UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

BRUNA GREGORY PALM

DETECTION AND INFERENCES IN NON-GAUSSIAN SIGNALS

Recife
2020

BRUNA GREGORY PALM

DETECTION AND INFERENCES IN NON-GAUSSIAN SIGNALS

Tese apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de Doutor em Estatística. Área de Concentração: Estatística aplicada

Orientador: Dr. Renato J. Cintra

Co-Orientador: Dr. Fábio M. Bayer

Recife

2020

Catalogação na fonte
Bibliotecária Arabelly Ascoli CRB4-2068

P171d Palm, Bruna Gregory
Detection and inferences in non-gaussian signals / Bruna Gregory Palm. – 2020.
157 f.: il. fig., tab.

Orientador: Renato J. Cintra
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN. Estatística. Recife, 2020.
Inclui referências e apêndices.

1. Detecção de mudanças. 2. Estimadores pontuais corrigidos. 3. Modelos de regressão. 4. Imagens SAR. I. Cintra, Renato J. (orientador). II. Título.

310 CDD (22. ed.) UFPE-CCEN 2020-55

BRUNA GREGORY PALM

DETECTION AND INFERENCE IN NON- GAUSSIAN SIGNALS

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Aprovada em: 19 de Fevereiro de 2020.

BANCA EXAMINADORA

Prof.(º) Renato José de Sobral Cintra
UFPE

Prof.(º) Francisco Cribari Neto
UFPE

Prof.(º) Abraão David Costa do Nascimento
UFPE

Prof.(º) Hélio Magalhães de Oliveira
UFPE

Prof.(º) Juliano Bandeira Lima
UFPE

To my parents, Ivalmir and Maria Helena

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisors Prof. Renato Cintra and Prof. Fábio Bayer for their guidance, safe orientation, patience, and valuable advice. You are examples of competence, ethics, and professionals to be followed. Dr. Cintra, from you, I have learned a lot about scientific research and science, and I am very grateful to have had this opportunity. Dr. Bayer, I want to thank you for being my academic model since 2012. If I got here, it is because you have crossed my path.

To my parents Ivalmir e Maria, who always encouraged my love for the study. They gave me the support I needed to be whatever I wanted. To my sister Joana, who undoubtedly is the brightest person in "Gregory Palm" family. Without your support, I wouldn't have come this far. I want to thank my whole family for always being present, in special my cousins Daniela, Débora, Camila, and Marcela. Definitively, you are my safe place.

To Prof. Mats Pettersson and Dr. Viet Thuy Vu for had welcomed me so well in Sweden. Also, I want to thank Prof. Renato Machado for allowing me to do a part of my doctorate at Blekinge Institute of Technology. This period was certainly one of the best of my life.

To my Brazilian and Swedish classmates and friends–you know who you are–for many and many hours of study shared. I am immensely grateful to you for your friendship, for the coffees, for so many good moments shared and memories that would never be erased from my mind.

To Prof. Hélio Magalhães, for all pedagogic and academic knowledge transmitted. Your advises during the teaching internship was of paramount importance to help me to be a better professional.

I thank to professors of the department of statistics at UFPE, especially Prof. Gauss Cordeiro, Prof. Francisco Cribari, Prof. Abraão Nascimento, and Prof. Klaus Vasconcellos for all the academic knowledge transmitted.

To Valéria Bittencourt, secretary of the Graduate Program in Statistics, for the competence, attention, and affection.

Finally, I would like to thank CAPES and CNPq for financial support.

ABSTRACT

Signal detection is a fundamental task in the field of signal and image processing, being pivotal for decision whether a signal is present or identify the different land cover type in synthetic aperture radar (SAR) images. Over the years, detection schemes have been developed assuming the Gaussian distribution. However, in the real world, most of signals are non-Gaussian, and the Gaussianity assumption may not be enough to model several practical contexts. In particular, quantized discrete-time sampled data and amplitude values of a SAR image pixels constitute clear examples of non-Gaussian data. Thus, in this thesis, we derived tools for non-Gaussian signals, such as (i) a new regression model based on the Rayleigh distribution; (ii) bias-adjusted estimators for the Rayleigh regression model parameters; (iii) a new two-dimensional autoregressive moving average model based on the Rayleigh distribution; (iv) a new time series model assuming the beta binomial distribution; and (v) the use of a stack of SAR images to obtain a ground scene prediction (GSP) image. The proposed Rayleigh regression model was applied in detection schemes of land cover type in SAR images and the obtained results were compared to the measurements from Gaussian-,Gamma-, and Weibull-based regression models. The Rayleigh regression model was the only model that could detect the difference among the three tested regions. The two-dimensional Rayleigh autoregressive moving average model were applied to detect changes in SAR images. For comparison purposes, we also obtained the detection results based on the two-dimensional Gaussian model. The proposed method detected 24 in a total of 25 military vehicles, while the Gaussian-based scheme detected only 16 military vehicles. The derived beta binomial autoregressive moving average model was employed in nonrandom signals detection showing a higher probability of detection and a lower probability of false alarm in comparison to the traditional Gaussian based methods. The obtained GPS image based on the median method was considered in a change detection algorithm displaying a probability of detection of 97% and a false alarm rate of $0.11/\text{km}^2$, when considering military vehicles concealed in a forest.

Keywords: Change detection. Improved point estimators. Regression models. SAR images. Two-dimensional models.

RESUMO

Em processamento de sinais e de imagens, detecção é um problema amplamente discutido na literatura, seja para detectar a presença de um sinal ou para identificar o tipo de solo em uma imagem de radar de abertura sintética (SAR). Ao longo dos anos, os métodos de detecção foram desenvolvidos assumindo distribuição gaussiana. Entretanto, em situações reais, os sinais são não gaussianos. Dois típicos exemplos de sinais tipicamente não gaussianos são os sinais digitais e os valores de amplitude em uma imagem SAR. Desta forma, na presente tese, são derivadas ferramentas para sinais não gaussianos, tais como: (i) um novo modelo de regressão baseado na distribuição Rayleigh; (ii) estimadores corrigidos para os parâmetros do modelo de regressão Rayleigh proposto; (iii) um novo modelo autorregressivo de médias moveis bidimensional baseado na distribuição Rayleigh; (iv) um novo modelo de séries temporais assumindo a distribuição beta binomial e (v) o uso de um pacote de images SAR para obter uma previsão sobre o verdadeiro terreno das imagens. O modelo de regressão proposto foi considerado em detecção do tipo de solo em images SAR e os resultados obtidos foram comparados com os modelos baseados nas distribuições gaussiana, gama e Weibull. O modelo de regressão Rayleigh foi o único modelo capaz de detectar diferença no tipo de solo das três áreas testadas. O modelo bidimensional proposto foi empregado na detecção de mudança em images SAR, e os resultados de detecção baseados no modelo bidimensional Gaussiano foram utilizados como critério de comparação. O modelo proposto detectou 24 dos 25 veículos militares presentes na imagem SAR, enquanto que o modelo Gaussiano detectou apenas 16 alvos. Ainda, o modelo beta binomial autorregressivo de média móveis derivado foi empregado em detecção de sinais não aleatórios apresentando maiores valores de probabilidade de detecção e menos taxas de falso alarme em comparação aos tradicionais métodos de detecção baseados na distribuição Gaussiana. Finalmente, a imagem predita baseada no método da mediana obtida considerando um pacote de imagens SAR foi utilizada em um algoritmo de detecção de mudanças apresentando probabilidade de detecção de veículos militares de 97% e taxa de falso alarme de $0.11/\text{km}^2$.

Palavras-chave: Detecção de mudanças. Estimadores pontuais corrigidos. Modelos de regressão. Imagens SAR. Modelos bidimensionais.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

T_L	Likelihood ratio test
T_R	Rao test
T_W	Wald test
βARMA	Beta autoregressive moving average
$\sqrt{\text{MSE}}$	Root mean squared error
P_d	Probability of detection
2D ARMA	Two-dimensional autoregressive moving average
2D RARMA	Two-dimensional Rayleigh autoregressive moving average
ACF	Autocorrelation function
AIC	Akaike's information criteria
AR	Autoregressive
ARMA	Autoregressive moving average
BBARMA	Beta binomial autoregressive moving average
BFGS	Broyden-Fletcher-Goldfarb-Shanno
CDA	Change detection algorithms
CI	Confidence interval
CMLE	Conditional maximum likelihood estimator
CR	Coverage rates
FAR	False alarm rate
GARMA	Generalized autoregressive moving average
GLM	Generalized linear models
GSP	Ground scene prediction
HH	Horizontal
HQ	Hannan and Quinn's information criteria
HV	Horizontal and vertical
IRBSN	Integrated Relative Bias Squared Norm
MA	Moving average
MAPE	Mean absolute percentage error
MdAE	Median absolute error
ML	Maximum likelihood
MLEs	Maximum likelihood estimators

MSE	Mean square error
pdf	Probability density function
pf	Probability function
RB%	Percentage relative bias
ROC	Receiver operating characteristic
SAR	Synthetic aperture radar
SIC	Schwartz's information criteria
UWB	Ultrawideband
VHF	Very-high frequency
VV	Vertical

CONTENTS

1 INTRODUCTION

In this chapter, we provide motivation and framework for the research topics investigated in this thesis. Additionally, we detail the thesis main goals and the document structure.

1.1 MOTIVATION

Signal detection is a major area of research in signal and image processing [3, 4]. In particular, the problem of detection whether a signal is present or not is a fundamental task in the field, being important for decision making and information extraction [4]. Over the years, detection schemes have been developed assuming that the additive noise is a Gaussian process [5, 6]. However, in the real world, signals are often non-Gaussian [7]. Therefore, the use of signal processing algorithms developed for Gaussian noise environments in situations where non-Gaussian data is present, typically result in significantly worse performance [8–12]. An example of a typical non-Gaussian signal is the quantized signal—here referred to as 'digital signals', since its signal amplitude is a discrete value [13]. Another scenario in which assuming Gaussianity may not be suitable is in the synthetic aperture radar (SAR) image processing [14].

In the context of SAR images processing, a usual problem is the identification and classification of land cover type in images or distinct targets [3, 15]. Frequently, the SAR image modeling is performed assuming constant parameters [16–18], i.e., considering homogeneity in the images. In cases where this assumption is not suitable, an alternative is to use a regression model, where each observation has one specific estimated mean. For this purpose, the Gaussian regression model is widely used [19–21]. However, regression models adopting suitable distributions, such as the Rayleigh distribution, can be used without assuming homogeneity in the images, generating accurate results in the detection of different land cover types from an image. The Rayleigh distribution is commonly used in signal and image processing, as demonstrated in [16, 17, 22–28]. Thus, a regression model assuming the Rayleigh distribution can be applied in the detection scheme of SAR images, yielding accurate results.

The Rayleigh distribution is widely considered in signal and image processing [17,24–28], being important in the context of synthetic aperture radar (SAR) image modeling, due to its good characterization of image pixel amplitude values [16–18, 29]. In the signal processing context, this distribution captures power variations on a wavelength scale and it is capable of modeling the measured amplitude of the received signal for systems with relatively large cells, called

macrocellular systems, and small-scale multipath fading [30]. Thus, the Rayleigh distribution is an important model in the context of image and signal processing, and consequently, an important topic of research.

The inferences of the Rayleigh regression model parameters are based on maximum likelihood estimation. The maximum likelihood estimators (MLEs) have good asymptotic properties in large signal lengths. However, in small signal lengths, the biases of the MLEs are on the order of N^{-1} [31], where N is the signal length. Thus, in small signal lengths, bias can become problematic [32]. Therefore, inferential corrections for small signal lengths become an important topic of study.

In SAR images, the content of information between pixels takes different levels of dependence, such as resolution spatial and temporal [18, 33]. Thus, a two-dimensional autoregressive moving average (2D ARMA) model can be used as a venue for dealing with such a problem. Indeed, the parametric representation of two-dimensional homogeneous random fields considering 2D ARMA models is frequently used in the literature for image processing [34, 34–48]. In special, the two dimensional autoregressive first-order model is commonly employed to represent image real scenarios [49, 50], being capable of representing different types of textures [34]. This model is an extension of the traditional autoregressive moving average (ARMA) model, widely used in time series analysis [43]. A two-dimensional autoregressive moving average model assuming the Rayleigh distribution can be used in the detection scheme of SAR images, yielding accurate results.

Additionally, digital signals can be defined as a discontinuous waveform with a finite range of levels [51]. The beta binomial distribution has been employed over the years to model bounded discrete values, as shown in [52–55]. Thus, the beta binomial distribution can be considered as a good candidate to fit digital signals. In addition, as signals can be defined as a function of time [56] and the signal amplitude of the digital signals is a discrete value [13], a time series model assuming the beta binomial distribution can be employed as a venue for developing detection methods tailored for such type of data. Besides quantized signal processing applications, the derived model can be used to fit any bounded count data observed over the time, such as the number of rainy days per time interval [57], number of defective products in one lot [58], or number of hospital admissions [59].

Change detection algorithm (CDA) has been widely considered over the years in the detection of distinct targets in SAR images [60–62]. In particular, the CDA is used to detect changes in a ground scene between distinct measurements in time, such as natural disasters like

flood and wildfire or man-made interference [63–65]. Generally, CDA can be simply obtained by the subtraction of two images (reference and surveillance), followed by a thresholding operation. However, an image stack can be considered instead of just two images in a CDA; such collection of images leads to improved detection performance, as discussed in [1]. An image stack is composed by images with similar heading and incidence angle of the illuminating platform [1]. Thus, the backscattering of SAR images in the stack is stable in time, i.e., the ground scene should be very similar, excepted by the changes. This way, a ground scene prediction (GSP) can be obtained based on the image stack. This new image can be used as a reference image in CDA yielding in high probability of detection and low false alarm rate.

1.2 MAIN GOALS

The main goal of this thesis is to address the open gaps in the literature in terms of (i) detection of different land cover types from a SAR image; (ii) improvement of point estimation; (iii) SAR image modeling; (iv) detection problem of signals embedded in noise from a quantized measured signal; and (v) improvement of CDA performance in a CDA framework.

Specifically, we aim at:

- proposing a regression model for non-Gaussian signals, assuming the Rayleigh distribution;
- introducing a change detector for the amplitude values of non-Gaussian SAR images;
- using corrections schemes to obtain bias-adjusted estimators for the Rayleigh regression model;
- deriving a two-dimensional Rayleigh autoregressive moving average (2D RARMA) model;
- introducing a ground type change detector for non-Gaussian SAR images based on the residual-based control charts of the two-dimensional model;
- proposing a modeling tool for non-Gaussian SAR images;
- deriving the beta binomial autoregressive moving average model (BBARMA) for non-Gaussian signals;
- introducing a signal detector based on the asymptotic properties of the BBARMA model parameter estimators;
- proposing a ground scene prediction SAR image based on an image stack;
- applying GSP imagery as input data in CDA.

1.3 ORGANIZATION OF THE THESIS

This thesis contains seven self-contained chapters, each with its own notation and terminology. Each chapter corresponds to the research progress obtained by the candidate as of the writing of this thesis.

In Chapter 2, we introduce a regression model for nonnegative signals. The discussed regression estimates the mean of Rayleigh distributed signals. For the derived model, we introduce parameter estimation, large data record results, and goodness-of-fit measures. Additionally, a change detector for the amplitude values of SAR images based on the asymptotic properties of the proposed Rayleigh regression model parameter estimators is discussed. The performance of maximum likelihood estimators of the derived model were evaluated using Monte Carlo simulations. Also, we present an application of the proposed detector in SAR image.

In Chapter 3, we present accurate point estimation strategies for the parameters of the Rayleigh regression model. In particular, the Cox and Snell's [66], Firth's [67], and parametric bootstrap [68] methods were considered to obtain bias-adjusted estimators. The performance of maximum likelihood estimators proposed were evaluated using Monte Carlo simulations.

In Chapter 4, we propose a two-dimensional ARMA model based on the Rayleigh distribution for non-Gaussian situations, where the observed output signal is asymmetric and measured continuously on the real positives values. For the derived model, we present conditional parameter estimation, large data record inference, a modeling tool, and residuals. Additionally, we introduce a ground type change detector for non-Gaussian SAR images based on control charts of the proposed spatial model residuals.

In Chapter 5, we discuss a new time series model for bounded discrete time values. This model can be applied for detection and inference in digital signals, estimating the mean of a beta binomial distributed variable observed over the time. For the introduced model, parameter estimation, large data record results, forecasting tools and diagnostic measures are discussed. In addition, we present a signal detector based on the asymptotic properties of the proposed model parameter estimators. The performance of the derived detector and of the model parameter estimators were evaluated using Monte Carlo simulations.

In Chapter 6, we consider traditional statistical techniques, such as autoregressive models, robust mean, median, intensity mean, and mean in a stack of SAR images to obtain a ground scene prediction. We evaluate the ground scene prediction images based on the presence of the true ground scene without change and checking if the predicted image preserves the ground

backscattering statistics of the images in the stack. The GSP image is considered as a reference image in a CDA, presenting competitive performance when compared with other recent results. Finally, in Chapter 7, some conclusions and ideas for further works are presented.

1.4 COMPUTATIONAL SUPPORT

All simulations results we present in this thesis were obtained using the Matlab software [69]. The plots from Chapter 2, 3, and 5 were produced using the `R` language [70]. In Chapters 4 and 6, the Matlab software [69] was considered to generate the images.

2 RAYLEIGH REGRESSION MODEL FOR GROUND TYPE DETECTION IN SAR IMAGERY

Abstract

This chapter proposes a regression model for nonnegative signals. The derived regression estimates the mean of Rayleigh distributed signals by a structure which includes a set of regressors and a link function. For the proposed model, we present: (i) parameter estimation; (ii) large data record results; and (iii) a detection technique. In this chapter, we present closed-form expressions for the score vector and Fisher information matrix. The introduced model is submitted to extensive Monte Carlo simulations and to measured data. The Monte Carlo simulations are used to evaluate the performance of maximum likelihood estimators. Also, an application is performed comparing the detection results of the proposed model with Gaussian-, Gamma-, and Weibull-based regression models in SAR images.

Keywords: Detection, Rayleigh distribution, regression model, reparameterized Rayleigh distribution, SAR images.

2.1 INTRODUCTION

The classical linear regression model is commonly employed to estimate an unknown and deterministic parameter vector $\boldsymbol{\beta}$ in the linear equation $\mathbf{y} = \mathbf{H}\boldsymbol{\beta} + \mathbf{e}$. The quantity $\mathbf{y}$ is defined as the observed output signal, $\mathbf{H}$ is a linear transformation, and $\mathbf{e}$ is a Gaussian noise vector [71]. However, in situations where the observed output signal is asymmetric, continuous, and nonnegative, as in Rayleigh distributed signals, inference methods based on the Gaussian assumption can lead to misleading results. Indeed, the Rayleigh distribution is widely used in signal and image processing, as in [17, 24–28].

One important application for the Rayleigh distribution is in the context of synthetic aperture radar (SAR) image modeling, where this distribution can be employed for characterizing amplitude values of single-look image pixels [16–18]. A common problem in SAR image processing is the identification and classification of distinct targets or land uses in images [3, 15]. Usually, these problems are treated assuming constant parameters—here referred to as 'homogeneity'— of the regions. However, the use of regression models adopting suitable

distributions without assuming homogeneity in the images can generate accurate results for the above SAR-related challenges, as presented by [72].

In this chapter, our goal is two-fold. First, we propose a regression model for non-Gaussian situations, where the observed output signal is asymmetric and measured continuously on the real positives values. For the proposed model, we introduce parameter estimation, large data record results, and goodness-of-fit measures. Second, we introduce a change detector for the amplitude values of non-Gaussian SAR images. Detection problems are commonly treated assuming Gaussian distribution to the signals. However, SAR images are usually non-Gaussian, prompting the use of the Rayleigh distribution to yield more accurate results for detection problems. Thus, the present chapter introduce a detector based on the asymptotic properties of the proposed Rayleigh regression model parameter estimators.

The chapter is organized as follows. In Section 2.2, we introduce the proposed model and present the score vector, and the goodness-of-fit measures. Section 2.3 shows the Fisher information matrix and the proposed detector. Section 2.4 presents Monte Carlo simulations and an application for SAR images. Finally, the conclusion of this work can be found in Section 2.5.

2.2 PROPOSED RAYLEIGH REGRESSION MODEL

Let Y be a random variable with Rayleigh distribution. Its probability density function (pdf) is given by [4], [18]:

$$p_Y(y;\sigma) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2}{2\sigma^2}\right), \quad y > 0,$$

where $\sigma > 0$ is the parameter. The mean and the variance of Y are given by

$$\mathrm{E}(Y) = \sigma\sqrt{\frac{\pi}{2}},$$
$$\mathrm{Var}(Y) = \sigma^2\left(\frac{4-\pi}{2}\right).$$

Although the Rayleigh density is commonly indexed by the parameter σ, regression models usually characterize the mean of the response signal [73], which has a more direct interpretation than σ. Thus, we consider a reparametrization of the Rayleigh distribution in terms of the mean of the response signal and its regression structure.

2.2.1 Reparametrization of the Rayleigh Distribution

Considering the parameterization $\mu = \sigma\sqrt{\frac{\pi}{2}}$, we have the following pdf of the mean-based Rayleigh distribution:

$$f_Y(y;\mu) = \frac{\pi y}{2\mu^2}\exp\left(-\frac{\pi y^2}{4\mu^2}\right), \quad y > 0, \tag{2.1}$$

where $\mu > 0$ is the mean parameter. The cumulative distribution function is given by

$$F_Y(y;\mu) = 1 - \exp\left(-\frac{\pi y^2}{4\mu^2}\right).$$

The quantile function, useful for generating pseudo-random occurrences in inversion method, is given by

$$Q_Y(u;\mu) = 2\mu\sqrt{\frac{-\log(1-u)}{\pi}}.$$

The mean and variance of Y are given, respectively, by

$$\begin{aligned} \mathrm{E}(Y) &= \mu, \\ \mathrm{Var}(Y) &= \mu^2\left(\frac{4}{\pi} - 1\right). \end{aligned}$$

2.2.2 Regression Model

Let $Y[1], Y[2], \ldots, Y[N]$ be independent random variables, where each $Y[n]$ assumes values $y[n]$ and follows the Rayleigh density in (2.1) with mean $\mu[n]$, $n = 1, 2, \ldots, N$. The proposed Rayleigh regression model is obtained by considering a linear predictor $\eta[n]$ for the mean of $Y[n]$ furnished by

$$\eta[n] = g(\mu[n]) = \sum_{i=1}^{r} \beta_i x_i[n], \quad n = 1, 2, \ldots, N, \tag{2.2}$$

where $r < N$ is the number of covariates considered in the model, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_r)^\top$ is a vector of unknown linear parameters, $\mathbf{x}[n] = (x_1[n], x_2[n], \ldots, x_r[n])^\top$ is a vector of deterministic independent variables, and $g(\cdot)$ is a strictly monotonic and twice differentiable link function where $g : \mathbb{R}^+ \to \mathbb{R}$. If an intercept is considered, then $x_1[n] = 1$. The link function $g(\cdot)$ relates the linear predictors $\eta[n]$ to the expected value $\mu[n]$ of data $y[n]$. When $\mu[n] > 0$, a common choice of link function is the log link $\log(\mu[n]) = \eta[n]$ with its inverse $\mu[n] = \exp(\eta[n])$, since $\exp(\eta[n]) > 0$. For instance, to fit the Rayleigh regression model in a SAR image, the pixel values should be vectorized based on a window of the interest pixels.

The proposed model is similar to the generalized linear models (GLM) [73], except for the fact that the Rayleigh density cannot be written in the canonical form of the exponential family of distributions. A regression model considering the Rayleigh distribution is also presented in [74]. However, the proposed model is based on the standard Rayleigh distribution parametrization. In addition, in this chapter, the maximum likelihood (ML) method [75] based on the reparametrized Rayleigh distribution is considered to obtain the regression parameters estimates, as presented in the next section.

2.2.3 Likelihood Inference

Parameter estimation of the Rayleigh regression model can be performed by the maximum likelihood method [75]. The ML estimates are given by

$$\widehat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} \ell(\boldsymbol{\beta}),$$

where $\ell(\boldsymbol{\beta})$ is the log-likelihood function of the parameters for the observed signal, defined as

$$\ell(\boldsymbol{\beta}) = \sum_{n=1}^{N} \ell[n](\mu[n]).$$

The quantity $\ell[n](\mu[n])$ is the logarithm of $f_Y(y[n];\mu[n])$ given by

$$\ell[n](\mu[n]) = \log[f_Y(y[n];\mu[n])] = \log\left(\frac{\pi}{2}\right) + \log(y[n]) - \log(\mu[n]^2) - \frac{\pi y[n]^2}{4\mu[n]^2},$$

where $\mu[n] = g^{-1}\left(\sum_{i=1}^{r} x_i[n]\beta_i\right)$.

The score vector, obtained by differentiating the log-likelihood function with respect to each unknown parameters β_i, is given by

$$U(\boldsymbol{\beta}) = \left(\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_1}, \frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_2}, \ldots, \frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_r}\right)^{\top}.$$

Then, invoking the chain rule, we have

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_i} = \sum_{n=1}^{N} \frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \frac{\partial \eta[n]}{\partial \beta_i},$$

where

$$\begin{aligned} \frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} &= \frac{\pi y[n]^2}{2\mu[n]^3} - \frac{2}{\mu[n]}, \\ \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} &= \frac{1}{g'(\mu[n])}, \\ \frac{\partial \eta[n]}{\partial \beta_i} &= x_i[n], \end{aligned} \tag{2.3}$$

and $g'(\cdot)$ is the first derivative of the link function $g(\cdot)$. In particular, for the log link function, $g(\mu[n]) = \log(\mu[n])$, we have $\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} = \mu[n]$.

In matrix form, the score vector can be written as

$$U(\boldsymbol{\beta}) = \mathbf{X}^\top \cdot \mathbf{T} \cdot \mathbf{v},$$

where $\mathbf{X}$ is an $N \times r$ matrix whose nth row is $\mathbf{x}[n]^\top$, $\mathbf{T} = \operatorname{diag}\left\{\frac{1}{g'(\mu[1])}, \frac{1}{g'(\mu[2])}, \ldots, \frac{1}{g'(\mu[N])}\right\}$ and $\mathbf{v} = \left(\frac{\pi y[1]^2}{2\mu[1]^3} - \frac{2}{\mu[1]}, \frac{\pi y[2]^2}{2\mu[2]^3} - \frac{2}{\mu[2]}, \ldots, \frac{\pi y[N]^2}{2\mu[N]^3} - \frac{2}{\mu[N]}\right)^\top$.

The maximum likelihood estimators (MLEs) for the Rayleigh regression parameters are obtained by solving the following nonlinear system:

$$U(\boldsymbol{\beta})|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = \mathbf{0}, \tag{2.4}$$

where $\mathbf{0}$ is the r-dimensional vector of zeros.

Solving (2.4) requires the use of nonlinear optimization algorithms. We adopted the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [76] for the numerical computations since the BFGS method requires just the first derivatives [76] and it is generally considered as the best-performing non-linear optimization method [77]. We suggest to use as initial point estimate for $\boldsymbol{\beta}$ the ordinary least squares estimate of $\boldsymbol{\beta}$, obtained from a linear regression of the transformed responses $g(y[1]), g(y[2]), \ldots, g([N])$ on $\mathbf{X}$.

Based on the MLE of $\boldsymbol{\beta}$, it is possible to obtain a MLE for μ, considering the invariance principle of the MLE [75], as $\widehat{\mu} = g^{-1}(\mathbf{X}\widehat{\boldsymbol{\beta}})$.

2.2.4 Goodness-of-fit Measures

In this section, diagnostic measures, such as the residual and the coefficient of determination, are presented to evaluate the correct adjustment of the proposed model. We considered the quantile residual as $r[n] = \Phi^{-1}\left(F(y[n]; \hat{\mu}[n])\right)$, where $\Phi^{-1}(\cdot)$ denotes the standard normal quantile function. The quantile residuals not only can detect poor fitting in regression models, but its distribution is also approximately standard normal [78].

The generalized coefficient of determination [79], which is a global measure of the goodness-of-fit, is given by

$$R^2 = 1 - \exp\left(-\frac{2}{N}\left[\ell(\widehat{\boldsymbol{\beta}}) - \ell(\mathbf{0})\right]\right),$$

where $\ell(\mathbf{0})$ is the maximized log-likelihood of the null model (without regressors) and $\ell(\widehat{\boldsymbol{\beta}})$ is the maximized log-likelihood of the fitted model. Note that $0 \leq R^2 \leq 1$ and it measure the

proportion of the variability of the observed output signal that can be explained by the fitted model [79]. Higher values of R^2 indicate better model fits.

2.3 DETECTION THEORY

It is possible to interpret a SAR image as a set of regions composed of possibly different types of probability laws [15]. The problem of correctly distinguishing between different regions in one image has been studied considering different statistical approaches. One approach to achieve this goal is the use of the hypothesis test, which allows for the computation of differences in the mean of the amplitude between two separate regions in a given image [3, 15]. In SAR image processing, this technique can also be considered for identification of land cover type, land cover change detection or classification, as shown in [80, 81].

2.3.1 Large Data Record Results

Under some mild regularity conditions [82], the MLEs are consistent and asymptotically ($N \to \infty$) normally distributed. Thus, for large data record,

$$\widehat{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}_r\left(\boldsymbol{\beta}, \mathbf{I}^{-1}(\boldsymbol{\beta})\right), \tag{2.5}$$

where $\overset{a}{\sim}$ denotes approximately distributed, $\mathbf{I}(\boldsymbol{\beta})$ is the Fisher information matrix and $\mathcal{N}_r$ is the r-dimensional Gaussian distribution with mean $\boldsymbol{\beta}$ and covariance matrix $\mathbf{I}^{-1}(\beta)$. Their asymptotic distribution can be used to construct confidence intervals [75] and hypothesis tests [75].

To obtain the Fisher information matrix we need to calculate the expectation of the negative value of the second-order partial derivatives of the log-likelihood function [75]. By applying the chain rule, the second-order derivatives of the $\ell(\boldsymbol{\beta})$ with respect to the β_i, $i = 1, 2, \ldots, r$, are given by

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_i \partial \beta_p} &= \sum_{n=1}^{N} \frac{d}{d\mu[n]} \left(\frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \right) \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \frac{\partial \eta[n]}{\partial \beta_p} \frac{\partial \eta[n]}{\partial \beta_i} \\ &= \sum_{n=1}^{N} \left(\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2} \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} + \frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} \frac{\partial}{\partial \mu[n]} \right. \\ &\quad \left. \times \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \right) \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \frac{\partial \eta[n]}{\partial \beta_p} \frac{\partial \eta[n]}{\partial \beta_i}, \quad i, p = 1, 2, \ldots, r.\end{aligned}$$

Note that taking expectation of (2.3), we have that

$$\mathrm{E}\left(\frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} \right) = 0.$$

In addition,

$$\frac{\partial \eta[n]}{\partial \beta_p} = x_p[n]$$

and

$$\frac{\partial \eta[n]}{\partial \beta_i} = x_i[n].$$

Thus,

$$\mathrm{E}\left[\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_i \partial \beta_p}\right] = \sum_{n=1}^{N}\left[\mathrm{E}\left(\frac{\mathrm{d}^2 \ell[n](\mu[n])}{\mathrm{d}\mu[n]^2}\right)\left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]}\right)^2 x_p[n]x_i[n]\right].$$

Now, differentiating (2.3), we obtain

$$\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2} = \frac{2}{\mu[n]^2} - \frac{3\pi y[n]^2}{2\mu[n]^4}.$$

Taking the expected value, we have

$$\mathrm{E}\left[\frac{\mathrm{d}^2 \ell[n](\mu[n])}{\mathrm{d}\mu[n]^2}\right] = -\frac{4}{\mu[n]^2}.$$

Finally, we have

$$\mathrm{E}\left[\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_i \partial \beta_p}\right] = \sum_{n=1}^{N}\left[-\frac{4}{\mu[n]^2}\left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]}\right)^2 x_p[n]x_i[n]\right].$$

In matrix form, the Fisher information matrix is given by

$$\mathbf{I}(\boldsymbol{\beta}) = \mathbf{X}^\top \cdot \mathbf{W} \cdot \mathbf{X},$$

where $\mathbf{W} = \mathrm{diag}\left\{\frac{4}{\mu[1]^2}\left(\frac{\mathrm{d}\mu[1]}{\mathrm{d}\eta[1]}\right)^2, \frac{4}{\mu[2]^2}\left(\frac{\mathrm{d}\mu[2]}{\mathrm{d}\eta[2]}\right)^2, \ldots, \frac{4}{\mu[N]^2}\left(\frac{\mathrm{d}\mu[N]}{\mathrm{d}\eta[N]}\right)^2\right\}$. Similar to the GLM models [73], if $\mathbf{X}$ is rank-deficient, $\left(\mathbf{X}^\top \cdot \mathbf{W} \cdot \mathbf{X}\right)^{-1}$ should be replaced by any generalized inverse.

2.3.2 Hypothesis Test

To test hypotheses over the regression parameters, we partition the parameter vector $\boldsymbol{\beta}$ as $(\boldsymbol{\beta}_I^\top, \boldsymbol{\beta}_J^\top)^\top$, where $\boldsymbol{\beta}_I$ is the vector of parameters of interest with dimension ν and $\boldsymbol{\beta}_J$ is the nuisance parameter vector with dimension $r-\nu$. The hypothesis of interest is $\mathscr{H}_0 : \boldsymbol{\beta}_I = \boldsymbol{\beta}_{I0}$ versus $\mathscr{H}_1 : \boldsymbol{\beta}_I \neq \boldsymbol{\beta}_{I0}$. Here, $\boldsymbol{\beta}_{I0}$ is a fixed column vector of dimension ν. The likelihood ratio

test (T_L), the Wald test (T_W), the and Rao test (T_R) statistics can be written, respectively, as [4]

$$
\begin{aligned}
T_L &= 2\left[\ell(\widehat{\boldsymbol{\beta}}) - \ell(\tilde{\boldsymbol{\beta}})\right], \\
T_W &= (\widehat{\boldsymbol{\beta}}_{I1} - \boldsymbol{\beta}_{I0})^\top \left(\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}}_1)\right]_{\beta_I\beta_I}\right)^{-1} (\widehat{\boldsymbol{\beta}}_{I1} - \boldsymbol{\beta}_{I0}), \\
T_R &= \left(\frac{\partial \ell}{\partial \boldsymbol{\beta}_I}\Big|_{\boldsymbol{\beta}=\tilde{\boldsymbol{\beta}}}\right)^\top \left(\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\beta}})\right]_{\beta_I\beta_I}\right)^{-1} \left(\frac{\partial \ell}{\partial \boldsymbol{\beta}_I}\Big|_{\boldsymbol{\beta}=\tilde{\boldsymbol{\beta}}}\right),
\end{aligned}
$$

where $\widehat{\boldsymbol{\beta}}_{I1}$ and $\widehat{\boldsymbol{\beta}}_1 = (\widehat{\boldsymbol{\beta}}_{I1}^\top, \widehat{\boldsymbol{\beta}}_{J1}^\top)^\top$ are the MLEs under $\mathscr{H}_1$ (unrestricted MLEs); $\boldsymbol{\beta}_{J0}$ is a fixed column vector of dimension $r-\nu$; $\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}})\right]_{\beta_I\beta_I}$ is a partition of $\mathbf{I}(\widehat{\boldsymbol{\beta}})$ limited to the estimates of interest; $\tilde{\boldsymbol{\beta}} = (\widehat{\boldsymbol{\beta}}_{I0}^\top, \widehat{\boldsymbol{\beta}}_{J0}^\top)^\top$ is the MLE under $\mathscr{H}_0$; $\ell(\tilde{\boldsymbol{\beta}})$ is the maximized log-likelihood of the model under $\mathscr{H}_0$; and $\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\beta}})\right]_{\beta_I\beta_I}$ is a partition of $\mathbf{I}(\tilde{\boldsymbol{\beta}})$ limited to the estimates of interest. As suggested in [83] and due to the convenience of the Wald test, since it requires only one estimation under alternative hypothesis, we selected the Wald test to perform the hypothesis test on the parameters.

From (2.5) and based on the consistency of the MLE, the T_W statistic has an asymptotically chi-squared distribution with ν degrees of freedom, χ^2_ν. The detection is performed by comparing the computed value of T_W with a threshold value γ obtained from the χ^2_ν distribution and the desired probability of false alarm [4].

We assume that the mean of the Rayleigh distributed signal presents different values depending on the ground type. To illustrate, consider a region of forest in an image. The detection of this type of ground can be obtained by fitting the following Rayleigh regression model

$$
g(\mu[n]) = \beta_1 + \beta_2 x_2[n] + \sum_{i=3}^{r} \beta_i x_i[n],
$$

where (i) β_1 is the intercept; (ii) $x_2[n]$ is a binary covariate equal to one if the region consists of forest and zero otherwise; and (iii) $x_i[n]$, $i = 3, 4, \ldots, r$, are any other covariates that can influence the mean of y. The detection problem is to distinguish between the hypotheses:

$$
\begin{cases}
\mathscr{H}_0 : \mu[n] = g^{-1}(\beta_1 + \sum_{i=3}^{r} \beta_i x_i[n]), \quad (\beta_2 = 0), \\
\mathscr{H}_1 : \mu[n] = g^{-1}(\beta_1 + \beta_2 x_2[n] + \sum_{i=3}^{r} \beta_i x_i[n]).
\end{cases}
\tag{2.6}
$$

To derive the detector, we can use the Wald test described above. We reject $\mathscr{H}_0$ when $T_W > \gamma$ [4]. In this situation, $\beta_2 \neq 0$ and the forest land use is detected. This technique can be considered to detect any type of ground in SAR images.

2.4 NUMERICAL RESULTS

This section presents Monte Carlo simulations and an empirical investigation in ground type detection in SAR images. Monte Carlo simulations were used to evaluate the MLE of the Rayleigh regression parameters. An application with real SAR data was considered to demonstrate the proposed detector.

2.4.1 Analysis with Simulated Data

The numerical results are based on the Rayleigh regression model with the structure of the mean given by (2.2) considering the log link function. The parameters were set as follows: $\beta_1 = 2$, $\beta_2 = -1$, and $\beta_3 = 1$ for Scenario 1, and $\beta_1 = 0.5$ and $\beta_2 = 0.15$ for Scenario 2. The covariates values were generated from the uniform distribution $(0,1)$ and kept constant in all Monte Carlo replications. The choice of parameters aims at capturing asymmetric distributions, i.e., values of skewness higher than zero. In each replication the inversion method was considered to generate $y[n]$ assuming the Rayleigh distribution with mean $\mu[n]$. The number of Monte Carlo replications, R, was set equal to $10,000$ and the signal lengths considered were $N \in \{25; 250; 1,000\}$. The structure of the Monte Carlo simulations is summarized in Algorithm 1.

Algorithm 1: Monte Carlo simulations for evaluation of the Rayleigh regression model parameter estimators.

Input: Vector of parameter $\boldsymbol{\beta}$ and signal length N.
Output: Results of the desirable figures of merit.

1: Suppose that the observed output signal $y[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\beta}$;
2: Generate $y[n]$ from $f_Y(\boldsymbol{\beta})$;
3: For each Monte Carlo replication, compute $\widehat{\boldsymbol{\beta}}$;
4: Repeat steps 2 and 3 a very large number R of times, obtaining: $\widehat{\boldsymbol{\beta}}[1], \widehat{\boldsymbol{\beta}}[2], \ldots, \widehat{\boldsymbol{\beta}}[R]$;
5: Use the estimates $\widehat{\boldsymbol{\beta}}[1], \widehat{\boldsymbol{\beta}}[2], \ldots, \widehat{\boldsymbol{\beta}}[R]$, to calculate the desired figures of metric (mean, relative bias, and mean square error).

We adopted the percentage relative bias (RB%) and the means square error (MSE) as figures of merit to numerically evaluate the proposed point estimators. Table 1 presents the simulation results for Scenarios 1 and 2. In general, we notice that the MLE of the Rayleigh regression model presented small values of percentage relative bias and mean square error. As expected, increasing N, the percentage relative bias and mean square error present lower values, which matches the consistence of the MLE.

Table 1 – Results of the Monte Carlo simulation of the point estimation for Scenarios 1 and 2. In particular, we considered the percentage relative bias (RB%) and the means square error (MSE) as figures of merit to numerically evaluate the proposed point estimators

	Scenario 1			Scenario 2	
Measures	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_1$	$\widehat{\beta}_2$
			$N = 25$		
Mean	1.9681	−1.0030	1.0040	0.4810	0.1472
RB(%)	1.5972	−0.3004	−0.4045	3.7913	1.8467
MSE	0.0909	0.1564	0.1533	0.0470	0.1421
			$N = 250$		
Mean	1.9971	−1.0016	1.0009	0.4984	0.1489
RB(%)	0.1450	−0.1600	−0.0900	0.3200	0.7333
MSE	0.0073	0.0121	0.0126	0.0041	0.0125
			$N = 1,000$		
Mean	1.9993	−1.0001	1.0002	0.4995	0.1502
RB(%)	0.0350	−0.0100	−0.0200	0.1000	−0.1333
MSE	0.0017	0.0030	0.0029	0.0010	0.0030

Source: Author (2020)

In Appendix A, we present simulation results for point parameter estimation of the Rayleigh regression model, in which the parameters were set as: (i) $\beta_1 = 0.5$, $\beta_2 = -0.5$, and $\beta_3 = 0.3$ for Scenario 3; and (ii) $\beta_1 = 0.5$, $\beta_2 = 0.15$, and $\beta_3 = 1$, for Scenarios 4 and 5. Additionally, we also present the simulation numerical results of Scenarios 1 and 2 for $N \in \{16; 25; 49; 250; 500; 1,000\}$.

2.4.2 Analysis with Real Data

The SAR image considered in this application was produced by CARABAS II [84], a Swedish ultrawideband (UWB) very-high frequency (VHF) SAR system. The system uses horizontal (HH) polarization. All information related to the data can be found in [2, 84] and the images are available in [85]. The ground scene of the selected image is dominated by pine forest, fences, power lines, military vehicles, and roads; a lake is also present [84].

Figure 1 shows the three different regions representing forest, lake, and military vehicle imagery; referred to as Regions A1, A2, and A3, respectively. These regions were submitted to the proposed modeling and detector. The model is specified for the mean of the response signal using an intercept ($x_1[n] = 1$) and two dummy variables ($x_2[n]$ and $x_3[n]$) representing

each tested region, as

$$g(\mu[n]) = \beta_1 + \beta_2 x_2[n] + \beta_3 x_3[n].$$

The response signal is composed of the amplitude values of the pixels of the Regions A1, A2, and A3. The variable $x_2[n]$ is defined as one for Region A2 and zero for the rest. The variable $x_3[n]$ is defined as one for Region A3 and zero for the others. Region A1 is represented when $x_2[n] = 0$ and $x_3[n] = 0$. In summary, the Rayleigh regression model is fitted in a SAR image following Algorithm 2.

Algorithm 2: Fitting a Rayleigh regression model in a SAR image.

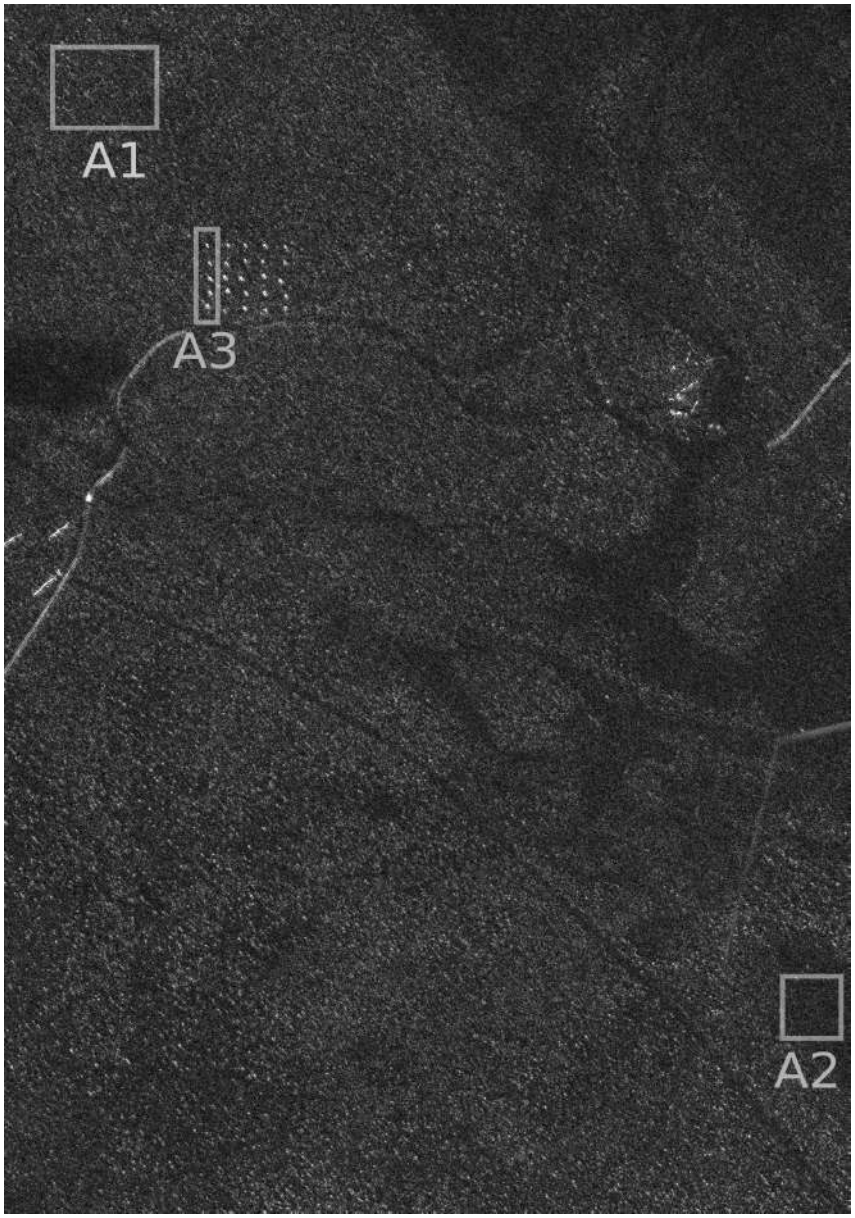
Input: Amplitude pixel window of a SAR image.
Output: Estimated parameters, fitted values, and residual of the Rayleigh regression model.
1: Select a window of pixels of each tested region;
2: Create the response signal with the pixels amplitude values of each tested region;
3: Fit the Rayleigh regression model.

For comparison purposes, we also fitted the standard Gaussian regression model, the GLM with Gamma distribution, and the Weibull regression model [86] to the Regions A1, A2, and A3. Detection with Gaussian distribution is widely discussed in literature and the Gamma and Weibull distribution are also used in SAR images, as in [62,87]. The estimated parameters for the considered models are given in Table 2. In the Rayleigh regression model, the mean response presents a negative relationship with $x_2[n]$ and positive relationship with $x_3[n]$. Additionally, we notice that the lake and the target regions led to mean responses which are 12.05% lower and 194.00% higher than the mean response from the forest region, respectively.

The R^2 values of the fitted models show that the Rayleigh regression model can explain 70.96% of the variation in $y[n]$, while the Gamma GLM, Gaussian, and Weibull regression models can explain just 30.09%, 15.28%, and 32.51%, respectively. The Rayleigh regression model can explain better the variations in $y[n]$ in comparative to Gamma GLM, Weibull, and Gaussian regression model. Figure 2 presents the residuals of the Rayleigh regression model. As expected, the residuals present values close to zero for 98.81% of the observations and approximately standard normal distribution (p-value = 0.113 in the Kolmogorov-Smirnov test [88]). The blue circles in Figure 2a are related to the pixels where the targets were deployed.

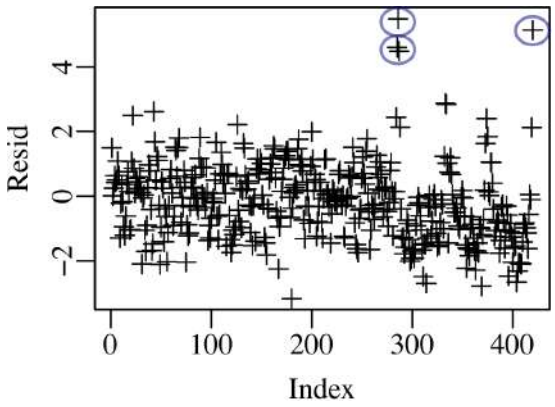
It is possible to define a detector for this specific regression model, based on (2.6). The detection problem in this image is based on computing the difference in the behavior among the

Figure 1 – CARABAS II single-look image used in the regression models showing the tested regions. Regions A1, A2, and A3 represent a forest, a lake, and an area containing military vehicles, respectively.

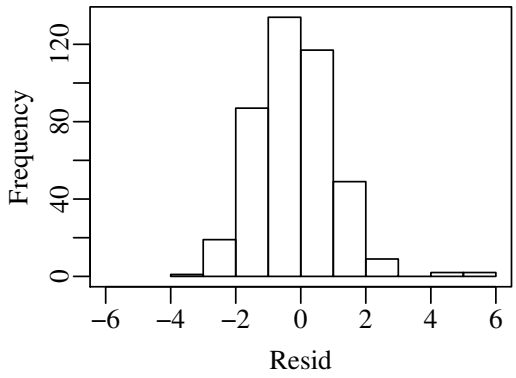


Source: Author (2020)

Figure 2 – Residual plots for the Rayleigh regression model. Blue circles are related to the pixels where targets were deployed.



(a) Residuals vs. index



(b) Histogram

Source: Author (2020)

tested regions. With the p-values of the Wald test presented in Table 2, we can verify that all variables in the Rayleigh regression model are significant for a probability of false alarm equal to 0.05. Hence, the null hypothesis in (2.6) can be rejected, indicating a correct detection of the land type. In contrast, the variable $x_2[n]$ is not significant for the Gamma GLM, Gaussian, and Weibull regression models, i.e., the Gaussian-, Gamma-, and Weibull-based detections can not distinguish the lake region from the other regions.

The Rayleigh distribution is a particular case of the Weibull distribution. However, the definitions of the Rayleigh regression model and the Weibull regression model considered in this study are different. Our model estimates the observed signal mean, while the employed Weibull regression model estimates the time to an event of interest, considering survival data. This fact may explain the results superiority of the Rayleigh regression model in terms of detection. The sample means of the lake and forest regions are more similar to each other ($\overline{X}_{\text{forest}} = 0.1267$ and $\overline{X}_{\text{lake}} = 0.1148$) when compared to the region with the vehicles ($\overline{X}_{\text{vehicles}} = 0.2863$). Thus, distinguishing the mean of the vehicle region is easier, even if the model is not the most suitable

Table 2 – Fitted regression models for Regions A1, A2, and A3. All variables in the Rayleigh regression model are significant for a probability of false alarm equal to 0.05, indicating a correct detection of the land type.

	Estimate	Standard Error	Detection (p-value)
Rayleigh regression model			
$\widehat{\beta}_1$	-2.0623	0.0445	< 0.001
$\widehat{\beta}_2$	-0.1280	0.0599	$\mathbf{0.0325}$
$\widehat{\beta}_3$	1.0784	0.0616	< 0.001
$R^2 = 0.7096$			
Gaussian regression model			
$\widehat{\beta}_1$	0.12683	0.01646	< 0.001
$\widehat{\beta}_2$	-0.01201	0.02213	0.588
$\widehat{\beta}_3$	0.15948	0.02277	< 0.001
$R^2 = 0.1528$			
Gamma GLM			
$\widehat{\beta}_1$	7.8844	0.5209	< 0.001
$\widehat{\beta}_2$	0.8248	0.7341	0.262
$\widehat{\beta}_3$	-4.3917	0.5657	< 0.001
$R^2 = 0.3009$			
Weibull regression model			
$\widehat{\beta}_1$	-1.9939	0.0583	< 0.001
$\widehat{\beta}_2$	-0.1157	0.0778	0.1373
$\widehat{\beta}_3$	0.9583	0.0815	< 0.001
$R^2 = 0.3251$			

Source: Author (2020)

to the data. Thus, the proposed Rayleigh regression model can be used for detecting differences in SAR image regions yielding more accurate results when compared to the competing regression models.

2.5 CONCLUSION

This chapter introduced a new regression model for nonnegative signals. The proposed Rayleigh regression model assumes that the mean of the Rayleigh distributed signal follows a regression structure involving covariates, unknown parameters, and a link function. An inference approach for the model parameters is introduced and diagnostic tools are discussed. We also presented Fisher information matrix, asymptotic properties of the MLE, and a detector useful

to detect differences in SAR image regions. In the Monte Carlo simulations, the MLE of the Rayleigh regression model showed small values of percentage relative bias and mean square error. An application of the Rayleigh regression model to distinguish between different regions in a SAR image was presented and discussed, showing more accurate detection results when compared with the measurements from Gaussian-, Gamma-, and Weibull-based regression models.

3 IMPROVED POINT ESTIMATION FOR THE RAYLEIGH REGRESSION MODEL

Abstract

The Rayleigh regression model was recently proposed for modeling amplitude values of synthetic aperture radar (SAR) image pixels. However, inferences from such model are based on the maximum likelihood estimators, which can be biased for small signal lengths. The Rayleigh regression model for SAR images often take into account small pixel windows, which may lead to inaccurate results. In this letter, we introduce bias-adjusted estimators tailored for the Rayleigh regression model and based on: (i) the Cox and Snell's method; (ii) the Firth's scheme; and (iii) the parametric bootstrap method. We present Monte Carlo simulations to evaluate the performance of the proposed estimators. Numerical results show that the bias-adjusted estimators yield nearly unbiased estimates.

Keywords: Bias correction, Rayleigh regression model, SAR images, small signal lengths inferences

3.1 INTRODUCTION

The classical linear regression model is widely employed to estimate an unknown and deterministic parameter vector assuming the Gaussian distribution [71]. However, practical contexts often exhibit non-Gaussian behavior [7]. An alternative to the Gaussian model is provided by the Rayleigh distribution [89] which is capable of characterizing asymmetric, continuous, and nonnegative signals, such as the amplitude values of synthetic aperture radar (SAR) image pixels [14, 16–18, 27, 90–93]. A usual problem in SAR image processing is the classification and identification of distinct land uses or target regions [3, 15]. Generally, these problems are treated assuming constant parameters, i.e., assuming homogeneity of the regions. However, when such assumption is not suitable [72], regression modeling can be considered where each observation has one specific estimated mean [73]. The Rayleigh regression model was proposed and discussed in [94], where a methodology for point estimation, large data record results, and goodness-of-fit measures were presented and discussed in the context of SAR image detection.

Parameter inference based on the Rayleigh regression model can be achieved by means of the maximum likelihood estimation (MLE), inheriting its good asymptotic properties for large

signal lengths. However, if the signal length N is small, then the maximum likelihood estimators present a bias in the order of N^{-1}, which can be regarded as problematic [31, 32]. For instance, in [94], the detection of different land cover types in SAR images was performed based on pixel windows of more than 126 pixels. However, if smaller windows, such as 3×3, are selected for Rayleigh regression model parameter estimation, then the obtained results can be severely biased [32].

An approach to address this issue is by means of inferential correction [32]. Three widely bias-adjusted methods are the Cox and Snell's [66], Firth's [67], and parametric bootstrap [68] schemes. The Cox and Snell's method is an analytical approach used to obtain second order corrected estimators for a variety of models, such as the generalized linear models [95], the beta regression model [96, 97], the overdispersed generalized linear models [98], and the extreme-value regression model [99]. The Firth's method is a preventive method [67] of bias reduction which modify the score function before obtaining parameter estimates based on the analytical second order biases of the maximum likelihood estimators [67]. The Firth's method was considered to obtain bias-adjusted estimators for several models, such as (i) the binomial regression model [100]; (ii) the multinomial logistic regression model [101]; and (iii) the Kumaraswamy distribution [102]. Finally, the bootstrap method is a computationally intensive method based on resampling, being suitable for inferential corrections when N is small [68]. The bootstrap-based bias-adjusted estimators are considered in the beta regression models [96], in the Birnbaum-Saunders distribution [103], and in the autoregressive time series [104].

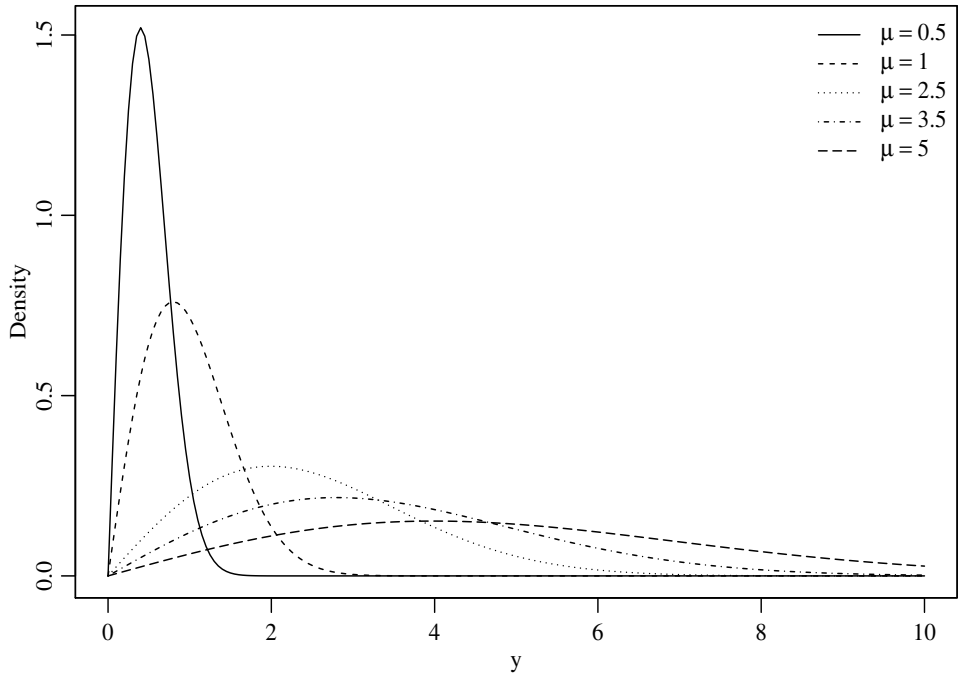
To the best of our knowledge, the literature lacks bias-adjusted estimators for the parameters of the Rayleigh regression model. In this chapter, our chief goal is to obtain accurate point estimation approaches to address this literature gap.

This chapter is organized as follows. In Section 3.2, we describe the Rayleigh regression model. Section 3.3 shows the bias correction estimators using both numerical and analytical correction schemes. Section 3.4 presents Monte Carlo simulations results. Finally, we conclude this work in Section 3.5.

3.2 THE RAYLEIGH REGRESSION MODEL

The Rayleigh regression model was proposed in [94] and can be defined as follows. Let Y be a Rayleigh distributed random variable with mean parameter μ. The probability density

Figure 3 – Rayleigh probability densities functions for $\mu \in \{0.5, 1, 2.5, 3.5, 5\}$.



Source: Author (2020)

function of the mean-based Rayleigh distribution is given by [94]

$$f_Y(y;\mu) = \frac{\pi y}{2\mu^2} \exp\left(-\frac{\pi y^2}{4\mu^2}\right), \tag{3.1}$$

where $\mu > 0$ and $y > 0$ is the observed signal value. The mean and variance of Y are given, respectively, by $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \mu^2\left(\frac{4}{\pi} - 1\right)$.

The Rayleigh probability density function is very flexible for nonnegative signals, as can be verified in Figure 3. Additionally, we note that higher mean values imply in lower asymmetry and kurtosis values.

Let $Y[1], Y[2], \ldots, Y[N]$ be independent random variables, where each $Y[n]$ assumes values $y[n]$ and follows the Rayleigh density in (3.1) with mean $\mu[n]$, $n = 1, 2, \ldots, N$. The Rayleigh regression model is defined assuming that the mean of the observed output signal $Y[n]$ can be written as [94]

$$\eta[n] = g(\mu[n]) = \sum_{i=1}^{k} \beta_i x_i[n], \quad n = 1, 2, \ldots, N,$$

where $k < N$ is the number of covariates considered in the model, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)^\top$ is a vector of unknown linear parameters, $\mathbf{x}[n] = (x_1[n], x_2[n], \ldots, x_k[n])^\top$ is a vector of independent input variables, $\eta[n]$ is the linear predictor, and $g : \mathbb{R}^+ \to \mathbb{R}$ is a strictly monotonic and twice differentiable link function.

Parameter estimation can be performed using the maximum likelihood method, as discussed in [94]. The estimated vector $\widehat{\boldsymbol{\beta}}$ is obtained by maximizing the logarithm of the likelihood function. The log-likelihood function of the parameter vector $\boldsymbol{\beta}$ for the observed signal is $\ell(\boldsymbol{\beta}) = \sum_{n=1}^{N} \ell[n](\mu[n])$, where

$$\ell[n](\mu[n]) = \log\left(\frac{\pi}{2}\right) + \log(y[n]) - \log(\mu[n]^2) - \frac{\pi y[n]^2}{4\mu[n]^2}.$$

The score vector can be written as

$$U(\boldsymbol{\beta}) = \mathbf{X}^\top \cdot \mathbf{T} \cdot \mathbf{v},$$

where $\mathbf{X}$ is an $N \times r$ matrix whose nth row is $\mathbf{x}[n]^\top$, $\mathbf{T} = \operatorname{diag}\left\{\frac{1}{g'(\mu[1])}, \frac{1}{g'(\mu[2])}, \ldots, \frac{1}{g'(\mu[N])}\right\}$ and $\mathbf{v} = \left(\frac{\pi y[1]^2}{2\mu[1]^3} - \frac{2}{\mu[1]}, \frac{\pi y[2]^2}{2\mu[2]^3} - \frac{2}{\mu[2]}, \ldots, \frac{\pi y[N]^2}{2\mu[N]^3} - \frac{2}{\mu[N]}\right)^\top$. Finally, the Fisher information matrix is given by

$$\mathbf{I}(\boldsymbol{\beta}) = \mathbf{X}^\top \cdot \mathbf{W} \cdot \mathbf{X},$$

where $\mathbf{W} = \operatorname{diag}\left\{\frac{4}{\mu[1]^2}\left(\frac{d\mu[1]}{d\eta[1]}\right)^2, \ldots, \frac{4}{\mu[N]^2}\left(\frac{d\mu[N]}{d\eta[N]}\right)^2\right\}$. Further mathematical properties, including large data record results, are detailed in [94].

3.3 BIAS CORRECTION OF MAXIMUM LIKELIHOOD ESTIMATORS

Generally, for small N, maximum likelihood estimators may be biased to their true parametric values [32]. The bias of the estimator $\widehat{\boldsymbol{\beta}}$ can be expressed as [75]

$$B(\widehat{\boldsymbol{\beta}}) = \mathrm{E}[\widehat{\boldsymbol{\beta}}] - \boldsymbol{\beta}.$$

Cox and Snell's bias correction formula [66] can be used to obtain the second order bias of the maximum likelihood estimators, considering the inverse of the Fisher information matrix and cumulants of log-likelihood derivatives up to third order with respect to the unknown parameters [31]. The Cox and Snell's bias correction formula for the ath component of $\widehat{\boldsymbol{\beta}}$ is given by [66]

$$B(\widehat{\beta}_a) = \sum_{r,s,u} \kappa^{ar}\kappa^{su}\left\{\kappa_{rs}^{(u)} - \frac{1}{2}\kappa_{rsu}\right\},$$

where $\kappa_{rs} = \mathrm{E}\left(\frac{\partial^2 \ell}{\partial\beta_r \partial\beta_s}\right)$, $\kappa_{rs}^{(u)} = \frac{\partial \kappa_{rs}}{\partial \beta_u}$, $\kappa_{rsu} = \mathrm{E}\left(\frac{\partial^3 \ell}{\partial\beta_r \partial\beta_s \partial\beta_u}\right)$, $-\kappa^{ar}$ and $-\kappa^{su}$ are the (a,r) and (s,u) elements of the inverse of the Fisher information matrix, respectively. The cumulants obtained

for the Rayleigh regression model can be found in the Appendix B. The second order bias of $\widehat{\boldsymbol{\beta}}$ is given by

$$B(\widehat{\boldsymbol{\beta}}) = \mathbf{I}^{-1}(\widehat{\boldsymbol{\beta}}) \cdot \mathbf{X}^{\top} \cdot \mathbf{W} \cdot \delta,$$

where δ is the main diagonal of $\mathbf{X} \cdot \mathbf{I}^{-1}(\boldsymbol{\beta}) \cdot \mathbf{X}^{\top}$ and

$$\mathbf{W} = \operatorname{diag}\left[-\frac{2}{\mu[n]^3} \left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \right)^3 - \frac{2}{\mu[n]^2} \left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \right)^2 \frac{\partial}{\partial \mu[n]} \left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \right) \right].$$

Replacing the unknown parameters by their maximum likelihood estimators, we have the MLE $\widehat{B}(\widehat{\boldsymbol{\beta}})$ of $B(\widehat{\boldsymbol{\beta}})$. Thus, and bias-adjusted estimators can be derived removing $\widehat{B}(\widehat{\boldsymbol{\beta}})$ of $\widehat{\boldsymbol{\beta}}$ [32]. Hence, corrected estimators based on the Cox and Snell's method, $\widetilde{\boldsymbol{\beta}}$, are obtained as follows [66]

$$\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - \widehat{B}(\widehat{\boldsymbol{\beta}}). \tag{3.2}$$

The Firth's method removes the second-order bias by modifying the original score function $U(\boldsymbol{\beta})$ according to [67]

$$U^*(\boldsymbol{\beta}) = U(\boldsymbol{\beta}) - \mathbf{I}(\boldsymbol{\beta}) \cdot B(\widehat{\boldsymbol{\beta}}). \tag{3.3}$$

The roots of the modified score function $U^*(\boldsymbol{\beta})$ constitute the corrected estimator $\widehat{\boldsymbol{\beta}}^{\star}$ according to the Firth's method.

In the bootstrap bias correction method, the bias estimation $\widehat{B}(\widehat{\boldsymbol{\beta}})$ is numerically obtained through Monte Carlo simulations. A bootstrap estimate of the bias can be obtained by

$$\widehat{B}_{\text{boot}}(\widehat{\boldsymbol{\beta}}) = \bar{\boldsymbol{\beta}}^{*} - \widehat{\boldsymbol{\beta}}, \tag{3.4}$$

where $\bar{\boldsymbol{\beta}}^{*} = \frac{1}{R} \sum_{b=1}^{R} \widehat{\boldsymbol{\beta}}_b$, R is the number of bootstrap replications, and $\widehat{\boldsymbol{\beta}}_b$ is the estimated values of $\boldsymbol{\beta}$ in each bootstrap replication. Thus, the corrected estimator based on the bootstrap method is given by [68, 105]

$$\widehat{\boldsymbol{\beta}}^{*} = \widehat{\boldsymbol{\beta}} - \widehat{B}_{\text{boot}}(\widehat{\boldsymbol{\beta}}) = \widehat{\boldsymbol{\beta}} - (\bar{\boldsymbol{\beta}}^{*} - \widehat{\boldsymbol{\beta}}) = 2\widehat{\boldsymbol{\beta}} - \bar{\boldsymbol{\beta}}^{*}.$$

The above described bias-adjusted estimators share the same asymptotic properties with the usual maximum likelihood estimators but are less biased for small N [106].

3.4 NUMERICAL RESULTS

Monte Carlo simulations were employed to evaluate the original maximum likelihood estimators performance of the Rayleigh regression model parameters and their bias-adjusted versions. The parameter estimates were obtained by maximizing the log-likelihood function considering the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method [76] with analytic first derivatives.

Simulations were performed under two scenarios. Each scenario aimed at capturing asymmetric distributions. For Scenario 1, we selected the following parameters values: $\beta_1 = 0.5$, $\beta_2 = 0.5$, $\beta_3 = 1$, and covariates generated from the binomial distribution; whereas, for Scenario 2, we adopted $\beta_1 = 2.5$, $\beta_2 = 1.5$, and covariates generated from the Rayleigh distribution. Considering the adopted parameters, we have for Scenarios 1 and 2, skewness about 4 and 3, respectively. The covariates values were kept constant for all Monte Carlo replications and the log link function was employed.

The inversion method was used in each replication to generate $y[n]$ assuming the Rayleigh distribution with mean $\mu[n]$. The number of Monte Carlo and bootstrap replications were set equal to 5,000 and 1,000, respectively, and the signal lengths considered were $N \in \{9;25;49\}$. Such blocklengths are popular choices of window sizes in SAR image processing [93, 107, 108]. The Monte Carlo simulations are summarized in Algorithm 3.

Algorithm 3: Monte Carlo simulations for the bias-adjusted estimators of the Rayleigh regression model parameters.

Input: Vector of parameter $\boldsymbol{\beta}$ and signal length N.
Output: Results of the desirable figures of merit.

1: Suppose that the observed output signal $y[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\beta}$;
2: Generate $y[n]$ from $f_Y(\boldsymbol{\beta})$;
3: For each Monte Carlo replication, compute $\widehat{\boldsymbol{\beta}}$ and $\widehat{B}(\widehat{\boldsymbol{\beta}})$;
4: Based on $\widehat{B}(\widehat{\boldsymbol{\beta}})$, obtain the Cox and Snell's and Firth's bias-adjusted estimates;
5: Generate bootstrap samples from $f_Y(\widehat{\boldsymbol{\beta}})$ and compute $\widehat{\boldsymbol{\beta}}_b$;
6: For each bootstrap replication, compute $\widehat{\boldsymbol{\beta}}_b$;
7: Repeat steps 5 and 6 a very large number R of times, obtaining: $(\widehat{\boldsymbol{\beta}}_b[1], \widehat{\boldsymbol{\beta}}_b[2], \ldots, \widehat{\boldsymbol{\beta}}_b[R])$;
8: Obtain the bias-adjusted estimates based on $\bar{\boldsymbol{\beta}}^*$;
9: Repeat steps 2–8 a very large number B of times, obtaining: $(\widehat{\boldsymbol{\beta}}[1], \widehat{\boldsymbol{\beta}}[2], \ldots, \widehat{\boldsymbol{\beta}}[B])$, $(\widetilde{\boldsymbol{\beta}}[1], \widetilde{\boldsymbol{\beta}}[2], \ldots, \widetilde{\boldsymbol{\beta}}[B])$, $(\widehat{\boldsymbol{\beta}}^{\star}[1], \widehat{\boldsymbol{\beta}}^{\star}[2], \ldots, \widehat{\boldsymbol{\beta}}^{\star}[B])$, and $(\widehat{\boldsymbol{\beta}}^{*}[1], \widehat{\boldsymbol{\beta}}^{*}[2], \ldots, \widehat{\boldsymbol{\beta}}^{*}[B])$;
10: Use the computed estimates to calculate the desired figures of metric (mean, relative bias, and mean square error).

The percentage relative bias (RB%) and the root mean squared error ($\sqrt{\text{MSE}}$) were adopted as figures of merit to numerically evaluate the point estimators. Table 3 presents the simulation results for point estimation of the Rayleigh regression model parameters for Scenarios 1 and 2. In general, we notice that the maximum likelihood estimators can be strongly biased for small N. For instance, for Scenario 1 and $N = 9$, the relative biases of the maximum likelihood estimators are approximately (in absolute values) 10%, 24%, and 1% for $\widehat{\beta}_1$, $\widehat{\beta}_2$, and $\widehat{\beta}_3$, respectively. For this same signal length, the relative bias of $\widetilde{\beta}_2$, $\widehat{\beta}_2^{\star}$, and $\widehat{\beta}_2^{*}$ are (in absolute values) 9%, 9%, and 7%, respectively.

The bias-corrected estimators present values closer to the true parameters when compared to the maximum likelihood estimates and have similar performance in terms of relative bias and mean square error. Additionally, the root mean square errors of the corrected estimators are usually smaller than the uncorrected estimators and they decrease when N increase, as expected.

To evaluate the overall performances of the four estimators for each value of N, we employed the integrated relative bias squared norm (IRBSN) figure of merit [109], which is defined as

$$\text{IRBSN} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\text{RB}(\widehat{\beta}_i)^2},$$

where $\text{RB}(\widehat{\beta}_i)$, $i = 1,2,\ldots,k$, correspond to the values of RB% of each estimator. The values of IRBSN for Scenarios 1 and 2 are given in Table 4. The corrected estimators excel in terms of IRBSN. Additionally, among the evaluated estimators, the ones obtained by the Firth's method present the smallest values of IRBSN in five of the six evaluated scenarios.

3.5 CONCLUSIONS

This chapter introduced bias-adjusted estimators for the Rayleigh regression model parameters. In particular, we employed the Cox and Snell's, Firth's, and parametric bootstrap methods to obtain the corrected estimators. The numerical evaluation of the maximum likelihood estimators and bias-adjusted estimators was performed considering Monte Carlo simulations. In general, the discussed bias-adjusted estimators outperformed maximum likelihood estimators in terms of relative bias and root mean square error; being the estimators based on the Firth's method the best performing approach. In conclusion, we recommend the use of corrected estimators based on the Firth's method to fit the Rayleigh regression model for small signal lengths.

Table 3 – Results of the Monte Carlo simulation for point estimation of Scenarios 1 and 2. The Cox and Snell ($\widetilde{\boldsymbol{\beta}}$), Firth ($\widehat{\boldsymbol{\beta}}^{\star}$), and bootstrap ($\widehat{\boldsymbol{\beta}}^{*}$) estimators are linked to Equations (3.2), (3.3), and (3.4), respectively. We highlighted the smallest percentage relative bias values for each signal length

		$\widehat{\beta}_1$	$\widetilde{\beta}_1$	$\widehat{\beta}_1^{\star}$	$\widehat{\beta}_1^{*}$	$\widehat{\beta}_2$	$\widetilde{\beta}_2$	$\widehat{\beta}_2^{\star}$	$\widehat{\beta}_2^{*}$	$\widehat{\beta}_3$	$\widetilde{\beta}_3$	$\widehat{\beta}_3^{\star}$	$\widehat{\beta}_3^{*}$
					$N = 9$								
Scenario 1	RB(%)	−10.0546	0.9665	**−0.4822**	1.0152	−24.3255	−8.9817	−8.8812	**−7.1089**	−1.4237	**1.0249**	2.5234	1.2595
	$\sqrt{\text{MSE}}$	0.7083	0.7049	0.7068	0.7051	0.7960	0.7927	0.7924	0.7922	0.8083	0.8055	0.8087	0.8062
Scenario 2	RB(%)	−1.6082	−0.0595	−0.1941	**−0.0252**	−2.4411	−0.5491	**0.2736**	−0.4622	–	–	–	–
	$\sqrt{\text{MSE}}$	0.4625	0.4599	0.4610	0.4600	0.6979	0.6959	0.6980	0.6961	–	–	–	–
					$N = 25$								
Scenario 1	RB(%)	−1.2113	1.1810	**0.6506**	1.4531	−13.9903	−4.1510	**−2.1974**	−4.5960	−1.8112	−0.3642	**−0.0567**	0.4670
	$\sqrt{\text{MSE}}$	0.3319	0.3317	0.3317	0.3316	0.3568	0.3562	0.3563	0.3565	0.3554	0.3548	0.3548	0.3549
Scenario 2	RB(%)	−0.2640	0.1544	**0.1150**	0.1812	−1.3497	−0.3138	**−0.1566**	−0.3695	–	–	–	–
	$\sqrt{\text{MSE}}$	0.2319	0.2318	0.2319	0.2319	0.3439	0.3432	0.3434	0.3432	–	–	–	–
					$N = 49$								
Scenario 1	RB(%)	−0.9191	**−0.0037**	−0.1914	0.0997	−7.0768	−1.3383	**−0.6904**	−1.5395	−0.6448	0.2236	0.3203	**0.1924**
	$\sqrt{\text{MSE}}$	0.2226	0.2225	0.2225	0.2226	0.2381	0.2380	0.2380	0.2381	0.2365	0.2362	0.2362	0.2364
Scenario 2	RB(%)	−0.1297	0.0580	**0.0462**	0.0716	−0.7194	−0.1266	**−0.0821**	−0.1552	–	–	–	–
	$\sqrt{\text{MSE}}$	0.1588	0.1588	0.1588	0.1589	0.2306	0.2304	0.2304	0.2305	–	–	–	–

Source: Author (2020)

Table 4 – Integrated relative bias squared norm results. Best results are highlighted

	MLE	Cox and Snell	Firth	Bootstrap
		Scenario 1		
$N = 9$	15.2190	5.2490	5.3378	**2.8764**
$N = 25$	8.1747	2.5005	**1.3235**	2.7960
$N = 49$	4.1369	0.7834	**0.4531**	1.0368
		Scenario 2		
$N = 9$	2.0670	0.3905	**0.2372**	0.2910
$N = 25$	0.9725	0.2473	**0.1374**	0.2910
$N = 49$	0.5169	0.0985	**0.0666**	0.1209

Source: Author (2020)

4 2D RAYLEIGH AUTOREGRESSIVE MOVING AVERAGE MODEL FOR CHANGE DETECTION IN SAR IMAGERY

Abstract

Two-dimensional (2D) autoregressive moving average (ARMA) models are commonly applied to describe real-world image data, usually assuming Gaussian or symmetric noise. However, real-world data often present signals that are non-Gaussian, presenting asymmetrical distributions and strictly positive values. In particular, SAR images are known to be well characterized by the Rayleigh distribution. In this context, this chapter introduces an ARMA model tailored for 2D Rayleigh-distributed data—the 2D RARMA model. The 2D RARMA model is derived and conditional likelihood inferences are discussed. The proposed model was submitted to extensive Monte Carlo simulations in order to evaluate the performance of the conditional maximum likelihood estimators and the proposed change detector in non-Gaussian images. Moreover, in the context of SAR image processing, two comprehensive numerical experiments were performed comparing the detection and image modeling results of the proposed model with traditional 2D ARMA models and competing methods in the literature.

Keywords: ARMA modeling, change detection, Rayleigh distribution, SAR images, two-dimensional models.

4.1 INTRODUCTION

The parametric representation of two-dimensional homogeneous random fields considering two-dimensional (2D) autoregressive moving average (ARMA) models is frequently adopted for image processing [34–48, 110], including (i) modeling [110, 111]; (ii) compression [112]; (iii) encoding [113]; and (iv) restoration [114–118]. The 2D ARMA model is an spatial extension of the classical one-dimensional ARMA model [119], and is often employed in edge detection [120] and stochastic texture analysis [121].

In particular, the two-dimensional autoregressive first-order model is commonly applied to describe real-world image data [49, 50], representing different types of textures [34]. Theoretical details of the two-dimensional autoregressive first-order model, such as properties, correlation structure, and maximum likelihood estimators of the parameters can be found in [43]. The

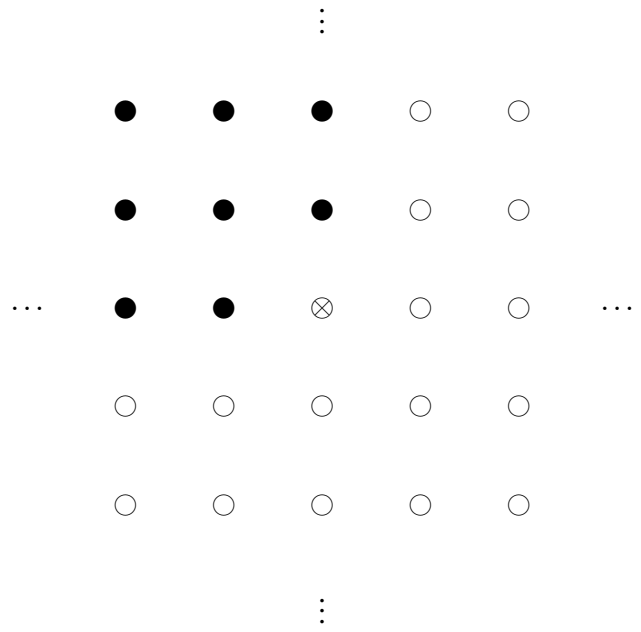
ARMA model is preferable over the autoregressive (AR) and moving average (MA) model, since it provides more effective models for homogeneous random fields [39, 122]. In fact, ARMA models can better characterize the power in the spectral domain representation when compared to AR or MA models [35].

ARMA modeling usually assumes Gaussian or symmetric noise [35, 123]. Indeed, the Gaussianity hypothesis has been widely considered in statistical signal processing [123], remote sensing analysis [124, 125], and detection theory [4, 126, 127]. However, actual measured signals often present non-Gaussian properties [7, 128–130], such as asymmetrical distributions and strictly positive values. An alternative approach for modeling such type of data is the use of the Rayleigh distribution [131]. This distribution is considered in signal and image processing [17, 24–28], being important in the context of synthetic aperture radar (SAR) image modeling, due to its good characterization of image pixel amplitude values [16–18, 29].

To the best of our knowledge, a two-dimensional ARMA model assuming the Rayleigh distribution is not present in the literature and this chapter aims at proposing a first treatment on the topic. Our goal is two-fold. First, we derive a two-dimensional ARMA model for non-Gaussian data, where the observed signal is asymmetric and strictly positive. For the proposed model, we introduce parameter estimation, large data record inference, an image modeling tool based on the derived spatial model estimated parameters, and the quantile residuals. Second, the proposed detector is based on the control charts of the derived spatial model residuals. Control charts have been used before in signal processing for wavelet-based shrinkage [132] and ground type change detection for SAR imagery [133]. The introduced detector was applied to non-Gaussian SAR image ground type change detection.

This chapter is organized as follows. In Section 4.2, we describe the proposed spatial model, provide maximum likelihood estimators, and present a hypothesis testing methodology. Section 4.3 details the introduced image modeling and proposed a ground type change detection algorithm. Section 4.4 presents Monte Carlo simulations and two empirical analyses of the derived change detector applied to SAR images. Section 4.5 brings final remarks and concludes the chapter.

Figure 4 – The strongly causal region at $[n,m]$. The evaluated pixel is represented by $\otimes$ and $\bullet$ is the considered neighborhood.



Source: Author (2020)

4.2 THE PROPOSED MODEL

4.2.1 Mathematical Setup

Recently, a regression model [94] and an one-dimensional ARMA model [131] based on the Rayleigh distribution have been proposed. The Rayleigh ARMA (RARMA) model introduced in [131] relates the mean of one-dimensional signal to a linear predictor through a strictly monotonic, twice differentiable link function $g(\cdot)$, where $g:\mathbb{R}^+\to\mathbb{R}$. The goal of this section is to extend the one-dimensional model presented in [131] and introduce to the 2D case.

Let $Y[n,m]$, $n=1,2,\ldots,N$, $m=1,2,\ldots,M$, be a Rayleigh distributed random variable representing the pixels of an $N\times M$ image; and let $y[n,m]$ be the realization of the observed signal $Y[n,m]$. Additionally, let $S[n,m]=\{[k,l]\in\mathbb{Z}^2:1\leq k\leq n,1\leq l\leq m\}-\{[n,m]\}$ be the strongly causal region at $[n,m]$ [34], as illustrated in Figure 4.

Considering the mean-based parametrization of $Y[n,m]$ proposed in [94], we have that the conditional density of $Y[n,m]$, given $S[n,m]$, is furnished by

$$f_Y(y[n,m]\mid S[n,m])=\frac{\pi y[n,m]}{2\mu[n,m]^2}\exp\left(-\frac{\pi y[n,m]^2}{4\mu[n,m]^2}\right).$$

The cumulative distribution function is given by

$$F_Y(y[n,m]\mid S[n,m])=1-\exp\left(-\frac{\pi y[n,m]^2}{4\mu[n,m]^2}\right).$$

The conditional mean and conditional variance of $Y[n,m]$, given $S[n,m]$, are, respectively, furnished by

$$\begin{aligned}\mathrm{E}(Y[n,m] \mid S[n,m]) &= \mu[n,m],\\ \mathrm{Var}(Y[n,m] \mid S[n,m]) &= \mu[n,m]^2\left(\frac{4}{\pi}-1\right).\end{aligned}$$

4.2.2 The Model

The proposed 2D Rayleigh autoregressive and moving average model, hereafter referred to as 2D RARMA, is defined according to

$$\Phi(z_1,z_2)g(y[n,m]) = \beta + \Theta(z_1,z_2)e[n,m], \tag{4.1}$$

where

$$\begin{aligned}\Phi(z_1,z_2) &= 1 - \sum_{i=0}^{p}\sum_{j=0}^{p}\phi_{(i,j)}z_1^{-i}z_2^{-j},\\ \Theta(z_1,z_2) &= 1 + \sum_{k=0}^{q}\sum_{l=0}^{q}\theta_{(k,l)}z_1^{-k}z_2^{-l},\end{aligned}$$

and $\beta \in \mathbb{R}$ is an intercept; the quantities z_1^{-i}, z_2^{-j}, z_1^{-k}, and z_2^{-l} are the spatial delays; p and q are the orders of the model; the quantities $\phi_{(i,j)}$, $i,j = 0,1,\ldots,p$, and $\theta_{(k,l)}$, $k,l = 0,1,\ldots,q$, are, respectively, the autoregressive and moving average parameters estimated based on the image pixels; $e[n,m] = g(y[n,m]) - g(\mu[n,m])$ is the moving average error term; and $g(\mu[n,m]) = \eta[n,m]$ the linear predictor. As suggested in [43], we assume $\phi_{(0,0)} = \theta_{(0,0)} = 0$. Replacing the quantities described above in (4.1), the 2D RARMA (p,q) model can be rewritten as

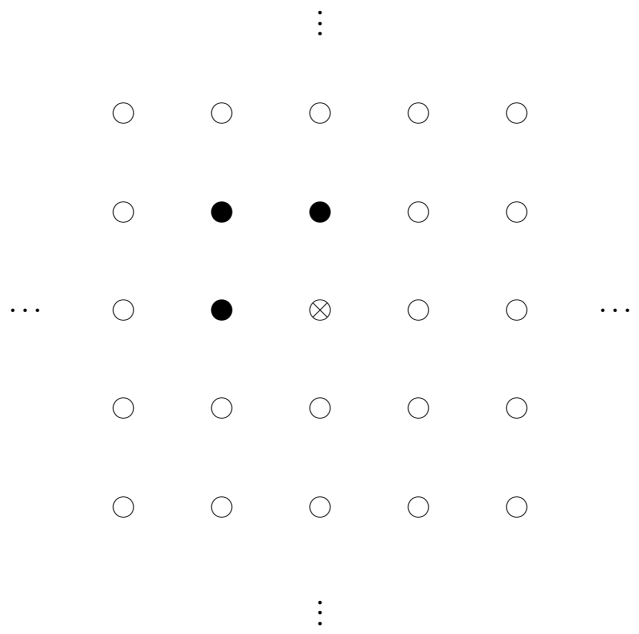
$$g(\mu[n,m]) = \beta + \sum_{i=0}^{p}\sum_{j=0}^{p}\phi_{(i,j)}g(y[n-i,m-j]) + \sum_{k=0}^{q}\sum_{l=0}^{q}\theta_{(k,l)}e[n-k,m-l].$$

Figure 5 depicts the considered neighbors pixels in a 2D RARMA$(1,1)$ model [34]. As usual [134], we require that the AR terms be such that the related characteristic polynomial does not have unit roots and the AR and MA characteristic polynomials possess no common roots.

4.2.3 Conditional Maximum Likelihood Estimation

The estimation of the 2D RARMA(p,q) model parameters can be realized by maximizing the logarithm of the conditional likelihood function [135]. Let $\boldsymbol{\gamma} = (\beta, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top)^\top$ be the parameter vector where $\boldsymbol{\phi} = (\phi_{(0,1)}, \phi_{(0,2)}, \ldots, \phi_{(p,p)})^\top$ and $\boldsymbol{\theta} = (\theta_{(0,1)}, \theta_{(0,2)}, \ldots, \theta_{(q,q)})^\top$,

Figure 5 – Example of the neighborhood used in a 2D RARMA$(1,1)$ model. The evaluated pixel is represented by $\otimes$ and $\bullet$ is the considered neighborhood.



Source: Author (2020)

with dimensions $(p+1)^2-1$ and $(q+1)^2-1$, respectively. The log-likelihood function for $\boldsymbol{\gamma}$, conditional on the $w=\max(p,q)$ preliminary observations, is given by

$$\ell(\boldsymbol{\gamma}) = \sum_{n=w+1}^{N}\sum_{m=w+1}^{M} \log f_Y(y[n,m] \mid S[n,m]) = \sum_{n=w+1}^{N}\sum_{m=w+1}^{M} \ell[n,m](\mu[n,m]),$$

where

$$\ell[n,m](\mu[n,m]) = \log\left(\frac{\pi}{2}\right) + \log(y[n,m]) - \log(\mu[n,m]^2) - \frac{\pi y[n,m]^2}{4\mu[n,m]^2}.$$

The conditional maximum likelihood estimators (CMLE), $\widehat{\boldsymbol{\gamma}}$, can be obtained by solving

$$\mathbf{U}(\boldsymbol{\gamma}) = \frac{\partial \ell}{\partial \boldsymbol{\gamma}^\top} = \left(\frac{\partial \ell}{\partial \beta}, \frac{\partial \ell}{\partial \boldsymbol{\phi}^\top}, \frac{\partial \ell}{\partial \boldsymbol{\theta}^\top},\right)^\top = \mathbf{0}, \tag{4.2}$$

where $\mathbf{U}(\boldsymbol{\gamma})$ is the score vector and $\mathbf{0}$ is the vector of zeros of dimension $(p+1)^2+(q+1)^2-1$. Computing the derivatives in (4.2), we obtain

$$\frac{\partial \ell}{\partial \boldsymbol{\gamma}} = \sum_{n=w+1}^{N}\sum_{m=w+1}^{M} \frac{\partial \ell[n,m](\mu[n,m])}{\partial \mu[n,m]} \frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]} \frac{\partial \eta[n,m]}{\partial \boldsymbol{\gamma}}.$$

Note that

$$\frac{\mathrm{d}\,\ell[n,m](\mu[n,m])}{\mathrm{d}\,\mu[n,m]} = \frac{\pi y[n,m]^2}{2\mu[n,m]^3} - \frac{2}{\mu[n,m]},$$

and

$$\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]} = \frac{1}{g'(\mu[n,m])},$$

where $g'(\cdot)$ is the first derivative of the selected link function $g(\cdot)$. Thus, we can write

$$\frac{\partial \ell}{\partial \boldsymbol{\gamma}} = \sum_{n=w+1}^{N} \sum_{m=w+1}^{M} \left(\frac{\pi y[n,m]^2}{2\mu[n,m]^3} - \frac{2}{\mu[n,m]} \right) \frac{1}{g'(\mu[n,m])} \frac{\partial \eta[n,m]}{\partial \boldsymbol{\gamma}},$$

where

$$\begin{aligned}
\frac{\partial \eta[n,m]}{\partial \beta} &= 1 - \sum_{s=0}^{q} \sum_{t=0}^{q} \theta_{(s,t)} \frac{\partial \eta[n-s,m-t]}{\partial \beta}, \\
\frac{\partial \eta[n,m]}{\partial \phi_{(i,j)}} &= g(y[n-i,m-j]) - \sum_{s=0}^{q} \sum_{t=0}^{q} \theta_{(s,t)} \frac{\partial \eta[n-s,m-t]}{\partial \phi_{(i,j)}}, \\
\frac{\partial \eta[n,m]}{\partial \theta_{(k,l)}} &= g(y[n-k,m-l]) - g(\mu[n-k,m-l]) - \sum_{s=0}^{q} \sum_{t=0}^{q} \theta_{(s,t)} \frac{\partial \eta[n-s,m-t]}{\partial \theta_{(k,l)}},
\end{aligned}$$

for $(i,j) \in \{0,1,\ldots,p\}^2 - \{(0,0)\}$ and $(l,k) \in \{0,1,\ldots,q\}^2 - \{(0,0)\}$.

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [76] with analytic first derivatives was adopted as the nonlinear optimization algorithm [136] to solve (4.2). The BFGS method was selected due to its superior performance for non-linear optimization [77]. The initial values for the constant (β) and the autoregressive ($\boldsymbol{\phi}$) parameters were derived from the ordinary least squares estimate associated to the linear regression model. The response vector is $(g(y[w+1,w+1]), g(y[w+1,w+2]), \ldots, g(y[N,M]))^\top$ and the covariate matrix is given by

$$\begin{bmatrix}
1 & g(y[w,w-1]) & g(y[w-1,w]) & \cdots & g(y[w-p,w-p]) \\
1 & g(y[w,w]) & g(y[w,w]) & \cdots & g(y[w-p+1,w-p+1]) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & g(y[N,M-1]) & g(y[N-1,M]) & \cdots & g(y[N-p,M-p])
\end{bmatrix}.$$

As suggested in [131], we set $\boldsymbol{\theta} = \mathbf{0}$.

4.2.4 Large Data Record Inference

Based on the consistency of the CMLE and on the asymptotic distribution of $\widehat{\boldsymbol{\gamma}}$, for large data record ($N \to \infty$ and $M \to \infty$), we have that [75, 82]

$$\widehat{\boldsymbol{\gamma}} \overset{a}{\sim} \mathcal{N}_u \left(\boldsymbol{\gamma}, \mathbf{I}^{-1}(\boldsymbol{\gamma}) \right),$$

where $\overset{a}{\sim}$ denotes approximately distributed, $u = (p+1)^2 + (q+1)^2 - 1$, and $\mathcal{N}_u$ is the u-dimensional Gaussian distribution with mean $\boldsymbol{\gamma}$ and covariance matrix $\mathbf{I}^{-1}(\gamma)$. The conditional Fisher information matrix, $\mathbf{I}(\gamma)$, is discussed in detail in the Appendix C.

To derive a hypothesis testing methodology tailored for the 2D RARMA model parameters, the likelihood-based detection theory [4, 75] can be considered. Let $\boldsymbol{\gamma}$ be partitioned in

a parameter vector of interest $\boldsymbol{\gamma}_I$ of dimension ν, and a vector of nuisance parameters $\boldsymbol{\gamma}_J$ of dimension $[(p+1)^2+(q+1)^2-1]-\nu$ [4]. In addition, $\mathscr{H}_0 : \boldsymbol{\gamma}_I = \boldsymbol{\gamma}_{I_0}$ is the hypothesis of interest and $\mathscr{H}_1 : \boldsymbol{\gamma}_I \neq \boldsymbol{\gamma}_{I_0}$ the alternative hypothesis, where $\boldsymbol{\gamma}_{I_0}$ is a fixed column vector of dimension ν. The likelihood ratio test (T_L), Wald test (T_W), the and the Rao test (T_R) statistics can be written as [4]

$$
\begin{aligned}
T_L &= 2\left[\ell(\widehat{\boldsymbol{\gamma}}) - \ell(\tilde{\boldsymbol{\gamma}})\right], \\
T_W &= (\widehat{\boldsymbol{\gamma}}_{I1} - \boldsymbol{\gamma}_{I0})^\top \left(\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\gamma}}_1)\right]_{\gamma_I\gamma_I}\right)^{-1} (\widehat{\boldsymbol{\gamma}}_{I1} - \boldsymbol{\gamma}_{I0}), \\
T_R &= \left(\frac{\partial \ell}{\partial \boldsymbol{\gamma}_I}\Big|_{\boldsymbol{\gamma}=\tilde{\boldsymbol{\gamma}}}\right)^\top \left(\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\gamma}})\right]_{\gamma_I\gamma_I}\right)^{-1} \left(\frac{\partial \ell}{\partial \boldsymbol{\gamma}_I}\Big|_{\boldsymbol{\gamma}=\tilde{\boldsymbol{\gamma}}}\right),
\end{aligned}
$$

where $\widehat{\boldsymbol{\gamma}}_{I1}$ and $\widehat{\boldsymbol{\gamma}}_1 = (\widehat{\boldsymbol{\gamma}}_{I1}^\top, \widehat{\boldsymbol{\gamma}}_{J1}^\top)^\top$ are the CMLEs under $\mathscr{H}_1$ (unrestricted MLEs); $\boldsymbol{\gamma}_{J0}$ is a fixed column vector of dimension $r-\nu$; $\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\gamma}})\right]_{\gamma_I\gamma_I}$ is a partition of $\mathbf{I}(\widehat{\boldsymbol{\gamma}})$ limited to the estimates of interest; $\tilde{\boldsymbol{\gamma}} = (\widehat{\boldsymbol{\gamma}}_{I0}^\top, \widehat{\boldsymbol{\gamma}}_{J0}^\top)^\top$ is the MLE under $\mathscr{H}_0$; $\ell(\widehat{\boldsymbol{\gamma}})$ is the maximized log-likelihood of the fitted model; $\ell(\tilde{\boldsymbol{\gamma}})$ is the maximized log-likelihood of the model under $\mathscr{H}_0$; and $\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\gamma}})\right]_{\gamma_I\gamma_I}$ is a partition of $\mathbf{I}(\tilde{\boldsymbol{\gamma}})$ limited to the estimates of interest. As suggested in [83] and due to the convenience of the Wald test, since it requires only one estimation under alternative hypothesis, we selected the Wald test to perform the hypothesis test on the parameters. Under $\mathscr{H}_0$, the test statistic, T_W, asymptotically follows the chi-square distribution with ν degrees of freedom, χ^2_ν [4]. The hypothesis test consists of comparing the computed value of T_W with a threshold value, ε, which is obtained based on the χ^2_ν distribution and the desired probability of false alarm [4].

To test the overall significance of a fitted model, we considered the following hypotheses

$$
\begin{cases}
\mathscr{H}_0 : \boldsymbol{\gamma}^\star = \mathbf{0}, \\
\mathscr{H}_1 : \boldsymbol{\gamma}^\star \neq \mathbf{0},
\end{cases}
\tag{4.3}
$$

where $\boldsymbol{\gamma}^\star = (\boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top)^\top$. Using the Wald test described above, we reject $\mathscr{H}_0$ when $T_W > \varepsilon$ [4]. In this situation, $\boldsymbol{\gamma}^\star \neq \mathbf{0}$, indicating that at least some of the autoregressive and moving average parameters are nonzero and the spatial correlation among the pixels is significant.

4.3 IMAGE MODELING AND CHANGE DETECTION

In this section, we propose an image modeling and change detection tool based on the proposed 2D RARMA model. For such, we introduce the estimated values of $\mu[n,m]$ and present the residuals of the 2D RARMA model.

4.3.1 Image Modeling

The modeled image is obtained by applying the estimated values of $\mu[n,m]$, $\widehat{\mu}[n,m]$, in the 2D RARMA(p,q) model structure, given by (4.1), and evaluating it at $\widehat{\boldsymbol{\gamma}}$. Therefore, the fitted signal is given by

$$\widehat{\mu}[n,m] = g^{-1}\big(\widehat{\beta} + \sum_{i=0}^{p}\sum_{j=0}^{p}\widehat{\phi}_{(i,j)}g(y[n-i,m-j]) + \sum_{k=0}^{q}\sum_{l=0}^{q}\widehat{\theta}_{(k,l)}e[n-k,m-l]\big), \tag{4.4}$$

where $n = w+1, w+2, \ldots, N$ and $m = w+1, w+2, \ldots, M$. Hence, similar to the one-dimensional model, the image border is not included in the modeling process, since the resulting fitted image has $(N-w)\times(M-w)$ pixels.

4.3.2 Change Detector

Residuals can be useful for performing a diagnostic analysis of the fitted model and can be defined as a function of the observed and predicted values [137, 138]. We employed the quantile residuals [78], defined as

$$r[n,m] = \Phi^{-1}\left(F(y[n,m] \mid S[n,m])\right),$$

where Φ^{-1} denotes the standard normal quantile function. When the model is correctly fitted, the residuals are approximately Gaussian distributed with zero mean and unit variance [78, 137].

Large values of $r[n,m]$ can be interpreted as changes in the image behavior. To capture such variations of the residual values, we adopted the use of control charts. Since the residuals have approximately unitary variance, the control chart detects an image change if the residual value is outside the control limit $\pm L$. We adopted $L = 3$, since it is expected that residuals are randomly distributed around zero and inside the interval $[-3,3]$, about 99.7% of the observations, since $2\Phi(L) - 1|_{L=3} \approx 99.7\%$ [132, 133]. If the residual value is outside this range, the analyzed pixel is understood to differ from the expected behavior according to the 2D RARMA model fitted in the region of interest and, consequently, some change might have occurred.

Notice that the proposed model relies on neighboring pixels from the northwest direction, as shown in Figure 5. Thus, to take into account the other directions in an omnidirectional manner, thus ensuring that all surrounding pixels are considered, the 2D RARMA fitting is also applied to the 90°, 180°, and 270° rotated region of interest to capture information from the versions of the southwest, southeast, and northeast directions. Results are combined according to the morphological union of the resulting binary images.

To further increase the performance of the proposed change detector, a post-processing step using mathematical morphological operations, such as erosion, dilation, opening, and closing operations, can be considered [139]. Such operations aim at (i) removing small spurious pixel groups which are regarded as noise and (ii) preventing the splitting of the interest targets into multiple substructures [139]. The resulting data is the detected image. The proposed change detection method is summarized in Algorithm 4.

Algorithm 4: Change detection method based on the 2D RARMA(p,q) model

Input: Rayleigh distributed image $\mathbf{X}_{\text{input}}$

Output: Change detection image $\mathbf{X}_{\text{detected}}$

1) Select region of interest $\mathbf{X}_{\text{selected}} \subset \mathbf{X}_{\text{input}}$ which change detection is to be tested against.

2) Fit the 2D RARMA(p,q) model for the following images:

$\mathbf{X}_0 = \mathbf{X}_{\text{selected}}$

$\mathbf{X}_k = \mathbf{rot90}(\mathbf{X}_{k-1})$,

for $k = 1,2,3$, where $\mathbf{rot90}(\cdot)$ rotates its argument counterclockwise by 90 degrees.

3) For each resulting fitted image, compute residuals $r_k[n,m]$ relative to $\mathbf{X}_{\text{input}}$.

4) Obtain four binary images as follows

if $(r_k[n,m] \leq -3)$ or $(r_k[n,m] \geq 3)$ **then**

$\tilde{X}_k[n,m] \leftarrow 1$

else

$\tilde{X}_k[n,m] \leftarrow 0$

end if

for $k = 0,1,2,3$.

5) Compute binary image from the following pixel-wise Boolean union: $\tilde{\mathbf{X}} \leftarrow \bigcup_{k=0}^{3} \tilde{\mathbf{X}}_k$.

6) Apply morphological operators as a final post-processing step:

$\mathbf{X}_{\text{detected}} \leftarrow \text{post-processing}(\tilde{\mathbf{X}})$.

4.4 NUMERICAL RESULTS

In this section, we aim at evaluating the CMLE of the 2D RARMA model parameters and assessing the performance of the proposed image modeling and change detector. For such, the proposed analyses were performed in the context of SAR image processing. We performed three numerical experiments: (i) a simulated data analysis to assess the proposed estimators and (ii) two computations aiming at ground type detection based on actual SAR images.

4.4.1 Analysis with Simulated Data

Rayleigh distributed data $y[n,m]$ were generated by the inversion method [140] with mean given by (4.1) and logarithm link function. We considered simulations under three sce-

narios: (i) 2D RARMA$(1,0)$ model, (ii) 2D RARMA$(2,0)$ model, and (iii) 2D RARMA$(1,1)$ model. The parameter values were selected based on estimated values of the 2D RARMA model parameters from a SAR image forest region acquired by the airbone CARABAS II system [84], a Swedish ultrawideband (UWB) very-high frequency (VHF) SAR device operating at horizontal (HH) polarization. Details related to the data can be found in [84]. The obtained numerical values of the parameters were $\beta = -0.2031$, $\phi_{(0,1)} = 0.4562$, $\phi_{(1,0)} = 0.4523$, and $\phi_{(1,1)} = -0.1054$, for the 2D RARMA$(1,0)$ model; $\beta = -0.7528$, $\phi_{(0,1)} = 0.4025$, $\phi_{(0,2)} = -0.0861$, $\phi_{(1,0)} = 0.3331$, $\phi_{(1,1)} = 0.0390$, $\phi_{(1,2)} = -0.0764$, $\phi_{(2,0)} = -0.0249$, $\phi_{(2,1)} = -0.0424$, and $\phi_{(2,2)} = 0.0717$, for the 2D RARMA$(2,0)$ model; and $\beta = 0.3569$, $\phi_{(0,1)} = 0.2155$, $\phi_{(1,0)} = 0.2032$, $\phi_{(1,1)} = 0.1500$, $\theta_{(0,1)} = 0.1529$, $\theta_{(1,0)} = 0.1744$, and $\theta_{(1,1)} = 0.1998$, for the 2D RARMA$(1,1)$ model. The number of Monte Carlo replications was set to $1,000$ and the sizes of the synthetic images were $\{10 \times 10; 20 \times 20; 40 \times 40; 80 \times 80\}$.

In order to numerically evaluate the point estimators, we computed the mean, percentage relative bias (RB%), and mean square error (MSE) of the CMLE. Tables 5, 6, and 7 present the simulation results for 2D RARMA$(1,0)$, 2D RARMA$(2,0)$, and 2D RARMA$(1,1)$ models, respectively. As expected, both bias and MSE figures improve when larger images are considered. This behavior is in agreement with the asymptotic property (consistency) of the CMLE. Convergence failures were absent for all considered scenarios. In contrast to the traditional 2D ARMA model, the proposed model avoids the estimation problem of the MA parameters, as discussed in [35, 141]. The proposed model estimates the AR and MA terms simultaneously; and $\widehat{\boldsymbol{\phi}}$ and $\widehat{\boldsymbol{\theta}}$ present closer values of RB(%) for all considered synthetic images. The image size of 40×40 was sufficiently large for accurate inference in 2D RARMA$(1,0)$ model, i.e., MSE and RB(%) values close to zero. On the other hand, the 2D RARMA$(1,1)$ model shows accurate inference results for an image size of 80×80 pixels. The evaluated 2D RARMA$(1,1)$ scenario presents positive percentage relative bias for the AR estimated parameters. On the other hand, the MA estimated parameters have negative values of relative bias. The Monte Carlo simulations are summarized in Algorithm 5.

To evaluate the performance of the ground type change detection proposed Algorithm 4 in a synthetic image, we generated an image according to the 2D RARMA$(1,0)$ model described above; the synthetic image is shown in 6a. We added three white rectangles that are the changes to be detected in the synthetic image. Hence, rectangle regions are expected to trigger a detection, suggesting a ground type change. In the post-processing step in Algorithm 4, we considered two mathematical morphology steps, namely, closing and opening operations. The dilation

Algorithm 5: Monte Carlo simulations for evaluation of the 2D RARMA(p,q) model parameter estimators.

Input: Vector of parameter $\boldsymbol{\gamma}$ and signal length N.
Output: Results of the desirable figures of merit.

1: Suppose that the observed image $y[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\gamma}$;
2: Generate $y[n]$ from $f_Y(\boldsymbol{\gamma})$;
3: For each Monte Carlo replication, compute $\widehat{\boldsymbol{\gamma}}$;
4: Repeat steps 2 and 3 a very large number R of times, obtaining: $\widehat{\boldsymbol{\gamma}}[1], \widehat{\boldsymbol{\gamma}}[2], \ldots, \widehat{\boldsymbol{\gamma}}[R]$;
5: Use the estimates $\widehat{\boldsymbol{\gamma}}[1], \widehat{\boldsymbol{\gamma}}[2], \ldots, \widehat{\boldsymbol{\gamma}}[R]$, to calculate the desired measures (mean, bias, confidence intervals, and mean square error).

Table 5 – Simulation results on point estimation of the 2D RARMA$(1,0)$ model

Measures	Mean	RB(%)	MSE
$N = M = 10$			
$\widehat{\beta}$	-0.2629	29.4370	0.0244
$\widehat{\phi}_{(0,1)}$	0.4385	-3.8767	0.0071
$\widehat{\phi}_{(1,0)}$	0.4339	-4.0692	0.0073
$\widehat{\phi}_{(1,1)}$	-0.0998	-5.2979	0.0081
$N = M = 20$			
$\widehat{\beta}$	-0.2160	6.3468	0.0051
$\widehat{\phi}_{(0,1)}$	0.4523	-0.8514	0.0014
$\widehat{\phi}_{(1,0)}$	0.4492	-0.6841	0.0014
$\widehat{\phi}_{(1,1)}$	-0.1053	-0.0223	0.0016
$N = M = 40$			
$\widehat{\beta}$	-0.2063	1.5865	0.0012
$\widehat{\phi}_{(0,1)}$	0.4551	-0.2409	0.0003
$\widehat{\phi}_{(1,0)}$	0.4514	-0.2021	0.0003
$\widehat{\phi}_{(1,1)}$	-0.1049	-0.4706	0.0004
$N = M = 80$			
$\widehat{\beta}$	-0.2043	0.6020	0.0003
$\widehat{\phi}_{(0,1)}$	0.4560	-0.0464	0.0001
$\widehat{\phi}_{(1,0)}$	0.45164	-0.1504	0.0001
$\widehat{\phi}_{(1,1)}$	-0.1052	-0.1450	0.0001

Source: Author (2020)

Table 6 – Simulation results on point estimation of the 2D RARMA$(2,0)$ model

Measures	Mean	RB(%)	MSE
	$N = M = 10$		
$\widehat{\beta}$	-0.9386	24.6750	0.2104
$\widehat{\phi}_{(0,1)}$	0.3903	-3.0300	0.0123
$\widehat{\phi}_{(0,2)}$	-0.1002	16.3707	0.0114
$\widehat{\phi}_{(1,0)}$	0.3197	-4.0359	0.0130
$\widehat{\phi}_{(1,1)}$	0.0335	-14.1328	0.0155
$\widehat{\phi}_{(1,2)}$	-0.0694	-9.2135	0.0126
$\widehat{\phi}_{(2,0)}$	-0.0469	88.5405	0.0118
$\widehat{\phi}_{(2,1)}$	-0.0362	-14.6376	0.0135
$\widehat{\phi}_{(2,2)}$	0.0590	-17.6877	0.0123
	$N = M = 20$		
$\widehat{\beta}$	-0.7879	4.6637	0.0287
$\widehat{\phi}_{(0,1)}$	0.4011	-0.3541	0.0020
$\widehat{\phi}_{(0,2)}$	-0.0872	1.2690	0.0018
$\widehat{\phi}_{(1,0)}$	0.3307	-0.7331	0.0019
$\widehat{\phi}_{(1,1)}$	0.0363	-6.9410	0.0025
$\widehat{\phi}_{(1,2)}$	-0.0754	-1.2448	0.0021
$\widehat{\phi}_{(2,0)}$	-0.0279	12.0074	0.0019
$\widehat{\phi}_{(2,1)}$	-0.0408	-3.7711	0.0022
$\widehat{\phi}_{(2,2)}$	0.0674	-6.0087	0.0019
	$N = M = 40$		
$\widehat{\beta}$	-0.7586	0.7675	0.0057
$\widehat{\phi}_{(0,1)}$	0.4023	-0.0443	0.0004
$\widehat{\phi}_{(0,2)}$	-0.0858	-0.3960	0.0004
$\widehat{\phi}_{(1,0)}$	0.3321	-0.2947	0.0004
$\widehat{\phi}_{(1,1)}$	0.0388	-0.4656	0.0006
$\widehat{\phi}_{(1,2)}$	-0.0762	-0.2910	0.0005
$\widehat{\phi}_{(2,0)}$	-0.0248	-0.2843	0.0004
$\widehat{\phi}_{(2,1)}$	-0.0422	-0.4756	0.0005
$\widehat{\phi}_{(2,2)}$	0.0705	-1.6577	0.0004
	$N = M = 80$		
$\widehat{\beta}$	-0.7540	0.1607	0.0015
$\widehat{\phi}_{(0,1)}$	0.4026	0.0281	0.0001
$\widehat{\phi}_{(0,2)}$	-0.0862	0.1668	0.0001
$\widehat{\phi}_{(1,0)}$	0.3333	0.0707	0.0001
$\widehat{\phi}_{(1,1)}$	0.0384	-1.5823	0.0001
$\widehat{\phi}_{(1,2)}$	-0.0763	-0.1536	0.0001
$\widehat{\phi}_{(2,0)}$	-0.0251	0.8129	0.0001
$\widehat{\phi}_{(2,1)}$	-0.0421	-0.7604	0.0001
$\widehat{\phi}_{(2,2)}$	0.0715	-0.2946	0.0001

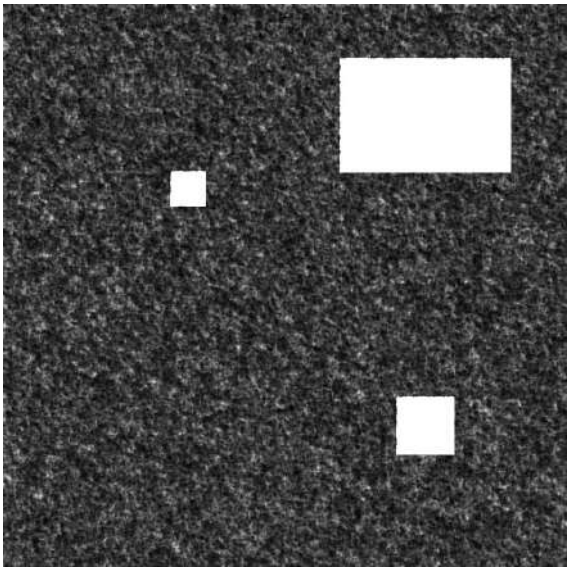
Source: Author (2020)

Table 7 – Simulation results on point estimation of the 2D RARMA$(1,1)$ model

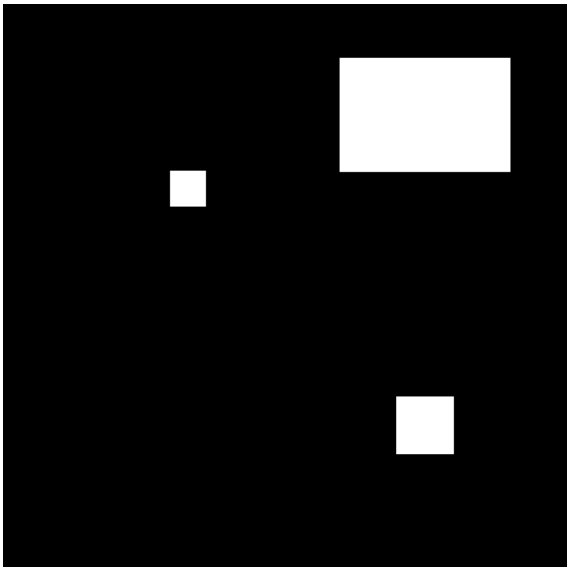
Measures	Mean	RB(%)	MSE
	$N = M = 10$		
$\widehat{\beta}$	0.2834	-20.5927	0.0223
$\widehat{\phi}_{(0,1)}$	0.2278	5.7302	0.0405
$\widehat{\phi}_{(1,0)}$	0.2266	11.5074	0.0388
$\widehat{\phi}_{(1,1)}$	0.1758	17.1830	0.0483
$\widehat{\theta}_{(0,1)}$	0.1371	-10.3477	0.0583
$\widehat{\theta}_{(1,0)}$	0.1413	-18.9945	0.0604
$\widehat{\theta}_{(1,1)}$	0.1582	-20.8355	0.0493
	$N = M = 20$		
$\widehat{\beta}$	0.3312	-7.2026	0.0036
$\widehat{\phi}_{(0,1)}$	0.2246	4.2177	0.0078
$\widehat{\phi}_{(1,0)}$	0.2124	4.5210	0.0081
$\widehat{\phi}_{(1,1)}$	0.1689	12.6150	0.0095
$\widehat{\theta}_{(0,1)}$	0.1428	-6.5797	0.0090
$\widehat{\theta}_{(1,0)}$	0.1673	-4.0788	0.0089
$\widehat{\theta}_{(1,1)}$	0.1753	-12.2650	0.0061
	$N = M = 40$		
$\widehat{\beta}$	0.3456	-3.1658	0.0009
$\widehat{\phi}_{(0,1)}$	0.2195	1.8514	0.0018
$\widehat{\phi}_{(1,0)}$	0.2077	2.2177	0.0019
$\widehat{\phi}_{(1,1)}$	0.1610	7.3560	0.0023
$\widehat{\theta}_{(0,1)}$	0.1499	-1.9723	0.0020
$\widehat{\theta}_{(1,0)}$	0.1711	-1.9205	0.0020
$\widehat{\theta}_{(1,1)}$	0.1878	-5.9962	0.0013
	$N = M = 80$		
$\widehat{\beta}$	0.3514	-1.5381	0.0002
$\widehat{\phi}_{(0,1)}$	0.2184	1.3555	0.0005
$\widehat{\phi}_{(1,0)}$	0.2059	1.3082	0.0005
$\widehat{\phi}_{(1,1)}$	0.1549	3.2791	0.0006
$\widehat{\theta}_{(0,1)}$	0.1503	-1.7221	0.0005
$\widehat{\theta}_{(1,0)}$	0.1717	-1.5643	0.0005
$\widehat{\theta}_{(1,1)}$	0.1932	-3.3146	0.0003

Source: Author (2020)

Figure 6 – Synthetic and detected image based on the 2D RARMA$(1,0)$ model. The white rectangles are related to the changes.



(a) Synthetic image



(b) Detected image

Source: Author (2020)

considered in both steps used a 10×10 pixel square structuring element, since the evaluated areas are bigger than 10×10 pixels. Figure 6b presents the change detection results. Our proposed model detected the three targets and no one false alarm, validating the proposed change detector in a synthetic image.

4.4.2 Analysis with Real Data

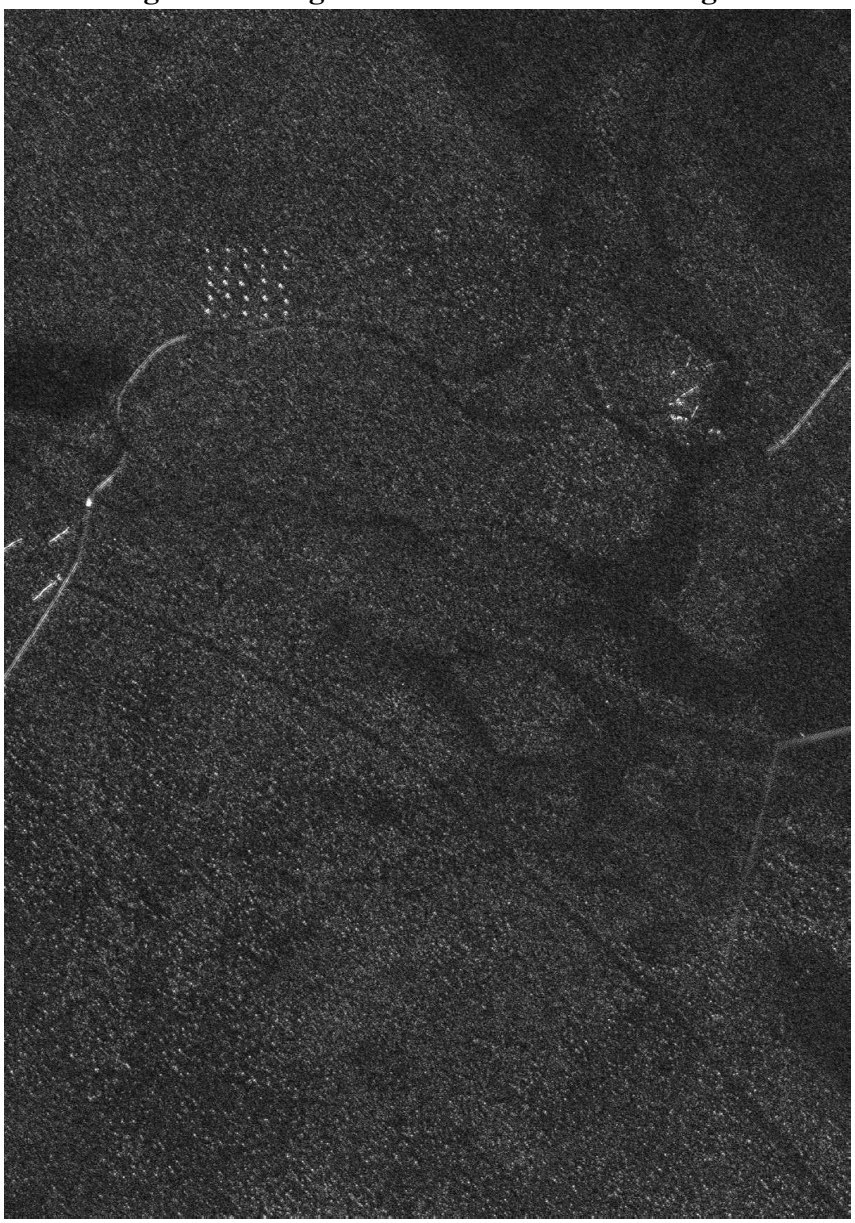
4.4.2.1 CARABAS II

The SAR image considered in this experiment was collected by the CARABAS II system described in the previous subsection. The ground scene of the selected image is dominated by pine forest, a lake, and 25 military vehicles [84]. The forest and lake regions characterize most of the image area and they follow a homogeneous pattern. The military vehicles deployed in the SAR scene [84] introduce more representative behavior changing when compared to the forest and lake regions. For instance, the sample mean value of forest, lake and military vehicles areas are about 0.1267, 0.1148, and 0.2863, respectively, i.e., the sample mean value of military vehicles region is roughly three times the forest and lake regions. Thus, only changes related to the target areas are expected. The considered SAR image in this experiment is shown in Figure 7.

The model selection was based on an exhaustive search aiming at the maximization of detection results. The search space was restricted to models with $(p,q) \in \{0,1,2\}$ and the size of the considered region of interest was $N = M \in \{10,20,40,80\}$. As a result, we obtained the following optimal parameters: (i) $p = q = 1$ and (ii) $N = M = 80$. The selected region of interest in this section was forest, as presented in Figure 8. For the post-processing step, we followed the methodology defined in [84], where we considered two morphology operations, namely, an opening operation followed by a dilation. The opening uses a 3×3 pixel square structuring element, whose size is determined by the system resolution; the dilation considers a 7×7 pixel structuring element which is linked to the approximate size of the military vehicles. Table 8 shows the estimated parameters for each rotated image, as described in the second step of the Algorithm 4. The neighborhood considered for each rotated image is shown in Figure 9. The Wald test p-value can be found in Table 8, showing that the spatial autocorrelation is significant for a probability of false alarm equal to 0.05.

Figures 10 and 11 show the residual and detected images for the four rotated images, respectively; the binary image from the pixel-wise Boolean union is presented in Figure 12. For comparison purposes, we also obtained the detection results based on the 2D ARMA$(1,1)$ model. Detection results for both 2D RARMA$(1,1)$ and 2D ARMA$(1,1)$ models can be found in Figures 13a and 13b, respectively. The proposed method detected 24 military vehicles and five false alarms. In contrast, the 2D ARMA$(1,1)$ model can only detect 16 military vehicles and two false alarms.

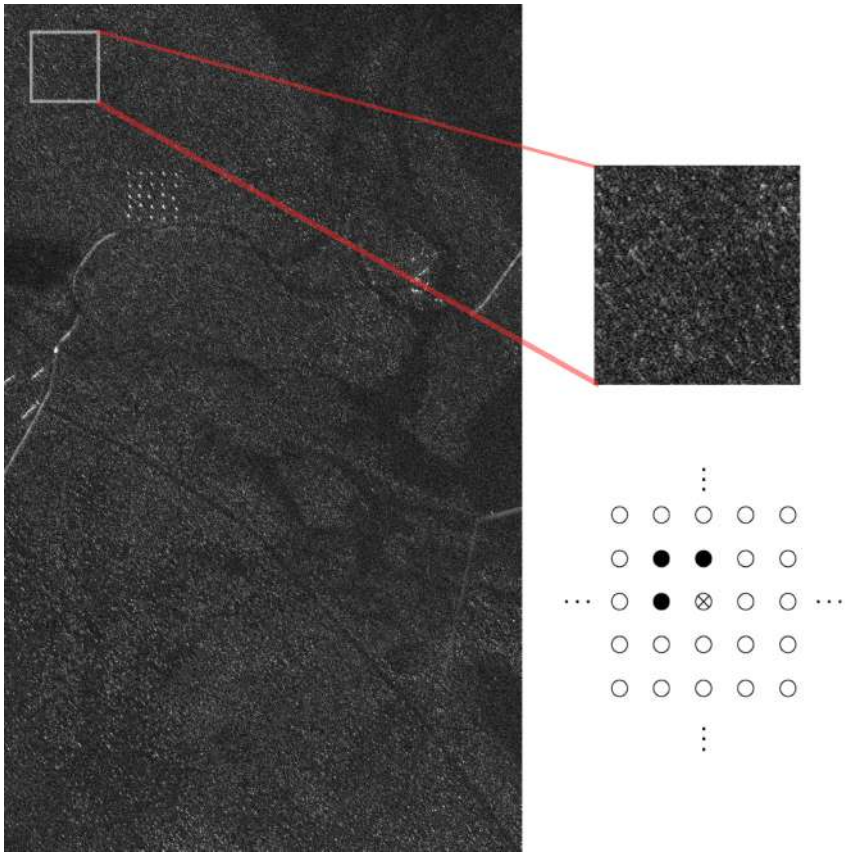
Figure 7 – Original CARABAS II SAR image.



Source: Author (2020)

We also compared the proposed methodology with three different competing approaches: (i) constant false alarm rate filtering combined with likelihood ratio test assuming the Gaussian distribution [84], (ii) a statistical hypothesis test for wavelength-resolution incoherent SAR change detection based on the bivariate gamma distribution [62]; and (iii) a statistical hypothesis test considering the bivariate Gaussian distribution [1]. Differently from the above methods, our detection scheme requires only one input image for analysis; whereas two images are demanded in [84], [62], respectively, and three, in [1]. Despite requiring much less assumptions and data information (only one image scene look vs several image scene looks) when compared to [1,62,84], the proposed 2D RARMA$(1,1)$ model performance was very close to the competing

Figure 8 – Original CARABAS II SAR image, region of interest, and the neighborhood considered to fit the 2D RARMA model.



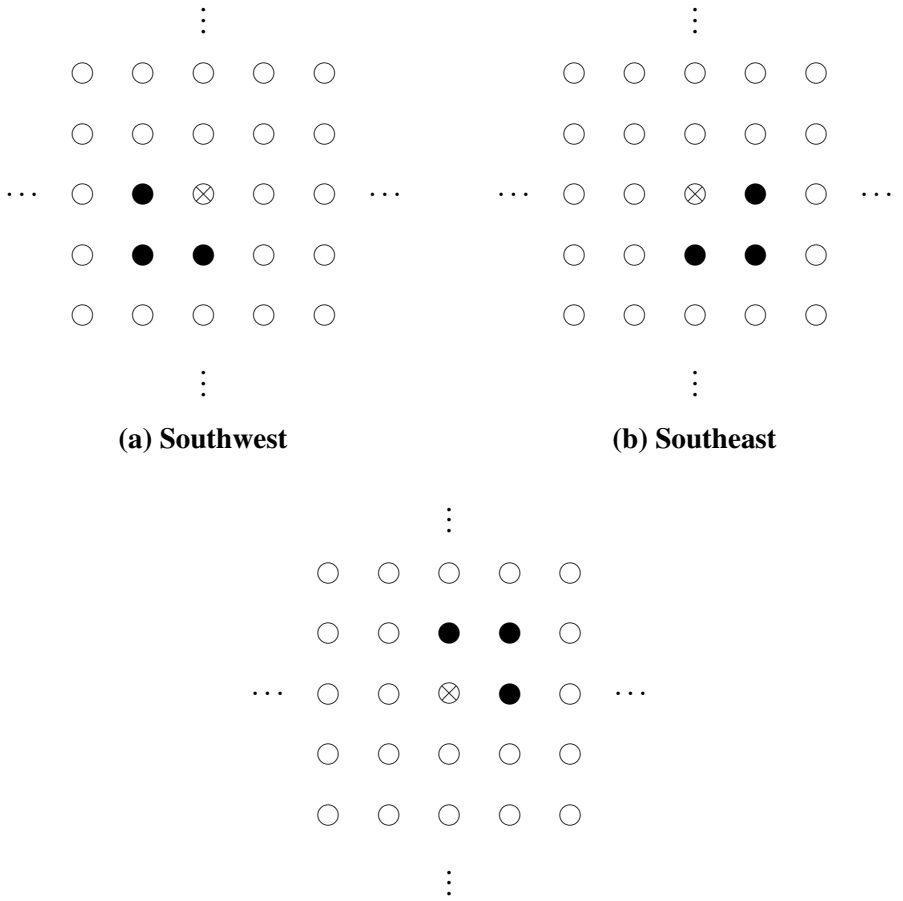
Source: Author (2020)

Table 8 – Estimated parameters and p-values of the 2D RARMA$(1,1)$ model for the SAR image of HH associated polarization channel

	Rotated Image			
	Northwest	Southwest	Southeast	Northeast
$\widehat{\beta}$	-1.2274	-1.1146	-1.1986	-1.2076
$\widehat{\phi}_{(0,1)}$	0.1723	0.1659	0.2218	0.1912
$\widehat{\phi}_{(1,0)}$	0.1526	0.2206	0.1572	0.1616
$\widehat{\phi}_{(1,1)}$	0.0675	0.0512	0.0294	0.0387
$\widehat{\theta}_{(0,1)}$	0.1773	0.1263	0.1305	0.1127
$\widehat{\theta}_{(1,0)}$	0.1646	0.1208	0.1808	0.1685
$\widehat{\theta}_{(1,1)}$	0.1935	-0.0691	0.2064	-0.0461
p-value	< 0.001	0.0012	< 0.001	0.001

Source: Author (2020)

Figure 9 – The neighborhood considered for each rotated image.



(a) Southwest **(b) Southeast**

(c) Northeast

Source: Author (2020)

methods: only one less detection hit; and 3 to 5 more false alarms.

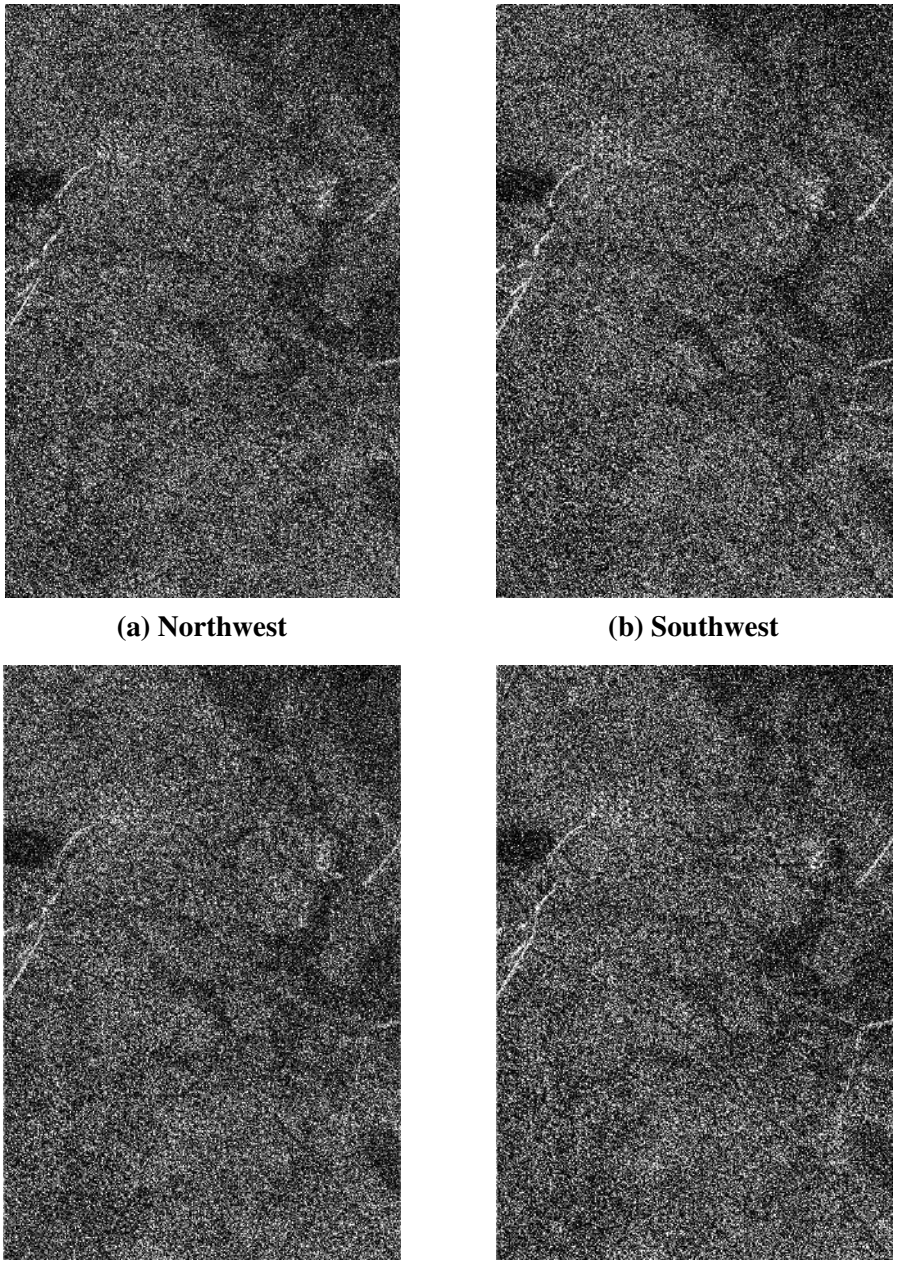
To further compare the image modeling performance of the evaluated models, we employed two traditional figures of merit: the mean square error (MSE) [142] and the mean absolute percentage error (MAPE) [142], which can be defined as follows

$$\mathrm{MSE} = \frac{1}{N \cdot M} \sum_{n=1}^{N} \sum_{m=1}^{M} (y[n,m] - \widehat{\mu}[n,m])^2,$$

$$\mathrm{MAPE} = \frac{1}{N \cdot M} \sum_{n=1}^{N} \sum_{m=1}^{M} \frac{|y[n,m] - \widehat{\mu}[n,m]|}{|y[n,m]|}.$$

The quality adjustment measures are expected to be as close to zero as possible. Table 9 summarizes the results of the quality adjustment measures for 2D RARMA$(1,1)$ and 2D ARMA$(1,1)$ models. The 2D RARMA$(1,1)$ model excels in terms of MSE and MAPE measures.

Figure 10 – Residual images of the four considered neighborhoods.



(a) Northwest **(b) Southwest**

(c) Southeast **(d) Northeast**

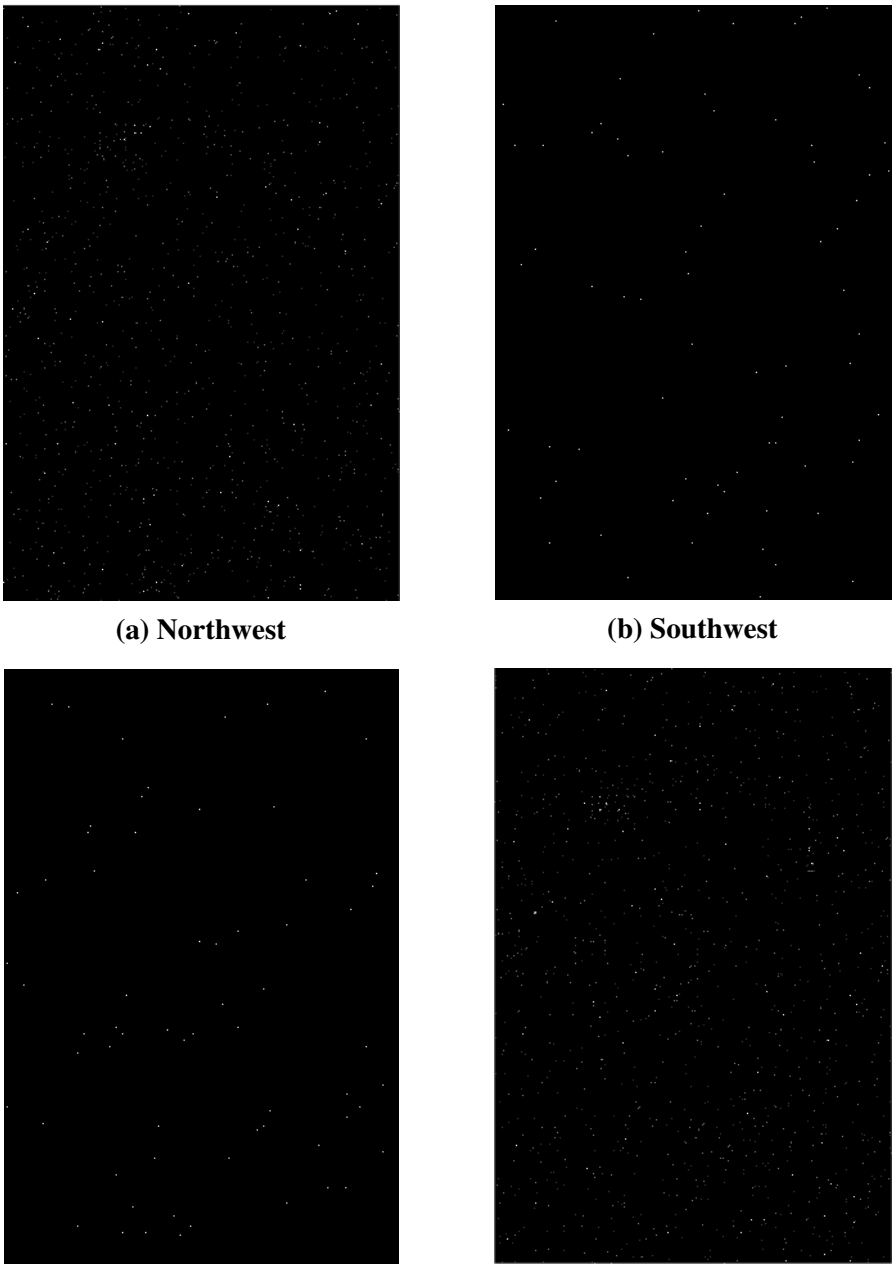
Source: Author (2020)

Table 9 – Measures of quality of the fitted CARABAS II SAR image based on 2D RARMA$(1,1)$ and 2D ARMA$(1,1)$ models

	Model	
	2D RARMA$(1,1)$	2D ARMA$(1,1)$
MSE	0.0562	0.1241
MAPE	0.4277	0.7499

Source: Author (2020)

Figure 11 – Detected images of the four considered neighborhoods.

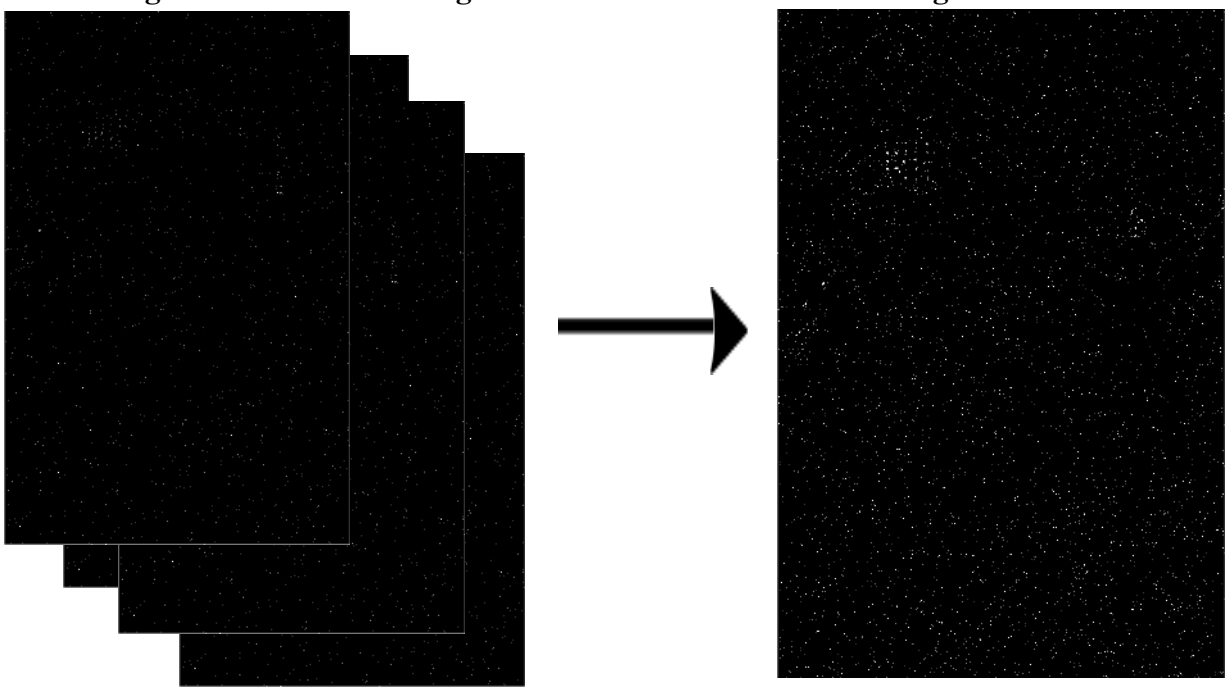


(a) Northwest **(b) Southwest**

(c) Southeast **(d) Northeast**

Source: Author (2020)

Figure 12 – Detected image based on the four considered neighborhoods.



Source: Author (2020)

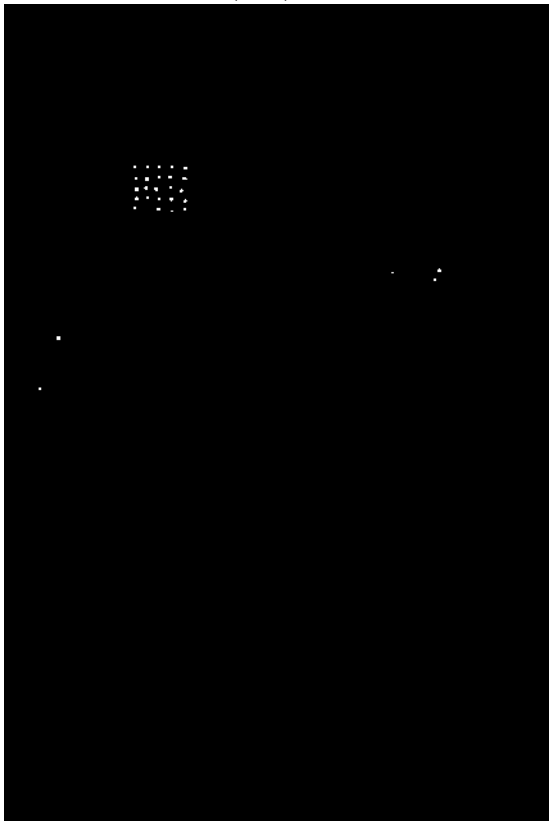
4.4.2.2 San Francisco Bay

The considered SAR image considered in this section is the San Francisco Bay, obtained by the AIRSAR sensor with four looks [15]. Figure 14 shows the amplitude data of the 200×350 San Francisco Bay image associated to the HH polarization channel. The ground scene of the evaluated image is dominated by ocean (dark ground–top and left part of the image), forest (gray ground), and urban area (light ground–bottom) [143].

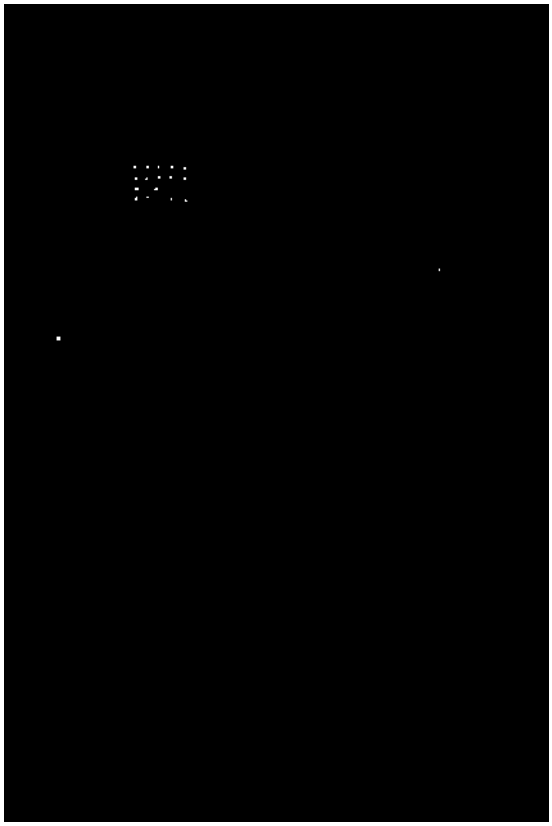
To perform the ground type change detection in the San Francisco SAR image, we set the following parameters in Algorithm 4 adopting the same methodology described in the previous subsection. We have that (i) $p = 1$ and $q = 0$; (ii) $N = M = 40$. The region of interest in this section is linked to the ocean area. Hence, non-ocean regions (forest and urban areas) are expected to trigger a detection, suggesting a ground type change. In the post-processing step, we considered two mathematical morphology steps, namely, closing and opening operations. The dilation considered in both steps used a 10×10 pixel square structuring element, since the evaluated areas are bigger than 10×10 pixels. Figures 15 and 16 show the residual and detected images for the four rotated images, respectively.

Because $q = 0$, we have that $\boldsymbol{\theta} = \mathbf{0}$, and therefore, $\boldsymbol{\gamma}^{\star} = (\boldsymbol{\phi}^{\top}, \mathbf{0}^{\top})^{\top}$ in (4.3). The estimated parameters for each rotated image, as described in the second step of the Algorithm 4, are shown in Table 10. The p-values of the Wald test are also reported in Table 10, indicating that the

Figure 13 – Detected image based on 2D RARMA$(1,0)$ and 2D ARMA$(1,0)$ models. The white dots are the detected pixels. The proposed method detected 24 military vehicles while the 2D ARMA$(1,1)$ model can only detect 16 military vehicles.



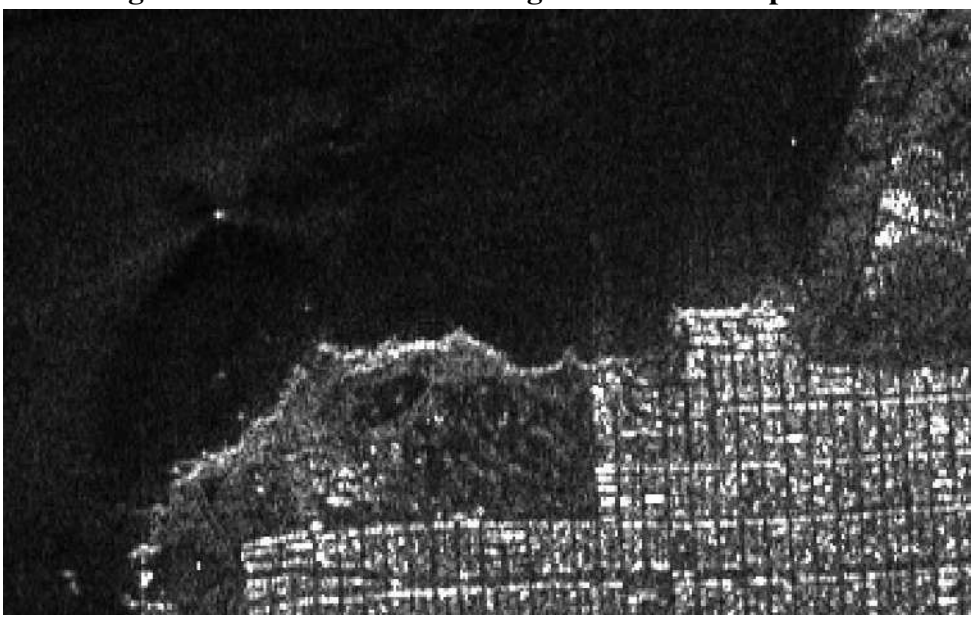
(a) 2D RARMA$(1,0)$ model



(b) 2D ARMA$(1,0)$ model

Source: Author (2020)

Figure 14 – Original San Francisco SAR image HH associated polarization channel.



Source: Author (2020)

Table 10 – Estimated parameters and p-values of the 2D RARMA$(1,0)$ model for the SAR image of HH associated polarization channel

	Fit			
	Northwest	Southwest	Southeast	Northeast
$\widehat{\beta}$	-1.2078	-1.1879	-1.2109	-1.1095
$\widehat{\phi}_{(0,1)}$	0.1408	0.4388	0.1516	0.4350
$\widehat{\phi}_{(1,0)}$	0.4412	0.1236	0.4389	0.1608
$\widehat{\phi}_{(1,1)}$	-0.0023	0.0247	-0.0071	0.0238
p-value	< 0.001	< 0.001	< 0.001	< 0.001

Source: Author (2020)

spatial autocorrelation is significant for a probability of false alarm equal to 0.05.

The ground type detection results can be found in Figure 17. Detection results were compared to the ones based on the 2D ARMA$(1,0)$. The results for the detectors base on the 2D RARMA and 2D ARMA can be found in Figures 17a and 17b, respectively. Both evaluated detectors identified the difference among the ocean ground type from the urban and forest areas in the tested SAR image, excepted for the area highlighted by the blue circle in Figure 17b. The original San Francisco and detected images related to horizontal and vertical (HV) and vertical (VV) polarization channels can be found in Appendix D. Both evaluated detectors identified the difference among the ocean ground type from the urban and forest areas in the tested SAR image HV associated polarization channel. On the other hand, in the image associated to the VV channel, the proposed detection method can not distinguish the ocean ground type from

Figure 15 – Residual images of the four considered neighborhoods.

(a) Northwest

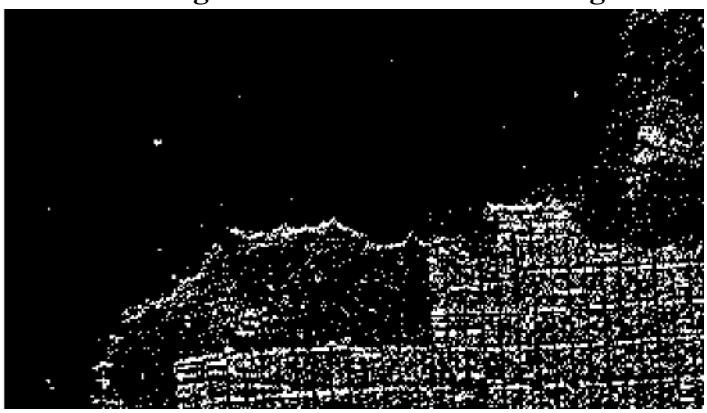
(b) Southwest

(c) Southeast

(d) Northeast

Source: Author (2020)

Figure 16 – Detected images of the four considered neighborhoods.



(a) Northwest

(b) Southwest

(c) Southeast

(d) Northeast

Source: Author (2020)

Table 11 – Measures of quality of the fitted San Francisco SAR image based on 2D RARMA$(1,0)$ and 2D ARMA$(1,0)$ models

	Model	
	2D RARMA$(1,0)$	2D ARMA$(1,0)$
MSE	0.2255	0.3191
MAPE	0.3405	0.9711

Source: Author (2020)

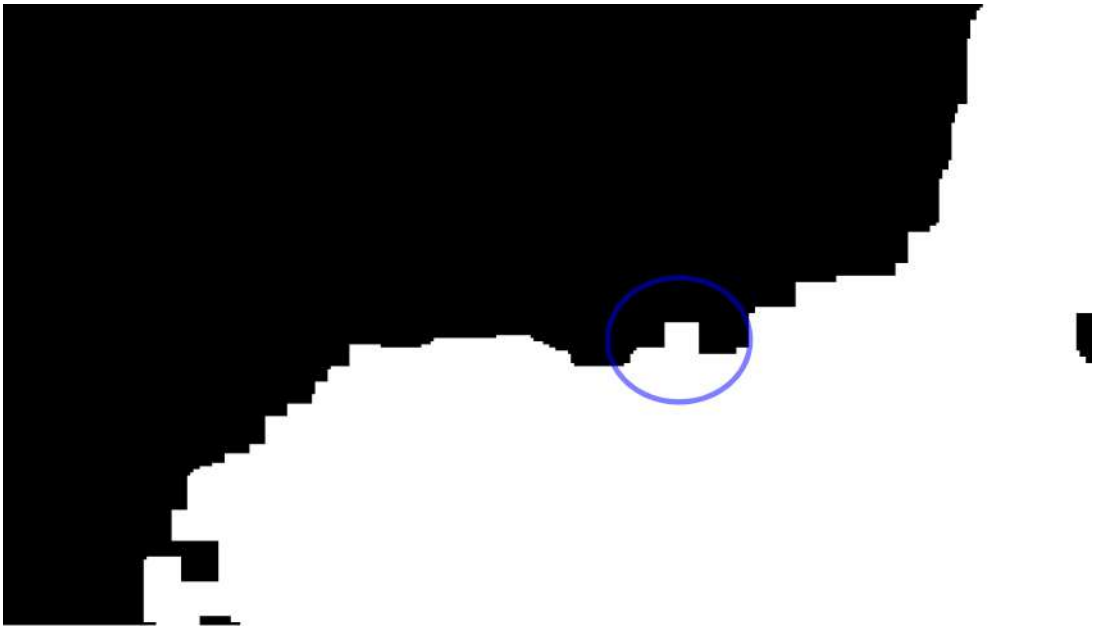
the forest area, while the Gaussian-based detector can not distinguish the ocean ground type from the forest and urban areas. These results may be related to the fact that the SAR image VV associated polarization channel has more noise in the ocean area when compared to the HH and HV polarization channels. Finally, we computed the MSE and MAPE figures of merit to evaluate the fitted image. The 2D RARMA$(1,0)$ model outperforms the alternative model in term of MSE and MAPE measures, as can be verified in Table 11. This experiment shows the resilience of the proposed change detection algorithm, since it presented reliable results of ground type change detection and modeling in an multi-look image.

4.5 CONCLUSIONS

In this chapter, we proposed the 2D RARMA model and derived an image change detector considering residual-based control charts. We introduced an inference approach for the model parameters, the conditional Fisher information matrix, and the asymptotic properties of the CMLE. Monte Carlo simulations were used to evaluate the performance of the CMLE. The proposed model was applied for ground type change detection in SAR images, showing competitive results when compared to 2D ARMA models. Moreover, although the proposed approach requires much less information when compared to [1,62,84], it offered detection results very close to the figures reported in [1,62,84] for the CARABAS II SAR image. The proposed model is presented as a suitable tool for image modeling and change detection in the context of Rayleigh distributed data, in general, and SAR image processing, in particular.

Figure 17 – Detected image associated to the HH polarization channel based on the 2D RARMA$(1,0)$ and 2D ARMA$(1,0)$ models. The white area is related to the detected changes and the blue circle is linked to the difference in the detection results.

(a) 2D RARMA$(1,0)$ model]



(b) 2D ARMA$(1,0)$ model

Source: Author (2020)

5 SIGNAL DETECTION BASED ON THE BETA BINOMIAL AUTOREGRESSIVE MOVING AVERAGE MODEL

Abstract

This chapter proposes the beta binomial autoregressive moving average model (BBARMA) for detection in digital signals. The BBARMA model estimates the mean of a beta binomial distributed variable observed over the time by a dynamic structure including: (i) autoregressive and moving average terms; (ii) a set of regressors; and (iii) a link function. Besides introducing the new model, we develop parameter estimation and detection tools. In particular, we provide closed-form expressions for the conditional information matrix. The proposed model was submitted to extensive Monte Carlo simulations in order to evaluate the performance of the conditional maximum likelihood estimators and of the proposed detector. The proposed detector outperforms the usual ARMA- and Gaussian-based detectors for sinusoidal signal detection.

Keywords: ARMA filter, beta binomial distribution, detection, digital signal, time series.

5.1 INTRODUCTION

Signal detection is a fundamental task in the field of signal processing, being pivotal for decision making and information extraction [4]. Over the years, several detection methods have been developed assuming additive Gaussian noise [4–6, 126, 127, 144–156], with decision criteria based on continuous-time waveforms [4]. In contrast, real-world data often present non-Gaussian signals [7, 128–130] and the Gaussianity assumption may not be enough to model several practical contexts, as illustrated in [155–163]. Furthermore, signal processing systems operate under quantized discrete-time sampled data. Quantized discrete-time signals—here referred to as 'digital signals'—constitute a clear example of non-Gaussian data and their estimation and detection have been attracting attention over the past years [86, 126, 156, 164–179]. For instance, in [180], the authors examine signal detection and scale estimation from unlabeled quantized samples. Quantizer design for weak-signal detection under arbitrary binary channel in generalized Gaussian noise is studied in [156]. In [169], the detection problem of sparse stochastic signals with quantized measurements is addressed. A maximum likelihood detector for quantized pulse-amplitude modulation signal is proposed in [171]. An investigation on the

effects of data quantization in constant false alarm rate signal detection is discussed in [172]. The performance degradation from one-bit quantized detection is discussed in [126].

Quantized signals represent data according to a finite number of discrete values observed over time (e.g., 256 amplitude levels in 8-bit quantization) [13, 51, 56, 181]. Thus, the application of Gaussian-based detectors and hypothesis tests to non-Gaussian processes may lead to suboptimal or erroneous detectors [148, 182]. Although unsatisfactory, a better approach is to consider a data transformation that maps the measured signal from its original distribution into the Gaussian distribution [83]. This method might offer limited results because the transformed data should be interpreted in terms of the transformed signal mean, not in terms of the measured data mean [83, 183, 184].

In this context, the beta binomial distribution offers another way to model the mean of quantized discrete-time signals. The beta binomial distribution has been used over the years to model bounded discrete values [52–55], arising as a natural candidate for digital signals modeling. Nevertheless, to the best of our knowledge, a time series model based on the beta binomial distribution capable of addressing the detection problem for digital signal is absent in the literature. In other words, we are pursuing models and detection methods under the following hypothesis: (i) the measured data is quantized, (ii) the data follow a beta binomial distribution, and (iii) outliers are not present.

Our goal is twofold. First, we introduce a time series model for quantized amplitude data, which estimates the mean of beta binomial distributed signals. The sought model consists of autoregressive and moving average terms, a set of regressors, and a link function. For the proposed beta binomial autoregressive moving average (BBARMA) model, we introduce parameter estimation and the conditional observed information matrix. Second, we present a signal detector based on the asymptotic properties of the sought model parameter estimators. The proposed detector is suitable for identifying the presence of particular signals from beta binomial distributed quantized measured data.

The chapter is organized as follows. In Section 5.2, we provide the mathematical formalism of the derived model. Section 5.3 shows the proposed detection theory, presenting the conditional observed information matrix, a hypothesis test, and the implied signal detector. Section 5.4 presents Monte Carlo simulations for evaluating the derived conditional maximum likelihood and assessing the performance of the proposed detector for simulated digital signals. Finally, Section 5.5 concludes the work.

5.2 THE PROPOSED MODEL

The beta binomial model was proposed in [185], but the idea of this distribution goes back to E. Pearson [186]. The beta binomial distribution is a composition of the beta and the binomial distribution, where the variable of interest, Y, is a random variable with binomial distribution, where the probability of success follows a beta distribution [187]. The beta binomial probability function (pf) is given by [188]

$$p_Y(y;K,a,b) = \frac{\Gamma(a+b)\Gamma(K+1)}{\Gamma(a)\Gamma(b)\Gamma(y+1)\Gamma(K-y+1)}\frac{\Gamma(a+y)\Gamma(K-y+b)}{\Gamma(K+a+b)},$$

where a,b are strictly positive numbers, $\Gamma(\cdot)$ is the gamma function [189], and K is a positive integer. The quantity $y = 0,1,2,\ldots,K$ can be interpreted as an observed signal value and K, as its maximum value. The mean and the variance of Y are, respectively,

$$\begin{aligned} \mathrm{E}(Y) &= \frac{a}{a+b}K, \\ \mathrm{Var}(Y) &= \frac{ab}{(a+b)^2(a+b+1)}\left(K^2+(a+b)K\right), \end{aligned}$$

where $\mu = a/(a+b)$ which can be understood as the mean of Y/K. In the next section, we introduce a dynamic model to fit the mean parameter.

5.2.1 Time Series Model

In order to define the BBARMA model, we consider a new parameterization in the beta binomial distribution based on (i) a precision parameter, φ, and (ii) the quantity μ. The proposed parameters satisfy the following relations: $a = \mu\varphi$ and $b = (1-\mu)\varphi$.

Let $\{Y[n]\}_{n\in\mathbb{Z}}$ be a stochastic process, where each $Y[n]$ assumes values $y[n]$ between zero and K. Let $\mathscr{F}[n]$ be the sigma-field generated by past observations $\{\ldots,y[n-2],y[n-1],y[n]\}$. Assume that, conditionally to the previous information set $\mathscr{F}[n-1]$, each $Y[n]$ is distributed according to the beta binomial distribution with parameters $\mu[n]$ and φ, where $\mu[n]$ is the conditional mean of $Y[n]/K$. The conditional probability function of $Y[n]$ given $\mathscr{F}[n-1]$ is defined as

$$\begin{aligned} f_Y(y[n] \mid \mathscr{F}[n-1]) = {} & \frac{\Gamma(K+1)}{\Gamma(y[n]+1)\Gamma(K-y[n]+1)} \\ & \cdot\frac{\Gamma(\varphi)\Gamma(y[n]+\mu[n]\varphi)\Gamma(K-y[n]+(1-\mu[n])\varphi)}{\Gamma(K+\varphi)\Gamma(\mu[n]\varphi)\Gamma((1-\mu[n])\varphi)}. \end{aligned}$$

The conditional mean and conditional variance of $Y[n]$, given $\mathscr{F}[n-1]$, are respectively given by

$$\begin{aligned}\mathrm{E}(Y[n] \mid \mathscr{F}[n-1]) &= \mu[n]K,\\ \mathrm{Var}(Y[n] \mid \mathscr{F}[n-1]) &= (\mu[n]-\mu[n]^2)K\frac{K+\varphi}{1+\varphi}.\end{aligned}$$

The beta binomial probability function is very flexible, as shown in Figure 18. For small values of φ, the beta binomial distribution can present decreasing, increasing-decreasing, and upside-down bathtub shapes. On the other hand, for large values of φ, the beta binomial distribution accommodates more symmetric distributions, i.e., data around the mean.

Similar to generalized linear models (GLM) [73], the generalized autoregressive moving average (GARMA) models [190], and the beta autoregressive moving average (βARMA) model [191, 192], the BBARMA model relates the mean $\mu[n]$ to one linear predictor, $\eta[n]$ [73]. This relation is based on a strictly monotonic and twice differentiable function, $g(\cdot)$, called link function [73]. Popular choices for the link functions are the logit [193], the probit [73], and the complementary log-log [190, 191, 193], defined, respectively, as [32]

$$\begin{aligned}\text{Logit: } g(\mu[n]) &= \log\left(\frac{\mu[n]}{1-\mu[n]}\right),\\ \text{Probit: } g(\mu[n]) &= \Phi^{-1}(\mu[n]),\\ \text{Complementary log-log: } g(\mu[n]) &= \log\{-\log(1-\mu[n])\},\end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Following the GLM mathematical formalism [73], we have that

$$g(\mu[n]) = \eta[n] = \mathbf{x}^\top[n]\boldsymbol{\beta},$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_l)^\top$ is the set of unknown parameters and $\mathbf{x}[n] = (x_1[n], x_2[n], \ldots, x_l[n])^\top$, $n = 1, 2, \ldots, N$, is the vector of the covariates with $l < N$. Therefore, similar to the GARMA and βARMA model, a dynamical general model for $\mu[n]$ is given by [190]

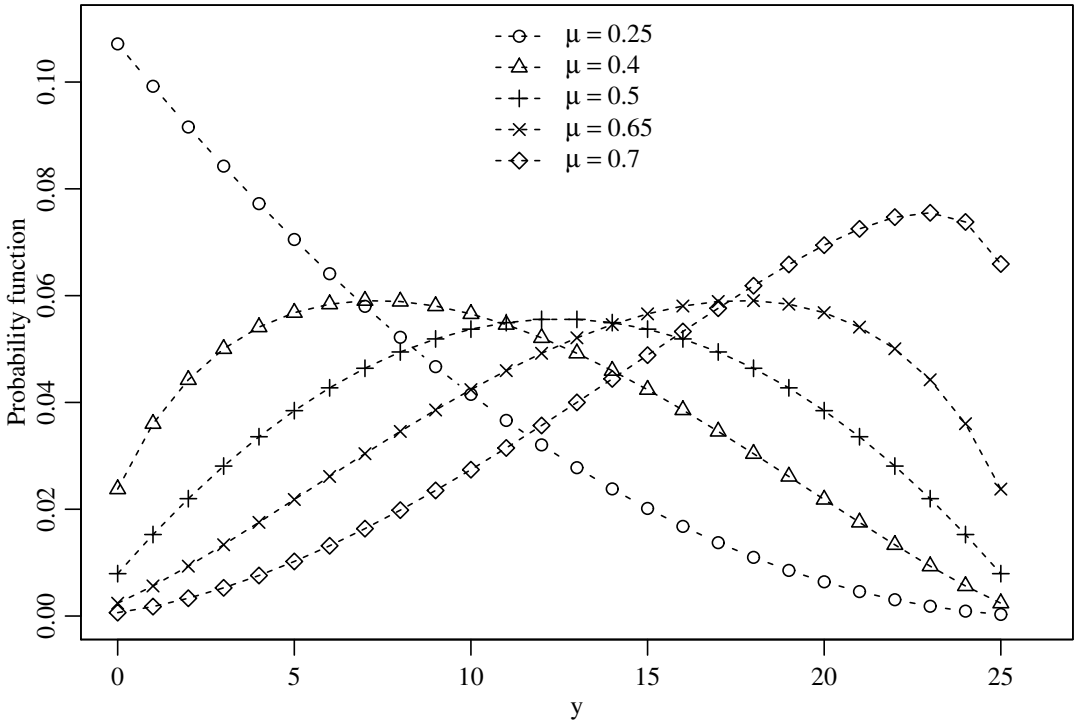
$$g(\mu[n]) = \eta[n] = \mathbf{x}^\top[n]\boldsymbol{\beta} + \boldsymbol{\tau}[n],$$

where

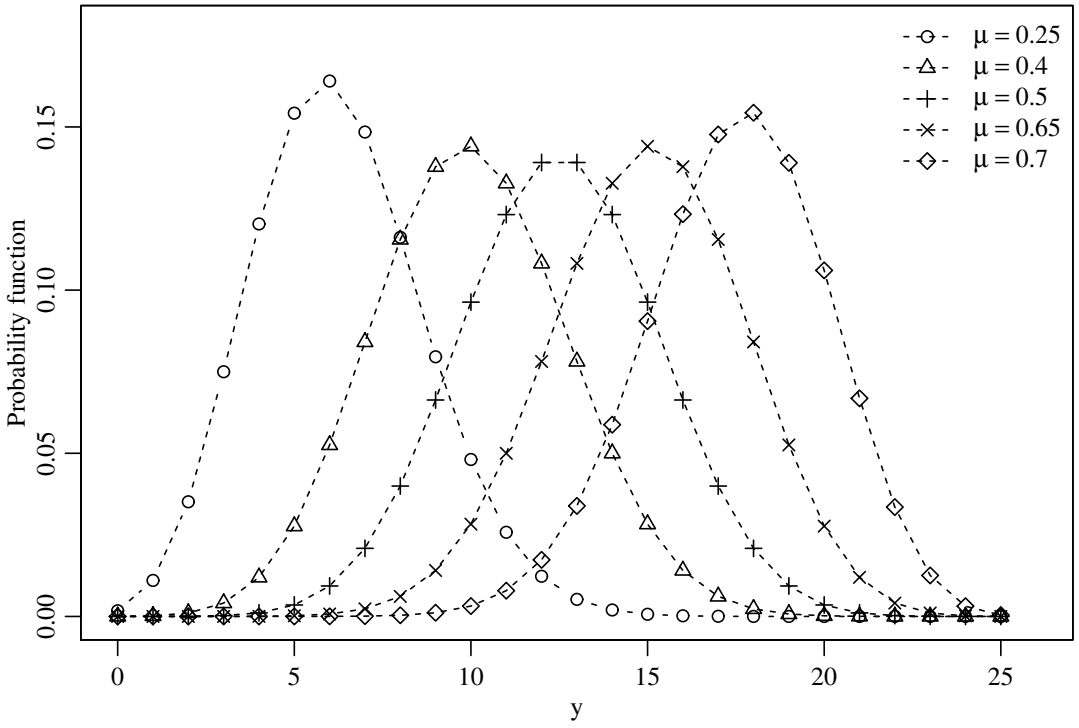
$$\boldsymbol{\tau}[n] = \sum_{i=1}^{p} \phi_i \mathscr{A}(y[n-i]) + \sum_{j=1}^{q} \theta_j \mathscr{M}(y[n-j], \mu[n-j]), \tag{5.1}$$

where p and q are the orders of the model; and the autoregressive and moving average terms are represented by the functions $\mathscr{A}(\cdot)$ and $\mathscr{M}(\cdot)$, respectively. The quantities ϕ_i, $i = 1, 2, \ldots, p$,

Figure 18 – Beta binomial probability functions for different values of μ and (a) $\varphi = 4$, (b) $\varphi = 100$.



(a) $\varphi = 4$



(b) $\varphi = 100$

Source: Author (2020)

and θ_j, $j = 1,2,\ldots,q$, are the autoregressive and moving average parameters, respectively. By including an intercept $\zeta \in \mathbb{R}$, we extend the model described in (5.1), yielding

$$g(\mu[n]) = \eta[n] = \zeta + \mathbf{x}^\top[n]\boldsymbol{\beta} + \sum_{i=1}^{p} \phi_i y^\star[n-i] + \sum_{j=1}^{q} \theta_j \{y^\star[n-j] - \mu[n-j]\}, \tag{5.2}$$

where $y^\star[n] = y[n]/K$ is the scaled observed signal to ensure its values at the same range of $\mu[n]$.

5.2.2 Conditional Likelihood Estimation

The estimation of BBARMA(p,q) model parameters can be realized by maximizing the logarithm of the conditional likelihood function [135, 190]. Let the vector parameters be $\boldsymbol{\gamma} = (\zeta, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top, \varphi)^\top$, with $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_l)^\top$, $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)^\top$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_q)^\top$. The log-likelihood function for the parameter vector $\boldsymbol{\gamma}$ conditional to the $m = \max(p,q)$ preliminary observations is given by

$$\ell = \sum_{n=m+1}^{N} \log f_Y(y[n] \mid \mathscr{F}[n-1]) = \sum_{n=m+1}^{N} \ell[n](\mu[n], \varphi), \tag{5.3}$$

where

$$\begin{aligned}\ell[n](\mu[n],\varphi) =& \log\Gamma(K+1) - \log\Gamma(y[n]+1) + \log\Gamma(\varphi) - \log\Gamma(K-y[n]+1) - \log\Gamma(\mu[n]\varphi) \\ &+ \log\Gamma(y[n]+\mu[n]\varphi) - \log\Gamma(K+\varphi) \\ &+ \log\Gamma(K-y[n]+(1-\mu[n])\varphi) - \log\Gamma((1-\mu[n])\varphi) + \log\Gamma(\varphi).\end{aligned}$$

The conditional maximum likelihood estimators (CMLE), $\widehat{\boldsymbol{\gamma}}$, can be obtained from the score vector $\mathbf{U}(\boldsymbol{\gamma})$ by solving

$$\mathbf{U}(\boldsymbol{\gamma}) = \frac{\partial \ell}{\partial \boldsymbol{\gamma}^\top} = \left(\frac{\partial \ell}{\partial \zeta}, \frac{\partial \ell}{\partial \boldsymbol{\beta}^\top}, \frac{\partial \ell}{\partial \boldsymbol{\phi}^\top}, \frac{\partial \ell}{\partial \boldsymbol{\theta}^\top}, \frac{\partial \ell}{\partial \varphi} \right)^\top = \mathbf{0}, \tag{5.4}$$

where $\mathbf{0}$ is a vector of zeros with dimension $p+q+l+2$. By the chain rule, for $\gamma_j \neq \varphi$, $j = 1,2,\ldots,p+q+l+1$, we have

$$\frac{\partial \ell}{\partial \gamma_j} = \sum_{n=m+1}^{N} \frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \gamma_j}.$$

Note that

$$\frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]} = \varphi \Upsilon[n],$$

where

$$\Upsilon[n] = \varphi\left[\psi(y[n]+\mu[n]\varphi) - \psi(K-y[n]+(1-\mu[n])\varphi) - \psi(\mu[n]\varphi) + \psi((1-\mu[n])\varphi)\right],$$

and

$$\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} = \frac{1}{g'(\mu[n])},$$

where $\psi(\cdot)$ is the digamma function, i.e., $\psi(z) = \frac{\mathrm{d}\log\Gamma(z)}{\mathrm{d}\,z}$, for $z > 0$, and $g'(\cdot)$ is the first derivative of the link function $g(\cdot)$. In particular, for the logit link function, $g(\mu[n]) = \log\left(\frac{\mu[n]}{1-\mu[n]}\right)$, we have $g'(\mu[n]) = (\mu[n](1-\mu[n]))^{-1}$.

Additionally, similar to the βARMA model [192], we have

$$\begin{aligned}
\frac{\partial\,\eta[n]}{\partial\,\zeta} &= 1 - \sum_{s=1}^{q} \theta_s \frac{1}{g'(\mu[n-s])} \frac{\partial\,\eta[n-s]}{\partial\,\zeta}, \\
\frac{\partial\,\eta[n]}{\partial\,\beta_k} &= \mathbf{x}^{\top}[n] - \sum_{s=1}^{q} \theta_s \frac{1}{g'(\mu[n-s])} \frac{\partial\,\eta[n-s]}{\partial\,\beta_k}, \\
\frac{\partial\,\eta[n]}{\partial\,\phi_i} &= y^{\star}[n-i] - \sum_{s=1}^{q} \theta_s \frac{1}{g'(\mu[n-s])} \frac{\partial\,\eta[n-s]}{\partial\,\phi_i}, \\
\frac{\partial\,\eta[n]}{\partial\,\theta_j} &= y^{\star}[n-j] - \mu[n-j] - \sum_{s=1}^{q} \theta_s \frac{1}{g'(\mu[n-s])} \frac{\partial\,\eta[n-s]}{\partial\,\theta_j},
\end{aligned}$$

for $k = 1,2,\ldots,l$, $i = 1,2,\ldots,p$, and $j = 1,2,\ldots,q$.

The score function with respect to φ is given by

$$\begin{aligned}
\frac{\partial\,\ell}{\partial\,\varphi} = \sum_{n=m+1}^{N} \mu[n]\left[\psi\left(y[n] + \mu[n]\varphi\right) - \psi\left(\mu[n]\varphi\right)\right] + (1-\mu[n])\left[\psi\left(K - y[n] + (1-\mu[n])\varphi\right)\right. \\
\left. - \psi\left((1-\mu[n])\varphi\right)\right] + \psi(\varphi) - \psi(K+\varphi).
\end{aligned}$$

The solution of (5.4) has no closed-form, thus, nonlinear optimization algorithms, such as Newton or quasi-Newton [136], are necessary. We selected the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [76] with analytic first derivatives, due to its superior performance for non-linear optimization method [77], requiring only the first derivatives [76]. The initial values for the constant (ζ), the autoregressive ($\boldsymbol{\phi}$) parameters, and the regressors ($\boldsymbol{\beta}$) were derived from the ordinary least squares estimate associated to the linear regression. The response vector is $(y^{\star}[m+1], y^{\star}[m+2], \ldots, y^{\star}[N])^{\top}$ and the covariate matrix is given by

$$\begin{bmatrix}
1 & x_1[m] & x_2[m] & \ldots & x_l[m] & y^{*}[m] & y^{*}[m-1] & \cdots & y^{*}[m-p+1] \\
1 & x_1[m+1] & x_2[m+1] & \ldots & x_l[m+1] & y^{*}[m+1] & y^{*}[m] & \cdots & y^{*}[m-p+2] \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_1[N] & x_2[N] & \ldots & x_l[N] & y^{*}[N-1] & y^{*}[N-2] & \cdots & y^{*}[N-p]
\end{bmatrix}.$$

For the initial values, we adopted $\boldsymbol{\theta} = \mathbf{0}$, as suggested in [134, 194] and $\varphi = 1$, following [195]. Diagnostic measures and forecasting tools are introduced in Appendix E.

5.3 DIGITAL SIGNAL DETECTION THEORY

The goal of this section is to introduce a detector based on a binary hypothesis test tailored for the BBARMA model. For such, we derive the conditional observed information matrix and the asymptotic properties of the conditional parameter estimators of the BBARMA model.

5.3.1 Conditional Observed Information Matrix

In a full exponential family model, the expected and observed information matrices are asymptotically identical [75]. However, because the beta binomial distribution is not in the exponential family, the expected and observed information matrix differ, the latter being preferable for hypothesis testing when $N \to \infty$ [75, 196].

The conditional observed information matrix is given by the negative of the second order derivatives of the conditional log-likelihood, which is given by

$$\begin{aligned}\frac{\partial^2 \ell}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\delta}} &= \sum_{n=m+1}^{N} \frac{\partial}{\partial \boldsymbol{\lambda}} \left(\frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \right) \\ &= \sum_{n=m+1}^{N} \left[\frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \mu[n]^2} \left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \right)^2 \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} + \frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]} \frac{\mathrm{d}^2\mu[n]}{\mathrm{d}\eta[n]^2} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} \right. \\ &\quad \left. + \frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial^2 \eta[n]}{\partial \boldsymbol{\delta} \partial \boldsymbol{\lambda}} \right],\end{aligned} \tag{5.5}$$

where $\boldsymbol{\lambda} = (\zeta, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top)^\top$ and $\boldsymbol{\delta} = (\zeta, \boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \boldsymbol{\theta}^\top)^\top$. The derivatives $\frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n]}$, $\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]}$, $\frac{\partial \eta[n]}{\partial \boldsymbol{\delta}}$, and $\frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}}$ are given in Section 5.2.2. The second order derivative of $\ell[n](\mu[n],\varphi)$ with respect to $\mu[n]$ is given by

$$\begin{aligned}\frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \mu[n]^2} =& \varphi^2 \big\{ \psi'(y[n] + \mu[n]\varphi) + \psi'(K - y[n] + (1-\mu[n])\varphi) - \psi'((1-\mu[n])\varphi) \\ & - \psi'(\mu[n]\varphi) \big\},\end{aligned}$$

where $\psi'(\cdot)$ is the first derivative of the digamma function, i.e., the trigamma function [189]. Note that

$$\frac{\mathrm{d}^2\mu[n]}{\mathrm{d}\eta[n]^2} = -\frac{g''(\mu[n])}{(g'(\mu[n]))^2}.$$

Additionally, the second order derivative of $\frac{\partial \eta[n]}{\partial \boldsymbol{\delta}}$ with respect to ζ, ϕ, and β, for $\delta_i \neq \theta_j$, where $i = 1,2,\ldots,(p+l+1)$ and $j = 1,2,\ldots,q$, is equal to zero. On the other hand,

$$\begin{aligned}
\frac{\partial^2 \eta[n]}{\partial \boldsymbol{\delta} \partial \theta_j} =& -\frac{d\mu[n-j]}{d\eta[n-j]}\frac{\partial \eta[n-j]}{\partial \boldsymbol{\delta}} - \sum_{s=1}^{q} \theta_s \frac{\partial \eta[n-s]}{\partial \boldsymbol{\delta}} \frac{d^2\mu[n-s]}{d\eta[n-s]^2}\frac{\partial \eta[n-s]}{\partial \theta_j} \\
& - \sum_{s=1}^{q} \theta_s \frac{d\mu[n-s]}{d\eta[n-s]}\frac{\partial^2 \eta[n-s]}{\partial \boldsymbol{\delta} \partial \theta_j}, \\
\frac{\partial^2 \eta[n]}{\partial \theta_k \partial \theta_j} =& -\frac{d\mu[n-j]}{d\eta[n-j]}\frac{\partial \eta[n-j]}{\partial \theta_k} - \frac{d\mu[n-k]}{d\eta[n-k]}\frac{\partial \eta[n-k]}{\partial \theta_j} \\
& - \sum_{s=1}^{q} \theta_s \frac{\partial \eta[n-s]}{\partial \theta_j} \frac{d^2\mu[n-s]}{d\eta[n-s]^2}\frac{\partial \eta[n-s]}{\partial \theta_k} - \sum_{s=1}^{q} \theta_s \frac{d\mu[n-s]}{d\eta[n-s]}\frac{\partial^2 \eta[n-s]}{\partial \theta_k \partial \theta_j},
\end{aligned} \tag{5.6}$$

for $j,k = 1,2,\ldots,q$. Appendix F provides mathematical details on the score vector, (5.5), and (5.6).

The derivative of $\frac{\partial \ell}{\partial \boldsymbol{\delta}}$ with respect to φ is given by

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\delta} \partial \varphi} = \sum_{n=m+1}^{N} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{1}{g'(\mu[n])} \frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n] \partial \varphi},$$

where

$$\begin{aligned}
\frac{\partial \ell[n](\mu[n],\varphi)}{\partial \mu[n] \partial \varphi} =& \varphi \left\{ (1-\mu[n]) \left[\psi'\left((1-\mu[n])\varphi\right) - \psi'\left(K - y[n] + (1-\mu[n])\varphi\right) \right] \right. \\
& \left. + \mu[n] \left[\psi'\left(y[n] + \mu[n]\varphi\right) - \psi'\left(\mu[n]\varphi\right) \right] \right\} + \Upsilon[n].
\end{aligned}$$

The second order derivative of $\ell[n](\mu[n],\varphi)$ with respect to φ is given by

$$\begin{aligned}
\frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \varphi^2} =& \psi'(\varphi) + \mu[n]^2 \left[\psi'(y[n] + \mu[n]\varphi) - \psi'(\mu[n]\varphi) \right] \\
& + (1-\mu[n])^2 \left[\psi'(K - y[n] + (1-\mu[n])\varphi) - \psi'((1-\mu[n])\varphi) \right].
\end{aligned}$$

To facilitate the presentation of the conditional observed information matrix, we introduce the following auxiliary vectors and matrices. Let $\mathbf{1}$ be the $(N-m)\times 1$ vector of ones, $\mathbf{T} = \text{diag}\{1/g'(\mu[m+1]), 1/g'(\mu[m+2]), \ldots, 1/g'(\mu[N])\}$, $\mathbf{a} = \left(\frac{\partial \eta[m+1]}{\partial \zeta}, \frac{\partial \eta[m+2]}{\partial \zeta}, \ldots, \frac{\partial \eta[N]}{\partial \zeta} \right)^\top$, and $\boldsymbol{\Upsilon} = (\Upsilon[m+1], \Upsilon[m+2], \ldots, \Upsilon[N])^\top$, $\Upsilon^*[n] = \frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \mu[n]^2}$, $\Upsilon_\varphi[n] = \frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \mu[n] \partial \varphi}$, $\Phi^*[n] = \frac{\partial^2 \ell[n](\mu[n],\varphi)}{\partial \varphi^2}$, $\kappa[n] = \frac{g''(\mu[n])}{(g'(\mu[n]))^2}$, $\xi[n] = \varphi\kappa[n]\Upsilon[n] - \Upsilon^*[n]\mathbf{T}^2$,

$$\mathbf{A} = \left(\frac{\partial^2 \eta[m+1]}{\partial \zeta \partial \boldsymbol{\theta}}, \frac{\partial^2 \eta[m+2]}{\partial \zeta \partial \boldsymbol{\theta}}, \ldots, \frac{\partial^2 \eta[N]}{\partial \zeta \partial \boldsymbol{\theta}} \right)^\top,$$

$$\mathscr{W} = (\varphi\Upsilon[m+1]\mathbf{T}, \varphi\Upsilon[m+2]\mathbf{T}, \ldots, \varphi\Upsilon[N]\mathbf{T}),$$

$$\mathbf{L} = \text{diag}\left\{\Phi^*[m+1], \Phi^*[m+2], \ldots, \Phi^*[N]\right\},$$

$$\mathbf{D} = \text{diag}\left\{\Upsilon_\varphi[m+1], \Upsilon_\varphi[m+2], \ldots, \Upsilon_\varphi[N]\right\},$$

$$\mathbf{W} = \text{diag}\left\{\xi[m+1], \xi[m+2], \ldots, \xi[N]\right\}.$$

Additionally, $\mathbf{M}[i,j]=\frac{\partial \eta[i+m]}{\partial \beta_j}$, $\mathbf{P}[i,j]=\frac{\partial \eta[i+m]}{\partial \phi_j}$, $\mathbf{R}[i,j]=\frac{\partial \eta[i+m]}{\partial \theta_j}$, $\mathscr{M}[i,j]=\frac{\partial^2 \eta[i+m]}{\partial \beta_i \partial \theta_j}$, $\mathscr{P}[i,j]=\frac{\partial^2 \eta[i+m]}{\partial \phi_i \partial \theta_j}$, and $\mathscr{R}[i,j]=\frac{\partial^2 \eta[i+m]}{\partial \theta_i \partial \theta_j}$. The matrices $\mathbf{M}[\cdot,\cdot]$, $\mathbf{P}[\cdot,\cdot]$, $\mathbf{R}[\cdot,\cdot]$, $\mathscr{M}[\cdot,\cdot]$, $\mathscr{P}[\cdot,\cdot]$, and $\mathscr{R}[\cdot,\cdot]$ are of dimensions $(N-m)\times l$, $(N-m)\times p$, $(N-m)\times q$, $(N-m)\times l$, $(N-m)\times p$, and $(N-m)\times q$, respectively. Based on (5.5) and on the above expressions, the conditional observed information matrix is given by

$$\mathbf{I}(\boldsymbol{\gamma})=\begin{bmatrix} I_{(\zeta,\zeta)} & \mathbf{I}_{(\zeta,\beta)} & \mathbf{I}_{(\zeta,\phi)} & \mathbf{I}_{(\zeta,\theta)} & \mathbf{I}_{(\zeta,\varphi)} \\ \mathbf{I}_{(\beta,\zeta)} & \mathbf{I}_{(\beta,\beta)} & \mathbf{I}_{(\beta,\phi)} & \mathbf{I}_{(\beta,\theta)} & \mathbf{I}_{(\beta,\varphi)} \\ \mathbf{I}_{(\phi,\zeta)} & \mathbf{I}_{(\phi,\beta)} & \mathbf{I}_{(\phi,\phi)} & \mathbf{I}_{(\phi,\theta)} & \mathbf{I}_{(\phi,\varphi)} \\ \mathbf{I}_{(\theta,\zeta)} & \mathbf{I}_{(\theta,\beta)} & \mathbf{I}_{(\theta,\phi)} & \mathbf{I}_{(\theta,\theta)} & \mathbf{I}_{(\theta,\varphi)} \\ \mathbf{I}_{(\varphi,\zeta)} & \mathbf{I}_{(\varphi,\beta)} & \mathbf{I}_{(\varphi,\phi)} & \mathbf{I}_{(\varphi,\theta)} & I_{(\varphi,\varphi)} \end{bmatrix},$$

where $I_{(\zeta,\zeta)}=\mathbf{a}^\top\mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\zeta,\beta)}=\mathbf{I}_{(\beta,\zeta)}^\top=\mathbf{M}^\top\mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\zeta,\phi)}=\mathbf{I}_{(\phi,\zeta)}^\top=\mathbf{P}^\top\mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\zeta,\theta)}=\mathbf{I}_{(\theta,\zeta)}^\top=\mathbf{R}^\top\mathbf{W}\mathbf{a}-\mathscr{W}\mathbf{A}$, $\mathbf{I}_{(\zeta,\varphi)}=\mathbf{I}_{(\varphi,\zeta)}^\top=-\mathbf{a}^\top\mathbf{T}\mathbf{D}\mathbf{1}$, $\mathbf{I}_{(\beta,\beta)}=\mathbf{M}^\top\mathbf{W}\mathbf{M}$, $\mathbf{I}_{(\beta,\phi)}=\mathbf{I}_{(\phi,\beta)}^\top=\mathbf{P}^\top\mathbf{W}\mathbf{M}$, $\mathbf{I}_{(\beta,\theta)}=\mathbf{I}_{(\theta,\beta)}^\top=\mathbf{R}^\top\mathbf{W}\mathbf{M}-\mathscr{W}\mathscr{M}$, $\mathbf{I}_{(\beta,\varphi)}=\mathbf{I}_{(\varphi,\beta)}^\top=-\mathbf{M}^\top\mathbf{T}\mathbf{D}\mathbf{1}$, $\mathbf{I}_{(\phi,\phi)}=\mathbf{P}^\top\mathbf{W}\mathbf{P}$, $\mathbf{I}_{(\phi,\theta)}=\mathbf{I}_{(\theta,\phi)}^\top=\mathbf{R}^\top\mathbf{W}\mathbf{P}-\mathscr{W}\mathscr{P}$, $\mathbf{I}_{(\phi,\varphi)}=\mathbf{I}_{(\varphi,\phi)}^\top=-\mathbf{P}^\top\mathbf{T}\mathbf{D}\mathbf{1}$, $\mathbf{I}_{(\theta,\theta)}=\mathbf{R}^\top\mathbf{W}\mathbf{R}-\mathscr{W}\mathscr{R}$, $\mathbf{I}_{(\theta,\varphi)}=\mathbf{I}_{(\varphi,\theta)}^\top=-\mathbf{R}^\top\mathbf{T}\mathbf{D}\mathbf{1}$, and $I_{(\varphi,\varphi)}=-\operatorname{tr}(\mathbf{L})$, where $\operatorname{tr}(\cdot)$ is the trace function.

Based on the consistency of the CMLE and on the asymptotic distribution of $\widehat{\boldsymbol{\gamma}}$ $(N\to\infty)$, we have that [75, 82]

$$\widehat{\boldsymbol{\gamma}}\overset{a}{\sim}\mathscr{N}_u\left(\boldsymbol{\gamma},\mathbf{I}^{-1}(\boldsymbol{\gamma})\right),$$

where $\overset{a}{\sim}$ denotes approximately distributed, $u=p+q+l+2$, $\mathbf{I}(\boldsymbol{\gamma})$ is the Fisher information matrix and $\mathscr{N}_u$ is the u-dimensional Gaussian distribution with mean $\boldsymbol{\gamma}$ and covariance matrix $\mathbf{I}^{-1}(\gamma)$.

5.3.2 Hypothesis Test

Let the parameter vector $\boldsymbol{\gamma}$ be partitioned in a parameter vector of interest $\boldsymbol{\gamma}_I$, of dimension ν, and a vector of nuisance parameters, $\boldsymbol{\gamma}_J$, of dimension $r-\nu$, $r=1,2,\ldots,p+q+l+2$ [4]. In addition, $\mathscr{H}_0:\boldsymbol{\gamma}_I=\boldsymbol{\gamma}_{I_0}$ is the hypothesis of interest and $\mathscr{H}_1:\boldsymbol{\gamma}_I\neq\boldsymbol{\gamma}_{I_0}$ the alternative hypothesis, where $\boldsymbol{\gamma}_{I_0}$ is a fixed column vector of dimension ν. The likelihood ratio (T_L), the Wald (T_W), and

the Rao (T_R) test statistics can be written as [4]

$$
\begin{aligned}
T_L &= 2\left[\ell(\widehat{\boldsymbol{\gamma}}) - \ell(\tilde{\boldsymbol{\gamma}})\right], \\
T_W &= (\widehat{\boldsymbol{\gamma}}_{I1} - \boldsymbol{\gamma}_{I0})^{\top} \left(\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\gamma}}_1)\right]_{\gamma_I \gamma_I}\right)^{-1} (\widehat{\boldsymbol{\gamma}}_{I1} - \boldsymbol{\gamma}_{I0}), \\
T_R &= \left(\frac{\partial \ell}{\partial \boldsymbol{\gamma}_I}\Big|_{\boldsymbol{\gamma}=\tilde{\boldsymbol{\gamma}}}\right)^{\top} \left(\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\gamma}})\right]_{\gamma_I \gamma_I}\right)^{-1} \left(\frac{\partial \ell}{\partial \boldsymbol{\gamma}_I}\Big|_{\boldsymbol{\gamma}=\tilde{\boldsymbol{\gamma}}}\right),
\end{aligned}
$$

where $\widehat{\boldsymbol{\gamma}}_{I1}$ and $\widehat{\boldsymbol{\gamma}}_1 = (\widehat{\boldsymbol{\gamma}}_{I1}^{\top}, \widehat{\boldsymbol{\gamma}}_{J1}^{\top})^{\top}$ are the MLEs under $\mathscr{H}_1$ (unrestricted MLEs); $\boldsymbol{\gamma}_{J0}$ is a fixed column vector of dimension $r-\nu$; $\left[\mathbf{I}^{-1}(\widehat{\boldsymbol{\gamma}})\right]_{\gamma_I \gamma_I}$ is a partition of $\mathbf{I}(\widehat{\boldsymbol{\gamma}})$ limited to the estimates of interest; $\tilde{\boldsymbol{\gamma}} = (\widehat{\boldsymbol{\gamma}}_{I0}^{\top}, \widehat{\boldsymbol{\gamma}}_{J0}^{\top})^{\top}$ is the MLE under $\mathscr{H}_0$; $\ell(\widehat{\boldsymbol{\gamma}})$ is the maximized log-likelihood of the fitted model; $\ell(\tilde{\boldsymbol{\gamma}})$ is the maximized log-likelihood of the model under $\mathscr{H}_0$; and $\left[\mathbf{I}^{-1}(\tilde{\boldsymbol{\gamma}})\right]_{\gamma_I \gamma_I}$ is a partition of $\mathbf{I}(\tilde{\boldsymbol{\gamma}})$ limited to the estimates of interest. As suggested in [83] and due to the convenience of the Wald test, since it requires only one estimation under alternative hypothesis, we selected the Wald test to perform the hypothesis test on the parameters.

Under $\mathscr{H}_0$, the test statistic, T_W, has asymptotically chi-squared distribution with ν degrees of freedom, χ^2_ν. Thus, the proposed detector consists of comparing the computed value of T_W with a threshold value ε. The threshold value is obtained from the χ^2_ν distribution and the desired probability of false alarm [4].

To illustrate the above approach, we consider the problem of detecting a signal $s[n]$ embedded in noise from a measured signal $y[n]$. For such, we have the following BBARMA model:

$$
g(\mu[n]) = \zeta + s[n]\beta_1 + \sum_{i=1}^{p} \phi_i y^{\star}[n-i] + \sum_{j=1}^{q} \theta_j \{y^{\star}[n-j] - \mu[n-j]\}, \tag{5.7}
$$

where β_1 is the unknown amplitude of the signal [4]. To detect whether $s[n]$ is present, we have the following hypotheses:

$$
\begin{cases}
\mathscr{H}_0 : \beta_1 = 0, \\
\mathscr{H}_1 : \beta_1 \neq 0.
\end{cases} \tag{5.8}
$$

The detector is derived using the Wald test described above. We reject $\mathscr{H}_0$ when $T_W > \varepsilon$ [4]. In this situation, $\beta_1 \neq 0$, indicating the presence of the signal.

5.4 NUMERICAL RESULTS

In this section, we aim at evaluating the CMLE of the parameters of the BBARMA model and assessing the performance of the proposed detector. For such, computational experiments based on Monte Carlo simulations were considered.

5.4.1 Evaluation of the CMLE

Signals $y[n]$ were generated from the beta binomial distribution by the acceptance-rejection method [197] with mean given by (5.2), logit link function, $K = 255$ (8-bit signals), without covariates in the simulations. We considered simulations under two scenarios. Scenario 1 employs $\mu \approx 0.9$ (asymmetric distribution) and Scenario 2 adopts $\mu \approx 0.5$ (almost symmetric distribution). For such, the selected parameters were $\zeta = 1$, $\phi_1 = 1$, $\theta = 0$, and $\varphi = 20$, for Scenario 1, and $\zeta = 0.2$, $\phi_1 = 0.5$, $\theta_1 = 0.3$, and $\varphi = 15$, for Scenario 2. The number of Monte Carlo replications was set to $10,000$ and the signal lengths considered were $N \in \{150; 300; 500\}$. In order to numerically evaluate the point estimators, we computed the mean, bias, and mean square error (MSE) of the CMLE.

For the evaluation of interval estimation, we calculated the coverage rates (CR) of the confidence interval (CI) with confidence $100(1-\alpha)\%$. The CR is derived based on the CMLE asymptotic distribution and it is defined as

$$\left[\widehat{\gamma}_i - z_{1-\alpha/2}\sqrt{\mathbf{I}_{ii}^{-1}(\widehat{\boldsymbol{\gamma}})}; \widehat{\gamma}_i + z_{1-\alpha/2}\sqrt{\mathbf{I}_{ii}^{-1}(\widehat{\boldsymbol{\gamma}})}\right],$$

where γ_i, $i = 1, 2, \ldots, l+p+q+2$, denotes the ith component of $\boldsymbol{\gamma}$, $\mathbf{I}_{ii}^{-1}(\widehat{\boldsymbol{\gamma}})$ is the ith element of the diagonal of $\mathbf{I}^{-1}(\widehat{\boldsymbol{\gamma}})$, α is the significance level, and z_ρ is the ρth quantile of the standard normal distribution. For each Monte Carlo replication, we computed the CI and interrogated whether the CI contains the true parameter or not. The CR is given by the percentage of replications for which the parameter is in the CI. The Monte Carlo simulations are summarized in Algorithm 6.

Algorithm 6: Monte Carlo simulations for evaluation of the BBARMA(p,q) model parameter estimators.

Input: Vector of parameter $\boldsymbol{\gamma}$ and signal length N.
Output: Results of the desirable measures.
1: Suppose that the observed output signal $y[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\gamma}$;
2: Generate $y[n]$ from $f_Y(\boldsymbol{\gamma})$;
3: For each Monte Carlo replication, compute $\widehat{\boldsymbol{\gamma}}$;
4: Repeat steps 2 and 3 a large number R of times, obtaining: $\widehat{\boldsymbol{\gamma}}[1], \widehat{\boldsymbol{\gamma}}[2], \ldots, \widehat{\boldsymbol{\gamma}}[R]$;
5: Use the estimates $\widehat{\boldsymbol{\gamma}}[1], \widehat{\boldsymbol{\gamma}}[2], \ldots, \widehat{\boldsymbol{\gamma}}[R]$, to calculate the desired measures (mean, bias, confidence intervals, and mean square error).

Tables 12 and 13 present the simulation results for Scenarios 1 and 2, respectively. As expected, both bias and MSE figures decrease to zero as N grows. This behavior agrees with

Table 12 – Simulation results on point and interval estimation of the BBARMA$(1,0)$ model, considering a significance level for $\alpha = 10\%$

Measures	$\widehat{\zeta}$	$\widehat{\phi}_1$	$\widehat{\varphi}$
	$N = 150$		
Mean	1.0852	0.9036	20.6377
Bias	0.0852	−0.0964	0.6377
MSE	0.3091	0.4104	7.8154
CR	0.9045	0.9045	0.9015
	$N = 300$		
Mean	1.0451	0.9493	20.3392
Bias	0.0451	−0.0507	0.3392
MSE	0.1538	0.2046	3.5039
CR	0.9011	0.9013	0.9039
	$N = 500$		
Mean	1.0344	0.9611	20.2202
Bias	0.0344	−0.0389	0.2202
MSE	0.0929	0.1235	2.0854
CR	0.8980	0.8964	0.8996

Source: Author (2020)

the asymptotic property (consistency) of the CMLE. The CR values of the BBARMA$(1,0)$ model are close to the nominal value of 0.90, for all considered N. In accordance with the literature [194,198], the BBARMA$(1,1)$ model presents CR values close to 0.90 for larger signal lengths. Convergence failures were absent for all the tested scenarios.

5.4.2 Evaluation of the Proposed Detector

Considering the same simulations parameters as detailed in the previous subsection, we aim at assessing the performance of the introduced detector showed in (5.8). For such, in (5.7), we adopted $p = q = 1$. The interest signal, $s[n]$, was selected as $s[n] = \cos(2\pi f_0 n)$, where f_0 is the signal frequency. This is the classical sinusoidal detection problem which is present in many fields, such as radar, sonar, and communication systems [4].

For each Monte Carlo replication, we fitted the BBARMA model in a simulated digital signal $y[n]$ as follows

$$g(\mu[n]) = \zeta + s[n]\beta_1 + \phi_1 y^{\star}[n-1] + \theta_1\{y^{\star}[n-1] - \mu[n-1]\}.$$

We considered two scenarios to obtain $y[n]$: (i) Scenario I employed $\zeta = 0.2$, $\beta_1 = 0.5$, $\phi_1 = 0.5$, $\theta_1 = 0.3$, $\varphi = 15$, and $f_0 = 0.5$; and (ii) Scenario II set $\zeta = 1$, $\beta_1 = 0.1$, $\phi_1 = 2$, $\theta_1 = 1$, $\varphi =$

Table 13 – Simulation results on point and interval estimation of the BBARMA$(1,1)$ model, considering a significance level for $\alpha = 10\%$

Measures	$\widehat{\zeta}$	$\widehat{\phi}_1$	$\widehat{\theta}_1$	$\widehat{\varphi}$
		$N = 150$		
Mean	0.2987	0.3428	0.4614	15.5923
Bias	0.0987	−0.1572	0.1614	0.5923
MSE	0.9607	2.4550	2.5562	4.1275
CR	0.7601	0.7593	0.7523	0.8949
		$N = 300$		
Mean	0.2629	0.4007	0.3971	15.2820
Bias	0.0629	−0.0993	0.0971	0.2820
MSE	0.6209	1.5866	1.6179	1.7988
CR	0.8105	0.8085	0.8064	0.8998
		$N = 500$		
Mean	0.2393	0.4375	0.3602	15.1728
Bias	0.0393	−0.0625	0.0602	0.1728
MSE	0.4106	1.0514	1.0704	1.0238
CR	0.8405	0.8405	0.8367	0.9031

Source: Author (2020)

50, and $f_0 = 0.7$. Scenarios I and II employed $\mu \approx 0.5$ and $\mu \approx 0.9$, respectively. Additionally, we considered $\beta_1 < 1$, since β_1 values higher than one resulted in similar performance of the evaluated detection methods. The levels of significance were

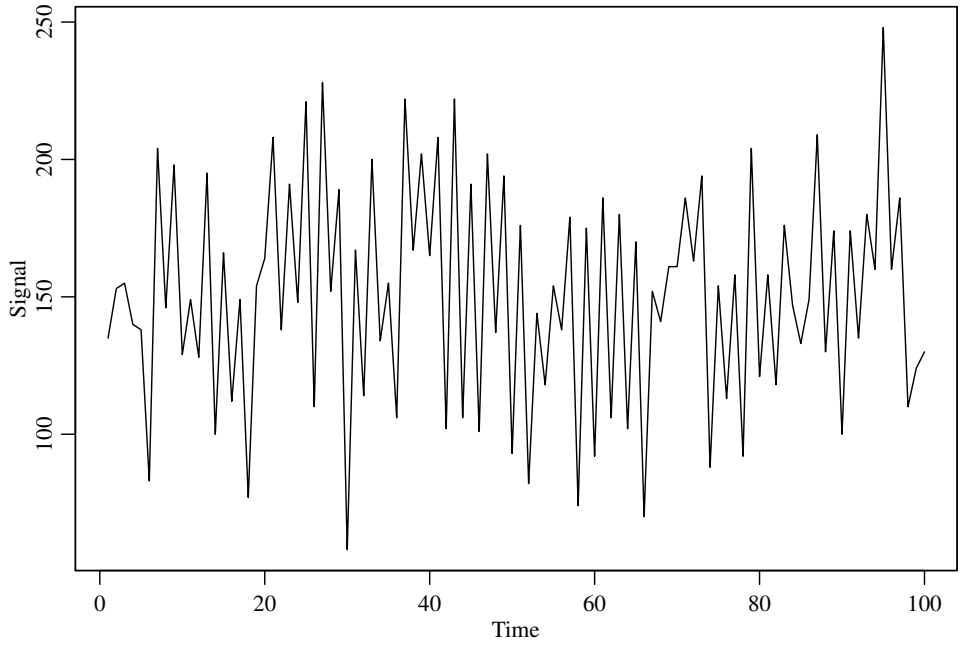
$$\alpha \in \{0; 0.05; 0.1; 0.15; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 1\},$$

the number of Monte Carlo replications equals $5,000$, and $N = 100$. Figure 19 shows a typical realization of the simulated signals, considering Scenarios I and II.

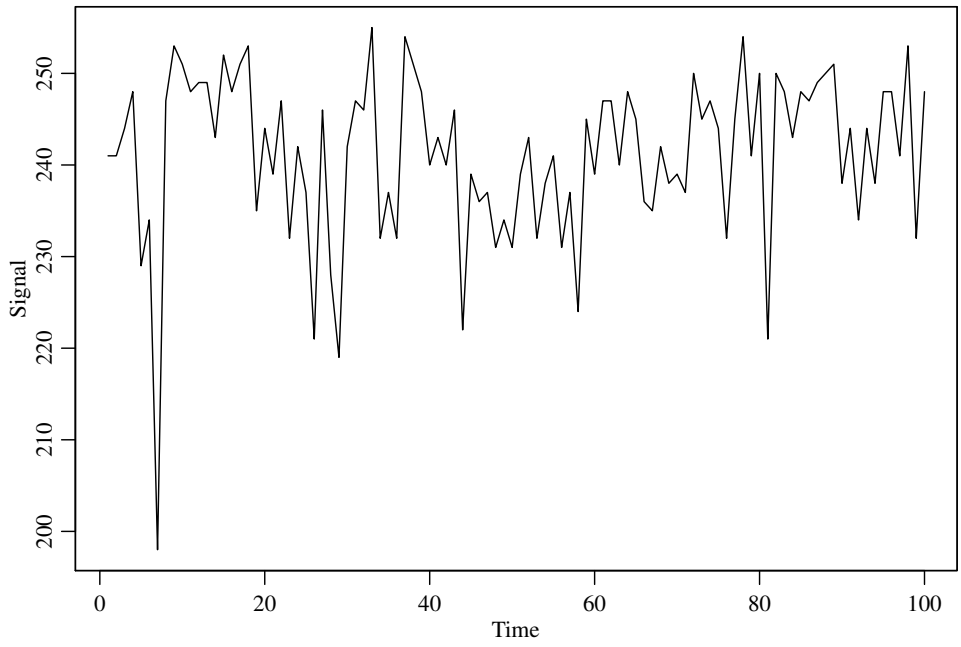
Based on (5.8), a signal is detected when the amplitude $\beta_1 \neq 0$ and the null hypothesis in (5.8) is rejected. To estimate the probability of detection, we computed the proportion of Monte Carlo replications in which the null hypothesis was rejected. We used the empirical size of the hypothesis test, which is computed according to the following steps: (i) for each Monte Carlo replication, generate a signal, $y^*[n]$, without covariates; (ii) fit the BBARMA model for $y^*[n]$, including the parameter β_1; and (iii) compute the percentage of replications that the null hypothesis is rejected.

We compared the proposed detector with the widely popular ARMA- and Gaussian-based detectors [4]. Figures 20 and 21 present the receiver operating characteristic (ROC) curves [199] of the detection results, showing the probability of detection versus the estimated probability of false alarm. In summary, the BBARMA ROC curve was computed following the Algorithm 7.

Figure 19 – Simulated digital signals tested in the proposed detector, considering Scenarios I and II.



(a) Simulated digital signal - Scenario I



(b) Simulated digital signal - Scenario II

Source: Author (2020)

Algorithm 7: Monte Carlo simulations for evaluation of the proposed detector.

Input: Vector of parameter $\boldsymbol{\gamma}$ and the signal length N.
Output: Estimated probability of detection.

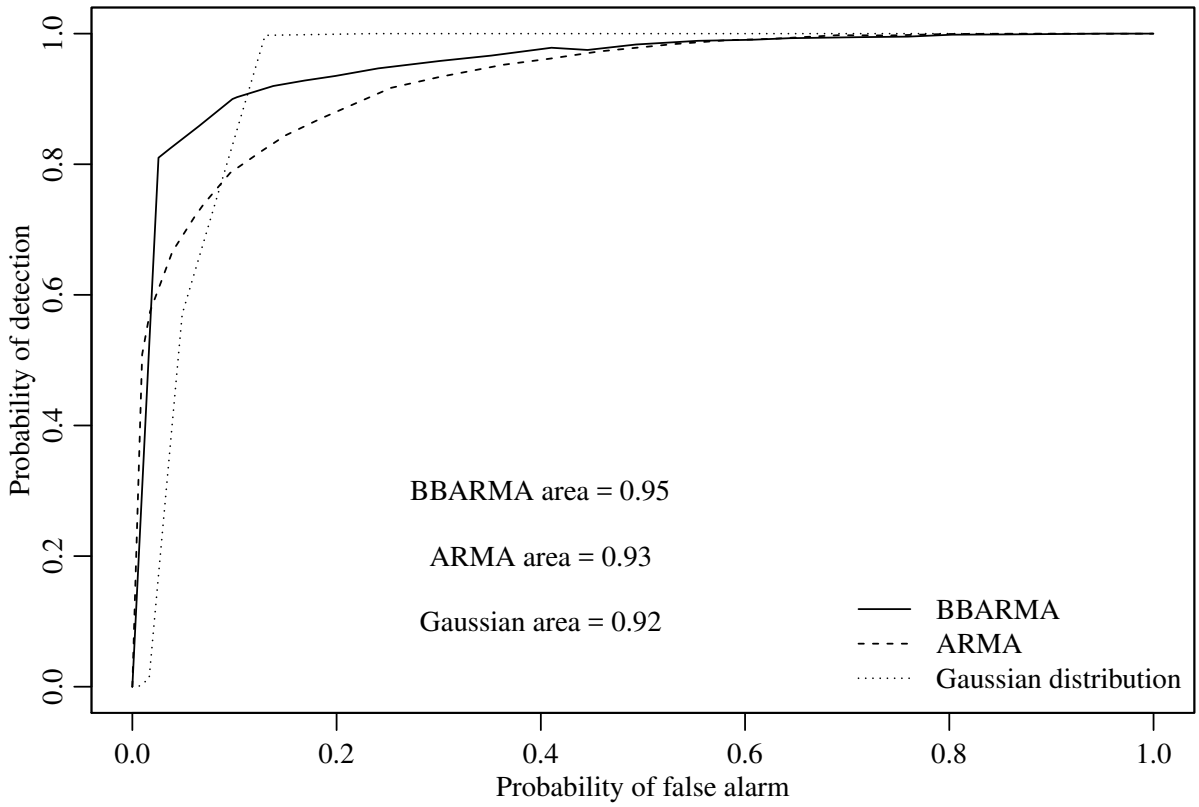
1: Suppose that the observed output signal $y[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\gamma} = (\zeta, \beta_1, \phi_1, \theta_1, \varphi)^\top$;
2: Suppose that the observed output signal $y^*[n]$ follows a distribution f_Y with parametric vector $\boldsymbol{\gamma}^* = (\zeta, \phi_1, \theta_1, \varphi)^\top$;
3: Generate $y[n]$ from $f_Y(\boldsymbol{\gamma})$ and $y^*[n]$ from $f_Y(\boldsymbol{\gamma}^*)$;
4: For each Monte Carlo replication, compute $\widehat{\boldsymbol{\gamma}}$ considering $y[n]$ and $y^*[n]$, obtaining $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$, respectively;
6: Compute the p-values of $\widehat{\beta}_1$ and $\widehat{\beta}_1^*$;
7: Repeat steps 3 to 6 a very large number R of times;
8: Count the percentage of replications that $\widehat{\beta}_1^* \neq 0$, and call this $\widehat{\alpha}$ (empirical size of the hypothesis test);
9: Count the percentage of replications that $\widehat{\beta}_1 \neq 0$, obtaining the probability of detection.

The proposed BBARMA detector outperformed the competing methods in terms of probability of detection and estimated probability of false alarm in both considered scenarios. The area under the ROC curves for the ARMA detector was 2.10% and 9.72% lower when compared with ROC curve of the proposed detector under Scenarios I and II, respectively; for the Gaussian detector, the ROC area values were also smaller, being 1.06% and 36.11% lower, for Scenario I and II, respectively. Additionally, in Scenario I, for a probability of detection equal 0.90, the estimated probabilities of false alarm is about to 0.10 for the BBARMA- and ARMA-based detectors and about 0.30 for the Gaussian detection method; in Scenario II, for a detection probability of 0.90, the estimated probabilities of false alarms are about 0.73, 0.85, and 0.90 for the BBARMA-, ARMA-, and Gaussian-based detectors, respectively.

5.5 CONCLUSION

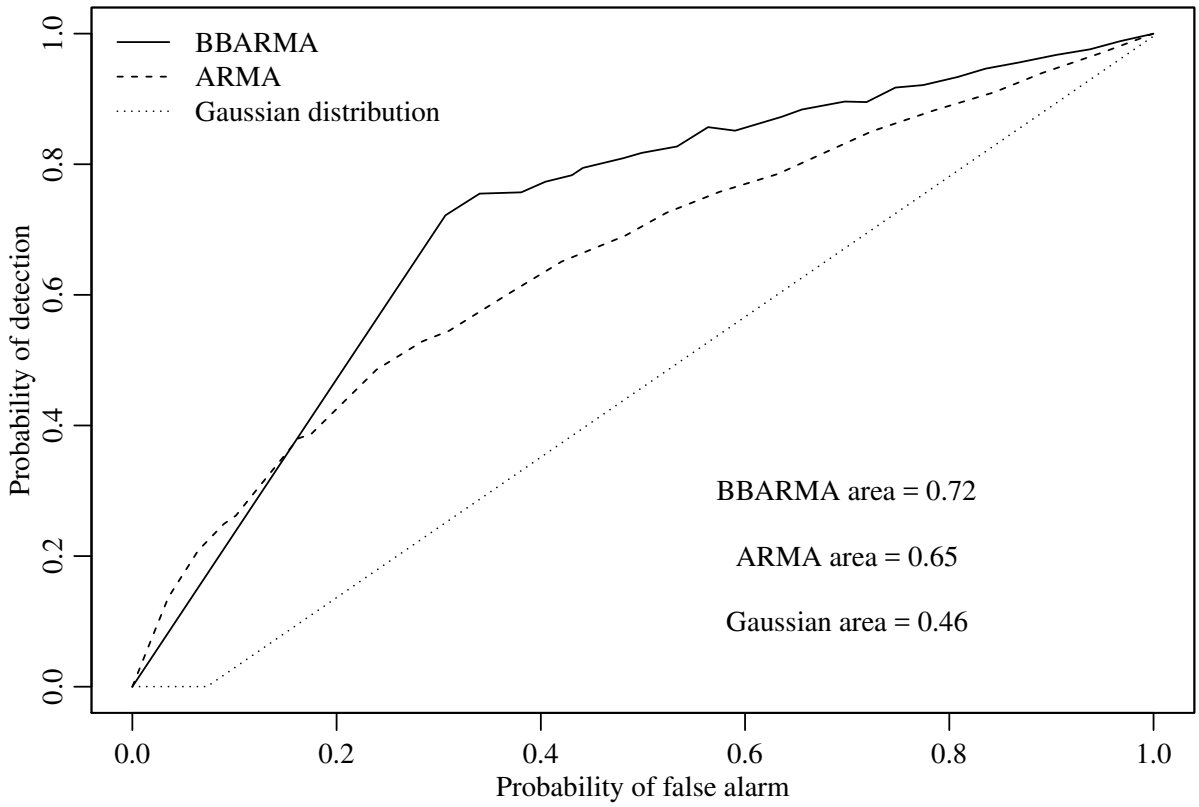
In this chapter, we derived the BBARMA model and a signal detector based on the asymptotic properties of the discussed model parameter estimators. We introduced an inference approach for the model parameters, the conditional observed information matrix, and the asymptotic properties of the CMLE. Monte Carlo simulations were used as a tool to evaluate the performance of the CMLE and of the proposed signal detector, indicating the consistency of the CMLE. The proposed BBARMA detector could outperform the ARMA- and Gaussian-based detectors in the evaluated scenarios. The proposed model is presented as a suitable tool for quantized signal detection.

Figure 20 – ROCs of the detection results for Scenario I, comparing the BBARMA-, ARMA-, and Gaussian-based detectors.



Source: Author (2020)

Figure 21 – ROCs of the detection results for Scenario II, comparing the BBARMA-, ARMA-, and Gaussian-based detectors.



Source: Author (2020)

6 WAVELENGTH-RESOLUTION SAR GROUND SCENE PREDICTION BASED ON IMAGE STACK

Abstract

This chapter presents five different statistical methods for ground scene prediction (GSP) in wavelength-resolution synthetic aperture radar (SAR) images. The predictions are based on image stacks, which are composed of images from the same scene acquired at different instants with the same flight geometry. The considered methods for obtaining the ground scene prediction include (i) autoregressive models; (ii) trimmed mean; (iii) median; (iv) intensity mean; and (v) mean. It is expected that the predicted image presents the true ground scene without change and preserves the ground backscattering pattern. The median method provided the most accurate representation of the true ground. To show the applicability of the GSP, a change detection algorithm was considered using the median ground scene as a reference image. As a result, the median method displayed the probability of detection of 97% and a false alarm rate of $0.11/\text{km}^2$, when considering military vehicles concealed in a forest.

Keywords: CARABAS II, ground scene prediction, image stack, multi-pass, SAR images.

6.1 INTRODUCTION

Common tasks in synthetic aperture radar (SAR) statistical image processing include the identification and classification of distinct ground type [3,15,81,94,200], modeling [90,201–203], and change detection [2, 80, 204, 205]. In special, wavelength-resolution low-frequency SAR systems are useful for natural disasters monitoring, foliage-penetrating applications, and detection of concealed targets [63]

The wavelength-resolution SAR system is usually associated with ultrawideband (UWB) radar signal and ultrawidebeam antenna [206]. With such, the maximum resolution is achieved and it is in the order of radar signal wavelength. Additionally, available UWB SAR systems only work at low frequencies. One essential feature of wavelength-resolution SAR systems is that the speckle noise does not influence the acquired images since it is likely that only a single scatter is present in the resolution cell. Additionally, small scatterers present in the ground area of interest do not contribute to the backscattering for low-frequency radar systems. Thus,

small structures, such as tree branches and leaves, are not shown in SAR images [207]. Because large scatterers are associated to low frequency components, they tend to be less influenced by environmental effects and are stable in time. Hence, by using multi-passes with identical heading and incidence angle of the illuminating platform at a given ground area, an image package with similar statistics can be obtained [1]. In [93], it is discussed clutter statistical models for stacks of very-high-frequency (VHF) wavelength-resolution SAR images. The SAR image stacks are a frequent topic of study for SAR systems with high resolution [208–210]. However, the literature lacks the use of large image stacks for wavelength-resolution SAR for change detection applications.

Recently, a study using a small stack of multi-pass wavelength-resolution SAR images for change detection was introduced in [1]. Usually, SAR change detection algorithms (CDA) are designed for two images (reference and surveillance) and used to detect changes in a ground scene between distinct measurements in time. The changes on the ground scene can be the result of man-made interference, such as facilities and building, or natural disasters, such as floods, wildfire, and deforestation [63–65]. However, an image stack can be considered instead of just two images in a CDA; such a collection of images leads to obtain a ground scene prediction (GSP) and improved detection performance, since more knowledge about the ground clutter can be obtained [1]. This information is used to eliminate clutter and noise in the surveillance image [1], and consequently, enhancing CDA results.

In [211], the autoregressive (AR) model was employed as a preliminary study to obtain a GSP based on a single wavelength-resolution SAR image stack. The resulting predicted image was submitted as input data to a change detection algorithm, based only on subtraction, thresholding, and morphological operations. The CDA in [211] corresponds to the detection analysis step of the CDA used in [84]. Despite its simplicity, the change detection results in [211] were competitive when compared with the ones recently presented in [1, 212].

Multi-pass SAR images cannot be equidistantly observed over time, the noise across the image stack is not related to the time order. As a consequence, the use of a time series model, commonly employed in statistical signal processing [213–216], may not be the most suitable approach to obtain a GSP. Additionally, the backscattering of the images in the stack is stable in time, i.e., a sequence of pixels for each position follows a similar pattern, and changes in such behavior are understood as outliers. Thus, an image filtering considering robust statistical methods, such as trimmed mean and median [217, 218], might be better candidates to obtain a ground scene prediction. These approaches can provide an accurate prediction of the ground

scene, avoid the time order problem, and exclude the pixels that do not follow the sequence pattern. Indeed, the median and the trimmed mean filters are traditionally used to remove impulse noise from an image [34, 219–225].

To the best of our knowledge, the study in [211] is the only work related to the ground scene prediction for wavelength-resolution SAR image stacks. This chapter extends the results presented in [211] with four other statistical methods to predict a ground scene for three SAR image stacks, since statistical methods are commonly employed in SAR image processing [80, 81, 90, 94, 200–205]. The selected statistical methods to obtain the prediction image are (i) autoregressive models; (ii) trimmed mean; (iii) median; (iv) intensity mean; and (v) mean. The predicted ground scene methods are sought to preserve the ground backscattering statistical characteristics of the images in the stack and presents predicted pixel values closer to the original images. It is expected that the predicted images represent the true ground scenes, allowing applications, such as monitoring of forested areas and natural disasters.

To illustrate a possible application of the sought GSP, the median ground scene was used as a reference image in a change detection algorithm based on the detection analysis step of the CDA presented in [84], which was evaluated in terms of target detection probability and false alarm rate. The results reported in [1, 2] were adopted as the reference model for comparison.

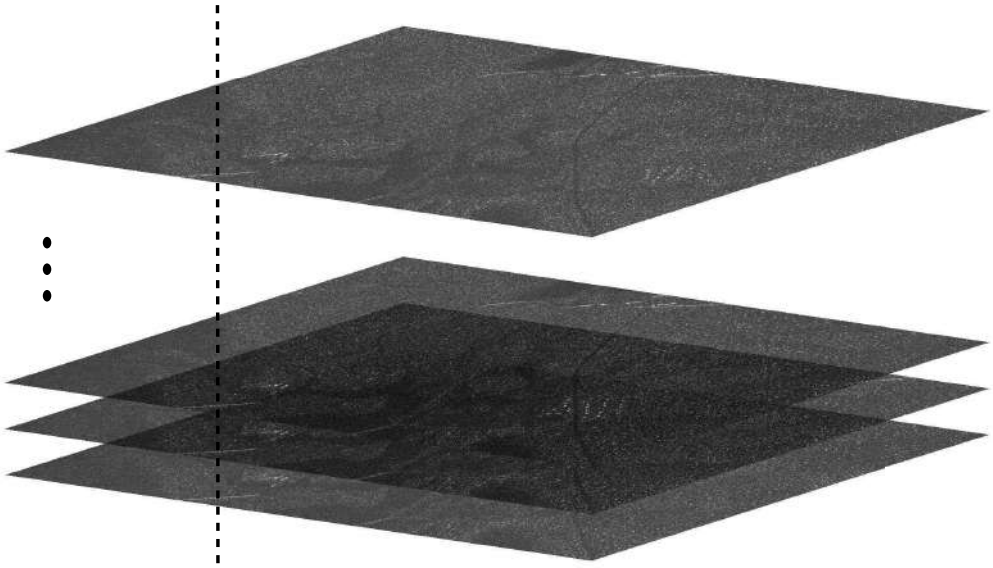
The chapter is organized as follows. In Section 6.2, we describe a suite of selected statistical methods for ground scene prediction. Section 6.3 presents experimental results, including a description of the considered data set and the ground scene prediction results. Then, a change detection method based on the discussed GSP approaches is introduced. Finally, Section 6.4 concludes the chapter.

6.2 GROUND SCENE PREDICTION

As discussed in [93], an image stack is composed of images with similar heading and incidence angle of the same illuminating platform. As a consequence of this similarity, the SAR images in the stack are very similar and stable in time. Thus, a sequence of each pixel position can be extracted from the stack, as illustrated in Figure 22.

The data set considered in this chapter is composed of wavelength-resolution SAR images, i.e., the resolution of the SAR image is in the order of the radar signal wavelength [207]. Therefore, there may only be a single scatter in the resolution cell. As a consequence, the considered images are not affected by speckle noise, which is typically a strong source of noise

Figure 22 – Stack of images to be considered in GSP. The methods should be applied for each pixel position, as evidenced by the vertical line.



Source: Author (2020)

in SAR images in higher frequency bands. Thus, the backscattering from the image stack is stable in time, allowing an accurate GSP.

We consider five statistical methods to obtain the ground scene predictions. The techniques are applied to a sequence of pixels as described in the following.

6.2.1 AR Model

The AR model was adopted to compute the GSP, which can be defined as [82]

$$y[n] = -\sum_{k=1}^{p} a[k]y[n-k] + u[n], \quad n = 1,2,\ldots,N,$$

where $y[n]$ is the value of each pixel in one image, N is the number of images in the stack, $a[k]$ are the autoregressive terms, $u[n]$ is white noise, and p is the order of the model [4].

The autoregressive terms $a[k]$ can be estimated by the Yule-Walker method [138]. Hence, the estimated autoregressive terms $\widehat{a}[k]$ are the solutions of the equation system, given by [4]

$$\begin{bmatrix} r_{yy}[0] & r_{yy}[1] & \ldots & r_{yy}[p-1] \\ r_{yy}[1] & r_{yy}[0] & \ldots & r_{yy}[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{yy}[p-1] & r_{yy}[p-2] & \ldots & r_{yy}[0] \end{bmatrix} \begin{bmatrix} a[1] \\ a[2] \\ \vdots \\ a[p] \end{bmatrix} = - \begin{bmatrix} r_{yy}[1] \\ r_{yy}[2] \\ \vdots \\ r_{yy}[p] \end{bmatrix}, \tag{6.1}$$

where $r_{yy}[\cdot]$ is the sample autocorrelation function. Information about large sample distributions of the Yule-Walker estimator, order selection, and confidence regions for the coefficients can be

found in [138].

Considering the estimated autoregressive terms obtained by (6.1), it is possible to forecast h steps ahead with the AR model as [135]

$$\widehat{y}[N+h] = -\sum_{k=1}^{p} \widehat{a}[k]y[N+h-k].$$

The ground scene prediction image is obtained by forecasting the one-step ahead ($h = 1$) pixel value for each pixel in the image.

6.2.2 Trimmed Mean, Median, and Mean

For SAR images whose backscattering is stable in time, robust methods can be applied to obtain a GSP. We consider the trimmed mean to obtain a GSP, which is given by

$$\bar{y}_{\text{tm}} = \frac{2}{N-2m} \sum_{n=m+1}^{N-m} y^{\star}[n],$$

where $y^{\star}[n]$ is the ordered sequence of $y[n]$, $m = (N-1)\alpha$, and $\alpha \in [0, 1/2)$ [217,218]. If $\alpha = 0$ or $\alpha \to 0.5$, then the trimmed mean corresponds to the sample mean and median, respectively [217], which are considered as methods for GSP derivation.

6.2.3 Intensity Mean

We also use the intensity mean for obtaining ground scene predictions, given by

$$\bar{y}_{\text{im}} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} y[n]^2}.$$

Compared to other statistical methods, the intensity mean has the advantage of providing physical interpretation about the image reflection. However, the intensities values contribute evenly to the prediction results, which can be strongly affected by the changes in the ground scene [217].

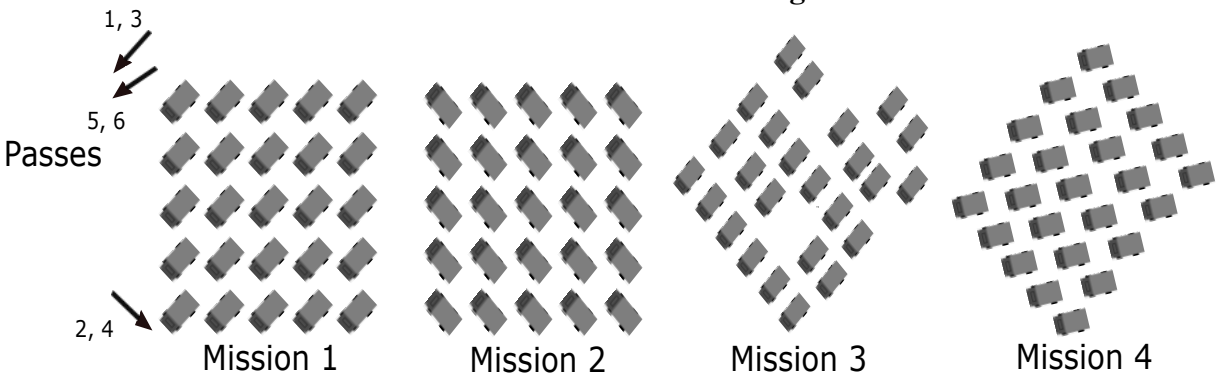
6.3 EXPERIMENTAL RESULTS

In this section, we present the results obtained from the discussed ground scene prediction methods and describe an approach for change detection based on such methods.

6.3.1 Data Description

In this study, we considered a data set obtained from CARABAS II, a Swedish UWB VHF SAR system system whose images are available in [85]. The data set was divided into

Figure 23 – Missions 1, 2, 3, and 4 and passes 1, 2, 3, 4, 5, and 6 considered in CARABAS II images.



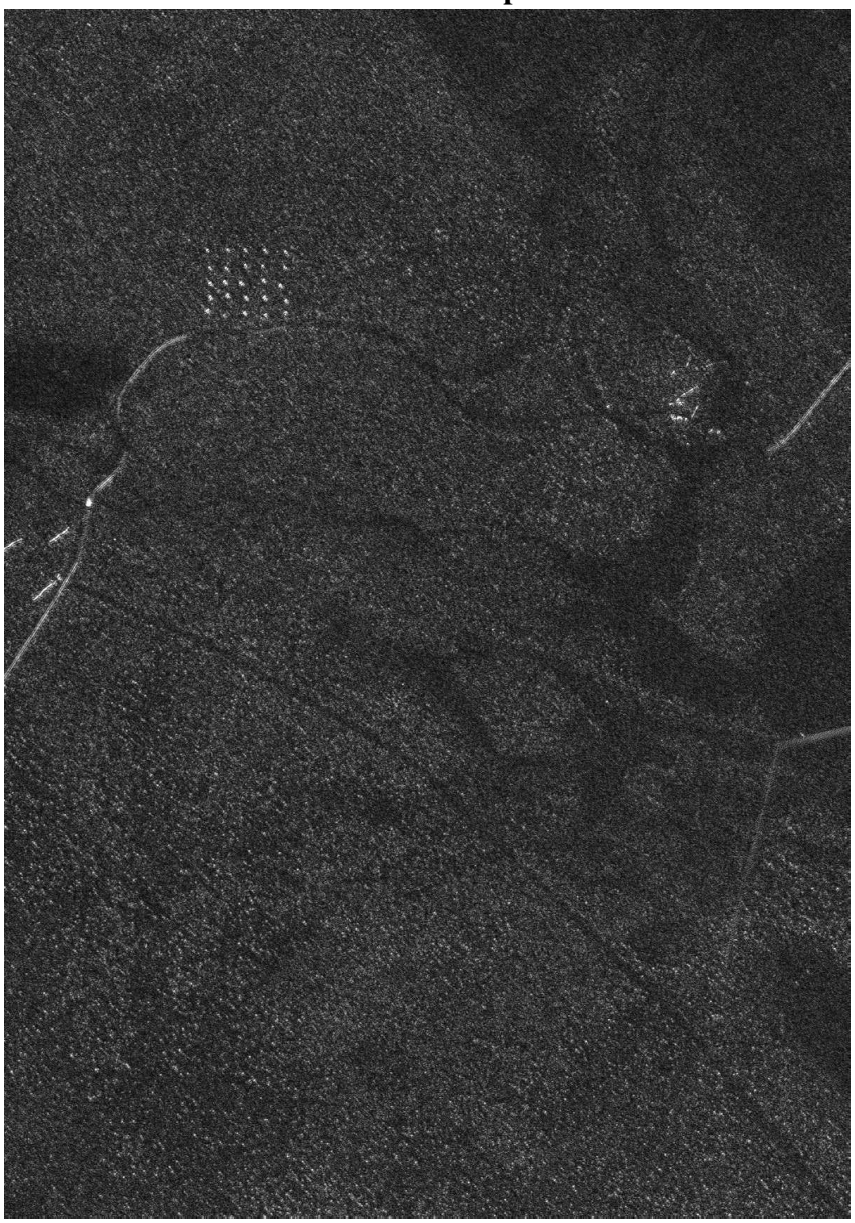
Source: Author (2020)

three stacks with eight images each, i.e., two out of six passes have identical flight headings. Two passes have a flight heading of 255°, two of 135°, and two of 230°, and the heading is defined as 0° pointing towards the north with clock-wise increasing heading. The images in the stacks have the same flight geometry but are associated with four different targets deployments (missions 1 to 4) in the ground scene, as specified in Figure 23. Hence, with four missions and six passes for each mission, there is 24 magnitude single-look SAR images. The images cover a scene of size 2 km $\times$ 3 km and are georeferenced to the Swedish reference system RR92 [2, 84].

The first stack is composed of images corresponding to flight passes 1 and 3; the second stack, with passes 2 and 4; and the last stack is composed of images associated with passes 5 and 6. In all images, the backscattering was stable in time, and only target changes are expected within the image stacks.

Each image is represented as a matrix of 3000 $\times$ 2000 pixels, corresponding to an area of 6 km^2. As reported in [2], the spatial resolution of CARABAS II is 2.5 m in azimuth and 2.5 m in range. The ground scene is dominated by boreal forest with pine trees. Fences, power lines, and roads were also present in the scene. Military vehicles were deployed in the SAR scene and placed uniformly, in a manner to facilitate their identifications in the tests [84]. Each image has 25 targets with three different sizes and the spacing between the vehicles was about 50 meters. For illustration, one image of Stack 1 is shown in Figure 24. In this image the vehicles were obscured by foliage and deployed in the top left of the scene and were oriented in a south-western heading. This deployment corresponds to mission 1. In missions 2, 3, and 4, these vehicles were deployed in other locations and were oriented in a north-western, south-western, and western heading, respectively [2, 84].

Figure 24 – CARABAS II mission one and pass one associated SAR image.



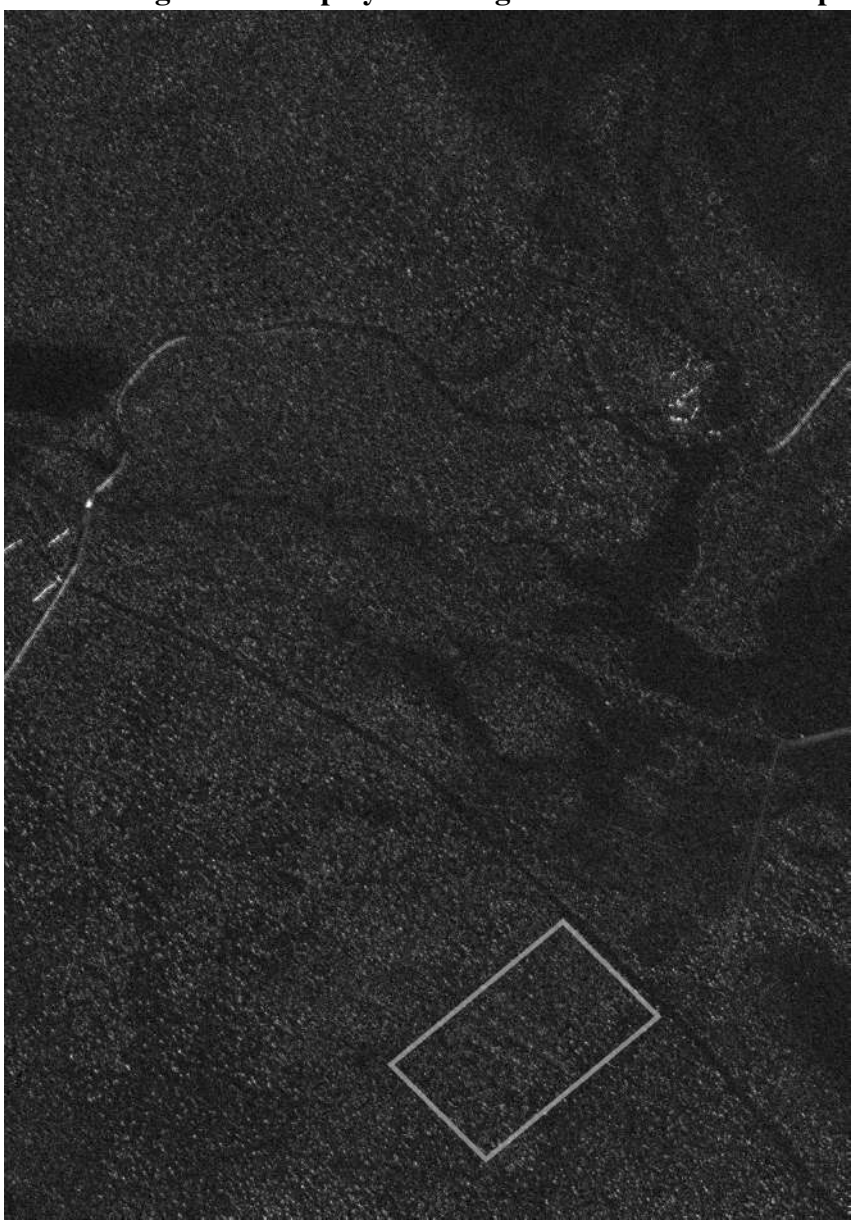
Source: Author (2020)

6.3.2 Ground Scene Prediction Evaluation

The AR model parameter estimation requires (i) fitting 6,000,000 models (one fit for each pixel) in each stack and (ii) evaluating the best model for each pixel sequence. Such demands lead to a significant computational burden. For simplicity, we considered $p = 1$ in the AR model. Within the image stack, the two images related to the targets have the highest pixel values in the areas where the targets were deployed. Thus, to compute the trimmed mean, we considered $m = 2$ ($\alpha \approx 0.3$), expecting to remove the pixels related to targets.

Figures 25-29 show the ground scene prediction for Stack 1, considering the discussed methods. Deployed targets are visually present in the ground scene images predicted with the

Figure 25 – Ground scene prediction image for Stack 1 based on the AR model. The areas highlighted by rectangles in the images indicate the regions where the targets were deployed during the measurement campaign.

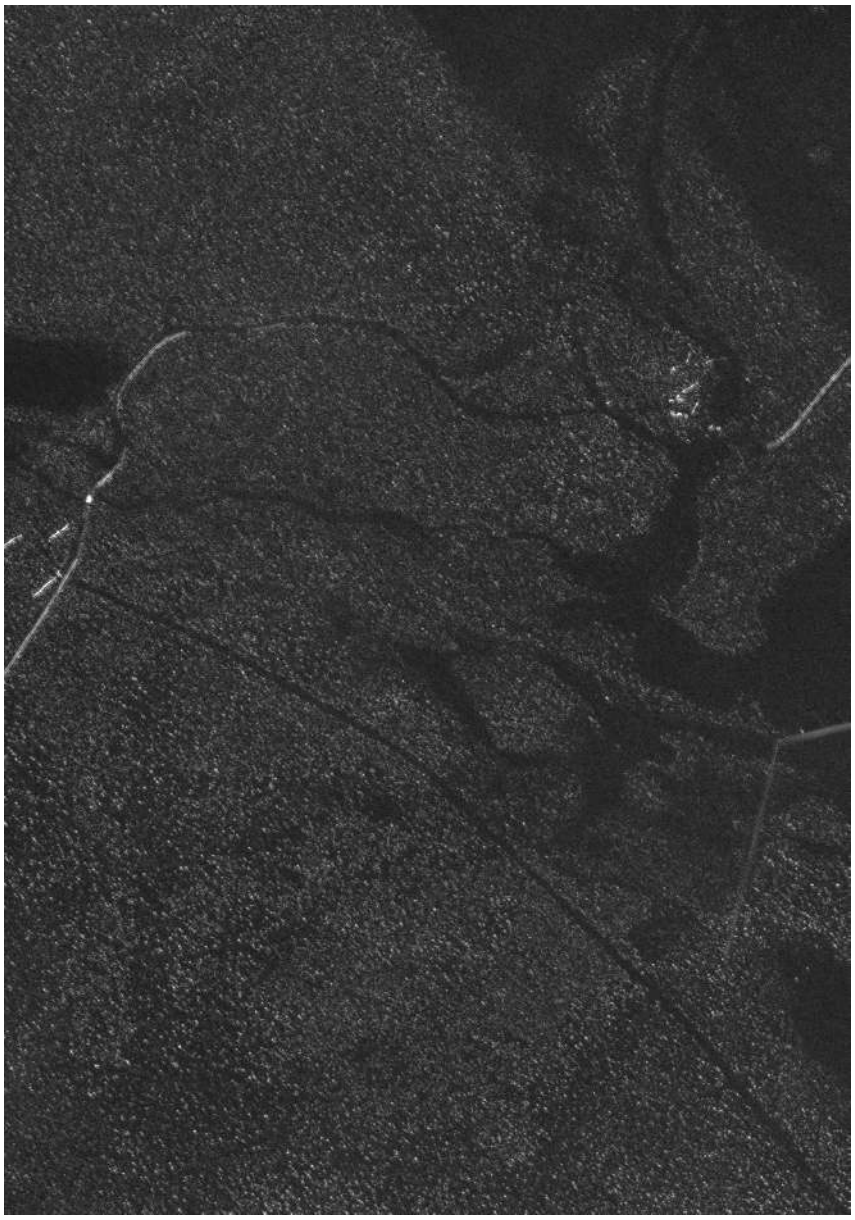


Source: Author (2020)

AR model, mean, and intensity mean methods. However, the targets images are absent in the images predicted with trimmed mean and median. The areas highlighted by rectangles in the images indicate the regions where the targets were deployed during the measurement campaign. With such visual analysis, the trimmed mean and median methods show better performance, i.e., better prediction of the ground scene. For brevity, we limited our presentation in this chapter to the GSP images from the Stack 1, which is representative of all considered stacks. The GSP images from Stack 2 and 3 can be found in Appendix G.

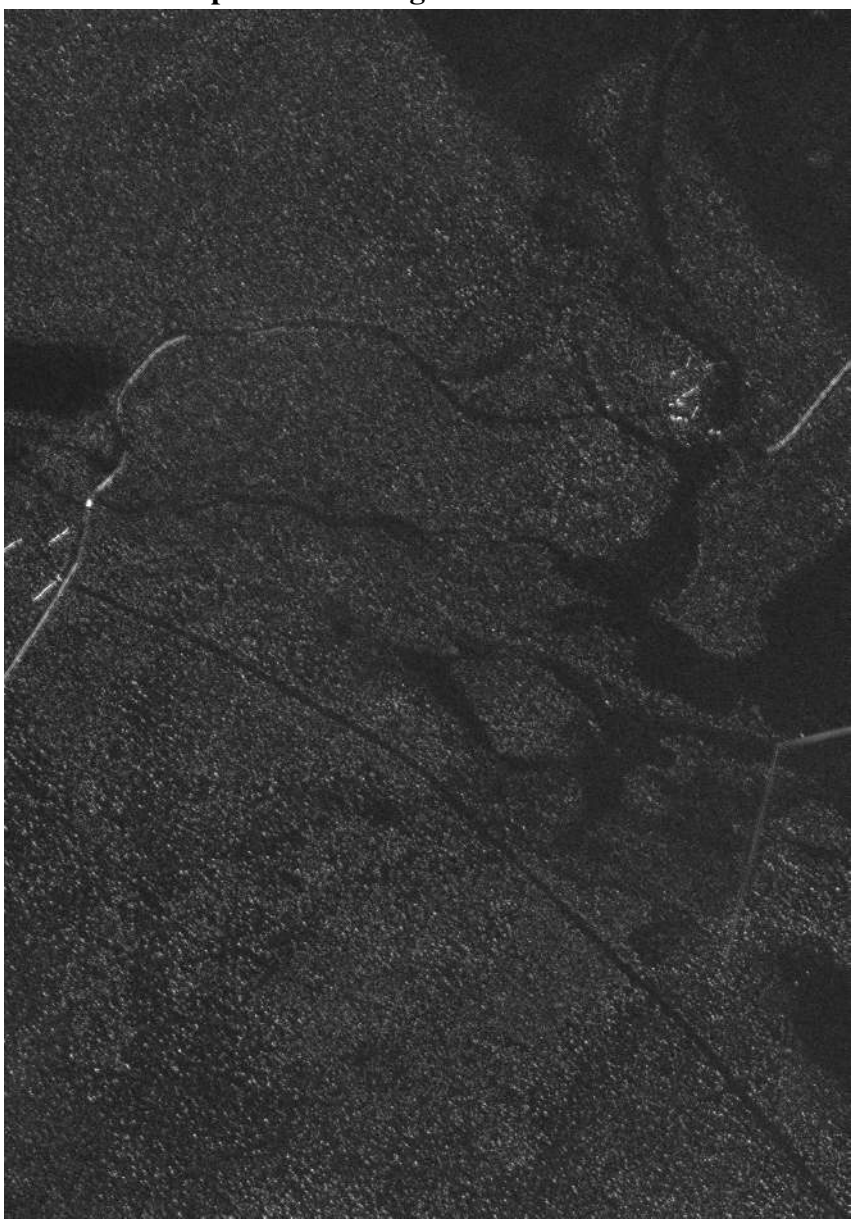
Table 14 displays descriptive statistics of the employed images, such as average, standard

Figure 26 – Ground scene prediction image for Stack 1 based on the trimmed mean method.



Source: Author (2020)

Figure 27 – Ground scene prediction image for Stack 1 based on the median method.



Source: Author (2020)

deviation, skewness, and kurtosis. It is desirable that a GSP presents not only a good visual representation of the true ground but also preserves the statistical characteristics of the image of interest. In Table 14, we highlighted the two best methods according to each considered measure. In the majority of the scenarios, the AR model and median methods outperformed the remaining methods.

To evaluate the difference between the ground scene prediction methods, we computed some standard quality adjustment measures. The criteria are the mean square error (MSE), mean absolute percentage error (MAPE), and median absolute error (MdAE), which can be defined as

Figure 28 – Ground scene prediction image for Stack 1 based on the mean method. The areas highlighted by rectangles in the images indicate the regions where the targets were deployed during the measurement campaign.

Source: Author (2020)

Figure 29 – Ground scene prediction image for Stack 1 based on the intensity mean method. The areas highlighted by rectangles in the images indicate the regions where the targets were deployed during the measurement campaign.

Source: Author (2020)

Table 14 – Average, standard deviation, skewness, and kurtosis of one interest image and the ground scene prediction. The interest image in Stack 1, 2, and 3, is the image of mission 1 and pass 1, 2, and 5, respectively. The two values of each measure that yielded the closest values with the interest image are highlighted

	Average	Standard deviation	Skewness	Kurtosis
Stack 1				
Interest image	0.1442	0.0894	1.8597	14.1740
AR model	0.1101	**0.0725**	**2.1120**	**13.5190**
Trimmed mean	**0.1430**	0.0680	2.9051	21.2919
Median	**0.1424**	**0.0688**	**2.8231**	**20.4990**
Mean	0.1467	0.0663	3.0516	22.8448
Intensity mean	0.1592	0.0667	3.0090	22.8725
Stack 2				
Interest image	0.1373	0.0968	2.9345	30.5666
AR model	0.0997	0.0784	**3.6398**	**40.9991**
Trimmed mean	**0.1344**	**0.0806**	4.4488	55.4260
Median	0.1339	**0.0812**	**4.3664**	**53.9367**
Mean	**0.1376**	0.0792	4.6022	58.3558
Intensity mean	0.1485	0.0792	4.5487	57.8894
Stack 3				
Interest image	0.1451	0.0905	1.8583	14.0932
AR model	0.0997	**0.0683**	**2.2034**	**14.6539**
Trimmed mean	**0.1372**	0.0665	2.8811	22.0954
Median	0.1366	**0.0674**	**2.8090**	**21.3242**
Mean	**0.1410**	0.0646	2.9582	22.9540
Intensity mean	0.1534	0.0655	2.9170	22.9794

Source: Author (2020)

follows [142]

$$\text{MSE} = \frac{1}{Q}\sum_{q=1}^{Q}(x[q]-\widehat{x}[q])^2,$$

$$\text{MAPE} = \frac{1}{Q}\sum_{q=1}^{Q}\frac{|x[q]-\widehat{x}[q]|}{|x[q]|},$$

$$\text{MdAE} = \text{Median}\left(|x[q]-\widehat{x}[q]|\right), \quad q = 1,2,\ldots,Q,$$

where $x[q]$ and $\widehat{x}[q]$ are the pixel values of the interest and predicted images respectively, Q is the number of pixels, and Median$(\cdot)$ is the median value of $|x[q]-\widehat{x}[q]|$, for $q = 1,2,\ldots,Q$. These goodness-of-fit measures are usually considered to compare different methods applied to the same data set [142]. They are expected to be as close to zero as possible.

For the quality adjustment measures, the target regions in the image were excluded since

Table 15 – Measures of quality of the ground scene prediction image. The interest image in Stack 1, 2, and 3 is the image of mission 1 and pass 1, 2, and 5, respectively. We highlighted the values of each quality adjustment measure that yielded the smallest values

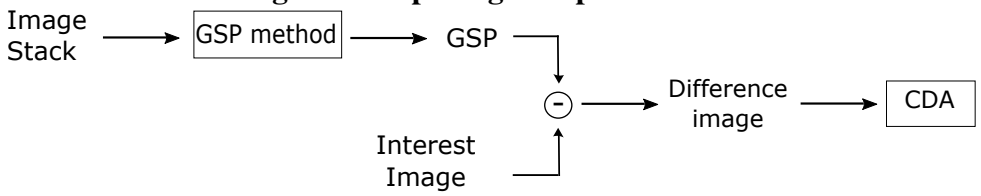
		MSE	MAPE	MdAE
Stack 1	AR model	0.0077	0.6756	0.0548
	Trimmed mean	**0.0036**	0.6187	0.0364
	Median	0.0037	**0.6125**	**0.0351**
	Mean	0**.0036**	0.6489	0.0401
	Intensity mean	0.0039	0.7505	0.0426
Stack 2	AR model	0.0068	0.6450	0.0502
	Trimmed mean	**0.0030**	0.5971	0.0326
	Median	0.0031	**0.5912**	**0.0315**
	Mean	**0.0030**	0.6254	0.0359
	Intensity mean	0.0032	0.7204	0.0378
Stack 3	AR model	0.0083	0.6337	0.0557
	Trimmed mean	0.0037	0.5809	0.0357
	Median	0.0038	**0.5751**	**0.0346**
	Mean	0**.0036**	0.6104	0.0392
	Intensity mean	0.0037	0.7011	0.0410

Source: Author (2020)

we expect to obtain an accurate ground scene prediction, and no target deployment should influence the measurements. Table 15 summarizes the results of the quality adjustment measures for the five considered statistical methods, and the best measurements are highlighted. The mean method presents the best performance according to MSE measurements, while the median method excels in terms of MAPE and MdAE measures in all the stacks. However, the MSE values obtained with the mean and median methods are similar. The results provided in Tables 14 and 15 consider the same reference image of each stack. Regardless of the selected image, the median method presented good performance according to MAPE, MdAE, and statistics measures.

Based on visual inspection, statistical characteristics, and quality adjustment measures, the median method yields the most reliable prediction among the considered methods. Therefore, we separate the predicted images from the median method as reference images in the change detection algorithm detailed in the next section.

Figure 30 – Processing scheme for change detection. The CDA is performed applying thresholding and morphological operations in the difference image.



Source: Author (2020)

6.3.3 Application in Change Detection

The change detection method used in this chapter applied the processing scheme given in Figure 30. Firstly, an image stack is processed by a desirable GSP method furnishing the GSP image. The changes are simply obtained with the subtraction of the image of interest (surveillance image) from the GSP image (reference image). For change detection, we applied thresholding to the difference image and then used morphological operations for false alarm minimization.

Two examples of subtraction images are shown in Figures 31 and 32. Figure 31 highlights the deployed targets while Figure 32 focuses on the targets and the back-lobe structures. A comparison between the difference image shown in Figure 32 to the related GSP image suggests that the back-lobe structures are related to issues in the SAR system and image formation algorithm.

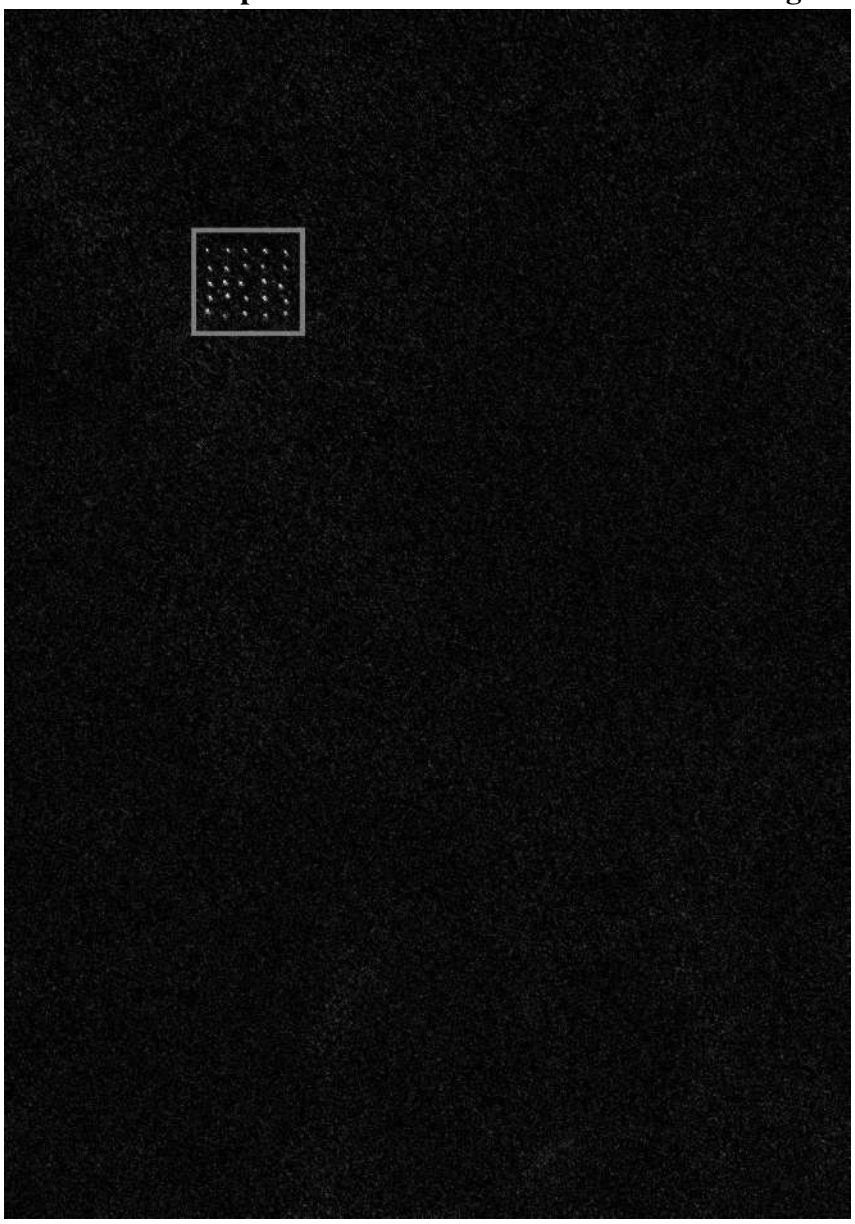
The employed CDA [84] consists of two mathematical morphology steps. First, an opening operation [139, 226] aimed at removing small pixels, which are regarded as noise. The second step is a dilation that prevents the splitting of the interest targets in multiple substructures. The first step uses a 3×3 pixel square structuring element, whose size is determined by the system resolution; the second step considers a 7×7 pixel structuring element, which is linked to the approximate size of the targets (about 10×10 pixels).

Figure 33 shows the pixels values of the image given in Figure 31 in vectorized form. In general, the subtracted image pixels values are randomly distributed in $(-0.4, 0.4)$. The extreme values in Figure 33 are the pixels where the targets are deployed. As discussed in [207], the distribution of the values of the subtraction image of CARABAS II approximately follows the Gaussian distribution and the regions where no change occurs are stable. Thus, the threshold (λ) can be simply chosen as [207],

$$C = \frac{\lambda - \widehat{\mu}}{\widehat{\sigma}},$$

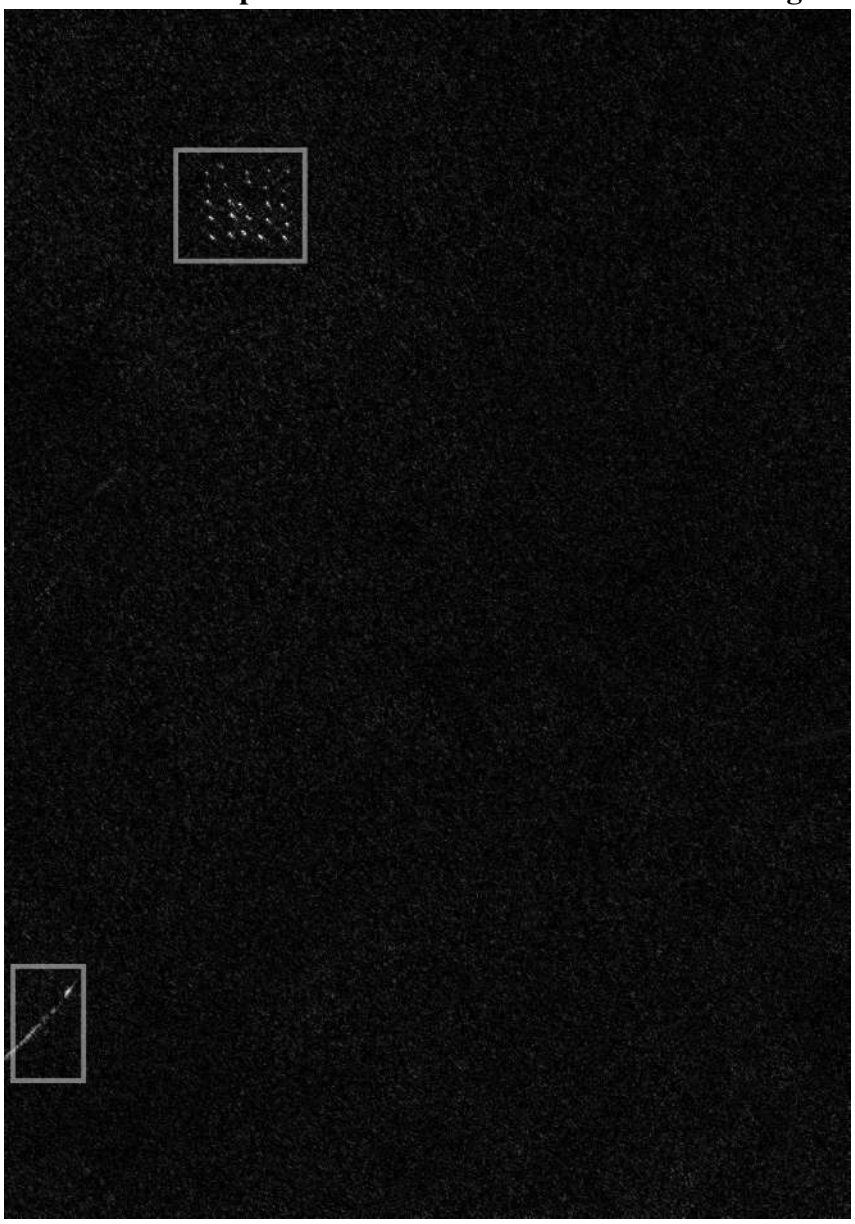
where C is a constant, $\widehat{\mu}$ is the estimated mean, and $\widehat{\sigma}$ is the estimated standard deviation of the

Figure 31 – Subtraction of an interest image from the median ground scene prediction for pass one and mission one associated image.



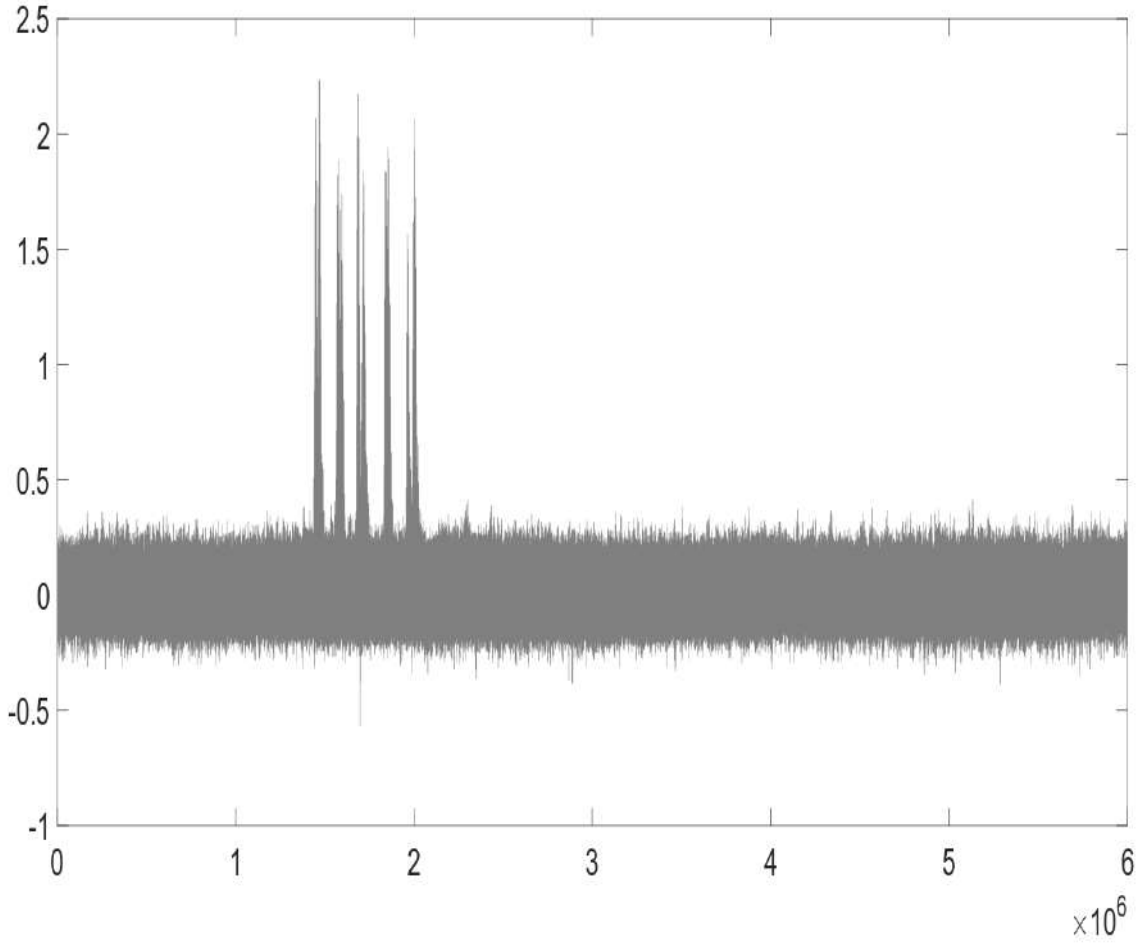
Source: Author (2020)

Figure 32 – Subtraction of an interest image from the median ground scene prediction for pass one and mission two associated image.



Source: Author (2020)

Figure 33 – Result of the subtraction of the ground scene prediction image from the image obtained from mission one and pass one.



Source: Author (2020)

considered amplitude pixels in the image. For the evaluation, we set $C \in \{2;3;4;5;6\}$, resulting in different values false alarm rates (FAR), which range from full detection to almost null false alarm rate.

The performance of change detection was evaluated by the probability of detection (P_d) and FAR. The quantity P_d was obtained from the ratio between the number of detected targets and the total numbers of known targets, while FAR is defined by the number of false alarms detected per square kilometer [84].

Table 16 presents the change detection results for $C = 5$. Among 600 deployed vehicles in the missions, 579 were correctly detected. There is 22 detected objects that can not be related to any vehicle and were considered to be false alarms. Thus, the detection probability is about 97%, while the false alarm rate is $0.15/\text{km}^2$ (total of $144/\text{km}^2$). Ten of the 22 false alarms are related to the back-lobe structures, i.e., they are not actually false alarms and may stem from system and image formation issues. Additionally, in general, the undetected targets are related to missions

Table 16 – Change detection results obtained with $C = 5$

Case of Interest		Number of	Detected	P_d	Number of
Mission	Pass	known targets	Targets		false alarms
1	1	25	25	1.00	0
2	1	25	25	1.00	3
3	1	25	25	1.00	0
4	1	25	23	0.92	2
1	2	25	25	1.00	0
2	2	25	25	1.00	1
3	2	25	25	1.00	2
4	2	25	23	0.92	1
1	3	25	25	1.00	2
2	3	25	23	0.92	0
3	3	25	25	1.00	3
4	3	25	23	0.92	0
1	4	25	25	1.00	0
2	4	25	25	1.00	0
3	4	25	25	1.00	1
4	4	25	23	0.92	0
1	5	25	25	1.00	0
2	5	25	15	0.60	6
3	5	25	25	1.00	0
4	5	25	24	0.96	0
1	6	25	25	1.00	0
2	6	25	25	1.00	1
3	6	25	25	1.00	0
4	6	25	25	1.00	0
Total		600	579	0.97	22

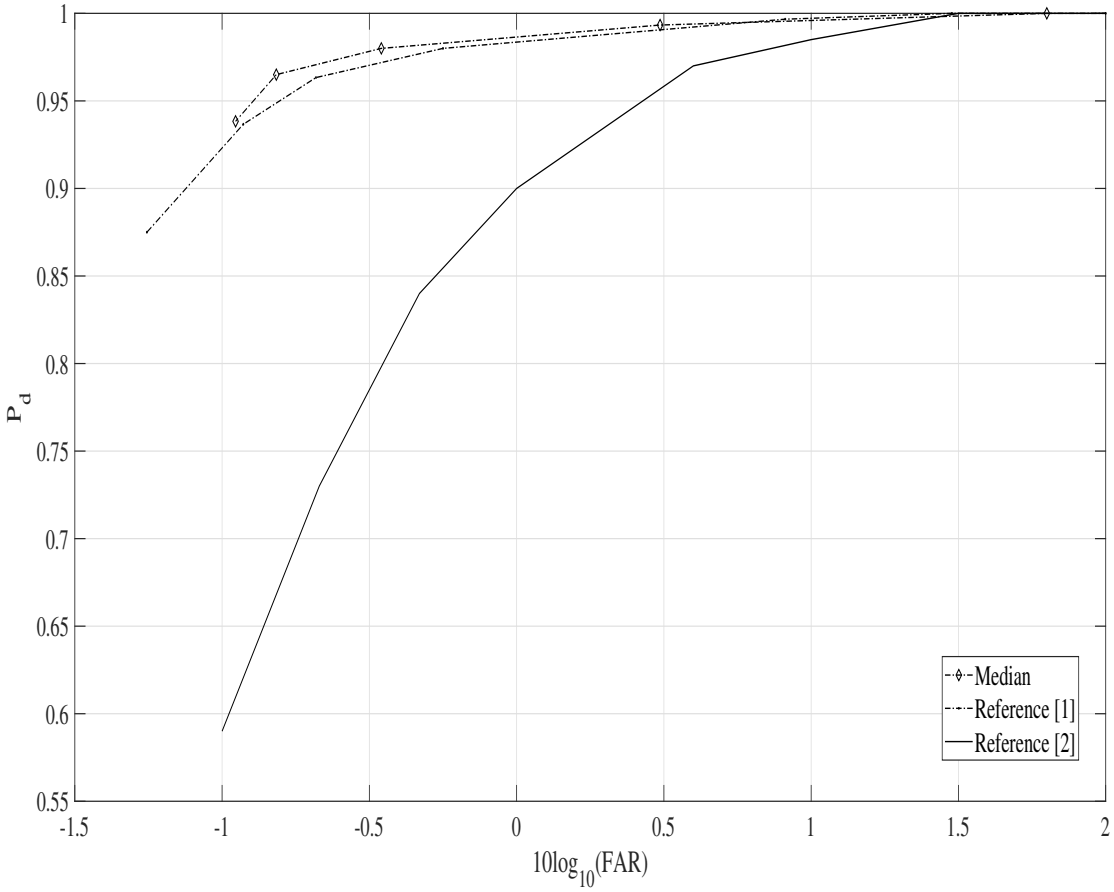
Source: Author (2020)

2 and 4. These undetected military vehicles are more difficult to detect since they have the smaller sizes and magnitude values, and consequently, pixel values closer to the forest ones. In accordance with [1, 2], the image associated to mission two and pass five has the highest number of false alarms and smallest number of detected targets.

Figure 34 presents the receiver operating characteristic (ROC) curves [199] of the change detection results, showing the probability of detection versus the false alarm rates for the different evaluated values of C. We compared the change detection results obtained from the proposed method with the results described in [2] and [1]. The proposed method excels in terms of probability of detection and false alarm rate in comparison to [1, 2].

For example, for a detection probability of 98%, our proposed change detection method presents $\log_{10}(\text{FAR})$ about -0.5, while [1, 2] have $\log_{10}(\text{FAR})$ about 1.4 and -0.3, respectively. For $\log_{10}(\text{FAR}) = -0.9$, i.e., a very low FAR, the probability detection given by [2] drops to 60%,

Figure 34 – ROC curves obtained with the CDA with the background predicted scene as the reference image compared with the best ROC curves extracted from [1, 2].



Source: Author (2020)

while our proposal still maintains the probability of detection more than 90%. The detection probability of our proposed method and [1] reach 100% with $\log_{10}(\text{FAR}) \approx 1$, while [2] has full detection for $\log_{10}(\text{FAR}) \approx 1.5$. Additionally, detection probability improvements of our method compared to [1] are found in the range of $(0.93, 0.98)$. For example, for a probability of detection of 0.97%, our proposed change detection method presents $\log_{10}(\text{FAR})$ about -0.8, while [1] has $\log_{10}(\text{FAR}) \approx -0.2$.

6.4 CONCLUSION

In this chapter, we presented five methods to obtain ground scene prediction of SAR images based on image stack. The experimental results revealed that among the considered techniques, the median method yielded the most accurate ground prediction. The statistical characteristics of the obtained GSP image were similar to the image of interest. Moreover, the

median method excels in terms of quality adjustment measures, and the changes in the image stack were not visually presented in the predicted image. We demonstrated the suitability of the proposed GSP method by presenting competitive performance when compared with recently published results.

7 CONCLUDING REMARKS AND FUTURE WORKS

In this chapter, some concluding remarks, contributions, and directions for future works are presented.

7.1 CONCLUDING REMARKS

In Chapter 2, we introduced a new regression model for nonnegative signals. The derived regression model assumes that the mean of the Rayleigh distributed signal follows a regression structure involving covariates, unknown parameters, and a link function. For the proposed model, we introduced an inference approach for the model parameters and discussed diagnostic tools, such as coefficient of determination and residual. Additionally, Fisher information matrix, asymptotic proprieties of the MLE, and a detector useful to detect differences in SAR image regions were presented. The parameter estimators of the proposed model was evaluated by Monte Carlo simulations, showing small values of percentage relative bias and mean square error. An application of the derived Rayleigh regression model was considered in SAR images to distinguish different regions. The Rayleigh regression model presented more accurate detection results when compared with the Gaussian-, Gamma-, and Weibull-based regression models.

In Chapter 3, we derived bias-adjusted estimators for the parameters of the Rayleigh regression model. In particular, we considered the Cox and Snell's, Firth's, and parametric bootstrap method to obtain the corrected estimators. In the Monte Carlo simulations, the MLEs of the parameters of the Rayleigh regression model displayed bigger values of relative bias and root mean square error when compared to the proposed corrected estimators. The estimators derived by the Firth's method outperformed the others bias-adjusted evaluated estimators.

In Chapter 4, we derived the 2D RARMA model and a SAR image ground type change detector considering residual-based control charts. Additionally, we introduced an inference approach for model parameters, the conditional Fisher information matrix, and the asymptotic properties of the CMLE. Monte Carlo simulations were used as a tool to evaluate the performance of the CMLE. Two investigations of the proposed model to distinguish between ground type changes in SAR images were presented and discussed, showing competitive detection results when compared with traditional 2D ARMA models.

In Chapter 5, we proposed that the mean of a beta binomial distributed signal in time follows a dynamic model involving covariates, unknown parameters, and a link function. We

developed parameter estimation and detection tools. For the introduced model, we discussed the conditional observed information matrix and asymptotic proprieties of the CMLE. The CLMEs of the BBARMA model parameters were evaluated by Monte Carlo simulations, showing small values of percentage relative bias and mean square error. Finally, we discussed an application of the BBARMA model to detect digital signals, presenting more accurate detection results than the Gaussian-, ARMA-, and beta binomial-based detection.

Finally, in Chapter 6, we considered five different statistical methods to obtain a ground scene prediction of SAR image based on image stack. We evaluated the descriptive statistics of the GSP images. Also, a visual inspection trying to verify the presence of changes in the ground scene and some quality adjustment measures were considered. The median method was elected as the most accurate method to obtain the ground scene prediction for SAR images. The ground scene prediction obtained by the median method was applied in a CDA, presenting competitive performance when compared with other recent results.

7.2 CONTRIBUTIONS

The contributions of this thesis are:

1. Palm, B.G., Bayer, F.M., Cintra, R.J., Pettersson, M.I. and Machado, R., 2019. Rayleigh Regression Model for Ground Type Detection in SAR Imagery. *IEEE Geoscience and Remote Sensing Letters*, 16(10), pp.1660-1664.

2. Palm, B.G., Alves, D.I., Vu, V.T., Pettersson, M.I., Bayer, F.M., Cintra, R.J., Machado, R., Dammert, P. and Hellsten, H., 2020. Wavelength-Resolution SAR Ground Scene Prediction Based on Image Stack. *Sensors*, 20(7), 2008.

3. Palm, B.G., Alves, D.I., Vu, V.T., Pettersson, M.I., Bayer, F.M., Cintra, R.J., Machado, R., Dammert, P. and Hellsten, H., 2018, October. Autoregressive model for multi-pass SAR change detection based on image stacks. In *Image and Signal Processing for Remote Sensing XXIV*. International Society for Optics and Photonics.

The manuscripts are:

1. Palm, B.G., Bayer, F.M., Cintra, R.J. Improved Point Estimation for the Rayleigh Regression Model.

2. Palm, B.G., Bayer, F.M., Cintra, R.J. 2D Rayleigh Autoregressive Moving Average Model for Change Detection in SAR Imagery.

3. Palm, B.G., Bayer, F.M., Cintra, R.J. Beta Binomial Autoregressive Moving Average Model for Digital Signals Detection and Inference.

7.3 FUTURE WORKS

As next steps of research, we have the following immediate goals:

1. developing bias-corrected estimators, using bootstrap and Cox-Snell correction, for the 2D Rayleigh autoregressive moving average model;
2. using robust methods for the estimation of 2D Rayleigh autoregressive moving average model parameters in presence of outliers;
3. studying spectral analysis of the proposed models;
4. extending the beta binomial autoregressive moving average model to a two-dimensional model;
5. deriving two-dimensional models for other distribution classes;
6. studying model selection criteria for the derived models;
7. applying the developed methods in the most different real applications in signal and image processing;
8. studying the stability and stationary regions of the proposed models.

BIBLIOGRAPHY

1 VU, V. T. Wavelength-resolution SAR incoherent change detection based on image stack. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 14, n. 7, p. 1012–1016, 2017.

2 ULANDER, L. M. *et al.* Change detection for low-frequency SAR ground surveillance. **IEEE Proceedings-Radar, Sonar and Navigation**, IEEE, v. 152, n. 6, p. 413–420, 2005.

3 INGLADA, J.; MERCIER, G. A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 45, n. 5, p. 1432–1445, 2007.

4 KAY, S. M. **Fundamentals of statistical signal processing: Detection theory**. [S.l.]: Prentice Hall, 1998. II.

5 YUCEK, T.; ARSLAN, H. A survey of spectrum sensing algorithms for cognitive radio applications. **IEEE Communications Curveys & Tutorials**, IEEE, v. 11, n. 1, p. 116–130, 2009.

6 GARTH, L. M.; POOR, H. V. Detection of non-Gaussian signals: A paradigm for modern statistical signal processing. **Proceedings of the IEEE**, IEEE, v. 82, n. 7, p. 1061–1095, 1994.

7 AL-SMADI, A.; ALSHAMALI, A. Fitting ARMA models to linear non-Gaussian processes using higher order statistics. **Signal Processing**, Elsevier, v. 82, n. 11, p. 1789–1793, 2002.

8 CHAVALI, V. G.; SILVA, C. R. D. Maximum-likelihood classification of digital amplitude-phase modulated signals in flat fading non-Gaussian channels. **IEEE Transactions on Communications**, IEEE, v. 59, n. 8, p. 2051–2056, 2011.

9 NASSAR, M. *et al.* Mitigating near-field interference in laptop embedded wireless transceivers. **Journal of Signal Processing Systems**, Springer, v. 63, n. 1, p. 1–12, 2011.

10 WANG, X.; POOR, H. V. Robust multiuser detection in non-Gaussian channels. **IEEE Transactions on Signal Processing**, IEEE, v. 47, n. 2, p. 289–305, 1999.

11 TSIHRINTZIS, G. A.; NIKIAS, C. L. Performance of optimum and suboptimum receivers in the presence of impulsive noise modeled as an alpha-stable process. **IEEE Transactions on Communications**, IEEE, v. 43, n. 234, p. 904–914, 1995.

12 VASTOLA, K. Threshold detection in narrow-band non-Gaussian noise. **IEEE Transactions on Communications**, IEEE, v. 32, n. 2, p. 134–139, 1984.

13 OPPENHEIM, A. V.; SCHAFER, R. W. **Discrete-time Signal Processing**. Third. Upper Saddle River: Pearson, 2010. (Prentice-Hall Signal Processing Series).

14 GOMES, N. R. *et al.* Comparison of the Rayleigh and k-distributions for application in incoherent change detection. **IEEE Geoscience and Remote Sensing Letters**, IEEE, 2018.

15 CINTRA, R. J.; FRERY, A. C.; NASCIMENTO, A. D. Parametric and nonparametric tests for speckled imagery. **Pattern Analysis and Applications**, Springer, v. 16, n. 2, p. 141–161, 2013.

16 KURUOGLU, E. E.; ZERUBIA, J. Modeling SAR images with a generalization of the Rayleigh distribution. **IEEE Transactions on Image Processing**, IEEE, v. 13, n. 4, p. 527–533, 2004.

17 OLIVER, C.; QUEGAN, S. **Understanding synthetic aperture radar images**. [S.l.]: SciTech Publishing, 2004.

18 JACKSON, J. A.; MOSES, R. L. A model for generating synthetic VHF SAR forest clutter images. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 45, n. 3, 2009.

19 FRANSSON, J. E. *et al.* Detection of storm-damaged forested areas using airborne CARABAS-II VHF SAR image data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 40, n. 10, p. 2170–2175, 2002.

20 MAGNUSSON, M.; FRANSSON, J. E. Combining carabas-ii vhf sar and landsat tm satellite data for estimation of forest stem volume. In: IEEE. **IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium**. [S.l.], 2004. v. 4, p. 2327–2331.

21 FRANSSON, J. E. *et al.* Estimation of stem volume in hemi-boreal forests using airborne low-frequency Synthetic Aperture Radar and lidar data. In: IEEE. **2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS**. [S.l.], 2013. p. 161–164.

22 SARTI, A. *et al.* Maximum likelihood segmentation of ultrasound images with Rayleigh distribution. **IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control**, IEEE, v. 52, n. 6, p. 947–960, 2005.

23 GUPTA, S.; CHAUHAN, R.; SAXENA, S. Locally adaptive wavelet domain Bayesian processor for denoising medical ultrasound images using Speckle modelling based on Rayleigh distribution. **IEE Proceedings-Vision, Image and Signal Processing**, IET, v. 152, n. 1, p. 129–135, 2005.

24 TARICCO, G. On the convergence of multipath fading channel gains to the Rayleigh distribution. **IEEE Wireless Communications Letters**, IEEE, v. 4, n. 5, p. 549–552, 2015.

25 ZANETTI, M.; BOVOLO, F.; BRUZZONE, L. Rayleigh-rice mixture parameter estimation via EM algorithm for change detection in multispectral images. **IEEE Transactions on Image Processing**, IEEE, v. 24, n. 12, p. 5004–5016, 2015.

26 GOMES, N. R. *et al.* Likelihood ratio test for incoherent wavelength-resolution SAR change detection. In: **2016 CIE International Conference on Radar (RADAR)**. [S.l.]: IEEE, 2017. p. 1–4.

27 SUMAIYA, M. N.; KUMARI, R. S. S. Unsupervised change detection of flood affected areas in SAR images using Rayleigh-based Bayesian thresholding. **IET Radar, Sonar & Navigation**, IET, v. 12, n. 5, p. 515–522, 2018.

28 LAMPROPOULOS, G. *et al.* High resolution radar clutter statistics. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 35, n. 1, p. 43–60, 1999.

29 YUE, D.-X. *et al.* A generalized Gaussian coherent scatterer model for correlated SAR texture. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, 2019.

30 SCHWARTZ, M. **Mobile wireless communications**. [S.l.]: Cambridge University Press, 2005.

31 CORDEIRO, G. M.; KLEIN, R. Bias correction in ARMA models. **Statistics & Probability Letters**, Elsevier, v. 19, n. 3, p. 169–176, 1994.

32 CORDEIRO, G. M.; CRIBARI-NETO, F. **An Introduction to Bartlett Correction and Bias Reduction**. [S.l.]: Springer, 2014.

33 TJØSTHEIM, D. Statistical spatial series modelling. **Advances in Applied Probability**, Cambridge University Press, v. 10, n. 1, p. 130–154, 1978.

34 BUSTOS, O. H. *et al.* Asymptotic behavior of RA-estimates in autoregressive 2D processes. **Journal of Statistical Planning and Inference**, Elsevier, v. 139, n. 10, p. 3649–3664, 2009.

35 KIZILKAYA, A. *et al.* ARMA-cepstrum recursion algorithm for the estimation of the MA parameters of 2-D ARMA models. **Multidimensional Systems and Signal Processing**, Springer, v. 16, n. 4, p. 397–415, 2005.

36 SANT'ANNA, S. *et al.* Textural method to evaluate speckle noise filters based on AR-2D models using P-band radar data. In: IEEE. **IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium**. [S.l.], 2001. v. 3, p. 1506–1508.

37 KASHYAP, R.; CHELLAPPA, R. Estimation and choice of neighbors in spatial-interaction models of images. **IEEE Transactions on Information Theory**, IEEE, v. 29, n. 1, p. 60–72, 1983.

38 KASHYAP, R. Characterization and estimation of two-dimensional ARMA models. **IEEE Transactions on Information Theory**, IEEE, v. 30, n. 5, p. 736–745, 1984.

39 ZHANG, X.-D. On the estimation of two-dimensional moving average parameters. **IEEE Transactions on Automatic Control**, IEEE, v. 36, n. 10, p. 1196–1199, 1991.

40 CHOI, B.; POLITIS, D. N. Modeling 2-D AR processes with various regions of support. **IEEE Transactions on Signal Processing**, IEEE, v. 55, n. 5, p. 1696–1707, 2007.

41 DELP, E. J.; KASHYAP, R. L.; MITCHELI, O. R. Image data compression using autoregressive time series models. **Pattern Recognition**, Elsevier, v. 11, n. 5-6, p. 313–323, 1979.

42 BOUMAN, C. A. **Model based image processing**. [S.l.]: Purdue University, 2013.

43 BASU, S.; REINSEL, G. C. Properties of the spatial unilateral first-order ARMA model. **Advances in Applied Probability**, Cambridge University Press, v. 25, n. 3, p. 631–648, 1993.

44 BOULEMNADJEL, A.; HACHOUF, F.; KHARFOUCHI, S. GMM estimation of 2D-RCA models with applications to texture image classification. **IEEE Transactions on Image Processing**, IEEE, v. 25, n. 2, p. 528–539, 2015.

45 PASCUAL, J. P. *et al.* Adaptive radar detection algorithm based on an autoregressive GARCH-2D clutter model. **IEEE Transactions on Signal Processing**, IEEE, v. 62, n. 15, p. 3822–3832, 2014.

46 SARKAR, A.; SHARMA, K.; SONAK, R. V. A new approach for subset 2-D AR model identification for describing textures. **IEEE Transactions on Image Processing**, IEEE, v. 6, n. 3, p. 407–413, 1997.

47 ZIELINSKI, J.; BOUAYNAYA, N.; SCHONFELD, D. Two-dimensional ARMA modeling for breast cancer detection and classification. In: IEEE. **2010 International Conference on Signal Processing and Communications (SPCOM)**. [S.l.], 2010. p. 1–4.

48 KIZILKAYA, A. Computation of the exact Cramer–Rao lower bound for the parameters of a nonsymmetric half-plane 2-D ARMA model. **Digital Signal Processing**, Elsevier, v. 18, n. 5, p. 835–843, 2008.

49 KASHYAP, R. L.; EOM, K.-B. Robust image modeling techniques with an image restoration application. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, IEEE, v. 36, n. 8, p. 1313–1325, 1988.

50 BENNETT, J.; KHOTANZAD, A. Maximum likelihood estimation methods for multispectral random field image models. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 21, n. 6, p. 537–543, 1999.

51 SMILLIE, G. **Analogue and digital communication techniques**. [S.l.]: Elsevier, 1999.

52 GRIFFITHS, D. Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. **Biometrics**, Wiley Online Library, p. 637–648, 1973.

53 LEE, J. C.; S., D. J. Bayesian estimation and prediction for the beta-binomial model. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 5, n. 3, p. 357–367, 1987.

54 WERNER, G. Á.; HANKA, L. Using the beta-binomial distribution for the analysis of biometric identification. In: IEEE. **2015 IEEE 13th International Symposium on Intelligent Systems and Informatics (SISY)**. [S.l.], 2015. p. 209–215.

55 WAGNER, B.; RIGGS, P.; MIKULICH-GILBERTSON, S. The importance of distribution-choice in modeling substance use data: a comparison of negative binomial, beta binomial, and zero-inflated distributions. **The American Journal of Drug and Alcohol Abuse**, Taylor & Francis, v. 41, n. 6, p. 489–497, 2015.

56 HADDAD, R. A.; PARSONS, T. W. **Digital signal processing: Theory, applications, and hardware**. [S.l.]: Computer Science Press, 1991.

57 GOUVEIA, S. *et al.* A full ARMA model for counts with bounded support and its application to rainy-days time series. **Stochastic Environmental Research and Risk Assessment**, Springer, v. 32, n. 9, p. 2495–2514, 2018.

58 NJIKE, A. N.; PELLERIN, R.; KENNE, J. P. Simultaneous control of maintenance and production rates of a manufacturing system with defective products. **Journal of Intelligent Manufacturing**, Springer, v. 23, n. 2, p. 323–332, 2012.

59 ALBARRACIN, O. Y. E.; ALENCAR, A. P.; HO, L. L. CUSUM chart to monitor autocorrelated counts using negative binomial GARMA model. **Statistical Methods in Medical Research**, SAGE Publications, v. 27, n. 9, p. 2859–2871, 2018.

60 WHITE, R. G. Change detection in sar imagery. **International Journal of Remote Sensing**, Taylor & Francis, v. 12, n. 2, p. 339–360, 1991.

61 ULANDER, L. *et al.* Performance of the CARABAS-II VHF-band synthetic aperture radar. In: IEEE. **IEEE International Geoscience and Remote Sensing Symposium**. [S.l.], 2001. v. 1, p. 129–131.

62 VU, V. T. *et al.* Bivariate gamma distribution for wavelength-resolution SAR change detection. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, n. 99, p. 1–9, 2018.

63 ULANDER, L. M. *et al.* Performance of VHF-band SAR change detection for wide-area surveillance of concealed ground targets. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Algorithms for Synthetic Aperture Radar Imagery XI**. [S.l.]: SPIE, 2004. v. 5427, p. 259–271.

64 FOLKESSON, K.; SMITH-JONFORSEN, G.; ULANDER, L. M. Model-based compensation of topographic effects for improved stem-volume retrieval from CARABAS-II VHF-band SAR images. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 47, n. 4, p. 1045–1055, 2009.

65 ULANDER, L. *et al.* Mapping of wind-thrown forests using the VHF-band CARABAS-II SAR. In: IEEE. **IEEE International Symposium on Geoscience and Remote Sensing**. [S.l.], 2006. p. 3684–3687.

66 COX, D. R.; SNELL, E. J. A general definition of residuals. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 30, n. 2, p. 248–265, 1968.

67 FIRTH, D. Bias reduction of maximum likelihood estimates. **Biometrika**, Oxford University Press, v. 80, n. 1, p. 27–38, 1993.

68 EFRON, B. Bootstrap methods: Another look at the jackknife. **The Annals of Statistics**, Springer, v. 7, n. 1, p. 1–26, 1979.

69 MATLAB. **Version 9.4 (R2018a)**. Natick, Massachusetts: The MathWorks Inc., 2019.

70 R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2019. Disponível em: <http://www.R-project.org/>.

71 WIESEL, A.; ELDAR, Y. C.; YEREDOR, A. Linear regression with Gaussian model uncertainty: Algorithms and bounds. **IEEE Transactions on Signal Processing**, IEEE, v. 56, n. 6, p. 2194–2205, 2008.

72 WANG, H.; OUCHI, K. Accuracy of the *k*-distribution regression model for forest biomass estimation by high-resolution polarimetric SAR: Comparison of model estimation and field data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 46, n. 4, p. 1058–1064, 2008.

73 MCCULLAGH, P.; NELDER, J. **Generalized linear models**. 2nd. ed. [S.l.]: Chapman and Hall, 1989.

74 AMINZADEH, M. S. Approximate 1-sided tolerance limits for future observations for the Rayleigh distribution, using regression. **IEEE Transactions on Reliability**, IEEE, v. 42, n. 4, p. 625–630, 1993.

75 PAWITAN, Y. **In all likelihood: Statistical modelling and inference using likelihood**. [S.l.]: Oxford, 2001. P. 528.

76 PRESS, W. *et al.* **Numerical recipes in C: The art of scientific computing**. 2. ed. [S.l.]: Cambridge University Press, 1992.

77 MITTELHAMMER, R. C.; JUDGE, G. G.; MILLER, D. J. **Econometric Foundations**. [S.l.]: Cambridge University Press, 2000.

78 DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 5, n. 3, p. 236–244, 1996.

79 NAGELKERKE, N. J. *et al.* A note on a general definition of the coefficient of determination. **Biometrika**, Oxford University Press, v. 78, n. 3, p. 691–692, 1991.

80 MERCIER, G.; MOSER, G.; SERPICO, S. B. Conditional copulas for change detection in heterogeneous remote sensing images. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 46, n. 5, p. 1428–1441, 2008.

81 HOEKMAN, D. H.; QUIRIONES, M. J. Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian Amazon. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 38, n. 2, p. 685–696, 2000.

82 KAY, S. M. **Fundamentals of statistical signal processing: Estimation theory**. Kingston: Prentice Hall PTR, 1993.

83 CRIBARI-NETO, F.; ZEILEIS, A. Beta regression in R. **Journal of Statistical Software**, Foundation for Open Access Statistics, v. 34, n. 2, p. 1–24, 2010.

84 LUNDBERG, M. *et al.* A challenge problem for detection of targets in foliage. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Proceedings in Algorithms for Synthetic Aperture Radar Imagery XIII**. [S.l.], 2006. v. 6237, p. 1–13.

85 SDMS. **Sensor Data Management System public web site**. 2018. Disponível em: <https://www.sdms.afrl.af.mil/index.php>.

86 ZHANG, Z. Parametric regression model for survival data: Weibull regression model as an example. **Annals of Translational Medicine**, AME Publications, v. 4, n. 24, 2016.

87 WANG, H. *et al.* In search of the statistical properties of high-resolution polarimetric SAR data for the measurements of forest biomass beyond the RCS saturation limits. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 3, n. 4, p. 495–499, 2006.

88 STEPHENS, M. A. EDF statistics for goodness of fit and some comparisons. **Journal of the American Statistical Association**, Taylor & Francis, v. 69, n. 347, p. 730–737, 1974.

89 PAPOULIS, A.; PILLAI, S. U. **Probability, random variables, and stochastic processes**. [S.l.]: Tata McGraw-Hill Education, 2002.

90 ELTOFT, T.; HOGDA, K. A. Non-gaussian signal statistics in ocean SAR imagery. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 36, n. 2, p. 562–575, 1998.

91 ACHIM, A.; KURUOGLU, E. E.; ZERUBIA, J. SAR image filtering based on the heavy-tailed Rayleigh model. **IEEE Transactions on Image Processing**, IEEE, v. 15, n. 9, p. 2686–2693, 2006.

92 PENG, Q.; ZHAO, L. SAR image filtering based on the Cauchy–Rayleigh mixture model. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 11, n. 5, p. 960–964, 2013.

93 ALVES, D. I. *et al.* A statistical analysis for wavelength-resolution SAR image stacks. **IEEE Geoscience and Remote Sensing Letters (Early Access)**, IEEE, 2019.

94 PALM, B. G. *et al.* Rayleigh regression model for ground type detection in SAR imagery. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 16, n. 10, p. 1660–1664, 2019.

95 CORDEIRO, G. M.; MCCULLAGH, P. Bias correction in generalized linear models. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 53, n. 3, p. 629–643, 1991.

96 OSPINA, R.; CRIBARI-NETO, F.; VASCONCELLOS, K. L. P. Improved point and intervalar estimation for a beta regression model. **Computational Statistics & Data Analysis**, Elsevier, v. 51, p. 960–981, 2006.

97 VASCONCELLOS, K. L.; CRIBARI-NETO, F. Improved maximum likelihood estimation in a new class of beta regression models. **Brazilian Journal of Probability and Statistics**, JSTOR, p. 13–31, 2005.

98 CORDEIRO, G. M.; BOTTER, D. A. Second-order biases of maximum likelihood estimates in overdispersed generalized linear models. **Statistics & Probability Letters**, Elsevier, v. 55, n. 3, p. 269–280, 2001.

99 BARRETO-SOUZA, W.; VASCONCELLOS, K. L. Bias and skewness in a general extreme-value regression model. **Computational Statistics & Data Analysis**, Elsevier, v. 55, n. 3, p. 1379–1393, 2011.

100 HEINZE, G.; SCHEMPER, M. A solution to the problem of separation in logistic regression. **Statistics in Medicine**, Wiley Online Library, v. 21, n. 16, p. 2409–2419, 2002.

101 BULL, S. B.; MAK, C.; GREENWOOD, C. M. A modified score function estimator for multinomial logistic regression in small samples. **Computational Statistics & Data Analysis**, Elsevier, v. 39, n. 1, p. 57–74, 2002.

102 LEMONTE, A. J. Improved point estimation for the Kumaraswamy distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 81, n. 12, p. 1971–1982, 2011.

103 LEMONTE, A. J.; SIMAS, A. B.; CRIBARI-NETO, F. Bootstrap-based improved estimators for the two-parameter birnbaum–saunders distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 78, n. 1, p. 37–49, 2008.

104 KIM, J. H. Forecasting autoregressive time series with bias-corrected parameter estimators. **International Journal of Forecasting**, Elsevier, v. 19, n. 3, p. 493–502, 2003.

105 DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. [S.l.]: Cambridge University Press, 1997.

106 EFRON, B.; TIBSHIRANI, R. J. **An Introduction to the Bootstrap**. [S.l.]: Monographs on Statistics and Applied Probability 57, 1994.

107 PARK, J.-M.; SONG, W.-J.; PEARLMAN, W. Speckle filtering of SAR images based on adaptive windowing. **IEE Proceedings–Vision, Image and Signal Processing**, IET, v. 146, n. 4, p. 191–197, 1999.

108 LI, G.-T. *et al.* SAR image despeckling using a space-domain filter with alterable window. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 10, n. 2, p. 263–267, 2012.

109 CRIBARI-NETO, F.; VASCONCELLOS, K. L. Nearly unbiased maximum likelihood estimation for the beta distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 72, n. 2, p. 107–118, 2002.

110 BUSTOS, O.; OJEDA, S.; VALLEJOS, R. Spatial ARMA models and its applications to image filtering. **Brazilian Journal of Probability and Statistics**, JSTOR, p. 141–165, 2009.

111 ROSENFELD, A. **Image modeling**. [S.l.]: Academic Press, 2014.

112 NIJIM, Y. W.; STEARNS, S. D.; MIKHAEL, W. B. Lossless compression of images employing a linear IIR model. In: IEEE. **1996 IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World. ISCAS 96**. [S.l.], 1996. v. 2, p. 305–308.

113 CHUNG, Y.-S.; KANEFSKY, M. On 2-D recursive LMS algorithms using ARMA prediction for ADPCM encoding of images. **IEEE Transactions on Image Processing**, IEEE, v. 1, n. 3, p. 416–422, 1992.

114 TEKALP, A.; KAUFMAN, H.; WOODS, J. Identification of image and blur parameters for the restoration of noncausal blurs. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, IEEE, v. 34, n. 4, p. 963–972, 1986.

115 ARGENTI, F. *et al.* Vector median deblurring filter for colour image restoration. **Electronics Letters**, IET, v. 27, n. 21, p. 1899–1900, 1991.

116 BIEMOND, J.; PUTTEN, F. V. D.; WOODS, J. Identification and restoration of images with symmetric noncausal blurs. **IEEE Transactions on Circuits and Systems**, IEEE, v. 35, n. 4, p. 385–393, 1988.

117 KAUFMAN, H.; TEKALP, A. M. Survey of estimation techniques in image restoration. **IEEE Control Systems Magazine**, IEEE, v. 11, n. 1, p. 16–24, 1991.

118 CHAPARRO, L.; JURY, E. Rational approximation of 2-D linear discrete systems. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, IEEE, v. 30, n. 5, p. 780–787, 1982.

119 BOX, G.; JENKINS, G. M.; REINSEL, G. **Time series analysis: Forecasting and control**. [S.l.]: Hardcover, John Wiley & Sons, 2008.

120 TAO, K. Adaptive image smoothing algorithms for edge and texture preservation. In: IEEE. **ICASSP'84. IEEE International Conference on Acoustics, Speech, and Signal Processing**. [S.l.], 1984. v. 9, p. 287–290.

121 HALL, T. E.; GIANNAKIS, G. B. Bispectral analysis and model validation of texture images. **IEEE Transactions on Image Processing**, IEEE, v. 4, n. 7, p. 996–1009, 1995.

122 CADZOW, J.; OGINO, K. Two-dimensional spectral estimation. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, IEEE, v. 29, n. 3, p. 396–401, 1981.

123 ZOUBIR, A. M. *et al.* **Robust statistics for signal processing**. [S.l.]: Cambridge University Press, 2018.

124 ZHAO, B. *et al.* A spatial Gaussian mixture model for optical remote sensing image clustering. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 9, n. 12, p. 5748–5759, 2016.

125 MORALES-ALVAREZ, P. *et al.* Remote sensing image classification with large-scale Gaussian processes. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 56, n. 2, p. 1103–1114, 2017.

126 WILLETT, P.; SWASZEK, P. F.; BLUM, R. S. The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise. **IEEE Transactions on Signal Processing**, IEEE, v. 48, n. 12, p. 3266–3279, 2000.

127 XUE, J.; XU, S.; SHUI, P. Near-optimum coherent CFAR detection of radar targets in compound-Gaussian clutter with inverse Gaussian texture. **Signal Processing**, Elsevier, v. 166, p. 107236, 2020.

128 MARGOOSIAN, A.; ABOUEI, J.; PLATANIOTIS, K. N. An accurate kernelized energy detection in Gaussian and non-Gaussian/impulsive noises. **IEEE Transactions on Signal Processing**, IEEE, v. 63, n. 21, p. 5621–5636, 2015.

129 LI, C. *et al.* PHD and CPHD filtering with unknown detection probability. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 14, p. 3784–3798, 2018.

130 LIU, J. *et al.* Persymmetric adaptive detection of distributed targets in compound-Gaussian sea clutter with Gamma texture. **Signal Processing**, Elsevier, v. 152, p. 340–349, 2018.

131 BAYER, F. M. *et al.* A novel Rayleigh dynamical model for remote sensing data interpretation. **Submitted to IEEE Transactions on Geoscience and Remote Sensing**, 2019.

132 BAYER, F. M.; KOZAKEVICIUS, A. J.; CINTRA, R. J. An iterative wavelet threshold for signal denoising. **Signal Processing**, Elsevier, v. 162, p. 10–20, 2019.

133 BROOKS, E. B. *et al.* On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 52, n. 6, p. 3316–3332, 2013.

134 BAYER, F. M.; BAYER, D. M.; PUMI, G. Kumaraswamy autoregressive moving average models for double bounded environmental data. **Journal of Hydrology**, Elsevier, v. 555, p. 385–396, 2017.

135 BROCKWELL, P. J.; DAVIS, R. A. **Introduction to time series and forecasting**. [S.l.]: Springer, 2016.

136 NOCEDAL, J.; WRIGHT, S. J. **Numerical optimization**. [S.l.]: Springer, 1999.

137 KEDEM, B.; FOKIANOS, K. **Regression models for time series analysis**. [S.l.]: John Wiley & Sons, 2005. v. 488.

138 BROCKWELL, P. J.; DAVIS, R. A. **Time series: Theory and methods**. [S.l.]: Springer Science & Business Media, 2013.

139 GONZALEZ, R.; WOODS, R.; EDDINE, S. **Digital Image Processing using MATLAB**. [S.l.]: Gatesmark publishing, LLC, 2009.

140 FISHMAN, G. **Monte Carlo: Concepts, algorithms, and applications**. [S.l.]: Springer Science & Business Media, 2013.

141 LIM, J. S. Two-dimensional signal and image processing. **Englewood Cliffs, NJ, Prentice Hall, 1990, 710 p.**, 1990.

142 HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, Elsevier, v. 22, n. 4, p. 679–688, 2006.

143 NASCIMENTO, A. D. *et al.* Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 7, n. 2, p. 648–663, 2013.

144 GINI, F.; GRECO, M. Suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 35, n. 3, p. 1095–1104, 1999.

145 WANG, X.; CHEN, R. Adaptive Bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise. **IEEE Transactions on Signal Processing**, IEEE, v. 48, n. 7, p. 2013–2028, 2000.

146 KAY, S. Can detectability be improved by adding noise? **IEEE Signal Processing Letters**, IEEE, v. 7, n. 1, p. 8–10, 2000.

147 CONTE, E.; MAIO, A. D. Exploiting persymmetry for CFAR detection in compound-Gaussian clutter. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 39, n. 2, p. 719–724, 2003.

148 SUN, Y.; WILLETT, P.; SWASZEK, P. F. A non-Gaussian problem that arises in fused detection in clutter. **IEEE Signal Processing Letters**, IEEE, v. 11, n. 2, p. 189–192, 2004.

149 JIN, Y.; FRIEDLANDER, B. A CFAR adaptive subspace detector for second-order Gaussian signals. **IEEE Transactions on Signal Processing**, IEEE, v. 53, n. 3, p. 871–884, 2005.

150 TAN, P. H.; RASMUSSEN, L. K. Asymptotically optimal nonlinear MMSE multiuser detection based on multivariate Gaussian approximation. **IEEE Transactions on Communications**, IEEE, v. 54, n. 8, p. 1427–1428, 2006.

151 NOVEY, M.; ADALI, T.; ROY, A. Circularity and Gaussianity detection using the complex generalized Gaussian distribution. **IEEE Signal Processing Letters**, IEEE, v. 16, n. 11, p. 993–996, 2009.

152 SALA-ALVAREZ, J.; VAZQUEZ-VILAR, G.; LÓPEZ-VALCARCE, R. Multiantenna GLR detection of rank-one signals with known power spectrum in white noise with unknown spatial correlation. **IEEE Transactions on Signal Processing**, IEEE, v. 60, n. 6, p. 3065–3078, 2012.

153 YANG, X. *et al.* CFAR detection of moving range-spread target in white Gaussian noise using waveform contrast. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 13, n. 2, p. 282–286, 2016.

154 PIMENTEL, M. A. *et al.* A review of novelty detection. **Signal Processing**, Elsevier, v. 99, p. 215–249, 2014.

155 WANG, G. *et al.* Noise-robust line detection using normalized and adaptive second-order anisotropic Gaussian kernels. **Signal Processing**, Elsevier, v. 160, p. 252–262, 2019.

156 WANG, G.; ZHU, J.; XU, Z. Asymptotically optimal one-bit quantizer design for weak-signal detection in generalized Gaussian noise and lossy binary communication channel. **Signal Processing**, Elsevier, v. 154, p. 207–216, 2019.

157 DO, M. N.; VETTERLI, M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. **IEEE Transactions on Image Processing**, IEEE, v. 11, n. 2, p. 146–158, 2002.

158 GAZOR, S.; ZHANG, W. Speech probability distribution. **IEEE Signal Processing Letters**, IEEE, v. 10, n. 7, p. 204–207, 2003.

159 NOVEY, M.; ADALI, T.; ROY, A. A complex generalized Gaussian distribution—characterization, generation, and estimation. **IEEE Transactions on Signal Processing**, IEEE, v. 58, n. 3, p. 1427–1433, 2009.

160 BARANIUK, R.; CEVHER, V.; WAKIN, M. B. Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective. **Proceedings of the IEEE**, IEEE, v. 98, n. 6, p. 959–971, 2010.

161 JACQUES, L.; HAMMOND, D. K.; FADILI, J. M. Dequantizing compressed sensing: When oversampling and non-Gaussian constraints combine. **IEEE Transactions on Information Theory**, IEEE, v. 57, n. 1, p. 559–571, 2010.

162 GUO, C.; DAVIES, M. E. Sample distortion for compressed imaging. **IEEE Transactions on Signal Processing**, IEEE, v. 61, n. 24, p. 6431–6442, 2013.

163 POOR, H. V.; THOMAS, J. Locally optimum detection of discrete-time stochastic signals in non-Gaussian noise. **The Journal of the Acoustical Society of America**, ASA, v. 63, n. 1, p. 75–80, 1978.

164 POOR, H. V. Fine quantization in signal detection and estimation. **IEEE Transactions on Information Theory**, IEEE, v. 34, n. 5, p. 960–972, 1988.

165 LI, H. Distributed adaptive quantization and estimation for wireless sensor networks. In: IEEE. **2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07**. [S.l.], 2007. v. 3, p. III–533.

166 RIBEIRO, A.; GIANNAKIS, G. B. Non-parametric distributed quantization-estimation using wireless sensor networks. In: IEEE. **Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.** [S.l.], 2005. v. 4, p. iv–61.

167 PAN, Y. *et al.* Noise enhancement in robust estimation of location. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 8, p. 1953–1966, 2018.

168 ZHU, J. *et al.* Parameter estimation from quantized observations in multiplicative noise environments. **IEEE Transactions on Signal Processing**, IEEE, v. 63, n. 15, p. 4037–4050, 2015.

169 WANG, X. *et al.* Distributed detection of sparse stochastic signals with quantized measurements: The generalized gaussian case. **IEEE Transactions on Signal Processing**, IEEE, v. 67, n. 18, p. 4886–4898, 2019.

170 WANG, X.; LI, G.; VARSHNEY, P. K. Detection of sparse stochastic signals with quantized measurements in sensor networks. **IEEE Transactions on Signal Processing**, IEEE, v. 67, n. 8, p. 2210–2220, 2019.

171 FOSCHINI, G.; GITLIN, R.; WEINSTEIN, S. Optimum detection of quantized PAM data signals. **IEEE Transactions on Communications**, IEEE, v. 24, n. 12, p. 1301–1309, 1976.

172 GANDHI, P. P. Data quantization effects in CFAR signal detection. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 32, n. 4, p. 1277–1289, 1996.

173 ZHANG, J. *et al.* A fundamental limitation on maximum parameter dimension for accurate estimation with quantized data. **IEEE Transactions on Information Theory**, IEEE, v. 64, n. 9, p. 6180–6195, 2018.

174 ZHANG, J. *et al.* Asymptotically optimum distributed estimation in the presence of attacks. **IEEE Transactions on Signal Processing**, IEEE, v. 63, n. 5, p. 1086–1101, 2014.

175 CIUONZO, D.; MAIO, A. D.; ROSSI, P. S. A systematic framework for composite hypothesis testing of independent Bernoulli trials. **IEEE Signal Processing Letters**, IEEE, v. 22, n. 9, p. 1249–1253, 2015.

176 FARIAS, R. C.; BROSSIER, J.-M. Scalar quantization for estimation: From an asymptotic design to a practical solution. **IEEE Transactions on Signal Processing**, IEEE, v. 62, n. 11, p. 2860–2870, 2014.

177 FARIAS, R. C.; MOISAN, E.; BROSSIER, J.-M. Optimal asymmetric binary quantization for estimation under symmetrically distributed noise. **IEEE Signal Processing Letters**, IEEE, v. 21, n. 5, p. 523–526, 2014.

178 ZHU, J. *et al.* Maximum likelihood estimation from sign measurements with sensing matrix perturbation. **IEEE Transactions on Signal Processing**, IEEE, v. 62, n. 15, p. 3741–3753, 2014.

179 WANG, G. *et al.* Signal amplitude estimation and detection from unlabeled binary quantized samples. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 16, p. 4291–4303, 2018.

180 WANG, G. *et al.* Signal amplitude estimation and detection from unlabeled binary quantized samples. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 16, p. 4291–4303, 2018.

181 DUECK, R. K. **Digital design with CPLD applications and VHDL**. [S.l.]: Granta Books, 2001.

182 SCHWARTZ, S. C.; THOMAS, J. B.; WEGMAN, E. J. **Topics in non-Gaussian signal processing**. [S.l.]: Springer, 1989.

183 GRILLENZONI, C. Forecasting unstable and nonstationary time series. **International Journal of Forecasting**, Elsevier, v. 14, n. 4, p. 469–482, 1998.

184 WHITE, H. **Asymptotic theory for econometricians**. [S.l.]: Academic press, 1984.

185 SKELLAM, I. G. A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. **Journal of the Royal Statistical Society. Series B (Methodological)**, Royal Statistical Society, B, n. 10, p. 257–261, 1948.

186 PEARSON, E. S. Bayes' theorem, examined in the light of experimental sampling. **Biometrika**, Oxford University Press, v. 17, n. 3/4, p. 388–442, 1925.

187 FORCINA, A.; FRANCONI, L. Regression analysis with the beta-binomial distribution. **Rivista di Statistica Applicata**, Italian Journal of Applied Statistics, v. 21, n. 1, 1988.

188 BIBBY, B. M.; VÆTH, M. The two-dimensional beta binomial distribution. **Statistics & Probability Letters**, Elsevier, v. 81, n. 7, p. 884–891, 2011.

189 ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions with formulas, graphs, and mathematical tables**. [S.l.]: Dover, New York, 1972. v. 9.

190 BENJAMIN, M. A.; RIGBY, R. A.; STASINOPOULOS, D. M. Generalized autoregressive moving average models. **Journal of the American Statistical Association**, Taylor & Francis, v. 98, n. 461, p. 214–223, 2003.

191 ROCHA, A. V.; CRIBARI-NETO, F. Beta autoregressive moving average models. **Test**, Springer, v. 18, n. 3, p. 529–545, 2009.

192 ROCHA, A. V.; CRIBARI-NETO, F. Erratum to: Beta autoregressive moving average models. **Test**, Springer, v. 26, n. 2, p. 451–459, 2017.

193 KOENKER, R.; YOON, J. Parametric links for binary choice models: A Fisherian-Bayesian colloquy. **Journal of Econometrics**, Elsevier, v. 152, n. 2, p. 120–130, 2009.

194 PALM, B. G.; BAYER, F. M. Bootstrap-based inferential improvements in beta autoregressive moving average model. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 47, n. 4, p. 977–996, 2018.

195 STASINOPOULOS, D. M.; RIGBY, R. A. *et al.* Generalized additive models for location scale and shape (GAMLSS) in R. **Journal of Statistical Software**, Foundation for Open Access Statistics, v. 23, n. 7, p. 1–46, 2007.

196 EFRON, B.; HINKLEY, D. V. Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. **Biometrika**, Oxford University Press, v. 65, n. 3, p. 457–483, 1978.

197 CASELLA, G.; ROBERT, C. P.; WELLS, M. T. Generalized accept-reject sampling schemes. **Lecture Notes-Monograph Series**, JSTOR, p. 342–347, 2004.

198 ANSLEY, C. F.; NEWBOLD, P. Finite sample properties of estimators for autorregressive moving average models. **Journal of Econometrics**, Elsevier, v. 13, p. 159–183, 1980.

199 METZ, C. E. Basic principles of ROC analysis. In: ELSEVIER. **Seminars in Nuclear Medicine**. [S.l.], 1978. v. 8, n. 4, p. 283–298.

200 TISON, C. *et al.* A new statistical model for markovian classification of urban areas in high-resolution SAR images. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 42, n. 10, p. 2046–2057, 2004.

201 SPORTOUCHE, H.; NICOLAS, J.-M.; TUPIN, F. Mimic capacity of fisher and generalized gamma distributions for high-resolution SAR image statistical modeling. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 10, n. 12, p. 5695–5711, 2017.

202 AMIRMAZLAGHANI, M.; AMINDAVAR, H.; MOGHADDAMJOO, A. Speckle suppression in SAR images using the 2-D GARCH model. **IEEE Transactions on Image Processing**, IEEE, v. 18, n. 2, p. 250–259, 2009.

203 BELCHER, D. P.; CANNON, P. S. Ionospheric effects on synthetic aperture radar (SAR) clutter statistics. **IET Radar, Sonar & Navigation**, IET, v. 7, n. 9, p. 1004–1011, 2013.

204 GUDNASON, J.; CUI, J.; BROOKES, M. HRR automatic target recognition from superresolution scattering center features. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 45, n. 4, 2009.

205 ZHENG, Y. *et al.* Using combined difference image and *k*-means clustering for SAR image change detection. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 11, n. 3, p. 691–695, 2014.

206 HELLSTEN, H. *et al.* Development of VHF CARABAS II SAR. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Radar Sensor Technology**. [S.l.]: SPIE, 1996. v. 2747, p. 48–61.

207 MACHADO, R. *et al.* The stability of UWB low-frequency SAR images. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 13, n. 8, p. 1114–1118, 2016.

208 BASELICE, F.; FERRAIOLI, G.; PASCAZIO, V. Markovian change detection of urban areas using very high resolution complex SAR images. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 11, n. 5, p. 995–999, 2013.

209 MONTAZERI, S. *et al.* Three-dimensional deformation monitoring of urban infrastructure by tomographic SAR using multitrack TerraSAR-X data stacks. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 12, p. 6868–6878, 2016.

210 WANG, Y.; ZHU, X. X.; BAMLER, R. An efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 11, n. 7, p. 1250–1254, 2014.

211 PALM, B. G. *et al.* Autoregressive model for multi-pass SAR change detection based on image stacks. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Proceedings–Image and Signal Processing for Remote Sensing XXIV**. [S.l.]: SPIE, 2018. v. 10789.

212 VU, V. T. *et al.* False alarm reduction in wavelength-resolution SAR change detection using adaptive noise canceler. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 55, n. 1, p. 591–599, 2017.

213 GHIRMAI, T. Representing a cascade of complex Gaussian AR models by a single Laplace AR model. **IEEE Signal Processing Letters**, IEEE, v. 22, n. 1, p. 110–114, 2015.

214 LIU, B.; REJU, V. G.; KHONG, A. W. A linear source recovery method for underdetermined mixtures of uncorrelated AR-model signals without sparseness. **IEEE Transactions on Signal Processing**, IEEE, v. 62, n. 19, p. 4947–4958, 2014.

215 BISCAINHO, L. W. AR model estimation from quantized signals. **IEEE Signal Processing Letters**, IEEE, v. 11, n. 2, p. 183–185, 2004.

216 MILENKOVIC, P. Glottal inverse filtering by joint estimation of an AR system with a linear input model. **IEEE transactions on acoustics, speech, and signal processing**, IEEE, v. 34, n. 1, p. 28–42, 1986.

217 MARONNA, R. A. *et al.* **Robust statistics: Theory and methods (with R)**. [S.l.]: Wiley, 2018.

218 HAMPEL, F. R. *et al.* **Robust statistics: The approach based on influence functions**. [S.l.]: John Wiley & Sons, 2011. v. 196.

219 WANG, Z.; ZHANG, D. Progressive switching median filter for the removal of impulse noise from highly corrupted images. **IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing**, IEEE, v. 46, n. 1, p. 78–80, 1999.

220 KIRCHNER, M.; FRIDRICH, J. On detection of median filtering in digital images. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Media forensics and security II**. [S.l.], 2010. v. 7541, p. 754110.

221 CHEN, C.; NI, J.; HUANG, J. Blind detection of median filtering in digital images: A difference domain based approach. **IEEE Transactions on Image Processing**, IEEE, v. 22, n. 12, p. 4699–4710, 2013.

222 ZHANG, Y. *et al.* Revealing the traces of median filtering using high-order local ternary patterns. **IEEE Signal Processing Letters**, IEEE, v. 21, n. 3, p. 275–279, 2014.

223 CHEN, J. *et al.* Median filtering forensics based on convolutional neural networks. **IEEE Signal Processing Letters**, IEEE, v. 22, n. 11, p. 1849–1853, 2015.

224 OTEN, R.; FIGUEIREDO, R. J. de. Adaptive alpha-trimmed mean filters under deviations from assumed noise model. **IEEE Transactions on Image Processing**, IEEE, v. 13, n. 5, p. 627–639, 2004.

225 AHMED, F.; DAS, S. Removal of high-density salt-and-pepper noise in images with an iterative adaptive fuzzy filter using alpha-trimmed mean. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 22, n. 5, p. 1352–1358, 2013.

226 GONZALEZ, R. C.; WOODS, R. **Digital image processing**. [S.l.: s.n.], 2008.

227 AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, IEEE, v. 19, n. 6, p. 716–723, 1974.

228 SCHWARZ, G. *et al.* Estimating the dimension of a model. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978.

229 HANNAN, E. J.; QUINN, B. G. The determination of the order of an autoregression. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, v. 41, n. 2, p. 190–195, 1979.

230 ANDERSON, R. L. Distribution of the serial correlation coefficient. **The Annals of Mathematical Statistics**, JSTOR, v. 13, n. 1, p. 1–13, 1942.

231 BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, American Statistical Association, v. 65, n. 332, p. 1509–1526, 1970.

232 LJUNG, G. M.; BOX, G. E. P. On a measure of a lack of fit in time series models. **Biometrika**, Oxford University Press, v. 65, n. 2, p. 297–303, 1978.

APPENDIX A – Numerical Results for Point Estimation of the Parameters of the Rayleigh Regression Model

In this appendix, we show numerical results for the estimators of the Rayleigh regression model parameters presented in Chapter 2. Table 17 presents the simulation results for point estimation of the parameters of the Rayleigh regression model for the Scenarios 1 and 2 presented in Chapter 2, for $N \in \{16; 25; 49; 250; 500; 1,000\}$. In Table 18, we present the numerical results for the Scenarios 3, 4 and 5. In these scenarios, parameters were adopted as follows: $\beta_1 = 0.5$, $\beta_2 = -0.5$, and $\beta_3 = 0.3$, for Scenario 3, and $\beta_1 = 0.5$, $\beta_2 = 0.15$, and $\beta_3 = 1$, for Scenarios 4 and 5. The covariates were generated from uniform $(0, 1)$, binomial, and Rayleigh distribution, respectively, for $N \in \{25; 250; 1,000\}$.

Table 17 – Results of the Monte Carlo simulation of the point estimation for Scenarios 1 and 2

	Scenario 1			Scenario 2	
Measures	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_1$	$\widehat{\beta}_2$
			$N = 16$		
Mean	1.9562	−1.0065	0.9976	0.4634	0.1553
RB(%)	2.1891	−0.6541	0.2417	7.3184	−3.5448
MSE	0.1579	0.2708	0.2737	0.0808	0.2462
			$N = 25$		
Mean	1.9681	−1.0030	1.0040	0.4810	0.1472
RB(%)	1.5972	−0.3004	−0.4045	3.7913	1.8467
MSE	0.0909	0.1564	0.1533	0.0470	0.1421
			$N = 49$		
Mean	1.9829	−0.9974	1.0010	0.4914	0.1472
RB(%)	0.8563	0.2635	−0.1006	1.7234	1.8394
MSE	0.0403	0.0690	0.0687	0.0222	0.0669
			$N = 250$		
Mean	1.9971	−1.0016	1.0009	0.4984	0.1489
RB(%)	0.1450	−0.1600	−0.0900	0.3200	0.7333
MSE	0.0073	0.0121	0.0126	0.0041	0.0125
			$N = 500$		
Mean	1.9992	−1.0008	0.9995	0.4988	0.1507
RB(%)	0.0400	−0.0800	0.0500	0.2400	−0.4667
MSE	0.0036	0.0062	0.0061	0.0020	0.0060
			$N = 1,000$		
Mean	1.9993	−1.0001	1.0002	0.4995	0.1502
RB(%)	0.0350	−0.0100	−0.0200	0.1000	−0.1333
MSE	0.0017	0.0030	0.0029	0.0010	0.0030

Source: Author (2020)

Table 18 – Results of the Monte Carlo simulation of the point estimation for Scenarios 3, 4 and 5

	Scenario 3			Scenario 4			Scenario 5		
Measures	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$
$N = 25$									
Mean	0.4681	−0.5030	0.3040	0.4861	0.1300	0.9863	0.4866	0.1351	0.9871
RB(%)	6.3888	−0.6008	−1.3483	2.7818	13.3524	1.3658	2.6810	9.9392	1.2895
MSE	0.0909	0.1564	0.1533	0.0211	0.0263	0.0266	0.1103	0.1285	0.1294
$N = 250$									
Mean	0.4971	−0.5016	0.3009	0.4981	0.1487	0.9991	0.4975	0.1486	0.9999
RB(%)	0.5708	−0.3165	−0.2857	0.3897	0.8663	0.0945	0.4929	0.9582	0.0091
MSE	0.0073	0.0121	0.0126	0.0018	0.0018	0.0018	0.0085	0.0097	0.0095
$N = 1{,}000$									
Mean	0.4995	−0.5008	0.2996	0.4997	0.1495	0.9994	0.5001	0.1492	0.9993
RB(%)	0.1015	−0.1603	0.1200	0.0534	0.3562	0.0620	−0.0258	0.5218	0.0716
MSE	0.0021	0.0024	0.0023	0.0004	0.0004	0.0004	0.0021	0.0024	0.0024

Source: Author (2020)

APPENDIX B – Cumulants of Second and Third Order of the Parameters of the Rayleigh Regression Model

In this appendix, we present the cumulants of second and third order used to derive the Cox and Snell's and Firth's corrected estimators. The cumulants are obtained using the chain rule, as

$$\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_i} = \sum_{n=1}^{N} \frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} \frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} \frac{\partial \eta[n]}{\partial \beta_i},$$

for $i = 1, 2, \ldots, k$. From [94], we have

$$\begin{aligned}
\frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]} &= \frac{\pi y[n]^2}{2\mu[n]^3} - \frac{2}{\mu[n]}, \\
\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]} &= \frac{1}{g'(\mu[n])}, \\
\frac{\partial \eta[n]}{\partial \beta_i} &= x_i[n], \\
\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2} &= \frac{2}{\mu[n]^2} - \frac{3\pi y[n]^2}{2\mu[n]^4}, \\
\mathrm{E}\left[\frac{\mathrm{d}^2\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]^2}\right] &= -\frac{4}{\mu[n]^2}.
\end{aligned}$$

Note that

$$\begin{aligned}
\frac{\partial^3 \ell[n](\mu[n])}{\partial \mu[n]^3} &= \frac{\mathrm{d}}{\mathrm{d}\,\mu[n]} \left(\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2}\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\,\mu[n]} \left(\frac{2}{\mu[n]^2} - \frac{3\pi y[n]^2}{2\mu[n]^4}\right) \\
&= \frac{6\pi y[n]^2}{\mu[n]^5} - \frac{4}{\mu[n]^3}.
\end{aligned}$$

Taking the expected value of the derivative above, we have

$$\mathrm{E}\left(\frac{\partial^3 \ell[n](\mu[n])}{\partial \mu[n]^3}\right) = \frac{24}{\mu[n]^3} - \frac{4}{\mu[n]^3} = \frac{20}{\mu[n]^3}.$$

From [94], the second order cumulant is given by

$$\kappa_{rs} = \mathrm{E}\left[\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_r \partial \beta_s}\right] = \sum_{n=1}^{N} \left[-\frac{4}{\mu[n]^2} \left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2 x_s[n] x_r[n]\right].$$

Differentiating the second order cumulant with respect β_u, we can obtain $\kappa_{rs}^{(u)}$ as

$$\begin{aligned}
\frac{\partial \kappa_{rs}}{\partial \beta_u} &= \frac{\mathrm{d}}{\mathrm{d}\,\mu[n]}\left[\sum_{n=1}^{N} -\frac{4}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\right]\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\frac{\partial \eta[n]}{\partial \beta_u}x_s[n]x_r[n] \\
&= \sum_{n=1}^{N}\left[\frac{8}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2 - \frac{4}{\mu[n]^2}\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\right]\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}x_u[n]x_s[n]x_r[n] \\
&= \sum_{n=1}^{N}\left[\frac{8}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2 - \frac{4}{\mu[n]^2}\left(2\times\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right.\right. \\
&\left.\left.\times\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}x_u[n]x_s[n]x_r[n] \\
&= \sum_{n=1}^{N}\left[\frac{8}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{8}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\right. \\
&\left.\times\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]x_u[n]x_s[n]x_r[n].
\end{aligned}$$

The third order derivatives of the log-likelihood function is

$$\begin{aligned}
\frac{\partial^3 \ell(\boldsymbol{\beta})}{\partial \beta_r \partial \beta_s \partial \beta_u} &= \frac{\partial}{\partial \mu[n]}\left[\sum_{n=1}^{N}\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2 + \frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]}\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right] \\
&\times\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\frac{\partial \eta[n]}{\partial \beta_u}x_r[n]x_s[n] \\
&= \sum_{n=1}^{N}\left[\frac{\partial^3 \ell[n](\mu[n])}{\partial \mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2 + \frac{\partial}{\partial \mu[n]}\times\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2} + \frac{\partial^2 \ell[n](\mu[n])}{\partial \mu[n]^2}\right. \\
&\times\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right) + \frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\frac{\partial}{\partial \mu[n]}\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\left.\frac{\mathrm{d}\,\ell[n](\mu[n])}{\mathrm{d}\,\mu[n]}\right] \\
&\times\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}x_u[n]x_r[n]x_s[n].
\end{aligned}$$

Taking the expected value, we obtain the third order cumulant

$$\begin{aligned}
\kappa_{rsu} &= \sum_{n=1}^{N}\left[\frac{20}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 + 2\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\left(-\frac{4}{\mu[n]^2}\right) - \frac{4}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\right. \\
&\left.\times\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]x_u[n]x_r[n]x_s[n] \\
&= \sum_{n=1}^{N}\left[\frac{20}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{12}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]x_u[n]x_r[n]x_s[n].
\end{aligned}$$

From the above expressions, we have that

$$\begin{aligned}
\kappa_{rs}^{(u)} - \frac{1}{2}\kappa_{rsu} &= \sum_{n=1}^{N}\left\{\left[\frac{8}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{8}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]\right. \\
&\left.-\frac{1}{2}\left[\frac{20}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{12}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]\right\}x_u[n]x_s[n]x_r[n] \\
&= \sum_{n=1}^{N}\left[-\frac{2}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{2}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial \mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)\right]x_u[n]x_s[n]x_r[n].
\end{aligned}$$

Now, it is possible to compute the second order biases of the MLEs in the Rayleigh regression model as

$$\sum_{r,s,u} \kappa^{ar}\kappa^{su}\left\{\kappa_{rs}^{(u)} - \frac{1}{2}\kappa_{rsu}\right\} = \sum_{n=1}^{N} w[n]\sum_{r}\kappa^{ar}x_r[n]\sum_{s,u}x_s[n]\kappa^{su}x_u[n],$$

where

$$w[n] = -\frac{2}{\mu[n]^3}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^3 - \frac{2}{\mu[n]^2}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right)^2\frac{\partial}{\partial\mu[n]}\left(\frac{\mathrm{d}\,\mu[n]}{\mathrm{d}\,\eta[n]}\right).$$

In particular, for $g(\mu[n]) = \log(\mu[n])$, we have $w[n] = -4$. Note that

$$\sum_{n=1}^{N} w[n]\sum_{r}\kappa^{ar}x_r[n]\sum_{s,u}x_s[n]\kappa^{su}x_u[n] = e_a^{\top}\mathbf{I}^{-1}(\boldsymbol{\beta})\sum_{n=1}^{N} w[n]x[n]\left(x^{\top}[n]\mathbf{I}^{-1}(\boldsymbol{\beta})x[n]\right),$$

where e_a is defined as the ath column vector of the $k \times k$ identity matrix. Then,

$$\sum_{r,s,u} \kappa^{ar}\kappa^{su}\left\{\kappa_{rs}^{(u)} - \frac{1}{2}\kappa_{rsu}\right\} = e_a^{\top}\mathbf{I}^{-1}(\boldsymbol{\beta})\mathbf{X}^{\top}\mathbf{W}\delta.$$

APPENDIX C – Conditional Fisher Information Matrix for the 2D RARMA Model

In this appendix, we provide the conditional Fisher information matrix of the 2D RARMA model described in Chapter 4, which is given by the expectation of the negative value of the second-order partial derivatives of the log-likelihood function, which is defined as follows

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\gamma})}{\partial\gamma_i\partial\gamma_j} &= \sum_{n=1}^{N}\sum_{m=1}^{M}\frac{\mathrm{d}}{\mathrm{d}\,\mu[n,m]}\left(\frac{\mathrm{d}\,\ell[n,m](\mu[n,m])}{\mathrm{d}\,\mu[n,m]}\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}\right)\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}\frac{\partial\eta[n,m]}{\partial\gamma_j}\frac{\partial\eta[n,m]}{\partial\gamma_i}\\ &= \sum_{n=1}^{N}\sum_{m=1}^{M}\left(\frac{\partial^2\ell[n,m](\mu[n,m])}{\partial\mu[n,m]^2}\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}+\frac{\mathrm{d}\,\ell[n,m](\mu[n])}{\mathrm{d}\,\mu[n,m]}\frac{\partial}{\partial\mu[n,m]}\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}\right)\\ &\cdot\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}\frac{\partial\eta[n,m]}{\partial\gamma_j}\frac{\partial\eta[n,m]}{\partial\gamma_i}.\end{aligned}$$

As shown in [94], we have that

$$\mathrm{E}\left(\mathrm{d}\,\ell[n,m](\mu[n,m])/\mathrm{d}\,\mu[n,m]\mid S[n,m]\right)=0.$$

Thus,

$$\begin{aligned}\mathrm{E}\left(\frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_i\partial\gamma_j}\middle| S[n,m]\right) &= \sum_{n=1}^{N}\sum_{m=1}^{M}\mathrm{E}\left(\frac{\partial^2\ell[n,m](\mu[n,m])}{\partial\mu[n,m]^2}\right)\\ &\cdot\left(\frac{\mathrm{d}\,\mu[n,m]}{\mathrm{d}\,\eta[n,m]}\right)^2\frac{\partial\eta[n,m]}{\partial\gamma_j}\frac{\partial\eta[n,m]}{\partial\gamma_i}.\end{aligned}$$

The derivatives of $\mathrm{d}\,\ell[n,m](\mu[n,m])/\mathrm{d}\,\mu[n,m]$, $\mathrm{d}\,\mu[n,m]/\mathrm{d}\,\eta[n,m]$, and $\partial\eta[n,m]/\partial\boldsymbol{\gamma}$ have been defined in Section 4.2.3. Now, the second derivative of $\mathrm{d}\,\ell[n,m](\mu[n,m])/\mathrm{d}\,\mu[n,m]$ is given by

$$\frac{\partial^2\ell[n,m](\mu[n,m])}{\partial\mu[n,m]^2}=\frac{2}{\mu[n,m]^2}-\frac{3\pi y[n,m]^2}{2\mu[n,m]^4}.$$

Taking the expected value, we have

$$\mathrm{E}\left[\frac{\mathrm{d}^2\,\ell[n,m](\mu[n,m])}{\mathrm{d}\,\mu[n,m]^2}\right]=-\frac{4}{\mu[n,m]^2}.$$

Thus,

$$\mathrm{E}\left[\frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\gamma_i\partial\gamma_j}\right]=\sum_{n=1}^{N}\sum_{m=1}^{M}-\frac{4}{\mu[n,m]^2}\left(\frac{1}{g'(\mu[n,m]}\right)^2\frac{\partial\eta[n,m]}{\partial\gamma_j}\frac{\partial\eta[n,m]}{\partial\gamma_i}.$$

The conditional Fisher information matrix is given by

$$\mathbf{I}(\boldsymbol{\gamma})=-\begin{bmatrix} I_{(\beta,\beta)} & \mathbf{I}_{(\beta,\phi)} & \mathbf{I}_{(\beta,\theta)}\\ \mathbf{I}_{(\phi,\beta)} & \mathbf{I}_{(\phi,\phi)} & \mathbf{I}_{(\phi,\theta)}\\ \mathbf{I}_{(\theta,\beta)} & \mathbf{I}_{(\theta,\phi)} & \mathbf{I}_{(\theta,\theta)}\end{bmatrix},$$

where $I_{(\beta,\beta)} = \mathbf{a}^\top \mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\beta,\phi)} = \mathbf{P}^\top \mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\beta,\theta)} = \mathbf{R}^\top \mathbf{W}\mathbf{a}$, $\mathbf{I}_{(\phi,\beta)} = \mathbf{a}^\top \mathbf{W}\mathbf{P}$, $\mathbf{I}_{(\phi,\phi)} = \mathbf{P}^\top \mathbf{W}\mathbf{P}$, $\mathbf{I}_{(\phi,\theta)} = \mathbf{R}^\top \mathbf{W}\mathbf{P}$, $\mathbf{I}_{(\theta,\alpha)} = \mathbf{a}^\top \mathbf{W}\mathbf{R}$, $\mathbf{I}_{(\theta,\phi)} = \mathbf{P}^\top \mathbf{W}\mathbf{R}$, $\mathbf{I}_{(\theta,\theta)} = \mathbf{R}^\top \mathbf{W}\mathbf{R}$. The matrices $\mathbf{P}$ and $\mathbf{R}$ are of dimensions $(N \cdot M - w) \times (p+1)^2 - 1$, and $(N \cdot M - w) \times (q+1)^2 - 1$, respectively, with (i, j)th elements given by

$$\mathbf{P}[i,j] = \frac{\partial \eta[i+w, j+w]}{\partial \phi_{(i,j)}},$$
$$\mathbf{R}[i,j] = \frac{\partial \eta[i+w, j+w]}{\partial \theta_{(i,j)}}.$$

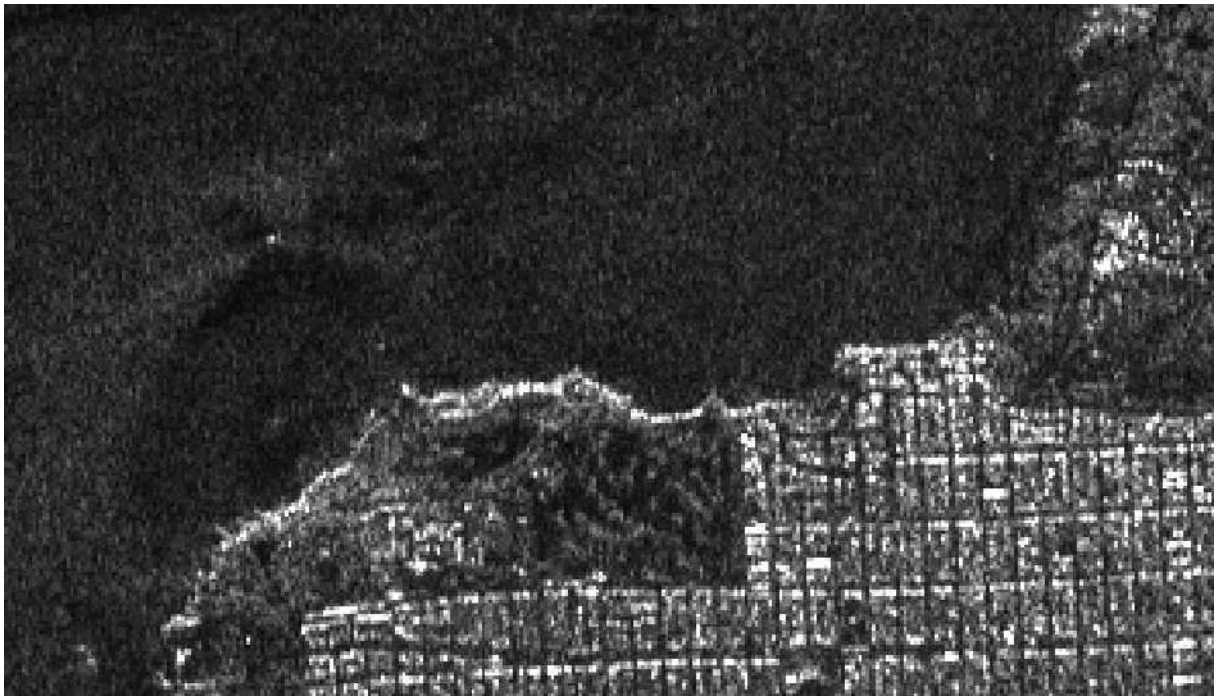
Finally, we have $\mathbf{W} = \text{diag}\left\{ \frac{4}{\mu[1,1]^2}\left(\frac{\mathrm{d}\mu[1,1]}{\mathrm{d}\eta[1,1]}\right)^2, \frac{4}{\mu[1,2]^2}\left(\frac{\mathrm{d}\mu[1,2]}{\mathrm{d}\eta[1,2]}\right)^2, \ldots, \frac{4}{\mu[N,M]^2}\left(\frac{\mathrm{d}\mu[N,M]}{\mathrm{d}\eta[N,M]}\right)^2 \right\}$ and $\mathbf{a} = \left(\frac{\partial \eta[n+1,m+1]}{\partial \beta}, \frac{\partial \eta[n+2,m+2]}{\partial \beta}, \ldots, \frac{\partial \eta[N,M]}{\partial \beta}\right)^\top$.

APPENDIX D – Detection Results Considering the San Francisco SAR Image HV and VV Associated Polarization Channels Based on the 2D ARMA$(1,0)$ Model and the 2D RARMA$(1,0)$ Model

In this appendix, we show the detection results considering the San Francisco SAR image HV and VV associated polarization channels based on 2D ARMA$(1,0)$ and 2D RARMA$(1,0)$ models. Figure 35 presents the San Francisco SAR images HV and VV associated polarization channel. Additionally, Figures 36 and 37 show detection results for HV and VV polarization channels, considering the 2D ARMA$(1,0)$ and 2D RARMA$(1,0)$ models, respectively.

Figure 35 – Original San Francisco SAR image HV and VV associated polarization channels.

(a) HV polarization channel



(b) VV polarization channel

Source: Author (2020)

Figure 36 – Negative detected image HV associated polarization channel based on 2D RARMA$(1,0)$ and 2D ARMA$(1,0)$ models.

(a) 2D RARMA$(1,0)$ model

(b) 2D ARMA$(1,0)$ model

Source: Author (2020)

Figure 37 – Detected image VV associated polarization channel based on 2D RARMA$(1,0)$ and 2D ARMA$(1,0)$ models.

(a) 2D RARMA$(1,0)$ model

(b) 2D ARMA$(1,0)$ model

Source: Author (2020)

APPENDIX E – Tools and Diagnostic Analysis for the BBARMA Model

In this appendix, diagnostic measures and forecasting tools are introduced for the BBARMA model. The data forecasting for the BBARMA(p,q) model can be produced using the CMLE of $\boldsymbol{\gamma}$, $\widehat{\boldsymbol{\gamma}}$, to obtain estimates $\widehat{\mu}[n]$, for $\mu[n]$. The mean response estimate at $N+h$, with $h=1,2,\ldots,H$, where H is the forecast horizon, is given by

$$\widehat{\mu}[N+h]=g^{-1}\left(\widehat{\zeta}+\mathbf{x}^{\top}[n]\widehat{\boldsymbol{\beta}}+\sum_{i=1}^{p}\widehat{\phi}_i\{y^*[N+h-i]\}+\sum_{j=1}^{q}\widehat{\theta}_j\{r^*[N+h-j]\}\right),$$

where

$$y^*[N+h-i]=\begin{cases}\widehat{\mu}[N+h-i], & \text{if } i<h,\\ y^{\star}[N+h-i], & \text{if } i\geq h,\end{cases}$$

$$r^*[N+h-j]=\begin{cases}0, & \text{if } j<h,\\ \widehat{r}[N+h-j], & \text{if } j\geq h,\end{cases}$$

and $\widehat{r}[n]=y^{\star}[n]-\widehat{\mu}[n]$. The quantity $\widehat{\mu}[n]\in(0,1)$ can be mapped to the discrete set $\{0,1,2,\ldots,K\}$ by means of round$(\widehat{\mu}[n]\cdot K)$, where round$(\cdot)$ is a round function. Note that no parameter restrictions are required for fitting or forecasting based on BBARMA model.

The diagnostic measures are useful to evaluate the performance of the adjusted model. The correct adjustment of the proposed model is important to obtain accurate out-of-signal forecasting. For the BBARMA model selections, we adopted the following information criteria: Akaike's (AIC) [227], Schwartz's (SIC) [228], and Hannan and Quinn's (HQ) [229]. Residuals are useful for performing the diagnostic analysis of the fitted model and can be defined as a function of the observed and predicted values of the model [137].

Different types of residuals are considered in literature for several classes of models, such as ordinary residuals, standardized residuals and some residuals in the predictor scale. We employed the standardized ordinary residuals

$$\varepsilon[n]=\frac{y^{\star}[n]-\widehat{\mu}[n]}{\sqrt{\widehat{\mathrm{Var}}(y[n])}},$$

where $\widehat{\mathrm{Var}}(y[n])=K\widehat{\mu}[n](1-\widehat{\mu}[n])\left[\frac{K+\widehat{\varphi}}{1+\widehat{\varphi}}\right]$.

A good model adjustment is indicate by zero mean and constant variance of the standardized residual [137]. Also, it is expected that the autocorrelation and partial autocorrelation

and conditional heteroscedasticity in the series of residuals are absent [119]. The residual autocorrelation function (ACF) is given by

$$\widehat{\rho}_k = \frac{\sum_{n=m+1}^{N-k} (\varepsilon[n] - \bar{\varepsilon})(\varepsilon[n+k] - \bar{\varepsilon})}{\sum_{n=m+1}^{N-k} (\varepsilon[n] - \bar{\varepsilon})^2}, \quad k = 0, 1, \ldots,$$

where $\bar{\varepsilon} = (N-m)^{-1} \sum_{n=m+1}^{N} \varepsilon[n]$. The distribution of $\widehat{\rho}_k$ is approximately normal with zero mean and variance $1/(N-m)$, for $i > 1$ and $N \to \infty$ [119, 137, 230]. Box-Pierce [231], Ljung-Box [232], and the ACF plot are useful to verify if autocorrelation and conditional heteroscedasticity in the series of residuals are absent.

APPENDIX F – Conditional Observed Information Matrix for the BBARMA Model

In this appendix, we provide mathematical details for the score vector, (5.5), and (5.6) presented in Chapter 5. The derivative of ℓ with respect to ζ is given by

$$\frac{\partial \ell}{\partial \zeta} = \varphi \sum_{n=m+1}^{N} \Upsilon[n] \frac{1}{g'(\mu[n])} \frac{\partial \eta[n]}{\partial \zeta}.$$

For the derivative of ℓ with respect to β_k, considering $k = 1, 2, \ldots, l$, we have that

$$\frac{\partial \ell}{\partial \beta_k} = \varphi \sum_{n=m+1}^{N} \Upsilon[n] \frac{1}{g'(\mu[n])} \frac{\partial \eta[n]}{\partial \beta_k}.$$

The score function for the parameter $\boldsymbol{\phi}$, given by the derivative of ℓ with respect to ϕ_i, for $i = 1, 2, \ldots, p$, is given by

$$\frac{\partial \ell}{\partial \phi_i} = \varphi \sum_{n=m+1}^{N} \Upsilon[n] \frac{1}{g'(\mu[n])} \frac{\partial \eta[n]}{\partial \phi_i}.$$

The derivative of ℓ with respect to θ_j, for $j = 1, 2, \ldots, q$, is given by

$$\frac{\partial \ell}{\partial \theta_j} = \varphi \sum_{n=m+1}^{N} \Upsilon[n] \frac{1}{g'(\mu[n])} \frac{\partial \eta[n]}{\partial \theta_j}.$$

In matrix form, the score vector can be written as $\mathbf{U}(\boldsymbol{\gamma}) = (U_\zeta, \mathbf{U}_{\boldsymbol{\beta}}, \mathbf{U}_{\boldsymbol{\phi}}, \mathbf{U}_{\boldsymbol{\theta}}, U_\varphi)^\top$, where

$$\begin{aligned}
U_\zeta &= \varphi \mathbf{a}^\top \mathbf{T} \boldsymbol{\Upsilon}, \\
\mathbf{U}_{\boldsymbol{\beta}} &= \varphi \mathbf{M}^\top \mathbf{T} \boldsymbol{\Upsilon}, \\
\mathbf{U}_{\boldsymbol{\phi}} &= \varphi \mathbf{P}^\top \mathbf{T} \boldsymbol{\Upsilon}, \\
\mathbf{U}_{\boldsymbol{\theta}} &= \varphi \mathbf{R}^\top \mathbf{T} \boldsymbol{\Upsilon}, \\
U_\varphi &= \frac{\partial \ell}{\partial \varphi}.
\end{aligned}$$

From (5.5), we have that

$$\begin{aligned}
\frac{\partial^2 \ell}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\delta}} &= \sum_{n=m+1}^{N} \frac{\partial}{\partial \boldsymbol{\lambda}} \left(\frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \right) \\
&= \sum_{n=m+1}^{N} \left[\frac{\partial^2 \ell[n](\mu[n], \varphi)}{\partial \mu[n]^2} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} + \frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\partial}{\partial \boldsymbol{\lambda}} \left(\frac{d\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \right) \right] \\
&= \sum_{n=m+1}^{N} \left[\frac{\partial^2 \ell[n](\mu[n], \varphi)}{\partial \mu[n]^2} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} + \frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\mathrm{d}^2 \mu[n]}{\mathrm{d}\eta[n]^2} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} \right. \\
&\quad \left. + \frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial^2 \eta[n]}{\partial \boldsymbol{\delta} \partial \boldsymbol{\lambda}} \right] \\
&= \sum_{n=m+1}^{N} \left[\frac{\partial^2 \ell[n](\mu[n], \varphi)}{\partial \mu[n]^2} \left(\frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \right)^2 \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} + \frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\mathrm{d}^2 \mu[n]}{\mathrm{d}\eta[n]^2} \frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \frac{\partial \eta[n]}{\partial \boldsymbol{\lambda}} \right. \\
&\quad \left. + \frac{\partial \ell[n](\mu[n], \varphi)}{\partial \mu[n]} \frac{\mathrm{d}\mu[n]}{\mathrm{d}\eta[n]} \frac{\partial^2 \eta[n]}{\partial \boldsymbol{\delta} \partial \boldsymbol{\lambda}} \right].
\end{aligned}$$

Additionally, from (5.6), note that

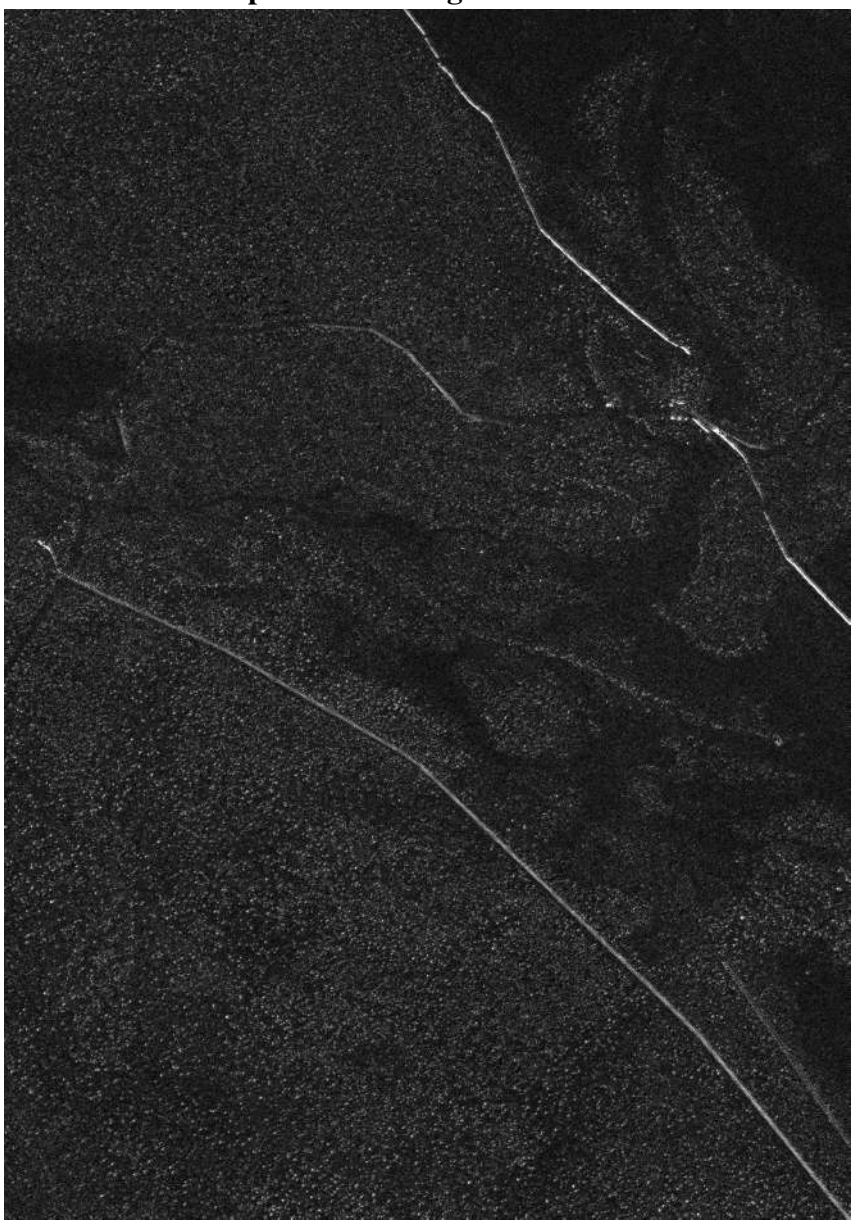
$$\begin{aligned}
\frac{\partial^2 \eta[n]}{\partial \boldsymbol{\delta} \partial \theta_j} &= \frac{\partial}{\partial \theta_j} \left(\frac{\partial \eta[n]}{\partial \boldsymbol{\delta}} \right) = -\frac{\mathrm{d}\mu[n-j]}{\mathrm{d}\eta[n-j]} \frac{\partial \eta[n-j]}{\partial \boldsymbol{\delta}} - \sum_{s=1}^{q} \theta_s \frac{\partial}{\partial \theta_j} \left(\frac{\mathrm{d}\mu[n-s]}{\mathrm{d}\eta[n-s]} \frac{\partial \eta[n-s]}{\partial \boldsymbol{\delta}} \right) \\
&= -\frac{\mathrm{d}\mu[n-j]}{\mathrm{d}\eta[n-j]} \frac{\partial \eta[n-j]}{\partial \boldsymbol{\delta}} - \sum_{s=1}^{q} \theta_s \frac{\partial \eta[n-s]}{\partial \boldsymbol{\delta}} \frac{\mathrm{d}^2 \mu[n-s]}{\mathrm{d}\eta[n-s]^2} \frac{\partial \eta[n-s]}{\partial \theta_j} \\
&\quad - \sum_{s=1}^{q} \theta_s \frac{\mathrm{d}\mu[n-s]}{\mathrm{d}\eta[n-s]} \frac{\partial^2 \eta[n-s]}{\partial \boldsymbol{\delta} \partial \theta_j}, \\
\frac{\partial^2 \eta[n]}{\partial \theta_k \partial \theta_j} &= \frac{\partial}{\partial \theta_j} \left(\frac{\partial \eta[n]}{\partial \theta_k} \right) = -\frac{\mathrm{d}\mu[n-j]}{\mathrm{d}\eta[n-j]} \frac{\partial \eta[n-j]}{\partial \theta_k} - \frac{\mathrm{d}\mu[n-k]}{\mathrm{d}\eta[n-k]} \frac{\partial \eta[n-k]}{\partial \theta_j} \\
&\quad - \sum_{s=1}^{q} \theta_s \frac{\partial}{\partial \theta_j} \left(\frac{\mathrm{d}\mu[n-s]}{\mathrm{d}\eta[n-s]} \times \frac{\partial \eta[n-s]}{\partial \theta_k} \right) \\
&= -\frac{\mathrm{d}\mu[n-j]}{\mathrm{d}\eta[n-j]} \frac{\partial \eta[n-j]}{\partial \theta_k} - \frac{\mathrm{d}\mu[n-k]}{\mathrm{d}\eta[n-k]} \frac{\partial \eta[n-k]}{\partial \theta_j} - \sum_{s=1}^{q} \theta_s \frac{\partial \eta[n-s]}{\partial \theta_j} \frac{\mathrm{d}^2 \mu[n-s]}{\mathrm{d}\eta[n-s]^2} \frac{\partial \eta[n-s]}{\partial \theta_k} \\
&\quad - \sum_{s=1}^{q} \theta_s \frac{\mathrm{d}\mu[n-s]}{\mathrm{d}\eta[n-s]} \frac{\partial^2 \eta[n-s]}{\partial \theta_k \partial \theta_j},
\end{aligned}$$

for $j = k = 1, 2, \ldots, q$.

APPENDIX G – Ground Scene Prediction Images from Stack 2 and 3

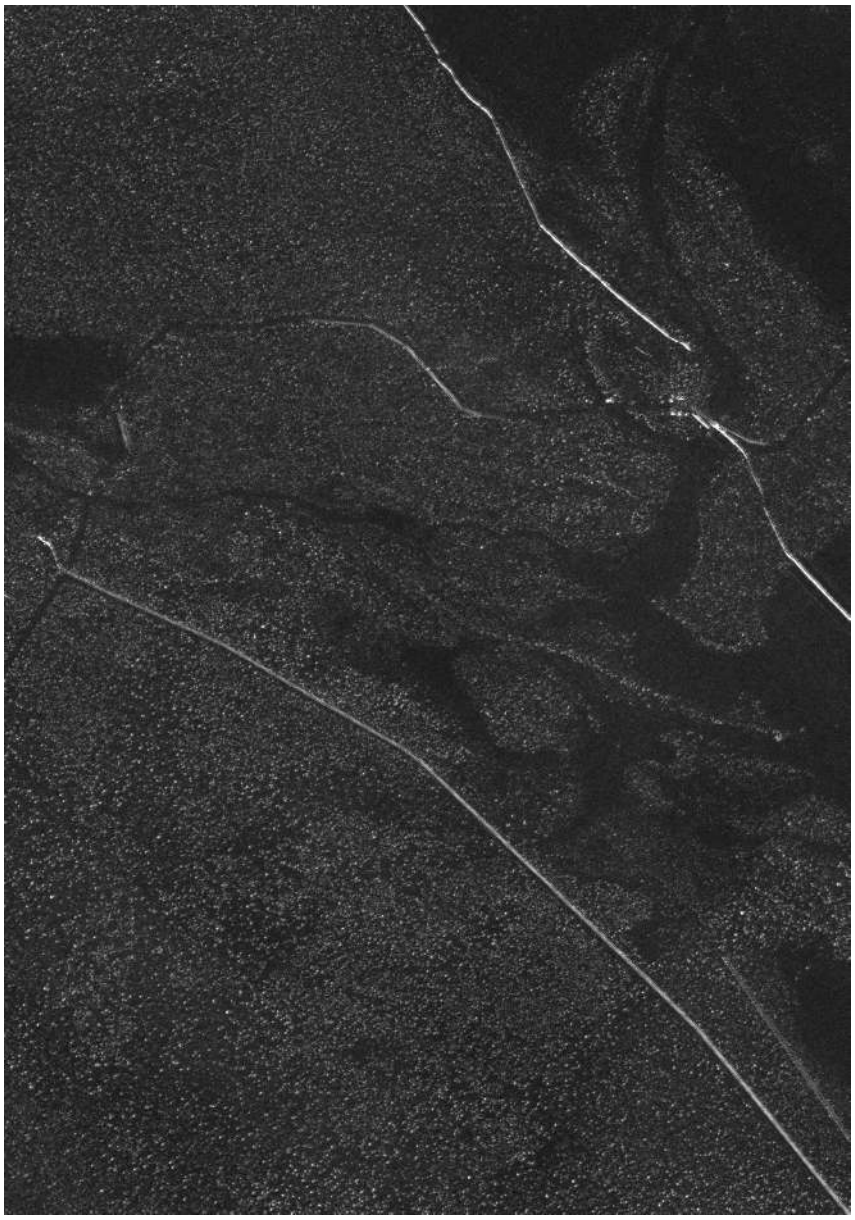
In this appendix, we present the ground scene prediction images from Stack 2 and 3, considering the methods discussed in Chapter 6. Figures 38-42 and 43-47 show the ground scene prediction images from Stack 2 and 3, respectively.

Figure 38 – Ground scene prediction image for Stack 2 based on the AR model.



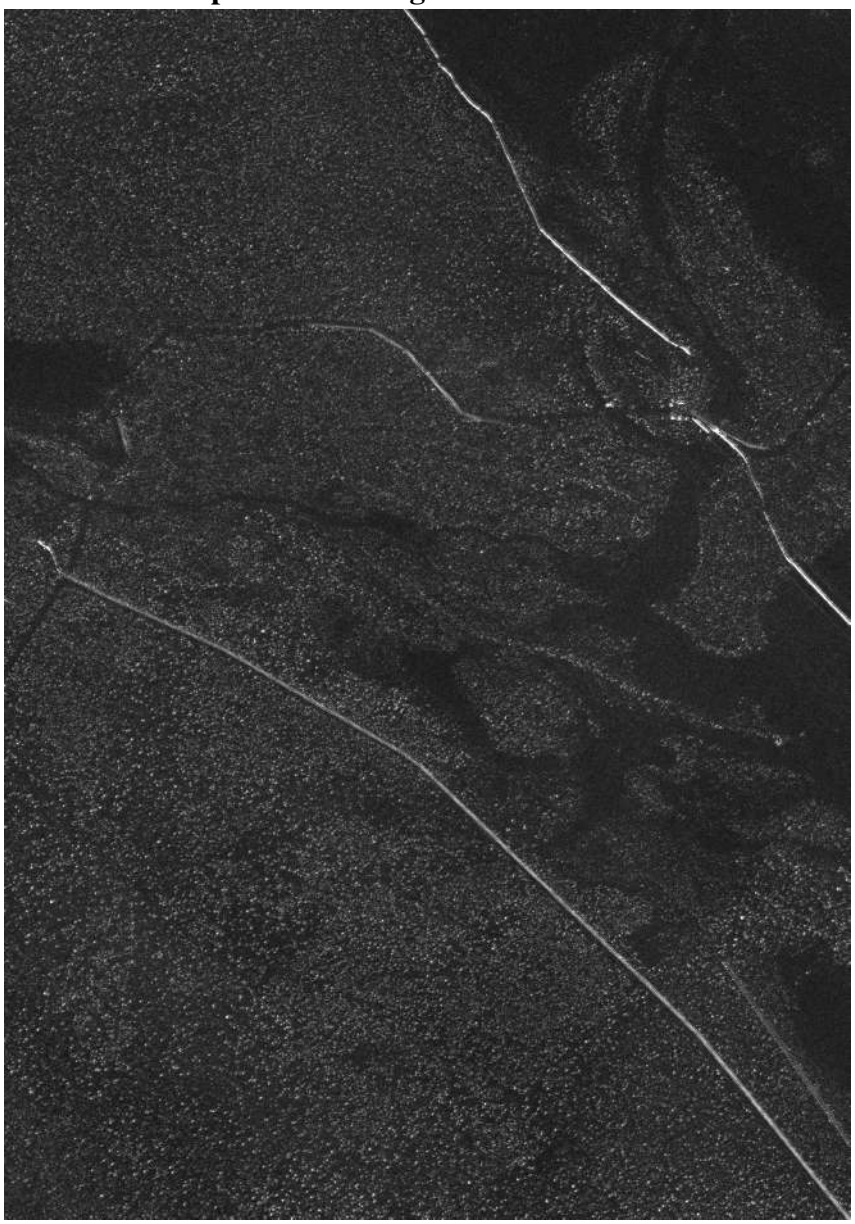
Source: Author (2020)

Figure 39 – Ground scene prediction image for Stack 2 based on the trimmed mean method.



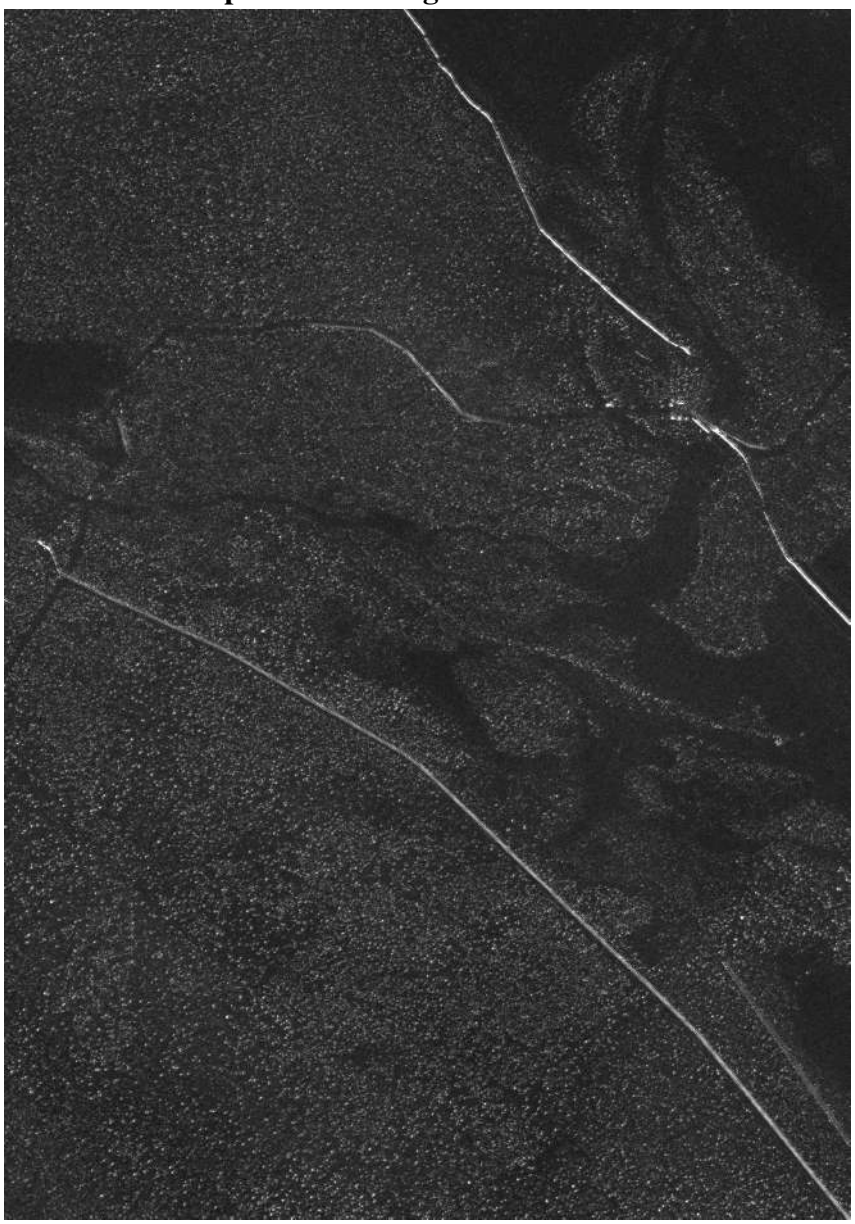
Source: Author (2020)

Figure 40 – Ground scene prediction image for Stack 2 based on the median method.



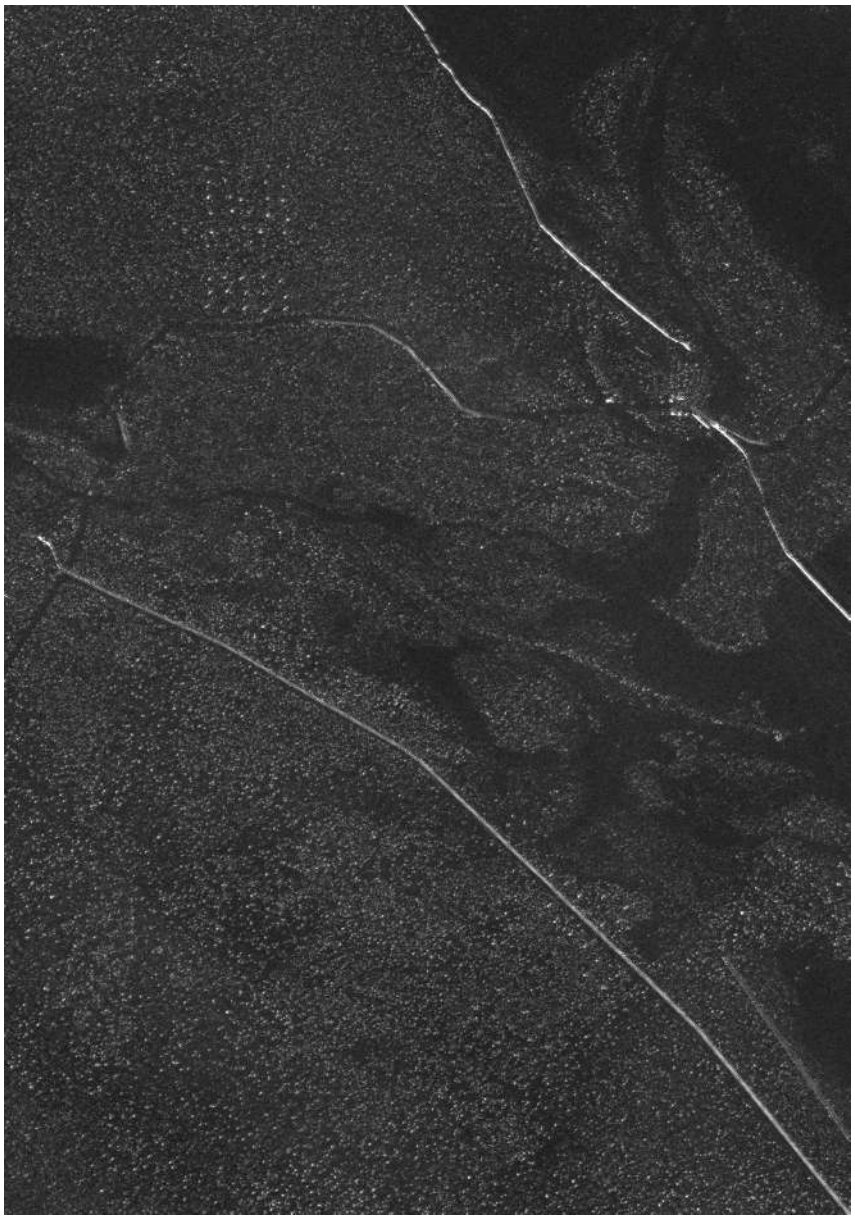
Source: Author (2020)

Figure 41 – Ground scene prediction image for Stack 2 based on the mean method.



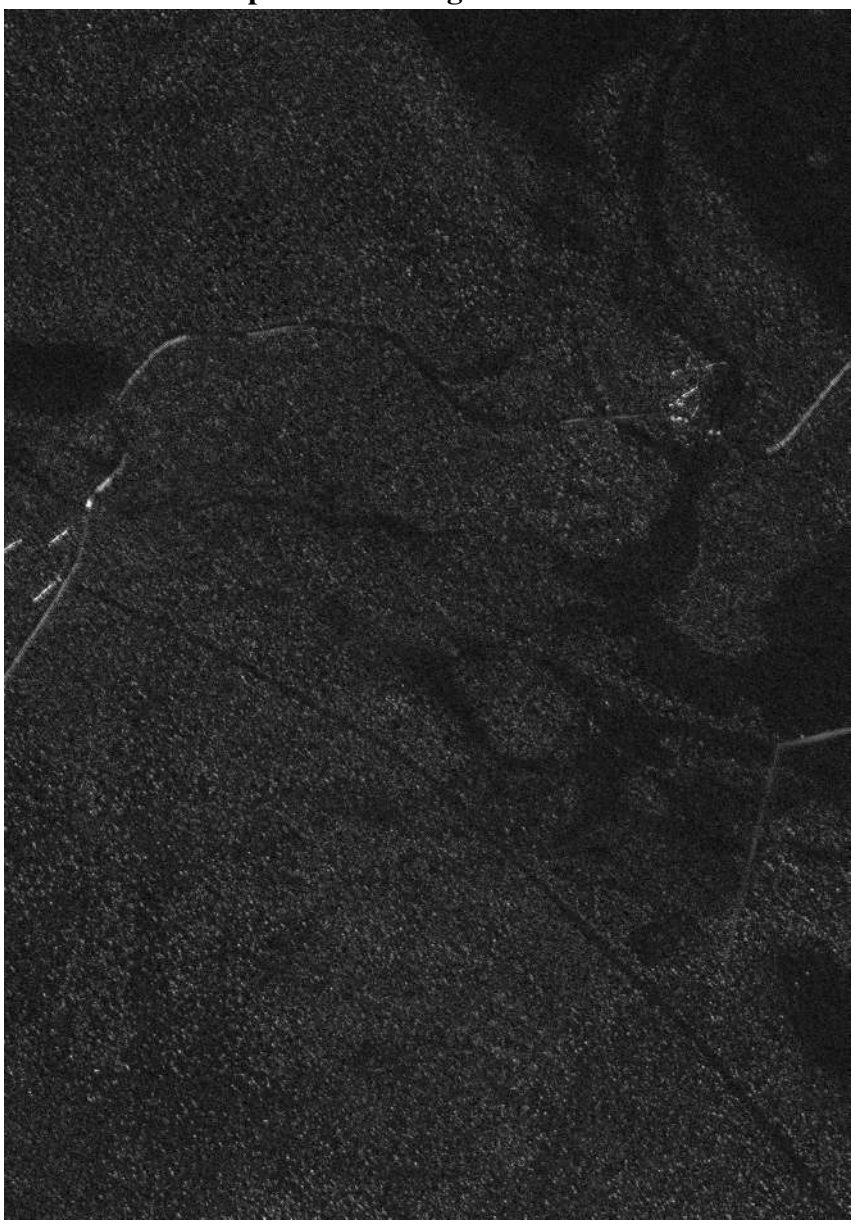
Source: Author (2020)

Figure 42 – Ground scene prediction image for Stack 2 based on the intensity mean method.



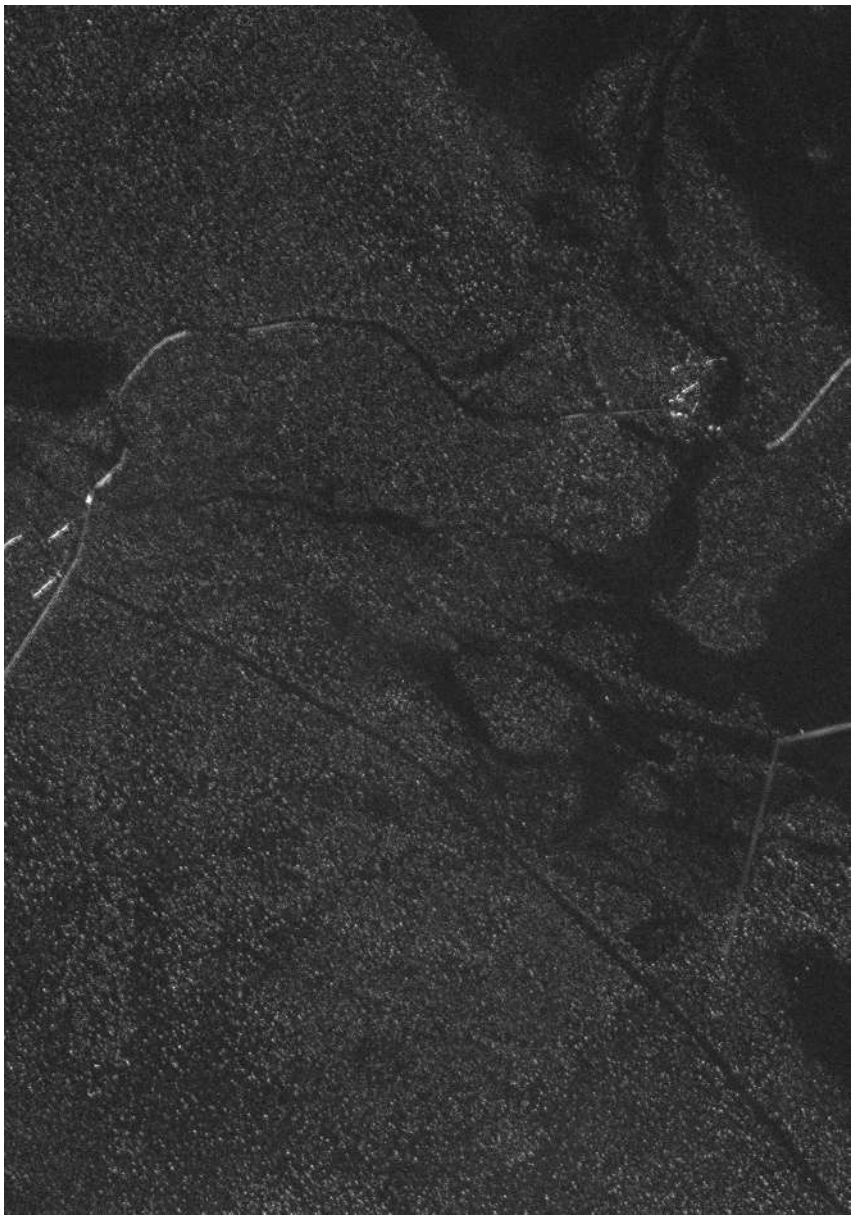
Source: Author (2020)

Figure 43 – Ground scene prediction image for Stack 3 based on the AR model.



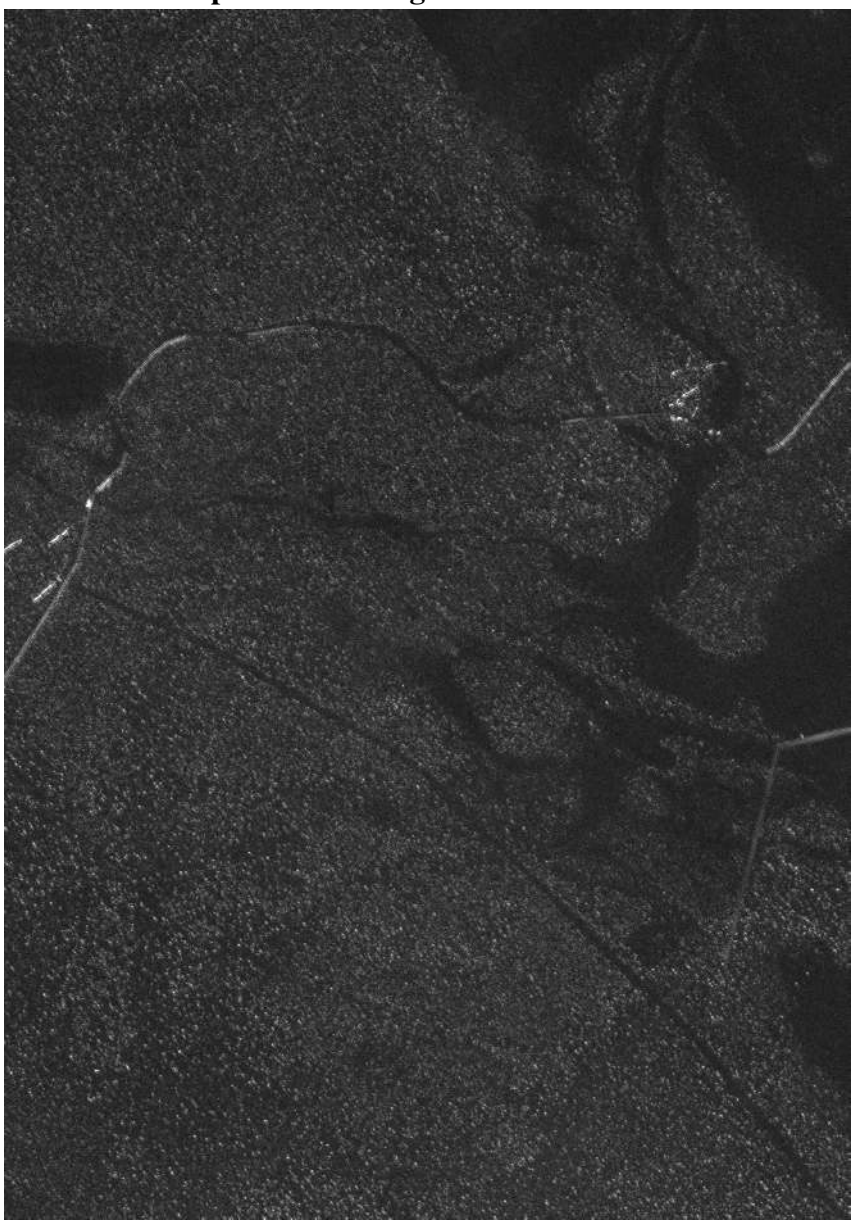
Source: Author (2020)

Figure 44 – Ground scene prediction image for Stack 3 based on the trimmed mean method.



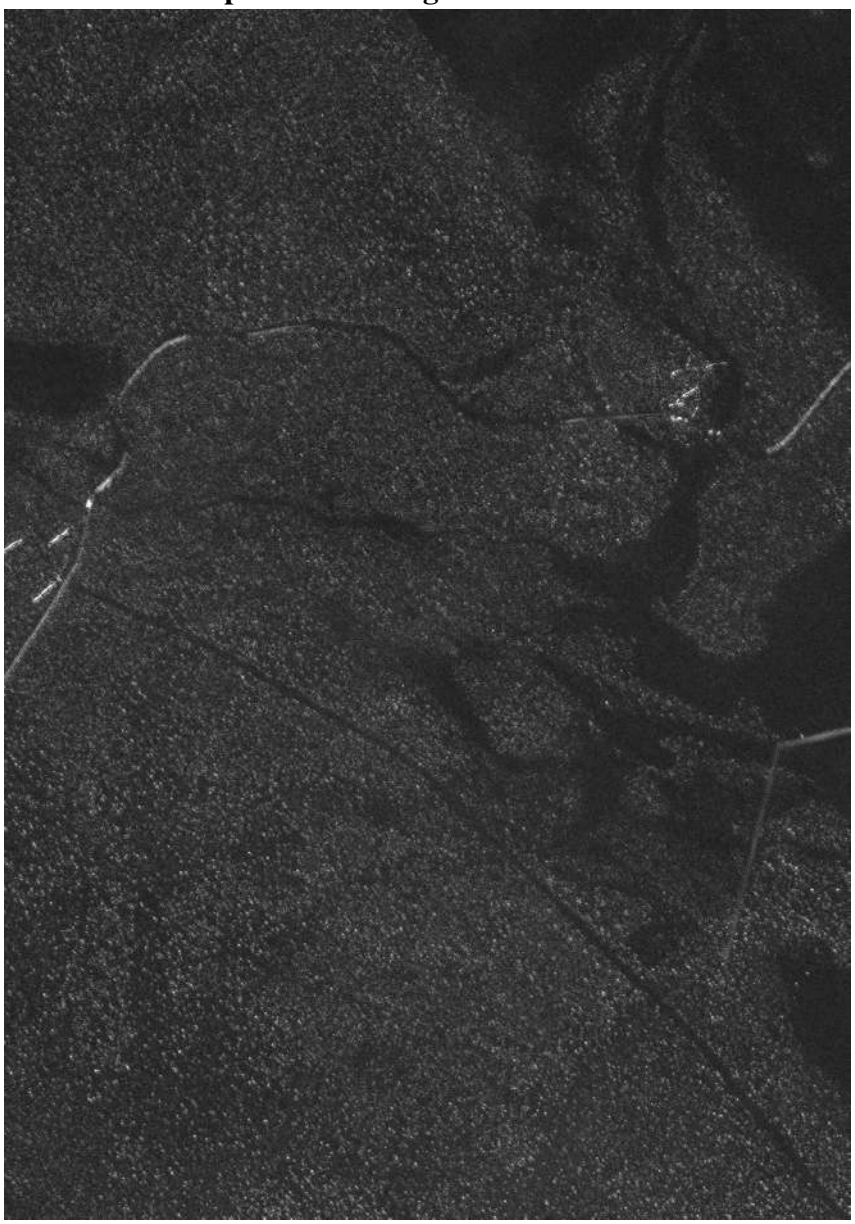
Source: Author (2020)

Figure 45 – Ground scene prediction image for Stack 3 based on the median method.



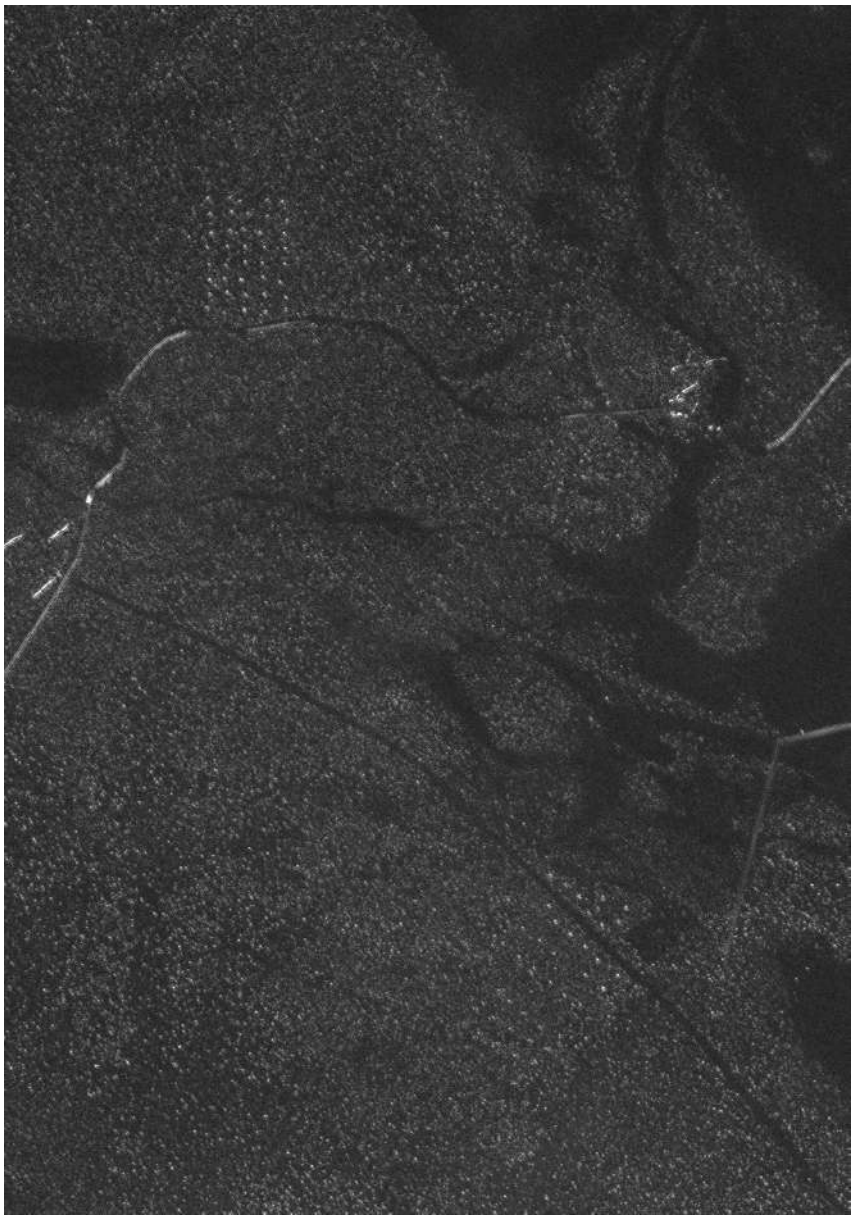
Source: Author (2020)

Figure 46 – Ground scene prediction image for Stack 3 based on the mean method.



Source: Author (2020)

Figure 47 – Ground scene prediction image for Stack 3 based on the intensity mean method.



Source: Author (2020)