

Response to Reviewer-3 Comments

1. Summary

Thank you very much for taking the time to review this manuscript. Please find the detailed responses below and the corresponding revisions/corrections highlighted/in track changes in the re-submitted files.

2. Point-by-point response to Comments and Suggestions for Authors

Comments 1: It would be valuable to consider incorporating statistical significance testing (e.g., t-tests, ANOVA) to support the performance comparisons between models. For instance, the F1CPA difference between ChaMTeC and TimeMixer on the WaterLog* dataset is relatively modest, and readers may benefit from understanding whether such differences are statistically meaningful. Including significance indicators in Tables 2–5 could enhance the robustness of the evaluation.

Response 1: Thank you for pointing this out. We agree with this comment. Therefore, we have explain this part as:

First of all, in line with your suggestion, one-way ANOVA test was applied to evaluate the performance differences of all models. Since the p-values obtained from ANOVA and t-tests were greater than 0.05, it was determined that there was no statistically significant difference between the models. This situation shows that different models generally offer similar F1CPA performances. Therefore, we think that these tests will yield more meaningful results on a newly obtained WaterLog-like dataset where different cyber scenarios are applied, not for this study. We think that this valuable idea of yours will contribute to our new study.

Comments 2: A brief discussion of which types of anomalies are most challenging for the model, and why, would enrich the analysis and offer useful insights into the framework’s limitations and areas for future improvement.

Response 2: Agree.

In line with your recommendation, relevant additions have been made to the Conclusion section.

Comments 3: Finally, elaborating on the individual contributions of each model component (e.g., the role of inverted embedding or the MSE-based feedback loop) through ablation results or discussion would help clarify the design choices and their impact on performance.

Response 3: Agree.

In line with your recommendation, relevant additions have been made to the Conclusion section.