INTRODUCTION

An important requirement of tropospheric propagation models is their ability to predict the statistics of propagation parameters across a large probability range. Special choices of reference periods may also be required e.g. the “any” month or “worst” moth period defined by the ITU [1]. Another important requirement to the model, regards its ability to quantify the variability of the propagation, with respect to its long-term expectations, when observed during shorter observation periods. The nature of such variability of radiowave propagation associated with atmospheric effects is the subject of study of this paper. In this paper we consider the modeling of an important class of statistics of atmospheric propagation parameters, namely the statistics of the so-called monthly ‘time fraction of excess (TFE). The work in this paper is based on theoretical modeling as well as analysis of a very large amount of data on the monthly TFE with respect to two diverse propagation processes (slant-path rain attenuation effects and ducting effects on transhorizon paths). The data used are those selected by [2] from the set given by the European COST205 and COST210 projects, in [3] and [4]. The selected data comprises a total of 92 site-years of observations obtained in some 9 different climatic zones of Europe.

The paper starts by studying and analysing the general properties of the random process of the monthly time fraction of excess. The statistical characterisation of the monthly TFE is treated whereby various aspects of this random process are analysed (independence, cyclo-stationarity, seasonality etc.). In the second part of the paper the focus is on obtaining good (approximate) modeling of the statistical distribution of the monthly TFE. For this three alternative candidate models were studied as to their suitability.

Finally we discuss the application of the model to risk and reliability type of analysis of telecommunication systems and give an example of insurance type of analysis for a telecommunication system.

THE RANDOM PROCESS X OF THE MONTHLY TIME-FRACTION OF EXCESS

As noted in the introduction a simplified but useful way to quantify the performance of a radio link is to evaluate the behaviour of the fractional duration \( X_{m,j} \) in each month that a (propagation) parameter \( a \), relevant to performance degradation, exceeds a given threshold value \( a \):

\[
X_{m,j}(a) = \text{fraction of time that } a \geq a \text{ within month nr. } m \text{ of year nr. } j \\
(m = 1, 2, ..., 12; j = ..., 1981, 1982, \ldots) 
\]

Note: in this paper a random variable is indicated by bold-italic typeface.

\( X_{m,j} \), also known as the monthly time fraction of excess (TFE), usually exhibits a very complex behaviour and can be looked at as a random process \( X \). From a physical point of view we can consider this random process \( X \) as representing the family (ensemble) of time dependent outcomes of all possible (hypothetical) radio links with equal link parameters (e.g. frequency, polarisation, pathlength) located in a specific geographical zone. The area of such a geographical zone is assumed to be large enough to accommodate many radio links but small enough to ensure homogeneous climatic properties. There may be mathematical-statistical objections to this interpretation, but in any case such an interpretation is in accordance with the more
operational aspect of the usage of statistical data in empirical science. Here we assume that ergodicity applies, i.e. that if we observe a radio-link long enough and derive its statistical properties, then these properties are valid as well for 'similar' links in the 'same' climatic zone.

As in the case with many meteorological effects, many atmospheric propagation effects exhibit seasonal dependence. Even longer cyclical periods are in principle possible. In [2] the basic nature of the process of $X$ has been investigated, using the propagation data with respect to slant-path attenuation and ducting mentioned in the introduction. The analysis using these measured data leads to the conclusion of a random process that is cyclo-stationary and exhibits independence. The proof of these two aspects is, admittedly, more circumstantial than rigorous due to the limitations of the data. The longest contiguous observation time of the available data was 9 years so those cycles longer than 9 years cannot be detected in this study. The independence has been conjectured from the, in general, observed lack of correlation between the samples; for TFEs associated with lower thresholds values some correlation between subsequent months have been observed. Nevertheless these two properties can be taken as simplifying approximations.

THE PARENT DISTRIBUTION $P_X$ OF $X$

The assumptions of cyclo-stationarity and independence means that the statistics of $X$ is fully determined by 12 statistical distributions (one for each month of the year):

$$P_{X|m}(X) = \text{probability}(X \geq X \mid \text{in month nr. } m) \quad (2)$$

To obtain a working model we must apply some further simplification. This is because the available propagation data are usually not enough for characterising the 12 distributions $P_{X|m}$. E.g. the data mentioned in the introduction is at most only sufficient for quantifying the seasonal dependence of the average of $X$ for some parts of Europe. Furthermore, seasonal dependence may be very different in different parts of the world, so that a generally applicable statistical model cannot be expected.

To circumvent this problem we shall use a simplified model that is based on the average, or aggregate, statistical distribution:

$$P_X(X) = \frac{1}{12} \sum_{m=1}^{12} P_{X|m}(X) = \text{probability}(X \geq X) \quad (3)$$

If necessary seasonality effects can be approximated using the dual population model first introduced by [5]. In this model $M$ months are considered to produce $X > 0$. The other $12 - M$ month only produce $X = 0$.

Three models for $P_X$ are given in literature. The first model is the Conditional Exponential (CEX) model [6]:

$$P_X(X) = C_0 e^{-X/C_1} \quad (4)$$

for $X > 0$; the model has been extended by [2] to allow for $C_0 > 1$. The second model is the Conditional Lognormal (CLN) [7]:

$$P_X(X) = C_0 \int_{\frac{\ln(x)}{C_2}}^{\infty} \frac{dt}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2}t^2}}{C_1} \quad (5)$$

for $X > 0$. The third model is the Shifted Gamma (SGM) model [2], [8]:

$$P_X(X) = \int_{\frac{X}{C_1}}^{\infty} \frac{dt}{\Gamma(C_2)} \frac{C_2^{-1} e^{-t}}{t^{C_2}} \quad (6)$$
for $X > 0$. $\Gamma$ is the well-known gamma function. $\kappa$ is the shift parameter with:

$$\kappa = C_2(1 - \sqrt{C_2})H(1 - C_2)$$  \hspace{1cm} (7)

where $H$ is the unit-step function: $H(X) = 1$ for $X \geq 0$ and $H(X) = 0$ for $X < 0$. $H$ is also known as the Heavyside function.

$C_0$ is also called the conditionality or location parameter, $C_1$ the scale parameter and $C_2$ the shape parameter; they must all be positive in value. Note that an effective conditionality parameter for the SGM model can also be defined e.i. $C_0 = \lim \varepsilon \to 0 P_X(|\varepsilon|)$.

In [2] various classical statistical tests (Chi-square, regression) were carried out using the aforementioned data to pick out the best model. The results were inconclusive. The problem stems from the fact the individual sample cumulative distribution function (s.c.d.f.) is too small (usually much less than 50 sample points with $X > 0$) to allow severe testing. To circumvent this problem a normalisation method was introduced by [2], [9] which allows of pooling of the whole data set (leading to some 3500 sample points with $X > 0$). The normalisation consist of, for each s.c.d.f., transforming $X$ to the variate $V$

$$V = 1 - P_X(X, C) / C_0$$  \hspace{1cm} (8)

$C = (C_0, C_1, C_2)$ are the model parameter values estimated from the s.c.d.f.. If $C$ is perfectly known then $V$ is uniformly distributed (in the range 0 – 1); provided of course that the model is correct. The grand pool of data from all the s.c.d.f., after transformation and pooling, will then also be uniformly distributed. This was the assumption of the testing performed by [9]. However, because the parameter $C$ cannot be perfectly determined $V$ will not be a perfect uniform random variate. In [2] this problem is dealt by determining the expected distribution through application of Monte Carlo simulation (which then includes the effects of imperfect estimation of $C$). By comparing the grand-pool distribution of $V$ obtained from the simulations with that obtained from measurement we can then determine the suitability of each model. The table below summarises the in this way obtained performance of each model in terms of the mean and the root mean square (RMS) of natural logarithm of the ratio of observed to model (simulated) probability distribution (restricted to non-zero values of $V$):

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean (%)</th>
<th>RMS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEX</td>
<td>-0.6</td>
<td>7.8</td>
</tr>
<tr>
<td>CLN</td>
<td>-12.5</td>
<td>27.4</td>
</tr>
<tr>
<td>SGM</td>
<td>3.4</td>
<td>5.5</td>
</tr>
</tbody>
</table>

This table shows clearly that the SGM model has the best performance with the CEX model a close second. The performance of the CLN model is not so good.

**APPLICATION OF THE MODEL**

In [2] it is shown that the simplified model discussed above, for the random process $X$, can used successfully to calculate/analyse many parameters with regard to variability (e.g. ITU worst-month quotient $Q_1$, or the calculation of return periods associated with the variability of worst-month resp. annual distributions). Here we show a novel usage of this model in the application of risk assessment. We consider the situation of a provider who wants to set up a (radio) communication system with many users and with each user having a separate link. The links have similar propagation properties but are statistically independent. The special service being provided is that for any monthly period in which the monthly TFE $X$ exceeds the promised long-term worst-month $<W>$ value the victim is given a financial compensation:

$$\begin{cases}
& z \times (X - <W>) / <W> \quad \text{when } X < <W> \\
& 0 \quad \text{when } X \leq <W>
\end{cases}$$  \hspace{1cm} (9)

Note: the worst-month $W$ in a given year is the maximum of $X$ in that year. $<W>$ is the average of all $W$ over all years. Of course many alternative compensation schemes may be devised. However, the above scheme seems to be a reasonable proposal for a compensation scheme: it produces compensation in proportion with the excess outage deviation from the
promised value ($<W>$); obviously a necessary requirement for an honest compensation scheme. The normalisation to $<W>$ seems also to be reasonable, it furthermore simplifies our analysis (by eliminating the scale parameter $C_1$). Of course in the end the service provider might want to make the choice of the value of $z$ dependent on many other factors (e.g. so as to avoid excessive loss in say unfavourable propagation regions). The expected payment per user per month is:

$$\text{cost} = z \int_{<W>} \frac{dX}{f_X(X - <W>/ <W>)}$$

(10)

where $f_X$ is the probability density function of $X$. Using the CEX model we can readily obtain an analytical expression for the expected cost per user per month. The results are given by Figure 1 that shows the dependence of $\text{cost}$ on $C_0$ (for several values of the seasonality index $M$). From this figure two interesting features can be seen. In the first place we can see that the largest cost value is obtained for the smallest value of $C_0$, here we have a cost of $z/12$ per user per month (or simply $z/12$ per user per year). In the second place we have a large range of $C_0$ (0.5 - 2) where the $\text{cost}$ is practically constant $= 0.0167 z$ per user per month (or equivalently $0.2 z$ per user per year). For $C_0$ below 0.5 the cost rapidly increases with decreasing value of $C_0$. This rapid increase of cost is not very obvious at the onset of this exercise. This demonstrates the necessity of properly analysing the problem at hand with the aid of good models.

![Figure 1 Expected cost per user versus $C_0$.](image)

ACKNOWLEDGEMENTS

I would like to thank all the people who worked for the European cooperative projects COST205 and COST210. Without their efforts the valuable and extensive propagation database used here would not be available. I would also like to thank prof.dr. Gert Brussaard for his encouragement and support.

REFERENCES