Radio-wave Propagation Prediction Model Tuning of Land Cover Effects

Magdalena Ballot, *Member, IEEE*

Abstract—**The effect of land cover is incorporated in the radio propagation prediction algorithm of Q-Rap. It is implemented by optimizing both the effective height of the land cover, hence affecting obstruction-loss calculations, and by the addition of terms to the basic transmission loss algorithm. A complete set of separate coefficients to these terms are determined for each land cover type. The optimization method improves the standard deviation of the error from 9.6 dB to 6.3 dB for measurements and predictions done at 390 MHz. This is an improvement of 3.3 dB over the original model that comprise of the free space loss equation with obstruction loss calculations for multiple knife edges. At this frequency the correlation coefficient between the measured and predicted values improved from 79.5% to 85.6%. At 2145 MHz the optimization method improves the standard deviation of the error from 16.2 dB to 8.6 dB and the correlation coefficient between the measurements and predicted values from 56.2% to 70.5%. The use of the correlation coefficient between the measured and predicted signal values in addition to the standard deviation of the error and mean error as criteria to be used when evaluating propagation prediction models is also proposed in this paper. A basis for best practices in the tuning of propagation prediction algorithms in radio planning tools using semi-empirical models is presented.**

Index Terms— **open source radio planning tool, Q-Rap, radiowave propagation, radio propagation model tuning, radio propagation terrain factors, radio systems planning, UHF propagation.**
.

I. INTRODUCTION

ACCURATE radio propagation prediction remains the cornerstone of radio system planning, particularly when new systems are to be installed. The effects of land cover, land use or clutter, as it is often called, on radio propagation have been widely discussed in the past [1], [2] and these effects are also considered in newer empirical and semi-empirical models [3] - [7]. To account for land cover effects and possible other reasons for differences between the signal strength predicted by a radio planning tool and the measured signal strength, most radio planning tools allow for the 'tuning' of the model to improve the accuracy of the propagation prediction. This is done even where the model can be considered deterministic. The motivation for the 'tuning' of propagation models is to improve the accuracy of propagation prediction within the context of requiring practical, cost and time effective radio propagation prediction where elaborate models are not considered worthwhile when weighting (potential) improved accuracy and the cost and time required against one another. This paper presents a number of methods to account for the land cover and improvement of the propagation prediction accuracy of the open source radio planning tool Q-Rap (www.QRap.org.za). The methods can be applied to a wide range of planning tools where the basic propagation prediction algorithm takes the form of an empirical radio propagation path loss formula combined with obstruction/diffraction losses. The methods for calculating the obstruction losses are the Deygout, the Epstein Petersen or Bullington methods for calculating the loss due to multiple knife edges. Examples of planning tools that implement such methods are *Asset* from *Aircom International* and *Atoll* from *Forsk* [8].

Two methods are employed to optimize the implementation of the effect of the land cover in the radio-wave propagation prediction algorithm. The first is a steepest descent method [9] used to determine the effective heights of the different land cover types to achieve the best possible correlation coefficient. The results of this study confirm that the inclusion of the height of the land cover have a significant impact on the prediction of shorter paths as is typically the case at higher frequencies, where the obstruction of the radio waves results primary from the land cover such as buildings or trees. In the data set used resultant land cover height is the order of 4 m in around 38% of the total data set and around 70% of the 2145 MHz set.

The second method is a simple, but elegant method of minimizing the mean square error (MSE) to determine the optimum coefficients of the additional terms added. The error here is considered the difference between the measurements and predictions. The results of a number of numerical experiments are presented in Section 0. These results provide a basis for suggestions of best practices in the tuning of prediction algorithms in Radio Planning Tools.

II. BACKGROUND

A. Prediction Algorithm

The propagation prediction algorithm of Q-Rap [10] has as its input a two dimensional path profile representing the heights at each point/pixel along the path between the transmitter and receiver, at regular intervals. It implements the Deygout method for multiple knife-edge diffraction [11] and accounts for rounded hills according to one of three methods, depending on the curvature of the profile at the point of the obstruction. For smaller curvatures, where these formulas apply, the formulas presented by Doughety and Maloney [12] are used. Where the path is best described by a spherical earth, the method recommended in [13] is used and for curvatures that do not fall in either of the applicable areas of the above two methods, the method described by Hacking [14] is used. The basic path loss is assumed to be free space loss. The algorithm also implements the effective earth model allowing the user to enter the effective k-factor to account for the change of the refractive index of the troposphere with height. More sophisticated methods for radio propagation prediction are described in literature [4] and [15] - [17] however these methods depend on detailed information of the terrain which includes building heights, street widths or very high resolution terrain data which can be expensive and time consuming to obtain. While these methods provide improved accuracy, the cost and effort are not always considered worthwhile and operators often opt for measurement based planning and optimization [18], [19]. The continued research into empirical prediction models, in for example [3], [5] - [7] also bears evidence to the restricted uptake of deterministic models requiring high resolution data.

Free access to all is one of the main drivers behind open-source developments such as the Q-Rap project; with that in mind the work in this paper was done using data that is freely available. The original NASA Shuttle Radar Topographic Mission (SRTM) digital elevation data (DEM) are maintained and made available by the CGIAR Consortium for Spatial Information [20]. The data is available for most of the world at 3 arc second, which is around 90 m. The land cover data used is the result of the National Land Cover 2000 project funded and implemented by the CSIR and the Agricultural Research Council of South Africa. The dataset that was released in 2005 has a 30 m resolution [21]. The predictions were done using a resolution of 30 m with linear interpolation used on the DEM.

B. Measurement Data

The measurements used in this paper were supplied by the Radio Technical Services of the South African Police Services (SAPS) and by Mobile Telephone Networks South Africa (MTN SA). All measurements comprised of the received signal strength and a GPS-localization of the measurement. All relevant information regarding the transmitting site was also available. The measurements received from the SAPS were done to confirm the coverage of the TETRA network in Gauteng between 2006 and 2008. The assumed transmitting frequency is 390 MHz. The measurement instrument was a MTH800 TETRA mobile from Motorola, with the quarter-wave monopole antenna mounted in the middle of the roof of a sedan car. A receive antenna height of 1.5 m was used. To simulate the effect of a handheld radio in a car, a 6 dB attenuator was inserted between the antenna and the mobile. The handover mechanism was turned off during the measurements to capture the coverage of each site separately. All the details of the site configurations were made available and were used in the signal strength predictions.

Three sets of measurements were supplied by MTN SA. All these measurements were originally done for the purposes of the tuning the operator's radio propagation prediction models. The first measurements were done in 2001 at 945 MHz in the GSM band. The second set was done in the UMTS band at 2145 MHz in 2005 and the last set was captured for model tuning in the 1800 MHz band in 2011, with most measurements done at 1817.4 MHz. The receive configuration was a quarter wave monopole mounted on top of a micro-bus or vehicle of similar height; a receive antenna height of 2 m was used. The transmit configurations varied but the details were captured when the predictions of the received signal strengths were made.

The number of transmitting sites used in this study was 117, of which 48 was in the 390 MHz set, 18 in the 945 MHz set, 24 in the 1800 MHz set and 24 in the 2145MHz set. Measurements for which the logged location was within the minimum of 25 times the wavelength and 25 m from one another was averaged before importing. This was done to attempt some degree of small scale averaging in particularly the 390 MHz set according to the recommendations in [22]. While, particularly in the 390 MHz measurement set, no comments on the certainty of how accurately the measurements capture the local mean can be made, the errors should balance one another out in the large set of data as there is no reason to believe that there is a systematic error. The data was sub-sampled to reduce the number of samples imported. Only samples differing with more than approximately 60 m in distance from one another was imported. As there was a disproportionate large number of measurements available in the 390 MHz band, samples in this set was taken at least 164 m apart.

To identify the noise floor present where the measurements were done, a histogram with respect to the received signal strength of each set of measurements was generated beforehand. Fig. 1 shows the histograms of the respective measurement sets. Measurements assumed to be within the noise floor or at the sensitivity level of the respective receivers were deleted as the measurement value does not represent the level of the signal at that point, but rather of the value of the noise received; these measurements are considered invalid. Measurements below -107 dBm were deleted from the 390 MHz set, in the 945 MHz and 1820 MHz sets values below -106 dBm were discarded, while in the 2145 MHz set only values above -115 dBm were considered.

Table Table I provides details on the composition of the measurements; numbers are provided as a percentage of the total of 261605 measurements. It was compiled after above mentioned invalid measurements were deleted. By far the

greatest majority of measurements were made where No Line of Sight (NLOS) exist.
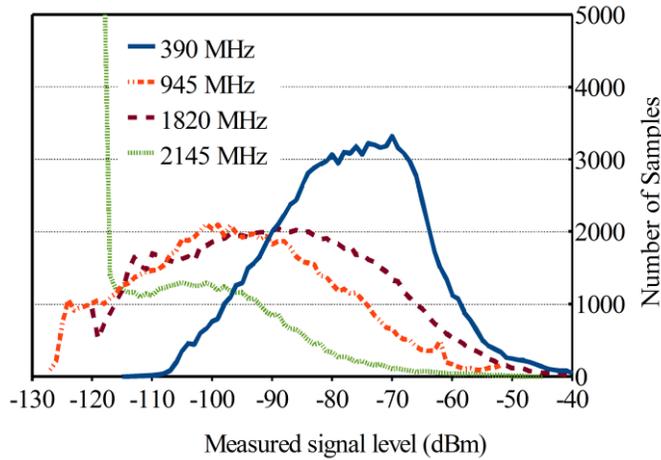

Fig. 1. Histogram of measurements indicating the noise floor and/or receiver sensitivity.

TABLE I
FREQUENCY COMPOSITION OF MEASUREMENTS USED. THE EXISTENCE OF LINE OF SIGHT (LOS) WAS DETERMINED AFTER HEIGHT OPTIMIZATION.

Set	LOS	NLOS	Total
390 MHz	1.43%	38.64%	40.07%
945 MHz	0.74%	18.27%	19.01%
1820 MHz	1.92%	23.43%	25.35%
2145 MHz	1.43%	14.15%	15.58%
Total:	5.52%	94.48%	100.00%

To provide an indication of the geographical distribution of the measurements, Fig. 2 was included. The boundaries of the Gauteng Province are indicated in the black lines. The province spans around 185 km by 185 km. The measurements for the higher frequencies was mainly in and around the two major cities, Johannesburg and Pretoria, while the 390 MHz covers the entire province. Since only the Gauteng province is shown not all of the 945 MHz samples are shown in Fig. 2.

Table II provides insight to the land cover types represented in the measurement set. The land cover data set used distinguishes between forty nine (49) land cover types. The *profiles* indicate the percentage of the samples that are classified to be primary influenced by a specific land cover type, while the *pixels* indicate what percentage of the pixels along all the transmission profiles belonged to a specific land cover type. The types that are well represented are shown in the table. *Formal suburbs* occur most often. *Natural grassland*

occur often in the measurement sets done at the lower frequencies, while *Commercial mercantile* is well represented in the measurements done in the higher frequencies.
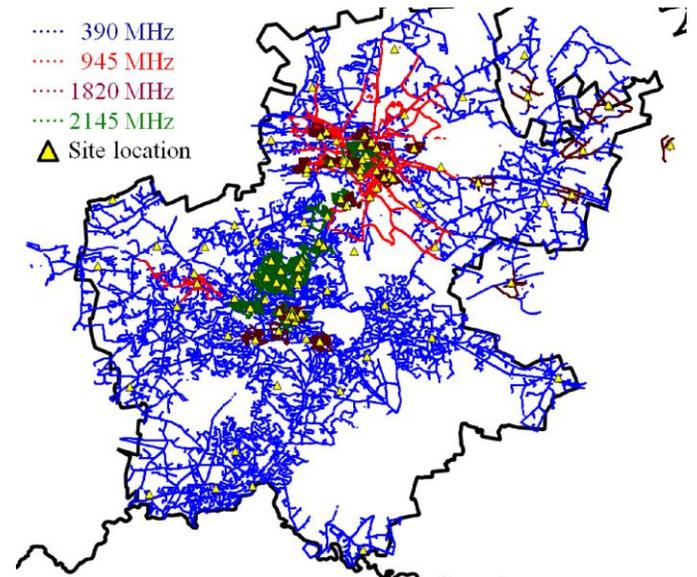

Fig. 2. Area map of the Gauteng Province indicating the spread of the measurements

Fig. 3 is a histogram with respect to distance of the different measurements sets. Most of the 2145MHz were within a 5km range while the majority of the 390MHz measurements were further than 5km from the site.
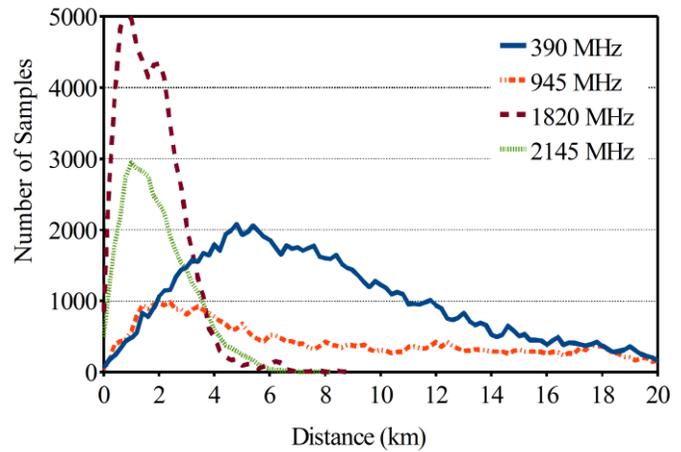

Fig. 3. Histogram of the number of measurements against the distance from the site

TABLE II
LAND COVER COMPOSITION OF MEASUREMENTS USED.

Land Cover Classification	All		390 MHz		900 MHz		1800 MHz		2145 MHz	
	Profiles	Pixels	Profiles	Pixels	Profiles	Pixels	Profiles	Pixels	Profiles	Pixels
Woodland	5.3%	6.8%	7.3%	8.2%	8.0%	6.2%	3.0%	3.7%	0.8%	0.6%
Thicket Bushland High Fynbos	9.0%	11.9%	10.9%	12.0%	15.7%	14.4%	4.3%	5.7%	3.8%	5.1%
Natural Grassland	20.5%	31.2%	33.4%	35.9%	22.0%	31.9%	8.6%	11.6%	4.8%	5.6%
Planted Grassland	1.6%	1.8%	1.0%	1.1%	0.9%	1.7%	2.0%	4.8%	3.1%	6.0%
Cultivated temp. Commercial Dry	6.2%	9.7%	10.9%	13.9%	3.4%	4.3%	4.2%	6.1%	0.5%	1.1%
Formal Suburbs	26.5%	13.5%	12.0%	8.5%	23.7%	13.2%	34.7%	26.5%	53.8%	49.5%
Urban Flatland	1.6%	0.5%	0.0%	0.0%	0.0%	0.0%	5.4%	4.8%	1.6%	1.6%
Commercial Mercantile	10.0%	4.0%	1.5%	1.3%	3.4%	2.4%	21.6%	19.6%	21.3%	19.4%

III. OPTIMIZATION METHODS EMPLOYED

A. Steepest Decent Method for Clutter Height Optimization

Because the base model of Q-Rap is a two-dimensional model, the degree of obstruction of the first Fresnel-zone, and hence the effective clutter height, depends on the density of the clutter type. While different frequencies have different Fresnel-zone radii for a given path length, on average, the same average/effective height exists. The main difference between different frequencies is that the variance of the degree of obstruction, and hence of the predicted signal strength, is higher at higher frequencies. Where deciduous trees are dominant in the land cover type, the season in which the measurements were taken has an effect on the effective land cover height optimized from the measurement set. Due of the changes in vegetation from area to area, the area also has an effect on the effective land cover type. In the South African context, the land cover height in younger suburbs is mostly determined by the height of the roofs, while in older suburbs the height of the trees determines the effective height of this land cover type. This section describes the method used to determine the effective land cover height that will result in the highest correlation coefficient between the measured and predicted signal levels.

The steepest descent method is well described in the literature [9]. The function to minimize for this problem is

$$f(\mathbf{h}) = 100 \cdot (1 - C(\mathbf{h})), \tag{1}$$

where $\mathbf{h}$ is the height vector indicating the heights of each land cover classification. The correlation coefficient between measurements and predictions of the received signal level is denoted $C(\mathbf{h})$ and is calculated using the following equation:

$$C(\mathbf{h}) = \frac{I \sum_{i=1}^{I} m_i p_i(\mathbf{h}) - \sum_{i=1}^{I} m_i \cdot \sum_{i=1}^{I} p_i(\mathbf{h})}{\sqrt{I \sum_{i=1}^{I} p_i^2(\mathbf{h}) - \left(\sum_{i=1}^{I} p_i(\mathbf{h})\right)^2} \cdot \sqrt{I \sum_{i=1}^{I} m_i^2 - \left(\sum_{i=1}^{I} m_i\right)^2}}, \tag{2}$$

with I the number of samples used, and m_i and p_i the measured value and the predicted value at each point. The predicted value is a function of the heights assigned to each clutter type, which makes the correlation coefficient a function of the height vector. The steepest descent method finds the minimum of any function. The function described in (1) represents the percentage of non-correlation. The minimum of $f(\mathbf{h})$ corresponds to the maximum of the correlation coefficient.

The gradient vector of $f(\mathbf{h})$ comprises of the first partial derivatives of $f(\mathbf{h})$ with respect to the height of each clutter type. The gradient vector is approximated as

$$\frac{\partial f(\mathbf{h})}{\partial h_i} \approx \frac{f(\mathbf{h})\big|_{h_i = H_{ni} + \Delta H} - f(\mathbf{h})\big|_{h_i = H_{ni}}}{\Delta H}, \tag{3}$$

with ΔH a small increase in the height, and H_{ni} the current height of the particular clutter type, i. Each next height estimation, n, is given by

$$\mathbf{h}_{n+1} = \mathbf{h}_n - \frac{\nabla f(\mathbf{h}_n)}{\|\nabla f(\mathbf{h}_n)\|} \lambda_n, \tag{4}$$

with the step size, λ_n, estimated from

$$\lambda_n = \frac{\lambda_a \cdot \partial^2 F(\lambda)/\partial^2 \lambda - \partial F(\lambda)/\partial \lambda}{\partial^2 F(\lambda)/\partial^2 \lambda}, \tag{5}$$

where $F(\lambda)$ is defined as

$$F(\lambda) = f(\mathbf{h} - \nabla f(\mathbf{h})/\|\nabla f(\mathbf{h})\| \cdot \lambda), \tag{6}$$

and its first derivate is approximated as

$$\frac{\partial F(\lambda)}{\partial \lambda} \approx \frac{F(\lambda_a) - F(0)}{\lambda_a}. \tag{7}$$

The second derivate of $F(\lambda)$ is estimated as

$$\frac{\partial^2 F(\lambda)}{\partial^2 \lambda} \approx \frac{(F(\lambda_b) - F(\lambda_a))/(\lambda_b - \lambda_a) - (F(\lambda_a) - F(0))/\lambda_a}{\lambda_a}. \tag{8}$$

The variables λ_a and λ_b are small step-sizes or height changes from the current evaluation point, λ_n, used to numerically estimate the first and second derivatives of $F(\lambda)$. If the second derivative is positive, it indicates that $F(\lambda_n)$ is a minimum, else it indicates that $F(\lambda_n)$ is a maximum in that interval and the next value of λ_n is rather selected from λ_a and λ_b depending on which one resulted in the lowest value of $F(\lambda)$. An attempt is made to select λ_a and λ_b small, but such that the minimum lies between them. In the implemented height optimization algorithm the next λ_a is the same as the previous λ_n. The choice of λ_b depends on whether $F(\lambda_a)$ is greater or smaller than that of $F(0)$. If $F(0) < F(\lambda_a)$, λ_b is the negative of λ_a otherwise it is twice the value of λ_a [9].

The aim of the optimization of the correlation coefficient is to identify the heights of the clutter which will result in an improved accuracy for any prediction used after the optimization process. While the above approach results in a fairly good improvement in the correlation coefficient and reasonable convergence towards an improved correlation coefficient, the heights of the clutter types that are not well represented in the data-set converge very slowly to the optimal height for that clutter type, as is evident from the results in Section V.A. If a specific clutter type has few occurrences in the prediction calculations it has a very low impact on the correlation coefficient, resulting in very small changes in the height of such a clutter type with each optimization step. To counter the small effect of ill represented clutter types on the correlation coefficient, the step size, λ_{ni}, for each clutter type, i, is augmented as

$$\lambda_{ni} = \lambda_n N_{tot}/N_i, \tag{9}$$

where N_{tot} is the total number of occurrences in all the predictions of the data set and N_i is the number of points along all profile paths where clutter type i was present. This means that rather than using one step size for all clutter types, each clutter type will have its own step size. We refer to this latter

method as the *augmented steepest descent* method in later discussions.

The steepest descent method was used until it could make no further improvement after which the implemented algorithm resorts to an exhaustive search where the height of each clutter type is changed to find any improvement and implement such changes. These height changes are swopped between an increase and decrease for each clutter type. If an improvement was made the new height is assumed other was the height of the clutter type remained unchanged. The magnitude of the experimental height changes is decreased as steps continue. Particularly, in the implemented height optimization algorithm an evaluation of

$$f_{ni} = f(\mathbf{h})\big|_{h_i = h_{ni} + \Delta H_n} \qquad (10)$$

is made for each clutter type. If f_{ni} is smaller than f_{min}, the previous minimum value found of $f(\mathbf{h})$ then

$$f_{\min} = f_{ni}. \qquad (11)$$

If no improvement can be made with the particular ΔH_n, then a new value is used

$$\Delta H_{n+1} = -\alpha \cdot \Delta H_n, \qquad (12)$$

with α any value *between* 0.5 and 1. The intention of the above step is both to change sign, i.e. search on the other side of the current value, and to reduce the granularity of the search. In the work done for this paper a value of 0.8 was chosen.

B. Minimizing the Mean Square Error in Determining the Coefficients of Terms

Yang and Shi [23] describe a rather simple but elegant way to optimize the empirical radio propagation prediction equation they used. They determine the coefficients of two additional terms to the COST231-Hata model by determining the minimum of the mean square error (MSE) between the original model and the measurements they took. It simply involves setting the partial derivatives of the (quadratic) MSE with respect to each coefficient equal to zero. This results in two linear equations that can be solved simultaneously. For the work done in this paper the principle is expanded by applying it to more than two terms.

In this paper an attempt is made to improve the propagation prediction model described in Section II.A. The final predicted value at point i is in the form

$$L_i = p_i + \sum_{n=0}^{N} k_n T_{ni}, \qquad (13)$$

with p_i the value of the original prediction model as described in Section II.A, k_n, the coefficients that are to be determined and T_{ni}, the value of the terms described in what follows at each point i. N is the number of terms added.

The choice of additional terms to use when doing model tuning can have a significant impact on the improvement that can be made by changing the coefficients of the terms. The first term is a simple offset

$$T_0 = 1. \qquad (14)$$

The multiplying constant, k_0, hence represents the offset. Most empirical models include the term

$$T_1 = \log_{10}(d). \qquad (15)$$

The distance, d, between the transmitter and receiver is in kilometers. The term

$$T_2 = \log_{10}(f) \qquad (16)$$

is also found in most empirical models. The operating frequency, f, is in MHz. The following terms occur in the Okumura-Hata model and all the models that built on it:

$$T_3 = \log(h_{tx}) \cdot \log_{10}(d) \qquad (17)$$

$$T_4 = \log(h_{tx}), \qquad (18)$$

with h_{tx} the height in meters of the transmitting base station.

In the particular data set used, a large portion of the clutter encountered are dominated by trees. The inclusion of some kind of term to represent the vegetation loss is considered wise. This term is here referenced to as the Vegetation term,

$$T_5 = f^x. \qquad (19)$$

Weissberger [1] suggested the exponent, x, to be 0.284, while reference [5] suggests a value of 0.5. Section V.B investigates the effect of the exponent. Note that no attempt is made to determine the distance that the signal travels through the vegetation as it is assumed that most of it will be in the last few meters where the clutter height will determine the distance and that the distance will be constant within a clutter type.

In a number of planning tools [24] the obstruction-diffraction component of the calculations can also be emphasized or de-emphasized. The effect of this on the MSE and correlation coefficient is investigated in Section 0. A seventh term is then added, namely

$$T_6 = L_{diff}, \qquad (20)$$

with L_{diff} the effective obstruction loss. When this term is included and treated separately, the loss, p_i in (13) is just the free space loss at point i.

Terms containing the mobile station height were not considered as it is assumed to be constant in the measurement set and inclusion of such a term will only be a duplication of the offset, T_o. In many empirical and semi-empirical models terms containing the height of clutter are used. These are mostly to account for the obstruction-diffraction effect. As is described in Section II.A, the prediction algorithm of Q-Rap already includes obstruction-diffraction components, hence such terms were also not included. Terms requiring street orientation or street width were also excluded as that information is not available in the land cover data-set used. Including all the additional terms, the MSE, denoted E, which is to be minimize, is

$$E = \sum_{i=1}^{I} \left(m_i - \left(p_i + \sum_{n=0}^{N} k_n T_{ni} \right) \right)^2 \bigg/ I, \qquad (21)$$

with m_i the measured signal strength value at point i, p_i is the predicted value as resulting from the core propagation prediction algorithm, T_{ni} is the value of term n at point i, k_n is the coefficient of the nth term, I is the number of measurements and N is the number of terms. To determine the optimum value of k_n, we determine the first partial derivative of the MSE with respect to each coefficient, k_m and equate that to zero:

$$\frac{\partial E}{\partial k_m} = -\sum_{i=1}^{I} \frac{2T_{mi}}{I}\left(m_i - p_i - \sum_{n=0}^{N} k_n T_{ni}\right) = 0. \quad (22)$$

This reduces to

$$\sum_{i=1}^{I}(m_i - p_i)T_{mi} = \sum_{n=0}^{N} k_n \sum_{i=1}^{I} T_{ni}T_{mi} \quad (23)$$

for each term. Note that the second partial derivative is

$$\frac{\partial^2 E}{\partial k_m^2} = \sum_{i=1}^{I} \frac{2T_{mi}^2}{I} \quad (24)$$

which will always be positive, meaning that the solution of equation (22) will always result in a minimum.

Solving the following matrix equation will provide the coefficients that will result in the lowest MSE

$$\mathbf{T} \cdot \mathbf{k} = \mathbf{R}. \quad (25)$$

The elements of $\mathbf{T}$, which is a NxN matrix, are

$$T_{nm} = \sum_{i=1}^{I} T_{ni}T_{mi} \quad (26)$$

and

$$R_m = \sum_{i=1}^{I}(m_i - p_i)T_{mi} \quad (27)$$

and $\mathbf{k}$ is the solution vector containing the coefficients for the minimum MSE. The software library, Eigen 3 [25] is used to solve the matrix.

IV. EVALUATION CRITERIA FOR RADIO-WAVE PROPAGATION PREDICTION MODELS

The criteria most often used for evaluating radio-wave propagation prediction models are in essence the mean difference between the measured and predicted signal level values,

$$\mu = \frac{1}{I}\sum_{i=1}^{I}(m_i - p_i), \quad (28)$$

and the standard deviation of that difference,

$$\sigma = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left((m_i - p_i) - \mu\right)^2}. \quad (29)$$

The root mean square error (RMSE) and mean square error (MSE) are often used, but does not really provide information additional to the mean error and standard deviation. Particularly, when tuning of models is done and the mean error often becomes zero, the MSE represents the variance which is the square of the standard deviation.

The correlation coefficient between the measured and predicted values provides a quantitative measure of the closeness of predicted value to measured value which is

separate to the standard deviation and the mean error. To illustrate this the results of three different optimization processes are compared below and presented in Table III.

TABLE III:
RESULTS ILLUSTRATING ADDITIONAL INFORMATION PROVIDED BY
CORRELATION COEFFICIENT

Description	μ	σ	C
Raw Model	22.9 dB	16.2 dB	56.2%
Empirical	-0.2 dB	10.7 dB	48.3%
Height Optimized	2.7 dB	23.0 dB	63.8%
Fully Optimized	-0.0 dB	8.6 dB	70.5%

μ is the mean value of the difference between the measurements and predictions in dB.

σ is the standard deviation of the difference between the measurements and predictions in dB.

C is the correlation coefficient of the predictions with the measurements.

The 2145 MHz set is used for this illustration. The *raw model* is the unaltered and pre-optimized propagation prediction model as is described in Section II.A, not using any information on the land cover. While the use of the *empirical model*, which does not make use of any terrain information, results in an insignificant mean difference, μ, between the measurements and the predictions and a lower standard deviation of that difference, σ, most radio planners would feel that the round circles that result from using an empirical model do not really provide much information. More information on the *empirical model* can be found in Section V.D .The *height optimized* model is the original model adding the optimized heights of the land cover at each point. The method of optimization used is described in Section III.A. When the effective height of each of the clutter types is altered to improve the correlation coefficient, C, the prediction follows the measurement more closely. However, it is often found that, when the degree of obstruction is predicted incorrectly the error that is made can be significant, resulting in a higher standard deviation of the error. This is particularly true at higher frequencies where predicting the signal level often amounts to predicting whether or not there exists line of site. Most radio planners would prefer the fully optimized model where both the standard deviation and the correlation coefficient are improved. In this model both the heights of the land cover types and the coefficients of the different terms were optimized. The latter was done using the method described in Section III.B

Over and above presenting the mean error and the standard deviation developers of deterministic models include graphs to show how well the predicted signal levels follow that of the measured values [16]. Sometimes coverage plots are used to convince users of the superiority of the models [26]. Using the correlation coefficient between the measurements and predictions, in addition to the standard deviation and mean of the difference between the measurements and predictions, provides a quantitative alternative for presenting the closeness of the predictions to the measurements.

V. NUMERICAL INVESTIGATIONS AND RESULTS

A. Comparison of Height Optimization Methods

In the interest of calculation speed, the 2145 MHz set was chosen to investigate the difference between the two approaches for optimizing the clutter heights presented in Section III.A. Fig. 4 presents the convergence of the two methods before an exhaustive search is started, where one can compare the resultant cost function value with the *minimum* that was obtained after numerous searches were run. It is clear that the augmented method converges quicker to the achieved minimum.
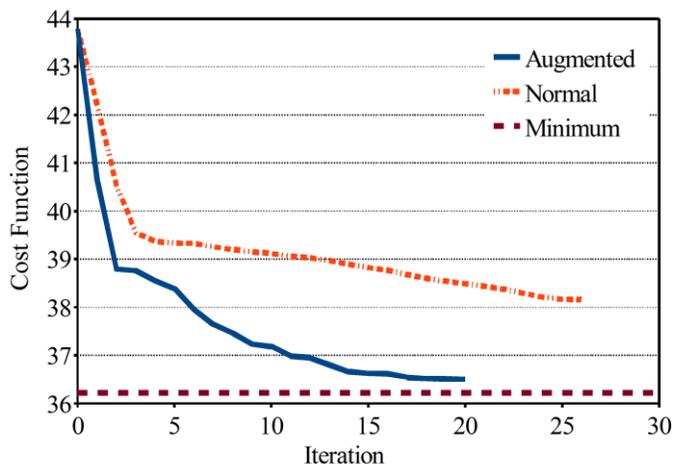
Fig. 4. Convergence of steepest decent methods for optimizing heights

When the final clutter heights are compared to the ones found in the minimum one finds that the augmented method determine the height of the clutter types that are well represented (more than 10%) closer to that of the minimum. It has differences with the heights of the minimum of less than 3 cm while the normal steepest descent method has errors just less than 1 m in these cases. When the rest of the heights are compared, the augmented method provide heights that are on average around 50% closer to that of the minimum. The normal method tends to underestimate all heights, while the augmented method overestimate the height of some clutter types.

B. Selecting the best exponent for the vegetation term.

For this investigation the entire set of data was used. The exponent, x, in (19) was altered and the standard deviation of the difference between the measured and predicted signal level values as well as the correlation coefficient of the predictions with the measurements were compared. Comparing the results when the exponent was altered from 0.284 to 0.5 revealed a very small change in both the correlation coefficient between the measurements and predictions and the standard deviation of the difference of the two. The difference in the standard deviation is less than 0.08% while the correlation coefficient of the two results differed with less than 0.05%. When altering the exponent from 0 (which is equivalent to omitting the term) to 10 resulted in a difference of 5.3% in the standard deviation and the correlation coefficient differed by less than 2.7%. In all cases the higher exponent resulted in better results. A larger

number and wider range of frequencies might prove more useful for investigating the best exponent. For the remainder of the investigation an exponent of 0.5 was used.

C. Investigating the resultant heights of the height optimization process

Table IV provides the heights of the better presented land cover types in the data set. Most of the heights are what one would expect, except the height of the *planted grasslands*. Investigation revealed that the areas classified as *planted grasslands* are sports grounds and particularly golf courses with rows of trees planted between the fairways. It is these trees that cause the additional obstruction of the radio wave.

TABLE IV
RESULTANT HEIGHTS OF THE HEIGHT OPTIMIZATION METHOD USING ALL THE MEASUREMENTS

#	Land Cover Classification	Height (m)
2	Woodland	1.6
3	Thicket Bushland High Fynbos	1.1
6	Natural Grassland	0.2
7	Planted Grassland	3.7
27	Cultivated temp. Commercial Dry	0.1
32	Formal Suburbs	3.9
33	Urban Flatland	4.6
43	Commercial Mercantile	3.9

Fig. 5 illustrates the effect that adding the additional land cover height have on the path profile of a radio link. If the height of the clutter is not added, the link analysis shows that there exists a line of sight between the transmitter and the receiver, while when one takes into consideration that closer to the transmitter (between 0 km and almost 0.8 km) the land cover is *formal suburban* with a height of almost 4 m it becomes apparent that no line of sight exists. The predicted radio propagation loss for the two scenarios differ with 16 dB.
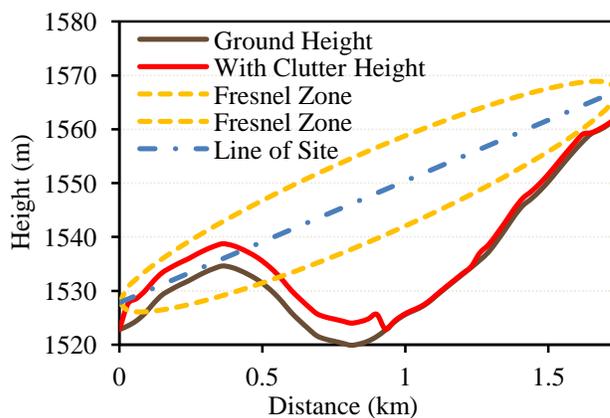

Fig. 5. Path profile illustrating the effect of the additional height of the Land Cover

D. Results of Implementing Different Tuning Methods

Table V lists some the optimization methods or combinations that were investigated to determine the effectiveness of different tuning methods. In most cases, the method was first implemented and tuned based on all

measurements and then applied to each set. The method was then also applied and optimized using the particular frequency set and then applied only to that set. Methods 1 to 3 are empirical methods not making use of any height data, not that of the clutter nor that of the terrain. Specifically, it means that in (21) to (27) the p_i values are set to zero. Method 4 is the unaltered base method implemented in Q-Rap before any tuning. In method 5 only the height is optimized as described in Section III.A. In methods 6 and 7 the effect of optimizing the entire set of coefficients, including L_{diff}, as described in Section III.B, was investigated. Methods 8 and 9 combine both of these methods. In methods 12 to 15 the effect of having only one set of coefficients per measurement set rather than having a set of coefficients for each clutter type was investigated. Methods 14 and 15 represent best the tuning freedom that is most often available in radio planning tools, which is one set of full coefficients for the entire set and an offset per clutter type [24, $4.3.3]. In method 16 the same height, which was optimized using all the measurements, was used for all frequency sets, but the coefficients for each clutter type was tuned per frequency set.

As a first verification, a simple Lee-type empirical model was tuned for the entire set of measurements resulting in the following empirical model for the path loss in decibel:

$$L = 37.4 + 24.6 \log \log(d) + 25.9 \log(f), \quad (30)$$

with the distance, d, in km and the frequency, f, in MHz. The resulting correlation coefficient for the entire set was 52.6% and the standard deviation was 12.6 dB. The empirical model of which the results are presented as method 1 in Table V is:

$$L = 150 + 58.8 \log(d) - 8.83 \log(f)$$
$$- 21.5 \log(h_{tx}) \log(d) - 0.654 f^{0.5} - 19.9 \log(h_{tx}), \quad (31)$$

with h_{tx} the height of the base-station transmit antenna in meter.

When comparing the results of the different optimization approaches in Table V, it is not surprising that the method which has the best results is where both the height and each term, including the obstruction-diffraction term, was optimized per land cover type and per frequency (method 9). Method 9 has the best correlation coefficient and the lowest standard deviation. A comparison of methods 5 and 7 reveals that optimizing all the different terms per land cover type only, provides better results than just optimizing the height for each land cover type. A comparison of methods 14 and 15 with methods 6 and 7 reveals that there is a significant benefit in tuning all the parameters per land cover type rather than just adding an offset for each. The inclusion of the freedom to change the emphasis on the obstruction-diffraction component

(T_6) seems to be vital when one compares methods 10 and 11 to methods 8 and 9. Using the same height for all frequency sets had a limited effect when methods 9 and 16 are compared, this confirms what was postulated in Section III.A, namely that the optimum height for a particular land cover type is independent from the frequency. Methods 12 and 13 represents the scenario where tuning is done ignoring any land cover information. Comparing these results with that of methods 8 and 9 provides one with an indication of the value of including the land cover information."

The basic model in Q-Rap provides good results for the lower frequencies with the longer paths as the terrain data represents the path profile well. The higher frequencies benefit most from including the land cover type as most of the obstruction in these sets with the shorter paths comes from the land cover itself.

Note that, in all cases, except for methods 4 and 5, the mean error is zero and, as we are dealing with tuning of propagation models, it is trivial to make the mean error zero by adding an offset equal to the negative of the mean error.

VI. CONCLUSIONS

The optimization methods described in this paper provide means for obtaining very good, for lower frequencies, to reasonably good, for higher frequency, radio propagation predictions. The standard deviation of the difference between the measured and predicted signal levels compares well with what was achieved in the ray-tracing results presented in [16].

A comparison of the standard deviation of the difference between the predicted and measured signal level and the correlation coefficient between the two, when using different prediction models, indicates that in order to evaluate a prediction model the mean difference, the standard deviation of the difference as well as the correlation between the measured and predicted signal levels should be used in the evaluation of prediction models.

Different models or methods of optimizing the prediction model to include the effect of the land cover was investigated. Where the base model is very similar to the one implemented in Q-Rap the best practice is to optimize the entire set of coefficients for each clutter type and include the optimized effective land cover height in the predictions.

TABLE V
RESULTS OF DIFFERENT APPROACHES IN OPTIMIZING THE INCLUSION OF THE EFFECTS OF THE LAND COVER

Ref Num	Method Description		Measurement set									
			All		390MHz		945MHz		1820MHz		2145MHz	
			σ	C	σ	C	σ	C	σ	C	σ	C
1	Empirical model	all	12.0	57.4%	9.9	59.6%	12.7	38.5%	11.6	50.4%	11.0	48.4%
2		set	12.0	57.4%	9.7	60.5%	12.0	35.5%	11.0	53.8%	10.7	48.3%
3	Empirical model with offsets per land cover type.	set	11.5	62.5%	9.4	63.6%	10.9	43.6%	9.9	65.5%	10.4	51.5%
4	Base Q-Rap model without alterations.		13.7	66.2%	9.6	79.5%	14.1	69.2%	15.0	52.6%	16.2	56.2%
5	**After** height optimization only.	set	16.2	71.0%	9.6	80.9%	13.7	73.3%	22.8	62.9%	23.0	63.8%
6	Full model per land cover type **before** height alterations allowing	all	9.0	78.2%	6.6	84.0%	8.7	72.7%	9.8	68.3%	9.9	61.9%
7	changes in obstruction-diffraction coefficient.	set	9.0	78.2%	6.4	85.0%	7.9	74.2%	9.0	72.2%	9.1	66.1%
8	Full model per land cover type **after** height alterations allowing	all	8.9	78.8%	6.6	84.3%	8.7	71.8%	9.6	70.8%	9.4	66.5%
9	changes in obstruction-diffraction coefficient.	set	8.9	78.8%	6.3	85.6%	7.6	75.7%	8.7	74.4%	8.6	70.5%
10	Full model per land cover type **after** height alterations **not** allowing	all	15.2	67.3%	9.8	78.0%	14.2	69.6%	17.5	57.1%	20.7	58.3%
11	changes in obstruction-diffraction coefficient.	set	15.2	67.3%	9.3	80.6%	12.8	71.3%	19.3	50.1%	21.7	58.2%
12	Global model **before** height change allowing changes in obstruction-	all	10.4	70.5%	7.2	81.7%	10.6	61.7%	11.1	57.7%	10.3	61.0%
13	diffraction coefficient.	set	10.4	70.5%	6.9	82.8%	10.0	61.3%	10.5	60.0%	9.6	63.9%
14	Global model **before** height alterations allowing alteration of	all	9.8	74.1%	7.1	82.2%	9.1	68.4%	10.3	64.1%	10.1	62.1%
15	obstruction-diffraction term with **offsets.**	set	9.8	74.1%	6.5	84.5%	8.7	68.5%	9.4	69.2%	9.3	63.9%
16	Full model per land cover type using the same height	set	8.9	78.8%	6.4	85.3%	7.9	74.2%	8.8	74.1%	8.7	70.2%

[*] The standard deviation of the difference between the measurements and predictions is indicated by σ in dB

[†] C is the correlation coefficient of the predictions with the measurements.

REFERENCES

[1] M. A. Weissberger, "An initial critical summary of models for predicting the attenuation of radio waves by trees," Electromagnetic Compatibility Analysis Center, Department of Defence USA, Annapolis, Maryland, USA, CAC-TR-81-101, 1982.

[2] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Trans. Veh. Technol.*, vol. VT-29, no. 3, pp. 317 - 325, Aug. 1980.

[3] J. Milanovic, S. Rimac-Drlje, and I. Majerski, "Comparison of Propagation Models Accuracy for WiMAX on 3.5 GHz," in *ICECS*, vol. 17, 2007, pp. 111-114.

[4] K. L. Chee, S. A. Torrico, and T. Kurner, "Radiowave Propagation Prediction in Vegetated Residential Environments," *IEEE Trans. Veh. Technol.*, vol. 62, no. 2, pp. 486-499, Feb. 2013.

[5] P. Phetsri, A. Sungkhapong, and S. Phaiboon, "Micro cell path loss prediction through hilly-forest terrain: A case study in south of Thailand," in *AIMSEC*, 2011, pp. 4670-4673.

[6] V. S. Abhayawardhana, I. J. Wassell, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *Proc. IEEE VTC 2005-Spring*, 2005, pp. 73-77.

[7] X. Cheng, J. Guan, and Chao K., "A WCDMA Propagation Model Based on COST-231-Walfish-Ikegami Pattern for Typical Residential Area," in *Proc. WiCOM 2012*, Beijing, 2012, pp. 1-3.

[8] H. Parsian, "Comparison of Asset and Atoll Cellular Planning tools for LTE Network Planning," Dept. Communication and Networking, Aalto Univ. Finland, Finland, Master Thesis 2012.

[9] J. A. Snyman, *Practical Mathematical Optimisation: an introduction to Basic optimization Theory and Classical and New Gradient-Based Algorithms.*: Springer Science+Business Media, 2005.

[10] M. Ballot, cPthlssp.cpp, [Online] Available: www.QRap.org.za and http://sourceforge.net/p/qrap/code/HEAD/tree/.

[11] J. Deygout, "Multiple Knife-Edge Diffraction of Microwaves," *IEEE Trans. Antennas Propag.*, vol. AP-14, no. 4, pp. 480-489, Jul. 1966.

[12] H. T. Dougherty and L. J. Maloney, "Application of Diffractions by convex Surfaces to Irregular Terrain Situations," *Radio Science Journal of Research*, vol. Vol. 68D, no. 2, pp. 239-250, Feb. 1964.

[13] ITU-R Radiocommunications Sector of International Telecommunications Union, "Propagation by Diffraction," Recommendation ITU-R P. 526-12, 2012.

[14] K. Hacking, "U.H.F. propagation over rounded hills," *Proc. IEE*, vol. 117, no. 3, pp. 499-511, Mar. 1970.

[15] Y. Lostanlen and Y. Corre, "Three-Dimensional Urban EM Wave Propagation Model for Radio Network Planning and Optimization Over Large Areas," *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3112-3123, Sep. 2009.

[16] S-C. Kim et al., "Radio Propagation Measurements and Prediction Using Three-Dimensional Ray Tracing in Urban Environments at 908MHz and 1.9MHz.," *IEEE Trans. Veh. Technol.*, vol. 48, no. 3, pp. 931-946, May 1999.

[17] R. Mathar, M. Reyer, and M. Schmeink, "A Cube oriented Ray Launching Algorithm for 3D Urban Field Strength Prediction," in *ICC*, 2007, pp. 5034-5039.

[18] M. Snyman, "The Challenges of Defining an Accurate Network Model to Provide Data," in *IIR Conference*, London, 1999.

[19] V. Wille, A. Kuurne, S. Burden, G. Dunn, and R. Barco, "Simulations and trial results for mobile measurement based frequency planning in GERAN networks," in *Proc. IEEE VTC 2002-Fall*, 2002.

[20] CGIAR Consortium for Spatial Information, "SRTM 90m Digital Elevation Data," 2008.

[21] CSIR and Argicultural Research Council South Africa, National Land Cover 2000 Dataset, 2005.

[22] J. D. Parsons, *The Mobile Radio Propagation Channel, 2nd Edition.*: Ed. John Wiley & Sons LTD, 2000.

[23] W. Yang and M. Shi, "A Linear Least Square Method of propagation Model Tuning for 3G Radio Network Planning," in *ICNC*, 2008, pp. 150-154.

[24] F. R. Gatarayiha, "Improved Radio Propagation Modeling for Cellular Networks in a Rolling Hill Environment.," Dept. Elec. Eng, TUT, Pretoria, South Africa, M.Tech Thesis February 2006.

[25] Eigen 3, http://eigen.tuxfamily.org.

[26] M. Coinchon, A-P. Salovaara, and J-F. Wagen, "The impact of radio propagation predictions on urban UMTS planning.," in *International Zurich Seminar on Broadband Communications, 2002. Access, Transmission, Networking.*, Zurich, 2002, pp. 32-1 - 32-6.

Magdalena Ballot received the B. Eng (electron) (cum laude) and M. Eng (computational electromagnetism) (cum laude) from the University of Pretoria, South Africa, in 1991 and 1995 respectively.

From 1995-1997 she worked as a Radio Planning Engineer at the Council for Industrial and Scientific Research. Between 1997 and 2002 she work at Mobile Telephone Networks (MTN) South Africa as a Radio Wave Propagation Prediction and Frequency Planning Specialist and later as Assistant and Acting Manager of Radio Network Optimization. Since 2003 she is employed as Senior Lecturer at the University of Pretoria in the Department of Electrical, Electronic and Computer Engineering.