# AN INFORMATION-THEORETIC APPROACH FOR EVALUATING CATCHMENT SCALE PROCESS RELATIONSHIPS

**Gowri Reghunath**[1] and Pradeep P. Mujumdar[1,2]

[1]Interdisciplinary Centre for Water Research, Indian Institute of Science, Bengaluru, India (gowri3041993@gmail.com)

[2]Department of Civil Engineering, Indian Institute of Science, Bengaluru, India



## 1. Introduction

- Hydrological responses of a catchment are governed by complex interactions of processes, and they exhibit non-linear behaviour at all scales.
- To have a detailed understanding of the catchment behaviour, it becomes imperative to explore its components and processes and identify their intricate relationships.
- Information-theoretic (IT) measures can help in developing deeper insights into the hydrological process relationships.

## 2. Research Objectives

- To disentangle hydrological process relationships and capture the non-linear catchment behaviour by applying IT measures.
- To identify the significant catchment attributes affecting various hydrological processes and their relationships which can help in building better modelling strategies.

## 3. Study Area & Metholodogy

- Study Area: Cauvery River Basin, Peninsular India.
- Basin area: approx. 85600 sq. kms.
- Major land use clasee is agriculture.
- Annual average rainfall varies from 500 to 3000 mm.
- Major soil types are black, red, lateritic and alluvial soils.
- Large number of small and large scale interventions.
- Large scale shift in land use and land cover over the past decades.

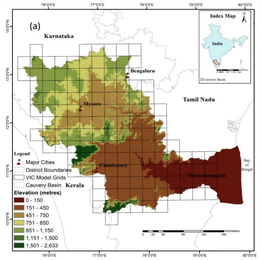


Fig 1: Study Area – Cauvery River Basin, India

### Methodology

Grid-based Variable Infiltration Capacity (VIC) model is employed at a spatial resolution of 0.25 x 0.25-degree over the study area at a daily time scale.

| | |
|---|---|
| Estimate | Normalized Shannon Entropy for model-simulated water balance components |
| ↓ | |
| Estimate | Mutual information (MI) and MI-based non-linear correlations between all the combinations of water balance components to capture process relationships |
| ↓ | |
| Identify | Influence of catchment attributes on IT measures by a correlation analysis |

## 4. Results: Information-Theoretic Measures

Shannon Entropy for a variable X: $H(X) = -\sum_x p(x) \log p(x)$

Mutual Information for X, Y: $I(X,Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$

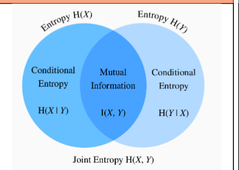


Fig 2: Relationship of MI with joint entropy and conditional entropy. Source: www.d2l.ai.

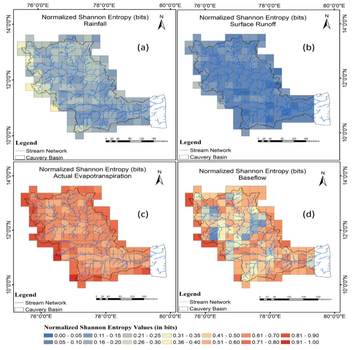


Fig 3: Spatial variation of Normalized Shannon Entropy (H) [unit: bits] computed for Rainfall (P) and Surface Runoff (Q), Actual Evapotranspiration (AET), and Baseflow (BF), respectively.

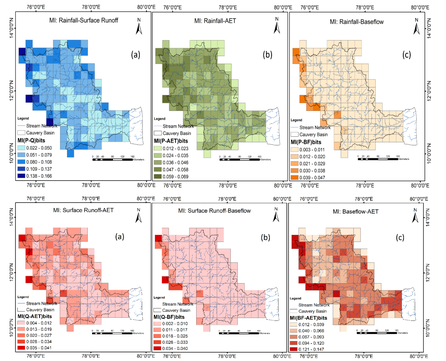


Fig 4: Spatial variations of Mutual Information values. Border grids on the western side of the basin show higher MI values.

- Maximum information content is shared between rainfall and surface runoff.
- Surface runoff becomes highly predictable through the knowledge of input rainfall.
- BF and AET share a high information content - stronger influence of land use and geological features on the AET and baseflow processes.

**Are conventional correlation measures adequate in evaluating process relationships?**

- Non-Linear Correlation Coefficient – Joe's Formula

$$NLCC = \sqrt{\left(1 - \exp(-2I)\right)}$$

- P and Q have high linear correlations; other pairs of variables exhibit stronger non-linear dependencies.
- Stronger non-linear associations are found in regions with higher altitudes.

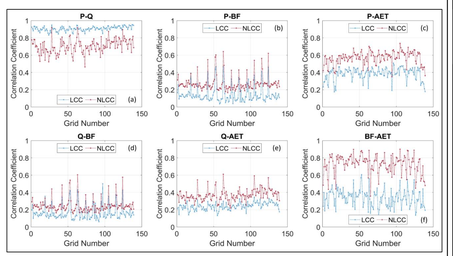


Fig 5: Linear vs Non-linear Correlations

## 5. Results: Correlations to Catchment Attributes

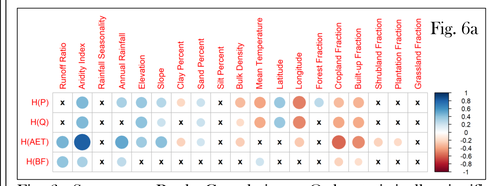


Fig. 6a

Fig 6: Spearman Rank Correlations. Only statistically significant values ($p < 0.05$) are shown. 'x' mark represents statistically insignificant values.

IT measures can be considered as a substitute for many of the catchment co-variates.

- There can be complex interdependencies between catchment attributes and various hydrological processes.
- Attributes related to climate, soil and land use are selected.

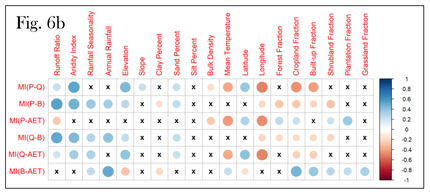


Fig. 6b

- These observations can be insightful in determining the adequacy of the model structure to simulate other hydrological fluxes other than surface runoff in a regionally complex catchment.

## 6. Concluding Remarks

- Information theory serves as a powerful tool in computing the information content of a variable as well as the amount of information one variable provides about another.
- IT measures can be applied to both linearly and non-linearly related variables and do not require prior assumptions on underlying distributions.
- Mutual information captures non-linear dependence and is advantageous over conventional statistical measures.
- For developing further insights into the complex catchment behaviour, the application of higher dimensional and multi-variate information measures can to be explored.

## 7. Acknowledgements

The authors express their gratitude for the financial support received from the National Supercomputing Mission (NSM) project "Urban Modelling: Development of Multi-sectorial Simulation Lab and Science-Based Decision Support Framework to Address Urban Environmental Issues (Project Sanction Number: MeitY/R&D/HPC/2(1)/2014))" through the Department of Sciences and Technology (DST) and the Ministry of Electronics and Information Technology (MeitY), which enabled their participation in EGU2023. The authors also acknowledge the support from DST and MeitY.