

Log-Domain Statistical Modeling of S-Band Rain Clutter for Surveillance Radar Applications

Neuton Severo, Renato Machado, Leandro Pralon, and Bruno Pompeo

Abstract—Rain clutter degrades surveillance radar performance by masking weak targets and increasing false alarms. Classical clutter models are formulated in the linear-amplitude domain, yet the large dynamic range of rain returns motivates a logarithmic transformation for more tractable statistical characterization. This paper analyzes the statistical characteristics of log-transformed rain clutter using 372 samples from an S-band surveillance radar dataset. Six candidate distributions (Log-Gamma, Log-Nakagami, Log-Rice, Log-Rayleigh, Log-Weibull, and Log-K) are fitted via maximum likelihood estimation and compared using the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the Kolmogorov–Smirnov and Cramér–von Mises criteria. The results support log-domain modeling for improved CFAR detection and motivate the development of a compound distribution to provide more precise modeling for S-band rain clutter, an approach that can be further extrapolated to other frequency bands.

Keywords—Rain clutter, S-band surveillance radar, Log-domain analysis, Goodness-of-fit, CFAR detection, Model selection

I. INTRODUCTION

Radar systems are widely employed for surveillance and target detection because they can operate in the absence of visible light. Nevertheless, heavy precipitation can severely compromise radar performance: rain returns may both trigger false detections and attenuate echoes from actual targets, posing a significant challenge for robust target detection [1], particularly in S-band systems.

In radar terminology, any undesired detection is referred to as clutter. Unlike stationary clutter (caused by fixed structures such as buildings or terrain features, which can be effectively suppressed through Moving Target Indication (MTI) or Doppler filtering), non-stationary clutter poses a considerably more difficult challenge [2]. Rain clutter, in particular, is one of the most disruptive non-stationary sources, as it can arise at any location and time, and its characteristics are governed by a complex interplay of atmospheric, geographic, and electromagnetic factors. Precipitation remains a highly variable and unpredictable natural phenomenon because it is driven by environmental factors like temperature, humidity, wind speed, and raindrop size distribution. The way we observe this event,

however, depends on radar design choices, specifically the frequency and polarization of the transmitted signal, which dictate how the waves interact with the rain [2].

Effective rain clutter suppression in surveillance radars has been the subject of extensive research, evolving from classical statistical modeling to modern machine learning and model-independent frameworks. While traditional Gaussian models are often insufficient for describing heavy precipitation, recent studies have demonstrated the effectiveness of the Nakagami- m distribution for broad weather clutter characterization [3], and the Weibull distribution for capturing the high-dynamic range of returns in stormy conditions [1]. Furthermore, for high-resolution surveillance, compound-Gaussian models (CGM) using gamma or log-normal textures have been proposed to adapt to varying rainfall rates [4]. Despite these advancements, the performance of parametric methods tends to degrade when the assumed model does not accurately reflect the actual clutter statistics. To mitigate this, recent works have explored model-independent alternatives, such as filtering based on the Spectral Flatness Measure of the received signal [5] and the integration of Convolutional Neural Networks (CNNs), which avoids explicit distributional assumptions altogether [6].

Regardless of the suppression strategy employed, accurate statistical characterization of rain clutter is a prerequisite for the design of robust detection algorithms, particularly those based on Constant False Alarm Rate (CFAR) frameworks [1]. The amplitude distribution of rain clutter returns is inherently right-skewed, complicating direct parametric fitting, as illustrated in Fig. 1 (left), where the candidate distributions fail to capture the empirical data across the entire support accurately. A logarithmic transformation compresses the dynamic range and yields a more symmetric distribution, substantially improving the parametric fits, as shown in Fig. 1 (right).

In this context, the present work investigates the statistical behavior of log-transformed rain clutter returns collected by the SABER M200 VIGILANTE S-band surveillance radar, depicted in Fig. 2. Six candidate distributions (Gamma, Nakagami- m , Rice, Rayleigh, Weibull, and K) are fitted to the transformed data and compared using AIC, BIC, Kolmogorov–Smirnov, and Cramér–von Mises criteria across 372 independent samples.

The remainder of this paper is organized as follows. Section II describes the data acquisition and methodology, including the log-transformation framework and the evaluated distributions. Section III presents the experimental results. Section IV concludes the work and outlines directions for future research.

Neuton Severo and Renato Machado are with the Warfare Competence Centre, Aeronautics Institute of Technology, São José dos Campos, SP, Brazil. Leandro Pralon and Bruno Pompeo are with the Brazilian Army Technology Center. E-mail: rmachado@ita.br, neuton.neto.102560@ga.ita.br, pralon.leandro@eb.mil.br, pompeo.bruno@eb.mil.br. This work was partly supported by the Brazilian National Council for Scientific and Technological Development (CNPq), the Ministry of Science, Technology and Innovation (MCTI), and the São Paulo Research Foundation (FAPESP) under Grant 2020/09838-0 (BIOS—Brazilian Institute of Data Science).

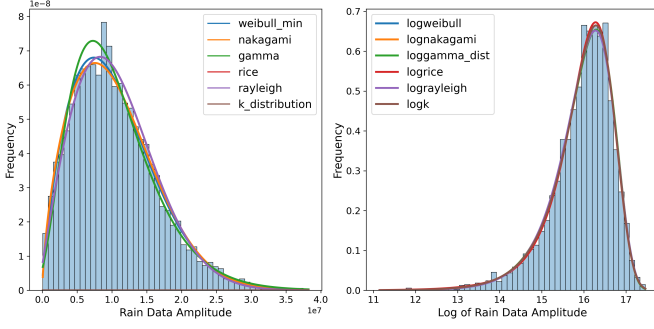


Fig. 1. Fitted distributions overlaid on rain clutter histograms.



Fig. 2. SABER M200 VIGILANTE (reproduced from [5]).

II. DATA ACQUISITION AND METHODOLOGY

The 372 samples were collected during a single measurement campaign in Campinas, São Paulo, Brazil, using the SABER M200 VIGILANTE S-band surveillance radar (Fig. 2), whose phased-array antenna electronically steers the beam across four azimuthal positions per scan. Each sample comprises the range cells identified as containing precipitation returns for a given beam position and scan, with sample sizes ranging from 38 to 7904 (median $n = 2812$, mean $n = 3057$). The variation in sample size reflects the spatial distribution of precipitation across the four beam directions: positions intercepting the core of the rain cell yield more returns, while those grazing the periphery yield fewer.

To model the statistical behavior of heavy rain amplitude logarithms, six continuous probability distributions were evaluated: the Weibull, Nakagami- m , K , Gamma, Rice, and Rayleigh. The Weibull and Nakagami- m distributions have been employed to model weather clutter amplitude in several studies [3], [7], [8]. The K -distribution and Gamma distribution have also been applied to radar clutter modeling, particularly in scenarios involving compound scattering mechanisms [9]. The Rice and Rayleigh distributions are fundamental models in radar signal processing: the Rayleigh distribution describes the envelope of thermal noise or clutter composed of many independent scatterers of comparable strength, whereas the Rice distribution extends this model to scenarios where a dominant deterministic component [2].

A. Log-Domain Transformation

Let X be a non-negative continuous random variable with PDF $f_X(x)$, $x \in (0, \infty)$. The log-amplitude variable $Y = \ln(X)$ has support on $(-\infty, \infty)$ and, by the standard change-of-variables formula [10], its PDF is given by

$$f_Y(y) = f_X(e^y) e^y, \quad y \in \mathbb{R}, \quad (1)$$

since the inverse mapping $X = e^Y$ has Jacobian $|dx/dy| = e^y > 0$. Equation (1) is the only transformation required. The six log-domain PDFs listed below are obtained by direct substitution of each parent PDF f_X .

B. Candidate Distributions

The six parent distributions and their physical motivations are as follows.

a) *Gamma distribution*: With shape $\alpha > 0$ and scale $\beta > 0$,

$$f_X(x; \alpha, \beta) = \frac{x^{\alpha-1}}{\Gamma(\alpha) \beta^\alpha} \exp\left(-\frac{x}{\beta}\right), \quad x \geq 0, \quad (2)$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ is the Gamma function. The Gamma family is widely used in compound-scattering models.

b) *Nakagami- m distribution*: With fading parameter $m \geq \frac{1}{2}$ and spread $\Omega > 0$,

$$f_X(x; m, \Omega) = \frac{2 m^m}{\Gamma(m) \Omega^m} x^{2m-1} \exp\left(-\frac{m}{\Omega} x^2\right), \quad x \geq 0. \quad (3)$$

The Nakagami- m distribution models fading channels with varying degrees of severity and has been applied to the modeling of weather clutter amplitude [3], [7].

c) *Rice distribution*: With line-of-sight parameter $\nu \geq 0$ and diffuse scale $\sigma > 0$,

$$f_X(x; \nu, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \nu^2}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right), \quad x \geq 0, \quad (4)$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order zero. The Rice distribution arises when exists a dominant deterministic component.

d) *Rayleigh distribution*: With scale $\sigma > 0$,

$$f_X(x; \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \geq 0. \quad (5)$$

The Rayleigh distribution is the envelope of two i.i.d. zero-mean Gaussian components. It is a special case of both the Rice ($\nu = 0$) and Nakagami- m ($m = 1$) families, and is included to test whether the additional parameters of those families are statistically justified.

e) *Weibull distribution*: With shape $k > 0$ and scale $\lambda > 0$,

$$f_X(x; k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{x}{\lambda}\right)^k\right], \quad x \geq 0. \quad (6)$$

The Weibull distribution offers considerable shape flexibility and has been employed in several weather-clutter studies [7], [8].

f) *K-distribution*.: With shape $\nu > 0$ and scale $\alpha > 0$,

$$f_X(x; \nu, \alpha) = \frac{2\alpha}{\Gamma(\nu)} \left(\frac{\alpha x}{2} \right)^\nu K_{\nu-1}(\alpha x), \quad x \geq 0, \quad (7)$$

where $K_{\nu-1}(\cdot)$ is the modified Bessel function of the second kind of order $\nu - 1$.

Table I collects the log-domain PDFs obtained by substituting each parent f_X into (1).

TABLE I

LOG-DOMAIN PDFs $f_Y(y)$ OBTAINED VIA EQ. (1). ALL EXPRESSIONS ARE VALID FOR $y \in (-\infty, \infty)$.

Model	$f_Y(y)$
Log-Gamma	$\frac{e^{\alpha y}}{\Gamma(\alpha) \beta^\alpha} \exp\left(-\frac{e^y}{\beta}\right)$
Log-Nakagami	$\frac{2 m^m}{\Gamma(m) \Omega^m} e^{2my} \exp\left(-\frac{m}{\Omega} e^{2y}\right)$
Log-Rice	$\frac{e^{2y}}{\sigma^2} \exp\left(-\frac{e^{2y} + \nu^2}{2\sigma^2}\right) I_0\left(\frac{e^y \nu}{\sigma^2}\right)$
Log-Rayleigh	$\frac{e^{2y}}{\sigma^2} \exp\left(-\frac{e^{2y}}{2\sigma^2}\right)$
Log-Weibull	$\frac{k}{\lambda^k} e^{ky} \exp\left[-\left(\frac{e^y}{\lambda}\right)^k\right]$
Log-K	$\frac{2\alpha^{\nu+1}}{\Gamma(\nu)} \frac{e^{y\nu+1}}{2^\nu} K_{\nu-1}(\alpha e^y)$

Note that the e^y Jacobian factor has been absorbed into the exponential terms where possible, yielding more compact expressions. The Log-Rayleigh PDF is recovered from the Log-Rice PDF by setting $\nu = 0$, preserving the nesting structure of the parent families in the log domain.

C. Parameter Estimation and Model Selection

The parameters of all six distributions were estimated using Maximum Likelihood Estimation (MLE). To evaluate and compare the fitted models, four goodness-of-fit and model selection statistics were computed, where smaller values indicate a better fit.

- **Akaike Information Criterion (AIC)**: Defined as

$$\text{AIC} = -2 \ln(\hat{L}) + 2k. \quad (8)$$

AIC balances goodness-of-fit against model complexity by penalizing the number of estimated parameters k , where $\hat{L}$ is the maximized likelihood of the model. It estimates relative information loss and favors parsimonious models. Selection is based on the relative difference ΔAIC between models: if $\Delta\text{AIC} < 1$, the models are considered nearly equivalent, whereas a difference larger than $1 \sim 2$ indicates a significant statistical preference for the lower-value model [11].

- **Bayesian Information Criterion (BIC)**: Defined as

$$\text{BIC} = k \ln(n) - 2 \ln(\hat{L}). \quad (9)$$

BIC penalizes each parameter by $\ln(n)$ rather than AIC's fixed factor of 2, making it increasingly conservative as the sample size n grows. This penalty arises from

a Laplace approximation to the marginal likelihood, grounding BIC as a large-sample approximation to the Bayes factor [12]. Unlike AIC, BIC is asymptotically consistent though the approximation and requires n to be much larger than k .

- **Kolmogorov–Smirnov (KS) statistic**: Defined as

$$D_n = \sup_x |F_n(x) - F(x)|. \quad (10)$$

The KS statistic measures the maximum absolute distance between the empirical CDF $F_n(x)$ and the theoretical CDF $F(x)$. It is particularly sensitive to deviations near the center of the distribution, where the empirical CDF has the highest density of observations. The null hypothesis of distributional agreement is rejected when D_n exceeds a critical value that depends on the sample size n and the significance level α .

- **Cramér–von Mises (CvM) statistic**: Defined as

$$W^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x). \quad (11)$$

The CvM statistic computes the integrated squared difference between $F_n(x)$ and $F(x)$, aggregating discrepancies across the entire support. Unlike the KS statistic, which captures only the single worst-case deviation, CvM accumulates all departures and thus provides greater sensitivity to systematic deviations in the tails.

AIC and BIC assess model quality through penalized likelihood, while KS and CvM directly evaluate distributional fit.

III. RESULTS

For each of the 372 samples, the six candidate distributions were ranked from 1 (best) to 6 (worst) according to each statistic independently. An overall combined rank was computed as the arithmetic mean of the four individual ranks, as

$$\bar{R}_i = \frac{1}{4} \left(R_i^{(\text{AIC})} + R_i^{(\text{BIC})} + R_i^{(\text{KS})} + R_i^{(\text{CvM})} \right), \quad (12)$$

where $\bar{R}_i$ denotes the mean rank of distribution i for a given sample. The distribution with the lowest $\bar{R}_i$ is selected as optimal for that sample.

This rank-aggregation scheme is adopted for two reasons. First, the four statistics are measured on fundamentally different scales and units: AIC and BIC are penalized log-likelihoods whose magnitude depends on sample size and the number of parameters, whereas KS and CvM are bounded distance measures between CDFs. Averaging raw statistical values would therefore be dominated by whichever criterion has the largest numerical range. Ranks map all four criteria onto a common ordinal scale of 1 to 6, eliminating the need for ad-hoc normalization or unit conversion. Second, equal weighting is chosen as the least assumption-laden aggregation rule: in the absence of a specific downstream application that would justify prioritizing one criterion over another (e.g., favoring parsimony for real-time CFAR implementation, or favoring tail accuracy for low false-alarm design), assigning uniform weights avoids introducing subjective bias into the model selection. We note, however, that the individual per-criterion results are also reported separately in Tables II

and III, allowing the reader to re-weight the criteria according to their own application requirements.

A. Average Ranks

Table II presents the mean rank of each distribution across all 372 samples for each evaluation criterion. A rank of 1 would indicate that a distribution was always the best; 6 would indicate it was always the worst.

TABLE II

AVERAGE RANKS BY EVALUATION STATISTIC (LOWER IS BETTER). BOLD VALUES INDICATE THE BEST-PERFORMING DISTRIBUTION FOR EACH CRITERION.

Distribution	AIC	BIC	CvM	KS	Mean Rank
Log-Gamma	1.13	1.03	2.28	2.47	1.73
Log-Nakagami	2.75	3.07	2.33	2.47	2.65
Log-Rice	3.35	3.61	2.88	2.76	3.15
Log-Rayleigh	2.91	2.42	3.76	3.64	3.18
Log-Weibull	5.05	5.05	3.78	3.68	4.39
Log- K	5.81	5.83	5.97	5.98	5.90

The Log-Gamma distribution achieves the lowest average rank across all four criteria (AIC, BIC, CvM, and KS) and the lowest overall mean rank of 1.73, establishing it as the most consistent performer. The Log-Nakagami distribution ranks second overall with a mean rank of 2.65, followed closely by Log-Rice (3.15) and Log-Rayleigh (3.18). The Log-Weibull and Log- K distributions consistently occupy the lowest positions, with mean ranks of 4.39 and 5.90, respectively.

B. Win Counts

Table III reports the number of times each distribution achieved rank 1 (best) for each statistic across the 372 samples.

TABLE III

NUMBER OF FIRST-PLACE RANKINGS PER STATISTIC AND OVERALL WIN PERCENTAGE.

Distribution	AIC	BIC	CvM	KS	Total	Win %
Log-Gamma	324	362	104	95	885	59.5
Log-Rice	48	10	117	122	297	20.0
Log-Nakagami	0	0	101	97	198	13.3
Log-Weibull	0	0	29	29	58	3.9
Log-Rayleigh	0	0	19	28	47	3.2
Log- K	0	0	2	1	3	0.2

The Log-Gamma distribution dominates the information-theoretic criteria, winning 324 of 372 AIC comparisons (87.1%) and 362 of 372 BIC comparisons (97.3%). Under the empirical goodness-of-fit statistics, the competition is more balanced: the Log-Rice distribution leads in both CvM (117 wins) and KS (122 wins), followed by Log-Gamma (104 and 95, respectively) and Log-Nakagami (101 and 97). Overall, Log-Gamma wins 59.5% of all comparisons, followed by Log-Rice at 20.0% and Log-Nakagami at 13.3%.

TABLE IV

NUMBER OF SAMPLES (OUT OF 372) WHERE EACH DISTRIBUTION ACHIEVES THE LOWEST COMBINED RANK.

Distribution	Samples Selected	Percentage (%)
Log-Gamma	247	66.4
Log-Rice	66	17.7
Log-Rayleigh	30	8.1
Log-Nakagami	29	7.8
Log-Weibull	0	0.0
Log- K	0	0.0

C. Sample-Level Selection

Table IV presents the number of samples in which each distribution was selected as optimal based on the lowest combined mean rank.

The Log-Gamma distribution is selected as the best-fitting model for 247 of 372 samples (66.4%), confirming its dominance observed in previous analyses. The Log-Rice distribution is selected in 66 samples (17.7%), indicating that it provides a competitive fit for a non-negligible subset of the data, particularly those samples where the empirical goodness-of-fit criteria outweigh the information-theoretic advantage of Log-Gamma.

D. Goodness-of-Fit Visualization

To complement the quantitative evaluation, QQ-plots were constructed for a representative sample, as shown in Fig. 3. A QQ-plot is a graphical tool that compares the empirical quantiles of the observed data, sorted in ascending order, against the theoretical quantiles computed from the fitted distribution [13]. If the candidate model adequately describes the data, the points should fall approximately along the 45-degree reference line. Systematic departures from this line reveal specific regions where the model fails to capture the underlying distribution: curved patterns suggest a mismatch in skewness, while deviations in the extremities indicate that the tails of the fitted distribution are heavier or lighter than those of the empirical data.

The Log-Gamma and Log-Nakagami distributions show the closest alignment to the reference line across all quantiles, confirming their strong performance observed in the quantitative analysis. The Log-Rice distribution also exhibits good agreement, with only slight deviations in the lower tail. The Log-Weibull and Log-Rayleigh distributions present systematic departures in the lower quantiles, with scattered points deviating below the reference line in the 12–14 range. The Log- K distribution exhibits the worst fit among the candidates, with substantial scatter in the lower tail, where empirical quantiles fall below the theoretical reference line, and a concave curvature throughout the mid-range quantiles, suggesting that the Log- K model fails to adequately capture the lower-tail behavior of the sample. These visual observations are consistent with the quantitative results in Tables II–IV.

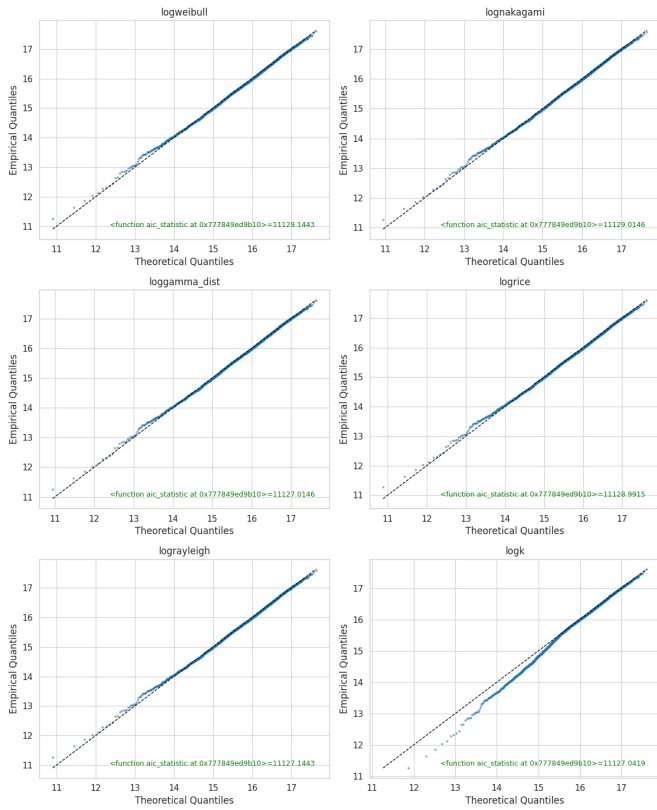


Fig. 3. QQ-plots of a representative log-transformed rain clutter sample for the six candidate distributions.

IV. CONCLUSION

This paper investigated the statistical modeling of rain clutter in S-band surveillance radar using log-transformed amplitude data. Six candidate distributions (Log-Gamma, Log-Nakagami, Log-Rice, Log-Rayleigh, Log-Weibull, and Log- K) were fitted and compared across 372 independent samples using AIC, BIC, KS, and CvM criteria. The results reveal a distinct performance split between the evaluation metrics. The Log-Gamma distribution consistently outperformed the alternatives in terms of parsimony, achieving the lowest overall mean rank of 1.73 and winning 87.1% of AIC and 97.3% of BIC comparisons (see Tables II and III). In contrast, the Log-Rice and Log-Nakagami distributions achieved notable results under pure goodness-of-fit tests, with the former leading the win counts for both CvM (117 wins) and KS (122 wins), suggesting a more precise characterization of the empirical curve when model complexity is not penalized (see Table III). As the dataset originates from a single precipitation event observed at one geographic location, generalization of these findings to different rainfall intensities, climatic regions, and radar frequencies remains to be validated.

In summary, while the Log-Gamma distribution provides the most accurate and parsimonious general-purpose model, the notable performance of the Rice and Nakagami families in empirical tests highlights their strong competitive utility for specific clutter scenarios. Future work will focus on the development of a compound distribution, using a Bayesian

approach analogous to the derivation of the K -distribution, to analyze whether it can capture the influence of these distributions jointly. The proposed distribution will be evaluated against existing models in terms of numerical stability and its impact on the detection probability.

REFERENCES

- [1] S. Sayama, "Detection of Target and Suppression of Sea and Weather Clutter in Stormy Weather by Weibull/CFAR," *IEEE Transactions on Electrical and Electronic Engineering (TEEE)*, vol. 16, no. 2, pp. 180–187, 2021. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/tee.23284>
- [2] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar*. SciTech Publishing, 2010.
- [3] S. Sayama, "Amplitude Statistics of Nakagami-m Distributed Weather Clutter Observed by X-Band Marine Radar," *IEEE Transactions on Electrical and Electronic Engineering*, vol. 16, no. 3, pp. 349–354, 2021. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/tee.23303>
- [4] H.-L. Su, X.-J. Zhang, Y.-H. Li, Y.-H. Wang, P.-J. Zou, and P.-L. Shui, "Rain clutter modelling and performance assessment of road traffic surveillance radars," *IET Radar, Sonar & Navigation*, vol. 19, no. 1, p. e70008, 2025.
- [5] N. Severo, P. Sega, L. Pralon, and B. Pompeo, "Rain clutter mitigation based on the spectral flatness measure in surveillance radar," in *XLIII Brazilian Symposium on Telecommunications and Signal Processing (SBrT 2025)*. SBrT, 2025, p. 5.
- [6] C. Zhan, S. Zhang, C. Sun, and S. Chen, "Anti-rain clutter interference method for millimeter-wave radar based on convolutional neural network," *Remote Sensing*, vol. 16, no. 20, p. 3907, 2024.
- [7] M. Sekine, T. Musha, Y. Tomita, T. Hagiwara, T. Irabu, and E. Kiuchi, "On Weibull-Distributed Weather Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 6, pp. 824–830, 1979.
- [8] M. Olivera and C. Heard, "Increases in the extreme rainfall events: Using the Weibull distribution," *Environmetrics*, vol. 30, p. e2532, 10 2018.
- [9] S. Gupta, I. Varga, M. Hein, and T. Dallmann, "Modelling of rain clutter based on empirical data for automotive radar validation," in *2025 19th European Conference on Antennas and Propagation (EuCAP)*, 2025, pp. 1–5.
- [10] J. P. de Almeida e Albuquerque, J. M. P. Fortes, and W. A. Finamore, *Probabilidade, Variáveis Aleatórias e Processos Estocásticos*, 2nd ed. Rio de Janeiro: Interciência, 2008.
- [11] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics*. Springer Netherlands, 1986. [Online]. Available: <https://books.google.com.br/books?id=KhnvAAAAMAAJ>
- [12] E. Wit, E. v. d. Heuvel, and J.-W. Romeijn, "All models are wrong...: an introduction to model uncertainty," *Statistica Neerlandica*, 2012. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9574.2012.00530.x>
- [13] C. Ford, "Understanding QQ plots," University of Virginia Library, Research Data Services, 2015, accessed: 2026-04-28. [Online]. Available: <https://library.virginia.edu/data/articles/understanding-q-q-plots>