Validation of the modified SPARC snow retrieval from historical NOAA AVHRR data over the European Alps based on Landsat TM images

Master's Thesis

Faculty of Science
University of Bern

presented by
Simon Albrecht
2011

Supervisor:
Dr. Stefan Wunderle

Advisor:
Dipl. Geogr. Fabia Hüsler

Remote Sensing Research Group
Institute of Geography and Oeschger Centre for Climate Change Research

Acknowledgment

I would like to thank all the people who supported me and contributed to the realization of this project.
In particular I would like to express my gratitude to Dr. Stefan Wunderle, the supervisor and head of the Remote Sensing Research Group, for enabling this thesis, his excellent assistance and for giving me the time to reflect upon my future plans. He guided me through this year with helpful advice and his rich scientific experience.
In addition, I would like to thank my advisor, Fabia Hüsler. She introduced me to the field of IDL programming with a lot of patience and was never too busy to answer all my questions. I always enjoyed our teamwork.
Very special thanks also go to the rest of the RSGB, namely Céline, Dominic, Emanuele, Jan, Jeff, Rafael and Thomas who helped make this year not only interesting but also very amusing. Christoph Neuhaus and his efforts to keep my computer running at all times should not be forgotten.
Thanks are also owing to Sophie and Lara for their language assistance.
Last but not least my thanks go to my family who supported me throughout all the years and encouraged me in difficult times of my education.

Abstract

The Alpine Region is a highly sensitive ecosystem and will suffer from above-average changes in the course of climate change (IPCC 2007c). Snow cover plays a key role in this context as it influences the radiation and energy budget, the water balance, surface properties and also human activities. Mapping snow cover adequately and analyzing it over time is therefore of major importance. Since 1978, the NOAA AVHRR sensor system has been operational and is therefore best suitable for climate related long-term research. The Remote Sensing Research Group at the University of Bern (RSGB) has collected and archived a unique set of such data providing daily coverage of the European Alps. This data set serves as the basis for a 25-year snow time series, which is currently being processed at the RSGB.
With this background, a validation study is presented, where selected scenes of the AVHRR snow mask (1.1km spatial resolution) calculated with the Separation of Pixels using Aggregated Rating over Canada (SPARC) algorithm by Khlopenkov and Trishchenko (2007) are validated with high-resolution Landsat images (30m resolution). According to this, the objectives of the thesis are:

- Quantification of accuracy and bias of the AVHRR snow mask taking the different land cover types and elevation zones into account
- Modification of the SPARC score thresholds for algorithm improvement
- Investigation of the factors affecting the SPARC classification accuracy

For the calculation of the Landsat reference snow masks an adapted version of the NDSI/NDVI[1]-algorithm by Hall et al. (1998) is implemented. These snow masks, fitted to the AVHRR resolution, serve as the ground truth for the validation of the SPARC algorithm. In a first step, the accuracy (=ACC) and the BIAS are calculated with the default values of SPARC by comparing the class of each AVHRR pixel with the corresponding Landsat pixel (pixel-by-pixel approach). The results show a substantial underestimation of snow cover (overall BIAS=0.77) with an overall ACC of 78%, whereby most of the disagreement is found in the transition zone. While most of the scenes have ACC values between 75% and

[1]NDSI: algorithm for snow detection. NDVI: algorithm for detection of vegetation.

83%, the early morning satellites N12, N15 and December scenes show poorer results and are therefore excluded from the threshold decision process. In a second step, the B-score, responsible for the differentiation between snow and snowfree areas, is shifted stepwise from b-3 to b-11. Best ACC values are found with a value of b-6. With this modification an ACC gain of 9.46% for non-forested areas and 1.94% for forested areas results in an overall improvement of 5.21%. The underestimation could be corrected with a slight tendency to overestimation (overall BIAS=1.03). In addition, ACC is generally lower if forest cover is present. A comparison with the snow product from MODIS data (MOD10A1) shows that with this threshold adjustment the AVHRR snow mask almost, i.e. with -1 to -2%, reaches the quality of the state-of-the-art sensor system for snow cover mapping.

Several factors are found which influence the ACC of the snow masks. Negative correlations are found with the scene satellite zenith angle (r^2=0.71) and the sun zenith angle (r^2=0.21). The ACC is also influenced by the date of scene acquisition. While scenes from December show low ACC values, the performance increases throughout the winter (r^2=0.65). Furthermore, for forest pixels it seems that the lower the snow extent of the scene is, the higher the resulting ACC tends to be (r^2=0.20). The opposite is true for forest-free areas (positive correlation with r^2=0.64). Moreover, a dependency is found between the altitudinal distribution of the forest pixels and ACC. The latter is lowest between 1000 and 2000m above sea level, while most of the forest pixels are present in this zone.

Furthermore, a subpixel approach is presented where the resolution of the Landsat image is kept at 30m. With this, a fractional snow cover area can be calculated for each AVHRR pixel, which corresponds to 27 x 38 Landsat pixels. It can be shown that a similar behavior of ACC and BIAS compared to the pixel-by-pixel approach occurs, although the optimal values of the B-score differ slightly.

For the future long-term development of the algorithm it should be pointed out that in the course of climate change, uncertainties of the algorithm ACC could be shifted temporally. Such examples are classification weaknesses owing to shallow snow layers in the beginning of the winter, whereby these weaknesses are likely to be shifted to later months due to a shortening of the snow season. Another example is the expected future rise of the snow line leading to a dislocation of the transition zone, where correct classification is difficult.

The potential of the snow time series for climate research is enormous. It is therefore essential that the continuation is guaranteed, also with future satellites.

Contents

List of Figures

List of Tables

Abbreviations

ACC	Accuracy (as defined with equation 4.5)
asl	above sea level
AVHRR	Advanced Very High Resolution Radiometer
BH	Backward scattering Hemisphere
BRDF	Bidirectional Reflectance Distribution Function
DN	Digital Number
FH	Forward scattering Hemisphere
IDL	Interactive Data Language
IPCC	Intergovernmental Panel on Climate Change
IR	Infrared
Landsat (E)TM	Landsat (Enhanced) Thematic Mapper
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NDSI	Normalized Difference Snow Index
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
NOAA	National Oceanic and Atmospheric Administration
NPOC	Swiss National Point of Contact for Satellite Images
NSIDC	National Snow and Ice Data Center
OcCC	Organe consultatif sur les changements climatiques
RMSE	Root Mean Square Error
ROI	Region Of Interest
SCA	Snow Cover Area
SLF	Institute for Snow and Avalanche Research
SPARC	Separation of Pixels using Aggregated Rating over Canada
SWE	Snow Water Equivalent
SWIR	Short-wave Infrared
TIR	Thermal Infrared
TOA	Top of Atmosphere
UNEP	United Nations Environment Programme

USGS	United States Geological Survey
VIS	Visible
WSL	Swiss Federal Institute for Forest, Snow and Landscape Research

Chapter 1

Introduction

In its Fourth Assessment Report the Intergovernmental Panel on Climate Change (IPCC) states that the Alpine Region as a highly sensitive ecosystem will suffer from above-average impacts in the course of climate change (IPCC 2007c). This is mainly related to the high fraction of snow and ice and the sensitive feedbacks related to these land cover types. Snow is of crucial importance due to its influence on radiation and energy budget, water balance, surface properties and also on human activities. It acts as a water storage and affects water supply and runoff rate, which is essential for the energy industry. Furthermore, transportation, agriculture and winter sports rely on snow conditions (UNEP 2007). Mapping snow cover adequately and analyzing it over time is therefore of major importance for the entire Alpine Region and the science community. In order to study long-term developments of the Earth's snow extent, reliable snow time series are required. For this specific purpose, satellite remote sensing has received increasing attention as its history has reached a climate-relevant length.

1.1 The Relevance of Alpine Snow Cover Mapping

To emphasize the relevance of alpine snow cover mapping, an introduction to the influence of snow on the climate system and to climate change in the Alpine Region over the last decades is given.

Since 1970 the mean temperature has increased by approximately 1.5°C in Switzerland (UNEP 2007). This corresponds to a 1.5 times higher warming intensity compared to the whole of the Northern Hemisphere. The Alps in general are warming three times faster than the global average. An increase in winter temperature of roughly 1°C to 3°C from 1990 by 2050 is projected. Several studies substantiate this trend (Appenzeller et al. 2008, Begert et al. 2005, Scherrer et al. 2004). Since snow cover controls the Earth's albedo and is

responsible for the largest annual and interannual differences in surface reflectivity among all surface types (Armstrong and Brodzik 2001), these accelerated temperature changes in combination with the snow-albedo feedback lead to a continual decrease in snow cover extent and duration. Regions in the transition zone, where snow accumulates close to its melting point, are sensitive already to small temperature fluctuations. This is particularly the case in low elevation zones with shallow snow layers (IPCC 2007b). The mean altitude of this transition zone will also change since the snow line in the European Alps is expected to rise about 150m for every 1.0°C increase in winter temperature (Beniston 2003).

Snow accumulation and melt are directly coupled with air and soil surface temperature, precipitation, wind and surface relief. Whereas precipitation governs the overall amount of snow, air temperature determines not only the phase of precipitation, i.e. snow or rain, but also the rate of snow melt (UNEP 2007). Warmer winters with possible increases in precipitation will raise the snow line, reduce the overall snow cover and decrease the summer runoff (Marinucci et al. 1995). Related to this, in combination with the knowledge about the snow water equivalent (SWE: amount of water stored as snow) and the melting rate, snow cover is a crucial contributor for the assessment of water supply and flooding potential. This is also relevant for climate models as realistic snow cover data are essential for simulating surface energy balance as well as predicting winter water storage and runoff (UNEP 2007).

Seasonal snow cover is of additional relevance as it acts as a radiative sink and influences the transfer of heat and moisture between surface and atmosphere due to the low thermal conductivity (=insulator) (Gustafsson et al. 2001). Not least, snow cover also affects the large-scale atmospheric circulation and is a sensitive indicator of regional climate variability and changes (UNEP 2007).

Nevertheless, the advisory body on climate research to the Swiss federal government, OcCC (Organe consultatif sur les changements climatiques), emphasizes in its report that snow is guided by short-term weather developments and is, therefore, a component that is hard to assess. The uncertainties related to snow cover complicate the assessment of the development. Here, remote sensing techniques in combination with digital elevation models, numeric models and a systematic monitoring play an important role (OcCC 2008). The objective of the present thesis is to further reduce these uncertainties and to contribute to a better understanding of the snow cover developments over the last decades.

As already mentioned, snow cover and the ongoing climatic changes are also of socio-economic relevance. Mountain tourism and particularly winter mountain tourism depend on the snow conditions. In the period between 1975 and 1999 the mean snow extent was reduced by 50% above 440m and by 15% above 2200m in Switzerland, whereby serious uncertainties in snow reliability are expected (IPCC 2007a). An increase of the minimum winter temperature by 2°C, for example, would reduce the snow cover duration on Säntis (2500m) by 50 days when keeping the precipitation constant. Under assumption of a precipitation increase of 50% and same warming, a duration reduction of 30 days is expected. In the critical zones in Austria

(below 600m in winter and 1400m in spring), as a further example, every degree Celsius warming results in a shortening of the ski sport season by 4 weeks during the main season (IPCC 2007a). In the Swiss context OcCC (2008) mentions that until 2100, the classic winter sports are not expected to be cost-effective anymore in elevation zones lower than 1500m and advises that a diversification of the mountain tourism has to happen.

The high relevance of snow cover on the different levels described above and the ongoing changes of the last decades emphasize the need to study the snow cover developments by means of accurate time series. Remote sensing satellite systems are predestinated for snow mapping since they are able to capture information of the Earth's surface simultaneously on large scales and therefore have an advantage over in-situ observations. Several different sensors with distinct characteristics are available (Rees 2006). The choice of the sensor depends on the parameters to be analyzed, cloud occurrence as well as on the complexity of the terrain and is always a tradeoff between temporal, spatial and spectral resolution. A huge set of techniques have been used to detect snow parameters such as snow albedo, snowpack temperature, grain size and density, SWE and snow extent. At an intermediate spatial resolution of around 1km, such as NOAA AVHRR used for this project, the wider swath of these systems allows for a high temporal resolution, which is a prerequisite in snow cover analysis on a daily basis (Rees 2006).
Before the objectives of the thesis are introduced, the topic of remote sensing of snow is put into the scientific context.

1.2 The Scientific Context of the Thesis

A large number of snow retrievals from different polar orbiting as well as geostationary satellites have been developed to contribute to the knowledge of seasonal snow cover (Siljamo and Hyvarinen 2011, Parajka and Bloeschl 2006, Foppa et al. 2004). The immediate need for continuous historical daily snow cover products on a moderate resolution (1km) is identified as well by the Global Climate Observing System (GCOS) which advances such products for the purpose of climate change monitoring (GCOS 2006). Up today, the Advanced Very High Resolution Radiometer (AVHRR) is the only sensor which allows to study surface processes from space on a daily basis of more than 25 years. Although several other snow cover time series from AVHRR at different spatial scales have become available, an AVHRR snow time series of the European Alps on full resolution is still missing (Huesler et al. 2012).
The present thesis should serve as a supporting input towards such a time series in the framework of the study from Huesler et al. (2012) at the University of Bern.
A broad range of snow retrievals from AVHRR have been developed but none of these seems to be in widespread use. A common approach is the Normalized Difference Snow Index, NDSI

(Hall et al. 2002), which is based on the fact that snow is highly reflective in the visible part of the electromagnetic spectrum but absorbs most of the light in the near-infrared region (NIR) at 1.6μm (see also theory section 2.4). Hence, the NDSI is only applicable to sensors which detect information of these spectral regions. Since only the latest generation is able to provide this information and to switch between channel 3A (centered at 1.6μm) and 3B (centered at 3.9μm), the NDSI is not implementable for historical AVHRR data. In order to address this issue Khlopenkov and Trishchenko (2007) developed the Separation of Pixels using Aggregated Rating over Canada (SPARC) algorithm, specifically for historical AVHRR data. The advantage of this snow retrieval is that it is applicable to all sensor generations since it is able to deal with both channels 3A and 3B. In addition, its high robustness against calibration uncertainties due to the scaling approach makes it preferable for long-term studies (Huesler et al. 2012). This is the reason why the SPARC algorithm is used for the unique 25-year snow time series of alpine snow cover mapping. Khlopenkov and Trishchenko (2007) demonstrated its applicability for Canada. However, for the European Alps the classification thresholds have to be modified and the complexity of the terrain in challenging combination with the high amount of forest has to be considered. This is the framework where the present master's thesis is located. The default settings of SPARC are applied to an alpine study area and the accuracy of the resulting snow mask is analyzed. In a further step, the improvement potential is assessed and possible adaptations of the score thresholds are implemented. With this validation procedure valuable information for the improvement of the snow time series shall be delivered to achieve an accuracy as high as possible.

The accuracy has to be related to reference data, which is taken as the ground truth. However, the choice of this reference is not straightforward. In a previous study, conducted by the Remote Sensing Research Group, Foppa et al. (2004) developed a sub-pixel algorithm by means of linear spectral unmixing to optimize the predictability of snow cover in the heterogeneous terrain of the Alps. The algorithm, using channel 3A to calculate fractional snow cover areas, was later validated based on high-resolution ASTER data (Foppa et al. 2007). This concept of validating a product from moderate resolution sensors with reference images based on high-resolution data is applied to the thesis at hand. Here, 30m-Landsat data are used. Since 500m-resolution MODIS data (launched in 1999) have become state-of-the-art in snow cover mapping in recent years, this sensor is also a preferable validation source. This is emphasized by Parajka and Bloeschl (2006), who found an average accuracy of the MODIS snow product (MOD10A1) of 95% for the Alpine Region. For this reason, the AVHRR snow mask is also compared with this MODIS snow product to show how well the SPARC classification works.

A remaining issue in remote sensing of snow are classification difficulties emerging over forest areas since shadow effects and the canopy modify the reflectance. Based on MODIS data, Hall et al. (1998) presented an improved algorithm to address this challenging task by coupling the NDSI with the Normalized Difference Vegetation Index (NDVI). The Landsat image

serving as the reference for the validation of this thesis (see section 4.1.2.2) is performed in analogy to this approach.
Besides satellite based validation, the use of ground-based snow observations is another common validation approach. A large number of literature is available related to such methods (Parajka and Bloeschl 2006, Simic et al. 2004 Maurer et al. 2003, Romanov et al. 2002). However, the lacking representativeness of these stations for a wider area in the heterogeneous terrain of the Alpine Region is the main reason for not considering such validation data in the framework of this thesis.

1.3 Objectives of the Thesis

The developments described above emphasize the need of ongoing research in the Alps and the suitability of remote sensing techniques for these purposes. As a member of the consortium of the HyperSwiss-Net Project[1], the RSGB is responsible for the snow and climate modules. In this framework, Huesler et al. (2011) are currently processing a NOAA AVHRR time series of the European Alps for climate change studies. This data set serves as the basis for a 25-year snow time series calculated with a modified version of the SPARC algorithm. The series is validated by using webcam images, a snow cover model and MODIS data. With an additional validation based on high-resolution Landsat TM/ETM images, this master's thesis should serve as a supporting input for the assessment of the performance and the improvement potential of the SPARC snow retrieval. The time series could later be used for a variety of climatological applications, such as for long-term climate change analyses or for regional climate models as well as for hydrological runoff models. For any kind of application, a high accuracy of snow mapping is crucial. Therefore, a substantial part of the thesis is focused on quality studies related to land cover types (forest area in particular) and the transition zone of snow cover.
In this context the main objectives of the thesis are:

- Quantification of accuracy and bias of the NOAA AVHRR snow mask taking the different land cover types and elevation zones into account
- Modification of the SPARC score thresholds for algorithm improvement
- Investigation of the factors affecting the SPARC classification accuracy

The focus will be on scenes of the winter season, since it is most representative for snow cover and since usually snow only lies above the tree line, which inhibits an analysis of the SPARC classification performance in forest areas.

[1] `https://hyperswissnet.wiki.geo.uzh.ch/Project` Accessed: August 13, 2011.

A basic idea for validation was to pick one winter and study the development of the accuracy from the built-up of the snow layer via peak and ablation phase. Another was the investigation of intersensor consistency. With that, several different years would have been considered and analyzed. Due to a lack of systematic Landsat reference data for the whole period of the 25-year time series, these approaches were not realizable. Therefore as many winter Landsat scenes as could be found, were collected with the aim to cover all available sensors (see the result in Table 3.2).
The choices of the thresholds and the algorithm in general are focused on time series consistency rather than on single scene accuracy. This means that an algorithm generating consistent results over the whole investigation period (i.e. the time series) is given priority over a perfect snow cover representation for a single scene. It is obvious that the results of this approach cannot be compared directly with studies validating their product with one single reference scene. The criteria which lead to the final Landsat and corresponding AVHRR scenes is described in section 3.5.
The whole methodology of the thesis is implemented by means of the Interactive Data Language (IDL). Appendix C gives an overview of all relevant routines written in the framework of the realization of these objectives.

1.4 The NOAA AVHRR Sensor System

The Advanced Very High Resolution Radiometer (AVHRR) is carried on board the Polar Orbiting Environmental Satellites (POES) of the National Oceanic and Atmospheric Administration (NOAA). The following information about this system is mainly based on the official NOAA webpage[2]. It is the longest-lived series of Earth observing satellites and is therefore best suitable for climate related research. The unique length of operational AVHRR data in combination with daily global coverage from two satellites, calibrated thermal data and the low costs makes it interesting for a broad range of other applications such as oceanography, terrestrial sciences (such as NDVI, snow cover mapping, monitoring of forest fires) or meteorology. AVHRR, first launched with TIROS-N in 1978 as a four channel radiometer for meteorological applications (AVHRR/1), is sensitive to the visible, near-infrared and the thermal infrared of the electromagnetic spectrum. It was followed by AVHRR/2 carried on NOAA-7 (launched in 1981) with an additional fifth channel. Since 1998 AVHRR/3, first carried on NOAA-15, has been operating with six channels. Table 1.1 summarizes the channel specifications of the three AVHRR sensor generations. At approximately 833km above the Earth's surface the satellite orbits the globe 14 times each day with a swath width of 2400km. The satellites with even numbers cover the morning passes with equatorial crossing at local

[2] `www.ngdc.noaa.gov/` Accessed: June 10, 2011.

solar times of 07:30 and the odd-numbered cross the equator in the afternoon at 14:30 (NOAA Web[2]).

Table 1.1: Channel specifications of the three AVHRR sensor versions (NOAA[a]).

Channel	**AVHRR/1 (μm)**	**AVHRR/2 (μm)**	**AVHRR/3 (μm)**
1	0.58-0.68	0.58-0.68	0.50-0.68
2	0.725-1.10	0.725-1.10	0.725-1.0
3A	-	-	1.58-1.64
3B	3.55-3.93	3.55-3.93	3.55-3.93
4	10.50-11.50	10.30-11.30	10.30-11.30
5	Ch4 rep.	11.50-12.50	11.50-12.50

[a] `http://www2.ncdc.noaa.gov/`

Whereas channels 1 and 2 detect reflected solar radiation (VIS, NIR), channels 4 and 5 are sensitive to the infrared windows measuring the emitted terrestrial longwave radiation. For AVHRR/1 and AVHRR/2 channel 3, centered at 3.7μm measures in both spectra. For the latest AVHRR generation, channel 2 is narrowed to reduce water absorbtion and to enhance the effect of the reflectance of green vegetation. Furthermore, channel 3 is able to be switched from channel 3A (1,6μm) during daylight and 3B (3.7μm) during night for cloud identification and improved snow detection. However, due to system problems channel 3A is available for N17 only (Huesler et al. 2011).
Further details regarding the AVHRR instrument are covered by Cracknell (1997) and the NOAA User's Guides Kidwell (1998), Goodrum et al. (2006) and Robel (2009).

1.5 Thesis Outline

After an introduction to the theory of remote sensing of snow in chapter 2, the study area, the data and the methods used for the validation process are described in chapters 3 and 4. The results of the thesis are presented and directly discussed in chapter 5 and put in context of related studies. Finally, chapter 6 concludes the main findings and gives an outlook on future work.

Chapter 2

Remote Sensing of Snow

Remote Sensing is defined as *"the gathering of information about an objective without physical contact"* and is usually referred to as *"airborne or spaceborne observations using electromagnetic radiation"* (Rees 2006).

2.1 A Brief History of Remote Sensing

The history of remote sensing is also the history of photography. After the first photographs taken by Daguerre and Niepce in 1839, photography was used for topographic purposes in the following year. In 1858, balloons served as a tool to take pictures of large areas. With the beginning of aviation, acquisition of data over selected areas and under controlled conditions became possible and in 1909 the first photographs from airplanes were taken. After World War II and the mid-1960s, NASA sponsored a large number of studies for applications of color infrared and multispectral photography. This led to the launch of the Landsat satellites in the 1970s (see section 3.3) although systematic Earth observations from space already began in 1960 with the launch of TIROS-1, the first meteorological satellite (Elachi and Van Zyl 2006).

2.2 Remote Sensing in General

One distinguishes between passive remote sensing, where natural radiation is detected (e.g. the AVHRR system belongs to this type) or active remote sensing, where the radiation is generated by the system itself. The natural radiation is composed of reflected solar radiation in the visible and near-infrared parts of the electromagnetic spectrum at roughly 0.35-2.5μm and thermally emitted radiation (Figure 2.1 top). The range of the latter depends on the temperature of the emitting body. The dominant wavelength for a body with a temperature of 273K (0°C) lies at around 11μm. This type of radiation can be detected in the thermal infrared part, typically between 8μm and 14μm as well as in the microwave region at 1cm to

1m. Due to the chemical composition of the atmosphere, the spectrum between 14μm and 1cm is mainly blocked for radiation (see Figure 2.1 bottom) and cannot be used for remote sensing (Lillesand et al. 2004).

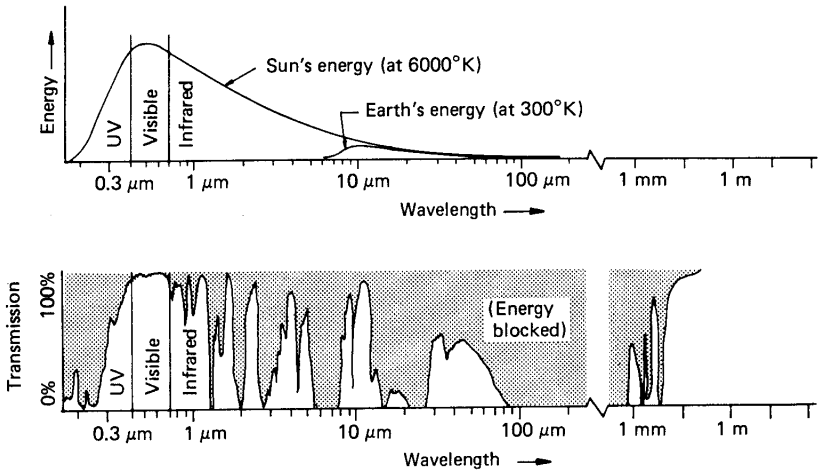


Figure 2.1: The electromagnetic spectrum of light. The energy of light depending on the wavelength (top) and the atmospheric transmittance at the corresponding wavelengths (bottom). Grey shaded area represents the wavelength regions at which radiation is blocked by the atmosphere (Lillesand et al. 2004).

Passive remote sensing systems measure *the amount of radiation reaching the sensor in a particular waveband* (Rees 2006), i.e. the radiance. In combination with the information about the incident radiation, the so-called reflectance of the surface can be calculated (see chapter 4.1). After correcting the atmospheric influence, the value of this reflectance and its spatial and temporal variations can be interpreted.
Remote sensing systems that detect thermal radiation measure the radiance as a brightness temperature, which is *the temperature of a perfect emitter (a so-called black body) that would produce the same amount of radiation* (Rees 2006). The brightness temperature is dependent on the physical temperature of the Earth's surface and its emissivity (dependent on the material). Knowing these two quantities allows for the calculation of the actual physical temperature (Rees 2006).
There are also active remote sensing systems, whereby two main types are defined. Ranging instruments are mainly used for topography information. The distance from the sensor to the Earth's surface is detected by measuring the travel time of radiation down to the surface and back again. It is used for example for laser profiler or radar altimeters. The second type measures surface reflectances, similar to passive systems. However, instead of using solar irradiance it illuminates the surface and receives the modified signal from the surface. The big advantage of this system, mainly used for imaging radars in the microwave region of the electromagnetic spectrum, is that the characteristics of the incident radiation, like the wavelength of polarization, can be controlled by the system itself (Rees 2006).

2.3 Operational Remote Sensing of Snow

First remote sensing snow observations date back to the first image of the TIROS-1 satellite in 1960. Since 1966 optical snow cover monitoring and since 1987 also passive microwave imagery have been developed. The detection range of optical systems (visible and infrared) allows for much higher spatial resolutions due to the higher energy of light in this spectrum compared to the microwave region. However, systems in this range also have their limits since they are not applicable during nighttime or cloud cover. Especially the latter is a limitation in terms of snow cover mapping in the Alpine Region during the winter season (Rees 2006). In addition, high repeating rates are required to cover the Earth with only small time gaps since snow underlies a high degree of variability on a daily basis. The complex topography and the high degree of forest cover are additional common problems in remote sensing of snow (Hall et al. 1998).
A broad range of operational snow products from remote sensing is available and accessible via the Web. The National Snow and Ice Data Center (NSIDC)[1], for example, archives and distributes operational snow and ice data products from different satellite platforms. In 1997, the Interactive Multisensor Snow and Ice Mapping System (IMS) became operational, which provides 24km and 4km snow and ice maps of the Northern Hemisphere. In addition, the Rutgers University Global Snow Lab (GSL)[2] offers climate data, as well as daily, weekly and monthly global snow products. Via PolarView[3], supported by the European Space Agency (ESA) and the European Commission with participation from the Canadian Space Agency, daily snow maps of fractional snow cover for Scandinavia, the Baltic Sea Area, and snow cover area maps of Central Europe and the Alps are available. Additional data can be expected from the GlobSnow project[4], which is aimed to implement an operational near-real time (NRT) snow information service that will provide daily snow maps for hydrological, meteorological, and climate research purposes in addition to the already existing snow extent and snow water equivalent products. The satellite based information is thereby combined with ground-based weather stations. Not least, based on daily NOAA AVHRR data, the Remote Sensing Research Group of the University of Bern facilitates operational snow maps, which are regularly delivered to the Swiss Federal Institute for Snow and Avalanches (SLF) for avalanche prediction.

All these products have to deal in some way or other with the special characteristics of a snow-covered surface. The theoretical background of this surface type is discussed in the following section.

[1] http://nsidc.org/
[2] http://climate.rutgers.edu/snowcover/
[3] http://www.polarview.org/
[4] http://www.globsnow.info/

2.4 Physical Properties of Snow

Fresh dry snow is highly reflective over the whole visible part of the electromagnetic spectrum (approximately 0.4-0.65μm) and therefore appears white to the human eye. The dielectric properties of ice and the fact that snow contains up to 10^9 particles per cubic meter are responsible for that behavior. The high absorbtion length[5] of visible-wavelength radiation in ice in the order of 10m prevents radiation from penetrating through thick snow packs. However, since the sun light is able to pass the upper most centimeters (~10cm) of the snow layer, mixed reflectance information is generated over shallow snow packs, which may lead to misclassification. Due to the large amount of air-ice interactions of a photon when traveling through a thicker snow pack, it is almost certain that it is scattered back. This also implies that the reflection coefficient changes with changing grain size. Whereas differences are small in the visible part of the spectrum, with larger grain diameters the reflectance decreases (and the absorbtion increases) due to the reduced air-ice interactions and thus reduced scattering. In the near-infrared region the reflectance declines strongly, reaches a minimum at around 1.5μm and stays at a low level with longer wavelengths in the thermal infrared region (see Figure 2.2). Multispectral remote sensing analysis takes advantage of this different reflectance

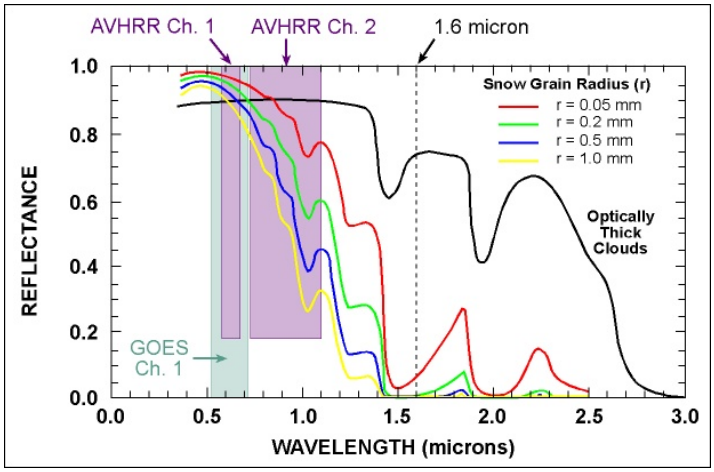


Figure 2.2: Reflectance values for snow with different grain radii in the visible and infrared region of the electromagnetic spectrum of light (colored curves), reflectance curve for clouds in black and corresponding coverage of the AVHRR channels 1, 2 and 3A (at 1.6μm). Source: NSIDC.[a]

[a] `http://nsidc.org/data/docs/daac/nsidc0066_avhrr_5km.gd.html` Accessed July 16, 2011.

behavior in the different regions of the electromagnetic spectrum. Whereas channels 1 and 2 of the AVHRR sensor are used to separate snow from other land cover types, channel 3A is predestinated for the discrimination of clouds and snow. Channels 3B, 4 and 5 are mainly

[5] The absorption length is defined as *"the distance through which the radiation must travel in order for its intensity to be reduced by a factor of* e *as a result of absorption alone."* (Rees 2006)

designed for (sea) surface temperature analysis.[6] The reflectance of a snow pack is subject to changes as it ages. Dust or soot are deposited and decrease the reflectance from fresh dry snow with albedo values higher than 90% to 40% or even 20% for dirty snow. The melting process causes the liquid water content to increase. Although the presence of liquid water has little direct effect on the snow reflectance, there is an indirect effect. Thus, it induces clustering of the ice crystals, which leads to larger grain sizes and hence to a lower reflectance (Rees 2006).
An additional characteristic of a snow pack is that it is anisotropic forward scattering (Luo et al. 2008) and therefore a larger sun zenith (or illumination) angle lead to larger reflectance values. The effect of this feature on the snow classification accuracy is analyzed in section 5.2.2 (Rees 2006).

Based on this theoretical background, the remote sensing snow classification approaches are designed.

2.5 Snow Classification Approaches

2.5.1 NDSI/NDVI approach

In the visible part of the spectrum the reflectances of snow and clouds are nearly equally high with a reflectance value of roughly 0.9, at least for wavelengths between 0.4-1.0μm. In the NIR, however, they diverge up to a maximum located at around 1.6μm. Whereas the reflectance of snow declines substantially, the cloud reflectance remains on a high level (see again Figure 2.2). This spectral behavior is widely used by snow classification schemes such as the normalized difference snow index (NDSI) (Hall et al. 1998). This snow index can be applied to all sensor systems that capture the visible as well as the near-infrared part of the electromagnetic spectrum separately in two channels. Using these two bands allows additionally for topographic corrections considering that variations in the sun-ground surface-sensor geometry are affected by the same error factor for all channels. By using a ratio of two spectral bands this error is eliminated. Analogous to the normalized difference vegetation index (NDVI, see below), the NDSI is calculated with resulting values between 0 and 1:

$$NDSI = \frac{VIS - NIR}{VIS + NIR} \tag{2.1}$$

whereby an NDSI of 0.4 is generally taken to classify snow (Hall et al. 1998). The NIR corresponds to SWIR for Landsat (see 3.1 in the data chapter 3).
Several snow classification algorithms, such as the Landsat or MODIS snow algorithm (Hall et al. 1998, 2002), also implement an NDVI to reduce the effect of vegetation, mainly forest,

[6] `http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html` Accessed July 17, 2011.

since certain forest types generate NDSI values below 0.4 (see section 4.1.2.2). It makes use of the characteristics that vegetation has its absorbtion maximum in the red portion of the electromagnetic spectrum (0.6-0.7μm) and its reflectance maximum in the near infrared region of the electromagnetic spectrum (0.75-1.5μm):

$$NDVI = \frac{NIR - red}{NIR + red} \tag{2.2}$$

2.5.2 SPARC algorithm

In contrast to this NDSI/NDVI-threshold approach, the SPARC snow retrieval for AVHRR data (Khlopenkov and Trishchenko 2007) consists of an aggregated rating of several tests. The results of these tests are accumulated and the calculated final rating classifies the pixel as clear-sky, partly cloudy, cloud-shadowed, or as snow- or ice-covered. Thereby, three major scores are calculated: the B-score derived from the reflectance brightness test in channel 1, the R-score derived from the reflectance test in channel 3 (A or B) and the T-score derived from the brightness temperature test in channel 4. This approach is a substantial advantage compared to other methods which use, for example, yes/no-decisions. In that case, the risk of misclassification is increased since once a pixel passes a test with 'no' it is not considered anymore.

The algorithm was basically introduced as a cloud detection scheme in the framework of the new AVHRR data processing system at the Canada Centre for Remote Sensing (CCRS) to study long-term landmass changes over Canada at high spatial and temporal resolution regarding climate change issues. Since clouds and snow have similar reflectance properties in certain channels (especially channel 1 - VIS), every cloud algorithm also contains specific snow classification tests which can be used for snow mask calculations. In analogy to the NDSI, it is also able to deal with both AVHRR channels 3A and 3B (Khlopenkov and Trishchenko 2007).

The detailed implementation of the SPARC algorithm is explained in the methods section 4.

Chapter 3

Data

3.1 Study Area

The selection of the study area was undertaken according to several criteria. Since research and data archiving of the Remote Sensing Research Group Bern are focused on the Alpine Region, a subset of this area was chosen. In addition, it should contain valley structures as well as high plateaus ranging from the low lands up to the high mountains whereby the structures are preferably clearly identifiable. Moreover, the predefined paths and rows of the Landsat 5/7 tracks limited the options since the study area has to be entirely part of one Landsat scene as overlaps would complicate the validation process. Figure 3.1 illustrates that Switzerland, for example, is covered only by three orbit paths (196 to 194, W to E) and rows (27 to 29, N to S). Therefore, the Landsat center image of the Swiss Alps (195/28) is chosen not least because it contains the Rhone Valley, which is clearly visible also on a coarse resolution.

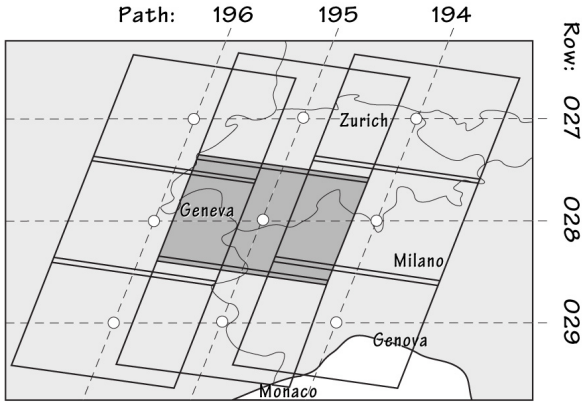


Figure 3.1: Paths and rows of Landsat orbits covering the Alpine Region. 8 of 9 ordered Landsat images cover Path 194/Row 28 shown in grey (EURIMAGE 2011).

The Landsat image spans an area of ~35'000km^2 (190 x 183km) between the southern part

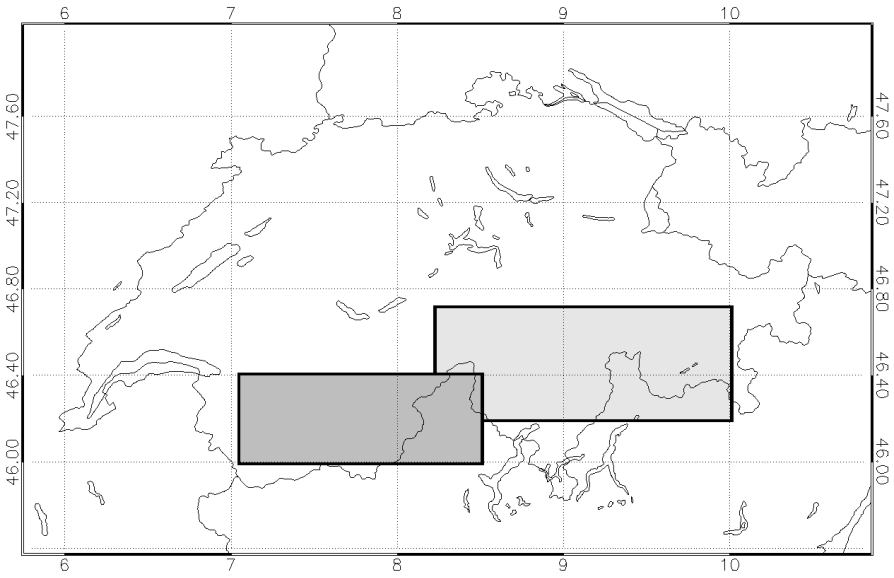


Figure 3.2: Overview of the study area represented by the dark-grey box. The light-grey box represents the study area only for scene L7 2001/12/14.

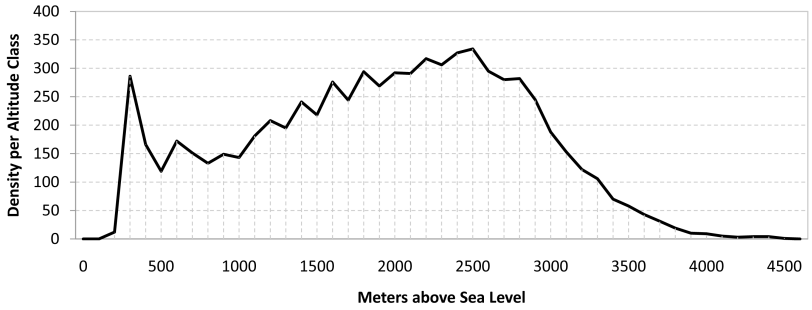


Figure 3.3: Altitude distribution of the study area.

of Lake Neuchatel in the Swiss Canton Vaud in the upper left corner (47.02N, 5.78E), the Valais Alps in its centre and the Italian Piemont (45.04 N, 9.08 E) in the lower right corner. Figure 3.5 visualizes an example of such a scene. The further specification of the subset was determined by cloud contamination. The lowest common denominator, i.e. the area where all Landsat scenes were cloud free, was chosen. Figure 3.2 shows the final selection of the study area. It covers a rectangle of around 100km in x- and 40km in y-direction. It mainly lies on Swiss terrain and consists of the surroundings of the Rhone Valley from Martigny to the river source in the central Alps. Expressed in coordinates, this is referred to an area from 46.40N / 7.06E in the upper left corner to 46.00N / 8.50E in the lower right corner. The graph of Figure 3.3 depicts the altitude distribution of the study subset emphasizing that, apart from the lowest terrains below 300m above sea level (=asl), all elevation zones up to roughly 4000m asl are covered. It ranges from 190m to 4428m asl with a mean elevation of 1883m asl.

There is one exception which has to be considered. Scene L7 2001/12/14 covers Path 194 / Row 28 instead of 195/28. Thus, it is shifted one path to the east. Since hardly any scenes were available for the first months of winter in the Alps, namely November and December, it was decided to analyze also this scene. The subset is somewhat bigger with 130km x 50km.

3.2 NOAA AVHRR Data Archive

The NOAA AVHRR sensor system providing daily global coverage with six available channels (for the latest generation) and a horizontal resolution of 1.1km was already introduced in section 1.4. In addition, the data is pre-processed according to the NOAA User's Guides Kidwell (1998), Goodrum et al. (2006) and Robel (2009), including calibration, georeferencing based on a feature-matching algorithm, orthorectification and a cloud masking.
Since 1981 the Remote Sensing Research Group at the University of Bern receives and archives daily AVHRR data. In cooperation with the Freie Universität Berlin, the Comprehensive Large Array data Stewardship System and the German Aerospace Center, a time series of AVHRR data providing full coverage of the Alpine Region has been collected since 1984. Figure 3.4

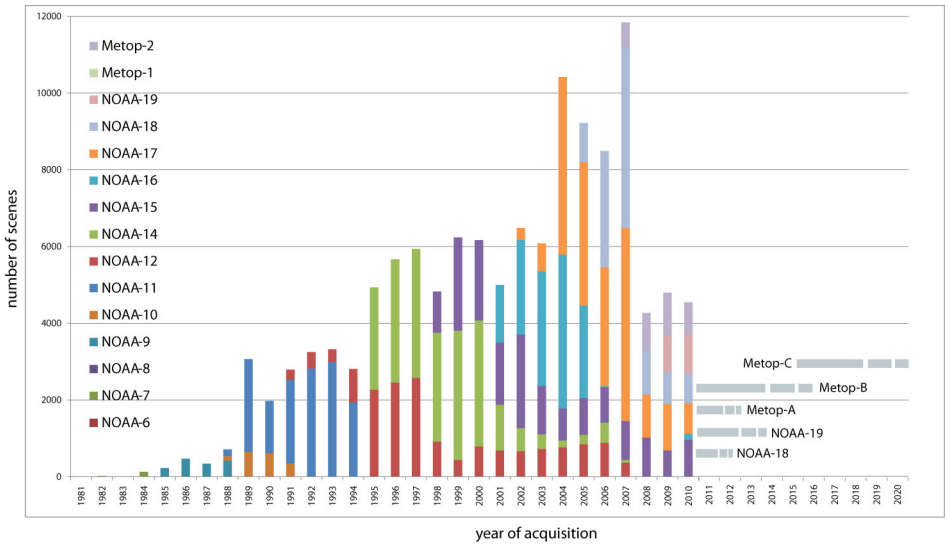


Figure 3.4: The AVHRR data archive. Planned missions are represented by grey horizontal bars (Huesler et al. 2011).

gives an overview of the data collected at the University of Bern, which consists of 124'000 files. Due to the poor data availability in the initial years, a spatially and temporally complete coverage is only available after 1985. The number of files per year increases and peaks in 2007 when six AVHRR platforms were in orbit. The overlapping operation time periods of the different platforms allow for intersatellite comparison in order to achieve a homogeneous time series with high quality. In order to maintain the AVHRR system and to improve climate change monitoring, new satellites have been launched and new missions are planned (Huesler

et al. 2011).
On the basis of this data set, the AVHRR snow masks for the validation are calculated with the SPARC algorithm (see section 4.2.1 for details).

3.3 Landsat

In the framework of the Apollo mission, photographs of the Earth's land surface were taken from space for the first time and served as the initial point for Landsat. At that time, no satellites were available that monitored the Earth's surface, although weather satellites had been observing the atmosphere since 1960. Despite opposition regarding secrecy and geopolitical concerns of "watching" Earth, Landsat 1 was first launched in 1972 by NASA (see more about Landsat history at `http://landsat.gsfc.nasa.gov/about/history.html`). Hence, Landsat is the longest-running commercial earth-observation satellite program. Since 1972 additional five satellites have been operational (EURIMAGE 2011). Due to the length of the available data sets, the Landsat satellites are best applicable for climate related applications and therefore also for the validation of the SPARC algorithm for the 25-year AVHRR time series of snow cover. However, the repeat time of only 16 days is a limitation, especially during winter when extensive cloud cover occurs. Here, scenes from Landsat 5 (launched in March 1985) and Landsat 7 (April 1999) carrying the Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+), respectively, are used. Both provide data of seven multi-spectral bands with 30m horizontal resolution (see two examples in Figure 3.5). Landsat 7 has an

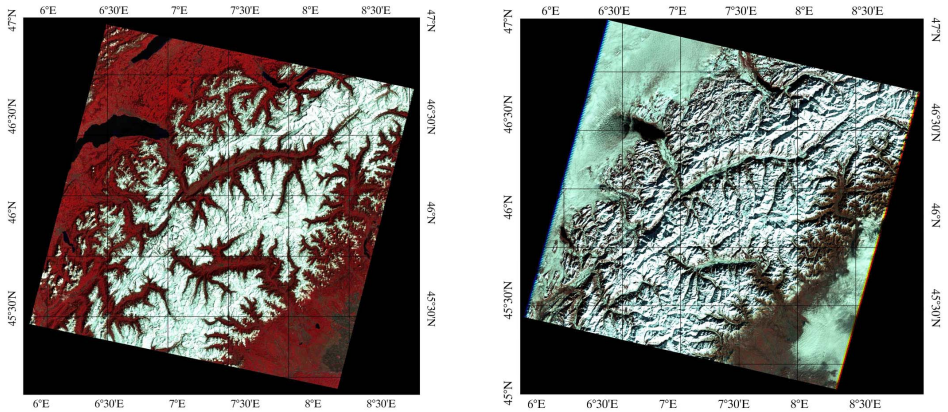


Figure 3.5: IR-images (i.e. channels 4/3/2 for R/G/B) of Landsat 5 1995/05/03 from NPOC/EURIMAGE (left) and Landsat 7 2003/02/10 from USGS (right). The different snow conditions and vegetation stages between winter and spring are clearly visible.

additional panchromatic channel with 15m resolution (see Table 3.1). The sun-synchronous and near-polar orbiting satellites cover the Earth from a flight altitude of 705km with a swath width of 183km and pass the equator between 9:30 and 10:00 local time (descending passes). Sun elevation and azimuth are depending on the day of year and vary between 18° and 53°

and 138° and 160°, respectively. These orbit parameters result in a repeat cycle of 16 days (Chander et al. 2009). Consequently, this coarse temporal resolution limits the availability of scenes, especially during winter when frequent cloud cover occurs and hence, reduces the number of possible AVHRR scenes to be analyzed.

Table 3.1: Bands and its wavelengths for Landsat 5 TM and Landsat 7 ETM+. All bands with 30m resolution. 15m-panchromatic channel 8 is only available for Landsat 7 (EURIMAGE 2011).

Band	**Region**	**Wavelength**
Band 1	Blue	0.45-0.52
Band 2	Green	0.52-0.60
Band 3	Red	0.63-0.69
Band 4	Near Infrared	0.76-0.90
Band 5	Short-wave Infrared	1.55-1.75
Band 6	Thermal Infrared	10.42-12.50
Band 7	Short-wave Infrared	2.08-2.35
Band 8	Panchromatic (Landsat 7 only)	.52-.90

The data mainly originates from the web portal of the U.S. Geological Survey (USGS[1]) (free of charge). Additionally, two commercial scenes were ordered from the Swiss National Point of Contact for Satellite Images (NPOC) and from EURIMAGE via the Federal Office of Topography (swisstopo). All images are radiometrically calibrated, georeferenced and terrain corrected (orthorectified). As for AVHRR, atmospheric correction is not included in the Landsat processing.

3.4 MODIS

The Moderate Resolution Imaging Spectroradiometer, MODIS, was first launched on board the Terra satellite in December 1999 by NASA, followed by a second instrument carried on Aqua in May 2002[2]. Terra crosses the equator in the morning from North to South (10:30 am local time) and Aqua in opposite direction in the afternoon (1:30 pm local time). Therewith, MODIS is able to cover the entire Earth's surface every 1-2 days with a spatial resolution of 500m detecting reflectances in 36 spectral channels. These features are a substantial improvement over those of AVHRR, especially because of the higher spatial and radiometric resolution. This enables a more effective cloud discrimination and the detection of snow under vegetation canopies. The latter is mainly related to the implementation of a combined NDSI/NDVI snow algorithm. Furthermore, the instrument is used in a broad field of observations in the atmo-

[1] `http://landsat.usgs.gov/`

[2] `http://aqua.nasa.gov/` and `http://terra.nasa.gov/` Accessed: May 11, 2011.

sphere, the ocean and on the land-surface. Included measures are for example cloud cover, sea-surface temperature and chlorophyll, land cover changes, land-surface temperature and vegetation properties. Therefore MODIS is expected to contribute to a better understanding of the Earth's system and its spheres. A large set of studies, which used MODIS for validation purposes, is available. Moreover, MODIS is of major importance for the development of validated Earth system models in terms of climate change.
The MODIS/Terra daily snow cover product used here, MOD10A1 (Hall et al. 2002), is available from February 2000 to present, freely accessed through the Distributed Active Archive Center of the National Snow and Ice Data Center, NSIDC[3]. Therefore, validation with this snow mask is only possible for scenes acquired after this date. The snow mask is reprojected and resampled to the 1.1km-resolution of AVHRR to allow for a pixel-by-pixel comparison. In analogy to the basic concept of Landsat snow classification (see section 4.1) the underlying algorithm employs an NDSI and an NDVI including certain additional tests. Thus, both the MODIS and the Landsat snow mask are expected to be similar.

3.5 Available Scenes

As already mentioned above, several criteria reduced the number of usable scenes for the study. Basically, these are:

- substantial lack of available Landsat winter scenes between 1990-1999 from USGS
- early morning overpasses cause heavy shadows in winter (e.g. N15)
- cloud cover over large parts of the images
- time gap between Landsat and AVHRR: In some cases perfect Landsat scenes were useless due to partly/completely cloud covered AVHRR images taken a few hours later
- corrupt AVHRR scenes and
- the price for commercial scenes.

Finally, Table 3.2 lists the scenes for the analysis and the corresponding geometries, whereby NOAA 11, 12, 14, 15, 16 and 17 are covered. In total, 9 Landsat images and 17 corresponding AVHRR scenes were found. Efforts were made to find feasible scenes for NOAA09 as well. Unfortunately, the only usable Landsat image corresponding to that time period could not be used due to a completely noisy AVHRR image (1987/01/21). The table also contains information about the snow extent of the Landsat image, referred to the previously defined study area. It lies between 42% for L7 2001/12/21 and 78% for L7 2003/02/10. For details

[3] www.nsidc.org

about the calculation of this snow information see section 4.1.2.2.
In a first selection only satellite zenith angles of less than 40° were considered since it is expected that the misclassification will increase with increasing viewing angle. With too flat angles, the information of one pixel might correspond to three or more pixels. The accuracy comparison between angles greater or smaller than 40° in section 5.2.1 shows that this assumption could be confirmed.

Table 3.2: Overview of analyzed Landsat and corresponding AVHRR scenes. If not specified by footnote, Landsat scenes are ordered from USGS free of charge. Snow extent [%] is calculated based on the Landsat snow mask and is referred to the study area only. Angle information [°] corresponds to the AVHRR center pixel of the study area. For L7 2001/12/14 and later, MODIS data is available for comparison.

			Satellite		Sun	
Landsat	**Snow extent [%]**	**NOAA AVHRR**	**Zenith**	**Azimuth**	**Zenith**	**Azimuth**
L5 1993/03/10[a]	65.02	N11, 14:25 UTC	10.7	75	62	228
L5 1995/05/03[a]	52.38	N12, 08:06 UTC	53.1	-69	52	109
		N14, 12:34 UTC	0.3	77	34	210
L7 2000/03/21	50.73	N14, 14:51 UTC	3.4	-103	62	237
		N15, 07:38 UTC	4.7	-77	69	113
L7 2001/12/14	59.72	N16, 12:38 UTC	3.3	77	72	199
L7 2001/12/21	41.90	N16, 1304 UTC	22.7	-100	73	202
L7 2002/01/06	51.52	N16, 11:52 UTC	55.8	67	69	184
		N16, 13:34 UTC	51.8	-95	74	208
L7 2003/01/25	71.37	N16, 11:57 UTC	56.1	67	65	183
		N16, 13:38 UTC	51.4	-95	70	209
		N17, 09:48 UTC	37.6	98	70	151
L7 2003/02/10	77.93	N16, 12:19 UTC	40.2	71	61	189
		N17, 10:28 UTC	16.0	-75	63	159
L7 2003/02/26	69.72	N16, 12:42 UTC	15.9	75	57	197
		N17, 09:28 UTC	54.7	94	63	142
		N17, 11:08 UTC	54.3	-68	56	170

[a]Source: NPOC

3.6 The Forest Mask

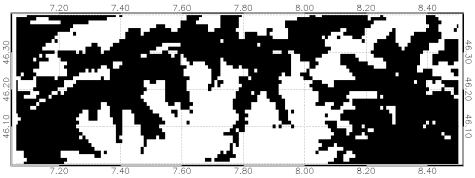


Figure 3.6: Forest mask of the study area. Forest area in black.

In order to study the algorithm accuracy in forested areas a mask was calculated from the Global Land Cover 2000 classification map (GLC2000, Hartley et al. 2006) containing all the forest classes present in the study area. These were Nr. 2: Tree cover, broadleaved, deciduous, closed; Nr. 4: Tree cover, needle-leaved, evergreen and Nr. 6: Tree cover, mixed leaf type. Figure 3.6 illustrates the resulting mask and emphasizes the high fraction of 55% of forest cover in the study area.

Chapter 4

Methods

4.1 Pre-processing steps for Landsat Data

4.1.1 Conversion from Digital Numbers (DN) to Top-Of-Atmosphere (TOA) reflectances

First, the Landsat images have to be prepared to generate the reference snow mask. The ordered data was delivered in the form of calibrated digital numbers (DN) whereby the values are scaled from the lowest reflectance value in the image (DN=1) to the highest reflectance value (DN=255). To make these values comparable to other scenes and to obtain a physical meaning, the data has to be converted to top-of-atmosphere (TOA) reflectances. These reflectance values are a prerequisite to apply the algorithms responsible for the calculation of the snow mask. The whole process with the required constants for all Landsat sensors and an overview of the radiometric calibration procedure are described in Chander et al. (2009). The following data conversion is based on this reference. For data ordered from the USGS, the required parameters for each individual scene are included in the metadata file (_MTL.txt). For scenes ordered from NPOC, the information is taken from the calibration parameter files (.CPF) provided by the USGS[1].
The ETM+ sensor of Landsat 7 operates either in low- or high-gain state. This allows for a maximization of the sensors' 8-bit radiometric resolution without producing oversaturation. Therefore, bright surfaces, which have a higher dynamic range but low sensitivity, are detected in low-gain mode, and dark surfaces, with lower dynamic range but high sensitivity, are detected in high-gain mode. This fact has to be considered when converting the data to TOA reflectances. The conversion procedure includes a conversion to at-sensor spectral radiance and, in a second step, to top-of-atmosphere (TOA) reflectances.

[1] `http://landsat.usgs.gov/science_L5_cpf.php` Accessed: March 23, 2011

4.1.1.1 Conversion to at-sensor spectral radiance

Calibration of the Landsat sensors results in calibrated digital numbers (DN's). A crucial step to convert the data into a physical meaningful radiometric scale is the conversion to at-sensor spectral radiance. This conversion requires knowledge about the lower and upper limit of the original rescaling factors (either 0 or 1 and 254 or 255, respectively). The procedure can be described with the following equation:

$$L_\lambda = \left(\frac{LMAX_\lambda - LMIN_\lambda}{Q_{calmax} - Q_{calmin}}\right)(Q_{cal} - Q_{calmin}) + LMIN_\lambda \tag{4.1}$$

where

L_λ = Spectral radiance at the sensor's aperture [W/(m^2 sr μm)]

Q_{cal} = Quantized calibrated pixel value [DN]

Q_{calmin}/Q_{calmax} = Minimum/Maximum quantized calibrated pixel value corresponding to $LMIX_\lambda/LMAX_\lambda$ [DN]

$LMIN_\lambda/LMAX_\lambda$ = Spectral at-sensor radiance scaled to $Q_{calmin/max}$ [W/(m^2 sr μm)].

4.1.1.2 Conversion to TOA reflectances

There are three advantages when using TOA reflectances. Firstly, it removes scene-to-scene differences caused by different solar zenith angles. Secondly, spectral band differences arising from different values of the exoatmospheric solar irradiance vanish. Thirdly, the variation of the Earth-Sun distance, which takes place on a daily basis, is corrected. These corrections are achieved by calculating:

$$\rho_\lambda = \frac{\pi \cdot L_\lambda \cdot d^2}{ESUN_\lambda \cdot \cos\theta_s} \tag{4.2}$$

where

ρ_λ = Planetary TOA reflectance [unitless]

π = Mathematical constant equal to 3.14159 [unitless]

L_λ = Spectral radiance at the sensor's aperture [W/(m^2 sr μm)] (see eq. 4.1)

d = Earth-Sun distance [astronomical units]

$ESUN_\lambda$ = Mean exoatmospheric solar irradiance [W/(m$^2\mu$m)]

θ_S = Solar zenith angle [degrees] = sine of sun elevation angle.

After this conversion, the Landsat images are ready for the calculation of the snow mask.

4.1.2 Landsat snow classification

4.1.2.1 Supervised classification

Due to the high resolution of the Landsat images it is possible to carry out two approaches to obtain a reference snow mask: a) the supervised classification and b) the NDSI/NDVI-threshold approach. As a first method, a supervised classification is performed using the

remote sensing software ENVI (ITT Industries). Manually, three so-called endmembers are defined (visually via computer screen) with several selected areas corresponding to: snow, snowfree and cloudy area. Depending on the properties of each of these areas, ENVI classifies every pixel as one of the groups. Despite the good results of this approach (see Figure 4.1), it is not further considered. This is due to great disadvantage of the procedure as it has to be repeated manually for every single Landsat scene and cannot be implemented into an automatic mode. A much faster method is to use the NDSI/NDVI threshold approach according to Hall et al. (1998). Here, however, without an additional cloud mask, cloud discrimination is not possible and investigations are limited to clear-sky scenes.
To illustrate the different outcomes of the two classification approaches, the differences are discussed in section 4.1.2.3 after introducing also the NDSI/NDVI-threshold approach.

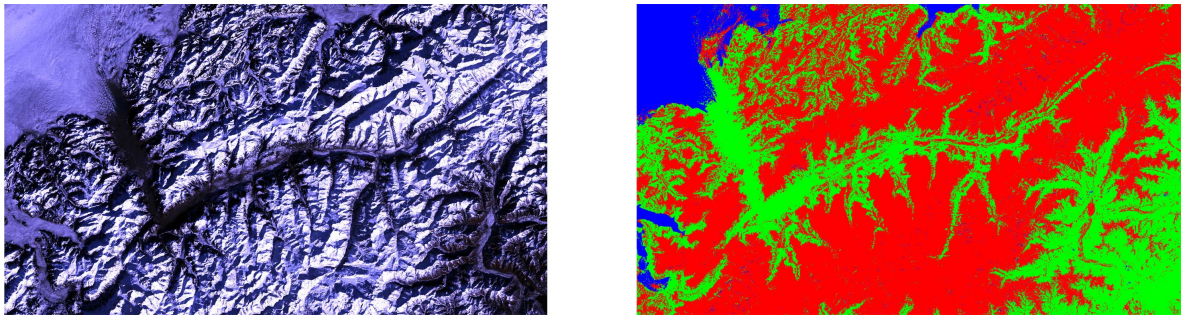


Figure 4.1: RGB image (channels 3/2/1) of Landsat 5 1987/02/06 (left) and corresponding result of the supervised snow classification (right) above the Rhone Valley. Snow in red, snowfree area in green, clouds in blue.

4.1.2.2 NDSI/NDVI threshold classification approach

Due to disadvantages of the supervised snow classification, the NDSI approach according to Hall et al. (1998) is applied. Based on the Landsat channel files converted to TOA reflectances, a binary snow mask is calculated by means of a multispectral imagery approach. It takes advantage of the spectral properties of snow and clouds in the visible (VIS) and the near infrared region (NIR) of the electromagnetic spectrum (see section 2.5.1). For Landsat TM and ETM+ the NDSI is calculated as follows:

$$NDSI = \frac{r_2 - r_5}{r_2 + r_5} \tag{4.3}$$

where r_2 defines the reflectance in band 2 (centered at 0.57μm) and r_5 is band 5 reflectance (centered at 1.65μm).

In addition to an NDSI $\geq$0.4, the band 4 reflectance has to be higher than 0.11 to distinguish between water and snow. Band 2 is set to be higher than 0.1 to exclude very low reflectances caused by certain forest types. Moreover, the algorithm considers the fact that most forests

generate NDSI values lower than 0.4. Therefore, the classification is combined with an additional threshold for the NDVI (>0.1) represented by a simplified triangle shown in Figure 4.2. In analogy to the NDSI, the NDVI is calculated for Landsat with bands 3 and 4.

Before the implementation of this algorithm, it is tested whether a simple approach with

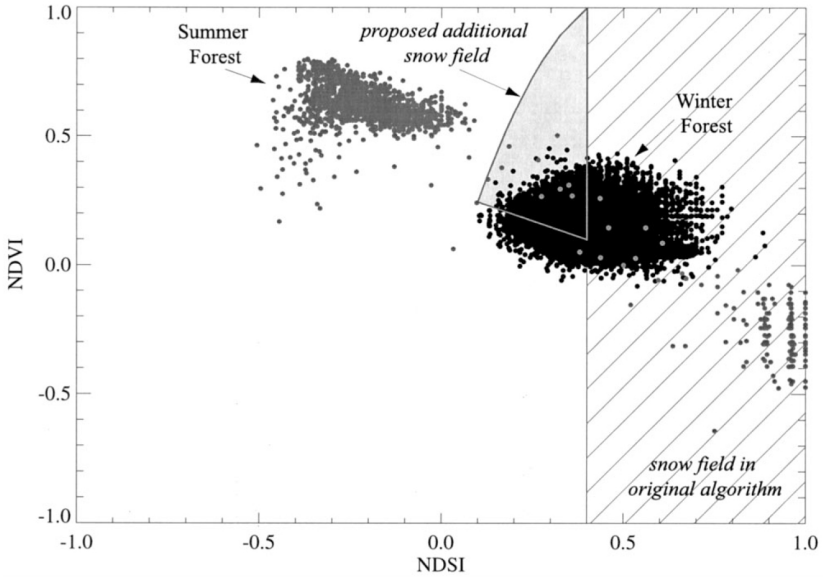


Figure 4.2: NDSI/NDVI snow classification. The hatched area represents pixels classified as snow when only using an NDSI-threshold. The shaded area depicts the additional field from Klein et al. (1998) also classified as snow.

an NDSI>0.4 only would be sufficient and how the different criteria change the result. For each setting, the quality of the outputs is assessed visually by comparing the generated snow mask with the original Landsat scene (i.e. RGB set to 4/3/2 or 3/2/1). Although the optimal value of the NDSI-threshold varies seasonally, a value of 0.4 seems to be appropriate for the winter season. However, the quality of the results is poor since a lot of totally snowfree pixels are classified as snow-covered. Especially forested areas cause extensive problems. Hence, the band 2 criterion (>0.10) seems to be required (see below). Consequently, the NDSI is then extended with this criterion, which yields better results for patchy snow-covered and for snowfree pixels. In comparison with the entire approach of Hall et al. (1998) including an NDVI, however, the quality of the snow mask in forested areas remains poor. Therefore it is concluded that the original algorithm shall be implemented with all criteria and, if necessary, these are modified to improve the quality of the binary Landsat snow mask.

One of the main problems with snow classification in mountainous areas, and especially in the winter season, is that large shadowed areas are present. In these areas the reflectance in the VIS is extremely low. Hence, values of snow-covered shadow pixels lower than 0.2 are often not classified as snow anymore. The following adaptations are also used to address these problems.

Band 2: On the one hand, lowering the original threshold of 0.10 yields a better result in dark shadowed areas (i.e. more snow pixels in shadows are classified correctly). On the other hand, however, this causes other dark surfaces such as the Rhone River or forested areas to be spuriously classified as snow, which Hall et al. (1998) intent to avoid by using a threshold for band 2. Here, completely wrong classifications are weighted stronger than perfect performance in the shadows. Therefore, the threshold is adapted to a value of 0.12.
Band 4: Using a band 4 threshold to exclude water bodies causes similar problems for shadows as did band 2. There are several shadowed snow areas with values b4 < 0.11. Since most water surfaces do not exceed the band 2 threshold, and therefore are already included there, the band 4 criterion is omitted. Moreover, the only large water body present in the study area is the Rhone River, where only few classification errors occur. Hence, the few river pixels wrongly classified as snow are accepted in order to avoid an extensive non-detection of snow in shadows. This compensates, at least to some extent, the poorer shadow result due to the adjustment of the band 2 threshold.
Finally, Table 4.1 gives an overview of the adapted Landsat snow mask threshold approach.

Table 4.1: Comparison of NDVI/NDSI snow mask algorithm parameters according to Hall et al. (1998) and modified thresholds for this thesis.

	Hall et al. (1998)	Master's Thesis
NDSI	> 0.40	> 0.40
Band 2	> 0.10	> 0.12
Band 4	> 0.11	-
NDSI/NDVI	Modelled:	Set to:
	$\sim$NDSI > 0.1, NDVI > 0.1	NDSI > 0.1, NDVI > 0.1

The explanations above emphasize the problems occurring with dark mountain shadows. Since these areas could even not perfectly be represented by the reference Landsat snow mask, the detailed analysis of the SPARC accuracy in shadow areas has to be dismissed.

4.1.2.3 Comparison of the two Landsat classification approaches

In order to obtain an impression of the differences between the supervised classification and the NDSI/NDVI approach, the two resulting snow masks are compared in the following. The comparison shows an overall agreement of 86.62% and a BIAS of 1.1, which indicates that the supervised snow mask classifies much more pixel as snow than the NDSI/NDVI approach does. This coincides with Figure 4.3, where much more red pixels than green pixels are present. Red pixels correspond to areas where only the supervised approach classifies snow, green pixels are related to snow pixels only detected by the NDSI/NDVI approach. Most of the disagreement emerges from the transition zone where the two approaches deal differently with thin or patchy

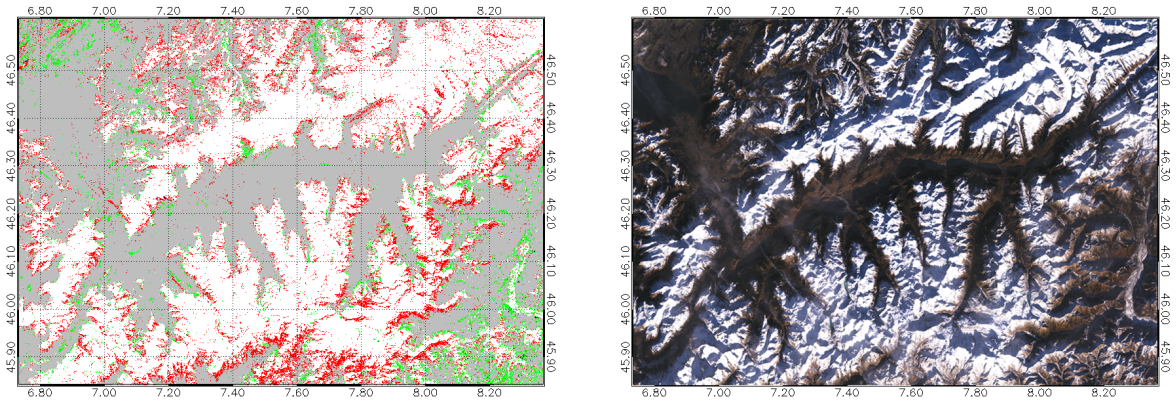


Figure 4.3: Difference between the snow mask generated with supervised classification and the snow mask calculated with the NDSI/NDVI approach (left) based on Landsat 7 from 2001/12/21 (right, RGB image). White and grey: agreement for snow and snowfree pixels, respectively. Red: pixels classified as snow only by supervised classification. Green: pixels classified as snow only by NDSI/NDVI approach.

snow layers. This feature appears again in the whole validation process, since also the SPARC algorithm is constructed in an other way than the approach for Landsat. On the one hand, the supervised approach seems to perform better in shadows. It often appears in the image that with the NDSI/NDVI approach the border of sun/shadow area is also taken as the boarder of snow/snowfree area. On the other hand this approach is preferable with very thin snow layers, which are nearly not represented in the supervised snow mask. Most of the green pixels are attributed to this fact. However, it has to be considered that for the supervised approach, all the different areas of an image have to be defined manually to either snow, snowfree or clouds. It is therefore essential to work precisely and it can not be excluded that for example the thinnest snow layers were taken into account sufficiently.
To conclude which one of the two approaches works better is difficult. With the supervised classification all different kinds of surfaces can be payed special attention and it might be preferable for a single scene. As already mentioned in section 4.1.2.1, the NDSI/NDVI approach is much faster (despite minor weaknesses in shadowed areas) and can be implemented into an automatic mode. It is therefore highly recommended when using a larger number of scenes.

4.1.2.4 Resampling

To analyze AVHRR with high-resolution Landsat images on a pixel-by-pixel basis, a reprojection of the Landsat snow mask to the geographical lat/long coordinate system and resampling (Nearest Neighbor) to the 0.01° x 0.01° grid resolution of AVHRR are required.

4.2 Pre-processing Steps for AVHRR Data

4.2.1 AVHRR snow classification with SPARC

The SPARC algorithm was already introduced in previous chapters (1.2 and 2.5.2). Here, the detailed procedure (see flowchart in Figure 4.4) and the thresholds for the snow classification according to Khlopenkov and Trishchenko (2007) and the modified version for the European Alps (Huesler et al. 2012) are explained.

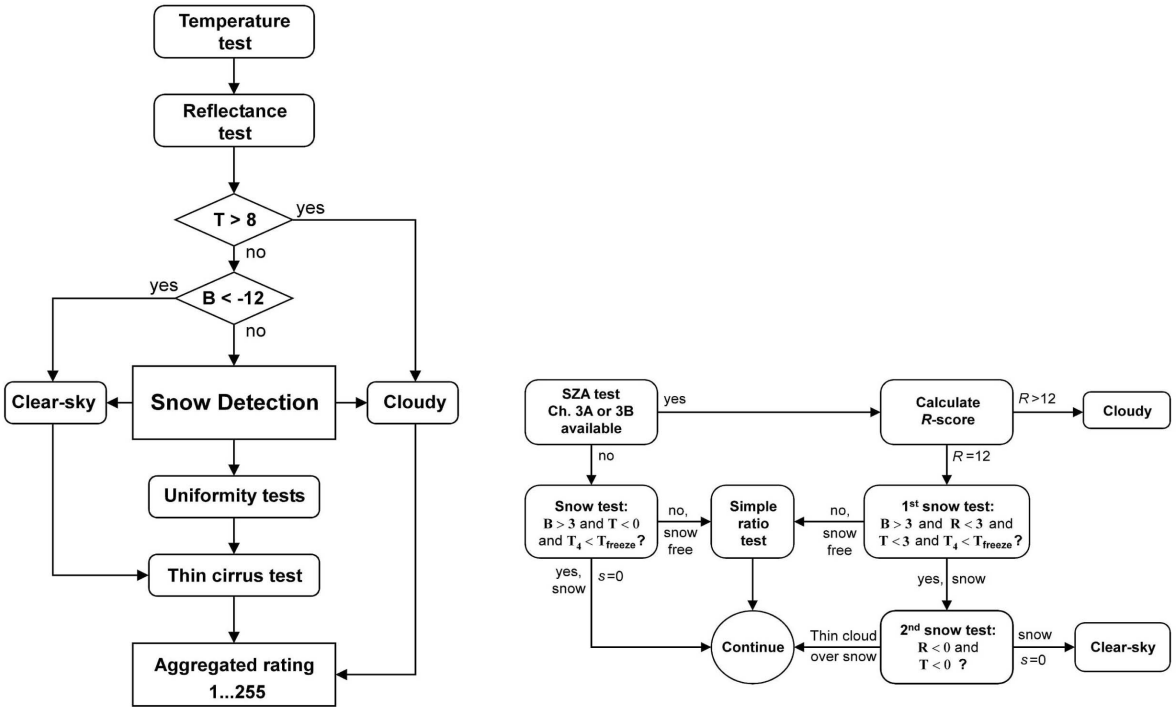


Figure 4.4: Total flowchart of the SPARC algorithm (left) and detailed flowchart of the box "Snow Detection" (right) (Khlopenkov and Trishchenko 2007).

The core of SPARC is composed of three major scores:

- B-score derived from reflectance brightness test in channel 1 (VIS: centered at 0.63μm)
- R-score derived from reflectance test in channel 3 (A or B) (NIR: channel 3A centered at 1.6μm or channel 3B centered at 3.9μm)
- T-score derived from brightness temperature test in channel 4 (TIR: centered at 10.8μm, combined with reanalysis skin temperature data).

Whereas the B- and the R-score are calculated directly from the radiometric data, the T-score has to be combined with surface skin temperature information. For the European Alps this data set is based on the daily gridded observation data set by Haylock et al. (2008) for the

years 1984-2001 and the downscaled COSMO data set available since 2002 (Marsigli et al. 2005). The overall rating leading to the defined class of the pixels is found by summing these scores:

$$F = T + B + R + \Sigma_i A_i. \tag{4.4}$$

Including an offset for each score which generates negative values for clear-sky and positive values for cloud-covered scenes, makes the summation of equation 4.4 possible. Thus, pixels with higher cloud cover have higher ratings. For pixels with ratings close to zero, additional tests are executed (see A_i in eq. 4.4) including a simple ratio test in channel 1 and 2, a uniformity test for reflectance as well as thermal uniformity and the thin cirrus test calculating the difference in brightness temperature between channel 4 and 5. Figure 4.4 (right) shows the procedure of the part in SPARC responsible for snow detection. If the R-score is higher than the cloud threshold (set to 12) then the pixel is classified as cloudy. For R-scores smaller than this limit the other two basic snow tests are applied. First, to be considered snow-covered, a pixel has to be sufficiently bright (B_{score} >3.0 over land or > -2.0 for water pixels), its R-score has to be smaller than 3.0 to remove cloud contamination and the brightness temperature has to be in the range of the surface (T_{score} <3.0) and lower than the freezing point T_{freeze} to eliminate ice/cold cloud. If the pixel is classified as snow, it has to pass a second stage to detect the presence of thin clouds or haze over snow. However, this second step was not implemented in Huesler et al. (2012), since for the snow mask it is irrelevant whether there are thin clouds or haze above the snow surface.
In addition to the original SPARC algorithm, Huesler et al. (2012) implemented several extensions. A digital elevation model (DEM) with information about the sun zenith angle at the time of acquisition was used for topographic modeling (with the imaging software ENVI, ITT Industries). With this, potentially shadowed pixels were identified and snow was classified for these pixels when lying above the mean snow altitude calculated independently for each snow mask. In addition, a land cover correction was applied for evergreen forest vegetation derived from GLC2000 (Hartley et al. 2006) (see section 3.6). It is based on the assumption that forest pixels with underlying snow have slightly lower reflectance values when compared to pure snow pixels areas and higher values compared to snowfree. Therefore the B-score was adjusted in forested areas to b-8. However, the land cover corrected snow mask is not analyzed in the framework of this thesis.
These snow masks are then used as the basis for the SPARC validation. Before these methods are presented, the following section addresses the problem of the geolocational shift, which is assumed to occur in the AVHRR data set.

4.2.2 "Correcting" geolocational shift

A difficulty when comparing different satellites is that the terrain corrections are not based on the same approaches. In the framework of the pixel-by-pixel validation, this may result in comparing pixels which do not correspond to each other. Thus, a systematic error expressed as misclassification is generated which cannot be related to a failure of the algorithm. Due to a poor orthorectification during the pre-processing chain of AVHRR, a shift of up to two pixels in North/South and in East/West direction is observed for 15 of total 17 scenes. Hence, before analyzing the behavior of SPARC, this incongruity related to the Landsat reference has to be corrected. For that purpose, the AVHRR snow mask is shifted ±2 pixels in x- and in y-direction leading to 25 possibilities. The shifted positions are chosen where the snow mask showed the highest agreement (=accuracy, see section 4.3) with the Landsat snow mask. With the same approach MODIS is analyzed against Landsat to allow for a consistent comparison. The scene from May 3th, 1995 is used as an example to show the effect of the geolocational shift between AVHRR and Landsat. Figure 4.5 depicts difference maps illustrating the classification difference between the two snowmasks. The analysis of the detailed spatial validation by means

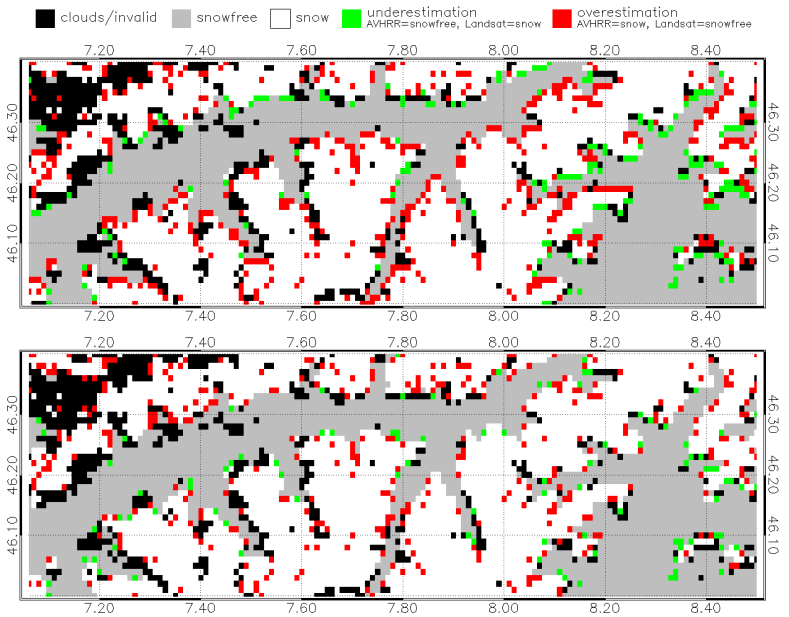


Figure 4.5: Difference map for scene N14, 95/05/03, 12:34 UTC with geolocational shift (top) and after correction (bottom).

of the difference maps follows in section 5.1. At the southern slopes (e.g. in the Rhone Valley or the Ossola Valley) SPARC underestimates snow cover (green color). On the other hand, on the northern slopes red color is dominant, which indicates a too optimistic snow classification. Since this green/red pattern is observed systematically over the whole image, it is assumed to indicate a shift by one pixel in the y-direction. To compensate this, the whole AVHRR

image is shifted by one pixel. This geolocational "correction" results in an accuracy gain of 2-5% (depending on scene and B-score setting). Table 4.2 lists the implemented shifting for all scenes. As can be seen, apart from two exceptions (N16, 03/01/25, 11:57 UTC and N17, 03/02/10, 10:28 UTC) all scenes have to be corrected. Despite the achieved improvement it has to be emphasized that it certainly is a temporary solution. If the problem is located at the AVHRR images, further research seems to be needed to correct this shift systematically. Unfortunately, no systematic pattern or any correlation with the orbit parameters could be found at a first view. The fact that also the MODIS images yielded better results when shifting them by one pixel indicates that also the Landsat reference might not have experienced a perfect orthorectification.

Table 4.2: Summary of all applied shift corrections (i.e. position, where the highest ACC occurred). Values >0 indicate eastward shift in x- and northward shift in y-direction. Values <0 indicate westward shift in x- and southward shift in y-direction.

Scene	**Shift x/y**	**Scene**	**Shift x/y**
N11, 93/03/10, 14:25 UTC	1/1	N16, 03/01/25, 11:57 UTC	0/0
N12, 95/05/03, 08:06 UTC	1/-1	N16, 03/01/25, 13:38 UTC	2/0
N14, 95/05/03, 12:34 UTC	0/1	N17, 03/01/25, 09:48 UTC	0/1
N14, 00/03/21, 14:51 UTC	0/1	*MODIS 03/01/25*	*0/1*
N15, 00/03/21, 07:38 UTC	-1/0	N16, 03/02/10, 12:19 UTC	1/0
N16, 01/12/14, 12:38 UTC	1/0	N17, 03/02/10, 10:28 UTC	0/0
MODIS 01/12/12	*0/1*	*MODIS 03/02/10*	*0/1*
N16, 01/12/21, 13:04 UTC	0/1	N16, 03/02/26, 12:42 UTC	2/1
MODIS 01/12/21	*0/1*	N17, 03/02/26, 09:28 UTC	-1/1
N16, 02/01/06, 11:52 UTC	1/-1	N17, 03/02/26, 11:08 UTC	-1/1
N16, 02/01/06, 13:34 UTC	2/1	*MODIS 03/02/26*	*0/1*
MODIS 02/01/06	*0/1*		

After these preparations of the AVHRR and Landsat data sets, the actual validation process of the SPARC algorithm can be started.

4.3 SPARC Validation

The equal study subset for both the resampled Landsat reference and the AVHRR image was already defined in section 3.1. Now, the both snow masks are compared on a pixel-by-pixel basis. To quantify the accuracy of these binary products, two-dimensional contingency tables and their skill measures are used (see Table 4.3) (Zappa 2008). The frequency of correctly classified snow pixels by SPARC can be related to those of Landsat (= true observations) and

vice versa for misclassified pixels. Two of the skill measures are applied and discussed in the following section.

Table 4.3: The 2 x 2 contingency table (Zappa 2008). S_1 = AVHRR snow, S_0 = AVHRR snowfree, O_1 = Landsat snow, O_o = Landsat snowfree.

	S_1	S_0	$\sum$
O_1	n_{11}	n_{01}	n_{x1}
O_0	n_{10}	n_{00}	n_{x0}
$\sum$	n_{1x}	n_{0x}	n_{xx}

Accuracy. The accuracy (=ACC, henceforward) is calculated by dividing the number of misclassified pixels by the total number of valid pixels in the study area (snow/snowfree).

$$ACC = \frac{(n_{11} + n_{00})}{n_{xx}} \tag{4.5}$$

BIAS. The BIAS is a quantitative measure of the relative frequency of the observed (Landsat) and predicted (AVHRR) events. Values above 1 indicate that SPARC classifies too many pixels as snow (=overestimation), values below 1 represent an underestimation. A value of 1 is referred to equal frequency of snow/snowfree pixels for the AVHRR and the Landsat image. In fact, the BIAS is the numeric result of the green and red colored pixels plotted in the difference maps (see chapter 5). These maps show for each AVHRR pixel, whether it matches the classification of the corresponding Landsat pixel or not. If not, it is colorized. If the AVHRR pixel is classified as snowfree but the Landsat pixel is defined as snow, then it is marked with green color for underestimation. The opposite holds for overestimated AVHRR snow pixels, which are flagged with red color. Consequently, more green than red pixels in the difference map indicate underestimation of the scene, which results in a BIAS below 1 and vice versa for BIAS values above 1.

$$BIAS = \frac{n_{1x}}{n_{x1}} \tag{4.6}$$

4.3.1 Algorithm improvement with threshold adaptation

The validation in general and the improvement of the algorithm are mainly focused on the variation of the B-score derived from the reflectance brightness test in channel 1 since this score is responsible for the differentiation between snow and snowfree areas. Snow masks calculated with the default value of b+3 are compared with variable values from b-3 to b-11. The aim of this approach is to identify the peak ACC. It is expected that the ACC values change by changing the B-score and peak at a certain, i.e. optimal, threshold. For this purpose, histograms are generated to illustrate the distribution of the pixels according to their scores and the development of the ACC when shifting the B-score is plotted. With

these, one is able to make conclusions about the improvement potential of the algorithm.
A minor influence is expected from the reflectance test in channel 3 (R-score), which is used for the discrimination of clouds and snow when limiting the R-score to small values. However, this score is also analyzed, even though in less detail since an additional cloud masking procedure is carried out, which also includes a snow-cloud discrimination part. Analogous to the B-score, the behavior of the classification ACC is investigated when varying the score from r-1 to r+10.
The T-score, which is responsible for the elimination of ice/cold clouds, is not expected to have an essential influence on the snow classification. Therefore its behavior is not analyzed.

4.3.2 Factors influencing the accuracy

After the threshold adaptation, correlation analyses by applying linear regression are performed to detect possible sources of the remaining errors, such as ACC dependencies on satellite and sun zenith angles, snow extent, forest distribution and on resolution differences of the original images of AVHRR and Landsat. An additional potential factor influencing the ACC are changing illumination and viewing conditions. This issue is addressed in the following section.

4.3.2.1 BRDF effect

The signal observed at the satellite sensor is a function of the amount of reflected radiance depending on the directions of sun irradiance and (satellite) viewing geometry. The so-called BRDF effect is responsible for the fact that the same surface looks differently and generates distinct reflectance values when the sun or/and the satellite change their position (Lillesand et al. 2004). This is expected to have an influence on the snow classification ACC and is therefore also a potential factor to be investigated. The position of the satellite relative to the sun can be expressed by means of the sun-satellite relative azimuth angle φ. It is defined as:

$$\varphi = \varphi_{sun} - \varphi_{sat} \tag{4.7}$$

where φ_{sun} is the sun azimuth and φ_{sat} the satellite azimuth.
In order to compensate the BRDF effect, Luo et al. (2008) suggest to separate the pixels into two groups: a) a forward scattering hemisphere (FH) with sun-satellite relative azimuth angle $\varphi \in$ 90°-270°, where the position of the satellite is in opposite direction of the sun and b) a backward scattering hemisphere (BH) with $\varphi \in$ 0°-90° and 270°-360°, where the satellite is located in the same direction as the sun. As already introduced in section 2.4 snow is strongly forward scattering.

Unfortunately, the available scenes and the small dimensions of the study area were not suitable for this purpose. The small sample sizes, namely, six scenes only with backward scattering, eight scenes with forward scattering and two scenes with both (=nadir passed the image) inhibited a reasonable analysis. Therefore the influence of the BRDF effect is not further considered here. Just for completeness it is mentioned that at a first look the scenes with backward scattering seem to be classified slightly better (ACC=79.78% with b+3) than those with forward scattering (ACC=76.92%). However, it certainly requires further research as it is expected to influence the ACC. This is emphasized by the findings of Luo et al. (2008) who concluded that the differentiation of the two categories BH and FH can substantially reduce the reflectance variability in clear-sky images.

4.4 Snow Fraction Analysis

The validation process described in the previous sections is based on Landsat data resampled to the resolution of AVHRR to allow for a pixel-by-pixel comparison. In addition to this, an approach based on subpixel ACC is implemented. By keeping the Landsat resolution at 30m it is possible to calculate a fractional snow cover area (SCA) for every AVHRR pixel, whereby one AVHRR pixel corresponds to 27 x 38 Landsat pixels. When assuming that SPARC classifies a pixel as snow if the SCA is greater or equal 50% (i.e. a binary snowmask), then one can study how many pixels classified as snow lie above this thresholds. Thus, these pixels correspond to correct classifications. The same holds for pixels classified as snowfree with SCA<50%. The optimal B-score threshold defined as the B-score where the sum of snow pixels with SCA≥50% and snowfree pixels with SCA<50% reaches its maximum, is then compared with the results from the pixel-by-pixel validation. The optimum would be a sum of 200%. In an additional approach, the B-score threshold is kept constant at a certain value (i.e. namely at the B-score threshold, which is found to be the best in the course of the pixel-by-pixel validation) and the SCA-threshold is changed so that the ACC of the snow mask is maximized. The comparison of these results with the pixel-by-pixel approach is expected to give an idea of the resampling error of the Landsat image on the one hand and to assess at which SCA SPARC classifies a pixel as snow-covered on the other hand.

This method can also be used to calculate customized snow cover maps. The WSL (Swiss Federal Institute for Forest, Snow and Landscape Research) Institute for Snow and Avalanche Research SLF, for instance, is interested in pixels which are certainly snowfree (e.g. pixels with SCA<20%). Such a map is presented in section 5.3.2 of the Results Chapter.

Chapter 5

Results and Discussion

5.1 SPARC Validation

The validation of the AVHRR snow mask with the default values of SPARC and the applied improvement steps are shown and discussed in the following sections. As an example, the scene N17 2003/02/10, 10:28 is chosen to illustrate the applied steps and the results.

5.1.1 Accuracy of SPARC with default values

First, the resulting snow masks of the SPARC algorithm are analyzed for the default value of the B-score threshold (b+3) to have a first impression of the classification ACC. Table 5.1 lists the ACC of all considered scenes. Apart from one exception (N14, 95/05/03, 12:34 UTC), the total ACC for the AVHRR scenes ranges from 72% to 84% with a mean of 78.65%. The good performance of the N14 scene, which classifies over 90% correctly, might be a direct consequence of the fact with the retreat of snow to higher altitudes, less and less snow pixels lie in the forest zone, where snow is hard to detect. The BIAS of 0.77 (averaged over all scenes) shows that SPARC underestimates snow cover. This discrepancy between AVHRR and Landsat can also be illustrated in difference maps (see Figure 5.1 top), where much more green pixels than red pixels are present. Green pixels are referred to underestimation (= AVHRR classifies snowfree, Landsat classifies snow), red pixels correspond to overestimation (= AVHRR classifies snow, Landsat classifies snowfree). The valley structures are evident, which implies that most of the differences are located in the transition zones from snowfree areas to snow when ascending to higher altitudes. One the one hand this is related to the different snow algorithms of AVHRR and Landsat, i.e. the different treatment of mixed pixels, and on the other hand to forest cover, where snow under trees is hard to detect (Huesler et al. 2012). Looking at the difference maps of all scenes does not lead to any conclusion about a systematic behavior of misclassification (see Appendix A). While some scenes show most of their underestimation in the Rhone Valley (Switzerland) as well as in the Ossola Valley (Italy)

Table 5.1: Results for ACC with B-score default value b+3. ♯p corresponds to the number of valid pixels (i.e. snow/snowfree, no clouds). Scene N16, 01/12/14 has an above-average pixel number since the ROI is larger, scene N16, 01/12/21 has a below-average pixel number since it is highly cloud contaminated.

			ACC [%]		
Scene	**♯p**	**Bias**	**Total**	**Non-Forest**	**Forest**
N11, 93/03/10, 14:25 UTC	5108	0.845	81.60	84.93	78.94
N12, 95/05/03, 08:06 UTC	4611	0.703	77.86	73.93	80.75
N14, 95/05/03, 12:34 UTC	5305	1.003	91.40	91.55	91.28
N14, 00/03/21, 14:51 UTC	5032	0.740	82.31	75.81	86.74
N15, 00/03/21, 07:38 UTC	5259	0.609	73.84	64.77	80.81
N16, 01/12/14, 12:38 UTC	8510	0.593	72.93	68.93	76.37
MODIS 01/12/14	*7441*	*1.023*	*88.52*	*88.73*	*88.32*
N16, 01/12/21, 1304 UTC	2817	0.516	73.80	57.51	84.23
MODIS 01/12/21	*4138*	*0.974*	*83.57*	*79.18*	*86.48*
N16, 02/01/06, 11:52 UTC	5829	0.660	73.49	66.82	78.62
N16, 02/01/06, 13:34 UTC	4753	0.818	73.51	72.78	74.26
MODIS 02/01/06	*5076*	*1.105*	*82.62*	*81.96*	*83.17*
N16, 03/01/25, 11:57 UTC	5815	0.705	72.61	75.08	70.78
N16, 03/01/25, 13:38 UTC	5792	0.875	81.79	88.98	76.41
N17, 03/01/25, 09:48 UTC	5945	0.863	82.46	85.48	80.18
MODIS 03/01/25	*5271*	*1.086*	*86.23*	*92.66*	*80.69*
N16, 03/02/10, 12:19 UTC	5876	0.793	77.03	83.01	72.48
N17, 03/02/10, 10:28 UTC	5888	0.795	79.14	84.20	75.34
MODIS 03/02/10	*5575*	*1.086*	*87.78*	*95.07*	*81.89*
N16, 03/02/26, 12:42 UTC	5605	0.819	80.59	86.88	75.46
N17, 03/02/26, 09:28 UTC	5513	0.974	84.84	93.57	77.49
N17, 03/02/26, 11:08 UTC	5623	0.768	73.02	76.26	70.43
MODIS 03/02/26	*5416*	*1.117*	*87.41*	*95.29*	*81.05*
Mean of all AVHRR scenes		0.77	78.36	78.26	78.27

and most of their overestimation in higher altitudes of smaller north-south orientated valleys, other scenes seem to have a slight east-west discrepancy, which could be a consequence of different sun and satellite viewing geometries. For certain scenes even no pattern is observed. Further research is needed here to analyze for example the influence of different aspects on the classification ACC. Huesler et al. (2012) mentioned in the course of their validation with MODIS data that there is no exceptional weakness for one aspect. Only slightly lower values were observed for north-facing slopes showing the lowest values for SPARC default settings. These could be improved in the adapted version by giving special consideration to shadowed areas.

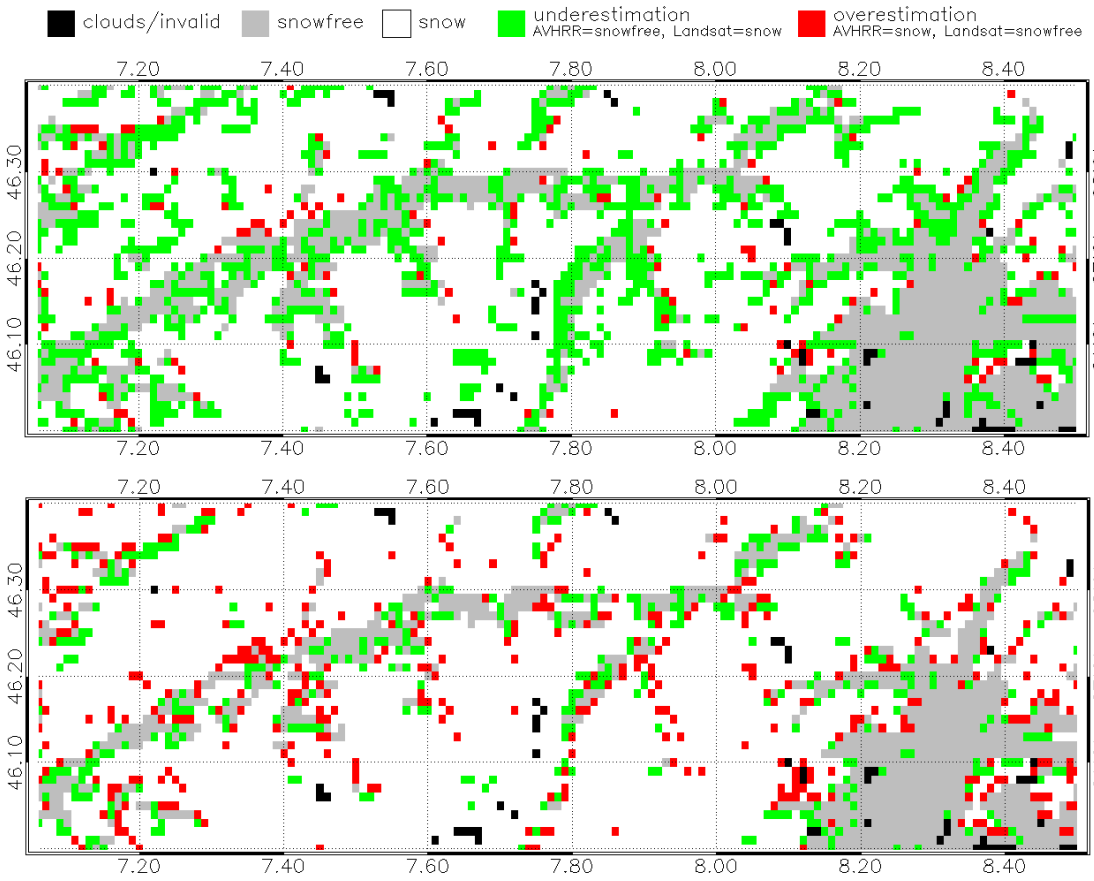


Figure 5.1: Difference maps for b+3 (top) and b-6 (bottom) for AVHRR scene N17, 03/02/10, 10:28 UTC referred to the resampled Landsat reference.

An additional source of differences could also be the time shift of the scene acquisition leading to different illumination conditions. Especially for afternoon satellites the time gap is evident (Landsat scenes are always acquired between 09:30 and 10:00 UTC). Moreover, it has to be considered that the transition zones, i.e. where snow accumulates close to its melting point, are sensitive already to small temperature fluctuations (IPCC 2007b). This means that in the course of climate change, this zone - or the uncertainty for this zone - also underlies a

certain fluctuation.
The MODIS product in comparison to Landsat performs much better with values between 83% and 89% and a BIAS of around 1.0 for all scenes with a slight tendency to an overestimation. This coincides with several studies analyzing MODIS snow products (see e.g. Hall et al. 2001).
Furthermore, Huesler et al. (2012) found that SPARC causes problems with early morning satellites (here N12 and N15), due to the flat sun angles causing large shadowed areas and changing the reflectance properties of snow. In order to minimize these BRDF effects (Lillesand et al. 2004) (see section 4.3.2) and extensive shadows (Salminen et al. 2009), N12 and N15 are excluded. Moreover, scenes from December are also found to have below average ACCs, probably due to extensive cloud cover and frequent fog occurrence. Indeed, as seen in Table 5.1 scenes covering these criteria have a relatively poor performance. Especially N15 and the two N16-scenes from December 2001 have low ACCs, even though they are not much lower than for example N16 from 02/01/06. The N12-scene is slightly below the average. Nevertheless, with the background of the findings of Huesler et al. (2012), it is decided that N12, N15 and December scenes are not considered any further here to avoid the risk of adulterating the adaptation process and the results.
ACC in forest-covered areas is also analyzed separately. Surprisingly, for scenes before 2003, the ACC is higher than the total. In section 5.2.4 this behavior is related to the dependency on snow extent. However, wether it is just a coincidence or a systematic behavior cannot be answered straightforwardly with such a limited number of scenes. The comparison with MODIS data shows that AVHRR experiences the same difficulties as the state-of-the-art satellite system. Nevertheless, the ACC values are on a much lower level than those of the MODIS snow mask.

As an improvement step, the underestimation of SPARC will be corrected in the following section. This is done by shifting the critical B-score to lower values, which leads to more pixels being classified as snow.

5.1.2 B-score threshold adaptation

The distribution of the snow and snowfree pixels over the spectrum of the B-score values from around -15 to +47 is depicted as histogram plot in Figure 5.2. Since the critical threshold is set to b+3 by the SPARC default settings, values to the right of the red threshold bar correspond to AVHRR snow pixels, either correctly classified (black curve) or misclassified (orange curve). Analogously, values below b+3 are represented by snowfree pixels (correct in blue, misclassified in green). A large number of very bright snow pixels, possibly oversaturated, is present and explains the big bar at around score +46. In addition, there are some pixels classified as snowfree although they have score values higher than the threshold. This might

be related to the fact that in the course of other tests (R-, T-scores) they were classified as snowfree. By means of these histogram plots, the potential of an improvement of the SPARC algorithm by adjusting the B-score can be illustrated. Shifting the threshold to the intersect of the curve of the misclassified snowfree pixels and the correctly classified snowfree pixels, results in a gain of the first (i.e. increase of ACC) and a loss of the latter (i.e. decrease of ACC). The overall gain, then, is the difference of these two (see grey area in Figure 5.2 top). A difficulty in finding the optimal threshold is that this intersect varies from scene to scene. Hence, it is not possible to set one universally valid value. However, it can be limited to a certain range.

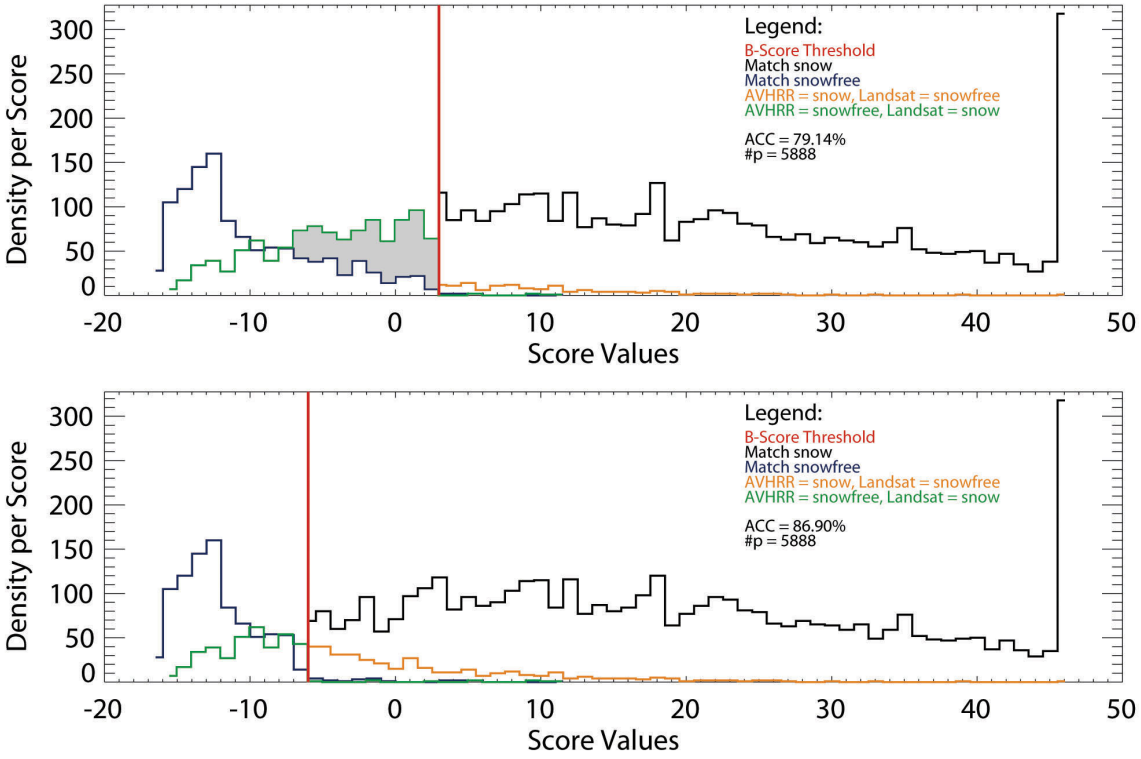


Figure 5.2: B-score histograms showing the distribution of the B-score values for all AVHRR pixels of the study area before improvement with b+3 (top) and after improvement with b-6 (bottom) for scene N17, 03/02/10, 10:28 UTC. The grey area represents the integral of the improvement potential when lowering the B-score. The optimum for this scene is at b-7.

To find the peak value, i.e. the optimal B-score threshold where all scenes have their highest ACC, the development of the ACC with different B-scores is plotted in Figure 5.3. Due to the findings of Huesler et al. (2012) and the results going to be presented in section 5.2.1 only scenes with satellite zenith angels smaller than 40° are considered (including N16, 03/02/10, 12:19 UTC with a slightly larger angle of 40.2°). The three lowest graphs correspond to the problematic scenes discussed already above (N15 and the two scenes from December), which

confirms the procedure to remove these images. Note that N12 is missing in the figure because of its satellite zenith angle of 53.1°. Figure 5.3 also underlines the variable scene-to-scene behavior and the improvement potential related to the default value (b+3). As the threshold is lowered, the ACC increases, peaks at a certain point and drops after that again. Values lower than b-9 lead to poorer results for all scenes, namely to an overestimation of snow cover. When

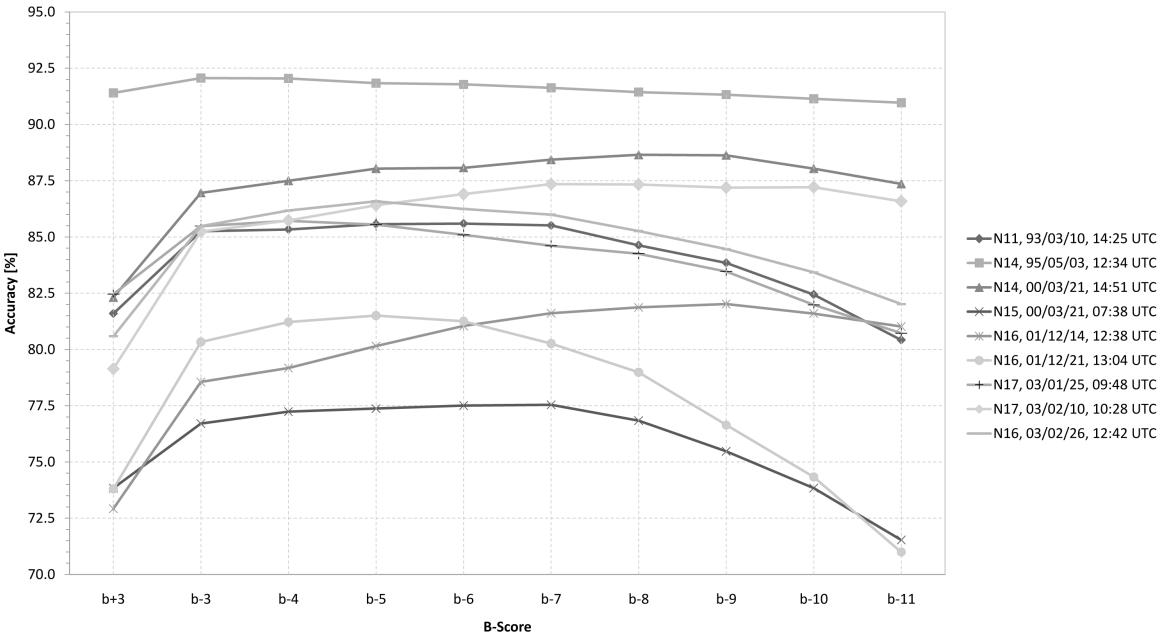


Figure 5.3: Development of the SPARC algorithm ACC depending on different B-score thresholds. Only scenes with satellite zenith angles <40° are displayed.

shifting the score threshold in a range of b-4 to b-8 the differences generally do not change the ACC substantially. Only the two N16 December scenes show higher variabilities. This insensitivity of SPARC concerning the application of different thresholds is a major advantage of this algorithm as it makes it very robust. Consequently, the algorithm is best applicable for climate related long-term applications. The curve of N14, 95/05/03, for example, which already displays a high ACC, stays nearly constant. It also shows that an increasing snow line in spring has a positive effect on the classification ACC since less forest-covered area is affected by snow cover. Finally, the optimal thresholds of each scene are averaged and yield a mean optimal B-score threshold of b-6. The new difference map of the example scene (N17 2003/02/10, 10:28) shows a less colored image and many of the underestimated pixels vanish (see Figure 5.1 bottom). The BIAS lies close to 1 (1.001), which means that the remaining errors are nearly equally distributed.

The implementation of the new SPARC B-score threshold leads to the ACC values listed in Table 5.2. Not all scenes experienced an equally good improvement. The last row shows the reached improvement in numbers averaged over all scenes. The difference from the results

Table 5.2: Final results for ACC after threshold adaptation for B-score b-6. ♯p corresponds to the number of valid pixels (i.e. snow/snowfree, no clouds). Scene 01/12/14 has above-average pixel number since the ROI is larger, scene 01/12/21 has below-average pixel number since it is highly cloud contaminated.

			ACC [%]		
Scene	**♯p**	**Bias**	**Total**	**Non-Forest**	**Forest**
N11, 93/03/10, 14:25 UTC	5108	1.074	85.59	92.02	80.45
N12, 95/05/03, 08:06 UTC	4611	1.105	83.17	86.91	80.41
N14, 95/05/03, 12:34 UTC	5305	1.084	91.78	92.50	91.18
N14, 00/03/21, 14:51 UTC	5032	0.983	88.08	86.70	89.01
N15, 00/03/21, 07:38 UTC	5259	0.824	77.51	70.50	82.90
N16, 01/12/14, 12:38 UTC	8510	0.821	81.05	81.22	80.89
MODIS 01/12/14	*7441*	*1.023*	*88.52*	*88.73*	*88.32*
N16, 01/12/21, 13:04 UTC	2817	0.931	81.26	74.61	85.51
MODIS 01/12/21	*4138*	*0.974*	*83.57*	*79.18*	*86.48*
N16, 02/01/06, 11:52 UTC	5829	1.085	81.35	82.82	80.22
N16, 02/01/06, 13:34 UTC	4753	1.103	77.36	82.28	72.34
MODIS 02/01/06	*5076*	*1.105*	*82.62*	*81.96*	*83.17*
N16, 03/01/25, 11:57 UTC	5815	0.994	82.80	90.46	77.17
N16, 03/01/25, 13:38 UTC	5792	1.095	82.84	92.94	75.29
N17, 03/01/25, 09:48 UTC	5945	1.056	85.10	91.94	79.94
MODIS 03/01/25	*5271*	*1.086*	*86.23*	*92.66*	*80.69*
N16, 03/02/10, 12:19 UTC	5876	1.018	85.62	93.42	79.69
N17, 03/02/10, 10:28 UTC	5888	1.001	86.91	93.75	81.77
MODIS 03/02/10	*5575*	*1.086*	*87.78*	*95.07*	*81.89*
N16, 03/02/26, 12:42 UTC	5605	1.077	86.24	93.80	80.09
N17, 03/02/26, 09:28 UTC	5513	1.150	83.44	94.60	74.05
N17, 03/02/26, 11:08 UTC	5623	1.101	80.69	90.81	72.57
MODIS 03/02/26	*5416*	*1.117*	*87.41*	*95.29*	*81.05*
Mean of all AVHRR scenes		1.03	83.57	87.72	80.21
Mean difference to b+3		**+0.26**	**+5.21**	**+9.46**	**+1.94**

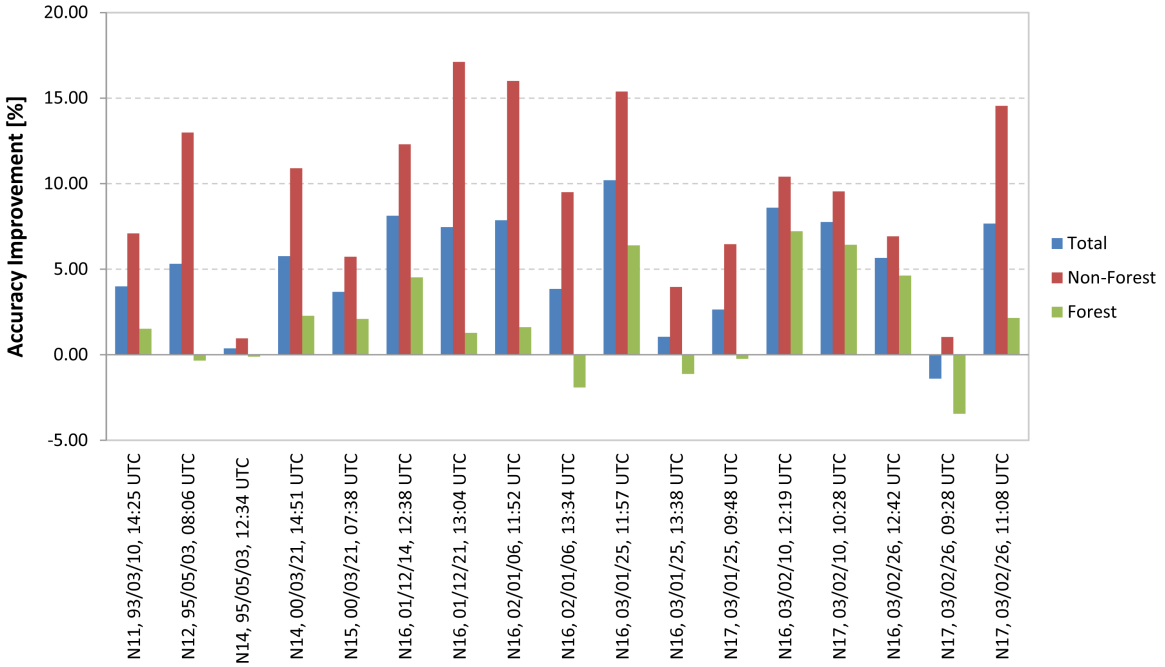


Figure 5.4: ACC improvement for the SPARC classification from the default B-score threshold b+3 to b-6 split into values for all pixels (total), non-forest pixels and forest pixels.

with the default value b+3 to those with b-6 is much larger for forest-free areas than for forested regions. Whereas the ACC on non-forest pixels is, averaged over all scenes, improved by 9.46%, forest-covered pixels only show an improvement of 1.94%. Taking into account the high amount of forest in the study area, this explains the moderate overall gain of the SPARC classification ACC of 5.21%. The improvement from the default value to b-6 is illustrated in Figure 5.4 for all scenes. It also shows that not all scenes benefit from the threshold adaptation in the same amount, or benefit at all. In some cases, ACC on forest-covered areas is reduced. Nevertheless, negatively affected scenes have satellite zenith angles greater than 50° (apart from N17, 03/01/25 with a very small ACC reduction) and only for scene N17, 2003/03/26, 09:28 UTC a negative effect on the total ACC can be observed. For all scenes that fulfill the criterion of the satellite zenith angle ($\leq$40°), the threshold adaptation lead to an improvement of the ACC.

For some dates, where more than one scene per day is acquired, such as 1995/05/03 or 2003/02/26, there are substantial differences between the corresponding AVHRR scenes. They amount to 8.6% and 5.5%, respectively. These differences are assumed to be related to the influencing factors discussed in section 5.2.

The B-score shift brings SPARC close to the MODIS snow mask. In some cases, the latter performs only 1-2% better. Such scenes are N16, 01/12/21, 13:04; N16, 02/01/06, 11:52; N17, 03/01/25, 09:48, N16, 03/02/10, 12:19; N17, 03/02/10, 10:28 and N16, 03/02/26, 12:42. These examples emphasize the quality of the result. Although MODIS does not generate a

perfect snow mask as well, it is a very good indicator for the quality of an algorithm. The highest deviations are observed with N16, 01/12/14, 12:38 (-7.84%) and N17, 03/02/26, 11:08 (-6.72%).

Up to now, all land cover types were treated equally, even though the reflectance depends on the surface properties. Especially, snow under a forest canopy affects the reflectance and leads to mixed pixel information. The following subsection presents a possible approach to address this problem.

5.1.2.1 Land cover correction

Due to their different reflectance properties, forest-covered and forest-free areas usually have different optimal thresholds. A separate treatment, i.e. a correction of the score thresholds for different land cover types, is expected to yield better results. To implement an improved land cover correction, the data plotted in Figure 5.3 is split into forest-covered and forest-free pixels. Figure 5.5 illustrates how the ACC changes by shifting the B-score for this case. The details of each scene are of minor importance here since the figure should simply emphasize the difference between forest and non-forest pixels. Again, only scenes with satellite viewing

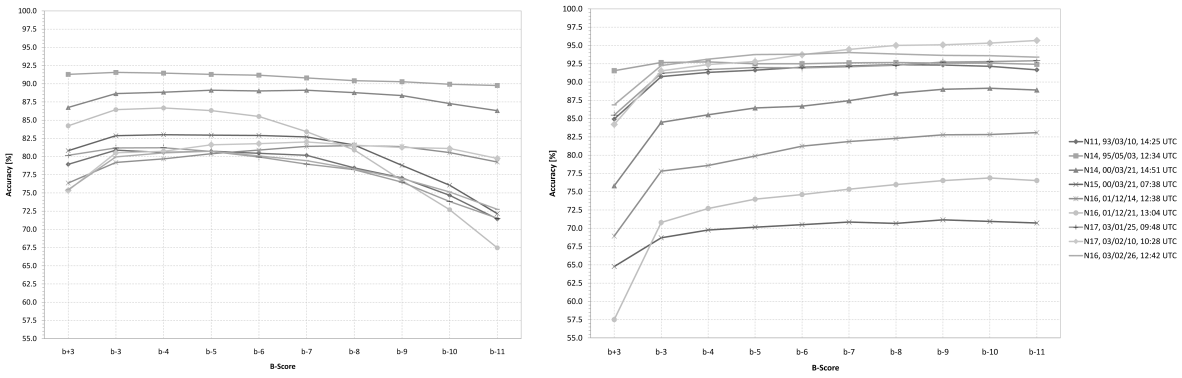


Figure 5.5: Development of the SPARC algorithm ACC depending on different B-score thresholds for forest-covered pixels (left) and non-forest pixels (right). Only scenes with satellite zenith angles $<40°$ are displayed.

angles $\leq 40°$ (without N15 and December scenes) are considered for the definition of the peak score. Whereas non-forested pixels show best results at around b-9, forest pixels experience the highest ACC at b-5. Contrary to the expectations discussed in section 4.2.1 with the concept of a SPARC land cover correction, forest pixels do not have lower scores as assumed due to the dark canopy and the resulting lower reflectance. Therefore, SPARC is run again with a B-score of b-9, and for forest pixels a land cover correction, e.g. a B-score of b-5 is applied. Since this optimal threshold is a "mean value" it is obvious that it is not optimal for all scenes. For certain scenes, these settings lead to a slight improvement of the snow mask of around 1-2%. For others, no improvement can be observed and no systematic behavior is identifiable. A first test during the work of Huesler et al. (2012) with the snow time series

validated with MODIS showed that the results are much poorer with these settings. This seems reasonable since the thresholds were expected to be inverted (i.e. higher scores for forest-free pixels compared to forest pixels). It is therefore concluded that possibly not the land cover type "forest" is the influencing factor for the behavior of the unexpected peak score values but the geographical location of the forest pixels per se. The transition zone from snow to snowfree areas is difficult to be classified correctly (see section 5.1.1 and Figure 5.1). If this zone is located in the wooded belt, then the ACC of forest pixels is also limited due to these conditions. This problem is addressed again in section 5.2.6.
It has become clear that applying a powerful land cover correction is difficult and further research is needed. With this simple B-score changes it seems not possible to improve the result. A weakness of the chosen approach could be that snow and snowfree pixels were not treated separately. Only forest pixels with underlying snow are expected to have slightly higher reflectance values compared to snowfree conditions. By simply analyzing the peak score of forest and non-forest areas, mixed results are the consequence. However, it shows that there is a potential to influence the ACC with a land cover correction.

5.1.3 R-Score threshold adaptation

The B-score in SPARC is responsible for the differentiation between snow and snowfree land. In addition, the R-score derived from the reflectance test in channel 3 separates clouds from snow. Therefore, this score is not expected to have an influence on the discrimination between snow and snowfree pixels and hence on the snow classification ACC. Indeed, the analysis shows that varying the initial R-score of r+3 from r-1 to r+10 does not affect the ACC positively. From r+1 to r+10 the ACC changes only in a range of $\pm$0.1-0.5%. This difference was found too low to be considered for further validation. However, the results clearly show that the R-scores of r+0 and especially r-1 yield substantially poorer result and that the default value of r+3 is well chosen. For almost every scene this value generates the best snow mask.

5.2 Discussion of Factors influencing the Accuracy

Providing explanations for the results yielded with the threshold adaptations to b-6 in the previous sections is not straightforward since the SPARC classification ACC is influenced by several factors and every scene has its special characteristics. Therefore, the following discussion is focused on such factors and degrees of influences are estimated.

5.2.1 Satellite zenith angle

The snow time series of Huesler et al. (2011) does not contain satellite zenith angles larger than 40° since they are expected to have a negative influence on the classification ACC. In this

context, the scenes' satellite zenith angles are correlated with its resulting ACC (based on a B-score of b-6). Figure 5.6 clearly shows that the SPARC performance decreases with increasing viewing angles with a correlation of -0.84, which results in an r^2=0.71 with a root mean square error (RMSE) of 1.88%. The highest ACC is reached with the nadir scene (viewing angle of 0.3°) from May 3th, 1995 (N14). Whereas scenes with angles at around 10° to 15° have equal ACC values to these between 30° and 40°, zenith angles above 50° cause substantially poorer results with a minimum ACC below 80% for the N16-scene in January 2002. Based on these findings, keeping the angle threshold at 40° can be recommended since this does not decrease the SPARC classification performance. A further decrease of the critical allowed angle is not necessary. Consequently, less scenes have to be excluded from the time series and hence the robustness is enhanced.

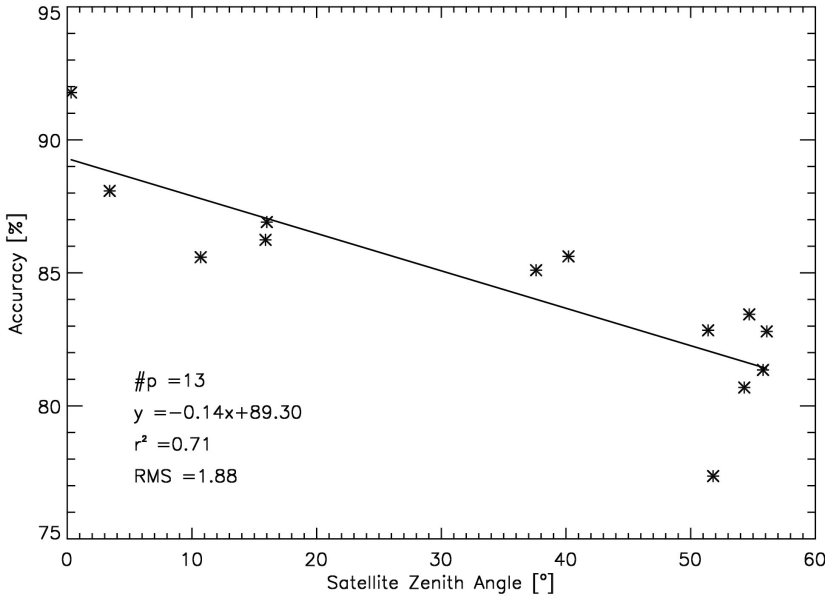


Figure 5.6: Correlation of satellite zenith angles and SPARC classification ACC.

5.2.2 Sun zenith angle

Similar to the satellite angles, the illumination conditions of the sun might influence the ACC of the SPARC classification. Figure 5.7 (left) shows that there is a dependency with r^2=0.50 and a RMSE of 2.48%. According to that, the ACC decreases with an increasing sun zenith angle (r=-0.71). This is comprehensible when considering that a flatter sun angle changes the irradiance and causes, for example, more shadows, which often lead to classification problems. Salminen et al. (2009) emphasize that reflectance of snow in shadow can be as low as 0.18. For Landsat scenes similar values were found in section 4.1.2.2. Since sun angles >55° are generally not representative for winter sun geometries over the Alps (Huesler et al. 2012), the nearly perfectly classified spring scene N14, 95/05/03, 12:34 UTC with an angle of 34° is

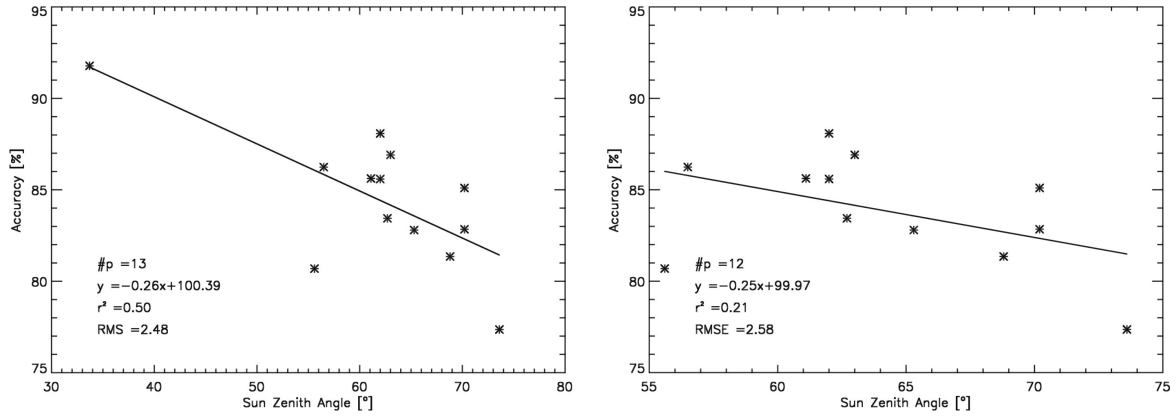


Figure 5.7: Correlation of sun zenith angles and SPARC classification ACC (left). Right: The 'outlier' from May, 1995 with an angle of 34° and an ACC of 91.78% is removed.

excluded from the correlation. After the removal r^2 is reduced to 0.21 (see Figure 5.7, right). Only a weak correlation remains and whereas the sun zenith changes by nearly 25°, the ACC variability stays in a moderate range of around 10%-points. A much larger set of scenes is required to draw any conclusions about the strength of the correlation for the whole time series because it cannot be excluded that there are more scenes available with angles between 30° and 50° and an ACC between 85% and 90%. In that case, the N14-scene from May 1995 would not be an exception and r^2 would be high again at a level of the left plot. If such a strong dependency is present, a threshold will have to be introduced which removes scenes with too large sun zenith angles. In any case, Khlopenkov and Trishchenko (2007) limit the sun zenith angle for SPARC depending on the brightness temperature of AVHRR channel 5 T_5 to:

$$\Theta_0^{max} = 89.5° - 0.000025(T_5 - 223)^3. \tag{5.1}$$

Nevertheless, these results imply that there seems to be a much smaller influence of the sun zenith angle than of the satellite viewing angle. This conclusion is also important in the context of satellite orbit drift. The drift causes a shift in overpass time over the lifetime of a sensor and consequently changes the sun zenith angle. If it can be verified with a much larger sample that the sun zenith angle is of minor importance, then this could suggest that the orbital drift has a minor influence on the classification ACC and does therefore not lead to an artificial trend in the time series (Huesler et al. 2012). However, this is a first assumption and requires further research.

5.2.3 Date of acquisition

In a next step it is also analyzed if the classification ACC depends on the date of acquisition since the ACC is expected to alter with changing snow conditions during the winter season.

Especially during the beginning of the winter and during snowmelt Simic et al. (2004) found reduced ACC values with MODIS and NOAA data. By sorting the snow masks calculated with a B-score of b-6 from December to May, a correlation of +0.8 (r^2=0.65) is found. This implies that the SPARC performance improves throughout the winter. As the snow layer evolves, the structure becomes clearer and is therefore easier to detect. Simic et al. (2004) also showed that the MODIS snow classification ACC (validated with Landsat ETM+) improves with increasing snow depth. This is of particular importance at the beginning of the accumulation period and towards the end of snow melt. A similar behavior is conceivable for AVHRR even though this cannot be confirmed due to missing snow depth measurements. Nevertheless,

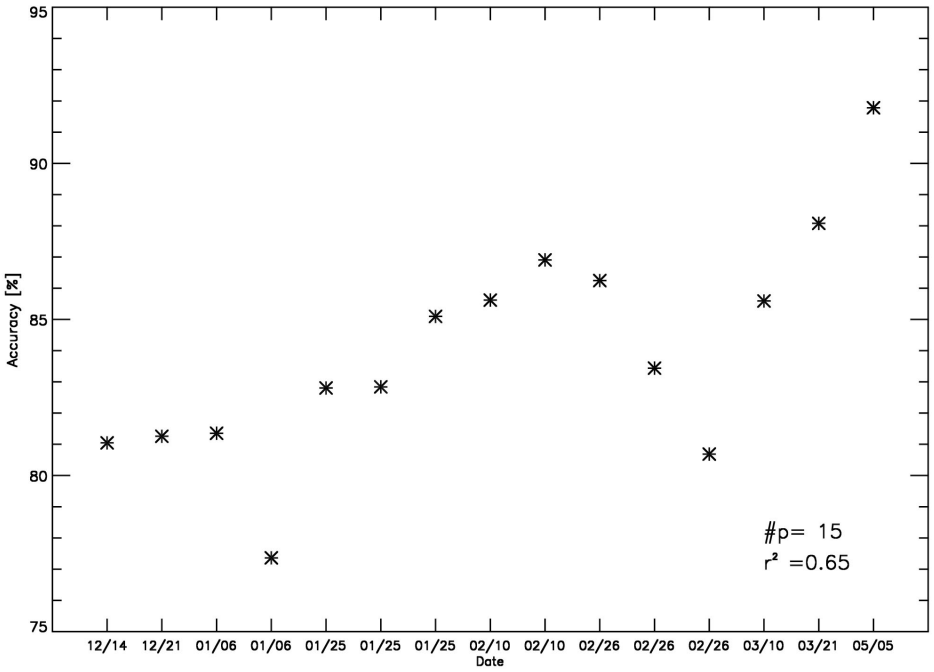


Figure 5.8: Correlation of date of scene acquisition and SPARC classification ACC.

several points have to be considered when making conclusions about the ACC of the 25-year time series based on these findings. The result is not based on a time series, which means that the more or less randomly chosen scenes are not proven to be representative for the respective time of year or put in other words, wether the snowpack of a scene conforms with the typical properties expected for the season during which the image was acquired. As an example the images from 1993/03/10 and 2000/03/21 show that, although both are acquired in March with a difference of eleven days, the first has a snow extent of 65%, while the latter has 51% snow cover. The two December scenes also have a large difference in snow extent of 18%. Furthermore, it cannot be concluded that the linear trend continues after May. Also summer and autumn scenes would have to be analyzed.

The ACC variability throughout the winter has to be considered for example when making conclusions about a potential shift of the snow line in the course of climate change. Thus,

uncertainties of the algorithm ACC which are related to climate depending factors and occur in certain months (e.g. in November or December due to shallow snow layers), could be shifted to other months in the future. This means that it cannot be concluded that the problem of the low snow depth will be present today in the same months like in future decades.

5.2.4 Snow extent

An image where the area is almost completely covered with snow may be easier to classify than a patchy scene where snow and snowfree areas frequently change with space. Nevertheless, the analysis of the available scenes did not show a substantial correlation (r^2=0.03) between classification ACC (snow mask at b-6) and snow extent. However, splitting the snow mask into forest and non-forest area leads to an interesting result. It may explain the fact already discussed in section 5.1.2 that scenes before 2003 have a higher ACC with b+3 for forest-covered pixels (than for non-forest). This might be related to snow extent since the analysis shows that for forest pixels it seems: the lower the snow extent is, the higher the resulting ACC tends to be (r=-0.65). This result is reasonable since snow-covered trees generate mixed information (if the trees are not completely snow-covered). The pixels are brighter than a snowfree pixel but darker than a snow-covered pixel. This causes for SPARC a potential for misclassification that is increased with increasing snow extent. For non-forest pixels the opposite is true. The higher the snow extent is, the higher the resulting ACC tends to be (r=0.63). The scenes before 2003 have substantially lower snow extent than the rest of the available scenes and consequently forested areas show even a higher ACC than forest-free areas. In the course of the threshold adaptation to b-6, non-forest areas are corrected more successfully (see Table 5.2) and gained higher ACC values than forest-free areas. The correlation coefficients shift to r=0.80 for uncovered pixels and r=-0.45 for forested areas. However, these dependencies are only relevant when splitting forest and non-forest pixels. As already described above, no correlation between snow extent and ACC is found when analyzing all pixels together (r^2=0.03).
An important remark has to be made for the analysis of forest-covered pixels. As already mentioned, for example, in section 5.1.1 or 5.1.2.1, the low ACC of this land cover type could eventually also be related to the fact that it is mostly located in the transition zone, where differences could simply be based on the distinct treatment of mixed pixels by the two different algorithms. This problem appears again in the following section.

5.2.5 Elevation zone and forest cover

From the results related to forest cover discussed in the previous sections it is expected that the ACC in the critical wooded belt between 1000m and 2000m asl is reduced and correlated with the number of forest-covered pixels. The lower graph of Figure 5.9 shows the

distribution of all forest-covered pixels of the study area by altitude and the upper illustrates the classification ACC for each altitude level for the example scene N17, 03/02/10, 10:28 UTC. At a first glance, the relationship is evident, even though it has to be considered also here

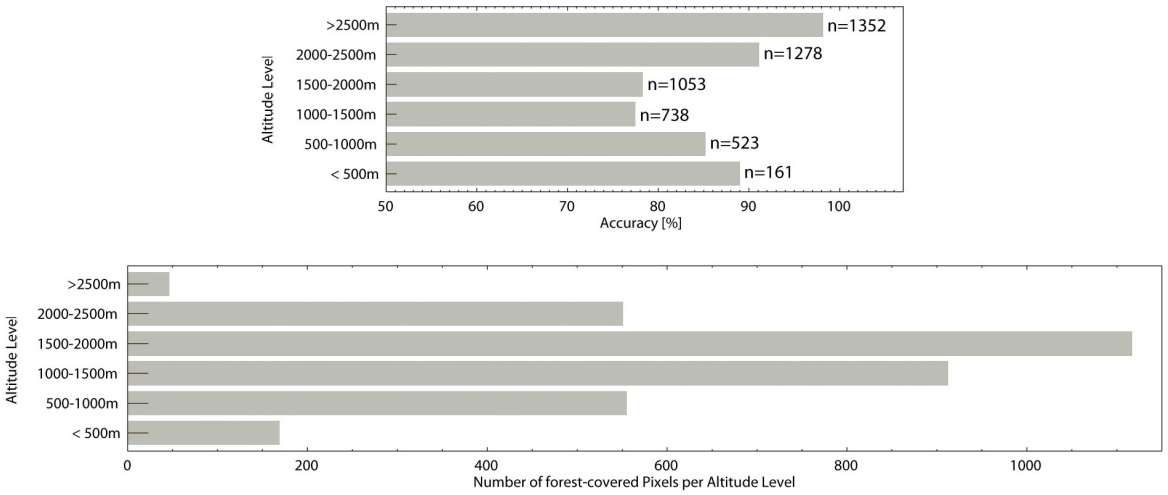


Figure 5.9: SPARC classification ACC per altitude level and corresponding number of pixels for scene N17, 03/02/10, 10:28 UTC (top). Distribution of the number of forest-covered pixels in the study area for same altitude levels (bottom). Snow mask calculated with a B-score of b-6.

that the snow line, i.e. the transition zone, often lies in this zone between 1000m and 2000m asl as well (see following section 5.2.6). A correlation between ACC and number of forest pixels of -0.95 with b+3 is found. Thus, 90% of the ACC variability can be explained by the distribution of the forest pixels with altitude. After the threshold adaptations the correlation between forest cover and ACC per altitude level is reduced from r^2=0.90 to r^2=0.81, which indicates that the negative effect of the forest could be reduced slightly. Nevertheless, the correlation remains high. This is not surprising considering the results displayed in Table 5.2 (ACC results for b-6). It is related to the fact that the gained improvement was much larger for non-forested areas than for forested areas. However, a look at Figure 5.10, which depicts the r^2-values for all scenes, shows that for scenes before 2003 the correlation is higher *after* the threshold shift to b-6. This is a consequence of the fact that scenes before 2003 have higher ACC values with b+3 for forest pixels than for non-forest pixels. This may again be related to the ACC dependency on snow extent (see section 5.2.4). It is obvious that if the forest pixels are better classified than the forest-free pixels the dependency to the number of snow pixels when rising to higher altitudes is smaller.

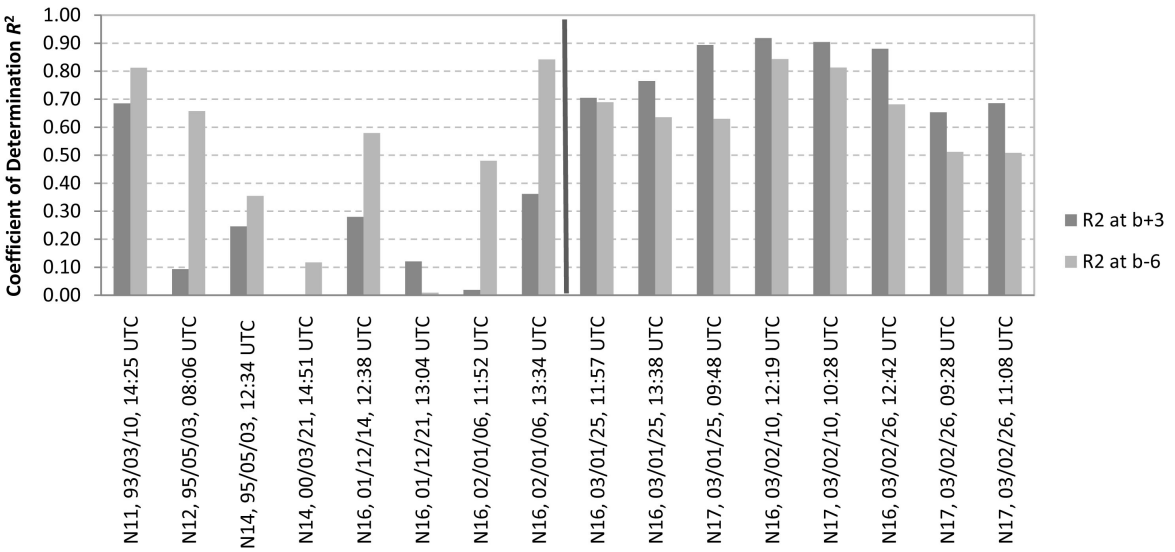


Figure 5.10: Overview of the correlation between forest-covered pixels and ACC per altitude level for all scenes.

5.2.6 ACC in forests related to the snow line/transition zone

The knowledge found in section 5.2.5 of the correlation between forest cover and ACC depending on the elevation zone is crucial, since:

- The position of the snow line influences the ACC of the algorithm due to this dependency. During the peak of winter, the snow line lies below the wooded belt. In spring, when snow cover retreats to the height mountains and less and less pixels are affected by 'problematic' forest cover, the ACC is expected to increase. Related to this, Foppa (2006) found that forest cover becomes a marginal factor influencing the ACC as the snowmelt season advances since the number of snow-covered forest pixels is reduced. However, possibly due to the small number of available scenes, this assumption cannot be confirmed systematically here with the analyzed data set.

- Since SPARC has been implemented by the RSGB mainly for long-term climate applications, the ACC of the snow mask might change with time. Hence, the ACC is influenced not only on a seasonal but also on an interdecadal scale. The expected rise of the snow line as a consequence of climate change (see for example OcCC 2008 or Beniston 2003) will also affect the ACC in the same way as described at the first point. Of course, this is a general problem of remote sensing of snow and is not only related to SPARC.

How these changes will exactly influence the ACC in the future is a challenging question. A more sophisticated analysis of the snow line changes over the last 25 years from other sources compared to the AVHRR snow time series could lead to an indication of the developments occurring in the future. Such sources could be higher resolved satellite images such as MODIS

or a reliable long-term in-situ snow data set, an other example. For MODIS, however, it has to be considered that data is only available from February 2000 to present. Previous work about the topic of snow line estimates from AVHRR data has been done for example by Wunderle et al. (2002). However, it has to be considered that with the coarse spatial resolution of 1.1km only very strong changes in the snow line would become visible. As already discussed in chapter 1, for every 1.0°C increase in winter temperature the snow line rises by 150m (Beniston 2003).

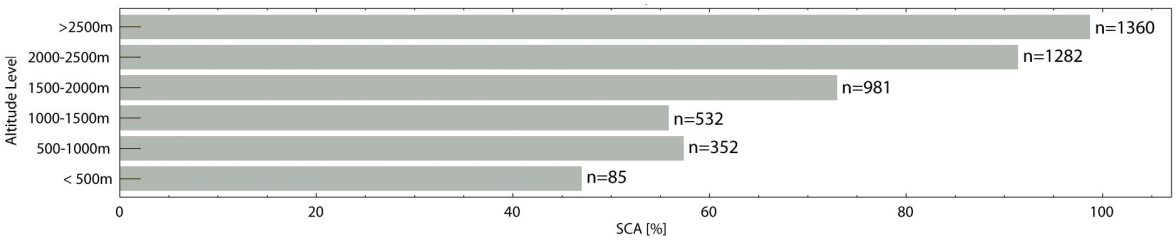


Figure 5.11: Snow-covered area on different altitude levels (y-axis in m asl) for N17, 03/02/10, 10:28 UTC.

Besides the altitude distribution of ACC and forest pixels, the percentage snow cover area for the same altitude levels it plotted in Figure 5.11. The initial aim of this figure was to assess the snow line and with this also the transition zone by using the definition of Seidel et al. (1997), who defined the altitude of the snow line as the belt where 50% snow cover is exceeded. For the example scene it would lie between 500-1000m asl. Unfortunately, the classes of 500m steps seem to be too coarse to make reliable conclusions. Nevertheless, Appendix B, which shows the altitude plots for all scenes, is used to find a possible relationship between the ACC distribution per altitude and the corresponding snow cover. It seems that for the scenes of the year 2003, which have above-average snow extent of roughly $\geq$70% and consequently have more snow in lower elevation zones, the reduced ACC values are located in the lower altitudes, compared to the other scenes. Reduced values seem to more be pronounced between 500-1500m asl rather than between 1500-2000m asl, where scenes before 2003 show more likely reduced values. The example scene (shown in the previous chapter 5.2.5) lies somewhere in the middle.

It has to be considered that this approach can just be seen as an additional indicator of the ACC variability and further research on this topic with numeric results is required. Nevertheless, it emphasizes the assumption that possibly not the forest cover per se but the difficulties emerging from the transition zone (mixed pixels) located at the same altitude levels are the driving factor of the reduced ACC signal in a certain elevation range (see also the Concluding Remarks 5.4 and the Outlook 6.2).

5.2.7 Known snow classification problems from MODIS

When generating snow masks by means of satellite remote sensing systems, further factors apart from those discussed above might influence the results. The ACC differences may occur, for example, due to differences in the sensor system differences itself. Several difficulties are known from MODIS, the state-of-the-art sensor for snow cover mapping. The consideration of obstacles already described in literature (see next paragraph) may also be transferred to the AVHRR system. In the following, such difficulties are discussed, by no means however claiming completeness.

Literature on several MODIS validation studies is available, each reaching different conclusions since the assessment of the ACC is not straightforward. Hall and Riggs (2007) observed that it is not possible to provide one universally valid number for ACC since it depends on several factors such as time, season, land cover and topography. The same conclusions were made in this study using AVHRR data. Every scene has its special characteristics and in combination with the small number of available data it is difficult to find a common pattern. In Hall and Riggs (2007) an overall ACC of around 93% was summarized with lower ACC in forested areas, complex topography and under thin-snow conditions. With respect to specific land cover types, Simic et al. (2004) and earlier also Hall et al. (2001) found the lowest ACC (80%) to occur in areas covered with evergreen forests. This also coincides with the analysis of the AVHRR data carried out in this project. Figure 5.12 depicts the distribution of the misclassified pixels on the the land cover types. Out of 771 misclassified pixels of scene N17,

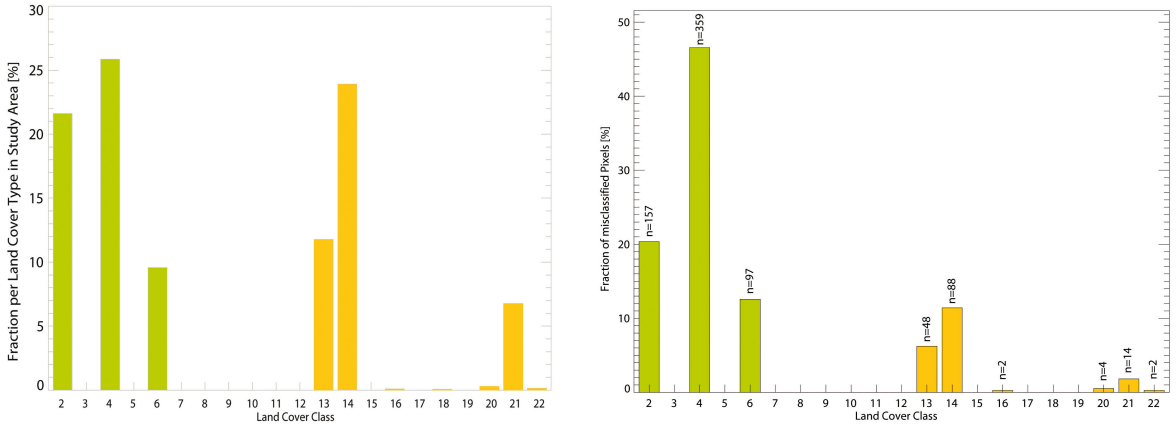


Figure 5.12: Fractions of the present land cover types in the study area (left) and distribution of the misclassified pixels over these land cover types for scene N17, 03/02/10, 10:28 UTC. 2=broadleaved, 4=needle-leaved and 6=mixed-leaved forest cover. 13=Herbaceous cover, 14=Sparse shrub cover, 16=cultivated and managed areas, 20=Water bodies, 21=Snow and ice, 22=Artificial surfaces.

03/02/10, 10:28 UTC 80% are owing to the land cover types 2, 4 and 6 which correspond to forest, although the study area is only roughly 55% forest-covered. 17.6% are related to

herbaceous and shrub cover, which hold roughly 36% of the land cover type fraction in the study area. A negligible amount of misclassified pixels correspond to water bodies, snow and ice and artificial surfaces.

Huang et al. (2011) validated MODIS snow-cover products with Landsat, especially for mountainous areas and obtained ACC values of 95% for plain and 88.1% for mountainous areas. The combination of low snow depth and high forest cover decreased the ACC below 82%. Considering that the study area of this project is located as well in a high mountain region with a large amount of forest cover (more than half of the study subset is covered with forest) this might explain the lower overall ACC of the validation results.

The findings of Huang et al. (2011) and also Simic et al. (2004) with respect to snow depth (see also section 5.2.3) are an interesting point. To recall, they found that most of the disagreement between MODIS and Landsat originates from snow depths lower than 4cm. Although it was not possible to test this with AVHRR, a connection might also exist. Looking back at Figure 5.8, which shows the correlation between date of image acquisition and ACC, leads to the assumption that at the beginning of the winter the snow layer is thin in large areas of the lower altitudes and might decrease the ACC.

Another factor possibly responsible for a fraction of the ACC differences between Landsat and AVHRR (or MODIS and AVHRR) is the problem of the different spatial resolutions. The coarse horizontal resolution of MODIS (500m), compared to Landsat (30m), leads to mixed MODIS pixels in the snow cover edge and patch snow areas (Hall et al. 2002). The binary MODIS algorithm classifies pixels with ∼50% snow cover as snow. Hence, for these mixed pixels, MODIS will overestimate snow cover. This is also observed here during the validation process. Apart from scene 01/12/21, MODIS shows a BIAS above 1.0, which indicates overestimation compared to Landsat. Figure 5.13 illustrates this problem with a patchy Landsat scene (left) and the nearly continuously snow-covered corresponding MODIS snow mask (right), even though it has to be considered that the Landsat image is not resampled to the coarser resolution of MODIS. However, a similar behavior can be observed in the study area

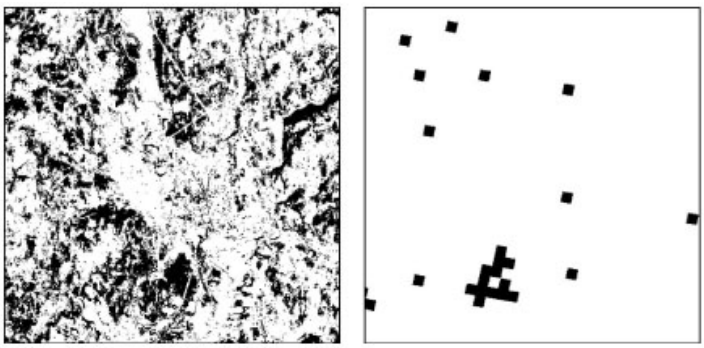


Figure 5.13: Landsat 7 ETM+ snow map (left) and corresponding MODIS snow mask showing an area near Keene, New Hampshire, USA. Both at original spatial resolution, i.e. 30m for Landsat, 500m for MODIS. The coarser resolution of MODIS leads to a high degree of generalization (Hall et al. 2002).

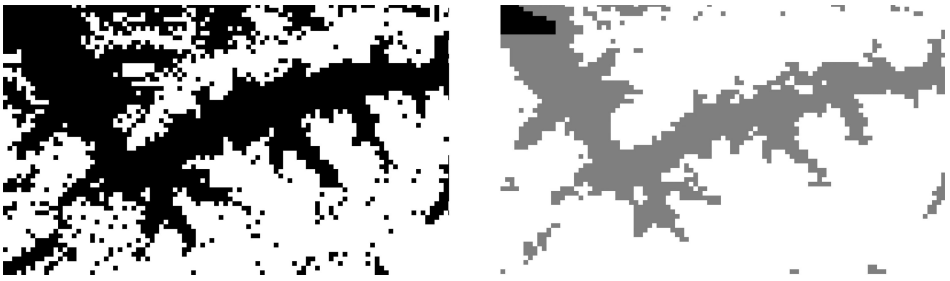


Figure 5.14: Resampled Landsat 5 TM snow map (left) and corresponding N14-AVHRR snow mask (right) showing the Rhone Valley at Martigny, Switzerland on May 3th, 1995. Both at a horizontal resolution of 1.1km. Snow pixels in white, snowfree pixels in black (left) or grey (right). Right black pixels correspond to invalid/cloud pixels. Landsat shows a much patchier snow map, even though the difference to the right AVHRR image is much smaller than in Figure 5.13 since both have same resolutions after the resampling of the Landsat snow mask.

with AVHRR data (see Figure 5.14), where both images have a 1.1km-resolution. The fact that the Landsat snow mask is calculated on the basis of the original 30m-resolution data and resampled *afterwards* to the 1.1km-resolution of AVHRR causes a certain proportion of the mosaic structure to remain. The AVHRR sensor, in contrast, is not able to detect structures on such a fine resolution and generalizes the Earth's surface a priori. Nevertheless, it is difficult to estimate the influence of this feature. The overestimation of the AVHRR snow mask is only observed *after* the threshold adaptation and might be a simple consequence of an 'overcorrection' of the thresholds. Too many other factors could influence the results to make a clear conclusion.

5.3 The Subpixel Scale

Up to here, the whole validation process was based on resampled Landsat data to allow for a pixel-by-pixel comparison. In the following, another approach with 30m-resolution Landsat data and its applicability are discussed.

5.3.1 Snow fraction analysis

Keeping the Landsat resolution at the original 30m and comparing it with the 1.1km-resolution of AVHRR allows for a validation approach based on a subpixel ACC. Nevertheless, it has to be considered that after reprojecting the Landsat image onto the lat/long grid to fit the AVHRR image, its pixels have a non-quadratic shape ($\sim$0.00038° x 0.00027°), which might be a further source of error. On the other hand, the error generated by the resizing of the Landsat pixels to the size of the AVHRR pixels vanishes. With this approach, though, each AVHRR pixel corresponds to a region of interest (ROI) of 27 x 38 Landsat pixels. The result is a fractional snow cover area map, where for every AVHRR pixel a fractional SCA between

0-100% snow cover is calculated from the number of snow-covered Landsat pixels. Figure 5.15 shows such an SCA-map for the N17 scene from February 2003.

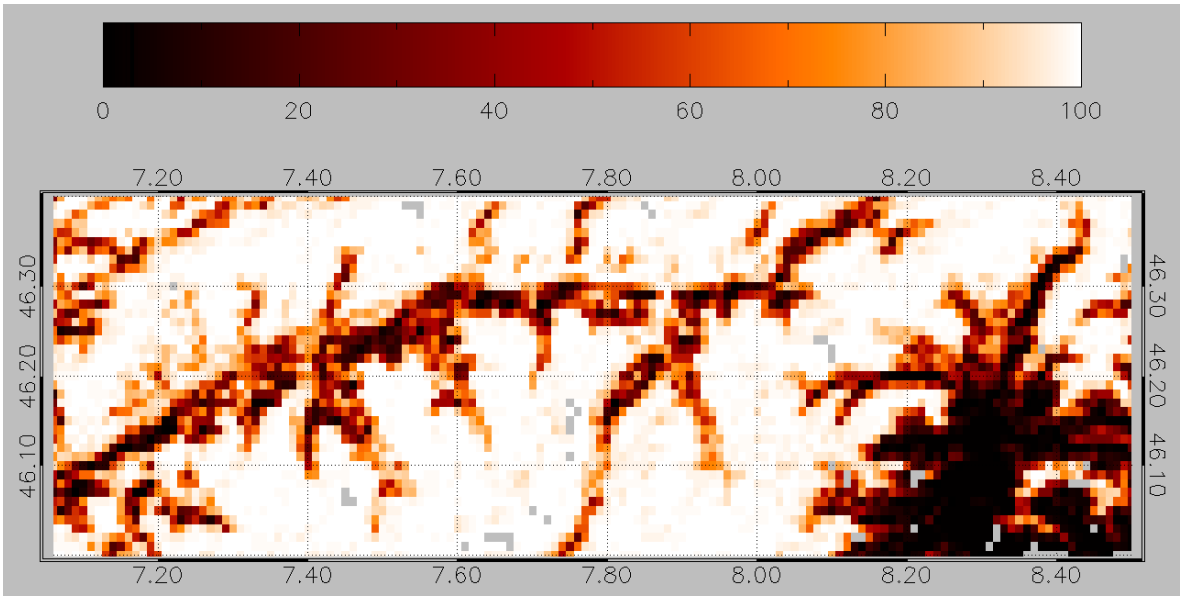


Figure 5.15: Snow cover area (SCA) map of the AVHRR study subset derived from Landsat for scene N17, 03/02/10, 10:28 UTC. The color bar indicates fractional SCA per pixel from 0-100%. Pixels in grey represent invalid data.

Figure 5.16 (top) depicts the snow cover area (SCA) of the AVHRR snow mask pixels derived from Landsat with B-score default b+3. It is assumed that SPARC classifies a pixel as snow if the SCA is greater or equal 50%. The snow pixels are classified nearly perfectly, e.g. 97.92% of all AVHRR pixels classified as snow have a higher SCA than 50%. For snowfree pixels, SPARC yields a much poorer result. Only 50.85% of the pixels classified as snowfree have lower SCA than 50%. Whereas the green bars (snow) follow an ideal logarithmic shape with few pixels below the 50%-threshold bar, the fraction of the yellow bars (snowfree) above 50% is way too high. Nearly 500 pixels which were classified as snowfree have 90-100% SCA according to the Landsat reference. This coincides with the underestimation found in section 5.1.1. Figure 5.16 (bottom) shows the result of the adapted threshold case of b-6. Although the underestimation could be corrected, a certain misclassification remains (see bars at 90-100%). The fraction for the snow pixels is slightly reduced to 94.00% and for the snowfree pixels it is enhanced to 72.76% but remains on a moderate level. 6% of the snow pixels have a snow fraction of less than half the pixel and even 27.24% of the snowfree pixels have more than 50% fractional snow cover. These values correspond to the signal observed in the difference maps, where, despite threshold shifting, a certain misclassification remained. Since this specific scene is roughly 78% snow-covered, the lower values for the snowfree pixels are of minor importance.
However, if we define the best threshold as the B-score where the sum of snow pixels with SCA$\geq$50% and snowfree pixels with SCA$<$50% reaches its maximum, i.e. 200% in an optimal

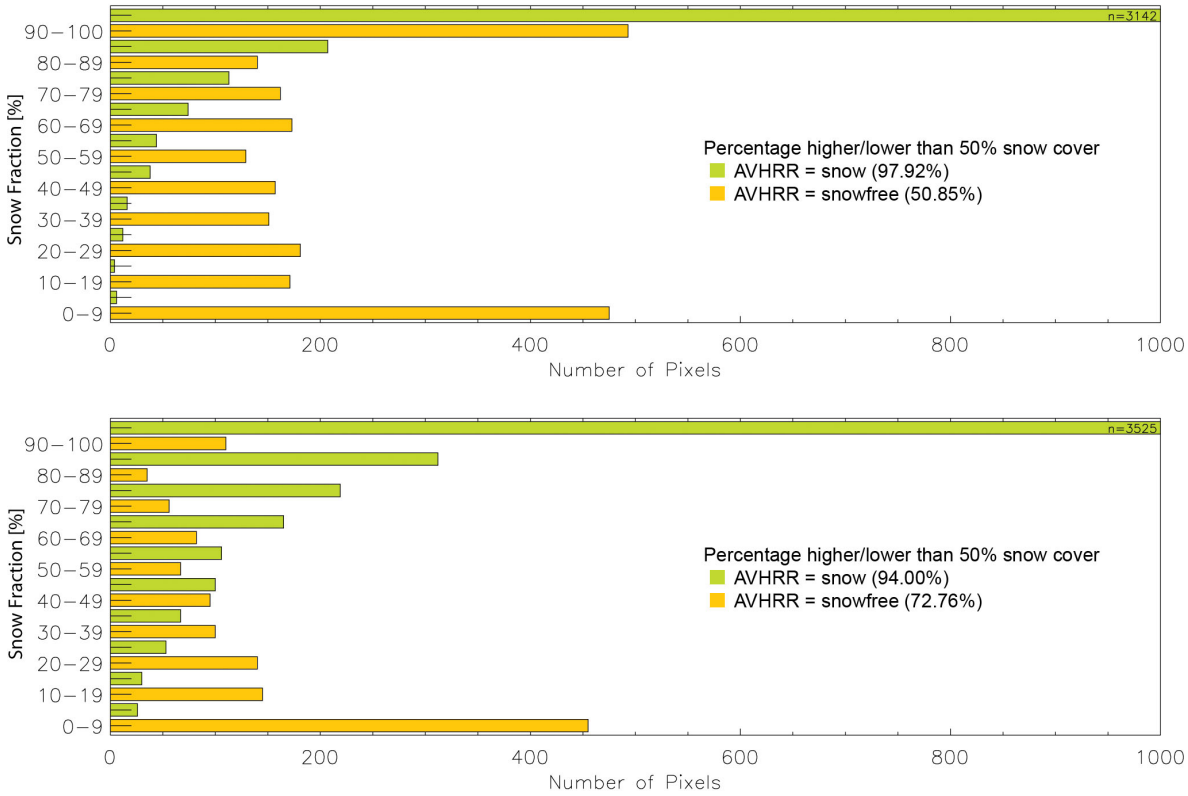


Figure 5.16: Snow cover area (SCA) of the AVHRR snow mask pixels of N17, 03/02/10, 10:28 UTC derived from Landsat with B-score default of b+3 (top) and b-6 (bottom) divided in 10 SCA-classes. Number of pixels on x-axis, fractional SCA per pixel on y-axis.

case, it can be concluded that the optimal B-score threshold value for this scene lies somewhat lower compared to the pixel-by-pixel approach. The sum is usually maximized at the point where the smallest difference is found between correctly classified pixels above or below the SCA-threshold. To recall, in Figure 5.3 the best ACC for this specific scene was found with a B-score of b-7. In contrary to that, applying the SCA-approach with the definition of the sum maximum, best values are observed with b-11 for a 50% SCA-threshold. At b+3 the sum amounts to 148.77%, at b-6 to 166.76% and finally at b-11 to 176.96%, where the values for snow and snowfree pixels amount to 89.24% and 87.73%, respectively. As can be seen, the difference is marginal. Averaged over all scenes, the B-score threshold is in the range of b-8 to b-9. Using only the valid scenes (i.e. no N12, N15, December scenes and satellite angles >40°) it is at b-10. Since the B-score values only range to b-11 it is not verifiable if, for scenes with a maximum at b-11, this actually is the real maximum. Hence, the resulting optimal threshold could even be lower. From the pixel-by-pixel approach it is known that values lower than b-9 lead to an enhanced BIAS (overestimation of snow cover.) Figure 5.17 of the example scene with b-11 shows that now both the green and the yellow curves have the expected logarithmic shape. Nevertheless, the overestimation is present as a large number of snow pixels are below

the 50%-threshold bar. This difference from b-11 to b-6 of the pixel-by-pixel approach (see

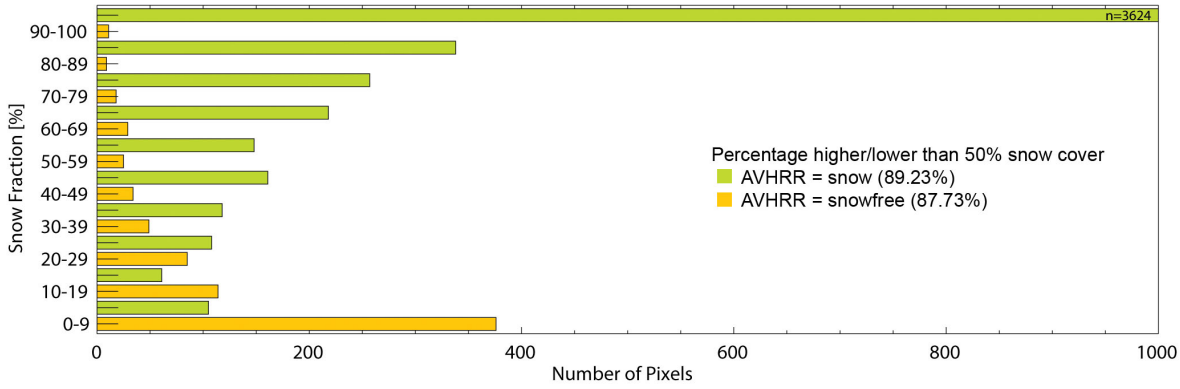


Figure 5.17: Snow cover area (SCA) of the AVHRR snow mask pixels of N17,03/02/10, 10:28 UTC derived from Landsat with a B-score of b-11 divided in 10 SCA-classes. Number of pixels on x-axis, fractional SCA per pixel on y-axis.

section 5.1.2) might be a consequence of the fact that SPARC does not classify snow at the threshold of SCA≥50% but lower or higher. Therefore an additional approach is implemented even though the two approaches can not be compared directly. Instead of keeping the SCA-threshold of 50% constant and changing the B-score, the SCA-value is modified so that the sum of snow pixels with SCA≥x and snowfree pixels with SCA<x are maximized at b-6. By doing this, the value can be derived at which SPARC classifies a pixel as snow-covered. It becomes apparent that this value lies in the mean at 58% SCA with a strong dependency on snow extent. This means that if an AVHRR snow mask is calculated from Landsat, snow should be classified when the pixel exceeds 58% fractional snow cover. Nevertheless, the correlation with snow extent has to be considered. While scenes with snow extent around 50% have best values at a 30-40% SCA-threshold, for scenes with snow extent up to roughly 80% the threshold leading to best results lies at 70% SCA.
Due to this dependency it is not possible to set a reliable SCA-threshold valid for all scenes. Furthermore, it has to be considered that the SPARC algorithm is not best applicable for such snow fraction calculations since the scores are not directly proportional to the snow fraction (Huesler et al. 2012). A more sophisticated approach is required to address this goal. A common method is the linear spectral mixture algorithm which was used, for example, by Foppa et al. (2004) for operational subpixel snow mapping with NOAA AVHRR data. This algorithm is, however, only applicable to sensors with channel 3A, namely N17 and a few scenes of N16. Nevertheless, the presented approach gave an indication of the ACC behavior when keeping the Landsat resolution at 30m and changing either the snow-covered area or the B-score threshold.

5.3.2 An application example

The subpixel approach can also be used for customized snow cover maps. The WSL (Swiss Federal Institute for Forest, Snow and Landscape Research) Institute for Snow and Avalanche Research SLF, for example, is interested in special snow maps where pixels which are proven to be absolutely snowfree are defined. Therefore the critical value is lowered from 50% SCA to 20%. Figure 5.18 shows an example of the resulting map.

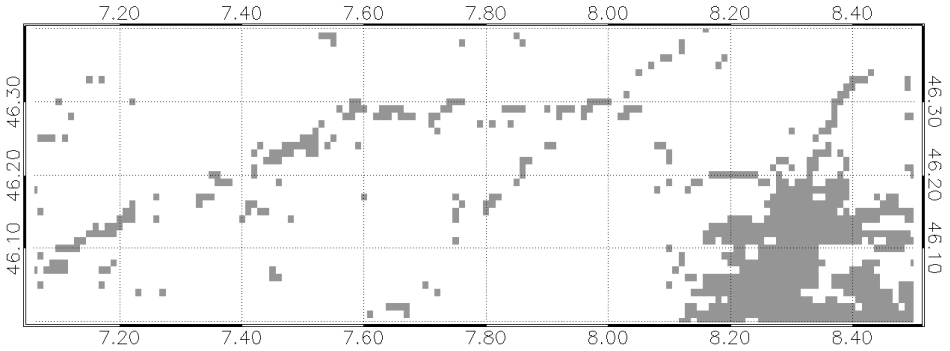


Figure 5.18: Snow cover area map for the study area showing pixels in grey with less than 20% of snow cover.

5.4 Concluding Remarks

During the whole validation process, the small number of available scenes was a limitation and conclusions about influencing factors have to be handled with care. Moreover, the small study area and consequently the small pixel sample size compared to the dimensions of the whole alpine snow time series might have overestimated the effects of certain errors. Such an example is the reduced overall accuracy due to the high fraction of pixels located in forests and the transition zone. Nevertheless, the choice of this specific study area was on purpose since the principal objective of the thesis was to have a closer look directly at the points where the largest deficits of the algorithm are supposed to be located. That both the transition zone and the forest area mainly occur in the same region, further complicates the analysis. Whether the weaknesses in certain elevation ranges are based on a deficit of SPARC or if they are just a consequence of the different algorithms of Landsat and AVHRR by dealing differently with mixed pixels is hard to assess. It is certainly a mixture of both. If the SPARC snow retrieval is validated over the whole Alpine Region, the classification performance is expected to be higher since certain errors, especially those related to the land cover types and transition zones will be smoothed. In addition, the analyzed scenes were, in a sense, randomly picked from the snow time series, since the search criteria were influenced by several different factors, such as for example the Landsat data availability. The fact that, nevertheless, the results coincide to a large extent with the findings of Huesler et al. (2012), supports the quality of

both studies. It is clear that the found "optimal" mean B-score threshold, for instance, is not the best value for every scene as the snow character varies from image to image. However, as already mentioned in the introduction, the improvement process of SPARC is focused on long-term consistency rather than on single scene accuracy. By means of the Results Chapter 5, this objective was proven to be achieved.

Chapter 6

Conclusion and Outlook

6.1 Summary of the Results and Conclusion

In the framework of the validation of the modified Separation of Pixels using Aggregated Rating over Canada (SPARC) snow retrieval the accuracy (=ACC) and the BIAS of the AVHRR snowmask could be quantified. It was shown that the SPARC algorithm underestimates snow cover substantially when running with default values. Analyses of the algorithm at different scores were used to find the thresholds leading to best ACC values. Whereas the R-score had a negligible influence on the classification performance, an optimal B-score threshold was found at b-6, whereby an ACC gain of +9.46% for non-forested areas and +1.94% for forested areas resulting in an overall improvement of +5.21% was achieved. Forest-covered pixels generally revealed lower ACC values than forest-free pixels. The underestimation could be corrected with a slight tendency to overestimation. A comparison with the MODIS snow product (MOD10A1), which is a very robust reference, showed that with this threshold adjustment the AVHRR snow mask almost (i.e. with -1 to -2%) reached the quality of the state-of-the-art sensor system for snow cover mapping.
Despite the modification leading to an improvement of the algorithm, the AVHRR snowmask still did not perfectly coincide with the Landsat snow map. Most of the differences are located in the transition zones (mixed pixels). This is related to the different snow algorithms of AVHRR and Landsat, i.e. the different treatment of mixed pixels, and to forest cover, where snow under trees is hard to detect. Several other factors were found to influence the ACC of the snow mask. However, the small number of analyzed scenes (13 to 17) confuses the issue of drawing conclusions about the quality of the whole time series. Negative correlations with the ACC were present with the scene satellite zenith angle and the sun zenith angle, whereby the influence of the first is much stronger. The ACC was also found to be influenced by the date of scene acquisition. While scenes from December showed low ACC values, the performance increased throughout the winter, which coincides with several studies. Snow extent also seems

to be relevant, at least when analyzing forest and non-forest pixels separately. For forest pixels the lower the snow extent of the scene was, the higher the resulting ACC tended to be. The opposite was true for forest-free areas (positive correlation). Here, the latter influence was much stronger. Moreover, a dependency was found between the altitudinal distribution of the forest pixels and the ACC. The highest number of forest pixels was present where the lowest ACC occurred, namely above 500m asl and below 2000m asl. This result is possibly also affected by the degree of snow extent and the related position of the snow line.
Thus, the dominant factors which affect the SPARC performance could be presented. Nevertheless, the variety of all these factors and the special characteristic of each scene complicate the assessment of the reasons why certain scenes have higher ACC values than others. The most dominant ones are certainly the satellite zenith angle and the classification difficulties in forest and transition zones. It was shown in the course of this thesis that especially the latter two require special consideration on the way to a further increase of the quality of SPARC.
The snow fraction analysis revealed that the presented subpixel approach is not preferable to achieve this objective, especially because the scores are not directly proportional to the snow fraction. However, a rough number of fractional snow cover could be estimated at which SPARC classifies a pixel as snow.
The modified SPARC snow retrieval is particularly developed to study climate change in the European Alps. As for all long-term applications a high level of time consistency is required in order to minimize artificial trends. Therefore it was emphasized that in the course of climate change uncertainties of the algorithm ACC which are related to climate depending factors could change in the future. Such examples are classification weaknesses owing to shallow snow layers in the beginning of the winter or a future rise of the snow line and the transition zone.

Overall, the principal objectives to assess and to quantify the quality of the SPARC classification performance could be achieved. It was shown that the SPARC snow retrieval is well suitable for snow cover mapping in the European Alps. For the first time, the algorithm was tested in a small area with a complex terrain, which unfolds existing weaknesses. This gave the opportunity to study directly how the modification of the score thresholds (mainly the B-score) changes the ACC. Finally, an improvement setting for the algorithm leading to an ACC gain could be presented. This is of particular importance towards a consistent high quality for the aimed 25-year snow time series of the University of Bern. Nevertheless, the following 'remaining' points are recommended to be addressed in order to further improve this time series.

6.2 Outlook

Based on the findings of this thesis, future work in the field of snow mapping with NOAA AVHRR should be focused on the following points:

- The improvements with the modification of the SPARC score thresholds could be shown and the limit of the potential is possibly reached. For snow classification in forest areas, in contrast, the potential of improvement by applying a land cover correction is still present although it is a difficult task and, up to today, no universally implementable solution seems to be available. Nevertheless, the findings of section 5.1.2.1 indicate that by treating forest-covered and forest-free areas differently and analyzing the reflectance behavior of forest in combination with snow cover, the classification ACC can be influenced.

- A separate treatment should not only be made with forest/non-forest areas. In a first step, snow and snowfree pixels should be separated from each other and be divided afterwards into forest/non-forest classes. With this approach more detailed conclusions about how forest pixels contribute to the ACC are expected or at least the results might be better explicable. It is possible, for example, that the ACC of the scene 1995/05/03 is so high because the snow line has already receded to higher altitudes and less snow pixels are located in the wooded belt.

- Several results indicated that most of the differences between the two algorithm products, i.e. the Landsat and the AVHRR snow mask, are located in the transition zone. In the course of future work, it should be verified how SPARC deals exactly with mixed pixel information and if there is a possibility for improvement.

- As discussed in section 5.2, morning satellites (N12 and N15) as well as scenes with large satellite zenith angles are currently excluded from the 25-year snow time series. This leads to a reduction of available scenes and hence to a reduction of robustness of the time series. Efforts to correct the negative effects of these features could allow the use of those scenes. In combination with topographic modeling by means of a digital elevation model and sun angle information (as applied by Huesler et al. 2012), the detection of snow in shadows may be improved and lead to an inclusion of the morning satellites.
  The problem of the large satellite viewing angles should be addressed in a first step by analyzing more scenes. The optimal outcome would be the presence of a systematic error that could be corrected, for example the presence of a constant reduction factor of the ACC depending on the viewing angle.

- More scenes are also required to analyze the ACC behavior with changing sun zenith angles, which is relevant, for example, in the context of satellite orbit drift over the

lifetime of a sensor. The results of section 5.2.2 emphasize the need of additional scenes. If it turns out that the ACC is actually depending on the sun angles, this may be corrected with a Bidirectional Reflectance Distribution Function (BRDF), a function of illumination and viewing geometry which is based on the fact that objects look differently when viewing from different angles and illuminated from different directions (Luo et al. 2008). Certainly this is a challenging task in the complex topography of the Alps.

- Furthermore, Huesler et al. (2012) have undertaken efforts to address snow classification weaknesses in shadow areas. A validation of these areas was deliberately excluded from this thesis since even the Landsat reference snow mask did not perfectly represent snow in shadows. Also here, shadow modeling could help improve the reference image, which could then be used for specific shadow validation. This could for example be useful for the morning satellites already discussed above.

- It was shown by Simic et al. (2004) and Huang et al. (2011) that the snow classification performance depends on snow depth and improves with a thickening of the snow layer. A similar development is observed here, although it has not been proven that snow depth actually is a driving factor. However, if it were so, then the SPARC algorithm should be combined with reliable snow depth data. This allows the fitting of the algorithm to the actual snow conditions, which is expected to lead to an ACC improvement of the snow mask. The lacking representativeness of ground station snow depth data for the area of an AVHRR pixel is certainly a challenge.

- In the framework of this master's thesis it was not possible to assess the classification error reaching from poor cloud masking. The fact that only clear-sky Landsat scenes were used and AVHRR cloud pixels were also not considered, reduces the effect of poor cloud classification. Only AVHRR cloud pixels misclassified as snow could not be captured. A project for the replacement of the poorer performing CASPR cloud mask with the SPARC cloud mask is currently in progress at the RSGB and will contribute to a further increase of the data quality.

- Not least, the presented geolocational shift (see section 4.2.2) has to be investigated. The fact that the MODIS snow mask had to be corrected as well, suggests that the error does not necessarily need to be located at the original AVHRR image but also at the Landsat reference. However, special attention should be payed to this since the whole AVHRR data archive and its applications will be affected.

The increasing interest in remote sensing techniques for long-term studies will continue to grow as remote sensing develops further. The validation study presented here is a contribution to the 25-year snow time series of the European Alps derived from historical AVHRR data. It is highly desirable that this series is continued in the future with the ongoing AVHRR mission since the enormous potential of such a data set for climate change studies is evident.

References

Appenzeller, C., M. Beyert, E. Zenklusen, and S. C. Scherrer (2008). Monitoring climate at Jungfraujoch in the high Swiss Alpine region. *Science of the total Environment 391*(2-3), 262–268. International Conference Celebrating the 75th Anniversary of the High Altitude Research Station Jungfraujoch, Interlaken, Switzerland, Sep 11-13, 2006.

Armstrong, R. and M. Brodzik (2001). Recent Northern Hemisphere snow extent: A comparison of data derived from visible and microwave satellite sensors. *Geophysical Research Letters 28*(19), 3673–3676.

Begert, M., T. Schlegel, and W. Kirchhofer (2005). Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000. *International Journal of Climatology 25*(1), 65–80.

Beniston, M. (2003). Climatic change in mountain regions: A review of possible impacts. *Climatic Change 59*(1-2), 5–31.

Chander, G., B. L. Markham, and D. L. Helder (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment 113*(5), 893–903.

Cracknell, A. P. (1997). *The Advanced Very High Resolution Radiometer (AVHRR)*. London: Taylor.

Elachi, C. and J. Van Zyl (2006). *Introduction to the physics and techniques of remote sensing. Second edition*. Wiley-Interscience.

EURIMAGE (accessed, 2011). Landsat - Products and Services. Decades of continuous medium-resolution data. Technical report, EURIMAGE - Multimission Satellite Data. www.eurimage.com.

Foppa, N. (2006). *Improved methods for operational remote sensing of snow cover using NOAA AVHRR data - exemplified for the European Alps*. Ph. D. thesis, Philosphisch-naturwissenschaftliche Fakultät der Universität Bern - Geographisches Institut.

Foppa, N., A. Hauser, D. Oesch, S. Wunderle, and R. Meister (2007). Validation of operational AVHRR subpixel snow retrievals over the European Alps based on ASTER data. *International Journal of Remote Sensing 28*(21), 4841–4865.

Foppa, N., S. Wunderle, A. Hauser, D. Oesch, and F. Kuchen (2004). Operational sub-pixel snow mapping over the Alps with NOAA AVHRR data. In Fohn, PMB (Ed.), *Annals of Glaciology*, Volume 38, pp. 245–252. International Symposium on Snow and Avalanches, Davos, Switzerland, Jun 02-06, 2003.

GCOS (2006). Systematic observation requirements for satellite-based products for climate. Supplemental details to the satellite-based component of the implementation plan for the global observing system for climate in support of the UNFCCC. In *GCOS-107, 107, Available online at: www.wmo.int/pages/prog/gcos/publications (accessed May 24, 2009)*.

Goodrum, G., K. Kidwell, and W. Winston (2006). NOAA KLM user's guide. http://www.ncdc.noaa.gov/oa/pod-guide/ncdc/docs/intro.htm. Technical report, National Environmental Satellite, Data, and Information Service (NESDIS).

Gustafsson, D., M. Stahli, and P. Jansson (2001). The surface energy balance of a snow cover: comparing measurements to two different simulation models. *Theoretical and applied Climatology 70*(1-4), 81–96.

Hall, D., J. Foster, V. Salomonson, A. Klein, and J. Chien (2001). Development of a technique to assess snow-cover mapping errors from space. *IEEE Transactions on Geoscience and Remote Sensing 39*(2), 432–438.

Hall, D., J. Foster, D. Verbyla, A. Klein, and C. Benson (1998). Assessment of snow-cover mapping accuracy in a variety of vegetation-cover densities in central Alaska. *Remote Sensing of Environment 66*(2), 129–137.

Hall, D., G. Riggs, V. Salomonson, N. DiGirolamo, and K. Bayr (2002). MODIS snow-cover products. *Remote Sensing of Environment 83*(1-2), 181–194.

Hall, D. K. and G. A. Riggs (2007). Accuracy assessment of the MODIS snow products. *Hydrological Processes 21*(12), 1534–1547. 63rd Eastern Snow Conference, Newark, DE, Jun 07-09, 2006.

Hartley, A., J. Pekel, L. Ledwith, J.-L. Champeaux, E. De Badts, and S. Bartalev (2006). Glc2000 database. Technical report, European Commision Joint Research Centre.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research-Atmospheres 113*(D20), 12.

Huang, X., T. Liang, X. Zhang, and Z. Guo (2011). Validation of MODIS snow cover products using Landsat and ground measurements during the 2001-2005 snow seasons over northern Xinjiang, China. *International Journal of Remote Sensing 32*(1), 133–152.

Huesler, F., F. Fontana, C. Neuhaus, J. Musial, S. Wunderle, and M. Riffler (submitted, 2011). AVHRR and processing facility at the University of Bern. A comprehensive 1-km satellite data set for climate change studies. *EARSeL eproceedings*.

Huesler, F., T. Jonas, S. Albrecht, and S. Wunderle (in prep., 2012). Validation of a snow cover retrieval from historical AVHRR data over the European Alps. *Journal of Geophysical Research*.

IPCC (2007a). *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, Eds. Cambridge University Press, Cambridge, UK, 976pp.

IPCC (2007b). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.). Cambridge University Press, Cambridge, UK, 976pp.

IPCC (2007c). Summary for policymakers. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Khlopenkov, K. V. and A. P. Trishchenko (2007). SPARC: New cloud, snow, and cloud shadow detection scheme for historical 1-km AVHRR data over Canada. *Journal of Atmospheric and Oceanic Technology 24*(3), 322–343.

Kidwell, K. (1998). *NOAA polar orbiter data user's guide (TIROS-N, NOAA-6, -7, -8, -9,-10, -11, -12, -13 and -14).*

Klein, A., D. Hall, and G. Riggs (1998). Improving snow cover mapping in forests through the use of a canopy reflectance model. *Hydrological Processes 12*(10-11), 1723–1744.

Lillesand, T., R. Kiefer, and J. Chipman (2004). *Remote Sensing and Image Interpretation. Fifth edition.* John Wiley & Sons, Ltd.

Luo, Y., A. P. Trishchenko, and K. V. Khlopenkov (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment 112*(12), 4167–4185.

Marinucci, M., F. Giorgi, M. Beniston, M. Wild, P. Tschuck, A. Ohmura, and A. Bernasconi (1995). High-resolution simulations of January and July climate over the Western Alpine Region with a nested regional modeling system. *Theoretical and applied Climatology 51*(3), 119–138.

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella (2005). The COSMO-LEPS mesoscale ensemble system: validation of the methodology and verification. *Nonlinear Processes in Geophysics 12*(4), 527–536.

Maurer, E., J. Rhoads, R. Dubayah, and D. Lettenmaier (2003). Evaluation of the snow-covered area data product from MODIS. *Hydrological Processes 17*(1), 59–71.

OcCC (2008). Das Klima ändert - was nun? Der neue UN-Klimabericht (IPCC 2007) und die wichtigsten Ergebnisse aus Sicht der Schweiz. Technical report, OcCC - Organe consultatif sur les changements climatiques, Bern.

Parajka, J. and G. Bloeschl (2006). Validation of MODIS snow cover images over Austria. *Hydrology and Earth System Sciences 10*(5), 679–689.

Rees, W. (2006). *Remote Sensing of Snow and Ice*. Taylor & Francis Group.

Robel, J. (2009). *NOAA KLM User's Guide with NOAA-N,-P Supplement.*

Romanov, P., G. Gutman, and I. Csiszar (2002). Satellite-derived snow cover maps for North America: Accuracy assessment. In J. Fellous, J. LeMarshall, B. Choudhury, M. Menenti, L. Paxton, and S. Gupta (Eds.), *Earth's Atmosphere, Ocean and Surface Studies*, Number 11 in Advances in Space Research, pp. 2455–2460. A0 1, A0 4-A0 5, A3 1, C1 1 and C2 6 Symposia of the COSPAR Scientific Commissions A and C held at the 33rd COSPAR Scientific Assembly, Warsaw, Poland, Jul, 2000.

Salminen, M., J. Pulliainen, S. Metsamaki, A. Kontu, and H. Suokanerva (2009). The behaviour of snow and snow-free surface reflectance in boreal forests: Implications to the performance of snow covered area monitoring. *Remote Sensing of Environment 113*(5), 907–918.

Scherrer, S., C. Appenzeller, and M. Laternser (2004). Trends in Swiss Alpine snow days: The role of local- and large-scale climate variability. *Geophysical Research Letters 31*(13), 1–4.

Seidel, K., C. Ehrler, J. Martinec, and O. Turpin (1997). Derivation of statistical snow line from high resolution snow cover mapping. In Wunderle, S (Ed.), *Proceedings of the EARSeL Workshop on Remote Sensing of Land Ice and Snow*, pp. 31–36. EARSeL Workshop on Remote Sensing of Land Ice and Snow, Univ Freiburg, Freiburg, Germany, Apr 17-18, 1997.

Siljamo, N. and O. Hyvarinen (2011). New Geostationary Satellite-Based Snow-Cover Algorithm. *Journal of applied Meteorology and Climatology 50*(6), 1275–1290.

Simic, A., R. Fernandes, R. Brown, P. Romanov, and W. Park (2004). Validation of VEGETATION, MODIS, and GOES plus SSM/I snow-cover products over Canada based on surface snow depth observations. *Hydrological Processes 18*(6), 1089–1104. 60th Eastern Snow Conference (ESC), Sherbrooke, Canada, 2003.

UNEP (2007). Snow. In *Global Outlook for Ice and Snow*. Barry, R.G., Armstrong, R., Callaghan, T., Cherry, J., Gearheard, S., Nolin, A., Russell, D., Zöckler, Chr.

Wunderle, S., M. Droz, and H. Kleindienst (2002). Spatial and temporal analysis of the snow line in the Alps based on NOAA-AVHRR data. In *Geographica Helvetica*, Volume 3 (Jg. 57). Verband Geographie Schweiz (ASG) & Geographisch-Ethnographische Gesellschaft Zürich (GEGZ).

Zappa, M. (2008). Objective quantitative spatial verification of distributed snow cover simulations - an experiment for the whole of Switzerland. *Hydrological Sciences Journal 53*(1), 179–191.

Appendix A

Difference Maps

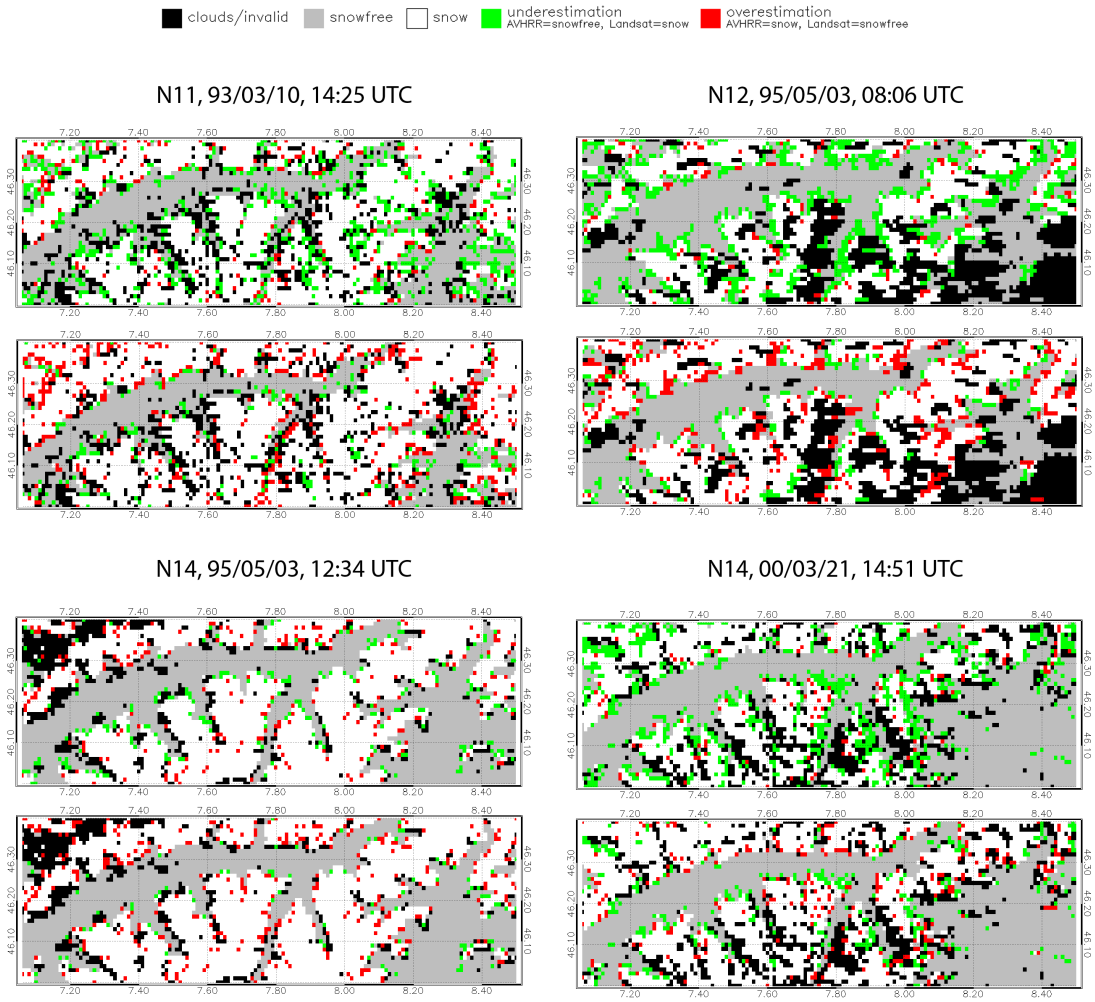
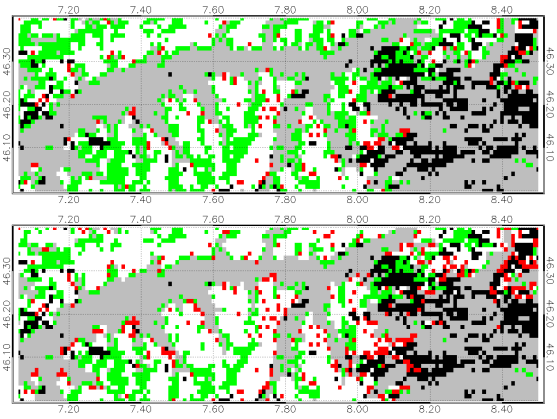
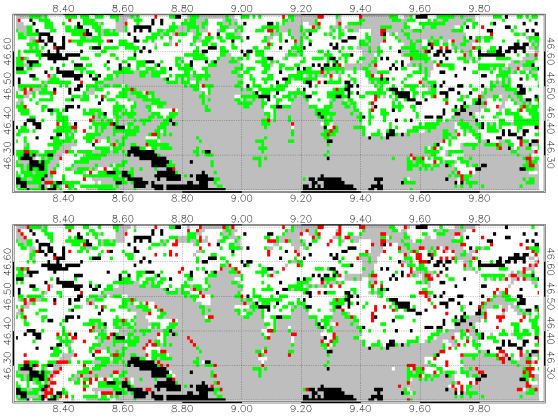


Figure A.1: Difference maps for B-scores b+3 (top) and b-6 (bottom) referred to the resampled Landsat reference.

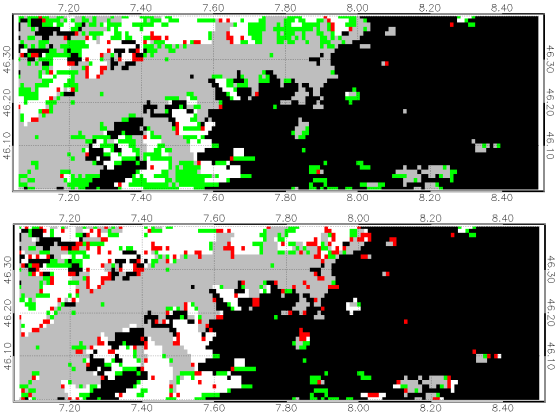
N15, 00/03/21, 07:38 UTC



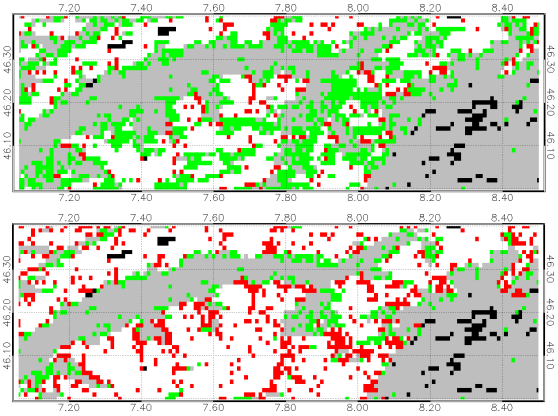
N16, 01/12/14, 12:38 UTC



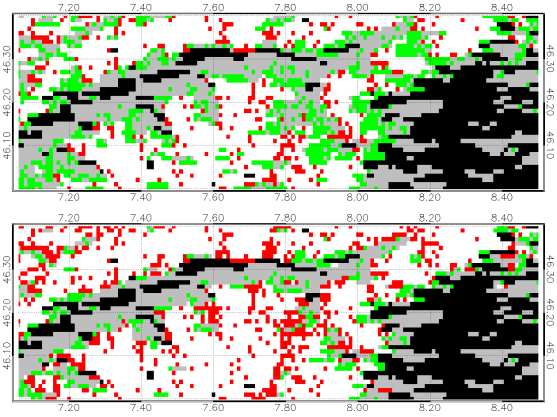
N16, 01/12/21, 13:04 UTC



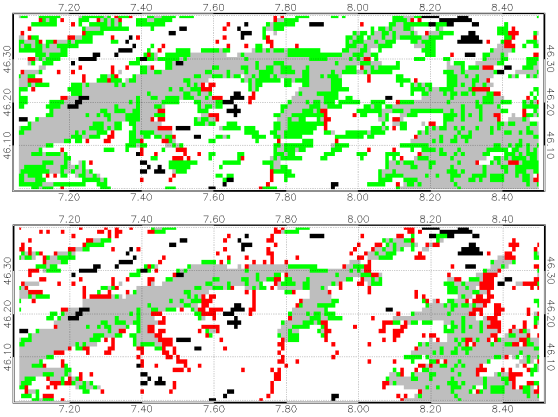
N16, 02/01/06, 11:52 UTC



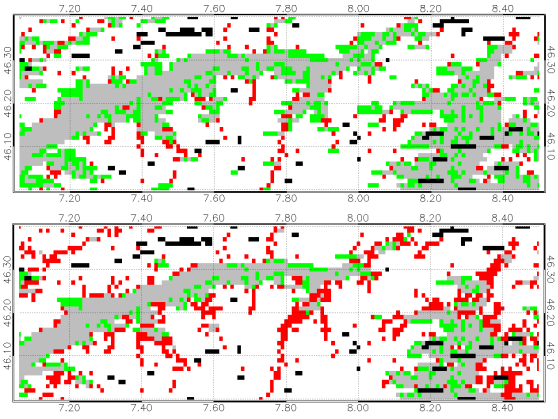
N16, 02/01/06, 13:34 UTC



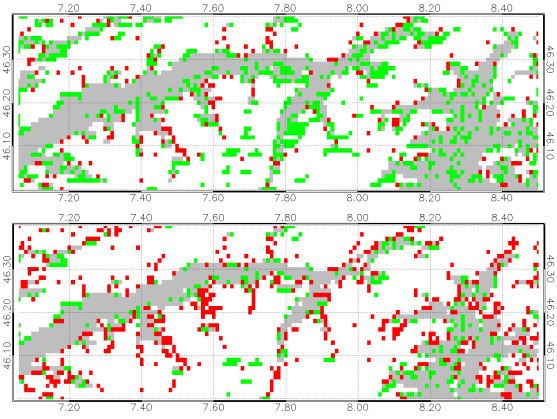
N16, 03/01/25, 11:57 UTC



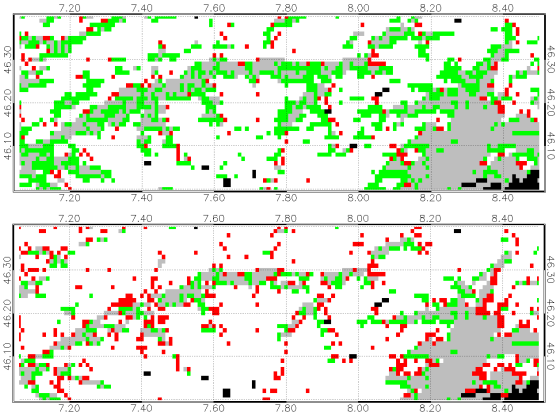
N16, 03/01/25, 13:38 UTC



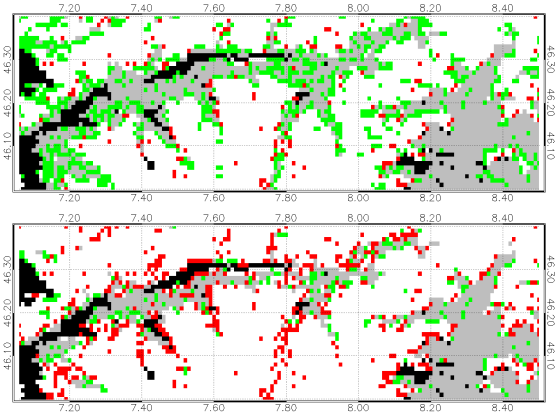
N17, 03/01/25, 09:48 UTC



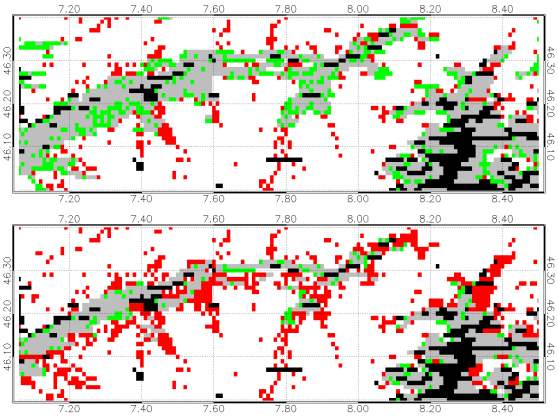
N16, 03/02/10, 12:19 UTC



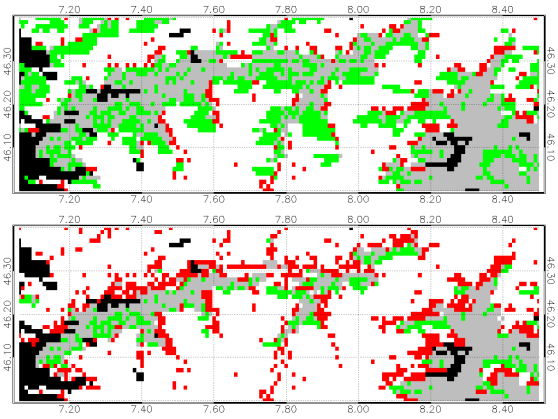
N16, 03/02/26, 12:42 UTC



N17, 03/02/26, 09:28 UTC



N17, 03/02/26, 11:08 UTC



Appendix B

Altitude Dependency of ACC

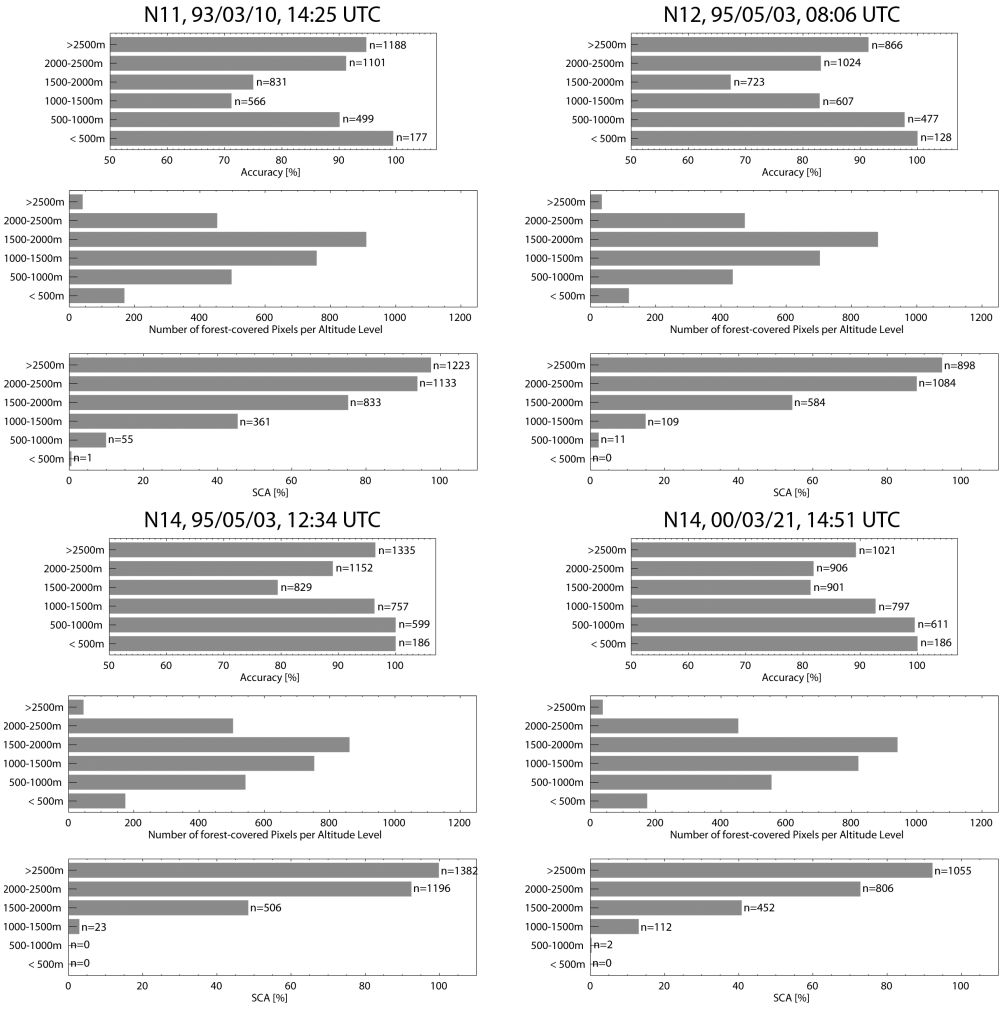
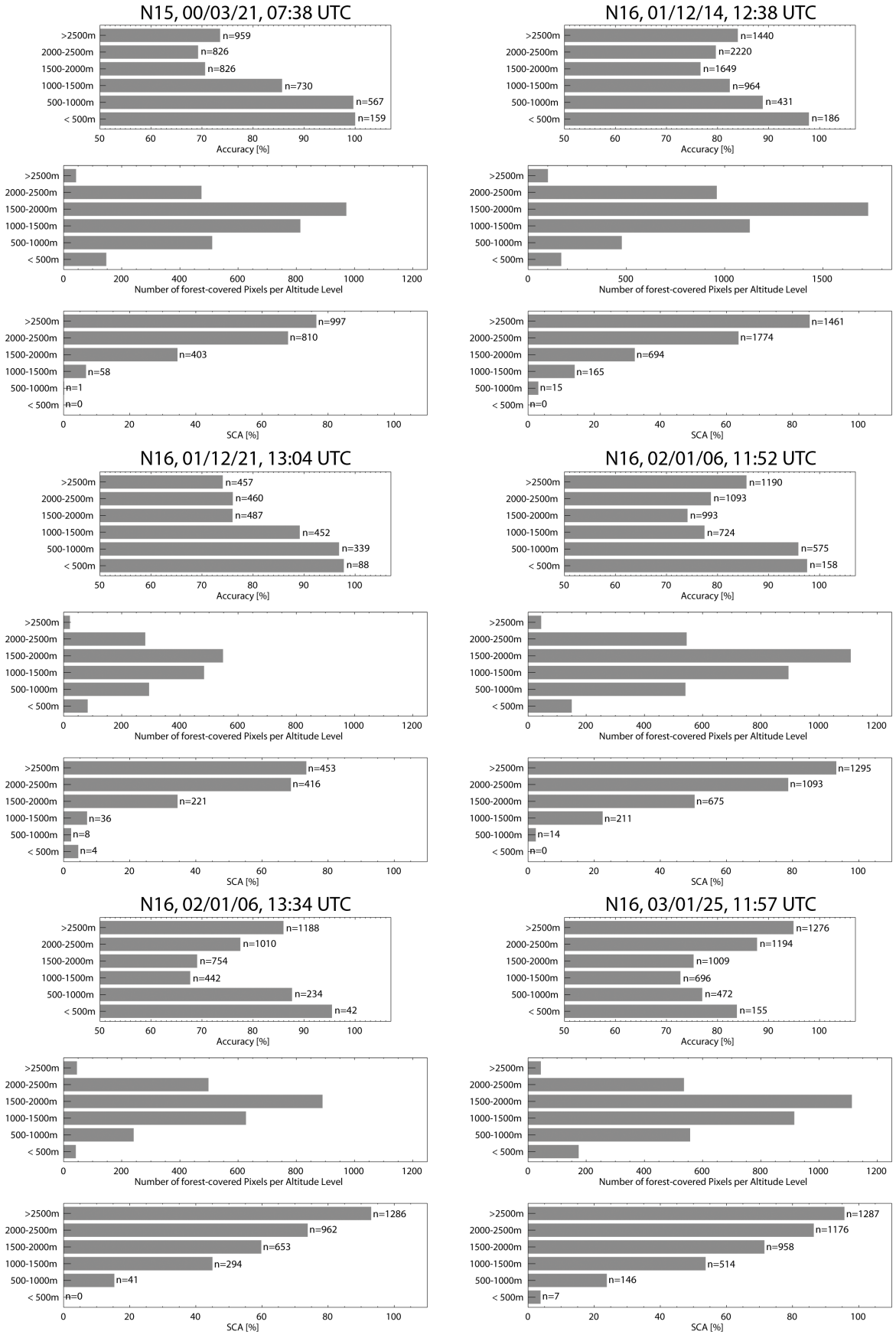
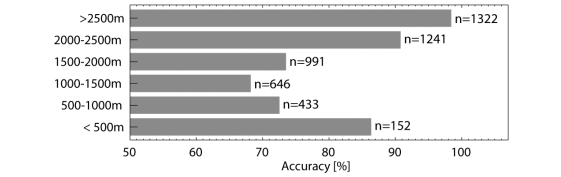
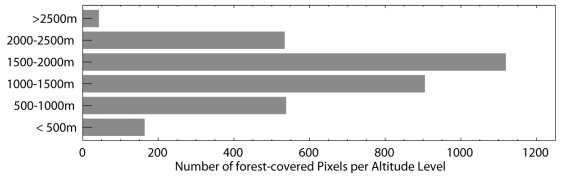


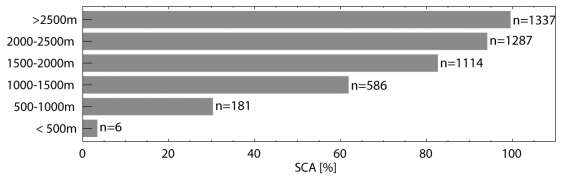
Figure B.1: SPARC classification ACC per altitude level (y-axis in m asl) and corresponding number of pixels (top). Distribution of the number of forest-covered pixels in the study area (middle) and snow covered area (bottom) for same altitude levels.


N15, 00/03/21, 07:38 UTC
>2500m n=959
2000-2500m n=826
1500-2000m n=826
1000-1500m n=730
500-1000m n=567
< 500m n=159
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=997
2000-2500m n=810
1500-2000m n=403
1000-1500m n=58
500-1000m n=1
< 500m n=0
SCA [%]
N16, 01/12/14, 12:38 UTC
>2500m n=1440
2000-2500m n=2220
1500-2000m n=1649
1000-1500m n=964
500-1000m n=431
< 500m n=186
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=1461
2000-2500m n=1774
1500-2000m n=694
1000-1500m n=165
500-1000m n=15
< 500m n=0
SCA [%]
N16, 01/12/21, 13:04 UTC
>2500m n=457
2000-2500m n=460
1500-2000m n=487
1000-1500m n=452
500-1000m n=339
< 500m n=88
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=453
2000-2500m n=416
1500-2000m n=221
1000-1500m n=36
500-1000m n=8
< 500m n=4
SCA [%]
N16, 02/01/06, 11:52 UTC
>2500m n=1190
2000-2500m n=1093
1500-2000m n=993
1000-1500m n=724
500-1000m n=575
< 500m n=158
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=1295
2000-2500m n=1093
1500-2000m n=675
1000-1500m n=211
500-1000m n=14
< 500m n=0
SCA [%]
N16, 02/01/06, 13:34 UTC
>2500m n=1188
2000-2500m n=1010
1500-2000m n=754
1000-1500m n=442
500-1000m n=234
< 500m n=42
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=1286
2000-2500m n=962
1500-2000m n=653
1000-1500m n=294
500-1000m n=41
< 500m n=0
SCA [%]
N16, 03/01/25, 11:57 UTC
>2500m n=1276
2000-2500m n=1194
1500-2000m n=1009
1000-1500m n=696
500-1000m n=472
< 500m n=155
Accuracy [%]
Number of forest-covered Pixels per Altitude Level
>2500m n=1287
2000-2500m n=1176
1500-2000m n=958
1000-1500m n=514
500-1000m n=146
< 500m n=7
SCA [%]

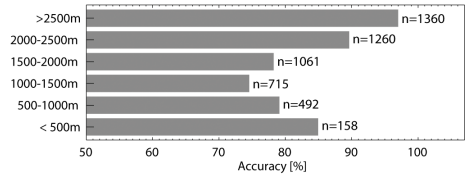
N16, 03/01/25, 13:38 UTC

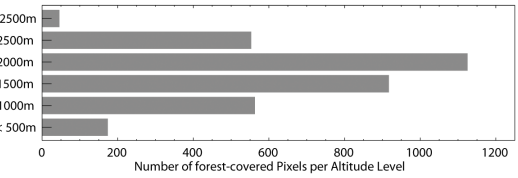


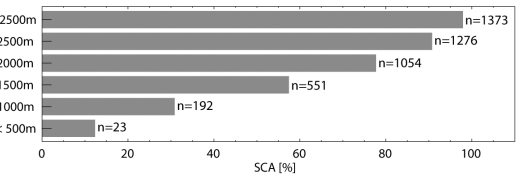




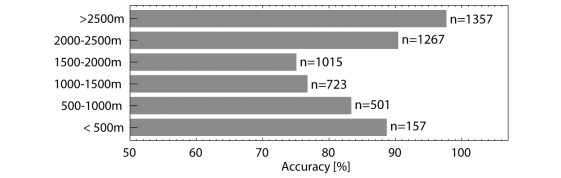
N17, 03/01/25, 09:48 UTC

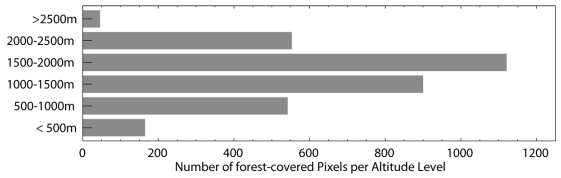


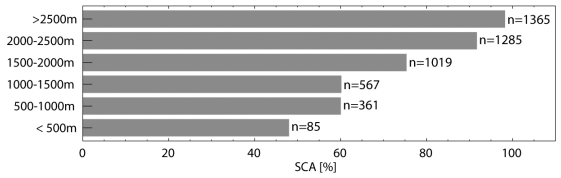




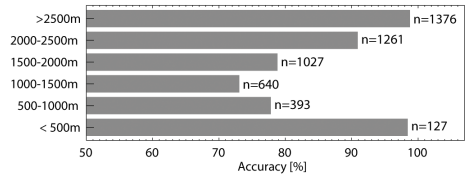
N16, 03/02/10, 12:19 UTC

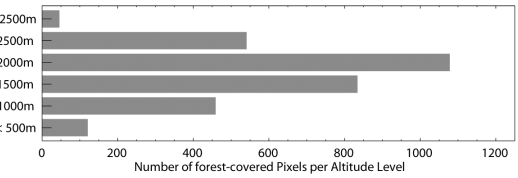


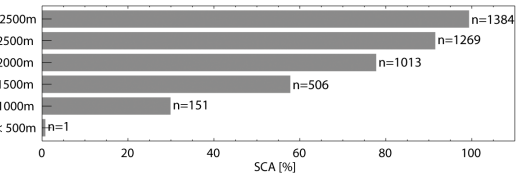




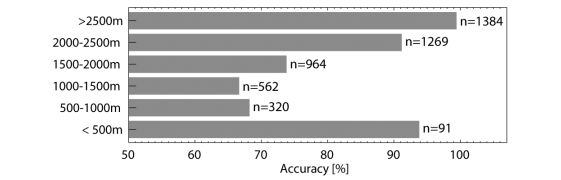
N16, 03/02/26, 12:42 UTC

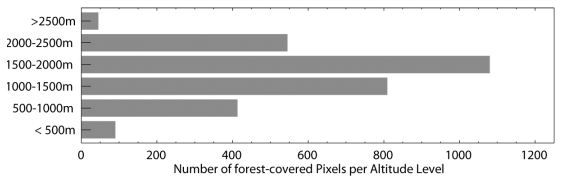


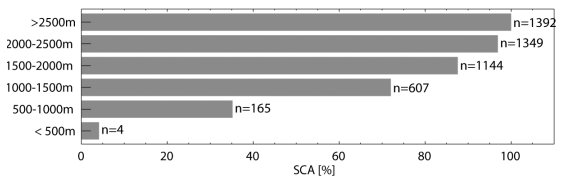




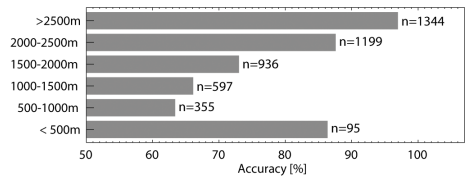
N17, 03/02/26, 09:28 UTC

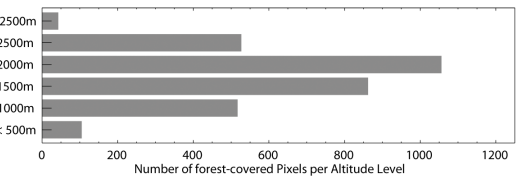


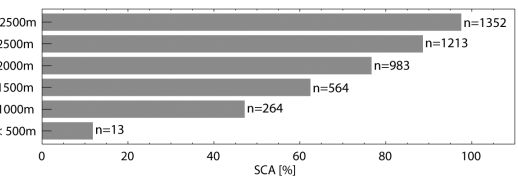




N17, 03/02/26, 11:08 UTC







Appendix C

IDL Routines

Pre-processing

- **landsat_reflectance.pro** Converts the Landsat data from digital number to top-of-atmosphere reflectance values according to Chander et al. (2009). For data ordered from USGS.
- **landsat_reflectance_npoc.pro** Same as landsat_reflectance.pro but for data ordered via NPOC since the required conversion parameters are read out differently.
- **landsat_ndsi.pro** The routine for Landsat snow mask calculation based on NDSI/NDVI approach according to Hall et al. (1998).
- **comparing_landsat.pro** Comparison of the snow mask generated by means of supervised classification (in ENVI) with the snow mask calculated by means of NDSI/NDVI-approach.
- **findshift.pro** Used to find the dimensions of geolocational shift.
- **read_geometry.pro** Reads the scene geometry (satellite zenith and azimuth, sun zenith and azimuth) out of the AVHRR geometry files *scene*_geom.ers, which is the basis for the geometry correlation procedures.

Analysis

- **validation.pro** The core of the master's thesis validation including all relevant steps for the ACC and BIAS results and plotting procedures.
- **validation_nonres.pro** A routine to perform the calculations for subpixel approach using non-resampled 30m-resolution Landsat data.
- **datecorrel.pro** Calculation and plot of the dependency between date of scene acquisition and SPARC classification ACC.

- **forestcorrelation.pro** Calculates the dependency between number of forest pixels and SPARC classification ACC for different elevation zones.
- **geometrycorrelation.pro** Calculates and plots the dependencies between satellite as well as sun zenith angle and ACC.

Declaration

under Art. 28 Para. 2 RSL 05

Last, first name:	Albrecht Simon
Matriculation number:	05-057-658
Programme:	Master of Science in Climate Sciences
	Bachelor ☐ Master ✕ Dissertation ☐
Thesis title:	Validation of the modified SPARC snow retrieval from historical NOAA AVHRR data over the European Alps based on Landsat TM images
Thesis supervisor:	Dr. Stefan Wunderle

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person, except where due acknowledgement has been made in the text. In accordance with academic rules and ethical conduct, I have fully cited and referenced all material and results that are not original to this work. I am well aware of the fact that, on the basis of Article 36 Paragraph 1 Letter o of the University Law of 5 September 1996, the Senate is entitled to deny the title awarded on the basis of this work if proven otherwise.

Bern, 18.08.2011
Place, date

.....
Simon Albrecht