



Comparison of Clustered Fields

Caren Marzban (1,2), Xiaochuan Du (2), and Scott Sandgathe (1)

(1) University of Washington, Applied Physics Lab, Seattle, WA, United States (marzban@stat.washington.edu), (2) Univ. of Washington, Department of Statistics, Seattle, WA 98195 USA

One approach to the verification of spatial/gridded forecasts involves identifying clusters in the forecast and observed fields, and then comparing the two clusterings. Given that clustering itself is often a stochastic process with random outcomes, the comparison of clusterings is important not only in verification (i.e. comparison of the clusterings of two fields - observed and forecast) but also in the assessment of the consistency of clusterings of a single field.

A number of methods for comparing clusterings have already been proposed in the statistics and machine learning communities. Many of these methods involve quantities which resemble scalar measures of verification of non-spatial fields commonly used in the verification of binary forecasts and observations in meteorology. Here, the similarities are clarified, and some of the methods are applied to several data sets including forecasts from Weather Research and Forecasting, the PSU/NCAR mesoscale model (known as MM5), and data sets from the Mesoscale Verification Intercomparison over Complex Terrain (MesoVICT) project.