Project Description – Project Proposals

Ben Marzeion, University of Bremen

A future-ready Open Global Glacier Model (OGGM)

Project Description

1 State of the art and preliminary work

Glaciers form prominent features of many landscapes, and their global shrinking has become an icon of climate change. Glaciers are a source of geohazards (Richardson and Reynolds, 2000), are important regulators of water availability for downstream populations in many regions of the world (Huss, 2011; Huss and Hock, 2015; Immerzeel et al., 2012; Kaser et al., 2010), and are a major contributor to sea-level rise (e.g. Church et al., 2013; Gregory et al., 2013; Gardner et al., 2013; Cazenave et al., 2018).

There are more than 200.000 mountain glaciers in the world. Before the release of the first global inventory of glacier outlines in 2012 (the Randolph Glacier Inventory, Pfeffer et al., 2014), estimates of global glacier evolution were based on extrapolated observations and scaling relationships. Since then, a new branch of glaciology ("global scale glaciology") emerged to estimate (i) the volume of all glaciers worldwide (e.g. Huss and Farinotti, 2012; Grinsted, 2013; Farinotti et al., 2019) and (ii) their past and future evolution (e.g. Radić and Hock, 2011, 2014; Giesen and Oerlemans, 2012, 2013; Huss and Hock, 2015; Marzeion et al., 2012, 2015, and additional references herein)

Past glacier change can be estimated from direct glaciological measurements, e.g. of mass balances, but these have been performed on only about 300 glaciers worldwide. In recent years, great improvements have been made in large scale processing of satellite data, leading to geodetic mass-balance estimates at the regional scale covering periods of about a decade, sometimes more (e.g. Brun et al., 2017; Braun et al., 2019; Dussaillant et al., 2019). These estimates are neither temporally nor spatially complete, and therefore observation-based global estimates rely on a combination of observational products and statistical methods of variable complexity (Kaser et al., 2006; Zemp et al., 2019).

Numerical models are the only way to estimate glacier change outside the observational period (Giesen and Oerlemans, 2013; Huss and Hock, 2015; Marzeion et al., 2012, 2014; Slangen et al., 2012; Zekollari et al., 2019). Here, the evolution of glaciers is described by physical or statistical relationships implemented in a numerical model. Initial (e.g., glacier geometry, ice thickness) and boundary conditions (e.g., climate forcing) are obtained either from observations or other models. The family of regional to global numerical models of glacier evolution is growing rapidly: in a recent model comparison, Hock et al. (2019) compared the output of 6 different numerical models, and this number is likely to double in the upcoming phase of *GlacierMIP* (acronyms and software packages - in italic in the text - are listed at the end of this document).

The applicant‘s working group is widely recognized in the field of global scale glaciology and numerical modeling thanks to pioneering publications (see Sect. 1.1 and citations herein). Today, the universities of Bremen and Innsbruck jointly lead the development of the first global scale, open-source numerical model of glacier evolution including ice dynamics: the Open Global Glacier Model (oggm.org, Maussion et al., 2019).

OGGM is able to simulate the past and future evolution of almost all of the world's glaciers. It is a community model, which means that any interested individual or research group can add to, or enhance parts of, the model. It is the only global-scale model that is open source, which has attracted many users and contributors from external institutions. It forms the basis of several research publications (see Sect. 1.1) and research projects and proposals (Section Sect. 1.2) originating from international teams in Germany, Austria, France, Belgium, USA, UK, India, and China.

OGGM belongs to the larger family of "glacier centric models", where each glacier is simulated independently from the others (the approach chosen by most glacier models to date). The model is modular, with different modules ("tasks") addressing various elements of the modeling chain: pre-processing of static data (e.g. glacier outlines, digital elevation models), pre-processing of time-varying data (climate time series, validation data), ice thickness estimation, mass-balance computation, ice dynamics, glacier evolution and data post-processing. The current set of modules is described in detail in Maussion et al. (2019) and the online documentation.

Very few global models reach a similar level of physical complexity. Only the model by Huss and Hock (2015) and Zekollari et al. (2019) aims for similar complexity and has been a driver for many advances in global scale glaciology in recent years. It is, however, not open source and written in two different proprietary languages (IDL and Matlab). A recent model (*PyGEM*) developed by D. Rounce became open source, but currently is mostly limited to solving the mass balance of glaciers. Additional unpublished models are likely in development, but OGGM is undoubtedly one of the most modern and complete global glacier models available. It is unprecedented in terms of open development, modular structure and testing practices.

Precursor versions of OGGM were written in the Matlab language and were not freely available (Marzeion et al., 2012). Further, the precursor model by Marzeion et al. (2012) relied on statistical scaling laws to describe glacier area and length change. Since 2014, this model has been completely re-written in the Python language to become OGGM, and it has been greatly extended to explicitly resolve the dynamics of ice flow in glaciers of different shapes. As is frequent for scientific models, the code base has grown organically. Diverse recent additions include a glacier frontal ablation module (Recinos et al., 2019) and a method for reconstructing 20th century glacier evolution (Eis et al., 2019).

At the time of writing, the model has become the centerpiece of numerous PhD projects (we count at least 8) as well as of running, proposed or planned research projects (at

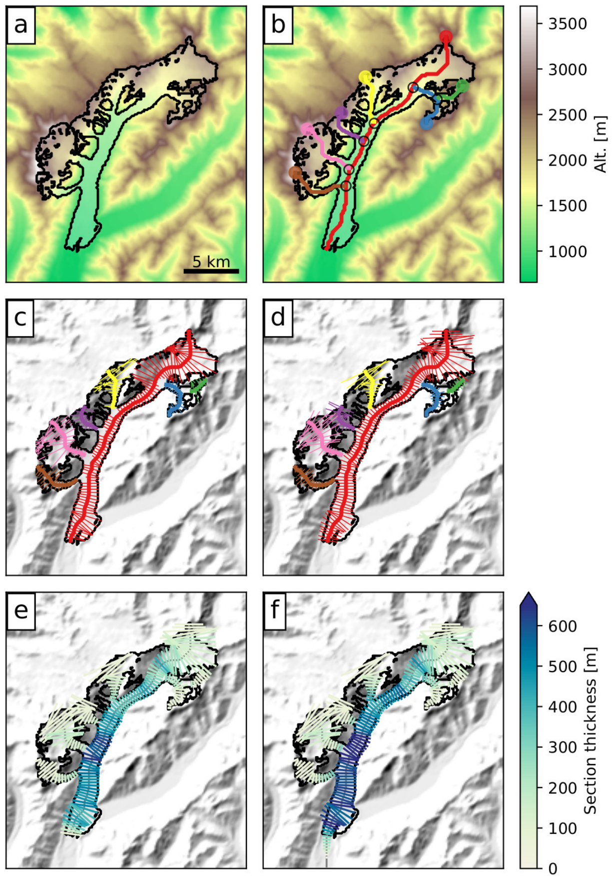


Figure 1: Example of the OGGM workflow applied to Tasman Glacier, New Zealand: (a) topographical data preprocessing; (b) computation of the flow lines; (c) geometrical glacier width determination; (d) width correction according to catchment areas and altitude-area distribution; (e) ice thickness inversion; and (f) random 100-year-long glacier evolution run leading to a glacier advance. See Maussion et al. (2019) for details.

least 15). While being a tremendous endorsement for the model, this growth raises new challenges for model development and structure (“Fit for Re-Use”), and puts an even larger pressure on the robustness and quality of the model (“Fit for Purpose”). These challenges are common to many large scientific software projects, but we discuss some of them in the context of OGGM here:

- The OGGM code base has grown and continues to grow organically, at the pace at which new features are added to the model. Future innovation will be slowed down and maybe even impeded by the so-called “technical debt”: code that works but would need considerable refactoring to adapt to new ideas and to allow further development. This refactoring (process of restructuring existing computer code) is not without costs: it takes considerable time, is not rewarded by traditional academic measures such as publications, and needs to be carefully conducted in order not to introduce new bugs into the software.
- Accumulating technical debt makes testing the software and ensuring the validity of its results very challenging or impossible. Certain functions might have been written in the past to take into account situations which are no longer an issue today, adding unnecessary complexity to the code. Well structured, well documented code is also easier to read, to test and is more likely to be reviewed by peers.
- As the software grows and increases in complexity, it becomes more difficult for new users to understand the model’s structure and application to their specific research problem. Furthermore, it is even harder for new users to envision a potential contribution to the code base, which will slow down future model development.
- A large part of routine software development maintenance currently relies on one or two main developers. The time of these main developers (who are scientists) is limited, and therefore the development process is slowed down. The way to deal with this issue is to increase the number of people with in-depth knowledge of at least a part of the code base in order to release the burden on the other core developers.
- The typical model contributors (PhD students and post docs) are employed on fixed-term contracts, causing a high turnover, making transmission of knowledge particularly difficult. Here, the solution is to rely on well-written and sustainable documentation, which will complement traditional peer-mentoring. Once a project has attained a sufficient number of expert users, this knowledge is also more sustainable and more likely to be shared.

In this project proposal, we seek dedicated funding to address the challenges outlined above. By modernizing its code structure, by implementing innovative testing methods and by offering interactive and attractive tutorials and documentation, we aim to make of OGGM a recognized standard for regional to global glacier change assessment and a flagship project for similar endeavors in the geosciences.

1.1 Project-related publications

Publications originating from previous model versions (Marzeion et al., 2012, model):

Gregory, J. M., White, N. J., Church, J. A., Bierkens, M. F. P., Box, J. E., Van Den Broeke, M. R., Cogley, J. G., Fettweis, X., Hanna, E., Huybrechts, P., Konikow, L. F., Leclercq, P. W., Marzeion, B., Oerlemans, J., Tamisiea, M. E., Wada, Y., Wake, L. M. and Van De Wal, R. S. W.: Twentieth-century global-mean sea level rise: Is the whole greater than the sum of the parts?, J. Clim., 26(13), 4476–4499, doi:10.1175/JCLI-D-12-00319.1, 2013.

Hock, R., Bliss, A., Marzeion, B., Giesen, R. H., Hirabayashi, Y., Huss, M., Radic, V. and Slangen, A. B. A.: GlacierMIP-A model intercomparison of global-scale glacier mass-balance models and projections, J. Glaciol., 65(251), 453–467, doi:10.1017/jog.2019.22, 2019.

Levermann, A., Clark, P. U., Marzeion, B., Milne, G. a, Pollard, D., Radic, V. and Robinson, A.: The multimillennial sea-level commitment of global warming., Proc. Natl. Acad. Sci. U. S.

A., 110(34), 13745–50, doi:10.1073/pnas.1219414110, 2013.
Marzeion, B., Jarosch, a. H. and Hofer, M.: Past and future sea-level change from the surface mass balance of glaciers, Cryosph., 6(6), 1295–1322, doi:10.5194/tc-6-1295-2012, 2012.
Marzeion, B., Cogley, J. G., Richter, K. and Parkes, D.: Attribution of global glacier mass loss to anthropogenic and natural causes, Science (345), 919-921, doi:10.1126/science.1254702, 2014.
Marzeion, B., Kaser, G., Maussion, F. and Champollion, N.: Limited influence of climate change mitigation on short-term glacier mass loss, Nat. Clim. Chang., 8, doi:10.1038/s41558-018-0093-1, 2018.
Parkes, D. and Marzeion, B.: Twentieth-century contribution to sea-level rise from uncharted glaciers, Nature, 563(7732), 551–554, doi:10.1038/s41586-018-0687-9, 2018.

Publications originating from the OGGM model:

Goosse, H., Barriat, P.-Y., Dalaiden, Q., Klein, F., Marzeion, B., Maussion, F., Pelucchi, P. and Vlug, A.: Testing the consistency between changes in simulated climate and Alpine glacier length over the past millennium, Clim. Past, 14(8), 1119–1133, doi:10.5194/cp-14-1119-2018, 2018.
Maussion, F., Butenko, A., Champollion, N., Dusch, M., Eis, J., Fourteau, K., Gregor, P., Jarosch, A. H., Landmann, J., Oesterle, F., Recinos, B., Rothenpieler, T., Vlug, A., Wild, C. T. and Marzeion, B.: The Open Global Glacier Model (OGGM) v1.1, Geosci. Model Dev., 12(3), 909–931, doi:10.5194/gmd-12-909-2019, 2019.
Recinos, B., Maussion, F., Rothenpieler, T. and Marzeion, B.: Impact of frontal ablation on the ice thickness estimation of marine-terminating glaciers in Alaska, Cryosph., 13(10), 2657–2672, doi:10.5194/tc-13-2657-2019, 2019.

1.2 Research projects making active use of OGGM

DFG financed:
MA 6966/1-1: Predictability and Attribution of Regional Sea Level Change Caused by Glacier Mass Change
MA 6966/1-2: Allocating Responsibility for Regional, Glacier-Related Sea-Level Change
MA 6966/4-1: Reconstruction of climate-glacier interaction on a centennial time scale
GRK-1904 (sub-project): Feedbacks between ocean circulation in the Canadian Arctic Archipelago and projected glacier mass change
FO 1269/1-1: Global glacier mass balance prediction on Seasonal and Decadal scale

Other funding agencies:
The Upper Grindelwald Glacier as indicator for Holocene climate variability (University of Innsbruck, Austria, funding agency: TWF)
Modelling of glacier length changes in the Alps on the base of tree-ring based temperature reconstructions for the last 2500 years (University of Innsbruck, Austria, funding agency: University of Innsbruck)
Modelling glacier changes over the past millennium (UC Louvain, Belgium, funding agency: Fonds de la Recherche Scientifique)
Projections of global glacier melt under low-end warming scenarios (University of Bremen, funding agency: BMBF FKZ: 01LS1602A)
Ocean-ice Interaction of peripheral Greenland Glaciers (University of Bremen, funding agency: BMBF FKZ: 03F0778)
Evaluating sediment Delivery Impacts on Reservoirs in changing climaTe and society across scales and sectors (University of Hannover, funding agency: BMBF FKZ: 01LS1902B)

Submitted (non-exhaustive list):
Impact of iceberg calving on global glacier change (Massachusetts Institute of Technology, USA)
Quantifying contributions from glaciers and terrestrial hydrology on recent and future sea level change (University of Alaska Fairbanks, USA)
Impacts Assessment to Support Water Resources Management and Climate Change Adaptation for China (University of Edinburgh, UK)
Estimation of the 21 st century contribution of Greenland's peripheral Glaciers to Sea Level through a synergistic Approach between model and satellite Data (University of Grenoble, France)

1.3 Current and potential use of the OGGM model

The model currently has a user base estimated to 10 or more research groups worldwide: Climate Lab (University of Bremen), Department of Atmospheric and Cryospheric Sciences (University of Innsbruck), Institute of Hydrology and Water Resources Management (Leibniz Universität Hannover), Earth and Life Institute (Université catholique de Louvain), Geography Department (Humboldt Universität zu Berlin), Institute of Geography (FAU Erlangen-Nürnberg), Department of Geography (University of Northern British Columbia), Indian Institute of Technology (Delhi), Institute of Tibetan Plateau Research (Chinese Academy of Sciences), and the Glaciers Group (Massachusetts Institute of Technology). Most of these research groups are in an early adoption phase, either testing the model for their research and/or applying for research proposals based on the model, further increasing the need for help and support from the current OGGM maintainers.

The potential user base is likely to be more than an order of magnitude larger. OGGM is used for research as well as for education (http://edu.oggm.org). It can be used to address pressing scientific and societal challenges, such as the estimation of the volume of all glaciers worldwide and their future contribution to sea-level change, their role as fresh-water reservoir, as indicator of past climate change, and much more. When further developed and advertised to its full potential, it will become a tool for other scientific disciplines, such as hydrology, earth system modelling, hazard assessment, and possibly also for the public and private sectors (e.g., hydropower and water resource management).

2 Objectives and work programme

2.1 Anticipated total duration of the project

The project is designed to run for a **total duration of 2 years**. To implement the project goals, we aim to employ a scientist at the post-doc level, either a software developer with demonstrated experience in scientific software development or a research scientist with demonstrated qualifications in software development. We plan to involve international partners from various disciplines (Sect. 5.3) to provide external guidance on our development strategy. One of the main roles of our collaborators will be to guide the development strategy and make sure that the model stays compatible with real use-cases from the community.

2.2 Objectives

In this project, we aim to:

A. **Develop state-of-the art software testing protocols using unit and integration testing**, and run a continuous integration suite **monitoring the model results and code changes over time**. This monitoring infrastructure will not only test current and future code for bugs, it will also track model results at large scale and send alerts to model developers when unexpected changes in model output occur after changes in the code. This will form the basis from which objective B can be reached.
B. **Apply modern software development techniques** (object oriented and modular programming) to refactor the current code base and **considerably reduce the "technical debt"** gathered by the software over the years. Our main objective is to **make the model as modular as possible** and to **encourage external contributions of innovative ideas within the OGGM framework**. This refactoring will occur in a parallel development branch of the model first, in an attempt to prevent backwards incompatible changes.
C. **Increase the model capacity to ingest and make use of the steadily growing wealth of glaciological data available for calibration and validation**, in order to **develop a protocol to estimate model uncertainties in an automated manner**. This is a pressing requirement as the number of users (and therefore the pressure on testing the validity of the model results in various situations) is increasing.
D. Adapt the existing code base to be **ready for the next evolution in numerical modelling and scientific computing: online interactive computing**, either on cloud or on high-performance computing infrastructures. In this new framework, **OGGM will be run using only a web browser**: this will considerably reduce the technical barrier to run the model and encourage further open exchange between users. By making OGGM ready for the cloud, we will take advantage of existing infrastructures and online data sets, thus further increasing model adoption by users without access to supercomputers. Furthermore, the use of **container technologies** will provide **reproducible and self-documented computing environments**.
E. **Extend the model documentation and develop a communication strategy to engage users and developers worldwide, and in developing countries in particular**. This will include software documentation, but also the development of interactive tutorials and the organization of training workshops.

Project "non-goals"

- This project does not aim to make of OGGM the single venue for future regional or global glaciological studies. We strongly believe in the usefulness of model inter-comparisons and on the necessity to have various ways to solve the same problem, for the sake of scientific repeatability and replicability. However, we will encourage model developers to make their model operable within the OGGM workflow, in order to facilitate and standardize future model inter-comparisons.
- This project does not aim to make significant improvements to the thematic scientific modules of OGGM. Not because it isn't necessary, but because the focus of this project should be bound to the aims listed above: operability, testing, data handling, robustness, reproducibility and documentation.

2.3 Work programme incl. proposed research methods

Each of this project's work packages (WPA, WPB, …) is associated to the main objectives listed above.

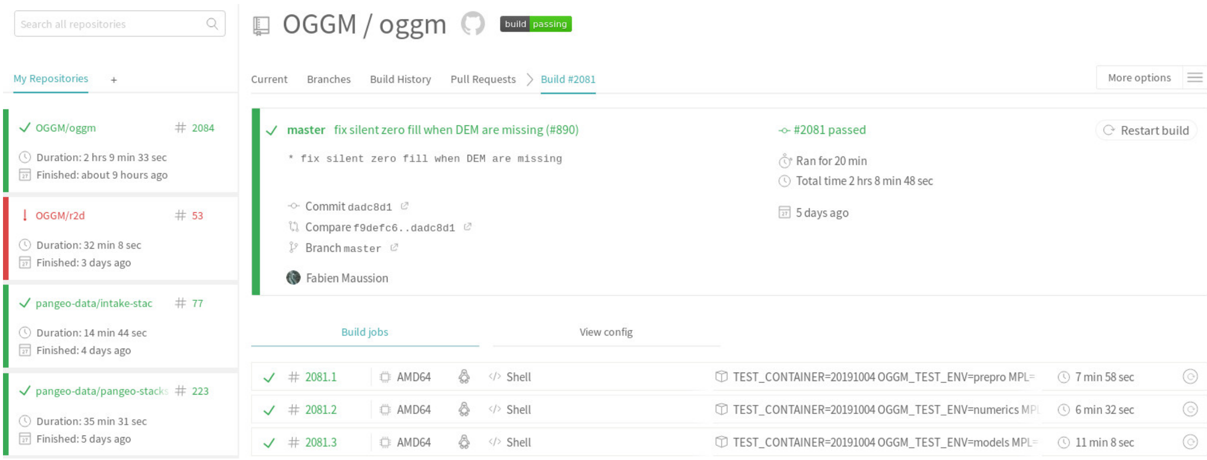


Figure 2: Screenshot of the Travis CI service running the OGGM test suite after a recent pull-request (see Glossary). These tests are run automatically on the cloud using a containerized environment.

WPA: Software testing protocols

In this WP, we will significantly update and extend the software testing code of OGGM. The current test suite is designed to run automatically on freely available resources thanks to the *Travis CI* service (this runs our test suite on dedicated virtual machines every time we make a change to the OGGM code base, a service called “continuous integration”, Fig. 2)

In a first step, we will revisit the available unit tests to use the testing framework *pytest* instead of the python standard *unittest* used currently. This will allow for faster and more flexible testing routines, and is likely to reduce the number of tests: indeed, past organic development lead to an unnecessary multiplication of the unit tests as a result of a “copy-paste” ad-hoc development: these tests could be merged for faster and clearer code. In this process, we will increase the so-called “code coverage”, i.e. the degree to which the source code is executed when the test suite runs. Our objective will be to reach more than 90% coverage, ensuring that most lines of code are visited at least once by our test suite: further coverage will require more extensive tests, as explained below. Where applicable, we will apply the concept of “property-based testing” by using the software package *hypothesis*. The current integration tests (checking that the program runs through a full modeling task chain) will also be revisited for better performance.

In a second step, we will drastically increase the number of cases tested by our test suite. Because a full OGGM test suite would incur high computational and data costs, we selected only a handful of test cases to be run in our current continuous integration suite (e.g. Hintereisferner in the Alps, Columbia glacier in Alaska). To achieve a better test coverage, we will need to run a more comprehensive test suite on our own computer resources instead: this will allow to run much more data intensive use cases, but will require the development of specific tools to do so (so called “hooks”, scripts being run automatically after a change in code, and reporting back after completion, similar to Fig. 2).

In a third and final step, we will develop a set of model tracking tools aiming at the automatic detection of unintended changes in model results and performance. This will aim to compensate for a major and unavoidable challenge of scientific software testing: Unlike traditional software, where the desired outcome of an action is known beforehand and therefore testable, numerical models often produce numbers that cannot be verified unequivocally. For example, the ice thickness of Hintereisferner used in our current test suite is not perfectly known. Our test cases are therefore checking that the model produces “reasonable” values according to a subjective assessment. There is no simple solution to this problem, other than validating smaller bits

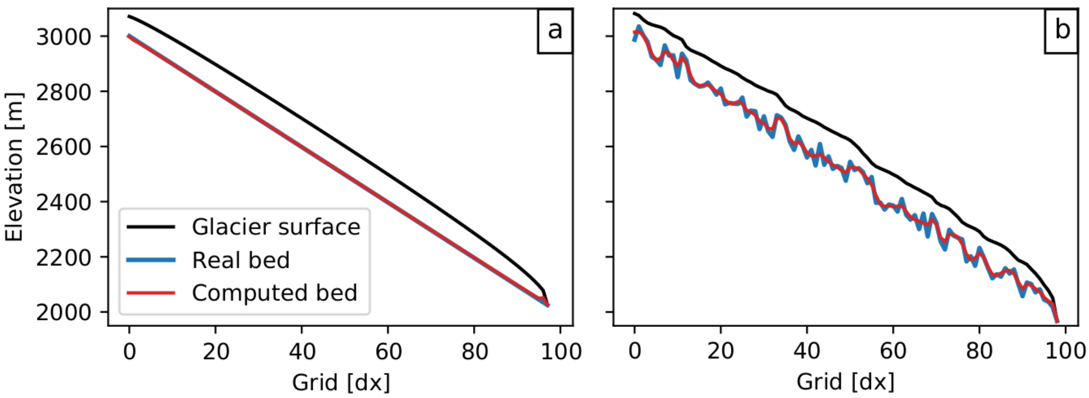


Figure 3: Example of "round-trips" tests used in the OGGM test suite. We compute the bed topography from the surface elevation obtained from a flow-line model applied to a predefined bed topography. See Maussion et al. (2019) for details.

of code which can be verified (e.g. comparing the simplified model output with an analytical solution) or by using different methods and verifying that they produce the same result (e.g. with "round-trips" or "perfect crimes", in which we check that the inverse model is able to reproduce a case generated by the forward model, as in Fig. 3).

For this project, we propose to add a series of benchmarks to monitor the repeatability and reproducibility of model results over time. These benchmarks can take the form of a standardized experiment on a large number of glaciers (> 100), where the simulation outcome is stored for later benchmarking of future model or configuration changes. Another kind of benchmark will ensure that the performance of certain modules (e.g. the mass-balance or ice thickness inversion module) remains constant or improves with time. These benchmarks will then be displayed online for anyone, and an automated script will send warning emails if an error or change occurs. A prototype of such a monitoring platform is already available (*oggm-crossval*).

Specific deliverables of WPA:

- improved and computationally efficient test suite
- increased code coverage and automated tracking of unintended side effects of code changes
- increased confidence in the model results

WPB: Code refactoring, modularity, and reduction of the technical debt

The early development of OGGM was realized step-by-step, towards the main goal of simulating global glacier change. For example, the development started by the automated computation of glacier centerlines using a geometric approach (Kienholz et al., 2014). These were then converted to another format by the task computing ice thickness. Format conversion occurred later in the development process and, as is now clear, would be unnecessary if the previous step had been written to generate the correct format. Similarly, the input climate data are currently parsed two times in the OGGM code base: once for the calibration of the mass-balance parameters, and a second time to provide mass-balance for the dynamical model. Here again, this code could be refactored to be used by both tasks without alteration. These are typical examples of easily diagnosed "code smells", which could be corrected with relatively little effort. However, there are two main obstacles preventing it: (i) any change in code is likely to introduce new errors (that is, one needs to be confident that the test suite is covering the parts that needs changing) and (ii) lack of time. Indeed, the rewards of such a refactoring have no visible impact on the model output (in the short term), and it is therefore unlikely to be conducted by a scientist focused on using the model for scientific applications.

In this WP, the hired software developer will work in close collaboration with the authors of the original code to discuss and implement a new structure based on modern principles of software design. Particular attention will be given to re-usability and readability, using object oriented programming where applicable. This refactoring will happen after the improvements in testing protocols (WPA) to minimize the risk of introducing new bugs.

A second major objective of this WP will be to increase the modularity of OGGM. By "modularity", we mean the capacity of the model to offer various solutions for a specific problem. We identify several key areas for such a task:

Surface mass balance: The user should have the choice between various ways to compute the surface (climatic) mass balance of glaciers, while still using the same ice dynamics model. OGGM currently allows this by providing an interface (a "class" in object-oriented programming) that model developers have to comply with in order to couple their mass balance model to the dynamical model. This has proven very stable and powerful, but has one drawback: it was developed with the current default OGGM model in mind and does not allow for more advanced models (e.g. with time steps shorter than one month, or with the ability to compute a spatially distributed mass balance on a 2D grid). In particular, it should be possible for models with a fundamentally different approach to calibration to be "pluggable" in the OGGM workflow. A perfect use case is given by the *PyGEM* model developed by our project collaborator D. Rounce, which should be able to use the pre-processing and glacier evolution modeling capabilities of OGGM with the mass balance module of *PyGEM*, ideally with as little modification as possible in both code bases.

Frontal ablation: The calving of icebergs is an important mechanism of glacier mass loss in many regions of the world, at high latitudes in particular. Including frontal ablation strongly influences the flux of ice and therefore modeled estimates of glacier volume (Recinos et al., 2019) and change (Bassis and Ultee, 2019). In a companion project led by our project collaborator L. Ultee, we plan to couple the *SERMeQ* model for frontal ablation (Ultee and Bassis, 2016, 2017) to OGGM. While the actual coupling is outside of this project's scope, making OGGM compatible with such extensions should be a general goal for the model, since changes in the model structure will probably be necessary to accommodate for this new process.

Glacier evolution: Currently, the model uses a network of interconnected flowlines to simulate the dynamics of glaciers and the evolution of their volume and geometry. OGGM should allow to use various approaches for the estimation of glacier change: either with simpler scaling models (e.g. Marzeion et al., 2012), simpler glacier geometry treatment (e.g. Huss and Hock, 2015), or even with a 3D ice dynamics model (e.g. Jarosch, 2008). It should also be possible to use different numerical solvers for the dynamical core (alternative solvers are currently used for testing purposes but are not documented). Again, implementing solvers is not the primary scope of our project: instead, in this WP, we will provide template infrastructures to implement alternative glacier evolution models with the example case of the scaling model by Marzeion et al. (2012) (implemented but undocumented) and with the simpler geometry treatment by Huss and Hock (2015) (not yet implemented).

Additional glacier processes: With recent improvements in computational methods and data availability, processes which were until recently not taken into account by global glacier models are now becoming applicable at large scales. This is the case for frontal ablation (see above), but also for other effects such as debris cover. While the influence of debris cover can be treated in the mass balance module (Kraaijenbrink et al., 2017), the transport of debris by glacier ice is tied to the dynamics of ice flow (Wirbel et al., 2018). Within this project, we would like to prepare OGGM for these future evolutions.

In comparison to more complex and interleaved coupled models, such as general circulation or ice-sheet models, modularity can be achieved at much lower development costs in OGGM. One main reason is the organization of the model workflow in "tasks", which are applied sequentially and individually for each glacier. Each task reads the necessary input from disk and writes its output to disk. This means that the various glacier evolution models to be implemented will have to agree on the format of the input and output data, and to comply with a specific calling syntax in Python, but do not have to comply to e.g. a specific multi-processing library or even a specific programming language.

In this WP, the role of the hired scientist will be to: (i) standardize and document these data formats, (ii) provide template implementations of simple alternative modules from which model developers can build upon, and (iii) provide support for future attempts to couple models to the OGGM dynamical core.

Specific deliverables of WPB:

- less error-prone and more reusable code
- facilitated access to the OGGM internals for present and future OGGM developers, with lower risk of inadvertent code duplication
- increased modularity via standardized interfaces and input/output formats

WPC: Data integration, model validation and uncertainty estimates

Since the first version of the model in 2012, tremendous advances have been made in remote sensing of glacier changes. In particular, the amount of geodetic mass-balance estimates from space has increased drastically, and homogeneous decadal estimates are now available at the continental scale (Brun et al., 2017; Braun et al., 2019; Dussaillant et al., 2019). Here we provide a non-exhaustive list of data sets that can be used by models like OGGM, together with their data source and their current state of use in OGGM:

Glacier outlines Source: *RGI*, *GLIMS*, individual studies. OGGM supports any glacier inventory, but users have to make their data resemble the RGI beforehand.

Digital Elevation Models (DEM) Source: various providers, not centralized. OGGM supports 10 different data sets available at the regional and global scale.

Gridded climate data Source: various providers, not centralized. OGGM supports various sources such as CRU, CMIP5, HISTALP. More sources are available and should be supported (e.g., ERA5).

Traditional mass-balance time-series Source: various providers, centralized in the *FoG* database at *WGMS*. Used in production by OGGM for calibration/validation.

Point glacier thickness measurements Source: various providers, centralized in the *GlaThiDa* database at *WGMS*. Used by OGGM in individual studies, but not standardized and undocumented.

Glacier length changes Source: various providers, centralized in the FoG database at WGMS and the "Leclercq" database (Leclercq et al., 2014). Used by OGGM in individual studies but not standardized and undocumented.

Geodetic mass-balance estimates Source: various providers, partly centralized at *WGMS* but often scattered as new publications become available. Highly inhomogeneous in space and time, probably requiring manual processing. Currently not used in OGGM.

Glacier surface velocities . Source: various, not centralized. Highly inhomogeneous in space and time, probably requiring manual processing. Currently not used in OGGM.

Making use of these numerous data sources and the process of combining these data with our model ("data assimilation") is a project per se and is outside the scope of this proposal (in fact, several planned or running projects plan to make use of these satellite products). However, **OGGM should be able to automatically access these data sets and download them for**

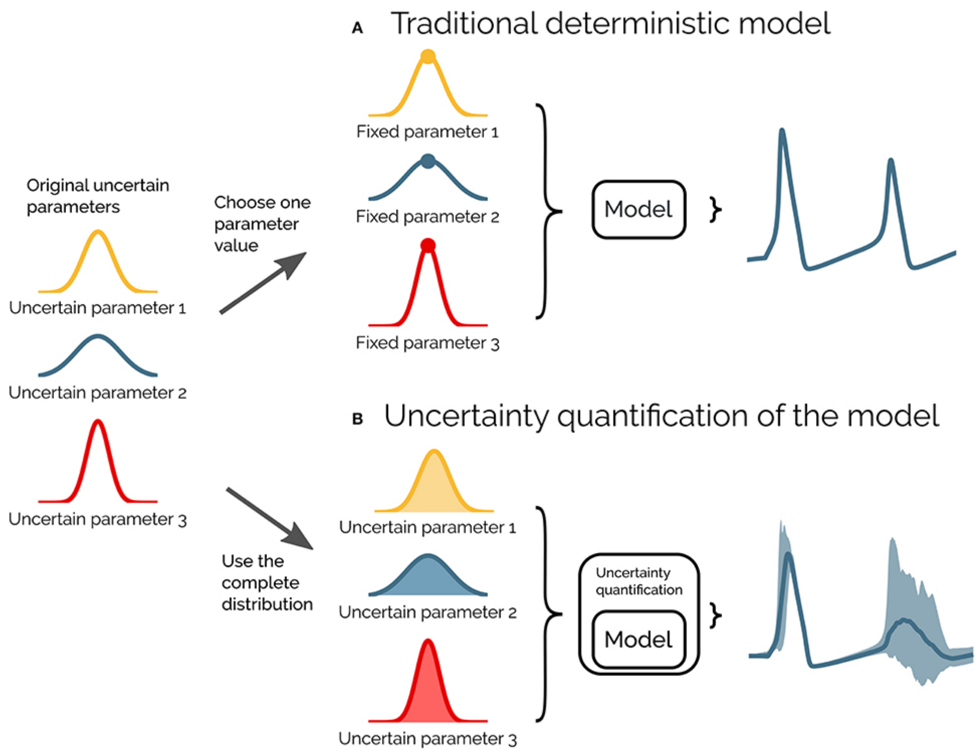


Figure 4: Illustration of uncertainty quantification of a deterministic model. (A) A traditional deterministic model where each input parameter has a chosen fixed value, and we get a single output of the model (gray). (B) An uncertainty quantification of the model takes the distributions of the input parameters into account, and the output of the model becomes a range of possible values (light gray). Illustration from Tennøe et al. (2018).

the user on command. This is already partly the case (e.g. for the DEMs and the reference climate data) but again, this development has been incremental and not well organized: today, a large part of the "untested" code in OGGM is related to such download utility functions. In this WP, we will make use of the *intake* protocol to better disseminate these data to OGGM users and allowing them to explore the data sets available to them. This will considerably simplify the current code, and will allow to standardize data access within and outside OGGM.

Some of these data serve the purpose of **model calibration and validation**: e.g. the mass-balance timeseries, length and area changes, etc. Currently, model calibration and validation in OGGM is limited to traditional mass-balance data (see the *oggm-crossval* platform), but this code is not well documented and not transferable. Like the vast majority of geoscientific models, OGGM currently does not provide validation tools to its users. The main reason for this choice is that each application has very specific needs: in the case of OGGM, the model can be used for millennial scale simulations of glacier length, for projections of future glacier change, to estimate the volume of glaciers, etc. Each of these use cases requiring a different approach to calibration and validation. Up to a certain extent, we would like OGGM to take over some of these responsibilities and **develop a new "uncertainties" module within the OGGM framework**.

Where applicable, OGGM should provide semi-automated and documented tools to compare model simulations with the products listed above. More importantly, OGGM should provide interfaces for parameter uncertainty quantification and the resulting model uncertainty. Following the example of the *uncertainpy* package (Tennøe et al., 2018, tailored for the neurosciences), we can make use of existing general frameworks for uncertainty quantification such as *chaospy* (Feinberg and Langtangen, 2015) and apply them to the OGGM specific problem at hand (see Fig. 4). In particular, the uncertainty module within OGGM will solve implementation issues related to random parameter selection and optimised distributed Monte-Carlo simulations (both requiring in-depth knowledge of model structure), while still allowing the flexibility required for the user's unique use-case. Finally, emphasis will be put on documentation. Experience shows that new users are often unaware of the influence that certain parameters can have on model output. With short, clear examples, users will be alerted about the probable sensitivity of their simulations to selected parameters.

Specific deliverables of WPC:

- an *intake* data catalog for all open data sets accessible from OGGM
- improved documentation on available data and how to access them
- an "uncertainties" module within OGGM offering tools and visualisation tailored for uncertainty estimation problems with the OGGM framework.

WPD: Interactive computing and reproducible science

Using OGGM on multiple processors is relatively simple: each glacier can run independently from the others on a single processor (a so-called "embarrassingly parallel" algorithm). Therefore, the time needed to run simulations scales linearly with the number of processors and the number of glaciers/simulations to run[1]. This feature encourages exploratory computing workflows, in which a user will develop a simulation and test it on a couple of glaciers on their laptop first, and later run the full simulation on their institution's high performance cluster (HPC).

Unfortunately, installation problems, data availability and internet access issues (firewall), as well as technical barriers (the command line, linux) make the transition to HPC very difficult. Furthermore, each user must take care of the installation of OGGM dependencies, leading to possible reproducibility issues (different package versions leading to different model results).

Therefore, an ever growing community of scientists and developers envision the future of scientific computing to be very different from the traditional HPC set-up which is the standard today (see e.g. the *Pangeo* project). Ideally, a scientist should not have to worry about how to translate a script for HPC or about package installation. Recent advances in interactive computing are making such a change possible today: the *jupyter* project and its ecosystem enable to develop and **run computationally expensive simulations using a user-friendly development environment** (the "jupyter notebook"), and use a familiar interface (the "jupyter-hub") to log-in into cloud or HPC infrastructures.

OGGM is already using these technologies to run interactive tutorials on https://edu.oggm.org (see Fig. 5). We also provide limited resources for workshops and classes (these were used, e.g., for a 3-days workshop for university students in Peru in summer 2019 by L. Ultee). However, it is not yet possible to run memory- and CPU-expensive simulations with these tools. The main developments necessary to reach this goal are presented in our documentation[2] and are realizable within the frame of this WP. As a result of this work package, selected users **will have access to our HPC infrastructures via their web browser**. Our IT will have control over the installed packages and therefore ensure that the setup is compliant with the model test suite. Internally, this is achieved with the use of *containers* (software "capsules") which will also be provided for everyone to use on their own system if preferred. With such containers, HPC systems with the *singularity* software installed (this is the case in Bremen and Innsbruck) will be able to run OGGM at minimal installation cost.

[1] This is only partly true: larger glaciers need longer to run, but mostly this can be compensated by running large glaciers first and use the remaining free resources for the many small glaciers.

[2] https://docs.oggm.org/en/latest/oeps/oep--0002-cloud.html

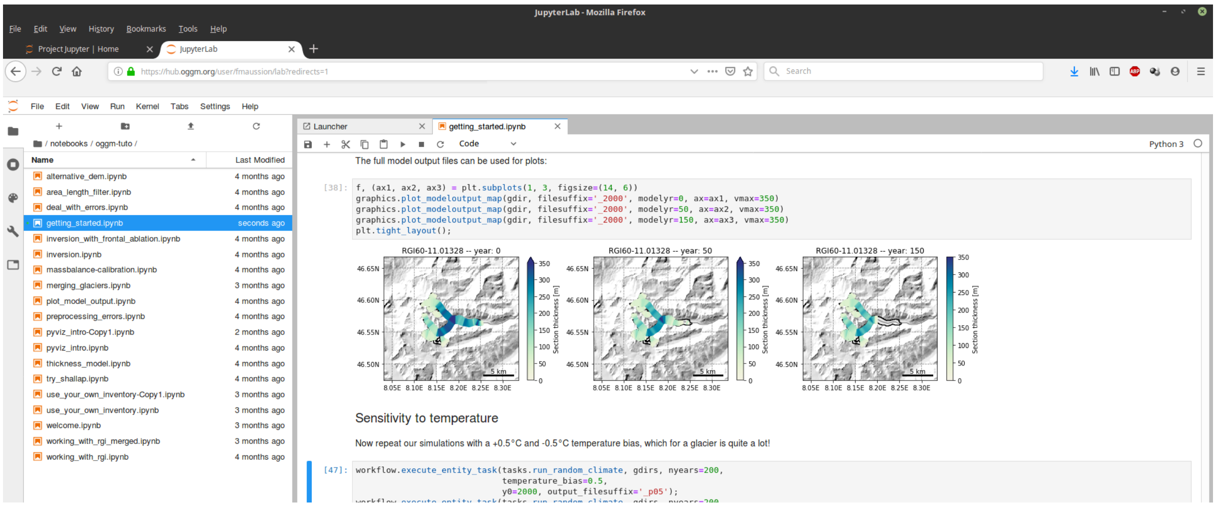


Figure 5: Screenshot of an OGGM tutorial running in https://hub.oggm.org. A "jupyter notebook" runs in "jupyterlab", which relies on "jupyterhub" to access virtual machines running on Google Cloud. The user doesn't need to know about these details, but can use a familiar environment on remote resources.

Specific deliverables of WPD:

- interactive platform to run OGGM in the web browser, for tutorials and for production runs
- free, documented and tested software containers to run OGGM without installation burden

WPE: Documentation, training and outreach

The documentation and outreach strategy of the OGGM project will rely on four pillars:

1. **Project website**: the project website, hosted at http://oggm.org contains general information about the project without too many technical details. It is the public facing part of the project and will be further developed within this project with blog posts, workshop announcements, etc.
2. **Online documentation**: the model's technical documentation (http://docs.oggm.org). It contains physical explanations as well as code examples, and the documentation of the OGGM internals (*API*). This documentation is generated automatically by our continuous integration service, ensuring that code and documentation remain synchronized (although it is still the developer's duty to ensure that the documentation is accurate). The documentation web page is the main venue to read about the model usage and its internals. To really learn how to use the model, we recommend a "learning by doing" approach and refer to our interactive tutorials.
3. **Interactive tutorials**: the OGGM tutorials (http://edu.oggm.org) are a collection of notebooks that can be either downloaded (for offline use) or run online interactively without installation. Currently, the tutorials are limited to basic functionality: in this WP, we will drastically increase the number of tutorials to cover many more advanced use cases, from the simulation set-up to the production on HPC resources.
4. **Training workshops**: in the course of the project, we will offer and advertise 3 training workshops which will be organised right after or during the main geoscientific conferences (EGU, AGU) in order to ensure a larger attendance. Ideally, we would offer a "sneak peek" into the model during the conferences for users to decide if the model is the right one for their use case, followed by a one day workshop for beginners. On the second year, expecting more attendees, we will offer two workshops in parallel: beginners and

advanced. In addition, we would like to offer a 3-days long workshop in a country of the Global South to promote open-source and scientific software capabilities. This workshop is envisioned to take place in either either Nepal, India, or Peru.

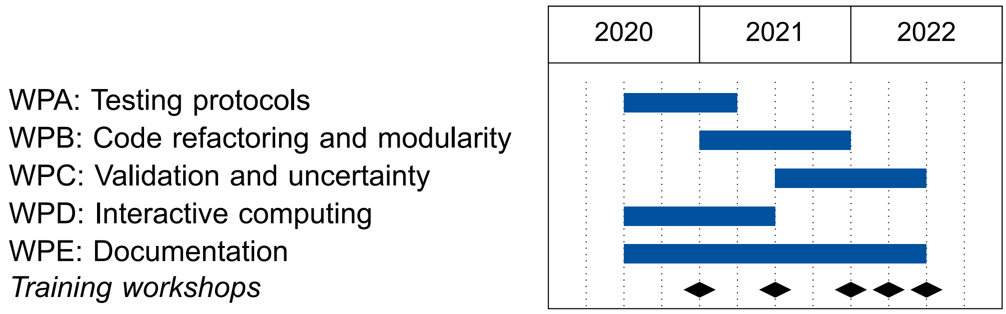
The proposed project will be pivotal to develop all four pillars by providing financial support and developer time: after the duration of the project, we expect all four pillars to continue to run on a low-maintenance, regular basis. We will make use of the traditional channels (mailing lists) and social media (e.g., Twitter) to advertise the model and related events.

Specific deliverables of WPE:

- accessible and up-to-date online documentation
- interactive online tutorials showcasing and explaining the model's workflow
- regular training workshops at conferences

Project management

We expect the management of this project to be straightforward, given that only one main person is involved in all WPs. We will make extensive use of online communication platforms (*Slack*, *GitHub*) to manage the code developments, and will ensure to be in frequent communication with our users in order to ensure a smooth transition to the newly developed model internals. The WPs to be addressed first will be WPA and WPD, because they form the basis of all later work. WPE (documentation and outreach) will run for the entire project duration. The training workshops will correspond to the main annual conferences.



2.4 Version control, software license and DOIs

All of OGGM's code is open-source. Developed under a LGPL v3 license, all contributors agreed to change to the more permissive MIT license in October 2019[3]. We run all code development operations and version control on the collaborative *GitHub* platform since 2016. Each time a new version is generated ("tagged"), the Zenodo platform generates a long-term archival of the code and provides a DOI. The most recent DOI is found at https://doi.org/10.5281/zenodo.597193.

2.5 Information on scientific and financial involvement of international cooperation partners

There is no direct financial involvement from our project partners, but a strong scientific involvement. For a list and description of our international and national collaborators, please refer to Sect. 5.3 ("Researchers with whom you have agreed to cooperate on this project").

[3] https://github.com/OGGM/oggm/issues/858

3 Bibliography

Bassis, J. N. and Ultee, L.: A Thin Film Viscoplastic Theory for Calving Glaciers: Toward a Bound on the Calving Rate of Glaciers, Journal of Geophysical Research: Earth Surface, 124, 2036–2055, doi: 10.1029/2019jf005160, 2019.

Braun, M. H., Malz, P., Sommer, C., et al.: Constraining glacier elevation and mass changes in South America, Nature Climate Change, 9, 130–136, doi: 10.1038/s41558-018-0375-7, 2019.

Brun, F., Berthier, E., Wagnon, P., et al.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geoscience, 10, 668–673, doi: 10.1038/ngeo2999, 2017.

Cazenave, A., Meyssignac, B., Ablain, M., et al.: Global sea-level budget 1993-present, Earth System Science Data, 10, 1551–1590, doi: 10.5194/essd-10-1551-2018, 2018.

Church, J., Clark, P., Cazenave, A., et al.: Sea Level Change, in: Climate Change 2013 - The Physical Science Basis, edited by Intergovernmental Panel on Climate Change, chap. 13, pp. 1137–1216, Cambridge University Press, Cambridge, doi: 10.1017/CBO9781107415324.026, 2013.

Dussaillant, I., Berthier, E., Brun, F., et al.: Two decades of glacier mass loss along the Andes, Nature Geoscience, 12, 802–808, doi: 10.1038/s41561-019-0432-5, 2019.

Eis, J., Maussion, F., and Marzeion, B.: Initialization of a global glacier model based on present-day glacier geometry and past climate information: an ensemble approach, The Cryosphere Discussions, pp. 1–21, doi: 10.5194/tc-2019-68, 2019.

Farinotti, D., Huss, M., Fürst, J. J., et al.: A consensus estimate for the ice thickness distribution of all glaciers on Earth, Nature Geoscience, 12, 168–173, doi: 10.1038/s41561-019-0300-3, 2019.

Feinberg, J. and Langtangen, H. P.: Chaospy: An open source tool for designing methods of uncertainty quantification, Journal of Computational Science, 11, 46–57, doi: 10.1016/j.jocs.2015.08.008, 2015.

Gardner, A. S., Moholdt, G., Cogley, J. G., et al.: A Reconciled Estimate of Glacier Contributions to Sea Level Rise: 2003 to 2009, Science., 340, 852–857, doi: 10.1126/science.1234532, 2013.

Giesen, R. H. and Oerlemans, J.: Calibration of a surface mass balance model for global-scale applications, The Cryosphere, 6, 1463–1481, doi: 10.5194/tc-6-1463-2012, 2012.

Giesen, R. H. and Oerlemans, J.: Climate-model induced differences in the 21st century global and regional glacier contributions to sea-level rise, Climate Dynamics, 41, 3283–3300, doi: 10.1007/s00382-013-1743-7, 2013.

Gregory, J. M., White, N. J., Church, J. A., et al.: Twentieth-century global-mean sea level rise: Is the whole greater than the sum of the parts?, Journal of Climate, 26, 4476–4499, doi: 10.1175/JCLI-D-12-00319.1, 2013.

Grinsted, A.: An estimate of global glacier volume, Cryosphere, 7, 141–151, doi: 10.5194/tc-7-141-2013, 2013.

Hock, R., Bliss, A., Marzeion, B. E., et al.: GlacierMIP-A model intercomparison of global-scale glacier mass-balance models and projections, Journal of Glaciology, 65, 453–467, doi: 10.1017/jog.2019.22, 2019.

Huss, M.: Present and future contribution of glacier storage change to runoff from macroscale drainage basins in Europe, Water Resources Research, 47, n/a–n/a, doi: 10.1029/2010WR010299, 2011.

Huss, M. and Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe, Journal of Geophysical Research: Earth Surface, 117, F04 010, doi: 10.1029/2012JF002523, 2012.

Huss, M. and Hock, R.: A new model for global glacier change and sea-level rise, Frontiers in Earth Science, 3, 1–22, doi: 10.3389/feart.2015.00054, 2015.

Immerzeel, W. W., van Beek, L. P. H., Konz, M., et al.: Hydrological response to climate change in a glacierized catchment in the Himalayas, Climatic Change, 110, 721–736, doi: 10.1007/s10584-011-0143-4, 2012.

Jarosch, A.: Icetools: A full Stokes finite element model for glaciers, Computers & Geosciences, 34, 1005–1014, doi: 10.1016/j.cageo.2007.06.012, 2008.

Kaser, G., Cogley, J. G., Dyurgerov, M. B., et al.: Mass balance of glaciers and ice caps: Consensus estimates for 1961–2004, Geophysical Research Letters, 33, L19 501, doi: 10.1029/2006GL027511, 2006.

Kaser, G., Grosshauser, M., and Marzeion, B.: Contribution potential of glaciers to water availability in different climate regimes, P. Natl. Acad. Sci. Usa., 107, 20 223–20 227, doi: 10.1073/pnas.1008162107, 2010.

Kienholz, C., Rich, J. L., Arendt, a. a., et al.: A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada, The Cryosphere, 8, 503–519, doi: 10.5194/tc-8-503-2014, 2014.

Kraaijenbrink, P. D., Bierkens, M. F., Lutz, A. F., et al.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers, Nature, 549, 257–260, doi: 10.1038/nature23878, 2017.
Leclercq, P. W., Oerlemans, J., Basagic, H. J., et al.: A data set of worldwide glacier length fluctuations, Cryosphere, 8, 659–672, doi: 10.5194/tc-8-659-2014, 2014.
Marzeion, B., Jarosch, a. H., and Hofer, M.: Past and future sea-level change from the surface mass balance of glaciers, The Cryosphere, 6, 1295–1322, doi: 10.5194/tc-6-1295-2012, 2012.
Marzeion, B., Cogley, J. G., Richter, K., et al.: Attribution of global glacier mass loss to anthropogenic and natural causes, Science, doi: 10.1126/science.1254702, 2014.
Marzeion, B., Leclercq, P. W., Cogley, J. G., et al.: Brief Communication: Global glacier mass loss reconstructions during the 20th century are consistent, The Cryosphere Discussions, 9, 3807–3820, doi: 10.5194/tcd-9-3807-2015, 2015.
Maussion, F., Butenko, A., Champollion, N., et al.: The Open Global Glacier Model (OGGM) v1.1, Geoscientific Model Development, 12, 909–931, doi: 10.5194/gmd-12-909-2019, 2019.
Pfeffer, W. T., Arendt, A. a., Bliss, A., et al.: The Randolph Glacier Inventory: a globally complete inventory of glaciers, Journal of Glaciology, 60, 537–552, doi: 10.3189/2014JoG13J176, 2014.
Radić, V. and Hock, R.: Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise, Nature Geoscience, 4, 91–94, doi: 10.1038/ngeo1052, 2011.
Radić, V. and Hock, R.: Glaciers in the Earth's Hydrological Cycle: Assessments of Glacier Mass and Runoff Changes on Global and Regional Scales, Surveys in Geophysics, 35, 813–837, doi: 10.1007/s10712-013-9262-y, 2014.
Recinos, B., Maussion, F., Rothenpieler, T., et al.: Impact of frontal ablation on the ice thickness estimation of marine-terminating glaciers in Alaska, The Cryosphere, 13, 2657–2672, doi: 10.5194/tc-13-2657-2019, 2019.
Richardson, S. D. and Reynolds, J. M.: An overview of glacial hazards in the Himalayas, Quaternary International, 65-66, 31–47, doi: 10.1016/S1040-6182(99)00035-X, 2000.
Slangen, A. B., Katsman, C. A., van de Wal, R. S., et al.: Towards regional projections of twenty-first century sea-level change based on IPCC SRES scenarios, Climate Dynamics, 38, 1191–1209, doi: 10.1007/s00382-011-1057-6, 2012.
Tennøe, S., Halnes, G., and Einevoll, G. T.: Uncertainpy: A Python Toolbox for Uncertainty Quantification and Sensitivity Analysis in Computational Neuroscience, Frontiers in Neuroinformatics, 12, 1–29, doi: 10.3389/fninf.2018.00049, 2018.
Ultee, L. and Bassis, J.: The future is Nye: An extension of the perfect plastic approximation to tidewater glaciers, Journal of Glaciology, 62, 1143–1152, doi: 10.1017/jog.2016.108, 2016.
Ultee, L. and Bassis, J. N.: A plastic network approach to model calving glacier advance and retreat, Frontiers in Earth Science, 5, 1–12, doi: 10.3389/feart.2017.00024, 2017.
Wirbel, A., Jarosch, A. H., and Nicholson, L.: Modelling debris transport within glaciers by advection in a full-Stokes ice flow model, Cryosphere, 12, 189–204, doi: 10.5194/tc-12-189-2018, 2018.
Zekollari, H., Huss, M., and Farinotti, D.: Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble, Cryosphere, 13, 1125–1146, doi: 10.5194/tc-13-1125-2019, 2019.
Zemp, M., Huss, M., Thibert, E., et al.: Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016, Nature, 568, 382–386, doi: 10.1038/s41586-019-1071-0, 2019.

4 Requested modules/funds

The total requested funds are: € 177'439

4.1 Basic Module

4.1.1 Funding for Staff

We request funding of a **full-time position at the post-doc level for the total duration of the project (two years)**. We will seek to hire either a software developer with demonstrated experience in scientific software development, or a research scientist with demonstrated qualifications in software development (shown e.g. by substantial contributions to open-source software). This person will be in charge of all WPs and will be in constant communication with OGGM users and developers. Total costs: € 144'000.

In addition, we request funding for a student assistant at the Bachelor or Master level for the total duration of the project, for 40 hours per month (€ 14.52 per hour). This student will be involved in the maintenance tasks of the projects such as the monitoring of the continuous integration suite, as well as other tasks such as data preparation, documentation, and helping with the organization of workshops and trainings. Total costs: € 13'939.

Total costs for Staff: € 157'939

4.1.2 Direct Project Costs

Travel Expenses (€ 6'000) We request funding for:

- two trips each year to Innsbruck for the project scientist (€ 500 per trip). The University of Innsbruck co-hosts the development of OGGM, and it is crucial for the project scientist to efficiently work and collaborate with people in both universities. Total costs: € 2'000.
- participation in two European conferences (€ 1'000 each) and one overseas conference (€ 2'000). These conferences will provide a platform to discuss and present our advances, as well as a perfect setting for OGGM training workshops (Sect. 4.2). Total costs: € 4'000.

Visiting Researchers (€ 6'000) We request funding for:

- a two-weeks long visit for each of our international project collaborators (Ultee, Rounce). These visits will serve the purpose of working on model coupling issues, and our collaborators will be asked to review the project developments to make sure that we continue to meet the needs of the user community. Total costs: € 5'000
- two one-week long visits to Bremen for each of our national and EU collaborators (Maussion, Förster). Total costs: € 1'000.

Project-related publication expenses (€ 1'500 We request the maximum allowed amount for publications (€ 750 per year). We expect the project to yield at least one direct publication in a specialized journal (e.g., Geoscientific Model Development) and many indirect publications thanks to the increased value and visibility of OGGM.

Total for Direct Project Costs: € 13'500

4.2 Module Workshop Funding

Training workshops for beginner and intermediate users of the OGGM model will be a fundamental element of our communication strategy. In order to minimize the costs and environmental impact of our workshops, we will organize them around large international conferences (EGU and AGU) and the regularly held, annual OGGM workshops. For these training workshops, we will rent two rooms for one days (beginner, intermediate) and ask the participants to cover their own travel expenses. We plan to offer three workshops at conferences and two workshops at OGGM meetings. Total costs (room rental and catering for 5 training days): € 2'500.

In addition, we would like to offer one 3 days workshop in a country of the Global South to promote open-source and scientific software capabilities in developing countries. This workshop could take place either in Nepal, India, or Peru. Two OGGM experts will travel to the workshop location (costs: € 2’000). Computer resources will be provided by our remote interactive computing resources (cloud or HPC), but we will need a room and a local organizing infrastructure (costs: € 1’500). Total costs: € 3’500.

Total costs for Workshops: € 6’000

5 Project requirements

5.1 Employment status information

Project PI: Ben Marzeion, Full Professor, Universität Bremen

5.2 Composition of the project group

We expect all current and future OGGM developers to be tightly involved in the proposed project. They will work closely with the hired scientist to make sure that the development closely follows the needs of both users and developers of the model. Furthermore, a permanent IT staff member in the applicant’s working group (Timo Rothenpieler) will be substantially involved in the project, in particular for WPA, WPD, WPE.

5.3 Cooperation with other researchers

5.3.1 Researchers with whom you have agreed to cooperate on this project

Fabien Maussion: Assistant Professor at the Department of Atmospheric and Cryospheric Sciences (ACINN), University of Innsbruck (Austria). Fabien Maussion is the main developer of OGGM and, together with Ben Marzeion, is co-leading the group of PhD students and Post Docs working on the project. He wrote parts of this proposal and will be co-PI on this project. He will be strongly involved in all aspects of project management and execution.

Lizz Ultee: Postdoctoral Associate at the Department of Earth, Atmospheric and Planetary Sciences (EAPS), MIT (Massachusetts, USA). Lizz Ultee is a glaciologist specializing in the dynamics of marine outlet glaciers. She is the main developer of the open-source calving model SERMeQ, which is planned to be coupled with OGGM in the near future. She will provide expert guidance for WPB (modularity) and WPE (documentation and training).

David Rounce: Postdoctoral Researcher at the Geophysical Institute, University of Fairbanks (USA). David Rounce is a glaciologist specializing on numerical modeling of glaciers at large scale. He is the main developer of the open-source glacier model PyGEM, which is planned to be coupled with OGGM in the near future. He will provide expert guidance in WPB (modularity) and WPC (data integration, validation and uncertainty estimates).

Kristian Förster Junior Professor at the Institute of Hydrology and Water Resources Management, Leibniz Universität Hannover (Germany). Kristian Förster is a hydrologist specializing in mountain hydrology. He is the PI of two projects making use of the OGGM model (DFG: FO 1269/1-1, BMBF: DIRT-X, FKZ 01LS1902B). He will provide expert guidance in WPB (modularity) and WPC (data integration, validation and uncertainty estimates).

5.3.2 Researchers with whom you have collaborated scientifically within the past three years

Listed below are the most relevant collaborations, i.e. first authors of papers I have co-authored, co-authors of papers I have first-authored, and partners in joint projects:

Surendra Adhikari, Dr. JPL, NASA, USA
Jonathan Bamber, Prof. Dr., Bristol University, UK
Michael Becht, Prof. Dr., Catholic University of Eichstätt, Germany
Matthias Braun, Prof. Dr., Friedrich-Alexander-Universität, Erlangen, Germany
Anny Cazenave, Prof. Dr., CNES and LEGOS, Toulouse, France
Gabriele Chiogna, Prof. Dr., Technical University of Munich, Germany
J. Graham Cogley, Prof. Dr., Trent University, Peterborough, Canada (deceased)
Markus Disse, Prof. Dr., Technical University of Munich, Germany
Brigitta Erschbamer, Prof. Dr. University of Innsbruck, Austria
Thomas Frederikse, Dr., Delft University of Technology, The Netherlands
Hugues Goosse, Prof. Dr., Catholic University of Leuven, Belgium
Florian Haas, Dr., Catholic University of Eichstätt, Germany
Wilfried Haeberli, Prof. Dr., University of Zurich, Switzerland
Tobias Heckmann, Dr., Catholic University of Eichstätt, Germany
Florian Herla, University of Innsbruck, Austria
Jochen Hinkel, Dr., Global Climate Forum, Berlin, Germany
Regine Hock, Prof. Dr., University of Alaska, Fairbanks, USA
Martin Horwarth, Prof. Dr., Technical University of Dresden, Germany
Alexander H. Jarosch, Prof. Dr., University of Iceland
Svetlana Jevrejeva, Dr., National Oceanography Centre, Liverpool, UK
Johnny A. Johannessen, Dr., Nansen Environmental and Remote Sensing Center, Norway
Torsten Kanzow, Prof. Dr., Alfred-Wegener-Institut, Bremerhaven, Germany
Georg Kaser, Prof. Dr., University of Innsbruck, Austria
Kirsty Langley, Dr., Asiaq Greenland Survey, Greenland
Paul Leclercq, Dr., University of Oslo, Norway
Marta Marcos, Dr., Mediterranean Institute for Advanced Studies, Mallorca, Spain
Fabien Maussion, Dr., University of Innsbruck, Austria
Daniele Melini, Dr., Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy
Matthias Mengel, Dr., Potsdam Institute for Climate Impact Research, Germany
Benoit Meyssignac, Dr., LEGOS, Toulouse, France
Martina Neuburger, Prof. Dr., Institute of Geography, University of Hamburg, Germany
Lindsey Nicholson, Dr., University of Innsbruck, Austria
J. Even Øie Nilsen, Dr., Nansen Environmental and Remote Sensing Center, Bergen, Norway
David Parkes, Dr., Catholic University of Leuven, Belgium
Frank Paul, Dr., University of Zurich, Switzerland
Norbert Pfeiffer, Prof. Dr., Technical University of Vienna, Austria
Monika Rhein, Prof. Dr., University of Bremen, Germany
Kristin Richter, Dr., University of Innsbruck, Austria
Riccardo Riva, Prof. Dr., Technical University of Delft, Netherlands
Christopher D. Roberts, Dr., ECMWF, Shinfield, UK
Gerard H. Roe, Prof. Dr., University of Washington, Seattle, WA, USA
Ursula Schauer, Prof. Dr., Alfred-Wegener-Institut, Bremerhaven, Germany
Aimée Slangen, Dr., NIOZ, Yerseke, Netherlands
Roderik S. W. van der Wal, Prof. Dr., Utrecht University, Netherlands
Ricarda Winkelmann, Prof. Dr., Potsdam Institute for Climate Impact Research, Potsdam, Germany

5.4 Scientific equipment

This project is very well prepared in terms of available resources and equipment. In addition to a full-time IT staff member at the University of Bremen, we have access to our own current small-size HPC and a planned medium-sized cluster dedicated to OGGM (early 2020). These HPC resources will cover all the needs for computational infrastructure for this project.

Furthermore, we have access to (and already make use of) several online services for international collaboration: *Slack* (project communication, management, teleconferences), *GitHub* (version control, code collaboration), *Google Cloud Compute Engine* (interactive computing, web hosting), *ReadTheDocs* (documentation).

Glossary, acronyms and tools mentioned in the proposal

API Application programming interface. In the case of OGGM, the name and documentation of the functions that users can apply in their simulations.
chaospy python package for performing uncertainty quantification using polynomial chaos expansions and Monte Carlo methods. Web: https://chaospy.readthedocs.io.
containers software "capsules" to download and run and where OGGM can run. We use the docker technology and already provide them to our users. Web: https://hub.docker.com/u/oggm.
singularity software able to run containers on HPC. It is more secure than Docker and therefore preferred by IT specialists on HPC. Web: https://sylabs.io.
FoG Fluctuations of Glaciers Database. Web: https://wgms.ch/data_databaseversions.
GitHub A platform for source code version control and collaborative development. Web: https://github.com.
GlacierMIP Glacier Model Intercomparison Project. Web: http://climate-cryosphere.org/activities/targeted/glaciermip.
GlaThiDa Glacier Thickness Database. Web: https://www.gtn-g.ch/data_catalogue_glathida.
GLIMS Global Land Ice Measurements from Space Database. Web: https://www.glims.org.
Google Cloud Compute Engine Cloud resources for computing and web hosting. Web: https://cloud.google.com/compute/.
hypothesis Property based testing. Web: https://hypothesis.readthedocs.io.
intake python package for finding, investigating, loading and disseminating data. Web: https://intake.readthedocs.io.
jupyter ecosystem of software tools enabling interactive computing on single and multiple computers. Web: https://jupyter.org.
oggm-crossval A continuous integration platform to monitor OGGM's mass-balance model performance. Web: https://cluster.klima.uni-bremen.de/~github/crossval.
Pangeo a community promoting open, reproducible, and scalable science. Web: http://pangeo.io.
pull-request A suggested change to the codebase. Maintainers can review the code and decide if changes are necessary before the code can be accepted. An example pull-request to the OGGM codebase: https://github.com/OGGM/oggm/pull/890.
PyGEM Python Glacier Evolution Model. Repository: https://github.com/drounce/PyGEM.
pytest A python package facilitating test development. Web: https://docs.pytest.org.
ReadTheDocs Building and hosting the OGGM documentation at each change in code. Web https://readthedocs.org, OGGM documentation: https://docs.oggm.org.
RGI Randolph Glacier Inventory (reference global dataset of glacier outlines). Web: https://www.glims.org/RGI/index.html.
SERMeQ Simple Estimator of Retreat Magnitude and Ice Flux (Q). Web: https://github.com/ehultee/plastic-networks.
Slack Collaboration platforms for teams. OGGM has access to a pro account. Web: www.slack.com.
Travis CI Continuous integration service used to build and test software projects hosted at GitHub. Web: https://travis-ci.org. It monitors the OGGM test suite at https://travis-ci.org/OGGM/oggm.
unittest The standard python testing framework (less flexible than pytest). Web: https://docs.python.org/3/library/unittest.html.
WGMS World Glacier Monitoring Service. Web: https://wgms.ch.