City–City Correlations to Introduce Galaxy–Galaxy Correlations

Daniel M. Smith, Jr.
Department of Biological and Physical Sciences, South Carolina State University
dsmith@scsu.edu

Keywords
Large-scale Structure, Cosmology Lab, Galaxy Distribution, Baryon Acoustic Oscillation, Galaxy–galaxy Correlation Function

Summary

The large-scale structure of the Universe, vividly displayed by the spatial distribution of galaxies, is characterised quantitatively by the two-point galaxy–galaxy correlation function. But the meaning of the correlation function is somewhat abstract because it does not have a ready analogy. This work computes the two-dimensional, two-point city–city correlation function for three populous regions of the United States, demonstrating that the city–city correlation function is analogous to the galaxy–galaxy correlation function determined from Sloan Digital Sky Survey data. City radii are analogous to galaxy cluster radii, and city-to-city distances are analogous to distances between galaxy clusters. Part of this work has been adapted for a lab suitable for non-experts.

Introduction

The large-scale structure (LSS) of the Universe — essentially the distribution of galaxies — is characterised by the two-point correlation function (Peebles, 1980). Its applications include determining the percentage of dark matter (Peacock et al., 2001; Hawkins et al., 2003) and demonstrating baryon acoustic oscillations (Eisenstein et al., 2005) as predicted by the dark energy–cold dark matter (ΛCDM) model of the Universe. However, the concept of the correlation function might seem obscure to the uninitiated. This motivated the current work to calculate, for three groups of cities in the United States of America, a two-dimensional, two-point correlation function that can be easily interpreted because of readily available city data.

This provides a tool for non-experts to interpret the three-dimensional two-point galaxy–galaxy correlation function calculated from Sloan Digital Sky Survey (SDSS) data or from other surveys. The question posed and answered in this work is the following: in what sense, mathematically, does the clustering of galaxies in Figure 1 resemble the clustering of cities in Figure 2?

Two-point correlation function

If galaxies were randomly distributed in the Universe, there would be a certain probability of finding two galaxies near each other. That probability is enhanced by the gravitational attraction between two galaxies, an enhancement called the two-point correlation function, represented by the dimensionless ξ. The method for calculating the two-point correlation function is clear in theory, if not in practice. First, generate a random catalogue of galaxies with the same spatial extent as the galaxy catalogue under analysis. Then determine the degree to which galaxy clustering of the actual catalogue is enhanced over that of the random catalogue.

City–city two-point correlation function

A method for calculating the galaxy–galaxy correlation function (Zehavi et al., 2002) was adapted to the problem of determining the city–city two-point correlation function for three groups of cities. The composite satellite image poster, *North America at Night* (Sullivan, 1993), was digitised, and cities represented by a blob of light were identified by comparison with a standard geographical map.

Three regions were chosen as representatives of clusters of cities. The criteria for choosing a region were:

1. That its shape must be square, for ease of analysis.
2. That two other non-overlapping regions could be chosen of the same size and approximate light density.

The Midwest, Southwest, and Southern regions of the United States chosen (highlighted squares in Figure 2) are approximately 414 000 square kilometres in area.

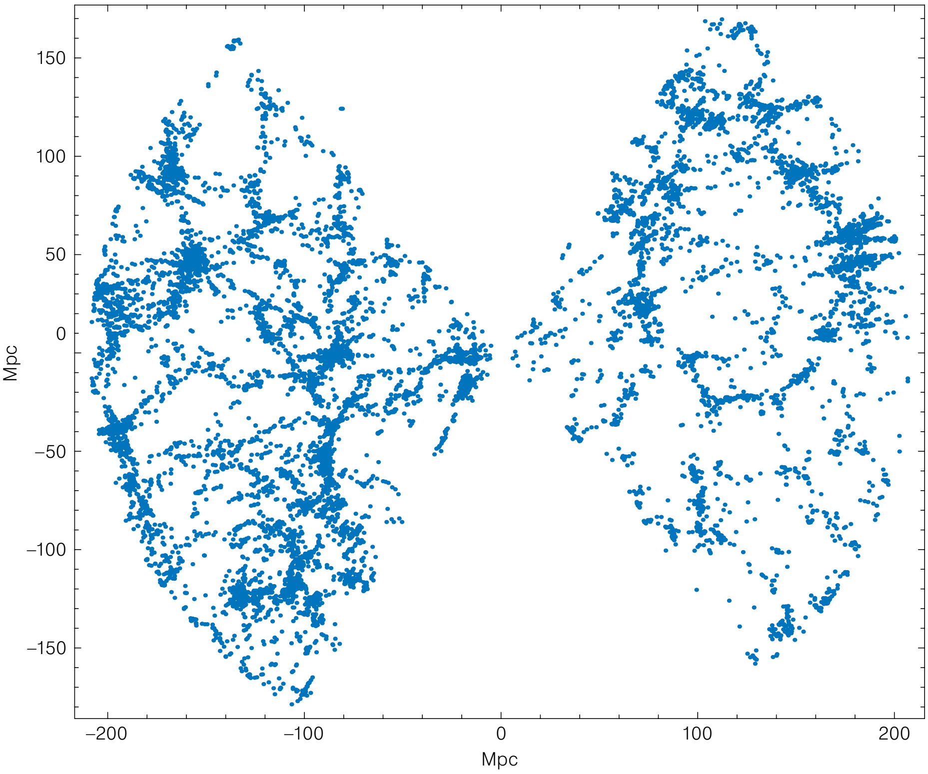

Figure 1*.* *Plot of 9659 galaxies from the SDSS Data Release 7 for 0 < z < 0.05 and –3° < dec < 3°.*

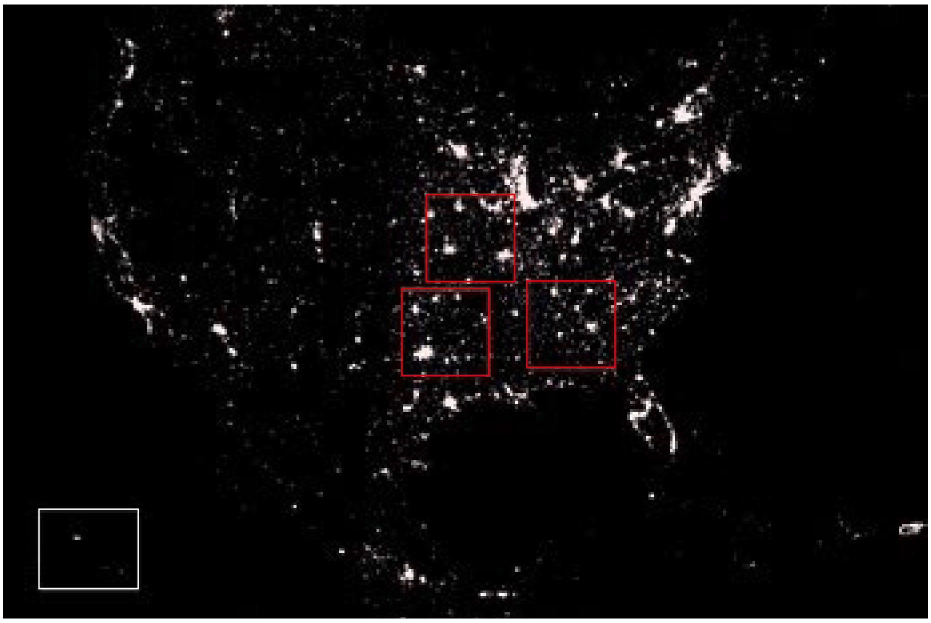

Figure 2*.* North America at Night *with the three squares indicating the cities grouped for analysis. Credit: Sullivan 1993.*

The correlation function is calculated for the Midwest cities, shown in Figure 3. The result is displayed in Figure 4 and is readily interpreted when compared to independent measurements.

City lights are initially positively correlated, a correlation that decreases and becomes negative as the distance from the city centre increases, but becomes positive again, reaching its peak ($\xi = 0.12$) at 389 ± 10 km when another city is encountered.

This value compares favourably to the 367 ± 132 kilometres determined by simply averaging all of the possible distances between all of the cities, using geographical data. Furthermore, a fit of the initial points of Figure 4 to a power law reveals that the correlation function decreases initially over a characteristic distance of 33 ± 0.8 kilometres, comparable to the average radius (32 ± 8 kilometres) of all of these Midwest cities as determined by making on-screen pixel measurements then converting to kilometres.

The same correlation function analysis is performed for the Southwest cities (Oklahoma City, and Tulsa, Oklahoma; Fayetteville and Little Rock, Arkansas; Dallas, Texas; and Shreveport, Louisiana), and Southern cities (Nashville, Knoxville, and Chattanooga, Tennessee; Birmingham and Montgomery, Alabama; and Atlanta, Georgia). Just as before, for each group there is a peak in the correlation function

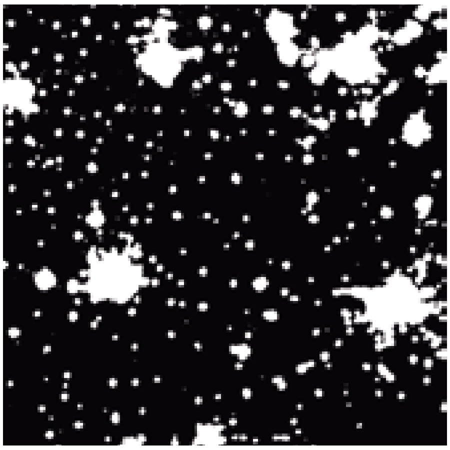

Figure 3*.* *The correlation function is calculated for this group of Midwest cities, the top-most square in Figure 2. Clockwise, from the upper-left corner: Omaha, Nebraska; Des Moines, Cedar Rapids, and Davenport, Iowa; and St. Louis, Springfield and Kansas City, Missouri.*

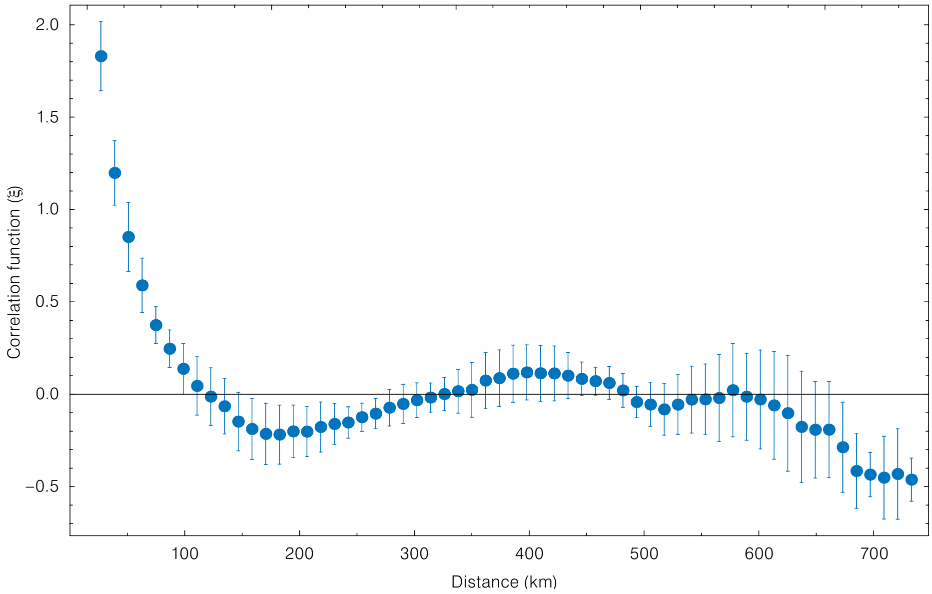

Figure 4. *Two-point correlation function for the Midwest cities in Figure 3.*

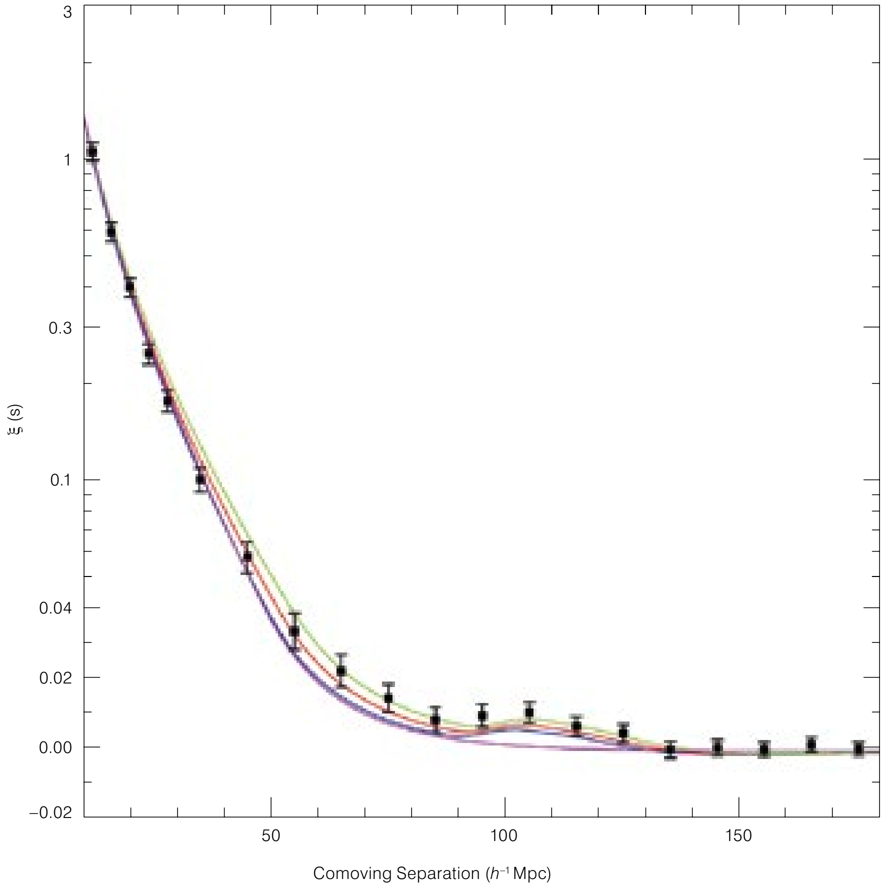

Figure 5. *Two-point correlation function for Luminous Red Galaxies from the SDSS survey. Curves are predictions for various fractions of dark matter. Adapted from Eisenstein et al. (2005).*

that corresponds to the average city–city distances, and the characteristic distance over which the correlation function decreases is the same as the average of the city radii. Average distances between cities, and average city radii for the three groups are summarised in Table 1.

Galaxy–galaxy two-point correlation function

Following the above interpretations, with a couple of caveats, the peak in the galaxy–galaxy correlation function of Figure 5 (Eisenstein et al., 2005) reveals an average distance between galaxy clusters of about 100 h^{-1} Mpc (≈ 143 Mpc for a Hubble parameter $h = 0.7$). And a galaxy cluster's average radius is approximately 8 h^{-1} Mpc (≈ 11 Mpc), according to the correlation function of Zehavi et al. (2002). Both values are from analyses of the Sloan Digital Sky Survey (SDSS) catalogue. Zehavi et al. (2002) includes a summary of correlation functions from other surveys.

The first caveat is that the correlation function peak in Figure 5 is for Luminous Red Galaxies (LRGs), which are particularly suited for galaxy clustering studies, but the galaxy cluster radii value is from a more general catalogue of galaxies. Zehavi et al. (2005) also gives the correlation function for LRGs.

The second caveat is that these galaxy–galaxy correlation functions do not represent real-space distances, but the co-moving distances used by cosmologists to give a meaningful measure of distance in an expanding Universe. These distances expand in step with the expanding Universe and depend on the redshift and the choice of the underlying cosmological parameters. These caveats are not detrimental to the analogy between the city–city and galaxy–galaxy correlation functions.

The peak of Figure 5, called the baryon acoustic peak, is of cosmological importance because it is an imprint of the oscillating plasma of baryons (protons, neutrons — essentially ordinary matter) that was coupled to photons in the young Universe via photon–electron scattering (Eisenstein & Bennett, 2008). When the Universe became cool enough, 380 000 years after the Big Bang, the baryons and

City Group	City–City Distance from ξ	Geographical City–City Distance	City Radius from ξ	Pixel Measure of City Radius
Midwest	389 ± 10 km	367 ± 132 km	33 ± 0.8 km	32 ± 8 km
Southwest	423 ± 10 km	346 ± 109 km	38.9 ± 1 km	35 ± 13 km
Southern	240 ± 6 km	249 ± 93 km	23 ± 0.5 km	23 ± 13 km

Table 1. City distances and radii as determined from the correlation function, ξ, and compared to direct pixel measurements converted to kilometres.

electrons combined, forming atoms, dominated by neutral hydrogen. The photons were now decoupled from matter and the Universe became transparent. The imprint of this event is carried by photons streaming freely to observers and is known as the Cosmic Microwave Background (CMB). The neutral atoms retained a frozen pattern of sound waves formed in the presence of dark matter. The resulting pattern of the density variations in both the dark matter and ordinary matter is reflected in the galaxy cluster distribution. So, in addition to the 100 h^{-1} Mpc cluster–cluster distance, the correlation function analysis also gives an independent estimate of the dark matter fraction of the Universe (Eisenstein et al., 2005).

Student laboratory

Some aspects of this work have been adapted to a lab in introductory astronomy (Smith Jr., 2012). Students are first introduced to astronomical distances by having them consult the web for typical distances to stars, and typical distances to galaxies. They are led to determine that galaxies are roughly a million times further away than stars.

Next, students are given the appropriate Structured Query Language (SQL) for downloading SDSS data sufficient for a wedge plot in Excel, consisting of about 9600 galaxies. This enables a visual comparison between the students' galaxy clustering plot and city clustering with images provided by the instructor.

Finally, students are given the data for the Midwest cities that enable them to plot the two-point correlation function so that it can be compared to the correlation function for galaxies provided by the instructor.

Summary

The two-point correlation function, used by cosmologists to describe the large-scale structure of the Universe, is explained by drawing an analogy between city clustering and galaxy clustering.

The analogy is demonstrated by calculating the two-point correlation function for night satellite images of three groups of cities in the USA, and comparing the results with geographical distances, and with direct pixel measurements of city radii, converted to kilometres. The comparison reveals that the characteristic distance of a correlation function can be interpreted as the average radius of a city group, and the correlation function peak is the average city–city distance.

Then similar features in the galaxy–galaxy correlation function can be similarly interpreted by non-experts: the characteristic distance is equivalent to the average radius of a galaxy cluster, and the peak's position, the baryon acoustic peak, is the distance between galaxy clusters. This cosmologically important peak is a relic of an oscillating plasma of coupled photons and baryons in the presence of dark matter just after protons and electrons combined to form neutral hydrogen and photons decoupled to form what became the CMB.

Acknowledgements

Support for this work has been provided by the National Science Foundation (NSF) Partnerships in Astronomy & Astrophysics Research and Education (PAARE) award AST-0750814.

References

Eisenstein, D. J. et al. 2005, The Astrophysical Journal, 633, 560

Eisenstein, D. J. & Bennett, C. L. 2008, Physics Today, April 2008, 44

Hawkins, E. et al. 2003, Monthly Notices of the Royal Astronomical Society, 346, 78

Peacock, J. A. et al. 2001, Nature, 410, 169

Peebles, P. J. E. 1980, *The Large-Scale Structure of the Universe*, Princeton (NJ: Princeton University Press)

Smith, Jr., D. M. 2012, *Cosmology for Non-Science Majors*, http://physics.scsu.edu/~dms/cosmology/simulations.html

Sullivan III, W. T. 1993, *North America at Night*, (Salt Lake City: Hansen Planetarium). A more recent version of this poster can be found at http://www.ngdc.noaa.gov/dmsp/night_light_posters.html

Zehavi, I. et al. 2002, The Astrophysical Journal, 571, 17

Zehavi, I. et al. 2005, The Astrophysical Journal, 621, 22

Biography

Daniel M. Smith, Jr. is a Professor of Physics at South Carolina State University with an interest in developing interactive labs and simulations to explain the results of cosmological research to non-science University students. His work can be found on the website Cosmology for Non-Science Majors: http://physics.scsu.edu/~dms/cosmology/home2.html