Chapter 5

Geostatistical Coverage Mapping

This chapter outlines an approach for robust coverage mapping using principled spatial sampling and geostatistical interpolation ("Kriging"). The chapter begins by providing some background on geostatistics for the uninitiated, since understanding the problem definition and assumptions are necessary to motivate the approach. Then, classic spatial sampling schemes are described along with more advanced multi-phase optimized sampling schemes (similar to those presented in chapter 8). Section 5.4 will walk though the fitting and mapping method proposed here as well as provide explanations and examples for the approach taken. This will set up the next chapter, 6, which descibes the application of these methods in two case studies mapping the coverage of production networks.

5.1 Geostatistics in a Nutshell

There are a number of textbooks that cover the topic of geostatistics in depth. The encyclopedic treatment by Cressie is a fine starting point [55]. However, Wackernagel's text [233] is more approachable for most topics and [118] provides a concise discussion of state-of-the-art and advanced geostatistical modeling techniques. Other texts that are less lucid but still worth mentioning are Ripley's very dense books [197, 196] and the thorough treatments in [46] and [104]. Also worth noting is [124], where Krige and Kleingold provide a history of the development of the field throughout the 1950's, 1960's, and 1970's.

5.1.1 A Random Field Called 'Z'

If we assume that there is a random field being modeled called Z, then the value of that field at a point in space $\mathbf{x}$ is $Z(\mathbf{x})$. The field can be defined in any dimension, but it is typically assumed that $\mathbf{x} \in \mathbb{R}^n$ with $n = 2$ or $n = 3$. The value at any point can be defined as the field mean (μ) plus some error ($\epsilon(\mathbf{x})$):

$$Z(\mathbf{x}) = \mu + \epsilon(\mathbf{x}) \tag{5.1}$$

5.1.2 The Variogram

Central to geostatistics is the variogram, a function that models the variance between two points in space as a function of the distance between them ($\mathbf{h}$). In the case of grid-sampled fields, the distance between measurements is a fixed lag distance. Randomized and optimized sampling schemes produce variable lag distances. The theoretical variogram, γ, is typically written as a function of the expected value of the squared difference between a given point value and a point some lag h way:

$$\gamma(\mathbf{h}) = \frac{1}{2}E[(Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x}))^2] \tag{5.2}$$

If it is known that the field is second order stationary (i.e., a measurement at the same point will not vary with time and the difference between two measurements at the same two points will not vary with time), then the covariance function (correllogram) is defined as:

$$C(\mathbf{h}) = E[(Z(\mathbf{x}) - \mu)(Z(\mathbf{x}+\mathbf{h}) - \mu)] = C(\mathbf{0}) - \gamma(\mathbf{h}) \tag{5.3}$$

However, second order stationarity is probably not a safe assumption for the radio environment (not without some effort to correct for temporal variation anyhow). With some set of measurements, an empirical variogram can be defined as the sum of squared differences for each observed lag distance h_i:

$$\gamma'(\mathbf{h}_i) = \frac{1}{2n}\sum_{j=1}^{n}(z(\mathbf{x}_j + \mathbf{h}_i) - z(\mathbf{x}_j))^2 \tag{5.4}$$

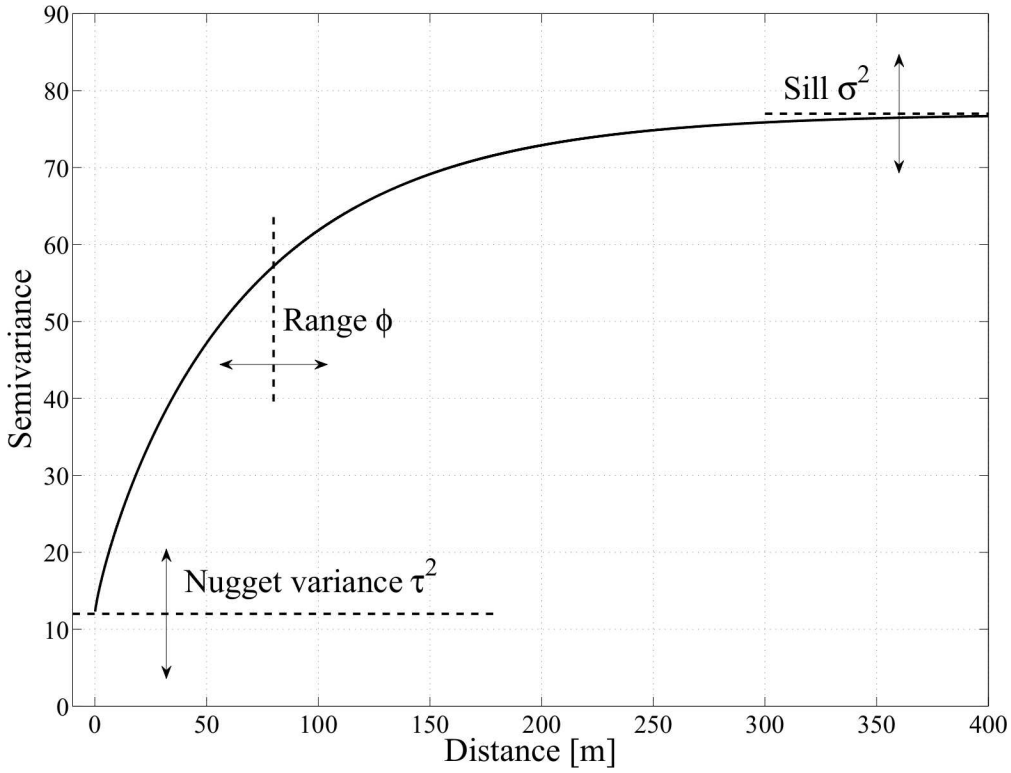


Figure 5.1: Explanation of variogram model parameters using the Matérn model as an example. Figure taken from [236].

A typical problem is to fit a variogram (or correllogram) given some number of measurements. There are a number of models that can be used for fitting. One example is the exponential model:

$$\gamma_{exp}(h) = \tau^2 + \sigma^2(1 - e^{-h/\phi}) \tag{5.5}$$

Figure 5.1 shows how the parameters of this model effect the shape of the fitted variogram. τ^2 is known as the nugget variance and is used to model discontinuity around the origin, by increasing or decreasing the starting threshold. It is so thusly named because this parameter helps model the likelihood of rare minerals ("nuggets") in geological surveys. σ^2 is known as the sill because it sets the maximum value of the semivariogram. Larger values of σ will increase the level at which the curve flattens out. Finally, the parameter ϕ acts as a scale and affects the overall shape of the curve. The value of ϕ determines the rate at which variance is expected to appear as a function of distance (lag) between points. There are a number of other models, such as the Gaussian, Cauchy, and Matérn models, which may or may not be the best fit depending on the data. [159] provides an accessible introduction to variogram fitting by walking through an example fitting. In that work, Olea discusses the pros and cons of various permissible functions, efforts to remove trend, nested models and anisotropic methods. In addition to the classic models, more advanced (and nested) models can be used so long as they are permissible. For instance, [178, 113] discuss the (optionally damped) cosine Hole-Effect model which is able to capture periodicities at varying scales within the semivariogram.

Variogram fitting can be carried out using a variety of methods. Historically, the method of moments has been used to derive a fitted variogram. However, Maximum Likelihood Estimator (MLE) and Least Squares (LS) methods have been used more recently with some substantial success. In [163], Pardo-Igúzquiza comes out as a proponent of the MLE method, claiming that its requisite assumption of an underlying Gaussian distribution is reasonable, and supports these conclusions with simulations. In [162], Pardo-Igúzquiza describes a software package to fit the variogram to one of several models using this MLE method, along with a modified ("restricted") MLE approach that avoids errors from simultaneously estimating the drift and covariance parameters of the variogram. In [112], Jian *et al.* argue for a Weighted Least

Squares (WLS) approach and suggest that the Akaike Information Criterion (AIC) be used as a goodness of fit metric. Later work by Lark [127] gives a more rigorous comparison of these two methods and finds that for both simulated and real data sets, the method of moments and MLE fitted models are not substantially different. Both methods are susceptable to distributional skew and outliers. However, for some specific cases, each approach outperforms the other. For instance, when nugget variance is relatively small and the correlation range of the data is large, method of moments performs better. In sum, Lark recommends that fits be made with both methods, and the resulting modes compared.

5.1.3 Kriging

"Ordinary" Kriging is an interpolation technique that predicts the unknown value at a new location ($Z(\mathbf{x}')$) from the weighted known values at neighboring locations ($\mathbf{x}_i$):

$$Z_K(\mathbf{x}') = \sum_{i=0}^{n} w_i Z(\mathbf{x}_i) \tag{5.6}$$

and, to determine the optimal weights ($\mathbf{w}$), we must minimize the estimation variance σ_E^2:

$$\sigma_E^2 = E[(Z_k(\mathbf{x}') - Z(\mathbf{x}')^2] \tag{5.7}$$

with

$$\sigma_E^2 = -\gamma(\mathbf{x}' - \mathbf{x}') - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + 2\sum_{i=1}^{n} w_i \gamma(\mathbf{x}_i - \mathbf{x}') \tag{5.8}$$

which leads to the following system of equations:

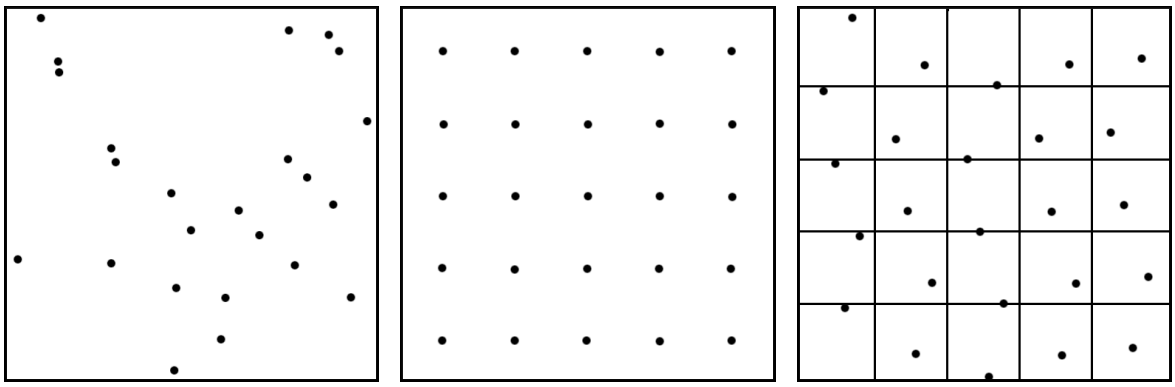
$$\begin{pmatrix} \gamma(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & \gamma(\mathbf{x}_1 - \mathbf{x}_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(\mathbf{x}_n - \mathbf{x}_1) & \cdots & \gamma(\mathbf{x}_n - \mathbf{x}_n) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma(\mathbf{x}_1 - \mathbf{x}_0) \\ \vdots \\ \gamma(\mathbf{x}_n - \mathbf{x}_0) \\ 1 \end{pmatrix} \tag{5.9}$$

where μ is called the Lagrange parameter. This interpolation is "exact", meaning that $Z_K(\mathbf{x}') = Z(\mathbf{x})$ if $\mathbf{x} = \mathbf{x}'$. This approach can be used in mapping by Kriging the value at each pixel position.

The quality of an interpolated field depends on the goodness of the fitted variogram (γ). In addition to this, there are a number of different ways to adapt Kriging to a specific data set. Anisotropic corrections are of particular interest for coverage mapping. This approach assumes that the field may require different statistics (i.e., a different variogram and possibly fitting method) in different directions from some point. There is also an entire branch of statistics dealing with multivariate analysis (i.e., co-Kriging). Depending on the importance of the time dimension, for instance, these directions may be of particular interest.

5.2 Spatial Sampling

Despite the many measurement-based approaches to path loss prediction and coverage mapping, no single work has looked at the important questions of *where* these measurements should be made and how many of them are needed. These questions of *where* and *how many* are at the center of this thesis. This section provides background on spatial sampling.



(a) Random Sample (b) Systematic Sample (c) Systematic Unaligned Sample

Figure 5.2: Examples of sampling schemes

5.2.1 Classic Sampling

Choosing an appropriate sampling scheme is exceptionally application dependent. The shape and variance of the field, as well as domain-specific knowledge about the process being modeled, must all be considered when selecting a sampling strategy. In [214], Stehman proposes a useful taxonomy of tradeoffs

1	Satisfies probability sampling protocol
2	Is simple to implement and analyze
3	Has low variance for estimates of high priority
4	Permits adequate variance estimation
5	Is spatially well distributed
6	Is cost effective

Table 5.1: Requirements of an appropriate sampling design

for spatial sampling schemes that is reproduced for reference in table 5.1.

SRS is the classic approach used in many spatial sampling problems. An example is given in figure 5.2(a). It is simple, straightforward, and well understood. In particular, SRS is convenient in that any consecutive subset of a simple random sample is also a simple random sample itself. This means that one can create a sample of points, and then test them consecutively until a statistical significance criterion is met.

Competing sampling schemes might include systematic (see figure 5.2(b)), systematic unaligned (see figure 5.2(c)), or stratified. Systematic sampling can be dangerous because it risks alignment bias. It is admittedly tempting when measuring coverage to align samples along an obvious geographic feature of cities, i.e., streets. However, this thesis argues that aligning samples along streets risks highly biasing results. Aside from degenerating to a type of one-sided stratified sampling, streets also have the capacity to act as RF waveguides (sometimes called “street canyons” in the literature). Systematic unaligned sampling can be a good compromise between SRS and systematic sampling as it is more robust against alignment bias, but guarantees an even distribution of sample points within the test area. Stratified sampling is typically used when there are differences and/or differences in variability in different areas. For instance, a municipality may wish to prioritize or set different performance and coverage criterion for different areas of a city.

For the purposes of geostatistical modeling (and Kriging), there are two important criteria that must be considered when selecting an intial sampling design. First, samples must cover the area to be sampled such that no two points are too far apart, which decreases interpolation resolution. And second, some number of samples must be taken at a variety of lags so that the variogram can be sufficiently estimated. In particular, clustered measurements are generally required to model small scale effects (i.e., variance from measurements separated by distances smaller than the lag distance). In [158], Olea investigates multiple

initial sampling schemes. In his approach, universal Kriging is used to select between several specific designs so that standard error is minimized. Olea strongly endorses stratified random sampling in this work, but it is not clear how well this mechanism works in other domains. In [245], Yfantis *et al.* study the efficiency of Kriging estimation for various types of sampling lattices. They find that, for the majority of cases, where the nugget effect is small relative to the variance, a triangular grid-based sample is the most efficient initial sampling scheme. In cases where the nugget variance is large and the linear sampling density is > 0.85 times the range, a hexagonal design is most efficient. The authors suggest that a small pilot sample be used to determine the empirical variogram, which can then chose an appropriate-density and grid-pattern sampling scheme for the initial sampling.

In addition to these works, if something is known about the underlying process and its variability, an optimization scheme can be used to select the best initial sample. For instance, in [230], van Groenigan et al., present a framework for Spatial Simulated Annealing (SSA) which uses a fitness function that either spreads points maximally, or chooses their lags according to a prescribed distribution. In SSA, points are varied randomly in a hill-climbing fashion so that an (at least locally) optimal sample is chosen. Additionally, if the variogram shape is known *a priori*, or a distribution of reasonable variogram parameters can be defined, then an initial sample can be chosen using SSA so as to minimize the summed or average point Kriging variance. Although related, this approach to optimizing the initial sample differs from the approach detailed in chapter 8, which seeks to optimize second-phase samples.

5.3 Interpolation

The question of interpolation is at the center of any measurement-based approach to coverage mapping. If measurements are collected at some number of points in a given region, what can be said about the points that have not been measured? Interpolation addresses this problem.

Besides general purpose spatial interpolation, there have been several papers that have attempted to develop interpolation strategies appropriate for wireless coverage mapping. In [52], Connelly *et al.* suggest a way to interpolate signal strength between RSS measurements using inverse distance weighting and claim less than 1 dB interpolation error. However, their minimal attempt to validate this, along with a lack of

realism in some of their assumptions (for instance, assuming propagation *stops* at 100 m), leaves one without much confidence in their method. In [56], Dall'Anese suggests a way to use distributed measurements from sensors to determine a sparsity-promoting WLS interpolated coverage map. This work is in the space of cognitive radios, so the authors assume that the location of sensors is not controllable and that the principle application is in empirically determining a safe transmit power for a given radio so as to avoid interfering with Primary User (PU)s. In [119], Konak proposes the use of ordinary Kriging over grid-sampled data for mapping coverage and shows that this approach can outperform a neural-network trained model presented in [150]. Finally, [161] provides a tutorial addressing the use of basic geostatistical interpolation for estimating radio-electric exposure levels. While not strictly the same as wireless network propagation, the approach is certainly relevant.

In addition to these works, there have been several recent publications by Riihijärvi and colleagues that discuss the use of spatial statistics to model radio propagation [195, 236]. As with [56], Riihijärvi's work is in the cognitive radio space, where the goal is to determine the signal at a given point from a PU so that a secondary user can choose when and where it is safe to transmit without interfering. Like [119], this work presumes a regular grid-based sample. Measurements are used to fit a semivariogram and several underlying functions are investigated. In [193], the authors suggest how this method can be used to more compactly store radio environment maps and in [194] the same authors look at how the placement of transmitters, terrain roughness, and assumed path loss effects the efficacy of the interpolated field. The theoretical work by Riihijärvi here is solid and is very inspiring, yet has two important limitations that this chapter (and the following) aims to address: (1) Riihijärvi does not evaluate the model with real measurements and hence it is difficult to say how well this approach would work in practice and (2) the work does not concern itself with where measurements are made and assumes simple grid-based sampling for measurement. The work of this thesis will build upon the work of Riihijärvi by making an empirical evaluation of these geostatistical techniques, applying them to the general case of coverage mapping, and exploring more advanced (optimized) sampling strategies.

As compared to alternative methods of interpolative mapping such as Inverse Distance Weighting (IDW), Kriging has three important benefits: (1) it is preceded by an analysis of the spatial structure of the

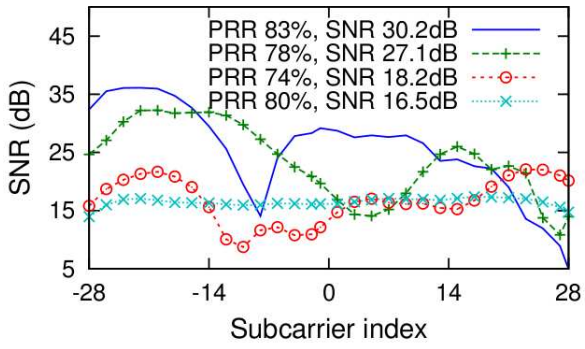


Figure 2: Channel gains on four links that perform about equally well at 52 Mbps. The more faded links require larger RSSIs (i.e., more transmit power) to achieve similar PRRs.

Figure 5.3: Example of frequency selective fading. Figure taken from [90].

data and thus an estimate of the average spatial variability of the data is integrated into the interpolation process *vis a vis* the variogram model, (2) it is an exact interpolation method, meaning that when data is available at a given point, the interpolated map has exactly that measured value at that point; and (3) since it is a robust statistical method, it provides a per-prediction indication of estimation standard error via the square root of the Kriging variance [233].

5.4 Method

This section describes the method for geostatistical mapping developed for the radio environment in this thesis. Of course, the method described here was not found on first trial and is the result of many false starts, mis-turns, and exciting discoveries. Although the description here is on the final method developed, some time will be given to explaining why particular design choices were made. The first subsection begins by discussing the important question of performance metrics and measurement, and the following sections will discuss the process of geostatistical mapping, which involves sampling design, spatial structure analysis, de-trending, variogram fitting, Kriging, and finally mapping and basic visualization.

5.4.1 Performance Metrics and Measurement

The question of what to measure or, which metric is the best predictor of network performance at a given point, is intimately tied to the efficacy of any coverage mapping system. The vast majority of research

has been done on measuring and predicting signal, noise, and SNR. However, there has been a vocal minority of researchers and network operators that have been claiming for some time that single value SNR is a poor predictor of network performance. In [132], Lee *et al.* propose a method for averaging signal strength measurements over a region of 20 to 40 wavelengths in order to average out small scale fading effects and obtain a more stable indicator of signal strength. By and large, this averaged SNR metric is what is used by cell network operators today.

5.4.1.1 Metrics for Simultaneous Multiple Streams

More recently, there has been some work to develop metrics appropriate for wide-band Orthogonal Frequency Division Multiplexing (OFDM) waveforms and Multiple Input Multiple Output (MIMO) streams. In [92], He and Torkelson present an "effective SNR" metric for OFDM systems that involves averaging and combining across subcarriers. The most recent paper in this thread is [90] by Halperin *et al.* In this work, the authors use 802.11n radios with a customized firmware to collect the Channel State Matrix (CSM), which contains SNR fading values for each subcarrier. This allows them to develop a practical metric of performance that takes into account frequency-selective fading effects, where some OFDM subcarriers are attenuated more significantly than others. Figure 5.3 gives an example of this phenomenon. The authors propose a metric called "effective SNR", which is the SNR value that would acheive the same bitrate for the given modulation scheme as the average bitrate across all the subcarriers:

$$\hat{r} = 1/N \sum_{i} BER_k(SNR_i) \tag{5.10}$$

$$SNR' = BER_k^{-1}(\hat{r}) \tag{5.11}$$

where SNR_i is the SNR of the i^{th} subcarrier, N is the number of subcarriers (52 for 802.11 OFDM implementations), $BER_k()$ is a function that computes the expected Bit Error Rate (BER) for a given modulation scheme k and $BER_k^{-1}()$ is the inverse function that produces the SNR required to achieve a given BER, $\hat{r}$ is the average BER across all subcarriers, and SNR' is the computed effective SNR. In their paper, the authors

show that this metric is better able to predict the performance of real wireless links than average (across the 10 MHz channel) SNR.

5.4.1.2 Higher-Layer Metrics

As is discussed in [92], BER is the ideal metric of performance for a given technology and propagation environment. However, in practice its collection is costly. In chapter 4, a state-based metric of performance that incorporates tests at multiple layers was discussed. In situations where the technology is fixed, application-layer tests are obviously the best metric of performance. However, in scenarios where measurements are being made in order to predict the propagation environment alone and should not be tied to a particular technology, modulation, waveform, or rate adaptation scheme, lower-level metrics are most useful (for instance, in cognitive Radio Environment Map (REM) applications). For this reason, to maintain generality this thesis will focus on lower-level metrics such as SNR and effective SNR, while performing due diligence to understand how and how well these metrics correlate with higher-layer metrics in each of the environments studied.

There have been some recent proposals which hint at a bountiful future for robust low-level metrics that can be collected with commodity hardware. For instance, in [22], Firooz *et al.* propose a way to use the GNU Radio Software Defined Radio (SDR) platform [179] to implement a Channel Impulse Response (CIR) metric based on an 802.11 transceiver, which can be used to measure delay spread. A practical method for measuring delay spread could be used to model multipath fading and ISI effects with high precision. Another promising example is [184], where Rayanchu *et al.* show that Commercial Off The Shelf Equipment (COTSE) hardware can be used to measure and model power from interfering stations and devices. Because the methods described here are agnostic to the underlying metric used, the mapping approach can be upgraded simultaneously as better methods are developed to estimate channel performance.

5.4.1.3 Estimating Channel Occupancy

Although the case studies in this thesis focus on the task of understanding the performance of a deployed network, it should be noted that the same techniques could be used to map an interfering network

or discover empty spectrum for Cognitive Radio (CR) applications. In [23], Anderson and Cameron discuss results of a spectrum survey around Annapolis, Maryland. They find that estimations of channel occupancy can vary substantially depending on how quickly the spectrum is scanned. In particular, fast scans tend to overestimate channel occupancy because they are effected by transient signals, while slow scans tend to underestimate occupancy. The authors also suggest that because some narrowband technologies operate at very low SNR values, highly sensitive receivers are required to detect low-power transmissions. Some exciting recent work by Rayanchu *et al.* has shown that it may be possible to collect sufficiently accurate information about cross and interband interference using COTSE commodity hardware [184].

5.4.2 Selecting a Sampling Density and Pattern

As was discussed in section 5.2, the spatial sampling literature suggests that an equilateral triangular uniform lattice is often the most efficient sampling strategy for two-dimensional spatial processes, and thus this strategy is utilized in this work. A uniform equilateral triangular lattice of a given lag h in meters is generated as described in algorithm 2. Choosing a sampling density, or lag h, requires consideration of a number of tradeoffs. Firstly, there are fundamental limits in terms of the smallest meaningful lag. In [208], Shin et al., suggest that there is substantial spatial autocorrelation of measurements (of IEEE 802.11 networks) within 1m. In [132], Lee *et al.* suggest averaging measurements within 20 to 40 wavelengths to avoid overfitting a model on the noise from fast fading. Finally, because a typical GPS unit has a working accuracy of between 1.5 and 10m (depending on the environment and weather), there is little benefit to sampling at a greater density than this, since the subsequent model fitting would be, in effect, only fitting noise from the locationing error. This distance works out to between 4.8 and 2.4m at 2.5 GHz. Hence, any measurement-based interpolation should not expect to produce a map with finer resolution than 20 to 40 wavelengths, in a scenario with fine locationing resolution, and no smaller than ≈ 5m in a scenario using commercial GPS for positioning.

Based on this reasoning, one might endeavor to take as many measurements as possible at a lag slightly above this lower bound. However, in many situations, measurements are costly to collect in terms of both time and money. It is one of the aims of this thesis to provide a sampling methodology that is relatively

Algorithm 2 Compute Equalateral Triangular Sample in 2-Space

$h \leftarrow$ desired lag in meters
$R \leftarrow$ the radius of the earth in meters
$step \leftarrow (h/R) * (180/\pi)$
$lngmin \leftarrow$ minimum longitude of bounding box
$lngmax \leftarrow$ maximum longitude of bounding box
$latmin \leftarrow$ minimum latitude of bounding box
$latmax \leftarrow$ maximum latitude of bounding box
$h \leftarrow latmax$
$w \leftarrow lngmin$
$nh \leftarrow 0$
$nw \leftarrow 0$
while $h > latmin$ **do**
    while $w < lngmax$ **do**
        $nw \leftarrow nw + 1$
        $w \leftarrow w + step$
    end while
    $nw \leftarrow 0$
    $w \leftarrow lngmin + step/2$ if nh is odd, otherwise $lngmin$
    $h \leftarrow h - step$
    $nh \leftarrow nh + 1$
end while

minimal in terms of the work required. For this reason, it is most desirable to select a sampling density that provides enough information to sufficiently model the important details of the network coverage, without requiring more work than is necessary. In order to understand how sampling density affects the ability to perform a meaningful fit, an experiment was performed using the data from the "pdx" drive-test described in 4.2.1.

In this experiment, grids are generated at multiple lag distances. For each point on the grid sample, the nearest measurement point is located for each of the 72 APs using a nearest neighbor algorithm. Sample points that are not within h meters of any measurement are discarded. To manage this task efficiently, measurements are inserted into a PostgreSQL database with PostGIS extensions [160, 88]. Figure 5.4 shows a portion of the sample, with those sample points that are within 40 wavelengths of a measurement highlighted.

Figure 5.5 shows the original measurements, as well as resampled measurements with varying lag distances for a representative AP called "pdx90". These figures plot the path loss in dB (calcuated by solving equation 2.1 for PL). From the top figure, 5.5a, which shows the raw measurements, it is clear to see the measurement bias due to drive-testing—all measurements fall in straight lines confined to streets. Each sample at a lesser resolution (larger lag) approximates the original data with decreasing fidelity. This can be seen clearly by inspecting both the spatial distribution and the shape of the value distribution (which appears to be lognormal, as we might expect). Although all four resamplings capture the basic value distribution, there is a clear loss of information when the lag is greater than 100 m.

Figure 5.6 gives another view of this data that is more common in the propagation modeling literature. Here, the observed relationship between path loss and distance is plotted for each sample. Each plot also provides a linear least squares regression fit to the data using the modified version of equation 2.8: $PL = 10\alpha log_{10})(d) + 20log_{10}(f) + 32.45 + \epsilon$ where α and ϵ are the fitted slope and intercept and correspond to the path loss exponent and offset. The residual error of the fit, which can be thought of as the variability due to fast and slow fades, is given as σ. The raw measurements show an α of approximately 2.2, an ϵ of 22.8 and a σ of approximately 8.5 dB. These are all reasonable and expected values for outdoor radio transmissions. As before, each successively sparse sample can be viewed as a reduced fidelity approximation of the underlying

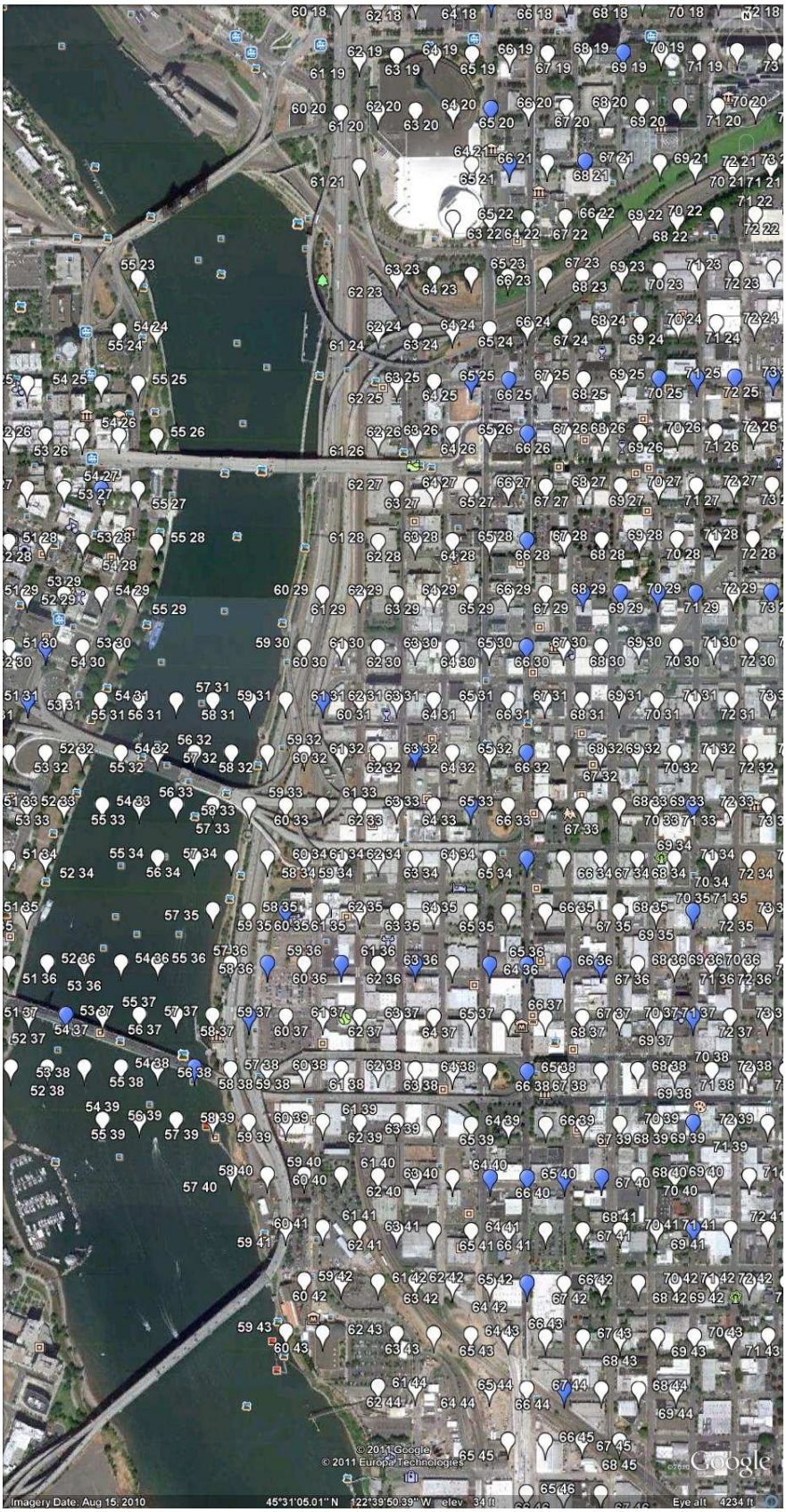


Figure 5.4: Example of uniform equilateral triangular sample with $h = 100m$. Markers highlighted blue (as opposed to white) are within 40 wavelengths (approximately 5 meters) of a measurement point.

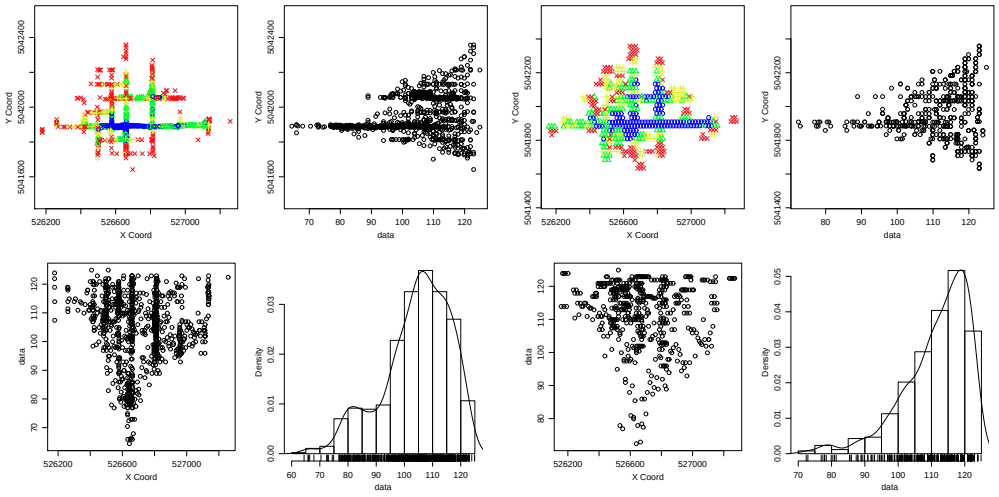
relationship and variability. At $h = 50m$ and $h = 100m$, the plots are very close. However as the lag is increased to 250 and 500 m, there are no longer sufficient samples to create an accurate picture. As a result, the fitted α, ϵ, and σ values radically underestimate those from the original measurements. Given this, a density of 100 m or less seems to sufficiently model the basic distance-attenuation relationship of this particular data set, and larger lag distances might radically underestimate the actual attentuation.

The next task is to attempt to characterize the semivariance of the measurements with a semivariogram model as described in section 5.1.2. Figure 5.7 shows the empirical variograms for the raw and resampled data. These semivariogram plots provide a final comparison of resampling densities. The raw data produces a semivariogram with nugget effect of approximately 25 dB (i.e., the y-axis crossing), range of 200 m (the location of the first peak) and sill of approximately 165 dB (the horizontal asymptote around the peak, which does not actually asymptote in this example). As before, the samplings less than and equal to $h = 100$ seem to reasonably well approximate this curve, while the other samplings deviate wildly. From the perspective of Kriging, none of these variograms are particularly well behaved as they do not have the characteristic shape and horizontal asymptote required of permissible semivariogram models; this will be addressed with more sophisticated fitting extensions below. Based on the results of this experiment, $h = 100$ appears to be a reasonable starting sampling density in the first case study. The process (and possible harm associated with) resampling of a biased sample to derive a uniform sample is the topic of chapter 7.

5.4.3 Krige the Residual

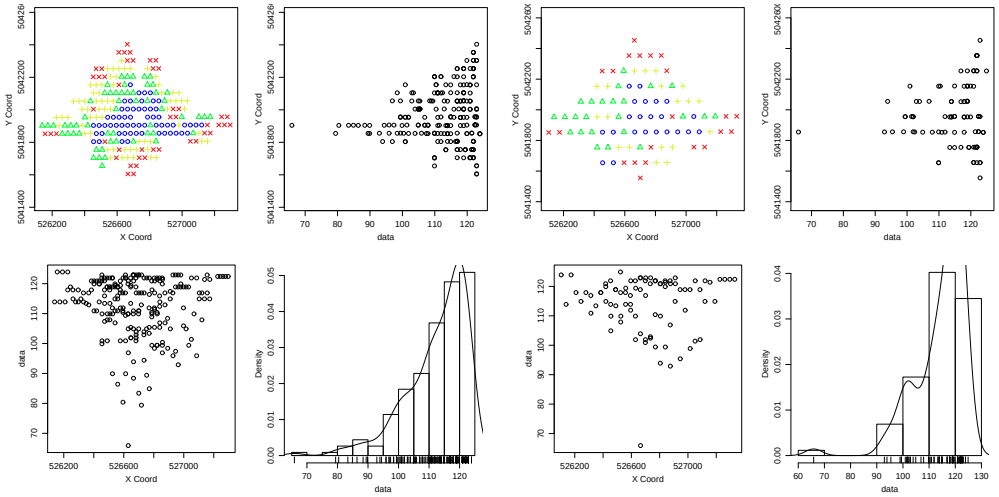
Assuming an initial sample has been obtained by either grid, random sampling, or resampling, the next task is to "de-trend" the data. In [159], Olea et al describe the importance of removing any sources of nonlinear trend from measurements so that the fitted (interpolated) field complies with the basic tenets of geostatistics. To this end, a hybrid approach is developed: a predictive (empirical) model is used to calculate the predicted path loss value at each measurement point. This prediction is subtracted from the actually observed value to obtain the residual, or excess path loss relative to the model predictions:

$$Z'(\mathbf{x}) = Z(\mathbf{x}) - P(\mathbf{x}) \tag{5.12}$$



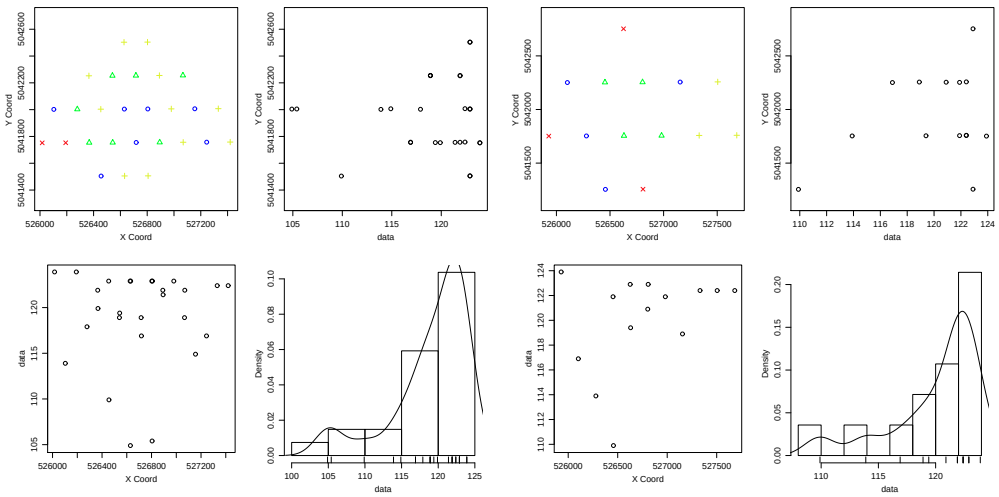
(a) Original Measurements

(b) h = 25 m



(c) h = 50 m

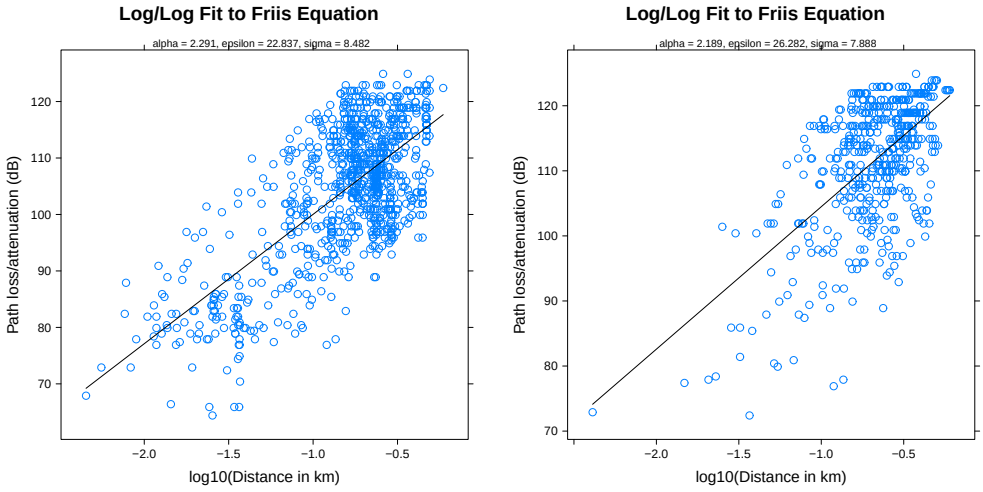
(d) h = 100 m



(e) h = 250 m

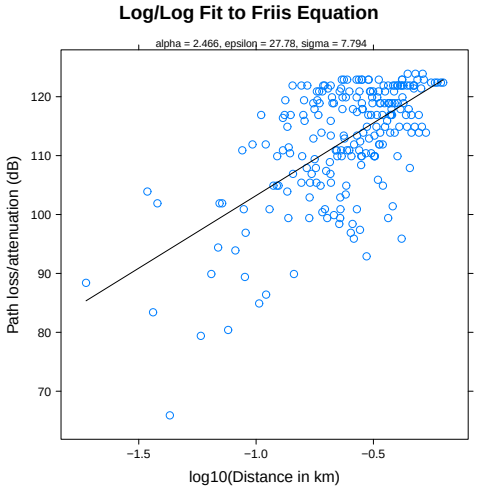
(f) h = 500 m

Figure 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

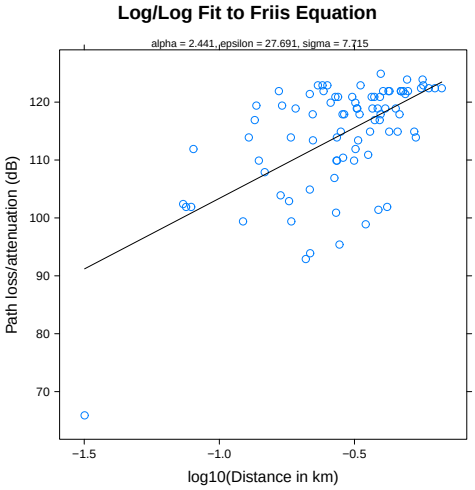


(a) Original Measurements

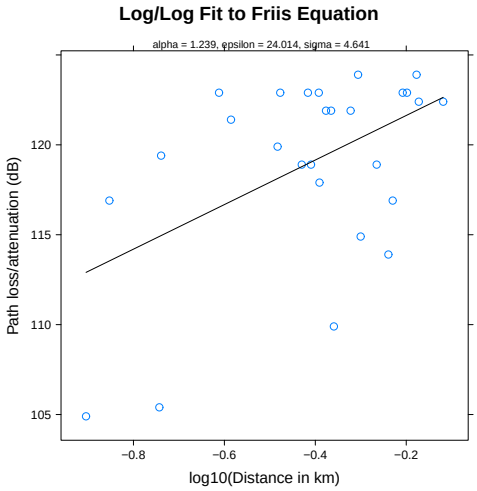
(b) h = 25 m



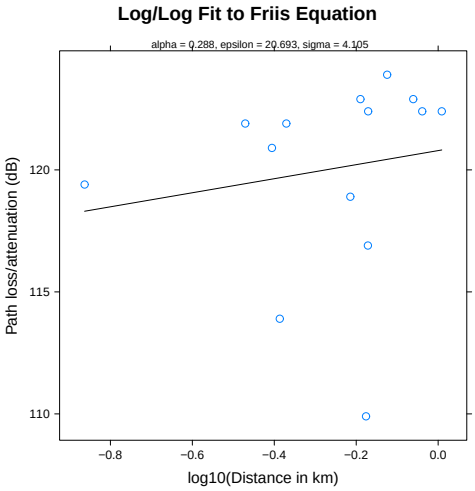
(c) h = 50 m



(d) h = 100 m

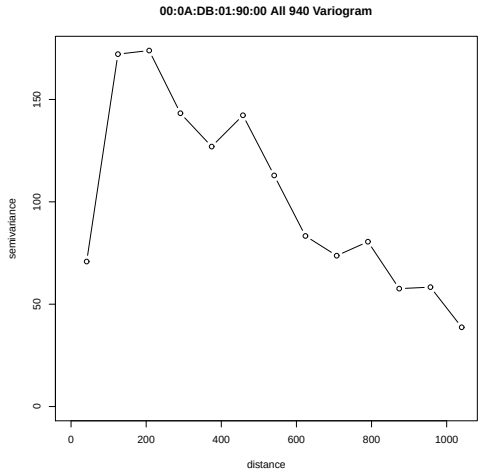


(e) h = 250 m

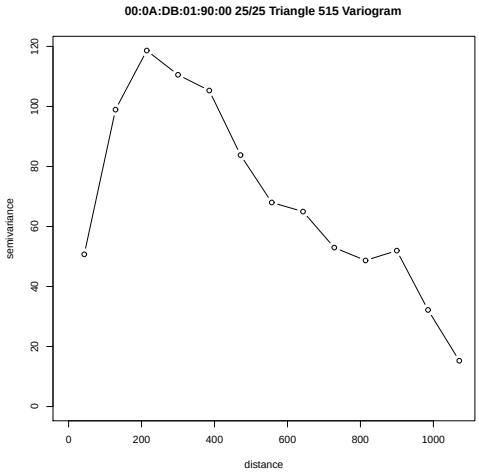


(f) h = 500 m

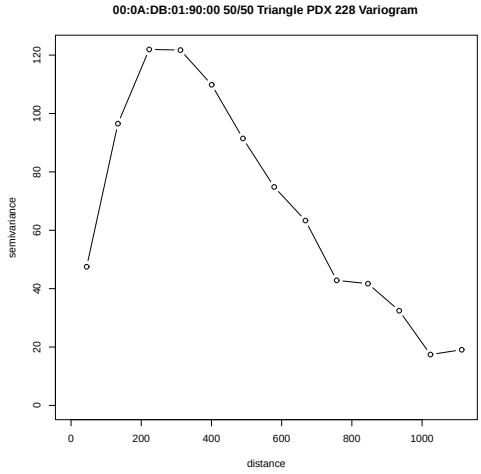
Figure 5.6: Path loss measurements for “pdx90” AP, both as-collected, and resampled at varying lag distances. Each figure plots the path loss as a function of distance on a log-log plot. A linear least squares regression fit line, and parameters are given.



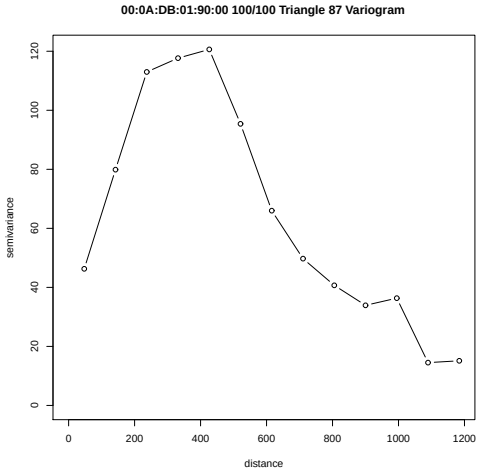
(a) Original Measurements



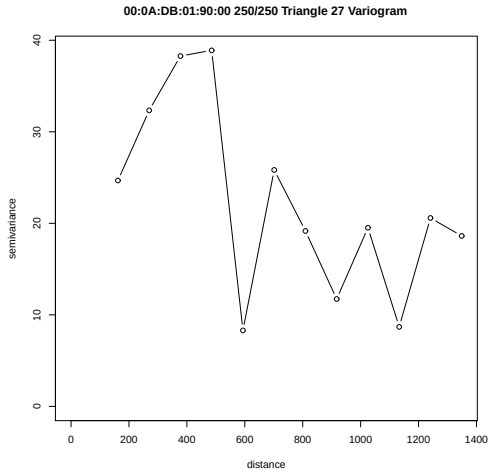
(b) h = 25 m



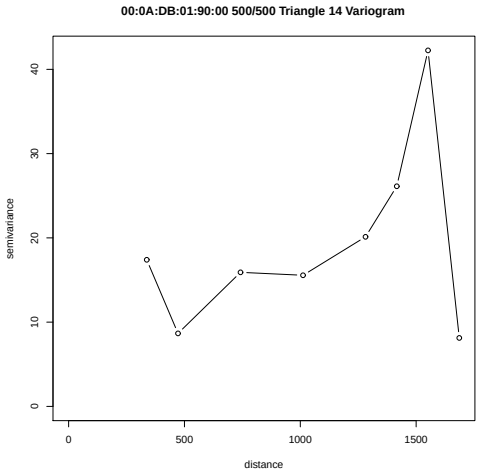
(c) h = 50 m



(d) h = 100 m



(e) h = 250 m



(f) h = 500 m

Figure 5.7: Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances.

where $Z'()$ is the residual (de-trended measurements) process, $Z()$ is the observed process and $P()$ is the model prediction. Chapter 3 showed that the best-case performance of *a priori* models, when tuned to their ideal parameters, was on the same scale as the residual error of a log-log fit to path loss using a small number of measurements. Given this, an empirical approach to modeling seems easy to advocate here. For the measurements from each access point, fitted α and ϵ values are determined so that the trend can be removed as follows:

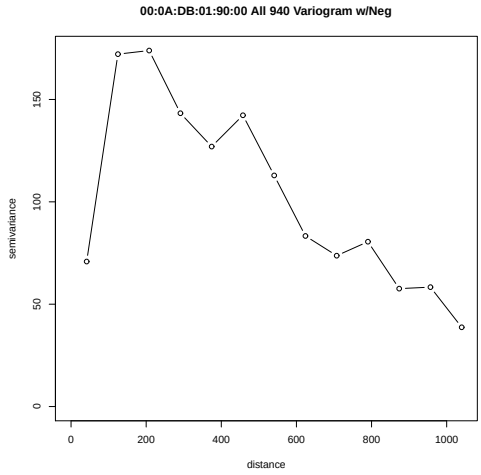
$$Z'(\mathbf{x}) = Z(\mathbf{x}) - (\alpha 10log)10(d) + 20log_{10}(f) + 32.45 + \epsilon) \tag{5.13}$$

the resulting de-trended observations can then be used to fit an empirical variogram as described in section 5.1.2. This is a pleasing approach to de-trending because it is entirely modular and extendable—the fitted log-log empirical model described here can be easily replaced with any other predictive model. In this way, the interpolation process can be viewed as careful way to correct for any remaining (environment-specific) model error, instead of as a complete replacement. As the state of the art in path loss modeling is advanced further, and models are able to make predictions closer to measurements, this improvement can be carried through to measurement-based interpolation in the process of de-trending as described here.

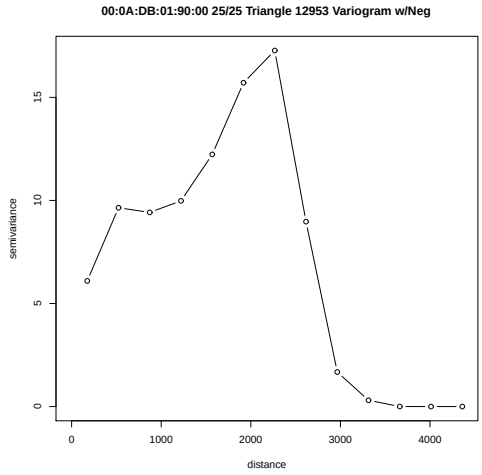
5.4.4 Variogram Fitting

Figures 5.8 and 5.9 show the resulting empirical variograms before and after de-trending the data for AP pdx90. The nugget tolerance (i.e., points within this distance are considered co-located) is set to 40λ per the classic averaging recommendations of Lee in [132]. As before, at larger lag distances the plotted variograms deviate wildly from ground truth. However, even using all measurements or a dense ($\leq 100m$) sampling, the variogram does not permit fitting with any of the permissible models described in [112]. In order to tame the variogram model, two more changes are necessary.

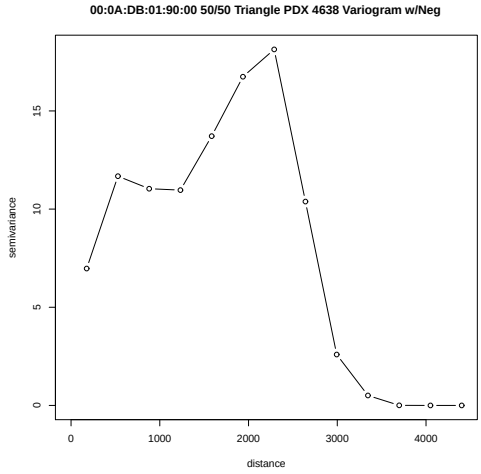
First, “negative results” must be included in the dataset: when a given point has been visited, and the instrument failed to observe a given AP at that point, it can then be inferred that the signal at that point must be lower than our receiver threshold. To assuage this, at points visited where there is no data for a given AP, a measurement with an unrealistically low value is used. The approximate noise floor (-95 dBm for most



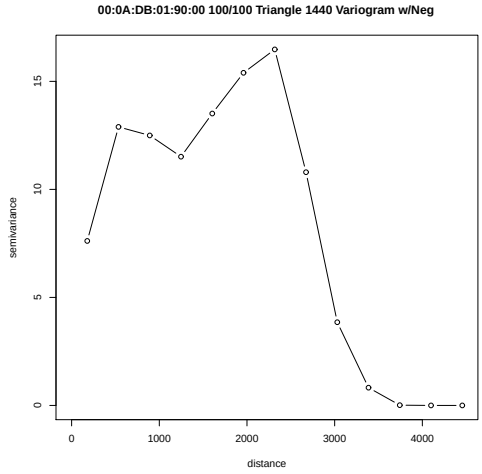
(a) Original Measurements



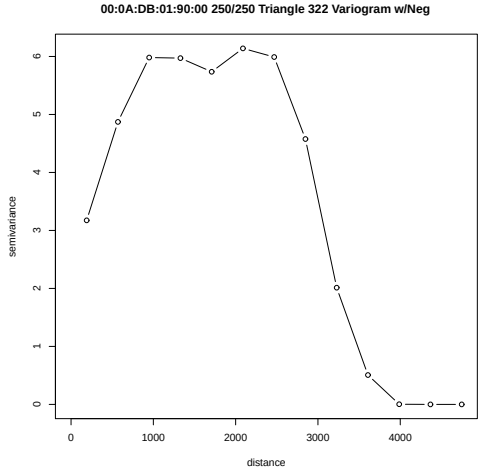
(b) h = 25 m



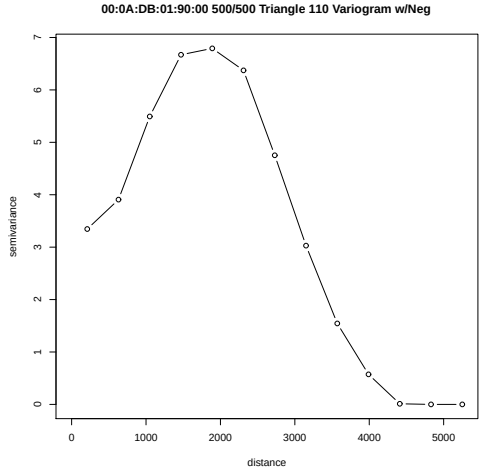
(c) h = 50 m



(d) h = 100 m

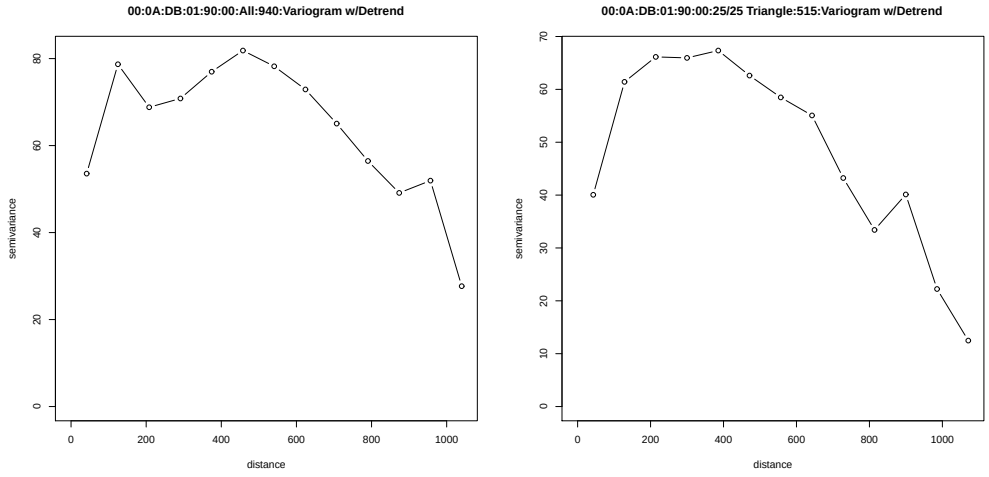


(e) h = 250 m



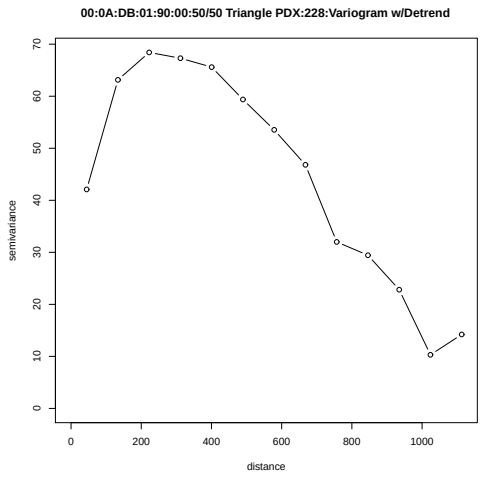
(f) h = 500 m

Figure 5.8: Empirical semivariograms of path loss for “pdx90” AP without de-trending.

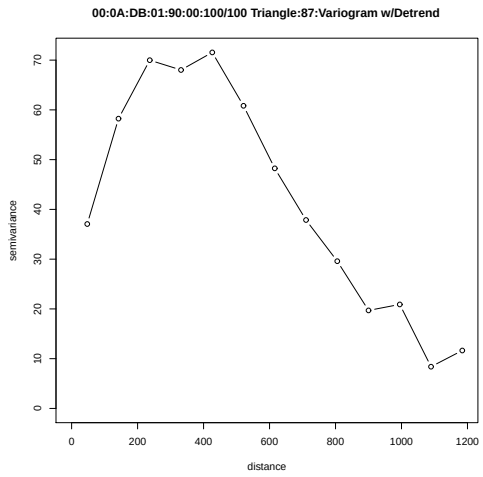


(a) Original Measurements

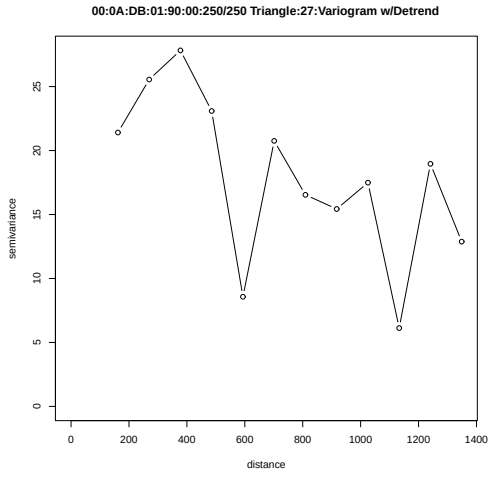
(b) h = 25 m



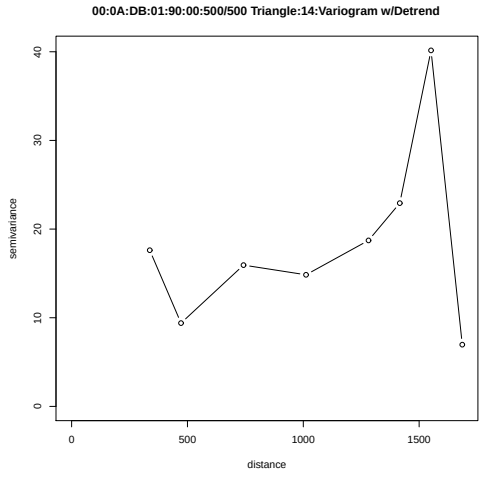
(c) h = 50 m



(d) h = 100 m



(e) h = 250 m



(f) h = 500 m

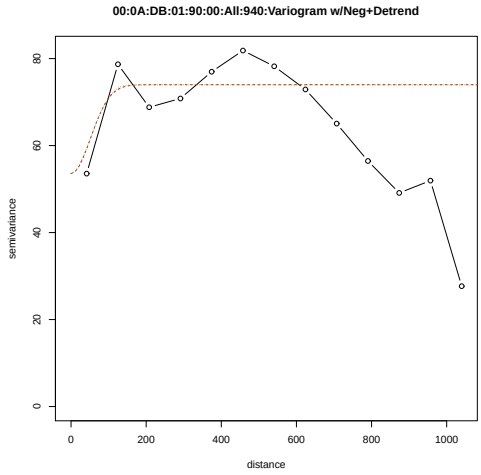
Figure 5.9: Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances with Frii’s freespace model used for de-trending.

transcievers at the frequency we're operating on) is a reasonable choice. This is an important extension: if a point has been visited, the lack of observation of a given AP carries information in itself. If no measurement is record at this point, there is substantial information loss. Instead, by recording a measurement at a very low value, the way the observed signal tapers towards the noise floor near the edge of each AP's coverage can be modeled explicitly. Figure 5.10 shows the resulting empirical semivariograms after this modification. These plots also show fit lines for Gaussian and cubic semivariogram models, which are poor.

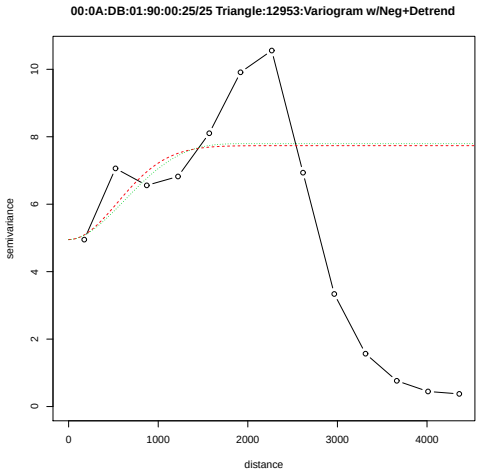
In large part the fits are poor because the empirical variogram models exhibit a large hump and then trail off afterward. This may be because observations at large lag distances are scarce because it is simply unlikely to observe a transmission further than say, 2 km, apart in a network of this kind. Hence, fits are truncated near the peak of the models and instead focus on fitting the portion of the curves where data is available and allow the remainder to lie under the asymptote. For this data set, several options were explored for truncation and it was found that truncating at $d = 800m$ worked well for all APs and nearly all resampling densities. Figure 5.11 shows the resulting empirical semivariograms and their fits with the cubic and Gaussian models. In general, cubic and Gaussian models fit well for nearly all datasets studied in this thesis. This is a pleasing result, since the cubic model is closely related to the spherical model, and the Gaussian model is closely related to the exponential model, both of which have been suggested as a reasonable choice in modeling the semivariograms of electromagnetic processes [119, 236]. Variogram fitting is accomplished with weighted least squares as described in [112] using the implementation available in the R package "geoR" [114].

5.4.5 Goodness of Fit

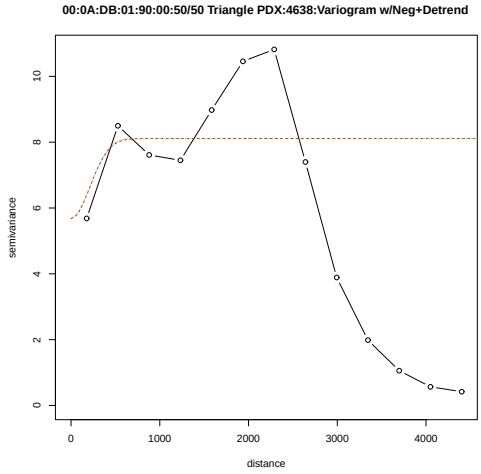
In the implementation used in the case studies below, this fitting process is automatic: fitting is performed using both the cubic and Gaussian models, with and without distance truncation, and with and without null measurements included. Then, the goodness of each fit is calculated using k-fold cross validation (typically with $k = 10$). For each fold a random sample of 20% of measurements is excluded, a fit is made with the remaining points, and the excluded points are predicted with the fit. The resulting RMSE and Mean Square Kriging Variance (MSKV) of these predictions are calculated and averaged across the k



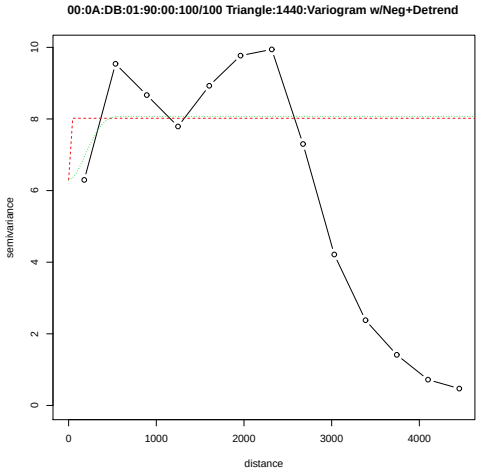
(a) Original Measurements



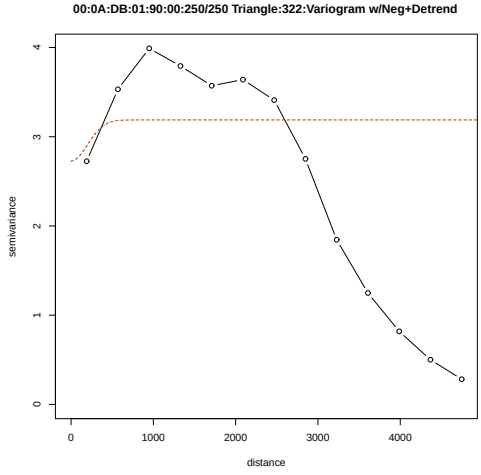
(b) h = 25 m



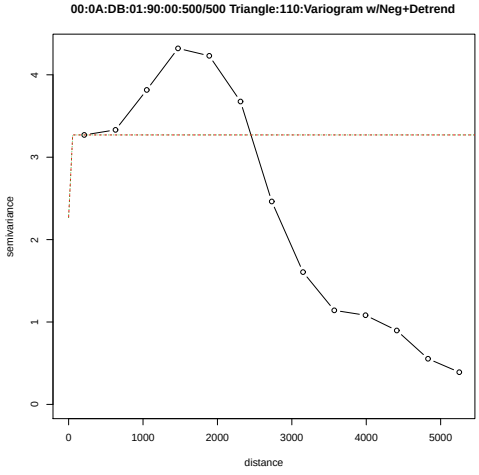
(c) h = 50 m



(d) h = 100 m



(e) h = 250 m



(f) h = 500 m

Figure 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.

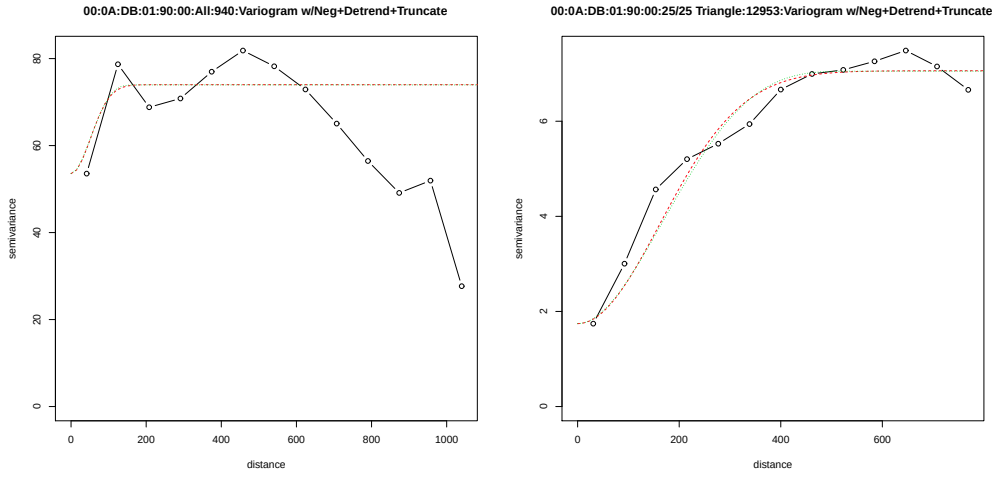
folds. The fit with the lowest mean RMSE is then used for Kriging. As a general rule of thumb, truncated fits that include null measurements are the best performing in our case-studies, so these are focussed on in the analysis. However, the split is fairly even between the Gaussian and cubic models.

5.4.6 Mapping with Ordinary Kriging

With fitted variogram models in hand, the next task is to generate a coverage map. This is a straightforward but computationally intense process that involves Kriging the value at each pixel in the mapping area. To accomplish this, Ordinary Kriging (OK) is applied as described in [119], using an implementation in the "geoR" R library [114].

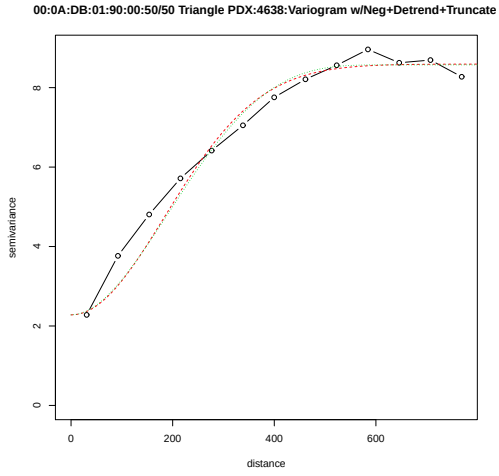
Two important questions that arise in mapping are: what resolution should be targeted, and what colormapping (visualization) strategy should be used? The computational demands of the process are intimately tied to the resolution, and maps of an arbitrarily fine resolution cannot be generated in a reasonable amount of time using basic Kriging methods (which, afterall, must solve a quadratic program of some size for each pixel!). In practice, a resolution of 0.05 Pixels per Meter (PPM) (one pixel for each 20x20m square) is good for quick estimates and a resolution of 0.2 PPM (one pixel for each 5x5m square) is the highest reasonable resolution. Because a nugget tolerance to 40 wavelengths (approximately 5m at 2.4 GHz) is used per the discussion in the previous section, it would not be meaningful to generate images at a higher resolution this. Calculating a map at the 0.2 PPM resolution can require a substantial amount of memory and fails for some maps with more data than the other resolutions. This stems from the fact that R requires its data structures to be allocated in contiguous blocks of memory and often a chunk of memory of the size required simply isn't available. For situations where a 0.2 PPM resolution map is intractable to generate, the 0.05 PPM resolution map can be used instead.

Using the de-trended measurements and fitted variogram as input, the Kriging process will produce a map matrix defining a map of excess path loss above or below the predictions of the fitted empirical model used for de-trending. These predictions are now added back to the Kriged map to create a map of the signal strength. Additionally, at each Kriged location the residual Kriging variance can be computed, which can be used to generate a map of residual uncertainty. Because Universal Transverse Mercator (UTM) coordinates

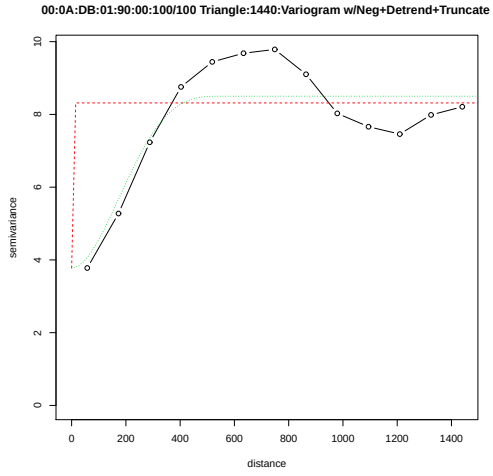


(a) Original Measurements

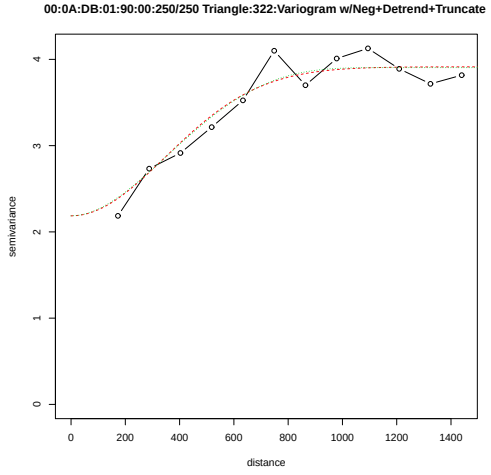
(b) h = 25 m



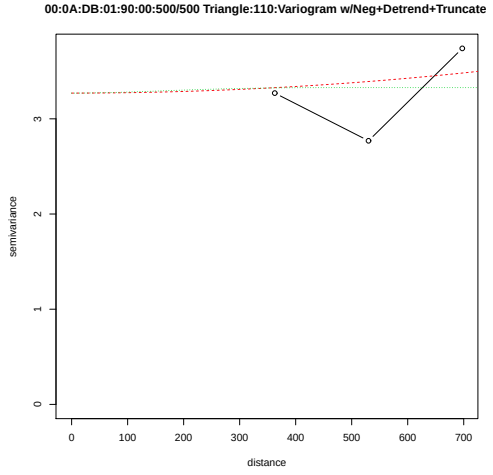
(c) h = 50 m



(d) h = 100 m



(e) h = 250 m



(f) h = 500 m

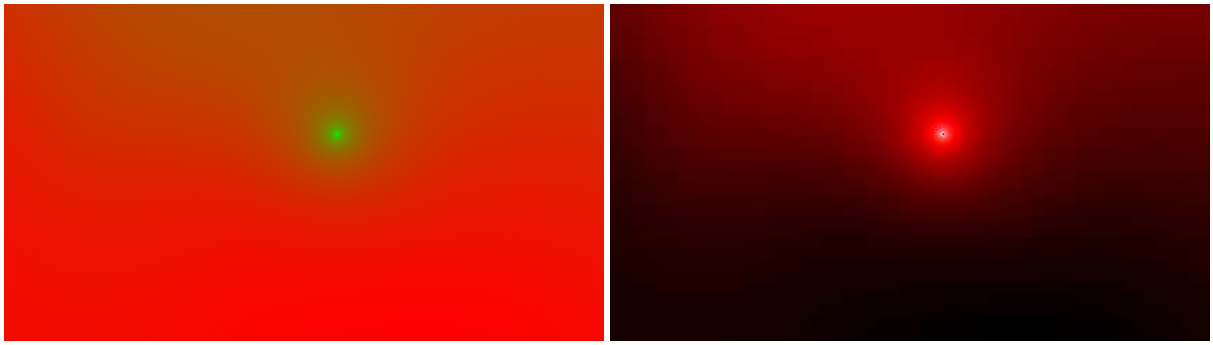
Figure 5.11: Empirical semivariograms of path loss for “pdx90” AP resampled at varying lag distances with Frii’s freespace model used for de-trending. Measurements at the noise floor have been inserted at points where an observation was unable to be made.

are used for all locations and mapping, the final map is simply a square meters-based grid with true north being upwards, as one would expect.

5.4.7 Visualization

These maps can produce different interpretations depending on the way values are combined, or even just through the color scheme. In [201], Rogowitz *et al.* show how choice of color map can dramatically effect the way data is interpreted and suggest careful choice of an appropriate color map for a given application that both varies luminance and saturation in addition to hue. Popular radio planning applications such as SPLAT! [136] and RadioMobile [54] color pixels on a map depending on the predicted SNR at that point. This color map can be overlayed on a terrain map to show their relationship (which is substantial because these tools use the ITM). In the planning tool proposed by Bartels *et al.* in [35], a logarithmic color encoding is used to show predicted coverage [152]. In [200], Robinson *et al.* take a simple domain-oriented visualization scheme and show coverage as a region of circles and coverage gaps (holes) as red x's. However, by and large there has been little substantial work studying visualization strategies for coverage maps. The maps in presented in this thesis take advantage of hue and luminance as suggested by [201], showing coverage on a color scale from (dark) black to (bright) red. By overlaying the luminance and saturation scale on the hue, contours not immediately visble in a simple hue-based (i.e., reg/green) interpolation can be readily seen. Figure 5.12 shows an example of the difference between these color maps using the example of WiMax coverage from a single BS on the CU campus.

A final example of how these maps might be used in practice is given in figures 5.13, and 5.14, which show Kriged maps here overlayed on Google Earth orthoimagery [82], Digital Terrain Model (DTM)s, and three-dimensional building models. In this example, the Kriged map is reprojected into the correct coordinate system and is placed at the correct coordinates in the Google Earth software using a Keyhole Markup Language (KML) file. An alpha channel is introduced to add partial transparency. Although a great deal of work could be devoted to domain-oriented visualization strategies for wireless coverage maps, simple interactive map overlays such as these, using widely available map tools like Google Earth, may be a strong first step in the direction of interactive mapping and planning tools for empirically derived wireless



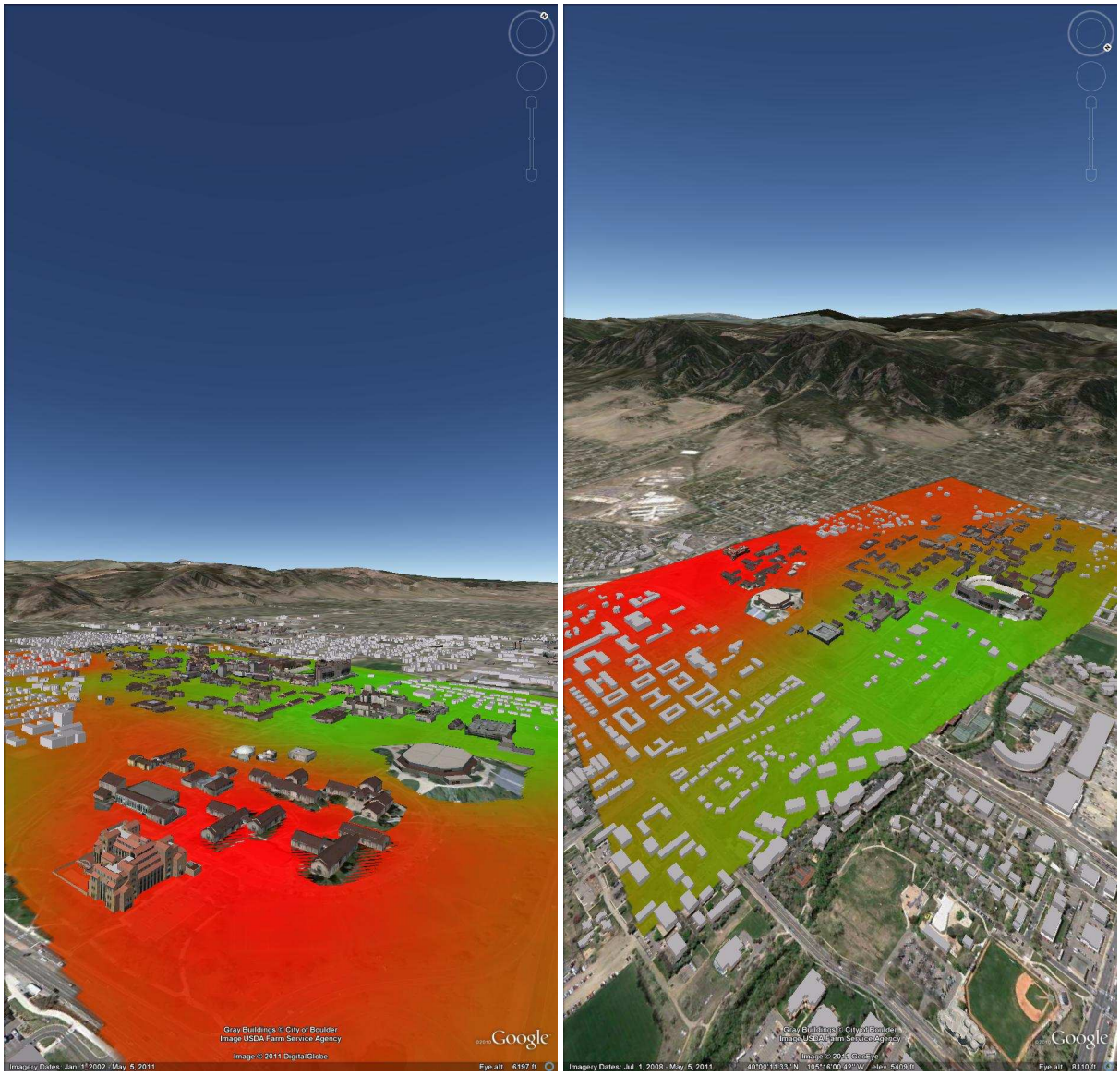
(a) Hue-Only Red/Green Map (b) Hue-Luminance Red/Black Map

Figure 5.12: Comparison of color maps using a Kriged map of WiMax CINR coverage for one BS at the University of Colorado.

coverage maps.

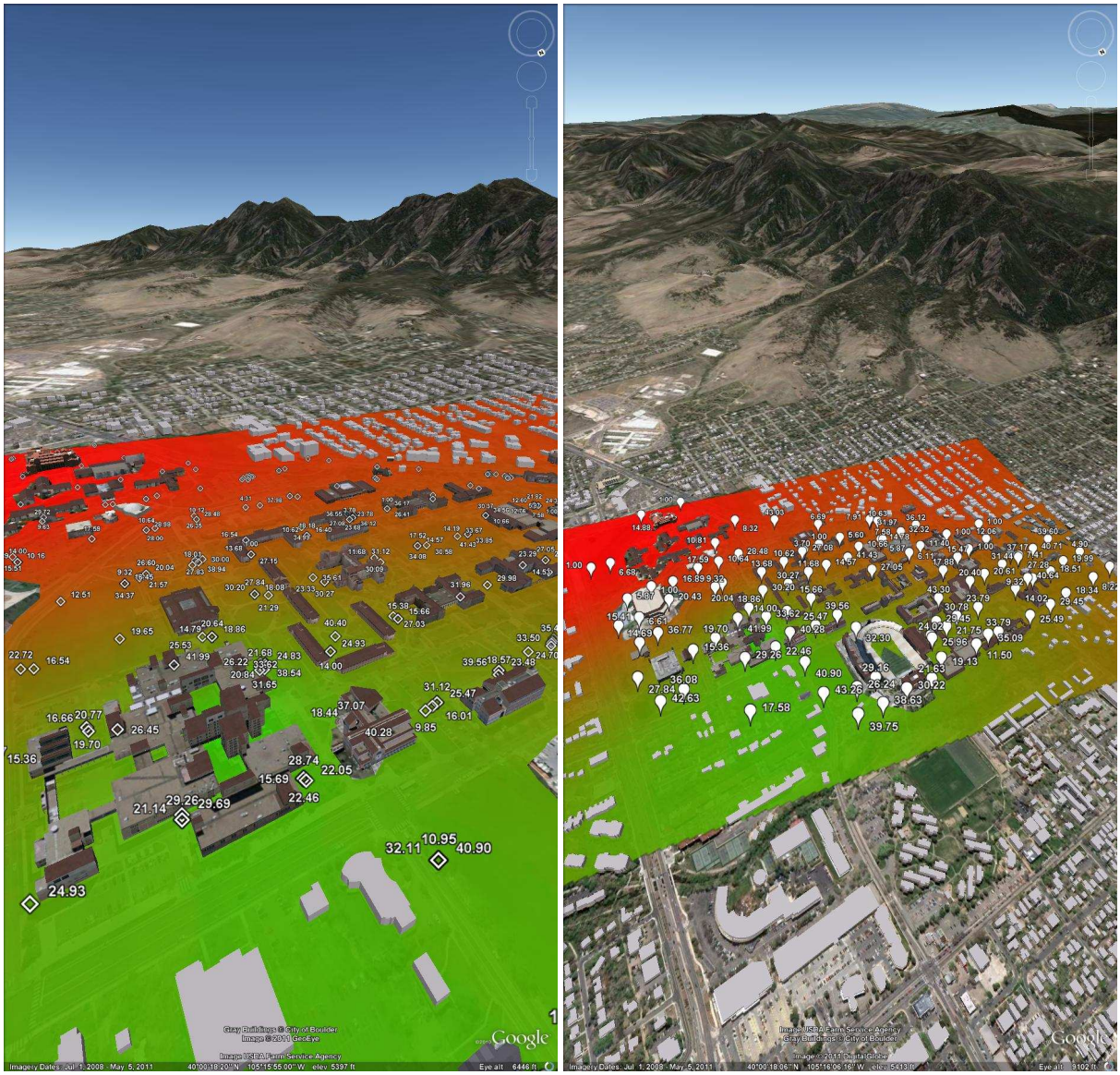
5.5 Summary and Conclusion

This chapter described the approach to geostatistical coverage mapping taken in this thesis, and practical adaptations made for the purpose of modeling the vagaries of RF propagation. Table 5.2 summarizes the best practices for geostatistical mapping derived in this thesis, and section C.5 provides source code implementing the core parts of the fitting and mapping functions. When performing variogram fitting, it was found that de-trending is necessary and can be accomplished by subtracting off predictions from a log/log fitted model. Null measurements can be included or not, using a constant low value where measurements could not be made. Variogram truncation is essential for a reasonable fit. The truncation value appears to be environment dependent, but largely consistent for APs in the same environment. Gaussian and cubic models perform well, with a nugget variance between zero and 40 wavelengths, and an imaging resolution of approximately 1 pixel per 5 meters (or 1 pixel per 20 when generating draft-quality maps). The approach described here utilized OK for Kriging and WLS for model fitting. However, Universal Kriging (UK) models, possibly utilizing anisotropic extensions are an interesting topic for future investigation. The next chapter will put the method proposed here to work in order to predict the coverage of production WiMax and LTE networks, as well as analyze the quantitative and qualitative efficacy of this approach to coverage mapping in the real world.



(a) (b)

Figure 5.13: Examples of coverage map (for CU WiMax cuEN node) overlayed on Google Earth orthoimagery, digital terrain, and 3D models.



(a) (b)

Figure 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.

Initial Sampling Design	100m triangular lattice with random clustered samples within 40 wavelengths
Second Phase Sampling	None. See chapter 8.
Unreachable Points	Take measurement at nearest accessible location avoiding systematic bias
Repeated Measurements	1-3 per (regular or cluster) location to model small scale temporal variation in isolation
Negative/Non Measurements	Include with constant “noise floor” value
Detrending	Subtract off Log/Log fit
Variogram Fitting	Weighted Least Squares (although MLE or Method Of Moments (MOM) are acceptable as well)
Variogram Model	Gaussian or cubic
Variogram Truncation	Yes, environment/technology dependent
Kriging	Ordinary Kriging (Universal Kriging is a topic for *Future Work*)
Anisotropic Modeling	None. A topic for *Future Work*.
Nugget Tolerance	0 (or up to 40 wavelengths to smooth out fast fading effects)
Prediction Resolution	0.05 (fast) - 0.2 (slow) pixels per meter

Table 5.2: Summary of derived best practices for geostatistical mapping of wireless network coverage.