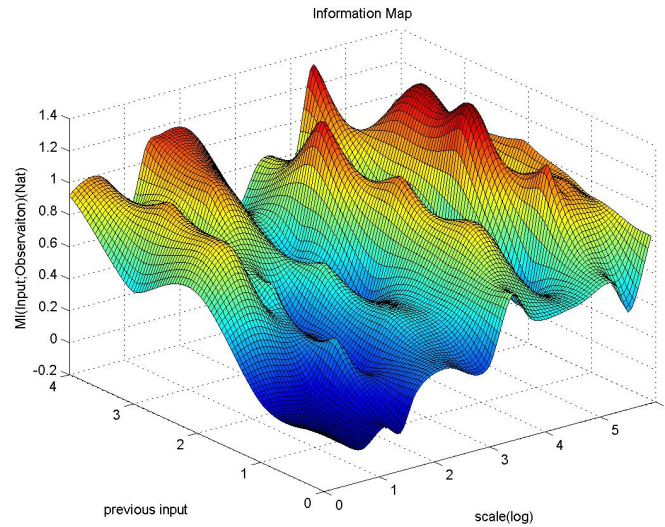


Difficulties In Hydrological Information Calculation

Pan Baoxiang

Research Clarification

The research we are doing is to exam the sufficiency of hydrological data for constructing conceptual hydrological models at different time scales. For example, we have distributed models at minute or hour scale, conceptual models like TOP model and Xinanjiang Model at hour or daily scale, water balance models at month scale and Budyko equation at annul time scale. We want to examine how these models co-exist at the border time scales, and how the iterative pattern (models at scales less than 1 year) transmits into non-iterative pattern (Budyko). The method we adapt is the epistemic aleatory uncertainty framework. Here is some primary result calculated for a experimental basin from MOPEX:



The blue area shows a deficiency of data provided for hydrological simulation, either due to ignorance of convergence or some auto-regressive pattern

of hydrological dynamics (I am not sure). We could tell that at a scale of e^4 days, the previous hydrological data remains little influence.

Problems

The puzzles that keeps me awake at night are from two aspects.

Scientific Considerations

The legality of Differential Entropy

Differential entropy could not represent the information content, which would cast great doubt on the definition of aleatory uncertainty.

Information terms without being normalized could not be compared with each other for simulation at different time scales.

Sample Re-cluster

The daily hydrological data for all the 54 year are not i.i.d for they fluctuates yearly, thus, to calculate the information terms, I have to choose the data from a same time of a year to form a sample space, which, not only reduces the sample number, but could not reflect the data adequacy along the whole time axis.

Technological Consideration

Though the result above is calculated with the method of ICA as recommended in Gong's paper, I doubt its legality. In the general formulation of ICA, the purpose is to transform an observed random vector X linearly into a random vector Y whose components are statistically as independent from each other as possible(Hyvarinen,1997), thus, the method would overrate the entropy for using $H(y_i)$ to represent $H(y_i|y_{rest})$ because the former is larger.

Furthermore, in application, the fast ICA method I encoded could not be implemented for a large part of the hydrological series.

Possible Solutions

I have been thinking about possible solutions for long, but with little achievements. I am trying to calculate the mutual information using the

method brought forward by Kraskov(2008). Could a fourier transform help to solve the problem of seasonal fluctuation? I have just been learning harmonic analysis and could not tell.

Sorry to take your time and hoping for some suggestions.