Antarstick: Extracting Snow Height From Time-Lapse Photography
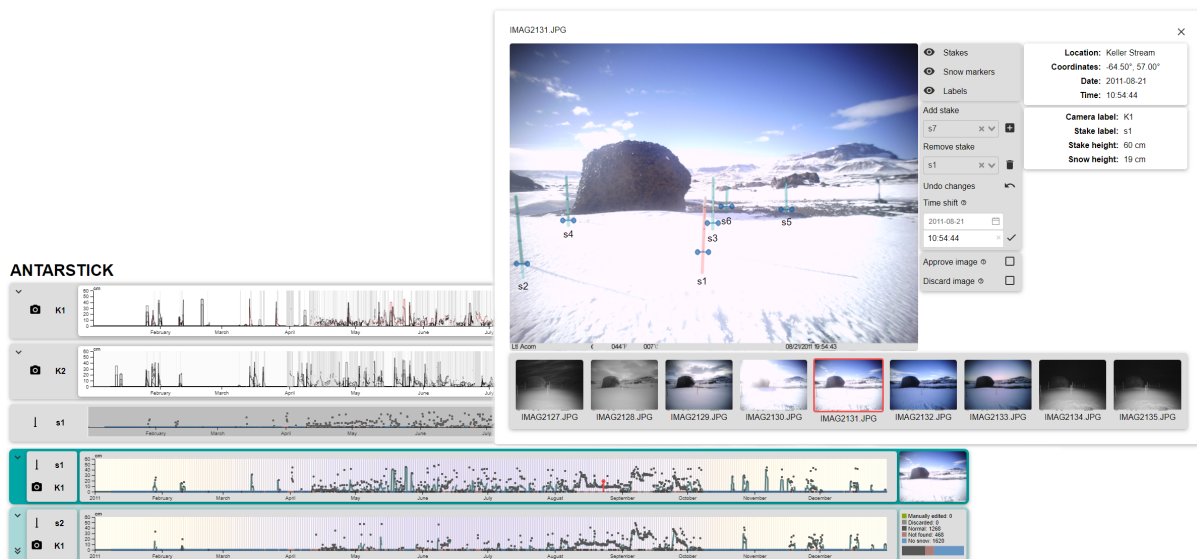
Matěj Lang ⓘ , Radoslav Mráz, Marek Trtík, Sergej Stoppel, Jan Byška ⓘ , and Barbora Kozlíková ⓘ

Figure 1: *An overview of the Antarstick tool, featuring the view of an example dataset. In the forefront, there is a window, where the user can edit the measured snow height values and see them in context. The background has aggregated views of all snow stakes in the area for easy navigation and comparison.*

Abstract
The evolution and accumulation of snow cover are among the most important characteristics influencing Antarctica's climate and biotopes. The changes in Antarctica are also substantially impacting global climate change. Therefore, detailed monitoring of snow evolution is key to understanding such changes. One way to conduct this monitoring is by installing trail cameras in a particular region and then processing the captured information. This option is affordable, but has some drawbacks, such as the fully automatic solution for the extraction of snow height from these images is not feasible. Therefore, it still requires human intervention, manually correcting the inaccurately extracted information. In this paper, we present Antarstick, a tool for visual guidance of the user to potentially wrong values extracted from poor-quality images and support for their interactive correction. This tool allows for much quicker and semi-automated processing of snow height from time-lapse photography.

CCS Concepts
• Human-centered computing → Visualization toolkits; Geographic visualization;

1. Introduction

Antarctica is one of the most important sites on our planet with respect to climate change. The amount of fallen snow is an important factor in the local hydrology, influencing the total amount of water present in the soil. Several peripheral areas of Antarctica are not formed by permafrost, and in summer, the snow and ice are melting

and nourishing the microorganisms and plants. The ability to measure the snow collected in an area and, subsequently, the amount of water in the system helps the researchers to understand the ecology of an area and its evolution over time [HCK*20].

Researchers implement various methods to measure the snowfall. These include radars (both terrestrial and satellite), lidars, or

ultrasound probes [LGG10]. The disadvantage of such systems is either their high complexity and purchase cost or the limitation of measuring only one place at a time.

A more affordable option is to use a set of trail cameras pointed at a location equipped with passive markers showing the snow height (snow stakes). The benefit of this setup is in its flexibility—researchers can include several cameras monitoring the area from various angles, which gives them a more robust system. It also allows them to visually verify the measurements to improve trust in data—this is virtually impossible with systems measuring only one place. On the other hand, this approach requires the development of image processing pipelines to get the measurements.

The time between frames needs to be up to a couple of hours to capture the snow evolution using trail cameras in sufficient detail. Our collaborators experimentally determined that taking a photo every 2½ hours is sufficient. This results in approximately 3500 images every year per camera that need to be analyzed. Fully manual data extraction at this scale is not practical, so we need to automate the process as much as possible. Based on our thorough evaluation of the existing methods, discussed in section 2, we decided to implement a combination of machine learning and morphological operations to detect the snow stakes in the image. However, as the weather often creates very unfavorable conditions, only a limited number of images can be correctly interpreted by a fully automatic solution. This results in a very noisy dataset with many outliers and missing values.

That led us to develop Antarstick, the tool for the exploration and manual correction of such datasets. Our tool is based on Shneiderman's mantra "Overview first, zoom and filter, details on demand" [Shn03]. It provides the user with an overview of the overall snow level from a measured area. The tool highlights the areas of the dataset that require manual treatment, which are usually diverging from the local average. When inspecting these outliers, the user can review the image and inspect the reason behind the outlier. They can repair the measurement manually or exclude it from the dataset.

Antarstick was designed in tight collaboration with the domain researchers from the Mendel Polar Station of Masaryk University, who were also testing the final product. Their evaluation is also summarized in this paper. Although the tool was primarily designed with domain-specific tasks in mind, the general concept of the tool is applicable to other areas, where time-lapse dataset is processed. We believe that our described methodology and provided source code (https://muni.cz/go/antarstick) can guide other visualization researchers in other domains.

The contributions of our paper to the visualization community can be thus summarized as:

- We propose a method for exploring and curating automatically processed time-lapse data.
- We designed visual representations of automatically processed data, their linking, and interactions with respect to the domain-specific requirements.
- We present a freely available solution with a sample dataset, demonstrating the proposed methodology and interaction concepts on a real-case scenario of snow-level extraction.

2. Related Work

Before delving into the details of methods and approaches relevant to our solution, it is worth mentioning our previous work [OHC*20], where we were also dealing with the data captured in the vicinity of the Mendel Polar Station in Antarctica. There, we primarily focused on designing a tool for exploring the correlation of various measurements, including wind speed and direction, soil and air temperature, and snow level. However, at that time, we used a straightforward solution for semi-automatic extraction of snow level. During the final testing of our previous tool, the experts expressed concerns regarding the robustness of the tool. They were interested in a more precise solution for measuring snow height, which initiated our further research presented in this paper.

2.1. Methods of Snow Measurement

The work of Lundberg et al. [LGG10] compiles a comprehensive list of snow-measuring techniques used in parts of Sweden, Norway, and Finland. The described techniques are used for measuring various parameters, such as depth of snow, density, and water equivalent. Many of the measuring techniques mentioned require a human operator at a site. Such methods are out of the question when used at a remote polar station, which is occupied only for several months a year. The researchers from Mendel Polar Station tested the ultrasonic transducer methods for the height measurement prior to installing the camera system and snow stakes. Their findings were that many of the measurements were imprecise due to wind-carried particles in the air since the area is very windy throughout the whole season. It was impossible to distinguish between the real value and the error, as they had only one number per measurement. Using the images from trail cameras can significantly improve the credibility of the measurement, as they show the data in the context it was measured.

Time-lapse photography has successfully been used to capture the presence of the snow and its height. The simpler task is to detect the presence of snow in the image. Millet et al. [MHW18] and Rüfenacht et al. [RBBS14] have used various clustering methods on a color image to identify the snow-covered percentage of a photo. Salvatori et al. [SPG*11] use histogram-based thresholding for detecting snow/no snow pixels. Fedorov et al. [FCFT16] evaluated the previous methods and compared them with several supervised machine-learning methods. They concluded that the machine learning approach yields generally more robust results with higher accuracy.

When measuring the snow height, there is usually an external marker required, typically in the form of a pole or *snow stake*. Several semi-automatic pipelines were devised [GPW13, OHC*20] to track these snow stakes in the images. In these approaches, the users determine the snow level in each photo by clicking on the stakes exactly where they are buried in the snow. The snow height is then automatically computed from the stake position and size that is defined a priori. The obvious problem with this approach is that the processing of large datasets is very time-consuming. Despite that, it is also common to process the images completely manually [LCW15, KHKN20].

To automate the measurement of the snow height, the first step is to find these snow stakes in the image automatically. An example

Figure 2: *Example images from our input datasets, demonstrating the variability in the image quality, caused by different weather conditions. (a),(b) Good quality day and night images; (c) color distortions caused by sun; (d),(e) day and night images during snowstorm; (f) camera is partially covered by snow.*

of this is the work of Kopp et al. [KTD19], using either image processing methods or a convolutional neural network to detect the snow stake. Once the stakes are detected, an automatic batch routine can analyze each photo and find how many pixels of the stake are covered with snow. This can be done by separating the snow-free pixels from the snow pixels based on the color [FW08] or counting visible height markers on the snow stakes [PHK*12]. However, the automatic measurement still suffers from various problems that require manual post-processing. The common issues include shifting of the camera housing and frost building up on the camera lenses, or reduced visibility conditions, such as fog or precipitation [PHK*12, BATDM21].

2.2. Cleansing and Visualizing Time-Dependent Data

The problem with automated methods is that they often create incorrect data that need to be manually revised. This is why we need data-cleansing solutions, that assist in the process of reviewing data.

Data cleansing is often the most time-consuming aspect of data analysis. Kandel et al. [KHP*11] provide examples of tasks required for cleaning a dataset, and Gschwandtner et al. [GGAM12] categorize types of errors and inconsistencies present in time-dependent data. There are many examples of software solutions used for visual analysis and subsequent data cleaning. Gschwandtner et al. [GE18] and Bernard et al. [BHR*19] presented tools for cleaning a general time-dependent dataset. Visplause [ASMP16] uses automated plausibility checks and multiple linked views to assess the quality of many time series at once. TimeCleanser [GAM*14] utilizes semi-automatic quality checks in combination with various representations, such as line charts, bar charts, and heatmaps, to detect data anomalies and provide users with an overview of the dataset quality. Finally, Shimabukuro et al. [SFdOL04] proposed a visual representation suitable for the comparison of a larger number of time-dependent climatology measures that can be helpful for identifying plausible values for missing data. The presented solutions are missing the crucial feature of exploring the underlying data, such as images in our case.

There are solutions developed for processing and labeling video data, i.e., for machine-learning applications. A tool by Kurzhals et al. [KHSW16] provides a rich environment for labeling a video feed into similar sequences and drawing them as scarf plots. The scarf plots are horizontal bands along the timeline that mark interesting sequences in the video. While this technique is suitable for video, the time-lapse feed has significant and sometimes uneven gaps between frames, which would result in very scarce representation.

In the following, we will discuss further approaches to the visual analysis of time series data that are the most relevant to our own application. For a more comprehensive overview, we kindly refer the reader to the survey of Aigner et al. [AMM*07] or to the book by Aigner et al. [AMST11]. When it comes to visualizing the development of the snow height over time, the most common approach is to use line charts [FW08, GPW13, LCW15, KHKN20] or scatterplots [PHK*12, KTD19]. Bögl et al. [BFG*17] presented an approach for analyzing patterns in periodic time series that can be potentially useful when depicting the snow evolution over time. However, the periodic encoding breaks the temporal coherence of the data, making it harder to spot outliers. Kopp et al. [KTD19]

used the color encoding in the background of the temporally ordered chart to encode periodic changes between day and night. These changes are important when assessing the quality of the images while analyzing time-lapse data.

2.3. Image Processing

As the pipeline in our solution first tries to automatically detect the snow stakes in the input images, we were also investigating the most suitable approaches to that. The most relevant ones are summarized in this section. We also sketch our selected solution here, as we are not considering that as novel to be presented as one of our contributions.

There are various image-processing techniques for detecting snow stakes, one of which is based on color channel analysis. Orange markers on a stake are used in [HM14] so that low values in the blue channel of an image reveal the stake's position. The approach [PHK*12] assumes the stakes are painted with alternating red and black bars of size 10 cm. And in [GPW13], stakes contain black bars separated 10 cm from each other. Convolutional network Mask R-CNN was used in [KTD19] for stake detection. The lowest pixel in an image classified by the network as belonging to the stake represents the point where it meets the snow.

3. Analysis of Requirements and Design Decisions

Within numerous interviews with our domain experts from the Mendel Polar Station, we discussed their issues with reading out the snow height from the series of images they are capturing in the vicinity of the station. Within the past years, they tried to find a suitable, at least semi-automatic solution, that would help them extract the required information from the images. However, in the end, they always ended with a purely manual approach, as the existing solutions were unreliable or did not fulfill their expectations.

As shown in Fig. 2, the combination of varying weather and light conditions, including snowstorms, day and night changes, sunlight glares, or similarity of the ground and sky for snowy images, is inevitably causing errors even in the most robust automatic solutions. Therefore, a certain level of human intervention is necessary to correct the snow height extracted from the problematic images.

As the manual annotation of all images would take days, the researchers always processed only a small subset of the images, and for the rest of the dataset, they performed simple interpolation. In this way, they were losing valuable information about sudden snowfalls and melting in shorter time periods, which could impact the ecosystem.

We concluded that the ideal tool for the researchers would help them visually explore the resulting values coming from the automatic measurement and manually correct the obviously erroneous values. In this way, the researchers would have means of understanding the quality of the whole dataset in a fraction of the time they originally spent on the analysis. They can easily fix it and understand the evolution of snow coverage in a given time period.

Based on these findings, we identified four high-level tasks that our envisioned tool should support:

T1 The tool should automatically process the whole dataset with only minimal user intervention.

T2 The tool should visualize the measured values and provide an indication of the credibility of the measurements.

T3 The tool should support the possibility of comparing the information about snow height evolution around a snow stake between the images taken from multiple cameras.

T4 The tool should allow the users to directly fix the erroneous records and store the corrected values in the dataset.

The first and main task is to alleviate the repetitive workload of the researchers when processing a large number of images. The user should not be forced to watch each individual image in the dataset and instead only set up the automatic extraction (**T1**). As it is apparent from the related work, the common workflow of measuring snow from an image is to mark the position of the snow stakes manually in the image, where the whole stake is visible. Another common technique is to define a region of interest, where the stake is expected so that the whole image is not searched. In the ideal case, the algorithm would find an image where the whole stake is visible and use it as a ground truth. This would automate the whole process, however the complexity of finding the reference image in a general dataset makes it hard to justify in a real-world application compared to its benefit. We believe the tasks for the user are simple enough and allow the system to be reasonably lightweight.

The second task requires giving a comprehensive overview of the measured data and then suggesting areas with potential errors (**T2**). For that, we need to already have an automatically processed dataset from **T1**. The dataset can be viewed from two perspectives: from the trail camera perspective, where there is information about snow height evolution around all stakes, and from the perspective of individual stakes, where we can compare the snow height seen from every camera that is capturing it. In the ideal case, the snow height measured from multiple cameras is identical. The stakes can be obscured by other objects, such as snow on the camera, so the ability to compare multiple camera sources is beneficial for verifying the data.

The appropriate visual support for the third task should help the user understand the relationships between data from one stake captured by multiple cameras and thus serve as an additional verification (**T3**). This requires the images to be taken ideally at the same time and if this is not achieved, we need to create a way of aligning images corresponding in time to each other. We also need to consider that the quality of the extracted information highly depends on the position of the stake on the image and its distance from the camera. It is evident that the readouts of those stakes far away and in the corner of the image will be of worse quality.

The fourth task aims to support fast and intuitive correction of incorrect records that the user found using the solutions defined by the three previous tasks (**T4**). The tool should present the measured data directly in the image and allow easy editing. The user should also have the option to completely discard an image if the quality is too poor.
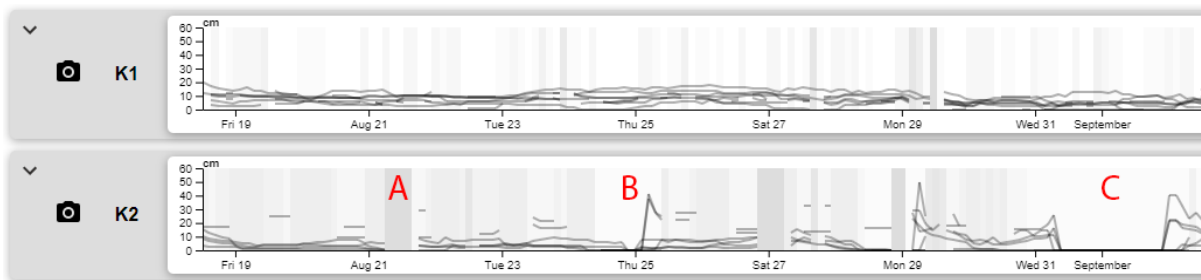
Figure 3: *Camera View with moving averages of individual stakes. The background rectangles show the missing data. The darker the color, the more stakes are missing at that point (A). The spike (B) suggests potentially erroneous data and the "sink" (C) shows the stretch without snow.*

4. Automated Snow Height Extraction

We developed an algorithm capable of producing an automatically measured set of images, with only minimal setup from the user. Since there is no standardized method, we developed our own one, using methods discussed in section 2. The full explanation of the algorithm can be found in work by Mráz [Mra21]. As our proposed algorithm utilizes a combination of already existing methods, we are not considering this as a novel contribution. However, we believe that it is worth to present for the sake of reproducibility.

In accordance with **T1**, we require only minimal user input when setting up the algorithm. We ask the user to select a representative photo, where all stakes are fully visible and mark their top and bottom. These positions now serve as a reference for the stake-finding algorithm. The user does this for each camera, and then they reference the same stake viewed from multiple cameras, so the visual representation across cameras is explicit. There is no more setup required from the user, and they can start the recognition.

First, we identify stakes using a histogram of oriented gradients [DT05] and a support vector machine [CVS95], which produces a set of rectangles indicating the presence of stakes in the image. We equalize the histograms in each rectangle. Then we detect the actual stakes contained in the computed rectangles by applying mathematical morphology [Soi03, Ser83] to convert them to binary images, where white pixels represent the stakes. Finally, we apply a version of Hough line transform [GMK99] to identify the lines in the binary images.

To avoid unnecessary processing of individual stakes, we evaluate if an image contains any snow at all. For this, we use a support vector machine classifier [CVS95] that was trained on manually labeled images (600 containing snow and 1800 without any snow). The image is processed either if it was recognized as an image with snow or if there was snow in the previous five images. This heuristic dramatically reduces the processing time.

In our evaluation, the processing of roughly 3600 images (3000 $\times$ 4000 px) in four threads took about 4.5 minutes (75 ms per image) on Core i7-4720HQ CPU. The threads took about 170 MB RAM of 8 GB available. The detection of snow/no-snow in the image is reasonably accurate, with about 97 % accuracy. The actual snow height detection was correct in 83 %. However, considering only images when the site was fully covered by snow, the height detection was correct in 92 %.

5. Antarstick Design

This section presents the details about the design of our proposed Antarstick tool. The tool consists of three main views. As the locations are captured from multiple angles, presenting all available information for comparison is necessary. The first visual representation the user encounters is the *Camera View* (Fig. 3), showing the snow height extracted at all visible stakes in the camera's field of view. The idea is to show the overall structure of the dataset throughout the given time period, represented by lines of moving averages. Exact values for individual stakes can be inspected in the *Stake View* (Fig. 4). It allows for checking every value and its relation to the moving average. If the stake is visible from more cameras, the *Stake View* presents this information as well (Fig. 4). The most detailed view, where the user can edit the incorrect detection, is the *Image Detail View* (Fig. 6). The interplay between these three proposed views guides the user to see the global overview of the dataset first, locate regions with possible problems, and then inspect them in detail.

5.1. Camera View

The purpose of *Camera View* is to aggregate all snow heights in one place. It consists of a set of line charts, where each chart corresponds to a single camera. The view was designed as a first overview of the dataset. The juxtaposed views of each camera serve as a comparison tool (**T3**), while the superimposition of the snow stakes shows the spread of snow height across the measured area. The naïve assumption is that all snow stakes should have the same snow height at a time. If there is a consistent spread among the heights in the graph, it indicates diverse conditions at each stake, pointing towards a potential trend. On the other hand, if there is a larger spread for a shorter time, that would indicate possibly erroneous values.

When designing the view, we needed to find a balance between complete information and readability. The initial solution of superimposing all heights as individual points resulted in an unreadable clutter. The discrete points also hide the fact that a measurement is missing, as is presented by Eaton et al. [EPD05]. We chose a line chart to show the information, for several reasons. The first is a visualization of sudden spikes in data (Fig. 3B). They could signify a massive snowstorm, or more probably, the value has been misinterpreted. In either case, it is something worth investigating, and the spikes are drawing attention to it.
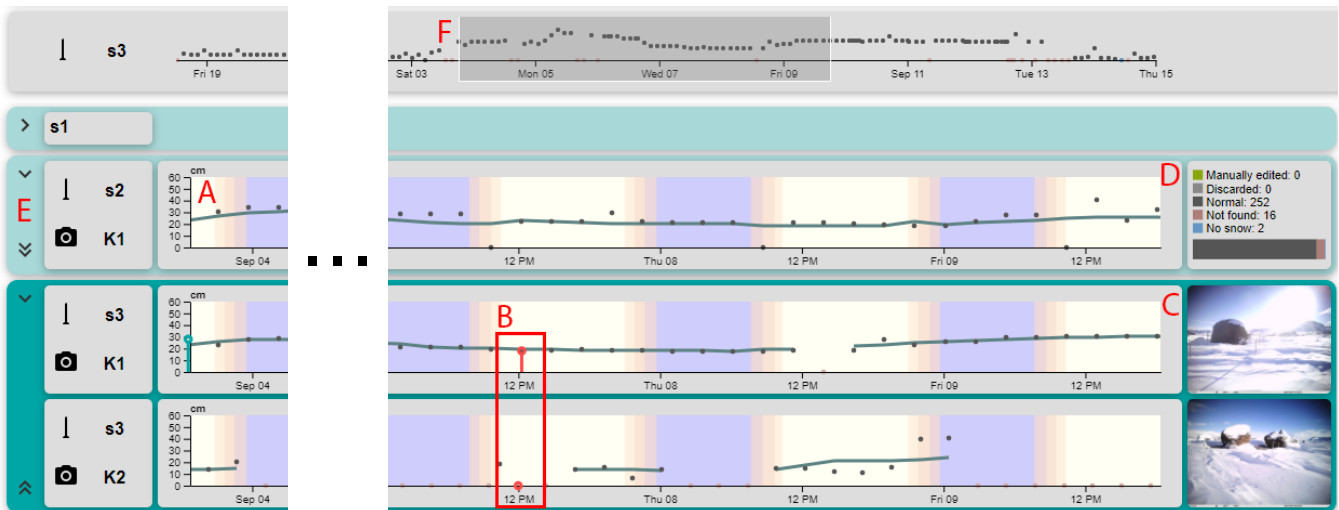
Figure 4: *The Stake View shows the data points and the moving average of one stake (A). When the stake is selected (B), the summary chart on the right (D) is replaced by a preview of the image (C). Chevrons on the left part (E) for showing and collapsing charts. The Overview Chart (F) shows the position and the zoom level with respect to the whole dataset.*

The second benefit of a line chart is the visualization of missing values since the line segments show continuity and end abruptly when the value is missing. To further enhance the visualization of missing values, we used the background to show rectangles of different shades. They encode the missing values at a given time. The darker the background, the more values are missing. When a longer series is missing, often due to the covering of the camera lens, the rectangles create wide dark bands, drawing attention to them (Fig. 3A). That way, we can guide the user to the low-quality parts of the dataset. Using them also reveals a difference between missing values and correct values without snow (Fig. 3C).

To further reinforce the interaction, we connected the *Camera View* with the *Stake View* in zoom and drag, so both views show the same time frame. The user can also select any line in the *Camera View* to open the corresponding *Stake View*.

5.2. Stake View

The *Stake View* (Fig. 4) provides a detailed view of a single stake, seen from multiple cameras (**T2**). The initial design featured a point chart, where each point is one measurement in time. That would suffice for a simple overview, however, there still were things to improve in terms of guidance.

First, the user couldn't distinguish between missing and zero values. We used the color and glyph channel of the point to encode this information, along with the image status (Fig. 5). On top of it, we also added the possibility to filter the points based on the same image status. The filtered points are semi-transparent and, therefore, less noticeable, but they are still accessible.

We use the *Stake View* to solve task **T3**, where we want to compare data from multiple cameras. Since we are using color and glyph channels to distinguish between states, it would be impractical to superimpose all views into the same chart. Instead, to reduce visual

clutter, we show all views juxtaposed. When selecting a point in one of the cameras, the closest point in the other camera time-wise is also selected (Fig. 4B). In case the other camera has a large gap in the measurements, the closest value will be at the edge of said gap and the user can disregard it for the comparison.

Similar to the discussion in subsection 5.1, the point chart is not well suited for finding outliers. We include a line chart in the form of a running average, which creates a proxy of an expected value. When there is an outlier, it shows up further from the average line, which guides the user towards it. We reused these running average lines in the *Camera View* to compare the stakes among themselves.

When using either the *Stake View* or the *Camera View*, the user can zoom in to inspect details in the timeline. It is easy to lose track of the scale and position, so we implemented the *Overview Chart* (Fig. 4F). It serves as a minimap, showing the position and visible range of the dataset. It is also connected with other views, so the user can drag and select the visible range from the *Overview Chart*.

The background shows the sun position, as the amount of light is one of the major factors that play into the visibility of the stakes. The user can easily see the day/night cycle and infer whether the outlier corresponds to a lack of light.

To the right of the chart, there is a window showing additional information. When no point is selected, it shows a distribution of points across states in a stacked bar chart (Fig. 4D), otherwise it shows the currently selected image (Fig. 4C). By clicking on the image preview, the user opens the *Image Detail View*.

5.3. Image Detail View

The purpose of this window is to inspect data and provide the tool for correcting it. We reserved most of the space for the original image, as it is the original purpose of this view (Fig. 6A). All found stakes
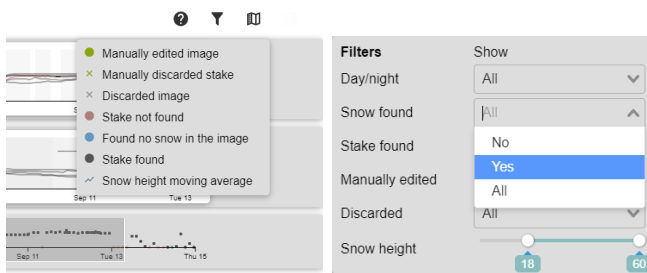
Figure 5: *Left: The legend shows different colors and glyphs of each point. Right: Available filtering options for images, stakes, and displayed snow height values. The filtered data points are semi-transparent in the Stake Views.*

are marked by an overlay showing the position of the stake and the snow height, all of which the user can hide (Fig. 6C). The user can move the stakes and the snow height, which directly changes the dataset underneath. Since the image-processing algorithm does not always find all stakes, it is possible to both add and remove stakes, as well as undo these changes.

When fixing such large datasets, it is easy for the user to get lost and review the same image multiple times. We prevent this waste of time by implementing a system of marking the images. The marks are then visible in the *Stake Views*. Any time the user edits the image, it is marked as "Approved," which changes the color of the point in the *Stake View* (see Fig. 5). The user can uncheck the mark at any moment to note the image is not fully processed. In the same way, the user can discard the image, which will still show up in the *Stake View* as a cross, but it will not be taken into account when computing the running average.

When fixing a larger sequence, it would be cumbersome to return to the main window and select the next image. Therefore, we included a small gallery of surrounding images (Fig. 6B) that also shows the "Approved"/"Discarded" marks.

6. Design Study

In order to understand the effectiveness of the proposed tool, we performed a design study to gain the user's opinions and insights. The study has been conducted as a series of semi-structured interviews, as described in [Off].

Data: We used a real dataset from Antarctica. Images were captured at a site on James Ross Island during the 2018 season. The measured area is a large boulder with eight snow stakes around it and can be seen in Fig. 2. It consists of over 7000 images taken by two trail cameras in roughly 2-hour intervals. Each image has been processed by our classifier, which tried to find all snow stakes and measure the snow height. This dataset has been used within the study with the experts.

Sample: We have conducted the study with five experts who regularly visit the Mendel Polar station in Antarctica and actively analyze the gathered data, including the trail camera images that form the input dataset of Antarstick. These experts were researchers,

all with experience with field measurements in Antarctica. Two of the participants are the original authors of the dataset, responsible for measuring snow at the Mendel Polar Station.

Procedure: In the study, the developers of the tool engaged in sessions with experts. These sessions were facilitated by two designers of the tool, meeting with one to two experts in a meeting room. Two of the five experts who participated joined remotely via video call, manipulating the tool through remote desktop software. Each session, spanning approximately one hour, was recorded with explicit consent from the participants. The first part comprised the introduction of the Antarstick tool, its purpose, and its functionality. After that, the experts were asked to interact with the application and narrate their mental model using the *thinking-aloud* method [Lew82]. We prepared a set of questions that we asked the participants to gain deeper insights during the tasks and afterward. The questions included their opinions on Antarstick, how they use individual visualizations, or how the visualizations work together. We also asked for useful, redundant, and missing features.

6.1. Results

In our evaluation, we were able to get many findings and remarks. We grouped them into several categories:

Workflow: The experts first inspected the general overview of the year provided by the *Camera Views* (Fig. 3), where they can see the distribution of missing data. Using that as navigation, the users located individual problems using the *Stake View* (Fig. 4).

We observed two different approaches when using the *Image Detail View* (Fig. 6) to correct data during the evaluation. The first approach is to edit the values by moving the markers to the correct position, which works well with images of lower quality (e.g., a quick drop in the height that was incorrectly measured in a blurred image). The second approach is to discard the incorrectly measured stake right away. This pruning approach allows for a quick improvement of the resulting dataset, removing large portions of erroneous but otherwise unimportant data. Both methods are supported by our design.

Attitude towards visualizations: The participants stated that even without any correction of the individual values, the initial view already gave them an understandable preview of the snow cover trends over the examined time period. They confirmed that the background highlight of missing data presented in the *Camera View* clearly suggests the critical parts where some user intervention is required. They also very much appreciated the possibility of accessing the original image using the *Image Detail view*. Lastly, the participants stated that the navigation in the charts and the zooming and data selection were intuitive. Especially the synchronized zooming was described as very useful.

Correction of errors: As we mentioned earlier, there were two possible workflows that the experts used. The first one, editing the stake position, requires the user to move the stake manually. The editing was regarded by the experts as easy and intuitive. The same goes for discarding and adding individual stakes, which is possible in the same window, in the *Image Detail View*. One user suggested the use of keyboard shortcuts.
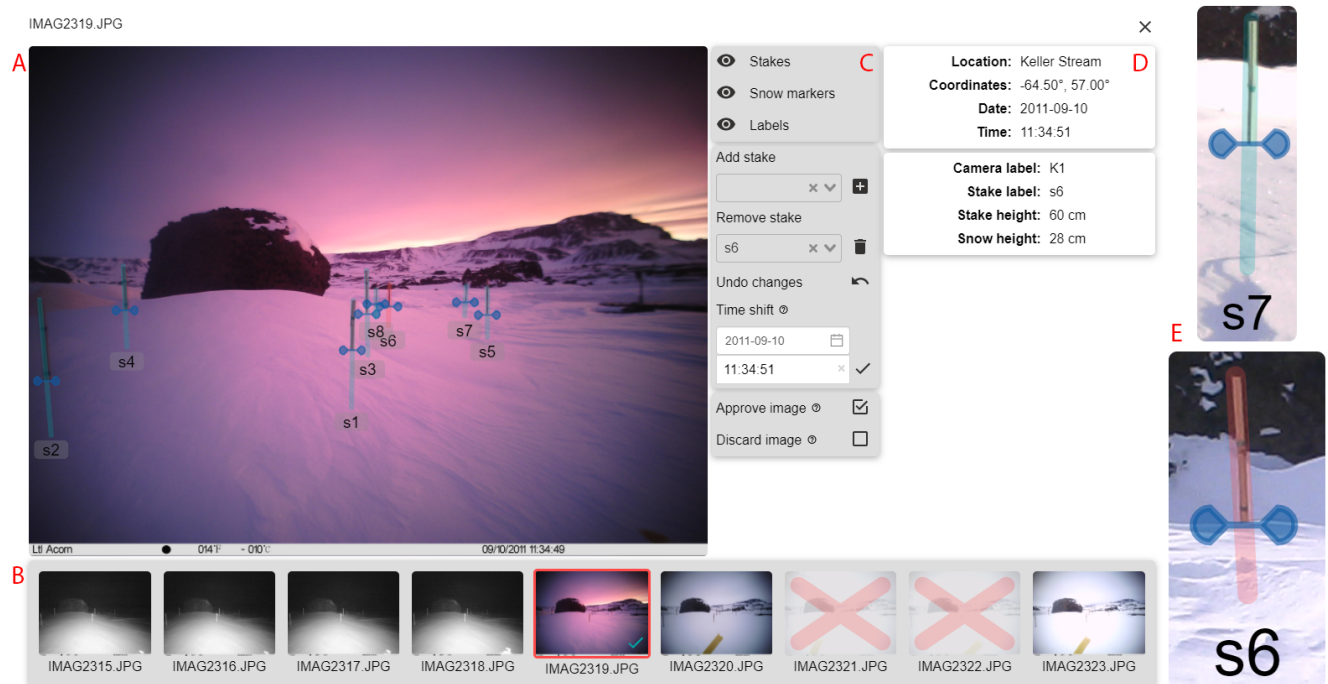
Figure 6: *The Image Detail View consists of a zoomable view of the source image with stake overlays (A). The image controls (C) offer tools to manipulate the stakes, and the image and stake info (D) show all available details. The gallery at the bottom (B) shows the neighboring images, providing important context and navigation. On the right, there are the stake overlays. The general stake overlay is on top, and the currently selected one is on the bottom (E).*

Limitations and new features: Some of the suggestions by users addressed the visualizations and the presentation of information, while others mentioned the user interface and interaction with the software. As for the visualizations, the users mentioned subtle discrepancies in the color scheme and other minor adjustments. For example, the reddish color selected for *Stake not found* state reminded them more of a ground color, which they connected more with the images, where there was no snow. The experts also had some requirements regarding the interaction, such as batch processing of many images at once. Usually, when the camera gets snowed in, the whole period can be discarded at once. That means the stake charts would need brushing and allow for removal without going into a detailed view. This will also complement the workflow that discards erroneous data instead of fixing them.

We believe that this evaluation revealed important insights, and we will use it to improve Antarstick to suit experts' needs. In the following section, we draw some of the conclusions for the subsequent work.

7. Discussion & Lessons Learned

After the design process, implementation, and user evaluation, we were able to summarize the findings into the following ideas. These are the lessons learned from the design study.

Users are more interested in correct data rather than complete data. When the users are cleaning the dataset, we have observed that they are not reluctant to discard data, especially when they do not see any significant change between the neighboring images. Therefore, it is important to provide the user with an easy way of discarding the data, be it individual measurements in the image or the image as a whole.

Thanks to the overlay of the stakes (Fig. 6E), it was trivial to identify and then adjust the measurement. The lesson learned is that when working with pictorial data, it is often preferable to show the user both the source data and the computed result.

Point chart with running average guides the users towards incorrect results. The users were able to find the outliers more easily when they had a reference line going through the chart. Similarly, they were able to spot other features of the line chart, such as split lines where the measurements were missing, or abrupt jumps in the data.

Background can be used to display the validity of the data. When designing the *Camera View*, we were aiming to summarize the overall trends in the dataset. We were able to utilize the background to show the number of valid data points, which clearly highlights the problematic segments. Our test subjects confirmed our initial assumption that it is an effective representation.

Batch processing based on brushing is required for efficient cleaning of data. During the evaluation, experts requested the option to select more images and change their parameters at once. This would speed up the cleaning significantly. However, it is not practical to allow every action to be done in a batch. In our case, the following edits are meaningful:

- discard the stake from selected photos,
- discard selected photos,
- set snow height at a stake to zero,
- mark images as no-snow (thus setting the values for all stakes in this image to zero).

Within the batch processing, setting the height of snow to a different value than zero is not a valid operation, since often the incorrect value is measured due to the shifted stake, and the image would still need to be checked manually. The same applies to setting the same height for every stake in the image (except for zero). In conclusion, it is practical to include batch editing, but the actions need to be well-designed.

Splitting the whole dataset into smaller parts can make it easier to work with. It is not usually crucial to clean the dataset as a whole. In our case, the images share context only locally, e.g., the camera gets snowed-in only for a couple of hours/days, or the snowstorm recedes overnight. The overview of the dataset is helpful as it provides a preliminary view of the whole measured period (i.e., year) and allows one to keep track of the edited data. However, the cleansing process itself does not require a detailed view of the entire year. While the application is designed to work with the whole year (over 7000 images), in some of our tests we tried to clean smaller portions of the data consisting of roughly one month. This made the visualizations less cluttered and it was easier to navigate in the shorter timeline. Arguably, if the users were able to finish cleansing faster, it would give them a greater sense of accomplishment and make the experience more enjoyable.

7.1. Generalization of the visualization

The proposed workflow of cleaning the automatically processed time-lapse data can be applied to other application domains as well. Especially in natural sciences, where long-term observations are quite common, time-lapse imaging techniques are a good way of studying changes. Combined with automatic processing, they can prove a powerful tool for researchers.

While we performed the evaluation with the experts, we have also discussed other possible usage scenarios of this tool. One of the interesting ideas was time-lapse measurements of the water level in lakes of James Ross Island. The water level measurement can be done in the same exact way as measuring snow, which suggests Antarstick would need little to no adjustments to process this kind of data. We would like to explore this application in the future.

One of the hypothetical usage scenarios could be the exploration of lichen growth. A commonly used technique is to take pictures of the colonies to compare the growth between years [SGP07, Ben90]. With the creation of an automatic measuring algorithm that would mark the grown area, the approach for cleaning the dataset can be utilized. Instead of a stake, every colony would have its own

chart, showing the measured area. Because the steady growth of the lichen is expected, every change in this trend would be easily detectable, and with a detailed view of every image, it could be instantly reviewed.

In more general terms, every time-lapse dataset with one or more scalar values extracted from it can be visualized using the principles described in our solution.

8. Conclusion and Future Work

In this paper, we introduced Antarstick, an interactive tool for the visual exploration and correction of information about the evolution of snow height values around stakes that were automatically extracted from time-lapse photographs. Antarstick was created in collaboration with experts in the field of environmental sciences who conduct yearly experiments in Antarctica. We asked five of them to evaluate the final version of Antarstick and its ability to aid their research. Based on the feedback we collected during the evaluation, we can conclude that the domain experts appreciated the features and options that the tool offers, as it allows them to quickly verify and correct large datasets of images, which was not possible to reach before without tedious manual work. The evaluation identified interesting directions for future improvements, coming from a new observation that the need for manual correction differs according to the amount of snow in a particular time period. We are planning to utilize this observation to improve the visual guidance in the next version of Antarstick. Moreover, while we demonstrated Antarstick's features and capabilities on a dataset that was parsed by our own classifier, the tool can be utilized for cleansing data produced by any approach that provides access to the detected stakes and the raw images. We also illustrated possible generalizations to cleaning other time-lapse image datasets. We believe that this will inspire other people to clean their data efficiently.

Acknowledgments

References

[AMM*07] AIGNER W., MIKSCH S., MÜLLER W., SCHUMANN H., TOMINSKI C.: Visualizing time-oriented data—a systematic view. *Computers & Graphics 31*, 3 (2007), 401–409. 3

[AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of Time-Oriented Data*. Springer-Verlag London, 2011. doi:10.1007/978-0-85729-079-3. 3

[ASMP16] ARBESSER C., SPECHTENHAUSER F., MÜHLBACHER T., PIRINGER H.: Visplause: Visual data quality assessment of many time series using plausibility checks. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2016), 641–650. 3

[BATDM21] BONGIO M., ARSLAN A. N., TANIS C. M., DE MICHELE C.: Snow depth time series retrieval by time-lapse photography: Finnish and italian case studies. *The Cryosphere 15*, 1 (2021), 369–387. 3

[Ben90] BENEDICT J. B.: Experiments on lichen growth. i. seasonal patterns and environmental controls. *Arctic and Alpine Research 22*, 3 (1990), 244–254. 9

[BFG*17] BÖGL M., FILZMOSER P., GSCHWANDTNER T., LAMMARSCH T., LEITE R. A., MIKSCH S., RIND A.: Cycle plot revisited: Multivariate outlier detection using a distance-based abstraction. *Computer Graphics Forum 36*, 3 (2017), 227–238. 3

[BHR*19] BERNARD J., HUTTER M., REINEMUTH H., PFEIFER H., BORS C., KOHLHAMMER J.: Visual-interactive preprocessing of multivariate time series data. *Computer Graphics Forum 38*, 3 (2019), 401–412. 3

[CVS95] CORTES C., VAPNIK V., SAITTA L.: Support vector networks. *International Journal on Machine Learning* (1995), 1–25. 5

[DT05] DALAL N., TRIGGS B.: Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) 1* (2005), 886–893 vol. 1. doi:10.1109/CVPR.2005.177. 5

[EPD05] EATON C., PLAISANT C., DRIZD T.: Visualizing missing data: Graph interpretation user study. In *Human-Computer Interaction-INTERACT 2005: IFIP TC13 International Conference, Rome, Italy, September 12-16, 2005. Proceedings 10* (2005), Springer, pp. 861–872. 5

[FCFT16] FEDOROV R., CAMERADA A., FRATERNALI P., TAGLIASACCHI M.: Estimating snow cover from publicly available images. *IEEE Transactions on Multimedia 18*, 6 (2016), 1187–1200. 2

[FW08] FLOYD W., WEILER M.: Measuring snow accumulation and ablation dynamics during rain-on-snow events: innovative measurement techniques. *Hydrological Processes: An International Journal 22*, 24 (2008), 4805–4812. 3

[GAM*14] GSCHWANDTNER T., AIGNER W., MIKSCH S., GÄRTNER J., KRIGLSTEIN S., POHL M., SUCHY N.: Timecleanser: A visual analytics approach for data cleansing of time-oriented data. *Proceedings of the 14th international conference on knowledge technologies and data-driven business* (2014), 1–8. 3

[GE18] GSCHWANDTNER T., ERHART O.: Know your enemy: Identifying quality problems of time series data. In *2018 IEEE Pacific Visualization Symposium (PacificVis)* (2018), IEEE, pp. 205–214. 3

[GGAM12] GSCHWANDTNER T., GÄRTNER J., AIGNER W., MIKSCH S.: A taxonomy of dirty time-oriented data. *International Conference on Availability, Reliability, and Security* (2012), 58–72. 3

[GMK99] GALAMHOS C., MATAS J., KITTLER J.: Progressive probabilistic hough transform for line detection. In *Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149)* (1999), vol. 1, IEEE, pp. 554–560. 5

[GPW13] GARVELMANN J., POHL S., WEILER M.: From observation to the quantification of snow processes with a time-lapse camera network. *Hydrology and Earth System Sciences 17*, 4 (2013), 1415–1429. doi:10.5194/hess-17-1415-2013. 2, 3, 4

[HCK*20] HRBÁČEK F., CANNONE N., KŇAŽKOVÁ M., MALFASI F., CONVEY P., GUGLIELMIN M.: Effect of climate and moss vegetation on ground surface temperature and the active layer among different biogeographical regions in Antarctica. *CATENA 190* (2020), 104562. doi:10.1016/j.catena.2020.104562. 1

[HM14] HEDRICK A., MARSHALL H.: Automated snow depth measurements in avalanche terrain using time-lapse photography. *2014 International Snow Science Workshop* (2014). doi:10.13140/2.1.2049.6646. 4

[KHKN20] KŇAŽKOVÁ M., HRBÁČEK F., KAVAN J., NÝVLT D.: Effect of hyaloclastite breccia boulders on meso-scale periglacial-aeolian landsystem in semi-arid Antarctic environment, James Ross Island, Antarctic Peninsula. *Cuadernos de Investigacion Geografica 46*, 1 (2020), 7–31. 2, 3

[KHP*11] KANDEL S., HEER J., PLAISANT C., KENNEDY J., VAN HAM F., RICHE N. H., WEAVER C., LEE B., BRODBECK D., BUONO P.: Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization 10*, 4 (2011), 271–288. 3

[KHSW16] KURZHALS K., HLAWATSCH M., SEEGER C., WEISKOPF D.: Visual analytics for mobile eye tracking. *IEEE transactions on visualization and computer graphics 23*, 1 (2016), 301–310. 3

[KTD19] KOPP M., TUO Y., DISSE M.: Fully automated snow depth measurements from time-lapse images applying a convolutional neural network. *Science of The Total Environment 697* (2019), 134213. doi:10.1016/j.scitotenv.2019.134213. 3, 4

[LCW15] LIU J.-F., CHEN R.-S., WANG G.: Snowline and snow cover monitoring at high spatial resolution in a mountainous river basin based on a time-lapse camera at a daily scale. *Journal of Mountain Science 12*, 1 (2015), 60–69. 2, 3

[Lew82] LEWIS C.: *Using the" thinking-aloud" method in cognitive interface design*. IBM TJ Watson Research Center Yorktown Heights, NY, 1982. 7

[LGG10] LUNDBERG A., GRANLUND N., GUSTAFSSON D.: Towards automated 'ground truth' snow measurements—a review of operational and new measurement methods for Sweden, Norway, and Finland. *Hydrological processes 24*, 14 (2010), 1955–1970. 2

[MHW18] MILLET P., HUWALD H., WEIJS S. V.: Extracting high resolution snow distribution information with inexpensive autonomous cameras. *EPiC Series in Engineering 3* (2018), 1397–1405. 2

[Mra21] MRAZ R.: *Analysis of images monitoring the snow level in Antarctica*. Master's thesis, Masaryk University, Faculty of Informatics, Brno, 2021. URL: https://is.muni.cz/th/fzrte/?lang=en. 5

[Off] OFFICE OF DATA, ANALYSIS, RESEARCH & EVALUATION ADMINISTRATION ON CHILDREN, YOUTH & FAMILIES: Qualitative research methods in program evaluation: considerations for federal staff. https://www.acf.hhs.gov/sites/default/files/documents/acyf/\qualitative_research_methods_in_program_evaluation.pdf. Accessed: 2021-09-06. 7

[OHC*20] ORÉMUŠ Z., HASSAN K. A., CHMELÍK J., KŇAŽKOVÁ M., BYŠKA J., RAIDOU R. G., KOZLÍKOVÁ B.: Pingu principles of interactive navigation for geospatial understanding. *2020 IEEE Pacific Visualization Symposium (PacificVis)* (2020), 216–225. doi:10.1109/PacificVis48177.2020.7567. 2

[PHK*12] PARAJKA J., HAAS P., KIRNBAUER R., JANSA J., BLÖSCHL G.: Potential of time-lapse photography of snow for hydrological purposes at the small catchment scale. *Hydrological Processes 26*, 22 (2012), 3327–3337. doi:10.1002/hyp.8389. 3, 4

[RBBS14] RÜFENACHT D., BROWN M., BEUTEL J., SÜSSTRUNK S.: Temporally consistent snow cover estimation from noisy, irregularly sampled measurements. *2014 International Conference on Computer Vision Theory and Applications (VISAPP) 2* (2014), 275–283. 2

[Ser83] SERRA J.: *Image Analysis and Mathematical Morphology*. Academic Press, Inc., USA, 1983. 5

[SFdOL04] SHIMABUKURO M. H., FLORES E. F., DE OLIVEIRA M. C. F., LEVKOWITZ H.: Coordinated views to assist exploration of spatio-temporal data: A case study. *Second International Conference on Coordinated and Multiple Views in Exploratory Visualization* (2004), 107–117. 3

[SGP07] SANCHO L. G., GREEN T. A., PINTADO A.: Slowest to fastest: extreme range in lichen growth rates supports their use as an indicator of climate change in antarctica. *Flora-Morphology, Distribution, Functional Ecology of Plants 202*, 8 (2007), 667–673. 9

[Shn03] SHNEIDERMAN B.: The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*. Elsevier, 2003, pp. 364–371. 2

[Soi03] SOILLE P.: *Morphological Image Analysis: Principles and Applications*, 2 ed. Springer-Verlag, Berlin, Heidelberg, 2003. 5

[SPG*11] SALVATORI R., PLINI P., GIUSTO M., VALT M., SALZANO R., MONTAGNOLI M., CAGNATI A., CREPAZ G., SIGISMONDI D.: Snow cover monitoring with images from digital camera systems. *Italian Journal of Remote Sensing 43*, 6 (2011). 2