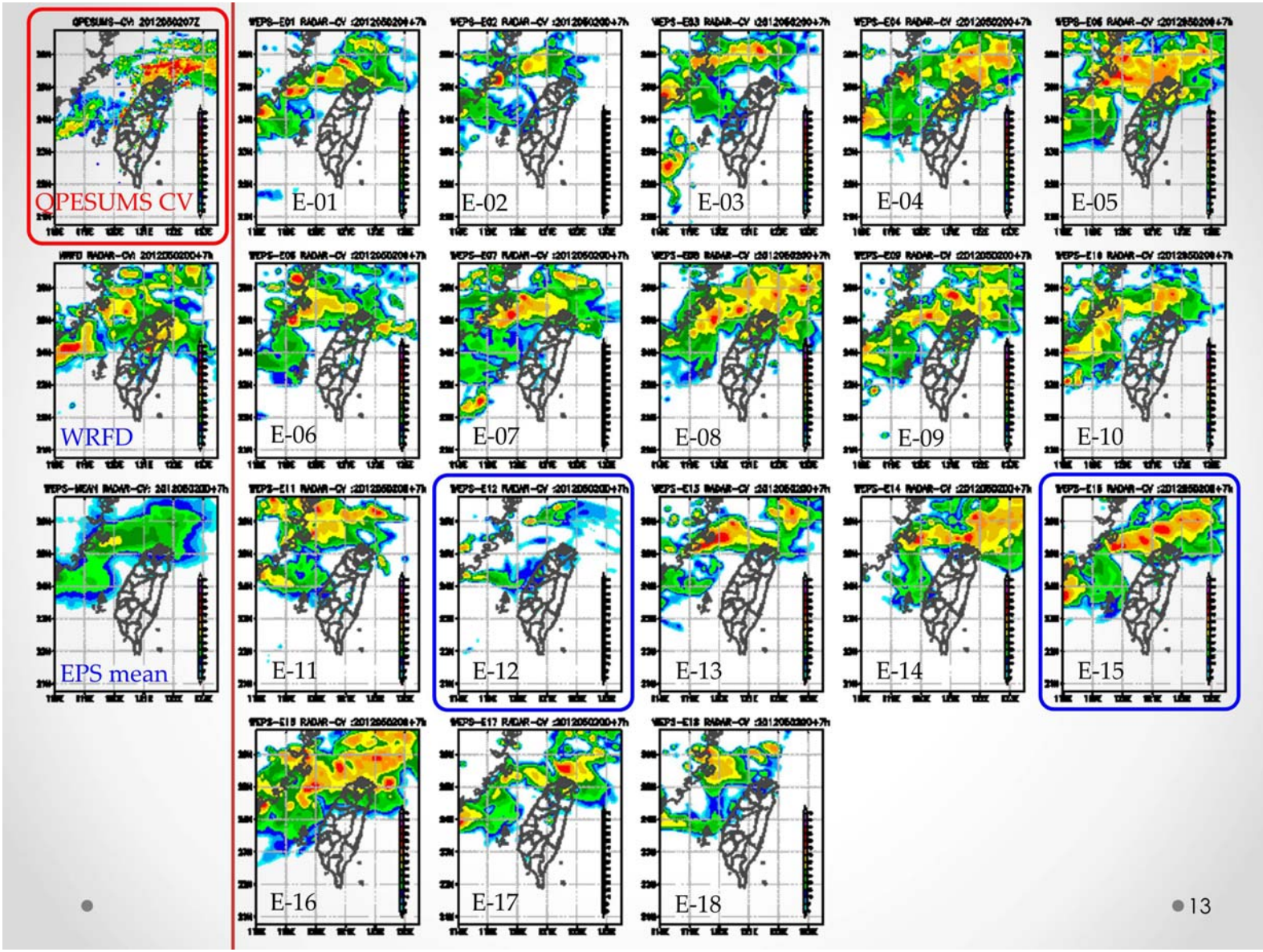
The Diagnostic Tool for Ensemble Prediction System using Tri-plots Methods

曾建翰　蔡禹明
中央氣象局
John Chien-Han Tseng Yu-Ming Tsai
Central Weather Bureau, Taipei, Taiwan

The Purpose

- Based on automatic, machine learning method to detect which ensemble member similar to true state (observation). Also, sorting the similarity.
  --mean error, root mean square method are too rigid to comparing the similarity

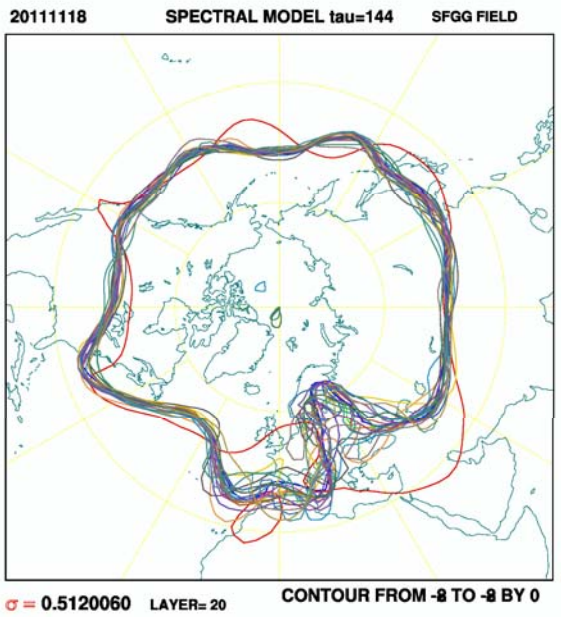
The Stamp Figures…

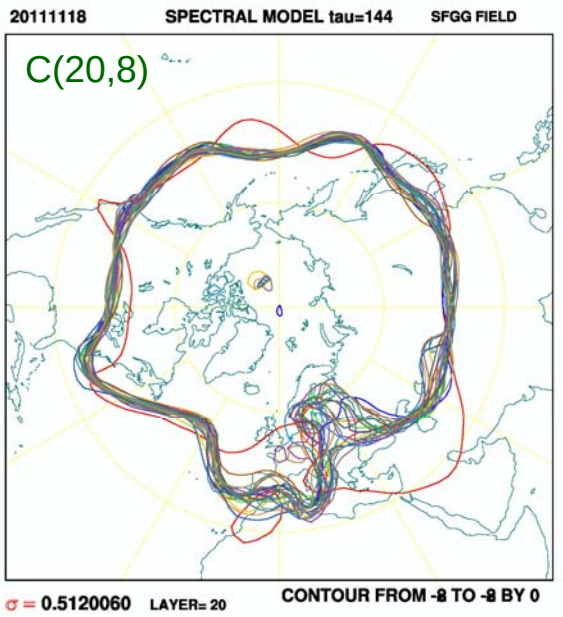


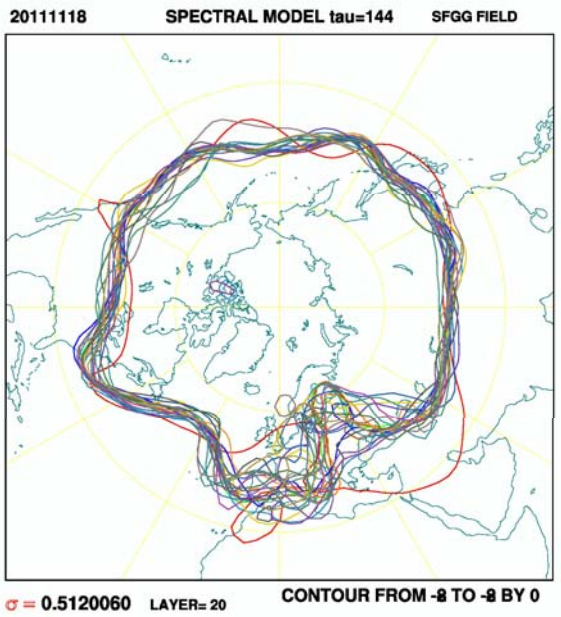
courtesy of 李志昕、洪景山

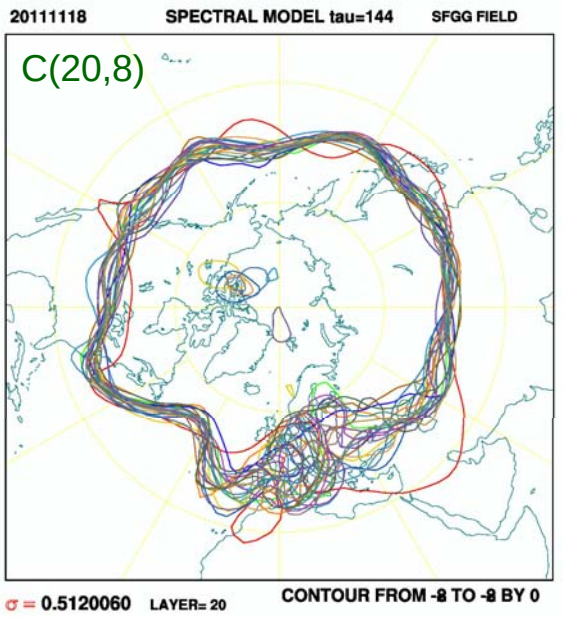
The Spaghetti…

SV x 3







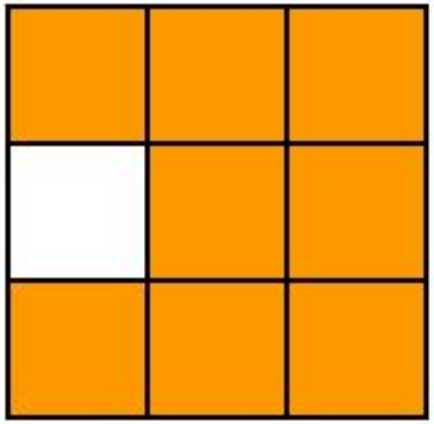


The Cases: Two Ensemble Systems

- The results from CWB WRF, the case time: 20120602 and 20120611
- CWB GFS dynamic T42L40 singular vectors, the case time: 20050215 and 20111118

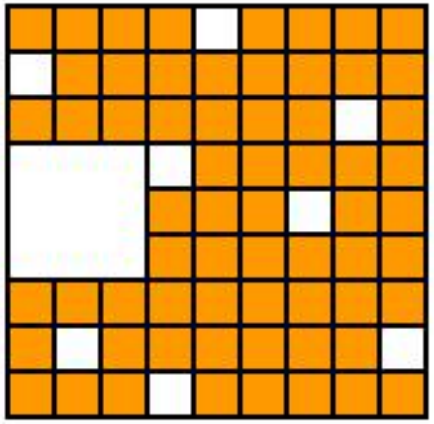
Fractal dimension

a



S_1

b



S_2

These two diagrams have same fractal dimension.

$$in\ S_1,\ N = 8,\ \varepsilon = \frac{1}{3}$$

$$in\ S_2,\ N = 8^2,\ \varepsilon = (\frac{1}{3})^2$$

$$it\ implies$$

$$d = \lim_{\varepsilon \to 0} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)} = \frac{\log(8^n)}{\log(3^n)} = \frac{n \log 8}{n \log 3} = \frac{\log 8}{\log 3}$$

Tri-plots: cross-plot, self-plot

Assuming there are two datasets *A* and *B*, and the cross-plot function is defined as:

$$Cross_{A,B}(r) = \log(\sum_i C_{A,\,i}\ C_{B,\,i})$$

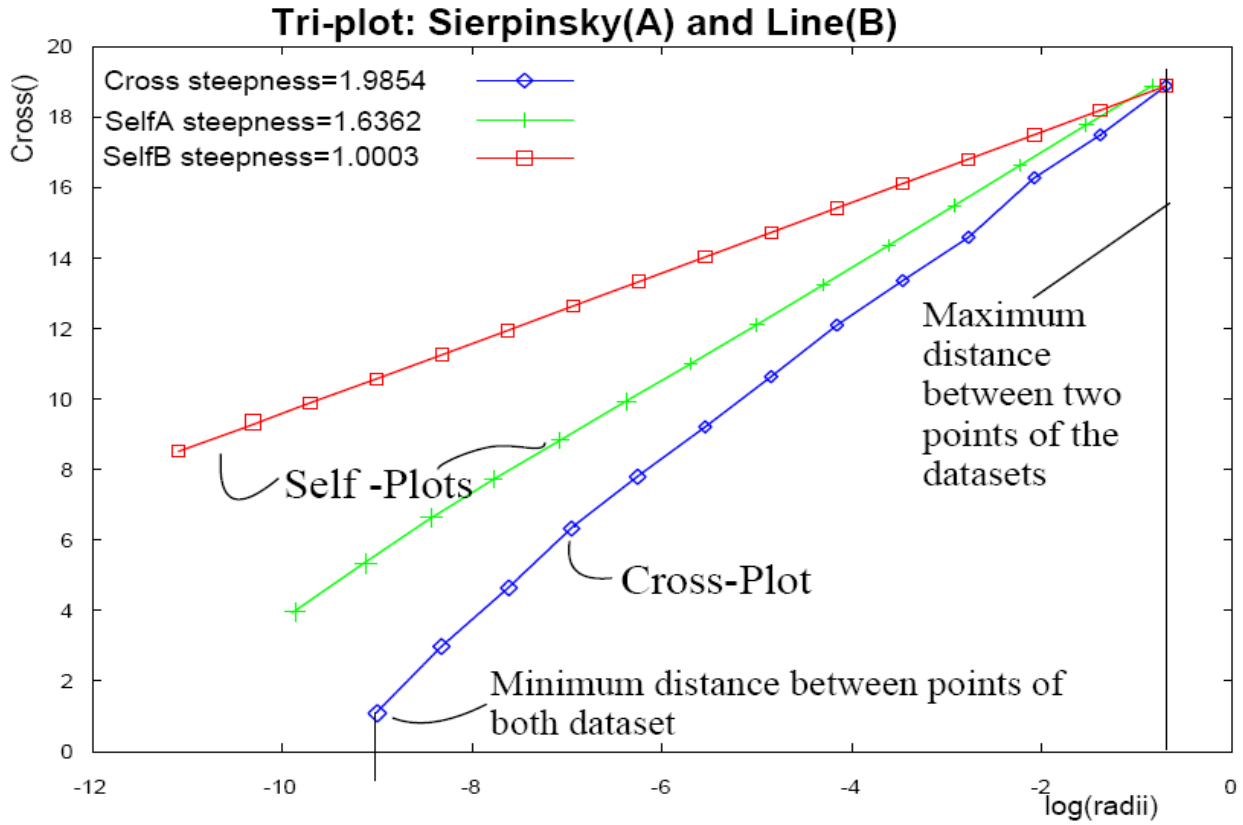
where $C_{A,\,i}$ ($C_{B,\,i}$) is the number of points from set *A*（*B*）in the *i*-th cell, and *r* is the distance of the pairs of points. Hence, the cross-plot function is proportional to the count of *A*-*B* pairs within distance *r*, and the cross-plot is the figure of the cross-plot function versus log(*r*).

Also, the self-plot function is defined as

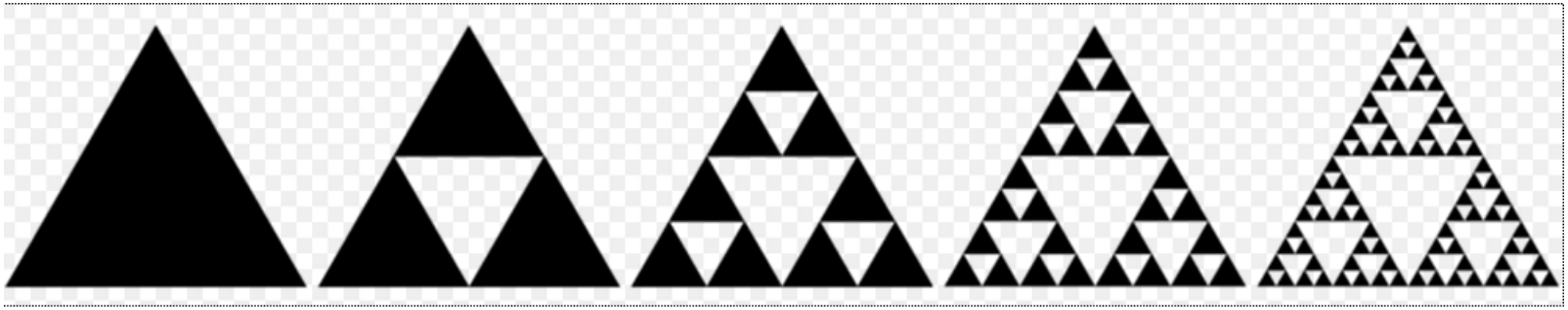
$$Self_A(r) = \log(\frac{\sum_i C_{A,\,i} \cdot (C_{A,\,i} - 1)}{2})$$

The typical tri-plots result

When we compare two datasets A and B,…

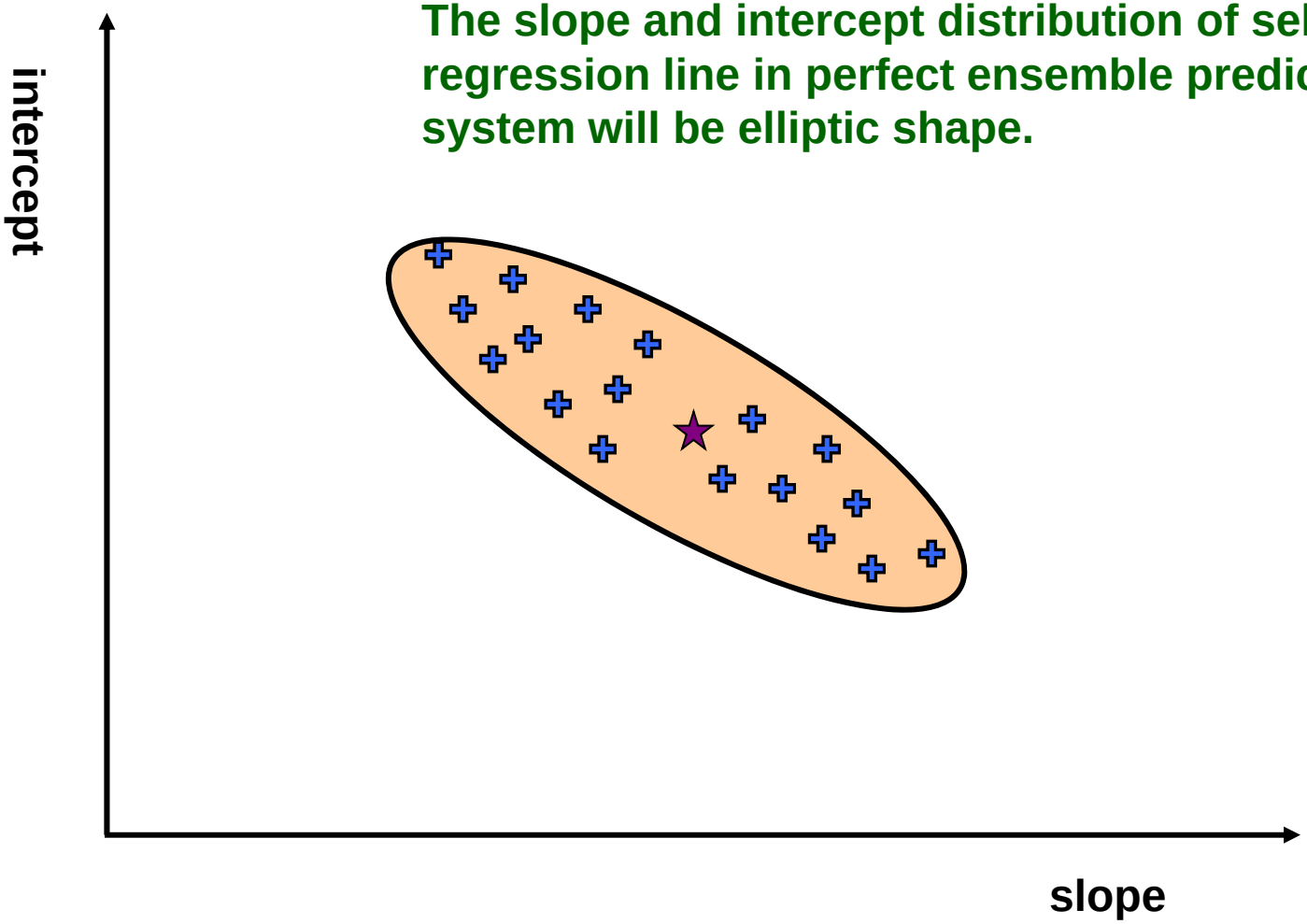


Adopt form Traina et al.(2001)



The Perfect Ensemble Self-Polts Distribution

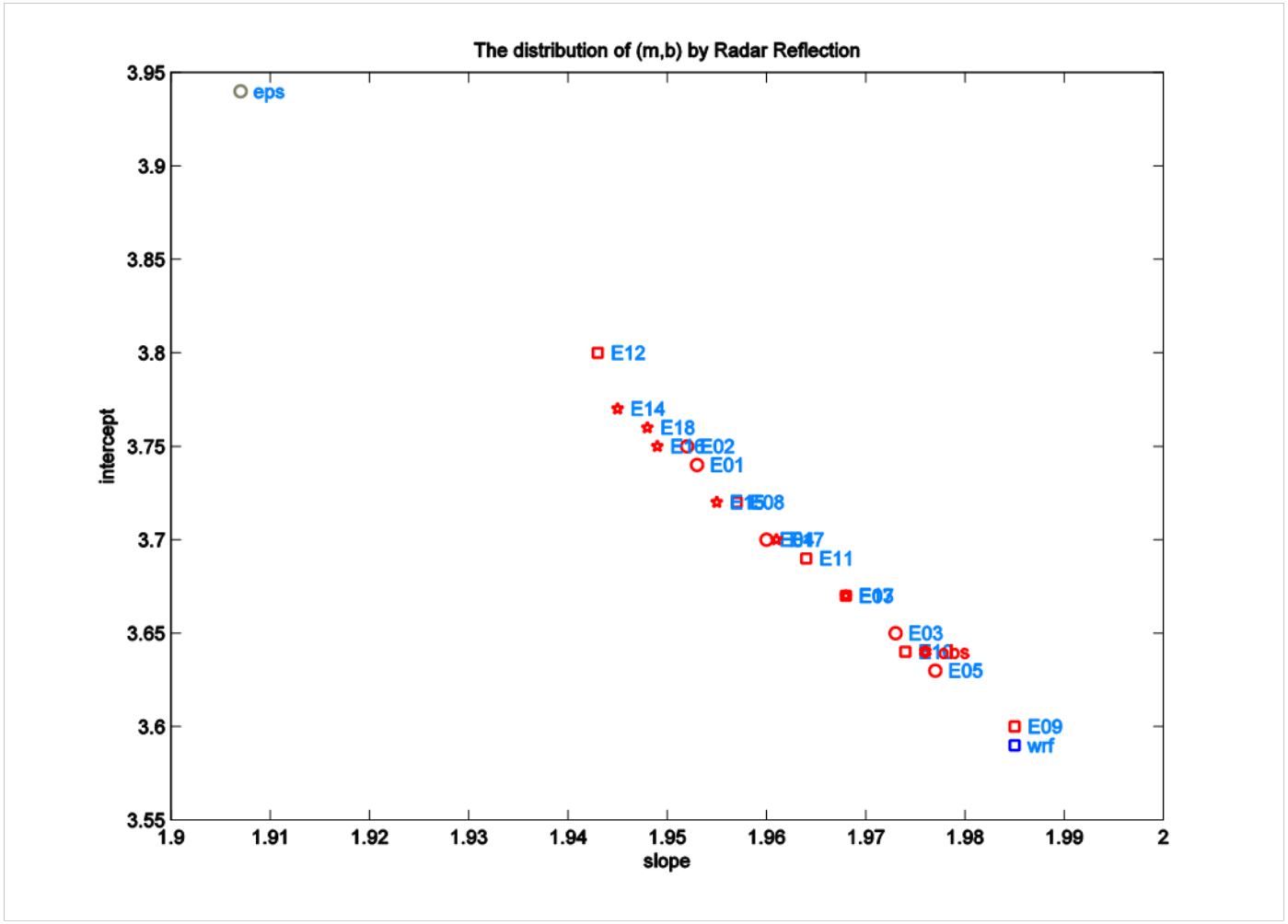
The slope and intercept distribution of self-plots regression line in perfect ensemble prediction system will be elliptic shape.



The Self-Plot Distribution of WRF Ensemble

WRF EPS, based on Tri-Plots，E03, E05, E10 are similar to the observation …

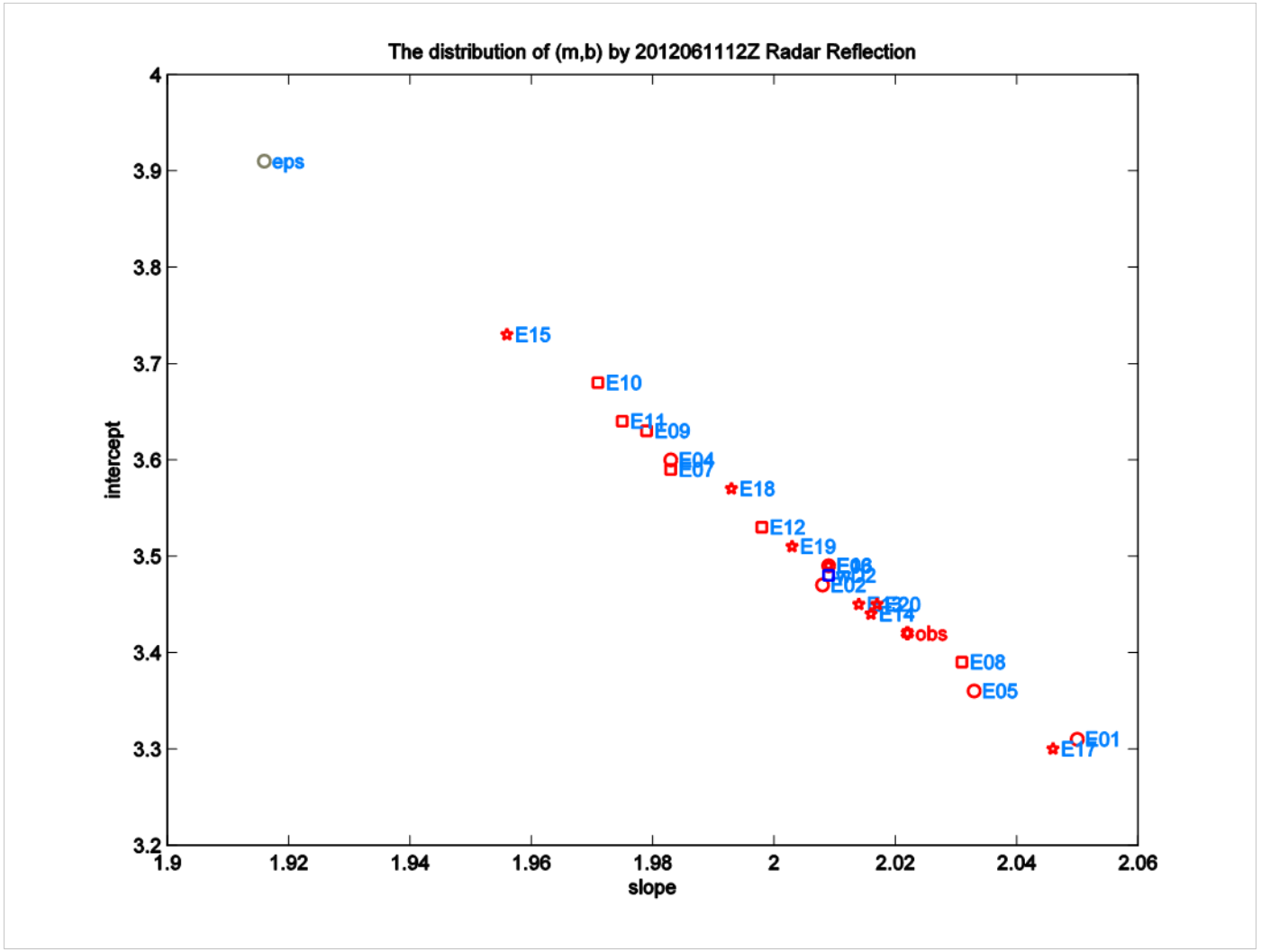
The intercept of the regression line



The slope of the regression line, Fractal dimension

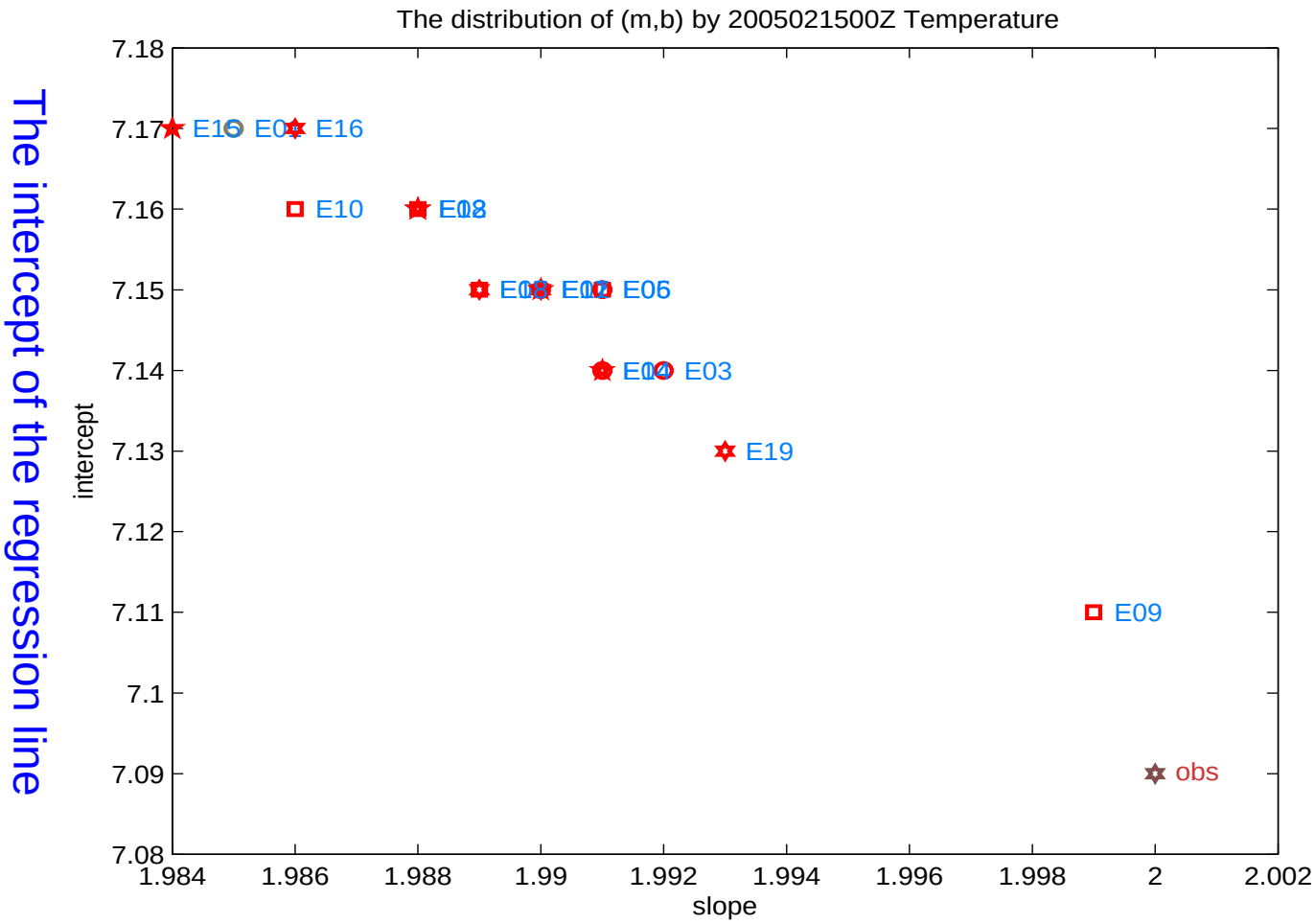
Tri-Plots based on DYM Ensemble(cont.)

Another case but similar distribution, and it seems too regular and reflecting the characteristic of the different physical schemes ensemble system …



Tri-Plots based on DYM Ensemble

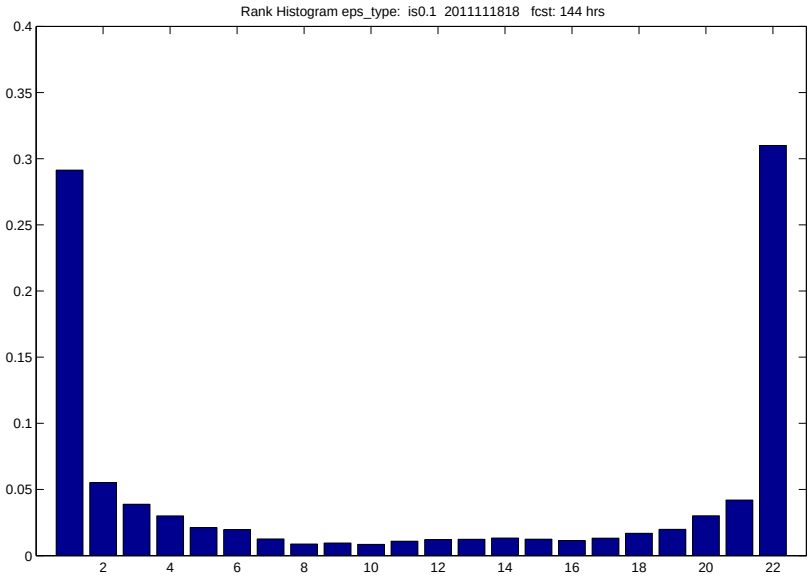
random error perturbation



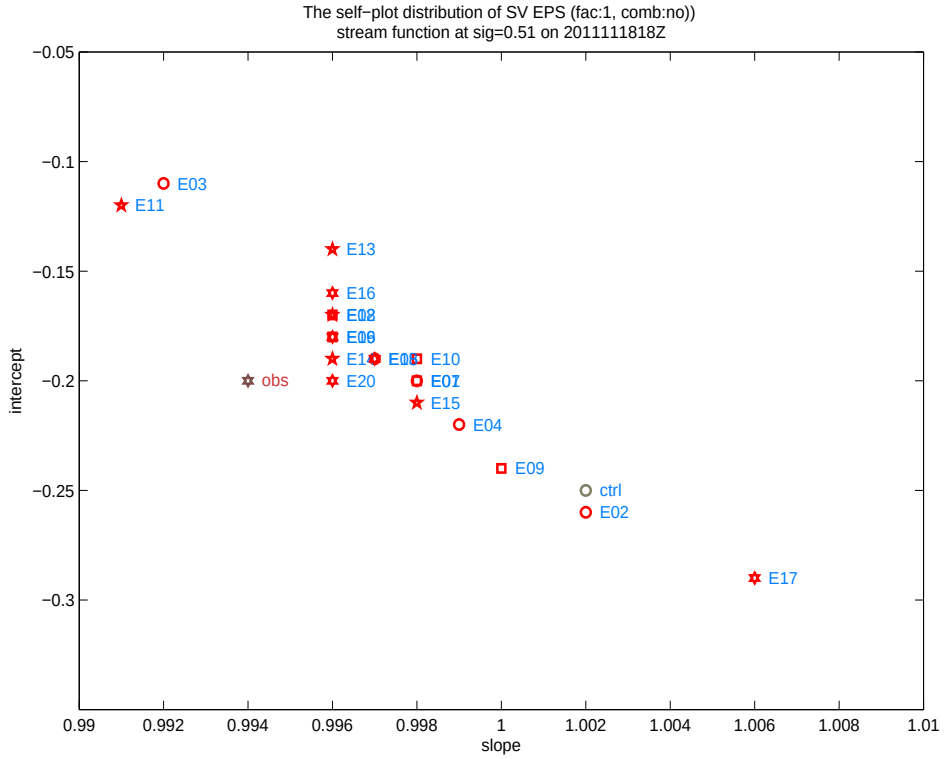
The intercept of the regression line

The slope of the regression line, Fractal dimension

EPS Diagnostic Tools ? (1)

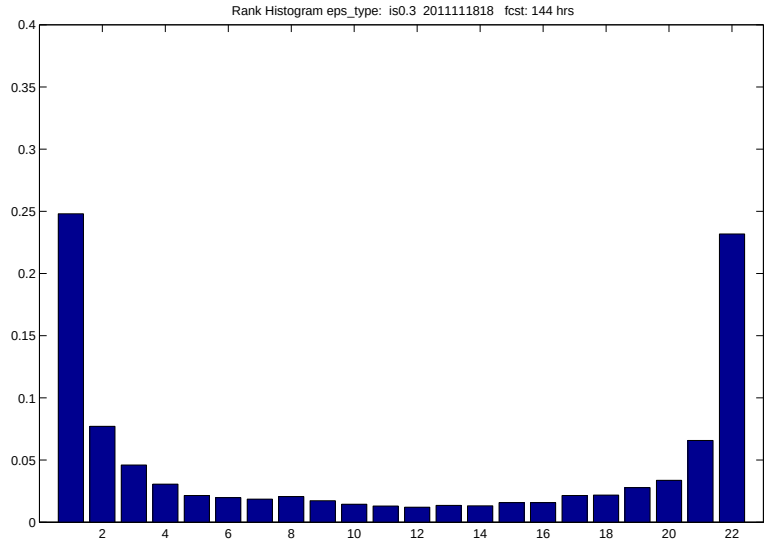


dym T42L40 Singular Vector Ensemble Prediction
optimum time: 12 hrs, stream function on *sig*=0.51
20 ensemble members + 1 deterministic run



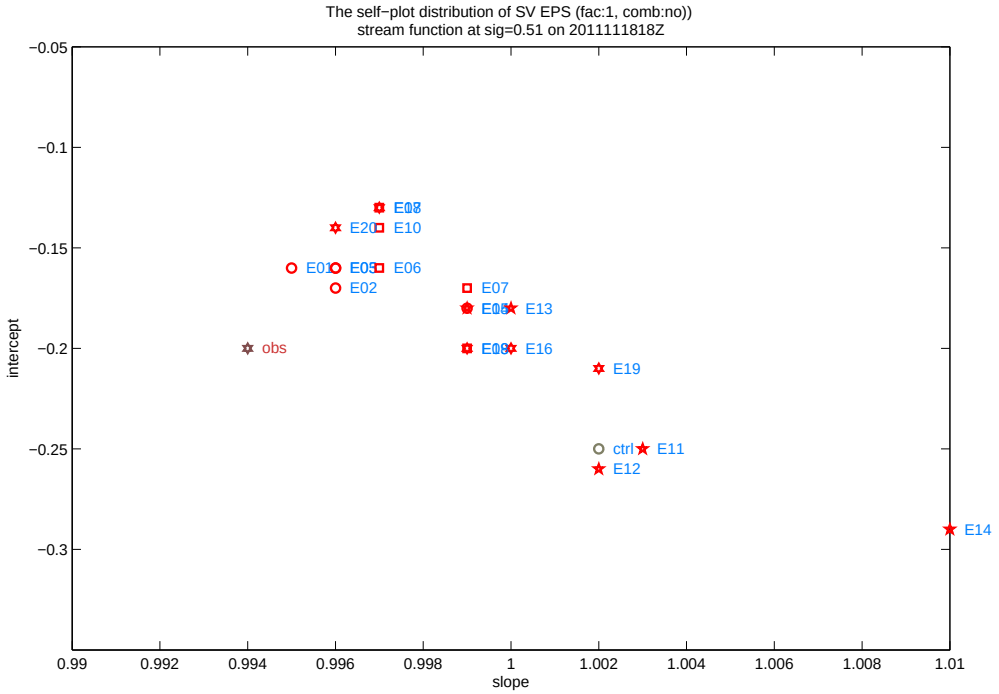
the spread is not enough, then how to do next…

EPS Diagnostic Tools ? (2)



dym T42L40 Singular Vector Ensemble Prediction
optimum time: 12hrs, stream function on *sig*=0.51
20 ensemble members + 1 deterministic run
the magnitude of SV times 3 !

when increasing the magnitude of the SVs, the spread extends little larger based on rank histogram, but never be seen in other diagnostic tool e.g. fractal dimension.



Conclusions

- Self-plots distribution points to which members are close to the true state and how close. Also, it shows the bias situation of the ensemble prediction system.
- estimate the similarity quantitatively and sort the rank of the similarity to the true state in necessary.
- Can measure the similarity of variable 3D structures; can measure two different variables at the same time.