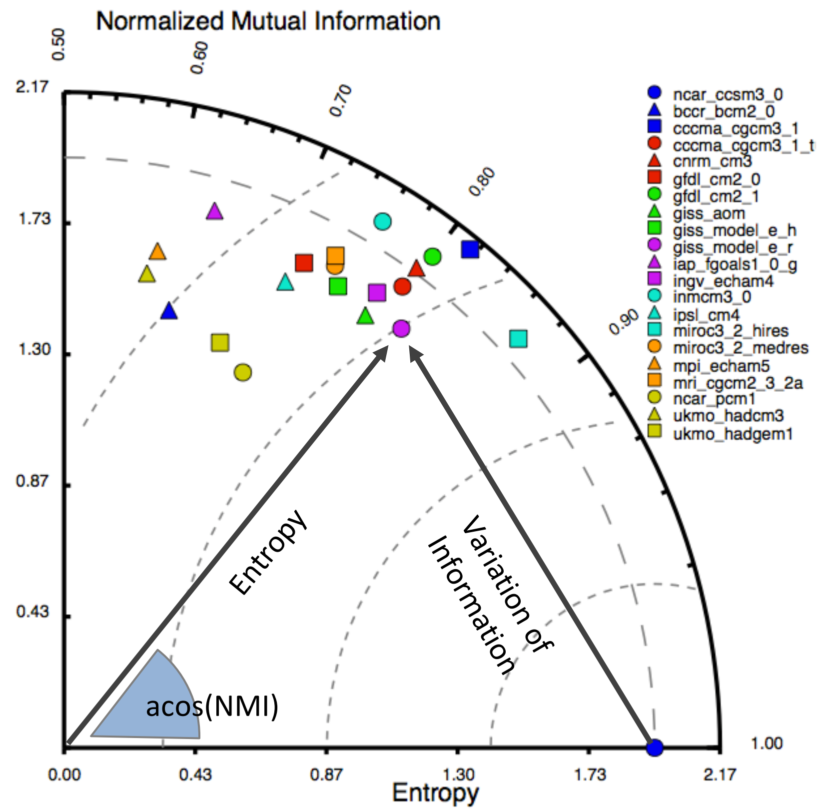
The
Mutual Information Diagram
for
Uncertainty Visualization

Carlos D. Correa and Peter Lindstrom

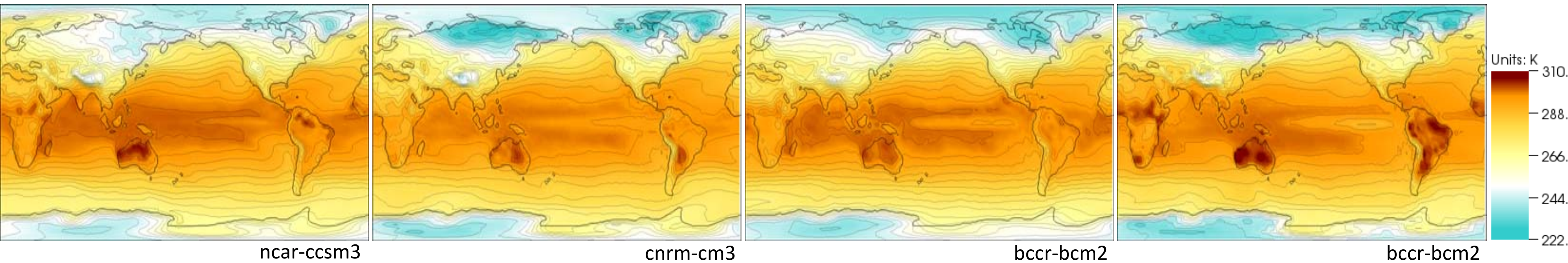
Center for Applied Scientific Computing - CASC

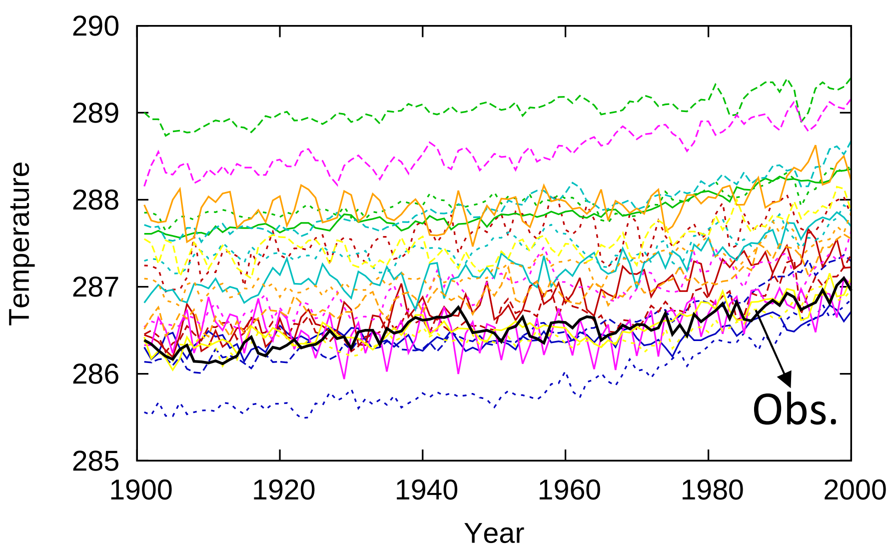
Lawrence Livermore National Lab

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.



Visualization should help compare models with observations



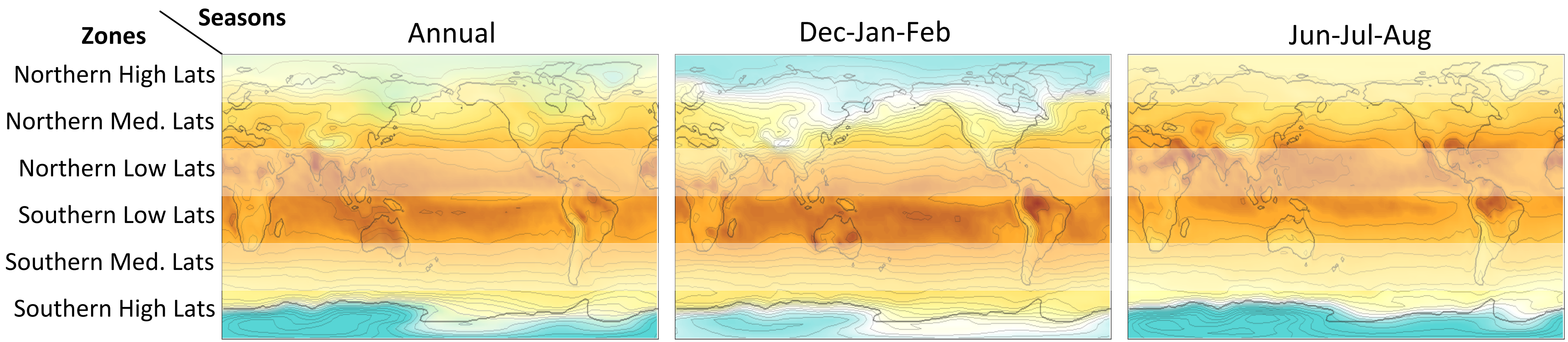


Average annual temperature 1900-2000 as predicted by various climate models.

Which model is more **similar** to a reference model or observations?

Trend plots often do not expose these aspects.

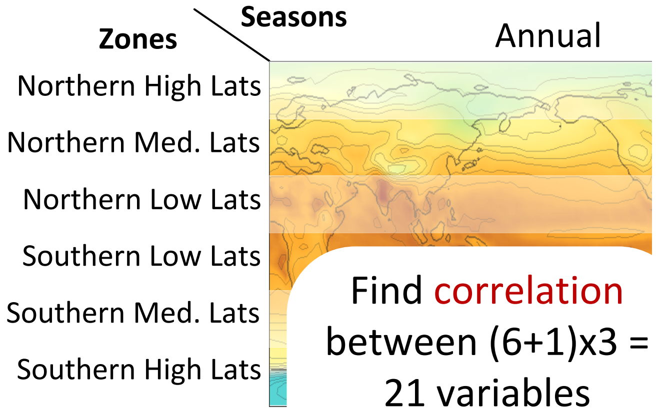
Visualization should help find correlations of similar outputs – important for uncertainty quantification



Divide ensembles in 6 latitude zones and 3 temporal averages

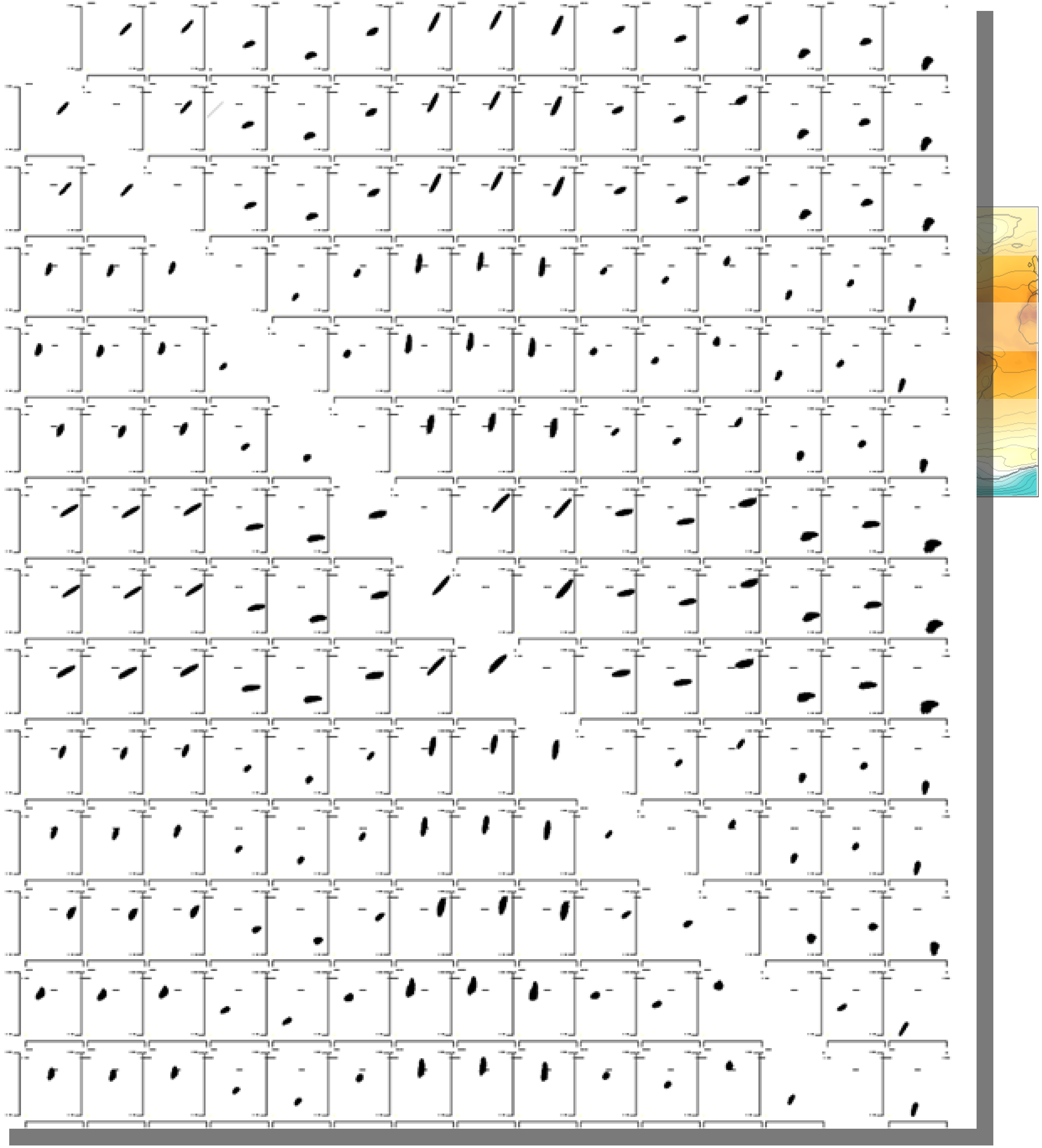
- Are there correlations across seasons or latitudes?
- Are there large discrepancies in the different outputs?

Visualization shoul
importa

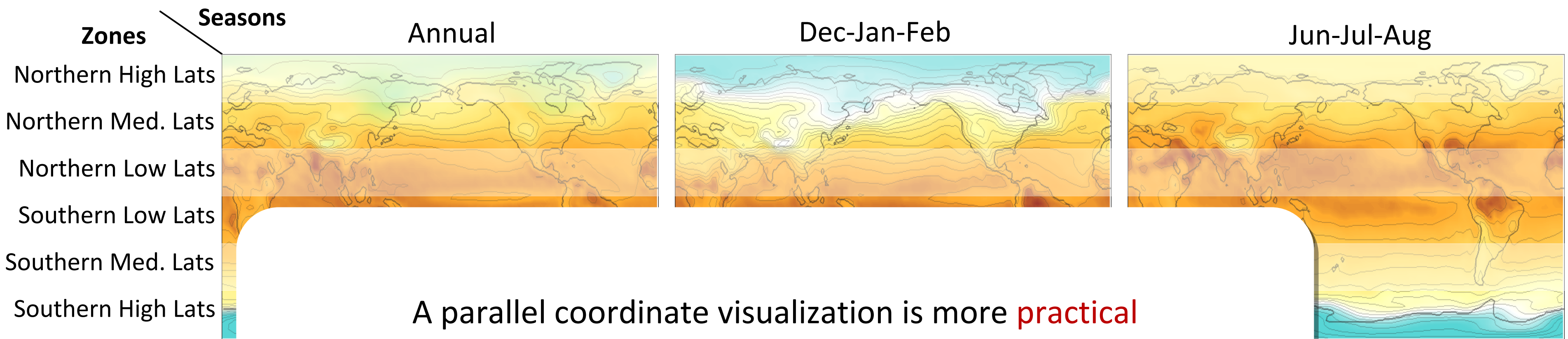


Find correlation between (6+1)x3 = 21 variables

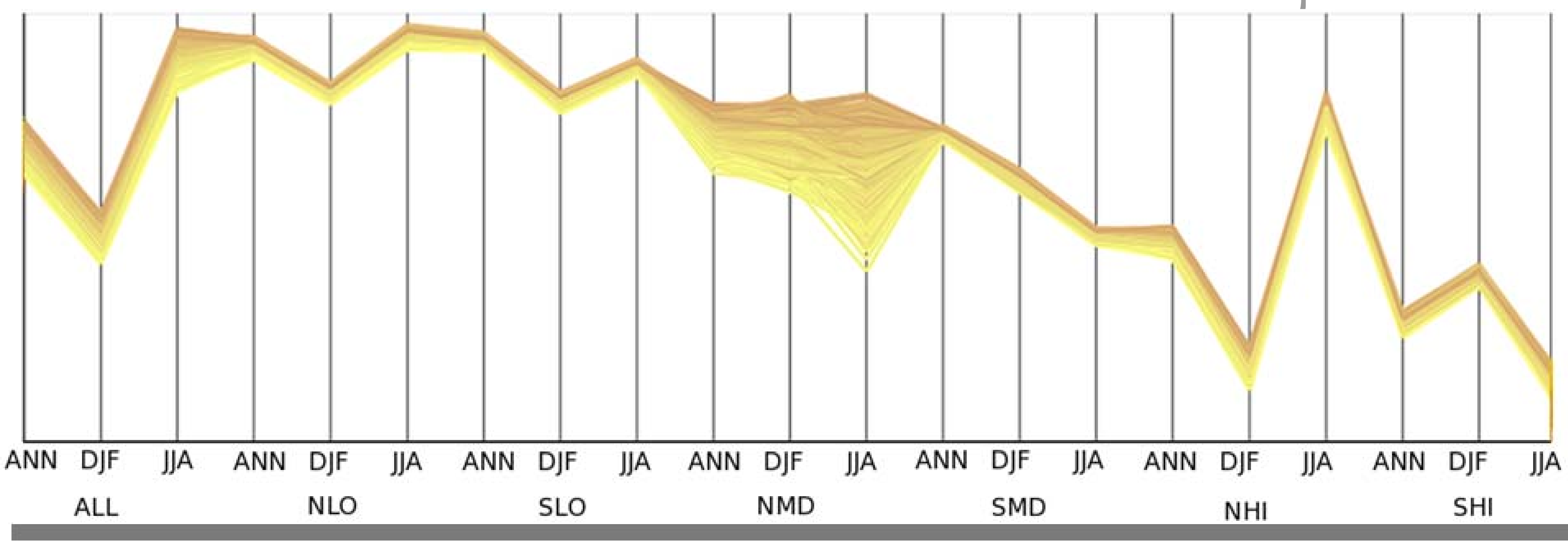
A scatterplot matrix becomes impractical for many outputs



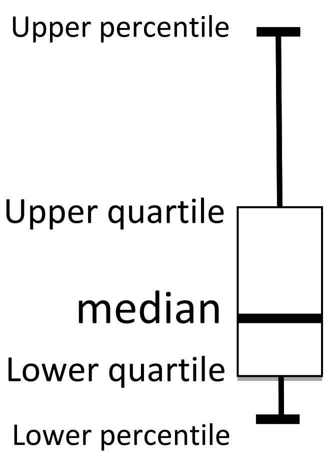
Visualization should help find correlations of similar outputs – important for uncertainty quantification

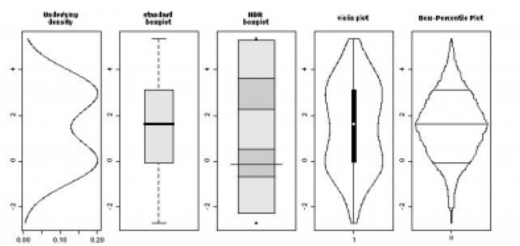


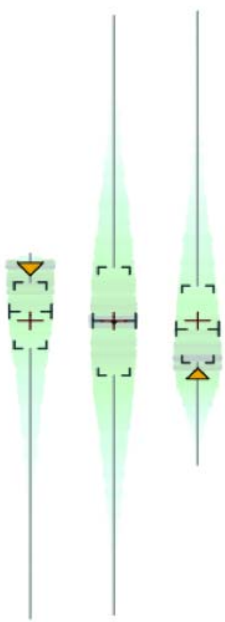
A parallel coordinate visualization is more practical
But only certain pairwise comparisons are possible

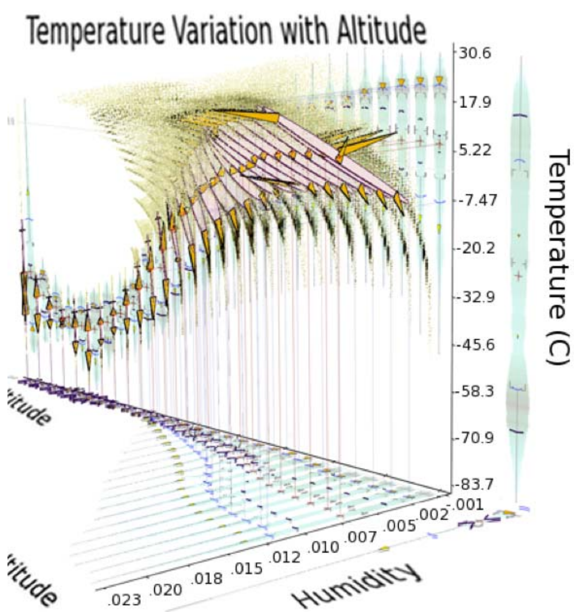


Visual Summaries





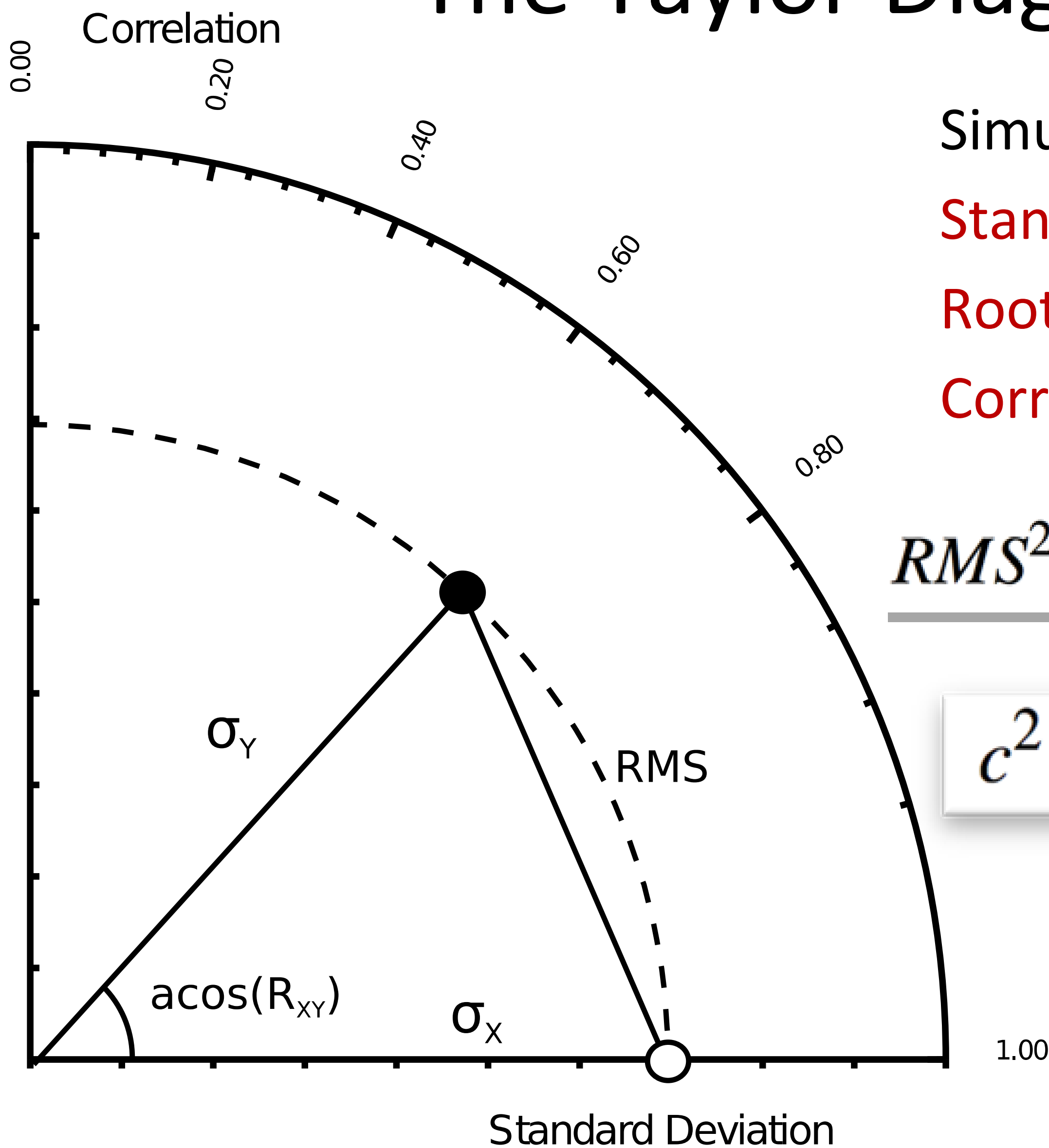




Potter et al. (Eurovis'2010)

- Represent directly summary quantities, e.g., mean, standard deviation, entropy.
- Box-plots and their many variants
- One plot per ensemble may result in clutter
- Visualizing several statistics simultaneously in a metric space: Taylor diagram

The Taylor Diagram



Simultaneously plots

Standard deviation,

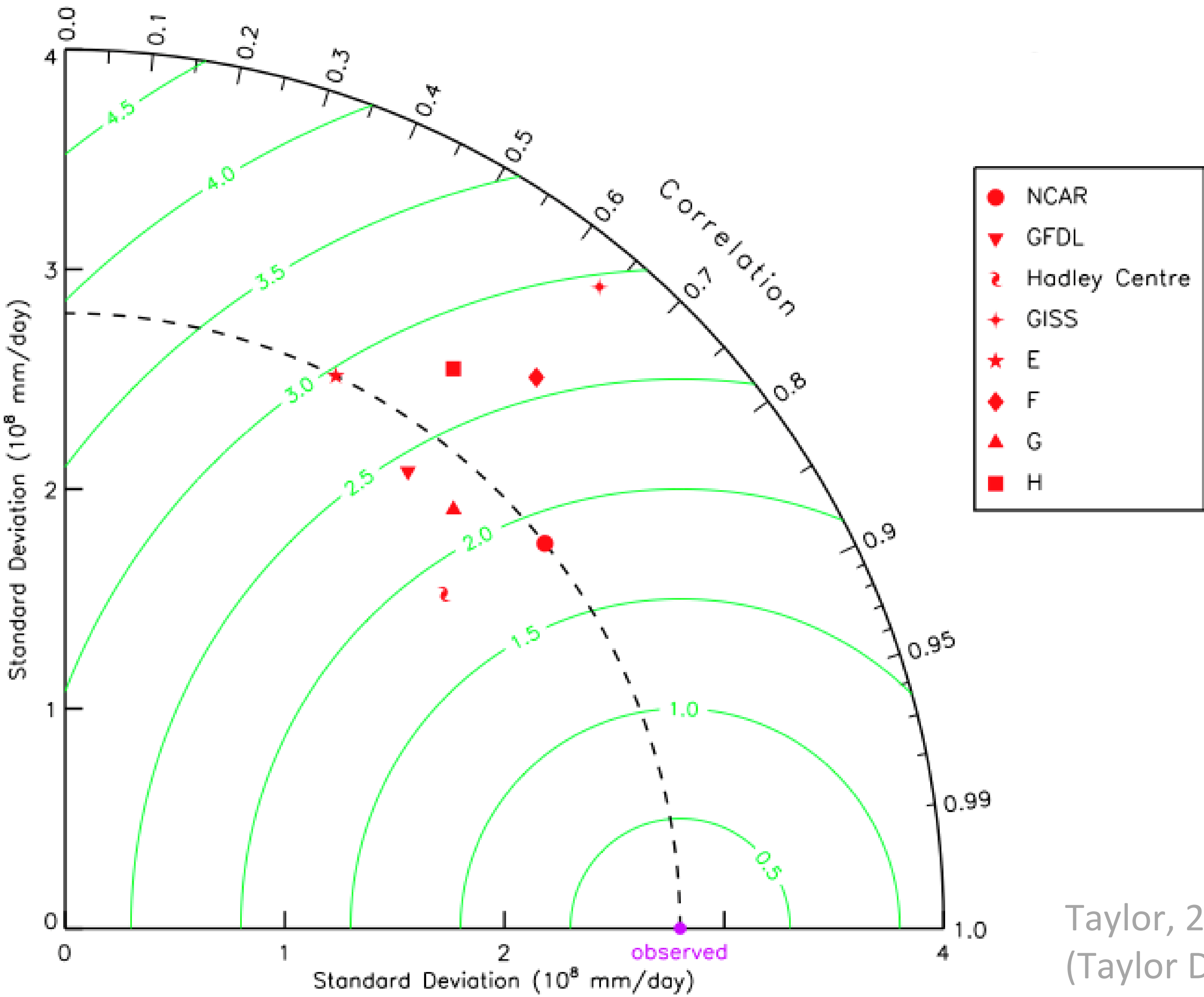
Root Mean Square Error and

Correlation R.

$$RMS^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_X \sigma_Y R_{XY}$$

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

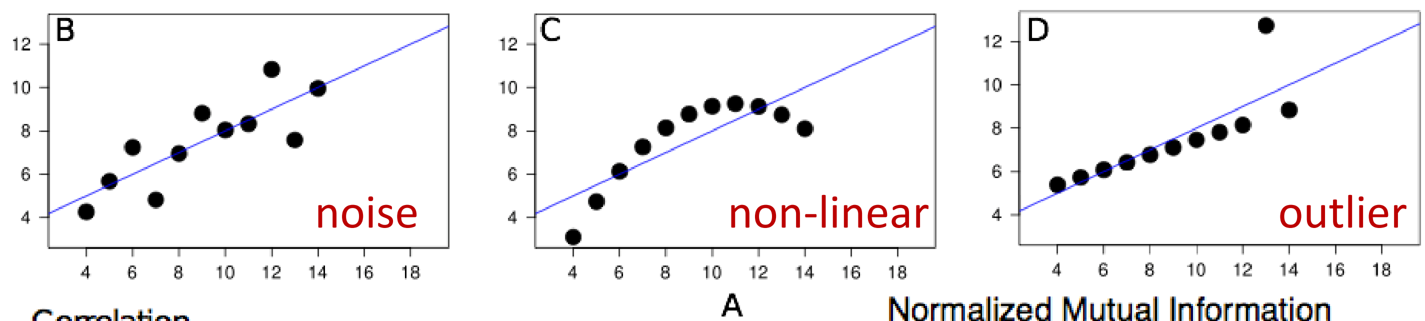
Applications of the Taylor Diagram

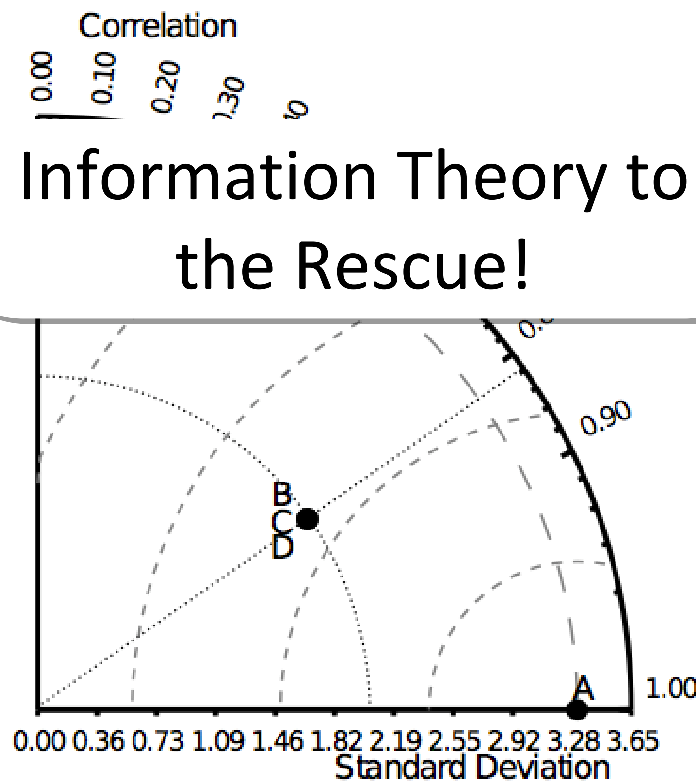


Taylor, 2005
(Taylor Diagram Primer)

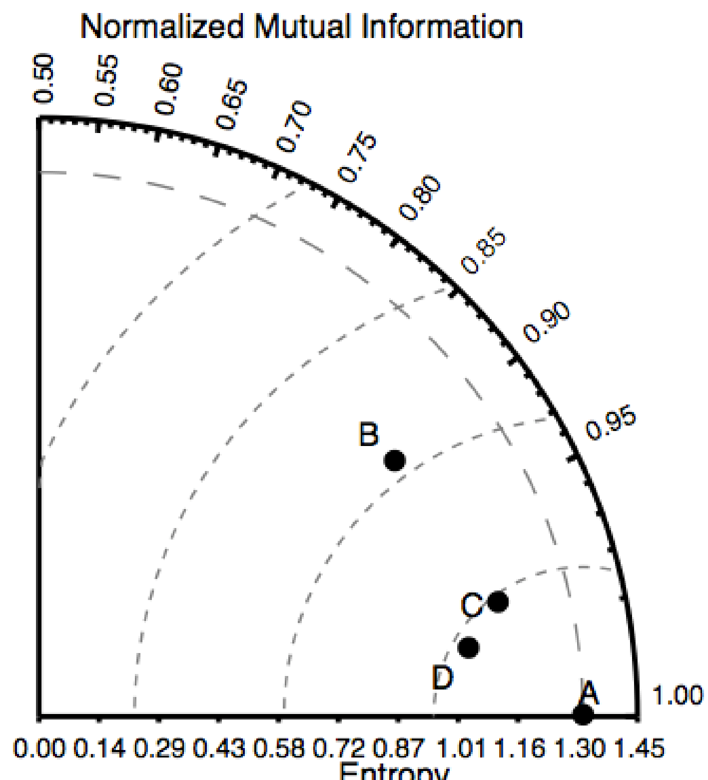
Anscombe's Trio

Variables B,C,D: same standard deviation and same correlation w.r.t. A

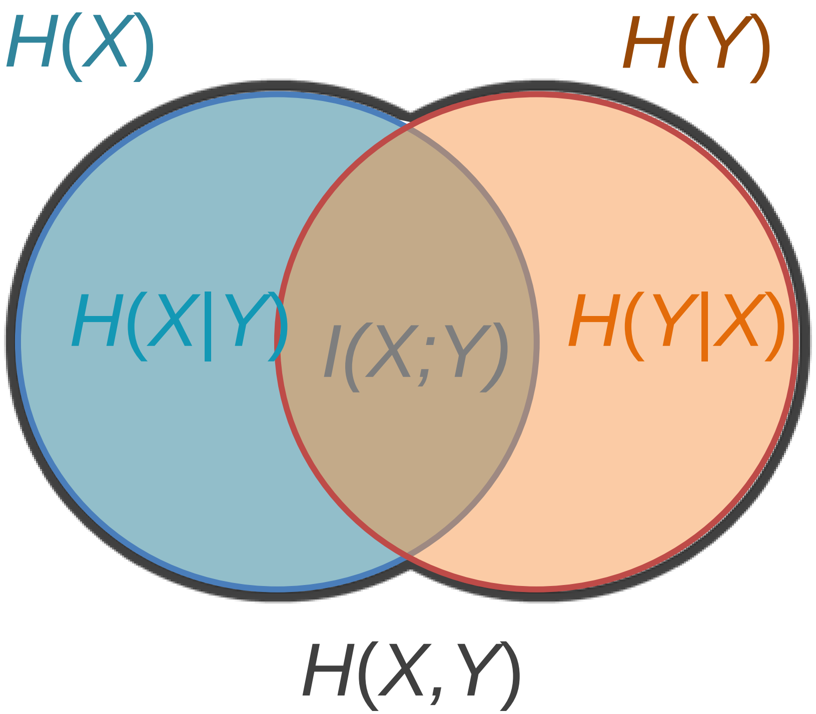




Information Theory to the Rescue!



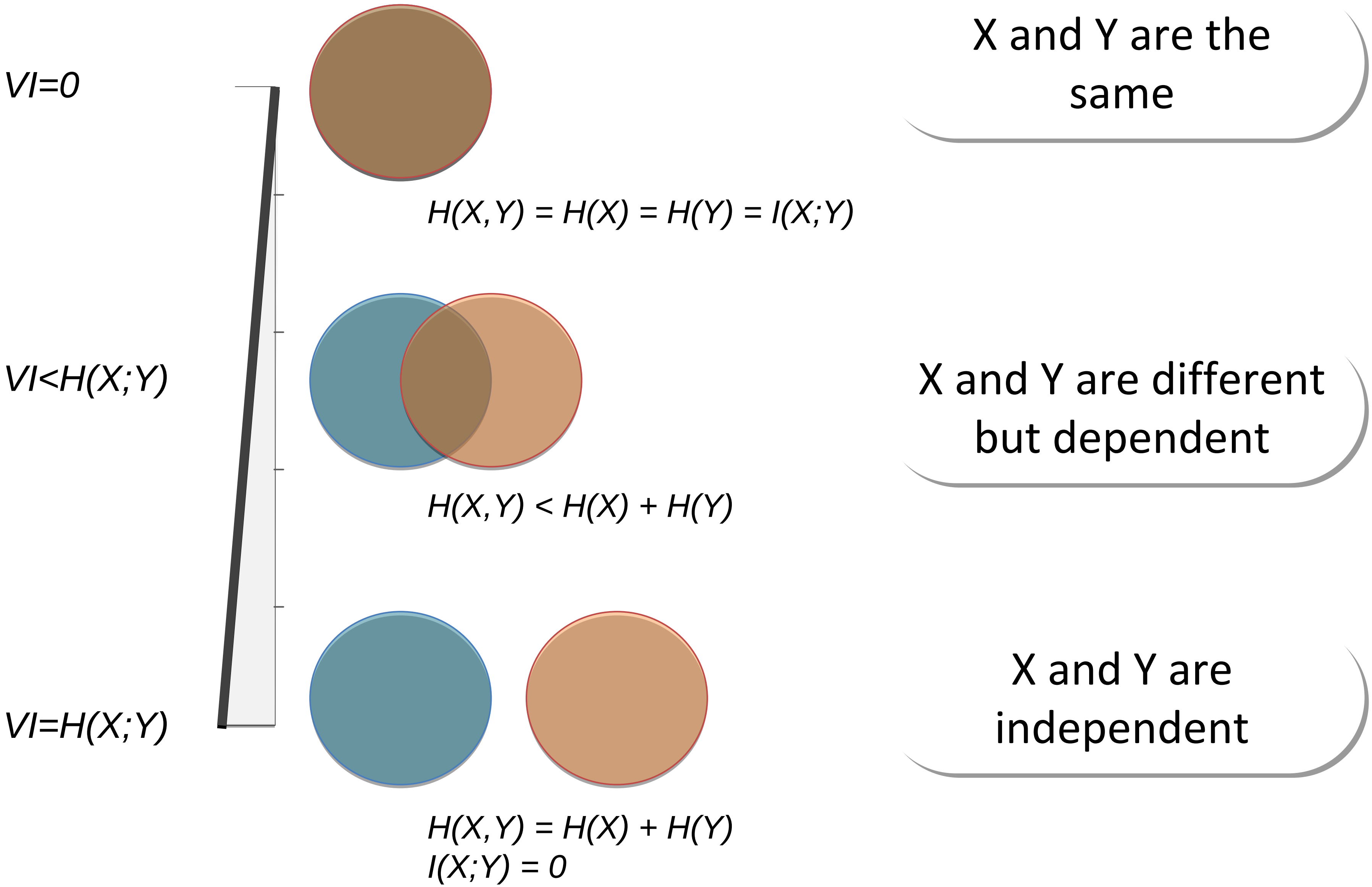
Information Theory Primer



$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

- Entropy H(X)
  - Measure of information uncertainty of X
- Joint Entropy H(X,Y)
  - Uncertainty of X,Y
- Conditional Entropy H(X|Y)
  - Uncertainty of X given that I know Y
- Mutual Information I(X;Y)
  - How much knowing X reduces the uncertainty of Y

Variation of Information $VI = H(X \mid Y) + H(Y \mid X)$



The Variation of Information VI: a measure of distance in information theory

$$VI(X,Y) = H(X) + H(Y) - 2I(X;Y)$$

$$RVI = \sqrt{VI}$$

$$h_X = \sqrt{H(X)}$$

$$h_Y = \sqrt{H(Y)}$$

$$RVI(X,Y)^2 = h_X^2 + h_Y^2 - 2I(X;Y)$$

$$\boxed{RVI(X,Y)^2 = h_X^2 + h_Y^2 - 2h_X h_Y \frac{I(X;Y)}{h_X h_Y}}$$

The Variation of Information VI: a measure of distance in information theory

$$VI(X,Y) = H(X) + H(Y) - 2I(X;Y)$$

$RVI = \sqrt{VI}$

$h_X = \sqrt{H(X)}$

$h_Y = \sqrt{H(Y)}$

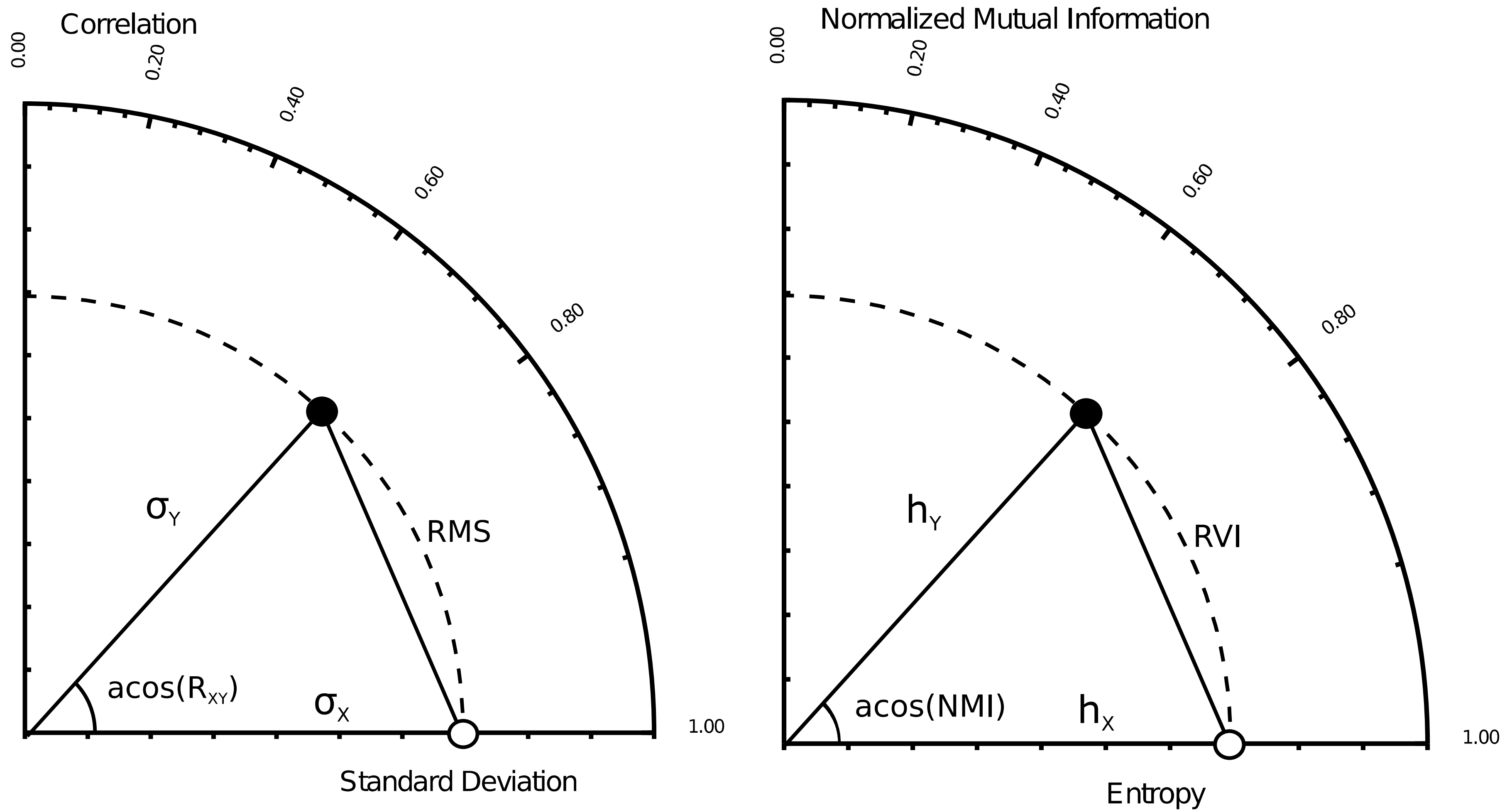
$$RVI(X,Y)^2 = h_X^2 + h_Y^2 - 2I(X;Y)$$

$$RVI(X,Y)^2 = h_X^2 + h_Y^2 - 2h_X h_Y \boxed{\frac{I(X;Y)}{h_X h_Y}}$$

Normalized Mutual Information (NMI)

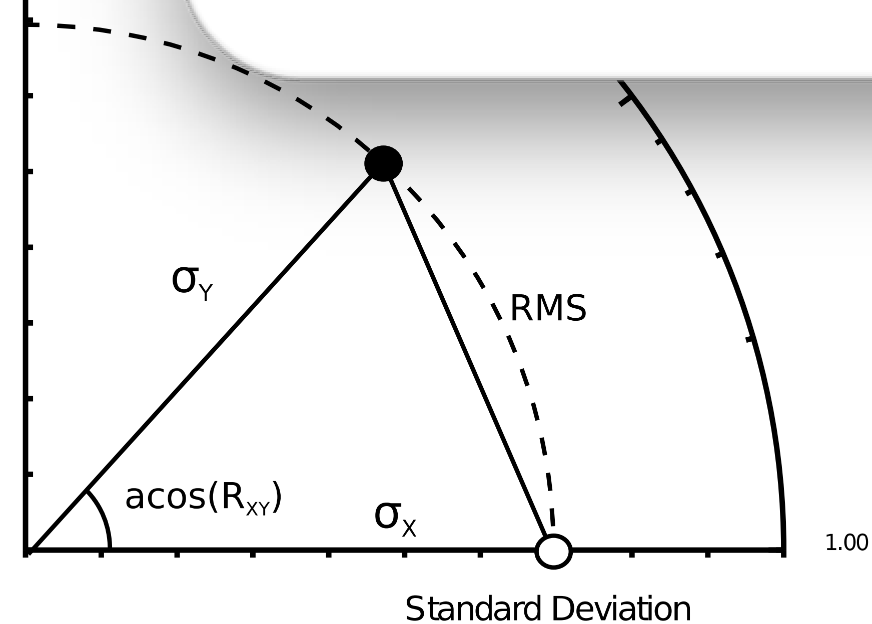
$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

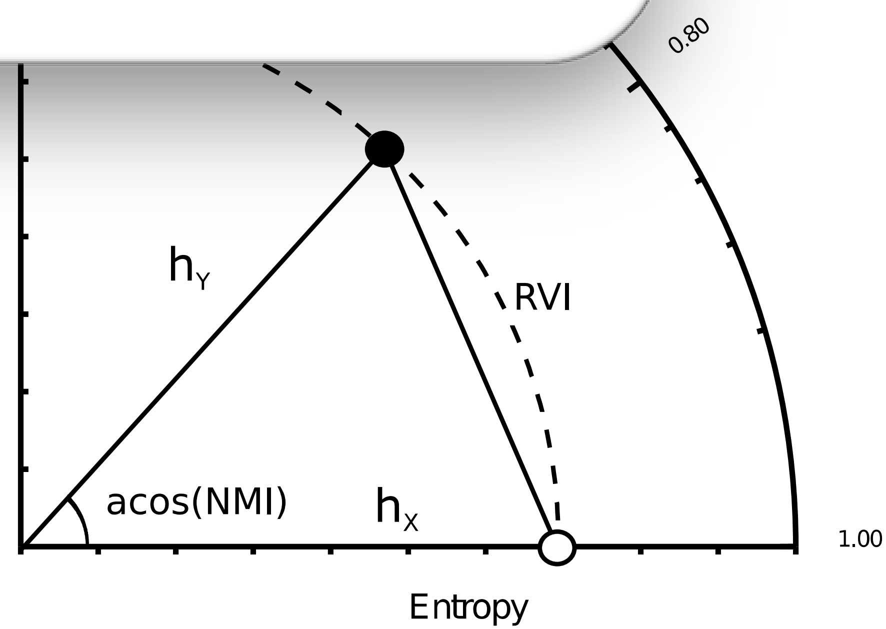
RVI Diagram


Correlation
0.00
0.20
0.40
0.60
0.80
1.00
σ_Y
RMS
acos(R_{XY})
σ_X
Standard Deviation
Normalized Mutual Information
0.00
0.20
0.40
0.60
0.80
1.00
h_Y
RVI
acos(NMI)
h_X
Entropy

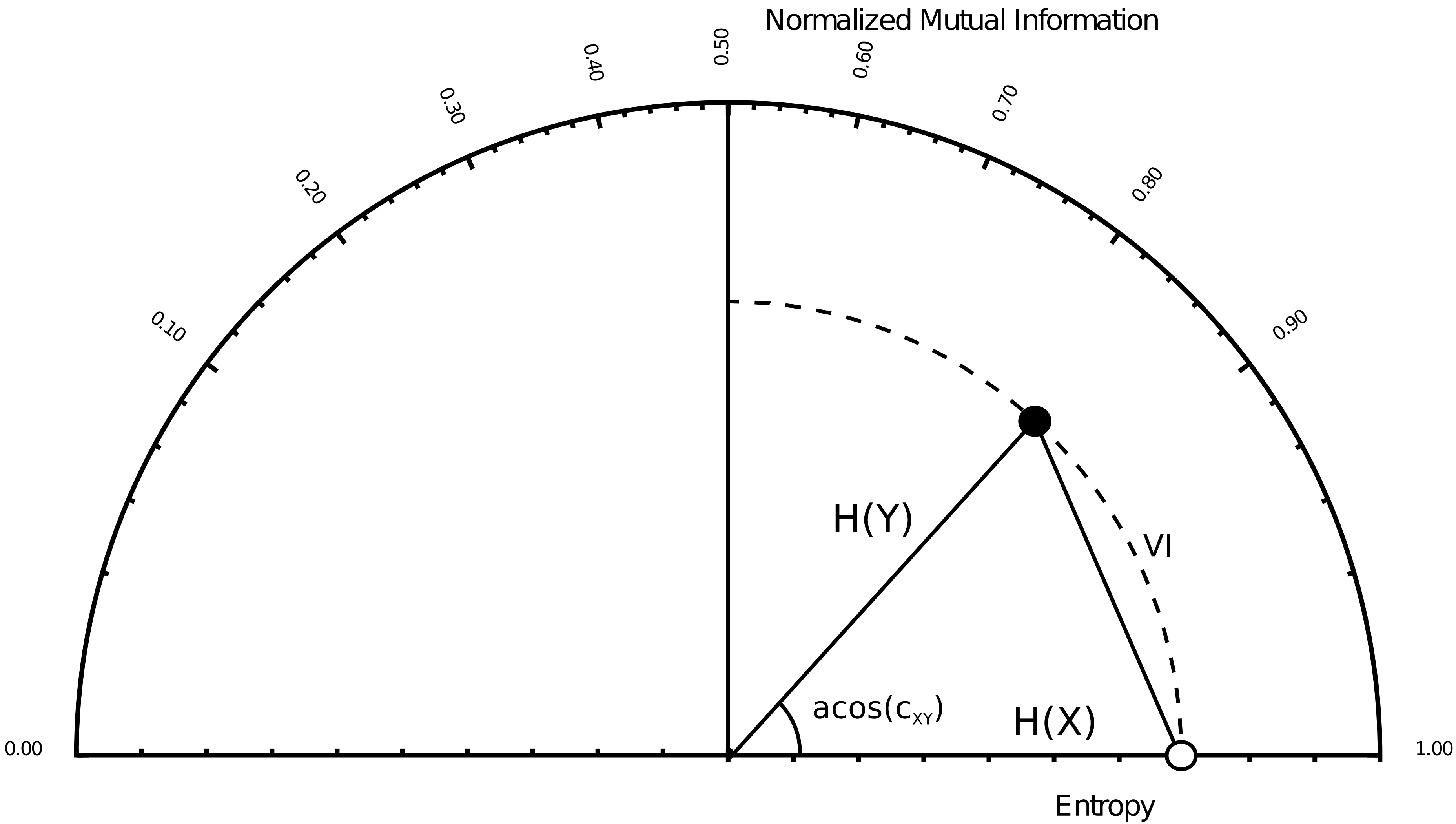
Equivalences

Statistics	$\Longleftrightarrow$	Information Theory
$\mathrm{RMS}(X,Y)$	$\Longleftrightarrow$	RVI $\sqrt{VI(X,Y)}$
Variance σ_X^2	$\Longleftrightarrow$	entropy $H(X)$
Covariance $cov(X,Y)$	$\Longleftrightarrow$	mutual information $I(X;Y)$
Correlation $R_{XY} = \frac{cov(X,Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}}$	$\Longleftrightarrow$	$NMI_{XY} = \frac{I(X;Y)}{\sqrt{H(X)H(Y)}}$



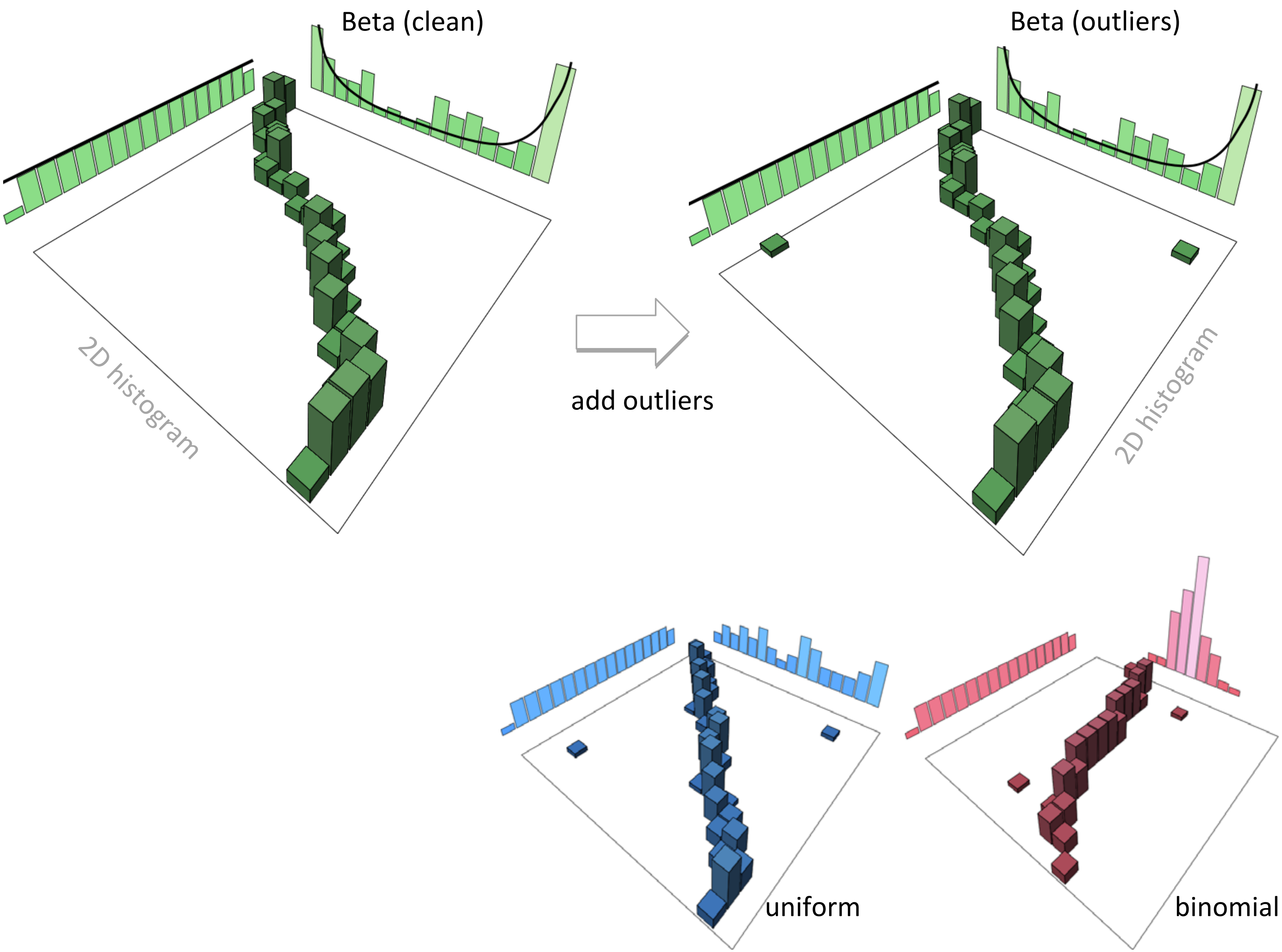


VI Diagram



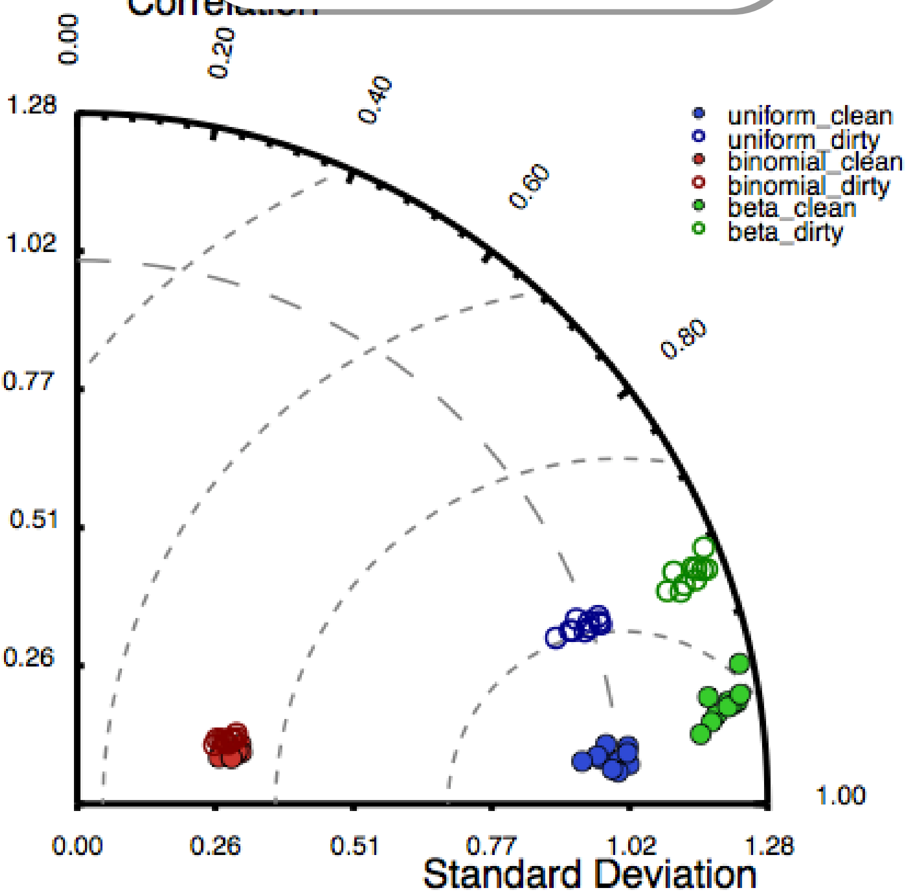
$$VI(X,Y)^2 = H(X)^2 + H(Y)^2 - 2H(X)H(Y)c_{XY}$$

Experiment of 2D distributions with outliers

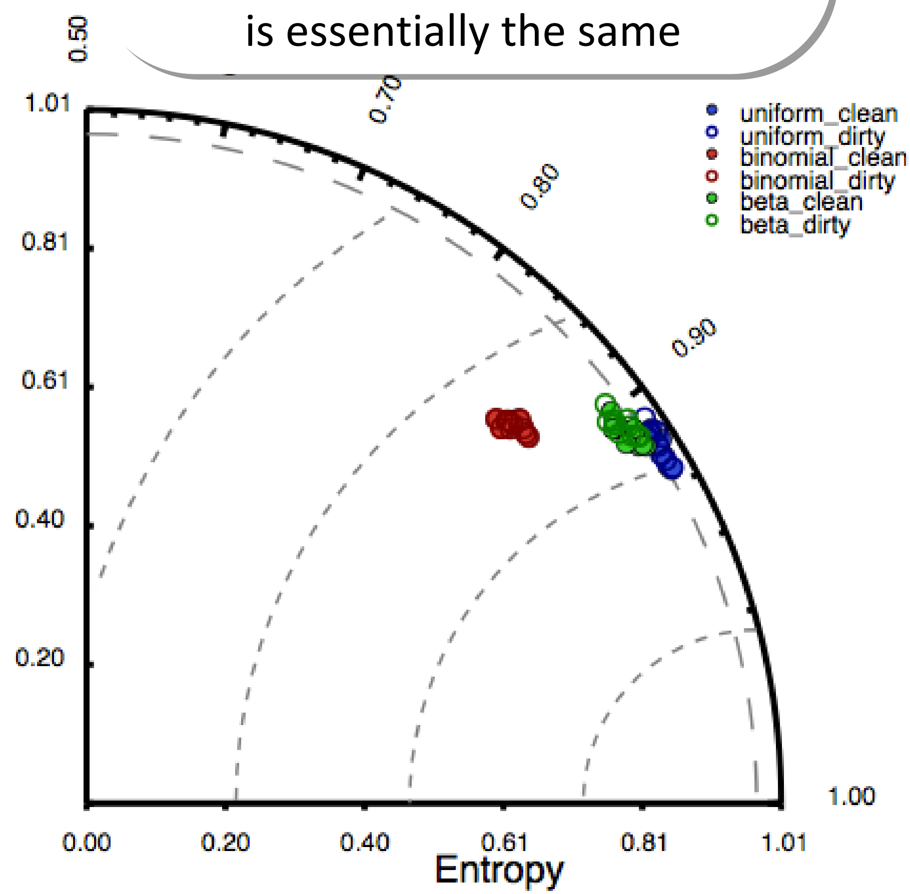


MI diagram is more resilient to outliers

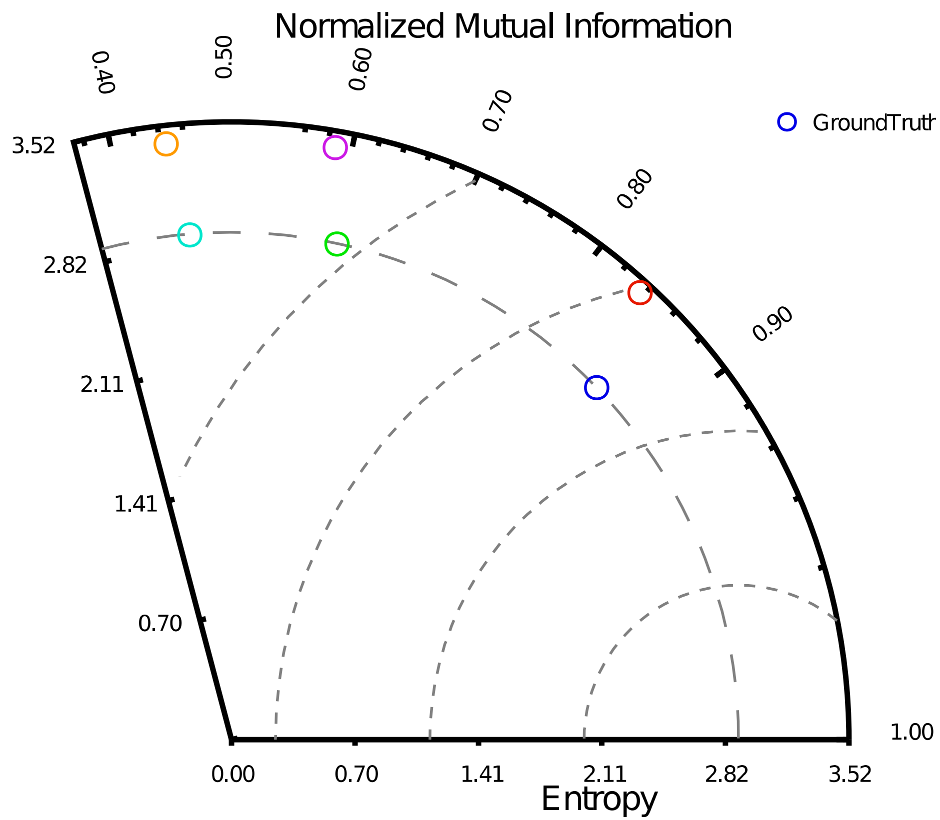
Outliers have a significant impact on correlation

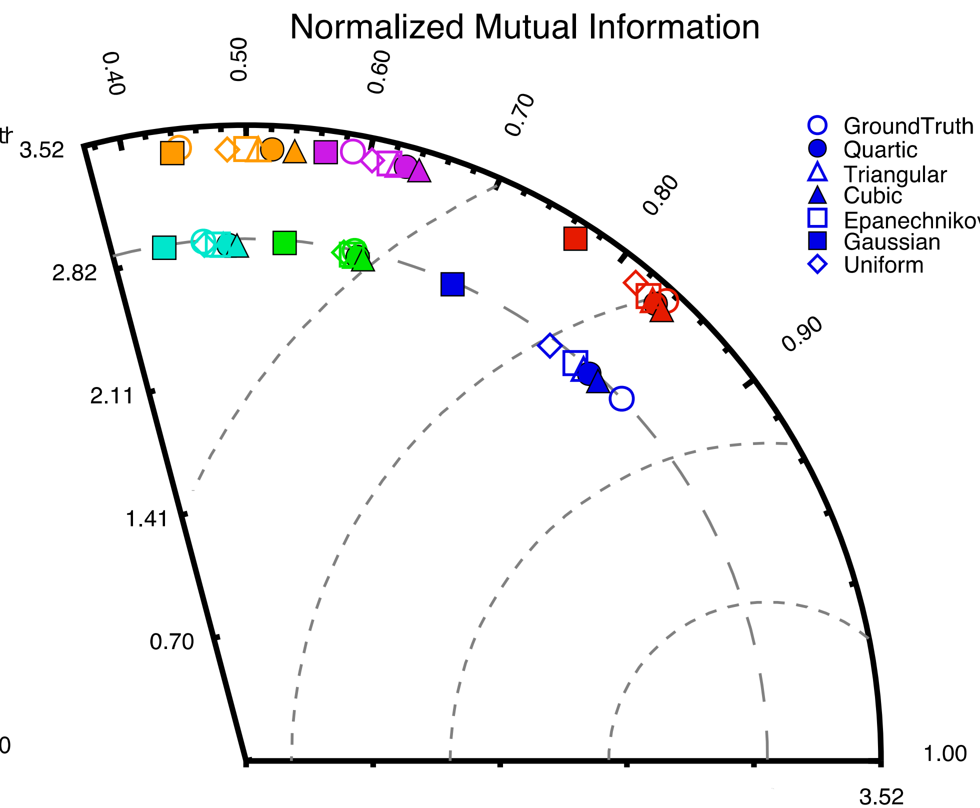


The information in both the "clean" and "dirty" distributions is essentially the same



Computing Entropy and Mutual Information *may* require estimation of underlying probability functions





N(sx=1.0, sy=0.5, R=0.99}

N(sx=1.0, sy=1.5, R=0.99}

N(sx=1.0, sy=0.5, R=0.95}

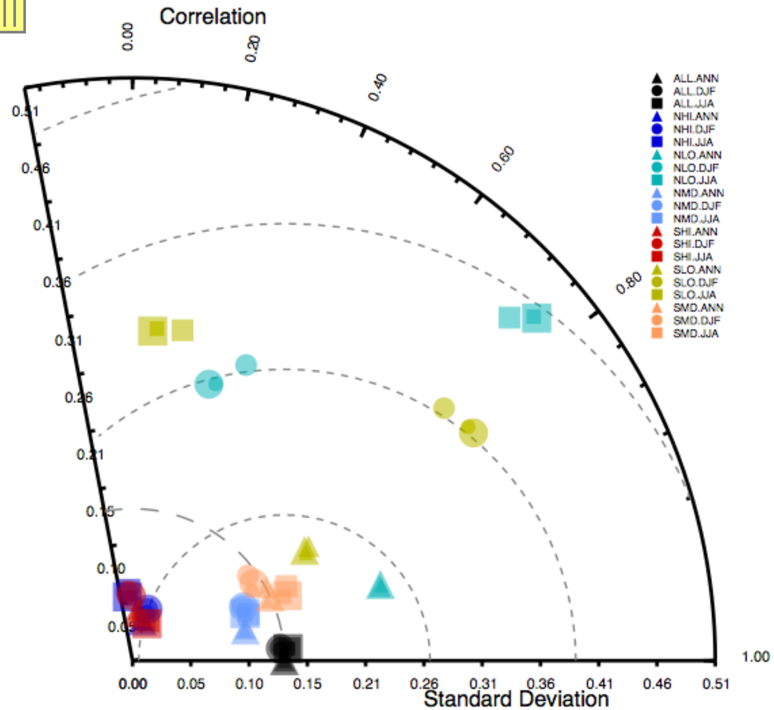
N(sx=1.0, sy=1.5, R=0.95}

N(sx=1.0, sy=0.5, R=0.90}

N(sx=1.0, sy=1.5, R=0.90}

Although there are differences, relative distances are consistent for each choice of kernel

Uncertainty Quantification in Climate Simulations

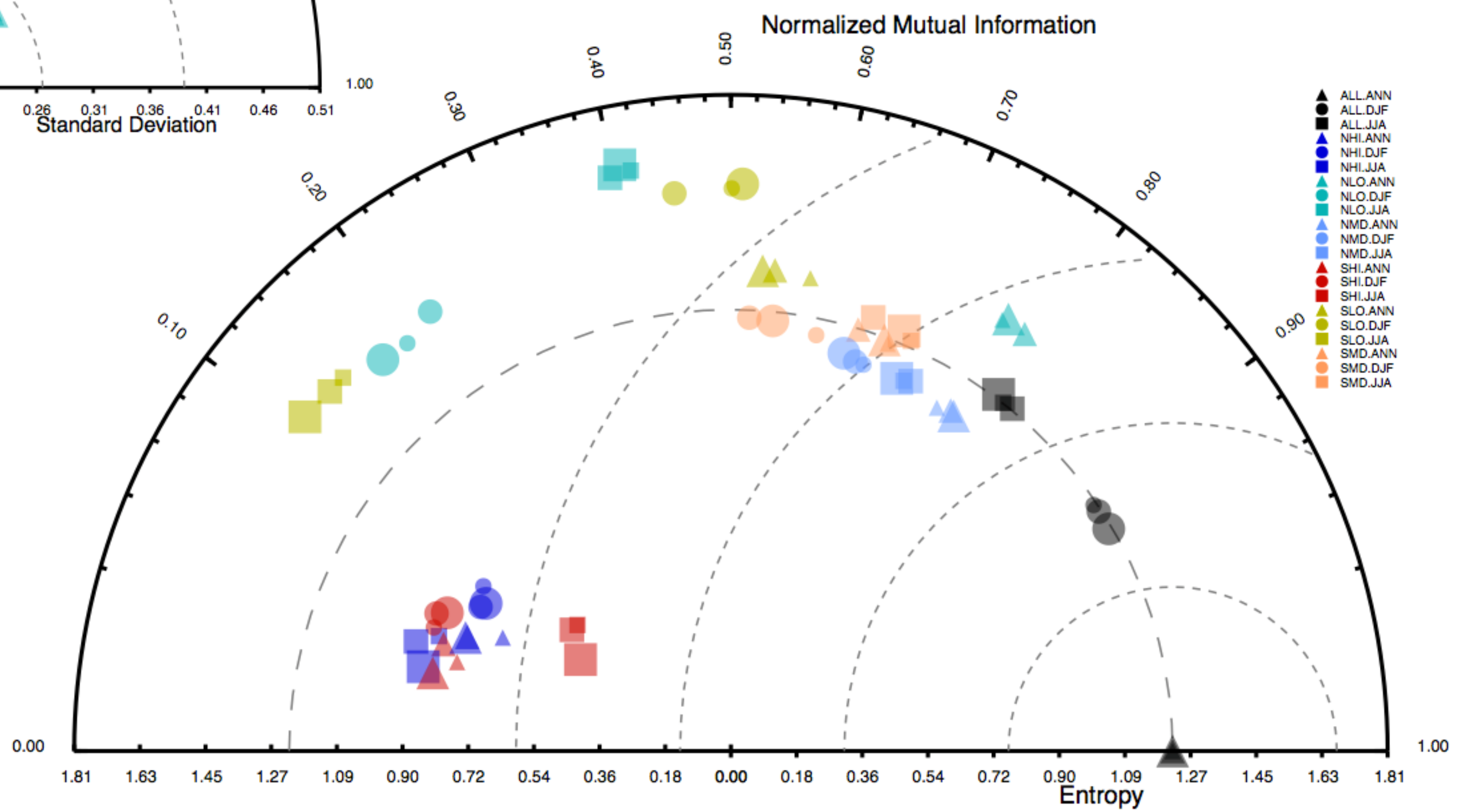


Precipitation average

7 Zonal averages (color)

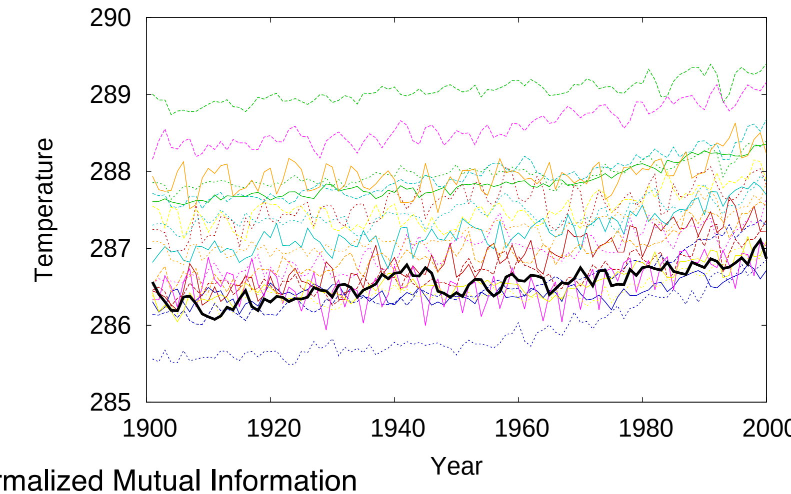
3 Temporal averages (shape)

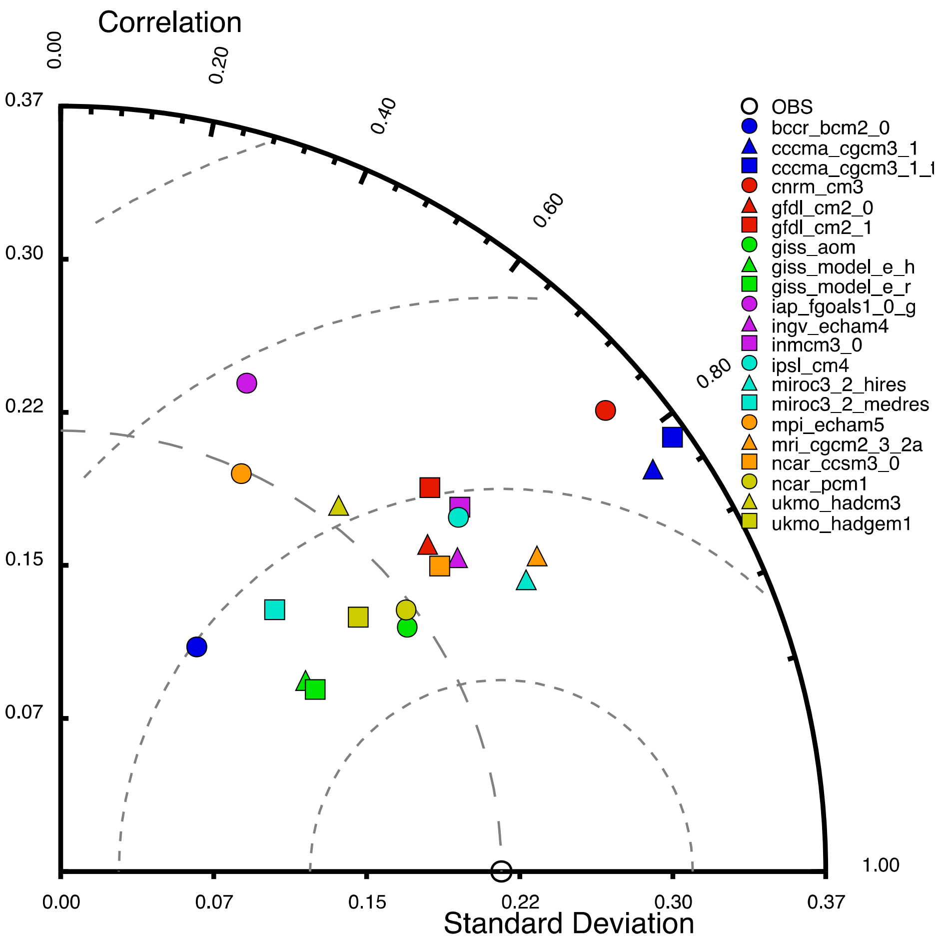
3 different ensemble sets (size)

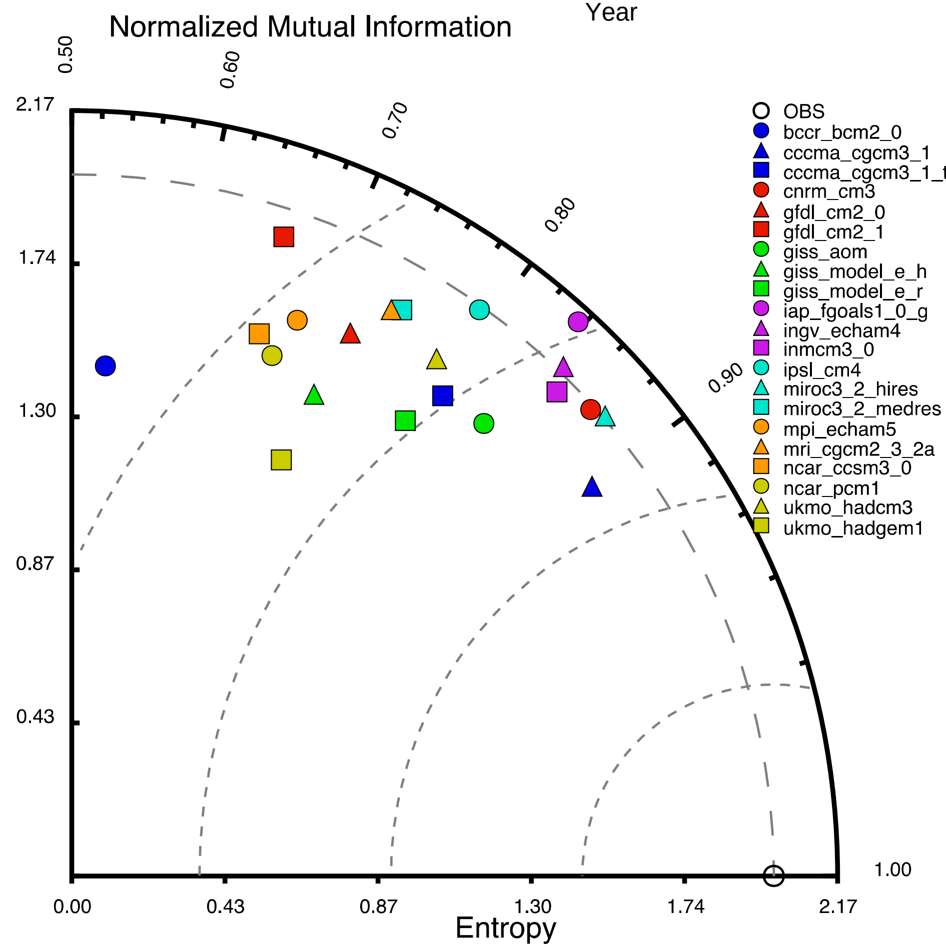


Intercomparison Studies

Annual mean temperature 1900-2000

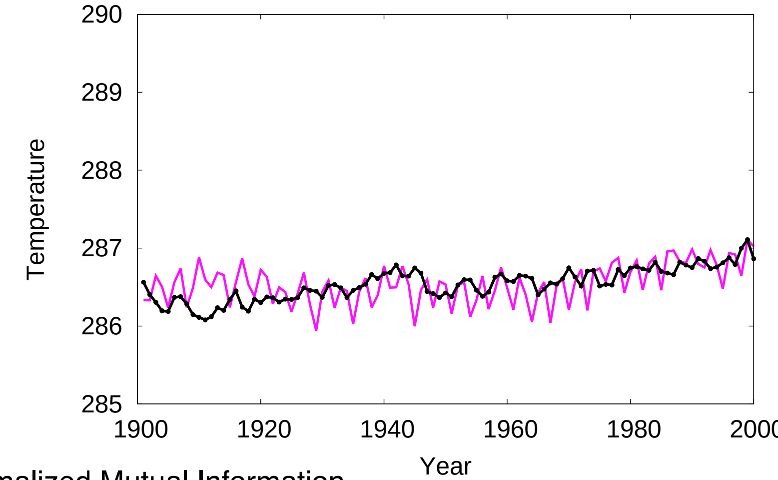


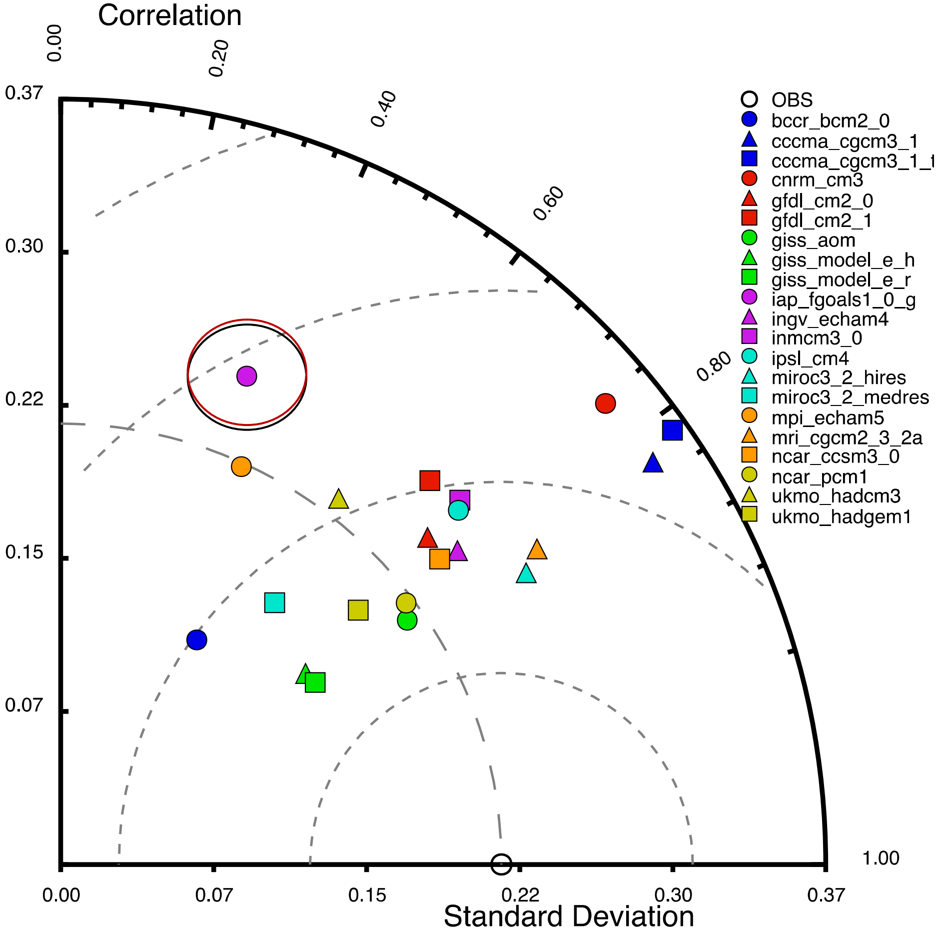


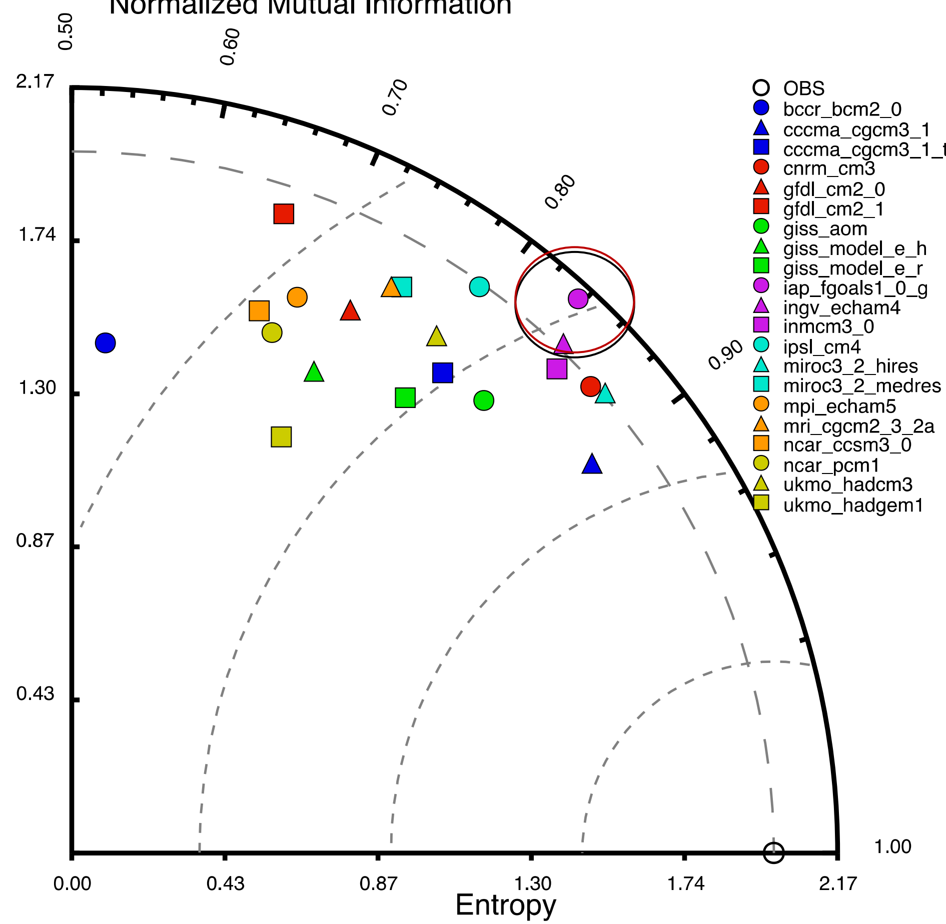


Intercomparison Studies

Annual mean temperature 1900-2000

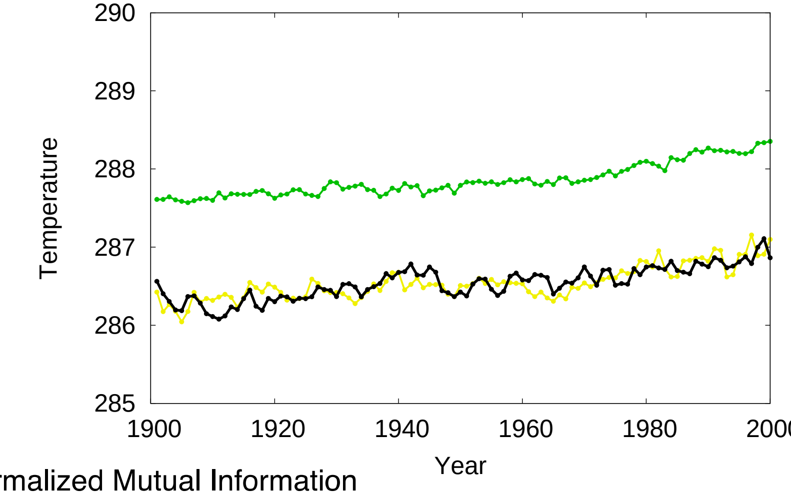


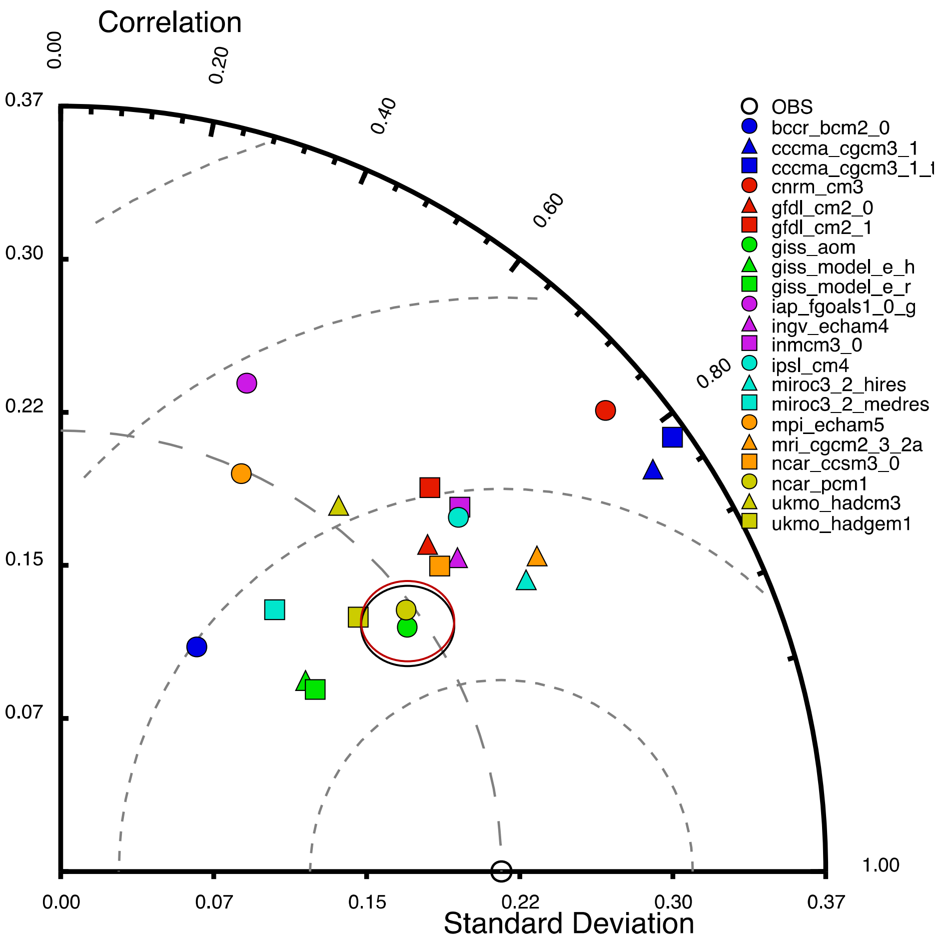


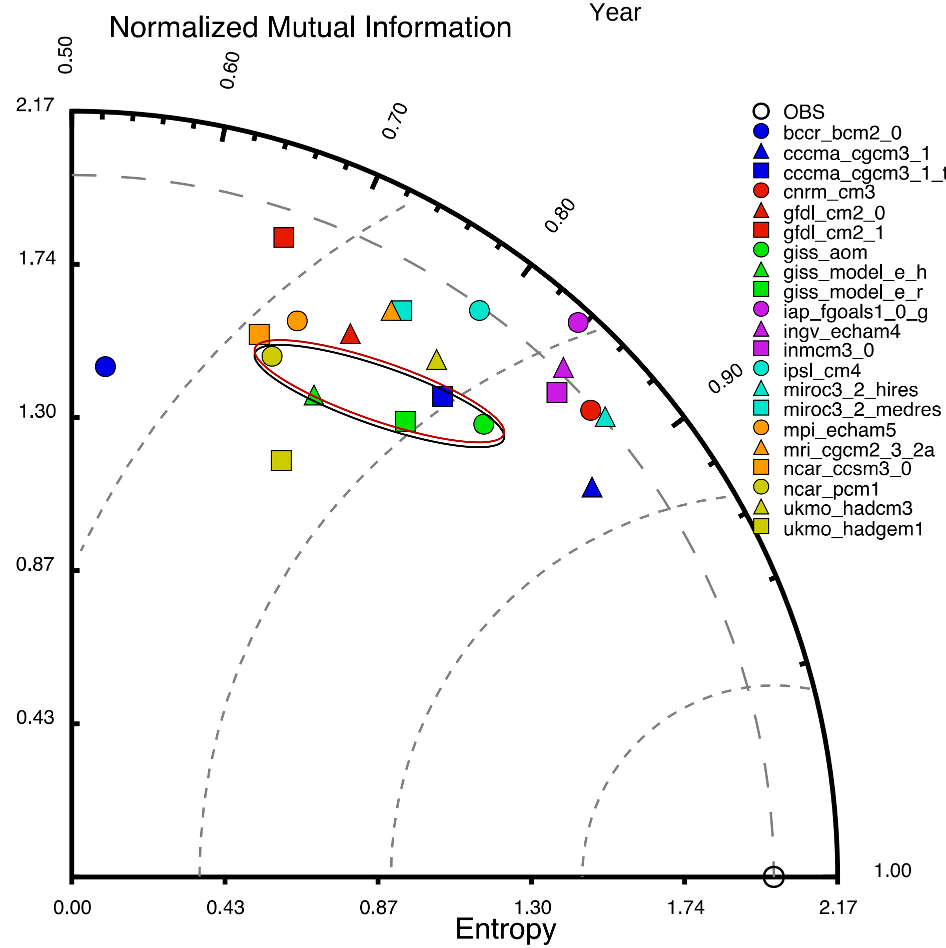


Intercomparison Studies

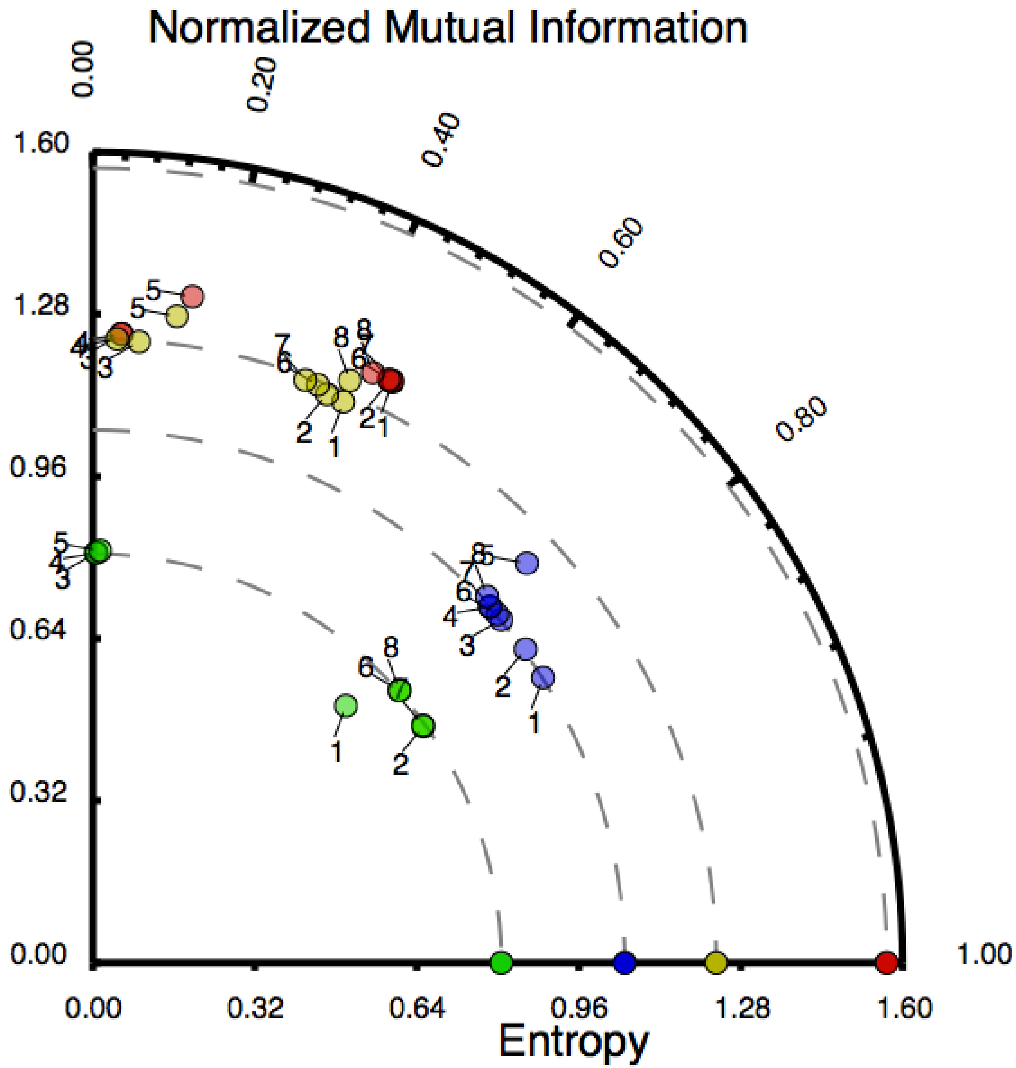
Annual mean temperature 1900-2000







MID applies to discrete data: useful when comparing Clustering Results



- Summarize study in clustering [Filippone et al. 2009]
- 8 different methods
- 4 classification problems

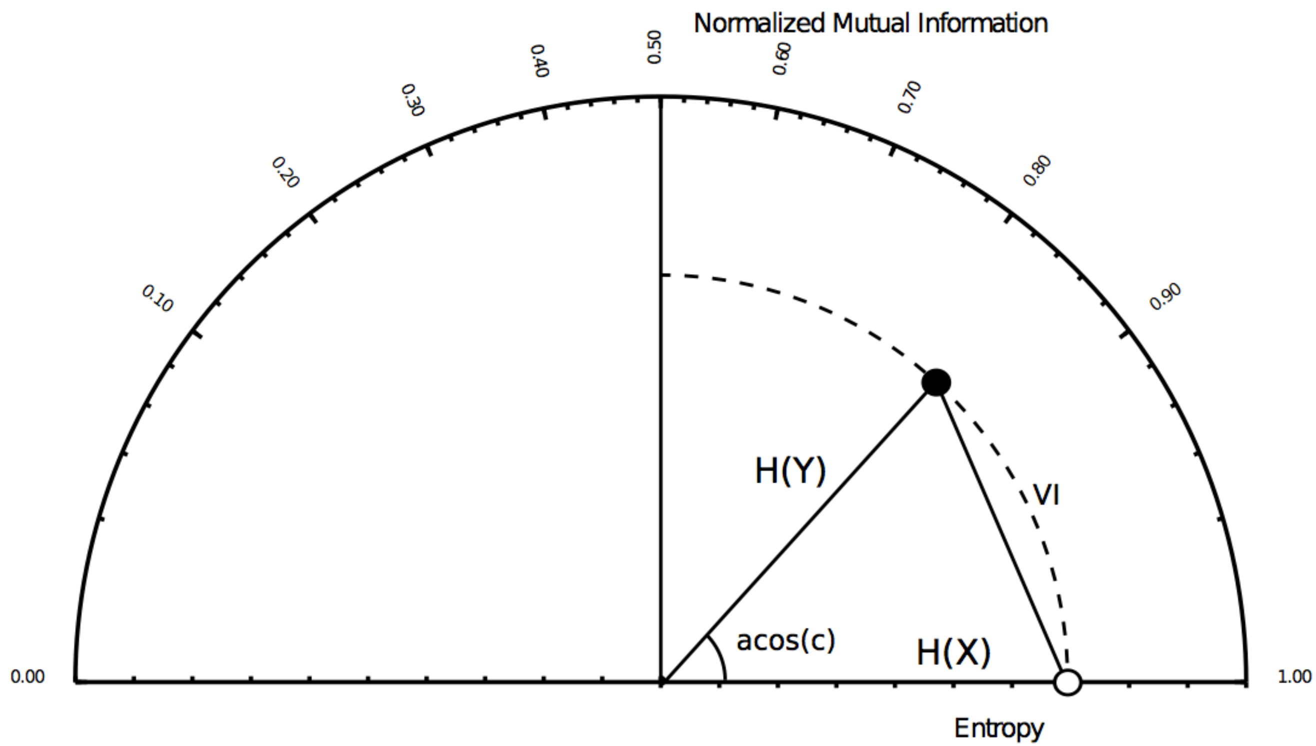
Concluding Remarks

- Taylor diagram:
  - easy to compute.
  - Well understood in geophysical sciences, climate.

- MI diagram:
  - Counterpart using information theory.
  - requires an estimation step that may introduce additional uncertainties.
  - extends nicely to categorical data, multi-variate distributions.
  - exposes non-linearities, difficult to see via (linear) correlation.

- More informed decisions when combining both diagrams.

Thanks!



Questions?