HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical Engineering

Kari Luojus

SNOW COVERED AREA ESTIMATION USING SPACE-BORNE RADAR

Master's Thesis which has been submitted for revising for the degree of Master of Science in Engineering.

Supervisor: Professor Jouni Pulliainen

Espoo, 1st of June 2004

PREFACE

The work presented in this Master's Thesis has been carried out in the Laboratory of Space Technology at Helsinki University of Technology under supervision of Professor Jouni Pulliainen.

First of all I want to thank Professor Jouni Pulliainen for supervision and for the extremely valuable support during the whole process of this work. I would also like to thank all the people working in the Laboratory of Space Technology for the great atmosphere and the inspirational working environment.

I would like to express my sincere gratitude for my family for the support and encouragement during my studies and my work. I am especially grateful for my mother Eeva-Liisa, my brother Timo, my father Tapani and my wife Mirja. All the support and encouragement has been appreciated.

I also thank EC for funding this study through the EnviSnow project.

Espoo, 1st of June 2004

Kari Luojus

Author: Kari Luojus

Name of the Thesis: Snow Covered Area Estimation Using Space-Borne Radar

Date: 1.6.2004

Number of Pages: 85

Department: Department of Electrical and Communications Engineering

Professorship: Space Technology

Supervisor: Professor Jouni Pulliainen

Abstract:

An accuracy analysis of the HUT Snow Covered Area estimation method is presented in this work. The main emphasis is to resolve the statistical accuracy of the estimation method. Other studied subjects include the effects of reference image selection, forest compensation, topography and satellite flight path on the SCA estimation method.

The method is studied by conducting the SCA estimation for 24 ERS-2 SAR intensity images for the boreal forest dominated test area in Northern Finland. The evaluation of the method is conducted by comparing the SCA estimation results with the available reference data.

Keywords:

Remote sensing, Snow Covered Area, snow melt season, radar, linear interpolation, forest compensation

Tekijä: Kari Luojus

Työn nimi: Snow Covered Area Estimation Using Space-Borne Radar

Päivämäärä: 1.6.2004

Sivumäärä: 85

Osasto: Sähkö- ja tietoliikennetekniikan osasto

Professuuri: S-92 Avaruustekniikka

Työn valvoja: Professori Jouni Pulliainen

Tiivistelmä:

Tässä työssä tutkitaan TKK:n Avaruustekniikan laboratoriossa kehitetyn lumenpeittoalan estimointimenetelmän tarkkuutta. Tärkeimpänä tavoitteena on menetelmän tilastollisen tarkkuuden selvittäminen. Toissijaisesti tutkitaan myös referenssikuvien valinnan, metsäkompensoinnin, maaston topografian ja satelliitin lentoradan vaikutuksia lumenpeittoalan estimointiin.

Menetelmää tutkitaan suorittamalla lumenpeittoalan estimointi 24:lle ERS-2 SAR satelliittikuvalle pohjois-Suomen havumetsävaltaiselle koealueelle. Menetelmän tarkkuuden arviointi suoritetaan vertaamalla saatuja estimointituloksia referenssiaineistoon.

Avainsanat:

Kaukokartoitus, lumenpeittoala, lumen sulamiskausi, tutka, lineaarinen interpolointi, metsäkompensointi

TABLE OF CONTENTS

LIST OF ACRONYMS AND SYMBOLS

DEM	Digital Elevation Model
ERS-2	European Remote Sensing Satellite - 2
ESA	European Space Agency
HUT	Helsinki University of Technology
Landsat TM	Landsat Thematic Mapper
MAE	Mean Absolute Error
PRI	Precision Image
RMSE	Root Mean Squared Error
SAR	Synthetic Aperture Radar
SCA	Snow Covered Area
SYKE	Finnish Environment Institute
WSFS	Watershed Simulation and Forecasting System

θ	angle of incidence in relation to the reference surface
κ_e	canopy extinction coefficient
λ	wavelength
σ°	radar backscattering coefficient
σ^o_{can}	forest canopy backscattering contribution
$\sigma^o_{ground,\ ref}$	reference backscattering coefficient from the snow-free ground
$\sigma^o_{ground,\ snow}$	reference backscattering coefficient from the wet snow covered ground
σ^o_{surf}	backscattering coefficient of the ground or snow layer
σ_v	canopy backscattering coefficient
τ_p	length of the transmitted pulse
χ	scalar variable describing the temporally varying scattering magnitude of forest canopy

a_0	constant coefficient, describing canopy extinction coefficient
A_{ill}	cross-area of illuminated target
B	bandwidth of transmitted signal
b_0	constant coefficient, describing canopy backscattering coefficient
c	the speed of light
E_i	incident electromagnetic field on target
E_s	scattered electromagnetic far field at the distance R from target
f	frequency
G	antenna gain
n	amount of samples
P_r	received power
P_t	transmitted power
R	distance between target scatterer and antenna
r_a	resolution element in ground range in azimuth direction
r_y	resolution element in ground range perpendicular to azimuth direction
t^2	two-way transmittivity through the forest canopy
V	forest stem volume
w_i	weighing factor for the amount of samples
x_i	estimated value
x	reference value
y_i	reference value

1 Introduction

The monitoring of our environment and nature is an invaluably important duty that needs to be addressed seriously. There are various traditional ways of collecting ground-based information on environment, but all the traditional ways suffer from serious constraints; large areas are extremely hard and sometimes even impossible to be monitored using conventional means. The gathering of information is usually expensive, and the temporal and spatial coverage is limited by the constraints of the physical observation network. The invaluable benefits of satellite remote sensing are obvious; extremely large areas can be monitored with ease. The measurements can be repeated in just daily intervals, depending on the satellite. The cost of repeated measurements depends on the satellite and the means of data distribution, but in most cases the measurements are considerably less expensive than conventional measurements yielding the same spatial or temporal coverage. Also, the delay between measurements and final data products is often smaller with satellite data than with conventional means. In operative use, the processed satellite data is usually available after a few hours from the satellite overpass.

Information on snow conditions is important for the inhabitants of the northern hemisphere. In Finland, snowy winter conditions dominate one third to almost one half of the annual weather conditions, depending on the location. The extensive gathering of information for large areas for long time periods is conducted by an extensive ground-based measurement network. These measurements can be augment and in some cases completely replaced by the means of remote sensing. Information on snow conditions is needed for environmental monitoring, weather forecasting and for hydropower industry needs. The information can be directly used in predicting and preventing floods and in optimising the hydropower industry operations. There is a distinct need from industry and the environmental point of view to observe the snow conditions during the snow melt season. This can be carried out using remote sensing data gathered using space-borne radars.

Remote sensing is based on electromagnetics, and electromagnetic wave propagation. When an electromagnetic wave is travelling through a medium, the properties of the medium determine the behaviour of the electromagnetic wave. Whenever there are changes in the medium, changes occur in the propagation of the electromagnetic wave. The two main situations that occur are the inhomogeneousities within the medium and the interface when a medium changes into another medium. The inhomogenousities in a medium cause a scattering phenomenon that is referred to as volume scattering. The magnitude of scattering is determined by the dielectric and geometrical properties of the inhomogeneousities. A transition from one medium to another medium causes a

phenomenon known as surface scattering. Surface scattering is determined by the electrical properties of the interface mediums. These two scattering mechanisms define the behaviour of electromagnetic waves. As far as remote sensing is concerned, we are speaking about studying the scatterings and reflections of electromagnetic waves caused by different objects. The most commonly known phenomenon is the reflection and scattering of sunlight, which is also the basis for human eyesight.

Remote sensing instruments are divided into active and passive instruments based on the functioning of the instrument. Passive instruments only receive electromagnetic waves; active instruments transmit and receive electromagnetic waves. With this classification the human eye is a passive optical remote sensing instrument. The passive remote sensing instruments operating at the optical and infrared regions require clear weather conditions for measuring the Earth's surface from space, and specifically, optical instruments are dependent on solar illumination. In contrast, microwave instruments have a major advantage in the ability to make observations in all weather conditions and without the illumination of the Sun. The instrument under study in this thesis is a Synthetic Aperture Radar (SAR). The synthetic aperture radar is an active microwave instrument operating at a radio frequency: 5.3 GHz in the case of ESA's ERS-2 SAR.

In this study, some of the most significant factors that affect remote sensing of snow cover are studied. The main focus will be on the remote sensing of snow during the snow melt season. The goal of the study is to asses the accuracy of a Helsinki University of Technology developed remote sensing method for Snow Covered Area estimation. The method is based on space-borne SAR intensity images. For this study the instrument under study is the ERS-2 SAR.

In Chapter 2 the snow covered area estimation method is explained.
Chapter 3 presents the reference data and the satellite data.
Chapter 4 presents the accuracy analysis of the HUT SCA estimation method.
Chapter 5 summarizes the work done.

This work has been carried out as part of the EC EnviSnow-project. One of the Envisnow-project goals was to study the statistical accuracy and the various error sources of the different snow covered area methods. This information is required for the operational use of the SCA methods.

2 Theory and Methods

The electromagnetic spectrum varies from radio waves to optical wavelengths and to gamma- and cosmic waves. The differences between the regions are defined by the frequency or the wavelength of the electromagnetic wave. They are related to each other by the equation:

$$\lambda = \frac{c}{f}, \tag{1}$$

where λ is wavelength, c is the speed of light and *f* is frequency. The visible light, or optical region is usually defined using the wavelength, and the radio frequency region is usually determined by the frequency of the wave. An example of optical remote sensing is the human eye. The human eye is a basic remote sensing instrument that functions within the spectrum of visible light. It interprets electromagnetic waves within the optical region and uses the information to generate an image. All remote sensing instruments function in the same basic way. A sensor gathers electromagnetic waves and an image formation unit, the brain in the human context, then generates the image from the sensory information.

There are several affecting factors that define the behaviour of electromagnetic waves. The most important one is the medium in the path of the propagating wave. The other affecting factors are the wavelength of the wave, the polarisation of the wave and the incidence angle that the wave is interacting with the target, i.e. the incidence angle of the wave in respect to the scattering surface. The different affecting factors each vary in respect to each other. The effect of the incidence angle is different on different wavelengths, and the properties of medium are dependent on the wavelength.

2.1 Theory of Space-borne Synthetic Aperture Radar

2.1.1 Synthetic Aperture Radar

Synthetic aperture radar is an advanced type of radar that functions by the same principle as a normal, real aperture radar; therefore the functioning of a real aperture radar is briefly explained. In this study the radar is monostatic; meaning that the transmitter and receiver are located in the same place. Radar is an active remote sensing instrument. The radar functions by sending radio frequency signals and measuring the backscattered signal from objects that are within the propagation path of the transmitted

wave. The amount of backscattering is measured as a ratio between the transmitted signal and the received signal. This ratio is known as the radar backscattering coefficient σ° and is formulated for an illuminated surface area by (Ulaby et al. 1982):

$$\sigma^\circ = \frac{P_r}{P_t}\frac{(4\pi)^3}{\lambda^2}\frac{1}{\int_A \frac{G^2}{R^4}dA} \approx \frac{P_r}{P_t}\frac{(4\pi)^3}{\lambda^2 G^2}\frac{R^4}{A_{ill}} \cong \frac{P_r}{P_t}\frac{(4\pi)^3}{\lambda^2 G^2}\frac{R^4}{r_a r_y}, \quad (2)$$

where P_r is the received power, P_t is transmitted power, λ is wavelength, R is distance between the target scatterer and the antenna, G is the antenna gain and A_{ill} is the cross-area of the illuminated target, which can also be expressed using the resolution elements r_a and r_y, which are explained below. The backscattering coefficient can also be formulated using the electromagnetic field variables instead of the transmitted and received power levels (Ulaby et al. 1982):

$$\sigma^\circ = \frac{4\pi R^2 \left|E_s\right|^2}{A_{ill}\left|E_i\right|^2}, \quad (3)$$

where E_s is the scattered electromagnetic far field at the distance R from the object, E_i is the incident electromagnetic field on the object, R is the distance between the target scattering surface and the antenna, and A_{ill} is the cross-area of the illuminated target.

The synthetic aperture radar functions by the same basic principles as normal radars, see e.g. (Ulaby et al. 1986). The radar transmits a signal and receives it to obtain the backscattering coefficient. But instead of taking only a single measurement from each target, the synthetic aperture radar takes multiple measurements from each target along the flight path. Measurements of a single target are spread over a certain distance along the flight path, which is called the synthetic aperture. When every target is covered by the synthetic aperture, the resolution acquired is equal to the resolution acquired by a real aperture antenna with the dimension of the synthetic aperture. When the measured backscattering is analysed, the different targets are identified from each other by the difference in their backscattering signatures. As illustrated in Figure 1, the different image elements in range direction are identified by the difference in wave propagation. Since the targets that are closer to the radar have a shorter distance, the backscattering from these targets are received earlier in the time domain than the backscattering from the more distant objects. The identification in the azimuth direction is made by the difference in the targets doppler frequency. Since the radar is moving in respect to the targets, the backscattered signatures from the targets are doppler shifted. The amount of doppler shift is proportional to the azimuth distance of the object, and this is used in the azimuth range interpretation, see Figure 1.

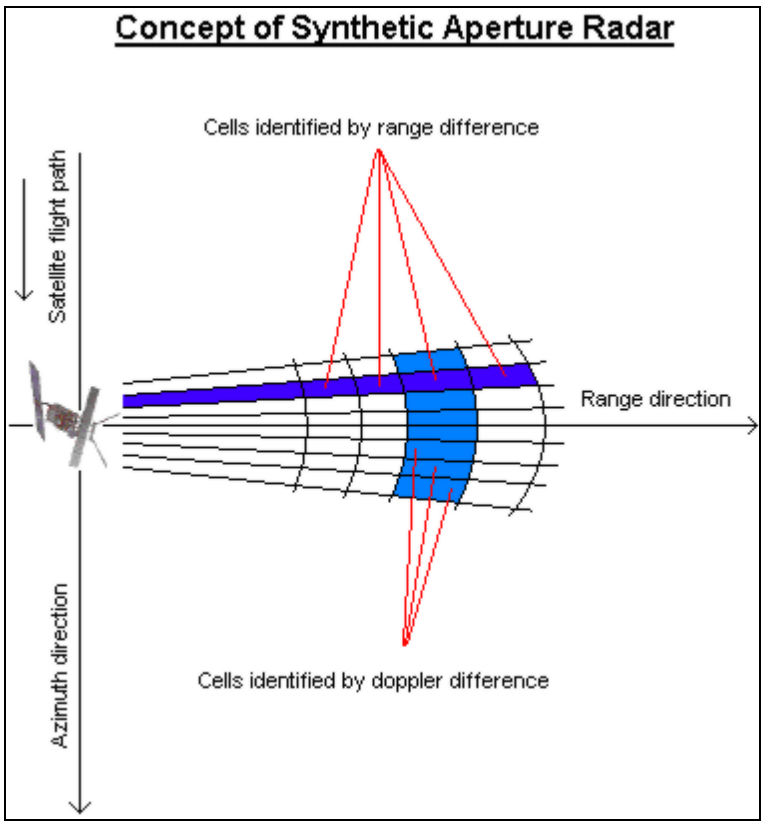

Figure 1. Illustration of the concept of Synthetic Aperture Radar

The dimensions of resolution elements are defined in Synthetic Aperture by the length of the transmitted signal and the physical length of the antenna. The resolution element in ground range in the azimuth direction r_a is given by (Ulaby et al. 1986):

$$r_a = \left(\frac{a_{hs}}{a_{hr}}\right)\frac{l}{2} \approx \frac{l}{2}, \tag{4}$$

where the ratio a_{hs} to a_{hr} is the relation between the aperture illumination taper factors, which is typically considered as 1 and l is the physical length of the antenna. The resolution element in ground range perpendicular to the azimuth direction r_y is given by (Ulaby et al. 1986):

$$r_y = \frac{c\tau_p}{2\sin\theta} = \frac{c}{2B\sin\theta}, \tag{5}$$

where c is the speed of light, τ_p is the length of the transmitted pulse, θ is the local incidence angle in relation to the reference surface and B is the bandwidth of the transmitted signal.

One problem in measuring the backscattering coefficient by SAR is the variation in the size of illuminated area caused by topographical variations. Topographical formations, hills and slopes cause the illuminated area to be either larger or smaller than the area for flat surface. This directly affects the backscattering coefficient, and must be considered in image processing. Topographical variations can also cause shadowing, foreshortening and layovering of image elements, see Figure 2. Foreshortening is occurring when a slope is facing towards the radar, shadowing occurs when a part of the backside of a slope is so steep that it cannot be seen by the radar. Layovering happens when a higher point of ground is closer to the radar than a lower point and is falsely interpreted as a closer image element. Thus the changes in topography affect the measured backscattering coefficients, and this is addressed in topography correction phase of SAR image processing.

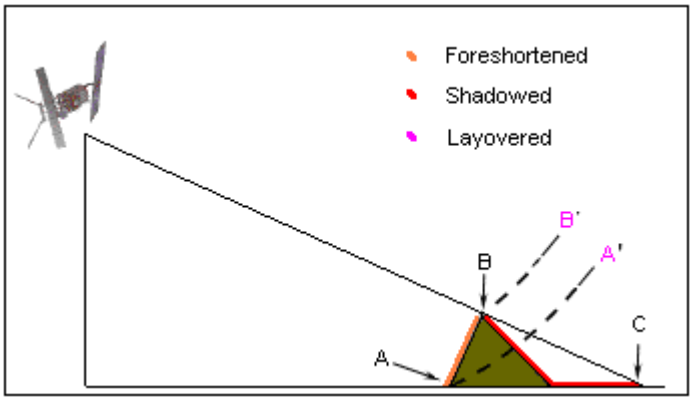

Figure 2. Illustration of the topographical errors caused by the imaging geometry. Foreshortening is seen when a slope is facing the radar, the area between A and B is foreshortened. Shadowing is seen when a slope is so steep that the surface cannot be seen, the area between B and C is shadowed. The layover happens when the apparent distance is shortened by a topographical formation. The distance of A and B are interpreted as A' and B' in the radar, thus B is falsely interpreted as a closer image element than A.

The backscattering coefficient is a descriptive variable that is straightforward to derive from the principles of radar functioning, and is therefore widely used in radar processing. The intensity images that are used in this study are directly derived from the measured backscattering coefficient, so the intensity stands for the intensity of the backscattering.

The radar images are formed by measuring different targets and by creating a spatial presentation from the measured backscattering data. The size of each image element is dependent on the antenna size and the radar implementation and the resolution of the overall image is determined by the size of image elements. The challenge with a real aperture radar is to obtain a good resolution for the image elements from distant

measurements, such as with satellite instruments. Since the size of the antenna in satellite instruments is limited, the resolution problem has to be solved by the radar implementation. The synthetic aperture radar is one solution for the resolution challenge.

2.1.2 The European Remote Sensing Satellite, ERS-2

Space-borne radars are in most ways similar to radars used in airborne applications. The biggest differences are found in the imaging geometry. The measuring distance in satellite instruments is vast when compared to airborne radars. The ERS-2 remote sensing satellite has an average altitude of 785 kilometres. With a looking angle of 23° this results in an 862.5 km average range to earth surface. The other significant difference is the orbital velocity of the satellite. Since the satellite is on an altitude of 785 km it has a velocity of nearly 26900 km/h. These two factors have to be considered in the radar design. From the operational point of view, the resulting factors that are of significance are the acquired resolution and the length of the satellite repeat cycle.

The resolution of the ERS-2 synthetic aperture radar is approximately 30 meters in both range and azimuth directions. The pixel spacing for the used PRI images is 12.5 meters. The frequency of the SAR is 5.3 GHz. The transmitted and received polarisations are vertical. The size of an image is typically 100 km x 100 km. The incidence angle varies from 19.35 to 26.50 degrees, and is at the middle of the swath 23 degrees.

The velocity of the satellite makes it circle the earth in 1 hour and 40 minutes, thus making 14 orbits and a fraction every day. The satellite is on a polar sun-synchronous orbit. This means that the satellite overpasses occur during the same local time at every location on earth. The repeat cycle of the satellite is 35 days; meaning that the satellite is in the exact same position above the earth every 35th day. In 35 days the satellite makes a total of 501 orbits. Although the repeat cycle is quite long, the satellite can usually measure the same area every four days, but with a difference in imaging geometry between the images. In addition to the synthetic aperture radar, the satellite also has a number of other remote sensing instruments aboard, see Figure 3. (ERS Missions)

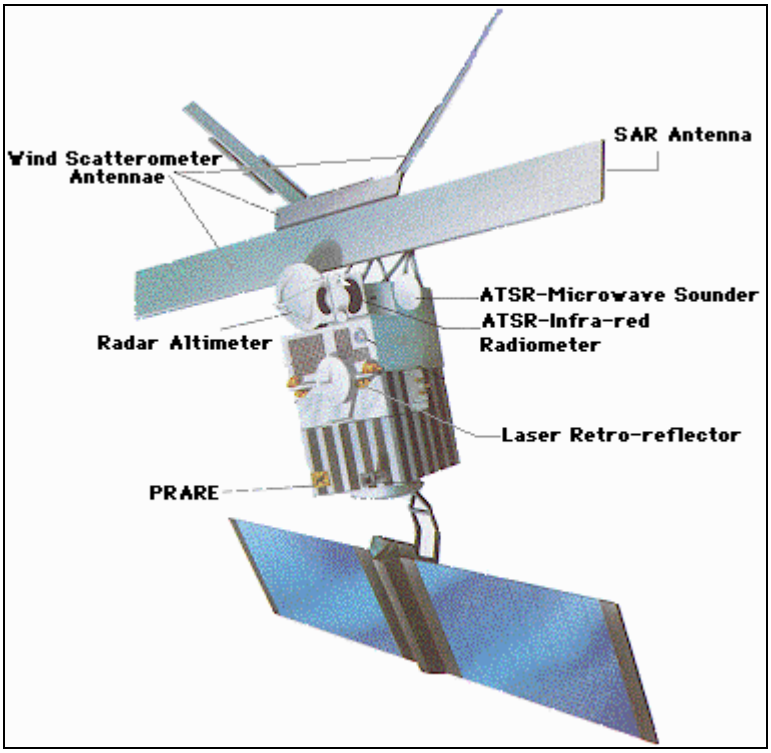

Figure 3. The ERS-2 Satellite. Image by ESA.

2.2 Remote Sensing of Snow and Snow Covered Area

Snow covered area (SCA) is a regional variable used in hydrology that indicates the percentage of the ground that is covered by snow. SCA is 100% when the ground is totally covered by snow, and 0% when the ground is completely snow-free. When the snow melt season starts the SCA is 100% and when all the snow has melted the SCA is 0%. The SCA value is used in monitoring the progress of the snow melt.

The capabilities of Space-borne SAR in snow monitoring has been widely studied (Rott 1984, Koskinen et al. 1994, Piesbergen at al. 1995, Guneriussen et al. 1996, Koskinen et al. 1997, Nagler and Rott 1998, Koskinen et al. 1999). The suitability of SAR for snow monitoring has been confirmed, and further studies are been made for snow monitoring algorithm development. The space-borne microwave radars are found to be feasible for snow monitoring especially during the snow melt season. However, current space-borne SAR systems are also known to have several limitations on snow monitoring applications (Koskinen et al. 1994, Piesbergen et al. 1995, Koskinen et al. 1997, Gunneriussen et al. 1996). The snow monitoring is based on changes in the backscattering properties of the snow pack and ground surface. The backscattering changes are typically well-observed between wet snow and dry snow, and between wet snow and bare ground, thus resulting in suitability for snow melt monitoring.

The discussion of snow monitoring from here on is considered for 5.3 GHz space-borne SAR, such as the instrument on board ERS-2 satellite; there are differences on backscattering on different frequencies and polarisations which are not discussed here. The snow covered area evaluation is based on the change of radar backscattering received from the snow and ground layer. During the melting period the snow layer is wet. The level of backscattering from wet snow is generally lower than that of dry snow or bare ground. The backscattering coefficient typically decreases with increasing snow wetness (Stiles et al. 1980, Ulaby et al. 1986, Fung 1994, Strozzi et al. 1997). When the melting season is progressing, the snow layer melts and bare ground is revealed from underneath. Bare ground typically has a much higher level of backscattering when compared with wet snow; so as the amount of bare ground increases, the amount of backscattering also increases. When all the snow has melted and only bare ground is visible, the backscattering level is highest. This temporal behaviour of backscattering is utilised for SCA evaluation, see Figure 4.

The HUT method for snow covered area estimation is based on the observations mentioned above: the backscattering level of wet snow and the level for snow-free ground are measured for a certain area. The levels of backscattering from wet snow and snow-free ground are marked as reference values for the respective situations. It is assumed that the amount of backscattering is near-linearly dependent on the amount of snow covered ground. Then by using the two reference images, one for wet snow and one for snow-free ground, the amount of snow covered area can be determined for an image taken during the snow melt season. The detailed method and the variables that affect the SCA estimation accuracy are described in the following chapters.

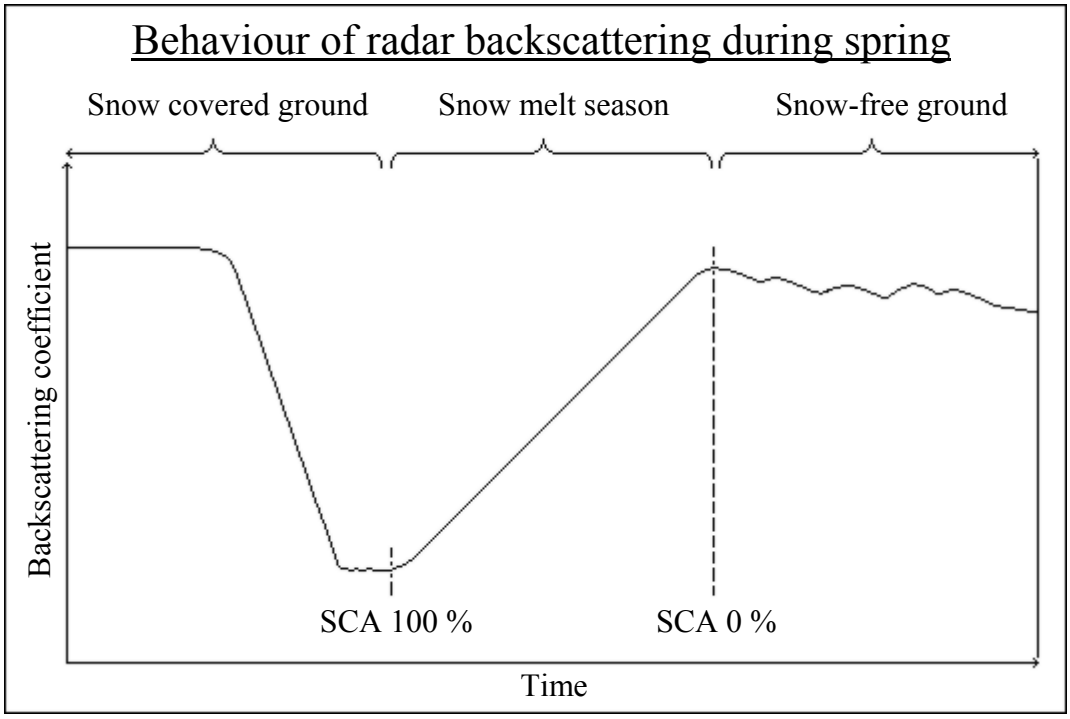

Figure 4. The behaviour of backscattering coefficient during the snow melt season.

2.3 HUT Method for Snow Covered Area Estimation

The Snow Covered Area (SCA) estimation method developed at HUT is a two step procedure for obtaining the SCA estimate from SAR intensity images during the snow melt season. The procedure requires the knowledge of the forest stem volume distribution of the target area and two reference images are needed for the SCA estimation. One reference image is required for describing the wet snow situation at the beginning of snow melt season and one for characterising the snow-free situation at the end of the snow melt season. The HUT method consists of two steps. The first step is the forest canopy compensation which is done using forest stem volume information. The second step is a pixel-wise linear interpolation algorithm that uses the reference images to calculate the SCA estimate.

2.3.1 Forest canopy compensation

The radar backscattering coefficient is a sum of backscattering signatures from different contributions. Some contributions are caused by the snow and ground, some are caused by the backscattering from forest and forest canopy. The contribution from forest canopy is not related to the SCA and it is therefore a source of error. This error can be minimised by using the forest compensation method. The method is based on the boreal forest backscattering model developed at HUT (Pulliainen et al. 1999, Pulliainen et al. 2001). Using the forest backscattering model the contribution of backscattering of forest canopy is calculated, and this contribution is reduced from the total observed backscattering coefficient. The corrected value is then used in the linear interpolation phase. The boreal forest backscattering model describes the backscattering coefficient (σ^o) as a function of stem volume:

$$\begin{aligned}\sigma^\circ(V,\chi) &= \sigma^\circ_{surf}\cdot\exp\left(\frac{-2\kappa_e(A(\chi))\cdot V}{\cos\theta}\right)+\frac{\sigma_V(B(\chi))\cos\theta}{2\kappa_e(A(\chi))}\left[1-\exp\left(\frac{-2\kappa_e(A(\chi))\cdot V}{\cos\theta}\right)\right] \\ &\equiv \sigma^\circ_{surf}\cdot t(V,\chi)^2+\sigma^\circ_{can}(V,\chi)\,,\end{aligned} \qquad (6)$$

where V is the forest stem volume [m³/ha], A(χ) defines the canopy extinction coefficient κ_e [1/m³/ha] and B(χ) defines the canopy backscattering coefficient σ_v [1/m³/ha], χ is a temporally varying scalar variable. σ^o_{surf} is the backscattering coefficient of the ground or snow layer, and θ is the angle of incidence. The first term of (6) defines the backscattering contribution from ground or snow layer σ^o_{surf} and the two-way transmittivity through the forest canopy t^2. The second term of (6) σ^o_{can} defines the forest canopy backscattering contribution.

The values for κ_e and σ_v are dependent on weather conditions and are obtained by:

$$\kappa_e(A(\chi)) = a_0 \cdot \chi \quad (7)$$
$$\sigma_v(B(\chi)) = b_0 \cdot \chi^2 \quad (8)$$

where constant coefficients $a_0 = 2.78 \cdot 10^{-3}$ ha/m³ and $b_0 = 9.99 \cdot 10^{-4}$ ha/m³. These values are derived in dry summer conditions where $\chi \approx 1$. Since the variable χ is dependant on weather conditions it needs to be solved separately for each observed case.

Using (6) – (8) the backscattering contribution for forest canopy can be solved when observations from different stem volume classes are made. The solving is done by non-linearly fitting the observed backscattering coefficients to the model, where the parameters χ and σ^o_{surf} are the variables to be optimised. For the optimisation the backscattering for each stem volume class needs to be calculated. The method used here is to gather information from a selected area thus receiving averaged values for each stem volume class and then fitting the averaged values with the model. The minimisation problem is written by:

$$\min_{\chi,\sigma^\circ_{surf}} \sum_{i=1}^{n} w_i \cdot \left(\left\langle \sigma^\circ_{OBSERVED,i} \right\rangle - \sigma^\circ(V_i, \chi, \sigma^\circ_{surf}) \right)^2 \quad (9)$$

where n is the number of volume stem classes, $\left\langle \sigma^\circ_{OBSERVED,i} \right\rangle$ is the observed backscattering coefficient for the stem volume class V_i. and σ^o is the calculated backscattering coefficient. The weighing factor w_i is used if the number of observations for the different stem volume classes are unevenly distributed.

When the minimisation has been successfully carried out and the parameters for the backscattering coefficient are known, the solved variables σ^o_{surf} and χ can be used with (6) to calculate the backscattering coefficient without the contribution from the forest canopy, corresponding to the backscattering with stem volume of 0 m³/ha. This is the forest compensated backscattering coefficient that can be used in the linear interpolation phase of SCA estimation, see Figure 5.

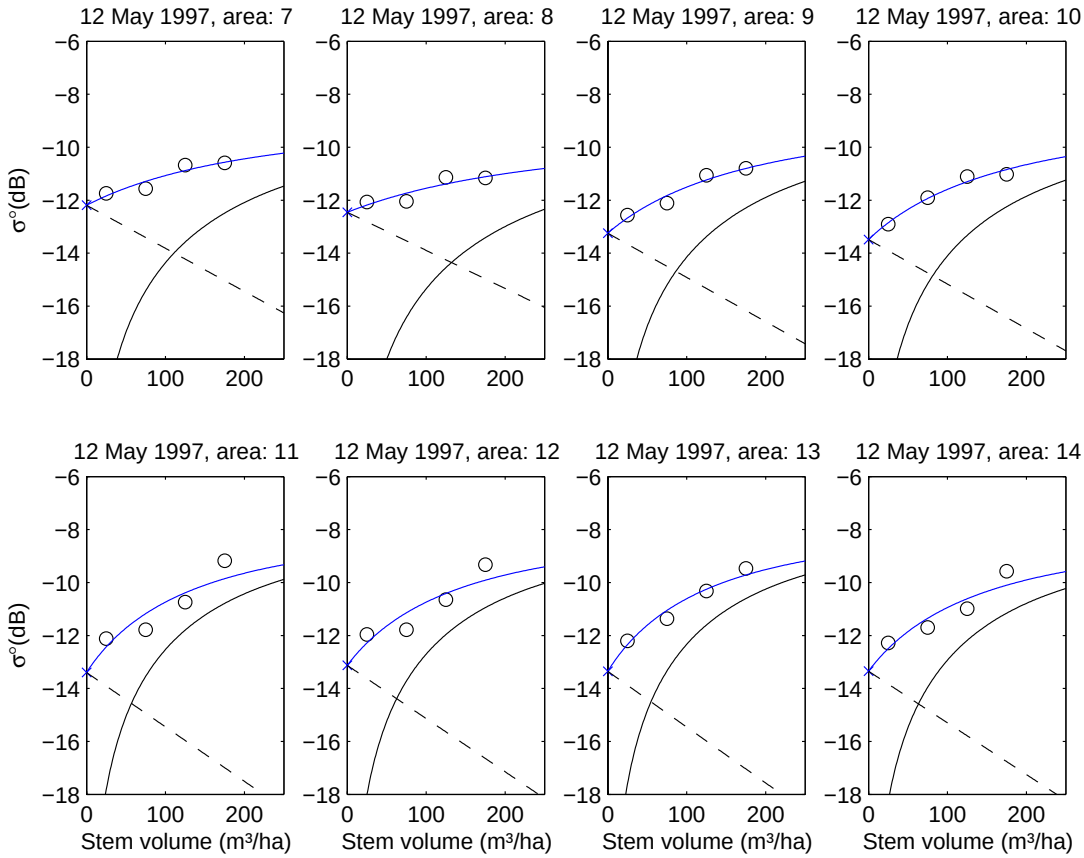

Figure 5. An example of the forest compensation. Contribution from forest canopy is drawn on solid black line, ground contribution is drawn on dashed black line and the total backscattering coefficient is drawn on solid blue line, the observed backscattering coefficients are drawn with circles. The drainage area number is also indicated.

2.3.2 Linear interpolation

The linear interpolation algorithm uses a pixel-wise comparison of the observed backscattering coefficient to the backscattering coefficients from the two reference situations (Koskinen et al. 1997). The outcome is the Snow Covered Area estimate. The estimate is a percentual figure that represents the amount of ground covered by snow. The estimate is obtained by:

$$SCA = 100\% \cdot \frac{\sigma^{\circ}_{surf} - \sigma^{\circ}_{ground,\,ref}}{\sigma^{\circ}_{snow,\,ref} - \sigma^{\circ}_{ground,\,ref}}, \qquad (10)$$

where σ^{o}_{surf} is the observed or estimated surface backscattering coefficient, $\sigma^{o}_{ground,\,ref}$ is the reference backscattering coefficient from the snow-free ground and $\sigma^{o}_{ground,\,snow}$ is the reference backscattering coefficient from the wet snow covered ground. In the case of forest compensated SCA evaluation, the used reference backscattering coefficients

are also forest compensated values. The linear interpolation algorithm is pixel-wise, but because of the speckle in SAR images the algorithm is more accurate when larger areas are averaged and the estimation is done for the averaged values. In this study the SCA estimates were calculated for drainage areas with several thousands of pixels.

In the HUT method, SCA estimates are calculated for two different cases. One estimate is obtained for the open areas, and one estimate for the forested areas. The open area estimation is obtained by calculating the average from the open area pixels and using the linear interpolation algorithm directly. The estimate for the forested areas are calculated by first using the forest compensation algorithm and then calculating the linear interpolation. So the whole method is using one step for open areas and two steps for forested areas. The outcome results in different SCA estimates for open and forested areas which is also desired by the operative hydrological models. In this study the benefits of forest compensation are analysed, and the overall accuracies for both the open and forested area estimates are calculated.

2.4 Variables Affecting the Accuracy of SCA Estimation

There are several factors that affect the accuracy of SCA estimation. Some of them are caused by the remote sensing nature of the estimation method and some variables are caused by the behaviour of the estimation method itself. The variables that are caused by the nature of remote sensing are related to the backscattering phenomenon. The backscattering coefficient of a certain image element is a sum of different backscattering contributions. The most important factors in the remote sensing of snow are listed here (Ulaby et al. 1986):

- Backscattering from the snow-air interface
- Backscattering from the snow layer
- Backscattering from the snow-ground interface
- Backscattering from the ground layer
- Backscattering from the forest canopy (on forested areas)

It can be concluded that the decisive backscattering effects are generated in the snow, ground and vegetation layers and in the interfaces between the different mediums. Since the goal of SCA estimation is to derive the amount of snow on the ground from the observations, the other contributors can be regarded as sources of error. There are also a number of other factors that affect the remote sensing of snow, which can be considered as constants in this study. They include the operating frequency of the radar, the incidence angle of the radar, the polarisation and other factors that are related to the electromagnetic wave characteristics. These factors can be regarded as constants, since the SCA estimation is conducted using images taken with the same instrument using the

same configuration for image acquisition. The incidence angle of the radar for each image element varies between different images. This variation is remarkable for large images, but for the ERS-2 PRI images used here, this variation is quite small and it is regarded as a constant. However the variation between images from ascending node and descending node is more obvious and thus that issue is covered in this study. The variables affecting the SCA estimation accuracy that are discussed in this study are considered in the next chapters in more detail.

2.4.1 Snow and ground layers

The key element in remote sensing of snow covered area is the behaviour of the observed variables: the snow and the ground. The backscattering from the observed variables dominates the total backscattering coefficient, so the changes in the observed variables are also the greatest possible error sources in the SCA estimation. When the snow or weather conditions differ from the conditions assumed by the SCA theory the accuracy of the SCA estimation can dramatically decrease. The backscattering from snow and ground can be divided into volume and surface scattering. The volume scattering effects are caused by the snow layer and the ground layer. The surface scattering is caused by the different interfaces: the air-snow, the air-ground and the snow-ground interfaces.

The dominating factors for the behaviour of the backscattering during the snow melt season are the snow wetness, the snow surface roughness and the ground wetness (Ulaby et al. 1986, Fung 1994, Shi and Dozier 1995). During the snow melt season the snow layer and the ground surface are expectedly wet. However if there are freezing periods during the snow melt season and the wetness of snow decreases, this results in increase in the backscattering coefficient and the accuracy of the SCA estimation is decreased. Therefore it is important to follow the overall weather conditions when calculating the SCA estimates. The backscattering from bare ground is dependent on the ground wetness, so the SCA estimation is accurate only during the snow melt season. The estimation ceases to function when all the snow has melted and the ground starts to dry. The changes in ground layer are minor for the backscattering coefficient during the snow melt period in comparison to the changes in the snow parameters.

2.4.2 Vegetation and forest canopy

The vegetation and forest canopy affect the backscattering coefficient by increasing the amount of volume scattering contribution. Vegetation and forest canopy also decrease the amount of backscattering from the snow and ground surfaces, thus decreasing the amount of information received from the observed variable, the amount of snow, see equation (6). The increase of backscattering caused by vegetation and forest canopy is

greater than the decrease of the backscattering from snow and ground layers, so that the total backscattering coefficient is increased. If this effect is neglected in the SCA estimation, the estimates received from forested areas are biased in relation to the amount of forest and vegetation in the area. This causes the SCA estimates for heavily forested areas to be more biased and results in more inaccurate estimates. The effect caused by forest canopy can be compensated if the geographical forest density distribution is known. The effect of surface vegetation is remarkably smaller when compared with the effect of forest canopy. Also the amount and distribution of the surface vegetation is generally not known, so therefore it is neglected in the SCA estimation.

The solution for the forest canopy compensation is described in Chapter 2.3.1. This method is used in determining the SCA estimates. The forest compensation improvement is analysed by comparing the SCA estimates calculated with the forest compensation to the SCA estimates calculated without the forest compensation.

2.4.3 Topographical variations

Topographical variations on earth are the source of imaging deformations that are encountered on all satellite images. The effects caused on SAR images are considered here. The near sides of hills are facing the radar and are shortened and the intensity of the backscattering from these are increased. The far sides of hills seem longer and the backscattering is weaker than from the surrounding surfaces. The apparent location of topographical formations are also shifted within the image. The objects that are on higher ground seem to be closer to the radar and the objects on lower terrain seem to be further away that they actually are. There are no problems when comparing two satellite images taken from exactly the same orientation, but when comparing images with different imaging geometry the differences between two images can be considerable. The ERS-2 satellite is taking SAR images to the right side of the flight path. Thus an image taken during the ascending node is seen from west and an image taken from the descending node is seen from east. The imaging geometry in correspondence to topography of imaged area is completely different from the different nodes. The opposite sides of the topographical formations are intensified or darkened, see Figure 6. This is a significant error source for the SCA estimation and needs to be addressed.

The effects for topographical image variations are minimised by the image processing software which is also used to geo-rectify the images. The image processing is done using a Digital Elevation Model DEM, thus the topographical information is used to correct for the errors.

Even though the images are corrected for topographical errors, there are still differences between images taken from different nodes, ascending or descending. The variation caused by topography is analysed by calculating the SCA estimates in cases with similar imaging geometries and comparing these to SCA estimates calculated with images on different imaging geometries. Also the amount of topographical variations within the test areas are estimated, so the SCA estimation accuracy can be assessed in relation to topographical variations.

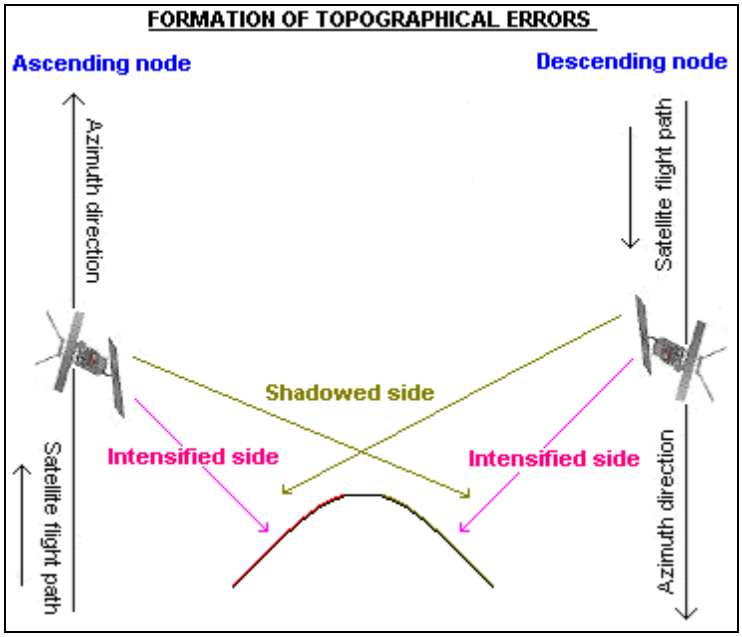

Figure 6. The formation of topographical errors. Different sides of topographical formations are intensified and shadowed, depending on the direction of satellite overpass.

2.4.4 Reference image selection

The HUT SCA estimation method uses a linear interpolation algorithm. In the linear interpolation the investigated image is compared with two reference images and the SCA estimate is therefore dependent on the reference images. The reference images need to be selected so that the first image describes typical backscattering conditions at the beginning of snow melt season and the other describes the end of the snow melt season. Since there are a number of ways to choose the reference images and the images affect the SCA estimation, it is important to study the effects of reference image selection. In this study different approaches to reference image selection are considered, and the effect of image selection to SCA estimation accuracy is analysed.

The reference image for the beginning of the snow melt season needs to be chosen so that the ground is totally covered by snow and the snow layer is wet. This results to a relatively low level of backscattering. The reference image for the end of the snow melt

season needs to be chosen so that all the snow has melted but the ground is still wet. This results to a relatively high level of backscattering.

The first goal of the reference image selection study is to find the images that yield the best SCA estimation results. When the best reference images have been acquired the next goal is to derive an universally usable method for the reference image selection. This is important for the overall usability of the HUT SCA estimation method.

An other important aspect is to analyse the variation in SCA estimates which is caused by varying backscattering signatures of reference images. This can be determined from the variation of the SCA estimates calculated using all the different reference image combinations and by comparing this variation with the results acquired with the best reference image combination.

2.5 Methods for Accuracy Analysis

The accuracy of SCA estimation is analysed by statistically comparing SCA estimates with reference data. The main variables that are examined include: root mean squared error (RMSE), mean absolute error (MAE), bias and correlation coefficients. The reference data is considered to be accurate concerning this study.

The root mean squared error is the key variable describing the overall statistical accuracy. It is calculated by:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x)^2}{n}}, \tag{11}$$

where x_i is the estimated value, x is the reference value and n is the amount of samples. The mean absolute error is acquired by:

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |x_i - x|}{n}, \tag{12}$$

where x_i is the estimated value, x is the reference value and n is the amount of samples. The bias is calculated by:

$$\mathrm{bias} = \frac{\sum_{i=1}^{n} x_i - x}{n}, \tag{13}$$

where x_i is the estimated value, x is the reference value and n is the amount of samples.

Linear-correlation coefficient used in analysis is calculated by (Bevington and Robinson 1992):

$$r = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\left[n\sum x_i^2 - \left(\sum x_i\right)^2\right]^{1/2}\left[n\sum y_i^2 - \left(\sum y_i\right)^2\right]^{1/2}}, \quad (14)$$

where n is the amount of samples, x_i is the estimated value and y_i is the reference value.

3 Test Area, Reference Data and Satellite Data

The SCA estimation was carried out for 24 ERS-2 SAR intensity images. Images were acquired for the snow melt period, and they include the years: 1997, 1998, 2000, 2001 and 2002. All the images were chosen so that they cover the selected test area in Northern Finland, the Lokka region. The test site consisted of fourteen different drainage basins, with a total area of 1160 km². The SCA estimates for the different SAR images were calculated separately for each drainage basin. The SCA estimation results were compared with reference data gathered through weather station observations and by using the Finnish Environment Institute SYKE's Watershed Simulation and Forecasting System WSFS (Vehviläinen, 1994).

The selected test site situates in Northern Finland, between Sodankylä and Lokka, see Figures 7 and 8. The test site is typical of the boreal forest: relatively sparse coniferous forest, with some open areas.

Figure 7. Geographical location of the Lokka test area.

3.1 The Lokka Test Area

The Lokka test area is divided into fourteen distinctive drainage basins that were analysed separately, see Figure 8. The discrimination between the different drainage basins is based on Finnish Environment Institute, SYKE's WSFS model. The different basins were analysed independently, and the information describing each area is presented in Table 1.

Figure 8. The fourteen drainage basins of the Lokka test area.

Table 1. The Land-usage distribution of the drainage basins.

Drainage basin	Amount of pixels	Location	Open area	Forested area	Open bog	Forested bog
1	12712	NW	5.0 %	55.9 %	17.4 %	21.2 %
2	15046	NE	10.5 %	70.6 %	7.2 %	11.7 %
3	5225	N	1.4 %	56.6 %	17.0 %	24.4 %
4	3155	NE	6.4 %	65.9 %	6.9 %	20.7 %
5	13334	N	2.6 %	60.4 %	17.1 %	18.9 %
6	10203	NW	3.3 %	55.2 %	11.8 %	28.4 %
7	6078	E	2.8 %	72.4 %	9.4 %	15.4 %
8	5833	W	2.6 %	46.6 %	22.9 %	26.0 %
9	3643	C	1.2 %	56.9 %	21.8 %	19.9 %
10	6544	C	3.4 %	71.0 %	14.9 %	10.3 %
11	7696	C	2.0 %	62.1 %	19.0 %	16.5 %
12	13970	S	4.5 %	63.0 %	17.7 %	13.8 %
13	6898	S	5.1 %	84.4 %	6.5 %	4.0 %
14	5790	S	3.9 %	49.8 %	15.9 %	23.2 %
Average	8376	-	3.9 %	62.2 %	14.7 %	18.2 %

As shown in Table 1, the drainage basins are dominated by forested areas and bogs. The basins are quite large with an average area of 83.8 km². The amount of open areas is quite small, and within the drainage basins there are virtually no urban areas. The average amount of water-pixels is about 1 %. The data analysis are conducted by combining the pixels of the open areas and open bogs for the open areas estimation, and by combining the forested areas with the forested bogs for the forested area estimation. The accuracy of SCA estimation without the bogs is separately studied.

An interesting question on the analysis of SCA estimation accuracy was the effect of topography. To assess this the amount of topography on the drainage basins had to be measured. The topographical variations of the drainage basins were calculated by two different methods. The first method was to calculate the standard deviation of elevation in each drainage basin. The amount of standard deviation would give information about the amount of topographical variation within the drainage basin. The second method was to calculate the average amount of difference in elevation for each pixel when compared with its four neighbouring pixels. This would result to information about the topographical variations calculated from the neighbourhood of every pixel. The calculated results are shown in Table 2.

Table 2. The amount of topographical variations calculated for the drainage basins.

Drainage basin	Amount of pixels	Standard Deviation [m]	Topographical variation in neighbourhood [m]	Overall amount of topography
1	12712	30.0	1.25	Average
2	15046	47.5	2.63	High
3	5225	17.7	1.68	Low
4	3155	35.3	1.82	High
5	13334	21.3	1.27	Low
6	10203	25.7	1.19	Average
7	6078	28.2	2.22	High
8	5833	13.4	0.80	Low
9	3643	14.9	1.35	Low
10	6544	22.4	2.02	Average
11	7696	19.6	1.49	Low
12	13970	18.9	1.51	Low
13	6898	34.3	2.60	High
14	5790	20.2	1.88	Average

As a further illustration of the topography; the Digital Elevation Model covering the test area is presented in Figure 9. The magnitude of topographical variations acquired by the calculations seem to correspond well with the visual interpretation of the DEM.

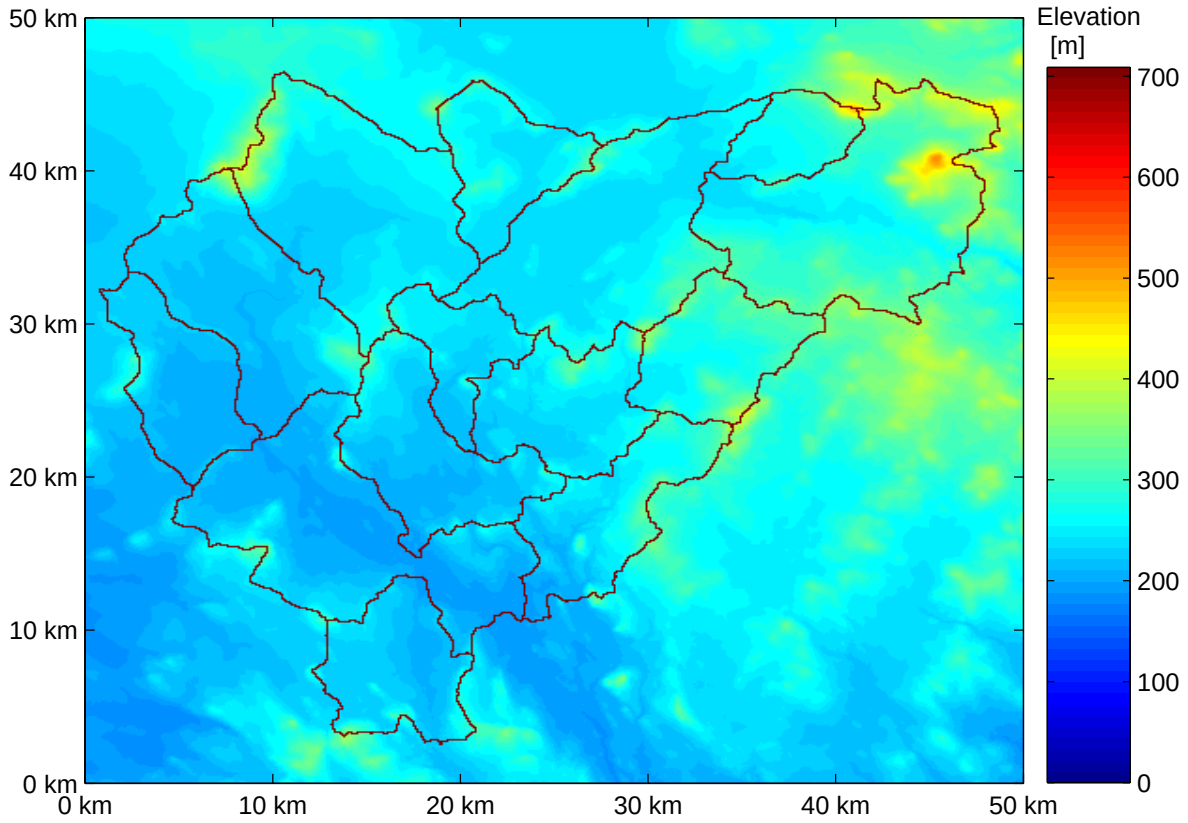

Figure 9. The Digital Elevation Model from the Lokka test area. Relative spatial scale shown in kilometres. The elevation is calculated in meters above the sea level.

3.2 Reference Data from the Test Area

There were two different sources for reference data for the test area. The reference data consisted of Finnish Environment Institute SYKE's WSFS results, and weather station observations during the corresponding time frame.

3.2.1 Reference data from Watershed Simulation and Forecasting System

The operative hydrological model WSFS simulates the hydrological cycle for all the land area of Finland (Vehviläinen 1994). Forecasts for discharge and water level are made for 5500 sub-drainage basins, to which the fourteen basins used in this study also belong. The model functions by using the precipitation and temperature information as input and calculating the water level and runoff for the drainage basins. As a by product of the simulations information on snow covered area is produced. The SCA functions as

a state variable in the simulations, and can also be used as an input for correcting simulations and forecasts of the WSFS model.

The reference data acquired from the WSFS covered the studied snow melt seasons completely. The acquired WSFS SCA data had separate values for all the fourteen drainage basins and separate values for the open areas and the forested areas. Since the WSFS SCA variable is one of the parameters that are adjusted through an optimisation procedure to produce overall watershed estimates, the reference SCA values are in some cases erroneous and this should to be taken into account when considering the results of the accuracy analysis.

3.2.2 Reference data from weather stations

There are more than five hundred weather stations situated around Finland that conduct daily weather observations. Located within the test area of Lokka is a climatological weather station that records the Snow Covered Area observations on a daily basis. The weather station is situated near the lake Lokka within the Drainage basin no. 3, see Figure 8. Since the weather station is located within the test area, and the area is homogenous, the observations can be assumed reasonably accurate for the whole test area. The reference data is temporally extensive, covering all the studied snow melt seasons. The obtained SCA observation is categorized by an universally used snow code. The different classes defined by the snow codes are described in Table 3. The category limits are used in the categorical accuracy analysis in Chapter 4.7.

Table 3. The weather station observation categories defined by the snow codes.

Snow code	**Description**	**SCA in open areas**	**SCA in forested areas**	**Categorical limits**
9	Dry snow 100%	100 %	100 %	0.95 – 1.00
7	Wet snow 100%	100 %	100 %	0.95 – 1.00
6	Wet snow >50%	50-100 %	50-100 %	0.50 – 1.00
5	Wet snow <50%	5-50 %	10-50 %	0.05 – 0.50
4	Wet snow only in forested areas	0-5 %	0-10 %	0.00 – 0.10
0, 1, 2	Snow-free	0 %	0 %	0.00 – 0.05

3.3 Space-borne SAR Data

The Space-borne SAR dataset consists of 24 different ERS-2 SAR PRI intensity images acquired during 1997-2002. The image resolution is 30 meters, and the raw images have a pixel spacing of 12.5 meters. The Digital elevation model used in the image processing has a resolution of 25 meters. The image processing was conducted using either Gamma Ltd. Software (Wegmüller et al. 1998) or an analytical correction method implemented with Imagine software. The Gamma software applies a correlation correction method for the topographical image correction. This results in very high position accuracy and also executes the topographical correction for the images (Wegmüller et al. 1998, Luojus et al. 2002). The accuracy of the analytical correction method is slightly weaker than that of Gamma software, but sufficient for the purposes of this study.

3.3.1 Processing and GEO-Rectification of SAR images and reference data

The SAR image processing was conducted using two different methods. The first method used for the images of 1997 and 1998 is an orbit parameter-based analytical method. The method uses the orbital parameters of the satellite and the DEM for analytically solving the topographical corrections for each image element. The information of DEM is also used for the Geo-rectification. The second method is a correlation correction method. On the first phase of the method a simulated SAR image of the target area is created using the orbital parameters of the ERS-2 satellite. In the second phase the simulated SAR image is correlated with the actual SAR image. Correction parameters for different sections of the SAR image are calculated based on correlation between the actual and the simulated images. The topographical variations on the image area are also considered in the correction phase, and these are included in the correction parameters. The calculated correction parameters are then used to geo-rectify the SAR image. Since the simulated SAR image is generated from the Digital Elevation Model of the target area and the correction parameters are calculated based on this image, the SAR image will be geo-rectified according to the DEM. The resulting projection is also determined by the DEM, for the images in this study it was Transverse Mercator, Finnish zone 3. The images are processed using the 12.5 meter pixel spacing, and after the geo-rectifying and topographical correction, the images are averaged into a 100 meter resolution. The averaging is done because of the reference material; the forest classification and the drainage basin classification are provided in a 100 meter resolution. An example of a processed SAR intensity image is shown in Figure 10.

The drainage basin information was provided by the Finnish Environment Institute. The drainage basin classification has a spatial resolution of 100 meters. The digital forest classification data have been produced by National Land Survey of Finland. The data

are based on the National Forest Inventory data of the Finnish Forest Research Institute. The Forest Inventory data are based on cartographic data, ground truth sampling and Landsat TM imagery, and it includes information on forest characteristics classified according to forest stem volume. In the forest classification the bog areas have been separated from the forested areas.

The overall image processing has several phases. After the SAR processing and the topography correction the images are averaged into a 100 meter resolution. Then each image is divided by the drainage basin classification. Finally the backscattering coefficient for each forest stem volume class is calculated. The result is an average backscattering coefficient for each stem volume class for each drainage basin. SCA estimates are then calculated using these averaged values by the equations (6) – (10).

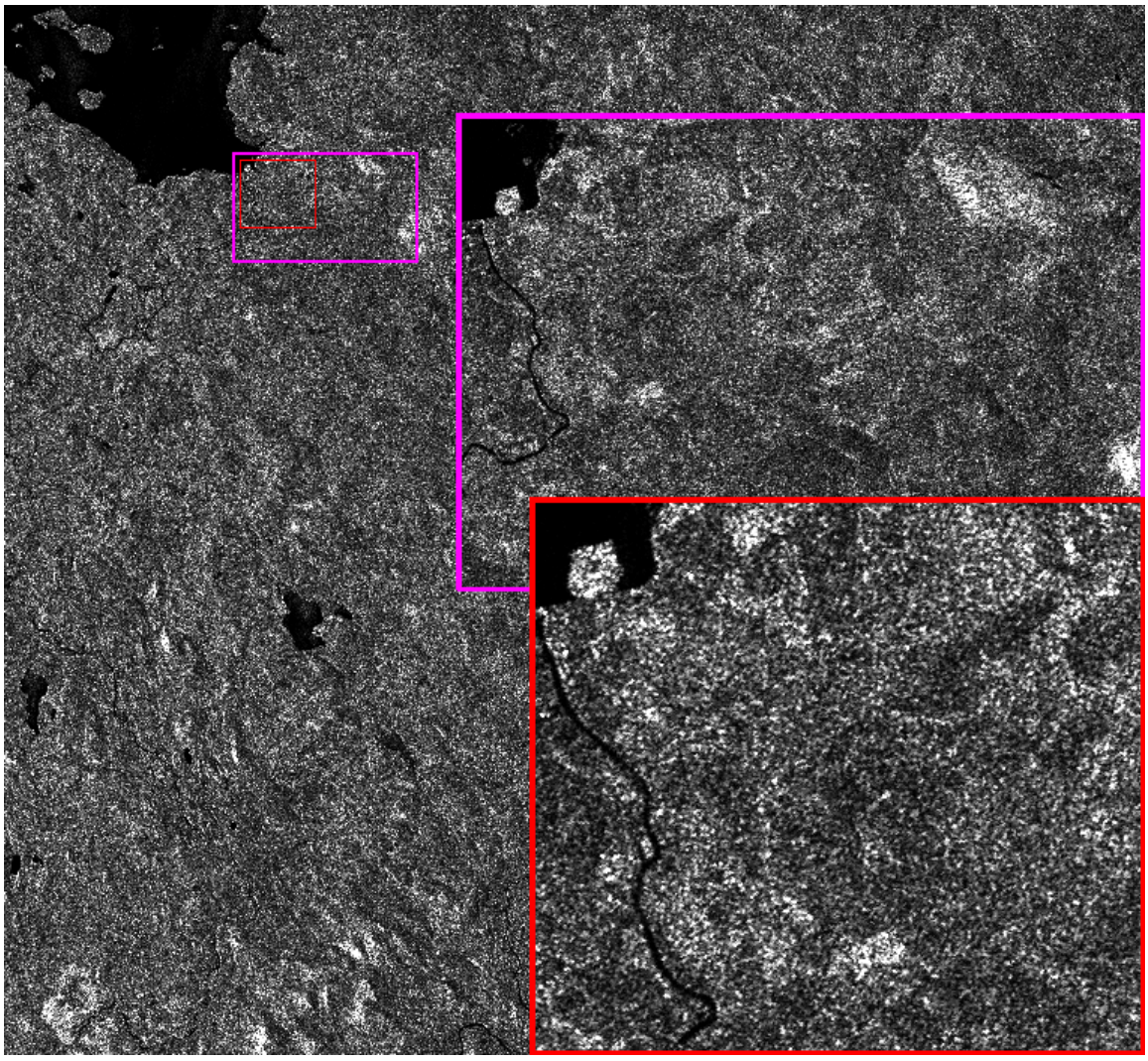

Figure 10. A processed and GEO-rectified ERS-2 SAR intensity image acquired on 2 June 2000. The image resolution for the caption with red frames is 12.5 meters. The largest image has been averaged into a resolution of 100 meters. The Lokka reservoir situates in the upper left corner of each image caption.

The satellite data are from the years 1997, 1998, 2000, 2001 and 2002. All the data and respective information are described in Tables 4 and 5. The images are taken from different orbital positions, so the different images cover slightly different areas. Some of the images only covered part of the test area, thus covering only some of the drainage basins or covering the drainage basins only partially. Since the backscattering coefficients are averaged values, it was decided that at least 10 % of the pixels of the drainage basin needs to be covered by the image, in order to accept the respective drainage basin data into the SCA analysis. Information on the drainage basin coverage is also presented in Table 4.

3.4 The Dataset Used in SCA Estimation

The available ERS-2 SAR images consisted of 24 images from years 1997, 1998, 2000, 2001 and 2002. The images from years 1997 and 1998 were processed with the orbit parameter-based analytical method (Rauste 1989), and the images from the years 2000, 2001 and 2002 were processed with the correlation method (Wegmüller et al. 1998). The images represent a large temporal window. Hence there are variations in the environmental conditions. These changes cause decrease in the SCA estimation accuracy because the reference images and the actual images that are interpreted can be from dissimilar weather and environmental conditions.

The reference image selection was performed according to weather conditions and the information on backscattering coefficients. The reference image for the beginning of the snow melt season was chosen so that the daily temperatures are remarkable above 0ºC, to ensure that the snow pack is wet. Moreover the snow pack should cover all or most of the ground. The weather station snow code for these reference images were either 6 or 7. The reference images for the snow-free ground were chosen close to the time when all the snow had been melted, and the ground was still wet. The weather station snow code range was quite large, ranging from 0 to 5. The Snow Covered Area reference value for open areas was in all cases 0, and for the forested areas it was in the range of 0 – 0.13. Thus according to reference data, the snow had melted completely or almost completely. After the preliminary selection by the weather information, the final selection was carried out by selecting the images that had the most prominent backscattering coefficients, complying with Figure 4. The SCA estimates were also calculated for the reference images, and were used in the accuracy analysis. The information of all the SAR images is collected in Tables 4 and 5.

For this study, the reference image selection was easy, since it could be carried out with all the needed information available, but this is not always the case. As in this study, data from different snow melt seasons can be used as the reference images, but if there are no suitable data available, the reference images need to be acquired from the same snow melt season as the evaluated data. This can sometimes be problematic, since the

satellite images usually need to be ordered some time in advance and the uncertainties of weather forecasts may produce undesirable surprises, i.e. the reference images for the beginning or the end of snow melt season may miss the optimal conditions.

Table 4. The ERS-2 SAR PRI dataset, containing images from 1997-2002. Information on the satellite flight direction, the acquisition time and drainage basin coverage are presented. Also the chosen reference images are shown on the table.

Date	Local Time	Direction	Drainage basins	Data info
YEAR 1997				
9 May	22:40	Ascending node	14	Ref. Image
12 May	22:40	Ascending node	14	Ref. Image
25 May	22:40	Ascending node	6	Data
28 May	22:40	Ascending node	14	Data
4 June	12:35	Descending node	14	Ref. Image
7 June	12:35	Descending node	14	Ref. Image
13 June	22:40	Ascending node	14	Ref. Image
16 June	22:40	Ascending node	14	Ref. Image
YEAR 1998				
1 May	12:35	Descending node	7	Data
7 May	12:35	Descending node	3	Data
13 May	22:40	Ascending node	14	Data
YEAR 2000				
5 May	12:35	Descending node	6	Ref. Image
14 May	22:40	Ascending node	6	Ref. Image
24 May	12:35	Descending node	14	Data
2 June	22:40	Ascending node	14	Ref. Image
YEAR 2001				
23 April	12:35	Descending node	14	Data
9 May	12:35	Descending node	14	Data
18 May	22:40	Ascending node	14	Ref. Image
25 May	12:35	Descending node	6	Data
28 May	12:35	Descending node	14	Data
YEAR 2002				
24 April	12:35	Descending node	14	Ref. Image
3 May	22:40	Ascending node	14	Ref. Image
13 May	12:35	Descending node	14	Data
29 May	12:35	Descending node	11	Data

Table 5. The reference data information of the dataset concerning the whole test area. The snow code is acquired from the Lokka weather station. The reference SCA values are acquired from the WSFS data.

Date	Julian day	Snow Code	SCA open	SCA forest	Average σ° open [dB]	Average σ° forest [dB]
YEAR 1997						
9 May	129	7	0.85	0.94	-11.44	-12.15
12 May	132	6	0.62	0.84	-12.08	-12.73
25 May	145	6	0.18	0.54	-10.23	-10.75
28 May	148	5	0.06	0.36	-9.16	-9.23
4 June	155	5	0	0.13	-7.03	-6.88
7 June	158	4	0	0.07	-5.65	-5.89
13 June	164	0	0	0	-6.08	-6.80
16 June	167	0	0	0	-7.19	-7.53
YEAR 1998						
1 May	121	7	0.78	0.90	-8.41	-9.18
7 May	127	6	0.72	0.86	-7.68	-7.27
13 May	133	6	0.19	0.54	-8.85	-9.48
YEAR 2000						
5 May	126	7	0.67	0.85	-13.14	-12.68
14 May	135	7	0.27	0.58	-10.88	-11.89
24 May	145	5	0	0.09	-7.62	-7.90
2 June	154	0	0	0.01	-5.87	-6.97
YEAR 2001						
23 April	113	7	1	1	-9.83	-10.25
9 May	129	5	0	0.12	-7.57	-7.91
18 May	138	2	0	0.02	-6.18	-7.61
25 May	145	0	0	0.11	-8.87	-9.99
28 May	148	0	0	0.02	-7.76	-8.58
YEAR 2002						
24 April	114	7	0.84	0.90	-11.11	-11.54
3 May	123	5	0	0.12	-6.34	-7.41
13 May	133	1	0	0.01	-7.63	-8.28
29 May	149	0	0	0	-8.07	-8.01

4 Results of SCA Estimation Using Space-borne SAR Data

4.1 The Effect of Reference Image Selection on SCA Estimation

The HUT Snow Covered Area estimation method heavily depends on the reference image selection. From the dataset of 24 images, five images were chosen as the reference for the wet snow situation, and seven images were chosen as the reference for the snow-free situation. The SCA estimation was also conducted for the reference images. Therefore these images were also part of the accuracy analysis. The first step of the analysis was to determine the overall accuracy of the SCA estimation. This was done by calculating the RMS-error (RMSE) between the reference data and the SCA estimates for every reference image combination. The calculations were made for each drainage basin and for the open areas and the forested areas independently. The open bogs and the forested bogs were included in the SCA estimation. The results of RMSE calculations for open areas, forested areas and for the average of open and forested areas are shown in Tables 6, 7 and 8 respectively.

Table 6. The RMSE for all reference image combinations for open areas.

Ref.image	**Snow-free ground**						
Wet snow	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	**18.5.2001**	3.5.2002
9.5.1997	0.223	0.330	0.346	0.267	0.262	0.221	0.240
12.5.1997	0.220	0.325	0.341	0.268	0.254	**0.213**	0.233
5.5.2000	0.243	0.341	0.323	0.249	0.288	0.247	0.260
14.5.2000	0.289	0.387	0.372	0.281	0.341	0.306	0.314
24.4.2002	0.232	0.341	0.357	0.282	0.269	0.224	0.246

Table 7. The RMSE for all reference image combinations for forested areas.

Ref.image	**Snow-free ground**						
Wet snow	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	**18.5.2001**	3.5.2002
9.5.1997	0.184	0.273	0.199	0.174	0.180	0.176	0.173
12.5.1997	0.182	0.263	0.195	0.181	0.179	**0.179**	0.174
5.5.2000	0.205	0.284	0.205	0.205	0.206	0.202	0.201
14.5.2000	0.221	0.298	0.224	0.221	0.227	0.219	0.220
24.4.2002	0.200	0.302	0.219	0.178	0.193	0.173	0.176

Table 8. The average of RMSE for open and forested areas, all reference image combinations.

Ref.image	**Snow-free ground**						
Wet snow	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.1997	**18.5.2001**	3.5.2002
9.5.1997	0.203	0.301	0.272	0.220	0.221	0.199	0.207
12.5.1997	0.201	0.294	0.268	0.225	0.217	**0.196**	0.204
5.5.2000	0.224	0.313	0.264	0.227	0.247	0.224	0.230
14.5.2000	0.255	0.342	0.298	0.251	0.284	0.262	0.267
24.4.2002	0.216	0.322	0.288	0.230	0.231	0.199	0.211

As shown in Table 8 the best results in average were acquired with reference images of 12th May 1997 and 18th May 2001. The amount of samples in RMSE calculations was 208 samples in average, ranging from 108 samples to 268 samples. Similar calculations with all the reference image combinations were conducted for mean absolute error, biasness and correlation coefficient, see Appendix A. These calculations confirmed that the best results were acquired with the above selected reference image pair. The overall effect of reference image selection for the SCA estimation was examined by comparing the results obtained with the best reference image pair with the average results obtained with all the reference image combinations. The average results were acquired by calculating the statistical figures by using all the samples of the dataset. The results are shown in Table 9.

Table 9. The accuracy characteristics for the best reference image pair in comparison with average values for all reference image combinations. Open and forested areas are studied independently.

	All combinations Open areas	**Best pair** Open areas	**All combinations** Forested areas	**Best pair** Forested areas
RMSE	0.278	0.213	0.204	0.179
Mean absolute error	0.209	0.137	0.154	0.119
Bias	+ 0.180	+ 0.095	+ 0.043	-0.049
Correlation coefficient	0.830	0.874	0.860	0.896
No. of samples	7280	268	7280	268

As seen in Table 9 the differences between the average results of all reference image combinations and the results of the best reference image pair are considerable. This is consistent with the predictions made for the overall functioning of the SCA estimation method. The higher performance obtained with best reference image pair is remarkable in all aspects of the statistical accuracy analysis. The RMSE is improved by 23 % in the

case of open areas, and 12 % in the case of forested areas. The improvement in mean absolute error is 34 % for open areas and 23 % for forested areas. The bias is improved by 47 % for open areas, but slightly weakened for forested areas. The linear correlation coefficients are also clearly improved in the case of the best reference image pair. The improved correlation coefficients are visualised by plotting the estimated SCA in respect with the reference SCA values for the cases of best reference image pairs and the cases of all reference image combinations. The visualisation is shown in Figure 11.

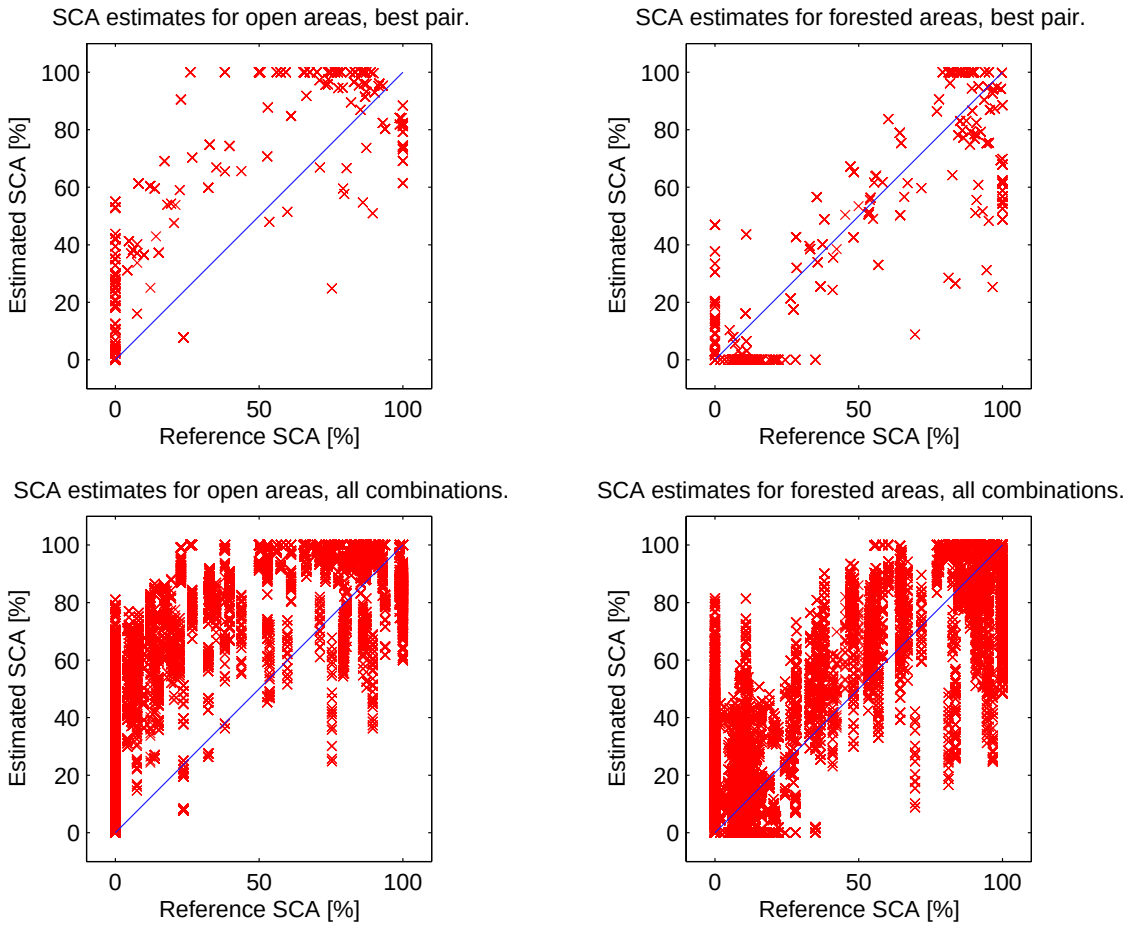

Figure 11. Estimated SCA values plotted in comparison with reference SCA values. The estimated SCA are plotted in respect with the reference SCA values for the cases of best reference image pairs and the cases of all reference image combinations. The amount of samples is 7280 for all reference image combinations and 268 samples for the best reference image pair.

As seen in Figure 11, the correlation plot seems to be somewhat scattered and biased in case of open areas. The correlation is obvious, especially for forested areas. The correlation plot for open areas seems to be worse, and a quite significant overestimation is observed. This overestimation could be caused by the fact that most of the reference images for the beginning of snow melt season had a SCA value significantly less than 100 %, see Table 5. This causes overestimation as can be confirmed by equation (10).

The average results of all combinations on forested areas are clearly better than the results for the open areas. One explanation for the better results of forested area SCA estimates could be the greater amount of pixels for each estimate, see Table 1. One explanation could be the fact that the forested areas are more stable for the variations of soil moisture, and thus the SCA estimation is more accurate. The behaviour of the results are also partly caused by the nature of the reference data. The reference data is obtained from the WSFS model simulations, and the SCA value is one of the parameters that are adjusted through an optimisation procedure to produce the overall watershed estimates. Therefore the reference SCA values are not exact values and this has to be taken into account.

The accuracy analysis is complemented by a distribution histogram of SCA estimation errors, shown in Figures 12 and 13. The histogram presents the amounts of errors classified by magnitude. The magnitudes are evenly distributed, ranging from –1 to +1 with an interval of 0.1, thus forming 20 distinct classes. From the distribution it is easy to see the overall biasness of the estimates and the amount of deviation within the dataset. The error distributions are drawn for all the reference image combinations and for the best reference image pair, separately for open and forested areas.

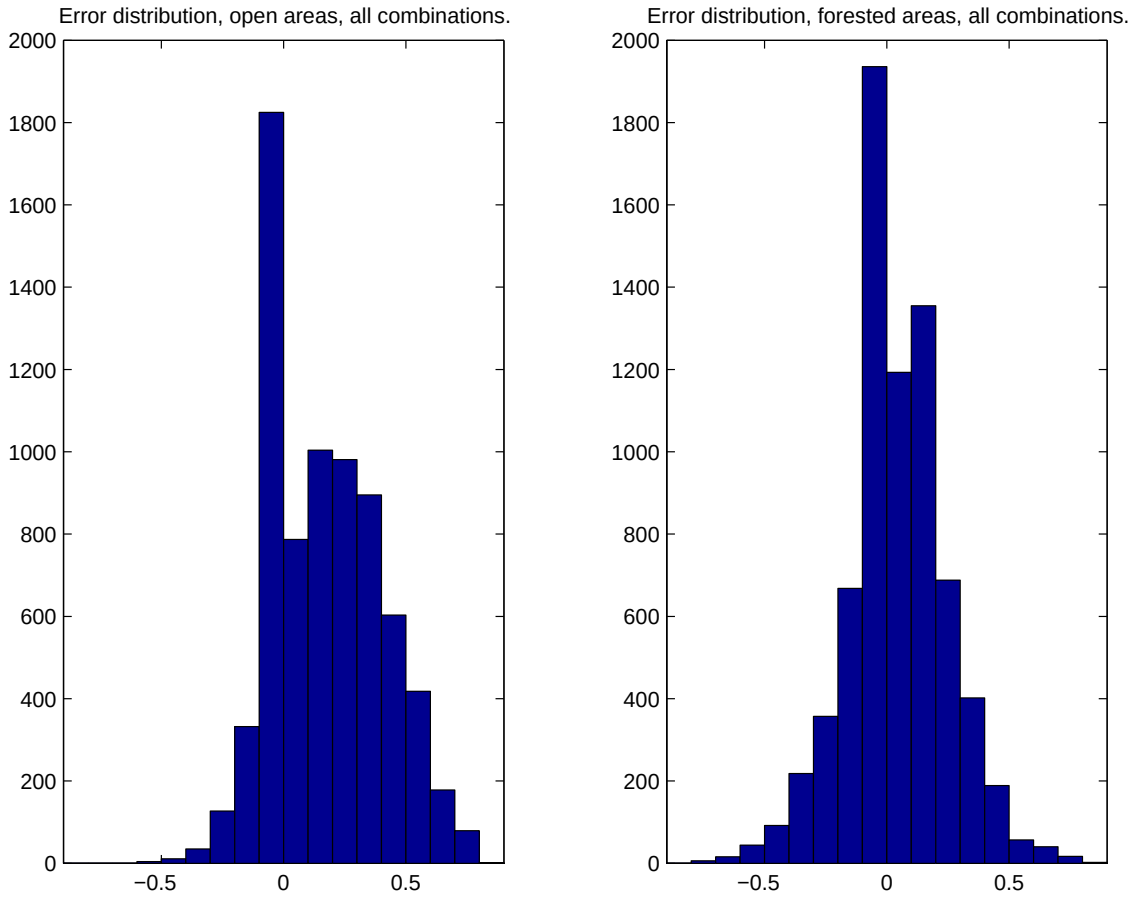

Figure 12. The error distribution of all reference image combinations for open and forested areas.

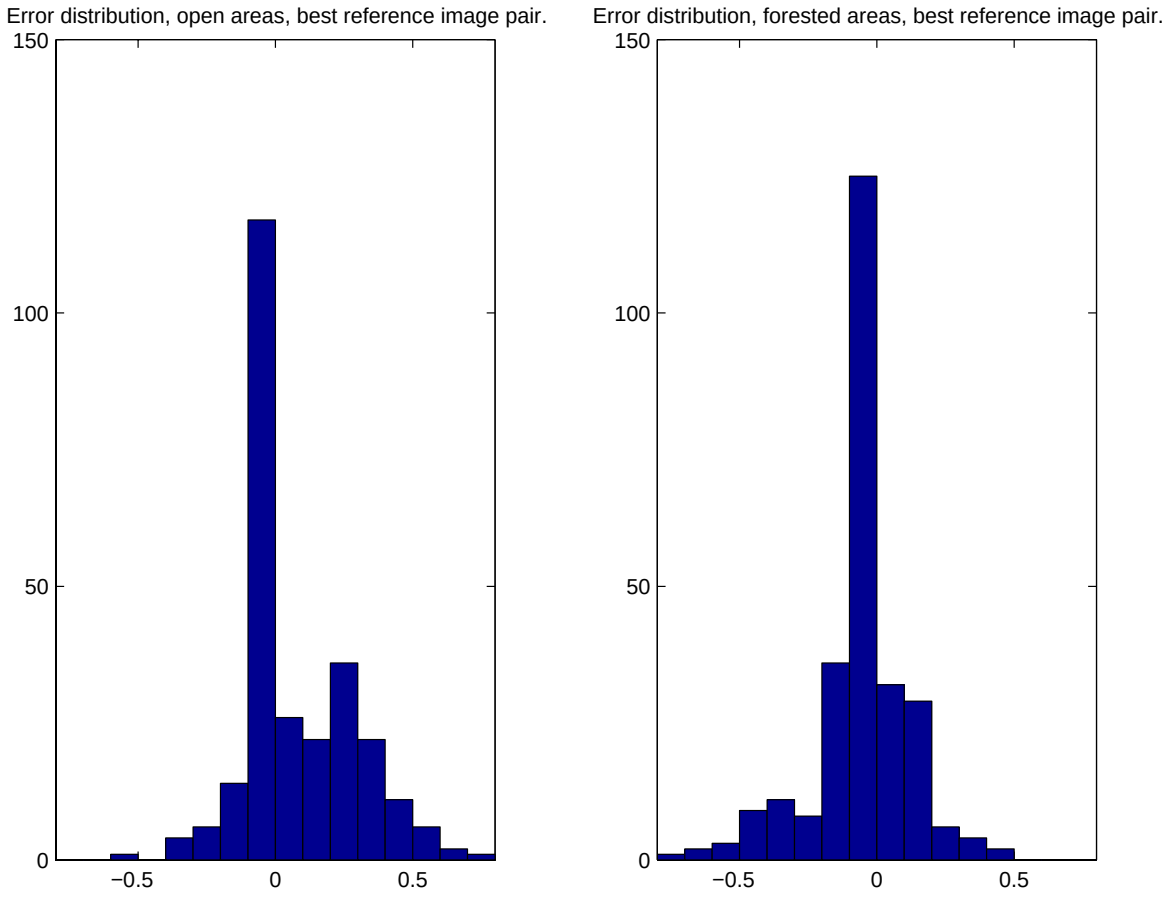

Figure 13. The error distribution of the best reference image for open and forested areas.

As seen from the Figures 12 and 13 the average bias is positive in the case of all reference image combinations. The bias for open areas is 0.180 and for forested areas 0.043. The bias for the best reference image is positive for open areas 0.095 and negative for forested areas -0.049. The standard deviation of the errors are quite remarkable, 0.217 for open areas and all combinations, 0.202 for forested areas and all combinations. The situation is slightly better in the case of the best reference image, the standard deviation for the error is 0.191 for open areas and 0.172 for forested areas.

The examination of SCA estimation methods dependence on reference image selection was confirmed by the results. When the SCA estimation is conducted with the best reference image pair the results are clearly improved. The RMSE for the best reference image pair is improved on average by 0.045 and the difference to worst reference pair case is 0.147. Similar behaviour found for all statistical parameters confirm the sensitivity of the HUT SCA estimation method to reference image selection.

4.2 The Universal Method for Reference Image Selection

An important part of the SCA estimation analysis was to find out if an universally usable method for acquiring the best possible reference image could be found. The goal of the study was to find out the criteria that are consistent for the best reference images in the study. The best reference image in the previous chapter was selected by comparing the RMSE results acquired with the different reference image combinations, and selecting the best candidate. The acquisition of the best pair is easy when the comparison can be conducted using reference data. However in an operational application the reference SCA data are not available, since the estimation is carried out in order to come up with the SCA information.

When the Tables 6, 7 and 8 are studied and average accuracies are calculated for each individual reference image, it can be easily seen that there are considerable differences between the best and worst reference images. In order to assess the statistical accuracy of the reference images an analysis was conducted separately for each of the reference images. This was carried out by calculating the average accuracy of all the SCA estimates with a certain reference image. For one wet snow reference image, all the seven snow-free reference images were gone through and the average accuracies were calculated. Thus resulting to an average accuracy for the certain reference image. The same procedure was conducted for all the reference images, so the accuracy results were obtained for each reference image. These results were then compared with the general information of each image, to find out if there was some clear indication of the conditions that produce the best reference images. The information of reference images and the accuracy analysis for them are collected in Table 10. The average accuracy parameters are also shown in Appendix A.

Table 10. Summary of the reference images. The images are sorted by the quality of RMSE and Mean Absolute Error, the best images listed first. Average precipitation and temperatures for the three days prior to image acquisition are shown for each case. The average backscattering coefficients for the images are also presented.

Image	Node	RMSE / MAE	Reference SCA	Day	General Information All info for 3 prior days of image
Wet Snow conditions, reference images for the beginning of the snow melt season. The images are sorted by average quality, best image first.					
12.5.1997	Asc	open: 0.265 / 0.199 fore: 0.193 / 0.154 ave: 0.229 / 0.176	open: 0.62 fore: 0.84 ave: 0.73	132	3.5 mm of rain in 3 prior days 3 day ave. temp: 5.2 ºC σ_{open}: -12.08 dB; σ_{forest}: -12.73 dB
9.5.1997	Asc	open: 0.270 / 0.301 fore: 0.194 / 0.157 ave: 0.232 / 0.179	open: 0.85 fore: 0.94 ave: 0.895	129	No rain in 3 prior days 3 day average temp: 5.2 ºC σ_{open}: -11.44 dB; σ_{forest}: -12.15 dB
24.4.2002	Desc	open: 0.279 / 0.209 fore: 0.206 / 0.170 ave: 0.242 / 0.189	open: 0.84 fore: 0.90 ave: 0.87	114	2.7 mm of rain in 3 prior days 3 day average temp: 5.5 ºC σ_{open}: -11.11 dB; σ_{forest}: -11.54 dB
5.5.2000	Desc	open: 0.279 / 0.206 fore: 0.216 / 0.172 ave: 0.247 / 0.189	open: 0.67 fore: 0.85 ave: 0.76	126	0.1 mm of rain in 3 prior days 3 day average temp: 4.1 ºC σ_{open}: -13.14 dB; σ_{forest}: -12.68 dB
14.5.2000	Asc	open: 0.327 / 0.240 fore: 0.233 / 0.177 ave: 0.280 / 0.208	open: 0.27 fore: 0.58 ave: 0.425	135	0.2 mm of rain in 3 prior days 3 day average temp: 0.3 ºC σ_{open}: -10.88 dB; σ_{forest}: -11.89 dB
Snow-free ground conditions, reference images for the end of the snow melt season. The images are sorted by quality, best image first.					
18.5.2001	Asc	open: 0.232 / 0.152 fore: 0.184 / 0.126 ave: 0.208 / 0.139	open: 0.00 fore: 0.02 ave: 0.01	138	5.5 mm of rain in 3 prior days 3 day average temp: 5.6 ºC σ_{open}: -6.18 dB; σ_{forest}: -7.61 dB
4.6.1997	Desc	open: 0.234 / 0.156 fore: 0.194 / 0.148 ave: 0.214 / 0.152	open: 0.00 fore: 0.13 ave: 0.065	155	No rain in 3 prior days 3 day average temp: 14.5 ºC σ_{open}: -7.19 dB; σ_{forest}: -7.53 dB
3.5.2002	Asc	open: 0.250 / 0.174 fore: 0.182 / 0.129 ave: 0.216 / 0.152	open: 0.00 fore: 0.12 ave: 0.06	123	1.2 mm of rain in 3 prior days 3 day average temp: 7.1 ºC σ_{open}: -6.34 dB; σ_{forest}: -7.41 dB
16.6.1997	Asc	open: 0.270 / 0.203 fore: 0.186 / 0.131 ave: 0.228 / 0.167	open: 0.00 fore: 0.00 ave: 0.00	167	0.3 mm of rain in 3 prior days 3 day average temp: 8.3 ºC σ_{open}: -7.03 dB; σ_{forest}: -6.88 dB
2.6.2000	Asc	open: 0.273 / 0.204 fore: 0.191 / 0.142 ave: 0.232 / 0.173	open: 0.00 fore: 0.01 ave: 0.005	154	4.8 mm of rain in 3 prior days 3 day average temp: 10.3 ºC σ_{open}: -5.87 dB; σ_{forest}: -6.97 dB
13.6.1997	Asc	open: 0.348 / 0.292 fore: 0.206 / 0.162 ave: 0.277 / 0.227	open: 0.00 fore: 0.00 ave: 0.00	164	0.2 mm of rain in 3 prior days 3 day average temp: 15.3 ºC σ_{open}: -6.08 dB; σ_{forest}: -6.80 dB
7.6.1997	Desc	open: 0.339 / 0.284 fore: 0.282 / 0.241 ave: 0.310 / 0.265	open: 0.00 fore: 0.07 ave: 0.035	158	No rain in 3 prior days 3 day average temp: 12.4 ºC σ_{open}: -5.65 dB; σ_{forest}: -5.89 dB

As seen in Table 10 the different reference images do have quite similar conditions when compared to each other within the corresponding category. The first impression is that the reference images are from quite similar conditions, but still the accuracy differences are quite evident. When the best and the worst reference images are compared there is a remarkable difference in results. The differences are even higher in the case of open areas when compared with the forested areas. So the analysis are conducted separately for the open areas and the forested areas and also separately for the wet snow and snow-free ground conditions.

The analysis for the reference images for the wet snow conditions is assessed first. The overall conditions on wet snow reference images were rather similar, all the temperatures were in the range of 4.1ºC and 5.5ºC, except for the image with the worst results, it had a three days average temperature of 0.3ºC, which seems to be too low. For the SCA estimation it is required that the average temperature is considerable higher than 0ºC to assure that the snow pack is wet during the image acquisition. It is interesting to see the variation in the Julian day of the images. They varied from 114th to 135th Julian day. Although at the 135th day the snow melt season had already started, and the results with this reference image were poor. Usually in Northern Finland the beginning of snow melt season situates in the beginning of May. The amount of precipitation was quite small prior all the reference image acquisitions, and there was no remarkable differences in the cases where there were some rain prior to the image when compared to the case with no rain prior to the image. It is still very important that there has not been any extreme conditions at or prior to the image acquisition time. For example a heavy rainfall could change the properties of the snow pack so that the SCA estimation could possibly be less accurate. The differences for the backscattering coefficients are the most obvious ones. In this dataset the best reference image has an average backscattering coefficient of -12.1 dB in the case of open areas. In the case of forested areas the best results are obtained with an average backscattering coefficient of -12.7 dB, it is however interesting to observe that one of the least accurate results for forested areas were gained with an image with an average backscattering coefficient of -12.7 dB. There are two possible explanations for this observation, one explanation is that the image is from an descending node, and this decreases the accuracy of the estimates. The other explanation is that the image has only six drainage areas that are covered, and these drainage areas are giving the least accurate SCA estimates, maybe because of topography or some other properties of the drainage areas that are covered. These observations are studied more closely on the next chapter. If both the topography and the satellite flight direction have an effect on SCA estimation, it could be that both of these factors are causing the behaviour. The reason for the inaccurate results obtained with the image of 14th May 2000 were caused by clearly other reasons, the image was taken when the snow melt season had already progressed too far.

As a conclusion for the reference image selection for the wet snow conditions, this analysis showed that the best SCA estimation results were achieved with backscattering coefficients of -12.1 dB and -12.7 dB for open areas and forested areas respectively. This information is valid for the dataset under consideration, and cannot be stated as an universal constant for all the cases. But the conditions that produce the best reference images are clearly defined: a wet snow layer that covers mostly all of the ground, and the mean temperature of the acquisition time needs to be well above 0ºC.

The reference images for the end of the snow melt season varied more than the images for the beginning of the season. The most interesting observation is that the best images for the open areas and forested areas were different. The best overall reference image had the best accuracy for the open areas but it had 2nd best results on forested areas. The best open area reference image had the 3rd best results in forested areas and in overall. Also the second best reference image for the open areas had the 5th best results in the forested areas. This behaviour is caused by the fact that can also be seen from the reference data. The open areas reach the snow-free conditions when there is still snow in the forests. So the reference image may be suitable for either the open areas or the forested areas, but not necessarily for both of the cases. The reference image should be from a situation where the snow has just melted from the open and forested areas and the ground is still wet. When the ground starts to dry the backscattering coefficient starts to decrease, and the SCA estimation results are also worsened. For this reason the best reference image is usually a trade-off in which there is still some snow in the forests and the open areas have already started to dry. The temperatures for the different images varied quite notably. There were temperatures higher than 15ºC and the best image had an average temperature of 5.6ºC. The differences in the temperatures did not correlate with the SCA estimation accuracy. The precipitations varied also somewhat between the different reference images. The best images had 5.5 mm of rain prior to image acquisition, the second best image with almost equally good results had no rain prior to the image acquisition. So the precipitation did not seem to correlate with the estimation accuracy either. The image acquisition days varied very much between the different images. One of the images had an acquisition day of 167th and one of the images was acquired on the 123rd Julian day. This variation shows the difficulty of predicting the end of snow melt season. It can vary by over a month in different years. For this reason it is necessary to observe the progress of the yearly snow melt season, and choose the reference image acquisition times according to the prevailing conditions.

The analysis showed that for this dataset the best results were seen with backscattering coefficients of - 6.2 dB for open areas and - 7.4 dB for forested areas. This information is valid for this dataset, and cannot be stated as an universal constant for different datasets. However the conditions for the best reference images are well defined: the image should be taken as close to the end of snow melt season as possible, so that the ground is still wet, and the forest should be as snow-free as possible.

4.3 The Effect of Topography on SCA Estimation

The effect of topography on SCA estimation was analysed by comparing the RMSE of the SCA estimates with the amount of topography on each drainage basin. The amount of topography on the drainage basins was measured by two different methods. The SCA to topography comparison was conducted with respect to both of the methods. The first approach was to calculate the standard deviation of elevation of each drainage basin. The standard deviation was then used as a measure of topographical variation for each drainage basin. The second method was to calculate the elevation difference between all the pixels of the drainage basin against their respective neighbours. The elevation difference was calculated for the four adjacent pixels of each drainage basin pixel (north, south, east and west), and an average of these values was calculated for each drainage basin. This method gives the magnitude of topographical variation calculated for the close neighbourhood of each pixel. The results for the amount of topographical variations are seen on Table 2. The main analysis for the effect of topography was conducted by comparing the RMSE of the SCA estimates with the amount of topographical variations for each drainage basin. This was carried out by calculating the linear correlation coefficient for the RMSE and topographical variation, see Table 11. The RMSE results for each drainage basin were calculated from all the SCA estimates of each basin by using the best reference image pair and by using all the reference image combinations.

Table 11. The linear correlation coefficients for comparison of topography with RMSE. Method 1 shows the comparison of RMSE of SCA estimates to standard deviation of elevation. Method 2 compares the RMSE of SCA estimates to topographical variation from the neighbourhood of each pixel.

	Method 1 Open areas	**Method 1** Forested areas	**Method 2** Open areas	**Method 2** Forested areas
Correlation, best pair	0.618	0.740	0.470	0.393
Correlation, all combinations	0.315	0.612	0.183	0.322

As seen in Table 11 the results obtained with the different methods are somewhat dissimilar. The acquired linear correlation coefficients show a considerable correlation between the standard deviation of elevation and the RMSE of SCA estimates. The correlations for the best reference image pair of 0.618 and 0.740 indicate clearly that the amount of error on the SCA estimates are correlated to the amount of standard deviation of elevation. The second method for analysis of topographical variation shows a remarkably smaller amount of correlation. So the amount of topographical variation calculated from the neighbourhood of each pixel does not correlate well with the

magnitude of RMSE of the SCA estimates. The differences when the best reference image pair is compared to all reference image combinations are remarkable, and it seems that the correlation is considerably weaker when calculated for all the reference image combinations. This could be caused by the fact that the variation of RMSE between the different drainage areas is reduced in the case of all reference image combinations, since also the poorer SCA estimates are included in the dataset. It can be concluded that the standard deviation of elevation correlates moderately well with the magnitude of RMSE of the SCA estimates. The correlations for the best reference image pair are visualised by showing the correlation plot for the case of standard deviation of elevation and the RMSE of the SCA estimates in Figure 14.

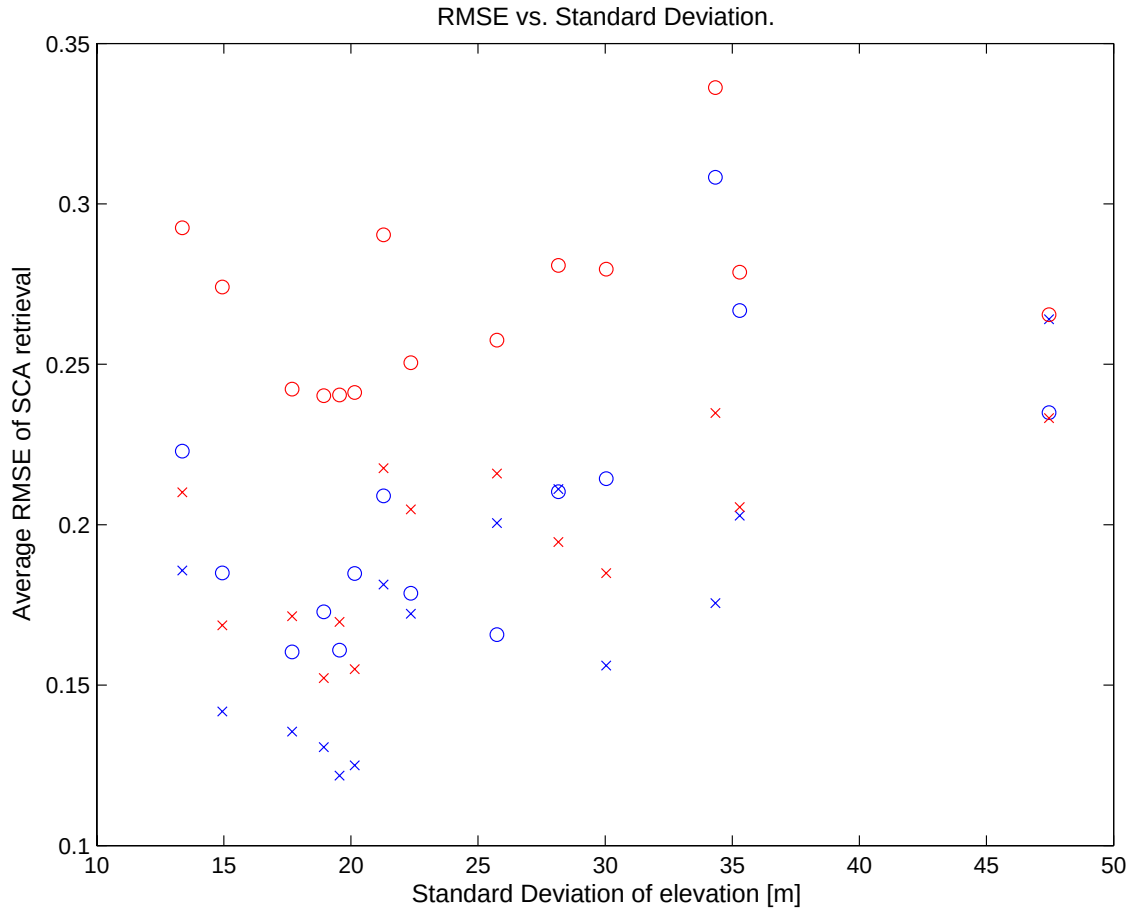

Figure 14. The RMSE of the SCA estimates plotted in comparison with the standard deviation of elevation in each drainage basin. The red colour presents all the reference image combinations, blue colour presents the best reference image pair. The circles present the open areas and the crosses present the forested areas.

The visualisation for the correlation of the amount of topographical variation calculated from the neighbourhood of each pixel with the amount of RMSE of the SCA estimates is shown in Figure 15.

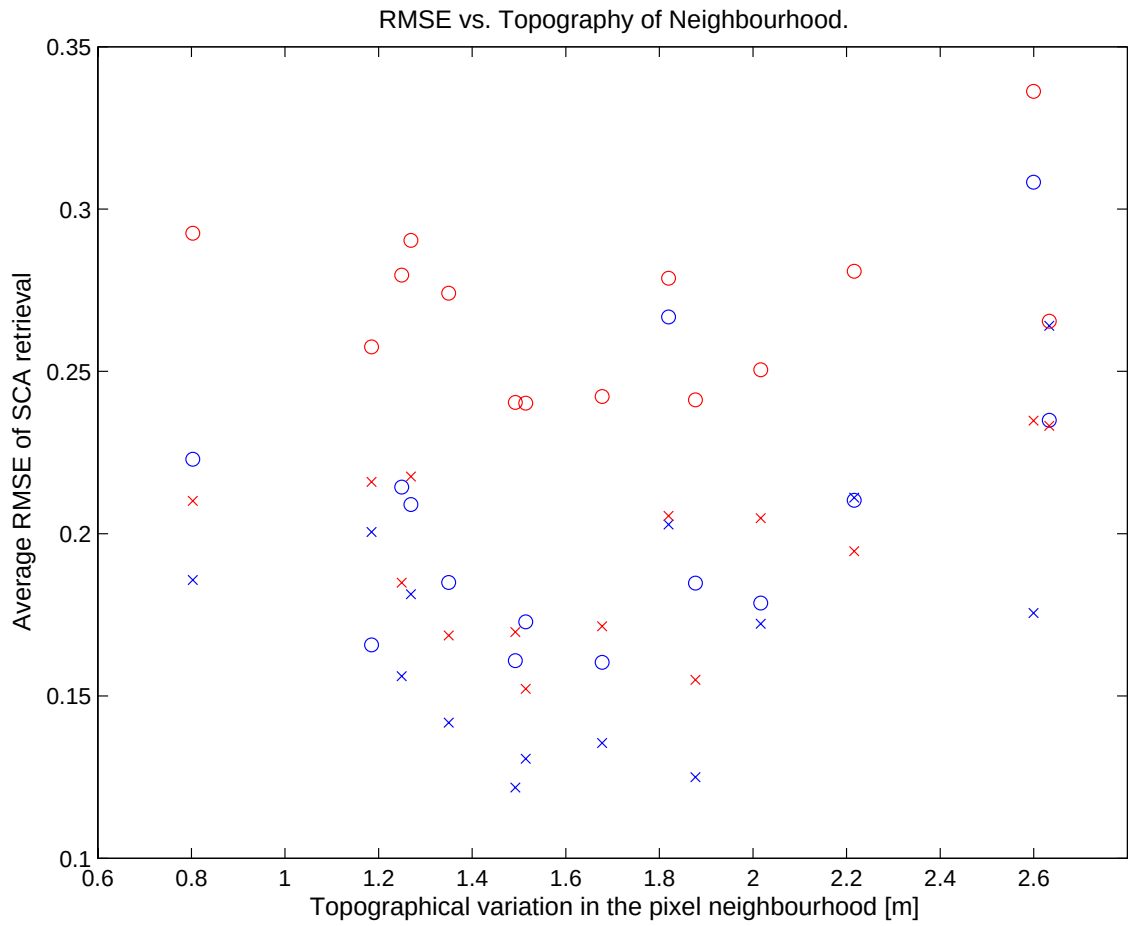

Figure 15. The RMSE of the SCA estimates plotted in comparison with the topographical variation of the neighbourhood in each drainage basin. The red colour presents all the reference image combinations, blue colour presents the best reference image pair. The circles present the open areas and the crosses present the forested areas.

As seen in Figure 15 the correlation between the RMSE and the neighbourhood topography is clearly weaker than that between the RMSE and the standard deviation of elevation.

The conclusion from these observations is that the topography of the estimated area has some effect on the SCA estimation accuracy. It can be seen from Figures 14 and 15 that there are also high RMSE results obtained for some areas with small topographical variations, although it seem to be that drainage basins with the highest topographical variations do have the highest level of RMSE. Although the small amount of drainage basins does not provide an extensive dataset, hence the statistical reliability of this study could be questioned. The correlation with the standard deviation of topography and the RMSE of the SCA estimates is considerable, and this information can be possibly used for the rough determination of SCA estimation accuracy.

In previous chapter, it was observed that one of the reference images, 5 May 2000, that covered only six drainage basins provided a clearly weaker estimation performance than the other images, even though it showed the same levels of average backscattering coefficient as the best reference image. One assumption for this behaviour is that the six drainage areas covered by the image would produce SCA estimates worse than average. The average RMSE for the six drainage basins (basins no: 2,4,5,7,10,13) of the image, was 0.218 and the overall average RMSE for all the drainage basins was 0.189, as calculated with the best reference image pair. This observation indicates, that these six drainage basins have some characteristics that yield performance worse than that obtained on average. Actually, their topographical variation is among the eight highest values. Additionally, in this case the reference image was from a different satellite flight direction than in most of the reference image pairs. Thus, the topographical variation combined with the different satellite flight direction seem to affect the SCA estimation accuracy. As a conclusion, it is justified to assume that the amount of topography affects the SCA estimation accuracy. The effect of satellite flight direction on the SCA estimation is studied in more detail in the following chapter.

4.4 The Effect of Satellite Flight Direction on SCA Estimation

The satellite images are taken from each applicable satellite overpass. The direction of the overpass is determined by the orbital parameters of the satellite. Some of the satellite images are taken from an ascending node and some images are taken from a descending node of the satellite orbit. As shown in Figure 6, the acquired satellite images from different nodes are slightly dissimilar. Thus, when the SCA estimation is conducted using images from different nodes it is supposed that the differences in the images are a source of error for the SCA estimation. The effect of satellite flight direction on SCA estimation is studied here. The analysis was made by comparing the SCA estimation accuracies determined in Chapter 4.1 with SCA estimation accuracies calculated using SCA estimation conducted with satellite images only from a certain, either ascending or descending node. The first step was the generation of suitable datasets containing only images from the same satellite flight direction. As seen from the Table 4 there are 11 images representing an ascending node and 13 images from a descending node. The SCA estimates in each dataset are calculated also for the reference images.

The first step was to calculate the SCA estimates for both datasets for all reference image combinations. The statistical analysis was again conducted in comparison with the WSFS reference data, and the key statistical parameters were calculated for the new cases. The results of the statistical analysis were then compared with the analysis for the complete dataset without the flight path separation discussed in Chapter 4.1.

4.4.1 The dataset for the ascending node

The dataset of ascending node consists of three reference images for wet snow conditions and five reference images for snow-free ground conditions The total amount of satellite images for the dataset is 11. The results of statistical analysis for the ascending node are shown in Table 12.

Table 12. The accuracy characteristics for the dataset of ascending node. The statistical analysis is shown for the best reference image pair in comparison with average values for all reference image combinations.

	Ascending Node			
	All combinations Open areas	**Best pair** Open areas	**All combinations** Forested areas	**Best pair** Forested areas
RMSE	0.313	0.247	0.132	0.100
Mean absolute error	0.239	0.167	0.093	0.075
Bias	+ 0.236	+ 0.160	+ 0.064	+ 0.011
Correlation coefficient	0.852	0.901	0.958	0.974
No. of samples	1480	123	1480	123

The Table 12 present the results for all the reference image combinations, and the results for the best reference image pair. The best reference image pair, 15 May 1997 and 18 May 2001, was chosen by the same method as in Chapter 4.1; by finding the pair that produced the best overall RMSE values and using this reference image pair for all the analysed variables, the extensive results for all the statistical parameters are shown in Appendix B. As seen in Table 12 the results for forested areas are very good, and clearly improved when compared with the results acquired without the satellite flight path separation. The results for the open areas are, however, weaker than the results without the flight path separation, see Table 9. This might be caused by the high amount of bias. The only variable that seems to improve for both the open and forested areas is the correlation coefficient.

The amount of samples for the dataset was considerably smaller than for the dataset without the flight path separation. The amount of samples was 1480 for all reference image combinations and 123 for the best reference image pair. The correlation coefficients for both open and forested areas are clearly improved when compared with results acquired without the satellite flight path separation. The improved correlation

coefficients are visualised by the correlation plot, Figure 16, where the high correlation obtained for forested areas is also demonstrated.

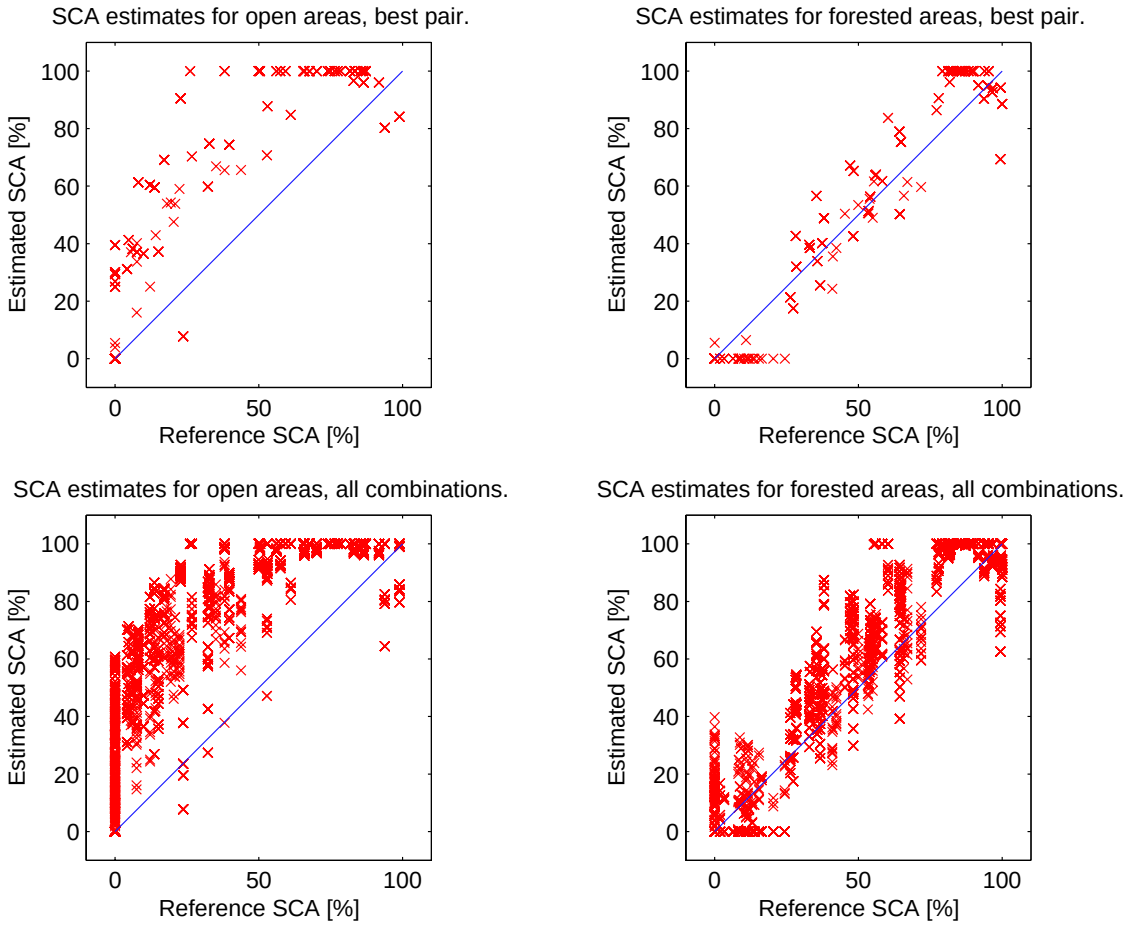

Figure 16. Estimated SCA values plotted in comparison with reference SCA values for the ascending node dataset. The amount of samples is 1480 for all reference image combinations and 123 samples for the best reference image pair.

The correlation plot of Figure 16, shows excellent estimation results for the forested areas. The linear correlation is very obvious and the error between the estimates and the reference values are very small. The plot for the open areas is more scattered, and the results show a significant overestimation in this case.

4.4.2 The dataset for the descending node

The dataset for the descending node contains two reference images for wet snow conditions and two reference images for snow-free ground conditions. The total amount of satellite images for the dataset is 13. The results of statistical analysis for the descending node are presented in Table 13.

Table 13. The accuracy characteristics for the dataset of descending node. The statistical analysis is shown for the best reference image pair in comparison with average values for all reference image combinations.

	Descending Node			
	All combinations Open areas	**Best pair** Open areas	**All combinations** Forested areas	**Best pair** Forested areas
RMSE	0.249	0.177	0.261	0.212
Mean absolute error	0.191	0.115	0.209	0.166
Bias	+ 0.146	+ 0.064	+ 0.098	+ 0.034
Correlation coefficient	0.857	0.914	0.799	0.858
No. of samples	430	145	430	145

The analysis was conducted in a similar manner to the ascending node. The best reference image pair for this dataset was 24 April 2002 and 4 June 1997. As seen in the Table 13 the amount of samples is much smaller than in the previous cases, only 430 samples in all reference image combinations and 145 samples for the best reference image pair, the extensive results for all the statistical parameters are shown in Appendix B. When the results for the dataset of descending node are compared with the results of the complete dataset a few observations are clear. The results for open areas are clearly improved, and the results for the forested areas are clearly weakened in this dataset when compared with the complete dataset. When comparing the open areas with the forested areas in this dataset, the open areas seem to have clearly better SCA estimation accuracy. In all the previous datasets; the complete dataset and the dataset for ascending node the forested areas have shown clearly better accuracy results than those for the open areas.

Thus, there seems to be some quite remarkable differences in the behaviour of the dataset of ascending and descending nodes. The differences are probably caused in some degree by the fact that the dataset of ascending node seems to contain the more suitable reference images for the SCA estimation method. When the reference image analysis was conducted in Chapter 4.2 the best images for the beginning and the end of the snow melt season were from the ascending node. The descending node reference images produced notably less accurate estimates than the best ascending node images.

The correlation plot for the descending node dataset compared with the complete dataset is drawn in Figure 17.

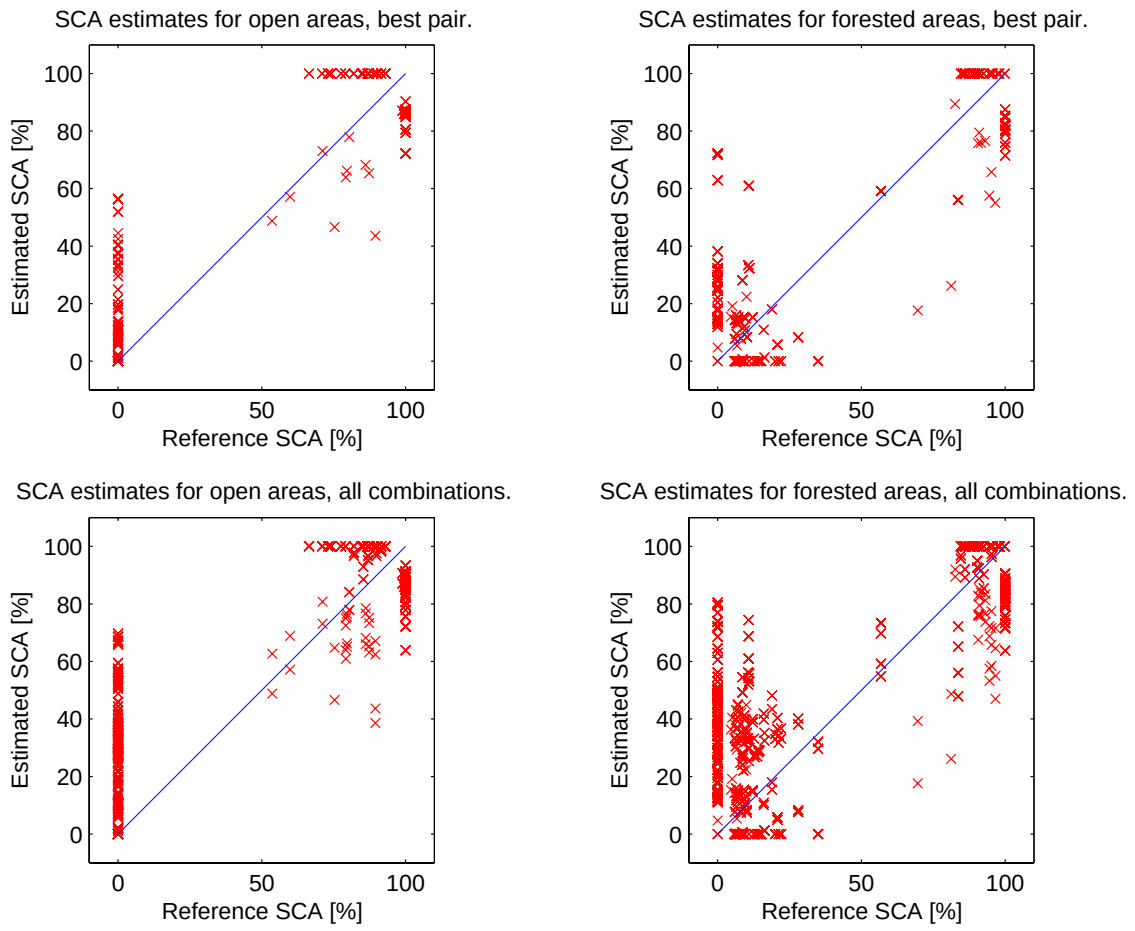

Figure 17. Estimated SCA values plotted in comparison with reference SCA values for the descending node dataset. The amount of samples is 430 for all reference image combinations and 145 samples for the best reference image pair.

The correlation plot of Figure 17, shows the estimation results for the descending node. The linear correlation is not very high and the error between the estimates and the reference values are quite remarkable. Both the open and forested area plots are considerably scattered. One interesting observation can be made from the very large scattering of estimates at the reference SCA of 0 %. This could be explained by the fact that when all the snow has been melted and the ground starts to dry the backscattering coefficient starts to decrease again. The result is that the SCA estimate shows that there is snow on the ground, when the ground has actually been dried since the snow melt season has already ended some time ago. One interesting observation can be made for the open areas near the SCA 100%. The WSFS reference data shows no values of 100%, which according to other reference data is not the case. Thus, at least in the open areas there are some considerable inaccuracies in the WSFS reference data.

4.4.3 The combined dataset of the flight path separated estimates

The combined dataset for the flight path separated estimates was acquired by estimating all the ascending node images with ascending node reference images, and by using the descending node reference images for the descending node estimation, and combining these for the statistical analysis. Thus, the average of 1480 ascending node and 430 descending node samples ware calculated for each statistical parameter, so the combined dataset contains all the information of both of the ascending and descending node datasets, the results are presented in Table 14.

Table 14. The accuracy characteristics comparison between the average values of satellite flight path separated dataset and the average values of the complete dataset.

	Complete dataset Open areas	**Flight paths separated** Open areas	**Complete dataset** Forested areas	**Flight paths separated** Forested areas
RMSE	0.278	0.298	0.204	0.161
Mean absolute error	0.209	0.228	0.154	0.119
Bias	+ 0.180	+ 0.216	+ 0.043	+ 0.072
Correlation coefficient	0.830	0.837	0.860	0.915
No. of samples	7280	1910	7280	1910

As seen in Table 14 the accuracy for open areas is slightly weakened in the flight path separated case. The RMSE, mean absolute error and bias show slightly poorer values, but the correlation coefficient is slightly improved. So the estimation for open areas does not seem to benefit from the flight path separation. However the results for forested areas show a major improvement in the case of flight path separation. The RMSE and mean absolute error are significantly improved, and the linear correlation coefficient also shows a remarkable improvement. The bias is slightly increased, but not in a remarkable way. As a summary, the improvement in forested areas seems quite promising.

The flight path separation appears to improve the accuracy of SCA estimates for the forested areas significantly, but in the case of open areas there is a slight decrease in SCA estimation accuracy. This is a difficult phenomenon to interpret, the difference in the open areas is quite small, the weakening in the estimation accuracy might be caused by the remarkably smaller amount of samples. However the significant improvement in the forested areas is even more difficult to explain. One hypothesis is that the forested areas appear differently from the different flight paths, and when the estimation is done

by using images from the same flight path the SCA estimation becomes more accurate. This could however be studied more thoroughly in the future.

The correlation plot of the combined ascending and descending node dataset compared with the complete dataset is shown in Figure 18.

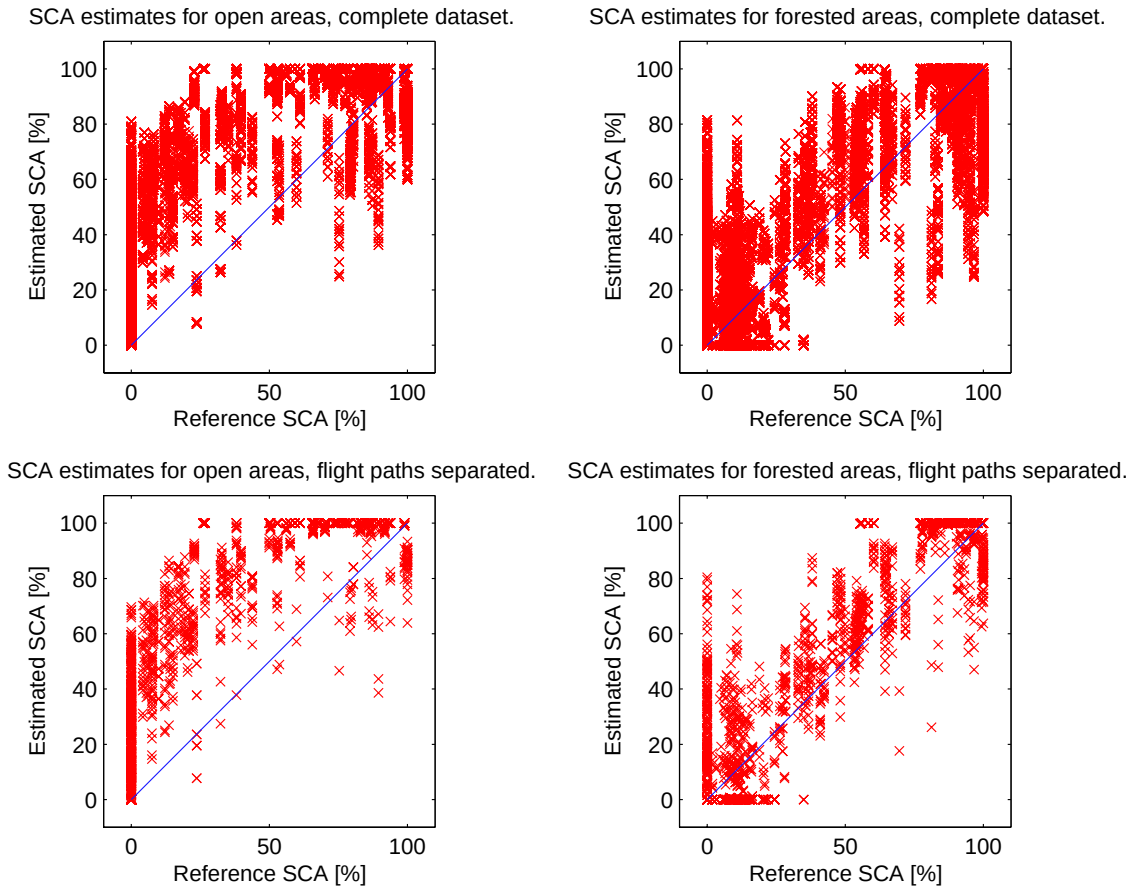

Figure 18. Comparison of the complete dataset with the combined ascending and descending node dataset. The amount of samples are 7280 for the complete and 1910 for the combined ascending and descending node dataset.

The correlation plot of Figure 18 shows that the SCA estimation on the open areas is quite similar in both the complete and flight path separated datasets. There seems to be some clear indications of overestimation in both cases, and both of the plots are quite scattered. The situation in the case of forested areas is different. The dataset for the flight path separated case is clearly better, and the average error is clearly smaller. The overall improvement in accuracy which was seen in Table 14, is quite obvious also in the scatter plot.

As a conclusion, the flight path separation slightly decreases the SCA estimation accuracy on the open areas. In the case of forested areas, the flight path separation improves the SCA estimation accuracy significantly.

4.5 The Effect of Forest Compensation on SCA Estimation

The forest compensation is an important feature of the HUT SCA estimation method. The purpose of it is to minimise the effect of forest canopy backscattering in order to increase the SCA estimation accuracy. The effect of forest compensation was analysed by calculating the SCA estimates without the forest compensation algorithm, and comparing those with the forest compensated estimates. So far, all the analyses have been conducted using the forest compensated estimates. The non forest compensated estimates were acquired by averaging the backscattering coefficients from the different forest stem volume classes by the amount of pixels. The SCA estimation was then conducted with these average values. The results of non forest compensated estimation were compared with the results using the forest compensation algorithm presented in Chapter 4.1. The results of the statistical analysis are shown in Table 15.

Table 15. The accuracy characteristics for forest compensated estimates when compared with estimates acquired without the forest compensation.

	All combinations No forest compensation	**All combinations** Using forest compensation	**Best pair** No forest compensation	**Best pair** Using forest compensation
RMSE	0.215	0.204	0.195	0.179
Mean absolute error	0.159	0.154	0.132	0.119
Bias	+ 0.039	+ 0.043	- 0.038	- 0.049
Correlation coefficient	0.841	0.860	0.871	0.896

Table 15 clearly shows that the forest compensation improves the estimation results. The improvement is seen in all statistical variables except the bias. The improvement with the forest compensation algorithm is not as large as was expected, but the improvement is all together considerable. The extensive results for all the statistical parameters are shown in Appendix C. The RMSE and the correlation coefficient show the most remarkable improvement, which however are positive signs. To analyse the behaviour of forest compensation the RMSEs for all the reference image pairs were calculated as in Chapter 4.1 and the behaviour of the compensation with different reference image pairs was analysed. The analysis is done by calculating the differences between the RMSE of each reference image pair, and these are shown in Table 16. The difference is calculated by subtracting the forest compensated RMSE value from the non-compensated value. Thus, the positive numbers show an improvement in forest compensation.

Table 16. The comparison of RMSE results with and without the forest compensation. The difference between the forest compensated and non compensated RMSEs for all the reference image combinations. A positive value shows that the forest compensated estimate is better, a negative value means that the non-compensated estimate is better.

Ref.image	**Snow-free ground**						
Wet snow	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.1997	**18.5.2001**	3.5.2002
9.5.1997	0.024	0.012	0.008	0.031	0.025	0.027	0.029
12.5.1997	0.022	0.009	0.007	0.031	0.022	0.024	0.026
5.5.2000	0.016	0.011	0.011	0.028	0.015	0.020	0.021
14.5.2000	0.000	-0.010	-0.005	0.031	-0.010	0.019	0.011
24.4.2002	0.024	0.015	0.005	0.026	0.026	0.032	0.029

As the Table 16 shows the forest compensation improves the SCA estimation in almost all cases. The average improvement in RMSE is 9.2%. The only cases where the non-compensated RMSE values are better, are from the reference image that produced the least accurate estimates, and was overall not very well suited for the SCA estimation to begin with. A scatter plot of forest compensated and non-compensated results is shown in Figure 19.

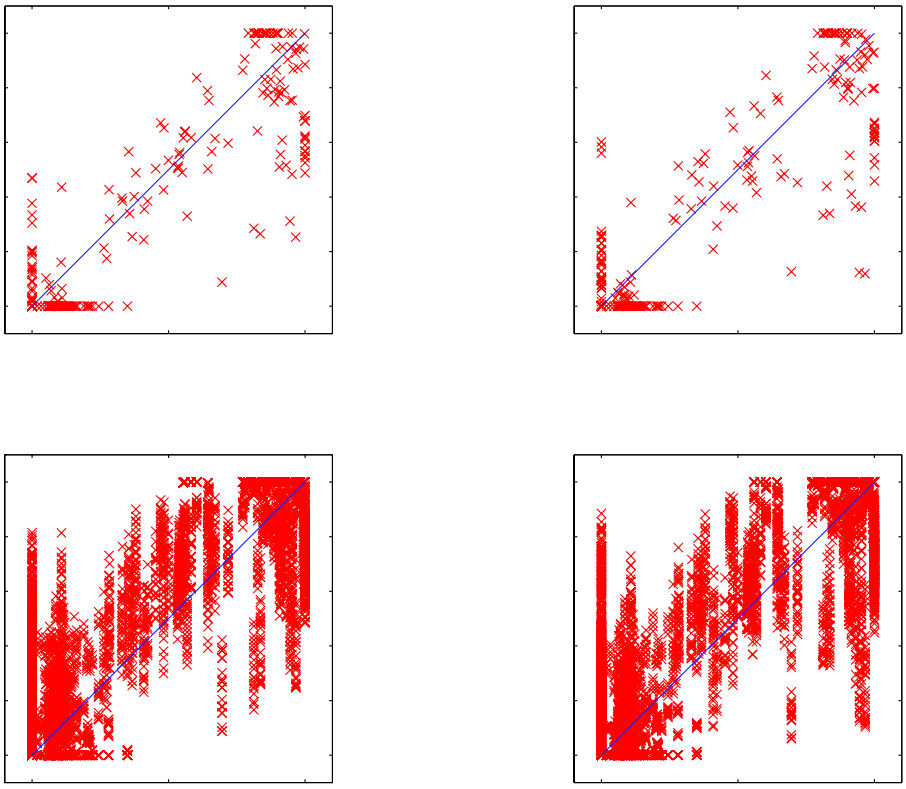

Figure 19. Forest compensated SCA estimates plotted in comparison with SCA estimates calculated without the forest compensation.

4.6 The Effect of Bogs on SCA Estimation

The land-usage classification for the test area consisted of two different classes for open and forested areas; open areas in mineral soil sites and open bogs, forested areas and forested bogs. The SCA estimation has been conducted using the combination of bog and normal pixels. One interesting question in the SCA estimation was the effect of the bogs on the estimation accuracy: would it increase or decrease if the estimation was conducted without using the bog pixels. In order to study this effect a new dataset without the bogs was constructed and the SCA estimation was conducted for this new dataset. The results for the new dataset were then compared with the results of the original dataset, acquired in Chapter 4.1. The most significant difference between the two datasets, the complete dataset and the dataset without the bogs was the amount of pixels on each drainage basin. As seen from Table 1, the dataset without the bogs has less pixels than the dataset with the bogs included. The results of SCA estimation with the different datasets are presented in Tables 17 and 18.

Table 17. The accuracy characteristics for complete open area dataset and the open area dataset without the bog pixels.

	Open Areas			
	All combinations Complete dataset	**All combinations** No bogs	**Best pair** Complete dataset	**Best pair** No bogs
RMSE	0.278	0.311	0.213	0.244
Mean absolute error	0.209	0.225	0.137	0.147
Bias	+ 0.180	+ 0.165	+ 0.095	+ 0.059
Correlation coefficient	0.830	0.744	0.874	0.799

Table 18. The accuracy characteristics for the complete forested area dataset and the forested area dataset without the bog pixels.

	Forested Areas			
	All combinations Complete dataset	**All combinations** No bogs	**Best pair** Complete dataset	**Best pair** No bogs
RMSE	0.204	0.223	0.179	0.206
Mean absolute error	0.154	0.163	0.119	0.131
Bias	+ 0.043	+ 0.031	- 0.049	- 0.047
Correlation coefficient	0.860	0.830	0.896	0.864

As seen in Tables 17 and 18 the results are obvious. The results acquired with the complete dataset are in all cases considerably more accurate than the results without the bogs. The differences are remarkable. The extensive results for all the statistical parameters are shown in Appendix D. The weaker estimation results of the dataset without the bogs might be caused, in some degree, by the smaller amount of samples in the dataset. But the difference in the amount of samples can not completely explain the large difference in the accuracies. Therefore it can be concluded that the SCA estimation should be conducted by including the bog pixels in the datasets. Hence, all other SCA estimations than those discussed in this section are performed using this approach; by including the bog pixels.

4.7 Categorical Analysis of the SCA Estimation Accuracy

The SCA estimation accuracy has so far been analysed by calculating the statistical error parameters. Another interesting way of determining the accuracy is to study the categorical accuracy of the SCA estimation. This can be studied using the weather station snow codes presented in section 3.2.2 and in Table 3. This is done by comparing the acquired SCA estimates with the weather station observations. If the SCA estimate is within the limits of a certain snow code, then a hit is recorded, if the SCA estimate is outside the limits a miss is recorded. Finally the ratio of hits and misses are calculated and a value describing the hit percentage is obtained. The obtained values are considered for the average of all the reference images and for the case of the best reference image. A similar comparison can also be made with the WSFS reference data. A limiting value demonstrating the error of the WSFS data is chosen. If the estimate is within the boundaries of WSFS-reference value plus or minus the limiting value then a hit occurs. Limiting values of 0.10, 0.15 and 0.20 were chosen for this study. The results are presented in Table 19.

Table 19. The results of categorical analysis. The SCA estimates compared with the weather Station Snow codes and the WSFS data with different limiting values.

	All combinations Open areas	**All combinations** Forested areas	**Best pair** Open areas	**Best pair** Forested areas
Weather station Snow Code	58.9 %	51.5 %	61.9 %	52.2 %
WSFS -/+ 0.10 limit	35.9 %	43.1 %	53.4 %	59.0 %
WSFS -/+ 0.15 limit	46.4 %	59.3 %	60.4 %	75.4 %
WSFS -/+ 0.20 limit	54.2 %	71.0 %	66.8 %	83.2 %

The best reference image pair was the same one that has been chosen in the Chapter 4.1. The average of all combinations were calculated by using all the reference image combinations that have been used before; five images for the beginning and seven images for the end of the snow melt season. As seen in Table 19, the agreement between the weather station snow codes and the SCA estimates are mediocre. The results with the best reference image pair are quite reasonable, but still not very good. The results with the WSFS data were also mediocre. The results obtained with the larger limiting values were better, but they probably overestimate the amount of error in the WSFS data.

4.8 A Visual Comparison of SCA Estimates with the WSFS Data

The SCA estimates are calculated with the HUT method for the corresponding days of the satellite images. However the SCA estimates from the WSFS data are available for every day during the snow melt season. Thus, the different data can be presented in a figure for the whole melting season. This is done for all the 14 drainage basins for each year of the dataset using the best reference image pair. The comparison for the reference data and the SCA estimates for all the estimated years are presented in Figures 20-24.

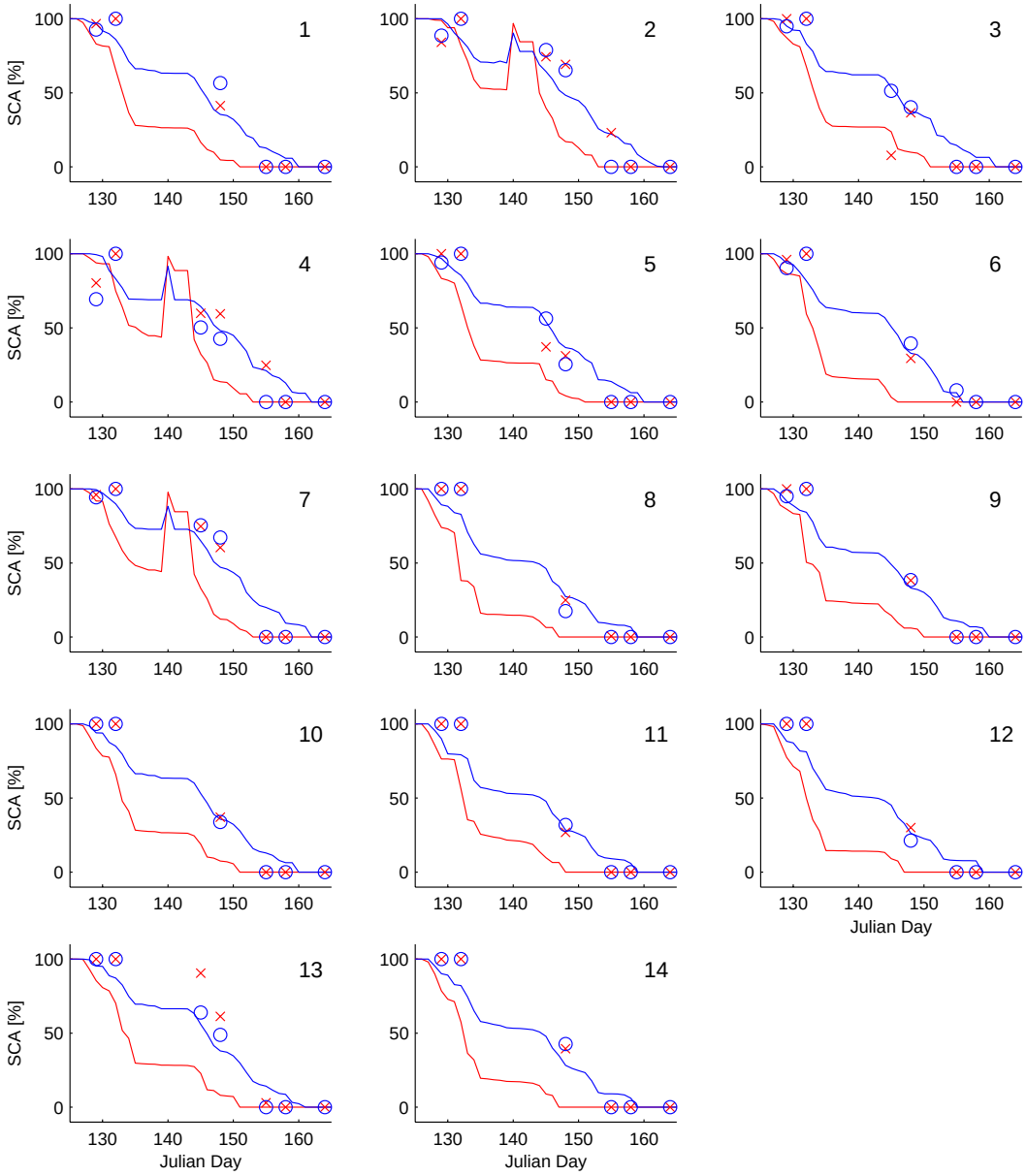

Figure 20. Estimated SCA values for 1997 compared with the reference SCA values. Solid lines present the WSFS reference data. Blue represents forested areas, red open areas. The SCA estimates are drawn by circles and crosses with respective colours. The number of each drainage basin is presented in the figure.

Figure 20 presents the snow melt season of 1997. The SCA estimates and the reference data seem to agree well. The overestimation for the open areas is apparent, and the high accuracy of the forested areas is well observed. In drainage areas 2, 4 and 7 there is increase in the SCA indicating snowfall, between Julian days 140 and 145, and the SCA estimates for these drainage areas are notably higher near this time period. The overall SCA estimation seems to succeed remarkably well for the year 1997. Also the high amount of images for this year present the capabilities of the SCA estimation method for the snow melt season monitoring well.

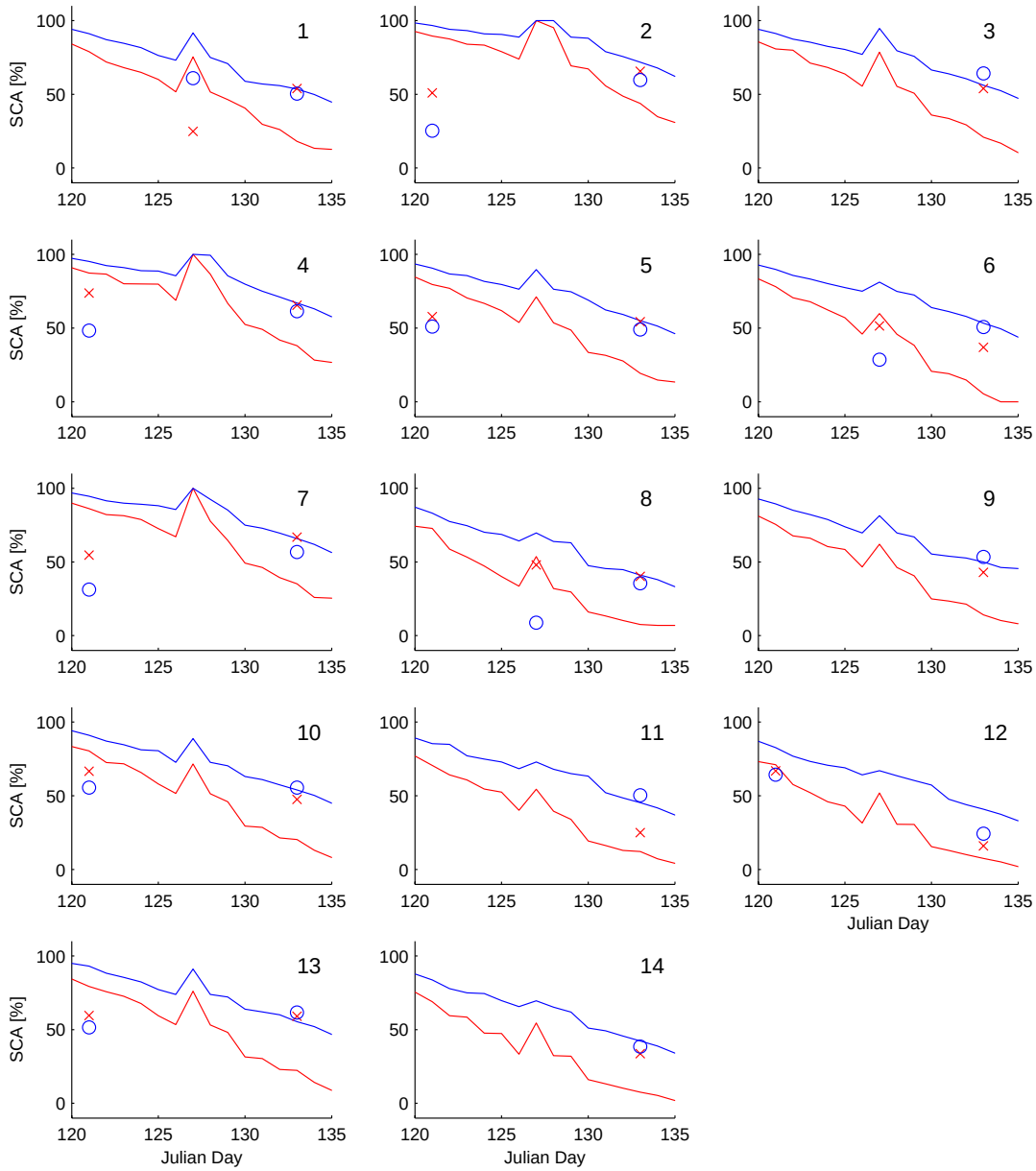

Figure 21. Estimated SCA values for 1998 compared with the reference SCA values. Solid lines present the WSFS reference data. Blue represents forested areas, red open areas. The SCA estimates are drawn by circles and crosses with respective colours. The number of each drainage basin is presented in the figure.

Figure 21 presents the snow melt season of 1998. Unfortunately, there are only few images for this year. The drainage areas 3, 9, 11 and 14 only have one sample, and the other ten drainage areas have only two samples. All the images are from the middle of the snow melt season, and the overall accuracy is not very good in these cases. For this year there seems to be equally large errors for both the open and forested areas.

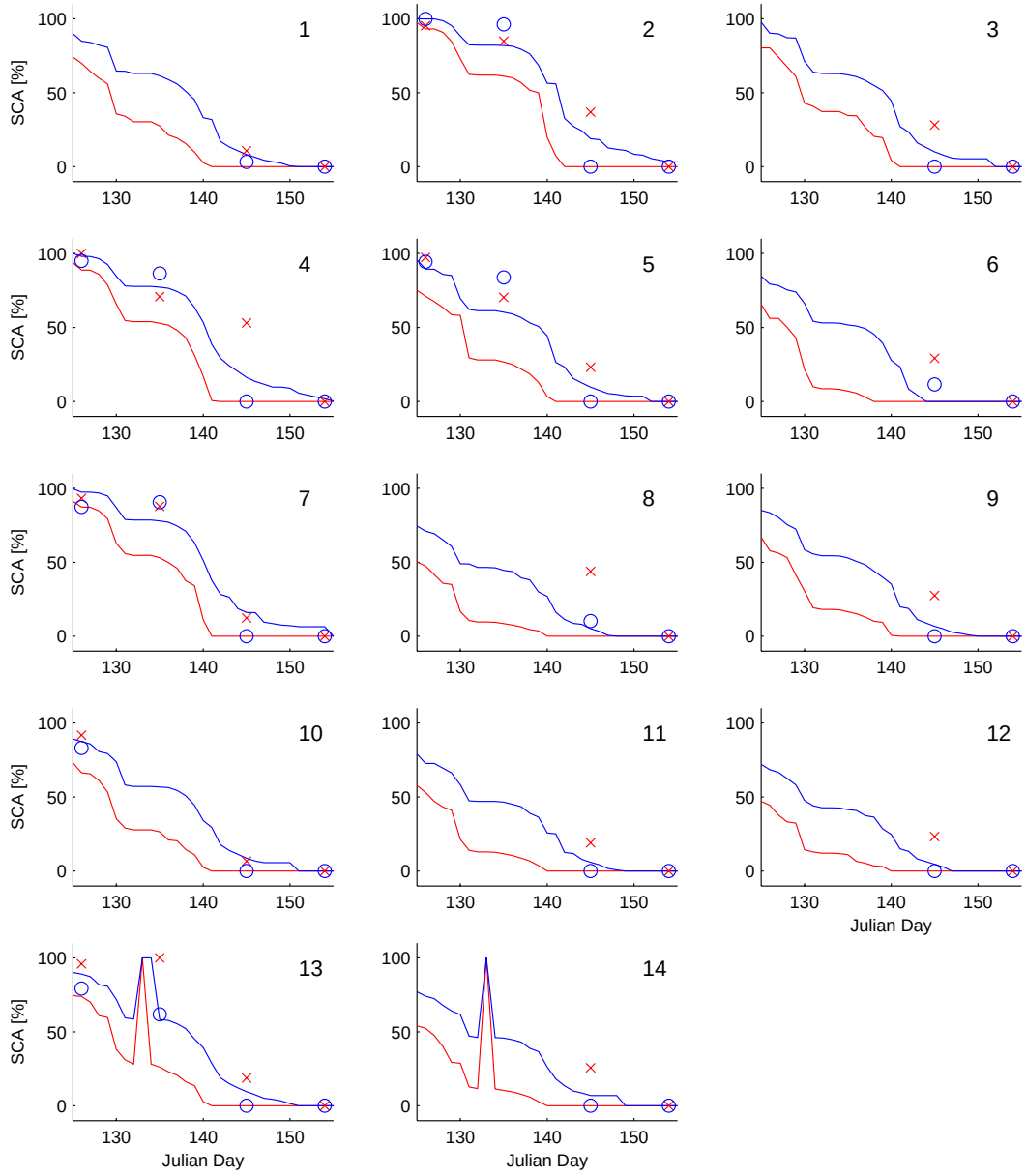

Figure 22. Estimated SCA values for 2000 compared with the reference SCA values. Solid lines present the WSFS reference data. Blue represents forested areas, red open areas. The SCA estimates are drawn by circles and crosses with respective colours. The number of each drainage basin is presented in the figure.

Figure 22 presents the snow melt season of 2000. The amount of samples vary quite much between the different drainage areas. Some areas have only two samples, and the best areas have four samples. The differences in the reference SCA data for the different drainage areas are remarkable, for the drainage areas 13 and 14 there seems to be a very notable increase in the SCA value between 130th and 135th Julian day, and the SCA estimates seem to observe the increased SCA very well. The SCA estimates seem to agree well with the reference data overall. Again the open areas are clearly over-estimated, and the forested areas seem to have reasonably good estimation results.

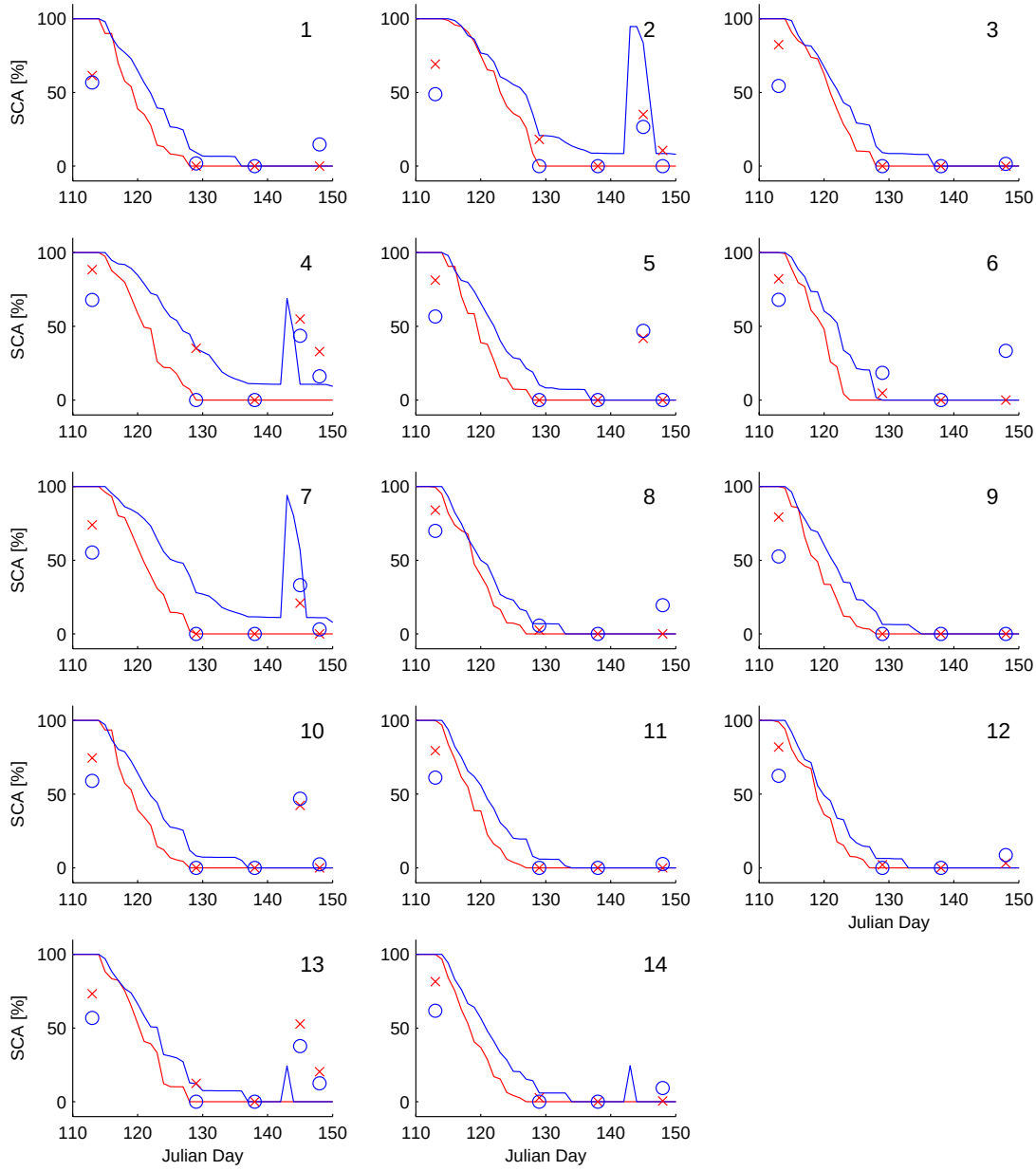

Figure 23. Estimated SCA values for 2001 compared with the reference SCA values. Solid lines present the WSFS reference data. Blue represents forested areas, red open areas. The SCA estimates are drawn by circles and crosses with respective colours. The number of each drainage basin is presented in the figure.

Figure 23 presents the snow melt season of 2001. The estimation for this year seems very promising. The snow melt season seems to end quite early, and the SCA estimates seem to agree with the reference data quite well. An interesting observation can be made at the beginning and at the end of the snow melt season. The first samples near the 110th Julian day are all underestimated. It seems that the backscattering value for this day has been somewhat higher than in the other SCA 100% situations. This causes the underestimation in the data points. Also at the end of the snow melt season, near the Julian day 150 the ground has clearly already started to dry, so the backscattering

coefficient has started to decrease, and this is interpreted as snow on the ground. Again there has been snowfall during the snow melt season which is seen as an increase in SCA data near the 145th Julian day. The reference data shows this increase in some drainage areas, and the estimates agree quite well. In some areas though there is a clear indication of an increased SCA in the estimate although the reference data shows a SCA of 0%. The reference is probably showing false values for these occasions, since the weather information for the time period agrees with snow fall.

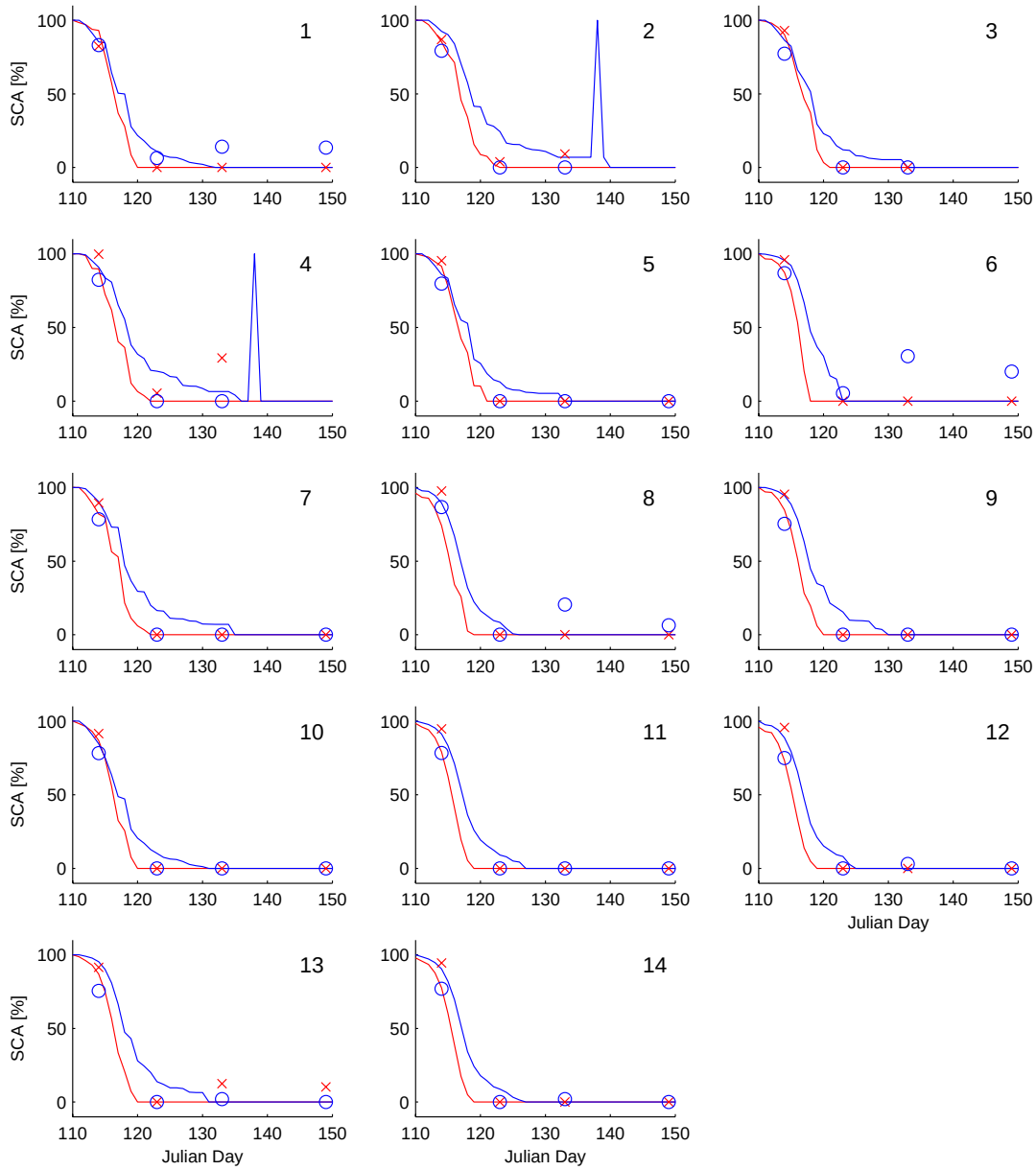

Figure 24. Estimated SCA values for 2002 compared with the reference SCA values. Solid lines present the WSFS reference data. Blue represents forested areas, red open areas. The SCA estimates are drawn by circles and crosses with respective colours The number of each drainage basin is presented in the figure.

Figure 24 presents the snow melt season of 2002. For this year the snow melt season happened very early in the spring, and most of the acquired images are from the end of the snow melt season. The agreement between the reference data and the SCA estimates is however quite good. The reference data shows again an increase in the SCA near the 140th Julian day, but unfortunately there are no satellite images from this time period. The SCA estimation seems to succeed very well for the snow melt season of 2002.

As a conclusion for the visual comparison of the SCA estimates, the reference data and the estimates seemed to agree quite well. The most notable observation was the overestimation of the open areas, which was already seen in the statistical analysis. The estimates for the forested areas seemed to agree very well with the reference data. The beginning and the end of the snow melt seasons were easily identified so the method seems to be quite suitable for the snow melt season monitoring.

A visualisation of the outcome of SCA estimation for 28 May 1997 is shown in Figure 25. The image is acquired at the middle of the snow melt season, so the variation in estimates on the different drainage areas is shown.

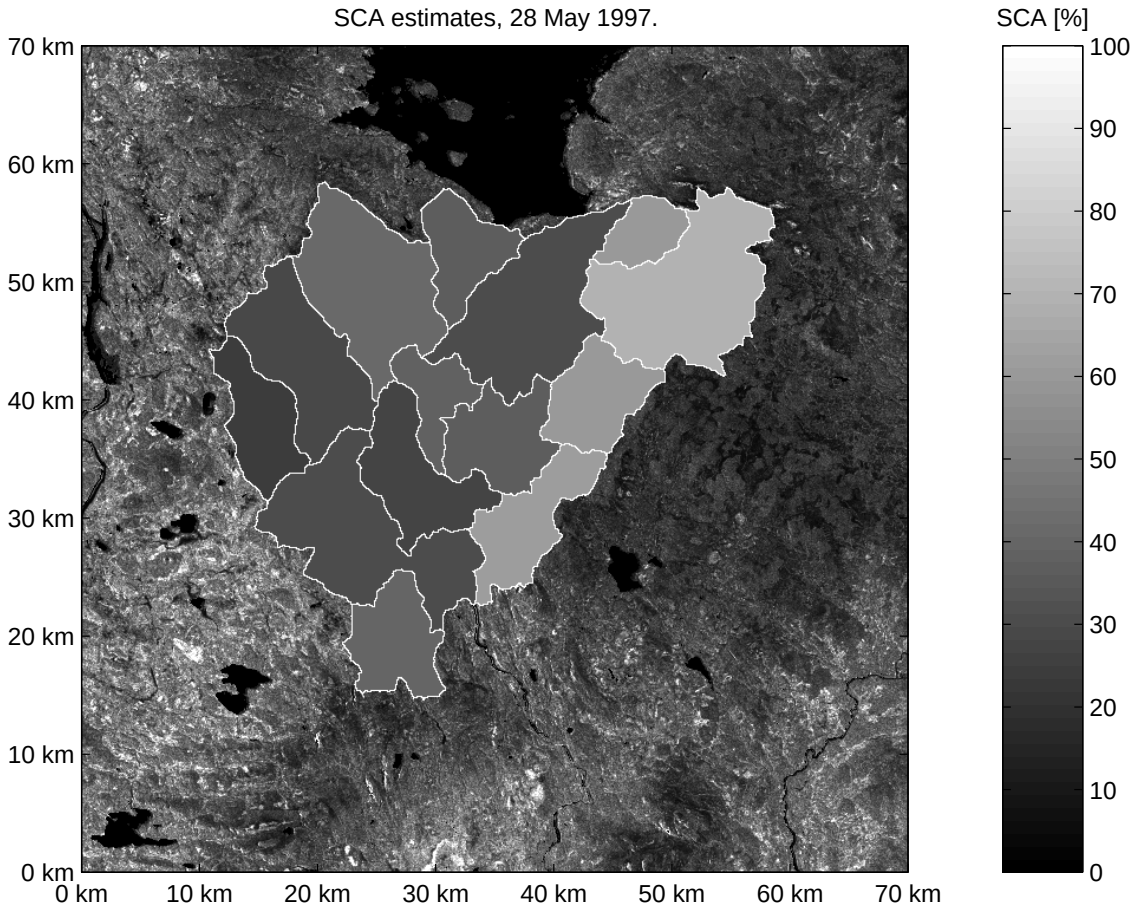

Figure 25. An example of SCA estimation for 28 May 1997 (148th Julian day). Relative spatial scale shown in kilometres. The SCA estimation is drawn on top of the ERS-2 SAR PRI image. The estimates are for open areas, and they are acquired with the best reference image pair.

5 Summary and Conclusions

The purpose of this work was to resolve the statistical accuracy of the HUT Snow Covered Area estimation method. This is required in order to integrate the SCA method for the SYKE's operational Watershed Simulation and Forecasting System. The statistical accuracy analysis was also an effective way in demonstrating the overall usability of the HUT SCA estimation method for the snow melt season monitoring. This work has proven the usability of the HUT method for the SCA monitoring, and has provided the required information on statistical accuracy.

The main goal of this study was to acquire the results for the statistical accuracy of the HUT Snow Covered Area estimation method. This has been accomplished concerning the dataset covering years 1997, 1998, 2000, 2001 and 2002. The dataset consisted of 24 images spread over five years. The images were divided into 14 distinct drainage areas covering a total area of 1160 km². The total amount of samples for the accuracy analysis was 7280, which provides a good basis for a comprehensive statistical analysis. The results for the statistical analyses are presented on pages 29-33.

The examination of the reference image selection was a significant aspect of the study. The importance of reference image selection was clearly proven. In addition the universal method for reference image selection was investigated. The conditions that are required for the reference image selection were clarified, but no well-defined universal method for the reference image selection was found. These results are presented on pages 34-37.

The effects of topography and satellite flight direction on the SCA estimation were studied, and the results were clear. The amount of topography did correlate with the SCA estimation accuracy, and the satellite flight direction seems to have an effect on the estimation. These aspects could be studied more thoroughly in future. The results for the effects of topography are presented on pages 38-41, and the results for the effects of satellite flight path are presented on pages 41-47.

The effect of the forest compensation algorithm was analysed on pages 48-49. The results were obvious; the forest compensation significantly increases the accuracy of the SCA estimation.

The effect of bogs on the SCA estimation was studied on pages 50-51. The obtained results show that the inclusion of bogs on the SCA estimation data is highly recommendable.

Categorical accuracy analysis of the SCA estimation method was conducted using two different reference datasets on pages 51-52. The results of categorical accuracy analysis were mediocre.

Lastly, the SCA estimation was compared with the reference data using a visual presentation of each snow melt season on pages 52-58. This evaluation proved the overall usability of the method for snow melt season monitoring.

References

Bevington, Philip R., Robinson, Keith D., “Data Reduction and Error Analysis for the Physical Sciences” , 2nd Edition, 1992, McGraw-Hill, Inc.

ERS Missions, ESA website: ”http://earth.esa.int/ers/”

Fung, A. (1994), ”Microwave Scattering and Emission Models and Their Applications”, Artech House.

Guneriussen, T., Johnsen, H., and Sand, K. (1996), “DEM corrected ERS-1 SAR data for snow monitoring”, Int. J. Remote Sens. 17(1).

Koskinen, J., Kurvonen, L., Jääskeläinen, V., and Hallikainen, M. (1994), ”Capability of radar and microwave radiometer to classify snow types in forested areas”, Proc. IGARSS’94, Pasadena, USA, 1283-1286.

Koskinen, J., Pulliainen, J., and Hallikainen, M. (1997), ”The use of ERS-1 SAR data in snow melt monitoring”, IEEE Trans. Geosci. Remote Sens. 35:601-610.

Koskinen, J., Metsämäki, S., Grandell, J., Jänne, S., Matikainen, L., and Hallikainen, M. (1999), ”Snow monitoring using radar and optical satellite data”, Remote Sensing Environ. 69:16-29.

Luojus, K., Hallikainen, M., Pulliainen, J., Heikkilä, S., Engdahl, M., “Topography corrected SAR image processing for ENVISAT ASAR data”, URSI/IEEE XXVII Convention on Radio Science, Espoo 2002.

Nagler, T., and Rott, H. (1998), ”SAR tools for snowmelt modeling in the project HydAlp”, Proc. IGARSS’98, Seattle, USA.

Piesbergen, J., Holecz, F., and Haefner, H. (1995), “Snow cover monitoring using multitemporal ERS-1 SAR data”, Proc. IGARSS’95, Florence, Italy, 1750-1752.

Pulliainen, J., Kurvonen, L., Hallikainen, M. (1999), ”Multi-temporal behavior of L- and C-band SAR observations of boreal forest”, IEEE Trans. Geosci. Remote Sensing, vol. 37, pp. 927-937.

Pulliainen, J., Koskinen, J., and Hallikainen, M. (2001), ”Compensation of forest canopy effects in the estimation of snow covered area from SAR data”, Proc. IGARSS'01, Sydney, Australia, 9-13 July 2001.

Rauste, Y. (1989). ”Methods for analysing SAR images”, VTT Research Reports 612, Technical Research Centre of Finland Espoo

Rott, H. (1984), “The analysis of backscattering properties from SAR data of mountain region”, IEEE J. Oceanic Eng. 9:347-355.

Shi, J., and Dozier, J. (1995), Inferring snow wetness using C-band data from SIR-C’s polarimetric synthetic aperture radar. IEEE Trans. Geosci. Remote Sens. 33: 905-914.

Stiles, W., and Ulaby, F. (1980), ”The active and passive microwave response to snow parameters: Part I – wetness”, J. Geophys. Research, (85):1037-1044.

Strozzi, T., Wiesmann, A., and Mätzler, C. (1997), ”Active microwave signatures of snow covers at 5.3 and 35 GHz”, Radio Science 32:479-495.

Ulaby, F., Moore, R., and Fung A. (1982), ”Microwave Remote Sensing, Active and Passive, Volume II”, Reading, Addison Wesley, 609 p.

Ulaby, F., Moore, R., and Fung, A. (1986), ”Microwave Remote Sensing, Active and Passive,Vol. III”, Artech House.

Vehviläinen, B. (1994), ”The watershed simulation and forecasting system in the National Board of Waters and Environment”, Publications of the Water and Environment Research Institute, No.17, National Board of Waters and the Environment, Finland.

Wegmüller, U., Werner, C., Strozzi, T. (1998), ”SAR Interferometric and Differential Interferometric Processing Chain”, Proceedings of IGARSS'98, Seattle, USA, 6-10 July 1998, vol.2, pp.1106-1108

Appendix A: Statistical Parameters of the Complete Dataset

COMPLETE DATASET - ROOT MEAN SQUARE ERROR

	RMSE, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.223	0.330	0.346	0.267	0.262	0.221	0.240	0.270
268	12.5.1997	0.220	0.325	0.341	0.268	0.254	**0.213**	0.233	**0.265**
128	5.5.2000	0.243	0.341	0.323	0.249	0.288	0.247	0.260	0.2789
108	14.5.2000	0.289	0.387	0.372	0.281	0.341	0.306	0.314	0.327
268	24.4.2002	0.232	0.341	0.357	0.282	0.269	0.224	0.246	0.2788
Average		0.234	0.339	0.348	0.270	0.273	**0.232**	0.250	

Average of all samples: 0.278

	RMSE, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.184	0.273	0.199	0.174	0.180	0.176	0.173	0.194
268	12.5.1997	0.182	0.263	0.195	0.181	0.179	**0.179**	0.174	**0.193**
128	5.5.2000	0.205	0.284	0.205	0.205	0.206	0.202	0.201	0.216
108	14.5.2000	0.221	0.298	0.224	0.221	0.227	0.219	0.220	0.233
268	24.4.2002	0.200	0.302	0.219	0.178	0.193	0.173	0.176	0.206
Average		0.194	0.282	0.206	0.186	0.191	0.184	**0.182**	

Average of all samples: 0.204

	Average RMSE								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.203	0.301	0.272	0.220	0.221	0.199	0.207	0.232
268	12.5.1997	0.201	0.294	0.268	0.225	0.217	**0.196**	0.204	**0.229**
128	5.5.2000	0.224	0.313	0.264	0.227	0.247	0.224	0.230	0.247
108	14.5.2000	0.255	0.342	0.298	0.251	0.284	0.262	0.267	0.280
268	24.4.2002	0.216	0.322	0.288	0.230	0.231	0.199	0.211	0.242
Average		0.214	0.310	0.277	0.228	0.232	**0.208**	0.216	

Average of all samples: 0.241

The chart shows the average RMSE values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

COMPLETE DATASET - MEAN ABSOLUTE ERROR

	Mean Absolute Error, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.149	0.277	0.293	0.205	0.193	0.141	0.165	0.203
268	12.5.1997	0.147	0.274	0.290	0.206	0.189	**0.137**	0.162	**0.201**
128	5.5.2000	0.166	0.286	0.266	0.172	0.222	0.173	0.187	0.210
108	14.5.2000	0.192	0.317	0.297	0.190	0.256	0.209	0.220	0.240
268	24.4.2002	0.155	0.287	0.302	0.216	0.199	0.144	0.170	0.211
Average		0.156	0.284	0.292	0.203	0.204	**0.152**	0.174	

Average of all samples: 0.209

	Mean Absolute Error, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.140	0.234	0.157	0.119	0.133	0.118	0.120	0.146
268	12.5.1997	0.138	0.226	0.153	0.124	0.132	**0.119**	0.120	**0.145**
128	5.5.2000	0.156	0.240	0.157	0.147	0.153	0.143	0.146	0.163
108	14.5.2000	0.163	0.247	0.167	0.159	0.165	0.155	0.159	0.174
268	24.4.2002	0.156	0.260	0.176	0.129	0.147	0.122	0.128	0.160
Average		0.148	0.241	0.162	0.131	0.142	**0.126**	0.129	

Average of all samples: 0.154

	Average Mean Absolute Error								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.144	0.256	0.225	0.162	0.163	0.129	0.143	0.175
268	12.5.1997	0.142	0.250	0.221	0.165	0.160	**0.128**	0.141	**0.173**
128	5.5.2000	0.161	0.263	0.211	0.159	0.188	0.158	0.166	0.187
108	14.5.2000	0.177	0.282	0.232	0.174	0.211	0.182	0.189	0.207
268	24.4.2002	0.155	0.274	0.239	0.173	0.173	0.133	0.149	0.185
Average		0.1521	0.263	0.227	0.167	0.173	**0.139**	0.1518	

Average of all samples: 0.182

The chart shows the average Mean Absolute Errors calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

COMPLETE DATASET - BIAS

	Bias, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.116	0.258	0.273	0.177	0.166	0.103	0.134	0.175
268	12.5.1997	0.107	0.251	0.266	0.169	0.158	**0.095**	0.126	**0.168**
128	5.5.2000	0.121	0.260	0.237	0.121	0.190	0.131	0.148	0.172
108	14.5.2000	0.164	0.302	0.279	0.162	0.238	0.184	0.197	0.218
268	24.4.2002	0.125	0.270	0.286	0.189	0.177	0.113	0.144	0.186
Average		0.122	0.264	0.271	0.170	0.177	**0.115**	0.143	

Average of all samples
0.180

	Bias, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.044	0.185	0.075	-0.021	0.031	-0.032	-0.015	0.038
268	12.5.1997	0.027	0.168	0.058	-0.038	0.015	**-0.049**	-0.031	**0.021**
128	5.5.2000	0.042	0.188	0.052	-0.040	0.028	-0.043	-0.012	0.031
108	14.5.2000	0.054	0.201	0.070	-0.026	0.046	-0.028	0.003	0.046
268	24.4.2002	0.081	0.225	0.113	0.014	0.067	0.002	0.019	0.074
Average		0.050	0.193	0.077	-0.019	0.037	-0.029	**-0.008**	

Average of all samples
0.043

	Average Bias								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.080	0.221	0.174	0.078	0.099	0.035	0.060	0.107
268	12.5.1997	0.067	0.210	0.162	0.066	0.087	**0.023**	0.048	**0.095**
128	5.5.2000	0.081	0.224	0.145	0.040	0.109	0.044	0.068	0.102
108	14.5.2000	0.109	0.251	0.175	0.068	0.142	0.078	0.100	0.132
268	24.4.2002	0.103	0.248	0.200	0.101	0.122	0.057	0.082	0.130
Average		0.086	0.229	0.174	0.075	0.107	**0.043**	0.068	

Average of all samples
0.112

The chart shows the average Bias values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

COMPLETE DATASET - CORRELATION COEFFICIENT

	Correlation coefficient, open areas							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
268	9.5.1997	0.873	0.827	0.815	0.849	0.850	0.869	0.861	
268	12.5.1997	0.869	0.826	0.813	0.833	0.852	**0.873**	0.862	
128	5.5.2000	0.850	0.812	0.817	0.839	0.831	0.853	0.844	
108	14.5.2000	0.828	0.779	0.779	0.837	0.797	0.815	0.810	Average of all samples
268	24.4.2002	0.872	0.825	0.814	0.836	0.854	0.878	0.865	0.830

	Correlation coefficient, forested areas							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
268	9.5.1997	0.888	0.859	0.879	0.898	0.891	0.898	0.899	
268	12.5.1997	0.885	0.859	0.876	0.890	0.887	**0.896**	0.897	
128	5.5.2000	0.861	0.837	0.863	0.870	0.858	0.875	0.869	
108	14.5.2000	0.846	0.819	0.844	0.854	0.836	0.858	0.850	Average of all samples
268	24.4.2002	0.887	0.854	0.877	0.901	0.891	0.908	0.905	0.860

The chart shows the average Correlation coefficients calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated by using the complete dataset. The chosen best reference image value is highlighted.

Appendix B: Statistical Parameters of Ascending and Descending Node Datasets

ASCENDING AND DESCENDING NODE DATASETS - ROOT MEAN SQUARE ERROR

ASCENDING NODE	Samples: 1480						
	RMSE, open areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.376	0.312	0.298	0.257	0.280	0.305
123	12.5.1997	0.371	0.311	0.291	**0.247**	0.272	**0.298**
50	14.5.2000	0.406	0.347	0.380	0.351	0.363	0.369
Average		0.379	0.318	0.309	**0.269**	0.290	**0.313**
	RMSE, forested areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.160	0.110	0.129	0.106	0.112	0.123
123	12.5.1997	0.149	0.110	0.120	**0.100**	0.103	**0.117**
50	14.5.2000	0.205	0.185	0.199	0.178	0.186	0.190
Average		0.163	0.123	0.137	**0.116**	0.121	**0.132**
	Average RMSE						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.268	0.211	0.214	0.181	0.196	0.214
123	12.5.1997	0.260	0.210	0.205	**0.174**	0.188	**0.207**
50	14.5.2000	0.305	0.266	0.289	0.265	0.274	0.280
Average		0.271	0.220	0.223	**0.192**	0.206	**0.222**

DESCENDING NODE	Samples: 430			
	RMSE, open areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.201	0.320	0.260
145	24.4.2002	**0.177**	0.309	**0.243**
Average		**0.185**	0.312	**0.249**
	RMSE, forested areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.234	0.294	0.264
145	24.4.2002	**0.212**	0.307	**0.260**
Average		**0.219**	0.303	**0.261**
	Average RMSE			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.217	0.307	0.262
145	24.4.2002	**0.194**	0.308	**0.251**
Average		**0.202**	0.308	**0.255**

The charts show the average RMSE values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The results for the ascending and descending node have been calculated separately.

ASCENDING AND DESCENDING NODE DATASETS - MEAN ABSOLUTE ERROR

ASCENDING NODE	Samples: 1480						
	Mean Absolute Error, open areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.321	0.248	0.226	0.173	0.200	0.234
123	12.5.1997	0.316	0.245	0.221	**0.167**	0.195	**0.229**
50	14.5.2000	0.327	0.254	0.291	0.247	0.261	0.276
Average		0.320	0.248	0.235	**0.183**	0.208	**0.239**
	Mean Absolute Error, forested areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.110	0.087	0.102	0.085	0.088	0.094
123	12.5.1997	0.131	0.078	0.098	**0.075**	0.081	**0.093**
50	14.5.2000	0.121	0.081	0.091	0.072	0.076	0.088
Average		0.121	0.082	0.098	**0.079**	0.083	**0.093**
	Average Mean Absolute Error						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.215	0.167	0.164	0.129	0.144	0.164
123	12.5.1997	0.224	0.161	0.160	**0.121**	0.138	**0.161**
50	14.5.2000	0.224	0.167	0.191	0.160	0.168	0.182
Average		0.220	0.165	0.167	**0.131**	0.146	**0.166**

DESCENDING NODE	Samples: 430			
	Mean Absolute Error, open areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.135	0.268	0.201
145	24.4.2002	**0.115**	0.258	**0.186**
Average		**0.121**	0.261	**0.191**
	Mean Absolute Error, forested areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.177	0.234	0.206
145	24.4.2002	**0.166**	0.255	**0.211**
Average		**0.170**	0.248	**0.209**
	Average Mean Absolute Error			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.156	0.251	0.204
145	24.4.2002	**0.140**	0.257	**0.198**
Average		**0.146**	0.255	**0.200**

The chart shows the average Mean Absolute Errors calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category. The results for the ascending and descending node have been calculated separately.

ASCENDING AND DESCENDING NODE DATASETS - BIAS

ASCENDING NODE		Samples: 1480					
		Bias, open areas					
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.321	0.248	0.226	0.170	0.200	0.233
123	12.5.1997	0.312	0.235	0.217	**0.160**	0.190	**0.223**
50	14.5.2000	0.327	0.254	0.291	0.247	0.261	0.276
Average		0.318	0.244	0.233	**0.179**	0.206	**0.236**
		Bias, forested areas					
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.128	0.047	0.087	0.030	0.048	0.068
123	12.5.1997	0.110	0.026	0.068	**0.011**	0.029	**0.049**
50	14.5.2000	0.139	0.059	0.115	0.054	0.083	0.090
Average		0.122	0.040	0.084	**0.026**	0.046	**0.064**
		Average Bias					
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.225	0.147	0.157	0.100	0.124	0.151
123	12.5.1997	0.211	0.130	0.143	**0.085**	0.109	**0.136**
50	14.5.2000	0.233	0.157	0.203	0.151	0.172	0.183
Average		0.220	0.142	0.159	**0.103**	0.126	**0.150**

DESCENDING NODE		Samples: 430		
	Bias, open areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.062	0.226	**0.144**
145	24.4.2002	**0.064**	0.230	0.147
Average		**0.063**	0.229	**0.146**
	Bias, forested areas			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	-0.011	0.145	**0.067**
145	24.4.2002	**0.034**	0.190	0.112
Average		**0.019**	0.176	**0.098**
	Average Bias			
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.026	0.186	**0.106**
145	24.4.2002	**0.049**	0.210	0.130
Average		**0.041**	0.202	**0.122**

The chart shows the average Bias values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category. The results for the ascending and descending node have been calculated separately.

ASCENDING AND DESCENDING NODE DATASETS - CORRELATION COEFFICIENT

ASCENDING NODE	Samples: 1480						
	Correlation coefficient, open areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.830	0.866	0.880	0.906	0.890	
123	12.5.1997	0.816	0.836	0.872	**0.901**	0.881	
50	14.5.2000	0.790	0.831	0.812	0.833	0.822	
Average							**0.852**
	Correlation coefficient, forested areas						
Samples	Images	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
123	9.5.1997	0.967	0.977	0.974	0.980	0.978	
123	12.5.1997	0.961	0.965	0.968	**0.974**	0.973	
50	14.5.2000	0.922	0.921	0.917	0.931	0.926	
Average							**0.958**

DESCENDING NODE	Samples: 430			
Correlation coefficient, open areas				
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.891	0.850	
145	24.4.2002	**0.914**	0.866	
Average				**0.857**
Correlation coefficient, forested areas				
Samples	Images	4.6.1997	7.6.1997	Average
70	5.5.2000	0.823	0.786	
145	24.4.2002	**0.858**	0.808	
Average				**0.799**

The chart shows the average Correlation coefficients calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated by using the complete dataset. The chosen best reference image value is highlighted. The results for the ascending and descending node have been calculated separately

Appendix C: Statistical Parameters of the Non-forest compensated Dataset

NON-FOREST COMPENSATED DATASET - ROOT MEAN SQUARE ERROR

	RMSE, Forested areas, Forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.184	0.273	0.199	0.174	0.180	0.176	0.173	0.194
268	12.5.1997	0.182	0.263	0.195	0.181	0.179	**0.179**	0.174	**0.193**
128	5.5.2000	0.205	0.284	0.205	0.205	0.206	0.202	0.201	0.216
108	14.5.2000	0.221	0.298	0.224	0.221	0.227	0.219	0.220	0.233
268	24.4.2002	0.200	0.302	0.219	0.178	0.193	0.173	0.176	0.206
Average		0.194	0.282	0.206	0.186	0.191	0.184	**0.182**	

Average of all samples
0.204

	RMSE, Forested areas, Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.208	0.285	0.207	0.206	0.205	0.204	0.202	0.216
268	12.5.1997	0.203	0.272	0.202	0.212	0.201	**0.202**	0.201	**0.213**
128	5.5.2000	0.221	0.295	0.217	0.233	0.222	0.222	0.221	0.233
108	14.5.2000	0.221	0.288	0.220	0.252	0.217	0.238	0.231	0.238
268	24.4.2002	0.224	0.317	0.223	0.204	0.219	0.205	0.204	0.228
Average		0.214	0.291	0.212	0.215	0.211	0.209	**0.208**	

Average of all samples
0.223

	Difference of Forest compensated and Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.024	0.012	0.008	0.031	0.025	0.027	0.029	0.022
268	12.5.1997	0.022	0.009	0.007	0.031	0.022	0.024	0.026	0.020
128	5.5.2000	0.016	0.011	0.011	0.028	0.015	0.020	0.021	0.018
108	14.5.2000	0.000	-0.010	-0.005	0.031	-0.010	0.019	0.011	0.005
268	24.4.2002	0.024	0.015	0.005	0.026	0.026	0.032	0.029	0.022
Average		0.020	0.009	0.006	0.029	0.020	0.026	0.025	

Average of all samples
0.019

The chart shows the average RMSE values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. A positive difference (black colour) shows that the forest compensated value is better.

NON-FOREST COMPENSATED DATASET - MEAN ABSOLUTE ERROR

	Mean Absolute Error, Forested areas, Forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.140	0.234	0.157	0.119	0.133	0.118	0.120	0.146
268	12.5.1997	0.138	0.226	0.153	0.124	0.132	**0.119**	0.120	**0.145**
128	5.5.2000	0.156	0.240	0.157	0.147	0.153	0.143	0.146	0.163
108	14.5.2000	0.163	0.247	0.167	0.159	0.165	0.155	0.159	0.174
268	24.4.2002	0.156	0.260	0.176	0.129	0.147	0.122	0.128	0.160
Average		0.148	0.241	0.162	0.131	0.142	**0.126**	0.129	

Average of all samples: 0.154

	Mean Absolute Error, Forested areas, Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.138	0.231	0.160	0.128	0.140	0.129	0.130	0.151
268	12.5.1997	0.137	0.224	0.157	0.131	0.139	**0.132**	0.131	**0.150**
128	5.5.2000	0.156	0.236	0.160	0.156	0.158	0.158	0.158	0.169
108	14.5.2000	0.168	0.248	0.172	0.169	0.171	0.169	0.169	0.181
268	24.4.2002	0.149	0.255	0.176	0.134	0.154	0.136	0.137	0.163
Average		0.146	0.238	0.164	**0.138**	0.149	0.139	0.139	

Average of all samples: 0.159

	Difference of Forest compensated and Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	-0.002	-0.003	0.003	0.009	0.007	0.011	0.009	0.005
268	12.5.1997	-0.001	-0.002	0.004	0.007	0.007	0.013	0.011	0.006
128	5.5.2000	0.000	-0.003	0.003	0.010	0.005	0.016	0.012	0.006
108	14.5.2000	0.005	0.001	0.005	0.010	0.005	0.014	0.011	0.007
268	24.4.2002	-0.007	-0.005	0.000	0.005	0.007	0.015	0.009	0.003
Average		-0.002	-0.003	0.003	0.008	0.007	0.013	0.010	

Average of all samples: 0.005

The chart shows the average Mean Absolute Errors calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. A positive difference (black colour) shows that the forest compensated value is better.

NON-FOREST COMPENSATED DATASET - BIAS

	Bias, Forested areas, Forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.044	0.185	0.075	-0.021	0.031	-0.032	-0.015	0.038
268	12.5.1997	0.027	0.168	0.058	-0.038	0.015	**-0.049**	-0.031	**0.021**
128	5.5.2000	0.042	0.188	0.052	-0.040	0.028	-0.043	-0.012	0.031
108	14.5.2000	0.054	0.201	0.070	-0.026	0.046	-0.028	0.003	0.046
268	24.4.2002	0.081	0.225	0.113	0.014	0.067	0.002	0.019	0.074
Average		0.050	0.193	0.077	-0.019	0.037	-0.029	**-0.008**	

Average of all samples
0.043

	Bias, Forested areas, Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	0.028	0.175	0.074	-0.035	0.039	-0.024	-0.015	0.035
268	12.5.1997	0.013	0.160	0.059	-0.050	0.024	**-0.038**	-0.030	**0.020**
128	5.5.2000	0.023	0.174	0.049	-0.058	0.033	-0.042	-0.015	0.023
108	14.5.2000	0.042	0.193	0.071	-0.040	0.053	-0.028	0.000	0.042
268	24.4.2002	0.061	0.213	0.110	-0.004	0.074	0.008	0.016	0.068
Average		0.033	0.183	0.076	-0.034	0.045	-0.022	**-0.009**	

Average of all samples
0.039

	Difference of Forest compensated and Non-forest compensated values								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
268	9.5.1997	-0.016	-0.010	-0.000	-0.014	0.008	0.009	-0.001	-0.003
268	12.5.1997	-0.014	-0.008	0.001	-0.012	0.010	0.011	0.001	-0.002
128	5.5.2000	-0.020	-0.014	-0.003	-0.018	0.006	0.000	-0.003	-0.007
108	14.5.2000	-0.012	-0.008	0.001	-0.014	0.007	-0.000	-0.003	-0.004
268	24.4.2002	-0.020	-0.012	-0.003	-0.017	0.006	0.006	-0.004	-0.006
Average		-0.017	-0.010	-0.001	-0.015	0.008	0.007	-0.001	

Average of all samples
-0.004

The chart shows the average Bias values calculated from all the SCA estimates obtained with the certain reference image pair combination Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. . The chosen best reference image value is highlighted. A positive difference (black colour) shows that the forest compensated value is better.

NON-FOREST COMPENSATED DATASET - CORRELATION COEFFICIENT

	Correlation coefficient, Forested areas, Forest compensated values							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
268	9.5.1997	0.888	0.859	0.879	0.898	0.891	0.898	0.899	
268	12.5.1997	0.885	0.859	0.876	0.890	0.887	**0.896**	0.897	
128	5.5.2000	0.861	0.837	0.863	0.870	0.858	0.875	0.869	
108	14.5.2000	0.846	0.819	0.844	0.854	0.836	0.858	0.850	Average of all samples
268	24.4.2002	0.887	0.854	0.877	0.901	0.891	0.908	0.905	0.860

	Correlation coefficient Forested areas, Non-forest compensated values							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
268	9.5.1997	0.876	0.851	0.865	0.877	0.874	0.877	0.878	
268	12.5.1997	0.872	0.848	0.862	0.871	0.869	**0.871**	0.873	
128	5.5.2000	0.841	0.819	0.841	0.839	0.835	0.836	0.835	
108	14.5.2000	0.824	0.799	0.823	0.824	0.816	0.821	0.819	Average of all samples
268	24.4.2002	0.877	0.845	0.864	0.881	0.872	0.878	0.880	0.841

	Difference of Forest compensated and Non-forest compensated values							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
268	9.5.1997	0.012	0.008	0.014	0.021	0.017	0.021	0.021	
268	12.5.1997	0.012	0.011	0.014	0.019	0.018	0.025	0.024	
128	5.5.2000	0.020	0.018	0.021	0.030	0.024	0.039	0.034	
108	14.5.2000	0.022	0.020	0.020	0.030	0.021	0.037	0.031	Average of all samples
268	24.4.2002	0.010	0.009	0.014	0.020	0.020	0.030	0.025	0.019

The chart shows the average Correlation coefficients calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated by using the complete dataset. The chosen best reference image value is highlighted. A positive difference (black colour) shows that the forest compensated value is better.

Appendix D: Statistical Parameters of the Dataset Without Bogs

DATASET WITHOUT BOGS - ROOT MEAN SQUARE ERROR

	RMSE, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.260	0.356	0.314	**0.244**	0.322	0.305	0.301	0.300
266	12.5.1997	0.247	0.338	0.293	0.244	0.302	0.285	0.281	**0.284**
127	5.5.2000	0.277	0.362	0.322	0.282	0.339	0.327	0.316	0.318
86	14.5.2000	0.342	0.427	0.405	0.495	0.404	0.400	0.386	0.408
266	24.4.2002	0.272	0.379	0.326	0.262	0.334	0.312	0.308	0.313
Average		**0.269**	0.364	0.320	0.275	0.329	0.312	0.307	

Average of all samples
0.311

	RMSE, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.208	0.285	0.207	**0.206**	0.205	0.204	0.202	0.217
266	12.5.1997	0.203	0.272	0.202	0.212	0.201	0.203	0.201	**0.214**
127	5.5.2000	0.221	0.295	0.217	0.233	0.222	0.222	0.221	0.233
86	14.5.2000	0.221	0.288	0.220	0.252	0.217	0.238	0.231	0.238
266	24.4.2002	0.223	0.317	0.223	0.204	0.219	0.204	0.204	0.228
Average		0.214	0.291	0.212	0.214	0.211	0.209	**0.207**	

Average of all samples
0.223

	Average RMSE								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.234	0.320	0.260	**0.225**	0.264	0.254	0.252	0.258
266	12.5.1997	0.225	0.305	0.248	0.228	0.252	0.244	0.241	**0.249**
127	5.5.2000	0.249	0.329	0.269	0.257	0.281	0.275	0.269	0.276
86	14.5.2000	0.281	0.357	0.312	0.373	0.311	0.319	0.309	0.323
266	24.4.2002	0.248	0.348	0.275	0.233	0.276	0.258	0.256	0.271
Average		**0.241**	0.328	0.266	0.245	0.270	0.261	0.257	

Average of all samples
0.267

The chart shows the average RMSE values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

DATASET WITHOUT BOGS - MEAN ABSOLUTE ERROR

	Mean Absolute Error, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.171	0.297	0.237	**0.147**	0.242	0.210	0.211	0.216
266	12.5.1997	0.162	0.285	0.224	0.145	0.230	0.198	0.198	**0.206**
127	5.5.2000	0.183	0.300	0.251	0.170	0.261	0.235	0.229	0.233
86	14.5.2000	0.223	0.343	0.307	0.328	0.298	0.279	0.268	0.292
266	24.4.2002	0.180	0.317	0.249	0.159	0.254	0.217	0.218	0.228
Average		0.177	0.304	0.244	**0.168**	0.249	0.218	0.217	

Average of all samples
0.225

	Mean Absolute Error, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.153	0.242	0.152	**0.131**	0.147	0.136	0.135	0.157
266	12.5.1997	0.150	0.232	0.148	0.135	0.145	0.135	0.135	**0.154**
127	5.5.2000	0.166	0.248	0.161	0.158	0.163	0.154	0.157	0.172
86	14.5.2000	0.165	0.243	0.165	0.173	0.160	0.166	0.164	0.177
266	24.4.2002	0.169	0.271	0.169	0.133	0.162	0.140	0.141	0.170
Average		0.159	0.248	0.158	**0.140**	0.154	0.142	0.142	

Average of all samples
0.163

	Average Mean Absolute Error								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.162	0.270	0.194	**0.139**	0.195	0.173	0.173	0.187
266	12.5.1997	0.156	0.258	0.186	0.140	0.187	0.167	0.167	**0.180**
127	5.5.2000	0.174	0.274	0.206	0.164	0.212	0.195	0.193	0.203
86	14.5.2000	0.194	0.293	0.236	0.251	0.229	0.222	0.216	0.234
266	24.4.2002	0.175	0.294	0.209	0.146	0.208	0.179	0.180	0.199
Average		0.168	0.276	0.2012	**0.154**	0.2014	0.180	0.179	

Average of all samples
0.194

The chart shows the average Mean Absolute Errors calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

DATASET WITHOUT BOGS - BIAS

	Bias, open areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.104	0.259	0.186	**0.059**	0.192	0.147	0.153	0.157
266	12.5.1997	0.084	0.240	0.169	0.041	0.174	0.130	0.135	**0.139**
127	5.5.2000	0.094	0.253	0.193	0.037	0.208	0.173	0.168	0.161
86	14.5.2000	0.162	0.311	0.266	0.106	0.257	0.226	0.218	0.221
266	24.4.2002	0.128	0.291	0.216	0.084	0.221	0.174	0.179	0.185
Average		0.109	0.266	0.197	**0.062**	0.203	0.159	0.162	

Average of all samples: 0.165

	Bias, forested areas								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.037	0.182	0.043	**-0.047**	0.032	-0.021	-0.019	0.030
266	12.5.1997	0.017	0.161	0.022	-0.069	0.011	-0.042	-0.039	0.009
127	5.5.2000	0.032	0.189	0.032	-0.060	0.024	-0.043	-0.022	0.022
86	14.5.2000	0.012	0.179	0.020	-0.076	-0.001	-0.059	-0.042	**0.005**
266	24.4.2002	0.078	0.227	0.084	-0.011	0.072	0.016	0.018	0.069
Average		0.040	0.189	0.045	-0.048	0.033	-0.023	**-0.017**	

Average of all samples: 0.031

	Average Bias								
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	Average
266	9.5.1997	0.071	0.221	0.115	**0.006**	0.112	0.063	0.067	0.093
266	12.5.1997	0.050	0.201	0.095	-0.014	0.093	0.044	0.048	**0.074**
127	5.5.2000	0.063	0.221	0.113	-0.011	0.116	0.065	0.073	0.091
86	14.5.2000	0.087	0.245	0.143	0.015	0.128	0.083	0.088	0.113
266	24.4.2002	0.103	0.259	0.150	0.037	0.146	0.095	0.098	0.127
Average		0.074	0.228	0.121	**0.007**	0.118	0.068	0.073	

Average of all samples: 0.098

The chart shows the average Bias values calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated in respect to the amount of samples. The chosen best reference image value is highlighted. The values in bold show the best value of each category.

DATASET WITHOUT BOGS - CORRELATION COEFFICIENT

	Correlation coefficient, open areas							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
266	9.5.1997	0.797	0.756	0.758	**0.799**	0.748	0.746	0.758	
266	12.5.1997	0.796	0.764	0.774	0.782	0.762	0.759	0.771	
127	5.5.2000	0.766	0.736	0.755	0.736	0.736	0.731	0.748	
86	14.5.2000	0.733	0.686	0.685	0.326	0.680	0.664	0.688	Average of all samples
266	24.4.2002	0.806	0.761	0.779	0.797	0.771	0.772	0.782	0.744

	Correlation coefficient, forested areas							
Samples	Image	4.6.1997	7.6.1997	13.6.1997	16.6.1997	2.6.2000	18.5.2001	3.5.2002	
266	9.5.1997	0.854	0.825	0.856	**0.864**	0.857	0.861	0.863	
266	12.5.1997	0.852	0.828	0.854	0.858	0.855	0.861	0.863	
127	5.5.2000	0.834	0.812	0.842	0.833	0.834	0.846	0.841	
86	14.5.2000	0.832	0.806	0.833	0.812	0.840	0.829	0.832	Average of all samples
266	24.4.2002	0.853	0.821	0.856	0.873	0.859	0.871	0.872	0.830

The chart shows the average Correlation coefficients calculated from all the SCA estimates obtained with the certain reference image pair combination. Rows indicate what reference images for the beginning of snow melt season (SCA 100%) are employed. The reference images for the end of snow melt season (SCA 0%) are situated in different columns. The column of samples shows the amount of SCA estimates for each pair on a certain row. The average values are calculated by using the complete dataset. The chosen best reference image value is highlighted.