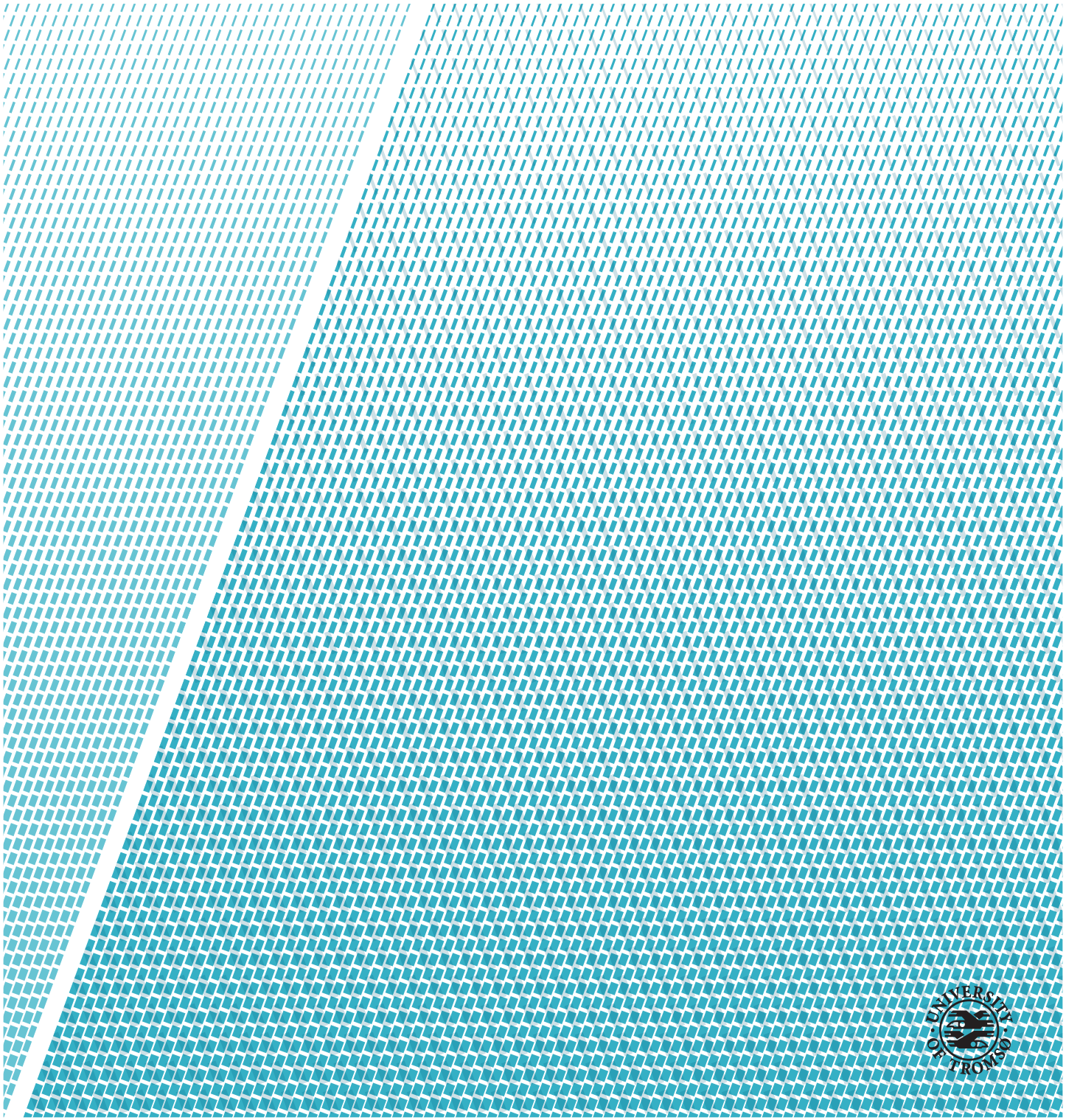
UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Science and Technology
Department of Physics and Technology

Evaluation of Speckle Variance, Higher Moments and Their Dependence on Incidence Angle

Jon Anders Hallaråker

Master thesis in EOM-3901 Master thesis in energy, climate and environment, Desember 2018



“Somewhere, something incredible is waiting to be known.”
–Carl Sagan

Abstract

This thesis investigates speckle properties from Sentinel-1 GRDM data to explore its dependency of incidence angle. The data is processed using the SNAP-tool such that a uniform region can be selected and estimation of speckle variance, skewness and kurtosis can be carried out. These speckle properties are calculated over ocean and sea ice for comparison and modeled to fit a regression line. Hypothesis testing is carried out to determine the statistical significance. Due to the GRDM product having different number of multi-looks depending on the swath, a normalization is done to fit the regression line.

It is shown that there is a slight decrease in variance over incidence angle with statistical significance for both ocean and sea ice. The response is however low. It is further shown that the trend in texture measurements such as skewness and kurtosis have a lower statistical significance and may be deemed to have little to no variation over incidence angle.

Acknowledgements

I would like to thank my supervisors Andrea Marinoni and Anthony Doulgeris for their assistance and guidance throughout this project. I would also like to thank my familiy for their support through these stressful times.

Contents

List of Figures

List of Tables

My list of definitions

SAR - Synthetic Aperture Radar
CIRFA - Center for Integrated
Remote sensing and Forecasting for Arctic Operations
UHF - Ultra High Frequency
EM - Electro Magnetic
ESA - European Space Agency
EW - Extra Wide (swwath width)
ID - Identifier
LUT - Look Up Tables
VV - Vertical Vertical
(Polarization - Transmitted and Received)
HV - Horizontal Vertical
(Polarization - Transmitted and Received)
RCS - Radar Cross Section
SNAP - Sentinel Application Platform
NRCS - Normalized Radar Cross Section
NESZ - Noise Equivalent Sigma Zero
GRD - Ground Range Detected
GRDM - Ground Range Detected Medium Resolution

1 Introduction

Synthetic aperture radar (SAR) has the ability to provide a continuous stream of information. In the arctic region, SAR has the ability to operate even with the challenges that this area has with harsh weather conditions and the polar nights with absence of sunlight during the winter. This widely used tool is however, not without issues. A very well known error is the intensity decay which occurs over the incidence angle [13]. Clustering based on intensity will therefore contain banding regions due to the decrease of intensity. The mean of the intensity tend to an exponential function and can be modeled to be linear in the log domain. This study aim to investigate the variance, skewness and kurtosis and its dependence on incidence angle, to see if modeling this can further improve classification in SAR scenes. This is done by investigating the speckle over incidence angle in uniform areas in different window-sizes and apply linear regression models to information gathered.

1.1 Motivation

The variance and higher moments of speckle have not to the authors knowledge been investigated over incidence angles. This thesis therefore aims to investigate speckle as a function of incidence angle to be used for modeling. Firstly, speckle properties will be explored over the whole scene and investigated over the different classes to see if these properties are dependent on the incidence angle and its response within each class. In the first simple log-normal case,

where the mean is assumed log-linear, the first step of the investigation will be the dependency of the variance. If it can be determined to be a correlation between the incidence angle and variance, further analysis will be done to find a model which later can be implemented into the algorithm provided by the Center of Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA). However, should it be the case where variance does not seem to be related, there are several other features that may be explored. The distribution for the log-intensity along an incident angle is assumed to follow a Gaussian distribution in [6]. However, this may not be the case for the speckle properties and will be further investigated. Different moments, such as skewness and texture may then be explored to be able to define a distribution applicable for modelling.

1.2 Objectives

As the goal of this thesis is to investigate the normalized central moments with respect to incidence angle, the objectives can therefore be listed as:

- Investigate the dependency of incident angle on speckle properties
- Develop a model of the different moments as a function of the incidence angle
- Implement the model into the algorithm provided by CIRFA if a strong response is found.

1.3 Contribution to the Field

This study provides an investigation of speckle over incidence angle as to improve classification. To the author's knowledge, earlier investigations have only considered the simplest case with Gaussian distribution and mean intensity value in order to prove the concept, while variance and higher moments have been left unexplored.

2

Basic Theory

2.1 Radar

One of the most commonly used sections of the electromagnetic spectrum that is utilized for remote sensing is the microwave region; which as mentioned extend from wavelengths of about 1 [mm] to 1 [m]. Both active and passive sensors are used to capture this region of electromagnetic waves, although the primary sensor is the active one.

Active Microwave Sensors are commonly referred to as radar devices. These devices transmit a microwave signal, and capture its reflection and use that to transform it to images of the Earth's surface.

2.1.1 Imaging Geometry

The configuration of an imaging radar system is shown in Figure 2.1. The antenna is pointed to either side of the nadir as to eliminate right-left ambiguities from two symmetric equidistant points. As the imaging system moves along in azimuth direction a continuous swath width [SW] is mapped and is given by:

$$SW \approx \frac{h\beta}{\cos^2 \psi} = \frac{\lambda h}{W \cos^2 \psi} \tag{2.1}$$

where β is the antenna beam width in elevation, W is the antenna width and

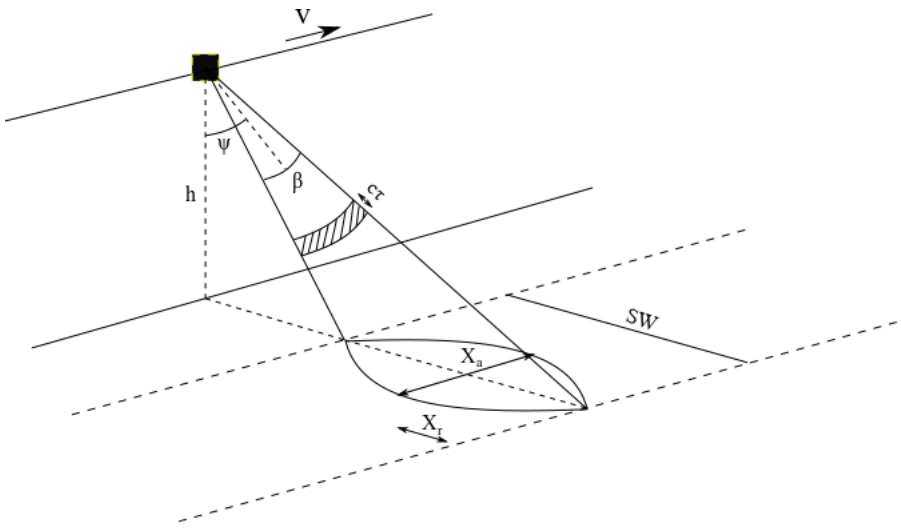


Figure 2.1: Imaging Geometry

ψ is the look angle. An assumption is made that $\beta \ll 1$ and does not consider the earth curvature.

2.1.2 Range Resolution

The range resolution is defined by the minimum distance that two points on a surface are still separable. The echos of the points will be separated in time given by:

$$\Delta t = \frac{2X_r}{c} \sin\psi \tag{2.2}$$

where X_r is the distance between the points, θ is the look angle and c is the speed of light. To find the absolute minimum distance, would mean to find the smallest time difference which is $\tau = 1/B$, where B is the band width

$$X_r = \frac{c\tau}{2\sin\psi} = \frac{c}{2B\sin\psi} \tag{2.3}$$

θ is the look angle and c is the speed of light. This is known as the ground range resolution. The resolution of two targets along the direct line of sight of the radar is called the slant range resolution of the radar and only differ from the ground range by the $\sin\psi$ term.

2.1.3 Azimuth Resolution

As in the range direction the azimuth resolution is defined as the two nearest separable points along the an azimuth line. This is equal to the width of the

antenna footprint as the echoes from all the points along a line spanning that width return at the same time. This is described by:

$$X_a = \frac{h\beta'}{\cos\psi} = \frac{h\lambda}{L\cos\psi} \tag{2.4}$$

where β' is the antenna beam width in the azimuth, λ is the wavelength and L is the antenna length.

2.1.4 Radar Equation

The power returned to the antenna is given by the radar equation:

$$P_r = \frac{\sigma G^2 P_t \lambda^2}{(4\pi)^3 R^4} \tag{2.5}$$

Where P_r is the power returned to the antenna, σ is the backscattering coefficient, G is the gain, P_t is the power transmitted and λ is the wavelength. σ is defined by the specific characteristics of the terrain surface. For image interpretation this parameter carries information about the landscape. It conveys information of the amount of energy scattered from a specific region in on the landscape and is measured by σ^0 and is called the radar cross section.

2.2 Synthetic Aperture Radar

Synthetic aperture radar is an extension of regular radar. It uses the theory that an object will remain within the beam over a longer interval of time as the system moves along the azimuth direction. By later reconstruction we can synthesize an artificially long antenna. Therefore, the azimuth resolution for a SAR imaging system is described by:

$$X_a = \frac{L}{2} \tag{2.6}$$

where L is the real antenna length.

2.3 Frequency

SAR operate in the microwave region, however, there are several sections: P, UHF, L, S, C, X, Ku, K and Ka band. They range from 0.3 GHz and up to 40

Band	Wavelengths
P	107-77 cm
UHF	100-30 cm
L	30-15 cm
S	15-7.5 cm
C	7.5-3.75 cm
X	3.75-2.40 cm
Ku	2.40-1.67 cm
K	1.67-1.18 cm
Ka	1.18-0.75 cm

Table 2.1: Frequency bands for SAR-remote sensing

GHz which is shown in Table 2.1. Depending on band used, they will interact with matter differently. This is of course due to the relative wavelength to the size of particles. Longer wavelengths which correspond to lower frequency, will have the ability to propagate further in mediums such as the atmosphere. The penetration depth is given by

$$e_p = \frac{\lambda\sqrt{\epsilon'}}{2\pi\epsilon''} \tag{2.7}$$

where λ is the wavelength of the electromagnetic wave and ϵ' and ϵ'' correspond to the real and complex part of the complex dielectric constant $\epsilon = \epsilon' + i\epsilon''$. The P, L, S, C and X bands penetrate the atmosphere easier than the shorter wavelengths.

2.4 Polarimetry

Polarization has a major role in remote sensing. All electromagnetic waves can be described by complex vectors [1]. Depending on the orientation of the electromagnetic wave, the information of the output may vary. Depending on the number of polarizations a system operates in, more information of the surface area can be interpreted.

The orientation of the electromagnetic wave is decided upon the electric field of said wave. If the electric field is moving in the horizontal plane, the electromagnetic wave is said to be horizontal. The same goes for every other orientation, such as vertical and elliptical.

In remote sensing, the system may transmit in one or more polarization and receive in another or the same polarizations. This is commonly abbreviated, e.g.: HH, HV, VH, VV. The first letter refer to the transmitted signal and the

last refers to the interpreted signal. A system may very operate with several of these, such as dual pol and quad pol.

2.5 Reflection, absorption and transmittance

When an electromagnetic wave is incident on a surface, some of the energy is reflected in the specular direction and some of it is scattered, depending on the surface. Specular reflection occurs when the surface is smooth relative to the wavelength. Diffusion or isotropic reflectance occurs when the surface is rough relative to the wavelength (J. Zyl, C. Elachi[1]). A surface which perfectly diffuses the incoming electromagnetic wave is called a Lambertian surface, named after Johan Henrich Lambert. This means that the radiance of the surface is equal for any angle of reflection θ to the surface normal (Slater, P. [3]). Specular reflection, diffuse scattering and spreading is illustrated in Figure 2.2.

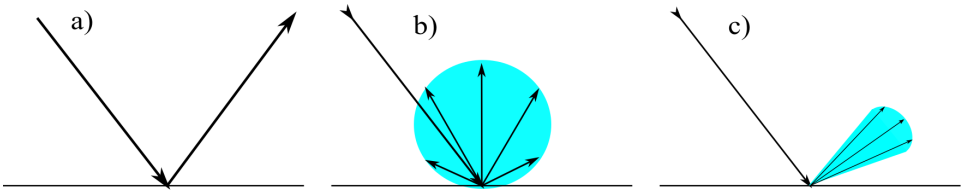


Figure 2.2: a) Left illustrate specular reflection. b) illustrate diffuse scattering. c) illustrate spreading.

Lambert's cosine law states that the flux per unit solid angle in any direction from a Lambertian surface varies with the cosine of the angle between the direction of the wave and the normal to the surface:

$$I_\theta = I_0 \cos\theta \tag{2.8}$$

Where I_0 is the incoming radiant flux, and I_θ is the outgoing radiant flux and angle θ of the surface normal.

Snell's law describe the relationship of incidence angles and the refracted energy and is given by:

$$n_1 \sin\theta_i = n_2 \sin\theta_t \tag{2.9}$$

where n_1 and n_2 is the refractive index of medium 1 and medium 2 respectively given by the Maxwell relations $n_r = \sqrt{\epsilon_r}$ and θ_i is the incidence angle, and θ_t is the transmission angle. Smooth surfaces relative to the incident wavelength λ cause the energy to be reflected in specular direction and the reflectivity is

given by equation 2.10. The reflection coefficient is a function of the refractive index and the incidence angle and is given by:

$$|R_h|^2 = \frac{\sin^2(\theta - \theta_t)}{\sin^2(\theta + \theta_t)} \tag{2.10}$$

where R_h is the reflection coefficient for horizontally polarized waves, θ is the wave incident on the surface and θ_t is the transmission angle. For vertically polarized incident waves the reflection coefficient is given by:

$$|R_v|^2 = \frac{\tan^2(\theta - \theta_t)}{\tan^2(\theta - \theta_t)} \tag{2.11}$$

This derivation is only valid for smooth surfaces, however in most cases the geometry of the surface is not smooth and will affect the relationship between the incident wave and the reflected wave.

Since the electromagnetic waves depend on the roughness of the surface, a criteria must be decided to establish if a surface is "smooth" or "rough" in relation to the wavelength. The most common criteria is the Rayleigh criterion defined as

$$h > \frac{\lambda}{8\cos\theta} \tag{2.12}$$

where h is the root mean squared height, λ is the radar wavelenght and θ is the incidence angle [1]. It is considered rough if the equality holds. A more accurate description is that a surface is considered smooth if $h < \frac{\lambda}{25\cos\theta}$ and rough if $h > \frac{\lambda}{4\cos\theta}$ while the values in between is considered to have intermediate roughness [29].

2.5.1 Double Bounce

Double bounce refers to the occurence where the EM wave comes in contact with two surfaces before it returns to the imaging sensor. When an EM wave interacts with surfaces as shown in Figure 2.3, there will be a 180 degree change in the VV channel.

2.5.2 Volume Scattering

Volume scattering occur when the EM wave bounces within a medium when there is variation of dielectric properties. When this occur the wavelength will also decide how far the signal propagate inside the medium as the penetration depth is strongly dependent on wavelength. Volume scattering is illustrated in Figure 2.4.

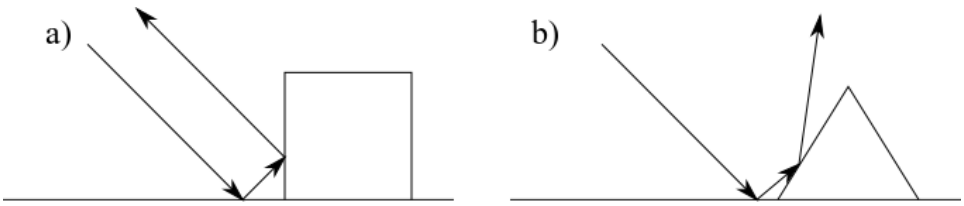


Figure 2.3: Double bounce, a) show double bounce on perpendicular surfaces, b) show double bounce on non-perpendicular surfaces.

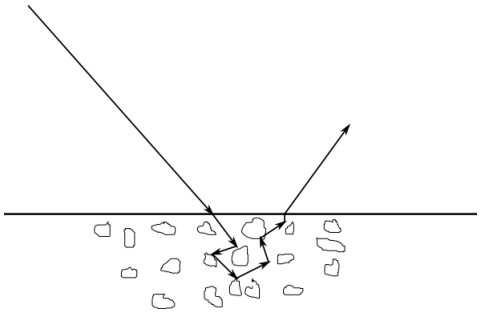


Figure 2.4: Volume Scattering

2.6 Speckle

As the radar pulse illuminates a point on the scene, the return signal includes scatter from several adjacent areas and thus includes additional information. The returns from each point add vectorially and lead to a single vector that represents the amplitude and phase of the total echo.

The phase is related to the distance between the scatterer and the sensor. A change in position will result in a difference in phase for each elementary vector, which in turn lead to change in the composite amplitude.

Therefore, different observations over the same area will result in variations of the amplitude V. To resolve this a method called multilooking is used to even out the differences. In essence, an image of homogeneous surface and constant backscatter cross section will show variations of intensity, called speckle. The effect of speckle can further mathematically be represented by:

$$V = \sum_{n=1}^{N_S} V_n e^{i\phi'_n} = V_e e^{i\phi'} = V_x + jV_y \tag{2.13}$$

This expression represents the voltage V due to N_S scatterers. V is the amplitude of the wave and ϕ' is the phase angle. V_e is the envelope of the signal, $V_x = V_e \cos\phi'$ and $V_y = V_e \sin\phi'$.

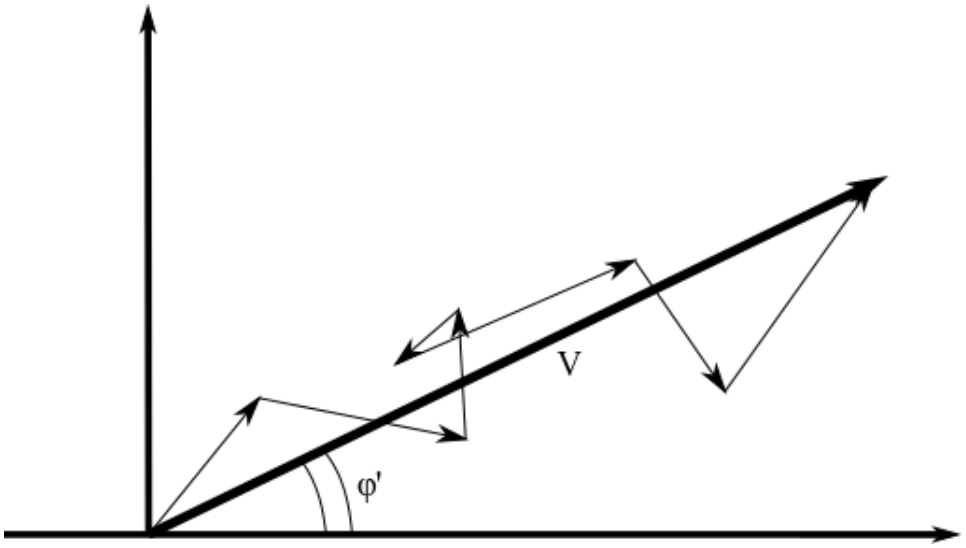


Figure 2.5: The composite return signal, this process is often called a random walk.

2.7 Radar Cross Section

The radar cross section (RCS) is a ratio which determines the detectability of an object. A larger radar cross section equals an object which is easier to detect. The radar cross section of any reflector is thought to be the projected area of an isotropic scatterer where it returns the EM wave with equivalent power [2]. The RCS can be expressed as

$$\sigma' = \frac{I_{recieved}}{I_{incident}} 4\pi R^2 \tag{2.14}$$

where I is the intensity and R is the range [7]. It is common to transform σ' to the normalized radar cross section (NRCS),

$$\sigma_0 = \frac{\sigma'}{A} \tag{2.15}$$

where A is the area of illumination.

2.8 Signal to noise ratio

SAR images will be affected by several forms of noise. It is therefore important to note that the signal received will have to be larger relative to the noise for the product to hold useful information. SAR images contain two forms of noise, additive and multiplicative. The additive noise is due to the equipment and

the limitations it hold with regard to the imaging system. An example of this is the thermal noise which occur in the satellites. This noise is known and can be considered to be white noise with Gaussian distributions.

Multiplicative noise is noise is an undesired random signal which is multiplied into a relevant signal during the imaging process. An example of this is the speckle effect which is discussed above and in later sections.

To be able to quantify the retrieved signal we have to be able to ascertain the quality of the signal. A measure of this is the signal to noise ratio defined as:

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{2.16}$$

where P_{signal} and P_{noise} is the power of the signal and noise respectively. Any ratio higher than 1, indicate more signal than noise. Increasing the power of the output signal will increase the signal to noise ratio. Due to the nature of noise in SAR imagery a noise-floor can be established which is a threshold of where the uncertainty of the true signal is to great. This noise floor vary from sensor to sensor as it is greatly dependent on each component.

2.9 Image Deformation

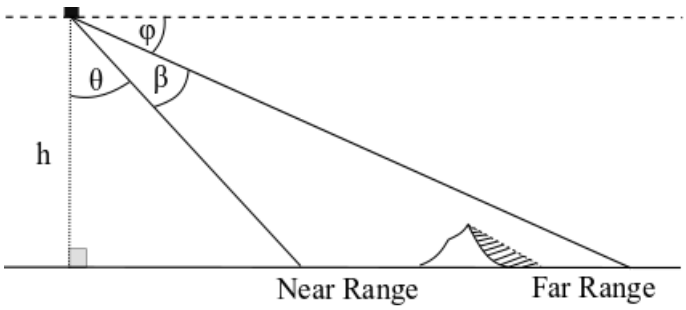


Figure 2.6: Shadow & depression angle

Depending on the depression angle and the steepness of the terrain some areas may not be illuminated by the radar beam, generating shadows. This is shown in the figure above. As the depression angle become smaller, the radar shadows become more severe.

As the radar system measures the time between transmitted signal and the

received signal in the slant-range, images may contain geometric artifacts. This can be shown in Figure 2.7 and 2.8.

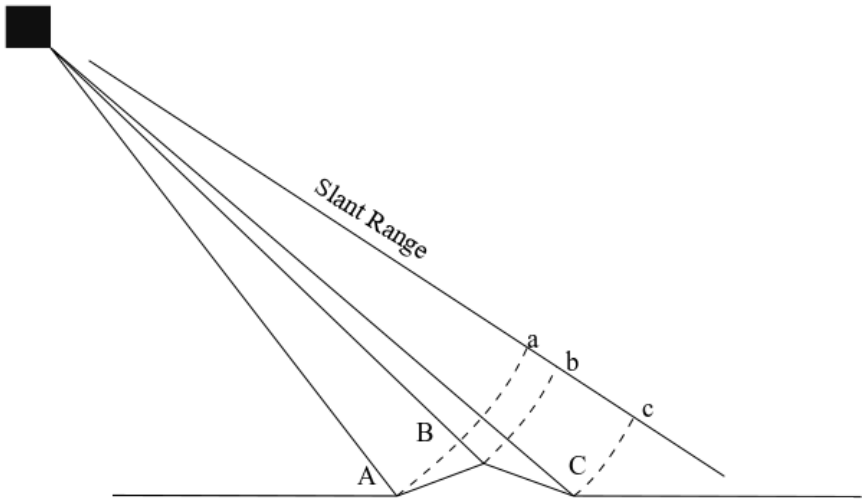


Figure 2.7: Radar foreshortening

Radar layover is when the top of an object is closer to the antenna than the base of the object, causing the radar echo from the top reach the antenna before the base. As the system measures distance with respect to time, the top will appear closer to the antenna than its base.

Radar foreshortening is when the geometric locations are the correct order but the echo from the base is perceived closer to the the echo from the top than it truly is.

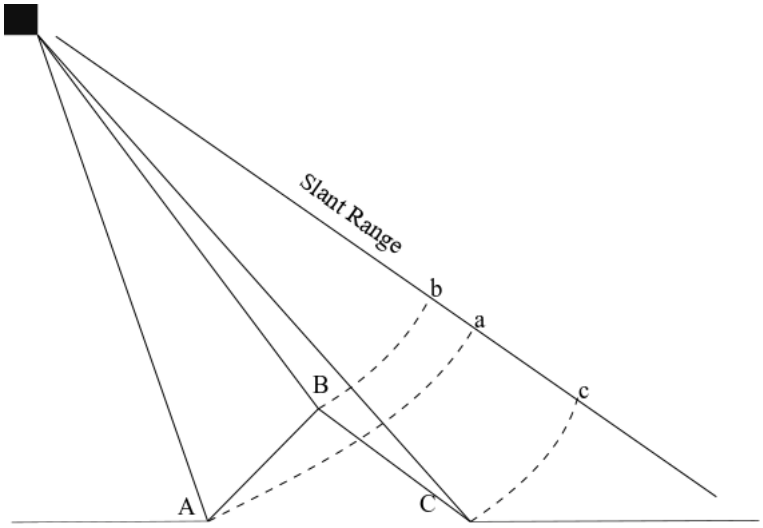


Figure 2.8: Radar layover

3

Statistics and Linear Regression

This chapter covers the theory needed for this thesis. It involves incidence angles and statistics necessary to investigate variance, skewness and kurtosis over incidence angles.

3.1 Theory on Incidence Angle

The brightness decay over incidence angle is a known issue in SAR images where the swath width is relatively large. In 1941 Minnaert [9] explained this phenomenon by studying the Moon, where he states that the Moon radiates more energy in the direction of the Sun, which is the source, than in the normal direction. This phenomenon can also be explained by Lamberts Cosine Law. The effect can be seen in Figure 3.1 which illustrates a typical decrease in brightness over incidence angle.

Minnaert modified Lamberts law to fit for non-ideal diffuse emitters [6, 9]:

$$I_{i_m} = I_0 \cos^{2k} \theta_i \tag{3.1}$$

where I_0 is the radiance in the normal direction, k is the Minnaert constant

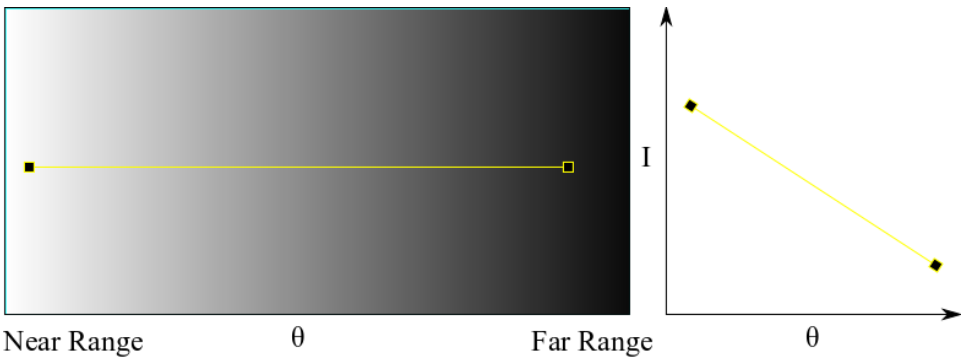


Figure 3.1: An illustration of a typical effect which can be seen in SAR images

which is a measure of surface roughness and θ_i is the incidence angle. The usage of this is that two surfaces have different decay rates based on roughness which can be utilized in classification.

Empirical evidence, indicate that the decay rate follows an exponential function [6]:

$$I_{i_2} = I_0 e^{-\theta_i/\theta_0} \tag{3.2}$$

where θ_0 is the decay rate and θ_i is the same incidence angle as above. This expression is linear in the logarithmic domain, such that the log-intensity becomes a linear function of the incidence angle.

$$I_{i_2}[dB] = a\theta_i + b \tag{3.3}$$

where $a = -1/\theta_0$ and $b = \log I_0$.

3.2 Speckle Statistics

For this project we need a further understanding of speckle than was discussed in Section 2.6.

The addition of N scatterers which result in either constructive or destructive addition, as mentioned in Chapter 2.6, can be assumed to be independent and identically Gaussian, by the central limit theorem [8].

The amplitude of the wave is defined as:

$$V_e = \sqrt{V_x^2 + V_y^2} \tag{3.4}$$

where V_e has a Rayleigh distribution. This is due to the randomness of the scattering points and the large number of scatteres for each pixel. By the central limit theorem V_x and V_y both have zero mean, has a Gaussian distribution and can be assumed independent [8]. The joint probability distribution can then be written as:

$$\begin{aligned} p(V_x, V_y) &= p(V_x)p(V_y) \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-V_x^2/2\sigma^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-V_y^2/2\sigma^2} = \frac{1}{2\pi\sigma^2} e^{-(V_x^2+V_y^2)/2\sigma^2} \end{aligned} \tag{3.5}$$

Since we are interested in the resulting sum of the phasors (V_e) we transform the above to polar coordinates:

$$p(V_e, \phi')\mathrm{d}V_e \, \mathrm{d}\phi' = p(V_x, V_y) \, \mathrm{d}V_x \, \mathrm{d}V_y \tag{3.6}$$

Since $V_x = V_e \cos\phi'$ and $V_y = V_e \sin\phi'$ and the Jacobian is given as

$$\mathrm{d}V_x \, \mathrm{d}V_y = \begin{vmatrix} \cos\phi' & -V_e \sin\phi' \\ \sin\phi' & V_e \cos\phi' \end{vmatrix} \mathrm{d}V_e \, \mathrm{d}\phi' = V_e \, \mathrm{d}V_e \, \mathrm{d}\phi' \tag{3.7}$$

Inserting equation 3.5 and 3.7 into 3.6 which result in the joint probability distribution function:

$$p(V_e, \phi') = \frac{V_e}{2\pi\sigma^2} e^{-V_e^2/2\sigma^2} \tag{3.8}$$

Integrating Equation 3.8 with respect to ϕ' over the interval $(-\pi, \pi)$ which is one rotation, the probability distribution function of the amplitude becomes:

$$p(V_e) = \frac{V_e}{\sigma^2} e^{-V_e^2/2\sigma^2}, \quad V_e \geq 0 \tag{3.9}$$

which is a Rayleigh distribution with mean $\mu_1(V_e) = \sigma\sqrt{\pi/2}$ and variance $Var_1(V_e) = (2 - \pi/2)\sigma^2$. However, this only applies to single look SAR images. For multi-look images the distribution changes which will be shown in a later section.

Since the distribution for V_e is Rayleigh distributed, the power P will be exponential distributed. Which is shown by:

$$\begin{aligned} P &= V_e^2 \\ \mathrm{d}P &= 2V_e \mathrm{d}V_e \end{aligned} \tag{3.10}$$

The distribution for the power then becomes:

$$p(P)\mathrm{d}P = p(V_e)\mathrm{d}V_e = \frac{1}{2\sigma^2}e^{-P/2\sigma^2}\mathrm{d}P \tag{3.11}$$

since $E(P) = E(V_e^2) = 2\sigma^2 = \mu_P$:

$$p(P) = \frac{1}{P}e^{-P/\mu_P} \tag{3.12}$$

which is exponentially distributed with mean $\mu_P = 2\sigma^2$ and variance Var(P) = μ_P^2.

3.2.1 Multi-look Speckle Statistic

Multi-looking is a process where the goal is to reduce the effect of speckle in the image, which is essentially averaging several independent estimates of reflectivity [8]. This is done by dividing the Doppler frequency spectrum into N segments commonly referred to as looks.

The intensity of the image with N looks can be be described by:

$$I_S = \frac{1}{N}\sum_{i=1}^{N}(V_x(i)^2 + V_y(i)^2) = \frac{1}{N}\sum_{i=1}^{N}V_{ei}^2 \tag{3.13}$$

where $V_x(i)$ and $V_y(i)$ are the real and imaginary components of the ith look. These parts are independently Gaussian distributed such that NI_S is Chi-square distributed with $2N$ degrees of freedom. Therefore the probability density function for a N-look intensity image follows the equation:

$$p(I_S) = \frac{I_S^{N-1}}{(N-1)!(\frac{2\sigma^2}{N})^N}e^{-NI_s/2\sigma^2} \tag{3.14}$$

which is a chi-squared distribution. When N becomes larger, it approaches a Gaussian distribution.

3.3 Speckle Noise model

Speckle can be considered to be a multiplicative noise. In image processing this is illustrated by:

$$y(r,c) = x(r,c)n(r,c) \tag{3.15}$$

where $y(r,c)$ is the pixel intensity with noise, $x(r,c)$ is the pixel intensity free of noise and $n(r,c)$ is the noise [8] with mean 1 and variance σ_v. The pixel

intensity and noise is assumed to be statistically independent. This means that $E[y] = E[x]$. The variance is defined as:

$$\begin{aligned} Var(y) &= E[(y - \bar{y})^2] = E[(x(\sigma_v - 1) + (x - \bar{x})^2)] \\ &= (Var(x) + \bar{x}^2)\sigma_v^2 + Var(x) \end{aligned} \tag{3.16}$$

If the area is homogeneous, this would mean that $Var(x) = 0$, which results in:

$$\sigma_v = \frac{\sqrt{Var(y)}}{\bar{x}} = \frac{\sqrt{Var(y)}}{\bar{y}} \tag{3.17}$$

This ratio is a measure of the speckle noise level and is dependant of the number of looks done in the image.

3.4 Linear Regression

Linear regression is used when it is known that there exist an inherent relationship among variables. As this relationship exist it may be used to predict values by creating a model which relates the response to the regressor [14]. This can be described by:

$$Y = \beta_0 + \beta_1 x + \epsilon \tag{3.18}$$

where β_0 is the intercept, β_1 is the slope and ϵ is the error term. There are several ways to calculate this relationship. One of the most used is the method of least squares, where the goal is to determine β_0 and β_1 such that the residual sum of squares is minimized, which can be described by:

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{3.19}$$

Differentiating SSE with respect to β_0 and β_1 and setting the partial derivatives to zero, the coefficients can then be calculated:

$$\frac{\partial(SSE)}{\partial\beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \tag{3.20}$$

$$\frac{\partial(SSE)}{\partial\beta_1} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i = 0 \tag{3.21}$$

Solving for β_0 and β_1 yields:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3.22}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \tag{3.23}$$

This can be extended to include several variables:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad \ldots, \beta_n x_n \epsilon \tag{3.24}$$

A special case for multiple linear regression is the polynomial regression where the variables are exponential.

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \quad \ldots \quad + \beta_n x^n + \epsilon \tag{3.25}$$

3.5 Hypothesis Testing

When performing a linear regression, it is important to determine if the variable explains the actual change in the data. To do this we have to establish a hypothesis test. A hypothesis is defined as a statement about a population parameter [15]. There exist several tests which can be utilized, among them are the T-test and F-test. Both tests use the principle where the two hypothesis are set:

$$\begin{aligned} H_0 \ &: \quad \beta_1 = \beta_{1,0} \\ H_1 \ &: \quad \beta_1 \neq \beta_{1,0} \end{aligned} \tag{3.26}$$

To test these hypotheses a test statistic is used. The test statistic is based on the test utilized. For a T-test the test statistic is

$$t_0 = \frac{\hat{\beta}_1 - \beta_{1,0}}{se(\hat{\beta}_1)} \tag{3.27}$$

where $\hat{\beta}$ is the estimated rate of the regressor for the fitted line, $\beta_{1,0}$ is the test value and $se(\hat{\beta}_1)$ is the standard error of $\hat{\beta}_1$. The standard error is defined as

$$se = \sqrt{\frac{\sigma_c^2}{n}} \tag{3.28}$$

This test statistic is often applied when the tested variable is normally distributed. The test statistic forces the observations to a t-distribution under the null-hypothesis.

The F-test follow the same principle as a t-test, where the tested variable has an F-distribution under the null-hypothesis. The test statistic is defined as:

$$F_0 = \frac{MS_R}{MS_E} = \frac{SSR/(K-1)}{SSE/(N-K)} \tag{3.29}$$

where MS_R is the regression mean square, MS_E is the mean square error, K is the number of groups and N is the sample size.

The F-tests are used in several cases where the most common are:

- Test if the means of data sets are equal [15, 16].
- Test the hypothesis that a regression model fits the data well[14, 16].
- That the data set in a regression analysis follows the simpler of two nested regression models [14, 16].

3.6 Correlation Coefficient

Another measure for determining the slope of a data set is the Correlation Coefficient [14]. It is the linear correlation of two variables. The coefficient takes a value between -1 and 1 where a negative value correspond to a negative linear correlation and a positive value correspond to a positive linear correlation. If the coefficient is 0 there is no linear correlation. The coefficient is defined as:

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \tag{3.30}$$

where $\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$ and σ_X and σ_Y is the standard deviation of X and Y respectively.

3.7 Central Moments

Central moments is a moment around a probability distributions from a random variables mean [14]. These values characterize the distribution of a random variable, such as the much known variance. The n-th central moment is defined as:

$$m_n = E[(X - E[X])^n] \tag{3.31}$$

The second central moment is equal to the variance:

$$m_2 = E[(X - (E[X])^2] = E[X^2] - E[X]^2 \tag{3.32}$$

The third and fourth moments are mostly used to get the standardized moments,

which is defined as

$$\tilde{\mu_k} = \frac{m_k}{\sigma^k} \tag{3.33}$$

where m_k is the k-th central moment and σ is the standard deviation.

The third moment is the skewness and holds information of the symmetry around a distribution. For a Gaussian distribution this value is zero as it is symmetric around its mean. Skewness can be described by:

$$\tilde{\mu_3} = \frac{m_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}} \tag{3.34}$$

The fourth moment is kurtosis. This standardized moment gives an indication of how much the "tails" of the distribution impact the total shape. The kurtosis of a univariate Gaussion distribution is 3.

$$\tilde{\mu_4} = \frac{m_4}{\sigma^4} = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^{4/2}} \tag{3.35}$$

3.8 Mixed Pixels and Variance

For classification, mixed pixels is something which needs to be considered. Mixed pixels are created when a pixel is not completely occupied by a single homogeneous area [24]. Therefore, when calculating variances in the image where mixed pixel occur, we may obtain a larger value. This larger value may then be an outlier for classification or a reason for error in linear regression.

3.9 Multivariate Gaussian Distribution

As mentioned in Chapter 2.6, given that we follow the Rayleigh speckle model, it can be assumed that the distribution of the intensities is Gaussian. The multivariate case results in the equation

$$p(x;\mu,\Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu)^{\mathbf{T}}\Sigma^{-1}(x-\mu)} \tag{3.36}$$

for a vector-valued random variable $X = [X_1, \ldots, X_n]^{\mathbf{T}}$, where μ is the mean and the Σ is the covariance matrix and $\mathbf{T}$ is the transpose of the matrix.

4

SAR Remote Sensing of Ocean and Sea Ice

In this chapter ocean and sea ice will be discussed with respect to remote sensing. Firstly a brief explanation of remote sensing on ocean will be given. Then a short summary of different sea ice types and its interaction with electromagnetic waves will be discussed.

4.1 Ocean

SAR is a widely used tool for oceanographic studies. It provides a tool where large areas can easily be analysed granting insight into several important fields such as wave and wind direction, ship detection and sea ice mapping [1]. One of the properties of electromagnetic waves from the SAR is that they only penetrate the surface by a few millimeters [25]. The SAR imagery thus only reflects the surface expression, and is very useful for applications where this is sought for.

Imaging of ocean is mainly dependant on Bragg resonance [26], except in the specular direction where specular scattering is dominant. Bragg resonance is the effect that explains the reflection of electromagnetic waves on periodic structures.

4.2 Sea Ice

The electromagnetic waves are dependent on the wavelength relative to the roughness of the surface. For sea ice, the roughness is strongly dependent on the weather conditions and the salinity of the surrounding sea. When the sea water begins to freeze small ice crystals [10] called frazil begins to form on the surface. The frazil crystals will push out the salt minerals which is called brine causing the nearby water to be saltier. As the frazil begin to gather, they may form different structures based on the outside sources of interaction. They are sorted as:

- Grease Ice: Grease ice is formed from frazil crystals in relatively calm waters. It is the coagulation of frazil ice crystals which has a dampening effect on the small capillary waves on the ocean surface [10]. This has a great impact on remote sensing, as calmer waters will have a low backscattering level and will appear as a dark spot in radar images.

- Nilas: Should the grease ice be allowed to continue to freeze, a relatively thin layer of ice will form and is called nilas if it is less than 10 cm thick. Nilas is further divided into dark nilas and light nilas which depends on the thickness of the ice. Furthermore, the brine drainage will also affect the visual interpretation of the nilas which is important for visual remote sensing. The pockets of brine surrounded by ice will begin to drain and be replaced by air, which changes the appearance from darker to lighter.

- Pancake Ice: In the presence of wind and swells, which is not unlikely in open waters, grease ice may break off forming small sheets of ice from 30[cm] to 3 [m]. With the continuous collision between these small sheets of ice they will form round and raised pancake formed ice sheets with rims 5-20 [cm]. Due to the rims they are able to reflect the electromagnetic radiation effectively and make them visible to the radar remote sensing system. Pancake ice may also be formed by slush or shuga, or any form of break-up in the ice formation under wind and wave conditions.

- Floes are relatively flat large pieces of ice, ranging from 20[m] across to up to 10[km] across.

Formations on the ice surface will vary with salinity and the amount of stress it has been exposed to. The dielectric constant is also stongly dependent on the salinity [23]. As newly formed and first year ice have a high salinity they will be very lossy, having low penetration depth and high reflectance. Snow will in a lot of cases be layered on top of the ice which will also affect the electromagnetic properties. However, in the microwave region the snow will

be mostly transparent unless the snow is wet. This is because non-frozen water has a low penetration depth such that during melting season snow may act as an opaque cover of the surface. Snow and ice may also be layered which will affect the penetration depth. Temperature also impact the penetration depth where lower temperature will increase the penetration depth.

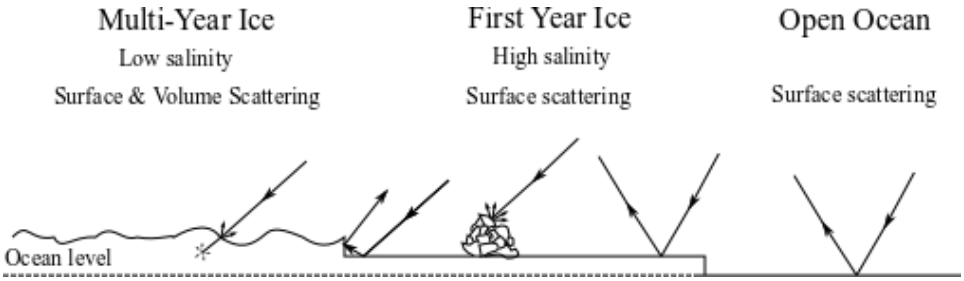


Figure 4.1: Illustration of scattering on multi-year ice, first year ice and ocean surface (based on Figure 3.6 in [23]).

5

Area of Study and Data Sets

This thesis is limited geographically to the arctic waters near Svalbard. This is chosen as it has sufficient sea ice and ocean exposure. The below sections describes the characteristics of the satellite and images used for the study.

5.1 Satellite

The images are taken with Sentinel 1 wide swath mode. Sentinel 1 SAR system works in C-band and has several modes in which it operates. The main modes are interferometric wide swath and wave. The interferometric wide swath mode is the main acquisition mode over land. The wave mode is used to determine the direction, wavelength and height of waves on the ocean [18].

There are two additional modes on Sentinel 1 which are the stripmap mode and the extra wide swath mode. The stripmap provide data with 80 km swath width with a resolution of 5x5m. Finally, the extra wide swath mode, which provide as indicated by the name, a product with wide coverage. This mode is very useful for maritime, ice and polar zone operational services [18]. Due to the large areas covered, the resolution is therefore reduced. The resolutions are shown in Table 5.1.

Mode	Resolution	Swath Width	Maximum NESZ
Stripmap	5x5m	80km	22dB
Interferometric Wide Swath	5x20m	250km	-22dB
Extra Wide Swath	20x40m	410km	-22dB
Wave	5x5m	Vignette Coverage	-22dB

Table 5.1: Sentinel-1 modes

5.2 Extra Wide Swath Mode

The Extra Wide Swath mode (EW) have 5 subswaths and is delivered in several product levels: 0, 1 and 2. Product level 0 is the raw data, and contains compressed and unfocused data. Level 1 data products can be separated into two categories, SLC (single-look complex) and GRD (Ground Range Detected). Level 2 products contain geophysical products derived from level 1.

This thesis uses mainly level 1 GRD products, where the data has been detected, multi-looked and projected to ground range. The data further consist of 5 sub-swaths of varying incidence angle. Depending on the end resolution, the processing of this data may differ from each sub-swath. For ground range detected medium resolution (GRDM), the first sub-swath, has been multi-looked 6x3 times in range and azimuth respectivly, while the rest have been multi-looked 6x2 times [19]. When calculating variance this is important as multi-looking specifically reduce variance, such that the different swaths will have an inherent difference unless a correction is made.

5.3 Images

Several images have been investigated and examined for use in the thesis. It was however difficult finding scenes where there was enough uniform areas for selection and calculations. Due to ocean surface having a low level of backscatter, and the noise floor seemingly having an effect on the variance, the image used over ocean need a high level of backscatter.

The final images used are all within the arctic circle and the file names are shown in Table 5.2. They are centered around Svalbard (See Figs. 5.1, 5.2 and 5.3).

Images Used	ID
S1A_EW_GRDM_1SDH_20151127T152113_20151127T152213_008792_00C8A0_77BA.SAFE	77BA
S1A_EW_GRDM_1SDH_20180224T050038_20180224T050143_020744_0238C0_4479.SAFE	4479

Table 5.2: Scenes and corresponding ID

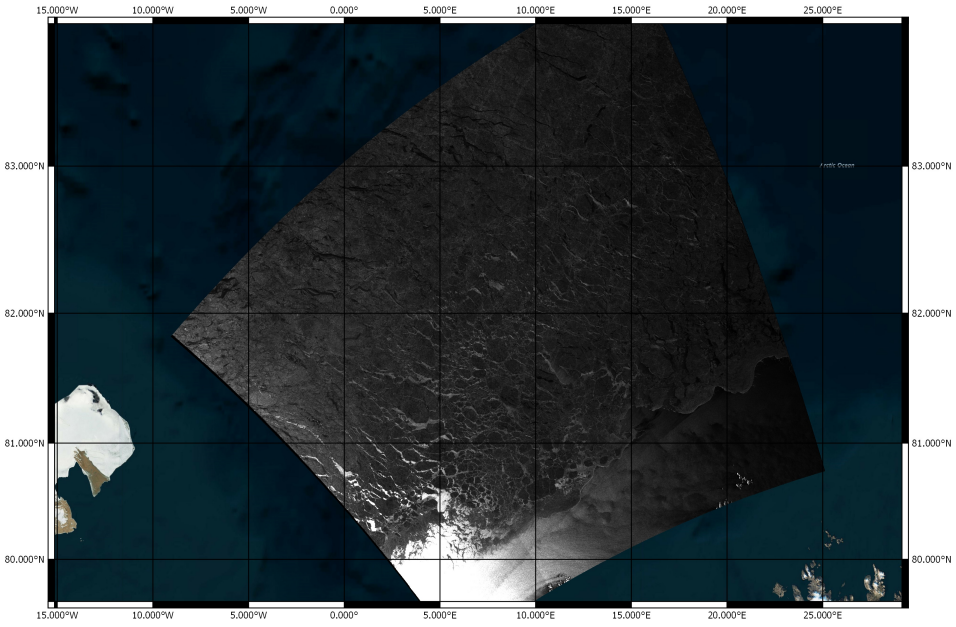


Figure 5.1: Image id 77BA

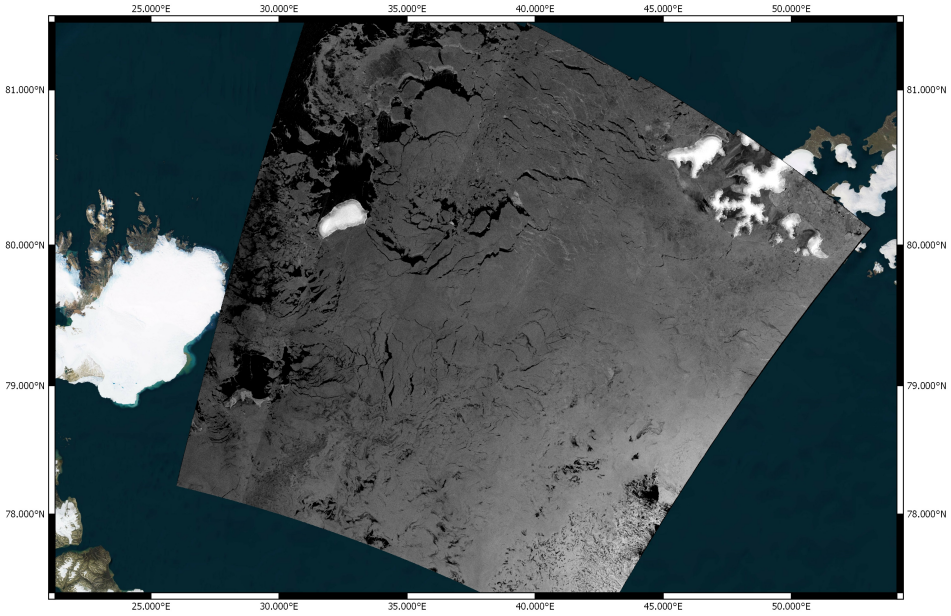


Figure 5.2: Image id 4479

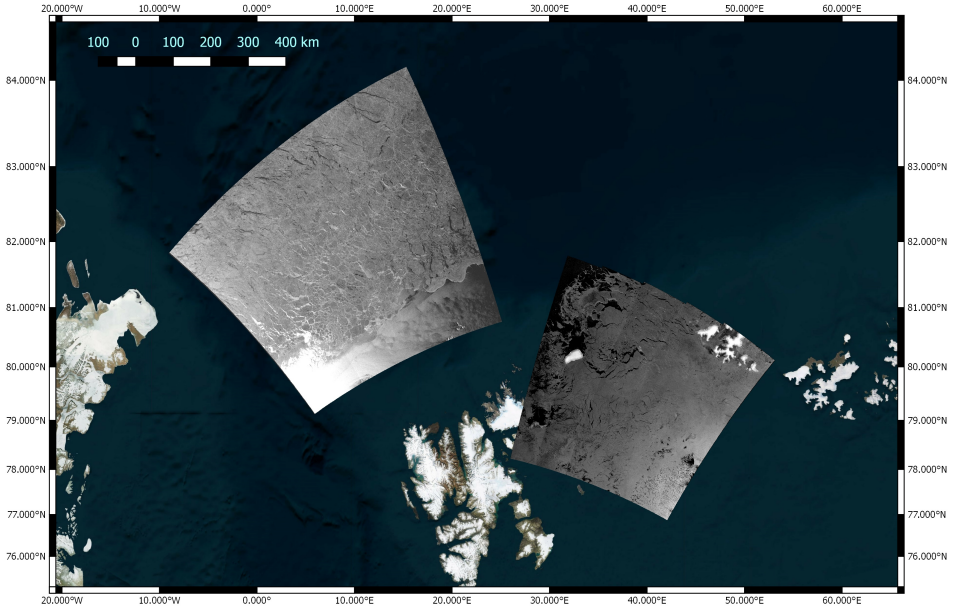


Figure 5.3: Image 77BA to the left and 4479 to the right, Svalbard in the bottom center

6 Method

The main parameters for this thesis are the speckle variance, skewness and kurtosis. This chapter considers the analysis methodology used to examine these features.

6.1 Linear Trend of Variance

The images provided by Sentinel-1 is taken in the EW mode, with incidence angle ranging from 18.9 - 47.0 degrees. The scenes are opened in SNAP for calibration and noise removal is applied using the built-in functions. After calibration the sigma naught (NESZ) is obtained from the intensity of the image. All images are then multi-looked with an additional 5 by 5 window, also using the built-in function. By multi-looking the distribution of the speckle will approach a Gaussian distribution as discussed in previous sections.

The first step is therefore to check that the previously obtained results are correct, as to establish a basis for continuation. This is that the decay rate of the mean follow an exponential distribution, which in the dB domain is linear as discussed in Chapter 3. Furthermore, the decay rate will be different for different classes such as ice and ocean water.

When the decay rate of the mean is established, further work on the variance can be carried out. This is done by selection of a region of interest which is as

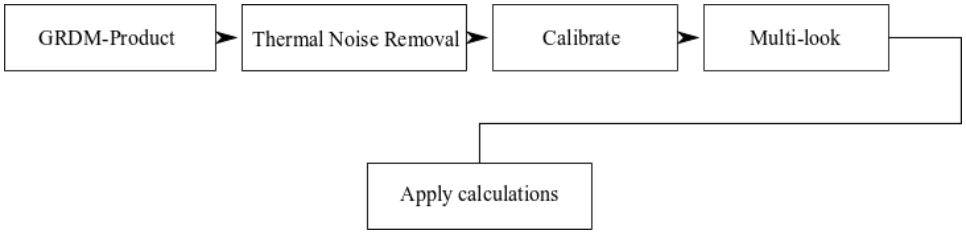


Figure 6.1: Processing Steps

uniform as possible. This is because the variance of a uniform area should be zero and therefore the remaining variance is due to speckle and not variation in the surface [8]. The variance is calculated in the dB image as the trend in mean in CIRFA's model is done in this domain. In the first image shown in 5.1, over Svalbard, sections of ocean is segmented out where the surface is uniform. Then the section for sea ice is selected for comparison. The same procedure is carried out for Figure 5.2 which is of newly formed sea ice. This image is selected as it is easier to select regions which is uniform to use for comparison.

To find the means, variances and higher moments, a moving window of a predefined size moves across the image calculating the values for the local area. If the window move close to the border of the selected area it may enclose regions outside the region of interest. The fix for this is to remove these values before calculation of variance and the higher moments. A minimum of 25 data points for each calculation is set to prevent outliers.

By plotting the variance and higher moments over incidence angle, the dependency can be visualized. With no dependence the trend will be flat over all angles, while if there is a dependence the trend can be found.

6.2 Processing of images

The images provided by ESA has some processing done to them the start as the images used are GRDM products as mentioned in Chapter 5.2. Although this may be the case, further processing must be used to extract useful information from the scene. These steps are shown in Figure 6.1.

6.2.1 Calibration

Calibration is the first step for to get usefull information. This relates the digital numbers from the satellite to a meaningful value [20]. The main goal for calibration is to address the issue of noise and gain term. Calibration in Sentinel-1 can be done by:

$$\sigma_0 = \frac{DN^2}{K^2} \tag{6.1}$$

where σ_0 is the backscatter coefficient, DN is the digital pixel amplitude value and K is extracted from the look-up tables (LUT), in the metadata.

6.2.2 Multi-looking

As mentioned, the GRDM products used in this study has been multi-looked to form the GRDM scenes. They are further Multi-looked to reduce the size of the file. The additional multi-looking is done by a 5x5 window averaging.

6.3 Regions of Interest

As mentioned the regions we are to look at have to be as homogeneous as possible to determine the variances dependence on incidence angle. From the regions the arithmetic mean and local variance can be calculated over the homogenous areas. The averages and variances are calculated from a box moving over the incidence angle of the decibel image.

$$\mu_a = \frac{1}{N}\sum_{i}^{N} g_i \tag{6.2}$$

where μ_a is the local mean and g_i is the pixel value in a N by N bounding box. The local variance is also calculated the same way by

$$\sigma_a^2 = \frac{1}{N-1}\sum_{i}^{N} (g_i - \mu_a)^2 \tag{6.3}$$

where g_i is the pixel value. When the moving box contain pixels which is not in the region selected, they are removed from consideration which leaves the local variance to be calculated from fewer points. Should the number of points used be reduced below 25 data points, it moves to the next location.

6.4 Sentinel 1

After careful selection of regions using Sentinel 1A and 1B products, there is a change in variance values in the transition from the individual sub-swaths. Every image investigated have an increase in variance at incidence angle 27 to 30 degrees. This jump in value is more noticeable when the number of multi-looking is kept low as the averaging do not blur it out. Multi-looking 5x5 times seem to preserve the "step" in variance. If multi-looking is done too many times this may indicate a upward trending slope which may not have anything to do with incidence angle.

From this the following can be done, either change the satellites or find a work-around that can be implemented. As this jump seemingly appear in all Sentinel 1 GRDM products, the investigation can be carried out for each sub-swath in the scenes to further investigate the change in variance.

The HV channel in Sentinel 1 is presented with azimuthal noise which is periodic in nature [21]. Most of this noise is removed through thermal noise removal. Low backscatter level is also an issue with the HV polariziation as it is near noise floor. To simplify the task at hand the main focus was put on the HH polarization channel where these issues are less prominent.

6.5 Hypothesis Test

After the speckle features are found and plotted against the incidence angle, we can test the statistical significance of the findings. This is done to substantiate whether there is a statistical significance to the slope of variance, skewness or kurtosis.

7

Results and Discussion

This chapter presents the results and findings of the thesis. It also shows the results from the different processing steps to produce the final results for linear regression and modelling of the speckle properties to incidence angle.

7.1 Variance

At higher incidence angles, the area on the ground will be larger, such that the number of speckle from scatters will be larger [22]. This should results in a decaying variance within the same class over the image. There is however, an importance to the level of backscatter. If the intensity of the image approaches the noise floor for the satellite, the variance seem to be severely infected by the noise floor. An example of this is shown in the Appendix A.1. The focus is therefore on images where the level of backscatter is high enough to provide reasonable security that the variance observed is due to speckle variation and not the noise floor. The region of interest can also be reduced to incidence angles where it is sufficient level of backscatter. The limit is set to intensity of -20 [dB] to be well above the noise floor.

7.1.1 Image 1

As mentioned in previous chapters, the areas where the variance is to be calculated has to be homogeneous such that the variance observed is not due to between-class variance. This will tend to an increase in local variance. As can be seen in Figure 7.1, there are at least two classes of ocean water due to local difference in wind speeds. The regions selected are the brighter areas as these regions are the most bountiful of the visual regions and spread across the whole scene. These regions can be seen in Figure 7.2.

The variance is calculated as mentioned in Chapter 6, by a "sliding box" within the selected regions. The resulting plot is shown in Figure 7.3. There is an obvious skip at incidence angle 28 degrees, where the first sub-swath ends. This is because the first sub-swath is multi-looked more than the others in GRDM products (see Section 5.2). This extra multi-looking, 18-looks compared to 12-looks, will therefore influence the variance values observed. This skip is not seen in the mean values using the same regions shown in Figure 7.4.

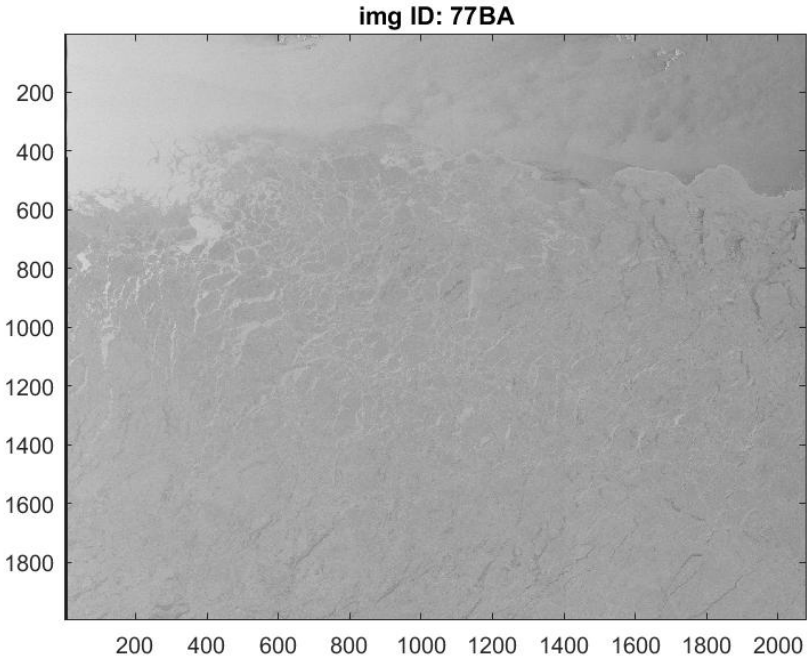


Figure 7.1: First image used where ocean water is in the top of the image and sea ice at the bottom.

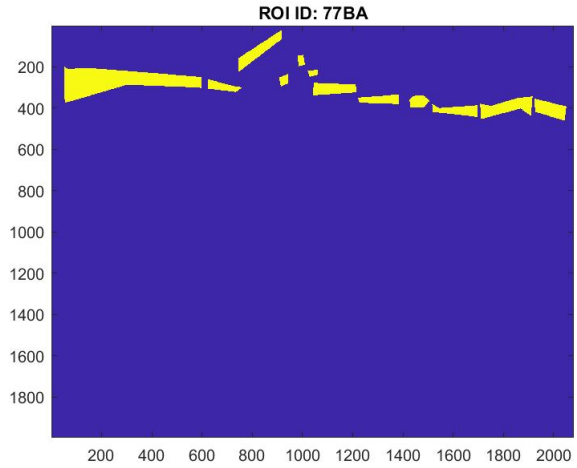


Figure 7.2: Region of interest in image id :77BA

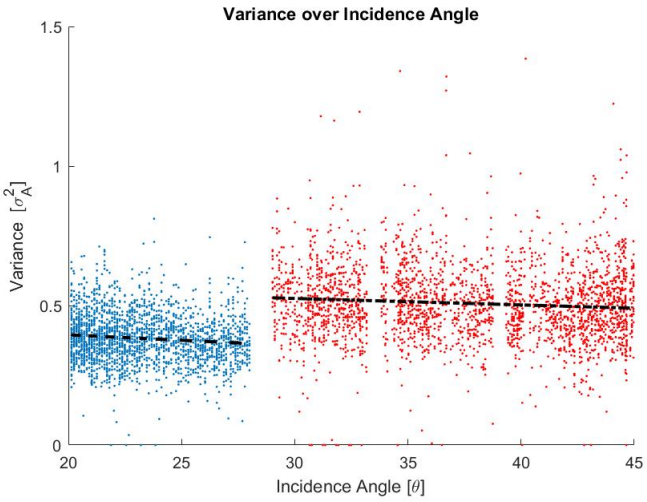


Figure 7.3: Scatter plot of variance over incidence angle. The blue dots are from the first subswath while the red is the from the rest. A regression line is fitted through the data for both cases.

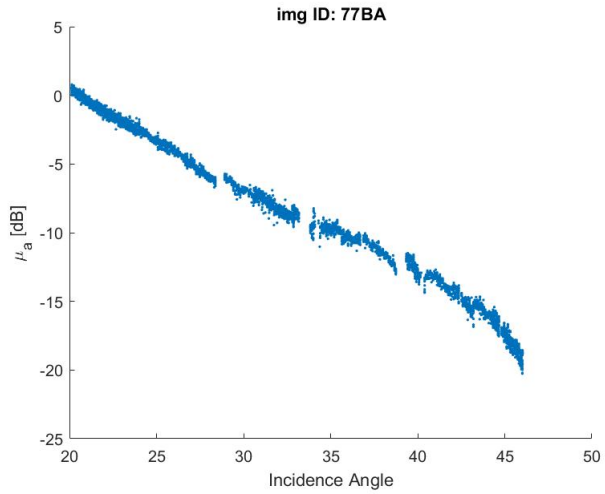


Figure 7.4: Mean of the logarithmic values caluclated in a 7x7 box plotted in [dB] to see the simple log-linear trend.

As described by de Vries [17], variance in the intensity of multi-looked images are related by

$$\text{var}(I_m) = \frac{\text{var(I)}}{L} \tag{7.1}$$

where $\text{var}(I_m)$ is the variance of multi-looked intensity with L looks and $\text{var}(I)$ is the variance of the intensity of the image.

An adjusted scatter plot, where the number of multi-looks in the first sub-swath is normalized to the number of looks in the others can be done with simple multiplication by $\frac{18}{12}$ and is shown in Figure 7.5.

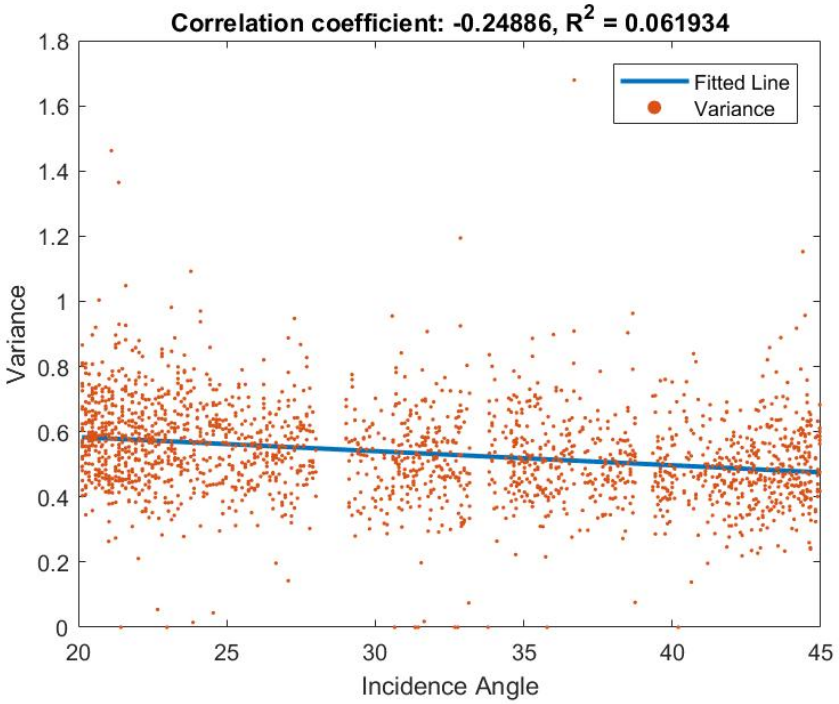


Figure 7.5: Corrected Sub-Swath

Figure 7.5 clearly shows a reduction in variance over incidence angle. The correlation coefficient shows us that there is a downward trend in the data set. The fitted regression line is also shown in Figure 7.5. However, the R^2 value tells us that the regression line does not model the variance of the data set well. Therefore by doing a hypothesis test of the slope we can more clearly determine if the variance is a function of incidence angle.

The hypothesis created is:

$$\begin{aligned} H_0 : &\quad \beta_1 = 0 \\ H_1 : &\quad \beta_1 \neq 0 \end{aligned} \tag{7.2}$$

The model tested is defined as

$$\sigma^2 = \beta_0 + \beta_1\theta \tag{7.3}$$

where σ^2 is the variance and θ is the incidence angle. The resulting statistics from the model is displayed in Table 7.1. From the t score we can reject the null hypothesis and state with a significance level of 0.001 that incidence angle impact the variance.

	Estimate (β)	Standard Error	t-score	p-value
Intercept	0.6704	0.0119	56.32	0
θ	-0.004335	0.0003698	-11.72	8.967×10^{-31}

Table 7.1: Hypothesis test statistics in 7x7 window

These results change depending on the window size of the sliding window box and the degrees of freedom. The sliding window does not use the same points twice and therefore the degrees of freedom is reduced when the window size increases. The window-sizes and t-score are shown Table 7.2.

Window-size	t-score	Degrees of Freedom
15x15	-5.515	455
30x30	-2.195	114
50x50	-0.5910	36

Table 7.2: Window-sizes and corresponding t-score

From the t-scores the null hypothesis is rejected for at a significance level of 0.001 for window-sizes of 7x7 and 15x15. However, the null hypothesis for window-size 30x30 it is rejected at a significance level of 0.05. And the 50x50 window-size the null hypothesis cannot be rejected. Due to the irregular shape of the selected region, the larger window-sizes are more likely to have empty data spots, such that the variance data points may not be too largely dependant on the window-size stated.

In the case of sea ice, getting a clear section where the sea ice is uniform proves difficult, as ridges and cracks will affect the variance greatly. As mentioned earlier, the type of sea ice also affect the dielectric properties of the ice [23]. The attempt to get clear sections large enough to calculate variance is displayed in Figure 7.6. The correlation coefficient, which is $r = -0.048394$, is small and negative which indicate that the slope has a slight downward trend. Fitting a regression line grants a slope of -0.0026284 which is shown in Table 7.3. The hypothesis test for this is displayed in Table 7.4.

From the t-score in Table 7.4, the null hypothesis that $\beta_1 = 0$ cannot be

	Estimate (β)	Standard Error	t-score	p-value
Intercept	1.0197	0.052294	19.498	0
θ	−0.0026284	0.0015181	−1.7314	0.083626

Table 7.3: Hypothesis test statistics for sea ice in 77BA in a 7x7 window

Window-size	t-score	Degrees of Freedom
7X7	-1.731	1277
15X15	-0.2097	273

Table 7.4: Window-sizes and corresponding t-score for Sea Ice

rejected at 0.05 significance level. However, as discussed in Section 3.8, mixed pixels provide a significant level of error when calculating variance and that different type of sea ice have different dielectric properties. Therefore, to more clearly represent sea ice, variance has been calculated in the scene with id 4479 where newly formed sea ice cover the entire span of incidence angles. As it is newly formed the cracks and ridges are less dominant in the image and uniform regions are more abundant. The findings for this is discussed in Section 7.1.2.

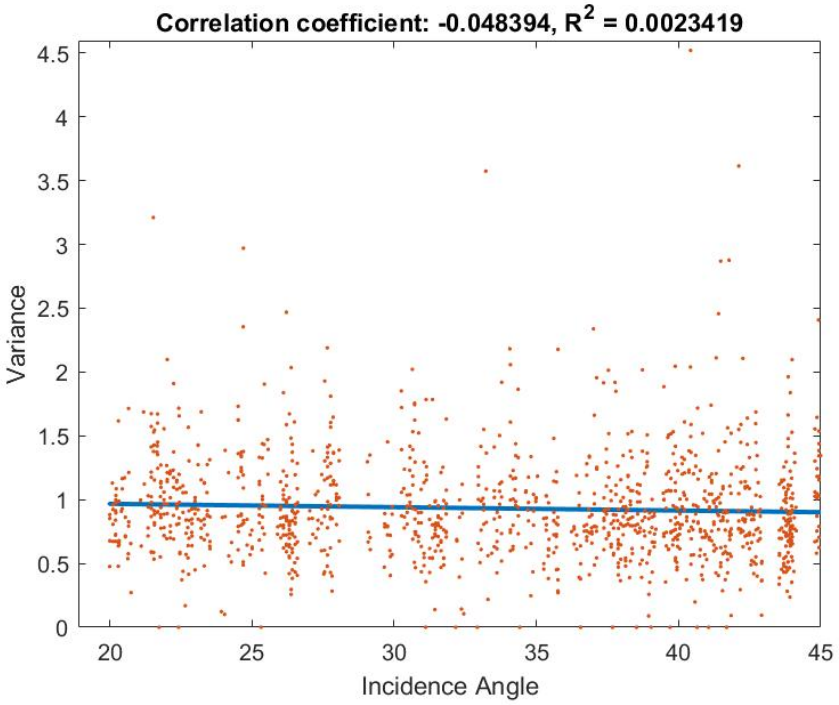
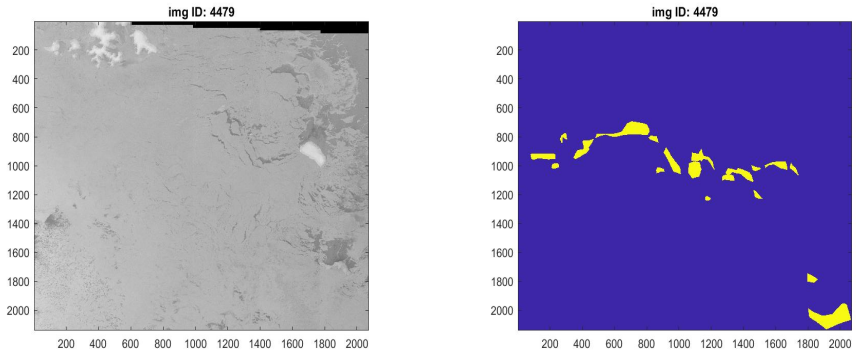


Figure 7.6: Corrected scatter plot of variance over sea ice



(a) **(b)**

Figure 7.7: a) Scene with sea ice covering all incidence angles. b) Regions selected over incidence angles.

7.1.2 Image 2

This image is used due to the selection of ice covered regions have uniform areas across the incidence angles and is illustrated in Figure 7.7a. After careful selection of regions which appear to be uniform, the variance is calculated, which is shown in Figure 7.8.

Pearsons Correlation coefficient is very small with a negative sign, indicating a sligtly negative slope, which was found for image 77BA as well. The regression line and its test statistic are found in Tables 7.5 and 7.6.

From these statistics the null hypothesis that the slope $\beta_1 = 0$, is rejected up to a significance level of 0.001 for the window size of 7x7, and 0.05 for a window size of 15x15.

	Estimate (β)	Standard Error	t-score	p-value
Intercept	0.5067	0.019347	26.189	0
θ	−0.0019525	0.0005883	−3.3189	0.00092661

Table 7.5: Hypothesis test statistics in 7x7 window

Window-size	t-score	Degrees of Freedom
7x7	-3.318948	1413
15x15	-2.212937	307
30x30	0.1104	89

Table 7.6: Window-sizes and corresponding t-score for Sea Ice in image 4479

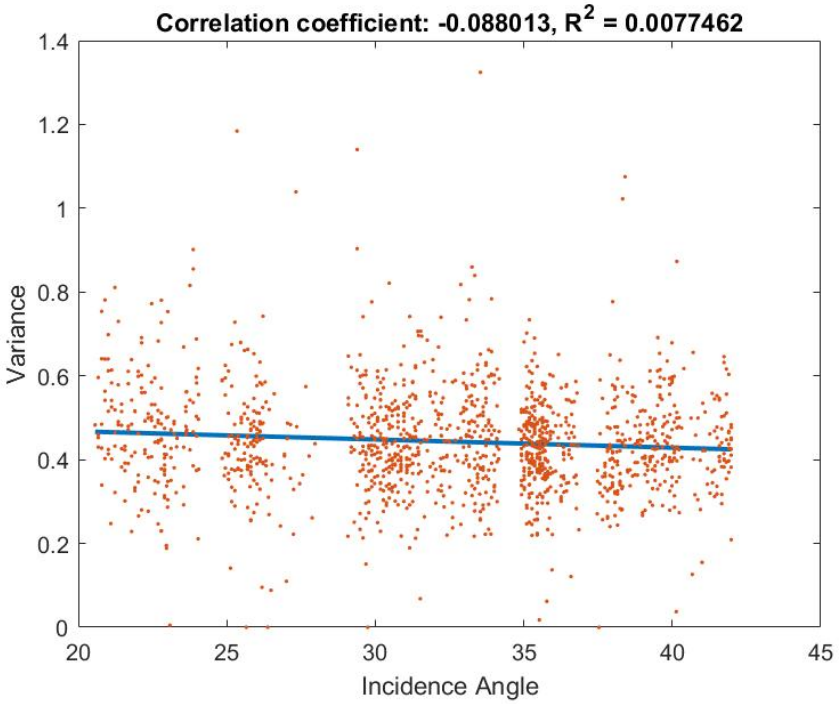


Figure 7.8: Variance over sea ice, id 4479

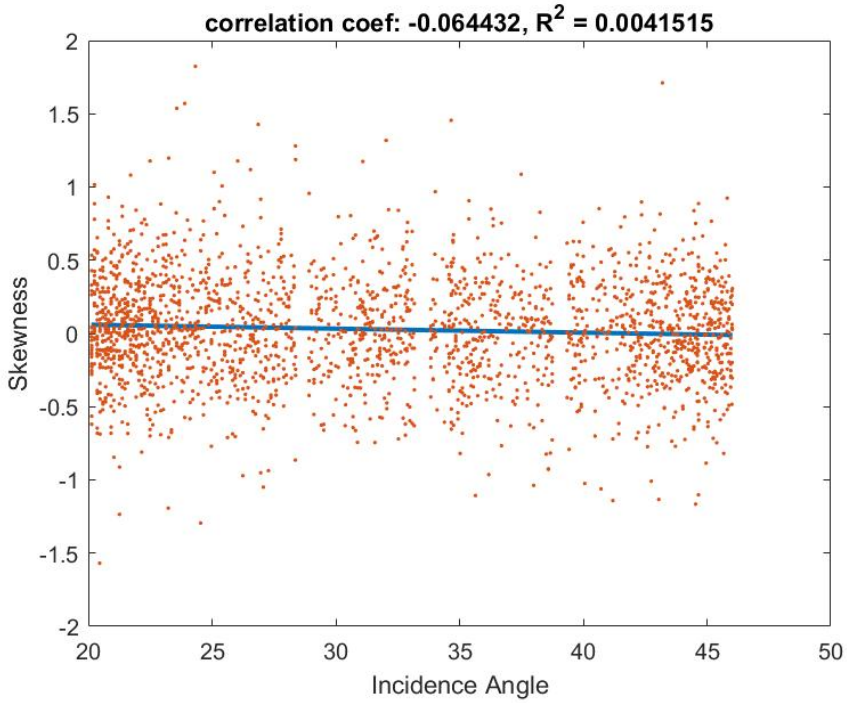


Figure 7.9: Skewness over incidence angle

7.2 Skewness

The skewness is the standarized third central moment and is calculated as described in Equation 3.33. As this is a function of the standard deviation it is natural to assume that the step observed due to the number of multi-looks would occur for skewness as well. The standard deviation of L number of looks is defined as:

$$\sigma_L = \frac{\sqrt{Var(I)}}{\sqrt{L}} \tag{7.4}$$

However, in Figure 7.9, the step cannot be seen. This would imply that the third central moment is affected by the same factor $\sqrt{L}$ as the standard deviation or that the slightly different amount of looks has no observable effect on the skewness. The third central moment is shown in Figure 7.10. As skewness is a measure of symmetry, multi-looking more times will make it more symmetrical, such that effect of more multi-looking is that it is more centered around zero. The step is more evident in the fourth central moment shown in Figure 7.12.

The skewness is centered around zero which is to be expected for a Gaussian

distribution. The correlation coefficient is slightly negative, indicating a negative slope. Running the same hypothesis test for this model on image 77BA we get a t score of −3.0647 shown in Table 7.7. From this we get a p-value of 0.0022049 which means we cannot reject the null hypothesis at a significance level of 0.001.

	Estimate (β)	Standard Error	t score	p-value
Intercept	0.11419	0.029376	3.8872	0.00010433
θ	−0.0027157	0.00088614	−3.0647	0.0022049

Table 7.7: Hypothesis test statistics in 7x7 window for skewness over ocean, id: 77BA

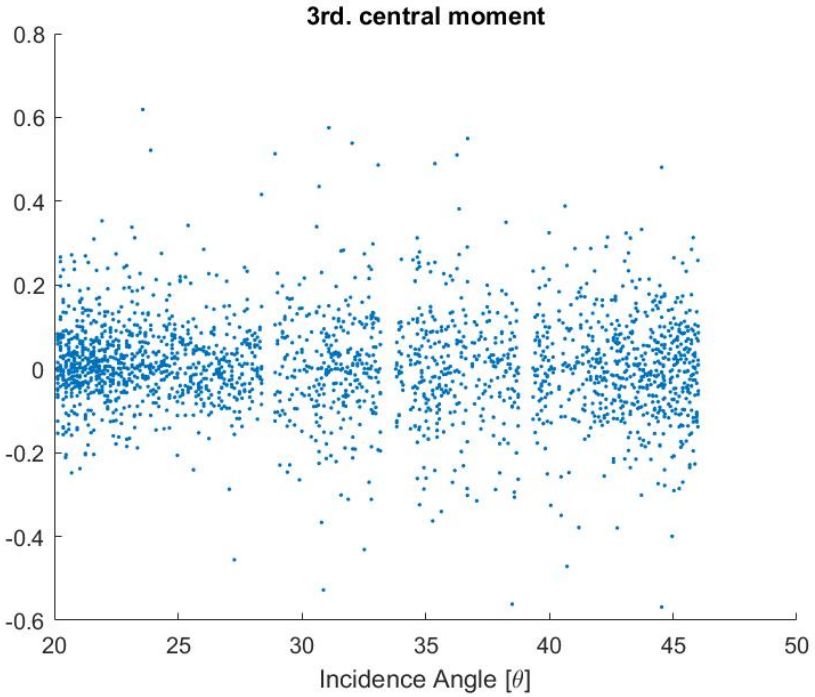


Figure 7.10: Third central moment. The values in the first sub-swath are more centered around zero.

7.3 Kurtosis

The step observed in variance can be observed in the fourth central moment as well, but as with the skewness, the normalization of the fourth central moment by the standard deviation seem to correct for this error. The kurtosis is calcuated as described by Equation 3.33.

The correlation coefficient is also slightly negative indicating a negative slope.

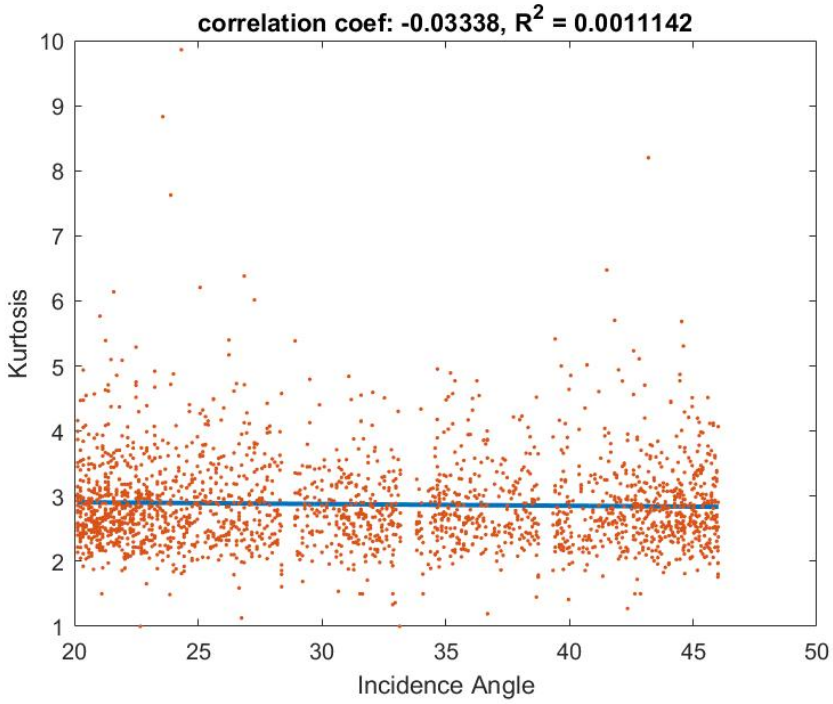


Figure 7.11: Kurtosis over ocean area in image 77BA

The hypothesis test that test if the null hypothesis $\beta_1 = 0$ grants the test statistics shown in Table 7.8. The t-score is -1.5853 granting a p-value of 0.11304 which mean that we cannot reject the null hypothesis and the kurtosis seem to not be affected by incidence angle.

	Estimate (β)	Standard Error	t score	p-value
Intercept	2.9607	0.056456	52.443	0
θ	-0.0026998	0.001703	-1.5853	0.11304

Table 7.8: Hypothesis test statistics in 7x7 window for Kurtosis over ocean, id: 77BA

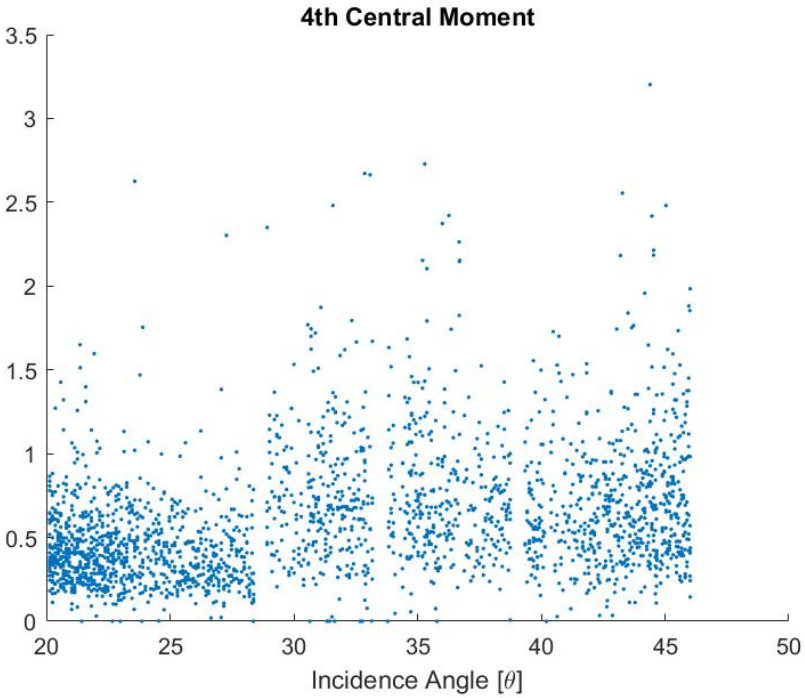


Figure 7.12: Fourth Central moment over ocean in image 77BA

7.4 Implication to modeling

When using the variance as a meassure, Sentinel-1 GRDM scenes the variance must first be pre-processed where the first swath is normalized to fit the rest with a multiplication of $\frac{18}{12}$. An overexagerated plot is shown in Figure 7.13 of the implication of not correcting for the number of multi-looks. The first sub-swath in Sentinel-1 is known for having issues [30], however as the scenes from this satellite is becoming increasingly used due to its availability, this error is important to note and may impact modeling and classification. It is further important to note that the texture parameters looked at in this thesis is unaffected by the step due to the normalization of the standard deviation of the same order. However, other texture parameters which is not normalized by the standard deviation may be impacted by the different number of looks.

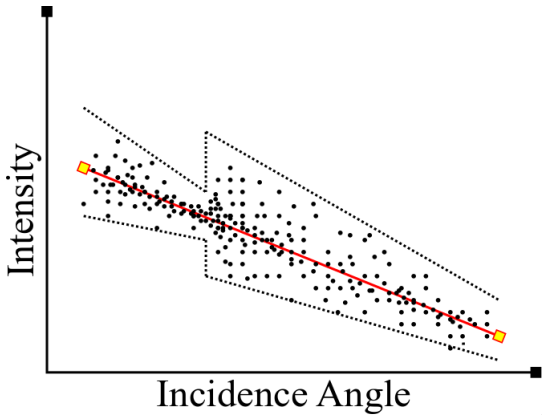


Figure 7.13: Model of intensity over incidence angle, step change of first swath indicated.

The variance has proven to be significant for ocean, and will slightly decrease with incidence angles. The rate of change is small, as illustrated by the regression lines in Figure 7.5, and has a slope of -0.004335. With incidence angles ranging from 20 to 45 degrees, the change will only differ by about 0.1084. This point is also demonstrated by increasing the window-sizes, where the statistical significance decreases as the window-sizes become larger. For window-sizes of 30x30 the null hypothesis that the slope is zero cannot be rejected at significance level of 0.025. The variance may therefore be constant when modeling as assumed by Doulgeris et al. [6] for window sizes 30x30 and larger.

The variance over sea ice in image 77BA has a lower valued slope than ocean, which is shown in Table 7.3. However, as we cannot reject the null hypothesis that the slope is zero at a 0.05 significance level, it is reasonable to let the variance be constant based on this sample.

Using the sample from image id 4479, the slope is still slightly negative. Furthermore, the null hypothesis can be rejected at 0.001 significance level for a window size of 7x7 and at 0.05 for a 15x15 window. At a value of −0.0019525 however, the change will not be very large resulting in a change of 0.0488 from 20 to 45 degrees. As with over ocean, the assumption that the variance is constant for window sizes above 15x15 for sea ice is reasonable.

From the results we see that increasing window-sizes reduce the effect of incidence angles on variance, skewness and kurtosis. When increasing the window-sizes the range of incidence angles increases which in turn reduces the resulting slope. The number of data points is also reduced and contributes to further reduce the certainty that the slope carry statistical significance. Therefore, by increasing the window-sizes the variance, skewness and kurtosis may therefore not contribute much to classification.

8

Conclusion

The results found are based on mainly two scenes. This may be too few to give a definite conclusions, but provide reasonable insight into the behaviour of speckle properties.

The results has shown that when calculating the variance from the scene in GRDM products from Sentinel-1, the first sub-swath needs to be normalized to fit the other swaths. Furthermore, the variance of speckle has shown to have a slight dependency on incidence angle where the dependency appear stronger over ocean than for sea ice.

A linear regression line is fitted to the data and a hypothesis test where the null hypothesis is that the slope is zero, can be rejected for window-sizes up to 50x50. For sea ice the slope value is smaller and show nearly no response to incidence angle. However, the null hypothesis for window-sizes of 7x7 can be rejected for the scenes used in this study with a high level of significance,but the significance level drops for larger window-sizes.

The step, which is observed in variance cannot be observed in the skewness-plot as the skewness formula is normalized to the variance and modeling with skewness is not affected by this error. Over ocean, skewness seem to have a slight dependence of incidence angle as can be described by the correlation coefficient, where the null hypothesis is rejected at a significance level of 0.001. This hold for window-sizes up to 15x15. For window-sizes 30x30 and above , the test statistics indicate that the slope might be zero.

Kurtosis is not affected by the step observed in variance either, and unlike the other moments, from the t-score show no indication of dependence of incidence angle. The null hypothesis cannot be rejected for all window-sizes.

9

Future Work

As mentioned in previous sections he step found in variance due to the different number of looks for the first sub-swath may affect some texture parameters unless normalized by the standard deviation and is something which should be investigated.

Since incidence angles show slight statistical relevance to the variance, implementing into a classification algorithm may further improve the results. However, the backscatter level need to be of a sufficient level above the noise floor as it seem to affect the variance significantly.

In this study, the HH channel is the only channel used which is due to the fact that the noise floor was too dominant to carry out further calculations on variance and higher moments with the polarimetric data at hand. Using different satellites with a lower noise floor and other pre-processing of data may also be used to confirm or verify findings.

Due to sea ice having different dielectric properties depending on the type, this may also affect the variance and the rate at which it decay. Therefore future studies may involve more accurate knowledge of the surface observed.

Appendices

A

Appendix A

A.1 Variance Near the Noise Floor

The noise floor is the sum of all sources of noise which impact the true signal of what is being monitored. In remote sensing, one such thing is the thermal noise from the imaging system. For uniform areas, the clean signal should have zero variance [8]. The signal to noise ratio is defined as

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{A.1}$$

where P_{signal} and P_{noise} is the power of the signal and noise respectively. With low signal no noise ratio, the variance within a area will then be mostly affected by noise. In Figure A.2 the variance is plotted as a function of incidence angle where the vertical lines are the region where the sub-swath changes. The observed peaks when the sub-swath changes may be cause because of several reasons. One is the interpolation of sub-swaths and the other is due to the noise floor. The noise floor of Sentinel-1 is well documented [21, 27, 28] and due to ocean having a low backscatter, the noise floor may dominate the variance.

Name:	ID
S1A_EW_GRDM_1SDH_20180830T045233_20180830T045337_023471_028E19_FDC6	FDC6

Table A.1: Name of scene with low backscatter.

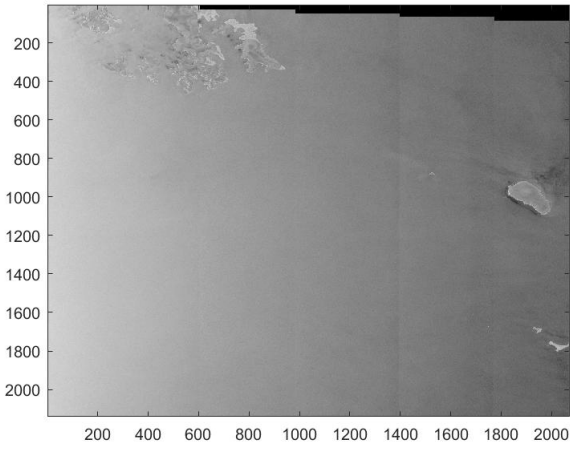


Figure A.1: Scene with only ocean with low backscatter.

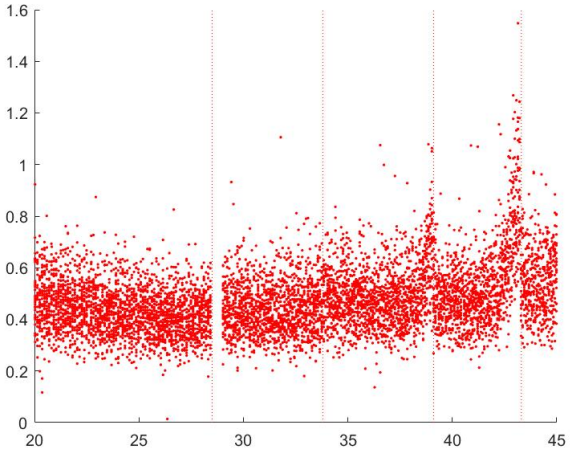


Figure A.2: Variance for low backscatter regions

B

Appendix B

Anova tables may provide some insight to the plots and the underlying variation of a model.

B.0.1 Ocean 77BA

	SumSq	DF	MeanSq	F	p-value
θ	2.5992	1	2.5992	135.74	1.8971e-30
Error	40.461	2113	0.019148		

Table B.1: Anova table for 7x7 window for Variance over ocean, id: 77BA

	SumSq	DF	MeanSq	F	p-value
θ	0.36361	1	0.36361	29.817	7.7658e-08
Error	5.6339	462	0.012195		

Table B.2: Anova table for 15x15 window for Variance over ocean, id: 77BA

B.0.2 Sea Ice 77BA

	SumSq	DF	MeanSq	F	p-value
θ	0.51882	1	0.51882	2.9977	0.083626
Error	221.01	1277	0.17307		

Table B.3: Anova table for 7x7 window for Variance over sea ice, id: 77BA

	SumSq	DF	MeanSq	F	p-value
θ	0.007092	1	0.007092	0.04398	0.8340
Error	44.02	273	0.1612		

Table B.4: Anova table for 15x15 window for Variance over sea ice, id: 77BA

Bibliography

[1] Elachi, Charles, and Jakob J. Van Zyl. *Introduction to the physics and techniques of remote sensing*. Vol. 28. John Wiley & Sons, 2006.

[2] Campbell, James B., and Randolph H. Wynne. *Introduction to remote sensing*. Guilford Press, 2011.

[3] P.Slater, *Remote Sensing, Optics and Optical systems*, Addison-Wesley Publishing Company, 1980

[4] Mäkynen, Marko, and Juha Karvonen. "*Incidence angle dependence of first-year sea ice backscattering coefficient in SENTINEL-1 SAR imagery over the Kara Sea.*" IEEE Transactions on Geoscience and Remote Sensing 55.11 (2017): 6170-6181.

[5] Lang, W., Zhang, P., Wu, J., Shen, Y., & Yang, X., "*Incidence Angle Correction of SAR Sea Ice Data Based on Locally Linear Mapping.*" IEEE Trans. Geoscience and Remote Sensing 54.6 (2016): 3188-3199.

[6] Doulgeris, A., Cristea, A., *Incorporating Incidence Angle Variation into SAR Image Segmentation*. UIT

[7] Woodhouse, Iain H. *Introduction to microwave remote sensing*. CRC press, 2017.

[8] Lee, Jong-Sen, and Eric Pottier. *Polarimetric radar imaging: from basics to applications*. CRC press, 2009.

[9] Minnaert, M., "*The Reciprocity Principle in Lunar Photometry*". The Astrophysical Joinral, vol. 93: 403, 1941

[10] Sandven, S., Johannessen, O., Kloster, K. (2006). *Sea Ice Monitoring by Remote Sensing*. 241 – 283. 10.1002/9780470027318.a2320.

[11] Fingas, M., Brown, C., *A Review of Oil Spill Remote Sensing*, Sensors 2018,

18, 91; doi:10.3390/s18010091

[12] Solberg, A. H., Brekke, C., & Husoy, P. O. (2007). *Oil spill detection in Radarsat and Envisat SAR images.* IEEE Transactions on Geoscience and Remote Sensing, 45(3), 746-755.

[13] Topouzelis, K., Singha, S., & Kitsiou, D. (2016). *Incidence angle normalization of Wide Swath SAR data for oceanographic applications.* Open Geosciences, 8(1), 450-464.

[14] Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2014). *Probability and statistics for engineers and scientists.* London: Pearson.

[15] Casella, G., & Berger, R. L. (2002). *Statistical inference*, (Vol. 2). Pacific Grove, CA: Duxbury.

[16] `https://en.wikipedia.org/wiki/F-test`, last visited: 06.11.2018

[17] De Vries, F. P. (1998). *Speckle reduction in SAR Imagery by various multi-look techniques*, (No. FEL-96-A015). FYSISCH EN ELEKTRONISCH LAB TNO THE HAGUE (NETHERLANDS).

[18] `https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-1/Instrument` European Space Agency Website, last visited: 10.11.2018

[19] Bourbigot, M., Johnsen, H., Piantan, R., *Sentinel-1 Product Definition*, 2016, Document Number: S1-RS-MDA-52-7440 S-1 MPC Nomenclature: DI-MPC-PB S-1 MPC Reference: MPC-0239

[20] Oliver, C., & Quegan, S. (2004). *Understanding synthetic aperture radar images.* SciTech Publishing.

[21] Karvonen, J., Babiker, M., Wagner, P., Aulicino, G., Parmiggiani, F.,*Sea ice classification with SAR* WP2, V1.2, 2017.

[22] Skrunes, S., Brekke, C., Jones, C. E., Espeseth, M. M., & Holt, B. (2018). *Effect of wind direction and incidence angle on polarimetric SAR observations of slicked and unslicked sea surfaces.* Remote Sensing of Environment, 213, 73-91.

[23] Jackson, C. R., & Apel, J. R. (2004). *Synthetic aperture radar: marine user's manual.*

[24] Choodarathnakara, A. L., Kumar, T. A., Koliwad, S., & Patil, C. G. (2012). *Mixed pixels: a challenge in remote sensing data classification for improving performance.* International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 1(9), pp-261.

[25] Gens, R. (2008). *Oceanographic applications of SAR remote sensing.* GI-Science & Remote Sensing, 45(3), 275-305.

[26] Valenzuela, G. R. (1978). *Theories for the interaction of electromagnetic and oceanic waves—A review.* Boundary-Layer Meteorology, 13(1-4), 61-85.

[27] Park, J. W., Korosov, A. A., Babiker, M., Sandven, S., & Won, J. S. (2018). *Efficient thermal noise removal for Sentinel-1 TOPSAR cross-polarization channel.* IEEE Transactions on Geoscience and Remote Sensing, 56(3), 1555-1565.

[28] Piantanida, R., *Thermal Denoising of Products Generated by the S-1 IPF*, Reference: MPC-0392, 2017, Nov.28.

[29] Peake, W. H., & Oliver, T. L. (1971). *The response of terrestrial surfaces at microwave frequencies* (No. ESL-2440-7). OHIO STATE UNIV COLUMBUS ELECTROSCIENCE LAB.

[30] Prats-Iraola, P., Nannini, M., Scheiber, R., De Zan, F., Wollstadt, S., Minati, F., ... & Siniscalchi, V. (2015, July). *Sentinel-1 assessment of the interferometric wide-swath mode.* In Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International (pp. 5247-5251). IEEE.