Chapter 9

Conclusion

This thesis began with the claim that *there are not good methods for determining how well a given network works over a given area and presenting this information in a meaninful way*. As a possible solution to this problem, the application of geostatistical mapping methods were proposed, adapting mathematics developed for geological mining applications to a new and vastly different domain. Ultimately, it was found that this is a reasonable application, and the robust spatial statistical methods used in geostatistics allow for the creation of coverage maps that embrace, rather than ignore, the spatiotemporal variability of the wireless channel. In the case studies presented above, geostatistical approaches were shown to produce maps with a fine accuracy and much better predictive performance than standard *a priori* models that do not use measurements, or simple measurement-based fitting. However, to focus only on the performance improvement is to miss the real value of the geostatisitcal methods: by implementing an appropriate sampling design, modeling the underlying spatial structure of the data, and using a statistical method, an interpolated map can be generated with a well defined notion of residual error: the prediction at each point is a distribution, not simply a value. Additionally, this robust coverage map can be produced using a reasonably small amount of easily obtained data (several hundred samples 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 order to enable these results, new mechanisms for measurement were developed and paired with statistically safe sampling methodologies and interpolation techniques. In addition, careful attention was paid to the comparative value of performance metrics, so that the resulting maps are not only well-fitting to the data, but also communicate meaningful information about the real performance, and underlying variability of that performance, at the interpolated locations.

In addition to this core work, several important tangential threads were investigated. In particular, several extensions were evaluated that provide features that would be useful were the methods proposed here to be widely adopted. First, the prospect of resampling was investigated to understand how a coverage map might be derived from measurements collected at locations where it is convenient to collect, but with some substantial sampling bias (for instance, in city streets). It was found that resampling of this data can help to alleviate bias and that the resampled data can be well-modeled with geostatistical techniques. In fact, the resulting coverage maps are as accurate at predicting coverage holes as state-of-the-art iterative heuristic refinement methods (e.g., [200]), with a nearly identical number of measurements. This is an exciting result because this performance is obtained with the same amount of effort, while producing a substantially richer coverage map, where each interpolated point is a value distribution instead of a binary value. Next, the prospect of crowd-sourced coverage mapping was investigated, where many volunteers might cooperate to collect the measurements for a coverage map. It was found that this may be a feasible approach to coverage mapping, if a sufficient fraction of the population inhabiting the mapped location is willing to participate. However, a case study using data collected with a production crowd-sourcing system showed that in practice this level of participation may not yet be present. Finally, as a way of refining and tuning the generated coverage maps, sample optimization was proposed and investigated. It was found that metaheuristic approaches to sample optimization perform well, and that an insightful second-phase sample can be found in several hours computation on a single computer. These additional optimized samples provide valuable insight into the coverage of a given transmitter by identifying areas where variance is high, near coverage thresholds for the network. Although quantitative improvements in predictive performance were shown to be small in a case study, this iterative optimized sampling strategy shows promise in this domain and deserves further investigation.

As is typical of the scientific process, this work has also brought to light a number of areas where future work is needed:

- Better methods are needed for visualization of wireless coverage maps. The color mapping and interactive map overlaying described in section 5.4.7 are only a beginning in terms of visualization strategies. In particular new methods are needed to draw out contours and highlight holes and deviations. Mapping systems that perform dimension reduction to simultaneously communicate value and variance are most needed. One can imagine a network planning tool that provides for interactive mapping and surveying of a network region, while interleaving GIS sources and orthoimagery smoothly. The method presented here could be easily integrated into such a system and the map data adapted to any such visualization method, however substantial further work is needed in order to understand which visualization strategies work best, and in which situations.

- Accurate, fast, and inexpensive measurement hardware are needed. The spectrum analyzer, drive-test software, and UE radio devices used in this study all presented substantial shortcomings. The most accurate tools were also very slow, cumbersome, and expensive. Meanwhile, COTSE-based devices are faster, but they provide fewer options for measurement, are generally closed to low-level driver modification and analysis, and produce noisy results that prove difficult to fit. Developing better mobile measurement hardware that is open and modifiable, easy to use both by experts and technicians, and provides accurate and useful metrics, would be a huge boon to the coverage mapping problem. One can imagine a "smart" measurement device that collects and actively guides the measurement process using the methods described here. Developing similar sensor systems for long-term spectrum sensing and mapping deployment would also have tremendous value as cognitive and whitespaces networking gains traction.

- The work in section 7.2 on crowd-sourced coverage mapping demonstrations that this area is ripe for further work, as evidenced by the fact that companies like OSM have already deployed software to collect this data with smart phones. However, it is still not clear the best way to cope with the relative sparsity and noise that is an intrinsic component of crowd-sourced data collection. Substantial work is needed to understand the practical accuracy of measurement using common mobile UE hardware, how to collect sufficient data without effecting the battery life of small mobile devices, and whether a sufficient fraction of users would be available to collect data in the regions it is needed most. Successful crowd-sourcing data collection projects like the “Test my ISP” project by the FCC are encouraging [50].

- The geostatistical coverage mapping method developed in this thesis makes use of standard Kriging approaches, like OK and omnidirectional models. However, more advanced Kriging methods might make way for further modeling gains. For instance, anisotropic models may offer a way to more finely model the coverage of transmitters with directional antennas. These models segment the azimuthal plane and fit each segment with a possibly different geostatistical model (and/or method). Higher-powered Kriging approaches, such as UK, which allow for the mean of the map to be modeled by an arbitrary function, or local Kriging which more carefully considers the “neighborhood” of measurements around a given pixel when fitting it, may offer additional gains. However, initial experimentation in this direction suggested that these approaches may also be needlessly over-powered (or myopic in the case of local Kriging) for the application to coverage mapping.

- In this work it was assumed that all areas within the region of interest are equally valuable to map. In practice, this is seldom the case; typically some areas receive more use or are more important for mapping than others. Identifying domain-appropriate stratified sampling designs, which sample some areas more densely than others is an interesting topic for future work. A similar approach could be used to guide multi-phase sample optimization, extending the methods proposed in chapter 8.

- Section 3.7 provided an analysis of the accuracy of commerical raytracing systems when predicting the propagation in one environment. Although the initial goal of this experiment was to determine the relationship between input data fidelity and raytracing prediction accuracy, it was found that these systems are not able to deal with high-resolution environmental data. And, even moderately complex data may lead to issues with computing knife-edge diffractions, leading to compounding errors. In general, better raytracing algorithms are needed that can make use of high resolution environment data, be it created through crowd-sourcing using design tools like Google SketchUp

[14], or painstakingly collected using arial scanning LiDaR. Being able to utilize this data will require substantial advances in terms of meaningful vectorization of point data, preprocessing, and UTD computation, perhaps utilizing substantial parallelism. And, once such a system is developed, serious work is needed in terms of verifying the accuracy of these ray-tracing systems and understanding the fundamental relationship this accuracy has with the fidelity of input data.

- Although some work was done here to parallelize geostatistical computation (e.g., the parallel Kriging variance implementation in section C.1), these methods are still quite computationally complex. Some trivial parallelization is possible, by simultanously mapping measurements from multiple APs, but the underlying functions, particularly those involving large matrix operations, could still be optimized and parallelized substantially. In order for geostatistical mapping methods of the sort proposed here to be widely integrated into desktop planning software, or even a handheld measurement device, some work will be needed to do finer parallelization on, e.g., General Purpose Graphical Processing Unit (GPGPU) hardware.

- Finally, The methods described here have been limited in their analysis to outdoor microcell networks operating in the Ultra High Frequency (UHF) band. Determining the efficacy of the methods when applied to other types of networks at different frequencies is an important area for further work. These methods could also be trivially adapted to indoor or three-dimensional mapping (e.g., in multi-floor buildings or Unmanned Aerial Vehicle (UAV) applications), however substantial work will be needed to determine their practical accuracy in these settings and whether domain-specific modifications will be necessary.

In sum, this thesis has provided a complete and functional system for mapping the coverage of a production wireless network. Although the results here cannot be extrapolated to any networking techology in any environment, these results appear to hold promise for the broad application of geostatistical mapping to the RF environment. All told, the future appears bountiful for additional work in this area. It is the humble hope of this thesis that the work done here will help enable exciting new technological solutions to the problems faced with wireless networks; in effect, helping to transform them from an amazing technology that “seldom works as well as one would like”, to a technology that is at the same time reliable, ubiquitous, and essential.