

## Research Article

# Parameter Estimation for the Field Strength of Radio Environment Maps

Zhisheng Gao, Yaoshun Li, and Chunzhi Xie

*School of Computer and Software Engineering, Xihua University, Chengdu 610039, China*

Correspondence should be addressed to Zhisheng Gao; [gzsds@189.cn](mailto:gzsds@189.cn)

Received 21 July 2017; Revised 13 October 2017; Accepted 1 November 2017; Published 29 November 2017

Academic Editor: Gianluigi Ferrari

Copyright © 2017 Zhisheng Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The parameters of a radio environment map play an important role in radio management and cognitive radio. In this paper, a method for estimating the parameters of the radio environment map based on the sensing data of monitoring nodes is presented. According to the principles of radio transmission signal intensity losses, a theoretical variogram model based on a propagation model is proposed, and the improved theoretical variation function is more in line with the attenuation of radio signal propagation. Furthermore, a weight variogram fitting method is proposed based on the characteristics of field strength parameter estimation. In contrast to the traditional method, this method is more closely related to the physical characteristics of the electromagnetic environment parameters, and the design of the variogram and fitting method is more in line with the spatial distribution of electromagnetic environment parameters. Experiments on real and simulation data show that the proposed method performs better than the state-of-the-art method.

## 1. Introduction

The radio environment map, which was first proposed by Zhao et al. [1], is mainly used in cognitive radio [2, 3]. A radio environment map is an integrated database that is used to describe the electromagnetic environment. Because of its wide range of applications for radio [4–8], it has been further extended. The radio environment map is a comprehensive database that contains many fields of information, such as the available spectrum profile, geographical features, rules, relevant laws and regulations, radio equipment situation, and expert experience [9]. The radio environment map is fundamental to the construction of a communication network, improving operation efficiency, and managing radio resources. Ojaniemi et al. [10] believe that the core content of the radio environment map is the field strength estimation, so accurately estimating the field strength of the radio signal in the geospatial space with a certain granularity is key. Pesko et al. [11] called this problem the construction of the radio frequency layer. In this paper, we call it radio environment map parameter estimation. This problem, especially since 2012, has been increasingly studied in depth [11–15].

Methods for estimating the parameters of the radio environment map can be divided into three categories [11]. The first are based on direct spatial interpolation, the second are based on the propagation model, and the third are the hybrid combinations of the first two methods. Propagation-based methods require a large amount of information, including the signal transmission source, latitude and longitude coordinates, antenna height, transmitting power, and even geographical information and climate information on the propagation path, which greatly limits the scope of application of this method. At the same time, because most propagation models are empirical models based on radio transmission, their universality is not strong. Ojaniemi et al. [10] showed that, under certain conditions, this method has lower prediction accuracy than spatial interpolation methods.

In recent years, the focus of research on radio environment map parameter estimation has transferred to spatial interpolation-based methods, especially methods based on geostatistics. In this kind of method, the measured values of ground truth are obtained by radio monitoring sensors, and then the spatial environment parameters of the remaining

locations are obtained using spatial interpolation estimation. Comparative studies of spatial interpolation methods were made in [15–17]. Comparative studies of the inverse distance weighted (IDW) method, Spline interpolation method, and Kriging interpolation method were presented in [15, 16, 18], and the IDW, gradient plus inverse distance squared (GIDS), and Kriging methods were compared in [17]. Prediction experiments on indoor and outdoor electromagnetic environments showed that the IDW method is more robust, while the Kriging method is the most accurate method. Reference [19] presented an approach that uses the spatial dependence of ground truth data and constructs the signal intensity map using Kriging. In [20], several spatial interpolation methods based on IDW were analyzed and used to estimate the spatial distribution of radio field strength. Reference [21] proposed a geostatistical method for the radio environment map and through an actual case study demonstrated that the method is superior to the method based on path loss model and data fitting. At the same time, this kind of method relies on the data collected by the monitoring sensor, so the distribution and quantity of the monitoring sensors affect the prediction accuracy of the radio environment map parameters. In [22], the relationship between the number of sensors and construction error of the radio environment map is analyzed in detail.

Existing studies show that the Kriging method is the best way to estimate the parameters of the radio environment map. However, the radio transmission process is affected by various factors such as the number of transmitting stations, geographical environment, and weather. In practice, the number of monitoring sensors is limited, so the data sampling points are sparsely distributed, which increases the difficulty of estimating the parameter space distribution. At the same time, because the Kriging algorithm is based on a variogram, its linear quadratic optimization is based on the assumption that the data set conforms to the normal distribution and meets the second-order stationary hypothesis or quasi-second-order stationary assumption. Therefore, a nonnormal distribution will affect the stability of the data and cause the variogram to produce a proportional effect. That is, it will improve the sill and nugget values and increase the estimation error [23]. To solve this problem, we propose a method to estimate the parameters of the radio environment map based on the radio propagation model and the Kriging method. This method retains the advantages of both the propagation model and the Kriging method and hence obtains better parameter space prediction precision than the single method. The main contributions of this paper are as follows: (1) Using the radio propagation model to improve the variogram of the Kriging algorithm, a new theoretical variogram model for radio environment map parameter estimation is proposed. (2) Based on the characteristics of radio signal propagation and data acquisition, a weighted optimization of the variogram is proposed, and particle swarm optimization (PSO) is applied to fit the modified variogram. The modified Kriging algorithm can hence be better adapted to the spatial distribution of the radio environment parameters.

The rest of the paper is organized as follows. Section 2 describes the related research on interpolation-based prediction of the radio environment. Section 3 introduces the improved Kriging method based on the electromagnetic propagation model and PSO-based weighting variogram fitting. Section 4 presents the results of some comparative experiments on real and simulation data to examine the effectiveness of the proposed method. Finally, Section 5 concludes this work.

## 2. Related Works

The IDW has been considered for radio environment map parameter space estimation in many studies [14–18, 20, 21, 24]. The estimated value of the forecast point parameter  $O_e$  can be calculated by the weighted sum of the actual observation values of nearby observation points. This method considers that the contributions of the observation points closer to the prediction point are greater; otherwise, the contribution is smaller, which can be expressed as follows:

$$O_e = \frac{\sum_{i=1}^n (d_i)^{-p} O_d(i)}{\sum_{i=1}^n (d_i)^{-p}}, \quad (1)$$

where  $O_d(i)$  is a sampled value of the actual parameter at the  $i$ th observation point,  $d_i$  is the Euclidean distance between the  $i$ th observation point and the predicted point, and  $p$  is a strength parameter that defines the decrease in weight as the distance increases. When  $p$  equals one, the method is called IDW, and when it equals two, the method is called the inverse distance squared weight.

Spline interpolation is another widely used method for estimating the parameters of the radio environment map [16, 20, 25]. In these methods, the Spline is generated using the actual measured values of all observation points to guarantee global smoothness, and then the parameter values of the predicted points are calculated using polynomial fitting.

The Kriging method is a method based on the spatial analysis of a variogram, which is an unbiased optimal estimation of regionalized variables over a finite area, and is considered to be the best method for estimating the parameters of the radio environment map [11–13, 13–21]. The Kriging method is divided into ordinary Kriging and universal Kriging depending on the existence of space field drift. Ordinary Kriging is more commonly used than universal Kriging [11]. The following is a description of the ordinary Kriging method.

For regionalized variable  $z(x)$ , the sample values for a series of observation points  $x_1, x_2, \dots, x_n$  are  $z(x_1), z(x_2), \dots, z(x_n)$ . Then, the estimated value  $z(x_t)$  of grid point  $x_t$  in a region can be estimated by a linear combination; that is,

$$z(x_t) = \sum_{i=1}^n \lambda_i \cdot z(x_i), \quad (2)$$

where  $\lambda_i$  is the  $i$ th weighting coefficient. According to the principle of optimal unbiased estimation, the value of  $\lambda_i$  should satisfy the following conditions:

$$\begin{aligned} E[z(x_t) - z'(x_t)] &= 0, \\ E[z(x_t) - z'(x_t)]^2 &= \min, \end{aligned} \quad (3)$$

where  $z'(x_t)$  is the real sample value. Assuming that  $z(x)$  satisfies the intrinsic hypothesis, then according to the Lagrange theorem, the ordinary Kriging equations can be expressed as follows:

$$\begin{aligned} \sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu &= \gamma(x_i, x_t), \quad i = 1, 2, \dots, n, \\ \sum_{i=1}^n \lambda_i &= 1, \end{aligned} \quad (4)$$

where  $\gamma(x_i, x_j)$  is the value of the variogram between sampling points  $x_i$  and  $x_j$  and  $\mu$  is the Lagrange constant. Weighting coefficient  $\lambda_i$  can be calculated by (4). When  $\lambda_i$  is substituted into (2), the estimation value  $z(x_t)$  of grid point  $x_t$  can be obtained.

The process of solving  $z(x_t)$  shows that the key of Kriging interpolation is how to obtain the best estimate of variogram  $\gamma(h)$ .

### 3. Proposed Method

**3.1. Improved Variogram for Parameter Estimation of Radio Environment Map.** In geostatistics, a variogram is a tool used to study the autocorrelation structure of regionalized variables. The value of a variogram function is only related to the distance between two regionalized variables. Larger values of the variogram indicate smaller autocorrelation. The variogram function is defined as follows [10]:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (z(x_i) - z(x_i + h))^2, \quad (5)$$

where  $n(h)$  is the number of pairs of observation data points with lag distance  $h$ ,  $z(x_i)$  is the value of the regionalized variable at position  $x_i$ , and  $z(x_i + h)$  is the value of the regionalized variable at a distance  $h$  from  $x_i$ . When the data distribution is relatively uniform, the basic lag distance can be equal to or slightly larger than the minimum distance between the observed data points. Alternatively, the basic lag distance can be obtained by comparing and analyzing the variability and stability of the experimental variogram of several candidate basic lag distances.

In practice, the most important parameter of the radio environment map is the signal radiation level in units of decibels (dB). If the Kriging algorithm is used directly, the expression of  $\gamma(h)$  can be simply obtained by (5) in units of  $\text{dB}^2$ . However, this is not consistent with the dB units of propagation loss that are calculated by the radio propagation model. Hence, the variogram is not dimensionally consistent

with the propagation model. We believe that the transmission loss of the propagation model represents the correlation between the two radio environment parameters. Therefore, to combine the variogram with the propagation model, the definition of the variation function used in traditional geostatistics is modified as follows:

$$\gamma(h) = \frac{1}{n(h)} \sqrt{\sum_{i=1}^{n(h)} (z(x_i) - z(x_i + h))^2}. \quad (6)$$

The newly defined variogram is called the parameter estimation variogram of the radio environment map, and the dimensions of the value calculated by the new variogram are consistent with the dimensions of the transmission loss obtained by the propagation model.

**3.2. Theoretical Variogram Model Based on Propagation Model.** It is necessary to use the theoretical variogram model to fit the actual variogram. The commonly used theoretical models for a variogram are the Gaussian, exponential, and spherical models. In practice, the most commonly used model is the spherical model proposed by Pesko et al. [11].

In this paper, two new theoretical variogram models are proposed based on the Longley–Rice model: one uses the Longley–Rice to model the theoretical variogram directly, and the other introduces free space transmission loss into the first model. The Longley–Rice model, also called the irregular terrain model [7], is mainly used to predict the median path loss over irregular terrain. The median value of the propagation loss in free space for different path lengths is calculated as follows:

$$\begin{aligned} A_{\text{ref}} &= \begin{cases} \max\left(0, A_{el} + k_1 d + k_2 \ln\left(\frac{d}{d_{LS}}\right)\right), & d_{\min} \leq d < d_{LS} \\ A_{ed} + m_d d, & d_{LS} \leq d < d_x \\ A_{es} + m_s d, & d \geq d_x, \end{cases} \quad (7) \end{aligned}$$

where  $d_{\min} \leq d < d_{LS}$  is the visual distance spread,  $d_{LS} \leq d < d_x$  is the diffraction propagation distance, and  $d \geq d_x$  is the scattering propagation distance. In addition,  $A_{el}$ ,  $A_{ed}$ , and  $A_{es}$  are propagation losses for sight, diffraction, and scattering in free space, respectively,  $k_1$  and  $k_2$  are propagation loss coefficients, and  $m_d$  and  $m_s$  are loss coefficients for diffraction and scattering, respectively. If the influence of the free space transmission loss is not taken into account, the loss on the whole transmission path can be expressed by (7).

In this paper, the loss prediction function of visual distance spread is used, and hence the propagation loss can be expressed as follows:

$$L = A_{el} + k_1 d + k_2 \log_e\left(\frac{d}{d_{LS}}\right), \quad (8)$$

where all other variables are defined as in (7).

For given parameters such as the heights of the transmitting and receiving antennae, the value of this function is

only related to distance  $d$ . Hence, the loss prediction can be rewritten as follows:

$$\gamma_{itm}(h) = a_1 + a_2d + a_3 \log_e \left( \frac{h}{\theta} + \varepsilon \right), \quad (9)$$

where  $h$  is the distance of two data points and  $\varepsilon$  is a very small constant, which prevents division by zero. Note that, in practice, the two data sampling points may be in the same coordinate position, and  $h$  is equal to 0 in this case. Coefficients  $a_1$ ,  $a_2$ , and  $a_3$  are coefficients to be determined. The value of  $\theta$  is set to 14,000, which is used to simulate the distance of sight.

If the effect of free space propagation loss is taken into account, the overall loss across the propagation path is

$$L = A_{\text{ref}} + 32.45 + 20 \lg d + 20 \lg f, \quad (10)$$

where  $32.45 + 20 \lg d + 20 \lg f$  is the loss of free space propagation,  $d$  is the propagation distance, and  $f$  is the emissive frequency. According to the same ideas above, (11) can be rewritten as

$$\begin{aligned} \gamma_{itmf}(h) = & a_1 + a_2d + a_3 \log_e \left( \frac{h}{\theta} + \varepsilon \right) \\ & + a_4 \log_{10}(h + \varepsilon), \end{aligned} \quad (11)$$

where  $a_4$  is a coefficient to be determined and the other variables are defined as in (9).

Equations (9) and (11) are the two theoretical variogram models proposed in this paper. The new models are more consistent with the parameter change behaviors in the radio environment map and more accurately reflect the relationship between parameter space changes.

**3.3. Weighted Fitting Algorithm for the Theoretical Variogram.** Using the ground truth data, the theoretical variogram is fitted and the undetermined coefficients in the model are obtained. In traditional methods, the least-squares method is mainly used to fit the function. Its fitness function is

$$F(j) = \sum_{i=1}^n [\gamma(h_{i,j}) - \gamma^*(h_{i,j})]^2, \quad (12)$$

where  $F(j)$  is the fitness function value of the  $j$ th variable,  $h_{i,j}$  is the  $i$ th lag of the  $j$ th variable,  $\gamma^*(h_{i,j})$  is the estimated value of variogram at position  $h_{i,j}$ , and  $\gamma(h_{i,j})$  is the real value of the variogram at position  $h_{i,j}$ . The disadvantage of this method is that it considers the contribution of all data to be equal without considering outliers and specific data points as well as the specificity of the radio environment parameters.

In practice, because of building occlusion and the effects of an uneven distribution of sampling nodes, abnormal noise exists. To overcome this problem, the method proposed in this paper increases the corresponding weight coefficient of the fitness function to strengthen or reduce some environmental factors or meet the distribution characteristics of the variogram. To address the problem of uneven sampling point distributions of the radio environment parameters, the first

weight coefficient  $\lambda_1 = N/N_i$  is introduced, where  $N_i$  is the number of sample point pairs that correspond to a certain lag distance and  $N$  is the total number of sample point pairs. The second weight addresses the inconsistencies and inaccuracies in the sampled point data, which is caused by the electromagnetic shadowing of buildings and reflections, multipaths, and radio propagation diffraction. For example, there could be some abnormally large or unusually small sampled values. To reduce the impacts of unreasonable sample points on the fitness function, the weight coefficient is  $\lambda_2 = \bar{\gamma}(h)/\gamma(h_i)$ , where  $\bar{\gamma}(h)$  is the mean value of the variogram and  $\gamma(h_i)$  is the value of the variogram at lag distance  $h_i$ . The third weight is added because point pairs with smaller lag distances often better reflect the degree of variability of regionalized variables. To increase the contribution of data point pairs with small lag distances, the proposed method adds weight coefficient  $\lambda_3 = \bar{h}/h_i$ , where  $\bar{h}$  is the mean value of the lag distance and  $h_i$  is the corresponding lag distance.

The final weight coefficient is the product  $\lambda = \lambda_1 \cdot \lambda_2 \cdot \lambda_3$ . Then,  $\lambda_i$  can be computed by

$$\lambda_i = \frac{N}{N_i} * \frac{\bar{\gamma}(h)}{\gamma(h_i)} * \frac{\bar{h}}{h_i}. \quad (13)$$

Hence, a new fitness function is obtained, expressed as follows:

$$F(j) = \sum_{i=1}^n \lambda_i [\gamma(h_{i,j}) - \gamma^*(h_{i,j})]^2. \quad (14)$$

For the fitness function defined by (14), the PSO algorithm is used to fit the weighted variogram. In the particle swarm, the position of the  $i$ th particle can be expressed as  $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ ,  $i = 1, \dots, N$ , where  $d$  is the dimension of the solution space and  $N$  is the number of particles. The previous most optimal position of the  $i$ th particle is denoted as  $p_{\text{best}i} = (p_{i1}, p_{i2}, \dots, p_{id})$  and the optimal position of the swarm is denoted as  $g_{\text{best}} = (g_{g1}, g_{g2}, \dots, g_{gd})$ . Each particle has a moving speed, and the moving speed of the  $i$ th particle is  $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$ . At each iteration, the particle velocity and position changes are updated by the following equation:

$$\begin{aligned} V_i^{k+1} = & w * V_i^k + C_1 * \text{rand}_1 * (p_{\text{best}i}^k - X_i^k) + C_2 \\ & * \text{rand}_2 * (g_{\text{best}}^k - X_i^k), \\ X_i^{k+1} = & X_i^k + V_i^{k+1}, \end{aligned} \quad (15)$$

where  $k$  is the number of iterations and  $C_1$  and  $C_2$  are learning factors (or acceleration coefficients) that determine the learning ability of each iteration of the algorithm.

**3.4. Radio Environment Map Field Strength Estimation Algorithm.** In this paper, an improved Kriging estimation algorithm for radio environment map parameters is proposed using the new variogram in (6) and the modified theoretical variogram model in (9) and (11). The algorithm includes the following main steps: (i) calculating the value of the variogram by sampling data; (ii) fitting the theoretical variogram

*Step 1*

- (1) Calculate distance matrix  $D[d_{ij}]$  and variogram value matrix  $G[g_{ij}]$ , where  $d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$  and  $g_{ij} = (z_i - z_j)^2$ .
  - (2)  $\text{lag} = \min(D)$ ,  $\text{lag\_max} = \text{round}(\max(D)/2/\text{lag})$ ,  $h = 1 : \text{lag} : \text{lag\_max}$ ,  $\text{LAGS} = \text{round}(D/\text{lag})$ .
  - (3) Calculate experimental variogram value  $R(i)$ ,
- ```

for i = 1 : lag_max
    h(i) = i * lag;
    SEL = (LAGS == i);
    N(i) = sum(sum(SEL == 1));
    R(i) = sqrt(sum(G(SEL))/(2 * N(i)));
end for

```

*Step 2*

The PSO algorithm is used to fit the theoretical variogram models  $\gamma_{itm}$  and  $\gamma_{itm_f}$  using the theoretical variogram models in Equations (9) and (12),

- (1) Initialization: the position and velocity of a particle in  $d$ -dimensional problem space is randomly generated.
- (2) Evaluation of particles: the fitness value of each particle is calculated using Equation (14).
- (3) Updating  $g_{best}$ : the particle fitness values are compared with the population optimal value  $g_{best}$ , and if the current value is better than  $g_{best}$ , the position of  $g_{best}$  is set to the current particle position.
- (4) Updating the particle: the velocities and positions of all particles are updated using Equations (15).
- (5) Stop condition: return to step (2) for  $S_{max}$  iterations to obtain optimal parameters  $a_1, a_2, a_3$ , and  $a_4$ . Then, use Equations (9) and (11) to obtain theoretical variogram models  $\gamma_{itm}$  and  $\gamma_{itm_f}$ .

*Step 3*

- (1)  $K$  and  $\kappa_t$  are calculated using the theoretical variogram model.
- (2) Equation (4) is written in matrix form  $K\lambda = \kappa_t$  to get  $\lambda = K^{-1}\kappa_t$ .
- (3) The estimated value is calculated by Equation (2):  $Z_t^* = \sum_{i=1}^n \lambda_i \cdot Z_i$ .

ALGORITHM 1: Radio environment map field strength estimation.

curve equation using PSO; and (iii) calculating the test weight parameters using the theoretical variogram curve equation. The complete process is shown in Algorithm 1. The inputs of the algorithm are sample point coordinates  $x_{n*1}$  and  $y_{n*1}$ , sampling value  $z_{n*1}$ , and coordinates  $(x_t, y_t)$  of the point  $t$  to be estimated. The output of the algorithm is the estimated value  $Z_t^*$ . In the algorithm, lag is the basic lag distance, lag\_max is the maximum multiple of the lag distance, and matrix vectors  $K$ ,  $\lambda$ , and  $\kappa_t$  are expressed, respectively, as follows:

$$K = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}, \quad (16)$$

$$\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix},$$

$$\kappa_t = \begin{bmatrix} \gamma_{1t} \\ \vdots \\ \gamma_{nt} \\ 1 \end{bmatrix},$$

where  $\gamma_{ij}$  is the value of the variogram between sampling points  $x_i$  and  $x_j$  and  $\mu$  is the Lagrange constant.

## 4. Experimental Classification Results and Analysis

In this paper, the proposed algorithm was compared with three kinds of mainstream algorithms, which are IDW [14–17, 20, 21, 24], Spline [16, 20, 25], and Kriging [11–21]. We tested them on two kinds of data sets and analyzed the performance of the algorithms through a variety of objective evaluation indexes.

*4.1. Objective Evaluation Indexes.* Five kinds of objective evaluation indexes are used to compare and analyze the estimation results of the various algorithms for the parameters of radio environment map, which are the maximum error (MAX\_ERR), the average error (AVE\_ERR), the average estimation error percentage (PAEE), the relative mean square error (RMSE), and the root mean square error (RMSPE).

## (1) PAEE

$$\text{PAEE} = \frac{1}{n \times \bar{z}} \sum_{i=1}^n (Z'_i - Z_i)^2 \times 100\%, \quad (17)$$

where  $\bar{z}$  is the mean value of all the sampling points,  $Z'_i$  is the predicted value at the position  $i$ , and  $Z_i$  is the sample value at the position  $i$ .

## (2) RMSE

$$\text{RMSE} = \frac{1}{n \times S^2} \sum_{i=1}^n (Z'_i - Z_i)^2, \quad (18)$$

where  $S^2$  is the variance of all sample data.

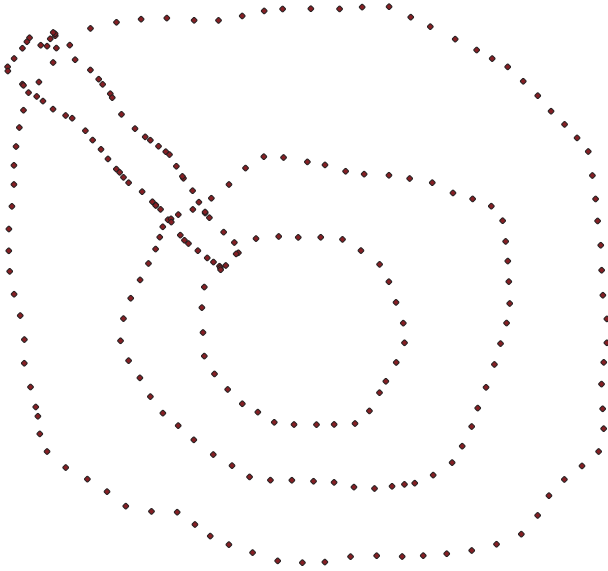


FIGURE 1: Data sampling position distribution map.

## (3) RMSPE

$$\text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z'_i - Z_i)^2}. \quad (19)$$

**4.2. Experimental Data.** Two sets of data are used to validate all these methods. One is real measured level data of FM radio FM 99.8 MHz and the max level data of bands 87–108 MHz and 1800–1900 MHz. The measuring terminal is a vehicle radio monitoring receiver, and the measurement location is located in Chengdu. The distribution of the sampling points was shown in Figure 1, and the specific parameters for real measured data were shown in Table 1. Another is radio signal simulation data which using the free space propagation model and the shadow model is a log normal model. The specific parameters for simulation data were shown in Table 2.

The real data set contains a total of 256 sampling points, and the simulation data set contains 1024 level data with range of 100 square kilometers. In order to compare and analyze the estimation results of different algorithms at different sampling granularities, we used 1/2 and 1/4 of the total data as the training data and the remaining data as the validation data. That is, when we used 1/4 data as the training data, the remaining 3/4 data was the validation test data.

**4.3. Analysis of Experiment Results**

**(1) 99.8 MHz Real Sampling Data.** This is the comparative testing on the actual acquisition level data of frequency 99.8 MHz. One-half of the data (128 sampling points in total) is used for training the model, and the remaining 1/2 of the data is used for testing and verification. The 5 evaluation results of the five kinds of algorithms are shown in Table 3.

TABLE 1: Parameters for real measured data.

|                           |                      |
|---------------------------|----------------------|
| Measurement area          | 28 Km * 28 Km        |
| Radio type                | FM radio band        |
| Broadcast                 | 99.8 MHz, 87–108 MHz |
| Number of sources         | 1 or N               |
| The average vehicle speed | About 80 Km/h        |

TABLE 2: Parameters for simulation data.

|                   |                  |
|-------------------|------------------|
| Calculated area   | 10 Km * 10 Km    |
| Numbers of points | 1024             |
| Carrier frequency | 101.7 MHz        |
| Path loss model   | Free space model |
| Shadow model      | Log normal model |

TABLE 3: Estimation results of level values for 99.8 MHz (1/2 training data).

|                  | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------|----------------|---------------|---------------|---------------|---------------|
| IDW              | 19.6973        | 6.3338        | 7.9729        | 0.5122        | 2.0172        |
| Spline           | 21.2641        | 4.2668        | 5.583         | 0.2512        | 0.9893        |
| Kriging          | 12.7859        | 3.6714        | 4.5996        | 0.1705        | 0.6714        |
| $\gamma_{itm}$   | 12.6277        | 3.6488        | 4.5090        | 0.1638        | 0.6452        |
| $\gamma_{itm_f}$ | <b>12.5643</b> | <b>3.5610</b> | <b>4.4359</b> | <b>0.1586</b> | <b>0.6244</b> |

TABLE 4: Estimation result of level values for 99.8 Mhz (1/4 training data).

|                  | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------|----------------|---------------|---------------|---------------|---------------|
| IDW              | 26.2765        | 6.3642        | 8.1205        | 0.5381        | 2.0736        |
| Spline           | 57.8689        | 4.9106        | 7.7875        | 0.4949        | 1.9070        |
| Kriging          | 16.0234        | 4.0510        | 5.1165        | 0.2136        | 0.8232        |
| $\gamma_{itm}$   | 16.7655        | 4.0749        | 5.0926        | 0.2116        | 0.8155        |
| $\gamma_{itm_f}$ | <b>15.8968</b> | <b>3.9979</b> | <b>5.0414</b> | <b>0.2070</b> | <b>0.8040</b> |

A quarter of the data (64 samples in total) is used to train the model, and the remaining 3/4 of the data is used for test validation. The results are shown in Table 4.

In order to reflect the estimation results of various algorithms directly, all the five kinds of algorithms use the same 1/4 training data, the results of which are compared with the same measured data. The comparisons of various algorithms were shown in Figure 2.

In the 1/2 data for training, there are 128 sampling points in the range of about 784 square kilometers. A sampling point covers an average of about 6 square kilometers. As could be seen from Table 3, the results of the prediction of our two methods and the Kriging method are relatively close, and the worst method is the IDW. For Spline and IDW, the former is significantly better than the latter on all indexes except the maximum error. The  $\gamma_{itm_f}$  method of our methods has achieved the best results on all the evaluation indexes. Compared with other methods, our methods have obvious advantages both in prediction accuracy and in prediction stability. The best average estimation error of our algorithms was 3.561Db. According to the research results of [6], it

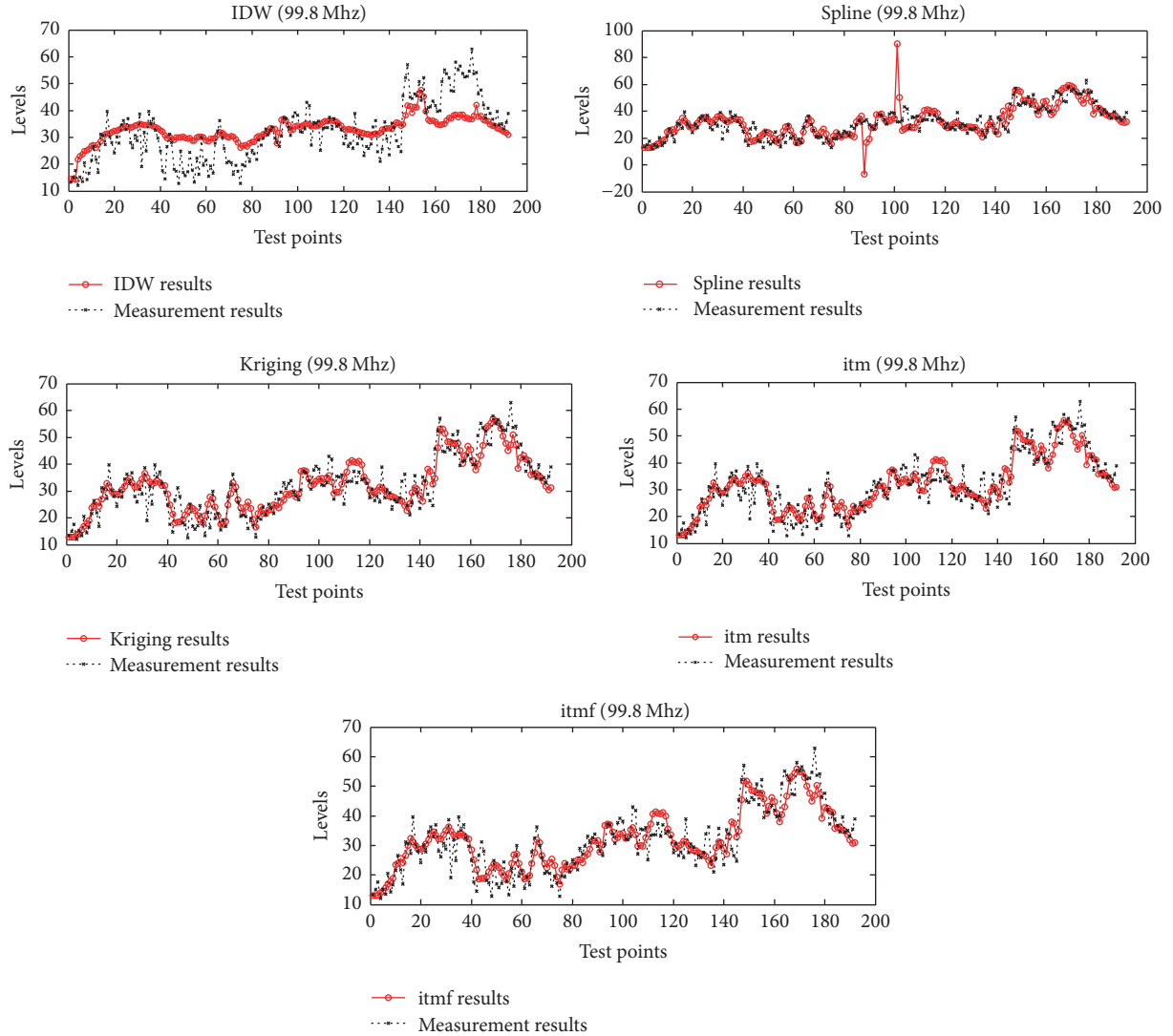


FIGURE 2: Estimation result of level values of 99.8 MHz (1/4 training data).

shows that our methods are very competitive. While in the 1/4 data for training, a sampling point covers about 12 square kilometers. From the results in Table 4, the methods based on Kriging system have better prediction and estimation effects. In particular, the prediction errors of Kriging based methods are all about 4 Db, indicating the effectiveness of these methods. Compared with the 1/2 training data, the increase of maximum error is more obvious. The second experiment showed that  $\gamma_{itmf}$  still has the best prediction results in all algorithms. As could be seen from Figure 2, although the IDW had obvious smoothing effect reflecting the overall trend of the data distribution, the prediction accuracy is the lowest and the estimation errors are very large in the ranges of 40–60 and 160–180. The prediction of Spline has two obvious outlier points, which indicated that the algorithm is sensitive to noise data. Compared with Kriging based methods, our methods yield significant improvements in the range of 140–180.

TABLE 5: Estimation results of max level values for 87–108 MHz (1/2 training data).

|                 | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|-----------------|----------------|---------------|---------------|---------------|---------------|
| IDW             | 12.2333        | 3.3063        | 4.3205        | 0.6093        | 0.2982        |
| Spline          | 12.3468        | 2.9376        | 3.8746        | 0.4900        | 0.2398        |
| Kriging         | 11.8977        | 2.6554        | 3.4731        | 0.3937        | 0.1927        |
| $\gamma_{itm}$  | 10.8254        | 2.5374        | 3.2675        | 0.3485        | 0.1706        |
| $\gamma_{itmf}$ | <b>10.8099</b> | <b>2.5368</b> | <b>3.2658</b> | <b>0.3481</b> | <b>0.1704</b> |

(2) 87–108 MHz Maximum Level of Real Sampling Data. This is a comparative testing on the actual acquisition max level data of band 87–108 MHz. One-half of the data (128 sampling points in total) was used for training the model, and the remaining was used for testing and verification. The 5 evaluation results of the five kinds of algorithms are shown in Table 5.

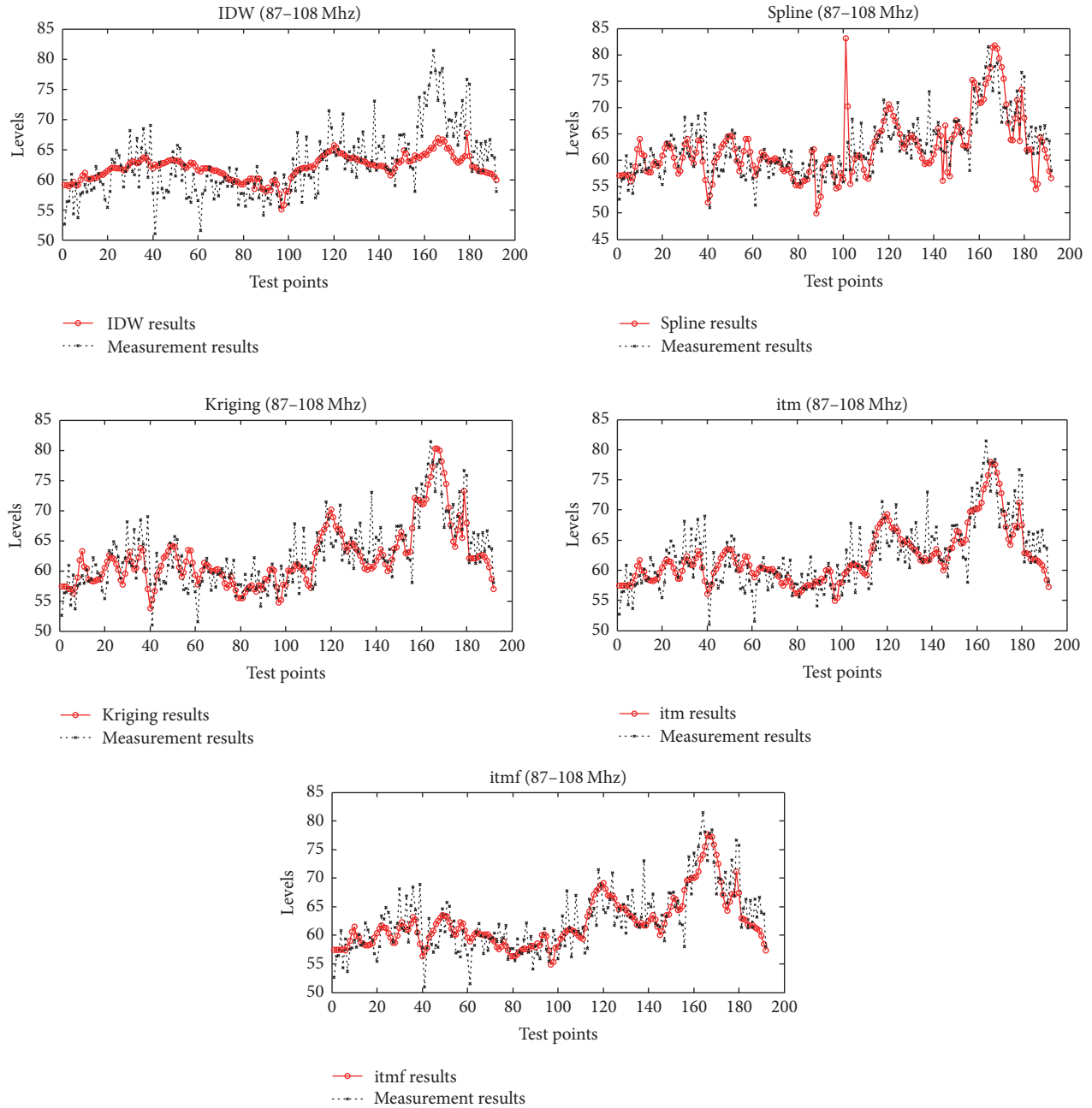


FIGURE 3: Estimation results of max level values of 87-108 MHz (1/4 training data).

A quarter of the data (64 samples in total) is used to train the model, and the remaining 3/4 of the data is used for test validation. The results are shown in Table 6.

In order to reflect the estimation results of various algorithms directly, all the five kinds of algorithms use the same 1/4 training data, the results of which are compared with the same measured data. The comparisons of various algorithms are shown in Figure 3.

The maximum signal strength in the frequency band is an important parameter of the radio environment map. In the large scale space, many signal sources constitute the spatial distribution of the maximum signal strength, so it is difficult to make estimation by using the radio propagation model.

TABLE 6: Estimation results of max level values for 87-108 MHz (1/4 training data).

|                 | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|-----------------|----------------|---------------|---------------|---------------|---------------|
| IDW             | 16.3241        | 3.3529        | 4.4390        | 0.6374        | 0.3152        |
| Spline          | 24.8394        | 3.1712        | 4.3034        | 0.5990        | 0.2962        |
| Kriging         | 12.6690        | 2.7459        | 3.4564        | 0.3864        | 0.1911        |
| $\gamma_{itm}$  | <b>10.2437</b> | 2.6877        | 3.4220        | 0.3788        | 0.1873        |
| $\gamma_{itmf}$ | 11.2699        | <b>2.6453</b> | <b>3.3709</b> | <b>0.3675</b> | <b>0.1817</b> |

The spatial interpolation method is more advantageous. The experimental results of 1/2 training data are shown in Table 5.



TABLE 7: Estimation results of max level values for 1800–1900 MHz (1/2 training data).

|                        | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------------|----------------|---------------|---------------|---------------|---------------|
| IDW                    | 22.4790        | 5.6729        | 7.5337        | 0.6532        | 1.3643        |
| Spline                 | 18.7473        | 4.9472        | 6.6059        | 0.5022        | 1.0489        |
| Kriging                | <b>17.2293</b> | 4.7758        | 6.3212        | 0.4598        | 0.9605        |
| $\gamma_{itm}$         | 18.2057        | 4.6203        | 6.1744        | 0.4387        | 0.9164        |
| $\gamma_{itm\text{f}}$ | 18.3114        | <b>4.6199</b> | <b>6.1344</b> | <b>0.4331</b> | <b>0.9046</b> |

TABLE 8: Estimation results of max level values for 1800–1900 MHz (1/4 training data).

|                        | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------------|----------------|---------------|---------------|---------------|---------------|
| IDW                    | 28.2658        | 5.9683        | 7.9227        | 0.7184        | 1.5180        |
| Spline                 | 30.1425        | 6.7523        | 9.0084        | 0.9287        | 1.9626        |
| Kriging                | 28.8990        | 6.2277        | 8.2594        | 0.7807        | 1.6498        |
| $\gamma_{itm}$         | 24.1963        | 5.5533        | 7.2744        | 0.6056        | 1.2797        |
| $\gamma_{itm\text{f}}$ | <b>23.9539</b> | <b>5.4968</b> | <b>7.2053</b> | <b>0.5942</b> | <b>1.2555</b> |

As could be seen from the table, all the methods have got good results. Surprisingly, compared to the radiation estimation of single source in experiments (1), the estimation of the maximum signal strength had a better accuracy, which was about 10 Db. Furthermore, for the average prediction error, the results of our methods are close to 2.6 Db. Similarly, for all the evaluation indexes, the proposed method  $\gamma_{itm\text{f}}$  achieves the best results, and our method increases about 10% compared to the Kriging. For the estimation of the maximum signal strength in frequency band, the prediction accuracy is not significantly reduced with the training data reduction (seen from Table 6). For the maximum error, the  $\gamma_{itm}$  algorithm gets the best result, and the  $\gamma_{itm\text{f}}$  algorithm achieves the best result on the rest of indexes. The Spline and IDW have a large maximum error. As could be seen from Figure 3, the Spline exhibits a significant error, which is consistent with the previous conclusion that the method has a weak ability to overcome noise. The proposed methods have the best prediction and estimation effect.

### (3) 1800–1900 MHz Maximum Level of Real Sampling Data.

This is a comparative testing on the actual acquisition max level data of band 1800–1900 MHz. One-half of the data (128 sampling points in total) is used for training the model, and the remaining is used for testing and verification. The 5 evaluation results of the five kinds of algorithms are shown in Table 7.

A quarter of the data (64 samples in total) is used to train the model, and the remaining 3/4 of the data is used for test validation. The results are shown in Table 8.

Same as the above, all the five kinds of algorithms use the same 1/4 training data, and the remaining data as the same measured data. The comparisons of various algorithms are shown in Figure 4.

The main services in the 1800–1900 MHz band are mobile communications, of which band has more complex electromagnetic environment, and the estimation of the parameters

TABLE 9: Estimation results of level values for 101.7 MHz (1/2 training data).

|                        | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------------|----------------|---------------|---------------|---------------|---------------|
| IDW                    | 26.6984        | 3.4255        | 4.4763        | 0.6184        | 0.3720        |
| Spline                 | 15.5874        | 3.0751        | 3.9084        | 0.4715        | 0.2836        |
| Kriging                | <b>12.1004</b> | 2.7197        | 3.3944        | 0.3556        | 0.2139        |
| $\gamma_{itm}$         | 12.8846        | 2.5712        | 3.2100        | 0.3180        | 0.1913        |
| $\gamma_{itm\text{f}}$ | 13.5290        | <b>2.5613</b> | <b>3.1978</b> | <b>0.3156</b> | <b>0.1899</b> |

TABLE 10: Estimation results of level values for 101.7 MHz (1/4 training data).

|                        | MAX_ERR        | AVE_ERR       | RMSPE         | RMSE          | PAEE          |
|------------------------|----------------|---------------|---------------|---------------|---------------|
| IDW                    | 26.6254        | 3.3326        | 4.4070        | 0.6072        | 0.3603        |
| Spline                 | 23.1643        | 3.6385        | 4.9289        | 0.7596        | 0.4507        |
| Kriging                | <b>11.0104</b> | 2.8034        | 3.5230        | 0.3881        | 0.2303        |
| $\gamma_{itm}$         | 11.5091        | 2.7089        | 3.3900        | 0.3593        | 0.2132        |
| $\gamma_{itm\text{f}}$ | 13.4562        | <b>2.6326</b> | <b>3.2835</b> | <b>0.3371</b> | <b>0.2000</b> |

of this band is more difficult. The experimental results of 1/2 training data are shown in Table 7. As could be seen from the table, all the methods have good results. The method Kriging has the best performance for MAX\_ERR index. For the remaining evaluation indexes, our method  $\gamma_{itm\text{f}}$  obtained the best results. In general, the results of the three methods of Kriging,  $\gamma_{itm}$ , and  $\gamma_{itm\text{f}}$  are relatively close, and the improved methods based on propagation model are superior to the traditional Kriging method. When the training set is reduced by half, the performance of all methods is significantly reduced (seen from Table 8). Among them, the predication performance of the Kriging based methods is reduced by nearly 20%. This is because the power of the mobile communication station is small, and the transmission distance is close, which causes the correlation between the sparse sampling points to be weak. However, in this case, our approach is significantly better than the traditional Kriging based approach, with all the indicators increasing by nearly 15%. It is worth mentioning that the IDW algorithm also obtains good results, indicating that the algorithm has good stability. As could be seen from Figure 4, the data set is difficult to predict, and in some individual positions, all methods have large errors. But the proposed methods have the best prediction and estimation effect, especially in the vicinity of the test point numbered 140, and the proposed methods are significantly better than the results of the Kriging method.

(4) 101.7 MHz Simulation Data. This was a comparative testing on the actual acquisition level data of frequency 101.7 MHz. One-half of the data (512 sampling points in total) is used for training the model and the remaining is used for testing and verification. The 5 evaluation results of the five kinds of algorithm are shown in Table 9.

A quarter of the data (256 samples in total) is used to train the model, and the remaining (768 testing points) is used for testing and validation. The results are shown in Table 10.

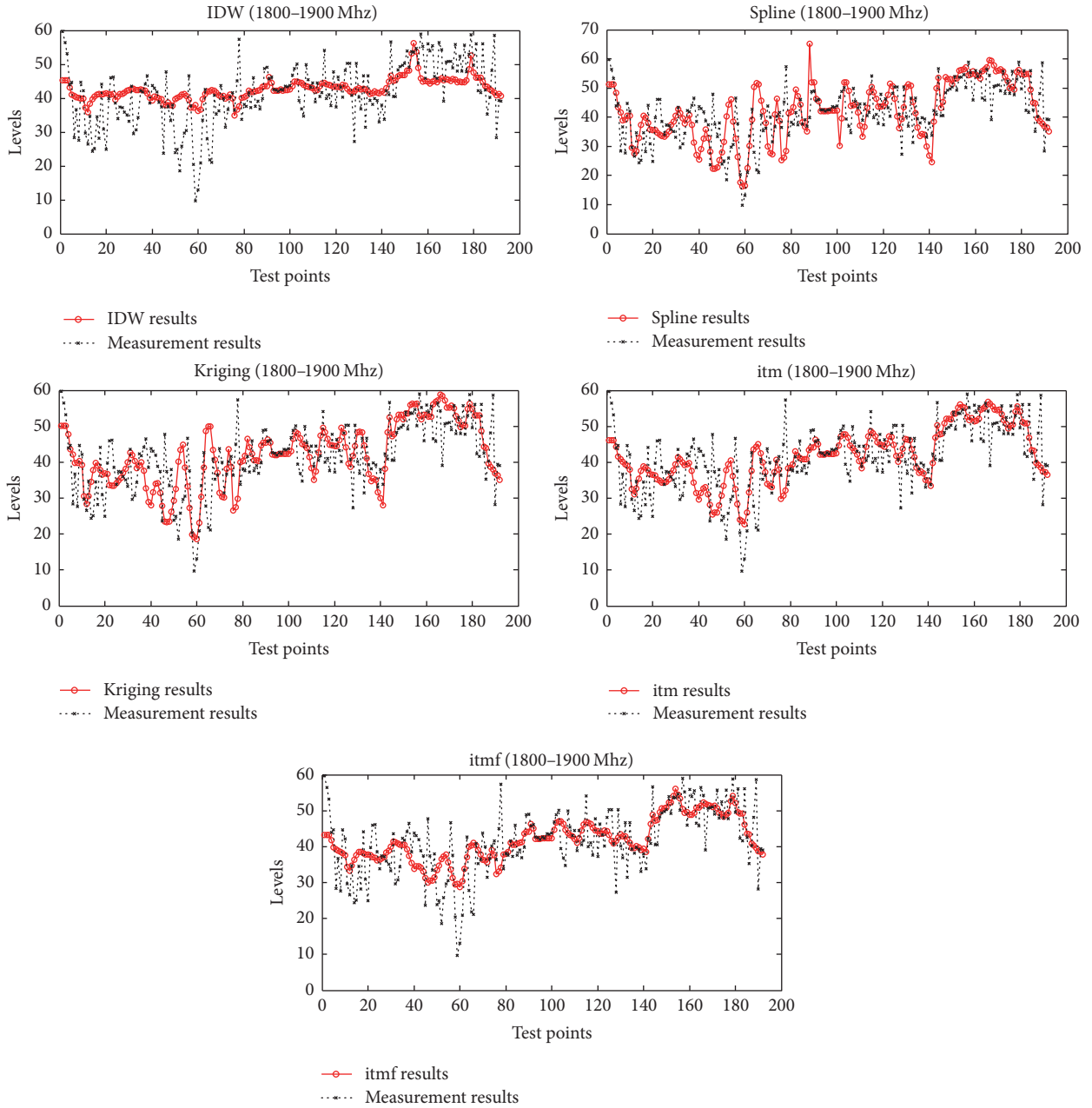


FIGURE 4: Estimation results of max level values of 1800–1900 MHz (1/4 training data).

In order to reflect the estimation results of various algorithms directly, all the five kinds of algorithms use the same 1/4 training data, the results of which are compared with the same measured data. The comparisons of various algorithms are shown in Figure 5.

We use 1/2 the simulation data for training model in our experiment, and each sampling point covers about 0.5 square kilometers. As can be seen from Table 9, the IDW has a very high maximum prediction error, and the Kriging obtains the best maximum error evaluation, while for the rest of the evaluation indexes, the  $\gamma_{itm}$  is about 10% higher than the

Kriging and about 20% higher than Spline and IDW. As can be seen from Table 10, the Kriging yields the best maximum error valuation, while, for the rest of the evaluation indexes, the  $\gamma_{itm}$  method obtains the best results. In 1/4 training data situation, the predictive performance of Spline declines more severely. As can be seen from Figure 5, IDW has a significant prediction error, while methods based on Kriging system have better prediction accuracy. In all compared methods, the proposed methods based on propagation model have achieved the highest accuracy. Because of the introduction of the shadow model, it is difficult to have a high prediction

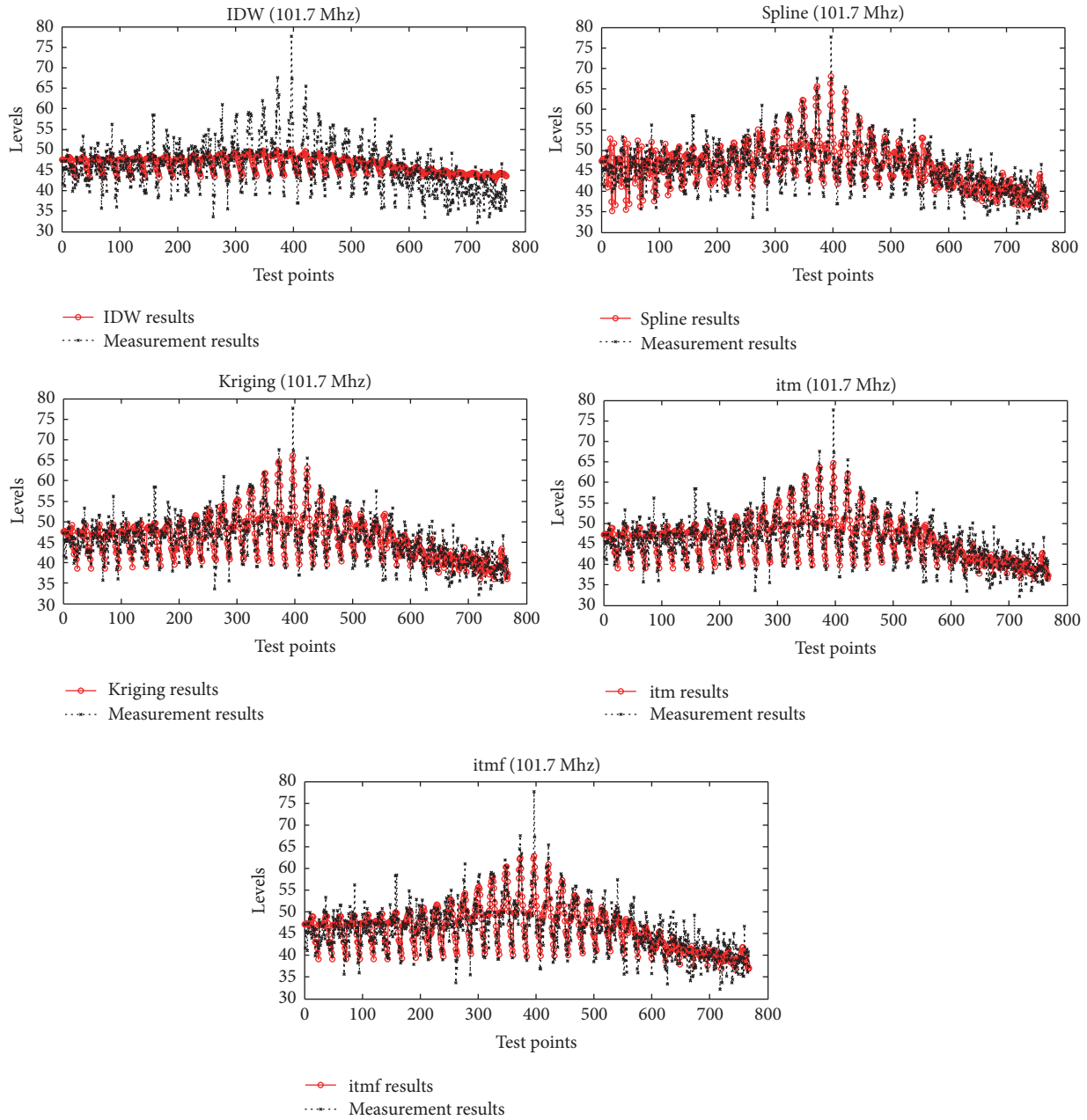


FIGURE 5: Estimation results of level values of 101.7 MHz (1/4 training data).

and estimation on the simulated data. It can be seen that the prediction accuracy of the simulation data is equivalent to that of the maximum signal strength of the frequency band.

### 5. Conclusion

In this paper, we proposed a spatial distribution prediction method for radio environment map parameters. It was shown that the IDW, Spline, and Kriging methods are the most effective methods to solve this problem, and, of these, Kriging is the best method. The main parameters of the radio environment map are the signal strength and other parameters are affected by it. In this study, based on the

Kriging approach, the definition of a variogram was improved based on the loss characteristics of radio propagation. A new variogram theoretical model was proposed in combination with a radio propagation model. Based on the characteristics of data sampling and signal propagation, a new weighted fitting method for variograms was also proposed. The new method is more suitable for the actual characteristics of radio environment map parameter prediction. Moreover, the proposed model is better adapted to the spatial correlation of radio environment parameters and has better prediction accuracy. Experiments on the signal strength data of a single frequency and the maximum signal strength data of a frequency band and simulation data prove these conclusions.

The evaluation indexes of our method are improved by about 10% on average compared with those of the conventional Kriging method.

Radio signal propagation in space is a complex process. There are some intractable problems such as building shadows, same-frequency adjacent channel interference, and multipath propagation. Therefore, to obtain better prediction accuracy, the terrain and station information to the model will be taken into consideration in our future research.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

Zhisheng Gao proposed the original idea, Yaoshun Li wrote the paper under the guidance of Zhisheng Gao, Yaoshun Li and Chunzhi Xie designed the experiment and provided all the figures, and Chunzhi Xie checked the manuscript. All authors read and approved the final manuscript.

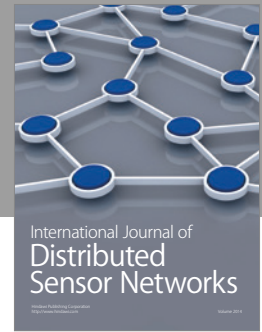
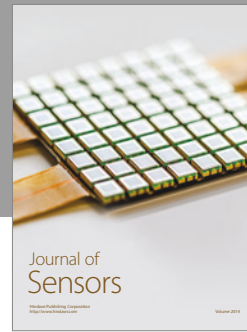
## Acknowledgments

This work has been partially supported by the Major Project of Education Department of Sichuan (Grant no. 14ZA0118), the Ministry of Education Chunhui project (Grant no. Z2016149), the Xihua University Key Laboratory Development Program (Grant nos. szjj2017-065, s2jj2014-050), and the Graduate Innovation Foundation of Xihua University (ycjj2017070).

## References

- [1] Y. Zhao, L. Morales, J. Gaeddert, K. K. Bae, J.-S. Um, and J. H. Reed, "Applying radio environment maps to cognitive wireless regional area networks," in *Proceedings of the 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, vol. 1, pp. 115–118, April 2007.
- [2] H. Yilmaz Birkan, T. Tugcu, F. Alagöz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 162–169, 2013.
- [3] S. Sodagari, "A Secure Radio Environment Map Database to Share Spectrum," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1298–1305, 2015.
- [4] N. Ezzati, H. Taheri, and T. Tugcu, "Optimised sensor network for transmitter localisation and radio environment mapping," *IET Communications*, vol. 10, no. 16, pp. 2170–2178, 2016.
- [5] J. Perez-Romero, A. Zalonis, L. Boukhatem et al., "On the use of radio environment maps for interference management in heterogeneous networks," *IEEE Communications Magazine*, vol. 53, no. 8, pp. 184–191, 2015.
- [6] A. Galindo-Serrano, B. Sayrac, S. Ben Jemaa, J. Riihijärvi, and P. Mähönen, "Automated coverage hole detection for cellular networks using radio environment maps," in *Proceedings of the 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, WiOpt 2013*, vol. 1, pp. 35–40, May 2013.
- [7] V. Atanasovski, J. Van De Beek, A. Dejonghe et al., "Constructing radio environment maps with heterogeneous spectrum sensors," in *Proceedings of the IEEE International Symposium on Dynamic Spectrum Access Networks, DySPAN 2011*, vol. 3, pp. 660–661, May 2011.
- [8] F. Paisana, Z. Khan, J. Lehtomaki, L. A. Dasilva, and R. Vuoltoniemi, "Exploring radio environment map architectures for spectrum sharing in the radar bands," in *Proceedings of the 23rd International Conference on Telecommunications, ICT 2016*, vol. 1, May 2016.
- [9] S. Subramani, T. Farnham, and M. Sooriyabandara, "Deployment and interface design considerations for radio environment maps," in *Proceedings of the IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2012*, vol. 1, pp. 480–487, October 2012.
- [10] J. Ojaniemi, J. Kalliovaara, A. Alam, J. Poikonen, and R. Wichman, "Optimal field measurement design for radio environment mapping," in *Proceedings of the 47th Annual Conference on Information Sciences and Systems, CISS 2013*, March 2013.
- [11] M. Pesko, T. Javornik, A. Košir, M. Štular, and M. Mohorčič, "Radio environment maps: The survey of construction methods," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 11, pp. 3789–3809, 2014.
- [12] M. Pesko, T. Javornik, L. Vidmar, A. Košir, M. Štular, and M. Mohorčič, "The indirect self-tuning method for constructing radio environment map using omnidirectional or directional transmitter antenna," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, pp. 1–12, 2015.
- [13] H. B. Yilmaz and T. T. Bogazici, "Location estimation-based radio environment map construction in fading channels," *Wireless Communications and Mobile Computing*, vol. 15, no. 3, pp. 561–570, 2015.
- [14] S. Ulaganathan, D. Deschrijver, M. Pakparvar et al., "Building accurate radio environment maps from multi-fidelity spectrum sensing data," *Wireless Networks*, vol. 22, no. 8, pp. 2551–2562, 2016.
- [15] G. Zou, K. Xue, D. Huang, C. Su, and J. Sun, "The comparison and study of small sample data spatial interpolation accuracy," in *Proceedings of the 6th International Conference on Natural Computation, ICNC'10*, vol. 4, pp. 1897–1900, August 2010.
- [16] M. Azpurua and K. dos Ramos, "A comparison of spatial interpolation methods for estimation of average electromagnetic field magnitude," *Progress in Electromagnetics Research M*, vol. 14, pp. 135–145, 2010.
- [17] M. Angjelinoski, V. Atanasovski, and L. Gavrilovska, "Comparative analysis of spatial interpolation methods for creating radio environment maps," in *Proceedings of the 19th Telecommunications Forum, TELFOR 2011*, vol. 1, pp. 334–337, November 2011.
- [18] K. Sato and T. Fujii, "Kriging-Based Interference Power Constraint: Integrated Design of the Radio Environment Map and Transmission Power," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 1, pp. 13–25, 2017.
- [19] J. Riihijärvi, P. Mähönen, M. Wellens, and M. Gordziel, "Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields," in *Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008*, vol. 1, September 2008.
- [20] D. Denkovski, V. Atanasovski, L. Gavrilovska, J. Riihijärvi, and P. Mähönen, "Reliability of a radio environment map: Case of spatial interpolation techniques," in *Proceedings of the 7th International ICST Conference on Cognitive Radio Oriented*

- Wireless Networks and Communications, CROWNCOM 2012*, vol. 1, pp. 248–253, June 2012.
- [21] C. Phillips, M. Ton, D. Sicker, and D. Grunwald, “Practical radio environment mapping with geostatistics,” in *Proceedings of the IEEE International Symposium on Dynamic Spectrum Access Networks, DYSPAN 2012*, vol. 1, pp. 422–433, October 2012.
- [22] Z. Wei, Q. Zhang, Z. Feng, W. Li, and T. A. Gulliver, “On the construction of Radio Environment Maps for Cognitive Radio Networks,” in *Proceedings of the IEEE Wireless Communications and Networking Conference, WCNC 2013*, vol. 1, pp. 4504–4509, April 2013.
- [23] X. H. Li, F. Yuan, C. Jia, M. M. Zhang, and T. F. Zhou, “Comparison of Geostatistical Interpolation Methods for Local Singularity Exponent Calculation,” *Scientia Geographica Sinica*, vol. 32, no. 2, pp. 136–142, 2012.
- [24] H. B. Yilmaz, *Cooperative spectrum sensing and radio environment map construction in cognitive radio networks*, Diss. Bogaziçi University, 2012.
- [25] G. Mateos, J.-A. Bazerque, and G. B. Giannakis, “Spline-based spectrum cartography for cognitive radios,” in *Proceedings of the Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 1025–1029, November 2009.



**Hindawi**

Submit your manuscripts at  
<https://www.hindawi.com>

