

The manuscript presents a clear, innovative, and highly relevant approach to addressing critical challenges in projecting extreme rainfall in tropical regions with complex topography. The integration of machine learning Variational Autoencode (VAE) with high-resolution large-eddy simulations (TaiwanVVM) to develop a process-based downscaling framework constitutes a valuable methodological contribution. The main argument of this study disentangling thermodynamic forcing from dynamics and demonstrating that upstream wind dynamics serve as the primary control of heavy rainfall types. It is supported by detailed multi-scale analyses.

Overall, the study demonstrates strong potential and offers important insights for the downscaling of topographic heavy rainfall. However, several issues require clarification and revision to enhance the clarity, reproducibility, and interpretability of the manuscript. My specific comments are as follows:

1. The objective of this study is to propose a downscaling method for Topographic Heavy Rainfalls over the Asian–Australian Monsoon Region by leveraging multiscale moisture dynamical control. However, the manuscript provides little information regarding the resulting downscaled resolution and its accuracy relative to topographic rainfall. This omission makes it difficult for readers to assess how the proposed method performs compared with existing GCM-based downscaling approaches.
2. Regarding Figure S1, it is unclear why the spatial differences between IMERG and CWA appear so large. Previous studies using gridded CWB datasets show more consistent spatial patterns between IMERG and CWB.
Reference: <https://www.sciencedirect.com/science/article/pii/S0169809518304666>
3. The introduction should elaborate on why Integrated Vapor Transport (IVT) is important in the downscaling process.
4. Although the VAE-based method shows improved correlation and reduced RMSE, the differences compared with EOF appear relatively modest (Figure S5). Have you considered comparing the clustering results between VAE and EOF?
5. The organization of figures in both the main text and the supplementary material is at times confusing. Several figures are referenced repeatedly across the Methods and Results sections, making it difficult for readers to follow—particularly in Section 2.2 (Methods) and Section 3.1 (Results). In addition, referring to certain figures in both the Results/Discussion and the Methods (e.g., line 141) does not enhance clarity and forces readers to jump back and forth.