Mars In-Class Discussion

The following is a presentation to elicit discussion of the article, Higher Order Neighbor Analysis of the Tartarus Colles Cone Group, Mars: The Application of Geographic Indices to the Understanding of Cone Pattern Evolution written by Mark Bishop

Outline of Discussion

I intend to:

Summarize the Article Purpose and Justification
List the Analysis Procedure Used in the Article
Identify Shortcomings of the Methodology
List the Main Conclusions of the Article
List students reactions

Summary of Article Purpose and Justification

In Mark Bishop's article, he uses higher order neighbor analysis to analyze whether the Tartarus Colles cone group is spatially random or clustered according to certain cone characteristics.

If clustered or clumpiness of cone groups are identified then we know there is a significant difference from the natural state of randomness.

This significance may lead to further examination of the data within that area as to whether it is volcanic in origin or whether amounts of water and/or ice exist.

Thus, further study of these possible clustered areas may lead to important research in the geologic, geomorphic, and geographic fields relating to Mars.

Analysis Methods Used in the Article

Mark Bishop picks a Poisson distribution and a method (Higher Order Neighbor Analysis) commonly used for this distribution to analyze the problem at hand. He proceeds to calculate the observed Euclidian distances form the center cone to other cones in the area. He then proceeds to calculate the R or ratio of the observed average distances over the expected average distances. He continues to calculate the standard error and the Z-score. Finally he uses the Kruskal-Wallis test and tabulates the results.

Steps to Complete the Analysis

Step one—Are Mars cone groups a countable or continuous variable. In general, the cones act as a single unit that in non-divisible therefore it belongs in the family of discrete distributions.

Some discrete distributions are Binomial, Poisson and Hyper-geometric

Question: For the analysis, the Poisson distribution was selected, why?

1. Infinite Number of Occurrences—The number of cones in the Tartarus Colles group can be infinite. In geometry a plane can have an infinite number of points in it because a point can be infinitely small.

2. Non-Negative Variable—There may be zero cones that exist within a space but not a negative number of them (Crater's not included).

3. Expected Mean Equals Expected Variance—The shape of the Poisson distribution is skewed to the right in such a way that the expected mean equals the expected variance.

4. Stochastic (Involving Randomness)—Cones have a chance of developing in any random order

Step two—Accurately place points on paper (i.e. Create a point layer in GIS according to an established projection.)

Step three—Compute distances form the center cone using the Euclidian distance formula not the Manhattan distance formula.

Step four—Compute R or the ratio of observed average distances to expected average distances.

Step five—Compute standard error.

Step six—Compute Z calc for the Gaussian or Normal distribution. Z calc is the difference between the observed average distances minus the expected average distance divided by the standard error.

We are testing whether out of randomness there is clumpiness in the spatial distribution of cones on Mars.

The anomaly in clumpiness therefore if the Z calc is outside of the Z value parameters we would not only reject the null hypothesis but there is a .025 significance of clumpiness.
.
There is also a probability of type one error or rejecting the claim of randomness when the claim is true.

Step six—Compute Kruskal-Wallis test. The test is used for the purpose of comparing the significance of R-values between higher ordered neighbors and is a non-parametric one way analysis of variance with used ranked data and suits the quantity and assumptions of the cone sample.

Shortcomings of the Methodology

MAUP-Modifiable Aerial Unit Problem

The overall implications of MAUP on this study is: if the aerial unit is drawn differently in terms of size and shape the results may be drastically different even though none of the individuals of which the cone group statistics are based have moved.

Conclusions of the Article

1. When all of the points were evaluated they demonstrated randomness
2. An outlier point exerted influence on the data analysis in the southern region. This outlier demonstrated the pervasive influence of the MAUP problem. When the outlier was dropped by one point there was an aerial decrease of 16%. Now the data set demonstrated randomness instead of clustered.
3. The difference in the spatial statistics may result from the categorization of these geomorphic features.
4. In the northern regions clustering was well defined
5. The trend of the neighbors progressively tends toward randomness.
6. Other summary points (1-7) are listed on page 82 of the article.

Student Reactions

All the questions posed to the class were answered in the discussion.

Question: After reading this article or after hearing this summary, does anyone have any ideas of other areas where nearest neighbor analysis can be studied or applied?

A student mentioned that nearest neighbor analysis is used in remote sensing.

Question: For the analysis, the Poisson distribution was selected, why?

A student said the distribution is discrete and countable.

Question: Why use the Euclidian distance over the Manhattan distance?

A student said Euclidian distances calculate as the bird flies.

Question: Name some other instances in projects completed at the University where you encountered the MAUP.

No one directly answered this one. My response was. A possible answer to this question may be the use of census data. Some such boundaries are census tracts, blocks and block groups.

Overall student agreed the paper was well written because the methodology was so explicit. But, they also noticed how the paper was all about the technique and not so much about earth shaking conclusions.

Other miscellaneous comment made by students and professor is:

The second order derivations could have sufficed in the calculation.

Formulas that have constants should leave them as whole number and not over simplify the equation. It the equation is oversimplified. Readers do not know what is constant and what is variable in the equation.