**INFSCI 2809: Spatial Data Analytics**

**Midterm Exam**

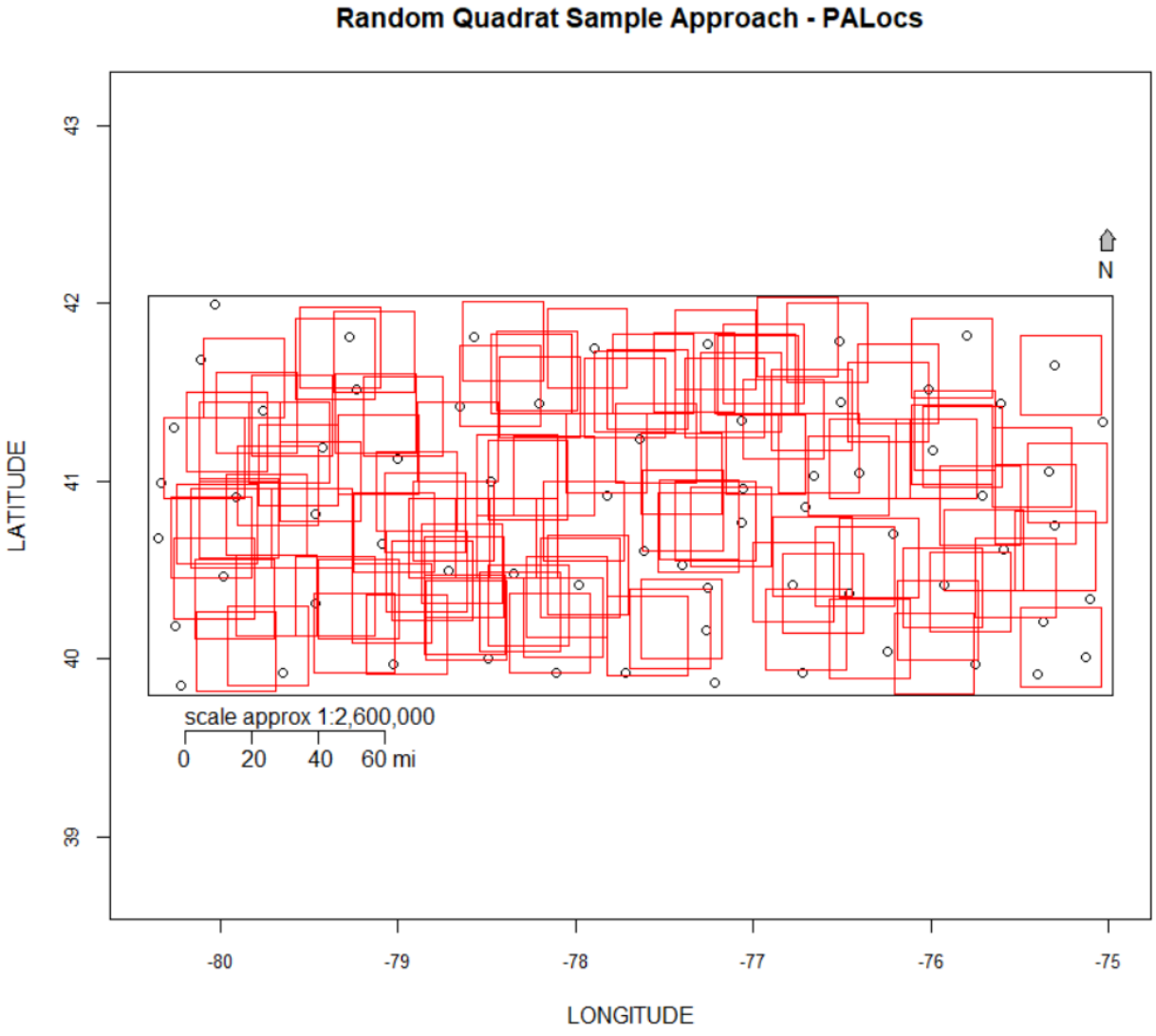
Student Name: Zechen Wang

PITTID: zew20@pitt.edu

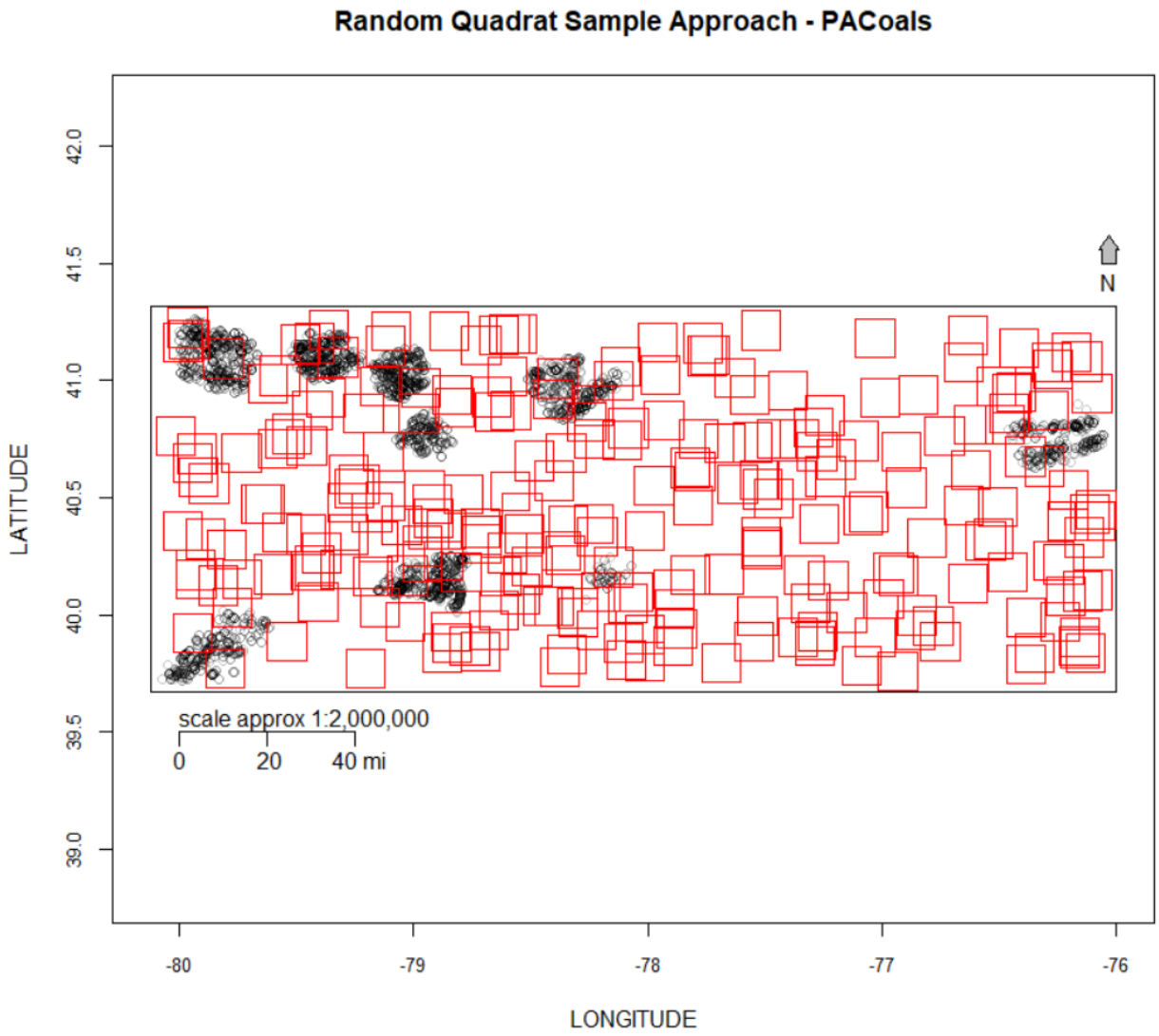
## PART A

I. Maps

The map for 'PALocs':



The map for ‘PACoals’:




II. Tables

The table for ‘PALocs’:

| | $K$ | $X$ | $K-\mu$ | $(K-\mu)^2$ | $X(K-\mu)^2$ |
|---|---|---|---|---|---|
| **1** | 0 | 22 | -1.1167 | 1.24694 | 27.4328 |
| **2** | 1 | 77 | -0.1167 | 0.01361 | 1.04806 |
| **3** | 2 | 20 | 0.88333 | 0.78028 | 15.6056 |
| **4** | 3 | 1 | 1.88333 | 3.54694 | 3.54694 |
| **Total** | | 120 | | | 47.6333 |

The table for ‘PACoals’:

| | $\mathbf{K}$ | $\mathbf{X}$ | $\mathbf{K-\mu}$ | $\mathbf{(K-\mu)^2}$ | $\mathbf{X(K-\mu)^2}$ |
|---|---|---|---|---|---|
| **1** | 0 | 202 | -16.424 | 269.748 | 54489.1 |
| **2** | 35 | 2 | 18.576 | 345.068 | 690.136 |
| **3** | 166 | 1 | 149.576 | 22373 | 22373 |
| **4** | 9 | 1 | -7.424 | 55.1158 | 55.1158 |
| **5** | 42 | 1 | 25.576 | 654.132 | 654.132 |
| **6** | 102 | 1 | 85.576 | 7323.25 | 7323.25 |
| **7** | 3 | 2 | -13.424 | 180.204 | 360.408 |
| **8** | 57 | 1 | 40.576 | 1646.41 | 1646.41 |
| **9** | 83 | 1 | 66.576 | 4432.36 | 4432.36 |
| **10** | 145 | 1 | 128.576 | 16531.8 | 16531.8 |
| **11** | 51 | 1 | 34.576 | 1195.5 | 1195.5 |
| **12** | 123 | 1 | 106.576 | 11358.4 | 11358.4 |
| **13** | 1 | 3 | -15.424 | 237.9 | 713.699 |
| **14** | 50 | 1 | 33.576 | 1127.35 | 1127.35 |
| **15** | 25 | 1 | 8.576 | 73.5478 | 73.5478 |
| **16** | 38 | 1 | 21.576 | 465.524 | 465.524 |
| **17** | 10 | 1 | -6.424 | 41.2678 | 41.2678 |
| **18** | 2 | 1 | -14.424 | 208.052 | 208.052 |
| **19** | 40 | 1 | 23.576 | 555.828 | 555.828 |
| **20** | 169 | 1 | 152.576 | 23279.4 | 23279.4 |
| **21** | 259 | 1 | 242.576 | 58843.1 | 58843.1 |
| **22** | 82 | 1 | 65.576 | 4300.21 | 4300.21 |
| **23** | 41 | 1 | 24.576 | 603.98 | 603.98 |
| **24** | 63 | 1 | 46.576 | 2169.32 | 2169.32 |
| **25** | 299 | 1 | 282.576 | 79849.2 | 79849.2 |
| **26** | 178 | 1 | 161.576 | 26106.8 | 26106.8 |
| **27** | 27 | 2 | 10.576 | 111.852 | 223.704 |
| **28** | 113 | 1 | 96.576 | 9326.92 | 9326.92 |
| **29** | 24 | 1 | 7.576 | 57.3958 | 57.3958 |
| **30** | 97 | 1 | 80.576 | 6492.49 | 6492.49 |
| **31** | 6 | 2 | -10.424 | 108.66 | 217.32 |
| **32** | 135 | 1 | 118.576 | 14060.3 | 14060.3 |
| **33** | 195 | 1 | 178.576 | 31889.4 | 31889.4 |
| **34** | 191 | 1 | 174.576 | 30476.8 | 30476.8 |
| **35** | 130 | 1 | 113.576 | 12899.5 | 12899.5 |
| **36** | 5 | 1 | -11.424 | 130.508 | 130.508 |
| **37** | 321 | 1 | 304.576 | 92766.5 | 92766.5 |
| **38** | 194 | 1 | 177.576 | 31533.2 | 31533.2 |

| | | | | | |
|---|---|---|---|---|---|
| **39** | 16 | 1 | -0.424 | 0.17978 | 0.17978 |
| **40** | 159 | 1 | 142.576 | 20327.9 | 20327.9 |
| **41** | 168 | 1 | 151.576 | 22975.3 | 22975.3 |
| **42** | 138 | 1 | 121.576 | 14780.7 | 14780.7 |
| **43** | 78 | 1 | 61.576 | 3791.6 | 3791.6 |
| **Total** | | 250 | | | 611397 |

III. Decisions

For 'PALocs':

- If I use large size and large number of quadrats, the VMR is less than 1.0, indicates an evenly spaced arrangement;
- If I use large size and small number of quadrats, the VMR is less than 1.0, indicates an evenly spaced arrangement;
- If I use small size and large number of quadrats, the VMR is around 1.0 (sometimes smaller than, sometimes larger than, depending on the random situation), indicates that most points are evenly spaced, but some points have a tendency toward clustering in the pattern;
- If I use small size and small number of quadrats, the VMR is around 1.0. In this situation, the random approach does not capture the feature of the map very well. Therefore, this result is less important than previous.
- **Conclude**: based on VMR, the data set is randomly scattered.

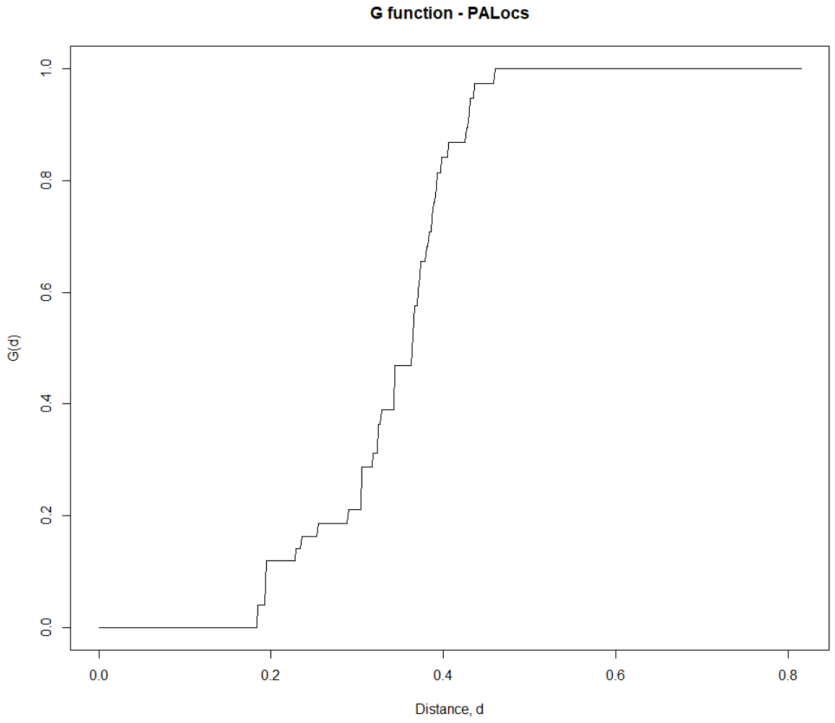
For 'PACoals':

- If I use large size and large number of quadrats, the VMR is larger than 1.0 (VMR > 300 in most cases), indicates a clustered arrangement;
- If I use large size and small number of quadrats, the VMR is larger than 1.0 (VMR > 200 in most cases), indicates a clustered arrangement;
- If I use small size and large number of quadrats, the VMR is larger than 1.0 (around 50 in most cases), indicates a clustered arrangement;
- If I use small size and small number of quadrats, the VMR is larger than 1.0 (fluctuate widely, but always larger than 1.0), indicates a clustered arrangement;
- **Conclude**: based on VMR, the data set is clustered.

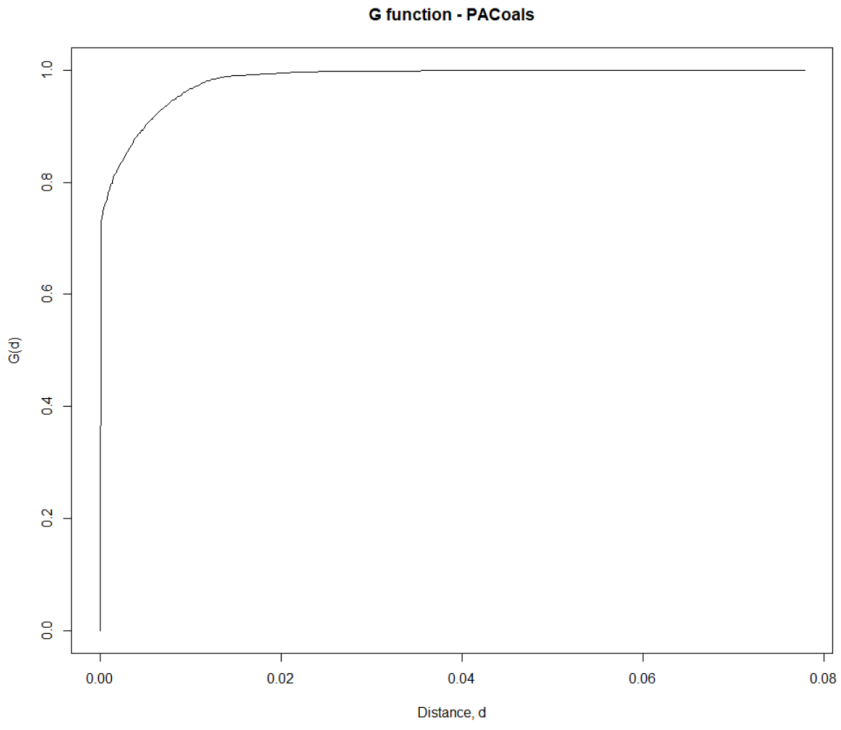
**PART B**

I. G function plots

For ‘PALocs’:

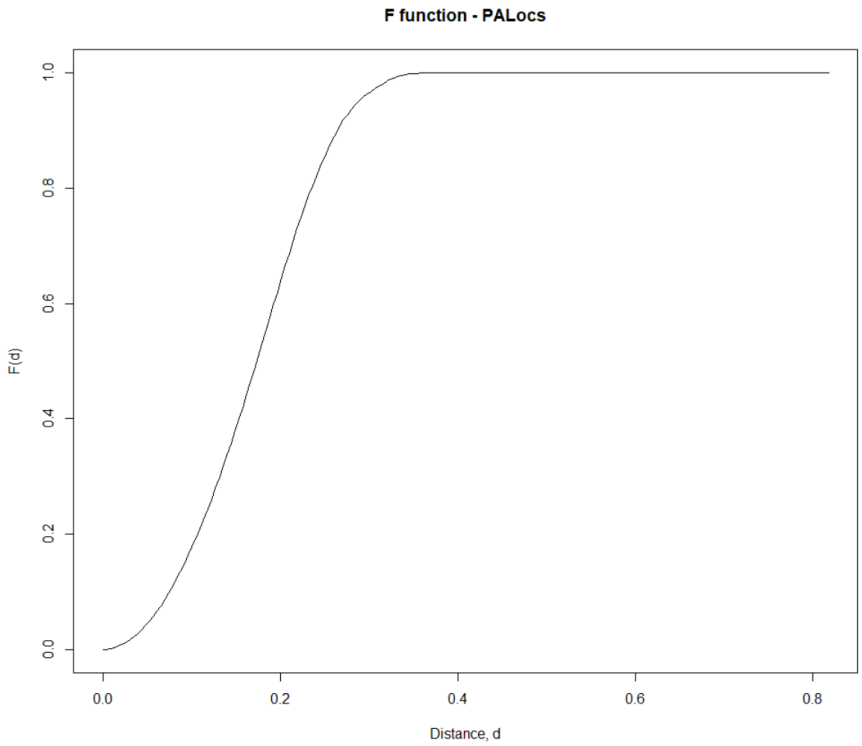



For ‘PACoals’:

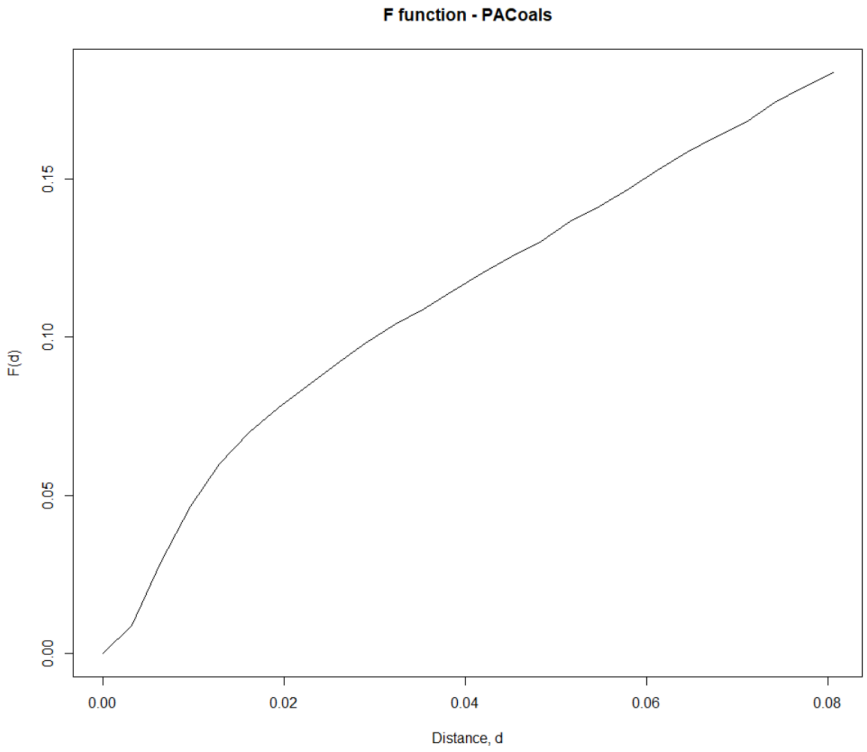


II. F function plots

For ‘PALocs’:



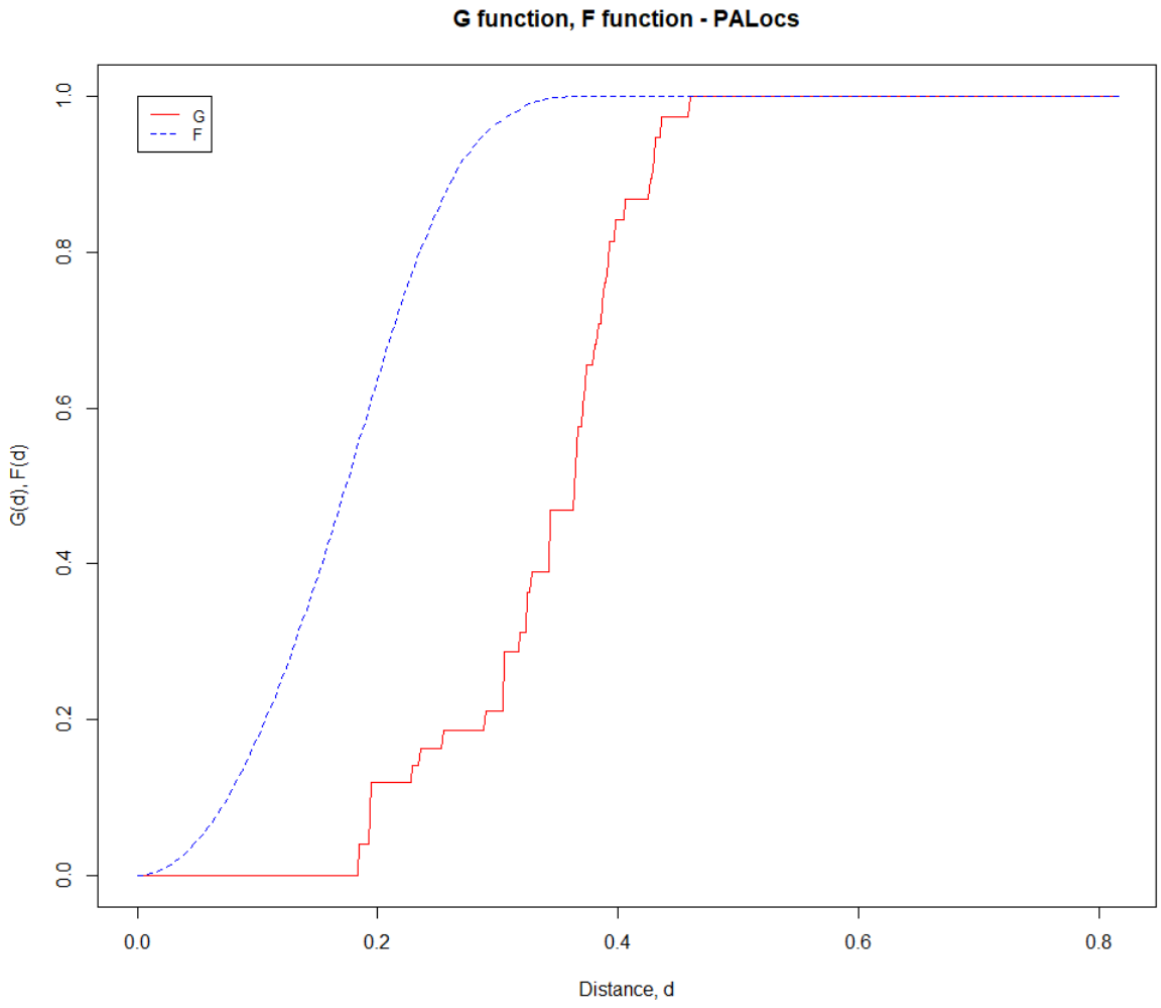

For ‘PACoals’:



III. Comparing the results of G and F function

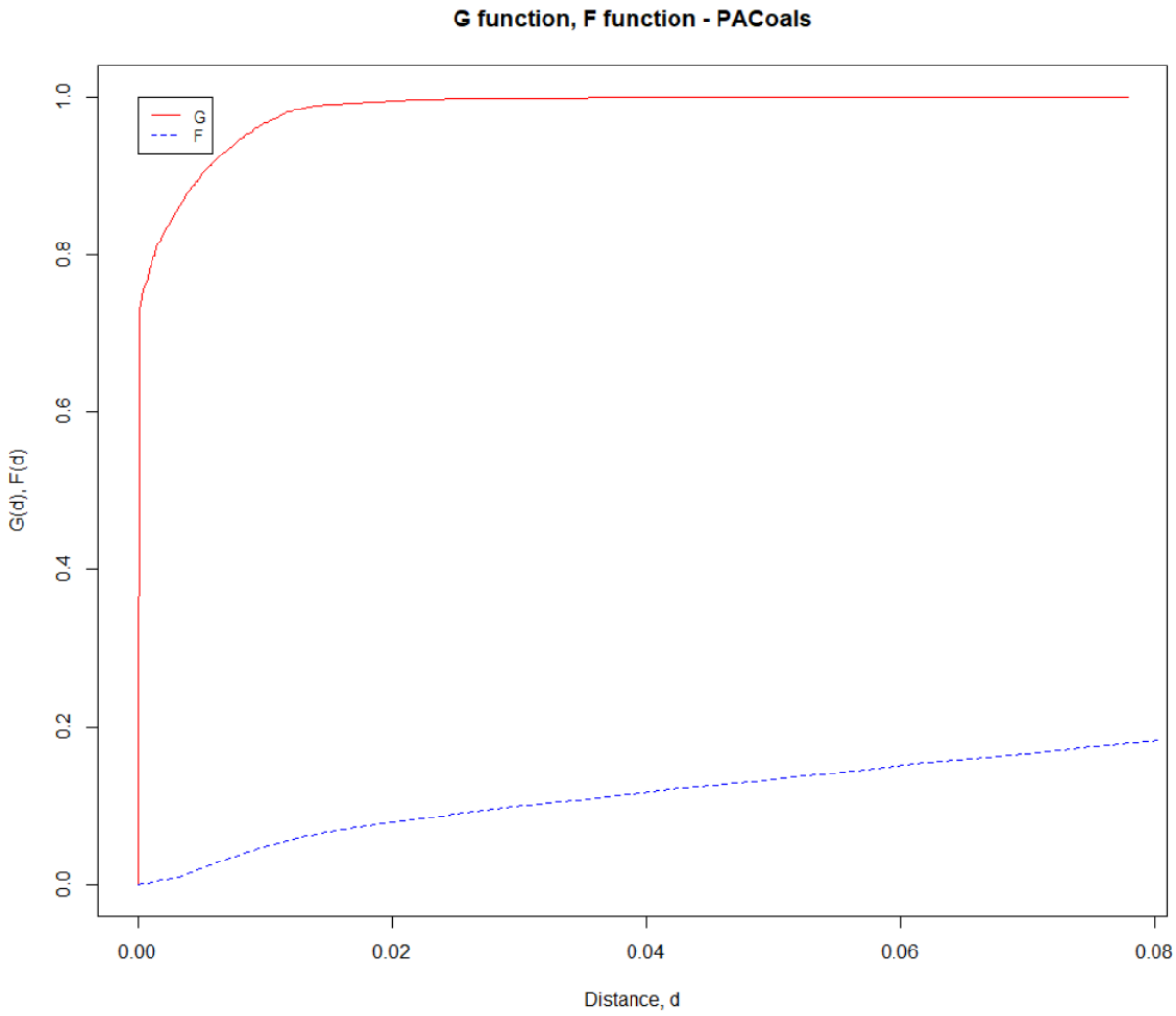
[**Basic knowledge**: if events are clustered in a corner of the study region, *G* **rises sharply at short distances** because many events have a very close nearest neighbor. The *F* function, on the other hand, it likely to **rise slowly at first**, but more rapidly at longer distances, because a good proportion of the study area is fairly empty, so that many locations are at quite long distances from the nearest event in the pattern. For evenly spaced patterns, the opposite is true.]

For 'PALocs':




Form the plot, F function rises faster than G function at the beginning, based on the basic knowledge, it indicates that the data set is evenly spaced.

For ‘PACoals’:




From the plot, *G* function rises sharply and much faster than *F* function, based on the basic knowledge, it indicates that the data set is clustered.