

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

Manuscript Number: egosphere-2025-3540

Renjie Zhou and Shiqi Liu. “A hybrid Kolmogorov-Arnold networks-based model with attention for predicting Arctic River streamflow”

Dear Professor Rohini Kumar,

Thank you and the anonymous reviewers for the positive comments and insightful feedback. We have carefully revised our manuscript egosphere-2025-3540 entitled “A hybrid Kolmogorov-Arnold networks-based model with attention for predicting Arctic River streamflow” after considering the comments made by you and the reviewers. New baseline comparison models, evaluation metrics and analysis have been added to the manuscript. The code and representative example datasets have been made openly accessible.

The followings are point-by-point responses to the comments.

Reviewer 1:

Zhou and Liu present a novel approach for a data-driven model for discharge modelling. It is based on a Kolmogorov-Arnold network combined with a Long-Short Term Memory (LSTM) model, an attention mechanism that includes a trigonometric depiction of seasonal patterns, as well as a physics-based constrain. The newly developed model aimed at improving the prediction of discharge within arctic areas with their special characteristics like perma frost and accumulation and melting of snow over longer periods. Therefore, the model was applied to the discharge data of the Kolyma River in Siberia and the prediction evaluated against the predictions of several other simpler models.

I have found the presented modelling approach to be a novel and valuable contribution to the hydrological modelling community. I believe it to be fitting for the scope of the

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

Journal. However, the presented manuscript needs work regarding the methodology section as well as the discussion.

Reply: We are grateful for the reviewer's positive feedback and constructive suggestions. We have thoroughly revised the manuscript, corrected errors, added references, addressed each comment, and provided the necessary clarifications as outlined below.

Major comments:

1. Line 30: I can't really support the statement that the presented framework is (better) suited for predicting Arctic River discharge under changing climate conditions. It is well likely that climate change impacts the respective catchments in a way that the general behaviour changes - which also alters how discharge forms. I then get to a model space where the model has to extrapolate - which data-driven models are unsuited for.

Reply: Implemented. We thank the reviewer for raising this point. We have revised the statement in the manuscript and added a short paragraph about its limitations. See Lines 27-29 in the abstract and 627-634 in the conclusion.

2. Line 137, Figure 1: I personally don't think the figure to be well chosen, as the important aspects are missing. I would rather use a fogure that shows the catchment itself with its topography.

Reply: Implemented. We thank the reviewer for this advice. The figure has been updated. See Figure 1 and Line 142.

3. Line 138-143: These lines are unnecessary here and probably can be deleted. All those things have already been said within the introduction and are explained over the methodology section anyways.

Reply: Implemented. These lines have been deleted and revised to increase the

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

flow. See Line 143.

4. Line 144-145: All steps, that are necessary for actual model runs should come after the model description. Otherwise, the order is confusing.

Reply: Implemented. It is reorganized to improve readability and clarity. See Lines 340-345 and 374-387.

5. Line 146-164: The description of the whole model structure should be done after the individual parts are explained. Figure 2 also should be moved there.

Reply: Implemented. We have reorganize the structure to introduce the individual parts before the whole structure. See Lines 346-387.

6. I do recommend the inclusion of an additional efficiency measure like KGE, that is complementary to the other ones and also incorporates different aspects of the discharge like bias for example. Please also cite and mention, which version of the KGE you use then.

Reply: Implemented. We have added KGE' (2012) as an additional evaluation metrics. Pictures, references and discussion are revised and updated accordingly. See Lines 319-340, 406-416, 436-438, 452-453, 575-580, 596-597 and 605-609.

7. Why does the methodology end here? Important parts that come up later within the results part are missing. The methodology should explain that the final model is compared to certain baseline models and how they distinguish from the new model presented here. Furthermore, the whole part is missing about how the model is trained on the data, with how many runs, ending criterion, hyper parameters and so on.

Reply: Implemented. A subsection of model implementation and training has been added here to introduce the model hyperparamters. See Lines 340-387,

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

Section 3.7.

8. Line 323-327: This is methodology and should not be within the results part - as it is missing within the methods section.

Reply: Implemented. This part has been removed from the results section. See Lines 363-365.

9. Line 328-329: As mentioned earlier, the baseline models cannot be newly introduced within the results.

Reply: Impelemented. The baseline models are introduced in the methods section. See Lines 365-366.

10. Line 343-344: You can't conduct boxplots. Do you mean you conducted the model application 10 times?

Reply: Implemented. We have rephrased the language to improve its clarity and readability. See Lines 366-370 and 426-428.

11. Line 357: Figure 6 y axis seems to be cut, the numbers are partly missing

Reply: Implemented. We thank the reviewer for pointing this out. The figure has been fixed. See Line 453 and Figure 6.

12. Line 361: I dont see how this represents the "spectrum of hydrological variability". From my understanding, it is more of a possibility to see, how the model performs if the data is only available in lesser resolution. How does this assess the depiction of the hydrological variability?

Reply: Implemented. We thank the reviewer for this important clarification. The reviewer is correct that our analysis examines model performance under varying flow conditions, from low to high discharge events. The corresponding description is rephrased for clarification. See Lines 455-457.

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

13. Line 405: Figure 8, are these for a aggregation period of 1 month?

Reply: We thank the reviewer for requesting this clarification. The box plots in Figure 8 show results aggregated across all forecasting time steps (1-12 months). See Lines 527-528 and 558-559 for clarification.

14. Line 407-415: This is all methodology and not results.

Reply: Implemented. We thank the reviewer for identifying this issue. The contents have been reorganized. See Lines 559-562.

15. Line 437-448: I dont think this part is really necessary here. The conclusion is not a whole summary of the paper, but points out the key findings again.

Reply: Impelemenened. This long paragraph is removed. See Lines 597-601.

16. Line 455-456: The river discharge has a long memory? The sentence does not make sense. I feel like there is a more thorough discussion necessary of why the model shows this behaviour regarding the model efficiency for different aggregation periods - where the reason must be within model structure and how it fits the discharge pattern over time.

Reply: Implemented. The sencence has been rephrased to avoid confusion. Also, a more thorough discussion has been added. See Lines 417-425.

17. I generally feel like the discussion part is lacking depth. While I personally recommend to separate results and discussion, you can keep both together if it makes sense overall. But in the current state, the results lack depth regarding the explanation of observed model behaviour. For example, line 462-463: has this been the same for the application of other models? Is this a common problem? Like this, a few more citations and comparisons to other studies would help putting the paper within a broader context.

Reply: Implemeneted. We have enhanced the results and discussion. In particular,

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

we added a more detailed explanation addressing the specific sentence raised by the reviewer. See Lines 471-480 and 581-595.

18. Also, I am currently missing a graphical depiction of the gauging curve and the simulated discharge. I believe a figure for that would help to give the reader an idea of how the model behaves, where it might deviate from gauging data and where it is strongly in congruence with it.

Reply: Implemented. A new graphic depiction of observed and simulated discharge has been added. See Lines 483-484 and Figure 8.

Minor comments:

19. Line 22: structure

Reply: Implemented. See Line 22.

20. Line 24: dominated by permafrost

Reply: Implemented. See Line 24.

21. Line 27: ...that these components improve the predictive performance.

Reply: Implemented. See Line 27.

22. Line 46: These temperature dependent transitions...?

Reply: Implemented. See Line 45.

23. Line 128-129: Why is there no citation for the Dataset?

Reply: Implemented. The data source and citation have been added to the manuscript. See Lines 133 and 139.

24. Line 178: 1) Input expansion

Reply: Implemented. See Line 156.

25. Line 183-185: Kolmogorov-Arnold theorem while avoiding the computational overhead

Reply: Implemented. See Line 163.

26. Line 195: GELU

Reply: Implemented. See Line 173.

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

27. Line 196: Figure 3 not referenced within the text.

Reply: Implemented. See Line 149.

28. Line 200: ...mechanism and a hidden state, an LSTM can efficiently regulate...

Reply: Implemented. See Line 178.

29. Line 209: The memory cell of an LSTM is primarily composed...

Reply: Implemented. See Line 187.

30. Line 240: "Q refers the discharge prediction using the context vector calculated from the context vector." It has to be "refers to" and what is "using the context vector calculated from the context vector" supposed to mean?

Reply: Implemented. It has been rephrased to improve clarity. See Line 218.

31. Line 273: I recommend a semicolon after water.

Reply: Implemented. See Line 273.

32. Line 279: caused by sources, such as model simplifications...

Reply: Implemented. See Line 279.

33. Line 285-286: Maybe its better to reformulate the sentence and describe alpha and beta as parameters that have to be fitted through model application?

Reply: Implemented. See Lines 285-288 and 361-363 in the manuscript and S1 in the Supplimentary Material.

34. Line 299: beneficial

Reply: Implemented. See Line 301.

35. Line 303-304: What is cited here? The Nash-Sutcliffe efficiency measure should be properly cited.

Reply: Implemented. See Line 308.

36. Line 330: I would recommend to implement the name RCPIKLA of the new model earlier, instead of within the results.

Reply: Implemented. See Line 106.

37. Line 396: change "better captures"

Reply: Implemented. See Line 385 and 482.

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

Reviewer 2:

1. One of the primary advantages of Kolmogorov-Arnold Networks is their enhanced interpretability compared to traditional MLPs. KAN is usually used to improve the interpretability of the relations between inputs and output, but there is no mention of that.

The manuscript fails to leverage or discuss this fundamental strength of KAN architecture. Specifically, there is no:

- Visualization of the learned univariate functions
- Symbolic regression analysis
- Interpretation of what relationships the KAN component discovered between hydrometeorological inputs and Arctic discharge
- Physical insights into the processes governing snowmelt-driven streamflow in permafrost regions

Include a dedicated subsection on KAN interpretability analysis containing:

- Visualization of learned activation functions for key input-output relationships
- Symbolic approximations of these functions where feasible (using symbolic regression tools available in KAN libraries)
- Physical interpretation of discovered patterns in the context of Arctic hydrology
- Comparison with known physical relationships in snowmelt hydrology from the literature

Reply: Implemented. We thank the reviewer for this insightful comment regarding the interpretability advantages of Kolmogorov-Arnold Networks. We have added a new subsection (section 4.3) to visualization of the learned univariate functions with symbolic regression analysis and discuss the interpretability of KAN. See Lines 484-520.

2. what are the hyperparameters (epochs, batch size, learning rate) and details of the architecture of the RNN, GRU and other neural nets used for comparison.

The manuscript lacks essential details for all baseline models (RNN, GRU, LSTM):

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

- No specification of hyperparameters (epochs, batch size, learning rate)
- No architectural details (number of layers, hidden units, activation functions)
- No information about initialization methods
- No training procedure details (optimizer type, learning rate schedules, dropout rates)
- No stopping criteria or early stopping procedures
- No hardware specifications or training times

Reply: Implemented. A subsection of model implementation and training has been added. See Lines 340-385.

3. Recent papers suggest that KAN based architectures outperform classical ANN based architectures. There should have been a comparison with KAN based LSTM, GRU and other neural nets. The manuscript only compares RCPIKLA (which uses KAN) against traditional ANN-based models (RNN, GRU, LSTM), not against KAN-enhanced versions of these baseline architectures.

The comparison with no physics informed constraints and no residual has been compared. However, the current experimental design still creates an attribution problem. Observed performance improvements could stem from:

- The KAN component specifically
- The attention mechanism
- The physics-informed constraints
- The residual compensation structure
- Seasonal trigonometric encoding
- Some synergistic combination of these components

Without proper ablation comparing LSTM-attention/KAN-LSTM/KAN-GRU versus RCPIKLA, the specific contribution of KAN remains unclear.

Reply: Implemented. We have added a new comparison model and a new evaluation

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

metric, discussions, and a new subsection 4.3 to demonstrate the contribution KAN. See Lines 442-452.

4. The manuscript describes a physics-informed constraint that imposes an upper limit on predicted snowmelt contribution but does not explain the asymmetric treatment of constraint violations.

The asymmetric design requires clear physical justification:

- Upper bound rationale: Snowmelt contribution physically cannot exceed available snow water equivalent - this is a hard constraint based on mass conservation
- Lower bound question: Are underpredictions physically plausible? Could incomplete melting, refreezing, or sublimation make them valid? Or do they indicate model failure to capture melt processes?
- Bias implications: Does the asymmetric penalty introduce systematic bias toward underprediction?

Reply: Implemented. The asymmetric physics constraint used in the manuscript represents a simplification of complex Arctic hydrological processes with available data. Discussions of rationale and justification regarding the upper bound rationale, lower bound question and bias implications have been added to the manuscript. See Lines 248-270 and 544-549.

5. Physics-informed neural networks fundamentally rely on balancing multiple loss terms through weighting parameters. The manuscript mentions α and β as weights for MSE loss and physics loss but does not report their values.

The manuscript must provide:

- Final α and β values used for all reported results
- Scenarios of hit and trials
- Search space explored

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

Reply: Impelemnted. A new subsection is added to the Supplementary Material to introduce the search process and justify the optimal choise. The values of α and β are explicitly stated in the manuscript. See Lines 361-363 in the manuscript and S1 in the Supplementary Material.

6. The manuscript lacks visualization of epoch-wise loss decomposition, which is important for assessment of convergence of all models. Without this analysis, it is impossible to assess whether the physics constraint meaningfully guides training or becomes negligible compared to the data-driven MSE loss.

Visualizing separate loss components reveals:

- Whether physics loss actually contributes to training or is overwhelmed by MSE loss
- Training stability and convergence behavior
- Potential issues: loss spikes, plateaus, phase transitions

Reply: We thank the reviewer for this suggestion to analyze training dynamics and loss component contributions. In the proposed model, a dual physics-guided approach with two components is implemented: 1) a snowpack layer, 2) a physics-informed loss constraint term. This manuscript included ablation analysis comparing models with and without both physics constraints. In Figure 10, it provided empirical validation and analyzed the results of RCPIKLA vs. RCKLA-no physics-informed components vs. PIKLA-no residual structure). We believe this analysis can address the reviewer's concern about the physics constraint's contribution.

7. Figure 6 (left): "y axis seems to be cut, the numbers are partly missing" - this affects readability and interpretation. Also, please check for spelling and grammatical errors throughout manuscript. Like a few spelling mistakes have been observed in abstract

Reply: Implemented. We thank the reviewer for pointing this out. We have carefully

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

reviewed the manuscript to correct the spelling mistakes. The plots have been updated. See Line 453 and Figure 6.

8. The physics-informed mechanism involves snow storage (S_t) and melt (M_t) terms that evolve over time. However, the manuscript does not specify:

- Initial values for S_0 and M_0 at the start of the simulation period
- How these initial conditions were integrated into the model?

Reply: We thank the reviewer for noting that the initial conditions were not explicitly stated. In our implementation, the snow storage (S_t) and melt (M_t) are initialized as zero at the beginning of each model input sequence for simplicity. Specifically, we set $S_0 = 0$ and $M_0 = 0$, and then update S_t and M_t recursively within the window based on precipitation and temperature. The computed term is integrated into the model by being added to the network-predicted discharge and by included in the physics-informed penalty term. See Lines 235-237. In the future, a continuous state carryover across windows that maintains snow storage between consecutive sequences could be considered for future work.

9. It is mentioned conducting 10 independent runs but provides unclear or incomplete reporting of variability in results. Fig8 represents the rmse and nse RCPIKLA variants with all predictions, what is the average RMSE over 10 runs, how much variation is observed over independent runs?

Additionally:

- Figure 8 shows results (RMSE and NSE for RCPIKLA variants) but it's unclear whether these represent single runs, mean values, or distributions

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

- No explicit reporting of mean \pm standard deviation for performance metrics
- No statistical significance testing comparing model variants

Reply: Implemented. We thank the reviewer for this important comment highlighting the need for comprehensive statistical reporting. We have substantially revised the figure and added statistical analysis to address the concerns. The figure caption is updated to explicitly state that each box plot aggregates results across forecasting horizons (1-12 months) and independent training runs, which produces 120 data points per model (12 time steps). See Lines 527-534, 550-554, 558-559 and Figure 10.

10. Figure 5 currently shows model predictions at 12 time intervals (representing different aggregation windows) but does not convey prediction uncertainty across the 10 independent runs. This limits the reader's ability to assess:

- Model reliability at different temporal scales
- Whether certain aggregation intervals show higher prediction variance

Reply: Implemented. We thank the reviewer for this valuable suggestion to quantify prediction uncertainty across forecasting horizons. To make run-to-run uncertainty visible, we summarize model performance across 10 independent training runs for each aggregation window. In the Supplementary Material, Tables S1–S3 report the mean, minimum, and maximum values values across runs for NSE, RMSE, and KGE' (2012) at each forecasting horizon. This should allow the readers to better evaluate model performance and reliability at different temporal scales and prediction variances, without overcrowding the main figure with multiple model curves. See Lines 393-394.

Summary:

The manuscript is recommended for publication if the above suggestions are addressed or answered.

Reply: We sincerely thank the anonymous reviewers for the positive comments and

Sam Houston State University

DEPARTMENT OF ENVIRONMENTAL AND GEOSCIENCES

constructive feedback. We have carefully revised the manuscript, added references and addressed comments in the manuscript.

Sincerely yours,

Renjie Zhou

Renjie Zhou
Associate Professor
Department of Environmental and Geosciences
Sam Houston State University
Huntsville, TX 77340