EESS 214 Project

Uncertainty in Remote Sensing Building Collapse Data: Can The Uncertainty Be Reduced With Geostatistics?

Steven Wong (uni: stywong)

June 5th, 2013

0. Table of Content

1. MOTIVATION

Natural disasters can impact a large area, while giving little (if any) prior warnings and taking place over a short timeframe. Both the scale and the suddenness of such phenomena render preparation for its relief difficult. Further, damages to city infrastructures can hinder the assessment of and the communication with regions affected most severely. For earthquakes specifically, damages to structures could range from cosmetic blemish to complete collapse. The ability to assess and map this range's geographic distribution with confidence can not only channel aid to appropriate regions, but also confirm if citizens can return safely to their homes, so as to prevent casualties from delayed post-earthquake collapses.

Two types of damage assessments can be conducted: (1) remote sensing, and (2) field based. Of the two, remote sensing is much quicker and lower in cost, but bares great uncertainties, and can only indicate if a building is either collapsed or not. For remote sensing, one would look at images of rooftops (figure1a), which shows neither the condition nor the number of buildings precisely; for field based assessments, on the other hand, one would go to each individual structure to assess its damage state. Nonetheless, remote sensing's quickness and cost benefit are substantial advantages, and this study wishes to determine if remote sensing data can be used with greater confidence, through quantifying and reducing its uncertainties with geostatistics.

For convenience, abbreviations are as follows: field based assessment is "FB"; remote sensing assessment, "RS".

figure1a: An example of remote sensing images: this one is taken to conduct damage assessment after the 2010 Haiti earthquake (source: Rochester Institute of Technology Chester F. Carlson Center for Imaging Science).

2. GOAL

This study wishes to implement geostatistics as a data-cleaning tool for RS data. To begin, spatially distributed building number and building collapse count are grouped into sub-regions, for each of which, a collapse percentage is computed. Assuming identical building vulnerability across the entire region, data points that are spatially closer should indicate similar collapse percentages, as ground motion intensity is spatially correlated. Specifically, if, say, three of the four points close together covey similar percentages, can we reject the fourth point, which is quite different? One example of rejection is taking the mean collapse percentage of an entire region, and eliminating all data values two standard deviations beyond the mean; however, such classical statistical method does not take into account of the data's spatial information. This study wishes to explore ways of using geostatistics to incorporate spatial information in data-cleaning.

The ultimate goal is to closer match RS data to actual collapse conditions, without using the latter to inform the former. This report highlights some of the findings from exploring the RS data at two different resolutions, and from conducting sample data-cleaning with geostatistics and with classical statistics. In the end, a comparability question remains: to what degree can RS data reflect reality, and whether random or systemic errors contribute more to the uncertainties in RS data?

3. DATASET

The dataset used in this study is from the Haiti January 12th, 2010 magnitude 7.0 earthquake (figure1), which contains both RS and FB. RS has about 250,000 spatially distributed binary data points, stating if a given structure has collapsed or not; FB has about 400,000 data points, stating in which of the six cumulative damage factor (CDF) is a given building (CDF ranges from no damage to complete collapse: 0-1,1-10, 10-30, 30-60, 60-100, and 100 percent). Since FB is much more precise than RS, FB is assumed as correct in this study.

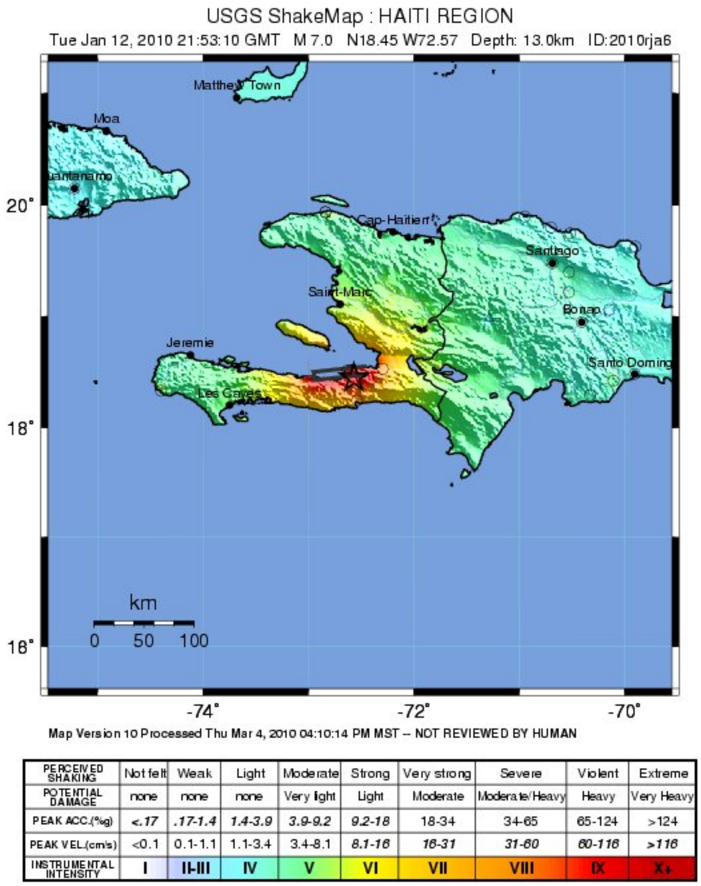


figure1b: Haiti earthquake intensity map on January 12th, 2010 (source: USGS).

4. METHOD

Step One: Data Pre-Processing
i. To increase computation efficiency, subdivide the FB and RS dataset into uniform grid boxes of two resolution ($250m^2$ and $50m^2$), compute the collapse percentage for each box, and place that value in the centroid of the box.
ii. For a more interesting comparison, zero into a geographic subset of the data that has a greater damage-exposure ratio. (This subset also happens to bring RS's aggregate building count closer to that of FB, potentially making the two more comparable). This area would be around Port-au-Prince, which is also close to the epicenter.

Step Two: Quantifying Uncertainty
i. Compare variograms for FB and RS, and at $250m^2$ and $50m^2$ resolution. Give particular attention to the nugget's potential division between micro-scale variability and measurement error. Due to various advantages(as illustrated later), the $50m^2$ resolution is chosen for the remainder of the study.
ii. Perform ordinary kriging on FB and continuous part kriging on RS; continuous part kriging is chosen over ordinary kriging for RS because RS is assumed to have measurement errors.
iii. Use Modified Mercalli Intensity (MMI), an earthquake intensity measure (figure1b), to perform universal kriging on RS, seeing if this additional information about the earthquake event can bring RS closer to FB.

Step Three: Reducing Uncertainty
i. Compare kriged FB against kriged RS result.
ii. For classical statistical data-cleaning, remove RS data two standard deviations beyond the dataset's mean, and re-krig with continuous part kriging.
iii. For geostatistical data-cleaning, remove RS data two kriged standard deviation beyond the kriged RS mean, and re-krig with continuous part kriging.

5. RESULTS

The following is a highlight of the study's observations; each is discussed further in next section.

i. Zeroing into a geographic subset of the raw data render FB and RS more comparable on the aggregate scale.

ii. Setting a building count threshold for data points reduces the nugget; this reduction should be all attributed to a reduction in measurement error, and a larger portion of what remains of the nugget should be micro-scale variability.

iii. Increasing the data resolution (from $250m^2$ to $50m^2$) increases the micro-scale variability portion of the nugget.

iv. Universal kriging with MMI earthquake intensity is not a viable option for this dataset, because the trend is sensitive to the binning method.

v. Compared to its classical statistics counterpart, geostatistical data-cleaning does not bring RS as close to FB on the aggregate scale (shown below), but does retain RS's original spatial characteristics better.

Removing Uncertainty						
					Remote Sensing Data within 2 Stdev	
	FB	FB_CPK	RS	RS_CPK	Variance from CPK	Variance from Classical Statistics
Number of Collapses	11936	12037	17085	17029	14819	14090
Percentage Collapse	8.78%	8.86%	13.25%	13.21%	11.49%	10.93%

6. RESULTS ANALYSIS AND DISCUSSION

i. Choosing a Comparable Site

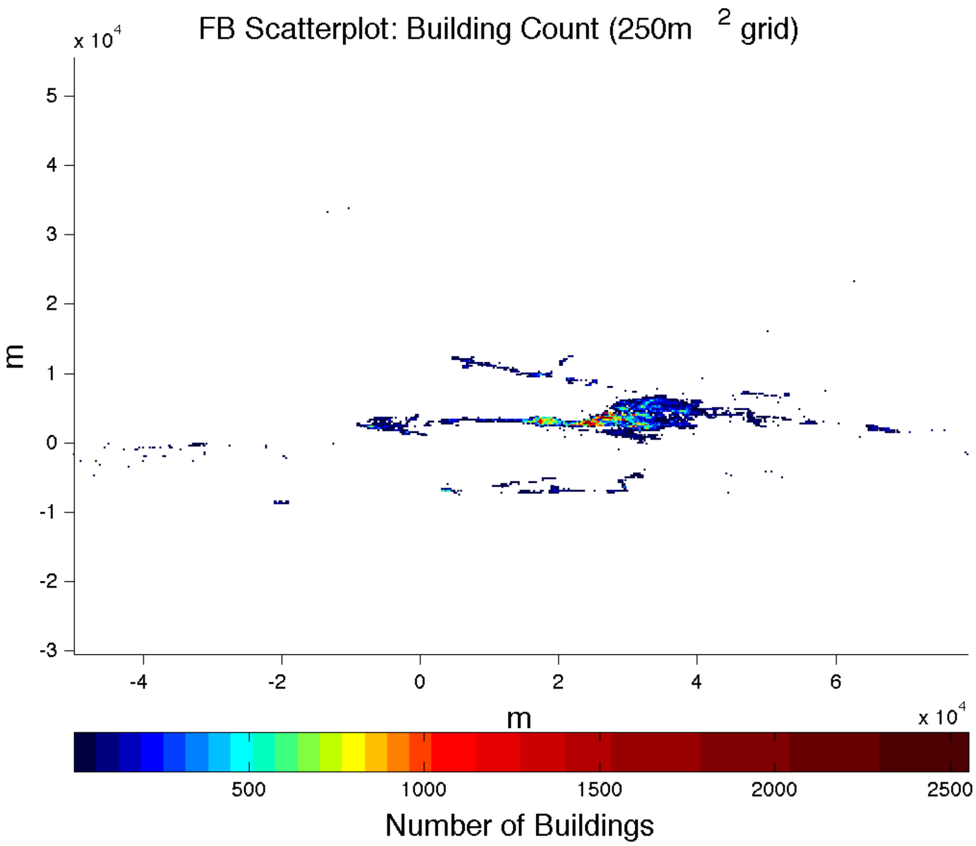


figure2a: FB building count, for the entire raw data domain.

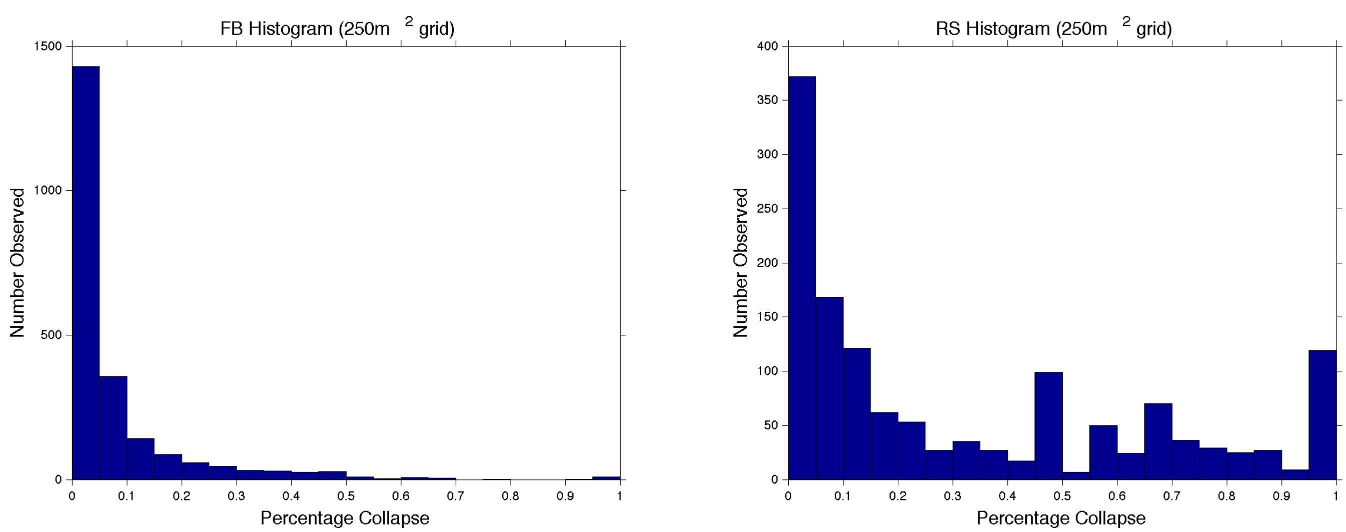


figure2b: Building collapse distribution for FB (left) and RS (right).

	FB (100CDF)	FB (>60CDF)	RS	RS/FB(100CDF)
Number of Buildings	405942	405942	277223	68.29%
Number of Collapses	31017	62703	29190	94.11%
Percentage Collapse	7.64%	15.45%	10.53%	

table1: FB and RS raw data statistics. (As seen in various tables in this report, there are slight value discrepancies between the $250m^2$ and $50m^2$ resolution dataset; for the sake of succinctness, this table refers to the $250m^2$ dataset, and for the entire raw data region).

To begin, a point of comparison between FB and RS needs to be established, as FB is in a range of six CDFs and RS is binary. Specifically, is it more reasonable to define "collapse" from the RS perspective as 100 CDF, or as above 60 CDF (the next increment down)? After all, above 60 percent damage state could look similar to total collapse from an aerial perspective. Comparing the aggregate percentage collapse, table1 shows FB at 100 CDF having a collapse percentage slightly closer to that of RS; the study will then use FB at 100 CDF.

Figure2a shows that building count is highly concentrated around the Port-au-Prince area for FB. The case is similar for RS, but table1 shows RS accounting for only 70 percent of the FB building count, and (now assuming FB 100 CDF and RS are comparable) overestimating the collapse percentage. Figure2b shows FB having an exponential distribution, whereas RS contains much "noise" beyond the 40 percent collapse mark.

By zeroing into that subarea of higher per grid building count (figure3a and figure3b), RS building counts get closer to that of FB, but its overestimation in collapse percentage remains (table2; as this report shows later, this gap is lessened after reducing the uncertainties in RS data). Figure3c and figure3d show RS collapse percentage distribution being closer in shape to that of FB, at both $250m^2$ and $50m^2$ resolution. This increased similarity is likely due to the choice of focusing on an area with higher building concentration, since lower building concentration is more prone to fluctuations in collapse percentage (explained further in the next section). Finally, because a substantial portion of the collapse percentage is zero, the data is not transformed. The zeroed-in subdomain is X(15km, 40km) by Y(0km,7.5km).

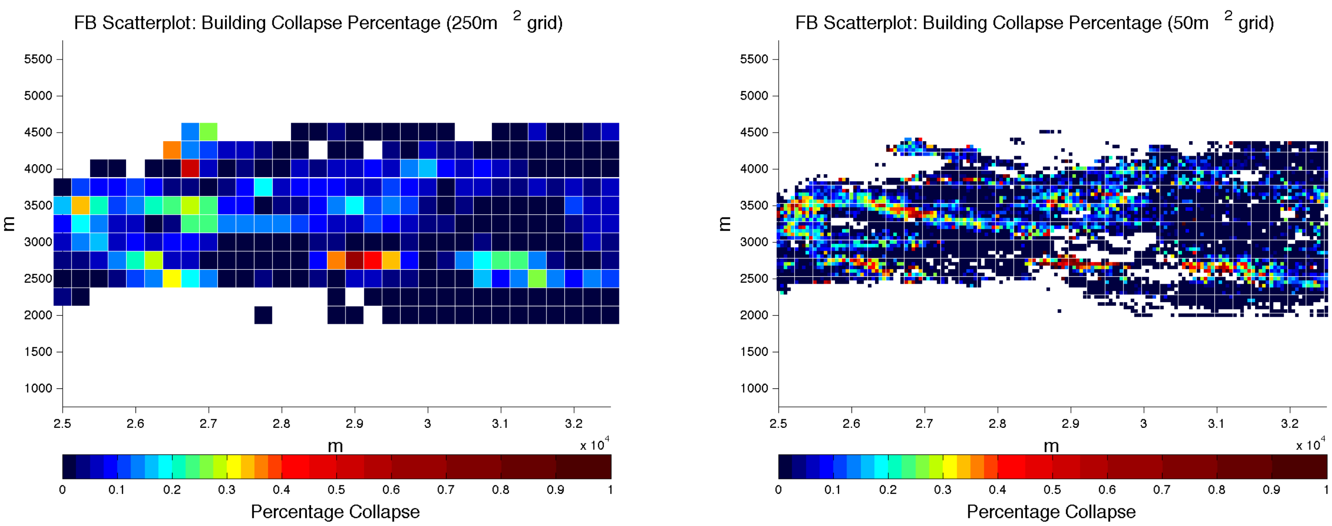


figure3a: FB building collapse percentage at $250m^2$ (left) and $50m^2$ (right) resolution

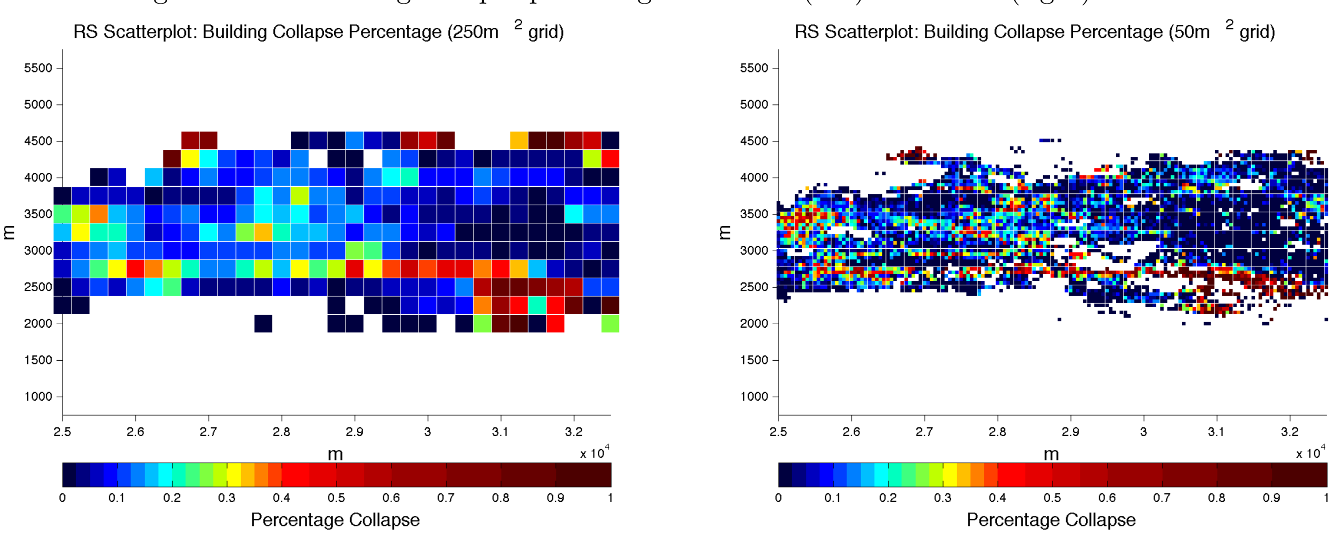


figure3b: RS building collapse percentage at $250m^2$ (left) and $50m^2$ (right) resolution

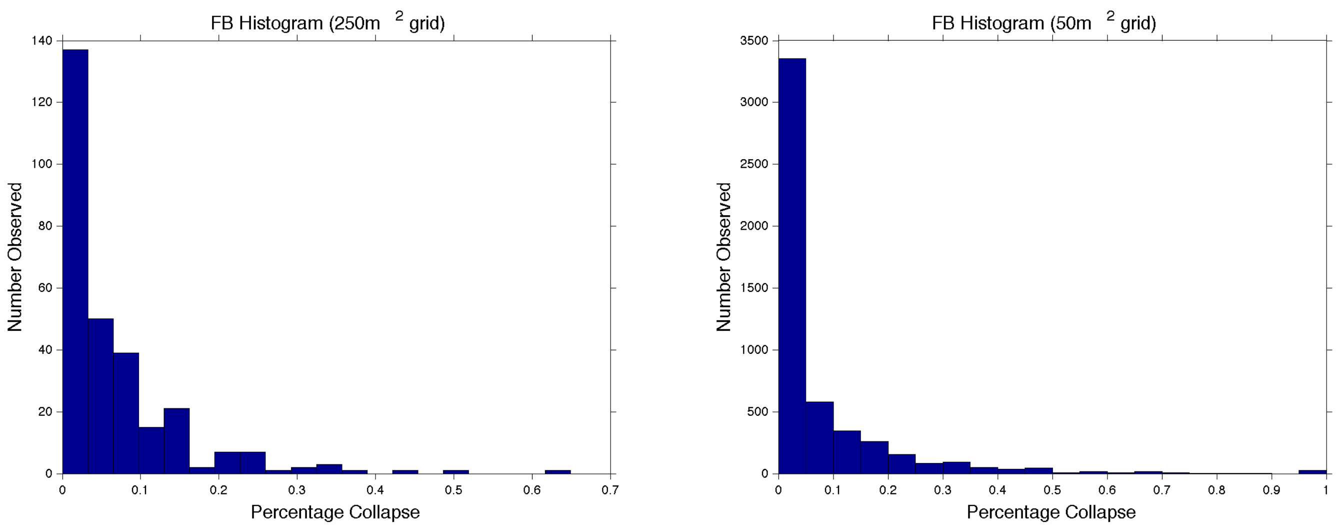


figure3c: FB building collapse distribution, at $250m^2$ (left) and $50m^2$ (right) resolution.

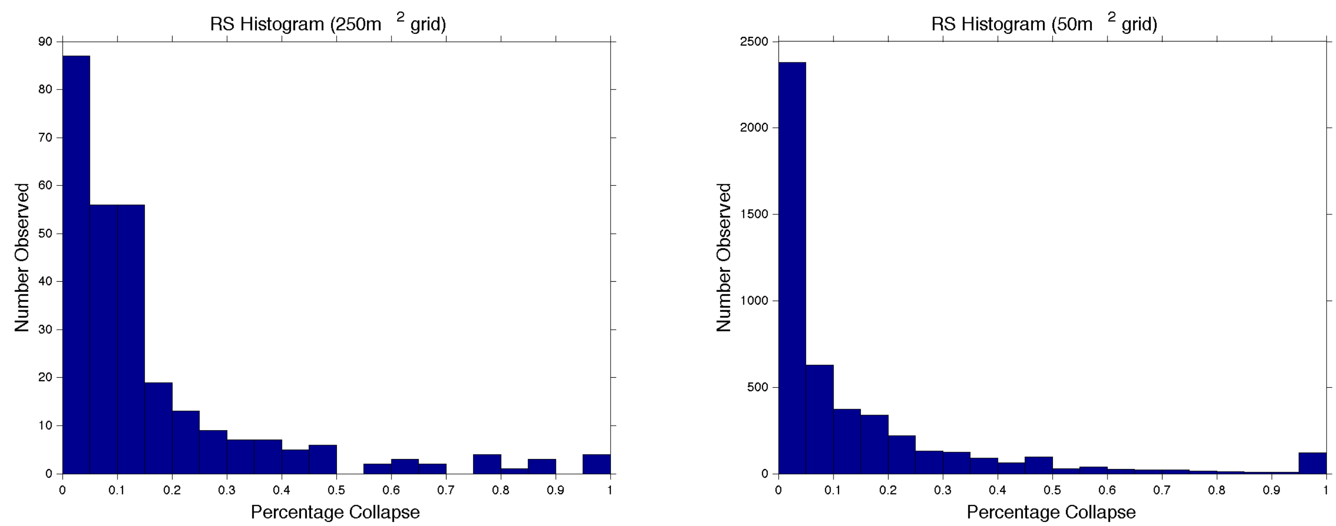


figure3d: RS building collapse distribution, at $250m^2$ (left) and $50m^2$ (right) resolution.

250m by 250m Grid			
Area of X(15km, 40km) by Y(0km,7.5km)			
	FB	RS	RS/FB
Number of Data Points	288	284	98.61%
Number of Buildings	140253	134578	95.95%
Number of Collapses	11913	17522	147.08%
Percentage Collapse	8.49%	13.02%	

50m by 50m Grid			
Area of X(15km, 40km) by Y(0km,7.5km)			
	FB	RS	RS/FB
Number of Data Points	5090	4750	93.32%
Number of Buildings	135907	129123	95.01%
Number of Collapses	11936	17146	143.65%
Percentage Collapse	8.78%	13.28%	

table2: FB and RS data for X(15km, 40km) by Y(0km,7.5km), at $250m^2$ (left) and $50m^2$ (right) resolution.

ii. Building Count Threshold

	250m by 250m Grid	50m by 50m Grid
	RS Data with Building Count Threshold	
	10 Buildings	1 Building
Number of Data Points	260	4582
Number of Buildings	134491	128955
Percentage of Pre-removal	99.94%	99.87%
Number of Collapses	17248	16979
Percentage Collapse	12.82%	13.17%

table2: FB building collapse percentage at $250m^2$ (left) and $50m^2$ (right) resolution

Table2 shows that about 100 percent of the RS building count remains, after setting a minimum of 10 buildings and 1 building threshold per grid for the $250m^2$ and $50m^2$ resolution data, respectively. The threshold of 10 for $250m^2$ is set arbitrarily; the threshold of 1 for $50m^2$ is set to match that of 10 in building count percentage change.

The thresholds are chosen to prevent RS data points with much lower building count having any weight in the analysis process. Indeed, collapse percentage data with a lower building count - and thus a lower denominator - is more sensitive to changes in the numerator, which represents both the collapse count and more importantly, the error in counting collapses.

Seen in figure4, for both of the $250m^2$ and $50m^2$ RS data, setting a threshold reduces the nuggets size, suggesting that setting a threshold is way of filtering out measurement error. FB data is not filtered because all of its values are assumed as correct - that the entirety of its nugget is micro-scale variability. The study will continue to use the "filtered" RS data.

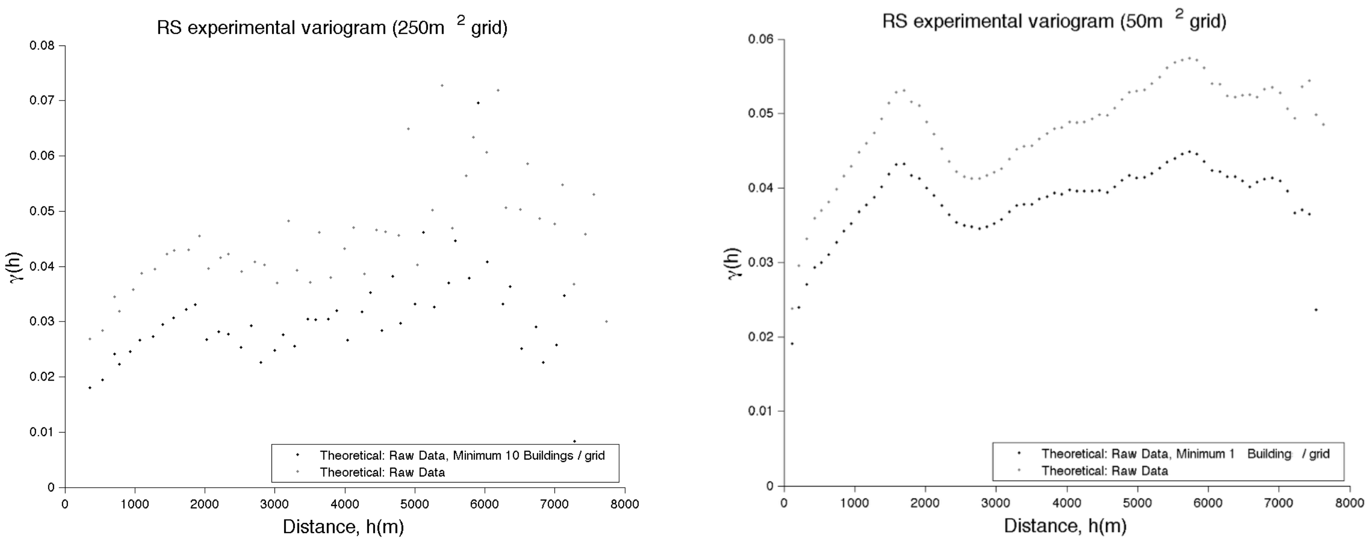


figure4: RS data sees a reduced nugget after setting a 10 building threshold, at $250m^2$ resolution (left); RS data at $50m^2$ resolution also sees a nugget reduction, with threshold of just 1 building (right).

iii. $250m^2$ *vs.* $50m^2$ *Data Resolution*

For all four dataset (FB and RS, and at $250m^2$ and $50m^2$), an exponential theoretical variogram is used. It should be noted that the y-axis of the FB graph is (0, 0.03), but is twice for RS, at (0, 0.06).

The $250m^2$ FB vairogram is almost entirely a nugget; its $50m^2$ resolution counterpart has a higher sill, a higher nugget and a short correlation length. It is expected that the nugget increases with higher resolution, as the intra-grid variability that would have otherwise been averaged at a lower resolution is now shown in the higher resolution (appendix illustrates further). Thus, the nugget increase should be attributed mostly to micro-scale variability.

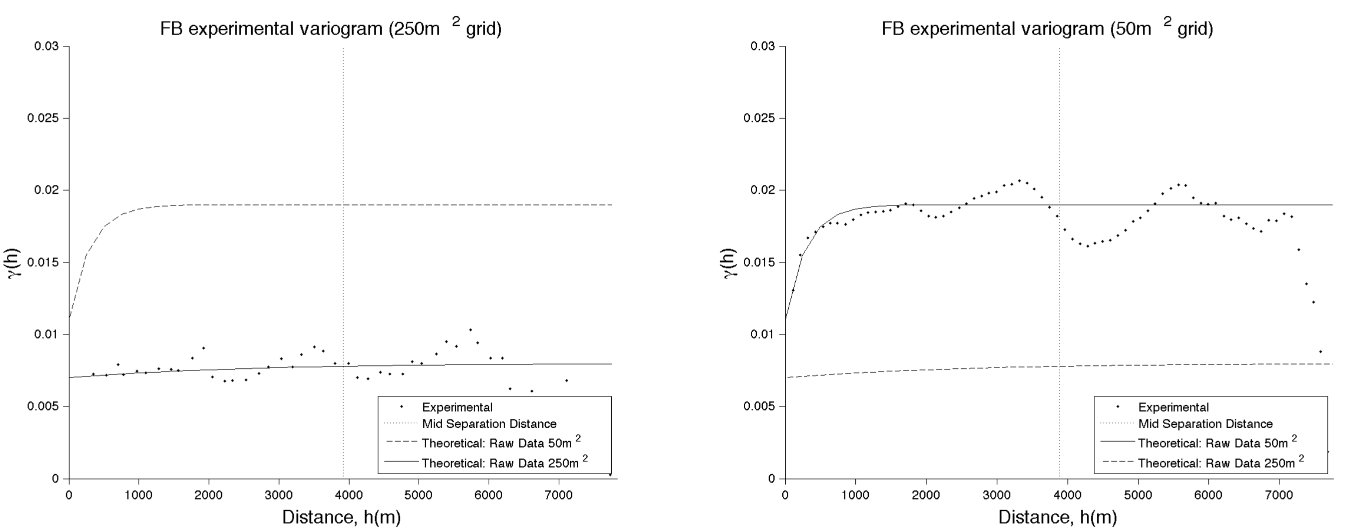


figure5a: FB variograms at $250m^2$ (left: nugget = 0.007, a = 2500, b = 0.008) and $50m^2$ (right: nugget = 0.011, a = 300, b = 0.019) resolution

Going from $250m^2$ to $50m^2$ resolution, RS yields a similar nugget increase as FB: $\Delta\gamma(h = 0) \sim 0.005$. RS's $250m^2$ resolution already has a relative sill and a correlation length, both of which are similar for its higher resolution.

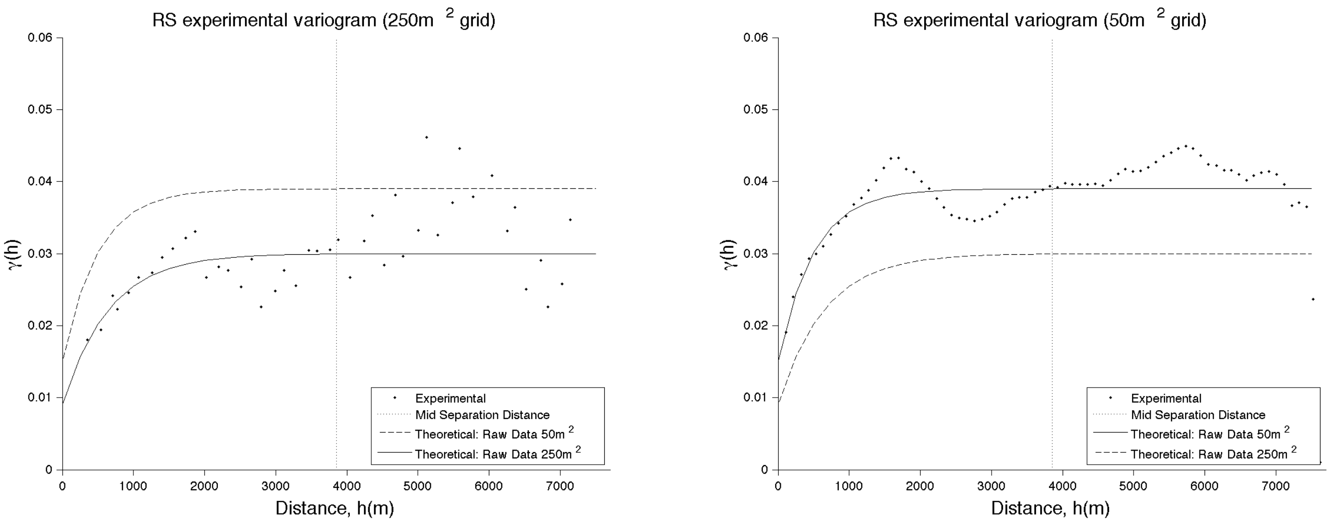


figure5b: RS variograms at $250m^2$ (left: nugget = 0.009, a = 650, b = 0.03) and $50m^2$ (right: nugget = 0.015, a = 500, b = 0.039) resolution

Overall, both FB and RS experimental variograms experience an abrupt change around h = 2,000m, which corresponds roughly to the the data range in the Y-direction. Thus, when fitting the variograms, focus is on the first 2,000m separation distance, so as to place equal emphasis on the Y- and X-direction. (Anisotropy is briefly explored, but because FB and RS data shows rather dissimilar major axis, anisotropy is not applied in this study).

Additionally, FB's mostly nugget variogram at $250m^2$ resolution suggests little spatial correlation at that resolution; indeed, even the $50m^2$ resolution only shows a relatively short correlation length of 900m, suggesting that any lower resolution could very well blur out the correlation. On the other hand, the fact that RS does show correlation at the same coarse resolution could imply a fundamental difference between FB and RS assessment: it is possible that during RS assessment, assessment experts are more inclined to label a given building as collapsed, after labeling more than a few immediately surrounding buildings as so. This influence might also explain RS's higher correlation length. This assessment influence could be less for FB, since each building is assessed in detail, and based on an objective checklist of criteria.

Since the variograms shows finer variabilities than the $250m^2$ resolution can capture, this study moves forward by using only the $50m^2$ data. Next, this study will determine if the variogram will further change from removing a potential trend.

iv. Universal Kriging with MMI

MMI, an indicator of structural damage, is explored as a potential source of trend. The MMI from USGS that is used here is an approximated indicator derived from ground motion measures (earthquake.usgs.gov/earthquakes/shakemap/background).

The MMI increment is chosen to be 0.25, and two ways of binning the percentage collapse for each MMI increment are explored: (1) simply averaging the percentages for all the data points within each MMI increment, and (2) dividing the total number of collapsed buildings by the total number of buildings within each MMI increment. The latter is arguable more true to the underlying data, whereas the former is more true to the data points used in this study.

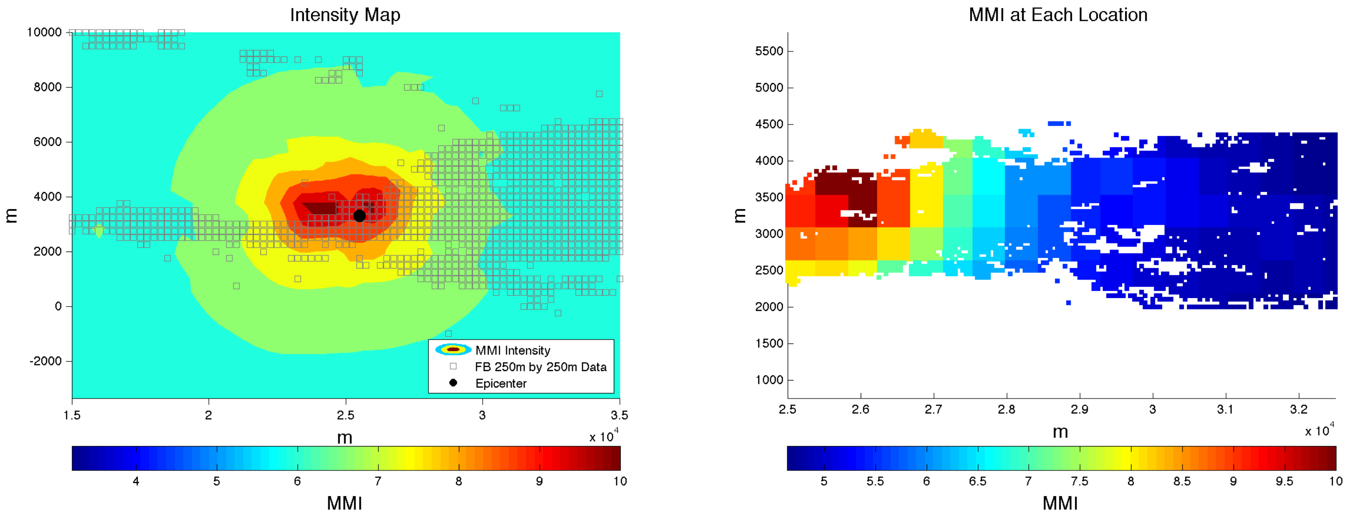


figure6a: MMI data from USGS (left) adjusted in location and scale to approximately fit the RS and FB data (right).

Most of the FB and RS data at the two resolutions show a 10 percentage collapse increase when MMI increases from 4 to 10 (figure6b). Yet, detrending neither FB nor RS data from MMI has a big effect on their variograms, or on the overall variances of the collapse percentage (table3). Further, depending on the binning method, the variance sometimes increases after detrending (table3 and figure6b).

This lack of variogram change is counterintuitive, since MMI is a product of earthquake intensity, as is building collapse. However, there are many other factors that could affect building collapses, all of which are assumed constant in this study, but are unlikely constant in reality - such as building types, inter-building distances and ground conditions. More importantly, the fact that the trend is sensitive to the binning method suggests a lack of trend fundamentally, as the lack of change in variance also suggests. Detrending, then, is not a viable option in this study.

Detrending									
	250m^2		50m^2			250m^2		50m^2	
(1) By Percentage	FB	RS	FB	RS	**(2) By Total Building**	FB	RS*	FB	RS*
Original Variance	0.0076	0.0282	0.0184	0.0378	Original Variance	0.0076	0.0282	0.0184	0.0378
Post-detrend Variance	0.007	0.0285	0.0179	0.0378	Post-detrend Variance	0.007	0.0293	0.018	0.038
*Trend is negative									

table3: Summary of detrending: very little changes to variances.

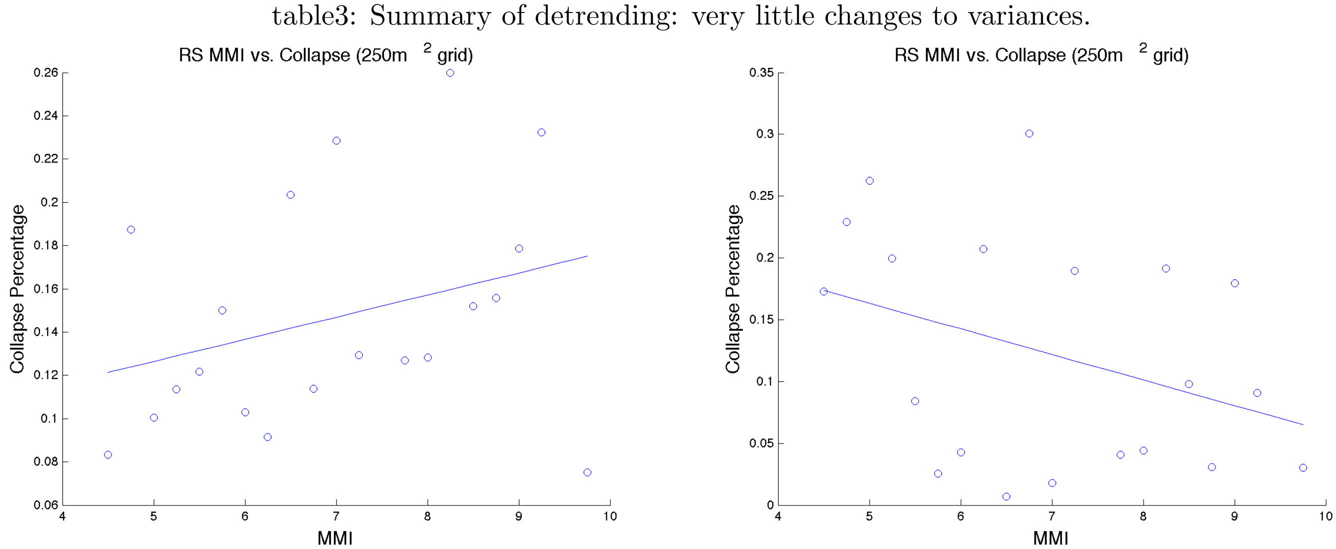


figure6b: An example of positive and negative trend: the 250m^2 resolution RS data has a positive trend if binned using method (1), but negative if binned using method (2)

v. Geostatistics vs. Classical Statistics in Reducing Uncertainties in Remote Sensing Data

Figure7 shows only slight resemblances between the FB and RS kriged results at the $50m^2$ resolution. These figures are, of course, indicative of their respective theoretical variograms and kriging method: (1) FB has some abrupt patches because ordinary kriging results collapse to the values of the raw data points at those points' locations; (2) RS appears smoother, in part because continuous part kriging allows for deviations from raw data point values, and in part because RS has a longer correlation length of 1500m, compared to FB's 900m (figure5a and figure5b).

Overall, RS appears to contain most of FB's spots with higher collapse percentages, though this is not the case vice versa. This visual supports the study's initial comparison, which shows RS having an overall higher collapse percentage than FB (table1). Next, this study will explore if data-cleaning with geostatistics and classical statistics can bring RS closer to FB.

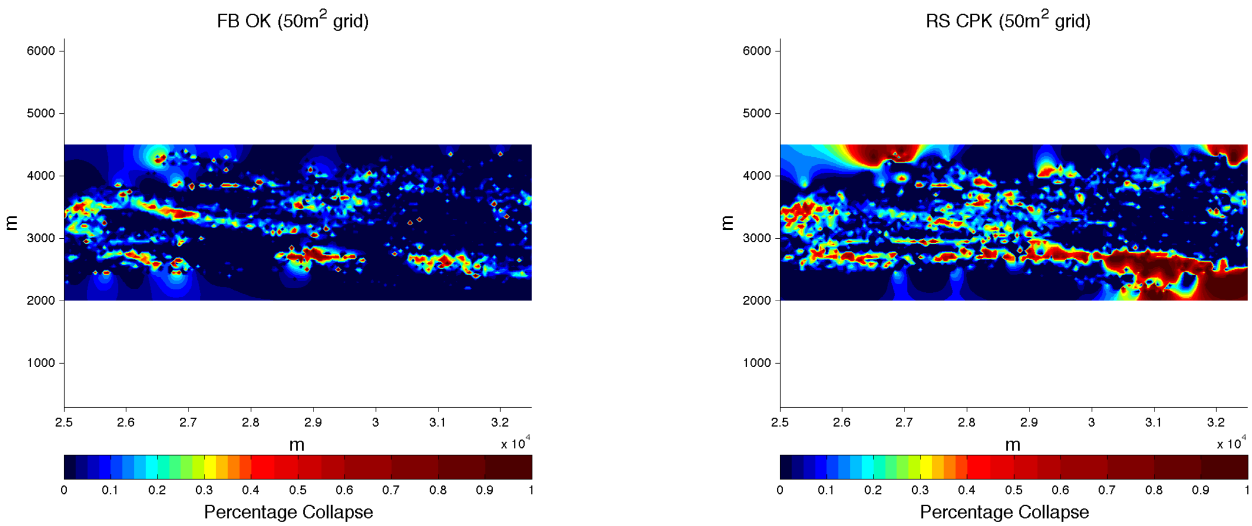


figure7: FB ordinary kriging (left) and RS continuous part kriging (right) at $50m^2$ resolution

Table4 is a duplicate of what was presented in the result summary, and shows that at an aggregate scale, data-cleaning with classical statistics improves RS better than with geostatistics. The total number of collapses is computed from multiplying the kriged percentages by the raw RS values. Nonetheless, the two method's difference is only 0.5 percent, and neither bridges the FB-RS discrepancy completely. Figure8a through figure8d compare the two different methods - illustrating, in the following order, (1) the data eliminated, (2) the variograms used, (3) the results from re-kriging the remaining data, and (4) the improvements made. These figures' captions will highlight their respective significance.

Removing Uncertainty						
					Remote Sensing Data within 2 Stdev	
	FB	FB_CPK	RS	RS_CPK	Variance from CPK	Variance from Classical Statistics
Number of Collapses	11936	12037	17085	17029	14819	14090
Percentage Collapse	8.78%	8.86%	13.25%	13.21%	11.49%	10.93%

table4: Summary result.

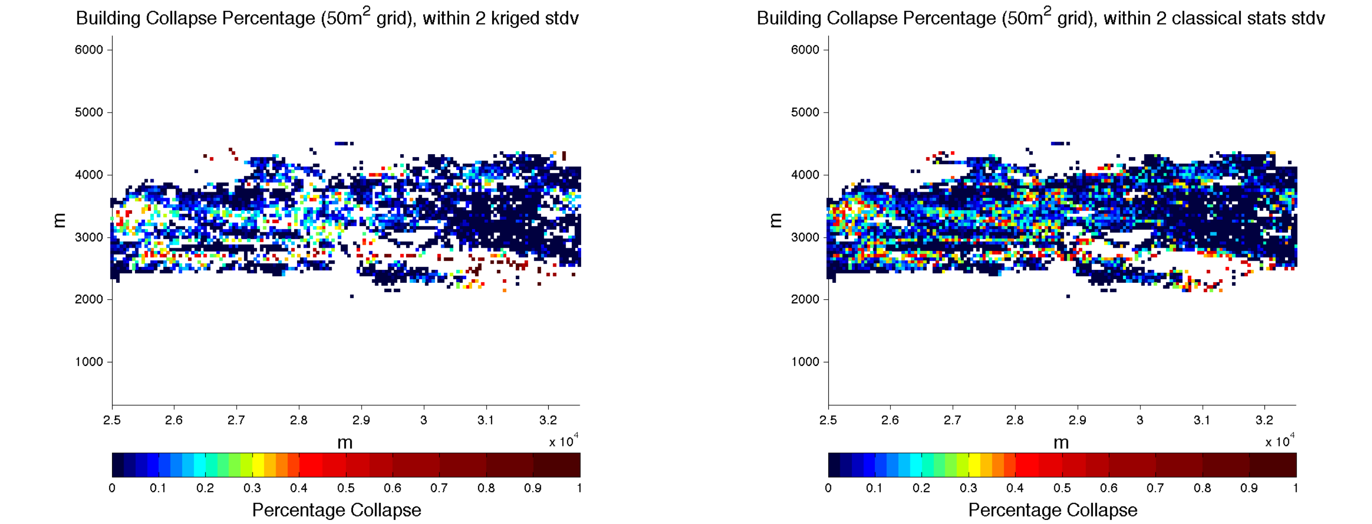


figure8a: A much smaller subset of the raw data is within two kriged standard deviations of the kriged mean (left), whereas much of the raw data remains within two standard deviations of the classical statistical mean (right). The difference is expected, because (1) the kriged mean is spatially variant and the kriged standard deviation is smaller compared to the aggregate standard deviation, and because (2) keeping two standard deviations within a classical statistical mean is by definition keeping roughly 95 percent of the raw data (though this is not strictly the case here, as the data has an exponential distriubtion).

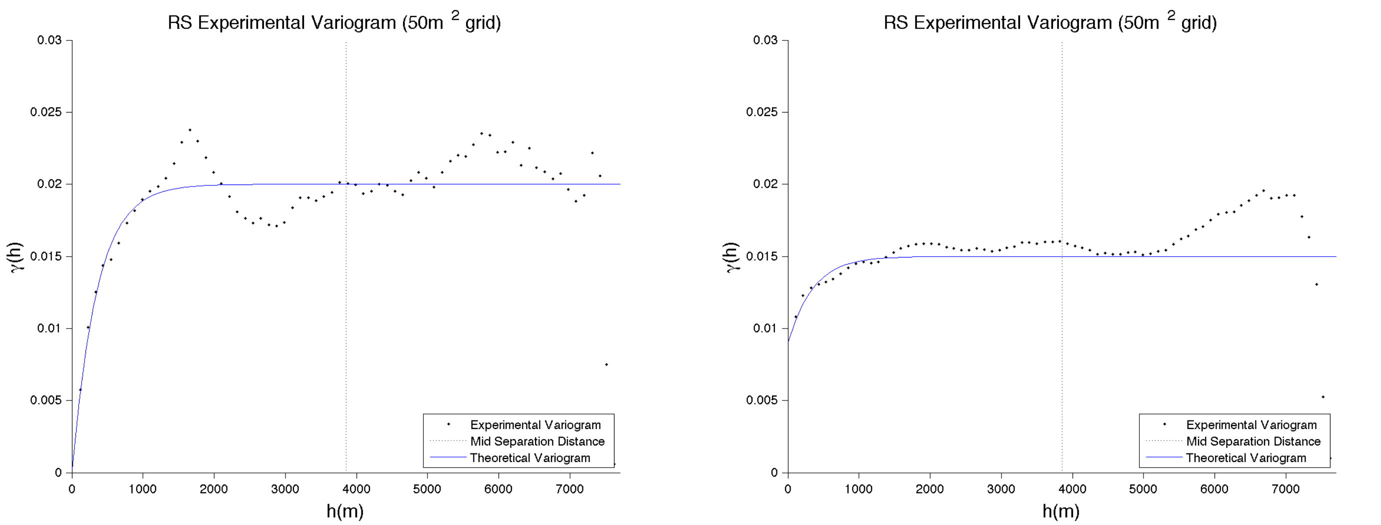


figure8b: The variogram structure of the remaining data is also reflective its data-cleaning process, especially when compared to the raw data (which has: nugget = 0.015, a = 500, b = 0.039). The geostatistics method brings the nugget essentially down to zero, and thereby eliminating all measurement error and micro-scale variability (left: nugget = 0, a = 400, b = 0.02). The relative sill decreases only a little for the geostatistical method (0.004), but quite a bit for the classical method (0.018). Having a sharp sill decrease, the classical method eliminates much of the data's spatial structure (right: nugget = 0.009; a = 350; b = 0.015). The change in correlation length for both methods is minor.

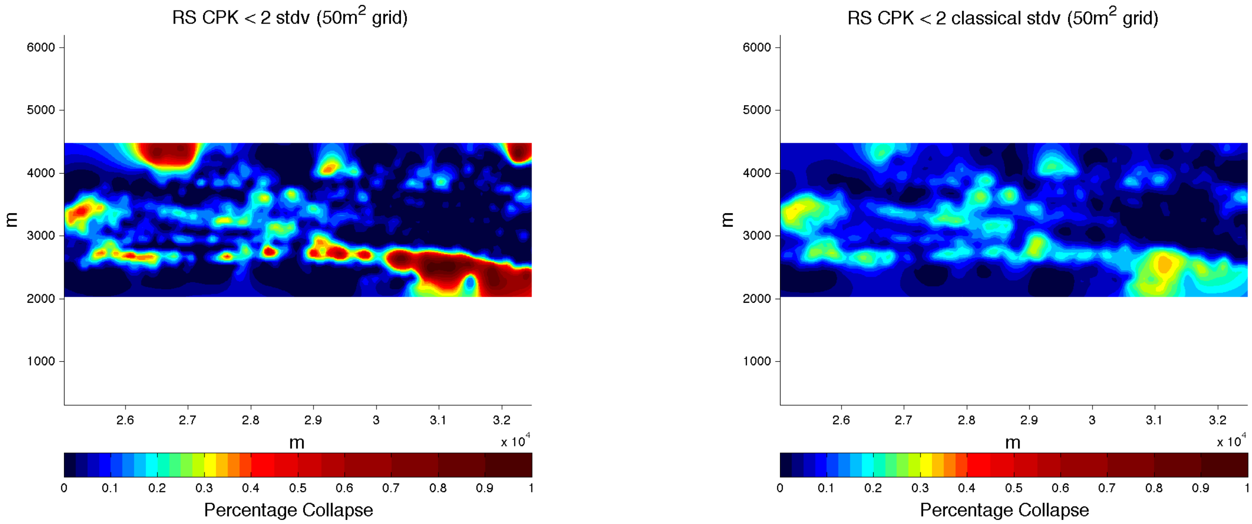


figure8c: The re-kriged results with figure8b's variograms restate those's variogram's characteristics: both smooth the map (as elimination of errors should do), but the geostatistics method keeps much of collapse percentages' spatial variations, while the classical method seems to have just leveled the area's collapse percentages.

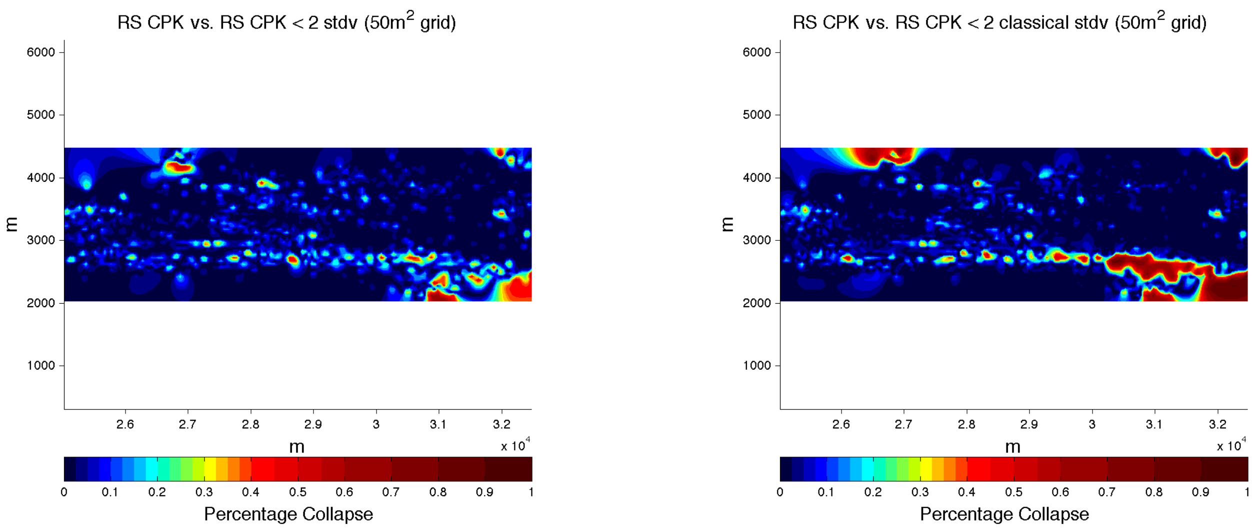


figure8d: The differences between pre- and post-data cleaning for the geostatistics method (left) and the classical method (right) are shown; the classical method eliminates more of the higher collapse percentage values that were originally present in the northwest and southeast region of the data domain.

The fact that reducing RS data's spatial structure (as is done by the classical statistics method) brings RS closer to FB on the aggregate scale begs the question if FB and RS are *spatially* comparable in the first place. Figure9a maps the difference between RS and FB - and this map being similar to RS's original distribution (figure7) suggests that RS and FB collapse concentrations are in different locations. Similarly, both FB and RS's short correlation length precludes much overlapping in the kriged values as well.

With this dissimilarity in mind, this study cannot state if the form of goestatistical data-cleaning presented is an appropriate method to reduce the uncertainties in RS data. However, it can perhaps be argued that a classical data-cleaning method is inappropriate, since it reduces the data's spatial structure. Further, because the collapse distribution is exponential and thus positively skewed, eliminating data beyond two standard deviations would lower the mean, and in this case, the FB collapse percentage just happens to be in the direction of this lowering. It is then no surprise that the result from the classical method is closer to FB (figure9c), than is the geostatistical method (figure9b).

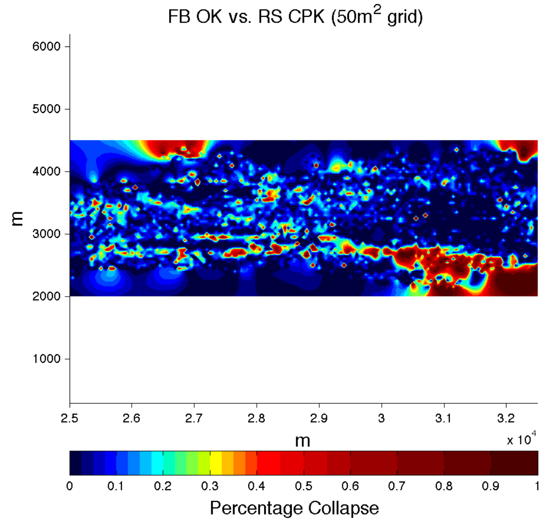


figure9a: Difference between RS's frist continuous part kriging and FB's ordinary kriging.

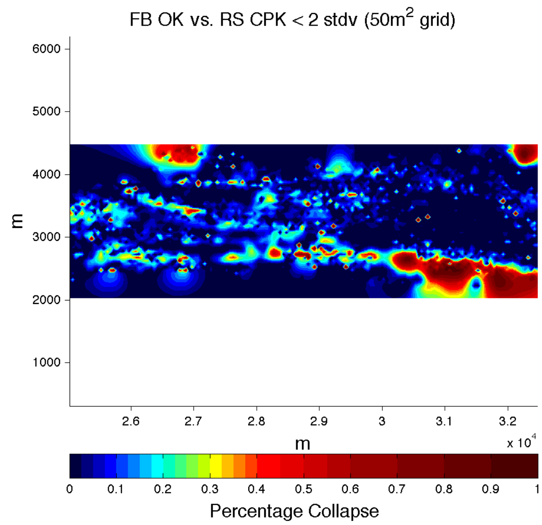


figure9b: Difference between RS's second continuous part krigingm, after geostatistical data-cleaning, and FB's ordinary kriging.

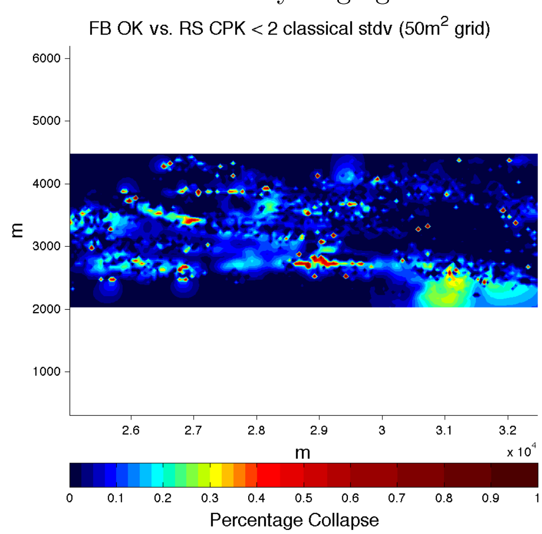


figure9c:Difference between RS's second continuous part kriging, after classical statistical data-cleaning, and FB's ordinary kriging.

7. CONCLUSION

This study shows that classical statistics is likely ill-suited for reducing the uncertainties in a spatially correlated dataset, as the method reduces the dataset's spatial structure. However, this does not mean that the alternative is strictly better: there remains the need to confirm if a combination of manual data cleaning by setting a threshold, and geostatistical data cleaning by kriging is optimal for reducing RS uncertainties. The former is saying that we do not want points containing too few data having any weight in the analysis; the latter is saying that points with a sizable number of data may still not be entirely accurate. Further, a central question remains: how comparable are FB and RS? If even after eliminating the nugget, RS still appears quite removed from FB, could we say that RS is simply systematically different from FB?

Part of that answer could potentially be found among the many other parameters involved in the study. In particular, choosing a different geographic domain or data resolution could very well have changed the outcome of this study. The sensitivity of these parameters remain unknown, so given more time, this study wishes to further explore the comparability issue by first understanding the influences of these parameters. Some sample future explorations include:

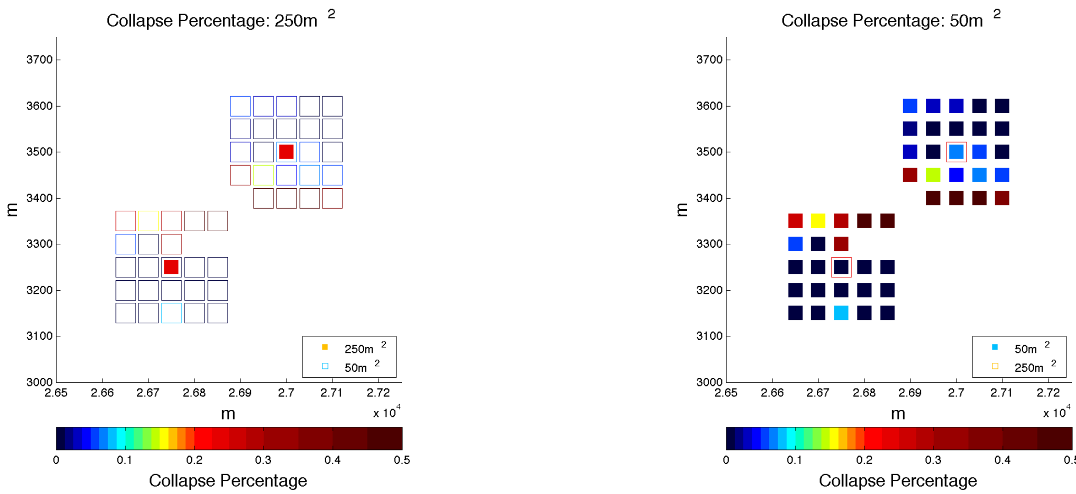
i. Weigh each data point proportional to its building count: data points with higher building counts should have less measurement error, and therefore should not be counted as equivalent to those with less building counts.

ii. Focus on an even smaller subset of the data: since the correlation length is rather short, and since the variogram appears to change form after reaching the Y-axis length, focusing on a smaller subset at higher resolution could potentially highlight finer variations, which would have been computationally impossible to achieve had the area been larger.

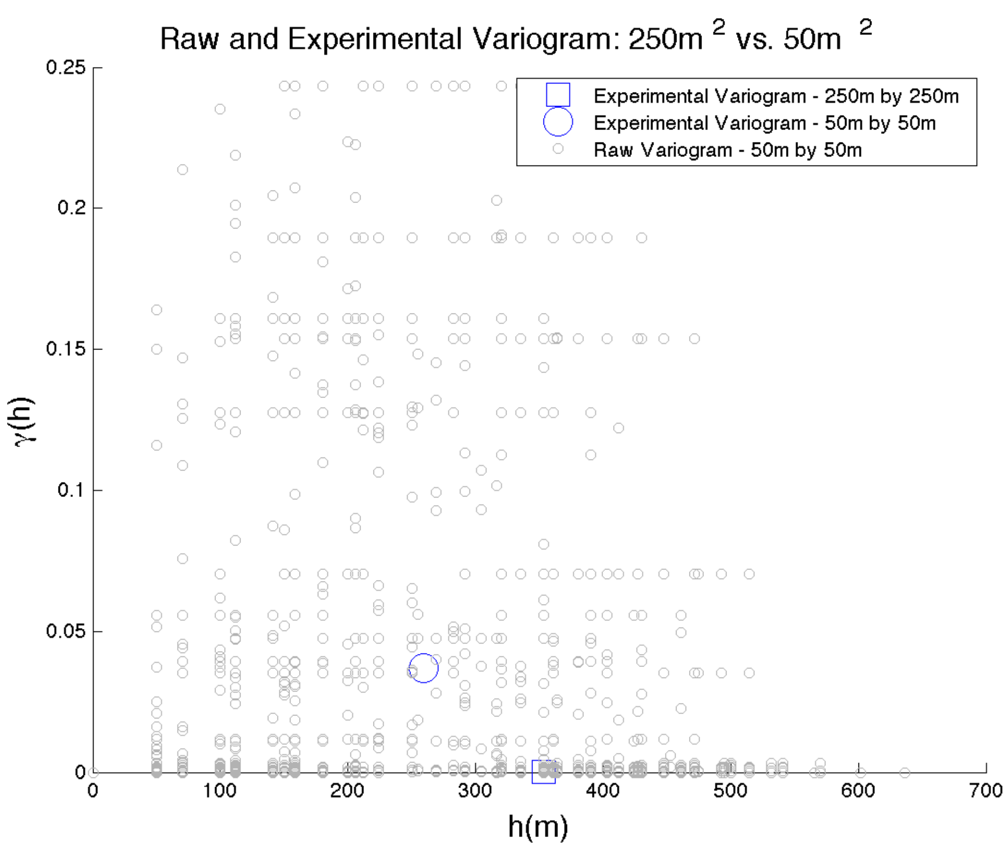
iii. In addition to the intensity measure shown in this study, incorporate other earthquake information, such as soil type and distance from the epicenter. Further, see if co-kriging would be better than universal kriging, since these earthquake information are also random variables (though arguably with less uncertainties than building collapse assessment).

A. APPENDIX: RESOLUTION STUDY

FigureA1 shows the content within two adjacent $250m^2$ resolution data points (left), when its resolution increases to $50m^2$ (right). Both are chosen from the FB dataset, and at the coordinates shown on the axis. The two points at the $250m^2$ resolution contain very similar values, and thus have a semivariance of about 0 (figureA2). However, those two points have intra-grid variations, and thus have a higher average semivariance at the higher $50m^2$ resolution. (Note: the color range is adjusted for these figures to range from 0 to 50 percent collapse, as suppose to 0 to 100, in order to highlight the value variations).



figureA1: FB building collapse percentage at $250m^2$ (left) and $50m^2$ (right) resolution



figureA2: FB collapse percentage raw and experimental variogram comparison.