
Peer reviewed version

Link to published version (if available):
10.1109/VTC.2001.956942

Link to publication record in Explore Bristol Research
PDF-document
Accurate High-Speed Urban Field Strength Predictions using a new Hybrid Statistical/Deterministic Modelling Technique

C. P. Michaelides and A. R. Nix

Center for Communications Research, University of Bristol,
Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, United Kingdom
E-mail: {Christos.Michaelides, Andy.Nix}@bristol.ac.uk

Abstract
To effectively dimension and rollout a cellular network, accurate site-specific field strength predictions are required. Traditionally, operators make use of high-speed statistical propagation models for each proposed base site. Today’s advanced cellular networks place a greater emphasis on urban microcellular deployment. It is well known that statistical models perform poorly in such environments. To overcome these limitations, the use of deterministic propagation modelling is common. However, although such models enhance prediction accuracy, they also require excessive computation time.

In this paper a new hybrid statistical/deterministic field strength prediction model is proposed that combines the speed of a statistical model with the site-specific accuracy of a deterministic ray model. Results indicate that the new hybrid model can produce very accurate field strength prediction grids with computing times reduced by one to two orders of magnitude. For the example considered in this paper, the computing time for a one square kilometer grid comprising 10,000 points (100 by 100) has been reduced from 20 mins to just over 4 minutes (running on an 800MHz Pentium III). For more complex ray models and/or higher grid densities, even higher speed-up factors can be achieved.

I. INTRODUCTION
A propagation model is a set of mathematical expressions and algorithms used to predict radio channel characteristics for a given environment. Prediction models can be empirical (statistical), theoretical (physical), or some combination of the two. Empirical models are based on measurements in the environments of interest, whereas theoretical models apply the fundamental principles of radio wave propagation to a given environment. Empirical modelling techniques have dominated the industry for the last 20 years and lie at the heart of today’s commercially available planning tools. These models have the advantage of unquestionable accuracy for the environments and locations in which measurements were taken. However, to form a propagation model the data must be generalized for different carrier frequencies, environments and antenna configurations. While useful empirical models exist for macrocellular environments, for microcells the required process of generalization is impractical (mainly due to the statistically small number of buildings involved).

Deterministic models combine physical modelling with site-specific environmental databases. Once appropriate building, terrain and foliage properties have been determined for the frequency of interest (normally through additional measurements), predictions can be generated for any environment and antenna configuration. Generality for deterministic models is not a problem and over the last 10 years a range of authors have demonstrated the accuracy of these techniques [1-8]. However, in order to achieve these high quality predictions, long run times are required (compared to empirical models). Commercially, the speed of deterministic models is often reduced by lowering the complexity of the algorithm (i.e. reducing the order of reflection/diffraction considered). Speed is now increased at the expense of prediction accuracy.

In this paper a hybrid propagation model is proposed that combines the processes of deterministic and statistical propagation modelling. This new approach enables high-speed field strength prediction grids to be generated in either microcellular or macro-cellular environments. While the partial use of statistical modelling implies that accuracy cannot be as high as a fully deterministic prediction, results indicate that they are dramatically superior to unaided statistical propagation models. More importantly, depending on the model settings and grid resolutions used, run times can be reduced by one to two orders of magnitude – making the model ideally suited for commercial radio planning applications [9-10].

Section II describes existing statistical and deterministic propagation modelling techniques. The newly proposed hybrid technique is explained in section III. The use of grid segmentation and local statistical parameter fitting is introduced. Analysis and comparison of prediction grids is used to further optimize key parameters in the hybrid model. In section IV the results of the hybrid model are compared with those of the full ray-tracing model and the COST 231 version of the Walfisch-Ikegami model. The paper concludes with a number of observations drawn from the work performed to date.

II. EXISTING MODELLING TECHNIQUES
A. Statistical Modelling Techniques
Statistical models have the key advantage that all environmental influences can be implicitly taken into account regardless of whether they can be recognized individually [11]. On the other hand, the accuracy of these models depends not only on the accuracy of the measurements, but also on the quality of the generalizations made and the similarities between the measurement environment and the environment under test.

In this paper one of the most widespread statistical propagation models, namely the COST 231 version of the Walfisch-Ikegami model (COST 231-WI) [11], is studied. This model has been extensively used in suburban and urban environments...
where building heights are assumed to be quasi-uniform. The model utilizes the theoretical Walfisch-Bertoni model [12] to obtain multiple screen forward diffraction losses for high base-station antenna heights. For low base-station antenna heights, the model uses parameters based on measurement data. The model considers free space path loss together with street losses based on diffraction and street orientation. However, in our implementation, site-specific street orientation data was not available and a fixed default value was applied. Steep transitions of path loss occur when the base station antenna height is close to the height of the surrounding building rooftops. Hence, the height accuracy of the base station antenna is especially significant if large prediction errors are to be avoided.

Figure 1 shows the predicted power grid for the test area of interest. The basestation (BS) was mounted on the top of the University’s Engineering building. The overall grid covers a service area of 1km by 1km with a resolution of 10m (i.e. 10,000 prediction points are calculated across the total service area). A carrier frequency of 1.8 GHz was assumed together with a fixed transmit power of 30 dBm. Vertically polarised dipole antennas were used at both the transmitter and receiver.

Due to the nature of statistical modelling techniques, the computation time for this model was relatively small when compared with deterministic methods (around 3 minutes on an 800MHz Pentium III using Matlab 5.3). A full discussion of the comparative run-times is given in section IV.

B. Deterministic Modelling Techniques

Deterministic models are based on well-established physical principles and can be applied to different environments without significantly affecting their accuracy [1-8]. In practice, the implementation of a deterministic model requires a sizeable database of environmental characteristics. For ray tracing models, detailed building, foliage, terrain and land usage databases are often required. The three-dimensional data shown in figure 1 was obtained via the photogrammetric analysis of stereo aerial photography [13]. Given the complexity of the algorithms and data involved, it is not uncommon for point predictions to take several seconds to compute.

For coverage plots requiring many thousands of point predictions, long run-times have prevented the widespread commercial application of deterministic models. In practice, these models are often simplified to achieve more acceptable run times; often with disastrous accuracy implications. In this paper our choice of deterministic model can be applied to both micro and macro-cellular environments. The model takes the form of a fully three-dimensional Radar Cross Section (RCS) scatter model with full support for multiple off-axis building/terrain scatter and diffraction [3]. However, the proposed hybrid solution can operate with any type of deterministic propagation model. Similar results could have been generated in combination with our previous and current microcellular models [1-2][8].

In general, work reported in the open literature indicates that carefully implemented deterministic propagation models offer exceptional accuracy [1-8]. For dense urban environments and locations with low mounted base sites, the accuracy of a well developed deterministic model is far higher than that of a well developed statistical model.

Computation times for deterministic models are generally exponentially related to the level of detail in the databases. Hence, the degree of detail influences both the computation time and the overall accuracy. Using the proposed hybrid model, only a small number of point-to-point predictions are required. This enables a complex deterministic model to be used together with a highly detailed database without seriously effecting the overall run time. In many deterministic models the level of database detail and/or the number of reflections and diffractions are sacrificed as a trade-off to improve run time (thus reducing the accuracy of the resulting predictions).
Figure 2 shows the resulting field strength prediction grid using the deterministic RCS model. As in Figure 1, the base site was assumed to lie on top of the University’s Engineering Building. The model was applied to an identical grid of 10,000 points. All other parameters remained constant, thus allowing a direct comparison between Figures 1 and 2. It can be seen that the results of Figure 1 dramatically under predict the received field strength for all locations in the service area. Given that previous work [3] has already demonstrated the accuracy of the RCS model, this implies that serious errors are generated using the statistical model. The predictions seen in Figure 2 are substantially superior to those of the unaided statistical model, however the computation time increased from around 3 minutes (Figure 1) to 20 minutes (Figure 2). We conclude that the RCS model is accurate but requires long computation times, whereas the statistical model is fast but offers poor prediction accuracy. Network operators require fast and accurate models for cell planning and site optimisation [9-10]. This conclusion provides clear motivation to develop a more accurate hybrid statistical/deterministic propagation model.

III. HYBRID TECHNIQUE

The hybrid model operates by locally tuning two received power versus distance statistical models, one for LoS and one for NLoS points. To improve accuracy, the coverage grid in automatically split into a number of local segments. Statistical models are then generated for each segment using data samples produced from the deterministic model.

A. LoS/NLoS Estimation

For each grid point, the algorithm must determine the appropriate segment number and whether LoS exists. This is performed using a topographic database containing terrain and building heights. The topographic database can be thought of as a two-dimensional array. In this array, each element corresponds to a certain point in the service area and its contents represent the building/terrain height above sea level. Using this database, the model reconstructs the ground profile information along the radial joining the Tx to the Rx. Since the radial may not pass through discrete data points, interpolation is used to determine the approximate heights involved.

B. Fitting of Equations

The hybrid technique locally fits a straight-line approximation for the LoS and NLoS received power versus distance using a small number of accurately generated point-to-point deterministic predictions. Given only a small number of points are required, the deterministic model should be configured to run in its most accurate mode (i.e. a complex database and taking into account a large number of reflections and diffractions). The model fits straight-line approximations in each grid segment using the following equations:

\[ P_{LOS} = C_{LOS} + n_{LOS} \log d \]  
\[ P_{NLOS} = C_{NLOS} + n_{NLOS} \log d \]

where \( P_{LOS} \) and \( P_{NLOS} \) represent the received power for LoS and NLoS points respectively, \( C_{LOS} \) and \( C_{NLOS} \) represent the intercept for the LoS and NLoS best fit equations respectively, and \( n_{LOS} \) and \( n_{NLOS} \) represent the gradient for the LoS and NLoS best fit equations respectively.

Equations 1 and 2 are now used to statistically predict the signal strength for each point in the service area. Although it is possible to fit equations 1 and 2 to the entire service area, this would ignore local factors such as building densities and heights, terrain variations, foliage densities, street widths and orientations. To incorporate this information, the coverage area is broken down into a number of local segments. Equations 1 and 2 are now optimised to fit a number of NLoS and LoS deterministic predictions in each segment. The deterministic model therefore considers local topographical information, which is then used to locally optimise LoS/NLoS power versus distance equations in each segment. The segmentation algorithm is described in more detail in the following section.

C. Segmentation of the total service area

The performance of the hybrid model is significantly improved by segmenting the total service area into a number of smaller regions. Smaller regions enable local factors to be tuned during the equation fitting process. Hence, best-fit equations for LoS and NLoS are produced separately for each segment in the service area. For simplicity, the desired coverage area is divided into a number of square segments. For comparison purposes, 1, 25 and 100 segments are considered here. Table 1 shows the standard deviation of the hybrid model (compared with the full deterministic model) for different segmentation sizes. In all cases, 800 point predictions (from a total of 10,000) were obtained using the deterministic model. The 800 prediction points were evenly distributed over the segments.

Analysis of Table 1 implies that 25 segments results in the lowest standard deviation compared with the full 10,000 point deterministic prediction. A single segment clearly does not allow sufficient local tuning of the hybrid model. While 100 segments enables a
high degree of local tuning, given that only 800 deterministic points were computed, this translates to just 8 predictions per segment (4 for LoS and 4 for NLoS). With such a small number of predictions per segment, the equation fitting process becomes unreliable.

### TABLE I
**Performance of Various Segmentation Schemes**

<table>
<thead>
<tr>
<th>Number of Segments</th>
<th>Size of Segment</th>
<th>Number of total ray-tracing points per segment</th>
<th>Standard Deviation (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1km by 1km</td>
<td>800</td>
<td>7.75</td>
</tr>
<tr>
<td>25</td>
<td>200m by 200m</td>
<td>32</td>
<td>6.55</td>
</tr>
<tr>
<td>100</td>
<td>100m by 100m</td>
<td>8</td>
<td>10.65</td>
</tr>
</tbody>
</table>

Assuming 25 segments over a 1km by 1km coverage area, each segment covers an area of 200m by 200m. If more deterministic prediction points were generated, a smaller segmentation size would result in a lower standard deviation.

### D. Optimization of Key Model Parameters

In this section the trade off between segment size and the number of deterministic prediction points is explored further. Figure 4 shows the standard deviation of the hybrid model as a function of segment number and deterministic prediction point number.

![Figure 4: Standard deviation versus segment number and maximum ray traced prediction points](image)

Not surprisingly, the lowest variance (3.4 dB) was observed when the largest number of ray-traced points was used together with the largest number of segments (i.e. 4,000 points or 40% of the total grid area spread over 100 segments). However, from Figure 4 it can be seen that excellent performance is also observed for 25 segments and 2,000 ray traced points (4.3dB). The speed of the hybrid model is obviously a function of the number of ray tracing predictions performed. When just 400 ray-traced points are used (4% of the grid area) the resulting predictions are relatively poor (10 dB for 25 segments). However for 800 points (8% of the grid area) and 25 segments the hybrid model achieves a respectable standard deviation of 6.3 dB. Figure 5 shows the resulting prediction grid for the hybrid model. Close inspection of Figure 5 reveals that much of the detailed structure (see figure 2) is still present.

![Figure 5: Power grid prediction (dBm) using the hybrid model at 1.8 GHz (800 ray-traced points, 25 segments)](image)

VI. COMPARATIVE RESULTS

For comparison purposes, results obtained using the hybrid model and the COST 231 Walfisch-Ikegami model are compared with those of the full ray-tracing model for an identical configuration. The hybrid model was configured using 800 points and 25 segments. Table II compares the run times and standard deviation of the models. For the case of the hybrid model, the processing time includes the time required by the deterministic model to calculate the required LoS and NLoS points.

<table>
<thead>
<tr>
<th>Model</th>
<th>Processing Time</th>
<th>Standard Deviation (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COST 231-WI</td>
<td>3 minutes 40 seconds</td>
<td>20.71</td>
</tr>
<tr>
<td>Hybrid Model</td>
<td>4 minutes 10 seconds</td>
<td>7.53</td>
</tr>
<tr>
<td>Ray-Tracing Model</td>
<td>20 minutes</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The statistical model offers a fast run time (which could be further improved via implementation in ‘C’), however the accuracy is extremely poor, and almost certainly unacceptable for planning purposes. The ray-tracing model produces a detailed output, however a run time of 20 minutes is required (increasing to well over an hour using the model’s most complex settings).
The hybrid model achieves a visually similar result to the deterministic model and is a vast improvement on the statistical approach. Run time for the hybrid model is now approximately 5 times less than that of the full deterministic model. Further speed enhancement could be obtained by coding the statistical portion in ‘C’.

Figure 6: Cumulative distribution function of Received Power in the Service Area for the three models

Figure 6 shows the cumulative distribution function of received power in the service area for the three models concerned. The full deterministic and hybrid models have near identical statistics, however the COST 231-WI model consistently and seriously under predicts the received power. The use of a small number of deterministic points can correct for errors in the statistical model, greatly improving the overall prediction quality.

V. CONCLUSIONS

In this paper a new hybrid field strength prediction technique has been reported for macro and microcellular environments. The model segments the total coverage area and then fits simple LoS and NLoS equations for power versus distance using a small number of deterministically generated points. For a dense urban microcellular environment, the COST 231-WI model is shown to offer a poor level of prediction, with a standard deviation relative to the deterministic model in excess of 20dB. For the hybrid model, good results were obtained when the ray tracer was used to predict approximately 4-8% of the grid points. These points were evenly distributed over 25 segments and further constrained to have an equal number of LoS and NLoS predictions. The resulting hybrid model prediction closely matched that of the full deterministic model with a computation time of approximately 4 minutes. With careful configuration, the hybrid model combines the accuracy of the deterministic approach with the speed advantages of a statistical model.

ACKNOWLEDGMENTS

The authors would like to acknowledge the help and support of Dr. Eustace Tameh. Christos Michaelides would also like to thank the institute of State Scholarships of Greece for providing the funding for the current project.

REFERENCES