

**Geostatistical Techniques for  
Practical Wireless Network Coverage Mapping**

by

**Caleb Timothy Phillips**

A.S., Portland Community College, 2002

B.S., Portland State University, 2006

M.S., University of Colorado, 2010

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Computer Science

2012

This thesis entitled:  
Geostatistical Techniques for  
Practical Wireless Network Coverage Mapping  
written by Caleb Timothy Phillips  
has been approved for the Department of Computer Science

---

Prof. Douglas Sicker

---

Prof. Dirk Grunwald

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Phillips, Caleb Timothy (Ph.D., Computer Science)

Geostatistical Techniques for

Practical Wireless Network Coverage Mapping

Thesis directed by Prof. Douglas Sicker

The problem of mapping the extent of “usable” coverage of an existing wireless network is important in a large number of applications, including communicating the abilities of the network to users, identifying coverage gaps and planning expansion, discovering opportunities for spectrum reuse, and determining possible sources of interference with other networks. This thesis addresses fundamental but unsolved problems of measurement-based wireless coverage mapping: where should measurements be made, how many are necessary, and what can be said about the coverage at points that have not been measured. To address these problems, this thesis advocates a geostatistical approach using optimized spatial sampling and ordinary Kriging. A complete system for coverage mapping is developed that systematically addresses measurement, sampling, spatial modeling, interpolation, and visualization. This geostatistical method is able to produce more accurate and robust coverage maps than the current state of the art methods, and is able to discover coverage holes as effectively as dedicated heuristic methods using a small number of measurements. Several important practical extensions are investigated: applying these methods to drive-test measurements which have been resampled to alleviate effects from sampling bias, and crowd-sourced coverage mapping applications where volunteer-collected measurements may be sparse or infrequent. The resulting maps can then be refined iteratively, and updated systematically over time using an optimized iterative sampling scheme. An extensive validation is performed using measurements of production WiFi, WiMax, GSM, and LTE networks in representative urban and suburban outdoor environments.

## **Dedication**

To my grandmother, Lila May Hiatt (Richards)

## **Acknowledgements**

Thanks first, to the denizens of the systems laboratory and most especially my “elder brothers” in research, Eric Anderson and Gary Yee. Thanks to Prof. Suresh Singh who provided me the first opportunity to do interesting research and walked with me both up mountains and hard research problems. Thanks to Russell Senior who taught me about patience and care in research and data collection and inspired me with his tireless rigor and curiosity. Thanks to my fellow volunteers at the Personal Telco Project who inspired me with their passion for wireless networks and free-as-in-freedom communication. Thanks to Mark Dailey who taught me to be a hacker and gave me my first real opportunity in industry. Thanks to my very supportive and helpful advisor, Prof. Doug Sicker, and my at-times informal advisor, Prof. Liz Bradley, as well as all of the wonderful teachers I have had along the way. Not least of all, the computer science teachers who got me started on this path: Dennis Heckman, Elizabeth Odekirk-Hash, Dave Blizzard, Melanie Mitchell, and Bart Massey. Finally, but most importantly, thanks to my loving family, wonderful partner Becky, amazing friends, big mountains, fast bicycles, hard climbs, loud music, delicious food, strong ales, and sunny Colorado days, all of which helped keep me sane and positive during the struggle.

## Contents

### Chapter

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation and Applications . . . . .	3
1.1.1	Cognitive Networks . . . . .	3
1.1.2	Self-optimizing Networks . . . . .	3
1.1.3	General Network Evaluation and Planning . . . . .	4
1.1.4	Contractual Verification . . . . .	5
1.1.5	Detecting Spurious Emissions . . . . .	5
1.2	The Case for Hybridization . . . . .	6
1.3	The Case for Geostatistics . . . . .	8
1.4	Thesis Statement . . . . .	8
1.5	Dissertation Outline . . . . .	9
<b>2</b>	<b>Path Loss Prediction</b>	<b>11</b>
2.1	Radio Propagation Basics . . . . .	14
2.1.1	Signal Propagation . . . . .	14
2.1.2	Path Loss . . . . .	15
2.2	Modeling Path Loss <i>A Priori</i> . . . . .	18
2.2.1	Theoretical/Foundational Models . . . . .	20
2.2.2	Basic Models . . . . .	23

2.2.3	Terrain Models . . . . .	36
2.2.4	Supplementary Models . . . . .	39
2.2.5	Stochastic Fading Models . . . . .	43
2.2.6	Many-Ray Models . . . . .	44
2.3	Modeling With Measurements . . . . .	46
2.3.1	Explicit Mapping . . . . .	47
2.3.2	Partition Models . . . . .	48
2.3.3	Iterative Heuristic Refinement . . . . .	48
2.3.4	Active Learning and Geostatistics . . . . .	49
2.4	Comparative Studies . . . . .	51
2.5	Discussion . . . . .	52
<b>3</b>	<b>Bounding the Error of Path Loss Prediction</b>	<b>55</b>
3.1	Measurement . . . . .	56
3.1.1	Packet-Based Measurements . . . . .	58
3.1.2	Rural Measurements . . . . .	60
3.1.3	Urban Measurements . . . . .	63
3.2	Implementation Details . . . . .	69
3.2.1	Terrain Databases . . . . .	70
3.2.2	Corrections for Hata-Okumura . . . . .	70
3.3	Method . . . . .	71
3.4	Simple Log-Domain Data Fitting . . . . .	72
3.5	Performance Metrics for Path Loss Prediction . . . . .	79
3.5.1	RMSE and SC-RMSE . . . . .	79
3.5.2	Competitive Success . . . . .	80
3.5.3	Individual Accuracy Relative to Spread . . . . .	80
3.5.4	Skewness . . . . .	81

3.5.5	Rank Correlation . . . . .	81
3.6	Results . . . . .	93
3.6.1	Explicit Parameter Fitting . . . . .	95
3.6.2	Factors Correlated with Error . . . . .	97
3.6.3	Distance-Hybrid Models . . . . .	97
3.6.4	Practical Interpretation . . . . .	101
3.6.5	Miscellaneous Observations . . . . .	103
3.7	Evaluation of Raytracing Systems . . . . .	104
3.7.1	Case Study: REMCOM Wireless Insight and Crowd Sourced Building Models . . .	106
3.7.2	Case Study: EDX SignalPro and Rectilinear Building Models . . . . .	107
3.7.3	Summary of Results . . . . .	109
3.8	Discussion . . . . .	111
<b>4</b>	<b>Large Area Coverage Testing</b>	<b>113</b>
4.1	Method . . . . .	114
4.2	Case Study: Portland, Oregon . . . . .	117
4.2.1	Method . . . . .	119
4.2.2	Results . . . . .	123
4.3	Discussion . . . . .	128
<b>5</b>	<b>Geostatistical Coverage Mapping</b>	<b>129</b>
5.1	Geostatistics in a Nutshell . . . . .	129
5.1.1	A Random Field Called 'Z' . . . . .	130
5.1.2	The Variogram . . . . .	130
5.1.3	Kriging . . . . .	133
5.2	Spatial Sampling . . . . .	134
5.2.1	Classic Sampling . . . . .	134
5.3	Interpolation . . . . .	136



5.4	Method . . . . .	138
5.4.1	Performance Metrics and Measurement . . . . .	138
5.4.2	Selecting a Sampling Density and Pattern . . . . .	141
5.4.3	Krige the Residual . . . . .	145
5.4.4	Variogram Fitting . . . . .	149
5.4.5	Goodness of Fit . . . . .	152
5.4.6	Mapping with Ordinary Kriging . . . . .	154
5.4.7	Visualization . . . . .	156
5.5	Summary and Conclusion . . . . .	158
<b>6</b>	<b>Case Studies</b>	<b>162</b>
6.1	Case Study: University of Colorado WiMax . . . . .	162
6.1.1	Measurement Apparatus and Procedure . . . . .	164
6.1.2	Comparison of Performance Metrics . . . . .	167
6.1.3	Possible Sources of Systematic Sampling Error . . . . .	171
6.1.4	Spatial Data Characterization and Variogram Fitting . . . . .	173
6.1.5	Mapping with Ordinary Kriging . . . . .	178
6.2	Map Combining . . . . .	184
6.2.1	Data Combining . . . . .	184
6.2.2	Map Combining . . . . .	184
6.3	Case Study: Verizon LTE Coverage on the University of Colorado Campus . . . . .	189
6.3.1	Measurement Apparatus and Procedure . . . . .	189
6.3.2	Performance Metrics . . . . .	192
6.3.3	Spatial Data Characterization and Variogram Fitting . . . . .	197
6.3.4	Mapping with Ordinary Kriging . . . . .	197
6.4	Modeling Small Scale Effects and Deviation from Stationarity . . . . .	205
6.5	Summary and Conclusions . . . . .	215

<b>7</b>	<b>Drive-test and Crowd-Sourced Coverage Mapping</b>	<b>219</b>
7.1	Drive Test Coverage Mapping . . . . .	220
7.1.1	Method . . . . .	220
7.1.2	Experiment . . . . .	221
7.1.3	Results . . . . .	223
7.1.4	Discussion . . . . .	228
7.2	Crowd-Sourced Coverage Mapping . . . . .	232
7.2.1	Feasibility Study using Simulated Mobility . . . . .	232
7.2.2	Case Study: Open Signal Maps . . . . .	238
7.3	Summary and Conclusion . . . . .	247
<b>8</b>	<b>Optimized Spatial Sampling</b>	<b>251</b>
8.1	Optimized Sampling . . . . .	251
8.2	Method and Implementation . . . . .	254
8.2.1	Specification of Measurement Boundaries . . . . .	255
8.2.2	Metrics of Sampling Gain . . . . .	258
8.2.3	Simple Random Sampling . . . . .	259
8.2.4	Second Phase Greedy Optimized Sampling . . . . .	260
8.2.5	Second Phase Spatial Simulated Annealing Optimized Sampling . . . . .	263
8.3	Parallel Spatial Simulated Annealing . . . . .	273
8.3.1	Experiment 1: $N = 10$ . . . . .	277
8.3.2	Experiment 2: $N = 25$ . . . . .	277
8.3.3	Experiment 2: $N = 50$ . . . . .	277
8.4	Case Studies: University of Colorado WiMax . . . . .	280
8.4.1	Excluding Null Measurements . . . . .	289
8.4.2	Phase Three Sample . . . . .	295
8.5	Summary and Conclusion . . . . .	295

<b>9 Conclusion</b>	300
<b>Bibliography</b>	306
<b>Appendix</b>	
<b>A EDAM: The Effective Directivity Antenna Model</b>	321
A.1 Directional Models . . . . .	322
A.2 Method . . . . .	323
A.2.1 Data Collection Procedure . . . . .	324
A.3 Measurements . . . . .	325
A.3.1 Experiments Performed . . . . .	325
A.3.2 Normalization . . . . .	327
A.3.3 Error Relative to the Reference . . . . .	327
A.3.4 Observations . . . . .	330
A.4 A New Model of Directionality . . . . .	336
A.4.1 Limitations of Orthogonal Models . . . . .	336
A.4.2 An Integrated Model . . . . .	339
A.4.3 Describing and Predicting Environments . . . . .	342
A.5 Simulation Process . . . . .	344
A.6 Summary and Conclusion . . . . .	345
<b>B The Stability of The ITM for Typical Problems</b>	347
B.1 Implementation . . . . .	348
B.2 Experiment . . . . .	348
B.3 Results . . . . .	349
B.4 Discussion . . . . .	349

<b>C</b>	<b>Reference Source Code</b>	<b>354</b>
C.1	Kriging Variance . . . . .	354
C.2	Path Loss Prediction . . . . .	355
C.3	Effective Signal to Noise Ratio . . . . .	379
C.4	Spatial Simulated Annealing . . . . .	384
C.5	Variogram Fitting and Kriging . . . . .	386
C.6	Anritsu National Instruments Interface . . . . .	390
<b>Acronyms</b>		<b>393</b>

## Tables

### Table

2.1	<i>A priori</i> models studied along with their categorization, required input, coverage remarks, relevant citations, and year of (initial) publication. . . . .	19
2.2	Commonly used mathematical symbols. . . . .	21
3.1	Summary of data sets . . . . .	57
3.2	Summary of results by data set . . . . .	78
4.1	Point test state categorization . . . . .	123
4.2	Random sample performance summary . . . . .	126
4.3	Summary of a selection of long-term test statistics. . . . .	126
5.1	Requirements of an appropriate sampling design . . . . .	135
5.2	Summary of derived best practices for geostatistical mapping of wireless network coverage. .	161
6.1	Specifications of five University of Colorado WiMax BSs. . . . .	164
6.2	Best fit statistics for variogram fitting of CU WiMax BSs. . . . .	179
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. . . . .	190
6.4	Table 1 of best fit statistics for variogram fitting of LTE data. . . . .	203
6.5	Table 2 of best fit statistics for variogram fitting of LTE data. . . . .	204

7.1	Best fit statistics for variogram fitting of resampled OSM AT&T data. . . . .	240
8.1	Summary of random sample validation and home prediction accuracy for phase 2 samples. .	285
8.2	Summary of cross validation and fit-statistics for phase 2 samples. . . . .	285
8.3	Summary of random sample validation and home prediction accuracy for phase 2 samples. .	292
8.4	Summary of cross validation and fit-statistics for phase 2 samples. . . . .	292
A.1	Summary of data sets. . . . .	328
A.2	Factors influencing fitted offset values, 16-bin case. . . . .	343
A.3	Summary of Data Derived Simulation Parameters: Gain-offset regression coefficient ( $K_{gain}$ ), offset residual std. error ( $S_{off}$ ), and signal strength residual std. error ( $S_{ss}$ ). . . .	344

## Figures

### Figure

2.1	Path loss model family tree. Individual models are shown as circles and categories as are shown as rectangles. Major categories are green. Minor categories are blue. . . . .	13
2.2	Horizontal and vertical radiation patterns for a (highly directional) 24 dBi parabolic dish antenna. Image taken from L-COM antenna specifications [125]. . . . .	15
2.3	Horizontal and vertical radiation patterns for 7 dBi colinear omnidirectional antenna. Image taken from L-COM antenna specifications [125]. . . . .	16
2.4	Schematic of link geometry used by basic models. . . . .	23
2.5	Schematic of link geometry used by the Flat-Edge family of basic models. . . . .	32
2.6	Schematic of link geometry used by terrain models. . . . .	37
2.7	Example of sectorized propagation model for a single transmitter using the Robinson model. The measured (oracle) coverage is given as blue circles. The predicted/fitted coverage is given as sector boundaries that are adjusted (pushed and pulled) by additional measurements. Figure taken from [200]. . . . .	50
2.8	Example of fitted step function to measurements for the Robinson method. Figure taken from [200]. . . . .	50
3.1	Linear fit to RSS error observed from commodity cards during calibration. . . . .	58
3.2	The largest of three disconnected sections of the network (80x100km). Link color indicates strength: blue implies strong, red implies weak. Backhaul nodes (mainly 5.8 GHz) are red and CPEs are light blue. . . . .	61

3.3	Visual schematic of three urban data sets. A: roof to roof measurements from CU WART (Wide Area Radio Testbed), B: ground (utility poles) to ground (mobile node) measurements in Portland, Oregon, C: roof to ground and ground to roof measurements from CU WART.	64
3.4	University of Colorado Wide Area Radio Testbed (CU-WART)	64
3.5	Google WiFi Network in Mountain View, California	66
3.6	TFA-Wireless Network measurements in Houston, Texas	68
3.7	Schematic showing the process of parallel computation of path loss predictions using many models and many links.	72
3.8	Explicit power law fits to Data. Fit parameters are provided on the plots.	73
3.9	Explicit power law fits to Data. Fit parameters are provided on the plots.	74
3.10	Number of samples required for naïve fit. Plots show fit standard error for fits increasing random samples and a horizontal line is given at the RMSE obtained for all points.	75
3.11	Number of samples required for naïve fit. Plots show fit standard error for fits increasing random samples and a horizontal line is given at the RMSE obtained for all points.	76
3.12	Schematic explaining error ( $\epsilon$ ) and spread-corrected error ( $\epsilon'$ ) in terms of measurement spread and measured and predicted median values.	80
3.13	Five metric results for <b>WMP</b> data set	82
3.14	Five metric results for <b>PDX</b> data set	83
3.15	Five metric results for <b>PDX/stumble</b> data set	84
3.16	Five metric results for <b>WART</b> data set	85
3.17	Five metric results for <b>COST231</b> data set	86
3.18	Five metric results for <b>Boulder/ptg</b> data set	87
3.19	Five metric results for <b>Boulder/gtp</b> data set	88
3.20	Five metric results for <b>Google</b> data set	89
3.21	Five metric results for <b>TFA</b> data set	90
3.22	Five metric results for all <b>Urban</b> data sets	91
3.23	Five metric results for all data sets combined	92



3.24	Explicit parameter fitting for the Allsebrook-Parsons and Flat-Edge model parameters. . . .	96
3.25	Correlation between model accuracy and link distance for each data set. Distance is bucketed by kilometer. . . . .	98
3.26	Correlation between model accuracy and link distance for each data set. Distance is bucketed by kilometer. . . . .	99
3.27	Correlation between model accuracy and link distance for each data set. Distance is bucketed by kilometer. . . . .	100
3.28	Comparison of predicted coverage maps for Portland, Oregon using two well-performing models, with and without same scale Gaussian error included. True green indicates predicted recieved signal at -30 dBm and true red indicates predicted received signal the noise floor (-95 dBm). Intermediary values are linearly interpolated between these two color values. . .	102
3.29	Correlation between predicted values and observed values using REMCOM ray-tracing software and WiMax data. The dotted line has a slope equal to 1, which the data points would fall upon if the predictions were perfect. Deviations from this line indicate the magnitude of error. Fit and correlation statistics are given for the aggregate (all APs) predictions. To simplify the plot, points where the prediction software refused to make a prediction have been censored, as well as locations where there was no signal observed. . . . .	108
3.30	Correlation between predicted values and observed values using EDX ray-tracing software and WiMax data. The dotted line has slope equal to 1, which the data points would fall upon if the predictions were perfect. Deviations from this line indicate the magnitude of error. Fit and correlation statistics are given for the aggregate (all APs) predictions. To simplify the plot, points where the prediction software refused to make a prediction have been censored, as well as locations where there was no signal observed. . . . .	110
4.1	Signal strength from APs in the POC area. Lighter dots (green) indicate stronger signal. . . .	118

4.2	Random locations and their categorization. Green (light grey) dots were tested, purple and orange (grey) were points within the POC that were excluded because they were inaccessible, and red (dark grey) were excluded because they were not within the POC. . . . .	120
4.3	Testing apparatus. A battery powered Netgear WGT634u wireless router outfitted with a GPS device, USB storage, speakers, and an enable key. . . . .	121
4.4	Contour map of p-values for an exact binomial test as a function of maximal distance to an AP (i.e., only concerning samples within some radius) and hypothesized coverage percent. p-values below $\alpha = 0.05$ reject the null hypothesis that the hypothesized coverage percent is possible given the observations. . . . .	125
4.5	Packet loss for long term test at site C as a function of time of day. Measurements are averaged across days and bucketed per hour. . . . .	127
5.1	Explanation of variogram model parameters using the Matérn model as an example. Figure taken from [236]. . . . .	131
5.2	Examples of sampling schemes . . . . .	134
5.3	Example of frequency selective fading. Figure taken from [90]. . . . .	138
5.4	Example of uniform equilateral triangular sample with $h = 100m$ . Markers highlighted blue (as opposed to white) are within 40 wavelengths (approximately 5 meters) of a measurement point. . . . .	144
5.5	Path loss measurements for “pdx90” AP, both as-collected and resampled at varying lag distances. Each figure contains four plots which show the spatial and value distribution of the processes . . . . .	146
5.6	Path loss measurements for “pdx90” AP, both as-collected, and resampled at varying lag distances. Each figure plots the path loss as a function of distance on a log-log plot. A linear least squares regression fit line, and parameters are given. . . . .	147
5.7	Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances. . . . .	148
5.8	Empirical semivariograms of path loss for “pdx90” AP without de-trending. . . . .	150

5.9	Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances with Frii’s freespace model used for de-trending. . . . .	151
5.10	Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances with offset- and slope-fitted Frii’s freespace model used for de-trending. Measurements at the noise floor have been inserted at points where an observation was unable to be made. . .	153
5.11	Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances with Frii’s freespace model used for de-trending. Measurements at the noise floor have been inserted at points where an observation was unable to be made. . . . .	155
5.12	Comparison of color maps using a Kriged map of WiMax CINR coverage for one BS at the University of Colorado. . . . .	157
5.13	Examples of coverage map (for CU WiMax cuEN node) overlayed on Google Earth orthoimagery, digital terrain, and 3D models. . . . .	159
5.14	Examples of coverage map (for CU WiMax cuEN node) overlayed on Google Earth orthoimagery, digital terrain, and 3D models with measurement locations and values. . . . .	160
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. . . . .	163
6.2	Diagram showing connectivity and specification of WiMax measurement cart devices. . . . .	165
6.3	Diagram of WiMax measurement cart. . . . .	165
6.4	Correlation between various metrics. . . . .	169
6.5	Correlation between WiMax throughput and metrics that are meaningfully correlated with application-layer performance. . . . .	170
6.6	Correlation between CINR and GPS accuracy . . . . .	172
6.7	Measurement location error and GPS accuracy as a function of the measurerer. . . . .	174
6.8	Measurements for CU WiMax cuEN (GENI) node. . . . .	175
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. . . . .	176

6.10	Empirical variogram and fits of four metrics for CU WiMax cuEN (GENI) node. . . . .	177
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. . . . .	181
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. . . . .	182
6.13	Binary coverage maps with the threshold of CINR=30 . . . . .	183
6.14	Empirical variogram and fits of four metrics for the combined CU WiMax measurements. . .	185
6.15	Kriged maps for combined CU WiMax measurements using the CINR metric. . . . .	185
6.16	Binary coverage map for the combined CU WiMax measurements. . . . .	185
6.17	Kriged maps for combined CU WiMax measurements using the CINR metric. . . . .	188
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. . . . .	188
6.19	Diagram showing connectivity and specification of LTE measurement cart. . . . .	190
6.20	Diagram of LTE measurement cart. . . . .	191
6.21	Correlation between upstream throughput performance and physical-layer metrics. . . . .	195
6.22	Correlation between physical-layer metrics. . . . .	196
6.23	Log/log fit of path loss for LTE BS 369. . . . .	198
6.24	LTE Measurements. . . . .	199
6.25	Measurements for LTE BS 369. . . . .	200
6.26	Variogram fits for all measurements combined. . . . .	201
6.27	Variogram fits for BS 369. . . . .	202
6.28	Kriged maps for BS 369. . . . .	206
6.29	Kriged maps for all measurements combined. . . . .	207
6.30	Map-combined maps for all measurements combined. . . . .	208
6.31	Threshold-based map-combined maps for all measurements combined. . . . .	209

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). . . . .	211
6.33	Amount of spread (variation) as a function of time elapsed between measurements using two different metrics. . . . .	212
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). . . . .	214
6.35	Mean absolute deviation as a function of time elapsed between measurements in seconds. There appears to be no discernable correlation. . . . .	216
6.36	Number of samples per bucket for time elapsed between repeated measurements. . . . .	217
7.1	Schematic describing validation process for drive-test coverage mapping. . . . .	222
7.2	Performance results for different resampling lags and careful and aggressive resampling using four metrics of interest. . . . .	225
7.3	ROC and DET curves for each dataset using “aggressive” resampling. . . . .	226
7.4	Hole prediction accuracy as a function of sampling density. . . . .	227
7.5	Map-combined (maximum combining) Kriged coverage map for PDX Data using best variogram and h=100m. . . . .	229
7.6	Map-combined (maximum combining) Kriged coverage map for Google Data using best variogram and h=100m. . . . .	230
7.7	Map-combined (maximum combining) Kriged coverage map for TFA Data using best variogram and h=100m. . . . .	231
7.8	Sample coverage as a function of percentage of participating users and length of experiment. The left-hand plots show the number of visits to each 5 meter cell on a log scale. The right-hand plots show a boolean map of cell coverage. . . . .	235
7.9	Sample coverage as a function of percentage of participating users and length of experiment. . . . .	236

7.10	Open Signal Maps measurements for AT&T network overlayed on Google orthoimagry. . . .	239
7.11	Performance results for OSM data at different resampling lags using four metrics of interest.	241
7.12	Performance results for OSM data at different resampling lags using four metrics of interest.	242
7.13	Performance results for OSM data at different resampling lags using four metrics of interest.	243
7.14	Performance results for OSM data at different resampling lags using four metrics of interest.	244
7.15	Map-combined maps using maximum-based combining for OSM data. . . . .	245
7.16	Map-combined threshold maps using maximum-based combining for OSM data. . . . .	246
7.17	Performance results for OSM data at different resampling lags using four metrics of interest.	248
7.18	ROC and DET curves for OSM measurements using “aggressive” resampling. . . . .	249
7.19	Hole prediction accuracy as a function of sampling density for OSM data. . . . .	249
8.1	Systematic (grid) sampling augmented with nested “random” samples. In the $N \times N$ case, points are placed on an equally spaced regular grid of $N \times N$ points and then $N$ additional points are sampled in the immediate neighborhood of $N$ grid points. This has the effect of creating an initial sample that both covers the region and has a range of lag distances between measurements for variogram estimation. Figure taken from [63]. . . . .	253
8.2	Optimized placement of second-phase samples according to the weighted Kriging variance proposed by Delmelle. Figure taken from [63]. . . . .	253
8.3	Convex hull around first-phase sample points for cuEN node. . . . .	256
8.4	Manual specification of boundary and unmeasureable polygons using the Quantum GIS software. . . . .	257
8.5	Performance gain from increasingly large simple random samples. . . . .	261
8.6	Maps used for sample optimization for cuEN node. . . . .	262
8.7	Improvement (gain) from iterative greedy sampling using the WPE map. . . . .	262
8.8	Greedy sample of 100 points. Preexisting sample locations are denoted with circles and second-phase (greedy) samples are filled (red) triangles. . . . .	264
8.9	Metropolis probability plot for spatial simulated annealing. . . . .	266

8.10 Improvement (gain) from spatial simulated annealing using the WPE map for sets of 10 or 25 additional points. . . . .	268
8.11 Optimized sample for 10 and 25 points after 1000 iterations. . . . .	269
8.12 Improvement (gain) from spatial simulated annealing using the WPE map for 50 additional points with increasing numbers of iterations . . . . .	270
8.13 Optimized sample using the WPE map for 50 additional points with increasing numbers of iterations . . . . .	271
8.14 Optimized sample using the WPE map for 50 additional points with increasing numbers of iterations, but with the ESNR metrics. . . . .	272
8.15 Optimization timeseries for $N = 25$ . . . . .	275
8.16 Distribution of gain for $N = 25$ . Register your sighting of Thesis Waldo at <a href="http://thesiswaldo.org">thesiswaldo.org</a> . Waldo image by Martin Handford, distributed by Classic Media Distribution Ltd.. . . . .	276
8.17 Parallel versus sequential performance for $N = 10$ . . . . .	278
8.18 Parallel versus sequential performance for $N = 25$ . . . . .	279
8.19 Distribution of gain for $N = 50$ with temperatures of 1000 and 2000. . . . .	281
8.20 Parallel versus sequential performance for $N = 50$ . . . . .	282
8.21 Optimized second-phase sample for cuEN node with 50 points, overlayed on Google Maps orthoimagery. . . . .	284
8.22 Comparison of second-phase threshold maps for CINR=20. . . . .	286
8.23 Comparison of second-phase coverage. . . . .	287
8.24 Comparison of second-phase variograms. . . . .	288
8.25 Measured second-phase sample for cuEN node with 50 points, overlayed on Google Maps orthoimagery. . . . .	290
8.26 Phase-2 Optimized Sample. . . . .	291
8.27 Comparison of second-phase threshold maps (threshold is CINR = 40 dB). . . . .	293
8.28 Comparison of second-phase coverage. . . . .	294
8.29 Phase-3 Optimized Sample. . . . .	296

8.30 Measured third-phase sample for cuEN node with 50 points, overlaid on Google Maps orthoimagery. . . . .	297
8.31 Comparison of second-phase threshold maps (threshold is CINR = 40 dB). . . . .	297
8.32 Comparison of second-phase coverage. . . . .	298
A.1 Probability Density Function (PDF) of percentage of dropped measurement packets in a given angle for all angles and all data sets. . . . .	331
A.2 Receiver side of measurement setup in floodplain . . . . .	331
A.3 Floorplan of office building used in Array-Indoor-A, Array-Indoor-B, Patch-Indoor-B, Patch-Indoor-C, Parabolic-Indoor-B, and Parabolic-IndoorC. . . . .	332
A.4 Comparison of signal strength patterns across different environments and antennas. . . . .	333
A.5 Probability Density Functions (PDFs) for the averaged error process (combined across mul- tiple traces) for each antenna type. . . . .	334
A.6 Cumulative Density Functions (CDFs) for the error process (combined across multiple traces) for each antenna type. . . . .	334
A.7 Probability Density Function (PDF) of p-values from testing the normality of the error pro- cess in each direction for each data set. In all cases, the null hypothesis (that the samples are normally distributed) can be confidently rejected. . . . .	335
A.8 Heatmap of p-values for the Mann-Whitney U-test which was run pairwise against the error from the reference pattern in each angle. This plot, which is for Patch-Indoor-A, was chosen as a representative. <i>All traces showed similar trends.</i> Darker values indicate very small p-values, meaning that the null hypothesis can be rejected with confidence. In this case, the null hypothesis is that the samples come from the same distribution. The Patch reference pattern is provided on the left for reference. . . . .	335
A.9 Differences between the orthogonal model and observed data in dB: $\hat{P}_{rx} - P_{rx}$ . . . . .	337
A.10 Mean error of orthogonal model for each observation point in the Array-Outdoor-A data set. The format is the same as in figure A.9. . . . .	338



A.11 Effect of increasing bin count (decreasing bin size) on modeling precision. . . . .	341
A.12 Residual error of the discrete offset model with 16 bins. . . . .	341
B.1 Box and whiskers plot of error as a function of precision. . . . .	351
B.2 Box and whiskers plot of error as a function of precision, showing only results for single- precision and greater arithmetic. . . . .	352
B.3 Running time of ITM algorithm as a function of precision. The 0-bit case is the machine- precision reference implementation. . . . .	353