Macrocell Path-Loss Prediction Using Artificial Neural Networks

Erik Östlin, *Member, IEEE*, Hans-Jürgen Zepernick, *Member, IEEE*, and Hajime Suzuki, *Member, IEEE*

***Abstract*—This paper presents and evaluates artificial neural network (ANN) models used for macrocell path-loss prediction. Measurement data obtained by utilizing the IS-95 pilot signal from a commercial code-division multiple-access (CDMA) mobile network in rural Australia are used to train and evaluate the models. A simple neuron model and feed-forward networks with different numbers of hidden layers and neurons are evaluated regarding their training time, prediction accuracy, and generalization properties. Furthermore, different backpropagation training algorithms, such as gradient descent and Levenberg–Marquardt, are evaluated. The artificial neural network inputs are chosen to be distance to base station, parameters easily obtained from terrain path profiles, land usage, and vegetation type and density near the receiving antenna. The path-loss prediction results obtained by using the ANN models are evaluated against different versions of the semi-terrain based propagation model Recommendation ITU-R P.1546 and the Okumura–Hata model. The statistical analysis shows that a non-complex ANN model performs very well compared with traditional propagation models with regard to prediction accuracy, complexity, and prediction time. The average ANN prediction results were 1) maximum error: 22 dB; 2) mean error: 0 dB; and 3) standard deviation: 7 dB. A multilayered feed-forward network trained using the standard backpropagation algorithm was compared with a neuron model trained using the Levenberg–Marquardt algorithm. It was found that the training time decreases from 150 000 to 10 iterations, while the prediction accuracy is maintained.**

***Index Terms*—Artificial neural network (ANN), backpropagation, feed-forward network, field strength measurements, Okumura–Hata (OH) model, path-loss prediction, point to area, Recommendation ITU-R P.1546.**

I. Introduction

A SUCCESSFUL rollout of a cellular mobile radio system largely depends on a well-planned and well-designed cellular structure. For every type of radio access technology, the cellular design has to be supported by the physical layer [1]. Thus, reliable and accurate models are crucial to the prediction of radio channel characteristics for where the cellular mobile radio system is to be deployed. In particular, large-scale fading characteristics have to be reliably predicted to enable the use of base stations (BSs) with optimized characteristics, such as position, height, transmitted power, and antenna pattern.

Traditionally, path-loss prediction models have been based on empirical and/or deterministic methods. Empirical models, such as the Bullington, Longley–Rice, Okumura, and Okumura–Hata (OH) models [2] are computationally efficient but may not be very accurate since they do not explicitly account for specific propagation phenomena. On the other hand, deterministic models, such as those based on the geometrical theory of diffraction [3], integral equation [4], and parabolic equation [5] can, depending on the topographic database resolution and accuracy, be very accurate but lack in computational efficiency. Therefore, artificial neural networks (ANNs) have been proposed to obtain prediction models that are more accurate than standard empirical models while being more computationally efficient than deterministic models. In recent years, ANNs have been shown to successfully perform path-loss predictions in rural [6], suburban [7], urban [8], and indoor [9] environments.

An ANN prediction model can be trained to perform well in environments similar to where the training data are collected. Therefore, to obtain an ANN model that is accurate and generalizes well, measurement data from many different environments should be used in the training process. The performance will also strongly depend on the chosen input parameters. Commonly, the well-known feed-forward structure is used together with the backpropagation training methodology incorporating the standard gradient descent algorithm. A drawback with multilayered feed-forward networks that contain numerous neurons in each layer is the required training time [10]. Furthermore, an overly complex ANN may lead to data overfitting and, hence, generalization problems [11].

Often, in the literature, the utilized feed-forward networks seem to be unnecessarily complex (e.g., three hidden layers and numerous neurons), or sometimes, the network size is not even mentioned at all. Therefore, in this paper, different-sized ANNs are trained by using different backpropagation training algorithms, such as gradient descent and Levenberg–Marquardt. Furthermore, different training data-selection strategies are incorporated to investigate what ANN complexity is needed to achieve high prediction accuracy while maintaining good generalization properties. The goal is to obtain an ANN that is not overly complex but still generalizes well and is accurate enough for the application of cellular mobile radio network

Manuscript received September 5, 2009; revised January 6, 2010 and March 31, 2010; accepted April 30, 2010. Date of publication May 18, 2010; date of current version July 16, 2010. This work was supported by the Commonwealth of Australia through the Cooperative Research Centre program. This paper was presented in part at the 60th IEEE Vehicular Technology Conference, Los Angeles, CA, September 2004. The review of this paper was coordinated by Prof. T. Kuerner.

E. Östlin and H.-J. Zepernick are with the Blekinge Institute of Technology, 372 25 Ronneby, Sweden (e-mail: erik.ostlin@bth.se; hans-jurgen.zepernick@bth.se).

H. Suzuki is with the Information and Communication Technologies Center, Commonwealth Scientific Industrial Research Organisation, Marsfield, NSW 2122, Australia (e-mail: hajime.suzuki@csiro.au).

Digital Object Identifier 10.1109/TVT.2010.2050502

planning. For this purpose, propagation measurements obtained by utilizing the IS-95 [12] pilot signal of a commercial code-division multiple-access (CDMA) mobile telephone network in rural Western Australia are used to train the ANN radio wave path-loss prediction models. This extensive collection of measurement data has previously been used to evaluate and propose modifications to the Recommendation ITU-R P.1546 [13]. One general finding is that a noncomplex ANN, such as the neuron model or a feed-forward network with one hidden layer and only a few neurons, will likely provide sufficient path-loss prediction accuracy for a typical rural macrocell radio network planning scenario. Furthermore, when training these ANNs using an efficient training method, such as the Levenberg–Marquardt algorithm, the required number of iterations in the training process may be reduced from approximately 150 000 to 10, which corresponds to reducing the training time from many hours to a few seconds. Therefore, when incorporating large amounts of training data, the training process will be much more efficient. It was also observed in the statistical analysis that feed-forward networks with several hidden layers and numerous neurons may lead to inferior generalization properties compared with the less-complex structures. This common phenomenon is most often caused by overtraining, i.e., the model performs very well on data that are similar to the training data but is not flexible enough to favorably adapt to data different from the training data. Therefore, the impact on prediction accuracy and generalization properties due to additional hidden layers and neurons was also investigated.

Our main contributions are summarized as follows. Different-sized feed-forward networks and training algorithms were evaluated regarding prediction accuracy, generalization properties, and training time. The aim was to investigate the need for multilayered feed-forward networks and if more efficient training methods, such as the Levenberg–Marquardt algorithm can be utilized to reduce the training time without introducing any negative effects on the path-loss prediction results. Compared with [14], this paper presents prediction results for an extended number of measurement routes, which greatly aided the evaluation of different ANNs' generalization properties. Furthermore, both early stopping (ES) and Bayesian regularization (BR) were incorporated and compared with regard to training time, prediction accuracy, and generalization properties. In addition to our previous work, the created ANN models also include land usage and vegetation input parameters that enable a fairer comparison with Model A [13], which includes some proposed modifications to the Recommendation ITU-R P.1546. Finally, all prediction results in this paper can directly be compared with our previously published Recommendation ITU-R P.1546, Model A, and OH prediction results [13].

The paper is organized as follows. Section II contains relevant ANN background information, such as the neuron model, the feed-forward network, the backpropagation, Levenberg–Marquardt algorithm, and methods to improve the ANN generalization properties. In Section III, the measurement procedure and equipment are introduced. In Section IV, the ANN's input parameters in relation to the problem setting of path-loss prediction for macrocell scenarios are provided. Sections V and VI contain the statistical analysis metric definitions and the ANN training and evaluation, respectively. Finally, conclusions and future work are presented in Section VII.

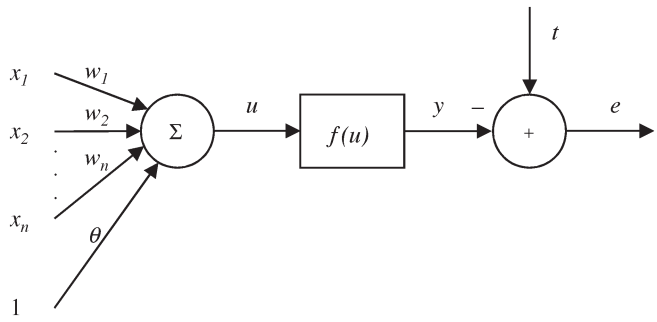


Fig. 1. Neuron model.

II. Artificial Neural Network

The ANN task is to find the best functional fit for a specified set of input–output pairs and interpolate and extrapolate for unknown data sets. In this paper, the application of ANNs for radio wave path-loss predictions can be compared with an approximation of a mathematical mildly non-linear noisy function (measurement data) given at a number of points (training set). To satisfactorily be able to predict the path-loss for different propagation scenarios, some type of regression that minimizes both the training error and the error of unknown inputs has to be performed. The motivation for this paper was to find the most suitable feed-forward structure for the application of path-loss prediction utilizing the measurement data collected in rural Australia and the chosen input parameters. First, the importance of the feed-forward network size was investigated regarding prediction accuracy and generalization properties. Second, the training algorithm selection was investigated with regard to training time, i.e., maintaining the prediction accuracy and generalization properties while minimizing the training time. In the sequel, the neuron model, feed-forward network, training, and generalization methods are introduced and described.

A. Neuron Model

Fig. 1 shows a simple neuron model. The neuron is presented with an input signal, i.e.,

$$\mathbf{x} = [\, x_1 \quad x_2 \quad \cdots \quad x_n \quad 1 \,]^T \tag{1}$$

and accordingly produces an output value, i.e.,

$$u = \mathbf{w}^T \mathbf{x} \tag{2}$$

where $(\cdot)^T$ denotes the transpose, and the neuron weights $\mathbf{w}$ are defined as

$$\mathbf{w} = [\, w_1 \quad w_2 \quad \cdots \quad w_n \quad \theta \,]^T. \tag{3}$$

Furthermore, to provide the possibility to shift the activation function $f(\cdot)$ to the left or right, an additional scalar bias parameter θ is added to the weights. The activation function is called the nonlinearity of the neuron model and can be any differentiable function. For this paper, the activation function

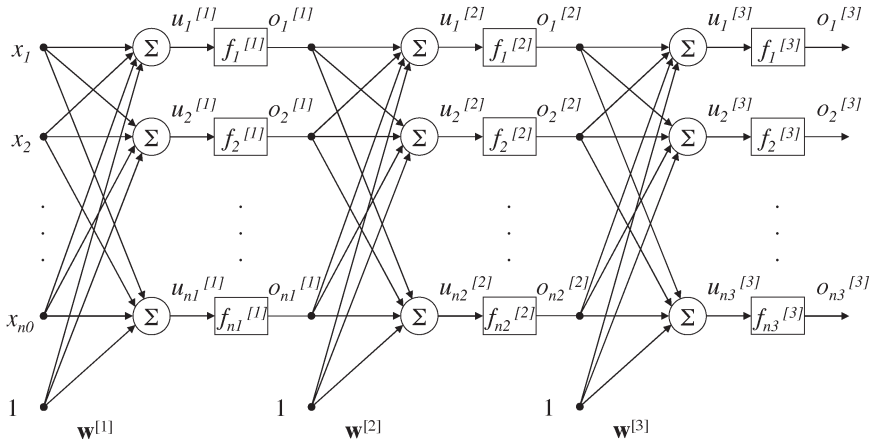


Fig. 2. Three-layer feed-forward ANN.

has been chosen to be the commonly used hyperbolic tangent sigmoid transfer function [15], which is defined as

$$f(u) = \frac{1 - \exp^u}{1 + \exp^u} \tag{4}$$

and compresses the output into the range -1 to 1. As can be seen in Fig. 1, the neuron output error e is calculated by subtracting the sigmoid output $f(u)$ from the target value t as

$$e = t - y. \tag{5}$$

The neuron model's objective is to minimize the output error e according to some optimization criteria, for example, minimizing the sum of the squared errors. In this paper, t and y represent the target path-loss value and the corresponding output path-loss value obtained in the training process, respectively.

B. Feed-Forward Network

The neuron model can be extended to a feed-forward network that incorporates several hidden layers, where each layer has a different number of neurons and weights [15]. Fig. 2 illustrates the network architecture for a three-layer feed-forward ANN. The multilayered feed-forward vectors and weights are defined in the same manner as that for the neuron model. The first hidden layer, the second hidden layer, and the output layer have n_1, n_2, and n_3 neurons, respectively, and the weight vectors $\mathbf{w}^{[1]}$, $\mathbf{w}^{[2]}$, and $\mathbf{w}^{[3]}$ all include a bias term θ. The feed-forward network notation is $\text{ANN}_{n_0-n_1-\cdots-n_k}$, where n_0, n_1, and n_k represent the number of inputs, the number of neurons in the first hidden layer, and the number of neurons in the output layer, respectively. As can be seen in Fig. 2, the ith activation function in the jth hidden layer is shown as $f_i^{[j]}$, and $o_i^{[j]}$ denotes the output from the ith neuron in the jth layer. The number of inputs, hidden layers, and neurons is chosen so that the model can provide an accurate approximation of the given problem. Usually, there is a tradeoff between capturing the complexity of the underlying function given by the training data and the model's ability to generalize to new inputs [16]. This behavior was sometimes observed in the training process and requires careful training data selection when the use of feed-forward networks with several layers and neurons is considered.

C. ANN Training

The training set should be representative of the problem the ANN is designed to solve. A properly trained ANN should be able to recognize whether a new input vector is similar to learned patterns and produce a similar result. Furthermore, when new unknown input parameters are presented to the ANN, it is expected to give an output using interpolation and also extrapolation if the input vectors exceed the parameter space used in the training process. For the multiple-input single-output (MISO) path-loss prediction ANN model, a supervised training process is incorporated [15]. The Q input–output training pairs are randomly chosen from the measurement data and are defined as

$$\{\mathbf{p}_1, t_1\}, \{\mathbf{p}_2, t_2\}, \ldots, \{\mathbf{p}_Q, t_Q\} \tag{6}$$

where $\mathbf{p}_q$ is an input vector, and t_q is the corresponding output. In the work reported here, we consider three different training approaches, which reveal the prediction accuracy and generalization properties of the ANNs (see Section IV). The ANN weights $\mathbf{w}$ are found by minimizing the mean square errors, i.e.,

$$\text{MSE} = \frac{1}{Q}\sum_{q=1}^{Q} e(q)^2. \tag{7}$$

In this paper, the well-known backpropagation (gradient descent) algorithm has been used as a benchmark to train the ANNs [17]. To give a considerable reduction in training time, other much faster training methods, such as the Levenberg–Marquardt algorithm [18] have also been utilized. In the supervised training process, the training algorithms were updated in batch mode, i.e., the ANN weights are renewed after a complete training set has been presented to the ANN [15]. Note that all input–output data pairs used in the training process are removed in the model evaluation.

The two following paragraphs introduce the gradient descent and Levenberg–Marquardt backpropagation algorithms.

Backpropagation and Gradient Descent: The standard implementation of backpropagation learning updates (i.e., ANN weights and biases) is to base the updates on the direction

of the steepest negative gradient descent. This is written in the form

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha_k \mathbf{g}_k \tag{8}$$

where $\mathbf{w}_k$ contains the current weights and biases at the kth iteration. The parameter α_k denotes the learning rate (constant), and $\mathbf{g}_k$ represents the current gradient vector. Some other common algorithms that incorporate additional methods to enable faster convergence can be found in, e.g., [19].

Levenberg–Marquardt Algorithm: The Newton algorithm is often used to obtain fast optimization characteristics. The weight update can be expressed as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mathbf{A}_k^{-1} \mathbf{g}_k \tag{9}$$

where $\mathbf{A}_k$ is the Hessian matrix [20] (second-order derivatives) of the performance index consisting of current weights and biases. Unfortunately, it is very complex and computationally expensive to calculate the Hessian matrix for a feed-forward network. An approach based on the quasi-Newton method is to compute, based on the local gradient, an approximative Hessian matrix at each algorithm iteration. The Levenberg–Marquardt method approaches the second-order behavior without the need to calculate or approximate the Hessian matrix, as in Newton and quasi-Newton algorithms, respectively [20]. For example, when training a feed-forward network, the Hessian matrix can be approximated using the Jacobian matrix $\mathbf{J}$ as [20]

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}. \tag{10}$$

The gradient $\mathbf{g}$ is then expressed as

$$\mathbf{g} = \mathbf{J}^T \mathbf{e} \tag{11}$$

where $\mathbf{e}$ is a vector that contains the network errors. The Jacobian matrix contains only first-order derivatives, with respect to weights and biases, and can, hence, be calculated using the standard backpropagation technique [18]. This approach is less complex and more computationally efficient than computing or approximating the Hessian matrix. The Newton-like weight update can now be expressed as

$$\mathbf{w}_{k+1} = \mathbf{w}_k - [\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}]^{-1} \mathbf{J}^T \mathbf{e} \tag{12}$$

where μ is an adaptive parameter that is decreased if the performance function is decreasing and increased if the performance function is increasing. The parameter $\mathbf{I}$ denotes the identity matrix. For large μ values, the Levenberg–Marquardt algorithm becomes the gradient descent algorithm (small step size). When μ is zero, the weight update method approximates the Newton method, which is faster and more accurate near an error minima [11]. The aim is therefore to shift toward the approximate Newton algorithm as soon as possible. The Levenberg–Marquardt algorithm is designed for training moderate-sized feed-forward neural networks, with less than 100 weights, and least squares problems that are approximately linear [11]. Hence, the algorithm is well suited for the task of training an ANN path-loss model. The Levenberg–Marquardt algorithm has been reported to train ANNs at a rate of 10 to 100 times faster (depending on the problem) than the standard gradient descent backpropagation algorithm [11], [18]. In Section VI, it is observed that the Levenberg–Marquardt algorithm is about 1000 times faster than the standard backpropagation algorithm, which potentially enables for a substantially larger amount of training data to be incorporated in the training process.

D. Generalization

For the purpose of achieving improved generalization properties, both ES [15] and BR [21], [22] have been evaluated utilizing the available measurement data. In this paper, ES was incorporated with the gradient descent algorithm to serve as the benchmark method. When trying to minimize the required training time while maintaining the ANN's generalization characteristics, both ES and BR were used with the Levenberg–Marquardt algorithm [11], [23]. To support the understanding of how to improve the generalization properties of ANN-based path-loss prediction models, the next two paragraphs explain the ES and BR methodologies, respectively.

ES: To utilize ES, the measurement data [13], [14], [24] are divided into three subsets (training, validation, and evaluation), where the training set is used for computing the gradient and the ANN weights. The errors obtained from the validation set are monitored during the training. In this paper, the number of input–output data pairs in the validation set is chosen to be 20% of the full training set [25]. When the network is starting to overfit the data from the training set, the errors obtained from the validation set usually start to increase. When the validation error has increased for a specified number of iterations, the training stops and the weights and biases at the minimum of the validation error are returned.

BR: In BR, a modified linear combination of squared errors and weights are minimized so that, at the end of training, the resulting ANN provides good generalization properties [21]–[23]. The method of BR does not require a validation data set and, therefore, only requires two data sets (training and evaluation). Hence, all training data can directly be used in the training process, which is preferable if only small amounts of training data are available. One drawback is that BR, in general, requires more iterations than ES to converge [11].

III. Measurement Procedure

The work presented in this paper utilizes mobile radio wave propagation measurements at a carrier frequency of 881.52 MHz that originate from two measurement campaigns performed in rural Western Australia in June 2003 [24] and May 2004 [14] for the training and evaluation of different ANN models. The measurements were conducted using a CDMA pilot scanner, which was developed and built collaboratively by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Australian Telecommunications Cooperative Research Centre (ATcrc). For the purpose of mobile measurements, the measurement system also includes a Global Positioning System receiver and an omni-directional receiving antenna, which, during the measurements, were placed on the roof of a car at a height of approximately 1.7 m above ground.

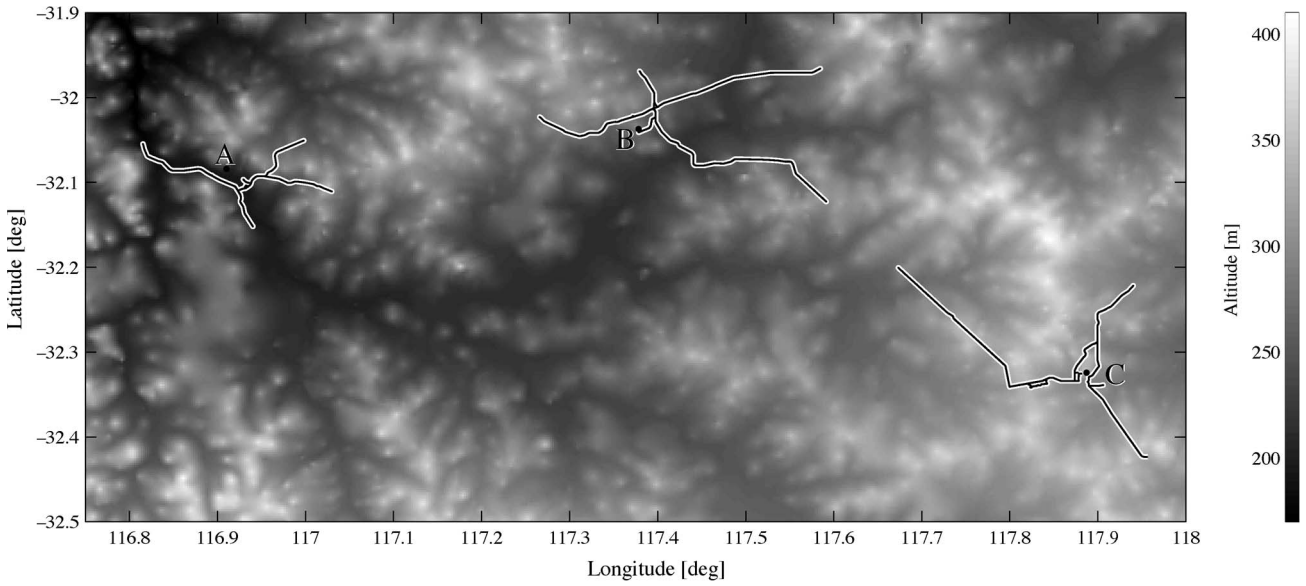


Fig. 3. DEM, with 9 s resolution, used for the ANN path-loss models. The measurement data originate from 19 routes and three BSs with omni-directional antenna characteristics in rural Western Australia [13]. From left to right, the BSs Beverley, Quairading, and Corrigin are marked with a dot and A, B, and C, respectively.

Without considering the influence of the roof of the car, the antenna's half-power beamwidth was measured to be close to 80°. The measurement system was controlled by custom software running on a laptop, which also logged all measurement data. The collected drive test data (19 routes) originate from three omni-directional BSs with macrocell characteristics [Beverley (A), Quairading (B), and Corrigin (C)] and cover more than 400 km of different rural terrains with different signal distributions. The measurement tracks from the area of interest plotted on a digital elevation map (DEM), together with BS positions and altitude shown in meters, can be seen in Fig. 3. At the time of the measurements, the heights for the three BSs (A, B, and C) were 30.5, 45, and 35 m, respectively. Note also that the maximum antenna separation distance was 24 km, and data originating from antenna separation distances less than 1 km were not analyzed. Hence, all measurements presented in this paper are considered to be performed in the far field. The correlation peak of the received pilot signal was sampled at 600 Hz. Typically, in these rural areas, mostly, only one pilot signal was received at the same time, hence limiting the effect of intercell interference. The BS positions were given by the register of radio communication licenses held at the Australian Communications and Media Authority. The BSs' transmitted pilot power, antenna height, and antenna pattern (omnidirectional) were provided by Telstra Wireless Access Services, Perth, Australia. Where possible, the BS transmitted power was verified by line-of-sight measurements near the BS.

The region from where the presented measurement results were obtained can be classified as open farmland with occasional large tree branches overhead (B and C). The measurement data originating from BS A are from a region with more vegetation and trees surrounding the receiving antenna. The roughness of the terrain in the three regions was estimated by the standard deviation of the terrain heights along the terrain path profile for every measurement position. The averaged value for each BS is given as follows: Beverley (A): 10.9 m, Quairading (B): 14.1 m, and Corrigin (C): 16.5 m. It should be noted that, in the measurements originating from BS B, the presence of a large built-up area including a huge grain-storage facility (>15 m height) has a noticeable impact on the measurements.

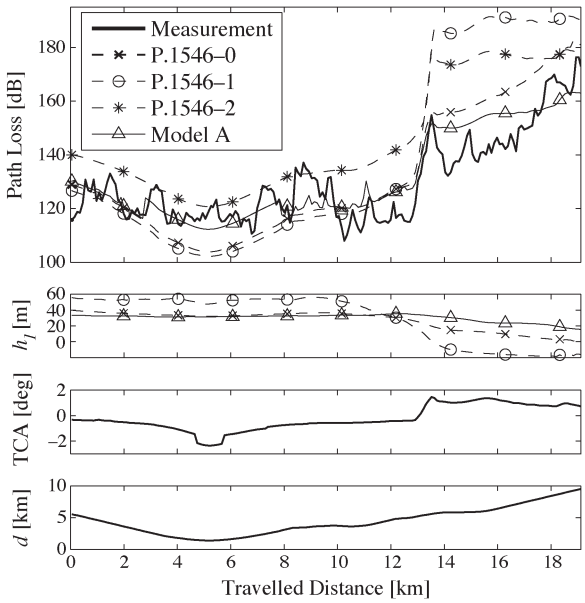


Fig. 4. P.1546 and Model A path-loss predictions for measurement 2003 A1 around the Beverley BS [13], where the transmitting/base antenna height is negative for some parts of the route (after 13 km traveled distance).

To be directly comparable with our previous journal paper [13] that compares different versions of the Recommendation ITU-R P.1546 model [26]–[28], the local average received powers used to train and evaluate the ANN models are computed by averaging signal measurements over a measurement track of 300 wavelengths, which corresponds to approximately 100 m. Note that all measurement path-loss results shown in Figs. 4–8 are averaged results.

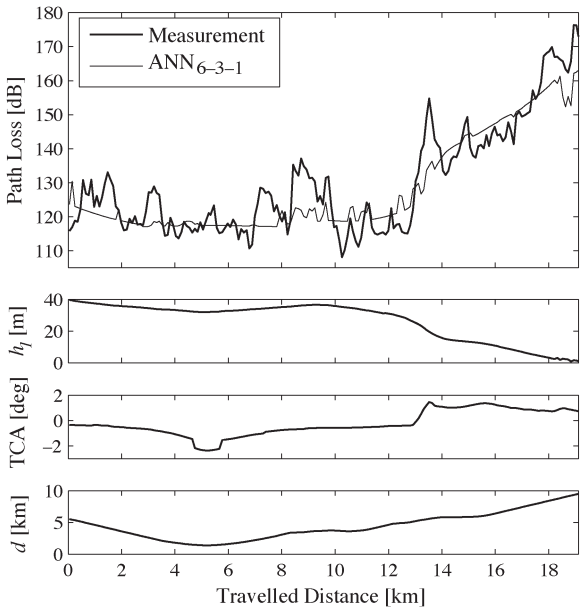


Fig. 5. Measurement 2003_A1 around the Beverley BS, where the mobile receiving antenna is on a slope positioned above the transmitting antenna.

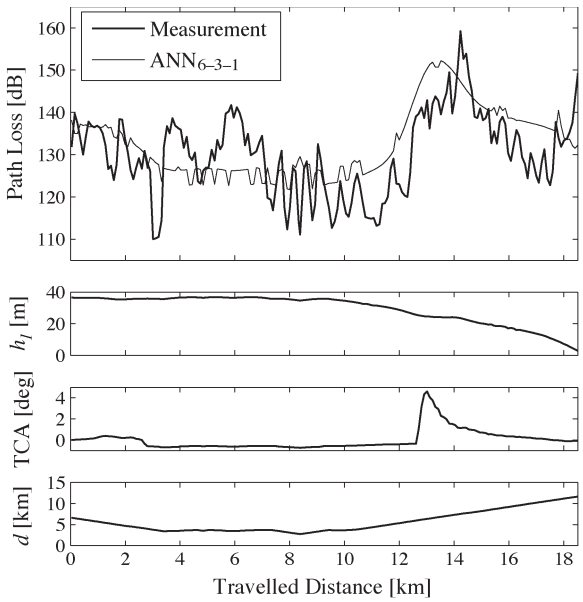


Fig. 6. Measurement 2004_A5 around the Beverley BS, where the transmitting/base antenna height is close to zero for parts of the route and a distinct peak in the TCA at 13 km traveled distance.

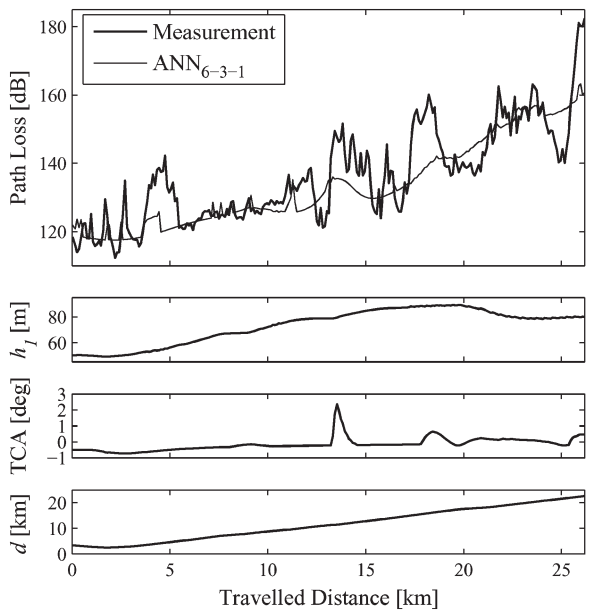


Fig. 7. Measurement 2004_B8 around the Quairading BS with a distinct peak in the TCA at 13 km traveled distance.

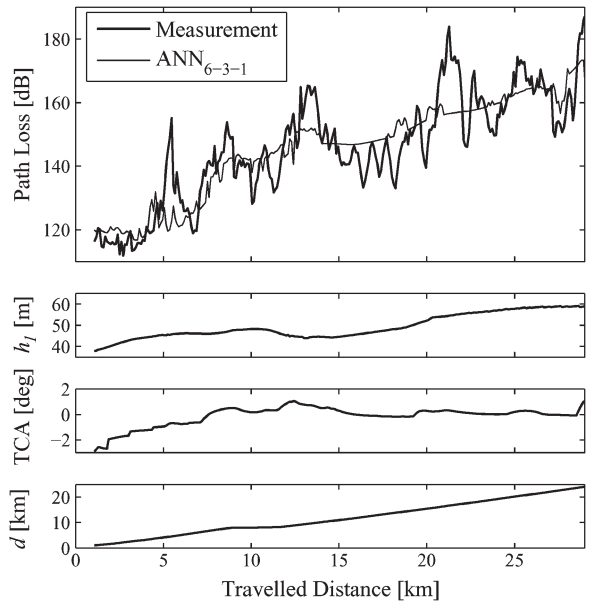


Fig. 8. Measurement 2003_C1 around Corrigin BS spanning over 1–24 km to the BS.

IV. Training and Prediction Parameters

In rural macrocell environments, path-loss strongly depends on transmitting and receiving the antenna-separation distance, the antenna heights, the terrain between the antennas, and the clutter surrounding the receiving antenna. In this paper, the ANN inputs are chosen to be the antenna-separation distance, the transmitting/base antenna height [26], the terrain clearance angle (TCA) [26], the terrain usage, the vegetation type (VT), and the vegetation density near the receiving antenna [13].

The terrain parameters are derived using a DEM with a 9 s resolution grid (approximately 250 m) supplied by Geoscience Australia (see Fig. 3). For an efficient ANN implementation, all input parameters are normalized to fall in the range between -1 and 1.

The following paragraphs provide the ANN input parameter definitions in order of importance for the path-loss prediction accuracy. More detailed descriptions of the input parameters can be found in [13].

Transmitting/Base Antenna Height h_1: The h_1 definition used in the ANN models is based on the definition of effective antenna height as in the Okumura model [29] and calculated as in the Recommendation ITU-R P.1546 [26]. Additional discussions regarding different transmitting/base antenna height definitions can be found in [13].

TCA: For land paths, a TCA correction may be added to increase the prediction accuracy, enabling obstacles close to the receiver site to be taken into account. The TCA is based on the angles relative to the horizontal between a line connecting the transmitting and receiving antennas and a line at the receiving antenna that clears all the terrain obstructions over a distance of up to 16 km, but not going beyond the base antenna [26].

Land Usage and Vegetation Information: To further enhance the ANN prediction accuracy, land usage and vegetation information may be incorporated in the ANN models. In [13], it was noted that the measured field strength is typically attenuated due to buildings near the BS (i.e., the receiving antenna is located below surrounding clutter height). Information regarding land usage (rural and built-up areas) was obtained by inspecting satellite photos and notes made during the measurement campaigns.

1) VT: The VT data used in the ANN analysis were provided by the Department of Agriculture, Western Australia. The data set contains information on two levels, namely, land-use data and VT (e.g., medium woodland and York gum). In the rural area where the measurements were performed, ten different VTs were identified [13]. For simplicity, these VTs were divided into three groups corresponding to woodland, shrublands, and no vegetation. The data extracted from the VT database correspond to the most occurring VT within a 100 m $\times$ 100 m area, which is typically only one type.

2) Vegetation density: The vegetation density data used in the analysis were provided by the CSIRO Mathematical and Information Sciences, and Land Monitor, Western Australia. The perennial crown density information is given on an accurate 25 m grid and is represented by a scalar value ranging from 0% to 100% corresponding to no vegetation and dense vegetation, respectively. The crown density metric corresponds to the estimated percentage of land covered by the tree crown [30]. The vegetation density near the receiver (VDN) corresponds to the average vegetation density within a 100 m $\times$ 100 m area surrounding the receiver.

V. Statistical Analysis Metrics

The first-order statistics, the correlation factor, and the average total hit rate error (AHRE) have been used to evaluate the results. Predicted and measured values are denoted p_i and m_i, respectively, and are given on a logarithmic (in decibels) scale. The prediction error ϵ_i is expressed as

$$\epsilon_i = p_i - m_i, \qquad i = 1, 2, \ldots, N \tag{13}$$

where N is the number of samples corresponding to 300 wavelengths, as described in Section III. The maximum prediction error, the mean prediction error, and the error standard deviation are denoted $\epsilon_{\max}$[dB], $\bar{\epsilon}$[dB], and σ_ϵ[dB], respectively.

The correlation coefficient is denoted r_ϵ and provides a measure of the degree of linear relationship between measured m_i and predicted p_i values.

As a complement to the first-order statistics and the correlation factor, the prediction accuracy is also quantified using the AHRE (in percentage), as introduced in [24]. This single-value metric is derived from the total hit rate (THR) curve described in [31]. The location-specific THR curve is used as a direct indication of the quality of the prediction model. Given a path-loss threshold L_T, if both predicted and measured path-loss values are greater than, less than, or equal to L_T, the prediction is regarded as correct irrespective of the deviation of the predicted from the measured value. The AHRE is the mean deviation from 100% THR and is expressed as

$$\text{AHRE} = \frac{1}{N_{L_T}} \sum_{L_T = L_{T,\min}}^{L_{T,\max}} 100\% - \text{THR}(L_T) \tag{14}$$

where L_T is the path-loss threshold, and N_{L_T} is the number of THR points. The thresholds $L_{T,\min}$ and $L_{T,\max}$ are chosen so that the AHRE may be interpreted as the area between the THR curve and 100%. The method is useful in assessing the validity of a model where coverage is simply determined by a threshold value. A small AHRE value indicates a good fit between predicted and estimated values.

VI. Training and Evaluation

In this section, different ANNs are evaluated using the measurement data introduced in Section III. Conveniently, the prediction evaluations in this paper can directly be compared with previously published results [13]. For easy comparison, Table I presents the result overview for the proposed Model A, which, among other things, addressed a shortcoming of the effective antenna height definition in the Recommendation ITU-R P.1546 and the well-known OH rural model. The statistical analysis average results show that Model A's predictions are more accurate compared with the predictions provided by the OH rural model for all statistical analysis metrics ($\epsilon_{\max}$, $\bar{\epsilon}$, σ_ϵ, r_ϵ, and AHRE). In particular, the maximum prediction error $\epsilon_{\max}$ is lower for all measurement routes, except for measurement 2004_B6, where Model A and the OH rural model provide 19.0 and 18.4 dB, respectively. Furthermore, Fig. 4 shows a path-loss prediction example for the different Recommendation ITU-R P.1546 versions and Model A. It was previously concluded that the P.1546 models, for certain scenarios, provide abnormal path-loss predictions due to negative transmitting/base antenna height definition h_1 values being produced. Model A therefore proposed a new h_1 definition that efficiently mitigates this problem. This P.1546 shortcoming and Model A's prediction accuracy improvement can clearly be seen in Fig. 4 after 13 km traveled distance.

From our previous publication [13], it is known that the obtained maximum errors produced by the different versions of the Recommendation ITU-R P.1546 (-0, -1, and -2) and the proposed Model A are 24.1, 53.2, 44.5, and 17.9 dB, respectively. In Section VI-C, Model A's prediction results are compared with the prediction results produced by different ANN models. Note that, in this paper, the original Recommendation ITU-R P.1546 definition of h_1 has been used as a training and prediction parameter for the ANNs.

This section is organized as follows. In Section VI-A, the ANNs are trained and evaluated using measurement data originating from the same BS (C). The training data are obtained

TABLE I
STATISTICS FOR MODEL A AND THE OH RURAL MODEL [13]

	ϵ_{max} [dB]		$\bar{\epsilon}$ [dB]		σ_ϵ [dB]		r_ϵ		AHRE [%]	
Meas.	Model A	OH	Model A	OH	Model A	OH	Model A	OH	Model A	OH
2003_A1	17.9	45.5	1.7	-12.4	7.7	11.4	0.89	0.75	9.5	18.4
2003_A2	18.2	34.2	-0.7	-8.5	8.5	9.5	0.67	0.49	13.8	16.5
2004_A3	17.4	24.0	3.4	-5.4	7.7	8.4	0.47	0.26	22.8	21.9
2004_A4	19.7	22.3	3.6	-5.1	8.1	9.2	0.39	0.13	23.8	23.0
2004_A5	27.3	31.6	4.7	-8.3	11.5	8.8	0.61	0.45	19.4	20.1
2003_B1	15.3	22.9	-2.2	-7.9	5.8	7.3	0.77	0.59	13.9	24.0
2003_B2	23.5	42.7	6.8	-4.1	9.3	10.3	0.82	0.78	12.1	9.6
2003_B3	31.2	49.2	-0.5	-10.1	10.2	12.0	0.75	0.62	10.7	14.4
2004_B4	29.2	37.0	-4.7	-13.6	6.8	9.2	0.80	0.59	12.0	24.5
2004_B5	25.3	33.3	1.9	-10.2	6.8	6.4	0.74	0.71	13.2	24.2
2004_B6	19.0	18.4	8.4	-5.5	5.8	5.9	0.59	0.58	26.9	22.4
2004_B7	18.7	26.4	-4.6	-13.6	6.9	6.7	0.58	0.59	19.6	30.2
2004_B8	22.0	41.4	2.8	-8.2	8.4	9.0	0.82	0.77	10.6	12.2
2004_B9	22.1	33.0	1.3	-8.1	9.5	9.7	0.71	0.58	14.1	16.5
2003_C1	31.1	45.4	-0.3	-15.8	8.3	9.8	0.87	0.84	8.2	17.4
2004_C2	24.5	30.4	4.9	-7.2	8.4	8.5	0.81	0.75	13.6	13.5
2004_C3	21.1	32.0	4.4	-4.9	8.0	8.6	0.76	0.68	12.4	11.2
2004_C4	27.6	32.7	4.7	-2.6	9.4	8.2	0.67	0.65	14.6	13.5
2004_C5	22.0	40.6	1.5	-12.8	8.2	8.7	0.86	0.80	9.2	15.9
Average	22.8	33.8	2.0	-8.7	8.2	8.8	0.71	0.61	14.8	18.4

from four routes, and the evaluation data are obtained from a different fifth route. In Section VI-B, the ANNs are trained with data from BS C (see Section VI-A) and are evaluated using measurement data originating from BS A. This training approach was performed to show that the ANNs can perform well in a different environment compared with from where the training data were gathered. Finally, Section VI-C presents a training and evaluation approach where the training data are obtained from one measurement route around each BS (A, B, and C). The training data are chosen so that the range covers most of the data used in the evaluation. Furthermore, the impact on training time, prediction accuracy, and generalization characteristics when performing the training using the Levenberg–Marquardt algorithm is presented.

A. Same Cell—Different Route

Here, the first ANN training/evaluation approach utilizes measurement data originating from the same omni-directional BS (C). Five measurements (different routes) have been obtained from this Corrigin BS. The measurement data cover a traveled distance of 129 km and consist of 1238 averaged pilot power samples, where each average power value has been estimated from measurement data spanning over 300 wavelengths. The ANN models are trained utilizing 500 randomly chosen input–output data pairs [see (6)] from four of the measurements (2003_C1, 2004_C2, 2004_C3, and 2004_C4) and evaluated using measurement data from the fifth measurement track (2004_C5).

Table II presents the numerical results for different-sized ANNs (ANN_{6-1}, ANN_{6-3-1}, and $\text{ANN}_{6-7-3-1}$). The numerical results for the training (TR) data correspond to the 500 data pairs taken from measurements 2003_C1, 2004_C2, 2004_C3, and 2004_C4. The numerical results for the evaluation (EV) data correspond to measurement 2004_C5 (201 samples). The

TABLE II
STATISTICAL ANALYSIS SUMMARY: SAME CELL—DIFFERENT ROUTE

Data	Metric	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$
TR	ϵ_{max} [dB]	27.0	31.9	32.5
	$\bar{\epsilon}$ [dB]	0.4	0.3	-0.1
	σ_ϵ [dB]	7.0	7.0	6.7
	r_ϵ	0.81	0.84	0.82
	AHRE [%]	10.2	9.2	10.3
EV	ϵ_{max} [dB]	25.6	23.1	22.2
	$\bar{\epsilon}$ [dB]	-4.2	-3.5	-2.8
	σ_ϵ [dB]	7.7	7.4	7.3
	r_ϵ	0.85	0.86	0.87
	AHRE [%]	9.5	8.8	8.5

results show that, for this scenario and for the TR data, increasing the number of hidden layers and neurons slightly decreases the prediction accuracy for some of the statistical analysis metrics ($\epsilon_{\max}$, σ_ϵ, and AHRE). In particular, it can be seen that the maximum error $\epsilon_{\max}$ increases from 27.0 to 32.5 dB, respectively, when the ANN is extended from ANN_{6-1} to $\text{ANN}_{6-7-3-1}$. For the EV data, it can be seen that extending the neuron model ANN_{6-1} to ANN_{6-3-1} and ANN_{6-3-1} slightly increases the prediction accuracy for all statistical analysis metrics. As an example, the standard deviation σ_ϵ decreases from 7.7 to 7.3 dB, respectively, when the ANN is extended from ANN_{6-1} to $\text{ANN}_{6-7-3-1}$. Note that the TR and EV data are taken from different measurement routes, and hence, some deviations due to specific topographical features shown in the analysis of the TR data are not present in the analysis of the EV data.

B. Different Cell

The different cell approach evaluates the ANNs, which are trained as described in the previous Section VI-A, using 188 input–output data pairs obtained from a different

TABLE III
STATISTICAL ANALYSIS SUMMARY: DIFFERENT CELL

Data	Metric	ANN_{6-1}	ANN_{6-3-1}	$ANN_{6-7-3-1}$
EV	ϵ_{max} [dB]	35.8	32.2	29.4
	$\bar{\epsilon}$ [dB]	-5.1	-4.9	-3.3
	σ_ϵ [dB]	9.3	8.3	8.1
	r_ϵ	0.87	0.87	0.86
	AHRE [%]	11.3	10.4	9.5

BS (2003_A1). The purpose of this training/evaluation approach is to see how well the different ANNs generalize. From previous investigations [13], [32]–[34], it was known that the propagation scenario that occurs after 13 km traveled distance, where the receiving antenna is on a slope above the BS, may be hard to predict, depending on how the transmitting/base antenna height is calculated and incorporated in the path-loss model.

Table III shows the prediction results for the different-sized ANNs using EV data from measurement 2003_A1. It can be seen that all statistical analysis metrics, except for the correlation factor r_ϵ, slightly improve when extending the ANN_{6-1} to $ANN_{6-7-3-1}$. Furthermore, for this scenario, the prediction accuracy somewhat increases with the additional number of hidden layers and neurons. It can be seen that the average error $\bar{\epsilon}$ decreases from -5.1 to -3.3 dB and the standard deviation σ_ϵ decreases from 9.3 to 8.1 dB when extending the ANN_{6-1} to $ANN_{6-7-3-1}$, respectively. It can be concluded that the ANN models generalize quite well when being evaluated using measurement data that originate from another BS and a different region. During the training, it was noted that the $ANN_{6-7-3-1}$ is much more sensitive to the training data compared with the ANN_{6-1} and ANN_{6-3-1}.

C. All Cells

In this approach, 700 randomly chosen input–output pairs originating from three BSs (2003_A1, 2003_B2, and 2003_C1) are used to train the ANNs. Together, these three measurements cover most of the input and output parameter range to be expected in the evaluation. Hence, the ANN's task is to interpolate and, to some degree, also extrapolate from the given input data. The 19 measurement routes cover a traveled distance of approximately 400 km and consist of 3823 averaged pilot power samples averaged over 300 wavelengths. Utilizing the backpropagation gradient descent algorithm with ES, the training process requires approximately 10 000, 100 000, and 150 000 iterations in batch mode for the ANN_{6-1}, ANN_{6-3-1}, and $ANN_{6-7-3-1}$ models, respectively.

Figs. 5–8 show the measured and predicted path-loss for four measurement routes (2003_A1, 2004_A5, 2004_B8, and 2003_C1). For clarity, only the prediction curves provided by the ANN_{6-3-1} model and the EV data are shown in the figures.

Fig. 5 shows the ANN path-loss prediction for measurement 2003_A1 (Beverley BS), where, for parts of the route, after approximately 13 km traveled distance, the mobile receiving antenna is on a slope above the BS. This particular scenario causes some severe overestimation (+40 dB; see Fig. 4 after 13 km traveled distance) of the path-loss when using the Recommendation ITU-R P.1546 model due to how the transmitting/base antenna height is defined and incorporated in the model [13]. Using an ANN model that has been trained with measurement data obtained from this specific scenario, it can be seen that the predicted path-loss accuracy is very good, even for the region after 13 km traveled distance.

Fig. 6 depicts the ANN path loss prediction for another measurement route around Beverley BS (2004_A5). At approximately 13 km traveled distance, a sharp peak in the TCA can be seen. For this specific scenario, the P.1546 model produces a correction overshoot (20–40 dB), which can result in unrealistic path-loss predictions [13]. Now, when using the ANN prediction model, the path-loss correction due to TCA is smoother and does not significantly overshoot.

Fig. 7 illustrates the ANN path loss prediction for measurement route 2004_B8 (Quairading BS). At around 4.5 km traveled distance, there is a huge grain-storage facility (>15 m height) that affects the measurements. Furthermore, at a traveled distance of approximately 13 km, a distinct TCA peak is present. In the path loss prediction, it can be seen that the ANN does not create an overshoot due to this peak, i.e., the response is relatively smooth.

Fig. 8 presents the ANN path loss prediction for a measurement route around Corrigin BS (2003_C1). It can be seen that the overall predictions provide good agreement with the measurements without introducing any major abnormalities. Note that, at approximately 5 and 22 km traveled distances, there are distinct increases in the measured path loss. These peaks are due to features that are not described by the topographical data and, hence, not reflected by the input parameters.

Tables IV–VI show the statistical analysis results for the evaluated ANN models and the 19 measurement routes. The numerical values given in Tables IV and V correspond to EV data. For comparison, the two bottom rows in all tables provide the average prediction results for TR and EV data ($\text{Average}^{\text{TR}}$ and $\text{Average}^{\text{EV}}$).

As a direct comparison with our previously published journal paper [13], Table I shows the obtained path loss predictions using Model A (our proposed modifications to Recommendation ITU-R P.1546) and the OH model. For Model A, the average maximum error $\epsilon_{\max}$ is approximately 23 dB, the mean error $\bar{\epsilon}$ is 2 dB, and the error standard deviation σ_ϵ is near 8 dB. For the same data, the OH model provides an average maximum error $\epsilon_{\max}$ of near 34 dB, the mean error $\bar{\epsilon}$ is approximately -9 dB, and the error standard deviation σ_ϵ is close to 9 dB.

Table IV presents the first-order statistical analysis results for the three different-sized ANN models. It can be seen that the average prediction results provided by the ANN models do not significantly differ. For the EV data ($\text{Average}^{\text{EV}}$), the average maximum error $\epsilon_{\max}$ is approximately 21–22 dB, the mean error $\bar{\epsilon}$ is close to zero, and the error standard deviation σ_ϵ is near 7 dB. By inspecting the $\text{Average}^{\text{TR}}$ and $\text{Average}^{\text{EV}}$ results in Table IV, it can be seen that extending the ANN structures results in greater prediction-accuracy improvements for the TR data compared with the EV data.

Table V shows the correlation coefficient and AHRE corresponding to the same ANNs as in Table IV. Here, as well, the prediction results are quite similar for the different ANN

TABLE IV
FIRST-ORDER STATISTICS FOR ANN_{6-1}, ANN_{6-3-1}, AND $\text{ANN}_{6-7-3-1}$ MODELS

Meas.	ϵ_{max} [dB]			$\bar{\epsilon}$ [dB]			σ_ϵ [dB]		
	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$
2003_A1	20.8	19.9	17.1	0.6	0.7	-0.8	7.2	7.7	6.2
2003_A2	19.4	16.7	18.1	-0.7	-1.1	-2.2	7.3	7.0	6.7
2004_A3	14.3	14.6	13.1	3.0	3.0	1.6	5.4	5.5	5.6
2004_A4	14.3	14.7	12.9	3.5	3.3	1.9	5.9	5.9	6.0
2004_A5	27.8	22.7	21.8	0.9	1.3	-0.8	9.7	9.8	8.7
2003_B1	16.3	12.1	16.0	2.1	0.6	1.1	7.2	6.4	6.6
2003_B2	19.6	19.7	20.4	4.5	2.5	2.4	7.9	7.5	8.0
2003_B3	34.9	27.2	33.5	-1.6	-1.0	-3.4	8.9	7.3	10.5
2004_B4	32.3	32.0	32.5	-3.7	-3.6	-3.8	8.5	7.3	7.5
2004_B5	28.6	28.2	28.8	-3.1	-3.8	-4.5	7.5	6.0	5.8
2004_B6	13.8	14.2	16.0	0.7	-0.6	-1.4	5.5	4.8	4.9
2004_B7	17.3	17.0	16.6	-2.7	-2.7	-2.0	5.6	5.8	5.6
2004_B8	21.5	22.3	21.8	-0.5	-2.1	-2.6	7.8	7.3	7.2
2004_B9	24.9	21.6	22.4	-1.3	-1.7	-2.0	7.5	6.9	6.9
2003_C1	24.9	33.6	30.9	-3.0	-0.4	1.2	8.0	9.1	8.8
2004_C2	24.5	24.1	22.5	0.6	2.6	3.3	8.7	8.1	7.1
2004_C3	25.0	16.6	15.1	1.4	2.8	2.3	7.3	6.5	6.4
2004_C4	19.2	21.5	20.3	5.6	5.3	5.2	6.2	6.2	6.2
2004_C5	24.0	22.9	19.8	-2.7	-0.3	0.3	7.9	8.0	7.6
$\text{Average}^{\text{TR}}$	26.0	25.7	22.8	0.4	0.3	-0.2	8.3	8.2	7.7
$\text{Average}^{\text{EV}}$	22.3	21.1	21.0	0.2	0.3	-0.2	7.4	7.0	7.0

TABLE V
CORRELATION COEFFICIENT AND AHRE FOR ANN_{6-1}, ANN_{6-3-1}, AND $\text{ANN}_{6-7-3-1}$ MODELS

Meas.	r_ϵ			AHRE [%]		
	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$	ANN_{6-1}	ANN_{6-3-1}	$\text{ANN}_{6-7-3-1}$
2003_A1	0.93	0.81	0.84	11.1	10.7	9.2
2003_A2	0.76	0.74	0.76	10.4	10.8	10.6
2004_A3	0.41	0.45	0.44	15.9	17.9	16.7
2004_A4	0.36	0.40	0.39	17.1	18.6	17.7
2004_A5	0.52	0.45	0.58	15.5	16.6	13.9
2003_B1	0.78	0.80	0.77	16.9	14.6	15.3
2003_B2	0.88	0.85	0.86	8.2	7.9	8.6
2003_B3	0.84	0.78	0.80	12.7	9.2	10.6
2004_B4	0.79	0.82	0.81	14.6	11.4	12.1
2004_B5	0.73	0.76	0.77	16.2	13.1	13.3
2004_B6	0.71	0.73	0.74	14.1	12.4	12.1
2004_B7	0.51	0.50	0.54	16.7	16.6	15.6
2004_B8	0.85	0.85	0.86	8.4	7.9	7.9
2004_B9	0.74	0.78	0.78	11.1	10.2	10.3
2003_C1	0.86	0.87	0.85	10.1	9.4	9.4
2004_C2	0.76	0.80	0.84	11.3	11.5	10.9
2004_C3	0.81	0.82	0.83	10.0	9.6	9.2
2004_C4	0.75	0.75	0.74	14.6	13.0	13.1
2004_C5	0.86	0.85	0.87	9.3	9.2	8.8
$\text{Average}^{\text{TR}}$	0.87	0.87	0.89	9.2	8.6	8.1
$\text{Average}^{\text{EV}}$	0.73	0.73	0.74	12.8	12.1	11.9

TABLE VI
NEURON MODEL STATISTICAL ANALYSIS—LEVENBERG–MARQUARDT TRAINING ALGORITHM INCORPORATING ES AND BR

	ϵ_{max} [dB]		$\bar{\epsilon}$ [dB]		σ_ϵ [dB]		r_ϵ		AHRE [%]	
	ANN_{6-1}^{ES}	ANN_{6-1}^{BR}	ANN_{6-1}^{ES}	ANN_{6-1}^{BR}	ANN_{6-1}^{ES}	ANN_{6-1}^{BR}	ANN_{6-1}^{ES}	ANN_{6-1}^{BR}	ANN_{6-1}^{ES}	ANN_{6-1}^{BR}
$\text{Average}^{\text{TR}}$	26.6	16.7	0.1	0.5	8.4	6.3	0.87	0.92	8.7	6.6
$\text{Average}^{\text{EV}}$	21.8	21.6	0.0	0.5	6.9	6.9	0.74	0.75	11.8	12.2

models. For the EV data ($\text{Average}^{\text{EV}}$), the correlation coefficient r_ϵ is greater than 0.7, and the AHRE is, on average, about 12% for all predictions. The $\text{Average}^{\text{TR}}$ and $\text{Average}^{\text{EV}}$ results presented in Table V also show the correlation coefficient r_ϵ and the AHRE to be significantly better for the TR data, compared with the EV data. Hence, the differences between the ANNs for

some of the predictions originate from how well the ANN is able to adapt to different scenarios.

The overall conclusion that can be drawn from Tables IV and V is that an ANN_{6-1} model (neuron model), for the given training approach and available measurement data, is sufficient to obtain an accurate path-loss prediction model that generalizes well.

Furthermore, the numerical values may directly be compared to the prediction results achieved for the Recommendation ITU-R P.1546 versions (P.1546-0, P.1546-1 and P.1546-2), the OH model, and Model A evaluated in [13]. Model A is a modified/improved version of the Recommendation ITU-R P.1546 in which a modified definition of transmitting/base antenna height and field-strength attenuation due to different land usage and vegetation are included. It is noted that the average prediction results obtained with the ANN_{6-1} model, for the available measurement data, are more accurate than the results achieved with Model A [13].

It should be noted that an $\mathrm{ANN}_{6-7-3-1}$ model requires 150 000 iterations (gradient descent and ES) in the backpropagation training process to converge to its optimum. Therefore, as a comparison, the Levenberg–Marquardt training algorithm was used to train the neuron model incorporating both ES and BR (ANN_{6-1}^{ES} and ANN_{6-1}^{BR}). For ES and BR stopping criteria, only approximately 10 and 20 iterations in batch mode are required, respectively, to train the ANN.

Table VI shows the average prediction results for the ANN_{6-1}^{ES} and ANN_{6-1}^{BR} models corresponding to both TR and EV data. It can be seen that the overall average path-loss prediction results are on par with the previous evaluation results shown in Tables IV and V. In summary, these simple path loss models provide accurate prediction results and generalize well. Note that, when evaluating the Levenberg–Marquardt ANNs using the TR data (700 input–output pairs from 2003_A1, 2003_B2, and 2003_C1), ANN_{6-1}^{BR} seems to be more accurate compared to ANN_{6-1}^{ES} (see Table VI). Finally, the average ANN_{6-1}^{ES} and ANN_{6-1}^{BR} prediction results for the EV data $(\mathrm{Average}^{\mathrm{EV}})$ are very much the same for all statistical analysis metrics. For this application, the differences would more likely originate from the used training data rather than the algorithm performance.

VII. Conclusion and Future Work

The work in this paper has investigated ANN path-loss prediction models for rural macrocell environments. A CDMA pilot scanner has been used to obtain measurement data in a commercial IS-95 mobile telephone network in rural Western Australia. The collected data and extracted terrain and topographical parameters have then been used to train and evaluate the different ANNs.

Different-sized ANNs, ranging from the neuron model to multilayered feed-forward networks, have been analyzed to determine the network complexity required to obtain accurate path-loss predictions. Furthermore, to reduce training time while maintaining prediction accuracy and generalization properties, faster training algorithms, such as Levenberg–Marquardt and ES and BR have been incorporated in the training process.

The ANN model evaluation using the first-order statistics, the correlation factor, and the AHRE metric shows that, for the given drive test data collected in rural Western Australia, using the given input parameters, the ANN models perform very well in comparison with other common path-loss models of similar complexity, such as the Recommendation ITU-R P.1546 and OH models.

It has been found that more complex ANNs (several hidden layers and neurons) do not considerably increase prediction accuracy. It has also been found during the training process that more complex ANNs may result in decreasing generalization properties.

Another important result was that larger feed-forward networks, due to overtraining, are more sensitive to the training data and may give less-accurate predictions when presented with inputs outside the training parameter space. On the other hand, when these more complex ANNs are presented with data similar to the training set, the predictions can be very accurate. However, the focus in this paper has been on model robustness and simplicity. When training an ANN_{6-1} using the Levenberg–Marquardt algorithm instead of an $\mathrm{ANN}_{6-7-3-1}$ using the standard backpropagation algorithm, the training time may be decreased from approximately 150 000 to 10 iterations. Hence, there is more available time to incorporate more training data. The most important aspect in creating an accurate and robust ANN path-loss model seems to be the training data selection. The input–output training data pairs should therefore describe as large a portion of the propagation problem to which the model is expected to be exposed.

In summary, it can be said that the trained neuron model provides more accurate prediction results than the Recommendation ITU-R P.1546, Model A, and the OH model. For the available measurement data, the average ANN prediction results were 1) maximum error: 22 dB; 2) mean error: 0 dB; and 3) standard deviation: 7 dB, correlation coefficient: 0.75, and AHRE: 12%. It should be noted that some of the large path loss variations, such as those due to blockage from large man-made objects, cannot be predicted due to lack of topographic databases.

Future work should incorporate careful selection of the training data, given that the training set describes a wider range of the propagation problem. This will increase the prediction accuracy and generalization characteristics of the ANN. It would also be very interesting to train an ANN using a very large set of measurement data, such as the data sets available to the International Telecommunications Union organization. A more diverse training data set would particularly be of benefit for the application of TCA correction, which, in the Recommendation ITU-R P.1546 for certain scenarios, seems to overcorrect the predictions [13]. It would also be of interest to compare the ANN predictions in this paper to a calibrated OH model, as described in [35].

Note that, since the publication of [13], a new version of the Recommendation ITU-R P.1546 has been released (P.1546-3 [36]). After evaluating the new P.1546-3 model, it was concluded that the overall prediction accuracy for these rural areas is slightly worse compared with the previous version P.1546-2. In particular, this is due to the simplified

definition of TCA and how the corresponding correction is applied.

ACKNOWLEDGMENT

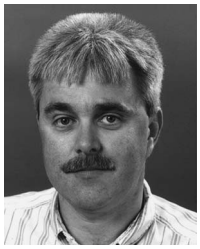
The authors would like to thank U. Engelke and A. Pollok for their assistance with the measurement campaigns, Telstra Wireless Access Services for providing the BS characteristics, J. Wallace and Dr. J. Chia of CSIRO Mathematical and Information Sciences and Land Monitor, Western Australia, for providing the digital vegetation density data, D. Shepherd of the Department of Agriculture, Western Australia, for making the VT database available, Dr. A. Weily for proofreading and providing constructive comments, which greatly improved the quality of this paper, and the anonymous reviewers for their valuable comments.

REFERENCES

[1] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, Dec. 2001.
[2] "Coverage prediction for mobile radio systems operating in the 800/900 MHz frequency range," *IEEE Trans. Veh. Technol.*, vol. 37, no. 1, pp. 3–72, Feb. 1988.
[3] R. J. Luebbers, "Propagation prediction for hilly terrain using GTD wedge diffraction," *IEEE Trans. Antennas Propag.*, vol. AP-32, no. 9, pp. 951–955, Sep. 1984.
[4] G. A. Hufford, "An integral equation approach to the problem of wave propagation over an irregular surface," *Q. J. Mech. Appl. Math.*, vol. 9, no. 4, pp. 391–404, Jan. 1952.
[5] C. A. Zelley and C. C. Constantinou, "A three-dimensional parabolic equation applied to VHF/UHF propagation over irregular terrain," *IEEE Trans. Antennas Propag.*, vol. 47, no. 10, pp. 1586–1596, Oct. 1999.
[6] K. E. Stocker and F. M. Landstorfer, "Empirical prediction of radiowave propagation by neural network simulator," *Electron. Lett.*, vol. 28, no. 12, pp. 1177–1178, Jun. 1992.
[7] I. Popescu, D. Nikitopoulos, P. Constantinou, and I. Nafornita, "ANN prediction models for outdoor environment," in *Proc. 17th IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Helsinki, Finland, Sep. 2006, pp. 1–5.
[8] Z. Stanković, B. Milovanović, M. Veljković, and A. Dordević, "The hybrid-neural empirical model for the electromagnetic field level prediction in urban environments," in *Proc. 7th Semin. Neural Netw. Appl. Elect. Eng.*, Belgrade, Serbia, Sep. 2004, pp. 189–192.
[9] A. Nešković, N. Nešković, and D. Paunović, "Indoor electric field level prediction model based on the artificial neural networks," *IEEE Commun. Lett.*, vol. 4, no. 6, pp. 190–192, Jun. 2000.
[10] J. D. Parsons, *The Mobile Radio Propagation Channel*. Chichester, U.K.: Wiley, 2000.
[11] H. Demuth and M. Beale, *Neural Network Toolbox*. Natick, MA: MATLAB, Jan. 2003.
[12] *Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System*, TIA/EIA/IS-95, Jul. 1993.
[13] E. Östlin, H. Suzuki, and H.-J. Zepernick, "Evaluation of the propagation model Recommendation ITU-R P.1546 for mobile services in rural Australia," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 38–51, Jan. 2008.
[14] E. Östlin, H.-J. Zepernick, and H. Suzuki, "Macrocell radio wave propagation prediction using an artificial neural network," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, Los Angeles, CA, Sep. 2004, vol. 1, pp. 57–61.
[15] R. Rojas, *Neural Networks: A Systematic Introduction*. Berlin, Germany: Springer-Verlag, 1996.
[16] M. T. Hagan and H. B. Demuth, "Neural networks for control," in *Proc. Amer. Control Conf.*, San Diego, CA, 1999, vol. 3, pp. 1642–1656.
[17] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*. Cambridge, MA: MIT Press, 1986.
[18] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.
[19] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*. Boston, MA: PWS, 1996.
[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.
[21] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, May 1992.
[22] D. J. C. MacKay, "A practical Bayesian framework for backpropagation networks," *Neural Comput.*, vol. 4, no. 3, pp. 448–472, May 1992.
[23] F. D. Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian regularization," in *Proc. Int. Joint Conf. Neural Netw.*, 1997, pp. 1930–1935.
[24] E. Östlin, H.-J. Zepernick, and H. Suzuki, "Evaluation of the new semi-terrain based propagation model Recommendation ITU-R P.1546," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, Orlando, FL, Oct. 2003, vol. 1, pp. 114–118.
[25] S. Haykin, *Neural Networks. A Comprehensive Foundation*. New York: Macmillan, 1994.
[26] *Method for Point-to-Area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*, ITU-R Recommendation P.1546, Oct. 2001.
[27] *Method for Point-to-Area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*, ITU-R Recommendation P.1546-1, Apr. 2003.
[28] *Method for Point-to-Area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*, ITU-R Recommendation P.1546-2, Sep. 2005.
[29] Y. Okumura, E. Ohmohri, T. Kawano, and K. Fukada, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elect. Commun. Lab.*, vol. 16, no. 9/10, pp. 825–873, Sep. 1968.
[30] S.-M. Lee, N. A. Clark, and P. A. Araman, "Automated methods of tree boundary extraction and foliage transparency estimation from digital imagery," in *Proc. 19th Biennial Workshop Color Photography Videography Airborne Imaging Resour. Assessment*, Logan, UT, Oct. 2003.
[31] A. S. Owadally, E. Montiel, and S. R. Saunders, "A comparison of the accuracy of propagation models using hit rate analysis," in *Proc. IEEE Semiannual Veh. Technol. Conf.*, Atlantic City, NJ, Oct. 2001, vol. 4, pp. 1979–1983.
[32] H. Suzuki and E. Östlin, "Effect on prediction accuracy using new definition of effective transmitting/base antenna height in Recommendation ITU-R P.1546-1—Comparison against measured data at 880 MHz in rural Australia," Sep. 2004, ITU-R Input Document 3K/39-E.
[33] H. Suzuki and E. Östlin, "Proposed revision to Recommendation ITU-R P.1546-1—Improved definition of effective transmitting/base antenna height," Sep. 2005, ITU-R Input Document 3K/72-E.
[34] E. Östlin, H. Suzuki, and H.-J. Zepernick, "Evaluation of a new effective antenna height definition in ITU-R Recommendation P.1546," in *Proc. 11th Asia-Pacific Conf. Commun.*, Perth, Australia, Oct. 2005, pp. 128–132.
[35] X. Liming and Y. Dacheng, "A recursive algorithm for radio propagation model calibration based on CDMA forward pilot channel," in *Proc. 14th IEEE Pers., Indoor, Mobile Radio Commun.*, Beijing, China, Sep. 2003, vol. 1, pp. 970–972.
[36] *Method for Point-to-Area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*, ITU-R Recommendation P.1546-3, Nov. 2007.

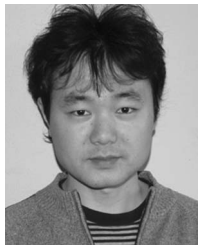
Erik Östlin (S'06–M'08) received the B.Sc. degree in electrical engineering from Högskolan Dalarna, Falun, Sweden, in 1999 and the M.Sc. degree in electrical engineering and the Ph.D. degree in applied signal processing from the Blekinge Institute of Technology, Ronneby, Sweden, in 2001 and 2009, respectively.

In 2002, he was invited to join the Australian Telecommunications Cooperative Research Centre (ATcrc) Wideband CDMA Scanner Project, Perth, Australia. Between April 2006 and February 2009, he was a Digital Signal Processing Engineer and Electro Acoustic Expert with Sensear Pty. Ltd., Perth. He is currently with the Blekinge Institute of Technology.



Hans-Jürgen Zepernick (M'94) received the Dipl.-Ing. degree in electrical engineering from the University of Siegen, Siegen, Germany, in 1987 and the Dr.-Ing. degree from the University of Hagen, Hagen, Germany, in 1994.

From 1987 to 1989, he was with the Radio and Radar Department of Siemens AG, Munich, Germany. He is currently a Professor and the Chair of radio communications with the School of Engineering, Blekinge Institute of Technology (BTH), Ronneby, Sweden. Prior to this appointment at BTH in October 2004, he held the positions of Professor of wireless communications with the Curtin University of Technology, Perth, Australia; Deputy Director of the Australian Telecommunications Research Institute (ATRI); and Associate Director of the Australian Telecommunications Cooperative Research Centre (ATcrc). He was the Leader of the ATcrc's wireless program and the Leader of its radio transmission technology project. He served on the Board of Management of ATRI and the Executive Research Committee of the ATcrc. He has been an active participant in the European FP7 Network of Excellence *EuroNF*, contributing to work packages on user-perceived quality of service. He is the author or a coauthor of some 150 technical papers and five book chapters in the areas of wireless and radio communications. He is also the lead author of the textbook entitled *Pseudo Random Signal Processing: Theory and Application* (Wiley, 2005). His research interests include radio channel characterization and modeling, radio transmission technologies, radio access techniques, mobile multimedia communications, and perceptual quality assessment of mobile multimedia services.



Hajime Suzuki (S'94–M'99) received the B.E. and M.E. degrees from the University of Electro-Communications, Tokyo, Japan, in 1993 and 1995, respectively, and the Ph.D. degree from the University of Technology, Sydney, Australia, in 1999.

In 1999, he joined the Division of Telecommunications and Industrial Physics, Commonwealth Scientific and Industrial Research Organisation, Marsfield, Australia, and was transferred to the Information and Communication Technologies Center in 2003. Currently, he is the Project Leader of the Adaptive Wireless Project, investigating the use of multiple-input multiple-output for high-data-rate and high-reliability wireless communications. He is an active participant of the Australian Radio Communication Study Group 3 and of the Study Group 3 (Radio Propagation), International Telecommunication Union Radiocommunication Sector (ITU-R), particularly of the Working Party 3K (point-to-area propagation).