Non-parametric and Parametric Modelling and Characterization of the Effective Earth Radius Factor for South Africa

A. M. Nyete and T. J. O. Afullo
Department of Electrical, Electronic and Computer Engineering
University of Kwa-Zulu Natal, South Africa

Abstract— The effective earth radius factor (k-factor) is an important parameter in the planning and design of both terrestrial microwave and UHF/VHF line-of-sight links. It is for this reason that the structure and variations in the effective earth radius factor in the first 200 m of the atmosphere is critical to radio link planners and optimization engineers alike. The atmospheric composition changes from time to time, place to place and even with height and hence the need for accurate determination and prediction of the k-factor. The four thirds (4/3) value given for the median k-factor in average temperate climate should only be used for gross planning but where data is available, the actual values of the same should be determined. This will ensure well designed links with minimum outage experienced due to k-factor related problems, i.e., diffraction (k-type) fading and thus avoiding expensive reverse engineering and optimization procedures. In this presentation, three years radiosonde measurements data sourced from the South African Weather Service (SAWS) has been processed and only k-factor statistics for the first 200 m above ground level considered for further analysis. Both non-parametric and parametric methods have been used to model solutions for the distribution characteristics of the k-factor across seven locations in South Africa. For the non-parametric approach, the kernel density estimate has been used. The so-called 'curve-fitting' method (Gaussian distribution modelling) has been used for the parametric technique. The Integral of Squared Error (ISE) has been applied to optimize the solutions in both cases. Using the foregoing procedures, both median ($k_{50\%}$) and k-factor values exceeded 99.9% of the time (k_e) have been determined. From the results, the kernel estimates are found to out-perform the curve-fitting method in terms of the ISE. Also, we observe that both techniques give very close values and so the error performance may be the only key performance indicator between both. The Rectangular kernel is observed to produce superior ISE performance in five out of the seven locations considered. It thus comes out as the more favourable kernel compared to the other three kernels used, i.e., Gaussian, Triangular and Epanechnikov kernel. It is also observed that with the optimum choice of the window width or bandwidth, h, there is little to choose from in terms of the kernel function, $K(k)$. Finally, we draw comparisons between the measured, curve-fitting and kernel values of the k-factor obtained.

1. INTRODUCTION

Clear air research advances in the recent past has been vigorous in South Africa. Palmer and Baker [1–4] have developed a cumulative distribution model for predicting the effective earth radius factor for South Africa. Afullo et al. [5–10] have also reported on the refractivity and k-factor statistics for Maun in Botswana and Durban in South Africa. For Botswana, they found the all year median value of the k-factor to be 1.1 and that of the effective value to be 0.61 for 0–200 m height range and 0.7 for the 0–500 m range. They also found median and effective k-factor values of 1.21 and 0.5 respectively for Durban, 0–500 m height range. Most recently, Fulgence [11] has also worked on refractivity and k-factor ranges for Central Africa; particularly in Rwanda and Tanzania. Also, in Nigeria, there has been a continued campaign for the determination of refractivity and refractivity gradient statistics mainly within the first 300 m of the atmosphere [12, 13].

2. ATMOSPHERIC RADIO REFRACTION

Electromagnetic waves propagating in the atmosphere do not travel in straight line but are bent (refracted) either towards or away from the earth's surface depending on the value refractive index. The refractive index, n, is defined as the ratio of the ratio of the velocity of the propagation of an electromagnetic wave in a vacuum to that of the velocity in a medium. It is given by [10]:

$$n = \frac{c}{v} = \sqrt{\mu\varepsilon} = 1 + N \times 10^{-6} \quad (1)$$

where c is the speed of a radio wave in a vacuum (free space), v is the speed of a radio wave in air, μ is the relative permeability, ε is the relative permittivity, N is the radio refractivity and n is the refractive index.

The radio refractivity changes with height above ground level. Its vertical gradient is particularly of great importance in the determination of the point k-factor statistics. The following equation is used to calculate the k-factor [2]:

$$k = \left[1 + \left(\frac{dN}{dh}\right)/50\pi\right]^{-1} \tag{2}$$

where k is the effective earth radius factor and $\frac{dN}{dh}$ is the vertical refractivity gradient.

3. PARAMETRIC AND NON-PARAMETRIC DATA TECHNIQUES

Parametric methods of data modelling are described by quantities commonly referred to as parameters which normally give a brief summary of the data structure. Of the many parameters used, the mean, median, standard deviation and variance always feature at the top. Such methods are therefore based on distributional assumptions of the data characteristics. The curve-fitting approach adopted in this paper, therefore falls in this category of data modelling.

On the other hand, non-parametric techniques make no assumptions as to the basic form of the variable under study. Of these, the two most common ones are the histograms and kernel density estimates. The kernel method performs better than the histogram and was actually introduced to counter the limitations associated with histograms. With kernel density estimation, a 'bump' is placed on each data point. The shape of this bump is determined by the kernel function, $K(k)$ and the overall data distribution is determined by the choice of the bandwidth, h. A very small value of h will result in estimates that are spiky in nature and difficult to interpret while a very big choice will result in over-smoothed estimates which will obscure fine details of the data structure, e.g., bi-modality.

4. MODELLING AND CHARACTERIZATION

Two different approaches were adopted for the task, i.e., the curve-fitting method and the kernel density estimation.

4.1. Curve-fitting Method

Previous work on the effective earth radius factor for Botswana by Afullo et al. [9] revealed that the distribution of the k-factor is bell-shaped, centred almost on a median value, u_k. From this observation, Afullo and Odedina [10] later developed an algorithm for modelling the probability density function (pdf) of the k-factor. They proposed the following pdf, $f(k)$ [10],

$$f(k) = Ae^{-\propto(k-\mu_k)^2} \tag{3}$$

They found that the relationship between A and $\propto$ is that of a normal (Gaussian) distribution. Using the algorithm in [10] and radiosonde data measurements obtained from South African Weather Service (SAWS), three-year Gaussian distribution models of the k-factor for seven locations in South Africa have been developed.

4.2. Kernel Density Estimation

Kernel density estimate of a variable k is given by [14, 15]:

$$f(k) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{k - X_i}{n}\right) \tag{4}$$

where h is the window width, bandwidth or smoothing parameter (depends on literature), n is the number of samples and X_i is the ith observation. Optimal kernel models are only possible when the value of the h chosen is such that the error performance (typically the Integral Squared Error, ISE) is at the minimum. The ISE is given by [6]:

$$\text{ISE} = \int_{-\infty}^{\infty} [f(k) - f^*(k)]^2 dk \tag{5}$$

Optimal bandwidth determination has been a subject of great research, but no single plug-in method will give the optimum value of h. Thus all the formulae for computing the optimum value of the window width will only give an estimate of the optimum h to start with. Several iterations may be the only sure way to get the global minimum error, thus the optimum value of h. Silverman's rule-of-thumb gives a good starting point for optimizing h. This rule is based on a Gaussian distribution assumption and is summarized by the following expression [14]:

$$h = (4/3n)^{0.2} * \sigma \tag{6}$$

where n is the number of samples and σ is a robust estimate of the sample standard deviation. The kernels used in this presentation and their efficiencies are shown in Table 1 below [6, 17].

Table 1: Kernels used and their efficiencies.

Kernel	Kernel function, $K(k)$	Efficiency (%)
Epanechnikov	$K(k) = \begin{cases} \frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}k^2\right), -\sqrt{5} \leq k \leq \sqrt{5} \\ 0, elsewhere \end{cases}$	100
Triangular	$K(k) = \begin{cases} \frac{15}{16}(1-k), -1 \leq k \leq 1 \\ 0, elsewhere \end{cases}$	98.6
Gaussian	$K(k) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{k^2}{2}\right)}, -\infty < k < \infty$	95.1
Rectangular	$K(k) = \begin{cases} \frac{1}{2}, -1 \leq k \leq 1 \\ 0, elsewhere \end{cases}$	93

5. RESULTS AND DISCUSSIONS

Results for the curve-fitting and kernel methods are presented. Figures 1–7 show the curve-fitting plots for the seven locations. Due to space considerations, only the kernel that produces the

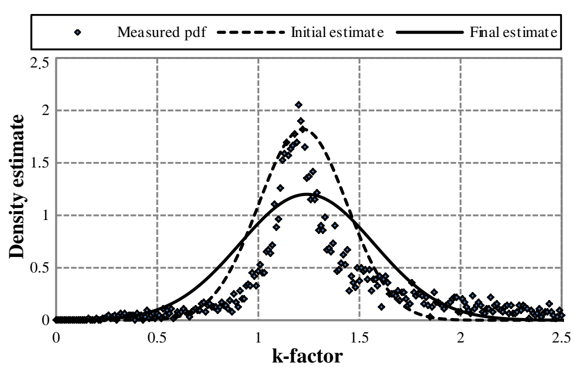

Figure 1: Bloemfontein Gaussian curve-fitting estimate, 200 m a.g.l.

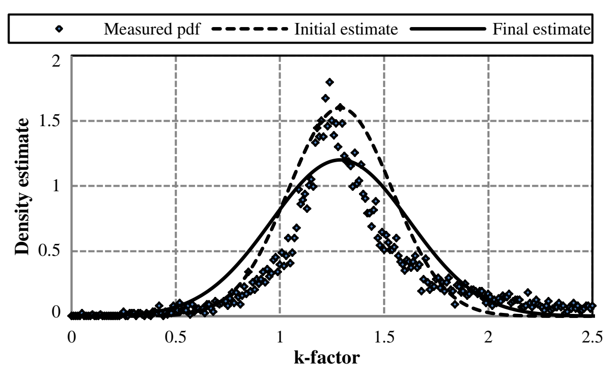

Figure 2: Cape Town Gaussian curve-fitting estimate, 200 m a.g.l.

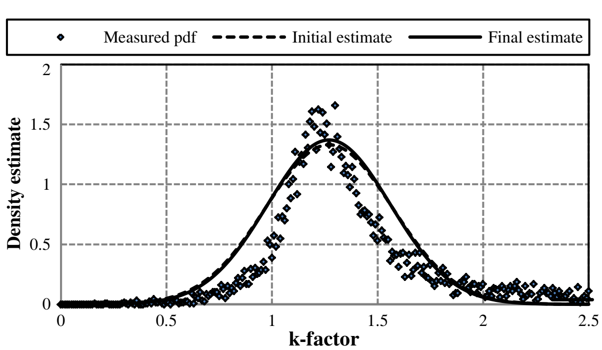

Figure 3: Durban Gaussian curve-fitting estimate, 200 m a.g.l.

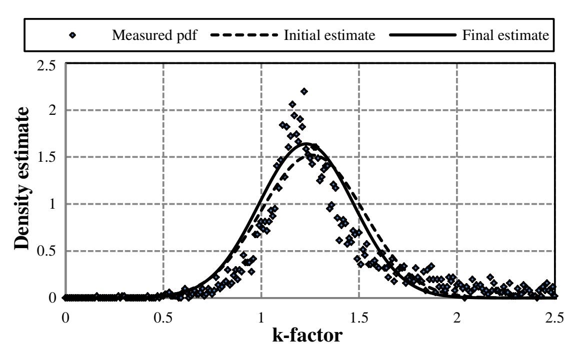

Figure 4: Polokwane Gaussian curve-fitting estimate, 200 m a.g.l.

best ISE performance for each location is plotted in this presentation. These plots are presented in Figures 8–14. Table 2 shows the Gaussian distribution models of the k-factor for the seven locations. From the results in Table 2, the curve-fitting median k-factor for Bloemfontein is found to be 1.24, while it is 1.29 and 1.27 for Cape Town and Durban respectively. It is found to be 1.23 for Polokwane, 1.19 for Pretoria, 1.17 for Upington and 1.20 for Bethlehem. A tabulation of the curve-fitting parameter A, median k-factor, μ_k and ISE values for the curve-fitting method are presented in Table 3.

From this table, Bethlehem is found to produce the worst error performance for the curve-fitting method while Durban is the best. The initial estimate is the first estimate obtained using values from the measured pdf. The final estimate represents the best estimate in that it is the one where minimum error is achieved. Kernel results of the median k-factor, h and ISE values are presented in Table 4. The kernel median k-factor is seen to vary between 1.21 to 1.22 for Bloemfontein, 1.26 to 1.27 for Cape Town, 1.25 to 1.26 for Durban, 1.25 to 1.26 for Polokwane, 1.20 to 1.21 for Pretoria,

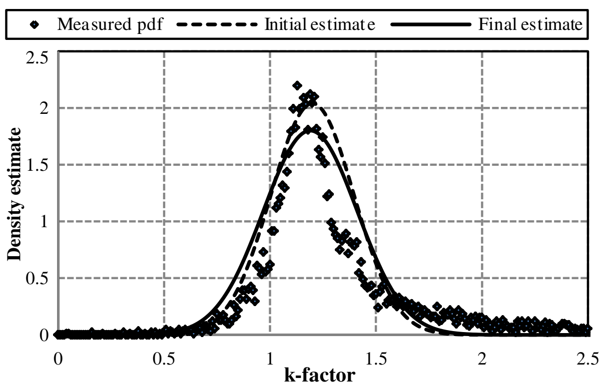

Figure 5: Pretoria Gaussian curve-fitting estimate, 200 m a.g.l.

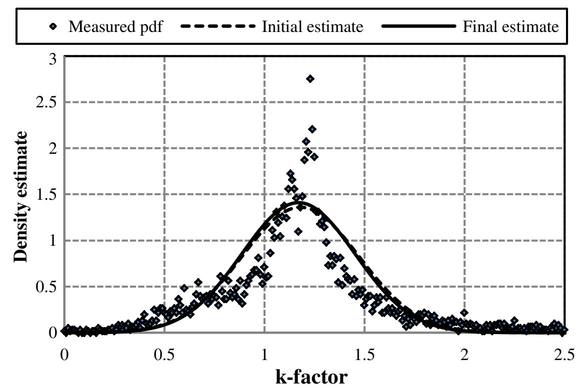

Figure 6: Upington Gaussian curve-fitting estimate, 200 m a.g.l.

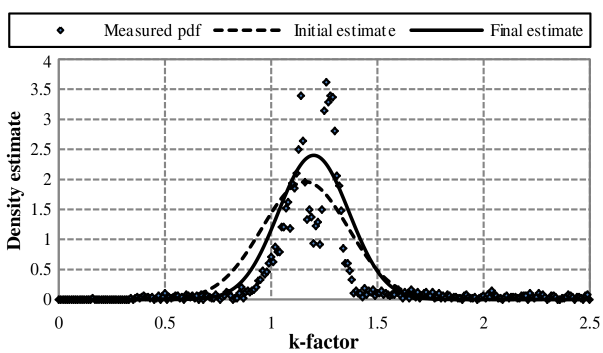

Figure 7: Bethlehem Gaussian curve-fitting estimate, 200 m a.g.l.

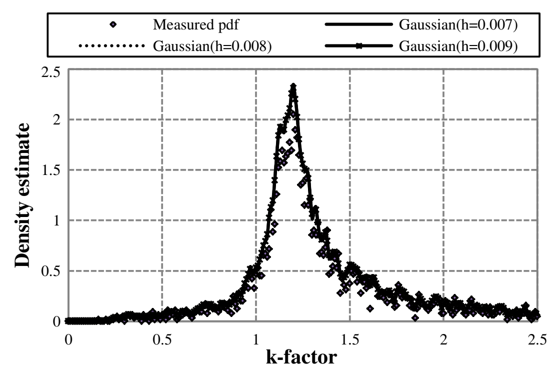

Figure 8: Bloemfontein Gaussian kernel density estimate, 200 m a.g.l.

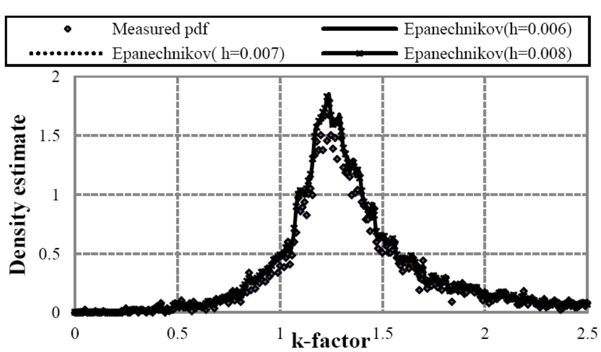

Figure 9: Cape Town Epanechnikov kernel density estimate, 200 m a.g.l.

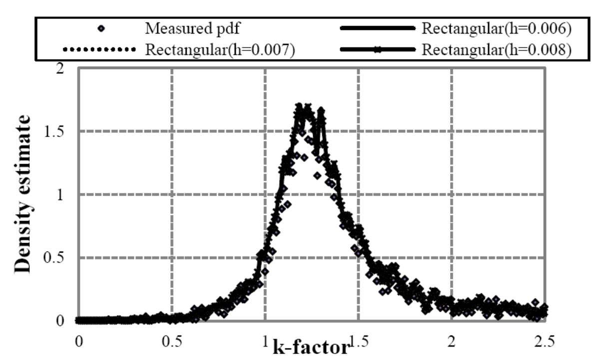

Figure 10: Durban rectangular kernel density estimate, 200 m a.g.l.

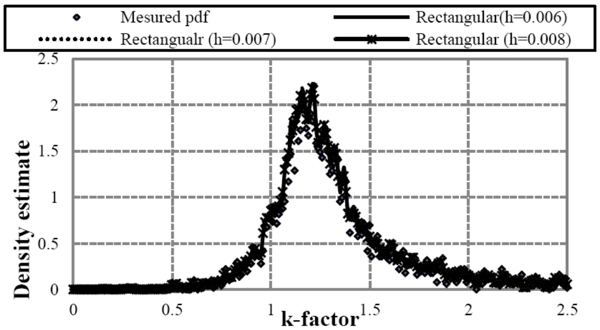

Figure 11: Polokwane rectangular kernel density estimate, 200 m a.g.l.

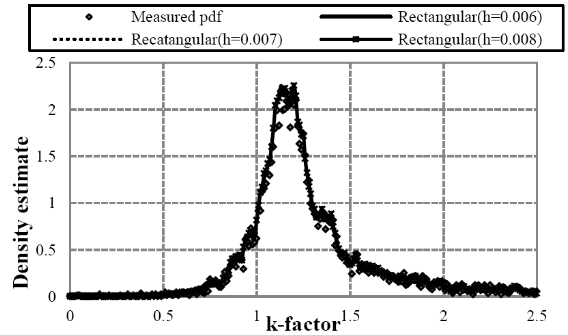

Figure 12: Pretoria-rectangular kernel density estimate, 200 m a.g.l.

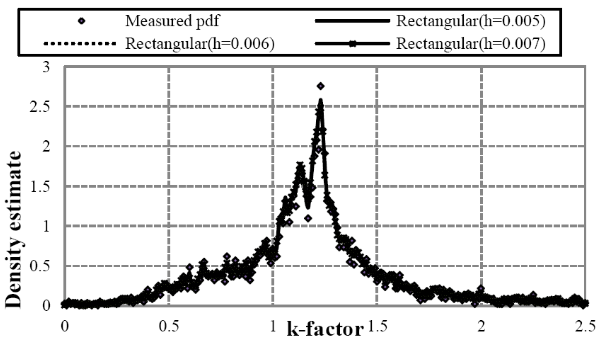

Figure 13: Upington rectangular kernel density estimate, 200 m a.g.l.

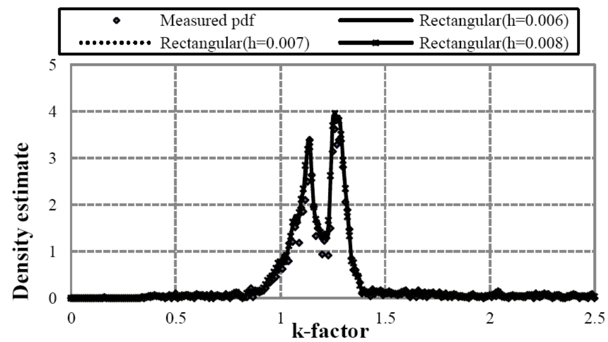

Figure 14: Bethlehem rectangular kernel density estimate, 200 m a.g.l.

Table 2: Three-year curve-fitting distribution models.

Location	Gaussian distribution Model
Bloemfontein	$1.2e^{-4.53(k-1.24)^2}$
Cape Town	$1.2e^{-4.53(k-1.29)^2}$
Durban	$1.37e^{-5.90(k-1.27)^2}$
Polokwane	$1.64e^{-8.45(k-1.23)^2}$
Pretoria	$1.8e^{-10.18(k-1.19)^2}$
Upington	$1.41e^{-6.25(k-1.17)^2}$
Bethlehem	$2.4e^{-18.10(k-1.20)^2}$

Table 3: Three-year curve-fitting parameters and ISE values.

Location	Initial estimates from measurements			Final estimates from curve fitting		
	u_k	A	ISE	u_k	A	ISE
Bloemfontein	1.22	1.82	0.23	1.24	1.2	0.19
Cape Town	1.29	1.60	0.13	1.29	1.2	0.11
Durban	1.27	1.33	0.102	1.27	1.37	0.101
Polokwane	1.26	1.52	0.11	1.23	1.64	0.09
Pretoria	1.20	2.04	0.14	1.19	1.8	0.13
Upington	1.18	1.36	0.14	1.17	1.41	0.13
Bethlehem	1.16	1.96	0.46	1.20	2.4	0.38

Table 4: Three-year kernel bandwidth, median k-factor and ISE values.

Locations	Bandwidth, h	Kernel							
		Gaussian		Triangular		Epanechnikov		Rectangular	
		Median	ISE	Median	ISE	Median	ISE	Median	ISE
Bloemfontein	0.007	1.21	0.0332	1.21	0.0334	1.21	0.0331	1.22	0.0334
	0.008	1.21	0.0330	1.21	0.0332	1.21	0.0332	1.22	0.0343
	0.009	1.21	0.0329	1.21	0.0333	1.21	0.0333	1.22	0.0349
Cape Town	0.006	1.27	0.0149	1.26	0.0147	1.26	0.0146	1.25	0.0147
	0.007	1.27	0.0147	1.26	0.0148	1.26	0.0148	1.25	0.0158
	0.008	1.27	0.0148	1.26	0.0150	1.26	0.0153	1.25	0.0160
Durban	0.006	1.26	0.0153	1.25	0.0153	1.26	0.0152	1.25	0.0149
	0.007	1.26	0.0151	1.25	0.0152	1.26	0.0153	1.25	0.0159
	0.008	1.26	0.0152	1.25	0.0154	1.26	0.0155	1.25	0.0163
Polokwane	0.006	1.26	0.0211	1.25	0.0205	1.25	0.0197	1.25	0.0193
	0.007	1.26	0.0203	1.25	0.0201	1.25	0.0198	1.25	0.0195
	0.008	1.26	0.0200	1.25	0.0200	1.25	0.0199	1.25	0.0208
Pretoria	0.006	1.20	0.0107	1.20	0.0107	1.21	0.0104	1.21	0.0099
	0.007	1.20	0.0105	1.20	0.0106	1.21	0.0105	1.21	0.0111
	0.008	1.20	0.0104	1.20	0.0107	1.21	0.0107	1.21	0.0112
Upington	0.005	1.20	0.0131	1.21	0.0129	1.20	0.0121	1.20	0.0113
	0.006	1.20	0.0126	1.21	0.0130	1.20	0.0127	1.20	0.0128
	0.007	1.20	0.0127	1.21	0.0130	1.20	0.0132	1.20	0.0149
Bethlehem	0.006	1.17	0.0567	1.18	0.0562	1.18	0.0555	1.17	0.0546
	0.007	1.17	0.0556	1.18	0.0560	1.18	0.0556	1.17	0.0576
	0.008	1.17	0.0552	1.18	0.0559	1.18	0.0563	1.17	0.0569

Table 5: Median ($k_{50\%}$) values compared, 200 m a.g.l.

Location	Measured	Curve-fitting	Kernel
Bloemfontein	1.22	1.24	1.21–1.22
Cape Town	1.29	1.29	1.26–1.27
Durban	1.27	1.27	1.25–1.26
Polokwane	1.26	1.23	1.25–1.26
Pretoria	1.2	1.19	1.20–1.21
Upington	1.18	1.17	1.19–1.20
Bethlehem	1.16	1.20	1.17–1.18

Table 6: k_{eff} ($k_{99.9\%}$) values compared, 200 m a.g.l.

Location	Measured	Curve-fitting	Kernel
Bloemfontein	0.51	0.53	0.52–0.53
Cape Town	0.49	0.51	0.50–0.51
Durban	0.53	0.55	0.55–0.56
Polokwane	0.63	0.64	0.63–0.64
Pretoria	0.66	0.68	0.67–0.68
Upington	0.49	0.50	0.52–0.53
Bethlehem	0.73	0.75	0.76–0.77

1.19 to 1.20 for Upington and finally 1.17 to 1.18 for Bethlehem. It is observed that the Rectangular kernel produces the best error performance in five of the seven locations, i.e., Durban, Polokwane, Pretoria, Upington and Bethlehem. The Gaussian kernel produces the best error performance for Bloemfontein while the Epanechnikov kernel is the best for Cape Town. It observed that, in the neighbourhood of the optimum h for each kernel, the values of the k-factor obtained are the same. The resulting plots are so close that discerning the difference between them is difficult. For each kernel, the optimum value of h is the one where the minimum error is obtained. Median k-factor results obtained from measurements, curve-fitting and kernel methods are tabulated in Table 5 for ease of comparison. Modelling k-factor values exceeded 99.9% of the time are compared against the measured ones in Table 6. From the curve-fitting models, the values are; 0.53 for Bloemfontein, 0.51 for Cape Town, 0.55 for Durban, 0.64 for Polokwane, 0.68 for Pretoria, 0.5 for Upington and 0.75 for Bethlehem. For the kernel models, the values vary between 0.52 to 0.53 for Bloemfontein, 0.5 to 0.51 for Cape Town, 0.55 to 0.56 for Durban, 0.63 to 0.64 for Polokwane, 0.67 to 0.68 for Pretoria, 0.52 to 0.53 for Upington and 0.76 to 0.77 for Bethlehem.

6. CONCLUSION

Both kernel and normal distribution models of the effective earth radius factor for South Africa have been determined. From these models, median and effective values of the k-factor have been obtained. These modelling results have also been compared with those obtained using measurements. It is observed that the curve-fitting, kernel and measured values obtained are quite close to each other. Also, the kernel models are found to follow the measured probability density function estimate as much as possible and as such their error performance is far much superior compared to the curve-fitting ones. The rectangular kernel is seen to produce progressively lower errors and is proposed as the best kernel for modelling the k-factor in South Africa. The results obtained in this presentation will go a long way in making sure that diffraction (k-type) fading is adequately addressed during link budgeting to counter any interference associated with k-factor variations for both UHF/VHF and terrestrial microwave links. In particular, both median and effective k-factor values obtained will be used by radio link planners to determine optimum antenna heights required to attain adequate path clearance in South Africa. The results will also serve as a benchmark for extension to cover the rest of the country by way of making predictions (interpolation) or more direct measurements to cover more areas for a longer period.

REFERENCES

1. Palmer, A. J. and D. C. Baker, "Predicting the long-term average of the effective earth radius factor for South Africa using ground based observations," *SAIEE Journal*, Vol. 97, No. 2, 182–185, June 2006.
2. Palmer, A. J. and D. C. Baker, "Predicting the monthly average cumulative distribution of the effective earth radius factor for South Africa," *SAIEE Journal*, Vol. 99, No. 2, 186–190, June 2006.
3. Palmer, A. J. and D. C. Baker, "A novel semi-empirical model of the effective earth radius factor," *IEEE Transactions on Broadcasting*, Vol. 52, No. 4, 557–565, December 2006.
4. Palmer, A. J. and D. C. Baker, "A proposed empirical model of the effective earth radius factor for telecommunications use in South Africa," *IEEE AFRICON*, 511–516, 2002.
5. Afullo, T. J., M. O. Adongo, and T. Motsoela, "Seasonal variation of radio refractivity in Botswana," *Proceedings of URSI Climpara'98 Conference on Climatic Parameters in Radio Wave Propagation Prediction*, 195–198, Ottawa, Canada, April 1998.
6. Afullo, T. J. and P. K. Odedina, "Effective earth radius factor measurement and modelling for radio link design in Botswana," *South Africa Institute of Electrical Engineers Research Journal*, Vol. 99, No. 3, 77–86, September 2008.
7. Afullo, T. J., T. Motsoela, and D. F. Molotsi, "Refractivity gradient and k-factor in Botswana," *Proceedings of the Third Regional Workshop on Radio Africa'99*, 107–110, Gaborone, Botswana, October 25–29, 1999.
8. Afullo, T. J. and P. K. Odedina, "On the k-factor distribution and diffraction fading for Southern Africa," *SAIEE Journal*, Vol. 97, No. 2, 172–181, 2006.
9. Afullo, T. J., T. Motsoela, and D. F. Molotsi, "Estimates of refractivity gradient and k-factor ranges for Botswana," *SAIEE Journal*, Vol. 92, 1–6, 2001.
10. Afullo, T. J. and P. K. Odedina, "Effective earth radius factor characterization for line-of-sight paths in Botswana," *IEEE AFRICON*, 227–232, 2004.
11. Fulgence, D., "Effective earth radius distribution for Central Africa: Rwanda and north-western Tanzania," MSc. Thesis, National University of Rwanda, August 2012.
12. Adediji, A. T. and M. O. Ajewole, "Vertical profile of the radio refractivity gradient in Akure, south-west Nigeria," *Progress In Electromagnetics Research C*, Vol. 4, 157–168, 2008.
13. Oyedum, O. D., "Radio refraction and atmospheric ray bending in Nigeria," *Nigerian Journal of Space Research*, No. 3, 99–110, November 2007.
14. Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Chapter 2–4, Chapman & Hall, 1990, ISBN 0-412-24620-1.
15. Sheather, S. J., "Density estimation," *Statistical Science*, Vol. 19, No. 4, 588–597, 2004.
16. Zucchini, W., *Applied Smoothing Techniques, Part 1: Kernel Density Estimation*, October 2003.
17. Ouyang, Z., "Univariate kernel density estimation," August 2005.