IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 5, 2019, accepted July 10, 2019, date of publication July 15, 2019, date of current version July 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928832

On the Performance of Neural Network Residual Kriging in Radio Environment Mapping

KOYA SATO[1], KEI INAGE[2], AND TAKEO FUJII[3], (Member, IEEE)

[1]Department of Electrical Engineering, Tokyo University of Science, Tokyo 125-8585, Japan
[2]Tokyo Metropolitan College of Industrial Technology, Tokyo 140-0011, Japan
[3]Advanced Wireless and Communication Research Center (AWCC), The University of Electro–Communications, Tokyo 182-8585, Japan

Corresponding author: Koya Sato (k_sato@ieee.org)

This work was supported by the Japan Society for the Promotion of Science through KAKENHI under Grant 16KK0124, Grant 16H02344, Grant 18H01439, and Grant 19K14988.

ABSTRACT This paper addresses the research question: *can feedforward neural network (FFNN)-based path loss modeling improve the accuracy of Kriging?* Radio propagation factors, which consist of path loss and shadowing, can accurately be obtained via crowdsourcing with Kriging. In most works on Kriging-aided radio environment mapping, measurement datasets are first regressed via linear path loss modeling to ensure spatial stationarity of the shadowing. However, in practical situations, the path loss often contains an anisotropy owing to terrain and obstacle effects. Thus, Kriging may not perform an optimal interpolation because of the errors in path loss modeling. In this paper, an FFNN is used for path loss modeling. Then, ordinary Kriging is applied to interpolate the shadowing. We first evaluate the performance of this method in a case where the transmitter is fixed. It is shown that this method does not improve Kriging in a large-scale and fixed transmitter system; although the FFNN outperforms OLS in path loss modeling. Then, this method is extended to distributed wireless networks where transmitters are arbitrarily located, such as in mobile ad hoc networks (MANETs) and vehicular ad hoc networks (VANETs). The results of a measurement-based experiment show that the FFNN is capable of improving Kriging in such a distributed network case.

INDEX TERMS Radio propagation, spatial statistics, crowdsourcing, neural network, Kriging, regression analysis.

I. INTRODUCTION

The growth in the demand for mobile communication systems has exponentially increased data traffic during the last decade. Because this exponential growth consumes finite spectrum resources, traditional spectrum utilization according to exclusive policies for spectrum allocation faces certain limitations. In addition to a shortage in frequencies exclusively assigned to the new system, problems such as the degradation of communication quality arise owing to traffic congestion in wireless local access network (WLAN) and cellular systems.

Currently, there are two countermeasures used to address such a problem:

- Dynamic spectrum sharing with existing systems [1].
- Improvement of spectrum utilization efficiency in existing systems via sophisticated low layer techniques (e.g., non-orthogonal multiple access (NOMA) [2]).

The associate editor coordinating the review of this manuscript and approving it for publication was Guan Gui.

In both these measures, it is important to improve the communication efficiency over the spatial domain. Because this requires accurate estimation of the available spectrum and the interference between transmitters, radio environment estimation is a key technology for the sustainable development of wireless communication systems.

A. KRIGING-BASED RADIO ENVIRONMENT MAP

One simple method of radio environment estimation for a system with a fixed transmitter is the use of the *radio environment map* (REM) [3]–[5]. An REM is generally defined as a map that indicates the average received power for each location. By storing an REM that has been constructed in advance based on actual measured values and accessing it via a cloud server, we can accurately predict the surrounding radio environment.

Because of its practicality, many researchers have investigated embedding an REM into wireless systems over the last decade, e.g., spectrum sharing over television white space

(TVWS) [6], coverage prediction in cellular networks [7], and communication quality prediction in WLANs [8].

Because the accuracy of the REM is directly related to the utilization efficiency of the spectrum [9], [10], many theoretical and experimental studies have been conducted. In particular, it is well known that *Kriging*, a spatial interpolation technique [11] with an actual measured dataset, can obtain an accurate REM.

Here, the average received signal power consists of two factors: the path loss and shadowing. Only path loss modeling cannot be used to estimate shadowing; thus, this method has a root mean squared error (RMSE) of approximately 6–8 [dB] [12], [13]. It is empirically known that shadowing has a spatial correlation [14]. Focusing on this characteristic, the shadowing can be estimated accurately by applying an appropriate weighted average to the observation dataset in the framework of spatial statistics.

Kriging is an interpolation method that minimizes the variance of estimation errors under the constraint of unbiased estimation. Kriging is subdivided according to the characteristics of a random variable because the optimality of the interpolation value is realized under the spatial constancy of the random variable. *Ordinary Kriging* is a fundamental Kriging that is often applied in REM construction. Ordinary Kriging realizes an optimal interpolation for the Gaussian process following a unique model in which both the expected value and semivariogram of the random variable have spatial stationarity.

B. PROBLEM OF KRIGING IN WIRELESS SYSTEMS

As the expected value of the received signal power has a spatial trend owing to the path loss, we first need to estimate the path loss and remove it from the dataset using regression analysis.

Path loss modeling often assumes a simple equation with a few parameters, for example,

$$L(d) = C + 10\eta \log_{10} d \quad [\mathrm{dB}], \tag{1}$$

where C [dB] is the constant factor, d [m] is the distance from the transmitter, and η is the path loss index. By applying ordinary least squares (OLS), we can model the path loss in a simple way. In many cases, such a tuned-path loss outperforms empirical models [13]. However, in fact, the path loss generally exhibits strong anisotropy because it is affected by several complex factors such as terrain, roads, and buildings. Additionally, this anisotropy becomes even stronger when the transmitter employs an array antenna. Therefore, path loss estimation assuming a simple model cannot fully obtain the spatial stationarity of the shadowing, and Kriging may not perform an optimal interpolation.

C. NEURAL NETWORK FOR RADIO PROPAGATION ESTIMATION AND OUR QUESTION

Some recent works have shown the outstanding performance of neural networks (NNs) in physical layers [15]–[18]. An NN is a well-known technique in the field of machine learning and is formed by propagating the output between multiple artificial neurons coupled together. By taking a weighted sum of activation functions appropriately designed based on a training dataset, a nonparametric estimation can be realized [19]. For example, the authors in [15] applied a deep recurrent NN for resource allocation in practical non-orthogonal multiple access systems. Additionally, reference [16] proposed a deep-NN-based fast beamforming design for downlink multiple-input and multiple-output (MIMO) systems.

As NNs can be used for regression analyses, they can help accurately model the path loss [20], [21]. From an optimization-based discussion, the authors in [20] concluded that a simple NN that employs a few (1–3) hidden layers can achieve a sufficient accuracy for path loss modeling. The authors in [21] evaluated the performance of a feedforward NN (FFNN) in path loss modeling over an ultrahigh frequency (UHF) band. The input layer of the FFNN proposed in [21] consisted of distance, frequency, information of structures, etc. From a measurement-based evaluation, they concluded that an FFNN can improve the accuracy of path loss modeling. Considering the problem of Kriging shown in Sect. I-B, these results raise a simple question: *can FFNN-based path loss modeling improve the accuracy of Kriging?*

D. OVERVIEW OF THIS PAPER AND MAIN CONTRIBUTIONS

In this study, to answer the above research question, we evaluate the practical performance of FFNN-aided Kriging in radio environment mapping. To this end, the path loss is first modeled with an FFNN from actual measured datasets. After the shadowing factors in the datasets are extracted, ordinary Kriging is performed for spatial interpolation. In the field of spatial statistics, such a procedure is called *neural network residual Kriging* (NNRK). NNRK was proposed in [22], where the authors applied it to analyze the air radiation dose and showed that FFNN-based regression can improve the accuracy of ordinary Kriging.

This study focuses on two situations as shown in Fig. 1(a)(b). We first consider a situation where a transmitter is fixed (Fig. 1(a)), such as in television and cellular systems; these are the main applications of the REM. Then, this discussion is extended to a more complex situations where transmitters are arbitrarily located, as shown at Fig. 1(b). This case considers mobile ad-hoc networks (MANETs), vehicular ad-hoc networks (VANETs), and device-to-device (D2D) communications. The main contributions of this study are summarized as follows.

- From a brief theoretical discussion and a numerical simulation, we first show that the effects of path loss modeling on Kriging strongly depend on the distance between the transmitter and the estimated location. Namely, as the communication distance decreases, the influence of errors in path loss modeling increases.

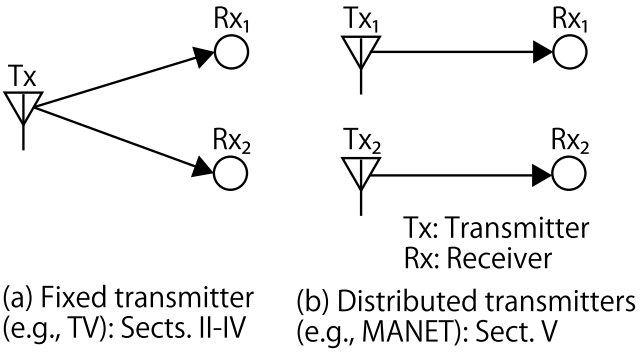

FIGURE 1. **This study focuses on two cases. The distributed network case consists of considerably more complex path loss compared with the fixed transmitter case; FFNN will significantly improve accuracy of path loss modeling in this situation.**

The following two contributions are based on this discussion (Sects. II-C and IV-A).

- It is shown that OLS with Kriging achieves a performance almost equal to that of NNRK in a large-scale and fixed transmitter system; although the FFNN outperforms OLS in path loss modeling. We demonstrate this using a dataset measured over TV bands (Sect. IV-B).
- We extend NNRK to a distributed case as shown in Fig. 1(b). The spatial correlation of shadowing can also be obtained in such a distributed situation [23]. Thus, Kriging-based radio environment mapping can be extended to such a case. Here, as the path loss exhibits considerably complex anisotropy according to the transmission location, we cannot develop an accurate model with an equation. Additionally, the communication distance is usually within a hundred meters; thus, the NN significantly improves the performance of Kriging. We show this using a dataset obtained from a vehicle-to-vehicle (V2V) communication system (Sect. V).

These discussions demonstrate how an FFNN can improve the performance of Kriging.

The reminder of this paper is organized as follows. In Sect. II, we introduce the principle of Kriging-based REM construction while focusing on a case where the transmitter is fixed. At the end of this section, we indicate the influence of path loss modeling in Kriging. Then, in Sect. III NNRK is described, and the performance of the NNRK is evaluated via both numerical simulation and experiment in Sect. IV. Then, Sect. V extends NNRK to a distributed situation. Finally, we conclude our work in Sect. VI.

II. OVERVIEW OF KRIGING-BASED RADIO ENVIRONMENT MAPPING

Methods for REM construction can be categorized into an empirical-model-based approach and crowdsourcing-based approach. In this paper, we focus on the latter approach because this usually outperforms the former in terms of accuracy. This section summarizes ordinary Kriging with OLS-based path loss modeling as a fundamental method in radio environment mapping. Note that this section considers the situation where a transmitter is fixed, as already shown in Fig. 1(a). In Sect. V, we extend this method to the distributed situation that is shown in Fig. 1(b).

A. RADIO PROPAGATION MODEL AND BRIEF PROCEDURE

We first summarize the task of Kriging based on a description of the radio propagation model. A block diagram of Kriging-based radio environment mapping is shown in Fig. 2. Although this study covers both fixed and distributed situations and both OLS and FFNN-based modeling, all methods in this study follow this diagram.

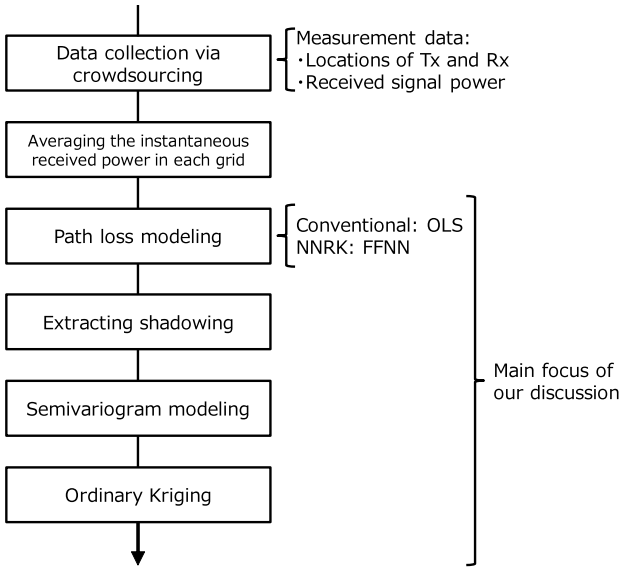

FIGURE 2. **Block diagram of radio environment mapping. Both fixed and distributed situations and both conventional Kriging (in Sect. II) and NNRK (in Sect. III) follow this procedure.**

The REM is usually stored in a cloud database. The database first collects information on the received signal power related to the received location. We define the REM as a map that stores two-dimensional coordinate information of the average received signal power focusing on a fixed transmitter. In this system, multiple nodes first measure the signal strength from a transmitter and report the data to a database. Then, the reported (instantaneous) values are averaged for each grid from several meters to several hundred meters to alleviate the effect of multipath fading. However, because the number of nodes and the area where observations can be performed are limited, the REM is missing teeth. Thus, we need to interpolate for missing-tooth information. Kriging is often used for this purpose.

In this section, we consider an REM construction where N values of the average received signals are given in the database. We define the dataset vector $\boldsymbol{y} = (P(\boldsymbol{x}_1), P(\boldsymbol{x}_2), \cdots, P(\boldsymbol{x}_N))^{\mathrm{T}}$, where $\boldsymbol{x}_i = (x_i, y_i)$ is the measurement location and $P(\boldsymbol{x}_i)$ [dBm] is the average received signal power. If the effect of multipath fading can be alleviated by averaging the instantaneous values, the average

received signal power at a given location $\boldsymbol{x}$ from the transmitter location $\boldsymbol{x}_{\mathrm{Tx}} = (x_{\mathrm{Tx}}, y_{\mathrm{Tx}})$ can be modeled as

$$P(\boldsymbol{x}) = \underbrace{P_C - L(\boldsymbol{x})}_{\triangleq \overline{P}(\boldsymbol{x})} + W(\boldsymbol{x}), \tag{2}$$

where P_C [dBm] is the location-independent scalar quantity that includes the transmission power, antenna gain, and the effect of frequency. $L(\boldsymbol{x})$ [dB] is the path loss between $\boldsymbol{x}$ and $\boldsymbol{x}_{\mathrm{Tx}}$. $W(\boldsymbol{x})$ [dB] is the log-normal shadowing with a standard deviation σ [dB]. As both L and W are location-dependent scalars [24], we express them as functions of the location vector $\boldsymbol{x}$. Here, we assume that the spatial correlation in shadowing follows a typical model [14],

$$\begin{aligned}\rho_{i,j} &= \frac{\mathrm{E}\left[\left(W(\boldsymbol{x_i}) - \mathrm{E}[W(\boldsymbol{x_i})]\right)\left(W(\boldsymbol{x_j}) - \mathrm{E}[W(\boldsymbol{x_j})]\right)\right]}{\sigma_i \sigma_j} \\ &\approx \exp\left(\frac{-||\boldsymbol{x}_i - \boldsymbol{x}_j|| \ln 2}{d_{\mathrm{cor}}}\right),\end{aligned} \tag{3}$$

where $||\cdot||$ is the Euclidean distance, σ_i is the standard deviation of $W(\boldsymbol{x}_i)$, d_{cor} [m] is the correlation distance, and $\mathrm{E}[\cdot]$ is the expected value of the random variable. The task of Kriging in this context is to interpolate the received signal power at an arbitrary location from the dataset $\boldsymbol{y}$.

B. ORDINARY KRIGING WITH OLS

Next, we describe ordinary Kriging with OLS, which has been well discussed in the community of spectrum sharing. Although Kriging is subdivided into some strategies according to the property of the target random variable Z, all methods interpolate by taking the weighted average of the measured values:

$$\hat{Z}(\boldsymbol{x}_0) = \sum_{i=1}^{N} \omega_i Z(\boldsymbol{x}_i), \tag{4}$$

where ω_i is the weight factor multiplied by $Z(\boldsymbol{x}_i)$, and $\hat{Z}(\boldsymbol{x}_0)$ is the interpolated $Z(\boldsymbol{x}_0)$. Kriging optimizes ω_i such that the variance of the estimation error is minimized.

Here, ordinary Kriging assumes that

- $\mathrm{E}[Z(\boldsymbol{x})] = const.$ over any $\boldsymbol{x}$.
- The spatial correlation property of $Z(\boldsymbol{x})$ is static over the entire measurement area.

Considering these assumptions, $Z(\boldsymbol{x})$ can be related to the shadowing $W(\boldsymbol{x})$ in Eq. (2). However, the received signal has a location-dependent trend factor $\overline{P}(\boldsymbol{x})$; i.e., $P(\boldsymbol{x}) \neq const.$. Thus, before applying ordinary Kriging, we need to estimate the model of $\overline{P}(\boldsymbol{x})$ and subtract it from $\boldsymbol{y}$ to gain the optimality of the interpolated value.

Usually, $\overline{P}(\boldsymbol{x})$ is assumed to follow a simple path loss model that depends on the distance $d \triangleq ||\boldsymbol{x} - \boldsymbol{x}_{\mathrm{Tx}}||$ [m],

$$\begin{aligned}\overline{P}(d) &= P_{\mathrm{Tx}} - L(d) \\ &= P_C - 10\eta \log_{10} d,\end{aligned} \tag{5}$$

where η is the constant path loss index. By applying OLS for $\boldsymbol{y}$, we can estimate both P_C and η and then extract $W(\boldsymbol{x}_i)$ from $\boldsymbol{y}$. Because ordinary Kriging should be applied for the extracted $W(\boldsymbol{x}_i)$ according to Eq. (4), the interpolation result at $\boldsymbol{x}_0$ can be derived by

$$\begin{aligned}\hat{P}(\boldsymbol{x}_0) = \hat{P}_C &- 10\hat{\eta}\log_{10}||\boldsymbol{x}_0 - \boldsymbol{x}_{\mathrm{Tx}}|| \\ &+ \sum_{i=1}^{N} \omega_i \underbrace{\left(P(\boldsymbol{x}_i) - \left(\hat{P}_C - 10\hat{\eta}\log_{10}||\boldsymbol{x}_i - \boldsymbol{x}_{\mathrm{Tx}}||\right)\right)}_{\triangleq \hat{W}(\boldsymbol{x}_i)},\end{aligned} \tag{6}$$

where $\hat{\eta}$ and $\hat{P}_C$ are the estimated η and P_C obtained by OLS.

Ordinary Kriging determines the optimum weights that minimize the variance of the estimation error $\sigma_k^2 = \mathrm{Var}[\hat{W}(\boldsymbol{x}_0) - W(\boldsymbol{x}_0)]$, where $\mathrm{Var}[\cdot]$ is the variance of the random variable. To achieve the best linear unbiased estimator (BLUE), this method determines the weights under the constraint $\sum_{i=1}^{N} \omega_i = 1$. Using the method of the Lagrange multiplier, the objective function can be written as

$$\phi(\omega_i, \mu) = \sigma_k^2 - 2\mu\left(\sum_{i=1}^{N} \omega_i - 1\right), \tag{7}$$

where μ is the Lagrange multiplier. Here, σ_k^2 can be written as follows [11]:

$$\sigma_k^2 = -\gamma(d_{0,0}) - \sum_{i=1}^{N}\sum_{j=1}^{N} \omega_i \omega_j \gamma(d_{i,j}) + 2\sum_{i=1}^{N} \omega_i \gamma(d_{i,0}), \tag{8}$$

where $d_{i,j} \triangleq ||\boldsymbol{x}_i - \boldsymbol{x}_j||$. In addition, γ is the semivariogram defined as

$$\gamma(d_{i,j}) = \frac{1}{2}\mathrm{Var}[\hat{W}(\boldsymbol{x}_i) - \hat{W}(\boldsymbol{x}_j)]. \tag{9}$$

By taking partial derivatives in Eq. (8) and by assuming that these results equal to zeros, we can obtain $N+1$ simultaneous equations:

$$\begin{pmatrix} \gamma(d_{1,1}) & \cdots & \gamma(d_{1,N}) & 1 \\ \gamma(d_{2,1}) & \cdots & \gamma(d_{2,N}) & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \gamma(d_{N,1}) & \cdots & \gamma(d_{N,N}) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_N \\ \mu \end{pmatrix} = \begin{pmatrix} \gamma(d_{1,0}) \\ \gamma(d_{2,0}) \\ \vdots \\ \gamma(d_{N,0}) \\ 1 \end{pmatrix}. \tag{10}$$

From the above simultaneous equations, the optimal ω_i can be derived.

C. WHEN DOES THE EFFECT OF ERRORS IN PATH LOSS MODELING BECOME SEVERE?

Next, we briefly discuss the effect of imperfect path loss modeling in the Kriging-based REM. Previous methods for REM construction assumed an isotropic path loss model, as shown in Eq. (5). Meanwhile, in many realistic situations, the path loss has anisotropy; e.g., η indicates a degree dependence, and the transmitter employs an antenna array. This difference between the assumed model and true characteristic

causes a difficulty in obtaining $\mathrm{E}[\hat{W}(\boldsymbol{x})] = const.$. Thus, the optimality of ω_i will be broken.

Let us consider a simple situation of path loss modeling to examine this effect. For simplicity, we assume a simple anisotropic path loss; the path loss follows Eq.(5) at a certain angle, and P_C and η depend on the horizontal angle between the transmitter and the receiver. Under this assumption, we discuss the error of path loss modeling over a straight line. Fig. 3(a) shows both true and estimated path loss curves. If OLS assuming the isotropic model is applied to model the path loss in such an anisotropic situation, the error becomes severe according to the anisotropy. In this figure, the error of path loss modeling at a given distance d can be written as

$$\varepsilon(d) = \left(P_C - 10\eta \log_{10} d\right) - \left(\hat{P}_C - 10\hat{\eta}\log_{10} d\right) \ [\mathrm{dB}]. \quad (11)$$

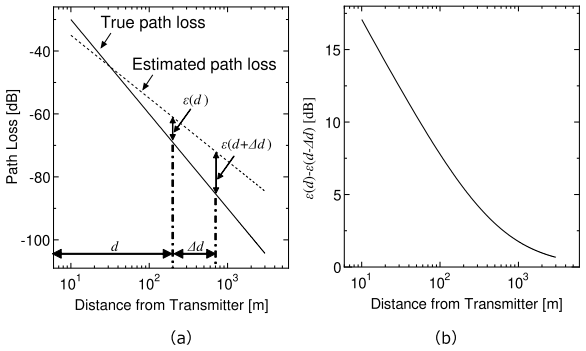

FIGURE 3. Example of path loss modeling where $\eta = 3.0$, $\hat{\eta} = 2.0$, $P_C = 0.0$, $\hat{P}_C = -15.0$, and $\Delta d = 500$). (a) True and estimated path loss curves, (b) $\varepsilon(d) - \varepsilon(d - \Delta d)$. The farther away is the transmitter, the greater is the probability that the regression result becomes stationary on the spatial axis. Near the transmitter, Kriging is strongly affected by the error of path loss modeling.

Here, we can achieve $\mathrm{E}[\hat{W}(\boldsymbol{x})] = const.$ spatially when the difference in the error of path loss modeling between two points is zero, i.e., $\varepsilon(d) - \varepsilon(d + \Delta d) = 0$. This condition then gives

$$\varepsilon(d) - \varepsilon(d + \Delta d) = 10(\eta - \hat{\eta})\log_{10}\frac{d + \Delta d}{d}. \quad (12)$$

As this equation approaches zero, ordinary Kriging will work properly. Here, we define Δd [m] as the maximum distance between the interpolated location and the dataset utilized for Kriging. Although the number of datasets for the interpolation can be increased by taking a large Δd, a too-large Δd may not efficiently improve the accuracy of Kriging because a pair of distant data are almost uncorrelated to each other. Thus, a sufficient Δd will be several hundred meters to several kilometers. A numerical example of Eq. (12) where $\Delta d = 500$ [m] is shown in Fig. 3(b). Even if OLS performs an imperfect path loss modeling, Eq. (12) approaches zero as the interpolated location leaves the transmitter because this value depends on $\log_{10}((d + \Delta d)/d)$. This brief discussion provides us with simple but important knowledge: the estimation error of the path loss strongly affects the optimality of ordinary Kriging if the interpolation is performed near the transmitter.

Note that we focused only on $\varepsilon(d) - \varepsilon(d + \Delta d)$, and the amount of ε itself is ignored in this discussion. This is because ordinary Kriging compensates $\varepsilon(d)$ if $\varepsilon(d) - \varepsilon(d + \Delta d) = 0$. Therefore, the influence of the error in path loss modeling can be neglected in the end even if $\varepsilon(d)$ is large. This characteristic is shown in Sect. IV via a numerical simulation (see Fig. 6).

III. NEURAL NETWORK RESIDUAL KRIGING

In Sect. II, we pointed out the effects of imperfect path loss modeling. The accuracy of REM is strongly affected by this imperfection; improving the path loss modeling may improve the accuracy of REM at locations near the transmitter. Path loss modeling has been widely discussed in the community of radio propagation. A number of path loss models, e.g., the Okumura-Hata model and COST-231, have been studied, and we can obtain an accurate path loss characteristic by choosing a suitable model [12]. Meanwhile, accurate path loss modeling requires a careful choice of model, and this difficulty often causes a mismatch between the selected model and the actual estimated area. In many situations, fitting a simple equation into an actual measured dataset outperforms the man-selected path loss model [13].

Some recent works showed the outstanding performance of NNs in path loss modeling [20], [21]. The use of NNs may improve the accuracy of REM. In this section, we present an NN-assisted REM construction called *NNRK*.

A. PATH LOSS MODELING VIA FFNN

An NN consists of artificial neurons. These neurons are connected by weighted links, and each neuron propagates a value obtained from the weight and its own nonlinear function. By adjusting the weights between the neurons from the dataset, we can perform model-free and accurate path loss modeling. Fig. 4 shows the FFNN we used in this paper. To fully take into account the effects of location dependence, we input both the transmitter location and the receiver location.[1] Because the path loss mainly depends on the communication distance and the azimuth, the receiver coordinate is first converted into d [m] and θ [deg]. Additionally, the output layer consists of the average received signal power over the logarithmic domain related to the input information.

Generally, we can easily gain the expression capability of NNs by employing a deep-layered structure [15]–[18]. Meanwhile, from an optimization-based discussion, the authors in [20] concluded that a few (1–3) hidden layers can achieve a sufficient accuracy for path loss modeling. The authors in [21] showed a similar result. According to these discussions, we exploit only one hidden layer.

[1] If the transmitter is fixed at a location, the input layer may need only the receiver location. Because Sect. V extends the NNRK to the distributed situation, we assume this four-dimensional input layer.

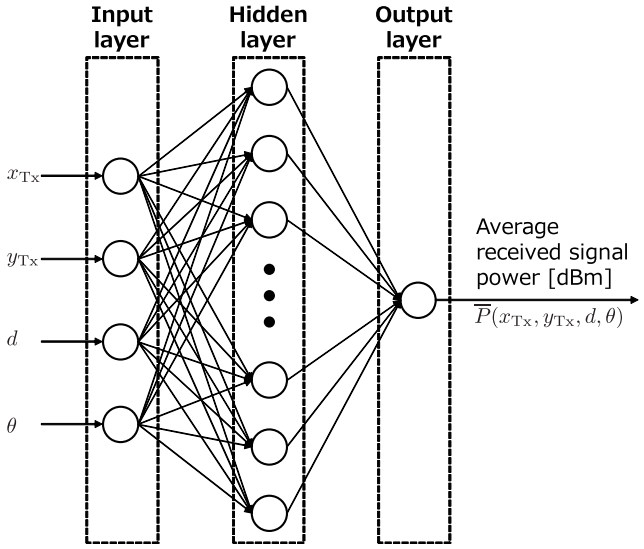

FIGURE 4. FFNN used for path loss modeling. Although a deep-layered structure can improve the capability of NNs [15]–[18], some authors have shown that a few (1–3) hidden layers can achieve a sufficient accuracy for path loss modeling [20], [21]. According to these results, this study uses only one hidden layer.

B. NEURAL NETWORK RESIDUAL KRIGING

Once the FFNN is constructed, the path loss can easily be obtained by using the FFNN, transmitter location, d, and θ; the FFNN can be expressed as a function of path loss $f(\boldsymbol{x}_{\mathrm{Tx}}, d, \theta)$. The procedure of NNRK is the same as the conventional method, except that FFNN is used for path loss modeling instead of OLS, as shown in Fig. 2. Thus, we can implement NNRK by modifying Eq. (6) as

$$\hat{P}(\boldsymbol{x}_0) = f(\boldsymbol{x}_{\mathrm{Tx}}, d_{0,\mathrm{Tx}}, \theta_{0,\mathrm{Tx}}) + \sum_{i=1}^{N} \omega_i \left(P(\boldsymbol{x}_i) - f(\boldsymbol{x}_{\mathrm{Tx}}, d_{i,\mathrm{Tx}}, \theta_{i,\mathrm{Tx}})\right), \quad (13)$$

where $\theta_{i,j} = \angle\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)$. Note that ω_i is calculated by solving Eq. (10) according to the manner of ordinary Kriging.

C. IMPLEMENTATION OF FFNN IN THIS PAPER

We implement the FFNNs using Python 3.6.5 with Chainer 4.5.0. The task of the FFNN in this context is to predict an accurate $\overline{P}(x_{\mathrm{Tx}}, y_{\mathrm{Tx}}, d, \theta)$, and this FFNN should be constructed via training dataset $\boldsymbol{y}$. We implement the FFNN according to the manner of mini-batch learning. As mentioned at the beginning of this section, we consider one hidden layer. All of the activation functions employ a rectified linear function (ReLU) that is defined as $h(x) = \max(0, x)$.

Before the learning, distance d_i and degree θ_i are extracted from $\boldsymbol{x}_{\mathrm{Tx}}$ and $\boldsymbol{x}_i$, and all input values are normalized to 0–1 in each dimension. In the learning process, $N_{\mathrm{mb}}(< N)$ mini-batch datasets are randomly selected from $\boldsymbol{y}$, and we input them into the FFNN at once. Then, the MSE between the estimated and measured datasets is calculated. The MSE is defined as

$$MSE = \frac{1}{N_{\mathrm{mb}}} \sum_{i=1}^{N_{\mathrm{mb}}} \left(p_i - \hat{p}_i\right)^2, \quad (14)$$

where $\hat{p}_i$ is the normalized output obtained from FFNN, and p_i is the normalized $P(\boldsymbol{x}_i)$. Next, the FFNN updates its own weight factors according to backpropagation-based gradient calculation and Adam [25], which is an efficient algorithm for stochastic optimization. The above procedures are iterated for N_{epoch} epochs.

IV. PERFORMANCE EVALUATION

This section presents performances of NNRK for the situation of a fixed transmitter. First, we verify the effect of imperfect path loss modeling via a numerical simulation by assuming a simple anisotropy. After that, an actual measured dataset is utilized for an experimental evaluation.

A. NUMERICAL SIMULATION

The simulation model is summarized in Fig. 5. A two-dimensional plane consisting of a square with a side length L [m] is used as the evaluation area. The transmitter is located on the coordinate $\boldsymbol{x}_{\mathrm{Tx}} = (0, L/2)$, and the REM focuses on this transmitter. In the first part of this simulation, we randomly select 1024 measurement locations based on a two-dimensional uniform distribution. If the received signal power exceeds a measurement threshold, then the value is counted for the dataset. All successfully received values are contained in the dataset vector $\boldsymbol{y}$, and its length is expressed as N. To evaluate the effect of the distance from the transmitter on the REM, we consider the evaluation location $\boldsymbol{x}_0 = (d_{0,\mathrm{Tx}}, L/2)$ where $d_{0,\mathrm{Tx}}$ [m] is the distance from the transmitter. Under the above conditions, the RMSE at $\boldsymbol{x}_0$ is evaluated via a Monte Carlo simulation.

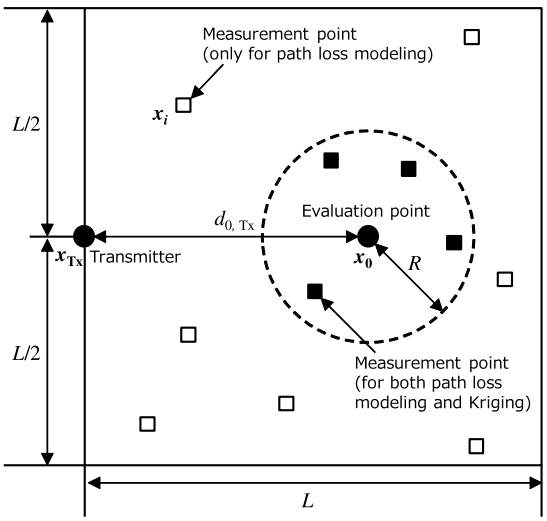

FIGURE 5. Simulation model.

TABLE 1. Simulation parameters.

Radio Propagation Characteristics	
Correlation distance d_{cor}	300[m]
Standard deviation of shadowing σ	8.0[dB]
Path loss index η	3.0
3 dB beamwidth $\theta_{3\mathrm{dB}}$	60[deg]
Maximum attenuation $A_{\max}$	20[dB]
Measurement configuration	
Area length L	5000[m]
Radius R	500[m]
Number of measurement locations	1024
Measurement threshold	-100[dBm]
Number of trials	10000
Hyperparameters in FFNN	
Number of units in hidden layer	100
Mini-batch size N_{mb}	256
Number of epochs N_{epoch}	500

After the measurement, the path loss is modeled from $\boldsymbol{y}$ via OLS or FFNN. Then, the shadowing at $\boldsymbol{x}_0$ is estimated by ordinary Kriging. In Kriging, we use only the dataset that is in a circle with radius R [m] where the center point is located at $\boldsymbol{x}_0$; we express this dataset by $\boldsymbol{y}^*$. Before applying ordinary Kriging, we generate empirical semivariograms from the estimated shadowing in $\boldsymbol{y}^*$ according to *binning* [11]. Then, the generated semivariograms are fitted to the exponential semivariogram model, which is defined as

$$\gamma\left(d_{i,j}\right) = \alpha_n^2 + \alpha_s^2 \left\{1 - \exp\left(-\frac{d_{i,j}}{\alpha_r}\right)\right\}, \tag{15}$$

where α_n^2, α_s^2, and α_r are *Nugget*, *Sill*, and *Range*, respectively. Note that we determine $\alpha_n^2 = 0$ in this simulation to avoid the event that Eq. (10) rarely has no result.

Radio propagation characteristics follow Eq. (2) and we do not consider any time-variant factors. Additionally, to express the anisotropic radio propagation simply, the transmitter employs beamforming; P_C in Eq. (2) is assumed to follow $P_C(\theta) = P_{\mathrm{Tx}} + G(\theta)$, where P_{Tx} [dBm] is the transmission power, θ ($-180° \leq \theta \leq 180°$) is the azimuth, and $G(\theta)$ [dBi] is the angle-dependent antenna gain. According to [26], we assume

$$G(\theta) = -\min\left(12\left(\frac{\theta}{\theta_{3\mathrm{dB}}}\right)^2, A_{\max}\right), \tag{16}$$

where $\theta_{3\mathrm{dB}}$ [deg] is the 3 dB beamwidth, and $A_{\max}$ [dB] is the maximum attenuation.

1) RMSE VS DISTANCE FROM TRANSMITTER

The impact of $\boldsymbol{d}_{0,\mathrm{Tx}}$ on the RMSE is shown in Fig. 6. Both OLS and FFNN are compared in this simulation. Additionally, we evaluated the performance where $\overline{P}(\boldsymbol{x})$ can be obtained perfectly. This means that path loss modeling was performed without any errors, and the Kriging fully optimizes the interpolation; thus, this value shows a maximum accuracy in this simulation. The RMSE of OLS-aided Kriging increases by 2.9 [dB] in $d_{0,\mathrm{Tx}} = 500$ [m] compared to the perfect case. Meanwhile, NNRK suppresses this degradation;

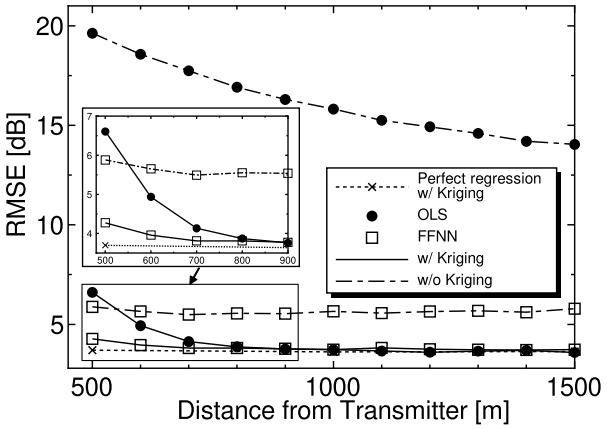

FIGURE 6. RMSE against distance from transmitter. The effects of path loss modeling strongly depend on the distance between the transmitter and the estimated location. As the communication distance decreases, the influence of errors in path loss modeling increases.

thus, FFNN can improve the accuracy of REM in the presence of anisotropic path loss. In both methods, the farther from the transmitter, the more the accuracies approach the perfect case because the effect of imperfect path loss modeling on Kriging can be ignored in far places, as previously discussed in Sect. II-C. We are interested in the results before and after applying Kriging in OLS. Before applying Kriging, the RMSE is over 15 [dB] in almost all regions and is *not* practical. Applying the Kriging compensates the error in path loss modeling, and near-optimal spatial interpolation can be performed.

OLS requires much lighter calculations than FFNN. If the REM is constructed in large-scale systems, such as television and cellular, OLS will be an appropriate selection for preprocessing in Kriging. By contrast, if the REM is constructed in small-scale systems, such as small-cell and WLAN, FFNN will be capable of improving the accuracy of REM.

2) RMSE VS. CORRELATION DISTANCE d_{cor}

The effects of the correlation distance of shadowing d_{cor} where $d_{0,\mathrm{Tx}} = 500, 2000$ [m] are shown in Fig. 7. d_{cor} depends on the obstacle structures, and the value is

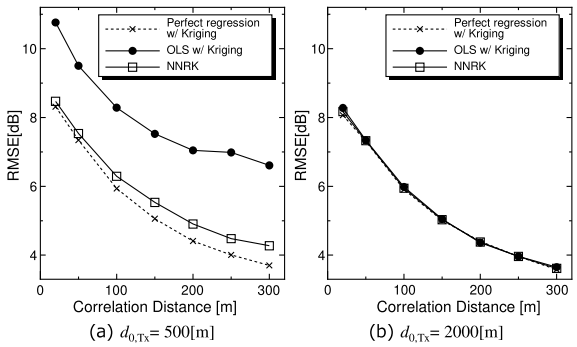

FIGURE 7. RMSE against correlation distance d_{cor}. Regardless of d_{cor}, the distance from the transmitter is a very important metric when considering if the FFNN should be applied.

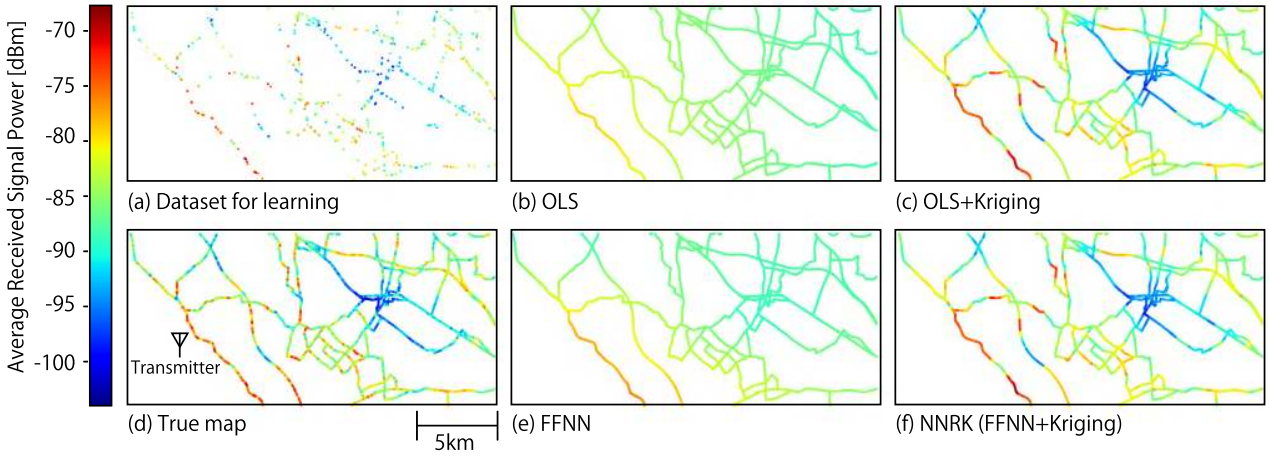

FIGURE 8. Example of REM where $N = 512$. OLS with Kriging and NNRK show no difference in their construction results, at least visually.

determined empirically (e.g., several tens of meters for urban and several hundreds of meters for suburban) [14]. Because highly spatial correlation allows Kriging to perform an accurate interpolation, the RMSE increases as d_{cor} becomes shorter. Meanwhile, in any d_{cor}, the effect of path loss modeling can be ignored away from the transmitter. Regardless of d_{cor}, the distance from the transmitter is a very important metric for considering whether we should apply the FFNN.

B. NUMERICAL EXAMPLE WITH ACTUAL MEASURED DATASET

We present a numerical example using an actual measured dataset that we used in [27], [28]. In [27], we conducted measurements over a five-day period in October 2013. The spectrum sensing function was implemented on a software-defined radio platform [Universal Software Radio Peripheral (USRP) N210] by the GNU Radio software and ran on a laptop computer. The TV signal was sampled using a fast Fourier transform with a sampling rate of 200 [kHz], and the number of samples was set to 2048. The noise floor of the USRP N210 for the above conditions is approximately -171 dBm/Hz in a temperature chamber. Five vehicles were used to perform signal measurements while driving on roads, and two sensing devices were installed in each vehicle, which made a total of 10 measurement devices. Each antenna was put on the roof of a vehicle at a height of 1.7 [m], and two antennas per vehicle were installed diagonally on the roof. The sensing results were stored on a laptop PC with the location information obtained using a USB-connected GPS unit (Garmin GPS18xUSB). Based on the above conditions, the television signals from the Kumagaya relay station were measured. This station is located in Kumagaya city, Saitama, Japan, which is a suburban area. We measured signals with a center frequency of 521.14 [MHz] that were vertically polarized and whose transmission consisted of the Integrated Services Digital Broadcasting-Terrestrial (ISDB-T) standard with an equivalent isotropically radiated power (EIRP) of 31 [W]. ISDB-T consists of 13 segments of orthogonal frequency division multiplexing (OFDM) signals, and each segment has 432 subcarriers; a channel has roughly 6 [MHz] of bandwidth. After the measurement campaign, to estimate the spatially distributed average received signal power, the prior datasets were averaged over spatial grids in which each side had a length of 10 [m].

In Fig. 8, we show an example of radio environment mapping. Fig. 8(d) plots all of the averaged datasets we used in this evaluation. The 22 332 grids contain average received signal values. In this example, 512 grids are randomly selected from these datasets and are utilized for learning. The learning datasets in this example are shown in Fig. 8(a), and Figs. 8(b), (c), (e), and (f) show the REM constructed via OLS, OLS-aided Kriging, FFNN, and NNRK, respectively. Although the FFNN-based path loss modeling well expresses the anisotropy of the path loss, both the OLS-aided Kriging and NNRK construct similar REMs, at least visually.

To compare the detailed accuracies of these methods, we perform a cross-validation-based evaluation. In this evaluation, several datasets are randomly selected from all of the datasets and are utilized for learning. Next, for the cross-validation, 100 grids are randomly selected from the datasets except the learning datasets, and the RMSE is calculated. These procedures are iterated 1000 times. Note that semivariogram modeling and Kriging use only datasets up to 500 [m] from the interpolation point.

Fig. 9 shows the CDFs of the RMSE where $N = 1024$. Although the FFNN statistically improves the performance of path loss modeling, the gap between FFNN and OLS becomes small after Kriging is applied. This is because the datasets were observed from several kilometers to several tens of kilometers from the transmitter. In such an environment, OLS assuming isotropic path loss can obtain sufficient accuracy, as shown in Fig. 6.

Fig. 10 shows 90-percentile RMSEs where $N = 512$, 1024, and 2048. Characteristics similar to those in Fig. 9 can be found for both $N = 512$ and 2048. If Kriging is applied to such a large-scale system, even if FFNN can improve

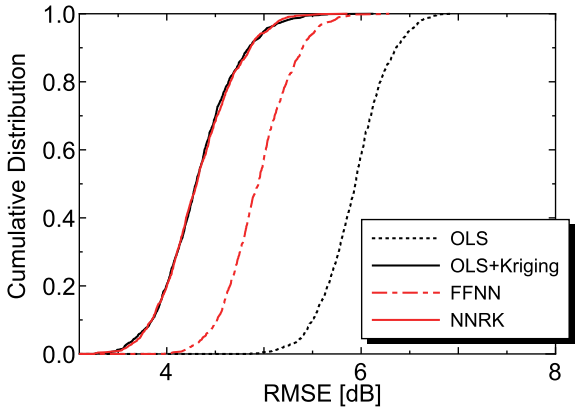

FIGURE 9. CDFs of RMSE over TV bands where $N = 1024$. Although FFNN improves the accuracy of path loss modeling, this effect is compensated after Kriging is applied in this case. Note that FFNN improves the accuracy of Kriging in the distributed case. This result is shown in Fig. 13.

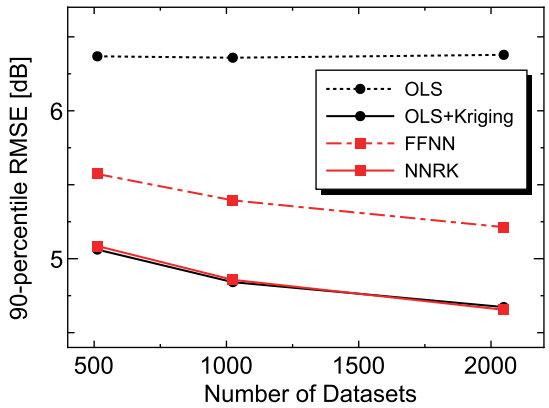

FIGURE 10. Ninetieth-percentile RMSE over TV bands.

the accuracy of the path loss modeling, OLS will achieve a sufficient accuracy after the Kriging is applied.

V. EXTENSION TO DISTRIBUTED WIRELESS NETWORKS

We have discussed the performance of NNRK in the situation where the transmitter is fixed. Even if the path loss has anisotropy, OLS with Kriging can perform accurate radio environment mapping in large-scale systems.

Next, we consider a situation where the transmitter has an arbitrary location, such as in MANET. The shadowing in such a distributed situation also has a spatial correlation [23]; thus, from the theoretical viewpoint, we can easily extend Kriging-based estimation to this situation. Meanwhile, because the path loss has too-complex anisotropy according to the transmission location, we cannot accurately model this with an equation. Because the communication range in such a situation is within several hundreds of meters in many cases, the error of path loss modeling may strongly degrade the accuracy of radio environment mapping, as shown in Sect. IV. In this section, we evaluate the performance of NNRK in the distributed situation.

A. DATA COLLECTION AND PREPROCESSING IN CLOUD

As in the fixed case, radio environment mapping in the distributed situation also starts from data collection via crowdsensing. Here, we assume that all packets include the transmission location. After peer-to-peer (P2P) wireless communication, the receiver records the received signal power that is related to both the receiver and transmitter locations, and reports it to the cloud server. After the server collects sufficient datasets from massive reporters, the time-variant factors such as multipath fading are removed from the datasets to conduct radio environment mapping in the manner of spatial statistics. In the fixed case, this process can be simply implemented by averaging the instantaneous signals by a grid. By contrast, in the distributed situation, simple averaging that focuses only on the receiver location cannot extract meaningful features because the datasets have arbitral coordinates for the transmission location. Thus, we also consider gridizing the transmitter locations. In this process, after both the transmitter and receiver locations are gridized, all datasets are classified to a pair of grids according to the transmitter and receiver locations. By averaging the datasets for each pair of transmit and receive grids, we can construct (tooth-missing) REMs for each transmit grid [29].

After this subsection, the received power represents the preprocessed value.

B. PRINCIPLE

Let us consider N datasets where the transmitter in the i-th link is located in $\boldsymbol{x}_{i,\mathrm{Tx}}$ and the paired receiver is in $\boldsymbol{x}_{i,\mathrm{Rx}}$. The average received signal power at $\boldsymbol{x}_{i,\mathrm{Rx}}$ can be expressed by redefining Eq. (2) as

$$P(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}}) = P_C - L(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}}) + W(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}}). \tag{17}$$

According to reference [23], the spatial correlation of shadowing in the distributed situation can be simply extended from Eq. (3). This characteristic can be derived by

$$\rho_{i,j} \approx \exp\left(-\frac{d_{(i,j),\mathrm{Tx}} + d_{(i,j),\mathrm{Rx}}}{d_{\mathrm{cor}}} \ln 2\right), \tag{18}$$

where $d_{(i,j),\mathrm{Tx}} = ||\boldsymbol{x}_{i,\mathrm{Tx}} - \boldsymbol{x}_{j,\mathrm{Tx}}||$ [m] and $d_{(i,j),\mathrm{Rx}} = ||\boldsymbol{x}_{i,\mathrm{Rx}} - \boldsymbol{x}_{j,\mathrm{Rx}}||$ [m]. The spatial correlation in the distributed situation depends on the moving distances of the transmitter and receiver. Considering this fact, the radio environment mapping derived in Sect. II and Fig. 2 can be applied in the distributed situation.

In the distributed situation, the task of radio environment mapping is to estimate the received signal power at a given link $P(\boldsymbol{x}_{0,\mathrm{Tx}}, \boldsymbol{x}_{0,\mathrm{Rx}})$ from the measurement vector $\boldsymbol{y} = \left(P(\boldsymbol{x}_{1,\mathrm{Tx}}, \boldsymbol{x}_{1,\mathrm{Rx}}), P(\boldsymbol{x}_{2,\mathrm{Tx}}, \boldsymbol{x}_{2,\mathrm{Rx}}), \cdots, P(\boldsymbol{x}_{N,\mathrm{Tx}}, \boldsymbol{x}_{N,\mathrm{Rx}})\right)^{\mathrm{T}}$. In the distributed situation, this task can be achieved by the following procedure.

1) The communication distance $d_i = ||\boldsymbol{x}_{i,\mathrm{Tx}} - \boldsymbol{x}_{i,\mathrm{Rx}}||$ and the horizontal angle $\theta_i = \angle(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}})$ are calculated from $\boldsymbol{y}$.

2) The path loss is modeled. If OLS is applied, the path loss is estimated from d_i and $P(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}})$. Note that we assume the path loss as $\overline{P}(d) = P_C - 10\eta \log_{10}(d + 1)$ because the communication distance often falls into zero in this grid-based architecture. If FFNN is applied, the path loss is learned from $\boldsymbol{x}_{i,\mathrm{Tx}}$, d_i, θ_i, and $P(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}})$.
3) Shadowing $W(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}})$ is extracted from $\boldsymbol{y}$.
4) A semivariogram is analyzed from $W(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}})$. Note that the semivariogram is expressed as the function of $d_{(i,j),\mathrm{Tx}} + d_{(i,j),\mathrm{Rx}}$; thus, the semivariogram can be derived as

$$\gamma(d_{(i,j),\mathrm{Tx}} + d_{(i,j),\mathrm{Rx}}) = \frac{1}{2}\mathrm{Var}[\hat{W}(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}}) - \hat{W}(\boldsymbol{x}_{j,\mathrm{Tx}}, \boldsymbol{x}_{j,\mathrm{Rx}})]. \quad (19)$$

5) Ordinary Kriging is applied. Because the semivariogram is defined as the function of $d_{(i,j),\mathrm{Tx}} + d_{(i,j),\mathrm{Rx}}$, the smaller this value, the larger the weight factor in Kriging. In this section, Kriging uses only datasets that satisfy $d_{(i,0),\mathrm{Tx}} + d_{(i,0),\mathrm{Rx}} \leq 100$ [m].

C. MEASUREMENT CAMPAIGN AND PREPROCESSING

To confirm the performance of NNRK in the distributed situation, we use the dataset that we measured using a V2V communication system in [29].

In this experiment, we measured the received signal power of the communication packet of the Dedicated Short Range Communications (DSRC), the current standard for Intelligent Transport Systems (ITS), with three vehicles. Each vehicle exploits an onboard unit (OBU: MK5 OBU, Cohda Wireless), and these run while communicating with each other. This OBU broadcasts 100 bytes and 400 bytes of packets including its own location at a rate of 200 packets per second. The received signal power related to the transmitter/receiver locations is recoded in the local storage; these data had been uploaded to the database implemented by MySQL. We performed the experiment at California Partners for Advanced Transportation Technology (PATH), Richmond, CA, USA, over two days. The detailed parameters of OBU are listed in Table 2.

TABLE 2. Parameters of OBU.

Equipment	MK5 OBU (Cohda Wireless)
Antenna	Two antennas for communication One antenna for GPS
Frequency	5890 [MHz]
Transmit power	24 [dBm]
Modulation	BPSK, QPSK, 16QAM
Coding rate	1/2
Transmit rate	200 [packets/s]
Packet length	100 and 400 [bytes]

Each OBU repeatedly transmits packets with different modulation formats and different packet lengths at a rate of 200 packets per second. Additionally, each OBU receives the packets transmitted by the other vehicles and stores communication logs. The log includes the packet ID, GPS information of both transmitter/receiver locations, received signal strength indicator (RSSI), and noise level of each antenna.

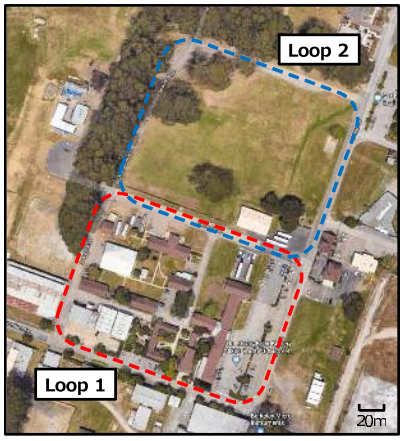

FIGURE 11. Measurement environment (acquired from Google Earth).

The test route is shown in Fig. 11. We chose two test fields named Loop 1 and Loop 2. The three vehicles run the outer circular road of each loop. One vehicle runs the loop in the clockwise direction, and the other two vehicles run the loop in the counterclockwise direction. Here, Loop 1 consists of many buildings; contrastingly, there are only a few buildings in Loop 2. Thus, these routes can demonstrate typical V2V communication environments that contain both line-of-site (LOS) and non-line-of-site (NLOS) conditions. We obtained approximately 15 000 000 instantaneous datasets in 4 h in each loop. Using these datasets, we constructed the average received signal power maps related to different transmission locations with 5-m grids. Note that the dataset for the performance evaluation was filtered so that the communication distance is within 100 m to remove the effect of the noise floor. The number of datasets after this filtering is 16 322.

D. RESULT

We perform a cross-validation-based evaluation that follows the same procedure in Sect. IV-B. Before showing the detailed performance, we plot an example of the theoretical semivariogram in Fig. 12. This figure was created from 1024 randomly selected datasets, and both OLS and FFNN use the same datasets. The semivariogram means $\frac{1}{2}\mathrm{Var}[\hat{W}(\boldsymbol{x}_{i,\mathrm{Tx}}, \boldsymbol{x}_{i,\mathrm{Rx}}) - \hat{W}(\boldsymbol{x}_{j,\mathrm{Tx}}, \boldsymbol{x}_{j,\mathrm{Rx}})]$, and Kriging assigns a large ω_i to the dataset that has a small semivariogram. In this figure, FFNN-based path loss modeling suppresses the semivariogram. This means that the FFNN allows Kriging to utilize data that is more distant in the interpolation. Fig. 13 shows CDFs of the RMSE where $N = 1024$. Unlike the evaluation in the situation of a fixed transmitter, the NNRK significantly improves the RMSE. Additionally, the 90-percentile RMSE is shown in Fig. 14. This tendency can be confirmed regardless of N.

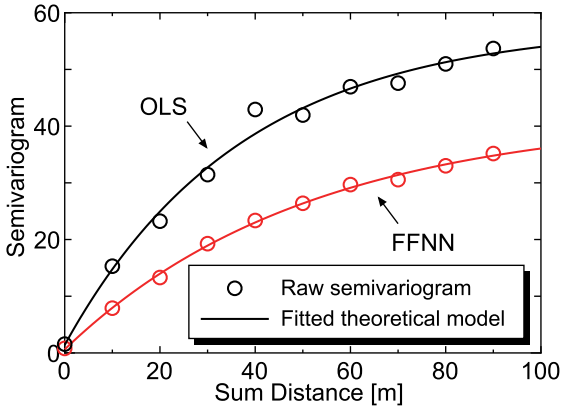

FIGURE 12. Example of semivariogram modeling in distributed situation. FFNN-based regression can suppress semivariogram of shadowing; this means that FFNN allows Kriging to utilize more distant data in interpolation.

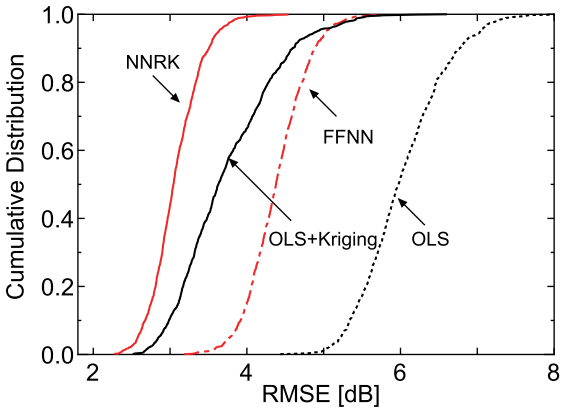

FIGURE 13. CDFs of RMSE over distributed situation. In this situation, FFNN-based path loss modeling significantly improves the accuracy of Kriging.

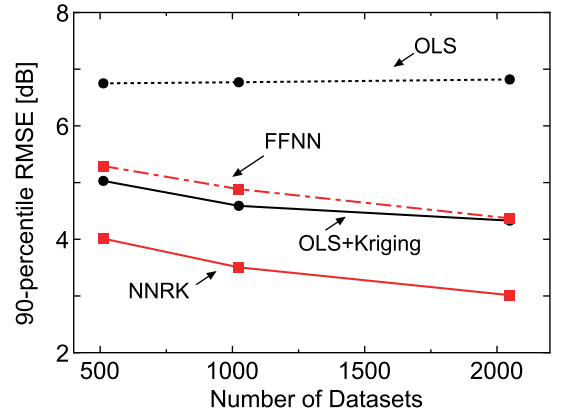

FIGURE 14. Ninetieth-percentile RMSE over distributed situation.

To clarify the effect of communication distance, we plot the RMSE vs. communication distance where $N = 1024$ in Fig. 15. In this evaluation, 100 grids are randomly selected from the datasets except the learning datasets, and each error is calculated in dB. After these procedures are iterated

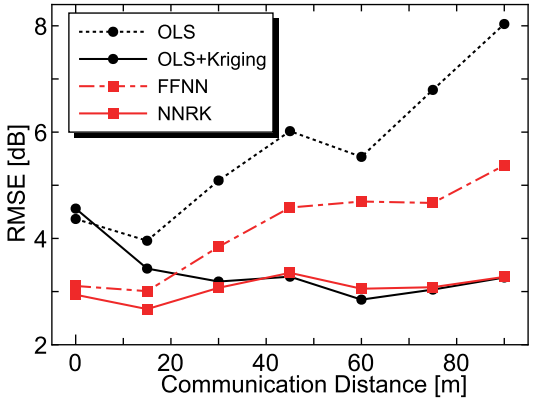

FIGURE 15. Effects of communication distance on RMSE where $N = 1024$. Even in a distributed environment, FFNN helps Kriging as communication distance is shorter.

1000 times, the error values are classified according to the communication distance, and each RMSE is calculated. It can be seen that FFNN improves the accuracy of Kriging as the communication distance is closer. In many distributed wireless systems, the communication distance is within hundreds of meters and the path loss is too complicated when compared with the situation of a fixed transmitter. Such a distributed network will be an application of FFNN-based path loss modeling.

VI. CONCLUSION

This paper evaluated the performance of NNRK in radio environment mapping. Specifically, we focused on whether FFNN-based path loss modeling improves the accuracy of radio environment mapping. The main novel knowledge obtained through this work is as follows:

- The effects of path loss modeling strongly depend on the distance between the transmitter and the estimated location. As the communication distance decreases, the influence of errors in path loss modeling increases.
- In a large-scale system, OLS with Kriging achieves an almost equal performance of NNRK, although FFNN outperforms OLS in path loss modeling. We demonstrated this fact with a dataset measured over TV bands.
- In distributed wireless networks, NNRK significantly improves the accuracy because such a system suffers from too-complex path loss and the communication distance is usually short. This was verified via the V2V system.

The communication distance and the complexity of path loss are useful criteria when deciding whether to adopt FFNN.

REFERENCES

[1] S. Bhattarai, J.-M. J. Park, B. Gao, K. Bian, and W. Lehr, "An overview of dynamic spectrum sharing: Ongoing initiatives, challenges, and a roadmap for future research," *IEEE Trans. Cogn. Commun. Netw.*, vol. 2, no. 2, pp. 110–128, Jun. 2016.

[2] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[3] 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 *Proc. 2nd IEEE Int. Symp. New Frontiers Dyn. Spectr. Access Netw.*, Apr. 2007, pp. 115–118.
[4] H. B. Yilmaz, T. Tugcu, F. Alagöz, and S. Bayhan, "Radio environment map as enabler for practical cognitive radio networks," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 162–169, Dec. 2013.
[5] S. Bi, J. Lyu, Z. Ding, and R. Zhang, "Engineering radio maps for wireless resource management," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 133–141, Apr. 2019.
[6] H. R. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," in *Proc. IEEE DySPAN*, Aachen, Germany, May. 2011, pp. 443–454.
[7] J. Perez-Romero, A. Zalonis, L. Boukhatem, A. Kliks, K. Koutlia, N. Dimitriou, and R. Kurda, "On the use of radio environment maps for interference management in heterogeneous networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 184–191, Aug. 2015.
[8] Z. El-friakh, A. M. Voicu, S. Shabani, L. Simić, and P. Mähönen, "Crowd-sourced indoor Wi-Fi REMs: Does the spatial interpolation method matter?" in *Proc. IEEE DySPAN*, Oct. 2018, pp. 1–10.
[9] K. Sato and T. Fujii, "Kriging-based interference power constraint: Integrated design of the radio environment map and transmission power ," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 1, pp. 13–25, Mar. 2017.
[10] K. Sato, K. Inage, and T. Fujii, "Modeling the Kriging-aided spatial spectrum sharing over log-normal channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 749–752, Jun. 2019.
[11] N. Cressie, *Statistics for Spatial Data*. Hoboken, NJ, USA: Wiley, 1993.
[12] C. Phillips, D. Sicker, and D. Grunwald, "A survey of wireless path loss prediction and coverage mapping methods," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 255–270, 1st Quart., 2013.
[13] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *Proc. IEEE DySPAN*, May 2011, pp. 71–82.
[14] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, Nov. 1991.
[15] M. Liu, T. Song, and G. Gui, "Deep cognitive perspective: Resource allocation for NOMA-based heterogeneous IoT with imperfect SIC," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2885–2894, Apr. 2019.
[16] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2018.
[17] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Commun.*, vol. 14, no. 11, pp. 92–111, 2017.
[18] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.
[19] C. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[20] S. P. Sotiroudis, S. K. Goudos, K. A. Gotsis, K. Siakavara, and J. N. Sahalos, "Application of a composite differential evolution algorithm in optimal neural network design for propagation path-loss prediction in mobile communication systems," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 364–367, 2013.
[21] M. Ayadi, A. B. Zineb, and S. Tabbane, "A UHF path loss model using learning machine for heterogeneous networks," *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3675–3683, Jul. 2017.
[22] M. Kanevsky, R. Arutyunyan, L. Bolshov, V. Demyanov, M. Maignan, "Artificial neural networks and spatial estimation of Chernobyl fallout," *Geoinformatics*, vol. 7, nos. 1–2, pp. 5–11, 1996.
[23] Z. Wang, E. K. Tameh, and A. R. Nix, "Joint shadowing process in urban peer-to-peer radio channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 52–64, Jan. 2008.
[24] A. J. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014, *arXiv:1412.6980*. [Online]. Available: https://arxiv.org/abs/1412.6980
[26] *Evaluation Methods for High Speed Downlink Packet Access (HSDPA)*, document RAN WG1 TDOC R1-00-0909, Jul. 2000.
[27] K. Sato, M. Kitamura, K. Inage, and T. Fujii, "Measurement-based spectrum database for flexible spectrum management," *IEICE Trans. Commun.*, vol. 98, no. 10, pp. 2004–2013, Oct. 2015.
[28] K. Sato, K. Inage, and T. Fujii, "Frequency correlation of shadowing over TV bands in suburban area," *Electron. Lett.*, vol. 54, no. 1, pp. 6–8, Jan. 2018.
[29] T. Fujii, "Smart spectrum management for V2X," in *Proc. IEEE DySPAN*, Oct. 2018, pp. 1–8.

KOYA SATO was born in Miyagi, Japan, in 1991. He received the B.E. degree in electrical engineering from Yamagata University, in 2013, and the M.E. and Ph.D. degrees from The University of Electro-Communications, in 2015 and 2018, respectively. From 2017 to 2018, he was a Research Fellow (DC2) with the Japan Society for the Promotion of Science. He is currently an Assistant Professor with the Tokyo University of Science. His current research interests include spectrum sharing, geostatistics, and radio propagation.

KEI INAGE was born in Chiba, Japan, in 1986. He received the B.E. and M.E. degrees in electronic engineering and the Ph.D. degree in communication engineering and informatics from The University of Electro-Communications, Japan, in 2009, 2011, and 2014, respectively. From 2012 to 2014, he was a Research Fellow (DC2) with the Japan Society for the Promotion of Science (JSPS). From 2014 to 2015, he was a Researcher with the Smart Wireless Laboratory, Wireless Network Research Institute, National Institute of Information and Communications Technology (NICT). From 2015 to 2018, he was an Assistant Professor in the Electrical and Electronics Engineering Program, Tokyo Metropolitan College of Industrial Technology, where he is currently an Associate Professor. His current research interests include spectrum sharing and ad hoc networks. He received the Young Researcher's Award from the IEICE, in 2012.

TAKEO FUJII was born in Tokyo, Japan, in 1974. He received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1997, 1999, and 2002, respectively, where he was a Research Associate with the Department of Information and Computer Science, from 2000 to 2002. From 2002 to 2006, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology. From 2006 to 2014, he was an Associate Professor with the Advanced Wireless Communication Research Center, The University of Electro-Communications, where he is currently a Professor. His current research interests include cognitive radio and ad-hoc wireless networks. He was a recipient of the Best Paper Award for the IEEE VTC 1999-Fall, the 2001 Active Research Award in Radio Communication Systems from the IEICE Technical Committee of RCS, the 2001 Ericsson Young Scientist Award, the Young Researcher's Award from IEICE, in 2004, the Young Researcher Study Encouragement Award from the IEICE Technical Committee of AN, in 2009, the Best Paper Award for IEEE CCNC 2013, and the IEICE Communication Society Best Paper Award, in 2016.