

Chapter 6

Case Studies

This chapter continues the discussion that began in chapter 5 where an approach to geostatistical coverage mapping for wireless networks was described. In this chapter, those methods will be applied in two real-world applications: coverage mapping of WiMax and LTE networks on the CU campus. In addition, sections 6.2 and 6.4 will discuss two important additional topics: map combining, and modeling deviations from stationarity.

6.1 Case Study: University of Colorado WiMax

This section describes the first case study conducted specifically for the purpose of evaluating the efficacy of Kriging-based coverage mapping. This is an important step, since the exploratory examples described in chapter 5 made use of drive-test data, which may be ill-suited for coverage mapping due to a sampling bias towards “street canyons”. The aim here is to map the coverage of five WiMax BSs deployed on the University of Colorado campus operating at 2.5 GHz within an educational spectrum license held by the University. An initial sample is taken on a uniform equalateral triangular lattice with a lag of 100 m. To constrain the data collection, measurements are confined to the main University of Colorado campus. Figure 6.1 shows the main campus along with points where samples were collected. The shape of the University is vaguely triangular, with the hypotenuse measuring 1.5 km and the shorter side measuring 1.1 km, giving a total measurement area of slightly more than $825m^2$.

Of the five WiMax BSs being studied, four are managed by the University of Colorado Office of Information and Technology (OIT) and primarily provide backhaul coverage to buses in and around Boulder

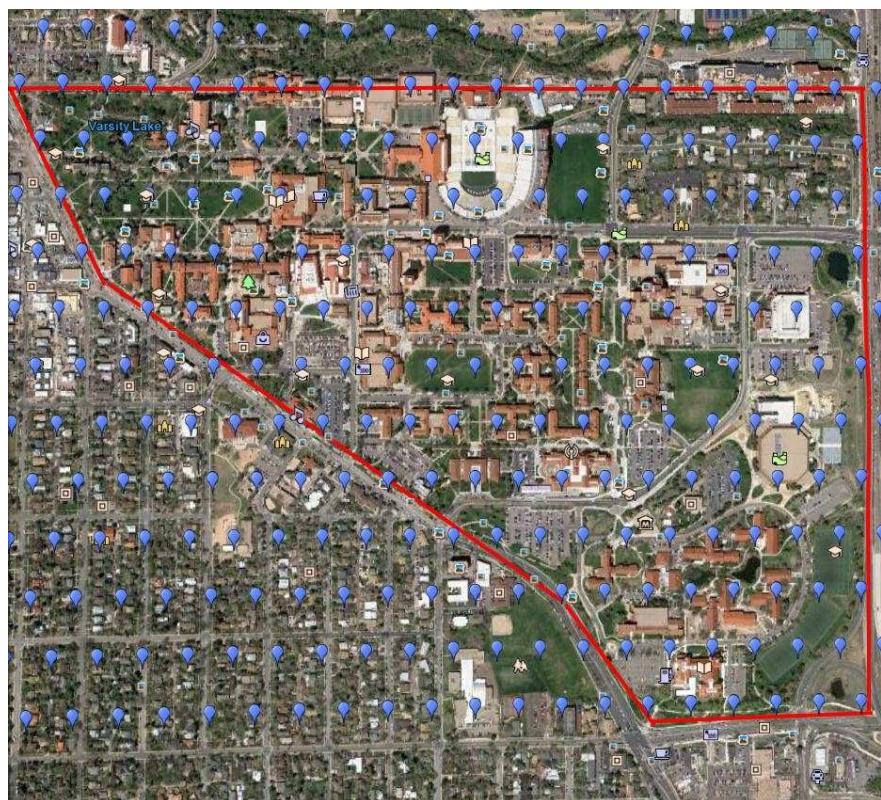


Figure 6.1: Map of University of Colorado and 100m uniform equalateral triangular sample. Measurements are limited to the main campus, which is outlined in red.

Name	Dir.	Freq.	Longitude	Latitude	Easting	Northing	AGL (m)
Gamow West (cuGW)	235	2530	-105.267778	40.008056	2017383.55	4582293.93	46
Gamow East (cuGE)	90	2520	-105.267778	40.008056	2017383.55	4582293.93	46
Eng East (cuEE)	120	2530	-105.263056	40.007222	2017808.74	4582284.17	34
Eng West (cuEW)	240	2510	-105.263333	40.007222	2017784.92	4582279.27	34
Eng North (cuEN)	0	2578	-105.263333	40.007222	2017784.92	4582279.27	34

Table 6.1: Specifications of five University of Colorado WiMax BSs.

[39]. The fifth is a Global Environment for Networking Innovation (GENI) testbed node used for research purposes [182]. Table 6.1 provides details about the location and configuration of each BS¹. All nodes use a channel bandwidth of 10 MHz, have 90-degree sector antenna (excepting the GENI node which has a 120 degree sector), and operate at a nominal transmit power of 40 dBm. Two BSs are deployed on the Gamow Physics Tower (pointing east and west) and three on the Engineering Center tower (pointing north, east, and west). The cuGW and cuEE nodes are deployed on the same frequency, while the other BSs each have their own frequency. Because the spectrum analyzer used for measurement has no way of differentiating between different BSs on the same frequency and will simply record the strongest measurement observed, these two BSs will be measured and modeled as if they were a single BS with two antennas.

6.1.1 Measurement Apparatus and Procedure

In order to make measurements in arbitrary locations, which might not be accessible with a large vehicle, a measurement apparatus was constructed especially for this data collection campaign, built into a small cart. The cart provides a stable platform on two wheels and can be connected to a bicycle or used as a hand-cart. To collect measurements, an Anritsu MS2721B portable spectrum analyzer is used. This analyzer is unique in that it is both battery-powered and portable, as well as having the ability to demodulate WiMax transmissions. A netbook laptop running Ubuntu Linux is connected to the spectrum analyzer with a single Category 5 (CAT5) crossover cable. This laptop controls the spectrum analyzer using a series of Virtual Instrument Software Architecture (VISA) commands, which creates the ability for measurement scripting on the laptop. Two GPS devices are used to record position, one connected to the spectrum analyzer and

¹ Unless otherwise specified, all latitude and longitude coordinates are given in WGS84/EPSC:4326 and UTM coordinates in EPSG:32160.

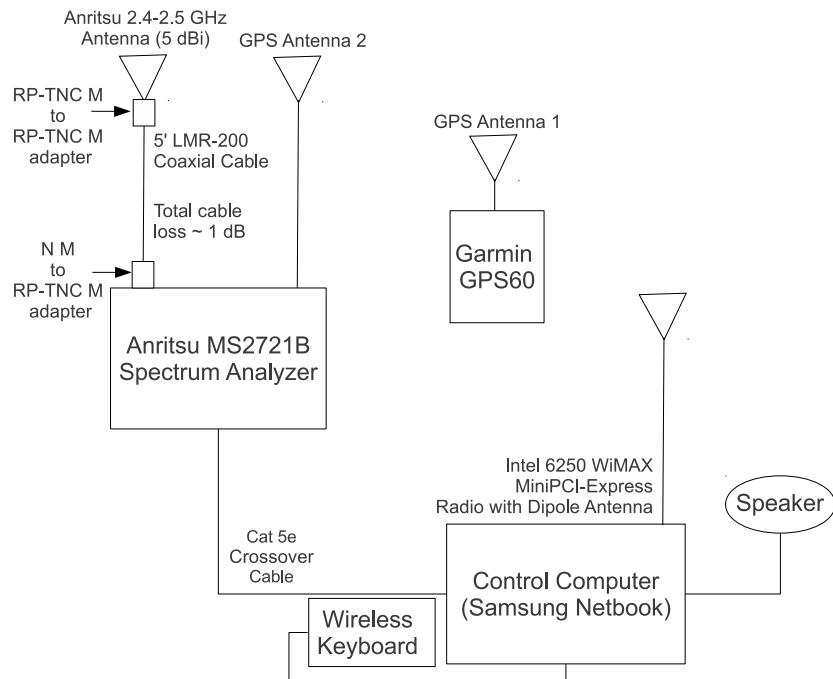


Figure 6.2: Diagram showing connectivity and specification of WiMax measurement cart devices.

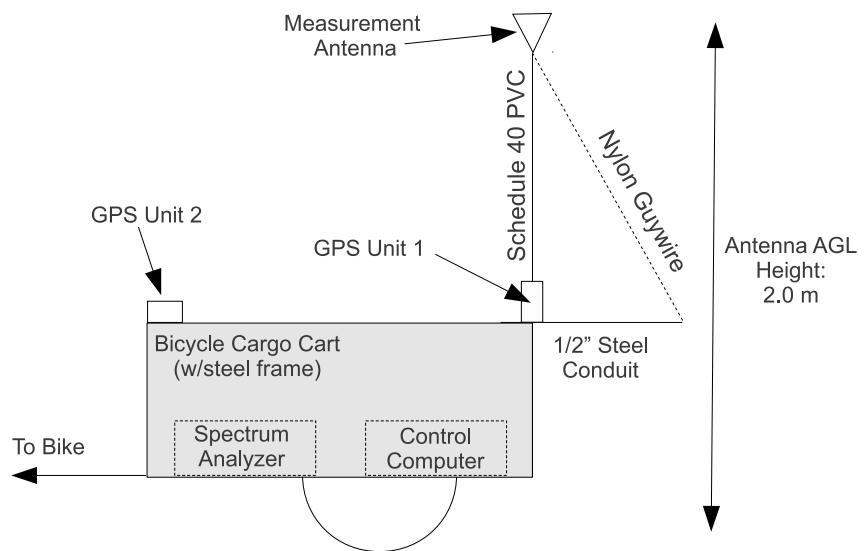


Figure 6.3: Diagram of WiMax measurement cart.

one a hand-held Garmin GPS60 device². The measurement antenna for the spectrum analyzer is raised 2 m from the ground using a piece of schedule-40 Polyvinyl Chloride (PVC) (non-conductive) pipe, and attached with plastic cable ties. Although the cart itself is conducting, care is made to ensure that no metallic objects are in close proximity to the elevated measurement antenna. Figure 6.3 and 6.2 show the design and layout of the measurement cart.

The measurement effort focuses on four important first-order metrics of channel performance: Carrier to Interference and Noise Ratio (CINR), Relative Constellation Error (RCE), Error Vector Magnitude (EVM), and subcarrier spectrum flatness. CINR provides a measurement of pure received power above noise, calculated from a clean carrier wave transmitted in the preamble of the WiMax frames. RCE and EVM quantify the amount of error in a binary or quaternary constellation plot, which provides a tight estimate of physical-layer error. Finally, subcarrier spectrum flatness is the amount of gain or attenuation on each of 52 (or more) subcarriers within the bandwidth relative to the mean signal strength. Using the spectrum flatness data, it is possible to calculate Effective Signal to Noise Ratio (ESNR), the metric shown in [90] to be a strong predictor of actual network performance (as compared to the more traditional metrics such as SNR and RSS).

Before beginning measurement, a policy must be defined for locating and measuring at sample sites. After some experimentation with direct location using a GPS device, a simple solution was chosen involving a printed map similar to 6.1. Each site is visited without any particular order. In the event that it is impossible to make a measurement at the site, either because it falls in an inaccessible (e.g., fenced) area or within a building, a measurement is made at the closest point (by straight line distance) that is measurable. Although there is some random error associated with locating points (due to GPS accuracy, point finding, and obstacles), this error is not harmfully aligned with any environmental feature and instead amounts to random jitter about the uniformly selected sample sites (which some spatial sampling studies have actually purposely advocated).

At each measurement location, a wireless keyboard is used to manage the control computer (which keeps the experimenter away from the apparatus, preventing them from interfering with the measurements

² A hand-held GPS device was chosen after finding the Anritsu's GPS reception to be weak and unreliable.

themselves) and the control computer provides feedback through an amplified speaker utilizing text-to-speech synthesis software. At each point, three repeated measurements were made of downstream system performance using various metrics. At a subset of points, additional clustered measurements were taken within a 40 wavelength radius of each true point. The combination of repeating measurements in time and space allows for accurate estimation and averaging of intrinsic channel variability due to small scale fading effects. The device first picks a given channel (carrier frequency) and then records all metrics for each measurement in turn. Then it switches to a different channel and repeats. While the device is performing measurements, the instrumenter uses the handheld GPS device to record the current position, sample location (each sample site is assigned a unique identifier), and GPS accuracy. At the end of a measurement effort (typically when the analyzer's battery is flat), the cart is returned to the lab for charging and data offload. The spectrum analyzer stores measurements in a proprietary, but plaintext, format that can be easily parsed.

6.1.2 Comparison of Performance Metrics

In this measurement campaign, several performance metrics besides the classic signal strength or SNR-equivalent metrics were collected. One question that naturally arises is: are these more robust metrics trivially correlated with simple and easy to collect metrics such as CINR? Figure 6.4 plots the relationship between CINR and each of the other metrics studied. RCE and EVM appear to be a simple (but nonlinear) function of CINR, at least as calculated by the spectrum analyzer used. There are several ways that EVM can be calculated from the constellation plot and observed power of constellation points, and it appears that the Anritsu spectrum analyzer is calculating EVM from CINR or vice versa. RCE is calculated directly from the EVM value and hence is equivalent. Given this, RCE and EVM do not provide novel information above and beyond what is provided by the CINR measurement. It is worth noting that in the process of data collection, a complete constellation plot is recorded for each measurement so these metrics could be calculated *ex post facto*. The relationship between ESNR and CINR is less trivial, especially for the lower (Phase Shift Keying (PSK) modulation based) bitrates. The higher bitrates, which use Quadrature Amplitude Modulation (QAM), tend to have a fairly well-defined linear correlation with CINR. This suggests that in cases where information about spectrum flatness is unavailable, ESNR₅₄ can be roughly approximated using

CINR measurements.

Knowing which metrics provide unique information about the channel, a natural followup question is: are these metrics correlated with application-layer performance? To answer this question, a measurement campaign was devised to perform throughput tests to the cuEN BS, the only one of the four BSs which is not in production use (and hence could produce clean throughput measurements), and was accessible at the time of measurement. In order to use a sufficiently diverse data set, measurements from a random sample of points around campus (which typically are NLOS to the antenna) were combined with measurements at regularly spaced intervals down a street in the LOS path of the BS antenna. At each point, a three-sample measurement was taken using the method described in the previous section. At the same time a netbook running Microsoft Windows XP, using a Accton Wireless Broadband Corporation (AWB) US210 wireless network adapter³ was used to collect a series of throughput measurements. Because it is difficult to fully saturate a high bandwidth link using an application-space packet generator (in Windows, especially), the choice was made to rapidly generate a large number of User Datagram Protocol (UDP) packets at a selection of packet sizes. The open-source tool nping [204], was used to generate 5,000 packets each using a payload of 80 bytes, 256 bytes, 512 bytes, 1024 bytes, and 1400 bytes. This results in 25,000 packets being transferred at the maximum rate possible, which amounts to 17,460,000 bytes, or 16.64 megabytes. To ensure that only the wireless system was being tested, the endpoint used was the BS Access Service Network (ASN) gateway controller, which was running Ubuntu Linux. Both the netbook and the ASN system collected a complete trace of sent and received packets using the tcpdump tool (or Windump in the case of Windows) [219, 221]. At the start of a throughput test, the measurer would attempt to make a connection to the BS, if a connection was obtained, the test would start by sending the upstream packets. Upon receipt of the last upstream packet, the daemon on the ASN server would wait five seconds and then begin sending an identical volley of packets downstream. The tracefiles were analyzed *ex post facto* to determine the loss rates and per-payload-size throughput.

To analyze how well physical-layer metrics collected with the spectrum analyzer predict application layer performance, an ANOVA was performed against each metric and upstream and downstream through-

³ This adapter was tested and approved by the GENI WiMax project.

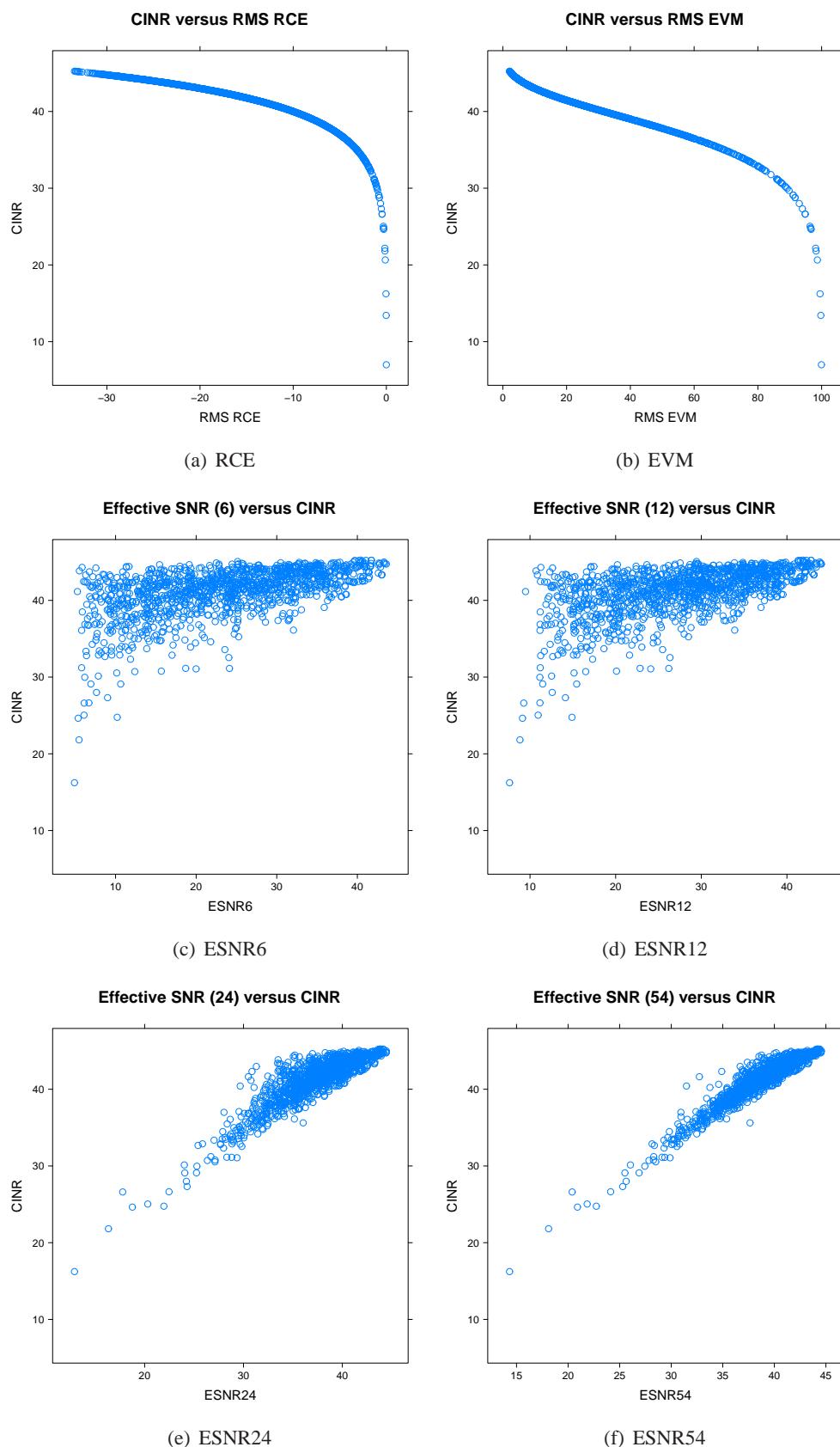


Figure 6.4: Correlation between various metrics.

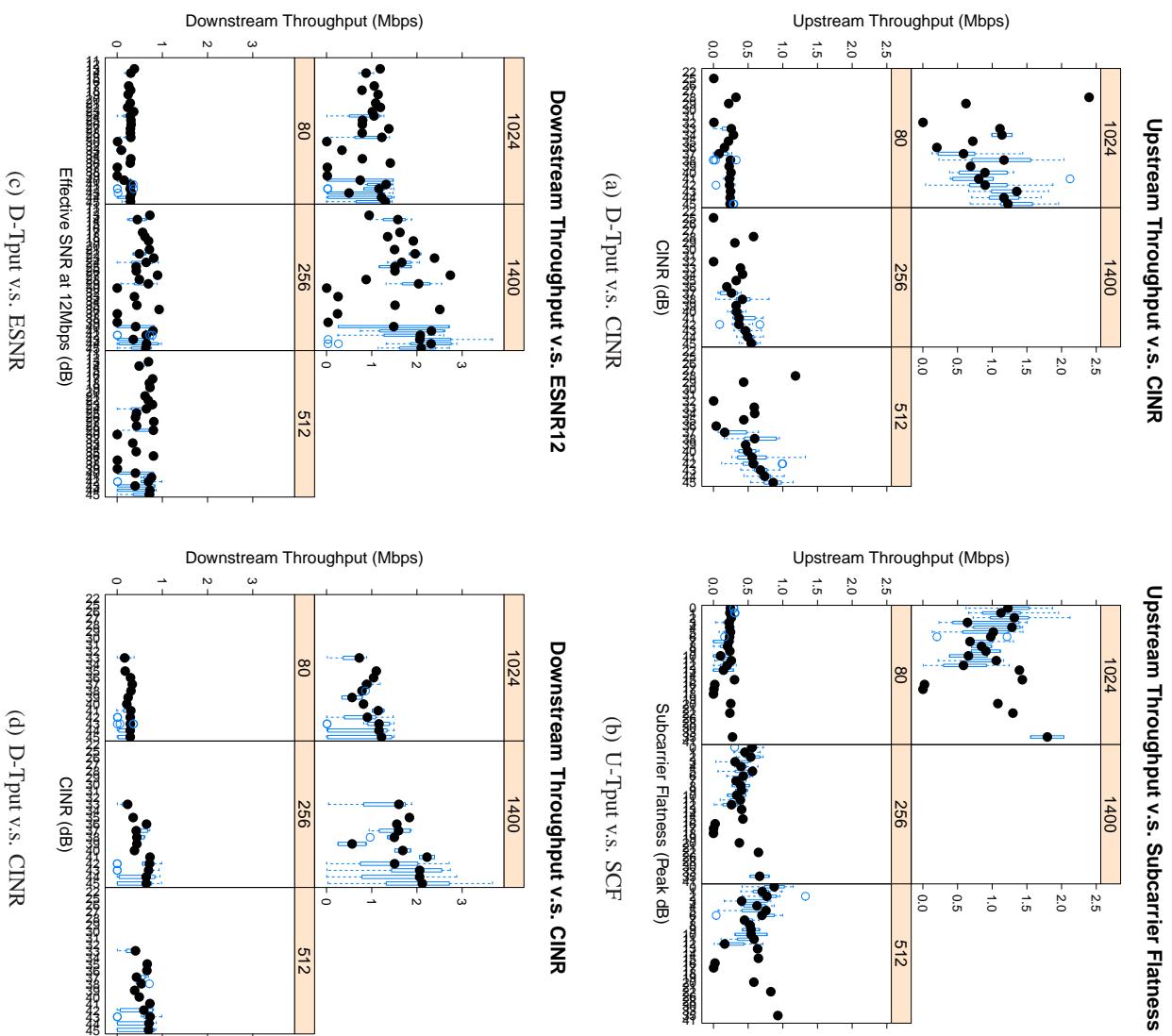


Figure 6.5: Correlation between WiMax throughput and metrics that are meaningfully correlated with application-layer performance.

put. For upstream throughput, the most significant correlations were present with CINR and Subcarrier Flatness (SCF), with a $p - value << 0.05$ for each and a $F - value$ of 28.10 and 12.45, respectively. Downstream measurements are more elusive—the only metric that has a significant correaltion with a $p - value < 0.05$ was ESNR, with an F-value of 4.86. Figure 6.5 shows the relationship of these metrics. Clearly there is a linear correlation in each case, however the correlations are very noisy. This noise is likely a result of unreliable components, in addition to the expected environmental noise. For instance, the ASN gateway server had a tendency to become loaded when generating or receiving a volly of packets and drop some frames. Similarly, the Winpcap library appears to be much slower than the linux-based equivalent and simply cannot generate packets at the same rate. Some packets, such as the smallest ones, were silently dropped by the AWB usb dongle (whose queue presumably overflowed when trying to send the 80 byte packets, and whose Maximum Transmission Unit (MTU) size appears to prohibit the 1400 byte packets). Despite this unavaoidable noise, there is a clear and distinct correlation present, which supports the use of CINR (which differs from SNR in that it is calculated from a clean carrier signal sent during the packet preamble) and spectrum-flatness-based metrics such as SCF and ESNR, supporting the findings of Halperin *et al.* in [90]. These results also suggest that, in terms of CINR, there is a clear division in performance above and below 40 dB, a threshold which can be used much in the same way Robinson et. al use 20 dB for WiFi networks when locating coverage “holes” [200].

As a final curiosity, GPS receiver accuracy is compared to CINR in figure 6.6. While making measurements, locations with the highest signal seemed to also have a high GPS precision (small accuracy value). This stands to reason since GPS also operates at 2.5 GHz and is effected by multipath interference from obstructions, in particular buildings and street canyons in urban environments. However, as shown by the figure, in practice the correlation is weak and noisy (certainly too noisy for planning purposes).

6.1.3 Possible Sources of Systematic Sampling Error

During the measurement campaign, three individuals used the cart to make measurements. Although all three measurers were collecting measurements using the same procedure, one possible source of systematic error is from the measurers themselves. Figure 6.7 shows the location error and GPS accuracy as

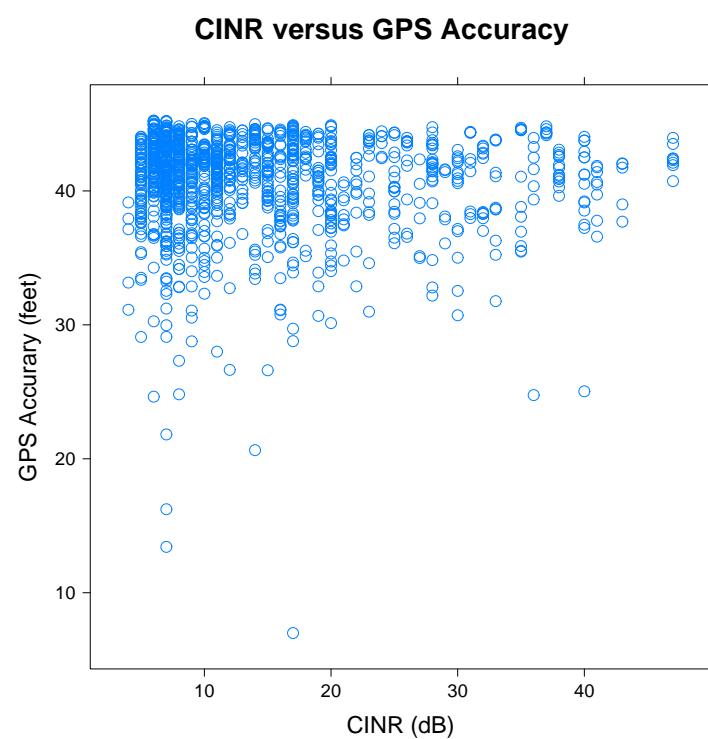


Figure 6.6: Correlation between CINR and GPS accuracy

a function of which measurer performed the measurements. There appears to be no discernable correlation and hence we do not need to correct for this bias in subsequent analysis. It is worth noting that some measurements are distant from their intended location. As discussed above, this occurs when a point is unreachable in practice. So long as the new measurement point is as close to the original measurement location as possible and there is no systematic error or systematic terrain alignment, these deviations should not effect the quality of the sample.

6.1.4 Spatial Data Characterization and Variogram Fitting

Figure 6.8 shows measurements taken for the cuEN BS, highlighting four metrics of interest: path loss computed from CINR, ESNR6 (ESNR for 6 Megabits per second (Mbps)), and ESNR54 (ESNR for 54 Mbps). The other four BSs produce similar plots. All four metrics produce a similar spatial distribution of values with large path loss or error values to the southwest and smaller (better) values to the north. All metrics have different value distributions, but the ESNR54 and CINR metrics appear to share the same basic skewed lognormal shape. Figure 6.8 shows the fitted relationship between path loss and distance for the three SNR-like metrics. The fits are not fantastic, but appear to at least account for some basic trend, which we can be removed to improve the efficacy of the Kriging process. The group of measurements around 155 dB are the “inferred” null measurements. For the SNR-like metrics, 1.0 is used for the null measurement and for EVM, 100 (i.e., 100% probability of error) is used.

After de-trending and accounting for “null” measurements, the next step is variogram fitting. Figure 6.10 shows the fitted variograms for cuEN using the same procedure as described in the previous section. As before, the fits are truncated. This time fits are truncated at 1.4 km, since this is approximately the width of the campus, and measurements further apart than that are unlikely (or erroneous). Because nugget variance was modeled explicitly with clustered measurements, the nugget tolerance is set to 0. Table 6.2 and provides the fitted variogram parameters. In this data set, the best fits are generally truncated, but without null samples at locations where a measurement was not observed. The best fits are split fairly evenly between Gaussian and cubic models.

As described in section 5.4.5, to determine the goodness of these fits a 10-fold cross validation is done

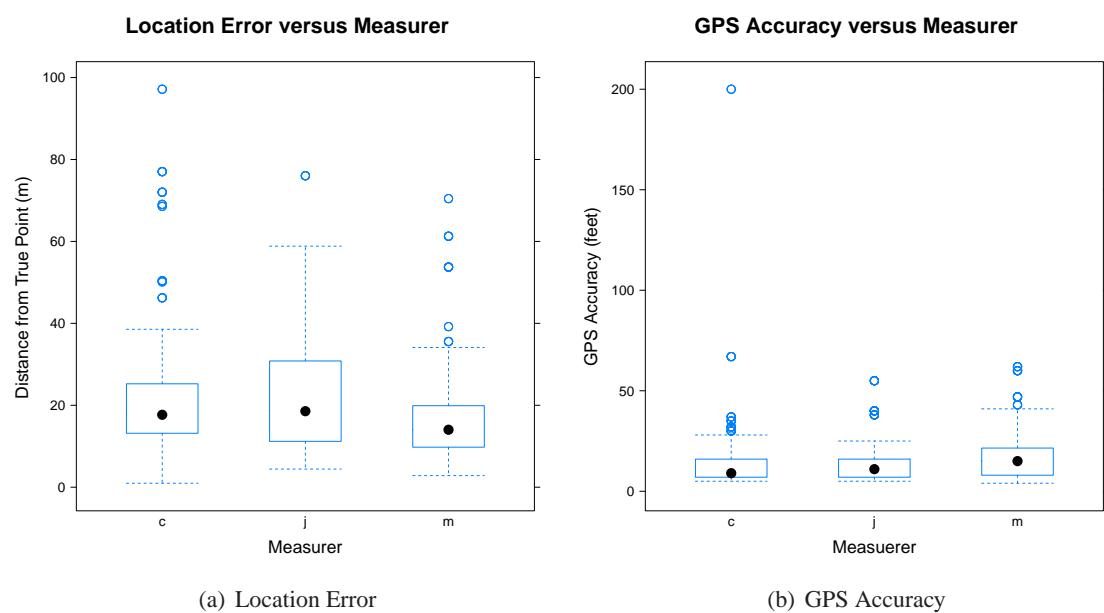


Figure 6.7: Measurement location error and GPS accuracy as a function of the measurerer.

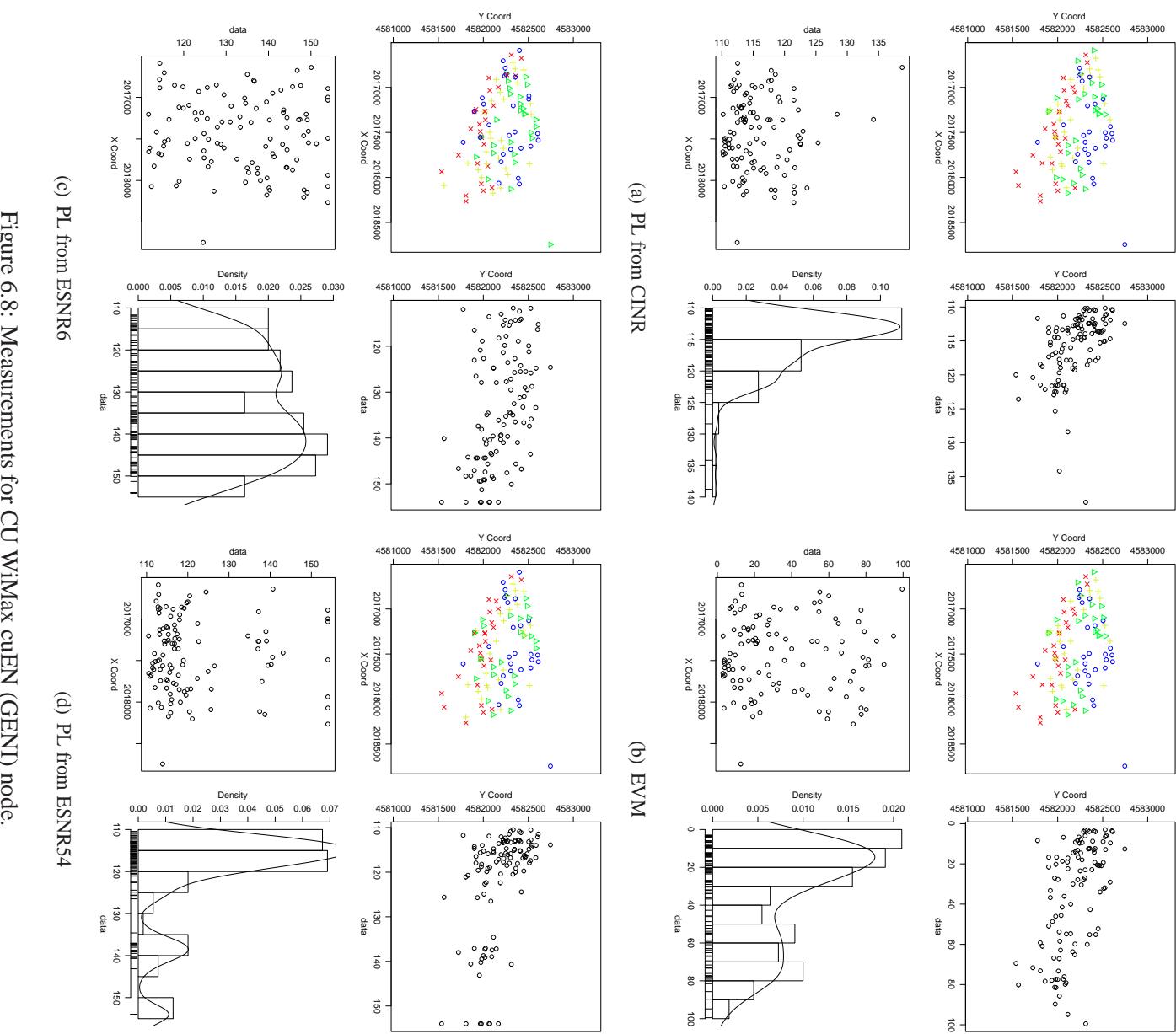


Figure 6.8: Measurements for CU WiMax cuEN (GENI) node.

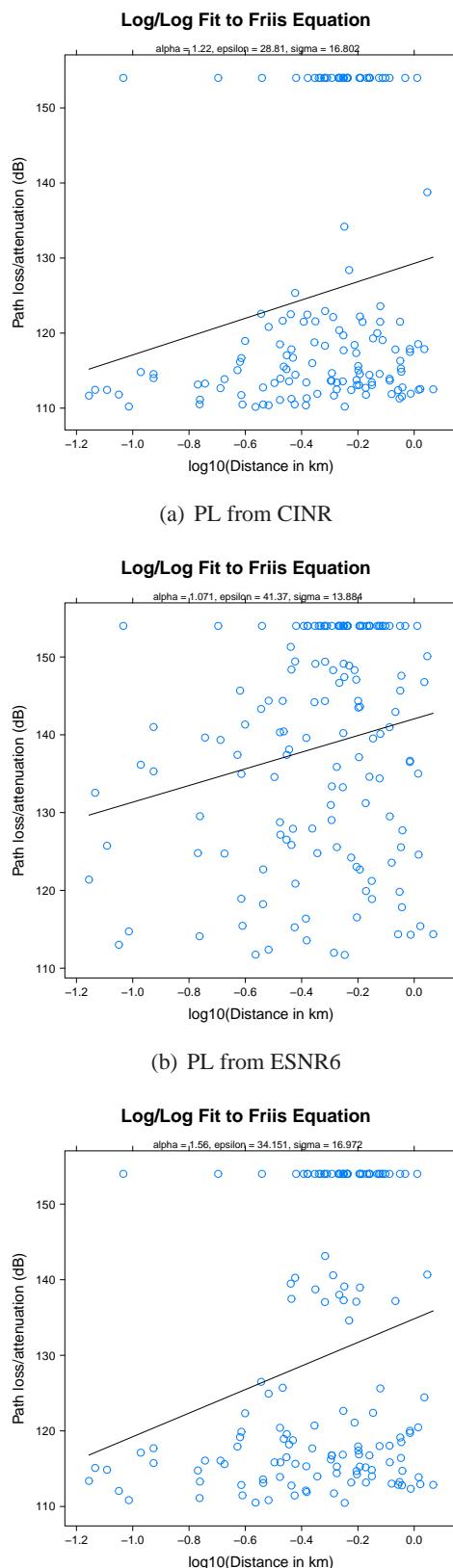


Figure 6.9: De-trending fits for the CU WiMax cuEN (GENI) node. Only the metrics that can be converted to path loss and de-trended (i.e., SNR and equivalents) are shown.

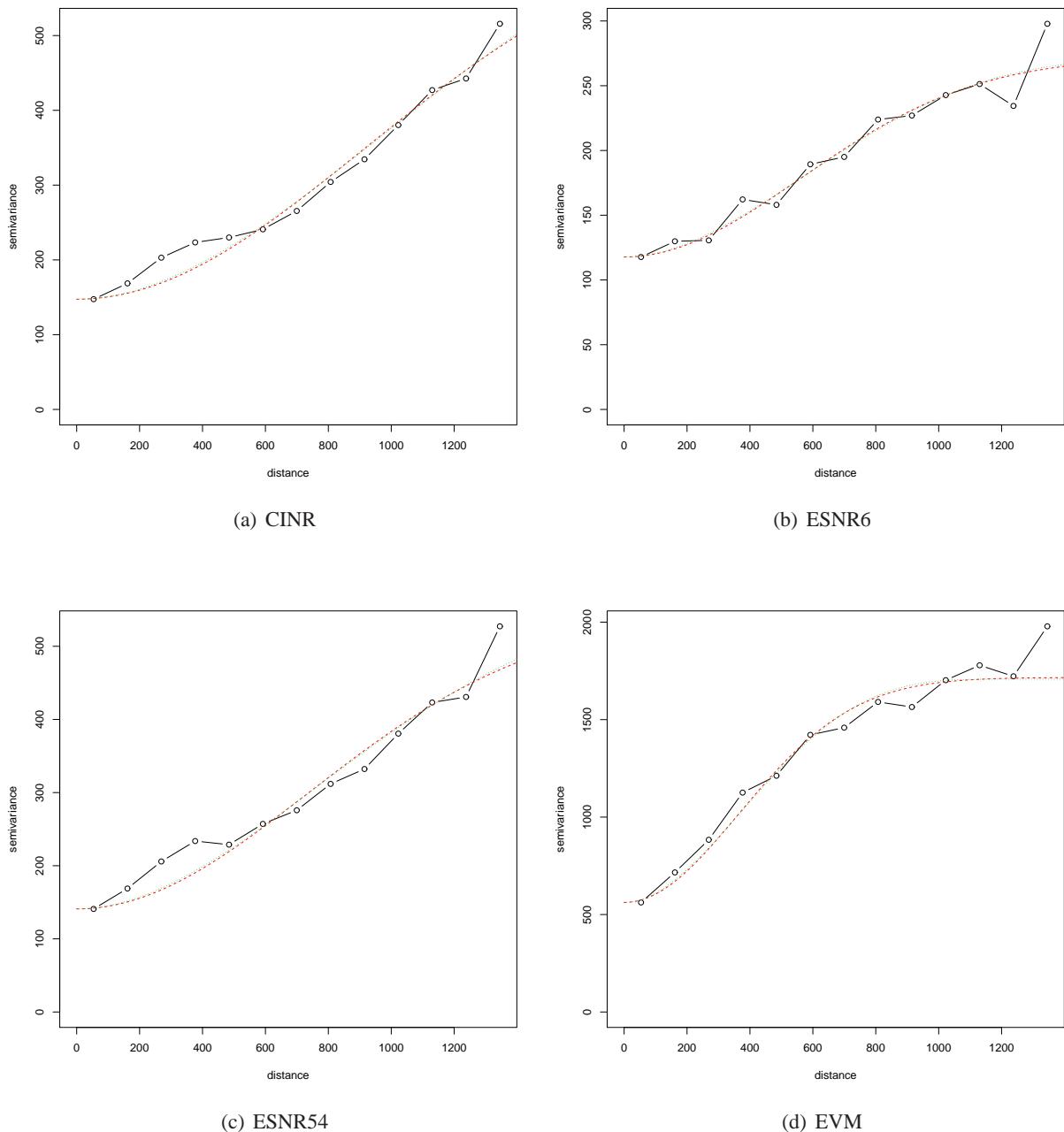


Figure 6.10: Empirical variogram and fits of four metrics for CU WiMax cuEN (GENI) node.

by random-exclusion of 10% of points, or a maximum of 50. This results in two metrics of predictive performance: RMSE (error of the predictions) and square-root of Kriging variance (residual error of the model). These metrics are averaged across the 10 folds. Overall, the fits for the CINR are quite good, achieving a RMSE of 2.03 for the cuGE BS, 4.09 for cuEN, and approximately 2.8 for the other two BS. This error is excellent, even by the standard of typical repeated-measures variance for outdoor urban environments [198]. The predictions for the ESNR metrics are less strong, ranging from 5 to 11 dB, which is still well under the typical performance of the *a priori* models described in chapter 3. Why ESNR fits less cleanly is not clear, although one explanation may be that the additional degrees of freedom it considers causes it to involve more intrinsic variability. The EVM metric, which was not de-trended, fits fairly poorly. Given that EVM can be computed from CINR directly, a better approach to mapping EVM is probably to map CINR and then translate the resulting map. The mean Kriging variance, which describes residual error in the model instead of predictive accuracy, tends to mirror the other metric. A final metric of improvement is given in the right-most column: the gain (reduction in error) as compared to the residual error of an explicit log/log fit to the measurements. The geostatistical fitting method produces a reduction in residual error for all metrics and all BSs.

6.1.5 Mapping with Ordinary Kriging

Figure 6.11 shows the resulting map for the cuEN node, which is a 120-degree sector propagating to the North. As might be expected, it appears to cover the 120-degree region to the North quite well and less well behind. While there is a clear difference between the predictions of the various metrics, the ESNR metrics are clearly predicting a less uniform propagation. The excess maps (before the de-trended values were added back), may actually provide the most useful information, since they show the deviation from a standard power-law relationship. There appears to be a very large positive excess directly behind the transmitter. This makes sense, as there is a large concrete wall preventing line of sight in that direction. As will be discussed in chapter 8, this is an area of particular interest. Interestingly, the ESNR maps push this high-point of excess further southeast, towards an area of campus shadowed by a hill, in the back-lobe of the antenna. The variance maps demonstrate the region of confidence for the map (the center red region)

BS	Metric	Model	ϕ	τ^2	σ^2	N	Trunc/Neg	Mean K-Var	Mean RMSE	Gain
cuEW	100wimaxEVM	gaussian	697.13	199.12	351.34	150	FALSE/FALSE	15.00	16.05	N/A
	100wimax	cubic	1839.69	3.99	19.38	150	FALSE/FALSE	2.16	2.75	17.54
cuEE/cuGW	100wimaxESNR54	cubic	2183.29	115.19	81.75	147	TRUE/FALSE	11.09	9.27	5.88
	100wimaxESNR6	cubic	1253.62	91.27	45.66	147	FALSE/FALSE	9.95	9.50	2.68
	100wimaxEVM	cubic	771.36	259.17	396.46	147	TRUE/FALSE	17.68	15.91	N/A
	100wimax	gaussian	541.94	8.48	9.30	147	TRUE/FALSE	3.04	2.87	12.65
cuGE	ESNR54	gaussian	2340.33	34.11	437.08	182	TRUE/FALSE	5.93	6.91	7.15
	ESNR6	gaussian	380.27	49.67	39.18	182	TRUE/FALSE	7.34	7.48	2.99
	EVM	gaussian	310.67	138.75	321.18	182	TRUE/FALSE	12.67	12.25	N/A
	CINR	cubic	1711.76	6.39	12.31	182	FALSE/FALSE	2.61	2.03	9.80
cuEN	ESNR54	cubic	1530.11	72.81	108.83	146	TRUE/FALSE	9.00	9.83	7.97
	ESNR6	gaussian	746.71	118.74	76.04	146	TRUE/FALSE	11.22	11.21	2.67
	EVM	cubic	751.21	444.98	357.14	146	FALSE/FALSE	22.84	21.11	N/A
	CINR	cubic	1304.05	14.22	20.04	146	TRUE/FALSE	4.00	4.09	12.80

Table 6.2: Best fit statistics for variogram fitting of CU WiMax BSs.

with confidence fading towards the borders where less information is available.

As another example, figure 6.12 shows the final maps for the cuGE node. This BS is a 90-degree sector pointing east, and as a result the propagation seems to favor that direction; however, there are clear and significant shadows to the west. This stands to reason, as this transmitter is high on a tower, so the region due west is not only in the back-lobe of the antenna, but may also be in the region under the half-power point (i.e., under the line of sight level with down-tilt) even for the back-lobe. Figure 6.13 shows a simplified view of the coverage of these BSs, where a threshold has been set at $CINR = 40$, based on the measurements described in section 6.1.2. From this simplified picture, it is clear to see the region of campus covered by each BS. This picture also highlights the relative complexity of the contours in the underlying maps, which are sometimes hard to discern in a colormap, even using both hue and luminance.

As a final metric of performance for these maps, each map is compared to a random sample of measurements taken around campus to see how well the maps are able to predict points *in between* the sample grid. For this experiment, 100 random sample locations were chosen and tested sequentially. Measurements were only made of the cuEN node for this test. When comparing those measurements to the predictions for the cuEN node, using the CINR metric alone, there is a RMSE of 4.71, slightly higher than that found with cross validation, but still quite good overall. If we look to the ability to predict “holes”, as was the goal of Robinson’s work in [200], the map achieves an accuracy of 69%, involving five false positives, and eight false negatives, out of 42 total test points. Since this map was produced using 110 unique measurements in a 2.54 km^2 area, it required 43.3 measurements per square km. As one point of comparison, at the same sampling density, Robinson’s iterative heuristic refinement proposal achieved approximately 80% accuracy. However, there are important semantic differences between the model presented here, which computes a smooth prediction of metric coverage over the entire region, and Robinson’s proposal, which is designed to find threshold boundaries in streets, using drive-test measurements. Given this, achieving a prediction accuracy only 10% less than Robinson’s proposal, while maintaining all of the other benefits of the geostatistical method, seems like an encouraging result.

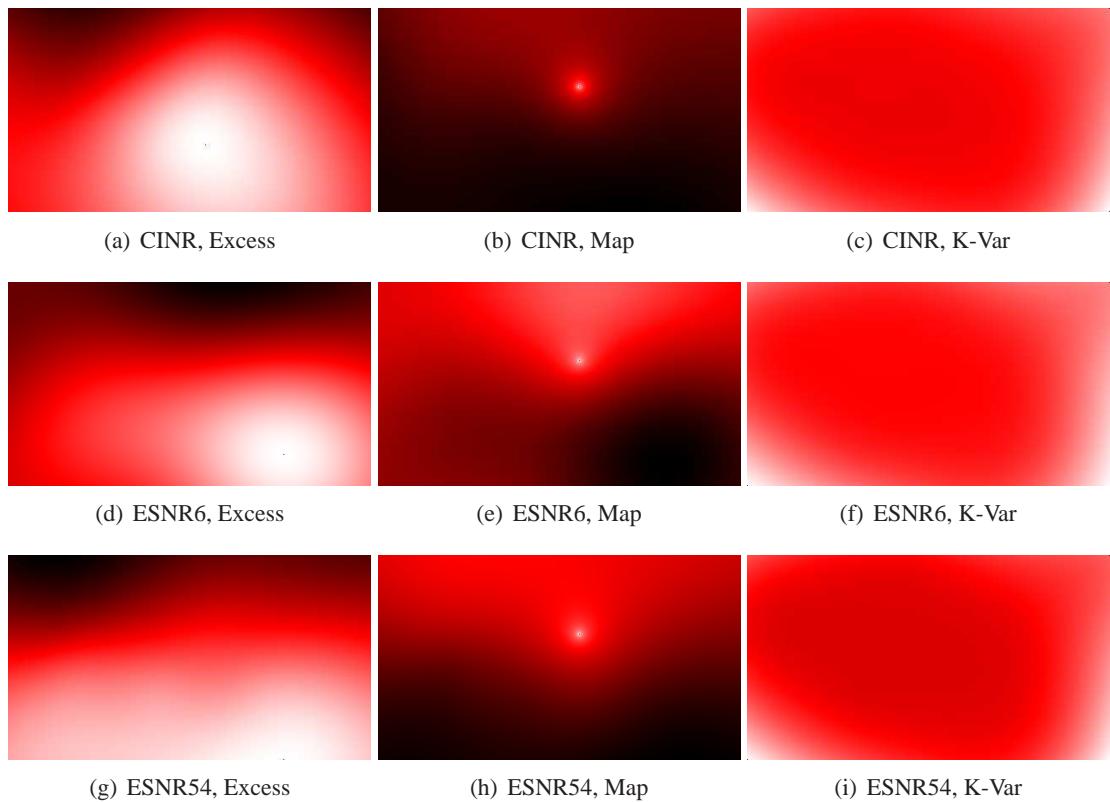


Figure 6.11: Maps for cuEN node. The left maps show the excess (residual after trend is removed). The center maps show the re-trended signal map. The right maps show the residual kriging variance of the other maps.

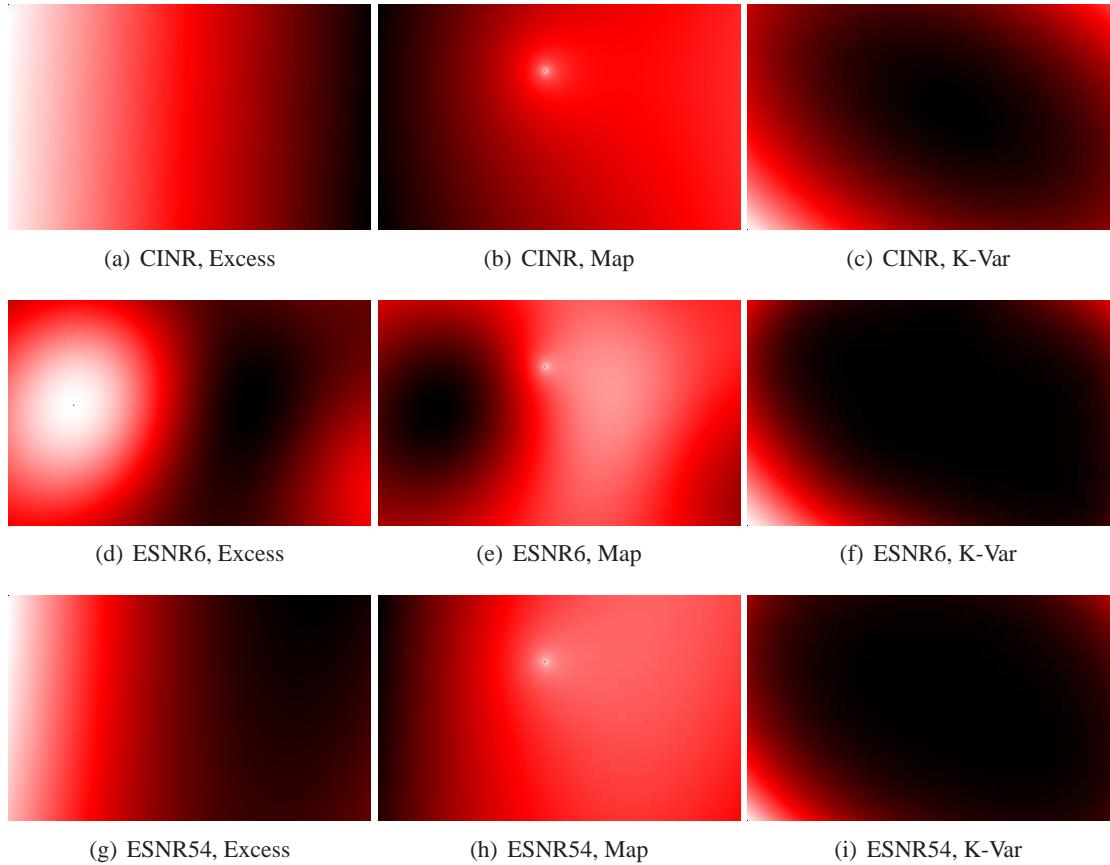
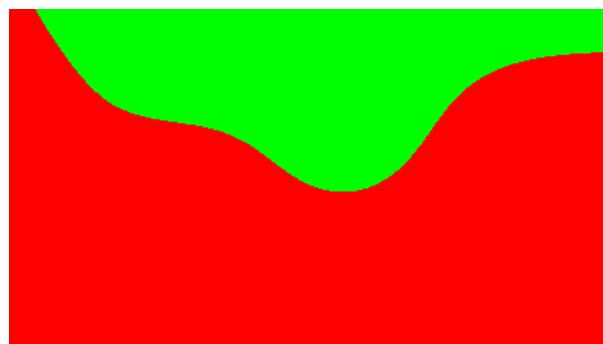
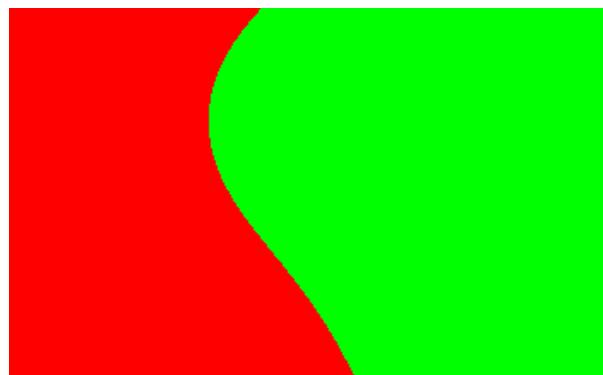


Figure 6.12: Maps for cuGE node. The left maps show the excess (residual after trend is removed). The center maps show the re-trended signal map. The right maps show the residual kriging variance of the other maps.



(a) cuEN



(b) cuGE

Figure 6.13: Binary coverage maps with the threshold of CINR=30

6.2 Map Combining

The previous section discussed how to generate per-BS maps of coverage using geostatistical techniques, guided by the example of the CU WiMax network. This section seeks to understand how measurements from multiple BSs can be combined to produce a composite map.

6.2.1 Data Combining

For a network that contains measurements from many BSs, a natural question is how a composite map can be created for the entire network. This can be accomplished either by fitting and Kriging the entire set of measurements or by fitting and Kriging measurements from each BS separately and then combining the resulting maps. The first approach is the most conceptually straightforward, but has some problems. Combining measurements from multiple APs may produce a map with a large amount of per-location variation, possibly with co-located points of drastically varying value. Exactly co-located measurements of differing value can produce unsolvable Kriging equations and must be “jittered” to create a solvable equation with a unique solution. In the end, this approach can result in a map that is difficult to interpret and has a large error margin. Consider figure 6.14 and 6.15, which show the fitted variograms and resulting maps for all the CU WiMax data combined, adding a jitter of up to 20 wavelengths to co-located measurements. The fitted variogram is relatively flat and has a large nugget variance at 250, implying an intrinsic variability at co-located points of more than 15 dB (the square root of the nugget variance). This is due to the fact that co-located points may be collected from different transmitters, and as such may have wildly differing values. However, this is not to say that the resulting map cannot be useful in spite of this basic error. Cross validation of this fit provides a mean RMSE of 3.72, a mean standard error (Kriging variance) of 3.79 and a gain of 17.64 dB over the residual error of a straight line fit through the measurements.

6.2.2 Map Combining

As an alternative to the data-combing method, Kruged maps can be combined *ex post facto*. Combining multiple Kruged maps that involve predictions in overlapping or partially-overlapping spatial regions

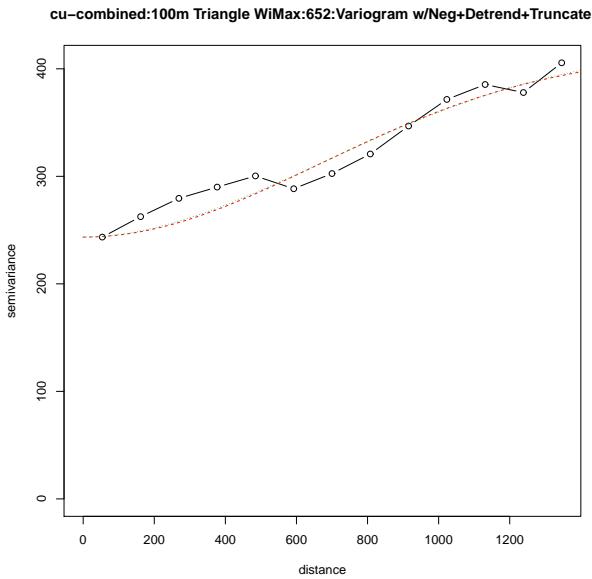


Figure 6.14: Empirical variogram and fits of four metrics for the combined CU WiMax measurements.

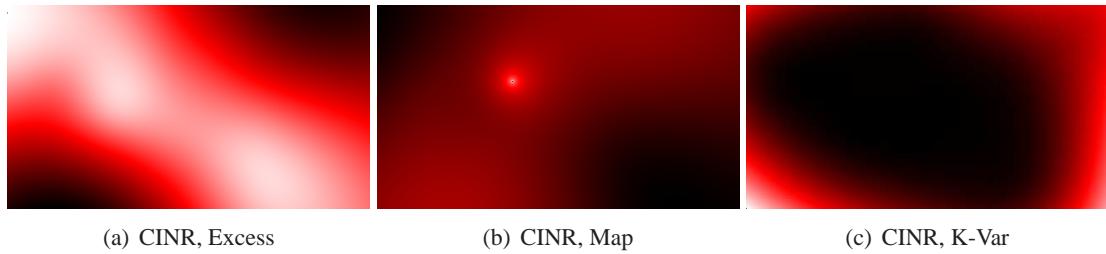


Figure 6.15: Kriged maps for combined CU WiMax measurements using the CINR metric.

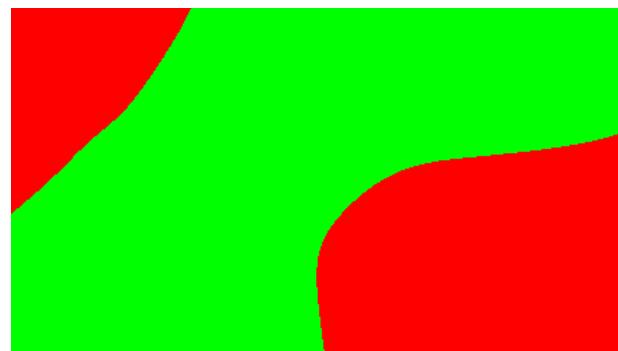


Figure 6.16: Binary coverage map for the combined CU WiMax measurements.

involves basic geospatial image tiling and combination. A basic two-pass method that first reads in all the map files to determine the total extent of the image, and then overlays the images, combining values at pixels as necessary, is used here. Algorithm 3 outlines this algorithm, where the result of computation is a matrix named “final” and the function “combine” is responsible for converting a vector of up to N values (some of which may be NULL). There are many maps that can be combined this way, the most obvious is to take the maximum value for SNR-like metrics or the minimum value for EVM-like metrics. In threshold-based combining, the number of transmitters whose interpolated signal is above 40 dB CINR (or below 60% in the case of the EVM metrics) is mapped.

Figure 6.17 shows the map-combined maps for the CU WiMax measurements. As compared to the data-combined maps, in these maps the strong signal associated with each of the two BS sites is clearly displayed with the Engineering Center Office Tower (ECOT) (right) tower showing a stronger signal since it houses three BSs instead of two. These maps show a more complex landscape of combined path-loss than the data-combined maps and bring to light some interesting observations. Chiefly, there appears to be two predominant shadows present, one due west of the Gamow (left) tower and one to the northwest of both towers. Although the source of these two shadows is unclear, and indeed it may not be possible to know for certain, it seems likely that the shadow due west of the Gamow tower is due to a misconfiguration of the downtilt on the cuGW node, since it should be covering this direction. However, since it is placed on top of an eight-story tower, the downtilt may be insufficient to cover the region directly west, below the antenna. Because this network is used to provide access primarily to buses, and one of the bus routes passes through this shadow on Colorado street, this may actually be a problem which deserves some attention. Figure 6.18 shows the threshold-based combining map where the gradient from red to green corresponds to the number of BSs providing at least a signal of 40 dB for CINR and 20 dB for ESNR. These maps differ substantially, and as a result are somewhat difficult to interpret out of context to the underlying environment. However, even at a high level these maps communicate areas lacking coverage (bright red), and areas of possible co- and cross-channel interference (bright green) where all four transmitters have strong signal.

Algorithm 3 Partial Overlay Map Tiling

```

1:  $N \leftarrow$  number of maps we are combining
2:  $minx \leftarrow$  minimum x coordinate of all maps
3:  $miny \leftarrow$  minimum y coordinate of all maps
4:  $maxx \leftarrow$  maximum x coordinate of all maps
5:  $maxy \leftarrow$  maximum y coordinate of all maps
6:  $res \leftarrow$  resolution of all maps in pixels per meter
7:  $width \leftarrow \lceil (maxx - minx) * res \rceil$ 
8:  $height \leftarrow \lceil (maxy - miny) * res \rceil$ 
9:  $mat \leftarrow$  a NULL-filled matrix of dimensions  $width \times height \times N$ 
10:  $n \leftarrow 0$ 
11: for each map  $m$  do
12:    $top \leftarrow$  topmost y coordinate of this map
13:    $left \leftarrow$  leftmost x coordinate of this map
14:    $xshift \leftarrow \lceil |top - maxy| * res \rceil$ 
15:    $yshift \leftarrow \lceil |left - minx| * res \rceil$ 
16:    $row \leftarrow height - xshift$                                  $\triangleright$  Loop over rows from top of image to the bottom
17:   for each row  $r$  in  $m$  do
18:      $row \leftarrow row - 1$                                       $\triangleright$  Loop over columns from left to right
19:      $col \leftarrow xshift$ 
20:     for each column  $c$  in  $r$  do
21:        $col \leftarrow col + 1$ 
22:        $mat[col, row, n] \leftarrow$  value at  $m[c, r]$ 
23:     end for
24:   end for
25:    $n \leftarrow n + 1$ 
26: end for
27:  $final \leftarrow$  a zero-filled matrix of dimensions  $width \times height$ 
28: for 0 to  $width$  as  $x$  do
29:   for 0 to  $height$  as  $y$  do
30:      $final[x, y] \leftarrow combine(mat[x, y])$ 
31:   end for
32: end for

```

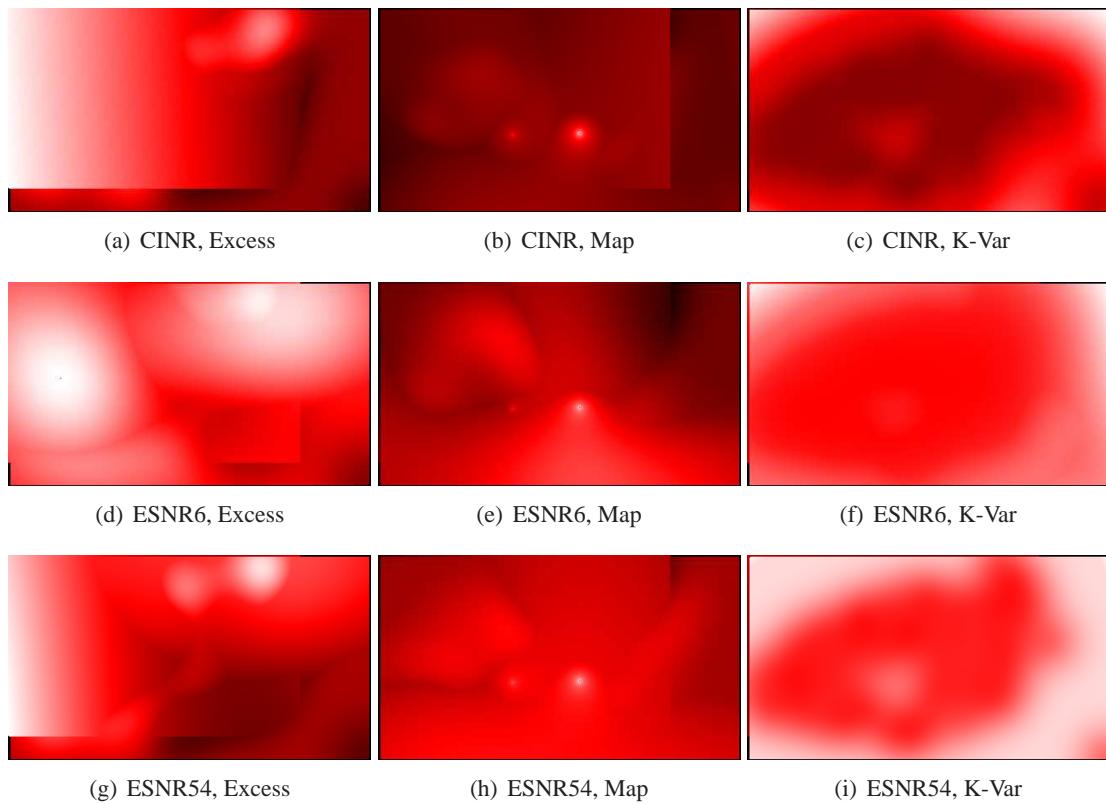


Figure 6.17: Kriged maps for combined CU WiMax measurements using the CINR metric.

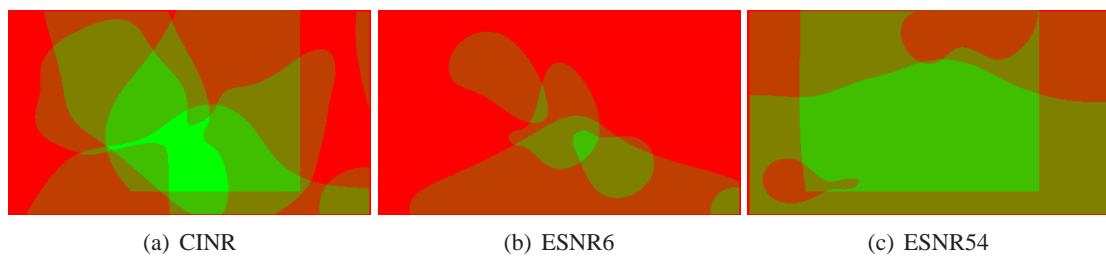


Figure 6.18: Boolean threshold-based maps for map-combined CU WiMax measurements. The CINR map uses a threshold of 40 dB and the ESNR maps use a threshold of 20 dB.

6.3 Case Study: Verizon LTE Coverage on the University of Colorado Campus

In this second case-study, the abilities of the geostatistical coverage mapping methodology are applied to a network utilizing a similar technology, however at a very different frequency: 700 MHz. In particular, the aim is to map the coverage of the Verizon Wireless Third Generation Partnership Project (3GPP) LTE network over the extent of the CU campus. LTE is a Fourth Generation (4G) compatible protocol that is also backwards-compatible with earlier Third Generation (3G) networks. LTE is favored by many to be the target of the next major cell technology roll out in the United States, and as such, is an especially interesting technology to study [186]. Verizon has deployed an early prototype LTE network operating around 700 MHz in central Boulder, Colorado, which allows the fairly unique analysis presented here. Table 6.3 shows the location and specifications of the LTE BSs [240].

6.3.1 Measurement Apparatus and Procedure

For the sake of consistency and comparability, measurements are made as before on a triangular lattice with a constant lag of 100 m. In addition, in this measurement campaign more care is paid to making systematic cluster measurements. Although the general measurement goals remain the same as in the previous study, the measurement apparatus and measurement procedure require some substantial adjustment for the LTE data collection. Figures 6.19 and 6.20 show the design of the LTE measurement apparatus. At the core of the measurement rig is a laptop computer running Windows XP and the JDS Uniphase (JDSU) E6474A network analysis and drive test software [53] connected to an LG VL600 LTE USB dongle. The JDSU software interfaces with the device to connect to the Verizon LTE network and collect real-time data about network performance.

As before, a hardcopy map is used to locate measurement points on a best-effort basis. Unlike the WiMax measurements, in this measurement campaign active measurements of upstream and downstream throughput are collected at each point as well. To accomplish this, the program iperf is used to perform a 60 second bidirectional Transmit Control Protocol (TCP) throughput test. The general measurement procedure is given in algorithm 4.

Site	Sector	Latitude	Longitude	Site Name	Azimuth	PCI Grp ID	PCI ID	OTA PCI	N Obs.
85	1	40.017486	-105.252212	Central Boulder	10	72	0	216	0
85	2	40.017486	-105.252212	Central Boulder	95	72	1	217	0
85	3	40.017486	-105.252212	Central Boulder	180	72	2	218	2
119	1	39.997778	-105.2615	CU Campus	335	78	0	234	42
119	2	39.997778	-105.2615	CU Campus	70	78	1	235	12
119	3	39.997778	-105.2615	CU Campus	180	78	2	236	0
138	1	40.004656	-105.260597	Buffs	290	123	0	369	60
138	3	40.004656	-105.260597	Buffs	205	123	2	371	55
294	1	39.984708	-105.233044	Table Mesa	340	80	0	240	0
294	2	39.984708	-105.233044	Table Mesa	110	80	1	241	0
294	3	39.984708	-105.233044	Table Mesa	240	80	2	242	0
391	1	40.007092	-105.276575	The Hill	350	75	0	225	22
391	2	40.007092	-105.276575	The Hill	90	75	1	226	30
391	3	40.007092	-105.276575	The Hill	220	75	2	227	0
392	1	40.016128	-105.259997	Walnut	0	73	0	219	0
392	2	40.016128	-105.259997	Walnut	180	73	1	220	27
392	3	40.016128	-105.259997	Walnut	270	73	2	221	25
492	1	40.018028	-105.277768	DT Boulder	0	74	0	222	0
492	2	40.018028	-105.277768	DT Boulder	160	74	1	223	33
492	3	40.018028	-105.277768	DT Boulder	260	74	2	224	0
644	1	39.995314	-105.233431	Baseline n Foot	0	71	0	213	0
644	2	39.995314	-105.233431	Baseline n Foot	90	71	1	214	0
644	3	39.995314	-105.233431	Baseline n Foot	270	71	2	215	0
650	1	40.007925	-105.268236	Whizzer	270	76	0	228	39
650	2	40.007925	-105.268236	Whizzer	90	76	1	229	31
650	3	40.007925	-105.268236	Whizzer	180	76	2	230	37
652	1	40.037164	-105.246139	Ara	270	59	0	177	0
652	2	40.037164	-105.246139	Ara	55	59	1	178	0
652	3	40.037164	-105.246139	Ara	180	59	2	179	0
694	1	40.008431	-105.2577	Bison	0	77	0	231	2
694	2	40.008431	-105.2577	Bison	120	77	1	232	1
694	3	40.008431	-105.2577	Bison	235	77	2	233	39

Table 6.3: Specification and location of Verizon LTE cell basestations around Boulder, Colorado. The column N Obs. provides a count of the number of times each BS was observed in our data collection.

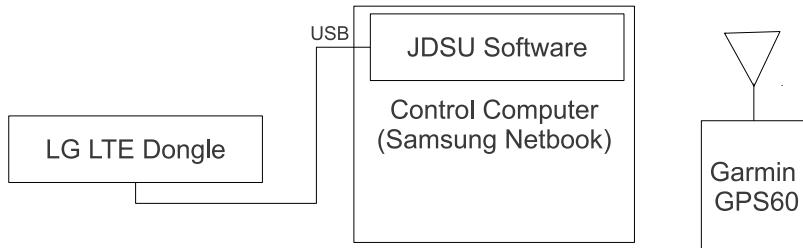


Figure 6.19: Diagram showing connectivity and specification of LTE measurement cart.

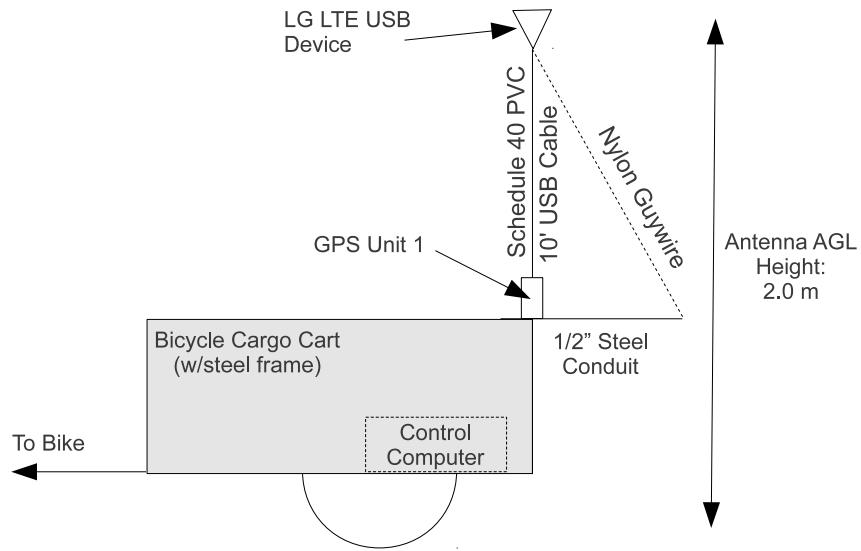


Figure 6.20: Diagram of LTE measurement cart.

Algorithm 4 Point testing procedure for LTE measurement

- 1: Locate a point to measure using the map. Try to get as close to the actual point as possible. If it is on/in a building or other obstacle, go to the nearest accessible outdoor location and do the measurement there instead. If the point is totally unreachable (i.e. closest point is 100m+ away), then circle the point and move on.
 - 2: Write down the ID of the point being measured (i.e., 12 4 0) for the 0th measurement at point 12 4.
 - 3: Using the hand held GPS, write down the current location and accuracy of the GPS position.
 - 4: Using the computers clock (or a synchronized watch), write down the current time.
 - 5: Click the green circle in the JDSU software to start a test.
 - 6: After 5 seconds, a command window will open behind the main window and will run a throughput test. After 60 more seconds, this window will close.
 - 7: Click the Stop button (or press Shift + F9) to stop the test.
 - 8: Mark off the location just measured on the map with an X.
 - 9: Every third point, perform 2 additional “cluster measurements” where you select a nearby point within an 8 meter radius of the original point (20 wavelengths at 700 MHz). These points should be more or less selected at random.
 - 10: Move to the next point.
-

At each point, the JDSU software collects a continuous stream of measurements using a large number of metrics. This data is stored in a proprietary file format that can be exported to a parseable American Standard Code for Information Interchange (ASCII) Comma Separated Value (CSV) file after measurement is completed.

LTE measurement differs from WiMax measurement in an important way: the measurement device is a functional LTE radio and as such will connect to the best network available (ranking by Reference Signal Received Power (RSRP)) in order to pass traffic. A given measurement point may be served exclusively by a single cell, or by a combination of cells. Hence, some number of measurements are available for most cells that serve the CU campus, but as a result of this behavior, the strongest cells receive the most measurements. And, the resulting measurement set is more descriptive of the combined coverage of the cells, rather than a complete model of any particular BS.

6.3.2 Performance Metrics

One question this study seeks to answer is which metric is the best indicator of performance for an LTE network of this kind. To this end, all metrics that are available through the JDSU software are collected. As was done in section 6.1.2, an analysis is then performed to determine how well or poorly they predict application-layer performance. [111] describes the physical-layer metrics that can be collected by the JDSU software and the 3GPP/LTE specification describes how many of these metrics are calculated [5]. The set of metrics collected is:

- Path Loss (PL): Path Loss in dB, calculated by the User Equipment (UE) using information from the BS.
- SNR: SNR in dB.
- RSSI: Three RSS measurements are available. It is not clear how they differ, so all three have been collected.
- Throughput: Upstream and downstream throughput are collected, however only an upstream

throughput test is performed, and the downstream channel is left unloaded. Throughput values are available at different layers, but are all simply a constant offset from one another.

- Transmit Power: Transmit power used by the UE, and the BS for the Sounding Reference Signal (SRS).
- Channel Quality Indicator (CQI): Values are available for the wide-band, sub-bands, and multiple code-words. All available are collected.
- Block Error Rate (BLER): Both Physical Downlink Shared Channel (PDSCH) and overall BLER are collected.

In order to understand how well the lower-level metrics predict the ultimate application-layer performance, an Analysis of Variance (AOV) is performed using the upstream throughput as the target and the physical-layer metrics as the factors to be tested. The results show that UE transmit power contributes most to higher-layer performance, and in fact, is inversely correlated with the performance, since the UE will choose a lower power when the upstream radio channel is good, and a higher power when it appears to be lossy. This factor produces an F-value of 385.6 and a p-value $<< 0.01$. Other significant factors include the PL (F-value = 90.6, p-value $<< 0.01$), SNR (F-value = 26.4, p-value $<< 0.01$), RSSI0 (F-value = 19.0, p-value $<< 0.01$), and distance from the BS (F-value = 19.5, p-value $<< 0.01$). Figure 6.21 shows the relationship between the best performing physical-layer performance metrics and upstream throughput.

It makes sense that the transmit power has a strong effect on the uplink throughput. And it stands to reason that PL would be a significant factor if the UE transmit power is, since the PL is used directly in the calculation of the transmit power. Similarly, the RSSI, SNR, and distance are all strongly correlated with the PL. However, this poses a problem: if the chosen transmit power is derived from a formula based on PL and CQI, and the transmit power affects the uplink throughput, then how can we know whether PL or CQI are themselves good predictors of channel quality, or whether they are being misused to tune the transmit power and hence bias the results towards or away from their values? Unfortunately, it is impossible to resolve this confound without performing modifications to the LTE protocol itself, and in this study we

have no control over the operators' infrastructure. Hence, this thesis proceeds with the assumption that the calculated PL (which is indeed determined using a channel-sounding mechanism) is the best performance predictor available, if it is not possible to make higher-layer performance tests directly. This is perhaps an unintuitive result, since PL does not consider non-flat fading across subcarriers (as CQI does), but is nonetheless the conclusion supported by the results.

A final question is how well correlated the various physical-layer metrics are *with each other*. Figure 6.22 shows the correlation between path loss and four other metrics. As expected, transmit power and RSSI appear to be strongly correlated with PL. This stands to reason because RSSI is likely calculated directly from path loss, and the UE transmit power is computed explicitly using path loss. The relationship between distance (from the connected BS) and SNR are less trivial, but there is still a clear and statistically significant correlation in both cases.

As could probably be inferred, the three RSS metrics are extremely similar, having a Pearson correlation coefficient of between 0.97 (for RSSI1 versus RSSI2) and 0.99 (for RSSI1 versus RSSI0). The differences between these RSS values is dominated by a small location shift, supporting the theory that these are measurements of the same signal using different methods (or hardware). As such, use of any of the three RSS values produced by the measurement software should be sufficient to model RSS. In fact, any two can be used to predict the third using a simple linear mixture model:

$$RSSI_0 = 0.38RSSI_1 + 0.62RSSI_2 + 7.13 \quad (6.1)$$

This model has a residual standard error of 1.5 dB, an R^2 of 0.98, and an F-statistic of 11,720. Similarly, the PL metric and the RSSI metrics are closely related:

$$RSSI_0 = 41.38 - 0.96PL \quad (6.2)$$

which implies that the EIRP of the BSs is approximately 41 dB. This fit has a residual standard error of 1.66, an R^2 of 0.98 and a F-statistic of 19,430 indicating an exceptionally tight fit.

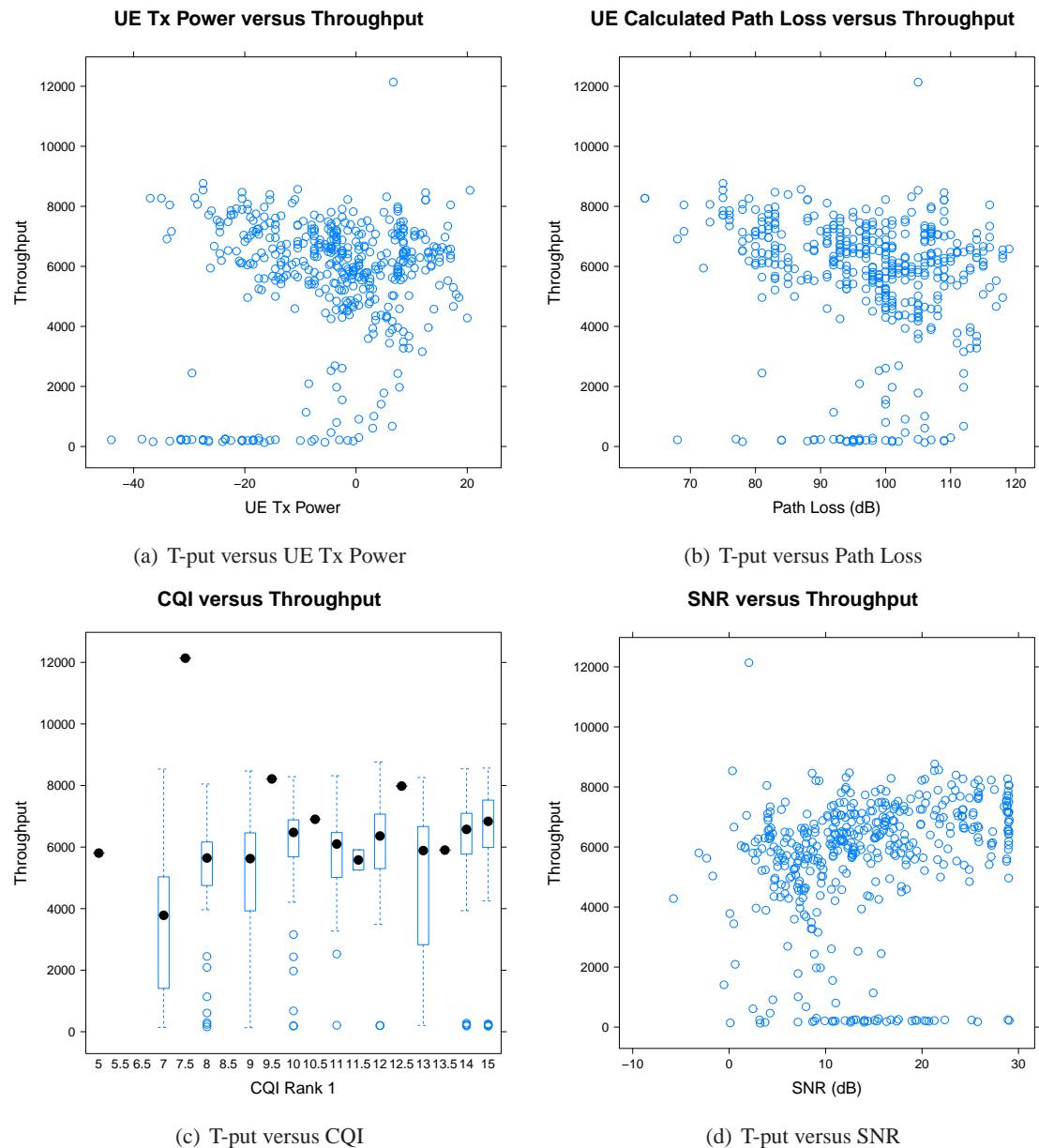


Figure 6.21: Correlation between upstream throughput performance and physical-layer metrics.

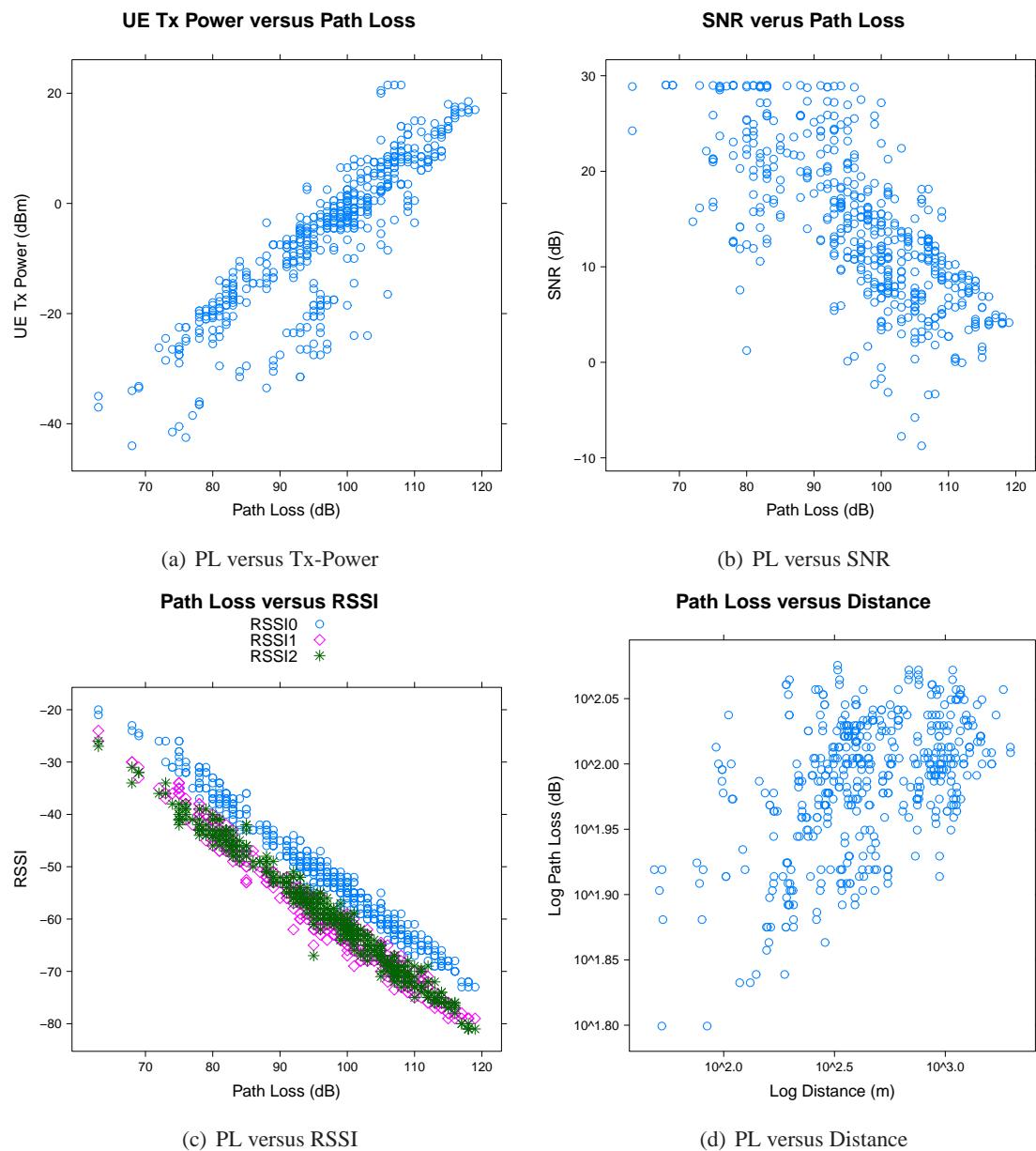


Figure 6.22: Correlation between physical-layer metrics.

6.3.3 Spatial Data Characterization and Variogram Fitting

In order to allow for comparability and to produce a more easily solveable Kriging matrix, the throughput measurements are normalized as such:

$$T_{kbps} = (T_{kbps} - \min(T_{kbps})) / (\max(T_{kbps}) - \min(T_{kbps})) \quad (6.3)$$

The result of this transformation is that the throughput value at any given point is actually the fraction of total observed throughput seen at that point.

Figures 6.24 and 6.25 show the LTE measurements for all BSs and for the most observed BS, 369, for each of the seven metrics. The combined measurements and the measurements for 369 appear to have a similar distributional shape and to some extent the measurements for 369 appear representative of the measurements as a whole. Figures 6.26 and 6.27 show the variogram fits for this data, and figure 6.23 shows the log/log fit of path loss versus distance. The fits for the combined data are much better, presumably due to the uneven sampling of the single-BS data due to the way BSs are selected for measurement by the measurement apparatus. This further confirms that for these LTE measurements, the combined measurement map is most meaningful. Tables 6.4 and 6.4 provide the fitted parameters for the combined data and a subset of interesting BSs.

6.3.4 Mapping with Ordinary Kriging

Ordinary Kriging proceeds as before, using the variogram models from section 6.3.3. For each combination of models and maps, an interpolated map is produced by Kriging the value at each pixel. Figures 6.28 and 6.29 show the final Kriged maps for the seven metrics for both BS 369 and all measurements combined. In the case of the maps in 6.28, the PL map shows a clear location and direction of transmission for the BS. The remaining physical-layer metrics provide a consistent picture of the abilities of this particular node, whose range of functional coverage is actually fairly limited. The application-layer (i.e., throughput) metrics are clear outliers. As can be seen in the plot of measurements in figure 6.25, upstream and downstream throughput tend to largely disagree—at spots where upstream throughput is at its peak, downstream

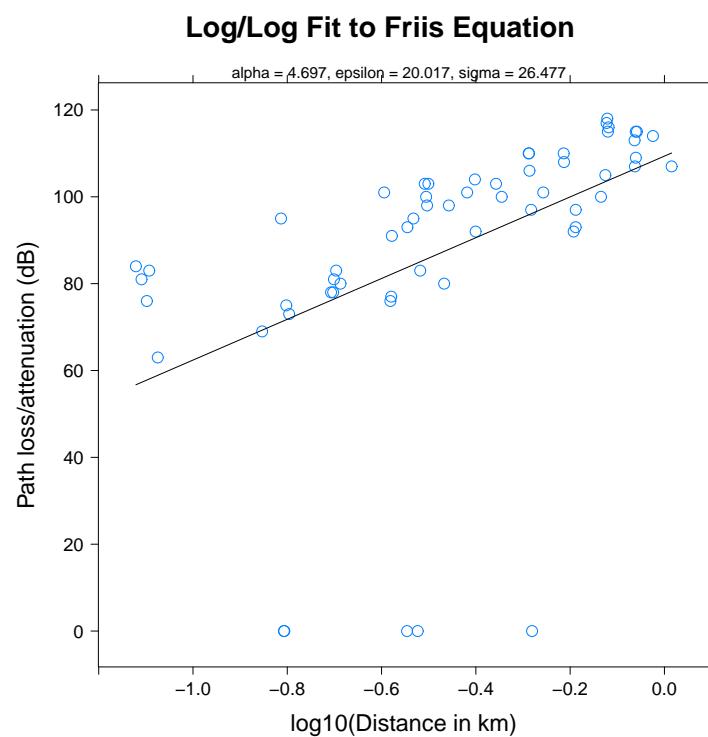


Figure 6.23: Log/log fit of path loss for LTE BS 369.

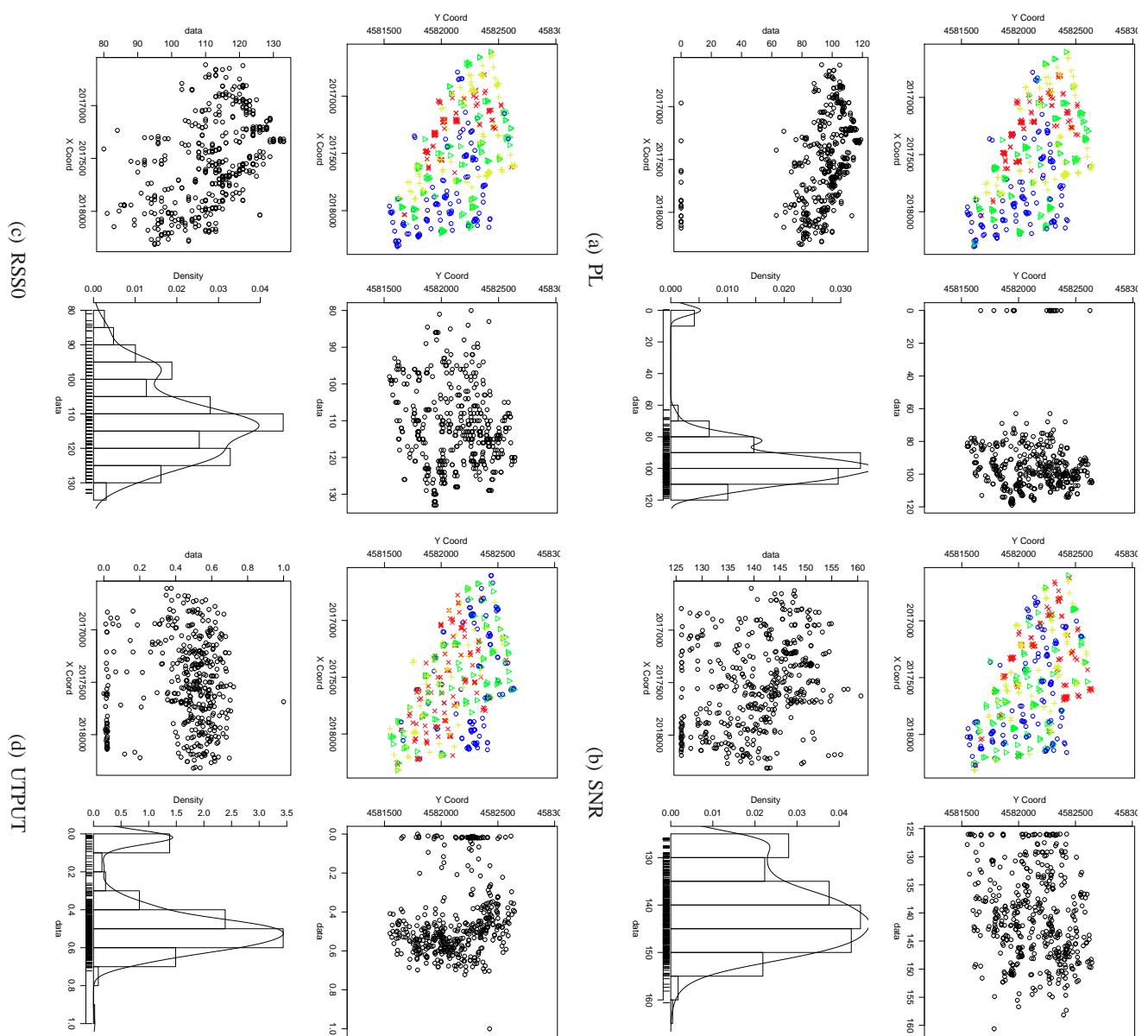


Figure 6.24: LTE Measurements.

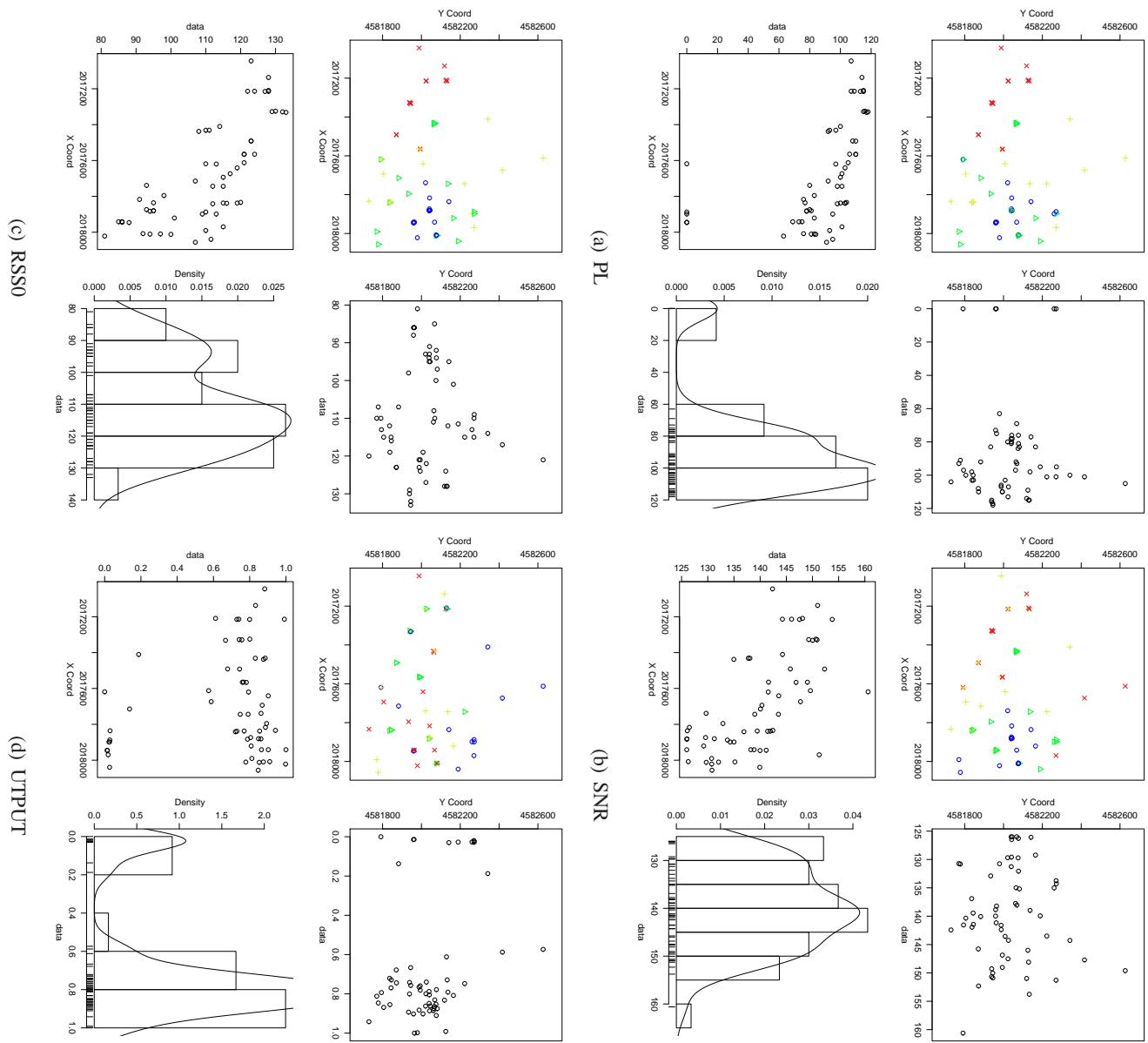


Figure 6.25: Measurements for LTE BS 369.

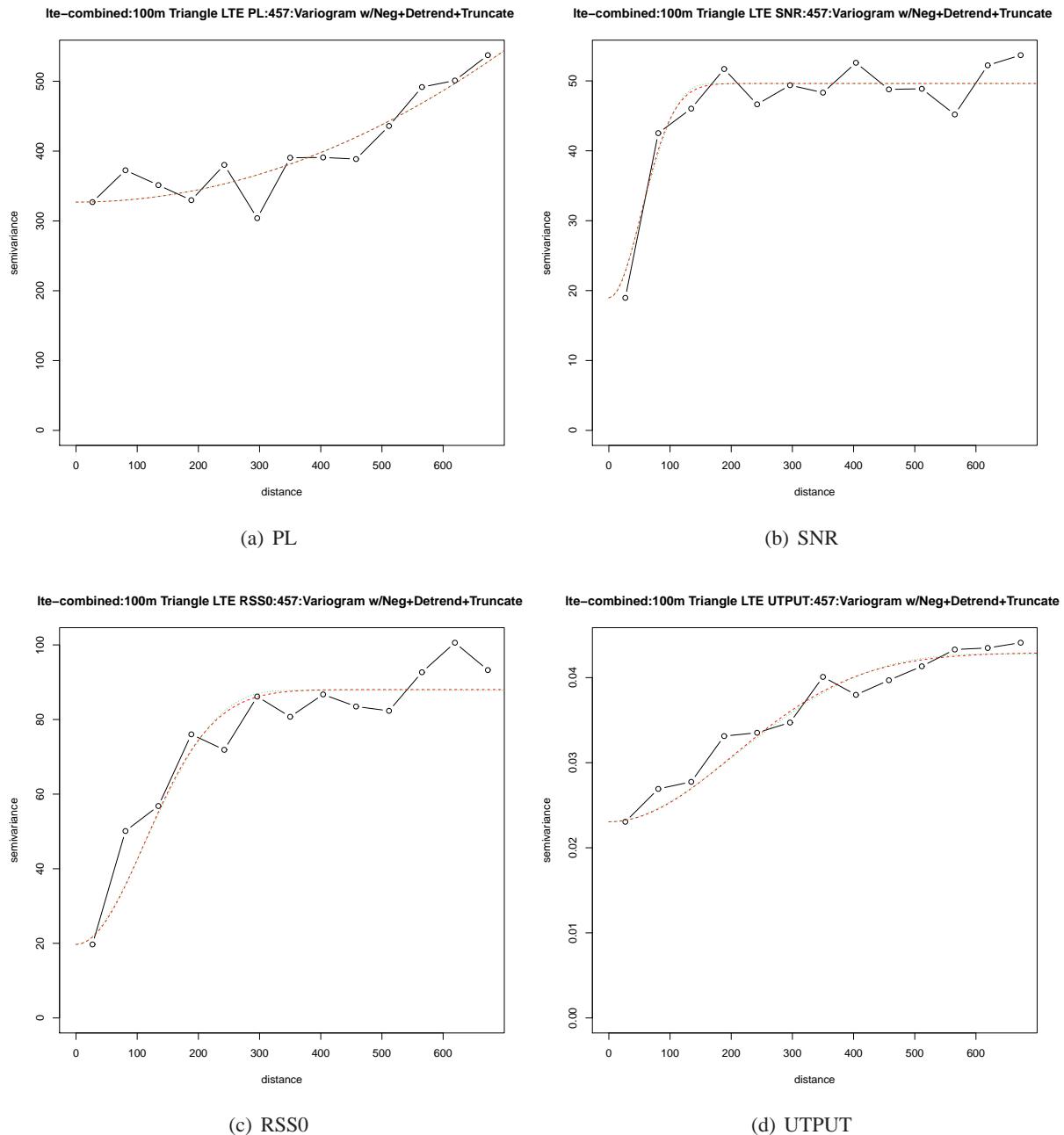


Figure 6.26: Variogram fits for all measurements combined.

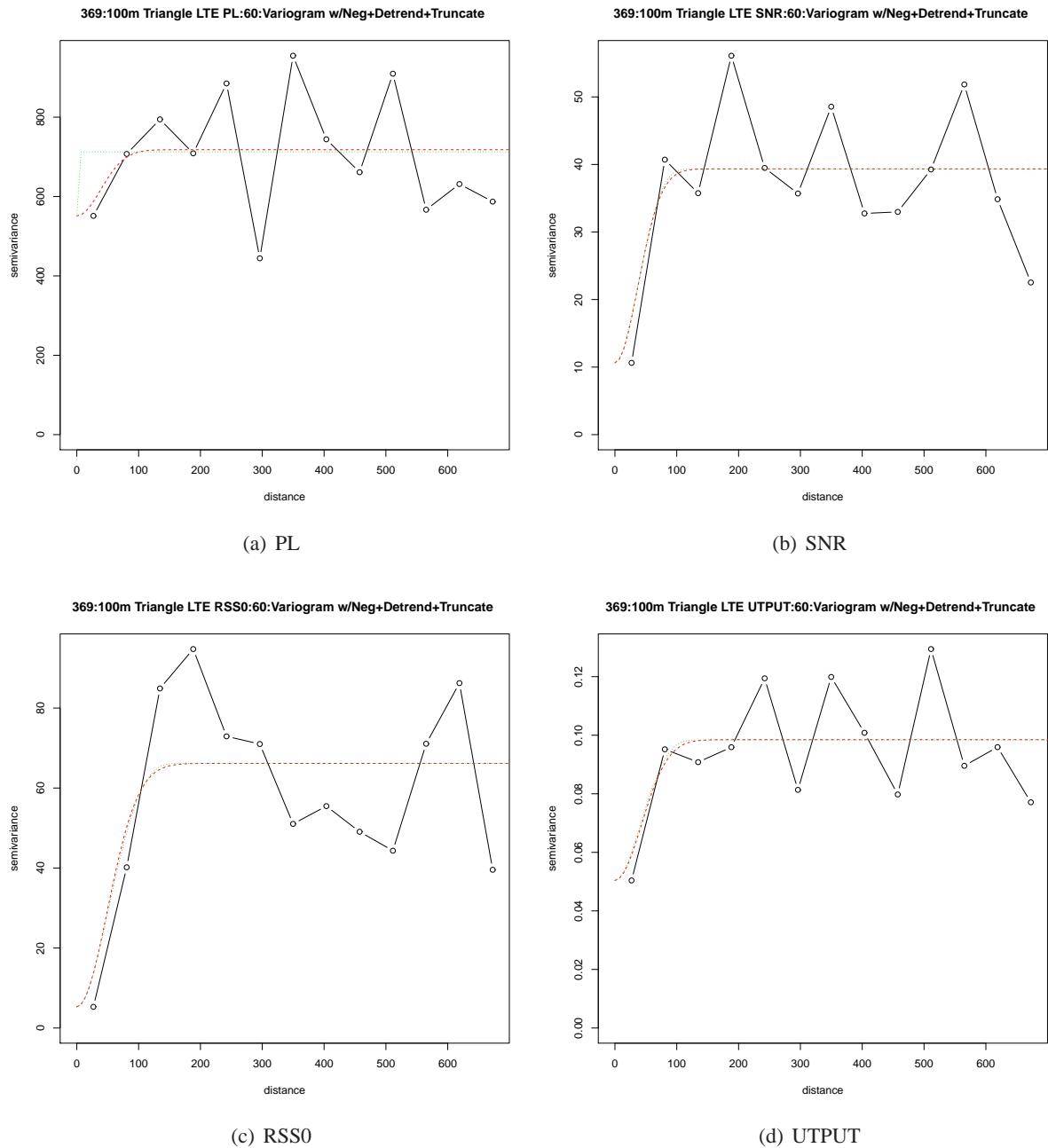


Figure 6.27: Variogram fits for BS 369.

AP Mac	Dataset	Model	ϕ	τ^2	σ^2	N	Trunc/Neg	Mean K-Var	Mean RMSE	Gain
220	100lteRSS0	cubic	421.14	14.35	16.87	27	FALSE/FALSE	4.58	4.05	0.97
220	100lteSNR	cubic	190.48	18.52	13.98	27	FALSE/FALSE	5.34	5.18	0.43
220	100lteUTPUT	Gaussian	176.34	0.06	0.04	27	TRUE/TRUE	0.29	0.25	-0.29
221	100ltePL	cubic	127.49	10.89	24.83	25	FALSE/FALSE	4.60	4.62	1.36
221	100lteRSS0	Gaussian	61.98	8.86	18.57	25	TRUE/TRUE	4.17	3.86	1.06
221	100lteSNR	Gaussian	0.00	10.61	18.60	25	FALSE/TRUE	5.28	3.89	0.24
223	100lteUTPUT	cubic	129.50	0.02	0.03	33	TRUE/FALSE	0.19	0.19	-0.19
225	100ltePL	cubic	860.53	2.55	9.67	22	TRUE/TRUE	1.84	2.02	0.53
225	100lteRSS0	Gaussian	194.24	1.82	4.21	22	TRUE/FALSE	1.64	1.71	0.48
225	100lteSNR	Gaussian	219.27	9.89	16.55	22	TRUE/FALSE	3.46	3.65	0.96
226	100lteRSS0	cubic	167.10	5.27	6.94	30	TRUE/TRUE	3.01	1.93	0.65
226	100lteSNR	cubic	323.82	10.61	46.58	30	FALSE/FALSE	4.64	3.65	2.52
226	100lteUTPUT	Gaussian	107.85	0.04	0.06	30	FALSE/TRUE	0.26	0.24	-0.26
228	100ltePL	cubic	106.64	9.85	85.51	39	TRUE/TRUE	6.89	6.01	2.72
228	100lteRSS0	cubic	105.90	8.12	80.15	39	TRUE/FALSE	6.08	6.78	3.18
228	100lteSNR	cubic	174.24	17.83	66.97	39	TRUE/FALSE	6.86	5.00	2.05
228	100lteUTPUT	Gaussian	191.37	0.02	0.05	39	FALSE/FALSE	0.17	0.19	-0.17
229	100ltePL	cubic	136.71	2.50	88.24	31	TRUE/TRUE	6.04	4.13	3.37
229	100lteRSS0	Gaussian	54.56	2.00	84.81	31	FALSE/TRUE	6.19	4.58	3.02
229	100lteSNR	cubic	146.18	3.02	43.50	31	FALSE/TRUE	4.58	3.29	2.16
229	100lteUTPUT	Gaussian	55.35	0.01	0.03	31	TRUE/TRUE	0.13	0.12	-0.13
230	100ltePL	cubic	96234.59	14.57	2743822.04	37	TRUE/TRUE	4.07	12.90	15.56
230	100lteRSS0	Gaussian	97.17	10.02	99.68	37	TRUE/TRUE	6.42	5.19	3.73
230	100lteSNR	Gaussian	74.13	11.24	31.12	37	TRUE/FALSE	5.32	5.07	1.10
230	100lteUTPUT	Gaussian	40389.84	0.04	312.68	37	FALSE/FALSE	0.20	0.19	-0.20
233	100ltePL	cubic	42.48	1.00	1014.75	39	FALSE/TRUE	23.36	19.92	8.88
233	100lteRSS0	cubic	118.82	2.50	45.15	39	TRUE/FALSE	4.34	4.82	2.46
233	100lteSNR	Gaussian	47.05	4.06	30.42	39	TRUE/FALSE	4.36	3.62	1.34
233	100lteUTPUT	Gaussian	146.82	0.01	0.18	39	FALSE/FALSE	0.19	0.17	-0.19

Table 6.4: Table 1 of best fit statistics for variogram fitting of LTE data.

AP Mac	Dataset	Model	ϕ	τ^2	σ^2	N	Trunc/Neg	Mean K-Var	Mean RMSE	Gain
234	100ltePL	cubic cubic Gaussian	343.61	1.25	68.95	42	FALSE/TRUE	3.65	6.32	4.39
234	100lteRSS0		387.17	9.25	57.03	42	FALSE/TRUE	4.51	5.11	3.25
234	100lteSNR		36.61	15.02	30.87	42	TRUE/TRUE	5.51	5.98	1.34
235	100ltePL	Gaussian cubic cubic cubic	0.00	20.02	1.62	12	TRUE/FALSE	4.86	4.46	0.36
235	100lteRSS0		326.60	18.70	6.03	12	TRUE/FALSE	4.83	4.74	1.19
235	100lteSNR		47.07	57.78	0.00	12	TRUE/FALSE	0.00	10.61	6.22
235	100lteUTPUT		47.07	0.09	0.00	12	TRUE/FALSE	0.00	0.54	0.00
369	100ltePL	Gaussian	0.00	599.99	89.53	60	FALSE/TRUE	26.48	21.92	-0.00
369	100lteRSS0	Gaussian	69.73	5.29	60.87	60	TRUE/FALSE	5.46	3.94	2.40
369	100lteSNR	cubic	178.08	18.11	19.76	60	FALSE/TRUE	5.38	4.52	0.75
369	100lteUTPUT	cubic	188.70	0.06	0.04	60	FALSE/TRUE	0.29	0.26	-0.29
371	100ltePL	cubic	996.66	313.27	213.95	55	FALSE/TRUE	18.53	11.82	1.23
371	100lteRSS0	Gaussian	118.86	7.76	68.13	55	FALSE/FALSE	4.56	4.50	3.84
371	100lteSNR	cubic	411.27	20.09	26.95	55	FALSE/TRUE	5.05	4.57	1.50
371	100lteUTPUT	cubic	1450.56	0.05	0.06	55	FALSE/FALSE	0.23	0.24	-0.23
lte-combined	100ltePL	cubic	87330.05	326.96	486883.61	457	TRUE/TRUE	18.16	17.81	3.16
lte-combined	100lteSNR	Gaussian	74.62	18.96	30.67	457	TRUE/TRUE	5.09	4.30	2.12

Table 6.5: Table 2 of best fit statistics for variogram fitting of LTE data.

throughput may be zero. The interpolated maps reflect this: upstream throughput interpolates to a largely positive value with holes around regions where traffic could not pass, whereas downstream throughput predicts mostly a bitrate of zero, excepting those few regions where downstream tests were successful. It is likely the case that these results are a function of a problem in testing throughput further up the network, either via traffic rate limiting or outright blocking. However, we do not have enough control over the Verizon network to differentiate possible sources of error here. Instead, we think this helps to further highlight the fickle nature of higher-layer tests.

Figure 6.29 shows the map with all measurements combined and table 6.5 provides the performance results for these fits. Although the differences in the measurement procedure results in fewer co-located measurements with radically different values (i.e., only the best BSs should receive measurements), the combined mapping process is still somewhat unpredictable as the interpolation must resolve unrealistically large differences in value between nearby points. This is especially visible in the K-Var plots, which show that variance is minimal only immediately around the points and then grows rapidly moving away from each measurement. A better map might be obtained if the UE hardware were able to collect independent measurements of each BS at each point (as was done in the WiMax tests). However, this isn't possible with the hardware used here and hence, there are some necessary limitations in terms of the resulting map. Were the data collected by Verizon themselves, it would be easy to address this limitation.

finally, figures 6.30 and 6.31 provide the map-combined maps for these measurements. The former is maximum-combined, and shows a very complex landscape for each metric. The latter provides threshold-combined maps using an SNR of 20 dB (or equivalent). The right-hand side shows which points are covered by this metric, and the left-hand side shows how many separate BSs cover the region with performance at least meeting the threshold value.

6.4 Modeling Small Scale Effects and Deviation from Stationarity

This section asks the question: how do measurements vary over small time scales and small distances? An underlying assumption of the Kriging process is that the process being modeled is stationary, meaning that the (fitted) mean is constant in both time and space. Clearly, this is a strong assumption that the (often

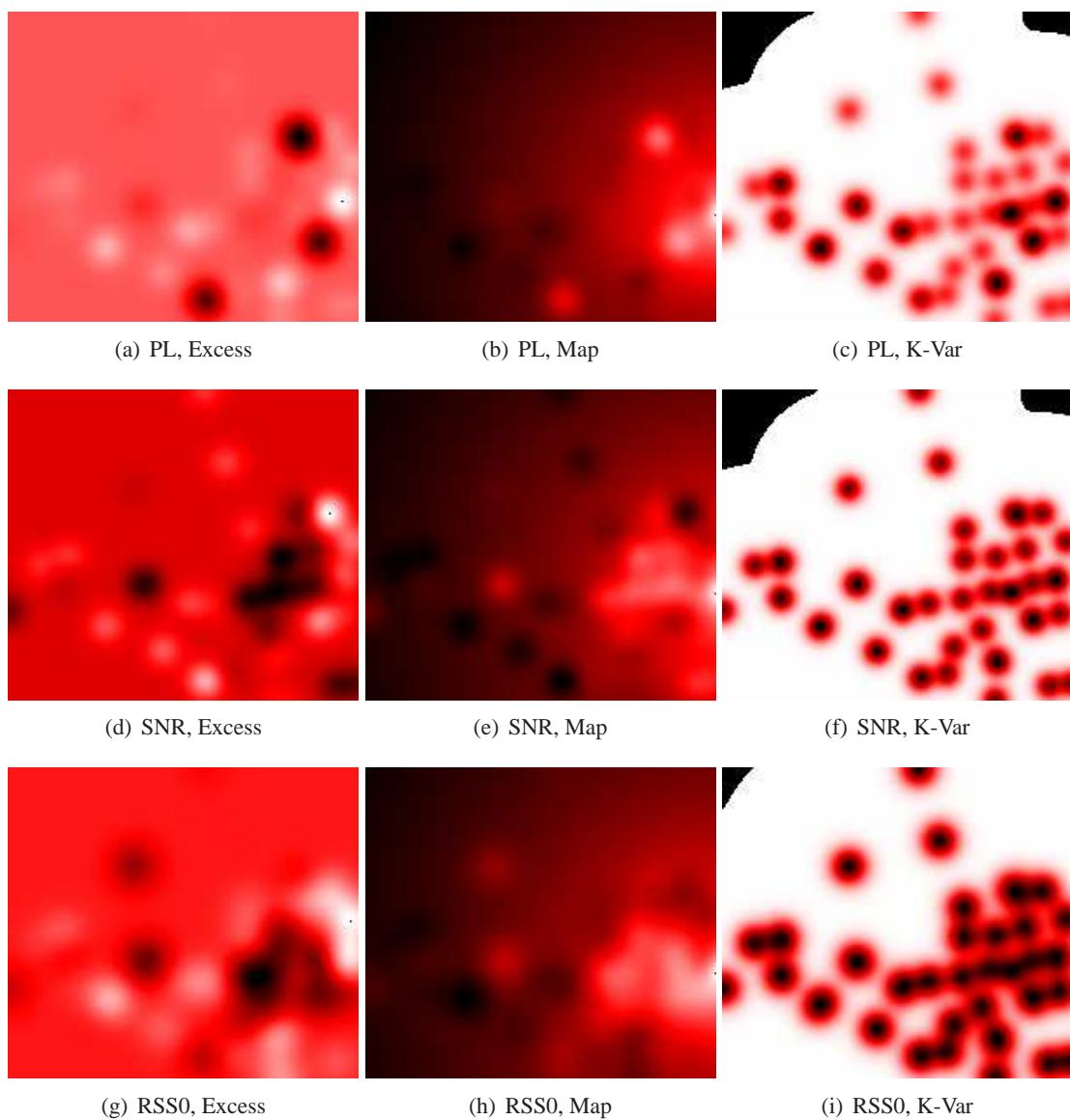


Figure 6.28: Kriged maps for BS 369.

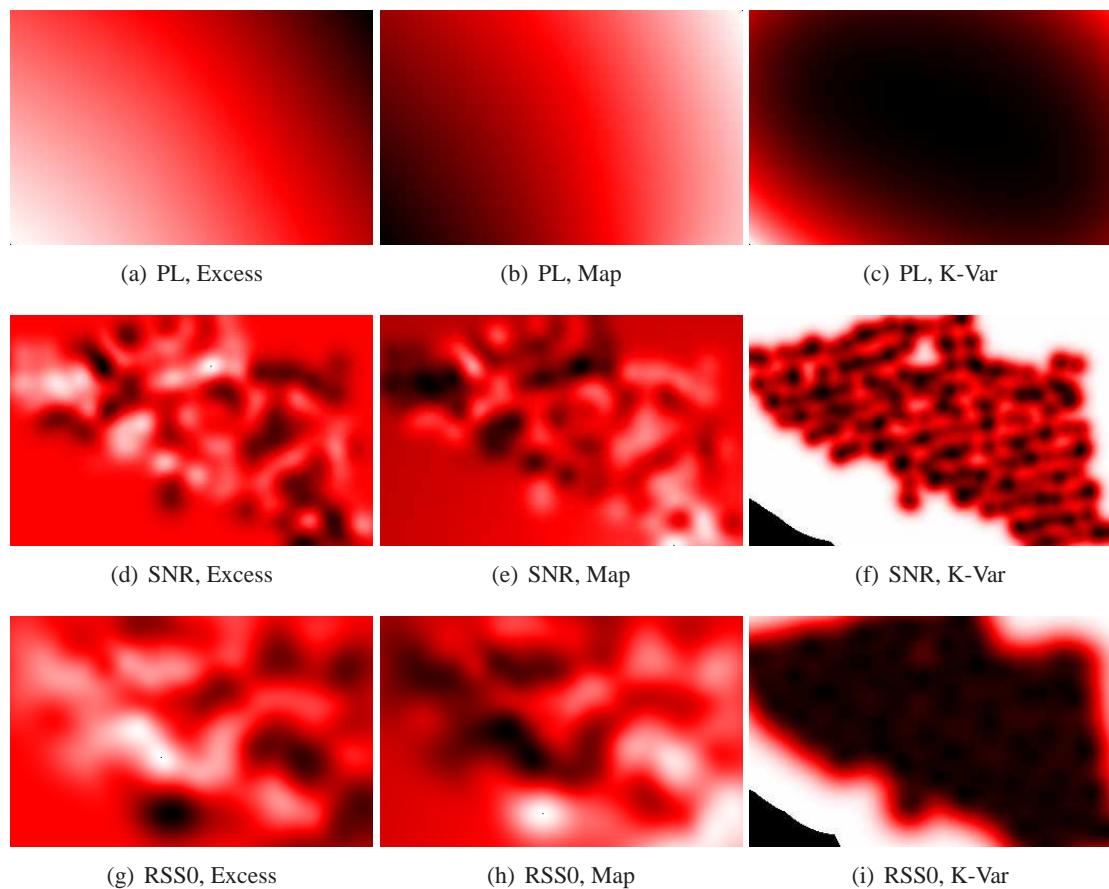


Figure 6.29: Kriged maps for all measurements combined.

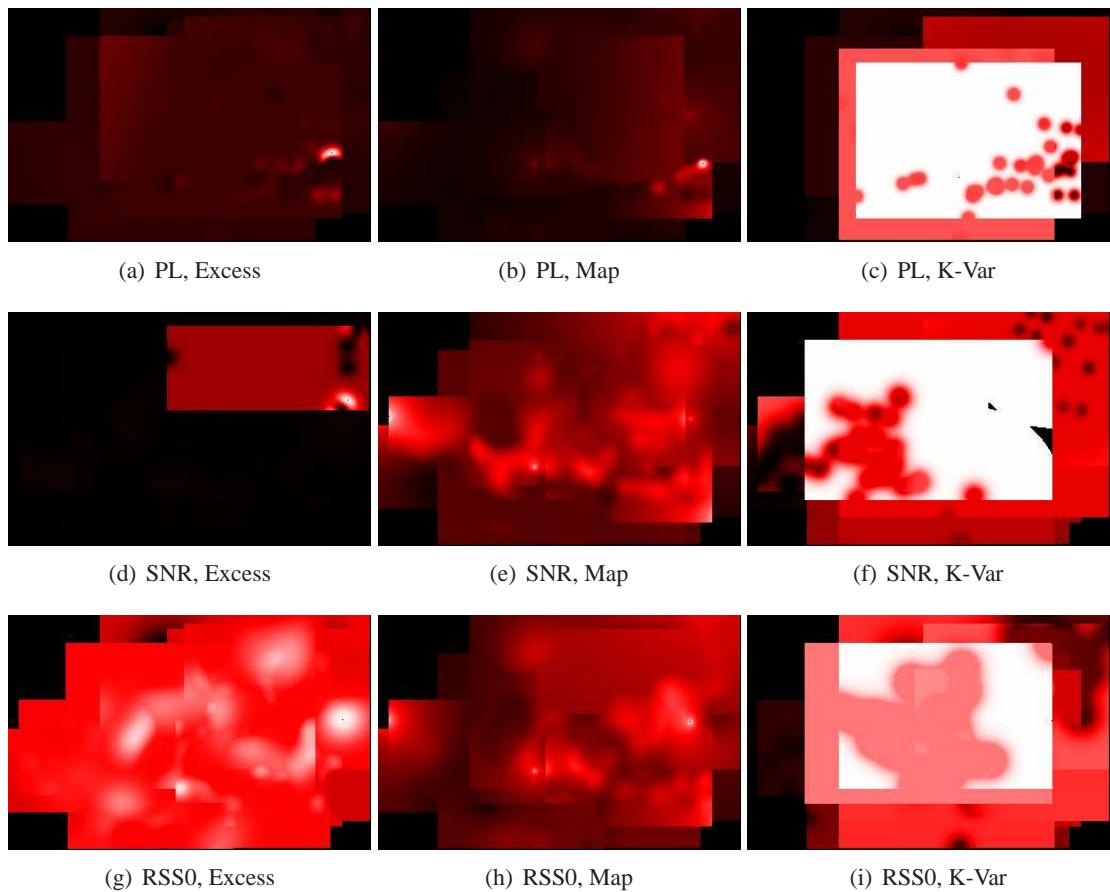


Figure 6.30: Map-combined maps for all measurements combined.

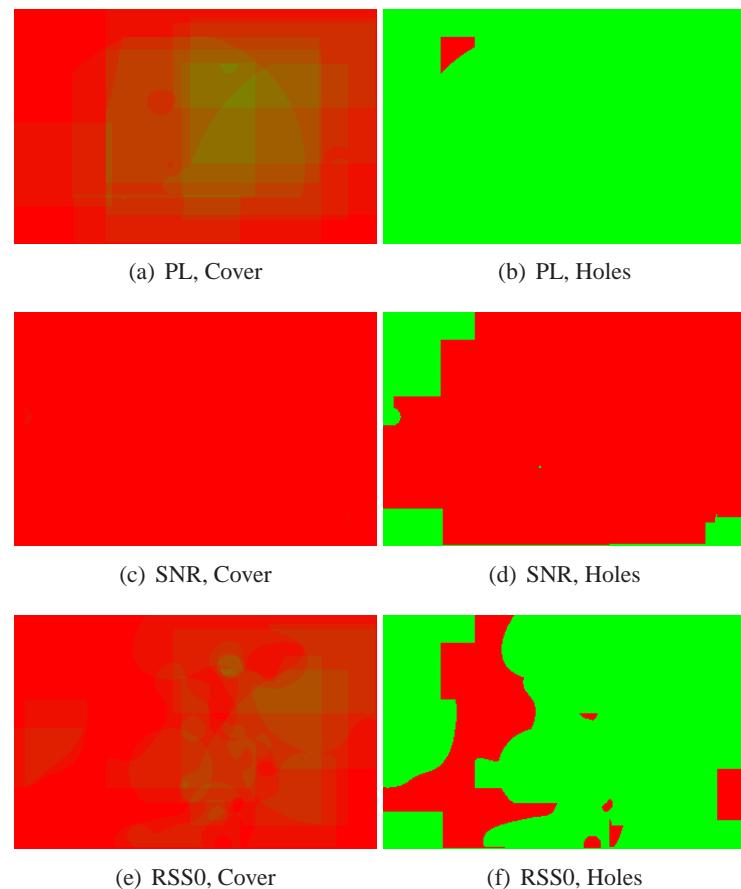
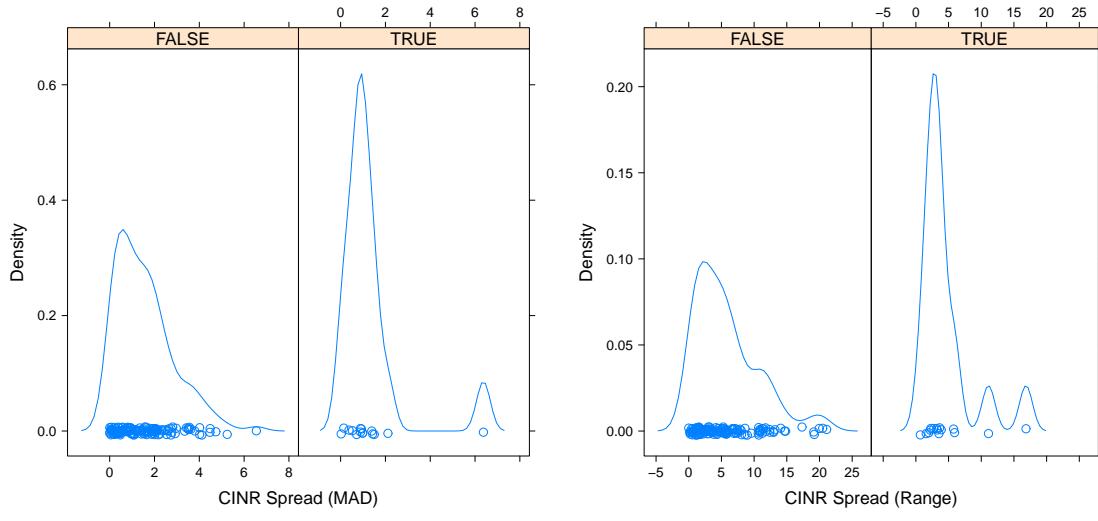
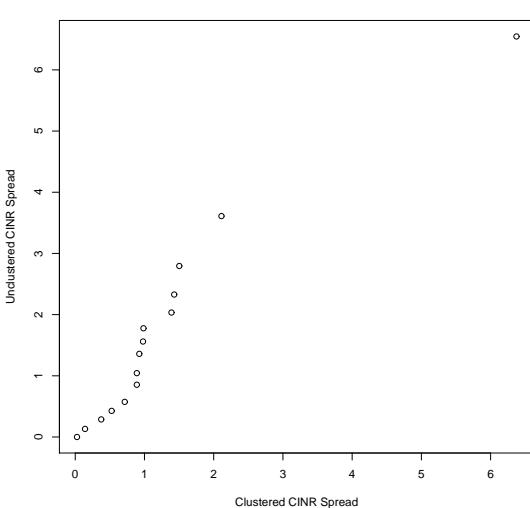


Figure 6.31: Threshold-based map-combined maps for all measurements combined.

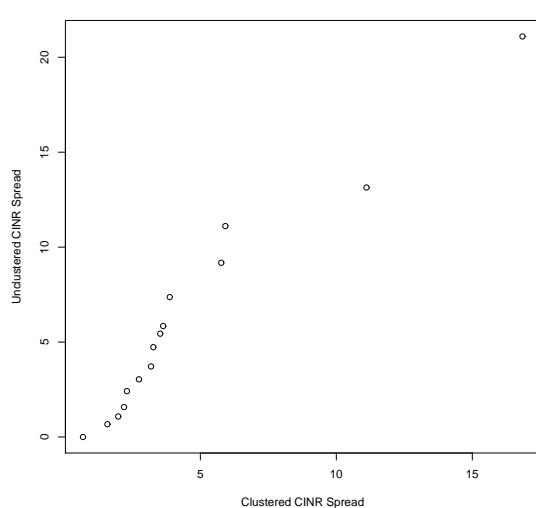
chaotic) radio environment is unlikely to obey. It is possible to loosen the stationarity assumption at the cost of substantial additional computational work, but in practice most users of Kriging processes opt to accept the implications of this assumption. By understanding how the radio environment changes over small time scales and small distances, a bound can be placed on repeated measurement variation and hence a bound on the implicit unavoidable error associated with the stationarity assumption.

As discussed in section 2.1, fading in the radio environment can be classified into small-scale and large-scale fades. Large-scale fades should be fairly constant over large distances and time, and hence are not troublesome—it is exactly the environment-specific large-scale fading effects that Kriging seeks to interpolate. However, small-scale fades can be highly varying in time and over small distances because they stem from multipath effects and (possibly mobile) scatterers. As a practical rule of thumb many experimenters average measurements within 40 wavelengths as a way to “average out” small-scale effects [132]. This section seeks to validate that standard practice as well as understand the scale of small-scale effects over short time scales.

After the initial CU WiMax measurement campaign, a second campaign was undertaken to collect data at clustered locations so that the small-scale (in space and time) variation can be compared to large-scale trends. To this end, a random subset of approximately 15 grid points were selected and at each point three complete measurements were taken at random locations within 40 wavelengths of the original grid point. Figure 6.32 shows the amount of measurement spread observed at these closely clustered locations versus the amount of measurement spread between repeated measurements at the same location along with comparative QQ-plots of the distributions. Here, two different definitions of spread are used that make no assumptions about the underlying distribution of the data: range, which is simply the total difference between the largest and smallest observation, and Median Absolute Deviation (MAD), which is the median of absolute deviations from the data’s median. Although the two distributions are not identical, they do appear to be Gaussian in both cases, with a similar central tendency and spread. Indeed, these distributions have been compared with a Welch two-sample t-test, two-sample Kolmogorov-Smirnov test, and Wilcoxon rank sum test (all of which test the null hypothesis that the difference in central tendency is significant), and none of the tests are willing to reject the null hypothesis that the data are drawn from the same distribution

CINR Spread(MAD) for Clustered/Nonclustered Measurements

QQ-Plot of Clustered/Nonclustered Spread(MAD)


(c) Range QQ-plot

QQ-Plot of Clustered/Nonclustered Spread(Range)


(d) MAD QQ-plot

Figure 6.32: Distribution of spread and comparative QQ-plots for measurements taken in within 40 wavelengths of each other (i.e., clustered) versus at the same point at different times (i.e., unclustered) for two different definitions of measurement spread (MAD and range).

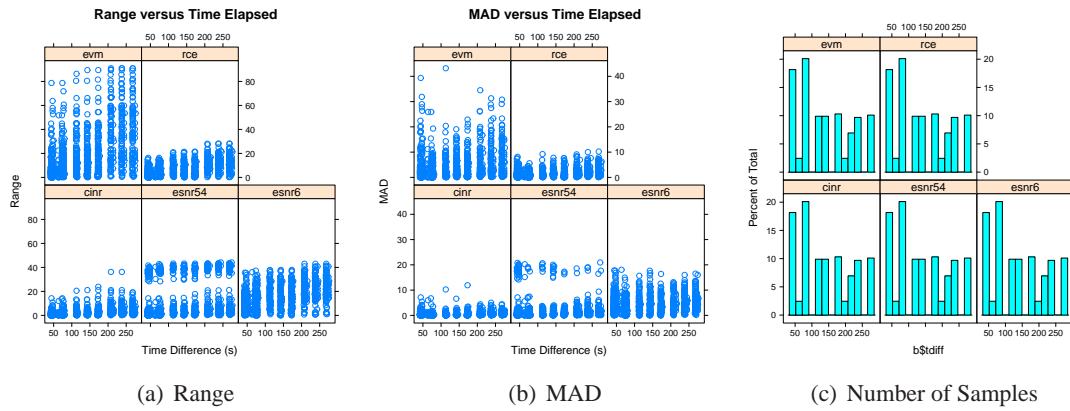


Figure 6.33: Amount of spread (variation) as a function of time elapsed between measurements using two different metrics.

(p-values are between 0.3 and 0.5). This result suggests that the amount of underlying small-scale fading in space can be sufficiently modeled by taking repeated measurements in time, at least in the environment studied here.

Figure 6.33 provides a different view of this intrinsic channel variability by plotting the amount of variation observed between repeated measurements taken at the same location as a function of time. This figure shows that the amount of variation appears to be fairly stable for all of the metrics over small time scales (several minutes). There is a slight increase in measurement spread observed for the RCE and EVM measurements, but this does not appear to be substantial, and may not be significant. Interestingly, the ESNR metrics appear to have more intrinsic variation than the simpler metrics, which may be due to the fact that these metrics take into account more degrees of freedom (i.e., independent fading on each subcarrier). Although it is likely that the radio environment is nonstationary at large time scales (days, weeks, and years), from these results it appears that the intrinsic variation is fairly stable on small time scales and hence a few repeated measurements are likely sufficient to characterize intrinsic variability.

As a further investigation, clustered measurements were systematically collected at every third sample location in the LTE data collection process. Figure 6.34 shows the distribution of spread at repeated measurements in space (right column (TRUE), measurements clustered within 40 lambda) and repeated measurements in time alone (left column (FALSE)). The time-delayed repeated measures variance for the

physical-layer metrics and is centered around zero (the mean and median are zero for the three RSS metrics and PL and 0.57 for SNR). Consistent with the maps, the throughput metrics experience greater variation at baseline. The variation for clustered measurements appears to be somewhat larger. This result appears to disagree with the seemingly same-scale variation observed between clustered and time-delayed repeated measurements for the WiMax data. This is especially striking because the lower carrier frequency of LTE (700 MHz versus 2.5 GHz) would seem to imply a more stable channel. However, a deeper look at the distribution of this error reveals the the majority of greater variation is due to outliers. In fact, the amount of variation is only 4.2 dB for the PL metric at the 80th percentile (0 dB in the median). Similarly, the variation is approximately 7 dB at the 80th percentile (\approx 4 dB at the median) for the RSS and SNR measurements. Still, this variation is larger than might be expected, and suggests two possible explanations:

- Because the WiMax measurements were made with a finely tuned and calibrated spectrum analyzer, and the LTE measurements were made with a COTSE USB radio, the measurement device is simply more noisy and this source of error is superimposed on the observed channel error.
- The spectrum is simply more variable at 700 MHz than it is at 2500 MHz in this environment (i.e., more random attenuation from location-specific scattering).

Resolving which of these two explanations is correct is not possible with the current measurements and would require a comparative evaluation of measurement hardware that is outside the scope of this thesis research. Instead, the assumption is taken that the baseline repeated measures variance for LTE measurement and mapping is between 2 and 4 dB for clustered repeated measurements and approximately zero for time-delayed repeated measures. This indicates that unlike in the case of the WiMax measurements, multiple measurements at closely clustered locations may be necessary to sufficiently model the intrinsic small scale variation (in addition to or replacing multiple measurements in time).

Finally, consider figures 6.35 and 6.36, which show the amount of variation as a function of distance between repeated measurements (in time), and the number of measurements observed at different time delays. As with the the WiMax measurements which show little or no trend in small time scales, this data seems to confirm that the amount of time between repeated measurements will not substantially increase the

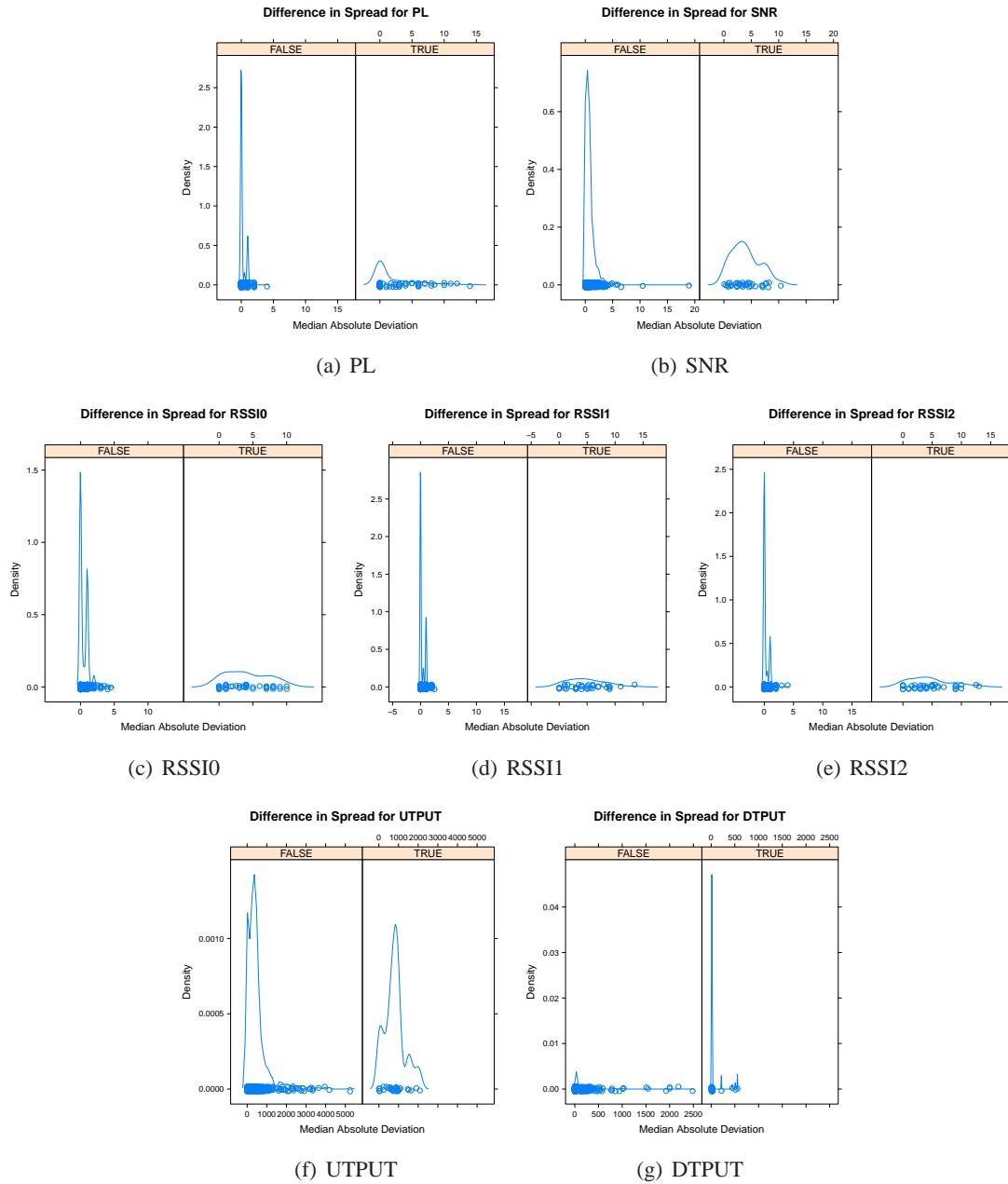


Figure 6.34: Comparison of mean absolute deviation for metrics both in the same place at different times (column FALSE, i.e. unclustered) and at different places and different times, but less than 40 wavelengths apart (column TRUE, i.e., clustered).

amount of variability observed. Instead, the number of measurements is a stronger predictor of the amount of observed variance, indicating that, in sum, a large number of co-located and clustered measurements may be necessary to model channel variability for LTE networks at 700 MHz. This includes, of course, the necessary reminder that hardware calibration (or lack thereof) may contribute (positively or negatively) to observed variance.

6.5 Summary and Conclusions

This chapter (in combination with chapter 5) has provided the first complete, real-world application of geostatistical modeling and interpolation to the problem of wireless coverage mapping. Although some other authors have proposed that geostatistical techniques may be appropriate for the domain, the work here is the first to actually apply the concepts and adapt them as necessary for the mapping of production networks. To analyze their efficacy, these techniques have been applied to two unique scenarios. This process has shown that robust coverage maps can be produced using a reasonably small amount of easily obtained data (several hundred samples on a hundred meter grid, for a space the size of a large university campus), which amounts to a tractable amount of routine “spade work” (approximately three days work for a single dedicated experimenter).

In the first scenario, extensive per-BS measurement of a 2.5 GHz WiMax network on the University of Colorado campus was performed. For that study specifically, this involves sampling on a 100m equilateral triangular grid, removing trend, truncating measurements, and interpolating using Kriging. A custom measurement apparatus was developed to allow for measurement at arbitrary locations with high precision equipment. In the second scenario, coverage of a 700 MHz Verizon LTE network on the University of Colorado campus was tested. As compared to the WiMax measurement campaign, measurement methods were further refined to include additional clustered samples to permit better modeling of small-scale (in time and space) stochastic fading effects. In both scenarios the result is a set of maps whose meaning and confidence is well defined and whose accuracy is better than what could be expected with simple measurement-based methods (powerlaw fitting) or standard predictive models. In general, an error reduction of at least several dBs over *a priori* models is obtained.

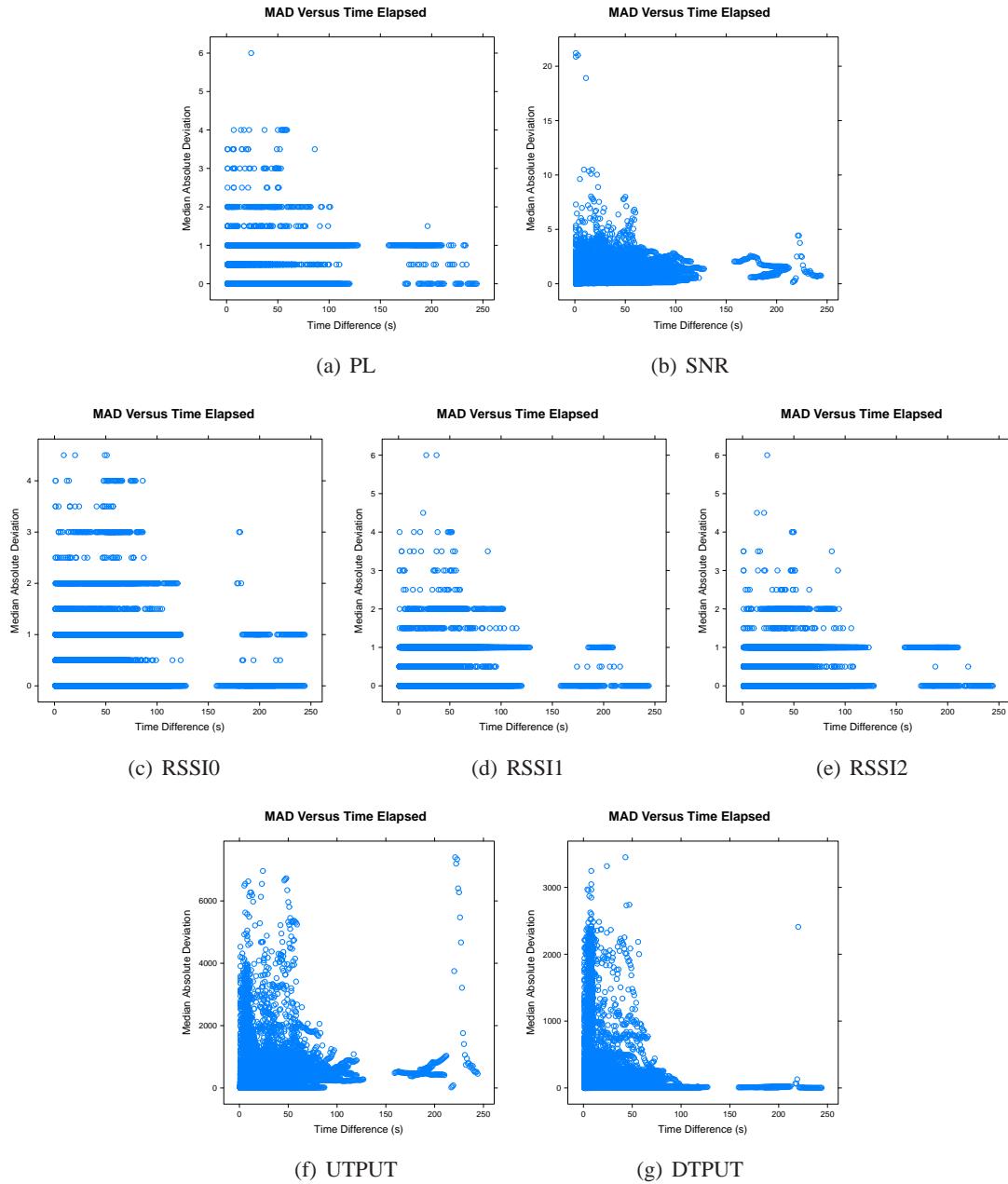


Figure 6.35: Mean absolute deviation as a function of time elapsed between measurements in seconds. There appears to be no discernable correlation.

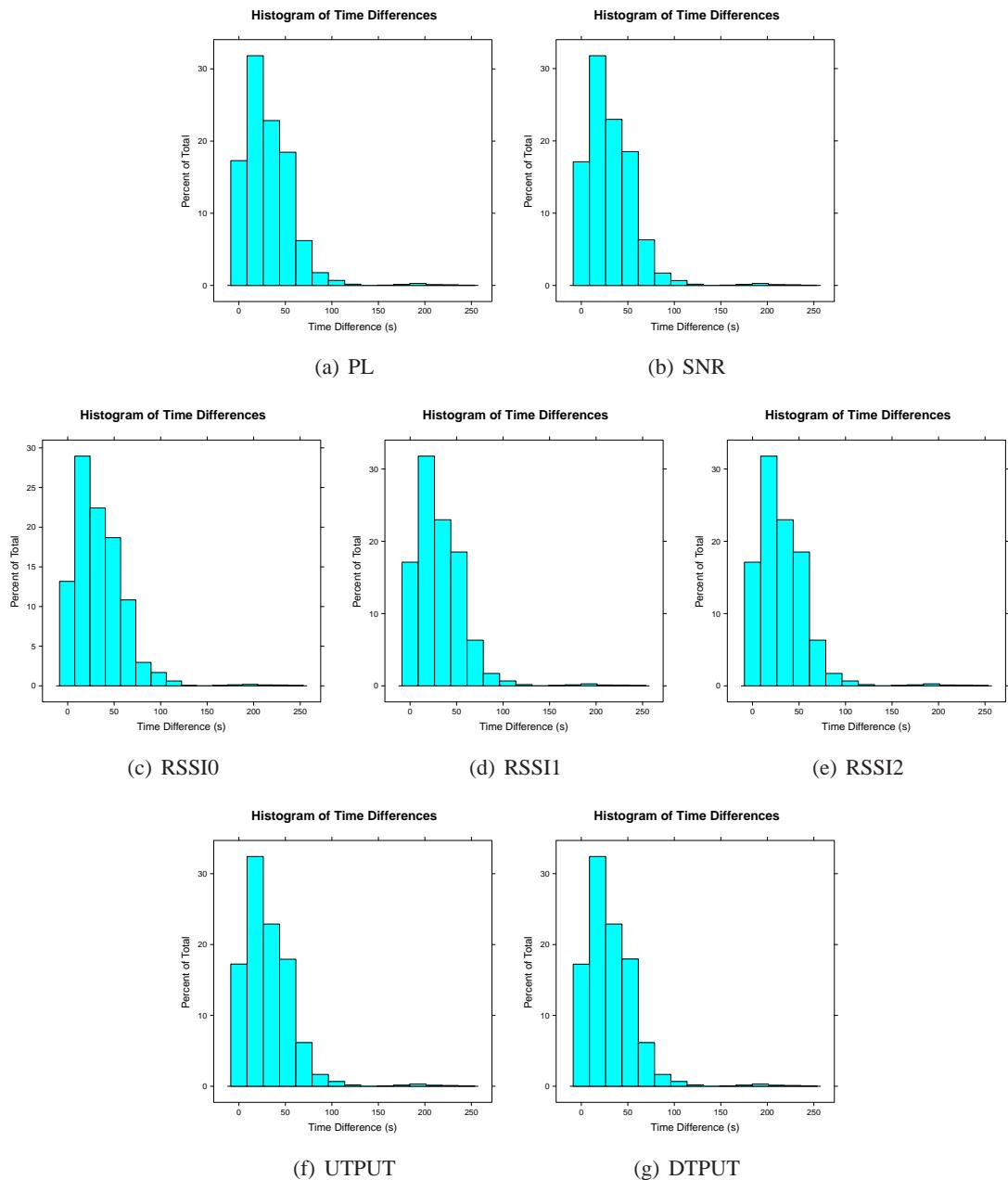


Figure 6.36: Number of samples per bucket for time elapsed between repeated measurements.

The next chapter will provide additional case studies focussing on the performance of these methods in scenarios where measurements are collected less rigorously, but with greater ease.