

Short-term prediction of rain attenuation level and volatility in Earth-to-Satellite links at EHF band

L. de Montera¹, C. Mallet¹, L. Barthès¹ and P. Golé¹

[1]{Centre d'étude des Environnements Terrestre et Planétaires, Vélizy-Villacoublay, France}

Correspondence to: L. de Montera (louis.demontera@cetp.ipsl.fr)

Abstract

This paper shows how nonlinear models originally developed in the finance field can be used to predict rain attenuation level and volatility in Earth-to-Satellite links operating at the Extremely High Frequencies band (EHF, 20-50 GHz). A common approach to solving this problem is to consider that the prediction error corresponds only to scintillations, whose variance is assumed to be constant. Nevertheless, this assumption does not seem to be realistic because of the heteroscedasticity of error time series: the variance of the prediction error is found to be time-varying and has to be modeled. Since rain attenuation time series behave similarly to certain stocks or foreign exchange rates, a switching ARIMA/GARCH model was implemented. The originality of this model is that not only the attenuation level, but also the error conditional distribution are predicted. It allows an accurate upper-bound of the future attenuation to be estimated in real time that minimizes the cost of Fade Mitigation Techniques (FMT) and therefore enables the communication system to reach a high percentage of availability. The performance of the switching ARIMA/GARCH model was estimated using a measurement database of the Olympus satellite 20/30 GHz beacons and this model is shown to outperform significantly other existing models.

The model also includes frequency scaling from the downlink frequency to the uplink frequency. The attenuation effects (gases, clouds and rain) are first separated with a neural network and then scaled using specific scaling factors. As to the resulting uplink prediction error, the error contribution of the frequency scaling step is shown to be larger than that of the downlink prediction, indicating that further study should focus on improving the accuracy of the scaling factor.

1 Introduction

An ever growing demand for increased capacity of communication channels has led the SATCOM industry to develop new satellite systems operating at frequencies above 20 GHz (Extremely High Frequencies band – EHF), where large bandwidths are available. In this context, a French military satellite, Syracuse 3, was launched in 2006 with 20 GHz downlink and 44 GHz uplink channels (Marsault et al., 2006). At these frequencies, and especially at 44 GHz, the attenuation can reach significant levels during rain events and it is no longer cost-effective to use a fixed power margin. Since rain events are sparse phenomena in space and time, some specific adaptive fade countermeasures, known as Fade Mitigation Techniques (FMT), were developed. The EHF propagation channel variations therefore have to be modelled, in order to trigger fade mitigation only when needed. The study originally concerns the use of the 44 GHz channel, but because the Syracuse 3 measurements are not available yet, the propagation channel model presented in this paper was developed with data collected during the previous Olympus 20/30 GHz propagation experiment (Sect. 2).

Whatever the FMT used, the prediction of attenuation a few seconds ahead (typically 10s) is needed due to the control loop reaction time. Nevertheless, direct models that estimate the attenuation from ground based meteorological data are not accurate enough since attenuation depends on the rain rate and also on the Drop Size Distribution (DSD) along the path that are not known. Such models can only be used to provide an estimation of yearly attenuation statistics (Dissanayake et al., 1997). Recently, weather radars (Núñez et al., 2006) and complex numerical weather models (Hodges & Watson, 2006) have been used to improve the accuracy of the attenuation retrieval and forecasting but the results are still approximate and these models are not appropriate for short-term prediction which is required in fixed satellite links.

This problem, as well as the need for synthetic attenuation time series for simulation purposes, had led to the development of statistical models that describe the attenuation process only through its past values. Since the uplink is more affected by attenuation than the downlink (because it generally operates at a higher frequency), the downlink attenuation is predicted and then scaled to the uplink frequency in order to estimate the future propagation conditions of the uplink. Various prediction models have been proposed, such as a linear regression (Dossi, 1990), a first order stochastic equation (Manning, 1990, 1991), Markov chains (Castanet et al., 2003; Fiebig, 2002), an adaptive linear filter (Grémont et al. 1999), a neural networks

(Chambers & Otung, 2005) or a model based on the fade slope (Van de Kamp, 2002). A comparison was performed with Olympus 20 GHz data (Sect. 5) and, in spite of the complexity involved in these models, none of them was found to provide a prediction method significantly better than persistence (i.e. considering that the process remains constant over the forecast interval). Indeed, the analysis of the attenuation time series shows that during rain events it behaves approximately like a random walk and that the autocorrelation of its increments is small over 10-second time lags (See fourth row in Fig.2, Sect.3). Besides, rain has been described as a Self Organized Critically (SOC) relaxation phenomenon and can be compared to avalanches or earthquakes (Peters et al., 2002), which explains the difficulty in predicting it.

Nevertheless, differentiating the attenuation time series shows volatility clustering and leads to a distribution with a large Kurtosis, and therefore behaves similarly to certain stocks and currency exchange rates. Thus the use of models originally developed for financial applications might be appropriate. Rain and financial assets forecasting are actually related fields of research because both phenomena have statistical properties corresponding to turbulent and fractal processes: intermittency, non-stationarity, non-linearity and scale invariance (concerning rain, see: Schertzer & Lovejoy, 1997, Bendjoudi & al., 1997, Lavergnat & Golé, 1998, concerning finance, see: Ghashghaie et al., 1996). A new modeling of attenuation based on a switching ARIMA/GARCH process is proposed (Sect. 3) with its corresponding multi-step prediction algorithm (Sect. 4). This type of statistical model has already been proven to model successfully other geophysical processes related to turbulence and fractals, such as wind speed (Ewing et al., 2005), river flow (Wang et al., 2005) as well as internet traffic (Zhou et al., 2006). Combined with a frequency scaling model adapted to the EHF band (Sect. 6), this model is shown to predict an accurate upper bound of the uplink attenuation using past values of the downlink attenuation, thus reducing the cost of the FMT for high levels of link availability.

The problem of frequency scaling is not trivial, because above 20 GHz, attenuation effects of atmospheric gases and clouds cannot be neglected. The physical components of attenuation therefore have to be separated. This step is performed with a neural network (Mallet et al., 2006) and the attenuation components are then scaled with specific scaling factors that are estimated from ground meteorological data (ITU-R, 2003; Liebe, 1989; Gibbins, 1986). An assessment of the contributions to the overall uplink prediction error shows that the error due to

the frequency scaling is larger than the one due to the downlink prediction (Sect. 7). This result indicates that further study should focus on improving the accuracy of the scaling factor.

2 Measurements

A database was created from attenuations of the Olympus satellite 20 GHz and 30 GHz beacons measured at Gometz-la-Ville, France, during 15 months at an elevation angle of 30° (Golé et al., 1994; OPEX, 1994). The data were recorded at a sampling frequency of 100 Hz and averaged over 1 second time intervals. Sky brightness temperatures measured with a ground radiometer were used to estimate the attenuation reference level during clear sky periods. The power received from the satellite was shifted to this reference level in order to retrieve the total attenuation of the atmosphere that includes the effects of gases, clouds and rain. Since the stochastic model presented in this paper has been developed for real time applications, the data sampled at 1Hz are unfiltered.

A technical incident on the satellite led to the failure of the Olympus 30 GHz beacon on 10/10/1992, and as a consequence, attenuation measurements at this frequency only lasted five months. However, as much as 67 attenuation events were selected, for a total of 550 hours, including at least 57 hours of rainfall (the criterion being that the 20 GHz attenuation level exceeds a threshold of 1.5 dB, Cf. Sect. 3). Figure 1 shows an example of attenuation time series caused by rain events on the 20 GHz beacon. The Olympus propagation experiment also includes measurements of ground meteorological data (temperature, pressure and humidity) that are useful to accurately split attenuation into its physical components (Sect. 6).

3 The switching ARIMA-GARCH Model

Examining the down-link attenuation time series and its autocorrelation suggests that it can be described as a switching ARIMA process (Auto Regressive Integrated Moving Average) with GARCH errors (Generalized Auto Regressive Conditionally Heteroscedastic) (Engle, 1982; Bollerslev, 1986). This chapter describes the identification process and the method for estimating its parameters.

The first difficulty in modeling the attenuation process comes from the intermittency of rain. Rain is a very sparse phenomenon that occurs typically only 3% of the time. When rain is

present, the attenuation is expected to be volatile (i.e. its variations can be large in short periods of time). At the contrary, when rain is not present, the attenuation is only due to atmospheric gases and clouds and its variations are expected to be smoother. Thus, if these two different situations are not treated separately, the parameters of the model will be characteristic of the drought periods that represent 97% of the time! A solution to solving this problem is to separate ‘smooth’ periods (clear sky, clouds, and/or light rain) from ‘volatile’ periods (significant rain events) with a threshold. Since rain has a much stronger attenuation effect than gases and clouds, it is assumed that ‘volatile’ periods are occurring only when the attenuation exceed a threshold T . A neural network trained to separate the attenuation components (Mallet et al., 2006) is presented in Sect. 6. The output of this neural network shows that, at 20 GHz and a 30° elevation angle, rain becomes the major contributor to attenuation above 1.5 dB (see Fig. 8, Sect. 6). The threshold T is therefore set at 1.5 dB.

The thresholding yields two statistically coherent time series, one containing the ‘volatile’ periods and the other the ‘smooth’ periods (first row in Fig. 2). It was found that their autocorrelations remain closed from 1 during important time lags (second row in Fig. 2), which indicates that both processes behave like random walks (Box & Jenkins, 1976). The time series are therefore non-stationary and a pre-processing is required. The classical method to achieve stationarity consists in performing successive differentiations. One differentiation is found to be enough, because the resulting time series are roughly stationary (third row in Fig. 2): they have zero mean values regardless of the time.

Nevertheless, the differentiated time series also present periods with large variations, indicating that their variances are not constant. This phenomenon, called volatility clustering or heteroscedasticity in finance, shows that the differentiated time series still present a second-order non-stationarity. Therefore, the Wold decomposition theorem (Wei, 2005, p. 271) that requires the assumption of second-order stationarity cannot be applied theoretically. However, the differentiated time series are not far from stationarity, so that they can be modeled as classical ARMA processes in a first step and the problem of heteroscedasticity can be addressed in a second step. The identification of the ARMA model's orders, performed by the so called corner method (Wei, 2005, p. 133), yielded the following equations:

$$\Delta A_t = A_t - A_{t-1} \tag{1}$$

$$\begin{cases} \Delta A_t = \phi'_1 \Delta A_{t-1} + \phi'_2 \Delta A_{t-2} + \varepsilon'_t + \theta'_1 \varepsilon'_{t-1} + \theta'_2 \varepsilon'_{t-2} & \text{if } A_t \geq T \\ \Delta A_t = \phi''_1 \Delta A_{t-1} + \varepsilon''_t + \theta''_1 \varepsilon''_{t-1} + \theta''_2 \varepsilon''_{t-2} & \text{if } A_t < T \end{cases} \quad (2)$$

with:

$$\begin{cases} \Delta \hat{A}_t = \phi'_1 \Delta A_{t-1} + \phi'_2 \Delta A_{t-2} + \theta'_1 \varepsilon'_{t-1} + \theta'_2 \varepsilon'_{t-2} & \text{if } A_t \geq T \\ \Delta \hat{A}_t = \phi''_1 \Delta A_{t-1} + \theta''_1 \varepsilon''_{t-1} + \theta''_2 \varepsilon''_{t-2} & \text{if } A_t < T \end{cases} \quad (3)$$

$$\begin{cases} \varepsilon'_t = \Delta A_t - \Delta \hat{A}_t & \text{if } A_t \geq T \\ \varepsilon''_t = \Delta A_t - \Delta \hat{A}_t & \text{if } A_t < T \end{cases} \quad (4)$$

Where A_t is the attenuation, ΔA_t , the differentiated attenuation, $\Delta \hat{A}_t$, the predicted differentiated attenuation one second forward, and ε'_t and ε''_t , the errors of the ARMA predictors. The time t is in seconds and the subscripts $t-n$ means n seconds before t . The parameters of the ARMA models ($\phi'_1, \phi'_2, \theta'_1, \theta'_2$ and $\phi''_1, \theta''_1, \theta''_2$) are estimated with a least squares algorithm (Ljung, 1999). The model that combines the differentiation and the ARMA process is called ARIMA(p, i, q) (Auto Regressive Integrated Moving Average), where p is the order of the autoregressive part of the model, i the number of differentiation, and q the order of the moving average part of the model. The attenuation time series is thus modeled by an ARIMA(2,1,2) during rain periods ($A_t \geq T$) and by an ARIMA(2,1,1) during clear sky or cloud periods ($A_t < T$). This result is in agreement with Portsmouth University's model (Grémont et al., 1999) in which an adaptive ARMA(3,3) is presented (equivalent to ARMA(2,1,2) in terms of orders). Whenever the attenuation crosses the threshold T , the model has to switch and a transition period may occur due to the recursiveness of the models. This effect can be avoided by running both ARIMAs regardless of the rain or clear sky periods, threshold T only being used for switching between the two model outputs.

The ARIMA models would be valid if their errors were white noises. Although their errors do not seem to be autocorrelated processes, they are not white noises because their conditional variances are not constant (first row in Fig. 3, the error time series exhibit periods with large variations). As mentioned previously, the linear ARIMA can only model the information contained in the autocorrelation of the differentiated time series which is very low (fourth row in Fig. 2) and is not capable of modeling heteroscedasticity. Because of this problem, the confidence intervals of the ARIMA predictions, assumed to be constant, will not be accurate. Moreover, the least squares algorithm used to estimate the ARIMA parameters is biased when

the error variance is not constant. A solution consists in modeling heteroscedasticity as a nonlinear relationship between consecutive errors. Heteroscedasticity, which leads to a large Kurtosis in the error distributions (second row in Fig. 3), is a feature often observed in financial time series and various nonlinear models have been developed to describe this type of process. The significant value over 10-second time lags of the autocorrelation of the realized variance, that is defined as the square of the error, suggests that a classical GARCH(1,1) model might be appropriate (third row in Fig. 3):

$$\begin{cases} \varepsilon'_t = \eta'_t \cdot \sigma'_t \\ \varepsilon''_t = \eta''_t \cdot \sigma''_t \end{cases} \quad (5)$$

$$\begin{cases} \sigma'^2_t = \omega' + \alpha' \cdot \varepsilon'^2_{t-1} + \beta' \cdot \sigma'^2_{t-1} \\ \sigma''^2_t = \omega'' + \alpha'' \cdot \varepsilon''^2_{t-1} + \beta'' \cdot \sigma''^2_{t-1} \end{cases} \quad (6)$$

Where σ'_t and σ''_t are the conditional standard deviations of the errors of the ARMA predictors, η'_t and η''_t , the normalized errors and $(\omega', \alpha', \beta', \omega'', \alpha'', \beta'')$, the parameters of the GARCH models. The main assumption of GARCH models is that the error is a Gaussian white noise with a variance equal to 1 (corresponding to η'_t and η''_t) modulated by a time-varying standard deviation (corresponding to σ'_t and σ''_t). Figure 4 shows an example of conditional standard deviation time series estimated from GARCH compared with the error time series. The standard deviations σ'_t and σ''_t are said to be *conditional* because they are time-varying and depend on the past values of time series. They do not correspond to the *absolute* standard deviation estimated with the central limit theorem. Thus, once the errors are normalized by the conditional standard deviation, the resulting processes η'_t and η''_t are assumed to be identically-distributed Gaussian variables, which allows estimating the GARCH parameters through maximum likelihood. However, in the case of rain attenuation, since the errors depend on the ARIMA parameters which are biased by heteroscedasticity, the GARCH parameters must be estimated jointly with the ARIMA parameters. Finally, the solution consists in initializing the ARIMA parameters by means of the least squares algorithm, with standard values for the GARCH parameters ($\alpha = 0.05$ and $\beta = 0.85$), and refining both the ARIMA and GARCH parameters through maximum likelihood (Hamilton, 1994). One half of the database was used as a learning database and the following parameters were found:

$$\begin{cases} \phi'_1 = 1.1924 \\ \phi'_2 = -0.2309 \\ \theta'_1 = -1.5938 \\ \theta'_2 = 0.6281 \end{cases} \quad \text{and} \quad \begin{cases} \phi''_1 = 0.1659 \\ \theta''_1 = -0.8046 \\ \theta''_2 = -0.1064 \end{cases} \quad (7)$$

$$\begin{cases} \omega' = 5.15 \text{ e} - 5 \\ \alpha' = 0.0674 \\ \beta' = 0.9306 \end{cases} \quad \text{and} \quad \begin{cases} \omega'' = 1.2 \text{ e} - 5 \\ \alpha'' = 0.0331 \\ \beta'' = 0.9649 \end{cases} \quad (8)$$

The analysis of the normalized errors shows that their conditional variances are constant (first row in Fig. 5) and equal to one. Moreover, their distributions are found to be Gaussian (second row in Fig. 5). Since it was also found that neither the normalized errors nor the squared normalized errors are autocorrelated processes (third and fourth row in Fig. 5), the latter are indeed second order Gaussian white noises. As a consequence, the switching ARIMA-GARCH model is validated. The fact that the squared normalized errors are not autocorrelated processes confirms the appropriateness of the GARCH model for representing nonlinear dependency relationships due to the squared errors. The next step is to develop a k -step forward algorithm based on the switching ARIMA/GARCH model that is able to accurately predict in real time the future attenuation level and the distribution of its error.

4 Multi-step prediction

The attenuation process being integrated one time, a first step in the k -step forward algorithm is to write the attenuation at time $t+k$ using the differentiated time series:

$$A_{t+k} = A_t + \sum_{i=1}^k \Delta A_{t+i} \quad (9)$$

Therefore, the predicted attenuation is:

$$\hat{A}_{t+k} = A_t + \sum_{i=1}^k \Delta \hat{A}_{t+i} \quad (10)$$

The prediction of the differentiated terms can be achieved by iterating i times the ARMA predictors (Eq. (3), Sect. 3). These iterations require the one second forward errors of the ARMA predictors which are not known after time t in real time. The future errors are therefore replaced by their expected values which is zero. Nevertheless, the first two iterations involve

deterministic past errors ($\varepsilon'_t, \varepsilon'_{t-1}, \varepsilon''_t, \varepsilon''_{t-1}$) that are known at the time of forecasting. After these first two iterations, the recurrence equation (Eq. (3)) is thus simplified. In order to simplify the equations, the dependency on threshold T is omitted, and only the case where $A_t > T$ will be discussed in the following:

$$\begin{cases} \Delta \hat{A}_{t+1} = \phi_1 \Delta A_t + \phi_2 \Delta A_{t-1} + \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} \\ \Delta \hat{A}_{t+2} = \phi_1 \Delta \hat{A}_{t+1} + \phi_2 \Delta A_t + \theta_2 \varepsilon_t \\ \Delta \hat{A}_{t+i} = \phi_1 \Delta \hat{A}_{t+i-1} + \phi_2 \Delta \hat{A}_{t+i-2} \quad \text{for } i > 2 \end{cases} \quad (11)$$

The k -step forward prediction of the attenuation can thus be calculated.

As to the prediction of the variance of the prediction error, Eq. (2) must be iterated i times to obtain the differentiated terms of Eq. (9), but this time without considering that the future errors are zero because, contrary to their expected values, their expected variances are not zero. Since Eq. (2) is linear, it is possible to separate the iteration of the unknown errors terms from the iteration performed previously with Eq. (11). Thus, the overall result of the iteration is the sum of the previously calculated term $\Delta \hat{A}_{t+i}$ with a linear combination of the future errors.

$$\Delta A_{t+i} = \Delta \hat{A}_{t+i} + \sum_{j=1}^i \lambda_{i,j} \varepsilon_{t+j} \quad (12)$$

This calculus of the coefficients ($\lambda_{i,1}, \dots, \lambda_{i,i}$) is performed by iterating i times Eq. (2) with suppressing all the deterministic past terms ($\varepsilon_t, \varepsilon_{t-1}, \Delta A_t, \Delta A_{t-1}$) in the first two iterations, since only the future unknown errors are considered. This problem is equivalent to finding the explicit form of the i th term of a sequence U_i defined by the following recurrence equation:

$$\begin{cases} U_1 = \varepsilon_{t+1} \\ U_2 = \phi_1 U_1 + \varepsilon_{t+2} + \theta_1 \varepsilon_{t+1} \\ U_i = \phi_1 U_{i-1} + \phi_2 U_{i-2} + \varepsilon_{t+i} + \theta_1 \varepsilon_{t+i-1} + \theta_2 \varepsilon_{t+i-2} \quad \text{for } i > 2 \end{cases} \quad (13)$$

Due to the complexity of this recurrence equation, the coefficients ($\lambda_{i,1}, \dots, \lambda_{i,i}$) have no simple analytic expression and have to be computed digitally. Then, by substituting Eq. (12) in Eq. (9):

$$A_{t+k} = At + \sum_{i=1}^k \Delta \hat{A}_{t+k} + \sum_{i=1}^k \left(\sum_{j=1}^i \lambda_{i,j} \varepsilon_{t+j} \right) \quad (14)$$

And then replacing the first two terms with Eq. (10):

$$A_{t+k} = \hat{A}_{t+k} + \sum_{i=1}^k \left(\sum_{j=1}^i \lambda_{i,j} \varepsilon_{t+j} \right) \quad (15)$$

The k -step forward prediction error is thus obtained:

$$e_{t+k} = A_{t+k} - \hat{A}_{t+k} = \sum_{i=1}^k \left(\sum_{j=1}^i \lambda_{i,j} \varepsilon_{t+j} \right) = \sum_{j=1}^k \left(\sum_{i=j}^k \lambda_{i,j} \right) \varepsilon_{t+j} \quad (16)$$

And its conditional variance is:

$$V_c(e_{t+k}) = V_c \left(\sum_{j=1}^k \mu_j \varepsilon_{t+j} \right) \quad \text{where} \quad \mu_j = \sum_{i=j}^k \lambda_{i,j} \quad (17)$$

$V_c(\cdot)$ denotes the conditional variance operator. Since the errors are not autocorrelated (Cf. Sect. 3), the conditional variance operator can be distributed:

$$V_c(e_{t+k}) = \sum_{j=1}^k \mu_j^2 V_c(\varepsilon_{t+j}) \quad (18)$$

Since σ_{t+j} is by definition the conditional standard deviation of ε_{t+j} , Eq. (18) is equivalent to:

$$V_c(e_{t+k}) = \sum_{j=1}^k \mu_j^2 \hat{\sigma}_{t+j}^2 \quad (19)$$

The one second forward conditional standard deviations σ_{t+j} of the future errors can be predicted using the GARCH recurrence relations (Eq. (6), Sect. 3). Since the errors are centered, for $j \geq 2$, the expected value of ε_{t+j}^2 is σ_{t+j}^2 and Eq. (6) can be simplified to:

$$\begin{cases} \sigma_{t+1}^2 = \omega + \alpha \varepsilon_t^2 + \beta \sigma_t^2 \\ \hat{\sigma}_{t+j}^2 = \omega + (\alpha + \beta) \hat{\sigma}_{t+j-1}^2 \quad \text{for } j > 1 \end{cases} \quad (20)$$

By iterating Eq. (20) and then substituting it in Eq. (19), the predicted variance of the prediction error is finally obtained:

$$\hat{V}_c(e_{t+k}) = \sum_{j=1}^k \mu_j^2 \left(\omega \frac{1 - (\alpha + \beta)^{j-1}}{1 - (\alpha + \beta)} + (\alpha + \beta)^{j-1} \sigma_{t+1}^2 \right) \quad (21)$$

5 Comparison and results

A comparison has been performed with five other prediction models over a 10-second forecast interval. The simplest model, called persistence, relies on the assumption that the attenuation level remains constant over the forecast interval. The Portsmouth University model (Grémont et al. 1999) is an adaptive ARMA(3,3) whose parameters are updated in real time by a Recursive Least Squares (RLS) algorithm (Ljung, 1999). The ADALINE model (Chambers & Otung, 2005) is a single neuron with a linear activation function. The learning parameter was empirically set at 10^{-4} . The NASA model is based on a non-linear Markov filtering (Manning, 1990, 1991). The main assumption of this model is that the normalized logarithm of the attenuation follows a first-order stochastic equation. The noise variance parameter was set at 0.1 dB. The ‘Two-Sample’ model (Van de Kamp, 2002; Bolea-Alamañac et al., 2003) is based on statistical study of the fade slope. This model is presented with a 0.1 Hz cutoff frequency filtering to remove scintillations. However, this preprocessing is not suitable in real time because the causal filter's phase shifting effect lowers the model's performance. For this reason, filtering is used to determine the parameters but not in the real time comparison. The parameters used are reported in Van de Kamp (2002).

Since certain models have been specifically developed for rain periods, the results of the comparison can be influenced by the percentage of clear sky or cloud periods in the OLYMPUS 20 GHz database. Therefore, the performances of the various models were computed only over ‘volatile’ periods (20 GHz attenuation greater than 1.5 dB). One half of the database was used as a learning database and the other half was used for the comparison. Figure 6 shows the Root Mean Square Errors (RMSE) of the prediction models according to the attenuation level. In spite of their complexity, all the models perform quite similarly to persistence and it is not possible to decide between them.

The RMSE is the criterion conventionally used to compare the different methods because the prediction error was assumed to be due to scintillations with a constant variance. In this case, the error margin is fixed and has no influence on the comparison. However, this assumption does not seem realistic for various reasons. Firstly, the scintillations variance is unlikely to remain constant over more than a few minutes because it has a strong dependence with meteorological data (Grémont et al. 1999; Baxter et al., 2003). Secondly, the variance of the fade slope distribution is shown to increase with the attenuation level (Van de Kamp, 2002). Thirdly, the high frequencies components of attenuation contain not only scintillations, but also

fast variations of rain attenuation whose variance is not expected to be constant (Garcia et al., 2002; Baxter & Upton, 2005). This analysis is confirmed by the fact that the errors time series are found to be heteroscedastic. Thus, the RMSE criterion is not entirely suitable to this problem because it does not take into account the error margin that has to be added to the prediction and therefore does not measure the ability of the model to give the upper bound of future attenuation. The predicted upper bound of the attenuation at time $t+k$, noted \bar{A}_{t+k} , which is the prediction plus the adaptive error margin, has to be taken into account in the comparison criterion:

$$\bar{A}_{t+k}(P) = \hat{A}_{t+k} + \bar{M}_{t+k}(P) \quad (22)$$

Where \hat{A}_{t+k} is the predicted attenuation and $\bar{M}_{t+k}(P)$ the adaptive error margin. The latter depends on the required link availability P , which is conventionally defined as the percentage of time during which the measured attenuation level does not exceed the predicted upper bound. Although the prediction error variance has already been estimated according to Eq. (21) (Sect. 4), the calculation of the error margin is not straightforward because the whole conditional cumulative distribution of the prediction error is needed. Equation (16) (Sect. 4) shows that the k -step forward prediction error is a linear combination of non-correlated centered Gaussian variables $(\varepsilon_{t+1}, \dots, \varepsilon_{t+k})$. Its distribution is therefore also a centered Gaussian variable. Thus, the error margin needed to reach a given availability P can be computed using the inverse Gauss error function $erf^{-1}(\cdot)$:

$$\bar{M}_{t+k}(P) = erf^{-1}\left(\frac{P}{100} - \frac{1}{2}\right) \sqrt{2\hat{V}_c(e_{t+k})} \quad (23)$$

Since the error margin cannot be neglected in the assessment of the model performance, a new comparison criterion must be used. The mean cost C is defined, in dB, as the mean overestimation of the attenuation level by its predicted upper bound and depends on the required availability:

$$C(P) = \frac{1}{N} \int H(\bar{A}_{t+k}(P) - A_{t+k}) (\bar{A}_{t+k}(P) - A_{t+k}) dt \quad (24)$$

Where N , in seconds, is the duration of the attenuation time series and $H(\cdot)$ the Heaviside step function. The mean cost C allows the accuracy of the predicted upper bound to be compared with the link availability that is reached.

As far as the existing prediction models are concerned, it is assumed that the error margin is constant, except for the Portsmouth University and ONERA models in which an error margin model is provided. The ONERA model, based on a statistical study, proposes a hyperbolic secant distribution for the prediction error whose parameters depend on the attenuation level. The Portsmouth University model assumes that the prediction error is due to scintillations whose variance is constant over a one-minute time interval and is therefore computed over the 60 last samples. The cost/availability relationships presented in Fig. 7 show that, compared to the other models, the switching ARIMA-GARCH model reduces the mean cost of the FMT system by nearly 30% for a required availability of 99% during rain events.

6 Frequency scaling

Frequency scaling is necessary because typically the prediction is performed based on the down-link attenuation and must be scaled to the up-link frequency. In the case of the OLYMPUS-Gometz-la-Ville link, the attenuation measurements of the 20 GHz and 30 GHz satellite beacons are used to simulate such a system. At these frequencies, the effect of the atmospheric gases (oxygen and water vapor) and clouds cannot be neglected as they can cause up to 1.5 dB and 4 dB of attenuation, respectively, during strong rain events (Barthes et al., 2006). Since each component of attenuation has a specific frequency scaling factor, a separation of the physical effects must be carried out. A neural network has been developed at CETP for this purpose (Mallet et al., 2006). The input data are the 20 GHz attenuation and ground meteorological data (pressure, temperature and humidity), which improve the performance of the neural network. The outputs provide an estimation of the contributions of each component to the overall measured attenuation. They are called the ‘normalized weights’ (denoted W_{gas} , W_{clouds} and W_{rain}) and are defined by the following equations:

$$\begin{cases} A_{rain} = W_{rain} \times A_{total} \\ A_{clouds} = W_{clouds} \times A_{total} \\ A_{gaz} = W_{gaz} \times A_{total} \end{cases} \quad (25)$$

Atmospheric profiles of meteorological data from the European Centre for Medium-Range Weather Forecasts (ECMWF) were used to create the learning database. Gas attenuation was computed from Liebe's model (Liebe et al., 1993), the cloud attenuation from the Rayleigh diffusion model (Liebe et al., 1993) and the rain attenuation from Mie's scattering model (Mie,

1908). In the case of the Olympus-Gometz-la-Ville link (20 GHz, 30° elevation angle) the separation of the effects yields the normalized weights presented in Fig. 8 (published in Barthes et al. (2006)).

Once the separation of the effects has been carried out, the frequency scaling can be performed.

The scaling factor is defined as:

$$K = \frac{A(F_2)}{A(F_1)} \quad (25)$$

Where $F_1=20$ GHz and $F_2=30$ GHz in this study. The uplink attenuation can be written as follows when using the specific scaling factors of each component K_{gas} , K_{clouds} and K_{rain} :

$$A(F_2) = K_{gas} \cdot A_{gas}(F_1) + K_{clouds} \cdot A_{clouds}(F_1) + K_{rain} \cdot A_{rain}(F_1) \quad (26)$$

Then, with the component weights given by the neural network (Eq. (24)):

$$A(F_2) = (K_{gas} \cdot W_{gas} + K_{clouds} \cdot W_{clouds} + K_{rain} \cdot W_{rain}) A(F_1) \quad (27)$$

And finally, by identification with Eq. (25):

$$K = W_{gas} \cdot K_{gas} + W_{clouds} \cdot K_{clouds} + W_{rain} \cdot K_{rain} \quad (28)$$

The calculation of the specific scaling factors is simplified by the fact that the unknown effective length of the link is eliminated in the ratio of the attenuations. The scaling factor of the gas component is therefore a function of the specific attenuations of gases (denoted γ):

$$K_{gas} = \frac{\gamma_{O_2}(F_2) + \gamma_{H_2O}(F_2)}{\gamma_{O_2}(F_1) + \gamma_{H_2O}(F_1)} \quad (29)$$

The specific attenuations of gases are estimated from the ITU-R model (Gibbins, 1986), which uses ground meteorological data and models of standard atmospheric profiles. The cloud specific attenuation is also given by the ITU-R model (Liebe, 1989). The ratio eliminates the unknown cloud water content, and the scaling factor of the cloud attenuation only depends on the ratio of the frequencies with a second order dependency on temperature. The scaling factor of the rain component is given in ITU-R (2003):

$$K_{rain} = \left(\frac{\phi(F_2)}{\phi(F_1)} \right)^{\phi(F_1, F_2, A(F_1))} \quad (30)$$

$$\phi(F) = \left(\frac{F^2}{1 + 10^{-4} F^2} \right) \quad (31)$$

$$\varphi = 1 - 112.10^{-5} \left(\frac{\phi(F_2)}{\phi(F_1)} \right)^{0.5} (\phi(F_1)A(F_1))^{0.55} \quad (32)$$

An estimation of the scaling factor, noted \hat{K}_t , can be calculated with the method that combines separation of the effects and specific scaling of each component. The error, noted δ_t , is the difference between the estimation of the scaling factor and its measured value:

$$K_t = \hat{K}_t + \delta_t \quad (33)$$

Figure 9 presents the distribution of the scaling factor error computed with the Olympus 20/30 GHz database. The error is found to be a centered Gaussian random variable with a standard deviation equal to 0.15. It is interesting to notice that the estimation error of the scaling factor does not depend on the level of the attenuation. Therefore, the estimation error of the uplink attenuation that is due to the frequency scaling remains constant in percentage. Since the scaling factor is typically 2 during rain events at these frequencies, the uncertainty of the uplink attenuation is around 7%, regardless of the attenuation level.

7 Uplink attenuation upper bound

The switching ARIMA/GARCH multi-step algorithm provides an accurate prediction of the downlink attenuation and of the distribution of the prediction error. Nevertheless, frequency scaling increases the prediction error resulting from the estimation of the frequency scaling factor. In the following, $e_{t+k}(F_1)$ will denote the error of the downlink attenuation prediction and $e_{t+k}(F_2)$ will denote the prediction error of the uplink attenuation prediction after frequency scaling. Considering that the frequency scaling factor remains constant over the forecast interval, by definition, the uplink attenuation is (Eq. (25), Sect.6):

$$A_{t+k}(F_2) = K_t \cdot A_{t+k}(F_1) \quad (34)$$

And can be written as:

$$A_{t+k}(F_2) = (\hat{K}_t + \delta_t) \cdot (\hat{A}_{t+k}(F_1) + e_{t+k}(F_1)) \quad (35)$$

where δ_t is the error made on the estimation of the scaling factor. Since the cross error term is negligible, Eq. (35) can be simplified as:

$$A_{t+k}(F_2) = \hat{K}_t \cdot \hat{A}_{t+k}(F_1) + \hat{A}_{t+k}(F_1) \cdot \delta_t + \hat{K}_t \cdot e_{t+k}(F_1) \quad (36)$$

Let $\hat{A}_{t+k}(F_2)$ be the up-link attenuation prediction. The attenuation at time $t+k$ can also be written as:

$$A_{t+k}(F_2) = \hat{A}_{t+k}(F_2) + e_{t+k}(F_2) \quad (37)$$

The identification of the deterministic and stochastic terms of Eq. (36) and Eq. (37) leads to the following expressions:

$$\hat{A}_{t+k}(F_2) = \hat{K}_t \cdot \hat{A}_{t+k}(F_1) \quad (38)$$

$$e_{t+k}(F_2) = \hat{A}_{t+k}(F_1) \cdot \delta_t + \hat{K}_t \cdot e_{t+k}(F_1) \quad (39)$$

Here, $e_{t+k}(F_1)$ is a centered Gaussian variable because it is a sum of independent centered Gaussian variables (cf. Sect. 5). Moreover, Fig. 9 (Sect.6) shows that δ_t is also a centered Gaussian variable. The linear combination of both errors, $e_{t+k}(F_2)$, is therefore also a centered Gaussian variable. Since there is no mutual dependency between these two errors because the scaling factor error is not related to the prediction error, their variances can be added:

$$V_c(e_{t+k}(F_2)) = \hat{A}_{t+k}(F_1)^2 \cdot V(\delta_t) + \hat{K}_t^2 \cdot V_c(e_{t+k}(F_1)) \quad (40)$$

Here, $V_c(e_{t+k}(F_1))$ is given by the multi-step algorithm in Eq. (21) (Sect. 4) and the value of $V(\delta_t)$ in the case of the Olympus 20/30 GHz database was reported in Sect.6. The uplink error margin needed to reach a given availability P is then computed using the inverse Gauss error function $erf^{-1}(\cdot)$:

$$\bar{M}_{t+k}(F_2) = erf^{-1}\left(\frac{P}{100} - \frac{1}{2}\right) \sqrt{2V_c(e_{t+k}(F_2))} \quad (41)$$

Finally, the predicted uplink upper bound can be derived from Eq. (38) and Eq. (41):

$$\bar{A}_{t+k}(F_2) = \hat{A}_{t+k}(F_2) + \bar{M}_{t+k}(F_2) \quad (42)$$

Figure 10 shows the performance of the switching ARIMA/GARCH model when frequency scaling is involved. The results are computed from the 20/30 GHz Olympus beacon data (only

‘volatile’ periods, $A_t(F_1) \geq 1.5$ dB) given in terms of cost/availability. It is not possible to compare these results with other existing models that combine the prediction of the downlink attenuation and a frequency scaling method, as no such models were found in the literature. However, Fig. 10 indicates that the predicted upper bound overestimates the measured 30 GHz attenuation by less than 2 dB on average for an availability reaching 99% of the time during rain events, which is quite interesting for FMT system designers.

In order to know which part of the prediction system is the major contributor to the uplink prediction error, the contributions of the downlink prediction error and the frequency scaling error (respectively, the first term and second term of Eq. (40)) have been estimated. Since these errors are not correlated, their RMSEs can be computed separately. The results (Fig. 11) show that under 6 dB attenuation at 30 GHz, both contributions are similar. Nevertheless, above this value, the contribution of the frequency scaling error is increasingly larger than the one of the downlink prediction error. Thus, at this stage, the estimation of the scaling factor seems to be more critical than the improvement of the prediction model. This result is based on 20/30 GHz Olympus data (30° elevation angle) and has to be confirmed with the 20/44 GHz Syracuse 3 measurements (17° elevation angle).

8 Conclusion

When modeling the prediction error volatility, it has been shown that the switching ARIMA-GARCH model outperformed other existing models. Using this model, the cost of the FMT can be reduced and high levels of link availability can be reached. This result can be explained by the fact that the prediction error contains not only scintillations with a constant variance, but also a fast fluctuating rain attenuation component that cannot be modeled by linear prediction methods. More generally, the nonlinear ARIMA/GARCH model, originally developed for finance applications, is also of interest for geophysical time series where fractal or turbulent underlying processes are involved, specifically those for which the distribution of differentiated time series has a large Kurtosis and which have a significant autocorrelation of its realized variance.

A frequency scaling method has been added to the switching ARIMA/GARCH model in order to predict the uplink channel propagation conditions from the downlink, which usually operates at a lower frequency. The error made on the estimation of the frequency scaling factor only

increases the prediction error reasonably, but its contribution is found to be larger than the one of the downlink prediction error, thus indicating that further studies should focus on improving the accuracy of the frequency scaling factor.

New measurements from the Syracuse 3 20/44 GHz propagation experiment should be soon available and will allow validating the presented model and checking its performance with different datasets, different frequencies and a different elevation angle (17°).

Acknowledgements

The authors thank the French DGA and especially Thierry Marsault from CELAR for the founding of this study.

References

- Barthes L., Mallet C., Brisseau O.: A neural network model for the separation of atmospheric effects on attenuation: Application to frequency scaling, *Radio Sci.*, Vol. 41, No. 4, 4012, doi: 10.1029/2005RS003310, 2006.
- Baxter, P. D., Upton, G. J. G., and Eden, D.: Measuring rain-fade-slope: objective evaluation of filtering techniques, *Proceeding of the International Workshop of COST Actions 272 and 280*, 67-74, 2003.
- Baxter P.D. and Upton G.J.G.: Differentiating noisy radio communications signals, wavelet estimation of a derivative in the presence of heteroskedastic noise, *Applied Statistics*, Vol. 54(4), pp. 753-767, 2005.
- Beguín J.-M., Gouriéroux C., Monfort A.: Identification of a mixed autoregressive-moving average process : The corner method, in O.D. Anderson, Ed., 'Time Series', Amsterdam, pp. 423-436, 1980.
- Bendjoudi H., Hubert P., Schertzer D., Lovejoy S.: Interpretation multifractale des courbes intensité-durée-fréquence des précipitations, *Multifractal point of view on rainfall intensity-duration-frequency curves*. *C. R. Acad. Sci. Paris, II*, 325, pp. 323-326, 1997.
- Bollerslev T., Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, pp. 307–327, 1986.

Box, G. E. P. and Jenkins, J. M.: Time series analysis: Forecasting and control, San Francisco, Holden-Day, 1976.

Bolea-Alamañac A., Bousquet M., Castanet L., Van de Kamp M.M.J.L.: Implementation of short-term prediction models in fade mitigation techniques control loops, COST 272/280 Workshop., Noordwijk, The Netherlands, ESA/ESTEC, ESTEC, PM5-067, 26-28 May 2003.

Castanet L., Deloues T., Lemorton J.: Channel modelling based on N-state Markov chains for SatCom systems simulation, IEE conf. Pub., Issue 49,1 Vol. 1, pp. 119 - 112. ICAP 2003, Exceter, UK, 31 March - 3 April 2003.

Chambers A.P., Otung L.E.: Neural network approach to short-term fade prediction on satellite links, Electronics Letters, 10th November 2005, Vol. 41, No. 23, pp.1990-1292.

Dissanayake A., Allnut J., Haidara F., A prediction model that combines rain attenuation and other propagation impairments along Earth-satellite paths, IEEE Trans. on Antennas and Propagation, Vol.45 , Issue10, pp. 1546-1558, 1997.

Dossi L.: Real-time prediction of attenuation for applications to fade countermeasures in satellite communications, Electronics Letters, Vol. 26, No. 4, pp. 250-251, 1990.

Engle, R.: Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation, Econometrica, 50, pp. 987–1008, 1982.

Ewing B.T., Kruse B.J., Schroeder J.L.: Time series analysis of wind speed with time-varying turbulence, Environmetrics, vol. 17, Issue 2, pp. 119-127, 2005.

Fiebig U.-C.: A Time-Series Generator Modelling Rain Fading, Proc. Open Symposium on Propagation and Remote Sensing, URSI Commission F, Garmisch-Partenkirchen, 2002.

Garcia P., Riera J.M., Benarroch A.: Statistics of dry and wet scintillation in Madrid using Italsat 50 GHz beacon, COST 280, 1st International workshop, PM3013, July 2002.

Gibbins C. J.: Improved algorithms for the determination of specific attenuation at sea level by dry air and water vapor, in the frequency range 1-350 GHz, Radio science, vol. 21(6), pp. 949-954, 1986.

Ghashghaie S., Breyman W., Peinke J., Talkner P., Dodge Y.: Turbulent cascades in foreign exchange markets, Nature 381, pp. 767 – 770, 1996.

Grémont B., Philip M., Galois P. and Bate S.: Comparative analysis and Performance of two Predictive Fade Detection Schemes for Ka-band Fade Countermeasures, IEEE Journal on selected areas in communications, vol.17, No. 2, pp. 180-192, Feb. 1999.

Golé P., Lavergnat J., Ulmer-Moll A.-M., Vernet M. : Les résultats de l'expérience OLYMPUS France Telecom FTR& D, technical report NT/CETP/001, 1994.

Hamilton J.D.: Time Series Analysis, Princeton University Press, 1994.

Hodges DD., Watson R.J.: Initial comparisons of forecast attenuation and beacon measurements at 20 and 40 GHz, Proceedings of the European Conference on Antennas and Propagation (EUCAP 2006), Nice, France, November 2006, p. 356.1, ESA SP-626.

ITU-R: Propagation data and prediction methods required for the design of Earth-space telecommunication systems, Recommendations of the ITU-R, Rec. P.618-8, 2003.

Lavergnat J., P. Golé, A Stochastic Raindrop Time Distribution Model, Journal of Applied Meteorology, 37, pp. 805-805, 1998.

Liebe H.J., MPM - An atmospheric millimeter-wave propagation model, Int. J. Infr. Mill. Waves 10, pp. 631-650, 1989.

Liebe H. J., Hufford G. A., and Cotton M. G., Propagation modeling of moist air and suspended water/ice particles below 1000 GHz, paper presented at AGARD 52nd Specialists Meeting of Electromagnetic Wave Propagation, Advis. Group for Aerosp. Res. and Dev., Palma de Mallorca, Spain, 17–21 May 1993.

Ljung L., System identification - theory for the user, 2nd edition, Prentice-Hall, 1999.

Mallet C., Barthes L., Marsault T., A neural network model for the separation of atmospheric effects on attenuation statistics, Proceedings of the European Conference on Antennas and Propagation (EUCAP 2006), Nice, France, November, 2006, ESA SP-626.

Manning R.M., A Unified statistical Rain Attenuation Model for Communication Link Fade Predictions and Optimal Stochastic Fade Control Design Using a Location Dependent Rain Statistic Database, Int. J. Satellite Commun., vol. 8, pp. 11-30, 1990.

Manning R.M., A statistical rain attenuation prediction model with application to the Advanced Communication Technology Satellite Project, part III: a stochastic rain fade control algorithm for satellite link power via nonlinear Markov filtering theory, NASA TM-100243, 1991.

Marsault T., Hermant J. D., Bouyer F. et al.: EHF Propagation Experiment with Syracuse 3 satellite, EuCAP 2006 (ESA SP-626), p. 357.1, 6-10 November 2006, Nice, France.

Mie G.: Contributions to the optics of turbid media, particularly of colloidal metal solutions, *Ann. Physik*, 25(3), pp. 377-445, 1908.

Núñez A., Pastoriza V., Mariño P., Fontán F.P., Fiebig U.-C.: Cellular automata for predicting three-state rain rate and rain attenuation field dynamics, Proceedings of the European Conference on Antennas and Propagation (EUCAP), Nice, 2006, ESA SP-626.

OPEX, Second workshop of the OLYMPUS propagation experimenters, vol. 1: reference book on attenuation measurement and prediction, Noordwijk, 8-10 Nov. 1994.

Peters O., Hertlein C., and Christensen K.: A complexity view of rainfall. *Phys. Rev. Lett.* 88, 018701, 1-4, 2002.

Schertzer D., Lovejoy S.: Universal Multifractals Do Exist!; Comments on "A Statistical Analysis of Mesoscale Rainfall as a Random Cascade". *Journal of Applied Meteorology*, 36, pp. 1296-1303, 1997.

Van de Kamp M.M.J.L.: Short-term prediction of rain attenuation using two samples, *Electronics Letters*, vol.38 (23), pp.1476-1477, 2002.

Wang W., Van Gelder P. H. A. J. M., Vrijling J. K. and Ma J.: Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes, *Nonlinear Processes in Geophysics*, 12, pp. 55–66, 2005.

Wei W.W.S.: *Time Series Analysis : Univariate and Multivariate Methods*, 2nd edn., Addison Wesley, 2005.

Zhou B., He D., and Sun Z.: Traffic modeling and prediction using ARIMA/GARCH model, in Proc. of 2nd EuroNGI Conference on Next Generation Internet Design and Engineering, 2006.

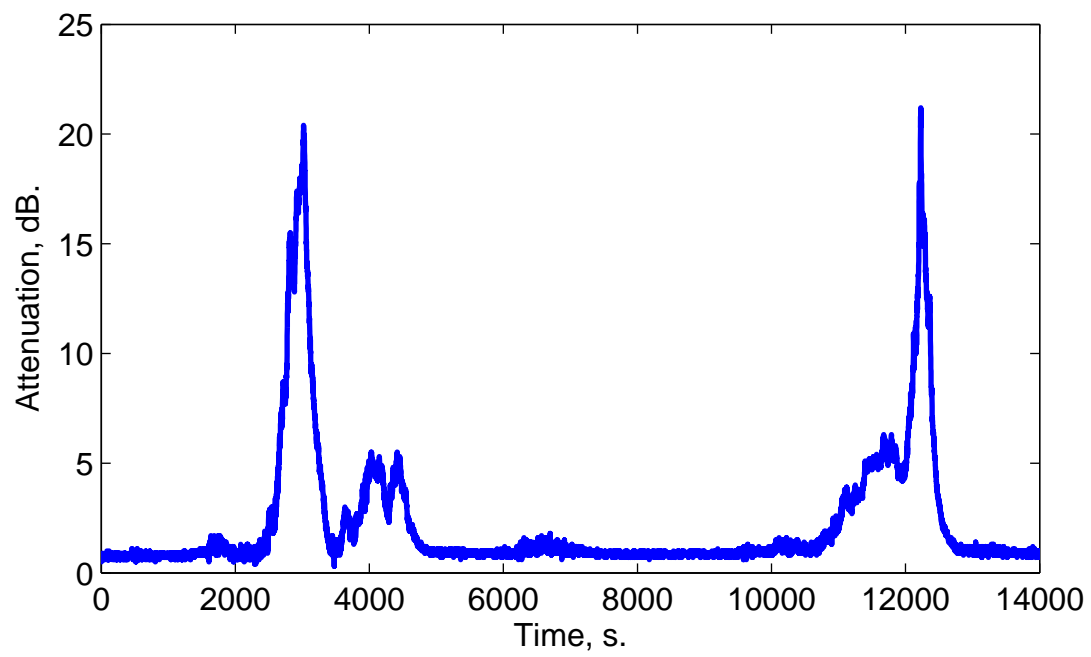


Figure 1. 20 GHz attenuation time series caused by rain events occurring on 30/06/1992.

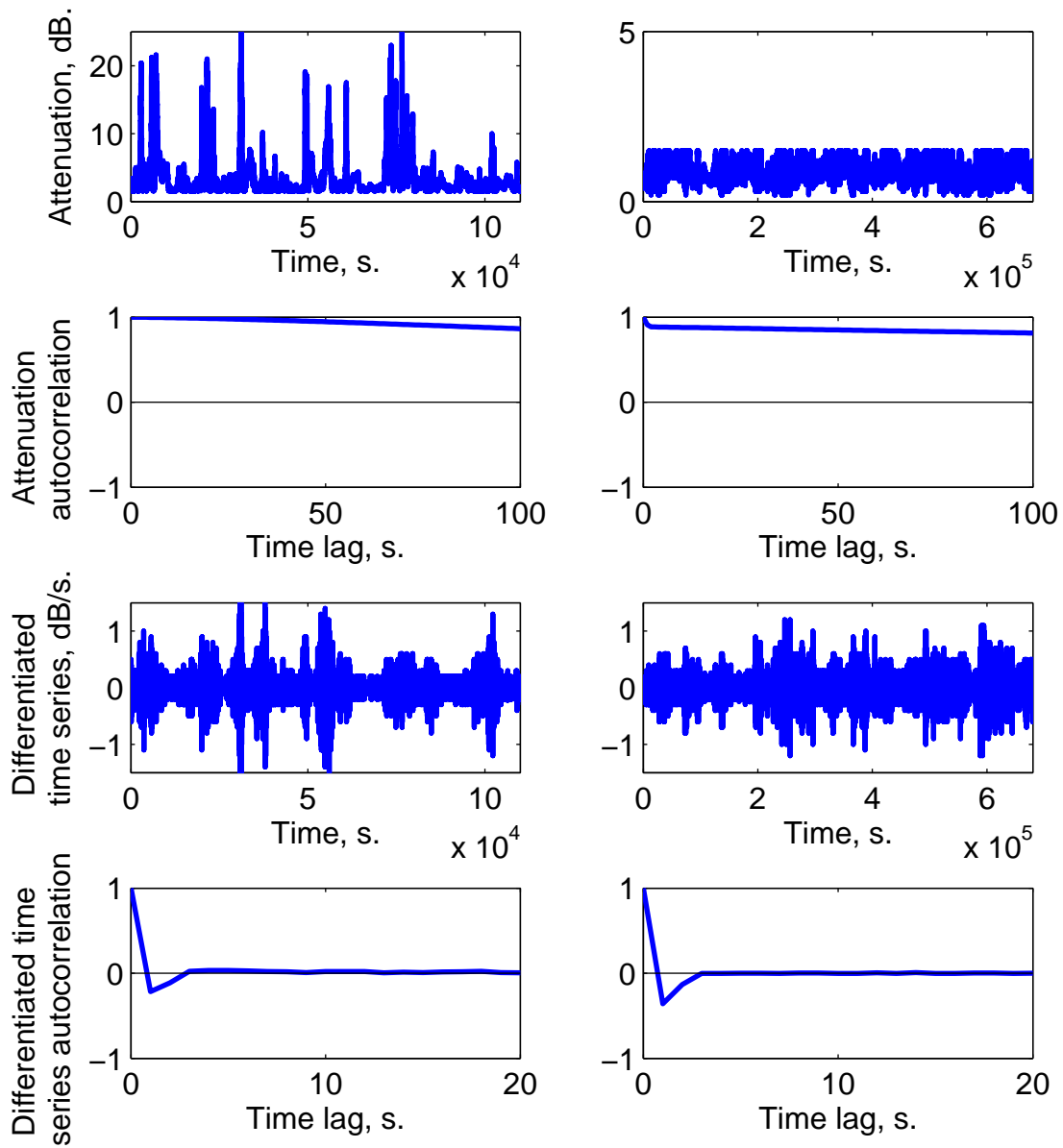


Figure 2. Attenuation time series and differentiated times series with corresponding autocorrelations (from top to bottom: attenuation time series, attenuation time series autocorrelation, differentiated time series, differentiated time series autocorrelation; from left to right: ‘volatile’ periods ($A_t \geq T$), ‘smooth’ periods ($A_t < T$)).

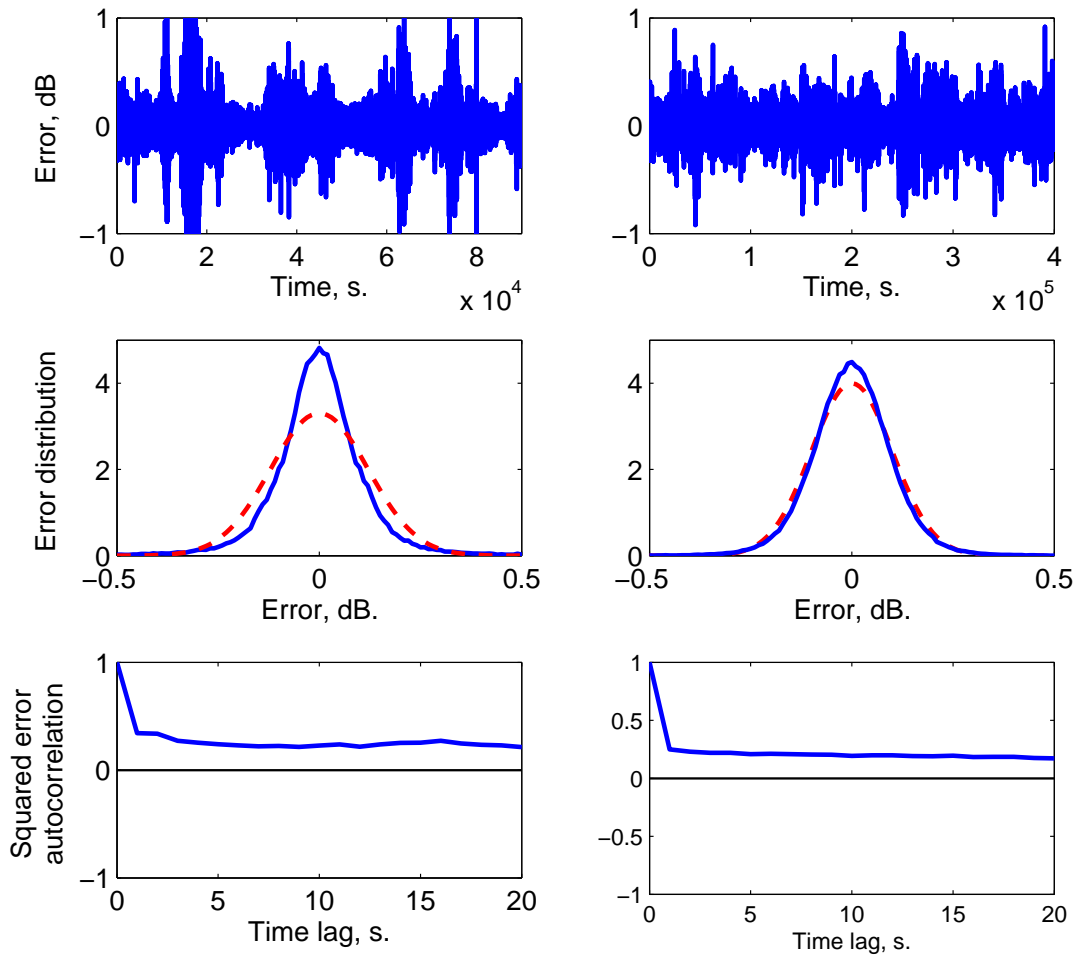


Figure 3. Statistical properties of the prediction errors (from top to bottom: error time series, error distributions compared with Gaussians, square error autocorrelations; from left to right: ‘volatile’ periods ($A_t \geq T$), ‘smooth’ periods ($A_t < T$))

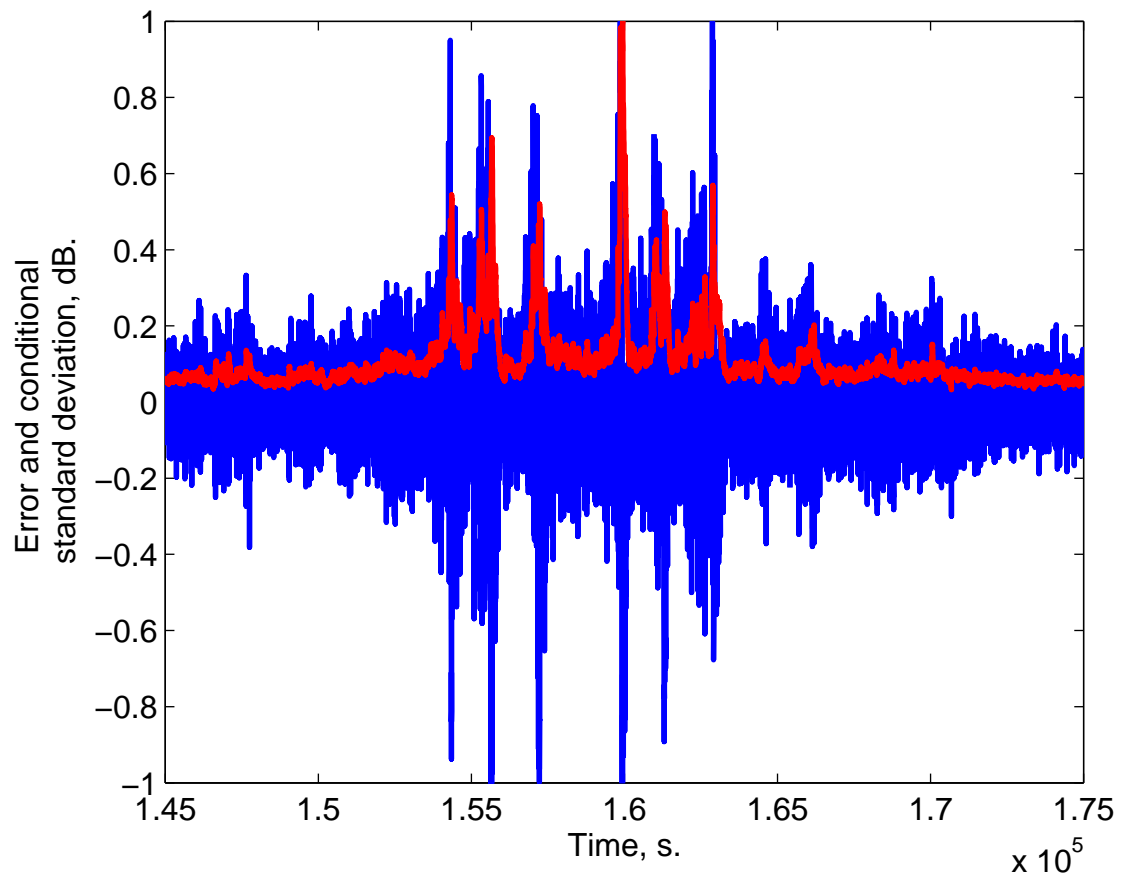


Figure 4. Example of error time series with the conditional standard deviation estimated according to GARCH.

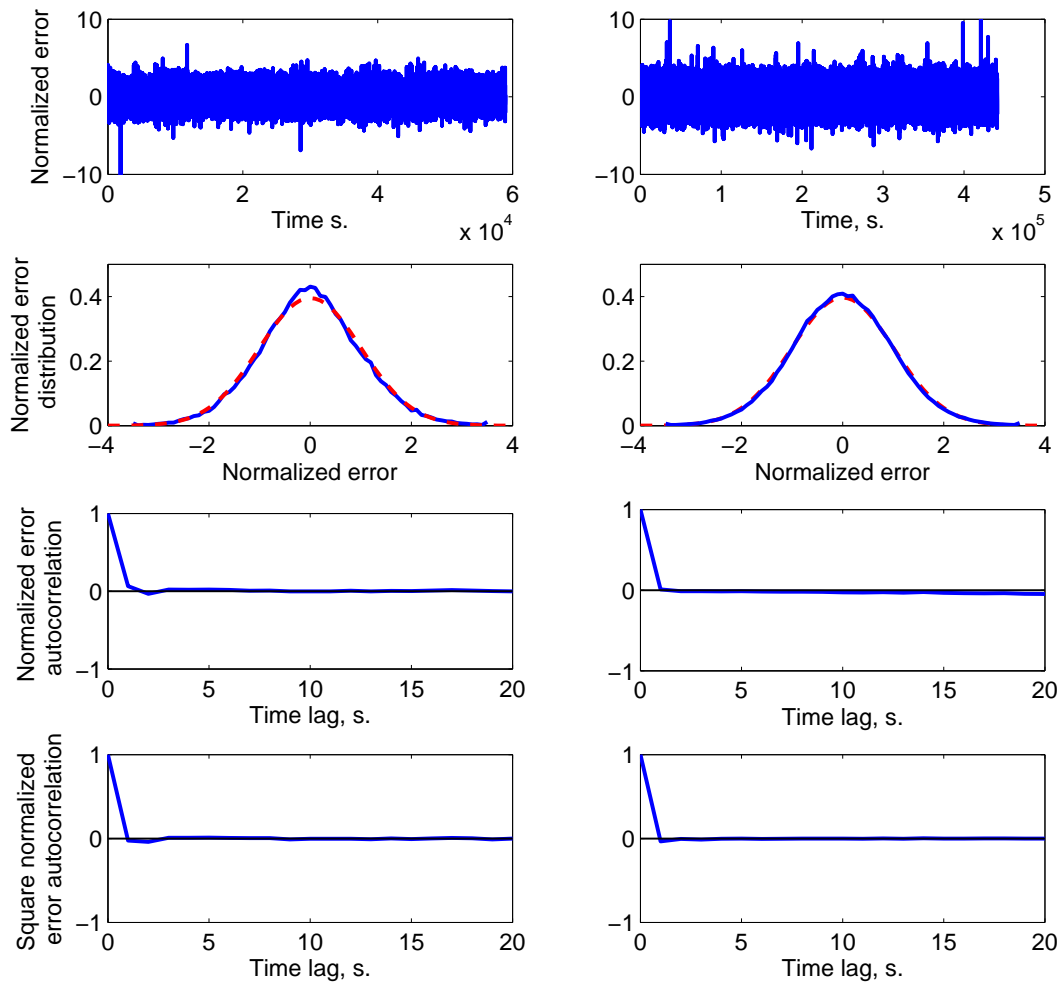


Figure 5. Statistical properties of the normalized prediction errors estimated according to GARCH (from top to bottom: error time series, error distributions compared with Gaussians, error autocorrelations, square error autocorrelations; from left to right: ‘volatile’ periods ($A_t \geq T$), ‘smooth’ periods ($A_t < T$)).

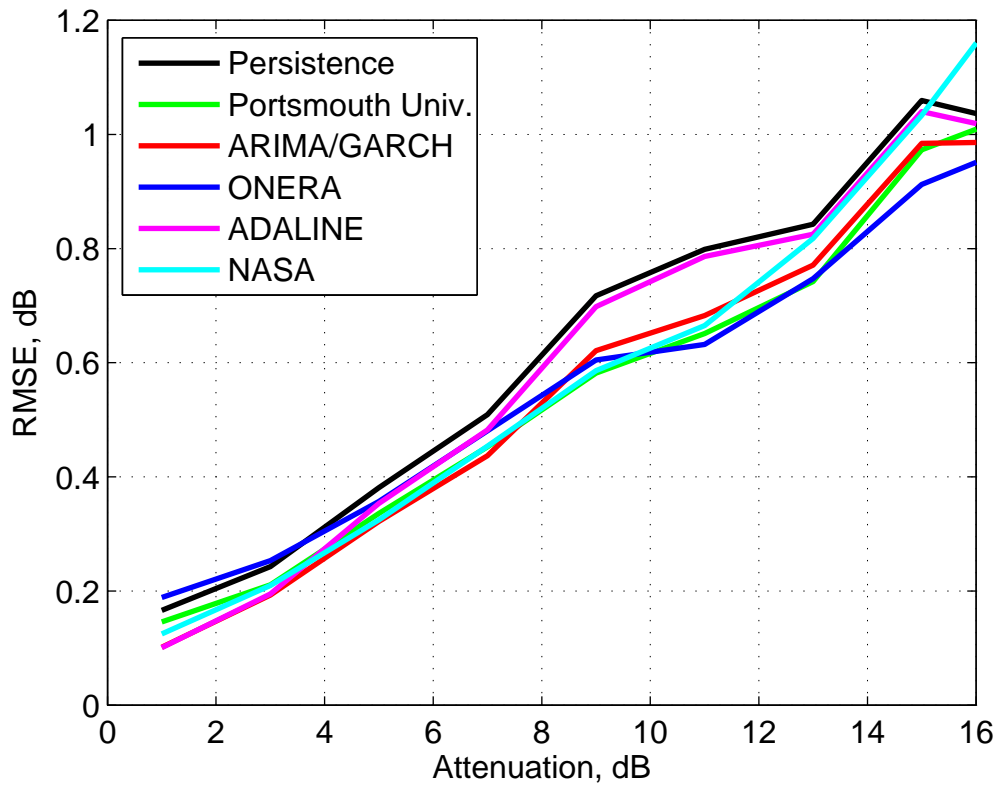


Figure 6. RMSE of the prediction models according to the attenuation level computed from the 20 GHz Olympus data.

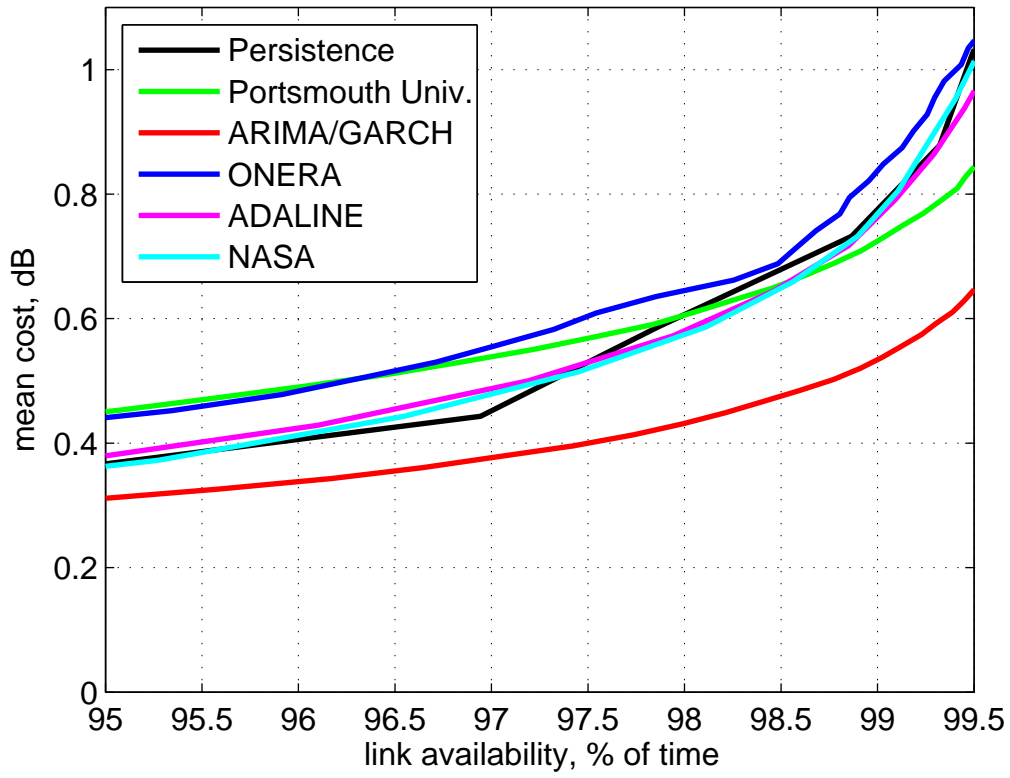


Figure 7. Cost/availability performance of the prediction models computed from the 20 GHz Olympus data.

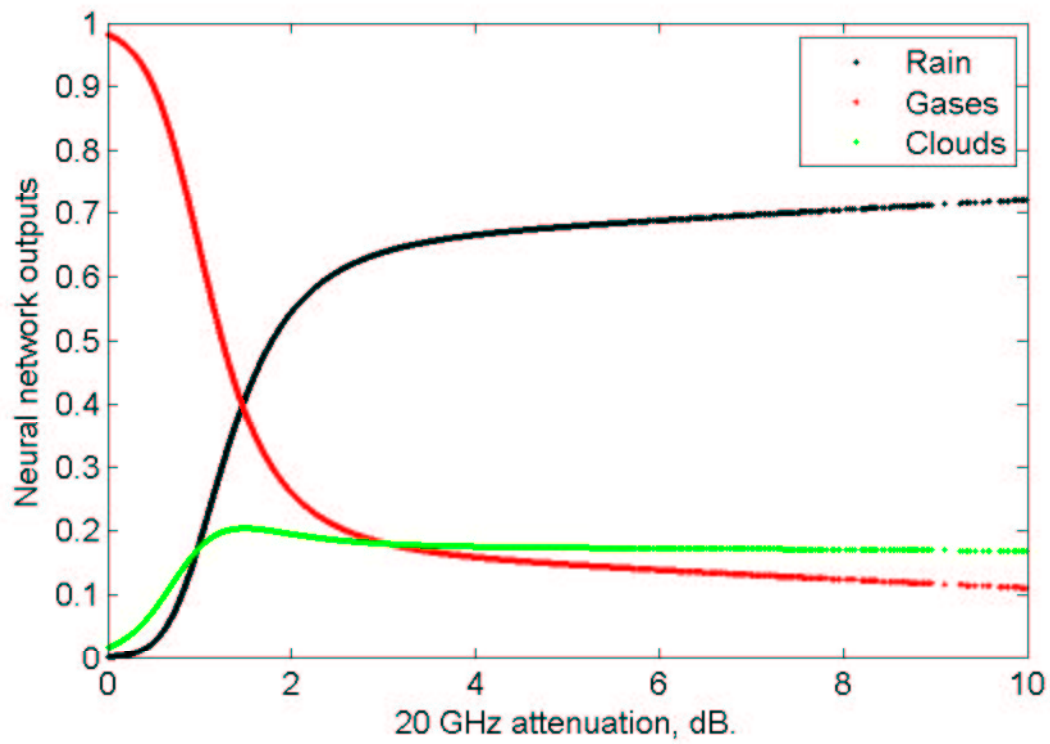


Figure 8. Normalized contribution of gases, clouds and rain to attenuation estimated by the neural network for Olympus-Gometz-la-Ville link (Figure published in Barthes et al. (2006)).

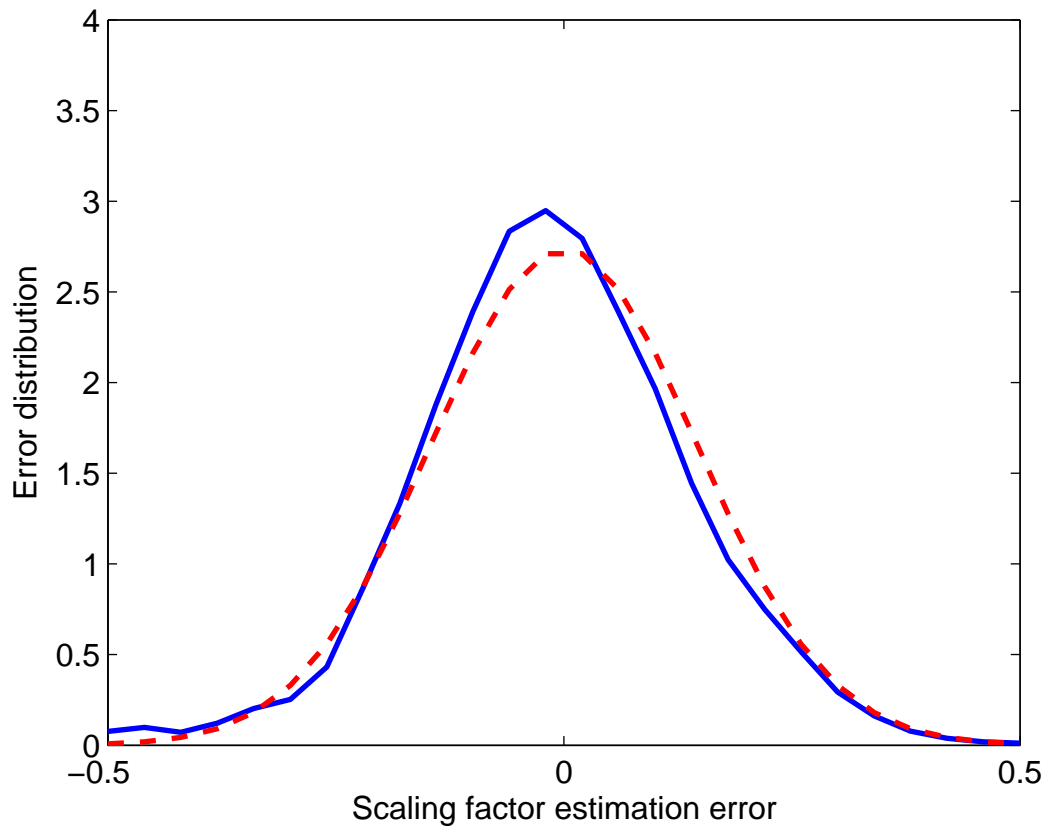


Figure 9. Distribution of the frequency scaling factor error compared with a Gaussian distribution.

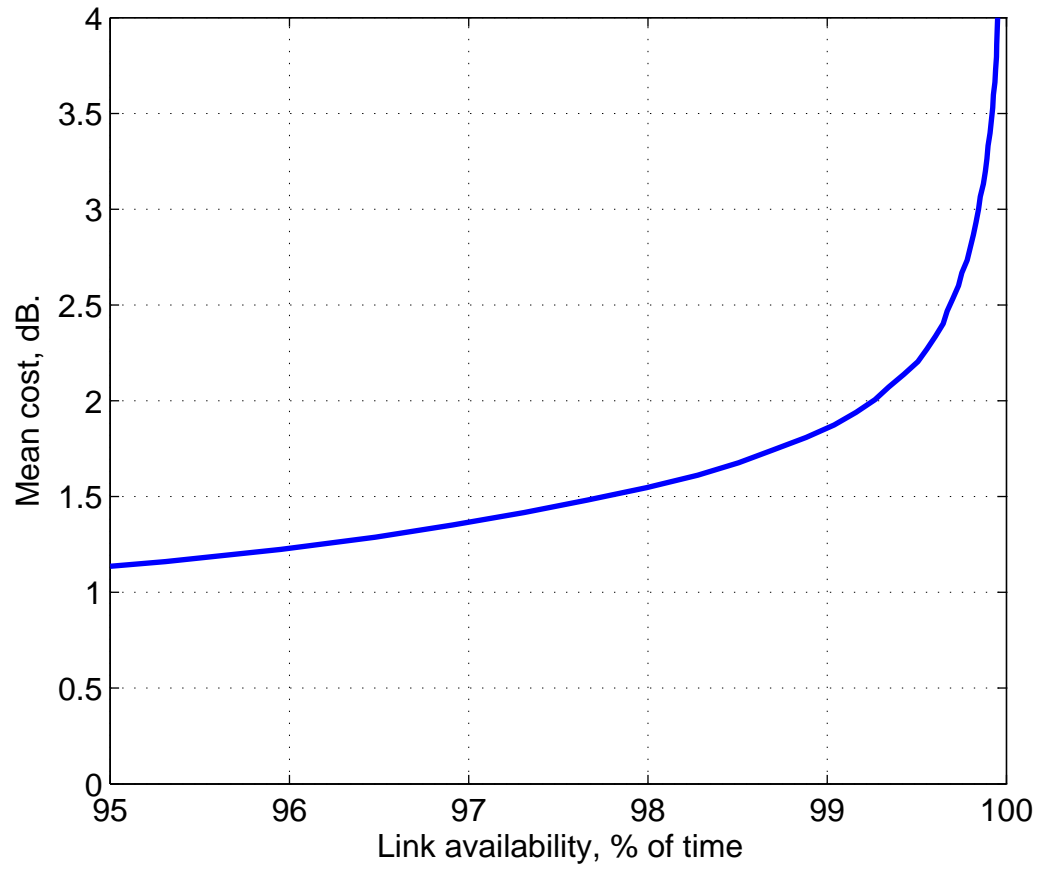


Figure 10. Cost/availability performance of the switching ARIMA/GARCH model including frequency scaling computed from the 20/30 GHz Olympus data.

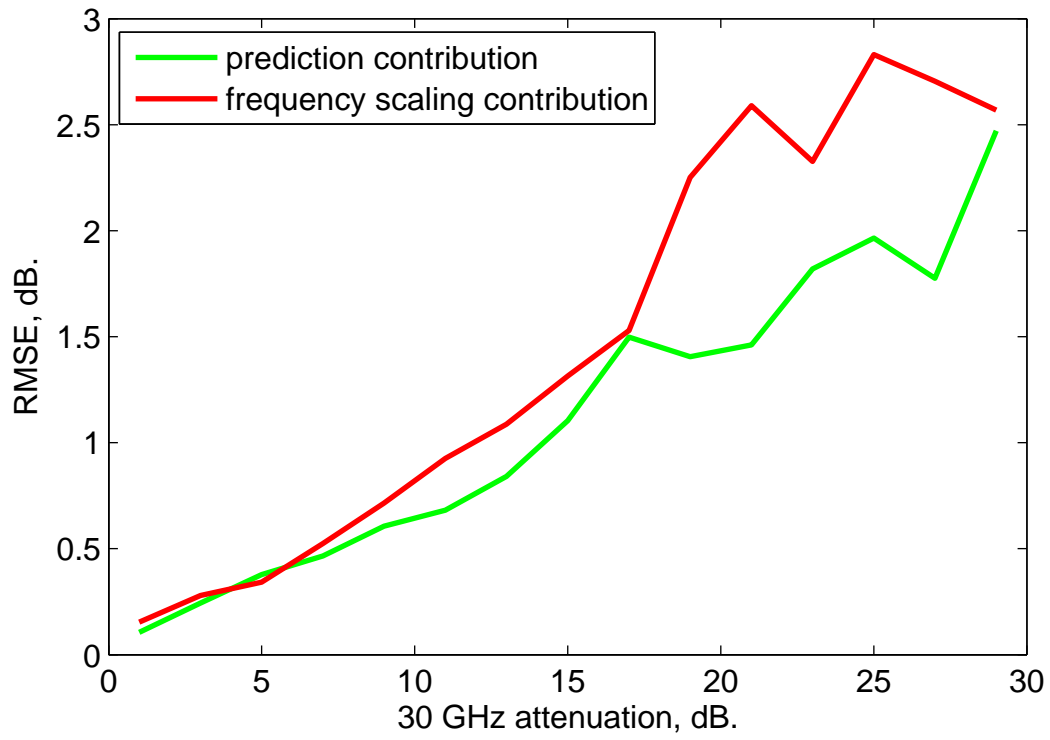


Figure 11. Contributions of the frequency scaling error and of the downlink prediction error to the uplink attenuation prediction error.