

# Machine Learning for Stochastic Flood Model Hydrograph Typing

**Elise Madonna**, Title, Bureau of Reclamation, Lakewood, CO, emadonna@usbr.gov  
**Drew Allan Loney**, PhD PE, Bureau of Reclamation, Lakewood, CO, dloney@usbr.gov  
**Amanda Stone**, PE, Bureau of Reclamation, Lakewood, CO, astone@usbr.gov

## Abstract

Stochastic rainfall-runoff models are at the forefront of hydrologic modeling state-of-practice. These models have increasingly been used by the Reclamation Technical Service Center (TSC) to estimate flood magnitudes and associated return periods, along with uncertainty, for detailed flood hazard studies such as issue evaluations (IEs) and corrective action studies (CASs). Stochastic rainfall-runoff models simulate many thousands of potential flood realizations spanning the frequency space to estimate probabilistic floods and support risk-informed decision making. Peak, volume, and maximum reservoir elevation characteristics of these thousands of flood realizations are then used to produce flood frequency curves to support understanding of flood risk at the facility.

One challenge that arises with these large datasets is when hydrographs must be used for additional analyses beyond determination of existing hydrologic loads, such as design, modification, or operational changes. In these scenarios, working with a smaller number of hydrographs becomes necessary. The process of selecting a representative subset of hydrographs is currently manual, time consuming, and dependent upon the judgement of the person tasked with selection.

To address the challenge of selection of representative hydrographs, an automated hydrograph classification workflow was developed using a two-stage classification procedure: self-organizing map (SOM) machine learning (ML) followed by mean shift clustering. The SOM method groups hydrographs by evaluating their similarity in the shape and magnitude. The SOM groups are further refined with the mean shift clustering operation to yield a small number of hydrograph clusters that are representative of the range of behavior at a site. The developed ML workflow is an automated process with minimal user input that runs rapidly and scales to the number of hydrographs produced by stochastic rainfall/runoff models. The ML hydrograph classification workflow was tested across multiple gage and model instances. In each of these cases, the ML workflow was robust and produced a hydrograph classification that is representative hydrologic variability of the site.

## Introduction

Stochastic rainfall-runoff models are considered a best practice in hydrologic modeling and have been increasingly used for dam safety decision-making (Bureau of Reclamation, 2003b; Sorooshian et al., 2008). Rainfall-runoff modeling has been used extensively to assess flood magnitudes and associated return periods at facilities, but incorporation of stochastic methodologies is more recent. Stochastic methods in rainfall runoff modeling treat inputs and

parameters as variables rather than set values. This allows for numerous combinations of input values to simulate some of the variability that is inherent in a hydrologic response. Stochastic rainfall-runoff methods create a large set of “realizations” of potential floods in a basin, but also introduce the complexity of a much larger set of hydrographs with which to work for decision-making. The purpose of this study is to develop a methodology that could be applied to a set of stochastically generated hydrographs to identify a smaller subset of representative hydrograph realizations to simplify the decision-making process.

Rainfall-runoff models combine precipitation-frequency estimates with a hydrologic model to evaluate runoff timing and magnitude (Beven, 2012; Devia et al., 2015; Sitterson et al., 2018). In contrast to streamflow measurements that are fixed to the precipitation events within the historical record, rainfall-runoff methods can be used for historical reanalysis as well as to explore behavior outside of the historical record (Dutta et al., 2012; Moore et al., 2001; W. Wang et al., 2021). Such exploration can include changing the precipitation magnitude, precipitation timing, antecedent basin conditions, or water management decisions. A hydrologic model used for a rainfall-runoff process would be calibrated/validated over the historical record and then driven with the varied conditions to understand the potential response of the basin. The flexibility, accuracy, and utility of rainfall-runoff methods have made them a key component across Reclamation water management and dam safety activities (Bureau of Reclamation, 2016, 2019).

When combined with a stochastic framework, a rainfall-runoff model can be used to estimate the likelihood, magnitude, and timing of flows with statistical rigor (Bureau of Reclamation, 2003b; Marco et al., 1993; Sorooshian et al., 2008). Probabilities are formulated for rainfall magnitude, rainfall distribution, and various antecedent conditions, including soil moisture and channel conditions, prior to application in the stochastic model (Tabari, 2019). These distributions are repeatedly sampled to create initial and boundary conditions for the rainfall-runoff model which is solved to obtain a hydrograph realization. By sampling the probability distributions many thousands of times, a flow-frequency or volume-frequency curve can be estimated with appropriate statistical assumptions. However, a peak streamflow or volume frequency curve distills the response of the basin to a single scalar representing the model solution. As useful and important to Reclamation are the hydrograph realizations resulting from the stochastic runs that give the flow as a function of time during the events.

Hydrograph shape and magnitude are used to understand hydrologic events and categorize the basin response. The flow magnitude, the number of peaks, and the spacing between peaks characterize the interactions of physical processes that govern the hydrology of a basin (Singh, 1997). A hydrograph is also utilized in flood operations and routing (Federal Emergency Management Agency, 2020). The shape and magnitude of a hydrograph determines how a facility operator may release water for emergency response. Flood routings use hydrographs to determine inundation depths and durations across a facility (Bureau of Reclamation, 2003a). Having a set of hydrograph realizations that is representative of the range of possible basin behavior is particularly important for these latter cases to understand and plan for scenarios at a facility.

As the number of hydrographs increases with the number of stochastic runs, it becomes increasingly challenging to understand and use the hydrograph realizations for things such as design and inundation mapping. Current Reclamation hydrograph classification methods call for a hydrologist to review and manually identify hydrographs for use. Facing several thousand hydrograph realizations, a manual process is time consuming and subject to the judgement of the analyst. While classifying hydrograph realizations could reduce the stochastic hydrograph set to a representative sample, current manual hydrograph classification is not feasible to enable such a reduction. A new classification tool is therefore necessary to support hydrograph classification for both improving hydrologic understanding and flood routings.

## Methods

The proposed classification approach consists of two steps: a SOM based approach for initial classification and a traditional clustering approach for subsequent refinement. The classification approaches are complementary. SOM methods are particularly useful for sorting hydrograph realizations into representative categories because they can classify large amounts of non-linear, high-dimensional data into a lower dimension (Holman, 2018; Kohonen, 2013). The ability of the method to capture details in high dimensional data was relevant due to the importance of hydrograph shape in the classification application cases. Additionally, SOMs generate a more sophisticated output layout than other clustering methods by creating an output node placement map that is indicative of similar nodes (Holman, 2018; Kohonen, 1982). A clustering algorithm is used subsequent to the SOM classification to further consolidate the SOM groups into a smaller, more hydrologically meaningful number of categories. This process follows the same fundamental concept as hierarchical clustering, which uses multiple clustering algorithms sequentially to achieve an optimal classification.

### SOM

SOMs function as single layer NNs that use an unsupervised learning algorithm, meaning that the algorithm discovers patterns in input data for which the correct categorization is unknown a priori by the user (Kohonen, 2001). Since their initial publication, SOMs have been broadly applied and adapted across a variety of use cases (Barreto, 2007; Kalteh et al., 2008; Kohonen et al., 1996). SOMs create a two-dimensional  $m$  by  $n$  gridded output map where  $m$  and  $n$  are dimensions that have been specified by the user (Kohonen, 2001). Each grid node contains a weight vector representing the hydrograph realization vectors that are most similar to it (Kohonen, 2013). The node weight vectors are initialized with random values and then trained by comparing every vector from the input hydrograph realizations one at a time to the SOM weight vectors until a steady state is achieved or until the user-specified number of training iterations has been reached. Each input realization vector is assigned to a node by calculating the Euclidean distance between it and every SOM node weight vector, then choosing the node that minimizes that distance. When a realization vector is assigned to a node, the weight vectors of the node and its neighbors get updated to become more similar to the assigned realization vector. The weight vector is therefore representative of the data assigned to it at the end of the SOM training (Lin & Wang, 2006). The nodes are organized throughout the output map in a way that reflects their similarity (Kohonen, 2013). The most similar realization vectors should be contained in nodes that are adjacent in the output map. This effort utilized version 2.2.9 of MiniSom package in Python (Giuseppe, 2018).

The SOM classification included two additional scalar parameters – the number of peaks and largest peak value in each hydrograph – in addition to hydrograph timeseries. Both of these parameters are calculated from the hydrograph realization prior to training the SOM and are appended to the hydrograph to form a vector representing the realization. To maintain scaling between the additional parameters and hydrograph timeseries, each is scaled to the timeseries so as to not bias the distance metrics.

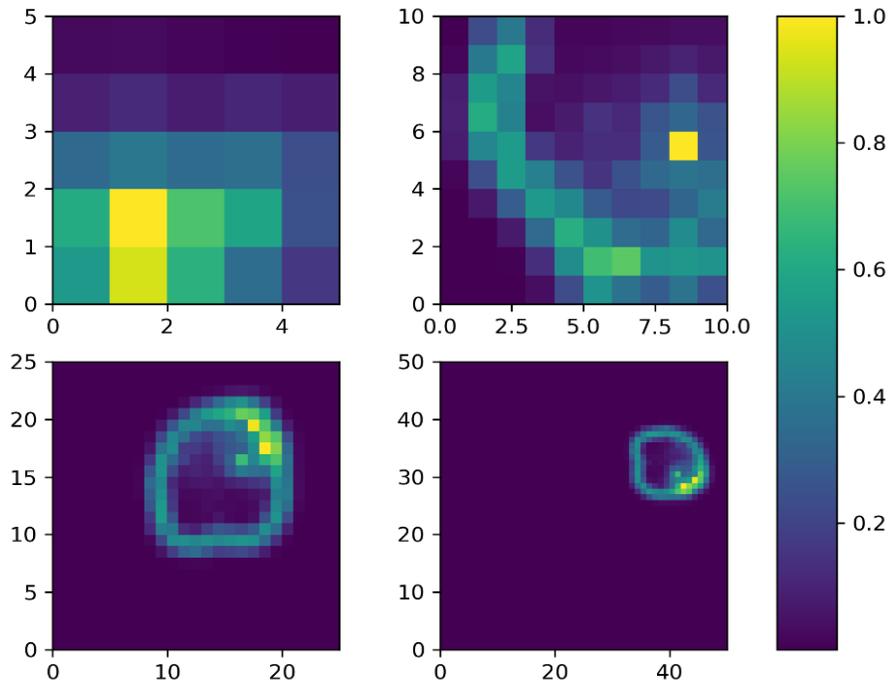
The peak streamflow in each hydrograph realization is calculated by taking a maximum of the hydrograph timeseries. The maximum is then normalized by first dividing by the mean streamflow of the hydrograph realization and then multiplying by the mean streamflow of all hydrograph realizations. This places the peak value on the same scale as the mean of the input

data while preserving the distribution of the peak values. Including this variable improved the ability of the SOM to partition the hydrograph realizations based on magnitude.

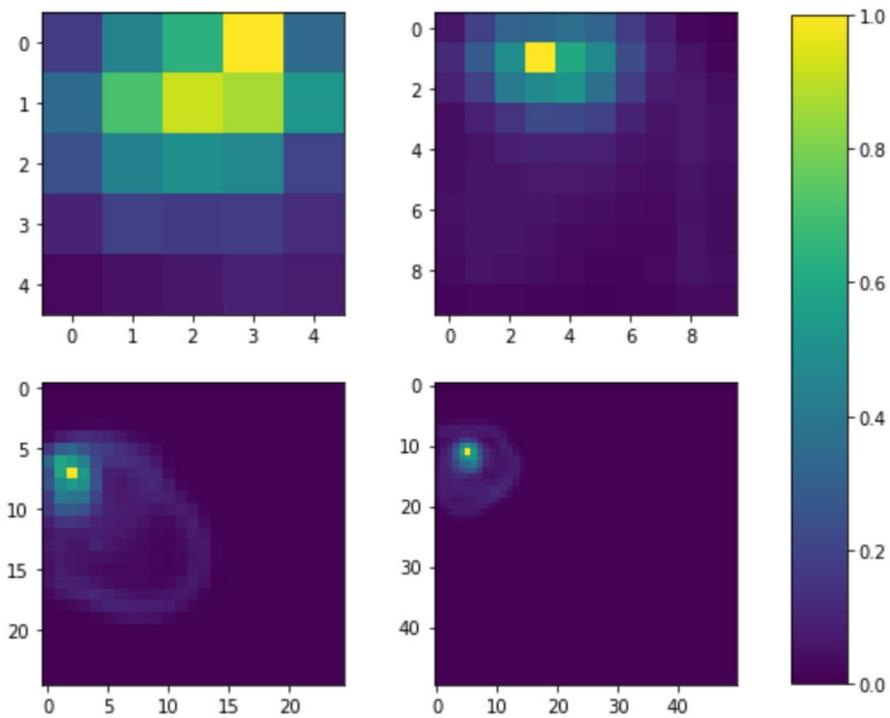
The number of peaks in each hydrograph is calculated in reference to the 85<sup>th</sup> percentile of the hydrograph realization. The 85<sup>th</sup> percentile is subtracted from the hydrograph realization, and the number of sign changes in the resulting series is used to count the number of peaks. Referencing the number of peaks to the 85<sup>th</sup> percentile suppresses small peaks from the count that are less relevant to the classification. The number of peaks is normalized for each hydrograph realization by dividing by the maximum number of peaks in any hydrograph realization and then multiplying by the overall mean streamflow of all hydrograph realizations. This again places the number of peaks on the same scale as the mean of the input data while preserving the count distribution. Including this variable improved the ability of the SOM to discriminate the shape of the hydrographs.

A 25 by 25 grid was set as the default SOM dimensionality, giving 625 possible nodes to utilize in the classification. The shape of the SOM space was not specified a priori and allowed to evolve naturally from the training process. The SOM dimensions were determined through a sensitivity analysis on the distance map evaluated across the test datasets, as shown in Figures 1 and 2. The distance map represents the normalized sum of the Euclidian distance between the nodal weight vectors of a node and its neighbors. The neighborhood function used to relate nodes was taken as a Gaussian distribution. Changes in distance between nodal weight vectors indicate how well resolved and trained the SOM classification is based on grid size and training iterations. As one increases the number of nodes available in the grid, the SOM algorithm will reach a size at which additional nodes no longer affect the classification. When using an unsupervised learning algorithm without any a priori knowledge of the correct classification structure, over resolving the SOM allows the algorithm to utilize an arbitrary number of classification groups without artificially forcing the classification. This size allows the SOM enough nodes in the output to fully capture representative characteristics of the hydrographs, while also limiting the number of empty cells present in the output. Although this size results in more categories than are likely possible as meaningful hydrologic groups, the number of SOM groups is further refined in later stages of the classification process. Additionally, this resolution simplifies the use of the classification process by eliminating need for the user to specify SOM dimensions that best fit each new dataset.

Specifying the number of training iterations is also important when constructing the SOM. With each training iteration, the model goes through every hydrograph realization, calculating the closest node by distance and updating the nodal weight vectors. The number of training iterations dictates the total number of times the algorithm loops through the hydrograph realizations to update the nodes. Increasing the number of iterations improves the classification accuracy of the SOM but also requires additional compute time. These two considerations are balanced by setting a tolerance threshold (a percent difference of 0.1%) between the results of the current and previous training iterations, and then testing how many iterations are necessary to be within the threshold. It was determined that using 2,500 training iterations is sufficient to meet the tolerance threshold using the test datasets. This number of iterations is therefore used as the default value in the classification process to train SOMs, but it may be updated by the user if it proves insufficient for a specific case.



**Figure 1.** Visualizations of distance maps for SOM dimensions of 5 x 5, 10 x 10, 25 x 25, and 50 x 50 from left to right for El Vado SEFM data.

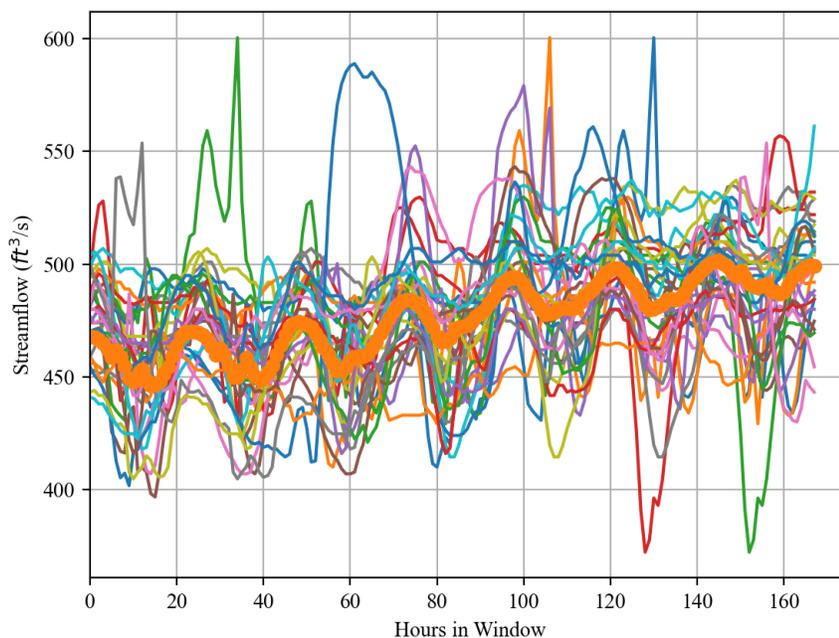


**Figure 2.** Visualizations of distance maps for SOM dimensions of 5 x 5, 10 x 10, 25 x 25, and 50 x 50 from left to right for Colorado River gage data.

While SOMs have proven to be powerful tools for classifications of this type, there are important limitations to the use of the method and interpretation of its output. First among these are resolution and accuracy. The dimensionality of the SOM is proportional to the resolution of the SOM. Because the number of output nodes is specified as an input to the algorithm, care must be taken not to generate a SOM that is either too small to encompass the entirety of variability among the input data or so large that the majority of SOM nodes do not contain any realization vectors (C  r  ghino & Park, 2009). An overfitted model will most likely have many nodes that each only represent one or two input vectors, while an underfitted model will lump the majority of input vectors into several nodes.

However, determining the correct SOM resolution for an accurate classification remains challenging, particularly for individuals without a data analytics background. This classification process therefore purposefully overfits the SOM map to reduce the problem from thousands of hydrograph realizations to less than a thousand SOM weight vectors that are representative of the grouped hydrographs. This is then used by a second clustering operation, outlined subsequently, to further reduce the number of groupings. This developed two-step workflow is thought to be more robust across different input datasets and minimizes user input. Additionally, it provides for a large reduction in the data dimensionality, which improves the skill of clustering algorithms.

The performance of the SOMs can be assessed visually by generating plots for each SOM node and observing how well each node weight vector represents characteristics of the assigned the hydrographs. Figure 3 shows an example of how the SOM node weight (shown as a thick line on top of the others) represents the hydrographs contained within the node (with each individual hydrograph shown as a thin line) for the USGS Colorado River dataset (U.S. Geological Survey, 2016). Plots of this type were used to assess the quality of the SOM hydrograph classification when individually applied to each test dataset.



**Figure 3.** An example of the hydrograph plots used to assess quality of the SOM’s fit to the data. The individual hydrographs are shown as thin lines, and the weight vector of the SOM is shown as a thicker line over the top of the hydrographs. Data is taken from the USGS Colorado River test dataset (U.S. Geological Survey, 2016).

## Clustering

Mean shift clustering operates by using kernel density and centroid estimation to evaluate similarity. This clustering method represents the parameter space as continuous and iteratively shifts each datapoint until it is as close as possible to the mean of the nearest kernel density estimation surface peak that represents a cluster (Comaniciu & Meer, 2002; Pedregosa, F. et al., 2021). Between iterations, the kernel density surface is updated, and the centroids of the clusters are recalculated. The algorithm is converged when the difference in the centroids of the clusters is within a tolerance between iterations.

Mean shift clustering was chosen over the other clustering methods because it is not necessary to specify a clustering depth or number of clusters prior to the analysis. Mean shift clustering models also allow the user to specify and adjust a radial bandwidth function which represents the datapoints to indirectly influence the number of clusters. However, many mean shift clustering implementations also provide a means to estimate a bandwidth using various heuristics. Because the clustering model is unsupervised, clustering performance is assessed visually by generating plots and metrics for each cluster and observing how distinctly each cluster represents hydrograph characteristics.

The mean shift clustering algorithm utilizes the output of the SOM classification to reduce the dimensionality of its input data and as an initial classification. Unlike the SOM, which uses the hydrograph timeseries in addition to scalar parameters, the mean shift clustering algorithm is not given the hydrograph timeseries as an input and uses only scalar parameters. The mean shift clustering algorithm incorporates the distance map from the SOM and, from the nodal weight vectors, the area under the curve (AUC), the number of peaks, and the largest peak magnitude. The weight vectors of each SOM node are utilized rather than the hydrographs directly because, through the SOM training process, the weight vectors are updated to reflect the hydrograph realizations that have been assigned to the SOM node. The behavior of many input hydrograph realizations can therefore be described by a single nodal weight vector. Additionally, the SOM distance map contains information about the hydrograph shape as similar features are placed more closely on the map. Use of the distance map in this manner reduces the dimensionality of including the hydrograph shape from a high dimensional problem based on the number of measurements in the timeseries to a single scalar that can serve as a proxy for the same information. The SOM distance map therefore provides the clustering with the primary magnitude and shape information.

AUC for the nodal weight vector is calculated by using trapezoidal integration. The AUC is analogous to the volume of the hydrograph realizations being represented by the weight vector and was introduced into the clustering algorithm to better distinguish the integrated magnitude of the nodal weight vector.

The number of nodal weight vector peaks and largest magnitude peak are calculated similar to the SOM inputs. These parameters are retained in the mean shift clustering analysis to improve discrimination of the hydrograph realization shape. While the distance map gives the relative overall similarity between two nodal weight vectors, it does not provide specific information about the shape of the nodal weight vector. These two additional scalar parameters improve the ability of the mean shift algorithm to discriminate hydrograph shape when different shapes are closely spaced in the SOM distance map.

In contrast to the calculated parameters which are input to the SOM, the mean shift clustering parameters are not scaled to the mean of the hydrographs. Instead, a standard scalar is used to normalize each parameter from zero to one, with the maximum value in the individual parameter

being set equal to one. Given the dimensionality of the SOM, clustering will have at most 625 inputs. In practice however, the number of inputs from the SOM are much lower. Among the datasets tested, the SOM categorizes input hydrographs into a minimum of 46 nodes and a maximum of 214 nodes.

The sensitivity of the clustering output to the bandwidth was examined for two test datasets. An initial bandwidth is estimated as the average maximum distance among randomly ordered groups of datapoints estimated using k-Nearest Neighbor (kNN) (Pedregosa, F. et al., 2021). The kNN searches for 30% the number of datapoints as neighbors and uses a Euclidean distance metric. Because a large number of randomized groupings are used to obtain an average distance, the sensitivity to kNN initialization is minimized.

This maintained a constant initial bandwidth for each dataset to which a bandwidth correction was applied. The correction factor is defined as:

$$\beta' = \frac{\beta}{k}$$

where  $\beta$  is the initial bandwidth,  $k$  is a constant correction factor, and  $\beta'$  is the updated bandwidth estimate. Larger correction factors therefore produce a smaller bandwidth and will tend to increase the number of clusters produced by the mean shift algorithm. Sensitivity to the bandwidth was determined by keeping a constant initial bandwidth and progressively changing values of the correction factor.

A default bandwidth correction of 1.25 is used in the mean shift clustering algorithm because it resulted in a moderate number of clusters for all examined datasets. Additionally, upon visual inspection of the clustering, meaningful regime differences in hydrograph characteristics could be identified among the clusters without a large number of clusters due to overfitting.

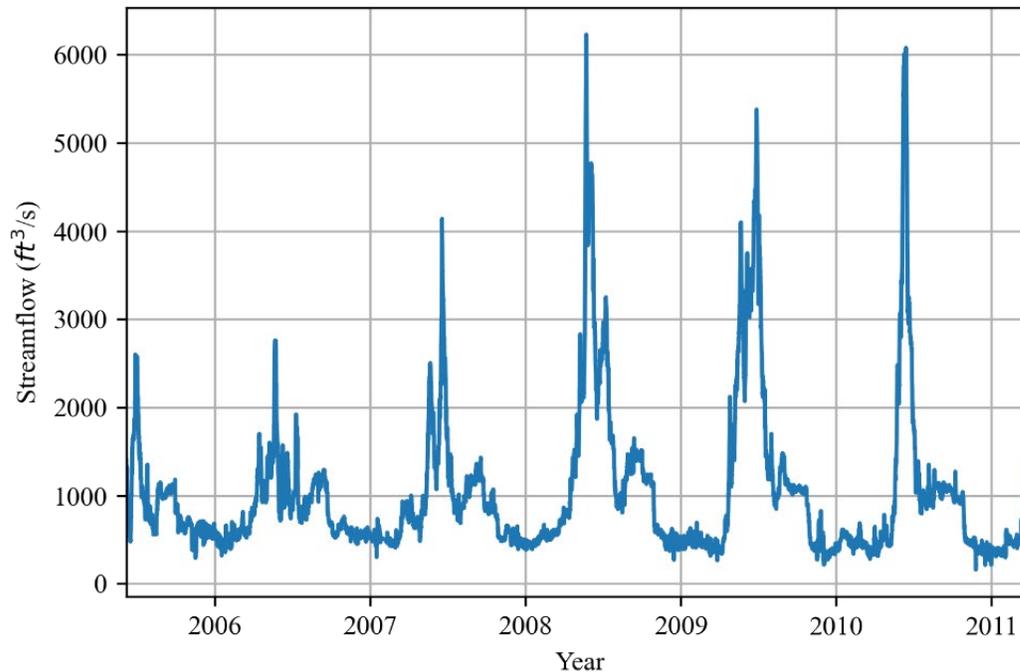
## Results

### Observed Dataset Results

The first set of observed gage data is from United States Geological Survey (USGS) site 09058000 on the Colorado River near Kremmling, Colorado (U.S. Geological Survey, 2016). This dataset contains streamflow measurements taken every 15 minutes from June, 2005 through June, 2021, and was retrieved from the USGS water data database on June 29<sup>th</sup>, 2021. The streamflow values in this dataset range from 158 ft<sup>3</sup>/s to 6230 ft<sup>3</sup>/s and have an average around 1030 ft<sup>3</sup>/s. Figure 4 plots the streamflow over the analyzed period. This dataset is considered to be an example of moderate to high flow gage data along a major river system. A seven-day window was arbitrarily utilized with a non-sliding window, meaning each day occurs uniquely in a single hydrograph. Both the size and type of window are parameters that can be changed within the workflow.

Figure 5 gives shows favorable results in that each cluster is clearly differentiated in either magnitude or shape (or both) from all other clusters. The SOM and mean shift clustering models have together been able to consolidate the hydrographs into a relatively small number of final categories which are all distinct from one another for this dataset. The clusters partitioned strongly on hydrograph volume with shape being a secondary characteristic, as evidenced by

several clusters containing hydrographs with both increasing and decreasing mean slope. If hydrograph shape is an important consideration, an increase in the bandwidth correction factor would be necessary to increase the number of output clusters.



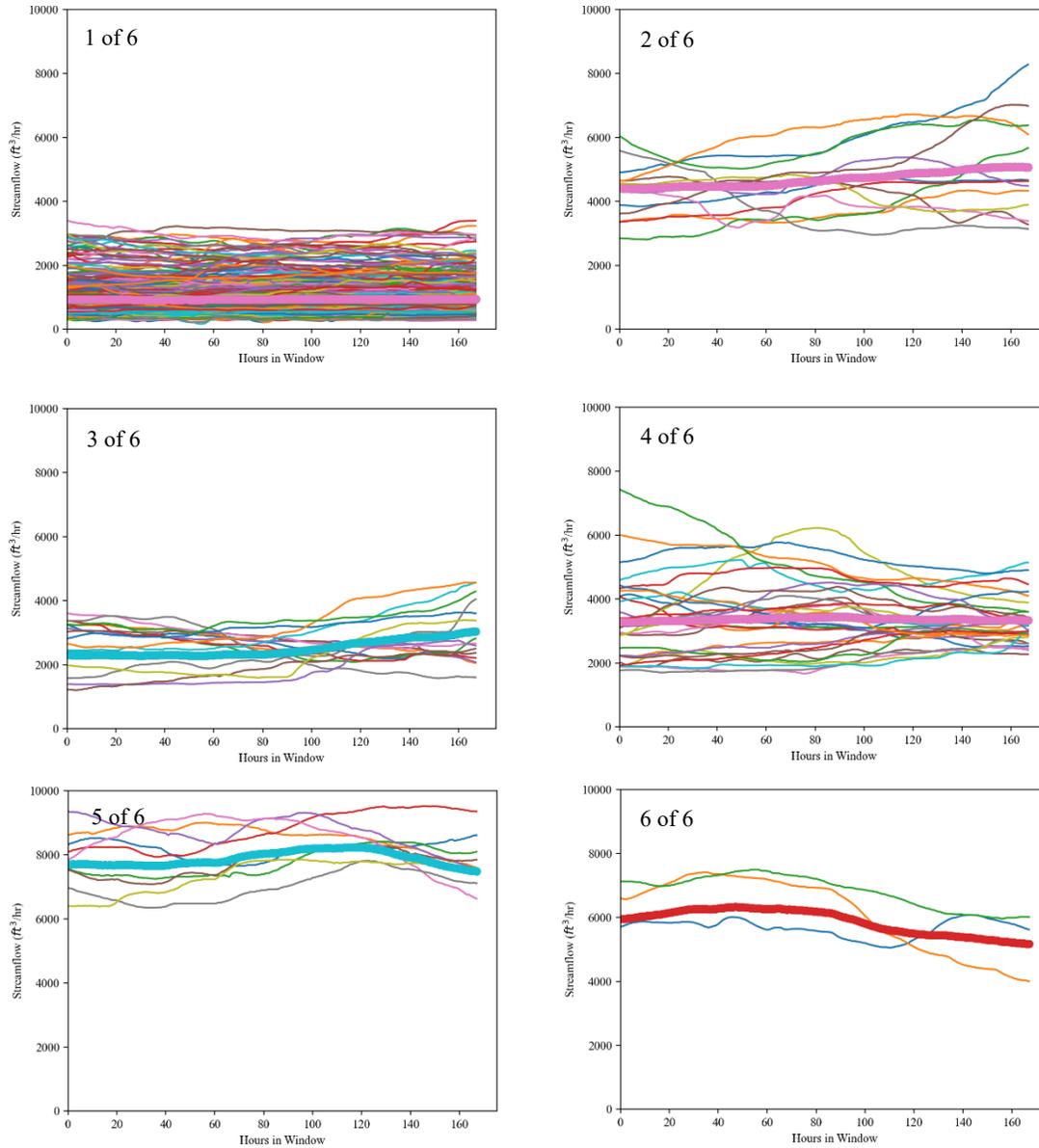
**Figure 4.** A plot of the USGS Colorado River streamflow dataset over the period of record considered in the classification process (U.S. Geological Survey, 2016).

The SOM uses several factors when it is classifying the input hydrographs as illustrated in Figure 5. The hydrograph magnitude and shape over time are the most intuitive factors from these plots; however, the SOM also uses other patterns and more complex factors which are less easy to observe such as the overall shape of the hydrograph. Additional SOM node weight plots for each dataset were generated when analyzing the dataset. These plots show that the SOM performs well to categorize the hydrograph realizations into representative groups. A full description of the clustering can be found in Bureau of Reclamation, 2022.

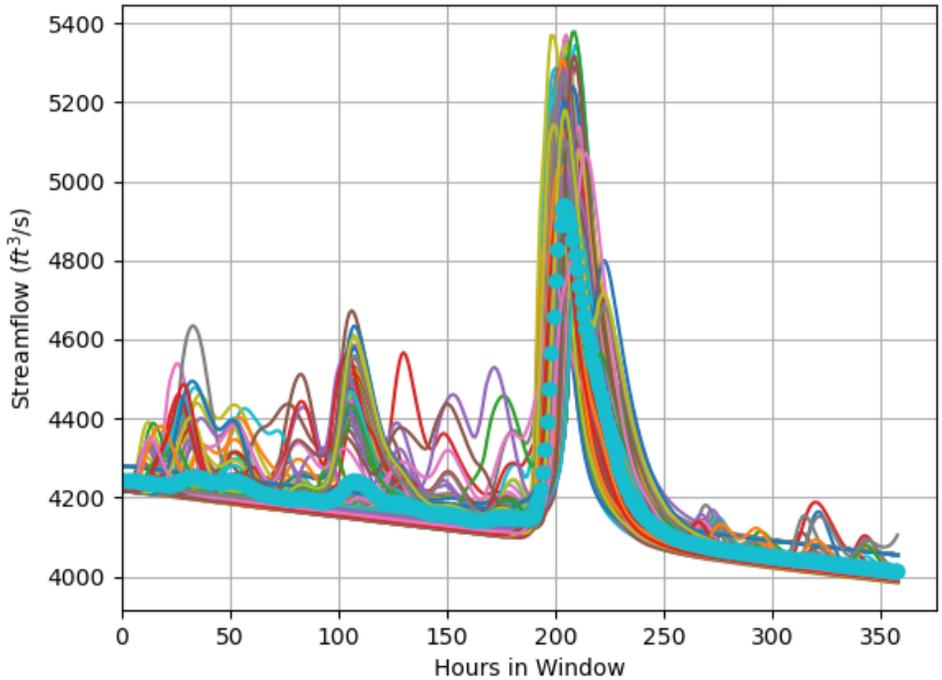
### **Simulated Data Results**

The simulated dataset used in testing was a set of hydrographs generated by the Stochastic Event Flood Model (SEFM) for the El Vado Lake in New Mexico (Bureau of Reclamation, 2016). Hourly streamflow values were generated for each 15-day hydrograph, and the data ranged from 3981 ft<sup>3</sup>/s to 55,230 ft<sup>3</sup>/s with an average of 4,409 ft<sup>3</sup>/s. These values are selected from a case in which Heron Dam is at maximum capacity, and all inflows are passed immediately downstream. Additionally, while the flows in the observed datasets tend to gradually increase or decrease and have multiple peaks at varying locations, the hydrographs in the simulated datasets mostly contain a small number of significant peaks and otherwise remain at a baseflow. Figure 6 shows an example of one SOM cell containing SEFM hydrograph data. The 20,000 total hydrographs in the SEFM data

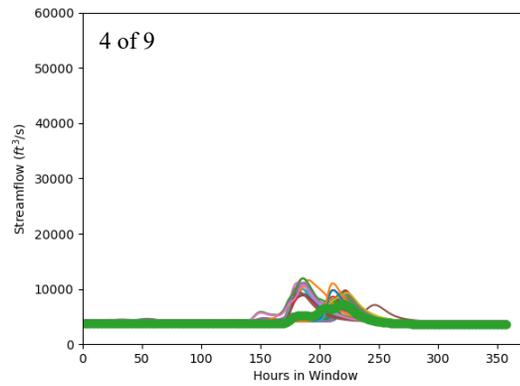
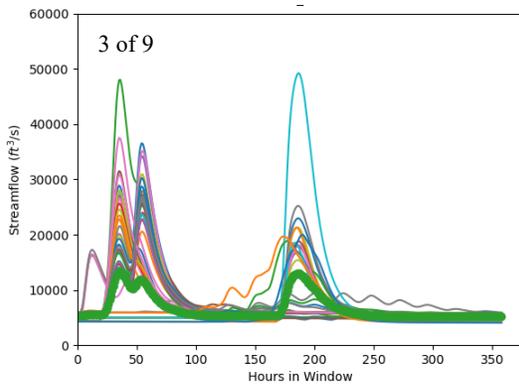
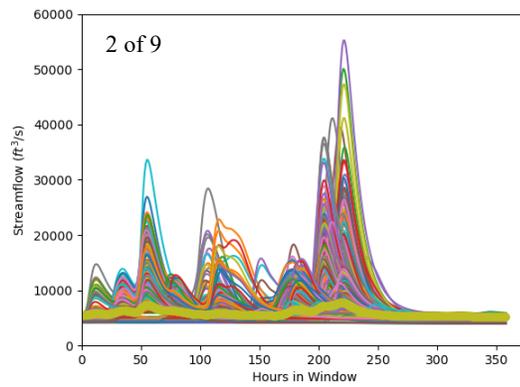
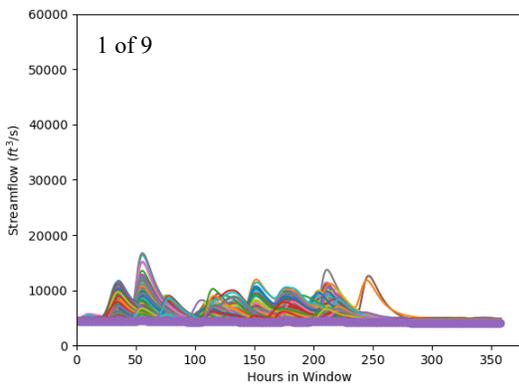
were classified into 80 SOM cells and nine final cluster categories, the latter shown in Figure 7. The window utilized the full period of the SEFM hydrographs.

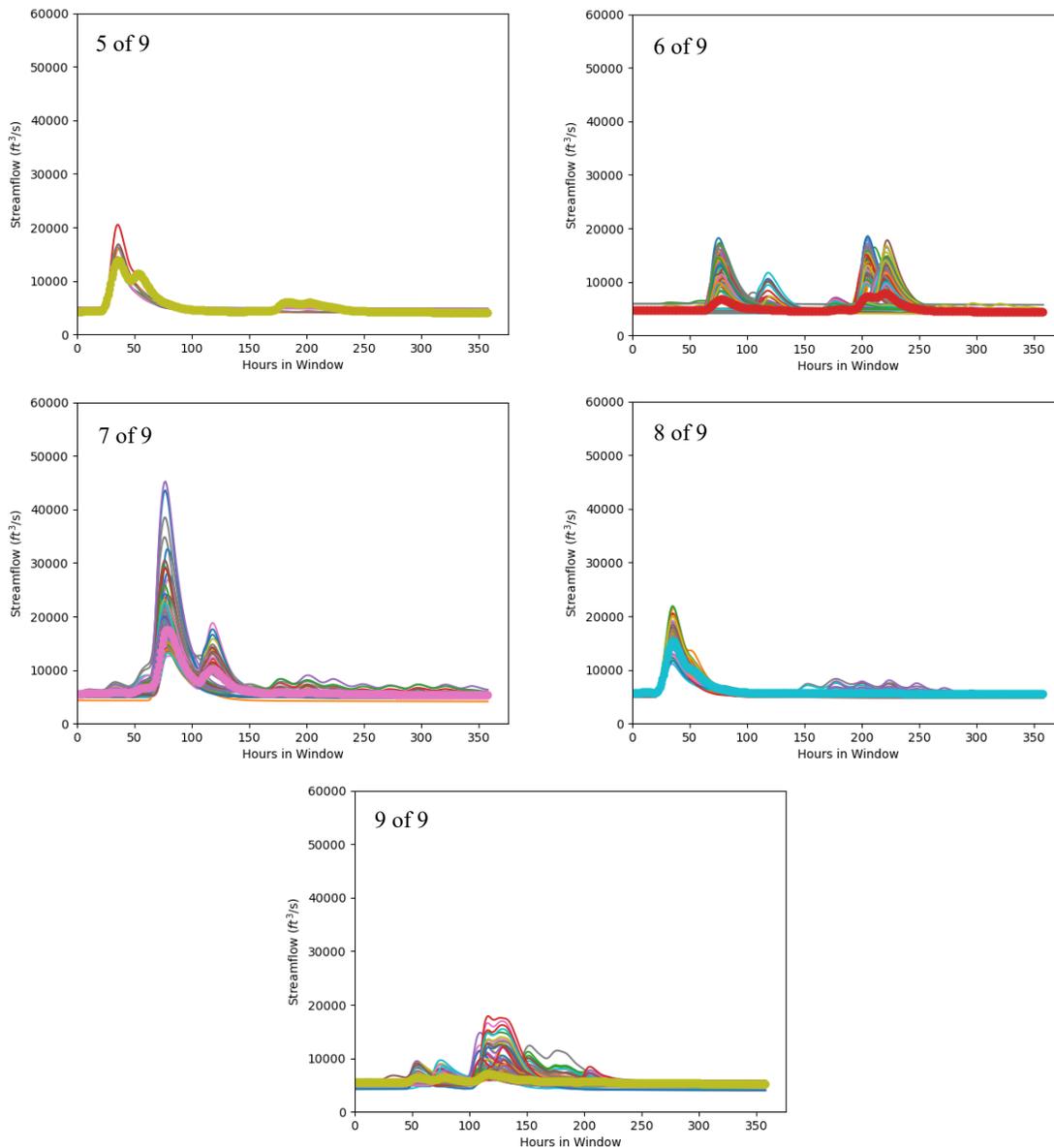


**Figure 5.** Cluster plots generated for USGS Colorado River data. Hydrographs from each cluster are shown as thin lines and the average weight vector for all SOM cells in the cluster is shown as a thicker line over the hydrograph lines.



**Figure 6.** One example of a SOM cell containing hydrographs from SEFM El Vado Lake data (Bureau of Reclamation, 2016). The individual hydrographs contained in the cell are shown as thin lines and the representative weight vector of the cell is shown as a thicker line on top of the individual hydrographs.





**Figure 7.** Cluster plots generated for SEFM El Vado data with fixed ranges to highlight differences in flow magnitude between the clusters. Hydrographs from each cluster are shown as thin lines and the average weight vector for all SOM cells in the cluster is shown as a thicker line over the hydrograph lines.

Although the input hydrographs from this simulated data differ significantly from the those of the observed gage data, the SOM and mean shift clustering methods maintain a high degree of classification skill. The differences in hydrograph shape and magnitude remain well captured. This is likely helped by the smoothness of the modeled hydrograph giving emphasis to small changes in the hydrograph shape. The degree of classification skill gives confidence in the overall classification workflow across different hydrograph sources.

## Conclusions

This work developed a process in which two algorithms are used to classify hydrograph realizations by similarity, providing a means to summarize many thousands of hydrograph realizations into a few representative groups that describe the overall streamflow behavior. The first classification employed a SOM ML algorithm for a broad first classification and dimensionality reduction. The second classification used the SOM outputs within the mean shift clustering algorithm to further reduce the number of classification groups. Validation of the approach was done by a qualitative assessment across test datasets representing a range of observed and simulated hydrologic behavior. Overall compute time across the investigated datasets was less than thirty minutes, although time will increase as a function of the number of provided hydrographs.

Multiple use cases are envisioned for the hydrograph classification process. The primary use is anticipated to be within stochastic hydrology applications. A reduction to representative hydrographs from a larger set of hydrograph realizations will facilitate better communication of hydrologic behavior. Additionally, it will simplify subsequent analyses that use the hydrographs as input while helping to ensure that the selected hydrographs are representative of the full range of basin behavior. Another anticipated use case is analysis of observed gage information to better understand the range of historical behavior. This can be useful in the construction of stochastic hydrologic models, to ensure that the simulated output is representative, and for water management, to better understand the pattern of common events. Across use cases, the expectation is that the user will select a hydrograph from within each cluster that is suitable for their task.

This effort improves hydrograph classification capability, and future efforts could provide additional refinement. While several representative test datasets were used to configure the default classifications parameters, this does not encompass the full range of input data. The classification process should be subject to continuing validation as it is applied to more data. If the classification skill becomes unsatisfactory, the SOM size or mean shift bandwidth default parameterization should be updated with additional logic based on the identified deficiencies. Additionally, the current approach is focused on classifying pre-existing hydrograph realizations. This workflow could readily be expanded to produce hydrograph realizations that are representative of the cluster. This could improve computation time by reducing the need for rainfall/runoff modeling to explore a particular hydrologic regime.

## References

- Barreto, G. A. (2007). Time Series Prediction with the Self-Organizing Map: A Review. In B. Hammer & P. Hitzler (Eds.), *Perspectives of Neural-Symbolic Integration* (pp. 135–158). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-73954-8\\_6](https://doi.org/10.1007/978-3-540-73954-8_6)
- Beven, K. J. (2012). *Rainfall-Runoff Modelling: The Primer* (2nd ed.). Wiley-Blackwell.
- Bureau of Reclamation. (2003a). *Probabilistic Extreme Flood Hydrographs That Use PaleoFlood Data for Dam Safety Applications* (Dam Safety Office No. 03–03). Department of the Interior.

- Bureau of Reclamation. (2003b). Stochastic Modeling Methods (Dam Safety Office No. 03–04). Department of the Interior.
- Bureau of Reclamation. (2016). El Vado Dam Hydrologic Hazard for Corrective Action Study (Technical Memorandum No. 8250-2016–014). Technical Services Center.
- Bureau of Reclamation. (2019). Stochastic Flood Frequency Analysis at Folsom Dam: (Technical Memorandum ENV-2019-070; p. 232). Bureau of Reclamation.
- Bureau of Reclamation. (2022). Investigating Methods for Stochastic Flood Model Hydrograph Extraction (Technical Memorandum ENV-2022-59). Technical Services Center.
- C er ghino, R., & Park, Y.-S. (2009). Review of the Self-Organizing Map (SOM) approach in water resources: Commentary. *Environmental Modelling & Software*, 24(8), 945–947. <https://doi.org/10.1016/j.envsoft.2009.01.008>
- Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 603–619. <https://doi.org/10.1109/34.1000236>
- Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A Review on Hydrological Models. *Aquatic Procedia*, 4, 1001–1007. <https://doi.org/10.1016/j.aqpro.2015.02.126>
- Dutta, D., Welsh, W. D., Vaze, J., Kim, S. S. H., & Nicholls, D. (2012). A Comparative Evaluation of Short-Term Streamflow Forecasting Using Time Series Analysis and Rainfall-Runoff Models in eWater Source. *Water Resources Management*, 26(15), 4397–4415. <https://doi.org/10.1007/s11269-012-0151-9>
- Federal Emergency Management Agency. (2020). Guidance for Flood Analysis and Mapping (Guidance Document No. 85). Department of Homeland Security. [https://www.fema.gov/sites/default/files/documents/fema\\_flood-profiles-guidance.pdf](https://www.fema.gov/sites/default/files/documents/fema_flood-profiles-guidance.pdf)
- Giuseppe Vettigli. (2018). *MiniSom: Minimalistic and NumPy-based implementation of the Self Organizing Map (2.2.9)* [Python]. <https://github.com/JustGlowing/minisom>
- Holman, K. D. (2018). Characterizing Antecedent Conditions Prior to Annual Maximum Flood Events in a High-Elevation Watershed Using Self-Organizing Maps. *Journal of Hydrometeorology*, 19(11), 1721–1730. <https://doi.org/10.1175/JHM-D-17-0229.1>
- Kalteh, A. M., Hjorth, P., & Berndtsson, R. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environmental Modelling & Software*, 23(7), 835–845. <https://doi.org/10.1016/j.envsoft.2007.10.001>
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1), 59–69. <https://doi.org/10.1007/BF00337288>
- Kohonen, T. (2001). *Self-Organizing Maps* (3rd ed., Vol. 30). Springer, Berlin, Heidelberg.
- Kohonen, T. (2013). Essentials of the self-organizing map. Twenty-Fifth Anniversay Commemorative Issue, 37, 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>
- Kohonen, T., Oja, E., Simula, O., Visa, A., & Kangas, J. (1996). Engineering applications of the self-organizing map. *Proceedings of the IEEE*, 84(10), 1358–1384. <https://doi.org/10.1109/5.537105>

- Lin, G.-F., & Wang, C.-M. (2006). Performing cluster analysis and discrimination analysis of hydrological factors in one step. *Advances in Water Resources*, 29(11), 1573–1585. <https://doi.org/10.1016/j.advwatres.2005.11.008>
- Marco, J. B., Harboe, R., & Salas, J. D. (1993). *Stochastic Hydrology and its Use in Water Resources Systems Simulation and Optimization* (Vol. 237). Springer, Dordrecht.
- Moore, R. J., Bell, V. A., & UK Environment Agency. (2001). Comparison of rainfall-runoff models for flood forecasting—Part 1: Literature review of models. UK Environment Agency.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2021). Scikit-learn: Machine Learning in Python (1.0.1) [Python]. <https://scikit-learn.org/stable/index.html>
- Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11(12), 1649–1669. [https://doi.org/10.1002/\(SICI\)1099-1085\(19971015\)11:12<1649::AID-HYP495>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1099-1085(19971015)11:12<1649::AID-HYP495>3.0.CO;2-1)
- Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Avant, B., & Muche, M. (2018). An Overview of Rainfall-Runoff Model Types. *International Congress on Environmental Modelling and Software.*, 41, 10.
- Sorooshian, S., Hsu, K., Coppola, E., Tomassetti, B., Verdecchia, M., & Visconti, G. (2008). Hydrological modelling and the water cycle: Coupling the atmospheric and hydrological models. Springer.
- Tabari, H. (2019). Statistical Analysis and Stochastic Modelling of Hydrological Extremes. *Water*, 11(9), 1861. <https://doi.org/10.3390/w11091861>
- U.S. Geological Survey. (2016). USGS Water Data for the Nation—Gage 09058000. U.S. Geological Survey. <https://waterdata.usgs.gov/monitoring-location/09058000/#parameterCode=00065&period=P7D>
- Wang, W., Liu, J., Li, C., Liu, Y., & Yu, F. (2021). Data Assimilation for Rainfall-Runoff Prediction Based on Coupled Atmospheric-Hydrologic Systems with Variable Complexity. *Remote Sensing*, 13(4), 595. <https://doi.org/10.3390/rs13040595>