



Linear and Nonlinear Mutual Information

- References:

- Smith R., 2015: A mutual information approach to calculating nonlinearity, *ISI J. Rapid Dissem. Stat. Res.*, doi: [10.1002/sta4.96](https://doi.org/10.1002/sta4.96).
- Hsu, H., and P. A. Dirmeyer, 2021: Nonlinearity and multivariate dependencies in land-atmosphere coupling. *Water Resour. Res.*, **57**, e2020WR028179, doi: [10.1029/2020WR028179](https://doi.org/10.1029/2020WR028179).

- Principle:

- Total mutual information (MI) between one or more source variables X_s and a target variable X_{tar} can be partitioned into linear and nonlinear parts. For a given set of X_s and X_{tar} , the procedure is:
 1. Fit a linear regression model in terms of predicting X_{tar} given X_s ;
 2. Obtain \hat{X}_{tar} as the fitted values of X_{tar} and define the nonlinear residual $X'_{tar} = X_{tar} - \hat{X}_{tar}$;
 3. Normalize X'_{tar} by the quantile normalization based on the value of X_{tar} (quantile normalization makes the distribution of X'_{tar} and X_{tar} identical in statistical properties; see Smith (2015) for a detailed discussion);
 4. Estimate the MI for both $I(X_s; X_{tar})$ and $I(X_s; X'_{tar})$. The quantity $I(X_s; X'_{tar})$ is the nonlinear dependency between X_s and X_{tar} . The linear dependency in terms of MI is the difference (total minus nonlinear): $I(X_s; X_{tar}) - I(X_s; X'_{tar})$.

- Data needs:

- As with MI and MMI, there is an underlying assumption that the terms are physically related. Thus, comparison of results among models or between models and observational data sets can be a powerful diagnostic tool.
- Can also be extended to other information theoretical metrics to decompose them into linear and nonlinear components, e.g., for temporal information partitioning (TIP; see: Hsu and Dirmeyer 2021).

- Observational data sources:

- Well suited to data where multiple components of the LoCo process chain are measured at the same location, such as flux tower data.

- Caveats:

- Causality between variables is not assured, particularly if all variables are synchronous in time.