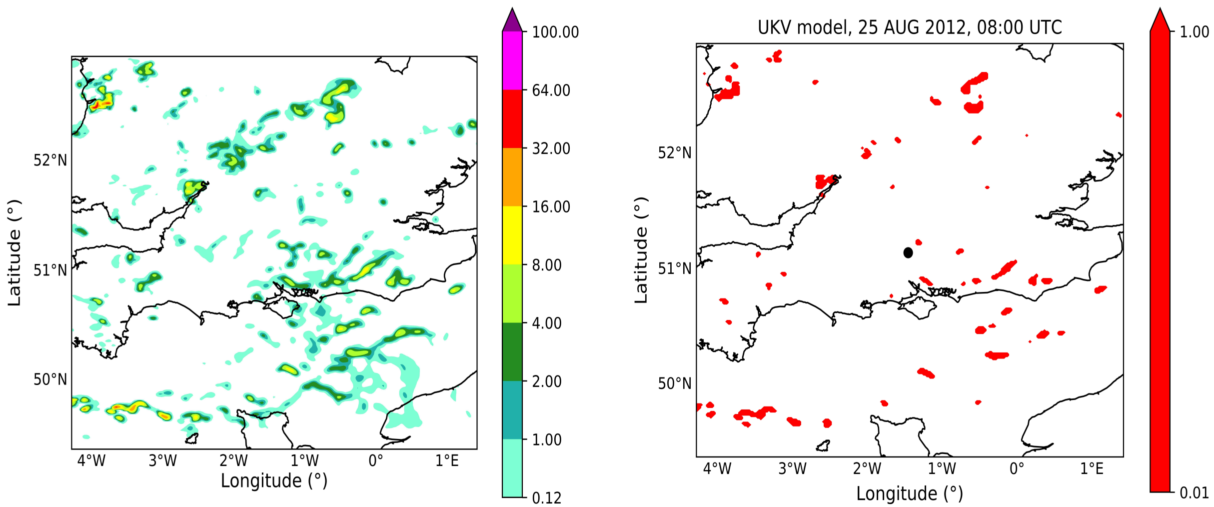
Beyond I_{org}: Advertisement for a suite of more robust organization measures

Bob Plant

Department of Meteorology, University of Reading



ParaCon Plenary

Met Office, 20-21 June 2022

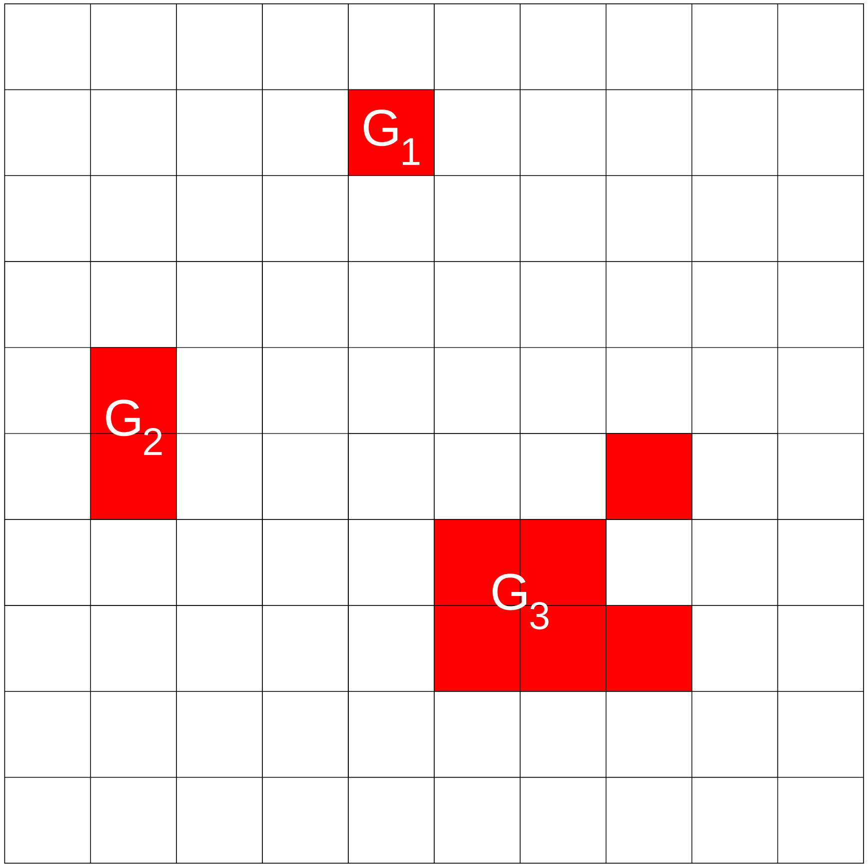
The University of Reading

Measuring Organization

- By organization, we will mean an arrangement of points that can be distinguished from the action of a Poisson process
- i.e., something other than scattering N points randomly within the area A
- Fortran code to calculate suite of metrics that test this and characterise the organization with options for
  - cloud definition
  - fixed or singly-periodic or bi-periodic bc's
  - corrections for finite size effects
  - corrections for finite domain effects

The University of Reading

Reduction to a set of points



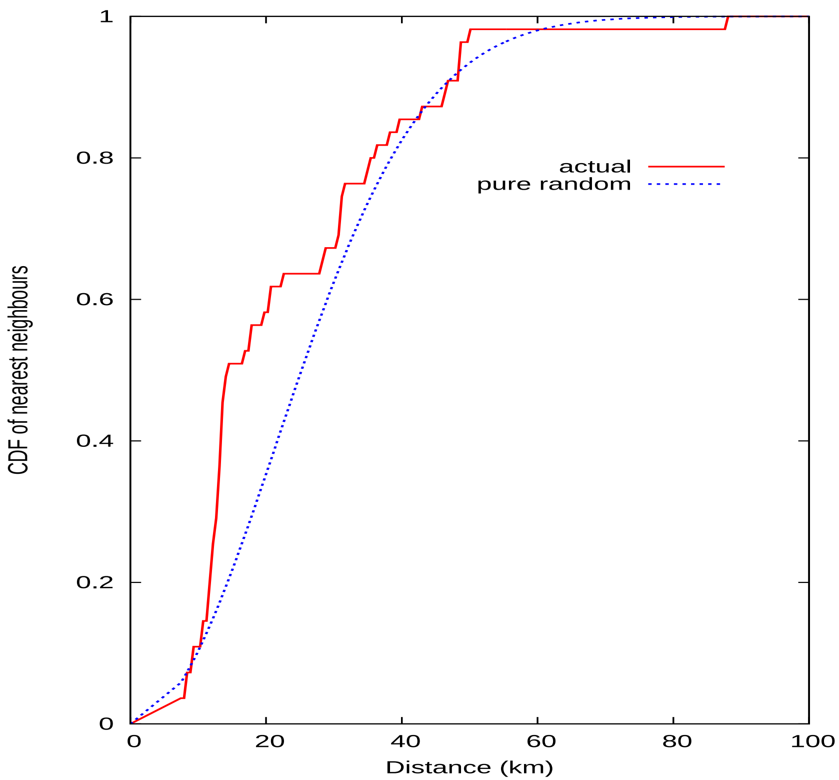
- Apply threshold
- Group the points into cloud objects
- Define a central point

$$\mathbf{x} = \frac{\sum_{i=1}^{P} \mathbf{x_i}}{P}$$

- Or: apply watershed algorithm (based on local maxima)

The University of Reading

I_{org} (Tompkins and Semie 2017)



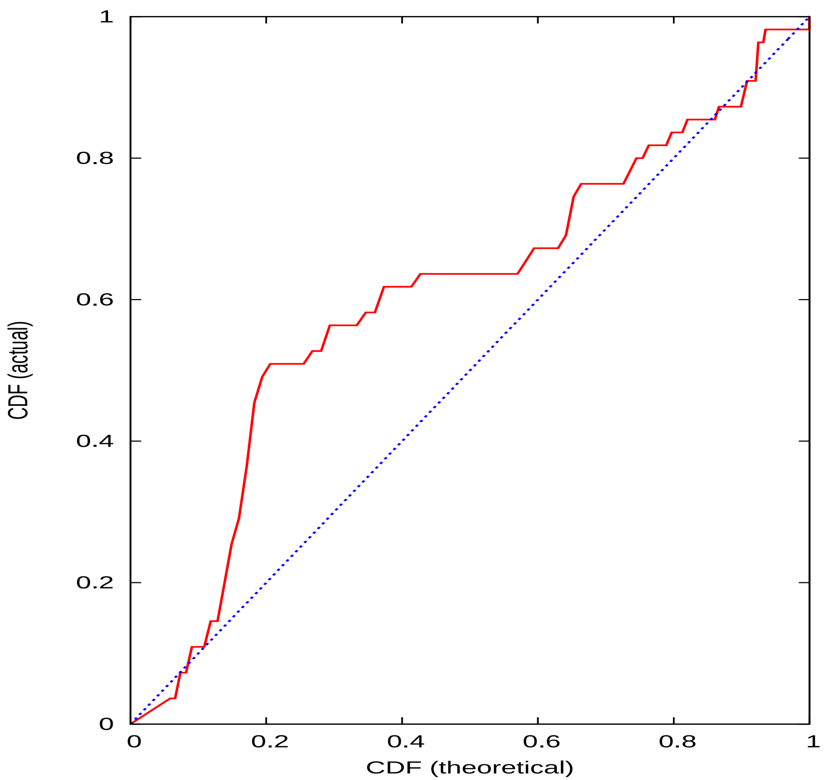
- From each cloud point find distance to its nearest neighbour
- Construct a CDF
- Compare to distribution for a random pattern

$$1 - \exp(-\lambda \pi r^2)$$

where $\lambda = N/A$

The University of Reading

I_{org} (Tompkins and Semie 2017)



- Plot actual (A) vs theoretical (T) CDF

$$I_{org} = \int_0^1 A(T)dT$$

- Random if $I_{org} = 0.5$ (for $A = T$)
- Clustered if $I_{org} > 0.5$
- Regular if $I_{org} < 0.5$
- In our example, $I_{org} = 0.604$

The University of Reading

Practical Issues

- Formula assumes points in infinite domain, not extended objects on finite domain
- Consider a Monte-Carlo approach
  - Relocate the clouds randomly in the area
  - Redraw until they all fit without touching
  - Calculate separation distances
  - Repeat many times (e.g. 1000)
- Sampling approach allows significance tests if required
- In our example, using MC reference increases I_{org} from 0.604 to 0.646

Boundary effects

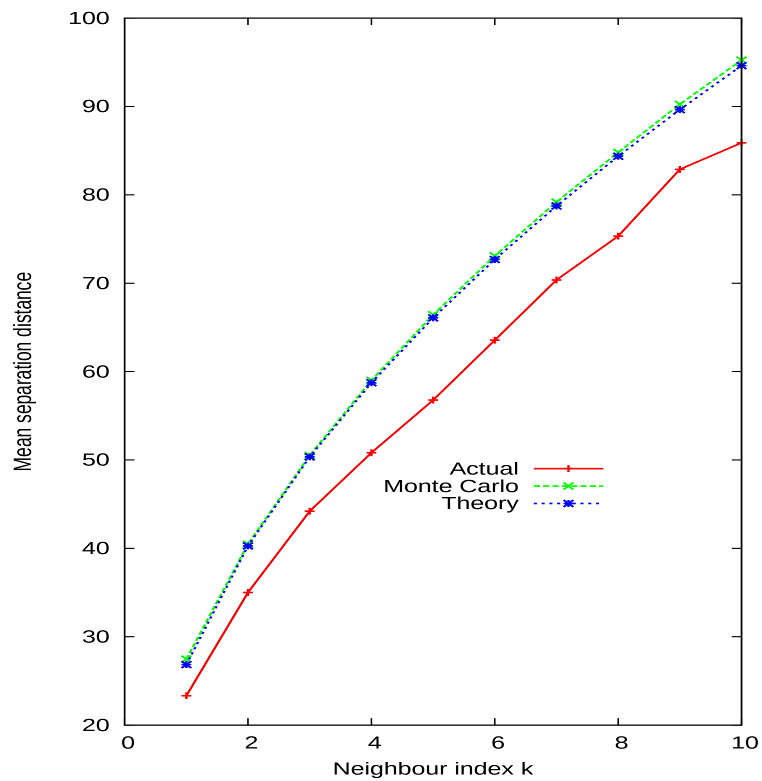
- I_{org} now well defined and 0.5 is meaningful
- But for a true best-estimate of the CDF on a finite domain, we must correct for clouds close to fixed boundary edges
- Let distance to nearest boundary be b_i
- If $b_i < r_i^{\text{nn}}$ there might be a nearer neighbour outside our domain

$$C(r) = \frac{\sum_i I(b_i > r > r_i^{\text{nn}})}{\sum_i I(b_i > r_i^{\text{nn}})}$$

where $I = 1$ if bracketed condition is true

Higher order measures

- I_{org} can be sensitive to cloud defintion (small fragments)
- Consider kth nearest neighbour distance
- Can generalize I_{org} in a natural way, using theory and/or Monte Carlo methods with corrections
- Increasing k focuses on the pattern at larger distances



The University of Reading

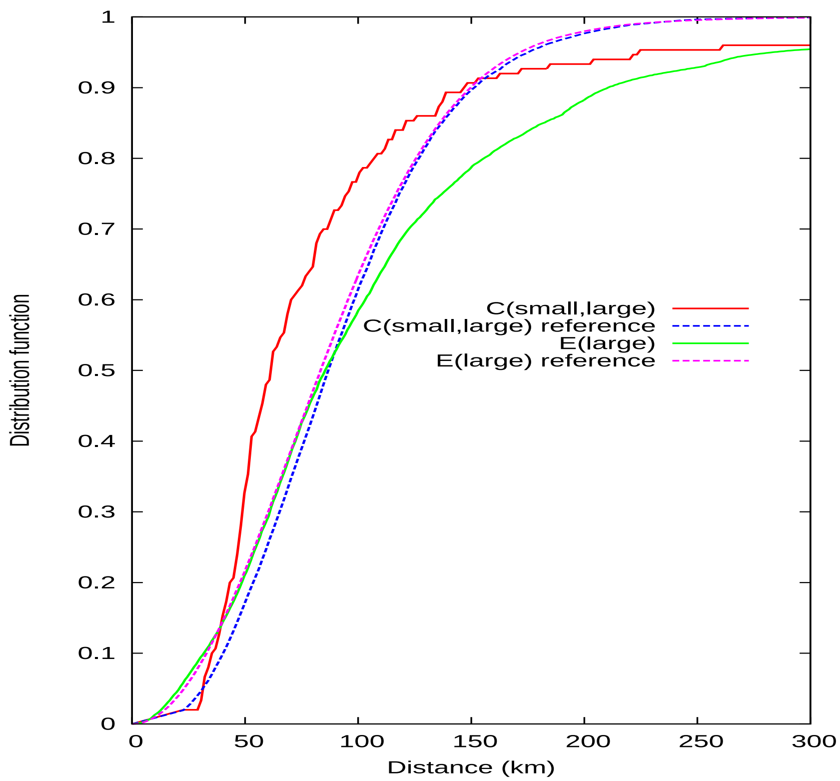
Empty space function

- Robust to cloud fragments very near other clouds
- Take M arbitrary points and measure the distance to the nearest cloud
- The CDF is the probability that a randomly-placed disc of radius r contains at least one cloud
- The theoretical distribution is the same as the nearest-neighbour distribution
- But better to obtain a Monte Carlo reference as before

Example of use

- Divide the clouds into two types (by area, rain rate, w...)
- We can analyse each type separately
- But we could also ask "starting from a cloud of type 1 what is the distance to the nearest type 2 cloud?"
- Let the corresponding mixed CDF be $C_{12}(r)$
- Is the pattern of the type 1 clouds related to the pattern of the type 2 clouds?
- If the two types are independent then $C_{12}(r) = E_2(r)$

An RCEmip example



- RCEmip snapshot split into smaller (2/3 of the total) and larger clouds
- $C_{12} \neq E_2 \neq$ reference
- Interpretation is that small clouds form clusters that are preferentially within the vicinity of the large ones

Summary

- A growing number of studies seek to measure organization
- Often using nearest-neighbour distances
- Finite cloud-size effects and finite domain effects can be important which means that a Monte-Carlo approach is needed for the unorganized reference process
- Other related measures are valuable to cross-check and enhance our interpretations (e.g. empty space functions more robust)
- Happy to share code if you want to explore