



GLIMS Analysis Tutorial

Bruce Raup
Siri Jodha Singh Khalsa

Contents

1	Introduction	2
2	Tools	3
3	Acquiring ASTER Imagery	3
4	Input Image(s)	3
5	Definition of a Glacier	4
6	Defining glacier outlines and their attributes	7
7	Definition of “left” and “right” of segments	9
8	Assigning GLIMS Glacier IDs	9
	Background	9
	Definition	10
	Choosing the Location for a GLIMS Glacier ID	10
9	Population of Other Fields (including Null values)	12
	Null values	12
	As-of time (source time) of glacier outlines	12
10	Creating a “mugshot” of a glacier or set of glaciers	12
11	Measurement uncertainty	13
12	Using GLIMSView as a filter	13
13	Working with multiple images	13
14	How the Ingest Process Works	14

1 Introduction

This document contains basic guidelines for preparing a glacier outline data set for insertion into the GLIMS Glacier Database. While other documents describe the details of the database design and other aspects of GLIMS, this document is designed to assist the process of doing a “GLIMS Analysis Session” and producing files that are optimally configured for ingest into the GLIMS Glacier Database. We also describe preparing an existing data set, such as one derived from maps, for GLIMS.

A GLIMS Analysis Session is defined as the activity of producing a set of glacier outlines, together with GLIMS attribute data, usually from a single image. A GLIMS Analysis (without the “Session”) is one snapshot of one

glacier. Each record in the Glacier_Dynamic table corresponds to one analysis. An Analysis Session typically contributes several new records to the Glacier_Dynamic table.

2 Tools

A variety of tools are being used in the GLIMS community to generate glacier outlines for GLIMS. For example, an analysis may be done using GIS tools such as ERDAS, ArcGIS or ENVI. In this case the analyst must create all the necessary shapefiles manually, or using GLIMSView, to contain the outlines plus the required metadata. These files are: session.shp, glaciers.shp, segments.shp, images.shp (with accompanying index and database files).¹

The GLIMS group at the USGS in Flagstaff, Arizona, has produced the software “GLIMSView”, with which a user can produce digitized glacier outlines, add appropriate metadata, and export the whole package into the GLIMS Data Transfer Format².

GLIMSView has a plug-in structure for adding functionality, such as image classification algorithms. It can also serve as a filter: a user can import outlines generated in another package (via the shapefile format, provided they are in lon/lat coordinates), then use GLIMSView to add the metadata and export into the transfer format. See Section 12 for details.

3 Acquiring ASTER Imagery

ASTER imagery can be searched for and ordered from at <http://edcimswww.cr.usgs.gov/pub/imswelcome/>. For information on FTP access to a pool of ASTER and Landsat 7 data provided by EROS Data Center, see <http://www.glims.org/MapsAndDocs/ftp.html>. GLIMS Regional Centers have free access to ASTER data.

4 Input Image(s)

Normally, the analyst will work on only one ASTER (or other) image during an analysis session. In this case, the image ID for that image should be input into the appropriate metadata field in GLIMSView, or the “images” shapefile. In some cases, however, it will be more pragmatic to work from a mosaic of two or more images. GLIMSView can open and work with a properly formatted mosaic image (such as a GeoTIFF image), but allows only one image ID to be entered. To identify multiple images for a single analysis

¹ See http://www.glims.org/MapsAndDocs/datatransfer/data_transfer_specification.html

² See <http://www.glims.org/glimsview/>

session, one must manually edit the “images” shapefile. Future versions of GLIMSVIEW may eliminate this need.

5 Definition of a Glacier

Through 2005, the GLIMS policy was to allow the analyst to decide whether and how to break up an ice mass with a branching structure into individual glaciers, because different branches of a given ice mass may historically have been treated as separate glaciers and given different names. However, this approach leads to inconsistencies in how glaciers are treated, which greatly complicates large-scale analysis of the GLIMS Glacier Database. It was recognized that a standard definition was needed: 1) to guide the delineation of, and separation into individual glaciers, bodies of ice that all glaciologists would agree is glacier ice, and 2) to guide the GLIMS analyst in determining what is glacier ice and what is not. As a result, we have decided upon the following definition of a “glacier” for the purposes of GLIMS. Note that the definition is tailored to remote sensing, and thus does not involve motion of ice. However, it is compliant with the World Glacier Monitoring Service (WGMS) standards with respect to Tier 5 sites as defined in the observing strategy of the Global Terrestrial Network for Glaciers (GTN-G). This definition is also not intended to be used in any sort of legal context. Given the limitations of current remote sensing technology, we recognize that this definition may lead to the inclusion, in certain cases, of what would generally be considered “perennial snow masses”. Definitions of “glacier” for other purposes outside of GLIMS exist elsewhere.

A glacier or perennial snow mass, identified by a single GLIMS glacier ID, consists of a body of ice and snow that is observed at the end of the melt season, or, in the case of tropical glaciers, after transient snow melts. This includes, at a minimum, all tributaries and connected feeders that contribute ice to the main glacier, plus all debris-covered ice. Excluded is all exposed ground, including nunataks. An ice shelf (see item 9 below) shall be considered as a separate glacier.

The following consequences and observations must be kept in mind:

1. Bodies of ice above the bergschrund that are connected to the glacier shall be considered part of the glacier, because they contribute snow (through avalanches) and ice (through creep flow) to the glacier.
2. A tributary in a glacier system that has historically been treated (and

- named) as a separate glacier should, within the GLIMS framework, be included as part of the glacier into which it flows. The name field for the glacier should be populated with all relevant names of tributaries.
3. Any steep rock walls that avalanche snow onto a glacier but do not retain snow themselves are NOT included as part of the glacier.
 4. A stagnant ice mass still in contact with a glacier is part of the glacier, even if it supports an old-growth forest.
 5. If no flow takes place between separate parts of a continuous ice mass, they should, in general, be treated as separate units, separated at the topographic divide. However, for practical purposes, such an ice mass may be analyzed as a unit at the analyst's discretion, if delineation of the flow divides is impossible or impractical. If the same system is analyzed in the same way later, it will have the same glacier ID, and can therefore be compared. If the system is analyzed in more detail later by breaking it into its component glaciers, those pieces will get new IDs (ID of system will be "parent icemass" ID for each part), and future analyses of those pieces, if done in the same way, will be comparable.
 6. All debris-covered parts of the glacier must be included.
 7. It is possible that an ice body that is detached from another may still contribute mass to the latter through ice avalanches, or it may no longer do so. It is practically impossible to tell which is the case from a single satellite image. Therefore, within GLIMS, adjacent but detached ice areas should, in general, be considered as different "glaciers", regardless of whether they contribute mass to the main glacier through snow or ice avalanches. However, at the analyst's discretion, detached ice masses may be included as parts of one glacier. This is similar to the situation described in 5 above. If the pieces are analyzed separately later, each piece should be given a new GLIMS ID, the old one being used as the "parent icemass" ID for all the pieces.
 8. What about the lower parts of lateral snowfields, whose extent varies from year to year? Map only at the end of the ablation period, preferably in a year without snow outside of the glaciers, to exclude seasonal snow. Then map everything that is connected to the glacier. If snowfields are identifiable, they should be disconnected from the main glacier. For hydrological purposes they can be included in the GLIMS Glacier Database under a separate GLIMS glacier ID, but they

- must be marked as a snowfield. Lateral glacier outlines that might be hidden by seasonal snow or by avalanches should be labeled as preliminary, or even the entire glacier can be excluded. Ice avalanche cones below a glacier terminus (dry-calving) are not a part of the glacier.
9. An ice shelf is typically ice downstream of the grounding zone of two or more glaciers that is floating on ocean water (e.g. the Getz Ice Shelf in Antarctica). Ice masses that form via direct accumulation at the surface, without significant ice input from the land surface (e.g. the Wilkins Ice Shelf), are also considered “ice shelves”. Ice tongues – floating ice extending from a single glacier (e.g. the Drygalski Ice Tongue) – are not considered ice shelves, and should be considered part of the glacier to which they are attached.
 10. Rock glaciers and heavily debris-covered glaciers tend to look similar, but their geneses are different. GLIMS does not currently treat rock glaciers.

Some points from the above are illustrated in Figure 1, which illustrates some decisions made by Frank Paul in his work with the Swiss Regional Center.

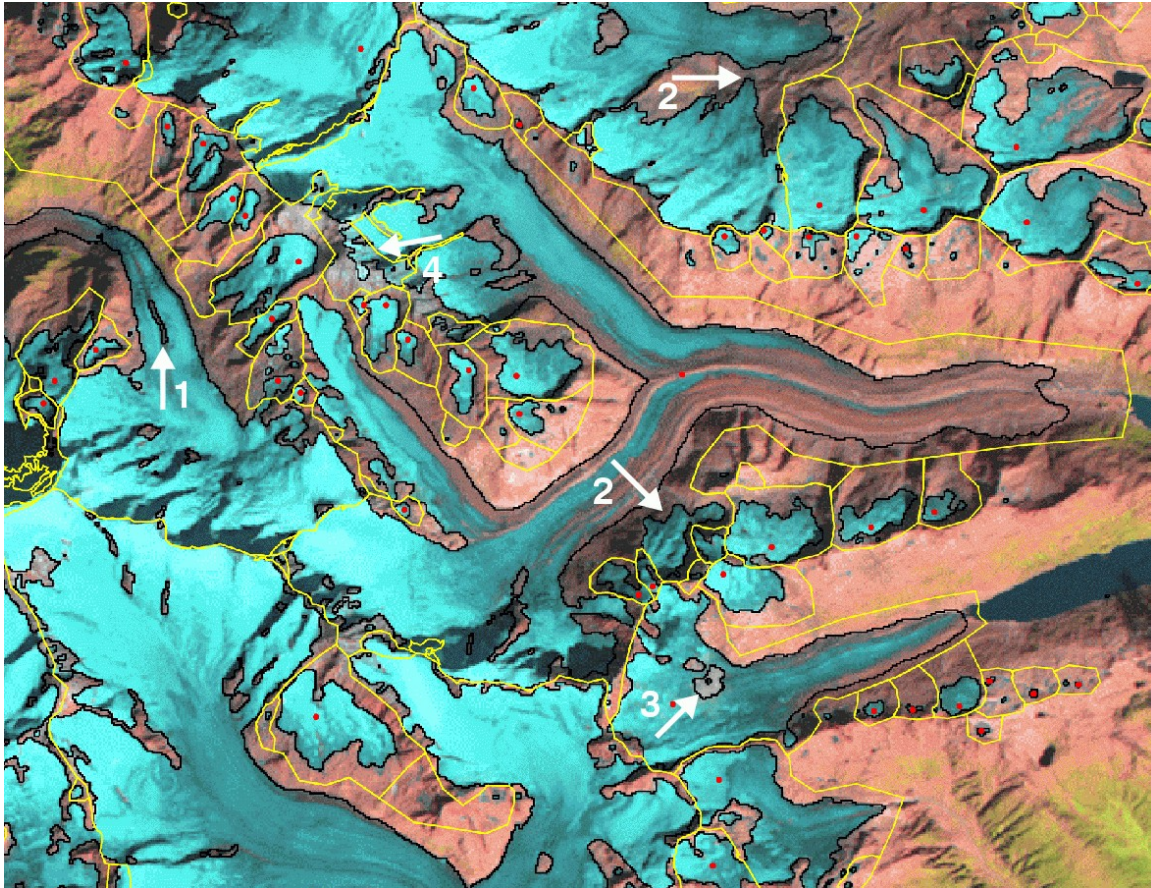


Figure 1: Landsat Thematic Mapper band 5,4,3 (red, green, blue) composite image of Unteraarglacier and surroundings, acquired 1998-08-31, with edited glacier outlines (black) and glacier basins (yellow) overlaid. The latter define ice divides and include all glacier parts that are related to a former glacier. Red dots mark basins that have been included in the new Swiss glacier inventory. Arrows denote: 1: Medial moraine outcrop (to be removed), 2: now disconnected glaciers, but included in the same basin for consistency, 3: a small cloud that hides a part of the glacier area (to be removed), 4: the ice divide is used here to correct for misclassified seasonal snow. Image and caption courtesy of Frank Paul.

6 Defining glacier outlines and their attributes

Two important considerations for producing a set of glacier outlines and metadata for GLIMS are 1) the data model, and 2) the file formats. The discussion below touches on both these aspects. For details on the GLIMS Data Transfer Format, see the specification at http://glims.org/MapsAndDocs/datatransfer/data_transfer_specification.html. That file describes the shapefiles and their attributes that form the basis for transferring data to the GLIMS Glacier Database.

To create an outline that conforms to the above definition, one should create one polygon (or series of segments) that circumscribes the entire glacier. Internal rock outcrops are excluded by producing outlines around them and labeling those outlines as internal rock. This can be done simply in GLIMSVIEW, or can be done with other tools. In the resulting “segments” shapefile, the “category” attribute should be “intrnl_rock” for internal rock segments. Internal rock polygons should be separate polygons (or collections of segments), not sub-parts of multi-part polygons. Glacier outlines should be of *glacier* boundaries, not *basin* boundaries.

Outlines can be single polygons, or, if the analyst wishes to supply more detailed attributes for different parts of a glacier outline, it may be composed of a number of segments. Each segment can have different attributes. In either case, each segment or polygon must have a “category” attribute that identify its type.

Possible values for the “category” attribute include “glac_bound”, “basin_bound”, “intrnl_rock”, “debris_cov”, “pro_lake”, “supra_lake”, “centerline”, and “snow_line”. Possible values for “segment_type” are “m” for measured and “a” for arbitrary (for when a portion of a boundary must be inferred, such as at the edge of an image, or the upper part of an outlet glacier or ice stream from an ice sheet). Additionally, segments may be labeled (attribute “label”) with values such as cld (cloud), shd (shadow), trm (terminus), div (divide), gdl (grounding line), lat (lateral margin), hed (boundary at head of glacier), and eod (edge of [image] data). Glacier polygons must be closed. That is, the set of segments that share a glacier ID and polygon type (e.g. glac_bound, intrnl_rock, debris_cov) must be combinable into one or more closed polygons. The same GLIMS glacier ID (see Section 8) is assigned to all polygons associated with that glacier, i.e. glacier outlines, centerlines,

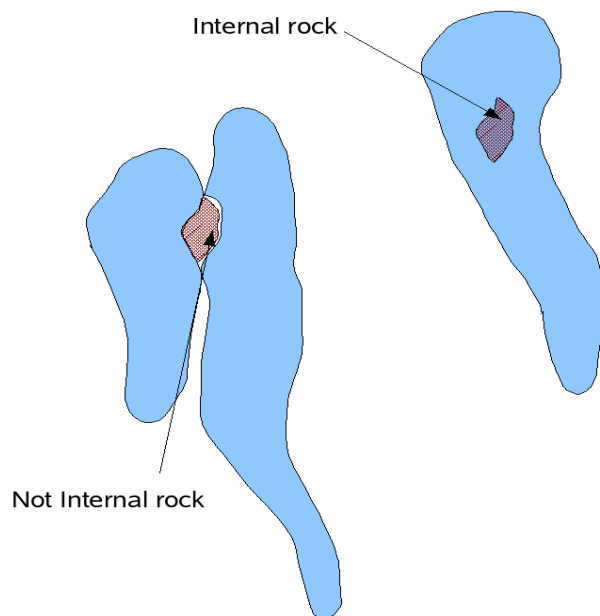


Figure 2: “Internal rock” polygons should only be included if the outcrop is completely internal to one glacier.

internal rock outlines all get the same glacier ID. If the analysis is done with a package other than GLIMSVIEW, be sure not to combine different line types, e.g. glacier outlines and internal rock outlines into a single multi-part polygon. GLIMSVIEW, and the database ingest software, cannot easily handle multi-part polygons.

Internal rock boundaries are only of interest to GLIMS if they are truly internal to a glacier boundary. If a rock outcrop is on the boundary between two glaciers, then it is not internal to either, and should not be explicitly represented in the database. These two cases are illustrated in Figure 2. The case at left shows two adjacent glaciers with a rock outcrop between. In this case, each glacier should have a glacier boundary, but no “intrnl_rock” segments. The case at right shows a single glacier with a nunatak. This glacier should have a “glac_bound” polygon and an “intrnl_rock” polygon, both associated with the same GLIMS glacier ID.

7 Definition of “left” and “right” of segments

For assignment of the attributes "segment_left_material", "segment_right_material", "seg_left_feature", and "seg_right_feature", one must know how “left” and “right” are defined. For this purpose, glacier outline polygons will be considered to go counter-clockwise (as viewed from above), and internal rock polygons will go clockwise (Figure 3).

This definition makes definition of “left” and “right” unambiguous, and has the side benefit of simplifying area calculations. Given a rock-ice boundary, ice will always be on the left, and rock will always be on the right.

Analysts do not need to worry about digitizing segments in any particular sequence, since the ingest software will impose the correct ordering when inserting the data into the GLIMS Glacier Database. However, the above convention must be kept in mind when assigning the feature and material attributes to segments. (Assignment of these attributes is optional.)

8 Assigning GLIMS Glacier IDs

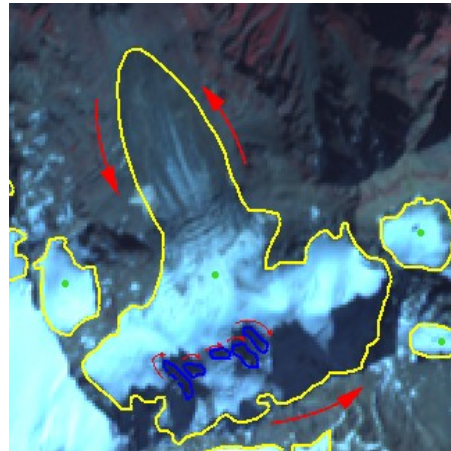


Figure 3: Illustrated circulation directions for glacier and internal rock boundaries define 'left' and 'right' for attributes.

Background

The ID associated with each glacier is the key that links the data in several parts of the database. The method used for assigning IDs should be fast and unambiguous. Use of an ID based on the World Glacier Monitoring Service (WGMS) system of counting drainages was considered. However, this would not work well in all areas, such as Greenland. An ID based on the latitude/longitude (L/L) position of a glacier is readily understandable by humans, well-defined over the entire earth, and is easily generated. (The database is designed to store WGMS IDs in addition for glaciers that have them. In all cases, though, the GLIMS ID is the primary key.) The uncertainty in the image position (for ASTER), based on orbit geometry, is 455 m per axis (3-sigma) (150 m for satellite position, plus 315 m for uncertainty in pointing)³. This is sufficient to allow unique IDs based on L/L since glaciers that are less than 455 m apart (center-to-center) are rare. For example, the small glaciers in the Pyrenees are on the order of 1 km apart.

Definition

The GLIMS glacier ID is of the form GnnnnnnEmmmmm[N|S], where [N|S] means “either N or S”, nnnnnn has the range [000000,359999], and mmmmm has the range [00000,90000]. The numeric values are in decimal degrees multiplied by 1000. Thus, the GLIMS ID of a glacier specifies a point on the ground, which should be chosen so that it can be identified with a glacier in an unambiguous way. As an example, the GLIMS ID for the Taku Glacier in southeast Alaska is “G225691E58672N”.

Choosing the Location for a GLIMS Glacier ID

A reasonable approach to take for deciding on the location of the GLIMS Glacier ID for a particular glacier is to compute the centroid of a polygon circumscribing the glacier by averaging the coordinates of the polygon, to come up with a single L/L position for the glacier. This approach is simple to implement, and may work well enough in many regions.

³From the ASTER document *End-to-End Data System Concept*

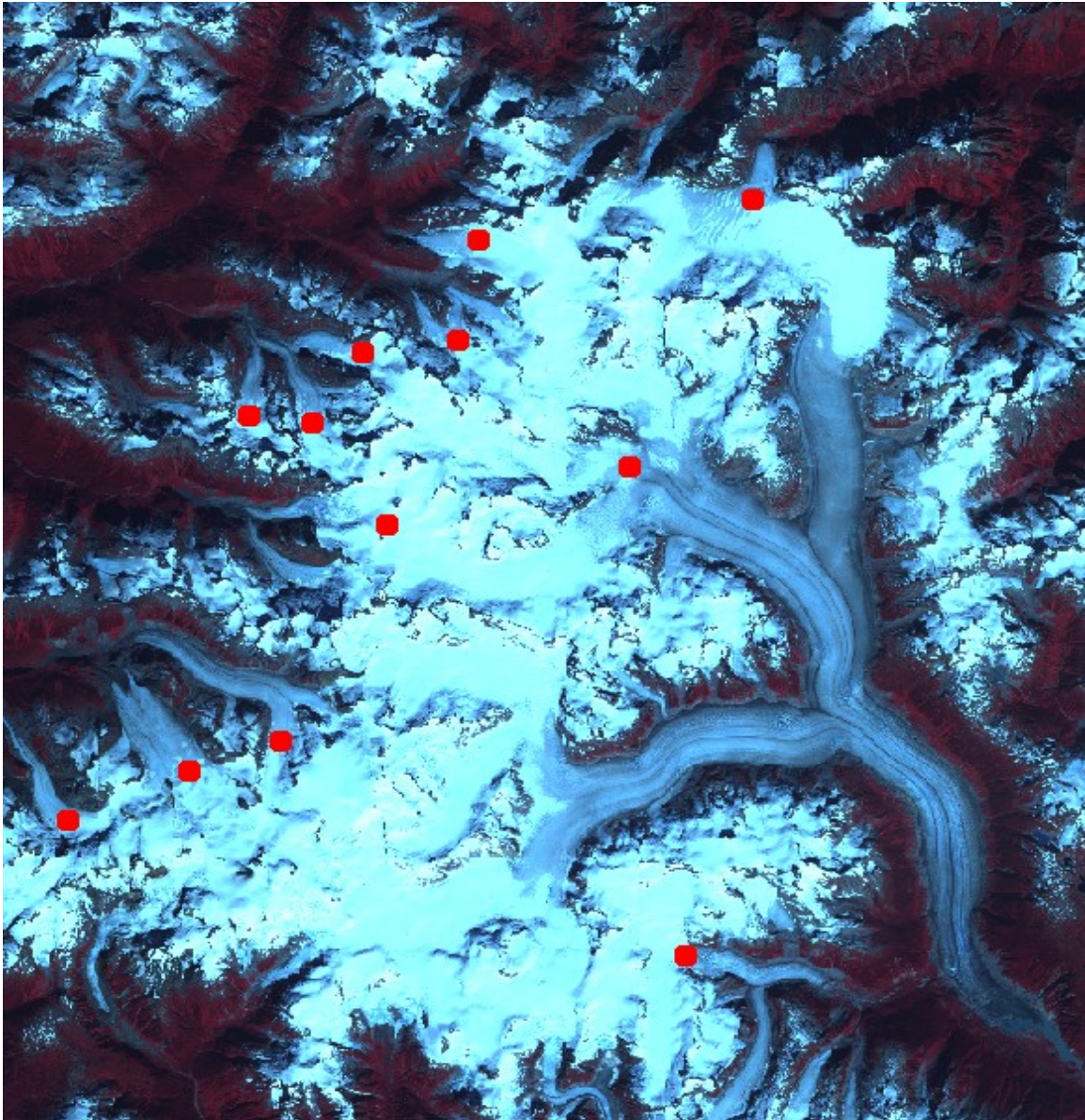


Figure 4 Example placements of GLIMS Glacier ID points on glacier visible in an ASTER image. Not all “glaciers” in the image have IDs placed; some small glaciers around the periphery have been omitted.

However, for glaciers that are retreating or may retreat markedly in the coming decades, it would be best to locate the ID point in the accumulation zone, on the approximate centerline, as shown in Figure 4. This method has the advantage of generating an ID that is likely to be easily identifiable with the glacier long into the future, but is difficult to automate.

When analyzing a glacier which already has an ID, one needs to use that ID in the current analysis. The method for doing this will vary depending on

what tools the analyst is using. In the future, we hope that GLIMSView will be able to query the GLIMS Glacier Database directly and present the likely ID(s) to the user.

9 Population of Other Fields (including Null values)

Null values

Fields that can be left empty should get an empty string (“”) as the value. For example, if the glacier does not have a name, rather than use "N/A", or “none”, simply leave the field blank.

As-of time (source time) of glacier outlines

The “as-of” time of glacier outlines are generally the acquisition time of the image from which the outlines were digitized. However, some data submitted to NSIDC for inclusion in the GLIMS Glacier Database has been based on maps, ground surveys, or photographs from various times. In such special cases, it is useful to know how the ingest software determines the as-of time. The precedence is as follows:

- src_time in “glaciers” shapefile
- timestamp from image specified in 'image_id1' in “glaciers” shapefile
- src_time in “session” shapefile
- timestamp from the image in the “images” shapefile

The as-of time stamp only needs to be specified in one of these places for a particular glacier. If it is specified in more than one place for a particular glacier, the first one found following the above order will be used.

10 Creating a “mugshot” of a glacier or set of glaciers

Each glacier outline should be identified with a GeoTIFF formatted “mugshot” image, when available, to show graphically what the glacier looked like at the time of image acquisition. This “mugshot” can be a sub-scene of the original image, or the full scene. In either case, full resolution is desired. The GeoTIFF format, which includes geolocation information, allows the GLIMS Web Map Server to show the glacier outlines over the image. While we prefer to receive such images in GeoTIFF format, we can accept a

number of formats that include geolocation information.

11 Measurement uncertainty

The GLIMS database is designed to store information about the measurement uncertainty associated with the glacier polygons. In the database, positional uncertainty (“CE90” -- CE90 of 20 m means 90% confidence that the given location lies within a circle of 20 m radius from the true location) is expressed as a distance in meters. There are two kinds of uncertainty that may be assigned to the points that form segments: global, and local. Global uncertainty is directly related to the accuracy of geolocation knowledge. Local uncertainty is the precision with which one can determine the position of a particular point on a glacier outline, for example.

It is important to indicate whether the analysis was done from orthorectified imagery or not. There is a Boolean field in the Segment table for this purpose (“orthocorrected”). This field, as well as others related to processing details, are now captured in the data submission interface, at <http://glims.colorado.edu/submissions/>.

12 Using GLIMSView as a filter

If glacier outlines are produced by some means other than digitization within GLIMSView, the analyst may use GLIMSView to transform those outlines and metadata into the required format for ingest into the GLIMS Glacier Database. First, define a new project and open the image that the outlines were digitized from. Then use the “import shapefile” function. **The data in the shapefile must be in longitude/latitude coordinates.** The outlines will be drawn on the image in green. Now the GlacierIDs must be defined, line types must be assigned to the polygons and then associated with the GlacierIDs. Finally, the session and image metadata are assigned in the respective fields.

In future versions of GLIMSView the user may be able to define the projection and spatial domain independent of an image.

Note that GLIMSView cannot work with shapefiles that contain multi-part polygons (polygons with holes, for example to represent internal rock boundaries). Thus, if you intend to use GLIMSView as a filter in this way, be sure to create separate polygons for objects (rocks, lakes) that are internal to the glacier boundaries.

13 Working with multiple images

Due to many possible circumstances, it may be desirable to derive glacier outlines from more than one image of a glacierized region. This may be due to cloud cover obscuring part of one image, or the ice mass being larger than one image. Further, the acquisition times of the images may differ from one another. If that is the case, then the analyst must decide whether to treat the analysis of both images as one "analysis session", or two separate analysis sessions. The criterion for this decision obviously is whether the two images are close enough in time so that the glaciers in that region have not changed detectably in the intervening time. If the analyst is confident that this is the case, he or she may choose to analyze the images as a unit (e.g. by mosaicking them together, or by combining the resulting glacier outlines after analysis). We leave this judgment call up to each analyst, since the critical time separation for images is likely to depend on the region and type of glacier system.

The GLIMS Glacier Database implements a many-to-many relationship between glacier analyses and images. That is, one glacier snapshot may be associated with many images, and one image may be associated with many glacier snapshots. The web-based interface to the database (<http://glims.colorado.edu/glacierdata/>) allows for constraining the viewed data to a specific time range. If a particular glacier analysis (snapshot) is associated with more than one image, and those images have different acquisition times, the latest image acquisition time will be taken as the "as-of" time for that glacier analysis. This will apply both for viewing data in the web interface, as well as for querying the data.

14 How the Ingest Process Works

After data in the transfer format are uploaded to NSIDC's GLIMS server, the files are examined for completeness, and the "ingest software" is run, using the uploaded shapefiles as input. The "glaciers" and "segments" shapefiles are opened, and as they are read, a data structure for each glacier -- identified by its GLIMS glacier ID -- is created. The "segments" file can be thought of as a big container of segments that are all mixed together. The ingest software "knows" how to group them only by GLIMS glacier ID. There is nothing in the segments file that indicates the order the segments should go in. So for each GLIMS glacier ID, the software finds all the segments that have that ID, and then further sorts by the "category" (glac_bound, intrnl_rock, etc.) attribute. For a given ID and category, it sorts the segments into topological order, making sure all segments connect end-to-end, and for entities that must be closed (glac_bound, intrnl_rock, etc.),

ensures that the segments form a closed polygon and circulate in the correct order. If the polygon is not closed, then the software explicitly closes it. The software then outputs SQL statements to input these data and metadata into the GLIMS Glacier Database.

In summary, in order for the ingest process to work, each record in the “segments” file must have, at minimum, a GLIMS glacier ID and a “category”. Internal rock polygons, debris-cover polygons, lake polygons, centerlines, and snow lines (categories “intrnl_rock”, “debris_cov”, “pro_lake”, “supra_lake”, “centerline”, “snow_line”) must all have the same GLIMS glacier ID as their corresponding glacier. See the GLIMS Data Transfer Specification for the details of data transfer format, http://www.glims.org/MapsAndDocs/datatransfer/data_transfer_specification.html.