STATISTICAL SEA CLUTTER MODELLING IN LITTORAL ENVIRONMENTS BY MEANS OF SEA WAVE SEGMENTATION

by

Manfred Dieter Strempel

Submitted in partial fulfillment of the requirements for the degree
Master of Engineering (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

February 2019

SUMMARY

STATISTICAL SEA CLUTTER MODELLING IN LITTORAL ENVIRONMENTS BY MEANS OF SEA WAVE SEGMENTATION

by

Manfred Dieter Strempel

Supervisor(s):	Prof. Dr. Johan P. de Villiers, Dr. Jacques E. Cilliers
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Electronic Engineering)
Keywords:	Radon transform, angle detection, small target detection, K-distribution, time-based correlation, periodic breaking wave, high resolution radar, South African coast, littoral environment

In this dissertation a combination of numerical and statistical methods are used to analyse and model periodic breaking ocean wave radar data in a littoral coastal zone. Analysis, based primarily on post-aligned parallel breaking waves, is considered for this research. Measurement data, obtained from the CSIR, is used to parametrise the model variables. Returns are corrected against the incoming wave angle, with respect to the radar, through a Radon transform analysis followed by a data alignment correction step. This ensures that the periodic breaking waves are then parallel to the time axis. The assumption is made that the fractional change in grazing angle has negligible effects on the clutter statistics. Subsequently statistical analysis is performed on the angular corrected measured data. Several different segments of a breaking wave are independently considered and analysed in terms of their distribution and correlation properties. Statistical results show that the best fitting distribution is a function of the sub-section of the sea wave and a function of the noise floor of the measuring radar. This is in part due to the radar signal not being able to reflect between the troughs of an ocean wave, because of the low grazing angle. Crest sections of an ocean wave tend to be distributed according to Rayleigh and K-distributions.

After data processing, statistical parameters are obtained, which allows for the successful simulation of statistically similar periodic breaking wave range-time clutter data. Angular data is captured for post skewing of the approaching ocean waves. Statistical correlation times and distribution parameters are analysed and recorded for use in the data generation process. Further datasets containing small rigid inflatable boats crossing the sea wave crest are compared with pure sea clutter datasets and discussed. Influences on the Radon transform and the resulting effects are also discussed.

Based on these models, methods and simulations, future work will include using the resultant model to develop improved detection algorithms, which measure and detect small targets in periodic breaking waves in littoral coastal zones. This could lead to integrating environmental wave structures, which may be independently tracked and categorised, into the algorithm design and ultimately lead to improving small vessel detection using low grazing angle coastal ground radar through classification.

LIST OF ABBREVIATIONS

AMF	Adaptive Matched Filter
ARMSCOR	Armaments Corporation of South Africa
CA-CFAR	Cell Averaging Constant False Alarm Rate
CFAR	Constant False Alarm Rate
CLS	Constrained Least Squares Filter
CSIR	Council for Scientific and Industrial Research (South Africa)
CUT	Cell Under Test
CW	Continuous Waveform
DoD	Department of Defence
DTFT	Discrete Time Fourier Transform
EEZ	Exclusive Economic Zone
EM	Electromagnetic
FFT	Fast Fourier Transform
FIR	Finite Impulse Response Filter
FS	Full Scale
GDP	Gross Domestic Product
GPS	Global Positioning System
HF	High Frequency Band
IPP	Inter Pulse Period
IQ	In-phase and Quadrature Components
KGLRT	Kelly's Generalised Likelihood Ratio Test
ML-MoM	Maximum Likelihood Method of Moments
MLE	Maximum Likelihood Estimator
MoM	Method of Moments
nm	Nautical Miles
NMF	Normalised Matched Filter

PCR	Pulse Compression Ratio
PDF	Probability Density Function
PRF	Pulse Repetition Frequency
PRI	Pulse Repetition Interval
PSD	Power Spectral Density
PW	Pulsed Waveform
q-q	quantile-quantile plot
RADAR	Radio Detection and Ranging
RCS	Radar Cross Section
RIB	Rigid-Inflatable Boat
SANDF	South African National Defence Force
SNR	Signal to Noise Ratio
UHF	Ultra High Frequency
UN	United Nations
UNCLOS	United Nations Convention on the Law of the Sea
VHF	Very High Frequency

ACKNOWLEDGEMENT

Gratitude is expressed towards the following organisations and institutions that provided funding and support through the LEDGER programme:

CSIR

our future through science

Council for Scientific and Industrial Research (CSIR)

defence

Department:
Defence
REPUBLIC OF SOUTH AFRICA

South African National Defence Force (SANDF)

ARMSCOR

Armaments Corporation of South Africa SOC Ltd

Armaments Corporation of South Africa SOC Ltd. (ARMSCOR)

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 PROBLEM STATEMENT

Rigid-inflatable boats (RIB) and other similar type vessels are difficult to detect due to their low radar cross section (RCS) and are easily hidden within sea clutter [2, 3]. Characteristics of importance, for detection, include the RCS of the boat (function of material type), absolute and relative size, radar grazing angle, reflection angle, polarisation, as well as the characteristics of the surrounding sea condition [4]. Periodic breaking wave sea clutter structures, in addition to the above characteristics, make small boats not readily detectable by common detection algorithms. They are easily treated as sea clutter and not processed correctly. Small boats typically have the response of a single radar resolution cell, which makes detection amongst measured anomalies extremely difficult in comparison to larger ships that return more cells. Ships of length greater than 30 m generally return a signal stronger than the sea clutter. Whereas for small boats ocean breaking waves and fronts can represent a larger and stronger signal, which effectively conceals the small vessel within the clutter [2].

In order to detect these types of vessels, a better understanding of periodic breaking waves in littoral environments is needed. This dissertation contains research into the statistical parameters, which are required to accurately simulate such an environment and analyse which anomalies can be detected therein.

The aim is that these methods will help researchers and practitioners to better understand and model such environments. This could lead to improved detection algorithms, aiding in improving detection of illegal fishing activities and or detection of debris for rescue operations. Overall this process has the intention to help improve border and coastal security by monitoring the shoreline more accurately at a lower resource cost [5].

1.1.1 Context of the problem

The South African Department of Defence (DoD) stated, in their 2015 Defence Review, critical tasks and goals relating to the safety and security of the country. The Defence Mission as stated in [6] is to "defend and protect South Africa, its territory, sovereignty and people and enhance peace security and development". Amongst this statement, is the need for the defence force to secure the air, land and sea borders of the country, of which the shoreline plays a formidable role essential to the economy [6, 7].

Goal 2 of the review is to safeguard South Africa against foreign and local threats. Task 4 of this goal, as outlined in the document, is to secure the borders. The borders consist of the airspace region, the land borders and the maritime areas. The country has an area of $1,219,090$ km^2 to secure, of which the land border has a length of a $4,471$ km, the an air border of $7,660$ km and a shoreline of $3,924$ km at the high water line, one of the longest shorelines in Africa [6, 7]. Further it is the responsibility of the country to also defend the exclusive economic zone (EEZ) along the shore, which includes Marion and Prince Edward Islands, as defined by the United Nations Convention on the Law of the Sea (UNCLOS) . Figure 1.1 shows the region of concern.

As mentioned in [6, 7], South Africa can be considered an island economy, 50% of its gross domestic product (GDP) is directly dependent on maritime based trade. This figure can be increased to 75% of the country's GDP, if one includes the dependency on oil and fuel required for economic productivity, which originates via sea from the middle east. This also includes an all important fishing industry, which provides an annual turnover of over R80 billion or 0.5% of the GDP [8].

There are principally four main maritime regions for the country to protect. The smallest and closest is the territorial waters measured as 12 nautical miles (nm) ($\approx$ 22.25 km) from the coast, followed by the contiguous zone (24 nm $\approx$ 44.45 km), the EEZ at 200 nm ($\approx$ 370 km) and finally the continental shelf at 350 nm ($\approx$ 556 km) [7, 9]. This is graphically represented in Figure 1.2.

South Africa lies along one of the primary international sea trading routes. These routes date back to the spice trading routes of Vasco da Gama's journey in 1497. The country is relatively isolated from its key Asian and European partners, and thus it is critical for South Africa to commit to providing safe passage and security for ships travelling through the region. The country is also obligated, by

Figure 1.1. Illustration of South African martime borders [7].

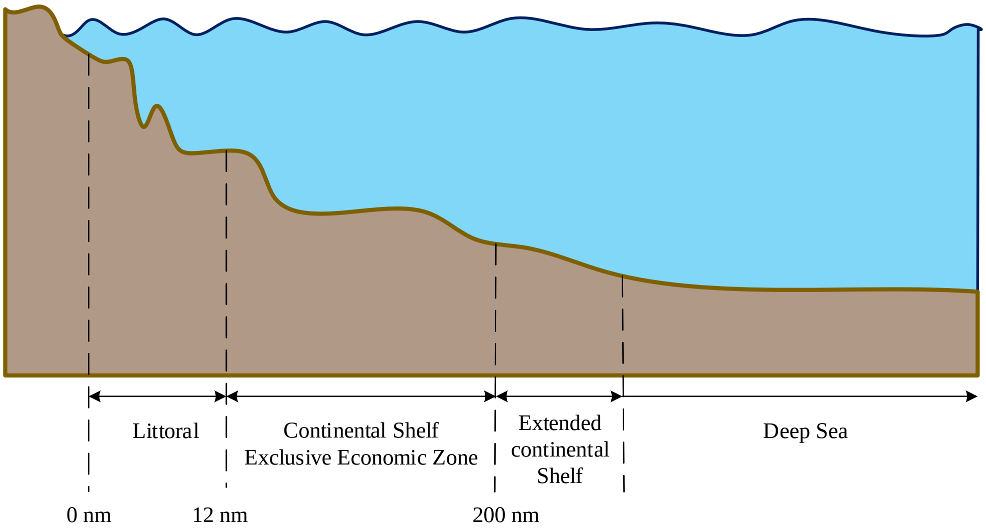


Figure 1.2. Illustration of South African maritime region of responsibility [7].

international agreement, to protect and care for the marine environment [6, 7]. Consequently it is of paramount importance to secure and protect the country’s maritime borders. This can only be achieved with reduced or low monitoring costs, such as effective detection and tracking via ground radar.

Accurate and intelligent monitoring of the shoreline is vital in addressing the following key concerns:

- peace, safety and security for the African continent and South Africa,
- piracy,
- smuggling of illegal weapons, contraband and people, and
- illegal exploitation of maritime resources.

The last point is of particular concern for this dissertation and as emphasised in [6, 7], "littoral states in the region generally lack the capacity to monitor, police and enforce maritime law". Consequently it is of utmost importance to implement the necessary new technologies to protect the country’s shoreline and economy and aid authorities to enforce relevant protection and policing acts.

Abalone poaching is a common maritime fishing infringement that occurs off the South African coast and one which is jeopardising the future of the local fishing industry. Poaching is a highly critical form of exploitation and occurs through contravening local maritime law as defined by the Marine Living Resources Act of 1998. Diving in restricted areas and exceeding the prescribed daily catch limits are difficult to regulate. Improved small boat detection will aid in the monitoring of off-shore activities and the protected fishing areas [10].

1.1.2 Research gap

The South African coastal oceanic littoral region, with regular occurring strong periodic breaking waves, have no optimised functional simulation models for periodically structured sea clutter. This is essential for the correct simulation of data, not only for analysis but also for the correct development of object detection algorithms within such strong clutter environments. Further, the generation of data may be used to verify such methods. The lack of such models limits the capabilities within the research community to expand and advance detection algorithms that match the sea conditions and sea clutter environment. This dissertation creates a first statistical representation of periodic breaking waves in a littoral environment typical to the South African coast.

1.2 RESEARCH OBJECTIVE AND QUESTIONS

The primary research objective of this dissertation is to develop the processing steps required to parametrise the model that is suggested for analysing and generating breaking sea wave clutter in a littoral environment. The model considers data distribution and correlation at individual sections of the breaking wave. This is justified as the signal from a radar with low grazing angle may be obscured by the wave structure itself. The goal is to reduce the cost of measuring radar data in littoral regions and generate accurate breaking wave sea clutter for further investigations and simulations.

The main research questions that are posed in this study include:

- Can recorded radar data within the littoral environment be accurately represented by statistical means?
- Can the statistical data then be used to generate accurate simulated data?
- What is the computational expense of such a process?
- What significant characteristics do each of the current methods offer and how are they related to one another?

1.3 HYPOTHESIS AND APPROACH

This dissertation investigates the littoral sea clutter amplitude distributions associated with individual segments within periodic ocean waves. Low grazing angle, X-band sea clutter data, obtained from the Council for Scientific and Industrial Research's (CSIR) online database was utilised for this study [11]. The dataset contains fixed frequency pulsed radar returns of low grazing angle sea clutter with a pulse repetition frequency (PRF) between 2.5 kHz and 5 kHz. The algorithmic processing involves correcting the data by means of rotation, such that the periodic wave phase structures within the data matrices align over time and are parallel to the longtime axis. Thereafter, the data and thus the wave structures, are segmented into multiple partitions, where each partition represents a section of the ocean wave phase in the range dimension. These segments are then fitted against multiple distributions typically used to model sea clutter. It is expected that the approach can accurately describe specific sections of the ocean wave with a reduced error between actual and estimated distributions. The improved probability density function (PDF) representation will allow for more accurate data analysis and data generation in littoral environments. This can then be used for the simulation and performance estimation as well as development of new detectors [4].

1.4 RESEARCH CONTRIBUTION

The above problem is still considered an open problem in literature [12]. This proposal suggests the generation of a sea clutter model for periodic breaking ocean waves in a littoral environment, which through future work and research, may lead to improved probability of detection of small vessels in such environments. One conference paper discussing the analysis steps of the suggested model has been published [4].

1.4.1 Research goals

The goals of this research are to:

- generate a statistically accurate model of breaking sea wave clutter in littoral environments,
- accurately represent breaking periodic waves inherent to the South African coastal environment,
- verify this method against actual measured radar data obtained off the coast of Cape Town.

1.5 OVERVIEW OF STUDY

In this dissertation, a brief review of present clutter models and detection algorithms are presented in the literature review of Chapter 2. This is followed by a basic summary of the operating principles of a radar system and its corresponding terminology, which is presented in Chapter 3. Chapter 4 presents the mathematical methods and algorithms used specific to the proposed algorithm. Some data illustration examples and further radar background are also provided where needed. The results are then presented and discussed in Chapter 5, followed by a summary and conclusion in Chapter 6. Appendix A presents results for further processed and analysed datasets that were not presented in Chapter 5.

CHAPTER 2   LITERATURE REVIEW

Clutter models and detection methods applicable to this dissertation are reviewed in Chapter 2. A summary of known clutter models with an emphasise on high resolution, low grazing angle radar data are presented in Section 2.1. Detector models based on statistical methods are reviewed in Section 2.2. The chapter is closed with some concluding remarks.

2.1   CLUTTER MODELS

The ability to analyse and model sea clutter is essential for accurate and reliable radar design and performance [13, 14, 15, 16, 17, 18]. Effective models allow for the optimisation of radar hardware and processing parameters during design, leading to more accurate and improved detection of targets [15, 17]. This is especially important for low RCS targets that are difficult to detect and measure. Much work has been completed in this area and several modelling methods based on statistical and electromagnetic principles have been investigated and analysed [16, 19].

Models can be broadly categorised into the following methods:

- distributional analysis with spatial correlation and amplitude statistics,
- fractal theory,
- chaos theory, and
- other processing methods specifically aimed at small target detection [16, 19].

Historically, low resolution radar systems typically show returns which can be stochastically categorised to be of a Gaussian nature [13, 15, 16, 20]. However, due to increased radar resolution, returns typically no longer follow these Gaussian based distributions and contain additional spiky measurements. These spikes were previously integrated out due to the lower resolution. Using high resolution data causes traditional Gaussian based detectors, such as the standard Gaussian based CFAR, to produce an increased number of false alarms [15, 20]. This occurs due to the wave front spikes being treated as possible targets, essentially decreasing the radar's performance and reliability to detect targets. Overall, the detection algorithms had to be adapted to account for the increase in resolution [15], which allows for the capturing of small targets previously not measurable.

Many distributions have been considered to accurately describe these high resolution radar returns, including the log-normal, Weibull, and the most commonly accepted K-distribution [13, 16, 19]. In order to address the increase in radar resolution Pusey presented an alternative to the Rayleigh model. Originating from the optical laser field, the K-distribution can be used for sea clutter distribution modelling [20]. Ward and Watts further illustrated that this distribution can be considered to be a phenomenological model [15, 21]. This essentially describes the electromagnetic reflections in such a way so that they can be traced back to a physical structure present within the waves [17, 18]. Ocean waves are considered to be the addition of two main types of sub-waves. Gravity waves make up for the large amplitudes and higher wavelengths, whereas capillary waves, caused by the water surface tension of the wave, add a random type of behaviour to the radar returns [17, 21]. Furthermore, it is shown that the amplitude of the capillary waves is also a function of the direction of the gravity wave. For example an approaching front travelling against the wind direction shows smaller capillary amplitudes than on the other side of the gravity crest travelling with the wind [17, 21]. Horizontally polarised returns are more pronounced to a limiting case of a log-normal distribution, for low grazing angles, than vertically polarised returns. The latter tends to be a limiting case of a Rayleigh distribution [13, 20]. All of these can, to an extent, be described by the K-distribution. Ward shows how to apply the K-distribution to coherent data [14, 21] and its applicability. The compound K-distribution further fluctuates the power of the K-distribution using an exponential distribution, thus improving the spatial correlation [19, 21]. In [22, 23], Watts accounts for time dependent frequency characteristics by using the power spectrum, mean Doppler spectrum, clutter to noise ratio, and normalised width of the spectrum as inputs to the model for single range bin. These are then expanded using a Gaussian variate for a weighted FIR to generate time-series correlated data.

More recently researchers suggest that sea clutter is better modelled as a chaotic process [24, 25]. This allows the clutter to be described by a non-linear deterministic system with a few degrees of freedom [26, 27]. The difficulty is thus to retrospectively track this time series data and fit it to a non-linear deterministic model. Methods include algorithms based on neural networks and multiple model neural networks, where predictions are then compared to measured data to determine detections [25, 28]. Chaotic forms of modelling do not easily integrate themselves into standard detection methods, which are commonly based on statistics [17]. Chaotic models are sensitive to initial conditions. In [29] an exponential sensitivity to initial conditions model is considered. Chaos theory includes the methods of fractal analysis, which implies a feedback loop together with an alternating loop to describe sea clutter. This allows high flexibility towards the input parameters and produces a certain statistical similarity to measured data [5, 30]. Chaotic models have been disputed by many researchers [27, 31].

2.2 DETECTION METHODS

Owing to the high resolution returns sea wave structures are resolved, making wave troughs and crests visible within the measured data [2]. This distributional information needs to be taken into account during the design of the detector algorithms such as the standard CFAR, so that targets can autonomously be detected within these structures [32, 33]. Statistical models have shown good results at modelling high resolution data [12, 13, 14, 15, 17]. The K-distribution has shown promising results as an underlying distribution for small target detection [2, 15, 18].

In [2], Panagopoulos summarises the common techniques for signal detection into the following categories:

- signal averaging,
- time-frequency representation, and
- morphological filtering.

Signal averaging relies on the property that noise is incoherent and the signal coherent. Under this condition and for a relatively stationary target in the given time-span, noise is averaged towards zero, whilst static signals are reinforced [2]. Even with the suggested forgetting factor, a relatively stationary target with respect to the wave velocity is assumed.

The time-frequency representation essentially filters out low frequencies due to the input window size, thus acting as a high pass filter. The method relies on targets and waves moving at different velocities

[2]. Frequency based analysis relies on the Doppler spectrum features in coherent data [32]. It often occurs that the Doppler of slow targets overlap with that of the wave crest, making detection difficult [32]. In [32], Wang expands on the traditional single dimensional Doppler amplitude based detection method, to improve small target detection, by instead considering the obtained multidimensional Doppler spectrum shape.

Morphological filtering is a non-linear image processing method to analyse textures [2]. This method assumes a target detection structure that is longer in time than the presented wave crest and a relatively stationary target for staring data.

For a well defined and known Gaussian distribution with known covariance, the optimal filter for training of the clutter distribution is given by a whitening-matched filter [12]. However, the distribution rarely fits the true world and the parameters are rarely known. Therefore, for a system to account for a changing environment with varying target distributions and characteristics, adaptive filters are commonly implemented. These allow the detection of targets within clutter of an unknown distribution.

Kelly's generalised likelihood ratio test (KGLRT) and the adaptive matched filter (AMF) are used for coherent data [12], where it is known that the clutter distribution is non-Gaussian and a reasonable assumption of the steering vector of the target can be made. The normalised matched filter (NMF) further improves on this as it only requires prior knowledge of the noise covariance matrix, allowing for relatively robust detection in multiple scenarios [12].

2.3 CONCLUDING REMARKS

The literature review presents known methods that are typically used to model high resolution radar data. These models are commonly suited to data with a clearly defined underlying distribution. The common distributions include the log-normal, Weibull, and the most commonly accepted K-distribution. Detection methods are also based on these models and their performance is thus dependant on the quality of radar returns and their correlation to the defined distribution.

CHAPTER 3 THEORETICAL BACKGROUND

Some of the primary theoretical knowledge used throughout this dissertation is summarised in this Chapter. Fundamental radar concepts are explained and statistical distributions, commonly associated with sea clutter, are introduced. Furthermore, specific methods for generating data and analysing the accuracy of statistical fits are presented. The chapter is concluded with a short discussion on one of the most common detection methods.

3.1 RADAR OVERVIEW

A radar emits electromagnetic (EM) waves from its transmitter via a specifically designed directional antenna. Echoing waves that have reflected back off the target, return to the receiver and are processed to obtain information such as range, velocity and angular position [1, 34].

The transmitted waveform is one of the primary characteristics which describe a radar system. There are two common types of waveforms, namely continuous waveforms (CW) and pulsed waveforms (PW). As the name suggests, the CW type continually transmits a signal such that both the receiver and transmitter are working concurrently. This is only possible with a two antenna system. Otherwise leakage of the high power transmitter, which typically produces signals in the range of hundreds of watts (10^2 W) to megawatts (10^6 W), into the ultra sensitive receiver, typically with a sensitivity in the region of milliwatts to nanowatts (10^{-3} W to 10^{-9} W), would render the received signal undetectable. This would require a costly receiver section with a high dynamic range [1].

The PW type emits short transmit pulses, followed by longer listening cycles where the system waits for the echoes to return, all via the same antenna. The transmitted pulse is of a short duration and denoted by the pulse width τ, which is usually around $\tau = 10^{-7}$ s. The time between pulses is proportionally much longer than the transmitted signal and is used by the radar system to listen to EM returns. The time difference between transmitted pulses and received pulses is then used to measure the one way distance to the target and summarises the basic operation of a radar system [1, 34].

This method is best described by the range equation (3.1),

$$R = \frac{c\Delta T}{2}, \tag{3.1}$$

where R is the range to the target in meters, calculated by the change in time, ΔT, it takes the transmitted signal (traveling at a speed of light $c \approx 3 \times 10^8\ \mathrm{m\,s^{-1}}$) to reflect off of the target. The division by two accounts for the two-way path time [34].

As mentioned, there are principally two types of radar antenna configurations. Those with separate transmission and reception antennas, and those with a single antenna together with a transmit/receive switch. If these antennas are in close proximity to each other the system can be considered monostatic; any systems where the antennas are significantly separated in space from each other are considered bistatic [34].

The key to improved reception and signal strength in monostatic configurations is the transmit/receive switch. This isolates the power sensitive receiver electronics from signal leakage originating from the powerful transmit stage, as shown in Figure 3.1. Upon reception of the signal, the signal is sent through a low noise amplifier and a matched filter. The signal is synced to the transmit oscillator so that phase information can also be determined. This analog signal is then sampled into the digital domain, where range and phase information is captured. From here subsequent signal processing can occur, and detection algorithms can be implemented [1].

An illustration of a typical system for measuring data is shown in Figure 3.1.

Another critical parameter for the description of a radar system is the frequency band. The operating frequency influences the radar's capability with regard to distance, resolution, weather dependency, required antenna size and complexity as well as and cost. Conventional radars use the frequency bands that lie between high-frequency (HF) and the W-band, as shown in the electromagnetic spectrum in Figure 3.2 [1, 34].

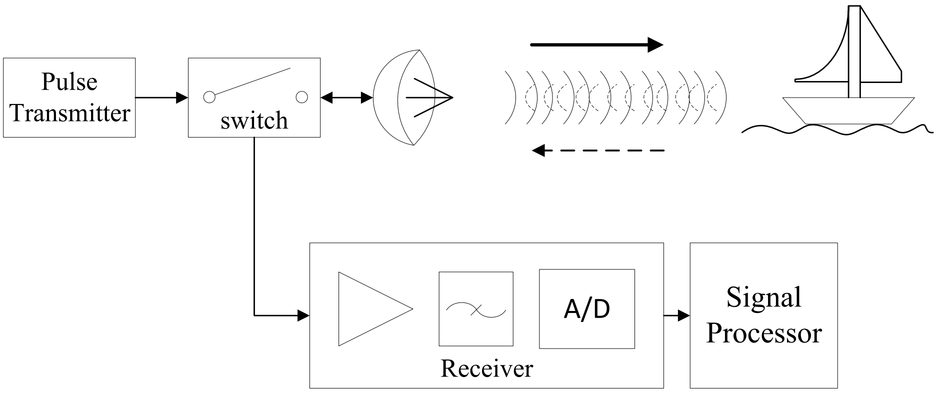


Figure 3.1. Principle elements illustrating a functioning radar system. Reproduced with permission from [1].

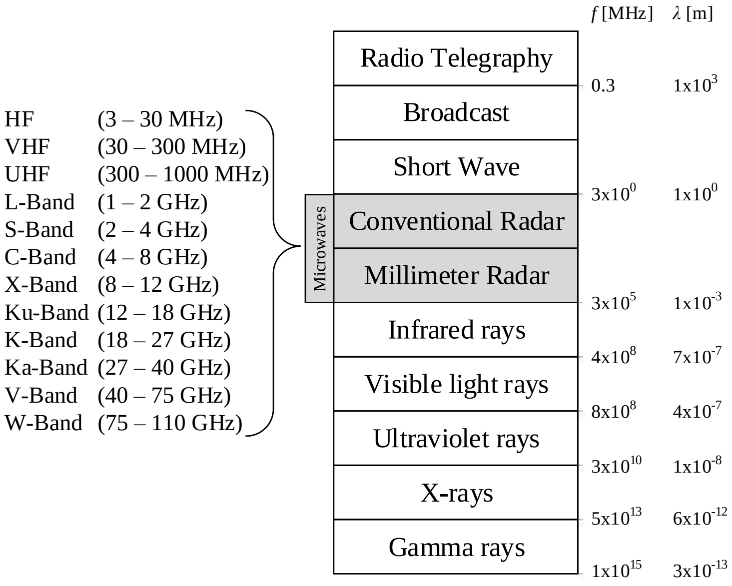


Figure 3.2. Electromagnetic spectrum showing common radar bands. Reproduced with permission from [1].

Non-coherent radar systems only measure amplitude returns, whereas coherent radar systems also capture phase information in addition to the amplitude. Typically non-coherent systems are used in cases where the target signal far surpasses that of the clutter signal. Coherent radars offer the user the ability to also determine motion characteristics by analysing the change in phase properties of the measured signal. For pulsed systems, the phase is measured on every return pulse (range bin) and compared against that of a local oscillator which is used to sync the transmit pulse with the received signal. Thus, any phase variations can easily be determined [1, 34].

3.1.1 Radar waveform principles

The propagation of a transmission wave oscillating at a carrier frequency can be defined by

$$E = E_0 \cos(kz - \omega t + \phi), \tag{3.2}$$

where E is the electromagnetic wave, E_0 the maximum wave amplitude, k the wave number, z the axis along which the wave propagates, ω the angular frequency, t the time in seconds , and ϕ the initial phase offset [1].

The wave number is related to the wavelength given by

$$k = \frac{2\pi}{\lambda}, \tag{3.3}$$

where λ is the wavelength in meters [1]. Further, the carrier frequency is related to the angular frequency defined by

$$\omega = 2\pi f, \tag{3.4}$$

where f is the carrier frequency in Hertz [1]. The wavelength and the frequency are related by the constant c (speed of light) defined by

$$\lambda f = c. \tag{3.5}$$

For a PW type radar, a pulse is propagated along an axis at a carrier frequency for a duration of τ [s]. This pulse is then repeated over time, essentially defining the inter pulse period (IPP) denoted by the symbol T, which is also known as the pulse repetition interval (PRI). A repeating sequence is commonly referred to as a pulse train [34]. Figure 3.3 illustrates this relationship.

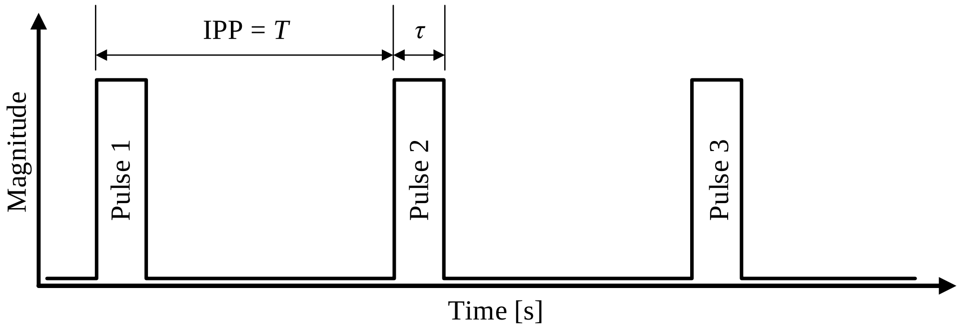


Figure 3.3. Transmission pulse train illustrating radar timing. Reproduced with permission from [1].

The pulse repetition frequency (PRF) can then simply be defined by the inverse of the PRI, given by

$$\begin{aligned} PRF &= \frac{1}{PRI} \\ &= \frac{1}{T}. \end{aligned} \tag{3.6}$$

A PW waveform is transmitted for a short duration as defined by τ. After transmission the transmitter is disconnected form the antenna and switched to the receiver. Thereafter the receiver waits and listens for a period of time, as shown in Figure 3.4 [1, 34].

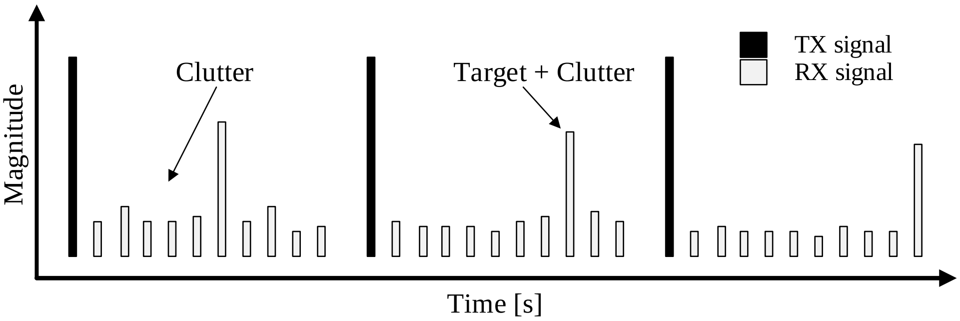


Figure 3.4. Illustration of radar return signals. Reproduced with permission from [1].

The PRF defines the number of transmit/receive cycles per second, usually expressed in Hz. Expanding on equation (3.6), the pulse width τ can be used to determine the duty factor d_t given by

$$\begin{aligned} d_t &= \frac{\tau}{\text{PRI}} \\ &= \tau \cdot \text{PRF}, \end{aligned} \tag{3.7}$$

thus allowing the calculation of the average transmitted power of the transmission signal. This can be expressed as

$$\begin{aligned} P_{avg} &= P_t \cdot d_t \\ &= P_t \cdot \tau \cdot \text{PRF}, \end{aligned} \tag{3.8}$$

where P_t is the radar peak power in Watts.

Furthermore, the energy of one pulse can be described by

$$\begin{aligned} E_p &= P_t \cdot d_t \\ &= P_{avg} \cdot T \\ &= \frac{P_{avg}}{f_r}. \end{aligned} \tag{3.9}$$

3.1.2 Range ambiguity and resolution

From Figure 3.4 it is clear that the listening time is finite (defined by $T - \tau$) and as such so is the maximum measurable range, R_{max}. Should a return occur at ΔT, it becomes unclear whether the return is due to the last transmitted pulse (τ_1) or the pulse before that (τ_0), making the response ambiguous.

In order to obtain unambiguous range measurements, the round trip time Δt must be less than the IPP. This ambiguity is typically avoided by ensuring that the PRI is long enough to include all radar returns of interest in the range under consideration. The maximum unambiguous range is then given by

$$\begin{aligned} R_{max} &\leq \frac{c \cdot \text{PRI}}{2} \\ &= \frac{c}{2 \cdot \text{PRF}}, \end{aligned} \tag{3.10}$$

where the division by 2 accounts for the two-way path delay.

Owing to the Nyquist-Shannon sampling theorem, the sampling frequency of the return pulses need to exceed that of the transmission pulse frequency, $f_{pulse} = \frac{1}{\tau}$ by a factor of at least 2 to avoid aliasing [34]. This is given by

$$\begin{aligned} f_{min} &= 2 \cdot \frac{1}{\tau} \\ &= 2 \cdot f_{pulse}. \end{aligned} \tag{3.11}$$

Every received sample during the listening period represents a distance in range. Discrete distances are referred to as range bins. The number of samples (which is a function of the pulse width τ) determines the range resolution ΔR and thus the individual size of a single range bin for the radar, given by the equation (3.12).

$$\begin{aligned} \Delta R &= \frac{c \cdot \tau}{2} \\ &= \frac{c}{2 \cdot B}, \end{aligned} \tag{3.12}$$

where B is the bandwidth of the transmitter ($B = \frac{1}{\tau}$) and is limited by the semiconductor electronics. The amount of energy that can be created during a short pulse period τ is limited to component dielectric strengths and operating temperatures. This in turn limits the maximum resolution that the radar system can evaluate [1, 34].

A method known as pulse compression increases the accuracy by modulating the pulse with a known waveform, typically frequency or phase modulated, thus allowing the receiver to identify overlapping targets within a return pulse. The improvement achieved by using such techniques, with respect to the traditional radar equation (3.17), is defined by the pulse compression ratio (PCR) [1, 34]. Measurements using these techniques were not considered in this dissertation.

3.1.3 Doppler shift

Relative motion between the receiver and the target is termed Doppler shift. This occurs when either the target or the radar (or both) travel towards or away from each other during the pulse period, thus effectively changing the frequency of the pulse. The difference in the frequency is then given by the Doppler frequency shift f_d and is described by equation (3.13) [1],

$$f_d \approx \frac{2v_r}{\lambda}, \tag{3.13}$$

where v_r is the target's relative velocity toward the radar. The relative velocity v_r and thus f_d are positive for targets approaching the radar and negative for targets moving away from the radar [1].

Doppler shift ambiguity occurs, as described by the Nyquist-Shannon sampling theorem, when the sampling rate of the velocity v_r is less than the PRF as defined by equations (3.14) and (3.15). Thus moving targets impose a maximum velocity which can be defined by $v_{r_{max}}$. The maximum non-ambiguous velocity is a function of the PRF and best described by (3.14) and (3.15).

$$f_{d_{max}} = \pm\frac{PRF}{2} \tag{3.14}$$

$$\begin{aligned} \text{PRF}_{min} &= 2 \cdot f_{d_{max}} \\ &= \frac{4v_{r_{max}}}{\tau} \end{aligned} \tag{3.15}$$

3.1.4 Radar cross section

The EM return wave is a function of multiple properties of the object it reflects off. The RCS of a target describes the signal strength of the reflected transmitted signal from the target. In its basic form, the RCS is given by

$$\sigma_{RCS} = \frac{P_r}{P_D} \cdot m^2, \tag{3.16}$$

where P_r is the power received off of the target and P_D is the transmitted power density. P_D is a function of range and thus attenuates over distance, m^2 represents the target's surface area [34].

The returned P_r value is a function of at least the following properties:

- the size of the target,
- the mechanical form of the target,
- the orientation and direction of the target, and
- the material properties of the target.

3.1.5 The radar equation

The radar equation describes the physical process and takes RCS, signal attenuation, transmission efficiency, SNR, power, frequency, and maximum range into account. There are many forms of the equation. These forms depend on which variable the equation is solved for and which radar configuration is given [1, 34]. The range, R, can be calculated from the radar equation as in (3.17)

$$R = \sqrt[4]{\frac{P_D \cdot G^2 \cdot \lambda^2 \cdot \sigma_{RCS}}{P_r \cdot (4\pi)^3}}, \tag{3.17}$$

where G is antenna gain (omnidirectional), which is dependent on design and operating frequency. The relation between distance R and power P_D shows the physical limitations and constraints commonly found in radar design.

Equation (3.17) can be extended to include further factors such as directional antennas, antenna loss, thermal noise, temperature and the Earth's surface in terms of reflection. However, for this dissertation it is deemed not necessary to include these as little reference to these design parameters is made with respect to the results, and thus forms only part of the background information.

Further important criteria for the design includes noise N, which is determined by the operating frequency and bandwidth, thus effectively determining the minimum detectable power S_{min}.

3.1.6 Clutter principles

Clutter represents any unwanted signal returns that are captured by the receiver. This includes noise and unwanted reflections off of objects that are not of interest. Thus, by definition, the classification of clutter itself is dependent on the function of the radar. The following groups of clutter are commonly considered [1]:

- Surface clutter: typically stationary masses or sea surface reflections that are large and have slow to no velocity. Is is described by the reflectivity given as a surface area in m^2.
- Volume clutter: typically unwanted objects found in the air between the radar and the object of interest, such as rain or snow (with the weather radar as an exception). Commonly described by the reflectivity defined by the volume and given in m^3.

These can further be categorised into clumped, spiky, correlated and discrete clutter [2].

Sea clutter is an inherently difficult form of clutter to detect, due to its inherent wind-based Doppler speed that occurs due to the ocean waves moving. Wave crests that are returned are either approaching or retracting from the radar origin over time, thus automatically influencing the Doppler. Objects of interest traveling at the same or similar speed along the crest of the wave are thus problematic to detect [1, 2].

3.1.6.1 Signal to noise and clutter ratio

The signal to noise and clutter ratio (SNR) is defined as the ratio between the wanted signal to the unwanted signal, best described by

$$\mathrm{SNR} = \frac{P_{signal}}{P_{noise}}, \tag{3.18}$$

where P_{signal} in this dissertation typically represent the power of the sea clutter and P_{noise} the power of the radar's ground noise [1].

In cases where the clutter is being analysed, commonly the signal then refers to the clutter signal and the noise refers to the radar's ground noise. In cases where a target (boat) is defined, then P_{signal} refers to the power of the reflected signal returned by the boat with respect to the unwanted background noise, which is sea clutter plus radar ground noise. In this document the analysis primarily focuses around the sea clutter signal P_{signal}, which refers to the clutter signal strength, with respect to the radar's ground noise P_{noise}, typically caused by the radar's amplification block.

Owing to the wide dynamic range of such signals, it is commonly preferred to represent these in decibels, given by

$$\mathrm{SNR}_{dB} = \frac{10 \cdot \log_{10}(P_{signal})}{10 \cdot \log_{10}(P_{noise})}. \tag{3.19}$$

3.2 SEA CLUTTER STATISTICS

In this dissertation, reference is regularly made to three statistical distributions. These are commonly used to describe the spectrum of the sea clutter signal contained within the measured radar returns [18, 19].

3.2.1 Log-normal distribution

The log-normal probability density function (PDF) is closely related to the normal Gaussian distribution. If $Y = \ln(X)$ and if Y is Gaussian, then X is said to be lognormally distributed [35, 36]. The PDF is defined by equation (3.20).

$$f_X(x|\sigma;\mu) = \frac{1}{x} \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(\frac{-(\log x - \mu)^2}{2\sigma^2}\right), \tag{3.20}$$

where σ represents the standard deviation of the logarithm and μ the mean. Figure 3.5 illustrates the PDF for multiple values of parameters μ and σ.

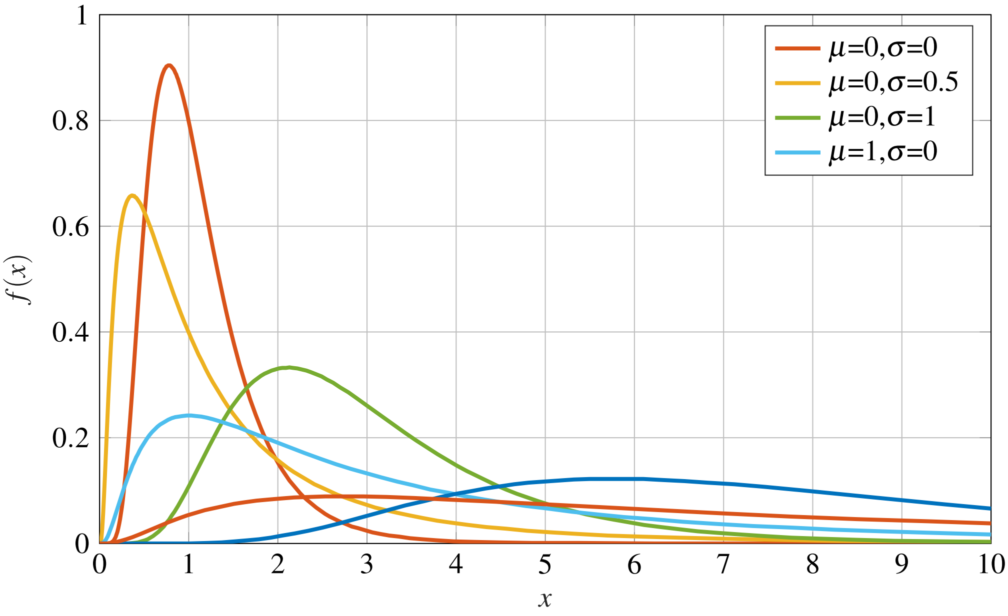


Figure 3.5. Log-normal probability density distribution function.

There are two common methods to estimate the parameters μ and σ for a given dataset. These are discussed below.

3.2.1.1 Maximum likelihood estimator

Used to estimate many distribution's parameters, the maximum likelihood estimator (MLE) maximises a derived likelihood function in order to parametrise the given distribution. For the log-normal PDF this is described by the function $\mathcal{L}(\mu,\sigma^2|x)$, which is given in [37, equation (2.1)] as

$$\mathcal{L}(\mu,\sigma^2|x) = (2\pi\sigma^2)^{\frac{-n}{2}} \cdot \prod_{i=1}^{n} x_i^{-1} \cdot \exp\left[\sum_{i=1}^{n} \frac{-(\ln(x_i)-\mu)^2}{2\sigma^2}\right]. \tag{3.21}$$

The solving of equation (3.21) as described in [37, Section 2.1], for μ and σ^2 is achieved by maximising the likelihood function and results in the maximum likelihood estimators given by the following two equations:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \ln(x_i)}{n}, \tag{3.22}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \left(\ln(x_i) - \frac{\sum_{i=1}^{n} (\ln x_i)}{n} \right)^2}{n}, \tag{3.23}$$

where $\hat{\mu}$ is the estimate of μ, $\hat{\sigma}^2$ is the estimate of σ^2, x_i is the sampled data at index i and n represents the total number of samples considered.

3.2.1.2 Method of moments

The method of moments (MoM) estimator determines the expected values for $E(x)$ and $E(x^2)$ for the given log-normally distributed data x and equates them to two known moment equations for the distribution as shown in [37]. In this case the expected values are then solved for moments m_1 and m_2, so that the parameter estimation is given to be [37]:

$$\hat{\mu} = -\frac{\ln(\sum_{i=1}^{n} x_i^2)}{2} + 2\ln\left(\sum_{i=1}^{n} x_i\right) - \frac{3}{2}\ln(n), \tag{3.24}$$

$$\hat{\sigma}^2 = \ln\left(\sum_{i=1}^{n} x_i^2\right) - 2\ln\left(\sum_{i=1}^{n} x_i\right) + ln(n), \tag{3.25}$$

where again $\hat{\mu}$ is the estimate of μ, $\hat{\sigma}^2$ is the estimate of σ^2, x_i is the sampled data at index i and n represents the total number of samples considered.

3.2.2 Rayleigh distribution

The Rayleigh distribution is a special case of the Weibull distribution. The Rayleigh distribution results from calculating the amplitude distribution of two uncorrelated, normally distributed equal variance zero mean components, such as real and imaginary components of a complex Gaussian signal. The distribution is given by equation (3.26) and is commonly used in communication theory to represent dense signal scattering due to multiple paths. It is also commonly used to describe ocean wave heights in oceanography [36].

$$f(x,\sigma) = \frac{x}{\sigma^2} \cdot \exp\left(\frac{-(x^2)}{2\sigma^2}\right), \qquad x \geq 0, \tag{3.26}$$

where σ is the scale parameter of the distribution.

Figure 3.6 illustrates the PDF for multiple σ parameters.

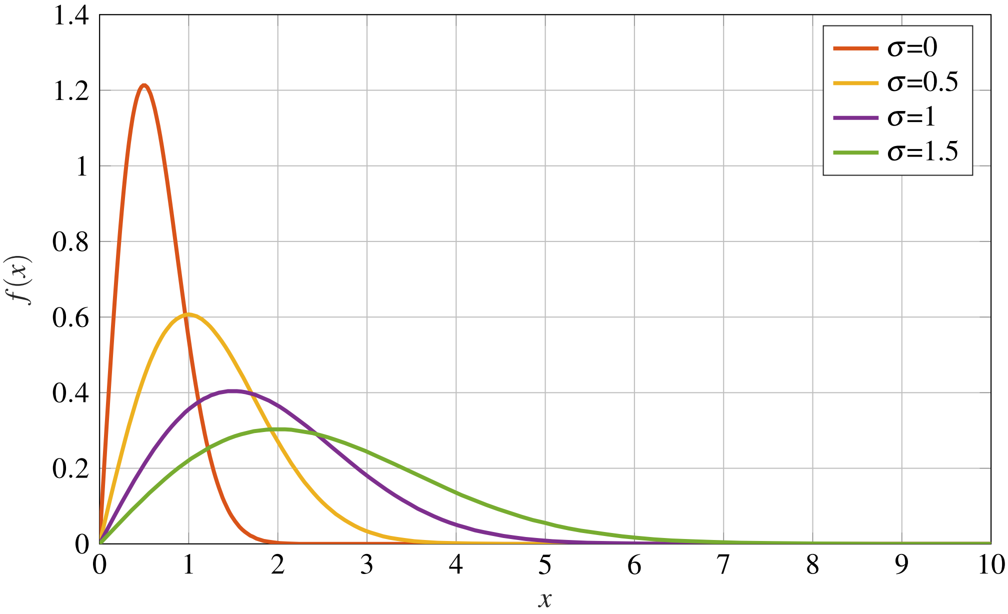


Figure 3.6. Rayleigh probability density distribution function.

3.2.2.1 Maximum likelihood estimator

The maximum likelihood estimator for the Rayleigh distribution creates an estimate for the parameter σ, so that the likelihood function of σ is maximised by solving its differential set equal to 0, the estimate is then given by equation (3.27)

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{2n}}. \tag{3.27}$$

3.2.3 K-distribution

The K-distribution has been shown to be a good estimator for maritime clutter, allowing for better representation of spikiness within the observed data compared to traditional Rayleigh, Weibull and log-normal distributions [4, 18]. Further, the K-distribution is the only distribution that can be traced back to the physical mechanisms underlying the observed data [4, 20, 38]. Two factors contribute to the stochastic nature of pulse returns for low grazing angle radars operating in a maritime environment. The first is the small capillary waves on the sea surface, which induce a speckle component in the pulse returns. The pulse returns have a relatively short decorrelation time (in the order of milliseconds) and are represented by a Rayleigh distribution [4]. The second is the result of the underlying sea swell, which induces a slow variation in the mean power level of the pulse returns. This can be associated with a wave structure over a sea surface and is represented by a generalized chi-distribution [38]. The K-distribution is derived from the Rayleigh distribution with a varying mean power implemented as a random variable that follows the chi-squared distribution.

The K-distribution is commonly given by equation (3.28) [39, 40, 41, 42],

$$f_{X1}(x) = \frac{2}{b\Gamma(u+1)}\left(\frac{x}{2b}\right)^{u+1} \cdot K_u\left(\frac{x}{b}\right), \tag{3.28}$$

where $x > 0$, $b > 0$ and $u > -1$. The standard gamma function is defined by $\Gamma(\cdot)$ and the modified Bessel function of order u is given by $K_u(\cdot)$, where u is denoted to represent the scale of the distribution and b is denoted to represent the shape of the distribution [41]. The texture of the K-distribution is typically referred to as a gamma distribution, which in turn is also chi-square distributed [39].

Alternative forms of the K-distribution can be found in literature and are shortly summarised.

The K-distribution definition given by [43] is represented by equation (3.29),

$$f_{X2}(x) = \frac{x^{(\alpha-1+\frac{N}{2})}2^{(2-\frac{N}{4}-\frac{\alpha}{2})}\sigma^{(\frac{N}{4}+\frac{\alpha}{2})}}{\Gamma(\frac{N}{2})\Gamma(\alpha)} K_{\alpha-\frac{N}{2}}(\sqrt{2\sigma}x). \tag{3.29}$$

Solving the equation for $N = 2$ it can easily be shown that equation (3.29) is equivalent to the third common form described by equation (3.30),

$$f_{X3}(x) = \frac{4x^{v}a^{\frac{v+1}{2}}}{\Gamma(v)} \cdot K_{v-1}(2x\sqrt{a}), \tag{3.30}$$

where $x > 0$, $a > 0$ and $v > 0$.

For completeness, equation (3.28) and equation (3.29) are solved with respect to the form used in this dissertation and given by equation (3.30). The results show the relationship between the scale and shape parameters which are summarised in the table below.

It can be shown that two further forms, given by equation (3.31) [15] and equation (3.32) [18], can also be traced back to the same definition. However, for this dissertation, these forms were not considered

Table 3.1. Parameter comparison for various K-distribution equation forms.

f_{X1}	f_{X2}
$b = \frac{1}{2\sqrt{a}}$	$\sigma = \frac{a}{\sqrt{2}}$
$u = v - 1$	$\alpha = v$

in depth and are illustrated only for completeness

$$f_{X4}(x) = \frac{\sqrt{2v}}{\sqrt{\mu}\cdot\Gamma(v)\cdot 2^{v-1}} \cdot \left(\sqrt{\frac{2v}{\mu}} z\right)^{v} \cdot K_{v-1}\left(\sqrt{\frac{2v}{\mu}} z\right) u(x) \qquad \text{and} \tag{3.31}$$

$$f_{X5}(x) = \frac{2b^{\frac{x+1}{2}} \cdot z^{\frac{x-1}{2}}}{\Gamma(v)} \cdot K_{v-1}\left(2\sqrt{bz}\right) u(x), \tag{3.32}$$

where $u(x)$ is the unit step function for $x > 0$.

Using the equation (3.30), a plot illustrating the influence of the scale v and shape a variables is shown in Figure 3.7.

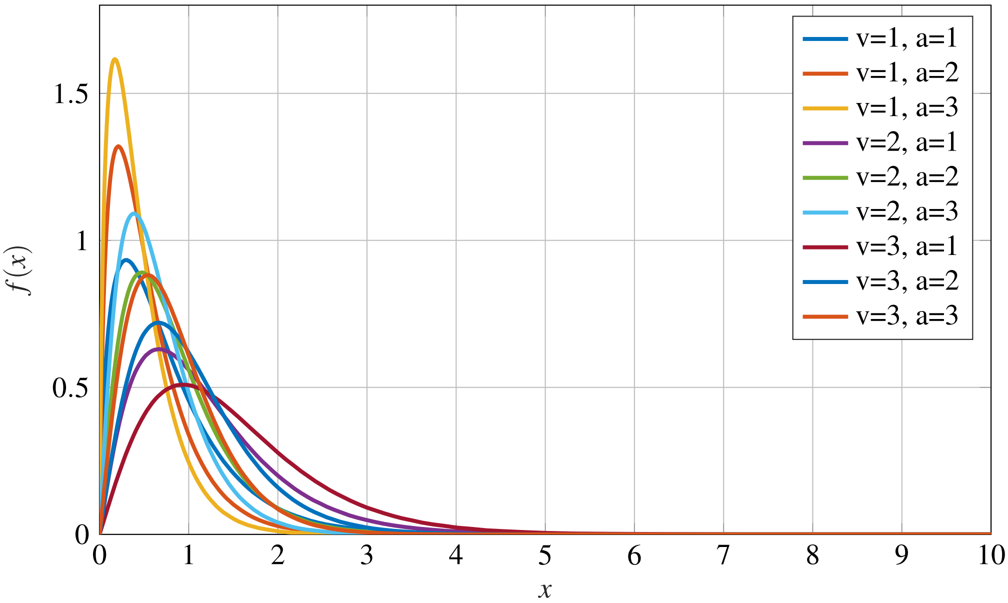


Figure 3.7. K-distribution probability density distribution function.

3.2.3.1 Maximum likelihood method of moments parameter estimation

The parameters, where v describes the scale and a the shape, are estimated using a technique described by Iskander in [40]. This method combines the maximum likelihood and the method of moments approaches to provide a computationally efficient estimate, in particular to prevent the necessity to numerically solve for the shape parameter v [4]. In [41], two moments create the relationship between the variable estimates $\hat{v}$ and $\hat{a}$. This is achieved by solving the equation as shown in [41, equation (13)] by Iskander, re-illustrated below for convenience in corrected form.

$$\begin{aligned} g_k(\hat{v}) &= \log\left(\frac{\Gamma(\hat{v}+1)}{\Gamma(\hat{v}+1+\frac{k}{2})}\right) + \frac{k\Psi(\hat{v}+1)}{2} \\ &= \frac{1}{N}\sum_{i=1}^{N}\log(x_i^k) - \log\left(\frac{1}{N}\sum_{i=1}^{N}x_i^k\right) - \frac{k\gamma}{2} + \log\left(\Gamma(1+\frac{k}{2})\right), \end{aligned} \tag{3.33}$$

where $k = 1.5$ is given to be optimal and $\gamma = 0.5772$ is defined to be Euler's constant.

Thus, as shown in [41], $\hat{v}$ can best be determined by equation (3.34),

$$\hat{v} = g_k^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\log(x_i^k) - \log\left(\frac{1}{N}\sum_{i=1}^{N}x_i^k\right) - \frac{k\gamma}{2} + \log\left(\Gamma(1+\frac{k}{2})\right)\right), \tag{3.34}$$

where g_k^{-1} denotes the inverse of equation (3.33).

Further, the estimation of parameter $\hat{a}$ is given by equation (3.35) [40]

$$\hat{a} = \frac{1}{2}\exp\left(\frac{\gamma - \Psi(\hat{v}+1)}{2} + \frac{1}{N}\sum_{i=1}^{N}\log(x_i)\right). \tag{3.35}$$

Figure 3.8 illustrates the estimation of the distribution parameters for generated data. The PDF is generated with parameters $v = 5.5$ and $a = 127$. Estimation using the MoM method is given by $v_{MoM} = 5.23$ and $a_{MoM} = 119.08$ and for the maximum likelihood method of moments (ML-MoM) method is given by $v_{ML-MoM} = 5.19$ and $a_{ML-MoM} = 118.13$.

An analysis of the estimation parameter range and its respective error towards the generated data is shown in Figure 3.9, illustrating the operating region for the ML-MoM technique as implemented for this dissertation.

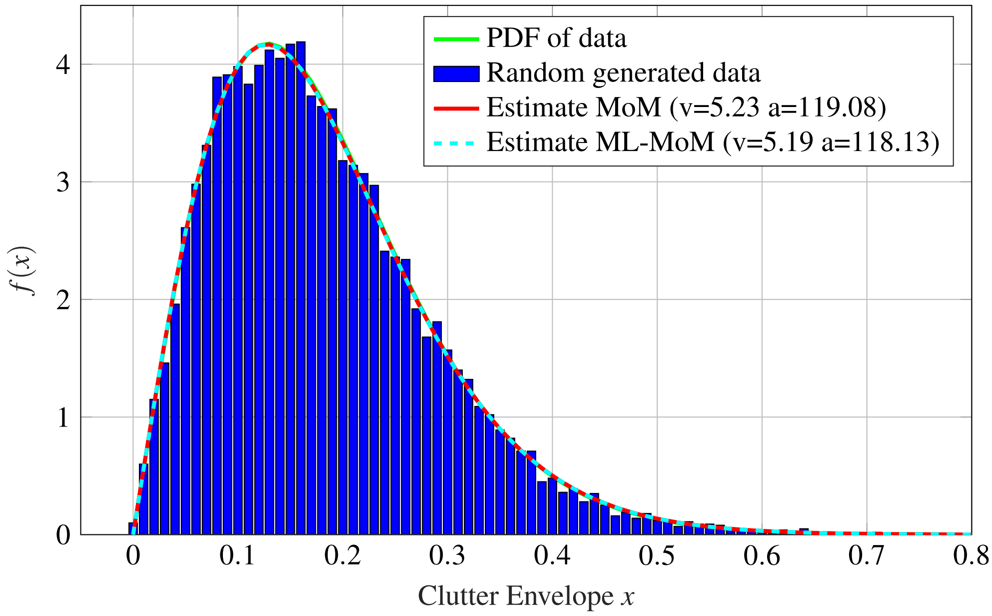
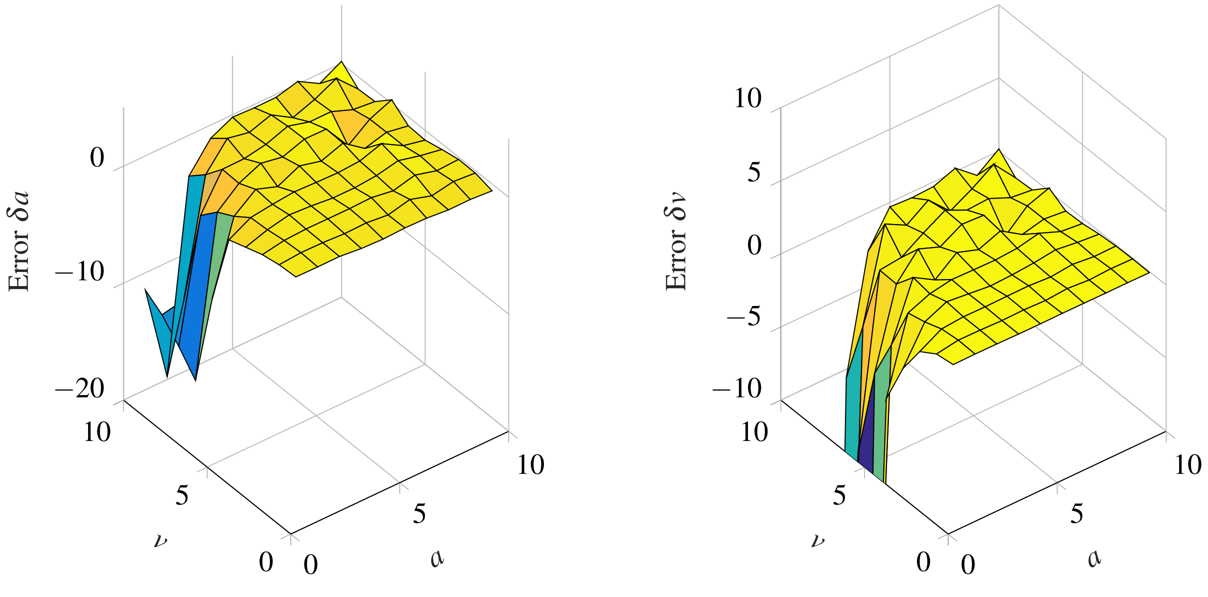


Figure 3.8. K-distribution PDF parameter estimation illustrating difference for each estimation parameter.



(a) **(b)**

Figure 3.9. K-distribution PDF parameter estimation error for (a) δa and (b) $\delta \nu$.

3.2.3.2 Rejection sampling

Accept-reject sampling is a rejection sampling method used to obtain random data from a given distribution, where no direct transform exists [44, 45, 46]. It relies on the sampling of a distribution larger than that of the desired distribution, for which a random sampling technique is defined. In this case a uniform distribution is used. When sampled data points fall within the 2D distribution of the desired distribution, the sample is accepted, else the sample is rejected and ignored. This is a form of Monte Carlo sampling, where the algorithm is repeated until a set number of accepted samples is obtained. Figure 3.10 illustrates the algorithm for 500 data points, showing the accept and reject regions.

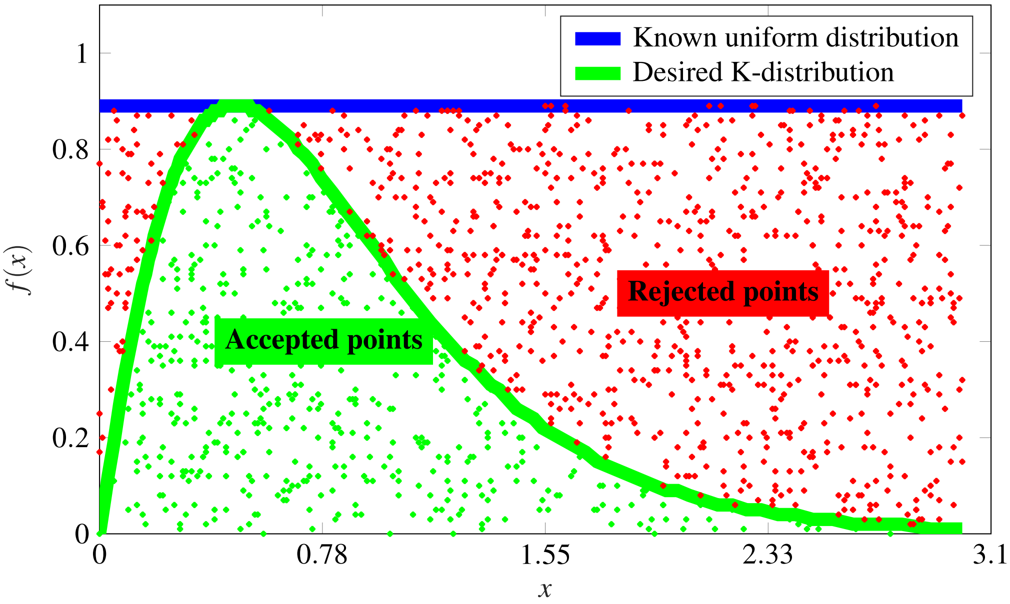


Figure 3.10. Distribution dataset illustrating accept reject sampling.

For n large samples the results can be compared against the desired distribution by the use of the histogram function. An example with 50k points is illustrated in Figure 3.11.

The performance of the sampling method is greatly dependent on the integral difference between the desired and known distributions. The larger the difference, the more samples are ignored and the less effective the algorithm becomes. The worst-case solution of a uniform distribution is used in this case as it easily provides a distribution which covers the desired distribution for all parameters v and a, with roughly 1 sampled data point for every 100 tries.

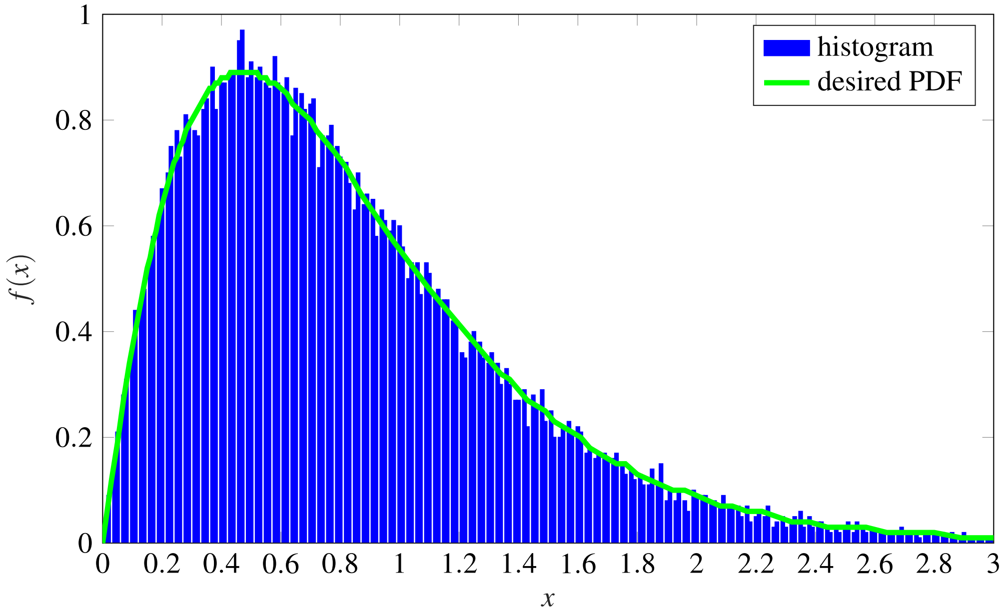


Figure 3.11. Verification of sampled data points' distribution.

3.2.3.3 Quantile-quantile Plots

In order to effectively evaluate the accuracy and goodness of fit for a selected distribution, the quantile-quantile (q-q) plot is presented. Similar to percentiles, the q-q plots n-Quantiles from the measured dataset against n-quantiles from the proposed PDF. A linear line with a slope of $m = 1$ indicates a good fit, whilst deviation from the slope indicates a poor fit. Even though the method does not provide an objective comparative number, it does provide good visual insight as to how well the head, tails and body fit the suggested distribution.

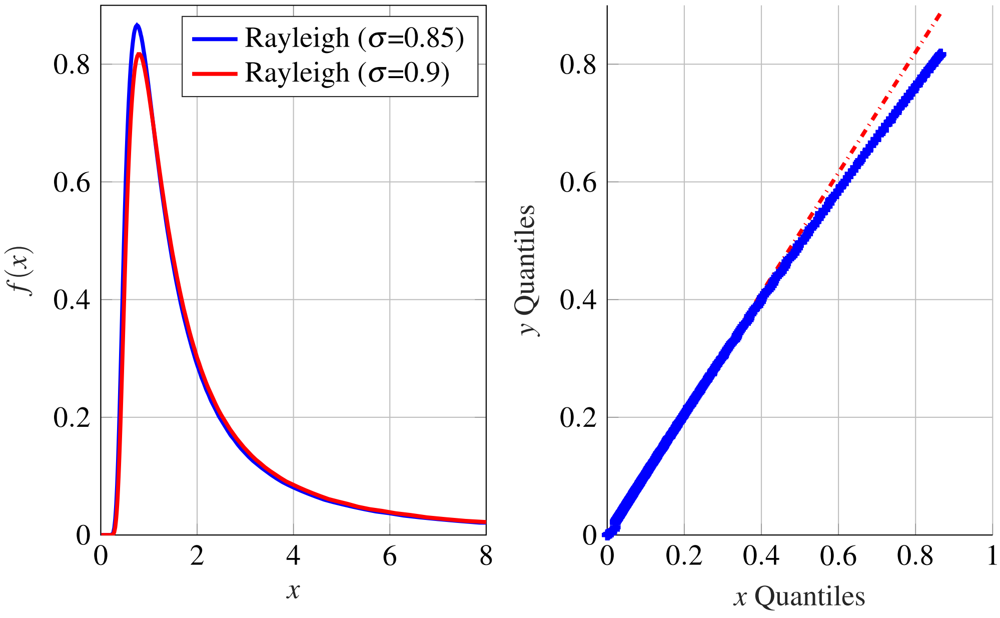
For this graphical technique, a quantile is defined by

$$Q(i) = F^{-1}\left(i - (\frac{1}{2}))/n\right) \qquad j = 1, ..., n \tag{3.36}$$

where n is the number of desired quantiles. In this dissertation n is equal to the number of sampled points allowing the highest accuracy [35, 47].

The q-q plot depicts the interdependency between sampled and theoretical quantiles. A linear plot along the $y = x$ line represents a close fitting dataset, whereas tail and head outliers towards the line indicate skewness or incorrect parametrisation of the PDF. Non similarity towards the straight line would represent a non applicable distribution [47].

An illustration of such plots is shown in Figures 3.12 and 3.13.



(a) **(b)**

Figure 3.12. Q-q plot illustrating (a) two differently parametrised Rayleigh distributions ($\sigma_a = 0.85$ and $\sigma_b = 0.9$) and (b) the corresponding q-q plot illustrating a deviation in the tail section.

As can be seen from Figure 3.12(a), two similar Rayleigh distributions are presented with marginal parameter difference. The q-q plot illustrates that the distributions are similar in that the one follows the slope of the other. However, the tail departs from the comparison distribution, indicated by the blue sample points. The deviation from the ideal line, shown in red, suggests bad parametrisation.

In Figure 3.13 a large difference between the two distributions is clearly visible. This is emphasized by the q-q plot, where the blue sample points deviate from the red slope line, suggesting that the two datasets are of different distribution types.

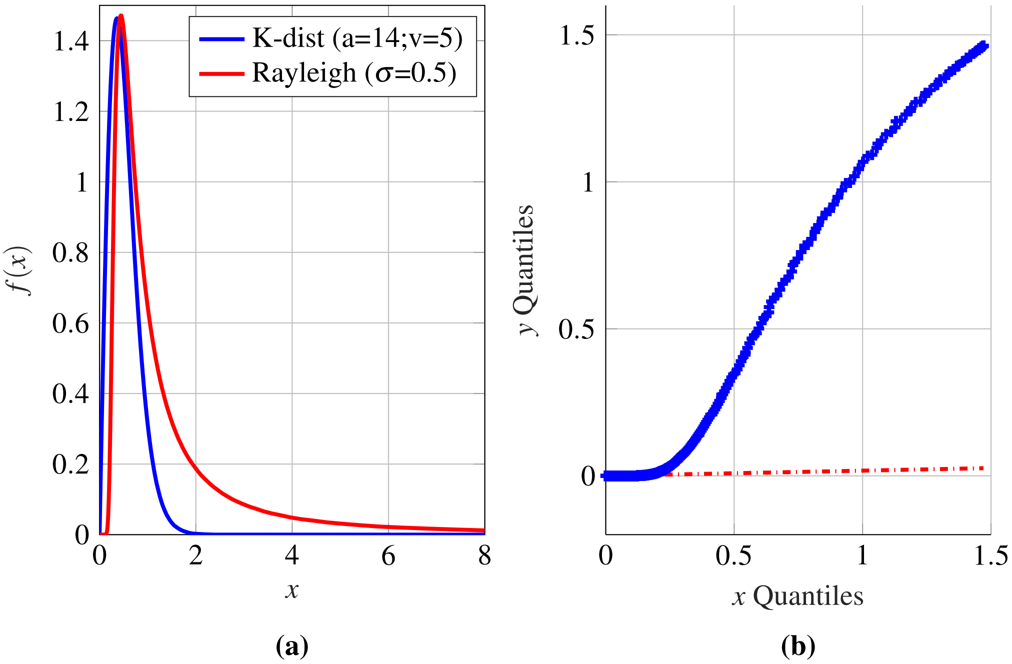


Figure 3.13. Q-q plot illustrating (a) strong difference between two distributions given by a K-distribution ($a = 14, v = 5$) and a Rayleigh distribution ($\sigma = 0.5$) and (b) the corresponding q-q plot illustrating a deviation from the slope line indicating a poor distribution choice.

3.3 DETECTION

3.3.1 Constant false alarm rate

A constant false alarm rate (CFAR) detector is an adaptive filtering technique used to detect targets. The detector is tasked with deciding whether radar returns are assigned to the target-plus-interference hypothesis (H_1 hypothesis) or the interference-only hypothesis (H_0 hypothesis).

A dynamic threshold is used to distinguish between noise and target information by comparing their respective amplitudes. The detector threshold is calculated to achieve the maximum probability of detection for a given SNR and a defined probability of a false alarm P_d. This is the most important and difficult parameter to specify for the detector.

A false alarm is defined to be an interference signal that exceeds the defined threshold and thus represents a false detection of a target where there is no actual target. For a given radar design, the probability of a false alarm is usually defined by the radar specifications.

In clutter environments the parameters of the assumed environmental distribution are generally unknown. Part of the CFAR algorithm is thus to sample data to estimate these parameters, consequently allowing the algorithm to dynamically change its threshold for the measured clutter. This allows clutter with varying strength to still be effectively filtered.

If the threshold is decreased, then the probability of detection increases along with the number of false detections. If the threshold is too high, then the probability of detection is very low and many actual targets will most likely be misinterpreted as clutter.

Typically, a radar configuration is setup so that a given probability of detection (or a given false alarm rate) is achieved. However, in the case of standard sea clutter, the measured spectrum is constantly changing in amplitude and distribution. Thus, an adaptive detection method is needed for each evaluation in order to allow the detector to adapt to a changing environment [1, 2].

3.3.2 Cell averaging constant false alarm rate

The most common form of adaptive sampling is cell averaging CFAR (CA-CFAR) . For this detector the noise floor is estimated by considering a number of cells around the cell of interest, defined as the cell under test (CUT) . Further, to avoid power leakage across the cells of interest and those for estimating the noise floor, guard cells around the CUT are inserted. This reduces the influence of the target on the leading and lagging cells for noise floor power estimation.

The noise floor power is estimated using the following equation:

$$P_i = \frac{1}{N} \cdot \sum_{i=1}^{N} abs(x_i), \tag{3.37}$$

where P_i denotes the power of the noise floor and N the number of training cells. Figure 3.14 represents the relation between CUT, Guard-Cells and Training-Cells.

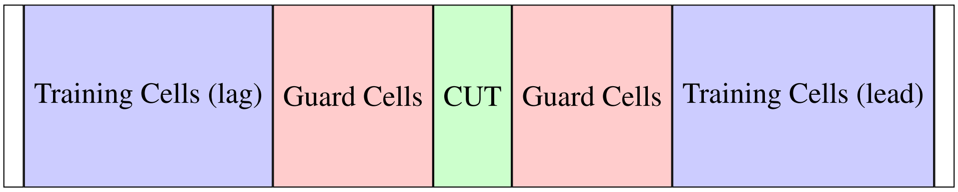


Figure 3.14. CFAR cell structure illustration.

Figure 3.15 illustrates the sensitivity and difficulty in determining the detector threshold T_{thres}. A PDF of the noise spectrum (red) vs the PDF of a target (blue) is illustrated. An ideal detector threshold is illustrated by the black dashed line. The area below the curve of the blue PDF of the target and to the right of the detector threshold is the probability of detection region P_D. A false detection would occur if an interference (clutter plus noise) sample occurs right of the threshold. The probability of such false alarm events (given by P_{FA}) is equal to the area under the red PDF curve (the distribution of clutter and noise) which is to the right of the threshold T_{thres}.

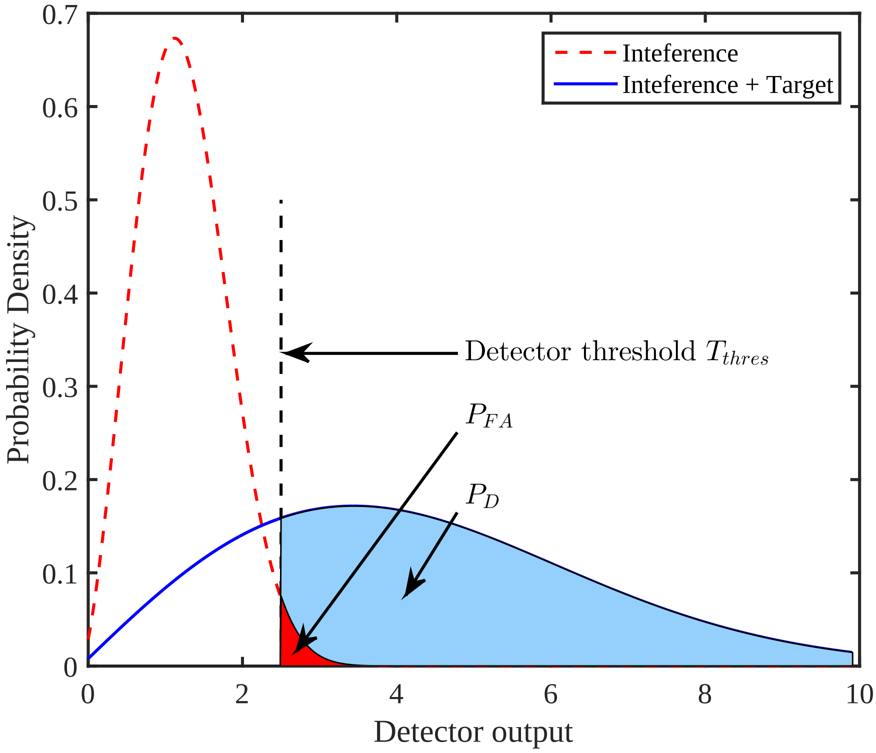


Figure 3.15. Illustration of threshold filtering with respect to the P_{FA} and P_D areas.

CHAPTER 4 METHOD

In this Chapter a short review of the methods implemented in the algorithm, presented in this dissertation, are provided. The methods are discussed using a generic dataset to aid in the descriptive function of each block below. Section 4.1 discusses an underlying assumption of the presented process. Section 4.2 provides an overview of the numerical and statistical methods that are implemented throughout this dissertation. These steps have the goal of analysing, parametrising and generating a low grazing angle sea clutter model for periodic breaking waves in a littoral coastal zone and are discussed in more detail in Sections 4.3 to 4.5. Essential to this is the angular correction of the approaching breaking waves prior to statistical analysis. This is discussed in Section 4.6 and 4.7. The measured radar returns are corrected such that the breaking waves are parallel to the x-axis. This allows individual phase segments of the ocean wave, i.e. crest and troughs, to be analysed separately. With the correction applied to a wider and more significant section of measured radar data, parametrisation provides insights into the alternating statistical structure of the measured sea waves as discussed in Section 4.8. Through the determination of appropriate parameters, a generation method is developed which allows for the output of statistically similar low grazing angle radar data.

4.1 ASSUMPTIONS

The approach angle of the breaking wave, with respect to the radar, is determined through a type of Hough transform known as the Radon transform. This integral transform effectively returns line intensities within a two dimensional matrix by integrating over all possible angles and positions. The transform is explained in detail in Section 4.4.

Given the determined angle, the data is shifted and corrected such that the breaking waves are now parallel to the x-axis. An assumption is made at this point, namely that the fractional change in grazing angle due to the angular correction of the measured data is considered negligible in terms of the clutter statistics. The maximum change is given by

$$\Delta\theta_{gr} = \arcsin\frac{h_r}{\Delta R}, \tag{4.1}$$

where h_r is the height of the radars antenna and ΔR is the change in range due to the shift in the data matrix.

After data processing, statistical parameters are obtained. Statistical correlation times and distribution parameters are recorded and used for the data generation process. Based on these models, methods and simulations, statically similar data can then be generated by applying part of the process in reverse.

4.2 ALGORITHM

A general overview of the algorithmic process, implemented throughout this dissertation, is shown in Figures 4.1 to 4.3. This first section (see Figure 4.1) of the algorithm concerns itself with the preparation of the measured radar returns for analysis and parametrisation. The individual processing blocks are first introduced in Figure 4.3 and then discussed in detail in the sections that follow.

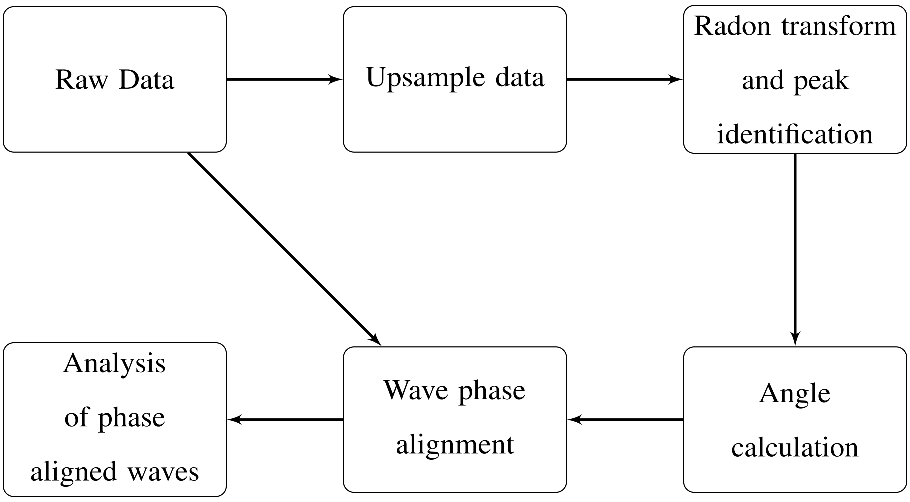


Figure 4.1. Process overview of angle correction algorithm.

Raw radar data refers to the digitally sampled complex measured radar returns obtained after the receiver stage from a coherent X-band radar with an I/Q intermediate frequency sampler. In order to make processing easier, it is common to organise the measured data in the form of a two dimensional matrix. As illustrated in Figure 3.4, the returns between pulses are recorded over time. Fast time returns are known as the measured returns that occur due to one transmission pulse. These returns indicate an increase in target distance (range) over time, thus the further the target is away from the radar the longer it takes until the signal is returned. Slow time refers to analysis of the same range distance bin over time, i.e. over multiple transmission pulses. At low grazing angles, these returns typically have a low number of resolution range bins (fast time), where individual gate sizes vary between 15 m and 45 m for the provided measurements. Slow time is merely limited to the recoding time.

In order to generate a square matrix of data, as needed by the Radon transform, either the number of returns (slow time) can be reduced to match the number of range bins, or the range resolution (fast time) can be increased through an up-sampling process. Although simpler, decreasing the number of returns produces insufficient accuracy. When the number of input samples to the transform is increased, the angular accuracy is improved as the transform has sufficient data. Thus for the square matrix an up-sampling process is implemented. The Radon transform block together with the angle calculation block is then responsible for determining the approaching angle of the ocean wave crests with respect to the radar. The wave phase alignment block can then correct the original measured data, which may be many times longer over slow time, such that all wave crests are parallel to the x-axis. This allows for simplified statical analysis of individual wave sections, which will be discussed in the next section. This step provides the basis for the statistical analysis of the remaining data.

The next section of the algorithm analyses and parametrises the angular corrected data, as illustrated in Figure 4.2. The processing block inputs the angular corrected and aligned measurement data, where the length of the data is theoretically unlimited. However, due to computational limits in terms of personally available processing and memory resources, radar data in excess of 30 s was not considered. Data of this length stored in complex double form results in an input matrix of $\approx$ 100 MB ($\frac{\approx 64\text{ bit}\times 2\times 64000\text{ samples}\times 96\text{ range bins}}{8}$). A radon transform of this data using a resolution of 0.1 degrees results in an output of $\approx$ 1GB. This limitation is further discussed in Section 4.4. Therefore, the data is firstly analysed for appropriate statistical distributions where, due to data from individual distributions, individual sections of the wave can be considered. These results are recorded and all necessary parameters for the generation are noted. Hereafter the correlation section investigates the correlation in fast time (i.e. between range bins) in order to determine parameters such as periodicity between ocean wave crests and troughs. Again these parameters are recorded and noted, as they are needed by the final processing section.

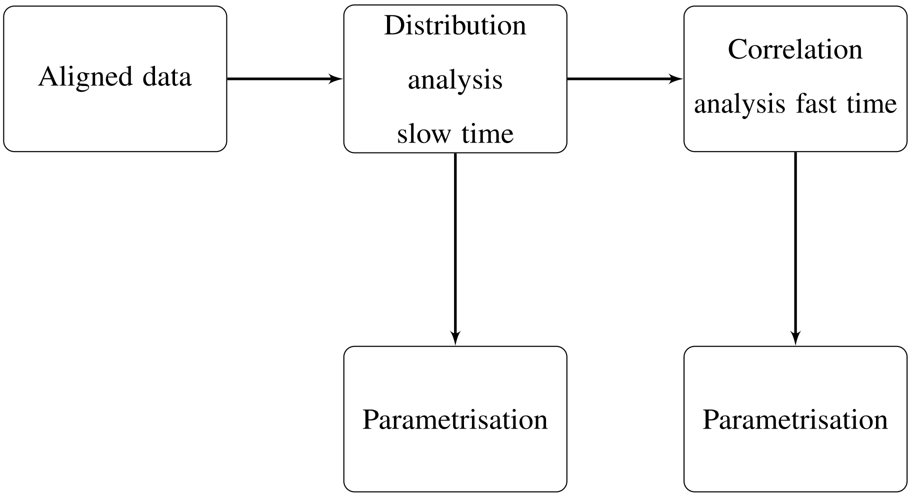


Figure 4.2. Process overview of the analysis and parametrisation algorithm.

Finally, after parametrisation, the entire process is reversed. The processing chain is shown in Figure 4.3. Fed by a random generator which generates independent complex samples, the wave generation process creates periodically structured sea clutter with a given statistical correlation and distribution. Finally, the generated data is then shifted to account for the ocean wave approach angle with respect to the radar position.

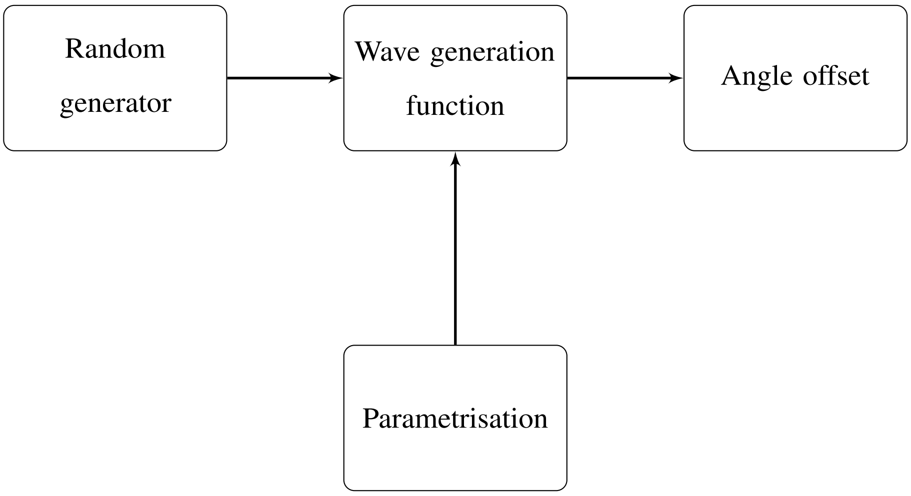


Figure 4.3. Process overview of data generation algorithm.

4.3 DATA UP-SAMPLING

In order to obtain unbiased and comparable results from the Radon transform, the input data needs to be a square matrix. Oblong input matrices cause varying integration lengths for different rotation angles, which is undesirable. This causes integration results that are not normalised with respect to each other and effectively distorts the output. This complication is explained in more detail in Section 4.4 of this chapter. The outcome is that square input data matrices are required.

4.3.1 Basic principle

The radar returns used for this simulation typically have a fast time width of 96 returns and a slow time in excess of minutes, which with a PRF of 2.5 kHz to 5 kHz, creates a highly rectangular data matrix (eg. 96×67334 for TFC15 data). In order to obtain square data matrices, the given data needs to be increased in resolution through interpolation. This concept is illustrated with a simplified matrix shown in the Figure 4.4. To address this challenge and obtain a square matrix for preprocessing, several methods to increase the fast time resolution are evaluated and include:

- zero order hold (block copy),
- first order hold (linear interpolation), and
- finite impulse response filter based interpolation (complex interpolation).

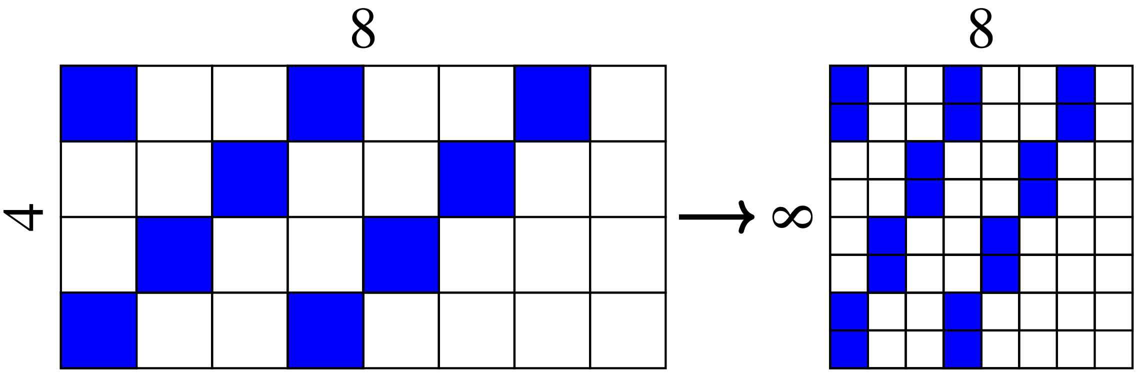


Figure 4.4. Creating a square matrix from measured data, where the individual squares indicate the matrix elements and size, and the blue highlighted cells illustrate a simple wave crest amplitude within the matrix.

4.3.2 Zero order hold filter

The zero order hold filter firstly determines the number of missing rows in the y-axis of the matrix for it to be square. These are then padded with zeros. Next a square pulse is convolved over the sampled data, effectively copying the existing data into the missing rows and making the output matrix square. The disadvantage of this method is that due to the replication of data over several fast time ranges, clear identifiable steps are produced in the output. Although the method is fully reversible, the processing leads to steps in the output of the Radon transform, identifiable as high frequency steps within the filter response.

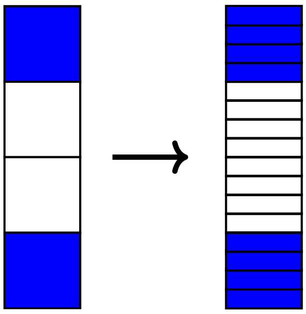


Figure 4.5. Zero order hold filter for a single time step within the measured data matrix illustrating the increase in fast time resolution.

4.3.3 First order hold filter

A more accurate method which causes less high frequency distortion is to use linear interpolation between the missing fast time rows. This method linearly estimates the values between the start and end values for the missing rows. Again, the data is first zero padded to represent the desired output resolution and then a triangular impulse response is convolved over the data. This method produces fewer high frequency steps due to the reduction of instantaneous changes in the slope of the curve. However, the complex data of the returns are not captured in the added rows. Further, this method assumes that the relationship between individual range bins is linear by nature.

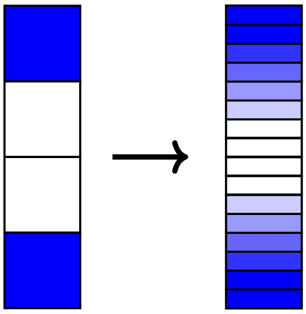


Figure 4.6. First order hold filter for a single time step within the measured data matrix, illustrating the linear transition in amplitude for an increased resolution in fast time.

4.3.4 FIR based filter

Zero and first order hold filters effectively increase the sampling rate of the discrete time signal. However, they also introduce several undesired distortions into the output signal. Thus an improved technique for increasing the number of discrete samples is required.

The fundamental principle underlying discrete time up-sampling is best described by Figure 4.7. This method also applies to the comprehensible methods already discussed. The discrete time input signal $x_d[n]$ is firstly expanded in time by inserting zeroes between each discrete time sample, described by $x_e[n]$. M describes the increase in the sampling rate. Finally, this expanded output is then filtered by a low pass filter which smooths the output result and produces the desired output signal $\hat{x}[n]$. All interpolations described thus far are therefore dependent of the implementation of the low pass filter.

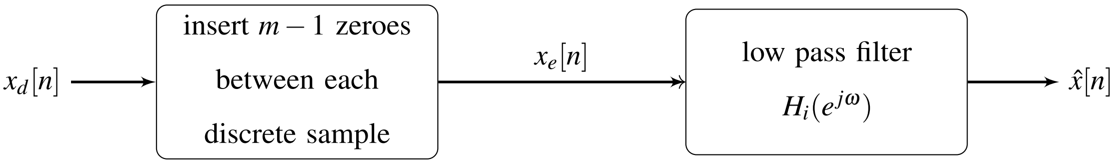


Figure 4.7. Discrete time signal interpolation process flow.

As stated $x_d[n]$ is the discrete time sampled signal that is returned from the radar. $x_e[n]$ is the zero expanded signal and can be described by Equation 4.2. $x_e[n]$ is equal to $x_d[n]$ only if $\frac{n}{m}$ is an integer, otherwise the signal is a zero. This equation can be rewritten in the form of Equation 4.3 by using the discrete time delta dirac function (Kronecker delta function) which is 1 only at $n = km$. M represents the increase in sampling frequency.

$$x_e[n] = \begin{cases} x_d\left[\frac{n}{m}\right] & \text{, where } \frac{n}{m} \in \mathbb{Z} \\ 0 & \text{, otherwise} \end{cases} \tag{4.2}$$

$$= \sum_{k=-\infty}^{\infty} x[k]\delta[n-km] \tag{4.3}$$

Using the definition of the discrete time fourier transform (DTFT) provided in Equation 4.4, the spectrum of $x_e[n]$ can be determined by inserting Equation 4.3 into Equation 4.4. This frequency spectrum of the expanded discrete time signal $X_e(\omega)$ is shown in Equation 4.5.

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \tag{4.4}$$

$$\begin{aligned} X_e(\omega) &= \sum_{n=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} x[k]\delta[n-km] \right) e^{-j\omega n} \\ &= \sum_{k=-\infty}^{\infty} x[k] \left(\sum_{n=-\infty}^{\infty} \delta[n-km]e^{-j\omega n} \right) \\ &= \sum_{k=-\infty}^{\infty} x[k]e^{-j\omega mk} \\ &= X_d(\omega m) \end{aligned} \tag{4.5}$$

From the result of Equation 4.5, the scaling in the frequency by a factor m is clearly visible with respect to the input signal. This result is also visually represented in the frequency domain in Figures 4.8 and 4.9.

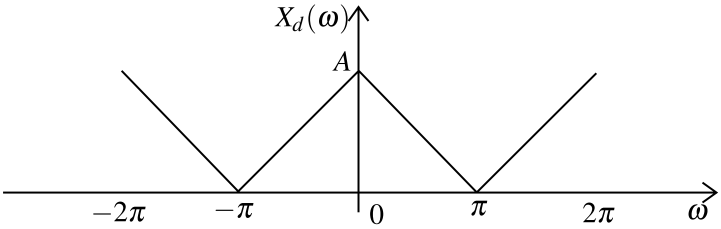


Figure 4.8. Illustrative spectrum of sampled signal $x_d[n]$ with no scaling.

For illustrative purposes a scaling factor of $m = 3$ was chosen for Figure 4.9. In order to obtain the original signal at a higher sampling frequency, the spectrum of the original signal must remain and the duplicates must be removed. This is achieved by filtering the data using a low pass filter.

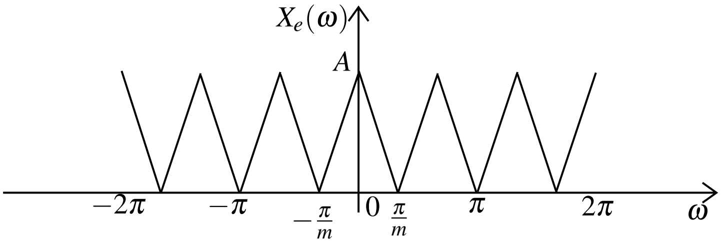


Figure 4.9. Scaling the input signal x_e by a factor of $m = 3$ shown in the frequency domain.

In order to isolate the periodic signal in the frequency domain and smooth the signal in the time domain, only every m-th pulse in the spectrum must remain after filtering. This is easily achieved by filtering the periodic signal with a low pass filter. The filter requires a passband over the region $-\frac{\pi}{m}$ to $\frac{\pi}{m}$ to isolate the desired repetition and with a gain of m to compensate for the amplitude change.

Applying an ideal low pass filter results in Figure 4.10. The red lines illustrate the ideal low pass filter and the black lines the resulting output spectrum $\hat{X}(\omega)$.

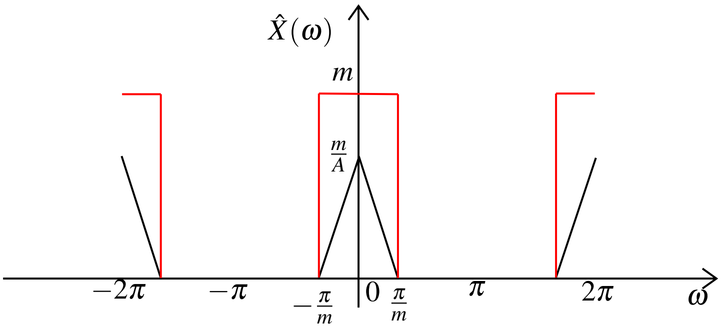


Figure 4.10. Low pass filter shown in red applied to $X_e(\omega)$ shown in black.

Such an ideal filter however cannot be implemented, especially as there is no provisioned gap between the spectrum for a passband to stopband transition. Thus a tradeoff between passband ripple, transition band width and stopband attenuation needs to me made.

4.3.4.1 FIR filter design of low pass

A finite impulse response (FIR) filter based on the windowing method is initially considered in this section. A Kaiser window approach specifies acceptable ripple for the passband and stopband, as well as the transition widths in order to specify the filter. The Kaiser window is then multiplied with the ideal impulse response for the given passband in order to generate realistic passband filter coefficients. The method was designed using an iterative process. Very small ripple specifications cause the side-lobes to increase, which causes the stopband to inject slight amounts of unwanted information into the passband, essentially resulting in an over-fit of the interpolated data.

In order to optimise this method a second approach, known as a constrained least squares filter (CLS), was considered. This filter considers the design of the FIR filter as an optimisation problem. First the cutoff frequencies for the passband are specified, then the ripple threshold. The transition width is then implicitly calculated. An error function, specified as the difference between the desired and actual filter, is minimised during design in order to find a filter that is closest to the design specifications. This method produces acceptable results with reduced over-fitting compared to interpolation.

4.3.4.2 Constrained least squares filter

Complex interpolation is based on filtering the data through a specially designed FIR filter based on the CLS method. The FIR filter results in a passband with a steep cut-off. Designed in the frequency domain, the filter is set to start slightly before the maximum sampling frequency fs_{max}, causing reduced reversal at the cost of slight information loss in the passband. This effectively ensures that only very little information from the outside stopband is within the final interpolated output data. This causes the interpolation function not to match all data points exactly, giving the system more flexibility and creating smoother output data. The original data is padded with zeros to expand to the desired resolution. In this case the resolution is increased 15 fold, from 96 range bins to $96 \times 15 = 1440$ range samples. The original zero padded data is convolved with the filter, thus interpolating the data.

Figure 4.11 illustrates the frequency response of the applied filter. Note the very steep cut-off frequency, preventing side-lobes from inserting non-relevant data into the filtered data in the form of over-fitted information. The ripple obtained by the optimised filter is shown in Figure 4.12 and estimated to be around $|H(\omega)_{Ripple_{p2p}}| = 0.01$ dB.

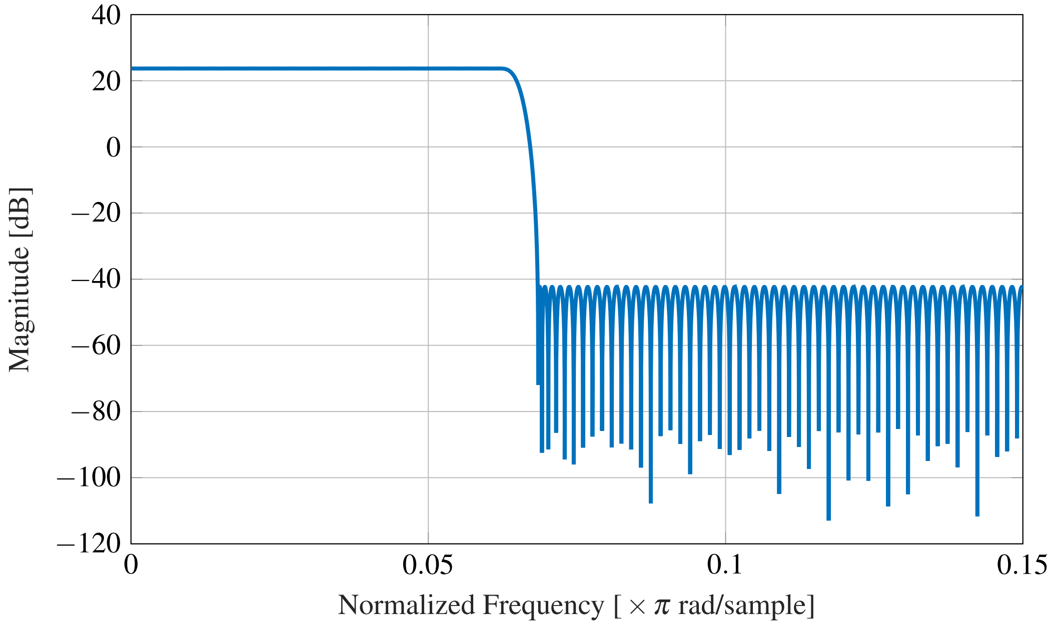


Figure 4.11. Frequency response of FIR filter illustrating a steep cut-off frequency with 60 dB drop.

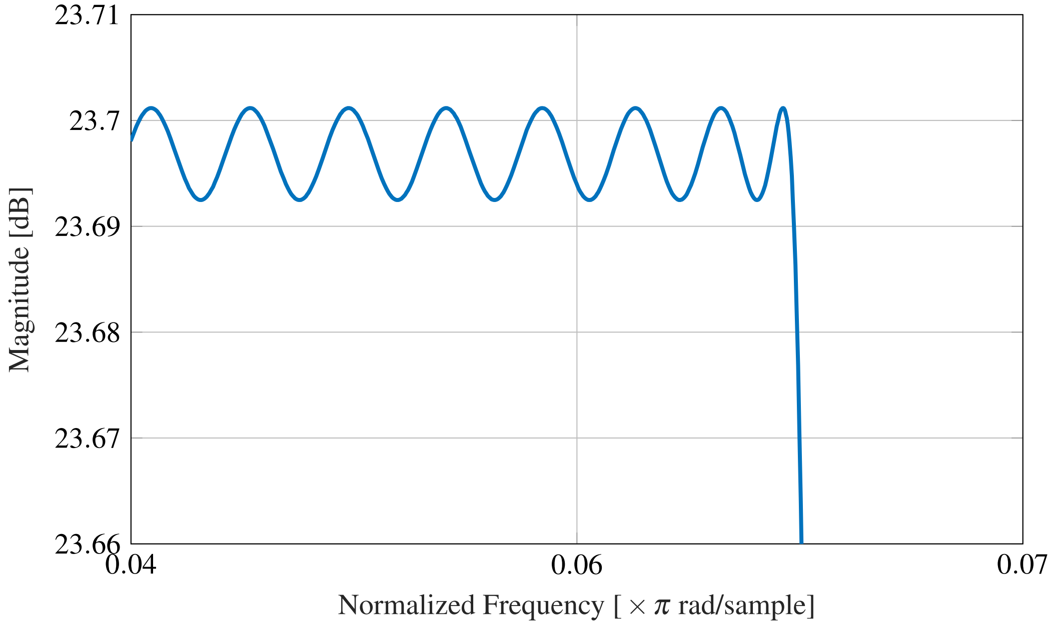


Figure 4.12. Ripple response of passband for the given FIR filter.

Figure 4.13 shows the original data points overlaid together with the interpolated data. It clearly illustrates the fitting accuracy, showing little to no over fitting due to the steep passband with a very low stopband response of the filter. Further, the data indicates the increased resolution obtained and the respective zero padding for a single range bin, in this case range bin number 45.

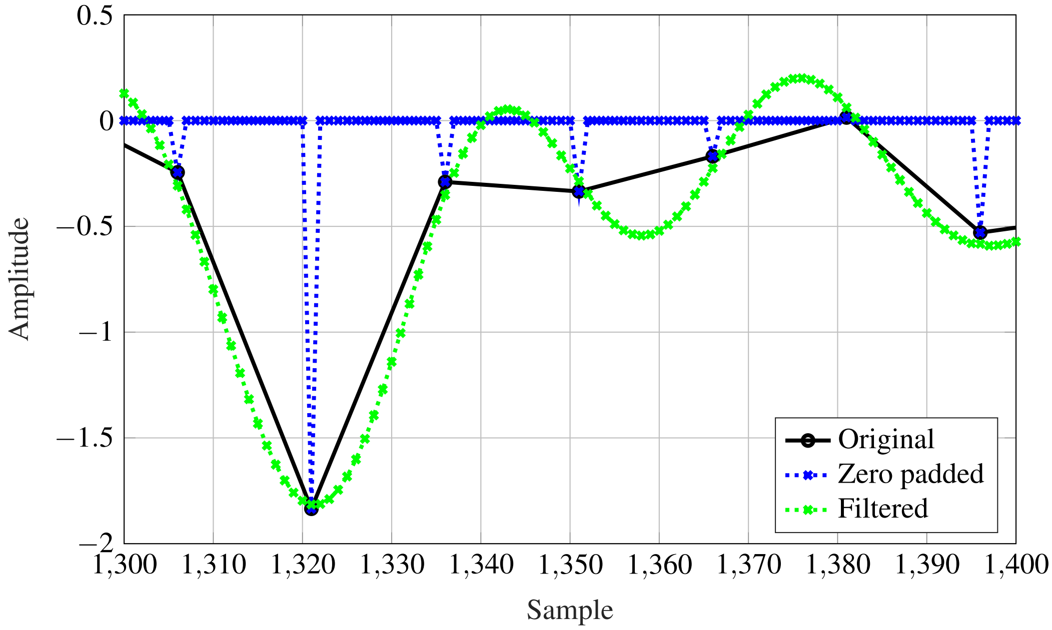


Figure 4.13. CFA16_002 CLS filter interpolated data points illustrating the filtered data matching the original sample points.

4.3.4.3 FIR output

The output of the generated data, plotted together with the original radar measured data on an in-phase and quadrature components (IQ) plot, shows the original data passing through the complex filter $|H(\omega)|$, where in-phase and quadrature components respectively refer to the real and imaginary components of the complex signal. This is shown in Figure 4.14 for the first 10 measured data points indicated in black, with red marking the start of the sequence. As mentioned above, due to the relaxed constraint on the maximum sampling frequency fs_{max}, not all data points are accurately pass through. This was done so that no over fitting occurs, preventing the filter's side-lobes from adding non-realistic data into the interpolated output results.

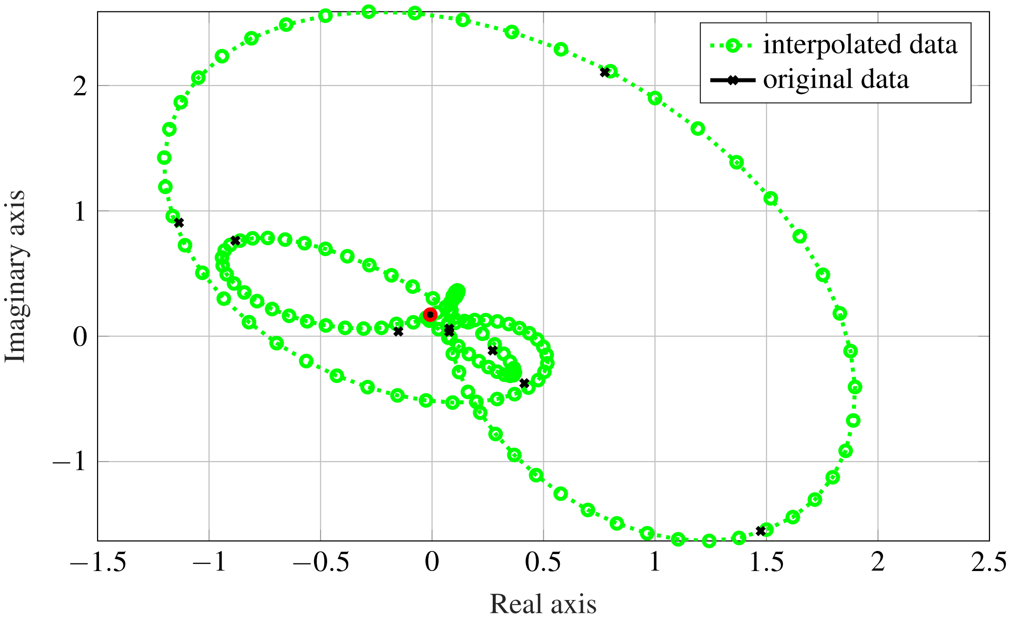
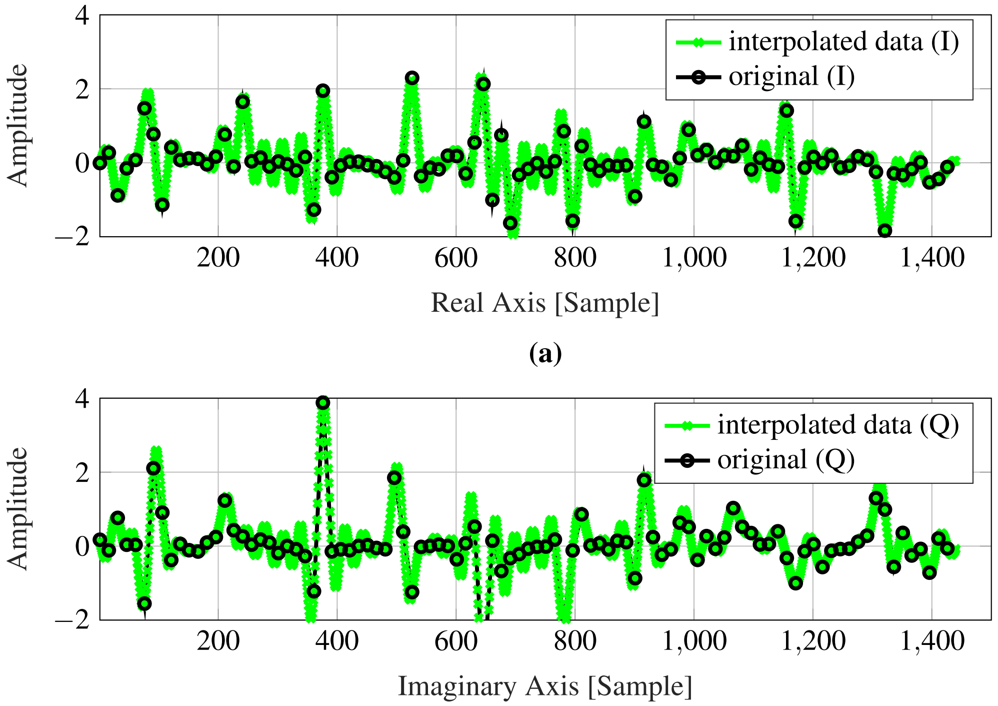


Figure 4.14. CFA16_002 IQ plot showing data points passing through the filter.

Figure 4.15 emphasises the real and imaginary components of the output result separately over time. The increased resolution data shown in green is plotted against the original data points shown in black. A section of 26 original resolution data points and 400 high resolution data points have been illustrated in Figure 4.15 in both the real and imaginary domains.



(a)

(b)

Figure 4.15. CFA16_002 output result from FIR based interpolation along (a) the real axis and (b) the imaginary axis.

4.3.5 Filter selection

In order to determine which filter method best suits the application, the fast Fourier transform (FFT) was applied to the interpolated data. This provides insights into the response spectrum of the filter and the effects the interpolation may have on the result spectrum for a given dataset. The FFT (with DC component) was performed over all 1440 range bins and then averaged to give a representative spectrum response as shown in Figure 4.16.

The difference in the spectrum is shown in Figure 4.17, illustrating the improved performance that the FIR filter offers over the other two filters. Side-lobe characteristics for the zero order based filter show recurring undesired harmonics at 10MHz intervals (given by 100 ns sampling time), whilst the FIR clearly shows improved response with respect to both the zero and first order hold filters. Based on these outcomes, the FIR was primarily considered for further processing. This selection was additionally verified by analysing the outcome of the algorithm when the other two filter types were

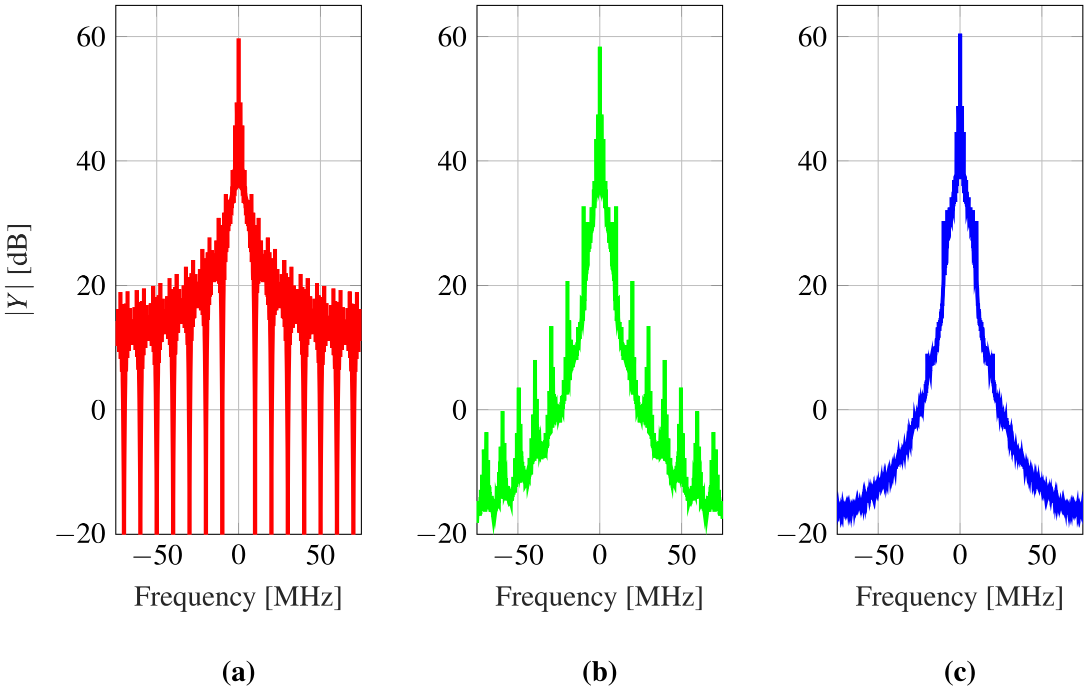


Figure 4.16. CFA16_002 comparison of filter response curves for (a) zero order (b) first order and (c) FIR based filters (full BW of filter is given by 150 MHz).

utilised. The analysis was conducted by comparing the results obtained from the corrected measured data step of the algorithm.

4.4 RADON TRANSFORM

A specific case of the Hough transform, the Radon transform is required in order to determine the angle at which the wave crests are approaching the radar over time. The Radon transform identifies the line-like ocean wave crest structures in the measured radar data over time and leads to the alignment of these within the output matrix, so that the crest structures are parallel to the x-axis.

4.4.1 Basic principle

In order to align the ocean wave crests to the x-axis in the output matrix, a Radon transform is applied to determine the angle of the sea wave towards the x-axis. Applied to this two dimensional matrix, which represents signal intensities over range and time, the Radon transform is broadly defined as

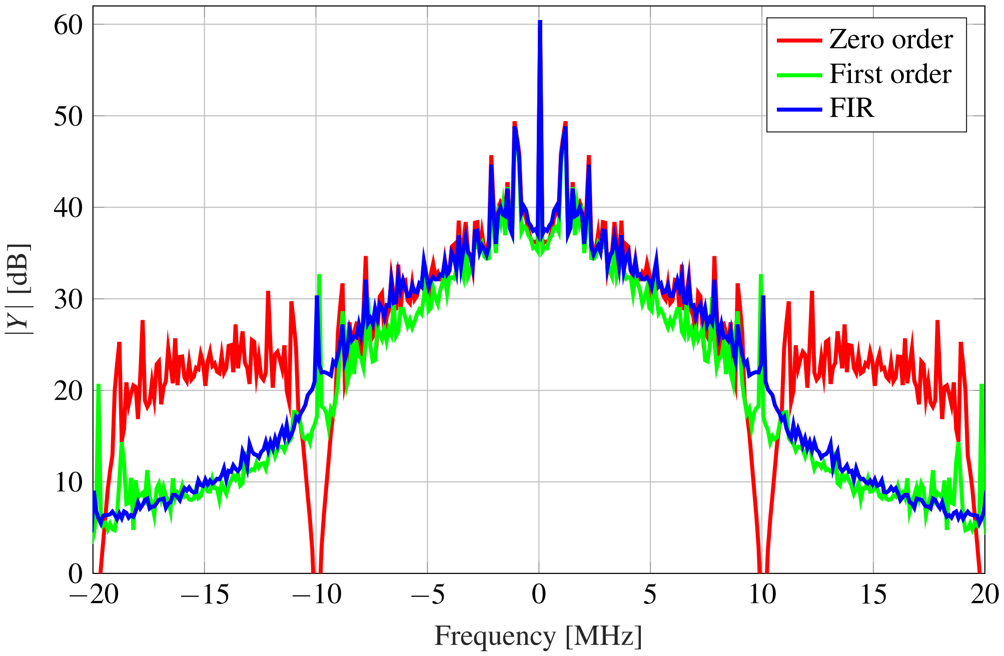


Figure 4.17. CFA16_002 average FFT magnitude response for a 40 MHz region after interpolation (full BW of filter is given by 150 MHz).

the intensity given by any line integral (or the projected function sum) at a given angle and offset to the origin of the transform output matrix. This projection is varied over 180° in order to analyse all possible rotations of the function data and over a given step size from origin. The intensity is taken to represent the strength of the function for a particular angle and a particular offset from the center point of the image [48].

In the case of a two dimensional image given by $f(x,y)$, the Radon transform can represent the image intensity over all angles and all depths of the image dataset, converting line qualities into point maps. This characteristic is what makes the Radon transform ideal for the detection and identification of lateral parallel waves. From this transform it is possible to calculate the equivalent line parameters over the image, allowing one to compute the motion vector of a given wave or boat, which is mathematically described by equation (4.8). Due to its symmetric nature, the transform is limited around the angle range as defined in equation (4.6).

$$\theta \in [0^\circ, 180^\circ) \tag{4.6}$$

where 180° is not included as it is identical in result to the angle 0°.

The Radon transform is given by

$$\rho = x\cos\theta + y\sin\theta \tag{4.7}$$

$$R(\rho,\theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\cdot\delta(\rho - x\cos\theta - y\sin\theta)\,dx\,dy, \tag{4.8}$$

where the parameters are illustrated in Figure 4.18 and δ is the Dirac delta function.

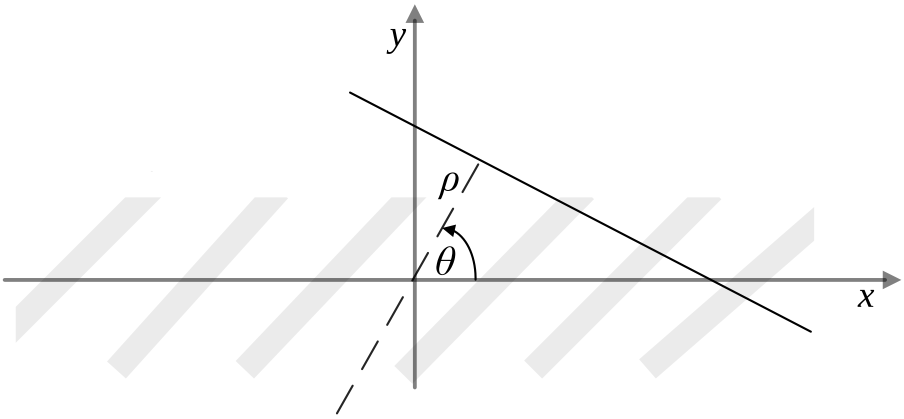


Figure 4.18. Functional illustration of the Radon transform integration lines, indicating the rotation θ and the perpendicular offset ρ from the center of the transform input matrix, where the shaded gray lines indicate amplitudes resembling those of ocean wave crests within the dataset.

Figure 4.18 depicts the function of the Radon transform. The block with grey bands illustrates periodic wave crest amplitudes in the data matrix. θ defines the angle of the dashed line and ρ the length of the dashed line with respect to the origin of the transform input matrix. The solid line indicates the line over which the integral is performed. Thus for a given θ and ρ a transform value is determined by integrating the data matrix values underlying the solid line. This is then completed for all angles and offsets for the entire input matrix.

The Radon method, due to its line based nature, offers considerable strength over similar image processing techniques which are commonly used to determined edges. These include techniques such as Canny and Sobel edge detection [49]. These methods rely on a differential of the input matrix. Based on the pixel differences between data points, regions are identified where possible edges may lie. After determining possible edges, an algorithm is required to identify and classify these edges into straight lines by linking possible neighbouring pixel points to possible lines [49]. In addition to this, clutter complicates the detection principle even further by adding severe measurement noise to the process. The Radon transform, on the other hand, identifies lines through integration, which produces acceptable results even in heavy noise situations, such as those seen in typical sea clutter data [50].

A subsection of the input data taken prior to upsampling, shows the highly rectangular matrix form of the matrix in Figure 4.19. The matrix is plotted on an equal axis system. The dimensions of this matrix are 1440 × 96 elements. An example of a Radon transform perfomed on this dataset is shown in Figure 4.20.

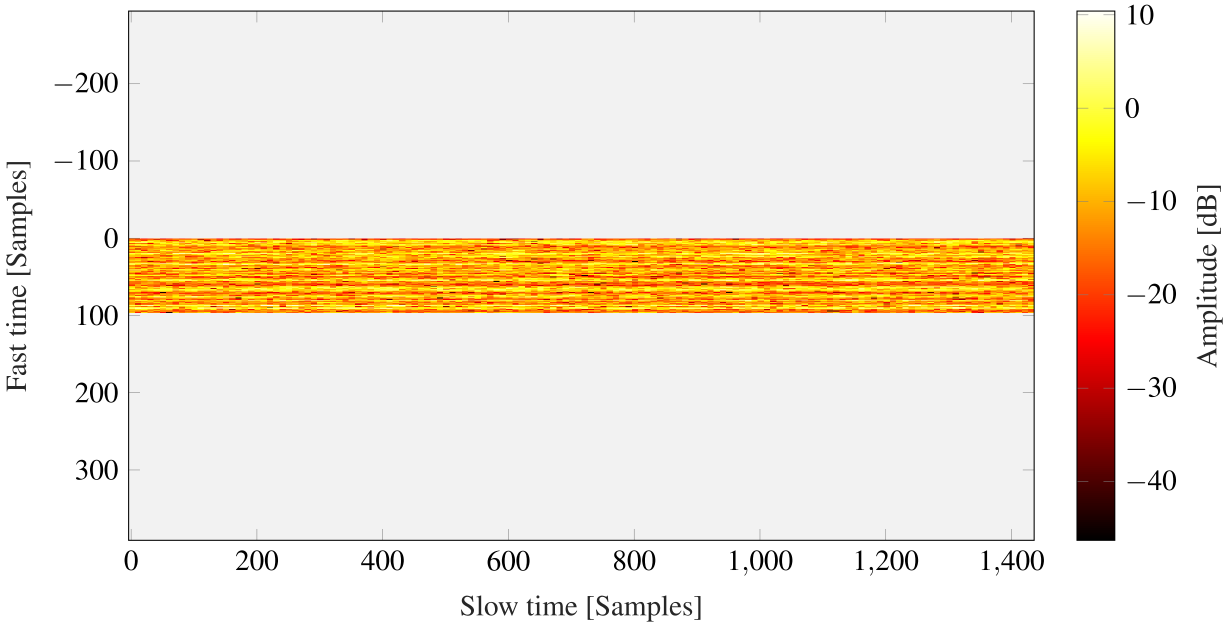


Figure 4.19. TFC15_025 rectangular matrix form illustrated on equal axis system.

In Figure 4.20 the produced Radon transform outputs some interesting points to note. Owing to the irregular input matrix form, strong side-lobes are identifiable. The ”X”-like shape represents these, where only the center of the ”X”-shape is actually of interest. The unwanted legs of the ”X” are caused by the length of the long side of the input matrix. The transform effectively rotates this edge during integration towards the y-axis. Even though only a small value is returned after integration, the center region of interest is also distorted by this effect.

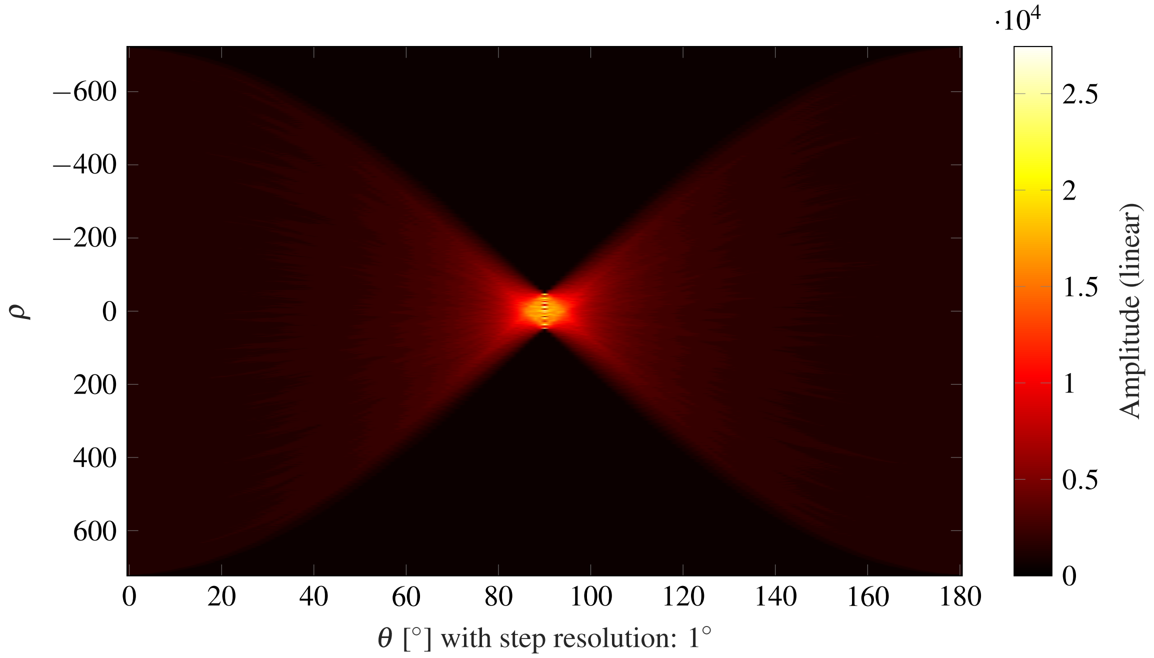


Figure 4.20. TFC15_025 Radon transform of rectangular data matrix $\theta_{step} = 1°$.

Due to the larger input matrix (see Figure 4.19), the output resolution is more sensitive to the size of θ_{step} as the y component of the input matrix is longer ($\tan(\theta) = \frac{y}{x}$). For the desired region of interest, a certain resolution is required to determine enough detail so that crests and troughs can be determined. This is directly related to the processing power required to determine the transform. Therefore an optimal θ_{step} size needs to be considered to transform the data with the required resolution.

Figure 4.21 shows an example where the transform is processed with a very low resolution (θ_{step} value large). The result cannot separate individual troughs and crests sufficiently. This is shown in contrast to the same input data matrix processed with a higher resolution (θ_{step} value small) as shown in Figure 4.22.

Figure 4.22 shows the data transformed with a smaller θ_{step} size at the cost of increased required processing power. With a step size larger than $\theta_{step} = 0.1°$, the individual wave and crest points tend to be difficult to differentiate. The required step size is also a function of the periodic frequency of the approaching waves, consequently this should also be considered.

In order to determine a good compromise between required resolution and processing power, relative resolutions were plotted against the time required to process the data above. This is shown in Figure

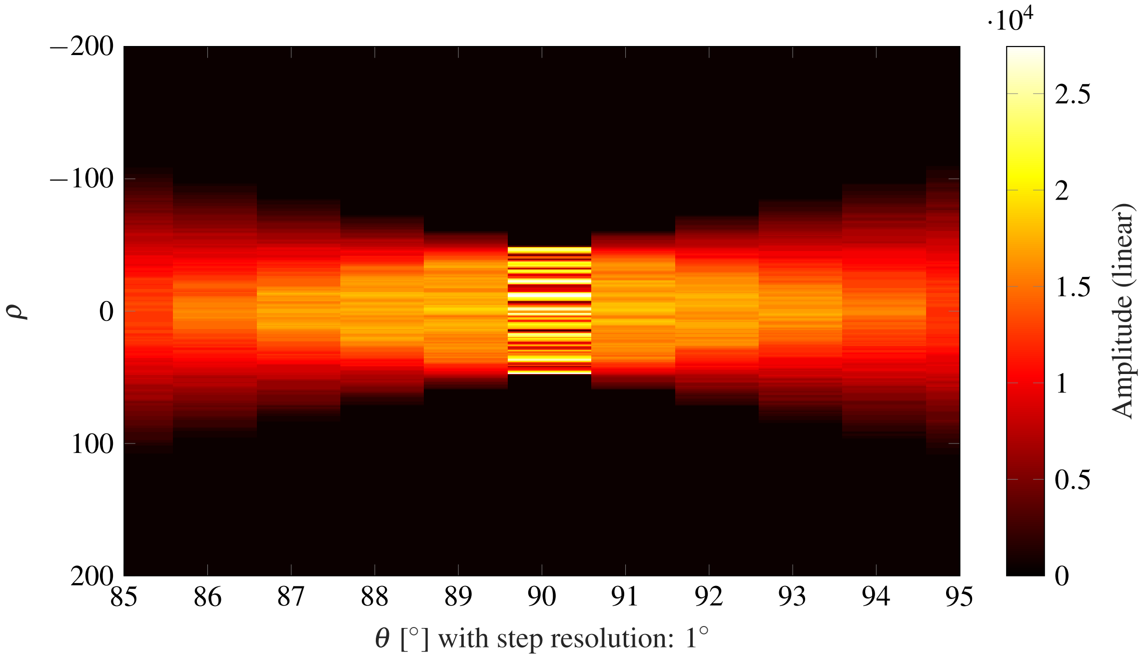


Figure 4.21. TFC15_025 resolution illustration with $\theta_{step} = 1°$.

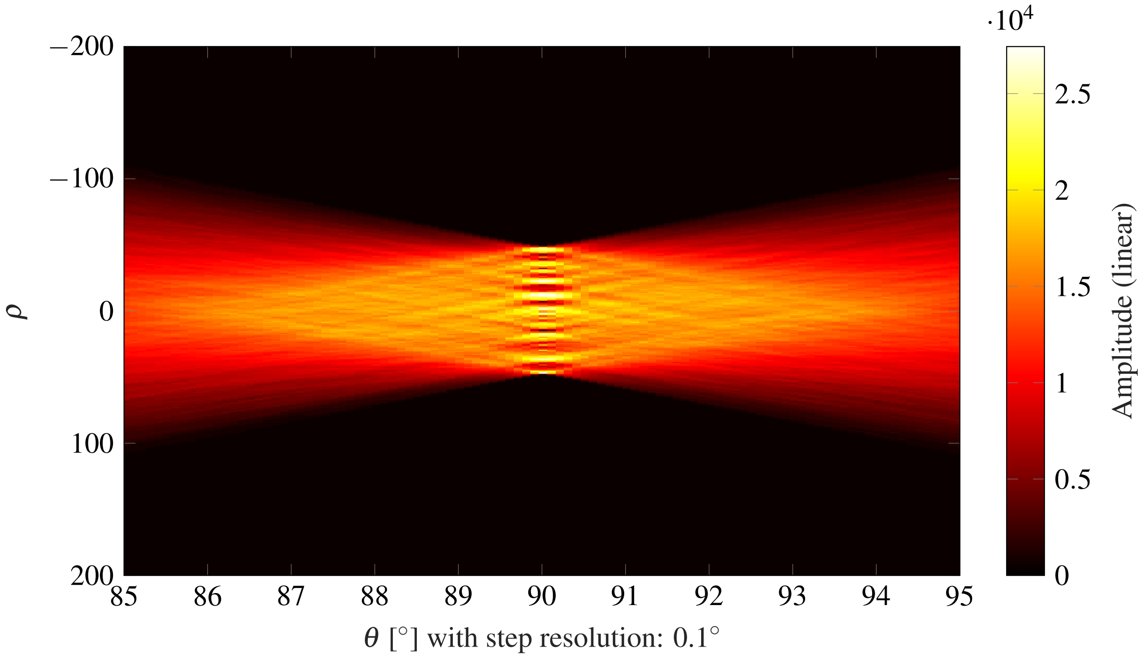


Figure 4.22. TFC15_025 resolution illustration with $\theta_{step} = 0.1°$.

4.23. The result indicates that an acceptable resolution is found at a step size of $\theta_{step} = 0.1°$, as indicated by the red marker in Figure 4.23.

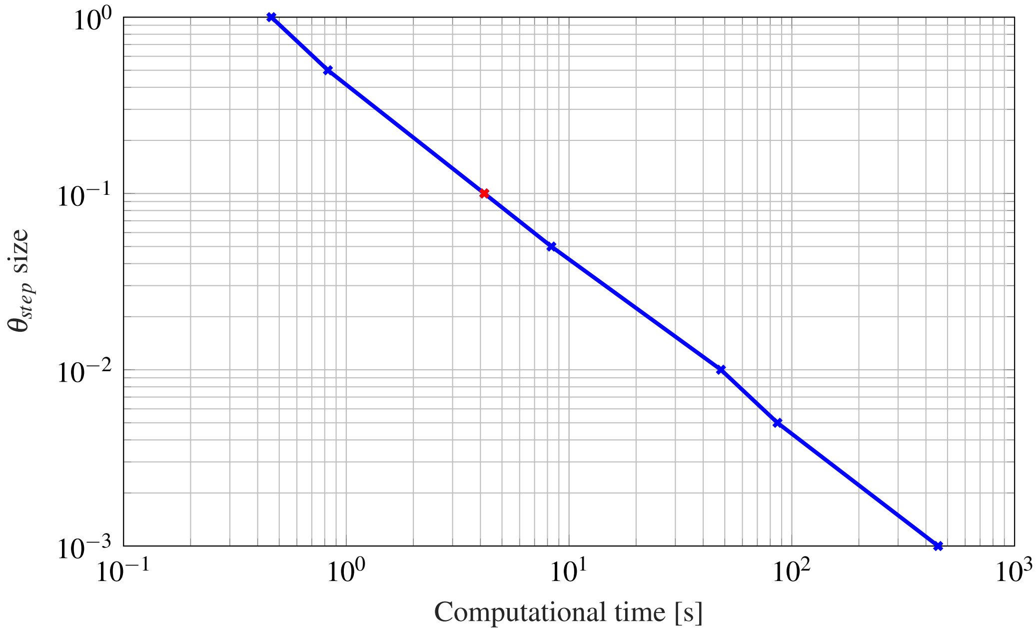


Figure 4.23. Resolution given by θ_{step} size vs computational time used to process a 1440 × 96 element matrix, with the red x marking the chosen step size which delivers acceptable performance.

4.4.2 Circular based Radon transform

A critical step in the presented algorithm is to use the FIR filter as described, to increase the fast time (y component) resolution of the matrix. Applying this filter, the data matrix is changed from an irregular shape with 1440 × 96 elements to square shape with 1440 × 1440 elements.

By ensuring that the input data is a square matrix, and then applying a circular mask to the input data, an improved processing of the transform can be achieved. Figure 4.24 illustrates the circular masking of the square Radon input matrix. In comparison to Figure 4.19, the dimensions of the matrix are now given to be square with 1440 × 1440 elements.

Further, owing to the circular input data where the outer mask areas are mapped to null-values, the integration process adds no additional components originating from the outer corners of the input matrix. This essentially normalises the output of the transform around the center point of the data and ensures that the integration length is the same for all values of the Radon transform output. These results show clearer output data, as illustrated in Figure 4.25, in contrast to the previous result.

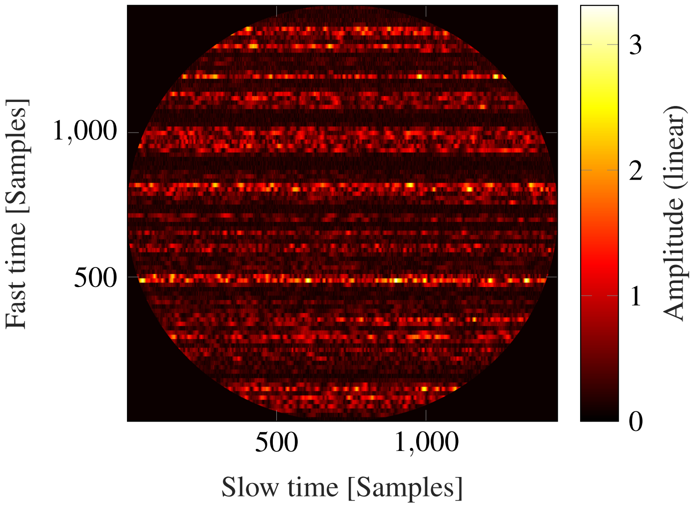


Figure 4.24. TFC15_025 Circular masking of square Radon input matrix.

Although independently developed for this dissertation, the filtering of the data so that the corner areas are discarded in the integral has also been described by [51] and additionally presented in [4].

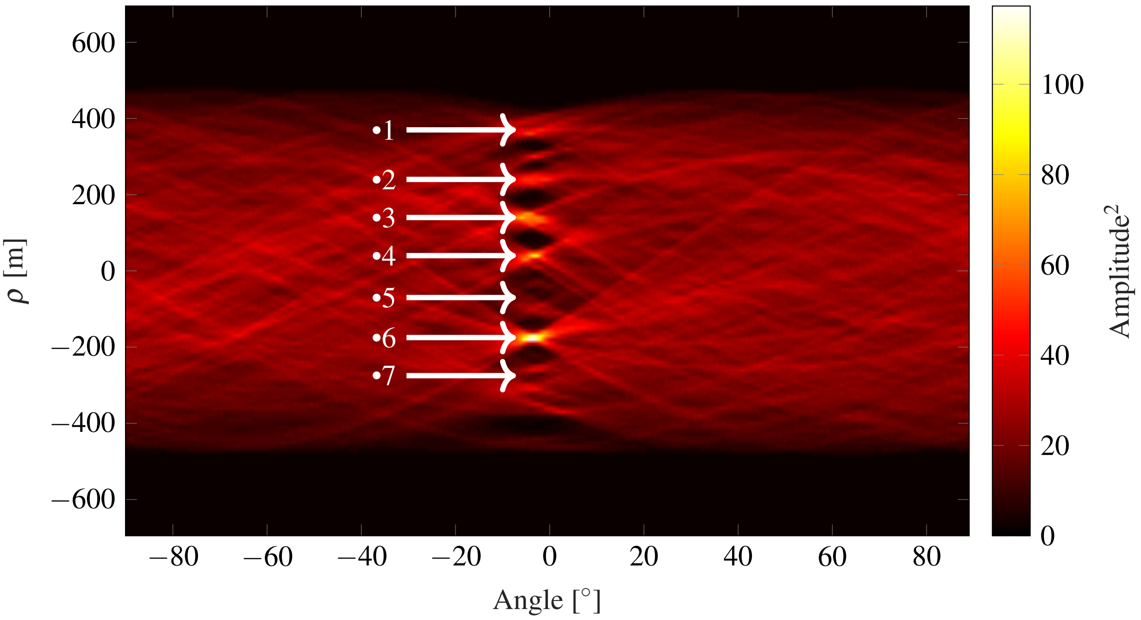


Figure 4.25. TFC15_025 Radon transform of circular input matrix highlighting regions of interest, where the numbered arrows indicate ocean wave crests.

In Figure 4.25, the numbered arrows indicate the ocean wave crests of the transformed data. The illustrated case shows absolute data, i.e. only positive values are shown. However, this can also be applied to negative data making the detection processes more complex, as one then needs to consider both minimum and maximum limits for threshold filtering.

4.5 EDGE DETECTION

In order to automatically identify the angles corresponding to the result of the Radon transform, an automatic edge detection processing block was added. The goal of the block is to locate the center of the Radon islands and return the respective angle from the data needed for further calculations. This can occur autonomously.

4.5.1 Basic principle

Edge detection, where the input is treated as a two dimensional image, locates edges and/or islands by analysing each individual pixel with respect to the pixels around it. Various detection methods were considered and two are discussed below.

In order to isolate the section required, amplitude threshold filtering, which produces binary coded data, is used. Data below the threshold is represented by a 0 and data above the threshold by a 1. This provides the algorithm with an input mask, essentially marking the regions of interest. The implementation of a corresponding threshold value is thus critical as this influences the output significantly. In this dissertation, an acceptable result was analytically found to be a mask that includes all data above the mean value of the entire dataset. This is illustrated for a simple Gaussian distribution in Figure 4.26.

In Figure 4.26, the spectrum is represented by the plotted distribution. From this distribution the mean intensity value is shown by the horizontally dashed line (green). The point where the horizontally dashed line leaves the right outer edge of the spectrum, results in the automated threshold value used for amplitude filtering. This filtering removes all non-desired amplitudes below this value. Subsequently the formula for the threshold value is given to be

$$T_{edge} = \max\left(f(x) > \text{mean}(f(x))\right), \tag{4.9}$$

where $f(x)$ is the returned transform spectrum.

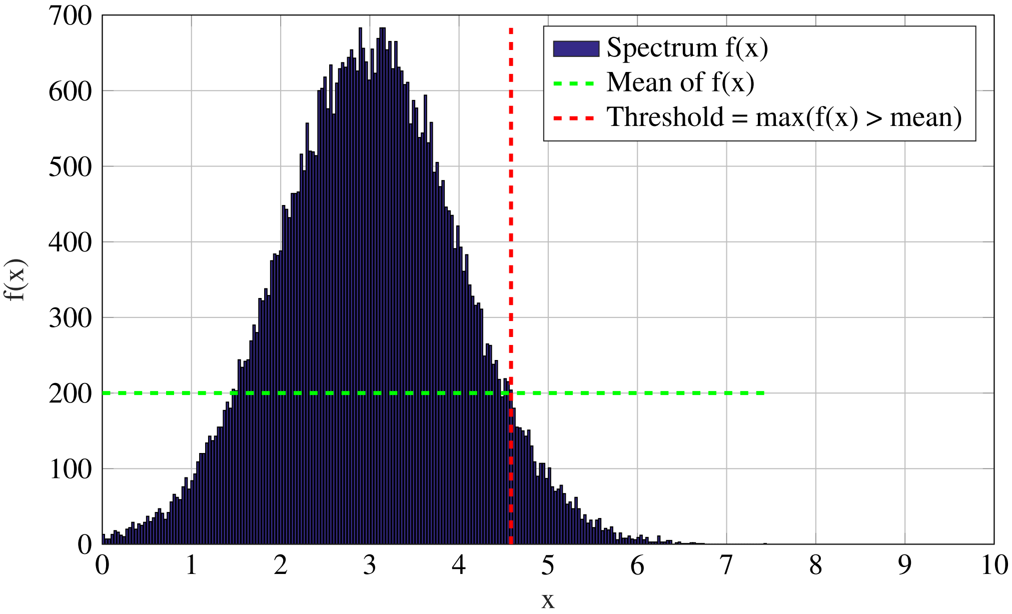


Figure 4.26. Maximum of the mean of the distribution of interest defining the filtering amplitude threshold given by the horizontal dashed line in red.

4.5.2 Canny edge detection

This process applies a gradient evaluation algorithm to a Gaussian based smoothed set of input data. The evaluation algorithm then further tracks the high gradient points and illustrates these as edges. The final edges are then threshold filtered to emphasise the identified sharp transitions. This filter is typically used in image processing applications. The filter is computationally expensive and not ideal for real time applications. Further, the identified edges in this application would still need to be categorised, requiring additional steps. Results illustrating this are shown in Figure 4.28.

4.5.3 Cluster segmentation detection

Blob detection focuses on the identification of actual islands of data. Using the threshold filter value, binary data of the Radon result is first obtained as discussed above. This binary data can then be analysed by a blob filter to successfully identify individual islands or blobs, which represent lines in the original measured data. The filter considers the data as an image and processes it pixel by pixel, effectively searching for connected pixels and then marking and tracking these as individual blobs.

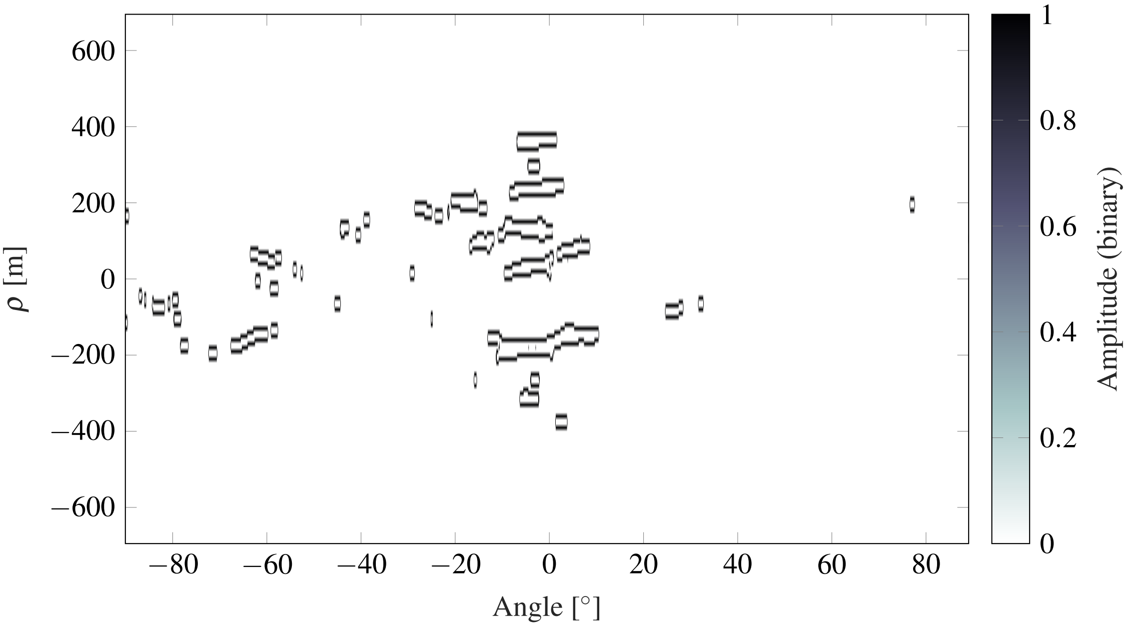


Figure 4.27. TFC15_025 canny edge detection results illustrating the high gradient edges.

Blobs are defined by a connectivity, which is only given if the identified pixel is directly next to the blob. In the case of no neighbouring pixels being set to a 1, a new blob is spawned.

Different techniques can be applied to find the center of a blob. In this dissertation, two primary methods known as centroid and weighted centroid locations are considered. The centroid algorithm calculates the center point to be the mean of the blob's x-axis and y-axis range. Whereas the weighted centroid further weights the centroid location based on the underlying measured data values, in order to create a weighted center value based on signal intensities. For the proposed method the weighted centroid was implemented.

Figure 4.28 illustrates the results the cluster segmentation algorithm produces when applied to the output of the radon transform.

4.6 ANGLE CALCULATION

The angle measured in the reduced matrix size (see Section 4.3), obtained by the Radon transform, needs to be corrected to account for the change in matrix size when it is later applied to the original measured data matrix. This is a simple trigonometric ratio that multiplies into the measured angle

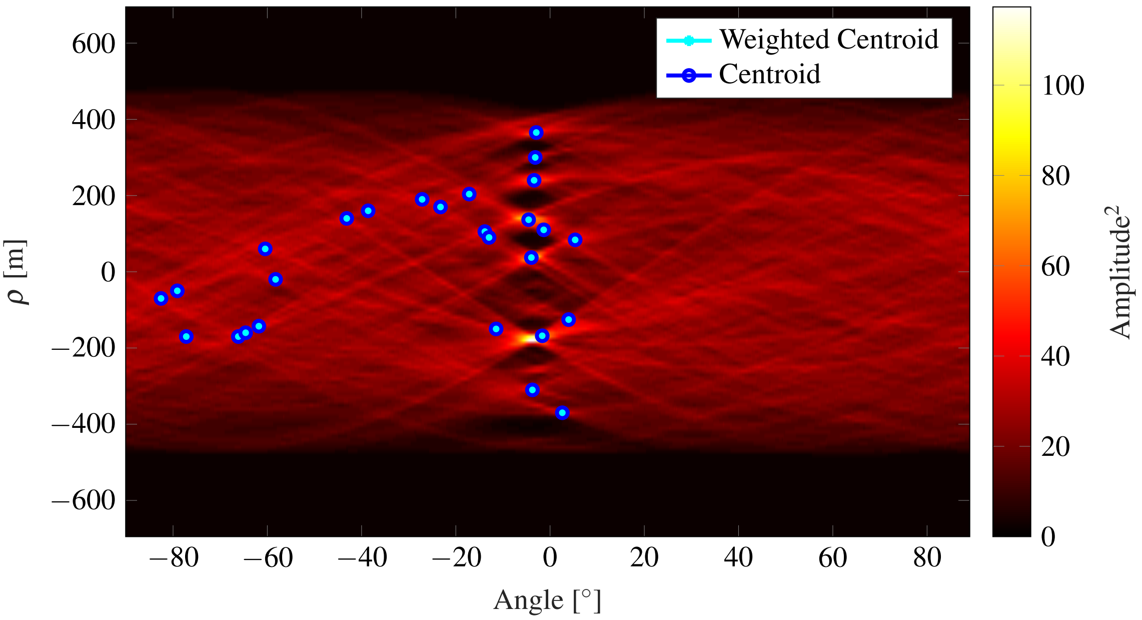


Figure 4.28. TFC15_025 cluster segmentation detection, where the identified clusters of interest are indicated by the blue marked centroids. The weighted centroids represent the center of the island biased by underlying amplitude intensities.

as determined for the square matrix. This step effectively compensates for the change in matrix size and allows the correction of the data to take place on the original measured data matrix. Owing to the difference in matrix size with a factor of 40 plus, the determined angle becomes very sensitive to small changes, with equation (4.10)

$$\theta_{corr} = \arctan\left(\frac{\tan(\theta_{avg} - 90^\circ)}{scale}\right), \tag{4.10}$$

where *scale* represent the multiplication ratio of the sampled data to the interpolated data in the y-axis.

4.7 WAVE PHASE ALIGNMENT

The alignment of the data towards the x-axis is implemented in the penultimate processing block of the first data-chain as given in Figure 4.1. Based on the computed angle from the previous step, the original full dataset is shifted, such that all x-units are moved y-units upwards/downwards (black arrows in Figure 4.29) on a linear scale within the matrix. This ratio is derived from the correction angle given by equation (4.10). This then produces the desired output of an ocean wave phase angular corrected data stream as illustrated in Figure 4.29. The number of y-units changed per x-unit count is given by equation (4.11)

$$C_{tD} = \frac{\tan(\theta_{corr}) \cdot t_s}{2}, \tag{4.11}$$

where C_{tD} represents the columns counted to drop/rise the data in the y-axis, t_s is the length of the data over slow time for which the correction is to be applied, and divide by 2 accounts for the rounding of the values to integers as needed by the algorithm. For given data length C_{tD} is to be equally dropped or added, depending on the sign of the angle from equation (4.10).

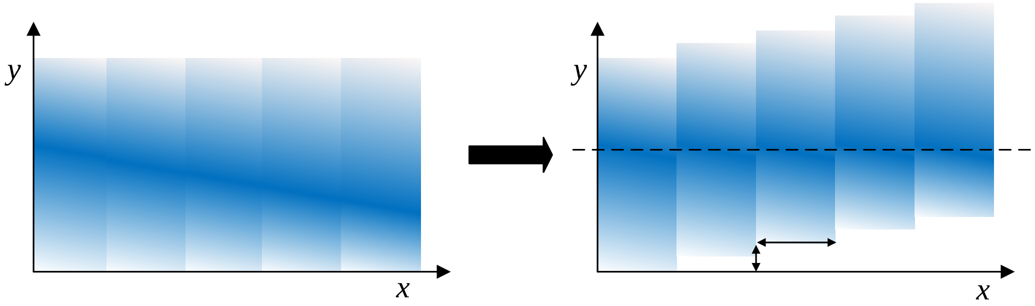


Figure 4.29. Illustration of the angle correction method, showing the drop/rise in the y-axis over time, where the measured amplitude of an ocean wave crest is represented by the intensity of the blue within the matrix.

4.8 STATISTICS

The analysis section of the processing chain as defined by Figure 4.2 relies on the properties and methods described in Section 3.2. Once the input matrix is aligned to the x-axis, the data is analysed based on the following two properties:

- probability distribution of the data, and
- data correlation.

4.8.1 Probability distributions

One can consider and analyse individual sub-sections of the wave, as shown by the indicated regions in Figure 4.30. The region between the oceanwave crests and troughs are split into equally spaced segments. These segments are then analysed individually.

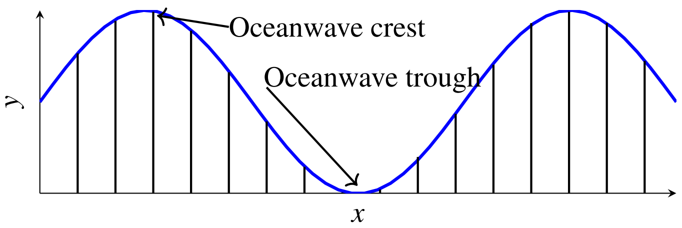


Figure 4.30. Ocean wave structural definitions illustrating the crest and trough regions of an ocean wave together with subdivisions which are statistically analysed over slow time for multiple samples.

Data is analysed and fitted against Rayleigh and K-distributions as previously described in Section 3.2. Individual sections of the wave can be considered, as the data is aligned to the x-axis. This allows insight into the measured returns based on the input section and type of wave. This approach is novel to the investigation of breaking waves in littoral environments and allows statistically significant analysis of a sub-section of a breaking wave using a high sample count of radar returns.

In the next section the following two distributions are considered with respect to correlation properties:

- Rayleigh distribution, and
- K-distribution.

The respective definitions for the above distributions were provided in Section 3.2.

4.8.2 Correlation

Correlation relates to the statistical relationship between the measured random variables. It provides insights into possible variable dependencies. In particular, the wave-based structure in itself presents a strong period and is consequently a correlated structure. Further analysis was conducted to determine if any additional relationships in range could be determined.

The analysis is performed by applying the discrete autocorrelation function, given by equation (4.12) for each individual slow time interval.

$$c_\tau = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n-\tau), \tag{4.12}$$

where $x(n)$ is the discrete signal at time n and τ the lag/lead time.

The generation method, of correlated data with a given distribution, varies for the given distribution.

4.8.2.1 Rayleigh distribution based correlated data

Two independent Gaussian random number generators are combined to generate random Gaussian complex numbers. The generated set of random values is then convolved with the filter $H(t)$. $H(t)$ is obtained by taking the inverse FFT of the square root of the power spectral density (PSD) for a given set of input data. The algorithm, presented in [52], is based on the originally proposed method by Smith in [53] for Rayleigh fading of a signal. Figure 4.31 illustrates a structural implementation of this.

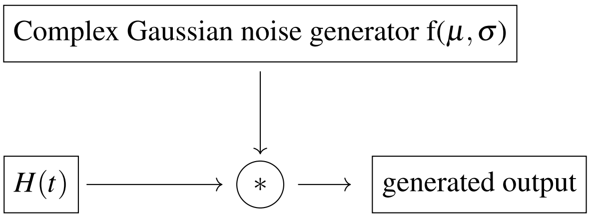


Figure 4.31. Generation of correlated Rayleigh data using filter $H(t)$ and complex Gaussian noise generator.

4.8.2.2 K-distribution based correlated data

In [54], Brotse describes a methodology for the generation of an arbitrary distribution with a given correlation. The process pre-distorts the correlation of the input distribution. This is done so that the desired distribution, with the desired correlations, can be obtained through the post distribution filter, which also alters the correlation. The process consists of a linear filter followed by a non-linear filter. The data is generated through three primary steps:

- determination of the non-linear zero memory filter,
- determination of the autocorrelation of the correlated Gaussian sequence, and

- determination of the linear memory filter.

First a Gaussian input distribution with narrow correlation is transformed to a uniform distribution through [54, equation (2)]. This corrupts the autocorrelation of the input data, whilst easily allowing the uniform distribution to be transformed to a desired output PDF. Thus it is required to pre-distort the correlation of the input data in such a way that the output generates the desired sequence in terms of distribution and correlation.

Due to the complexity of determining the pre-distortion coefficients, required by [54, equation (5)] and then [54, equation (11)], the generation of correlated K-distributed samples is considered out of scope for this document.

CHAPTER 5 RESULTS AND DISCUSSION

The results of the individual processing steps of the method described in Section 4.2 are presented in this Chapter. First the utilised datasets are described in detail in Section 5.1, followed by their evaluation with respect to the individual algorithm steps in Sections 5.2 to 5.6. Finally, the data generation process is presented together with the obtained results in Section 5.7. The chapter is concluded with Section 5.8 which remarks on datasets containing additional boat reflections and their potential impact on the algorithm.

5.1 RAW DATASETS

Low grazing angle, fixed frequency, pulsed radar, X-band sea clutter data was obtained from the Council for Scientific and Industrial Research's (CSIR) online database [11]. Six datasets were considered in detail. Three of these contain purely periodic sea clutter returns and three have the addition of a RIB crossing the periodic ocean waves. Only the first, non-boat based dataset (see Figure 5.1) is presented and discussed in this chapter. Appendix A contains the results of the remaining five datasets. Reference to these datasets is made at the respective section during the discussion.

The raw measured data from the radar is shown in Figure 5.1. The data reflects returns for a duration of 13 seconds, with 96 range bins, each with a range resolution of 15 m. For this dataset the radar uses a fixed frequency transmission of 6.9 GHz and a PRF of 5 kHz. The grazing angle for these measurements is given to be $\theta_g = [0.853^\circ, 1.27^\circ]$. Further details regarding the radar and this measurement dataset is obtainable from the CSIR database for the measurement CFA16_002 [11].An overview of the six datasets is provided in table 5.1.

For target data, measurements including the boat location information were also captured using a Global Positioning System (GPS) data logger onboard the RIB. This information was referenced to the datasets so that the travelled route could be overlaid. See [11] for additional details regarding the measurement database. For overview plots of the remaining five datasets in raw format see Figures A.1, A.16, A.31, A.46 and A.61.

Table 5.1. Dataset overview

Dataset	**Type**	**TX Freq.**	**PRF**	θ_g
CFA16_002	sea clutter	6.9 GHz	5.0 kHz	$0.853° - 1.27°$
CFA16_003	sea clutter	6.9 GHz	5.0 kHz	$0.427° - 0.525°$
CFA16_004	sea clutter	6.9 GHz	5.0 kHz	$0.427° - 0.525°$
TFC15_025	sea clutter + RIB	9.0 GHz	2.5 kHz	$0.628° - 0.838°$
TFC15_028	sea clutter + RIB	9.0 GHz	2.5 kHz	$0.492° - 0.619°$
TFC15_029	sea clutter + RIB	9.0 GHz	5.0 kHz	$0.445° - 0.551°$

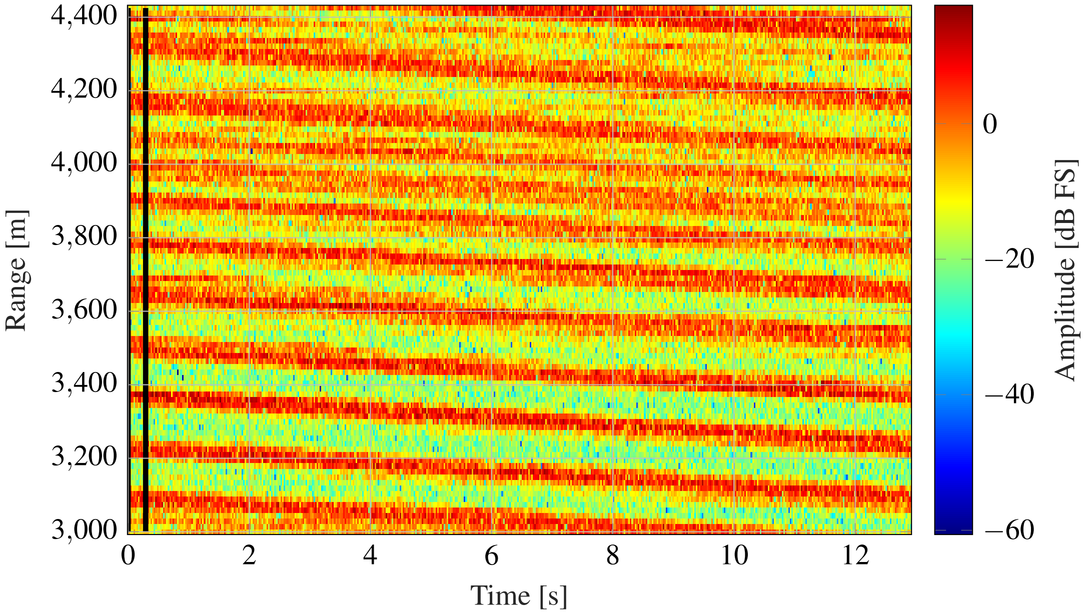


Figure 5.1. CFA16_002 range time plot of 13 s of measured data given in dB full scale (FS), where the black vertical lines indicate the region used for angle determination.

In order to determine and analyse the approaching periodic breaking wave angle, the first segment of the data, given by ≈ 0.5 seconds, is considered first. This region is indicated in Figure 5.1 via two vertical black lines marking the analysed region. These measurements are re-plotted separately in Figure 5.2, where the extreme angle of the approaching crests is not directly evident but still contained within the data. The data now represents a matrix of 96 range bins by 1440 time samples. The length of 1440 time samples was chosen, so that the interpolation later increases the range resolution by a factor of 15 (96 × 15 = 1440), thus essentially representing 1 m range resolution.

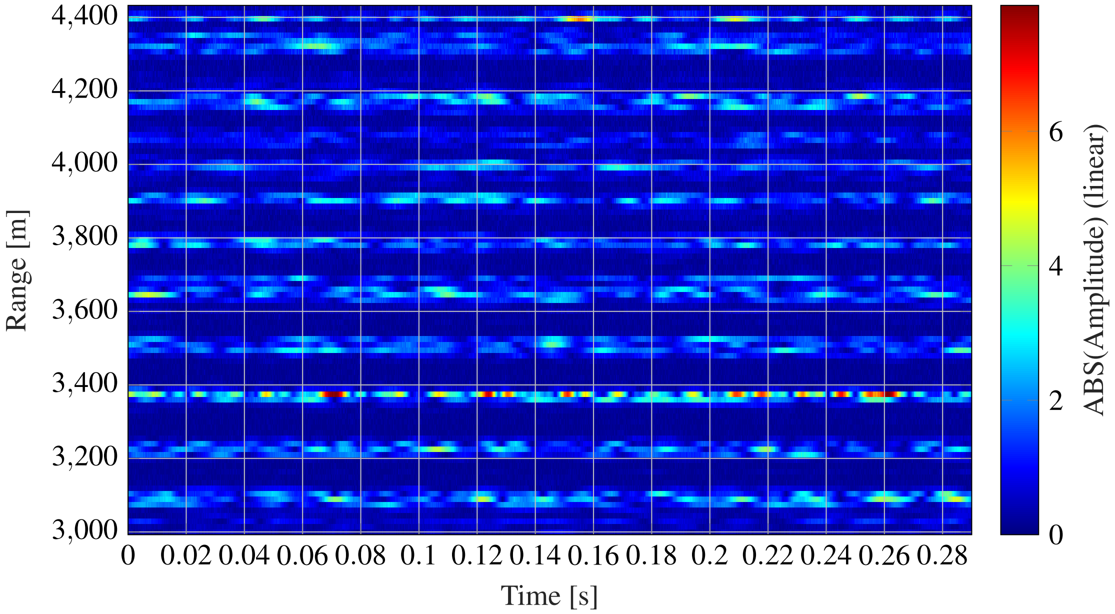
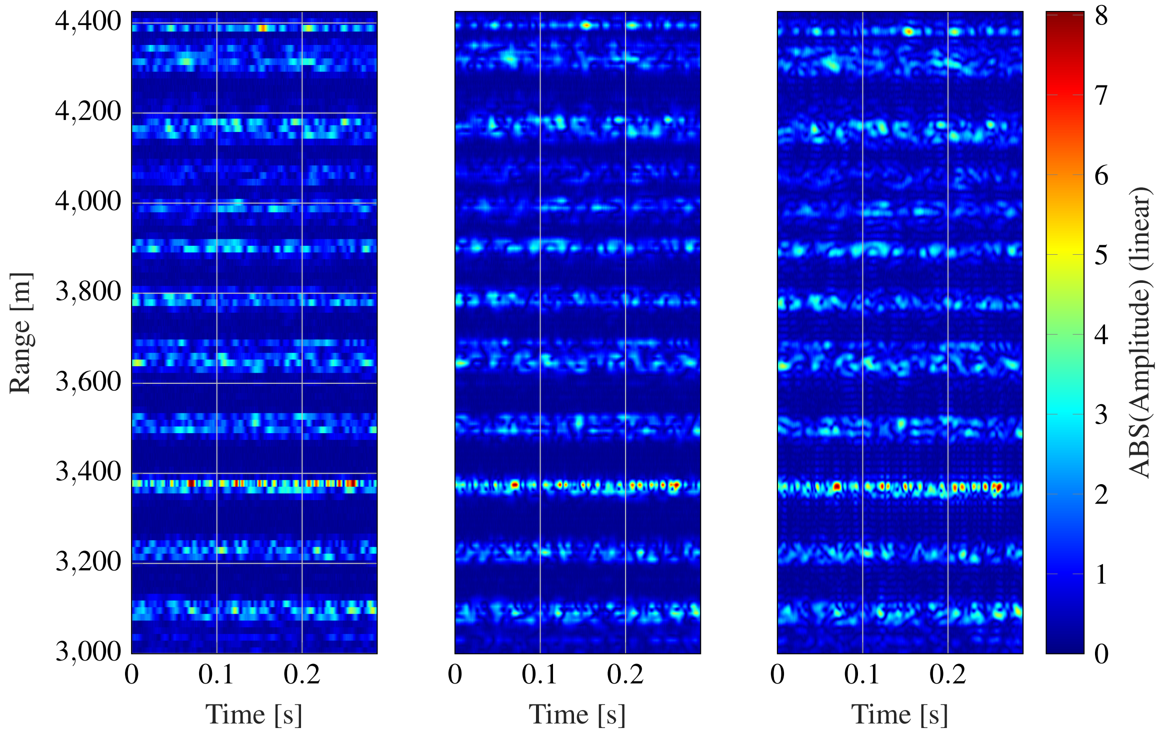


Figure 5.2. CFA16_002 range time plot of analysed section of data given by the time interval of [0 0.288] seconds.

5.2 DATA UPSAMPLING

The methods as described in Chapter 4.3, are now applied to the given data and compared. Figure 5.3 shows the results for the same input data across all three filters. The different frequency responses are identifiable in this figure. Figure 5.3(a) shows higher frequency components compared to the other two responses. Whereas Figure 5.3(c) shows smoother transitions containing fewer high frequency components.

Based on the discussion in the previous chapter with regard to the frequency response of the three types of filters and the results obtained from Figure 4.17, the next steps involved processing all data using only the FIR interpolation results. This output for the given dataset CFA16_002 is provided in Figure 5.4.



(a) **(b)** **(c)**

Figure 5.3. CFA16_002 post upscale filter output results based on filter types (a) zero order hold, (b) first order hold and (c) FIR based.

The output from Figure 5.4 illustrates the analysis section of the measurement data up-sampled using the FIR based interpolation filter. The filter interpolates 15 points between every input point as mentioned previously. The increased resolution output is then used to allow further processing. Structural patterns of individual wave crests are identifiable. Worth noting, as stated in Figure 4.1, is that the only goal of this step is to determine the angle with which to correct the original measured data so that the crests are parallel to the x-axis. Further processing steps will all take place on the original full dataset, so as not to influence the distributional properties through filtering.

After interpolation, a desired square matrix with 1440 range bins over 1440 time samples is obtained, thus making a square input matrix available. This data is illustrated in Figure 5.5(a).

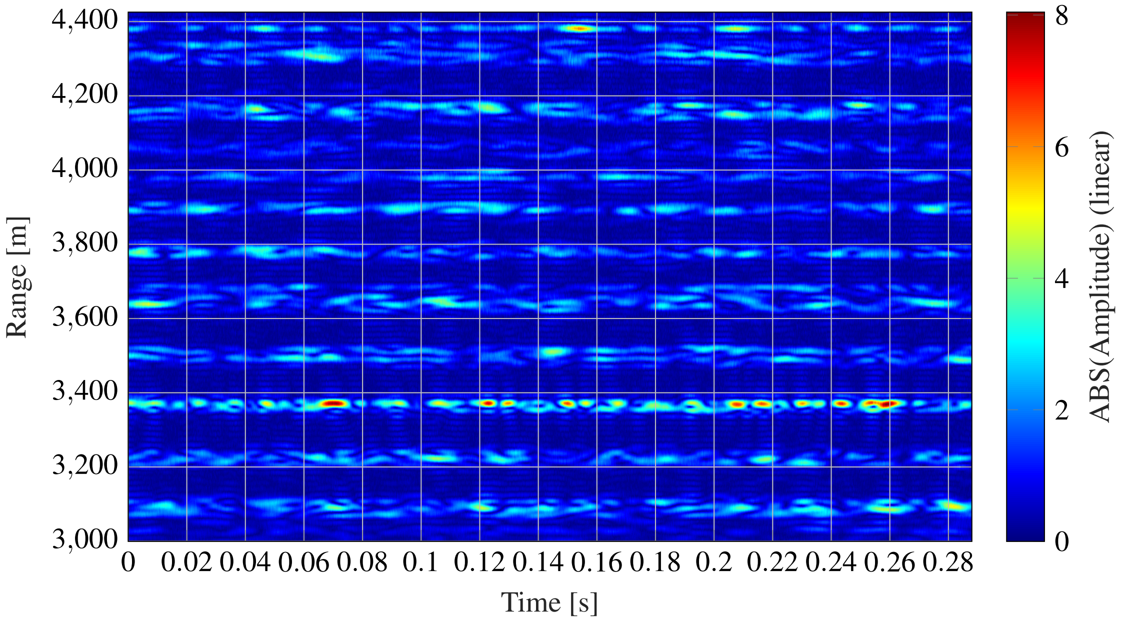


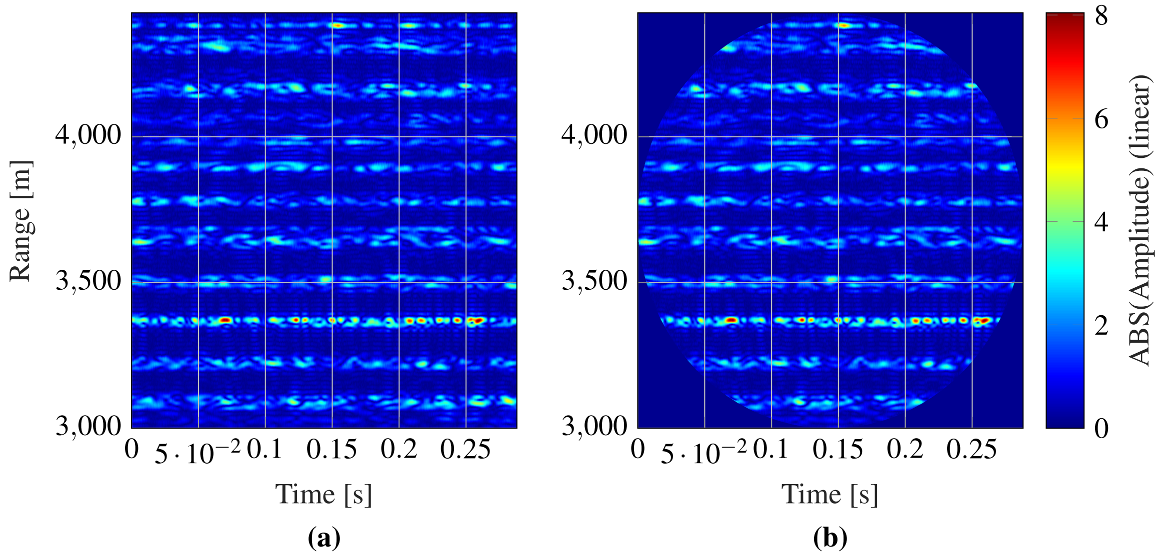
Figure 5.4. CFA16_002 up-sampled data using second order filter.

5.3 RADON TRANSFORM AND PEAK DETECTION

In order to avoid the effects of longer integration paths along diagonal directions of the image, as described in Section 4.4, the square matrix is then overlaid with a circular mask. Thus all data points outside the circle are designated as null values and do not add to the integral. This is equivalent to removing the outer measurement data and essentially normalising the integration length for the Radon transform over all angles θ and position offsets ρ. The mask applied prior to the transformation step is shown in Figure 5.5(b).

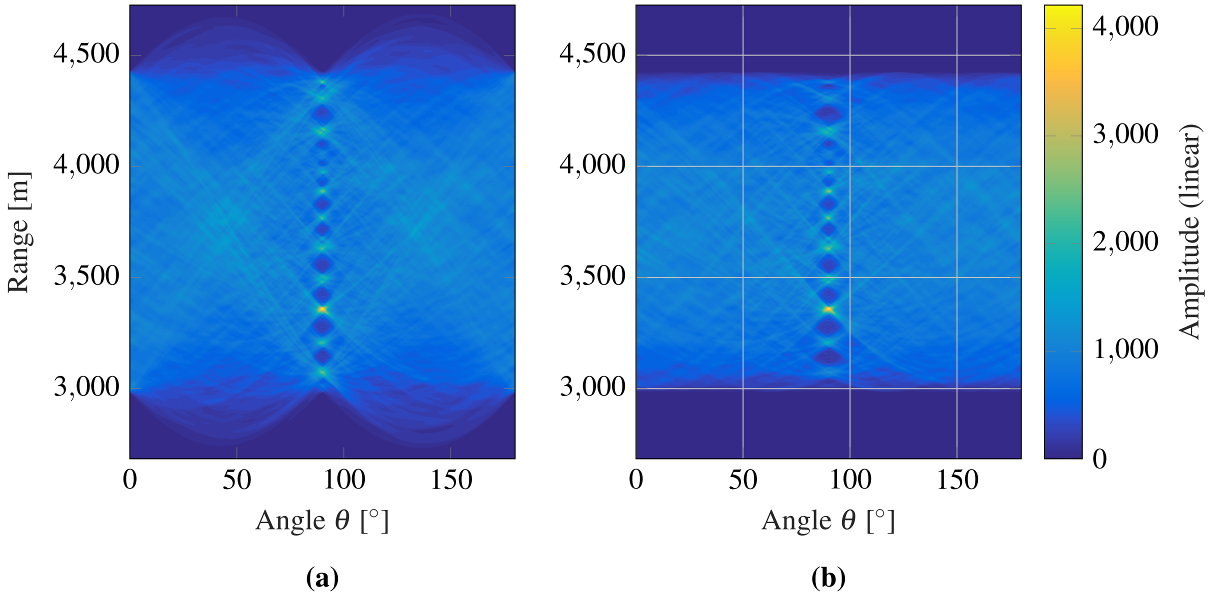
In order to illustrate the difference and necessity of the circular processing, the two inputs as given in Figure 5.5 were processed through a Radon transform and their results are shown in Figure 5.6. As can be seen in Figure 5.6(a), side-lobes are clearly reflected through the Radon transform onto the output data when the input is non-circular. This effect is due to the difference in integration length of the number of samples. Whereas, for the circular input data, the integration length is equal for all angles θ and offsets ρ, thus the result is normalised and the output improved.

From the circular data, a pattern for the periodic wave structure can already be seen. The offset from 90° on the x-axis determines the angle at which the wavefronts in the data are approaching over time.



(a) **(b)**

Figure 5.5. CFA16_002 non-circular vs circular input processing matrices for Radon transform, indicating the null data points on the outside of the circular mask.



(a) **(b)**

Figure 5.6. CFA16_002 Radon transform result illustrating the difference in output result for (a) non-circular and (b) circular input matrix.

A simple threshold filter is implemented to isolate the effective wave crests and troughs. In order to determine an effective processing threshold for the binary filter, the measured data distribution was analysed and the upper percentage was chosen as a filtering factor. Thus, the return data is limited to the upper returns based on intensity of the transform and corresponds to the strongest periodic wave

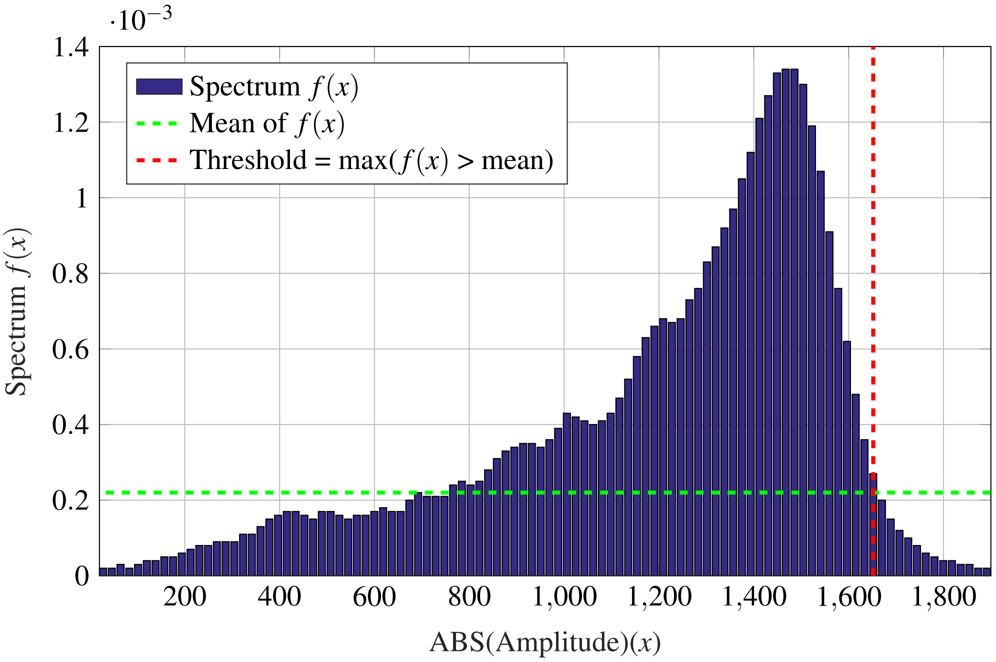


Figure 5.7. CFA16_002 threshold filtering on processed data showing the Radon spectrum $f(x)$ and respectively the automatically determined threshold T_{edge}.

structures in the measured data. This is shown in Figure 5.7 and after filtering in Figure 5.8. The threshold is determined by the equation (4.9).

The black data in Figure 5.8 is what remains after simple threshold filtering of the output of the Radon transform. This data is then subjected to a blob detection algorithm, where the centroids of the islands are located and their respective area is calculated. The algorithm for the detection of the islands is sufficient enough to detect the required information needed to determine the angle so that no further processing, such as shape classification and or filtering, needs to be implemented. In order to only include the larger intensity responses, the returned data objects are filtered based on their individual intensity standard deviations, which result from their underlying measurement data. This allows the returns below the 6σ limit to be excluded from further consideration.

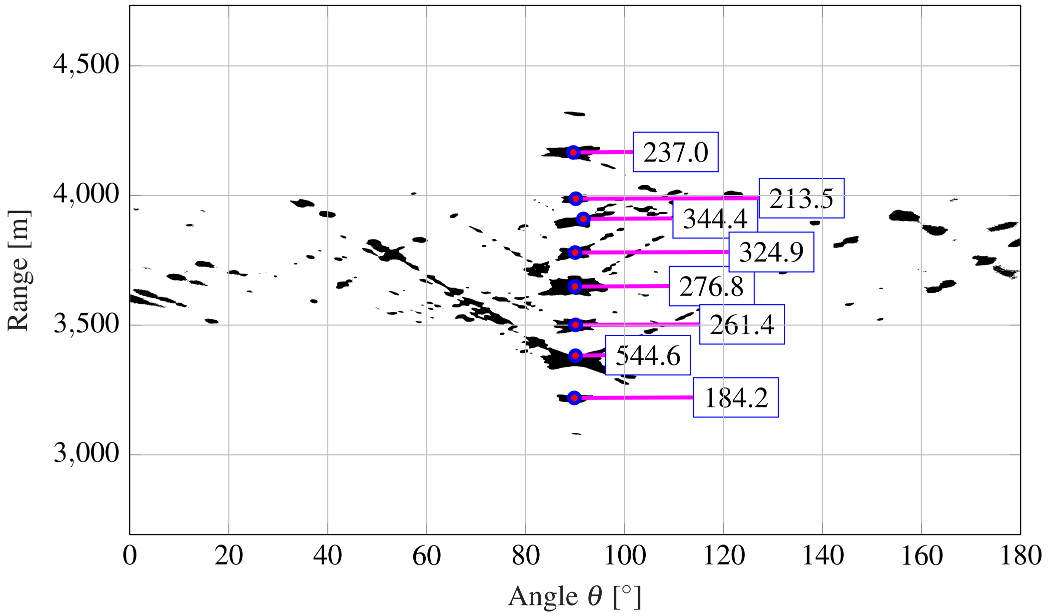


Figure 5.8. CFA16_002 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

5.4 ANGLE CALCULATION

The angle, with regards to the x-axis, is determined using the centroid locations. The average angle θ_{avg} of the detected islands equates to the angle offset for the square matrix. The region of interest is determined by examining the average summation as defined by equation (5.1). The summation plot of the transform determines the region of interest and is indicated by the red region in Figure 5.9. This parameter may optionally be used to restrict the angle response band obtained from the Radon transform, preventing outliers from obscuring the correction angle by being included in the angle average.

$$I_\theta = \sum_{i=0}^{\rho_{max}} f(\rho_i, \theta), \tag{5.1}$$

Taking the individually determined angles, an improved estimate can be obtained by using the median value instead of the average value. The median is advantageous as it accounts for a non-Gaussian distribution of angles, which due to the low number of samples and high sensitivity is beneficial as singular outliers may occur through the dynamics of measured data. This offset would then be

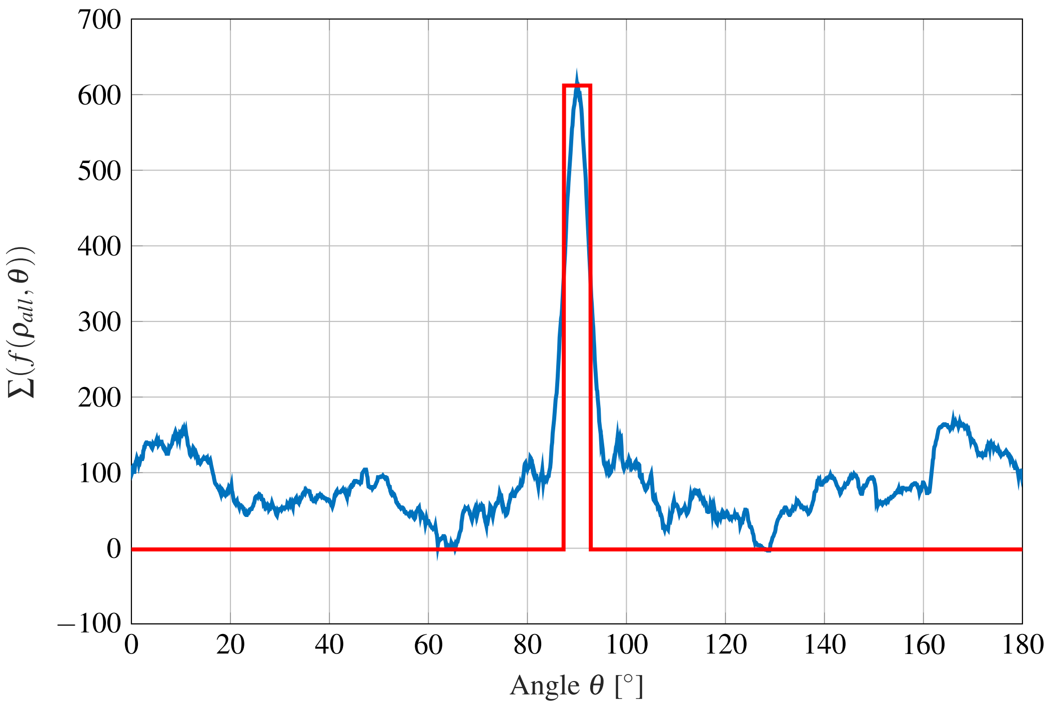


Figure 5.9. CFA16_002 peak detection region of interest and threshold filtering.

contained in the output and influence the angular result. This improvement is illustrated in Figure 5.10

Verification of these angles over the input data is achieved by using the inverse Radon transform to generate lines based on input points given by an angle θ_{in} and position ρ_{in} values. These results are overlaid on the original data and shown in Figure 5.11.

As can be seen from Figure 5.11, the two outer wave crest at a range of 3100 m and 4300 m were not determined. Although not detected, it was evaluated as non-critical since further processing in the chain does not rely on all individual wave crests being present, but only on the collective median angle being correct. As stated only the corrective angle is determined from this data.

A summation of the individual range bins over time is shown in Figure 5.12. The comparison illustrates the difference between the original data vs that of the processed data generated by the inverse Radon transform. It clearly shows the wave crest identification and isolation as acceptable. As indicated by Figure 5.12(d), the signal to noise isolation for the individual wave crests is clearly shown in Figure 5.13, which illustrates the regions with the determined wave crests. The amplitude of the wave crest at

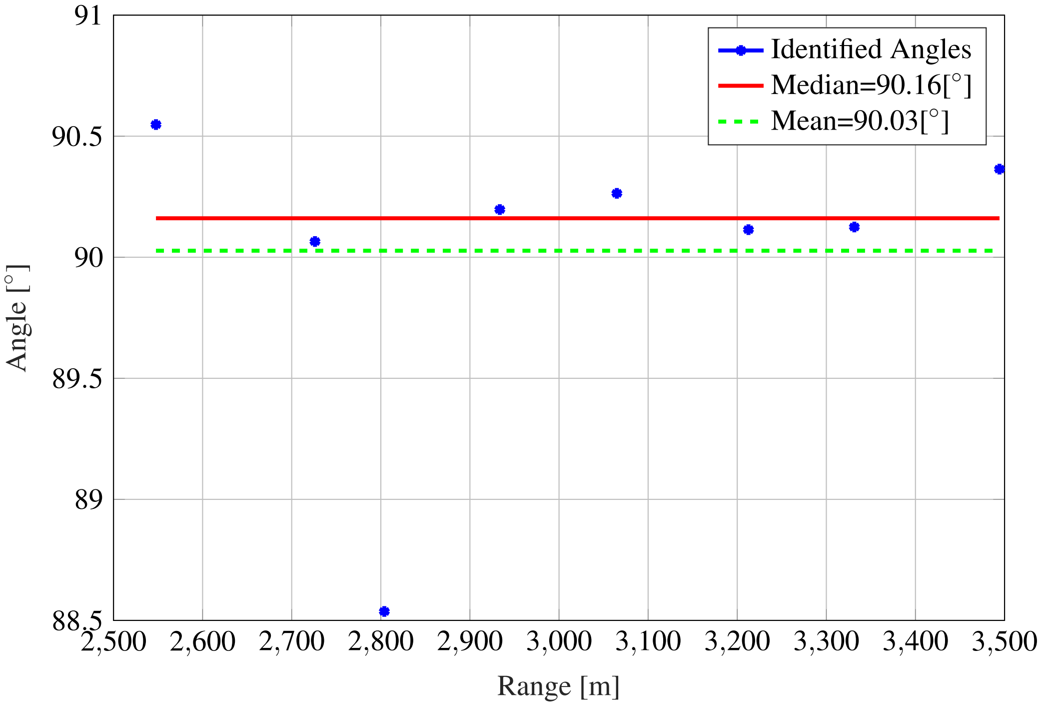


Figure 5.10. CFA16_002 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

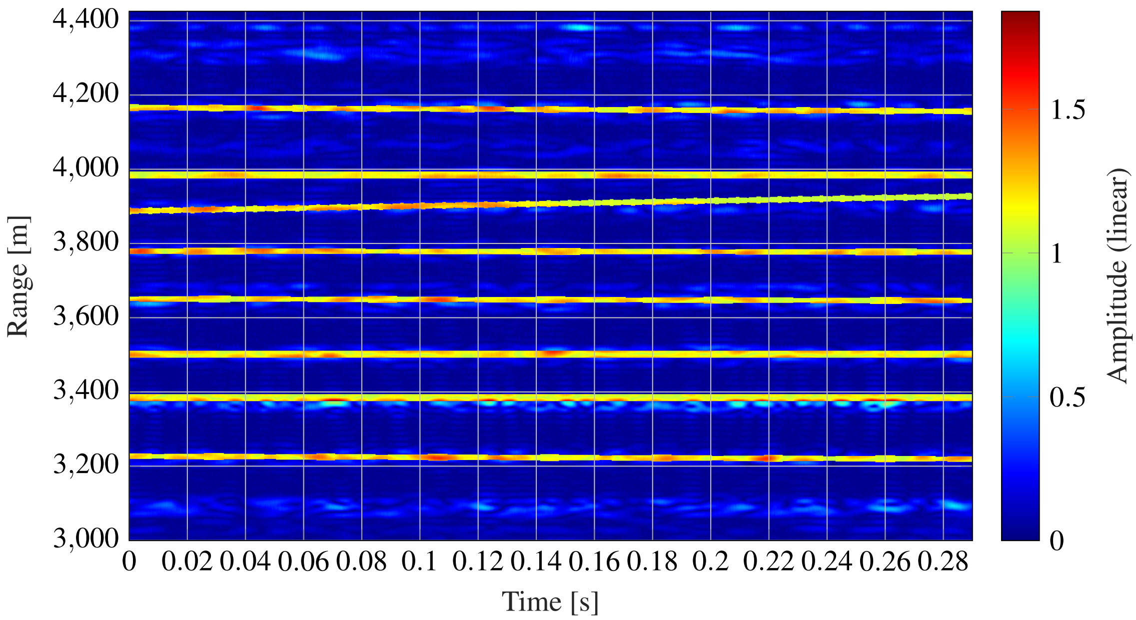


Figure 5.11. CFA16_002 overlaid generated data lines (yellow) originating from the inverse Radon transform illustrating good alignment and visually verifying the detected angles.

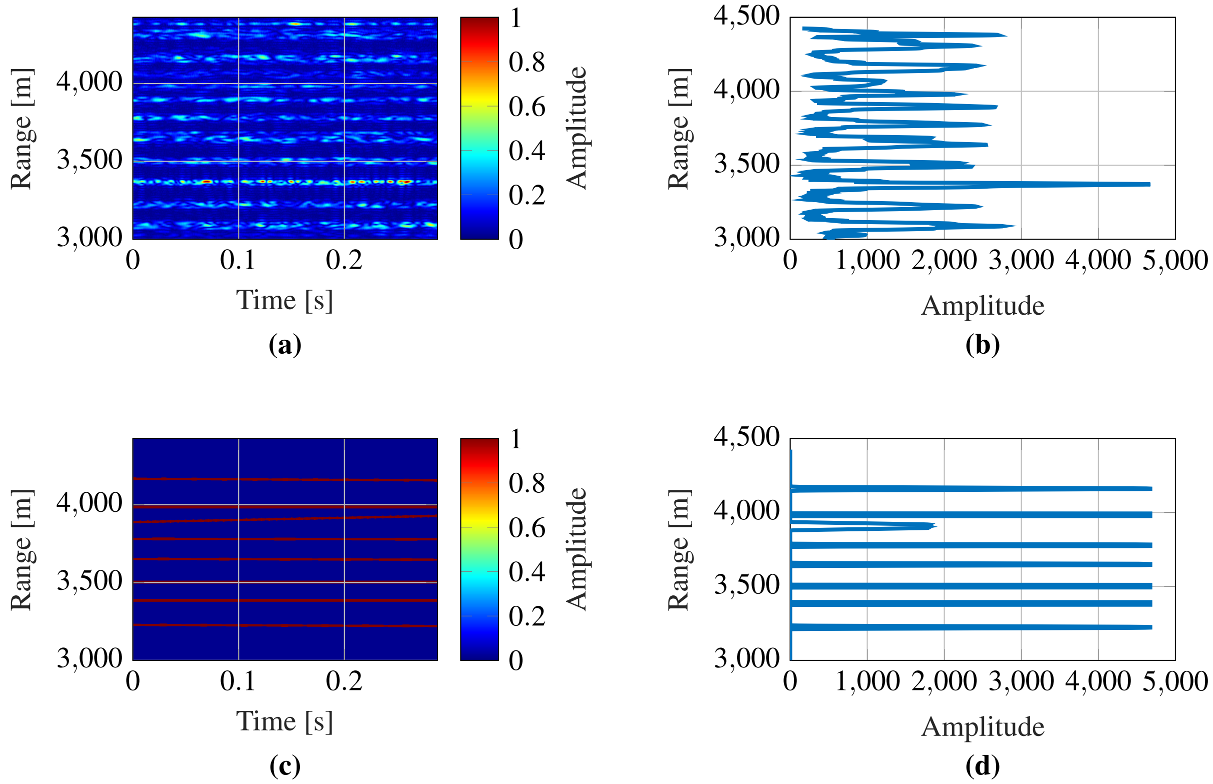


Figure 5.12. CFA16_002 Radon transform results showing (a) original data, (b) the summation of the original data in horizontal direction, (c) the overlaid generated line data from the inverse Radon transform, and (d) the summation of the generated inverse Radon transform line data in horizontal direction.

3900 m is lower due to the Radon transform identifying the incorrect angle. This effectively reduces the total sum of a given range bin as the angle crosses the other wave crests.

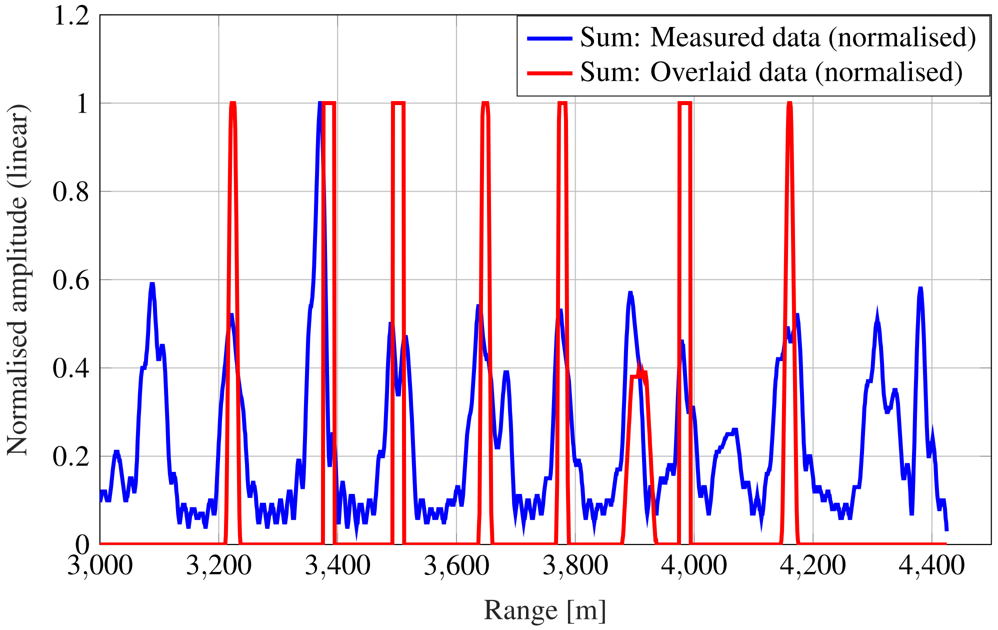


Figure 5.13. CFA16_002 measured vs generated summation data over range, normalised to 1 and illustrating alignment and detection.

5.5 WAVE PHASE ALIGNMENT

From the angles determined in Figure 5.10, for this case $\theta_{avg} = 90.1612^\circ$, the expanded angle for the entire dataset is calculated as per equation (4.10) and (4.11) and given to be

$$\begin{aligned} \theta_{corr} &= \arctan\left(\frac{\tan(90.1612-90)}{15}\right) \\ &= 0.0107^\circ, \end{aligned} \tag{5.2}$$

where *scale* is given to be 15 from interpolation. The value 15 was chosen so that the provided data interpolates to 1 m in range, and

$$\begin{aligned} C_{tD} &= \frac{\tan(\theta_{corr})\cdot 64580}{2} \\ &= \frac{12.1155}{2} \\ &\approx 6, \end{aligned} \tag{5.3}$$

with $t_s = 64580$ samples, the number of time samples for dataset CFA16_002 is $\approx$ 13 s at a PRF of 5 kHz.

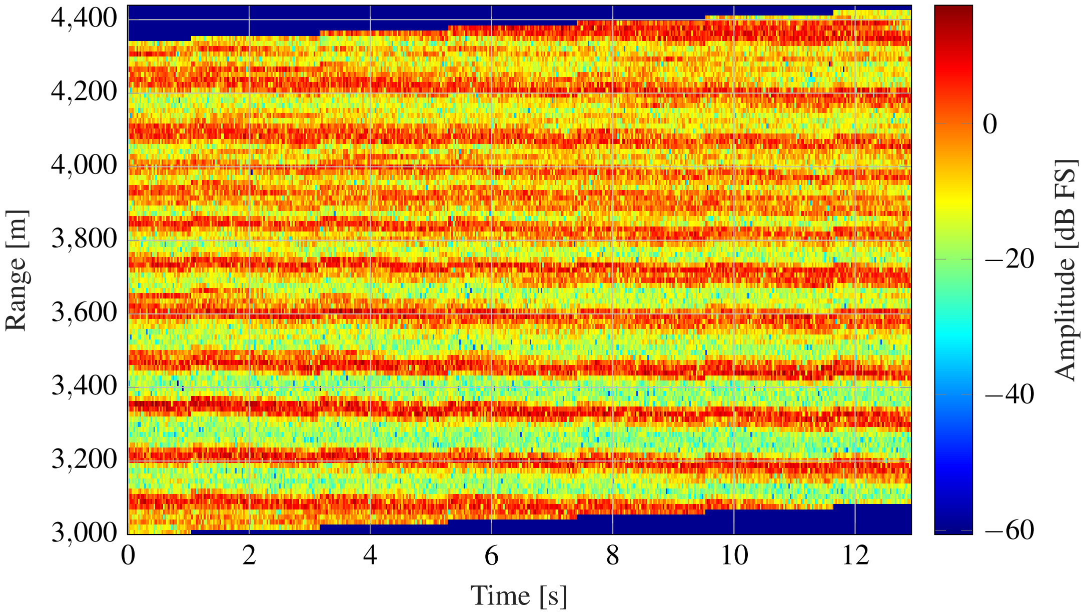


Figure 5.14. CFA16_002 output of wave-phase corrected data with x-axis aligned waves illustrating 6 clear steps which arise from the value determined by C_{tD}.

Applying this correction value to the original input data rectifies the approaching waves such that they are parallel to the x-axis. The correction steps are clearly shown when applied to the entire range of the original data as shown in Figure 5.14.

5.6 ANALYSIS OF ALIGNED WAVES

With all wave crests approaching the radar parallel to the time axis (x-axis), investigations into the statistical properties of the data can be conducted. This allows for further understanding of the wave structure and is also required for the parametrisation step needed for generating periodic breaking waves.

As previously mentioned, the method assumes that a fractional change in grazing angle due to the angular correction of the measured data is negligible in terms of the clutter statistics. For dataset

CFA16_002 this difference is calculated from equation (4.1) to be

$$\begin{aligned} \Delta\theta_{gr} &= \arcsin\frac{66}{6\cdot 15} \\ &\approx 0.014^\circ. \end{aligned} \tag{5.4}$$

5.6.1 Spectral analysis of aligned data

This section illustrates the spectral results obtained by analysing separate ranges of the corrected data, which has wave structures parallel to the x-axis. The statistical distribution analysis of a higher number of samples over the given wave structure section (range bin) provides insight into the radar returns and their respective distribution for separate parts of the ocean wave.

There are primarily two main components to consider during the analysis:

- spectral and distribution analysis over fast time, and
- correlation analysis over fast time.

The objective is to obtain parameters for the above characteristics, so that these can later be used during the generation process to parametrise the system.

After reshaping the two dimensional matrix into a vector, it is shown that the angular corrected PDF and the original data's PDF are identical, as shown in Figure 5.15. This merely verifies the data-alignment step and shows that the algorithm does not disregard any input data during the alignment process and is included purely for verification purposes.

Further, Figure 5.15 illustrates that the distribution of the entire (wave crest parallel to the x-axis) dataset considered against multiple fitting algorithms for both the Rayleigh and the K-distribution does not accurately fit the data. The plot is shown in the log-domain to illustrate the head and tail inaccuracies. Interestingly, the K-distribution fits the data better than the Rayleigh, although the number of samples is large.

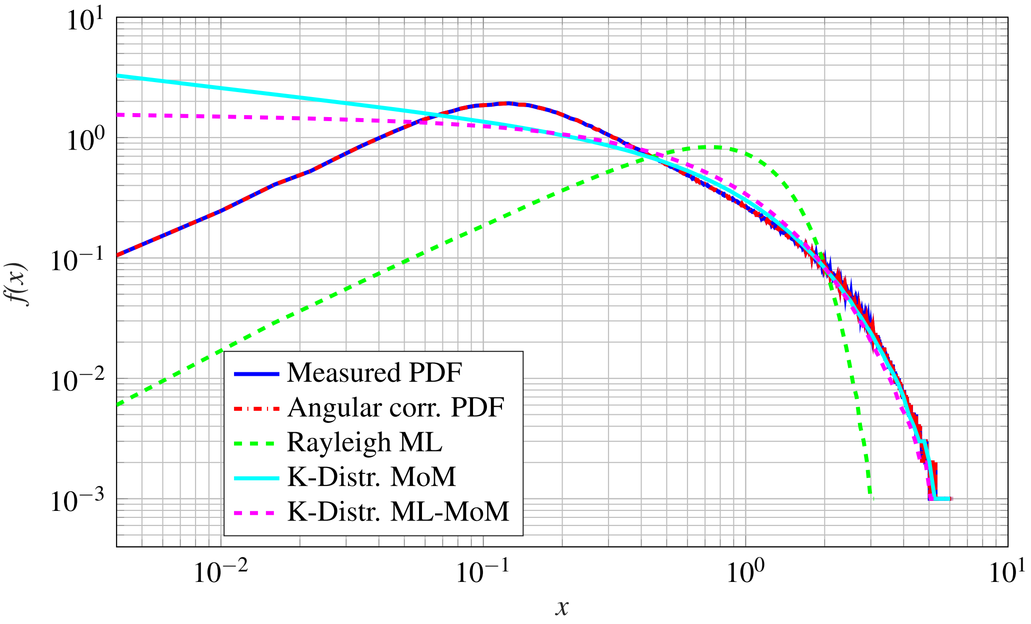
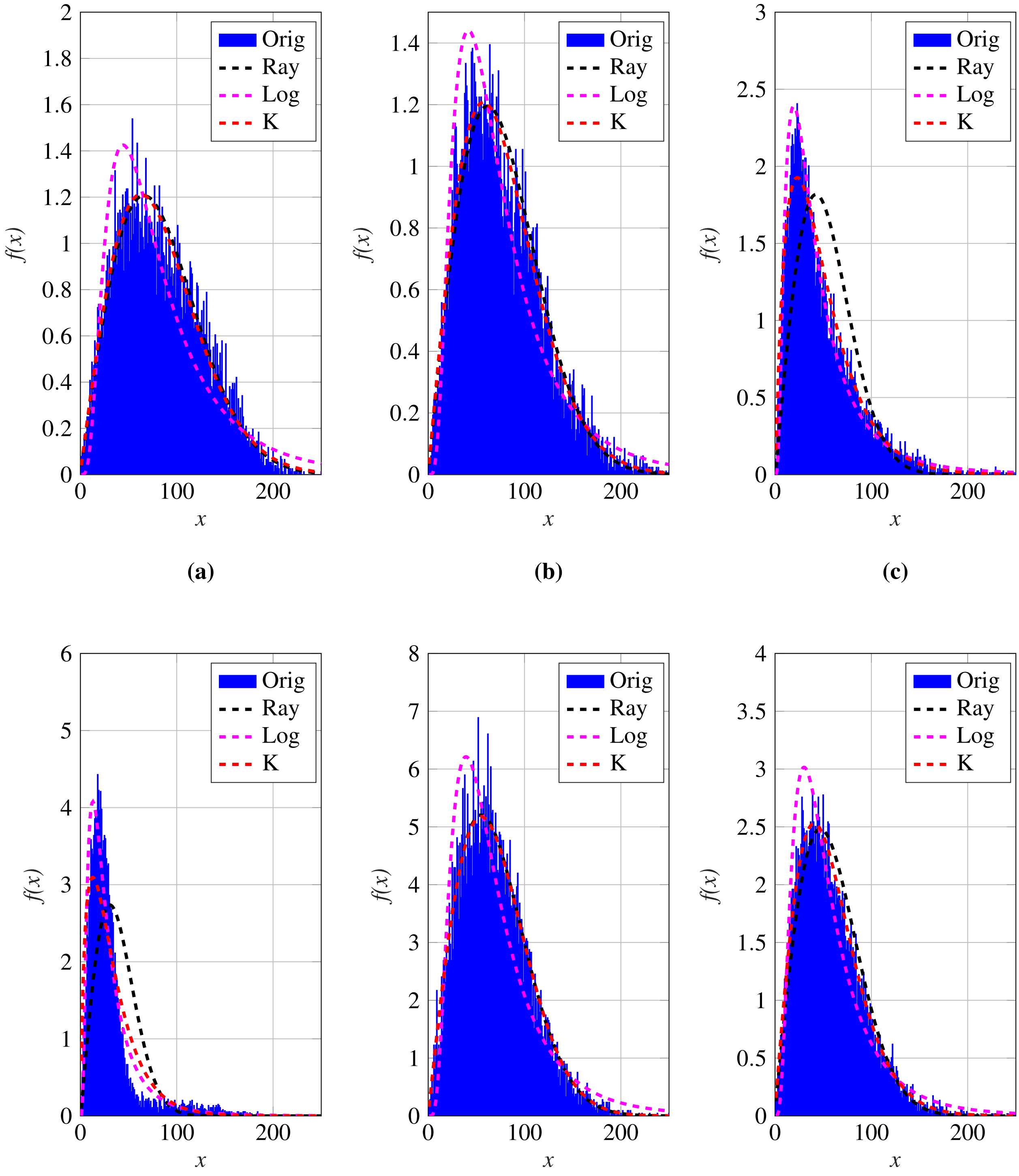


Figure 5.15. CFA16_002 PDF comparison of the vector reshaped output, verifying that the measured and angular aligned data is unchanged in amplitude, illustrating Rayleigh and K-distribution fits in the log-log domain.

5.6.2 Distribution fitting

The angular aligned data, where the ocean wave crests are parallel to the x-axis, is statistically analysed in terms of distribution and correlation. The data is split into sections over the fast time axis, so that each range bin (1 to 96) is individually analysed over several seconds of data ($\approx$ 13 s). Different distributions are fitted to the data ranges and their fitting parameters are compared and plotted against each other. For illustrative purposes a range of 180 m is plotted in Figure 5.16 and 5.17. Further plots can be found in Appendix A.

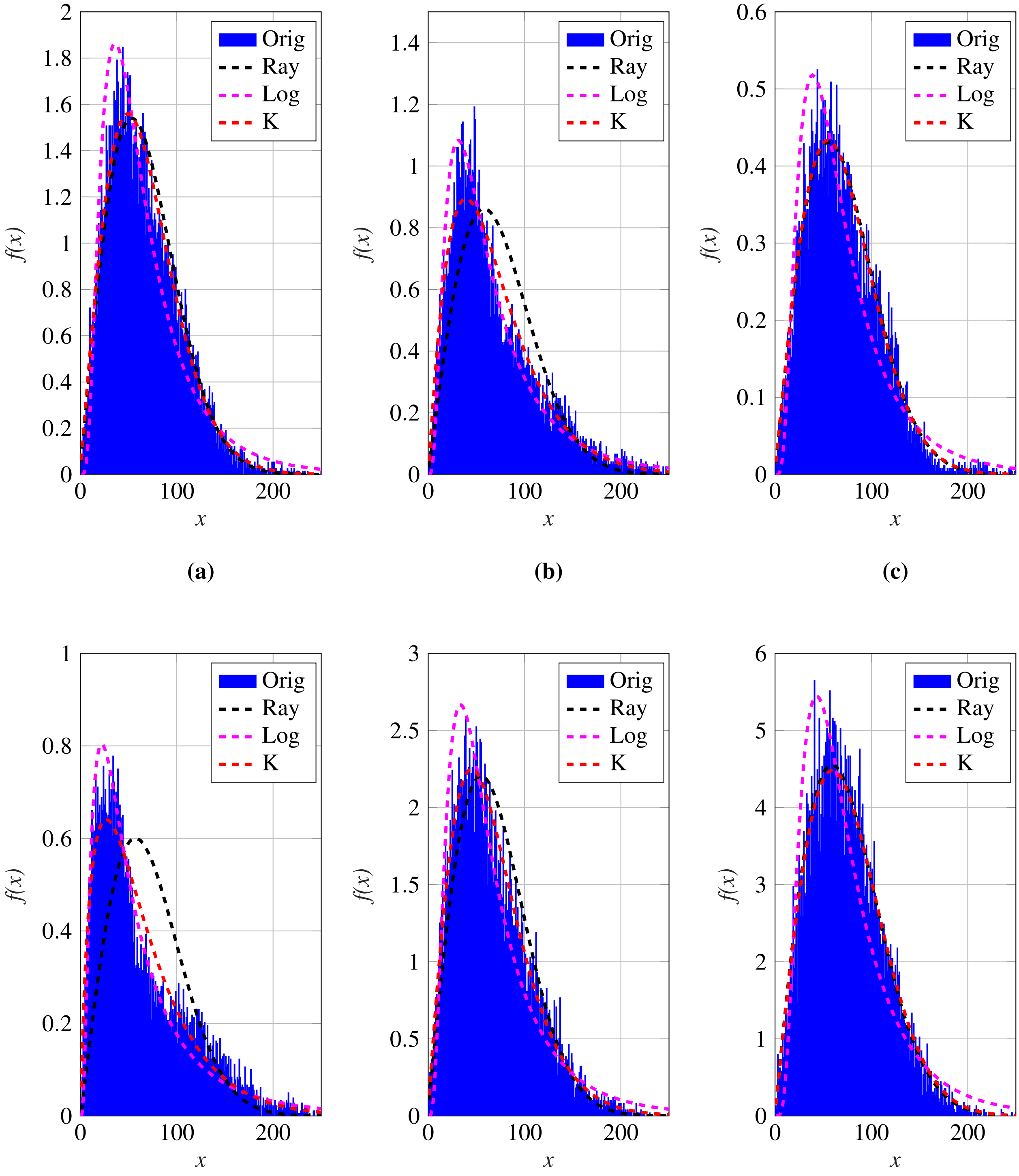
It can be observed, from Figure 5.16 and 5.17, that the K-distribution suits the dynamic range of the input data over multiple wave structures and points well in comparison to the other two distributions. As seen in these figures there are sections of the wave where the K-distribution performs similarly to the Rayleigh distribution. This is explained by the structure of the K-distribution, which for certain parameters is Rayleigh in nature. An overview of these fitting parameters for the shape and scale parameters are shown and discussed later (see Figures 5.26 - 5.28). These summaries provide an overview of the PDF performance and tendencies for different data ranges, effectively parametrising different sections of the periodic wave as it is changing from crests through to troughs.



(a) (b) (c)

(d) (e) (f)

Figure 5.16. CFA16_002 distribution fit example at range (a) 3652 m, (b) 3667 m, (c) 3682 m, (d) 3697 m, (e) 3712 m and (f) 3727 m, illustrating the periodic nature of the wave and the respective fitting of the distributions to different sections of the ocean wave structure.

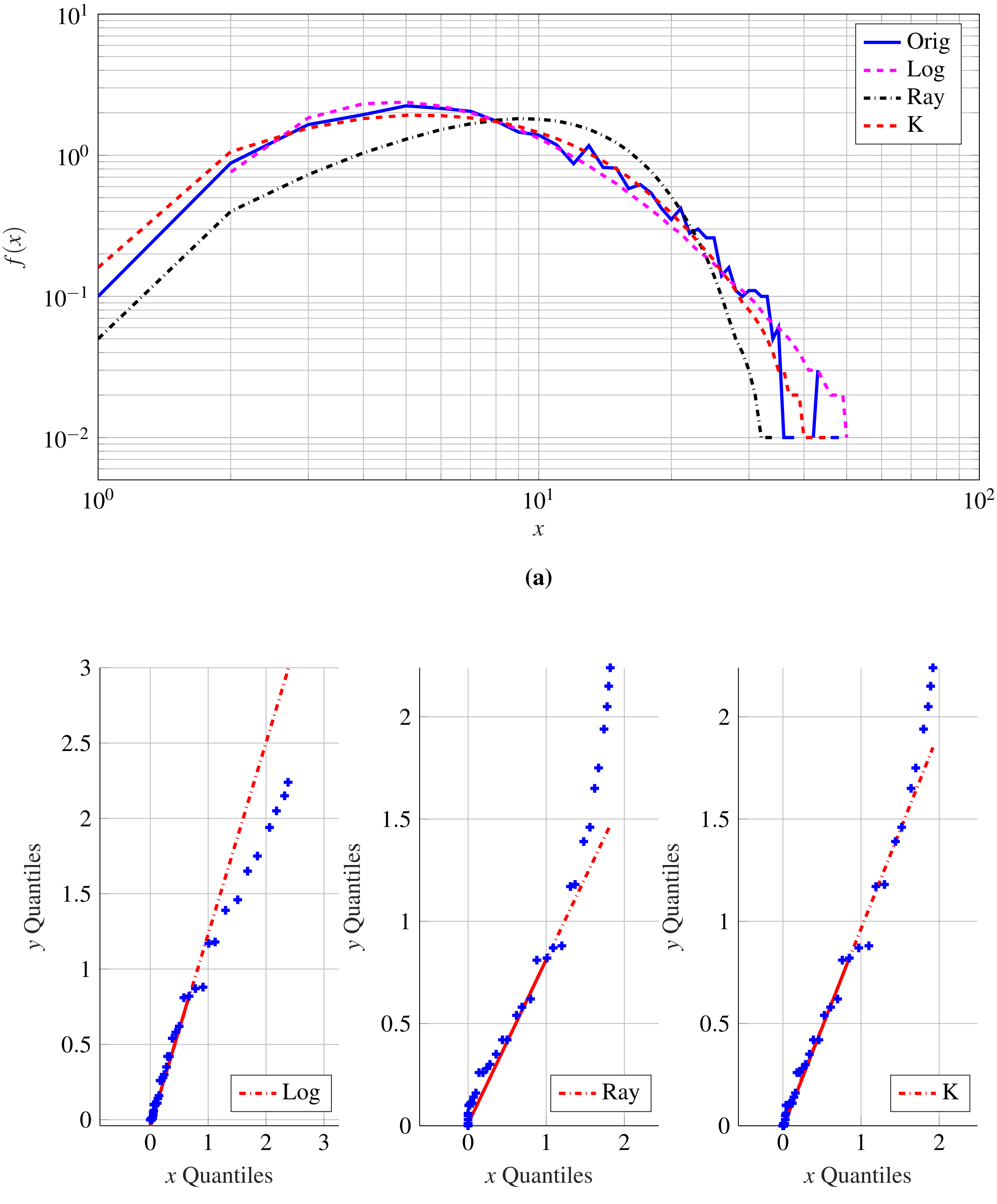


(a) **(b)** **(c)**

(d) **(e)** **(f)**

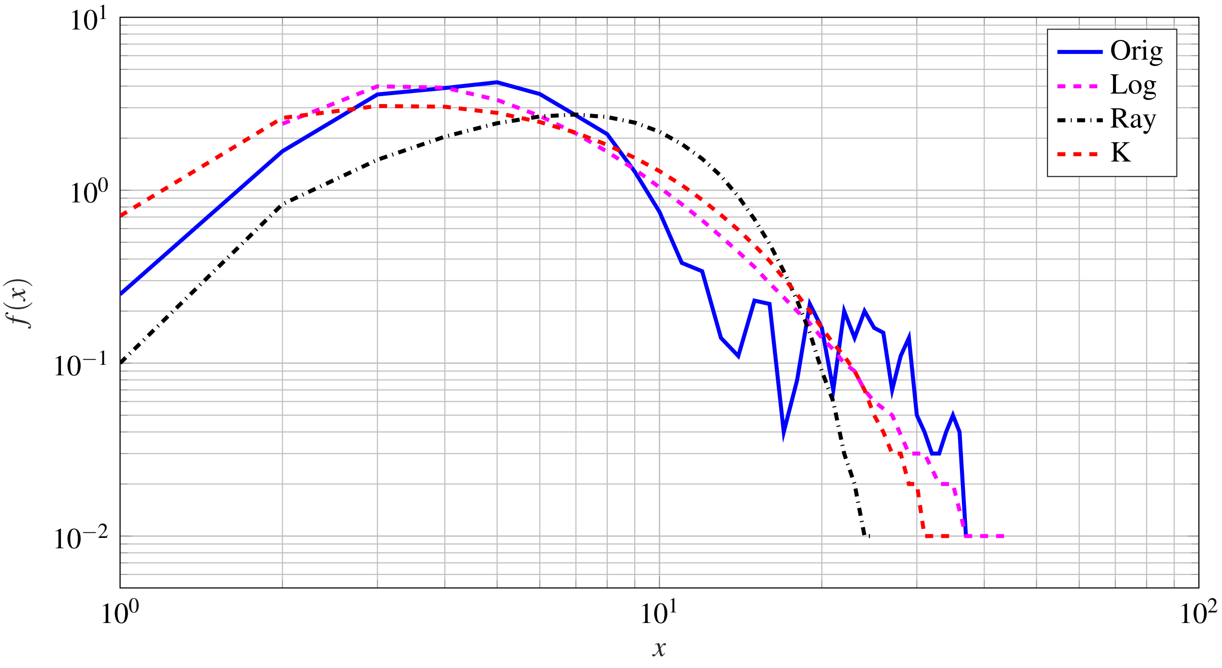
Figure 5.17. CFA16_002 distribution fit example at range (a) 3742 m, (b) 3757 m, (c) 3772 m, (d) 3787 m, (e) 3802 m and (f) 3817 m, illustrating the respective fitting of the distributions to the different sections of the ocean wave structure.

Further, in order to evaluate the goodness of the distributional fit, goodness of fits were visually observed and plotted. The PDF of the data, together with the fitted PDF of the log-normal, Rayleigh and K-distribution are also shown on the same figure in subplots (a) of Figures 5.18 - 5.25. The log-domain allows the inspection of the head and tail section of the distribution and provides insight into how well the distribution fits the given data. Further, the q-q plot illustrates the fitted distribution against the actual measured data using the method described in Section 3.2.3.3. This was performed for each range bin for the log-normal, Rayleigh and K-distribution. Figures 5.18 - 5.25 illustrate a visual analysis over a range distance of 135 m.

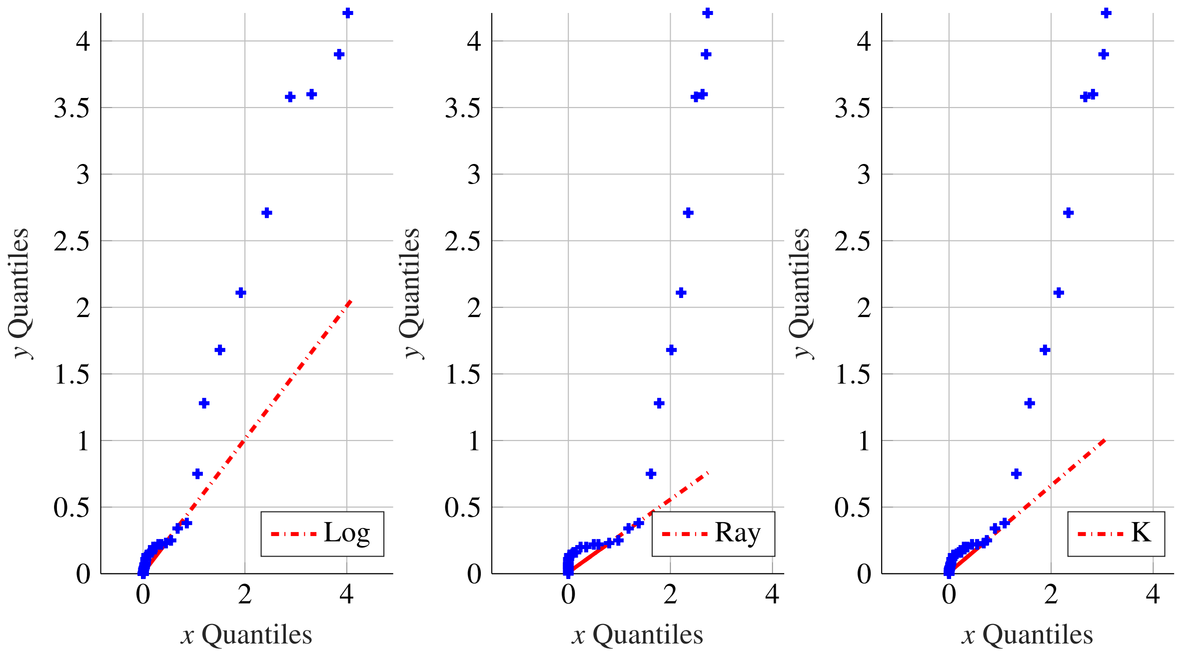


(b) **(c)** **(d)**

Figure 5.18. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3682 m (see reference Figure 5.16(c)), illustrating good K-distribution performance along the slope.

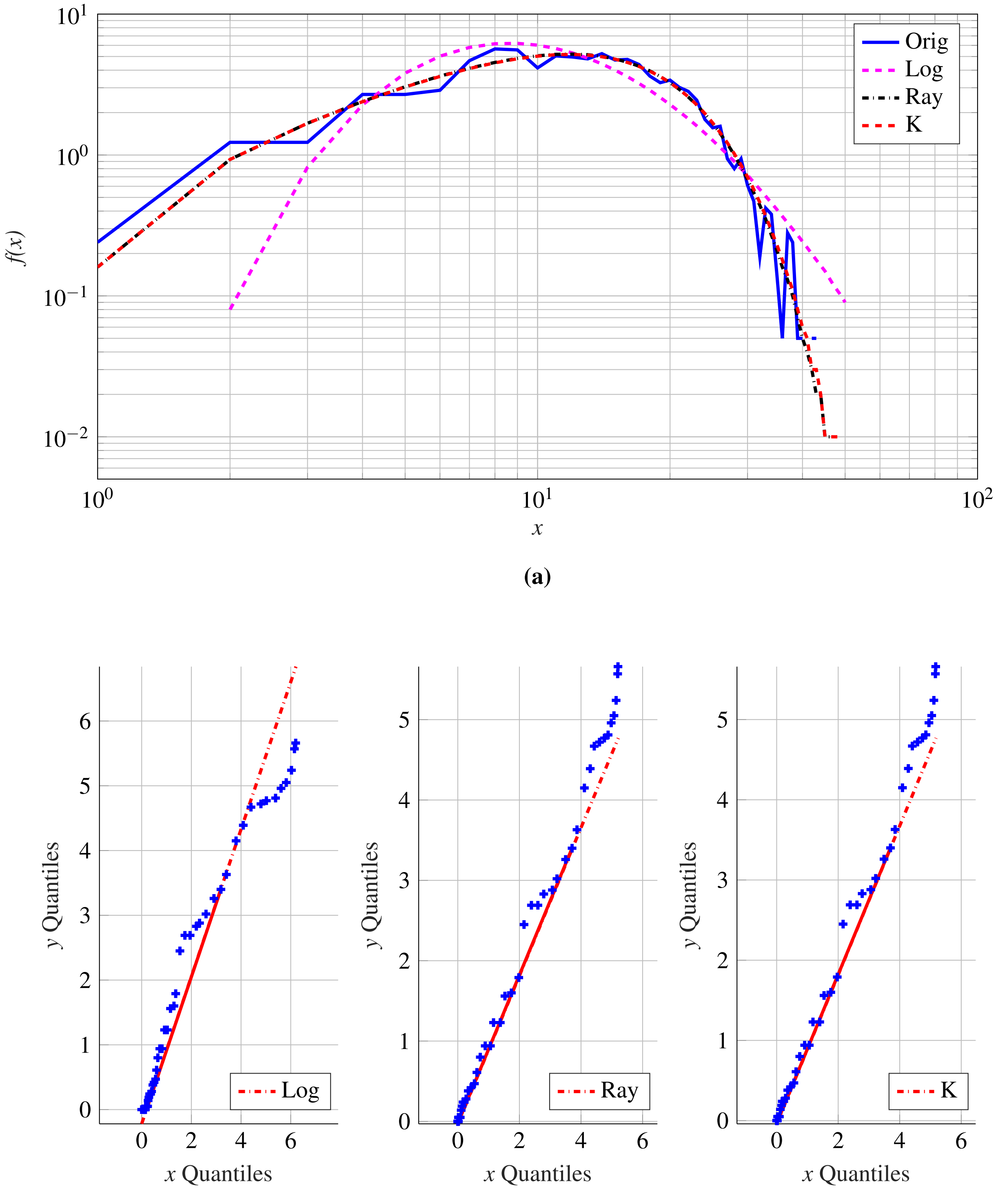


(a)



(b) **(c)** **(d)**

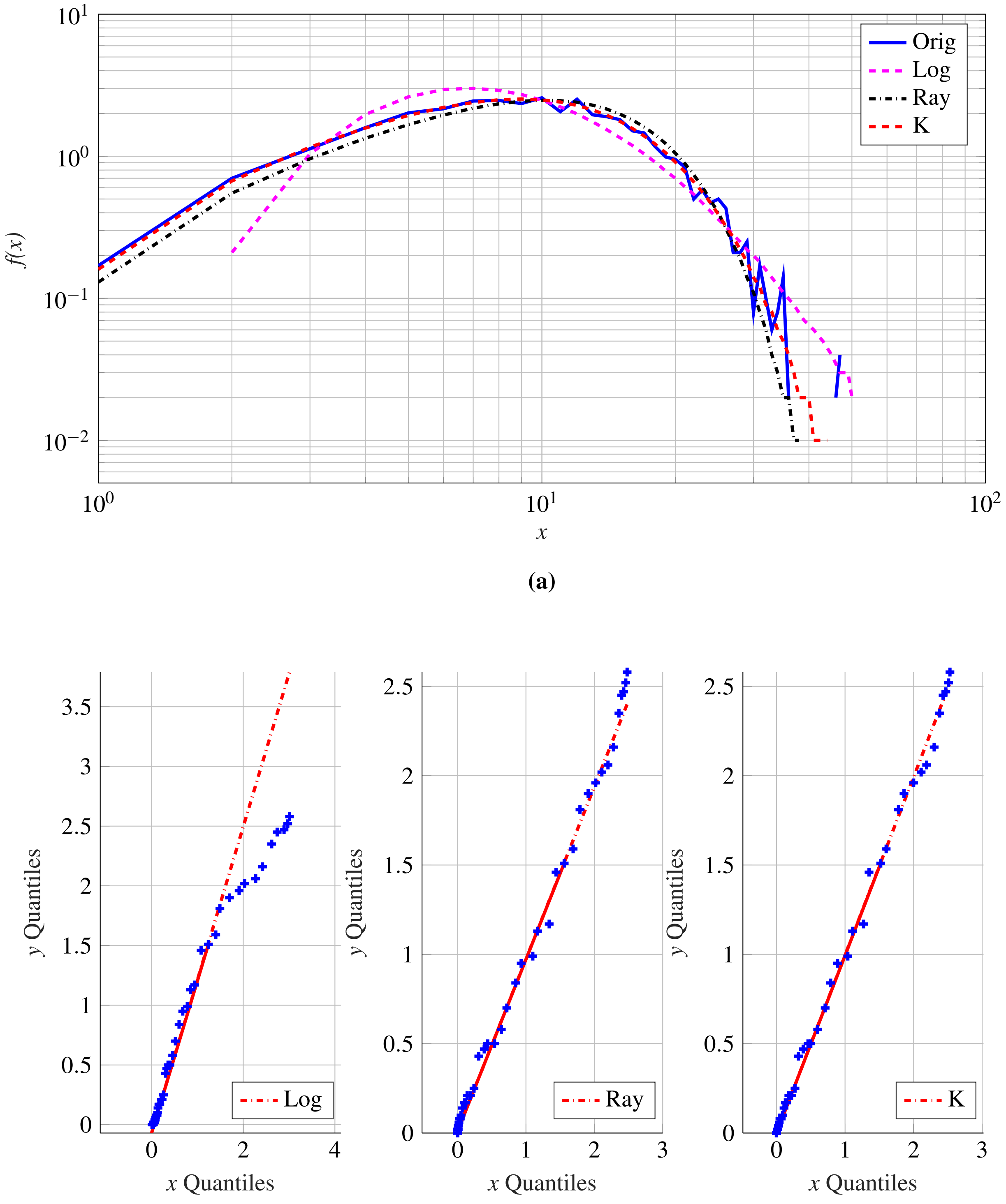
Figure 5.19. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3697 m (see reference Figure 5.16(d)) illustrating poor performance in the tail ends and deviating from the slope.



(a)

(b) **(c)** **(d)**

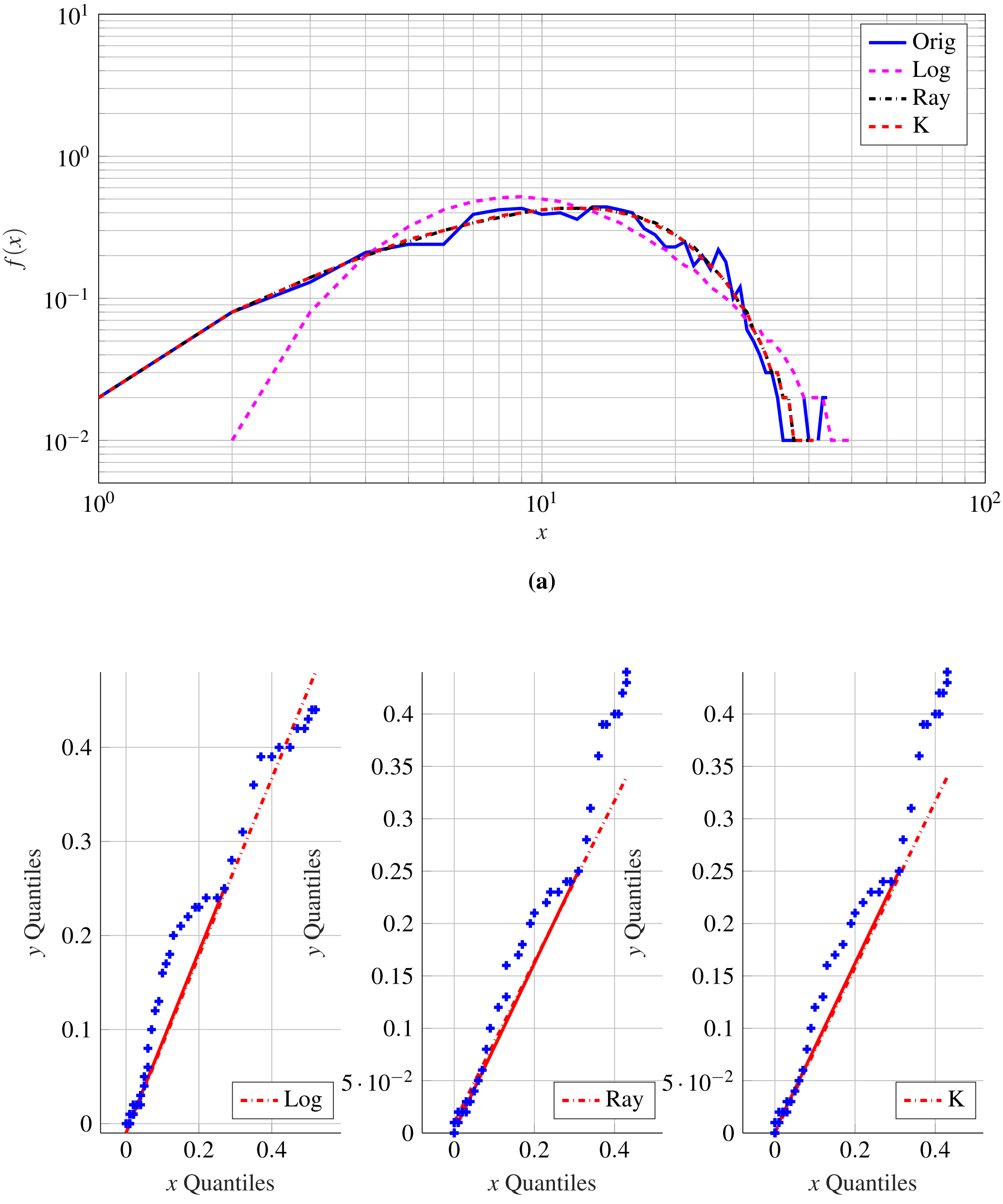
Figure 5.20. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3712 m (see reference Figure 5.16(e)) illustrating a similar performance between the Rayleigh and K-distribution.



(a)

(b) **(c)** **(d)**

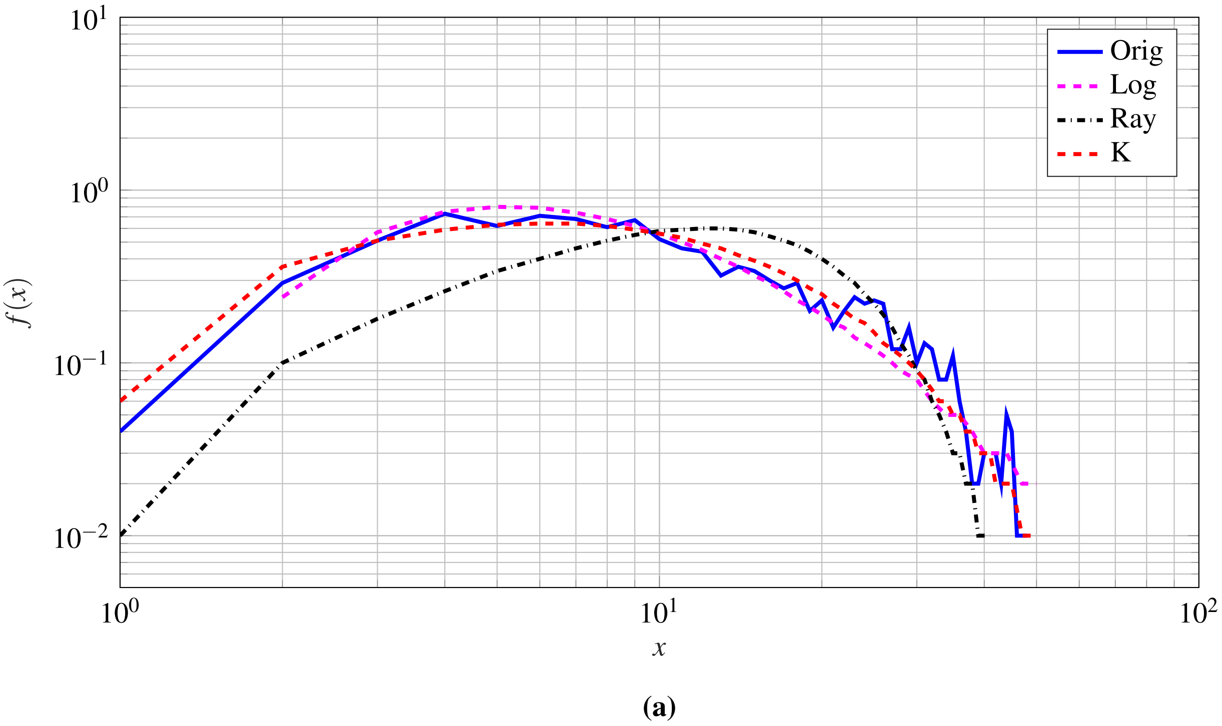
Figure 5.21. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3727 m (see reference Figure 5.16(f)) illustrating good performance by the Rayleigh and K-distributions.



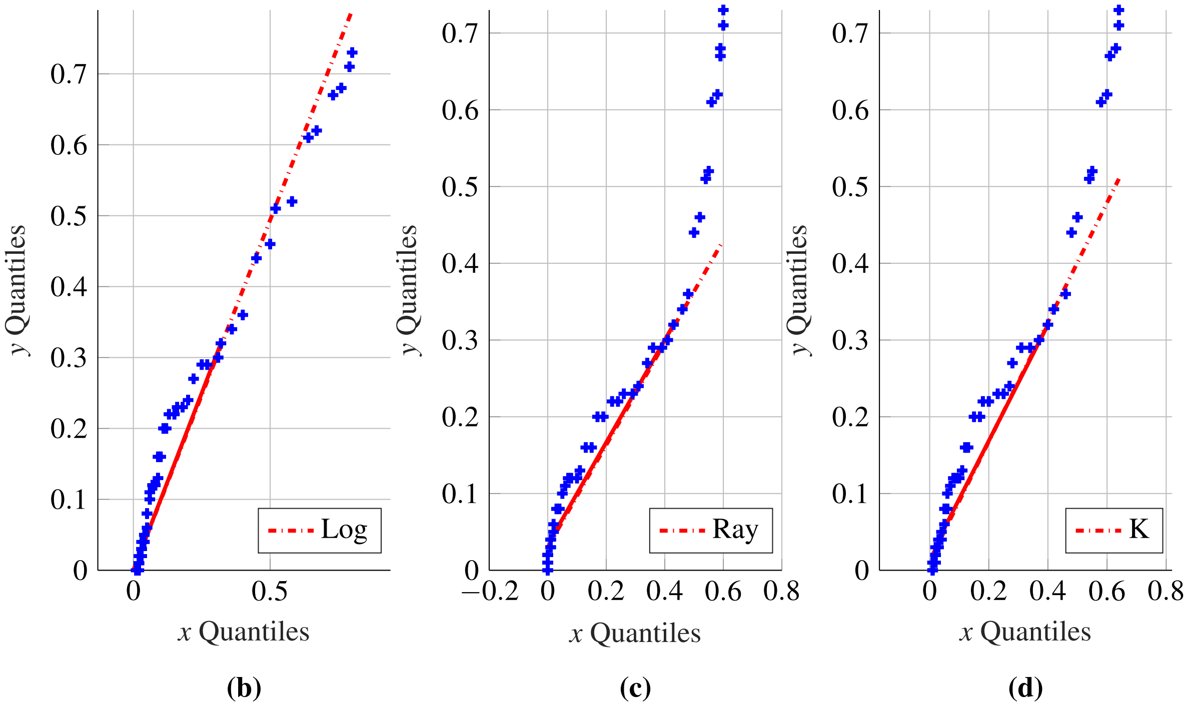
(a)

(b) (c) (d)

Figure 5.22. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3772 m (see reference Figure 5.17(c)) illustrating similar performance for all distributions.

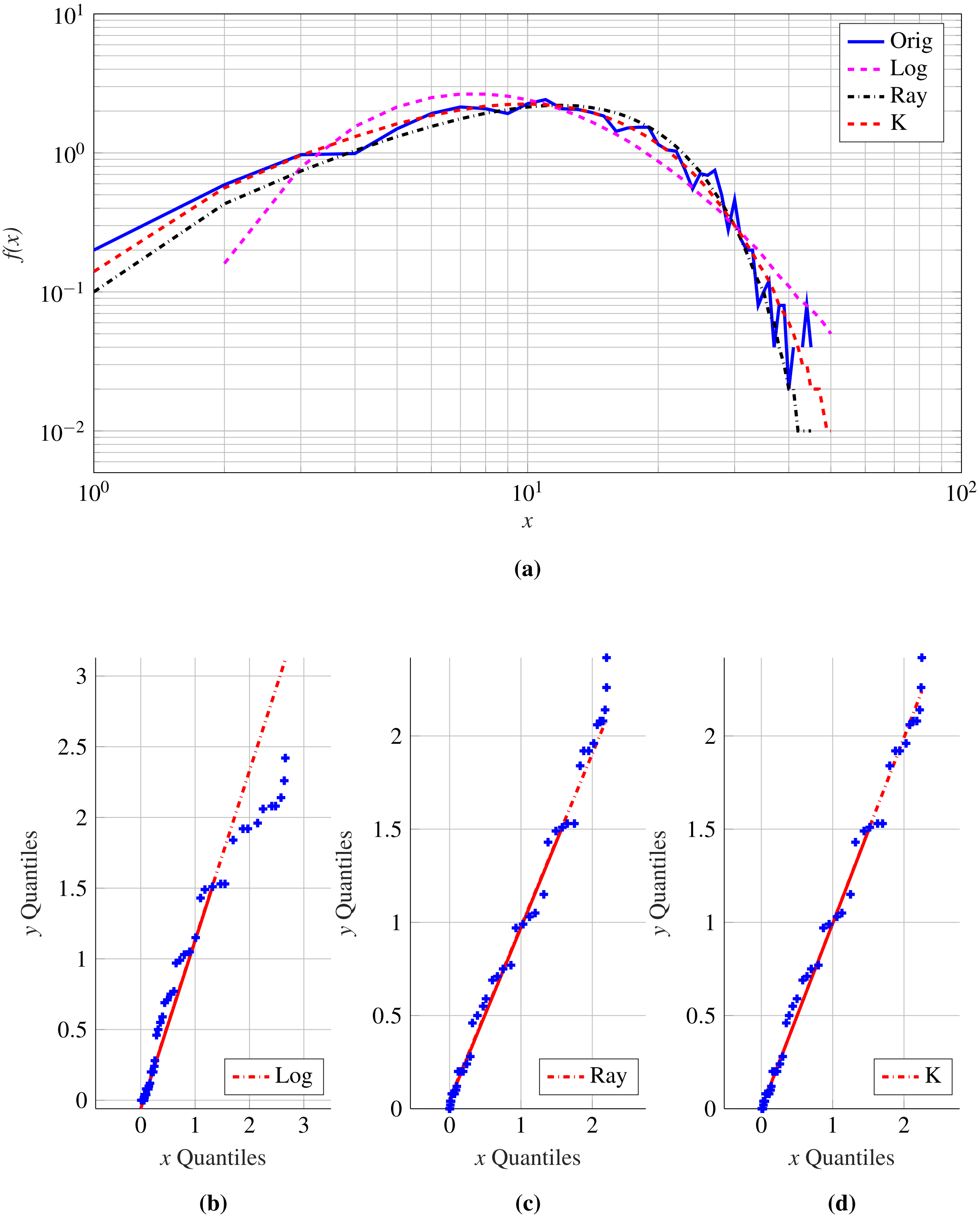


(a)



(b) **(c)** **(d)**

Figure 5.23. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3787 m (see reference Figure 5.17(d)) illustrating poor performance in the tail ends for Rayleigh and K-distributions.



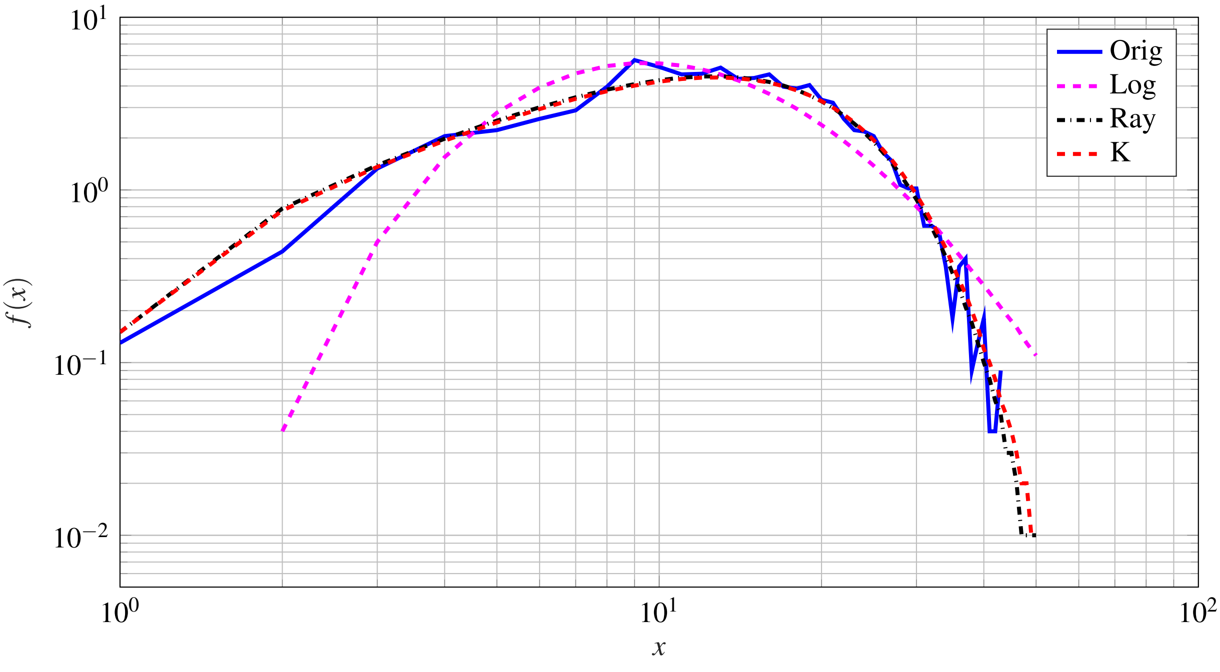
(a)

(b)

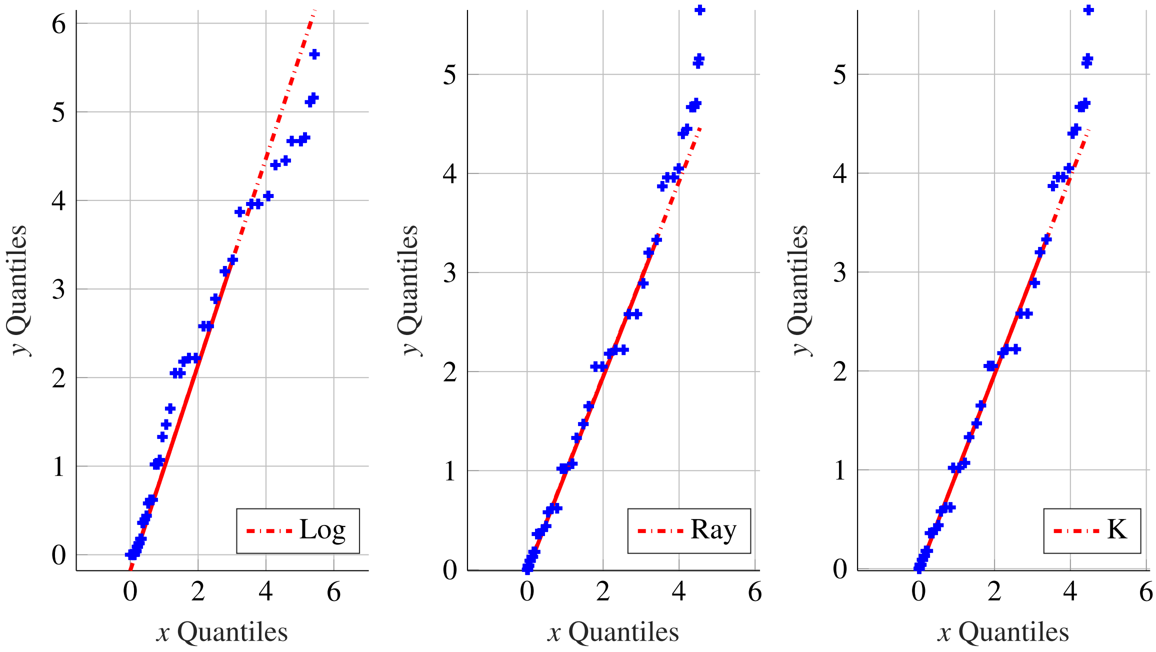
(c)

(d)

Figure 5.24. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3802 m (see reference Figure 5.17(e)) illustrating a similar performance by the Rayleigh and K-distributions.



(a)



(b) **(c)** **(d)**

Figure 5.25. CFA16_002 distribution log-log plot shown in (a) with goodness of fits, shown in (b) for log-normal, in (c) for Rayleigh, and in (d) for K-distributions at a range of 3817 m (see reference Figure 5.17(f))illustrating a similarly good performance by the Rayleigh and K-distributions.

The log-normal distribution, as suggested above, offers no sufficient fits to the provided data. This is supported by Figure 5.26, which illustrates the log-normal fitting parameters (μ and σ) over the entire analysed range for the given dataset. The q-q plots of Figures 5.18 to 5.25 show a large deviation from the ideal slope as shown in the respective (b) subplots of the figures. As expected this coincides with the (a) distribution plots shown in the same figures, where large sections of the data do not follow the PDF, especially the tail section of the PDF deviates from the data.

The Rayleigh and K-distributions on the other hand tend to show better fitting results, as the data points tend to follow the ideal slope line for longer periods (see Figures 5.20(c,d), 5.21(c,d), 5.24(c,d) and 5.25(c,d)). The parameters of the Rayleigh distribution over the full range are given by Figure 5.27. This again depicts the Rayleigh distribution parametrisation of the analysed dataset over the entire range. The periodic structure of the ocean waves are identifiable with the fitting parameters, thus showing how the PDF is adapting to different sections of the ocean wave. This provides insights that measured radar data is a function of what section of the wave is being returned. In ranges 3000 m to 3800 m clear periodic structures are identifiable. Thereafter, the crests tends to roll-off into one another, as illustrated in the range-time plot of Figure 5.1. After 4200 m the periodic structure returns.

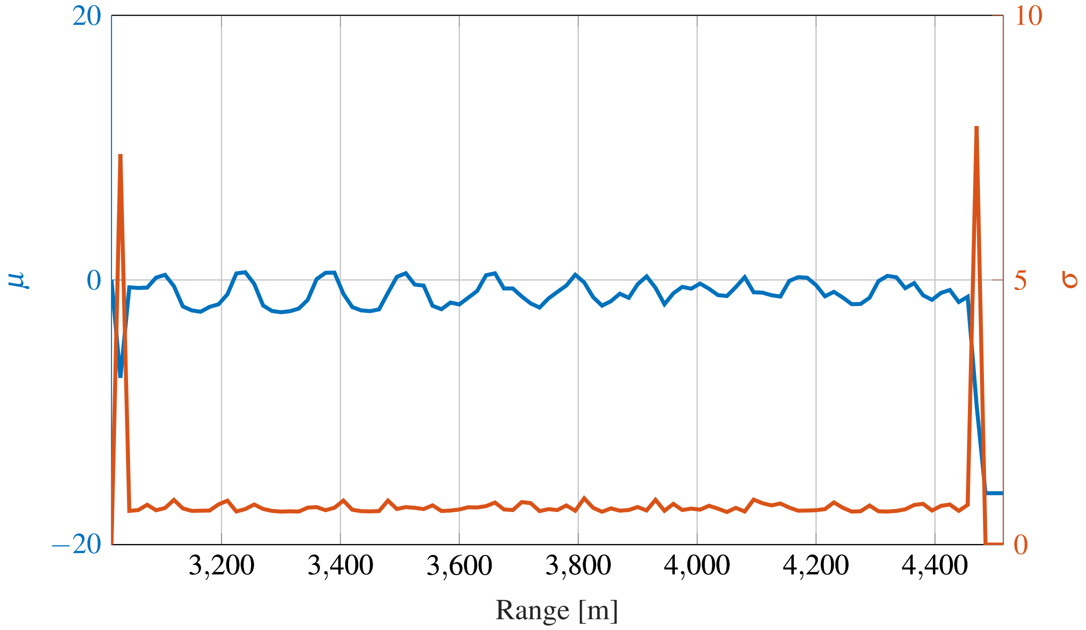


Figure 5.26. CFA16_002 log-normal distribution parameters plot illustrating the periodicity of the distribution fit over range.

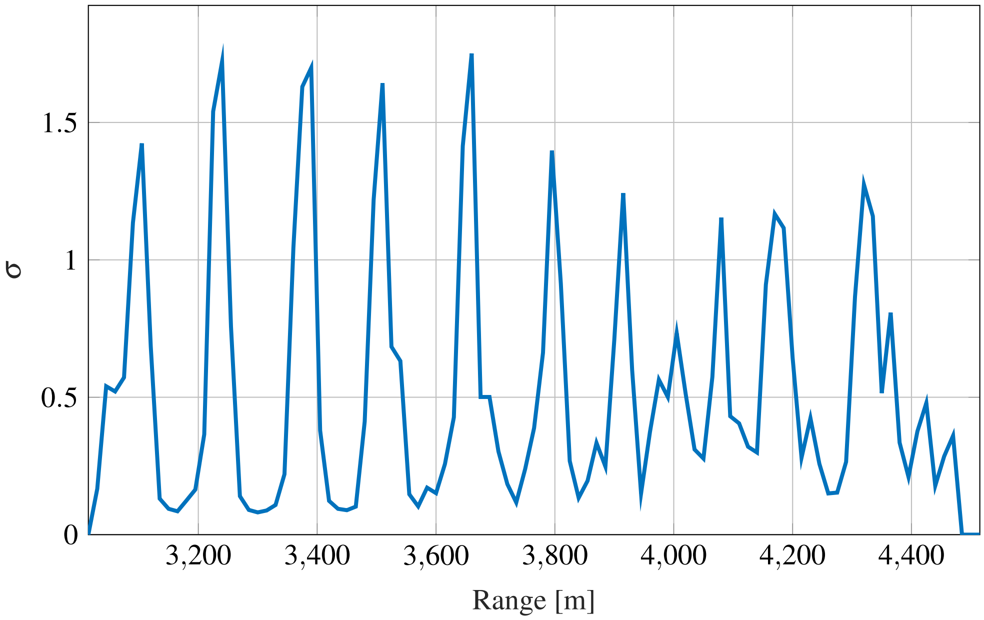


Figure 5.27. CFA16_002 PDF fit for Rayleigh distribution illustrating the σ parameter over range.

As mentioned in Section 3.2.3, the K-distribution is derived from the Rayleigh distribution with a varying mean power implemented as a random variable that follows the chi-distribution. Due to the compound structure of this distribution, the K-PDF fit handles the dynamic range of the analysed input measurements better over the entire range than the other two distributions. The K-distribution parameters are given in Figure 5.28 for the entire range of the analysed data. The shape parameter (using equation (3.30)) is given by the variable a and scale parameter by the variable v. As a tends towards ∞, which corresponds to a poor fit, the distribution becomes more Gaussian oriented. This implies that a relationship between a wave-phase section and the shape parameter exists. This is investigated below. The scale parameter v tends to follow the shape parameter a for poor K-distribution fits.

In order to evaluate these three distributions, an error based comparison is utilised to illustrate the effectiveness of each distribution against each other. The mean squared error between the actual data and generated data, based on the respective parametrised distribution, is provided in Figure 5.29. The log-normal distribution shows the largest error, whilst the Rayleigh and K-distributions show similarly good performance. The K-distribution offers slightly lower error bands, over the entire measurement range, than the Rayleigh distribution. The K-distribution also shows very good performance in the wave crest sections of the ocean wave. In order to illustrate the wave crest section in detail, where

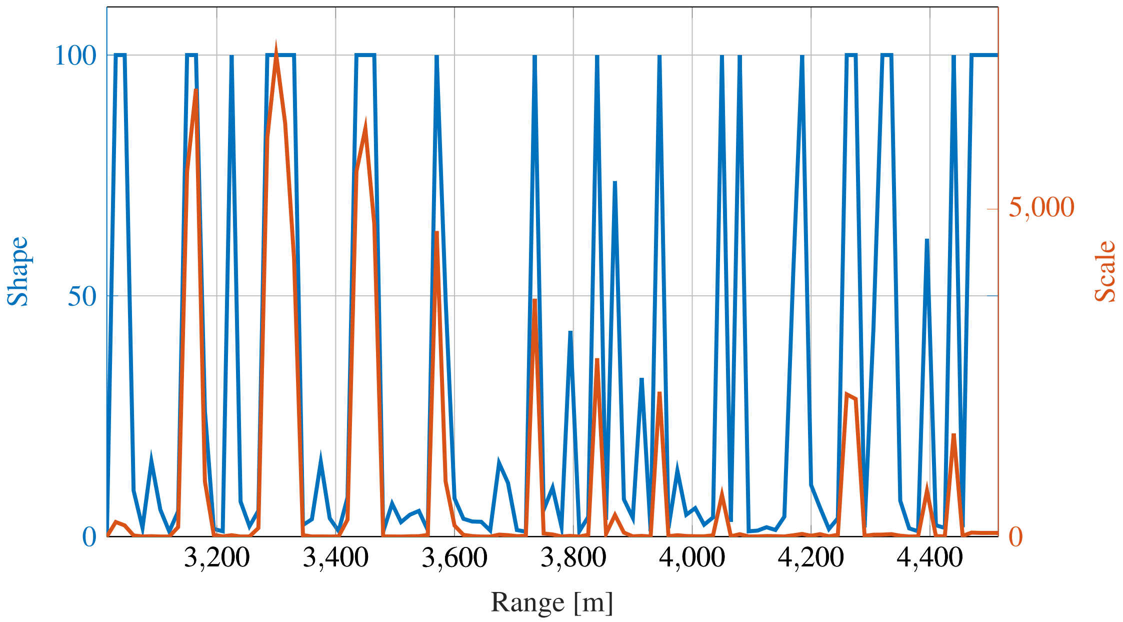


Figure 5.28. CFA16_002 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm, as large shape values correspond to poor fits and are better suited to Gaussian distributions.

the mean squared error is small, a log-domain plot of the same error band is shown for the three distributions in Figure 5.30.

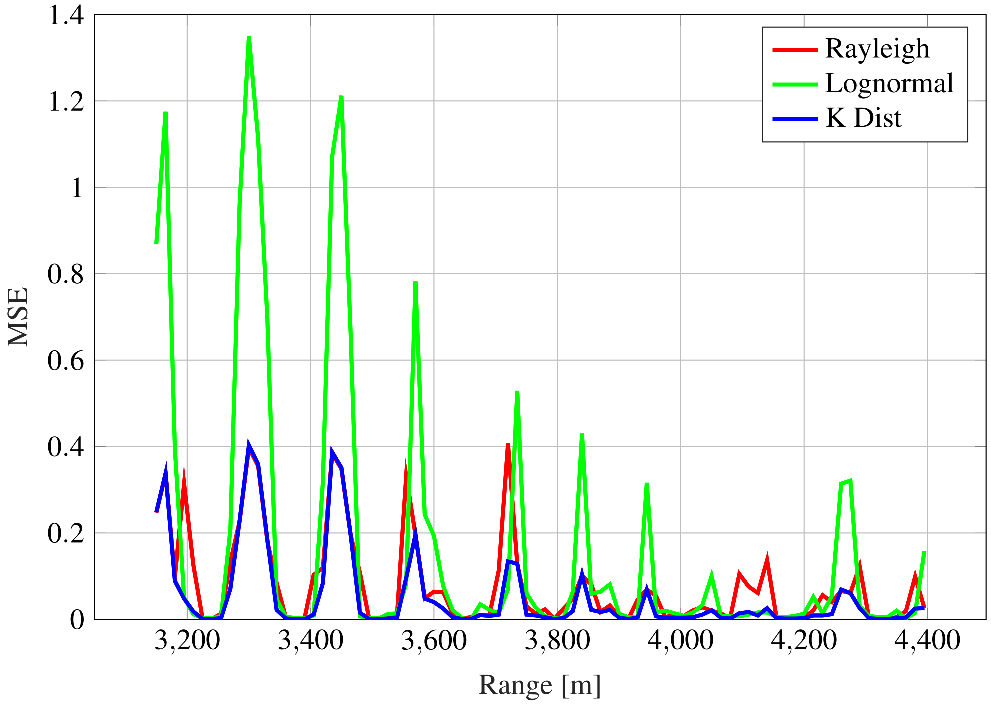


Figure 5.29. CFA16_002 parameter MSE comparing Rayleigh, log-normal and K-distributions over range, where higher errors relate to poor fits.

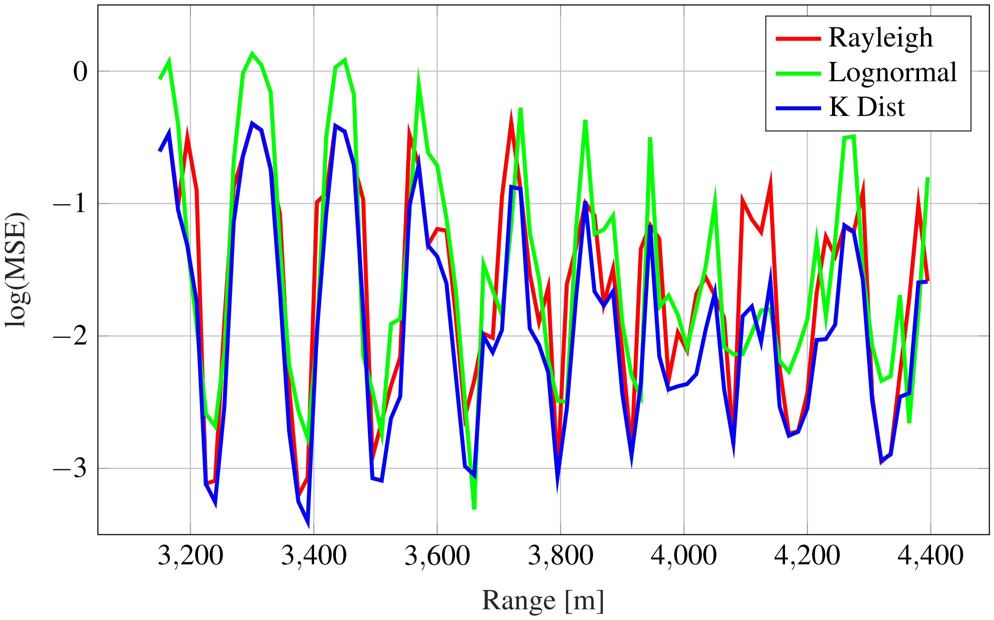
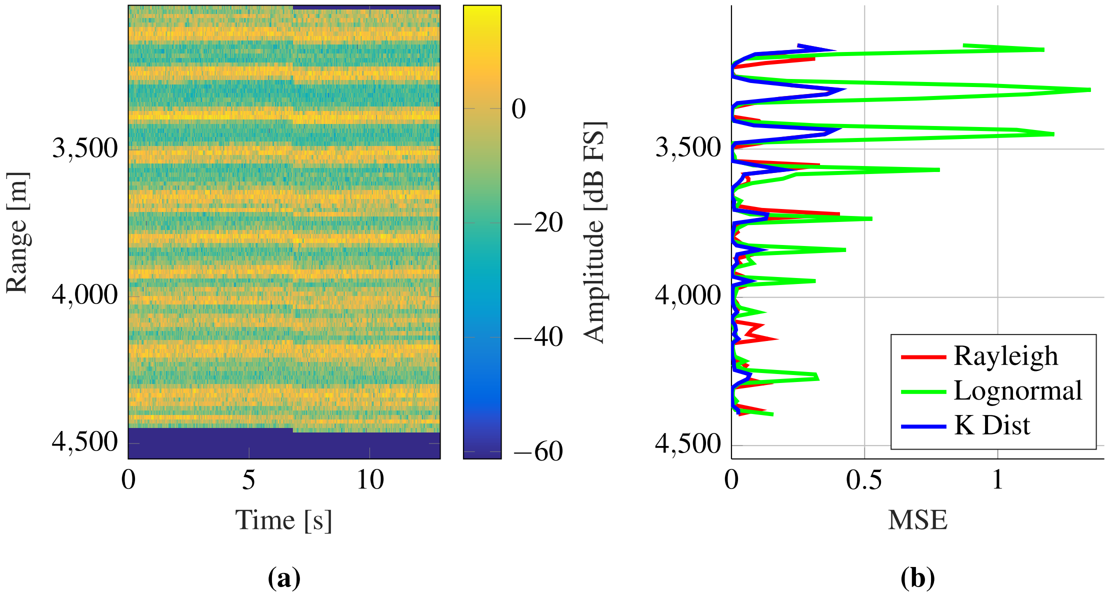


Figure 5.30. CFA16_002 parameter MSE comparison of distributions over range in log-domain, illustrating the small error obtained by good fits between parameter peaks.

In order to better understand how these parameters follow the structured periodic ocean wave, the distribution parameters were plotted alongside the original range-time measured radar data. This was done so that the correlation between radar amplitude, periodic wave and distribution parameters can be better understood. These are shown in Figures 5.31, 5.32 and 5.33. As the trough of the wave is reflected, the data is no longer K-distributed and tends towards the noise floor of the radar, implying a more Gaussian distribution. The K-distribution fits well to data that is reflected off the surface of the wave crest. The pattern of poor and well fitting sections, shown by the parameters, matches with the cyclic frequency of the wave crests within the data.



(a) **(b)**

Figure 5.31. CFA16_002 comparison showing alignment of (a) range-time measured radar data vs (b) distribution mean squared error, illustrating the periodic nature of good and poor fits with respect to sections of the ocean wave.

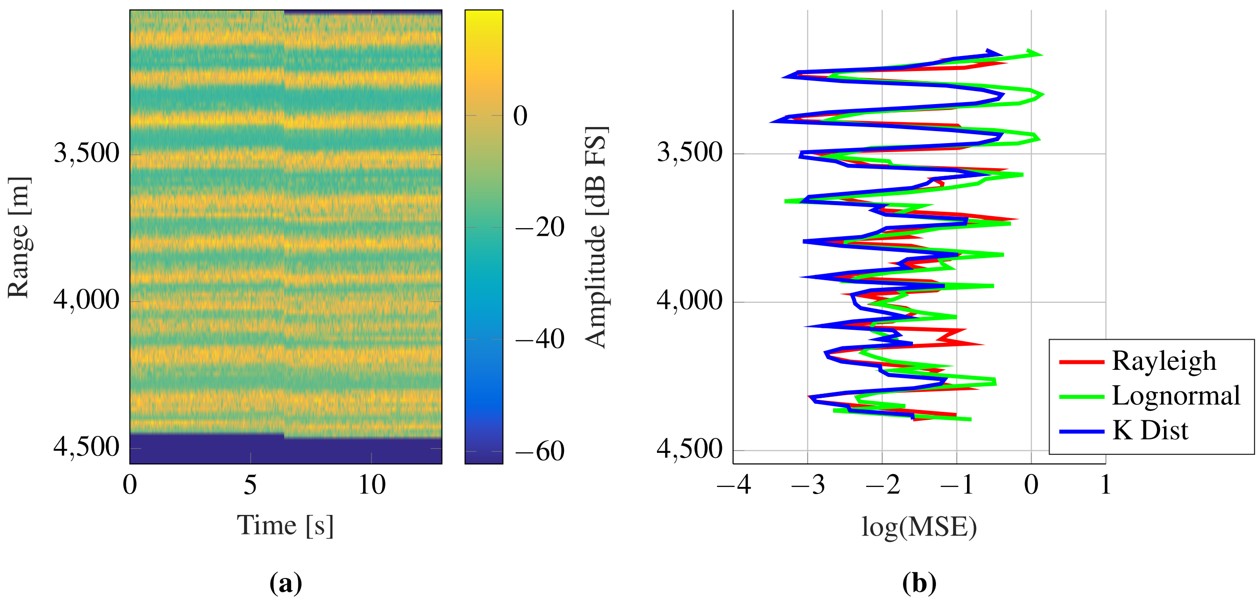


Figure 5.32. CFA16_002 comparison illustrating alignment of (a) range-time measured radar data vs (b) distribution mean squared error given in the log-domain to emphasize the good fitting sections.

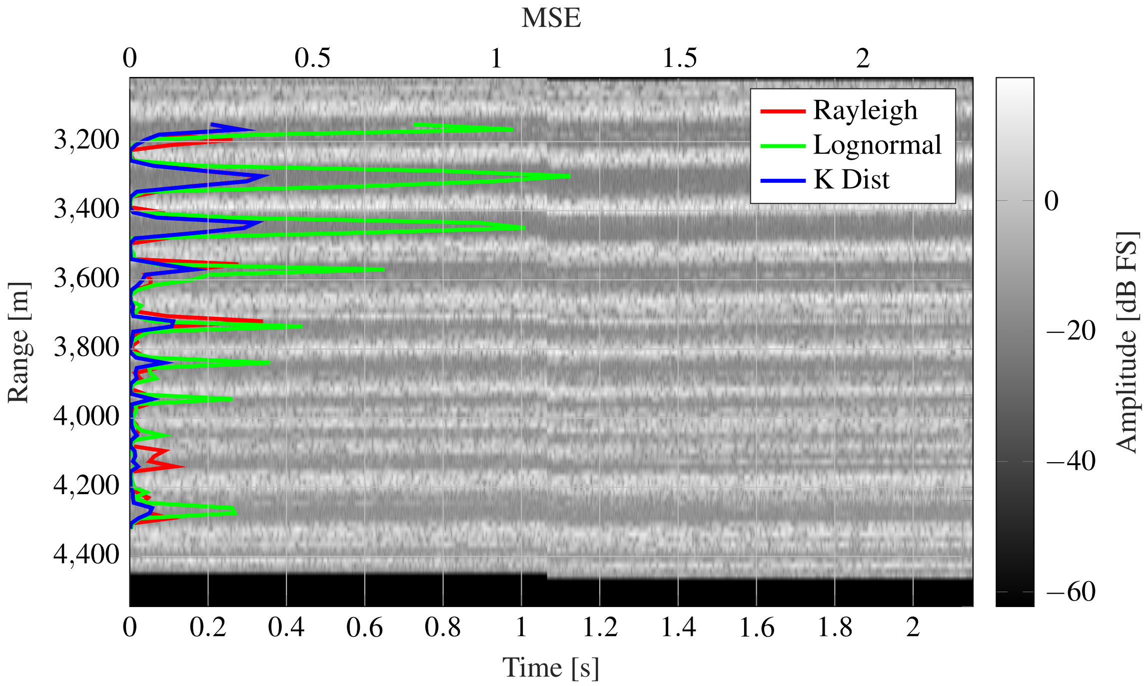


Figure 5.33. CFA16_002 comparison showing alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.

The fitting matches the measured data power with respect to estimated parameters. The peaks in the shape parameter correspond to troughs in the measured amplitude. This occurs where Gaussian noise dominates and causes the shape parameter estimator to tend towards ∞, whereas ocean wave crests which reflect higher intensity signals match the K-distribution. This results in low shape parameters and indicates non-Gaussian distributed data. Figure 5.34 presents the alignment over the entire dataset for convenience.

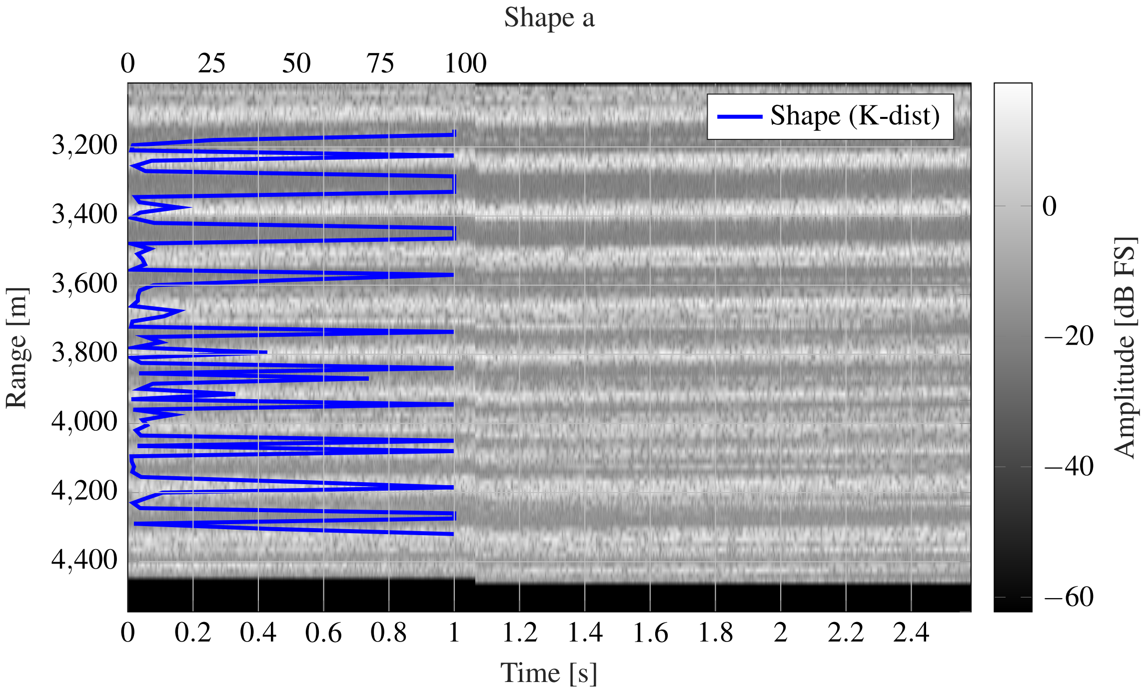


Figure 5.34. CFA16_002 range-time measured radar data aligned with shape parameters from K-distribution, illustrating trough and crest based parameters.

5.6.3 Correlation

Alongside the distributional properties, the correlation represents the other statistically important characteristic needed for the generation process. The correlation of each range bin over the full range of 96 bins is considered. Note that this is correlation in the fast time axis. The spectrum of the autocorrelation determines the spectrum width and thus the correlation, the wider the response spectrum the more uncorrelated the data. For a very narrow response a distinct frequency represents a strong correlation property. The frequency at which the correlation occurs is determined by the phasor, where positive frequencies represent a wave approaching the radar.

Figure 5.35, 5.36, 5.37 and 5.38 illustrates the correlation responses over a section of a 165 m. The correlation spectrum is broad for noise or Gaussian based responses, i.e. when the radar signal is returned from troughs, and narrower for crest based responses, where the periodicity of the breaking ocean wave reoccurs.

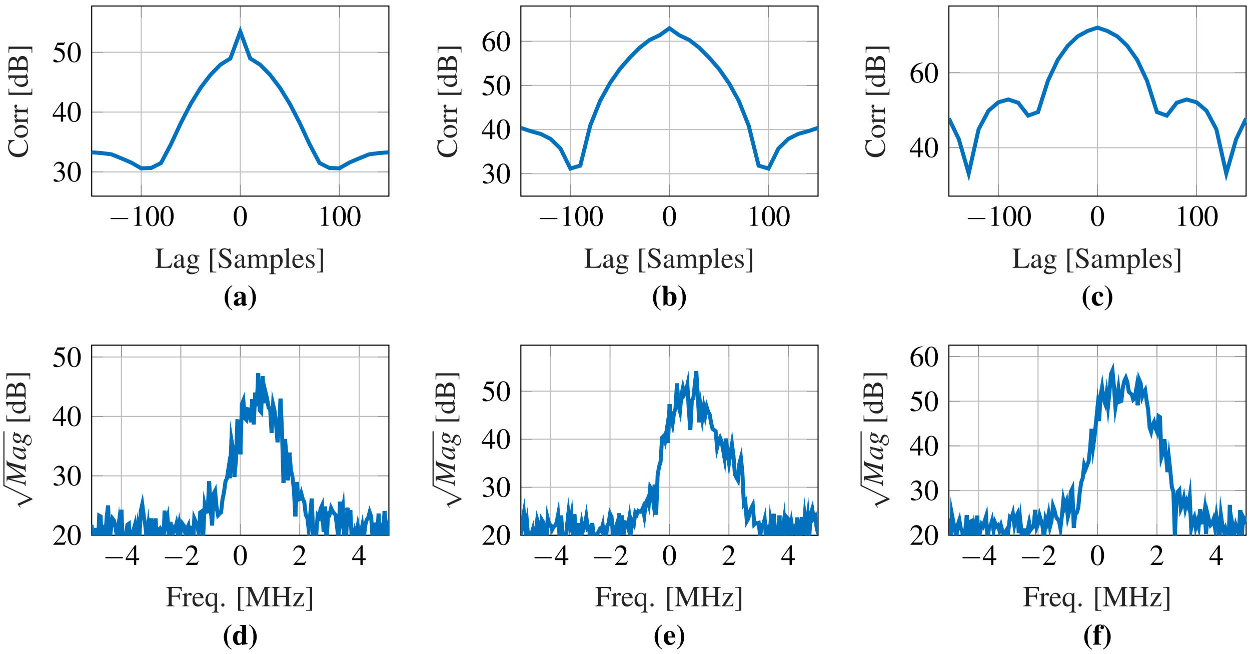


Figure 5.35. CFA16_002 correlation at range (a) 3622 m, (b) 3637 m and (c) 3652 m, where (d) to (f) illustrates a narrow response corresponding to ocean wave crests.

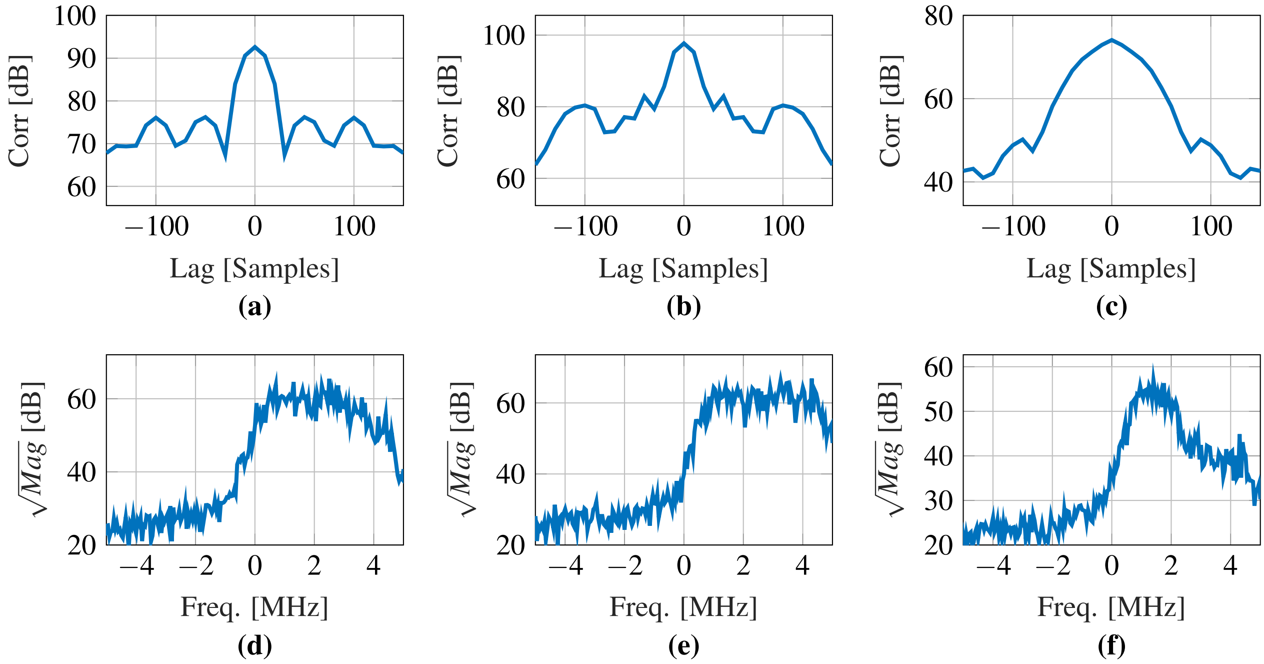


Figure 5.36. CFA16_002 correlation at range (a) 3667 m, (b) 3682 m and (c) 3697 m, where (d) and (e) illustrates a wide spectrum corresponding to ocean wave troughs whilst (f) tends towards the narrow spectrum of a crest.

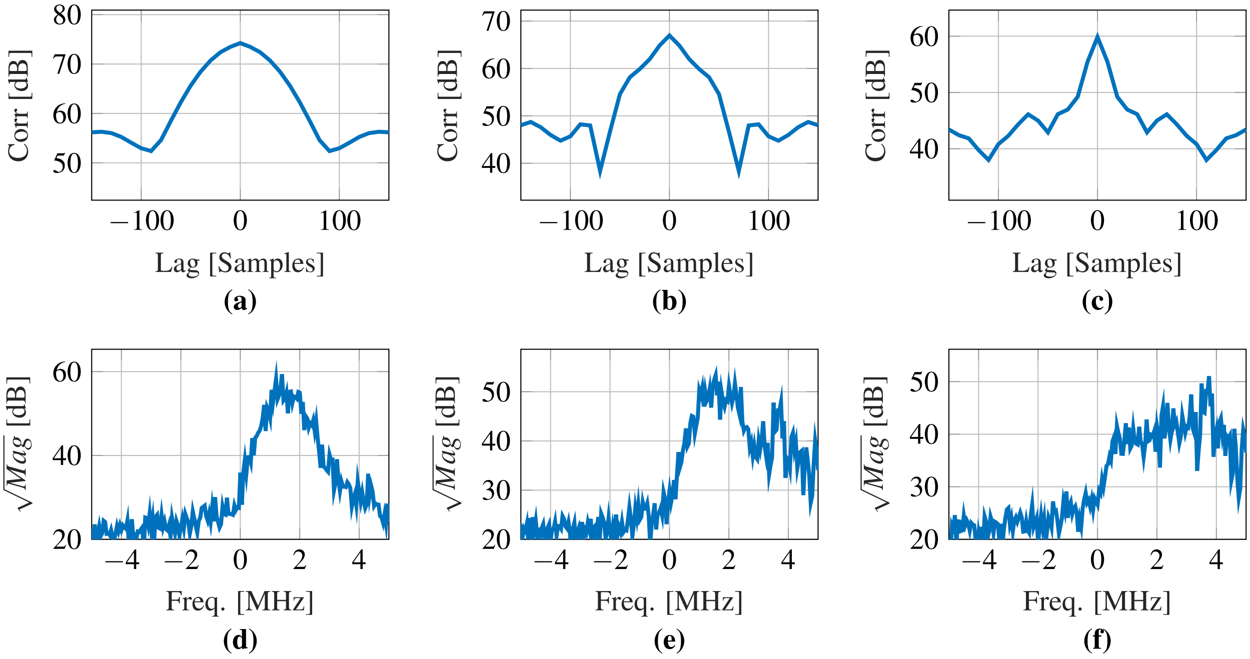


Figure 5.37. CFA16_002 correlation at range (a) 3712 m, (b) 3727 m and (c) 3742 m, where (d) illustrates a crest and (e) to (f) transitions back to a wide spectrum, which corresponds to ocean wave troughs.

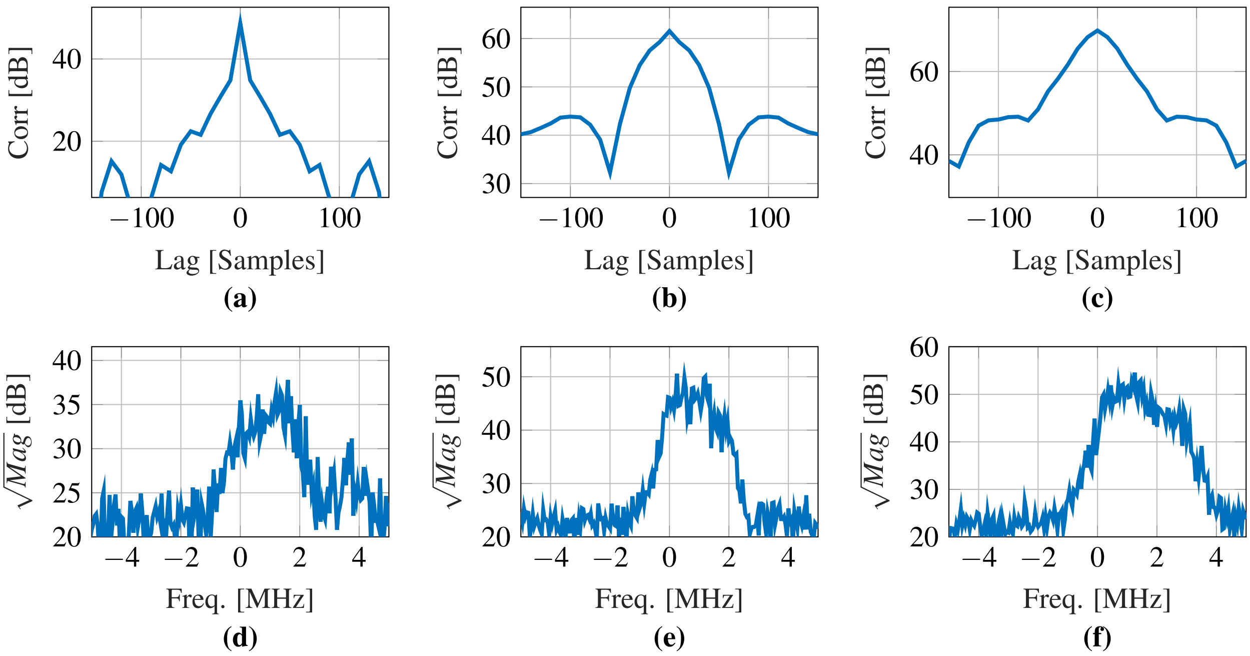


Figure 5.38. CFA16_002 correlation at range (a) 3757 m, (b) 3772 m and (c) 3787 m, where (d) illustrates a transition from trough to crest, and (e) and (f) illustrate the crest region again.

5.7 GENERATION PROCESS

Reversing the analysis processes and using the gained statistical insights, the generation process as described in Figure 4.3 can now be applied. For convenience this is re-plotted in Figure 5.39. First the random number generator outputs a sequence of zero mean values with a standard deviation of one. This is input into the respective generation function. For correlated Rayleigh see Section 4.8.2.1. K-distributed correlated data, see Section 4.8.2.2, is not simulated as it is considered out of scope for this document. Finally after the correlated and statistically distributed data has been generated, a desired ocean wave approach angle, for the range-time plot, can be applied.

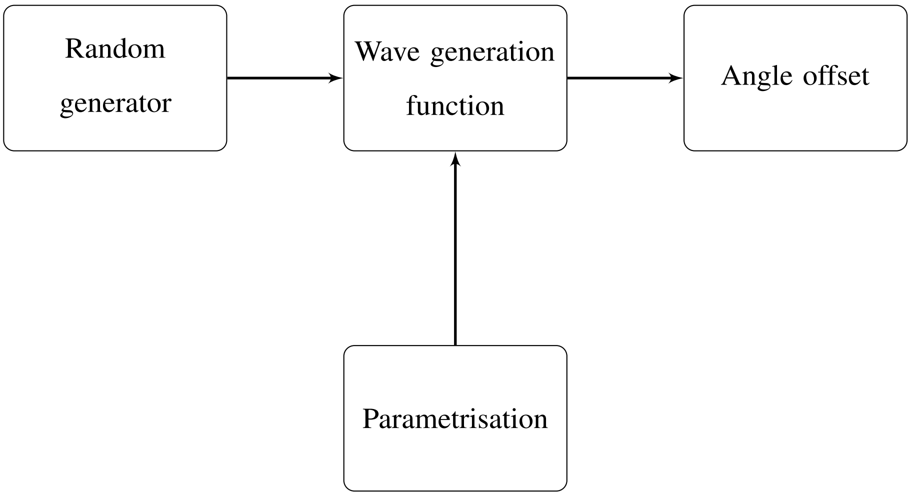


Figure 5.39. Process overview of data generation algorithm.

5.7.1 Generation of a Rayleigh parametrised system

First the corrected data is analysed and parametrised. In this particular case, an individual range bin is considered and analysed in order to evaluate and elaborate the method. The correlation of range bin 45, which corresponds to a range distance of 3667 m in the range-time plots above, is illustrated in Figure 5.40. The figure, given in dB, illustrates the relatively short correlation intervals of the range bin, suggesting, as discussed above, that the results stem from the ocean wave crest section. The power spectral density indicates that the wave is moving towards the radar with a positive Doppler frequency, whilst a negative frequency would indicate that the wave is moving away from the radar.

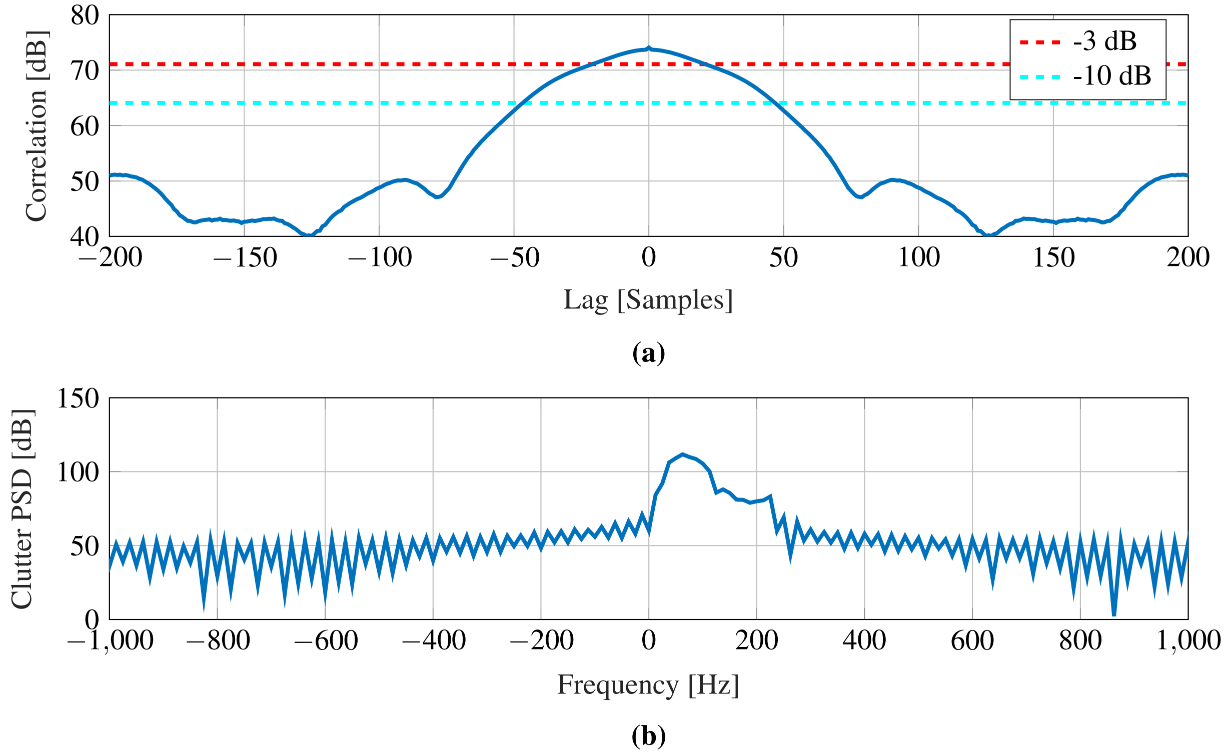
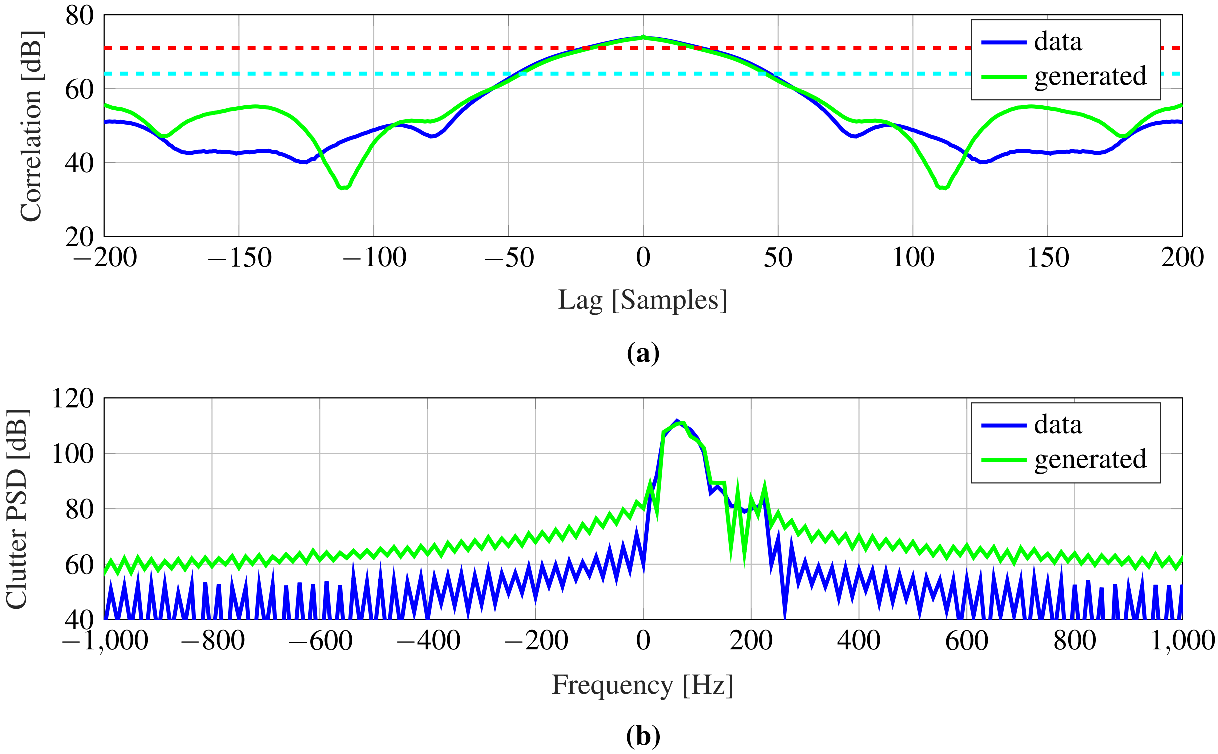


Figure 5.40. CFA16_002 PSD of angular corrected data at range 3667 m (bin = 45), with (a) showing a detailed view of the correlation around 0 lag and (b) illustrating the clutter spectral density.

Implementing the generation process mentioned in Section 4.8.2.1, sample data is generated for the same parameters to verify the generation steps. This is illustrated in Figure 5.41, where the output of the correlation and clutter power spectral density are compared. This method can then be expanded and applied to multiple range bins of input data in order to generate statistically correlated ocean wave data for multiple range bins.

The simulation of the data over time is implemented by applying a filter to data from a random complex number generator with

- mean $\mu = 0$,
- standard deviation $\sigma = 1$,
- ramp-up region to disregard filtered values of 1000 samples and
- and no correlation between I and Q based random values.

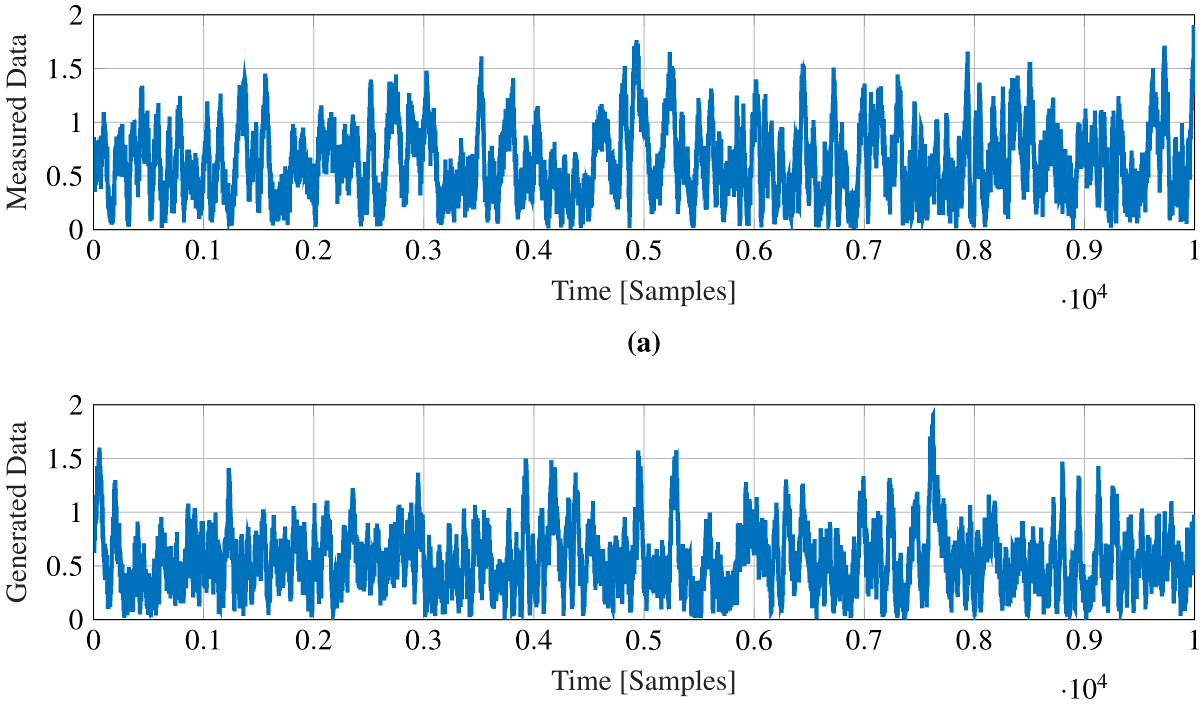


(a)

(b)

Figure 5.41. CFA16_002 PSD of generated and angular corrected data at range 3667 m (bin = 45), with (a) showing a detailed difference of the correlation around 0 lag and (b) illustrating the difference in clutter spectral density.

The random data is transformed as per Figure 4.31 producing a random sequence with given correlation and statistical distribution. Figure 5.42 compares the spectral content of the original input data to the generated output data in the time domain. This provides a mere optical comparison of the two datasets for illustrative purposes. From this figure the similarity between the generated and measured data is confirmed.



(a)

(b)

Figure 5.42. CFA16_002 (a) measured vs (b) generated data in the time domain for one range bin at range 3667 m (bin = 45).

5.7.2 Rayleigh based generation of 96 range bin data

Applying the foregoing method to all input range-bins, creates a set of output data with a given statistical and correlational similarity to the original input measurement. Figure 5.43 illustrates the dataset of the radar measured results that were angular corrected with the process described thus far. Figure 5.44 illustrates the generation of data using the distribution and correlation properties, obtained through parametrisation of the input data and the application of the random Gaussian number generator. The result is created for the same length of data, for purely comparative purposes. However, this process can easily be expanded for longer times by increasing the number of random generated input points, which the random function outputs.

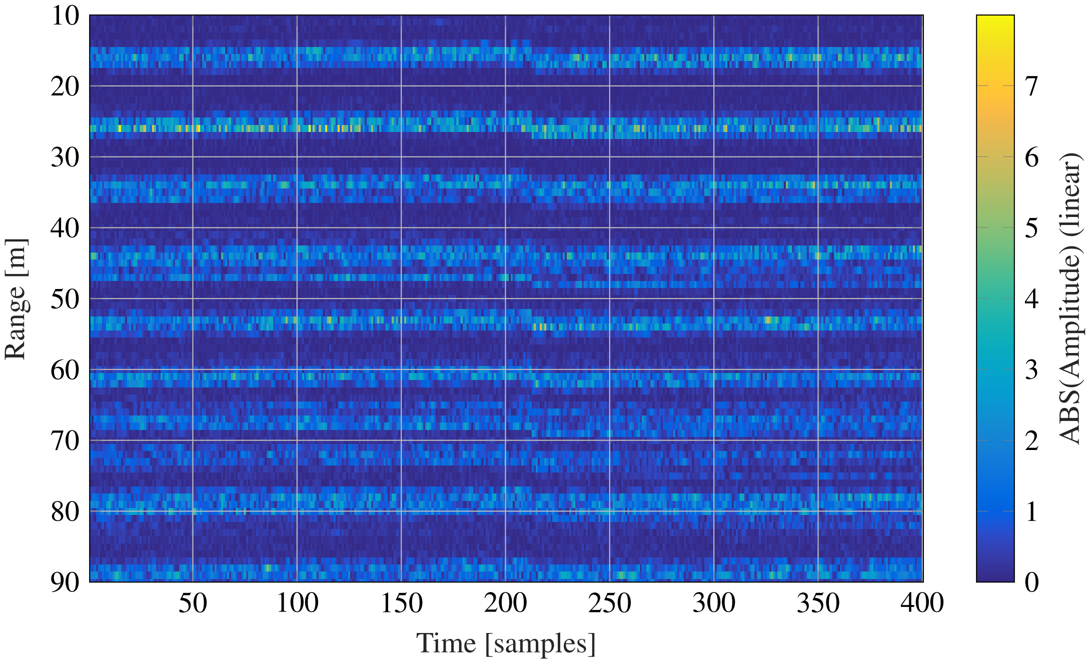


Figure 5.43. CFA16_002 range time plot of measured data after angle correction, illustrating correction step occurring at sample point 230.

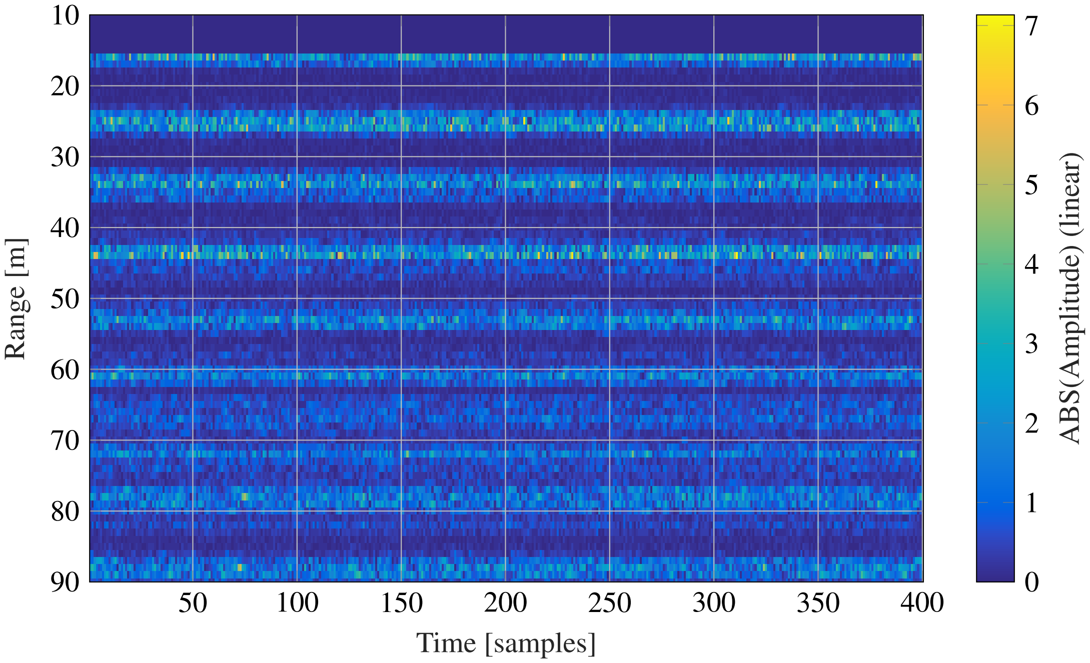


Figure 5.44. CFA16_002 range time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs (last step of angle counter correction not applied).

5.7.3 Comparison

In order to compare the generated results, the distributions and correlation of the generated data is compared with the original measured data. As the datasets are of the same length and range depth, these results can be compared directly to one another.

5.7.3.1 Distribution

The statistical distribution for 8 range bins (a total range of 120 m) illustrates that the Rayleigh distribution only partially fits the trough sections of the ocean wave well. Crest results, as discussed in the results above tend towards the K-distribution with a better performance. This is seen in Figures 5.45 to 5.47, in particular 5.47(a), illustrating a crest section of the ocean wave with poor performance towards the Rayleigh distribution.

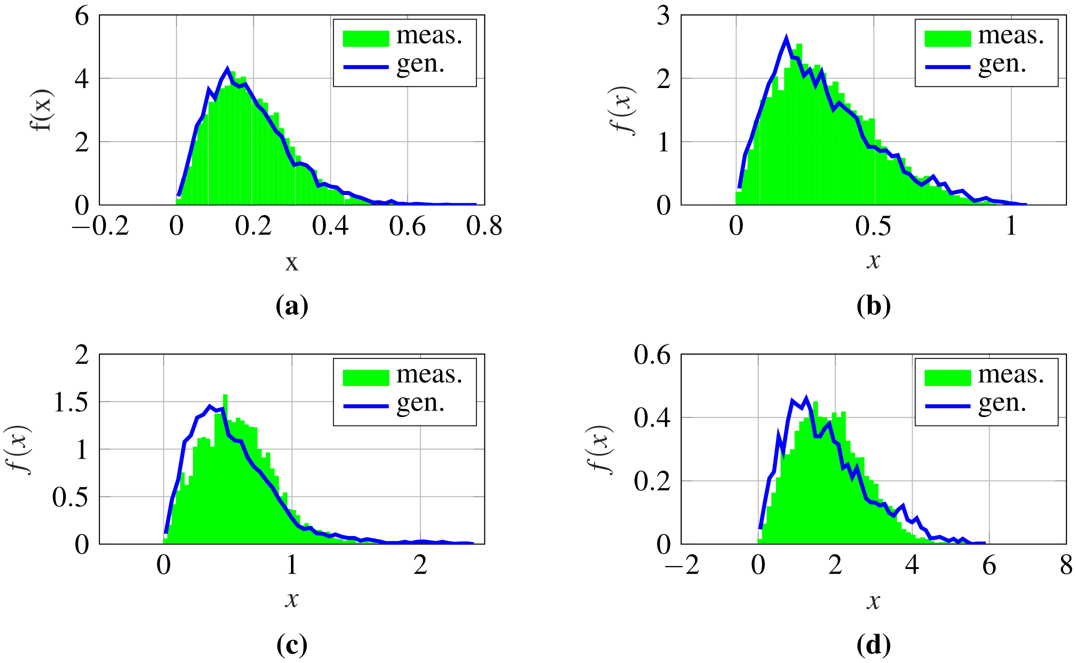


Figure 5.45. CFA16_002 distribution comparison at range bin (a) 40, (b) 41, (c) 42, and (d) 43, illustrating acceptable performance particularly in trough regions.

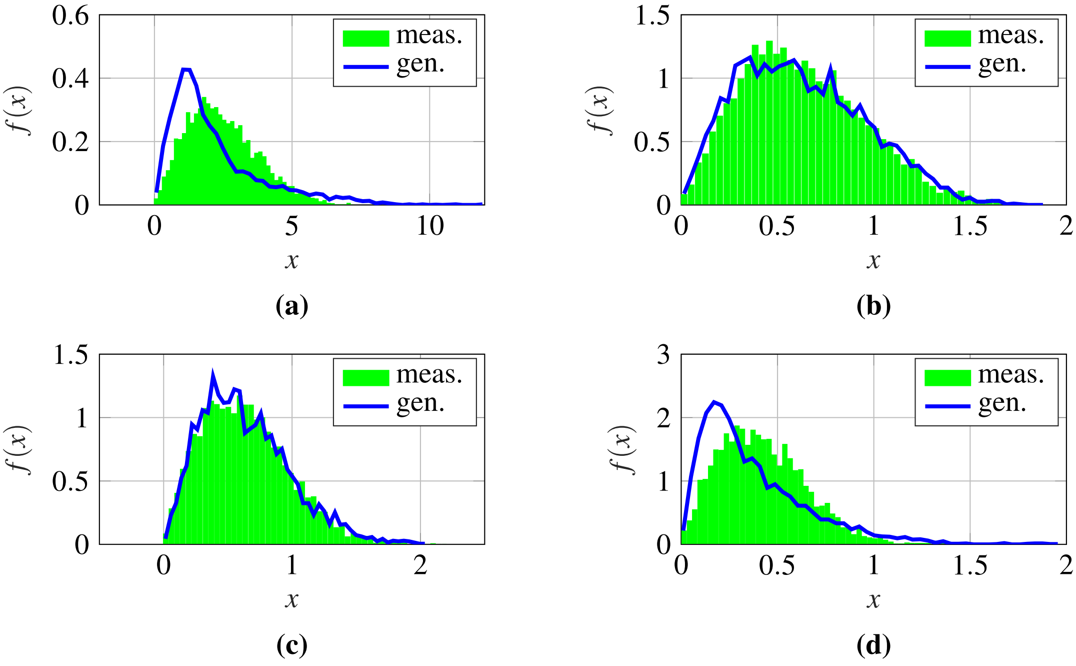


Figure 5.46. CFA16_002 distribution comparison at range bin (a) 44, (b) 45, (c) 46, and (d) 47, illustrating deviation for crests regions.

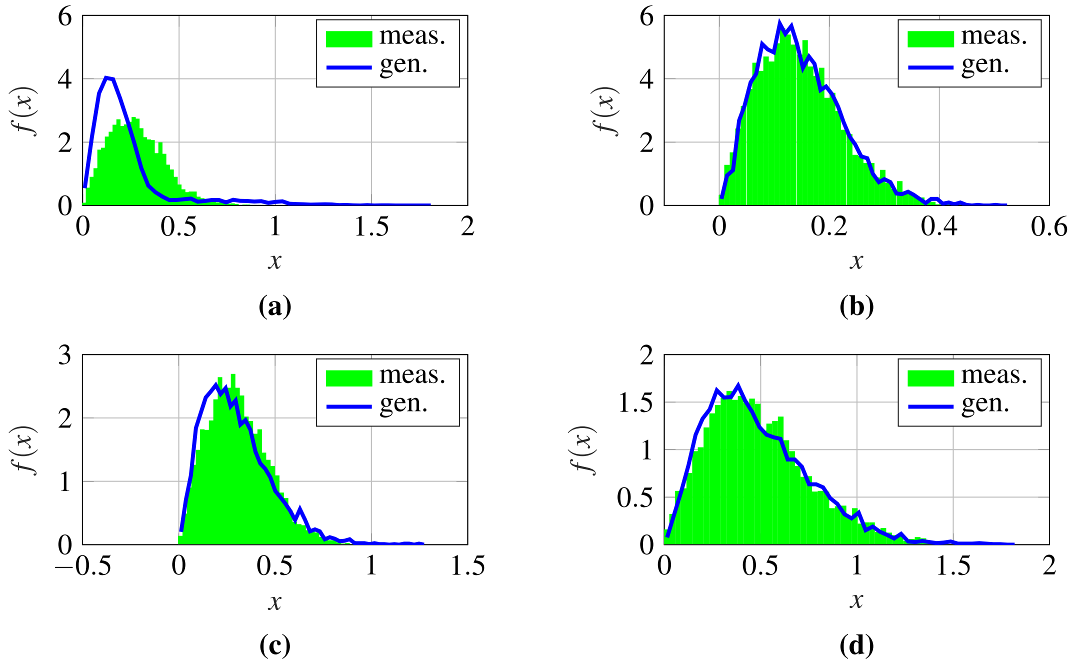


Figure 5.47. CFA16_002 distribution comparison at range bin (a) 48, (b) 49, (c) 50, and (d) 51, illustrating transition area from crest to trough regions of the ocean wave.

5.7.3.2 Correlation

The correlation shows that using the method described in Section 4.8.2.1 produces good results with close fitting and following outputs. This is seen through Figures 5.48 and 5.49.

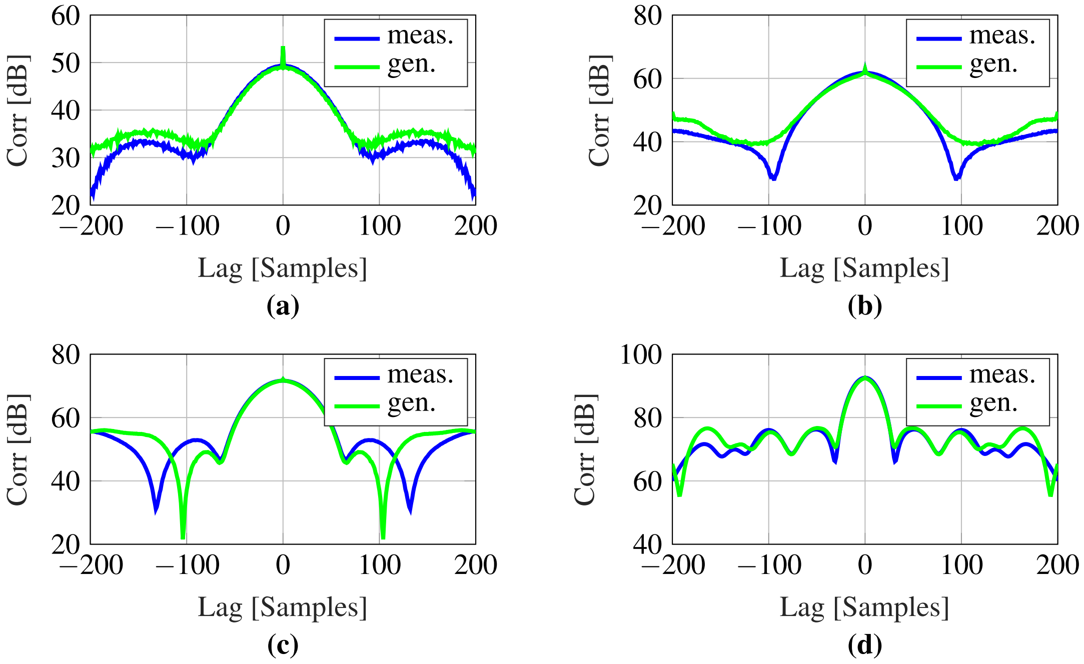
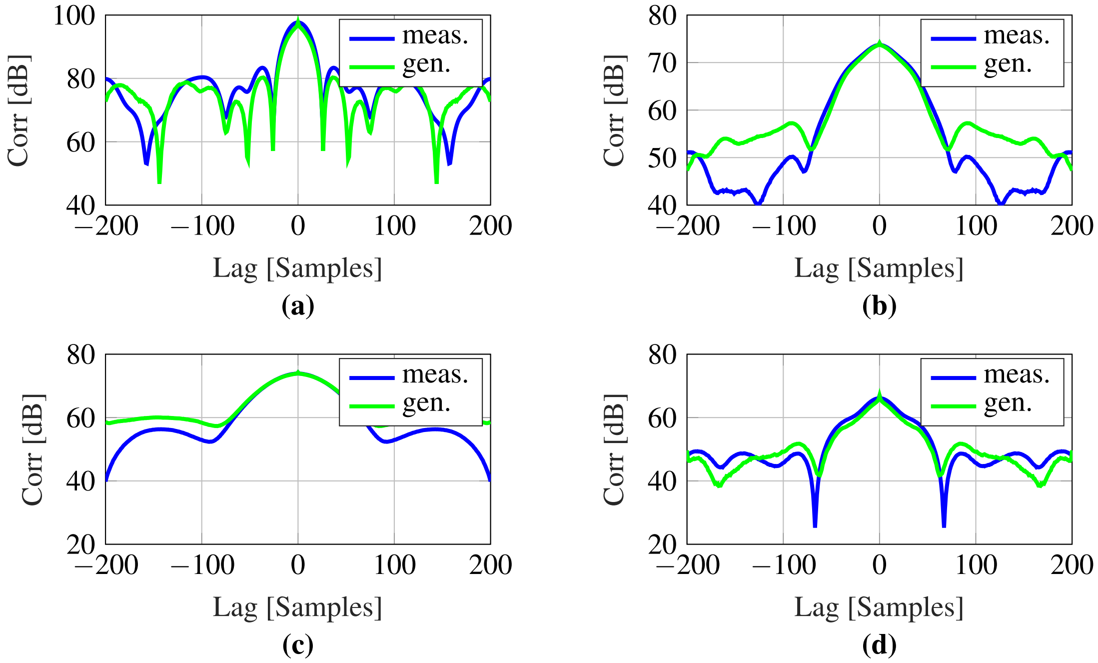


Figure 5.48. CFA16_002 correlation comparison at range bin (a) 40, (b) 41, (c) 42, and (d) 43, illustrating good performance.



(a) (b) (c) (d)

Figure 5.49. CFA16_002 correlation comparison at range bin (a) 44, (b) 45, (c) 46, and (d) 47, illustrating good performance.

5.8 DATASETS CONTAINING BOAT REFLECTIONS

Out of the six datasets analysed and processed, three contained a RIB crossing the breaking waves. These were datasets:

- TFC15_025,
- TFC15_028 and
- TFC15_029.

In order to identify the influence of the RIB on the parametrisation of the model, range-bins with strong boat reflections were considered. In order to analyse this occurrence, a section with strong reflections was identified. Strong reflections typically occur when the RIB crosses the ocean wave crest, as shown in Figure 5.50 between 7 and 8 seconds and 4680 range meters. This region is the highlighted area between the two black lines shown in Figure 5.50.

The region of interest, re-plotted in Figure 5.51, clearly shows the signal amplitude influence associated

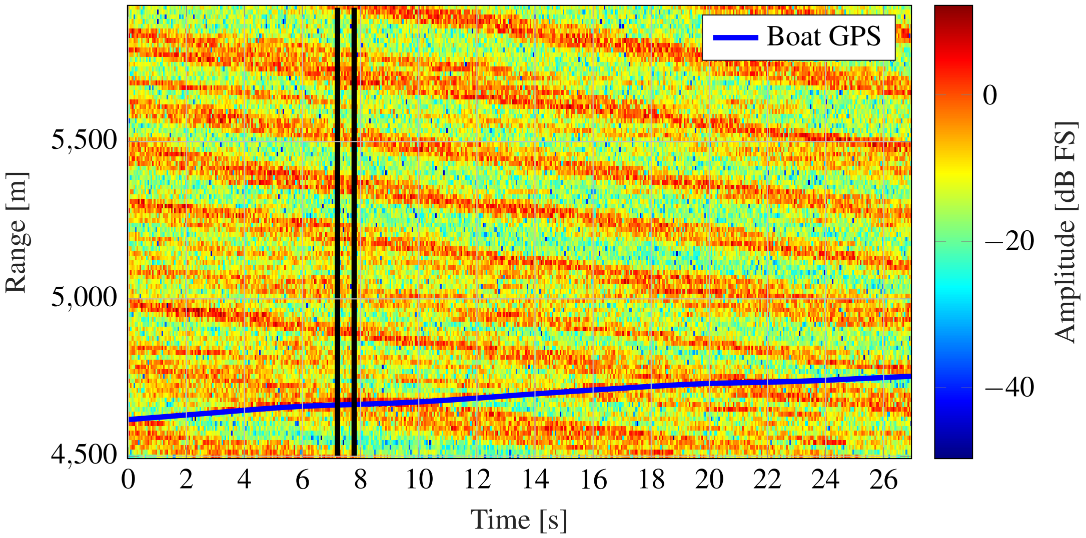


Figure 5.50. TFC15_025 Range time plot illustrating region with strong RIB reflections on ocean wave crest, where the section between the two black lines indicates the time region where a strong reflection occurs at 4680 m.

with the RIB crossing the ocean wave crest. The boat reflections can be identified in Figure 5.51 at 4680 m by the strong returns.

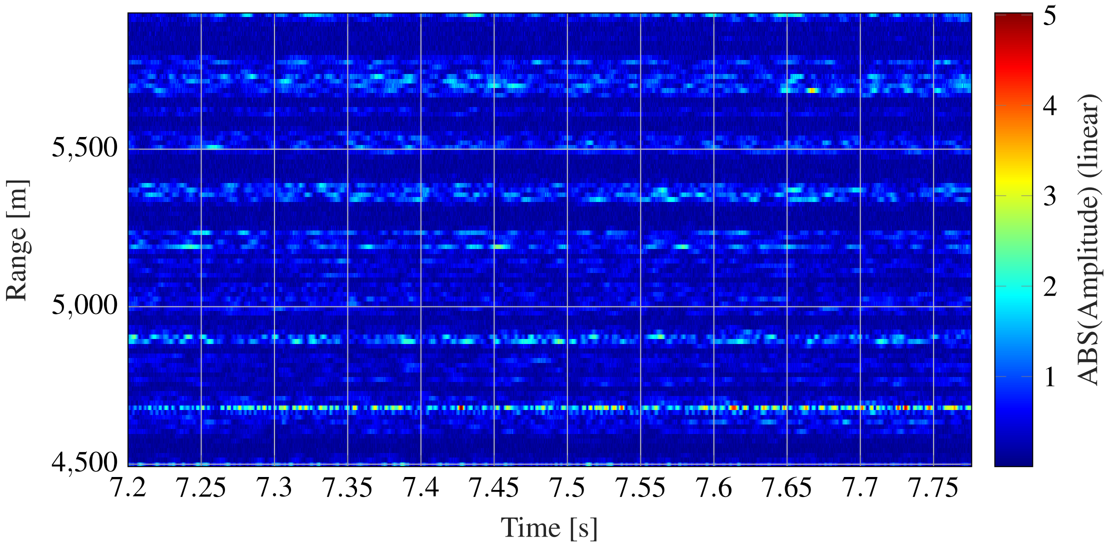


Figure 5.51. TFC15_025 input region illustrating strong boat reflections at a range of 4680 m and at a time of 7.2 to 7.8 seconds.

5.8.1 Effects

In the processing chain, the identification of the ocean wave crest angle is determined through the mean angles, which are obtained from the output of the Radon transform that is applied to the circular-masked input data. This input data matrix is shown in Figure 5.52 and again illustrates strong intensities at 4680 m which relate to the RIB.

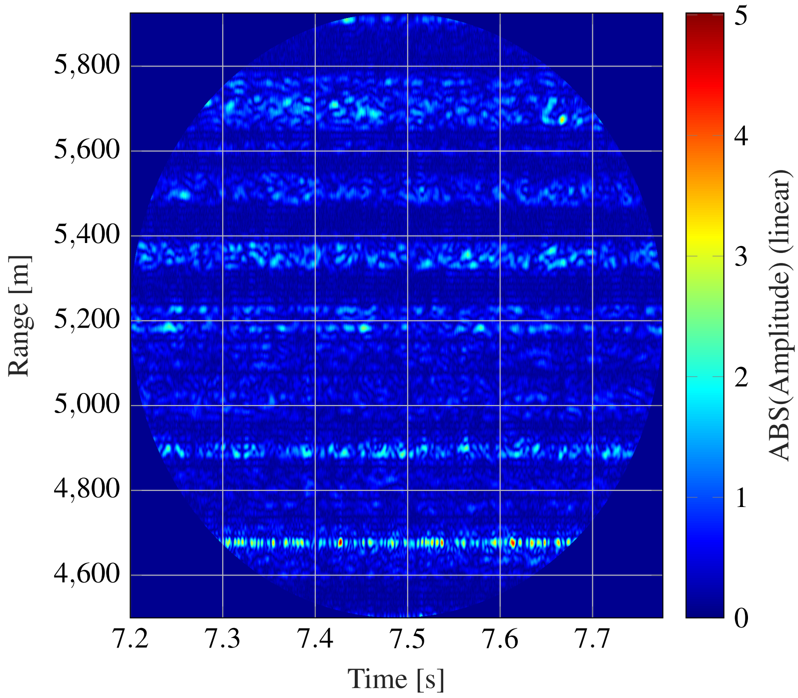


Figure 5.52. TFC15_025 circular input mask data for Radon transform illustrating strong signal returns from the RIB at 4680 m.

The output of the Radon transform indicates the boat structure through increased amplitude and width at the range of 4680 m. In order to elaborate the difference in amplitude from the transform's response, the squared output result is illustrated in Figure 5.53.

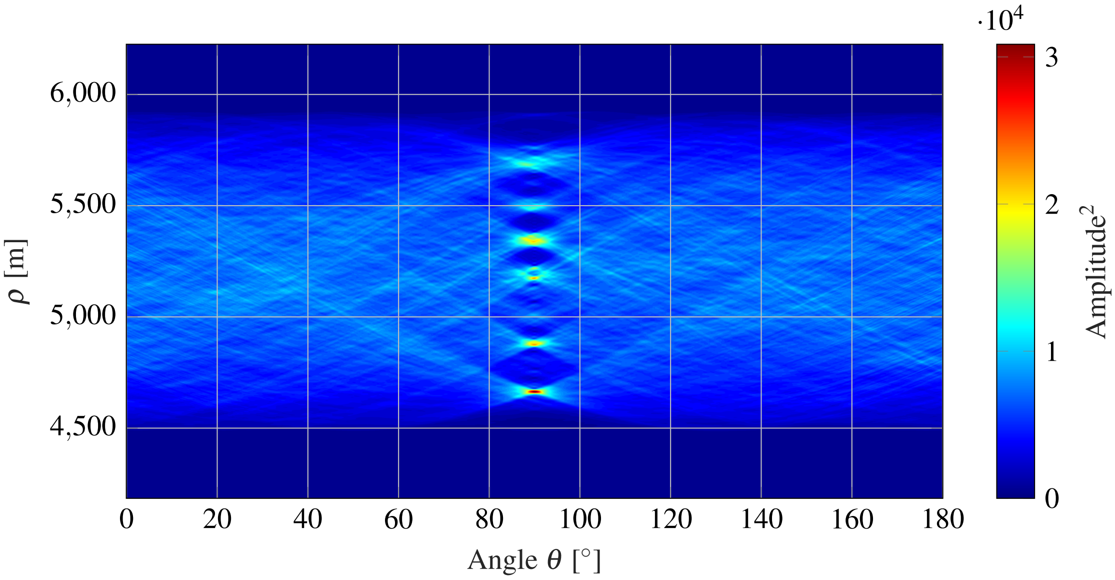


Figure 5.53. TFC15_025 Radon transform result of circular input matrix illustrating strong transform from RIB data at 4680 m and a wide angle spectrum.

Comparing the output angles, results show that the influence of the boat causes no significant change of the determined mean angle. This is supported by Figure 5.54. Further Figure 5.55 also illustrates that the largest deviation in the results is not related to the data obtained from the RIB at range 4680 m.

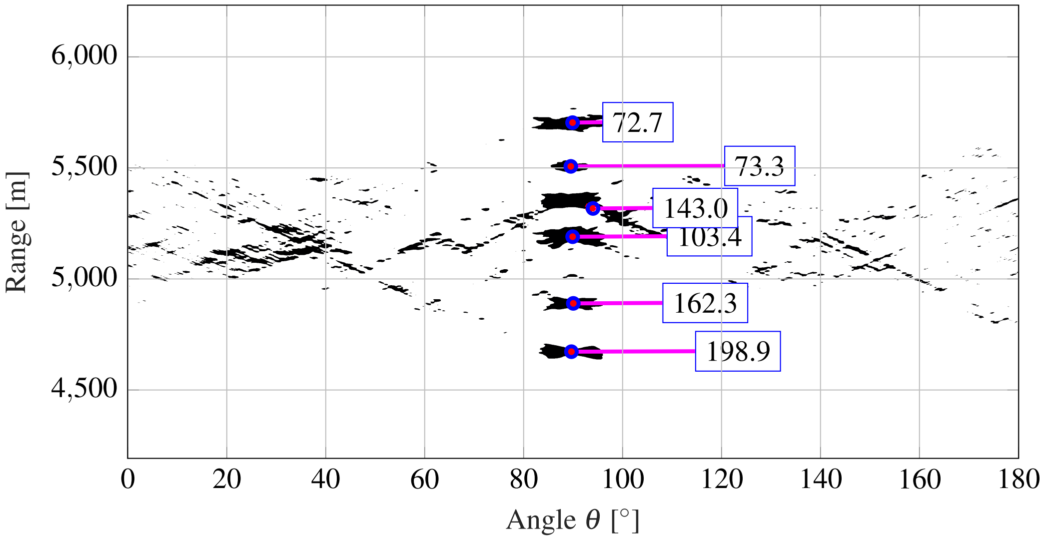


Figure 5.54. TFC15_025 illustrating results of angle determination illustrating the minimal effect of the RIB on the mean angle output as given by the detection at range 4680 m.

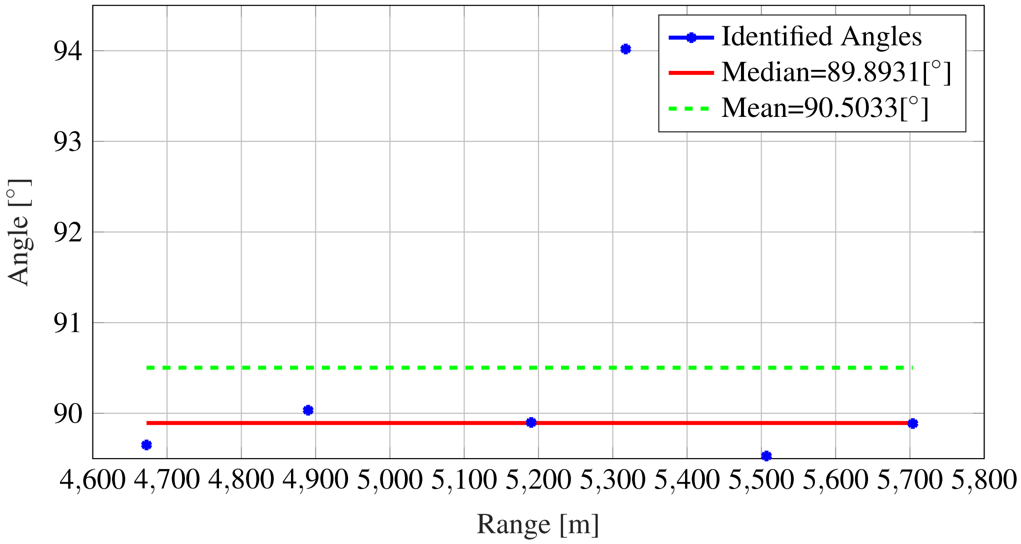


Figure 5.55. TFC15_025 illustrating results of angle determination illustrating the minimal effect of the RIB on the mean angle output as given by the detection at range 4680 m.

The result of the angular corrected ocean wave crest aligned to the *x*-axis is shown in Figure 5.56, illustrating good performance despite the input segment containing a strong reflection from a RIB crossing the ocean wave crest at a different approach angle to the ocean waves. This can be attributed to the small influence a single angle has on the algorithm together with the low impact a small RIB has on the measurement data. For larger boats (> 15 m) with stronger trails this may not be the case.

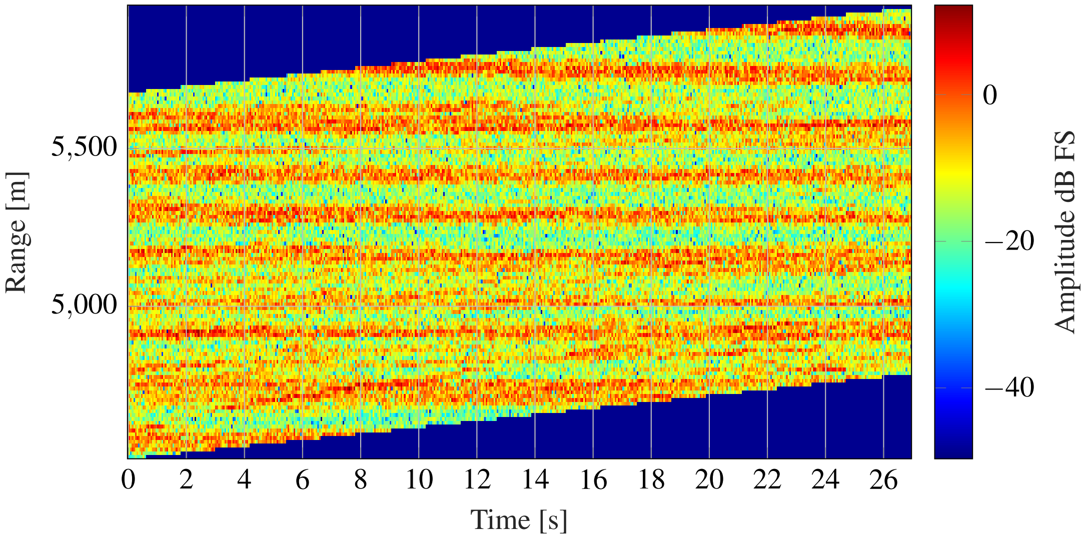
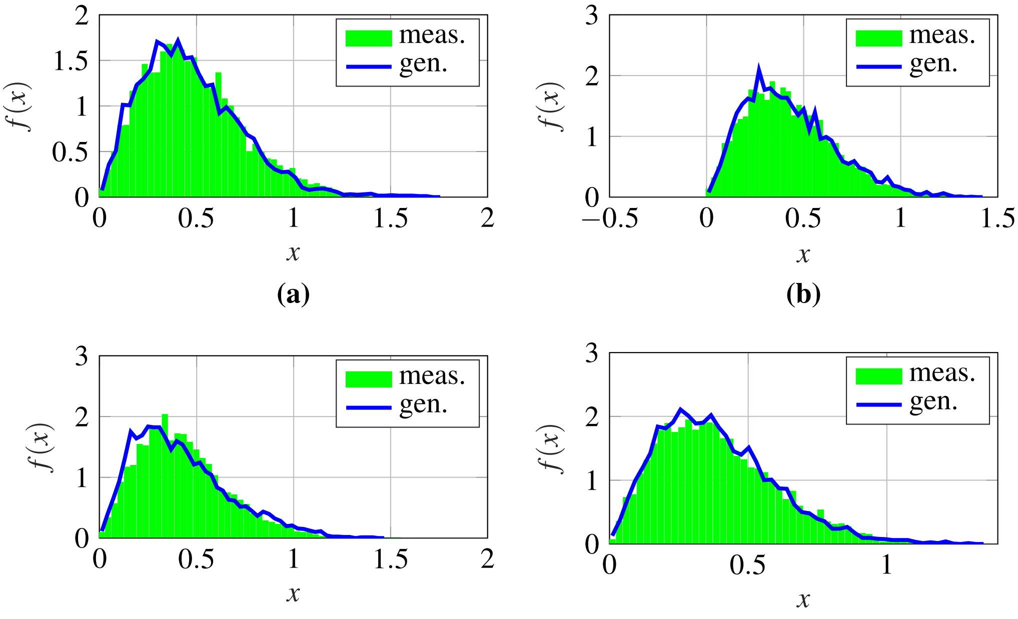


Figure 5.56. TFC15_025 output of wave-phase corrected data with *x*-axis aligned waves illustrating strong boat reflections at a range of 4680 m an at a time of 6 to 8 seconds.

In order to compare the distributions, the crest containing the boat was compared against similar ocean wave crest regions from the same angular corrected data matrix. These results are shown in Figures 5.57, 5.58 and 5.59.



(a) (b) (c) (d)

Figure 5.57. TFC15_025 Distribution with RIB crossing wave crests taken from range bins (a) 4665 m, (b) 4680 m, (c) 4695 m and (d) 4710 m.

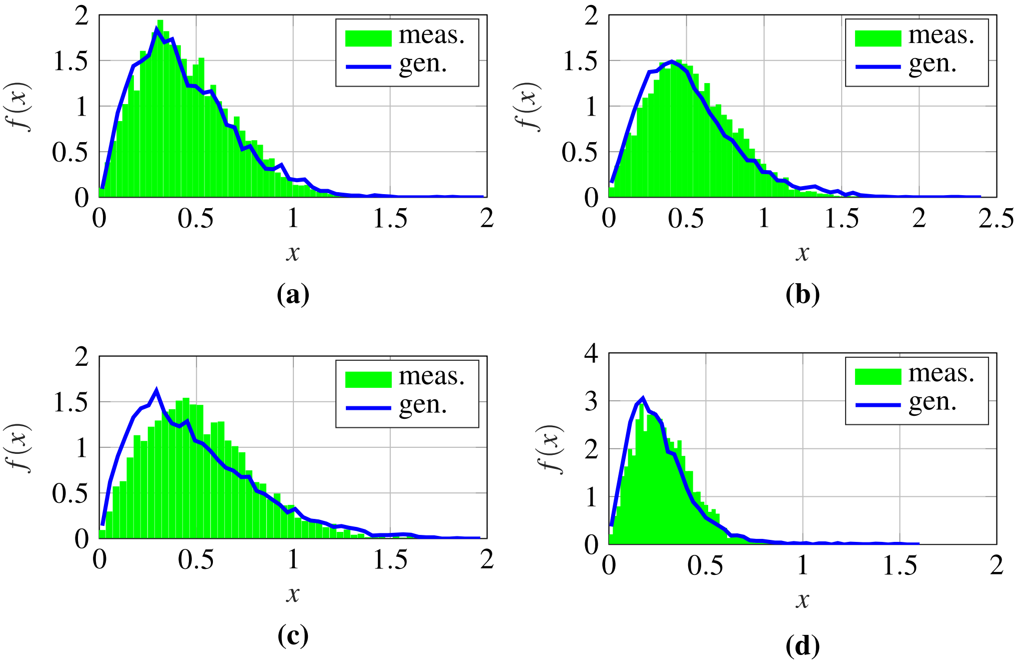


Figure 5.58. TFC15_025 Distribution of similar wave crest region containing no-boat data taken from range bins (a) 5010 m, (b) 5025 m, (c) 5040 m and (d) 5055 m.

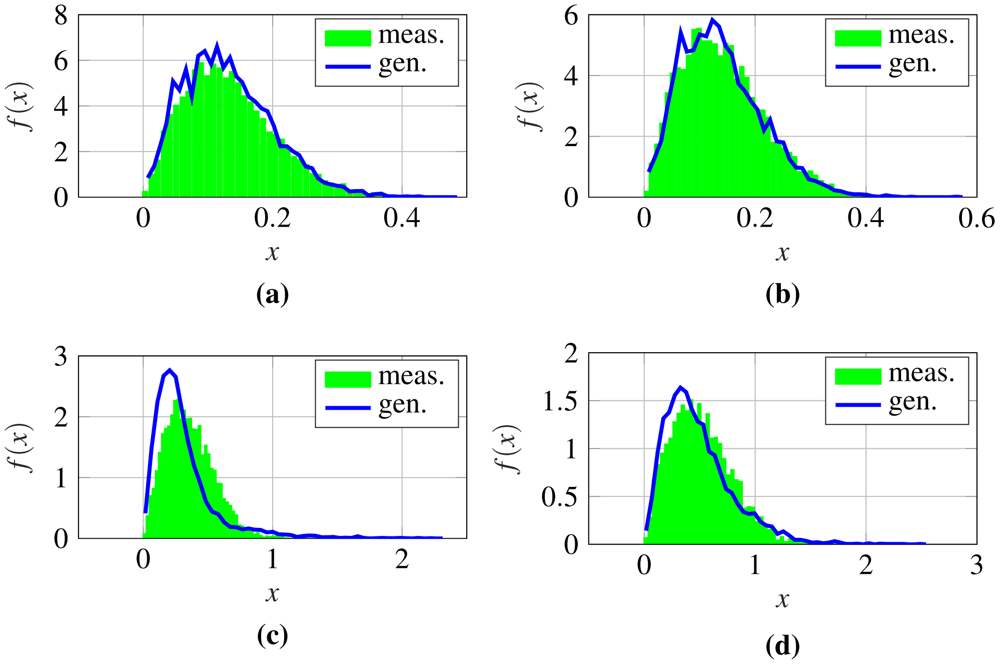


Figure 5.59. TFC15_025 Distribution of similar wave crest region containing no-boat data taken from range bins (a) 5310 m, (b) 5325 m, (c) 5340 m and (d) 5355 m.

These results illustrate the algorithm's robustness and stability for determining the approach angle, achieved through the parallel analysis of multiple ocean wave crests.

CHAPTER 6 CONCLUSION AND SUMMARY

6.1 SUMMARY

This dissertation introduced a statistical model for the simulation of periodic breaking sea waves in a littoral environment. In order to obtain a better understanding of periodic breaking waves in this environment, first a model was suggested that analyses the different segments of the ocean wave against the commonly used distributions for sea clutter. These distributions included the log-normal, the Rayleigh and the K-distribution. The model analysed existing low grazing angle, X-band radar data obtained from the CSIR, which was measured off the coast of Cape Town, South Africa.

Using this data, the parametrisations of the suggested model was possible. The analysis of the correlation between wave crests was also included in the model in order to accommodate the periodic nature of the parallel waves, typically seen in the measurements. This allowed for the generation of statistically similar periodic breaking wave data, which in turn was validated against the measurement data in terms of correlational and distributional properties. This elaborated method was applied to six datasets obtained from the CSIR, which demonstrated the robustness and flexibility of the suggested model.

The core of the analysis section of the model, is the alignment of the ocean wave crests to the x-axis of the data matrix. This is obtained through an adapted circular masked use of the Radon transform, resulting in an accurate determination of the approaching angle of the ocean waves with respect to the stationary radar position. Using this angle, the measured radar data can be shifted, so that the ocean wave crests are parallel to the x-axis. This step allows for the statistical analysis of individual segments of a given ocean wave, effectively statistically differentiating regions from the trough through to the crest of the wave.

This information was then applied to parametrise a simulation model, which can generate data with similar distributional properties generated from a standard Gaussian random noise function. Furthermore, correlation in fast time was considered in order to capture the periodic nature of the parallel wave crests. This is also captured and successfully implemented in the generation model. The Rayleigh distribution provides accurate statistical results, which together with the correlation properties, accurately generate data from a Gaussian random generator that is used as an input for the transformational filter.

It is shown that the influence of a RIB does not alter the output of the model. This is due to the minor impact the measurements have on the returns, further strengthened by the parallel detection scheme which offers robustness against outliers.

The suggested model provides insights into the statistics of the individual segments of the ocean wave and allows the simulation of similar data. Simulation is essential for further development of detection methods needed for such littoral ocean environments. This is critical for the implementation of low cost policing in these regions and could aid in the reduction of illegal fishing activities as well as improved border control.

6.2 CONCLUSION OF RESULTS

The results showed that measured radar data can be simulated for a littoral environment through parametrisation of the suggested model. The statistical similarity was confirmed through the comparison of generated and measured data and the visual use of the q-q plots. These plots were utilised to evaluate the accuracy of the applied distribution for the given segment of the ocean wave. This also provided insights, which support the theory that the returns of the ocean wave are based on reflections from the ocean wave crests. These crests tend to follow the Rayleigh and K-distribution. Troughs however, are not seen by the radar as they are obscured by the ocean wave crests. Thus trough data tend to follow the Gaussian distribution.

The generation model showed that it can accurately reproduce the data in an autonomous way for a variety of input data. The computational overhead is primarily related to the complexity and resolution of the Radon transform used for the approaching angle determination process. The advantages of the suggested model are the handling of periodic breaking waves through the treatment of individual segments of the ocean wave with different distributions. This also leads to the advantage of individual wave identification, which could in future be used to track individual waves and in turn be used for the detection of anomalies or deviations from common ocean wave movement.

REFERENCES

[1] M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar*. Raleigh, NC.: SciTech Publishing, Inc., 2010.

[2] S. Panagopoulos and J. J. Soraghan, “Small target detection in sea clutter,” *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1355–1361, 2004.

[3] H. Wensink, “On parametric detection of small targets in sea clutter,” in *Proc. Third Int. Conf. Inf. Fusion*, pp. MOC1/17–MOC1/24 vol.1, IEEE, 2000.

[4] M. Strempel, J. De Villiers, J. E. Cilliers, and A. McDonald, “Distribution analysis of segmented wave sea clutter in littoral environments,” in *2015 IEEE Radar Conf.*, pp. 133–138, IEEE, Oct. 2015.

[5] J. Hu, W. Tung, and J. Gao, “Detection of Low Observable Targets Within Sea Clutter by Structure Function Based Multifractal Analysis,” *IEEE Trans. Antennas Propag.*, vol. 54, pp. 136–143, Jan. 2006.

[6] Department of Defence, “South African Defence Review 2015,” tech. rep., Department of Defence South Africa, 2015.

[7] Department of Defence, “South African Defence Review 2012,” tech. rep., Department of Defence South Africa, 2012.

[8] L. van Niekerk, *Pocket Guide to South Africa 2011/12*. Government Communications (GCIS), 9th ed., 2011.

[9] F. Onuoha, "Sea piracy and maritime security in the Horn of Africa: The Somali coast and Gulf of Aden in perspective," *African Secur. Rev.*, vol. 18, pp. 31–44, Sept. 2009.

[10] M. Hauck, "A case study of abalone poaching in South Africa and its impact on fisheries management," *ICES J. Mar. Sci.*, vol. 56, pp. 1024–1032, Dec. 1999.

[11] P. Herselman and C. Baker, "Analysis of calibrated sea clutter and boat reflectivity data at C- and X-band in south african coastal waters," in *IET Int. Conf. Radar Syst. 2007*, vol. 27, pp. 140–140, IET, 2007.

[12] J. Carretero-Moya, J. Gismero-Menoyo, A. Asensio-Lopez, and A. Blanco-Del-Campo, "Small-Target Detection in High-Resolution Heterogeneous Sea-Clutter: An Empirical Analysis," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, pp. 1880–1898, Jul. 2011.

[13] H. Chan, "Radar sea-clutter at low grazing angles," *IEE Proc. F Radar Signal Process.*, vol. 137, no. 2, p. 102, 1990.

[14] C. Baker, "K-distributed coherent sea clutter," *IEE Proc. F Radar Signal Process.*, vol. 138, no. 2, p. 89, 1991.

[15] A. Farina, F. Gini, M. Greco, and L. Verrazzani, "High resolution sea clutter data: statistical analysis of recorded live data," *IEE Proc. - Radar, Sonar Navig.*, vol. 144, no. 3, p. 121, 1997.

[16] J. Hu, J. Gao, R. S. Lynch, and G. Chen, "Multi-scale Modeling Approach for Detecting LowObservable Targets within Sea Clutter," in *2008 IEEE Aerosp. Conf.*, pp. 1–7, IEEE, Mar. 2008.

[17] F. Totir, E. Radoi, L. Anton, C. Ioana, A. Serbanescu, and S. Stankovic, "Advanced Sea Clutter Models and their Usefulness for Target Detection," *MTA Rev.*, vol. 18, no. 3, pp. 257–272, 2008.

[18] Tough, Watts, and Ward, *Sea Clutter: Scattering, the K Distribution and Radar Performance*. Institution of Engineering and Technology, 2nd ed., Apr. 2013.

[19] S. Watts, “Radar sea clutter: Recent progress and future challenges,” in *2008 Int. Conf. Radar*, pp. 10–16, IEEE, Sept. 2008.

[20] E. Jakeman and P. Pusey, “A model for non-Rayleigh sea echo,” *IEEE Trans. Antennas Propag.*, vol. 24, pp. 806–814, Nov. 1976.

[21] K. Ward, C. Baker, and S. Watts, “Maritime surveillance radar. Part 1: Radar scattering from the ocean surface,” *IEE Proc. F Radar Signal Process.*, vol. 137, no. 2, p. 51, 1990.

[22] S. Watts, “A new method for the simulation of coherent sea clutter,” in *2011 IEEE RadarCon*, pp. 052–057, IEEE, May. 2011.

[23] S. Watts, “Modeling and Simulation of Coherent Sea Clutter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, pp. 3303–3317, Oct. 2012.

[24] H. Leung, “Experimental modeling of electromagnetic wave scattering from an ocean surface based on chaotic theory,” *Chaos, Solitons & Fractals*, vol. 2, pp. 25–43, Jan. 1992.

[25] Nan Xie, H. Leung, and Hing Chan, “A multiple-model prediction approach for sea clutter modeling,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, pp. 1491–1502, Jun. 2003.

[26] H. Leung, N. Dubash, and N. Xie, “Detection of small objects in clutter using a GA-RBF neural network,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 1, pp. 98–118, 2002.

[27] J. Gao and K. Yao, “Multifractal features of sea clutter,” in *Proc. 2002 IEEE Radar Conf. (IEEE Cat. No.02CH37322)*, no. 1, pp. 500–505, IEEE, 2002.

[28] L. Yujie, W. Wenguang, and S. Jinping, “Research of Small Target Detection within Sea Clutter Based on Chaos,” in *2009 Int. Conf. Environ. Sci. Inf. Appl. Technol.*, pp. 469–472, IEEE, Jul. 2009.

[29] S. Haykin and S. Puthusserypady, “Chaotic dynamics of sea clutter,” *Chaos An Interdiscip. J. Nonlinear Sci.*, vol. 7, pp. 777–802, Dec. 1997.

[30] F. Berizzi, E. Dalle Mese, and G. Pinelli, “One-dimensional fractal model of the sea surface,” *IEE Proc. - Radar, Sonar Navig.*, vol. 146, no. 1, p. 55, 1999.

[31] S. Haykin, R. Bakker, and B. Currie, “Uncovering nonlinear dynamics-the case study of sea clutter,” *Proc. IEEE*, vol. 90, pp. 860–881, May. 2002.

[32] X. Wang, J. Liu, and H. Liu, “Small Target Detection in Sea Clutter Based on Doppler Spectrum Features,” in *2006 CIE Int. Conf. Radar*, pp. 1–4, IEEE, Oct. 2006.

[33] S. Watts, “Radar detection prediction in sea clutter using the compound K-distribution model,” *IEE Proc. F Commun. Radar Signal Process.*, vol. 132, no. 7, p. 613, 1985.

[34] B. Mahafza, *Radar systems analysis and design using MATLAB*. Chapman & Hall, 2000.

[35] V. K. Rohatgi and A. K. M. E. Saleh, *An Introduction to Probability and Statistics*. Wiley Series in Probability and Statistics, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2nd ed., Sept. 2000.

[36] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. Mc Graw Hill, 4th ed., 2002.

[37] B. F. Ginos, *Parameter Estimation for the Lognormal Distribution*. Master of science, Brigham Young, 2009.

[38] I. Antipov, “Simulation of Sea Clutter Returns,” tech. rep., DSTO Electronic and Surveillance Research Laboratory, 1998.

[39] R. Raghavan, “A method for estimating parameters of K-distributed clutter,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, pp. 238–246, Mar. 1991.

[40] D. R. Iskander, A. M. Zoubir, and B. Boashash, "A Method for Estimating the Distribution Parameters of the K distribution," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1147–1151, 1999.

[41] D. R. Iskander and A. M. Zoubir, "Estimating the parameters of the K-distribution using the ML/MOM approach," in *IEEE TENCON - Digit. Signal Process. Appl*, pp. 769–774, IEEE, 1996.

[42] D. Iskander and A. Zoubir, "Estimation of the parameters of the K-distribution using higher order and fractional moments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 4, pp. 1453–1457, 1999.

[43] W. Roberts and S. Furui, "Maximum likelihood estimation of K-distribution parameters via the expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3303–3306, 2000.

[44] R. M. Neal, "Slice sampling: Rejoinder," *Ann. Stat.*, vol. 31, no. 3, pp. 758–767, 2003.

[45] L. Martino and J. Míguez, "Generalized rejection sampling schemes and applications in signal processing," *Signal Processing*, vol. 90, no. 11, pp. 2981–2995, 2010.

[46] G. Casella, C. P. Robert, and M. T. Wells, "Generalized Accept-Reject sampling schemes," *Lecture Notes Monograph Series*, vol. 45, pp. 342–347, 2004.

[47] A. C. Sauvain-dugerdil, B. Gakou, F. Berthé, A. W. Dieng, G. Ritschard, M. Lerch, R. J. Hyndman, and Y. Fan, "Computing Sample Quantiles in Statistical Packages," *Am. Stat.*, vol. 50, no. 4, pp. 361–365, 1996.

[48] Carsten Hoilund, *The Radon Transform*. PhD thesis, Aalborg University, 2007.

[49] X. Li, S. Zhang, X. Pan, P. Dale, and R. Cropp, "Straight road edge detection from high-resolution remote sensing images based on the ridgelet transform with the revised parallel-beam Radon transform," *Int. J. Remote Sens.*, vol. 31, pp. 5041–5059, Oct. 2010.

[50] P. Toft and J. Sørensen, *The Radon transform - theory and implementation.* PhD thesis, Technical University of Denmark, 1996.

[51] R. Pourreza, T. Banaee, H. Pourreza, and R. D. Kakhki, “A radon transform based approach for extraction of blood vessels in conjunctival images,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5317 LNAI, no. June 2014, pp. 948–956, 2008.

[52] D. Young and N. Beaulieu, “The generation of correlated Rayleigh random variates by inverse discrete Fourier transform,” *IEEE Trans. Commun.*, vol. 48, pp. 1114–1127, Jul. 2000.

[53] J. I. Smith, “A Computer Generated Multipath Fading Simulation for Mobile Radio,” *IEEE Trans. Veh. Technol.*, vol. 24, no. 3, pp. 39–40, 1975.

[54] N. Broste, “Digital generation of random sequences,” *IEEE Trans. Automat. Contr.*, vol. 16, pp. 213–214, Apr. 1971.

ADDENDUM A FURTHER RESULTS

The results for the remaining datasets, which were not shown in Chapter 5, are illustrated in this appendix.

A.1 DATASET OVERVIEW

An overview of the six datasets is provided in the Table 5.1 and is reproduced in Table A.1.

Table A.1. Dataset overview reproduced from Table 5.1.

Dataset	**Type**	**TX Freq.**	**PRF**	θ_g
CFA16_002	sea clutter	6.9 GHz	5 kHz	$0.853^\circ - 1.27^\circ$
CFA16_003	sea clutter	6.9 GHz	5 kHz	$0.427^\circ - 0.525^\circ$
CFA16_004	sea clutter	6.9 GHz	5 kHz	$0.427^\circ - 0.525^\circ$
TFC15_025	sea clutter + RIB	9 GHz	2.5 kHz	$0.628^\circ - 0.838^\circ$
TFC15_028	sea clutter + RIB	9 GHz	2.5 kHz	$0.492^\circ - 0.619^\circ$
TFC15_029	sea clutter + RIB	9 GHz	5 kHz	$0.445^\circ - 0.551^\circ$

Results for the datasets not discussed in Chapter 5 are presented. All key figures, required for the model, are presented and illustrate the robustness of the implemented model. Reference to these figures is made in the results and discussion chapter at the respective sections.

A.2 RESULTS DATASET 2: CFA16_003

A.2.1 Raw data

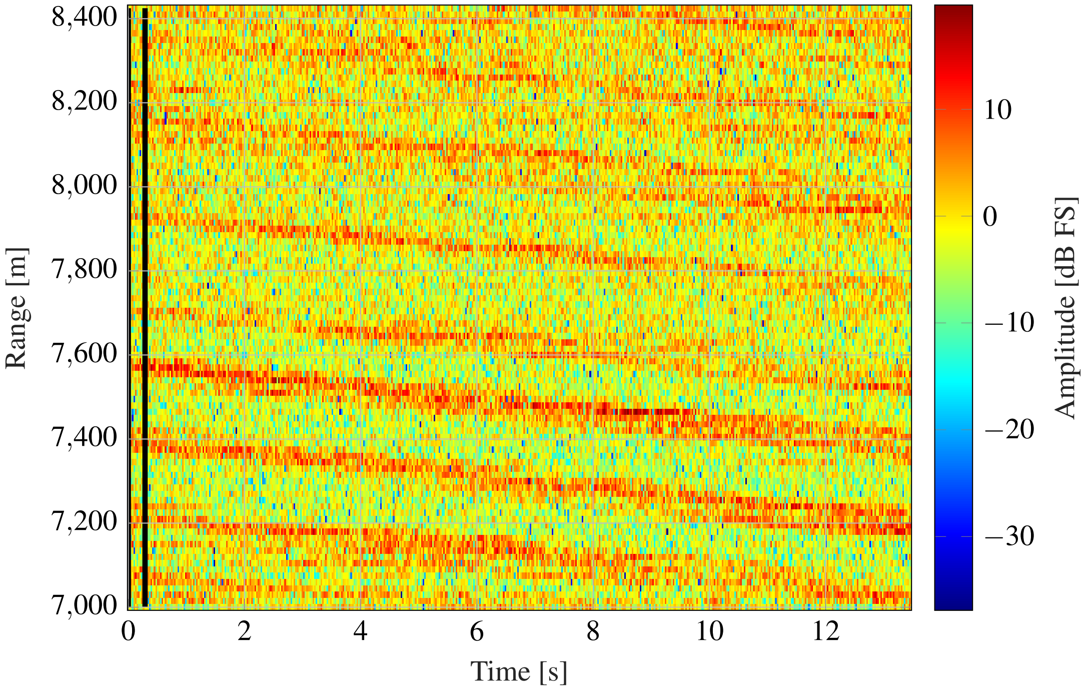
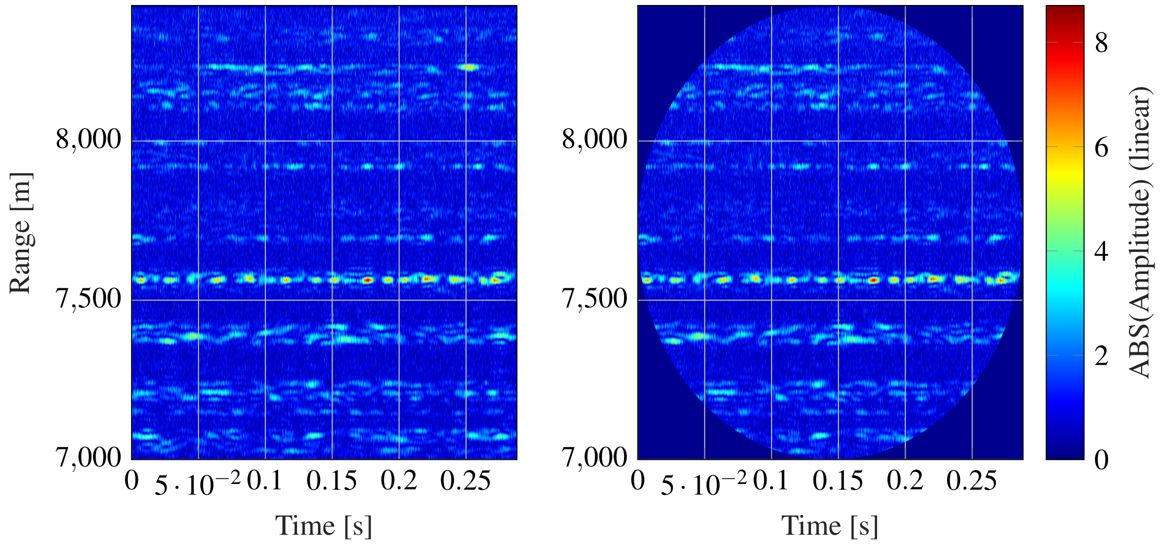


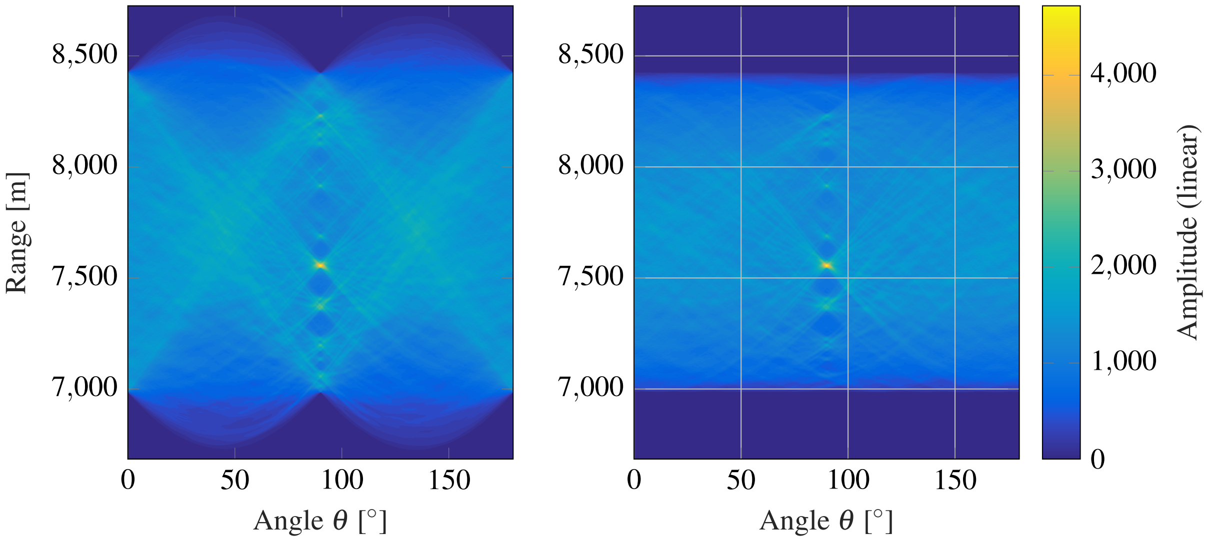
Figure A.1. CFA16_003 range-time plot of 13 s of measured data given in dB full scale, where the black vertical lines indicate the region used for angle determination.

A.2.2 Radon transform



(a) **(b)**

Figure A.2. CFA16_003 non-circular vs circular input processing matrices for the Radon transform, indicating the null data points on the outside of the circular mask.



(a) **(b)**

Figure A.3. CFA16_003 Radon transform result illustrating the output for the (a) non-circular and (b) circular input matrix.

A.2.3 Angle calculation

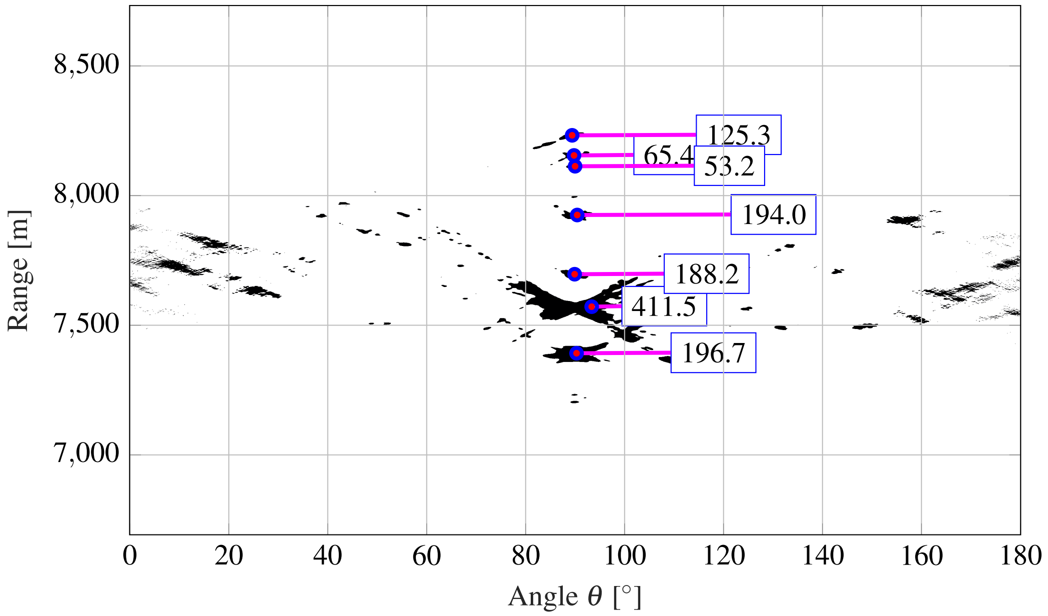


Figure A.4. CFA16_003 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

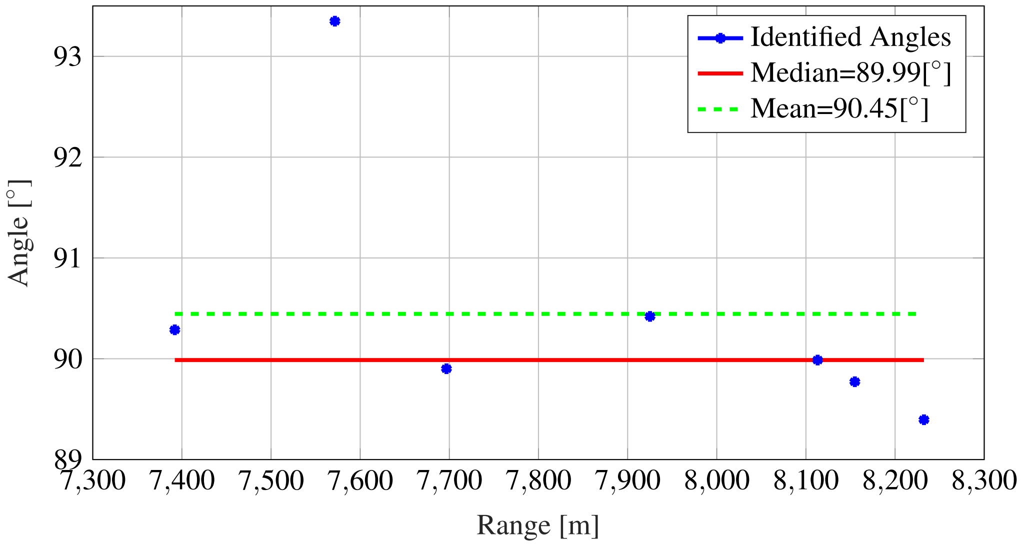


Figure A.5. CFA16_003 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

A.2.4 Wave phase alignment

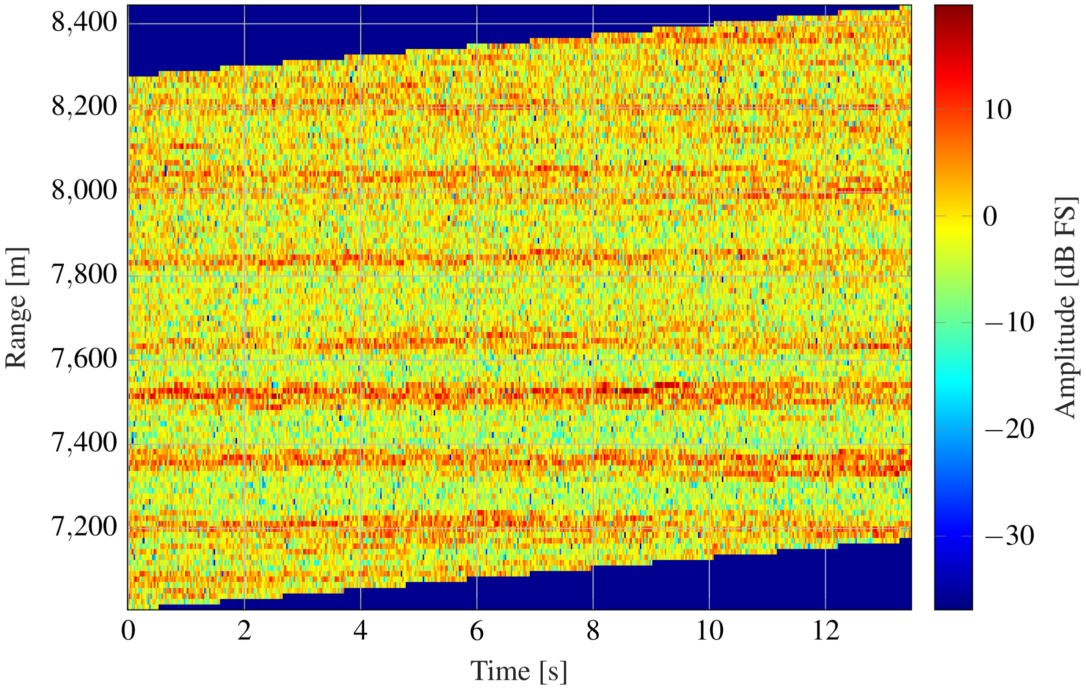
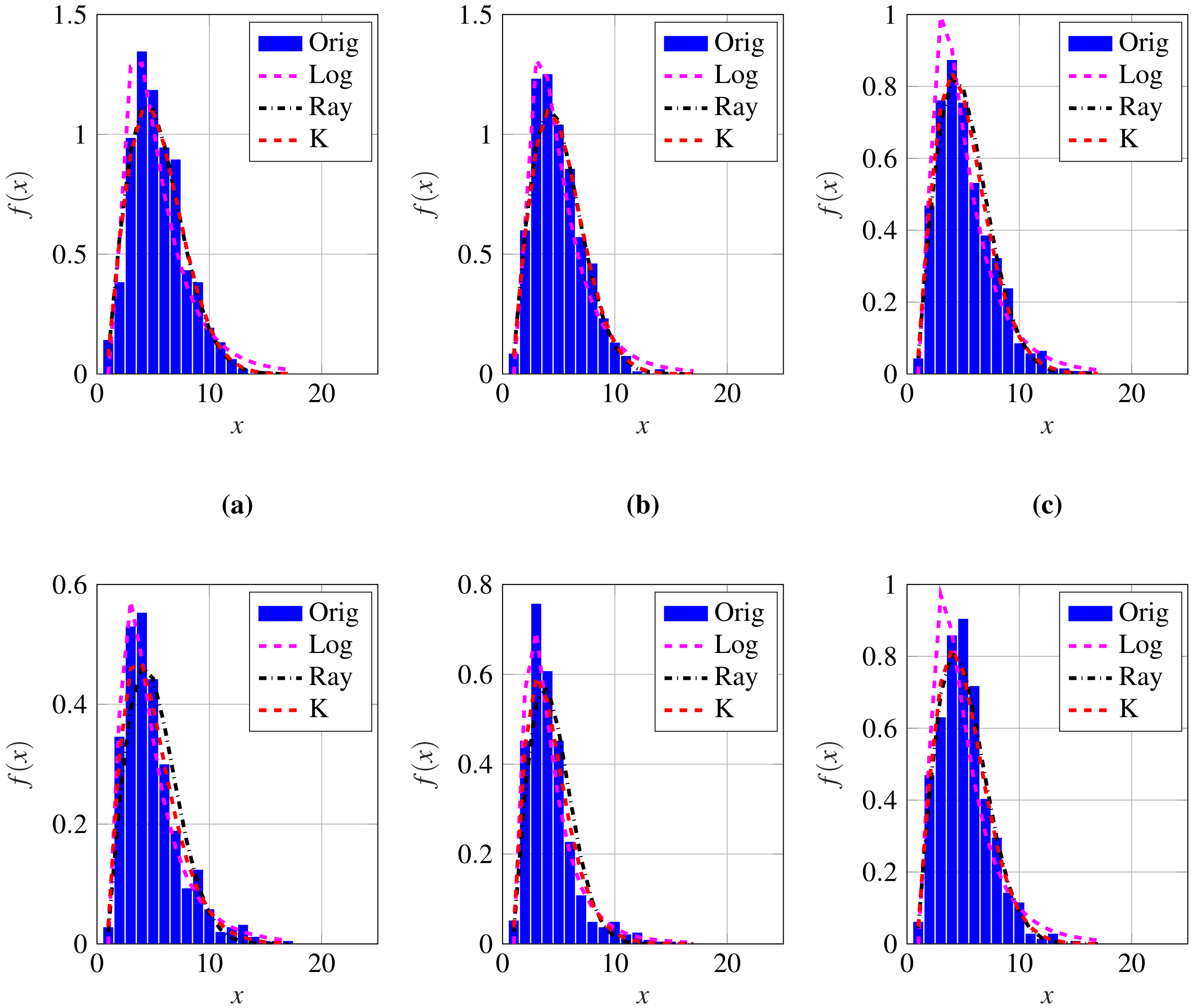


Figure A.6. CFA16_003 output of wave-phase corrected data with *x*-axis aligned waves illustrating clear steps.

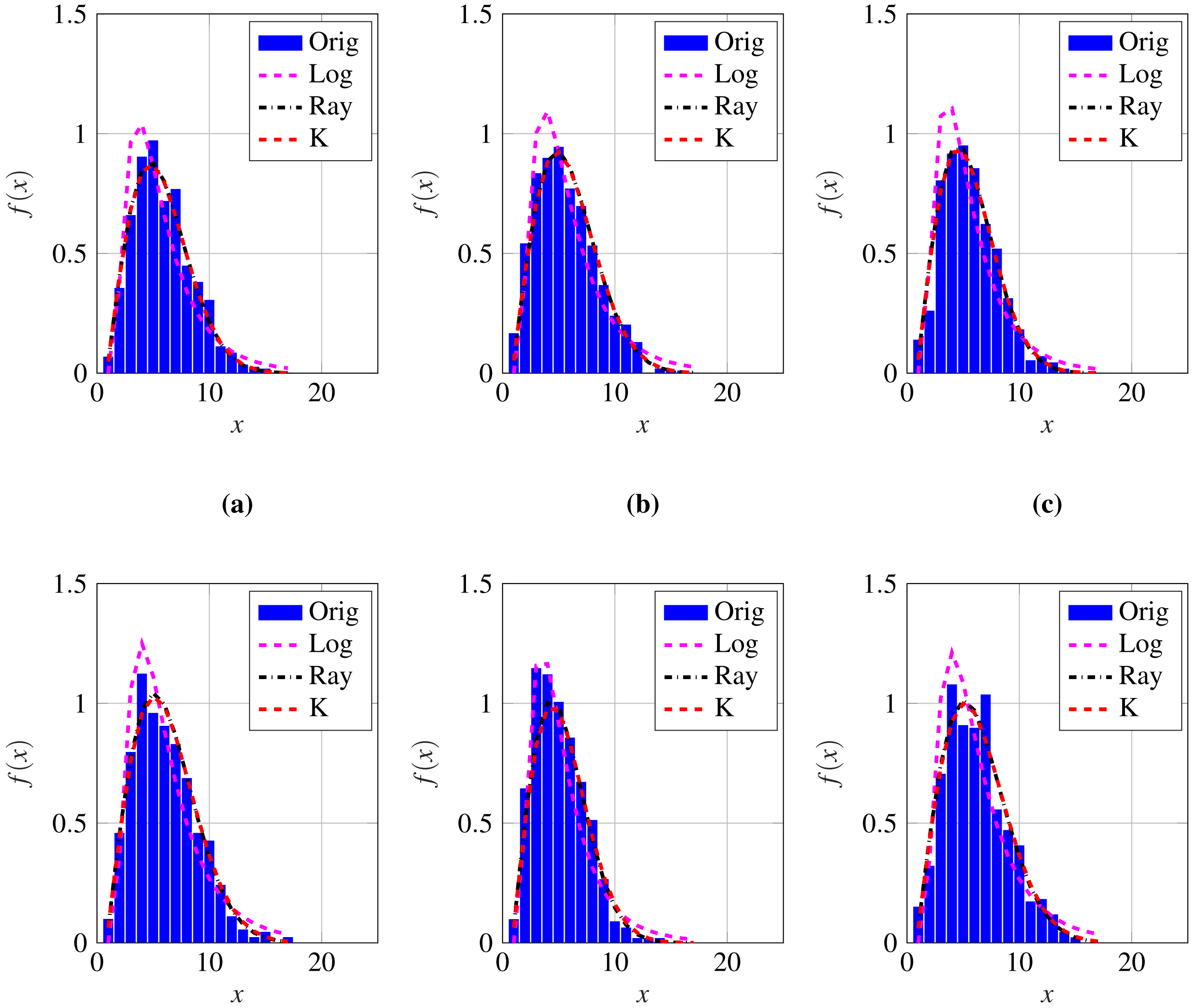
A.2.5 Distribution fitting



(a) (b) (c)

(d) (e) (f)

Figure A.7. CFA16_003 distribution fit example at range bin (a) 44, (b) 45, (c) 46, (d) 47, (e) 48, and (f) 49.



(a)

(b)

(c)

(d)

(e)

(f)

Figure A.8. CFA16_003 distribution fit example at range bin (a) 50, (b) 51, (c) 52, (d) 53, (e) 54, and (f) 55.

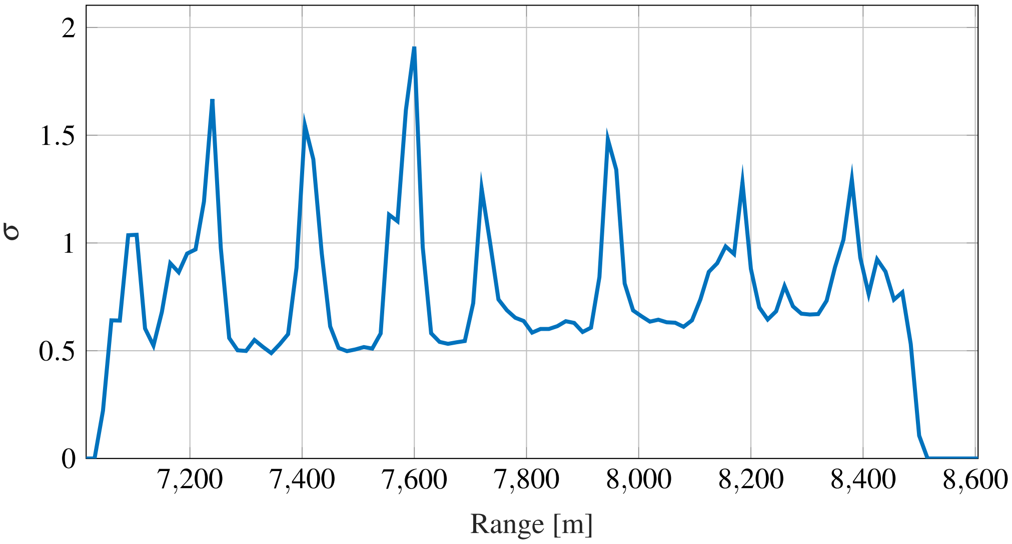


Figure A.9. CFA16_003 PDF fit for Rayleigh distribution illustrating the σ parameter over the entire response range.

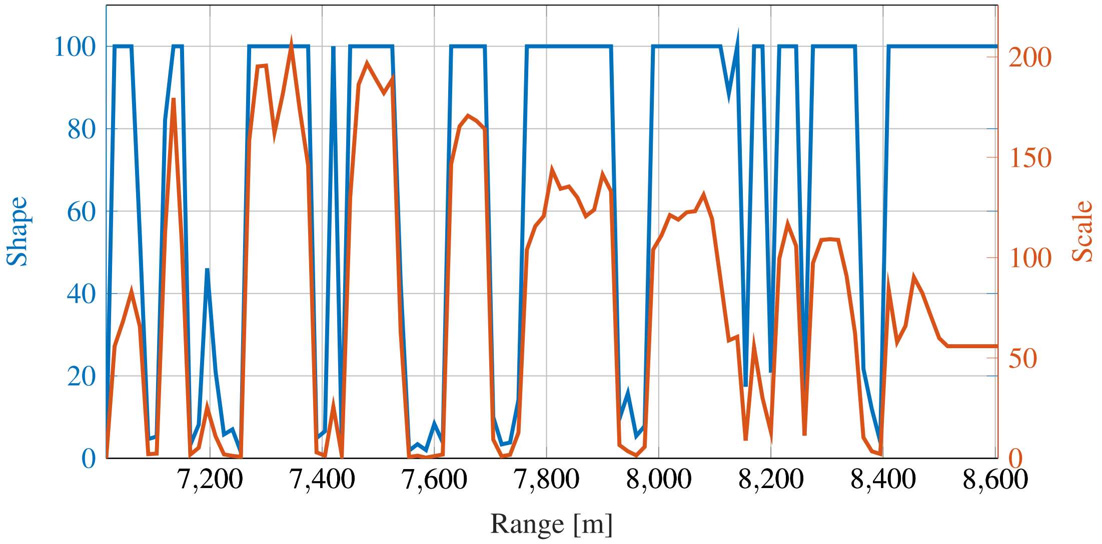


Figure A.10. CFA16_003 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm.

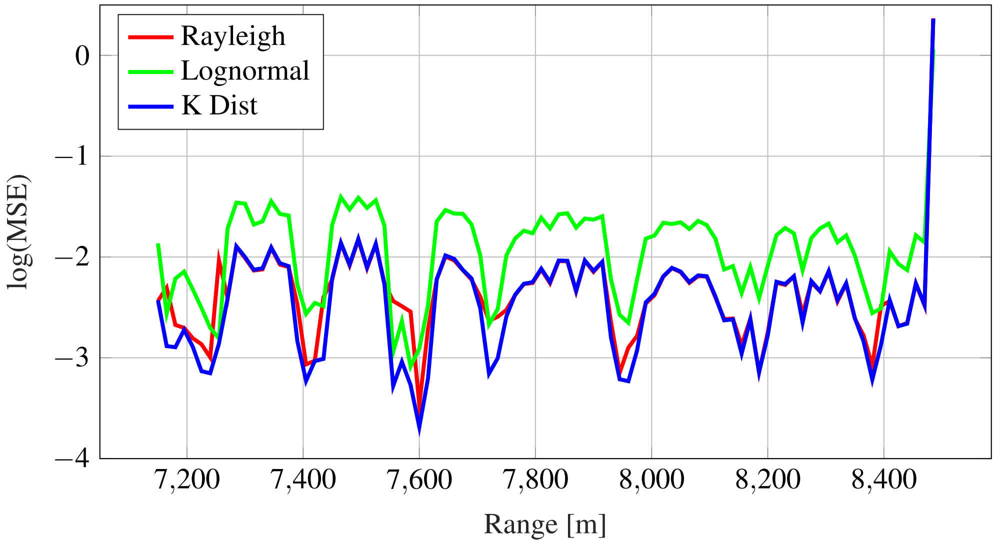


Figure A.11. CFA16_003 parameter MSE comparison of distributions over range in log-domain, illustrating the small error for good fits between parameter peaks.

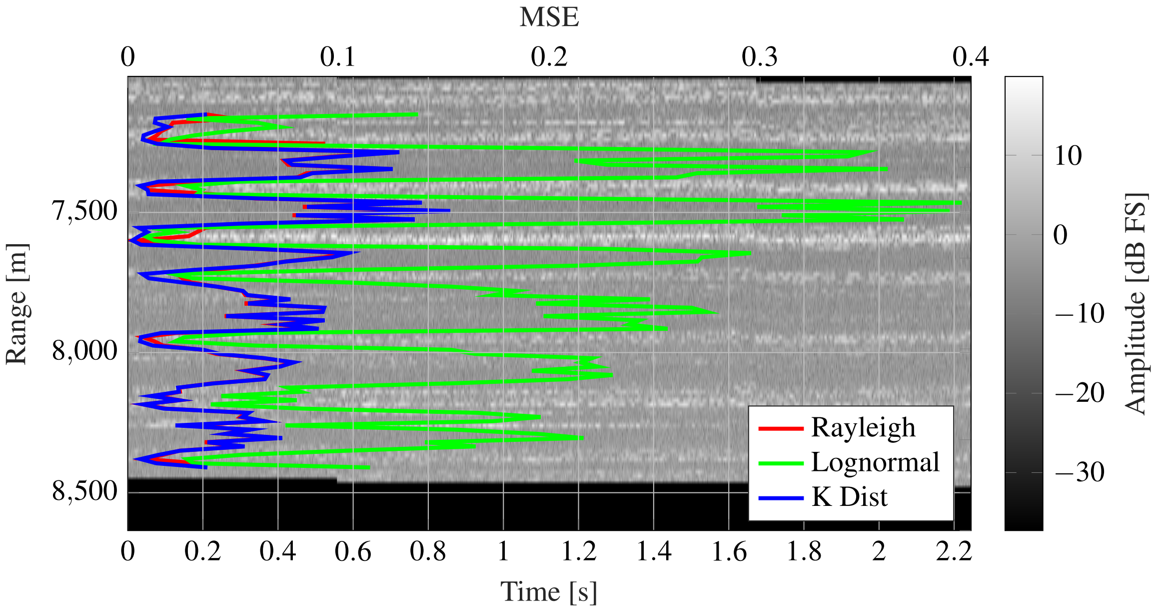


Figure A.12. CFA16_003 comparison illustrating alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.

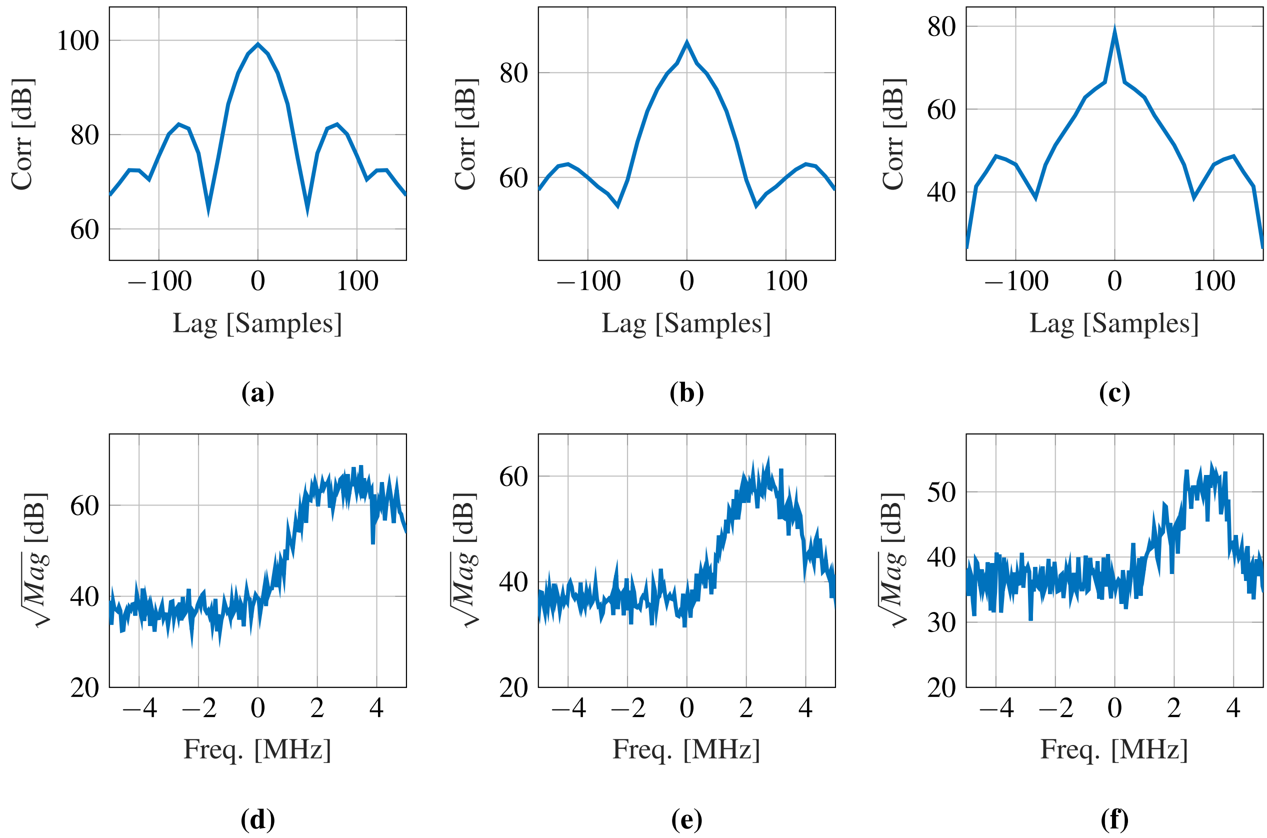


Figure A.13. CFA16_003 correlation at range bin (a) 42, (b) 43, and (c) 44, where (d) to (f) illustrate the corresponding PSD.

A.2.6 Rayleigh correlated output

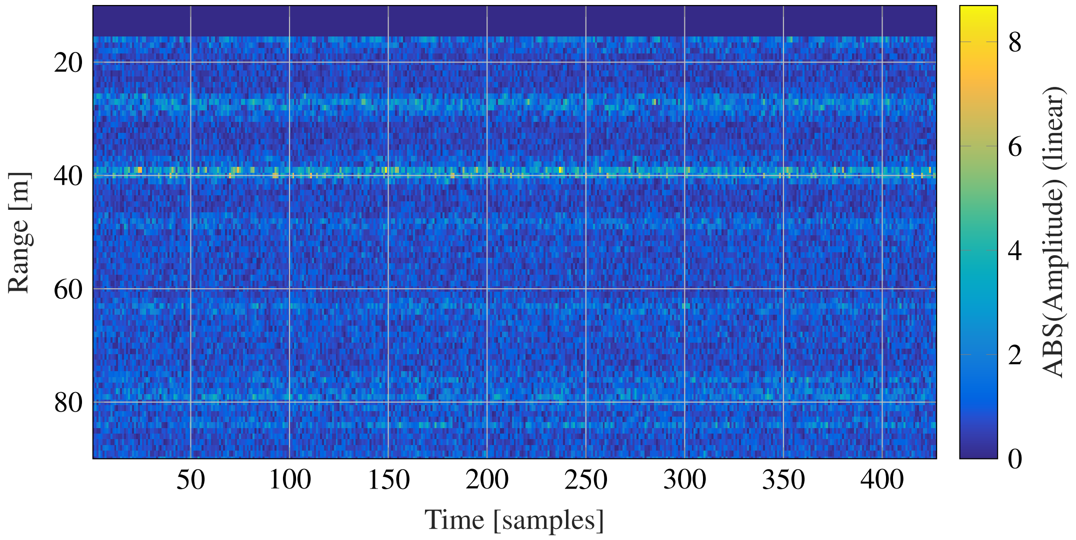
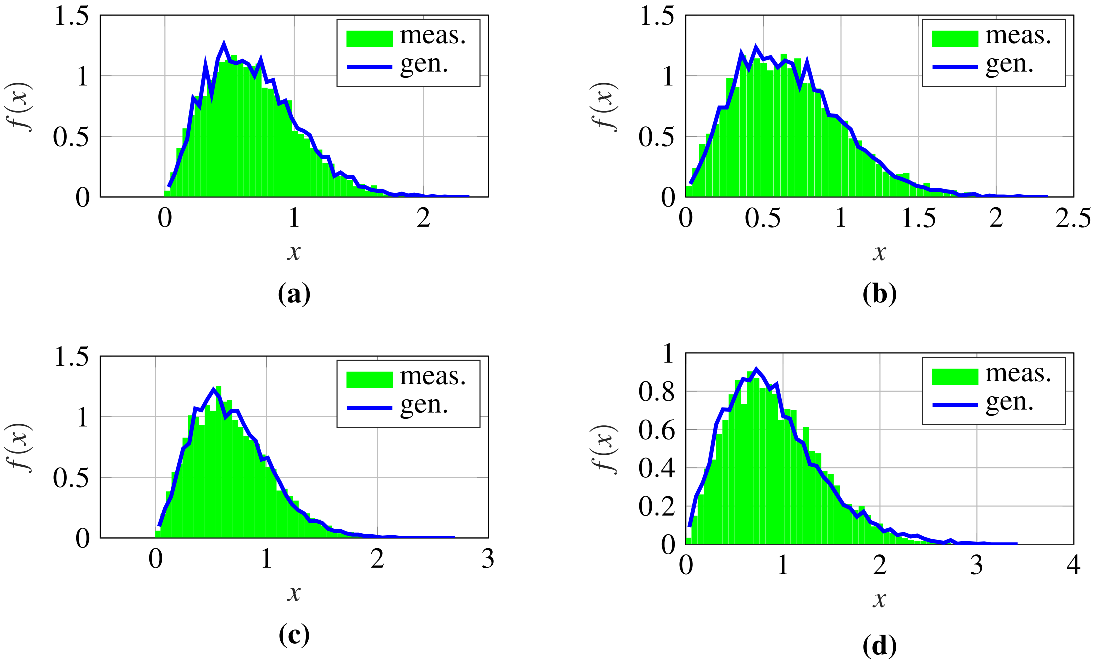


Figure A.14. CFA16_003 range-time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs.



(a)

(b)

(c)

(d)

Figure A.15. CFA16_003 Distribution comparison between measured and generated data at range bin (a) 44, (b) 45, (c) 46, and (d) 47.

A.3 RESULTS DATASET 3: CFA16_004

A.3.1 Raw data

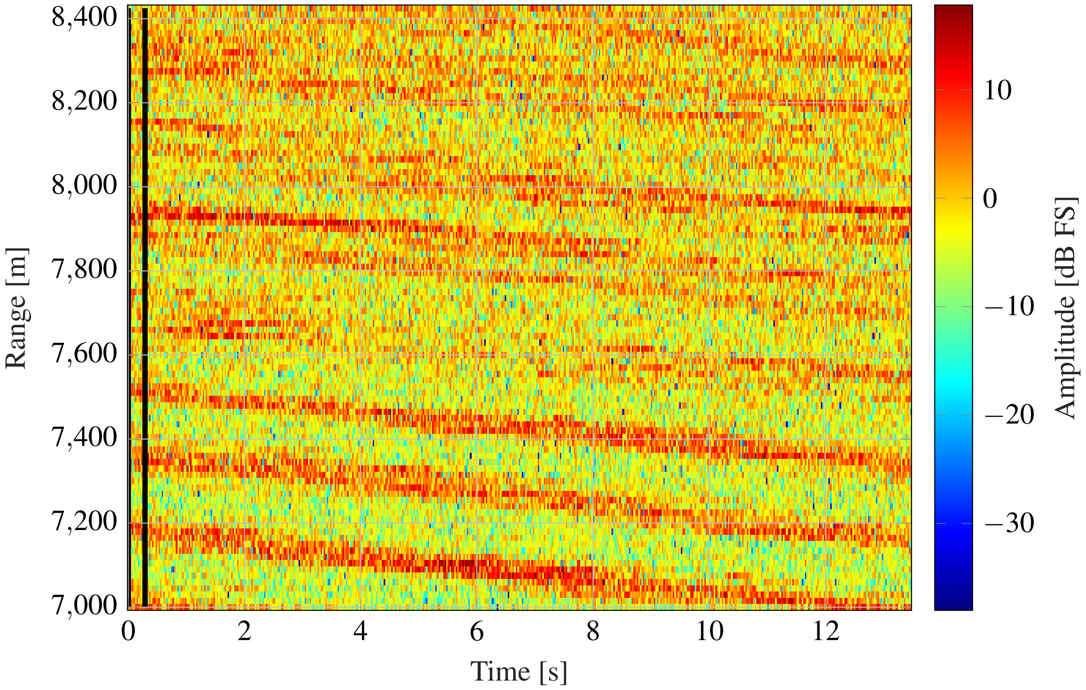


Figure A.16. CFA16_004 range-time plot of 13 s of measured data given in dB full scale, where the black vertical lines indicate the region used for angle determination.

A.3.2 Radon transform

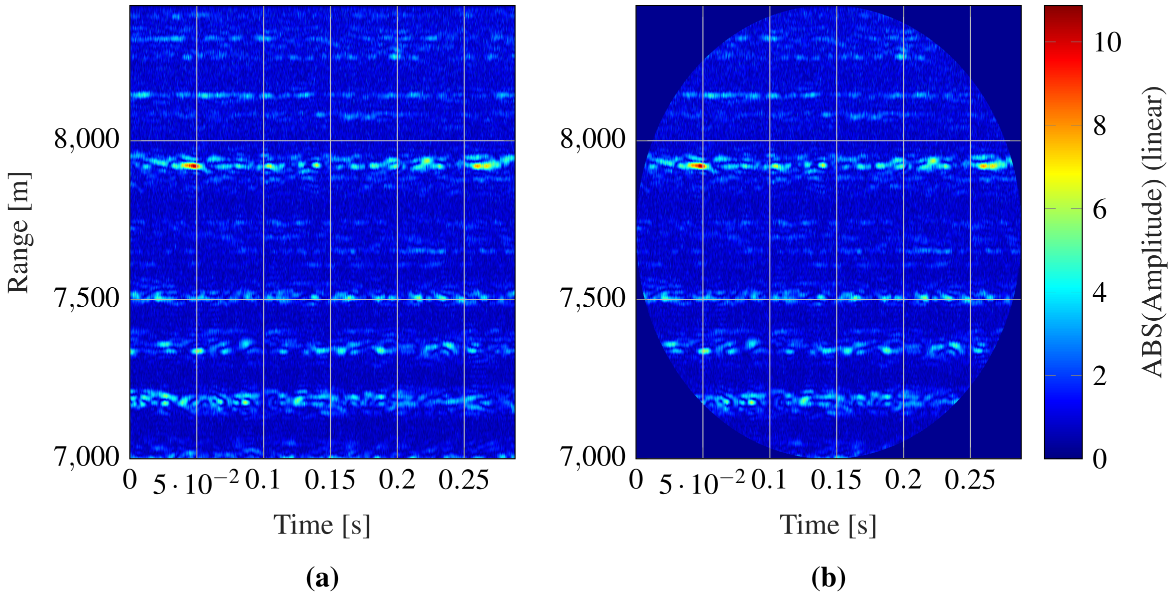


Figure A.17. CFA16_004 non-circular vs circular input processing matrices for the Radon transform, indicating the null data points on the outside of the circular mask.

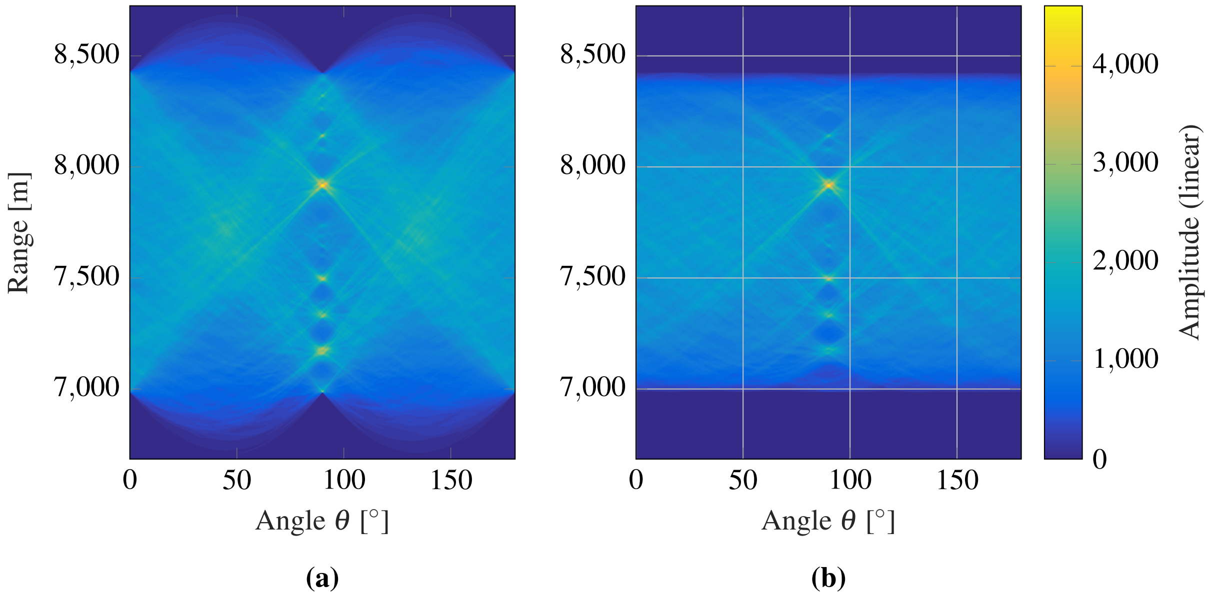


Figure A.18. CFA16_004 Radon transform result illustrating the output result difference for the (a) non-circular and (b) circular input matrix.

A.3.3 Angle calculation

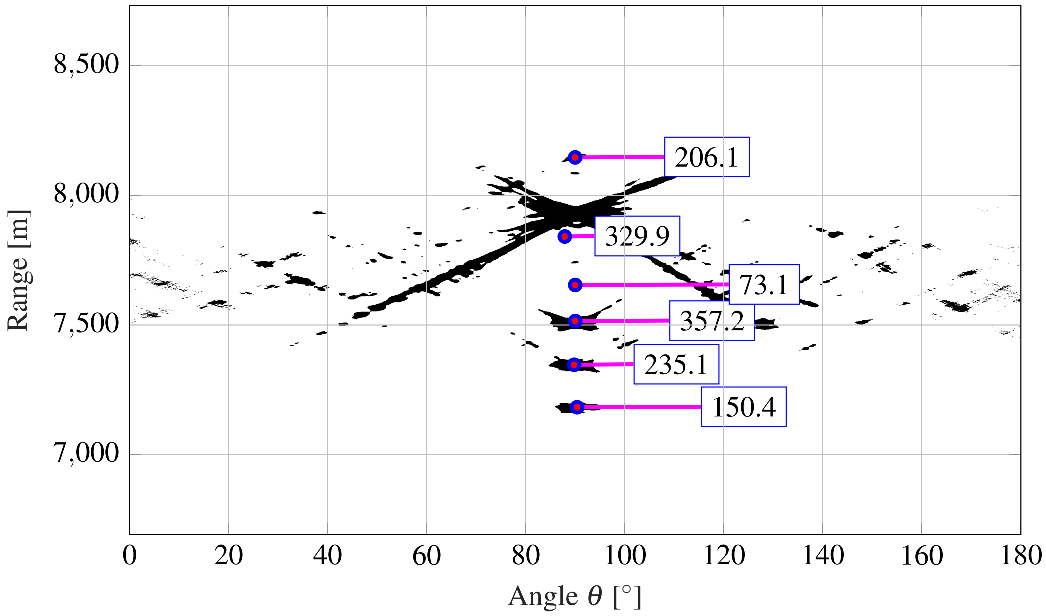


Figure A.19. CFA16_004 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

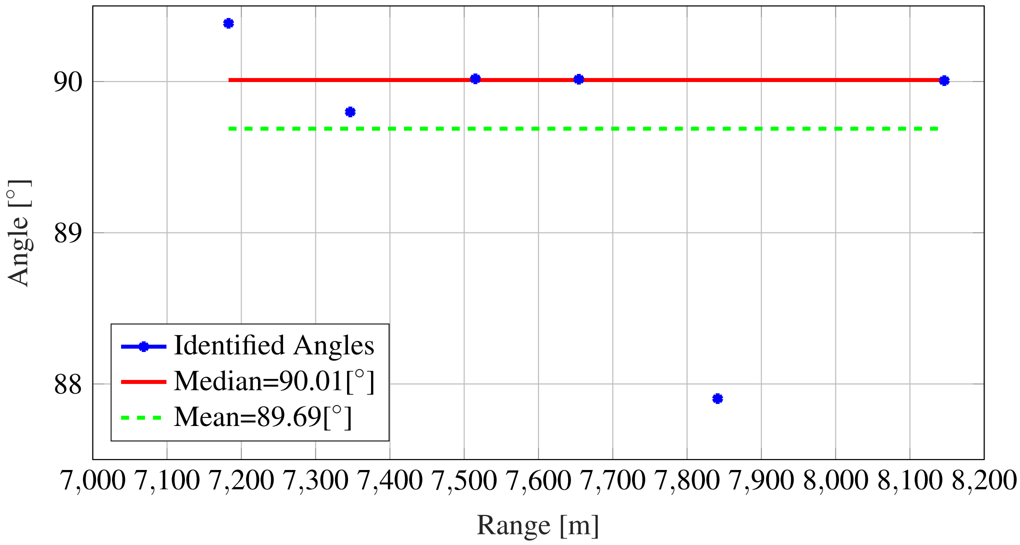


Figure A.20. CFA16_004 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

A.3.4 Wave phase alignment

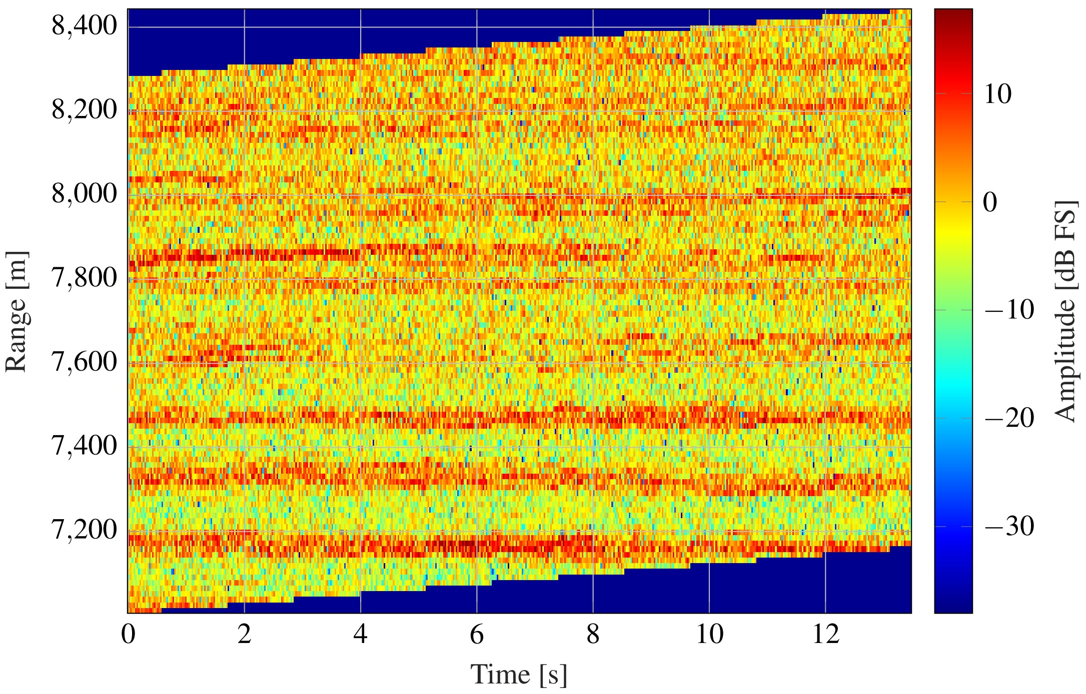
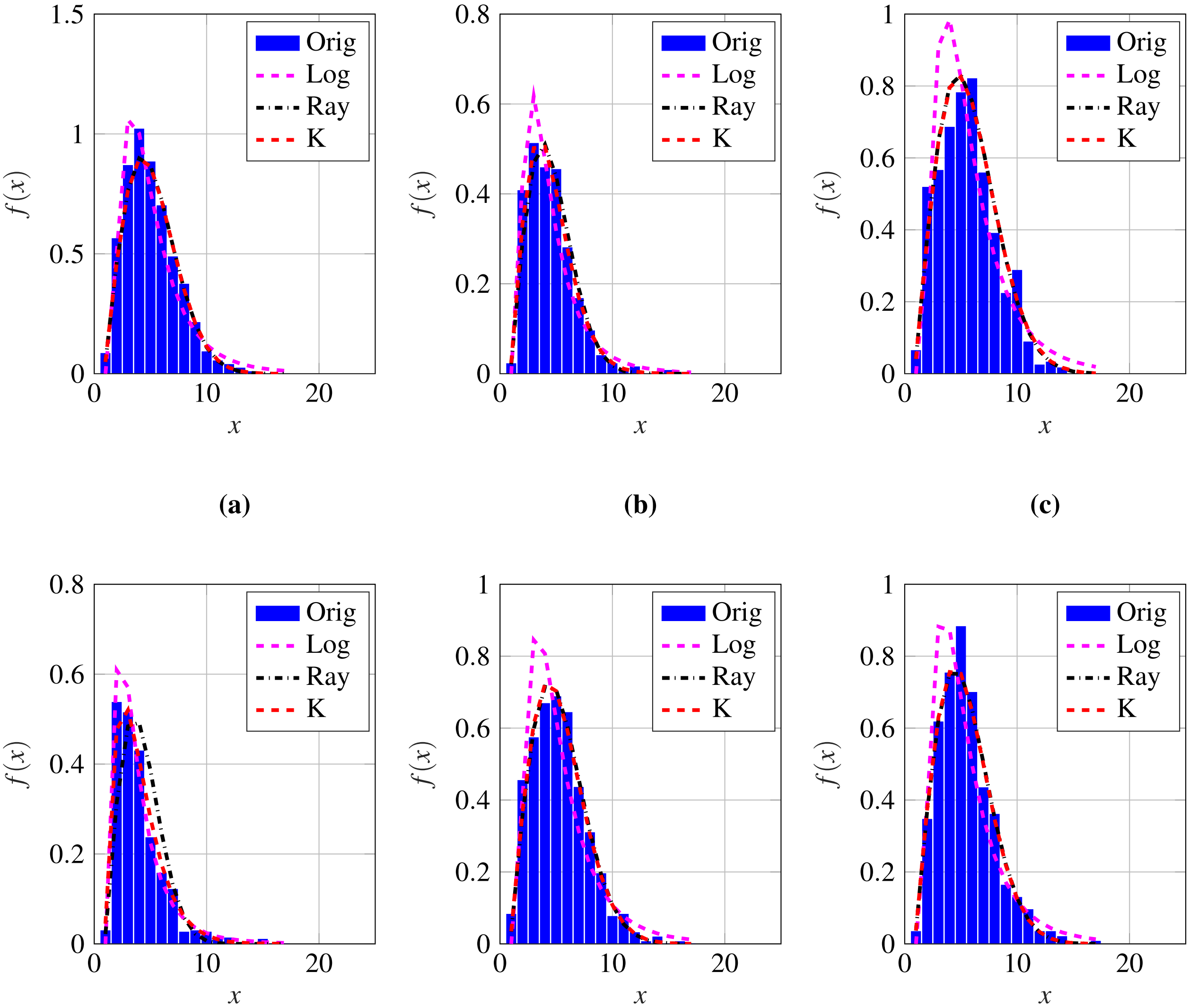


Figure A.21. CFA16_004 output of wave-phase corrected data with *x*-axis aligned waves illustrating clear steps.

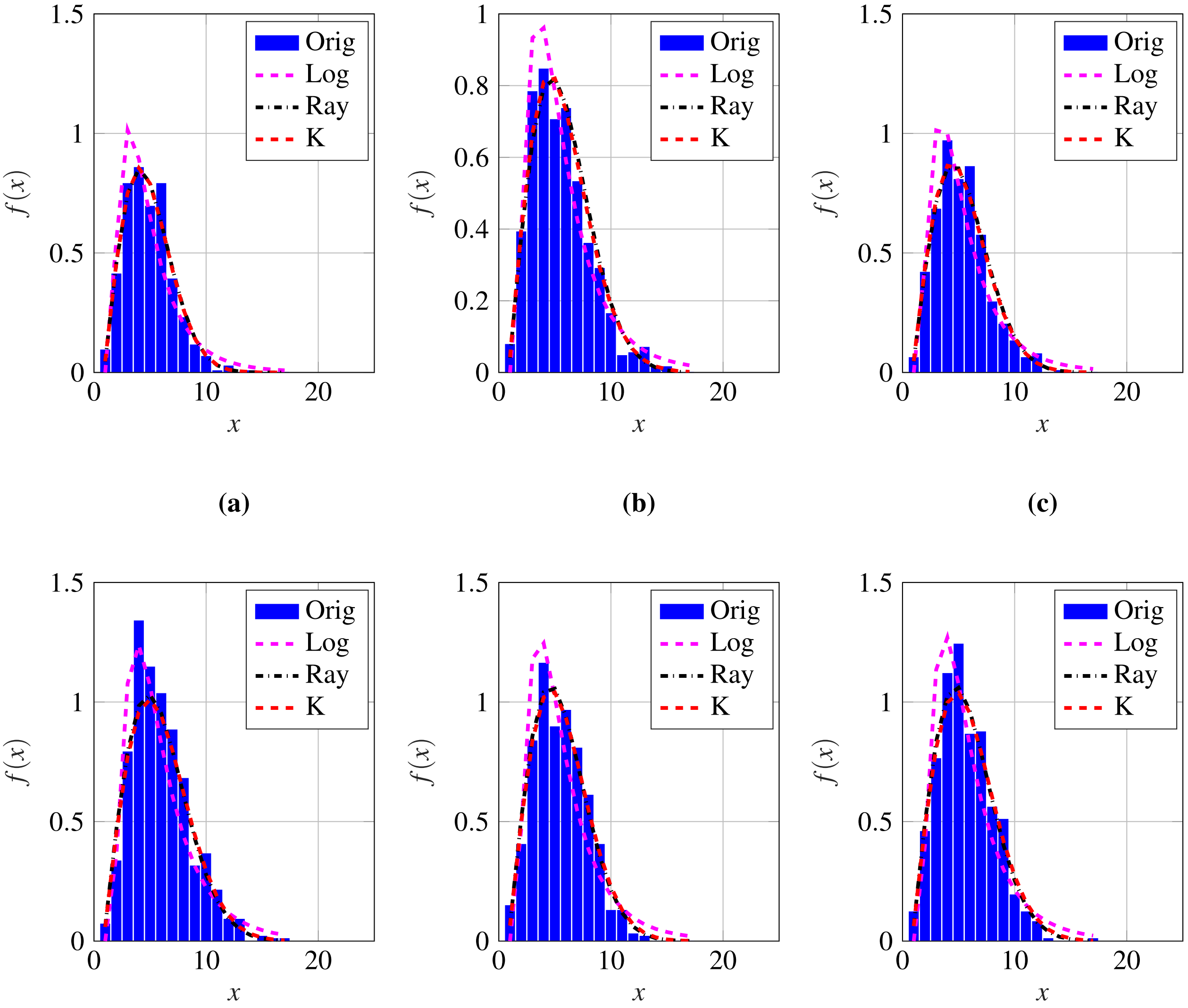
A.3.5 Distribution fitting



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.22. CFA16_004 distribution fit example at range bin (a) 44, (b) 45, (c) 46, (d) 47, (e) 48, and (f) 49.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.23. CFA16_004 distribution fit example at range bin (a) 50, (b) 51, (c) 52, (d) 53, (e) 54, and (f) 55.

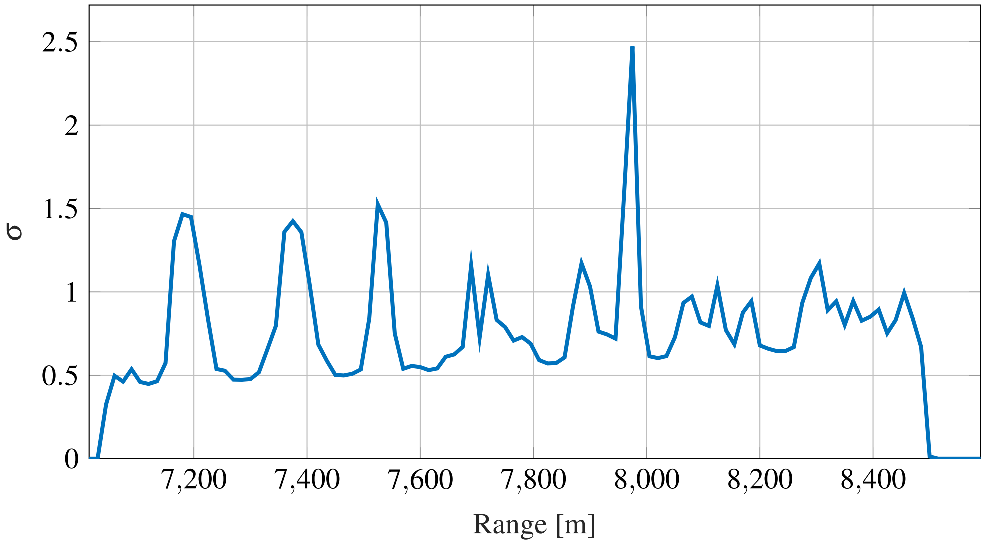


Figure A.24. CFA16_004 PDF fit for Rayleigh distribution illustrating the σ parameter over the entire response range.

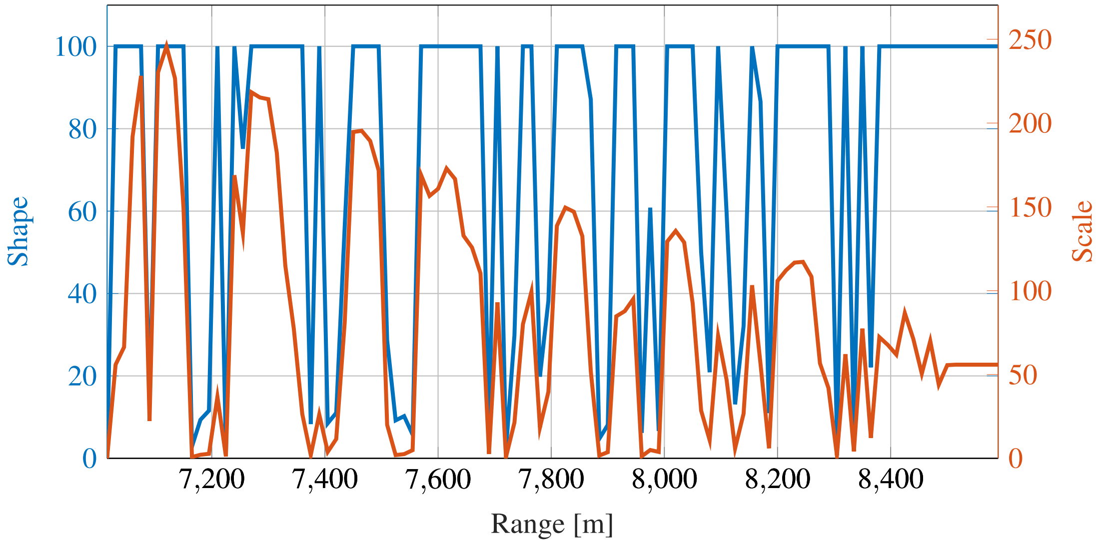


Figure A.25. CFA16_004 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm.

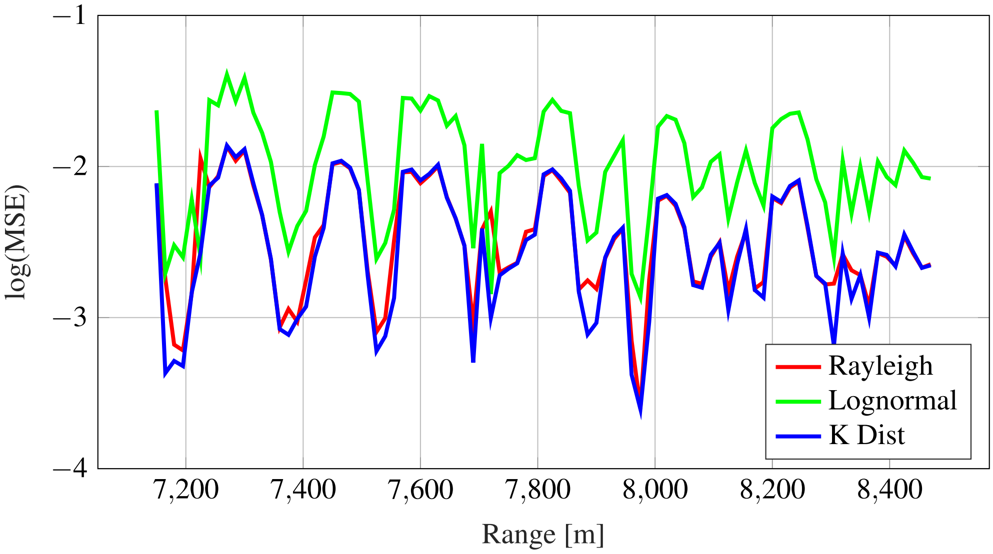


Figure A.26. CFA16_004 parameter MSE comparison of distributions over range in log-domain, illustrating the small error for good fits between parameter peaks.

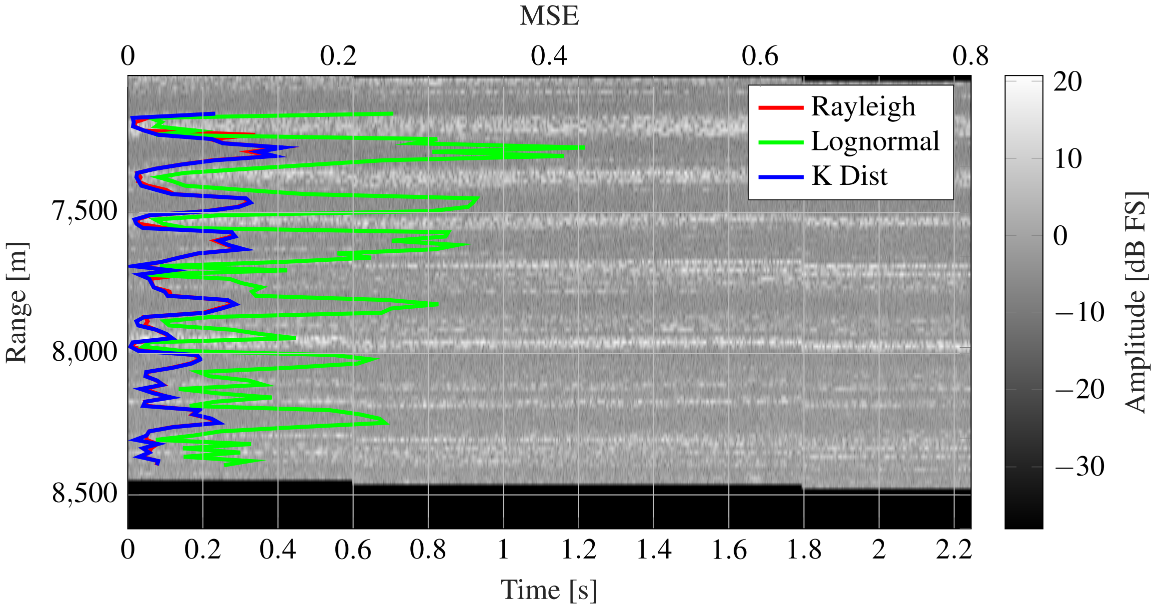
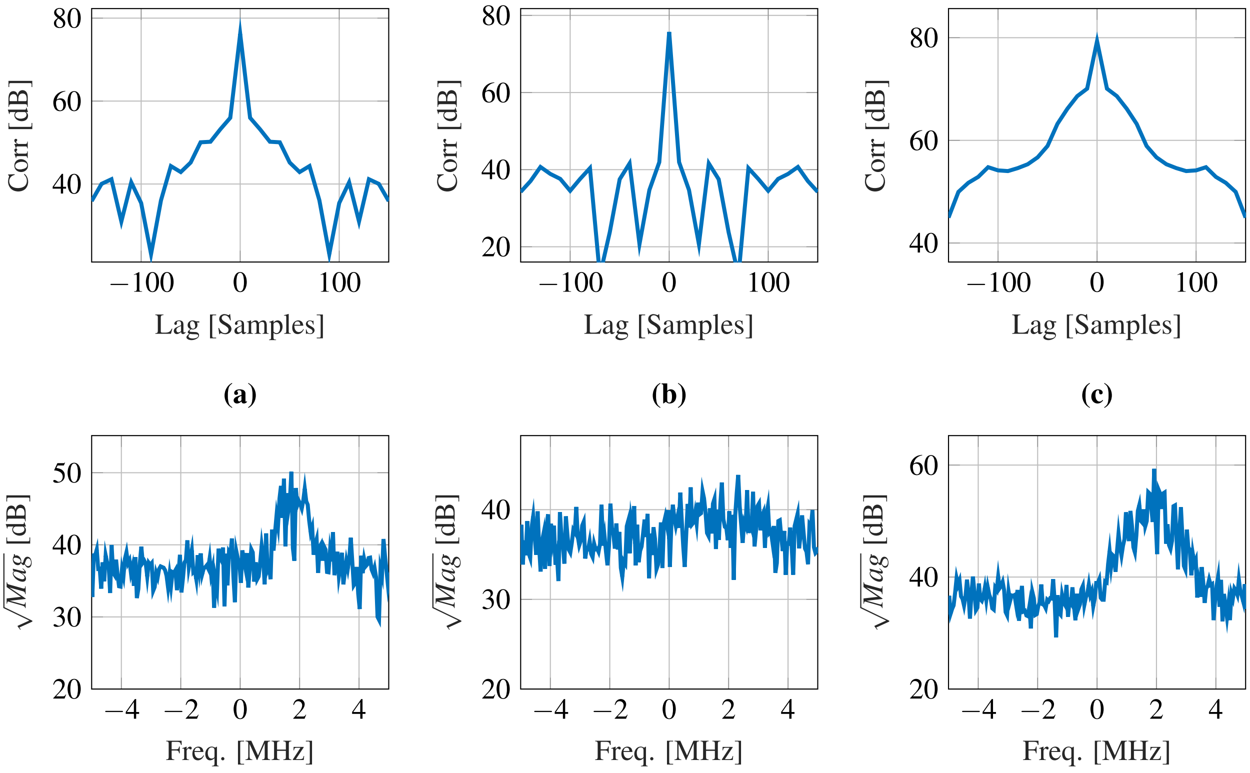


Figure A.27. CFA16_004 comparison illustrating alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.28. CFA16_004 correlation at range bin (a) 42, (b) 43, and (c) 44, where (d) to (f) illustrate the corresponding PSD.

A.3.6 Rayleigh correlated output

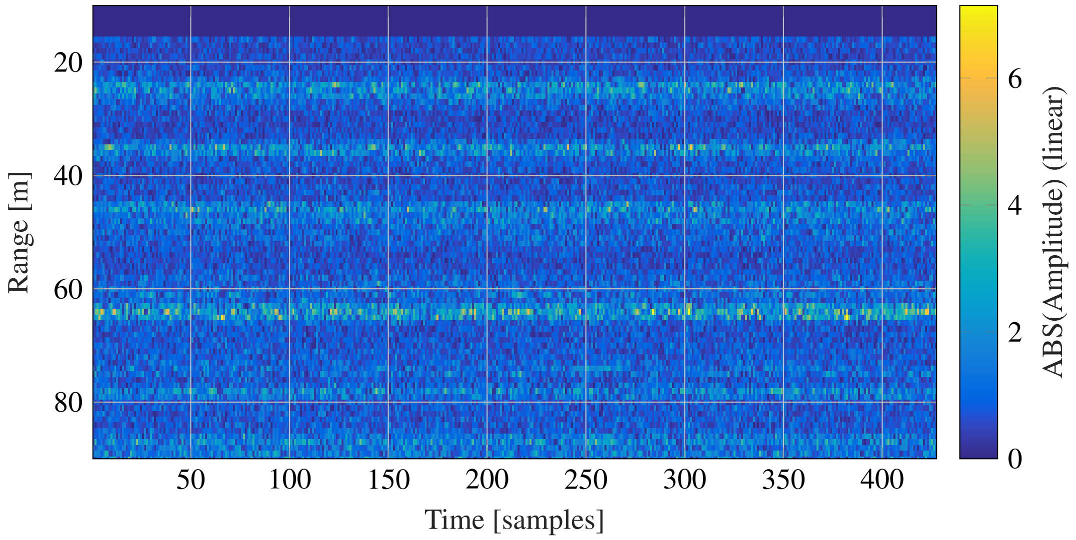
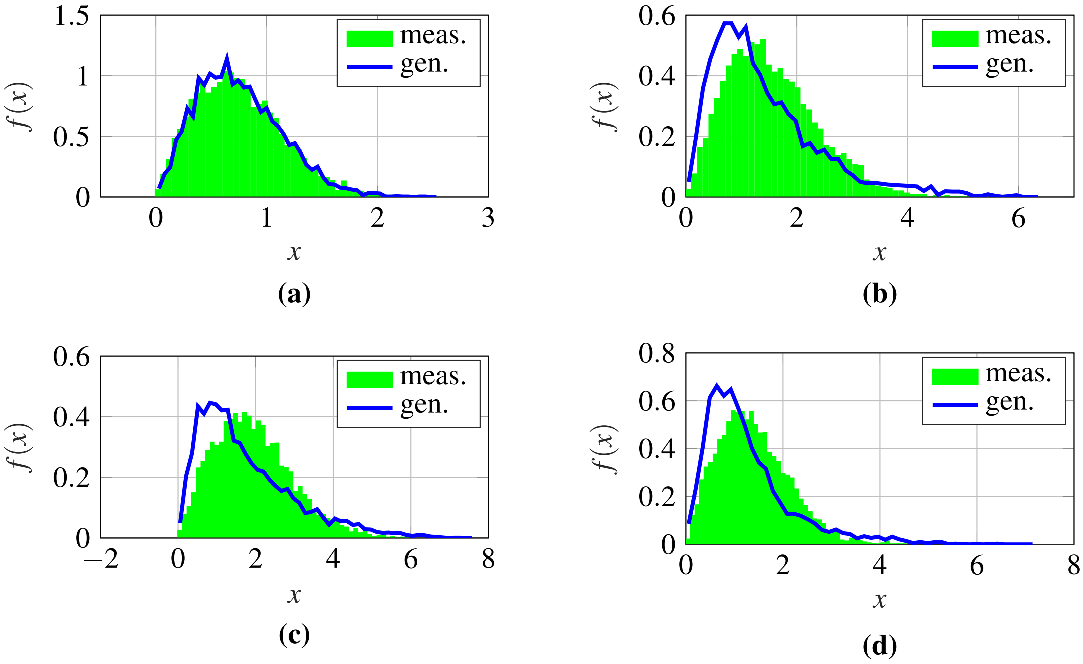


Figure A.29. CFA16_004 range-time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs.



(a)

(b)

(c)

(d)

Figure A.30. CFA16_004 Distribution comparison between measured and generated data at range bin (a) 44, (b) 45, (c) 46, and (d) 47.

A.4 RESULTS DATASET 4: TFC15_025

A.4.1 Raw data

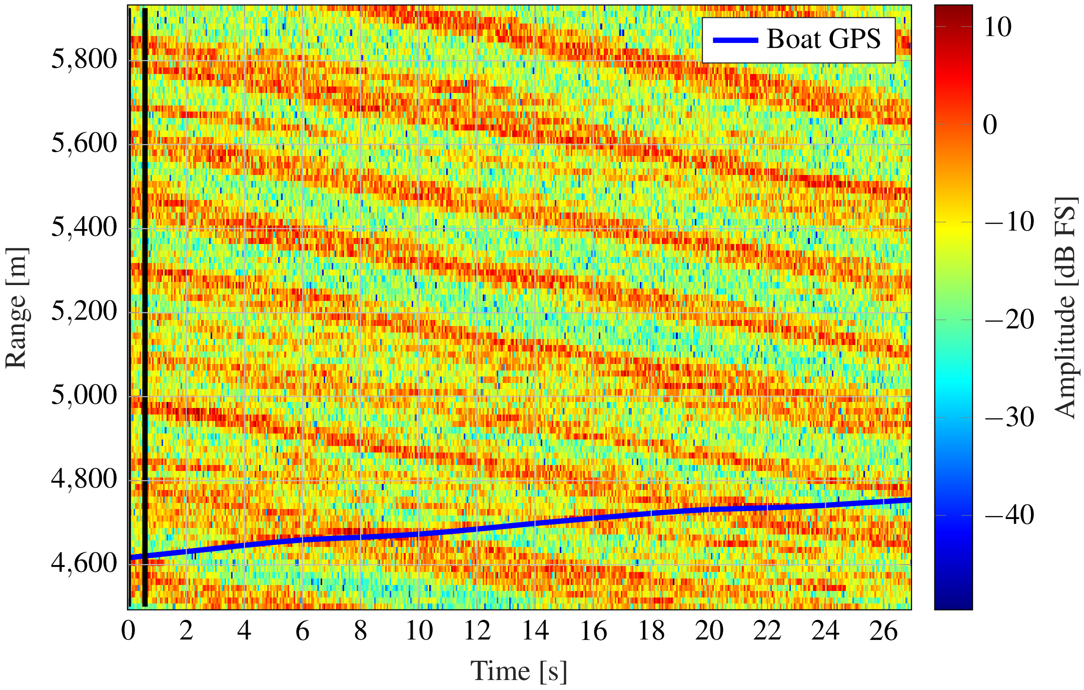
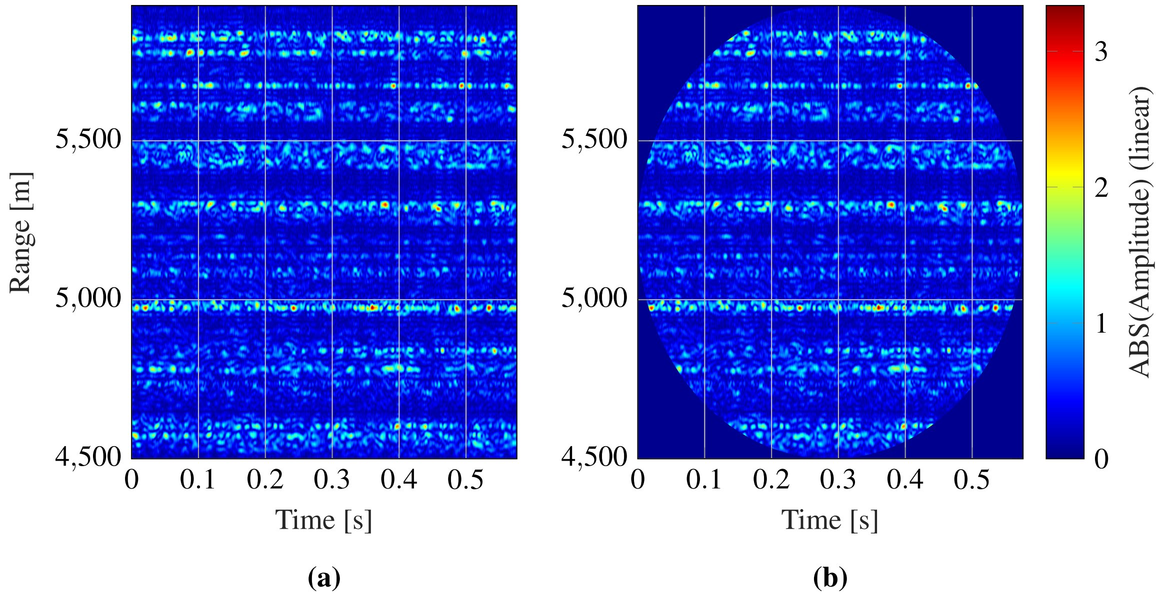


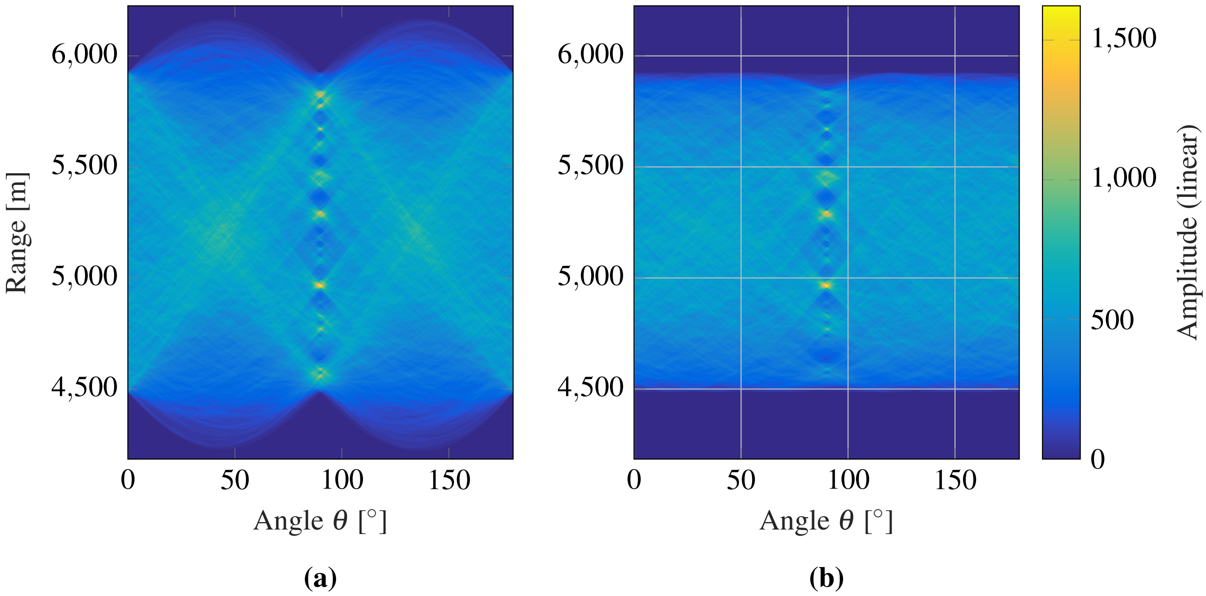
Figure A.31. TFC15_025 range-time plot of 26 s of measured data with boat information given in dB full scale, where the black vertical lines indicate the region used for angle determination.

A.4.2 Radon transform



(a) **(b)**

Figure A.32. TFC15_025 non-circular vs circular input processing matrices for Radon transform, indicating the null data points on the outside of the circular mask.



(a) **(b)**

Figure A.33. TFC15_025 Radon transform result illustrating the output result difference for the (a) non-circular and (b) circular input matrix.

A.4.3 Angle calculation

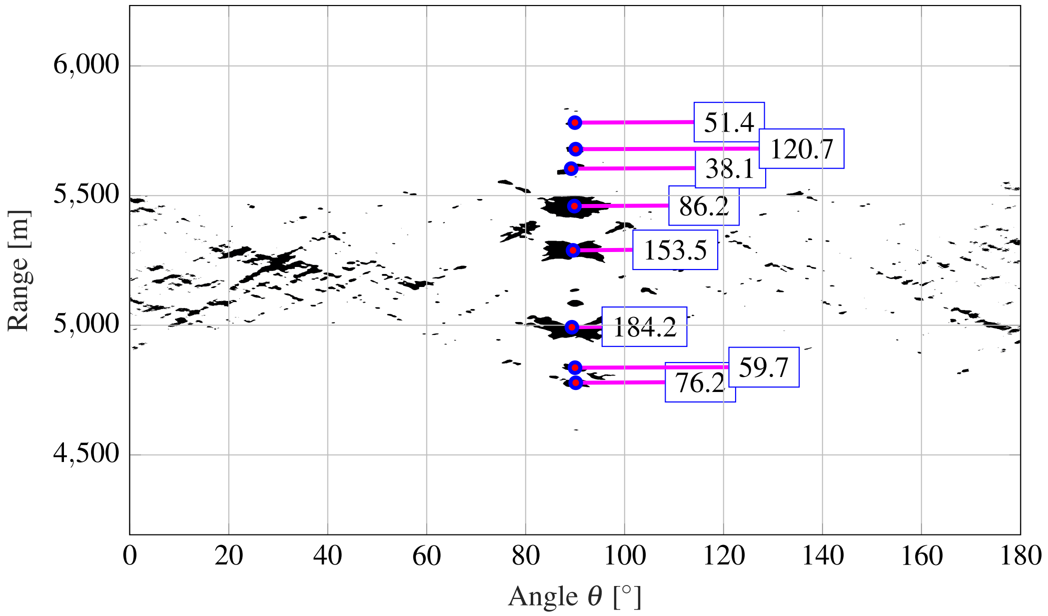


Figure A.34. TFC15_025 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

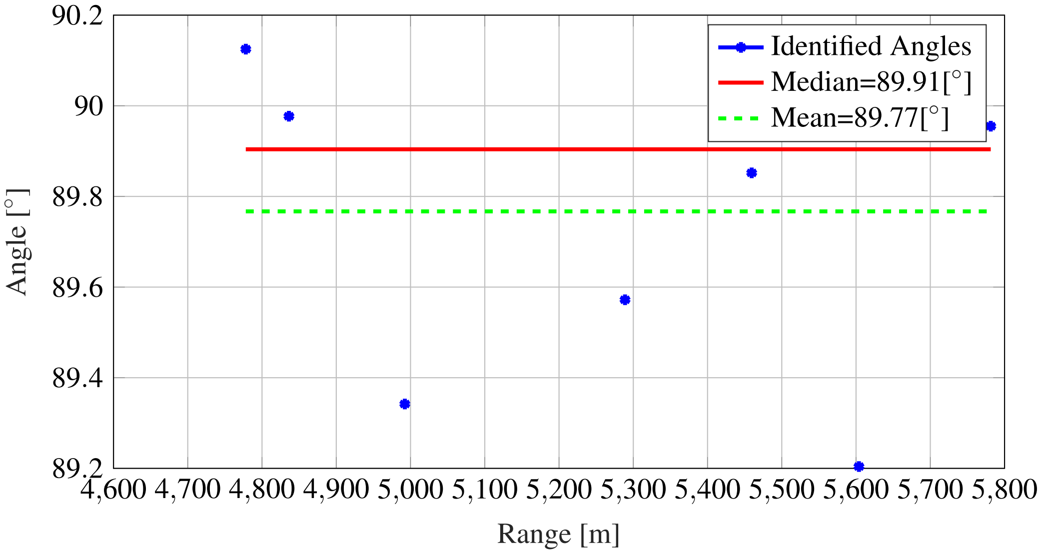


Figure A.35. TFC15_025 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

A.4.4 Wave phase alignment

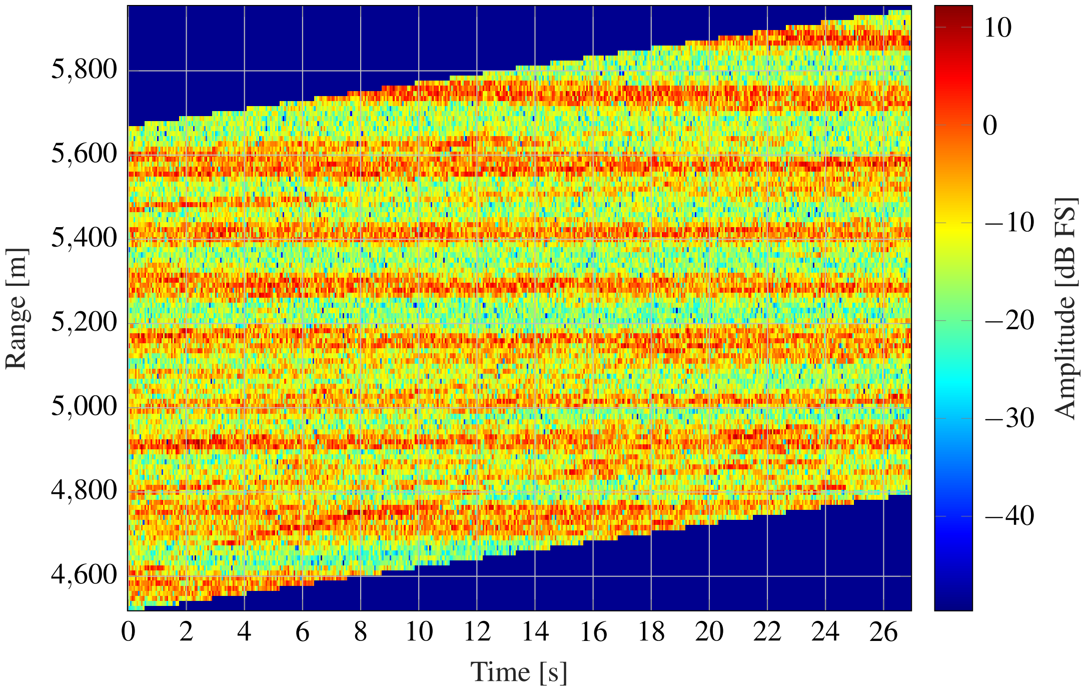
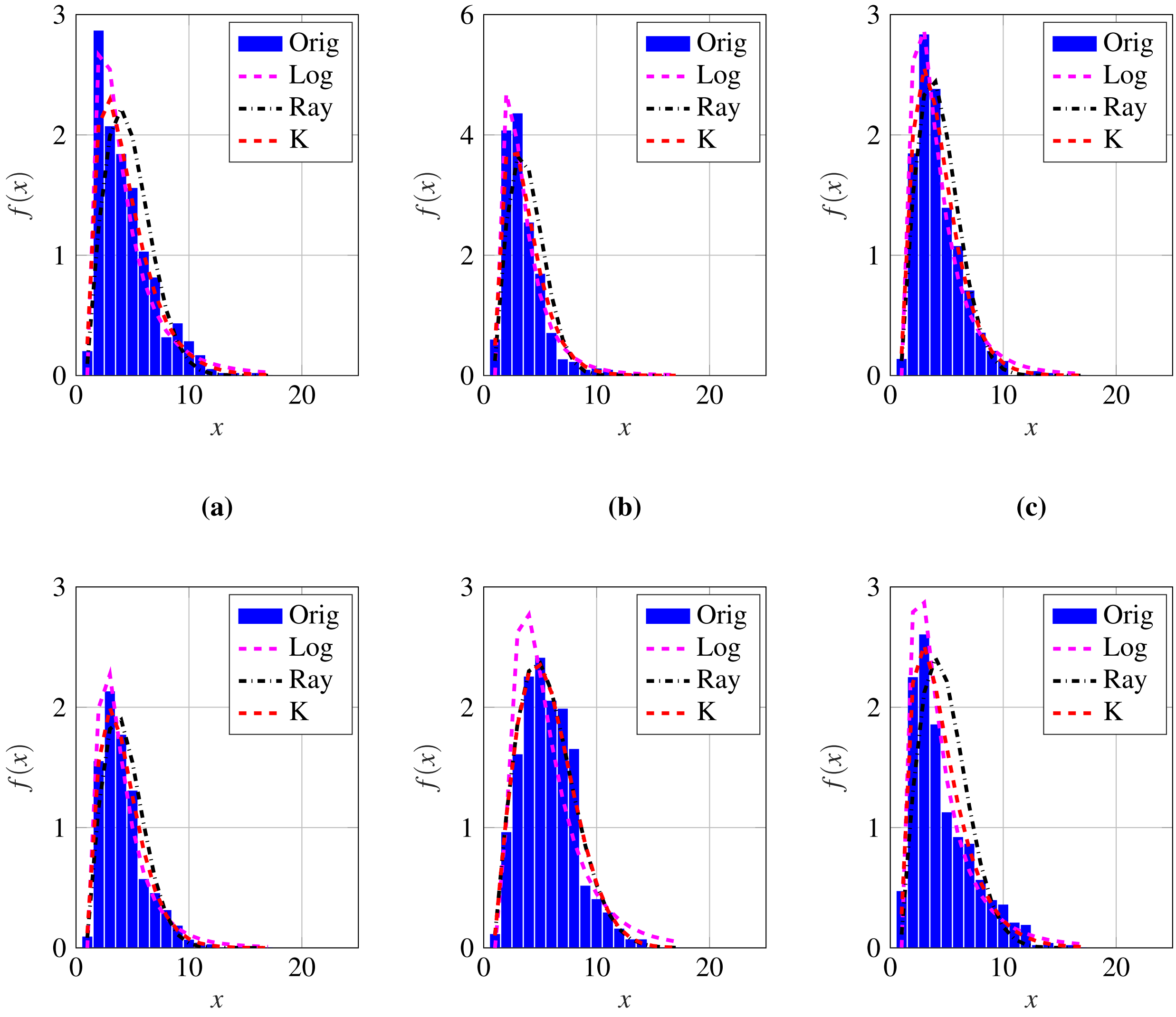


Figure A.36. TFC15_025 output of wave-phase corrected data with *x*-axis aligned waves illustrating clear steps.

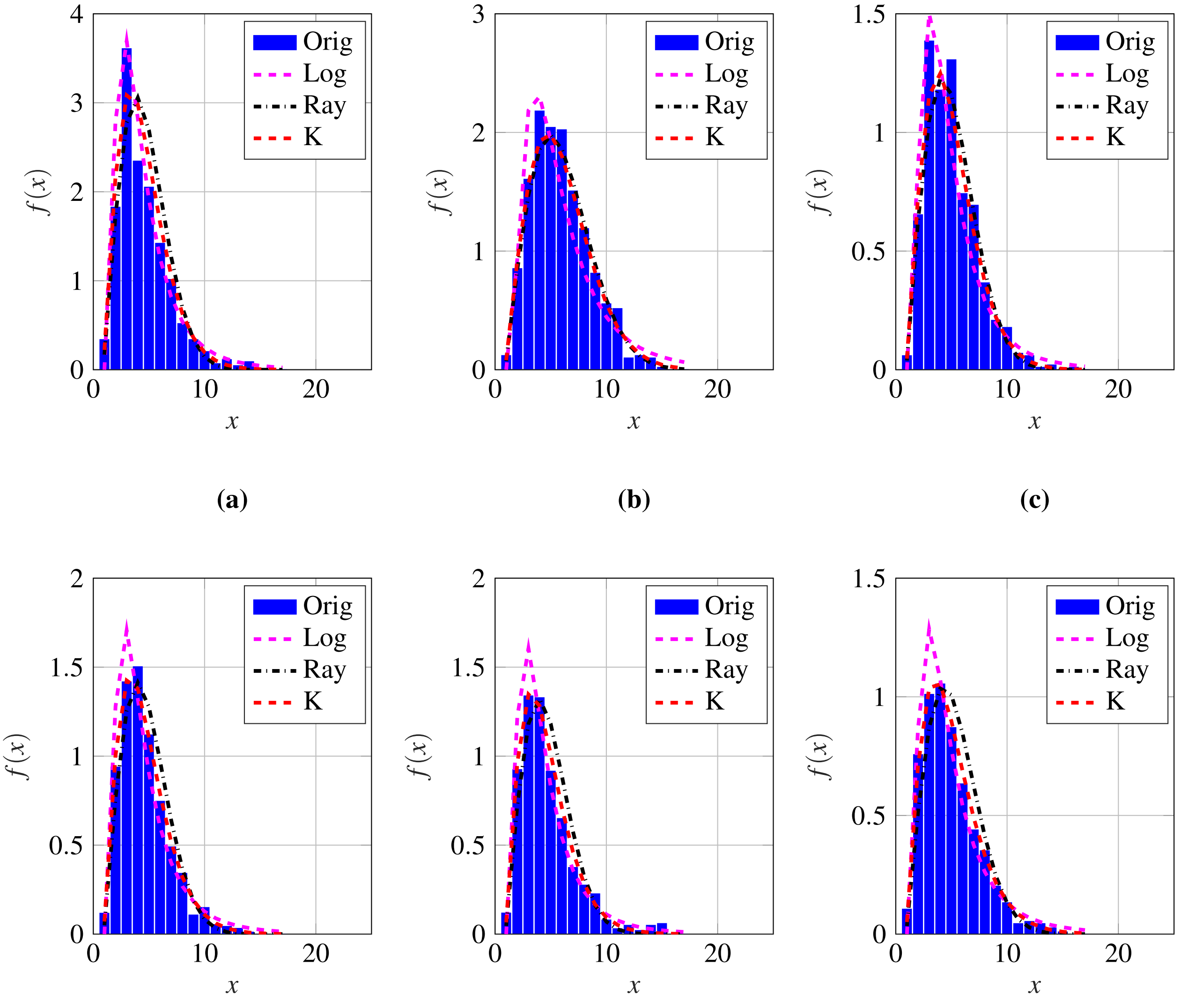
A.4.5 Distribution fitting



(a) (b) (c)

(d) (e) (f)

Figure A.37. TFC15_025 distribution fit example at range bin (a) 44, (b) 45, (c) 46, (d) 47, (e) 48, and (f) 49.



(a) (b) (c)

(d) (e) (f)

Figure A.38. TFC15_025 distribution fit example at range bin (a) 50, (b) 51, (c) 52, (d) 53, (e) 54, and (f) 55.

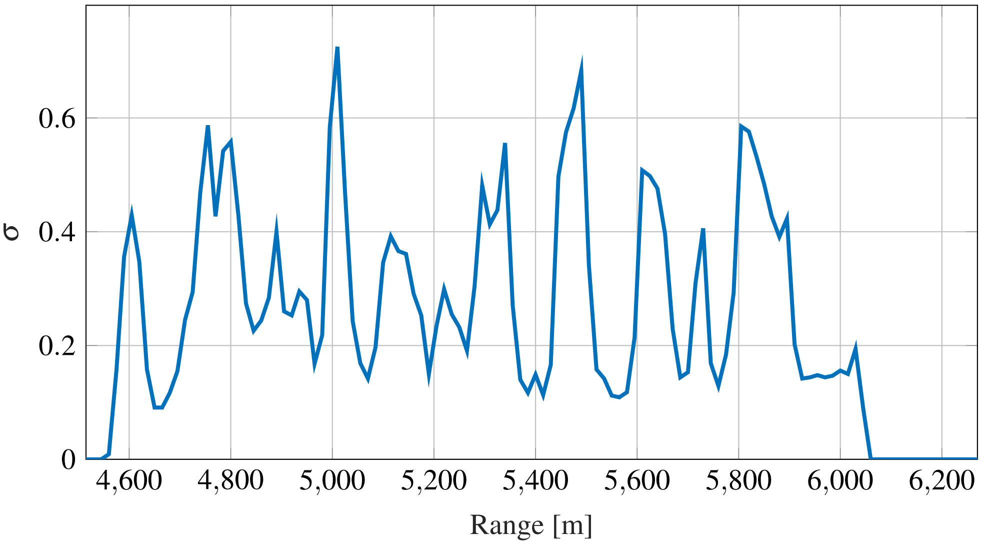


Figure A.39. TFC15_025 PDF fit for Rayleigh distribution illustrating the σ parameter over the entire response range.

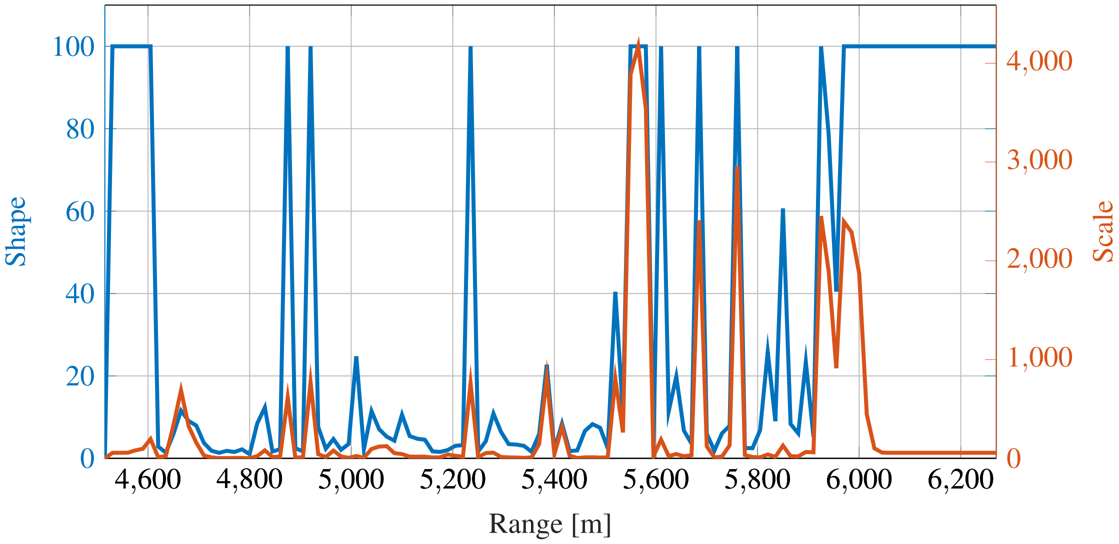


Figure A.40. TFC15_025 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm.

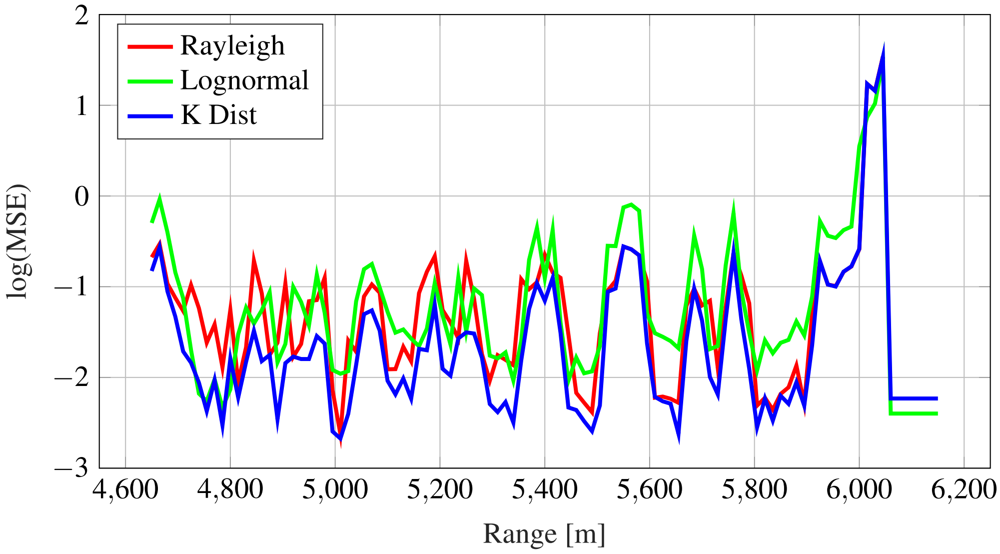


Figure A.41. TFC15_025 parameter MSE comparison of distributions over range in log-domain, illustrating the small error for good fits between parameter peaks.

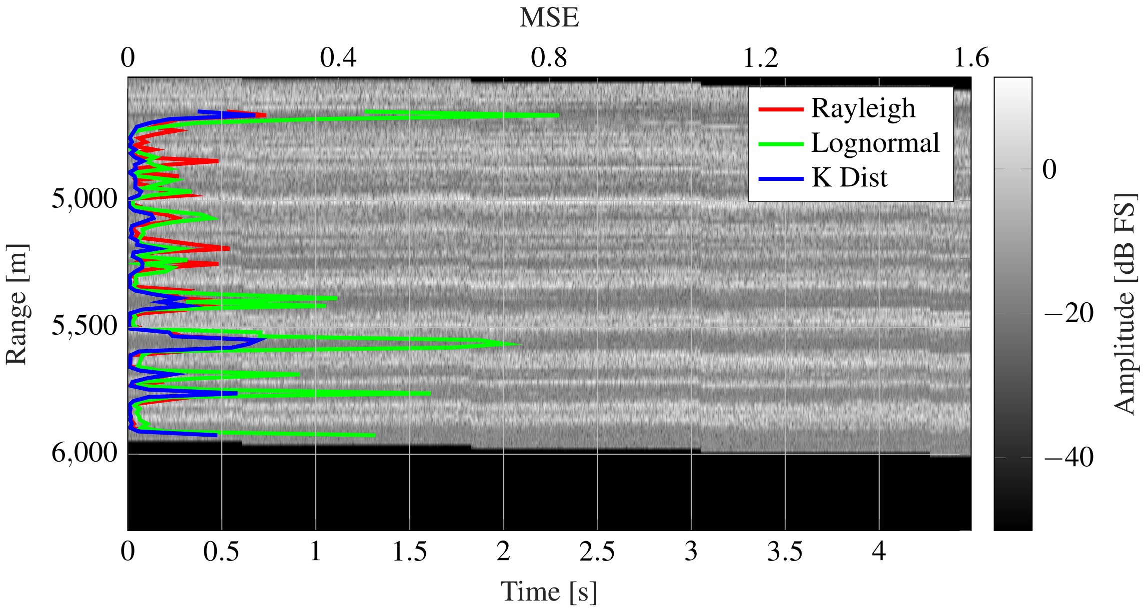
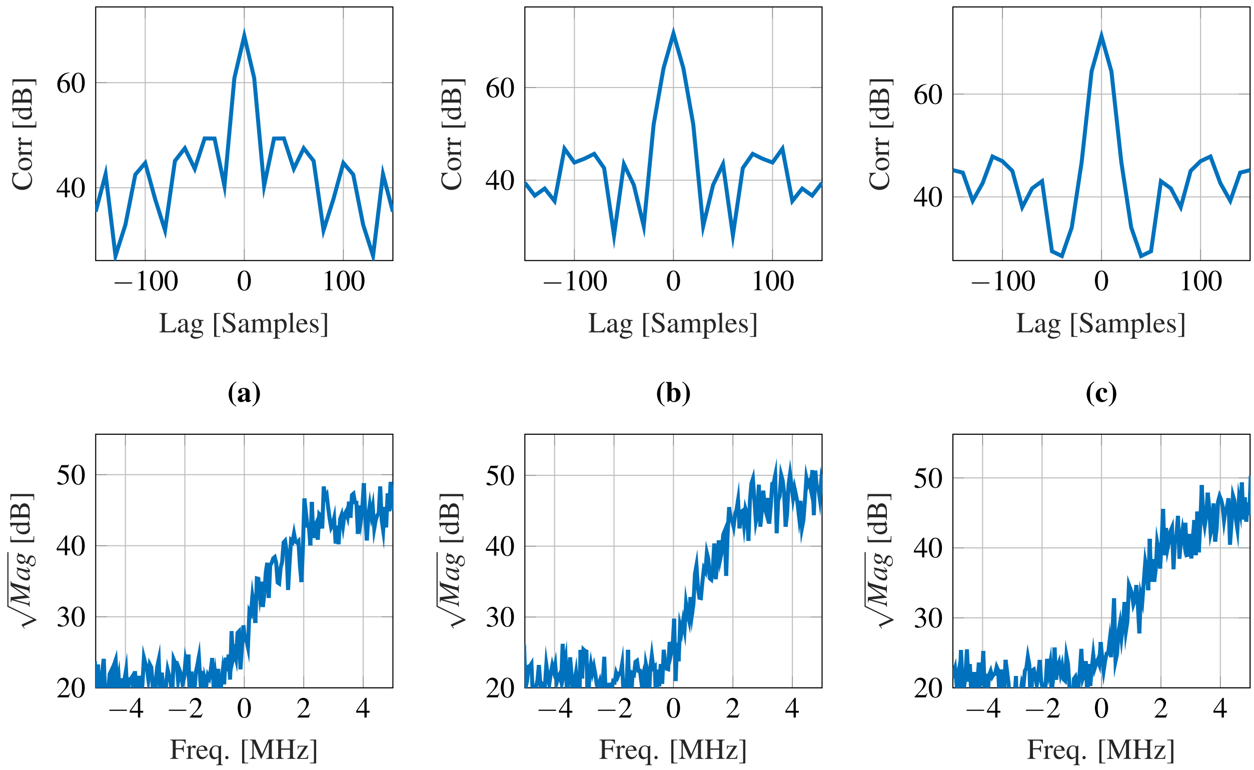


Figure A.42. TFC15_025 comparison illustrating alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.43. TFC15_025 correlation at range bin (a) 42, (b) 43, and (c) 44, where (d) to (f) illustrate the corresponding PSD.

A.4.6 Rayleigh correlated output

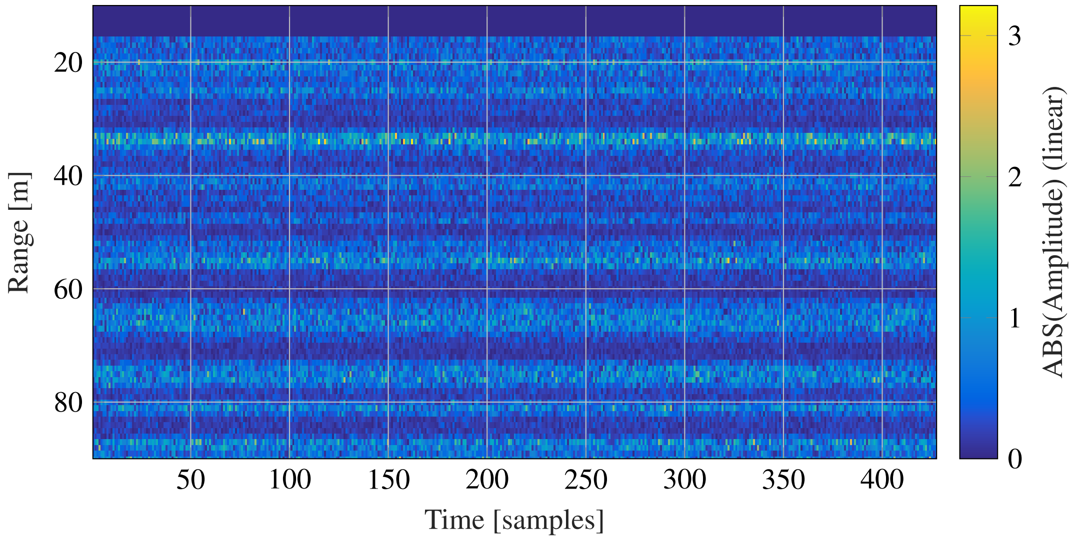
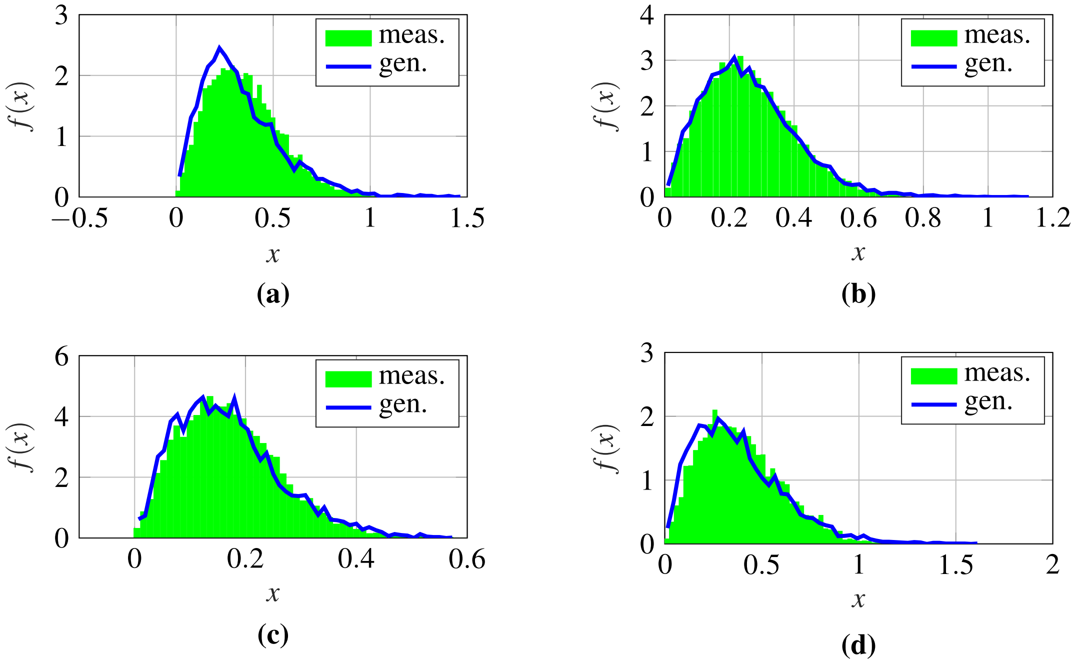


Figure A.44. TFC15_025 range-time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs.



(a) **(b)** **(c)** **(d)**

Figure A.45. TFC15_025 Distribution comparison between measured and generated data at range bin (a) 44, (b) 45, (c) 46, and (d) 47.

A.5 RESULTS DATASET 5: TFC15_028

A.5.1 Raw data

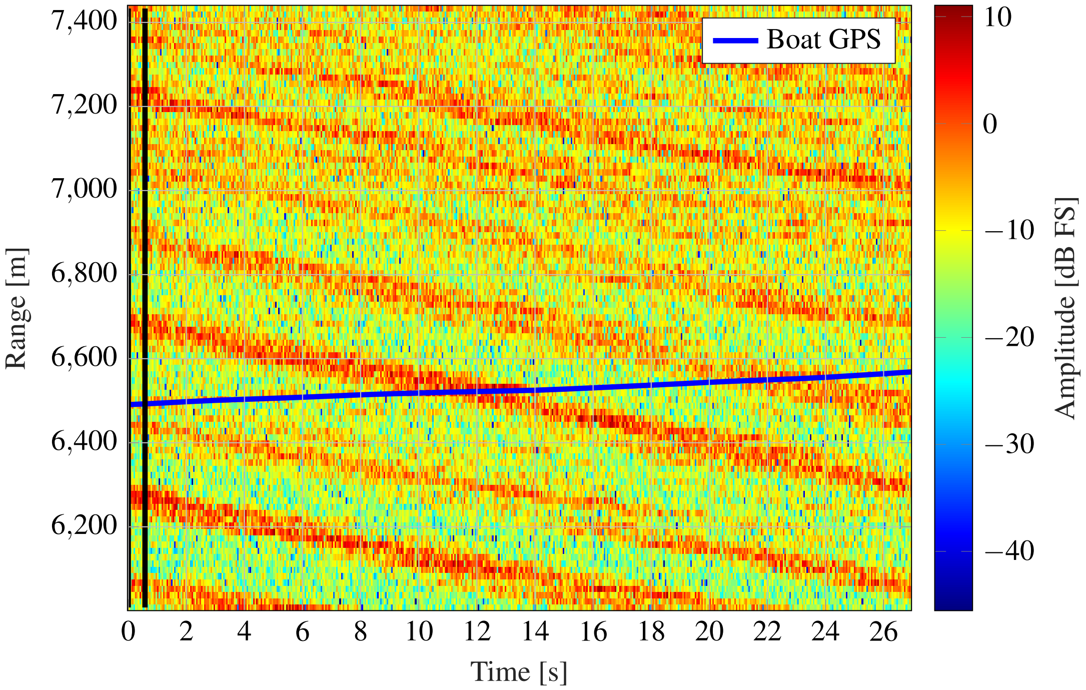
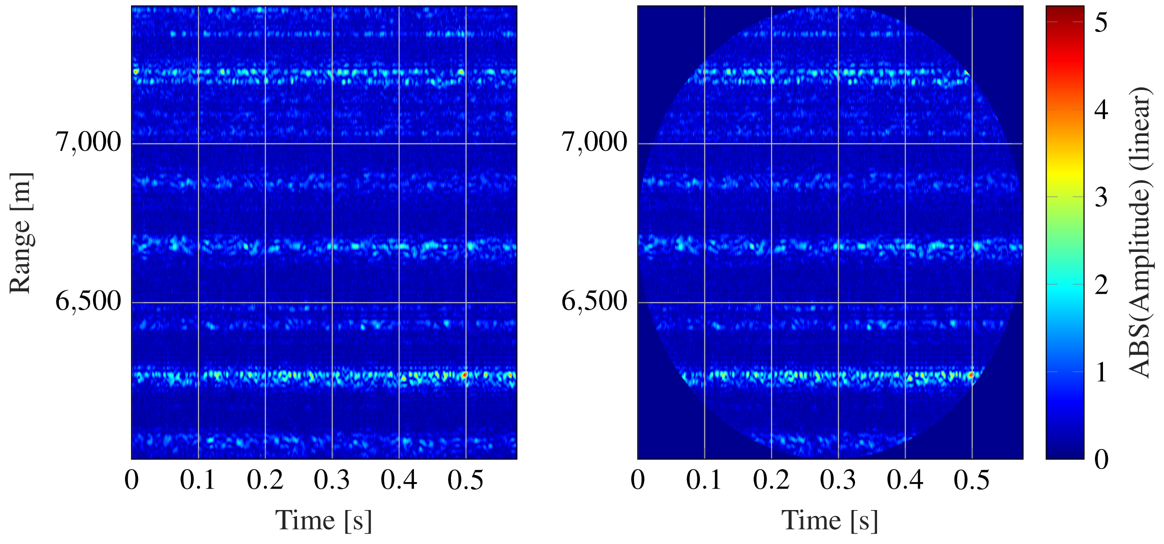


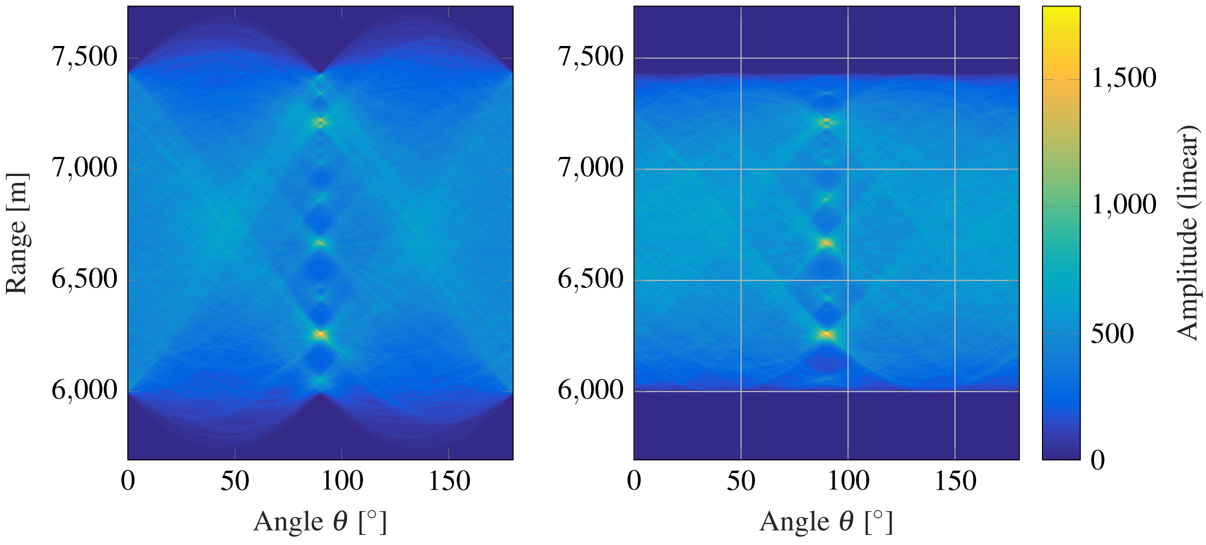
Figure A.46. TFC15_028 range-time plot of 26 s of measured data with boat information given in dB full scale, where the black vertical lines indicate the region used for angle determination.

A.5.2 Radon transform



(a) **(b)**

Figure A.47. TFC15_028 non-circular vs circular input processing matrices for Radon transform, indicating the null data points on the outside of the circular mask.



(a) **(b)**

Figure A.48. TFC15_028 Radon transform result illustrating the output result difference for the (a) non-circular and (b) circular input matrix.

A.5.3 Angle calculation

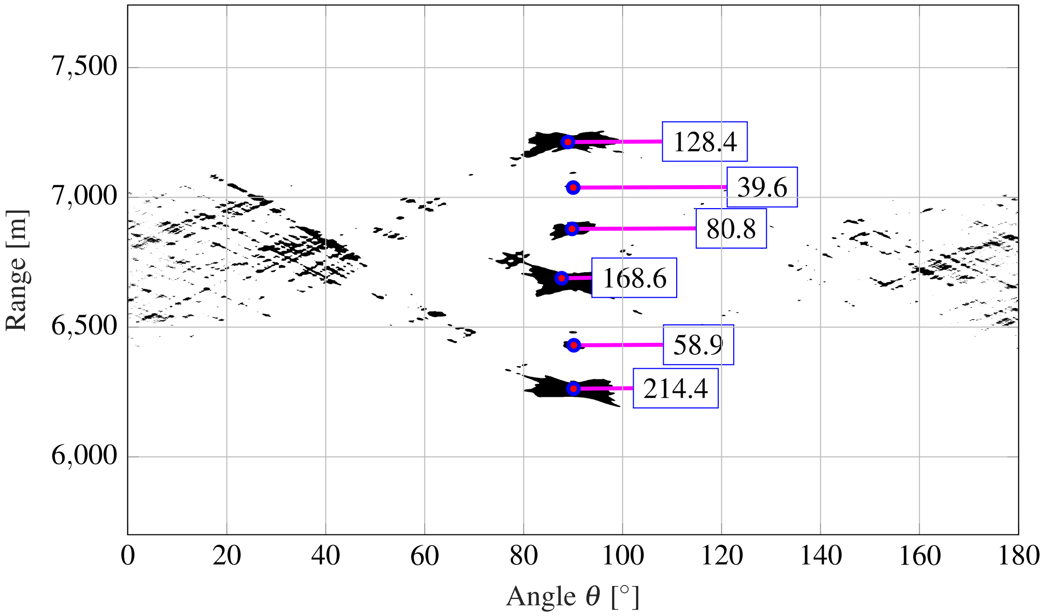


Figure A.49. TFC15_028 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

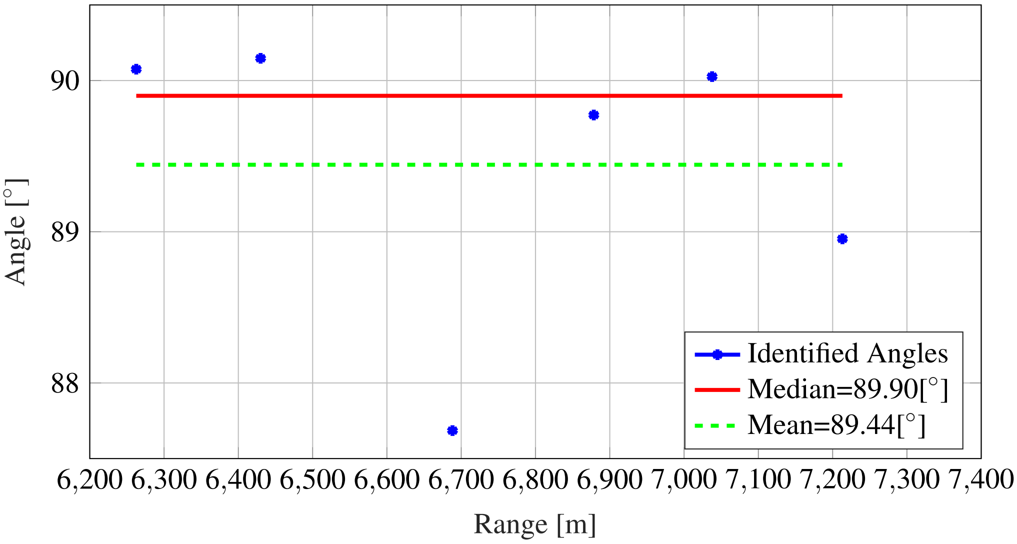


Figure A.50. TFC15_028 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

A.5.4 Wave phase alignment

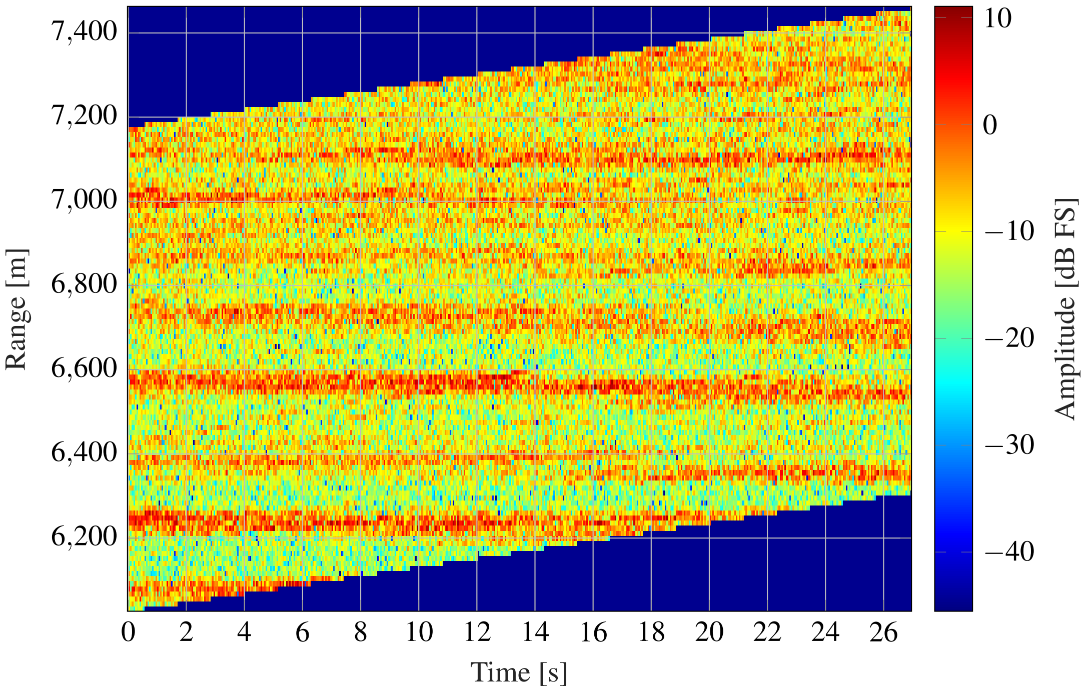
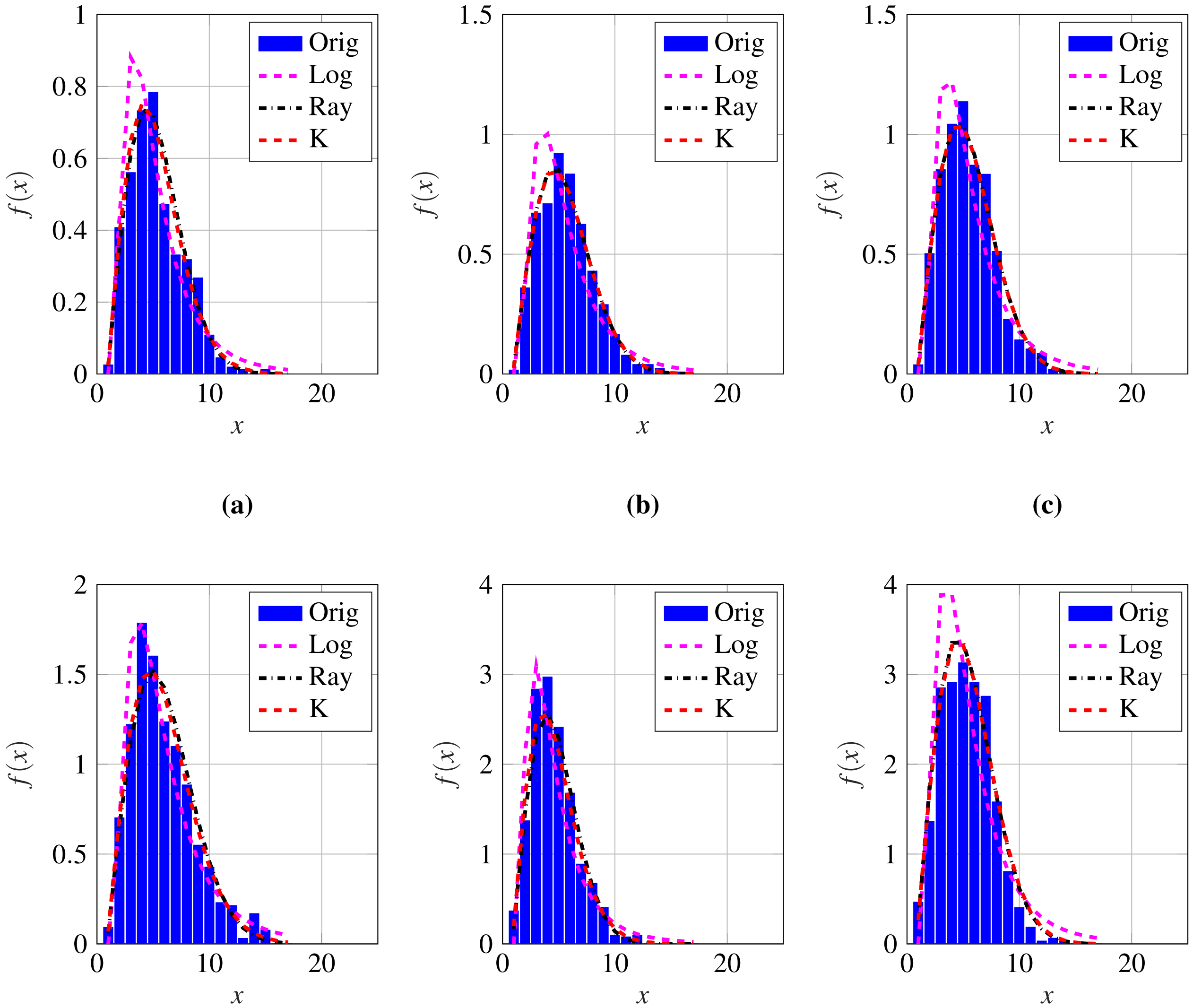


Figure A.51. TFC15_028 output of wave-phase corrected data with *x*-axis aligned waves illustrating clear steps.

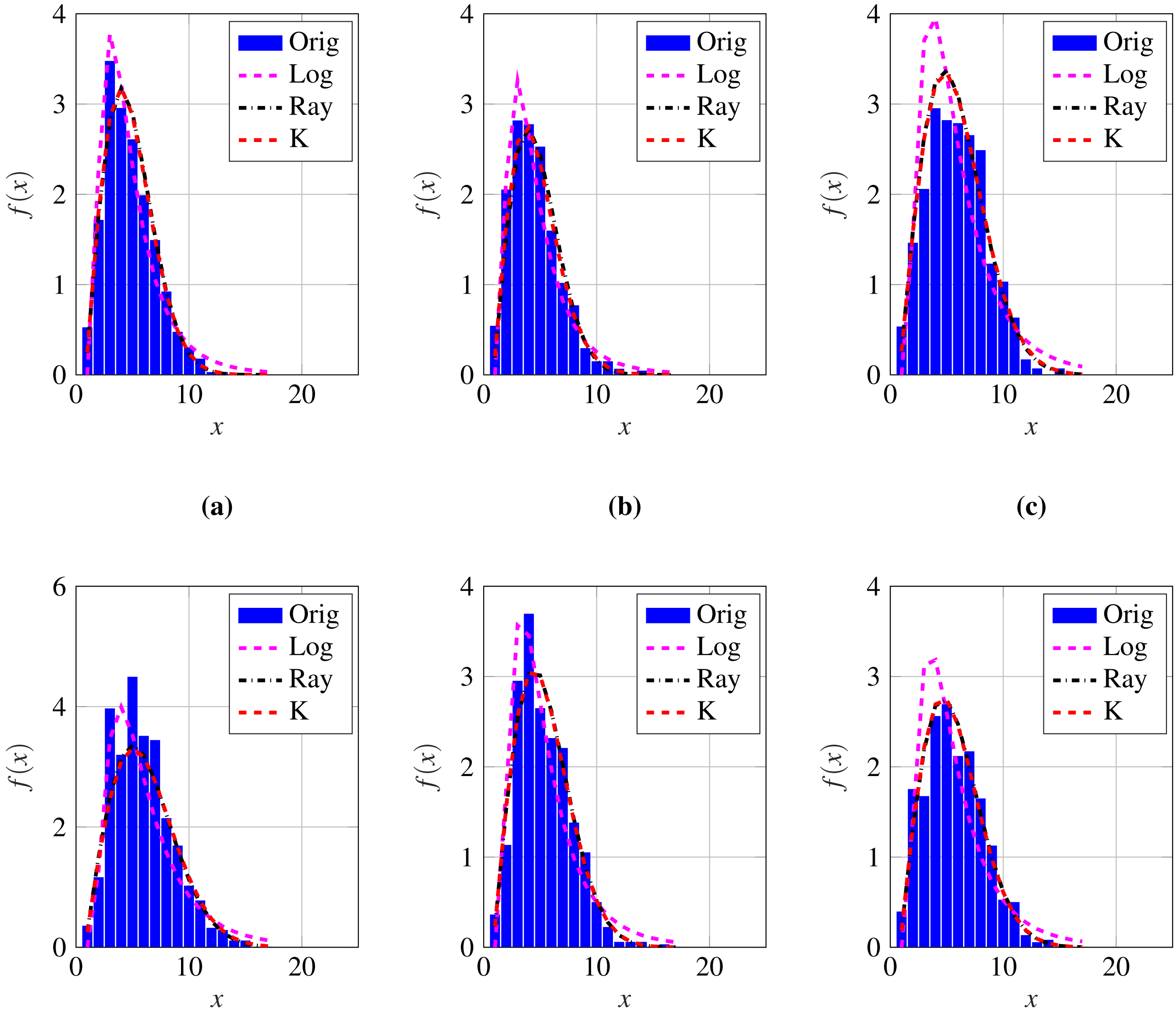
A.5.5 Distribution fitting



(a) (b) (c)

(d) (e) (f)

Figure A.52. TFC15_028 distribution fit example at range bin (a) 44, (b) 45, (c) 46, (d) 47, (e) 48, and (f) 49.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.53. TFC15_028 distribution fit example at range bin (a) 50, (b) 51, (c) 52, (d) 53, (e) 54, and (f) 55.

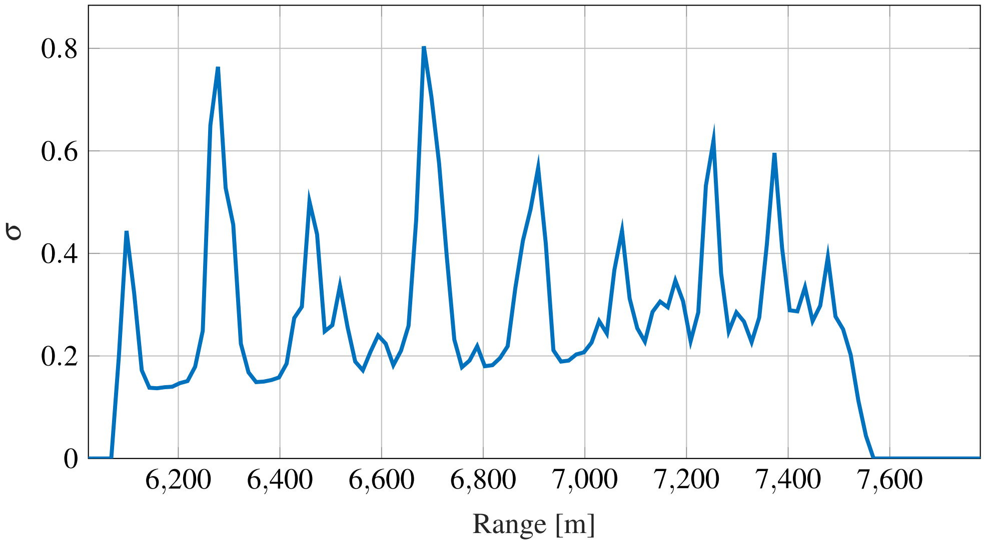


Figure A.54. TFC15_028 PDF fit for Rayleigh distribution illustrating the σ parameter over the entire response range.

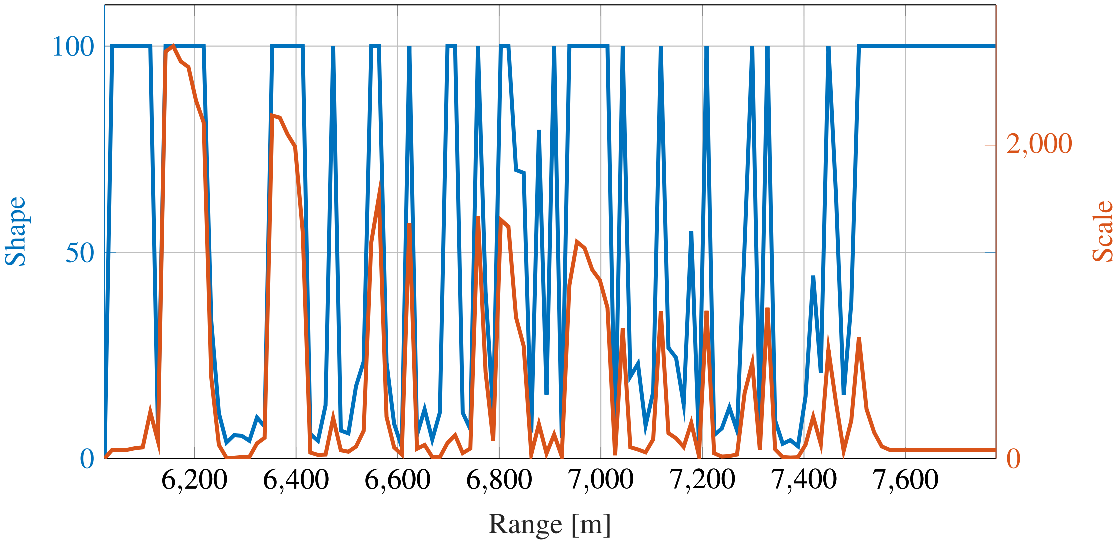


Figure A.55. TFC15_028 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm.

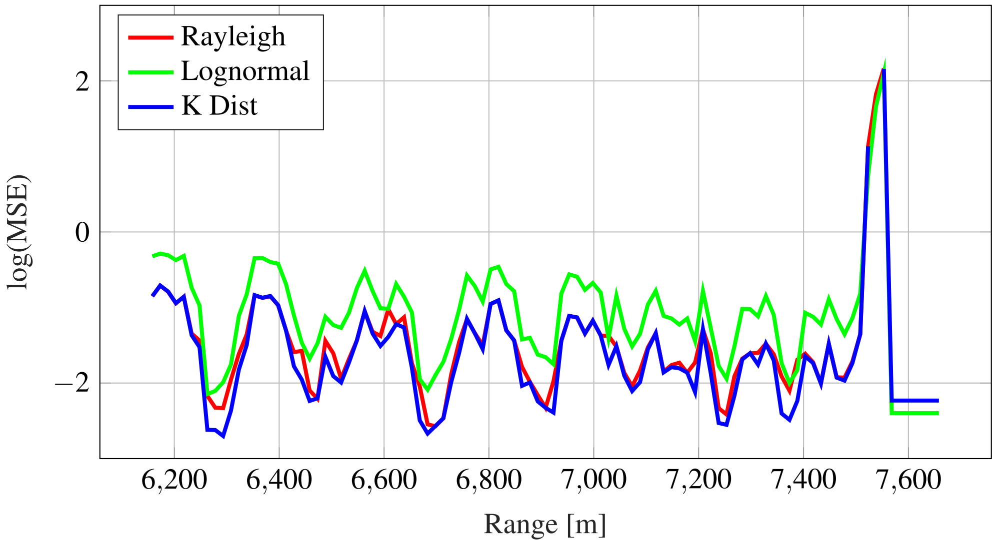


Figure A.56. TFC15_028 parameter MSE comparison of distributions over range in log-domain, illustrating the small error for good fits between parameter peaks.

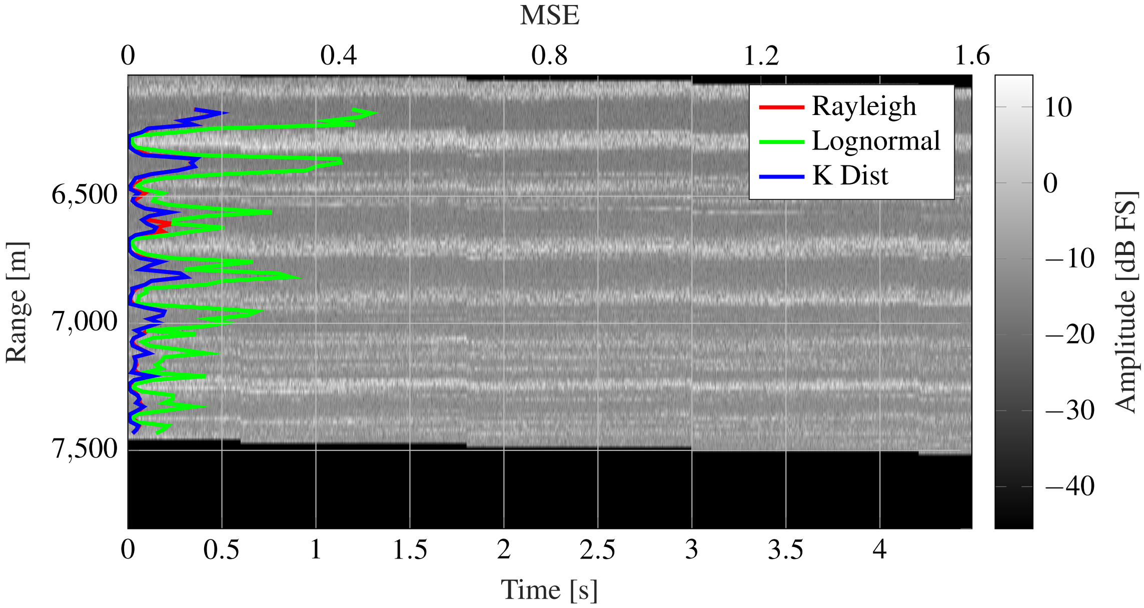


Figure A.57. TFC15_028 comparison illustrating alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.

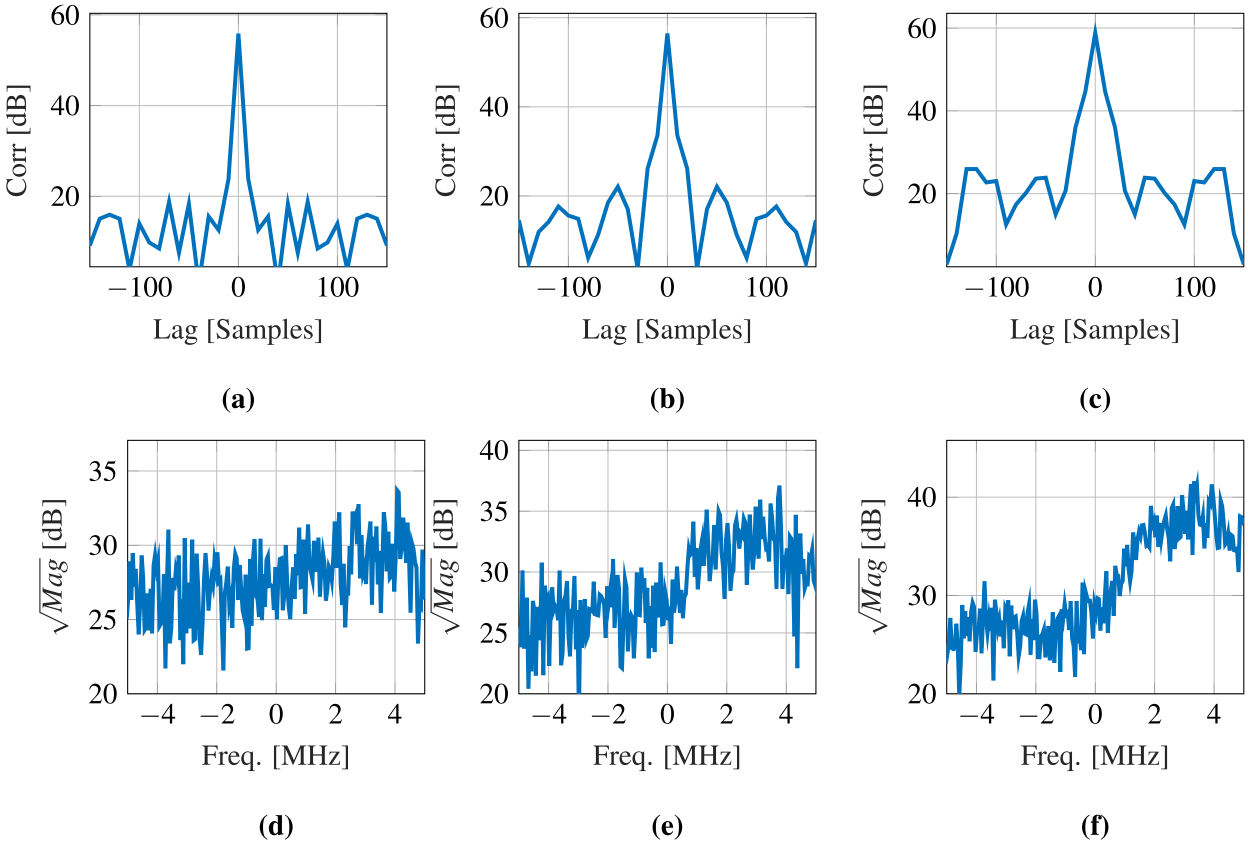


Figure A.58. TFC15_025 correlation at range bin (a) 42, (b) 43, and (c) 44, where (d) to (f) illustrate the corresponding PSD.

A.5.6 Rayleigh correlated output

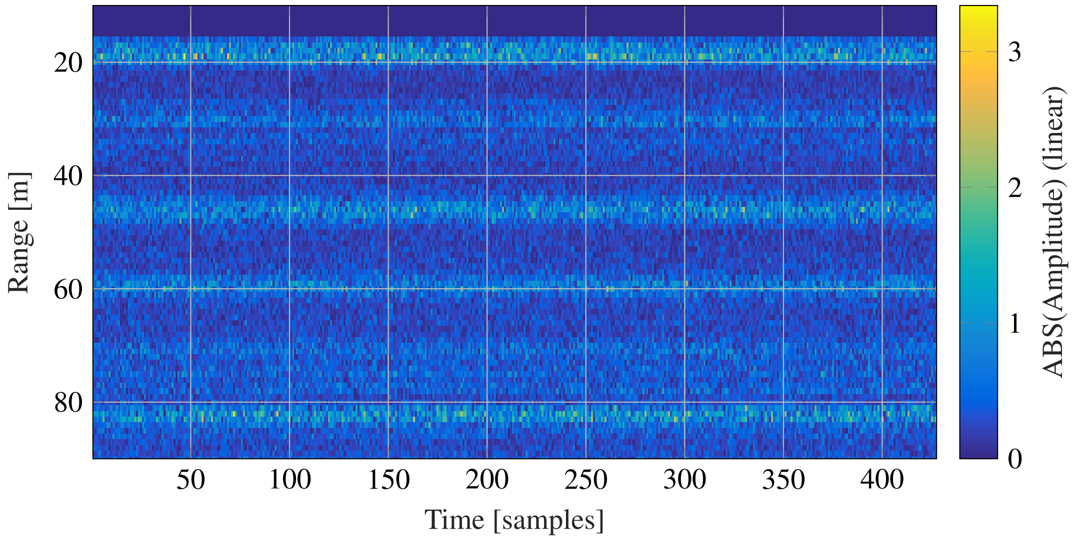


Figure A.59. TFC15_028 range-time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs.

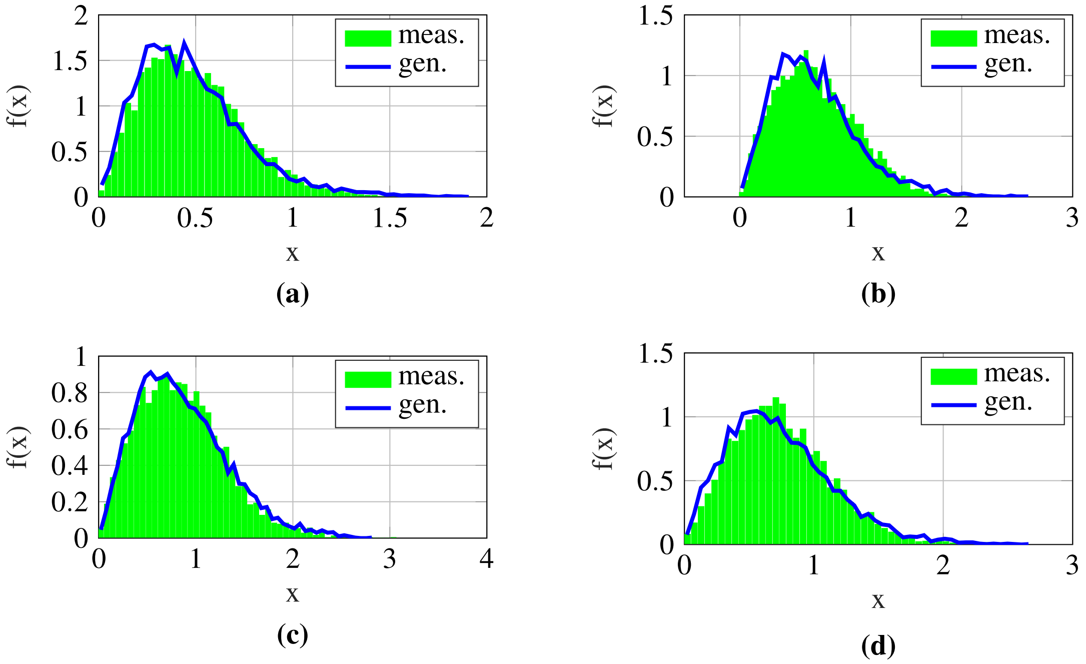


Figure A.60. TFC15_028 Distribution comparison between measured and generated data at range bin (a) 44, (b) 45, (c) 46, and (d) 47.

A.6 RESULTS DATASET 6: TFC15_029

A.6.1 Raw data

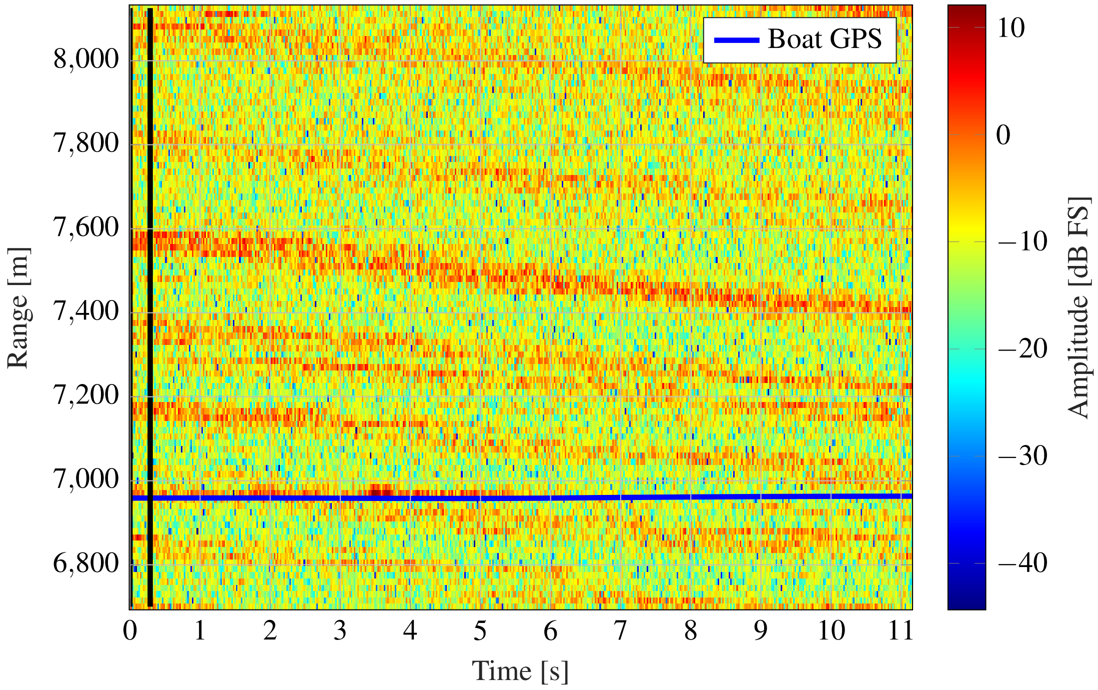
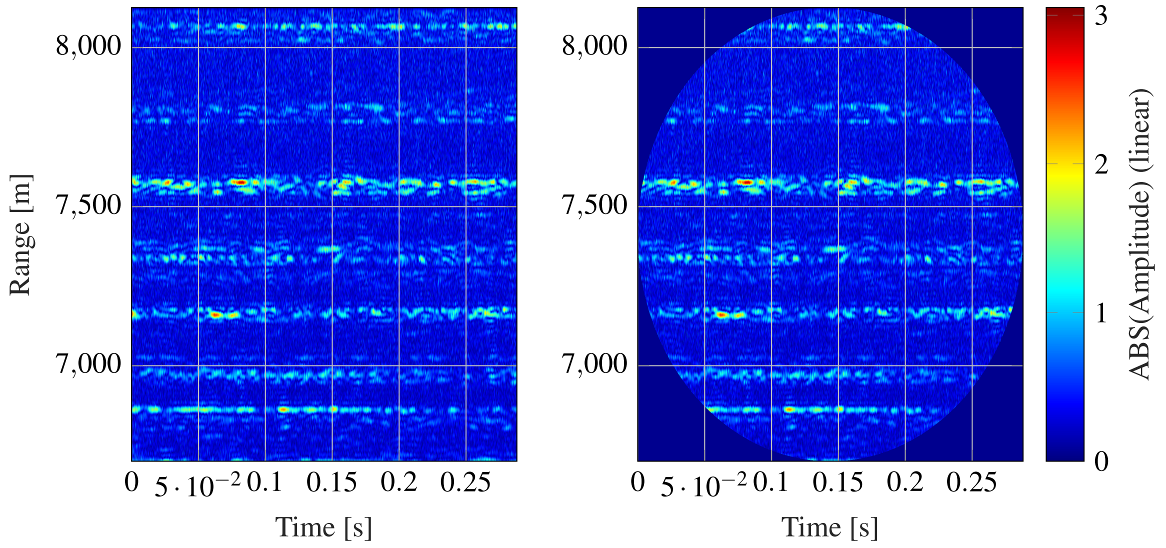


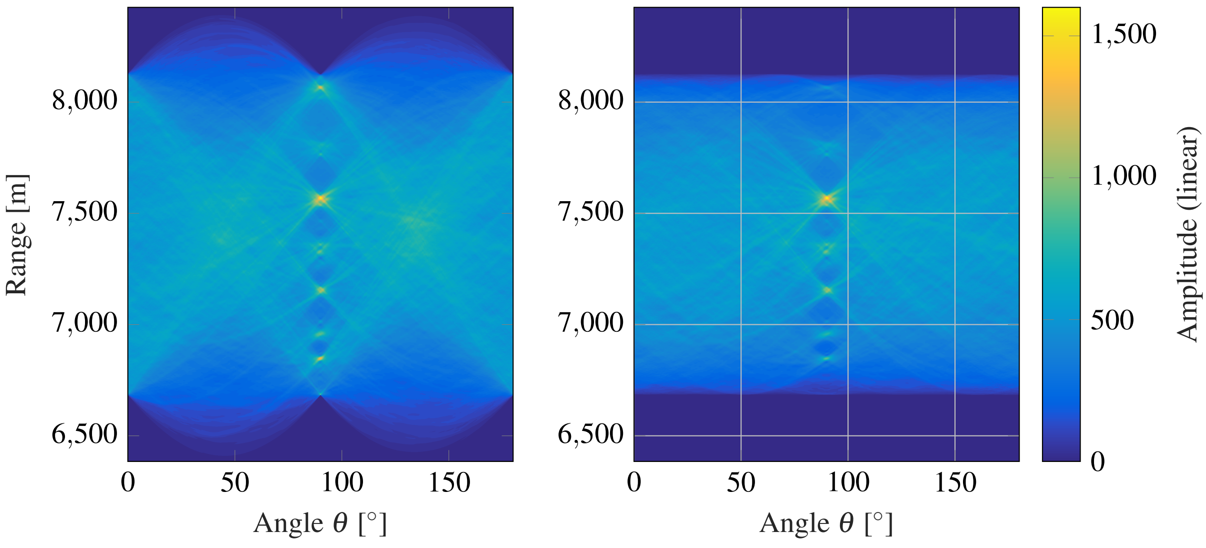
Figure A.61. TFC15_029 range-time plot of 26 s of measured data with boat information given in dB full scale, where the black vertical lines indicate the region used for angle determination.

A.6.2 Radon transform



(a) **(b)**

Figure A.62. TFC15_029 non-circular vs circular input processing matrices for Radon transform, indicating the null data points on the outside of the circular mask.



(a) **(b)**

Figure A.63. TFC15_029 Radon transform result illustrating the output result difference for the (a) non-circular and (b) circular input matrix.

A.6.3 Angle calculation

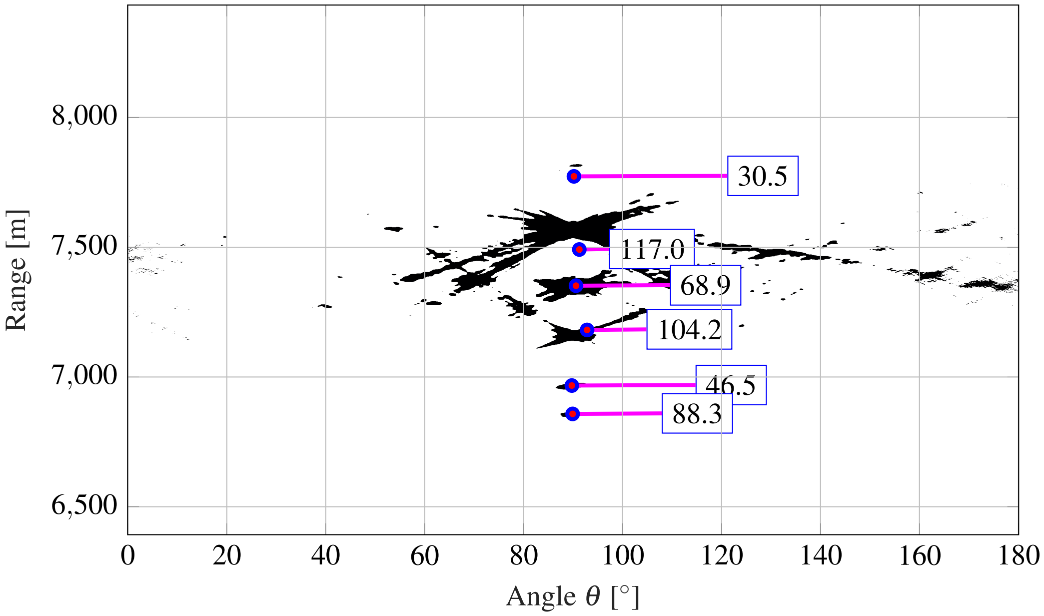


Figure A.64. TFC15_029 blob detection on Radon transform result illustrating the weighted centroids of the underlying data together with their individual amplitude intensities.

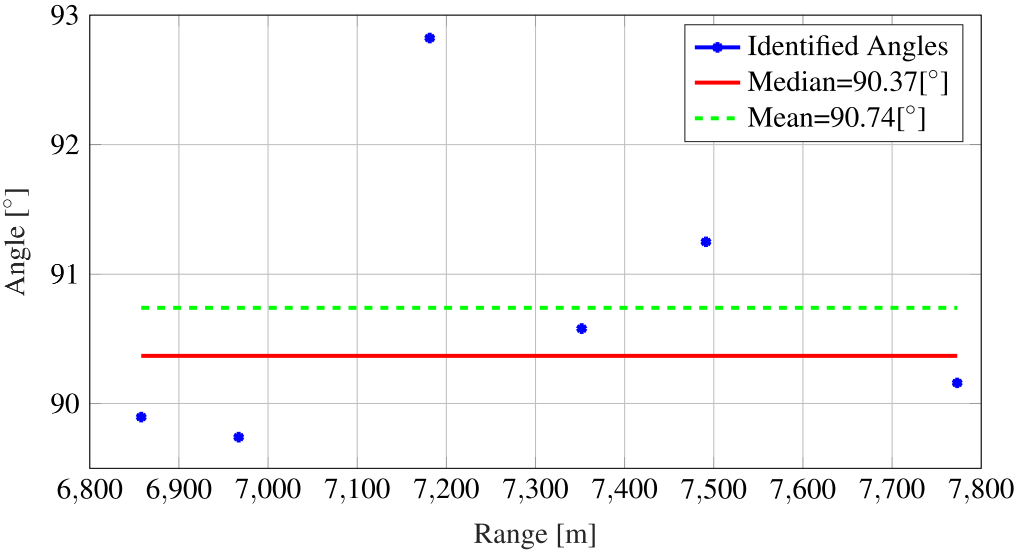


Figure A.65. TFC15_029 mean and median of determined angles, illustrating the advantage of the median to exclude outliers.

A.6.4 Wave phase alignment

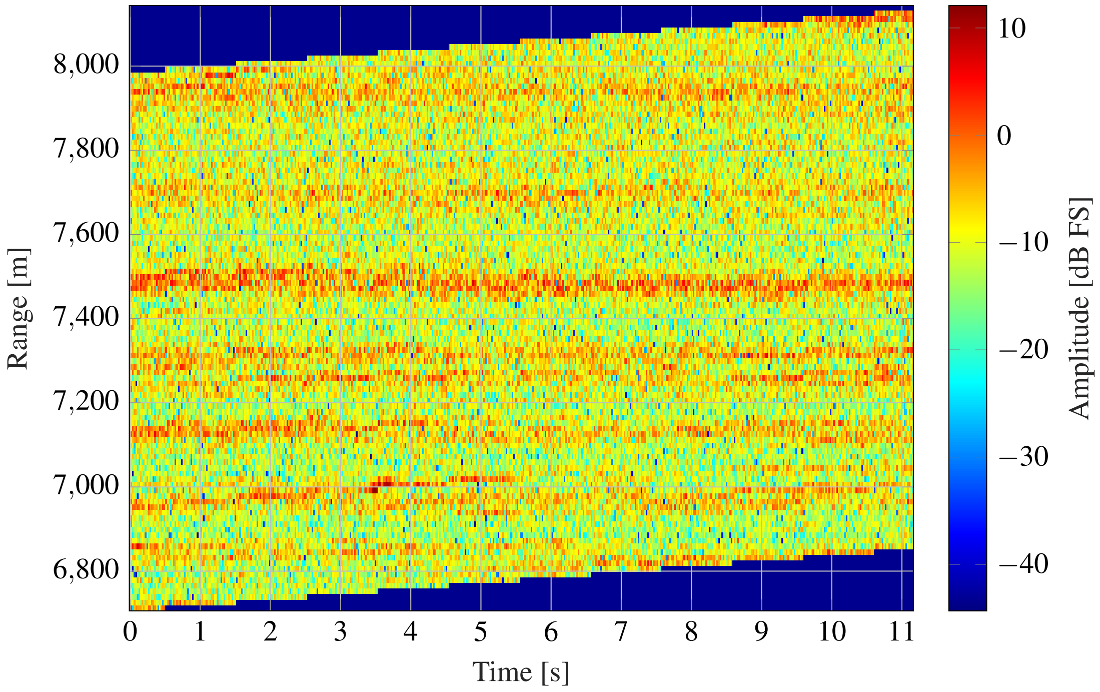
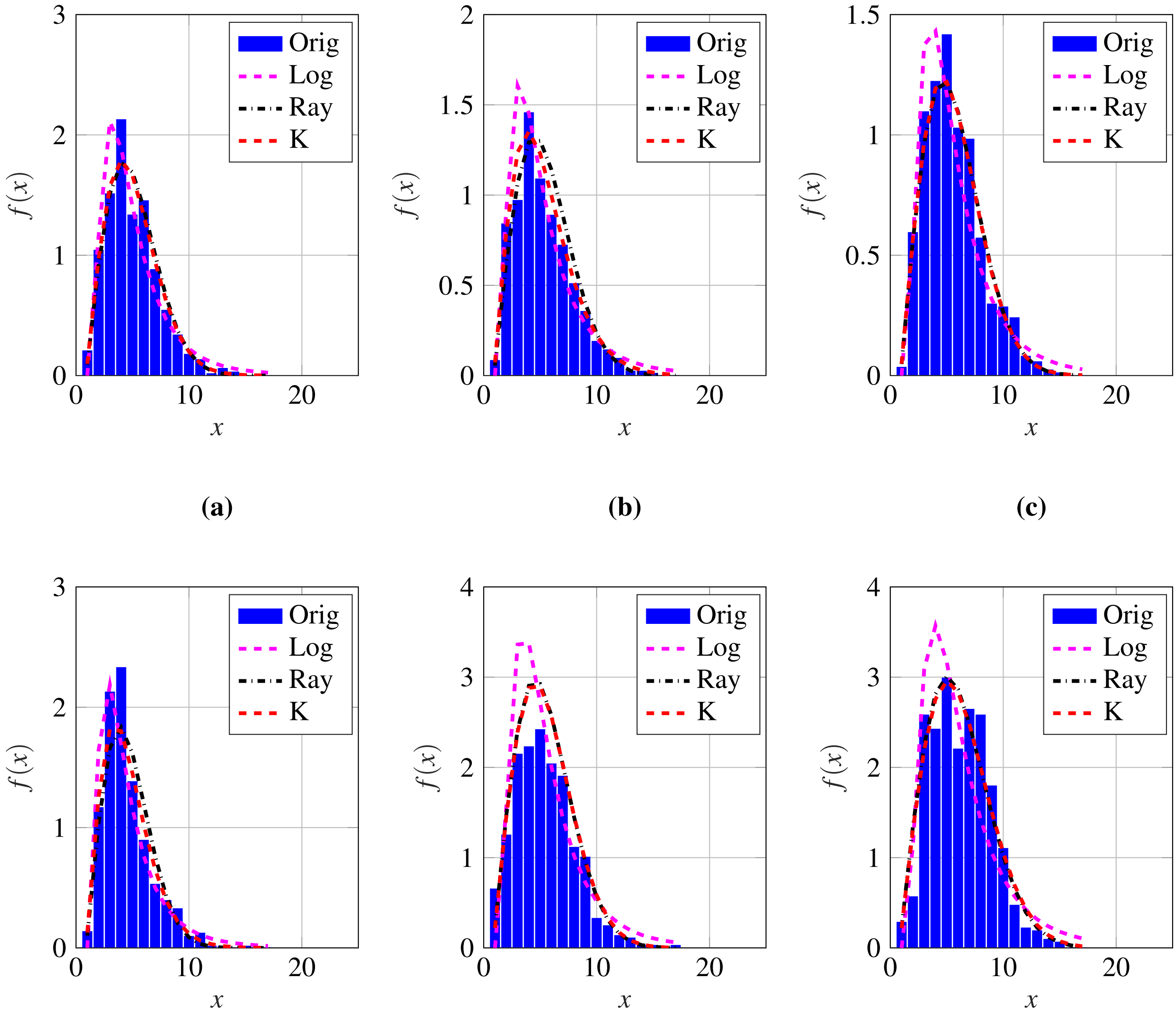


Figure A.66. TFC15_029 output of wave-phase corrected data with *x*-axis aligned waves illustrating clear steps.

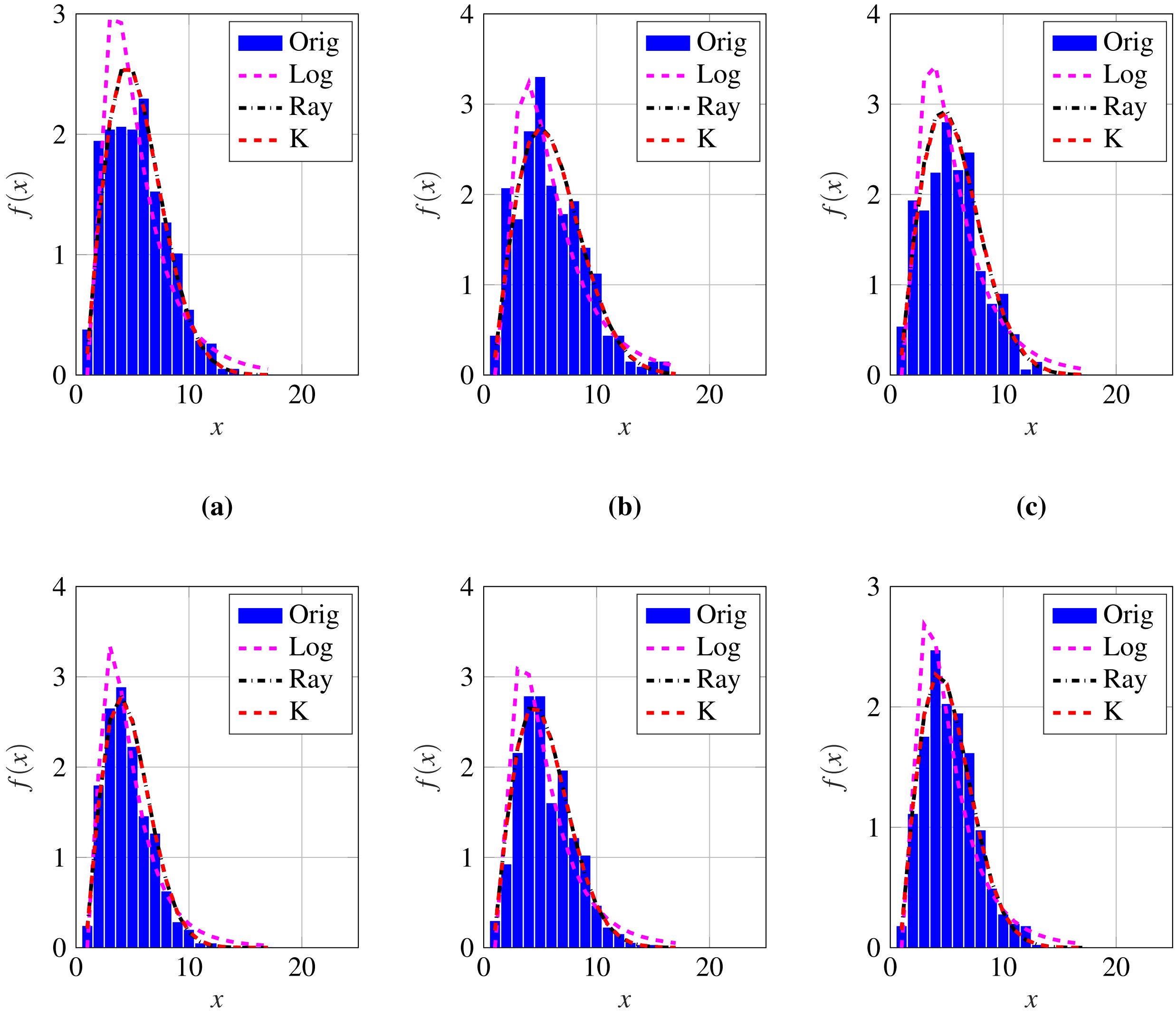
A.6.5 Distribution fitting



(a) (b) (c)

(d) (e) (f)

Figure A.67. TFC15_029 distribution fit example at range bin (a) 44, (b) 45, (c) 46, (d) 47, (e) 48, and (f) 49.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.68. TFC15_029 distribution fit example at range bin (a) 50, (b) 51, (c) 52, (d) 53, (e) 54, and (f) 55.

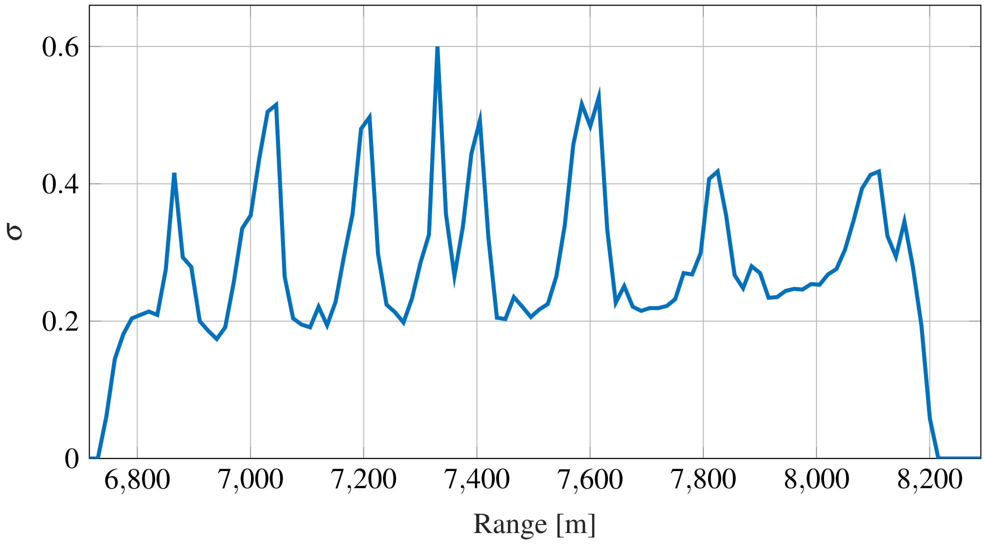


Figure A.69. TFC15_029 PDF fit for Rayleigh distribution illustrating the σ parameter over the entire response range.

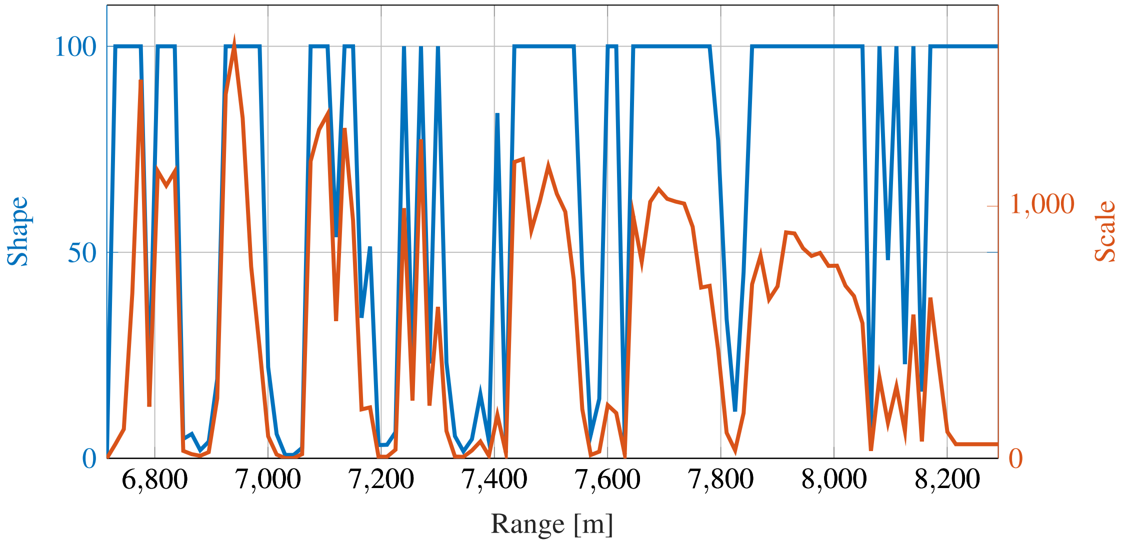


Figure A.70. TFC15_029 K-distribution fit parameters over entire measurement range, where the shape parameter a is limited to 100 digits in the fitting algorithm.

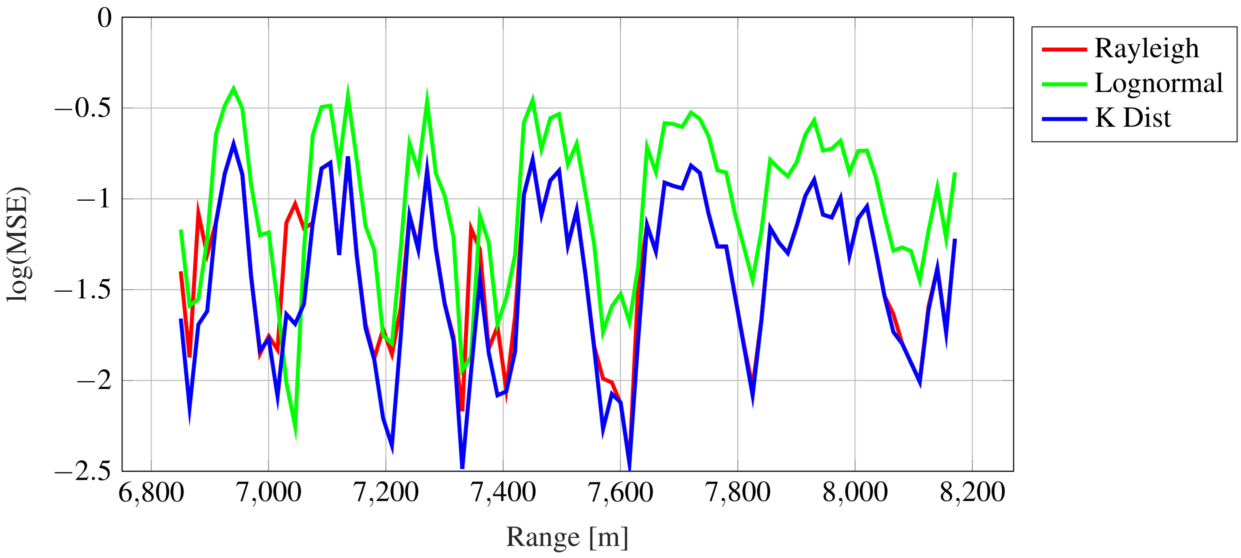


Figure A.71. TFC15_029 parameter MSE comparison of distributions over range in log-domain, illustrating the small error for good fits between parameter peaks.

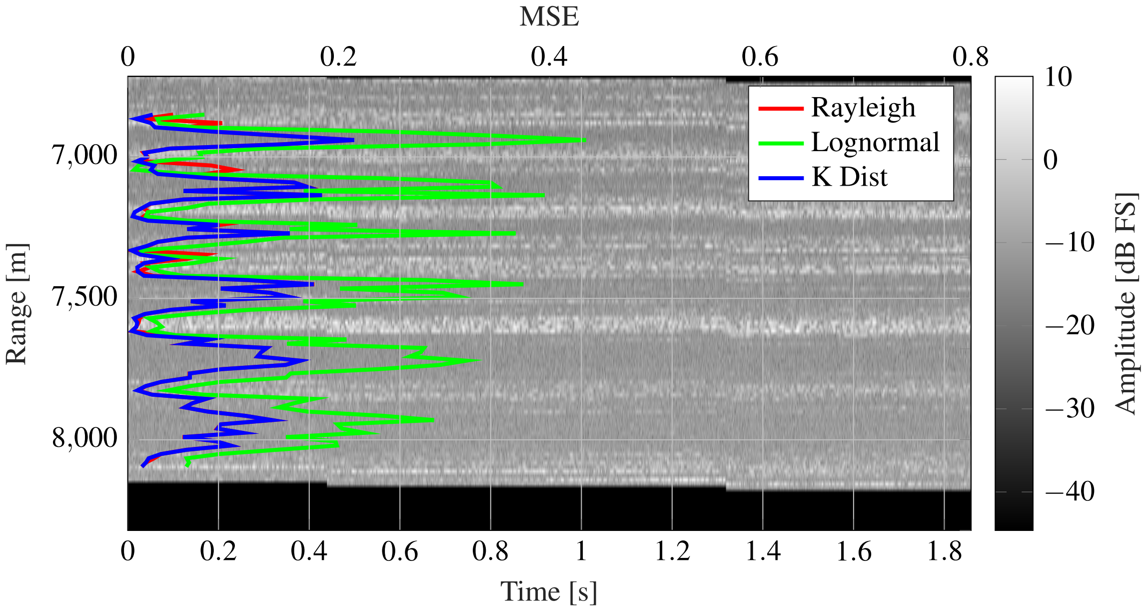
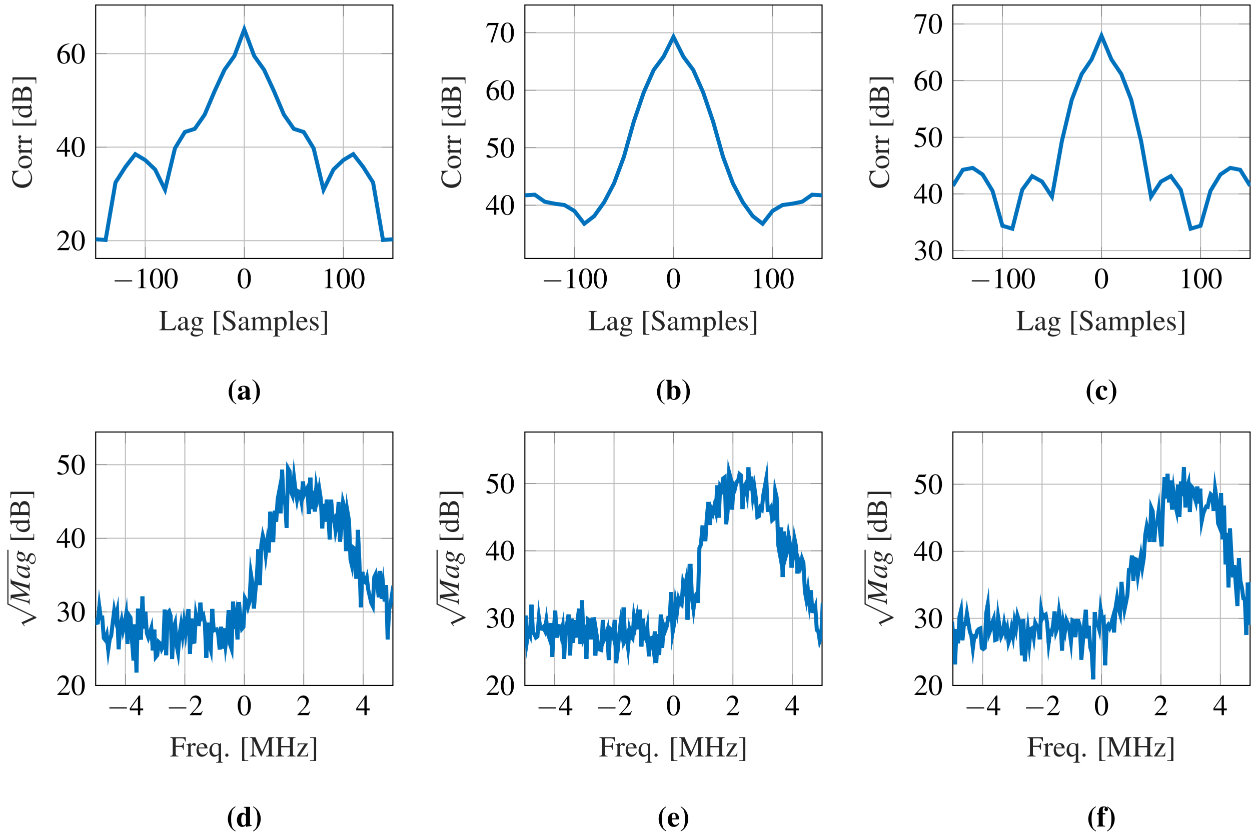


Figure A.72. TFC15_029 comparison illustrating alignment of range-time measured radar data vs distribution mean squared error, illustrating good and poor fit sections.



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure A.73. TFC15_029 correlation at range bin (a) 42, (b) 43, and (c) 44, where (d) to (f) illustrate the corresponding PSD.

A.6.6 Rayleigh correlated output

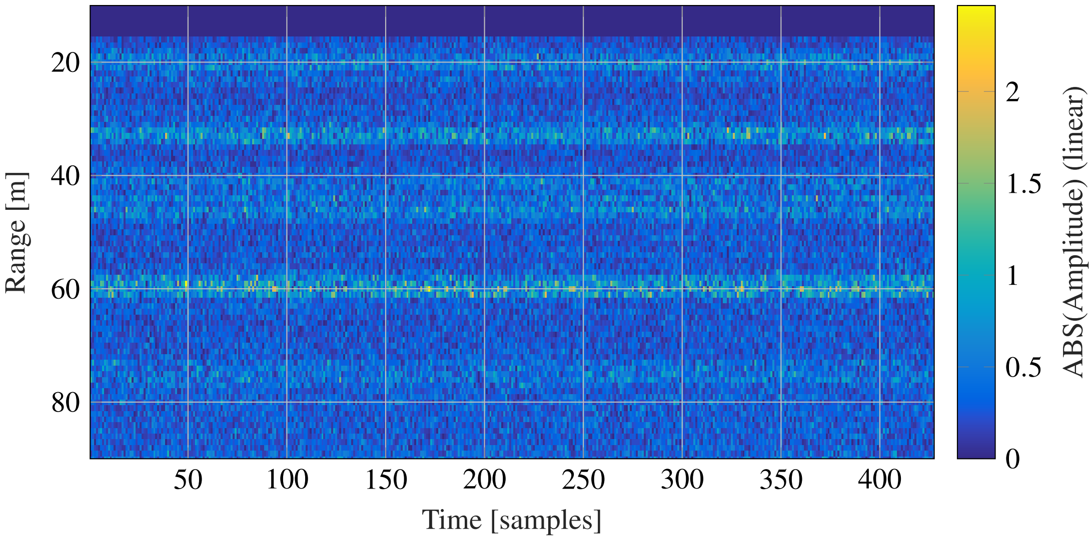
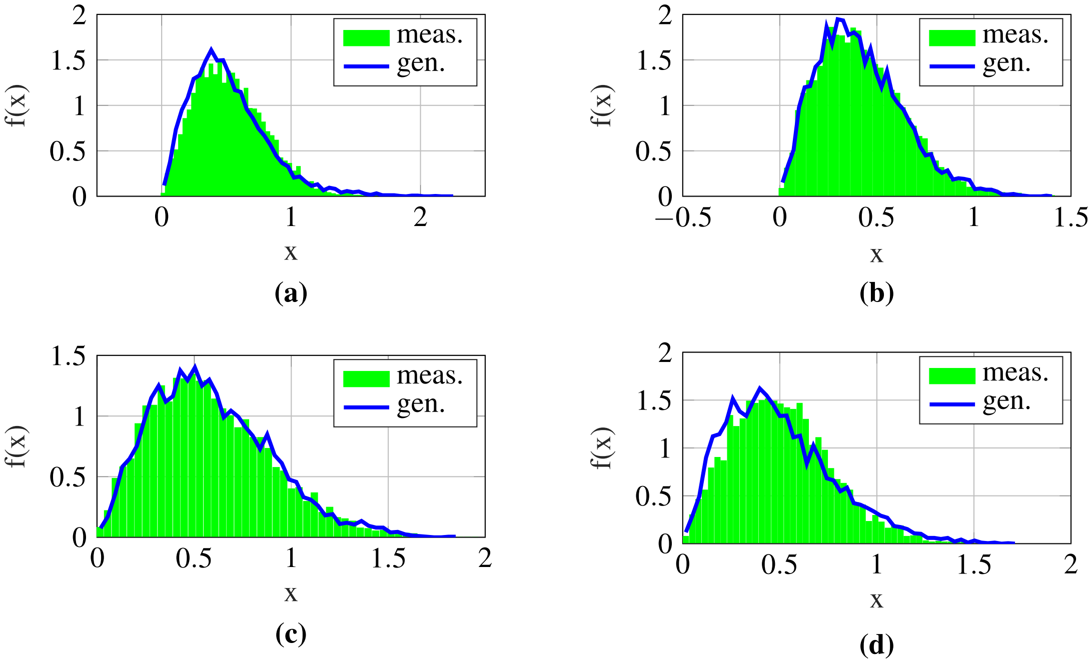


Figure A.74. TFC15_029 range-time plot of generated data based on Gaussian random variables and parametrised filter for Rayleigh correlated outputs.



(a)

(b)

(c)

(d)

Figure A.75. TFC15_029 Distribution comparison between measured and generated data at range bin (a) 44, (b) 45, (c) 46, and (d) 47.