Model-Based Interference Cartography and Visualization

P. N. Karthik, Raksha Ramakrishna, Geethu Joseph, Chandra R. Murthy, *Senior Member, IEEE*,
Joyson Sebastian, Neelesh B. Mehta, *Senior Member, IEEE*

***Abstract*—In this work, we present a tool to construct and visualize the spatio-temporal variations of power. A dataset of real-world power measurements is collected over a geographical area of interest. Relevant parameters of the environment such as the path loss exponent and the decorrelation time of the lognormal shadow fading are extracted from the dataset. Also, the average powers measured at a finite set of known locations are interpolated to obtain the average power distribution over the area. Using the parameters of the lognormal shadow fading, synthetic data with the same temporal behavior of the dataset is generated, and multiplied with the average power distribution. The resulting spatio-temporal power map is displayed on the screen through a graphical user interface developed in-house. The proposed approaches for interpolation and parameter extraction are validated using test datasets generated using the well-accepted modified Gudmundson model for the spatio-temporal correlation of lognormal shadow fading. We also undertake a comparative study of three different interpolation techniques: linear interpolation, inverse distance weighing and ordinary kriging. Further, we compare a model-based approach with a model-free approach for interpolation, and find that model-based ordinary kriging provides the best mean absolute percentage error performance.**

I. Introduction

Construction of spatio-temporal maps to visualize the variations in the radio frequency (RF) power over a geographical region of interest is referred to as spectrum cartography. An RF power map is used to enhance environment-awareness in a variety of applications like spectrum hole identification in cognitive radio networks, frequency planning and management, and spectrum policing. It requires one to learn about the RF propagation properties of the region from a (relatively) small number of power measurements at known locations and times, and use the information learned to predict the power patterns that might occur for any distribution of active transmitters in the area. A comprehensive study of various aspects of spectrum cartography is the objective of our work.

The concept of interference or spectrum cartography was introduced in [1], where the authors reconstruct a spatial map indicating the power levels over a geographical area by interpolating power measurements at a few known locations. Later, several RF map reconstruction algorithms have been studied in the literature. These algorithms are based on either a model-free or a model-based reconstruction approach using interpolation techniques such as kriging, thin splines and natural neighbor interpolation [2]–[4].

Researchers have also looked at the problem of mathematical modeling of RF propagation as a function of frequency and distance using experimental power measurements [5]–[8]. The modeling of path loss in an indoor environment through experimental power measurements is studied in [5]. A model for correlation in lognormal shadow fading is presented in [6] by comparing the measured and predicted correlation in the power values. Also, there are a few commercially available tools to visualize the power distribution patterns due to multiple transmitters operating over a region, e.g., [9]. The goal in these tools, in general, is to provide a (near) real-time view of the spectrum utilization over the area of interest.

While the studies in the literature investigate either the interpolation aspect or the mathematical modeling aspect of spectrum cartography, a collective study of these aspects does not exist, to the best of our knowledge. An integration of the parameter extraction, interpolation and reconstruction techniques on a single platform is the goal of this paper. The specific contributions of this paper are as follows:

- We develop a simulator that can incorporate real-world power measurements collected at known locations and times, to construct and visualize power maps for the region. The tool is capable of accepting parameters defining the radio propagation characteristics of the environment and constructing synthetic power maps for different configurations of the transmitters.
- We propose a procedure to obtain estimates of the aforementioned environment-dependent parameters (path loss exponent and the decorrelation time) using the real-world power measurements. We validate the procedure using test data generated using well-accepted spatio-temporally correlated lognormal shadow fading models.
- We interpolate the average power measurements at a finite set of known locations to obtain the distribution of the power over the operational area. We compare three methods for spatial interpolation: linear, inverse distance weighting, and ordinary kriging, by cross-validation using synthetically generated test data. We also contrast

The authors were with the Dept. of ECE, Indian Institute of Science, Bangalore 560-012, India, during the course of work (e-mails: {pnkarthik1992, raksharamakrishna13, gethuj, cmurthy1, joysonsebastian, neeleshbmehta}@gmail.com).

This work was financially supported in part by research grants from the Aerospace Networking Research Consortium (ANRC) and the Dept. of Electronics and Information Technology, Govt. of India. G. Joseph was supported by a fellowship from the Intel India PhD Fellowship Program.

978-1-5090-2361-5/16/$31.00 © 2016 IEEE

a model-free approach with a model-based approach, and find that the model-based ordinary kriging approach offers the best performance among the techniques.

As a by-product of the study, we obtain guidelines on the minimum number of measurements required to construct power maps of sufficient statistical accuracy for the indoor environment considered in this work. This inference is important because data collection is often the most laborious and time-consuming aspect of spectrum cartography.

The organization of the rest of the paper is as follows. In the next section, we describe the path loss and the lognormal shadow fading model employed for the model-based approach in this work. We then describe our interference visualization tool in detail in Section III. In Section IV, we present our proposed algorithms for extracting the relevant wireless channel parameters, and Section V describes the interpolation techniques and the validation of procedures developed. Numerical results are presented in Section VI, and some concluding remarks are given in Section VII.

II. Path Loss and Lognormal Shadowing Models

We assume that the received RF power $P(\mathbf{x}, t)$ in dBm at the location $\mathbf{x} \in \mathbb{R}^2$ at time t from access point (AP) is the sum of a path loss component $\text{PL}(\mathbf{x})$ and a shadow fading component $\chi(\mathbf{x}, t)$. The path loss component of the k^{th} AP is given by

$$\text{PL}_k(\mathbf{x}) = P_0 + 10\eta \log\left(\frac{d_0}{d_k(\mathbf{x})}\right), \; d_k \geq d_0, \qquad (1)$$

Here, $d_k(\mathbf{x})$ is the distance from location $\mathbf{x}$ to the k^{th} AP, P_0 is the transmit power of the AP in dBm, measured at a reference distance d_0 from the AP, and η is the path loss exponent. We assume that the shadow fading component in linear scale follows a lognormal distribution, and, hence, its value in dB is Gaussian distributed, i.e., $\chi(\mathbf{x}, t) \sim \mathcal{N}(0, \sigma^2)$, where σ^2 is the shadow fading variance in dB^2. Further, $\chi(\mathbf{x}, t)$ follows the well-accepted model for the spatio-temporal correlation of [7], [8]:

$$\mathbb{E}\left[\chi(\mathbf{x}, t)\chi(\mathbf{x} + \delta\mathbf{x}, t + \delta t)\right] = \sigma^2 \exp\left\{-\left(\frac{\|\delta\mathbf{x}\|}{d_{\text{corr}}} + \frac{|\delta t|}{t_{\text{corr}}}\right)\right\}, \qquad (2)$$

where $\|\delta\mathbf{x}\|$ is the Euclidean distance between the locations $\mathbf{x}$ and $\mathbf{x}+\delta\mathbf{x}$. The parameters d_{corr} and t_{corr} that specify the model are known as the decorrelation distance and decorrelation time, respectively.

III. The Interference Visualizer Tool

The high-level architecture of the interference visualizer tool developed in this work is shown in Figure 1. It comprises four components: the real-world dataset, the data preprocessing module which generates the interference interpolation files, the interference calculation module which fuses the interpolation files which is synthetically generated data in the form of files containing lognormal shadow fading, and the graphical user interface (GUI). The real-world dataset is comprised of received signal strength measurements collected over a floor of an office building. The data preprocessing module extracts the path loss exponent and decorrelation time from the dataset and generates data files containing interpolated power values. The interference calculation module uses the output of the data preprocessing module and synthetically generated lognormal shadow fading power values to compute the spatio-temporal variations in the power values over the region. These files are input to the interference calculation module, which displays the power maps on the screen through the GUI. We describe each of these components in detail below.

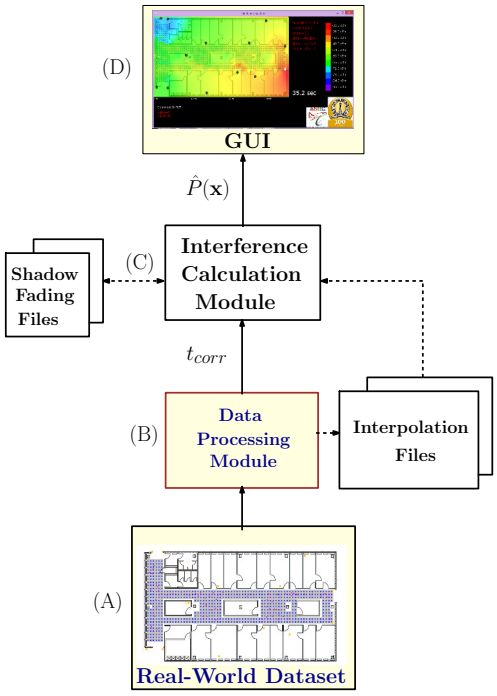

Fig. 1. Architecture of the interference visualizer tool comprising: (A) Real-world dataset, (B) data processing module that generates interpolated path loss values and t_{corr}, (C) interference calculation module that generates files of lognormal distributed shadow fading using t_{corr} and hence power map $\hat{P}(\mathbf{x})$, and (D) graphical user interface that displays the power map $\hat{P}(\mathbf{x})$.

A. Real-world Dataset

In this work, we use the Crawdad dataset [10], which contains data records in the form $(\mathbf{x}_i, t_j, P_{ij})$, for $i = \{1, 2, \ldots, N\}$ and $j = \{1, 2, \ldots, T\}$. Here $\mathbf{x}_i \in \mathbb{R}^2$ represents a location, t_j represents the timestamp at which the measurement is recorded, and P_{ij} is the corresponding measurement. In this dataset, there are $M = 14$ access points, and the data of $(\mathbf{x}_i, t_j, P_{ij})$ exists for each access point. At each location $\mathbf{x}_i$, there are $T = 110$ power values collected over time. After an interval of a few minutes, 110 power values are collected at the next location. Thus, the data collected at different locations are several minutes apart in time. Such measurements are acquired at $N = 580$ locations over an indoor region measuring 36 m $\times$ 15 m. Figure 2 shows a schematic of the floor plan of the office building where the power measurements are recorded. We note that although we

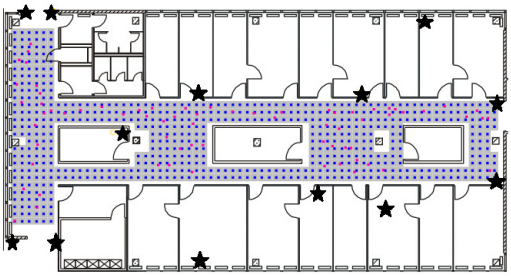

Fig. 2. Schematic of the floor plan from the dataset in [10]. Dots in the gray area indicate $N = 580$ measurement locations and stars mark the locations of the access points.

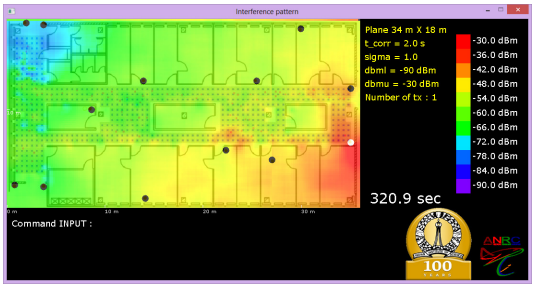

Fig. 3. A snapshot of the simulator window when a single AP is on. The active AP is shown in white; the inactive APs are shown as black dots.

use real-world power measurements collected in an indoor environment, our analyses and inference techniques are applicable to other environments as well.

B. Data Preprocessing Module

The data preprocessing module first computes the time-averaged power value at each location using the dataset. Then, it extracts the parameters of the wireless environment such as the path loss exponent and the decorrelation time. The parameter extraction procedures are described in Section IV. Further, the module interpolates the power values to estimate the distribution of received power values over the region. The three techniques for interpolation considered in this work are described in Section V. The decorrelation time and interpolated values are used by the interference calculation module to generate the power map.

C. Interference Calculation Module

The interference calculation module uses the spatial variation of the power as the baseline power at each location, which is computed in the data preprocessing module. The average power at a location is given by the sum of the interpolated power values in linear scale from all the active APs. The temporal variations in the power values are synthetically generated using the value of t_{corr} extracted from the dataset, following the modified Gudmundson model in (2), by passing temporally white shadow fading values through a filter programmed to achieve the required correlation between the entries. The power at each location is calculated as the sum of the average power value in dBm and the lognormal shadow fading power in dBm at the corresponding location.

D. Graphical User Interface

The front-end of the tool is a window that displays the spatio-temporal variations in the received signal power, and a side panel that provides information pertaining to the region for which the power map is displayed. Figure 3 shows a snapshot of the simulator window. The total power value at each location (in dBm) is fed to the GUI, which displays the power map on the screen by assigning appropriate colors to the individual power values based on a user-selected color scale.

The visualizer tool also has the flexibility to switch from the display of power maps using the real-world measurements to synthetic power maps simulating various environments characterized by different values of decorrelation time t_{corr}, shadow fading variance σ^2, and path loss exponent η. For this purpose, it has the provision to accept user inputs to change the values of these parameters through the command input block that is located at the bottom of the simulator window.

IV. Extraction of Wireless Channel Parameters

This section describes the estimation of the path loss exponent (η) and the decorrelation time (t_{corr}) of the lognormal shadow fading, using the dataset. As mentioned earlier, the dataset in [10] contains T measurements recorded over time at each location for each of M access points. The procedure for parameter extraction described below is for power values corresponding to a single access point, and the average of the estimates obtained from M sets of power measurements gives the estimate of η and t_{corr}.

A. Extraction of the Path Loss Exponent

We now present a procedure to obtain a minimum mean squared error (MMSE) estimate of η from the dataset corresponding to each access point. We first compute the mean of the power values at each location to average out the temporal variations in the received power. We model the resulting power values as the path loss component according to (1) contaminated by an additive error term z in dB, as follows:

$$\tilde{P}(\mathbf{x}_i) = P_0 + 10\eta \log\left(\frac{d_0}{d(\mathbf{x}_i)}\right) + z, \tag{3}$$

where $\tilde{P}(\mathbf{x}_i)$ is the time-averaged power value, in dBm and $d(\mathbf{x}_i)$ is the distance between the AP and the location $\mathbf{x}_i$, $i = 1, 2, \ldots, N$. To keep the notations simple, we omit the index corresponding to the access point. We choose d_0 as the distance between the AP and its nearest measurement location. The time average of power measurement at d_0 is taken as P_0. From (3), we estimate the variance ν^2 of the error term z as

$$\nu^2 = \frac{1}{N}\sum_{i=1}^{N}\left[\tilde{P}(\mathbf{x}_i) - P_0 - 10\eta\log\left(\frac{d_0}{d(\mathbf{x}_i)}\right)\right]^2.$$

We estimate η so as to minimize ν^2, i.e., we solve $\arg\min_\eta \nu^2$. This leads to the following MMSE estimate $\hat{\eta}$ of η:

$$\hat{\eta} = \frac{\sum_{i=1}^{N}\left(P_0 - \tilde{P}(\mathbf{x}_i)\right)\left(10\log\left(\frac{d(\mathbf{x}_i)}{d_0}\right)\right)}{\sum_{i=1}^{N}\left(10\log\left(\frac{d(\mathbf{x}_i)}{d_0}\right)\right)^2}.$$

B. Extraction of Decorrelation Time

This subsection deals with the procedure to obtain an MMSE estimate of t_{corr} from the dataset for each access point. We model the temporal correlation from (2) as:

$$\mathbb{E}[\chi(\mathbf{x},t)\chi(\mathbf{x},t+\delta t)] = \sigma^2 e^{-\left(\frac{|\delta t|}{t_{\text{corr}}}\right)}, \tag{4}$$

for $i = 1, 2, \ldots, N$. In (4), we vary δt to include each of $T = 110$ samples of lognormal shadow fading power available at a location. However, not just the lognormal shadow fading power but also the path loss contributes to the available received power measurements. We proceed by noting that

$$\begin{aligned}\mathbb{E}\left[\left(P(\mathbf{x},t) - P(\mathbf{x},t+\delta t)\right)^2\right] &= \mathbb{E}\left[\left(\chi(\mathbf{x},t) - \chi(\mathbf{x},t+\delta t)\right)^2\right] \\ &= 2\sigma^2\left(1 - e^{-\left(\frac{|\delta t|}{t_{\text{corr}}}\right)}\right),\end{aligned}$$

for $i = 1, 2, \ldots, N$. Let $C_{\delta t}$ be the empirical estimate of $\mathbb{E}\left[\left(P(\mathbf{x},t) - P(\mathbf{x},t+\delta t)\right)^2\right]$. Let $\mathcal{S}$ represent the set of all values of δt in the dataset. As before, we solve for the MMSE estimate $\hat{t}_{\text{corr}}$ of the decorrelation time, which is given by:

$$\hat{t}_{\text{corr}} = \frac{1}{|\mathcal{S}|}\sum_{\delta t \in \mathcal{S}} \frac{|\delta t|}{\ln\left(\frac{2\sigma^2}{2\sigma^2 - C_{\delta t}}\right)}.$$

To compute the estimate $\hat{t}_{\text{corr}}$ above, we need the value of the lognormal shadow fading variance, σ^2. This, in turn, is estimated from the Crawdad dataset. A natural approach is to first estimate the shadow fading values by computing the average received power at each location and removing the average value from the dataset. Then, one can simply regard the dataset as consisting of the shadow fading values, and compute the sample variance of the dataset, denoted by $\hat{\sigma}^2$, as an estimate of σ^2. However, upon using this procedure on the synthetic dataset with a given value of σ^2, we found that this procedure marginally under-estimates the value of σ^2, and that using a simple linear function of the estimate, as $1.06\hat{\sigma}^2 + 0.37$, yields better results. We follow the same procedure on the Crawdad dataset also to estimate σ^2.

Remark: The estimate t_{corr} from the dataset is of the order of a few seconds, while measurements at different locations are taken several minutes apart in time. Hence, from (2), the received power values at different locations are virtually completely decorrelated in time. Due to this, it is infeasible to use the dataset to estimate d_{corr}, as measurements at different locations at the same or closely spaced times are not available.

V. Interpolation and Sub-Sampling

In this section, we describe the techniques to reconstruct the spatial spectral map by interpolating the time-averaged values of power at each location denoted by $\tilde{P}(\mathbf{x}_i), i = 1, 2, \ldots, N$. We consider three interpolation techniques: linear interpolation, inverse distance weighting and ordinary kriging. These spatial interpolation techniques are popular in geostatistics [11], and have been considered for spectrum cartography in [1], [4], and [2]. We also explain the process of sub-sampling the dataset to obtain insights on the minimum number of measurements required to construct power maps for the region considered in the case study presented here.

A. Linear Interpolation

This is the simplest technique for interpolation, where the power at a given location $\mathbf{x}_0$ is obtained by linearly interpolating the values at nearby locations where power values are known. Using only nearby locations for interpolation is appropriate here, because the linear model is only accurate in the neighborhood of a given location. At the very least, we need to use the power values at the three nearest neighbors for linear interpolation (LI). However, the three nearest neighbors may not lead to a full rank system of linear regression equations, since measurements are taken at grid locations (for example, the nearest neighbors may lie on a straight line). In case the system is rank deficient, we use an iterative procedure, where we start with the three nearest neighbors, and add farther neighbors until we get a full rank system of equations.

B. Inverse Distance Weighting

In inverse distance weighting (IDW), the power measured at each location is assigned a weight according to its distance from the location where power is to be estimated, as follows:

$$w_i = \frac{1/\|\mathbf{x}_i - \mathbf{x}_0\|^k}{\sum_{i=1}^{N} 1/\|\mathbf{x}_i - \mathbf{x}_0\|^k},$$

where w_i is the weight assigned to the power value P_i measured at the i^{th} location, $\|\mathbf{x}_i - \mathbf{x}_0\|$ is the distance measured from the i^{th} location to the location $\mathbf{x}_0$ where power is to be estimated, N is the number of locations where power samples are collected and the exponent k is a constant. We use $k = 4$. The estimate $\hat{P}(\mathbf{x}_0)$ of the power at $\mathbf{x}_0$ is evaluated as a weighted linear combination of the known power samples: $\hat{P}(\mathbf{x}_0) = \sum_{i=1}^{N} w_i \tilde{P}(\mathbf{x_i})$.

C. Ordinary Kriging

In ordinary kriging (OK), an estimate of power value $\hat{P}(\mathbf{x}_0)$ at the location $\mathbf{x}_0$ is expressed as a convex combination of known values $\tilde{P}(\mathbf{x}_i),\ i = 1, 2, \ldots, N$ [12]. Thus,

$$\hat{P}(\mathbf{x}_0) = \sum_{i=1}^{N} w_i \tilde{P}(\mathbf{x}_i), \tag{5}$$

such that $\sum_{i=1}^{N} w_i = 1$. To obtain the weights w_i, an empirical variogram is computed to which a theoretical variogram is fit. Let $\mathcal{N}(h)$ denote the set of all pairs of locations separated by distance h. The empirical variogram $\gamma(h)$ is constructed as

$$\gamma(h) = \frac{1}{2|\mathcal{N}(h)|}\sum_{i,j \in \mathcal{N}(h)} (\tilde{P}(\mathbf{x}_i) - \tilde{P}(\mathbf{x}_j))^2,$$

where $|\mathcal{N}(h)|$ denotes number of elements in the set $\mathcal{N}(h)$.

Among the many possible choices for the theoretical variogram, we fit the Gaussian model that describes correlation as

an exponentially decreasing function of h [4]. The theoretical variogram is given as:

$$\bar{\gamma}(h) = c\left(1 - e^{-\frac{h^2}{a^2}}\right). \tag{6}$$

The constants c and a in (6) are referred to as the sill and the range of the theoretical variogram and are determined from the curve that best fits the empirical variogram constructed using the power values from the dataset. By using the theoretical variogram, the weights w_i in (5) are obtained by minimizing error variance of the power values with the constraint $\sum_{i=1}^{N} w_i = 1$. The error variance is obtained using (5) and (6) as

$$\begin{aligned}\text{var}(\hat{P}(\mathbf{x}_0) - P(\mathbf{x}_0)) &= \text{var}\left(\sum_{i=1}^{N} w_i P(\mathbf{x}_i) - P(\mathbf{x}_0)\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \bar{\gamma}\left(\|\mathbf{x}_j - \mathbf{x}_i\|\right) + \bar{\gamma}(0) \\ &\quad - 2\sum_{i=1}^{N} w_i \bar{\gamma}\left(\|\mathbf{x}_i - \mathbf{x}_0\|\right).\end{aligned}$$

This optimization problem can be solved using the Lagrange multiplier method. Differentiating the objective function of the unconstrained optimization gives the following system of linear equations:

$$\sum_{j=1}^{N} w_i \bar{\gamma}\left(\|\mathbf{x}_j - \mathbf{x}_i\|\right) + \mu = \bar{\gamma}\left(\|\mathbf{x}_i - \mathbf{x}_0\|\right), i = 1, 2, \ldots, N$$

where μ is a Lagrange multiplier. We solve for the $N+1$ unknowns including w_i, $i = 1, 2, \ldots, N$ and μ using the above N linear equations and the constraint $\sum_{i=1}^{N} w_i = 1$.

D. Model-Free vs. Model-Based Performance Evaluation

In order to evaluate the performance of the aforementioned interpolation techniques, we take two approaches: a model-free approach, and a model-based approach. In the model-free approach, we directly interpolate average power values in linear scale using LI and IDW, and in dBm using OK. In the model-based approach, we assume that the path loss component of power values available in the dataset follow the model in (1), and extract the lognormal shadow fading power values from the dataset by using the estimated path loss exponent along with the knowledge of the locations of the APs. Subsequently, we interpolate the lognormal shadow fading power values in dB scale using LI, IDW and OK.

E. Sub-Sampling

When the number of power measurements that can be collected is limited by the availability of resources, it is of practical interest to determine the minimum number of measurements that need to be used in order to construct power maps that are statistically faithful to those constructed using a larger number of measurements. In this work, we first divide the dataset into two subsets of equal size and designate them as the training and test datasets. From the training set, we pick various fractions of power measurements and obtain interpolated power values corresponding to the locations in the test dataset. The metric used to compare the interpolated and true power values is mean absolute percentage error (MAPE):

$$\text{MAPE} = \left(\frac{1}{|\mathcal{A}|}\sum_{i\in\mathcal{A}} \text{abs}\left(\frac{\hat{P}_i - P_i}{P_i}\right)\right) \times 100,$$

where $\mathcal{A}$ represents the test dataset, $|.|$ denotes the cardinality of the set, abs(.) denotes the absolute value, and $\hat{P}_i$ and P_i are the estimated and true values of power in dBm respectively. The numerical results of our sub-sampling experiments are presented in Section VI.

VI. Numerical Results

In this section, we present numerical results for validating the parameter extraction procedure proposed in Section IV. We also present a comparative study of the relative performance of the three interpolation techniques discussed in Section V. Thirdly, we study the effect of sub-sampling the dataset of measurements in the construction of the power maps.

A. Parameter Extraction Results

To validate the procedures on parameter extraction, a synthetically generated dataset similar to the Crawdad dataset is generated by assuming various values for η and t_{corr}. The parameter extraction procedures are applied on this synthetic dataset to obtain the estimates of η and t_{corr}, and the obtained values are compared with the values used for generating the dataset using the MAPE metric.

The following values are assumed for generating the artificial dataset comprising 580 locations identical to those in the Crawdad dataset and 10 time instants at each location: $\eta = 3.0$, $t_{\text{corr}} = 0.5$s and $\sigma^2 = 10$. The estimate of η obtained from the artificial dataset is found to be 2.9 with a mean absolute percentage error of 4.71% and that of t_{corr} is found to be 0.46 s with a mean absolute percentage error of 17.42%.

We proceed to apply the extraction procedures on the Crawdad dataset to extract η and t_{corr} corresponding to the indoor environment where the measurements were recorded. The estimate of η obtained from the Crawdad dataset is found to be 1.79, which lies within the nominal range of η reported in [13] for a single floor of an office building. The estimate of t_{corr} is found to be 4.26 s.

B. Interpolation and Sub-Sampling Results

In this subsection, the effect of choosing a subset of measurements from the dataset to construct power maps is investigated with respect to the number of measurements chosen and the interpolation technique used. Figure 4 shows the power maps constructed using 2%, 25%, 50% , and 100% of the measurements, respectively, from the dataset. An increase in the smoothness of the maps is observed as the number of points used to construct the maps is increased. Further, the improvement in the smoothness of the maps constructed using 25% or higher number of power measurements from the dataset is observed to be only marginal. Thus, the maps

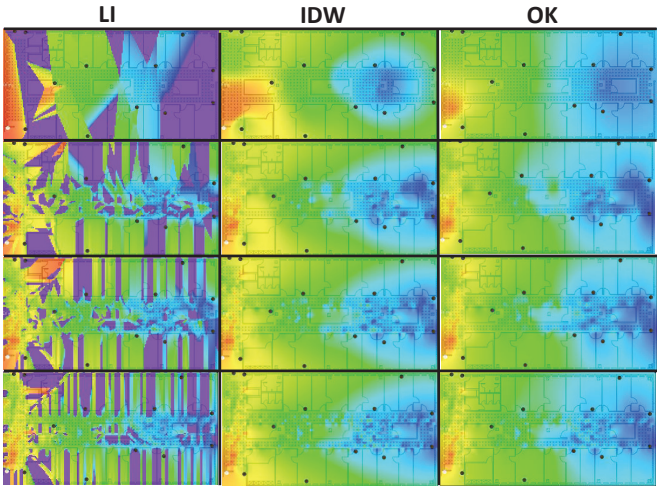

Fig. 4. Snapshots of the simulator comparing maps constructed using a subset of measurements from the dataset by employing linear interpolation, IDW and kriging. The maps in row 1 are constructed using 2% of measurements from the dataset, those in row 2 use 25%, those in row 3 use 50%, and those in row 4 use the entire dataset.

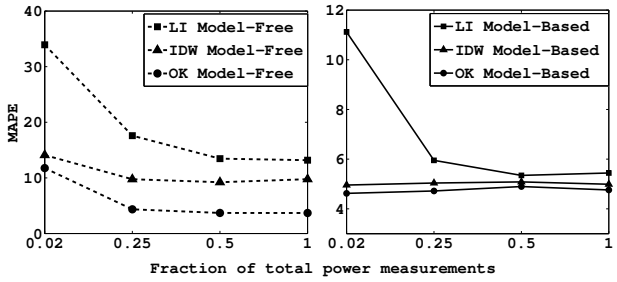

Fig. 5. Performance evaluation of interpolation techniques with respect to the fraction of power measurements chosen from the training dataset of 290 power measurements to construct power maps.

constructed using 25% of power measurements provide very nearly the same visual information as those constructed using all the power measurements from the dataset.

Finally, Figure 5 depicts the MAPE performances of LI, IDW and OK by using a subset of measurements from the dataset, where the MAPE is evaluated using the values in the synthetic dataset for each of the techniques, for the model-free and the model-based approaches. This figure serves as a guideline to pick the number of power measurements to be used for constructing power maps satisfying desired tolerable error values. When 2% of the power measurements are used to construct maps, OK is observed to perform the best for both the approaches. However, when 25% or higher number of measurements are used, the performance of OK is observed to be marginally better for the model-free approach than the model-based approach. Further, a reduction of 59% and 49% in error is observed in the performances of LI and IDW, respectively, by taking the model-based approach rather than the model-free approach. Hence, we see that the model-based approach in conjunction with OK offers the best fidelity in reconstruction among the techniques compared. We note that the guidelines given are based on the experimental results. The theoretical analysis showing the dependency of the number of measurements required on the environment-dependent parameters and error tolerance requires additional investigation and is beyond the scope of the paper.

VII. Conclusion

We developed a new tool to construct and visualize power maps depicting the spatio-temporal variations in RF power received from a set of transmitters deployed over a region of interest. A dataset containing real-world power measurements was used to extract the path loss exponent and decorrelation time for the region. Using only a finite number of measurements, power maps were constructed for the entire region by employing three interpolation techniques and the performance of each technique was evaluated by considering a model-free approach and a model-based approach. The effect of using a subset of measurements from the dataset on the visual quality of power maps was investigated and the minimum number of measurements to construct sufficiently accurate maps was obtained. The work considered here could be extended to different wireless environments by accounting for the attenuation by walls and other obstacles such as furniture in an indoor environment, in addition to the path loss and lognormal shadow fading.

References

[1] A. Alaya-Feki, S. Ben Jemaa, B. Sayrac, P. Houze, and E. Moulines, "Informed spectrum usage in cognitive radio networks: interference cartography," in *Proc. PIMRC*, Sept. 2000, pp. 1–5.

[2] S. Ureten, A. Yongacoglu, and E. Petriu, "A comparison of interference cartography generation techniques in cognitive radio networks," in *Proc. ICC*, June 2012, pp. 1879–1883.

[3] G. Mateos, J. Bazerque, and G. Giannakis, "Group-lasso on splines for spectrum cartography," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4648–4664, Oct. 2011.

[4] G. Boccolini, G. Hernandez-Penaloza, and B. Beferull-Lozano, "Wireless sensor network for spectrum cartography based on kriging interpolation," in *Proc. PIMRC*, Sept. 2012, pp. 1565–1570.

[5] J. Medbo and J. Berg, "Simple and accurate path loss modeling at 5 GHz in indoor environments with corridors," in *Proc. VTC (Fall)*, Sept. 2000, pp. 30–36.

[6] R. A. Valenzuela, D. Chizhik, and J. Ling, "Measured and predicted correlation between local average power and small scale fading in indoor wireless communication channels," in *Proc. VTC*, May 1998, pp. 2104–2108.

[7] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.

[8] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Comput. Netw.*, vol. 45, no. 3, pp. 245–259, Jun. 2004. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2004.03.007

[9] "Visiwave site survey tool." [Online]. Available: http://www.visiwave.com/

[10] T. King, S. Kopf, T. Haenselmann, C. Lubberger, and W. Effelsberg, "Crawdad data set mannheim/compass (v. 2008-04-11)," Apr. 2008. [Online]. Available: http:/crawdad.org/mannheim/compass/

[11] O. Falivene, L. Cabrera, R. Tolosana-Delgado, and A. Sáez, "Interpolation algorithm ranking using cross-validation and the role of smoothing effect: a coal zone example," *Computers & Geosciences*, vol. 36, no. 4, pp. 512–519, Apr. 2010.

[12] E. H. Isaaks, *An introduction to applied geostatistics*. Oxford University Press, 1989.

[13] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.