Validation of ERA5 Cloud Liquid Water Content Against Idealized Moist Adiabatic Process

Sun-Hsing Ho

December 12, 2025

Abstract

This project evaluates the representation of cloud liquid water content (clwc) in the ERA5 reanalysis monthly mean dataset by comparing it with a theoretical baseline calculated from idealized moist adiabatic process. The investigation covers a 20-year period (2005–2024) over a domain spanning East Asia and the Western Pacific (105°E–145°E, 0°–40°N). We assume mass conservation of total water following a parcel originating from 1000 hPa, neglecting precipitation processes. The results indicate a systematic negative bias in the theoretical model compared to ERA5. This bias is fundamentally attributed to the theoretical assumption of a reversible moist adiabatic process, whereas the reanalysis accounts for pseudo-adiabatic processes where precipitation removes water. A distinct seasonal cycle is observed, with larger deviations (more negative differences) occurring in summer due to efficient water removal in deep convection, while in winter, stratus exhibits smaller deviations.

1 Introduction

Clouds play a pivotal role in the Earth's radiation budget and hydrological cycle. Reanalysis datasets, such as ERA5, rely on complex parameterization schemes to estimate cloud properties, including cloud liquid water content (clwc). Validating these estimates is challenging due to the scarcity of direct in-situ observations.

This study proposes a thermodynamic approach to validation by comparing ERA5 output with a theoretical upper bound of liquid water content calculated using a moist adiabatic parcel method. The primary assumption is that the air parcel ascends adiabatically from the surface (1000 hPa) while conserving total water mass (i.e., no precipitation occurs). This adiabatic liquid water mixing ratio serves as a baseline to understand the existanace of "modeled clouds" in ERA5 model.

2 Data and Methodology

2.1 Data and Domain

The study utilizes monthly mean data from the ERA5 Reanalysis. The spatial domain is defined as 105°E to 145°E and 0° to 40°N, covering the subtropical Western Pacific, the South China Sea, and East Asia. The temporal domain spans 20 years, from January 2005 to December 2024.

2.2 Thermodynamic Calculations

The core methodology involves lifting an air parcel from the 1000 hPa pressure level. We first determine if the parcel reaches saturation at a target pressure level by comparing the target pressure with the lifting condensation level (LCL).

The temperature at the LCL (T_{LCL}) is approximated using the empirical formula mentioned in Barnes (1968):

$$T_{LCL} = T_d - (0.001296 \cdot Td + 0.1963) \cdot (T - T_d) \tag{1}$$

where T and T_d are the temperature and dew point temperature (both in Celsius) at the reference level, respectively.

For parcels that reach saturation (i.e., $P_{level} \leq P_{LCL}$), we calculate the theoretical liquid water mixing ratio (χ). The calculation is governed by the conservation of total water (Q) and the moist adiabatic relation. The conservation equation used is:

$$\frac{T}{p^{R_d/C}} \cdot \exp\left(\frac{L_v \cdot w_s}{C \cdot T}\right) = \text{constant} \tag{2}$$

where C is the effective specific heat capacity defined as:

$$C = C_p + w \cdot c_w \tag{3}$$

Here, C_p is the specific heat of dry air, c_w is the specific heat of liquid water, L_v is the latent heat of vaporization, R_d is the gas constant for dry air, and w_s is the saturation mixing ratio. The theoretical liquid water content is then derived as the difference between the total water (Q) and the saturation mixing ratio at the target level.

3 Results and Discussion

3.1 Source of Systematic Bias

The comparison between the calculated liquid water mixing ratio and the ERA5 CLWC reveals a widespread negative difference (ERA5 $-$ Calculated < 0). This systematic bias stems directly from the fundamental difference between the theoretical assumption and the physical reality modeled by ERA5.

Our calculation assumes a **reversible moist adiabatic process**, implying strictly conserved total water mass ($Q = \text{const}$). In this idealized scenario, all condensed water is retained within the parcel. However, the real atmosphere and the ERA5 model follow a **pseudo-moist adiabatic process**, where condensed water is continuously removed from the parcel via precipitation. Consequently, our theoretical value represents a "no-precipitation" upper bound, resulting in the calculated χ being consistently larger than the ERA5 CLWC, hence the negative difference.

Seasonal Variability and Physical Interpretation:

- **Winter:** The difference is smaller (i.e., less negative, closer to zero), indicating a smaller deviation between the theoretical model and ERA5. Winter conditions in the domain are characterized by weaker convection and a prevalence of stratus. These clouds produce less intense precipitation compared to deep convection. As a result, water removal is less efficient, and the assumption of mass conservation is less violated. The ERA5 clwc is therefore closer to the adiabatic total water, resulting in a smaller bias.

- **Summer:** The difference is larger (i.e., more negative), indicating a larger deviation between the theoretical model and ERA5. Summer has more cumulus which is dominated by the deep convection. These systems are highly efficient at converting cloud water into precipitation. The removal of mass is significant, causing the actual clwc in ERA5 to be far lower than the theoretical conserved value. This leads to a much larger deviation in summer.

Spatial Trend: Consistent with the temporal analysis, both the animation and time-series plots demonstrate that regions with weaker convection (or less frequent deep convection), such as higher latitudes or subsidence zones, exhibit smaller deviations (difference values closer to zero). Conversely, areas of intense convection show the largest negative differences, confirming that precipitation efficiency is the primary driver of the observed divergence.

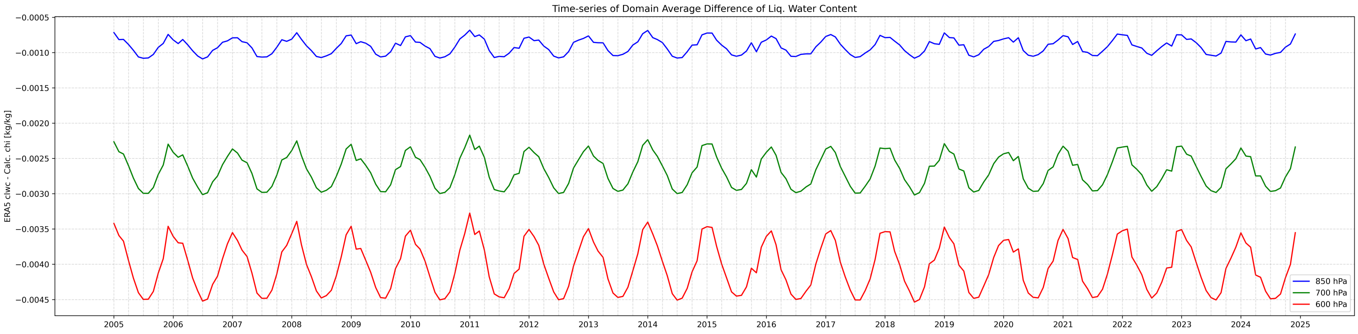

Figure 1: Fig1. The difference of calculated & ERA5 clwc

4 Literature Review and Theoretical Context

The discrepancy between idealized parcel calculations and model output also touches upon the fundamental challenges of representing convection in General Circulation Models (GCMs), particularly the issue of "modeled clouds" or "gridpoint storms."

Scinocca and McFarlane (2004) describe how "modeled clouds" often manifest as unnatural grid-scale precipitation events. These artifacts arise from a "spurious competition" between the model's Large-Scale Precipitation (LSP) scheme and its convective parameterization. When the convective parameterization is too weak to consume the available instability, the grid column becomes supersaturated, triggering the LSP scheme to condense moisture rapidly, behaving like a "Moist-Convective Adjustment" mechanism. This results in intense, unrealistic vertical motions and variance.

The root of this problem lies in the "artificial separation of scales," as discussed by Arakawa (2004). Conventional parameterizations assume a spectral gap between convective plumes and large-scale flow (fractional area $\sigma \ll 1$). However, as model resolutions increase into the "gray zone," this assumption fails, leading to a deadlock where the model struggles to choose between parameterized and explicit cloud representation.

To address this, Arakawa and Wu (2013) proposed the "Unified Parameterization." This framework explicitly treats the fractional convective cloudiness (σ) as a variable. By formulating eddy transport as a function of $\sigma(1-\sigma)$, the parameterization ensures a smooth transition:

$$\overline{w'h'} \propto \sigma(1-\sigma) \tag{4}$$

When the grid box is fully filled with clouds ($\sigma \to 1$), the parameterized eddy transport vanishes, allowing the explicit grid-scale dynamics to handle the transport naturally. Wu and Arakawa (2014) extended this to momentum and microphysics, ensuring that processes like condensation and evaporation are correctly partitioned between the updraft and the environment.

This theoretical context highlights that the "Cloud Liquid Water Content" in ERA5 is a product of complex interactions between these parameterized schemes. The systematic differences observed in our study likely reflect not just the neglect of precipitation in our simple model, but also the sophisticated (and sometimes artificial) adjustment mechanisms the reanalysis model employs to resolve cloudiness across different scales and stability regimes.

5 Conclusion

This project validates ERA5 clwc against a moist adiabatic benchmark. We confirm that the systematic negative bias in the difference (ERA5 − Calculated) arises from the theoretical assumption of a reversible moist adiabatic process (mass conservation), which neglects the water removal inherent in the pseudo-adiabatic processes (precipitation) captured by ERA5.

The analysis further reveals that this bias is modulated by convective intensity. In winter, dominated by stratus and weaker convection, precipitation removal is less efficient, leading to a smaller deviation from the conserved state. In contrast, summer convective regimes drive efficient precipitation, causing a significant departure from mass conservation and a larger deviation. Spatially, regions with weaker convection consistently show smaller deviations, reinforcing the conclusion that precipitation efficiency is the governing factor in the discrepancy between the idealized parcel theory and reanalysis data.

References

[1] Arakawa, A., 2004: The Cumulus Parameterization Problem: Past, Present, and Future. *J. Climate*, **17**, 2493–2525.
`https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2`

[2] Arakawa, A., and C. Wu, 2013: A Unified Representation of Deep Moist Convection in Numerical Modeling of the Atmosphere. Part I. *J. Atmos. Sci.*, **70**, 1977–1992.
`https://doi.org/10.1175/JAS-D-12-0330.1`

[3] Barnes, S. L., 1968: An Empirical Shortcut to the Calculation of Temperature and Pressure at the Lifted Condensation Level. *J. Appl. Meteor. Climatol.*, **7**, 511.
`https://doi.org/10.1175/1520-0450(1968)007<0511:AESTTC>2.0.CO;2`

[4] Scinocca, J. F., and N. A. McFarlane, 2004: The Variability of Modeled Tropical Precipitation. *J. Atmos. Sci.*, **61**, 1993–2015.
`https://doi.org/10.1175/1520-0469(2004)061<1993:TVOMTP>2.0.CO;2`

[5] Wu, C., and A. Arakawa, 2014: A Unified Representation of Deep Moist Convection in Numerical Modeling of the Atmosphere. Part II. *J. Atmos. Sci.*, **71**, 2089–2103.
`https://doi.org/10.1175/JAS-D-13-0382.1`