Sivarajah Mylevaganam

Alumnus, Spatial Sciences Laboratory, Texas A&M University, College Station, USA.

9)   Line 90-Line 94:

We **have some evidence** that there might be ways to construct training sets that could result in better models than simply training on all available streamflow data. We **do not have results that support** this directly.

These statements are not understood.

10)   Line 79-Line 84: Figure 5 shows how test period **performance increases as more basins are added to the training set.** Performance continues to increase up to the maximum size of the CAMELS data set (531 basins). In other words, even these 531 basins are most likely not enough to train optimal LSTM models for streamflow.
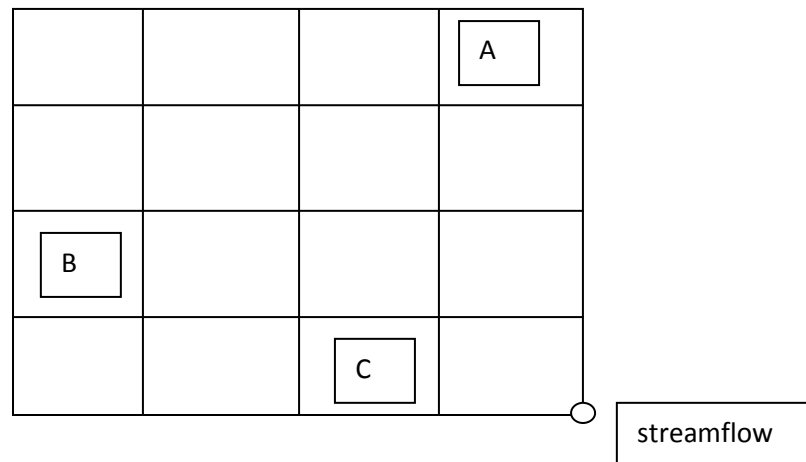
In my opinion, these statements do not make sense. To better understand your methodology, assume that we have 16 basins (not 531 basins as in your analysis) as shown in the figure in the attached PDF file. For simplicity, let us forget the shapes of the basins. Moreover, assume that you have added the basins A, B, and C to have a training set to derive the associated NSE. Likewise you add more basins to your training set to have an array of NSE values to show a plot like the one that you have shown in Fig.5. Based on this figure, is it meaningful to conclude that the **performance of the model increases as more basins are added to the training set?** From a hydrological point of view, does it make sense to have a streamflow that is sourced by basins A, B, and C? Would the stream network that sources the flow at a location of interest become discontinuous?

The NSE value that you have reported for a training set size of 100 basins (see Fig.5) may not actually represent the hydrology although the reported NSE value is very close to the NSE value for a training set size of 531 basins that may represent the actual hydrology. In ML models, you need to understand the

theory that governs the system of equations. An in-depth understanding on the system of equations and how they are formulated will lead to understand the physics. Do the catchment attributes that you have chosen in your analysis play a role in the NSE values that you have reported in Fig.5?

Acknowledgement and Disclaimer