

# **TECHNICAL NOTE: ANALYSIS OF CONCENTRATION-DISCHARGE HYSTERESIS LOOPS USING SELF-ORGANIZING MAPS**

## **RESPONSE TO REFEREES**

Referees' comments appear in black; authors' responses are in red. Verbatim excerpts from the manuscript are indented and shown in black.

### **AUTHORS' RESPONSE TO REFEREE 3**

The authors have worked extensively on the manuscript. I appreciate the additional figures and explanations. And based on these, I have a few additional suggestions for clarification / improvements.

Line numbers below refer to the updated version without tracked changes.

Comment 2

The additional explanations and figures are helpful!

However, the titles and structure are confusing and I think are not well structured / do not align with the figure(s):

- 2.2 Applying SOM for hysteresis analysis
- 2.3 Representing hysteresis loops for SOM
- 2.4 Quantifying similitude between loops: The distance function
- 2.5 SOM training and evaluation
  - 2.5.1 SOM training and hyperparameter optimization
  - 2.5.2 Hyperparameter optimization
- 2.6 Evaluation of a trained SOM for hysteresis analysis

We appreciate your feedback and agree that, as you suggest below, adopting a section structure that better aligns with Fig. 3 would improve the organization of this part of the manuscript. Accordingly, we have restructured Section 2. Importantly, the content remains mostly unchanged, aside from minor edits made to ensure smoother transitions, and the addition of a new short section (2.4 Analysis phase) following one of your comments below. The revised structure of Section 2 is as follows:

- 2.1 Brief description of SOM
- 2.2 A two-phase workflow for applying SOM in hysteresis analysis
- 2.3 Training and evaluation phase
  - 2.3.1 Input data: Representing hysteresis loops for SOM
  - 2.3.2 SOM training
    - 2.3.2.1 Training setup
    - 2.3.2.2 Training algorithm
  - 2.3.3 Hyperparameter optimization

### 2.3.4 Evaluation of a trained SOM

#### 2.4 Analysis phase

Note also the *training phase* is now called *Training and evaluation phase* to better reflect its purpose.

I see the following issues:

- Why is "Analysis Phase" (Fig 3) not explained in the current chapter 2?

The analysis phase is initially described at a high level in Section 2.2; however, we have now added a dedicated subsection (Section 2.4, Analysis phase) to ensure consistency with the rest of the section. This addition provides a more concrete description of the analysis phase after the SOM algorithm has been thoroughly explained. The new Section 2.4 reads as follows:

#### *2.4 Analysis phase*

*Once the SOM has been trained, it can be used to classify hysteresis loops in new datasets. For each new loop, the two-dimensional  $Q-C$  sequence is compared against all map prototypes using the same distance function employed during training (i.e., DTW). The prototype yielding the minimum distance is identified as the BMU, and the loop is assigned to the corresponding map node. Applying this procedure to all samples in a dataset enables the results to be visualized as heatmaps, as illustrated in Fig. 3. In addition to visualizing the frequency of loop-type assignments, associations between loop types and hydrologic variables can be explored by mapping the corresponding variable values onto the SOM lattice. This approach is demonstrated in detail in the proof-of-concept presented below.*

- 2.4 is the nested step described as "Find BMU" in Fig 4, right? Why does it have an own chapter?

The step "Find BMU" is correct. The former Section 2.4 described a critical element of the SOM, namely how the algorithm quantifies similarities between hysteresis loops. Moreover, as explained in that section, the distance function we selected deviates from the more conventional Euclidean distance, which is why we dedicated a separate subsection to this topic. In the restructured Section 2, this explanation is now included in Subsection 2.3.2.1 (Training setup), where we further justify the need for this detailed discussion as follows:

*"As a preliminary step prior to training, several components of the SOM algorithm must be specified, including a distance function, a decay function, and a neighborhood function, which together control how input samples are compared and organized within the low-dimensional lattice during SOM training. We adopt standard formulations for the decay and neighborhood functions commonly used in SOM applications, as described in Section 2.3.2.2 and shown in Fig. 4. However, we provide greater explanation regarding our choice of distance function, as its application to hysteresis analyses warrants additional consideration.*

- 2.5 contains 'evaluation' but then 2.6 deals again with the "Evaluation"

Thank you for spotting this discrepancy. "Evaluation of the trained SOM" should be part of the former Section 2.5 (Section 2.3 in the revised version). It is now labeled as Section 2.3.4 following the updated structure.

- 2.5.1 and 2.5.2 both contain hyperparameter optimization - this is confusing.

Thank you for spotting this discrepancy. Hyperparameter optimization is now addressed exclusively in Section 2.3.3.

I think it would make sense to simply follow the main headings of the figure subsections and have these as the titles. Also, refer to the figures more closely when explaining a step so that the reader understands to which figure and where in the figure you exactly refer to.

We have updated the section structure as suggested and added additional references to Fig. 4 in Section 2.3.2.2 (Training algorithm) to help readers better follow the training procedure.

l.180 "map size iteratively" - this is not clear in Fig 4. I assume N is not the number of iterations for trying different hyperparameters. It rather is the number of iterations in the loop? Suggest to refer to 'N' in the relevant subtitle and indicate where the number of iterations for the hyperparameters come in. Can you 'nest' Fig 5 into Fig 4 with a note (i.e. as you have done for Fig 4 and 5 in Figure 3)? l.248-259 suggest to add that the final selection of the map size is a 'manual process'

We have added a reference to Section 2.3.3 and Fig. 5 in this sentence to direct the reader to where the procedure for defining the map size is described in detail. We also updated Fig. 4 to clarify that it is the number of epochs that is defined (as described in the text), from which N is then calculated. Finally, we added a reference to Fig. 4 in Fig. 5 to more clearly show how the training process illustrated in Fig. 4 is executed multiple times as part of the hyperparameter tuning procedure shown in Fig. 5.

Comment 3

The explanations of the SOM terminology in the response is very helpful - particularly as it is all in one place. I suggest to add this to the manuscript or (perhaps better) to the supplementary material.

Thanks for the suggestion, we have added a glossary in the SI with the relevant terminology.

Comment 11

"Event delineation was supported by a custom-built application designed to label hydrologic time series." The event delineation is a very important step and I suggest to add a few sentences how the "custom-built application" is achieving the event detection.

We have made explicit in the manuscript that the custom-built application aids in **manual** delineation of hydrologic events. The criteria used to delineate events are explicitly stated in the manuscript (section 3.1.1):

*Watershed selection and event delineation followed the criteria outlined below:*

- *Only rainfall-dominated hydrologic events were included; heavily regulated rivers and snowmelt-dominated watersheds were excluded.*
- *Multi-peak events were split into single-peak events if turbidity receded fully before the onset of the subsequent peak.*

- *Irregular loops (such as complex loops) or loop shapes that did not conform to the typologies illustrated in Fig. 6 were excluded.*

Comment 13

Suggest to add the 'epochs' to the figure where it is relevant.

The definition of the number of epochs and how it is related to the number of iterations (N) is now included in Fig. 4

Comment 14

Yes, it is clearer now. If I understand this correctly, it would make sense from my point of view to make this clearer in Figure 3 by adding: "SOM \*unsupervised\* training and hyperparameter optimization" and then instead of "Consistent mapping of samples to prototypes" under the evaluation panel write something like "Comparison of trained SOM with manual loop labels of curated dataset"

Thank you for your suggestions. We do not believe the term “unsupervised” is necessary, as SOMs can only be trained in an unsupervised manner, which is made explicit in several places throughout the manuscript. Regarding the evaluation note, we have revised it as follows:

*“Comparison of sample types and associated BMUs to ensure consistency”*

which we believe better reflects the purpose of the evaluation phase.