Revisiting TV Coverage Estimation with Measurement-based Statistical Interpolation

Xuhang Ying, Chang Wook Kim, Sumit Roy
Department of Electrical Engineering, University of Washington
Email: {xhying, hyperkcw, sroy}@u.washington.edu

***Abstract*—Per FCC rules, secondary users in TV White Spaces must operate only within spectrum subject to a *no-harmful-interference* condition to existing primary receivers. In effect, this translates into a protection region around every TV transmitter, wherein secondary nodes must *not* transmit (on the same channel). This is implemented by requiring secondary users to consult a database prior to any channel access so as to obtain the local prohibited channels/spatial zones (or equivalently the free channels or White Spaces). Construction of such protection regions for a transmitter within a database has been done, mainly based on empirical propagation models that estimate the received signal strength at a location. Clearly, such model-based prediction is always of limited accuracy and should be supplemented by measurement based approaches that help validate and improve the database predictions. In this work, we present results based on applying spatial interpolation techniques (Kriging) to measurement data obtained in Seattle, WA. Our results have shown that empirical DBA models tend to over-estimate received signal strengths by not explicitly accounting for local obstructions, while measurement-based Kriging achieves consistently good performance. Furthermore, boundary estimation via Kriging achieves a type I error rate 46.1% lower than comparable DBA approach while keeping type II error rate under a low limit (5%) for a given service threshold (i.e., -84 dBm/6 MHz); this is also an improvement over a method using k-Nearest Neighbor for such estimation.**

I. Introduction

The exponential growth of data services such as mobile multimedia on devices like smartphones and tablets translates to a proportionate surge in a need for additional network bandwidth. However, within today's fixed spectrum allocation policy regime, a large portion of the licensed spectrum is temporarily unused or underutilized by *licensed (primary) users*. To address the need for improved spectrum utilization and efficiency, a promising technology called *Dynamic Spectrum Access (DSA)* has been proposed, which allows *unlicensed (secondary) users* to dynamically access locally unoccupied spectrum (i.e., so called *white spaces* (WS)) while subjecting them to a no-harmful-interference-to-primary-users condition.

This work was supported in part by NSF Grant AST-1443923.
Sumit Roy is currently with IIIT-D on sabbatical leave.

The first potential application of DSA principles was intended for the VHF/UHF TV bands, concurrent with the transition from analog-to-digital over the broadcast in the U.S. In 2008, the Federal Communication Commission (FCC) opened up (portions of) unused TV band spectrum [1]–[3]. Per FCC ruling [2], secondary users must obtain information regarding locally available spectrum by contacting a *database administrator (DBA)* to retrieve a list of available channels. A core output of DBA querying is the estimated coverage contours of primary sources (i.e., TV transmitters) based on suitable propagation models (notably, the FCC Curves [1] prescribed by the FCC for all current DBAs[1]). In any case, the contour is determined by comparing the (estimated) signal strength at any location to a preset threshold (e.g., 41 dBμV/m or -84 dBm for digital TV services[2]). Secondary TV band devices are not allowed to operate (transmit) within this region plus the vicinity - the so-called "no-talk" zone.

Concerns about local accuracy of such signal strength estimation based on propagation models include the fact that these do not account properly for *built environments* (e.g. buildings et al) in dense urban outdoors as well as indoors [6]. As a result, DBAs may not estimate coverage regions accurately, causing possible interference to primary users or missing WS opportunities. On the other hand, local spectrum sensing is generally more accurate, but is expensive and labor-intensive, making sampling at all locations over a large geographical region practically infeasible. This leads to a natural proposition given the respective advantages of both approaches - can these be effectively combined, namely, by conducting a small amount of local measurements over fringe areas of protected regions predicted by DBAs, using statistical interpolation techniques to estimate signal strengths at

[1]It is worthwhile noting that coverage prediction based on Longley-Rice Irregular Terrain Model (ITM) [4] as used in http://specobs.ee.washington.edu, overcomes several limitations of the FCC curves.

[2]This threshold is subject to calculation for DTV stations using the Longley-Rice methodology based on the receive system model in [1], [5]. FCC ignores actual waveform information (e.g., ATSC 1.0), which should be considered as a weakness.

unmeasured points, and finally refining protected contours based on interpolation results?

In [7], [8], authors applied well-established geostatistical interpolation technique called *Kriging* to coverage prediction, and showed that it is able to achieve competitive or better performance as compared to propagation models. While this prior work serves as the immediate inspiration, they limited themselves to signal interpolation and did not actually explore the problem of estimation of the TV coverage regions - which is the main issue of interest in protecting primary receivers.

In this work, we quantify and compare the performance of DBA model-based predictions (serving as a baseline) with measurement-based approaches that employ Kriging followed by classification for coverage region estimation. Towards this end, we conduct a vehicle-based measurement campaign over a 4.6km x 5km suburban region in Seattle, WA, to collect I/Q samples of 31 permissible UHF channels at 240 locations. Our primary contributions are as follows.

- As a baseline, we first quantify the received signal strength (RSS) prediction errors of DBA models – FCC Curves and Longley-Rice (LR) model. Our results show that these tend to *over-estimate* true RSS resulting from omitting man-made or natural local obstructions (e.g., trees, buildings etc.).
- We further quantify the boundary estimation performance in terms of type I/II error rates, using 3 approaches: a) Kriging on measurement data followed by estimation, b) Estimation on measurement data using k-Nearest Neighbor (k-NN) classifier, and c) Estimation on predicted field strengths using DBA models. Our result show that for a preset threshold (i.e., -84 dBm), Kriging reduces type I errors by 46.1% compared to DBA models while keeping type II error rate under 5%, which is an improvement with respect to the k-NN approach. More importantly, both measurement driven approaches performance *significantly* better than DBA models that suffer from high type I error rate.

The paper is organized as follows. In Section I-A, we present a review of recent Kriging applications in radio mapping. Basic Kriging background is provided in Section II, and our measurement campaign is described in Section III. We compare Kriging RSS and boundary estimation against DBA models in Section IV and V respectively, and conclude this study in Section VI.

A. Related Work

Radio mapping approaches rest on two broad thematic pillars - use of analytical propagation models for a-priori prediction of signal strength complemented by measurement-driven a-posteriori techniques. The performance of propagation models has been investigated earlier in literature [9], [10] and more recently, in [6] authors systematically analyzed the accuracy of numerous propagation models using a large dataset. Data driven spatial interpolation using statistical methods rests on the work of Daniel Krige originally developed for mining applications, and later applied to many other fields such as soil science, hydrosciences, as well as to wireless environment mapping [11]–[15].

The comparative use of predictive path loss models (such as FCC Curves and the LR model) and Kriging based approaches for signal strength estimation has only a sparse recent history. In [7], authors compared RSS estimation performance of the LR model against Ordinary Kriging (OK) for a single TV source; in [8], authors further analyzed the multi-transmitter scenario, and empirically demonstrated the advantages of Kriging. Unlike [14], we do not track (slow channel fading) variations and defer investigation of temporal interpolation to future work. In this work, we attempt a structurally similar exercise specifically for mapping of TV channels; we explore prediction accuracy of FCC curves (mandated by FCC for DBAs) and the Longley-Rice point-to-point model in predicting available channels. We also conduct a measurement campaign in TV bands for a typical suburban region in Seattle, WA of approx. 23 sq. km., and extend Kriging techniques for boundary estimation - for comparison against model based predictive techniques.

II. Method

The RSS (in dBm) at a point in a region of interest $D \in \mathbb{R}^2$ is modeled as a *random field* $Z(\mathbf{x})$, and any set of measured RSS values $\{Z(\mathbf{x}_i) : i = 1, 2, ..., n\}$ is a realization of the underlying random process. One of the popular Kriging techniques, Ordinary Kriging (OK), models this with a mean field μ and the residual $\delta(\mathbf{x})$:

$$Z(\mathbf{x}) = \mu + \delta(\mathbf{x}), \mathbf{x} \in D \tag{1}$$

where μ is an unknown constant. In Radio Mapping, one may interpret μ as being determined by the path loss and large-scale fading effects, and $\delta(\cdot)$ possible sampling errors. In OK, $Z(\mathbf{x})$ is further assumed to be *intrinsically stationary*, i.e.,

$$E[Z(\mathbf{x})] - E[Z(\mathbf{x}+\mathbf{h})] = 0 \tag{2}$$

$$E[(Z(\mathbf{x}) - Z(\mathbf{x}+\mathbf{h}))^2] = 2\gamma(\mathbf{h}) \tag{3}$$

where $\gamma(\mathbf{h})$ is called the *semivariogram*, and $\mathbf{h}$ the *distance lag* between two locations[3]. Generalized Kriging

[3] $\mathbf{h}$ is a vector with an amplitude and direction. A semivariogram considering directions is called an *anisotropic* semivariogram, otherwise, *isotropic*. In this paper, we consider isotropic semivarograms, leaving anisotropy our future work.

techniques drop this constant-mean assumption by applying a varying-mean model (e.g., Universal Kriging). In this work, we report on results using OK (and defer investigation of other Kriging methods to future work).

A. Semivariogram $\gamma(h)$

At the center of geo-statistical analysis lies $\gamma(h)$, a function representing the spatial continuity or correlation of the signal field, i.e., two points closer in space tend to have more similar RSS values than those farther apart.

1) Constructing empirical semivariogram $\hat{\gamma}(\mathbf{h})$ for measured data: In this study, we adopt the *Cressie-Hawkin (C-H) robust estimator* [16] given as below:

$$\hat{\gamma}(h) = \frac{1}{2} \cdot \frac{\left\{ \frac{1}{N(h)} \sum_{N(h)} (Z(\mathbf{x}_i) - Z(\mathbf{x}_j))^{1/2} \right\}^4}{(0.457 + 0.494/N(h))} \tag{4}$$

where $Z(\mathbf{x}_i)$ and $Z(\mathbf{x}_j)$ are values at locations $\mathbf{x}_i$ and $\mathbf{x}_j$, whose distance separation is *approximately* equal to $\mathbf{h}$, and $N(h)$ the total number of such pairs. Compared to the *classical estimator* used in [12], the C-H estimator is more robust to atypical observations and outliers [16].

2) Fitting $\hat{\gamma}(h)$ with parametric models: The next step is to fit $\hat{\gamma}(h)$ with parametric $\gamma(h)$ models, such as exponential, Gaussian, spherical and cubic models. For example, the exponential model is defined as:

$$\gamma(h) = a + (s - a)\left(1 - e^{-3h/r}\right), h \geq 0 \tag{5}$$

where a (i.e., the *nugget*) represents the discontinuity around the origin caused by sampling errors and small-scale variability, r (i.e., the *range*) the distance where the model first flattens out, and s (i.e., the *sill*) the limit of $\gamma(h)$ when h goes to infinity, which measures the maximum variance between two points that are far apart. Then the Weighted Least Squares (WLS) is performed to determine the best parameters for each model, using the available package in "geoR" [17].

3) Choosing the best fitted model: To find the best among competing fitted models, we cross-validate each fitted model with the same training dataset used for $\hat{\gamma}(h)$ estimation. A popular cross-validation (CV) technique called the *leave-one-out* CV works as follows: each and every one data point is excluded from the training dataset, whose value is then predicted by OK with the proposed model using the remaining data, and compared with its true value. The one with the smallest mean squared error (MSE) is selected as the best model.

B. Interpolation

In OK, the predicted value $\hat{Z}(\mathbf{x}_0)$ at an unmeasured location $\mathbf{x}_0$ is a linear combination of measured values at nearby locations $\{Z(\mathbf{x}_i) : i = 1, 2, ..., n\}$, namely,

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^{n} \omega_i Z(\mathbf{x}_i) \tag{6}$$

where weights $\{\omega_i\}$ are normalized, i.e. $\sum_{i=1}^{n} \omega_i = 1$.

The minimization of mean squared error $E[(\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0))^2]$ with respect to $\{\omega_i\}$ under the normalization constraint leads straightforwardly to a set of linear equations, solving which results in the optimal $\{\omega_i^*\}$:

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ ... \\ \omega_n \\ \lambda \end{pmatrix} = \begin{pmatrix} \gamma(d_{11}) & ... & \gamma(d_{1n}) & 1 \\ \gamma(d_{21}) & ... & \gamma(d_{2n}) & 1 \\ ... & ... & ... & ... \\ \gamma(d_{n1}) & ... & \gamma(d_{nn}) & 1 \\ 1 & ... & 1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} \gamma(d_{01}) \\ \gamma(d_{02}) \\ ... \\ \gamma(d_{0n}) \\ 1 \end{pmatrix} \tag{7}$$

where $d_{ij} \equiv |\mathbf{x}_i - \mathbf{x}_j|$, the Euclidean distance between locations i and j, and λ the Lagrange parameter used to enforce the normalization constraint. The minimized MSE is called the *Kriging variance*, representing prediction uncertainty at location $\mathbf{x}_0$. In general, the Kriging variance is smaller if the unmeasured location is close to nearby data points providing interpolation support.

III. Measurement Campaign

A vehicle-based measurement campaign was conducted over a 4.6km x 5km sub-urban area (Fig. 1) with a typical geographical layout (e.g., streets, trees, buildings etc.) in Seattle lasting two and a half days. Our main purpose is to evaluate the efficacy of Kriging in estimating RSS and coverage boundaries for TV transmitters. A total of 31 permissible 6-MHz UHF channels (CH 21 to 51) from 512 MHz to 698 MHz, were sampled at a total of 240 different locations.

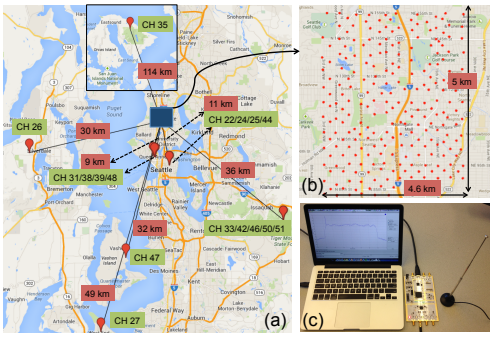


Fig. 1: (a) Locations of TV transmitters whose FCC Curves coverage regions cover the sampled regions. (b) 240 sample locations (in red dots). (c) Spectrum analyzer consisting of a laptop, a USRP B210 and an antenna.

A. Equipment and Setup

Our spectrum analyzer (Fig. 1) was constructed with a laptop, a USRP B210 operating over GNURadio [18], and an omni-directional digital TV antenna of a 3-dBi

gain, which was securely mounted on the top of a van. The antenna height was around 2.2 meters above ground.

To calibrate our spectrum analyzer, we fed tones with known signal power at a known frequency using a signal generator, recorded I/Q samples, estimated uncalibrated signal power at the corresponding bin, and computed the average offset based on 900 measurements. To further measure the noise level, we fed nothing to USRP and estimated the calibrated signal strength over a 6-MHz channel. The maximum signal strength (with noise only) out of 500 measurements was taken as the noise level.

B. Measurement Design

We adopted a grid sampling design and determined measurement locations as follows. First, we defined the region of interest, and generated a "uniform" (equilateral triangular) sample grid similar to [12]. To balance opportunity costs (e.g., labor, time etc.) and interpolation resolution, a grid spacing of approximate 350 meters was chosen, resulting in a total of 240 sample locations. However, due to various physical constraints (e.g., roads, buildings, trees etc.), we were not able to comply with the original sampling plan strictly, and hence the final layout appeared a little irregular at many places. That being said, our measurement campaign reflected the complexity of practical sampling over a large geographical region, as well as irregularity in obtained data. In practice, we did not follow any particular sampling order, and sampled in a way that was smooth and efficient.

C. Methodology

Our measurement consisted of two phases: i) I/Q sample collection, and ii) energy-detection-based post-processing. In the first phase, we prepared a Python script which tuned USRP to the center frequency of a 6-MHz channel, and sampled the signal at a rate of 8 MHz for 0.5 seconds. The main reason of over-sampling a channel was to minimize FFT edge effects by dropping the first and last 12.5% bins of obtained periodograms (up to a total of 25% bins, i.e., 2 MHz) later at the post-processing stage.

Raw I/Q samples for a single channel were then stored in shorts (i.e., two bytes for an I or Q sample), generating a data file of size 16 MB. After sampling a channel, the script waited for 0.5 seconds before switching to the next one to ensure successful tuning. Hence, it took approximately 30 seconds to perform a complete scan for 31 channels.

The script varied a key parameter called the USRP gain (ranging from 0 to 72) to adjust SNR when sampling signals, especially weak ones. However, if the signal was strong at the sampled location and the gain was too high, it would cause Analog-Digital-Conversion (ADC) saturation, distorting power spectral density (PSD) and RSS data. To avoid possible ADC saturation for any channel at any location, we set the gain to a relatively low value of 26. On the other hand, to measure RSS for CH 35 (weak signal) more accurately, we chose a higher gain of 36. Given the above USRP settings, the measured noise levels were -82.29 dBm and -92.92 dBm for gains of 26 and 36 respectively.

In the post-processing phase, we used Matlab to perform energy detection based on I/Q samples. Apart from over-sampling, we also applied Blackman-Harris windowing before performing FFT to further minimize the effect of spectral leakage. The number of FFT bins was 2048 for a 8-MHz band (before truncation), and the frequency resolution of was 3.9 KHz per bin. After dropping 25% of bins, a total of 1536 bins were obtained for a 6-MHz channel. Then we computed power spectral density (PSD) by averaging 100 periodograms, whose equivalent sensing duration was 25.6 ms[4]. Eventually, we could get a 240-by-3 matrix of output data for each channel containing (X, Y, RSS) at 240 locations[5].

IV. RSS Estimation

RSS estimation is the building block for predicting coverage regions of primary sources. In this section, we first demonstrate the application of OK to RSS prediction, and then quantify its performance improvement over DBA models (i.e., FCC Curves and Longley-Rice).

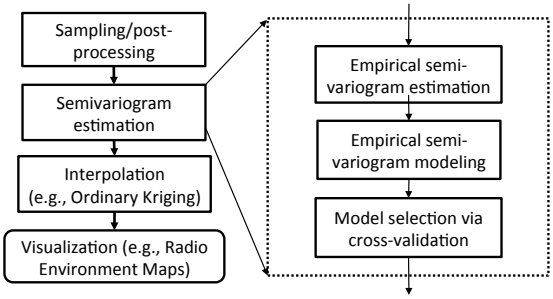


Fig. 2: General Kriging procedure.

A. An Ordinary Kriging Example

Fig. 2 illustrates a general Kriging procedure. After sampling and post-processing, the empirical semivariogram $\hat{\gamma}(\mathbf{h})$ (truncated, as shown in Fig. 3b) was estimated from measured data using the robust estimator in Eq. 4. The lag spacing was 0.35 km - same as the sample spacing - and 1/3 of the maximum pairwise distance

[4]Different from real-time spectrum sensing, sensing duration was not critical for Kriging applications. Hence, we chose a relatively long sensing duration to ensure good energy detection performance.

[5]Locations were converted from geo-coordinates with respect to a reference point.

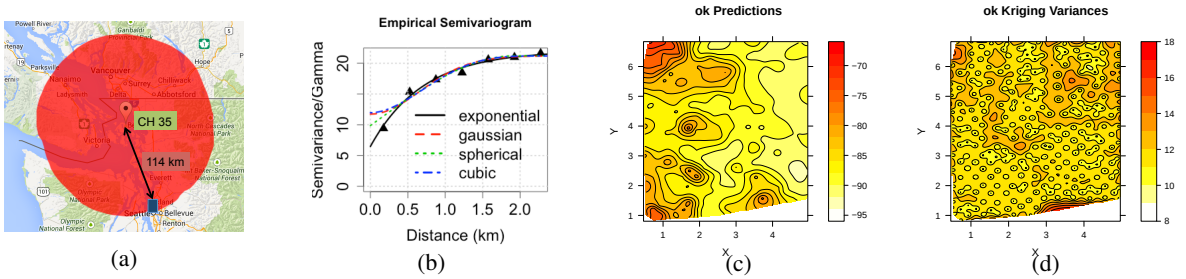


Fig. 3: (a) TV transmitter of CH 35 and its DBA coverage region (in red). (b) Empirical semivariogram and four fitted parametric models (i.e., exponential, Gaussian, spherical and cubic). Based on cross-validation results, the exponential model was chosen with a nugget of 6.48, a range of 2.11 and a sill of 22.02. (c) A map of RSS predictions. (d) A map of Kriging variances. The red indicates high values, while the white implies low values. Lines in (c)/(d) are contours representing different RSS/Kriging variance values (not shown due to limit of space).

was chosen as the maximum lag distance (i.e., 2.3 km). Next, four popular models, i.e., exponential, Gaussian, spherical and cubic models were fitted to $\hat{\gamma}(\mathbf{h})$, and the exponential one was selected based on CV results.

To visualize a Radio Map for a continuous region (i.e., estimating every location within the region), we first discretized it using a mesh of grid points at a particular resolution. For example, with a resolution of 0.03 km, we have a total of over 28,400 points for the target region. Then, we interpolated the RSS value for each grid point, and obtained its corresponding Kriging variance. Finally, we plotted both Kriging predictions and variances on maps as shown in Fig. 3c and 3d.

B. Kriging Versus DBA Models

To compare the performance between OK, Longley-Rice (LR) and FCC Curves (FC), four channels (e.g., CH 25, 31, 33, 50) were investigated with diverse channel characteristics (Table I), whose average SNR exceeded 15 dB so that noise power could be omitted. Representing the simplest method, OK's performance may serve as the lower-bound of the Kriging family. To ensure fair comparison, the following equation was used to convert fields strength predictions (dBμV/m) of DBA models to signal strengths (dBm):

$$P_{dBm} = E_{dB\mu V/m} - 20log f_{MHz} + G_{dBi} - 77.2 \quad (8)$$

where f_{MHz} is the center frequency of a channel and $G_{dBi} = 3\ dBi$ is the antenna gain.

The leave-one-out cross-validation technique was applied, and two metrics were adopted to quantify the prediction performance: i) the bias or mean error (ME) – the average difference between predicted and measured RSS (ground truth), and ii) the root mean squared error (RMSE) between the prediction and ground truth.

We have two observations from results in Table I. First, OK achieves a ME value very close to 0, mainly because it is an unbiased estimator based on local measurements. On the other hand, although FC causes a smaller ME than LR, both over-estimate RSS in general causing a ME as high as 31.59 and 27.27 dB respectively. Second, OK consistently produces a RMSE of 5 - 6 dB, which is significantly better than DBA models, primarily because that DBA models do not explicitly account for environmental obstructions (e.g. trees, buildings, etc.). However, DBA model performance differs from channel to channel, thus making OK gain channel-dependent.

V. Boundary Estimation

We first introduce two metrics used to measure boundary estimation performance, and then compare Kriging against DBA models and the popular k-Nearest Neighbor (k-NN) classifier, applied to our problem.

A. Performance Metrics

TV coverage regions are essentially decision regions (Fig. 3a) defined by their boundaries, based on which we may predict at a given location whether a channel is either *available* for unlicensed access (labeled with 1, TV service unavailable) or *occupied* (labeled with 0, TV service available). In DBAs, the channel availability label is determined via simple thresholding as follows:

$$\hat{L}_{DBA}(\mathbf{x}) = \begin{cases} 1, & \text{if } \hat{Z}_{DBA}(\mathbf{x}) < \Gamma \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

where $\hat{Z}_{DBA}(\mathbf{x})$ is the predicted field strength (dBμV/m), $\hat{L}_{DBA}(\mathbf{x})$ the predicted label, and Γ a preset full-power DTV service threshold (i.e., 41 dBμV/m).

Therefore, it is essentially a *classification* problem to determine channel availability at a location. Given a set of N measurements, we may evaluate the performance of different classification schemes or boundaries via leave-one-out cross-validation as follows. First, one data point is excluded from the entire dataset, whose label is predicted based on estimated boundaries derived from

CH	ERP	HAAT	Dist.	Mu	Sigma	ME (dB)			RMSE (dB)			OK Gain
	(dBW)	(m)	(km)	(dBm)	(dB)	OK	LR	FC	OK	LR	FC	Over FC
25	60.00	290	11.32	-50.34	7.11	0.02	31.59	27.27	5.15	32.45	28.10	22.95
31	58.45	218	9.48	-53.46	8.34	0.00	30.65	26.57	5.79	31.51	27.92	22.13
33	56.02	716	35.95	-68.86	6.59	-0.07	27.86	21.60	5.21	29.06	22.63	17.42
50	53.08	719	35.98	-52.06	8.87	0.00	14.99	8.95	5.72	19.78	12.75	7.03

TABLE I: Performance comparison between OK, Longley-Rice (LR) and FCC Curves (FC) via ME and RMSE. OK gain is defined as the amount of RMSE reduced by OK compred to FC. ERP stands for effective radiate power, and HAAT height above average terrain. Note that Kriging performs significantly better than DBA models.

the remaining data. If the measured RSS value at that location is treated as *ground truth*, its *true* label may then be obtained with respect to a preset threshold through thresholding. Comparing the predicted label against the true label, we get two types of errors:

- *Type I error*: a channel is predicted to be occupied (0), when it is actually available (1).
- *Type II error*: a channel is predicted to be available (1), when it is actually occupied (0).

Then we could repeat the above procedure for each data point, and count type I/II errors for a particular boundary estimation scheme and arrive at:

- *Type I error rate* ϵ_1:

$$\epsilon_1 = \frac{\text{No. of Type I Errors}}{\text{No. of True 1 Labels}} \tag{10}$$

- *Type II error rate* ϵ_2 :

$$\epsilon_2 = \frac{\text{No. of Type II Errors}}{\text{No. of True 0 Labels}} \tag{11}$$

Notably, the cost of the two types of errors is *asymmetric* per FCC ruling; we seek to reduce ϵ_1 (i.e., missing spectrum opportunities), while keeping ϵ_2 (i.e., possible interference) below an acceptable limit.

B. Methods

1) Method I – DBA Boundary: To first predict the field strength at a given location from a TV transmitter, a DBA model takes engineering parameters such as the distance between transmitter and receiver, channel, effective radiate power, antenna height and patterns, as well as terrain data as inputs (for Longley-Rice), and outputs the field strength (dBμV/m) without involving any local measurements.

To further determine boundaries of a transmitter, a DBA finds the *farthest* point for each azimuth value (0-360 degrees centered at TV transmitter location), whose predicted field strength is above the preset TV service threshold (i.e., 41 dBμV/m). By connecting those points with a continuous line, a DBA is able to obtain boundaries that define a TV coverage region.

2) Method II – k-NN Boundary: k-NN normally takes labels of measured (training) data with respect to a threshold Γ (i.e., -84 dBm) as true inputs, and directly predicts the label of an unmeasured (testing) location as follows. First, it searches for the k nearest neighbors for the given location based on Euclidean distance, and uses the majority label among k neighboring labels as the predicted label (any tie is broken randomly).

Note that k is user-defined, and the optimal value k^* is determined from the training dataset through cross validation as follows. First, the entire training set is randomly divided into M (e.g., 10) subsets. Then, for a given k, each and every subset are excluded, whose labels are predicted based on true labels of other four subsets. Then we could count type I/II errors, and compute ϵ_1/ϵ_2 accordingly. Different values of k from 1 to k_{max} (some large value) are examined, and the one with minimum ϵ_2 (with higher priority) is chosen as k^*.

Obviously, k-NN boundary estimation performance evaluated with the testing dataset in terms of ϵ_1 and ϵ_2 largely depends on Γ. Hence, one way to control k-NN boundaries is to use an adjusted threshold $\Gamma' = (1+\alpha)\Gamma$ to label training data instead of Γ, where $\alpha \geq 0$. For example, if $\Gamma = -84$ dBm and we want to make k-NN more conservative in predicting available channels, we may choose a lower $\Gamma' = -85$ dBm ($\alpha = 1.2\%$).

3) Method III- Kriging Boundary: We use Eq. 9 to produce a labeling rule for Kriging by recognizing the prediction uncertainty at a given location as follows:

$$\hat{L}_{Kri}(\mathbf{x}) = \begin{cases} 1, & \text{if } \hat{Z}_{Kri}(\mathbf{x}) < \Gamma - \lambda \cdot \sigma(\mathbf{x}) \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

where $\sigma(\mathbf{x})$ is the Kriging standard deviation, and $\lambda \geq 0$ a control parameter. Given the true value $Z(\mathbf{x}) > \Gamma$ (occupied) but $\hat{Z}_{Kri}(\mathbf{x}) < \Gamma$ (predicted to be available), $\lambda_{min} = (\Gamma - \hat{Z}_{Kri}(\mathbf{x}))/\sigma(\mathbf{x})$ is the minimum value to change the predicted label from 1 to 0 (no type II errors).

Note that objectives of α in k-NN and λ in Kriging are consistent: the predicted RSS at a given location has to be low enough to be labeled as 1 (available); otherwise, we would rather label it as 0 to avoid possible type II

errors (even at the cost of increased type I errors). Their effects will be explored in the next section.

C. Controlling k-NN and Kriging Boundaries

To understand the impact of α and λ in k-NN and Kriging boundaries, we investigated CH 35 in Seattle whose predicted boundary region via the FCC Curves falls within the measurement region. The TV service threshold Γ was set to be -84 dBm as specified in FCC [1]. Note that the dataset for CH 35 collected with a USRP gain of 36 were used, as they better captured signal power at different locations. We varied α and λ, and computed corresponding ϵ_1 and ϵ_2 with the approach described in Section V-A.

	$\alpha(\%)$	**0**	**2.0**	**4.0**	**6.0**	**7.0**
k-NN	ϵ_1 (%)	12.8	16.7	27.2	47.8	59.4
	ϵ_2(%)	43.3	38.3	18.3	8.3	5.0
	sum (%)	56.1	55.0	45.5	46.1	64.4
	λ	**0**	**0.40**	**0.80**	**1.03**	**1.34**
Kriging	ϵ_1 (%)	10.0	16.7	30.0	41.7	53.9
	ϵ_2 (%)	46.7	31.7	18.3	10.0	5.0
	sum (%)	56.7	48.4	48.3	51.7	58.9

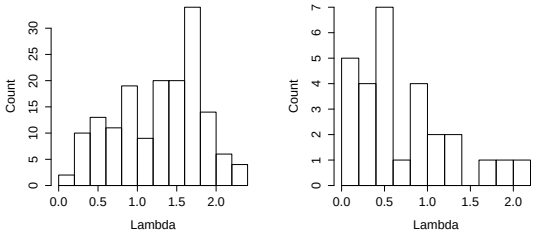
TABLE II: Impact of α and λ on ϵ_1/ϵ_2. $\Gamma = -84$ dBm. With increased α and λ, ϵ_2 of k-NN and Kriging boundaries is reduced, while ϵ_1 is increased.

We make two observations from Table II. First of all, by increasing α and λ, ϵ_2 may be effectively reduced. For instance, when $\alpha = 7.0\%$ and $\lambda = 1.34$, ϵ_2 of both k-NN and Kriging boundaries is reduced to 5%. This is because we want to be more conservative in predicting available channels, thus reducing type II errors as a consequence. Second, there exists a trade-off between ϵ_1 and ϵ_2 as shown in Fig. 4 which illustrates the impact of different λ values. For example, with $\lambda = 0.4$, we are able to reduce 9 type II errors but introduce 12 type I errors.

However, complete elimination of type II errors for Kriging boundaries requires $\lambda = 2.00$, which would introduce a large amount of type I errors. It implies that predicted RSS values at a few locations are much lower than the threshold and their true (measured) values, and it is essentially Kriging's under-estimation at those locations that causes type II errors. To determine whether they are outliers (possibly due to sampling errors), additional measurements may be conducted at the same and nearby locations to improve RSS estimation and subsequent boundary estimation.

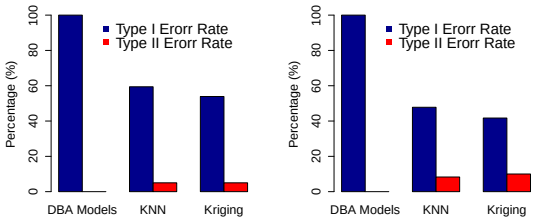
D. Performance Comparison

In this experiment, we compared the performance of Kriging, k-NN and DBA boundaries for CH 35. Based on both FCC Curves and Longley-Rice boundaries, the channel is occupied at all locations within the sampled region (i.e. predicted labels are always 0), leading to $\epsilon_2 = 0$ while $\epsilon_1 = 100\%$. To ensure fair comparison, we adjusted Kriging and k-NN boundary estimation in such a way that ϵ_2 was no larger than 5% or 10%, and compared corresponding ϵ_1 when $\Gamma = -84$ dBm.



(a) Introduced type I errors (b) Reduced type II errors

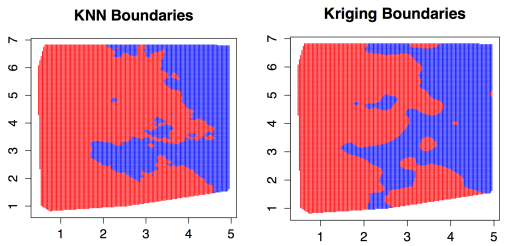
Fig. 4: Histogram of λ_{min} that changes a predicted label from 1 (available) to 0 (occupied) for $\Gamma = -84$ dBm. In (a), both predicted and true labels are 1, and changes in predicted labels from 1 to 0 introduce type I errors. In (b), predicted labels are 0 while the true labels are 1 (i.e. type II errors). Hence, changes in predicted labels from 1 to 0 eliminate type II errors.



(a) $\epsilon_2 \leq 5\%$ (b) $\epsilon_2 \leq 10\%$

Fig. 5: Comparison between DBA models, k-NN and Kriging boundaries for Γ = -84 dBm. (a) $\alpha = 7.0\%$, $\lambda = 1.34$. (b) $\alpha = 6.0\%$, $\lambda = 1.03$.

As shown in Fig 5, Kriging achieves *significant* performance improvement over DBA boundary prediction with gains of 46.1% and 58.3% in ϵ_1 for $\epsilon_2 \leq 5\%$ (i.e., fewer than 3 type II errors) and 10% (i.e., fewer than 6 type II errors) respectively. Although k-NN boundaries cause 6% higher ϵ_1 than Kriging, they also perform very well compared to DBA prediction. The primary reasons that Kriging and k-NN have very close performance are that i) both fundamentally use spatial (neighbor-based) information, and ii) are measurement-driven, which better captures the local environment. Nevertheless, Kriging uses more soft information (i.e. RSS values) than binary labels (as in k-NN) and from more than k neighbors.



(a) (b)

Fig. 6: Coverage regions (in red) defined by (a) k-NN ($\alpha = 7.0\%$) and (b) Kriging boundaries ($\lambda = 1.34$) for $\Gamma = -84$ dBm. The non-coverage region is in blue.

Fig. 6 illustrates k-NN and Kriging boundaries using all data points for $\Gamma = -84$ dBm. While the coverage region predicted are very similar, Kriging appears to draw boundaries more smoothly compared to k-NN. This is mainly because k-NN considers the nearest k neighbors, and thus two nearby locations could end up with different labels, if the majority of their respective neighborhoods have different labels. Furthermore, both tend to create red/blue "holes" due to data points with large/low RSS values that are very different from their neighbors. Hence, to determine whether they are true outliers or not, we may need to take additional measurements around those holes to improve boundary estimation.

E. Impact of TV Service Threshold

Despite the specified TV service threshold of -84 dBm by FCC, it would be interesting to compare k-NN and Kriging boundaries for different Γ values. The results are shown in Table III; despite fairly close performance, Kriging performs generally better than k-NN in estimating boundaries with a gain as high as 16.5%. More importantly, both achieve significant and consistent performance improvement over DBA models.

	Γ **(dBm)**	**-83**	**-84**	**-85**	**-86**
$\epsilon_2 \leq 5\%$	k**-NN**	53.9	59.4	58.3	74.8
	Kriging	56.0	53.9	52.6	58.3
	Gain	-2.1	5.5	5.7	16.5
$\epsilon_2 \leq 10\%$	k**-NN**	47.1	47.8	46.3	49.7
	Kriging	39.3	41.7	39.4	49.1
	Gain	7.8	6.1	6.9	0.6

TABLE III: ϵ_1 (%) of k-NN and Kriging boundaries for different service thresholds for a bounded ϵ_2. The gain is the reduced amount in ϵ_1 of Kriging compared to k-NN.

VI. Conclusion

In this study, we quantified RSS prediction errors of both DBA models and Kriging, and further compared the performance of Kriging boundaries against DBA and k-NN boundaries with data obtained at 240 locations over a 4.6km x 5km sub-urban area in Seattle. Our results have shown that since empirical DBA models do not take into account local obstructions, they tend to over-estimate RSS, while measurement-based Kriging achieves consistently good performance. Furthermore, Kriging boundaries achieves a type I error rate 46.1% and 6.5% lower than DBA and k-NN boundaries respectively while keeping type II error rate under a low limit (5%) for a given service threshold (i.e., -84 dBm), and such improvement also exists for different thresholds.

References

[1] FCC, "Second report and order and memorandum opinion and order," 2008.
[2] ——, "Second memorandum opinion and order," 2010.
[3] ——, "Third memorandum opinion and order," 2012.
[4] A. G. Longley and P. L. Rice, "Prediction of tropospheric radio transmission loss over irregular terrain. a computer method-1968," DTIC Document, Tech. Rep., 1968.
[5] FCC, "Oet bulletin no. 69, longley-rice methodology for evaluating tv coverage and interference," 2004.
[6] C. Phillips *et al.*, "Bounding the error of path loss models," in *Proc. 4th IEEE DYSPAN*, 2011, pp. 71–82.
[7] A. Achtzehn *et al.*, "Improving accuracy for tvws geolocation databases: Results from measurement-driven estimation approaches," in *Proc. 7th IEEE DYSPAN*, 2014, pp. 392–403.
[8] ——, "Improving coverage prediction for primary multi-transmitter networks operating in the tv whitespaces," in *Proc. 9th IEEE SECON*, 2012, pp. 623–631.
[9] V. Erceg *et al.*, "An empirically based path loss model for wireless channels in suburban environments," *Selected Areas in Communications, IEEE Journal on*, pp. 1205–1211, 1999.
[10] V. Abhayawardhana *et al.*, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *Vehicular Technology Conference, 2005.*, 2005, pp. 73–77.
[11] A. Konak, "A kriging approach to predicting coverage in wireless networks," *International Journal of Mobile Network Design and Innovation*, pp. 65–71, 2009.
[12] C. Phillips *et al.*, "Practical radio environment mapping with geostatistics," in *Proc. 5th IEEE DYSPAN*, 2012, pp. 422–433.
[13] J. Ojaniemi *et al.*, "A practical method for combining multivariate data in radio environment mapping," in *IEEE PIMRC 2013*.
[14] S.-J. Kim *et al.*, "Cooperative spectrum sensing for cognitive radios using kriged kalman filtering," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 1, pp. 24–36, 2011.
[15] H. B. Yilmaz and T. Tugcu, "Location estimation-based radio environment map construction in fading channels," *Wireless Communications and Mobile Computing*, 2013.
[16] N. Cressie and D. Hawkins, "Robust estimation of the variogram: I," *Journal of the International Association for Mathematical Geology*, pp. 115–125, 1980.
[17] P. J. Ribeiro Jr and P. J. Diggle, "geor: A package for geostatistical analysis," *R news*, vol. 1, no. 2, pp. 14–18, 2001.
[18] "Universal software radio peripheral b210," https://www.ettus.com/product/details/UB210-KIT, accessed: 2014-07-11.