A Spatial Take on U.S. K-12 Education and Mobility

Anita Mehrotra
Statistics 225, Institute for Applied Computational Science
Harvard University
Cambridge, MA 02138
anitamehrotra@fas.harvard.edu

Abstract

Intergenerational income mobility in the United States has long been studied as a function of educational and geographic characteristics. In this paper, we extend existing work by formalizing the informal visual methods often utilized to draw spatial conclusions. We explore the absolute mobility metrics computed by Chetty et al in their 2014 paper [6] on 40 million children born between 1980 and 1982, then conduct kriging on the contiguous U.S. We find low mobility in the Northwest and South, high mobility in the Midwest, and consistent standard errors across the nation. Anomalous trends of low mobility appear in southern Maine, New Mexico and at the border of Nebraska and South Dakota. We conclude that formal geostatistics models corroborate the trends identified in [6], and highlight less-obvious patterns. Furthermore, this work is the first of its kind and because it opens the door to future studies in which formal spatial statistics models are applied to intergenerational mobility - an area that, till now, has almost entirely been studied via a frequentist lens.

1 Introduction

In 2012, President Obama stated during his State of the Union Address that the future income of U.S. children depended heavily on the quality of their K-12 teachers. In fact, "a good teacher [could] increase the lifetime income of a classroom by over $\$250,000$ [13]." The reference was to a paper that had been published just three months earlier by Harvard economists Raj Chetty and John Friedman, and Columbia Business School professor Jonah Rockoff [7]. Chetty et al had measured the value-add (VA) of teachers - defined in terms of students' test scores - and utilized this information to quantify the impact that VA had on students' college attendance and projected income.

It was from this research that the Equality of Opportunity Project (EOP) was born. Designed to study U.S. educational inequalities and their impact on future leaders, EOP was spearheaded by four economists from Harvard University and the University of California-Berkeley. In January 2014, two new works were published. They studied the effect of education, mean parent income, racial segregation and teenage birthrates (among other features) on future *relative* and *absolute mobility* [6] and over time [5]. The conclusions were drawn by plotting a heat map of mobility measures by commuting zone (below), then visually inspecting the map in order to infer spatial trends.

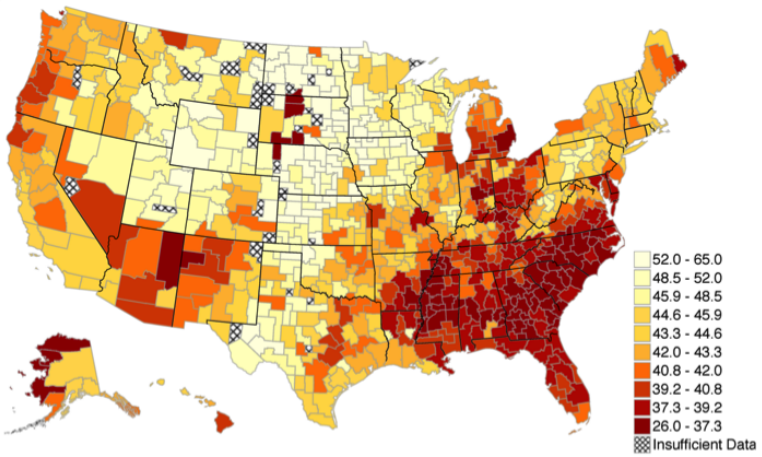

Figure 1: Absolute Mobility, as computed and studied by Chetty et al in their 2014 paper [3].

In this paper, we explore the absolute mobility data on nearly 3000 counties from a formal spatial perspective. We use spatial Bayesian techniques from geostatistics to answer the following question: What patterns insight do we glean in comparing the formal spatial results to the results that Chetty et al derive?

We first provide a literature overview and a brief explanation of the theory we use [Section 2]. The data is presented in more detail in [Section 3]. In [Section 4], we share our results and in [Section 5], we study the trends identified between relative and absolute mobility as well as my results compared to Chetty et al's. We conclude with possible next steps and the significance of our work [Section 6]. All of the code and figures referenced in this paper are publicly available [1], as is the data [8].

2 Background

2.1 Literature Overview

The study of intergenerational income mobility in the United States is by no means a new topic of interest. Furthermore, economics literature consistently suggest a relationship between the physical location of an individual and his/her future mobility.

More than twenty-five years ago, Gary Solon published his renowned work on "Integenerational mobility in the labor market [11]." The paper, which was incorporated into the classic *Handbook of Labor Economics* in 1999, used new empirical evidence to find that intergenerational mobility in long-run income was at least 0.4 – "dramatically less mobility than suggested" by previous research. Around that time, Princeton economists David Card and Alan Krueger (1990) were studying the impact of school quality on the "rate of return to education" for a subset of men born between 1920 and 1949 [3]. Using a variety of proxies like student-teacher ratio and relative teacher salaries, they found no evidence that parental income or education affected state-level rates of return.

Yet by the early- to mid-2000's "[t]he economic value of higher teacher quality" was being studied intently by researchers like Eric Hanushek [10]. He found that that there was direct relationship between how good a teacher was and the increase in marginal gains for that teacher's students.

In early 2014, Chetty et al used IRS income data for 40 million children in their paper, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States [6]." The computed mobility metrics were couched in terms of children's projected mean income relative to other children in the same birth cohort. To summarize the conditional expectation of a child's rank

[1] Code and high-resolution figures for this project can be found here: `https://github.com/anitameh/stats225-final-project`.

given her parents' rank, Chetty et al performed an ordinary linear regression and defined "relative mobility" as the estimated slope. Then, given a percentile p for the parent from the national parent income distribution, the researchers were able to calculate the expected rank of children, i.e. the child's "absolute mobility at percentile p [6]."

This paper takes [6] one step further by using kriging, a classic geostatistics model, applied to the mobility data in order to formally learn spatial trends. Though we initially hypothesize that these models may reveal different trends, we find instead that they robustly corroborate the trends identified in [6].

2.2 A Brief Theoretical Overview

We use Gaussian processes to model spatial variation in the absolute mobility data for 40 million children born between 1980 and 1982 in 2880 counties of the contiguous United States.

A Gaussian process (GP) is a generalization of a multivariate Gaussian distribution over infinitely many variables. As a result, it is fully specified by its infinitely-long mean vector and infinitely-large covariance matrix. GP's are used to specify distributions over multidimensional functions. In GPs, sample points that are close to one another have higher covariance i.e. have a greater influence on one another than points that are further apart. That is, the conditional distribution at a particular point (and subsequent points) may or may not be influenced by the conditional distributions at other sample data points.

Because the marginal of a Gaussian is a Gaussian, we can study GPs in terms of a finite subset of the mean vector and covariance matrix. The covariance matrix must be positive semidefinite, and is determined by a covariance function, often denoted κ, and which can take many forms, including, (squared) exponential, Gaussian, spherical, cubic, and the Matern.

The theoretical variogram measures the variance of the difference of two data points, and captures the degree of spatial dependence of a random field or stochastic process [4]. In this paper, the spatial random field is taken to be the set of standardized absolute mobility data points. It is often easier to work with the "semivariogram," i.e. half the variogram, which is what I use here. Formally,

$$\text{Variogram } \equiv 2\hat{\gamma}(h) = \mathbb{E}[Z(s) - Z(s+h)]^2 \tag{1}$$

$$\text{Semivariogram } \equiv \hat{\gamma}(h) \tag{2}$$

where $Z(s)$ is a stochastic process and h is the difference in the variance of two data points.

"Kriging" is the process in which interpolation of missing values for geostatistics data are modeled by a GP. Barring unreasonable assumptions for priors, kriging gives the best linear unbiased prediction of such values. As such, the "empirical variogram" uses existing data to construct the required covariance matrix [2].

3 Data

We use the absolute mobility metrics computed in [6] and obtained from Online Table 3 in [9]. Though the data used in [6] is for "commuting zones" - aggregations of, on average, four counties - we choose to study counties since all existing shapefiles are county-level. For plotting purposes, we use shapefiles from the U.S. Census Bureau [15-17]. We merge the two files in order to obtain a complete data set with latitude and longitude of centers of each county for each mobility metric.

We remove the non-contiguous portions of the U.S. (specifically, Hawaii, Alaska, Puerto Rico and the Virgin Islands), since kriging results are most reliable for non-disjoint surfaces. We then plot the absolute mobility per county. Lighter colors indicate higher absolute mobility values:

[2] The "sill," "nugget," and "practical range" are important concepts associated with determining which covariance function best models the empirical variogram. The sill is the asymptote of the semivariogram. The nugget can be thought of as an error term of sorts, and is the height of the jump/discontinuity that occurs at the origin. The practical range is the distance at which the semivariogram has reached 95% of the sill value.

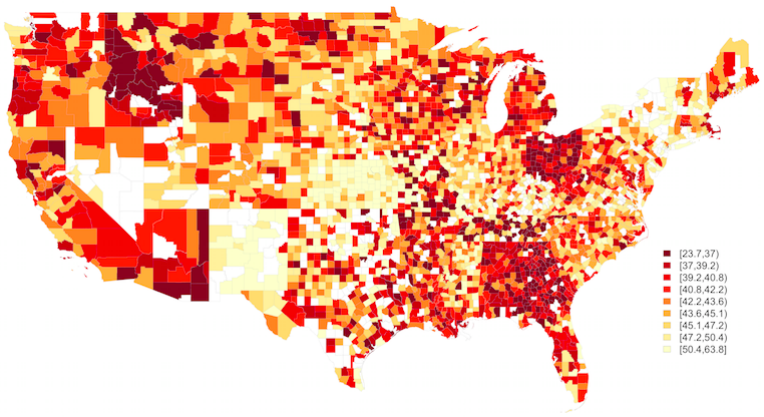

Figure 2: Absolute mobility by *county*, as computed by Chetty et al in their 2014 paper [3].

Notice that there are regions with missing data, largely in rural areas where Chetty et al were unable to obtain data or because few (or no) people lived there.

Next, we standardize the mobility data by subtracting out the mean, then dividing by the standard deviation. The standard deviation is proportional to the inverse of the square root of the sample size. The sample size refers to the number of children in the 1980-1982 birth cohort that are in each county. We therefore have the following:

$$z = \frac{x - \mu_x}{\sigma_x}$$
$$\Rightarrow z \propto (x - \mu_x) \cdot \sqrt{population}$$

where x is the original absolute mobility metric and z is the associated standardized absolute mobility. Omitting NA values, $\mu_x = 43.43497$. This gives the following plot:

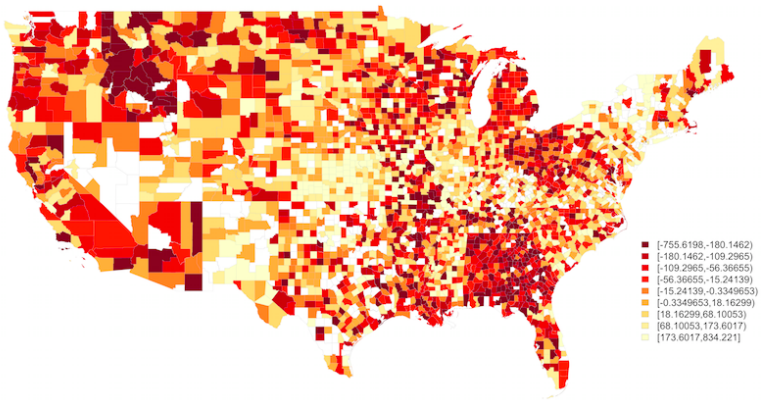

Figure 3: Standardized absolute mobility.

4 Methodology

We first begin by plotting the "variogram cloud." This provides a visual way of studying the rescaled mobility data points (from here on, referred to in the paper as the "rescaled data" or "data") in terms of distance versus semivariance. Next, we study how much of the data is captured over a range of different distance bins. This plot demonstrates that a majority of the data is captured by studying a subset, i.e. data points that are spatially separated by a maximum threshold distance. Using this information, we re-plot the variogram cloud with restricted distance.

Recall that county centers are separated by degrees of longitude and latitude. A "degree" is approximately equal to 67 miles. Thus, we define a "reasonably" close distance as counties that are less than or equal to 20 degrees apart i.e. approximately 1340 miles apart. This definition is corroborated by the plot in Figure 5, in which it is clear that 80% of the data is captured by points that are less than or equal to 20 degrees apart. Furthermore, we know that the furthest distance in the contiguous United States measures between California and Maine and is approximately 2630 miles. Thus our maximum distance threshold corresponds to counties that are separated by at most approximately half the "width" of the U.S.

The cloud plots also allow us to apply an "eyeball test" for stationarity [2]. With geostatistics data, "stationary" data are such that the difference in the variance of points will not change under translation of these points in space.

We next apply Cressie's robust variogram estimator to the standardized, stationary absolute mobility data, estimating parameters using both the unrestricted and restricted distance data. Recall from (1) that the classical estimator of the variogram is given by:

$$2\hat{\gamma}(h) \equiv \frac{1}{|N(h)|} \sum_{N(h)} (Z(s_i) - Z(s_j))^2, h \in \mathbb{R}^d$$

where $|N(h)|$ is the number of distinct sample pairs lagged by the vector h. Cressie's robust estimator takes *fourth* roots of the squared differences, which minimizes the effect of outliers on the remaining data [4, pg. 75]:

$$2\bar{\gamma}(h) \equiv \left[\frac{1}{|N(h)|} \sum_{N(h)} |Z(s_i) - Z(s_j)|^{1/2} \right]^4 / \left(0.457 + \frac{0.494}{|N(h)|} \right). \tag{3}$$

In each case, we fit the empirical variogram with six different covariance functions:

1. Spherical
2. Squared exponential
3. Wave
4. Exponential
5. Cubic
6. Circular

It is clear that the unrestricted distance empirical variogram is difficult to fit (notice its un-intuitive shape in Figure 7). Yet the restricted distance empirical variogram has a classic shape, and we find that three of the covariance functions model it equally well (Spherical, Exponential, and Circular).

We plot the variogram envelope with ten simulations for the restricted distance data with the estimated parameters from the Spherical covariance function. The resultant envelope (Figure 9) demonstrates how permutations in the mobility locations affects the variogram.

Finally, we conduct ordinary kriging on the contiguous U.S. states. We choose to implement "ordinary kriging," as opposed to simple or universal kriging, because we assume that the mean is constant only in the local neighborhood of each estimation point. "Simple kriging" is not a good choice for this data set because it assumes the expectation and covariance function to be known and constant over the entire domain.

5 Results

The variogram cloud for the data is shown here:

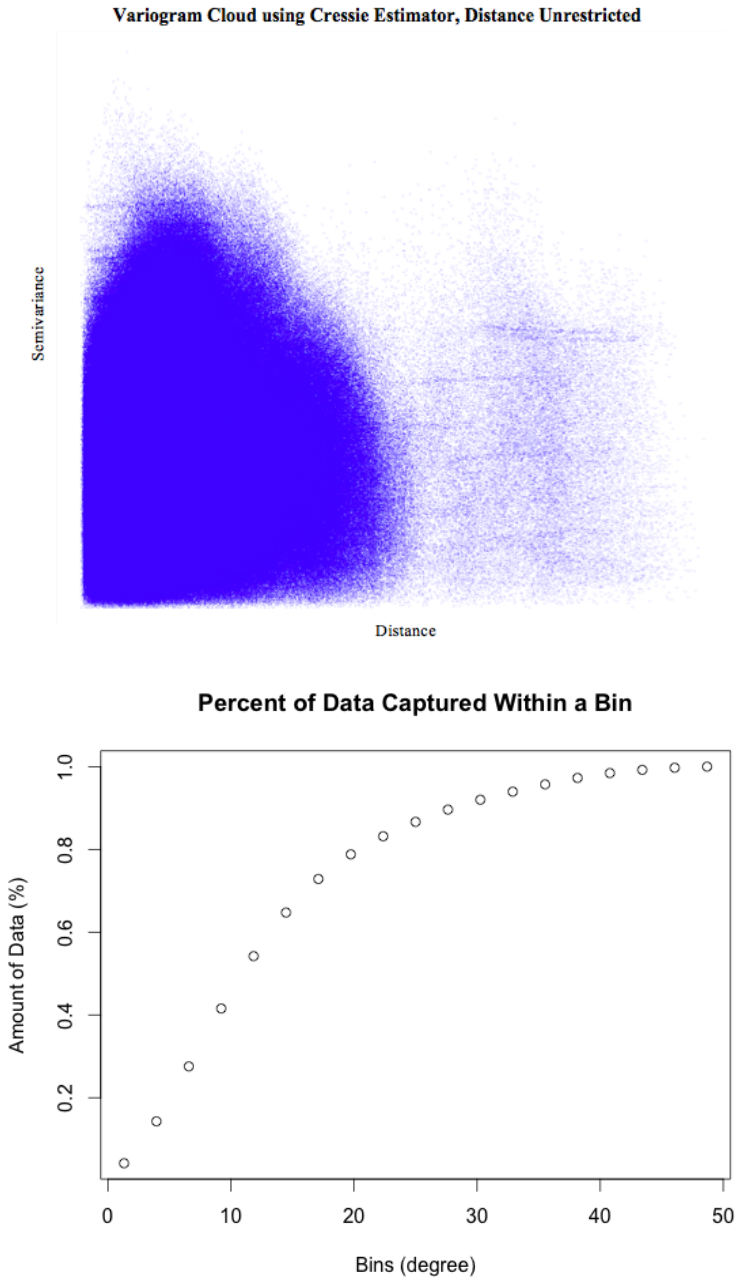

Figure 4: Unrestricted standardized absolute mobility variog cloud using robust estimator (top); Binned data (bottom).

The variogram cloud plot is rather uninformative, as the image is dominated and skewed because of outliers to the right. If we bin the data by distance and plot against how much data is included in those bins, we get the second figure above (numbers in Table 1). Notice counties separated by approximately 49 degrees or less capture 100% of the data, and that counties separated by 20 degrees or less capture approximately 80% of the data. We re-plot the variogram cloud, this time restricting the maximum distance to be 20 degrees.

In Figure 5, we see that the cloud appears to follow a uniform distribution. Furthermore, horizontal lines tend to appear, indicating that groups of counties (across a range of distances) have the same semivariance. Using the "eyeball test" discussed in [2], it appears that the data does not exhibit

Table 1: Percent of Data Captured

Binning Distance (deg)	**Data Captured (%)**	**Binning Distance (deg)**	**Data Captured (%)**
1.32	4.16	27.63	89.62
3.95	14.29	30.26	92.01
6.58	27.57	32.89	93.95
9.21	41.56	35.53	95.72
11.84	54.21	38.16	97.27
14.47	64.74	40.79	98.44
17.11	72.86	43.42	99.22
19.74	78.80	46.05	99.73
22.38	83.17	48.68	100
25	86.65		

nonstationarity. This makes sense because we took into account population size when standardizing the original mobility data, which is one way of combating nonstationarity. (See Appendix A for the variogram clouds of the non-standardized data.)

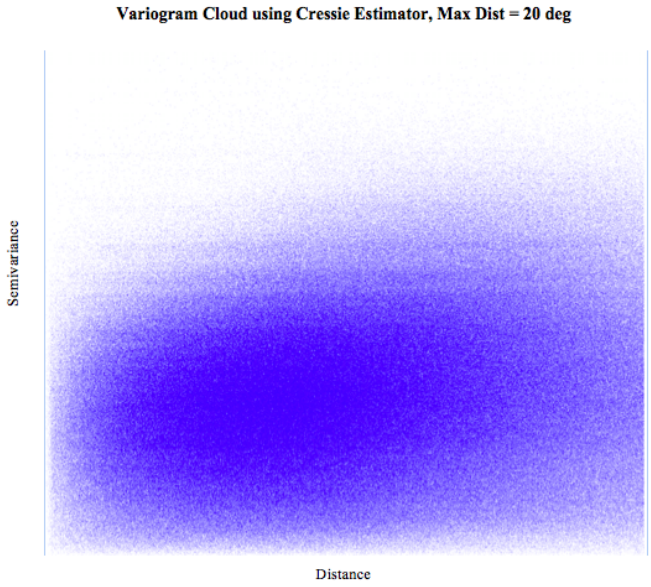

Figure 5: Restricted Standardized Absolute Mobility Variog Cloud using a Robust Estimator.

Figures 6 and 7 show the resulting fits of covariance functions to the empirical variograms for both the unrestricted and restricted distance data. It becomes even more evident that restricting the maximum distance is necessary in order to infer meaningful results: the restricted distance data appears to achieve local stationarity, and the covariance fits are dramatically more accurate.

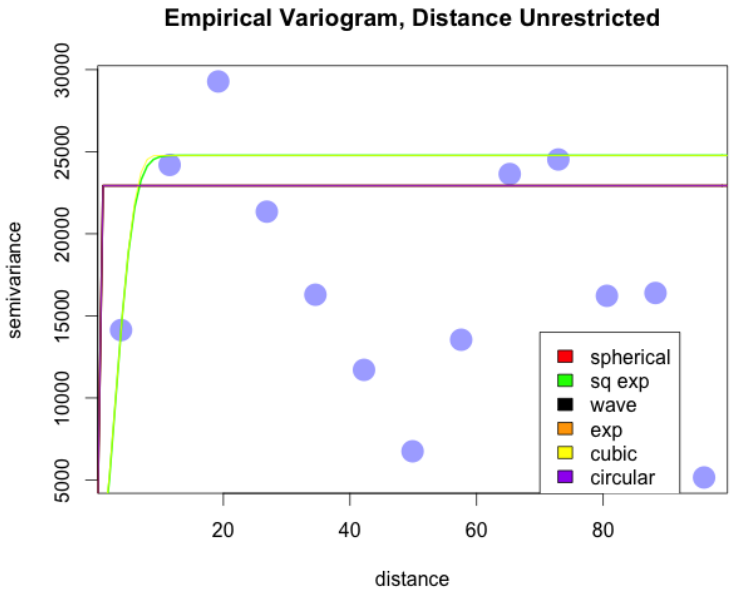

Figure 6: Variety of covariance functions for empirical variogram, All distances.

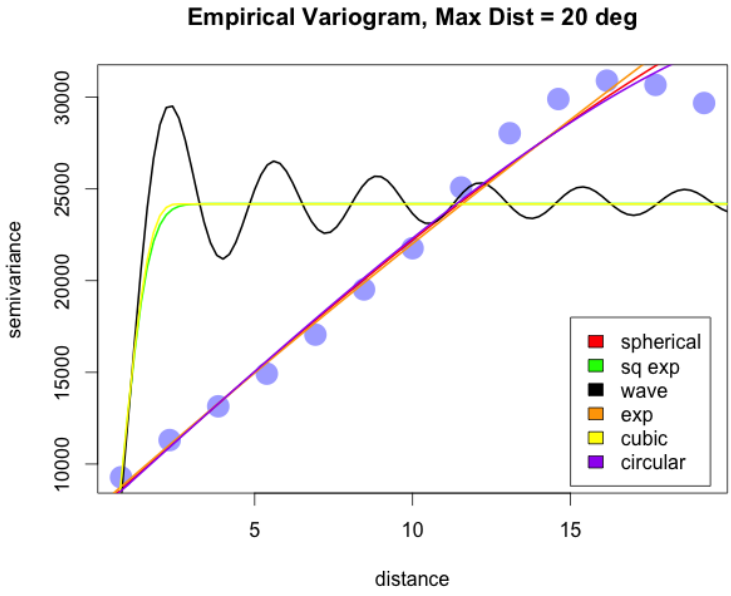

Figure 7: Variety of covariance functions for empirical variogram, Maximum distance=20 deg.

Figure 8 shows the spherical covariance function that best models the data. Recall that the spherical covariance function is defined as

$$C(h) = \sigma^2 \left(1 - \frac{3||h||}{2\phi} + ||h||^3 2\phi^3\right) \text{ for } \sigma^2 > 0, ||h|| \leq \phi.$$

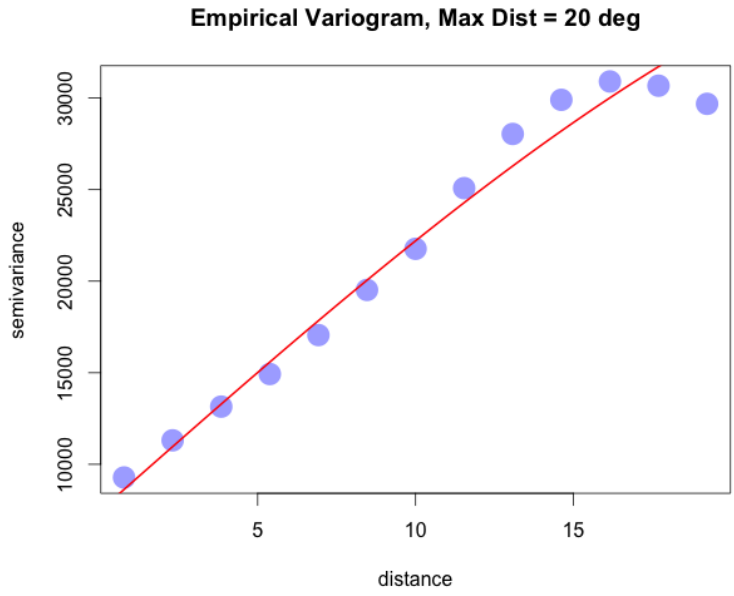

Figure 8: Spherical covariance function for empirical variogram.

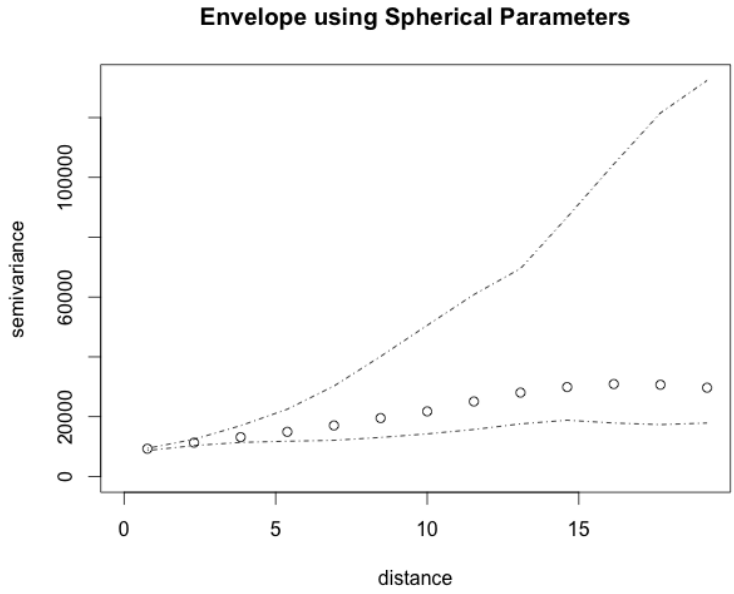

Figure 9: Envelope using 10 simulations.

The plot indicates that we can use an initial sill of 30000 and an initial range of 20. After fitting the variogram, the resulting nugget is ≈ 7455.1672, which makes sense given the plot above. The estimated parameters are:

$$\hat{\sigma}^2_{spherical} = 32794.5477$$
$$\hat{\phi}_{spherical} = 32.2863$$

By permuting the locations of the data points, we can generate an envelope that shows the upper and lower bound of the empirical variogram fit. The computation time took approximately one hour to perform ten simulations. I use this data to create an envelope (as opposed to the default of ninety-nine simulations) in Figure 9 above.

Finally, I conduct kriging and get the following surface (Figure 11) and standard errors plot (Figure 12).

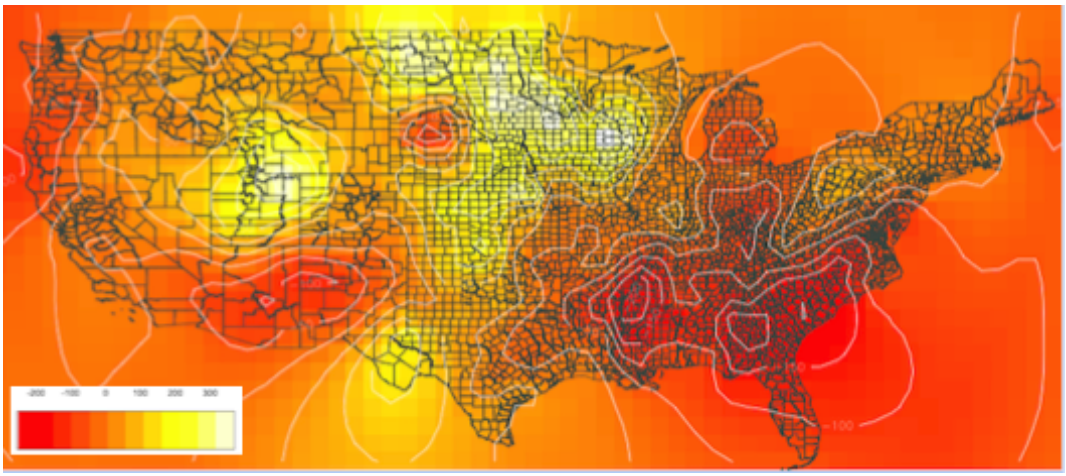

Figure 10: Kriged Surface.

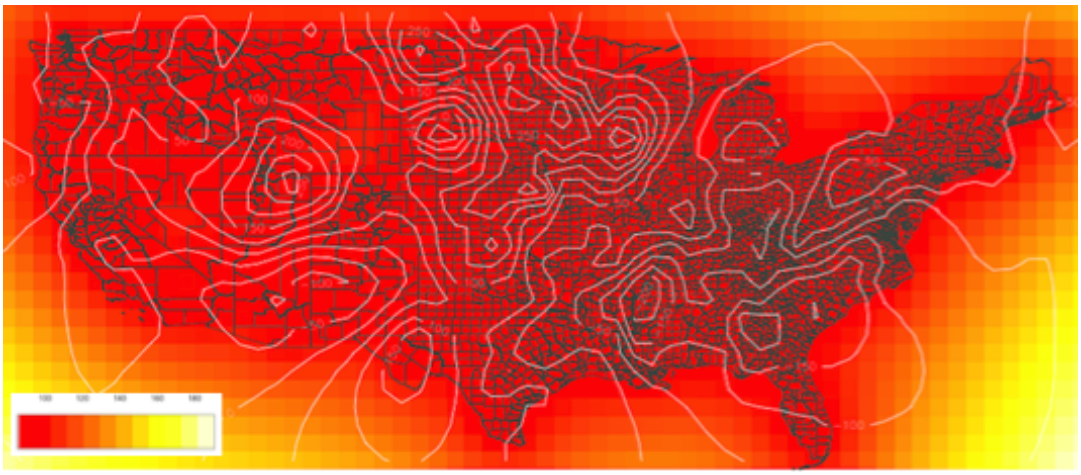

Figure 11: Standard Errors for kriged surface.

6 Analysis of Results

6.1 Initial Analysis

The estimated partial sill $\hat{\sigma}^2_{spherical} = 32794.5477$ is consistent with the plot in Figure 9. Since this is the asymptote of the semivariogram, it indicates that the variance of the differences between all possible data points, when spaced a constant distance apart, is approximately 60000. This means that the standard deviation is approximately $\sqrt{60000} \approx 244$, and indicates that most of the rescaled data points are within ≈ 244 of the mean. The estimate gives an idea of how much variation is

overall in the data after adjusting for possible correlation among the more closely-spaced points. In studying the map from Figure 3, it's apparent that the estimate makes intuitive sense as well: notice that all rescaled data values fall in the interval $[-500, 800]$ and that the analytical standard deviation is 160.104. Adjusting for especially close points, as is done when estimating the variogram, would thus contribute to an increase in the variance.

The estimated range, $\hat{\phi}_{spherical} = 32.2863$, is the greatest distance over which the rescaled mobility value at a point in the U.S. is related to the rescaled mobility value at another point. The estimated range, given the restricted distance data, seems both intuitive and reasonable.

In Figure 9, we generate the envelope by randomly allocating the rescaled data points to the spatial locations. The empirical variogram is then computed for each simulation using the same lags as for the variogram computed for the original data. The envelopes are computed by taking, at each lag, the maximum and minimum values of the variograms for the simulated data. The plot presents the minimum variogram value and maximum variogram value at each distance value.

Notice that the variance is very small when permuting points that are close to each other. This makes sense, because we expect points that are spatially near each other to be correlated and thus the variance in the distance to be small. Permuting points that are farther apart leads to high possible semivariance.

6.2 Kriged Surface

The kriged surface show low rescaled values in the Northwest and South. The darkest regions "valley" somewhere between -200 and -150. To get a mobility value in the range of the original data, we divide by the square root of the population of that county, then add 43. For example, Alachua, FL which is in the upper/middle portion of the state, has a sample size of 6953, which means that its absolute mobility is $\approx (-150/\sqrt{6953}) + 43 \approx 41.2$. Based on the scale of the original data, this is slightly below average mobility, consistent with what Chetty et al find.

Just as importantly, the consistent results in regions of the map in Figure 3 where there is little or no missing data, indicate that the kriged surface is accurately capturing the trends in areas where there is high data availability.

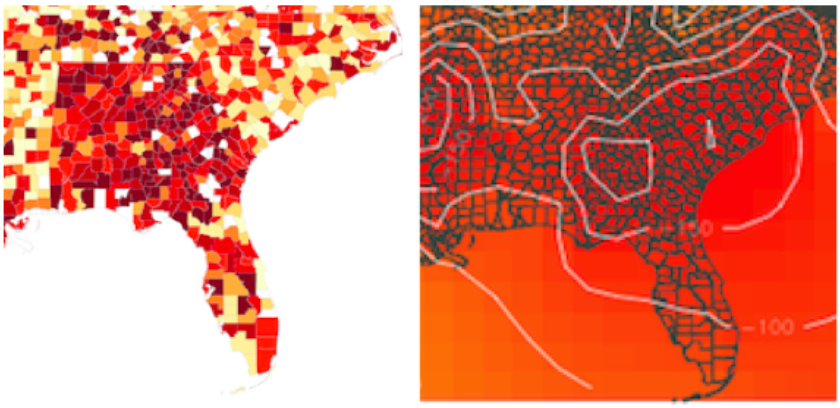

Figure 12: The kriged surface does a good job of capturing obvious trends.

In studying the kriged surface in Figure 13, we see clear regions of the U.S. with high rescaled values, specifically in the mid-West, Western regions of Texas and Northwest Utah/Southeast Wyoming. Interestingly, there are low rescaled values at the South Dakota-Nebraska border, in New Mexico and southern Maine (not pictured), suggesting low absolute mobility.

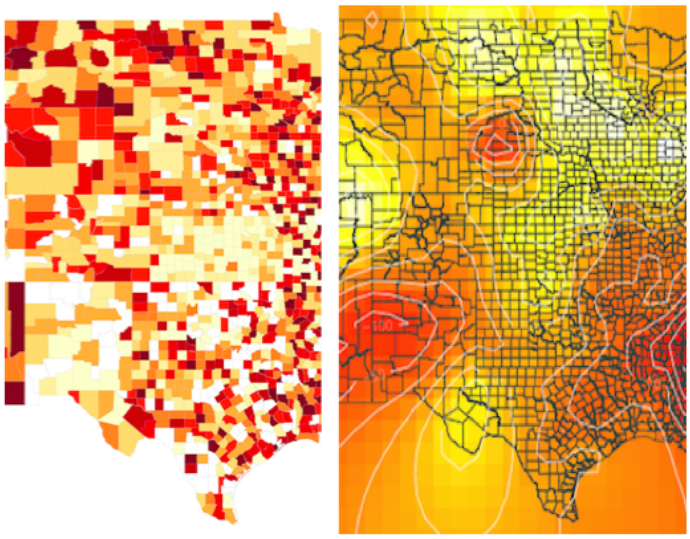

Figure 13: Interesting spatial patterns appear in the mid-West.

We dig into articles and databases that cover education and demographic statistics in northern Nebraska, and find that the student poverty rate is on average 14% from 2009 to 2012 - just below the national average. Approximately 40% of pre-K-12 students qualify for free and reduced-priced lunches - relatively high compared to the national average [10]. These facts suggest potential factors for the localized level of low mobility in that region.

Furthermore, we can see that the "peaks" and "valleys" evidenced in the kriging surface are underlying trends, and *not* solely a result of the missing data. Let's compare an area with high missing data to the kriged surface. Notice that although there are missing values in the Midwest and parts of Texas that also see high rescaled values in the kriged surface, there are also a large number of missing values in Nevada and New Mexico that have approximately 0 or low rescaled values in the kriged surface.

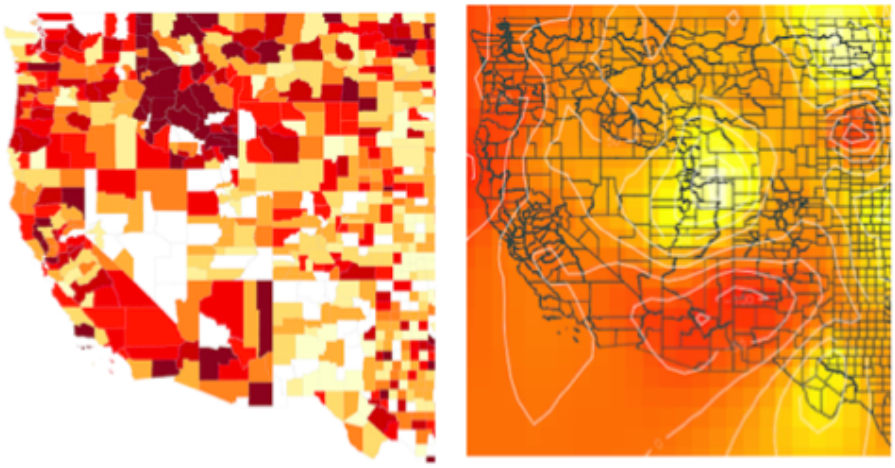

Figure 14: Missing values versus Kriged surface.

The low rescaled values in New Mexico are likely a result of overflow from the low rescaled mobility values in Arizona, which is its bordering state.

6.3 Comparison with Original Paper

Let's now compare the trends that emerge from our kriged surface with the spatial trends observed from Chetty et al's paper.

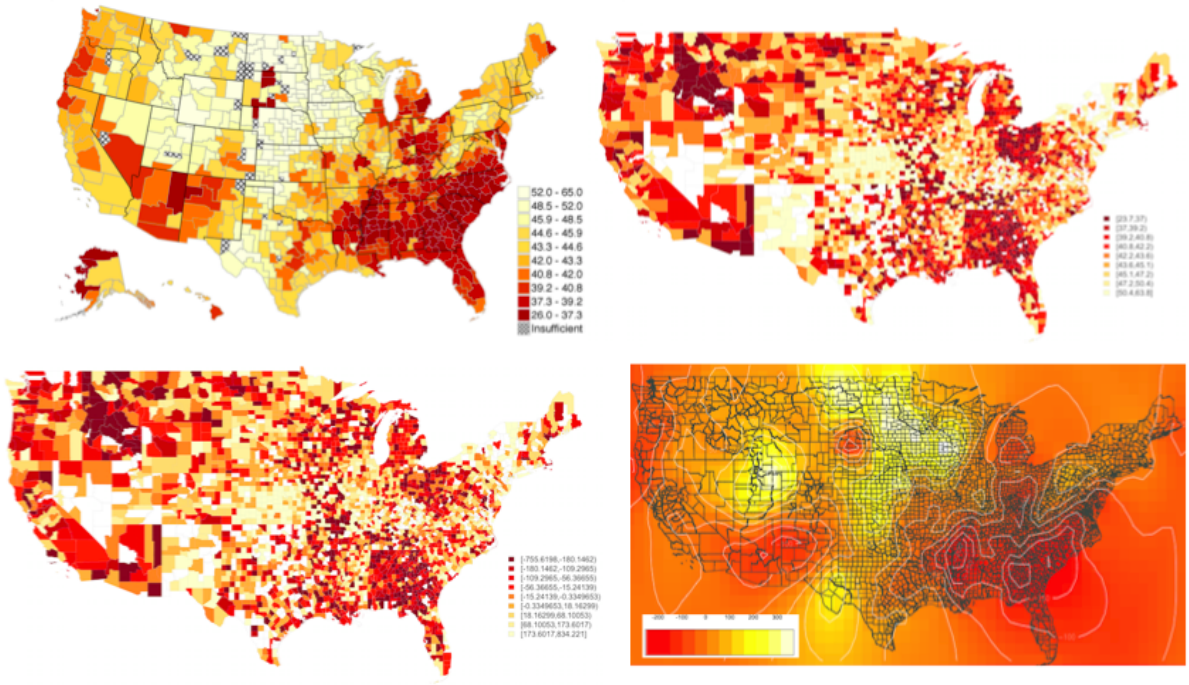

Figure 15: Original CZ plot of mobility (upper left), County plot of mobility (upper right), Rescaled mobility by county (lower left), Kriged surface (lower right).

Low rescaled values correspond to low original mobility values, as evidenced by the plots in Figure 15 above. We find similar trends in the kriged surface (lower right) as in the original map (upper left).

The South, the Pacific border of the Northwest and parts of the Southwest experience lower levels of absolute mobility relative to the rest of the nation. Many states in the South are also home to extreme racial segregation [6], and as recently featured in *The New York Times*, high drop-out rates [8]. The author, a native of northern California, has seen evidence of "beach towns" and rural portions of the Northwest (consider rural counties like Humboldt) with high crime rates, young marriages, and few people who migrate from the area, which (perhaps unsurprisingly) lead to low mobility. We surmise that southern-most parts of the Southwest, like Arizona, have low mobility due to its proximity to Mexico, evidenced most recently by the intractable illegal immigration debate.

Regions in the Midwest, on the other hand, experience higher values of absolute mobility. At first, we are surprised by this, given that most of the Breadbasket States are rural. Recall, however, that mobility is defined in terms of a child's income rank, *given his/her parent's income rate*. Furthermore, the Midwest is renowned for its educational system, with some of the top high schools in the nation routinely from northern Texas and Illinois. Thus, a child's future income may be higher relative to his/her parents', which translates to higher mobility. This is further substantiated if we study the time frame - parents' income was measured at the time that children were 16, which means that parent income was studied in the mid-1990's. This was the same time that some of the coldest Arctic temperatures struck the Midwest, which had a large negative impact on farmland values, and suggestively, on parents' income.

Our kriged surface picks out the low level of mobility on the border between South Dakota and Nebraska, which exist in the original map but is not as obvious. This is supported by recent policy research that finds a host of factors impacting upward mobility, in turn causing large variations in the Midwest [14].

Finally, the standard error plot indicates relatively consistent standard errors of approximately 100 across the entire nation, with intuitively higher error as we move away from the United States and into the Pacific and Atlantic Oceans.

We thus conclude that although no dramatically different spatial trends appear from my study, certain anomalies are highlighted more clearly. This shows that geostatistics models can, at the least, mathematically formalize the results derived in [3].

7 Conclusion

This project uses formal Bayesian spatial models to explore absolute mobility data of 40 million children born between 1980 and 1982 in counties of the contiguous United States, a la Chetty et al's 2014 paper on geographic trends in intergenerational mobility.

In testing a range of empirical variograms and conducting kriging, I find that similar trends arise across the U.S. as were identified informally in [6]. In particular, by using Cressie's robust variogram estimator, a spherical covariance function, and enforcing local stationarity, I am able to recreate nearly identical spatial trends.

Though our resulting kriged surface does not illuminate any new or vastly different spatial trends as we hoped it would, it does emphasize regions of anomalies that were less obvious in the original plot. For example, we find peaks in mobility in spots in the Midwest (amidst regions of low absolute mobility) .

Byy iterating several times through the process described here, the author learned a great deal about how to handle real-world, messy spatial data from a geostatistics perspective. Moreover, we think that the process of applying formal spatial models to such data and formalizing existing trends provides an opportunity for future work that could, for example, study the correlation between individual counties. Understanding which regions have the least correlation with each other can enable us to examine areas where there are large changes in mobility from one county to the next. Such research is critical because it can provide incredibly useful information to education policy makers.

References

[1] Bohling, Geoff. "Kriging." http://people.ku.edu/~gbohling/cpe940/Kriging.pdf. *C&PE 940*. Oct 2005.

[2] Bornn, Luke, Gavin Shaddick, and James V. Zidek. "Modeling nonstationarity processes through dimension expansion." American Statistical Association, Vol. 107 No. 497. Mar 2012. http://www.people.fas.harvard.edu/~bornn/Papers/Bornn2012a.pdf.

[3] Card, David and Alan Krueger. "Does school quality matter? Returns to education and the characteristics of public schools in the United States." http://www.nber.org/papers/w3358. *NBER*, Working Paper No. 3358. May 1990.

[4] Cressie, Noel A.C. *Statistics for Spatial Data*. Wiley, 1993.

[5] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. "Is the United States still the Land of Opportunity? Recent trends in intergenerational mobility." NBER Jan 2014. http://www.nber.org/papers/w19844.

[6] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. "Where is the Land of Opportunity? The geography of intergenerational mobility in the United States." http://obs.rc.fas.harvard.edu/chetty/mobility_geo.pdf. *NBER*, Jan 2014.

[7] Chetty, Raj, John N. Friedman, Jonah E. Rockoff. "The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood" NBER Dec 2011. http://www.nber.org/papers/w17699.

[8] Gabriel, Trip. "50 Years Into the War on Poverty, Hardship Hits Back." http://www.nytimes.com/2014/04/21/us/50-years-into-the-war-on-poverty-hardship-hits-back.html?_r=0. *The New York Times*, 20 April 2014.

[9] *Equality of Opportunity Raw Data Download*. http://www.equality-of-opportunity.org/index.php/data.

[10] New America Foundation, *Federal Education Budget Project*. http://febp.newamerica.net/k12/NE.

[11] Hanushek, Eric A. "The economic value of higher teacher quality." http://www.nber.org/papers/w16606. *NBER* Working Paper No. 16606. Dec 2010.

[12] Solon, Gary. "Intergenerational income mobility in the United States." http://www.jstor.org/discover/10.2307/2117312?uid=3739696&uid=2&uid=4&uid=3739256&sid=21103957174327. *The American Economic Review*, Vol. 82, No. 3. June 1992.

[13] Strauss, Valerie. "Obama on education in State of the Union address" 24 Jan 2012. http://www.washingtonpost.com/blogs/answer-sheet/post/obama-on-education-in-state-of-the, -union-address/2012/01/24/gIQAVfAwOQ_blog.html.

[14] Tomaka, Laura. "Brakes on economic mobility: New research points to a host of factors likely impacting upward mobility, causing large variations in Midwest." *The Council of State Governments: Midwest*, Policy and Research. http://www.csgmidwest.org/policyresearch/1213economic-mobility.aspx. Dec 2013.

[15] *U.S. Census Bureau County Shapefiles*. http://www.ers.usda.gov/data-products/commuting-zones-and-labor-market-areas.aspx#.UziXGq1dVy8. 2014.

[16] *U.S. Census Bureau Population Data*. https://www.census.gov//geo/reference/docs/cenpop2010/county/CenPop2010_Mean_CO.txt. 2014.

[17] *U.S. Census Bureau Location Data*.https://www.census.gov/geo/reference/centersofpop.html. 2014.

8 Appendix A

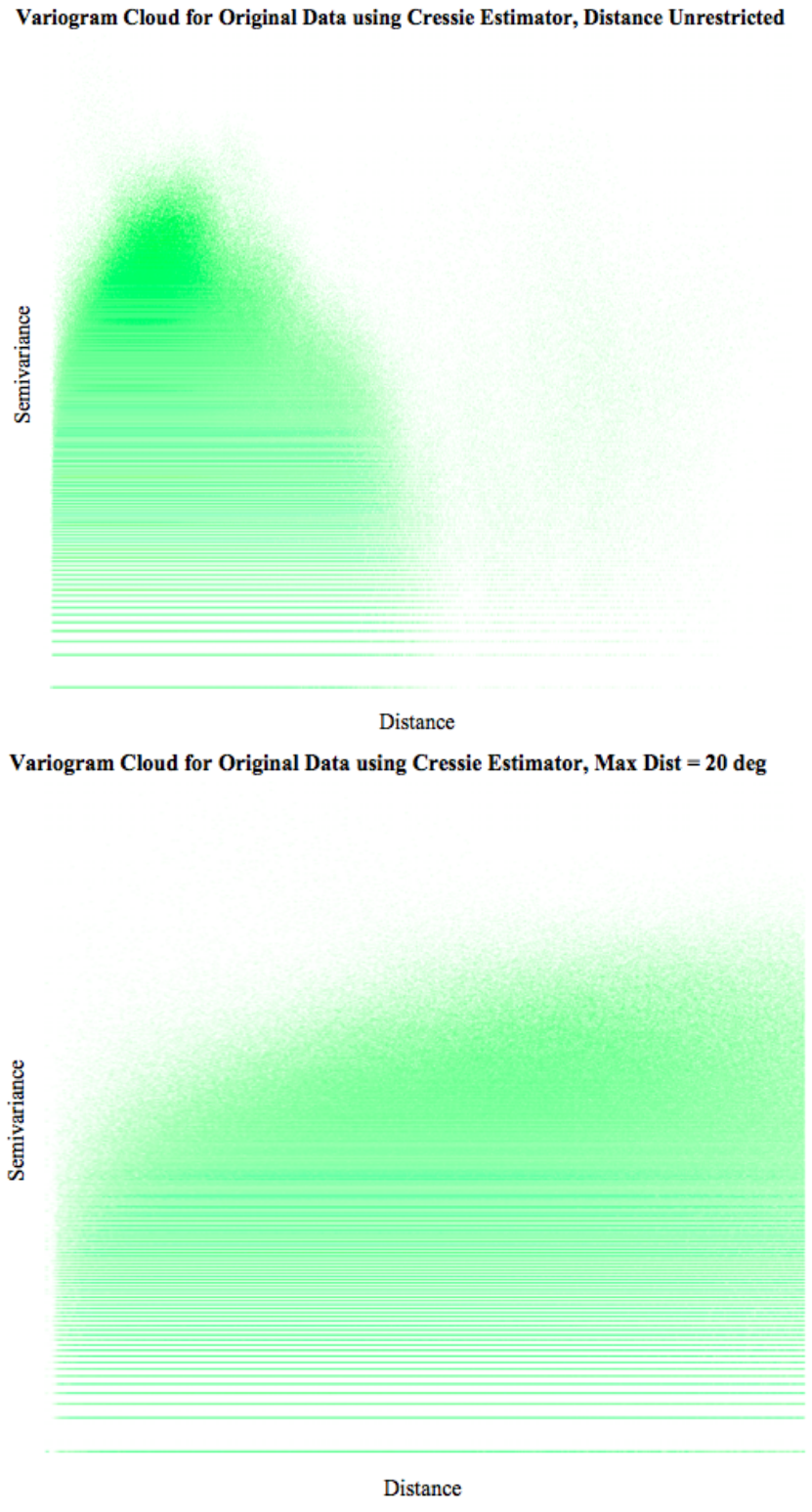

Figure 16: Variogram clouds for non-standardized data; Unrestricted (top) and Restricted (bottom) distances.