Product Validation Plan (PVP)

Prepared by: Glaciers_cci consortium
Contract: 4000101778/10/I-AM
Name: Glaciers_cci-D2.1_PVP
Version: 0.9
Date: 28.02.2012

Contact: Frank Paul
Department of Geography
University of Zurich
Winterthurerstr. 190
CH-8057 Zurich
frank.paul@geo.uzh.ch

Technical Officer: Stephen Plummer
ESA ESRIN

University of Zurich
University of Oslo
University of Bristol
GAMMA
enveo
University of Leeds
The work described in this report was done under ESA contract 4000101778/10/I-AM. Responsibility for the contents resides in the authors that prepared it.

Author team:
Thomas Nagler, Kilian Scharrer, Helmut Rott (Enveo), Frank Paul (GIUZ), Andrew Shepherd, Francesca Ticconi (SEEL), Tazio Strozzi (Gamma), Andreas Kääb, Christopher Nuth (GUIO)

Glaciers_cci Technical Officer at ESA:
Stephen Plummer
Summary

This document is the Product Validation Plan (PVP) of the ESA project Glaciers_cci. It outlines and schedules the validation and round robin (RR) activities undertaken by the project consortium and external participants for the three products generated within the project: (i) glacier area, (ii) glacier elevation change as observed by multi-temporal altimeter data and DEM differencing, and (iii) ice velocity.

Due to the different characteristics of each product, a detailed validation strategy is outlined for every product individually and the approaches can differ substantially between them. However, all products share a general problem of limited availability of datasets suitable for validation (especially in-situ data), which is the limiting factor for the level of validation that can be achieved for each of the products. Therefore, the project partners carefully selected the test sites based on availability and accessibility of validation data allowing the best possible validation for each product.

For glacier area, validation is closely related the RR exercises and follows a two-way approach including comparisons of the derived glacier area against a reference data set and multiple manual digitisations of the same glacier. For the RR, the latter will also be applied to very high resolution EO data that serve as reference datasets in the product validation.

Elevation change is divided into altimetry and DEM differencing. The validation strategy for altimetry also follows a two-way approach, with method and sensor inter-comparisons and – if available – comparisons of satellite altimeter products with airborne data. For DEM differencing, the strategy is generally based on comparing the individual satellite DEMs with higher accuracy and precision data sets to quantify systematic and random errors.

Validation of glacier velocity products will follow a three-step strategy where the performance of the matching algorithms applied to determine displacement vectors is evaluated and/or ice velocity products from different methods/sensors are compared. In some regions in-situ GPS measurements will serve as a reference.

The RR exercises for the three products will first be performed in a consortium-internal round, in order to identify potential problems and weaknesses of the strategy and to ensure a smooth cycle with the external participants. Therefore, we only present a preliminary framework of the RR tests in this document. The detailed structure, focus, protocol, etc. will evolve during the internal assessment.
Contents

1. Introduction ............................................................................................................................................. 6
   1.1. Purpose .............................................................................................................................................. 6
   1.2. Outline .............................................................................................................................................. 6

2. CCI project guidelines ............................................................................................................................. 7

3. Master Schedule ..................................................................................................................................... 8

4. Rules for unbiased validation and validation criteria ............................................................................... 10
   4.1. Glacier area ....................................................................................................................................... 10
   4.2. Elevation change – DEM differencing ............................................................................................ 10
   4.3. Elevation change – Altimetry ......................................................................................................... 11
   4.4. Ice velocity ....................................................................................................................................... 11

5. Planned validation activities for glacier area product .................................................................................. 12
   5.1. Background ....................................................................................................................................... 12
   5.2. Selected test sites and available validation data ............................................................................... 13
      5.2.1. Chugach Mts., Alaska (test site GA1) ....................................................................................... 14
      5.2.2. Ötztal, Alps (test site GA2) ....................................................................................................... 15
      5.2.3. Silvretta and Gotthard, Alps (test site GA3) ............................................................................ 16
      5.2.4. Kashmir, Himalaya (test site GA4) .......................................................................................... 16
   5.3. Validation strategy ............................................................................................................................. 17
      5.3.1. Product validation with reference data ..................................................................................... 18
      5.3.2. Relative inter-comparisons ..................................................................................................... 18

6. Planned validation activities for elevation change products ........................................................................ 19
   6.1. Selected test sites and available validation data ............................................................................... 20
      6.1.1. Devon Icecap (test site EC1) .................................................................................................... 21
      6.1.2. Austfonna Icecap (test site EC2) .............................................................................................. 23
      6.1.3. Findelen Glacier (test site EC3) ............................................................................................... 26
      6.1.4. Southern Alps, New Zealand (test site EC4) .......................................................................... 27
   6.2. Validation strategy and parameters ................................................................................................... 28
      6.2.1. DEM differencing ...................................................................................................................... 28
      6.2.2. Repeat altimetry (cross-over and repeat track) ....................................................................... 29

7. Planned validation activities for glacier velocity products ......................................................................... 31
   7.1. Selected test sites and available validation data ............................................................................... 32
      7.1.1. Breidamerkurjökull, Iceland (test site V1) ................................................................................. 32
      7.1.2. Kronebreen, Svalbard (test site V2) .......................................................................................... 34
      7.1.3. Nordaustlandet, Svalbard (test site V3) .................................................................................... 35
      7.1.4. Baltoro glacier, Karakoram (test site V4) ............................................................................... 38
7.2. Validation strategy ................................................................................................................. 38
  7.2.1. Internal method evaluation ............................................................................................... 39
  7.2.2. Method/sensor cross-comparisons ..................................................................................... 40
  7.2.3. Validation of EO products with in-situ velocity measurements (VT5) ......................... 41

8. Review process and decision sequence ....................................................................................... 42
  8.1. Glacier area ......................................................................................................................... 42
  8.2. Elevation change .................................................................................................................. 42
    8.2.1. Altimetry ....................................................................................................................... 42
    8.2.2. DEM differencing ........................................................................................................ 43
  8.3. Ice velocity ........................................................................................................................ 43

9. Resources summary and reference data sets ............................................................................. 44
  9.1. Glacier area ......................................................................................................................... 44
  9.2. Elevation change .................................................................................................................. 44
  9.3. Ice velocity ........................................................................................................................ 44

10. Validation documents and endorsements ............................................................................... 45

11. Round Robin exercises ......................................................................................................... 46
  11.1. Overall goals of round robin exercises ............................................................................. 46
  11.2. Tasks and responsibilities for the round robin activities .................................................. 48
  11.3. Schedule for the round robin experiments ....................................................................... 48
  11.4. Glacier area ....................................................................................................................... 48
    11.4.1. Goals of the round robin experiments ......................................................................... 48
    11.4.2. Definition of data sets ............................................................................................... 49
    11.4.3. Methods for inter-comparison and analysis ............................................................ 49
    11.4.4. Selection criteria ...................................................................................................... 50
  11.5. Elevation change - DEM differencing .............................................................................. 50
    11.5.1. Goals of the round robin ........................................................................................ 50
    11.5.2. Definition of data sets ............................................................................................ 50
    11.5.3. Methods for inter-comparison and analysis .......................................................... 51
    11.5.4. Selection criteria .................................................................................................... 51
  11.6. Elevation change – Altimetry ........................................................................................... 51
    11.6.1. Goals of the round robin ........................................................................................ 51
    11.6.2. Definition of data sets ............................................................................................ 52
    11.6.3. Methods for inter-comparison and analysis .......................................................... 52
    11.6.4. Selection criteria .................................................................................................... 52
  11.7. Ice Velocity ......................................................................................................................... 52
    11.7.1. Goals of the round robin experiments ....................................................................... 52
    11.7.2. Definition of data sets ............................................................................................ 52
    11.7.3. Methods for inter-comparison and analysis .......................................................... 54
    11.7.4. Selection criteria .................................................................................................... 54
  11.8. Round Robin Protocol .................................................................................................... 54

12. References ............................................................................................................................ 55

13. Acronyms ............................................................................................................................... 57
1. Introduction

1.1. Purpose

This report deals with the product validation plan (PVP) for the three main products generated within the Glaciers_cci project:

- glacier area (outline)
- glacier elevation change (altimetry and DEM differencing)
- glacier ice velocity

The PVP is the first deliverable of Task 2 (D2.1) with the overall goals algorithm development, inter-comparison and selection. The PVP has its focus on the selection criteria and describes for each of the three products how the performance and accuracy assessment of the algorithms is conducted from a more strategic point of view. It also describes the data sets to be used for the round robin (RR) and validation, and how independent validation of the products is guaranteed. Algorithms and parameters for characterizing the uncertainty of the products are specified following the recommendations from the CCI Project Guidelines (ESA, 2010) and are adapted to the three products. Furthermore, a formal validation procedure is proposed, which will be used within the round robin.

1.2. Outline

This document is divided in 13 chapters, starting with an Introduction (1), and short review of the ESA CCI validation guidelines with respect to the individual Glaciers_cci products (2). In section 3 we outline the schedule for the validation and round robin activities. Chapter 4 includes some general remarks on unbiased validation and validation criteria. The following three chapters (5, 6, and 7) describe the planned validation activities for the three Glaciers_cci products in detail. They follow an identical substructure where first the test sites and the available validation data are introduced, followed by a detailed description of the validation strategy and the derived validation parameters. Then we outline the review process and decision sequence (8), and summarise the required resources (9) for potential participants in the validation and RR process. The next chapter (10) gives an overview how the validation activities of the Glaciers_cci consortium will be documented. Chapter 11 comprehensively specifies the round robin exercises, again following an identical substructure for each of the three Glaciers_cci products. The final two sections list the references (12) and acronyms (13) used in the document.
2. CCI project guidelines

In response to the CCI project guidelines, we have summarised in Table 2.1 how the validation activities for each of the products will meet those requirements.

<table>
<thead>
<tr>
<th>No.</th>
<th>Recommendation</th>
<th>Area</th>
<th>Elev. Change</th>
<th>Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-1</td>
<td>All CCI projects should use the definition of validation approved by the CEOS-WGCV</td>
<td>Validation for all products is compliant with the definition.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-2</td>
<td>All CCI project Product Validation Plans (PVP) shall adhere to the three</td>
<td>All three rules for independence of validation are adopted for each</td>
<td>Accepted statistical measures will be applied, including estimated biases</td>
<td>The detailed outline of the validation strategy as well as the results will be publically</td>
</tr>
<tr>
<td></td>
<td>requirements regarding independence as specified below.</td>
<td>product.</td>
<td>between independent sensors, commission and omission errors associated with</td>
<td>available.</td>
</tr>
<tr>
<td>V-3</td>
<td>The CCI teams shall use established, community accepted, traceable validation</td>
<td>Accepted statistical measures will be applied (mean, std dev.) along</td>
<td>Accepted statistical measures will be applied, including estimated biases</td>
<td></td>
</tr>
<tr>
<td></td>
<td>protocols where they exist. If such protocols do not exist then CCI projects</td>
<td>with overlay of outlines and omission / commission errors.</td>
<td>will be applied, including estimated biases between independent sensors,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>may adapt existing protocols if appropriate and in any event shall offer their</td>
<td></td>
<td>commission and omission errors associated with the elevation trend</td>
<td></td>
</tr>
<tr>
<td></td>
<td>final protocol for future community acceptance.</td>
<td></td>
<td>measurements, and calculation of the root mean square departure of the data</td>
<td></td>
</tr>
<tr>
<td>V-4</td>
<td>Each CCI project shall select appropriate validation data to ensure that an</td>
<td>High-resolution satellite data and multiple digitizing will be used</td>
<td>Airborne altimeter data will be used as an independent dataset. Elevation</td>
<td>Test sites were selected based on the availability of data suitable for validation.</td>
</tr>
<tr>
<td></td>
<td>adequate level of validation (confidence) is applied to all output products.</td>
<td>for product validation.</td>
<td>change maps produced using different algorithms/ sensor measurements will be</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The level of validation (confidence) should be indicated in the output product.</td>
<td></td>
<td>used for a cross validation.</td>
<td></td>
</tr>
<tr>
<td>V-5</td>
<td>The CCI programme should hold a dedicated session (or workshop) on common</td>
<td>This will be announced when the results of the RR and validation are</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>validation infrastructure during (or prior to) the next co-location meeting.</td>
<td>available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V-6</td>
<td>The PVP shall fully describe the validation process for each CCI project. An</td>
<td>We will involve the GLIMS community and the CRG, as well as</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>independent international review board of experts should be invited to review</td>
<td>international experts for each Glaciers_cci product to give feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>the PVP of each project team. Each CCI project should involve experts from the</td>
<td>on the validation. Peer-reviewed joint papers (for each product) are</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CMUG throughout their validation activities. A CCI product will be deemed to</td>
<td>planned to document the results.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>be validated once all steps of the validation process documented in the PVP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>have been completed and documented accordingly.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 2.1: Validation in Glaciers_cci compared to the overall project guidelines.*
3. Master Schedule

The planned schedule of the round robin and validation activities is summarised in Table 3.1. The round robin will be carried out for each of the three Glaciers_cci products, although the schedule is slightly different depending on the expected workload for generating the data products and performing the validation/RR (Table 3.1). Due to previously gathered experiences (e.g. GLIMS analysis comparison experiments, GLACE) the RR experiments for glacier area can follow a tighter schedule and are planned to be finished by March 2012. The lessons learned in the glacier area RR will be used for carrying out the RR for elevation change and ice velocity. The round robin for all products will first be carried out internally in order to prepare advanced data packages and to optimize the RR protocol before asking external participants. The various tasks and their schedule are listed in detail in Table 3.1. We expect to revise the schedule depending on the contributions of the participants. The results of the RR will be compiled in the PVSAR document and are the basis for algorithm selection.

<table>
<thead>
<tr>
<th>Task \ Date</th>
<th>December 2011</th>
<th>January 2012</th>
<th>February 2012</th>
<th>March 2012</th>
<th>April 2012</th>
<th>May 2012</th>
<th>June 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Datasets prepared for product validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance of database</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation/Inter-comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis and Reporting of results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Datasets prepared for RR tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RR protocol v0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal RR test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revised RR protocol v1 (D2.4: RRDP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.1: Proposed master schedule for the validation (top) and round robin activities (below the grey line). The different colours denote the three products: brown=glacier area, blue=elevation change, green= velocity.
4. Rules for unbiased validation and validation criteria

4.1. Glacier area

The glacier area product that requires validation has two components: (a) the result of any automated method to map glacier outlines, and (b) the manual correction of debris-covered glacier parts. Independent validation for both is guaranteed by two measures: (i) The datasets that should be validated will be processed independently from the data used for validation, and (ii) the participating people do not know each other. We will further ask for assistance in results evaluation from people that were not involved in the RR and validation. For the RR we will not prescribe the algorithm to be applied. This will be up to the participants. The consortium will, however, assess systematic effects of a specific algorithm, e.g. influence of the selected threshold value and the median filter on glacier size. For the selection of the most suitable algorithm we will use the results of the evaluation (as an objective measure), but also seek for advice from the wider community when results are less clear.

The evaluation of algorithm performance will be based on standard statistical measures like mean values and standard deviations for each glacier (as derived by the participants or from multiple digitising). The mean values are then used as a reference value for comparison with the result of an automated algorithm. The differences between the two will be calculated and compared to the standard deviation obtained in creating the reference data set. Apart from these objective criteria, we will also use more subjective measures like the overlay of all digitised outlines. This will help to identify regions with the largest variability in interpretation and thus provide important information for 'illustrated guidelines for the analyst' that we intend to prepare as a result of the RR and validation experiments.

4.2. Elevation change – DEM differencing

The validation for the elevation change product from DEM differencing is based on a cross comparison of several independently generated DEMs by the EO team. We do not refer here to directly measured cumulative mass balances, as their spatial interpolation to the entire glacier and conversion to water equivalents gives variable results. Most important issues to be considered before the DEMs can be compared are: (a) proper co-registration and (b) consideration of resampling effects that can introduce a systematic bias with elevation (Paul, 2008). A second issue is then to focus on stable and flat ground (regions off-glacier) to derive differences not influenced by glacier changes. Over glaciers we expect to see considerable differences due to strong overall melt in the past two decades. However, some of the DEMs compared (ASTER and LIDAR) were acquired in the same week and will allow us to assess also DEM related differences over glacier surfaces. As all DEMs used in this validation experiment have been generated independently, validation results should be unbiased. The validation criteria applied will consider elevation differences as a quantitative measure, but also analysis of hill-shades for detection of artefacts and DEM specific structures.
4.3. **Elevation change – Altimetry**

Pure elevation change validation requires at least two elevation measurements that need to coincide temporally with the satellite acquisitions. These observations, which need to be of known accuracy (preferably higher) and ideally with equal or finer spatial resolution, will be used to determine the elevation change to compare them with the satellite product. This is very rarely the case, and therefore product validation will also rely on inter-algorithm and inter-sensor comparisons. The product validation therefore focuses on three components: (a) comparison of temporally consistent airborne elevation changes with satellite altimeter elevation changes, (b) comparison of elevation changes derived from different sensors (e.g. radar vs. laser), and (c) comparison of elevation changes derived from different algorithms (cross-track vs. repeat track, cross-over versus DEM differencing, repeat-track versus DEM differencing). The validation of (a) and (b) will occur internally, while (c) will also be part of the round robin. The validation strategy (a) ensures independence, since it is based on external data which will not be used during generation of the elevation change products derived from satellite altimeter data. The validation strategy (b) provides two independent elevation change products to be compared, because even if the same algorithm will be used to create them, it will be applied on two different data sets acquired by two different sensors. The same applies for for the validation strategy (c), which will provide two independent elevation change products derived from two different algorithms to a data set acquired by the same sensor.

The validation criteria is based upon the computation of the RMSE and the correlation coefficient, R², which gives the degree of goodness of fit especially when two quantities are matched. This applies to the comparison between (a) airborne elevation changes and satellite altimeter elevation changes, (b) estimates of elevation changes using different satellite sensors, and (c) estimates of elevation changes obtained by applying different algorithms to the same sensor data set. The unbiased validation is achieved by the comparison of independent data sets: (a) elevation changes determined from satellite sensors and airborne observations, and (b) using multi sensor techniques, namely radar versus laser altimeters. Independence in the inter-algorithm comparisons (c) is achieved by approaching independent experts. They will likely provide (after the round robin) elevation changes that differ from each other even when the same algorithm is applied.

4.4. **Ice velocity**

The rules for unbiased validation and validation criteria are outlined in detail in section 7 where we discuss the inherent drawbacks of the method which complicate the validation of glacier displacements from space against independent data with equal or better resolution, accuracy and precision. We conclude that the validation of glacier velocity products from repeat optical and SAR imagery cannot be based on external data alone, but will also be validated internally (i.e. from the product itself and from product inter-comparisons). Different experts participating in the RR will guarantee unbiased validation.
5. Planned validation activities for glacier area product

5.1. Background

Validation activities for the glacier area product will basically be two-fold:

i. Comparison against a reference data set (i.e. higher-resolution satellite data acquired at about the same date), and

ii. multiple digitisations of the same glacier entities using manual digitisation.

To avoid conflict with interpretation of which parts actually belong to a glacier, the glaciers are specifically selected and clearly marked. However, a number of challenges in identifying the glacier boundary will also be considered for the validation. This includes ice under debris cover and in shadow (either casted by the terrain or clouds), as well as rock outcrops and seasonal snow. Method (i) is widely accepted as a means to test the accuracy of automated glacier mapping from Landsat-type satellite data and is thus in particular useful for debris-free ice. Method (ii) is required as well, because debris-covered glacier parts are in general manually corrected by visual comparison with a contrast-enhanced background image.

Most important, glaciers on high-resolution images can only be digitised manually, as automated mapping requires a spectral band in the shortwave infrared (e.g. Paul et al., 2002). Due to the variability in interpreting the outlines and generalization effects, manual digitisation provides each time slightly different results. The best means to assess this analyst-internal variability in interpreting a glacier outline is to digitize the same glacier several times and calculate mean values and standard deviations. The mean value will serve as a reference for validating the Landsat-type data sets. This will form a part of the RR by performing multiple digitisations of the same glacier by the same person and of the same set of glaciers by different persons, using both types of data (see overview in Table 5.1).

<table>
<thead>
<tr>
<th>Activity</th>
<th>Sensor</th>
<th>Purpose</th>
<th>Responsible</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of algorithm</td>
<td>Landsat</td>
<td>Round robin</td>
<td>GLIMS, CRG</td>
<td>algorithm selection</td>
</tr>
<tr>
<td>Influence of threshold</td>
<td>Landsat</td>
<td>Validation</td>
<td>EO Team</td>
<td>sensitivity test</td>
</tr>
<tr>
<td>Influence of noise filter</td>
<td>Landsat</td>
<td>Validation</td>
<td>EO Team</td>
<td>sensitivity test</td>
</tr>
<tr>
<td>Correcting outlines</td>
<td>Landsat</td>
<td>Round robin</td>
<td>GLIMS, CRG</td>
<td>Editing exercise</td>
</tr>
<tr>
<td>Multiple digitisations</td>
<td>High res./Landsat</td>
<td>Round robin</td>
<td>GLIMS, CRG</td>
<td>many-to-one test</td>
</tr>
<tr>
<td>Digitizing outlines ones</td>
<td>High resolution</td>
<td>Validation</td>
<td>EO/CRG/GLIMS</td>
<td>one-to-many test</td>
</tr>
</tbody>
</table>

Table 5.1: Overview of the validation and round robin activities to be performed by the EO team, the CRG and the wider GLIMS community for glacier area.

As part of the validation, the results of the automated methods for glacier mapping will be assessed against these validation data sets. In the round robin, we will compare the results of different algorithms for glacier mapping as applied by the participants and the manual
corrections applied in a sub-region of the test region. In another region we will also assess the sensitivity of the mapping results on thresholds and filtering for two pre-selected algorithms. To ensure useful results, the activities will be performed by different groups (the project EO team, the CRG, external participants) with validation data or algorithm results not revealed. For these regions only the final images (RGB composites) rather than raw data of individual image bands will be provided (in Geotiff format). All activities are summarized in Table 5.1.

5.2. Selected test sites and available validation data

Before we explain the validation strategy in more detail, we introduce the selected study regions. As product validation will be performed in sub-regions of the round robin and because the round robin will also provide validation data, we will already introduce the round robin test sites here as well. Some validation data sets will only be uncovered after the round robin to guarantee independence of the assessments. The selection criteria for the validation sites result from the two ways of validation that will be performed. The manual digitisation of selected glaciers will be based on the same satellite images used for the glacier mapping in the round robin. They do thus have optimal mapping conditions, but impose a number of challenges like debris cover and cast shadow. The high-resolution images are selected according to the agreement of the mapping conditions (if possible acquired in the same week) compared to the original satellite image. The selected regions will cover glaciers of different size and type from the Himalaya, the Alps, and Alaska. An overview of all test regions and their characteristics is provided in Table 5.2.

<table>
<thead>
<tr>
<th>ID</th>
<th>Region</th>
<th>Scene/Glacier</th>
<th>Sensor</th>
<th>Date</th>
<th>Who?</th>
<th>Challenges</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA1</td>
<td>Alaska</td>
<td>066-017 (RR)</td>
<td>TM</td>
<td>06.09.09</td>
<td>Val</td>
<td>shadow, water, debris</td>
<td>sensitivity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8 glaciers</td>
<td>Quickbird</td>
<td>27.08.03</td>
<td>RR</td>
<td>debris</td>
<td>digitize 1 x</td>
</tr>
<tr>
<td>GA2</td>
<td>Alps (AU, IT)</td>
<td>193-027 (RR)</td>
<td>TM</td>
<td>30.07.03</td>
<td>RR</td>
<td>all</td>
<td>digitize 3 x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glacier 1</td>
<td>Ikonos</td>
<td>not public</td>
<td>Val</td>
<td>contrast</td>
<td>digitize 5 x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glacier 2</td>
<td>Ikonos</td>
<td>not public</td>
<td>Val</td>
<td>contrast</td>
<td>digitize 5 x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glacier 3</td>
<td>Ikonos</td>
<td>not public</td>
<td>Val</td>
<td>contrast</td>
<td>digitize 5 x</td>
</tr>
<tr>
<td>GA3</td>
<td>Alps (CH)</td>
<td>Geren</td>
<td>Aerial</td>
<td>2010</td>
<td>RR</td>
<td>debris</td>
<td>digitize 1 x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Futschöl</td>
<td>Aerial</td>
<td>04.10.09</td>
<td>RR</td>
<td>shadow</td>
<td>digitize 1 x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d’Urezzas</td>
<td>Aerial</td>
<td>04.10.09</td>
<td>RR</td>
<td>debris</td>
<td>digitize 1 x</td>
</tr>
<tr>
<td>GA4</td>
<td>Kashmir</td>
<td>14B-035 (RR)</td>
<td>ETM+</td>
<td>04.09.00</td>
<td>RR</td>
<td>all</td>
<td>algorithm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>52S-670</td>
<td>PALSAR</td>
<td>12.07.07</td>
<td>Val</td>
<td>image</td>
<td>interpretation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.08.07</td>
<td></td>
<td></td>
<td>validation</td>
</tr>
</tbody>
</table>

Table 5.2: Overview of validation and round robin regions (RR) for glacier area. In the column ‘Who?’ RR = round robin participants and val = validation team. In the ‘Challenges’ column all means debris, shadow, snow and low contrast.

Outlines for the selected glaciers on these images will be digitised by the EO team and members of the CRG to guarantee independence. Orthorectified high-resolution image data will be taken from Google Maps and interactively geolocated (pixel size and upper left corner). We have also evaluated other independent vector datasets and decided to digitise the outlines ourselves. Otherwise we would only reveal differences in the mapping conditions and interpretation rather than the accuracy of the algorithm (Paul et al., 2011).
5.2.1. Chugach Mts., Alaska (test site GA1)

The test site in Alaska is located in the western Chugach Mountains to the north of the large Miles Glacier. It is covered by a Landsat TM scene (path 66, row 17) that was acquired under near-optimal mapping conditions (Figure 5.1). A sub-region covered by a high-resolution Quickbird scene was selected to create a validation data set (Figure 5.2). The participants will be asked to digitize the outlines of eight glaciers on this image once.

![Figure 5.1: Test region GA1 in Alaska was selected for testing algorithm sensitivity.](image1)

![Figure 5.2: Validation site in region GA1 (Alaska) showing the eight glaciers to be digitised once from each participant. Screen shot taken from Google maps.](image2)
5.2.2. Ötztal, Alps (test site GA2)

Glaciers in the Ötztal are selected mainly for multiple digitisation experiments using the Landsat image shown in Figure 5.3 (ten glaciers, to be digitised three times by the participants) and three glaciers on high-resolution Ikonos images (to be digitised five times by the EO team) that will be shown later, in Figure 5.4 and Figure 5.5.

![Figure 5.3: Test region GA2 in the Alps was selected for multiple digitization experiments.](image)

![Figure 5.4: Two glaciers from the validation site GA2 in the Ötztal Alps as seen by Ikonos. (The high-resolution data of the selected glaciers will be shown after the RR in order to accomplish blind tests).](image)

![Figure 5.5: One glacier from the validation site GA2 in the Ötztal Alps as seen by Ikonos, selected for “blind validation”. (The high-resolution data of the selected glaciers will be shown after the RR).](image)
5.2.3. Silvretta and Gotthard, Alps (test site GA3)

Three further glaciers located in Switzerland have been selected for product validation, one in the Gotthard region (Figure 5.6) and two in the Silvretta group (Figure 5.7). They are covered by aerial photographs and need to be digitized once by each participant.

Figure 5.6: Geren-glacier in the Gotthard region as seen on an aerial photograph acquired in 2010. Some fresh snow covers the highest mountain peaks. Screenshot from Google maps.

Figure 5.7: Vadret d’Urezzas (left) and Vadret Futschöl (right) in the Silvretta-group as seen on aerial photographs. Screenshots are taken from Google maps.

5.2.4. Kashmir, Himalaya (test site GA4)

The test region GA4 will be used for algorithm intercomparison and is depicted in Figure 5.8. Glaciers in this region have to be mapped by all participants with their algorithm of choice. For a small sub-region, participants will be asked to correct glaciers for debris-cover. Results will be compared against glacier areas as derived from PALSAR coherence images.
Figure 5.8: Test region GA4 for algorithm inter-comparison in the western Himalaya (Kashmir). For the right part of the image PALSAR coherence scenes are available.

5.3. Validation strategy

Validation of the glacier area product means to determine the differences in glacier area derived from the applied algorithm to a reference data set. This is actually rather challenging, as the reference data set must have been acquired under the same mapping conditions in the same year, it must have a much better spatial resolution and, as a best option, must be derived with the same method. For high-resolution satellite data or aerial photography, the latter constraint cannot be fulfilled as these sensors do not have SWIR bands. Hence, the delineation has to be done manually and this gives results that are not reproducible. A data set used for validation should thus be a mean from multiple digitisations. The validation strategy considers these challenges by implementing single and multiple digitisations for several individual glaciers on very high (e.g. Ikonos-type) and high (e.g. Landsat-type) resolution images. This will be performed for glaciers of various sizes, in different regions and with different characteristics (debris, shadow).

A second part of the validation strategy is related to the overlay of the outlines as derived from different methods, sources and analysts. Though this is only a qualitative measure, it is the most important first step in product evaluation. It must be applied to check if all analysts have mapped the same glacier entities. But it is also the measure to determine common mapping problems, for example in the interpretation of debris-covered regions or those in cast shadow. We will thus give this part of the validation the same weight as the comparison of numbers. Figure 5.9 shows the components of the validation strategy schematically.
5.3.1. Product validation with reference data

We will create reference dataset for several regions, partly by the EO team, and partly by the CRG and the participants of the round robin. They will consist of (cf. Table 5.2):

a) digitisation (once) of 8 glaciers by the participants in region GA1 (Alaska) using Quickbird
b) multiple digitisations (three times) of glacier outlines (10 selected glaciers) by the participants in region GA2 (Ötztal) based on FCCs from Landsat
c) digitisation (once) of 3 glaciers by the participants in region GA3 (Switzerland) based on aerial photography
d) multiple digitization (three times) of 3 glaciers by the EO team and the CRG in region GA2 based on Ikonos images, and
e) delineation of glacier outlines for region GA4 (Kashmir) by the EO team using the Landsat data and PALSAR coherence images to improve the debris covered regions

For each region and dataset, mean glacier areas and standard deviations will be computed. Of course, all outlines will also be subject to a relative comparison by overlay of the outline. The glacier outlines that are created by the round robin for algorithm selection and sensitivity testing will then be compared against the validation data in a quantitative (area comparison) and qualitative way (outline overlay). The main region specific assessments are:

A) Test of the impact of the threshold value and median filter in Alaska (GA1) by the EO team
B) validation of raw and corrected algorithm results in GA2 (Austria) and GA3 (Switzerland) by the EO team, and
C) comparison of the algorithms applied by the participants in GA4 (Kashmir).

5.3.2. Relative inter-comparisons

The relative inter-comparisons (overlay of outlines) will be applied to all datasets described in section 5.3.1 (a) to (d) and (B) to (C). It is assumed that in many cases differences in the manual delineation will drive the differences rather than shortcomings in the automated mapping. Hence, these comparisons will provide the lessons to be learned. In a quantitative sense, we will also calculate omission and commission errors following standard approaches (e.g. Gjermundsen et al. 2011).
6. Planned validation activities for elevation change products

This section describes the validation plan for the two elevation change products to be generated within the project. The approaches for elevation change retrieval are (1) DEM-differencing and (2) elevation changes from satellite altimeter data.

(1) DEM-differencing

The validation of glacier elevation changes derived from space-borne DEMs is, in principle, challenging due to a lack of temporally coincident high-accuracy data (e.g. airborne LIDAR DEMs). Hence, an external validation consisting of a comparison between elevation changes derived from medium resolution/accuracy DEMs (from space-borne platforms) vs. elevation changes derived from high resolution/accuracy DEMs (from airborne platforms) is not practically feasible. We have thus developed a validation strategy that should answer the following three main questions:

What is the accuracy and precision of each specific DEM?

The accuracy and precision of any individual DEM can be assessed by comparison to a higher resolution/accuracy product if available. If this product is not temporally consistent, the comparison must be made using stable terrain (e.g. terrain assumed not to change with time). In general, the relative accuracy within a DEM will be improved with co-registration before comparison. The main goal here is thus to determine the relative internal precision of the specific DEM product (e.g. ASTER, SRTM, SPOT5-HRS). An additional component is to determine how topographic derivatives such as slope and aspect change with the DEM. This is especially important for calculating glacier-specific attributes from DEMs with known artefacts (e.g. ASTER GDEM).

What is the accuracy and precision of co-registering a DEM pair?

The main question here is to determine how well we can resolve and correct for linear horizontal and vertical discrepancies between the two DEMs. This may vary for each unique DEM pair, but is most likely related to specific products (SRTM, ASTER, etc). In addition, at least two methods are being used for co-registration and a validation of the co-registration entails comparing the methods, and checking specifically how the mean vertical adjustment varies between the approaches. This mean vertical bias may be one of the largest unaccounted biases that exist in glacier elevation changes.

Are there any potential scene/sensor/product specific biases that decrease the accuracy and can they be corrected for?

This component is likely the most difficult to resolve, but might be an important error source for glacier elevation changes. For example, there is significant discussion on whether an elevation dependent bias may occur for specific DEMs. Also, internal geo-location biases within satellite stereo-scenes caused by unresolved shaking of the instrument, jitter (LePrince et al., 2007) have also shown to transfer into the DEMs derived from those images (Berthier et al., 2007; Nuth and Kääb, 2011). All of these internal biases should be investigated to determine their potential error contribution.
(2) **Altimetry**

Validating the elevation change products derived from satellite repeat altimeter data requires, for both the cross-over and the repeat track algorithms, independent surface elevation measurements of at least two epochs in time coinciding with the altimeter measurements.

Validation data can be collected in-situ via fieldwork or it can be obtained from remote sensing. In the present validation plan, only the latter data will be considered and in particular airborne data will be used. The main reason of this choice is due to the fact that the accuracy of pulse-limited altimeters deteriorates with increasing surface slope. Thus, in-situ surface elevation measurements from small, easy to reach glaciers with high slopes are rarely useful in product validation and in-situ surface elevation measurements from the flat, high-elevation area of large Arctic ice caps are necessary.

For this purpose, the CryoSat Validation Experiment Campaign (CryoVEX-campaign) and the NASA Operation IceBridge mission provide airborne altimeter measurements. As regard as the first campaign, the datasets can be requested at the address ftp://ftp.cryosat.esa.int, whereas the NASA airborne data are freely downloadable at the web-address http://nsidc.org/data/icesat/index.html. The CryoVEX campaign operates both with a laser scanning altimeter (ALS) and with a radar altimeter system (ASIRAS). The NASA ‘Operation IceBridge’ mission flies several instruments, of which the most useful for the elevation change products validation are the Airborne Topographic Mapper (ATM) and the Airborne Topographic Mapper 3 (ATM3). They are scanning lidars developed and used by NASA for observing the changes of the Arctic and Antarctic icecaps and glaciers. Furthermore, the Laser Vegetation Imaging Sensor (LVIS), which is a scanning laser altimeter originally optimised for vegetation science, later also used in monitoring the vertical movements of ice masses due to the capacity of the instrument to use the so-called waveform-based measurement technique, which makes it especially useful for measuring extremely rough ice surfaces by analysing the shape of the returned pulse. And finally the Ku-Band Radar Altimeter, which is a radar whose generated signal penetrates through snow to measure the ice surface elevation. The validation will also be performed by the comparison of the elevation change products obtained using the different algorithms described in the ATBDv0 (Glaciers_cci, 2012) and also by the comparison of radar and laser altimeter products. Both of these comparisons provide information that is useful for the validation process. In case that further datasets will become available during the the project, they can also be used for the assessment of the elevation change products.

### 6.1. Selected test sites and available validation data

For glacier elevation changes, availability of suitable validation data is the largest constraint of the validation process. Thus, for altimetry the elevation change product validation will be carried out on several targets meeting the conditions of a suitable validation area - including at least Devon Ice Cap and Austfonna. For DEM differencing, we therefore restrict our validation set to one region where higher accuracy data is available. For the DEM differencing round robin, we choose a representative alpine glacier setting where all global datasets are available.
6.1.1. Devon Icecap (test site EC1)

The Devon Ice Cap is a large ice cap on Devon Island in Nunavut arctic Canada at 75° N, 82° W (it is represented by the black area in Figure 6.1). Covering an area of 13’700 km², the Devon Ice Cap is one of the largest ice caps on Earth. It consists of a 11’700 km² dome-shaped main ice cap and a 2000 km² western arm that is stagnant and dynamically-separate (Dowdeswell et al., 2004). The main ice cap has a maximum elevation of 1921 m and a maximum ice thickness of 880 m.

![Figure 6.1: Map showing the location of Devon Island and the Devon Ice Cap.](image)

Over this region, many laser and altimeter acquisitions are available. As an example, the left panel of Figure 6.2 shows the ground tracks of the radar altimeter (RA2) on board of ENVISAT during the period of May-June 2004 (ENVISAT worked on the same orbit of ERS-1 and ERS-2 until 22 October 2010, thus the ground tracks of ENVISAT are approximately overlapping the ground tracks of ERS-1 and ERS-2), whereas the right panel of Figure 6.2 shows ground tracks of the ICESat GLAS altimeter.

During the same period, CryoVEX airborne data were acquired. The left panel of Figure 6.3 shows the flown tracks of the Polar 4 aircraft of AWI with the Airborne Synthetic Interferometric Radar Altimeter System (ASIRAS) on board (in red), while the right panel of Figure 6.3 shows another acquisition collected on September 2004. In the same figure also the Differential Global Positioning System (DGPS) profiles are reported (in blue).
The CryoVEX surface elevation data are available from Devon Island for the years 2004 and 2006. Table 7 in the DBT2 Technical Note (Glaciers_cci, 2011) summarises the CryoVEX dataset, including date of acquisition, notes and the link where the data can be downloaded.

Additionally, Figure 6.4 shows the locations of the precise airborne laser surveys carried out over the past 20 years by NASA P-3. These surveys were repeated during the spring 2000 on a commercial Twin Otter platform. The survey trajectories were designed to cross the broadest and longest portions of the major ice caps. The DBT2 Technical Note (Glaciers_cci, 2011) reports on the dates of the NASA airborne campaigns.
6.1.2. Austfonna Icecap (test site EC2)

Austfonna is an ice cap covering 8120 km² of Nordauslandet (Figure 6.5), the northernmost island of the Svalbard archipelago at 79° N and 24° E (Dowdeswell, 1986). It is the fourth largest ice cap on Earth and the largest on Svalbard (Dowdeswell and Hagen, 2004).

Over this region, many radar altimeter acquisitions are available. The left panel of Figure 6.6 shows an example of the acquisitions carried out by the radar altimeter (RA2) on board of Envisat during the first 20 days of April 2004 (Envisat worked on the same orbit of ERS-1 and ERS-2 until 22 October 2010, thus the ground tracks of Envisat are approximately overlapping the ground tracks of ERS-1 and ERS-2). There are also a considerable number of data acquired by the ICESat altimeter from 2003 to 2009. The right panel of Figure 6.6 shows an example of the ground tracks of this satellite over Nordauslandet.

The total number of measurements varies considerably within each year, mainly due to the meteorological conditions and the fact that the signal absorption in optically thick clouds causes some profiles to be incomplete. Figure 6.7 shows the ICESat observation campaigns and the temporal distribution of data over glacier terrain in Svalbard between the 2003-2008 observation campaigns in Feb./Mar., May/June and Oct./Nov. There are most data from the winter campaigns and least data from the summer campaigns, reflecting the meteorological circumstances in Svalbard with more cloud cover in summer than winter.
Figure 6.5: Map showing Nordauslandet with the Vestfonna and Austfonna Ice Cap.

Figure 6.6: Envisat RA2 ground tracks from 01/04/2004 to 20/04/2004 (left) and ICESat GLAS ground tracks (right) over Nordauslandet and the Austfonna Ice Cap.

Figure 6.7: Total number of ICESat footprints in each acquisition campaign within the entire Svalbard archipelago. Figure modified from Moholdt et al., 2010.
CryoVEX airborne data were acquired over the region since 2003. The left panel of Figure 6.8 shows the ground track of the 2003 campaign, whereas the central and right panel show the flown tracks of the Polar 4 aircraft of AWI with the Airborne Synthetic Interferometric Radar Altimeter System (ASIRAS) on board (in red). In the same figures also the Differential Global Positioning System profiles are reported (in blue).

Table 8 in the DBT2 Technical Note (Glaciers_cci, 2011) summaries all the CryoVEX dataset over the Svalbard archipelago, including date of acquisition, notes and the link where the data can be downloaded. During the springs of 1996 and 2002, ice-surface elevation measurements were acquired using the ATM3 (Krabill et al., 2000). The flight lines over Svalbard are shown in Figure 6.9. The DBT2 Technical Note (Glaciers_cci, 2011) reports the data of the NASA airborne campaigns.

Figure 6.8: Flight tracks of CryoVEX -2003 (left) and CryoVEX 2004, 20th of April (middle) and 25th April 2004. DGPS (blue) and ASIRAS (red).

Figure 6.9: The Svalbard archipelago, showing the IceBridge flight lines (ATM3) flown during the spring of 1996 and 2002. 23rd and 24th May 1996 = yellow; 20th and 22nd and 23rd May 2002 = blue.
6.1.3. Findelen Glacier (test site EC3)

The test site Findelen glacier is located in the southern part of Switzerland at the border to Italy. The glacier is about 16 km$^2$ in size has a comparably flat accumulation region and a well developed tongue. The glacier forefield has a flat sedimentary basin that can be used for validation purposes (Fig. 6.10). This glacier was selected as mass balance measurements (with the direct glaciological method) were performed for several years and high-resolution DEMs from several sources (e.g. satellite data from ASTER and SRTM, air-borne data from photogrammetry and LIDAR, see Fig. 6.11) and dates (from 1985 to 2010) are available.

Figure 6.10: Findelen glacier as seen on Oct 2, 2011 by the automatic camera that is operated by the University of Fribourg. The snow line was extremely high at that date.

Figure 6.11: Hillshade of the LIDAR image of Findelen and adjacent Adler glacier acquired in Oct 4, 2009.
6.1.4. Southern Alps, New Zealand (test site EC4)

A region in the Southern Alps is chosen for the elevation change RR from DEM differencing. The main test glaciers that typically fall within an acquisition scene are Tasman Glacier, Murchison Glacier Franz Josef Glacier, and Fox Glacier (Fig. 6.12). These glaciers are high-elevation glaciers and experience large variability. The site is chosen, because of its representativeness as a high-alpine glacier type and the availability of global elevation datasets like SRTM and ASTER.

![Figure 6.12: ASTER image of the study site in New Zealand, blue lines denote recently derived glacier outlines (Gjermundsen et al. 2011). The largest ice mass in the northern centre of the scene contain Fox, Franz Josef, Tasman and Murchison glaciers.](image)
6.2. Validation strategy and parameters

The following section is sub-divided due to the different approaches used for generating the Glaciers_cci elevation change products.

6.2.1. DEM differencing

The validation of glacier elevation changes from DEM differencing is complicated by the coincident availability of both satellite and higher-accuracy DEMs. For most glaciers and glacier regions, the availability of higher accuracy and precision data is greatly limited. Therefore, the pure external validation of glacier elevation changes derived from medium resolution/accuracy DEMs with those generated from high resolution/accuracy DEMs is not possible for our study sites, and are thus excluded from further investigation. However, availability of a single high resolution/accuracy DEM (e.g. airborne LIDAR DEM) provides an alternate external validation strategy by comparing the individual satellite DEMs with higher accuracy and precision for validating the raw elevations rather than the elevation changes. If the data sets are temporally consistent, they can also be compared over glaciers. Most cases, however, will involve comparisons over stable terrain (e.g. temporally constant; off-glacier). It is also of importance to develop a validation plan for situations where high resolution/accuracy data is not available. The internal validation strategy involves comparison between three or more datasets, also using stable terrain outside of the glaciers.

6.2.1.1. Co-registration (pre-processing)

For all validation purposes, an initial pre-processing step of co-registration is required to remove any linear bias resulting from the geographic mis-match between two DEMs or elevation products. In the following, co-registration refers to the following procedures:

- Selection of stable terrain (all terrain outside glaciers) using the available glacier masks.
- Determine the 3D co-registration vector between the two DEMs using only stable terrain following the algorithm description in the ATBDv0 (Glaciers_cci, 2012), this may require multiple iterations.
- One DEM is shifted using the 3D co-registration vector to the reference DEM.
- The two DEMs are then differenced, where the standard deviation ($\sigma$) of elevation differences over stable terrain represents the random error. The linear vertical bias is (should be) zero after co-registration.

6.2.1.2. External validation

The external validation involves verification of a satellite DEM using a high accuracy DEM. The process begins with co-registration of the datasets and the resulting estimate of the random error of the satellite DEM. If the data is temporally consistent, the comparison can be made over the glacier to analyse/detect any glacier specific biases related to the acquisition strategy (i.e. radar wave penetration into snow/firm within the SRTM DEM). In most cases, however, the analysis will only be made over stable terrain (off-glacier terrain). After co-registration, visual analysis of the changes over stable terrain is performed to detect any internal scene biases that may exist. If detected, procedures for removal will be investigated. This step is however case specific and can therefore not be standardized such as the co-
registration pre-processing step. As a final step, the error budget of an elevation change between two medium resolution/accuracy DEMs can be quantified, in principle, by the combination of errors of the individual DEMs vs. the high resolution/accuracy DEMs.

Alternate components of the external validation involve comparing the topographic derivatives of slopes and aspects over glaciers. To perform this, we will calculate the variability and mean difference between topographic derivatives of the medium resolution DEMs (SRTM, ASTER, ASTER GDEM, etc.) and the high resolution DEMs. Of importance is to further understand how potential elevation changes are affecting the topographic derivatives, especially when incorporating them into the GLIMS database.

6.2.1.3. Internal validation

Internal validation refers to the quantification of random and systematic errors in elevation differences between the DEMs using the DEMs themselves. In principle, this is not different from the external validation, except that there may not be one DEM that is more accurate than the others. The two multi-temporal DEMs are first co-registered to remove a potential systematic linear bias (horizontal and vertical), as well as to provide an estimate of the random error. A third elevation dataset is then incorporated, typically of higher accuracy. This may be either ICESat if the distribution of acquisitions is sufficient over the DEM or national DEMs. Each of the individual DEMs is then co-registered to the third dataset. This process returns three co-registration vectors between the data products that should form a perfect triangle (vector sum). Any mismatch in a vertex of the triangle is an estimate of the remaining un-removed bias. In practice, this un-removed linear vertical bias can be added to the error budget for elevation change. In summary, the resulting parameters of internal validation are an estimate of the random error and an un-removed linear vertical systematic bias. Last, visual analysis of the changes over stable terrain is performed to detect any internal scene biases that may exist. If detected, then procedures for removal will be investigated. This step is also case specific and can therefore not be standardized.

6.2.1.4. Internal versus external validation

This part of the validation strategy is aiming to determine whether the quantified accuracy of elevation differences generated solely from satellite data (internal validation) is within the true accuracy as determined by comparisons with the high-accuracy DEMs (external validation). In particular, this test will ensure that the error budget as derived from the internal validation contains the magnitude of the true errors derived. It is particularly important to determine and analyze any bias over the glaciers in relation to the bias estimated using only the terrain.

6.2.2. Repeat altimetry (cross-over and repeat track)

The validation of the elevation change products derived from satellite repeat altimetry will be performed in two ways: 1) intercomparing the products obtained using different algorithms and/or sensors, and 2) using independent measurements.
6.2.2.1. **Intercomparison of the elevation change products**

This method consists in the intercomparison of elevation change products that result from elevation change maps, which can be generated using:

- different algorithms, that is cross-over, repeat track and DEM differencing
- different sensors or sources, that is laser and radar altimeter products.

The comparison will be based on the absolute differences of the two elevation change maps using the two different methods and/or sensors. The quantitative analysis will be provided by means of the root mean square error, RMSE, and the correlation coefficient, $R^2$.

When the two different sensors, radar versus laser, are used for the validation process, a number of challenges will occur mainly due to the fact that the two datasets need to be collected during the same time/period. In fact, the laser altimeter data are temporally sparser compared to the radar ones and in addition they are only available during the cloud free days of the operation periods. Then, laser and radar measurements will not always refer to the same location, but to two locations close to each other. Finally, if the snow surface is dry, laser and radar measure a different surface elevation due to the penetration of the radar beam, whereas in case of a snowfall or rapid densification of the upper snowpack due to liquid water, the laser measurement will be misleading due to its insensitivity to the density and moisture of the snowpack. As a consequence, the RMS value of the difference between radar and laser product can be interpreted as the maximum uncertainty of both instruments.

6.2.2.2. **Validation of the satellite altimeter products with airborne data**

For this comparison, the airborne validation data need to coincide temporally with the satellite overpasses and this is the largest constraint of the validation process. Once these data are available, they need to undergo the same processing applied to the satellite data in order to obtain elevation change values suitable for a direct comparison (this will include also the same projection):

- the airborne measurement needs to be processed in order to obtain the elevation value;
- the cross-point locations need to be located in order to apply the cross-over algorithm;
- the application of the repeat-track algorithm requires the resurvey of the same area at a different time. As the flight lines cannot repeat exactly, cross-track offsets should be accounted for so that the measurements are more directly comparable.

The comparison will be obtained making the absolute difference of the two elevation change products using the two different measurement techniques. The RMSE will be computed together with the correlation coefficient, $R^2$, which will provide a degree of goodness of the fitting between satellite elevation change estimations with respect to the airborne estimations. In addition, any airborne dataset should have the precision and the accuracy of the radar and laser measurements.
7. Planned validation activities for glacier velocity products

In this section the validation plan of glacier velocity products from time series of SAR and optical satellite data is given. Glacier motion retrieval from space-borne sensors is characterised by some inherent methodological drawbacks, which complicate the validation of glacier displacements from space against independent data with equal or better resolution, accuracy and precision. The main reasons are:

1. **Coincident observation of EO and validation data:** Glacier motion often follows diurnal to seasonal cycles and year-to-year variations, among others as a consequence of the varying sub-glacial hydrology, therefore glacier motion is highly variable temporally, at scales from hours to seasons and years. A strict validation of glacier velocity products would therefore require simultaneous acquisitions of product generation and validation data, which can only be achieved by continuous ground measurements.

2. **Adjusting spatial scales:** Glacier displacement measurements from repeat images require image windows to be compared, i.e. the motion of feature ensembles rather than single features is estimated. Therefore, the derived displacement is not representative for a certain finite point, but rather for an area. Further, this representativeness is not a strict analytical function of the real displacement field, but a statistical relation of it, its gradients, image features and contrast, as well as the tracking algorithm and its implementation.

3. **Observation of different velocity components:** Depending on the sensor (SAR, optical) and applied method, different components of the true 3D velocity are observed. For example, SAR interferometry is sensitive to the Line-Of-Sight (LOS) velocity component of the ice, but it is in-sensitive to along track motion. Image Cross correlation techniques using SAR data measure displacement in LOS and along track direction, but are also sensitive to net elevation changes (e.g. melting, vertical velocity components). In addition, the movement of the ice particles on a glacier does not follow the surface. In general, there is submergence in the accumulation area and emergence in the ablation area. Additionally, in summer, changes of the ice surface motion of glaciers are a combination of ice displacement and surface melt. In the field, ice velocity is measured with stakes at various depths or with continuous GPS on the surface. In order to validate and/or compare the products from various methods, the transformation to the same velocity component is a pre-requisite.

Further, it should be noted that image cross-correlation strictly provides displacements for the time period between the acquisitions used. Thus, the glacier velocity product is the mean velocity over the observation period, and does not take any velocity variations between the image acquisitions into account. This fact is of particular importance when analysing time series of glacier velocities.
In general, the validation of glacier velocity products from repeat optical and SAR imagery cannot be based on external data alone, but will also be validated internally (i.e. from the product itself and from product inter-comparisons). In addition to these quantitative tests for product validation, we will consider a qualitative visual interpretation of the derived velocity field by a glaciologist: These checks are subjective, but will rely on expert knowledge based on basic physical laws such as the incompressibility of ice and will include checks for unnatural outliers or other features in the flow field, checks for coherence and consistency, as well as checks for unnatural patterns and flow direction.

7.1. Selected test sites and available validation data

In this section the validation plan of ice velocity products from time series of SAR data and optical satellite data is given. Based on data availability, the test sites listed in Table 7.1 have been selected for validating the glacier velocity products.

<table>
<thead>
<tr>
<th>ID</th>
<th>Glacier (Region)</th>
<th>Applied technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>Breidamerkurjökull (Iceland)</td>
<td>InSAR  SAR - OT   Optical-OT In-situ</td>
</tr>
<tr>
<td>V2</td>
<td>Kronebreen (Svalbard)</td>
<td>X                  X                  X</td>
</tr>
<tr>
<td>V3</td>
<td>Austfonna/Vestfonna (Svalbard)</td>
<td>X                  X                  X                  X</td>
</tr>
<tr>
<td>V4</td>
<td>Baltoro glacier (Karakoram) (RR)</td>
<td>X                  X</td>
</tr>
</tbody>
</table>

Table 7.1: Overview of data availability for glacier velocity (OT – Offset tracking).

7.1.1. Breidamerkurjökull, Iceland (test site V1)

As a test site for validation of ice motion fields from SAR and optical data we selected the glacier Breidamerkurjökull, Iceland. Breidamerkurjökull is an outlet glacier of the ice cap Vatnajökull which is located at 64.13° N and 16.3° W, and has an elevation range from 200 m a.s.l. to approx. 2000 m a.s.l. The glacier drains from the plateau of the icecap towards the South. The central part of the outlet calves into the terminal lake Jökulsárlón.

Three permanent GPS stations have been installed by the University of Iceland for monitoring the ice motion continuously and can be used for product validation within the Glaciers_cci project. The approximate location of the 3 GPS stations is indicated as red circles in Figure 7.1. They are located in the lower part Breidamerkurjökull, close to the centre flow line.

The GPS stations BMJ3 and BMJ5 are devices comparable to consumer handheld GPS systems and are less accurate than BMJ2. For these data sets a manual quality control and averaging of data with a moving window of two days was performed to reduce noise and improve accuracy (E. Magnússon, personal comm., Jan. 2010). Due to the special set-up, GPS station BMJ2 is more accurate than the two others. The measured data are processed by means of a differential GPS (DGPS). After DGPS processing a one-hour moving-average filter was applied to the 15 seconds GPS acquisitions to reduce noise and improve accuracy. The time series of filtered GPS positions will be used to calculate the ice displacements for the time intervals corresponding to the EO data used for velocity extraction (Figure 7.2).
Figure 7.1: Overview of the validation site Breidamerkurjökull, Vatnajökull ice cap, Iceland. Red circles indicate the approximate location of the GPS stations. The inset shows the location of Breidamerkurjökull in the south-eastern part of Vatnajökull.

Figure 7.2: Example of the derived magnitude of ice motion from GPS measurements for the period August 2008 to February 2009. The magnitude of ice motion is filtered with one hour and 12 hour moving average filter.

In addition to the ground based GPS data, a series of very high resolution X-band SAR data from the German TerraSAR-X satellite is available to the Glaciers_cci project. The series comprises four images acquired between 4th August 2008 and 6th September 2008 with an 11 day repeat cycle. In agreement with the German Aerospace Center (DLR), the data can be used for validation purposes as well as for the Glaciers_cci round robin experiments.
7.1.2. Kronebreen, Svalbard (test site V2)

The Kronebreen glacier system is chosen as a validation site for all velocity algorithms from both radar and optical. The glacier is located in northwest Svalbard and is ~50 km long with an elevation range 0–1400 m a.s.l and an area of ~390 km². Kronebreen is one of the fastest flowing glaciers in Svalbard (non-surging) with speeds approaching 2-3 m day⁻¹ during summer (Kääb et al., 2005).

Continuous GPS measurements are available from September 2008 in the heavily crevassed parts of the glacier (Figure 7.3). The GPS system is code-based and typically acquires one measurement per hour (den Ouden et al., 2010). Therefore, temporal smoothing (i.e. 3 days) is typically required for significant results. Approximately 3-5 units may be operational within the period September 2008 up to the present. The variation in availability of data is due to the loss of units into crevasses. Above the heavily crevassed zone, one code based GPS has been operational since May 2008 (Figure 7.4). This station has been recently replaced with a differential GPS unit.

Figure 7.3: The location of the Kronebreen in Svalbard. The locations of continuous GPS stations are shown as black squares. In the background is an ASTER image of the region.
7.1.3. Nordaustlandet, Svalbard (test site V3)

As an additional test-site for validation of ice motion fields from SAR data and for the round robin exercise we selected Nordaustlandet. Nordaustlandet is the second largest island over Svalbard and has two major ice caps: Austfonna and Vestfonna. These ice caps represent one of the largest ice-covered areas in the Eurasian Arctic. Even though low temperatures and low balance gradients generally result in low flow rates on the glaciers of Svalbard, the Austfonna and Vestfonna ice caps are also characterized by fast-flow regions (velocities over several hundreds of meters per year) at several outlet glaciers.

During 2007, a geodetic campaign was launched on the Svalbard ice cap Vestfonna in order to estimate the velocity field within the frame of the IPY project KINNVIKA (Pohjola et al., 2011). The geodetic measurements were done both as static surveying of stakes and as continuously recording GPS stations. In particular, as a part of a mass balance survey programme, 35 mass balance stakes were drilled on Vestfonna in 2007 and maintained to 2010 during repeat spring campaigns using snow mobile transports (Figure 7.5). Pohjola et al. (2011) published the horizontal ice surface velocity component from the geodetic survey campaigns in 2007–2010, which serve as a benchmark dataset towards an integrated assessment of the flow and mass balance of Vestfonna and can be used within the Glaciers_cci for a validation of products and of the round robin results.

As a further validation data set for Nordaustlandet we will consider a surface velocity map computed with ERS-1/2 InSAR Tandem scenes acquired between December 1995 and January 1996 (see Figure 7, courtesy EU INTEGRAL project). InSAR processing of this data set yielded a comprehensive map of the ice velocities over most of the ice caps with errors assumed to smaller than about 2 cm/day (or 7 m/year). Use of this data set for the round robin exercise was discarded because all the current and future-planned SAR sensors with a global acquisition strategy cannot provide these short-time intervals required to maintain the interferometric coherence over ice.
Figure 7.5: Svalbard, Nordaustlandet and the ice cap Vestfonna. Elevation- and coast line contours are from the Norwegian Polar Institute DEM. The elevation contour spacing is 500 m. The black dots mark the position markers used for the DGPS surveys. The triangular symbols are the fixed points/base stations. The letter at each fixed point/base station refers to the first letter of its name. Inset (a) show the base stations and inset (b) show the topography of all the ice caps on Nordaustlandet (from Pohjola et al., 2011).

Figure 7.6: Velocity map derived from ERS-1/2 1-day repeat InSAR (Winter 95/96), which has been compared to geodetic measurements in the framework the ESA GlobGlacier project.
Test runs of the 2008 ALOS PALSAR data sets showed, that these image pairs are appropriate for the round robin exercise, providing both a dense coverage of velocity information and good performance with respect to the GPS survey. Additionally to ALOS PALSAR, other SAR data are considered for the round robin (compare to Table 11.2): an ENVISAT ASAR image pair with a time interval of 35 days and two ERS-1 image pairs with a 12 days repeat cycle which correspond to the orbit cycle of the upcoming Sentinel-1 SAR sensor. Even if preliminary results obtained within GlobGlacier with these two C-band sensors and time intervals are of lower quality with respect to the ALOS PALSAR or the ERS-1/2 InSAR results, they are in our opinion very valuable to investigate the performance of an algorithm within non-optimal conditions.

Dunse et al. (2011) present continuous GPS observations along the central flow-lines of the two fast flowing outlet glaciers Duvebreen and Basin-3 on Austfonna over the 2008–2010 period (Figure 7.7 left). The data show prominent summer speed-ups with ice-surface velocities as high as 240 % of the pre-summer mean (Figure 7.7 right). Also these published horizontal annual mean velocities for the periods May 2008–2009 and May 2009–2010, either based on all available daily values or the first and last day’s position at the beginning and end of the respective year, can be used within the Glaciers_cci for a validation of products and of the round robin results.

Table 10 and 11 in the DBT2 Technical Note (Glaciers_cci, 2011) summarise all data available for velocity production and validation for the test site Austfonna/Vestfonna, Svalbard.
7.1.4. Baltoro glacier, Karakoram (test site V4)

The Baltoro glacier in the Karakoram region (Pakistan) comprises numerous tributaries, has a length of 63 km, and approximately 70% of the glacier is debris-covered. It is chosen as a round robin site because of its representativeness for the alpine glacier type, debris cover, velocity and available data for inter-algorithm comparison (Quincey et al., 2009).

Figure 7.8: Location of Baltoro glacier in Pakistan and the numerous names of the tributary glaciers that make up Baltoro glacier (adapted from Quincey et al. 2009).

7.2. Validation strategy

Due to the problems described above, we will follow a three-step validation strategy with an increasing level of “validation quality”:

- Internal method evaluation
- Method/sensor cross-comparisons
  - Comparison of ice velocity products from different methods (same data but different algorithms) and different sensors (same algorithm but different data)
- Validation with in-situ observations

Table 7.2 gives an overview about the validation activities at the different test sites.

<table>
<thead>
<tr>
<th>ID</th>
<th>Glacier (Region)</th>
<th>Quality control (SNR, CC)</th>
<th>Validation methods</th>
<th>Internal</th>
<th>Comparisons</th>
<th>in situ</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>Breidamerkurjökull (Iceland)</td>
<td>X</td>
<td>VT1 VT2 VT3 VT4 VT5</td>
<td>X X X X X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>Kronebreen (Svalbard)</td>
<td>X</td>
<td>VT1 VT2 VT3 VT4 VT5</td>
<td>X X X X X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>Nordaustlandet (Svalbard)</td>
<td>X</td>
<td>VT1 VT2 VT3 VT4 VT5</td>
<td>X X X X X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>Baltoro glacier (Karakoram)</td>
<td>X</td>
<td>VT1 VT2 VT3 VT4 VT5</td>
<td>X X X X X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.2: Overview of product validation activities at test sites VT1-VT5 (see section 7.2).
7.2.1. Internal method evaluation

Most matching algorithms record measures describing the degree of similarity between the matching image windows, e.g. the magnitude of the correlation coefficient (CC) and the signal-to-noise ratio (SNR). While the CC depends on the matching Kernel, the SNR is the ratio of mean to standard deviation of a measurement (for further details see Glaciers_cci, 2012). These parameters are an indication for the reliability of an individual match.

The matching CC or SNR of individual displacement vectors (both optical and microwave) can be analysed for defining a threshold and filter the data accordingly. However, this threshold has to be set for every individual image set (varying ground conditions) and no global strategy can be employed for this task.

Additionally, the measure is not strict, i.e. bad matches might actually accurately reflect the true displacement, or vice-versa. As a consequence, the parameters are not sufficient for product validation as they only provide a relative measure for the algorithm performance. Therefore, we introduce two tests on this level of validation where the basic algorithm performance in terms of consistency and matching quality will be evaluated, namely backmatching (VT1), and velocity retrieval for stable ground (VT2). Both are outlined in the following two subsections.

7.2.1.1. Backmatching (VT1)

Errors in the displacement field derived from an image pair are produced when the point identified by the matching algorithm in the search window does not correspond to the one in the reference window. To measure the consistency of the matching process, the error in the correlation can be determined by so-called backmatching.

For this purpose matching is performed with image 1 as reference and image 2 as slave. Then, the process is reversed with image 2 as the master and image 1 as the slave. Any discrepancies between the results are evidence for erroneous matches. By comparing the resulting displacement fields, this test measures the consistency in the correlation with the following parameters:

- \( N \) Total Number of matches
- \( N_c \) Number of correct matches
- \( F = \frac{N_c}{N} \) fraction of correct backmatches

7.2.1.2. Retrieval over Stable Ground (VT2)

An approach which is widely used for quality control with image correlation, is matching of stable ground in the image set. This gives a good indication for the bias introduced by co-registration errors between the repeat images. After performing the matching for the entire region covered by the image pair, the results for the ice covered (moving) area will be separated from ice-free (stable) ground. The masking will be done using a polygon of the glacier outline. Buffers around the glacier polygon will be applied before extraction of stable ground for statistic calculation, or alternatively a final visual check for mis-classified stable terrain will be performed in order to avoid potential errors introduced by the area polygon.
The approach will be used to determine the systematic error of Glaciers_cci velocity products, where the following parameters will be recorded for the non-moving area:

- **N**  Total number of matches (only over stable ground)
- **RMSD**  Root mean square difference
- **sdev**  Standard deviation stable ground
- **K**  Kurtosis of distribution

### 7.2.2. Method/sensor cross-comparisons

Data availability for the individual validation sites drives the selection of tests performed on this level of validation activities. Here, we distinguish between the comparison of different methods (same data, but different algorithms) and the comparison of different sensors (same algorithm, but different data) as the former can be accomplished in the original sensor geometry and the latter requires geocoding and accurate transformation of the observed velocity component/spatial/temporal resolution.

#### 7.2.2.1. Comparison of different methods (same data, different algorithms) (VT3)

This test will give an indication about the overall performance and reliability of different offset tracking approaches. The method cross-comparisons will be performed at all test sites selected for validation (see Table 7.2). The comparisons will be performed internally (e.g. within the Glaciers_cci consortium) and play a very important role in preparing the external round robin experiments. A displacement field for one image pair of each sensor at each test site will be computed by Gamma, GUIO, and ENVEO, and fixed sizes for the search/reference window will be used. Therefore, the results can be compared directly where the following parameters will be utilised:

- **N**  Total number of matches
- **RMSD**  Root mean square difference
- **sdev**  Standard deviation
- **bias**  bias between velocity fields
- **R**  correlation coefficient

Note that for calculation of statistical parameters the number of matches has to be statistically significant.

#### 7.2.2.2. Comparison of different sensors (same algorithm, different data) (VT4)

Comparison of ice velocities estimated from image data of different sensors is a powerful way to learn about the relative accuracy of different datasets. However, a major drawback of these tests is related to the availability of more or less simultaneous acquisitions from different satellite systems, so that glacier-dynamical changes do not influence the validation. Furthermore, the datasets have to be adapted in order to represent the same component of the velocity vector, similar geometry as well as spatial resolution. This processing step contributes to the error budget and the precision of the transformations will be recorded.

The VT4 comparison will be performed at the test sites Breidamerkurjökull (V1), Kronebreen (V2), and Nordaustlandet (V3) where the cross-comparisons will additionally benefit from...
the availability of ground data. We will compare velocity fields from different SAR sensors at specified regions on the glaciers:

- V1 (Breidamerkurjökull):
  - ALOS-PALSAR versus TerraSAR-X
  - ALOS-PALSAR versus Landsat-7 ETM
- V2 (Kronebreen):
  - ALOS-PALSAR versus ENVISAT-ASAR
  - ALOS-PALSAR versus ASTER
- V3 (Nordaustlandet):
  - ALOS-PALSAR versus ENVISAT-ASAR
  - ALOS-PALSAR versus ERS-1 (12 day repeat)
- V4 (Baltoro):
  - ALOS-PALSAR versus Landsat TM

After transforming the different products to a common map projection, the discrepancies between the velocity fields will be described with the same parameters as in section 7.2.2.1.

7.2.3. Validation of EO products with in-situ velocity measurements (VT5)

The comparison to in-situ measured velocity data represents the highest level of validation achievable for satellite derived velocity data. Such comparisons can be performed at three test sites within the Glaciers_cci project:

- Test site V1 (Breidamerkurjökull, Iceland)
- Test site V2 (Kronebreen, Svalbard)
- Test site V3 (Nordaustlandet, Svalbard)

Two types of in situ velocity data will be used for the validation process: (1) Continuous GPS measurements at V1, V2 and V3, and (2) annually repeated measurements of survey stakes at V3 (see section 7.1.2 and 7.1.3 for details). However, as outlined in the beginning, the comparison of space-borne glacier velocity estimates with in-situ data is complicated by several issues. Though highly precise, the temporal and spatial representativeness of the GPS data compared to the area and time covered by the image data to be validated will vary and is not strictly known. In addition, the GPS data needs to be transformed to the velocity components observed by the corresponding feature tracking product. Therefore, the GPS inherent errors as well as the errors introduced by converting the GPS data in the sensor geometry and velocity component will propagate into the analysis and will be tracked and quantified. Note that for calculation of statistical parameters the number of in-situ data and corresponding EO observations has to be statistically significant. The comparison will be performed for the following parameters:

- \( N \)  Total number of in-situ data with corresponding EO observations
- \( \text{RMSD} \)  Root mean square difference
- \( \text{stdev} \)  Standard deviation
- \( \text{bias} \)  bias between in-situ and EO measurements
- \( R \)  correlation coefficient
8. **Review process and decision sequence**

8.1. **Glacier area**

As current algorithms used for glacier mapping are rather uncritical for product generation (the glacier outlines have to be visually controlled and corrected anyway), the objective validation criteria outlined above will likely be sufficient for a decision. This is also due to the long-term experience and the numerous comparison and validation experiments that have already been performed in the past decade. In this regard we are on solid ground. More critical is the manual editing part that has to deal with a wide range of possible interpretations. Even when high-resolution imagery (e.g. aerial photography) is available for comparison, it can be difficult to find a consensus on where a glacier boundary is located, in particular for debris-covered glaciers. In this case we foresee to give the results of the analysis along with the overlay of outlines for open discussion by the participants or even the wider GLIMS community. Such a discussion will, however, not be finished within a few months, but will likely be an ongoing activity. We think that the Glaciers_cci project can make an important step forward here, just by stimulating this discussion. We will thus invite the community to openly discuss the achieved results, give feedback and help to find a consensus. But we are also aware that for specific users of the products (e.g. global sea-level rise modelling) the discussion at the level of individual pixels will not matter. For them global completeness is much more important than the best possible accuracy of the product. The decision sequence will thus start with the objective criteria (area differences), continue with the more subjective details (outline overlay), and as a last step include community feedback on individual aspects of the comparison. This should allow us to define a best algorithm, provide a measure for product accuracy, and find consensus for specific glaciers with unclear outlines. The PVSAR will document the results and discussion of this analysis.

8.2. **Elevation change**

8.2.1. **Altimetry**

The validation of elevations by means of a comparison between airborne vs. satellite elevation change products and intercomparison between products obtained by different satellite sensors, will mainly be done internally, and documented. Inter-comparisons between satellite sensors and algorithms will be done both internally and externally from a group of participating specialists. All results from the different algorithms and sensors will be collected and drafted in a document that will be sent to all participants. This document will outline similarities and differences between the sensors/algorithms as well as error estimates derived from this. The Glaciers_cci team will then lead a discussion about the outcome of the inter-comparisons and promote discussions into why differences (if there are any) are occurring between the sensors/ algorithms. The decision sequence begins with the internal validation which then leads to the round robin and inter-method comparison. Once this is completed, the comparison between algorithms and sensors can be compiled as a draft and sent to all participants. This will then generate the
required discussion of the results and clarify the important learning aspects of this comparison.

8.2.2. DEM differencing

The validation of the global data products (i.e. SRTM, ASTER etc.) will be first completed internally using the higher resolution data. From this, co-registration and determination of the best possible mean bias correction is determined and all errors for the DEMs and the elevation changes are derived. All components of this internal validation will be drafted into an initial document. The decision sequence continues to the round robin where the specialists are approached and asked to provide their best estimates of co-registration and elevation change error estimates (Test site EC4, New Zealand). The data will be provided on a server, either as the original DEMs, orthophotos, or even raw images on request of participants. After analysis by the participants, results will be compiled by the Glaciers_cci team and then released to the specialists for analysis and discussion. Similarities and differences will be outlined between the different co-registration algorithms, possibly the different DEMs generated from the same imagery and the different error estimates. The focus will be on quality and efficiency of DEM co-registering and determination of a best practice for error quantification.

8.3. Ice velocity

The validation of ice velocity will be done internally and documented. All results from the different algorithms and sensors will be collected and drafted in a document that will be sent to all who participated. This document will outline similarities and differences between the sensors/algorithms as well as error estimates derived from this. The Glaciers_cci team will then lead a discussion about the outcome of the inter-comparisons and further promote discussions into why differences are occurring between the sensors / algorithms.

The decision sequence begins with the internal validation, which then leads to the round robin and inter-method comparison. Once this is completed, the comparison between methods and sensors can be compiled as a draft and sent to all participants. This will then generate the required discussion of the results and clarify the important learning aspects of this comparison. Focus will be on the cross-correlation algorithms rather than on pre-processing (e.g. global co-registration) or post-processing (filtering of outliers, etc.) and determination of a best practice for error quantification.
9. Resources summary and reference data sets

9.1. Glacier area

The Glaciers_cci EO team has carefully selected the datasets provided to the community for the RR and validation. The only requirement for participation consists in the availability of a Geographic Information System (GIS) or equivalent software that allows visualisation of Geotif files and digitised vector lines (shape files). Apart from a wide range of commercial products, also public domain software is available for this (e.g. GLIMSView, GRASS, QGIS). In this regard we hope that a large number of scientists will participate and send us their results.

The reference data sets will consist of orthorectified high-resolution satellite imagery (from Ikonos and Quickbird) as well as aerial photography in Geotif format. We will neither use field observations, nor vector outlines digitised by others as both do not fulfil the criteria for a scientifically sound validation (scale and interpolation problem). The data selected for the RR will be made freely available via the Glaciers_cci project website.

9.2. Elevation change

The software requirements for deriving elevation changes are typically a GIS (ArcGIS, GRASS etc.) or Remote Sensing Software (COSI-CORR, PCI Geomatica, ENVI etc.) and/or a mathematical scripting software (i.e. Matlab, Python, IDL, Scilab etc.) that allows the implementation of the algorithms and manipulation of matrices; i.e. (1) intersections/cross-overs for altimetry or (2) co-registration of DEMs and subsequent analysis over stable terrain. The software used by the participants are personally specific and restricted to what the individual researchers are familiar with and have access to. The round robin is directed to those who have particular experience working with altimetry/DEM differencing.

The reference datasets are those that are of finer (or at least known) accuracy and equal or better resolution than the satellite datasets of which we want to validate. For altimetry, this focuses on the NASA/CryoVEX airborne campaigns. For DEM differencing, a LIDAR DEM that will be considered the reference dataset in the validation of Findelen glacier and a national DEM is used for validation in New Zealand.

9.3. Ice velocity

The software requirement for generating velocity products from SAR data applying interferometry and offset tracking requires professional software (in-house, commercial, or open-source). Depending on the method and maturity of implementation public available software packages are DORIS, NEST, ROI_PAC; commercial software packages are e.g. GAMMA Software and SARScape. In order to participate in the round robin, suitable computer hardware for processing SAR data sets is also needed. The SAR data will be provided as Single-Look Complex (SLC) images.
The software used by the participants are personally specific and restricted to what the individual researchers are familiar with and have access to. For optical images, example software include COSI-CORR, IMCORR, or personally derived routines/functions in Matlab, IDL etc. Reference data sets include in-situ GPS measurements of ice motion. Other validation methods are based on intercomparison of ice motion fields from various sensors and methods. The availability of reference datasets is described in section 7.1 and in the DBT2 Technical Note (Glaciers_cci, 2011).

10. Validation documents and endorsements

Table 10.1 provides an overview on deliverables with information on product validation and the results of the algorithm selection. Apart from those that are already part of the project deliverables, we also seek for documenting the results in additional publications, such as a peer reviewed paper in a scientific journal (validation and RR results of all products) and ‘illustrated guidelines for the analyst’ (glacier area, overlay of outlines). Whereas the former will be prepared by the consortium, the latter will be prepared together with the interested community. We will seek for an open review process of all results achieved by informing the respective group of scientists (CRG, GLIMS, RR participants) when these documents are accessible.

As outlined above, apart from members of the CRG we will ask the wider GLIMS community and glaciologists via the GLIMS mailing list and Cryolist to participate in the RR and validation activities. If the results of the validation and RR activities of the individual Glaciers_cci products can be presented as papers and for example consensus will be reached for difficult glacier mapping conditions, the largest possible endorsement is achieved.

<table>
<thead>
<tr>
<th>Deliv. No.</th>
<th>Name</th>
<th>Date</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>D2.1</td>
<td>PVP</td>
<td>Dec. 2011</td>
<td>Outlines validation strategies and RR experiments</td>
</tr>
</tbody>
</table>
| D2.5      | PVSAR| March 2012 (GA)  
            | May 2012 (EC, IV)  | will summarise the measures and results of RR and validation activities and algorithm selection |
| D2.10     | ECR  | 2012       | describes uncertainty and error characterisation                        |
| D4.1      | PVR  | 2013       | product validation of a real dataset                                     |
| D4.2      | CAR  | 2013       | gives feedback on products from hydrological and climate modelling communities |
|           |      | end of 2012 | should describe the validation results                                   |
|           |      | 2013       | should help the analyst in delineating glaciers manually                 |

Table 10.1: Documents related to product validation for the Glaciers_cci products.
11. Round Robin exercises

11.1. Overall goals of round robin exercises

The round robin will serve as benchmark tests for algorithms available and / or proposed / developed during the project.

The final goal is to select the best performing algorithm in terms of various aspects:

- spatial coverage (i.e. number of successful matches in case of the velocity product)
- average accuracy
- robustness
- processing time and manual interaction
- software and hardware requirements, etc.

We intend to perform a consortium-internal round first, in order to identify potential problems and weaknesses of the strategy and to ensure a smooth cycle with the external participants later on.

A good design of these tests is crucial to maximise the number of participants as well as the quality of the feedback and therefore the positive impact of the Glaciers_cci results and beyond. Furthermore, we strongly believe that this two-step approach will help to speed up the later phases (analysis and report on the results) of the RR exercise, reducing the amount of time needed for the entire experiment. Due to these reasons, we will only present a preliminary framework of the RR tests in this document as the structure, focus, protocol etc., will change and evolve during the internal round.

Analysis of the results and their comparison to validation data will be performed internally. The RR participants will be provided with the EO and auxiliary datasets necessary for product generation, the validation data will in general not be distributed (apart from the area product which takes a dual approach). This will assure that the analysis of the RR tests will be conducted under the same conditions and follow the same rules. Furthermore, this avoids copyright issues with data held by consortium members (e.g. high resolution national DEMs).

<table>
<thead>
<tr>
<th>No.</th>
<th>Recommendation</th>
<th>Area</th>
<th>Elev. Change</th>
<th>Ice Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR-1</td>
<td>The meaning of the ‘best’ algorithm and of how to select it (evaluation protocol) has to be defined before the start of the round robin exercise. The definition of ‘best’ and the scope of the round robin exercise have to be specified in the PVP.</td>
<td>detailed in section 5</td>
<td>Detailed section 6</td>
<td>spatial coverage (count of successful matches), average accuracy, robustness, processing time (w.r.t. software, hardware, manual interaction, etc.)</td>
</tr>
<tr>
<td>No.</td>
<td>Recommendation</td>
<td>Area</td>
<td>Elev. Change</td>
<td>Ice Velocity</td>
</tr>
<tr>
<td>-----</td>
<td>----------------</td>
<td>------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>RR-2</strong></td>
<td>The round robin should be made at the beginning of the project based on objective criteria. There should be one or more iterations to show algorithm improvement throughout the project. The most objective algorithm selection would be based on blind testing to avoid any bias.</td>
<td>As the algorithms are already well established, the RR will only be performed once.</td>
<td>The RR will only be performed once.</td>
<td>as the algorithms are already well established, the RR will only be performed once</td>
</tr>
<tr>
<td><strong>RR-3</strong></td>
<td>Every CCI project has to perform a round robin exercise. In the exceptional case that a final algorithm has been pre-selected, component modules need to be tested also for this pre-selected algorithm. Furthermore, the pre-selection criteria should be in line with the CCI objectives.</td>
<td>The RR will cover a wide range of assessments, incl. algorithm selection, 1 to many &amp; many to 1 digitising</td>
<td>One algorithm is generally pre-selected, but tested against other non-analytical solutions. Component modules may be tested for efficiency.</td>
<td>compliant</td>
</tr>
<tr>
<td><strong>RR-4</strong></td>
<td>The same auxiliary and Level 1 data should be used in the processing, as well as the same reference data.</td>
<td>The L1T product from USGS is used for both</td>
<td>We may compare different DEMs generated from the same imagery.</td>
<td>compliant</td>
</tr>
<tr>
<td><strong>RR-5</strong></td>
<td>The round robin results need to be open and the algorithm must be well-documented and public, but the actual code does not need to be public.</td>
<td>ATBDs, protocols and results will all be public (website)</td>
<td>ATBDs, protocols and results will all be public (website)</td>
<td>The algorithms are well-documented and public.</td>
</tr>
<tr>
<td><strong>RR-6</strong></td>
<td>The algorithm selection should be made by an independent team that is not directly involved in the algorithm development, although of course the members of that team should be experts. The selection shall be made based on a round robin evaluation protocol developed beforehand and providing objective criteria.</td>
<td>The EO team for validation has no relation to the algorithms applied by the participants. Evaluation criteria are documented in the PVP.</td>
<td>Evaluation criteria are documented in the PVP</td>
<td>Internal validation is performed by the consortium using independent data sets. External participants are only involved in the RR.</td>
</tr>
<tr>
<td><strong>RR-7</strong></td>
<td>The algorithm selection should be made by an independent team that is not directly involved in the algorithm development, although of course the members of that team should be experts. The selection shall be made based on a round robin evaluation protocol developed beforehand and providing objective criteria.</td>
<td>The proposed RR and validation activities are all based on well established standards</td>
<td>The proposed RR and validation activities are all based on well established standards</td>
<td>Same as RR-6</td>
</tr>
</tbody>
</table>

*Table 11.1: Response to the recommendations for the round robin from the project guidelines.*
11.2. Tasks and responsibilities for the round robin activities

The selection of suitable RR sites, EO and validation data is done by the responsible partners in the consortium. The following partners are responsible for:

- glacier area: GIUZ
- elevation change – DEM differencing: GUIO
- elevation change – altimetry: SEEL
- velocity: GUIO (supported by ENVEO and Gamma)

The database for the RR experiments will be set up and maintained by ENVEO. ENVEO will also make sure that the RR participants can download the data sets for the corresponding experiment (this will require registration) and also upload the products and the completed RR protocols. The submission of useful results by any participant will be acknowledged with co-authorship in a related publication presenting the results.

Project partners will also contribute to the RR experiments (the internal RR).

- GIUZ: contributions to glacier area RR
- ENVEO: contributions to glacier area and velocity RR
- SEEL: contributions to elevation change (altimetry) RR
- GUIO: contributions to elevation change (DEM Differencing), velocity (optical), and glacier area RR
- GAMMA: contributions to velocity (SAR) RR

The validation, intercomparison of the products from the RR experiments, as well as the evaluation of algorithms will be done by the responsible partners (Glacier Area: GIUZ, Elevation Change (DEM differencing): GUIO; Elevation Change (altimetry): SEEL; Velocity (optical): GUIO, (SAR): Gamma. The results of the validation and RR experiments will be presented in the PVSAR which is lead by GIUZ, SEEL, and GUIO for the respective products.

11.3. Schedule for the round robin experiments

The schedule of round robin exercises is described in the Master schedule (section 3).

11.4. Glacier area

11.4.1. Goals of the round robin experiments

The round robin for glacier area has three main goals:

A) Comparison of different algorithms for glacier mapping (and selection of a most suitable one),
B) assessment of the interpretation of the glacier outlines by different analysts (to detect methodological variability), and
C) multiple delineations of the same glacier entity (to assess the accuracy of manual digitisations).
All three components should cover glaciers of different size, type and location to get the widest possible coverage of possible cases. On the other hand, the round robin should not take more than a few days to allow a wide community to participate. Hence, the sample of glaciers is restricted and strongly selected, but it will give nevertheless statistically significant results in regard to sample size and representativeness. The RR for the glacier area product will be open to all interested scientists to get the widest possible feedback. We hope that about 5 to 10 persons will participate in all experiments and maybe 20 for at least some of them. An overview to the recommendations for the RR from the project guidelines is provided in Table 11.1.

11.4.2. Definition of data sets

As for the validation data sets, the satellite scenes for the round robin will be provided in Geotif format. They consist of selected subsets of two different Landsat satellite scenes (from the Alps and the Karakoram) with already prepared RGB composites with bands 321, 432, and 543 as RGB. Additionally, the individual bands 1 to 5 are provided for the Karakoram scene to allow application of an automated mapping algorithm. Of course, participants can also decide to use manual delineation or to process the full scene, which is freely available from glovis.usgs.gov. As manual correction is an important part of achieving a high product accuracy, the participants will be asked to correct the glaciers in a small part of the scene and leave the other regions unchanged. The Landsat scene from the Alps will only be used for multiple digitizations of ten pre-selected glaciers. They will be used for an independent validation of the most suitable algorithm. For the Alaska test site eight glaciers on a high resolution Quickbird scene need to be digitised once by the participants. The EO team will perform a sensitivity analysis in this region. All glacier outlines have to be provided in shapefile format. The applied algorithm needs to be described in an additional document (.doc format). The schematic overview provided in Figure 5.9 and the description in section 5.3 gives further details. All regions for the RR are illustrated in section 5.2.

11.4.3. Methods for inter-comparison and analysis

The datasets generated in the round robin will allow us to assess:
- the performance of different algorithms for two test regions with challenging conditions,
- the interpretation of debris-covered glaciers based on manual corrections,
- the analysts internal accuracy from multiple digitisations of the same glacier, and
- the results of different analysts for the same glacier.

This should provide a broad overview on algorithm performance and analyst’s interpretation. Together with the validation datasets, results can also be assessed in a quantitative way, so that an objective evaluation is possible. The RR will also provide some of the datasets that will be used for product validation (e.g. by multiple digitizing of the same glaciers).
11.4.4. Selection criteria

The selection of the ‘best’ algorithm will be based on objective criteria (difference to a validation data set) as well as more subjective criteria (overlay of the digitised outlines). The latter is required, as it is likely that very similar total glacier areas result from outlines that are very different, as omission and commission errors will average out (random errors). It will thus be assessed if these deviations are randomly distributed or point to a systematic bias in interpretation. For the pure (uncorrected) algorithm results, the algorithm closest to a reference area will win the first round. But also here a second criterion will be applied, that will investigate where the differences are and if they need correction. For example, one algorithm might map a few pixels more in debris-covered regions, but less in regions with ice in shadow. As shadow regions should be properly mapped and debris-covered regions will be corrected anyway, the algorithm with the better performance in shadow will be selected. The criterion applied is based on the smallest workload required for post-processing. Further criteria that might be considered will also depend on the range of used algorithms:

- complexity of the calculation (required input data, pre-processing workload),
- robustness of the result in regard to slightly different thresholds (if applicable),
- processing speed and degree of automation, and
- special software requirements.

11.5. Elevation change - DEM differencing

11.5.1. Goals of the round robin

The round robin for DEM differencing has four main questions:
- Do the different DEM co-registration techniques arrive at the same solution?
- What is the efficiency of the different co-registration techniques and can the computational efficiency be improved?
- How well can the mean elevation adjustment between DEMs be determined? And does the error estimate encompass this mean bias.
- What are the consequences of using different DEMs from the same imagery?

This type of processing is time consuming and therefore is limited to one representative study site. The region of the Southern Alps, New Zealand is used as it represents well the alpine nature of many glacierized regions of the world, because it has variable glacier changes from which we compare, and because of the available regional subsets of global data sets (e.g. SRTM, GDEM) of the region.

11.5.2. Definition of data sets

DEM co-registration is tested using the globally available SRTM and ASTER DEMs. For the ASTER DEM, we will offer both raw images, as well as DEMs generated from the SilcAst software. For the SRTM, we will use the raw version without void fills. In addition, ICESat will serve as a reference dataset that may also be provided.
11.5.3. Methods for inter-comparison and analysis

For each participant, we will request a short summary of the methods applied as well as the results of the co-registration (i.e. the 3D co-registration vector) for each of the products. We will also ask the participants to check their co-registered products to ICESat to provide a mean and standard deviation of the differences over stable terrain. This will form the basis for inter-algorithm comparisons. Finally, the triangulation of the co-registration vectors between three datasets (i.e. two ASTER DEMs + one SRTM or one ASTER + one SRTM + one ICEsat) will provide an estimate of the mean un-removed bias between the datasets. The participants will be asked to provide an error estimate for a single elevation change point on the glacier and for the mean elevation change of the glaciers.

11.5.4. Selection criteria

The criteria used to evaluate the results of the round robin are:
- complexity of the calculation (required input data, workload),
- variation in the results of the different algorithms and justifiable reasons,
- robustness of the result (how does the result change when limited to certain sample sizes),
- processing speed and efficiency, and
- degree of automation.

The various error estimates derived by the partners in the round robin will also be analyzed in attempts to determine a means of best practice when quoting errors. In particular, it will be determined whether the given error estimate actually includes the accuracy of how well we can determine the vertical bias between DEMs.

11.6. Elevation change – Altimetry

11.6.1. Goals of the round robin

The round robin for elevation change products obtained from altimetry data has as a principal goal the comparison of the different algorithms available in literature for determining the elevation change and consequently the selection of a most suitable one.

The round robin should cover glaciers of different size, type and location to get the widest coverage of possible cases. On the other hand, when radar acquisitions are considered for the estimation of the elevation change, it is necessary to take into account the possibility of lags in the radar pulse tracking system of the altimeter. These are mainly due to larger topographic variations. As a consequence, the glaciers to consider for the RR activity should not present a high slope gradient. With respect to the size of the glaciers, this is limited by the large footprint of the radar on ground, which simply sets a lower limit of target size, and by the acquisition strategies of the several missions, which cause the data to be sparsely distributed in space and time. As a consequence, large glaciers and ice caps with an area in the order of hundreds of square kilometres are best suited for studying elevation changes. As for glacier area, this means that even a restricted, selective sample can be used to derive statistically significant results. The RR for the elevation change product will be made available to all interested parties to get the widest possible feedback.
11.6.2. Definition of data sets
The test sites for the round robin are the same as used for the validation activity and as for the validation data sets. The satellite data are freely available.

11.6.3. Methods for inter-comparison and analysis
The methods for the intercomparison and analysis are the same as those described in the section 6.2 for the validation process.

11.6.4. Selection criteria
The selection criteria will be based on:

a) the goodness of fit:
   - the smallest RMSE deviation of satellite derived elevation change trends with respect to the airborne ones
   - the highest $R^2$
   - values of the slope and intercept of the regression line between the two estimated datasets, which give an assessment of an over- or under estimation.

These criterions will fulfil the principles of objectivity;

b) the computational cost, i.e. time taken to process data;

c) the geographical extent, i.e. volume and distribution of data resulting from each algorithm/dataset.

11.7. Ice Velocity

11.7.1. Goals of the round robin experiments
The goal for the RR experiments for ice velocity is to find the best performing algorithm with respect to:

- spatial coverage (i.e. number of successful matches),
- average accuracy,
- robustness, and
- processing time (software, hardware, manual interaction).

11.7.2. Definition of data sets
It is planned to provide pre-processed datasets to the participants to keep their workload to a minimum. That means that the RR data packages for velocity will include co-registered image pairs/series subsetted to the extent of the test site. However, for SAR data the SLC image (not co-registered) will be provided as input data, as well as whole frames for the internal RR, possibly only sub-regions of the data afterwards for the external RR. The aim of this is to reduce the workload for the RR participants and to provide the same starting data sets (focus on algorithm inter-comparison).

The location of the glaciers in each EO data set will be specified and provided to RR participants. Further, we will add a quick-look of the results derived within the Glaciers_cci consortium. Table 11.2 lists the proposed RR experiments that are planned for ice velocity product (provided that sufficient external participants register to portion the workload).
<table>
<thead>
<tr>
<th>ID</th>
<th>Aim</th>
<th>Site Name</th>
<th>Method</th>
<th>Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR-V1</td>
<td>Optical Offset Tracking (Intercomparison of diff methods)</td>
<td>Baltoro, alpine and debris covered</td>
<td>Section 7.2.2.1;</td>
<td>DBT2 table 10, dataset V4D3</td>
</tr>
<tr>
<td>RR-V2</td>
<td>SAR Offset Tracking (Intercomparison of diff methods)</td>
<td>Baltoro, alpine and debris covered</td>
<td>Section 7.2.2.1</td>
<td>DBT2 table 10, dataset V4D1</td>
</tr>
<tr>
<td></td>
<td>Comparison SAR - Optical</td>
<td>Baltoro, alpine and debris</td>
<td>Section 7.2.2.2</td>
<td>Products from RR-V1 and RR-V2</td>
</tr>
<tr>
<td>RR-V4</td>
<td>SAR Offset Tracking (Intercomparison of different SAR sensors)</td>
<td>Nordaustlandet, polar ice cap</td>
<td>Section 7.2.2.2</td>
<td>DBT2 table 10, datasets V3D1 and V3D3</td>
</tr>
<tr>
<td></td>
<td>Absolute validation: SAR versus GPS</td>
<td>Nordaustlandet, polar ice cap</td>
<td>Section 7.2.3</td>
<td>Products from RR-V4; GPS data in DBT2 table 11</td>
</tr>
<tr>
<td>RR-V5</td>
<td>SAR Offset Tracking (Intercomparison of different SAR sensors)</td>
<td>Kronebreen, polar glacier</td>
<td>Section 7.2.2.2</td>
<td>DBT2 table 10, datasets V2D1, V2D3 and V2D4</td>
</tr>
<tr>
<td></td>
<td>Absolute validation: SAR versus GPS</td>
<td>Kronebreen, polar glacier</td>
<td>Section 7.2.3</td>
<td>Products from RR-V5 GPS data in DBT2 table 11</td>
</tr>
<tr>
<td>RR-V6</td>
<td>Optical Offset Tracking (Intercomparison diff. Algorithms)</td>
<td>Kronebreen, polar glacier</td>
<td>Section 7.2.2.2</td>
<td>DBT2 table 10, datasets V2D2 and V2D3</td>
</tr>
<tr>
<td></td>
<td>Absolute validation: Optical versus GPS</td>
<td>Kronebreen, polar glacier</td>
<td>Section 7.2.3</td>
<td>Products from RR-V6 GPS data in DBT2 table 11</td>
</tr>
<tr>
<td></td>
<td>Comparison SAR - Optical</td>
<td>Kronebreen, polar glacier</td>
<td>Section 7.2.2.2</td>
<td>Products from RR-V5 and RR-V6</td>
</tr>
<tr>
<td>RR-V7</td>
<td>SAR Offset Tracking (Comparison of different SAR sensors)</td>
<td>Breidamerkurjö kull, temperate glacier</td>
<td>Section 7.2.2.2</td>
<td>DBT2 table 10, datasets V1D1, V1D2</td>
</tr>
<tr>
<td></td>
<td>Absolute validation: SAR versus GPS</td>
<td>Breidamerkurjö kull, temperate glacier</td>
<td>Section 7.2.3</td>
<td>Products from RR-V7 GPS data in DBT2 table 11</td>
</tr>
<tr>
<td>RR-V8</td>
<td>Optical Offset Tracking (Intercomparison of diff methods)</td>
<td>Breidamerkurjö kull, temperate glacier</td>
<td>Section 7.2.3</td>
<td>DBT2 table 10, datasets V2D3</td>
</tr>
<tr>
<td></td>
<td>Absolute validation: Optical versus GPS</td>
<td>Breidamerkurjö kull, temperate glacier</td>
<td>Section 7.2.3</td>
<td>Products from RR-V8 GPS data in DBT2 table 11</td>
</tr>
<tr>
<td></td>
<td>Comparison SAR - Optical</td>
<td>Breidamerkurjö kull, temperate glacier</td>
<td>Section 7.2.2.2</td>
<td>Products from RR-V7 and RR-V8</td>
</tr>
</tbody>
</table>

Table 11.2: Round robin experiments (internal and external) for ice velocity and options for validation and product intercomparison. The DBT2 is published in Glaciers_cci (2011).
11.7.3. Methods for inter-comparison and analysis

We will use the same methods for the inter-comparison of the RR results that we have specified in sections 7.2.1, 7.2.2, and 7.2.3.

11.7.4. Selection criteria

Based on the number of participants in the RR and the tests that can be accomplished, we will analyse all returned results as indicated in Table 11.3. These cross-comparisons will be utilized to rank the performance of the different algorithms and approaches with respect to the criteria listed in section 11.7.1.

<table>
<thead>
<tr>
<th>Participant</th>
<th>RR-V1</th>
<th>RR-V2</th>
<th>RR-V3</th>
<th>. .</th>
<th>RR-V8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>Parameters as in section 7.2.2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 2</td>
<td>Parameters as in section 7.2.2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant 3</td>
<td></td>
<td>Parameters as in section 7.2.2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>. . .</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant X</td>
<td>Parameters as in section 7.2.2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 11.3: Schematic outline for the inter-comparisons of the RR results

11.8. Round Robin Protocol

A template of the round robin protocol for each of the products will be designed and distributed to the participants of the round robin experiments. The round robin protocol will include question on the:

- Input data used for generating products (EO data, auxiliary data, etc.)
- Algorithm used to generate the product (including links to detailed description of the algorithm and/or processing line; e.g. publications, manuals, etc.).
- System information
  - Description of process flow (flow chart or text that identifies the modules, processing sequence/logic, operator interaction/decisions needed)
  - Used Software Tools
  - Used Hardware (RAM, Disk, Processor)
  - Minimal software and hardware requirements (if known)
  - Processing time
- Quicklook of the generated product

The round robin protocol will be designed and adapted to the needs of each of the products. It will be tested internally by the project partners before it is sent out to the external participants of the RR experiments. This will ensure that all the required parameters and questions can be asked and collected.
12. References


13. Acronyms

ALOS  Advanced Land Observing Satellite
ALS  Laser Scanning Altimeter
ASAR  Advanced SAR
ASIRAS  Airborne SAR Interferometric Radar System
ASTER GDEM  ASTER Global Digital Elevation Model
ASTER  Advanced Spaceborne Thermal Emission and Reflection radiometer
ATBDv0  Algorithm Theoretical Basis Document version 0
ATM  Airborne Topographic Mapper
AWI  Alfred Wegener Institute
AWS  Automatic Weather Station

CC  Correlation Coefficient
CCI  Climate Change Initiative
CEOS-WGCV  Committee for Earth Observation Satellites – Working Group for Calibration / Validation
CMUG  Climate Modelling User Group
CRG  Climate Research Group
CryoVEX  Cryosat Validation Experiment

DARD  Data Access Requirement Document
DBT2  Database for Task 2
DEM  Digital Elevation Model
DGPS  Differential Global Positioning System
DInSAR  Differential Interferometric Synthetic Aperture Radar
DORIS  Delft Object-oriented Radar Interferometric Software

ECV  Essential Climate Variable
ENVISAT  Environmental Satellite
EO  Earth Observation
ERS  European Remote-sensing Satellite
ETM+  Enhanced Thematic Mapper plus

FCC  False Colour Composite

GIS  Geographic Information System
GLACE  GLims Analysis Comparison Experiment
GLIMS  Global Land Ice Measurements from Space
GPS  Global Positioning System
GRASS  Geographic Resources Analysis Support System

ICESat GLAS  ICESat Geoscience Laser Altimeter System
ICESat  Ice, Cloud, and land Elevation Satellite
InSAR  Interferometric SAR
IPY: International Polar Year

LIDAR: Light Detection and Ranging
LOS: Line Of Sight
LVIS: Laser Vegetation Imaging Sensor

NDSI: Normalized Difference Snow Index
NEST: Next ESA SAR Toolbox

PALSAR: Phased Array type L-band Synthetic Aperture Radar
PVP: Product Validation Plan
PVSAR: Product Validation and Algorithm Selection Report
QGIS: Quantum GIS

RA2: Radar Altimeter 2
RADAR: Radio Detection and Ranging
RAM: Random-Access-Memory
RGB: Red Green Blue
RMSD: Root Mean Square Difference
RMSE: Root Mean Square Error
ROI_PAC: Repeat Orbit Interferometry Package
RR: Round Robin
RRDP: Round Robin Data Package

SAR: Synthetic Aperture Radar
SLC: Single Look Complex
SNR: Signal-to-Noise Ratio
SPOT: System Pour l’Observation de la Terre
SPOT5-HRS: SPOT5- High Resolution Sensor
SRTM: Shuttle Radar Topography Mission
SWIR: Short-wave Infrared

TM: Thematic Mapper

USGS: United States Geological Survey