Spatial Statistics For Identifying Non-Primary And Saturated Impact Crater Populations

J.D. Riggs, S.J. Robbins, M.R. Kirchoff, E.B. Bierhaus, B.P. Weaver

Issues in Crater Studies and the Dating of Planetary Surfaces

May 19 - 22, 2015

NORTHWESTERN UNIVERSITY | SCHOOL OF PROFESSIONAL STUDIES

Presentation Outline

Introduction

- Cratering process assumptions
  1. Craters form stochastically through time
  2. Craters form randomly (mostly) across a surface
  3. Impact cratering is an ongoing process
- Two cratering processes tied to stochastic assumptions
  1. Secondary cratering
  2. Crater saturation
- Spatial statistics to identify secondaries and saturation
  1. Familiar methods
  2. Esoteric methods

Context

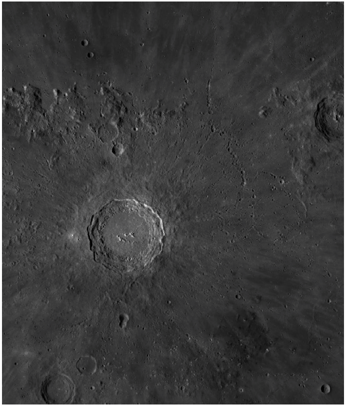
- Secondary cratering
  - Secondaries exhibit non-random patterns
    assumption 2 violation
  - Task is identification of secondaries from spatial statistics
- Crater saturation
  - New craters compromise the structure of an existing crater
  - Crater spatial population distributions static in time and space
    assumption 3 violation
- Focus of this presentation is on secondary cratering

Impacts

- Primary impacts
  - Complete spatial random (CSR) distributions (some deviation)
  - CSR follows a Poisson probability distribution function (pdf)
  - Poisson pdf is counts by area (univariate spatial density)
- Small crater impacts include
  - Primaries
  - Secondaries
  - Sesquinary secondaries (maybe)
- Primaries and sesquinary secondaries considered CSR
- Secondaries can cluster or appear as CSR

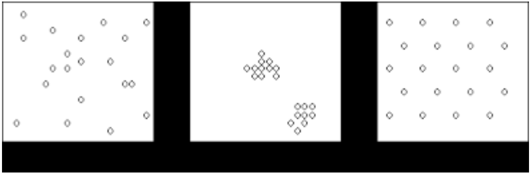
Secondary Impacts: Morphology

- Near primaries tend to be unambiguous
- Distant more challenging to identify
- Sesquinary secondaries usually indistinguishable from primaries



Spatial Distributions

- Left panel CSR
- Middle panel clusters
- Right panel regular (uniform)



Familiar Statistical Methods

- Region divided and CSFDs compared
- CSFD slopes
- Nearest neighbor distance (NND) with Z-statistic

CSFD Method

- Univariate spatial density comparisons among region subdivisions
- Sub-region differences suggest secondary contamination
- Similarities suggest no contamination or CSR ambiguity
- Steeper CSFD slopes (than primary SFDs) suggest secondaries

CSFD Method: Issues

- Dependent on region characteristics
- Sensitive to CSFD uncertainties
- Specificity poor
- Spatial distribution identification challenging
- Spatial correlation ignored
- CSFD comparison and slope, may be improved using CEDF

Distance Method

- Nearest Neighbor Distance (NND)
- Uses location (lat/lon), independent of region characteristics
- Under CSR, counts pdf transformed to univariate distance pdf

$$1 - \mathcal{P}(N = 0) = 1 - \frac{\rho \pi r^2}{0!} e^{-\rho \pi r^2}, \text{ NND} < r$$

- Distances are Z-scores so standard deviations from CSR
  - $Z < 0 \Rightarrow$ clustering
  - $Z = 0 \Rightarrow$ CSR
  - $Z > 0 \Rightarrow$ uniformity
  - Location uncertainty quantifiable

Distance Method: Issues

- Spatial correlation information lost
- Only smallest scale information, no information on all scales
- Only magnitude and direction of deviations from CSR
- Affected by boundaries, though mitigation possible

Less Familiar Methods

- Cluster methods
- Ripley's K-function
- Besag's L-function
- Pair Correlation Function

Cluster Analysis Methods

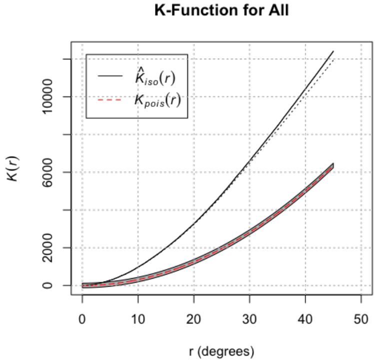
- Exploratory method for spatial data
- Groups craters by specified homogeneous properties
- Common techniques
  - Agglomerative hierarchy
  - K-means (K specified)
  - Categorical MLE (K estimated)
- Useful for visualizing spatial relationships
- Visualization of probabiity of a crater as a secondary

Spatial Point Methods

- Intensity function measure of counts / unit area: $\lambda = \frac{N}{|A|}$, (counts, N, for normalized area, $\mid A \mid$)
- Implementations using λ (specified) or $\hat{\lambda}$ (estimated)
  - Ripley's K-function
  - Besag's L-function
  - Two point correlation function
- These methods identify
  - CSR
  - Regularity
  - Clustering

Ripley's K-Function: $K(r) \equiv \lambda^{-1}\mathcal{E}$ (# nonrandom events $0 \leqslant r \leqslant h$)

- Captures spatial areal dependence (correlation)
- Red dashed curve is CSR with 0.05 uncertainty region $K(r) = \pi r^2, \;\; r \geqslant 0$
- Black solid curve shows clustering at scales $> 5^\circ$ $K(r) > \pi r^2$ (dashed curve has boundary correction)
- If black curve below CSR region, distribution regular $K(r) < \pi r^2$

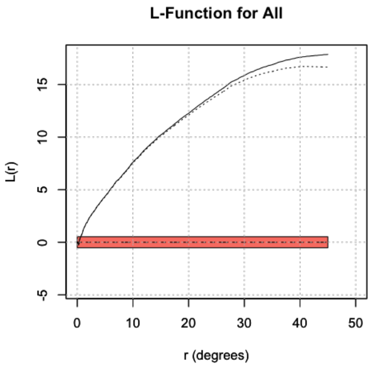


Ripley's K-Function

- K indepdent of region shape
- K has region boundary accountability
- Information exists on all scales
- Accounts for lat/lon uncertainties
- Accounts for spatial correlation

Besag's L-function: $L(r) = \sqrt{K(r)/\pi} - r, r \geqslant 0$

- Function of Ripley's K, has same information
- Plot easier to interpret than K-function
- Zero-line and uncertainty represents CSR
- Excursions above the zero-line imply clustering $> 5^\circ$
- Excursions below the zero-line imply regularity



Besag's L-function: $L^*_{12}(r; n_1, n_2)$

- Adapted to address clustering interactions
  Two or more populations, e.g., same-sized craters of primaries and secondaries

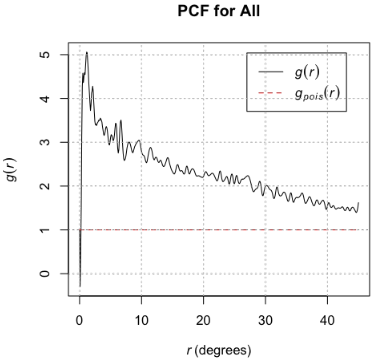
- Counts of 1 population assessed around counts of another population
  - >0 segregated
  - <0 aggregated
  - $=0$ no interaction

- Can help identify types in mixes of same-size primaries and secondaries

Pair Correlation Function

- Counts events in annuli rather than circles about a reference
- Related to the differential of Ripley's K-function
- Derives from the joint probability 2 events lie in infinitesimal annuli of areas dA_1 and dA_2 giving the correlation $\xi(r)$

$$\xi(r) = \frac{\lambda(r)}{\bar{\rho}^2} - 1$$



Summary

- Clustering methods for
  - Data exploration
  - Spatial dependence visualization
- Pair correlation function
  - Contains the same information as K- and L-functions
  - Requires binning, hence less desireable

Summary

- *K*- and *L*-functions
  - Like CEDF, not binnied
  - Unbaised estimation of spatial correlation
  - Spatial correlations at all scales
- Lattice spatial methods
  - Independent of lat/lon
  - Utilizes 2-D NND techniques that include spatial correlation
  - Untried? for secondary crater identification

Conclusions

- Existing implementations of spatial methods
  - Alone will not discriminate unambiguously whether a crater is a secondary
  - Mathematics can be reformulated to assign a probability a crater is a secondary
  - Mathematics can be reformulated to assign a crater as a secondary with specified uncertainty