This manuscript presents **MESMER-RCM**, a modular and computationally efficient probabilistic emulator for annual mean 2 m temperature at regional (EURO-CORDEX) scales. The approach is conceptually simple and interpretable: a Lasso-based deterministic mapping from nearby GCM grid points to each RCM grid point, combined with a data-driven, two-stage shrinkage estimator for the RCM residual covariance that enables sampling of spatially correlated internal variability. I think the illustrations in this manuscript is valuable, but there are still some issues need to be addressed before acceptance.

1. The authors need to clarify the MESMER method in detail, maybe in the supporting information. A schematic diagram is necessary.
2. The authors need to describe why you choose MESMER framework and its advantages compared with other generative AI models.
3. How to divide the training samples and testing samples in MESMER-RCM? And how many years for training and testing? In addition, please mark the time span of every dataset in Table S1.
4. Why the residual variability belongs to a multivariate Gaussian distribution? Is there any risk for underestimations under some extreme events? Please clarify this issue.
5. I think there should be another benchmark method to compare for further confirm the advantage of MESMER-RCM.
6. Could you provide more metrics on result evaluation? Such as CRPS.
7. Whether the residual covariance matrixes for different time ranges (e.g., 1979–2000, 2001–2050, 2051–2099) are various? If so, whether such variations are significant? Please clarify its effectiveness.