation.

5. Lines 70-80. *This empirical fact is an inherent property of seawater.* I disagree. Nearly all fluids a priori suffer from the same problem, as can easily be demonstrated.

6. Lines 82-83. *[...] results in inherent contradictions.* 'Contradiction' is a loaded word here, because all what the authors have established so far is that treating potential temperature and Conservative Temperature as conservative is only approximate, and that the approximation is a better one for the latter than for the former. Using the term 'contradiction' frames the problem as one that should be solved by logic alone, whereby an illogical approach can only be corrected by a logical one. In contrast, 'an approximation' can be improved by using a more accurate formulation, such as would be achieved by adding the missing non-conservation terms in the potential temperature equation and using the correct flux of potential temperature. In using the term 'contradiction', the authors signal their intention of rooting their arguments in a 'right' versus 'wrong', rather than by a direct demonstration based on comparing two EOS-80 based model using the incorrect and correct evolution equation for θ.

7. Lines 88-89. *even at the cost of introducing problems elsewhere* Why would we want to that when solutions exist to solve problems without adding new ones elsewhere?

8. LInes 94-96. *For example, the insistence that a model's temperature variable is potential temperature involves errors in the air-sea heat flux in some areas that are as large as the mean rate of global warming* This is quite a misleading statement, given that these errors are at least partly compensated by the error arising from neglecting the non-conservation of potential temperature, as suggested by the results of Tailleux (2015).

9. Lines 96-99. Heat is not a conservative property but total energy is. Why do the authors insist on conserving heat but not total energy? Why do they consider it is more logical or rational to conserve heat but not total energy?

10. Lines 105-108. *It is well known that in-situ temperature is not an appropriate measure of the "heat content" [...]* I find it very strange that the authors should discuss what is or what is not an appropriate measure of heat content in the absence of consensus on what should be the 'true' definition of heat content.

11. Lines 119 Section A.17 of IOC et al. (2010). It is interesting to see that this Appendix only discusses the diffusive part of the non-conservation of θ and Θ, completely overlooking the role of viscous dissipation, and that only the non-rigorous derivations of Graham and McDougall (2013) are cited when Tailleux (2010) gives the exact and explicit forms of non-conservation for the Navier-Stokes equations written in terms of both θ and Θ.

12. Line 141. Why is it unfortunate?

13. Line 149-150 [...] *has a mean non-conservation error in the global ocean of only about* $0.3\,\mathrm{mW.m^{-2}}$. As shown by Tailleux (2010) (his Equation (25)) and Tailleux

(2015) (his equation (26)), the exact expression for the non-conservative production/destruction of Θ is

$$\dot{\Theta}_{irr} = \frac{1}{c_p^0}\left[\varepsilon_k - \mathbf{F}_S \cdot \nabla\left(\mu - \frac{T\mu_r}{\theta}\right) - \mathbf{F}_\Theta \cdot \nabla\left(\frac{Tc_p^0}{\theta}\right)\right] \tag{1}$$

and is seen to include the viscous dissipation term ε_k, which the authors subsequently say is of the order of $3mW.m^{-2}$. Presumably, the $0.3mW/m^2$ only refer to the 'diffusive' nonconservation of Θ. The true nonconservation of Θ is therefore at least $3.3mW.m^{-2}$, which is only 5 times less than the $16mW.m^{-2}$ mean nonconservation of θ. This number increases to $10.3mW.m^{-2}$ if a more realistic estimate of $10mW.m^{-2}$ is used for viscous dissipation, as mentioned in my major comments section.

14. Re-reading Graham and McDougall (2013), I realise that the authors estimated the rate of viscous dissipation ε_k from the formula $\varepsilon_k = DN^2/\Gamma$. Their estimate corresponds to a surface integrated value of $3.10^{-3} \times 3.10^{14} = 0.9\,\mathrm{TW}$, which is way too small. Clearly, there is at least 3 TW of mechanical energy input due to the wind and surface buoyancy fluxes, not more. A more reasonable estimate of the total viscous dissipation is therefore closer to $10\,\mathrm{mW.m^{-2}}$, which is only a factor of 2 smaller than the nonconservation of θ.

15. Lines 163-164. Not really, given the above arguments. The authors are clearly applying double standards here.

16. Lines 165-167. This is a revisionist view of history, given that Conservative Temperature was introduced by McDougall (2003), long before the IAPWS group was formed, and is not actually part of the UNESCO endorsed part of TEOS-10 (as far as I am aware). This decision was Prof. McDougall's alone, and was not part of the TEOS-10 work. The SCOR/IAPSO WG127 was approved in 2005, and its first meeting took place in Warnmünde in May 2006.

17. Lines 180-183. The question of why potential temperature is non-conservative was actually answered earlier by Tailleux (2010), who showed that the non-conservation of any heat variable is dictated by the first law of thermodynamics (the law of total energy conservation). This is again a revisionist view of history where Graham and McDougall (2013) attempts to get credit for something that needs to be attributed to Tailleux (2010).

18. Lines 199 and below. How conservative is conservative temperature? As said above, this section only describes the nonconservation of Conservative Temperature arising from diffusive mixing and completely omits viscous dissipation. Given that the latter dominates, this section is at best misleading.

19. Line 382. What the authors call naive is what I call common sense. The authors should expect that many oceanographers will feel insulted here.

20. Line 448. I dispute that this is thermodynamic best practice if the authors fail to understand the results of Tailleux (2010) .

21. Lines 479. The authors confuse the terms 'contradiction' and 'approximation' - This is an idiosyncratic interpretation because this is not how idealised modelling should be viewed. Numerical modellers and oceanographers understand that potential temperature is non-conservative; it is therefore unfair to accuse them of assuming potential temperature to be conservative. Rather, they treat it as conservative because they assume that the small nonconservative terms and heat flux errors do not matter on the time scales generally considered. 'Contradiction' and 'approximation' are two completely different concepts, which it is crucial to distinguish in the present discussion, since the authors use the first interpretation in oder to be able to accuse EOS-80 based modelling practice as being illogical. The interpretation 'Models assume the potential temperature to be conservative' has been disproven and there is no reason to accuse ocean modellers of ignoring this result. On the other hand, the interpretation that 'models assume that the nonconservation of potential temperature is sufficiently small that it can be neglected in practice' has not been disproven yet. Indeed, disproving such an approach can only be achieved by running an EOS-80 based model with the incorrect and corrected equation for potential temperature. Such experiments are urgently needed so that we can stop with all the speculation.

22. Lines 735. It is not true that Conservative Temperature is consistent with the first law of thermodynamics, because it assumes that all the heat goes into heat and none into work. Indeed, it is well known from Lorenz's theory of available potential energy that there is about 0.5 TW of the surface buoyancy fluxes going into the production of available potential energy. This suggests that potential enthalpy includes APE — a work-like quantity — as well as heat, not just heat.

Literature

- Eden, C., et al 2014. Toward energetically consistent ocean models. JPO, 3160-3184.
- Feistel, R. et al., 2010. Thermophysical properties anomalies of Baltic seawater. Ocean Sciences, 6, 949–981.
- Tailleux, R., 2010. Identifying and quantifying nonconservative energy production/destruction terms in hydrostatic Boussinesq primitive equation models. Ocean Modelling, 34, 125-136. doi 10.1016/j.ocemod.2010.05.003
- Tailleux, R., 2015. Observational and energetics constraints on the non-conservation of potential/Conservative Temperature and implications for ocean modelling.