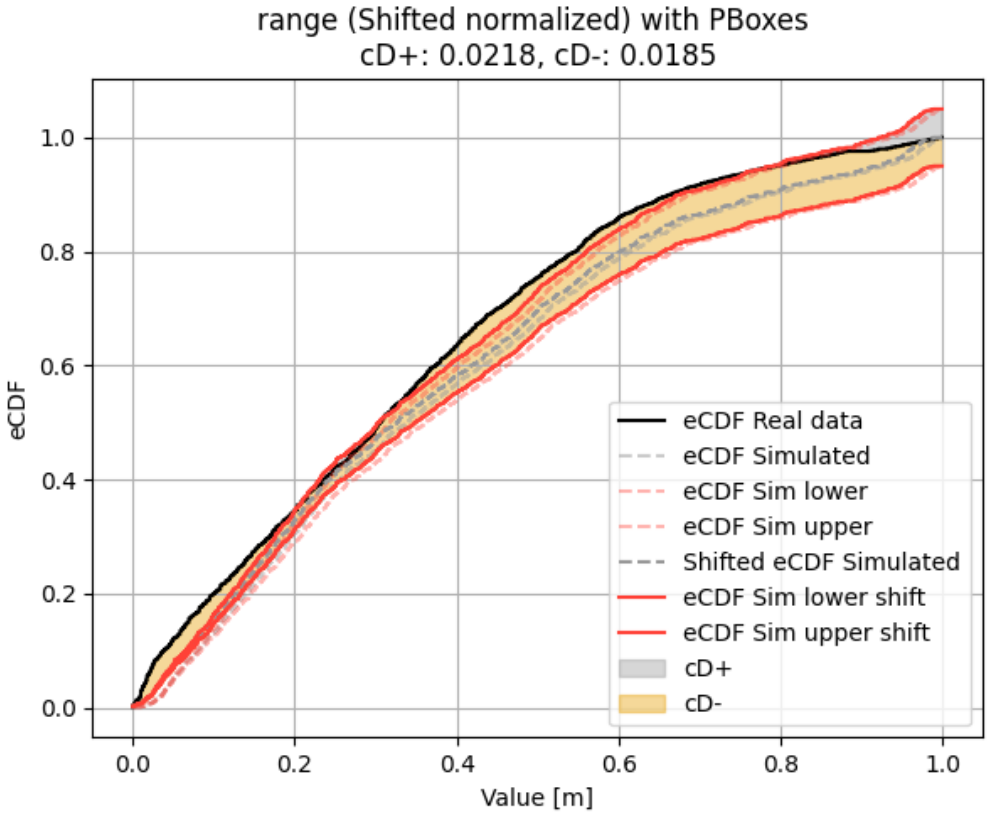

Statistical Evaluation of Radar Simulation Models Towards Real Data

A Method Implementation and Statistical Comparison

Master thesis in Complex Adaptive Systems and Engineering Mathematics & Computational Science

GUSTAV WIKMAN
WILLIAM NILSEN

DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Statistical Evaluation of Radar Simulation Models Towards Real Data

A Method Implementation and Statistical Comparison

GUSTAV WIKMAN
WILLIAM NILSEN

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Mathematical Sciences
Chalmers University of Technology
Gothenburg, Sweden 2025

Statistical Evaluation of Radar Simulation Models Towards Real Data
A Method Implementation and Statistical Comparison

GUSTAV WIKMAN
WILLIAM NILSEN

Supervisor: Sara Tell, Volvo Cars
Examiner: Serik Sagitov, Applied Mathematics and Statistics, Mathematical Sciences

Master's Thesis 2025
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: The figure shows the Empirical Cumulative Distribution Function of real data, shifted simulated data, and simulated data with corresponding positive and negative difference.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2025

Statistical Evaluation of Radar Simulation Models Towards Real Data
A Method Implementation and Statistical Comparison
GUSTAV WIKMAN
WILLIAM NILSEN
Department of Mathematical Sciences
Chalmers University of Technology

Abstract

This thesis aimed to create a method that could be used by Volvo Car Corporation (VCC) to statistically evaluate the truthfulness and accuracy of the built-in radar models used by VCC. Using pre-collected data from European New Car Assessment Programme (Euro NCAP) scenarios—specifically Car-to-Car Rear moving (CCRm) and Car-to-Bicyclist Nearside Adult Obstructed (CBNAO) — a method was developed to extract real world data, enabling the recreation of actual scenarios within a simulated environment. Furthermore, a comparison between the simulated and real data was conducted. This comparison was conducted using the statistical metric Double Validation Metric (DVM), which is a combination of the Area Validation Metric (AVM), Corrected Area Validation Metric (CAVM), and the model bias within the simulated data.

The developed method includes parsing of Hierarchical Data Format version 5 (HDF5) files, enabling reading and manipulating these files. It also features a Graphical User Interface (GUI) that reads one real file and one simulated file, visualizes the desired detection columns within a given timestamp interval, and a statistical evaluator that performs all the calculations, plots the Empirical Cumulative Distribution Function (eCDF), and analyzes their statistical characteristics.

The results demonstrated tendencies pointing towards both reliability and unreliability. Different parameters showed varying degrees of correlation for different scenarios, runs, and speeds. However, one clear trend was the high tendency for detection differences between the simulated model and the real model, with the simulated model outnumbering the real model. Based on this, the developed method can be used to some extent to relay information about whether the radar model is trustworthy. However, its effectiveness heavily depends on the specific area of usage.

In conclusion, the method effectively visualizes the detections, indicating where they occur and how they differ in number. While the model may not be suitable for final certification, it can be a valuable tool for statistical approximation of the radar model's performance.

Keywords: Statistical Evaluation, Statistical Metrics, Real Data, Simulated Data, Double Validation Metric, Radar Models, Euro NCAP scenarios.

Acknowledgements

We want to start by thanking our supervisor Sara Tell at Volvo Cars for all her guidance and support throughout this thesis. Without her help this work would not have been possible. We also want to thank William Millqvist for providing us with all logged data and appurtenant ground truth. Our thanks also goes to our examiner Professor Serik Sagitov at Chalmers University of Technology.

We would also like to take this moment to thank all of our family and friends for the enormous support and joy you have given us during the spring and through all the years.

Gustav Wikman and William Nilsen, Gothenburg, June 2025

Acronyms

AD&ADAS	Advanced Driver Assistance & Autonomous Driving Systems.
AEB	Automatic Emergency Breaking System.
AIC	Akaike Information Criterion.
APSS	Active Perception Sensor Methods.
AV	Autonomous Vehicle.
AVM	Area Validation Metric.
CAVM	Corrected Area Validation Metric.
CBNAO	Car-to-Bicyclist Nearside Adult Obstructed.
CCRm	Car-to-Car Rear moving.
CDF	Cumulative Distribution Function.
DVM	Double Validation Metric.
eCDF	Empirical Cumulative Distribution Function.
Euro NCAP	European New Car Assessment Programme.
FFT	Fast Fourier Transform.
FMCW	Frequency modulated Continous Wave.
FR	Front Radar.
GUI	Graphical User Interface.
GVT	Global Vehicle Target.
HDF5	Hierarchical Data Format version 5.
IF	Intermediate Frequency.
MLE	Maximum Likelihood Estimation.
NATM	New Assessment/Test Method for Automated Driving.
OSI	Open Simulation Interface.

P-Box	Probability Box.
PDF	Probability Density Function.
RCS	Radar Cross Section.
RWS	Real-World Scenarios.
RX antenna	Receiving Antenna.
SNR	Signal-to-noise ratio.
SRM	Simulated Radar Models.
TX antenna	Transmitting Antenna.
UNECE	United Nations Economic Commission for Europe.
VCC	Volvo Car Corporation.
VRU	Vulnerable Road User.
VUT	Vehicle Under Test.

Nomenclature

Below is the nomenclature of parameters and variables that have been used throughout this thesis.

Symbol	Description
bias	Bias

Parameters

Symbol	Unit	Description
ϕ	rad	Mounting angle
p_{lat}	m	Latitudinal position of the radar
p_{long}	m	Longitudinal position of the radar
c_0	$\frac{\text{m}}{\text{s}}$	Speed of light

Variables

Symbol	Unit	Description
α	rad	Azimuth angle
r	m	Radar range
ϵ	rad	Elevation angle
v_{rel}	$\frac{m}{s}$	Relative velocity
t	ms	Index for time
rcs	dBsm	Radar Cross Section
snr	dB	Signal-to-Noise ratio
box_h	m	Bounding box height
box_w	m	Bounding box width
box_l	m	Bounding box length
x	m	Longitudinal position
y	m	Latitudinal position
D+	-	Positive difference
cD+	-	Corrected positive difference
D-	-	Negative difference
cD-	-	Corrected negative difference
Δt	s	Round-trip delay
Δf	Hz	Measured frequency difference
$\frac{d(f)}{d(t)}$	-	Frequency shift per unit of time
R	m	Distance between antenna & reflecting object
F	-	Real data
$\widetilde{F}$	-	Simulated data
ζ	t	Currant time during a certain period
M	-	Number of data points

Contents

List of Figures

List of Tables

1
Introduction

The development of autonomous vehicles has been a main focus within the car industry for a long time, starting its journey during the 16^{th} century with Leonardo da Vinci's three-wheeled, self-propelled cart [1]. Throughout the years, the development has constantly been thriving, trying to find new designs, and improving and expanding safety features. To ensure that the more intelligent and complex cars being launched for line production aren't unsafe, certain demands have to be fulfilled, such as the work presented by United Nations Economic Commission for Europe (UNECE): *"New Assessment/Test Method for Automated Driving (NATM) – a framework, which introduces a multi-pillar approach for safety validation of automated driving"*, meaning they must pass certain test methods in both the physical and simulated world [2]. These validations apply to various autonomous driving features, including Active Perception Sensor Methods (APSS), such as radar systems. The APSS focused on in this project is the so-called Frequency modulated Continous Wave (FMCW) system [3].

Simulating different driving scenarios for developing Autonomous Vehicle (AV) is cost-effective and time-efficient [4]. By simulating real-life scenarios, such as those used by European New Car Assessment Programme (Euro NCAP), the likelihood that physical testing proceeds as expected increases, reducing the need to test every possible case to find minor faults rigorously. This approach allows for a greater number of scenarios to be evaluated efficiently. Thus, gaining a better understanding of the software within the radar and how the radar data affects different Advanced Driver Assistance & Autonomous Driving Systems (AD&ADAS) functions in the car. Additionally, using already collected data from different scenarios to simulate new, untried scenarios can save even more time and effort. However, this can only be trusted if it is known that the simulated radar systems align with the real world.

A method for determining how well the Simulated Radar Models (SRM) represents the Real-World Scenarios (RWS) needs to be developed to get this kind of insight. Radar sensors are crucial for AVs as they provide essential data for detecting and tracking objects, even in challenging environmental conditions such as rain, snow, fog, and dust. FMCW radar systems are particularly significant due to their ability to measure objects' distance, angle, and speed accurately. Automotive radars operate at millimeter wave (mmWave) frequencies, typically around 77 GHz, and are integral to the vehicle's perception system, helping it navigate safely by detecting obstacles and other cars [5].

By integrating advanced radar sensors, robust simulation models, and comprehensive safety validation frameworks, the development of autonomous vehicles continues to progress towards creating safer and more reliable self-driving cars.

1.1 Aim

The primary objective of this thesis is to develop a method to evaluate how representative the SRM used by Volvo Cars is towards RWS. This will be based on statistical metrics indicating the difference between the two datasets.

The intended outcome is to provide Volvo Cars with a satisfactory method to detect potential deviations and errors within the SRM, ensuring the reliability of their radar detection models.

1.2 Scope and Limitations

The work is carried out solely on scenarios defined by Euro NCAP influenced by UNECE, as they provide standards for which scenarios should be simulated and tested for radar applications to an acceptable level of safety. Additionally, Volvo Cars are already subject to these scenarios; hence, no new data collection has been done for this project.

Only pre-collected data from Volvo Car Corporation (VCC) is considered, and the investigation is limited to the Front Radar (FR). However, the results from this project should be replicable on all the radars fitted to Volvo Cars vehicles.

The supervisor will conduct simulations to save time and align with the thesis, which focuses on evaluating data from simulations rather than the simulations themselves.

1.3 Outline of thesis

The thesis contains six chapters. After Chapter 1, the introduction to the thesis, Chapter 2 contains all the necessary theory to understand the content mentioned and discussed within the thesis. To maintain a good structure, this chapter is split into six sections, each including several subsections. The first section mentions relevant facts about FMCW radars. The second section is about probability theory, and the third is about statistical metrics, explaining the foundational concepts for the evaluation. The fourth section discusses how the data is structured. The fifth section explains what a simulation is and how it is used throughout the thesis, while the sixth and final section covers the scenarios and their usage.

Chapter 3 contains the methodology and explains how all steps, from the beginning to the end, were conducted and why.

Next is Chapter 4, which displays all the results gathered during the final tests. These results are then discussed in the upcoming chapter, Chapter 5.

The final chapter is Chapter 6, which contains the conclusions that can be drawn from the thesis, along with some future advancements.

1.4 Disclaimer

Throughout the work of this project AI-tools; Microsoft Copilot, ChatGPT and Grammarly, has been used as help when reviewing text and code. The text and code have been cross-reviewed by the authors.

2
Theory

The following chapter contains all the necessary theory and concepts to understand this thesis. First, in section 2.1, the fundamentals of an FMCW radar will be presented, along with its associated parameters. Secondly, in section 2.2, information about the basics of probability used in the thesis, and how the Empirical Cumulative Distribution Function (eCDF) can be used to prime radar data. Thirdly, in section 2.3 statistical errors and uncertainties are introduced. Then in section 2.4, each step of the chosen metric, along with its theoretical background, is explained. Lastly, section 2.5 introduces the definition of scenarios.

2.1 Frequancy Modulated Continuous Wave Radars

FMCW radars are used in various areas, such as naval tactical navigation radars, smart ammunition sensors, and automotive radars [6]. These radars can measure both the distance and velocity of moving and stationary objects. This is achieved by continuously varying the transmitted signal's frequency with a modulating signal at a known rate over a fixed period. This is also called a *chirp* [7]. The modulation techniques used in these radars include sawtooth modulation, triangle modulation, step modulation, square wave modulation, and sine wave modulation, with sawtooth and triangular being the most common [8]. The signal used in this thesis is a linear chirp, commonly employed in FMCW radars [7]. Each radar frame updates approximately every 50 ms, with individual chirps transmitted at microsecond intervals within each frame.

2.1.1 Parameters

Within a typical automotive FMCW radar, various parameters can be calculated based on the FMCW signal gathered from the radar when in use. These parameters can then be utilized to calculate the location of objects. The following subsubsections (2.1.1.1 through 2.1.1.6) will dive into the theory of these parameters in greater detail.

2.1.1.1 Azimuth angle

The horizontal direction is expressed as the angular distance between the direction of the fixed point, front looking, and the direction of the object [9]. To calculate this angle the radar needs to have one Transmitting Antenna (TX antenna) and at least two Receiving Antenna (RX antenna). The TX antenna emits a chirp that reflects of an object and then being received by the RX antennas. Due to differences in distance, the retrieving signal may have to travel a little longer to one of the RX antenna, Δd. With this Δd, the phase difference can be calculated with Equation 2.1 [10]. This case is displayed in Figure 2.1.

$$\omega = \frac{2\pi\Delta d}{\lambda} \tag{2.1}$$

where λ is the wavelength and ω is the phase difference.

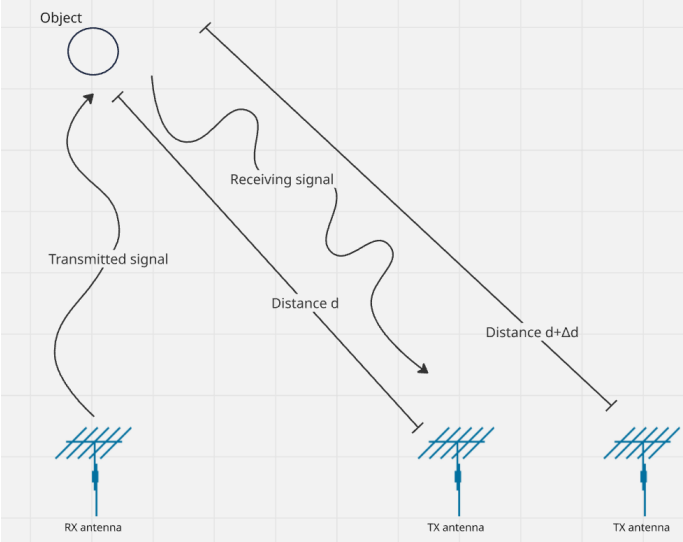

Figure 2.1: Showing how the RX antenna and TX antenna are places and their corresponding difference in distance d and d+Δd.

To easen the calculations, the receiving signal getting noted by the two RX antennas are assumed to be parallel, see Figure 2.2. Thus, the azimuth angle is calculated by Equation 2.2 [10].

$$\theta = \sin^{-1}(\frac{\lambda\omega}{2\pi \mathrm{d}}) \tag{2.2}$$

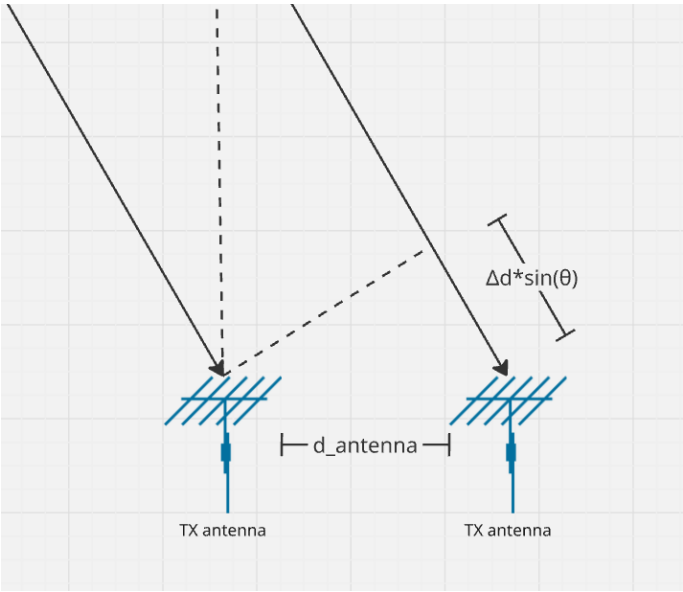

Figure 2.2: Displays the RX antennas and the receiving signals. θ represents the azimuth angle, $\mathrm{d}_{\mathrm{antenna}}$ is the distance between the two RX antennas, and Δd is the difference in distance for the receiving signal between the two RX antennas.

A clarifying image illustrating what the azimuth angle is can be seen below in Figure 2.3.

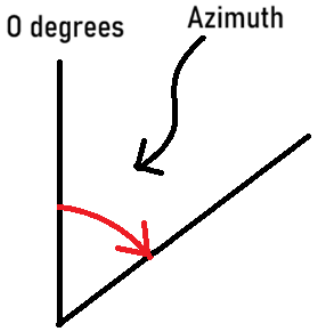

Figure 2.3: The 0°axis represents the forward-facing surface of the FMCW radar, with the azimuth angle (the red line) representing how far from the front face the radar will collect data. This interval is mirrored.

2.1.1.2 Angular resolution

This is the ability of a radar to distinguish between different objects located at different bearings[11, 12]. Having a low resolution radar, two close objects would be

recognized as one larger object [13].

By adding more RX antennas, the angular resolution increases. This is because a larger number of antennas in an array allows for finer spatial sampling of the incoming wavefronts, which enhances the array's ability to distinguish between signals arriving from closely spaced directions [13].

2.1.1.3 Range

For a radar to detect the range for different objects, the radar undergoes the following procedure:

1. The radar generates a chirp with the help of a synthesizer [14].
2. The RX antennas note the chirp [14].
3. The Intermediate Frequency (IF), being both the RX and TX signal, is digitized and converted into Analog-to-Digital data [14].
4. An Fast Fourier Transform (FFT) is performed on the Analog-to-Digital data where the location of the peaks in the FFT correspond to the range of the different objects [14].

2.1.1.4 Elevation angle

The elevation angle for a radar is the angle between the horizontal plane and the line of sight, measured in the vertical plane. The reference direction, or an elevation angle of zero degrees, is a horizontal line extending from the radar antenna towards the horizon.

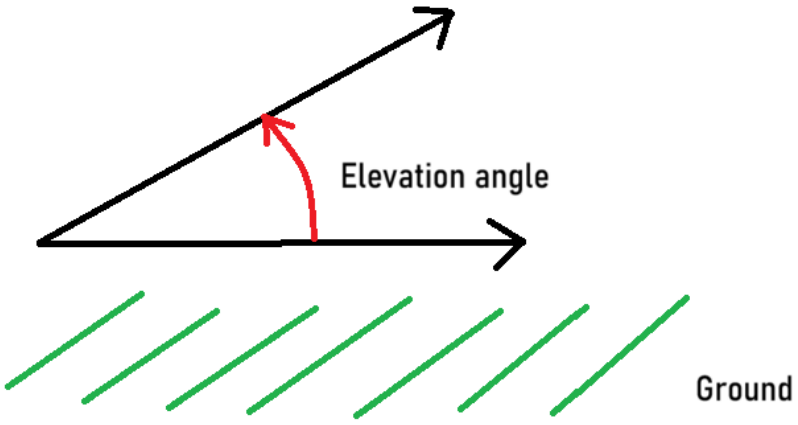

Figure 2.4: The elevation angle for a radar is the angle between the horizontal plane and the line of sight.

2.1.1.5 Relative velocity (Doppler velocity)

Doppler velocity, also known as relative velocity, refers to the measurement of the velocity of an object in the direction of the radar beam, based on the Doppler effect. The Doppler effect, or Doppler shift, is the change in frequency or wavelength of a

wave to an observer moving relative to the wave source. When applied to radar, this effect allows the determination of the speed at which an object moves towards or away from the radar [15].

Here are the characteristics of the Doppler velocity:

Approaching Object: If the object is moving towards the radar, the frequency of the returned signal increases, i.e., doppler is positive.

Receding Object: If the object is moving away from the radar, the frequency of the returned signal decreases, i.e., doppler is negative.

2.1.1.6 Radar Cross Section

Radar Cross Section (RCS) measures how effectively a target reflects and scatters incoming electromagnetic waves from the radar. It represents the equivalent area that would scatter the same amount of radar energy back to the radar receiver as the actual target does. Essentially, a larger RCS indicates that an object is more easily detected by radar [16, 17].

Below are some things that influence the RCS
Material: The composition of the target affects how much radar energy is reflected.

Size: Larger objects generally have a higher RCS, assuming all other factors such as shape, material, and orientation remain the same.

Shape: The geometry and orientation of the target relative to the radar beam can significantly impact the RCS.

Features: Smooth, flat surfaces reflect radar waves differently than rough or irregular surfaces.

Frequency: The radar signal's wavelength can also affect the RCS.

2.1.1.7 Signal-to-noise ratio

Signal-to-noise ratio (SNR) is a measure used in science and engineering to compare the level of a desired signal to the level of background noise. It is defined as the ratio of signal power to noise power. A higher SNR indicates a more precise and more easily detectable signal, while a lower SNR means the signal is more obscured by noise [18, 19].

2.2 Probability theory

Probability Theory is a branch of mathematics that deals with analyzing random phenomena. It provides a framework for quantifying uncertainty and predicting the

likelihood of various outcomes. Probability theory is fundamental to many fields, including statistics, finance, science, engineering, and everyday decision-making [20]. Below this section, the fundamentals of probability theory used throughout the thesis are mentioned and explained.

2.2.1 Cumulative Distribution Function

The Cumulative Distribution Function (CDF) is a fundamental concept in probability theory and statistics. It describes the probability that a real-valued random variable X will take a value less than or equal to a certain threshold x [21]. Mathematically, the CDF of a random variable X is defined as:

$$F_X(x) = P(X \leq x) \tag{2.3}$$

If X is a continuous random variable, the CDF can be written as:

$$F_x(x) = \int_{-\infty}^{x} f_X(t)dt \tag{2.4}$$

where f_X is the probability density function [21]. In cases where the random variable X is discrete, the CDF is expressed as a sum rather than an integral as [21]:

$$F_X(x) = \sum_{x_i \leq x} P(X = x_i) \tag{2.5}$$

For a continuous random variable, the derivative of the CDF (when it exists) yields the probability density function:

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2.6}$$

This relationship forms the basis for transitioning between the Probability Density Function (PDF) and CDF in the context of continuous distributions.

In practical applications, the CDF is instrumental in calculating probabilities over intervals. For example, the probability that X falls within the interval (a,b] is given by:

$$P(a < X \leq b) = F_X(b) - F_X(a) \tag{2.7}$$

2.2.2 Empirical Cumulative Distribution Function

eCDF is a non-parametric way to estimate the CDF of a random variable. The eCDF is a step function, meaning that it jumps up by $\frac{1}{N}$ at each of the observation points N [22, 23]. The mathematical expression of a step function defined over an arbitrary interval A, denoted by the indicator function χ_{A}, identifies whether an element x from the universal set (i.e., the domain under consideration) belongs to the set A [24]. This is formally expressed in Equation 2.8 below.

$$\chi_{\mathrm{A}} = \begin{cases} 1 \text{ if } x \in \mathrm{A} \\ 0 \text{ if } x \notin \mathrm{A} \end{cases} \tag{2.8}$$

The mathematical concept behind the eCDF is defined as follows:
$(x_1, x_2, \cdots, x_N)$ are all independent and identically distributed real random variables from a common CDF, $F(t)$. Furthermore, the eCDF, $\hat{F}$ is as follows:

$$\hat{F} = \frac{\text{number of elements in the sample} \leq t}{N} = \frac{1}{N}\sum_{i=1}^{N} = \mathbf{1}_{x_i \leq t} \tag{2.9}$$

In equation 2.9, $\mathbf{1}_A$ is a characteristic function with the same properties as χ_A in Equation 2.8 [25]. A is the corresponding subset to $\mathbf{1}_A$. $\mathbf{1}_{x_i \leq t}$ is a Bernoulli random variable, meaning that the variable is 1 with a probability of p and 0 with probability $q = 1 - p$ [26]. Therefore, $n\hat{F}$ has a Binomial distribution, meaning it has a mean of $nF(t)$ and variance $nF(t)(1 - F(t))$ and that $\hat{F}_n(t)$ is an unbiased estimator for $F(t)$ [22, 23].

2.2.3 Maximum Likelihood Estimation

The Maximum Likelihood Estimation (MLE) is a method used to estimate the parameters of an assumed probability distribution on a given sample set. This is attained by maximizing the likelihood function 2.10 or the log-likelihood function 2.11 of the dataset on the sample space and the hyperparameters of the assumed probability function [27]. Maximizing Equation 2.10 or 2.11 is equivalent to Equation 2.12. The point estimator in the weight space, all possible parameter values, that maximizes the likelihood function is called the maximum likelihood estimator [27].

Likelihood function:

$$\hat{L}(\theta|x) = \prod_{i}^{n} f(x_i|\theta) \tag{2.10}$$

Log-likelihood function:

$$\ln(\hat{L}(\theta \mid x)) = \sum_{i=1}^{n} \ln(f(x_i \mid \theta)) \tag{2.11}$$

Maximum function:

$$\hat{\theta} = \arg\max L(\theta \mid \mathbf{x}) \tag{2.12}$$

2.2.4 Coefficient of Determination

R-squared, denoted as R^2, is a statistical measure in the interval [0,1] that indicates the proportion of variance in the dependent variable explained by one or more independent variables in a regression model. 1 means that the modeled data has a perfect fit [28, 29].

$$SS_{\text{res}} = \sum_{i}(y_i - f_i)^2 \tag{2.13}$$

where y_i is the actual observed value, f_i is the predicted value, and SS_{res} is the residual sum of squares [29].

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2 \tag{2.14}$$

where y_i is the actual observed value, $\bar{y}$ is the mean of all observations, and SS_{tot} is the total sum of squares [29].

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{2.15}$$

where SS_{tot} is the total sum of squares and SS_{res} is the residual sum of squares [29].

2.2.5 Akaike Information Criterion

The Akaike Information Criterion (AIC) is an estimator of prediction error. Given a predefined set of probability density functions i.e. modeling distributions, the AIC scores the fit of each modeling distribution relative to the others according to equation 2.16 based on the maximized log likelihood estimation, $log(\hat{L})$, and number of parameters, k, in the distribution function, where the lowest AIC score gives the best distribution fit.

$$AIC = 2k - 2\ln(\hat{L}) \tag{2.16}$$

2.2.6 Aleatory Variability

Aleatory variability or uncertainty in data is the natural randomness of a system or a process and the reason why the value of one parameter can vary each time the exact same experiment is repeated. This uncertainty can not be decreased by gathering more data; it can only be observed and accounted for as a natural error or occurrence. [30, 31, 32].

2.2.7 Epistemic Uncertainty

Epistemic uncertainty is the technical and statistical variability within the model. This uncertainty is founded on data limitations and a lack of theoretical knowledge of the method. The epistemic uncertainty can be reduced by gathering more data and expertise about the modeled system or the method used for the model development. [30, 31, 32]. Since this uncertainty may be decreased, in contrast to aleatory variability, it is important to reduce it as much as possible.

2.3 Statistical Error and Uncertainty

Statistical metrics are quantitative measures used to summarize, describe, analyze data, understand data, identify trends, and make evidence-based decisions. They provide insights into the characteristics and patterns within a dataset, helping to make informed decisions based on empirical evidence. These metrics are fundamental in various fields, including economics, psychology, engineering, medicine, and

social sciences, to draw meaningful conclusions and take informed actions.

2.3.1 Probability Box

A Probability Box (P-Box) is defined by a lower CDF, $F_{X_{lower}}(x)$, and an upper CDF, $F_{X_{upper}}(x)$, such that:

$$F_{X_{lower}}(x) \leq F(x) \leq F_{X_{upper}}(x), \quad \forall x \in \mathbb{R} \tag{2.17}$$

Here, $F(x)$ represents the unknown true cumulative distribution function of a random variable X. The pair $[F_{X_{lower}}, F_{X_{upper}}]$ bounds a family of plausible CDFs consistent with the available data. These bounds can be interpreted similarly to a confidence band, enclosing all CDFs that are compatible with the sample data at a given confidence level.

The eCDF, used as the data-driven basis for constructing the P-Box, is defined as above in equation 2.9, section 2.2.2.

Since data sets may vary in size depending on the length of the scenario logs, a consistent method is needed to maintain a uniform uncertainty bound across all samples. To achieve this, the Kolmogorov-Smirnov test metric is used to define upper and lower bounds with a chosen confidence level (e.g., $\alpha = 0.05$). The resulting uniform uncertainty bound is applied across the entire eCDF, producing a distribution-free P-Box.

The lower and upper bounds are calculated as:

$$F_{lower}(x) = \max\left[\hat{F}_n(x) - D_{n,\alpha}, 0\right] \tag{2.18}$$

$$F_{upper}(x) = \min\left[\hat{F}_n(x) + D_{n,\alpha}, 1\right] \tag{2.19}$$

where the margin $D_{n,\alpha}$ is defined by:

$$D_{n,\alpha} = \frac{K_\alpha}{\sqrt{n}} \tag{2.20}$$

and $K_\alpha \approx 1.36$ is the critical value from the Kolmogorov–Smirnov distribution corresponding to a confidence level of $\alpha = 0.05$. The parameter n is the sample size [33].

Through the P-Box framework, epistemic uncertainty is captured by assigning an interval $[F_{X_{lower}}(x), F_{X_{upper}}(x)]$ for each value of $x \in \mathbb{R}$, as shown in Equation 2.17. The ECDF on which this construction is based is given in Equation 2.9. The bounds in Equations 2.18 and 2.19, using the margin defined in Equation 2.20, represent a statistically meaningful interval within which the true CDF is likely to lie.

When high-quality or abundant data is available, the interval between the bounds tends to be narrow, and the P-Box approaches a well-defined distribution. Conversely, with limited or uncertain data, the bounds widen, indicating reduced confidence in the estimated distribution. Without assuming a specific form for the underlying CDF, this construction is considered a distribution-free P-Box. It provides a general and flexible framework for modeling parameters subject to both aleatory (random) and epistemic (knowledge-based) uncertainty. [34].

2.3.2 Bias

Data bias refers to data being inaccurate or incomplete due to failing to accurately cover or reflect the overall opinion of a population [35]. Two variations of bias are:

Model scattering error: Refers to the inaccuracies that arise from the assumptions or parameters used in a scattering model—such as those modeling the behavior of light, radar, or other wave interactions with materials. These errors occur when the model fails to capture the true physical behavior of scattering, resulting in discrepancies between simulated predictions and observed measurements. Unlike the standard error, which measures variability due to sampling, model scattering error reflects the inherent limitations or inaccuracies of the model itself. Understanding and minimizing these errors is crucial to ensure that the model predictions reliably align with real-world data.

Model Bias: refers to a systematic error in the predictions made by a model, where the outputs consistently deviate from the true values in a particular direction. Unlike random errors, which fluctuate unpredictably, bias is a repeatable and consistent inaccuracy that stems from flaws in the model's assumptions, structure, or the data used during training or measurement. This kind of error affects the accuracy of a model by introducing a persistent tendency to underestimate or overestimate outcomes.

In both cases, bias is a consistent deviation from the true value. They specifically refer to systematic errors rather than random errors.

All of the above will consequently influence the accuracy of the data and require some manipulation for it to be considered representative.

2.4 Double Validation Metric

In the following part of this section, the different steps that the Double Validation Metric (DVM) contains will be explained in detail. In Figure 2.5, the process of the DVM is illustrated, providing a visual representation of the workflow and the sequence of operations involved.

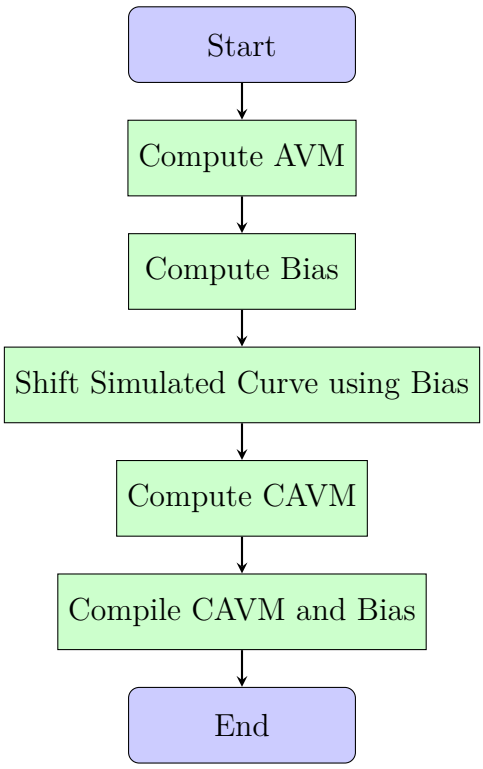

Figure 2.5: Flowchart of the DVM process. The steps include computing the AVM, estimating and applying bias, and calculating the CAVM

2.4.1 Area Validation Metric

The first basis of the DVM is the implementation of the Area Validation Metric (AVM). This metric, in essence, compares the area between the curves of model data and some ground truth data.

By calculating the eCDF for two datasets, the AVM can be defined as the integral of the absolute difference between their respective eCDF curves, as shown in Equation 2.21. This provides a quantitative metric for assessing their similarity.

$$d_{\mathrm{AVM}}(F, \widetilde{F}) = \int_{-\infty}^{\infty} \left| F(\zeta) - \tilde{F}(\zeta) \right| d\zeta \tag{2.21}$$

where $F(\zeta)$ is the eCDF curve for the real data and $\widetilde{F}(\zeta)$ is the eCDF curve of the simulated data [36].

From Equation 2.21 together with the fact that the cumulated probability $F(\zeta)$ for each timestep ζ is limited to values between [0,1] and dimensionless with M, a new equation emerges, Equation 2.22, resulting in the mean error of all M quantiles [36]. Additionally, the change in perspective and mapping the probability to the corresponding value of ζ results in $\zeta(F)$ and $\widetilde{\zeta}(F)$ is the quantile functions.

$$d_{\mathrm{AVM}}(F,\widetilde{F}) = \int_0^1 \left|\zeta(F) - \tilde{\zeta}(F)\right| dF = \frac{1}{M}\sum_{i=1}^{M}\left|\zeta(F_i) - \tilde{\zeta}(F_i)\right| \tag{2.22}$$

Furthermore, Equation 2.22 shows a strong resemblance to the mean error of all measurements and can therefore be expressed as Equation 2.23.

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n}\left|\zeta_i - \tilde{\zeta}_i\right| \tag{2.23}$$

2.4.2 Bias

Within the DVM framework, the model bias, denoted as d_{bias}, is computed according to Equation 2.24.

$$d_{\mathrm{bias}}(F,\tilde{F}) = d^- - d^+ \tag{2.24}$$

To account for this bias in subsequent evaluations, a correction can be applied using Equation 2.25, which adjusts the model output accordingly.

$$\tilde{F}_c(\zeta) = \tilde{F}(\zeta - d_{\mathrm{bias}}) = \tilde{F}(\zeta - (d^- - d^+)) \tag{2.25}$$

$\tilde{F}_c(\zeta)$ is the corrected simulation curve.

The quantity d^- represents the region where the simulated P-Box $\widetilde{F}$ lies below the real P-Box F, while d^+ corresponds to the region where $\widetilde{F}$ exceeds F.

2.4.2.1 Corrected Area Validation Metric

The limitation of DVM is that it lacks sensitivity against model scattering error when there is no overlap between the calculated sampled distributions, be it eCDF or P-Box. To compensate for this, one introduces Corrected Area Validation Metric (CAVM), evolving the AVM to validate model scattering error and model bias separately [37].

$$d_{\mathrm{CAVM}}(F,\tilde{F}) = d_{\mathrm{AVM}}(F,\tilde{F}_c) = d_c^- + d_c^+ \tag{2.26}$$

where d_{CAVM} is calculated as for AVM but with the newly shifted $\tilde{F}$ curve instead.

2.4.3 Double Validation Metric Result

DVM is a *quasimetric* as it is non-symmetric [37].
This statistical metric combines the uncertainty and error measurement from AVM with the model bias of the simulation [36]. Such a combination of model errors and uncertainties is helpful when performing evaluations of simulation models such as the one covered in this thesis. When the model attributes are unknown both areas of uncertainty needs to be accounted for.

The result gathered from the DVM is displayed in Equation 2.27 where the result is a combination of d_{bias} and d_{CAVM} which is calculated based on the simulated data, $\tilde{F}$, and real data, F.

$$d_{DVM}(F, \tilde{F}) = (d_{bias}(F, \tilde{F}), d_{CAVM}(F, \tilde{F})) \tag{2.27}$$

2.5 Scenario Definition

A definition by Cambridge Dictionary [38] describes a simulation as *"a situation or event that seems real but is not real, used especially to help people deal with such situations or events"*. In this project, the keywords *situations* or *events* are defined as **scenarios**.

The environments used in this project are those where data were collected from the physical car and later modeled in simulations. A simulation is essentially a replica of a real-world scenario, designed to mimic the behavior, processes, or outcomes of a system under specific conditions. For instance, if the physical car detects a large, stationary traffic sign on the right-hand side of the road, the same sign should appear in the simulation's detections. By using mathematical models and algorithms, simulations aim to reproduce similar detection data, offering valuable insights into complex systems and ensuring consistency between real-world and simulated environments.

3
Methods

In this chapter, the methodology used to conduct the thesis is defined. The thesis used two scenarios to validate the logged data against the simulated data. Specific information about how this was conducted can be found under section 3.1. Next, all the data management, including; target tracking, object identification, parsing of the data, filtering, and preprocessing, needed to be examined and understood to support future development. This is presented in the section 3.2 through 3.2.6. Furthermore, for a better understanding of where and when the radar detects objects, a visualization tool was created, and how this was conducted can be seen under section 3.3 and the subsection 3.3.1. Additionally, for the thesis to result in some form of validation of the radar model, a statistical evaluation was conducted, and further information about how the metric used was selected can be found in section 3.4 and how the evaluation were conducted can be found under section 3.5. Furthermore, on the next page the thesis outline can be seen i Figure 3.1.

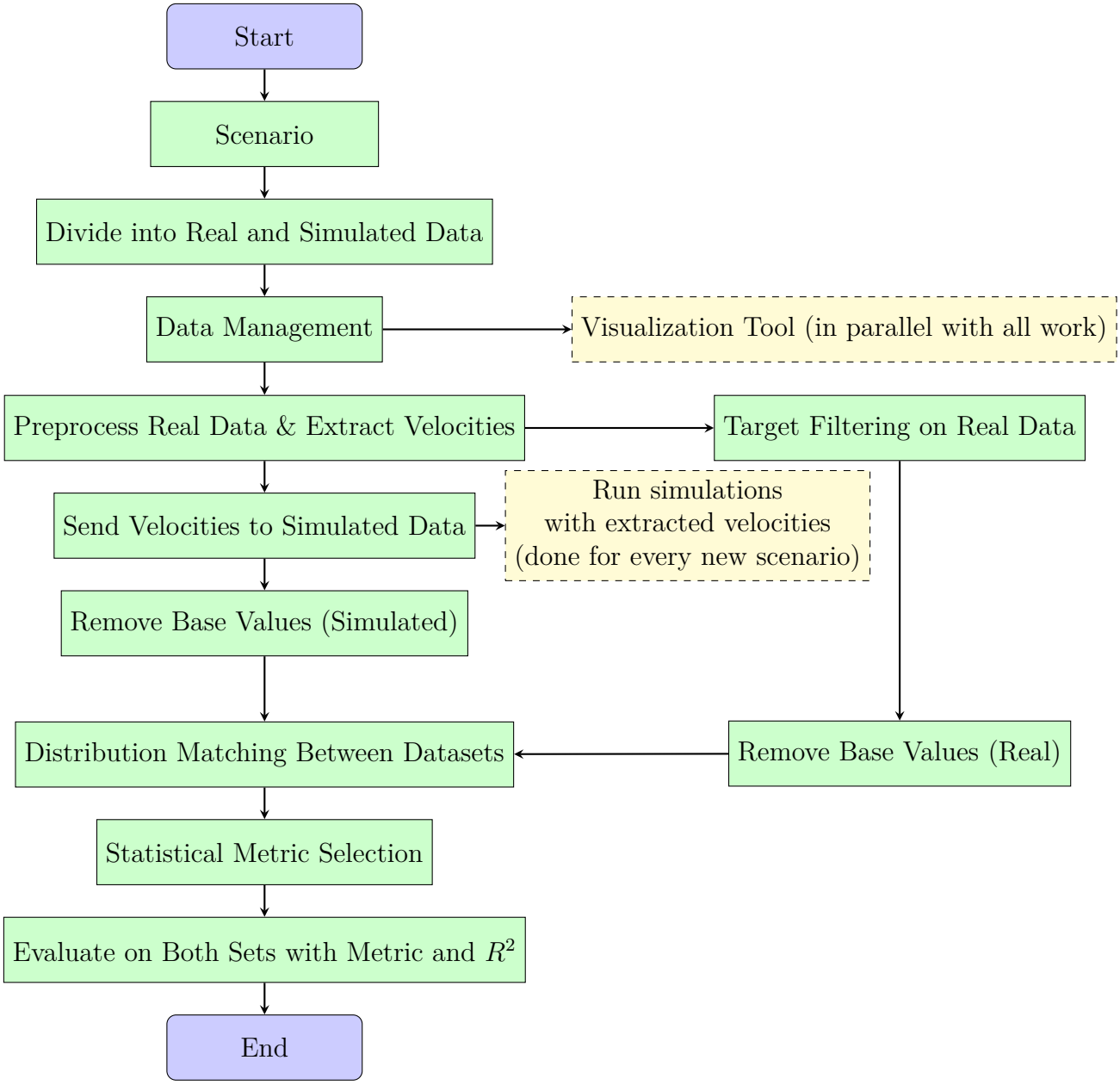

Figure 3.1: Thesis outline: The figure illustrates the workflow followed throughout the thesis. It shows how scenario selection serves as the foundation for both real and simulated data. This data is then preprocessed, extracted, parsed, and manipulated; irrelevant values are removed, and the distributions of the two datasets are aligned. Next, a statistical metric is selected based on a defined set of criteria. Finally, the data undergoes statistical evaluation using the metric chosen in the previous stage.

3.1 Scenario Selection

Based on the Euro NCAP criteria for developing AD&ADAS functions, the usage of simulation scenarios was chosen accordingly. For this project, simplicity was of importance as a starting point. Starting with a simpler case, as the one described in subsection 3.1.1. This provided insight into how the radar sensor system works in practice and a more precise comparison of synthetic and real detections. Furthermore, a more complex scenario was then used for further insight into how well the method performed. This is described in subsection 3.1.2. For this project, the statistical evaluation of the simulation models was intended to be applicable to any scenario. However, the simplicity of the scenarios and the use of pre-collected data made it practical to work with these specific cases.

3.1.1 Automatic Emergency Breaking System Car-to-Car Rear moving

This scenario is defined in the Euro NCAP protocol; "TEST PROTOCOL – AEB Car-to-Car systems, v4.3.1" [39] section 8.2.2.3, page 18. The so-called Car-to-Car Rear moving (CCRm) is a simple scenario within the subcategory - Automatic Emergency Breaking System (AEB) Car-to-Car systems. Here, a Volvo vehicle, Vehicle Under Test (VUT), is approaching another vehicle, Global Vehicle Target (GVT), traveling with a much lower velocity, from behind, causing the VUT AEB to engage. The scenario is further limited to events where the overlap between VUT and GVT is as large as possible, i.e., 100%, and the velocity of VUT is $30, 50, 60, 70 \frac{km}{h}$. The scenario subject to variations of VUT velocities is furthermore defined as sub-scenarios.

The CCRm is used as a starting point to evaluate the synthetic and real detection data from this scenario, which should overlap well and give a clear first insight into how the model performs.

3.1.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

The second scenario introduces more complexity to the method. Presented in section 7.3.3 on page 30 in the Euro NCAP protocol; "TEST PROTOCOL – AEB/LSS VRU systems" [40] is the Car-to-Bicyclist Nearside Adult Obstructed (CBNAO) scenario. A vehicle, VUT, is driving straight, passing two parked vehicles to the right. Before the VUT passes the parked vehicles, a bicycle appears from behind the parked vehicles to the right, causing the VUT AEB to engage. To simplify the scenario even further, the scenario can be split into a smaller sub-scenario, defined by VUT velocities of $10, 20, 30, 40, 50 \frac{km}{h}$.

This scenario contains more detection points and a level of complexity in terms of the trajectory and obstructed view of the Vulnerable Road User (VRU). Hence why it is chosen as the second scenario to be tested with the validation method. More detection data and a higher level of complexity allows for a more developed validation method with more parameters to take into account, which may provide a more diverse usage of the method.

3.2 Data Management

The data appears as a matrix with columns, representing the maximum number of detections per chirp and has the file format Hierarchical Data Format version 5 (HDF5), which as the name suggests is built up in a structured way. For easier accessibility and to navigate the files and to find all the desired parameters and features, the data were observed through a software called *HDFView 3.3.2* [41].

All HDF5 files contain head-folders grouping the main data bases for different systems in the vehicle, for example all radars. Within each head-folder are sub-folders containing various data for each system, this is where parameter data for the radars and detection data is stored, which can be parsed and used for the evaluation.

In addition to detection data, the HDF5 files contain target detection data. This includes properties like size, and other estimated attributes as well as ID numbers of objects detected by a variety of sensors on the car. Furthermore, object size can be represented using bounding boxes and objects get individual names through object IDs. Additional information on this can be found in subsection 3.2.1 and 3.2.2.

For this thesis to be conducted, the data is split into two different categories. Henceforth, the data will be referred to as:

Simulated data: Data simulated by personal in the AD&ADAS radar team at Volvo Cars.

Real data: Data collected from actual observations and measurements in real-world (controlled) environments by real cars, such as on test tracks.

3.2.1 Target tracking

When the data is collected, the software combines signals from several measuring instruments fitted to the cars. As the car observes objects during a driving session, data about these objects is stored for analysis. By combining these observations, a complete picture of tracked objects or "targets" can be created. Properties such as size and road status, i.e., velocity and positional data of the targets, are stored. Furthermore, these so-called boundary boxes contain information about the target's height, width, and length.

3.2.1.1 Reference Points

When the boundary boxes are estimated in the pre-processing, each object is assigned a reference point. As displayed in 3.2 the point of reference is located on on of the objects four sides, most often the side closest to VUT. Due to their location the object position has to be adjusted by 1/2 object length or width, depending on position of reference point.

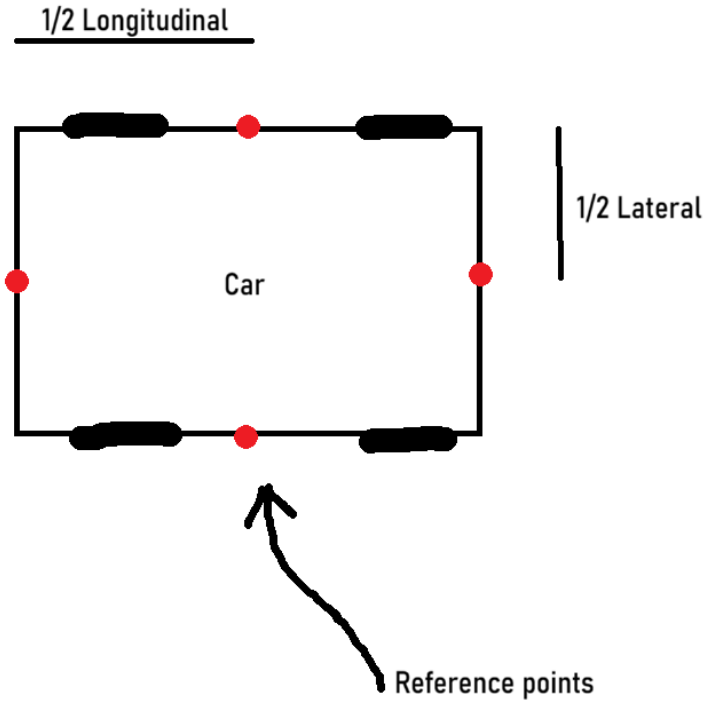

Figure 3.2: Displaying the reference points used for positioning the boundary box on the different objects.

3.2.2 Object ID

When the radar detects an object, it is assigned an object ID. This ID remains the same until the radar loses sight of the object; at this point, the ID is reset to the default value of 0. If the object remains detectable throughout the entire session, the column where the object is recorded is dedicated solely to that object and contains only the given ID number or 0. If the radar loses the object and then detects a new object, the new object may be assigned to the same column as an object detected earlier, but with a different ID.

Algorithm 1 serves as a guide for selectively evaluating objects by filtering out irrelevant ones. It streamlines the selection process by identifying only the desired objects, ensuring that only relevant data is processed.

Algorithm 1 Selecting the desired targets: Step-by-step on how to locate, cross-check and select the desired targets from the HDF5 file that will be evaluated.

1. Locate and open the object ID and one of the size attributes; height, length, and width, object position; latitudinal, longitudinal, and speed.
2. Do a cross-check between all variables to detect which column in the files corresponds to which object.
3. Enter the wished columns in a list.

3.2.3 Parsing

After familiarizing with the data, the data needed to be parsed, and some parameters calculated. All required parameters were extracted during the parsing - *azimuthal angle*, *range*, *elevation angle*, *Doppler Velocity*, SNR, and RCS. Besides the parameters, the lateral and longitudinal positions and mounting angle, in this case, for the placement of the FR, were extracted as well. Once all this was parsed, the x and y values of the detection could be calculated, allowing for the visualization of the detections in a 2D environment. Further information is found in section 3.3.

To calculate the x and y values, the following equations are used:

$$\begin{aligned} x &= r \cdot cos(\alpha + \phi) + p_{long} \\ y &= r \cdot sin(\alpha + \phi) + p_{lat} \end{aligned} \tag{3.1}$$

where r is the range from the radar [meters], α is azimuthal angle [radians], ϕ is the mounting angle of the radar on the car [radians] and p_{long}, p_{lat} are longitudinal and latitudinal mounting position of the radar on the car.

3.2.4 Filtering and Preprocessing Real Data for Simulation Comparison

To perform a fair scenario simulation, focusing on the critical settings specific to the situation, particularly where the scenario starts and ends, is essential.
In the majority of the logs, since data is collected during the entire drive cycle, there is an abundance of detections recorded during a run in the physical vehicle (real data), causing a mismatch with the specified scenario. This includes detections before the test environment starts and unwanted movements or positioning of the VUT. In the simulated environment, the defined scenario begins once played, making it as precise as possible, thus only containing the scenario.

Consequently, the real datasets must undergo substantial preprocessing before a meaningful comparison with the simulated datasets can be performed. While the scenarios defined by Euro NCAP are well-structured and require specific conditions to be met for accurate simulation, simulations offer the advantage of being conducted in controlled, "vacuum-like" environments, free from natural disturbances and unpredictable elements. In contrast, real data collection occurs in uncontrolled settings, often resulting in detections of irrelevant objects such as signs, road barriers, and terrain features unrelated to the intended scenario. These irrelevant detections, referred to as noise, complicate the analysis and can distort the evaluation. Therefore, filtering techniques are essential to isolate relevant information and ensure a justified and accurate comparison between the two datasets. Selecting only the relevant detections helps reduce or eliminate noise, streamlining the evaluation process and enhancing the reliability of the results. Hence, the following criteria must be fulfilled to ensure the integrity of the comparison.

1. Scenario target(s) can be identified. Follow algorithm 1 in subsection 3.2.2 to select the desired targets.
2. VUT driving according to scenario.
3. Target positioned correctly with scenario-specified velocity.
4. Non-zero velocity data for VUT to ensure that the vehicle has the correct speed, and marks can act as a start marker for the scenario.
5. Initial positional difference (range) between VUT and all included targets.
6. (If necessary) Select the desired start and end row index, slicing the data frames. This eliminates unnecessary data and unwanted targets embedded in the same column as desired objects at earlier or later timestamps.

Once all the requirements above are met, the data can be paired. Data frames containing the velocity vectors of the VUT and moving targets are created. If requirement six is not met, the data frames are sliced according to the desired start and end indices to ensure only the indices within the framework where all requirements are met are stored. If requirement six is met, the frames are already sliced as needed. This data is then sent to the simulation, where the scenario is simulated based on the logged velocities to replicate the factual scenario as closely as possible.

Positions for target objects in the simulated environment are calculated according to the scenario specifications. Based on the reference point tracked by the target tracking functions, the origin of the target vehicle (located in the center of the back wheel axis) can be determined based on object attributes (i.e., car model, pedestrian, or bicyclist).

3.2.5 The Radar Simulation Model

There are several types of sensor models with varying levels of accuracy (fidelity): ideal sensor models, statistical models, and physics-based models. An **ideal radar model** can use ground truth to propagate detections on various objects, with or without constraints such as maximum range or field of view.

A **statistical model** is based on probability and stochastic data. It can also use ground truth data, but will generate detections within a range based on probability derived from statistical data about the actual sensor.

Physics-based models aim to replicate actual physical phenomena, such as the reflection and transmission of a microwave signal.

The RWS evaluated in this project can be described as a mix between an ideal and a statistical sensor model. While taking in ground truth data describing the position, velocity, and other attribute information of various objects in the Open Simulation Interface (OSI) format. It then generates radar detections from statistical data gathered from the real sensor.

As mentioned in the limitation section 1.2, the simulation procedure is done externally, hence only briefly mentioned in this section.

3.2.5.1 Targeted Filtering and Detection Refinement

In addition to extracting the desired target(s) and the velocities, the detections and the corresponding parameter values needed to be extracted.

Since the estimated properties of the targets' bounding boxes are stored in the logged data package, as shown in subsection 3.2.1, the relevant detections within these bounding boxes can be extracted. The first step is to select the desired objects, which is done in a manner similar to the selection of velocity vectors described in subsection 3.2.4, using Algorithm 1 from subsection 3.2.2.

By knowing the height, length, and width of the objects of interest, and utilizing the results from Equation 3.1, all x and y coordinates lying within or arbitrarily close to the boundary boxes are considered the desired detections. If detections arbitrarily close to the box are not considered, there is a slim risk that these detections may not be noted due to fluctuations in the detections.

3.2.5.2 Removing standard values for Real Data

As mentioned in subsection 3.2.4, real data contains significantly more information than the simulated preprocessing, more cluttered surroundings, and longer recording. Consequently, preprocessing is required, as mentioned in subsubsection 3.2.5.1 and subsection 3.2.4, especially if the whole recording should be visualized without slicing the data.

Even after selecting the desired columns and removing unwanted objects, the fact that not all radar scans detect an object causes problems. When the radar does not receive a detection, the non-detections are filled with different standard values depending on the parameter being investigated. The visualization is representative and useful if all the standard values are filtered out and removed.

3.2.6 Distribution Matching

To determine whether the simulation is a trustworthy representation of the real data, eCDF can be utilized for both datasets. One can evaluate the overall distribution and identify deviations between the two datasets by comparing their individual eCDF.

Even after undergoing matching procedures, as mentioned in subsection 3.2.4 and 3.2.5.1, the final detection datasets from the defined scenarios might not match in length due to discrepancies between the datasets.
This depends both on different numbers detections being tracked by the physical sensor and the simulation model, but also timestamp dependencies. When the real

detection data is preprocessed the data is logged in timestamps of 25 ms, whereas the actual time between each chirp, as modeled and logged in the simulation is 50 ms. This leads to duplicates of some rows from the real sensor data.
To ensure the data arrays have the same length while maintaining their distribution and information, piecewise step interpolation is applied to the synthetic data. This method adds elements to the arrays, resulting in equal lengths and enabling a comparison of the distributions between the two datasets. The interpolation is based on the uniquely common samples from each sample set. Applying this also handles mismatch in time stamps between the physical sensor output table and the simulation output tables.

3.3 Visualisation

The visualization was developed with Python version 3.13.1 with the help of the standard Python interface to the Tcl/Tk GUI toolkit, *tkinter*, and the plotting library for Python, *matplotlib* [42, 43, 44].

Firstly, data visualization is mandatory to understand the underlying distributions and similarities or lack thereof between synthetic and real data. By scattering two sets of data for each parameter, one can get an initial comparison and underlying information about the data structure.

Secondly, understand where detected objects and structures are located from the car's point of view. For this, a Graphical User Interface (GUI) was created to enable data visualization. By plotting the calculated coordinates from Equation 3.1 of the detected objects in a 2D view, and with a deeper understanding of radar functionality and the distribution of different datasets, one can gain valuable insights into the detection patterns and data dispersion. The GUI also has a slider which makes it handy to move throughout the different timestamps and simultaneously see how the scatters propagate.

Displayed in Figure 3.3 is a frame from the visualization, where the colored dots are real detections at some timestamp captured on a physical radar sensor system on a Volvo Car's vehicle.

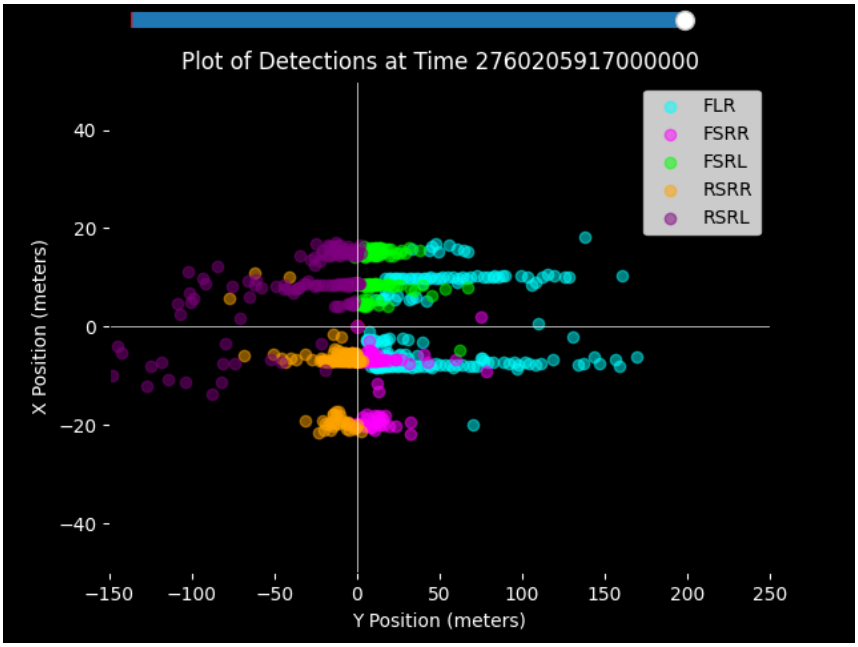

Figure 3.3: One frame from the visualization where the different colors represent different radars on the VUT. The VUT is moving towards the right and is seen from above.

3.3.1 Multi-visualization

To effectively compare the real and simulated datasets, an additional visualization was developed that plots both datasets within the same window. This combined visualization serves several important purposes.

Firstly, by plotting the real and simulated data, one can observe how the simulation propagates through time compared to the real data. This helps identify any discrepancies in the timing or sequence of events between the two datasets. This visualization assists in assessing whether the simulation has been accurately modeled. By comparing the detections and overlaying the patterns of the real and simulated data points, it is possible to determine if the simulation behaves as expected.

Secondly, the visualization also aids in evaluating the number of detections detected by the simulated model versus the real radar model. By displaying both datasets together, it becomes easy to see if the simulated radar model detects more or fewer objects than the real radar model and where they differ. This provides insights into the effectiveness and reliability of the simulation. Identifying where the simulation diverges from the real data makes it possible to pinpoint specific aspects and areas of the simulation that may need to be adjusted or improved.

Overall, this visualization tool is a powerful aid in validating and improving the scenario(s) being simulated. Differences and abnormalities between the two datasets can be seen and potentially fixed, ensuring that the simulated scenario accurately represents the real-world scenario. An example of how this visualization is used is shown below in Figure 3.4.

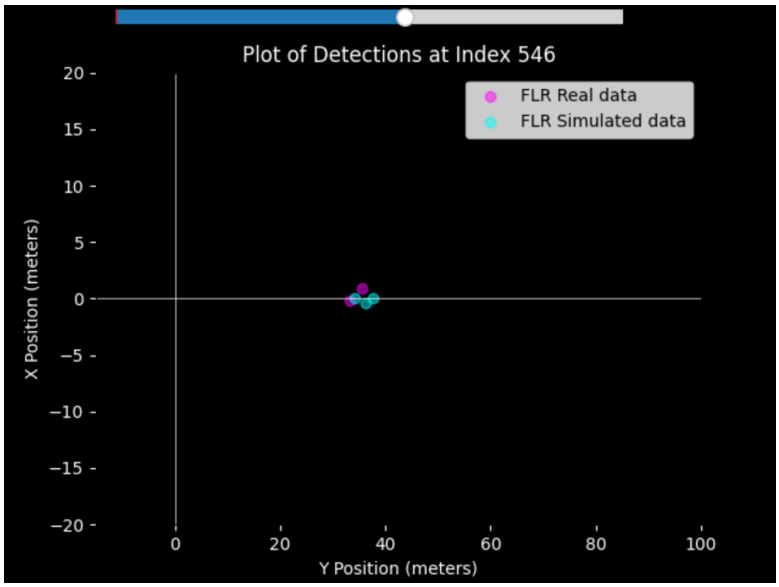

Figure 3.4: One frame from the visualization where the different colors represent two datasets (real and simulated) from detections by the FR. The VUT is moving to the right and is seen from above. During this visualization, the radar was located at $(0, 0)$ and then the car was tested under Euro NCAP AEB CCRm with the VUT having a speed of $30\frac{km}{h}$ and the GVT having a speed of $20\frac{km}{h}$. The VUT has a lateral overlap of 100%.

3.4 Statistical Metric Selection

Evaluation of APSS simulation models is a complex task. Since the software and parameters of the simulation model are undisclosed by the distributor, not only are deviations based on probabilistic distribution (aleatory uncertainty) to be accounted for. However, one also needs to consider epistemic uncertainty, which involves the lack of knowledge about the model structure or measurement errors [36]. With this in mind, a statistical metric needs to fulfill a set of criteria to work objectively well as an evaluation tool on APSS simulations.

Discussed by Elster et al. in the *"Introducing the double validation metric for radar sensor models"* [36] and the Phd thesis by Rosenberger et al.[37], the following criteria should apply to the evaluation metric of choice:

1. Metrics meet the mathematical properties of a metric as defined by Fréchet [45]. (Unbounded results)
2. Metrics are intuitive. (Plausible & output in unit of measurand)
3. Metrics apply to both deterministic and non-deterministic data.
4. Metrics are quantitative and objective. (No manually tuned parameters)
5. Metrics do not include acceptance criteria. (No Boolean output)
6. Metrics consider uncertainties. (Epistemic and aleatory)

7. Metrics define a confidence interval with respect to the number of measurement data points.

In addition to the above criteria, statistical evaluation metrics for APSS can be chosen according to their applicability on interface patterns such as detections or objects, scenario definitions, and specified uncertainties.

According to the investigations carried out by (Rosenberger et al.) [37] the best validated statistical metric for evaluating APSS simulation models is the AVM. This holds since it nearly fulfills all criteria above except one, criterion 6.

Besides the criteria mentioned above, an additional requirement is the ability to distinguish model bias and model scattering error. Therefore, Rosenberger introduced the DVM to distinguish the two components [37, 36]. Hence, the evaluation of the simulation model was limited to only using the DVM. How it works can be seen in section 2.4.

To utilize the entire AVM and make up for the aleatory uncertainty in the simulation model, the implementation of P-Box is introduced to the eCDF computations as explained in section 2.3.1. An arbitrary confidence level of 5% is added to create the upper and lower bounds of the eCDF's. This level of confidence may be altered further, for the statistical investigative nature of this thesis, an uncertainty of 5% should be adequate as it is a commonly used level of confidence within academic research.

For further reasoning and additional insight into the truthfulness of the simulated model, MLE and R-squared were used as described in section 2.2.3 and 2.2.4.

3.4.1 Integration method

To calculate the integrals of the two data sets for the AVM method, the Trapezoidal rule, as implemented in the **SciPy** Python package [46], is employed. Equation 3.2 describes the calculation [47].

$$\int_a^b f(x)dx \approx (b-a) \cdot \frac{1}{2}(f(a) + f(b)) \tag{3.2}$$

This is a simple and cost effective way of approximating integrals. Since the functions covered in this method take the form of step functions, approximating the integral into trapezoids will not deviate much from the true integrated area.

3.5 Metric Procedure

When it comes to the evaluation of the data, four main steps are performed:

1. Calculate the bias
2. Shift the simulated curve based on the bias
3. Calculate the CAVM
4. Sum the DVM with the results gathered by step 1 and step 3.

Once the two curves are plotted, as mentioned in the first paragraph in subsection 3.2.6, the AVM is used to calculate the areal difference between the two curves. Based on the numbers retrieved from the AVM, the bias is calculated according to Equation 2.24. This is done to account for the bias within the model and implies shifting the simulated curve based on the bias, without changing its characteristic. Once the simulated curve is shifted, the CAVM can be calculated in the same way as the AVM. Further, the DVM is complete since this is just a collection of both the bias and the CAVM.

3.5.1 Maximum Likehlihood Estimation

Maximum Likelihood Estimation was used to determine the expected distribution of the two datasets. Firstly, the information gathered from the MLE provided insight and a basis for discussing the results from the DVM between simulated and real radar detections. Secondly, offering a deeper understanding of the results by serving as an evaluation tool for the DVM, while comparing the simulated detections against real detections.

Since the parameter data theoretically belonged to different distributions, MLE was tested on a set of commonly used distributions. The fit was evaluated using the AIC, providing a comparative score of the fit's quality.

Based on the theoretical properties of the radar parameters and distribution theory [48, 49, 50], the following seven distributions were fitted against the data:

Table 3.1: Displaying the set of chosen theoretical distributions and their connections to the parameters in radar modeling.

Distribution	Properties	Radar application
1. Normal	• Common modeling distribution • Symmetric measurement errors • Variations around a mean	• Useful for modeling `Doppler Velocity` • Noise in range/velocity estimations
2. Log-Normal	• Positive-valued • Right skewed data	• Common for modeling (`rcs`) • Signal strength (`snr`)
3. Gamma [50]	• Generalization of exponential	• Suitable for modeling `snr` and `rcs` • Cluttered environments
4. Uniform	• Equally likely values within a range	• Baseline assumption for random angular data • `azimuthal angle` and `elevation angle`
5. Beta [49]	• Variables bounded between 0 and 1.	• Good fit for angular parameters • `azimuthal angle` or `elevation angle` • Normalized values
6. Weibull	• Highly flexible (Skewed lifetime/failure models)?	Modeling backscatter amplitude • Target reflectivity • `rcs`
7. Exponential [48]	• Time or distance between rare events	• Target detection distances • Fading signal properties

3.5.2 Correlation analysis by R^2

Correlation plots for the six parameters; *range*, *azimuthal angle*, SNR, RCS, *Doppler Velocity*, and *elevation angle*, were utilized for further analysis between the two datasets. These plots visually represent the direct comparison between the real and simulated data for each parameter, allowing for a detailed comparison of their correlations. By examining these plots, patterns, trends, and discrepancies between the datasets can be identified, providing insights into the accuracy and reliability of the models performance in replicating real-world conditions.

3.5.3 Number of detections

In addition to the six parameters retrieved from the radar data, a seventh parameter was created based on the number of detections the two sets of detection data received for each scenario and sub-scenario. The number of detections plays a vital role in determining the robustness of the radar models and whether they behave consistently during the runs. One key objective is to determine if the simulation model and the physical radar sensor collect the same number of detections throughout the entire run, revealing whether the model and sensor react similarly to different ranges and times.

4
Results

This chapter contains results from the two scenarios and the corresponding evaluations. The results are presented visually and numerically for each section, illustrating the output from the method created in this thesis. After each section, an interpretation of the results is presented, discussing the implications and significance of the findings.

In the interpretation sections, the strengths and limitations of the method are discussed, along with potential areas for improvement. The implications of the findings for future research and practical applications are also considered.

Additionally, the number of plots and tables generated by the two scenarios and their respective sub-scenarios is too large. Therefore, only one sub-scenario for each scenario is displayed in the results. The remaining plots and tables are in Appendix A - H.

4.1 Automatic Emergency Breaking System Car-to-Car Rear moving

In this section, the first scenario, CCRm, and one of the sub-scenarios, CCRm with VUT speed of $30\frac{km}{h}$, is displayed with the results based purely on the method created and explained in the methodology under chapter– 3.

The first two figures, figure 4.1 and figure 4.2, display each parameter along with three lines representing the simulated data, the shifted simulated data, the corresponding P-Box area, and real data. The first plot is non-normalized, while the second plot is normalized. The results show a clear trend regarding speed. As the speed of the VUT increases, the overall DVM also increases, except for the SNR, where the DVM decreases. Underneath, in Tables 4.1 and 4.2, the corresponding tables display all numbers obtained from the calculations done throughout the methodology.

Figure 4.5 presents the results from the MLE analysis. Generally, the same distributions are selected across all sub-scenarios for the real and simulated data, with minor variations. The most frequently chosen distributions are the *beta distribution*, *log-normal distribution*, and *weibull minimum distribution*. Across all VUT speeds, 10 out of 24 eCDF parameters share the same distribution between the real and

simulated data, while the remaining 14 parameters exhibit different distributions. The *beta distribution* is the most commonly identified related distribution.

Figure E.7 displays the real and simulated data correlation. Generally, it can be observed from the plots that the coefficient of determination (R^2) value sinks as the speed of the VUT increases. This indicates that the accuracy of the simulation decreases with higher speeds. Among the various metrics analyzed, the *range* is most significantly impacted by the increase in speed, showing the greatest deviation between the simulated and real data.

Finally, Figure 4.7 presents a plot showing the number of detections at each timestep for both datasets. The trends observed indicate that the simulated data consistently registers a higher number of detections compared to the real data at VUT speeds of $30\frac{km}{h}$ and $50\frac{km}{h}$ throughout the entire scenario. Opposite, for the other two speeds, the simulated data only shows more detections during the initial timestamps. In the later stages of the scenario, the real data surpasses the simulated data in the number of detections.

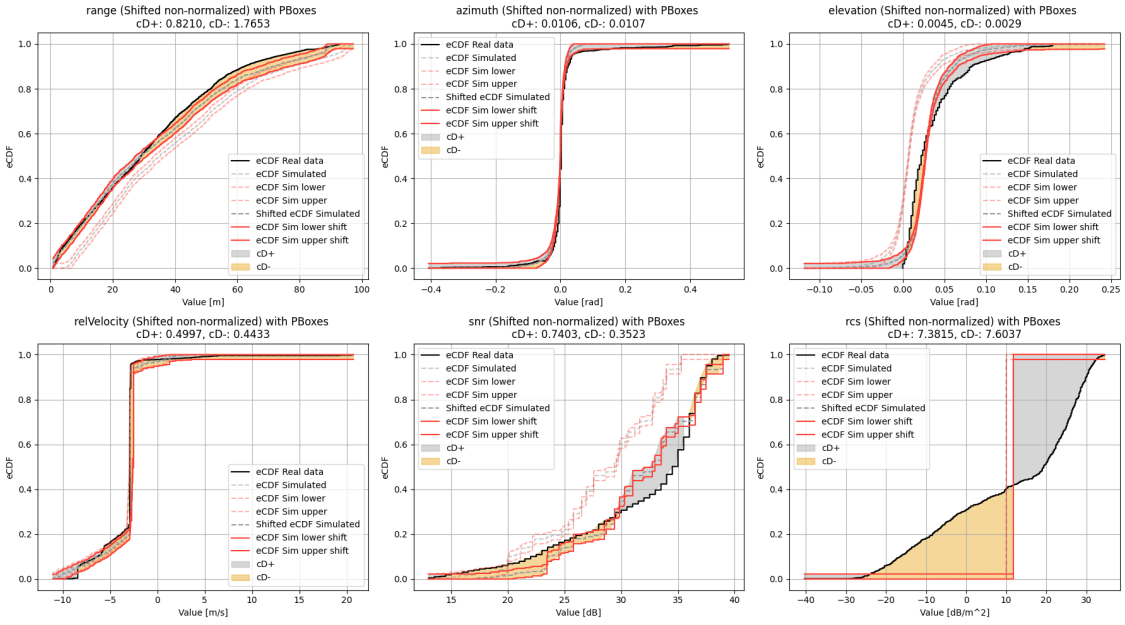

Figure 4.1: Shifted and non-normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $30\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

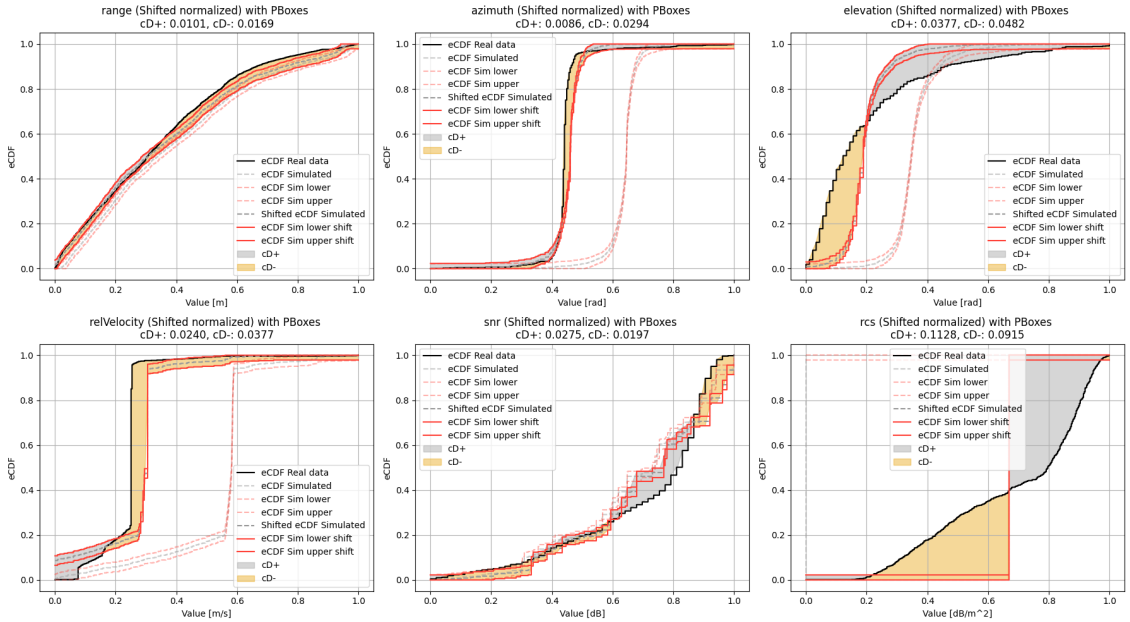

Figure 4.2: Shifted and normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $30\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

Table 4.1: Data from shifted non-normalized CCRm scenario with VUT speed of $30\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0105	5.9172	0.8210	1.7653	2.5863	5.9067	(2.5863, 5.9067)
azimuth angle	0.0103	0.0107	0.0106	0.0107	0.0213	0.0004	(0.0213, 0.0004)
elevation	0.0206	0.0004	0.0045	0.0029	0.0074	-0.0202	(0.0074, -0.0202)
Doppler Velocity	0.5789	0.2369	0.4997	0.4433	0.9430	-0.3420	(0.9430, -0.3420)
SNR	3.3821	0.0653	0.7403	0.3523	1.0925	-3.3168	(1.0925, -3.3168)
RCS	8.3449	6.9394	7.3815	7.6037	14.9851	-1.4055	(14.9851, -1.4055)

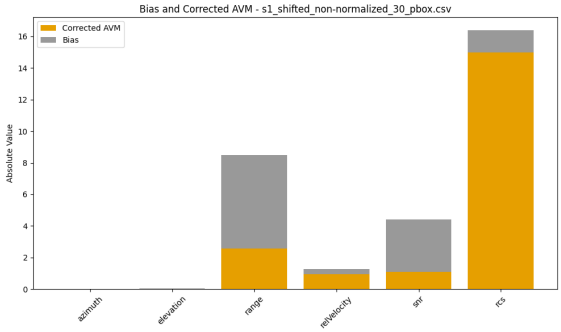

Figure 4.3: Corresponding stacked bar plots to the results from Table 4.1 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table 4.2: Data from shifted normalized CCRm scenario with VUT speed of $30\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0012	0.0326	0.0101	0.0169	0.0270	0.0314	(0.0270, 0.0314)
azimuth angle	0.0079	0.1941	0.0086	0.0294	0.0380	0.1861	(0.0380, 0.1861)
elevation	0.0094	0.1725	0.0377	0.0482	0.0859	0.1631	(0.0859, 0.1631)
Doppler Velocity	0.0053	0.2910	0.0240	0.0377	0.0616	0.2857	(0.0616, 0.2857)
SNR	0.0387	0.0129	0.0275	0.0197	0.0472	-0.0258	(0.0472, -0.0258)
RCS	0.6680	0.0002	0.1128	0.0915	0.2043	-0.6678	(0.2043, -0.6678)

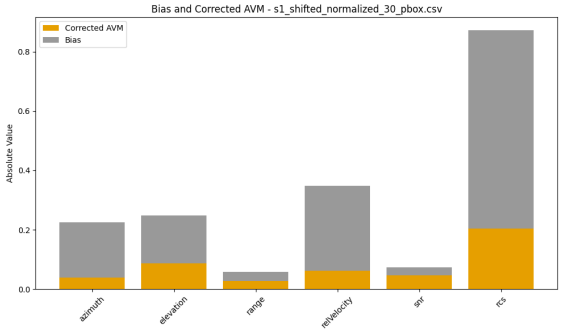

Figure 4.4: Corresponding stacked bar plots to the results from Table 4.2 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

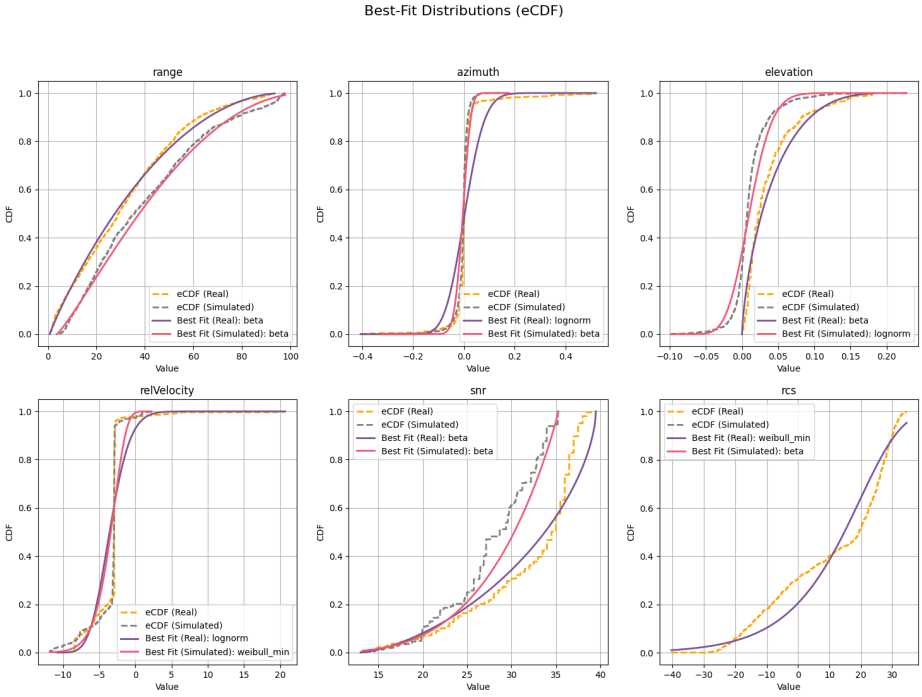

Figure 4.5: MLE plot for non-normalized and non-shifted data where VUT has a speed of $30\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

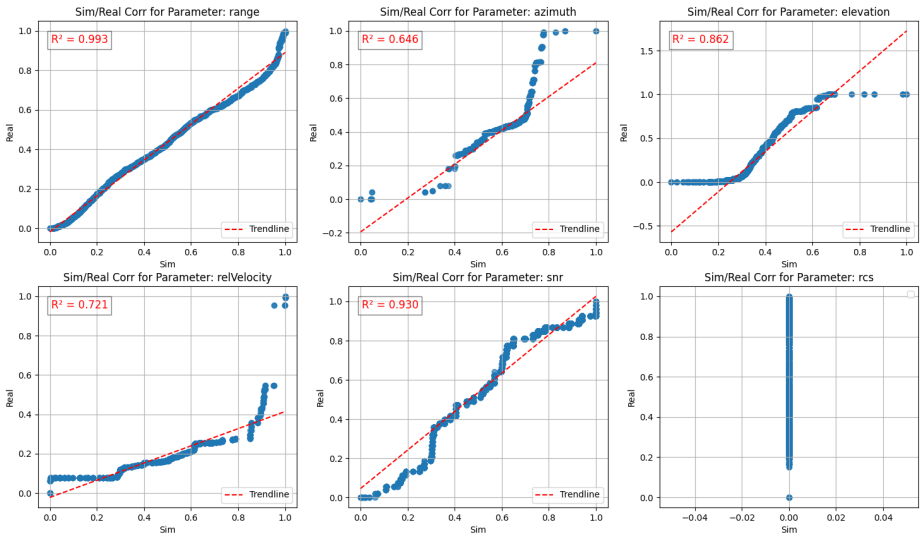

Figure 4.6: Correlation between the same parameter for simulated data against real data.

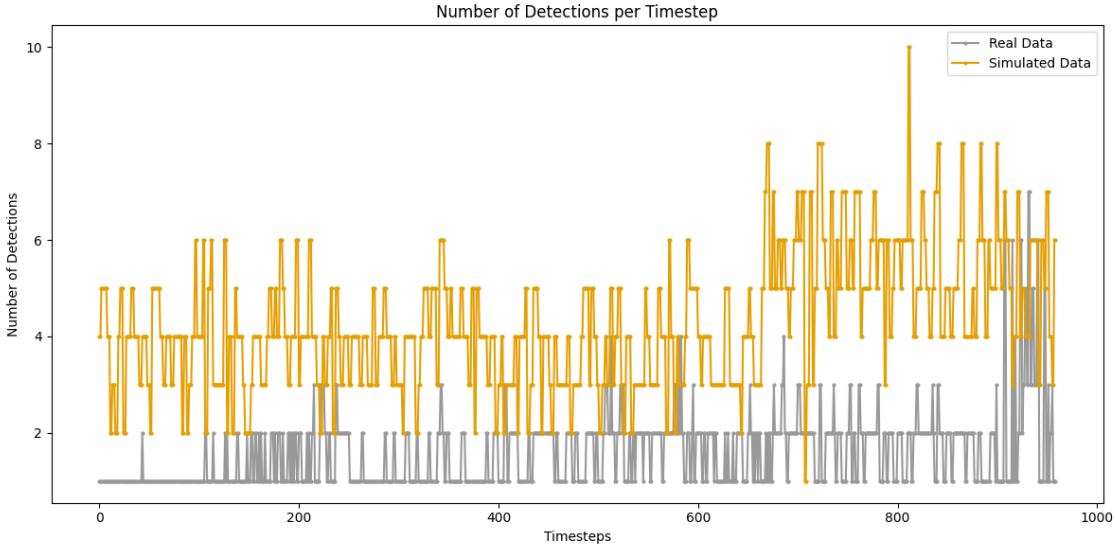

Figure 4.7: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CCRm with VUT speed $30\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

4.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

In this section, the second scenario, CBNAO, and one of the sub-scenarios, CBNAO with VUT speed of $10\frac{km}{h}$, is displayed. As for the second scenario, the results are based on the method, found in chapter 3, created within the thesis.

Figure 4.8 and Figure 4.9 display each parameter along with three lines representing the simulated data, the shifted simulated data, the corresponding P-Box area, and real data. The first plot is non-normalized, while the second plot is normalized. The results do not show any clear trends, but as the speed of the VUT increases, fluctuations can be seen for all parameters and their corresponding DVM values. The changes are small and remain nearly the same throughout all speeds. Underneath, in Tables 4.3 and 4.4, the corresponding tables display all numbers obtained from the calculations done throughout the methodology.

Figure 4.12 shows the result from the MLE analysis. Similarities between the most common distributions of the MLE plots of CCRm and CBNAO can be seen. The most common distributions are, as for the previous scenario, *beta distribution*, *log-normal distribution*, and *weibull minimum distribution*. A trend that can be seen for both *range* and SNR is that once a specific parameter has been fitted with, e.g, *beta distribution*, that fit remains through all scenarios with different VUT speeds. *Azimuth* and *elevation* have the same tendency, but changes fit at least once for the real data.

Figure 4.13 displays the real and simulated data correlation. The result that can be seen is that *range* and SNR have the same R^2 value throughout all runs. *Doppler Velocity* and RCS gets an improved R^2 as the speed increase, while *azimuth* and *elevation* obtain a lower R^2 values as the speed increases.

Finally, Figure 4.14 presents a plot showing the number of detections at each dataset's timestep. The observed trends are similar for CBNAO as for CCRm, indicating that the simulated data consistently registers a higher number of detections than the real data until the last timestep, where the simulated and real data approximately have the same number of detections. This holds for all the cases.

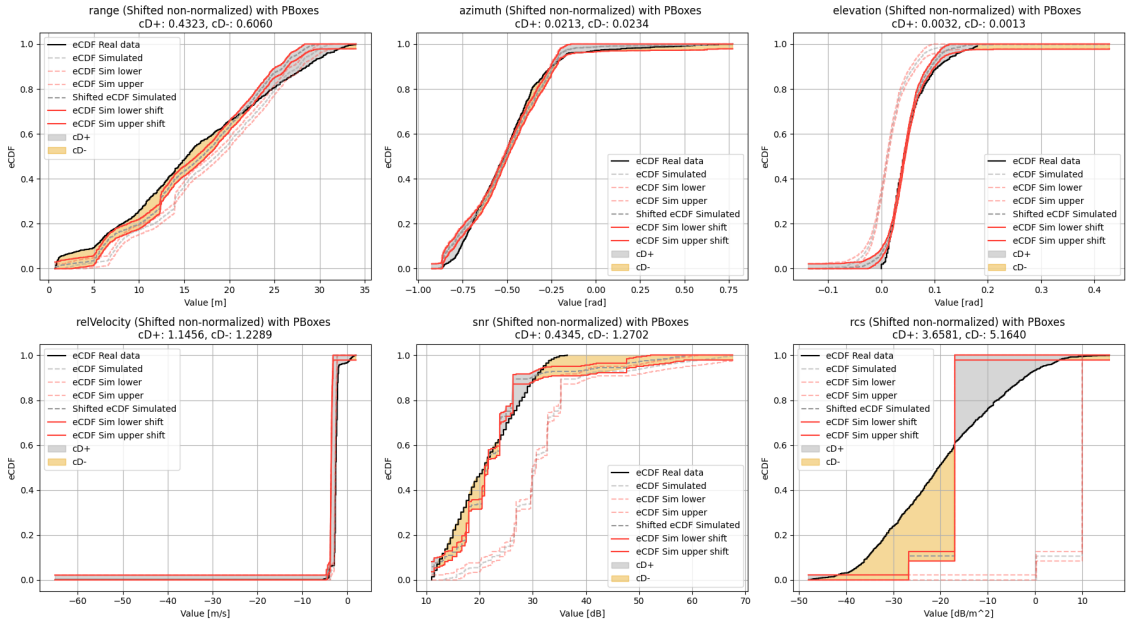

Figure 4.8: Shifted and non-normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $10\frac{km}{h}$.

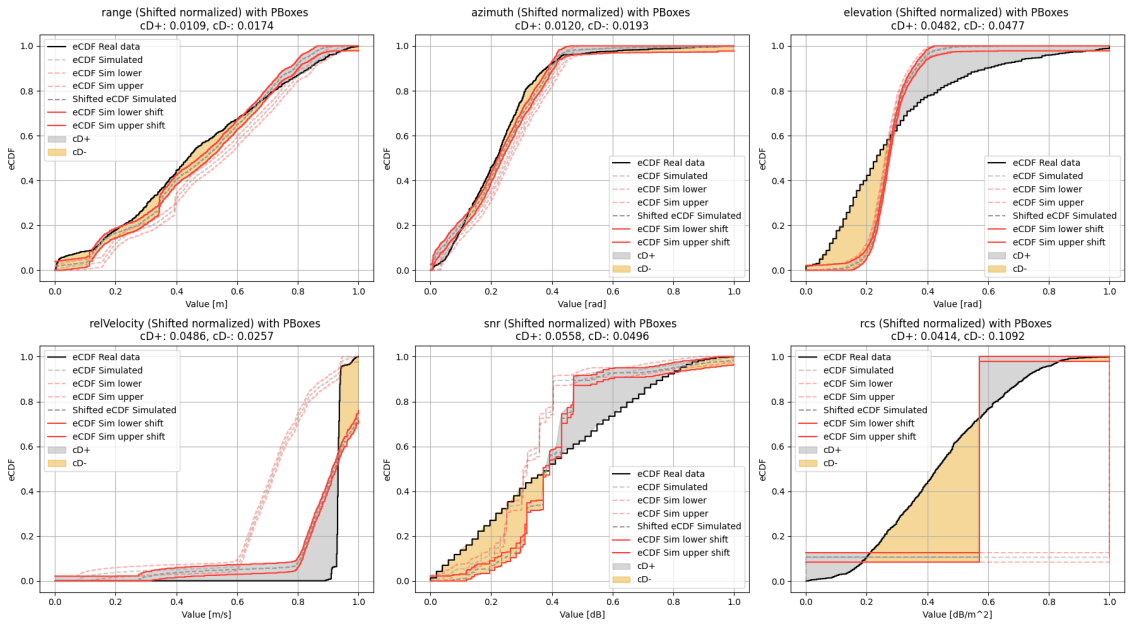

Figure 4.9: Shifted and normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $10\frac{km}{h}$.

Table 4.3: Data from shifted non-normalized CBNAO scenario with VUT speed of $10\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.1346	1.7351	0.4323	0.6060	1.0383	1.6005	(1.0383, 1.6005)
azimuth angle	0.0255	0.0210	0.0213	0.0234	0.0448	-0.0045	(0.0448, -0.0045)
elevation	0.0320	0.0002	0.0032	0.0013	0.0045	-0.0317	(0.0045, -0.0317)
Doppler Velocity	0.3103	1.2416	1.1456	1.2289	2.3745	0.9313	(2.3745, 0.9313)
SNR	0.0035	9.0984	0.4345	1.2702	1.7048	9.0949	(1.7048, 9.0949)
RCS	0.1665	27.3213	3.6581	5.1640	8.8221	27.1548	(8.8221, 27.1548)

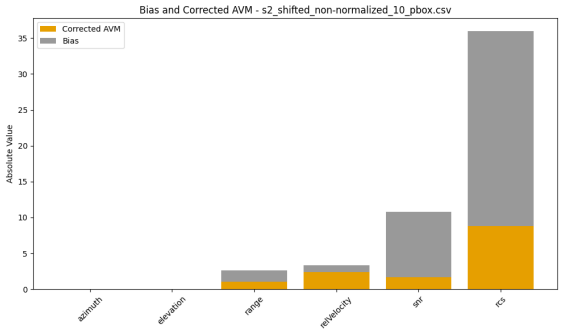

Figure 4.10: Corresponding stacked bar plots to the results from table 4.3 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table 4.4: Data from shifted normalized CBNAO scenario with VUT speed of $10\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0036	0.0553	0.0109	0.0174	0.0284	0.0517	(0.0284, 0.0517)
azimuth angle	0.0103	0.0389	0.0120	0.0193	0.0313	0.0286	(0.0313, 0.0286)
elevation	0.0519	0.0420	0.0482	0.0477	0.0959	-0.0100	(0.0959, -0.0100)
Doppler Velocity	0.1986	0.0033	0.0486	0.0257	0.0743	-0.1953	(0.0743, -0.1953)
SNR	0.0858	0.0212	0.0558	0.0496	0.1055	-0.0646	(0.1055, -0.0646)
RCS	0.0112	0.4431	0.0414	0.1092	0.1506	0.4319	(0.1506, 0.4319)

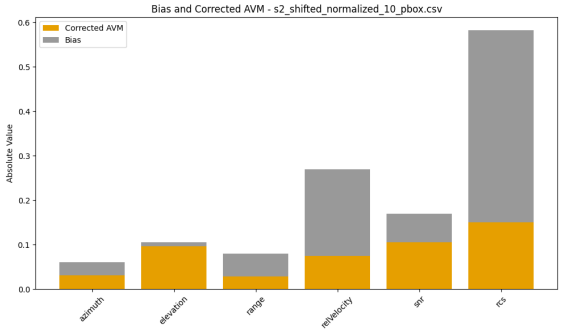

Figure 4.11: Corresponding stacked bar plots to the results from Table 4.4 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

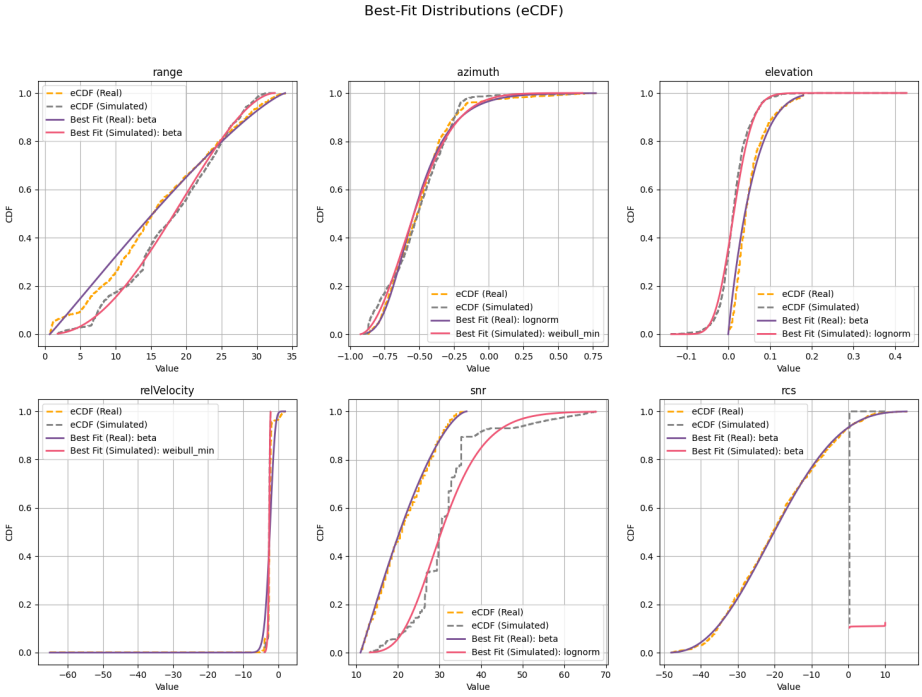

Figure 4.12: MLE plot for non-normalized and non-shifted data where VUT has a speed of $10\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

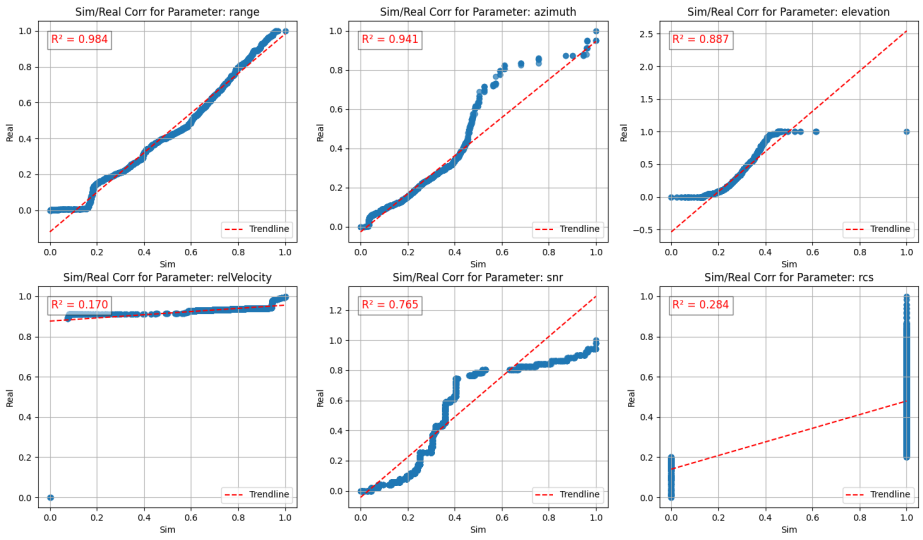

Figure 4.13: Correlation between the same parameter for simulated data against real data.

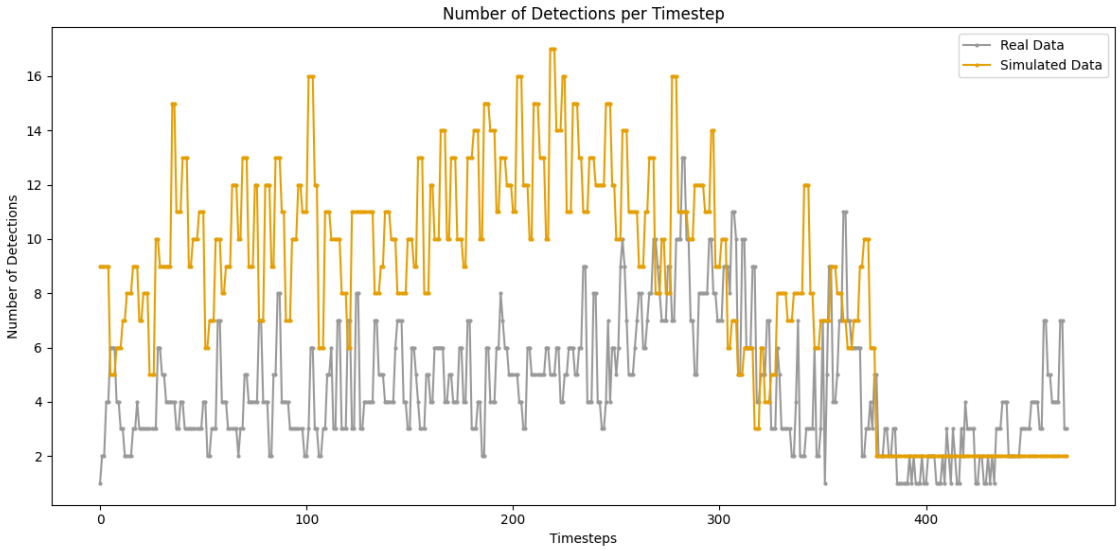

Figure 4.14: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CBNAO with VUT speed $10\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

4.3 Interpretation

When observing the tables summarizing the results from the DVM, there are clear variations between the parameters in each sub-scenario. Generally the modeled sensor is prone to bias more than model scattering error.

According to our DVM method on the CCRm scenario, we see a clear overlap in most of the radar parameters at lower velocities. Specifically, *range*, *azimuthal angle*, *elevation angle*, and *signal to noise ratio* show a good fit of the simulation applied to the real scenario.

A clear conclusion can be done regarding RCS. The simulation model is limited to one fixed value for each target detected by the simulated radar which seems to be predefined by the ground truth according to the modeled target. In the case of CCRm this is a smaller hatchback vehicle which have been determined to have a radar cross section of 10 dB/m^2. In reality the value fluctuates quite significantly. From figures 4.1 - 4.2, We observe that the radar-reported detected RCS of the vehicle in front is right-skewed around $25\frac{dB}{m^2}$, with a range of varying RCS values observed throughout the test run.

There is a deviation in number of detections between the physical sensor and the simulated scenario. The simulated sensor observes significantly more detections across the entire time of the test run. Number of detections from the physical sensor increases later on in the time series, as the detected objects comes closer to the radar and reflection surface increases, several detection points on one object can be observed, where as when they are further away only point detections can be observed for the targets. The simulated sensor shows no such discrepancy, several detections occurs on targets independent on range. In lower velocities, figures 4.7 4.14, one can observe that the overall profile of number of detections in both cases follow the same structure, (there are peaks of detections occurring at the same time and equivalently at the same range from target).

The correlation figure E.7 gives a good overview of the models performance to mimic detections from the physical sensor.

5
Discussion

This chapter discusses the interpretation of key elements presented in Chapter 4. The chapter is divided into two sections with associated subsections, with each section and subsection discussing a specific area of the results. The methodology will be the central focus, examining how the method was applied and how it might be improved based on the results obtained. Additionally, this chapter will explore the fundamentals of the findings and suggest potential areas for future research. Doing so aims to thoroughly understand the results and their significance within the method created.

5.1 Method Evaluation

This section presents the methodology results, with each subsection representing a part of the method.

5.1.1 Metric Evaluation

The statistical metric is the most fundamental thing about creating this method to evaluate one radar model against one radar, since the statistical method lays the foundation of how the output is evaluated, and whether the method can be used as an evaluation tool. Furthermore, as mentioned in section 3.4, the selection of the metric was based purely on the literature read. Moreover, the chosen metric met nearly all the criteria identified as important in the literature, and the DVM incorporates several critical internal components. These components include, among others, the integration of P-Box, as detailed in Section 2.3.1. The DVM utilizes the distribution-free P-Box which serves as a modeling parameter to epistemic uncertainty within the models, and since the radar model is based on distributions instead of raw data, and the method created is also based on distributions, the P-Box accounts for the aleatory uncertainty as well. Hence, DVM is a good metric for handling these uncertainties, as mentioned before in the Method and Theory.

However, as the other statistical metrics mentioned were neither tested nor evaluated due to the DVM meeting the majority of criteria, it is plausible that an alternative statistical metric could perform equally well or even surpass the performance of the one utilized in this study. More about this in Chapter 6.

As outlined in the methodology, additional evaluation beyond the metric was con-

ducted using correlation and MLE plots to identify distributional differences and highlight significant dissimilarities between the datasets. Additionally, detection counts were analyzed to further clarify these differences. These steps proved highly beneficial.

Interpreting the results from the MLE analysis reveals that the real and simulated data often share the same distribution. This suggests that a certain degree of alignment between the two datasets exists. However, there are notable cases where the estimated distributions differ, yet the DVM still yields favorable results. This inconsistency may occur since in DVM the distributions are empirical and non-parametric while MLE tries to fit the best approximated true distribution of the data. In many cases where the results from the MLE differ, the characteristics of the distribution curves are similar. This inconsistency comes from the fact that no absolute parameters are set for the distribution functions to be fitted to the two data sets. Hence some estimated distribution functions, shifted due to their parameters, may overlap well with others. With this in mind, while the MLE may not serve as a strict model accuracy indicator, it can still provide valuable guidance in identifying distributional differences and supporting broader conclusions about data similarity.

As mentioned in both Chapter 4 and in the paragraph above, an analysis of the number of detections was conducted to get an additional parameter besides the six original radar parameters, and to gain further insight into where and when the radar model and radar sensor might diverge or cohere. The result proved to be very helpful for the method as a whole, particularly when emphasizing how the simulated data captures more detections throughout both scenarios, especially at greater distances. Combining the multi-visualization for each scenario and their corresponding subscenario, and the line graph displayed in Figure 4.7, Figure G.1 -G.3, Figure 4.14, and Figure G.4 - G.7, the difference in detections get obvious. The simulated model has a continuous detection stream, only fluctuating by a few detections. Meanwhile, the real scenario has a much lower initial detection rate and a large detection cloud at the end, where the desired object gets within a few meters. Therefore, since the number of detections is dependent on range, this parameter acts as the limiting factor and the reliability of the *range* parameter is reduced due to the difference. This limitation suggests that the parameter's accuracy and trustworthiness could be significantly improved if the number of detections were consistent across both datasets. In other words, a more balanced detection count would allow the *range* parameter to reflect the true performance more accurately and make the difference between the real and simulated data deviate less. In summary, there is a big difference between the RWS and SRM.

Finally, the output results from the DVM exhibit variations in both CAVM and bias across different measurements. These changes are structural differences in sensor behavior and environmental conditions. While model bias can influence the results, it is generally easier to identify and compensate for compared to model scattering error.

Model scattering error, on the other hand, arises from assumptions in the scattering

model, as explained in section 2.3.2, which are more difficult to correct and can significantly degrade the reliability of the measurements. For this reason, model scattering error is considered a more critical issue when evaluating model performance.

Therefore, when assessing how well specific parameters perform, the magnitude of the model scattering error should be a primary consideration. A parameter with low bias but high scattering error may still represent unreliable results, whereas a parameter with a higher bias can often be corrected more effectively. This difference is essential when determining whether a given DVM value should be regarded as trustworthy or not.

5.1.2 Scenario

Two scenarios are considered in this project CCRm and CBNAO. First, the CCRm proved helpful as a starting ground in developing the method. Its simplicity led to understanding how to extract and visualize data properly and conduct the statistical analysis. The CBNAO scenario was meant to test the method on a more complex environment. Introducing more targets led to further development of the method, making it more versatile.

Positioning of the VUT and all target objects connected to each scenario is essential to create the simulated environments. Since, as of now, no ground truth of exact spatial positions for the included objects or the VUT, the method is prone to some uncertainty and margin of error which may impact the end result. Even though the statistical evaluation is based solely on the detections and not any specific scenario the setup is essential. Especially for more advanced scenarios such as CBNAO, to accurately mimic the experiments properly in the simulation. While this leads to more time spent on data handling, it strengthens the arguments for why a project like this is necessary. Generating accurate and reliable mimics of the real world in the simulations may lead to fewer physical experiments, where much time and resources can be saved in the development and testing process.

5.1.3 Parsing

In section 3.2.6, the concern of mismatched timestamps is raised. Since these mismatches do not occur consequently across the entire scenario data frame and do not alter the data too much, the decision to remove duplicates was ignored. The piecewise step interpolation proved useful in up-scaling the simulated detection data arrays to the same length as the real sensor data while not disturbing the distribution.

Finding where in the tables the scenario is fully defined, as mentioned in section 3.2.4, can be hard. This is based on ambiguities in the movement of the VUT. To make up for this, it is crucial to search for when the VUT has reached scenario-specific velocity and where the yaw rate is close to zero.

5.1.4 Object Finding

The object finding process relies on the target tracking function described in subsection 3.2.1, but due to the structure of the HDF5 file, there is potential for errors and slight divergence in identifying the correct objects for evaluation—particularly in complex scenarios where objects are closely spaced or moving rapidly. These discrepancies necessitate accounting for significant margins of error when interpreting the results. Because the simulations are based on real-world scenarios, an initial challenge lies in accurately matching the positions and velocity vectors of the objects and the VUT vehicle. Despite these challenges, the target tracking functions have proven useful in reasonably mapping the correct objects and their attributes.

However, utilizing bounding boxes introduced certain challenges, primarily due to inconsistencies in the reference points used for positioning, as mentioned in section 3.2.1.1. This leads to inaccuracies in the estimated positions of the bounding boxes, with observed shifts of up to half a car length in the latitudinal and longitudinal direction. To diminish this issue, a margin of error equivalent to half a car length was added when capturing detections within the bounding boxes. For the specific scenarios addressed in this project, this adjustment proved effective and ensured reliable detection coverage.

Nevertheless, this approach may not generalize well to more constrained environments, such as narrow roads or areas with high object density. In such cases, the added margin could overlap with adjacent objects or boundaries, potentially leading to false positives by logging unwanted items as valid detections. This limitation highlights the need for more precise localization methods or adaptive margin strategies when applying the evaluation framework in more complex or cluttered settings.

5.2 Result

When discussing the results, it becomes clear that there is no definitive or straightforward answer. This ambiguity arises primarily from the absence of a clear benchmark or reference that defines what comprises a "good" or "bad" outcome. Without such a standard, the results cannot be judged in absolutes. Instead, they serve as indicators, pointing toward certain trends, parameter sensitivities, or conditions under which specific parameters may prove useful or relevant.
To create some sort of benchmark, the method needs to be tested on more data and more scenarios and many similar tests. If one was to use the results from this project as a benchmark today, the absolute best overlaps from the DVM result could be use as a starting baseline for what is a good metric value, hence a good overlap between the simulated detection data and data from the physical sensor. Observing the results from the figures and tables in sections 4.1 and 4.2 and appendix D and H, one can determine that the overall best results obtains a normalized DVM value of 0.1-0.2. Setting this as a benchmark for a good score can be a start, as the method gets tested further this baseline value may increase or decrease depending on further findings. It is also important to draw the distinction between bias and model scattering error when interpreting the end result, a higher DVM value with majority of bias is preferable over a lower metric value with majority model scattering error.

The core value of the results, lies not in delivering a final verdict but in guiding further research. They highlight potential directions for optimization or improvements, suggesting which parameters might be worth focusing on and under what circumstances the parameters could be most effective. Ultimately, the interpretation and application of these findings depend heavily on the user's specific context and objectives. It is up to the practitioner to determine how the insights align with their goals and how they might be integrated into their workflow or decision-making process.

6
Conclusion

This thesis has examined the challenges associated with real experimental data and the various uncertainties that arise from it. Improving the reliability of sensor simulation models is undoubtedly an urgent matter and a critical objective for accelerating the development and testing of APSS as well as AD&ADAS functions. The method presented here represents a small but meaningful step in that direction. Applying the DVM has proven valuable in offering a high-level understanding of the simulation model's strengths and limitations. It also reveals trends related to radar parameter behavior and the model's ability to replicate real-world sensor responses. While not exhaustive, this approach provides a foundation for more refined evaluation techniques and underscores the importance of continued research in this domain.

As the scenario increases in complexity, involving more objects and greater velocity differences, the similarity between simulation data and real-world data decreases for most of the parameters measured in this thesis.

Parameter-wise, SNR, *relative velocity*, and RCS seemed to perform the worst, receiving the highest DVM values throughout the different scenario(s) and subscenario(s). Meanwhile *range* receives the lowest DVM value, followed by *elevation* and *azimuth,* which are equally good when comparing between the two scenarios.

Due to the absence of a clear reference for what constitutes a "good" result, it is difficult to determine whether the outcomes are satisfactory or not. While the method may not be suitable for formal certification purposes, it can still be valuable for identifying trends or uncovering insights into the reliability and performance of parameters such as *range*, *azimuth*, SNR, RCS, *relVelocity*, and *elevation*. This exploratory approach may help assess how trustworthy these measurements are in practice.

6.1 Further Work

In practice, the method created and used in this project can be utilized for statistical analysis of several scenarios and other sensor data - essentially, any context involving a RWS and a SRM. Even though this project focuses on the FR, the method may be applied to analyze the entire radar system on VCC vehicles, or which ever sensor model, as long as the parameters are adjusted accordingly.

As mentioned in the discussion, there may be alternative statistical metrics that could outperform the DVM metric. Evaluating and implementing different metrics could potentially provide a more certain answer, thereby improving the overall effectiveness of the method.

For applications on more complex scenarios than what is covered in this thesis, some improvements to the method should be made. Data handling has been the most time-consuming process which is directly correlated with the complexity of the scenario setup. Some areas could be automated, for example, a search algorithm for the target distinction.

Another improvement is to systematically store the parsed, comparative data frames in a database for easy access to perform deeper analysis and modeling, which goes beyond the scope of this project. This could also reduce runtime for the method when re-running scenarios.

The main focus for further improvements is to make up for the bias. Since the scattering error captured by the CAVM consists of unknown modeling errors, the magnitude of the error will remain. However, adapting the use case of the model according to the bias captured by the method may lead to more similarities between the SRM and RWS data, hence a greater reliability in this area of radar system testing.

Bibliography

[1] "A Brief History of Autonomous Vehicles – from Renaissance to Reality | Mobileye Blog," [Online; accessed 2025-02-04]. [Online]. Available: https://www.mobileye.com/blog/history-autonomous-vehicles-renaissance-to-reality/

[2] "Multi-pillar approach for safety validation of automated vehicles - Simcenter," Jan. 2024, [Online; accessed 2025-01-22]. [Online]. Available: https://blogs.sw.siemens.com/simcenter/multi-pillar-approach-for-safety-validation-of-automated-vehicles/

[3] A. Stove, "Linear fmcw radar techniques," *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, pp. 343–350, 1992. [Online]. Available: https://digital-library.theiet.org/doi/abs/10.1049/ip-f-2.1992.0048

[4] "Evidence of Multiple Benefits from Driving Simulator Training," [Online; accessed 2025-01-22]. [Online]. Available: https://viragesimulation.com/documentation/evidence-multiple-benefits-driving-simulator-training/

[5] "NVIDIA Autonomous Vehicles Technology," [Online; accessed 2025-04-22]. [Online]. Available: https://www.nvidia.com/en-us/solutions/autonomous-vehicles/

[6] A. Stove, "Linear fmcw radar techniques," *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, pp. 343–350, 1992. [Online]. Available: https://digital-library.theiet.org/doi/abs/10.1049/ip-f-2.1992.0048

[7] "Why is Chirp Used in Radar?" [Online; accessed 2025-02-14]. [Online]. Available: https://resources.system-analysis.cadence.com/blog/why-is-chirp-used-in-radar

[8] "What is a FMCW Radar? - everything RF," [Online; accessed 2025-01-30]. [Online]. Available: https://www.everythingrf.com/community/what-is-a-fmcw-radar

[9] "Definition of AZIMUTH," Jan. 2025, [Online; accessed 2025-02-11]. [Online]. Available: https://www.merriam-webster.com/dictionary/azimuth

[10] F. Engineer, "Fmcw radars lec 5: Angle estimation," https://www.youtube.com/watch?v=Ok2HarUxS8o, 2018, [Online; accessed 2025-05-19].

[11] "Azimuth Resolution - an overview | ScienceDirect Topics," [Online; accessed 2025-02-11]. [Online]. Available: https://www.sciencedirect.com/topics/computer-science/azimuth-resolution

[12] A. Khalique, *Navbasics : watchkeeping & electronic navigation. Vol. 3*, 2nd ed. Livingston: Witherby Seamanship Inter, 2011.

[13] MATLAB, "Measuring angles with fmcw radar | understanding radar principles," https://www.youtube.com/watch?v=GpXF4wVQ-L4&t=517s, 2022, [Online; accessed 2025-05-19].

[14] F. Engineer, "Introduction to mmwave sensing : Fmcw radars," https://www.youtube.com/watch?v=8cHACNNDWD8, 2018, [Online; accessed 2025-05-20].

[15] A. Helmenstine, "Doppler Effect Definition, Formula, and Examples," May 2023, [Online; accessed 2025-02-12]. [Online]. Available: https://sciencenotes.org/doppler-effect-definition-formula-and-examples/

[16] "Radar Cross-Section - Radartutorial," [Online; accessed 2025-02-12]. [Online]. Available: https://www.radartutorial.eu/02.basics/Frequency%20Modulated%20Continuous%20Wave%20Radar.en.html

[17] Wikipedia contributors, "Radar cross section — Wikipedia, the free encyclopedia," 2024, [Online; accessed 12-February-2025]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Radar_cross_section&oldid=1258616101

[18] ——, "Signal-to-noise ratio — Wikipedia, the free encyclopedia," 2024, [Online; accessed 12-February-2025]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Signal-to-noise_ratio&oldid=1265000650

[19] "What is Signal to Noise Ratio and How to calculate it?" Jul. 2024, [Online; accessed 2025-02-12]. [Online]. Available: https://resources.pcb.cadence.com/blog/2020-what-is-signal-to-noise-ratio-and-how-to-calculate-it

[20] "Probability Theory | Probability in Maths with Examples & Formulas," Jan. 2021, [Online; accessed 2025-03-03]. [Online]. Available: https://www.geeksforgeeks.org/probability-theory/

[21] Wikipedia contributors, "Cumulative distribution function — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Cumulative_distribution_function&oldid=1286248864, 2025, [Online; accessed 2025-05-20].

[22] ——, "Empirical distribution function — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Empirical_distribution_function&oldid=1277923994, 2025, [Online; accessed 24-March-2025].

[23] "Compute Empirical Cumulative Distribution Function in R," Mar. 2023, [Online; accessed 2025-03-24]. [Online]. Available: https://www.geeksforgeeks.org/compute-empirical-cumulative-distribution-function-in-r/

[24] Wikipedia contributors, "Step function — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Step_function&oldid=1276070402, 2025, [Online; accessed 24-March-2025].

[25] ——, "Indicator function — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Indicator_function&oldid=1289416227, 2025, [Online; accessed 2025-05-20].

[26] "Bernoulli Distribution - Definition, Formula, Graph, Examples," [Online; accessed 2025-05-20]. [Online]. Available: https://www.cuemath.com/data/bernoulli-distribution/

[27] Wikipedia contributors, "Maximum likelihood estimation — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Maximum_likelihood_estimation&oldid=1286989864, 2025, [Online; accessed 2025-05-02].

[28] "R-Squared: Definition, Calculation, and Interpretation." [Online]. Available: https://www.investopedia.com/terms/r/r-squared.asp

[29] Wikipedia contributors, "Coefficient of determination — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Coefficient_of_determination&oldid=1277871013, 2025, [Online; accessed 2025-05-21].

[30] U. of Memphis, "Aleatory variability and epistemic uncertainty," 2023. [Online]. Available: http://www.ce.memphis.edu/7137/PDFs/Abrahamson/C05.pdf

[31] N. Siu and J. Marble, "Aleatory vs. epistemic uncertainties: Principles and challenges," 2011. [Online]. Available: https://www.nrc.gov/docs/ML1120/ML112000408.pdf

[32] "Aleatoric uncertainty," https://www.sciencedirect.com/topics/engineering/aleatoric-uncertainty, 2021, accessed: 2025-05-20.

[33] Wikipedia contributors, "Kolmogorov–smirnov test," https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test, 2025, [Online; accessed 2025-05-21].

[34] M. Faes, M. Daub, and M. Beer, "Engineering analysis with imprecise probabilities: a state-of-the-art review on p-boxes," in *Proceedings of the 7th Asian-Pacific Symposium on Structural Reliability and Its Applications*. University of Tokyo, 2020.

[35] "Common Types of Data Bias (With Examples) | Pragmatic Institute," [Online; accessed 2025-03-21]. [Online]. Available: https://www.pragmaticinstitute.com/resources/articles/data/5-common-bias-affecting-your-data-analysis/

[36] L. Elster, P. Rosenberger, M. Holder, K. Mori, J. Staab, and S. Peters, "Introducing the double validation metric for radar sensor models," *Automotive and Engine Technology*, vol. 9, p. 6, Jun. 2024, [Online; accessed 2025-02-10]. [Online]. Available: https://doi.org/10.1007/s41104-024-00143-5

[37] P. Rosenberger, "Metrics for specification, validation, and uncertainty prediction for credibility in simulation of active perception sensor systems," Ph.D. dissertation, Technische Universität Darmstadt, Darmstadt, January 2023. [Online]. Available: http://tuprints.ulb.tu-darmstadt.de/23034/

[38] "Cambridge dictionary," https://dictionary.cambridge.org/dictionary/english/simulation, 2025, [Online; accessed 2025-02-27].

[39] "Crash avoidance frontal collisions," [Online; accessed 2025-05-20]. [Online]. Available: https://www.euroncap.com/media/80155/euro-ncap-aeb-c2c-test-protocol-v431.pdf

[40] "Test protocol – aeb/lss vru systems," [Online; accessed 2025-05-20]. [Online]. Available: https://www.euroncap.com/media/80156/euro-ncap-aeb-lss-vru-test-protocol-v451.pdf

[41] "HDFView 3.3.2," [Online; accessed 2025-03-27]. [Online]. Available: https://support.hdfgroup.org/downloads/hdfview/hdfview3_3_2.html

[42] "Python Release Python 3.13.1," [Online; accessed 2025-03-25]. [Online]. Available: https://www.python.org/downloads/release/python-3131/

[43] "tkinter — Python interface to Tcl/Tk," [Online; accessed 2025-03-25]. [Online]. Available: https://docs.python.org/3/library/tkinter.html

[44] "matplotlib: Python plotting package," [Online; accessed 2025-03-25]. [Online]. Available: https://matplotlib.org

[45] M. M. Fréchet, "Sur quelques points du calcul fonctionnel," *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, vol. 22, no. 1, pp. 1–72, 1906. [Online]. Available: https://doi.org/10.1007/BF03018603

[46] SciPy Developers, "scipy.integrate.trapezoid — scipy v1.11.1 manual," 2023, accessed: 2025-06-12. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.trapezoid.html

[47] Wikipedia contributors, "Trapezoidal rule — wikipedia, the free encyclopedia," 2025, accessed: 2025-06-12. [Online]. Available: https://en.wikipedia.org/wiki/Trapezoidal_rule

[48] F. E. Nathanson and J. P. Reilly, "Frequency agility for radar target detection and tracking," *APL Technical Digest*, vol. 9, no. 6, pp. 2–12, 1970. [Online]. Available: https://secwww.jhuapl.edu/techdigest/content/techdigest/pdf/APL-V09-N06/APL-09-06-Nathanson.pdf

[49] N. Petrov, E. Yiğit, O. Krasnov, and A. Yarovoy, "Radar calibration by corner reflectors with mass-production errors," in *Proceedings of the 18th European Radar Conference (EuRAD)*. IEEE, 2022, pp. 253–256. [Online]. Available: https://pure.tudelft.nl/ws/portalfiles/portal/142953117/Radar_Calibration_by_Corner_Reflectors_with_Mass_production_Errors.pdf

[50] M. S. Greco, "Radar clutter modeling," https://docenti.ing.unipi.it/m.greco/esami_lab/Radar/Clutter_modeling.pdf, [Online; accessed 2025-05-23].

A

Appendix

Appendix A contains all plots of the non-normalized AVM. There is one plot for each case for the different scenarios, with each subplot displaying the result for each parameter.

A.1 Automatic Emergency Breaking System Car-to-Car Rear moving

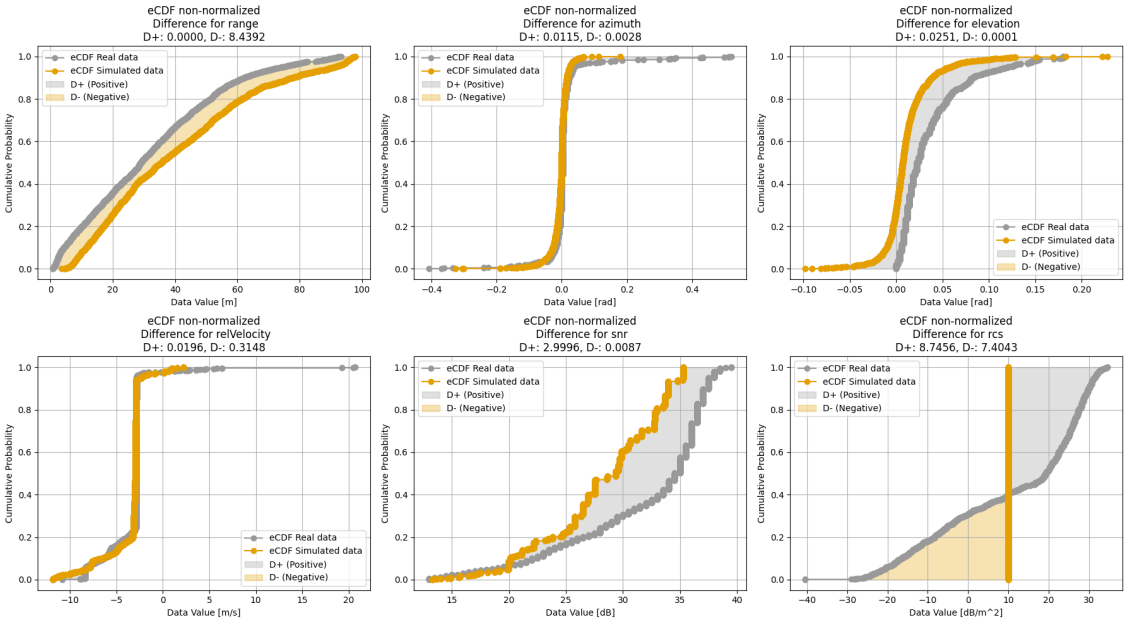

Figure A.1: AVM result for CCRm scenario. The orange curve represents the simulated data, and the gray curve represents the real data. During this scenario, VUT had a speed of $30\frac{km}{h}$ and GVT a speed of $20\frac{km}{h}$

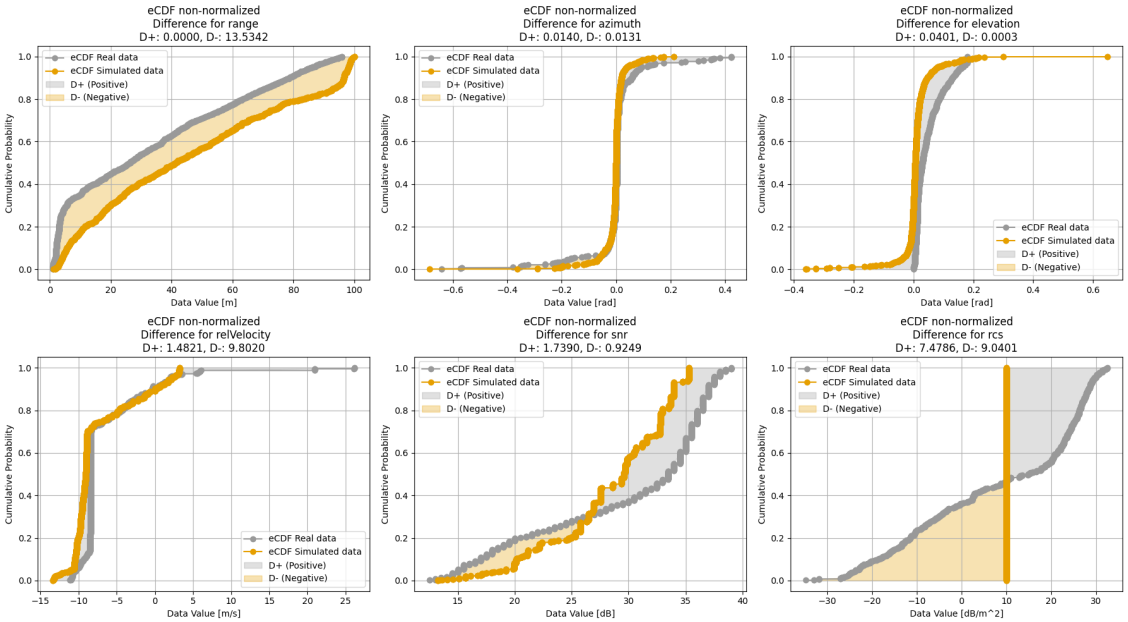

Figure A.2: An eCDF plot for non-normalized data where the VUT has a speed of $50\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves depend on the range of indices for the specific parameter.

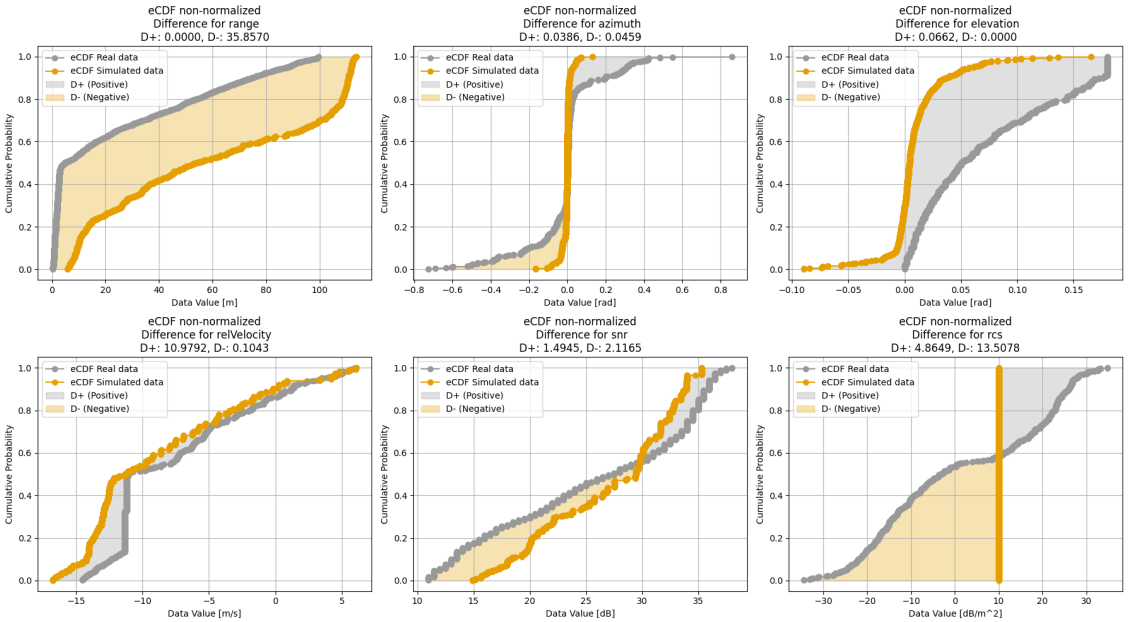

Figure A.3: An eCDF plot for non-normalized data where the VUT has a speed of $60\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves depend on the range of indices for the specific parameter.

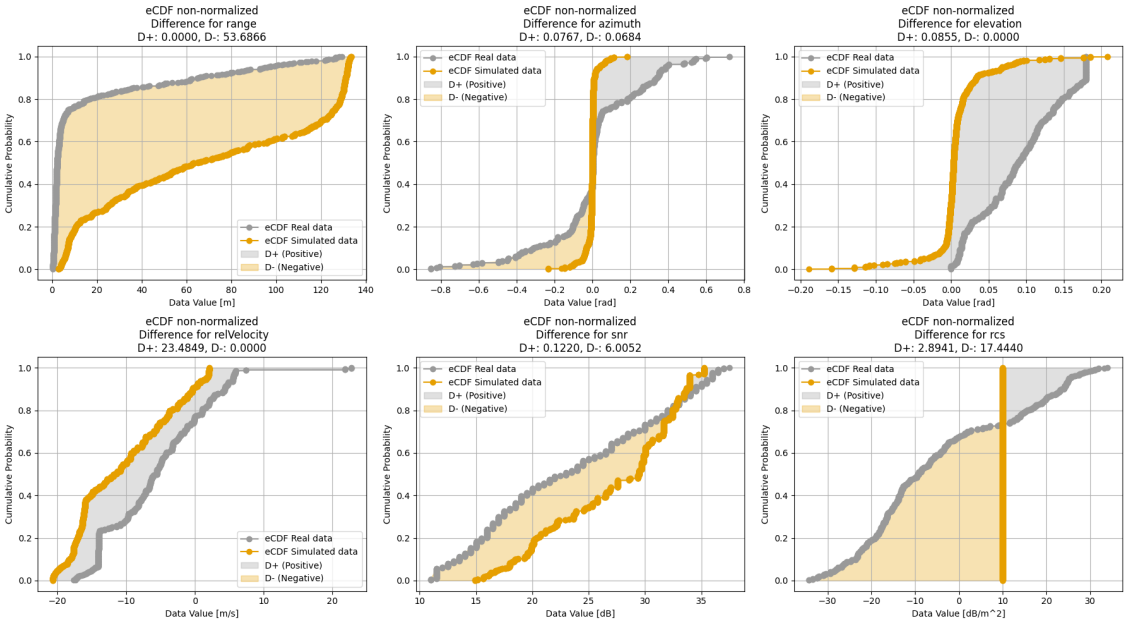

Figure A.4: An eCDF plot for non-normalized data where the VUT has a speed of $70\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves depend on the range of indices for the specific parameter.

A.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

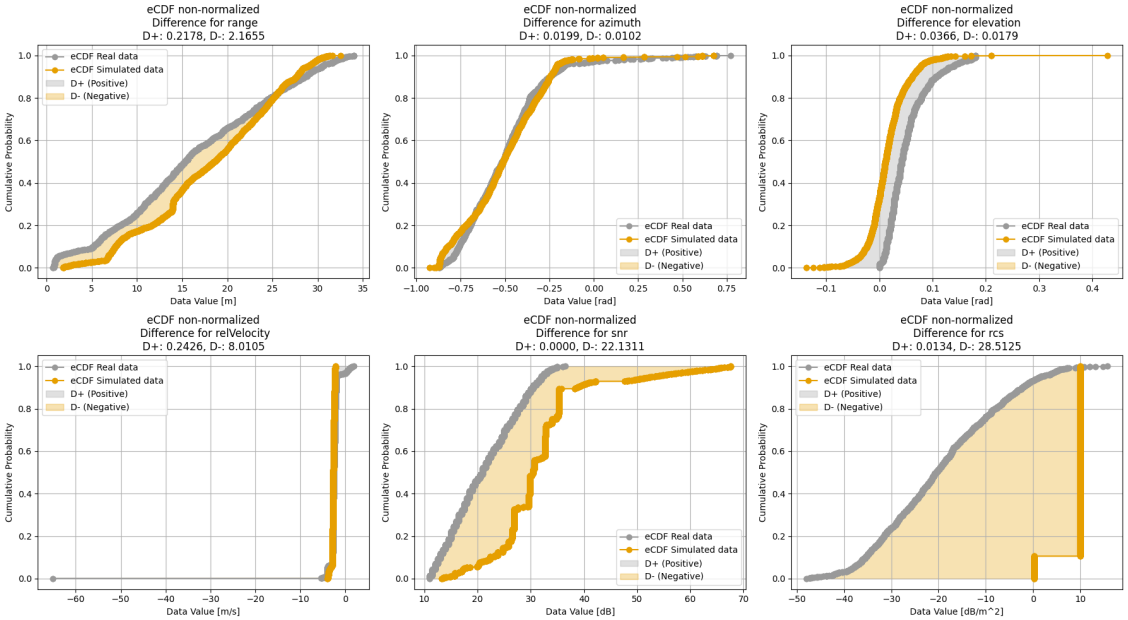

Figure A.5: AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $10\frac{km}{h}$.

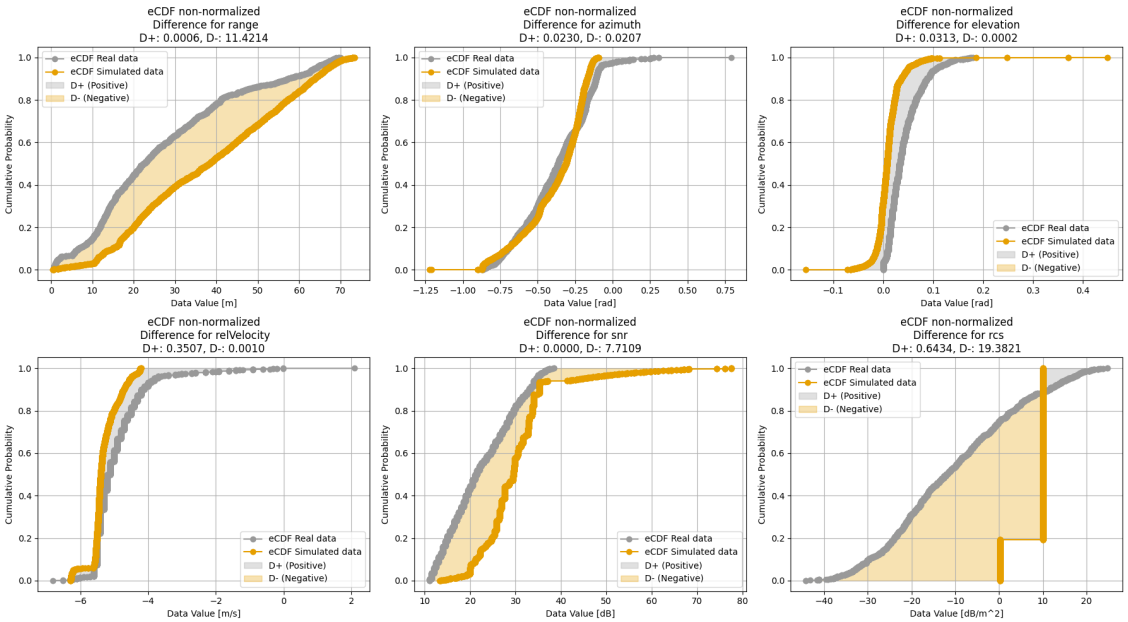

Figure A.6: AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $20\frac{km}{h}$.

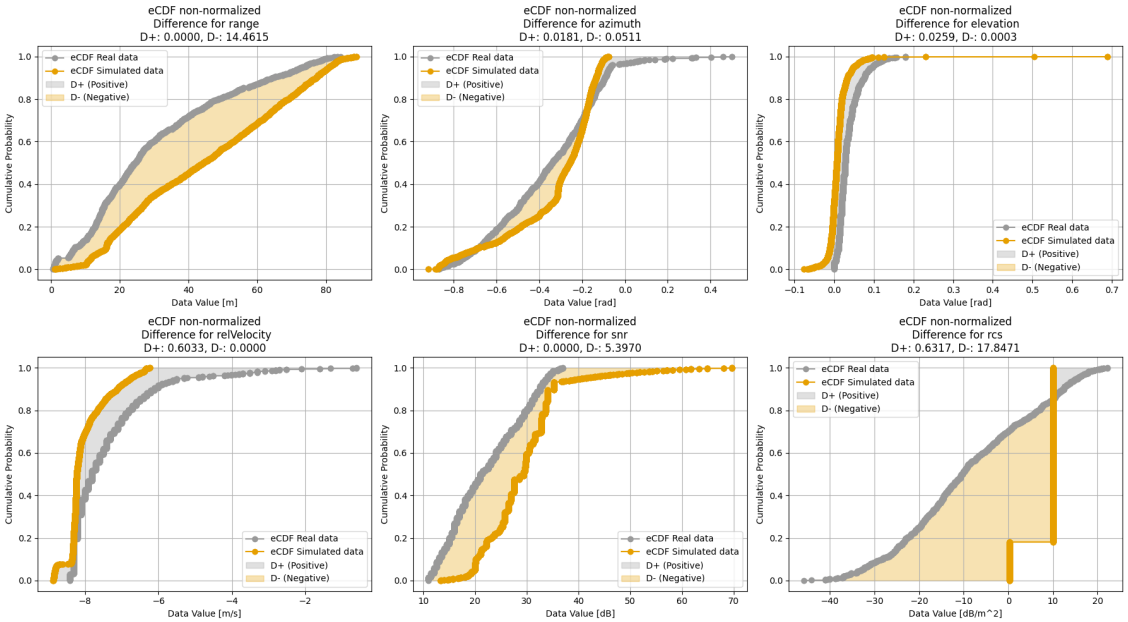

Figure A.7: AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $30\frac{km}{h}$.

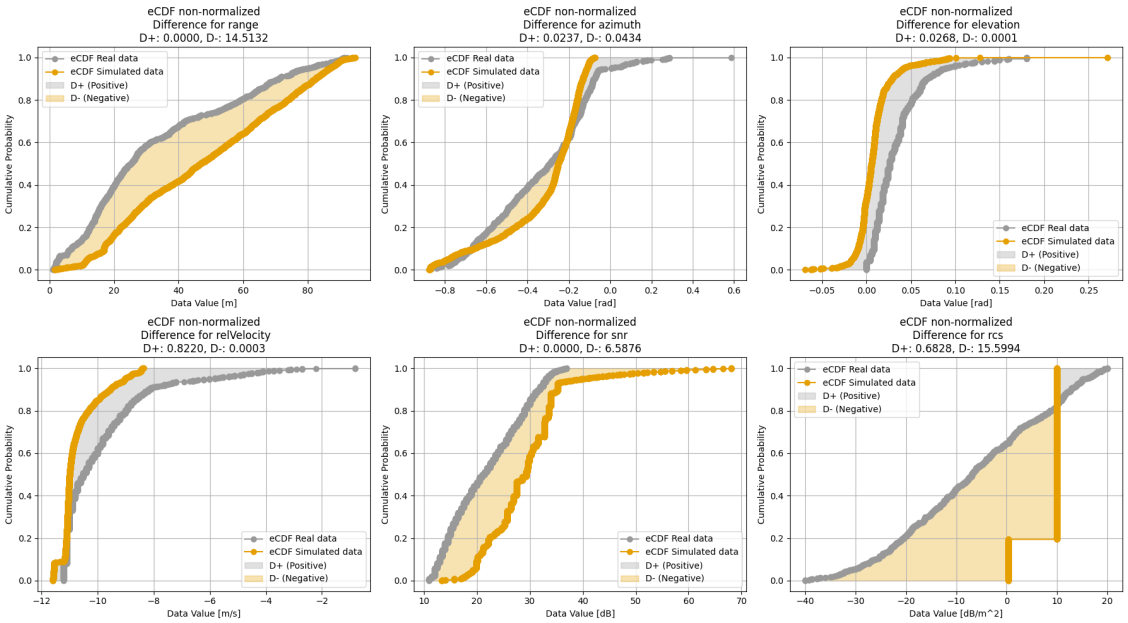

Figure A.8: AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $40\frac{km}{h}$.

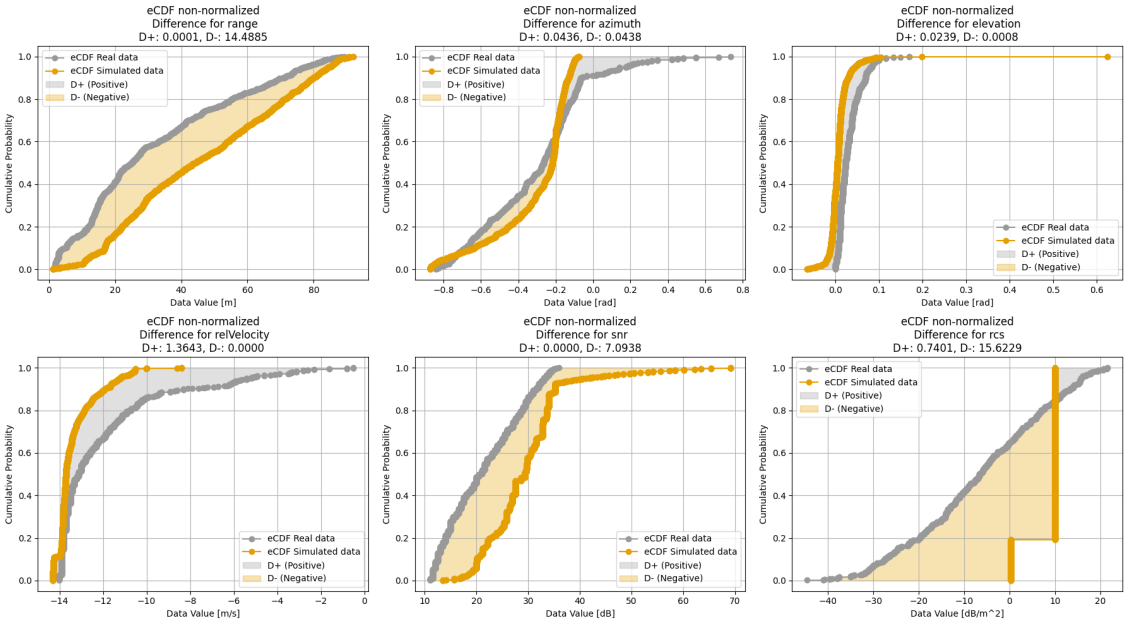

Figure A.9: AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $50\frac{km}{h}$.

B

Appendix

Appendix B contains all plots of the normalized AVM. There is one plot for each of the cases for the different scenarios.

B.1 Automatic Emergency Braking System Car-to-Car Rear moving

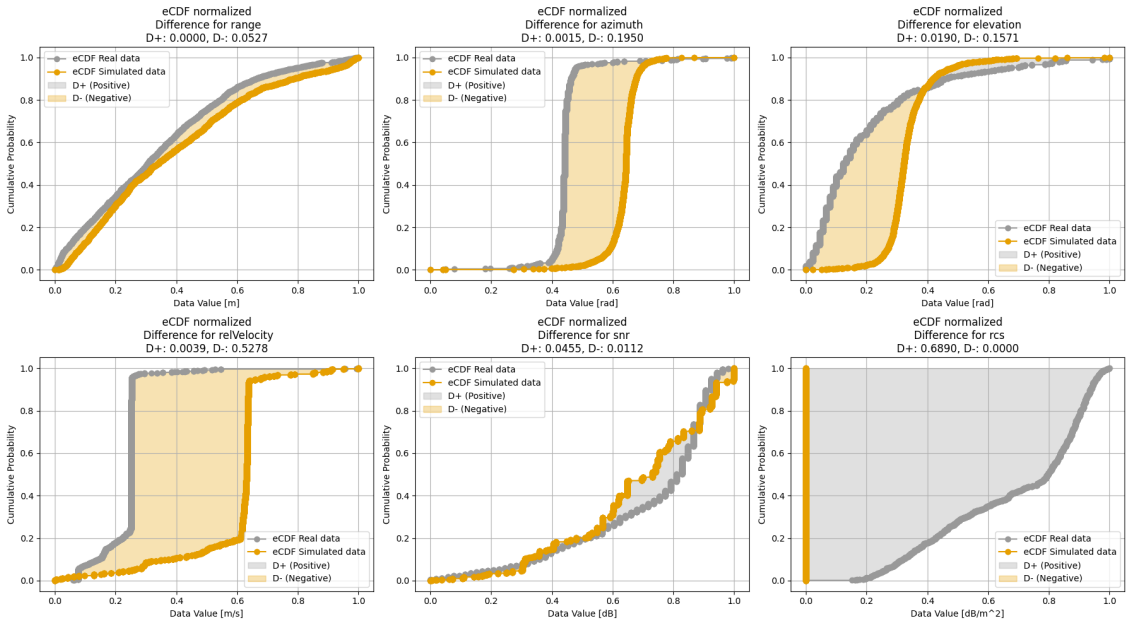

Figure B.1: Normalized AVM result for CCRm scenario. The orange curve represents the simulated data, and the gray curve represents the real data. During this scenario, VUT had a speed of $30\frac{km}{h}$ and GVT a speed of $20\frac{km}{h}$

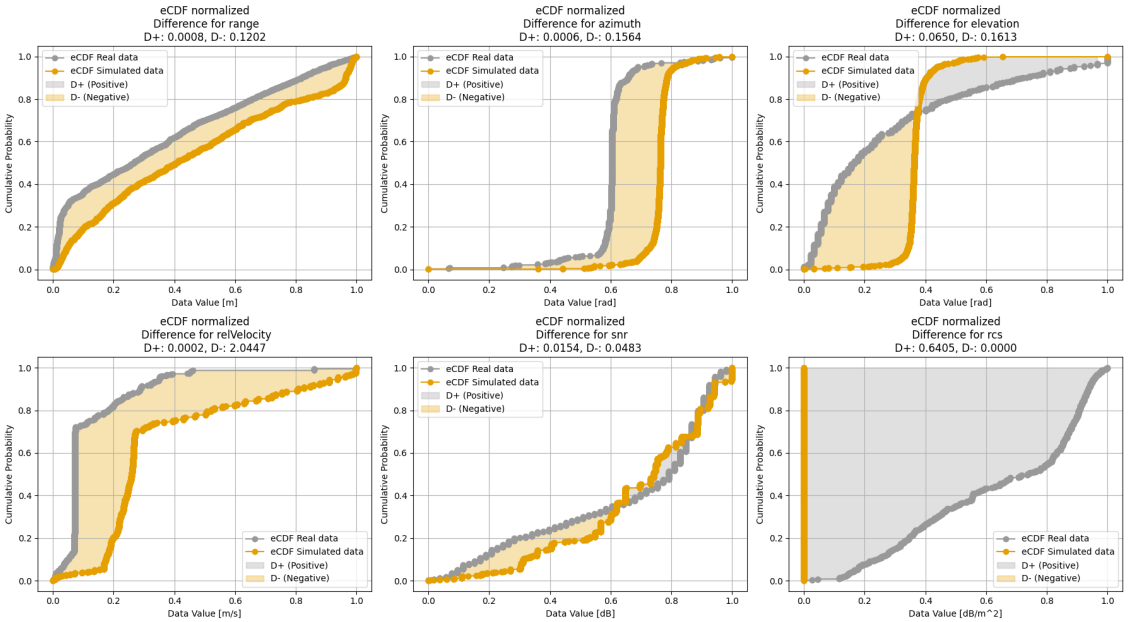

Figure B.2: An eCDF plot for normalized data where the VUT has a speed of $50\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves range from zero to one for all the parameters.

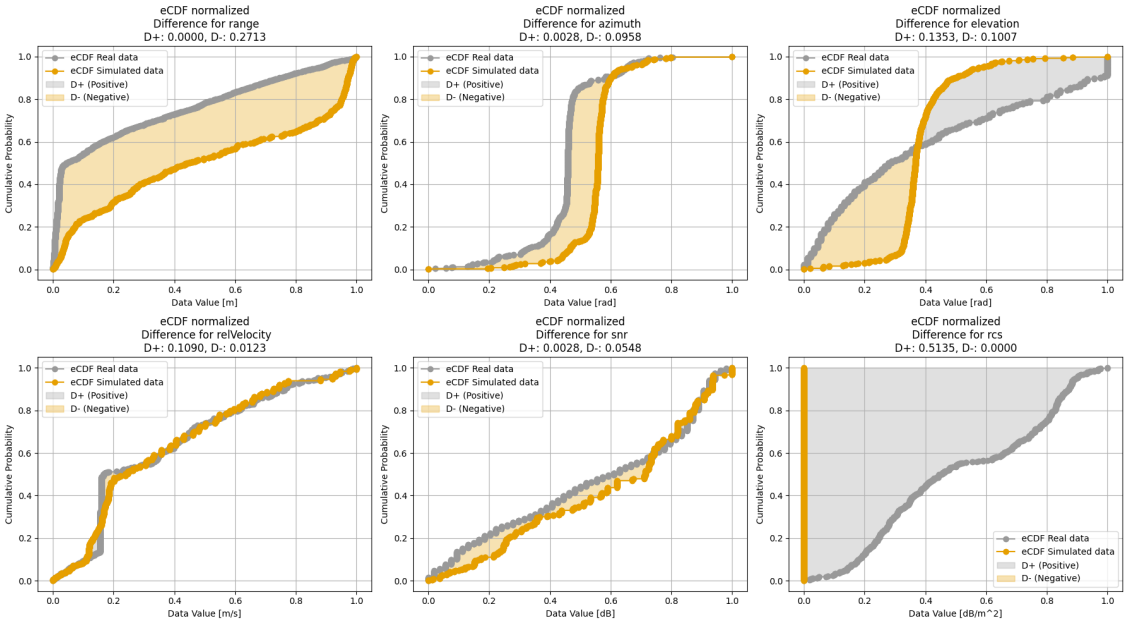

Figure B.3: An eCDF plot for normalized data where the VUT has a speed of $60\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves range from zero to one for all the parameters.

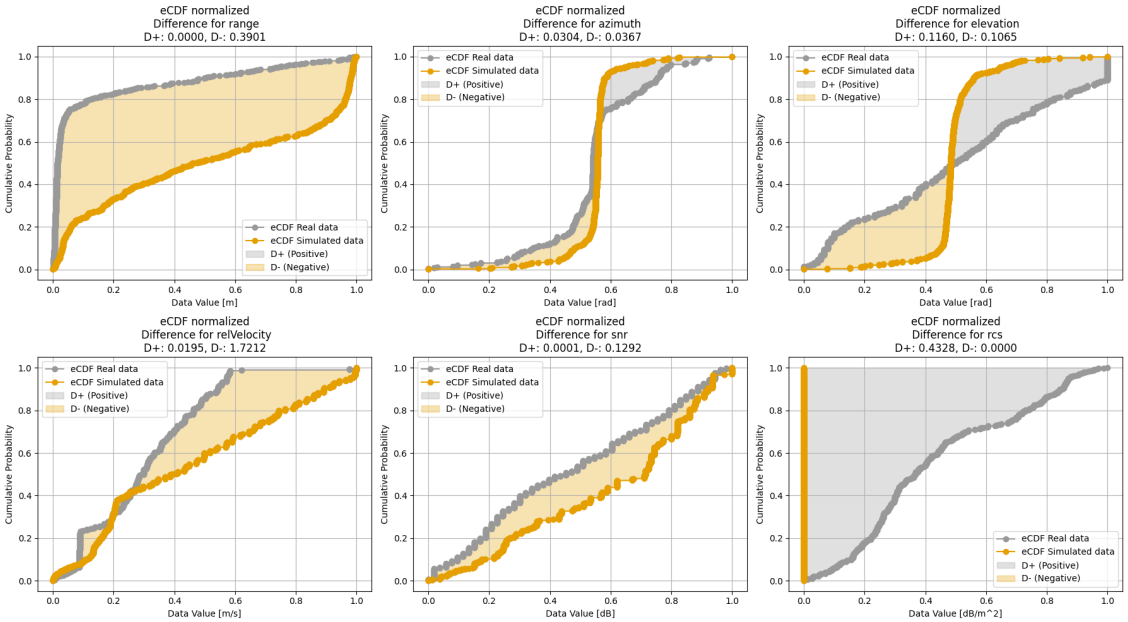

Figure B.4: An eCDF plot for normalized data where the VUT has a speed of $70\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ is shown. The orange curve represents the simulated data, while the gray curve represents the real data. When the orange curve lies below the gray curve, the area between them is shaded gray and represents $D+$. Conversely, when the orange curve is above the gray curve, the area between them is shaded orange, representing $D-$. These two areas together make up the AVM. The starting and ending points of the curves range from zero to one for all the parameters.

B.2 Automatic Emergency Braking System Car-to-Bicyclist Nearside Adult Obstructed

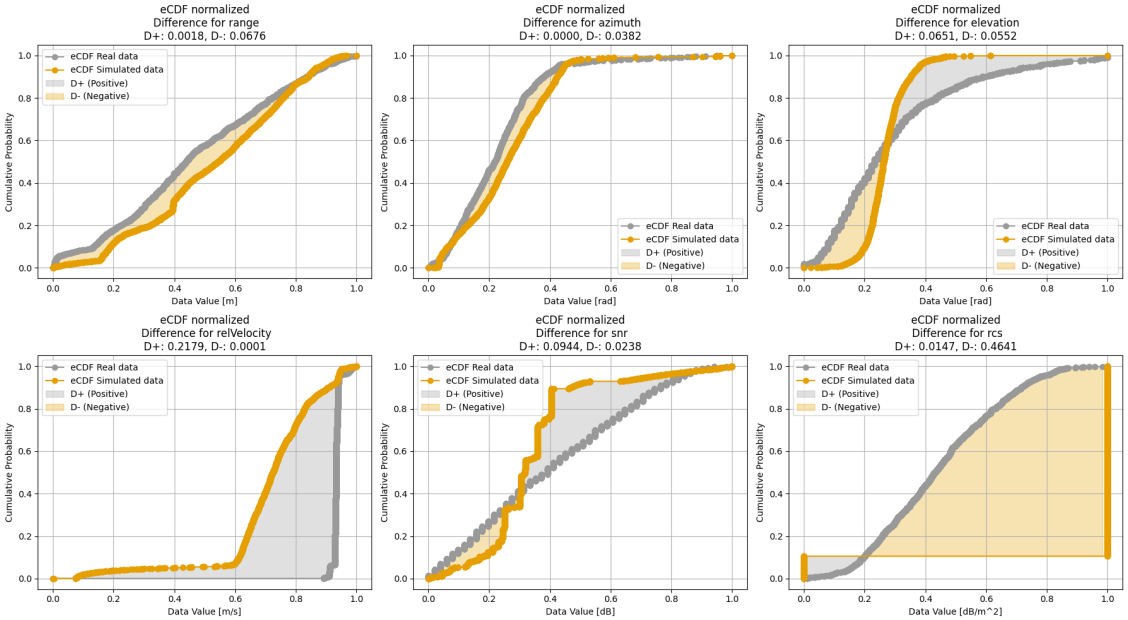

Figure B.5: Normalized AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, VUT had a speed of $10\frac{km}{h}$.

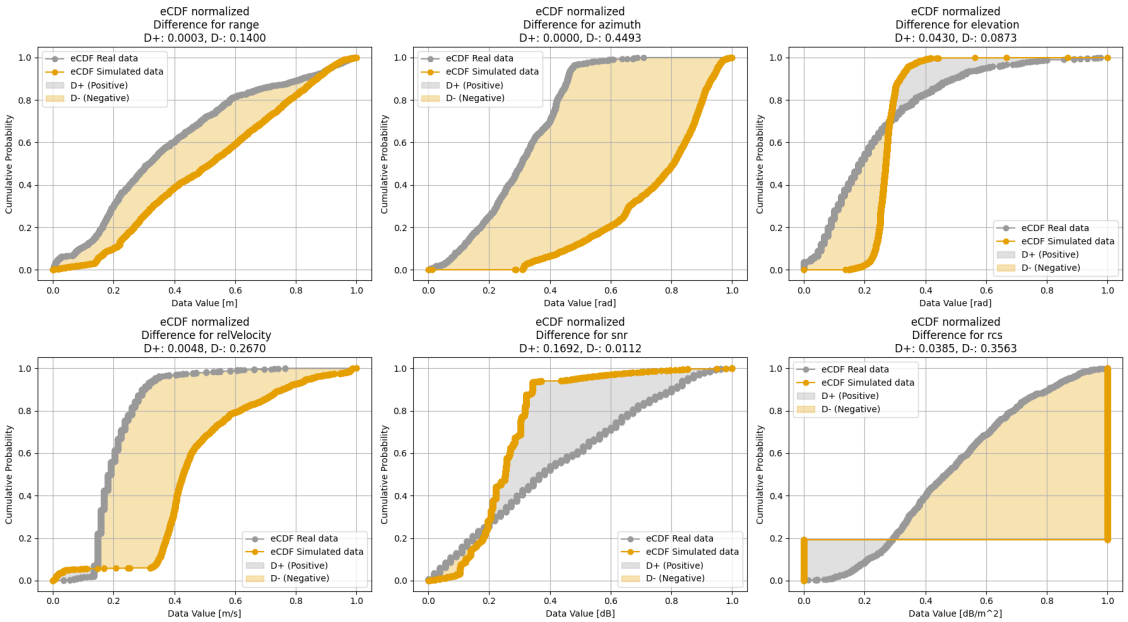

Figure B.6: Normalized AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, the VUT had a speed of $20\frac{km}{h}$.

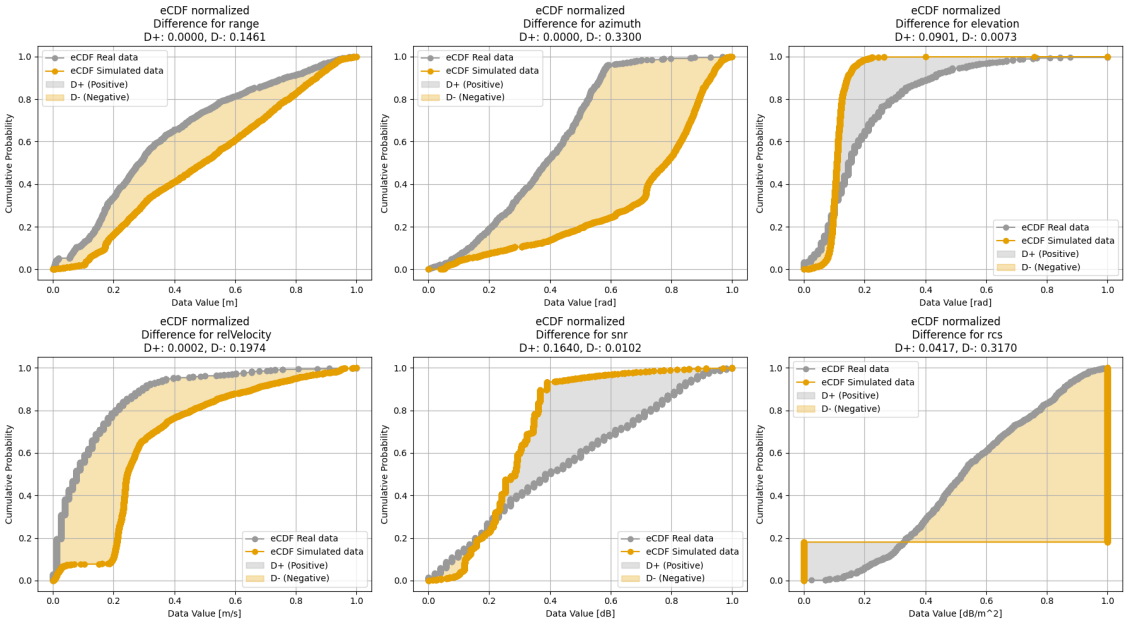

Figure B.7: Normalized AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, the VUT had a speed of $30\frac{km}{h}$.

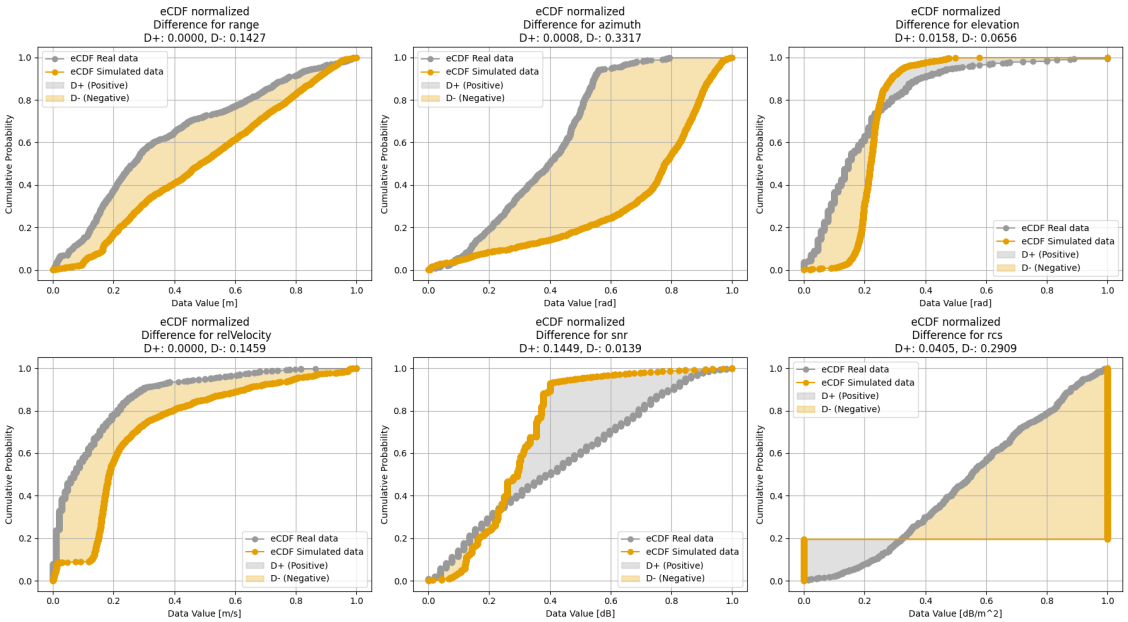

Figure B.8: Normalized AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, the VUT had a speed of $40\frac{km}{h}$.

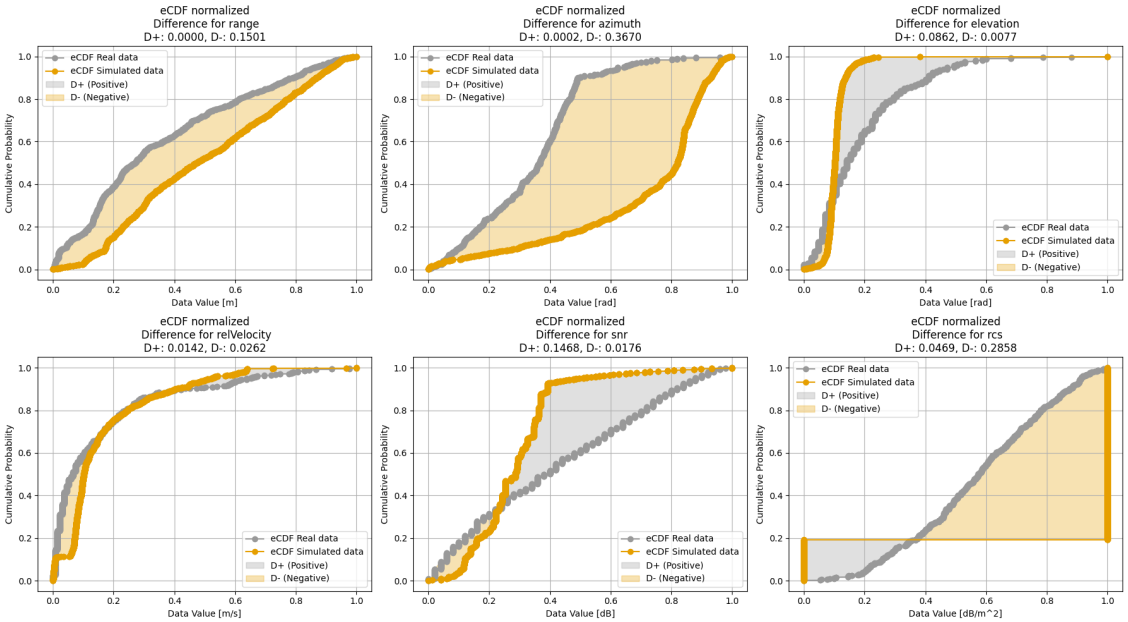

Figure B.9: Normalized AVM result for CBNAO scenario. The orange curve represents the simulated data, and the gray curve represents the real data for each parameter. During this scenario, the VUT had a speed of $50\frac{km}{h}$.

C
Appendix

Appendix C contains all plots of the non-normalized shifted curve for the CAVM. There is one plot for each of the cases for the different scenarios.

C.1 Automatic Emergency Breaking System Car-to-Car Rear moving

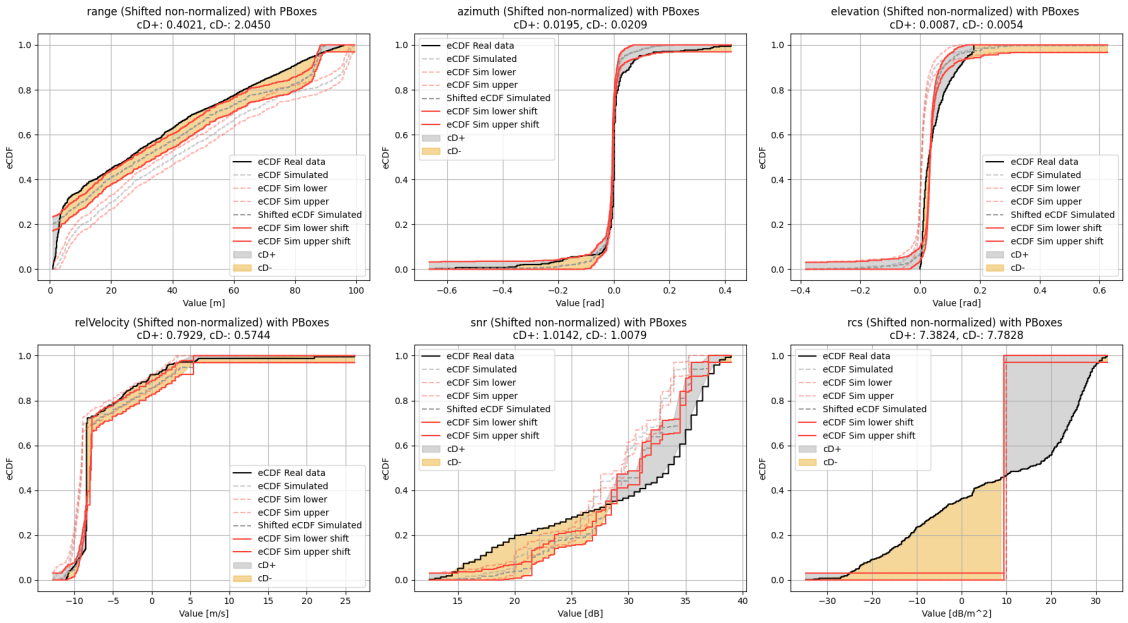

Figure C.1: Shifted and non-normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $50\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

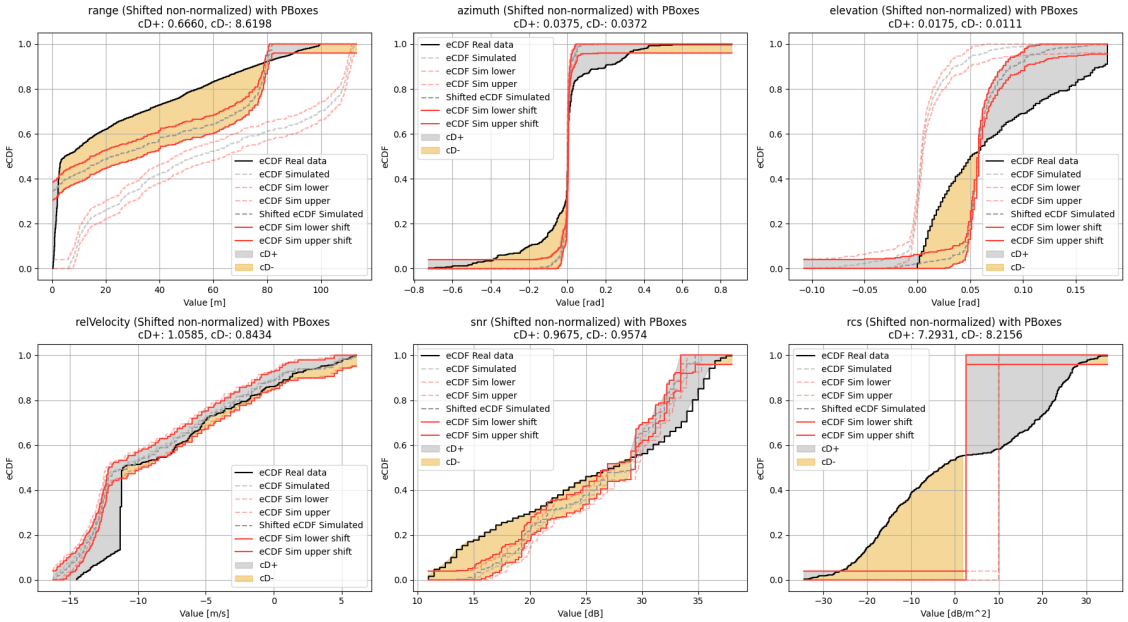

Figure C.2: Shifted and non-normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $60\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

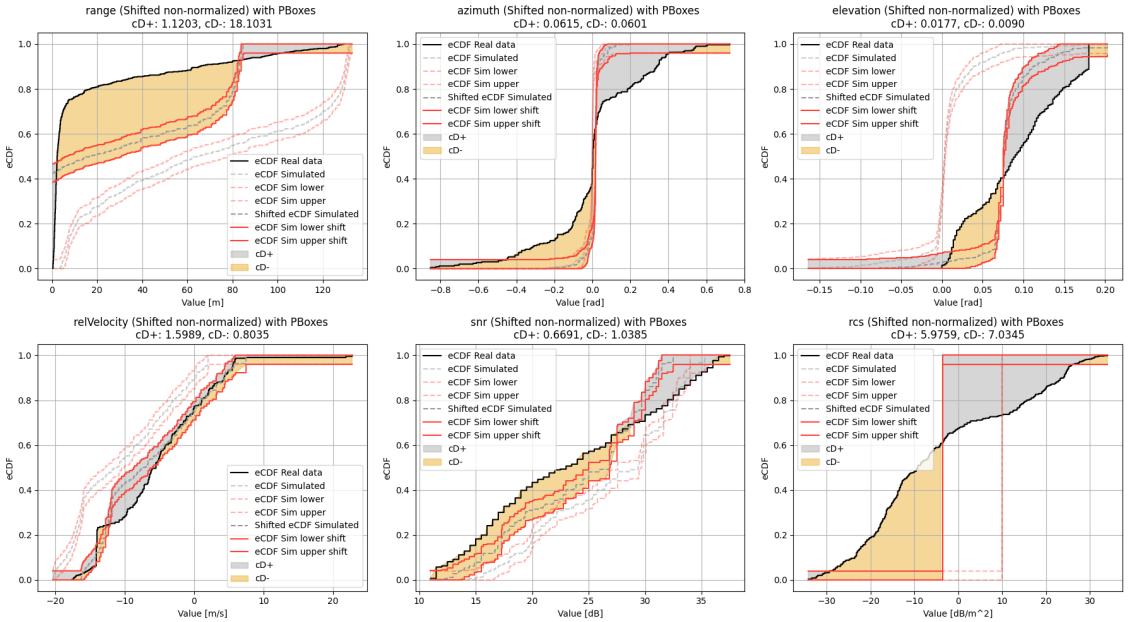

Figure C.3: Shifted and non-normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $70\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

C.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

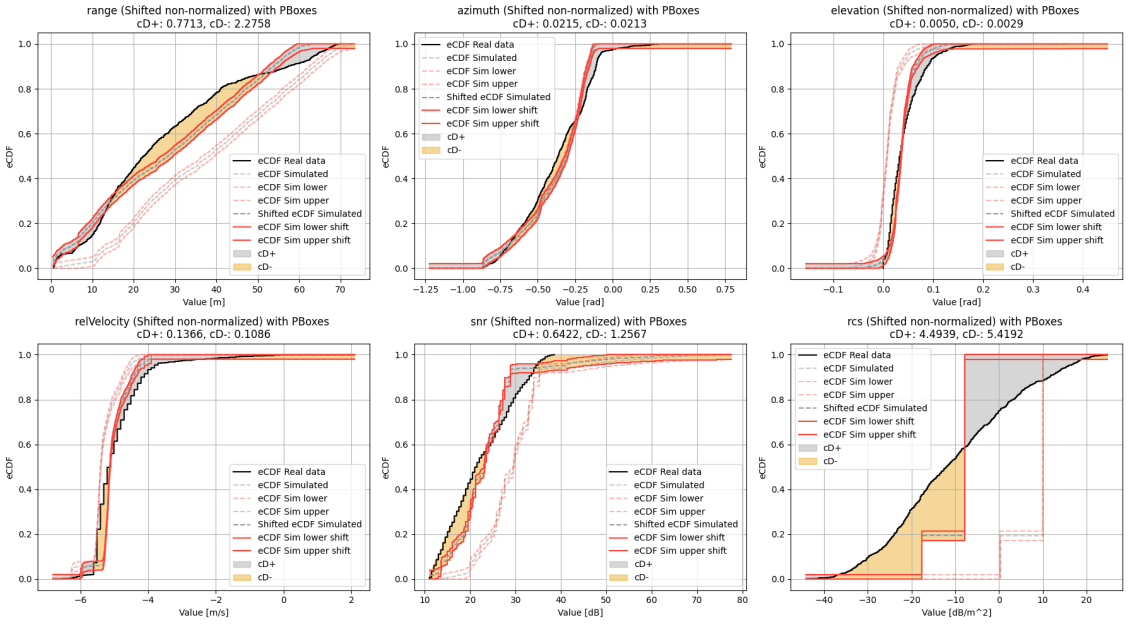

Figure C.4: Shifted and non-normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $20\frac{km}{h}$.

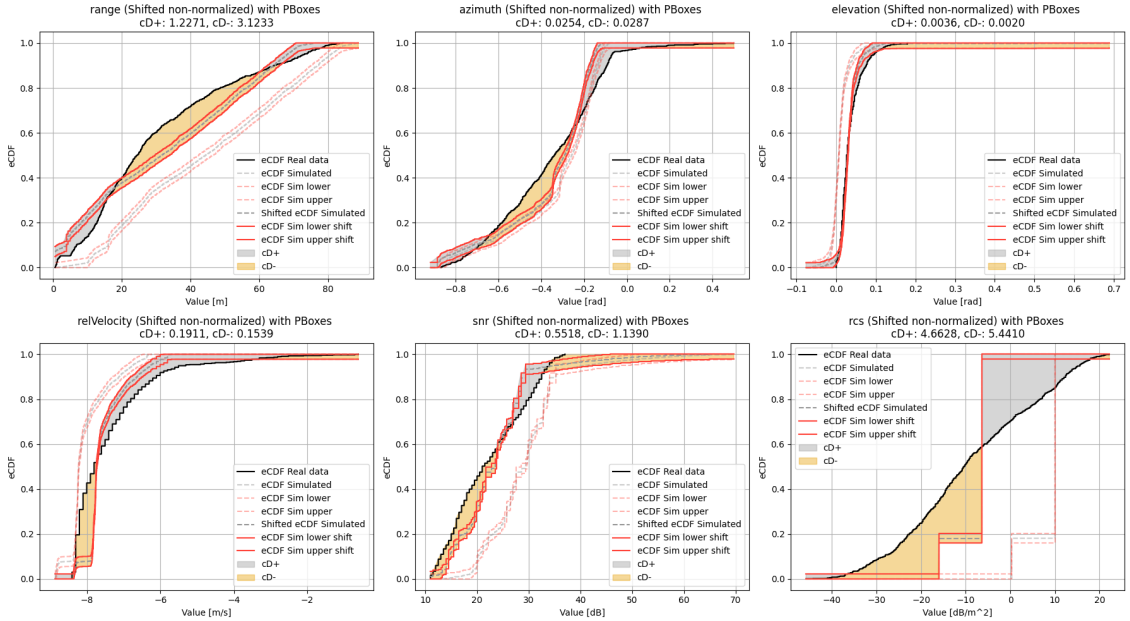

Figure C.5: Shifted and non-normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $30\frac{km}{h}$.

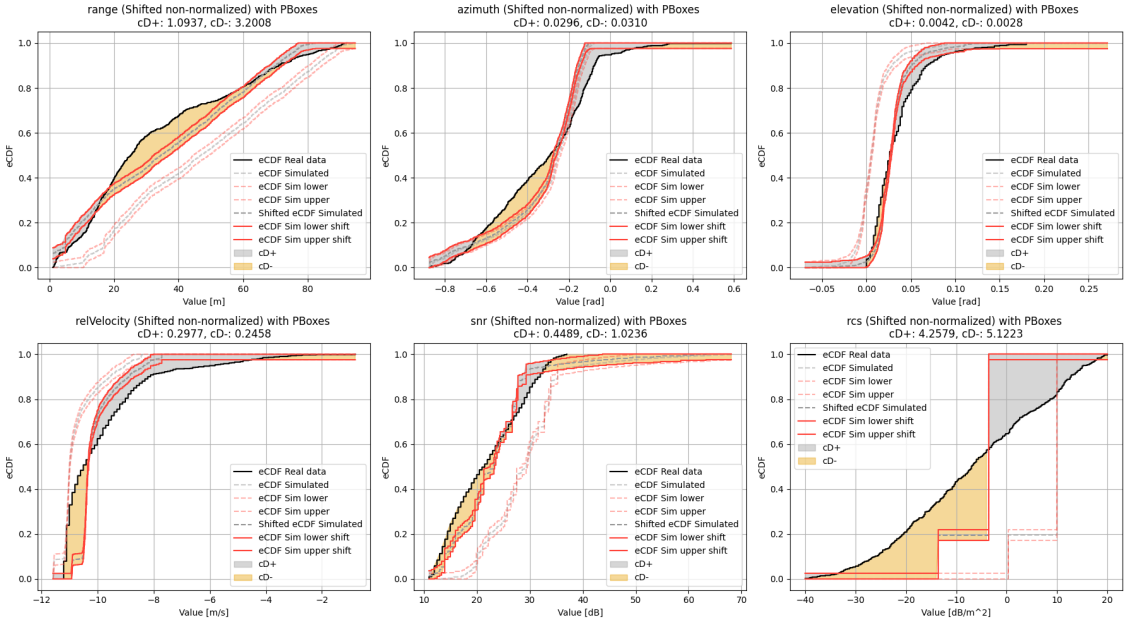

Figure C.6: Shifted and non-normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $40\frac{km}{h}$.

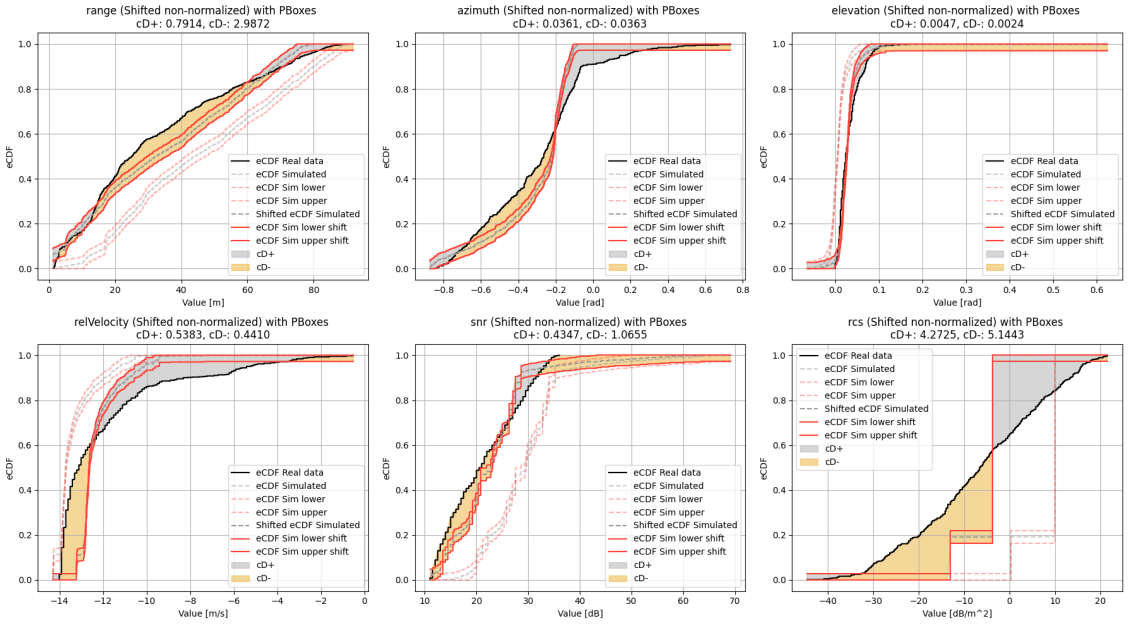

Figure C.7: Shifted and non-normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $50\frac{km}{h}$.

D

Appendix

Appendix D contains all plots of the normalized shifted curve for the CAVM. There is one plot for each of the cases for the different scenarios.

D.1 Automatic Emergency Breaking System Car-to-Car Rear moving

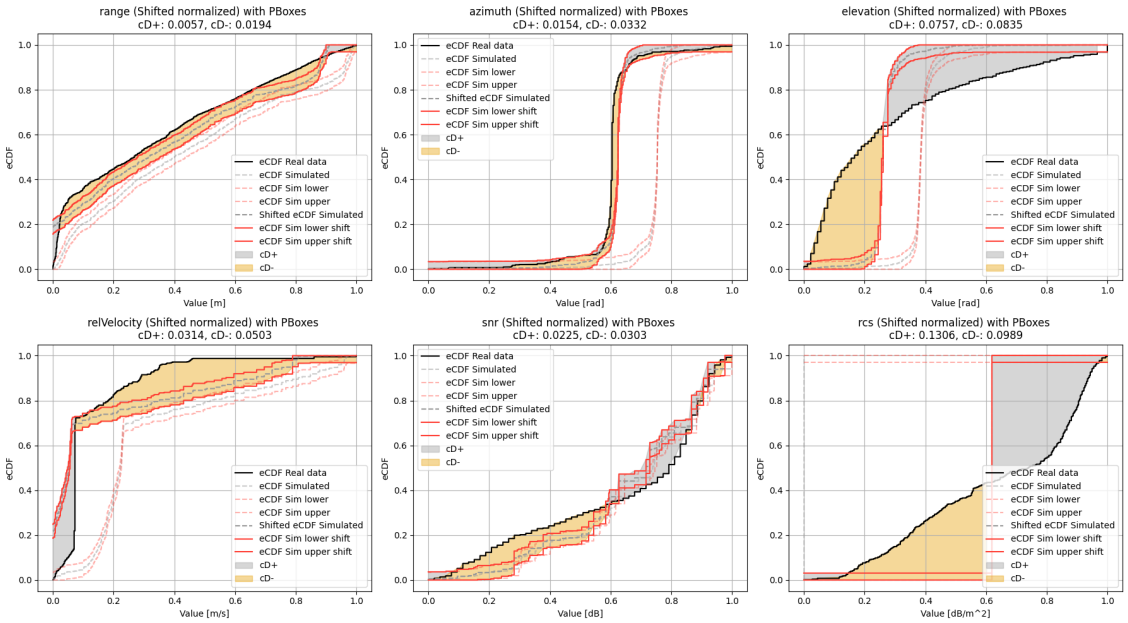

Figure D.1: Shifted and normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $50\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

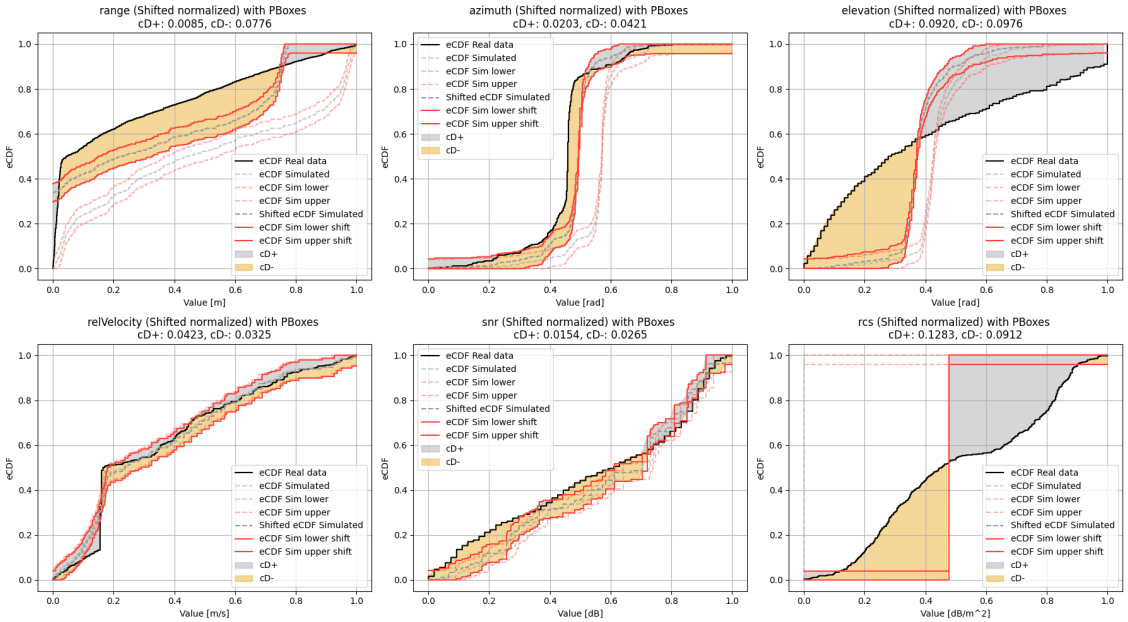

Figure D.2: Shifted and normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $60\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

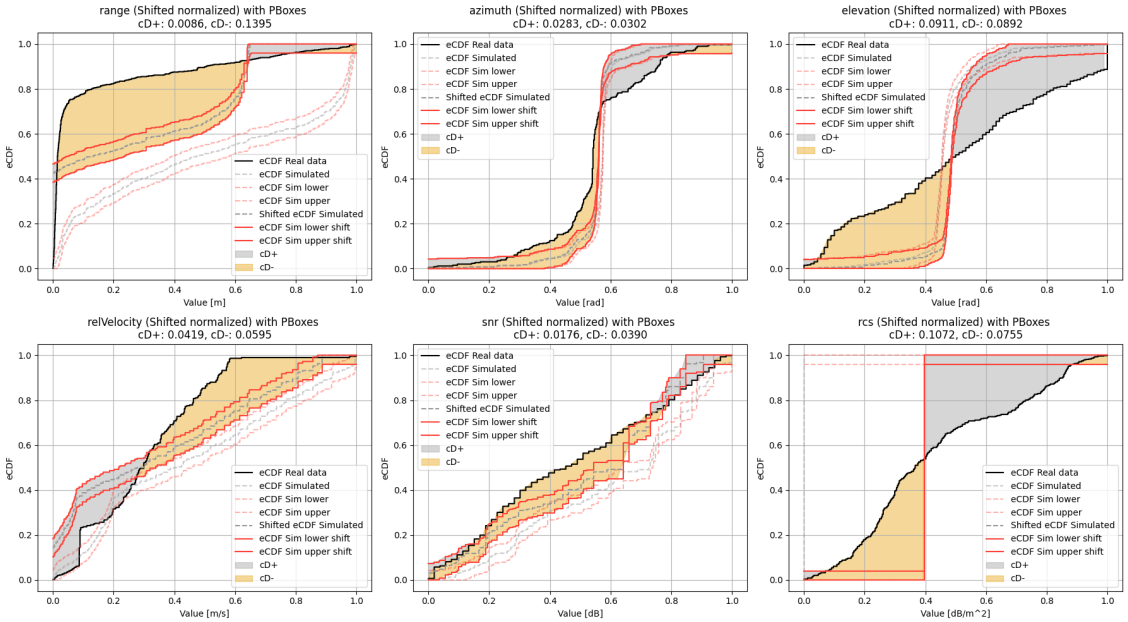

Figure D.3: Shifted and normalized DVM result for the CCRm scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $70\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$.

D.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

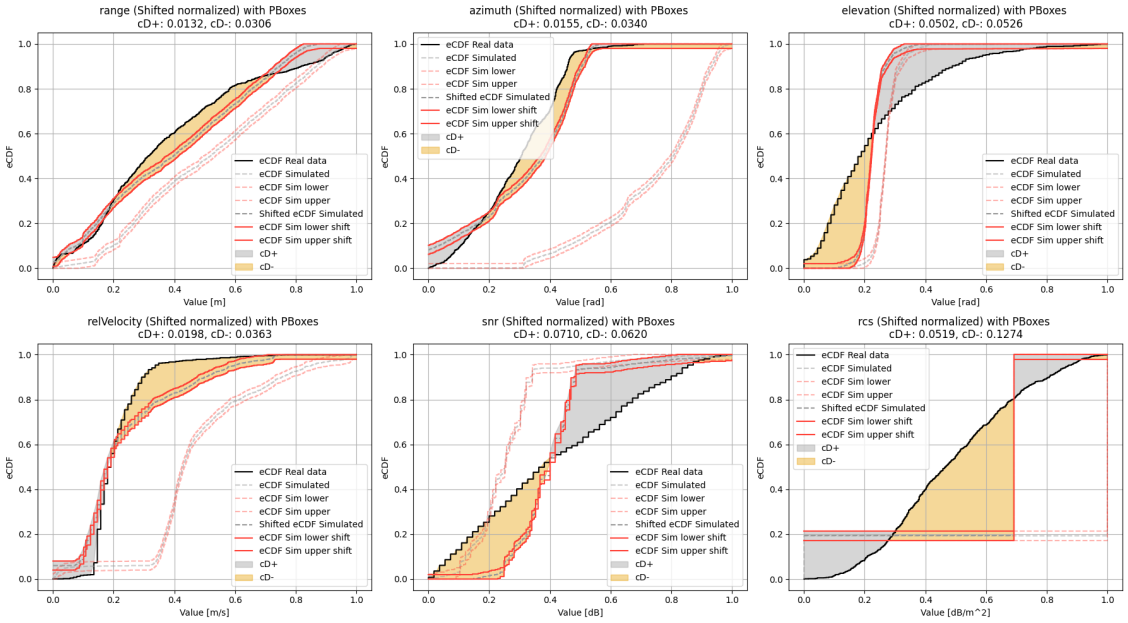

Figure D.4: Shifted and normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $20\frac{km}{h}$.

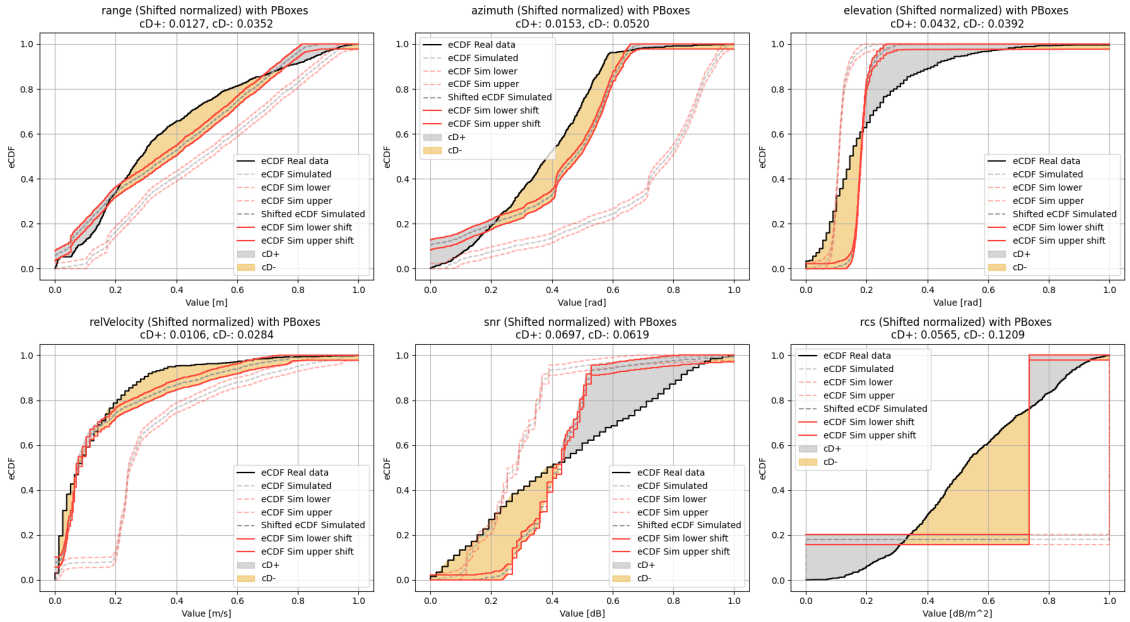

Figure D.5: Shifted and normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $30\frac{km}{h}$.

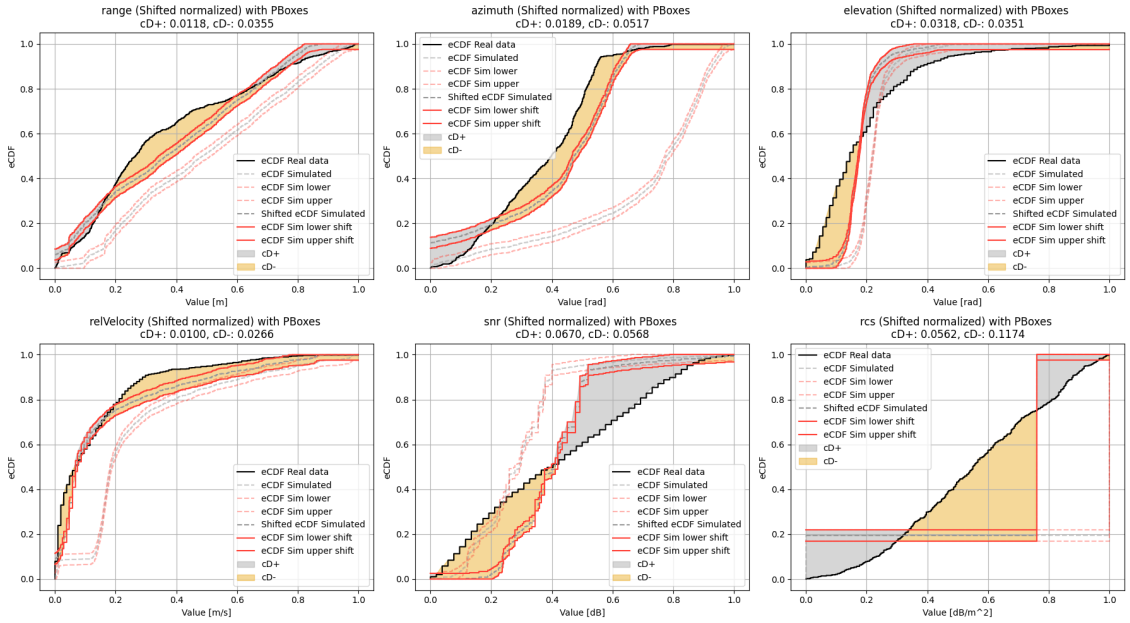

Figure D.6: Shifted and normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $40\frac{km}{h}$.

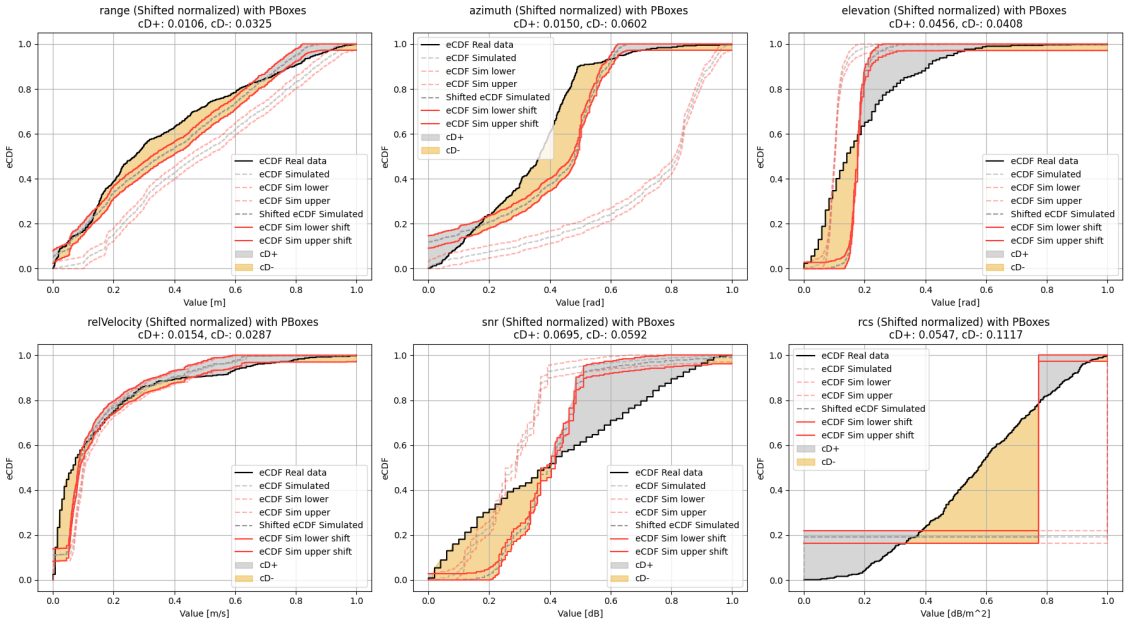

Figure D.7: Shifted and normalized DVM result for the CBNAO scenario. The dashed dark gray curve represents the shifted simulated data, the solid black line shows the real data, and the dashed light gray line corresponds to the non-shifted simulated data. The red solid and dashed lines indicate the shifted and non-shifted simulation bounds, respectively. During this scenario, the VUT had a speed of $50\frac{km}{h}$.

E

Appendix

Appendix E contains all the R-squared correlation plots. One plot is for each of the scenarios.

E.1 Automatic Emergency Breaking System Car-to-Car Rear moving

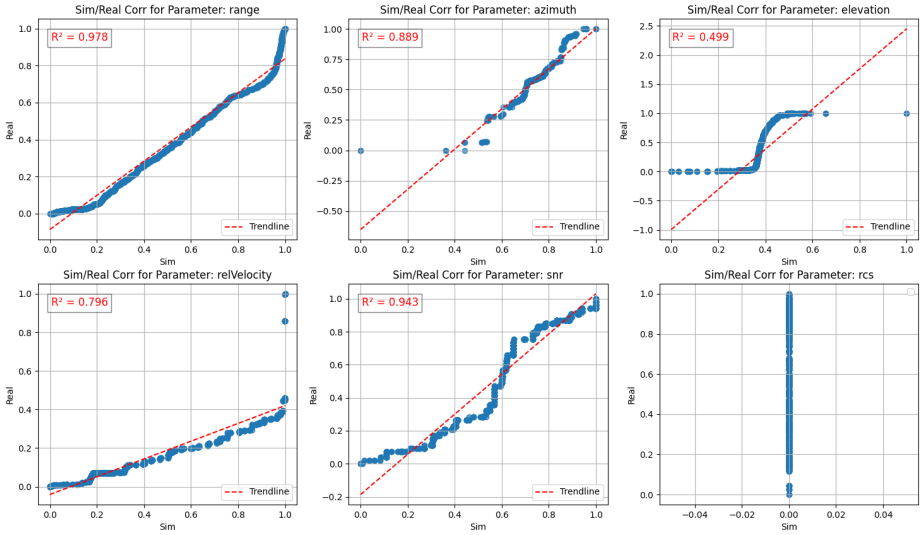

Figure E.1: Correlation plot illustrating the R^2 value for the different parameters for VUT speed $50\frac{km}{h}$ GVT speed of $20\frac{km}{h}$ for the CCRm scenario.

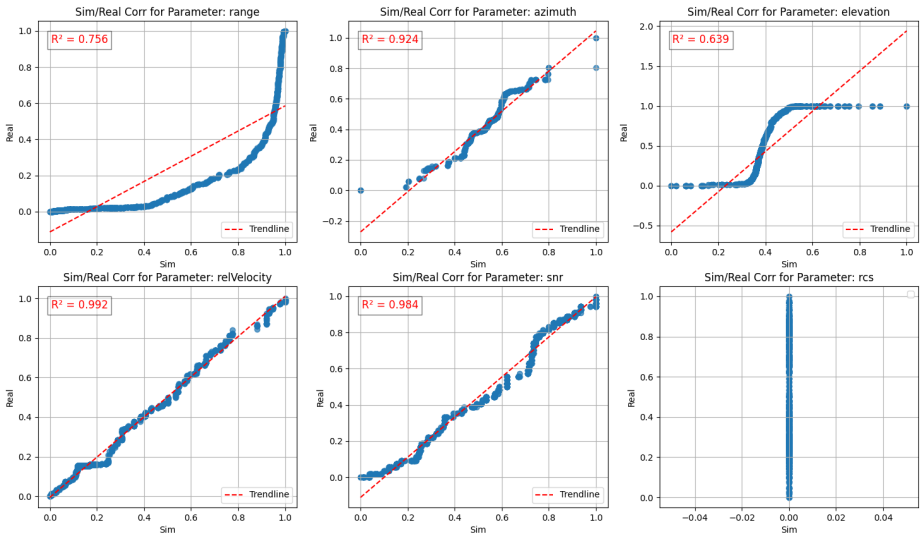

Figure E.2: Correlation plot illustrating the R^2 value for the different parameters for VUT speed $60\frac{km}{h}$ GVT speed of $20\frac{km}{h}$ for the CCRm scenario.

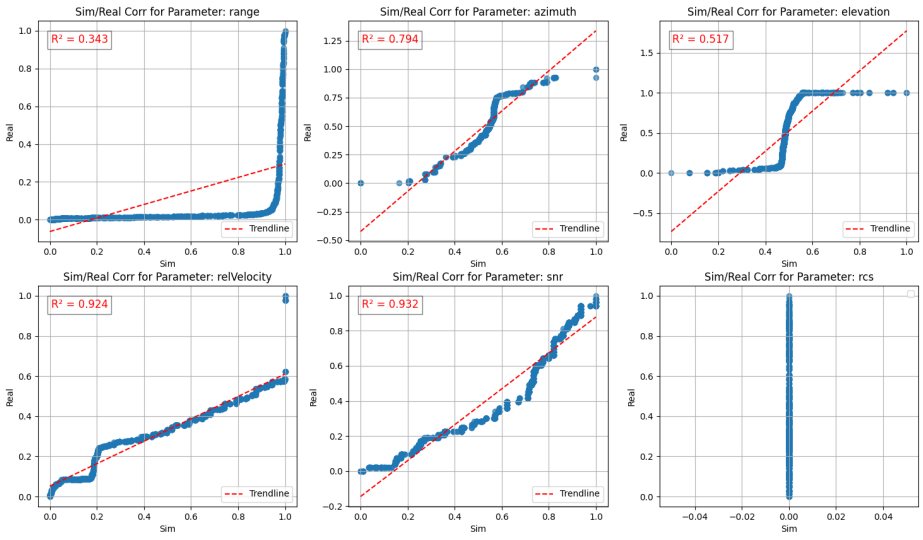

Figure E.3: Correlation plot illustrating the R^2 value for the different parameters for VUT speed $70\frac{km}{h}$ and GVT speed of $20\frac{km}{h}$ for the CCRm scenario.

E.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

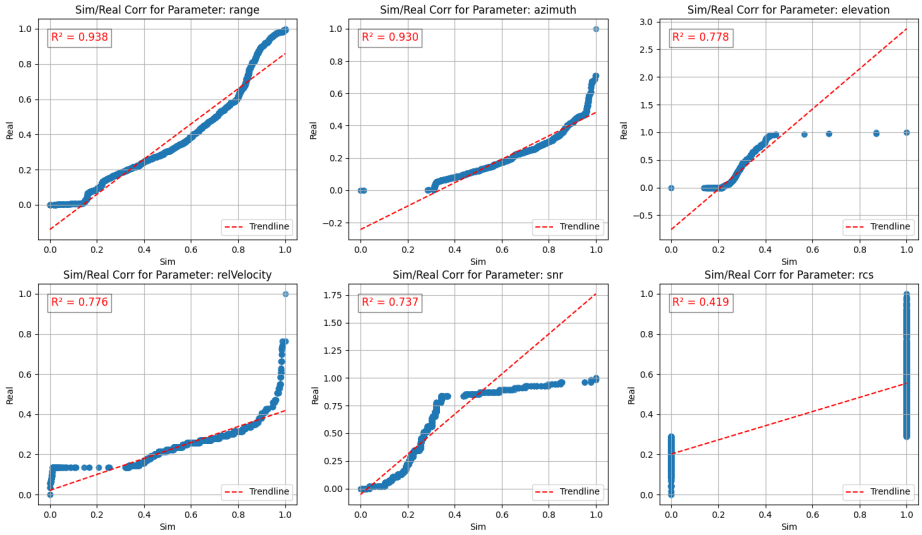

Figure E.4: Correlation between the same parameter for simulated data against real data.

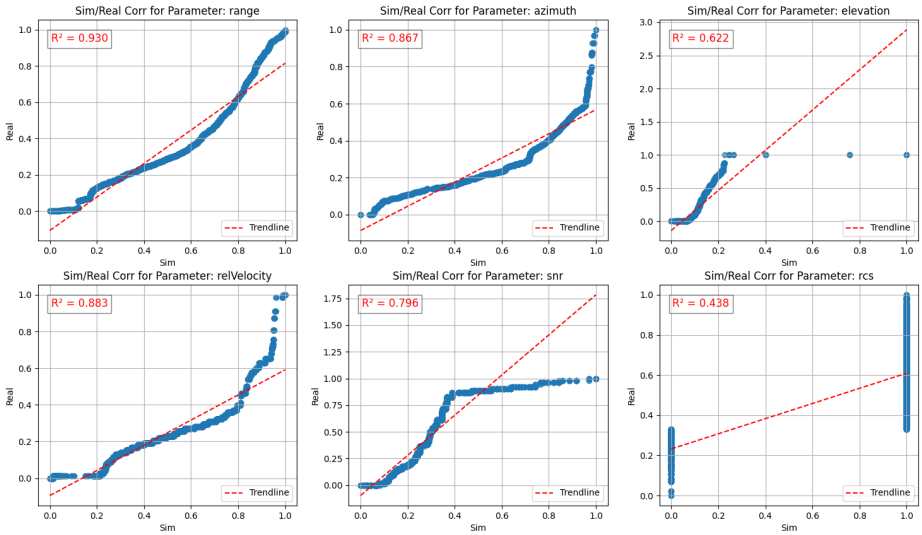

Figure E.5: Correlation between the same parameter for simulated data against real data.

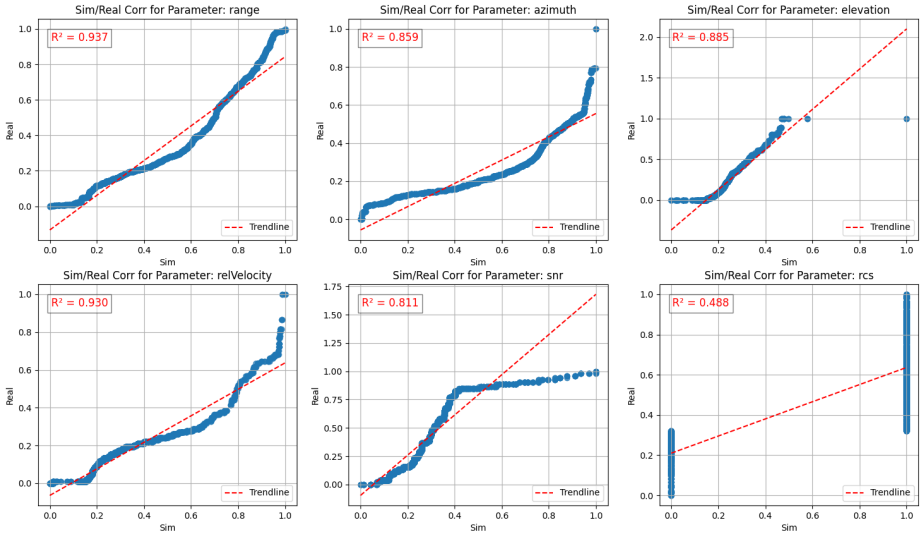

Figure E.6: Correlation between the same parameter for simulated data against real data.

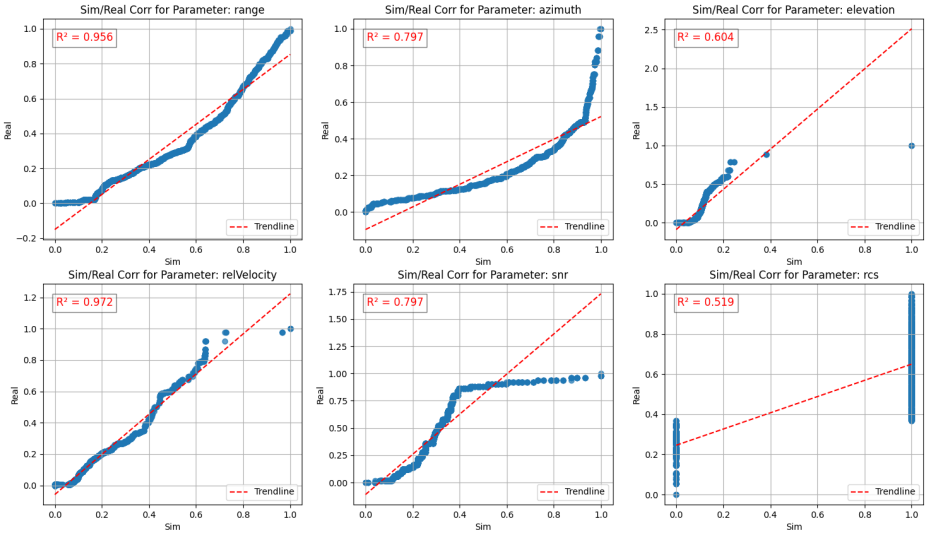

Figure E.7: Correlation between the same parameter for simulated data against real data.

F

Appendix

Appendix F contains all plots of the MLE for the non-normalized and non-shifted curve. There is one plot for each of the cases for the different scenarios.

F.1 Automatic Emergency Breaking System Car-to-Car Rear moving

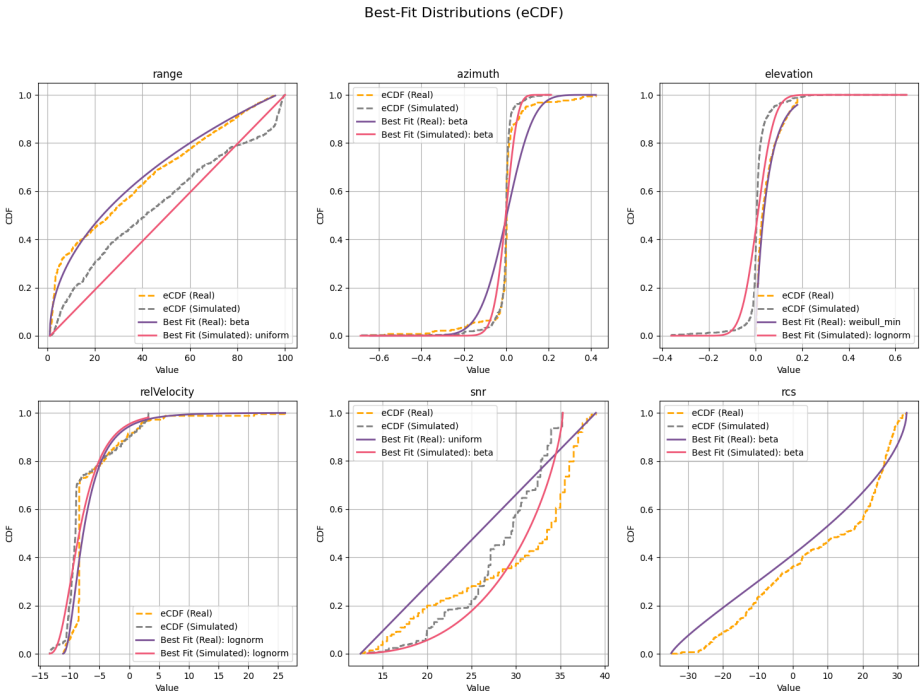

Figure F.1: MLE plot for non-normalized and non-shifted data where VUT has a speed of $50\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

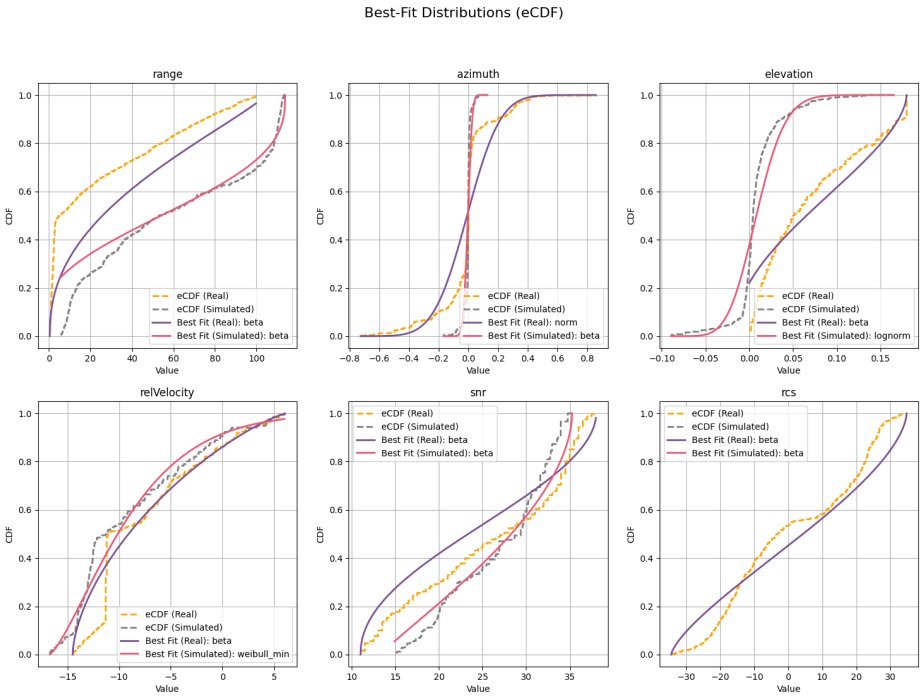

Figure F.2: MLE plot for non-normalized and non-shifted data where VUT has a speed of $60\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

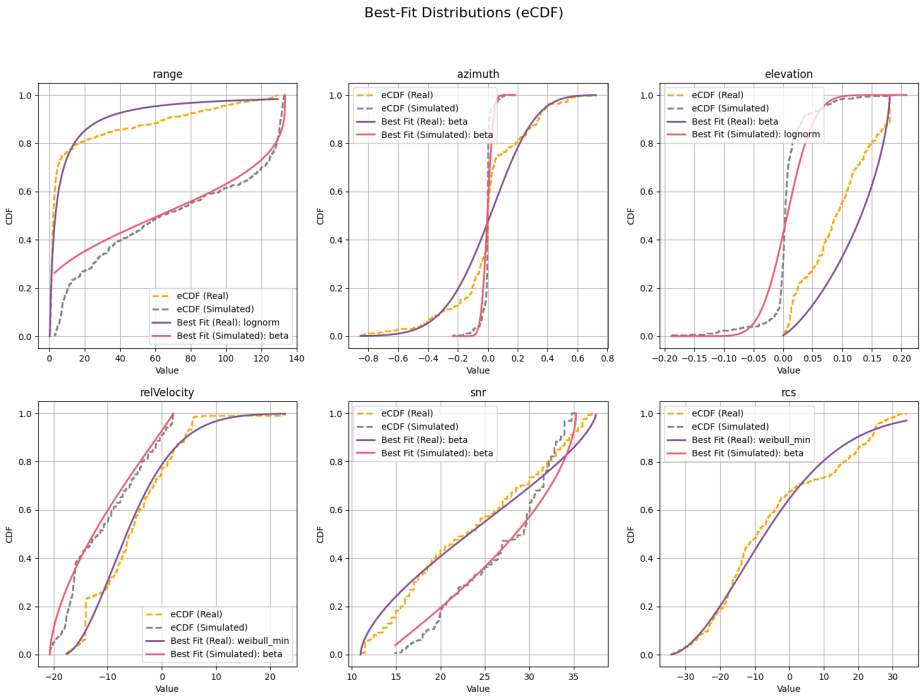

Figure F.3: MLE plot for non-normalized and non-shifted data where VUT has a speed of $70\frac{km}{h}$ and the GVT a speed of $20\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

F.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

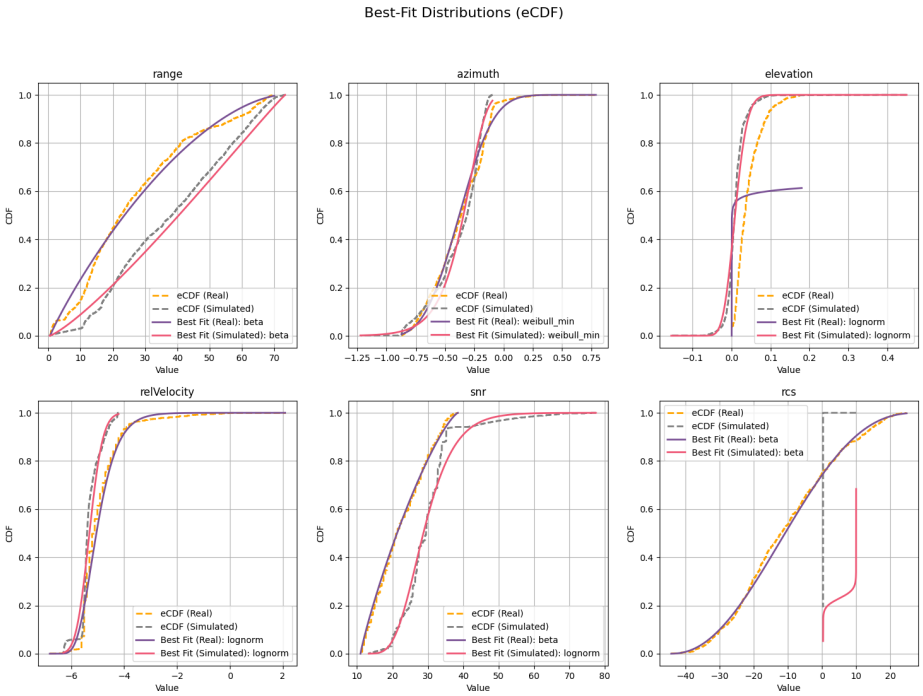

Figure F.4: MLE plot for non-normalized and non-shifted data where VUT has a speed of $20\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

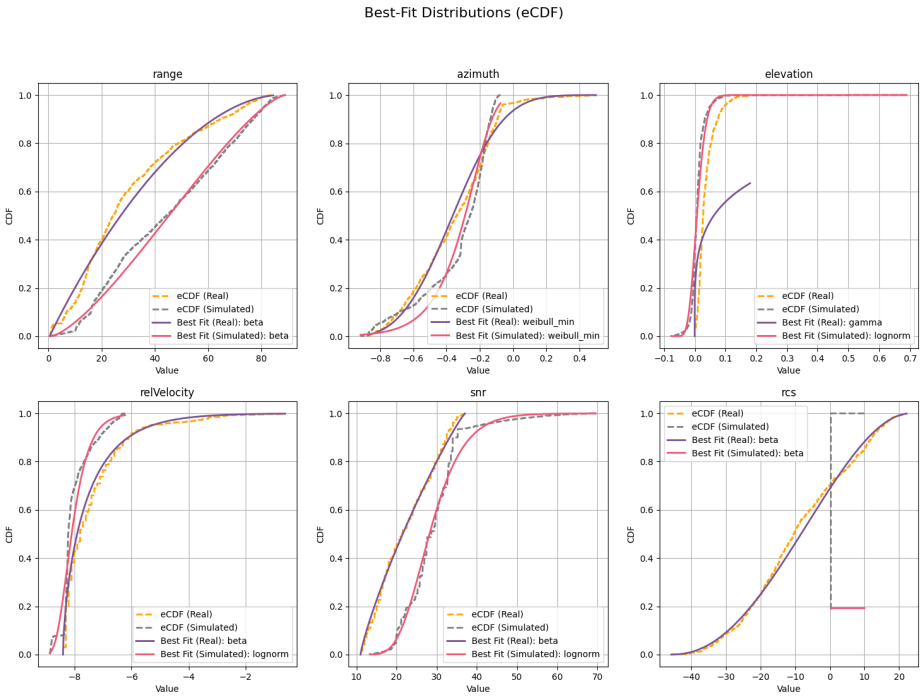

Figure F.5: MLE plot for non-normalized and non-shifted data where VUT has a speed of $30\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

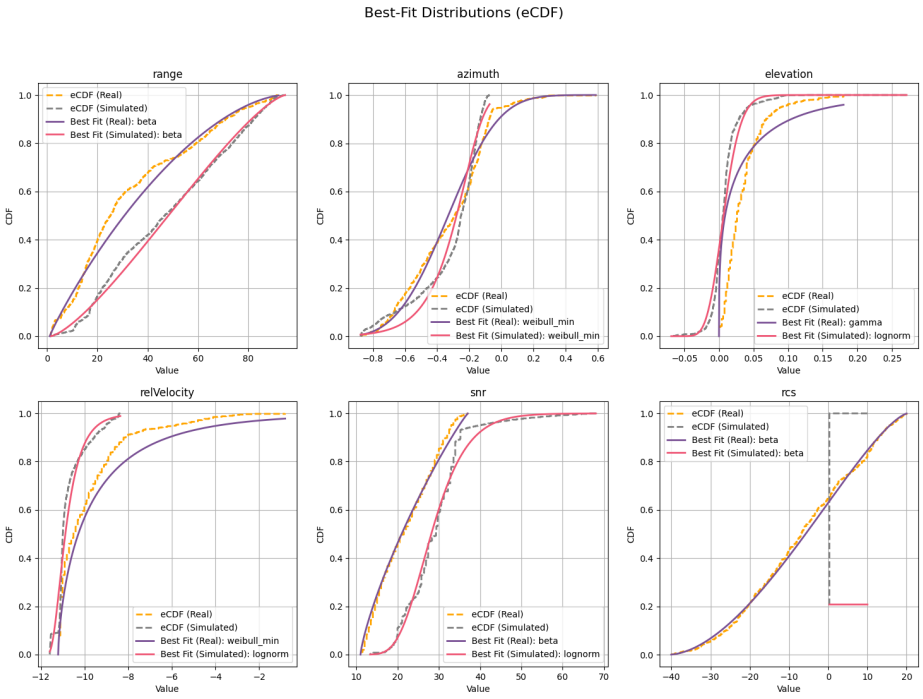

Figure F.6: MLE plot for non-normalized and non-shifted data where VUT has a speed of $40\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

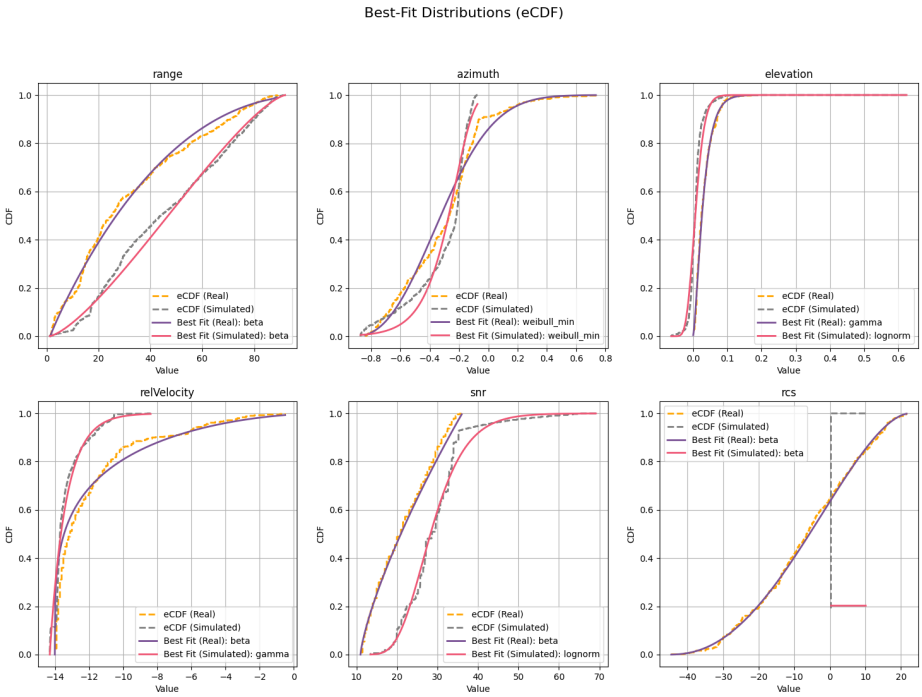

Figure F.7: MLE plot for non-normalized and non-shifted data where VUT has a speed of $50\frac{km}{h}$ illustrating the eCDF (gray curve) for each parameter with the best fit marked as a red dashed line.

G

Appendix

Appendix G contains all plots displaying the difference in the number of detections between the simulated and real data for the different scenarios.

G.1 Automatic Emergency Breaking System Car-to-Car Rear moving

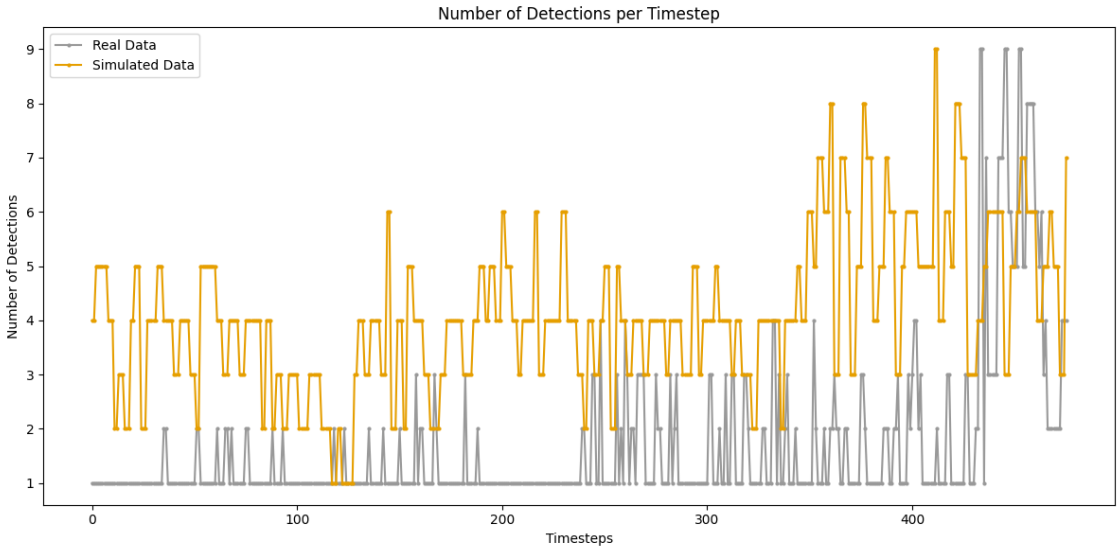

Figure G.1: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CCRm with VUT speed $50\frac{km}{h}$ and GVT speed of $20\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

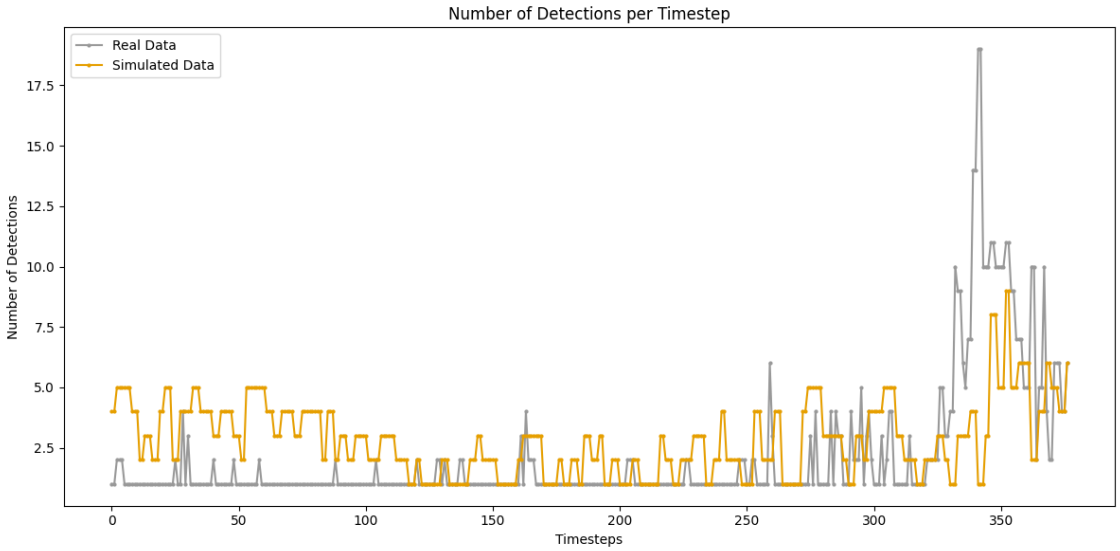

Figure G.2: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CCRm with VUT speed $60\frac{km}{h}$ and GVT speed of $20\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

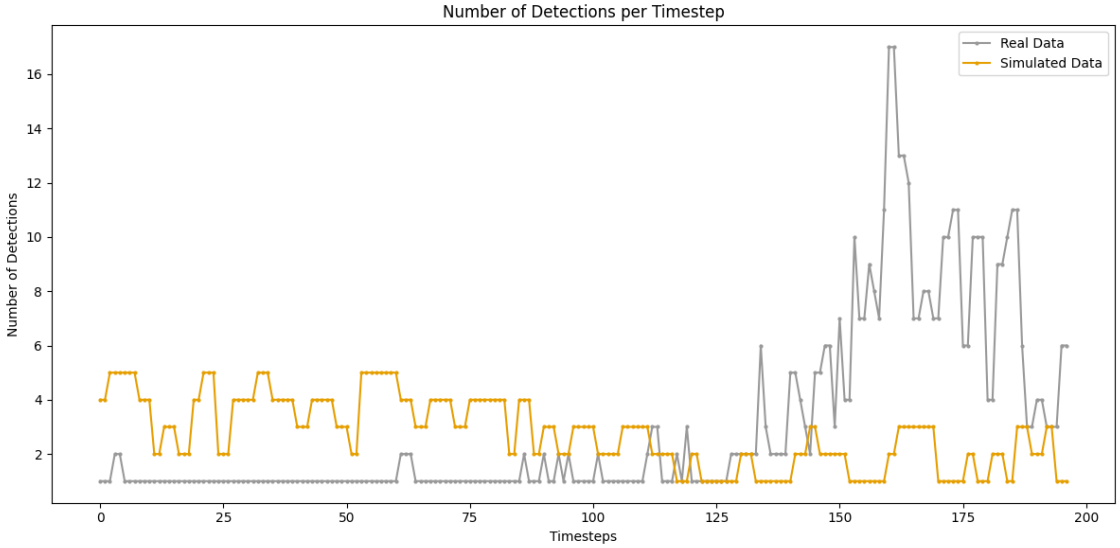

Figure G.3: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CCRm with VUT speed $70\frac{km}{h}$ and GVT speed of $20\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

G.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

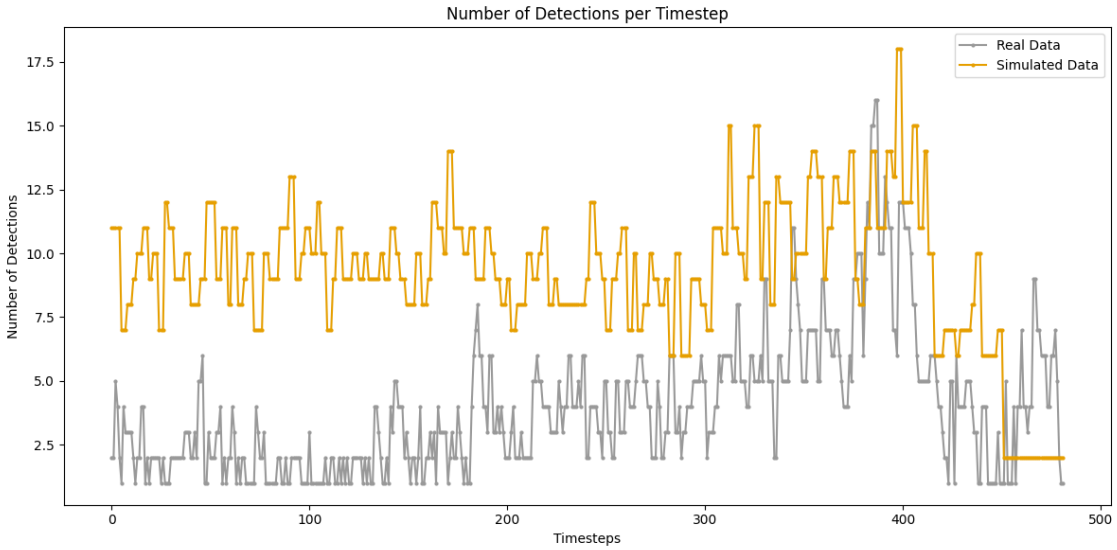

Figure G.4: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CBNAO with VUT speed $20\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

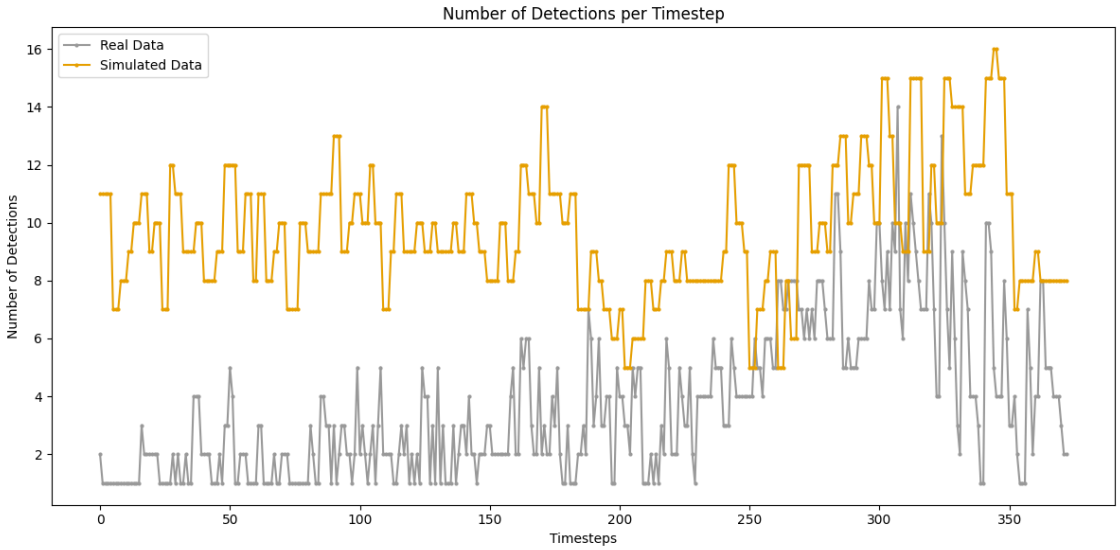

Figure G.5: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CBNAO with VUT speed $20\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

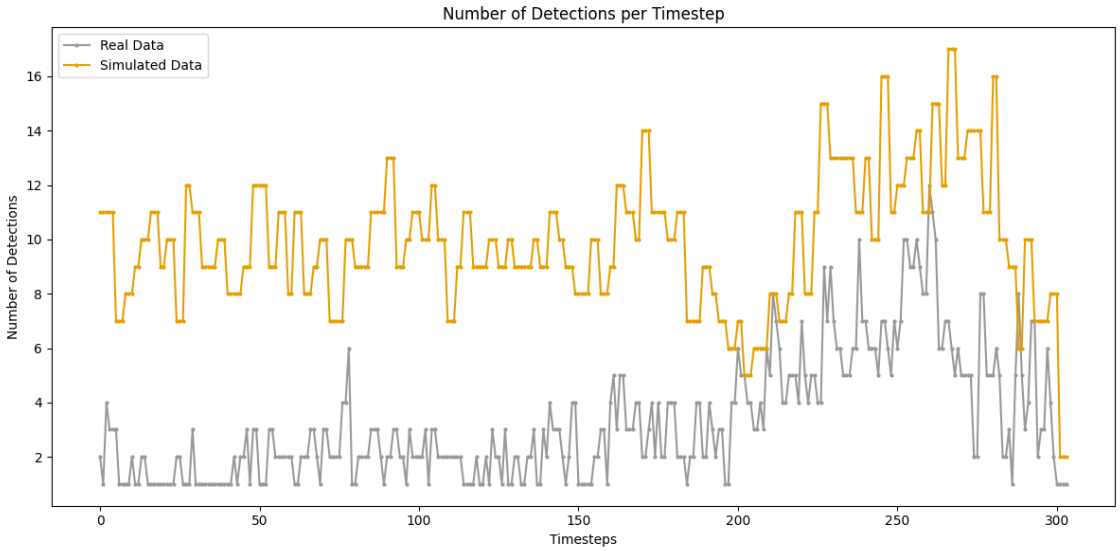

Figure G.6: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CBNAO with VUT speed $40\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

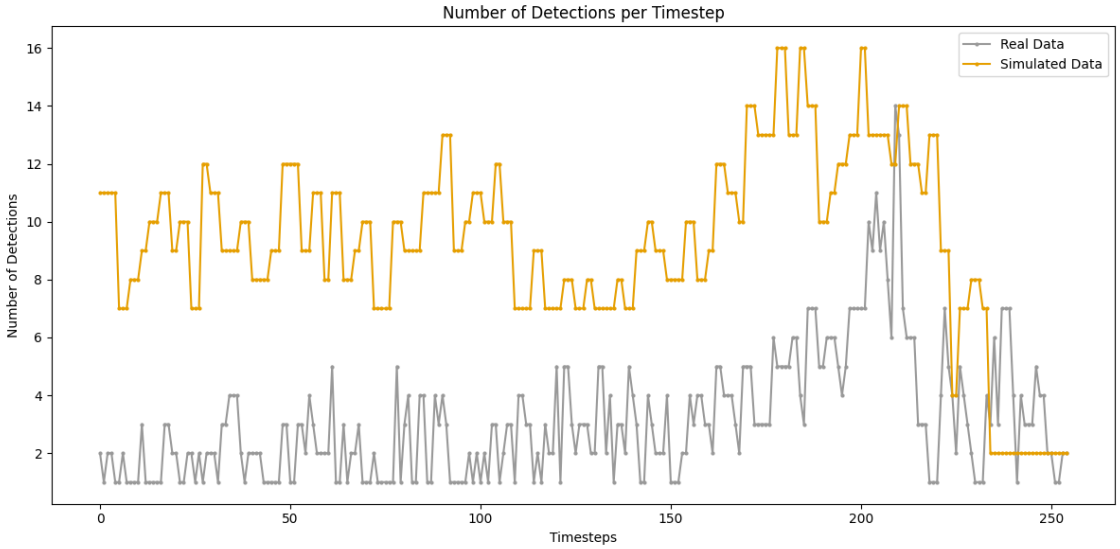

Figure G.7: A line graph illustrating the number of detections for both real and simulated datasets across each timestep for CBNAO with VUT speed $50\frac{km}{h}$. The orange line denotes the simulated data, while the gray line represents the real data.

H

Appendix

Appendix H contains all the result data for the different comparisons and the corresponding metrics: AVM, bias, positive difference, negative difference, DVM, and CAVM, all for both normalized and non-normalized, shifted and non-shifted. Below each table is a corresponding bar plot displaying the full magnitude of the DVM, with corrected AVM, (orange), and the **absolute value** of the bias, (gray).

H.1 Automatic Emergency Breaking System Car-to-Car Rear moving

Table H.1: Data from shifted non-normalized CCRm scenario with VUT speed of $50\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0055	9.7340	0.4021	2.0450	2.4471	9.7285	(2.4471, 9.7285)
azimuth	0.0168	0.0213	0.0195	0.0209	0.0404	0.0045	(0.0404, 0.0045)
elevation	0.0288	0.0010	0.0087	0.0054	0.0141	-0.0278	(0.0141, -0.0278)
relVelocity	1.5477	0.4122	0.7929	0.5744	1.3674	-1.1355	(1.3674, -1.1355)
snr	1.9308	0.6168	1.0142	1.0079	2.0221	-1.3140	(2.0221, -1.3140)
rcs	6.9762	8.1270	7.3824	7.7828	15.1652	1.1508	(15.1652, 1.1508)

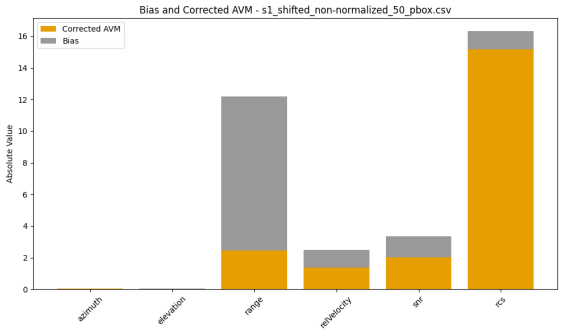

Figure H.1: Corresponding stacked bar plots to the results from table H.1 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.2: Data from shifted normalized CCRm scenario with VUT speed of $50\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0007	0.0859	0.0057	0.0194	0.0252	0.0851	(0.0252, 0.0851)
azimuth	0.0112	0.1427	0.0154	0.0332	0.0486	0.1314	(0.0486, 0.1314)
elevation	0.0404	0.1685	0.0757	0.0835	0.1592	0.1280	(0.1592, 0.1280)
relVelocity	0.0024	0.1736	0.0314	0.0503	0.0817	0.1713	(0.0817, 0.1713)
snr	0.0141	0.0374	0.0225	0.0303	0.0528	0.0233	(0.0528, 0.0233)
rcs	0.6101	0.0004	0.1306	0.0989	0.2295	-0.6097	(0.2295, -0.6097)

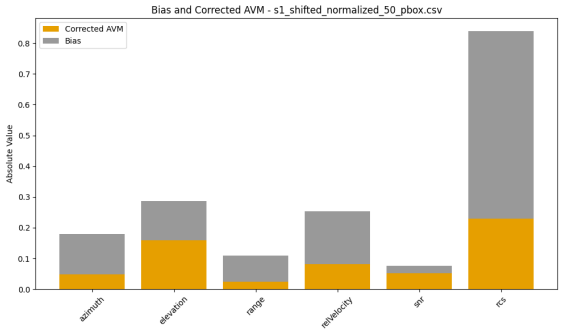

Figure H.2: Corresponding stacked bar plots to the results from table H.2 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.3: Data from shifted non-normalized CCRm scenario with VUT speed of $60\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0036	30.9124	0.6660	8.6198	9.2859	30.9088	(9.2859, 30.9088)
azimuth	0.0399	0.0351	0.0375	0.0372	0.0747	-0.0047	(0.0747, -0.0047)
elevation	0.0525	0.0000	0.0175	0.0111	0.0286	-0.0525	(0.0286, -0.0525)
relVelocity	1.0637	0.8665	1.0585	0.8434	1.9020	-0.1971	(1.9020, -0.1971)
snr	0.5953	1.2344	0.9675	0.9574	1.9249	0.6390	(1.9249, 0.6390)
rcs	4.3312	12.1483	7.2931	8.2156	15.5087	7.8171	(15.5087, 7.8171)

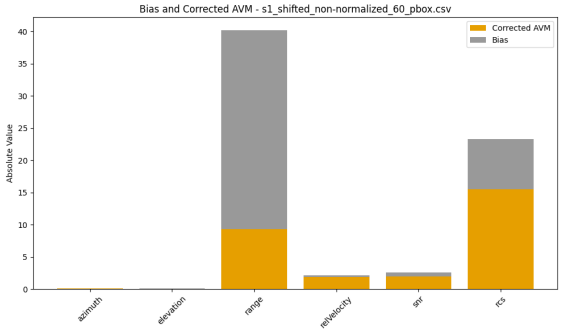

Figure H.3: Corresponding stacked bar plots to the results from table H.3 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.4: Data from shifted normalized CCRm scenario with VUT speed of $60\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0009	0.2214	0.0085	0.0776	0.0861	0.2205	(0.0861, 0.2205)
azimuth	0.0188	0.0967	0.0203	0.0421	0.0623	0.0778	(0.0623, 0.0778)
elevation	0.0718	0.1291	0.0920	0.0976	0.1896	0.0573	(0.1896, 0.0573)
relVelocity	0.0408	0.0353	0.0423	0.0325	0.0747	-0.0055	(0.0747, -0.0055)
snr	0.0111	0.0411	0.0154	0.0265	0.0419	0.0299	(0.0419, 0.0299)
rcs	0.4774	0.0015	0.1283	0.0912	0.2195	-0.4758	(0.2195, -0.4758)

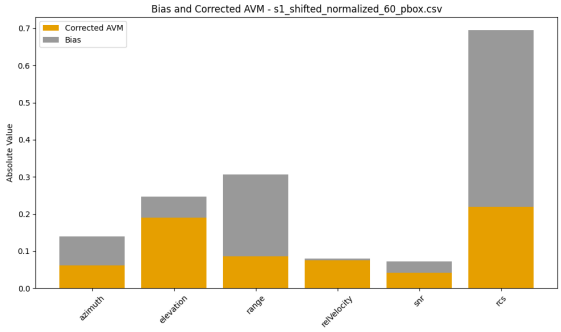

Figure H.4: Corresponding stacked bar plots to the results from table H.4 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.5: Data from shifted non-normalized CCRm scenario with VUT speed of $70\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0062	47.9793	1.1203	18.1031	19.2233	47.9730	(19.2233, 47.9730)
azimuth	0.0683	0.0513	0.0615	0.0601	0.1216	-0.0170	(0.1216, -0.0170)
elevation	0.0715	0.0000	0.0177	0.0090	0.0268	-0.0715	(0.0268, -0.0715)
relVelocity	4.1774	0.1411	1.5989	0.8035	2.4024	-4.0362	(2.4024, -4.0362)
snr	0.1341	2.9169	0.6691	1.0385	1.7076	2.7828	(1.7076, 2.7828)
rcs	2.3457	15.9958	5.9759	7.0345	13.0104	13.6501	(13.0104, 13.6501)

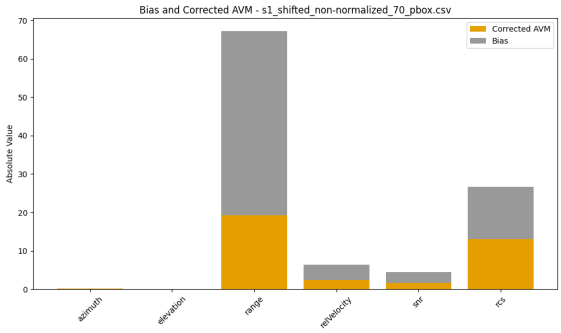

Figure H.5: Corresponding stacked bar plots to the results from table H.5 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.6: Data from shifted normalized CCRm scenario with VUT speed of $70\frac{km}{h}$ and a GVT speed of $20\frac{km}{h}$.

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0006	0.3476	0.0086	0.1395	0.1481	0.3470	(0.1481, 0.3470)
azimuth	0.0255	0.0388	0.0283	0.0302	0.0585	0.0134	(0.0585, 0.0134)
elevation	0.1064	0.0763	0.0911	0.0892	0.1803	-0.0301	(0.1803, -0.0301)
relVelocity	0.0053	0.1275	0.0419	0.0595	0.1014	0.1222	(0.1014, 0.1222)
snr	0.0040	0.1055	0.0176	0.0390	0.0566	0.1015	(0.0566, 0.1015)
rcs	0.3976	0.0018	0.1072	0.0755	0.1827	-0.3958	(0.1827, -0.3958)

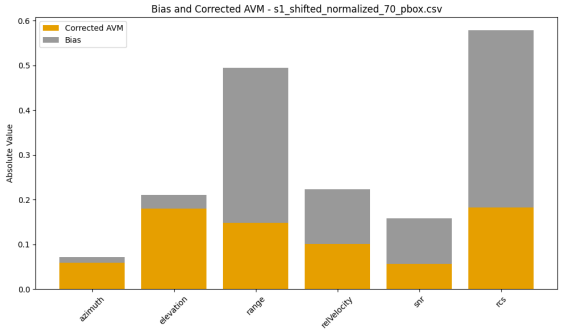

Figure H.6: Corresponding stacked bar plots to the results from table H.6 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

H.2 Automatic Emergency Breaking System Car-to-Bicyclist Nearside Adult Obstructed

Table H.7: Data from shifted non-normalized CBNAO scenario with VUT speed of $20\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0308	10.0307	0.7713	2.2758	3.0470	9.9999	(3.0470, 9.9999)
azimuth	0.0238	0.0185	0.0215	0.0213	0.0429	-0.0053	(0.0429, -0.0053)
elevation	0.0288	0.0005	0.0050	0.0029	0.0079	-0.0284	(0.0079, -0.0284)
relVelocity	0.3067	0.0338	0.1366	0.1086	0.2452	-0.2728	(0.2452, -0.2728)
snr	0.0018	6.4277	0.6422	1.2567	1.8988	6.4259	(1.8988, 6.4259)
rcs	0.5380	18.4991	4.4939	5.4192	9.9131	17.9611	(9.9131, 17.9611)

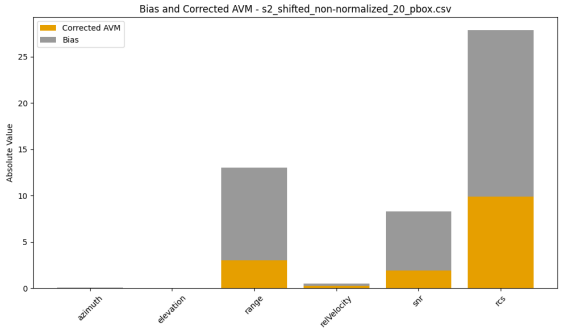

Figure H.7: Corresponding stacked bar plots to the results from table H.7 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.8: Data from shifted normalized CBNAO scenario with VUT speed of $20\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0028	0.1251	0.0132	0.0306	0.0438	0.1223	(0.0438, 0.1223)
azimuth	0.0013	0.4297	0.0155	0.0340	0.0495	0.4284	(0.0495, 0.4284)
elevation	0.0350	0.0825	0.0502	0.0526	0.1028	0.0475	(0.1028, 0.0475)
relVelocity	0.0040	0.2581	0.0198	0.0363	0.0562	0.2541	(0.0562, 0.2541)
snr	0.1523	0.0087	0.0710	0.0620	0.1329	-0.1436	(0.1329, -0.1436)
rcs	0.0333	0.3412	0.0519	0.1274	0.1793	0.3079	(0.1793, 0.3079)

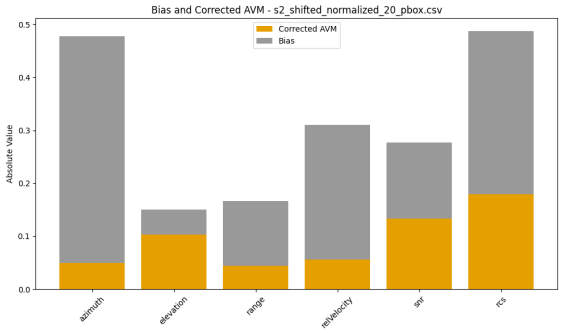

Figure H.8: Corresponding stacked bar plots to the results from table H.8 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.9: Data from shifted non-normalized CBNAO scenario with VUT speed of $30\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0520	12.7894	1.2271	3.1233	4.3504	12.7374	(4.3504, 12.7374)
azimuth	0.0151	0.0487	0.0254	0.0287	0.0541	0.0336	(0.0541, 0.0336)
elevation	0.0232	0.0005	0.0036	0.0020	0.0056	-0.0227	(0.0056, -0.0227)
relVelocity	0.5025	0.0193	0.1911	0.1539	0.3450	-0.4832	(0.3450, -0.4832)
snr	0.0040	5.9027	0.5518	1.1390	1.6907	5.8987	(1.6907, 5.8987)
rcs	0.5244	16.9637	4.6628	5.4410	10.1038	16.4393	(10.1038, 16.4393)

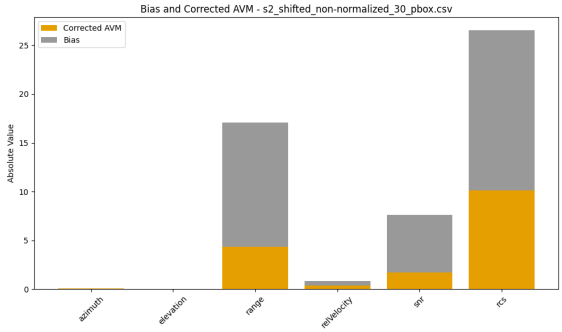

Figure H.9: Corresponding stacked bar plots to the results from table H.9 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.10: Data from shifted normalized CBNAO scenario with VUT speed of $30\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0026	0.1254	0.0127	0.0352	0.0479	0.1228	(0.0479, 0.1228)
azimuth	0.0024	0.3082	0.0153	0.0520	0.0672	0.3058	(0.0672, 0.3058)
elevation	0.0801	0.0076	0.0432	0.0392	0.0823	-0.0725	(0.0823, -0.0725)
relVelocity	0.0036	0.1774	0.0106	0.0284	0.0390	0.1738	(0.0390, 0.1738)
snr	0.1464	0.0072	0.0697	0.0619	0.1316	-0.1392	(0.1316, -0.1392)
rcs	0.0351	0.3019	0.0565	0.1209	0.1774	0.2669	(0.1774, 0.2669)

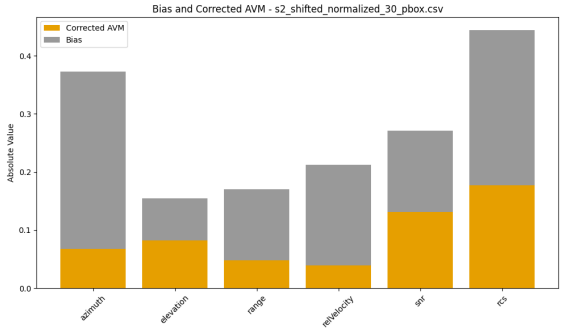

Figure H.10: Corresponding stacked bar plots to the results from table H.10 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.11: Data from shifted non-normalized CBNAO scenario with VUT speed of $40\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.1948	12.5890	1.0937	3.2008	4.2945	12.3942	(4.2945, 12.3942)
azimuth	0.0217	0.0412	0.0296	0.0310	0.0605	0.0195	(0.0605, 0.0195)
elevation	0.0223	0.0007	0.0042	0.0028	0.0070	-0.0216	(0.0070, -0.0216)
relVelocity	0.6773	0.0270	0.2977	0.2458	0.5435	-0.6503	(0.5435, -0.6503)
snr	0.0125	6.2000	0.4489	1.0236	1.4726	6.1876	(1.4726, 6.1876)
rcs	0.5851	14.5087	4.2579	5.1223	9.3802	13.9236	(9.3802, 13.9236)

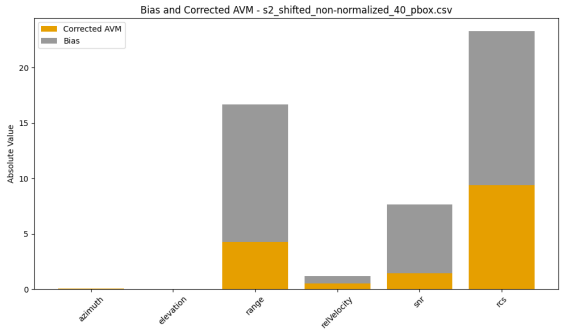

Figure H.11: Corresponding stacked bar plots to the results from table H.11 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.12: Data from shifted normalized CBNAO scenario with VUT speed of $40\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0027	0.1235	0.0118	0.0355	0.0473	0.1208	(0.0473, 0.1208)
azimuth	0.0036	0.3102	0.0189	0.0518	0.0706	0.3066	(0.0706, 0.3066)
elevation	0.0176	0.0640	0.0318	0.0351	0.0668	0.0464	(0.0668, 0.0464)
relVelocity	0.0046	0.1222	0.0100	0.0266	0.0366	0.1176	(0.0366, 0.1176)
snr	0.1267	0.0108	0.0670	0.0568	0.1238	-0.1159	(0.1238, -0.1159)
rcs	0.0339	0.2745	0.0562	0.1174	0.1737	0.2406	(0.1737, 0.2406)

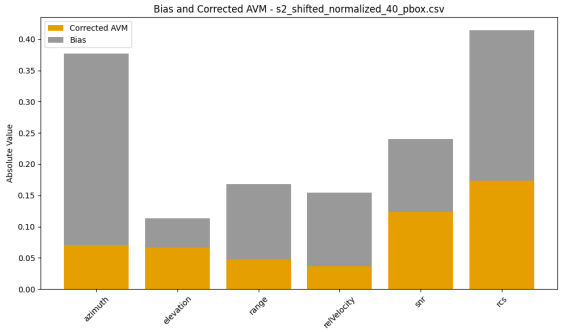

Figure H.12: Corresponding stacked bar plots to the results from table H.12 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.13: Data from shifted non-normalized CBNAO scenario with VUT speed of $50\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.1386	11.8844	0.7914	2.9872	3.7786	11.7458	(3.7786, 11.7458)
azimuth	0.0354	0.0375	0.0361	0.0363	0.0724	0.0021	(0.0724, 0.0021)
elevation	0.0217	0.0007	0.0047	0.0024	0.0071	-0.0210	(0.0071, -0.0210)
relVelocity	1.0967	0.0316	0.5383	0.4410	0.9793	-1.0652	(0.9793, -1.0652)
snr	0.0022	6.6995	0.4347	1.0655	1.5002	6.6973	(1.5002, 6.6973)
rcs	0.6769	14.5349	4.2725	5.1443	9.4168	13.8580	(9.4168, 13.8580)

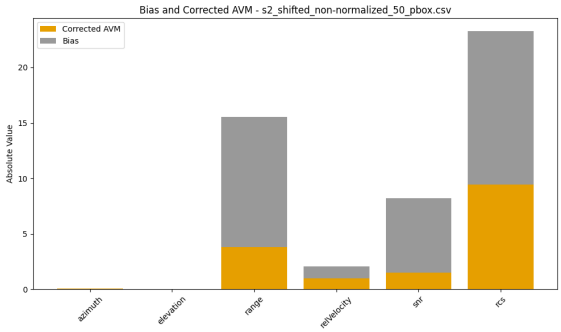

Figure H.13: Corresponding stacked bar plots to the results from table **??** illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

Table H.14: Data from shifted normalized CBNAO scenario with VUT speed of $50\frac{km}{h}$

Parameter	(D+)	(D-)	(cD+)	(cD-)	CAVM	Bias	DVM
range	0.0034	0.1226	0.0106	0.0325	0.0431	0.1192	(0.0431, 0.1192)
azimuth	0.0029	0.3432	0.0150	0.0602	0.0752	0.3403	(0.0752, 0.3403)
elevation	0.0821	0.0086	0.0456	0.0408	0.0863	-0.0735	(0.0863, -0.0735)
relVelocity	0.0173	0.0333	0.0154	0.0287	0.0442	0.0159	(0.0442, 0.0159)
snr	0.1298	0.0143	0.0695	0.0592	0.1288	-0.1154	(0.1288, -0.1154)
rcs	0.0387	0.2671	0.0547	0.1117	0.1664	0.2284	(0.1664, 0.2284)

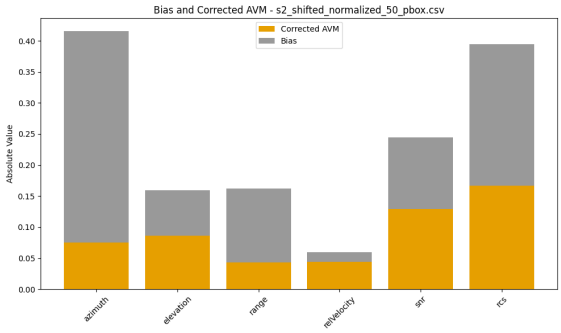

Figure H.14: Corresponding stacked bar plots to the results from table H.14 illustrating the result from the complete Double Validation Metric. The corrected AVM in orange and bias in gray. (Observe that the bias in the plot is in absolute value to visualize the size of the total error between simulations and real data.)

DEPARTMENT OF MATHEMATICAL SCIENCES
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY