Reconciling and improving formulations for thermodynamics and conservation principles in Earth System Models (ESMs)

Peter H Lauritzen, N.K.-r. Kevlahan, T. Toniazzo, C. Eldred, Thomas Dubos, A. Gassmann, V.E. Larson, C. Jablonowski, O. Guba, B. Shipway, et al.

Reconciling and improving formulations for thermodynamics and conservation principles in Earth System Models (ESMs)

P.H. Lauritzen[1], N.K.-R. Kevlahan[2], T. Toniazzo[3,14], C. Eldred[4], T. Dubos[5], A. Gassmann[6], V.E. Larson[7,10], C. Jablonowski[8], O. Guba[4], B. Shipway[9], B.E. Harrop[10], F. Lemarié[11], R. Tailleux[12], A.R. Herrington[1], W. Large[1], P.J. Rasch[10], A.S. Donahue[13], H. Wan[10], A. Conley[1], J.T. Bacmeister[1]

[1]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA
[2]Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada
[3]NORCE research and Bjerknes Centre for Climate Research, Norway
[4]Sandia National Laboratories, NM, USA
[5]Laboratoire de Météorologie Dynamique/IPSL, École Polytechnique, Palaiseau, France
[6]formerly: Leibniz-Institut fur Atmosphärenphysik, Kühlungsborn, Germany; now without affiliation
[7]Department of Mathematical Sciences, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin, USA
[8]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA
[9]Met Office, Fitzroy Rd, Exeter, UK
[10]Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA
[11]Univ Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France
[12]Dept of Meteorology, University of Reading, Reading, UK
[13]Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[14]Department of Meteorology, Stockholm University, Stockholm, Sweden

Key Points:

- Closing total energy budgets in ESMs without *ad hoc* fixers is a monumental task.
- Largest errors are from missing processes/terms, thermodynamic inconsistencies and dynamical core.
- Further research is needed on conservative discretizations, unified thermodynamics and missing processes.

Corresponding author: Peter Hjort Lauritzen, `pel@ucar.edu`

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2022MS003117.

Abstract

This paper provides a comprehensive derivation of the total energy equations for the atmospheric components of Earth system models (ESMs). The assumptions and approximations made in this derivation are motivated and discussed. In particular, it is emphasized that closing the energy budget is conceptually challenging and hard to achieve in practice without resorting to *ad hoc* fixers. As a concrete example, the energy budget terms are diagnosed in a realistic climate simulation using a global atmosphere model. The largest total energy errors in this example are spurious dynamical core energy dissipation, thermodynamic inconsistencies (e.g. coupling parameterizations with the host model) and missing processes/terms associated with falling precipitation and evaporation (e.g. enthalpy flux between components). The latter two errors are not, in general, reduced by increasing horizontal resolution. They are due to incomplete thermodynamic and dynamic formulations. Future research directions are proposed to reconcile and improve thermodynamics formulations and conservation principles.

Plain Language Summary

Earth System Models (ESMs) have numerous total energy budget errors. This article establishes the governing total energy equations for large-scale ESMs and assesses the energy budget errors in real-world simulations in a widely used climate model. To move towards a closed energy budget in ESMs, further research on total energy conserving discretizations (in the dynamical core), unified thermodynamics (through thermodynamic potentials/conserved variables) and missing processes is paramount. This research is especially important since some of the energy budget errors will not improve with higher spatial resolution and may even get worse.

1 Introduction

Earth system models (ESMs) are extremely complex, and represent a wide range of physical processes in separate component models (atmosphere, ocean, land surface, sea ice). In addition, each component has separate sub-modules that compute the effect of unresolved physical processes, such as cloud microphysics and radiation in atmosphere models. The Navier–Stokes equations that serve as the foundation for the ESMs used for climate change projections satisfy conservation laws for mass, energy, and momentum. These conservation laws, combined with the second law of thermodynamics, impose key constraints on the climate system. It is believed that accurate satisfaction of these constraints is critical for the integrity of long-term climate change simulations.

In practice, satisfying such constraints is complicated. Discretisation schemes tend to introduce aliasing errors, phase errors, unphysical modes, and spurious extrema. These errors may introduce spurious sources and sinks of energy, erroneous spectral transfers between energy modes, make positive definite concentrations negative etc. Some numerical schemes and generalized vertical coordinates require interpolation, which may produce similar errors.

The use of certain physical approximations (such as the anelastic or Boussinesq approximations) alter the form of the conserved energy quantities and thermodynamic potentials in (sometimes) subtle ways that need to be properly understood to avoid the introduction of spurious sinks and sources. Understanding how to endow such approximations with consistent energetics and thermodynamics is relatively recent, but has been progressing rapidly over the past decade, e.g., Ingersoll (2005), Pauluis (2008), Young (2010), Tailleux (2012), Brown et al. (2012), Vasil et al. (2013) and Dewar et al. (2016). Many of these advances, however, have yet to be widely adopted in ESMs. Additionally, the fact that physical parameterisations are developed independently may mean that different components of a model rely on inconsistent approximations for constants or con-

served quantities. For example, until quite recently NCAR's Community Atmosphere Model (CAM, Neale et al., 2010) used a total energy formula that was inconsistent with the sigma-pressure vertical coordinate (Williamson et al., 2015). Another example is when a physics subgrid scale model assumes that a certain quantity is conserved (e.g. heat based on liquid water potential temperature, θ_l) while the atmosphere model is expecting another quantity to be conserved (e.g. enthalpy). These sorts of inconsistencies can easily arise if the choices made by each sub-module, or component as a whole, are not documented and coupled consistently.

Another source of energy errors is that different components of an ESM often operate on different temporal scales and are coupled asynchronously. This may produce spurious energy transfers. Also, the fact that models cannot resolve scales below the model grid scale means that assumptions have to be made about the energy transfers between the resolved and unresolved motions. The physical validity of these assumptions is hard to assess, although emerging literature is making significant strides towards a better and more physical understanding of grid and sub-grid scale coupling (e.g., Gassmann & Herzog, 2015). The basic problem is that dissipative processes occur at the viscous scale which cannot be resolved with a finite grid size. Hence, energy propagates from resolved to unresolved scales (or vice versa) and is no longer visible to the resolved dynamics. It is therefore by no means straightforward to represent physically correct energy transfers. It is likely that designing prognostic sub-grid models that are coupled consistently with the resolved-scales is the only way to ensure physically realizable energy transfers for more complicated wave interactions.

Numerical dissipation, which is commonly used to stabilize fluid flow solvers, can produce physically unrealizable solutions. By physically unrealizable we mean that the second law of thermodynamics is not satisfied (i.e. there is "unphysical" up-gradient mixing). If one argues that the second law should be satisfied for numerical dissipation operators, then they must be positive definite. This is, for example, not the case for standard hyperdiffusion operators, but is true for carefully implemented Smagorinsky type (second-order) diffusion (Becker, 2001; Becker & Burkhardt, 2007; Schaefer-Rolffs & Becker, 2018)

Even though they continue to neglect the complications of grid/sub-grid scale coupling, research groups studying the Earth's energy imbalance using observations and reanalysis products have started developing advanced energy equations that include, for example, enthalpy fluxes associated with precipitation (Mayer et al., 2017; Trenberth & Fasullo, 2018; Kato et al., 2021). The global modeling community has been lagging behind in terms of incorporating these processes in their modeling systems. Kato et al. (2021) noted

> When models are constrained by ToA (Top of Atmosphere) radiation data products, models that conserve energy and water mass do not match an observed global mean precipitation rate (e.g. Held et al., 2019) because of, in part, the existence of a significant energy balance residual when satellite energy flux products are integrated.

Hence it is clear that even though the energy budget is closed, in part by using energy conserving formulations or the use of energy fixers where energy consistency is not possible, the energy balance still suffers from significant errors (Kato et al., 2011; L'Ecuyer et al., 2015). The causes of these errors are largely unknown, but are observed to be large over tropical oceans (Loeb et al., 2014; Kato et al., 2016; Harrop et al., 2022). Furthermore, Lucarini and Ragone (2011) analyzed pre-industrial simulations from a wide range of IPCC climate models, and found that most climate models featured biases of the order of 1 W/m^2 for the net global and the net atmospheric, oceanic, and land energy bal-

ances. They concluded that these imbalances are partly due to imperfect closure of the energy cycle in the fluid components.

As a result of these problems, demonstrating that ESMs are physically consistent is an enormously complicated problem that requires a holistic view and wide expertise that most individual scientists do not possess. In October 2019 we organized a week-long workshop at the Banff International Research Station (BIRS) in order to address some of these big-picture fundamental issues related to ESMs (BIRS, 2019). We brought together scientists involved in developing operational and research atmosphere models, physics packages, ocean models and couplers. Our meeting did not include representatives of the land surface and sea ice modeling communities, although they are clearly important. Participants included developers for the MOM6 (Adcroft et al., 2019), CROCO (Hilt et al., 2020) and NEMO (Madec et al., 2019) ocean models, the Community Atmosphere Model (CAM, Neale et al., 2010) which is part of the Community Earth System Model (CESM, Danabasoglu et al., 2020) from the National Center for Atmospheric Research (NCAR), the Energy Exascale Earth System Model (E3SM) (Golaz et al., 2019), the DYNAMICO (Dynamical Core on Icosahedral Grid), Dubos et al., 2015), the Unified Model (Walters et al., 2017), LFRic (S. Adams et al., 2019; Melvin et al., 2019) and ICON-IAP (Icosahedral Nonhydrostatic model at the Institute for Atmospheric Physics, Gassmann, 2013) atmosphere models, and the CLUBB (Cloud Layers Unified By Binormals) parameterization of turbulence and clouds in the atmosphere (Larson, 2017). Participants also included developers of Adaptive Mesh Refinement (AMR) methods for geophysical flows via the CHOMBO library (M. Adams et al., 2019; Ferguson et al., 2016, 2019) as well as WAVETRISK (Dubos & Kevlahan, 2013; Kevlahan & Dubos, 2019). Since the group was deliberately heterogeneous, the workshop began with a series of overview talks on each component (atmosphere, ocean, physics, couplers). This ensured that all participants started from a common knowledge base for each component. In addition to reviewing the basic structure and assumptions of each component, overview talks considered specific topics such as total energy errors in dynamics, energy conversion and internal entropy production, energy conservation in cloud and turbulence parameterizations, and ocean-atmosphere coupling.

The goals of the workshop were to identify the most important fundamental questions related to energy conservation in each component/sub-module and in the coupled system, and to propose prioritized future directions of research. We also decided to make some specific recommendations for best practices which take into account all components of ESMs. These classifications and recommendations were developed through extensive focused discussions involving all workshop participants during and after the workshop. With such a broad interdisciplinary group, it was challenging to assemble a coherent manuscript, and many iterations were necessary to produce this version. At an early stage we decided to focus on the atmosphere component, leaving out very important discussions on other components in the Earth system. Attempting to write a paper encompassing all components and discussions during the workshop would require leaving out important details that are paramount for understanding the complexity of energy budgets in ESMs. Also, any attempt at a "unified discussion" of energetics of all components would lead to a paper written at an undesirably high level of mathematical abstraction.

A first step towards a more accurate and comprehensive closure of energy budgets in ESMs is to formulate comprehensive energy equations and diagnose the different terms involved in the energy budget. In the spirit of reaching the broader modeling community, we have chosen to start with well-known large-scale equation sets (and associated assumptions) and then add complexity incrementally. Please note that this paper focuses on documenting possible sources of energy non-conservation, raising awareness about all the problems that may arise, as well as providing suggestions on how to address some of the issues identified. Hence, this paper is by no means an exhaustive analysis of the

overall problem, but rather an attempt at assembling the building blocks of a future, more general, approach.

The paper is structured as follows:

- **Theoretical energetics/budgets (Section 2)**
  First, the total energy equations for traditional (large-scale) climate models are derived. We start with the dry hydrostatic primitive equations (HPE) and gradually increase the thermodynamic complexity by first adding water vapor and then condensates to the HPE. Special attention is given to the derivation of enthalpy terms (and associated reference states), latent heat terms and surface flux terms. For these models, a detailed explanation of the approximations made in large-scale models can be included rigorously. An in-depth discussion is included of surface fluxes and the complications arising due to falling precipitation and/or water entering the atmosphere using a single-component fluid approach.
- **Energy budget errors of a climate model (Section 3)**
  After deriving the theoretical energy budget, we use the NCAR CAM climate model to estimate the magnitude of various terms in the total energy equation. This example highlights which energy terms are important and illustrates why it is difficult to close energy budgets when using the assumptions that are standard for large scale models (e.g. that all components of air have the same temperature).
- **Other energy budget errors (Section 4)**
  In addition to the energy budget errors and imbalances discussed in Section 3, we discuss the following additional energy errors:
  - Numerical truncation energy errors in dynamical cores (adiabatic).
  - Physics–dynamics coupling errors due to spatial and temporal discretization errors. This excludes inconsistencies at the level of the continuous equations.
  - Thermodynamic inconsistency energy errors:
    * As an illustration we discuss a specific example in some detail: coupling the CLUBB cloud parameterization package with the CAM climate model.
    * Mass "clipping" errors and energy
    * Thermodynamic and vertical coordinate inconsistencies between dynamical core and parameterizations
- **Summary and future directions (Section 5)**
  We end by summarizing the various energy budget imbalances discussed in this paper. This provides an opportunity to rank these errors in order of importance, and suggests where advances need to be made to reduce energy imbalances in a large-scale (traditional) climate model setting.
  Since the workshop participants also include experts in cutting edge mathematical areas that could lead to major advances in developing more energetically consistent models, we take the opportunity to provide accessible introductions to these new approaches. These include:
  - **Geometric mechanics formulations for Geophysical Fluid Dynamics (GFD)**
    A framework for deriving energy conserving equation sets beyond the hydrostatic primitive equations is given in Section 5.2.1.
  - **Structure-preserving discretizations**
    Strategies for eliminating or reducing energy discretization errors in the fluid flow solver (i.e. the dynamical core) using mimetic discretizations.
  - **Thermodynamic Potential**
    Section 5.2.3 introduces the concept of a thermodynamic potential in some detail. While the use of thermodynamic potentials is standard in the ocean community, due to the more complex equation of state for seawater (TEOS-10 standard; `http://www.teos-10.org/`), thermodynamic potentials are less commonly

used by the atmospheric community. The thermodynamic potential can be used to ensure thermodynamic consistency in modeling systems.

Several discussions and details have been moved from the main text to Appendices, with the intention of a better flow of the discussions in the main text. These Appendices include comprehensive detail and we hope they will serve as a standard reference for expert readers. The Appendices are:

A. *A note on averaging: unresolved and unrepresented total energy.* This is a huge, complicated and immensely important subject. Although it is not tackled in this paper, we feel it is important to introduce the reader to this subject.
B. *Derivation of the single-fluid energy equation and energy conservation equation in the primitive equation set with a generalized vertical coordinate in the presence of mass sources.* A detailed derivation of local energy equations for a single-fluid is given. Among other things, this Appendix serves the purpose of discussing how the geopotential and internal energy terms are converted into surface geopotential and enthalpy terms. The two pairs can not be conflated. Their relationship to the surface fluxes is also presented.
C. *Enthalpy and energy formulas using different reference states.* The energy formulas can be derived using different reference states and different communities tend to use different reference states. For completeness, we include the energy formulas using all possible reference states. The Appendix is divided into several subsections: (B1) Partial specific enthalpies for different reference states, (B2) Surface enthalpy flux, and (B3) Final energy equations with different reference states.
D. *Ocean enthalpy fluxes.* A brief discussion of the enthalpy boundary condition between atmosphere and ocean as well as their associated complications.
E. *Mechanical dissipation of kinetic energy and resultant heating.* The momentum exchange between the atmosphere and ocean is intriguing since it is not a material energy flux. The ocean and atmosphere exert mechanical work on each other through a stress tensor, but there is no transfer of kinetic energy via a flux. The energetics of that problem is discussed in some detail.
F. *Use of the barycentric velocity as a basis for consistently representing the frictional dissipation of rain in atmospheric models.* Large-scale models have many missing processes associated with water entering and leaving the column. One missing process is the frictional heating of falling precipitation, which can be represented using the barycentric velocity framework. This is not explicitly discussed in the literature and so we add a discussion on this approach in this Appendix.
G. *CAM setup.* Details of the climate model setup we are using in this paper to assess the magnitude of different terms in the energy equation.
H. *Sponge-layer diffusion and frictional heating.* This Appendix supports the discussion of sponge layers and energetics.
I. List of symbols appearing more than once in the main text to help the reader through the many equations.

2 Theoretical energetics and energy budgets

From a modeling perspective an ESM is a collection of components (atmosphere, ocean, sea ice, land surface, ...) that are coupled so that they exchange energy. The upper boundary of an ESM is the top of the atmosphere component and the lower boundary is either the sea floor or the lower boundary of the land component (see Figure 2). From first principles the Earth's total energy budget should be closed, meaning that the change in total energy of the Earth system, $\frac{\partial}{\partial t}\iiint \rho^{(all)}(E_{esm})dV$ is balanced by the flux of total energy through the upper ($\mathcal{F}^{(top)}$) and lower boundary ($\mathcal{F}^{(bottom)}$) of the Earth

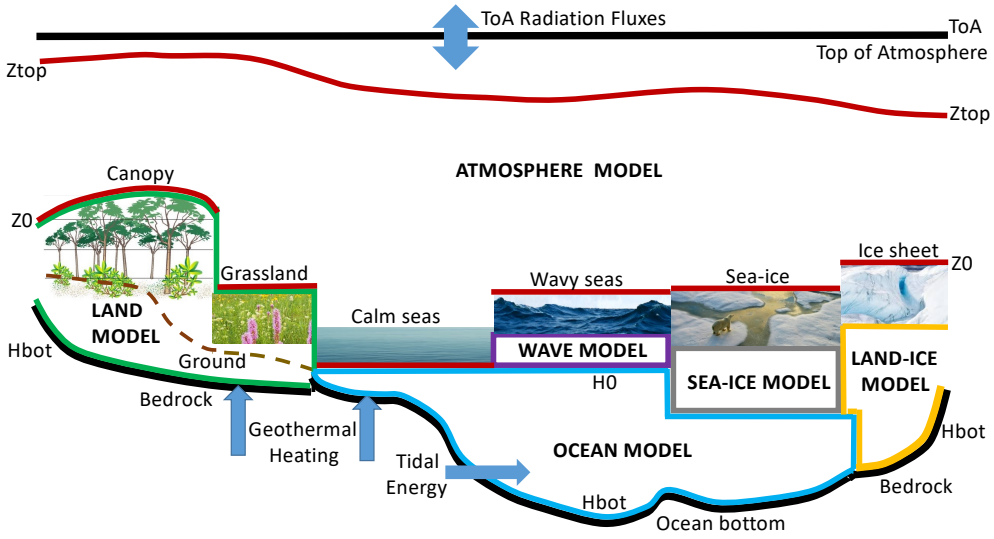


Figure 1. Vertical extents of the modeled Earth System (ToA to $Hbot$, black); and of its component models of the Atmosphere ($Ztop$ to $Z0$, red); the Ocean (surface, $H0$ to bottom, $Hbot$, blue); the Land ($Z0$ above ground to bedrock, $Hbot$); floating Sea-ice (grey); Land-ice (orange) and Surface Waves (purple). The external energy fluxes are Radiation, Geothermal and Tidal.

system, i.e.

$$\frac{\partial}{\partial t}\iiint \rho^{(all)}\left(E_{esm}\right)dV = -\oiint_{top}\mathcal{F}_{esm}^{(top)}dA + \oiint_{bottom}\mathcal{F}_{esm}^{(bottom)}dA. \tag{1}$$

where $\rho^{(all)}$ is the total density (defined in detail later). Here, the direction of the fluxes is defined positive upwards. The subscript esm refers to ESM. The upper boundary flux of energy $\mathcal{F}^{(top)}$ is the net radiative flux at the atmosphere model top. The lower boundary flux will either be the lower boundary flux in the ocean, land or ice component depending on geographical location. The flux could be due to geothermal heating or tidal energy (over ocean).

The total energy of the ESM, E_{esm}, can be divided into atmosphere ($'atm'$), ocean ($'ocn'$), etc. components $\mathcal{C} = ('atm','ocn',...)$

$$\frac{\partial}{\partial t}\iiint \rho^{(all)}\left(E_{esm}\right)dV = \sum_{\ell\in\mathcal{C}}\frac{\partial}{\partial t}\iiint \rho^{(all)}E_\ell\, dV, \tag{2}$$

with energy fluxes exchanged between the components (that cancel out) and outer boundaries of the domain (that do not cancel out). For the atmosphere, the energy budget equation can be written as

$$\frac{\partial}{\partial t}\iiint \rho^{(all)}\left(E_{atm}\right)dV = -\oiint \mathcal{F}_{atm}^{(top)}dA + \oiint \mathcal{F}_{atm}^{(bottom)}dA, \tag{3}$$

where $\mathcal{F}_{atm}^{(bottom)} = \mathcal{F}_{ocn}^{(top)}$ is the energy flux from the ocean into the atmosphere (for example, energy fluxes associated precipitation and evaporation). In this paper we do not consider land and ice components, but their energy fluxes at the interface with the atmosphere are part of equation (3). For the ocean the energy equation can be written as

$$\frac{\partial}{\partial t}\iiint \rho^{(all)}\left(E_{ocn}\right)dV = -\oiint \mathcal{F}_{ocn}^{(top)}dA + \oiint \mathcal{F}_{ocn}^{(bottom)}dA, \tag{4}$$

where as before the upward flux $\mathcal{F}_{ocn}^{(top)} = \mathcal{F}_{atm}^{(bottom)}$. Again, there are additional terms in (4). For example, energy fluxes from the land to the ocean (such as fresh water fluxes) are not discussed here. In most models the fluxes do not exactly match, as discussed in Section 3.

The total energy conserved by the governing equations of motion and associated thermodynamics is referred to as the *fluid equations of motion energy*, E_{feom}. The fluid equations of motion and the thermodynamics are usually approximated. For example, the fluid equations of motion may make the hydrostatic assumption, which means that the total energy of the system E_{atm} includes energy associated with motions neglected in the fluid equations of motion (e.g. non-hydrostatic motion, breaking gravity waves and turbulence). Other simplifications in the fluid equations of motion could include neglecting individual momentum equations for hydrometeors, and making the single temperature assumption, so that all components of moist air have the same temperature. Therefore, the total energy for the atmosphere may be divided into the energy associated with the fluid equations of motion and the energy associated with all motions and processes (such as radiation) not represented in the fluid equations of motion, E_{other}

$$E_{atm} = E_{feom} + E_{other}. \tag{5}$$

Alternatively, one could describe E_{other} as *unrepresented* energy, as it is not represented in the fluid equations of motion and their associated thermodynamics. We obviously do not have an exact expression for E_{other}. However, some parameterizations include approximations for the energies associated with some of the processes that we know must be in the true unrepresented energy E_{other}. As models are run at higher and higher resolutions, and use less approximated equation sets and associated thermodynamics, more and more processes are being represented in the fluid equations of motion energy rather than in the "other" category.

In addition to this prior argument for the continuous equation of motion, there is an even more complex problem. It is not possible to run ESMs at the small scales necessary to resolve all fluid and physical processes. We must therefore homogenize (i.e. average) processes smaller than about 50–100 km in operational climate models, and roughly 0.5–3 km for cutting edge convection-permitting global models. In particular, the smallest energetic turbulent scales of motion are $O(10^{-3})$ m, which means that the energy E will always have both a resolved and an unresolved component

$$E_{atm} = E^{(res)} + E^{(unres)}. \tag{6}$$

Things now become complicated and less well understood. Since this topic, though immensely important, is not the main focus of this paper, a (brief) discussion of the treatment of resolved and unresolved scales is provided in Appendix A. Henceforth, we will focus mainly on the unaveraged energy equations and only, when necessary, distinguish between resolved and unresolved scales (e.g. when discussing frictional heating and subgrid-scale energy reservoirs).

2.1 Unaveraged total energy equations: Atmosphere E_{feom}

The energy E_{feom} associated with the governing equations of motion is the energy of a fluid subject to gravity and rotation:

$$E_{feom} = K + U + \Phi. \tag{7}$$

The various equation sets discussed below have different expressions for kinetic energy K, internal energy U and geopotential Φ, which depend on the particular assumptions made and the choice of thermodynamics. Commonly made assumptions for geophysical fluid models are both geometric (traditional shallow-atmosphere, spherical-geoid) and

dynamical (quasi-hydrostatic, Boussinesq). In what follows, our emphasis will be almost entirely on the form of internal energy U and the associated thermodynamics, since this is usually the place where significant approximations are made. It is also where inconsistency is easily introduced. An important assumption we will always make is that of *local thermodynamic equilibrium*, which ensures that mean fluid quantities such as temperature and pressure are well-defined.

It is important to note here that in this section we discuss the total (conserved) energy for a given set of assumptions (e.g. hydrostatic, shallow atmosphere) for the fluid part of the atmosphere. If we consider only reversible dynamics, total energy in fact defines the associated processes that occur in the fluid (through the Lagrangian/Hamiltonian formulation). However, this same energy is also conserved if we introduce irreversible dynamics, which is a new set of processes, but with the same energy. For example, both the compressible Euler (inviscid) and compressible Navier–Stokes–Fourier equations (viscous) conserve the same total energy. However, the Navier–Stokes–Fourier equations have additional terms describing the irreversible (viscous) processes. This distinction is important since we are interested in physics-dynamics coupling, which necessarily involves both reversible and irreversible dynamics.

A formal general derivation of the energy conservation equations for the primitive equations is given in Appendix B. Here we take a more constructive approach, for didactical purposes and for application to special cases.

We start by defining the energy for the atmosphere component E_{feom} for commonly used equation sets. The energy formulae are for the non-averaged equations applying to all scales.

2.1.1 Total energy of the dry primitive equations

Many ESMs make use of the *primitive* equations on the sphere (essentially the Navier–Stokes equations in the hydrostatic and shallow atmosphere limits). We refer to this equation set as the hydrostatic primitive equations (HPE). The corresponding total energy is well-established in text books and the literature

$$E_{feom} = K + U + \Phi, \tag{8}$$

(e.g. Kasahara, 1974) where K is *horizontal* specific kinetic energy ($\frac{1}{2}\vec{v}\cdot\vec{v}$ where $\vec{v}$ is the horizontal velocity), U is the specific internal energy and Φ is the geopotential. Specific total energy $K + U + \Phi$ differs from that of a non-hydrostatic flow (obeying the standard Euler or Navier–Stokes equations) only in that vertical kinetic energy does not contribute. In the commonly used shallow atmosphere approximation $\Phi = gz$ where g is the (constant) gravitational acceleration. Note that there is an arbitrary choice of reference for the geopotential. The common reference for z is mean sea-level so that z is elevation above sea level.

To begin, let us consider a single phase, single component fluid, i.e. the dry atmosphere. Dry air is actually a mixture of many different gases, but it is assumed to have constant concentrations of each component, same temperature and velocity and thus is referred to as a single-component fluid.

For an ideal perfect gas the internal energy (up to a constant reference discussed later) is given by

$$U^{(d)} = c_v^{(d)} T, \qquad \text{(dry air = ideal perfect gas)}, \tag{9}$$

where $c_v^{(d)}$ is the specific heat at constant volume for dry air. We note that the terminology "specific heat capacity" in the physics literature is typically used to indicate an extensive quantity (amount of heat energy required by a substance to raise its temperature by $1K$) whereas "specific heat" is an intensive quantity (amount of heat energy

required by a unit of mass of a substance to raise its temperature $1K$); the meteorological literature is not always consistent with this terminology.

The energetics of the dry primitive equations should be unambiguous. However, we will highlight several sources of widespread ambiguities that arise even for these simple equations.

Ambiguity 1: It should be noted that specifying how U depends on T does not fully determine the thermodynamics used by the atmospheric model. In general, U is not a function of temperature alone. However, even when it is, the relationship $U = U(T)$ must be complemented by extra information, such as the perfect gas law

$$p = \rho^{(d)} R^{(d)} T, \tag{10}$$

where $R^{(d)} \equiv k_B/\mathcal{M}^{(d)}$ (k_B is the Boltzmann constant, $\mathcal{M}^{(d)}$ is the molar mass of dry air). This can be avoided if (specific) internal energy is regarded as a function of (specific) volume and (specific) entropy, in which case it is a thermodynamic potential that contains all information needed for the thermodynamics of the atmospheric fluid. This approach is developed further in subsection 5.2.3.

Ambiguity 2: Depending on the type of boundaries and the boundary conditions, total energy might not be conserved. Typically, in the atmosphere we assume a rigid lower boundary and either a rigid (constant-height) or a pressure top upper boundary. In the case of a rigid upper boundary, total energy will be conserved up to boundary fluxes, with the following expression

$$\frac{\partial}{\partial t} \iiint \left[K + c_v^{(d)} T + \Phi \right] \rho^{(d)} dA\, dz \quad = \quad 0,\ z_t \text{ constant} \tag{11}$$

(Kasahara, 1974) where $dA = r^2 \cos\varphi d\lambda d\varphi$ is an infinitesimal surface area element, z_t is the height at the model top and $\rho^{(d)}$ is the mass of dry air per unit volume of air (i.e. density). However, in the case of a pressure top, total energy is not conserved, even in the case of no boundary fluxes. Instead, the fluid exchanges energy with the environment through the motion of the upper surface. There is a conserved quantity (up to boundary fluxes), sometimes referred to as the total enthalpy, given by

$$\frac{\partial}{\partial t} \iiint \left[K + c_v^{(d)} T + \Phi \right] \rho^{(d)} dA\, dz + \frac{1}{g} \frac{\partial}{\partial t} \iint p_t \Phi_t dA \quad = \quad 0,\ p_t \text{ constant} \tag{12}$$

(Kasahara, 1974), where $p_t^{(d)}$ is the (dry) pressure at the model top and Φ_t is the geopotential height at the model top. The additional term accounts for the work done by the moving pressure top. Finally, (12) can be integrated by parts to yield

$$\frac{\partial}{\partial t} \iiint \left[K + c_p^{(d)} T + \Phi_s \right] dA\, \frac{dp^{(d)}}{g} \quad = \quad 0,\ p_t \text{ constant.} \tag{13}$$

where surface geopotential is $\Phi_s = g z_s$ for shallow atmosphere approximation, z_s is surface height, and $c_p^{(d)}$ is the specific heat at constant pressure for dry air. For a detailed derivation and discussion on the derivation of (13) see Appendix B1.

Apart from differences in the top boundary condition, a widespread ambiguity is that it is tempting to regard $K + c_p^{(d)} T + \Phi_s$ as total energy per unit mass in (13). This is incorrect for several reasons:

- The definition of energy per unit mass, or local energy, must include a term for internal energy, but specific enthalpy, which in the particular formulation above equals
$$h^{(d)} = c_p^{(d)} T, \text{ (enthalpy; ideal gas)} \tag{14}$$
is not a substitute for the internal energy. More generally,
$$h^{(d)} = U^{(d)} + \alpha^{(d)} p^{(d)}, \tag{15}$$

where $\alpha^{(d)}$ is the specific volume of dry air.

- The transformation that allows one to use $h^{(d)}$ instead of $U^{(d)}$ under the integral in (13) is possible only for the HPE with specific boundary conditions. Modifying the boundary conditions or the specific assumptions underlying the HPE invalidates this transformation. For instance, this transformation is not valid for deep-atmosphere equations, or shallow-atmosphere based non-traditional quasi-hydrostatic equations. However, depending on the upper boundary condition used either (11) or (12) will hold.

We will largely ignore these complications, however, and use (13) as our total energy expression in this section.

Ambiguity 3: The energy flux is NOT proportional to the energy density! For tracer transport, for example, a material quantity ψ is transported, and the flux of ψ is equal to ψ times mass flux so that the transport equation reads: $\frac{\partial}{\partial t}(\rho\psi) = -\nabla\cdot(\vec{v}\rho\psi)$. It is tempting to do the same with specific energy E_{feom}. This is, however, incorrect. The local energy density

$$\rho^{(d)}\left[K + c_v^{(d)}T + \Phi\right], \text{ (total energy density)} \tag{16}$$

is not modified only by its flux, but also by the work exerted by pressure forces. The flux of energy therefore is not $\left[K + c_v^{(d)}T + \Phi\right]$ times the mass flux. The flux divergence is

$$\nabla\cdot\left\{\vec{v}\rho^{(d)}\left[K + c_p^{(d)}T + \Phi\right]\right\}, \text{ (kinetic, geopotential, enthalpy flux)} \tag{17}$$

(see, e.g., the total energy equations for various vertical coordinates (5.8), (5.15), (5.18) that all have the same energy flux term in Kasahara, 1974). A manifestation of this point is that waves can transfer energy without transferring mass. These relations are completely general and do not hold just for the primitive equations, but also for non-hydrostatic or deep-atmosphere flows for example. Also, note that $K + c_p^{(d)}T + \Phi$ is the Bernoulli function; without the specific kinetic energy term the Bernoulli function is the dry static energy!

Ambiguity 4: When using a pressure-based vertical coordinate system where we have just pointed out that the integrand in the energy equation (13) is not local energy density, one may ask why most physics packages are formulated in terms of constant pressure and, if conserving energy, use (13)? The reason is that enthalpy is commonly used to redefine the first law of thermodynamics:

$$dU^{(d)} = -p^{(d)}d\alpha^{(d)} + dQ, \text{ (1st law of thermodynamics)} \tag{18}$$

where Q is heating per unit mass. From (15) and substituting (18) it is observed that for enthalpy we have

$$dh^{(d)} = dU^{(d)} + d\left(\alpha^{(d)}p^{(d)}\right) = dU^{(d)} + \alpha^{(d)}d\left(p^{(d)}\right) + p^{(d)}d\left(\alpha^{(d)}\right) = \alpha^{(d)}d\left(p^{(d)}\right) + dQ. \tag{19}$$

Hence we see that when pressure is constant and there is no heating $dQ = 0$ then enthalpy is conserved (Vallis, 2006). This simplifies conservation of energy during phase transitions (once the definition of enthalpy is extended to include water). Since the total energy in the formulation (13) includes only boundary terms that do not change under the constant pressure assumption, using enthalpy as the conserved thermodynamic variable automatically conserves total energy of the model.

Aside: Using (18) the prognostic thermodynamic equation used in models can be written in an "internal energy form" (heating under constant volume)

$$\frac{DU^{(d)}}{Dt} = \left.\frac{DQ}{Dt}\right|_{c_v^{(d)}T} - p^{(d)}\alpha^{(d)}\left(\nabla\cdot\vec{v} + \frac{\partial w}{\partial z}\right), \text{ (constant volume)} \tag{20}$$

where we have used that

$$\begin{aligned}\frac{D\alpha^{(d)}}{Dt} &= \frac{D}{Dt}\left(\frac{1}{\rho^{(d)}}\right), \\ &= -\frac{1}{\left(\rho^{(d)}\right)^2}\frac{D\rho^{(d)}}{Dt}, \\ &= \frac{1}{\rho^{(d)}}\left(\nabla\cdot\vec{v}+\frac{\partial w}{\partial z}\right), \text{ since } \frac{D\rho^{(d)}}{Dt} = -\rho^{(d)}\left(\nabla\cdot\vec{v}+\frac{\partial w}{\partial z}\right), \\ &= \alpha^{(d)}\left(\nabla\cdot\vec{v}+\frac{\partial w}{\partial z}\right),\end{aligned}$$

and where w is vertical velocity and $D/Dt = \partial/\partial t + \vec{v}\cdot\nabla + w\frac{\partial}{\partial z}$. Or the prognostic thermodynamic equation can be written in an "enthalpy form" (heating under constant pressure)

$$\frac{Dh^{(d)}}{Dt} = \left.\frac{DQ}{Dt}\right|_{c_p^{(d)}T} + \alpha^{(d)}\omega^{(d)}, \text{ (constant pressure)} \tag{21}$$

using (19) and $\omega^{(d)} \equiv Dp^{(d)}/Dt$.

Henceforth, unless otherwise stated, we will assume a pressure-based vertical coordinate, η, with the common assumption of constant pressure at the top boundary.

First, we remark that with a pressure-based vertical coordinate, the energy equation can be written more generally in terms of specific enthalpy, $U^{(d)}+p^{(d)}/\rho^{(d)}$, which for an ideal perfect gas (up to a constant reference) is

$$h^{(d)} \equiv U^{(d)} + p^{(d)}/\rho^{(d)} = c_p^{(d)}T, \qquad \text{(ideal gas)}. \tag{22}$$

i.e. (13) can be written as

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}E_{feom}dV = \frac{\partial}{\partial t}\iiint \rho^{(d)}\left(K + h^{(d)} + \Phi_s\right)dA\,dz, \tag{23}$$

where we integrate with respect to $\rho^{(d)}dz$ despite using a pressure-based vertical coordinate. The reason for using $\rho^{(d)}dz$ is to avoid the ambiguity of using p as a mass integration variable. In addition to the differences associated with the hydrostatic and non-hydrostatic approximations, p could either include or exclude all water, or all condensed water, or precipitating water etc. It seems better to avoid that ambiguity and extra complication by explicitly integrating over vertical length times all the relevant densities.

Large-scale models will include, at least, the gas phase of water (i.e. water vapor) in its equations of motion and thermodynamics in addition to dry air. Many readers will likely view this as a solved or trivial problem. However, this seemingly modest increase in complexity leads to significant energy inconsistencies in many current large-scale models (due to using the heat capacity for dry air also for water vapor).

2.1.2 Total energy of the primitive equations for moist gas (dry air and water vapor)

Most large-scale models make the following assumptions:

- **Temperature assumption**: Assumes that all constituents of the fluid (in this case, dry air and water vapor) have the same temperature.
- **Momentum assumption**: Assume that dry air and water vapor move with the same horizontal velocity (at this point we are not considering phase changes, precipitation and evaporation, so we may call $\vec{v}$ the *gaseous* velocity vector).

With these assumptions the addition of water vapor does not modify the general expression $K + U + \Phi$ for total energy, but it does lead to two changes. One is the addition of water vapor mass to the mass of the atmosphere and the other is the thermodynamic effect. These will be discussed in the following two paragraphs.

Mass effect: Density now includes the effect of water vapor so that

$$\rho^{(all)} = \sum_{\ell \in \mathcal{L}_{all}} \rho^{(\ell)} = \rho^{(d)} \left(\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \right), \tag{24}$$

where $\mathcal{L}_{all}$ is the set of species included in moist air (in this case dry air and water vapor) and $m^{(\ell)}$ is the dry mixing ratio defined by

$$m^{(\ell)} \equiv \frac{\rho^{(\ell)}}{\rho^{(d)}}, \tag{25}$$

where $\rho^{(\ell)}$ is the mass of species ℓ per unit volume of moist air. Note that $m^{(d)} \equiv 1$. The mass of air now includes contributions from all species in moist air. If $M^{(\ell)}(z)$ is the mass of species ℓ in a column of air above height z

$$M^{(\ell)}(z) = \int_{z'=z}^{z'=\infty} \rho^{(d)}\, m^{(\ell)}\, dz', \tag{26}$$

then the hydrostatic pressure at height z is

$$p(z) = g \sum_{\ell \in \mathcal{L}_{all}} M^{(\ell)}(z). \tag{27}$$

Equations (24), (26) and (27) are completely general and not just applicable to moist air consisting of dry air and water vapor.

It is noted that one usually formulates the above equations using a moist basis in terms of specific contents

$$q^{(\ell)} \equiv \frac{\rho^{(\ell)}}{\rho^{(all)}}. \tag{28}$$

If we assume that dry air and water vapor move with the same velocity the advantage of using a dry basis instead of a moist basis is that the mixing ratios do not need to be re-computed when, for example, water leaves the air parcel (precipitation). In this case ρ changes so $q^{(\ell)}$ must be recomputed. On the other hand, $\rho^{(d)}$ remains constant so $m^{(\ell)}$ does not need to be re-computed. This is discussed at length in Lauritzen et al. (2018). However, note that if using a barycentric velocity basis, the prognostic continuity equation for dry air $\rho^{(d)}$ will have additional source/sink terms based on the diffusive precipitation and turbulent mixing fluxes (discussed in Appendix F2 as well as Wacker & Herbert, 2003). These terms are, however, small.

Thermodynamic effect: The second effect of water vapor is to modify the equation of state $p = p(T, \rho^{(d)}, \rho^{(wv)})$ and, more generally, all thermodynamic relationships. A simple way to take water vapor into account is to regard dry air and water vapor as ideal perfect gases, and moist air as an ideal mixture of them. From first principles that:

- The perfect gas law holds in the form

$$p = \rho^{(d)} R^{(d)} T + \rho^{(wv)} R^{(wv)} T = \rho^{(all)} R^{(all)} T, \tag{29}$$

  where

$$R^{(all)} = \frac{\sum_{\ell \in \mathcal{L}_{all}} R^{(\ell)} m^{(\ell)}}{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}} \tag{30}$$

with $R^{(d)} \equiv k_B/\mathcal{M}^{(d)}$ and $R^{(wv)} \equiv k_B/\mathcal{M}^{(wv)}$, respectively ($k_B$ is the Boltzmann constant, $\mathcal{M}^{(\ell)}$ is the molar mass of gas component ℓ, where $\ell = d$ (dry air) or wv (water vapor). For non-gasses (i.e. condensates) $R^{(\ell)}$ is zero.

- The perfect gas law is often rewritten in the form

$$p = \rho^{(all)} R^{(d)} T_v, \tag{31}$$

with virtual temperature T_v defined as

$$T_v \equiv T \left(\frac{1 + \frac{R^{(wv)}}{R^{(d)}} m^{(wv)}}{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}} \right). \tag{32}$$

- Specific heat at constant pressure:

$$c_p^{(all)} = \frac{\sum_{\ell \in \mathcal{L}_{all}} c_p^{(\ell)} m^{(\ell)}}{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}}. \tag{33}$$

- Internal energy per unit mass $U^{(all)} = c_v^{(all)} T$ where

$$c_v^{(all)} = \frac{\sum_{\ell \in \mathcal{L}_{all}} c_v^{(\ell)} m^{(\ell)}}{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}}. \tag{34}$$

As a result, the primitive equations conserve the total energy (we reiterate that the integrand of (35) is *not* total specific energy)

$$\iiint \rho^{(all)} \left(E_{feom}\right) dA\, dz = \sum_{\ell \in \mathcal{L}_{gas}} \iiint \rho^{(d)} \left\{ m^{(\ell)} \left(K + c_p^{(\ell)} T + \Phi_s \right) \right\} dA\, dz, \tag{35}$$

where $\mathcal{L}_{gas}$ is the set of gaseous components of air (dry air and water vapor)

$$\mathcal{L}_{gas} \equiv \{d, wv\}, \tag{36}$$

The above expression in (35) decomposes the total energy into contributions from each component of moist air treated as an ideal mixture of perfect gases and condensate. For real moist air, e.g., Hermann et al. (2009) and Feistel et al. (2010), the expressions $c_p^{(\ell)} T$ would have to be replaced by partial enthalpies, which in general depend not only on temperature but also on pressure and composition.

The effect of water vapor is not always included in the specific heat at constant pressure, i.e. some models keep $c_p^{(\ell)} = c_p^{(d)}$. It is possible to derive energy-consistent primitive equations with that assumption. Although, $c_p^{(wv)}$ is about twice the dry air value of $c_p^{(d)}$ (see, e.g., Appendix G), the majority of moist air is dry air and not water vapor, and hence the assumption $c_p^{(\ell)} = c_p^{(d)}$ is not too severe. Many physics packages and dynamical cores use this simpler enthalpy formulation (22), in which case the total energy is

$$\iiint \rho^{(all)} \left(E_{feom}\right) dA\, dz \quad = \quad \iiint \rho^{(all)} \left[\left(K + c_p^{(d)} T + \Phi_s \right) \right] dA\, dz. \tag{37}$$

(this is the "$c_p^{(d)}$ total energy equation" used in many large-scale models)

While the above practice involves a slight loss of physical accuracy, it is energetically consistent since the equations conserve an energy. Note that we are considering a mixture of dry air and water vapor only, so there are no latent heat terms. When including condensates in moist air (and thereby phase transformations) the latent heat terms appear in the energy equation as discussed in detail in the next section. The energy inconsistency error occurring if dynamics uses (35) and physics (37) is discussed in section 4.5.1.

2.1.3 Total energy of the primitive equations for moist gas with condensates (dry air and water vapor and condensates)

Now let's add condensates to the mix! Condensates are not gases, so their inclusion in the thermodynamics is not straightforward. A series of assumptions are made in today's ESMs to keep the problem more tractable, rather than using a full multi-component fluid formulation (Bannon, 2002).

We extend the assumptions from the previous section to air that contains condensates:

- **Temperature assumption**: Assume that all constituents of the fluid have the same temperature.
- **Zero volume assumption**: Assume the condensates do not take up any volume.
- **Momentum assumption**: Assume that all components of moist air move with the same horizontal velocity $\vec{v}$.

The hydrostatic mass effect of condensates is straightforward. As mentioned in Section 2.1.2, the equations for total density (24), species masses ℓ (26) and hydrostatic pressure (27) are the same when including condensates. Typical condensates included in ESMs are cloud liquid cl, cloud ice ci, rain rn, snow sw and graupel gr. In this case the components of moist air in equations for $\mathcal{L}_{all}$ is given by

$$\mathcal{L}_{all} \equiv \{d, wv, cl, ci, rn, sw, gr\}\,. \tag{38}$$

Since $\mathcal{L}_{all}$ contains both gaseous and non-gaseous (condensates) we define the set of condensates only

$$\mathcal{L}_{cond} \equiv \{cl, ci, rn, sw, gr\}\,, \tag{39}$$

and it is also convenient to define the set of all water substances

$$\mathcal{L}_{H_2O} \equiv \{wv, cl, ci, rn, sw, gr\}\,. \tag{40}$$

The specific heat of liquid forms of water are assumed the same

$$c_p^{(cl)} = c_p^{(rn)} \equiv c_p^{(liq)}, \tag{41}$$

and similarly for frozen forms of water (cloud ice, graupel and snow)

$$c_p^{(ci)} = c_p^{(sn)} = c_p^{(gr)} \equiv c_p^{(ice)}. \tag{42}$$

For notational convenience we refer to the density of liquid forms of water as $\rho^{(liq)} = \rho^{(cl)} + \rho^{(rn)}$, and similarly for frozen forms of water $\rho^{(ice)} = \rho^{(ci)} + \rho^{(sn)} + \rho^{(gr)}$. All mixing ratios represent mass of water species with respect to the mass of dry air. If the model's microphysics scheme requires more forms of water they can easily be added as long as they are characterized correctly in terms of liquid and ice. For notational purposes it is convenient to denote

$$\rho^{(H_2O)} \equiv \rho^{(d)} \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}, \tag{43}$$

$$\rho^{(all)} \equiv \rho^{(d)} \sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}, \tag{44}$$

and similarly for the mixing ratios

$$m^{(H_2O)} \equiv \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}, \tag{45}$$

$$m^{(all)} \equiv \sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}. \tag{46}$$

Please note that even if we assume condensates to be incompressible, the partial pressures of dry air and water vapor are both affected by the weight (force exerted by the matter when it is in a gravitational field) of the condensates. We see this by integrating the hydrostatic balance equation

$$\frac{dp}{dz} = -g\rho^{(all)} = -g\rho^{(d)} \sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \tag{47}$$

(where $\rho^{(all)}$ includes condensates) and using Dalton's law of partial pressures to observe that

$$p^{(d)}(z) + p^{(wv)}(z) = g \sum_{\ell \in \mathcal{L}_{all}} M^{(\ell)}(z), \tag{48}$$

where $M^{(\ell)}$ is mass per unit area for species ℓ defined in (26). When there are no condensates present then

$$p^{(d)}(z) = gM^{(d)}(z) \qquad \text{and} \qquad p^{(wv)}(z) = gM^{(wv)}(z), \qquad \text{with } \mathcal{L}_{cond} \equiv \varnothing. \tag{49}$$

In the presence of condensates we can no longer pair terms on the left and right-hand side of (48) and separate the pressure into a dry and moist part (see, e.g., detailed discussion in Section 2.4.2 in Lauritzen et al., 2018).

The perfect gas law and virtual temperature take the same form as presented in (31) and (32), respectively, if one assumes that the physical volume occupied by condensates is negligible (for details on the validity of this approxiation see equations 1.71-3 in Staniforth et al., 2006). The virtual temperature is often approximated further

$$T_v = T\left(\frac{1 + \frac{R^{(wv)}}{R^{(d)}} m^{(wv)}}{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}}\right) \simeq T\left(\frac{1 + \frac{R^{(wv)}}{R^{(d)}} m^{(wv)}}{1 + m^{(wv)}}\right) = T\left[1 + \left(\frac{R^{(wv)}}{R^{(d)}} - 1\right) q^{(wv)}\right] \tag{50}$$

which is exact provided that moist air is only composed of dry air and water vapor (i.e. condensates are not thermodynamically active). To derive the right hand side of (50) we have used

$$\frac{1}{1 + m^{(wv)}} = \frac{1 + m^{(wv)} - m^{(wv)}}{1 + m^{(wv)}} = 1 - \frac{m^{(wv)}}{1 + m^{(wv)}} \simeq 1 - q^{(wv)} \tag{51}$$

which, again, assumes that total water is water vapor only.

However, if this expression for T_v is used in the perfect gas law $p = \rho^{(all)} R^{(d)} T_v$ then the resulting thermodynamics can be inconsistent, depending on how the equations are used. For example, diagnosing p from $p = \rho^{(all)} R^{(d)} T_v$ using an approximated T_v is, in general, inconsistent. A general framework for making consistent approximations is described in Section 5.2.3.

Relaxing the momentum and temperature assumption (Section 2.1.2) leads to a significant increase in complexity. However, if we do make the momentum and single temperature assumptions, the total specific energy takes the same form, $K + U^{(all)} + \Phi$, but now condensates contribute to $\rho^{(all)}$, the internal energy and enthalpy.

The specific enthalpy of an air constituent ℓ (we may call it a partial enthalpy) can be written in the form

$$h^{(\ell)}(T) = h_{00}^{(\ell)} + c_p^{(\ell)}(T - T_{00}), \tag{52}$$

where $h_{00}^{(\ell)}$ is a reference enthalpy usually chosen to be the specific enthalpy of air component ℓ at $T_{00} = 273.15K = 0^oC$ (it would be equally valid to use the triple point of water which is $T_{00} = 273.16273K$). Note that for ocean models the complete expression for enthalpy is a nonlinear function of pressure, temperature and salinity rather than

a simple linear function of temperature as used in atmosphere models, (52). The enthalpy of moist air can be written as

$$\rho^{(all)}U^{(all)} + p = \sum_{\ell \in \mathcal{L}_{all}} \rho^{(\ell)} \left[h_{00}^{(\ell)} + c_p^{(\ell)}(T - T_{00}) \right]. \tag{53}$$

Expanding the sum on the right-hand side of (53) we get

$$\begin{aligned} \rho^{(all)}U^{(all)} + p &= \rho^{(d)} c_p^{(d)} T + \rho^{(d)} \left(h_{00}^{(d)} - c_p^{(d)} T_{00} \right) + \rho^{(wv)} \left[h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00}) \right] + \\ & \rho^{(liq)} \left[h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}) \right] + \rho^{(ice)} \left[h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}) \right]. \end{aligned} \tag{54}$$

Only enthalpy changes have physical meaning. Hence, we rewrite (54) using Kirchhoff's equations for latent heats (note that latent heat release is when the heat is liberated to become sensible heat; here we are only talking about the physical measures and not yet the physical effect).

Latent heat of vaporization (liquid → water vapor):

$$L_v(T) = L_{v,00} + \left(c_p^{(wv)} - c_p^{(liq)} \right) (T - T_{00}) \text{, where } L_{v,00} \equiv h_{00}^{(wv)} - h_{00}^{(liq)}. \tag{55}$$

Latent heat of sublimation (solid → water vapor):

$$L_s(T) = L_{s,00} + \left(c_p^{(wv)} - c_p^{(ice)} \right) (T - T_{00}) \text{, where } L_{s,00} \equiv h_{00}^{(wv)} - h_{00}^{(ice)}. \tag{56}$$

Latent heat of fusion (solid → liquid):

$$L_f(T) = L_{f,00} + \left(c_p^{(liq)} - c_p^{(ice)} \right) (T - T_{00}) \text{, where } L_{f,00} \equiv h_{00}^{(liq)} - h_{00}^{(ice)}, \tag{57}$$

(see, e.g., p.114-5 in Emanuel, 1994). Note that the latent heat of fusion, $L_f(T)$, may also be written in terms of latent heat of vaporization and sublimation

$$L_f(T) = L_s(T) - L_v(T). \tag{58}$$

In (54), there is some liberty in choosing the reference state in terms of reference enthalpy $h_{00}^{(\ell)}$ and reference temperature T_{00} (we highlight this dependency on the left-hand side of the latent heat formulae in (55), (56) and (57)). That said, the constants entering the definitions of the specific enthalpies of the different phases of water are not independent of each other but must be such that $h^{(wv)}(T) - h^{(liq)}(T) = L_v(T)$ and $h^{(liq)}(T) - h^{(ice)}(T) = L_f(T)$ as is shown next.

If we use the enthalpy of ice $h_{00}^{(ice)}$ as our reference enthalpy then the enthalpy of ice is simply

$$h^{(ice)}(T) = h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}), \tag{59}$$

whereas the enthalpy of liquid, using (57), can be rewritten as

$$\begin{aligned} h^{(liq)}(T) &= h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}) - L_f(T) + L_f(T), \\ &= h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}) - \left[L_{f,00} + \left(c_p^{(liq)} - c_p^{(ice)} \right) (T - T_{00}) \right] + L_f(T), \\ &= h_{00}^{(liq)} + c_p^{(ice)}(T - T_{00}) - \left[h_{00}^{(liq)} - h_{00}^{(ice)} \right] + L_f(T), \\ &= h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}) + L_f(T), \end{aligned} \tag{60}$$

and similarly for the enthalpy of water vapor using the latent heat of sublimation

$$\begin{aligned} h^{(wv)}(T) &= h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00}) - L_s(T) + L_s(T), \\ &= h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}) + L_s(T) \end{aligned} \tag{61}$$

The specific enthalpies for a liquid and water vapor reference states are given in Appendix C1. Substituting (59), (60) and (61) into the total enthalpy formula (54) and rearranging terms yields

$$\begin{aligned}\rho^{(all)}U^{(all)} + p \quad &= \quad \rho^{(d)}c_p^{(d)}T + \rho^{(d)}\left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right) + \rho^{(H_2O)}\left[h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00})\right] \\ &\quad + \rho^{(wv)}L_s(T) + \rho^{(liq)}L_f(T),\end{aligned}$$

(ice reference state)

(see Appendix C1 for other reference states). As a result, the primitive equations, without fluxes through the top and bottom boundaries, conserve the total (globally integrated) energy

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}E_{feom}\,dA\,dz = 0, \tag{62}$$

where

$$\begin{aligned}E_{feom} = \quad & m^{(all)}\left(K + \Phi_s\right) + \left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right) + c_p^{(d)}T \\ & + m^{(H_2O)}\left[h_{00}^{(ice)} + c_p^{(ice)}\left(T - T_{00}\right)\right] + m^{(wv)}L_s(T) + m^{(liq)}L_f(T),\end{aligned} \tag{63}$$

(ice reference state)

for an ice reference state (see formulas for E_{feom} for other reference states in Appendix C1). We reiterate that E_{feom} is *not* total specific energy, i.e. E_{feom} must be integrated over the entire domain to constitute a global energy equation.

The dry air constant term $\left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right)$ on the right-hand side of (63) integrates to a constant globally if dry air mass is conserved and hence the time change of that term is zero. This term contributes an unavailable energy to the total energy and can thus be omitted in the total energy integral (as long as dry air is conserved). For notational clarity we will selectively include this term knowing that it is zero in the energy equation.

We also note that if we expand the latent heat terms on the right-hand side of (63) using (56) and (57) and re-arrange terms, (63) can be written as

$$\begin{aligned}E_{feom} &= \left(K + \Phi_s\right) + c_p^{(d)}T \\ &+ \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}\left[K + \Phi_s + c_p^{(\ell)}\left(T - T_{00}\right) + h_{00}^{(ice)}\right] + m^{(wv)}L_{s,00} + m^{(liq)}L_{f,00}.\end{aligned} \tag{64}$$

(ice reference state, $h_{00}^{(d)} - c_p^{(d)}T_{00} \equiv 0$)

that separates into a sum over all forms of water. Writing the energy equation in this form makes it clear that using variable latent heats leads to using the correct specific heat for each water species in the enthalpy terms, $c_p^{(\ell)}T$. This is not obvious in (63), where the enthalpy term uses the same specific heat based on reference state and the correct specific heat for each water species is hidden in the variable latent heat terms $L(T)$.

Perhaps a natural choice of reference enthalpy is $h_{00}^{(ice)} = 0$. While one can arbitrarily choose reference states and temperatures, they must be consistent between model components (unless accounted for in the model coupler) or if one compares energetics of two models (Mayer et al., 2017). Oceanographers have a standard (TEOS-10; `http://www.teos-10.org/`) where the thermodynamic standard for seawater defines the reference constants for the Gibbs function of seawater. It would make the most sense to define the reference constants for the atmospheric partial enthalpies to be consistent with the TEOS-10 choice of constants. That is, however, easier said than done as the liquid water enthalpy used by the Gibbs function strongly depends on pressure and is nonlinear in temperature (which is not the case for the atmosphere enthalpy formulation used here). Some discussion of this issue can be found in Appendix D. It is beyond the scope of this paper to enforce such consistency across components, although progress on that front does exist in the literature (Feistel et al., 2008).

2.1.4 Enthalpy flux at the surface

We have, so far, not considered energy fluxes into and out of the atmosphere, in particular, at the surface. Given (17) (cf also B18–B20), the total energy fluxes at the surface are divided into

$$F^{(h)}_{net} + F^{(\Phi)}_{net} + F^{(K)}_{net} \text{ (total energy flux at the surface)}, \tag{65}$$

which is the enthalpy, geopotential and kinetic energy fluxes, respectively. Each energy flux at the surface will be discussed in separate subsections starting with the enthalpy flux associated with mass transfer at the surface (this subsection).

Atmospheric models lose water vapor (dew, rime), liquid (rain) or ice water (snow, hail, ...) to the land, land-ice, sea-ice and the ocean. Similarly, the atmosphere receives water vapor (through evaporation and sublimation), snow (through snow drift, except from the ocean) and other forms of water from the land, land-ice, sea-ice or ocean surface. Not all of these processes are included in ESMs (usually just evaporation from the surface and precipitation in either frozen or liquid form from the atmosphere).

When, for example, water evaporates from the ocean the atmosphere gains energy (and mass) which is compensated for by ocean cooling due to the latent heat flux. Hence, this process occurs without any net change in the total energy of the coupled system. In addition to the latent heat flux there is also an enthalpy flux associated with the mass transfer between components that depends on the temperature of water. For evaporation, this temperature can naturally be assumed to be that of the ocean or land surface. However, for falling water the temperature at landfall is harder to compute (since falling precipitation is not prognostic in our current framework). More on this later.

For a mathematical description of the surface enthalpy flux exchange, we distinguish between enthalpy transfer from the atmosphere to the surface, denoted with subscript $a \to s$, and from surface to atmosphere, $s \to a$. Similarly, let the temperature of the atmosphere at the surface be denoted $T_{atm,s}$ and that of the surface itself be $T_{surf,s}$. Although there is a thin layer at the surface for which $T_{surf,s} = T_{atm,s}$ we distinguish between the two, since models usually operate with layer mean temperatures, in which case, the surface layer temperatures will be different between components.

The enthalpy flux from atmosphere to surface depends on the form the water has before it hits the surface. This is also true for enthalpy fluxes from surface to atmosphere (although this is simpler, as discussed in the previous paragraph). Assuming local thermodynamic equilibrium, the flux from the atmosphere to the surface is the mass-flux of each species $F^{(\ell)}_{a\to s}$ multiplied by its respective enthalpy at the atmospheric surface temperature, $T_{atm,s}$,

$$F^{(h)}_{a\to s}(T_{atm,s}) = \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{a\to s} h^{(\ell)}(T_{atm,s}). \tag{66}$$

In reality each form of water has a different temperature in which case the terms in the sum on the right-hand side of (66) would be $h^{(\ell)}(T^{(\ell)}_{atm,s})$ where $T^{(\ell)}_{atm,s}$ is the temperature of water species ℓ.

The superscript h refers to total enthalpy flux (F denotes flux). We are assuming that the dry air flux is zero, $F^{(d)}_{a\to s} \equiv 0$. An example of the liquid enthalpy flux is the rain flux $F^{(liq)}_{a\to s}$. The ice flux can be snow $F^{(ice)}_{a\to s} = F^{(sn)}_{a\to s}$, but may also include graupel and hail falling into the surface. Water vapor flux to the surface can be dew and rime. Substituting the partial enthalpy terms (59), (60) and (61) into (66) and re-arranging terms (just like we did in the derivation of (64)) yields

$$F^{(h)}_{a\to s} = \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{a\to s}\left[c^{(\ell)}_p\left(T_{atm,s} - T_{00}\right) + h^{(ice)}_{00}\right] + F^{(wv)}_{a\to s} L_{s,00} + F^{(liq)}_{a\to s} L_{f,00}. \tag{67}$$

Similarly, the flux from the surface to the atmosphere (exactly the same derivations as before but with $T = T_{surf,s}$ and flux reversed; $s \to a$ instead of $a \to s$)

$$\begin{aligned} F^{(h)}_{s\to a}(T_{surf,s}) &= \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{s\to a} h^{(\ell)}(T_{surf,s}), \\ &= \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{s\to a}\left[c^{(\ell)}_p\left(T_{surf,s}-T_{00}\right)+h^{(ice)}_{00}\right] + F^{(wv)}_{s\to a}L_{s,00} + F^{(liq)}_{s\to a}L_{f,00}. \end{aligned} \tag{68}$$

where $T_{surf,s}$ is the surface temperature in the surface. Water vapor can enter the atmosphere through evaporation and sublimation. The process with which liquid enters the atmosphere could be spray and ice can enter the atmosphere through snow drift.

The net flux from the surface to the atmosphere, $F^{(h)}_{net} \equiv F^{(h)}_{s\to a} - F^{(h)}_{a\to s}$, is then

$$\begin{aligned} F^{(h)}_{net} = \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{net}\left[h^{(ice)}_{00} - c^{(\ell)}_p T_{00}\right] + F^{(wv)}_{net}L_{s,00} + F^{(liq)}_{net}L_{f,00} + \\ \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{s\to a}c^{(\ell)}_p T_{surf,s} - \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{a\to s}c^{(\ell)}_p T_{atm,s}. \end{aligned} \tag{69}$$

Note that the net fluxes of mass balance the mass change in the column

$$\iint \left[F^{(\ell)}_{s\to a} - F^{(\ell)}_{a\to s}\right] dA = \frac{\partial}{\partial t}\iiint \rho^{(d)}\widehat{m}^{(\ell)} dA\, dz, \quad \ell \in \mathcal{L}_{H_2O}, \tag{70}$$

where $\partial\widehat{m}/\partial t$ is change in mixing ratio ℓ due to falling precipitation or evaporation and not phase changes, $\partial\breve{m}/\partial t$. Consequently, the constant enthalpy terms in the square brackets on the right-hand side of (64) exactly cancel the surface flux terms when integrated over the global domain and the time derivative is taken:

$$\begin{aligned} \frac{\partial}{\partial t}\iiint \rho^{(d)}\left[\widehat{m}^{(\ell)}\left(h^{(ice)}_{00} - c^{(\ell)}_p T_{00}\right)\right] dA\, dz &= \left(h^{(ice)}_{00} - c^{(\ell)}_p T_{00}\right)\frac{\partial}{\partial t}\iiint \rho^{(d)}\widehat{m}^{(\ell)} dA\, dz, \\ &= \left(h^{(ice)}_{00} - c^{(\ell)}_p T_{00}\right)\iint F^{(\ell)}_{net}\, dA, \end{aligned}$$

where $F^{(\ell)}_{net}$ is the net change of water species ℓ in the column, and thus the terms cancel regardless of the choice of reference enthalpy $h^{(ice)}_{00}$ and reference temperature T_{00}.

For simplicity we assume that the temperature at the interface is the same in both atmosphere and the surface, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$. In this case the enthalpy flux at the surface, (69), is vastly simplified,

$$F^{(h)}_{net} \approx \sum_{\ell\in\mathcal{L}_{H_2O}} F^{(\ell)}_{net}\left[c^{(\ell)}_p\left(\widetilde{T}_s - T_{00}\right) + h^{(ice)}_{00}\right] + F^{(wv)}_{net}L_{s,00} + F^{(liq)}_{net}L_{f,00} \tag{71}$$

(ice reference state, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

The surface enthalpy flux formulas for other reference states can be found in Appendix C2. Note that assuming that all forms of water enter and leave the atmosphere at the same temperature is questionable (this assumption is discussed in more detail in Section 2.1.13). Also, this surface enthalpy interface condition does not consider a surface component using a different definition of enthalpy. An example from ocean-atmosphere coupling is given in Appendix D. In this paper we will not pursue this more complicated interface condition further, however, the reader should be aware of the simplifications made in this section.

2.1.5 Kinetic energy fluxes and surface stress

The flux at the surface should specify the flux of kinetic, internal and potential energy. Thus far we have only discussed the enthalpy flux. The flux of horizontal kinetic energy due to falling precipitation and water entering the atmosphere can be written as

$$F^{(H_2O)}_{s\to a}K_{surf,s} - F^{(H_2O)}_{a\to s}K_{atm,s} \tag{72}$$

where $K_{atm,s}$ is the specific kinetic energy at the surface from the atmosphere side (or surface layer as we will discuss in a moment) and $K_{surf,s}$ is the horizontal specific kinetic energy of the water leaving and entering the atmosphere, respectively.

The complications associated with specifying the surface (horizontal) kinetic energy flux is discussed further in Section 2.1.13. For simplicity, let's assume that

$$F_{net}^{(K)} \equiv F_{net}^{(H_2O)} \widetilde{K}_s, \tag{73}$$

where $\widetilde{K}_s = K_{atm,s} = K_{surf,s}$ is a common horizontal kinetic energy at the surface. No models, as far as the authors are aware, transfer horizontal kinetic energy associated with falling precipitation and evaporation between components. At very high vertical resolution and using a no-slip boundary condition the winds should be zero at the surface making this terms zero. A detailed discussion of the subject of surface drag (mechanical stress) is beyond the scope of this paper, however, a brief discussion is given in Appendix E as it does relate to the energy budget in a substantial way.

2.1.6 Potential energy fluxes

Ideally the potential energy flux through the surface would require one to track the altitude at which each water molecule in the air evaporated from the surface, and subtract the geopotential when that water molecule left the atmosphere as precipitation. Thus precipitation takes potential energy out of the atmosphere *only* if it hits the ground at a different altitude than evaporation (analogous to hydropower in mountains). Lacking such information, our best guess is that the overwhelming majority of water has evaporated from the ocean and the flux can be approximated as

$$F_{net}^{(\Phi)} \approx F_{net}^{(H_2O)} \Phi_s. \tag{74}$$

Another argument for (74) is that most of the potential energy excess of precipitation with respect to Φ_s is dissipated aerodynamically when the drops/snowflakes fall through air (after conversion of that excess potential energy into kinetic energy of the droplets). Hence most of the $\Phi - \Phi_s$ energy is in fact left in the air. A way to consistently include that process in models (that assume a single temperature) is given in Appendix F.

Note that we have implicitly chosen sea-level as the reference height for computing potential energy. So, similarly to enthalpy, there is a reference state (height) for potential energy. No models, as far as the authors are aware, transfer potential energy between components.

2.1.7 Sensible/turbulent heat flux

In addition to the fluxes discussed so far, there is also a surface turbulent/sensible heat flux $F^{(turb)}$ at the surface which come about by direct contact interaction with the surface. It is typically parameterized based on measurements of turbulent transport at a height, z, in the atmosphere:

$$F^{(turb)} = \left(\rho^{(all)} c_p^{(d)}\right) St_H \|\vec{v}_h(z_1) - \vec{v}_{h,s}\| \left(\theta(z_1) - \theta_s\right) \tag{75}$$

with the Stanton number, St_H, depending on underlying model, z, and the modeled states. Although these dependencies vary, flux differences are generally much less than the uncertainty of (75), with the use of temperature difference $(T(z_1) - \widetilde{T}_s)$ one example. Nevertheless, it is important that each component sees the same sensible heat flux ($\mathcal{F}_{ocn}^{(top)} = \mathcal{F}_{atm}^{(bottom)}$ in (3) and (4)) and applies it consistently with the energy formulation used in the components (see the example in Section 4.3.2).

2.1.8 Bulk heating and cooling

In addition to the flux terms and heating/cooling terms associated with phase transformations (that are already represented in the enthalpy terms), there is *bulk* heating and cooling within the atmosphere. In particular, heating/cooling that is associated with radiation. Others could be turbulent or wave dissipation incl., e.g., ion friction in the mesosphere. The radiative heating/cooling is represented as the divergence of radiative fluxes and such a term should be included on the left-hand side of the energy equation. Integration of this term over a column results in boundary terms (top of atmosphere, ToA, and surface terms). These terms can be written as

$$F_{net}^{(rad)} = F_{SW,surf}^{(rad)} + F_{LW,surf}^{(rad)} - F_{ToA}^{(rad)}, \tag{76}$$

(e.g. Trenberth, 1997) where the net short-wave and long-wave radiation fluxes at the surface are $F_{SW,surf}^{(rad)}$ and $F_{LW,surf}^{(rad)}$, respectively, and $F_{ToA}^{(rad)}$ is the net radiative flux at the model top. The upward short-wave radiative flux at the surface is computed as $F_{SW,surf}^{(rad),(up)} = \alpha_s F_{SW,surf}^{(rad),(down)}$ with α_s the surface albedo and $F_{SW,surf}^{(rad),(down)}$ the downward shortwave radiation. The long-wave upward flux, $F_{LW,surf}^{(rad),(up)}$, is computed based on each underlying model's surface state and parameterizations for surface emissivity and for reflecting downward long-wave radiation from the atmosphere.

We re-iterate that the radiative fluxes represent a bulk source/sink of atmospheric energy, not a surface (or ToA) boundary term; these result from integrating the divergence of the radiative flux over the column. For notational clarity (and since this term is not discussed further), the sensible/turbulent heat flux and the bulk heating/cooling terms are lumped into one term, $F_{net}^{(turb,rad)}$, on the right-hand side of the total energy equation (and we omit the explicit radiative flux-divergence term on the left-hand side of the energy equation).

2.1.9 Final total energy equations (assuming $T_{atm,s} = T_{surf,s}$)

For notational brevity we assume that water entering/leaving the atmosphere have the same temperature, $T_{atm,s} = T_{surf,s}$, in the "final" energy equations (extending to the more general equation allowing for different temperatures for water entering or leaving the atmosphere is straight forward by using (69) instead of (71) for the net enthalpy flux). Adding all the terms described in the previous sections results in the following total energy equation

$$\begin{aligned}
&\frac{\partial}{\partial t}\iiint \rho^{(d)}\Bigg\{\sum_{\ell\in\mathcal{L}_{all}} m^{(\ell)}\left(K+\Phi_s\right) + c_p^{(d)}T + m^{(H_2O)}\left[c_p^{(ice)}\left(T-T_{00}\right) + h_{00}^{(ice)}\right] \\
&\qquad\qquad + m^{(wv)}L_s(T) + m^{(liq)}L_f(T)\Bigg\} dA\, dz \\
=\;& \iint \Bigg\{F_{net}^{(H_2O)}(\widetilde{K}_s + \Phi_s) + F_{net}^{(H_2O)}\left[c_p^{(ice)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(ice)}\right], \\
&\qquad\qquad + F_{net}^{(wv)}L_s(\widetilde{T}_s) + F_{net}^{(liq)}L_f(\widetilde{T}_s) + F_{net}^{(turb,rad)}\Bigg\} dA
\end{aligned} \tag{77}$$

(ice reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

or, equivalently, by expanding latent heat terms and re-arranging terms:

$$\begin{aligned}
&\frac{\partial}{\partial t}\iiint \rho^{(d)}\Bigg\{K + \Phi_s + c_p^{(d)}T + \sum_{\ell\in\mathcal{L}_{H_2O}} m^{(\ell)}\left[K + \Phi_s + c_p^{(\ell)}\left(T-T_{00}\right) + h_{00}^{(ice)}\right] \\
&\qquad\qquad + m^{(wv)}L_{s,00} + m^{(liq)}L_{f,00}\Bigg\} dA\, dz
\end{aligned}$$

$$= \iint \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)} \Big[\widetilde{K}_s + \Phi_s + c_p^{(\ell)} \Big(\widetilde{T}_s - T_{00} \Big) + h_{00}^{(ice)} \Big] + F_{net}^{(wv)} L_{s,00} + F_{net}^{(liq)} L_{f,00} + F_{net}^{(turb,rad)} \Bigg\} dA. \quad (78)$$

(ice reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

Final total energy equations using other reference states can be found in Appendix C3.

For the following discussions/observations the time-change of water species ℓ is separated into local phase changes and changes associated with water entering or leaving the column

$$\frac{\partial m^{(\ell)}}{\partial t} = \frac{\partial \breve{m}^{(\ell)}}{\partial t} + \frac{\partial \widehat{m}^{(\ell)}}{\partial t}, \quad (79)$$

respectively. Strictly speaking the $\widehat{(\cdot)}$ and $\breve{(\cdot)}$ notation should be $\widehat{\frac{\partial m}{\partial t}}^{(\ell)}$ and $\breve{\frac{\partial m}{\partial t}}^{(\ell)}$ since it is meant to describe two separate processes both of which change the constituent mass; the mass itself cannot be split into a part originating from phase changes and one originating from precipitation and surface exchanges that the notation could imply. For notational ease, however, we use the notation in (79) so that the time-derivative can stay outside the vertical integral in the energy equations.

Important observations to be made regarding the enthalpy reference state terms: First of all, note that the constant enthalpy term $h_{00}^{(ice)}$ cancels on the left and right-hand side of (78) since

$$\begin{aligned} \frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)} h_{00}^{(ice)} \Bigg\} dA\, dz &= h_{00}^{(ice)} \frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} \Big(\breve{m}^{(\ell)} + \widehat{m}^{(\ell)} \Big) \Bigg\} dA\, dz, \\ &= h_{00}^{(ice)} \frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} \widehat{m}^{(\ell)} \Bigg\} dA\, dz, \quad (80) \\ &= h_{00}^{(ice)} \sum_{\ell \in \mathcal{L}_{H_2O}} \iint F_{net}^{(\ell)} dA, \quad (81) \end{aligned}$$

In (80) we have decomposed $\partial m^{(\ell)}/\partial t$ into local phase changes and falling precipitation $\partial/\partial t \left(\breve{m}^{(\ell)} + \widehat{m}^{(\ell)} \right)$ (see (79)) and used that local phase changes just rearrange water while total water is conserved $\partial \breve{m}^{(H_2O)}/\partial t = 0$ and, in (81), we use that the mass-change in the column equals the flux into the column

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \widehat{m}^{(\ell)} dA = \iint F_{net}^{(\ell)} dA\, dz, \quad (82)$$

(falling precipitation and surface evaporation)

Second, note that, by making the common simplification of assuming that all heat capacities are that of dry air (89), the reference temperature terms on the left- and right-hand side of the energy equation (78) cancel since

$$\begin{aligned} \frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)} c_p^{(d)} T_{00} \Bigg\} dA\, dz &= c_p^{(d)} T_{00} \frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} \widehat{m}^{(\ell)} \Bigg\} dA\, dz, \quad (83) \\ &= c_p^{(d)} T_{00} \sum_{\ell \in \mathcal{L}_{H_2O}} \iint F_{net}^{(\ell)} dA, \\ &= c_p^{(d)} T_{00} \iint F_{net}^{(H_2O)} dA, \quad (84) \end{aligned}$$

$$(\forall \ell : \quad c_p^{(\ell)} = c_p^{(d)})$$

using the same manipulations as in deriving (81).

Third, if variable heat capacities are used then a similar equation only holds for falling precipitation and water entering the column

$$\frac{\partial}{\partial t}\sum_{\ell\in\mathcal{L}_{H_2O}}\iiint \rho^{(d)}\left\{\widehat{m}^{(\ell)}c_p^{(\ell)}T_{00}\right\}dA\,dz \quad = \quad \sum_{\ell\in\mathcal{L}_{H_2O}}\iint c_p^{(\ell)}T_{00}F_{net}^{(\ell)}dA, \tag{85}$$

(falling precipitation and evaporation)

whereas for phase changes

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\left\{\sum_{\ell\in\mathcal{L}_{H_2O}}\breve{m}^{(d)}c_p^{(\ell)}T_{00}\right\}dA\,dz \quad \neq \quad 0, \tag{86}$$

(phase changes only)

due to the different specific heat weighting. This term is consistent with the definitions of latent heat and the corresponding terms $\rho^{(d)}\sum_{\ell\in\mathcal{L}_{H_2O}}\breve{m}^{(d)}L_{x,00}$ when $T_{00}\neq 0$; it is in effect an enthalpy term asociated with the latent heats of water phases and implies no dependence on the chosen reference state (ice, liquid, vapor) or temperature T_{00}. For constant latent heats and phase-independent water heat capacities (e.g. $c_p^{(\ell)}\equiv c_p^{(d)}$ for all water species), the right-hand side of (86) is zero since $\partial\breve{m}^{(H_2O)}/\partial t = 0$. In the case of variable latent heats, and associated phase-specific heat capacities, if we were to change to a different reference state then the latent heat terms would change alongside with T_{00} so that effectively one would get the same "physical" result (i.e. the energy equation should not depend on reference state but the left-hand side and right-hand side separately do depend on reference state).

In the following we will discuss the total energy equations under various assumptions. The discussion has been split into several sections and will (in some sections) touch upon vast and complicated problems/subjects with extra detail diverted to Appendices. The sections are about energy conservation

- (section 2.1.10) under adiabatic conditions (dynamical cores),
- (section 2.1.11) under the presence of local phase changes due to (parameterized) microphysics (no falling precipitation, no re-evaporation, no momentum changes due to other unresolved processes, etc.),
- (section 2.1.12) under the presence of momentum sources and sinks due to parameterized processes such as boundary layer turbulence, gravity waves, etc. and
- (section 2.1.13) for falling precipitation/evaporation and surface fluxes.

2.1.10 Energy conservation in an adiabatic dynamical core (no phase changes and no fluxes)

Consider an adiabatic dynamical core with inert water species i.e. there are no phase transformations and/or fluxes so the global mass of each component of air is conserved. As a consequence, the latent heat terms and constant enthalpy terms integrate to zero in the total energy equation (62)

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\left[\sum_{\ell\in\mathcal{L}_{H_2O}} m^{(\ell)}\left(-c_p^{(\ell)}T_{00}+h_{00}^{(ice)}\right)+m^{(wv)}L_{s,00}+m^{(liq)}L_{f,00}\right]=0. \tag{87}$$

The total energy equation becomes

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}E_{feom}\,dA\,dz=\frac{\partial}{\partial t}\iiint \rho^{(d)}\Big[\sum_{\ell\in\mathcal{L}_{all}} m^{(\ell)}\left(K+\Phi_s+c_p^{(\ell)}T\right)\Big]dA\,dz=0. \tag{88}$$

(all enthalpy reference states, adiabatic dynamical core with inert water species)

consistent with the derivation in Lauritzen et al. (2018). Note that the energies discussed above are pseudo energies since constant terms in the energy have been discarded as they are zero when taking the time derivative.

It is common practice in large scale models to assume that the latent heat terms are constant which, by examination of (55), (56), (57), is equivalent to assuming that all components of moist air have the same specific heat,

$$c_p^{(\ell)} = c_p^{(d)} \text{ for } \ell \in \mathcal{L}_{H_2O}, \tag{89}$$

where the specific heat of dry air is used since that is the most abundant air component in the atmosphere. In this case, the total energy formula becomes

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} E_{feom}\, dA\, dz = \frac{\partial}{\partial t} \iiint \rho^{(d)} \Big[\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \left(K + \Phi_s + c_p^{(d)} T \right) \Big] dA\, dz = 0. \tag{90}$$

(all enthalpy reference states, adiabatic dynamical core with inert water species and $c_p^{(\ell)} \equiv c_p^{(d)}$) which is most easily derived by using (64). Ideally, the dynamical core should use an equation set that is consistent with the physics package. In this case this means assuming either variable, (88), or constant latent heats (90).

2.1.11 Local phase changes and total energy conservation

Parameterizations are, in general, not formulated such as to satisfy the conservation law in (62), but rather in terms of local thermodynamic variables that are conserved in the process under consideration (e.g., various forms of static energy; Emanuel, 1994). Nevertheless, parameterizations (e.g. micro physics) that rearrange water locally through phase transformations from one species to another, denoted $\frac{\partial \breve{m}^{(\ell)}}{\partial t}$, while preserving the total water content $\frac{\partial}{\partial t} \breve{m}^{(H_2O)} = 0$, should satisfy conservation of energy (62). In this case the temperature change associated with latent heat release should locally at each grid point satisfy

$$\frac{\partial}{\partial t} \left[\rho^{(d)} c_p^{(d)} T + \rho^{(d)} \breve{m}^{(H_2O)} c_p^{(ice)} T + \rho^{(d)} \breve{m}^{(wv)} L_s(T) + \rho^{(d)} \breve{m}^{(liq)} L_f(T) \right] = 0, \tag{91}$$

which, by expanding the latent heat terms and re-arranging terms, can be written as

$$\frac{\partial}{\partial t} \left[\rho^{(d)} \sum_{\ell \in \mathcal{L}_{all}} \breve{m}^{(\ell)} c_p^{(\ell)} \left(T - T_{00} \right) + \rho^{(d)} \breve{m}^{(wv)} L_{s,00} + \rho^{(d)} \breve{m}^{(liq)} L_{f,00} \right] = 0. \tag{92}$$

(ice reference; local phase changes; no falling precipitation or surface water source)

For notational clarity on the left-hand side of (92) we add the term $-\frac{\partial}{\partial t} \left(\rho^{(d)} c_p^{(d)} T_{00} \right) \equiv 0$ (so that all terms can be in the summation). These equations, perhaps most easily observed with (91), are consistent with the "moist atmospheric energy" in equation (12) in Mayer et al. (2017). Note that since $\breve{m}^{(H_2O)}$ is constant, the constant enthalpy, kinetic and Φ_s terms do not appear in (92). Also, the property that $\breve{m}^{(H_2O)}$ is constant implies that the vertical coordinate is Lagrangian hence, with hydrostatic balance, $\partial p(t, \eta)/\partial t = 0$, e.g. the vertical coordinate is pressure-based. Then (92) is a special case of the first law of thermodynamics $dh^{(all)} = \alpha^{(gas)} dp = 0$ when $dp = 0$.

Invoking the usual assumption/approximation that all forms of water have the same specific heat, $c_p^{(\ell)} = c_p^{(d)}$, in which case the latent heat becomes a constant (see (55), (56), (57)), then (92) becomes

$$c_p^{(d)} \left(\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \frac{\partial T}{\partial t} \right) = -\frac{\partial}{\partial t} \left(\breve{m}^{(wv)} L_{s,00} + \breve{m}^{(liq)} L_{f,00} \right). \tag{93}$$

(ice reference; local phase changes; no falling precipitation; $c_p^{(\ell)} = c_p^{(d)}$)

Any temperature change due to local phase changes (not falling precipitation or surface water sources) should satisfy this equation locally (in each grid point). A discussion of $\partial\widehat{m}/\partial t$ (e.g. falling precipitation, evaporation) is given in Section 2.1.13 once the surface fluxes have been discussed.

2.1.12 Local momentum sources/sinks and energy conservation

Sources and sinks of momentum (e.g. gravity wave parameterization, boundary layer turbulence schemes or other drag parameterizations) affect kinetic energy, and enforcing total energy conservation in their presence is not straightforward due to its interaction with sub-grid-scales. In particular, we would like to point out that a "naive" closure of the energy budget by transferring kinetic energy change into heat is, in general, **not** physically correct

$$\left(\sum_{\ell\in\mathcal{L}_{all}} m^{(\ell)} c_p^{(\ell)}\right)\frac{\partial T}{\partial t} \neq -\left(\sum_{\ell\in\mathcal{L}_{all}} m^{(\ell)}\right)\frac{\partial K}{\partial t}. \tag{94}$$

although sometimes applied in models (an example is given in 4.1.2). The rationale for turning the rate of change of resolved kinetic energy into an increase of temperature is that ultimately kinetic energy is converted into heat by frictional processes occurring at unresolved scales. However $\partial K/\partial t$ should in general be decomposed into a flux divergence and source, and *only* the latter is potentially convertible into heat. Let us explain in more detail.

Frictional dissipation of kinetic energy ultimately occurs at the molecular scale where it is turned into heating (a.k.a. frictional heating). However, at resolutions at which ESM's are run there are many processes resulting in a frictional force on momentum that are much more efficient than viscous effects for transferring momentum to/from sub-grid scales (e.g. small-scale turbulent eddies, "eddy friction" due to cumulus momentum mixing and breaking gravity waves). These processes are represented by a stress tensor in the equations of motion (e.g., Smagorinsky, 1963).

Since this discussion is focused on column physics, let's consider only frictional terms due to the vertical mixing of horizontal momentum (formally, the z-component of the divergence of the stress tensor τ_{ij}, see e.g., p.36 in Peixoto & Oort, 1992).; i.e. we are ignoring the cross terms in the stress tensor. The frictional force on momentum is the divergence of the stress tensor (in tensor notation $\vec{\tau}_z = (\tau_{zx}, \tau_{zy})$ in Cartesian coordinates) so in a physics column the horizontal momentum equations take the form

$$\frac{D\vec{v}}{Dt} = -\vec{F}_h, \tag{95}$$

$$= -\frac{1}{\rho^{(all)}}\frac{\partial}{\partial z}\left(\vec{\tau}_z\right), \tag{96}$$

(section 3.2.1 in Peixoto & Oort, 1992). Note that since this equation is for column physics the pressure gradient force, Coriolis force and horizontal advection of momentum terms do not appear. Now we can compute the (specific) kinetic energy equation by multiplying (96) with $\vec{v}$, using the chain rule for differentiation and re-arranging terms

$$\rho^{(all)}\frac{DK}{Dt} = \rho^{(all)}\frac{D}{Dt}\left(\frac{1}{2}\vec{v}^2\right) = \rho^{(all)}\vec{v}\cdot\frac{D\vec{v}}{Dt} = -\vec{v}\cdot\frac{\partial}{\partial z}\left(\vec{\tau}_z\right) = -\frac{\partial}{\partial z}\left(\vec{v}\cdot\vec{\tau}_z\right) + \vec{\tau}_z\cdot\frac{\partial\vec{v}}{\partial z}. \tag{97}$$

(mean kinetic energy equation in physics column)

Equation (97) is the mean kinetic energy equation as it describes the evolution of resolved-scale kinetic energy. The interaction between the sub-grid scale kinetic energy (often referred to as **turbulent kinetic energy**, **TKE**, but could also be other un-resolved wave

interactions such as gravity waves) and resolved scales is represented by the stress tensor terms (right-hand side of (97)).

Now lets integrate (97) over a layer. First it is noted that by repeated application of the chain rule for differentiation after expanding the material derivative and using the continuity equation for density, the left-hand side of (97) can be written as

$$\begin{aligned}\rho^{(all)}\frac{DK}{Dt} &= \frac{D}{Dt}\left[\rho^{(all)}K\right] - K\frac{D\rho^{(all)}}{Dt},\\ &= \frac{\partial}{\partial t}\left[\rho^{(all)}K\right] + \vec{v}\cdot\nabla\left[\rho^{(all)}K\right] + w\frac{\partial}{\partial z}\left[\rho^{(all)}K\right] - K\left\{-\rho^{(all)}\left[\nabla\vec{v} + \frac{\partial w}{\partial z}\right]\right\},\\ &= \frac{\partial}{\partial t}\left[\rho^{(all)}K\right] + \nabla\cdot\left[\rho^{(all)}K\vec{v}\right] + \frac{\partial}{\partial z}\left[\rho^{(all)}Kw\right],\end{aligned}$$

(for column parameterizations with no interactions between columns the material derivative is $D/Dt = \partial/\partial t + w\partial/\partial z$ rather than $D/Dt = \partial/\partial t + \vec{v}\cdot\nabla + w\partial/\partial z$; for completeness we retain the horizontal term) which, when integrated over a layer, becomes

$$\int_{layer}\rho^{(all)}\frac{DK}{Dt}dz = \frac{D}{Dt}\int_{layer}\rho^{(all)}K\,dz \tag{98}$$

As a results the mean kinetic energy equation, (97), integrated over a layer can be written as

$$\frac{D}{Dt}\int_{layer}\rho^{(all)}Kdz = \int_{layer}\left[-\frac{\partial}{\partial z}\left(\vec{v}\cdot\vec{\tau}_z\right) + \vec{\tau}_z\cdot\frac{\partial\vec{v}}{\partial z}\right]dz. \tag{99}$$

The first term on the right-hand side of (99) is in flux-form. Therefore, if we assume zero velocity or stress at the lower and upper boundaries (more on that in Section 2.1.5) then this term does not change kinetic energy in the column (only redistributes kinetic energy between layers). The second term represents the shear production of TKE, and is a term in the sub-grid scale (turbulent) kinetic energy equations (Lumley & Panofsky, 1964).

The crux, in terms of energy conservation, is whether TKE produced by shear should be converted into a resolved-scale temperature tendency or not. A more conceptual question is whether this conversion should be called viscous/frictional.

Regarding the first question, if subgrid energy reservoirs are neglected, total resolved energy should indeed be conserved, which can only be achieved by translating the shear production of TKE into an increase of temperature. Whether the underlying process is indeed viscous dissipation is more subtle. For example, by the second law of thermodynamics, conductive heat transfer in stably stratified turbulence increases entropy, and thus potential temperature, thereby increasing column-integrated enthalpy at a rate linked to the parameterized turbulent heat flux. The energy needed for this increase comes from the TKE. Only the fraction of TKE not consumed by heat transfer is dissipated to heat by viscosity, and conversely the parameterized heat flux cannot exceed an upper bound controlled by the shear production of TKE (Akmaev, 2008).

Thus, although a loss of resolved kinetic energy must be compensated somehow by an increase of (potential) temperature, the underlying small-scale processes are not purely viscous. An example of a rigorous closure in ocean models, in terms of energetics, is given in Eden (2016).

In the special case where the stress tensor represents an eddy momentum flux (downgradient diffusion typically found in turbulence and boundary layer schemes) and is parameterized as

$$\vec{\tau}_z \equiv -\rho^{(all)}\nu^{(all)}\frac{\partial\vec{v}}{\partial z},\ \text{(eddy diffusion parameterization)} \tag{100}$$

(e.g., Kieu, 2015) equation (99) becomes

$$\int_{layer} \rho^{(all)} \frac{\partial K}{\partial t} dz = \int_{layer} \left[\underbrace{\frac{\partial}{\partial z}\left(\rho^{(all)} \nu^{(all)} \frac{\partial K}{\partial z}\right)}_{\text{Redistribution of } \rho^{(all)} K} - \underbrace{\rho^{(all)} \nu^{(all)} \left(\frac{\partial \vec{v}}{\partial z}\right)^2}_{\text{Shear production of TKE}} \right] dz. \tag{101}$$

which has the same form as molecular friction. In this special case, it can be argued that the last term in (101) represents frictional/dissipative heating if $\nu^{(all)}$ is positive (note also that eddy diffusion with $\nu^{(all)} > 0$ fulfills the second law of thermodynamics (Schaefer-Rolffs & Becker, 2018)) and can be included in the (resolved-scale) thermodynamic equation for energy conservation (e.g., Bister & Emanuel, 1998):

$$\rho^{(d)} \left(\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} c_p^{(\ell)} \right) \frac{\partial T}{\partial t} = \rho^{(d)} \left(\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \nu^{(\ell)} \right) \left(\frac{\partial \vec{v}}{\partial z} \right)^2. \tag{102}$$

Note that the local change in the winds $\partial K/\partial t = \partial/\partial t \left(\frac{1}{2}\vec{v}^2\right)$ is not equal to the frictional heating term (right-hand side of (102)) since the energy-conservative redistribution of K (first term on left-had side of (101)) is missing.

In general, however, there is no guarantee that the last term on the right-hand side of (99) is negative; if positive (e.g. for sub-grid-scale gravity waves accelerating the flow) it would lead to "frictional cooling" if added as to the thermodynamic equation to close the energy budget, which makes no physical sense (Becker, 2001). Hence, in general, a sub-grid model is needed to close the energy budget since artificially closing the energy budget at resolved-scales can lead to unphysical cooling. This is related to the immensely complicated subject of the interaction between resolved and unresolved scales discussed briefly in Appendix A.

For a comprehensive discussion of entropy budgets in numerical models see Gassmann and Herzog (2015). The horizontal part of frictional heating is discussed in the context of dynamical cores in Section 4.1. Also, kinetic energy and the boundary condition (surface stress) is discussed in Section 2.1.5.

2.1.13 Energy conservation for falling precipitation/evaporation and surface fluxes

In Section 2.1.11 we discussed energy conservation for processes in the atmosphere not involving falling precipitation and evaporation. For water entering or leaving the vertical column the picture becomes much more complicated compared to, e.g., having a closed energy budget for phase changes.

Making the assumption that falling precipitation (that does not re-evaporate or undergo other microphysical interactions) instantly hits the surface after formation (i.e. that it does not interact with the atmosphere, so that it falls like an *insulated rain shaft*), the latent heat terms on the left and right-hand side of (77) should exactly match

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \left[\widehat{m}^{(wv)} L_s(T) + \widehat{m}^{(liq)} L_f(T) \right] dA\, dz = \iint \left[F_{net}^{(wv)} L_s(\widetilde{T}) + F_{net}^{(liq)} L_f(\widetilde{T}) \right] dA. \tag{103}$$

(ice reference; falling precipitation \w no interaction \w environment)

This equation states that the temperature $\widetilde{T}$ should be the temperature at which the falling precipitation / evaporation was formed for the atmosphere energy budget to be closed (under the single-temperature assumption, that is the temperature of the layer where the falling rain was formed). Note that evaporation is simpler since it enters from the surface (known temperature) directly into the lowest model level and does not fall through

the atmosphere. A similar budget holds for the enthalpy flux term that should balance the time change of the enthalpy terms on the left-hand side of the energy equation

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\widehat{m}^{(H_2O)}c_p^{(ice)}T\,dA\,dz = \iint F_{net}^{(H_2O)}c_p^{(ice)}\widetilde{T}dA. \tag{104}$$

(ice reference; falling precipitation \w no interaction \w environment)

as well as kinetic energy

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\left[\widehat{m}^{(H_2O)}K\right]dA\,dz = \iint \left[F_{net}^{(H_2O)}\widetilde{K}_s\right]dA. \tag{105}$$

(falling precipitation \w no interaction \w environment)

For (105) to be satisfied (i.e. closed energy budget) under the "insulated rain shaft" assumption then $\widetilde{K}_s$ should be the horizontal kinetic energy where the falling water is formed and surface winds for water entering the atmosphere. The potential energy budget is a little different

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\left[\widehat{m}^{(H_2O)}\Phi_s\right]dA\,dz = \iint \left[F_{net}^{(H_2O)}\widetilde{\Phi}_s\right]dA. \tag{106}$$

since it does not depend on temperature and wind. Also note that potential energy only changes if water exits the atmosphere at a different elevation than where it entered and vice versa (see section 2.1.6).

While (103), (104), (105) and (106) under the insulated rain shaft assumption are energetically consistent, important physical processes are omitted. For example, measurements and theory for the surface rainfall temperature show that it is approximately equal to the surface wet-bulb temperature (Byers et al., 1949; Kinzer & Gunn, 1951; Gosnell et al., 1995), which, on average, will be warmer than the temperature at which condensation occurs. When the droplets hit the surface their temperature can be different than the surface ambient air temperature or temperature of the actual surface. For example, Anderson et al. (1998) observed that in the Indo-Pacific warm pool precipitation is $\sim 5K$ colder than the sea surface temperature (SST). This means that specifying $\widetilde{T}$ accurately requires parameterizations for the processes that precipitation undergo while it falls through the atmosphere. In addition, the droplets exert horizontal drag on the mean flow as they fall through the atmosphere and water vapor entering the atmosphere is accelerated towards the lowest model level winds. Precipitation also impacts the sensible heat flux as studied/discussed, in the context of CAM, in Wei et al. (2014) and not discussed further in this paper. These processes (and possibly others) are generally not parameterized in climate models.

Frictional heating of falling precipitation: Let's first discuss the vertical component of kinetic energy of falling precipitation as this has received some attention in the literature without broad adoption in global models. The process is referred to as frictional heating of falling precipitation (Pauluis et al., 2000; Pauluis & Dias, 2012). As soon as a droplet starts to fall there is a loss of potential energy which is first converted into kinetic energy, which is afterwards turned into internal energy by friction. Representation of this process is possible (it is actually automatic) using the barycentric velocity framework. This is described in detail in Appendix F. Like in the case discussed above where shear production is not the same as molecular dissipation, here the mixing dissipation represents the true Stokes dissipation (which is likewise invisible). In Gassmann and Herzog (2015) this reinterpretation for numerical models has been given by analyzing the local entropy production of this process. Using the barycentric framework this problem can, to some extent, be considered solved. That said, this formulation assumes additionally that all species have the same temperature. There is some concern in the community about that assumption. Since global models usually do make the constant temperature assumption, using the barycentric framework would be a first step towards adding missing processes associated with precipitation and evaporation.

The horizontal momentum of falling precipitation has not been addressed (to the authors knowledge) in the literature. The falling droplets will, in general, exert a drag in the horizontal (slowing down the horizontal winds) as they fall through the atmosphere. Also, when the droplets hit the surface they will have a horizontal momentum component (even though the wind may be zero in a very thin layer at the surface, the lower model level will not have zero winds). Possibly the common Lagrangian frame of reference as described in Appendix F could automatically treat this problem.

In conclusion, when making the insulated rain-shaft assumption it is problematic to assign a temperature to the latent heat flux and enthalpy flux (right-hand side of (103) and (104), respectively). If one assumes that the heat capacities for all forms of water are the same, $c_p^{(d)}$, then the latent heats become constant and (103) no longer has a temperature dependency

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \left[\widehat{m}^{(wv)} L_{s,00} + \widehat{m}^{(liq)} L_{f,00} \right] dA\, dz = \iint \left[F_{net}^{(wv)} L_{s,00} + F_{net}^{(liq)} L_{f,00} \right] dA, \quad (107)$$

(ice reference; falling precipitation \w no interaction \w environment, $c_p^{(\ell)} = c_p^{(d)}$)

and is trivially satisfied in models that have a closed mass budget (82). Unfortunately, the enthalpy flux term (104) still has a temperature dependency, even when making the simplification $c_p^{(\ell)} = c_p^{(d)}$.

Most large scale models do not rigorously account for the energy terms associated with falling precipitation and water entering the atmosphere, other than through (constant) latent heat terms. In terms of equations, this means that terms on the left-hand side of the energy equation are not balanced by parameterizations in the atmosphere and right-hand side flux terms:

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \sum_{\ell \in \mathcal{L}_{H_2O}} \widehat{m}^{(\ell)} \left(K + \Phi_s + c_p^{(\ell)} T \right) dA\, dz \quad (108)$$

(any reference; falling precipitation \w no interaction \w environment)

The magnitude of these terms will be estimated in the next section.

3 Energy budget errors of a climate model

While the previous section is quite general in its assumptions (e.g., single temperature and velocity), in the next Section we discuss detailed issues that occur due to specific assumptions that are made in CAM in terms of energetics. Therefore details specific to CAM will be explained in some detail although they are not necessarily relevant to other modeling systems.

We present a concrete example of the energy budget, using simulation results from the CAM climate model. We focus on the energy budget in parameterization suites where the total energy budget should be closed in each column. In the dynamical core, where there is interaction between columns, the integrated energy equation only holds globally.

3.1 Physics (parameterization) vertically integrated energy budget

To keep the discussion tractable we assume a hydrostatic model with the total energy equation given by (77) that uses an ice reference state. This energy equation holds for the continuous state variables $\left(u, v, T, m^{(\ell)}, \rho^{(\ell)}\right)$. Although averaging the equations of motion, and hence also the energy equation, entails additional terms and assumptions (as discussed in Appendix A), we follow the common practice and require that (77) hold for the grid cell averaged state, $\left(\overline{u}, \overline{v}, \overline{T}, \overline{m}^{(\ell)}, \overline{\rho}^{(\ell)}\right)$. The implied model will only approximate energy conservation imperfectly; however such a model will converge towards an

exactly energy-conserving model in the limit where grid spacing and time-step tend to zero. At finite resolution, a consequence of this assumption is that, without a sub-grid model that can store energy (a sub-grid scale energy reservoir), energy on the resolved scales must be conserved. As discussed in Section 2.1.11 this is problematic since terms that interact with the sub-grid-scale must satisfy certain properties (e.g. down-gradient diffusion being turned into frictional heating at the resolved scales) for the averaged energy equation to be physically valid. However, not all processes satisfy this property (e.g. gravity wave parameterizations may accelerate the flow leading to "frictional cooling" to close the energy budget at resolved scales).

Before discussing the energy errors, it is insightful to first assess the magnitude of the energy budget terms in a real-world climate simulation. Again, to keep the discussion tractable, we use the CAM model, and hence make assumptions specific to CAM. We will also diagnose energy budget changes associated with relaxing certain assumptions, e.g. by introducing variable latent heats and fluxes associated with enthalpy, kinetic and potential energy.

Please note that in the following we focus on the vertically integrated energy budget in the CAM physics parameterizations. These represent all diabatic heat sources of the equations of motion. In a steady-state, or long-term time average, their vertical integral must match energy flowing in and out of the upper and lower boundaries in order for energy conservation to be satisfied. The horizontal transport we are going to neglect (and which is computed in the dynamical core of the model) in general gives by design a vanishing contribution to the global total energy budget, and its main role is simply to redistribute the heat of the diabatic sources, column by column, such as to yield a locally vanishing long-term time average. Therefore, errors in the column energy budgets due to physics parametrisations directly impact the dynamics, which in response produces local errors equal in magnitude and opposite in sign. Note that the integrated energy equations derived in section 2 can be applied to the partial (physics-only) budgets in the column, since the tendencies associated with transport are neglected. Local energy conservation equations are presented in Appendix B2.

In terms of neglected terms and approximations to terms in the energy equation, CAM physics makes the following assumptions:

- The latent heat terms are constant. This is equivalent to assuming that the heat capacities of all species are the same (in the case of CAM $c_p^{(\ell)} = c_p^{(d)}$, $\forall \ell$)
- The kinetic, enthalpy and surface-geopotential surface flux terms are neglected so that only the sum

$$-c_p^{(d)} T_{00} \overline{F}_{net}^{(H_2O)} + F_{net}^{(wv)} L_{s,00} + F_{net}^{(liq)} L_{f,00} + F_{net}^{(turb,rad)}. \tag{109}$$

  is retained on the right-hand side of the energy equation
- The ice phase represents zero reference enthalpy, $h_{00}^{(ice)} \equiv 0 J/kg^2$ (if it was non-zero the $h_{00}^{(ice)}$ terms on the left and right-hand side of the energy equation would cancel anyway as shown in (81)), and the reference temperature is $T_{00} = 0^\circ C$ (for reasons that will become clear that term is kept in the equations).

With these assumption the energy equation for each physics column can be written as

$$\frac{\partial}{\partial t} \int \overline{\rho}^{(d)} \left\{ \left(1 + \overline{m}^{(H_2O)}\right) \left[\overline{K} + \overline{\Phi}_s + c_p^{(d)} \left(\overline{T} - T_{00}\right)\right] + \overline{m}^{(wv)} L_{s,00} + \overline{m}^{(liq)} L_{f,00} \right\} dz$$
$$= -c_p^{(d)} T_{00} \overline{F}_{net}^{(H_2O)} + \overline{F}_{net}^{(wv)} L_{s,00} + \overline{F}_{net}^{(liq)} L_{f,00} + \overline{F}_{net}^{(turb,rad)}, \tag{110}$$

where for notational clarity we have re-introduced the term $\rho^{(d)} c_p^{(d)} T_{00}$ which integrates to zero anyway (see paragraph under (63)). Furthermore, in terms of temporal evolution of the energy equation, CAM makes the following assumptions:

- Only water vapor contributes to kinetic, potential and sensible heat: $m^{(H_2O)} \approx m^{(wv)}$. In other words, condensates are massless in terms of kinetic, internal and geopotential energy (in (110) that is the sum of the K, Φ_s and c_pT terms multiplied by $\rho^{(d)}m^{(H_2O)}$).
- In addition, the total water mass is assumed constant during physics updates. That means the water mass, $m^{(H_2O)}$, associated with (i.e. multiplying) specific kinetic (K), potential (Φ_s), and non-latent internal energy (c_pT) is regarded as constant over the time-step; however the water mass associated with latent heat is updated consistently with water mass conservation and phase changes. Keeping mass fixed is denoted with subscript $t = t^n$, i.e. keeping the mass of water species ℓ constant at its initial value when initiating physics parameterization computations is denoted $\rho^{(d)}_{t=t^n} m^{(\ell)}_{t=t^n}$.

These simplifications lead to the following energy equation that each CAM parameterization, in theory, should satisfy:

CAM parameterization total energy equation

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\left\{\left(1+\overline{m}^{(H_2O)}_{t=t^n}\right)\left[\overline{K}+\overline{\Phi}_s+c_p^{(d)}\left(\overline{T}-T_{00}\right)\right]+\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\right\}dz$$
$$= -c_p^{(d)}T_{00}\overline{F}^{(H_2O)}_{net}+\overline{F}^{(wv)}_{net}L_{s,00}+\overline{F}^{(liq)}_{net}L_{f,00}+\overline{F}^{(turb,rad)}_{net}, \quad (111)$$

(assume (109), $c_p^{(\ell)} = c_p^{(d)}$, $h_{00}^{(ice)} \equiv 0 J/kg^2$, $T_{00} = 0^\circ C$, $\overline{m}^{(H_2O)} = \overline{m}^{(wv)}_{t=t^n}$).

Since total water is kept fixed, the total water flux term $c_p^{(d)}T_{00}\overline{F}^{(H_2O)}_{net}$ and the time-derivative of $\overline{m}^{(H_2O)}_{t=t^n}c_p^{(d)}T_{00}$ are zero. These terms are kept for the coming discussion on dependency on reference temperature. Also, even though CAM uses $m^{(H_2O)} \approx m^{(wv)}$ we keep the notation $m^{(H_2O)}$ in (111) to facilitate the discussion below where we estimate the energy error associated with that assumption.

After the last parameterization the total water (in CAM's case $m^{(H_2O)}_{t=t^n} \approx m^{(wv)}_{t=t^n}$) is updated leading to an imbalance in (111) denoted $\Delta\mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}$ (subscript $\partial m^{(H_2O)}/\partial t$ refers to imbalance due to letting total water evolve in time)

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\left\{\left(1+\overline{m}^{(H_2O)}\right)\left[\overline{K}+\overline{\Phi}_s+c_p^{(d)}\left(\overline{T}-T_{00}\right)\right]+\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\right\}dz$$
$$-\Delta\mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = -c_p^{(d)}\overline{F}^{(H_2O)}_{net}T_{00}+\left\{\overline{F}^{(wv)}_{net}L_{s,00}+\overline{F}^{(liq)}_{net}L_{f,00}+\overline{F}^{(turb,rad)}_{net}\right\}. \quad (112)$$

("dry-mass adjustment" equation)

By subtracting (111) from (112) (and expanding the time-derivative using the chain rule for differentiation), it can be shown that the imbalance of updating total water is

$$\Delta\mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = \int\left\{\frac{\partial}{\partial t}\left[\rho^{(d)}\left(1+\overline{m}^{(H_2O)}\right)\right]\right\}\left(\overline{K}+\overline{\Phi}_s+c_p^{(d)}\overline{T}\right)dz. \quad (113)$$

("dry mass adjustment" - referred to as *dme_adjust* in CAM codes)

Caution: We remind the reader that the two last terms in the integrand on the right-hand side of (113) are NOT potential and internal energy (see ambiguity 3 in section 2.1.1). These terms came about by integration by parts and application of a specific boundary condition used in most pressure-based models. In this form the internal and geopotential energy terms are not separated. The two terms are enthalpy (which obviously has physical meaning) and (what we may refer to as) a surface geopotential term. If we consider falling precipitation; the change in geopotential energy (not surface geopotential term) is large (since $\Phi - \Phi_s$ is large where falling precipitation forms) compared to the

internal energy difference between where falling precipitation is formed and the ground value of the internal energy. But in (113) the enthalpy term is the largest. Hence it is important not to conflate the internal/potential energy terms with enthalpy and surface geopotial energy terms. A complete mathematical description of this "caution" in the context of CAM is given in Appendix B1.

If we split $\partial m/\partial t$ into phase changes / kinetic energy of condensates "staying in the atmosphere", $\partial \breve{m}/\partial t$, and falling precipitation/evaporation $\partial \widehat{m}/\partial t$ as discussed in Section 2.1.13, (113) can be written as

$$\Delta \mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = \Delta \breve{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} + \Delta \widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}, \tag{114}$$

where

$$\Delta \breve{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = \int \frac{\partial}{\partial t}\left[\rho^{(d)}\overline{\breve{m}}^{(H_2O)}\right]\left(\overline{K} + \overline{\Phi}_s + c_p^{(d)}\overline{T}\right) dz \tag{115}$$

("spurious phase change term")

and

$$\Delta \widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = \int \frac{\partial}{\partial t}\left[\rho^{(d)}\overline{\widehat{m}}^{(H_2O)}\right]\left(\overline{K} + \overline{\Phi}_s + c_p^{(d)}\overline{T}\right) dz \tag{116}$$

(total energy flux/change associated with falling precipitation - evaporation)

The first term in the imbalance, (115), would have been zero if all forms of water were included in $m^{(H_2O)}$. Since CAM only includes water vapor in total water, a phase change from water vapor to a condensate causes spurious reduction in total water, and vice versa for phase change to water vapor. This impacts kinetic, surface geopotential and internal energy. Hence, we refer to it as the "spurious phase change" term. The second term is the kinetic, surface geopotential and enthalpy change associated with falling precipitation and evaporation, $\Delta \widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}$. The total $\Delta \mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} = \Delta \mathcal{I}^{(CAM)}_{\partial m^{(wv)}/\partial t}$ is referred to as "dry mass adjustment" in the CAM model code since there is an adjustment to the dry-mass to conserve tracer mass and dry mass (see equation 3.65 in Neale et al., 2010). This is a quite common assumption in NWP and climate models as is also discussed in Section 7.4 in Catry et al. (2007) in connection with the barycentric formulation and mass fluxes.

A global energy fixer computes the global integral of this imbalance

$$\iint \Delta \mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}\, dA, \tag{117}$$

and restores energy conservation globally through a uniform temperature increase that exactly balances (117). It is noted that NorCESM, whose the atmospheric component is based on CAM, restores $\Delta \mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}}$ in each column (see section 2.1.2 in Guo et al., 2019) rather than using a global "fixer". The global energy fixer also fixes energy imbalances from the dynamical core, physics-dynamic inconsistencies and physics-dynamics coupling (Lauritzen & Williamson, 2019). This concludes the description of the CAM energy formula in CAM physics.

An important observation: Any total energy formula should be independent of reference state and, in particular, the reference temperature T_{00}. Since we are using constant latent heats, we see that the T_{00} terms on the left and right-hand side of (112) cancel regardless of the value of T_{00} (see derivations in (84)) if all forms of water are included in $m^{(H_2O)}$ and $F^{(H_2O)}_{net}$. So changing reference temperature does not alter the energy equation. Similarly for changing from, e.g., ice to liquid reference state. This can be illustrated through the following manipulations

$$\begin{aligned}
F^{(wv)}L_{s,00} + F^{(liq)}L_{f,00} &= F^{(wv)}\left(L_{v,00} + L_{f,00}\right) + F^{(liq)}L_{f,00}, \text{ using (58)} \\
&= F^{(wv)}L_{v,00} + L_{f,00}\left(F^{(wv)} + F^{(liq)}\right), \\
&= F^{(wv)}L_{v,00} + L_{f,00}\left(F^{(H_2O)} - F^{(ice)}\right), \\
&= F^{(wv)}L_{v,00} - F^{(ice)}L_{f,00} + F^{(H_2O)}L_{f,00}.
\end{aligned}$$

and doing the same manipulation with E_{feom}

$$m^{(wv)}L_{s,00} + m^{(liq)}L_{f,00} = m^{(wv)}L_{v,00} - m^{(ice)}L_{f,00} + m^{(H_2O)}L_{f,00}, \tag{118}$$

we see that the last term in both equations above will cancel in the column integral as in (104). Hence it does not matter which combination of latent heat terms (last two terms in (C5), (C7) or (C9)) are used along with associated reference temperature. This also applies to fluxes sent to the surface components as long as they are internally consistent with the surface component. For example, CAM uses an ice reference state and the ocean model uses a liquid reference state and it is consistent to let the two components internally use $F^{(wv)}L_{s,00}+F^{(liq)}L_{f,00}$ and $F^{(wv)}L_{v,00}-F^{(ice)}L_{f,00}$, respectively, with associated reference temperatures. For variable latent heats it becomes much more complicated as discussed later.

A caveat for the above argument (particular to CAM) is that only water vapor is included in the T_{00} terms on the left and right-hand side of (112) so they do not exactly cancel since the left-hand side term includes "spurious phase transformations" as described above. Hence there is a spurious dependence (although likely small) on the reference state in CAM due to only including water vapor in total water.

Next we would like to give the reader a sense of what the terms in the CAM total energy equation look like in real-world climate simulations. Figure 2 shows spatial plots of one-year average values of terms in (111). The terms are computed using inline diagnostics from an Atmospheric Model Intercomparison Project (AMIP)-like simulation cycling over year 2000 SSTs (more details on the CAM setup is given in Appendix G). Row 1 shows one-year averages of the left- and right-hand sides of (111) vertically integrated in each column as well as the residual. That is, (a) is the average time-tendency of total energy from all CAM parameterizations (assuming total water stays constant) and (b) shows the time-averaged fluxes into each column. Since CAM physics has been constructed (in theory) so that each parameterization satisfies (111), plot (a) and (b) should balance. This is shown in (c) to be true to within 10^{-6} W/m^2. The reader is reminded that we show terms in the total energy equation only for CAM physics processes, with vertically integrated energy tendencies. In steady-state, or equivalently for long-term time averages in climate equilibrium, the energy budget must be closed (Figure 2c), that is, the net energy flux into the column through its upper and lower boundaries (surface energy flux minus ToA energy flux) must be exactly balanced by the physics heat tendencies integrated over the column.

Figures 2 (d-f) show the break-down of total energy flux terms. That is, latent heat fluxes (d,e) and turbulent/sensible and radiative flux (f), respectively. Note that individual latent heat flux terms depend on the reference state (compare the last two terms in (C5), (C7) and (C9)), but the total latent heat balance (adding up left and right-hand side latent heat terms) does not (physically, latent heats are measurable quantities and should therefore not depend on reference state).

While (111) provides a closed energy budget, total water (which in the case of CAM is $m^{(H_2O)} \approx m^{(wv)}$) has not been updated for the kinetic, enthalpy and geopotential terms. Figure 3 shows the "dry-mass adjustment" resulting from updating water vapor in the total water terms. That is, $\Delta\mathcal{I}^{(CAM)}_{m^{(wv)}}$, which, due to only water vapor being included in total water in CAM, has two terms: $\Delta\breve{\mathcal{I}}_{\partial m^{(wv)}/\partial t}$ (spurious phase change term) and $\Delta\widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(wv)}/\partial t}$ (falling precipitation term), see (114). The Figure shows the breakdown of the contributions from the kinetic K, Φ_s and enthalpy ($c_p^{(d)}T$) terms to $\Delta\mathcal{I}^{(CAM)}_{m^{(wv)}}$. First, we note that the adjustment tendency is more than 0.3 W/m^2 globally, and over $\pm$30 W/m^2 locally, as observed in previous atmosphere studies (Guo et al., 2019; Lauritzen & Williamson, 2019; Harrop et al., 2022) and forced ocean simulations (see Appendix A4 in Griffies et al., 2014). While CAM does not have "explicit enthalpy fluxes" is it noted that they are effectively in the global energy fixer (but not necessarily with the correct sign since all

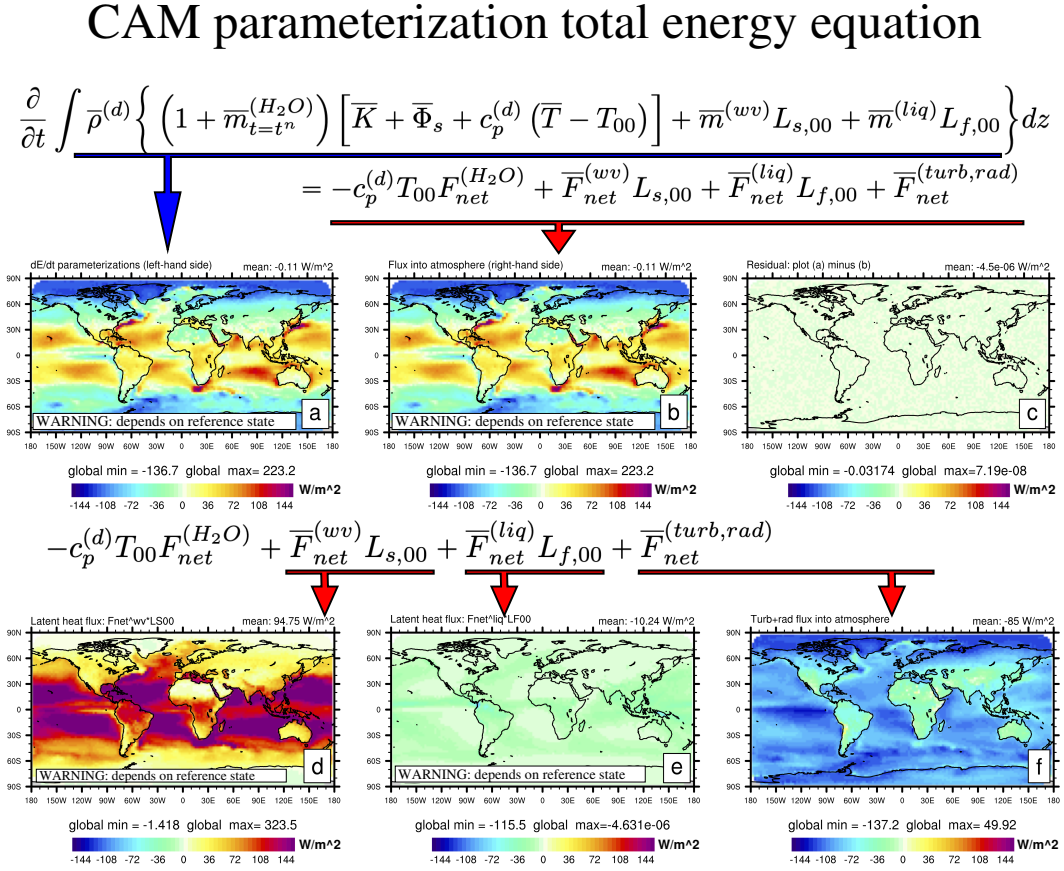


Figure 2. One year average of vertically integrated terms in the CAM parameterization energy budget equation based on an Atmospheric Model Intercomparison Project (AMIP)-like simulation cycling over year 2000 SST's. (a) Left-hand side of (111), (b) right-hand side of (111), (c) residual, (d-e) latent heat flux terms, (111), (f) turbulent/sensible and radiative flux $F^{(turb,rad)}$. Note that plots (a), (b), (d), and (e) depend on the specific reference state used in CAM (ice enthalpy reference state with $T_{00} = 0^{\circ}C$) whereas (c) does not. In the upper right corner of each plot is the global average of the term in question.

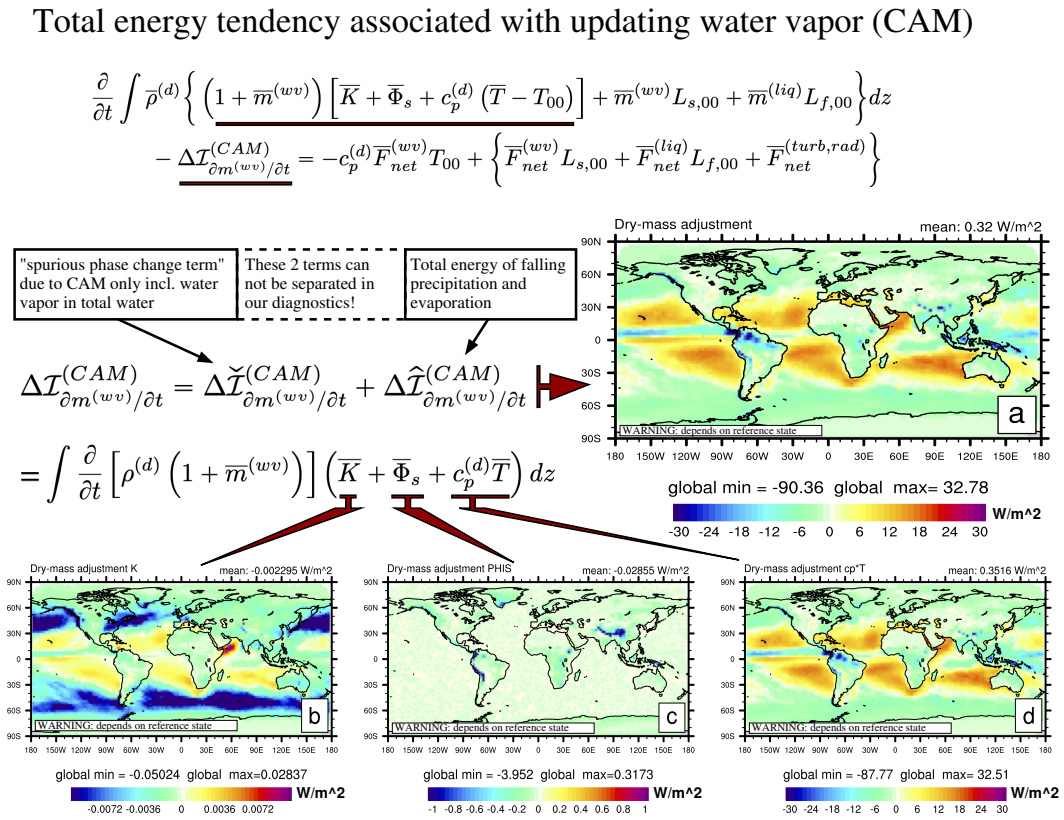


Figure 3. (a) One year average of the total energy budget imbalance in each physics column , $\Delta\mathcal{I}^{(CAM)}_{\partial m^{(wv)}}/\partial t$ $[W/m^2]$, associated with updating total water (which in CAM is only water vapor); see also (113). The global average of $\Delta\mathcal{I}^{(CAM)}_{\partial m^{(wv)}}/\partial t$ is shown in the upper right corner of each plot. The imbalance can be split into a "spurious phase change" term (115) and a term (116) associated with falling precipitation and evaporation. These two terms are, unfortunately, not separable in our diagnostics (the former is, however, likely small). (b,c,d) show the break-down of the kinetic, geopotential and enthalpy terms, respectively. Observe that the enthalpy term is two and three orders of magnitude larger than then kinetic and geopotential energy parts, respectively.

the energy stays in the atmosphere and is not communicated/exchanged with the surface).

In the following we assess the energy tendencies associated with relaxing assumptions in the CAM total energy equation. A natural first step is to include all forms of water in the total energy equation (111), so that there is no spurious phase change term. This leads to the following imbalance

$$\Delta\mathcal{I}^{(CAM)}_{m_{tn}^{(H_2O)}}=\int\left[\rho^{(d)}\left(\sum_{\ell\in\mathcal{L}_{cond}}\overline{m}^{(\ell)}_{tn}\right)\right]\frac{\partial}{\partial t}\left(\overline{K}+\overline{\Phi}_s+c_p^{(d)}\overline{T}\right)dz, \tag{119}$$

(total energy tendency due to non-precipitating condensates)

where $\mathcal{L}_{cond}$ is the set of condensates (non-gases). Basically, this is the total energy tendency of non-precipitating condensates, which in our experiments is, as expected, rather small with a mean of -0.0091 W/m^2 (see Figure 4(a)). Larger values are found in heavy precipitation zones. Note that the energy of this imbalance should stay in the atmosphere.

Henceforth we will neglect $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$ in our quantitative analysis, i.e. assume that CAM includes all forms of water in total water and not just water vapor so that the spurious tendency associated with water-vapor-mass changes in phase transitions are not included in the dry-mass adjustment in (112) (as well as a spurious dependence on reference temperature). Unfortunately, we cannot easily compute the energy tendency due to "spurious" phase transitions in the CAM code, but it seems likely to be similar (or smaller) than the energy tendency of non-precipitating condensates $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$.

If we now compute the dry-mass adjustment, including all water species in (111), we get the following imbalance:

$$\begin{aligned}
\Delta\mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t} &= \Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}} + \int \frac{\partial}{\partial t}\left[\rho^{(d)}\left(\sum_{\ell\in\mathcal{L}_{all}} \overline{m}^{(\ell)}\right)\right]\left(\overline{K} + \overline{\Phi}_s + c_p^{(d)}\overline{T}\right) dz, \\
&= \Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}} + \int \frac{\partial}{\partial t}\left[\rho^{(d)}\left(\sum_{\ell\in\mathcal{L}_{all}} \widehat{\overline{m}}^{(\ell)}\right)\right]\left(\overline{K} + \overline{\Phi}_s + c_p^{(d)}\overline{T}\right) dz, \\
&= \Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}} + \Delta\widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t},
\end{aligned}$$

where in the second line we have used that the time-change of the sum of all water species due to phase-changes is zero: $\frac{\partial}{\partial t}\left(\sum_{\ell\in\mathcal{L}_{all}} \overline{\overline{m}}^{(\ell)}\right) = 0$. The second-term on the right-hand side in the equation is the total energy flux associated with falling precipitation and evaporation. Again, we see the advantages of using all water species in total water in that the dry-mass adjustment does not have spurious energy adjustments due to phase changes. Figure 4 shows the two terms in $\Delta\mathcal{I}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}$: (a) $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$ and (b) $\Delta\widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}/\partial t}$. Note that $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$ is very small compared to the "dry-mass adjustment" term, therefore it may be neglected in this analysis, as stated above.

3.2 Boundary flux terms associated with falling precipitation and water entering atmosphere

In the previous section estimates of the energy changes in the column that either should be compensated for by missing boundary fluxes or missing processes in the atmosphere have been given. First, to understand and analyze the energy fluxes at the surface, it is useful to plot the water fluxes at the surface (since they are paramount to the energy flux terms). Figure 5 (a) shows the net water flux (with positive sign for water entering the atmosphere) and (b,c,d) show the breakdown of the flux into water vapor, liquid and ice fluxes, respectively. Not surprisingly, the liquid and water vapor fluxes dominate the net fluxes (on average).

In the CAM energy equation for each column the neglected flux terms for kinetic, geopotential and enthalpy, are

$$\overline{F}^{(H_2O)}_{net}\left[\overline{\widetilde{K}}_s + \overline{\Phi}_s + c_p^{(d)}\left(\overline{\widetilde{T}}_s - T_{00}\right)\right], \tag{120}$$

Note that in section 2.1.6 the gepotential flux term was approximated by substituting Φ with Φ_s. Essentially that means that we assume that the "$\Phi - \Phi_s$ energy" remains in the atmosphere. We argue that this assumption is physically defensible. For falling precipitation potential energy is converted to kinetic and then into heating as a result of microturbulence dissipation processes (frictional dissipation of falling precipitation; Appendix F).

As discussed in Section 2.1.13, it is problematic to specify the surface kinetic energy and temperature in (120). For the sake of simplicity, considering that surface-air kinetic energy is generally small, and that the surface air temperature often (not always!) is very close to the surface temperature, let's use the kinetic energy in the lowermost level

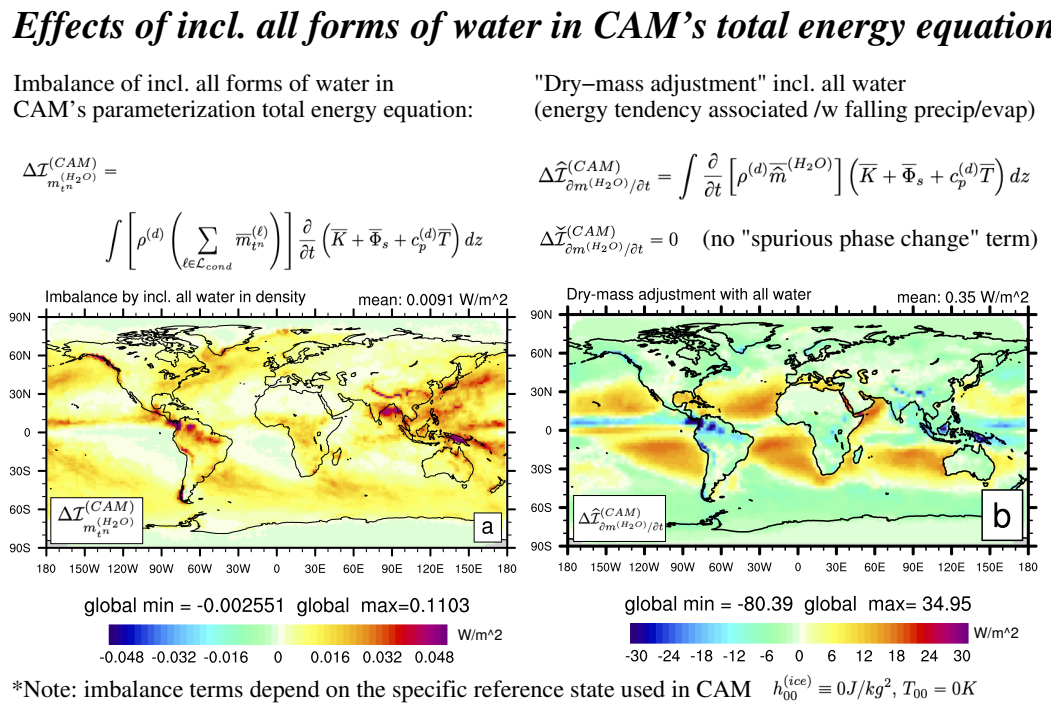


Figure 4. One year average of the total energy budget imbalance when including all forms of water in (a) the CAM parameterization energy equation, $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$ (111), and in (b) the "dry-mass adjustment" term $\Delta\widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}}/\partial t$, respectively. The global average of $\Delta\mathcal{I}^{(CAM)}_{m^{(H_2O)}_{tn}}$ and $\Delta\widehat{\mathcal{I}}^{(CAM)}_{\partial m^{(H_2O)}}/\partial t$ is shown in the upper right corner of each plot. Note that the total energy of condensates is small and the dry-mass adjustment term is comparable to the CAM dry-mass adjustment (Figure 3(a)) where only water vapor is in "total" water.

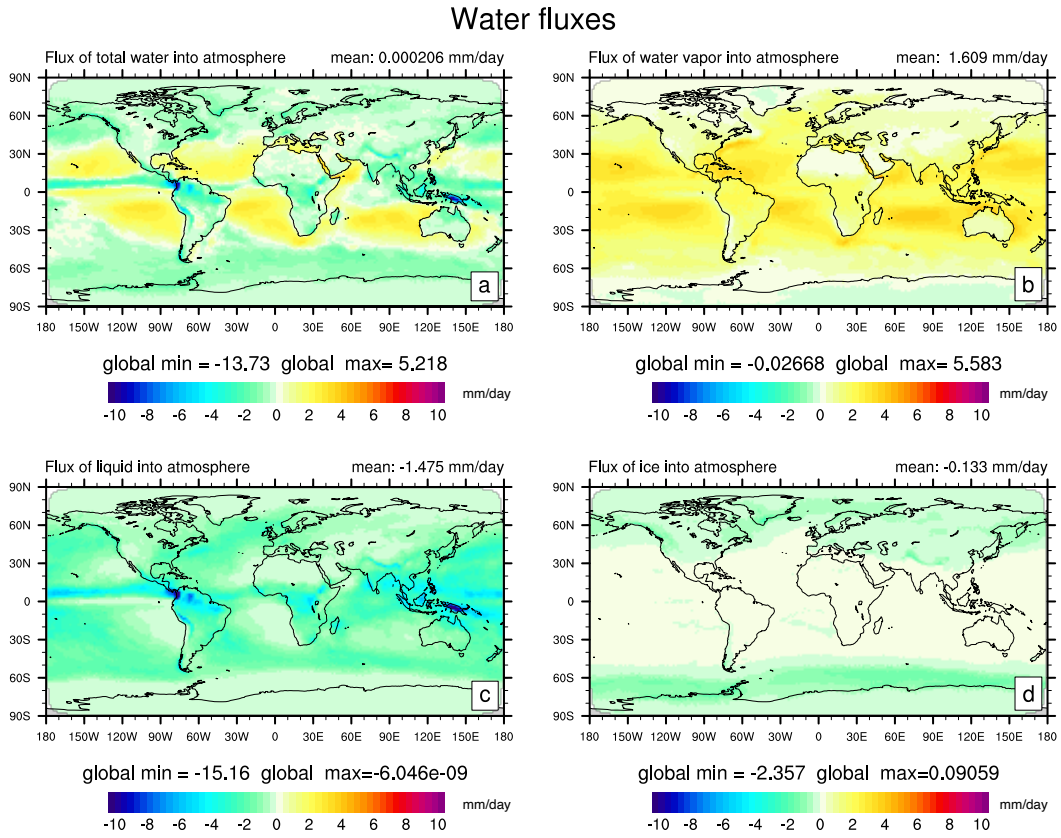


Figure 5. One year average surface fluxes of water (fluxes are positive for inflow into the atmosphere; units are mm/day; in SI units the flux is $kg/m^2/s$ which is converted to mm/day by using that $1cm^3$ of water is $1gm$ and that the length of day is $86400s$). (a) is total net water flux and (b,c,d) is the breakdown into net water vapor, liquid and ice flux, respectively. Note the similarities, in terms of pattern, between the total water flux and the *dry-mass adjustment* in Figures 3d and 4b which accounts for the total energy imbalances associated with water leaving/entering the column.

in the atmosphere, and similarly for temperature. Note that splitting the net material (water) fluxes into downward and upward with different temperatures can easily be done - see section 2.1.4 for the enthalpy flux.

Including these terms in the CAM energy equation, as well as using all forms of water in $m^{(H_2O)}$, yields

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\left\{\left(1+\overline{m}^{(H_2O)}\right)\left[\overline{K}+\overline{\Phi}_s+c_p^{(d)}\left(\overline{T}-T_{00}\right)\right]+\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\right\}dz$$
$$-\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}-\Delta\mathcal{I}_{m_{tn}^{(H_2O)}}=\overline{F}_{net}^{(H_2O)}\left[c_p^{(d)}\left(\widetilde{\overline{T}}_s-T_{00}\right)+\widetilde{\overline{K}}_s+\overline{\Phi}_s\right]$$
$$+\overline{F}_{net}^{(wv)}L_{s,00}+\overline{F}_{net}^{(liq)}L_{f,00}+\overline{F}_{net}^{(turb,rad)} \quad (121)$$

(CAM physics energy equation including surface fluxes of enthalpy, kinetic and geopotential energy)

where $\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}+\Delta\mathcal{I}_{m_{tn}^{(H_2O)}}$ is the imbalance required to close the energy budget. Now that the surface energy terms are not ignored, the imbalance terms are clearly defined (dropping "CAM" superscripts to indicate that the following equations no longer reflect specific CAM approximations)

$$\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}=\Delta\widehat{\mathcal{I}}^{(h)}_{\partial m^{(H_2O)}/\partial t}+\Delta\widehat{\mathcal{I}}^{(K)}_{\partial m^{(H_2O)}/\partial t}+\Delta\widehat{\mathcal{I}}^{(\Phi)}_{\partial m^{(H_2O)}/\partial t},$$

where

$$\Delta\widehat{\mathcal{I}}^{(h)}_{\partial m^{(H_2O)}/\partial t} = \int c_p^{(d)}\left(\overline{T}-T_{00}\right)\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}c_p^{(d)}\left(\widetilde{\overline{T}}_s-T_{00}\right) \quad (122)$$

$$\Delta\widehat{\mathcal{I}}^{(K)}_{\partial m^{(H_2O)}/\partial t} = \int \overline{K}\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\widetilde{\overline{K}}_s \quad (123)$$

$$\Delta\widehat{\mathcal{I}}^{(\Phi)}_{\partial m^{(H_2O)}/\partial t} = \int \overline{\Phi}_s\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\overline{\Phi}_s. \quad (124)$$

It is important to note that these imbalance terms do not depend on reference states. Before discussing the physical meaning of the imbalance terms, we note that including the extra flux terms on the right-hand side of (121) drastically reduces the imbalance in areas where the imbalances without those terms were large ($\Delta\mathcal{I}_{\partial m^{(wv)}/\partial t}$; compare Figure 3(a) and 6(c)). For example, the imbalance in evaporation zones over the ocean, such as off the west coast of South America, USA, Africa and Australia are reduced from $\sim$ 30 W/m^2 to $\sim$ 0.5 W/m^2.

Rows 1, 2 and 3 in Figure 7 shows the first and second term on the right-hand side of (122) (enthalpy), (123) (kinetic energy) and (124) (geopotential energy), respectively, as well as the difference (imbalance). Note that the separate terms depend on somewhat arbitrary reference states (so, on their own, they do not have physical meaning) whereas the residuals (column 3) are not (they have physical meaning).

Why are the residuals in (122), (123) and (124) not zero? A theoretical answer to this question was given in Section 2.1.13. We will repeat the argument here, but backed with values from our simulation. Let's start with the kinetic energy imbalance where, the "dry-mass adjustment" (Figure 7(d)) is an order of magnitude larger (on average) than the surface kinetic energy term (Figure 7(e)). This is due to the falling precipitation being formed higher up in the atmosphere where the winds (i.e. kinetic energy) tend to be larger than the surface winds (that are used to compute the surface K flux in this experiment). In CAM it is assumed that precipitation instantly falls to the ground (as if it was falling through an insulated rain shaft). Assuming that the precipitation does indeed hit the surface with a horizontal velocity component equal to the ambient atmosphere (lower level winds) then $\Delta\widehat{\mathcal{I}}^{(K)}_{\partial m^{(H_2O)}/\partial t}$ represents missing processes, that is, how the precipitation interacts with the ambient air as it falls through the atmosphere to reach the surface layer wind (for example, falling precipitation will tend to slow down resolved-scale horizontal winds).

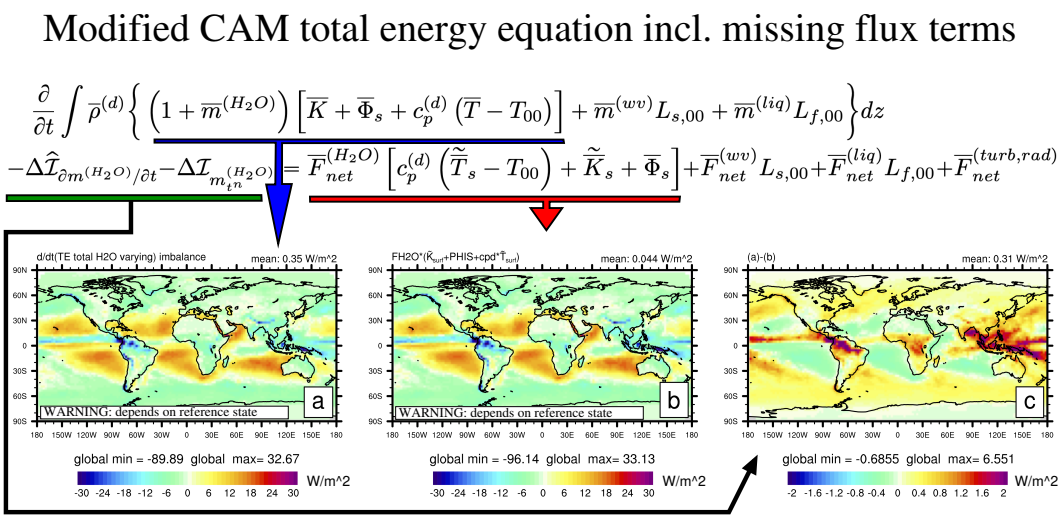


Figure 6. Modified (consistent) CAM total energy equation terms in W/m^2: (a) Imbalance introduced by "dry-mass adjustment" using all forms of water in the kinetic, geopotential and enthalpy terms, (b) missing flux terms and (c) is the difference between (a) and (b). Note that the imbalance is locally much reduced when using the modified total energy equation. Also, the imbalance does not depend on the reference state (as should always be the case).

The interpretation is similar for $\Delta\widehat{\mathcal{I}}^{(h)}$. The temperature at which precipitation was formed is different than the surface temperature and hence there is an imbalance (row 1 in Figure 7). The surface geopotential terms match well (Figure 7(i)) since they depend only on the column mass balance, which is closed in CAM. Please note that, as mentioned in detail in the "**Caution** paragraph" under (113), that this formulation approximates the geopotential flux in terms of Φ_s (see section 2.1.6) and therefore assumes that the $\Phi-\Phi_s$ energy of falling precipitation remains in the atmosphere (through conversion of potential energy to kinetic energy to frictional dissipation/heating).

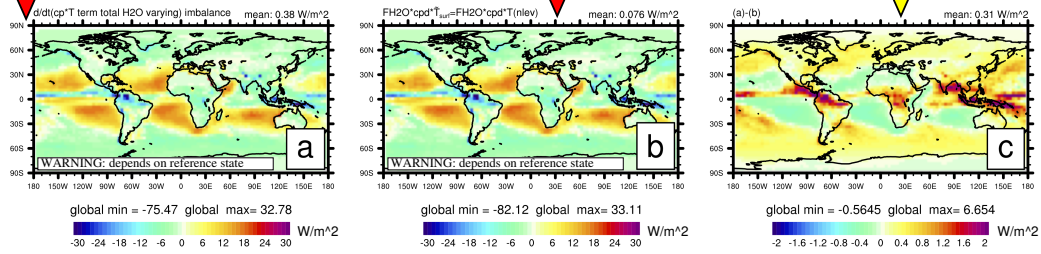
We note that cleverly choosing a surface wind and temperature that exactly matches the "dry mass adjustment" for kinetic energy, $\Delta\widehat{\mathcal{I}}^{(K)}_{\partial m^{(H_2O)}/\partial t}$ and enthalpy $\Delta\widehat{\mathcal{I}}^{(h)}_{\partial m^{(H_2O)}/\partial t}$, in the column would not necessarily be accurate since this assumption ignores processes for falling precipitation (Section 2.1.13). For example, a missing process relevant to the kinetic energy budget is the exchange of momentum between ambient air and falling precipitation as it falls through the atmosphere. This term is largest for high winds and large precipitation rates such as in hurricanes. Similarly, falling precipitation interacts with the environment as it falls (on average falling precipitation will tend to slow down the resolved-scale winds and the falling hydro-meteors tend to heat up the ambient air as they fall). Hence, artificially closing the energy budget by cleverly choosing surface winds and temperature potentially excludes important processes. The energy budget would be closed, but not entirely for the right reasons since all the energy (from where precipitation was formed) is passed to the surface and none stays in the atmosphere. The missing process of frictional heating of falling precipitation (conversion of potential energy to vertical kinetic energy and then to heating through Stokes friction) and its possible incorporation into a model like CAM has been discussed in detail in Appendix F.

The same argument can be made for the reverse process (transfer of water mass from the surface to the atmosphere), e.g., evaporation. At the air-sea (or -lake, or -soil/-canopy) momentum is exchanged in the mass transfer. In this case momentum transfer within the gas phase is extremely efficient and can be safely assumed to be complete.

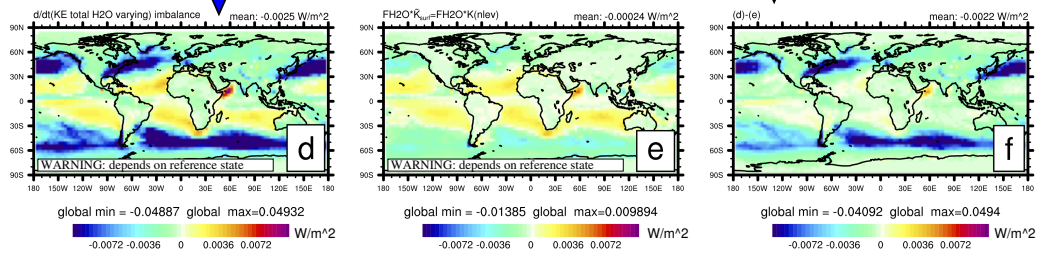
Modified (consistent) total energy equation assuming constant latent heats

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\left\{\left(1+\overline{m}^{(H_2O)}\right)\left[\overline{K}+\overline{\Phi}_s+c_p^{(d)}\left(\overline{T}-T_{00}\right)\right]+\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\right\}dz$$
$$-\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}-\Delta\widehat{\mathcal{I}}_{m_{tn}^{(H_2O)}}=\overline{F}_{net}^{(H_2O)}\left[c_p^{(d)}\left(\widetilde{\overline{T}}_s-T_{00}\right)+\widetilde{\overline{K}}_s+\overline{\Phi}_s\right]+\overline{F}_{net}^{(wv)}L_{s,00}+\overline{F}_{net}^{(liq)}L_{f,00}+\overline{F}_{net}^{(turb,rad)}$$

$$\int c_p^{(d)}\left(\overline{T}-T_{00}\right)\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}c_p^{(d)}\left(\widetilde{\overline{T}}_s-T_{00}\right)=\Delta\widehat{\mathcal{I}}^{(h)}_{\partial m^{(H_2O)}/\partial t}$$



$$\int \overline{K}\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\widetilde{\overline{K}}_s=\Delta\widehat{\mathcal{I}}^{(K)}_{\partial m^{(H_2O)}/\partial t}$$



$$\int \overline{\Phi}_s\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\widehat{\overline{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\overline{\Phi}_s=\Delta\widehat{\mathcal{I}}^{(\Phi)}_{\partial m^{(H_2O)}/\partial t}$$

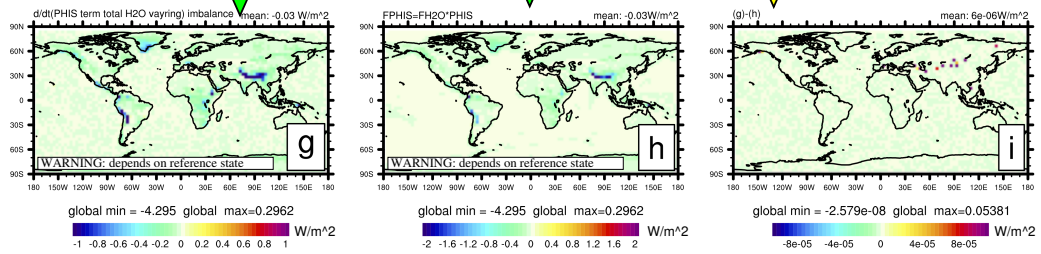


Figure 7. A breakdown of the imbalance term in W/m^2 (shown in Figure 6) associated with (row 1) enthalpy, (row 2) kinetic and (row 3) geopotential energy, respectively. Note that plot (i) should be exactly zero, however, due to truncation errors in computing the diagnostics (and mass clipping - see section 4.4) it is not round-off.

The temperature of water vapor from the surface is likely close to the known surface temperature.

3.3 Latent heats as a function of temperature

Now we move to the most complicated energy equation we will discuss here. Let's no longer assume constant latent heats. In this case the total energy equation becomes

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\Bigg\{\left(1+\overline{m}^{(H_2O)}\right)\left(\overline{K}+\overline{\Phi}_s\right)+c_p^{(d)}T+\sum_{\ell\in\mathcal{L}_{H_2O}}\overline{m}^{(\ell)}c_p^{(\ell)}\left(\overline{T}-T_{00}\right) +\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\Bigg\}dz-\Delta\breve{\mathcal{I}}_{L(T)}-\Delta\widehat{\mathcal{I}}_{L(T)}= -\sum_{\ell\in\mathcal{L}_{H_2O}}\overline{F}_{net}^{(\ell)}\left[c_p^{(\ell)}\left(\widetilde{\overline{T}}_s-T_{00}\right)+\widetilde{\overline{K}}_s\right]+\overline{F}_{net}^{(wv)}L_{s,00}+\overline{F}_{net}^{(liq)}L_{f,00}+\overline{F}_{net}^{(turb,rad)} \tag{125}$$

where the imbalance terms are split into water changes and heating, which stay in the atmosphere, $\Delta\breve{\mathcal{I}}_{L(T)}$, and falling precipitation/evaporation, $\Delta\widehat{\mathcal{I}}_{L(T)}$. These can each be divided into kinetic, enthalpy and geopotential energy imbalances (as done in our previous imbalance equations):

$$\Delta\breve{\mathcal{I}}_{L(T)}^{(K)} = \sum_{\ell\in\mathcal{L}_{all}}\frac{\partial}{\partial t}\int\rho_{t^n}^{(d)}m_{t^n}^{(\ell)}\left(K+c_p^{(\ell)}T\right)dz - \sum_{\ell\in\{'d',wv'\}}\frac{\partial}{\partial t}\int\rho_{t^n}^{(d)}m_{t^n}^{(\ell)}\left(K+c_p^{(d)}T\right)dz, \tag{126}$$

$$\Delta\breve{\mathcal{I}}_{L(T)}^{(h)} = \frac{\partial}{\partial t}\int\left\{\sum_{\ell\in\mathcal{L}_{all}}\overline{\rho}_{t^n}^{(d)}\overline{\breve{m}}_{t^n}^{(\ell)}c_p^{(\ell)}\overline{T}+\overline{\rho}^{(d)}\overline{\breve{m}}^{(wv)}L_{s,00}+\overline{\rho}^{(d)}\overline{\breve{m}}^{(liq)}L_{f,00}\right\}dz, -\int\sum_{\ell\in\{'d',wv'\}}\overline{\rho}_{t^n}^{(d)}\overline{\breve{m}}_{t^n}^{(\ell)}c_p^{(d)}\frac{\partial\overline{T}}{\partial t}-\frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(wv)}\right)L_{s,00}-\frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(liq)}\right)L_{f,00}dz, \tag{127}$$

$$\Delta\breve{\mathcal{I}}_{L(T)}^{(\Phi)} = 0, \tag{128}$$

and

$$\Delta\widehat{\mathcal{I}}_{L(T)}^{(K)} = \int\overline{K}\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\overline{\widehat{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\widetilde{\overline{K}}_s \tag{129}$$

$$\Delta\widehat{\mathcal{I}}_{L(T)}^{(\Phi)} = \int\overline{\Phi}_s\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\overline{\widehat{m}}^{(H_2O)}\right]dz-\overline{F}_{net}^{(H_2O)}\overline{\Phi}_s, \tag{130}$$

$$\Delta\widehat{\mathcal{I}}_{L(T)}^{(h)} = \sum_{\ell\in\mathcal{L}_{H_2O}}\left[\int c_p^{(\ell)}\left(\overline{T}-T_{00}\right)\frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\widehat{m}}^{(\ell)}\right)dz-\overline{F}_{net}^{(\ell)}c_p^{(\ell)}\left(\widetilde{\overline{T}}_s-T_{00}\right)\right] -\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}^{(h)}. \tag{131}$$

First of all we note that (129) and (130) are the same as $\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}^{(K)}$ and $\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}^{(\Phi)}$ in (123) and (124), respectively. Secondly, all the $\Delta\breve{\mathcal{I}}_{L(T)}^{(x)}$ terms in (126) and (127) appear because the heating from momentum mixing and phase changes assume constant latent heats in CAM. For example, the down-gradient diffusion of momentum should lead to a temperature change (frictional heating) in the column (not in each grid-point as discussed extensively in Section 2.1.12) that should satisfy

$$\sum_{\ell\in\mathcal{L}_{all}}\frac{\partial}{\partial t}\int\rho_{t^n}^{(d)}m_{t^n}^{(\ell)}\left(K+c_p^{(\ell)}T\right)dz=0, \tag{132}$$

whereas in CAM (with constant latent heats and only water vapor in total water) the heating due to kinetic energy change in the column satisfies

$$\sum_{\ell\in\{'d',{}'wv'\}} \frac{\partial}{\partial t}\int \rho_{t^n}^{(d)} m_{t^n}^{(\ell)} \left(K + c_p^{(d)} T\right) dz = 0. \tag{133}$$

The difference between (132) and (133) is the imbalance $\Delta\breve{\mathcal{I}}_{L(T)}^{(K)}$ (126). The enthalpy budget has a similar imbalance problem, although it is more complicated. When using variable latent heats the temperature change due to phase changes should satisfy

$$\frac{\partial}{\partial t}\sum_{\ell\in\mathcal{L}_{all}} \overline{\rho}^{(d)}\overline{\breve{m}}^{(\ell)} c_p^{(\ell)}\overline{T} + \frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(wv)}\right) L_{s,00} + \frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(liq)}\right) L_{f,00} = 0,$$

at each grid-point whereas CAM satisfies

$$\sum_{\ell\in\{'d',{}'wv'\}} \overline{\rho}_{t^n}^{(d)}\overline{\breve{m}}_{t^n}^{(\ell)} c_p^{(d)}\frac{\partial\overline{T}}{\partial t} + \frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(wv)}\right) L_{s,00} + \frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(liq)}\right) L_{f,00} = 0.$$

The difference between these two equations is the enthalpy imbalance $\Delta\breve{\mathcal{I}}_{L(T)}^{(h)}$ in (127) which, after expanding and re-arranging terms, can be written as

$$\begin{aligned}\Delta\breve{\mathcal{I}}_{L(T)}^{(h)} &= \int \overline{\rho}_{t^n}^{(d)}\left[\sum_{\ell\in\mathcal{L}_{H_2O}} \overline{m}_{t^n}^{(\ell)} c_p^{(\ell)}\frac{\partial\overline{T}}{\partial t} - \overline{m}_{t^n}^{(wv)} c_p^{(d)}\frac{\partial\overline{T}}{\partial t}\right] dz + \\ &\qquad \int \sum_{\ell\in\mathcal{L}_{H_2O}} \frac{\partial}{\partial t}\left(\overline{\rho}^{(d)}\overline{\breve{m}}^{(\ell)}\right) c_p^{(\ell)}\overline{T}\, dz, \\ &\approx \int \overline{\rho}_{t^n}^{(d)}\left[\sum_{\ell\in\mathcal{L}_{H_2O}} \overline{m}_{t^n}^{(\ell)} c_p^{(\ell)}\frac{\partial\overline{T}}{\partial t} - \overline{m}_{t^n}^{(wv)} c_p^{(d)}\frac{\partial\overline{T}}{\partial t}\right] dz.\end{aligned} \tag{134}$$

The first integral on the right-hand side of (134) is the imbalance due to not using the "correct" heat capacities for all forms of water, combined with the imbalance of only using water vapor in total water. The second integral would have been zero if all water species had the same specific heat since total water is constant for phase transformations. Since a phase change with variable latent heats is "weighted" with different heat capacities this extra term is non-zero. Unfortunately, we cannot compute this term in our diagnostics since we are not able to separate phase-changes from falling precipitation/evaporation. In general, however, the two terms are of opposite sign and of comparable magnitude. E.g. for pure freezing of e.g. supercooled cloud water, the first term is positive ($\sim L_{f,00}\frac{\partial\rho}{\partial t}^{(ice)}$), while the second is negative ($-\frac{\partial\rho}{\partial t}^{(ice)}(c_p^{(liq)}-c_p^{(ice)})T$). So estimating just one will result in an overestimation of the magnitude of the sum.

The global average energy imbalance resulting from using variable latent heats when excluding falling precipitation/evaporation, $\Delta\breve{\mathcal{I}}_{L(T)}$ (excluding the second term in (134) since it can't easily be diagnosed in CAM), is $\sim$ 0.26 W/m^2 (see Figure 8(a)) which, as mentioned in the previous paragraph, is likely an overestimation. The imbalance is more than 1 W/m^2 in heavy precipitation zones where there are large amounts of water vapor phase changing to cloud liquid. Since the specific heat of liquid water is about four times larger than the specific heat of dry air, it is not surprising that areas with large precipitation amounts have the largest signal. This imbalance should remain in the atmosphere and would be zero if CAM physics would have used variable latent heats.

Now let's consider the imbalance terms due to falling precipitation / evaporation. To compute the imbalance terms in (129) and (131) we use that

$$\Delta\widehat{\mathcal{I}}^{(x)} = \Delta\mathcal{I}^{(x)} - \Delta\breve{\mathcal{I}}^{(x)}. \tag{135}$$

Modified (consistent) total energy equation assuming variable latent heats

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}\left\{\underline{\left(1+\overline{m}^{(H_2O)}\right)\left(\overline{K}+\overline{\Phi}_s\right)+c_p^{(d)}T+\sum_{\ell\in\mathcal{L}_{H_2O}}\overline{m}^{(\ell)}c_p^{(\ell)}\left(\overline{T}-T_{00}\right)}+\overline{m}^{(wv)}L_{s,00}+\overline{m}^{(liq)}L_{f,00}\right\}dz$$

$$-\Delta\breve{\mathcal{I}}_{L(T)}-\Delta\widehat{\mathcal{I}}_{L(T)}=-\sum_{\ell\in\mathcal{L}_{H_2O}}\underline{\overline{F}_{net}^{(\ell)}\left[c_p^{(\ell)}\left(\widetilde{\overline{T}}_s-T_{00}\right)+\widetilde{\overline{K}}_s\right]}+\overline{F}_{net}^{(wv)}L_{s,00}+\overline{F}_{net}^{(liq)}L_{f,00}+\overline{F}_{net}^{(turb,rad)}$$

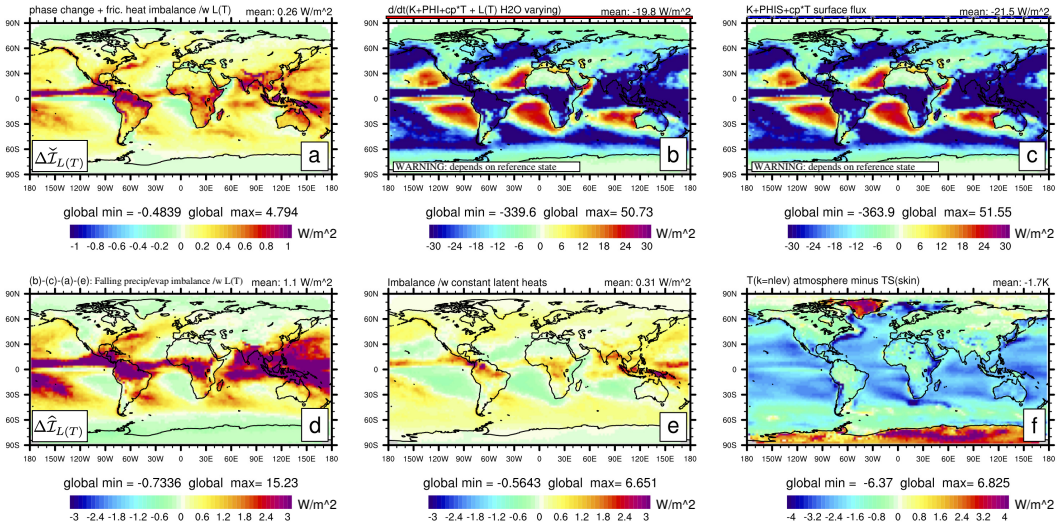


Figure 8. (a) Effect of using variable latent heats while keeping total water constant during physics updates in terms of energy imbalances in W/m^2. (b) Dry-mass adjustment when using variable latent heats. (c) Surface flux terms associated with kinetic, geopotential and enthalpy. (d) Estimated imbalance due to falling precipitation and evaporation. (e) Same as (d) but with constant latent heats (also shown on Figure 6(c)). (f) Difference between the lowest level atmosphere temperature and the skin temperature (which is SST over ocean).

Figure 8(b) shows the first term in the imbalance $\Delta\widehat{\mathcal{I}}_{L(T)}$

$$\sum_{\ell\in\mathcal{L}_{H_2O}}\int\left[\overline{K}+\overline{\Phi}_s+c_p^{(\ell)}\left(\overline{T}-T_{00}\right)\right]\frac{\partial}{\partial t}\left[\overline{\rho}^{(d)}\overline{\widehat{m}}^{(\ell)}\right]dz,$$

and Figure 8(c) the corresponding flux terms

$$\sum_{\ell\in\mathcal{L}_{H_2O}}\overline{F}_{net}^{(\ell)}\left[\widetilde{\overline{K}}_s+\overline{\Phi}_s+c_p^{(\ell)}\left(\widetilde{\overline{T}}_s-T_{00}\right)\right].$$

Note that Figures 8(b) and (c) depend on the specific reference temperature used in CAM. Compared to CAM, the enthalpy terms on the left and right-hand side of the energy equation are now weighted by their actual heat capacities. Compared to $c_p^{(d)}$, $c_p^{(wv)}$ is almost twice as large and $c_p^{(liq)}$ is about four times larger (Appendix G). We see that effect clearly in the Pacific ocean evaporation zones where the terms are 20–30 W/m^2 compared to 10–20 W/m^2 in default CAM with constant (dry air) latent heats. For heavy precipitation zones the imbalance increases by roughly a factor of four, as expected. The mean is now -20 W/m^2 as compared to 0.3 W/m^2 in default CAM. This is again due to the relative weighting of the different phases. In Figure 5 one sees that evaporation and precipitation roughly balance in terms of water flux amount. However, precipitation is weighted twice as much as evaporation in the enthalpy term because the specific heat difference makes the imbalance negative.

Now to the imbalance $\Delta\widehat{\mathcal{I}}_{L(T)}$ of falling precipitation / evaporation with variable latent heats. Note that it is rather complicated to compute $\Delta\widehat{\mathcal{I}}_{L(T)}$ since it is not only the difference between Figure 8(b) and (c) but we need to subtract plot (a), see (135), to subtract the phase change / heating that stays in the atmosphere. Also one needs to subtract plot (e), $\Delta\widehat{\mathcal{I}}_{\partial m^{(H_2O)}/\partial t}$ (this term appears since we are computing these diagnostics based on the modified CAM energy equation (121)). Compared to the constant latent heat case (Figure 8(e)), the imbalance is larger. There are several reasons for this. As mentioned above Figure 8(a) is likely overestimated. Second, the specific heat weighting described above, combined with the ambiguous choice of $\widetilde{\overline{T}}_s$ and $\widetilde{\overline{K}}_s$ leads to a larger dependence on the surface values of T and K. For example, in the tropics the falling precipitation is on average colder than the $\widetilde{\overline{T}}_s$ used here and the ocean is usually warmer than the atmosphere temperature (see Figure 8(f)). It is clear that using variable latent heats (which is inevitably more accurate from a physical stand point) exacerbates the issue of specifying a surface temperature for the surface enthalpy flux. It also highlights the increased need to include missing processes such as frictional heating of falling precipitation to specify more physically correct temperatures for falling precipitation as it exits the atmosphere.

An example of a variable versus constant latent heat inconsistency in a coupled ESM is given in Section 4 of Golaz et al. (2019). The imbalance in this case was approximately 0.5 W/m^2. To avoid spurious energy imbalances in the coupled climate system, the local inconsistency was fixed globally by adding a correction term to the sensible heat flux (Appendix A in Golaz et al., 2019).

The analysis above is based on annual average values and thereby omits temporal variations in the energy budget imbalances. Time dependence of energy imbalances will be discussed briefly in the following section in the context of the global energy fixer in CAM.

3.3.1 Boundary flux terms and their effect on the global energy fixer

The global energy fixer in CAM restores total energy conservation by fixing energy imbalances introduced by the dry-mass energy adjustment ($\Delta\mathcal{I}^{(CAM)}_{\partial m^{(wv)}/\partial t}$ term in (114)),

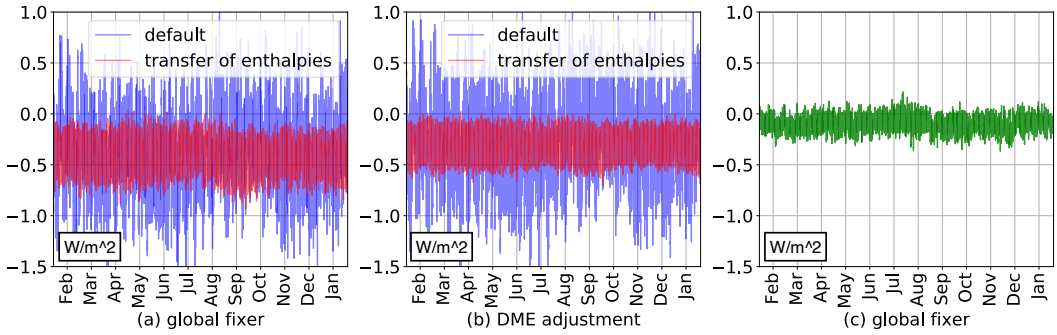


Figure 9. Time series of global energy fixer tendency (in W/m^2) for the three runs described in Section 3.3.1. The blue curves are for the default CAM configuration and the red curves for a run including enthalpy flux terms in the energy equation. The green curve is for a run with energy reset after the dry-mass adjustment, so that it only includes dynamical core and physics-dynamics coupling imbalances. Plot (a) is the total energy imbalance whereas (b) is only the imbalance due to the dry-mass adjustment. Comparing (a) and (b) it is clear that most of the variance in the default model is due to the dry-mass adjustment and including enthalpy fluxes at the surface (despite the ambiguities of choosing a temperature for water leaving/entering the column), the global variance can be reduced significantly. Plot (c) compared to (a) and (b) shows that the dynamical core (and physics-dynamics coupling) variance and mean are small in comparison (this result may be specific to the spectral-element dynamical core as its discretization is mimetic and hence inherently has energy conservation properties).

dynamical core energy truncation errors (discussed further in Section 4.1) and various physics-dynamics coupling errors (discussed further in Section 4.2). The purpose of this section is to show that the temporal statistics of the global fixer largely coincide with those of the dry-mass adjustment. The blue line on Figure 9(a) shows the global energy fixer at each time-step for a 1 year simulation (details of setup in Appendix G). The fixer has a strong diurnal cycle and oscillates between ~ -1.5 W/m^2 and ~ 1 W/m^2. The blue line on Figure 9(b) is the energy imbalance only for the dry-mass adjustment (-$\Delta\mathcal{I}^{(CAM)}_{\partial m^{(wv)}/\partial t}$ term in (114)) which, by eye, matches the energy fixer in Figure 9(a). The averages for the two blue lines is -0.42 W/m^2 and -0.32 W/m^2, respectively (the latter is consistent with Figure 3(a); note opposite sign since the Figure shows -$\Delta\mathcal{I}^{(CAM)}_{\partial m^{(wv)}/\partial t}$) and the standard deviations are 0.41 W/m^2 and 0.44 W/m^2, respectively. In short, most of the imbalance is due to the dry-mass adjustment approximation.

In the second run, we account for the enthalpy of precipitation and evaporation (the last term in equation (120)) in the total energy computation for the global fixer using temperature of the bottom atmospheric layer. Removing this term from the total energy of the atmosphere at the end of the physics step in CAM mimics the transfer of the enthalpy of precipitation from the atmosphere and transfer of the enthalpy of evaporation to the atmosphere. These transfers are currently missing from the model. In their absence, the global fixer recovers corresponding quantities within the atmosphere, with the net amount close to zero.

Although the global, averaged in time, amounts of enthalpies of precipitation and evaporation are close in magnitude, their variations in time are not small. These variations significantly affect the amount of energy distributed by the global fixer. As shown in Figures 9 (a) (red line representing the global energy fixer) and (b) (red line representing the dry-mass adjustment), accounting for the transfer of the enthalpies signif-

icantly reduces the variability in the global fixer and dry-mass adjustment, with standard deviations 0.13 and 0.10 W/m^2, respectively. The averages of these quantities, -0.35 and -0.25 W/m^2, respectively, are close to the default run, which confirms that the averaged effect of the transport of the enthalpies (or its absence) is small.

From Figures 9 (a) and (b) we conclude that, on average, the amount of energy redistributed by the global fixer is close to the amount of the energy flux introduced by the dry-mass adjustment. To confirm this, we perform a third run where we reset the global energy of the model to the energy after the dry-mass adjustment. In other words, in this run the energy fixer corrects only energy errors or fluxes due to the discretizations in the dynamical core and forcings that are applied in the dynamical core (Lauritzen & Williamson, 2019), not due to the dry-mass adjustment. In this case, the curve in Figure 9 (c) of the global fixer has much smaller variability, with standard deviation of 0.07 W/m^2 and average -0.10 W/m^2. Hence, the energy error associated with the dynamical core and physics–dynamics coupling combined is much smaller in magnitude and variability than the error associated with the dry-mass adjustment.

This concludes the analysis of (CAM) physics energy terms and imbalances. We now consider other energy budget errors in ESMs.

4 Other energy budget errors

4.1 Numerical truncation energy errors in dynamical core (adiabatic)

In the absence of space-time truncation errors, the dynamical core conserves the same energy that the continuous equations of motion conserve. In the case of the HPE system, the total energy equations were discussed in Section 2.1.10. However, no operational dynamical cores known to the authors conserves total energy, mostly due to explicit horizontal diffusion of the prognostic variables, filters and other sources of numerical energy error (e.g., Jablonowski & Williamson, 2011). At $\sim 1^\circ$ resolution dynamical cores typically dissipate energy at a rate of 1 W/m^2 (Lauritzen & Williamson, 2019). This energy dissipation can come from intrinsic dissipation (e.g. shape-preserving finite-volume operators) or explicit dissipation. The latter is discussed next.

4.1.1 Explicit diffusion operators and energy

Consider artificial Laplacian diffusion of momentum added to the momentum equations in a HPE model

$$\frac{\partial \vec{v}}{\partial t} = ... + \nu_2 \nabla_h^2 \vec{v}, \tag{136}$$

where ∇_h is applied along coordinate surfaces and ν_2 is an empirical second-order horizontal diffusion coefficient. In some models, this term is added to the equations of motion for numerical stability reasons to control spurious numerical noise at the grid scale. The diffusion term is formally similar to molecular viscosity (a physical process) although it is artificial. It differs from molecular viscosity in that it is not 3D and only applied along coordinate surfaces without a correction term to z-surfaces.

The associated kinetic energy equation (similar to (101)) is

$$\begin{aligned}
\frac{\partial K}{\partial t} = \vec{v} \cdot \frac{\partial \vec{v}}{\partial t} &= ... + \nu_2 \vec{v} \nabla_h^2 \vec{v}, \\
&= ... + \nu_2 \vec{v} \nabla_h \left(\nabla_h \vec{v} \right), \\
&= ... + \nu_2 \left\{ \nabla_h \left[\vec{v} \left(\nabla_h \vec{v} \right) \right] - \left(\nabla_h \vec{v} \right)^2 \right\}, \\
&= ... + \underbrace{\nu_2 \nabla_h^2 \left(K \right)}_{\text{Diffusion of } K} - \underbrace{\nu_2 \left(\nabla_h \vec{v} \right)^2}_{\text{Dissipation of } K},
\end{aligned} \tag{137}$$

using the chain rule for differentiation. The first term on the right-hand side of (137) redistributes kinetic energy (hence the global integral of that term is zero) and the second term denotes the dissipation of kinetic energy (always negative). For a closed energy budget the second term needs to be added to the thermodynamic equation as an artificial frictional heating mechanism. In spherical geometry and with hybrid vertical coordinates the dissipation term must be carefully formulated to guarantee down-gradient diffusion. If the diffusion/dissipation term in the momentum equations can be written as the divergence of a symmetric tensor then one can guarantee down-gradient diffusion (which can be turned into heat). The trick is to derive the symmetric stress tensor which has been done in Becker (2003) and Schaefer-Rolffs and Becker (2013) for Laplacian diffusion operators in a hybrid vertical coordinate system. Unfortunately, Laplacian diffusion operators are considered too diffusive at ESM scales and hence artificial diffusion operators are usually higher-order.

For higher-order operators (e.g., ∇_h^4) it is less obvious how to assign a physical meaning to the terms and separate them into diffusive and dissipative parts as in (137); or even more challenging in spherical terrain-following coordinates. In the absence of a better justified approach, it is common practice for dynamical cores used in climate modeling to add the entire kinetic energy increment due to $\nabla_h^4 \vec{v}$ as heating locally for an improved total energy budget (e.g., Lauritzen et al., 2018). This term is ~ 0.5 W/m^2 so it reduces the energy error by roughly $\sim 50\%$. Since it is not obvious how to separate its diffusive and dissipative parts, this "artificial frictional heating" is more like a local energy fixer than a correction derived from theory (i.e. deriving a symmetric stress tensor). Note that models that use implicit diffusion through limiters and filters in the numerical operators (e.g. Lin, 2004) do not have an explicit expression for diffusion/dissipation, and it is less obvious how to represent frictional heating due to momentum dissipation.

4.1.2 Sponge layer damping and energy conservation

An aspect that needs attention in weather and climate models is the choice of all dissipation processes and, in particular, the treatment of the upper boundary in the atmospheric dynamical core. Most often, dynamical cores apply artificial sponge-layer mechanisms near the model top to damp upwards traveling waves and reduce the wind speeds. These can grow substantially in the upper model domain which is, in particular, true for models with high model lids between 80-150 km. Examples of explicitly-applied sponge-layer dissipation mechanisms in the dynamical core portfolio of NCAR's CESM version 2.2 (CESM2.2) are Rayleigh friction, enhanced second-order horizontal divergence damping, or an energy- and mass-conserving 2-Δz filter in CAM's cubed-sphere finite-volume dynamical core FV3 (L. M. Harris et al., 2021). In addition, CAM's spectral transform Eulerian (EUL, Neale et al., 2010) and the Spectral Element (SE) dynamical cores (Lauritzen et al., 2018) apply enhanced second-order horizontal diffusion near the model top. Some models like ICON, ICON-IAP and MPAS use the gravity wave damping mechanism which is inherent in the vertical solver (Klemp et al., 2008).

Regardless of the chosen dissipation method, once it is applied to the velocity components or the related horizontal divergence and relative vorticity fields it reduces the kinetic energy of the flow. This triggers the question of whether lost kinetic energy should be converted to heat in an effort to conserve the total energy. And if yes, how? Various algorithms are possible for this energy conversion which can either target a local energy restoration or a global energy fix.

Here, we focus on the local energy restoration that is applied after an explicitly-applied dissipation process. In general, the kinetic energy loss due to dissipation $\delta K < 0$ can be assessed via the approach

$$\delta K \quad = \quad K_{after} - K_{before} \tag{138}$$

$$
\begin{aligned}
&= \frac{\rho}{2}\left(\left(\vec{v}+\delta\vec{v}\right)^2-\vec{v}^{\,2}\right)\\
&= \frac{\rho}{2}\left(\vec{v}^{\,2}+2\vec{v}\cdot\delta\vec{v}+\left(\delta\vec{v}\right)^2-\vec{v}^{\,2}\right)\\
&= \rho^{(all)}\,\vec{v}\cdot\delta\vec{v}+\frac{1}{2}\,\rho^{(all)}\,\delta\vec{v}\cdot\delta\vec{v} \qquad (139)
\end{aligned}
$$

which compares the kinetic energy at a chosen grid point after and before the dissipation is applied. The symbol $\rho^{(all)}$ stands for the density and is treated as a constant. This is typically justified since the velocity dissipation mechanisms do not change the mass or pressure at a grid point. In addition, $\vec{v}$ denotes the velocity vector, and $\delta\vec{v}$ describes the reduction of the velocity components. All components of $\delta\vec{v}$ are negative since the dissipation slows down the wind speeds. The exact form of the δK computation depends on the dissipation mechanism, time-stepping approach and the dynamical core design, e.g., whether a hydrostatic or nonhydrostatic design is chosen. The details are not important for the generic discussion here. However, a concrete example for Rayleigh friction is shown in Appendix H.

The kinetic energy change ($-\delta K > 0$) needs to get converted to heat which, as before, needs to take the design of the dynamical core into account. For example, when picking a hydrostatic HPE model with a pressure-based vertical coordinate and the horizontal velocity vector $\vec{v} = (u, v)^T$ with the zonal and meridional wind components u and v the conversion utilizes (assuming a pressure-based vertical coordinate)

$$
\rho^{(all)}\,c_p^{(all)}\,\delta T = \rho^{(all)}\,c_p^{(all)}\,(T_{conserve}-T) = -\delta K. \qquad (140)
$$

The resulting, slightly warmer temperature $T_{conserve}$ then replaces the temperature T locally via

$$
\begin{aligned}
T_{conserve} &= T-\frac{\delta K}{\rho^{(all)}\,c_p} \qquad (141)\\
&= T-\frac{1}{c_p}\left(\vec{v}\cdot\delta\vec{v}+\frac{1}{2}\,\delta\vec{v}\cdot\delta\vec{v}\right)\\
&= T-\frac{1}{c_p}\left(u\delta u+v\delta v+\frac{1}{2}\,(\delta u)^2+\frac{1}{2}\,(\delta v)^2\right) \qquad (142)
\end{aligned}
$$

in an effort to conserve the total energy at the chosen grid point. A nonhydrostatic dynamical core design will require the use of the 3D velocity vectors $\vec{v} = (u, v, w)^T$ and $\delta\vec{v} = (\delta u, \delta v, \delta w)^T$ which adds the contribution of the vertical velocity component w.

It is also possible to approach the problem differently and utilize the chain rule for differentiation to assess the kinetic energy change. This leads to

$$
\frac{\partial K}{\partial t} = \frac{\partial}{\partial t}\left(\frac{1}{2}\,\rho^{(all)}\,\vec{v}^{\,2}\right) = \rho^{(all)}\,\vec{v}\cdot\frac{\partial\vec{v}}{\partial t} \qquad (143)
$$

or equivalently

$$
\delta K = \rho^{(all)}\,\vec{v}\cdot\delta\vec{v}. \qquad (144)
$$

Equation (144) apparently differs from (139). However, in models decisions need to be made about the time discretization, and $\vec{v}$ can be chosen in various ways. For example, a centered-in-time expression $\widetilde{\vec{v}} = \vec{v} + \frac{1}{2}\delta\vec{v}$ can be selected which updates the velocity vector (before dissipation) with half the contribution from the dissipative process. This leads to the kinetic energy change

$$
\begin{aligned}
\delta K &= \rho^{(all)}\,\widetilde{\vec{v}}\cdot\delta\vec{v}\\
&= \rho^{(all)}\left(\vec{v}+\frac{1}{2}\delta\vec{v}\right)\cdot\delta\vec{v} \qquad (145)\\
&= \rho^{(all)}\,\vec{v}\cdot\delta\vec{v}+\frac{1}{2}\,\rho^{(all)}\,\delta\vec{v}\cdot\delta\vec{v} \qquad (146)
\end{aligned}
$$

which recovers the form of (139). The hydrostatic CAM EUL dynamical core, CAM's SE dynamical core (Lauritzen et al., 2018), and the SE version embedded in the Department of Energy (DOE) Energy Exascale Earth System Model version 1 (E3SMv1, Rasch et al., 2019) utilize (144) with the original $\vec{v}$ from a previous time level. They thereby apply the temperature correction

$$T_{conserve} = T - \frac{1}{c_p}\left(u\delta u + v\delta v\right) \tag{147}$$

to mimic frictional heating due to their hyper-diffusion mechanisms (note that Lauritzen et al. (2018) reported an additional factor of 1/2 for this temperature update (their (59)) which was a typographical error). In contrast, the CESM2.2 variant of SE implements the time-centered velocity approach $\widetilde{\vec{v}}$ which leads to the temperature correction shown in (142).

Whether frictional heating should be applied, especially at the local scale and not as a global total energy fixer, is debatable as it was outlined for the Laplacian diffusion operator above. When the dissipation is treated as a physical process one can find arguments for it. However, if the dissipation is simply used to prevent numerical artifacts, such as the reflection of waves at the model top, then frictional heating becomes questionable and might even lead to artificial temperature signals. This is demonstrated for a Rayleigh friction (RF) sponge layer in Figure 10 that shows the time-mean and horizontal-mean temperature of the FV3 dynamical core (L. M. Harris et al., 2021) in a dry Held–Suarez test case configuration without topography (Held & Suarez, 1994). In general, the Held–Suarez test replaces the physical parameterization package of an atmospheric model with a Newtonian temperature relaxation and Rayleigh friction at low-lying levels below 700 hPa. This thereby mimics the effects of radiation and the planetary boundary layer mixing, and enables climate-like model assessments of a dynamical core.

In the example below, FV3's resolution was set to C24 (400 km) with 64 vertical levels that reach up to about 100 km. The vertical grid placement utilized the level configuration of NCAR's Whole Atmosphere Model WACCM6 with 70 levels (Gettelman et al., 2019), but without reaching up to WACCM6's model top at about 140 km. Rayleigh friction can formally be represented by $\delta\vec{v} = -k\vec{v}\Delta t$ where Δt is the length of the time step and k is the RF damping coefficient. In FV3 RF is activated above a user-chosen cutoff pressure p_c and utilizes the user-selected maximum RF timescale τ. We choose $p_c = 10$ Pa and the three τ timescales 1, 3 and 5 days for demonstration purposes. These settings activate the Rayleigh friction in the top 14 levels between 10^{-1}–10^{-4} hPa with increasing strength. The vertical profiles of the damping strength k are shown in Figure 10(b), and the underlying equation is documented in L. M. Harris et al. (2021) and Appendix H. Appendix H also highlights that a fully time-implicit algorithm is used to compute the induced frictional heating in FV3.

In FV3, local frictional heating due to RF in the sponge layer is applied by default. Here, we turn the frictional heating either on or off to document its impact on the time-mean horizontal-mean temperature in the upper atmosphere. This is demonstrated in Figure 10(a) that shows the 6-month temperature averages in the upper domain after an initial 12-month spin-up period. This spin-up period is needed to reach an equilibrium with the Held–Suarez forcing when starting from an isothermal state at rest, as done here. Figure 10(a) documents the model response to three RF damping timescales with and without frictional heating. The figure shows that the damping timescale τ has a decisive impact on the mean temperature in the sponge layer above 0.1 hPa. In general, stronger RF damping with shorter τ timescales and thereby increased Rayleigh friction coefficients k (see Figure 10(b)) leads to enhanced frictional heating effects in FV3. This is reflected by the higher mean temperatures with increasing RF strength (dashed curves). This is expected behavior and not the focus of our discussion. Here we highlight the large temperature differences for an identical τ setting that are solely due to the application or non-application of frictional heating. These are depicted by the dashed and solid curves

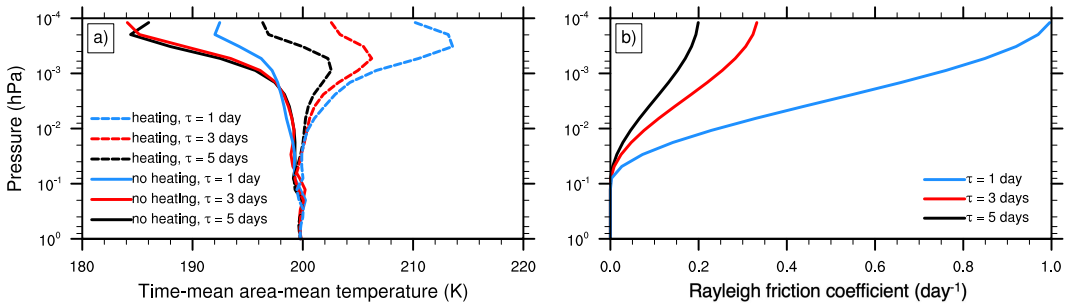


Figure 10. (a) 6-month-mean horizontal-mean vertical temperature profiles of six Held-Suarez simulations with the FV3 dynamical core at the resolution C24L64 (400 km, 100 km model top, CESM framework). Three maximum damping timescales τ are depicted with (dashed) and without (solid) frictional heating due to Rayleigh friction. (b) Vertical profiles of the Rayleigh friction damping coefficient k for the three maximum timescales τ and $p_c = 0.1$ hPa.

with identical colors. Differences of up to 15-20 K are apparent near the model top. These temperature differences due to frictional heating become even more pronounced above 10^{-4} hPa. If, for example, WACCM6's full vertical domain with a model lid around 140 km is utilized temperature differences of over 100 K were found near the model top due to frictional heating (not shown). However, a high sensitivity to other FV3 sponge-layer mechanisms like an enhanced second-order divergence damping (used by default in FV3 in the two topmost layers) and an optional 2-Δz filter was also found. In the experiments here the 2-Δz filter is not activated and the enhanced second-order divergence damping mostly affects the topmost kinks of the temperature profiles in Figure 10(a). The general spread between the simulations with and without frictional heating remains unaffected by the top-layer divergence damping.

None of the depicted simulations with and without frictional heating can be considered a reference solution for the Held-Suarez test. The Held–Suarez test relaxes the temperature above 100 hPa to 200 K so that the closeness to the 200 K isotherm can be used to judge the temperature profiles. The FV3 simulations start deviating from this isotherm rather quickly once the Rayleigh friction is activated above 0.1 hPa. This illustrates the impact of frictional heating in FV3, raises awareness, and encourages continued discussions about energy conservation and the use of frictional heating processes in atmospheric models.

In conclusion, we would like to highlight that Rayleigh damping has no physical meaning since it is ad-hoc in form and not justifiable on the basis of physical theory. Hence we argue that there should be no physical heating/cooling term associated to it in the thermodynamic equation. Also, Laplacian damping introduced near the upper boundary to prevent reflection of waves is purely artificial (introduced due to an artificial upper boundary) and does not represent a physical dissipation process. Therefore, the momentum that is being dissipated should not be converted into a heating/cooling term in the thermodynamic equation.

4.1.3 Energy truncation errors associated with discretizing the inviscid HPEs

In addition to artificial diffusion operators, the choice of discretization of the inviscid equations of motion often affect the energy budget. Nevertheless, energy conserving numerical methods, called mimetic discretizations, do exist (e.g., Taylor & Fournier, 2010; Dubos et al., 2015; Gassmann, 2013; Taylor et al., 2020; Eldred et al., 2019) (see

Section 5.2.1). Briefly, one discretizes the Hamiltonian and the Poisson brackets themselves, rather than the associated discrete equations of motion. The discrete equations of motion are then derived from the discrete Hamiltonian and Poisson brackets. If the Poisson brackets are discretized in a way that preserves the key elements of their geometric structure (anti-symmetry and some of the null space), then the resulting equations of motion will inherit the desired conservation properties. The resulting Poisson system of ODEs can then be integrated in time using an energy-conserving time integrator, such as a discrete gradient method (Cohen & Hairer, 2011). This latter step is rarely taken (although see Eldred et al., 2019), since these integrators are in general fully implicit and therefore too expensive for practical atmospheric models. In addition, the energy time-truncation errors for inviscid dynamics are negligible (~ -0.005 W/m^2), making it hard to justify the increased computation expense of energy-conserving time-stepping. In contrast, spatial mimetic discretizations are currently being used in some operational climate dynamical cores (e.g., Taylor, 2011; Taylor & Fournier, 2010; Dubos et al., 2015; Gassmann, 2013; Taylor et al., 2020; Melvin et al., 2019).

In the absence of time-truncation errors, vertical remapping errors and explicit diffusion the spectral-element model conserves total energy. It has been found necessary for practical applications to employ explicit diffusion operators that dissipate energy, because the nonlinear (inertial) energy cascade extends beyond the resolved scales. The advantage of mimetic methods is that they help distinguish between the reversible and irreversible parts of a model.

In contrast, energy-conserving vertical remapping has been found to be problematic. Hence, at least for the immediate foreseeable future, dynamical cores (even those using mimetic discretizations) require a total energy fixer to close their energy budget. Nevertheless, mimetic discretizations can reduce the amount of spurious energy that must be handled in the fixer. Figure 9(c) shows the global energy fixer for the spectral-element dynamical core (and physics dynamics coupling) as a function of time-step. Compared to the dry-mass adjustment, the spurious tendency from the dynamical core (and physics-dynamics coupling) is much less (~ -0.1 W/m^2) in terms of standard deviation and mean.

4.2 Physics-dynamics coupling errors (spatial and temporal)

Assuming that physics and dynamics are based on the same equations of motion and thermodynamics, how physics and dynamics are coupled has a significant impact on the energy and mass conservation properties. Adding physics tendencies throughout the dynamical core time-integration can produce energy errors (discussed in detail in Lauritzen & Williamson, 2019). The applicability of parallel coupling of physics and dynamics is limited by mass and energy errors (Donahue & Caldwell, 2020). If physics tendencies are interpolated from a physics grid to a dynamical core grid (and the two grids differ), then the interpolation can lead to energy errors (Herrington et al., 2018). There can also be physics–dynamics coupling errors if the continuous equations of the dynamical core conserve a total energy that is different from the one conserved in the physics. This could be characterized as a consistency error. An example is given in Lauritzen and Williamson (2019). For an extensive discussion on physics–dynamics coupling please see Gross et al. (2018). An example of similar issues in water mass conservation can be found in Zhang et al. (2018). Since coupling errors have already been discussed in the literature, and the scope of this paper must be limited, these errors will not be discussed further.

4.3 Thermodynamic conserved variable inconsistency leading to total energy errors

Richard Feynmann remarked that (Feynman et al., 1989),

> The subject of thermodynamics is complicated because there are so many different ways of describing the same thing ... with respect to the internal energy $U^{(all)}$, we might say that it depends on the temperature and volume, if those are the variables we have chosen - but we might also say that it depends on the temperature and pressure, or the pressure and volume, and so on.

Not surprisingly various different conserved variables have been used in moist thermodynamics (see, e.g., Emanuel, 1994), and parameterizations may adopt one rather than another based on what makes most sense for that particular parameterization. This may lead to inconsistencies with the total energy formula. These inconsistencies occur when a parameterization is based on a conserved variable which is different from that of the host model. From first principles both sets of conserved variables are correct, but they do not lead to the same energy conservation formula.

Confused? Below is an example from the CLUBB parameterization of turbulence and clouds (Larson, 2017), which has been implemented in CAM (Bogenschutz et al., 2013). This example illustrates the intricacies of this problem, which are likely to be present in some form or extent in many comprehensive parameterization suites (as far as the authors are aware). Please note that this is not a discretization error, but rather an error arising from the choice of conserved variables. Since this subject has received little attention in the literature, it will be discussed in some detail in the following section.

4.3.1 An example from CAM: CLUBB

The CAM interface with parameterizations conserves (111). That is, the total enthalpy plus kinetic energy in each column satisfies

$$\frac{\partial}{\partial t}\int \overline{\rho}^{(d)}_{t=t^n}(1+\overline{m}^{(wv)}_{t=t^n})\left[c_p^{(d)}\overline{T}+\overline{K}-L_{v,00}\overline{m}^{(liq)}-L_{s,00}\overline{m}^{(ice)}\right]dz=\overline{F}^{(rhs)}, \tag{148}$$

$$\text{(assume (109), } c_p^{(\ell)}=c_p^{(d)},\ \widetilde{h}^{(wv)}_{00}\equiv 0 J/kg^2,\ \overline{m}^{(H_2O)}=\overline{m}^{(wv)},\ \overline{m}^{(H_2O)}=\overline{m}^{(wv)}_{t=t^n}).$$

(Williamson et al., 2015) where, for notational clarity, the right-hand side terms in (111) are denoted $\overline{F}^{(rhs)}$ and, in order to facilitate a comparison of enthalpy with θ_l below, a water vapor reference state is assumed (despite CAM using an ice reference state). The thermodynamic potential (which in CAM is a conserved quantity) is used to update temperature as a result of heating in CAM parameterizations (assuming a water vapor reference state). This thermodynamic potential is

$$h^{(CAM)}_{h^{(wv)}_{00}}\equiv c_p^{(d)}T-L_{v,00}\overline{m}^{(liq)}-L_{s,00}\overline{m}^{(ice)}. \tag{149}$$

Also note that since CLUBB conserves water,

$$C\equiv L_{s,00}m^{(wv)}+L_{s,00}m^{(liq)}+L_{s,00}m^{(ice)}=\text{constant} \tag{150}$$

which, when added to (149), yields

$$h^{(CAM)}_{h^{(ice)}_{00}}=h^{(CAM)}_{h^{(wv)}_{00}}+C=c_p^{(d)}T+L_{s,00}\overline{m}^{(wv)}+L_{f,00}\overline{m}^{(liq)}, \tag{151}$$

(CAM thermodynamic potential; ice reference state)

after rearranging terms and using (58). The right-hand side of (151) is now the enthalpy with respect to an ice reference state, which is what is used in CAM. Hence, as long as total water is conserved, one can seamlessly switch between reference states (up to a constant which disappears in the time-derivative). Note that (150) does not hold in each grid-point if vertical mixing is included. It does, however, hold in each column since vertical mixing conserves mass.

CLUBB is a model based on the anelastic approximation for the fluid equations (Larson, 2017). Rather than enthalpy, CLUBB's prognostic thermodynamic variable is liquid water potential energy, θ_l, which is conserved under adiabatic transformations in the relevant approximations. The use of a different thermodynamic variable in CLUBB (θ_l) than in CAM ($h^{(CAM)}_{h^{(wv)}_{00}}$) leads to two potential problems.

First, the mapping between θ_l and $h^{(CAM)}_{h^{(wv)}_{00}}$ needs to be handled correctly (see Section 4.3.2). Secondly, even if the mapping is handled correctly, conservation of θ_l violates conservation of $h^{(CAM)}_{h^{(wv)}_{00}}$ because the two variables are weighted differently in the vertical by the Exner function (discussed below). Because of such an inconsistency, the combined model (CAM-CLUBB) conserves neither of the two variables and the discrepancy needs to be resolved using an energy fixer. This situation arose because of a historical accident, namely, CLUBB was originally developed in the context of a different model that does conserve θ_l and was only later ported to CAM. The same problem also occurs in the Energy Exascale ESM (E3SM, Golaz et al., 2019).

In detail, CLUBB transports an approximate form of the conserved moist potential temperature θ_l (see Tripoli & Cotton, 1981; Cotton et al., 2011), which is defined as

$$\theta_l \equiv T\Pi^{-1} - \frac{L_{v,00}}{c_p^{(d)}}\Pi^{-1} m^{(liq)}, \tag{152}$$

where Π is the Exner function, which is purely a function of pressure. CLUBB then returns to CAM the following tendency of θ_l,

$$\overline{\rho}^{(d)}(1+\overline{m}^{(wv)})\left.\frac{\partial \overline{\theta_l}}{\partial t}\right|_{\text{CLUBB}} = -\frac{\partial}{\partial z}\left[\overline{\rho}^{(d)}(1+\overline{m}^{(wv)})\,\overline{w'\theta_l'}\right], \tag{153}$$

which, if integrated in the vertical, yields zero, aside from fluxes at the upper and lower boundaries of the atmosphere. In (153) the $(\cdot)'$ variables are deviations from the grid cell mean values and w is vertical velocity.

Therefore, if CLUBB is advanced one time step, then the profiles of θ_l before and after the call to CLUBB are related by

$$\int \left(\overline{\theta_l}\right)_{t^{n+1}} \overline{\rho}^{(d)}_{t^n}(1+\overline{m}^{(wv)}_{t^n})\,dz = \int \left(\overline{\theta_l}\right)_{t^n} \overline{\rho}^{(d)}_{t^n}(1+\overline{m}^{(wv)}_{t^n})\,dz, \tag{154}$$

or

$$\int \Delta\overline{\theta_l}\,\overline{\rho}^{(d)}_{t^n}(1+\overline{m}^{(wv)}_{t^n})\,dz = 0, \tag{155}$$

where $\Delta\overline{\theta_l} \equiv \left(\overline{\theta_l}\right)_{t^{n+1}} - \left(\overline{\theta_l}\right)_{t^n}$, given the usual CAM assumptions of total water and pressure staying constant during physics updates. Substituting the definition of θ_l (152) into (155), and multiplying by $c_p^{(d)}$, we find

$$\int \frac{1}{\Pi_{\,t^n}}\left(c_p^{(d)}\Delta\overline{T} - L_{v,00}\Delta\overline{m}^{(liq)}\right)\overline{\rho}^{(d)}_{t^n}(1+\overline{m}^{(wv)}_{t^n})\,dz = 0. \tag{156}$$

Now, consider the host model CAM. Its conserved prognostic thermodynamic variable is enthalpy, (149), which should be updated according to (148). Since CLUBB does not alter $m^{(ice)}$ until CLUBB's diffusion acts on it in a different part of the code, we will omit the last term in (148) and take

$$h^{(CAM)}_{h^{(wv)}_{00}} \equiv c_p^{(d)}T - L_{v,00}m^{(liq)}. \tag{157}$$

If CLUBB were to conserve $h^{(CAM)}_{h^{(wv)}_{00}}$, then advancing CLUBB one time step would yield

$$\int \Delta\overline{h^{(CAM)}_{h^{(wv)}_{00}}}\,\overline{\rho}^{(d)}(1+\overline{m}^{(wv)}_{t^n})\,dz = 0, \tag{158}$$

in the absence of fluxes through the top and bottom of domain. We write this in terms of T by using the definition of $h^{(CAM)}_{h^{(wv)}_{00}}$ (157), which yields

$$\int (c_p^{(d)} \Delta \overline{T} - L_{v,00} \Delta \overline{m}^{(liq)}) \, \overline{\rho}^{(d)}_{t^n} (1 + \overline{m}^{(wv)}_{t^n}) \, dz = 0. \tag{159}$$

What CLUBB does (see equation (156)) is not what is required for conservation of CAM's $h^{(CAM)}_{h^{(wv)}_{00}}$ (see equation (159)). CLUBB's integral differs because it includes Π^{-1}, which preferentially weights CLUBB's tendencies in the upper part of the atmosphere. The resulting energy imbalance in each column is shown in Figure 11(c). The global average tendency is rather large ~ 0.4 W/m^2, and similar in magnitude to other leading energy imbalances/errors in the system. Currently, this imbalance is restored in each column as a uniform heating increment (but only in the layers where CLUBB is active). To avoid this inconsistency in CAM, CLUBB could use $h^{(CAM)}_{h^{(wv)}_{00}}$ rather than θ_l as its thermodynamic potential.

Aside: Figure 11(a) shows the total imbalance and (b) the kinetic energy imbalance in CAM-CLUBB. The kinetic energy imbalance is non-zero since CLUBB does not include frictional heating (also discussed in the context of dynamical cores). The kinetic energy budget would be closed if local heating were applied where momentum is dissipated (through mixing/turbulence). This missing heating is large (~ 1.8 W/m^2).

4.3.2 Thermodynamic inconsistency in sensible heat flux in CAM-CLUBB

Another thermodynamic inconsistency in CAM-CLUBB is that the sensible heat flux, $F^{(turb)}$, provided by the host model, CAM, is passed incorrectly to CLUBB.

Neglect kinetic energy (i.e. assume for the moment that CLUBB does not alter winds), neglect radiation, and assume that there are no phase changes. Then CAM's energy equation reduces to:

$$\frac{\partial}{\partial t} \int \left\{ \overline{\rho}^{(d)} \left[1 + \overline{m}^{(wv)}_{t=t^n} \right] c_p^{(d)} \overline{T} \right\} dz = \overline{F}^{(turb)}_{net}. \tag{160}$$

In contrast, CLUBB conserves

$$\frac{\partial}{\partial t} \int \left\{ \overline{\rho}^{(d)} \left[1 + \overline{m}^{(wv)}_{t=t^n} \right] \theta_\ell \right\} dz = \overline{\rho}^{(d)} \left[1 + \overline{m}^{(wv)}_{t=t^n} \right] \overline{w' \theta'_l} \Big|_{\text{surface}} \equiv \overline{\mathcal{F}}^{(turb)}_{net}. \tag{161}$$

That is, CLUBB conserves a potential temperature variable rather than temperature. In the absence of phase changes, (161) becomes

$$\frac{\partial}{\partial t} \int \left\{ \overline{\rho}^{(d)} \left[1 + \overline{m}^{(wv)}_{t=t^n} \right] \frac{T}{\Pi} \right\} dz = \overline{\mathcal{F}}^{(turb)}_{net}. \tag{162}$$

In order for the sensible heat flux to result in the same temperature change in the column, then the CAM turbulent heat flux should be scaled by the surface Exner function, Π_s, so that

$$\overline{\mathcal{F}}^{(turb)}_{net} = \frac{\overline{F}^{(turb)}_{net}}{c_p^{(d)} \Pi_s}, \tag{163}$$

(in terms of the actual CAM CLUBB interface code, the actual flux passed to CLUBB is expected to be in units of $[K/m^2]$, and so $\overline{F}^{(turb)}_{net}$ is divided by the air density at the surface).

In the current coupling of CLUBB with CAM, the following flux is erroneously passed to CLUBB

$$\frac{\overline{F}^{(turb)}_{net}}{c_p^{(d)}}, \tag{164}$$

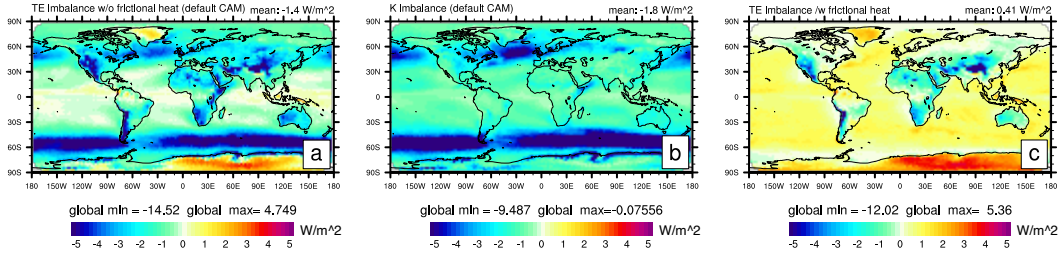


Figure 11. (a) Total energy imbalance between CAM and CLUBB energies, (b) kinetic energy imbalance in CLUBB and (c) energy imbalance excluding kinetic energy imbalance (as if CLUBB included frictional heating that locally converts dissipated momentum into heating - currently done columnwise with fixer). Figure (c) is plot (a) minus (b), i.e. the energy imbalance as a consequence of differences in thermodynamic potential in CAM (enthalpy) and CLUBB (liquid water potential temperature). All units are in W/m^2.

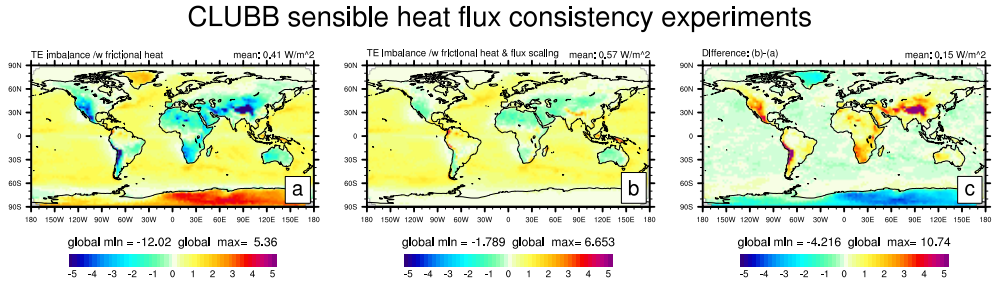


Figure 12. One-year average of total energy imbalance for CLUBB in W/m^2 (a) including frictional heating and default CAM sensible heat flux passed to CLUBB (see equation (164)), (b) including frictional heating and scaling the sensible heat flux for consistency (see equation (163)) and (c) the difference between the first two panels.

which is missing the Exner function in the denominator. By consistently scaling the sensible heat flux, the imbalance between CAM and CLUBB energy is reduced, as shown in Figure 12. It shows a one-year average of the total energy imbalance for CLUBB when including frictional heating and using the default sensible heat flux (164) formulation. Note that the imbalances are large over high orography, where $1/\Pi_s$ is large. By using the more consistent formulation of the sensible heat flux (163), the imbalance over orography is significantly reduced. Since the spurious biases over high orography were negative the imbalance in terms of the global average is increased from $\sim$ 0.41 W/m^2 to $\sim$ 0.57 W/m^2. Discussing the impact of this change to the climate is beyond the scope of this paper.

4.4 Mass clipping errors

Parameterizations (e.g. vertical diffusion) or the dynamical core may produce small negative mixing ratios of water vapor and/or other forms of water. Usually these negative value are clipped by simply setting the mixing ratios to zero. This is seen by the model as a spurious source of water and thereby impacts total energy. That said, these errors have been found to be very small, but will show up in an accurate budget analysis. The error appears as a bias (always positive), so it could accumulate over time in long runs. This topic is discussed further in Zhang et al. (2018).

4.5 Inconsistencies between dynamical core and parameterizations

The dynamical core may be based on equations of motion and thermodynamics different than those assumed in the parameterization suite.

4.5.1 Example 1: Latent heat inconsistency

NCAR's spectral-element dynamical core uses the same HPE equations as CAM physics, except that it assumes variable latent heats. This means that the enthalpy terms, and therefore the total energy formula, differ as discussed in detail in Section 2.1.10. This leads to a ~ 0.6 W/m^2 imbalance. This inconsistency can be alleviated by scaling the temperature increments, ΔT, accordingly

$$\frac{\Delta T c_p^{(d)}}{c_p^{(all)}}, \tag{165}$$

where the generalized $c_p^{(all)}$ is given in (33). Hereby the energy increment associated with the enthalpy terms are the same in physics and dynamics. This approach is also taken in FV3 (L. M. Harris et al., 2021).

4.5.2 Example 2: Different vertical coordinates in physics and dynamics

Many non-hydrostatic models are formulated using a height-based vertical coordinate (e.g. Skamarock et al., 2012). Leaving aside the non-hydrostatic effect, the total energy differs between a z-based HPE system and pressure based HPE system due to the different boundary conditions (constant z and p at model top, respectively); see equations (11) and (13). Heating increments, ΔQ_k, added under constant pressure and constant volume assumptions lead to different temperature increments ($\Delta Q/c_p^{(all)}$ and $\Delta Q/c_v^{(all)}$, respectively). A scaling of the temperature increments by $c_p^{(all)}/c_v^{(all)}$ can alleviate this inconsistency. There are other inconsistencies in this context (e.g. related to top boundary condition) that are beyond the scope of this paper to discuss.

5 Summary and future directions

5.1 Summary

A summary of total energy errors discussed in this paper is given below:

- Dynamical core total energy errors (sections 2.1.10 and 4.1) can be large ($\sim 0.1-1$ W/m^2) at large-scale model resolutions ($\sim 1^\circ$; Lauritzen & Williamson, 2019). Errors are expected to decrease as horizontal/vertical resolution increases. At least for lower resolution applications, further research in energy conserving numerical methods is highly recommended (sections 5.2.1 and 5.2.2).
- Temporal physics–dynamics coupling errors (where tendencies are added throughout the dynamical core time loop) can be large (~ 0.4 W/m^2, Lauritzen & Williamson, 2019, see also section 4.2), but can be reduced by clever physics–dynamics coupling methods (e.g. mass-weighting of momentum and enthalpy tendencies; see section 4.5). Errors, in general, are expected to decrease with shorter physics time-steps and adequate weighting of the tendencies for energy conservation.
- Spatial physics–dynamics coupling errors due to mapping tendencies between grids (if applicable) are small (~ -0.1 W/m^2, Herrington et al., 2018). This error will decrease with increased resolution (see section 4.5).
- Physics–dynamical errors due to the fact that the energy associated with the continuous equations of the dynamical core is different than the energy of the physics (e.g. due to different a-principio approximations) can be large when the dynam-

ical core uses variable latent heats and physics does not (~ 0.5 W/m^2, Lauritzen & Williamson, 2019). These errors can be alleviated by clever scaling of tendencies (e.g., equation (165); see section 4.5)).

- Enthalpy tendencies associated with falling precipitation and water entering the atmosphere are large (~ 0.3 W/m^2) when using constant latent heats and even larger (~ 1.1 W/m^2) when using variable latent heats. Locally, the errors can be orders of magnitude larger. This error, in general, is not expected to decrease with increased resolution. In fact, larger water contents at higher resolution may make the issue worse. This error is extensively discussed in Section 3.
- Kinetic and potential energies associated with falling precipitation (and evaporation or other forms of water entering the atmosphere) are small ($\sim< 0.02$ W/m^2). This error is expected to increase with increased resolution for the same reasons as for enthalpy flux. This error is extensively discussed in Section 3.
- Errors associated with not including all forms of water in pressure/mass are small ($\sim< 0.01$ W/m^2). Local errors could increase with increased resolution as the water content locally is larger at higher resolutions. This error is extensively discussed in Section 3.
- Thermodynamic inconsistency errors in parameterizations: These imbalances are, of course, specific to the inconsistency in question. For example, we showed that the inconsistency between CLUBB using moist potential temperature θ_l as its conserved variable and CAM using enthalpy leads to ~ 0.4 W/m^2 imbalance. Similarly, inconsistencies in surface sensible heat flux between the host model and the parameterization (in this CLUBB) can lead to significant errors/imbalances in the total energy budget. This error is extensively discussed in Section 4.3.

For a closed energy budget that fully accounts for falling precipitation (and water entering the atmosphere) it is necessary to specify temperatures and kinetic energies of water entering/leaving the atmosphere (for the surface flux terms). This has also been noted in several other studies assimilating observations and diagnosing energy budgets (Trenberth & Fasullo, 2018; Kato et al., 2021). This is problematic (as discussed here in detail) since falling precipitation is formed away from surface and the temperature and winds can be quite different from those near the surface where, ultimately, falling precipitation leaves the atmosphere. Complicating the matter is that large-scale models use a single temperature for all components of moist air. This is especially an issue for larger hydrometeors (e.g. hail) that tend to have appreciably different temperatures than the air around them and the supercooled droplets that they accrete. This is because riming releases latent heat that can't be diffused away quickly. As these large hydrometeors become included in advanced microphysics schemes, such as *PUMAS* (Gettelman & Morrison, 2015), the error of using the assumption of a single temperature may be worth considering. Also, climate models do not typically parameterize processes associated with falling precipitation (frictional heating of falling precipitation, drag exerted by falling hydrometeors, etc.).

The total energy budget errors associated with these processes can be large and we recommend that this issue be addressed as a high priority in climate and weather models. We have shown that including a surface enthalpy flux (though difficult to accurately compute) can reduce the variance of energy errors/imbalances significantly (see Figure 9) but, maybe, not entirely for the right reasons. A framework for representing frictional heating of falling precipitation can be achieved by using the barycentric velocity framework as outlined in detail in Appendix F. It still makes use of the single temperature assumption that most ESMs follow, so using the barycentric velocity framework could be a first "incremental" step towards a more realistic and better closure of the total energy budget.

Given these challenges in rigorously closing energy budgets, it is unlikely that the community will be able to entirely remove energy fixers in operational models anytime

soon. We recommend that, if an energy budget cannot be closed locally, global energy fixers should be used.

Global fixers do not seem to affect the climate negatively, and they restore the global energy budget so that coupled climate simulations do not have spurious drifts. Trying to restore energy conservation with a local fixer can lead to incorrect solutions (e.g. Williamson et al., 2009). As a rule of thumb, a small global heating increment (i.e. a global energy fixer) seems to be less detrimental. That said, some local errors in current models (e.g. CAM) are entirely fixable. These include adding frictional heating to compensate for down-gradient diffusion of kinetic energy. Note that this would not actually be a fixer, but would emulate a physical process.

5.2 Future directions

In this section we conclude the paper by suggesting possible future directions and associated major problems that have begun to receive some attention in the literature. This list is not exhaustive. It is based on countless discussion between the BIRS workshop participants that inherently could be biased due to the lack of participation from experts in all aspects of Earth System Modeling.

Global models are moving away from the HPE set to equations of motion that are less approximated, and even incorporate some irreversible processes into the dynamical core. As this transition is made, it becomes increasingly tedious and error-prone to ensure that fundamental physical principles are satisfied. These physical principles include energy conservation and entropy conservation by reversible processes, and entropy generation by irreversible processes. Verifying these physical principles is especially challenging when approximations are made at the level of the equations of motion. In this section we give a brief discussion of three related developments that aim to make this process simpler, and provide a unifying approach:

- **Geometric mechanics formulations** (i.e. Lagrangian or Hamiltonian) to ensure physically-consistent equations of motion.
- **Structure-preserving mimetic discretizations** based on these geometric structures. These discretizations are inherently energy-conserving.
- **Thermodynamic potentials** to ensure consistent thermodynamics.

This section also includes short notes on averaging operators and the available and non-available components of energy.

5.2.1 Geometric mechanics formulations for GFD

Rather than working at the level of the equations of motion, a more powerful approach is to instead work with a geometric mechanics formulation, such as a Lagrangian (Holm et al., 2002, 1998) or Hamiltonian (Salmon, 1998; Dubos & Tort, 2014; Bannon, 2003) approach. This is commonly done in other areas of physics (it is at the heart of modern physics), and it can also be done for GFD.

Briefly, a variational approach to fluid dynamics states that the equations of motion arise from the minimization (through the calculus of variations) of the *action functional* $\mathcal{S} = \int_{t^1}^{t^2} \mathcal{L}[\mathbf{x}]$

$$\delta\mathcal{S} = \delta \int_{t^1}^{t^2} \mathcal{L}[\mathbf{x}] = 0 \tag{166}$$

where $\mathbf{x}$ is the set of predicted variables (for example: density, fluid velocity, specific entropy in the case of the dry Euler equations) and $\mathcal{L}[\mathbf{x}]$ is the *Lagrangian*.

By using the appropriate Lagrangian, one can recover all the consistent equation sets used in GFD for both single and multicomponent fluids. More specifically, it is possible to incorporate (Tort & Dubos, 2014; Eldred & Gay-Balmaz, 2021):

- *Geometric* (shallow/traditional atmosphere, spherical geoid, etc.) and *dynamical* (quasi-hydrostatic, etc.) approximations.
- Arbitrary *equations of state*, since the *thermodynamic potential* (see below) determines the internal energy $U^{(all)}$ (in fact, internal energy is one of the possible choices for thermodynamic potential).
- Sound-proofing *constraints* that lead to *semi-compressible* equations: anelastic, Boussinesq, pseudo-incompressible, semi-hydrostatic, etc.
- *Free surface boundary surfaces*, such as a constant pressure upper boundary.

In all cases, by construction the resulting equations of motion conserve energy. The variational approach can also be used to derive the associated Hamiltonian formulation (Salmon, 1998; Dubos & Tort, 2014; Bannon, 2003), which is useful for (amongst other things) the development of numerical schemes that have no spurious energy conversions and therefore discretely conserve energy (see below).

In addition to providing a unified framework for reversible (entropy-conserving) dynamics, geometric mechanics formulations (both variational (Gay-Balmaz, 2019) and bracket-based (Eldred & Gay-Balmaz, 2020)) have recently been extended to incorporate irreversible (entropy-generating) processes such as viscous dissipation. This includes pioneering work of Gassmann and Herzog (2015) and Gassmann (2018) on the treatment of physics parameterizations by analogy with molecular-scale irreversible processes, so that they have the same functional form. This has provided, for the first time, a full model (dynamical core + physics) with a consistent treatment of the first and second laws of thermodynamics that does not require an energy fixer to correct mismatches between the dynamical core and the physics, and has a physically plausible treatment of entropy generation by physics parameterizations. Of course, treating subgrid processes by analogy with irreversible processes can have limitations and lead to an unphysical representation (despite being energetically and thermodynamically consistent), especially since many subgrid processes are actually unresolved reversible dynamics. Some key examples of this for the atmosphere are gravity wave drag and convection. Nevertheless, the work in Gassmann and Herzog (2015) and Gassmann (2018) represents an important first step towards energetically and thermodynamically consistent physics prameterizations.

5.2.2 Structure-preserving discretizations

Many of the important properties that the equations used in GFD possess, such as energy conservation, arise fundamentally from characteristics of the geometric structures that underlie them. For example, energy-conservation can be associated with temporal symmetry of the Lagrangian in the variational formulation, or anti-symmetry of the Poisson brackets in the Hamiltonian formulation. A structure-preserving (or mimetic) discretization preserves these properties in the discrete sense, to produce numerical models that have the same properties as the continuous equations. This enables, for example, dynamical cores that inherently conserve energy for the correct physical reasons (e.g. reversible transformations between different energy reservoirs). Structure-preserving discretizations have a long history in atmospheric dynamical cores, going back to the celebrated Arakawa and Lamb scheme (Arakawa & Lamb, 1981). Some recent work in this this area for variational formulations is found in Gawlik and Gay-Balmaz (2021); Bauer and Gay-Balmaz (2019); Brecht et al. (2019) and for Hamiltonian formulations in Dubos et al. (2015), Gassmann (2013), Eldred and Randall (2017), Taylor et al. (2020), and Wimmer et al. (2020). It has even been extended to incorporate irreversible processes in Gassmann and Herzog (2015) and Gassmann (2018). Both research models such as ICON-IAP and

"production/operational" models such as CAM-SE (used in this study), E3SM and DYNAMICO make use of this approach. In addition, both ICON-IAP and the MPI-M/DWD ICON model (Zängl et al., 2015) makes use of the barycentric velocity framework although a simplified lower boundary condition is used in ICON ($w = 0$; as a consequence ICON erroneously conserves the total mass rather than dry mass) and the diffusive fluxes of the non-precipitating species are neglected.

5.2.3 Thermodynamic potentials

Given the issues concerning thermodynamic consistency, we would like to point the reader to advances in this area as well. In the atmospheric dynamics literature it is common to work directly with equations of state like (29) and potential temperature definitions. However, this approach becomes increasingly complicated and error-prone as complexity increases, and it is easy to lose thermodynamic consistency if approximations are made to these expressions individually. This leads to violations of the first and/or second laws of thermodynamics, which is an undesirable scenario.

The correct approach is to start more fundamentally, with a *thermodynamic potential* (such as *internal energy*, *enthalpy* or *Gibbs function*) regarded as a function of *canonical state variables* (e.g. pressure, temperature and composition for the Gibbs function), and use this thermodynamic potential to derive the equations of state and other required thermodynamic equations or quantities (i.e. the thermodynamics of the model). Then, when making approximations, or introducing complications such as condensates, or using a temperature-dependent $c_p^{(all)}$, this can be done at the level of the thermodynamic potential. The thermodynamics are then re-derived from the new thermodynamic potential. This avoids inconsistencies, and provides a unified framework for the thermodynamics of a given system.

This thermodynamic potential approach is well-known in the ocean modelling community, since realistic ocean models require an accurate approximation of the complex thermodynamics of seawater. This is achieved consistently through the definition of an approximate, but accurate, Gibbs potential (Feistel, 2008; Feistel et al., 2010; Thuburn, 2017; Hermann et al., 2009). This approach is less common in the atmospheric modeling community, but some recent work has investigated thermodynamic potentials for moist air. This includes an approach based on the Gibbs function for moist air with liquid water (Thuburn, 2017, 2018), a comprehensive study of multiple thermodynamic potentials for moist air with liquid and frozen water (including a consistent treatment of the commonly used constant κ approximation; Eldred et al., 2022), and an approach based on the internal energy for moist air with liquid and frozen water (that also includes non-equilibrium thermodynamic processes; Bowen & Thuburn, 2022a, 2022b). Bowen and Thuburn (2022a) found that using Gibbs function is viable for moist air containing vapor and liquid water but problematic to extend to include the ice phase (due to ambiguity at the triple-point). They show that using internal energy potential resolves this issue.

Oceanographers have a new thermodynamic standard (TEOS-10; `http://www.teos10.org/`) formulating all possible thermodynamic properties of seawater in terms of a Gibbs function and its various partial derivatives. As a result, all thermodynamic quantities are consistent with each other, which was not the case in earlier formulations. The Gibbs function contains four arbitrary constants that have been specified in a way that fixes the value of the freshwater component of specific enthalpy $h^{(W)}$. Atmospheric modelers should be aware of how oceanographers have set the value of enthalpy for seawater, in order to define the value of the specific enthalpy for water vapor to ensure that $h^{(wv)} - h^{(W)}$ is equal to an acceptable form of the latent heat of evaporation. How to achieve consistent definitions of liquid water and moist air enthalpies has been discussed in detail in Feistel et al. (2008).

5.2.4 Dissipation of mechanical energy

An important issue highlighted in this paper is the need to understand the pathways to dissipation of mechanical energy. This is required to determine when the energy loss should contribute to the Joule heating associated with viscous dissipation in the temperature equation and when it should enhance the vertical turbulent transfer of heat and other tracers (resulting in change in the gravitational potential energy of the fluid, but with no contribution to viscous dissipation). In the oceanographic community, such an idea has motivated the development of "energetically consistent models", e.g., Eden (2016), whereby explicit descriptions of the sub-grid-scale energy transfers and of their interactions with the resolved energy reservoirs are sought. One possible way to achieve this is to develop separate evolution equations for the turbulent kinetic energy and the available potential energy. This approach is being explored in the atmospheric boundary layer community. Further research in this area is facilitated by the fact that a fully general local theory of available potential energy for multi-component compressible stratified fluids is now available.

5.2.5 Averaging operators

The energetics of a model depends *a priori* on the exact interpretation of the coarse-grained prognostic variables. In particular, it depends on the averaging operator used to construct the coarse-grained equations satisfied by the model. In practice, however, the precise form of this averaging operator is often unspecified. In oceanography, there is increased interest in coarse-grained equations obtained from some form of Lagrangian averaging, such as the thickness-weighed averaged equations, e.g. Young (2012). In the atmosphere, however, Favre-averaging may be a more natural way to account for the large variations in density. To a large extent, the present paper has approached the energetics of atmospheric models by treating the model variables in a similar way as their non-averaged form. A more consistent and physically-based approach would require to have a clear sense of the exact way in which the coarse-graining has been performed.

5.2.6 Available and non-available components of energy

For simplicity, the present paper considered only the total potential energy of the fluid. In particular, it did not attempt to develop separate budgets for its available and non-available components as originally defined by Lorenz (1955). Physically, however, the available and non-available forms of potential energy have very distinct signatures, so that it might be desirable in the future to extend the present analysis to the various forms of potential energy, by using the newly developed exact available potential energy frameworks of Novak and Tailleux (2018) and Tailleux (2018), although significant conceptual barriers remain to incorporating moisture in such frameworks, e.g., B. L. Harris and Tailleux (2018).

Appendix A A note on averaging: unresolved and unrepresented total energy

A rigorous definition of the resolved and unresolved scales requires a complete equation set valid for all energetic scales of motion and processes present in the atmosphere. Unfortunately, we do not have such a set of equations, derived without approximation from first principles. Nevertheless, we can consider the most comprehensive system currently available. We then need to define an appropriate averaging operator to separate the resolved and unresolved scales. Unfortunately, there is not a unique choice for such an operator in general. The most developed example is large eddy simulation (LES), which assumes an averaging scale in the inertial range of fully developed turbulence (Lesieur & Metais, 1996). Reynolds-averaged Navier—Stokes (RANS) models are less accurate,

but are computationally cheap and easy to implement. They use Reynolds averaging to decompose a flow into its time-averaged and fluctuating (turbulent) components and provide an approximate solution for the time-averaged flow.

Fluid turbulence is a particularly straightforward case since we have a complete equation set (the Navier–Stokes equations) valid over all energetic scales of motion. We also have a numerical method, direct numerical simulation (DNS) to solve the unapproximated Navier–Stokes equations. DNS provides a benchmark for evaluating and validating various approximate models. In turbulence modeling an appropriate filter is used to partition the Navier–Stokes equations into resolved unresolved components. Because the Navier–Stokes equations are nonlinear, this filtering results in a residual Reynolds stress-like term that represents the effect of the unresolved fluctuations on the resolved large scales. Modeling these Reynolds stresses is known as the "closure" problem, since an additional model is required to close the large scale equations. The best-known closure model is Smagorinsky eddy viscosity, which was proposed in the 1960s for atmospheric flows (Smagorinsky, 1963). LES models are now well-developed, and include a range of sophisticated filtering methods (e.g. dynamic LES, which uses the Germano identity (Germano, 1992) to measure energy flux across the cut-off scale). However, LES requires that the cut-off scale is in the scale-free inertial range, and that the equations of motion are the same at all scales of motion. LES may not be directly applicable to atmospheric dynamics because of the interaction between fluid dynamics and thermodynamics.

We have a pretty good idea of how averaging and subgrid scale modeling (i.e. parameterization) works for fluid turbulence. However, when dealing with anything more complicated than the Navier–Stokes equations the correct approach is not obvious. Any attempt to describe the average state of the atmosphere from an approximate system of equations is bound to be problematic. The interaction between resolved and unresolved scales in multiphysics simulations is a vast, largely undeveloped challenge for ESMs.

In particular, developing appropriate averaging operators and subgrid scale models for thermodynamics is problematic. For simplicity, we assume that the average internal energy is the energy of the averaged variables substituted into the continuous formula

$$\overline{U^{(all)}} := \overline{c_v^{(all)} T} \approx \overline{c}_v^{(all)} \overline{T}, \tag{A1}$$

where the generalized specific heat $c_v^{(all)}$ is the continuous formula replaced with dynamical core prognostic state values (i.e. from the resolved fluid dynamics)

$$\overline{c}_v^{(all)} = \frac{\sum_{\ell \in \mathcal{L}_{all}} c_v^{(\ell)} \overline{m}^{(\ell)}}{\sum_{\ell \in \mathcal{L}_{all}} \overline{m}^{(\ell)}}. \tag{A2}$$

where $\overline{T}$, and $\overline{m}^{(\ell)}$ are the prognostic variables (temperature and dry mixing ratio, respectively) defined as cell-averaged values on the discretization grid. The set $\mathcal{L}_{all}$ is the components of moist air (more details in Section 2.1.2). Things are now complicated because we are combining different kinds of filtering and approximation: spatial averaging, temporal averaging, linearization, and parameterization of unresolved physics. In addition, as with Reynolds stresses, nonlinear terms in the resolved equations do not commute with averaging (i.e. small scale processes drive large scale motions). These subgrid scale terms are modeled as functions of the resolved variables, which results in parameterizations that are distinct from parameterizations of fully unresolved processes (e.g. cloud microphysics).

As grid resolution increases, we eventually cross a threshold where the basic nature of the filtering, and the resolved equations, changes. This happens because processes that had been completely unresolved become partially resolved. They are therefore associated with equations that must be solved, in addition to the unresolved scales that must still be parameterized.

One well-known example of such a resolution threshold is the "gray zone" problem of deep convection problems, where convection that was completely unresolved (i.e. fully parameterized) at larger scales suddenly becomes partially resolved at scales of $O(10)$ km. In addition, the statistical assumptions of some parameterizations are no longer satisfied. This means that in practice it is not possible to simply increase the resolution of ESMs as more powerful computers become available. A significant jump in resolution requires time-consuming tuning of the parameterizations and, in some cases, adding new equations to the dynamical core. Had the parameterizations been formulated in terms of an averaging operator and a well-defined closure then the transition would be seamless.

One way of dealing with changing resolutions is to use so-called "scale aware" parameterizations (Arakawa & Wu, 2013; Freitas et al., 2018; Arakawa et al., 2011; Park, 2014). In this approach the parameterizations modify themselves based on information provided by the dynamical core about the current local scale.

Given the challenges of developing consistent averaged equations, henceforth we focus on the unaveraged total energy equations, with the exception of Section 2.1.12 on kinetic energy dissipation and heating which inevitably involves turbulent (sub-grid scale) kinetic energy.

Appendix B Derivation of the single-fluid energy equation and energy conservation equation in the primitive equation set with a generalized vertical coordinate in the presence of mass sources

B1 Local energy equations

The fluid energy equation is obtained from the two separate equations for the enthalpy and the kinetic energy of the fluid:

$$\frac{dh}{dt} - \alpha \frac{dp}{dt} = Q \tag{B1}$$

$$\frac{d}{dt}\left(\frac{1}{2}\vec{v}^2\right) = (\alpha \mathbf{F} - \alpha \nabla p - \nabla \Phi) \cdot \mathbf{v} \tag{B2}$$

Equation (B1) expresses the first principle of thermodynamics, and (B2) is the scalar product of the Euler equations with the 3-dimensional fluid velocity $\mathbf{v} := (\vec{v}, w)$. (Similarly, we use the three-dimensional gradient operator $\nabla := (\vec{\nabla}, \partial/\partial z)$). Both equations are expressed per unit mass of fluid with specific volume $\alpha =: 1/\rho$, pressure p, and specific enthalpy h. Note that for a fluid composed of several, non interacting constituents, $1/\alpha = \sum_\ell 1/\alpha^{(\ell)}$; likewise, p is the sum of the partial pressures. Expressions for h are given in Appendix C1. The r.h.s. term Q is the heat acquired by the fluid, per unit mass and per unit time, from external sources (e.g. radiation). In (B2), Φ is the geopotential (a combination of the Earth's gravitational potential and the centrifugal potential, calculated on geoid surfaces), and $\mathbf{F}$ represents other forces acting on the fluid (which might include viscous and other dissipative forces); in a multi-constituent fluid, $\mathbf{F} = \sum_\ell \mathbf{F}^{(\ell)}$ where $\mathbf{F}^{(\ell)}$ is the force acting on constituent ℓ. (Note that here we do not discuss cases when different constituents might have different three-dimensional velocities, i.e. we adhere to a strict single-fluid framework; this assumption is partially relaxed in Appendix F.)

Summation of (B1)-(B2) results in the evolution equation:

$$\frac{d\epsilon}{dt} = \alpha \frac{\partial p}{\partial t} + \frac{\partial \Phi}{\partial t} + \alpha \mathbf{F} \cdot \mathbf{v} + Q\,, \tag{B3}$$

where

$$\epsilon := \frac{1}{2}v^2 + \Phi + h\,, \tag{B4}$$

is the Bernoulli function of the fluid (section 4.8 Gill, 1982), and we have used the relation

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla \,. \tag{B5}$$

Note (B3) is not valid for every choice of coordinates: for example, the pressure force has to be a spatial gradient of the pressure. Hence (B3) is only valid for time-fixed (time-independent) coordinates e.g. terrain-following z-based coordinates. Hence the partial time derivatives in (B3) should be understood to be taken at fixed z. In a coordinate transformation where z is replaced by a generalized vertical coordinate ('GVC'),

$$\{t, x, y, z\} \rightarrow \{t, x, y, \eta\} \quad \left(\frac{\partial \eta}{\partial z} \neq 0 \;\; \forall \{t, x, y, z\}\right), \tag{B6}$$

where η is a C^1 function of $\{t, x, y, z\}$ strictly monotonic in z, these terms become (cf. equation (3.6) in Kasahara, 1974)

$$\alpha \frac{\partial p}{\partial t} + \frac{\partial \Phi}{\partial t} = \alpha \left.\frac{\partial p}{\partial t}\right|_\eta + \left.\frac{\partial \Phi}{\partial t}\right|_\eta - \left(\frac{\partial z}{\partial \eta}\right)^{-1} \left.\frac{\partial z}{\partial t}\right|_\eta \left(\alpha \frac{\partial p}{\partial \eta} + \frac{\partial \Phi}{\partial \eta}\right) . \tag{B7}$$

In this section, as in the rest of the paper, we make the assumption of hydrostatic equilibrium, i.e. that the vertical ($z-$) acceleration of the fluid is always negligible in comparison to other inertial, pressure, viscous, turbulent or external forces. This assumption breaks down at the scales where buoyancy acceleration becomes important, e.g. in topographic gravity waves and in local circulations driven by atmospheric convection. In global models, such motions are heavily parameterized and their associated energy is mostly unaccounted for (see also Appendix A and Appendix E). In a GVC system (B6), the hydrostatic condition reads (cf. equation (3.20) of Kasahara, 1974)

$$\alpha \frac{\partial p}{\partial \eta} + \frac{\partial \Phi}{\partial \eta} = 0 \,, \tag{B8}$$

so that the last term in parenthesis in (B7) vanishes identically. In other words, under the hydrostatic assumption, the fluid energy evolution equation (B3) is valid in form for any GVC choice (B6). Note that, in GVC, the total derivative operator (B5) is expressed in the natural form:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla}_\eta + \dot{\eta} \partial_\eta \,, \tag{B9}$$

where the arrow symbols indicate 2-D vectors in the horizontal plane only, ∇_η is formed of partial derivatives taken at constant η, and $\dot{\eta}$ is the total time derivative of η, related to the vertical velocity w as (cf. equation (3.12) of Kasahara, 1974):

$$\dot{\eta} = \frac{\partial \eta}{\partial z} \left[w - \left(\left.\frac{\partial}{\partial t}\right|_\eta + \vec{v} \cdot \vec{\nabla}_\eta \right) z \right] . \tag{B10}$$

Frequently, hydrostatic pressure is used as a GVC, e.g. in (13). In such cases, (B8) is the *definition* of the pressure *coordinate*, which needs, in general, to be distinguished from the pressure *state variable*. Only if condition (B8) is *assumed* to be valid for the pressure variable, as we do here, may the two be confused and a single symbol, p, used for both.

Finally, we combine (B3) with the equation of mass conservation to obtain a conservative form of the energy equation. In local Cartesian coordinates, mass conservation is expressed as

$$\frac{1}{\alpha} \frac{d\alpha}{dt} = \nabla \cdot \mathbf{v} - \alpha \dot{\mu} \,, \tag{B11}$$

where $\dot{\mu}$ is the apparent source/sink of fluid mass per unit volume, usually associated in the atmosphere with removal of water by precipitation and deposition processes, or

with addition of water by evaporation and sublimation from the surface. In CAM, where condensates are approximated to be mass-less, $\dot{\mu}$ in fact includes all processes with phase changes to or from water vapor.

In GVC, the divergence term on the r.h.s. of (B11) reads (cf. equations (3.16) and (3.17) of Kasahara, 1974)

$$\nabla \cdot \mathbf{v} = \vec{\nabla}_\eta \cdot \vec{v} + \partial_\eta \dot{\eta} + (\partial_\eta z)^{-1} \frac{d}{dt} (\partial_\eta z) \,, \tag{B12}$$

so that (B11) becomes

$$\frac{d}{dt} \left(\partial_\eta z \, \alpha^{-1}\right) + \partial_\eta z \, \alpha^{-1} \left(\vec{\nabla} \cdot \vec{v} + \partial_\eta \dot{\eta}\right) = \partial_\eta z \, \dot{\mu} \,. \tag{B13}$$

Note that (B13) is valid (physically or mathematically) irrespective of the hydrostatic assumption. If we now make that assumption, then can multiply (B3) by $\partial_\eta z \, \alpha^{-1}$ and sum it with (B13) multiplied by ϵ to obtain the total energy equation in formally conservative form:

$$\frac{\partial}{\partial t} \left(\partial_\eta z \, \alpha^{-1} \epsilon\right) + \nabla_\eta \cdot \left(\partial_\eta z \, \alpha^{-1} \mathbf{v}_\eta \epsilon\right) = \partial_\eta z \, \alpha^{-1} \left(\alpha \partial_t p + \partial_t \Phi\right) + \partial_\eta z \, \alpha^{-1} Q + \partial_\eta z \, \mathbf{F} \cdot \mathbf{v} + \partial_\eta z \, \dot{\mu} \epsilon \,, \tag{B14}$$

where we have introduced the shorthands $\nabla_\eta := (\vec{\nabla}, \partial_\eta)$ and $\mathbf{v}_\eta := (\vec{v}, \dot{\eta})$. Note that for sake of readeability we have dropped the subscripts η from the partial time derivative operators. In the following, we will also drop them from the shorthands of three-component vectors and operators. (B14) is formally valid for a general geopotential field $\Phi = \Phi(t, x, y, z)$. However, consistency with the primitive equation set used here, which neglects spherical geometry terms in the momentum equations under the assumption of a thin-shell atmosphere stratified approximately over equilibrium geoid surfaces, in fact requires the approximation that Φ be linearly proportional to z, $\Phi = gz$. Together with the hydrostatic approximation, this implies

$$\alpha^{-1} = -\frac{1}{g} \frac{\partial_\eta p}{\partial_\eta z} \tag{B15}$$

so that (B14) becomes

$$\frac{\partial}{\partial t} \left(\partial_\eta p \, \epsilon\right) + \nabla \cdot \left(\partial_\eta p \, \mathbf{v} \epsilon\right) = \partial_\eta p \, \left(\alpha \partial_t p + \partial_t \Phi\right) + \partial_\eta p \, Q + \partial_\eta p \, \alpha \mathbf{F} \cdot \mathbf{v} + \partial_\eta p \, \dot{q} \epsilon \,, \tag{B16}$$

where $\alpha \dot{\mu} =: \dot{q}$, and q may be interpreted to be the specific humidity for either total water (in general) or for water vapor only (CAM assumptions) since in all cases dry air mass is conserved. Next, we explore various form of (B16) integrated over a global or limited horizontal domain and over the entire vertical domain.

B2 Energy conservation equations

We write out (B16) as

$$\partial_t \left(u_m \partial_\eta p\right) + \vec{\nabla} \cdot \left(\vec{v} \epsilon \partial_\eta p\right) + \partial_\eta \left(\dot{\eta} \epsilon \partial_\eta p\right) = \partial_\eta \left(p \partial_t \Phi\right) + \left(\alpha F \cdot v + Q + \dot{q} \epsilon\right) \partial_\eta p \tag{B17}$$

where $u_m := \epsilon - \alpha p$ is the total (kinetic + thermal + latent + potential) energy per unit mass of air.

Integration of (B17) between the top of the model domain, $\eta = \eta_t$, and the surface, $\eta = \eta_s$, and over a horizontal domain $\mathcal{A}$, yields for the total energy change:

$$\begin{aligned} \partial_t \int_{\mathcal{A}} dS \int_t^s \partial_\eta p \, d\eta \, u_m \quad = \quad & - \int_{\mathcal{A}} dS \, p_t \, \partial_t \left[\Phi(\eta_t)\right] - \oint_{\delta \mathcal{A}} \int_t^s \partial_\eta p \, d\eta \, \epsilon \, \vec{v} \times d\vec{s} \\ & + \int_{\mathcal{A}} dS \, \dot{h}_s + \int_{\mathcal{A}} dS \int_t^s \partial_\eta p \, d\eta \, \left(\alpha \mathbf{F} \cdot \mathbf{v} + Q + \dot{q} \epsilon\right) \,, \end{aligned} \tag{B18}$$

where dS is the area element in $\mathcal{A}$, and in the second term on the r.h.s. $d\vec{s} = (dx, dy)$ is the line element of the boundary line integral over $\delta\mathcal{A}$ (taken anti-clockwise) and $\vec{v}\times d\vec{s} = v_x dy - v_y dx$. The first term on the right-hand side is the pressure work on the top boundary; the second is the advection of total enthalpy out of the domain; the third represents turbulent surface fluxes of enthalpy, $\dot{h}_s := -c_p\overline{\omega' T'}|_s - L\overline{\omega' q'}|_s$; and the fourth and last term represents the contributions from internal sources and sinks of mass, energy, and momentum, e.g. evaporation/condensation processes, radiation, and wave breaking. In obtaining (B18), we have assumed $\dot{\eta}_t = 0$ and $\partial_t \Phi_s = 0$, viz no flow out of the top of the vertical domain, and time-invariant surface geopotential. In the case of a fixed lid, $\partial_t p_t = 0$, Equation (B18) can be also be written for the sum of kinetic energy and enthalpy $\epsilon_m := \epsilon - \Phi$, by using $p_t\, \partial_t \left[\Phi(\eta_t)\right] = \partial_t \left[p_t\, \Phi(\eta_t)\right] = \partial_t \int_s^t \partial_\eta \left(p\Phi\right) + \Phi_s \partial_t p_s$ and using the hydrostatic relation for $\partial_\eta \Phi$; this yields:

$$\begin{aligned} \partial_t \int_{\mathcal{A}} dS \int_t^s \partial_\eta p\, d\eta\, \epsilon_m \;\; &= \;\; -\int_{\mathcal{A}} dS\, \Phi_s\, \partial_t p_s \; - \oint_{\delta\mathcal{A}} \int_t^s \partial_\eta p\, d\eta\, \epsilon\, \vec{v} \times d\vec{s} \\ &\quad + \int_{\mathcal{A}} dS\, \dot{h}_s \; + \int_{\mathcal{A}} dS \int_t^s \partial_\eta p\, d\eta\; \left(\alpha \mathbf{F}\cdot\mathbf{v} + Q + \dot{q}\epsilon\right) \; . \quad \text{(B19)} \end{aligned}$$

Equation (B19) is the form used in CAM to compute and check the atmospheric energy (rectius: enthalpy) budget. It is useful to gather the terms involving the geopotential Φ, by substituting $\partial_t p_s = \int_t^s \partial_t \partial_\eta p d\eta$ and using the continuity equation, and obtain finally:

$$\begin{aligned} \partial_t \int_{\mathcal{A}} dS \int_t^s \partial_\eta p\, d\eta\, \epsilon_m \;\; &= \;\; -\oint_{\delta\mathcal{A}} \int_t^s \partial_\eta p\, d\eta\, \epsilon_z \vec{v} \times d\vec{s} \; - \nabla\Phi_s \cdot \int_{\mathcal{A}} dS \int_t^s \vec{v}\, \partial_\eta p d\eta \\ &\quad + \int_{\mathcal{A}} dS\, \dot{h}_s \; + \int_{\mathcal{A}} dS \int_t^s \partial_\eta p\, d\eta\; \left(\alpha \mathbf{F}\cdot\mathbf{v} + Q + \dot{q}\epsilon_z\right) \; , \quad \text{(B20)} \end{aligned}$$

where $\epsilon_z := \epsilon_m + (\Phi - \Phi_s)$ only involves the geopotential height with respect to the ground. The first two terms on the right-hand side depend on the horizontal winds only (including up- and down-slope mass transfers), and in numerical models, where $\mathcal{A}$ may be taken to represent a single "column" of grid-cells, these two tendency terms are calculated in the dynamical core, i.e. the numerical solver for the adiabatic fluid equations without sources of mass, momentum or energy. The third and fourth terms in (B20) are associated with the physical parameterizations. Considering only these two terms on the r.h.s., we can now make the connection with Equations (110), (121), and (125).

First, we note that, except for the term in Φ_s, the l.h.s. of all those equations express the l.h.s. of (B20) with different forms of the thermodynamic enthalpy $h = \epsilon_m - K$ of the fluid and with different decompositions of the mass term $\rho^{(d)}\left(1 + \overline{m}^{(H_2O)}\right) = \partial_\eta p\, d\eta$. The Φ_s term is expressed as part of $\partial_\eta p\, d\eta \dot{q}\epsilon_z$ in (B20) by making use of (B11). The term proportional to $\mathbf{F}\cdot\mathbf{v}$ is not part of (110), (121), and (125); it is discussed in Section 4.1 and in Appendix E. So we only have to discuss, on the r.h.s. of (B20), the surface integral of the sensible and latent heat fluxes $\dot{h}_s$, and the mass integrals (over $\partial_\eta p d\eta = g/\alpha dz$) of Q and of $\dot{q}\epsilon$. Note that the last two are identical with the volume integrals of Q/α and $\dot{\mu}\epsilon$. Part of the diabatic heating per unit volume, Q/α, is associated with the convergence of radiation fluxes, $-\nabla\cdot\mathbf{F}^{(rad)}$, which in numerical models is approximated as $-\partial F_z^{(rad)}/\partial z$. The vertical integral over the atmospheric column thus results in a boundary flux term $F_{net}^{(rad)} = F_z^{(rad)}(z_s) - F_z^{(rad)}(z_t)$, integrated over $\mathcal{A}$. In (110), (121), and (125) this is implicitly summed to the integral, over $\mathcal{A}$, of $\dot{h}_s$, i.e. $F^{(turb)}(z_s)$, in the r.h.s. term $\overline{F}_{net}^{(turb,rad)}$.

While Q might contain other diabatic heating terms, e.g. mesospheric ion drag or the dissipation of vertically propagating gravity waves, it does not include latent heating terms, which are all accounted for in the expressions for the thermodynamic enthalpy of the fluid h. There are therefore no explicit terms associated with evaporation/sublimation

or condensation/deposition processes in (B20). However, to understand, within the present formalism, the role of precipitation or other processes involving exchanges of material with the surface, which change the composition of the fluid (e.g. by only removing liquid water, but not dry air), we have to apply (B20) to individual fluid components under dynamic and thermodynamic equilibrium assumptions.

The resulting energy conservation equation is a sum of equations of the form (B20) with the pressure replaced by the partial pressure and the specific enthalpy replaced by the specific enthalpy of each substance, velocities and temperatures being shared by all components. The volume integral of $\dot{\mu}\epsilon$ then becomes the volume integral of $\sum_\ell \dot{\mu}^{(\ell)}\epsilon^{(\ell)}$ over all species, ℓ, considered to be part of the fluid motion. This source term is associated with all the *local* enthalpy, kinetic energy and geopotential of the relevant fluid component as a fraction of its mass is considered to join or leave the atmospheric fluid.

The fluid equations thus only describe the process of "joining" or "leaving" the fluid, and the associated energy change. Additional processes that occur in the lead-up, or as a consequence, of mass changes have to be described separately in an ad-hoc fashion. Physically, the fractional mass source of each fluid component does not just appear or disappear. In the case of precipitation, it travels vertically and exits the atmospheric domain at its surface boundary, and similarly for water vapor that joins the atmosphere as a result of surface evaporation.

A zeroth-order assumption is that there is no interaction whatsoever between the newly acquired, or newly lost, mass and that persisting in the atmospheric fluid. In that case the mass source term, $\dot{\mu}^{(\ell)}$, may be described as the divergence of a vertical mass flux of the fluid-decoupled part of that components, $-\partial F^{(\ell)}\partial z$. If all the divergence were localized at one altitude z, the vertical integral

$$\int_t^s dz \dot{\mu}^{(\ell)}\epsilon^{(\ell)} = F^{(\ell)}_{net}(z_s)\epsilon^{(\ell)}(z_p)\,, \tag{B21}$$

where z_p is where precipitation is formed. However, this would be an unrealistic assumption, because local thermodynamic equilibration of e.g. rain drops with the surrounding air is generally faster than the drop crossing time as they fall to the ground. Similarly, aerodynamic friction slows the drops down to much less than the free-fall speed (in vacuum).

Thus, a much more plausible physical approximation would allow rain drops to exchange with the atmosphere all of their energy that is in excess of the water enthalpy equilibrated at the temperature of the lowest atmospheric layer. Effectively, this represents an additional, ad-hoc contribution to Q, say $Q^{(\ell)}_p$, such that

$$\int_t^s dz \left(\dot{\mu}^{(\ell)}\epsilon^{(\ell)} + Q^{(\ell)}_p/\alpha\right) = F^{(\ell)}_{net}(z_s)\epsilon^{(\ell)}(z_s)\,. \tag{B22}$$

Because now all precipitation is associated with the same, surface, specific enthalpy, its altitude of provenance no longer matters and (B22) holds without any assumption on the shape of $\dot{\mu}(z)$.

Summing (B22) for all atmospheric fluid species (ℓ) finally yields the r.h.s. of Equations (110), (121), and (125).

Note that assumptions like (B22) are not in any way necessary. If explicit expressions are used for $Q^{(\ell)}_p$, based on more detailed physical modeling of the interactions between precipitation and atmospheric air, as done for example in Appendix F, then also the corresponding terms $\dot{\mu}^{(\ell)}\epsilon^{(\ell)}$ must be kept in the energy equations of the fluid.

Appendix C Enthalpy and energy formulas using different reference states

Here we give explicit expressions for the terms forming the enthalpy function of the air-water fluid used in the thermodynamic equation (B1). In doing so, we adopt simplified expressions of h which are generally sufficient to account for the leading errors in energy conservation of numerical atmosphere models like CAM.

The most general formulation of the thermodynamic equations for a fluid of variable density, pressure, temperature and composition is by means of the Gibbs function, G, with total differential given by $dG = -sdT + \alpha dp + \mu dr$ where we indicate with s the fluid entropy and with μ and r the chemical potential and the mixing ratio of e.g. water vapour (Feistel et al. (2010), Thuburn (2017), (Eldred et al., 2022), Bowen and Thuburn (2022a)). Thermodynamic equilibrium states are characterized by a minimum of G on T, p isosurfaces (Adkins (1968), section 10.2). The differential of G shows conjugate variable pairs $(T, -s)$, (p, α), and (r, μ), and h is formally obtained as the Legendre transform w.r.t. the first pair, $h := g + Ts$. Hence, from general principles, h is a conserved quantity under isobaric, isothermal transformations.

We note that an advantage of using the form (B1) of the thermodynamic equation, compared to e.g. an evolution equation for the entropy or the Gibbs function, is that there is no need to explicitly represent terms arising from changes in chemical potential associated with compositional changes.

This is of particular advantage when neglecting the effects of significant departures from thermodynamic equilibrium in e.g. phase changes of water, a common situation in the atmosphere. Equation (B1) is general and the source term Q only includes external heating sources such as radiation flux-divergence or viscous (and, in the Reynolds-average equation set, turbulent) dissipation. Therefore, the thermodynamic approximations made in this paper are contained within the expression used for h.

Following Thuburn (2017) and Bowen and Thuburn (2022a), we use a linearized form of the enthalpy function, based on the two assumptions:

1. specific heats are constants independent of temperature
2. the specific volume of condensates (liquid water and ice) is zero.

These approximations are generally very good in the range of temperatures and pressures typical of atmospheric conditions, and they allow us to avoid significant formal complications which could distract from the main scope of this paper. Please see Thuburn (2017); Bowen and Thuburn (2022a) and Bowen and Thuburn (2022b) for alternative, general formulations that incorporate non-zero condensate volume but keep temperature-independent specific heats. The resulting expressions follow below.

C1 Partial specific enthalpies for different reference states

This Appendix section provides formulas for enthalpy and energy using different reference enthalpy states (see Section 2.1.3). Specific enthalpies are given by

$$\begin{aligned} h^{(wv)}(T) &= h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00}), \\ h^{(liq)}(T) &= h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00}) - L_v(T), \\ h^{(ice)}(T) &= h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00}) - L_s(T), \end{aligned}$$

(water vapor reference state)

$$h^{(wv)}(T) = h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}) + L_v(T), \tag{C1}$$

$$h^{(liq)}(T) = h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}), \tag{C2}$$
$$h^{(ice)}(T) = h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00}) - L_f(T), \tag{C3}$$

(liquid reference state)

$$\begin{aligned} h^{(wv)}(T) &= h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}) + L_s(T), \\ h^{(liq)}(T) &= h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}) + L_f(T), \\ h^{(ice)}(T) &= h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00}), \end{aligned}$$

(ice reference state)

for water vapor, liquid and ice reference states. Substituting these expressions into the total enthalpy formula (54) and rearranging terms yields

$$\begin{aligned} \rho^{(all)}U^{(all)} + p &= \rho^{(d)}c_p^{(d)}T + \rho^{(d)}\left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right) + \rho^{(H_2O)}\left[h_{00}^{(wv)} + c_p^{(wv)}(T - T_{00})\right] \\ &\quad -\rho^{(liq)}L_v(T) - \rho^{(ice)}L_s(T), \end{aligned}$$

(water vapor reference state)

$$\begin{aligned} \rho^{(all)}U^{(all)} + p &= \rho^{(d)}c_p^{(d)}T + \rho^{(d)}\left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right) + \rho^{(H_2O)}\left[h_{00}^{(liq)} + c_p^{(liq)}(T - T_{00})\right] \\ &\quad +\rho^{(wv)}L_v(T) - \rho^{(ice)}L_f(T), \end{aligned}$$

(liquid reference state)

$$\begin{aligned} \rho^{(all)}U^{(all)} + p &= \rho^{(d)}c_p^{(d)}T + \rho^{(d)}\left(h_{00}^{(d)} - c_p^{(d)}T_{00}\right) + \rho^{(H_2O)}\left[h_{00}^{(ice)} + c_p^{(ice)}(T - T_{00})\right] \\ &\quad +\rho^{(wv)}L_s(T) + \rho^{(liq)}L_f(T), \end{aligned}$$

(ice reference state)

for water vapor, liquid and ice as reference states, respectively. As a result the primitive equations, without fluxes through the top and bottom boundaries, conserve the total (globally integrated) energy given in (62) which can be formulated in terms of various reference states

$$\begin{aligned} E_{feom} &= m^{(all)}\left(K + \Phi_s\right) + c_p^{(d)}T \\ &\quad +m^{(H_2O)}\left[h_{00}^{(wv)} + c_p^{(wv)}\left(T - T_{00}\right)\right] - m^{(liq)}L_v(T) - m^{(ice)}L_s(T), \\ &= \left(K + \Phi_s\right) + c_p^{(d)}T \\ &\quad + \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}\left[K + \Phi_s + c_p^{(\ell)}\left(T - T_{00}\right) + h_{00}^{(wv)}\right] - m^{(liq)}L_{v,00} - m^{(ice)}L_{s,00}, \end{aligned}$$

(water vapor reference state, $h_{00}^{(d)} - c_p^{(d)}T_{00} \equiv 0$)

$$\begin{aligned} E_{feom} &= m^{(all)}\left(K + \Phi_s\right) + c_p^{(d)}T \\ &\quad +m^{(H_2O)}\left[h_{00}^{(liq)} + c_p^{(liq)}\left(T - T_{00}\right)\right] + m^{(wv)}L_v(T) - m^{(ice)}L_f(T), \\ &= \left(K + \Phi_s\right) + c_p^{(d)}T \\ &\quad + \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}\left[K + \Phi_s + c_p^{(\ell)}\left(T - T_{00}\right) + h_{00}^{(liq)}\right] + m^{(wv)}L_{v,00} - m^{(ice)}L_{f,00}, \end{aligned}$$

(liquid reference state, $h_{00}^{(d)} - c_p^{(d)}T_{00} \equiv 0$)

$$
\begin{aligned}
E_{feom} &= m^{(all)}\left(K+\Phi_s\right)+c_p^{(d)}T \\
&\quad +m^{(H_2O)}\left[h_{00}^{(ice)}+c_p^{(ice)}\left(T-T_{00}\right)\right]+m^{(wv)}L_s(T)+m^{(liq)}L_f(T), \\
&= \left(K+\Phi_s\right)+c_p^{(d)}T \\
&\quad +\sum_{\ell\in\mathcal{L}_{H_2O}} m^{(\ell)}\left[K+\Phi_s+c_p^{(\ell)}\left(T-T_{00}\right)+h_{00}^{(ice)}\right]+m^{(wv)}L_{s,00}+m^{(liq)}L_{f,00},
\end{aligned}
$$

(ice reference state, $h_{00}^{(d)}-c_p^{(d)}T_{00}\equiv 0$)

C2 Surface enthalpy flux

Net enthalpy fluxes at the surface can be written as

$$
\begin{aligned}
F_{net}^{(h)} &= F_{net}^{(H_2O)}\left(h_{00}^{(wv)}-c_p^{(wv)}T_{00}\right)+c_p^{(wv)}\Big[F_{s\to a}^{(wv)}T_{surf,s}-F_{a\to s}^{(wv)}T_{atm,s}+ \\
&\quad F_{s\to a}^{(liq)}T_{surf,s}-F_{a\to s}^{(liq)}T_{atm,s}+F_{s\to a}^{(ice)}T_{surf,s}-F_{a\to s}^{(ice)}T_{atm,s}\Big] \\
&\quad -F_{s\to a}^{(liq)}L_v(T_{surf,s})+F_{a\to s}^{(liq)}L_v(T_{atm,s})-F_{s\to a}^{(ice)}L_s(T_{surf,s})+F_{a\to s}^{(ice)}L_s(T_{atm,s}) \\
&= \sum_{\ell\in\mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[h_{00}^{(wv)}-c_p^{(\ell)}T_{00}\right]-F_{net}^{(liq)}L_{v,00}-F_{net}^{(ice)}L_{s,00}+ \\
&\quad \sum_{\ell\in\mathcal{L}_{H_2O}} F_{s\to a}^{(\ell)}c_p^{(\ell)}T_{surf,s}-\sum_{\ell\in\mathcal{L}_{H_2O}} F_{a\to s}^{(\ell)}c_p^{(\ell)}T_{atm,s}.
\end{aligned}
$$

(water vapor reference state)

$$
\begin{aligned}
F_{net}^{(h)} &= F_{net}^{(H_2O)}\left(h_{00}^{(liq)}-c_p^{(liq)}T_{00}\right)+c_p^{(liq)}\Big[F_{s\to a}^{(wv)}T_{surf,s}-F_{a\to s}^{(wv)}T_{atm,s}+ \\
&\quad F_{s\to a}^{(liq)}T_{surf,s}-F_{a\to s}^{(liq)}T_{atm,s}+F_{s\to a}^{(ice)}T_{surf,s}-F_{a\to s}^{(ice)}T_{atm,s}\Big]+ \\
&\quad F_{s\to a}^{(wv)}L_v(T_{surf,s})-F_{a\to s}^{(wv)}L_v(T_{atm,s})-F_{s\to a}^{(ice)}L_f(T_{surf,s})+F_{a\to s}^{(ice)}L_f(T_{atm,s}) \\
&= \sum_{\ell\in\mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[h_{00}^{(liq)}-c_p^{(\ell)}T_{00}\right]+F_{net}^{(wv)}L_{v,00}-F_{net}^{(ice)}L_{f,00}+ \\
&\quad \sum_{\ell\in\mathcal{L}_{H_2O}} F_{s\to a}^{(\ell)}c_p^{(\ell)}T_{surf,s}-\sum_{\ell\in\mathcal{L}_{H_2O}} F_{a\to s}^{(\ell)}c_p^{(\ell)}T_{atm,s}.
\end{aligned}
$$

(liquid reference state)

$$
\begin{aligned}
F_{net}^{(h)} &= F_{net}^{(H_2O)}\left(h_{00}^{(ice)}-c_p^{(ice)}T_{00}\right)+c_p^{(ice)}\Big[F_{s\to a}^{(wv)}T_{surf,s}-F_{a\to s}^{(wv)}T_{atm,s}+ \\
&\quad F_{s\to a}^{(liq)}T_{surf,s}-F_{a\to s}^{(liq)}T_{atm,s}+F_{s\to a}^{(ice)}T_{surf,s}-F_{a\to s}^{(ice)}T_{atm,s}\Big]+ \\
&\quad F_{s\to a}^{(wv)}L_s(T_{surf,s})-F_{a\to s}^{(wv)}L_s(T_{atm,s})+F_{s\to a}^{(ice)}L_f(T_{surf,s})-F_{a\to s}^{(ice)}L_f(T_{atm,s}) \\
&= \sum_{\ell\in\mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[h_{00}^{(ice)}-c_p^{(\ell)}T_{00}\right]+F_{net}^{(wv)}L_{s,00}+F_{net}^{(liq)}L_{f,00}+ \\
&\quad \sum_{\ell\in\mathcal{L}_{H_2O}} F_{s\to a}^{(\ell)}c_p^{(\ell)}T_{surf,s}-\sum_{\ell\in\mathcal{L}_{H_2O}} F_{a\to s}^{(\ell)}c_p^{(\ell)}T_{atm,s}.
\end{aligned}
$$

(ice reference state)

If the interface temperature is the same in both atmosphere and the surface, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$ theformulas are vastly simplified

$$F_{net}^{(h)} \approx F_{net}^{(H_2O)}\left[c_p^{(wv)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(wv)}\right] - F_{net}^{(liq)} L_v(\widetilde{T}_s) - F_{net}^{(ice)} L_s(\widetilde{T}_s) \quad \text{(C4)}$$

$$\approx \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[c_p^{(\ell)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(wv)}\right] - F_{net}^{(liq)} L_{v,00} - F_{net}^{(ice)} L_{s,00} \quad \text{(C5)}$$

(water vapor reference state, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

$$F_{net}^{(h)} \approx F_{net}^{(H_2O)}\left[c_p^{(liq)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(liq)}\right] + F_{net}^{(wv)} L_v(\widetilde{T}_s) - F_{net}^{(ice)} L_f(\widetilde{T}_s) \quad \text{(C6)}$$

$$\approx \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[c_p^{(\ell)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(liq)}\right] + F_{net}^{(wv)} L_{v,00} - F_{net}^{(ice)} L_{f,00} \quad \text{(C7)}$$

(liquid reference state, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

and

$$F_{net}^{(h)} \approx F_{net}^{(H_2O)}\left[c_p^{(ice)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(ice)}\right] + F_{net}^{(wv)} L_s(\widetilde{T}_s) + F_{net}^{(liq)} L_f(\widetilde{T}_s) \quad \text{(C8)}$$

$$\approx \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[c_p^{(\ell)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(ice)}\right] + F_{net}^{(wv)} L_{s,00} + F_{net}^{(liq)} L_{f,00}, \quad \text{(C9)}$$

(ice reference state, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

respectively.

C3 Final energy equations with different reference states

The total energy equation with different reference enthalpies are

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\Bigg\{\sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)}\left(K + \Phi_s\right) + c_p^{(d)} T + m^{(H_2O)}\left[c_p^{(wv)}\left(T - T_{00}\right) + h_{00}^{(wv)}\right] - m^{(liq)} L_v(T) - m^{(ice)} L_s(T)\Bigg\} dA\, dz =$$

$$\iint \Bigg\{F_{net}^{(H_2O)}(\widetilde{K}_s + \Phi_s) + F_{net}^{(H_2O)}\left[c_p^{(wv)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(wv)}\right] - F_{net}^{(liq)} L_v(\widetilde{T}_s) - F_{net}^{(ice)} L_s(\widetilde{T}_s) + F_{net}^{(turb,rad)}\Bigg\} dA. \quad \text{(C10)}$$

(water vapor reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

or, equivalently, by expanding latent heat terms and re-arranging terms:

$$\frac{\partial}{\partial t}\iiint \rho^{(d)}\Bigg\{K + \Phi_s + c_p^{(d)} T + \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)}\left[K + \Phi_s + c_p^{(\ell)}\left(T - T_{00}\right) + h_{00}^{(wv)}\right] - m^{(liq)} L_{v,00} - m^{(ice)} L_{s,00}\Bigg\} dA\, dz$$

$$= \iint \Bigg\{\sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)}\left[\widetilde{K}_s + \Phi_s + c_p^{(\ell)}\left(\widetilde{T}_s - T_{00}\right) + h_{00}^{(wv)}\right] - F_{net}^{(liq)} L_{v,00} - F_{net}^{(ice)} L_{s,00} + F_{net}^{(turb,rad)}\Bigg\} dA. \quad \text{(C11)}$$

(water vapor reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

For liquid and ice as a reference states the total energy equation takes the form

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \left(K + \Phi_s\right) + c_p^{(d)} T + m^{(H_2O)} \left[c_p^{(liq)} \left(T - T_{00}\right) + h_{00}^{(liq)} \right] + m^{(wv)} L_v(T) - m^{(ice)} L_f(T) \Bigg\} dA\, dz$$
$$= \iint \Bigg\{ F_{net}^{(H_2O)} (\widetilde{K}_s + \Phi_s) + F_{net}^{(H_2O)} \left[c_p^{(liq)} \left(\widetilde{T}_s - T_{00} \right) + h_{00}^{(liq)} \right] + F_{net}^{(wv)} L_v(\widetilde{T}_s) - F_{net}^{(ice)} L_f(\widetilde{T}_s) + F_{net}^{(turb,rad)} \Bigg\} dA. \quad \text{(C12)}$$

(liquid reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

or, equivalently, by expanding latent heat terms and re-arranging terms:

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ K + \Phi_s + c_p^{(d)} T + \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)} \left[K + \Phi_s + c_p^{(\ell)} \left(T - T_{00}\right) + h_{00}^{(wv)} \right] + m^{(wv)} L_{v,00} - m^{(ice)} L_{f,00} \Bigg\} dA\, dz$$
$$= \iint \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)} \left[\widetilde{K}_s + \Phi_s + c_p^{(\ell)} \left(\widetilde{T}_s - T_{00} \right) + h_{00}^{(wv)} \right] + F_{net}^{(wv)} L_{v,00} - F_{net}^{(ice)} L_{f,00} + F_{net}^{(turb,rad)} \Bigg\} dA. \quad \text{(C13)}$$

(liquid reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ \sum_{\ell \in \mathcal{L}_{all}} m^{(\ell)} \left(K + \Phi_s\right) + c_p^{(d)} T + m^{(H_2O)} \left[c_p^{(ice)} \left(T - T_{00}\right) + h_{00}^{(ice)} \right] + m^{(wv)} L_s(T) + m^{(liq)} L_f(T) \Bigg\} dA\, dz$$
$$= \iint \Bigg\{ F_{net}^{(H_2O)} (\widetilde{K}_s + \Phi_s) + F_{net}^{(H_2O)} \left[c_p^{(ice)} + \left(\widetilde{T}_s - T_{00} \right) + h_{00}^{(ice)} \right] + F_{net}^{(wv)} L_s(\widetilde{T}_s) + F_{net}^{(liq)} L_f(\widetilde{T}_s) + F_{net}^{(turb,rad)} \Bigg\} dA, \quad \text{(C14)}$$

(ice reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

or, equivalently, by expanding latent heat terms and re-arranging terms:

$$\frac{\partial}{\partial t} \iiint \rho^{(d)} \Bigg\{ K + \Phi_s + c_p^{(d)} T + \sum_{\ell \in \mathcal{L}_{H_2O}} m^{(\ell)} \left[K + \Phi_s + c_p^{(\ell)} \left(T - T_{00}\right) + h_{00}^{(ice)} \right] + m^{(wv)} L_{s,00} + m^{(liq)} L_{f,00} \Bigg\} dA\, dz$$
$$= \iint \Bigg\{ \sum_{\ell \in \mathcal{L}_{H_2O}} F_{net}^{(\ell)} \left[\widetilde{K}_s + \Phi_s + c_p^{(\ell)} \left(\widetilde{T}_s - T_{00} \right) + h_{00}^{(ice)} \right] + F_{net}^{(wv)} L_{s,00} + F_{net}^{(liq)} L_{f,00} + F_{net}^{(turb,rad)} \Bigg\} dA. \quad \text{(C15)}$$

(ice reference enthalpy, $\widetilde{T}_s \equiv T_{atm,s} = T_{surf,s}$)

Appendix D Ocean enthalpy fluxes

Let's briefly consider the interface condition for atmosphere-ocean coupling:

$$\underbrace{h^{(sw)}F_{net}^{(H_2O)} + \mathcal{F}_{ocn}^{(heat)}}_{\text{ocean}} = \underbrace{h^{(wv)}F_{net}^{(wv)} + h^{(liq)}F_{net}^{(liq)} + h^{(ice)}F_{net}^{(ice)} + F_{net}^{(turb,rad)}}_{\text{atmosphere}}, \tag{D1}$$

where $h^{(sw)}$ is the enthalpy of seawater and $\mathcal{F}_{ocn}^{(heat)}$ is the net heat flux out of the ocean (which is the heat flux that the ocean "sees"). Note that h is, in general, temperature dependent $h^{(\ell)} = h(T^{(\ell)})$, $\ell =' wv',' liq',' ice'$ and the temperatures of each species $T^{(\ell)}$ are, in general, different for different forms of water. The enthalpy of seawater, $h^{(sw)}$, is a much more complicated function involving pressure and a nonlinear dependence on temperature (TEOS-10; `http://www.teos10.org/`). In the literature $F_{net}^{(wv)}$ is usually denoted $\mathcal{E}$ and refers only to evaporation. Precipitation is denoted $\mathcal{P} = -F_{net}^{(liq)} - F_{net}^{(ice)}$, which is usually divided into rain and snow $\mathcal{P} = \mathcal{P}^{(rain)} + \mathcal{P}^{(snow)}$. In this case (D1) becomes

$$h^{(sw)}\left(\mathcal{E} - \mathcal{P}\right) + \mathcal{F}_{ocn}^{(heat)} = h^{(wv)}\mathcal{E} - h^{(liq)}\mathcal{P}^{(rain)} - h^{(ice)}\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)}, \tag{D2}$$

meaning that the net heat flux out of the ocean is

$$\mathcal{F}_{ocn}^{(heat)} = \left[h^{(wv)} - h^{(sw)}\right]\mathcal{E} - \left[h^{(liq)} - h^{(sw)}\right]\mathcal{P}^{(rain)} - \left[h^{(ice)} - h^{(sw)}\right]\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)}. \tag{D3}$$

For reasons that will become clear, let's rewrite (D3) as

$$\mathcal{F}_{ocn}^{(heat)} = \left[h^{(wv)} - h^{(liq)}\right]\mathcal{E} - \left[h^{(ice)} - h^{(liq)}\right]\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)} + , \left[h^{(liq)} - h^{(sw)}\right]\left(\mathcal{E} - \mathcal{P}\right). \tag{D4}$$

If we assume a single interface temperature and a liquid reference state, then

$$L_v(\widetilde{T}_s) = h^{(wv)}(\widetilde{T}_s) - h^{(liq)}(\widetilde{T}_s), \tag{D5}$$

$$L_f(\widetilde{T}_s) = -\left[h^{(ice)}(\widetilde{T}_s) - h^{(liq)}(\widetilde{T}_s)\right], \tag{D6}$$

(assume liquid reference state: (C1), (C2), (C3))

and (D4) can be written as

$$\mathcal{F}_{ocn}^{(heat)} = L_v(\widetilde{T}_s)\mathcal{E} + L_f(\widetilde{T}_s)\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)} + \left[h^{(liq)} - h^{(sw)}\right]\left(\mathcal{E} - \mathcal{P}\right), \tag{D7}$$

$$= c_p^{(wv)}\left(\widetilde{T}_s - T_{00}\right)\mathcal{E} + c_p^{(liq)}\left(\widetilde{T}_s - T_{00}\right)\left[\mathcal{P}^{(snow)} - \mathcal{E}\right] - \left[c_p^{(ice)}\left(\widetilde{T}_s - T_{00}\right)\right]\mathcal{P}^{(snow)}, + \mathcal{E}L_{v,00} - \mathcal{P}^{(snow)}L_{f,00} + F_{net}^{(turb,rad)} + \left[h^{(liq)} - h^{(sw)}\right]\left(\mathcal{E} - \mathcal{P}\right), \tag{D8}$$

(assuming liquid reference and interface temperatures the same: $T_{atm,s} = T_{surf,s} = \widetilde{T}_s$)

where in (D8) the latent heat terms have been expanded (using a liquid reference state) and the terms have been re-arranged. Assuming constant latent heats (equivalent to assuming all the heat capacities are the same), and assuming that $h^{(liq)} = h^{(sw)}$, then (D8) simplifies to

$$\mathcal{F}_{ocn}^{(heat)} \approx L_{v,00}\mathcal{E} - L_{f,00}\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)}. \tag{D9}$$

(assume $L_v(T) \equiv L_{v,00}$, $L_f(T) \equiv L_{f,00}$, and $h^{(liq)} \equiv h^{(sw)}$)

If one specifies sea surface water to have the temperature $T^{(ocn)}$, evaporated water vapor to have the same temperature as the sea surface $T^{(ocn)}$ and all forms of precipitation to have the same temperature $T^{(atm)}$, then the more general formula (D3) must be used

$$\mathcal{F}_{ocn}^{(heat)} = h^{(wv)}\left(T^{(ocn)}\right)\mathcal{E} - h^{(liq)}\left(T^{(atm)}\right)\mathcal{P}^{(rain)} - h^{(ice)}\left(T^{(atm)}\right)\mathcal{P}^{(snow)} + F_{net}^{(turb,rad)} - h^{(sw)}\left(T^{(ocn)}\right)\left(\mathcal{E} - \mathcal{P}\right), \tag{D10}$$

which is the interface condition used in Mayer et al. (2017) (their equation (18)).

Appendix E Mechanical dissipation of kinetic energy and resultant heating

The winds in the atmosphere and the surface (e.g. currents in the ocean) exchange momentum through a surface stress $(\vec{\tau}_z)_{atm,s}$ and $(\vec{\tau}_z)_{surf,s}$, respectively. For both fluids this stress vector may be interpreted as having the relevant components of the stress tensor at the interface. The stress tensor was introduced in Section 2.1.12 and discussed in terms of internal stresses and energy conservation. In terms of the interaction between atmosphere and surface, each component is responsible for constructing a stress profile that then can be used to redistribute momentum and any associated dissipative heating. In general, the two stress vectors have to match at the interface; except when the latter is constituted by a third dynamical component, for example ocean waves, which has no explicit spatial representation (and therefore no explicit boundary with the atmosphere and/or the ocean).

The total rate of change of kinetic energy in each atmospheric layer due to internal stresses was given in (99). Integrating over the entire column we get

$$\int_{z_s}^{z_t} \frac{\partial}{\partial t}\left(\rho^{(all)} K\right) dz = (\vec{v}\cdot\vec{\tau}_z)\big|_{z=z_s} - (\vec{v}\cdot\vec{\tau}_z)\big|_{z=z_t} + \int_{z_s}^{z_t} \vec{\tau}_z \cdot \frac{\partial \vec{v}}{\partial z} dz. \tag{E1}$$

As discussed in Section 2.1.12, the last term represents dissipation of kinetic energy in the interior of the atmosphere due to Reynolds stresses (see e.g. Landau and Lifschitz (1968) section 16, Peixoto and Oort (1992) section 3.2.1, Duran et al. (2018)), but also includes unresolved turbulence or wave-breaking processes, and possibly numerical diffusion terms (these various sources will determine the form of $\vec{\tau}_z$). This kinetic energy loss term is normally accounted for and balanced by sensible heating in individual physics parameterizations (e.g. for the planetary boundary layer scheme and for gravity wave drag). The second term on the right-hand side of (E1) represents the work performed on air by the stress at the upper boundary. For high-top models the stress at the upper boundary makes a very minor contribution (for the purposes of the general circulation) arising only from interactions with the geomagnetic field and the solar wind (on the other hand, for the ocean, it represents the source of mechanical energy for almost its entire circulation). For a closed atmospheric energy budget it is therefore desirable to ensure that a total stress $\tau_z = 0$ is applied at the top of the computational domain. That said, the lower the model top, the worse the approximation of vanishing stress at the model top becomes. For example, in the lower thermosphere local tendencies caused by gravity wave propagation are still important. However, even in this case the total work remains negligible in the global average. In numerical models the top layers usually are subject to increased artificial damping to avoid wave reflection from the model top. These artificial numerical stresses at the upper boundary may result in a significant non-physical contribution to the column energy budget which is not balanced by physical terms (more on this in Section 4.1.2).

At the lower boundary of the atmosphere, in contrast, the corresponding term on the right-hand side of (E1) vanishes only if the boundary condition $\vec{v}_s \equiv \vec{v}(z = z_s) = 0$ applies. This is assumed true over solid ground (land, ice sheets etc). Over the ocean, surface currents imply $\vec{v}_s \neq 0$ and the common formulation for $\vec{\tau}_z\big|_{z=z_s}$ is

$$\vec{\tau}_z\big|_{z=z_s} = -C_D \|\vec{v}(z_1) - \vec{v}_s\| (\vec{v}(z_1) - \vec{v}_s) \tag{E2}$$

where C_D is a drag coefficient which depends on the sea-state, near-surface atmospheric quantities and $z = z_1$ the thickness of the surface boundary layer. The thickness of the surface boundary layer is usually assumed to be the altitude of the lowest grid point of the atmospheric model (i.e. $\vec{v}(z_1)$ is the lowest dynamically resolved wind). The ocean surface velocity dependence (a.k.a. the current feedback) in the surface wind stress (E2) is not systematically taken into account in climate models. However, such dependence leads to a reduction of roughly 20%–35% in the rate of kinetic energy input (or equivalently in the work done by the wind) to the ocean circulation (Duhaut & Straub, 2006).

This modulation of the energy transfer does not significantly affect the atmospheric component, but has large impacts on the oceanic surface eddy kinetic energy (Renault et al., 2016). In addition to the current feedback, the surface wave field also plays a significant role in the air-sea momentum exchange. In the current generation of climate models, wave-induced processes are often ignored and all of the surface wind-stress goes into direct forcing of the surface currents, such that $(\vec{\tau}_z)_{ocean,s} = \vec{\tau}_z\big|_{z=z_s}$ (i.e. no gain or loss of net kinetic energy by surface waves). This is an assumption, since part of the work performed by the atmosphere is absorbed by the ocean surface wave field and contributes to the growth and maintenance of surface waves.

We also note that global atmosphere models normally do not include the interface layer between atmospheric domain and the surface explicitly in the computational domain. This has implications for the energy budget, because the lowest resolved air speed (whether in the dynamical core or in the sub-grid parameterizations) is non-zero even over solid ground — and in general it is different from the velocity of the surface (e.g. the ocean current). The surface stress therefore appears to perform work, which is lost from the global budget. This work is, in fact, entirely transformed into heat within the interface layer, which is defined as the thin boundary layer of air near the surface within which both heating and stress are zero and the corresponding vertical fluxes of heat and momentum constant. For momentum, the implication is that within this layer energy is dissipated according to the last term on the r.h.s. of Equation (E1), which across the surface layer integrates to $\vec{\tau_z}|_{z=z_s}\cdot\vec{v}|_1$, where $\vec{v}|_1$ is the lowest dynamically resolved wind. As expected, this matches the stress work, and results in a noticeable heating term, which is positive definite everywhere and thus contributes to a appreciable global mean. One way to represent this process in an energetically consistent way (so that the energy dissipation is computed and added as heating, or, equivalently, as shear production) is described in Becker and Burkhardt (2007). We note that computing the linear (constant coefficient) Laplacian diffusion operator requires ensuring that the correction terms are not spurious sources of energy (e.g. angular momentum). Nonlinear diffusion requires even more terms, and extending this approach to terrain-following coordinates requires even more terms (see, respectively, equation (18), (17) and Section 3 in Becker & Burkhardt, 2007).

Appendix F Use of the barycentric velocity as a basis for consistently representing the frictional dissipation of rain in atmospheric models

This section makes use of vector notation using bold-face fonts. For example, the barycentric velocity vector is denoted $\mathbf{v}$ and is a three-dimensional velocity vector, whereas the main part of the paper is using $\vec{v}$ to refer to the horizontal velocity vector.

F1 Introduction

Pauluis et al. (2000) and Pauluis and Dias (2012) estimated the frictional dissipation due to Stokes friction between precipitation and gas to be globally comparable to the frictional dissipation due to the turbulent kinetic energy cascade. Locally in the tropics, the dissipation due to falling rain might dominate and could be as large as 10 W/m^2. This process must be represented in numerical models. It is argued here that this process is automatically included into the model if the barycentric (center of mass) velocity is used in the model equations.

When representing thermodynamic processes from the perspective of the resolved model equations, we have to take into account that normal and tangential stresses in the vicinity of falling precipitation are not resolved. Microphysics equations are valid for droplets of specific sizes for which the Stokes friction law can be accessed directly. Contemporary parameterization schemes for microphysics deliver an integrated precipitation flux $\mathbf{P}$ over

all droplet sizes to the coarse grained model equations. The precipitation flux can be interpreted to be the difference between the bulk velocity of the precipitation, superscript *prec*, and the gaseous part of the air, superscript *gas*, hence $\mathbf{P} = \rho^{(prec)}(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)})$. A bulk terminal velocity characterizing all droplets is then defined as $-V^T\mathbf{k} = \mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}$, where $\mathbf{k}$ is the vertical unit vector.

The terminal velocity of a droplet is obtained from the equilibrium of the Stokes friction force, the gravitational force and the buoyancy force. At microscales, when the shear zones of the droplets is resolved by a model, the friction force is larger than the buoyancy force. At the coarse grained macroscales, however, this Stokes friction force is not directly accessible. Therefore, the equilibrium is represented as a balance between the gravitational force and a buoyancy (pressure) force. This shift in interpretation correctly models the energy transfer processes. The lost potential energy of a droplet is first converted into kinetic energy and then converted into internal energy by friction. An alternative representation is that the conversion of kinetic energy into heat is also achieved by the pressure gradient force. In the following it will be shown that this process is indeed recovered when using the barycentric velocity of an air parcel and the associated kinetic energy.

The concept of the barycentric velocity is documented in de Groot and Mazur (1984); Zdunkowski and Bott (2004); Degrauwe et al. (2016); Catry et al. (2007); Gassmann and Herzog (2008); Wacker et al. (2006). A macroscopic air parcel is considered to consist of diverse species ℓ, which may have different bulk velocities. The barycentric velocity is then defined to be

$$\mathbf{v} = \sum_{\ell \in \mathcal{L}_{all}} \rho^{(\ell)}\mathbf{v}^{(\ell)}/\rho^{(all)} = \sum_{\ell \in \mathcal{L}_{all}} q^{(\ell)}\mathbf{v}^{(\ell)}, \tag{F1}$$

where $\rho^{(all)} = \sum_{\ell \in \mathcal{L}_{all}} \rho^{(\ell)}$ is the total density. The second expression uses the specific contents $q^{(\ell)} = \rho^{(\ell)}/\rho^{(all)}$ with $\sum_{\ell \in \mathcal{L}_{all}} q^{(\ell)} = 1$ in the definition of the barycentric velocity. This means that sedimenting condensates have downward relative velocities with respect to the barycentric velocity. The non-sedimenting gaseous and cloud constituents move slightly upwards compared to the barycentric velocity (Wacker & Herbert, 2003). Note that even though this framework allows for different horizontal velocities between, e.g. gas and precipitation, the horizontal velocity components of $\mathbf{v}$ are usually assumed the same.

Diffusive fluxes are defined as the difference between the actual flux of a given species and the barycentric flux

$$\mathbf{J}^{(\ell)} = \rho^{(\ell)}(\mathbf{v}^{(\ell)} - \mathbf{v}), \text{ (diffusive flux).}$$

Note that the diffusive flux in a barycentric velocity frame must not be confused with molecular diffusion fluxes, and also not with turbulent diffusion fluxes arising from Reynolds averaging/filtering. By definition, the sum of all diffusive fluxes is zero

$$\sum_{\ell \in \mathcal{L}_{all}} \mathbf{J}^{(\ell)} = 0. \tag{F2}$$

It might seem that the sum of the diffusive fluxes should not vanish at the surface (Bott, 2008). If it rains, there, a vanishing flux sum would be associated with an artificial flux of dry air through the surface. In fact, the sum vanishes as well by definition at the surface. Rather the barycentric velocity is non-zero there. When evaporation and precipitation take place at the same time, we can separate the velocity of water vapor into two distinct contributions, one which is equal to the dry air velocity, and a second part that differs from the dry air velocity and only appears if dry air and water vapor mix $\mathbf{v}^{(wv)} = \mathbf{v}^{(d)} + \mathbf{v}^{(wv)}_{mix}$. Then the surface barycentric mass flux is

$$\rho^{(all)}\mathbf{v}_{surf} = \rho^{(d)}\mathbf{v}^{(d)} + \rho^{(wv)}\mathbf{v}^{(d)} + \rho^{(wv)}\mathbf{v}^{(wv)}_{mix} + \rho^{(prec)}\mathbf{v}^{(prec)}$$

$$
\begin{aligned}
&= \underbrace{\rho^{(prec)}\mathbf{v}^{(prec)} - \rho^{(gas)}\mathbf{v}^{(d)}}_{=\mathbf{P}} + \underbrace{\rho^{(wv)}\mathbf{v}^{(wv)}_{mix} - \rho^{(d)}\mathbf{v}^{(d)}}_{=\mathbf{E}} + \underbrace{(2\rho^{(d)} + \rho^{(wv)})\mathbf{v}^{(d)}}_{=0} \\
&= \mathbf{P} + \mathbf{E}.
\end{aligned}
$$

Clearly, the surface dry air velocity vanishes, $\mathbf{v}^{(d)} = 0$. The here given form is only chosen for the sake of making the definitions of $\mathbf{P}$ and $\mathbf{E}$ obvious, which are both always formally defined with the help of mass flux differences. In the above equation it is assumed that suspended condensates move with the dry air velocity and hence do not show up in the boundary condition. This boundary condition is applicable to all prognostic equations, especially also for the momentum equation, which means that momentum leaves the atmosphere if it rains.

F2 Prognostic equations for species

The continuity equation for each specie can be written in two different flavors (note that these are three-dimensional continuity equations. The gradient operator ∇ is three-dimensional!)

$$
\begin{aligned}
\frac{\partial \rho^{(\ell)}}{\partial t} &= -\nabla \cdot \left(\rho^{(\ell)}\mathbf{v}^{(\ell)}\right) + S^{(\ell)}, \\
&= -\nabla \cdot \left(\rho^{(\ell)}\mathbf{v} + \mathbf{J}^{(\ell)}\right) + S^{(\ell)},
\end{aligned} \tag{F3}
$$

(see section 2a in Wacker et al., 2006), where the second form is using the barycentric perspective. $S^{(\ell)}$ is the source/sink term. Note that total mass is conserved, so

$$
\sum_{\ell \in \mathcal{L}_{all}} S^{(\ell)} = 0, . \tag{F4}
$$

We note that the constituent equation can also be written in Lagrangian form

$$
\rho^{(all)} \frac{Dq^{(\ell)}}{Dt} = -\nabla \cdot \mathbf{J}^{(\ell)} + S^{(\ell)}, \tag{F5}
$$

where

$$
\frac{D}{Dt} = \partial_t + \mathbf{v} \cdot \nabla \text{ (barycentric material derivative).} \tag{F6}
$$

Often, models do not use a prognostic form for the equations of falling precipitation. Rather, the precipitation is thought to fall down immediately. This is expressed as $0 = -\nabla \cdot \mathbf{J}^{(\ell)} + S^{(\ell)}$. Hence, the barycentric framework is easily applicable for this case, since the required rule $\sum_{\ell \in \mathcal{L}} \mathbf{J}^{(\ell)} = 0$ is untouched by this computational simplification.

Summing the continuity equations for all components gives

$$
\begin{aligned}
\frac{\partial \rho^{(all)}}{\partial t} &= \sum_{\ell \in \mathcal{L}_{all}} \frac{\partial \rho^{(\ell)}}{\partial t}, \\
&= -\sum_{\ell \in \mathcal{L}_{all}} \left[\nabla \cdot \left(\rho^{(\ell)}\mathbf{v} + \mathbf{J}^{(\ell)}\right) + S^{(\ell)}\right], \\
&= -\nabla \cdot (\rho^{(all)}\mathbf{v}).
\end{aligned} \tag{F7}
$$

As said above, at the surface, the barycentric velocity is the sum of the precipitation flux and the evaporation flux. Hence, the lower boundary conditions for the continuity equation are easily obtained when applying the barycentric framework. The general and intuitive form of the continuity equation (F7) allows to derive its primitive equation formulation as we are used to it. Specifically then, the total density is replaced with $\rho^{(all)} = -\partial p/\partial \Phi$.

Specifically, we can also give the continuity equations for the species independent of the gas velocity, which is likewise the velocity of the other non-precipitating species

like cloud ice and cloud water. For comparison, we repeat here the two forms for the continuity equation of the non-precipitating constituents, hence the sum of the gas and the non-precipitating liquid and frozen parts

$$\frac{\partial \rho^{(non-prec)}}{\partial t} = -\nabla \cdot (\rho^{(non-prec)} \mathbf{v}^{(gas)}) + \mathcal{S}^{(non-prec)}, \tag{F8}$$

$$\frac{\partial \rho^{(non-prec)}}{\partial t} = -\nabla \cdot (\rho^{(non-prec)} \mathbf{v} - q^{(non-prec)} \mathbf{P}) + \mathcal{S}^{(non-prec)}, \tag{F9}$$

where the first equation assumes $\mathbf{v}^{(gas)}$ to be known. The second equation assumes an available $\mathbf{v}$ and uses the diffusive flux of the non-precipitating species as

$$\begin{aligned} \mathbf{J}^{(non-prec)} &= \rho^{(non-prec)} \mathbf{v}^{(gas)} - q^{(non-prec)} \rho^{(all)} \mathbf{v} \\ &= \rho^{(non-prec)} \mathbf{v}^{(gas)} - q^{(non-prec)} (\rho^{(non-prec)} \mathbf{v}^{(gas)} + \rho^{(prec)} \mathbf{v}^{(prec)}) \\ &= \rho^{(non-prec)} \mathbf{v}^{(gas)} - q^{(non-prec)} (\rho^{(non-prec)} \mathbf{v}^{(gas)} + \mathbf{P} + \rho^{(prec)} \mathbf{v}^{(gas)}) \\ &= -q^{(non-prec)} \mathbf{P} \end{aligned}$$

Likewise, the possible prognostic equations for the precipitation are found to be

$$\frac{\partial \rho^{(prec)}}{\partial t} = -\nabla \cdot (\rho^{(prec)} \mathbf{v}^{(gas)} + \mathbf{P}) + \mathcal{S}^{(prec)}, \tag{F10}$$

$$\frac{\partial \rho^{(prec)}}{\partial t} = -\nabla \cdot (\rho^{(prec)} \mathbf{v} + q^{(non-prec)} \mathbf{P}) + \mathcal{S}^{(prec)}. \tag{F11}$$

The second equation using the barycentric framework includes the diffusive precipitation flux as $\mathbf{J}^{(prec)} = q^{(non-prec)} \mathbf{P}$.

The different velocities are related according to

$$\mathbf{v} = \mathbf{v}^{(gas)} + \mathbf{P}/\rho^{(all)}. \tag{F12}$$

Therefore, modelers may choose either prognostic equation for the species in their numerical model. For instance, when modeling passive tracer transport, $\mathbf{v}^{(gas)}$ is a more natural choice for the advective velocity than performing a two-step process consisting of an advection with the barycentric velocity and a correction with the diffusive velocity. A common counterargument against the use of barycentric velocities is that we cannot measure them directly. However, if the precipitation flux, the wind velocity — and hence the velocity of the non-precipitating part — and the density are known, the recovery of the barycentric velocity is straightforward.

F3 Prognostic equation for momentum and the kinetic energy budget

Let us write the resolved momentum equations for the gaseous part, the precipitating part and the cloud part independently

$$\rho^{(all)} \frac{D}{Dt} q^{(gas)} \mathbf{v}^{(gas)} = -\nabla p - \rho^{(gas)} \nabla \Phi - \mathbf{F}_{gp} - \mathbf{F}_{gc}, \tag{F13}$$

$$\rho^{(all)} \frac{D}{Dt} q^{(prec)} \mathbf{v}^{(prec)} = -\rho^{(prec)} \nabla \Phi + \mathbf{F}_{gp}, \tag{F14}$$

$$\rho^{(all)} \frac{D}{Dt} q^{(cloud)} \mathbf{v}^{(gas)} = -\rho^{(cloud)} \nabla \Phi + \mathbf{F}_{gc}, \tag{F15}$$

where turbulent diffusion terms and the Coriolis term are omitted (we could add them later). The momentum exchange forces $\mathbf{F}$ subsume all forces which are not tractable by the coarse-grained model. The force $\mathbf{F}_{gp}$ is exerted with opposite signs to the moist air around the precipitation particles and the bulk precipitation velocity. The force $\mathbf{F}_{gc}$ accounts for a similar interplay between gas and cloud particles. Several possible reasons exist for such forces. The most obvious is the frictional force, but also, momentum might

be exchanged because of condensation or evaporation. The forces $\mathbf{F}$ are not known, and they are not provided by any microphysics parameterization of the coarse-grained model. At the microphysics level, the comparable Stokes friction depends on the radius of the droplets and the shear zones formed around these droplets. However, on the macroscopic level these radii are not known and the shear zones are not represented.

The assumption of a common — namely barycentric — advective velocity $\mathbf{v}$ is necessary here, because the forces $\pm\mathbf{F}$ are relative forces between the gas and the droplets. Therefore, all species need a common Lagrangian frame of reference for their motion.

The forces $\mathbf{F}$ are not explicit in the barycentric momentum equation,

$$\rho^{(all)} \frac{D}{Dt}\mathbf{v} = -\nabla p - \rho^{(all)} \nabla \Phi. \tag{F16}$$

Here, again, possible additional turbulent diffusion and Coriolis terms are omitted. The fact that the forces $\mathbf{F}$ are not explicit means that using the barycentric velocity as the prognostic variable is an attractive choice. An apparent complication is that the barycentric velocity is needed for advection in this momentum equation.

Let us derive the kinetic energy change of an air parcel moving with the barycentric velocity. This is obtained by multiplying (F16) by the barycentric velocity

$$\begin{aligned}
&(q^{(non-prec)}\mathbf{v}^{(gas)} + q^{(prec)}\mathbf{v}^{(prec)}) \cdot \rho^{(all)} \frac{D}{Dt}\mathbf{v} = \\
&\underbrace{-\mathbf{v}^{(gas)} \cdot \nabla p}_{\text{pressure work}} \underbrace{-\rho^{(non-prec)}\mathbf{v}^{(gas)} \cdot \nabla\Phi - \rho^{(prec)}\mathbf{v}^{(prec)} \cdot \nabla\Phi}_{K \leftrightarrow \text{potential energy}} \underbrace{-q^{(prec)}(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}) \cdot \nabla p}_{K \leftrightarrow \text{internal forces precip-air}}
\end{aligned} \tag{F17}$$

.

The first three terms on the right represent individual energy changes which are attributable to known physical processes. The first term is the pressure work term that converts kinetic energy into internal energy via the compression work $-\mathbf{v}^{(gas)}\cdot\nabla p = -\nabla\cdot(\mathbf{v}^{(gas)}p) + p\nabla\cdot\mathbf{v}^{(gas)}$. The two next terms describe the exchange of kinetic energy with the potential energy, i.e. the rising or falling air parcels. The fourth term represents the kinetic energy change due internal forces between the precipitation and the moist air. This term is identified as an effect of internal friction because the other physical processes are already accounted for. Forces between the cloud droplets and the gas do not result in a kinetic energy change, because the gas and the cloud droplets have the same velocity.

Let us now devote our attention to the last term in (F17). Assuming that the hydrostatic relation dominates,

$$\mathbf{k}\cdot\nabla p = \frac{\partial p}{\partial z} \approx -\rho^{(all)} g, \tag{F18}$$

and taking the definition of the terminal velocity $-V^T\mathbf{k} = \mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}$, the macroscopic effect of Stokes friction can be approximated as

$$-q^{(prec)}\nabla p\cdot(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}) \approx -q^{(prec)}\rho^{(all)} g V^T = -\rho^{(prec)} g V^T \text{ (macr. effect of Stokes friction).}$$

It is then clear that this term represents dissipation of kinetic energy

$$|-\rho^{(prec)} V^T g| > 0. \tag{F19}$$

Pauluis et al. (2000) denotes this as the frictional dissipation of rain. In conclusion we note that this dissipation is only accessible if the barycentric velocity is chosen to be prognostic. This is because, otherwise, the particular energy budget (F17) would not have been obtained.

In a hydrostatic primitive equation modeling framework, the left hand side of (F17) would be set to zero. This means that kinetic energy of vertical motions is not stored

in a kinetic energy reservoir, but pretends to be converted directly between potential energy and internal energy. Besides this, the physical processes are represented in the same manner in a primitive-hydrostatic modeling environment as they are in a non-hydrostatic modeling environment.

F4 Internal energy and enthalpy equations

After having analyzed the loss of kinetic energy due to falling rain, we have to find/derive the terms where the respective gain of internal energy shows up as a positive temperature tendency. We will cover both the constant height formulation (internal energy) and the constant pressure formulation (enthalpy).

The internal energy equation (20) using the barycentric velocity formulation can be written as

$$\rho^{(all)}\frac{DU^{(all)}}{Dt} = - \underbrace{p\nabla\cdot\mathbf{v}}_{\text{compression term}} -\nabla\cdot\left[\sum_{\ell\in\mathcal{L}_{all}} \underbrace{h^{(\ell)}(T)\mathbf{J}^{(\ell)}}_{\text{diffusive enthalpy flux of the constituents}}\right] - \Gamma. \tag{F20}$$

(Gassmann & Herzog, 2015) where Γ represents three terms not central for this discussion (divergences of radiation and sensible heat fluxes, turbulent shear production term) and $U^{(all)}$ is the specific internal energy (see (22))

$$U^{(all)} = \sum_{\ell\in\mathcal{L}_{all}} q^{(\ell)}\left[h_{00}^{(\ell)} + c_p^{(\ell)}(T-T_{00})\right] - \frac{p^{(d)}}{\rho^{(all)}} - \frac{p^{(wv)}}{\rho^{(all)}}.$$

The internal energy equation combines temperature and constituent changes and we are now after the temperature change. Before starting, let us replace the pressures in the above internal energy specification by their expression of the ideal gas law, hence $p^{(d)}/\rho^{(all)} = R^{(d)}q^{(d)}T$ and $p^{(wv)}/\rho^{(all)} = R^{(wv)}q^{(wv)}T$. For gaseous components we have $c_v^{(d)} = c_p^{(d)} - R^{(d)}$ and $c_v^{(wv)} = c_p^{(wv)} - R^{(wv)}$, whereas for the condensates the heat capacities at constant volume and constant pressure are formally equal, $c_v^{(\ell)} = c_p^{(\ell)}$ for $\ell \in \mathcal{L}_{cond}$. That is,

$$\begin{aligned}
\frac{p^{(d)}}{\rho^{(all)}} + \frac{p^{(wv)}}{\rho^{(all)}} &= \frac{R^{(d)}\rho^{(d)}T}{\rho^{(all)}} + \frac{R^{(wv)}\rho^{(wv)}T}{\rho^{(all)}}, \\
&= R^{(d)}q^{(d)}T + R^{(wv)}q^{(wv)}T, \\
&= \left[c_p^{(d)} - c_v^{(d)}\right]q^{(d)}T + \left[c_p^{(wv)} - c_v^{(wv)}\right]q^{(wv)}T, \\
&= \sum_{\ell\in\mathcal{L}_{all}}\left[c_p^{(\ell)} - c_v^{(\ell)}\right]q^{(\ell)}T, \text{ using } c_v^{(\ell)} = c_p^{(\ell)} \text{ for } \ell\in\mathcal{L}_{cond}
\end{aligned}$$

Now, we use that to rewrite the left-hand side of (F20)

$$\begin{aligned}
\rho^{(all)}\frac{DU^{(all)}}{Dt} &= \rho^{(all)}\frac{D}{Dt}\left(\sum_{\ell\in\mathcal{L}_{all}} q^{(\ell)}c_v^{(\ell)}T + (h_{00}^{(\ell)} - c_p^{(\ell)}T_{00})q^{(\ell)}\right), \\
&= \rho^{(all)}\sum_{\ell\in\mathcal{L}_{all}}\left\{q^{(\ell)}c_v^{(\ell)}\frac{DT}{Dt} + (Tc_v^{(\ell)} + h_{00}^{(\ell)} - c_p^{(\ell)}T_{00})\frac{Dq^{(\ell)}}{Dt}\right\}, \text{ (chain rule)} \\
&= \sum_{\ell\in\mathcal{L}_{all}}\left\{\rho^{(\ell)}c_v^{(\ell)}\frac{DT}{Dt} - (Tc_v^{(\ell)} + h_{00}^{(\ell)} - c_p^{(\ell)}T_{00})(\nabla\cdot\mathbf{J}^{(\ell)} - S^{(\ell)})\right\}, \qquad \text{(F21)}
\end{aligned}$$

using the constituent equation (F5). Second, the diffusive enthalpy flux term can be written as

$$-\nabla\cdot\left[\sum_{\ell\in\mathcal{L}_{all}} h^{(\ell)}(T)\mathbf{J}^{(\ell)}\right] = -\sum_{\ell\in\mathcal{L}_{all}}\nabla\cdot\left\{\left[h_{00}^{(\ell)} + c_p^{(\ell)}(T-T_{00})\right]\mathbf{J}^{(\ell)}\right\},$$

$$= \quad -\sum_{\ell\in\mathcal{L}_{all}} \left\{\nabla\cdot\left[c_p^{(\ell)}T\mathbf{J}^{(\ell)}\right] + (h_{00}^{(\ell)} - c_p^{(\ell)}T_{00})\nabla\cdot\mathbf{J}^{(\ell)}\right\} \quad \text{(F22)}$$

Now, it is clear that the last term in (F22) cancels with a part of (F21). Substituting (F21) on the left-hand and the rewrite (F22) on the right hand side of the internal energy equation (F20) yields after rearranging, using again $c_v^{(\ell)} = c_p^{(\ell)} - R^{(\ell)}$ and the ideal gas law

$$c_v^{(all)}\rho^{(all)}\frac{DT}{Dt} = -p\nabla\cdot\mathbf{v} - \Gamma \underbrace{-\sum_{\ell\in\mathcal{L}_{all}}(h_{00} - c_p^{(\ell)}T_{00} + c_v^{(\ell)}T)S^{(\ell)}}_{\text{latent heating}}$$
$$\underbrace{-\nabla\cdot\sum_{\ell\in\mathcal{L}_{gas}}\frac{p^{(\ell)}}{\rho^{(\ell)}}\mathbf{J}^{(\ell)}}_{\text{diffusive pressure flux}} \underbrace{-\sum_{\ell\in\mathcal{L}_{all}} c_v^{(\ell)}\mathbf{J}^{(\ell)}\cdot\nabla T}_{\text{adv. of } T \text{ with diffusive constituent fluxes}} . \quad \text{(F23)}$$

The important terms are the compression term and the diffusive pressure flux term. For the latter we do not yet have an interpretation and, indeed, there is no stand-alone interpretation as we shall see soon. We can only understand the frictional heating due to falling rain when we consider both terms together. To demonstrate this, recall the definitions of the previous sections: $\mathbf{P} = \rho^{(prec)}(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}) = -\rho^{(prec)}V^T\mathbf{k}$ and $\mathbf{J}^{(gas)} = -q^{(gas)}\mathbf{P}$. Using them to express the diffusive pressure flux yields

$$\sum_{\ell\in\mathcal{L}_{gas}}\frac{p^{(\ell)}}{\rho^{(\ell)}}\mathbf{J}^{(\ell)} = -\frac{p\mathbf{P}}{\rho^{(all)}} = pq^{(prec)}(\mathbf{v}^{(gas)} - \mathbf{v}^{(prec)}) = pq^{(prec)}V^T\mathbf{k} \quad \text{(F24)}$$

Now, use the definition of the barycentric velocity $\mathbf{v} = q^{(non-prec)}\mathbf{v}^{(gas)} + q^{(prec)}\mathbf{v}^{(prec)}$ in the compression term and differentiate the diffusive pressure flux term by parts. This gives after some cancellations and using $q^{(non-prec)} + q^{(prec)} = 1$

$$\begin{aligned} -p\nabla\cdot\mathbf{v} - \nabla\cdot\left(\sum_{\ell\in\mathcal{L}_{gas}}\frac{p^{(\ell)}}{\rho^{(\ell)}}\mathbf{J}^{(\ell)}\right) &= -p\nabla\cdot\mathbf{v}^{(gas)} + q^{(prec)}(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)})\cdot\nabla p \\ &= -p\nabla\cdot\mathbf{v}^{(gas)} + \frac{\mathbf{P}}{\rho^{(all)}}\cdot\nabla p \\ &\approx -p\nabla\cdot\mathbf{v}^{(gas)} + \rho^{(prec)}V^T g \end{aligned}$$

where the last term is again found by using the hydrostatic approximation. With this term we have finally discovered the place where the frictional dissipation term shows up in the temperature equation. Only the gaseous components contribute to the compression term, but because the compression term shows up in the original equation with the barycentric velocity, a part of it constitutes, together with the diffusive pressure flux term, the frictional heating due to the falling precipitation. Comparing the result with the kinetic energy loss (F17), it becomes clear that total energy conservation is obtained with this formulation. The lost kinetic energy shows up exactly, with the opposite sign, in the internal energy equation, and hence gives rise to an increase of temperature.

It is interesting to give an estimate of the relative orders of magnitude of the dissipative heating term due to precipitation and the last term in (F23), which accounts for the differential advection of the species besides the barycentric advection. When considering only rain fluxes below a cloud, it becomes

$$-\sum_{\ell\in\mathcal{L}_{all}} c_v^{(\ell)}\mathbf{J}^{(\ell)}\cdot\nabla T = -\mathbf{P}^{(rain)}(q^{(d)}(c_p^{(liq)} - c_v^{(d)}) - q^{(wv)}(c_p^{(liq)} - c_v^{(wv)}))\cdot\nabla T$$

Assuming the worst case of a neutral stratification with a vertical temperature gradient of $-g/c_p^{(all)}$, and considering the thermodynamic parameters given in Appendix F,

the differential advection leads to a cooling which is about three times larger than the frictional heating effect. This calls into question the assumption of identical heat capacities for all the species, because in this case, the last term of (F23) would vanish. The physical mechanism behind this cooling is that the specific heat of liquid or frozen species is considerably higher than the specific heat of the gases. Therefore, the net transportation of heat is downward, even though the sum of the diffusive fluxes vanishes. Thus, there is a small cooling even without evaporation of rain.

What are the implications for a temperature equation in the height based coordinate system? Assume that the precipitation fluxes $\mathbf{P}^{(\ell)}$ are the output of some parameterisation scheme. Then, equation (F23) can be cast in the following form

$$c_v^{(all)}\rho^{(all)}\frac{DT}{Dt} = -p\nabla\cdot\mathbf{v} - \Gamma' - \nabla\cdot\left(\sum_{\ell\in\mathcal{L}_{prec}}\left(-\frac{p}{\rho^{(all)}}\mathbf{P}^{(\ell)}\right) + \sum_{\ell\in\mathcal{L}_{gas}} R^{(\ell)}T\mathbf{J}^{(\ell)}_{mix}\right) - \sum_{\ell\in\mathcal{L}_{all}} c_v^{(\ell)}\mathbf{J}^{(\ell)}\cdot\nabla T. \tag{F25}$$

where Γ' now also comprises the latent heating term. For the last term, the precipitation fluxes have to be converted into the diffusive fluxes. Note that in this case all non-sedimenting constituents, which might also be cloud water or cloud ice in addition to the gas, share the same relative velocity. Hence we set $\mathbf{J}^{(\ell)} = -q^{(\ell)}\mathbf{P}$ for every non-precipitating constituent and $\mathbf{J}^{(\ell)} = q^{(non-prec)}\mathbf{P}$ for a precipitating component. This must be done independently for different precipitation fluxes, say those of rain, snow, etc. There might also be fluxes of gaseous components due to turbulent mixing. These are represented in the second term under the divergence operator. Finally, the diffusive fluxes in the last term might be summed up from quite different contributions.

Writing the thermodynamic equation (21) in enthalpy form – as it is convenient in the primitive equation's modeling framework – we get

$$c_p^{(all)}\rho^{(all)}\frac{DT}{Dt} = \frac{Dp}{Dt} - \Gamma \underbrace{- \sum_{\ell\in\mathcal{L}_{all}} (h_{00} + c_p^{(\ell)}(T - T_{00}))S^{(\ell)}}_{\text{latent heating}} \underbrace{- \sum_{\ell\in\mathcal{L}_{all}} c_p^{(\ell)}\mathbf{J}^{(\ell)}\cdot\nabla T}_{\text{adv. of } T \text{ with diffusive constituent fluxes}}, \tag{F26}$$

It is remarkable that no additional term appears, compared with the existing formulation. This is because the barycentric vertical pressure velocity is not "just" the gas pressure vertical velocity when taking the barycentric framework into account

$$\begin{aligned}
\frac{Dp}{Dt} &= \frac{\partial p}{\partial t} + \mathbf{v}\cdot\nabla p, \\
&= \frac{\partial p}{\partial t} + \sum_{\ell\in\mathcal{L}_{all}} q^{(\ell)}\mathbf{v}^{(\ell)}\cdot\nabla p, \\
&= \frac{\partial p}{\partial t} + \sum_{\ell\neq' prec'} q^{(\ell)}\mathbf{v}^{(gas)}\cdot\nabla p + q^{(prec)}\mathbf{v}^{(prec)}\cdot\nabla p, \\
&= \frac{\partial p}{\partial t} + \sum_{\ell\in\mathcal{L}_{all}} q^{(\ell)}\mathbf{v}^{(gas)}\cdot\nabla p + q^{(prec)}\left(\mathbf{v}^{(prec)} - \mathbf{v}^{(gas)}\right)\cdot\nabla p, \\
&= \underbrace{\frac{\partial p}{\partial t} + \mathbf{v}^{(gas)}\cdot\nabla p}_{=\omega^{(gas)}} - q^{(prec)}V^T\mathbf{k}\cdot\nabla p, \\
&= \frac{\partial p}{\partial t} + \mathbf{v}^{(gas)}\cdot\nabla p + \rho^{(prec)}V^T g,
\end{aligned}$$

The last term is here the frictional dissipation term. In contrast to the internal energy form (F25), the frictional heating due to falling precipitation is completely hidden in the pressure vertical velocity. This velocity becomes larger when precipitation is falling, indicating a stronger apparent subsidence.

The internal energy (F25) and enthalpy (F26) forms of the temperature equation assign the heating to all ingredients of the air parcel to the extent that they all have the same temperature, but different heat capacities and different mass concentrations. It lies at the very heart of our current understanding of non-equilibrium thermodynamics that a single temperature is assigned to the air parcel consisting of the diverse species. This is an idealization which is heavily challenged, but an alternative would require multiple temperature equations for multiple species. Since we describe droplet size distributions with only bulk quantities and smaller droplets certainly have different temperatures than bigger droplets, an alternative approach quickly becomes confusing.

In section 2.1.13 we speculated that the barycentric framework could possibly also handle horizontal velocity deviations/friction for falling precipitation. In that case we would have to define a terminal velocity so that it not only has a vertical component, but also a horizontal component. The barycentric viewpoint already formally also accounts for the horizontal momentum of falling precipitation. The challenge is that in the microphysics parameterization scheme it would be unclear how to compute a terminal velocity in the horizontal direction as that is not governed by Stokes law and the empirical formula used there.

Appendix G CAM setup

The simulations with CAM use the `cam_energy_analysis` branch of `https://github.com/PeterHjortLauritzen/CAM` which is a fork from CAM tag *cam*6_2_026 (tag in `cam_developement` branch of `https://github.com/ESCOMP/CAM`). The simulations use the CAM-SE-CSLAM dynamical core (Lauritzen et al., 2017) based on a dry-mass vertical coordinate (Lauritzen et al., 2018) and coupling to physics using a finite-volume physics grid (Herrington et al., 2018, 2019). A run-script to replicate the simulations can be found in the `cam_energy_analysis` code base under the *my_scripts* directory.

Reproducibility information on the simulations for Figure 9 is stored in `git@github.com:oksanaguba/cam-enthalpy-runs.git`. The runs are based on the same branch `cam_energy_analysis` as described above.

The following constants in the energy equation are used in CAM:

$$
\begin{aligned}
c_p^{(d)} &= 1004.64\, J\, K^{-1}\, kg^{-1} \\
c_p^{(wv)} &= 1810.00\, J\, K^{-1}\, kg^{-1} \\
c_p^{(liq)} &= 4188.00\, J\, K^{-1}\, kg^{-1} \\
c_p^{(ice)} &= 2117.27\, J\, K^{-1}\, kg^{-1} \\
L_{v,00} &= 2501000.0\, J\, kg^{-1} \\
L_{s,00} &= 2834700.0\, J\, kg^{-1} \\
L_{f,00} &= 333700.0\, J\, kg^{-1}
\end{aligned}
$$

This version of the spectral-element dynamical core differs from the CESM2.2 release version (Lauritzen et al., 2018) in that an FV3 vertical remapping algorithm is used (which reduced vertical energy conservation due to higher-order numerics) and reference states are subtracted from hyperviscosity operators to reduce spurious noise for flow over orography (Liu et al., 2022) and thereby also energy errors.

Appendix H Sponge-layer diffusion and frictional heating

As discussed in Section 4, frictional heating is often implemented in the dynamical cores of ESMs after the application of second-order or higher-order horizontal diffusion processes and sponge-layer dissipation mechanisms, such as Rayleigh friction, near the model top. These processes are represented on the right-hand-side of the momen-

tum equation as also shown for the Laplacian diffusion in (136). In addition, horizontal fourth-order hyperdiffusion and Rayleigh friction are represented respectively by

$$\frac{\partial \vec{v}}{\partial t} = \cdots - \nu_4 \nabla^4 \vec{v} \quad \text{and} \tag{H1}$$

$$\frac{\partial \vec{v}}{\partial t} = \cdots - k\,\vec{v}, \tag{H2}$$

where ν_4 is a fourth-order horizontal diffusion coefficient (see also the discussion in Jablonowski and Williamson (2011)) and k is the Rayleigh friction damping coefficient.

Section 4.1.2 highlighted a particular application of the sponge-layer Rayleigh friction in the FV3 dynamical core (L. M. Harris et al., 2021). FV3 lies at the heart of the Unified Forecast System (UFS) developed by the National Oceanic and Atmospheric Administration (NOAA) and is an optional dynamical core in NCAR's CESM2.2 framework. We therefore provide additional insight into the implementation details to foster the reproducibility of the FV3 results. They were generated via CESM2.2's "Simpler Models" hierarchy which includes the dry Held-Suarez test (Held & Suarez, 1994) as the "FHS94 compset" with analytic initial conditions. FV3's Rayleigh friction in the sponge layer is implemented in a time-implicit way. When using the acceleration of the zonal wind component $\partial u/\partial t = \cdots - ku$ as an example this leads to

$$\frac{u^{n+1} - u^n}{\Delta t} = -k\,u^{n+1}, \tag{H3}$$

$$u^{n+1} = \frac{u^n}{1 + k\,\Delta t}, \tag{H4}$$

where $n+1$ and n denote the updated (via the RF process) and current time levels, respectively. The time-implicit update of the meridional wind component v follows the same principle and is given by

$$v^{n+1} = \frac{v^n}{1 + k\,\Delta t}. \tag{H5}$$

The loss of kinetic energy due to Rayleigh friction in the hydrostatic FV3 configuration with the horizontal velocity vector $\vec{v} = (u, v)^T$ can then be computed via

$$\begin{aligned}
\delta K &= K_h^{n+1} - K_h^n, \\
&= \frac{\rho}{2}\left[\left(\vec{v}_h^{\,n+1}\right)^2 - \left(\vec{v}_h^{\,n}\right)^2\right], \\
&= \frac{\rho}{2}\left[u^{n+1}u^{n+1} + v^{n+1}v^{n+1} - u^n u^n + v^n v^n\right], \\
&= \frac{\rho}{2}\left[\left(\frac{u^n}{1 + k\,\Delta t}\right)^2 + \left(\frac{v^n}{1 + k\,\Delta t}\right)^2 - \left(u^n\right)^2 - \left(v^n\right)^2\right], \\
&= \frac{\rho}{2}\left[\left(\left(u^n\right)^2 + \left(v^n\right)^2\right)\left(\frac{1}{(1 + k\Delta t)^2} - 1\right)\right], \\
&= K_h^n \left(\frac{1}{(1 + k\Delta t)^2} - 1\right).
\end{aligned} \tag{H6}$$

This kinetic energy loss is applied as frictional heating (-δK) as shown in (140). It leads to the expression

$$\begin{aligned}
T^{n+1} &= T^n - \frac{\delta K}{\rho^{(all)}\, c_p^{(d)}}, \\
&= T^n + \frac{K_h^n}{c_p}\left(1 - \frac{1}{(1 + k\Delta t)^2}\right),
\end{aligned} \tag{H7}$$

which describes the local update of the temperature, here labeled as time level $n + 1$. A dry specific heat is used since there is negligible moisture in the sponge layer domain.

L. M. Harris et al. (2021) show the equivalent temperature update equation due to frictional heating (their equation (8.17)) for the nonhydrostatic FV3 configuration. It is

$$\begin{aligned} T^{n+1} &= T^n - \frac{\delta K_{3D}}{\rho^{(all)}\, c_v^{(d)}}, \\ &= T^n + \frac{K_{3D}^n}{c_v^{(d)}} \left(1 - \frac{1}{(1+k\Delta t)^2} \right), \end{aligned} \tag{H8}$$

where K_{3D} is the kinetic energy that takes all three velocity components into account. This heat conversion utilizes the dry specific heat at constant volume $c_v^{(d)}$ in FV3.

The choice of the Rayleigh friction damping coefficient in FV3 is

$$k = \frac{1}{\tau} \left[\sin \left(\frac{\pi}{2} \frac{\ln(p_c/p)}{\ln(p_c/p_t)} \right) \right]^2, \tag{H9}$$

where p symbolizes the pressure at the chosen grid point. The symbol p_c stands for the user-selected RF cutoff pressure, p_t is the pressure at the model top, and τ is the maximum RF timescale which needs to be expressed in units of seconds. For the FV3 C24L64 simulations in Section 4.1.2 the following settings were selected: $p_t = 9.04{\times}10^{-5}$ hPa, $p_c = 0.1$ hPa and $\tau = 1$, 3, and 5 days.

Appendix I Notation

Below is list of variables appearing more than once in the main part of this paper:

symbol	description	unit
$\overline{(\cdot)}$	cell averaged/mean value	
$(\cdot)'$	deviation from cell averaged/mean value	
α	specific volume of moist air	m^3/kg
$\alpha^{(d)}$	specific volume of dry air	m^3/kg
$c_p^{(\ell)}$	specific heat at constant pressure of species ℓ	$J/K/kg$
$c_v^{(\ell)}$	specific heat at constant volume of species ℓ	$J/K/kg$
D/DT	material derivative	$1/s$
E_x	specific energy of x where x='atm' atmosphere and x='ESM' Earth System model	J/kg
E_{feom}	specific energy of fluid equations of motion	J/kg
$F_{net}^{(\ell)}$	net flux of water species ℓ into the atmosphere	$kg/m^2/s$
$F_{a\to s}^{(\ell)}$	flux of water species ℓ from atmosphere to surface	$kg/m^2/s$
$F_{s\to a}^{(\ell)}$	flux of water species ℓ from surface to atmosphere	$kg/m^2/s$
$F_{net}^{(turb,rad)}$	radiative and sensible/turbulent fluxes into atmosphere (76)	$J/m^2/s$
$\mathcal{F}_{net}^{(turb)}$	sensible/turbulent fluxes into atmosphere consistent with CLUBB	$J/m^2/s$
$F_{net}^{(h)}$	net enthalpy flux into atmosphere	$J/s/m^2$
$F_{a\to s}^{(h)}$	enthalpy flux from atmosphere to surface	$J/s/m^2$
$F_{s\to a}^{(h)}$	enthalpy flux from surface to atmosphere	$J/s/m^2$
$F_{net}^{(\Phi)}$	net potential energy flux into atmosphere	$J/s/m^2$
$F_{net}^{(K)}$	net (horizontal) kinetic energy flux into atmosphere	$J/s/m^2$
$\vec{F}_h$	frictional force on velocity vector	m/s^2
$\vec{\tau}_z$	vertical component of stress tensor	m^2/s^2
g	gravitational constant	$m^3/kg/s^2$
$h^{(\ell)}$	partial enthalpy of air constituent ℓ	J/kg
$h_{00}^{(\ell)}$	reference enthalpy for water phase ℓ	J/kg
k	vertical unit vector	
K	specific horizontal kinetic energy ($\equiv \frac{1}{2}\vec{v}^2$)	m^2/s^2
$K_{surf,s}$	specific horizontal kinetic energy in surface at surface	m^2/s^2
$K_{atm,s}$	specific horizontal kinetic energy in atmosphere at surface	m^2/s^2
$\widetilde{K}_s$	specific horizontal kinetic energy at surface	m^2/s^2
δK	specific horizontal kinetic energy increment	m^2/s^2
$m^{(\ell)}$	dry mixing ratio of constituent ℓ ($\equiv \rho^{(\ell)}/\rho^{(d)}$)	kg/kg
$m^{(all)}$	sum of all dry mixing ratios of moist air	kg/kg
$\frac{\partial \hat{m}}{\partial t}^{(\ell)}$	rate of phase change to falling precipitation or evaporation	$kg/kg/s$
$\frac{\partial \hat{m}}{\partial t}^{(\ell)}$	rate of phase change to/from suspended condensate	$kg/kg/s$
$M^{(\ell)}(z)$	mass per unit area at height z for component ℓ	kg/m^2
$L_{f,00}$	latent heat of fusion	J/kg
$L_{s,00}$	latent heat of sublimation	J/kg
$L_{v,00}$	latent heat of vaporization	J/kg
$\mathcal{L}_{all}$	set of all constituents of moist air	
$\mathcal{L}_{H_2O}$	set of all water species in moist air	
$\mathcal{L}_{cond}$	set of all condensates in moist air	
Φ	geopotential	m^2/s^2
Φ_s	surface geopotential	m^2/s^2
$\rho^{(\ell)}$	density of air constituent ℓ	kg/m^3
$\rho^{(all)}$	density of all constituents of moist air	kg/m^3

symbol	description	unit
p	moist pressure	Pa
p_t	moist pressure at model top	Pa
$p^{(d)}$	partial pressure of dry air	Pa
$p^{(wv)}$	partial pressure of water vapor	Pa
Π	Exner pressure	
Π_s	Exner pressure at surface	
$q^{(\ell)}$	specific/moist mixing ratio of ℓ	kg/kg
Q	heating per unit mass	J/kg
R	generalized ideal gas constant	$J/K/mol$
$R^{(\ell)}$	ideal gas constant for $\ell = `d', `wv'$	$J/K/mol$
$(\cdot)_{t^n}$	variable held fixed during physics update	
T	temperature	K
ΔT	temperature increment	K
ΔT_k	temperature increment in atmosphere level k	K
T_{00}	reference temperature	K
$T_{atm,s}$	temperature in atmosphere at surface	K
$T_{surf,s}$	temperature in surface at surface	K
$\widetilde{T}_s$	single temperature for water leaving/entering atmosphere	K
T_v	virtual temperature	K
δT	temperature increment	K
θ	potential temperature	K
θ_l	moist potential temperature	K
$U^{(all)}$	specific internal energy of moist air	J/kg
$U^{(d)}$	specific internal energy of dry air	J/kg
$\vec{v}$	horizontal velocity vector (u, v)	m/s
δu	zonal velocity component increment	m/s
δv	meridional velocity component increment	m/s
$\mathbf{v}$	3D barycentric velocity vector (only used in Appendix F)	m/s
V^T	bulk terminal vertical velocity of falling precipitation (used in barycentric framework)	m/s
ν_2	horizontal second-order diffusion coefficient	m^2/s
w	vertical velocity	m/s
ω	vertical pressure velocity ($\equiv Dp/Dt$)	Pa/s
z	height above surface geoid	m
z_t	height of model top	m
η	standard hybrid pressure-based vertical coordinate	
ΔI	used for various imbalance terms in energy equation (see text for details)	J/s or W/m^2

Acknowledgments

We are especially grateful that the Banff International Research Station (BIRS) approved our proposal for the week-long workshop Physics-Dynamics Coupling in ESMs (19w5153). Without this support, this paper would not exist. BIRS provided a beautiful location and ideal facilities for a true workshop that encouraged lively scientific discussion and debate on the fundamental challenges facing Earth Systems models.

We thank the reviewers and editors for taking the time to review and handle this unusually long and complex paper. It has taken many iterations (and countless Zoom meetings during the pandemic) to cast our ideas/discussions/solutions into publishable form and we are extremely grateful that the reviewers and editors have appreciation for

that. Initially we were also planning on covering ocean models which would have made the paper even longer and more complex!

We thank BIRS participants who contributed to the many discussions at the BIRS meeting but ended up not contributing to this paper (A.J. Adcroft, H. Johansen, D.A. Randall and R. Kumar).

We thank Richard Neale (NCAR) for suggesting the *insulated rain-shaft* terminology. P.H. Lauritzen acknowledges funding support by NSF under Cooperative Agreement No. 1852977. Computing resources (doi:10.5065/D6RX99HX) were provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and other agencies. N.K.-R. Kevlahan acknowledges funding from an NSERC Discovery grant and allocation of computer time from Compute Ontario (`computeontario.ca`) and Compute Canada (`www.computecanada.ca`). T. Toniazzo acknowledges funding from the Norwegian Research Council (project "INES", # 270061, and "KeyClim", # 295046) and from the European Research Council (project "highECS" # 770765). C. Jablonowski acknowledges the support from the NOAA grant NA17OAR4320152 (127). O. Guba acknowledges funding support from the Energy Exascale ESM (E3SM) project, funded by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research (BER) and by the DOE Office of Science, Advanced Scientific Computing Research (ASCR) Program under the Scientific Discovery through Advanced Computing (SciDAC 4) ASCR/BER Partnership Program. O. Guba's research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; and used a high-performance computing cluster provided by the BER ESMing program and operated by the Laboratory Computing Resource Center at Argonne National Laboratory. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Almut Gassmann was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 274762653 – TRR 181. A. Donahue's contribution was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 IM Release LLNL-JRNL-827510. B. Harrop's contribution was supported by the U.S. Department of Energy Office of Science Biological and Environmental Research (BER) as part of the Regional and Global Climate Modeling program. H. Wan's contribution was supported by the U.S. DOE's SciDAC program via a partnership in Earth System Model Development between the Office of Biological and Environmental Research and the Office of Advanced Scientific Computing Research. Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC05-76RL01830.

The data for the energy analysis is available here `https://doi.org/10.5281/zenodo.6819581`.

References

Adams, M., Colella, P., Graves, D. T., Johnson, J. N., Johansen, H. S., Keen, N. D., ... Straalen, B. V. (2019). Chombo software package for amr applications - design document [Computer software manual]. (Technical Report LBNL-6616E)

Adams, S., Ford, R., Hambley, M., Hobson, J., Kavčič, I., Maynard, C., ... Wong, R. (2019). LFRic: Meeting the challenges of scalability and performance portability in weather and climate models. *Journal of Parallel and Distributed Computing*, *132*, 383-396. doi: https://doi.org/10.1016/j.jpdc.2019.02.007

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., ... Zhang, R. (2019). The GFDL global ocean and sea ice model om4.0: Model description and simulation features. *J. Adv. Model. Earth Syst.*, *11*, 3167-3211. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001726` doi: 10.1029/2019MS001726

Adkins, C. J. (1968). *Equilibrium thermodynamics*. McGraw-Hill.

Akmaev, R. A. (2008). On the energetics of maximum-entropy temperature profiles. *Quart. J. Roy. Meteor. Soc.*, *134*(630), 187-197. doi: https://doi.org/10.1002/qj.209

Anderson, S., Hinton, A., & Weller, R. A. (1998). Moored observations of precipitation temperature. *J. Atmos. Oceanic Technol.*, *15*, 979-986. doi: https://doi.org/10.1175/1520-0426(1998)015⟨0979:MOOPT⟩2.0.CO;2

Arakawa, A., Jung, J.-H., & Wu, C.-M. (2011). Toward unification of the multiscale modeling of the atmosphere. *Atmospheric Chem. Phys.*, *11*(8), 3731–3742. doi: 10.5194/acp-11-3731-2011

Arakawa, A., & Lamb, V. R. (1981). A potential enstrophy and energy conserving scheme for the shallow water equations. *Mon. Wea. Rev.*, *109*, 18-36. doi: 10.1175/1520-0493(1981)109⟨0018:APEAEC⟩2.0.CO;2

Arakawa, A., & Wu, C.-M. (2013). A unified representation of deep moist convection in numerical modeling of the atmosphere. part i. *J. Atmos. Sci.*, *70*(7), 1977–1992. doi: 10.1175/JAS-D-12-0330.1

Bannon, P. R. (2002). Theoretical foundations for models of moist convection. *J. Atm. Sci.*, *59*, 1967–1982. doi: 10.1175/1520-0469(2002)059%3C1967:TFFMOM%3E2.0.CO;2

Bannon, P. R. (2003). Hamiltonian description of idealized binary geophysical fluids. *J. Atmos. Sci.*, *60*(22), 2809-2819. doi: 10.1175/1520-0469(2003)060⟨2809:HDOIBG⟩2.0.CO;2

Bauer, W., & Gay-Balmaz, F. (2019). Variational discretization framework for geophysical flow models. In F. Nielsen & F. Barbaresco (Eds.), *Geometric science of information* (pp. 523–531). Cham: Springer International Publishing. doi: https://doi.org/10.1007/978-3-030-26980-7_54

Becker, E. (2001). Symmetric stress tensor formulation of horizontal momentum diffusion in global models of atmospheric circulation. *J. Atmos. Sci.*, *58*(3), 269–282. doi: 10.1175/1520-0469(2001)058⟨0269:SSTFOH⟩2.0.CO;2

Becker, E. (2003). Frictional heating in global climate models. *Mon. Wea. Rev.*, *131*(3), 508 - 520. doi: 10.1175/1520-0493(2003)131⟨0508:FHIGCM⟩2.0.CO;2

Becker, E., & Burkhardt, U. (2007). Nonlinear horizontal diffusion for GCMs. *Mon. Wea. Rev.*, *135*(4), 1439 - 1454. doi: 10.1175/MWR3348.1

BIRS. (2019). *Physics–dynamics coupling in Earth System Models (19w5153).* `https://www.birs.ca/events/2019/5-day-workshops/19w5153`. (Accessed: 2021-09-28)

Bister, M., & Emanuel, K. A. (1998). Dissipative heating and hurricane intensity. *Meteorology and Atmospheric Physics*, *65*, 233–240. doi: https://doi.org/10.1007/BF01030791

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., & Schanen, D. P. (2013). Higher-order turbulence closure and its impact on climate simulations in the Community Atmosphere Model. *J. Climate*, *26*, 9655–9676. doi: 10.1175/JCLID-13-00075.1

Bott, A. (2008). Theoretical considerations on the mass and energy consistent treatment of precipitation in cloudy atmospheres. *Atmospheric Research*, *89*, 262–269. doi: 10.1016/j.atmosres.2008.02.010

Bowen, P., & Thuburn, J. (2022a). Consistent and flexible thermodynamics in atmospheric models using internal energy as a thermodynamic potential. Part I: Equilibrium regime. *Quart. J. Roy. Meteor. Soc.*.

Bowen, P., & Thuburn, J. (2022b). Consistent and flexible thermodynamics in atmospheric models using internal energy as a thermodynamic potential. Part II: Non-equilibrium regime. *Quart. J. Roy. Meteor. Soc.*.

Brecht, R., Bauer, W., Bihlo, A., Gay-Balmaz, F., & MacLachlan, S. (2019). Variational integrator for the rotating shallow-water equations on the sphere. *Quart. J. Roy. Meteor. Soc.*, *145*(720), 1070-1088. doi: 10.1002/qj.3477

Brown, B. B., Vasil, G. M., & Zweibel, E. G. (2012). Energy conservation and gravity waves in sound-proof treatements of stellar interiors. part i. anelastic approximations. *The Astrophysical Journal*, *756*(109). doi: http://dx.doi.org/10.1088/0004-637X/756/2/109

Byers, H. R., Moses, H., & Harney, P. J. (1949, feb). Measurement of rain temperature. *J. Meteorol.*, *6*(1), 51–55. doi: 10.1175/1520-0469(1949)006⟨0051:MORT⟩2.0.CO;2

Catry, B., Geleyn, J.-F., Tudor, M., Bénard, P., & Trojáakováa, A. (2007). Flux-conservative thermodynamic equations in a mass-weighted framework. *Tellus A: Dynamic Meteorology and Oceanography*, *59*(1), 71-79. doi: 10.1111/j.1600-0870.2006.00212.x

Cohen, D., & Hairer, E. (2011, Mar 01). Linear energy-preserving integrators for poisson systems. *BIT Numerical Mathematics*, *51*(1), 91–101. doi: 10.1007/s10543-011-0310-z

Cotton, W. R., Bryan, G. H., & Van den Heever, S. C. (2011). *Storm and cloud dynamics: the dynamics of clouds and precipitating mesoscale systems* (2. ed ed.) (No. 99). Amsterdam: Acad. Press. (OCLC: 837354620)

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., ... others (2020). The Community Earth System Model version 2 (CESM2). *J. Adv. Model. Earth Syst.*, *12*(2), e2019MS001916. doi: https://doi.org/10.1029/2019MS001916

Degrauwe, D., Seity, Y., Bouyssel, F., & Termonia, P. (2016). Generalization and application of the flux-conservative thermodynamic equations in the AROME model of the ALADIN system. *Geosci. Model Dev.*, *9*(6), 2129–2142. doi: 10.5194/gmd-9-2129-2016

de Groot, S. R., & Mazur, P. (1984). *Non-equilibrium thermodynamics*. New York: Dover Publications, Inc. doi: https://doi.org/10.1063/1.3050930

Dewar, W. K., Schoonover, J., McDougall, T. J., & Klein, R. (2016). Semicompressible ocean thermodynamics and boussinesq energy conservation. *Fluids*, *1*(9). doi: https://doi.org/10.3390/fluids1020009

Donahue, A. S., & Caldwell, P. M. (2020). Performance and accuracy implications of parallel split physics-dynamics coupling in the energy exascale Earth system atmosphere model. *J. Adv. Model. Earth Syst.*, *12*(7), e2020MS002080. doi: https://doi.org/10.1029/2020MS002080

Dubos, T., Dubey, S., Tort, M., Mittal, R., Meurdesoif, Y., & Hourdin, F. (2015). Dynamico-1.0, an icosahedral hydrostatic dynamical core designed for consistency and versatility. *Geosci. Model Dev.*, *8*(10), 3131–3150. doi: 10.5194/gmd-8-3131-2015

Dubos, T., & Kevlahan, N. K.-R. (2013). A conservative adaptive wavelet method for the shallow water equations on staggered grids. *Q.J.R. Meteorol. Soc.*, *139*, 1997–2020. doi: 10.1002/qj.2097

Dubos, T., & Tort, M. (2014). Equations of atmospheric motion in non-Eulerian vertical coordinates: Vector-invariant form and quasi-Hamiltonian formulation. *Mon. Wea. Rev.*, *142*(10), 3860-3880. doi: 10.1175/MWR-D-14-00069.1

Duhaut, T. H. A., & Straub, D. N. (2006). Wind stress dependence on ocean surface velocity: Implications for mechanical energy input to ocean circulation. *J.*

Phys. Oceanogr., *36*(2), 202 - 211. doi: 10.1175/JPO2842.1

Duran, I. B., Geley, J.-F., Vana, F., Schmidli, J., & Brozkova, R. (2018). A turbulence scheme with two prognostic turbulence energies. *J. Atmos. Sci.*, *75*, 3381–3402. doi: 10.1175/JAS-D-18-0026.1

Eden, C. (2016, 2). Closing the energy cycle in an ocean model. *Ocean Modelling*, *101*. doi: 10.1016/j.ocemod.2016.02.005

Eldred, C., Dubos, T., & Kritsikis, E. (2019). A quasi-Hamiltonian discretization of the thermal shallow water equations. *J. Comput. Phys.*, *379*, 1 - 31. doi: https://doi.org/10.1016/j.jcp.2018.10.038

Eldred, C., & Gay-Balmaz, F. (2020, sep). Single and double generator bracket formulations of multicomponent fluids with irreversible processes. *Journal of Physics A: Mathematical and Theoretical*, *53*(39), 395701. doi: 10.1088/1751-8121/ab91d3

Eldred, C., & Gay-Balmaz, F. (2021, aug). Thermodynamically consistent semi-compressible fluids: a variational perspective. *Journal of Physics A: Mathematical and Theoretical*, *54*(34), 345701. doi: 10.1088/1751-8121/ac1384

Eldred, C., & Randall, D. (2017). Total energy and potential enstrophy conserving schemes for the shallow water equations using hamiltonian methods – part 1: Derivation and properties. *Geosci. Model Dev.*, *10*(2), 791–810. doi: 10.5194/gmd-10-791-2017

Eldred, C., Taylor, M., & Guba, O. (2022). Thermodynamically consistent versions of approximations used in modelling moist air. *Quart. J. Roy. Meteor. Soc.*. (revising) doi: https://doi.org/10.48550/arXiv.2204.00534

Emanuel, K. (1994). *Atmospheric convection*. Oxford University Press. doi: https://doi.org/10.1002/qj.49712152516

Feistel, R. (2008). A gibbs function for seawater thermodynamics for -6 to 80^oc and salinity up to 120 g kg^{--1}. *Deep-Sea Res I*, *55*(12), 1639 - 1671. doi: https://doi.org/10.1016/j.dsr.2008.07.004

Feistel, R., Wright, D. G., Kretzschmar, H.-J., Hagen, E., Herrmann, S., & Span, R. (2010). Thermodynamic properties of sea air. *Ocean Science*, *6*, 91–141. doi: https://doi.org/10.5194/os-6-91-2010

Feistel, R., Wright, D. G., Miyagawa, K., Harvey, A. H., Hruby, J., Jackett, D. R., & McDougall, T. J. (2008). Mutually consistent thermodynamic potentials for fluid water, ice, and seawater: a new standard for oceanography. *Ocean Sci.*, *4*, 275–291. doi: https://doi.org/10.5194/os-4-275-2008

Ferguson, J., Jablonowski, C., & Johansen, H. (2019). Assessing Adaptive Mesh Refinement (AMR) in a Forced Shallow-Water Model with Moisture. *Mon. Wea. Rev.*, *147*, 3673–3692. doi: https://doi.org/10.1175/MWR-D-18-0392.1

Ferguson, J., Jablonowski, C., Johansen, H., McCorquodale, P., Colella, P., & Ullrich, P. A. (2016). Analyzing the Adaptive Mesh Refinement (AMR) Characteristics of a High-Order 2D Cubed-Sphere Shallow-Water Model. *Mon. Wea. Rev.*, *144*, 4641–4666. doi: https://doi.org/10.1175/MWR-D-16-0197.1

Feynman, R., Leighton, R., & Sands, M. (1989). *Volume 1 of the Feynman Lectures on Physics*. Addison-Wesley. doi: https://doi.org/10.1119/1.1972241

Freitas, S. R., Grell, G. A., Molod, A., Thompson, M. A., Putman, W. M., Santos e Silva, C. M., & Souza, E. P. (2018). Assessing the grell-freitas convection parameterization in the NASA GEOS modeling system. *J. Adv. Model. Earth Syst.*, *10*(6), 1266–1289. doi: 10.1029/2017MS001251

Gassmann, A. (2013). A global hexagonal c-grid non-hydrostatic dynamical core (icon-iap) designed for energetic consistency. *Quart. J. Roy. Meteor. Soc.*, *139*(670), 152–175. doi: 10.1002/qj.1960

Gassmann, A. (2018). Entropy production due to subgrid-scale thermal fluxes with application to breaking gravity waves. *Quart. J. Roy. Meteor. Soc.*, *144*(711), 499-510. doi: 10.1002/qj.3221

Gassmann, A., & Herzog, H.-J. (2008). Towards a consistent numerical compressible

non-hydrostatic model using generalized Hamiltonian tools. *Q. J. R. Meteorol. Soc.*, *134*, 1597-1613. doi: https://doi.org/10.1002/qj.297

Gassmann, A., & Herzog, H.-J. (2015). How is local material entropy production represented in a numerical model? *Quart. J. Roy. Meteor. Soc.*, *141*(688), 854-869. doi: 10.1002/qj.2404

Gawlik, E. S., & Gay-Balmaz, F. (2021). A variational finite element discretization of compressible flow. *Foundations of Computational Mathematics*, *21*(4), 961–1001. doi: https://doi.org/10.1007/s10208-020-09473-w

Gay-Balmaz, F. (2019). A variational derivation of the thermodynamics of a moist atmosphere with rain process and its pseudoincompressible approximation. *Geophysical & Astrophysical Fluid Dynamics*, *113*(5-6), 428-465. doi: 10.1080/03091929.2019.1570505

Germano, M. (1992). Turbulence: the filtering approach. *Journal of Fluid Mechanics*, *238*, 325–336. doi: 10.1017/S0022112092001733

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., ... Randel, W. J. (2019). The Whole Atmosphere Community Climate Model Version 6 (WACCM6). *J. Geophys. Res. Atmos.*, *124*(23), 12380–12403. doi: 10.1029/2019JD030943

Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for global models. part i: Off-line tests and comparison with other schemes. *J. Climate*, *28*(3), 1268 - 1287. doi: 10.1175/JCLI-D-14-00102.1

Gill, A. E. (1982). *Atmosphere-ocean dynamics*. Academic Press, New York.

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., ... Zhu, Q. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *J. Adv. Model. Earth Syst.*, *11*(7), 2089-2129. doi: https://doi.org/10.1029/2018MS001603

Gosnell, R., Fairall, C. W., & Webster, P. J. (1995). The sensible heat of rainfall in the tropical ocean. *J. Geo. Res.: Atmospheres*, *100*(C9), 18437. doi: 10.1029/95JC01833

Griffies, S. M., Yin, J., Durack, P. J., Goddard, P., Bates, S. C., Behrens, E., ... Zhang, X. (2014). An assessment of global and regional sea level for years 1993?2007 in a suite of interannual core-ii simulations. *Ocean Modelling*, *78*, 35-89. doi: https://doi.org/10.1016/j.ocemod.2014.03.004

Gross, M., Wan, H., Rasch, P. J., Caldwell, P. M., Williamson, D. L., Klocke, D., ... Leung, R. (2018). Physics-dynamics coupling in weather, climate and earth system models: Challenges and recent progress. *Mon. Wea. Rev.*, *146*, 3505-3544. doi: 10.1175/MWR-D-17-0345.1

Guo, C., Bentsen, M., Bethke, I., Ilicak, M., Tjiputra, J., Toniazzo, T., ... Otterå, O. H. (2019). Description and evaluation of noresm1-f: a fast version of the norwegian earth system model (noresm). *Geosci. Model Dev.*, *12*(1), 343–362. doi: 10.5194/gmd-12-343-2019

Harris, B. L., & Tailleux, R. (2018). Assessment of algorithms for computing moist available potential energy. *Quart. J. Roy. Meteor. Soc.*, *144*, 1501–1510. doi: https://doi.org/10.1002/qj.3297

Harris, L. M., Chen, X., Putman, W., Zhou, L., & Chen, J.-H. (2021, June). *A Scientific Description of the GFDL Finite-Volume Cubed-Sphere Dynamical Core* (Technical Memorandum No. GFDL2021001). Princeton, New Jersey: Geophysical Fluid Dynamics Laboratory. (109 pp., available from https://github.com/NOAA-GFDL/GFDL_atmos_cubed_sphere/tree/master/docs)

Harrop, B. E., Pritchard, M., Parishani, H., Gettelman, A., Hagos, S., Lauritzen, P. H., ... Sakaguchi, K. (2022). Conservation of dry air, water, and energy in CAM and its potential impact on tropical rainfall. *J. Climate*. doi: 10.1175/JCLI-D-21-0512.1

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J.,

... Zadeh, N. (2019). Structure and performance of GFDL's CM4.0 climate model. *J. Adv. Model. Earth Syst.*, *11*(11), 3691-3727. doi: https://doi.org/10.1029/2019MS001829

Held, I. M., & Suarez, M. J. (1994). A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteorol. Soc.*, *73*, 1825-1830. doi: 10.1175/1520-0477(1994)075⟨1825:APFTIO⟩2.0.CO;2

Hermann, S., Kretzschmar, H.-J., & Gatley, D. P. (2009). Thermodynamic properties of real moist air, dry air, steam, water, and ice (rp-1485). *HVAC&R Research*, *15*, 961–686. doi: https://doi.org/10.1080/10789669.2009.10390874

Herrington, A. R., Lauritzen, P. H., Reed, K. A., Goldhaber, S., & Eaton, B. E. (2019). Exploring a lower-resolution physics grid in CAM-SE-CSLAM. *J. Adv. Model. Earth Syst.*, 1894– 1916. doi: 10.1029/2019MS001684

Herrington, A. R., Lauritzen, P. H., Taylor, M. A., Goldhaber, S., Eaton, B. E., Reed, K. A., & Ullrich, P. A. (2018). Physics-dynamics coupling with element-based high-order Galerkin methods: quasi equal-area physics grid. *Mon. Wea. Rev.*, *134*, 3610-3624. doi: 10.1175/MWR3360.1

Hilt, M., Auclair, F., Benshila, R., Bordois, L., Capet, X., Debreu, L., ... Roblou, L. (2020). Numerical modelling of hydraulic control, solitary waves and primary instabilities in the strait of gibraltar. *Ocean Modell.*, *151*, 101642. doi: https://doi.org/10.1016/j.ocemod.2020.101642

Holm, D. D., Marsden, J. E., & Ratiu, T. S. (1998). The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, *137*(1), 1 - 81. doi: 10.1006/aima.1998.1721

Holm, D. D., Marsden, J. E., & Ratiu, T. S. (2002). The Euler-Poincaré equations in geophysical fluid dynamics. *Large-scale atmosphere-ocean dynamics*, *Vol. II*, 251–300.

Ingersoll, A. P. (2005). Boussinesq and anelastic approximations revisited: potential energy release during thermobaric instability. *J. Phys. Oceanogr.*, *35*, 1359–1369. doi: https://doi.org/10.1175/JPO2756.1

Jablonowski, C., & Williamson, D. L. (2011). The pros and cons of diffusion, filters and fixers in atmospheric general circulation models. In P. H. Lauritzen, C. Jablonowski, M. Taylor, & R. Nair (Eds.), *Numerical techniques for global atmospheric models* (pp. 381–493). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-11640-7_13

Kasahara, A. (1974). Various vertical coordinate systems used for numerical weather prediction. *Mon. Wea. Rev.*, *102*(7), 509-522. doi: https://doi.org/10.1175/1520-0493(1974)102⟨0509:VVCSUF⟩2.0.CO;2

Kato, S., Loeb, N. G., Fasullo, J. T., Trenberth, K. E., Lauritzen, P. H., Rose, F. G., ... Satoh, M. (2021). Regional energy and water budget of a precipitating atmosphere over ocean. *J. Climate*, *34*(11), 4189 - 4205. doi: 10.1175/JCLI-D-20-0175.1

Kato, S., Rose, F. G., Sun-Mack, S., Miller, W. F., Chen, Y., Rutan, D. A., ... Collins, W. D. (2011). Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *J. Geophys. Res. Atmos.*, *116*(D19). doi: https://doi.org/10.1029/2011JD016050

Kato, S., Xu, K.-M., Wong, T., Loeb, N. G., Rose, F. G., Trenberth, K. E., & Thorsen, T. J. (2016). Investigation of the residual in column-integrated atmospheric energy balance using cloud objects. *J. Climate*, *29*(20), 7435 - 7452. doi: 10.1175/JCLI-D-15-0782.1

Kevlahan, N. K.-R., & Dubos, T. (2019). WAVETRISK-1.0: an adaptive wavelet hydrostatic dynamical core. *Geosci. Model Dev.*, *12*, 4901–4921. doi: 10.5194/gmd-12-4901-2019

Kieu, C. (2015). Revisiting dissipative heating in tropical cyclone maximum poten-

tial intensity. *Quart. J. Roy. Meteor. Soc.*, *141*(692), 2497-2504. doi: https://doi.org/10.1002/qj.2534

Kinzer, G. D., & Gunn, R. (1951, apr). The evaporation, temperature and thermal relaxation-time of freely falling waterdrops. *J. Meteorol.*, *8*(2), 71–83. doi: 10.1175/1520-0469(1951)008⟨0071:TETATR⟩2.0.CO;2

Klemp, J. B., Dudhia, J., & Hassiotis, A. D. (2008). An upper gravity-wave absorbing layer for NWP applications. *Mon. Wea. Rev.*, *136*(10), 3987 - 4004. doi: 10.1175/2008MWR2596.1

Landau, L. D., & Lifschitz, E. M. (1968). *Fluid mechanics.* Pergamon Press.

Larson, V. E. (2017). *CLUBB-SILHS: A parameterization of subgrid variability in the atmosphere.* (arXiv preprint arXiv:1711.03675v4)

Lauritzen, P. H., Nair, R., Herrington, A., Callaghan, P., Goldhaber, S., Dennis, J., ... Dubos, T. (2018). NCAR release of CAM-SE in CESM2.0: A reformulation of the spectral-element dynamical core in dry-mass vertical coordinates with comprehensive treatment of condensates and energy. *J. Adv. Model. Earth Syst.*, *10*(7), 1537-1570. doi: 10.1029/2017MS001257

Lauritzen, P. H., Taylor, M. A., Overfelt, J., Ullrich, P. A., Nair, R. D., Goldhaber, S., & Kelly, R. (2017). CAM-SE-CSLAM: Consistent coupling of a conservative semi-lagrangian finite-volume method with spectral element dynamics. *Mon. Wea. Rev.*, *145*(3), 833-855. doi: 10.1175/MWR-D-16-0258.1

Lauritzen, P. H., & Williamson, D. L. (2019). A total energy error analysis of dynamical cores and physics-dynamics coupling in the Community Atmosphere Model (CAM). *J. Adv. Model. Earth Syst.*, *11*(5), 1309-1328. doi: 10.1029/2018MS001549

L'Ecuyer, T. S., Beaudoing, H. K., Rodell, M., Olson, W., Lin, B., Kato, S., ... and K. Hilburn, D. P. L. (2015). The observed state of the energy budget in the early twenty-first century. *J. Climate*, *28*(21), 8319 - 8346. doi: 10.1175/JCLI-D-14-00556.1

Lesieur, M., & Metais, O. (1996). New trends in large-eddy simulations of turbulence. *ANNUAL REVIEW OF FLUID MECHANICS*, *28*, 45-82. doi: 10.1146/annurev.fl.28.010196.000401

Lin, S.-J. (2004). A 'vertically Lagrangian' finite-volume dynamical core for global models. *Mon. Wea. Rev.*, *132*, 2293-2307. doi: https://doi.org/10.1175/1520-0493(2004)132⟨2293:AVLFDC⟩2.0.CO;2

Liu, H.-L., Lauritzen, P. H., Vitt, F., & Goldhaber, S. (2022). Thermospheric and ionospheric effects by gravity waves from the lower atmosphere. *Earth and Space Science Open Archive*, 39. doi: 10.1002/essoar.10511744.1

Loeb, N. G., Rutan, D. A., Kato, S., & Wang, W. (2014). Observing interannual variations in hadley circulation atmospheric diabatic heating and circulation strength. *J. Climate*, *27*(11), 4139 - 4158. doi: 10.1175/JCLI-D-13-00656.1

Lorenz, E. (1955). Available potential energy and the maintenance of the general circulation. *Tellus*, *7*, 138–157. doi: https://doi.org/10.1111/j.2153-3490.1955.tb01148.x

Lucarini, V., & Ragone, F. (2011). Energetics of climate models: Net energy balance and meridional enthalpy transport. *Reviews of Geophysics*, *49*(1). doi: https://doi.org/10.1029/2009RG000323

Lumley, J., & Panofsky, H. (1964). *The structure of atmospheric turbulence.* New York, Interscience Publishers. doi: https://doi.org/10.1002/qj.49709138926

Madec, G., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., ... Samson, G. (2019, October). *Nemo ocean engine.* Zenodo. doi: 10.5281/zenodo.3878122

Mayer, M., Haimberger, L., Edwards, J. M., & Hyder, P. (2017). Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. *J. Climate*, *30*(22), 9225-9246. doi: 10.1175/JCLI-D-17-0137.1

Melvin, T., Benacchio, T., Shipway, B., Wood, N., Thuburn, J., & Cotter, C. (2019).

A mixed finite-element, finite-volume, semi-implicit discretization for atmospheric dynamics: Cartesian geometry. *Quart. J. Roy. Meteor. Soc.*, *145*(724), 2835–2853. doi: https://doi.org/10.1002/qj.3501

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., ... Taylor, M. A. (2010, June). *Description of the NCAR Community Atmosphere Model (CAM 5.0)* (NCAR Technical Note Nos. NCAR/TN-486+STR). Boulder, Colorado: National Center of Atmospheric Research. (268 pp., available from http://www.cesm.ucar.edu/models/cesm1.0/cam/)

Novak, L., & Tailleux, R. (2018). On the local view of atmospheric available potential energy. *J. Atmos. Sci.*, 1891–1907. doi: 10.1175/JAS-D-17-0330.1

Park, S. (2014). A unified convection scheme (UNICON). part II: Simulation. *J. Atmos. Sci.*, *71*(11), 3931–3973. doi: 10.1175/JAS-D-13-0234.1

Pauluis, O. (2008). Thermodynamic consistency of the anelastic approximation for a moist atmosphere. *J. Atmos. Sci.*, *65*, 2719–2729. doi: https://doi.org/10.1175/2007JAS2475.1

Pauluis, O., Balaji, V., & Held, I. M. (2000). Frictional dissipation in a precipitating atmosphere. *J. Atmos. Sci.*, *57*(7), 989-994. doi: 10.1175/1520-0469(2000)057⟨0989:FDIAPA⟩2.0.CO;2

Pauluis, O., & Dias, J. (2012). Satellite estimates of precipitation-induced dissipation in the atmosphere. *Science*, *335*, 953–956. doi: 10.1126/science.1215869

Peixoto, J. P., & Oort, M. (1992). *Physics of climate*. American Institute of Physics. doi: https://doi.org/10.1063/1.2809772

Rasch, P. J., Xie, S., Ma, P.-L., Lin, W., Wang, H., Tang, Q., ... Yang, Y. (2019). An Overview of the Atmospheric Component of the Energy Exascale Earth System Model. *J. Adv. Model. Earth Syst.*, *11*(8), 2377–2411. doi: 10.1029/2019MS001629

Renault, L., Molemaker, M. J., McWilliams, J. C., Shchepetkin, A. F., Lemarié, F., Chelton, D., ... Hall, A. (2016). Modulation of wind work by oceanic current interaction with the atmosphere. *J. Phys. Oceanogr.*, *46*(6), 1685 - 1704. doi: 10.1175/JPO-D-15-0232.1

Salmon, R. (1998). *Lectures on geophysical fluid dynamics*. Oxford University Press.

Schaefer-Rolffs, U., & Becker, E. (2013). Horizontal momentum diffusion in GCMs using the dynamic Smagorinsky model. *Mon. Wea. Rev.*, *141*(3), 887 - 899. doi: 10.1175/MWR-D-12-00101.1

Schaefer-Rolffs, U., & Becker, E. (2018). Scale-invariant formulation of momentum diffusion for high-resolution atmospheric circulation models. *Mon. Wea. Rev.*, *146*(4), 1045 - 1062. doi: 10.1175/MWR-D-17-0216.1

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L., Park, S.-H., & Ringler, T. D. (2012). A multi-scale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering. *Mon. Wea. Rev.*, *240*, 3090-3105. doi: 10.1175/MWR-D-11-00215.1

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Wea. Rev.*, *91*(3), 99 - 164. doi: 10.1175/1520-0493(1963)091⟨0099:GCEWTP⟩2.3.CO;2

Staniforth, A., White, A., Wood, N., Thuburn, J., Zerroukat, M., Cordero, E., ... Diamantakis, . M. (2006). Joy of U.M 6.3 - model formulation. *UK Met Office Technical Note*, *15*. Retrieved from `http://research.metoffice.gov.uk/research/nwp/publications/papers/unified_model/umdp15_v6.3.pdf`

Tailleux, R. (2012). Thermodynamics/dynamics coupling in weakly compressible turbulent stratified fluids. *ISRN Thermodynamics*, *2012*, 1-15. doi: https://doi.org/10.5402/2012/609701

Tailleux, R. (2018). Local available energetics of multicomponent compressible stratified fluids. *J. Fluid Mech*, *842*(R1). doi: 10.1017/jfm.2018.196

Taylor, M. A. (2011). Conservation of mass and energy for the moist atmospheric primitive equations on unstructured grids, in: P.H. Lauritzen, R.D. Nair, C.

Jablonowski, M. Taylor (Eds.), Numerical techniques for global atmospheric models. *Lecture Notes in Computational Science and Engineering, Springer, 2010, in press.*, *80*, 357-380. doi: 10.1007/978-3-642-11640-7_12

Taylor, M. A., & Fournier, A. (2010). A compatible and conservative spectral element method on unstructured grids. *J. Comput. Phys.*, *229*(17), 5879 - 5895. doi: https://doi.org/10.1016/j.jcp.2010.04.008

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldred, C. (2020). An energy consistent discretization of the nonhydrostatic equations in primitive variables. *J. Adv. Model. Earth Syst.*, *12*(1), e2019MS001783. doi: https://doi.org/10.1029/2019MS001783

Thuburn, J. (2017). Use of the Gibbs thermodynamic potential to express the equation of state in atmospheric models. *Quart. J. Roy. Meteor. Soc.*, *143*(704), 1185-1196. doi: https://doi.org/10.1002/qj.3020

Thuburn, J. (2018). Corrigendum: Use of the Gibbs thermodynamic potential to express the equation of state in atmospheric models. *Quart. J. Roy. Meteor. Soc.*, *144*(711), 632-632. doi: https://doi.org/10.1002/qj.3232

Tort, M., & Dubos, T. (2014). Usual approximations to the equations of atmospheric motion: A variational perspective. *J. Atmos. Sci.*, *71*(7), 2452-2466. doi: 10.1175/JAS-D-13-0339.1

Trenberth, K. E. (1997). Using atmospheric budgets as a constraint on surface fluxes. *J. Climate*, *10*(11), 2796 - 2809. doi: 10.1175/1520-0442(1997)010⟨2796:UABAAC⟩2.0.CO;2

Trenberth, K. E., & Fasullo, J. T. (2018). Applications of an updated atmospheric energetics formulation. *J. Climate*, *31*(16), 6263-6279. doi: 10.1175/JCLI-D-17-0838.1

Tripoli, G., & Cotton, W. (1981). The use of ice-liquid water potential temperature as a thermodynamic variable in deep atmospheric models. *Mon. Wea. Rev.*, *109*, 1094–1102. doi: https://doi.org/10.1175/1520-0493(1981)109⟨1094:TUOLLW⟩2.0.CO;2

Vallis, G. K. (2006). *Atmospheric and oceanic fluid dynamics.* Cambridge University Press, Cambridge. doi: https://doi.org/10.1017/9781107588417

Vasil, G. M., Lecoanet, D., Brown, B. P., Wood, T. S., & Zweibel, E. G. (2013). Energy conservation and gravity waves in sound-proof treatments of stellar interios. II. Lagrangian constrained analysis. *The Astrophysical Journal*, *773*(169). doi: http://dx.doi.org/10.1088/0004-637X/773/2/169

Wacker, U., Frisius, T., & Herbert, F. (2006, 10). Evaporation and Precipitation Surface Effects in Local Mass Continuity Laws of Moist Air. *J. Atmos. Sci.*, *63*(10), 2642-2652. doi: 10.1175/JAS3754.1

Wacker, U., & Herbert, F. (2003). Continuity equations as expressions for local balances of masses in cloudy air. *Tellus A*, *55*(3), 247-254. doi: https://doi.org/10.1034/j.1600-0870.2003.00019.x

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., ... others (2017). The Met Office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geosci. Model Dev.*, *10*(4), 1487–1520. doi: https://doi.org/10.5194/gmd-10-1487-2017

Wei, N., Dai, Y., Zhang, M., Zhou, L., Ji, D., Zhu, S., & Wang, L. (2014). Impact of precipitation-induced sensible heat on the simulation of land-surface air temperature. *J. Adv. Model. Earth Syst.*, *6*(4), 1311-1320. doi: https://doi.org/10.1002/2014MS000322

Williamson, D. L., Olson, J. G., Hannay, C., Toniazzo, T., Taylor, M., & Yudin, V. (2015). Energy considerations in the Community Atmosphere Model (CAM). *J. Adv. Model. Earth Syst.*, *7*(3), 1178–1188. doi: 10.1002/2015MS000448

Williamson, D. L., Olson, J. G., & Jablonowski, C. (2009). Two dynamical core formulation flaws exposed by a baroclinic instability test case. *Mon. Wea. Rev.*, *137*(2), 790 - 796. doi: 10.1175/2008MWR2587.1

Wimmer, G. A., Cotter, C. J., & Bauer, W. (2020). Energy conserving upwinded compatible finite element schemes for the rotating shallow water equations. *J. Comput. Phys.*, *401*, 109016. doi: https://doi.org/10.1016/j.jcp.2019.109016

Young, W. R. (2010, 02). Dynamic Enthalpy, Conservative Temperature, and the Seawater Boussinesq Approximation. *J. Phys. Oceanogr.*, *40*(2), 394-400. doi: 10.1175/2009JPO4294.1

Young, W. R. (2012). An exact thickness-weighted average formulation of the boussinesq equations. *J. Phys. Oceanogr.*, 692–707. doi: https://doi.org/10.1175/JPO-D-11-0102.1

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quart. J. Roy. Meteor. Soc.*, *141*(687), 563-579. doi: https://doi.org/10.1002/qj.2378

Zdunkowski, W., & Bott, A. (2004). *Thermodynamics of the atmosphere – A course in theoretical meteorology.* Cambridge: Cambridge University Press.

Zhang, K., Rasch, P. J., Taylor, M. A., Wan, H., Leung, R., Ma, P.-L., ... Xie, S. (2018). Impact of numerical choices on water conservation in the E3SM atmosphere model version 1 (EAMv1). *Geosci. Model Dev.*, *11*(5), 1971–1988. doi: 10.5194/gmd-11-1971-2018