This study addresses the important issue of budget analysis closure and provides a precise budget calculation tool that can benefit the numerical modeling community. When reading the abstract, one of my concerns is that the main concept and results (including the comparison with simplified approximations, i.e., using a lower-order advection operator than the model) resemble those in Chen et al. (2020). However, after reading the entire manuscript, I think this article does contain substantial new content, and WRFlux has its unique usefulness and applicability. The practicability of Cartesian coordinate transformation is also clear in the example simulation with non-uniform orography and therefore strongly sloped $\eta$ surfaces at low levels. In all, the advantage of using WRFlux is well exhibited, and the illustration materials are nicely presented. Still, I have some comments and suggestions:

Specific comments

- **Title:** Suggest changing "energy" and "mass" to "potential temperature" and "moisture", respectively. Most results are about the $\theta$ tendency budget, which is related but not equivalent to "energy budget" in the strict sense. The current title may misleadingly imply an Earth's energy budget, moist static energy budget, turbulence kinetic energy (TKE) or total energy budget, etc. Furthermore, I don't think the result of the "mass budget" is ever shown.
- **Abstract:** I suggest reconstructing the abstract to emphasize the unique features and applicability of WRFlux to differentiate this work from Chen et al. (2020), such as the retrieval of the resolved turbulence component, transformation to the easy-to-interpret Cartesian coordinate, etc., all of which are particularly useful for large-eddy simulations or simulations over non-uniform topography.
- **Introduction:** It is important to mention that there have been some attempts to achieve a precise budget retrieval in the exact WRF model (Chen et al. 2020 is not the first), e.g., Lehner (2012, https://collections.lib.utah.edu/ark:/87278/s61n8fxw), Moisseeva and Steyn (2014, https://doi.org/10.5194/acp-14-13471-2014) and Potter et al. (2018, https://doi.org/10.1029/2018JD029427), etc.
L35-38: I’m not sure how the advective form may “hinder interpretation when internal processes dominate the budget of the control volume, which can be a single grid cell or a larger volume”? Can you please elaborate?

Chen et al. (2020) noted that a closed budget for \( w \) is challenging, partially because \( w \) equation is implicit, coupled with the geopotential tendency equation, and partially due to the variable’s inherent rapid variation on small scales. How does WRFlux overcome this issue for the \( w \) budget? It is also surprising to see that WRFlux performs the best for \( w \) compared to all the other variables and that the accuracy (in terms of NRMS and r99; Table 1) is much higher for the momentum than for the thermodynamic variables (\( \theta \) and \( q_v \)) in your case. Do you have any idea why? Is this result case-dependent or relevant to the physical design of WRFlux?

It would help readers follow Section 2 more easily if the connections between subsections are made more explicit and some terminologies are explained more precisely.

- For example, I find it confusing that “Eq. 11 is equivalent to Eq. 8” (L124-128), but “Eq. 8 cannot be closed numerically in contrast to Eq.11” (L128). This confusion may be related to how you define the term “close” here. Suppose you are referring to “budget closure”. In that case, my understanding is that it depends on the accuracy of the conducted budget calculation, not the format of the chosen equation to which the budget analysis is applied. Do you mean that because Eq. 11 is numerically consistent with the equation used in the model, the left-hand-side terms of Eq. 11 represent more accurately the actual tendency simulated by the model? And therefore, a budget analysis using Eq. 11 can achieve a higher degree of budget closure? Please clarify.

- Also, I suggest reminding readers at the beginning of Section 2.4 that Eq. 11 (i.e., Eq. 13 after discretization) will be used for the precise budget calculation. It would also be helpful to specify which terms in Eq. 11 can be decomposed into the mean advective and turbulent components (all right-hand-side terms excluding \( S \) or all right-hand-side terms excluding only some partial \( S \), i.e., the subgrid physics?)

L113-L114: “Numerical consistent” --> “Numerical consistency”. Also, placing this sentence (“Numerically... after discretization.”) here seems odd. You are neither showing the discretization nor discussing the budget closure in the following lines yet.

L130: It is still unclear to me why recalculating the vertical velocity is desired. If Eqs. 6 and 9 are also used in WRF, doesn’t it make more sense to use WRF’s prognostic to be numerically consistent with the model? Unless this is relevant to the issue I mentioned in comment #4?

L191-194: Does this part follow exactly Chen et al. (2020)?

L195: “The fluxes and all... are averaged in time during model integration”. Averaged over the entire simulation? Or over a user-specified time window? If so, what are your suggested value and the physical reason behind it?

L203-204: “Since WRF does not actually solve the continuity equation...” This may be misleading. WRF does solve the continuity equation, just not in the same format as expressed in your Eq. (12), i.e., in terms of the 3D variable \( \rho \) but column dry air mass per unit area (2D \( \mu_d \)). Note that \( \mu_d \) is indeed advanced with time in the WRF dynamical solver. I suggest changing the sentence to something like “Since WRF does not directly solve for \( \rho \) but the integrated column dry air mass...”

L244: “The setup leads to ... a very dry atmosphere, therefore moist processes are neglected.” Not sure if I overlook something but the initial setup of the moisture field is not mentioned?

L246: “We calculate full \( \theta \)-tendencies and decompose them into resolved turbulence, subgrid-scale turbulence and mean advective.” What about the rest of the retrieved budget terms, such as “physical parameterizations and numerical diffusion and damping”? To clarify, it may help to move the sentence “Since no microphysics scheme is activated and the simplified radiation scheme only affects the surface energy balance, the heat budget in the atmosphere only consists of resolved advection and subgrid-scale diffusion” (L261-263) here and mention that for general applications, other grid-resolved parameterized physics terms are possible and categorized as
additional budget components.

- Figure 1: Panels a-c show the total turbulence (trb), which is the sum of the resolved and subgrid-scale components. I'm interested in their individual contributions. E.g., What is the relative magnitude between these two components? Do they have similar spatial distributions with the same signs, or do they offset each other? Considering that one of the distinctive features of this tool from previous WRF budget retrieval works is the decomposition into mean and turbulent components, I suggest strengthening the relevant discussion.

- L283: “...in the alternative form of the equation (Eq. 8), the correction term for the time derivative is almost negligible” I’m lost. This is not shown or can be inferred by figure 2?

- L290-295: Have you checked if the sum of all the budget components in this alternative analysis is in close balance with the sum of your Eq. 11(13) in Fig. 1?

Technical corrections:

- L113, L124, ...: Replace “equation” with “Eq.” here. Please check the rest of the manuscript for consistency.
- L134: “energy” --> “potential temperature”
- L154-155: “Since... to derive.” This sentence is confusing. Suggest changing to “Although the momentum variables are staggered differently from the thermodynamic variables, their discretized equations can be derived analogously...”
- L226: “ridge-to-ridge” --> “bottom-to-bottom” or “valley-to-valley” ?
- Figure 4 legend: I believe the legends “WRFlux (Eq. 14)” should be replaced by “WRFlux (Eq. 13)”, “Eq. 15” by “Eq. 14”, and “Eq. 16” by “Eq. 15”.
- Figure 5 subtitles for each panel: Same issue as above.
- Table 1: Same issue as above.