My experiences with reproducible research:
Assessing glacier area and volume/mass changes

Tobias Bolch, Glaciology and Geomorphology Group (3G)
Contact: tobias.bolch@geo.uzh.ch
• How would you define reproducibility?
Have you ever tried, failed, or succeeded to reproduce another researcher’s results?

Example: Glacier area / outlines:

Glacier area of High Asia:

• Dyurgerov and Meier (2005): 116,180 km²

• Randolph Glacier Inventory (Pfeffer et al. 2014): 119,878 ± 9201 km²

• GAMDAM (Nuimura et al. 2015): 91,263 ± 13 689 km²

What are the reasons for the differences?
Glacier changes? Different data sources? Different methods? Different definition?
### Таблица 24

Распределение ледников Памиро-Алая по размерам их площади

<table>
<thead>
<tr>
<th>Районы</th>
<th>Менее 1,1 км²</th>
<th>1,1—5,0 км²</th>
<th>5,1—10,0 км²</th>
<th>10,1—25,0 км²</th>
<th>Более 25,0 км²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Северный Памир</td>
<td>357</td>
<td>121,8</td>
<td>147</td>
<td>336,9</td>
<td>19</td>
</tr>
<tr>
<td>Центральный и Западный Памир</td>
<td>1 775</td>
<td>527,3</td>
<td>322</td>
<td>709,3</td>
<td>39</td>
</tr>
<tr>
<td>Юго-Западный Памир</td>
<td>1 566</td>
<td>411,0</td>
<td>267</td>
<td>563,1</td>
<td>32</td>
</tr>
<tr>
<td>Восточный Памир</td>
<td>1 783</td>
<td>516,1</td>
<td>308</td>
<td>610,7</td>
<td>21</td>
</tr>
<tr>
<td>Всего по Памиру:</td>
<td>5 481</td>
<td>1 576,2</td>
<td>1 044</td>
<td>2 220,0</td>
<td>111</td>
</tr>
<tr>
<td>Гиссаро-Алай</td>
<td>3 440</td>
<td>951,0</td>
<td>409</td>
<td>810,6</td>
<td>28</td>
</tr>
<tr>
<td>Памиро-Алай</td>
<td>8 921</td>
<td>2 527,2</td>
<td>1 453</td>
<td>3 030,6</td>
<td>139</td>
</tr>
</tbody>
</table>

Dolgushin & Osipova, 1989
Inventories for the Himalaya

Nuimura et al. 2015
Changes of Gangotri Glacier / Garwhal Himalaya

Annual retreat rate of Gangotri Glacier by various authors
Mass change (m w.e. a⁻¹):
1970 – 2007: -0.32 ± 0.08
2002 – 2007: -0.79 ± 0.52

Gardelle et al. 2013:
1999 – 2011: -0.41 ± 0.21

Nuimura et al. 2012
2000 – 2008: -0.45 ± 0.60

Bolch et al. 2011, TC
• What kind of measures do you / the community take to improve the reproducibility of your research?

• And which possibilities do you see to improve the reproducibility of your research / the research of the community?

✓ Clear documentation of methods (incl. uncertainty assessment) and utilized data
✓ Making codes and results (e.g. glacier outlines, dh/dt data) freely available
✓ Standardized and individually adjusted metadata
✓ Making data and codes available as prerequisite for a publication

• What or which tools assist you in making your research reproducible?

✓ Commonly agreed recommendations/guidelines
✓ Databases to store the data (best with support)

Major problem: Time and pressure to publish...
Are researchers in your field making an effort to make their research reproducible?

- In general little willingness to share data freely and to make the research reproducible.
- In many cases it is hardly possible to reproduce the results.
- However, good exceptions exist and I feel an increasing willingness for data sharing and reproducibility.
- Many researchers are now willing to submit their data to international databases, e.g. GLIMS initiative (with a standardized database for glacier outlines) WGMS (here at GIUZ, talk will follow)
- One of the major journals in our field requires now a statement about how to access the data utilized in the study.
A tentative list of minimum requirements for a submission from a Regional Center is as follows:

- glacier outline
- GLIMS ID (based on the lat/lon location of a "centerpoint" on the glacier)
- Data source
- Date and time of analysis
- Analyst's name
- Analyst's institution
- Description of processing, including algorithms

The GLIMS initiative started in 1999, however a globally complete glacier inventory was only available in 2012 (RGI),

The GLIMS database is still not globally complete... Why?

- Technical issues but also lack of credit (The own paper is not cited, „just“ GLIMS).
Where in the research workflow do you see potential issues:

• data collection (gathering, field/lab work),
• preprocessing, analysis, (documentation),
• paper writing (documenting/log/reasoning, figure/table/text reproduction)?
Which experiences with failures/successes in making research reproducible could be helpful for other people at the institute?

- Use automated /standardized methods to produce data.
- Describe methods short but precise (use supplement if required).
- Good figures help a lot to understand and be able to interpret (and partly also reproduce methods/results).
- When using/citing data /numbers try to find the original sources.
- Establish an international standardized database and motivate colleagues to submit their data.
- Make the submission as easy as possible and keep the requirements for data submissions low.
- Ensure that the work done receives credits.
- Talk to / contact colleagues (was quite successful for RGI).
- Act as an editor / reviewer.