Questions

Mapping the true spatial form of cities and settlements (as opposed to administrative boundaries) provides a functional representation that informs our understanding the underlying processes.

Mapping the past growth and evolution of cities and settlements informs our understanding of how they may grow and evolve in the future.

Static

How does the spatial form of agglomerations differ from isolated cities?
How are cities and settlements distributed on the landscape?

Dynamic

Where, and at what rate, is urban growth and development occurring now?
Does urban growth conform to the predictions of economic theory?
Is urban growth predictable?

Do systems of human settlements exhibit spatial scaling analogous to the power law scaling observed for cities as defined by population?

Zipf’s Law

A power law with an exponent of ~1.

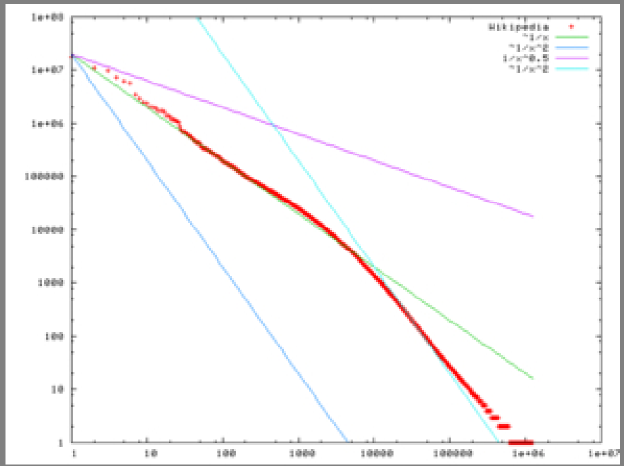
“Zipf's law, named after the Harvard linguistic professor George Kingsley Zipf (1902-1950),is the observation that frequency of occurrence of some event (P), as a function of the rank (i) when the rank is determined by the above frequency of occurrence, is a power-law function $P_i \sim 1/i^a$ with the exponent a close to unity (1).” http://www.nslij-genetics.org/wli/zipf/

“Zipf's law states that given some corpus of natural language utterances, the frequency of any word is inversely proportional to its rank in the frequency table. Thus the most frequent word will occur approximately twice as often as the second most frequent word, which occurs twice as often as the fourth most frequent word, etc.” http://en.wikipedia.org/wiki/Zipf's_law

The Problem: Inconsistency. Sometimes it works, sometimes it doesn’t. Hence, the controversy. Within individual countries, “King Cities” are often underpredicted. At global scales, the largest cities are overpredicted. Exponents often vary among countries and through time.

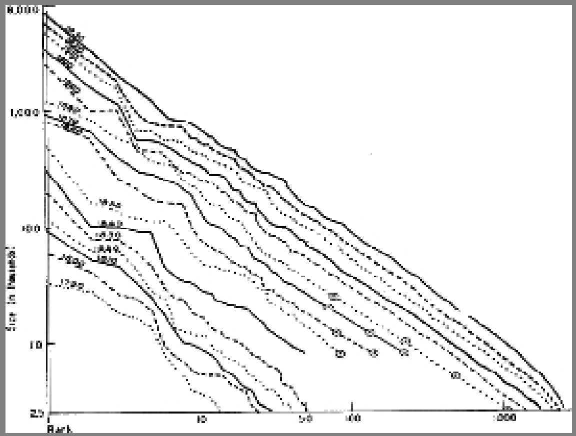
Why? Varying definitions of urban. Administrative Fragmentation. Arbitrary criteria for fit.

Words



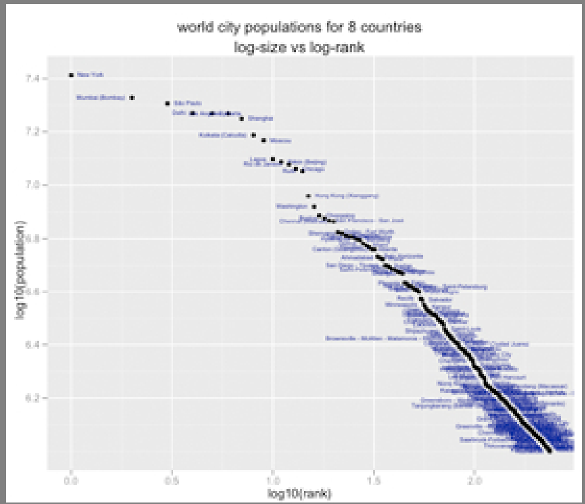
Rank-frequency plot of word frequency in Wikipedia (11/27/2006). http://en.wikipedia.org/wiki/Zipf's_law

U.S. Cities



Original Zipf plot using incorporated places in the U.S.A. from 1790 to 1930.

World Cities



http://anyall.org/blog/2009/05/zipfs-law-and-world-city-populations/

Zipf Testimonials

What Zipf said:

George Zipf: *"Specialization of enterprise, conditioned by the various advantages offered by a non-homogeneous terrain, naturally presupposes an exchange of goods…"* G. K. Zipf (1949) Human Behavior and the Principle of Least-Effort. Addison-Wesley

What the luminaries say:

Xavier Gabaix *"Zipf's law for cities is one of the most **conspicuous** empirical facts in economics – or in the social sciences generally."* Quarterly Journal of Economics 114 (1999), pp. 739..

Paul Krugman: "*...rank-size rule is a **major embarrassment** for economic theory: one of the strongest statistical relationships we know, lacking any clear basis in theory*"
Development, Geography, and Economic Theory, 1994, pp. 44.

Steven Strogatz: *"...a **beautiful** law of collective organization that links urban studies to zoology. It reveals Manhattan and a mouse to be variations on a single structural theme."* http://opinionator.blogs.nytimes.com/2009/05/19/math-and-the-city/

Cosma Shalizi: *"So you think you have a power law. Well, isn't that **special**."*
http://www.cscs.umich.edu/~crshalizi/weblog/491.html

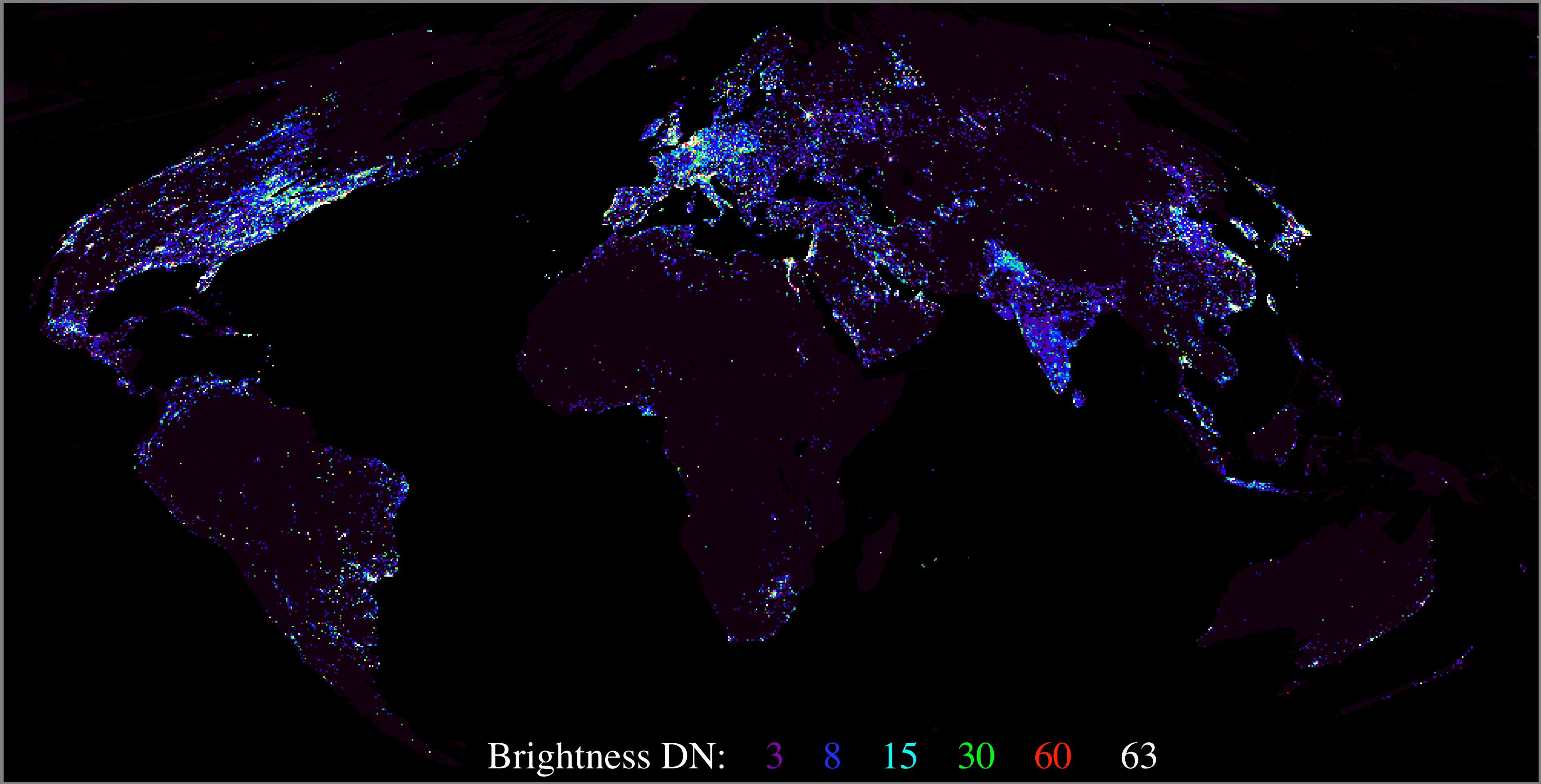
Frank Small: *"That's **weird** - but **interesting**."* Personal Communication

Data & Strategy

Quantify changes in spatial distribution, and intensity, of anthropogenic land surface modification using:

1) New inter-calibrated DMSP stable night light composites between 1992 and 2008. *Spatial resolution ~2.5 km.*
2) LandScan 2008 ambient population model of spatially disaggregated census data. *Spatial resolution ~1 km*
3) GPW3 gridded population data of maximally dispersed census data *Spatial resolution variable μ = 18 km*
4) GRUMP partially disaggregated population data *Spatial resolution variable - gridded at ~2.5 km*
5) MODIS-derived urban land cover mask *Sensor resolution 500 m Minimum mapping unit > 1 km^2*

Segment each data set at multiple thresholds & calculate rank-size distributions of contiguous segments.

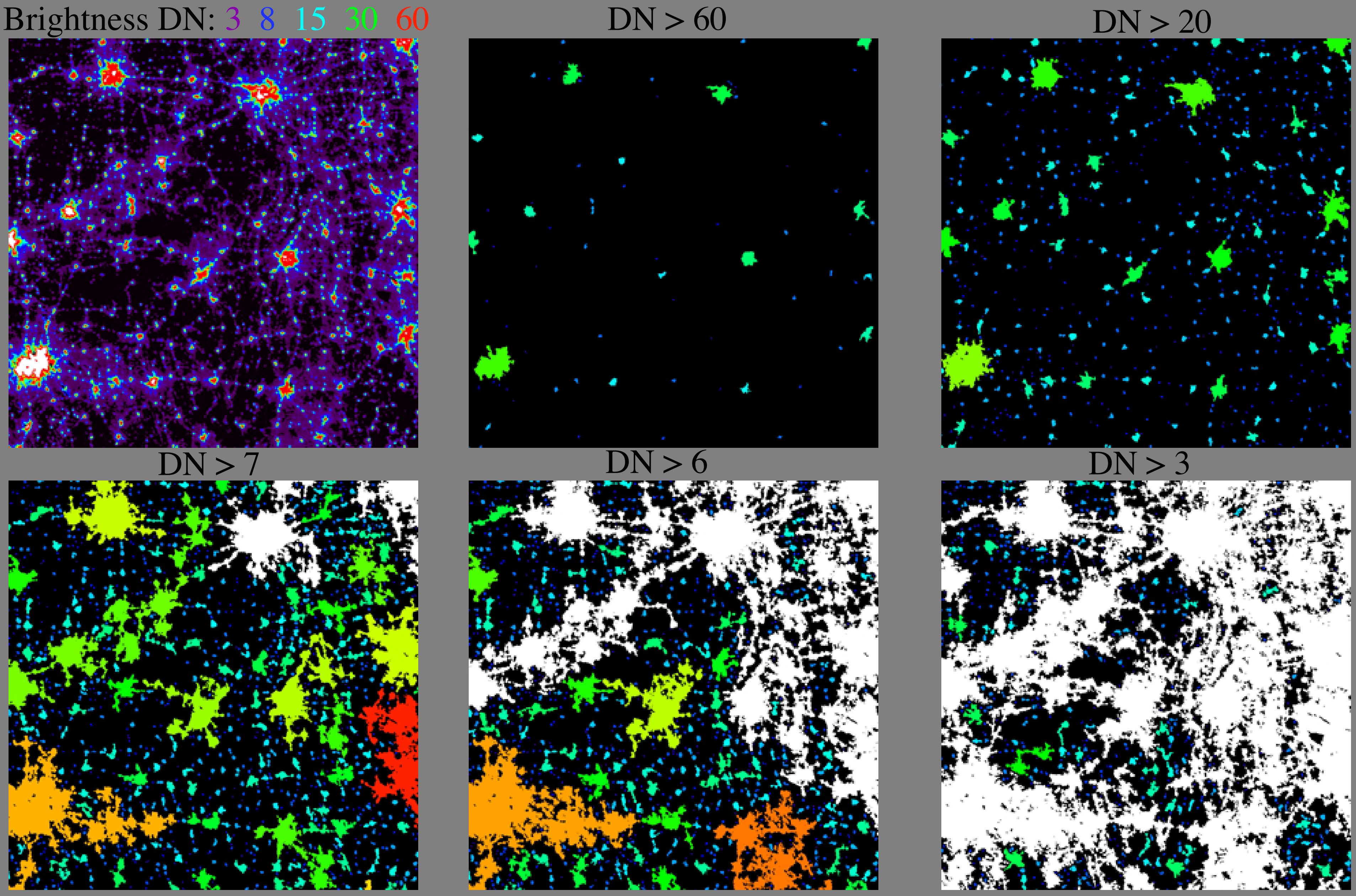


Spatial Connectivity of Continuous Fields

1) Quantify spatial extent and intensity of development by brightness thresholding.

2) Quantify size distribution of spatially contiguous patches of developed land.

Patch Size: *1 1.5 2 3.5 4.25 4.5 5 5.5 6* ***Log$_{10}$(km^2)***



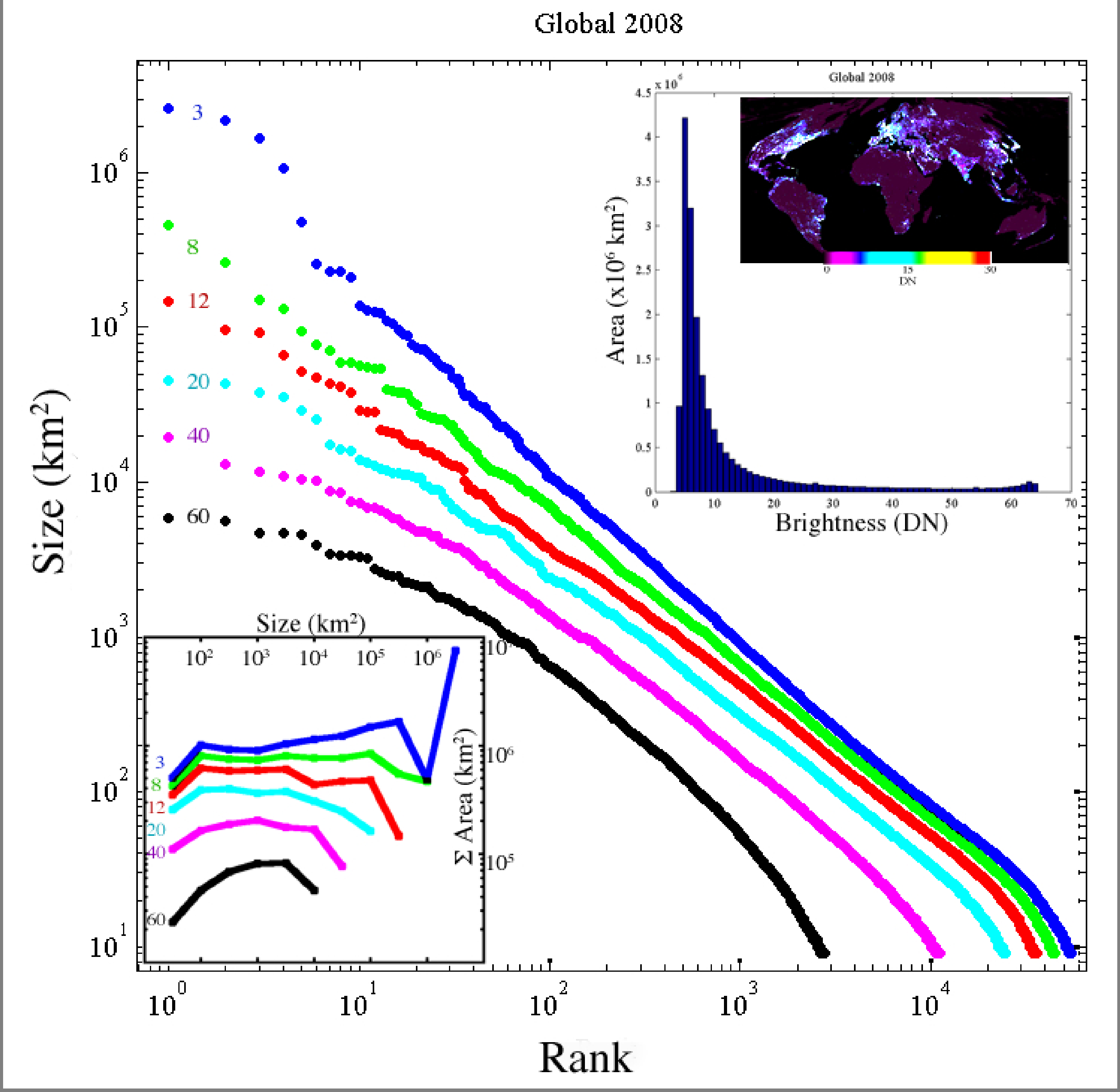
Effects of Low Light Thresholding on Size Distributions

Most lighted area is dim DN < 20

Low thresholds have largest effect on size distribution and patch contiguity.

Low thresholds produce linear rank-size in Log_{10} space - even in upper tail - but with *varying slopes*.

Higher thresholds affect both ends of distribution: *attenuating* small lights & *fragmenting* large clusters



Is it really a power law?

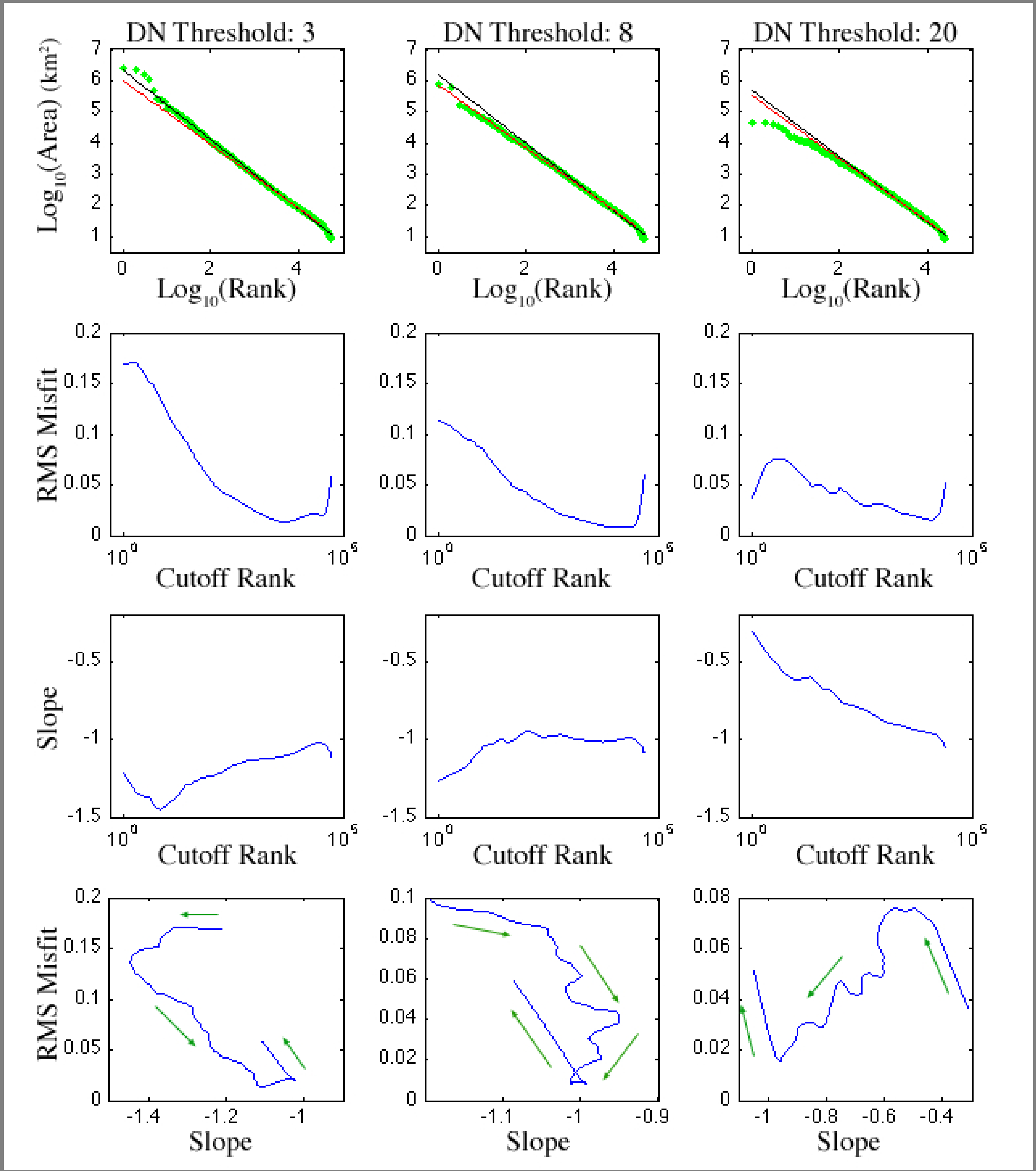
Naive estimation (OLS in Log_{10}) gives reasonable (but inaccurate) results.

Roll-off in lower tail torques slope down: Fit upper tail above some cutoff.

Find optimal cutoff from RMS misfit

Use Hill plots to track exponent

What exponent minimizes RMS misfit?
How much of the tail is linear?

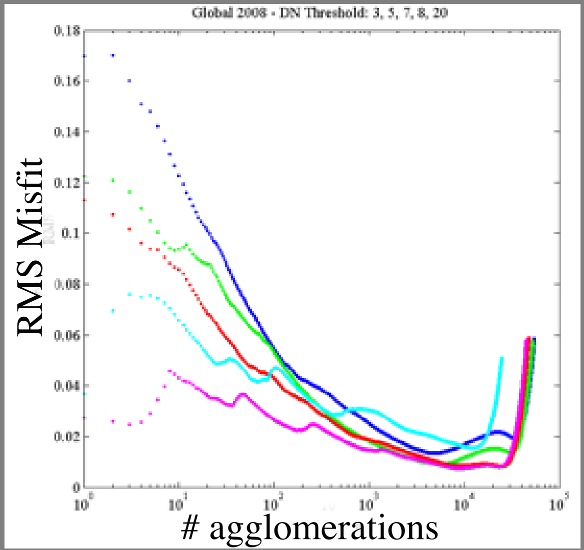


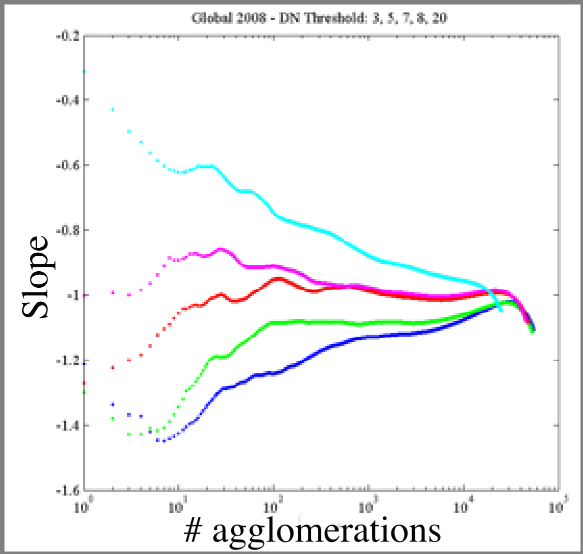
Weird Scenes in Parameter Space

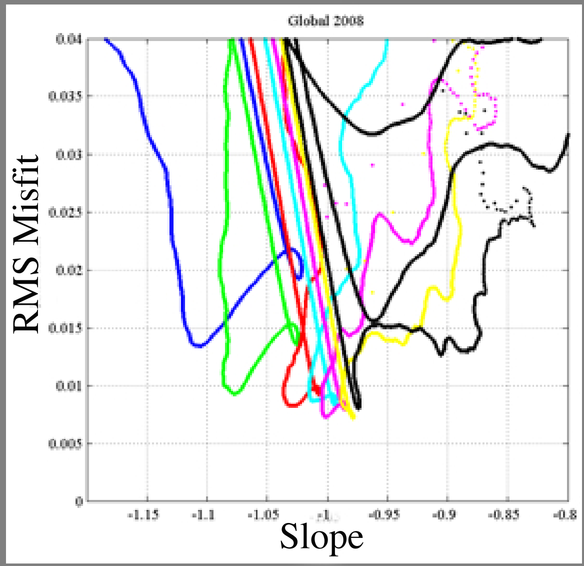
Tracking RMS misfit and Slope for increasing upper tail size quantifies incremental effects of non-linear scale ranges.

3 Interesting Observations

Multiple local minima for thresholds DN<8 & DN>12
Smallest misfits coincide with well-defined minima near Zipf exponent (-1)
Transitional exponents near Zipf (8>DN>12) have small misfit & slopes near -1 over wide range of tail sizes up to divergence in lower tail.



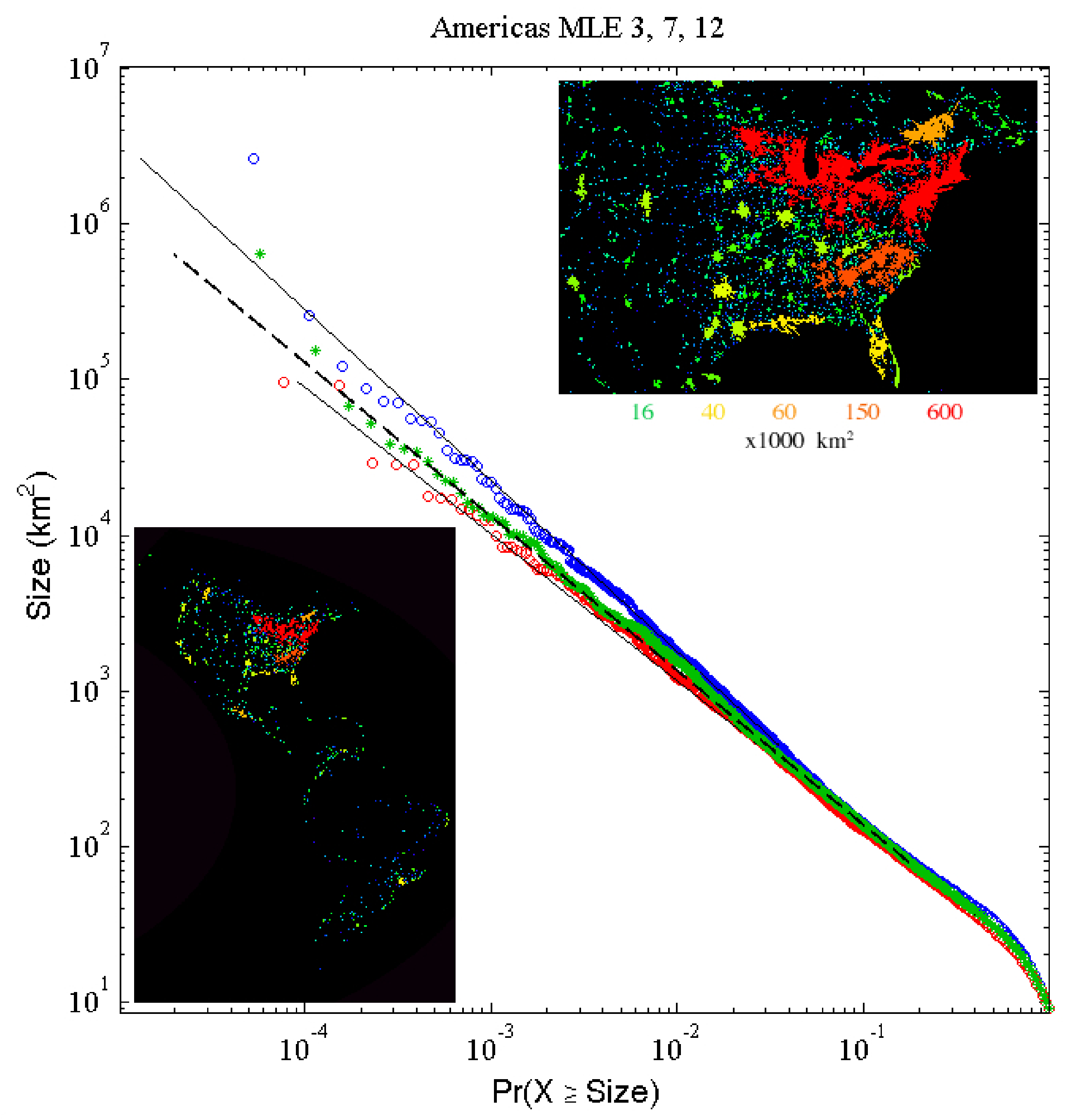




Zipf Distributions - Americas

Good linear fit in upper tail - except for the largest agglomeration

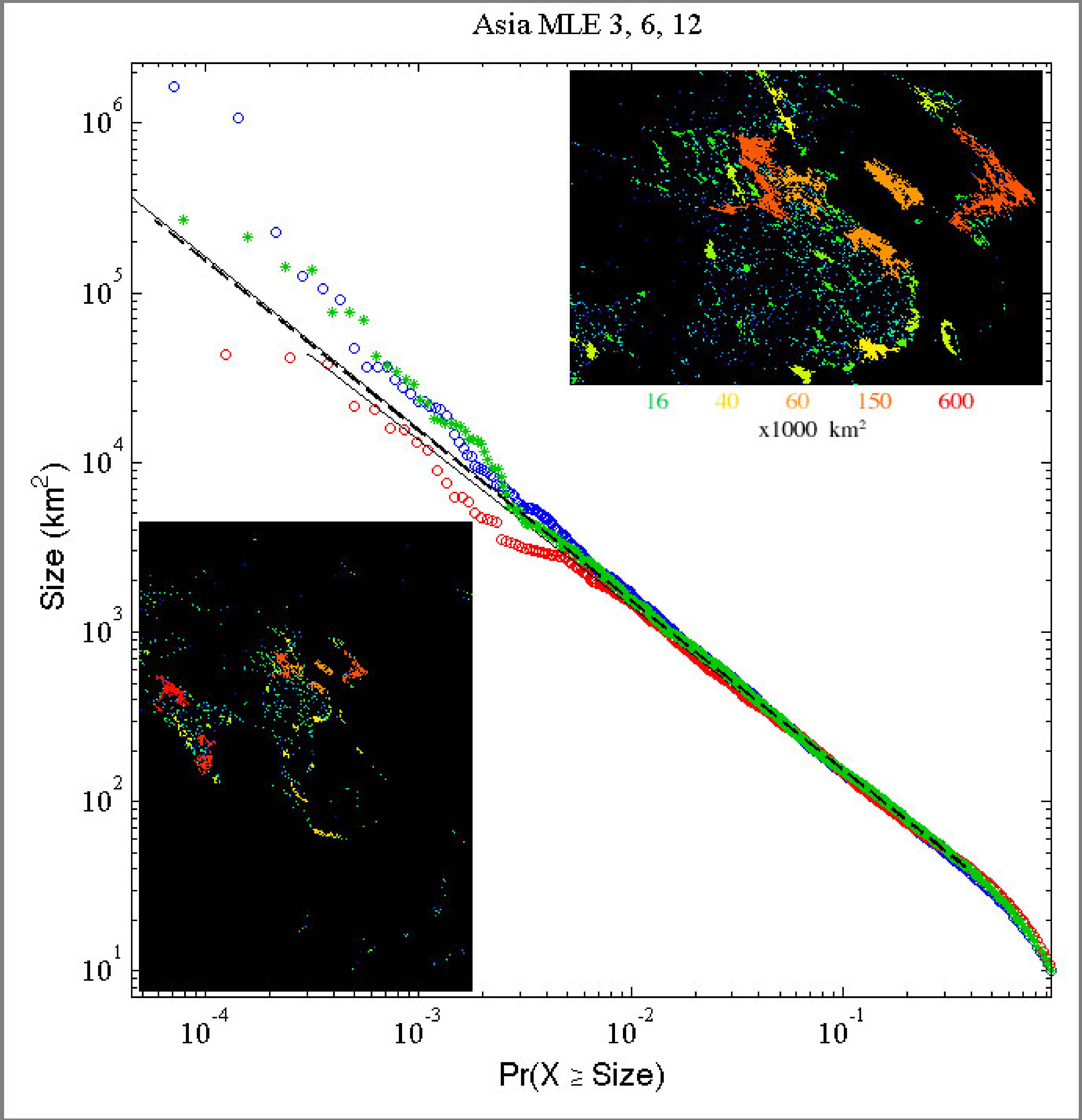
Pronounced change in slope across Zipf exponent



Zipf Distributions - Asia & Oceania

Upper tail not well fit.
Many excedences.
Divergent upper tail for higher & lower thresholds.

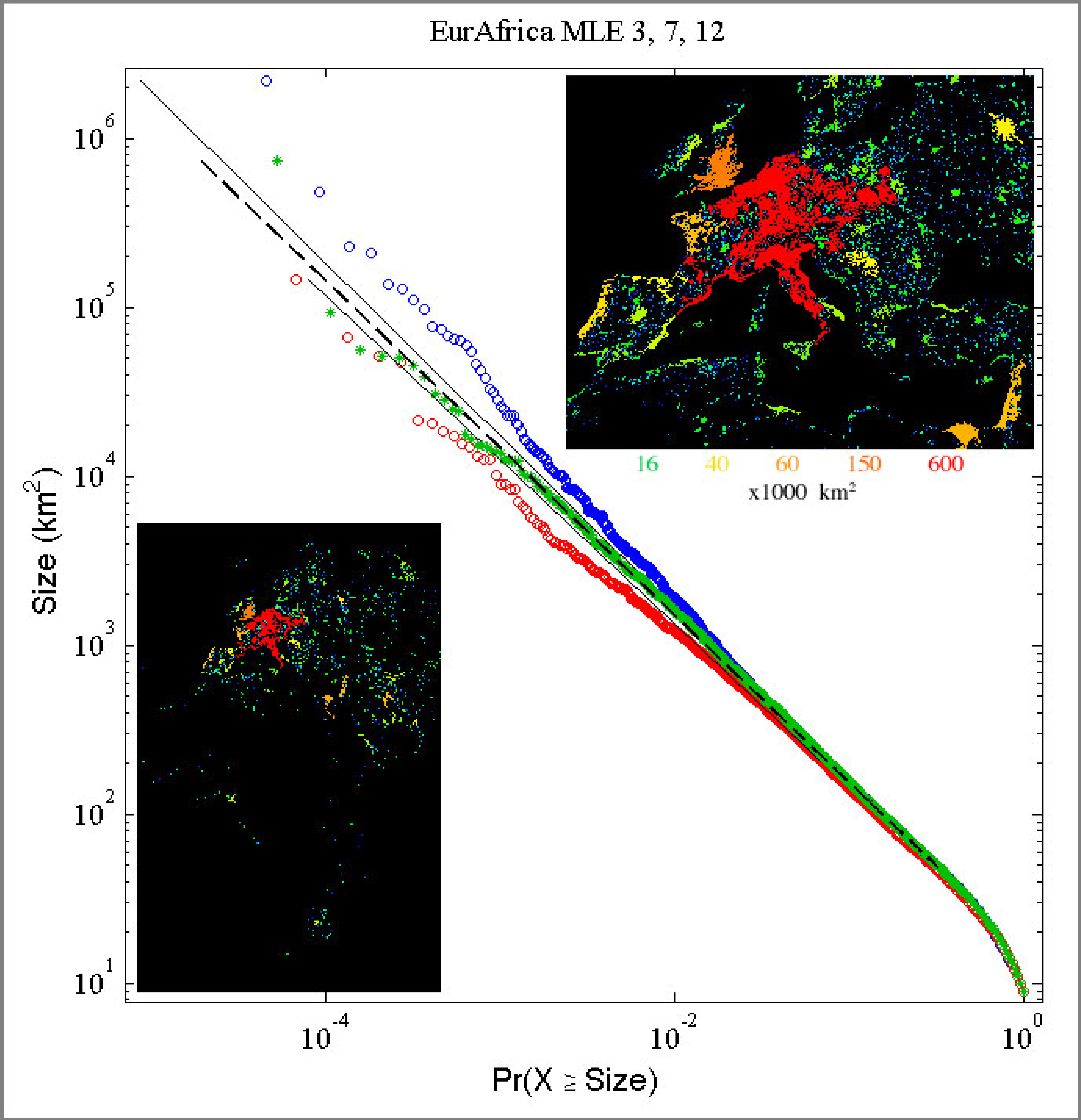
Little change in slope across Zipf exponent



Zipf Distributions - Eurasia/Middle East/Africa

Good linear fit for Zipf case (except largest) but divergent upper tail for higher & lower thresholds.

Little change in slope across Zipf exponent



Global Zipf Distributions

Upper tail very well fit.
for Zipf case.
No excedences.
Linear upper tail for
higher & lower thresholds.
Few excedences even for
DN = 3.

Large change in slope
across Zipf exponent