A Visual Analysis Concept for the Validation of Geoscientific Simulation Models

Andrea Unger[1], Sven Schulte[2], Volker Klemann[1,3], and Doris Dransch[1]

[1]GFZ German Research Centre for Geosciences, Potsdam, Germany,
{unger,volkerk,dransch}@gfz-potsdam.de
[2]Magdeburg-Stendal University of Applied Sciences, Germany,
sven.schulte@hs-magdeburg.de
[3]National Oceanography Centre, Liverpool, UK, volman@noc.ac.uk

Geoscientific modeling and simulation helps to improve our understanding of the complex Earth system. During the modeling process, validation of the geoscientific model is an essential step. In validation, it is determined whether the model output shows sufficient agreement with observation data. Measures for this agreement are called goodness of fit. In the geosciences, analyzing the goodness of fit is challenging due to its manifold dependencies: 1) The goodness of fit depends on the model parameterization, whose precise values are not known. 2) The goodness of fit varies in space and time due to the spatio-temporal dimension of geoscientific models. 3) The significance of the goodness of fit is affected by resolution and preciseness of available observational data. 4) The correlation between goodness of fit and underlying modeled and observed values is ambiguous. In this paper, we introduce a visual analysis concept that targets these challenges in the validation of geoscientific models – specifically focusing on applications where observation data is sparse, unevenly distributed in space and time, and imprecise, which hinders a rigorous analytical approach. Our concept, developed in close cooperation with Earth system modelers, addresses the four challenges by four tailored visualization components. The tight linking of these components supports a twofold interactive drill-down in model parameter space and in the set of data samples, which facilitates the exploration of the numerous dependencies of the goodness of fit. We exemplify our visualization concept for geoscientific modeling of glacial isostatic adjustments in the last 100,000 years, validated against sea levels indicators – a prominent example for sparse and imprecise observation data. An initial use case and feedback from Earth system modelers indicate that our visualization concept is a valuable complement to the range of validation methods.

1 Introduction

The Earth is a complex system, which can be described by the interplay of multiple simultaneous processes. A deeper understanding of this complexity is promoted by modeling and simulation. Towards the long-term objective to build a comprehensive computational model of the Earth system, geoscientists today make efforts to model the individual geoprocesses that contribute to it.

Building models of such real-world processes is a complex task that comprises various steps. Validation is an essential building block to ensure that the model correctly reproduces the real world process under study. To this end, automatically derived goodness of fit measures are commonly employed that determine the agreement between modeled data and observations of the process under study. In this regard, a number of challenges occur that are characteristic for the Earth sciences. These challenges include: the dependency of the goodness of fit on the model's parameterization, spatio-temporal variations of the goodness of fit, often limited resolution and precision of observation data, and an ambiguous relation between goodness of fit and underlying modeled and observed data. Especially limited resolution and precision of observation data hinder a rigorous analytical validation, which accounts for all these challenges, in numerous geoscientific applications. The limited observability of many processes on Earth often requires to rely on observation data that is sparse, unevenly sampled in space and time, or imprecise, e.g., when data is derived from interpretation rather than measurement.

To support a validation of geoscientific models in applications where a rigorous analytical approach is not feasible, we introduce a tailored visual analysis concept in this paper that makes the analytically derived goodness of fit assessable within an interactive visual interface. Our concept addresses the validation challenges by specific visualization components. A synthesized view on the individual aspects is accomplished by a twofold selection concept that supports a simultaneous exploration within model parameter space and within underlying data samples. Our solution has been derived in close cooperation with researchers from Earth system modeling. It extends previous work [11], in which we introduced the idea to support validation in geoscientific modeling and simulation with interactive visual analysis. The initial design study lead to a compact spatio-temporal visualization to assess the goodness of fit. Here, we specify and significantly extend this visualization strategy by a combination of interactively linked views, interactive aggregation, and parameter exploration of model ensembles. Consequently, the prior visualization now serves as one component within a larger visual analysis concept and has been adapted to refined analysis requirements.

The concrete implementation of our general concept depends on the application at hand. Data considered in the validation of geoscientific models – involving model parameterizations, model outputs, observation data, and automatically derived goodness of fit values – can considerably differentiate in dimensionality and scale among different modeling applications. We exemplify our concept for the application of modeling glacial isostatic adjustments. A use case and feedback from geoscientists indicate the potential of our visual analysis concept to substantially support validation in Earth system modeling.

2 Exemplary application: Modeling glacial isostatic adjustments

Glacial isostatic adjustment (GIA) terms the response of the solid earth to glacial loading processes. Most striking is the present uplift of northern Sweden by about 1 cm/yr after the melting of the Fennoscandian ice sheet 20,000 yr before present (e.g., [30]). The still ongoing process is governed by the response of the elastic lithosphere and the back-flow of displaced viscous mantle material, which hinders an instantaneous response to the unloading [40]. The deformation process contributes a linear signal to a large number of geodetic observations. A prominent example for this is the interpretation of gravity data from the GRACE satellite mission, which was designed for the research of climate processes [31]. It determines mass redistributions at the Earth's surface by its gravitational signature. Effects due to other processes – like GIA – have to be corrected for.

The only observable that covers the whole time length of the GIA process are sea level indicators. They record sea level variations over thousands of years, allowing for a better understand-

ing and quantification of this process. Sea level indicators are fossil samples, archaeological constructions or other features, which allow, often only imprecisely, to reconstruct the local sea level at the time of formation or activity. This imprecision is due to dating or the character of the sample. Also, sea level indications are relative, with respect to both varying sea level and varying elevation of the land mass [16].

3 Challenges in the Validation of Geoscientic Models

Validation is an essential building block in model development [35]. It refers to " building the right model" [6], that is, to ensure that the model reproduces the real-world process under study sufficiently with respect to the modeling objectives.

The goal of this section is to describe the validation approach that we refer to throughout this paper. General challenges of this analysis process in the Earth sciences are exemplified with respect to validating GIA modeling against sea level indicators.

3.1 Analytical Validation Approach

The validation approach followed in this paper belongs to the group of data validation methods [6], which perform a comparison of model output with observation data. It involves a computationally derived measure for the agreement between modeled and observed values, denoted as goodness of fit f [21]. Numerous methods exists to determine the goodness of fit [20]. Due to limited availability and preciseness of observation data, we restrict our approach to a point wise comparison of individual modeled and observed values (Fig. 1) This approach requires that the model output follows the sampling of observation data. For each sample i, the local goodness of fit f_i is derived as the agreement between corresponding modeled and observed value. Naturally, also the sampling of f_i follows the observation data. An aggregation of f_i is performed to derive an average goodness of fit f of a model. In general, f is a value in the range $[0,1]$, with $f = 1$ indicating a full agreement of model output and observation data, and $f = 0$ indicating no agreement.
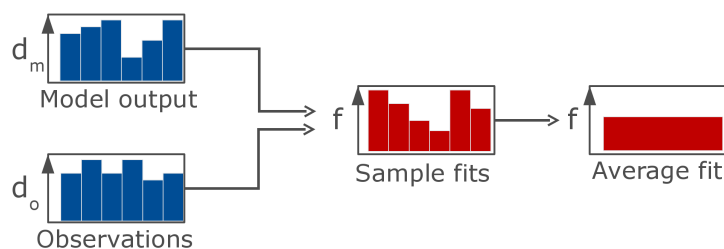

Figure 1: Analytical approach to determine the model's goodness of fit. Modeled and observed values (d refers to relevant variable) are compared point by point, resulting in a set of sample fits f_i. These are aggregated in a single value f that characterizes the model's goodness of fit.

3.2 Analysis Challenges

The goodness of fit of a geoscientific model depends on four different aspects, which we discuss in the following. We illustrate these challenges by our exemplary application of validating modeled glacial isostatic adjustments in the past against sea level indicators.

3.2.1 Dependency of Goodness of Fit on Model Parameterization

A major interest in the described validation approach is to determine the model's goodness of fit for potential values of model parameters. Model parameters describe the factors that have a major impact on the system behavior. A common approach to address the challenging task of parameter quantification in Earth system modeling is to evaluate the goodness of fit for potential parameter values.

Exemplified: Model parameters in GIA modeling In our GIA model, three input parameters are evaluated, which describe how the Earth surface deforms due to ice masses. The parameters describe the thickness of the lithosphere L, and the viscosities of upper Earth mantle UM and lower Earth mantle LM. Each parameter is given as a scalar value.

3.2.2 Variations of Goodness of Fit in Space and Time

Commonly, a process that contributes to the Earth system takes different effect on different spatial locations on Earth and in different time spans. An important aspect of Earth system modeling is to reproduce these spatio-temporal variations. But multiple model parameterizations match these variations with differing accuracy in space and time. This is reflected by spatio-temporal variations of the goodness of fit.

Exemplified: Spatio-temporal variations of goodness of fit The GIA model estimates sea level variations in the last 100,000 years due to glaciation and deglaciation. These processes, caused by global climate changes cause over time, induce temporal changes in sea level heights. GIA model output reproduces sea level height in different time spans and regions with diverging accuracy. Also, the model does not regard all processes that contribute to variations in sea level height. But all these processes are captured in the observation data.

3.2.3 Resolution and Preciseness of Observation Data

The availability of meaningful observational data is a restraining factor in data validation. Spatio-temporal coverage and preciseness are determined by the applied observation method. In the last years, technological advances have massively facilitated the observation of the Earth system. Despite the availability of novel technologies such as satellite missions and sensor networks, which produce data with before unknown spatio-temporal resolution and precision, the observability of many processes of the Earth system remains limited. Here, scientists have to rely on sparse, unevenly distributed point sets with limited precision of values. This is especially true for geoprocesses in the past, as it is the case in the application at hand.

Exemplified: Observation data derived from sea level indicators To validate GIA modeling, fossil samples that give indications about former sea level are the only available source of real-world observations. These sea level indicators exemplify limitations in data distribution and preciseness very well:

- The spatio-temporal distribution is sparse and uneven.
  Sea level indicators are restraint to their find spot and their span of life. Geologists search for sea level indicators near assumed past seashores. Hence, the available amount of samples in these geographical areas is higher than in other areas. Fig. 2 shows the distribution in space and time for our exemplary data.
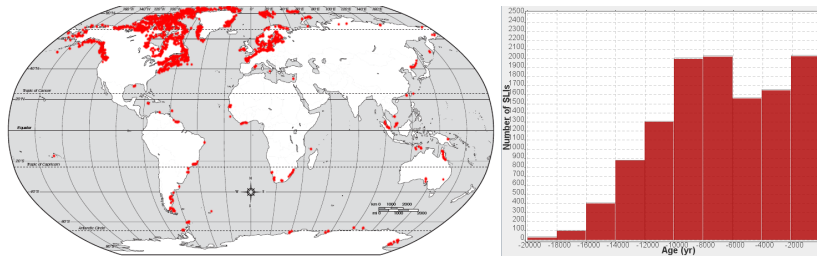
Figure 2: Distribution of sea level indicators in space (red dots on map on the left) and in time (temporal histogram on the right).

- Temporal locations and derived sea level heights are imprecise.
  While the geolocation of a fossil sample is determined from its find spot with relative precision, the life span is bounded by an aging method. Inevitable dating errors limit the fossil's age by a probability function. Further, as the denotation says, sea level indicators provide indications of sea level height rather than measurements. Depending on geological knowledge of the kind of fossil (sea shells, peat, corals), sea level height can be bounded in one or two directions. In the result, their are three different kinds of indications of the sea level height: upper bounded ranges, lower bounded ranges, and double bounded ranges (Fig. 3).
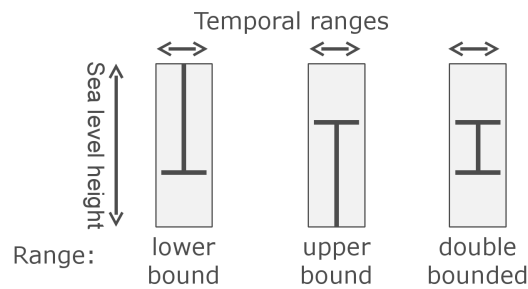

Figure 3: Impreciseness of observation data. Each observation data has an associated temporal range and value range. Three types of value ranges occur: Lower bounded ranges, upper bounded ranges, and double bounded ranges.

3.2.4 Dependency of Goodness of Fit on Underlying Values

The goodness of fit results from an analytical comparison of modeled and observed values, but modeled and observed value cannot be recovered from the goodness of fit. For the analysis though, the two-sided relation between goodness of fit and underlying values is an important information in the assessment of the goodness of fit.

Exemplified: Goodness of fit for observed bounds and modeled values The dual impreciseness of the observation data (in age range and observed value range) demands a tailored analytical method to determine the goodness of fit. The applied method is based on fuzzy logic [16], which leads to fit values f_i in the continuous range $[0, 1]$. Fig. 4 illustrates the resulting mapping of modeled and observed value to goodness of fit: If the modeled value falls into the bounded range given by the sea level indicator, then $f_i = 1$, otherwise the goodness of fit depends on the distance to the bounded value range. f_i is ambiguous in two ways: For $f_i = 1$, the

5

analyst does not know the position of the modeled value within the value range, which can lead to irrelevant goodness of fit values for half bounded ranges. For $f_i < 1$, it is not known whether the modeled value is an over- or underestimate of the observed sea level height.
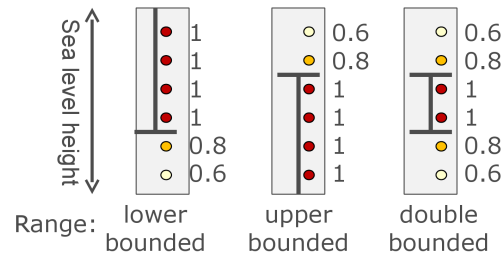

Figure 4: Deriving the goodness of fit for different types of value ranges provided by observation data and corresponding modeled values. Dots indicate potential modeled values, their color indicates the resulting goodness of fit as labeled.

3.3 Visualization Requirements

In the following, we present the requirements for a visualization concept that supports the validation of geoscientific models. The requirements were derived in an iterative process, comprising unstructured interviews with Earth system modelers and the evaluation of preliminary prototopes, which triggered further discussions and a constant sharpening of the analysis demands. The ongoing cooperation helped to substantially broaden our understanding of visualization requirements compared to prior work [11]. Our previous findings are now subsumed in visualization requirement VR 2.

VR1 Exploration of goodness of fit in model parameter space

A central requirement for successful validation is the assessment of variations of the goodness of fit that are due to varying model parameterizations. This includes the following tasks:

- Explore distribution of goodness of fit in model parameter space
- Identify influence of individual parameters or combinations of parameters on goodness of fit

VR2 Exploration of spatio-temporal variations of goodness of fit

Spatio-temporal variations of the goodness of fit are induced by varying accuracy of modeled values – or other geoprocesses reflected in observed values. These variations need to be assessed on appropriate abstraction levels, as the scope of analysis can be global or restraint to a local region, which involves to:

- Evaluate spatio-temporal variations of goodness of fit on appropriate abstraction levels, according to the current analysis scope
- Evaluate the significance of goodness of fit in regions or time spans by implicit geoscientific knowledge about inferring geoprocesses

VR3 Exploration of observation data

Preciseness and resolution of observation data constrain the significance of the goodness of fit, which needs to be assessed by exploring the observation data. The analysis involves to:

- Explore spatio-temporal distribution of observation data

- Explore preciseness of observation data
- Evaluate significance of observation samples by additional geological knowledge

VR4 Exploration of relation between goodness of fit and data values

The ambiguous relation between goodness of fit values and underlying modeled and observed values needs to be clarified in validation. Hence, the following tasks need to be supported:

- Identify modeled and observed values for the goodness of fit of a single sample or a group of samples
- Compare modeled and observed values manually

VR5 Combined view on dependencies of goodness of fit

The manifold dependencies of the goodness of fit are closely related and cannot be isolated from one another. A synthesized view is required to support the following objectives:

- Explore dependency of spatio-temporal variations on model parameterization
- Explore dependency of goodness of fit in model parameter space for a sample subset, which groups samples, e.g., by spatial location, temporal range, or type of fossil
- Determine significance of goodness of fit in spatial-temporal regions based on coverage and preciseness of observation data

4 Related Work

The goal of our work is to provide a visual analysis concept that targets specific challenges in the validation of geoscientific models. Thereby, we want to narrow the gap between state-of-the-art methods of interactive visual analysis and their application in Earth system modeling. This general gap still exists, as it was recently shown in a survey by Tominski et al. [33] among 76 scientists from the closely related field of climate research. Nevertheless, the last years have shown increased efforts in bridging the gap by developing tailored visual analysis methods that consider domain-specific challenges.

In Earth system modeling, Ahrens et al. [1] address the verification of simulation models by comparing various modeling outputs. Here, the model output arises from different code implementations of the same model, which agrees with the objective of verification to check whether a conceptual model has been implemented correctly. In contrast, validation answers the question whether the correct conceptual model has been built.

Validation in our sense considers a model ensemble that arises from unknown model parameterizations. Various visual analysis approaches target such model ensembles in Earth system modeling or closely related fields. Applied to climate research, Potter et al. [27] and Kehrer et al. [14, 18, 15] employ multiple linked views for the analysis of spatio-temporal ensemble data. Model ensembles in weather forecasting are addressed by Sanyal et al. [28]. An alternative approach to explore parameter spaces in Earth system modeling that employs self organizing maps is provided by Boschetti et al. [8]. Addressing the objective of deriving fast response strategies in flooding hazards, which also strongly relies on the dependency of model output on model parameterization, Waser et al. [38, 39] focus on interactive steering of simulation. These approaches link multiple model parameterizations and their spatio-temporal model output. However, the goodness of fit derived from comparison with observation data – a major requirement in our intended data validation – is not considered. A rare example for a comparative visualization of model output and observation data is presented by Nocke et al. [25]. Focusing solely on observation data, Yuan et al. [41] present visual analysis methods for gaining insight in earthquake-related data from different sources, Köhler et al. [17] employ self organizing maps to extract features from seismic data.

The presented related work encouraged our research, as it presents powerful visual analysis methods for Earth system modeling and closely related domains. Visual support in the modeling and simulation cycle in other domains has been presented by Matkovic et al. [24, 23] for engineering applications and Unger and Schumann [36] and Schulz et al. [29] for cell biology. While being valuable for the respective analysis objectives, the presented approaches do not support data validation in Earth system modeling as intended by our application – which requires a simultaneous consideration of various parameterizations and outputs of a model ensemble, observation data, and computationally determined goodness of fit values. More closely related to this challenge, Piringer et al. [26] target the validation of regression models, with an application in car engine engineering, thus not explicitly considering spatio-temporal context.

Even though we are not aware of a visual analysis approach that supports a combined consideration of all our visualization requirements, prior work targets individual visualization requirements that we identified in Section 3.3. Regarding the exploration of parameter spaces, Bruckner and Möller [9] turn towards visual effects design, Franken [13] towards exploration of algorithm parameter spaces, and Berger et al. [7] allow the exploration of continuous parameter spaces. Torsney-Weir et al. [34] focus on the application of parameter finding in image exploration. Their analysis involves the comparison of model output to ground-truth images, thereby determining the goodness of a parameterization. Despite the fact that a different application domain is targeted, evaluating the goodness of model outputs in parameter space is closely related to the domain task targeted in our work.

Another relevant visualization requirement in our application is the exploration of spatio-temporal context of data. Guidelines and numerous techniques for the visualization of geospatial data [22, 12] and for temporal data [2] are provided in the literature. Still, the visualization of geospatio-temporal data is a challenge that currently gains much attention [5, 4]. According to Chung et al. [10], the spatio-temporal context in our application can be described as event data. Visual analytics tools that employ multiple linked views for such event data in geospatial and temporal context have recently been presented by Tomaszewski and MacEachren [32] with respect to crisis management and Andrienko et al. [3] for movement data.

5 Visual Analysis concept

In this section, we introduce our visual analysis concept for the validation of geoscientific models. We argue that the complexity of the visualization requirements described in Section 3.3 can best be met by multiple visualization components that focus on specific requirements. The components are integrated in one visual interface and interactively linked. A concise overview of our visual analysis concept is given in Fig. 5. An exemplary realization in the application context of modeling glacial isostatic adjustments is described in the following Section 6. We separate concept and realization because the choice of appropriate visual representations depends on the data characteristics in the individual application context. This includes, for instance, dimensionality of the parameter space or spatio-temporal coverage, preciseness, and volume of observation data.

5.1 Visualization components

To address visualization requirements VR 1 to VR 4 identified in Section 3.3, we define one visualization component for each requirement. Each component provides suitable visual representations that operate on the corresponding data portions.

VC1 Exploration of goodness of fit in model parameter space.
Following the analytical goal to understand the relation between goodness of fit and
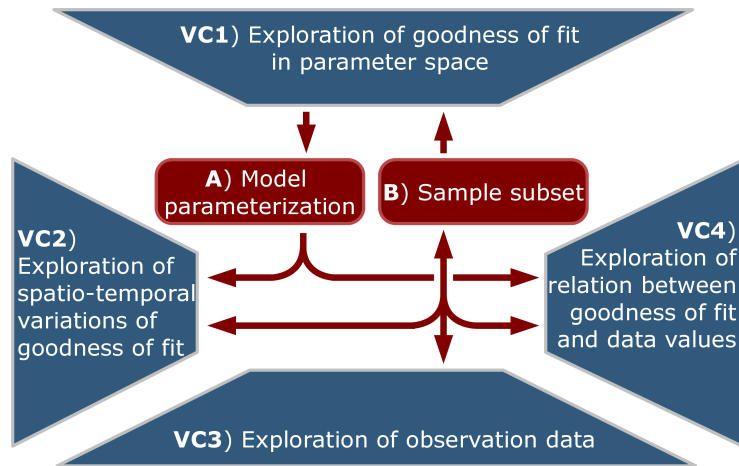
Figure 5: Visual analysis concept for validation of geoscientific simulation models. Four visualization components VC1 to VC4 (blue), addressing the visualization requirements VR1 to VR4, are integrated in one visual interface. The exploration of their combined effect, as demanded by VR5, is supported by a twofold interaction concept (in red) that supports selection in model parameter space and within the set of data samples.

model parameterization, we define one component that communicates variations within the model parameter space by high-level characterizations of the goodness of fit.

VC2 Exploration of spatio-temporal variations of goodness of fit.
The component supports the detailed assessment of variations of the goodness of fit in space and time; the exploration is restraint to a single model parameterization. The visual representation provides appropriate abstraction levels to explore variations in space and time.

VC3 Exploration of observation data.
Resolution and preciseness of observation data are addressed by this visualization component.

VC4 Exploration of relation between goodness of fit and data values.
The ambiguous relation of goodness of fit to underlying observed and modeled values is revealed by a direct visualization of the values along with the corresponding goodness of fit.

5.2 Interactive linking by twofold selection

To support the combined exploration of the manifold dependencies of the goodness of fit, as required by VR 5, the components are interactively linked. This linking is based on a twofold selection concept, which manages the flow of information among the components.

A One model parameterization is analyzed in detail.
Exploration of the goodness of fit in model parameter space in component VC1 relies on high-level descriptions of the goodness of fit. A detailed exploration of the goodness of fit for all model parameterizations simultaneously is not suitable, due to the complex information that needs to be understood. For these detailed analyses, a consecutive exploration of model parameterizations is more feasible. To this end, the selection of a single parameterization in visualization component VC1 is forwarded to components VC2 to explore

9

spatio-temporal variations of the goodness of fit and VC4 to evaluate relation between data values and goodness of fit values.

B A subset of samples is analyzed in model parameter space.
Depending on current analysis objectives, the dependency of the goodness of fit on the model parameterization needs to be analyzed for a specific subset of samples. The subset can be based on a spatial region or a certain time span. Also, other data characteristics may be decisive, such as precision, value ranges or additional geological information given for observation data. We provide a flexible selection mechanism that allows the composition of the sample subset based on all these criteria. The sample subset is mutually determined in components VC2, VC3, and VC4 and forwarded to VC1 to evaluate the average goodness of fit for the selected sample subset in model parameter space.

The dual interactive drill-down within the model parameter space and within the set of data samples provides a highly flexible exploration of the manifold dependencies of the goodness of fit.

6 Views

In the following, we introduce a realization of our visualization concept for the application of validating GIA models.

6.1 Visualization Components

6.1.1 Exploration of Goodness of Fit in Model Parameter Space

A suitable visualization of the goodness of fit in model parameter space mainly depends on space dimensionality and the number of model parameterizations. Further, the sampling of the parameter space (regular vs. random sampling) plays a role.

Exploring the parameter space is closely related to the problem of visualizing multivariate data. A standard approach in this regard is a scatterplot matrix. In our application, the regular sampling of the parameter space leads to overplotting of model parameterizations in a scatterplot matrix, which hinders a clear identification of individual model parameterizations. HyperSlice [37] overcomes this limitation by visualizing orthogonal 2-D slices through the n-dimensional parameter space. The 2-D slices are defined by an n-d focal point, which is interactively adapted to navigate through parameter space. In recent related work [26, 7], this idea was the basis for powerful visualization methods to explore model parameter spaces.

In our application, the low number of parameters (3) as well as the low number of values of each parameter (≤ 5) allow us to provide complete representations of the parameter space, as described in the following. A first overview on the model parameter space is given by a scatterplot matrix. Using transparency, overlapping parameterizations become apparent (view A in Fig. 6). The goodness of fit can be explored upon interactive selection of a scatterplot from the matrix, which opens 2-D slices for each value of the remaining parameter. Color in 2-D slices indicates the high-level abstraction of the goodness of fit related to each model parameterization. As a suitable characterization of the goodness of fit in our application, we employ the average goodness of fit of the current sample subset. 2-D slices are accompanied by a direct visualization of the goodness of fit in 3-D parameter space, which is equipped with interactively controlled view transformations (Fig. 6, view C). Graphical primitives in these views are automatically sized according to the sampling density, thus being decreased if more parameterizations need to be shown.
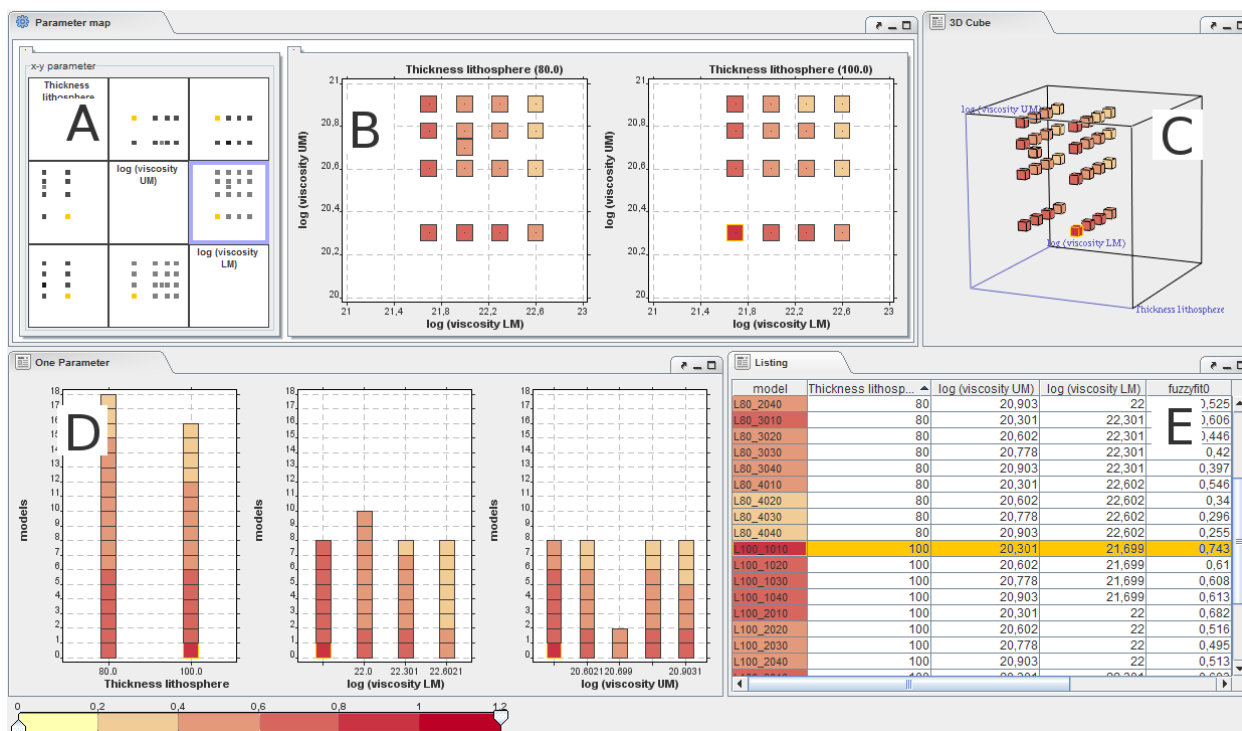
Figure 6: Exploration of goodness of fit in model parameter space. The goodness of fit is mapped to color according to the color scale at the bottom. A - A scatterplot matrix provides an overview of the model parameter space. B - 2-D slices through parameter space support the exploration of dependencies on two parameters. C - The distribution of the goodness of fit is directly visualized in the 3-D parameter space. D - 1-D plots for each parameter convey variances of the goodness of fit across parameter values. E - A table provides a detailed representation of the model parameter space together with the respective fits.

In addition, plots that communicate the dependency of the goodness of fit on single parameters have been considered useful by Earth system modelers. Similar to histograms, these 1-D plots (Fig. 6, view D) map one model parameter on the x axis, and model parameterizations are ordered by goodness of fit along the y axis, which nicely conveys how goodness of fit varies within one parameter. The low number of parameters allows for one plot each. Finally, a table provides a detailed view on all model parameterizations and the associated goodness of fit (conveyed by color of the label column, see Fig. 6, view E), which is valuable for the analyst as it provides detailed representation of the model parameter space together with the respective fits. By combining these views, tailored views for various analysis goals are available.

To fulfill the component's role in our twofold selection concept, brushing of individual model parameterizations for detailed inspection in other visualization components is supported in 2-D slices, 1-D plots, and detail table. The selected parameterization is visually highlighted across all these views, which are further visually linked by a consistent color scale for the goodness of fit.

2-D slices and 3-D view specifically account for the dimensionality of the parameter space in the application example. Nevertheless, our visualization is adaptable to higher numbers of parameters and a more densely sampled parameter space by including a HyperSlice visualization (instead of a scatterplot matrix), which requires the definition of an n-dimensional focal point and its interactive manipulation. Our remaining views, such as 1-D plots and detail table, readily adapt to more parameters (if we speak of about a dozen parameters).

6.1.2 Exploration of Spatio-Temporal Variations of Goodness of Fit

The component visualizes spatio-temporal variations of the goodness of fit for one model parameterization. In general, geospatial distributions can be shown as either globes or maps. We include both alternatives, as they reveal their strengths for different objectives: Maps are suited for an instantaneous global overview of the spatial distribution, while a globe avoids spatial distortions, which are inevitable in map projections. The samples' spatial locations on the globe (or map) are visualized as dots colored according to the sample's goodness of fit, as shown in Fig 7.

The visualization of individual samples is not sufficient, as these do neither convey trends nor support the comparison of spatial regions or time spans. To convey variations of the goodness of fit on appropriate spatio-temporal abstraction levels, we group samples into geologically meaningful regions of different spatial extent. These abstraction levels are visualized by bounding boxes on the Earth surface. For each spatial group, the goodness of fit is characterized by the average of all sample fits that are currently included in the selection, visualized by the color of the bounding box.

Accounting for variations in space *and* time, temporal variations within regions are communicated by additional symbols, which are associated with each region. The symbol employs the visual metaphor of population pyramids: Time is shown vertically elapsing from top to bottom. For each temporal interval, the number of samples is shown. The number of samples with fit $f = 1$ is shown by bar length to the right, the number of samples with $f < 1$ by bar length to the left. This supports a fast identification of the agreement between observed and modeled values. Color is used to differentiate fits values. Like a rotated stacked histogram, the color distribution within the left bar corresponds to the distribution of fuzzy fits.
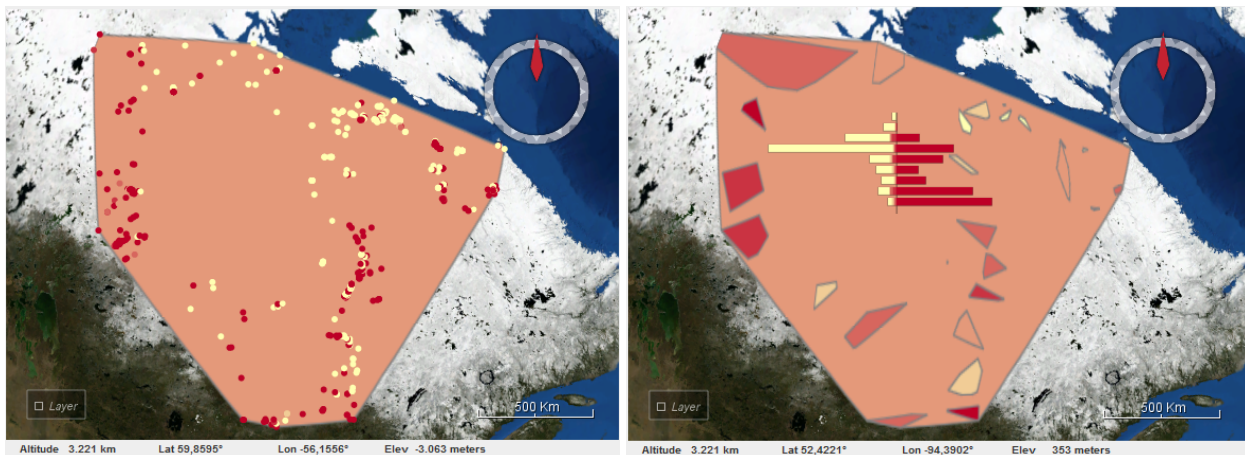

Figure 7: Exploration of spatio-temporal variations of goodness of fit for one model parameterization. Left: Individual samples are shown in the region of interest (here: Hudson Bay). The color corresponds to the goodness of fit of each sample, the color of the bounding box to the average goodness of fit of the region. Right: Average goodness of fit for geologically meaningful spatial regions are shown. The symbol communicates temporal variations in goodness of fit within the region.

The visualization component extends our previous visualization of spatio-temporal variations presented in [11] as follows: Dots (indicating spatial distribution) and symbols (indicating temporal distributions) now communicate goodness of fit values rather than an observation-data related attribute value. In conjunction with the likewise novel mapping of a region's average fit to the color of its bounding box, our visualization now provides additional detail about the variations of the goodness of fit in space and time.

The analyst can adapt the visual representation according to his current objectives. This includes zoom and rotation of the globe according to the spatial extent of interest, definition of the temporal range and its subdivision, and the interactively controlled combination of simultaneously shown spatial and temporal abstraction levels. Further, the selected sample subset can be adjusted by adding or removing individual samples as well as whole regions with a mouse click. Additional information, such as labels and data values, are available as tooltips for samples and regions.

6.1.3 Exploration of Observation Data

The component supports the assessment of spatio-temporal resolution and preciseness of observation data. Impreciseness is expressed by age range and sea level range, which can be manually constrained further based on additional geological information, such as the type of fossil. Such information is provided by additional attributes. Our example observation data provides about 30 attributes for the samples.

A synoptic view on all these facets is gained by a combination of views that emphasize different aspects of the data (see Fig. 8):
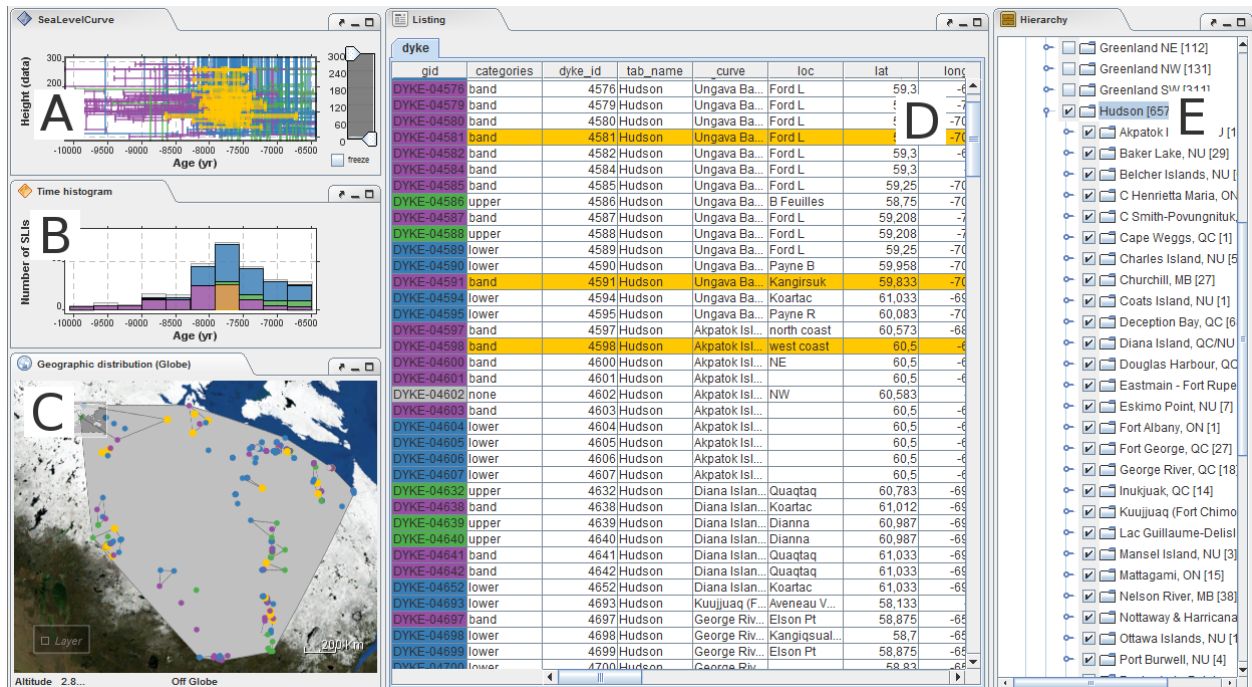

Figure 8: Multiple linked views concept to analyze unevenly distributed and imprecise observation data. Each view offers a different perspective on the data. A - Data view: A time value plot shows observed value ranges along with imprecise age ranges of samples. B - Temporal View: A temporal histogram communicates temporal distribution of samples. C - Geospatial View: A map (or globe) shows the uneven coverage of data samples in geospatial context. D - Attribute View: Additional attributes provided in a table convey important geological information. E - Hierarchy View: An explicit visualization of the spatial hierarchy supports fast overview and selection of observations. Color used throughout views indicates the categorization of observation samples by their type of value range.

A *Data View:* The impreciseness of observation samples in age and sea level range is shown in a time value plot – a standard representation in the assessment of sea level indicators.

Our visualization makes use of established graphical primitives to convey the dual impreciseness. Accounting for the relevance of the type of value range (upper bounded, lower bounded or double bounded), they are discerned by color for instantaneous visual categorization.

B *Temporal View:* The temporal distribution of samples is given in a histogram, which groups them into time intervals. Categories of observation samples are conveyed by color.

C *Geospatial View:* The spatial distribution of samples is provided by colored dots in a globe. Additionally, bounding boxes group the observation samples into geologically meaningful regions.

D *Attribute View:* The imprecise sea level range of an observation sample can be constraint based on additional attributes, comprising scalar values, categorical data or textual descriptions. A table containing one row for each sample and one row for each attribute provides these details. The categorization of samples is visually encoded in the color of the label column.

E *Hierarchy View:* Observations are organized in a spatial hierarchy. Its explicit representation in a tree hierarchy provides a fast and compact overview over available data and facilitates the selection of data of interest.

These views are enriched by interactive methods to retrieve details on demand, including tooltips, zooming and panning. Views can also be filtered to show only selected data items. Further, brushing mechanisms realize one side of our twofold interaction concept, the selection of a relevant sample subset. Beyond individual samples, sets of samples can be brushed at once, which are grouped by, i.e., temporal range and height range (brushing a rectangular region in view A), by temporal range and sample type (brushing a bar in view B), by region (brushing a bounding box in view C or selecting a region in view E), or by any other attribute (brushing an interval of rows in view D upon interactively triggered sorting by any attribute of interest). The current selection is shared among all views and highlighted by color (orange in Fig 8).

Giving access to the numerous facets of the observation data, these views provide the basis for a free definition of relevant sample subsets.

6.1.4 Exploration of Relation between Goodness of Fit and Data Values

The component addresses observed and modeled values as well as related goodness of fit values with respect to the current sample subset and the current model parameterization.

The visual comparison of modeled and observed values is based on the time value plot introduced in Section 6.1.3, as it provides tailored graphical primitives to handle impreciseness in observation data. The time value plot is accompanied with additional visual representations for modeled values and goodness of fit (Fig. 9). Modeled values are communicated by graphical primitives (small squares). The visual connection between modeled and observed value of a sample is maintained by their mapping onto the same position on the time axis. Deviations in sea level height between modeled and observed value are shown on the vertical axis. The close relation between goodness of fit and modeled value is conveyed by the color of the graphical primitive.

Interactive features like tooltips, highlighting of related observed and model values and zooming in by an adaptation of the visible time range and height range are provided to discern samples that visually overlap due to similar ages. Nevertheless, direct visual comparison is generally limited to a relatively small set of samples. For larger sample sets, the time value plot view rather conveys value ranges than values of individual samples.
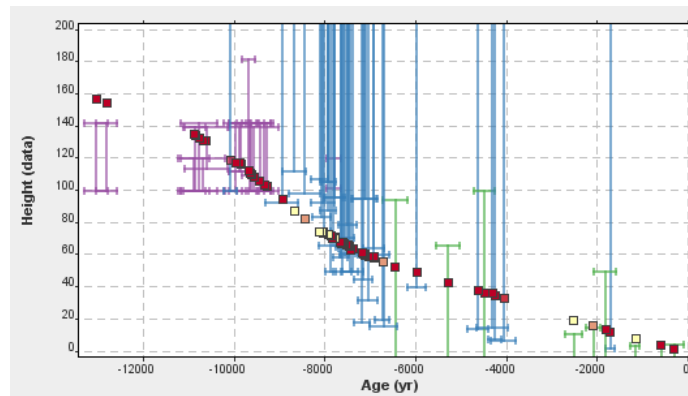
Figure 9: Visual exploration of relation between goodness of fit and underlying values in a time value plot. Observed value ranges are visually encoded by vertical lines. The horizontal line at the bounds of the value range indicates the age range. Modeled values are depicted by small squares, colored according to the goodness of fit that resulted from the analytical comparison of modeled and observed value.

6.2 Linking and Interaction

The visualization components are arranged and linked within one visual interface. The previous discussion of visualization components proposed similar visual encodings for different components: spatial-temporal representations are used for both the exploration of observation data and of the goodness of fit, time value plots are applied to evaluate the preciseness of observation data and to explore relations between data values and goodness of fit. To keep the visual interface as compact as possible, we include only one spatial visualization and one time value plot, equipped with interactive facilities to adapt the visualized facets of the data. The resulting combination of views is illustrated in Fig. 10. Views can be arranged flexibly on demand.

Our twofold interaction concept propagates selected data among the visualization components as follows (see also Fig. 10). In model parameter space, the user brushes a model parameterization (VC1). Spatio-temporal view (VC2) and comparison view (VC4) are instantly adapted to provide details about the goodness of fit of this model parameterization beyond the average value. This is one side of our twofold interaction concept. The second part is the selection of a sample subset. In data space, the user brushes a sample subset in spatio-temporal view (VC2), in views on observation data (VC3), and in comparison view (VC4). Each adaption of the sample subset leads to a recalculation of average goodness of fit values in model parameter space (VC1). Thus, the ability to iteratively adapt selected model parameterization and selected sample subset forms a strong interaction mechanism to explore the manifold dependencies of the goodness of fit.

Visual highlighting of selections across visualization components provides guidance on selected data. The selected model parameterization is highlighted in model parameter space (VC1). The selected sample subset is highlighted in spatio-temporal view (VC2), all views on observation data (VC3) and comparison view (VC4). For further visual linking, similar visual representations of the same data are shared among all views: Goodness of fit values (of individual samples and sample subsets) as well as main categories of observation data are discerned by color. This leads to two distinct global color schemes: a sequential color scale for the goodness of fit (in range $[0, 1]$) and a categorical color scale for the three main categories of observation data. A consistent visual encoding throughout all views is also used for sizes of subsets (as used in the histogram or by the symbols), which are generally encoded as bar lengths.
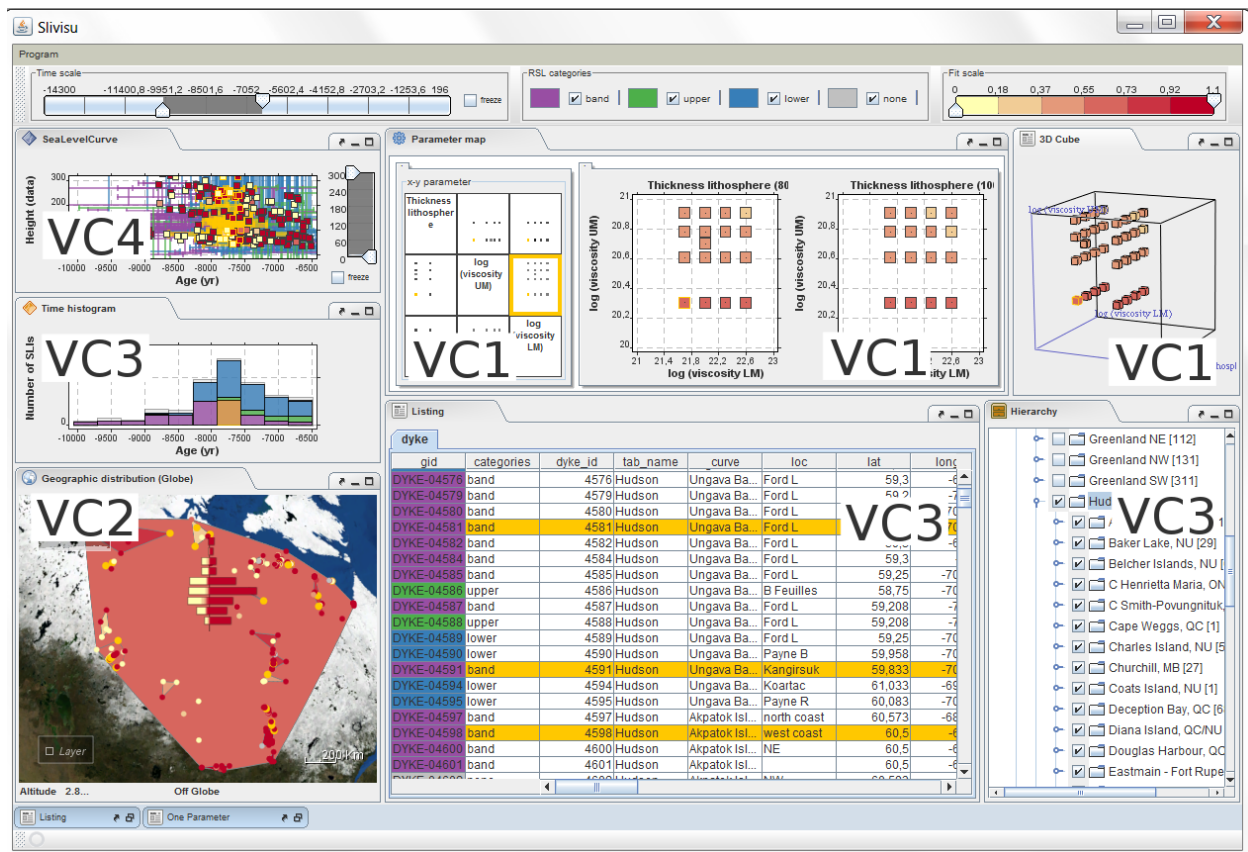
Figure 10: Visual interface that integrates all views and links them with our twofold interaction concept. Views are visually linked by two global color maps: A categorical color scale for categorization of observation data by type of value range (top middle) and a sequential color scale for goodness of fit (top right).

7 Results

Our approach has been readily applied as a method of validation in modeling glacial isostatic adjustments (GIA). In this section, we demonstrate its use and present feedback from our collaborating Earth system modelers.

7.1 Use case

The goal of the visual analysis conducted in this use case is to understand how the goodness of fit of the GIA model depends on the Earth characteristics, which are described by the three model parameters. The analysis serves two purposes: First, it is to determine if the model parameters' effect on the goodness of fit agrees with expected behavior. The second purpose is to narrow the range of suitable model parameterizations by the identification of the best goodness of fit.

Definition of spatial region of interest. Not all regions on Earth are suited for this analysis. Often, variations in sea level height are induced by the interplay of numerous processes. To validate the GIA model, the analysis focuses on a region where the process dominates variations in sea level: Hudson Bay in North America. During the last glaciation phase, the area was covered by ice masses with an assumed thickness of three to four kilometers. As a consequence,
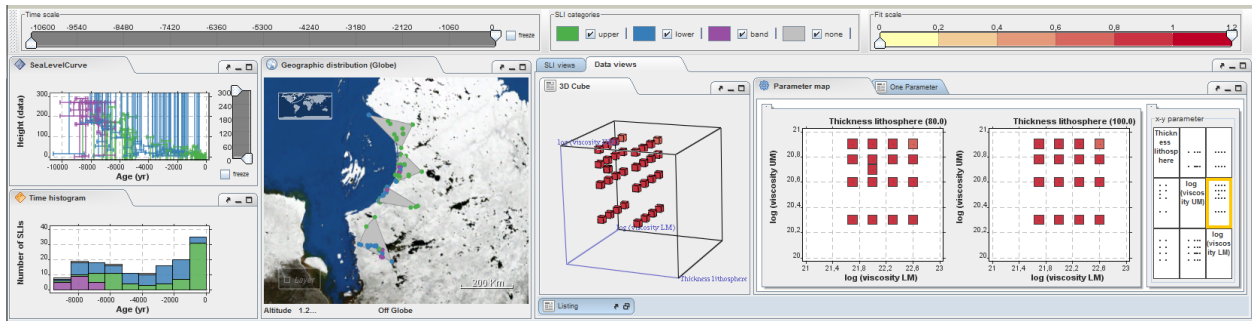
(a) Exploration of observation data in central area of a former ice shield. For initial sample set, a low variability in goodness of fit in model parameter space is observed.
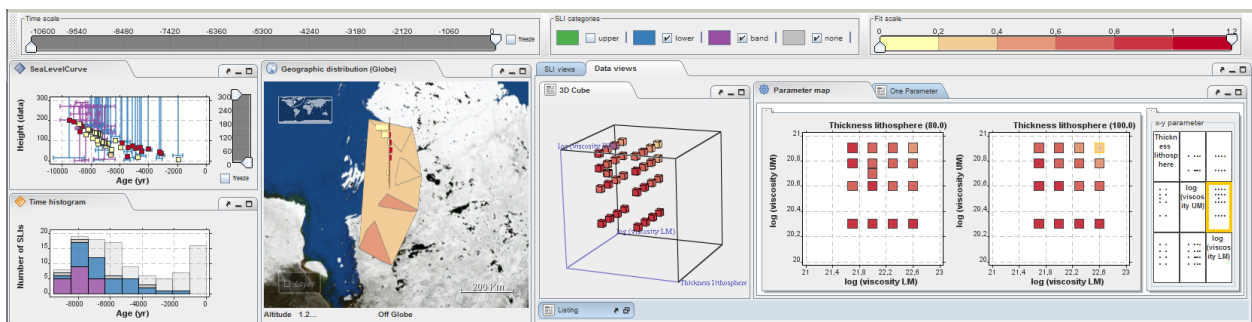
(b) The initial sample set is reduced to relevant samples. Goodness of fit now shows noticeable variations in model parameter space: With increasing values for *UM* and *LM*, the goodness of fit decreases, due to an underestimation of sea level height. This can be seen from the detailed inspection of the model parameterization with worst fit (views on the right). *L* (thickness of lithosphere) has no visible impact on the goodness of fit.
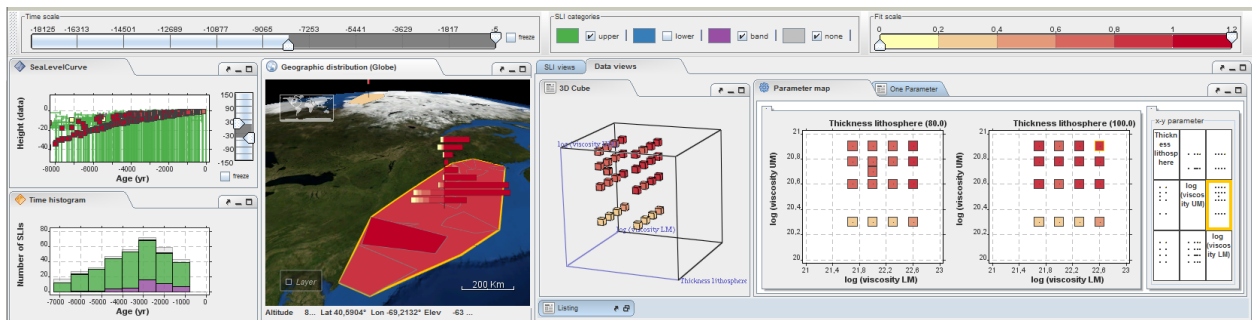
(c) Exploration of goodness of fit in model parameter space in boundary region of former ice shield, for an interactively defined relevant sample set. A strong dependency of goodness of fit on *UM* is apparent, while *LM* plays little role. Slight impact of *L* (thickness of lithosphere) can be seen. The best fitting parameterization in this region showed worst fitting for previously regarded region (compare to (b)).

Figure 11: Use case for validation of GIA modeling: Exploration of goodness of fit in model parameter space for central and boundary regions of a former ice shield.

17

the region was strongly deformed. The basin was lowered by approximately one kilometer. When ice masses melted in the time span from of 20,000 to 8,000 years before present, a general uplift of the region was caused that continues until present.

Initial inspection of observation data. The analysis starts with an evaluation of the observation data in Hudson Bay (Fig. 11a, left side). From the globe, an uneven distribution of the data along the east coast of Hudson Bay can be seen. In addition, the temporal histogram (bottom left) reveals great differences in the presence of samples over time. Most samples are given for the time span 1,000 years to present. However, only more imprecise half-bounded values are available in this time span. "Valuable" samples with double bounded ranges (in purple) are present in significant number in the time range 10,000 to 6,000 years before present.

Initial inspection of model ensemble. After the initial inspection of observation data, the model ensemble is regarded that is validated against observation data. Our visualization shows the distribution of model parameterizations in parameter space (Fig. 11a, right side). The two parameters that characterize the upper and lower mantle viscosities of the Earth (UM and LM) are sampled by four values each, the value for the lithosphere thickness L by two values. This gives 4*4*2 = 32 model parameterizations plus one additional parameterization that represents a standard model. The goodness of fit in model parameter space is expected to noticeably vary for UM and LM, while L should show little influence. To check whether the model agrees with the expectations, the average goodness of fit of each model parameterization is initially derived from all samples along Hudson Bay. Almost all model parameterizations have a similar goodness of fit, with values between 0.6 and 0.8.

Refinement of sample subset. Inspecting the goodness of fit in parameter space for individual categories of observation data reveals that trends in parameter space are apparent for double bounded and lower bounded ranges, but not for upper bounded samples – these samples thus generally agree with all considered model parameterizations. A detailed inspections of individual parameterizations, selected across parameter space, confirms the assumption that many upper limited ranges are too imprecise. They cannot be used to differentiate the goodness of fit in model parameter space. In the following, such irrelevant samples are discarded by a combination of automatic identification of samples with high goodness of fit for all model parameterizations and their manual filtering based on geological knowledge. For the most part, the resulting sample set comprises lower bounded and double bounded values (Fig. 11b, left side). Also, only few relevant samples in the range from 6,500 years to present remain.

Assessment of the goodness of fit in model parameter space and identification of best goodness of fit. For the resulting sample subset, it is apparent that the goodness of fit depends on M_L and M_U, but not on L (Fig. 11b, right side). This agrees with the above stated expectations. Further, the goodness of fit generally increases for lower values of both M_U and M_L. An inspection of individual model parameterizations shows why: A comparison of modeled and observed values for the best fitting parameterization and for the worst fitting parameterization shows that the latter generally underestimates sea level heights. Consequently, model parameterizations that best agree with observed sea level ranges in this region have low values for M_L and M_U.

Exploration of spatio-temporal variations in goodness of fit. These findings are compared to other regions, which are differently exposed to the studied geoprocess. As an example, we look at a boundary region of the ice shield that covered Hudson Bay (Fig. 11c). Here, the

influence of L should be more apparent, while M_U should have less impact. After interactively deriving a relevant sample subset in the region, the goodness of fit shows indeed a noticeable influence of the lithosphere thickness L (Fig. 11c, right side). Also, as expected, M_L influences the goodness of fit, while the effect of M_U is negligible. Regarding suitable values for model parameters, a different parameterization is preferable in this region compared to the previously inspected Hudson Bay. This indicates that constant global parameter values for the Earth characterization do not sufficiently reproduce the impact of glacial isostatic adjustments.

7.2 Discussion and Feedback

Before our visual analysis concept was available for model validation, the variability of the sample quality, especially due to their indicative meaning or type, hindered a rigorous evaluation by an aggregated fit. As an alternative, a bottom-up strategy often used in validation of GIA modeling is based on the analysis of modeled values for locations represented by sea-level curves [19]. These are small sets of samples, which commonly comprise a few dozen samples grouped by their similar spatial location and comparable exposure to sea level variations. Curve by curve, observation data is evaluated and presented in separated static plots for space and time. The goodness of fit in parameter space is derived for individual curves to assess how they respond. This tedious analysis is performed for only a few pre-selected curves, thereby discarding many relevant samples.

The close linkage between observation data, model data, and goodness of fit as provided by our visual analysis concept allows a better understanding and improvement of the evaluation strategy. As one result, the construction of relevant sample sets is largely simplified by our concept. With our synthesized view on the numerous aspects in the data, all potentially relevant samples in an area of interest – which are not anymore constraint to individual sea-level curves – can be interactively inspected and filtered based on the analyst's assessment. From these sample sets, the determination of meaningful goodness of fit values becomes possible.

Meaningful goodness of fit values are the basis to target various analysis objectives, which could not be addressed adequately before. This includes the evaluation of the dependency on model parameters, whose exact value is unknown. In this regard, our tool provides the ability to narrow the range of suitable model parameters based on the best goodness of fit. Suitable parameters can further be identified with respect to different geospatial regions that show deviating response to glacial isostatic adjustments. In a broader sense, examining the variabilities in parameter space, in geographical space, and in time helps to answer the question if the model assumptions are sufficient to reproduce the studied process, or if a revision or improvement of the modeling is necessary.

In addition to these validation related objectives, the visual data representation allows a better communication of model results and sea level indicator data between research groups.

8 Conclusion

Our main contribution is a specific visual analysis concept for validation tasks in Earth system modeling, where a combination of challenges hinders a rigorous analytical approach. These challenges include unknown model parameter values, spatio-temporal variations, sparse and imprecise observation data with uneven spatio-temporal distribution, and ambiguous relations between computationally derived comparison results to underlying modeled and observed values. In close cooperation with Earth system modelers, we identified visualization requirements to target these challenges. The requirements are fulfilled by tailored visualization components that are linked by a twofold selection.

The realization of the visual analysis concept for the application of modeling glacial isostatic adjustments has already become a valuable member of the analysts' toolbox. For this exemplary application, our solution substantially broadens the analysis capabilities in model validation. Before, validation was only performed for small, predefined sample sets, whose characteristics were known. Now, for the first time, a comprehensive validation of the model is supported, as all dependencies of the automatically derived goodness of fit can be interactively explored and related to each other.

Our visualization is adoptable to other modeling applications where identical challenges occur and limited availability of observation data makes purely analytical approaches not feasible. Other examples of sparsely sampled observation data are geophysical point measurements – like tide gauges, geothermal heat flow measurements, or seismological measurements – , biological markers, or agricultural data. Such terrestrial observations remain the dominant data source for investigation of processes and validation of simulation models in numerous applications in disciplines such as geoscience, climatology, and regional policy. Beyond, we will investigate how our visualization concepts can be adopted to models and observation data with other dimensionality and scale. Regarding model parameterizations, higher-dimensional parameter spaces, larger model ensembles, and stochastically sampled parameter space will be considered.

Feedback has shown that the tool is well suited to provide insight into the relation between observed and modeled data values on the one hand and resulting goodness of fit values on the other hand. With respect to this capability, we will further investigate how our visual analysis concept can be employed for the development and evaluation of suitable goodness of fit measures.

Acknowledgments

References

[1] J. Ahrens, K. Heitmann, M. Petersen, J. Woodring, S. Williams, P. Fasel, C. Ahrens, C.-H. Hsu, and B. Geveci. Verifying Scientific Simulations via Comparative and Quantitative Visualization. *IEEE Computer Graphics and Applications*, 30:16–28, 2010.

[2] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of Time-Oriented Data*. Springer, 2011.

[3] G. Andrienko, N. Adrienko, M. Mladenov, M. Mock, and C. Politz. Identifying Place Histories from Activity Traces with an Eye to Parameter Impact. *IEEE Transactions on Visualization and Computer Graphics*, 18(5):675 –688, 2012.

[4] G. Andrienko, N. Andrienko, U. Demsar, D. Dransch, J. Dykes, S. I. Fabrikant, M. Jern, M.-J. Kraak, H. Schumann, and C. Tominski. Space, time and visual analytics. *International Journal of Geographical Information Science*, 24(10):1577–1600, 2010.

[5] N. Andrienko and G. Andrienko. *Exploratory Analysis of Spatial and Temporal Data - A Systematic Approach*. Springer, 2006.

[6] O. Balci. Verification, validation, and accreditation. In *WSC '98: Proceedings of the 30th conference on Winter simulation*, pages 41–4, Los Alamitos, CA, USA, 1998. IEEE Computer Society Press.

[7] W. Berger, H. Piringer, P. Filzmoser, and E. Gröller. Uncertainty-aware exploration of continuous parameter spaces using multivariate prediction. *Published in Computer Graphics Forum*, 30(3):pp. 911 – 920, 2011. Best Paper Award.

[8] F. Boschetti, C. Wijns, and L. Moresi. Effective exploration and visualization of geological parameter space. *Geochemistry Geophysics Geosystems*, 4:1086–1096, 2003.

[9] S. Bruckner and T. Möller. Result-Driven Exploration of Simulation Parameter Spaces for Visual Effects Design. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1468–1476, 2010.

[10] W. Chung, H. Chen, L. G. Chaboya, C. D. O'Toole, and H. Atabakhsh. Evaluating event visualization: a usability study of COPLINK spatio-temporal visualizer. *International Journal of Human-Computer Studies*, 62(1):127–157, 2005.

[11] D. Dransch, P. Köthur, S. Schulte, V. Klemann, and H. Dobslaw. Assessing the quality of geoscientific simulation models with visual analytics methods - a design study. *International Journal of Geographical Information Science*, 24(10):1459–1479, 2010.

[12] J. Dykes, A. MacEachren, and M.-J. Kraak. *Exploring Geovisualization*. International Cartographic Association, 2005.

[13] N. Franken. Visual exploration of algorithm parameter space. In *Evolutionary Computation, 2009. CEC '09. IEEE Congress on*, pages 389 –398, 2009.

[14] J. Kehrer, F. Ladstädter, P. Muigg, H. Doleisch, A. Steiner, and H. Hauser. Hypothesis Generation in Climate Research with Interactive Visual Data Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1579–1586, 2008.

[15] J. Kehrer, P. Muigg, H. Doleisch, and H. Hauser. Interactive Visual Analysis of Heterogeneous Scientific Data across an Interface. *IEEE Transactions on Visualization and Computer Graphics*, 17(7):9340–0946, 2011.

[16] V. Klemann and D. Wolf. Using Fuzzy Logic for the Analysis of Sea-level Indicators with Respect to Glacial-isostatic Adjustment: An Application to the Richmond-Gulf Region, Hudson Bay. *Pure and Applied Geophysics*, 164:683–696, 2007.

[17] A. Köhler, M. Ohrnberger, and F. Scherbaum. Unsupervised feature selection and general pattern discovery using Self-Organizing Maps for gaining insights into the nature of seismic wavefields. *Computers and Geosciences*, 35(9):1757–1767, 2009.

[18] F. Ladstädter, A. K. Steiner, B. C. Lackner, B. Pirscher, G. Kirchengast, J. Kehrer, H. Hauser, P. Muigg, and H. Doleisch. Exploration of Climate Data Using Interactive Visualization. *Journal of Atmospheric and Oceanic Technology*, 27(4):667–679, 2010.

[19] K. Lambeck, C. Smither, and P. Johnston. Sea-level change, glacial rebound and mantle viscosity for northern Europe. *Geophysical Journal International*, 134:102–144, 1998.

[20] D. R. Legates and G. J. McCabe Jr. Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35:233–241, 1999.

[21] D. R. Lehmann. Validity and Goodness of Fit in Data Analysis. *Advances in Consumer Research*, 2:741–750, 1975.

[22] A. M. MacEachren. *How Maps Work: Representation, Visualization, and Design*. The Guilford Press, 1995.

[23] K. Matkovic, D. Gracanin, M. Jelovic, A. Ammer, A. Lez, and H. Hauser. Interactive Visual Analysis of Multiple Simulation Runs Using the Simulation Model View: Understanding and Tuning of an Electronic Unit Injector. *IEEE Transactions on Visualization and Computer Graphics*, 16:1449–1457, 2010.

[24] K. Matkovic, D. Gracanin, M. Jelovic, and H. Hauser. Interactive Visual Steering - Rapid Visual Prototyping of a Common Rail Injection System. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1699 –1706, 2008.

[25] T. Nocke, M. Flechsig, and U. Bohm. Visual exploration and evaluation of climate-related simulation data. In *Proc. of the 2007 Winter Simulation Conference*, pages 703–711, 2007.

[26] H. Piringer, W. Berger, and J. Krasser. HyperMoVal: Interactive Visual Validation of Regression Models for Real-Time Simulation. *Computer Graphics Forum*, 29(3):pp. 983–992, 2010.

[27] K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, and C. R. Johhson. Ensemble-vis: A framework for the statistical visualization of ensemble data. In *IEEE Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes.*, pages 233–240, 2009.

[28] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead. Noodles: A Tool for Visualization of Numerical Weather Model Ensemble Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 16:1421–1430, 2010.

[29] H.-J. Schulz, A. M. Uhrmacher, and H. Schumann. Visual analytics for stochastic simulation in cell biology. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, i-KNOW '11, pages 48:1–48:8, New York, NY, USA, 2011. ACM.

[30] H. Steffen and P. Wu. Glacial isostatic adjustment in Fennoscandia – A review of data and modeling. *Journal of Geodynamics*, 52:169–204, 2011.

[31] B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber. The Gravity Recovery and Climate Experiment: Mission Overview and Eearly Results. *Geophysical Research Letters*, 31:L09607, 2004.

[32] B. Tomaszewski and A. MacEachren. Geo-historical context support for information foraging and sensemaking: Conceptual model, implementation, and assessment. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 139 –146, 2010.

[33] C. Tominski, J. F. Donges, and T. Nocke. Information Visualization in Climate Research. In *Proceedings of the International Conference Information Visualisation (IV)*, pages 298–305. IEEE Computer Society, 2011.

[34] T. Torsney-Weir, A. Saad, T. Möller, H. Hege, B. Weber, J. Verbavatz, and S. Bergner. Tuner: Principled Parameter Finding for Image Segmentation Algorithms Using Visual Response Surface Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1892 –1901, 2011.

[35] T. Trucano, L. Swiler, T. Igusa, W. Oberkampf, and M. Pilch. Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering & System Safety*, 91:1331–1357, 2006.

[36] A. Unger and H. Schumann. Visual support for the understanding of simulation processes. In *PACIFICVIS '09: Proceedings of the 2009 IEEE Pacific Visualization Symposium*, pages 57–64, Washington, DC, USA, 2009. IEEE Computer Society.

[37] J. J. van Wijk and R. van Liere. Hyperslice: visualization of scalar functions of many variables. In *Proceedings of the 4th conference on Visualization '93*, VIS '93, pages 119–125, Washington, DC, USA, 1993. IEEE Computer Society.

[38] J. Waser, R. Fuchs, H. Ribicic, B. Schindler, G. Bloschl, and E. Gröller. World Lines. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1458–1467, 2010.

[39] J. Waser, H. Ribicic, R. Fuchs, C. Hirsch, B. Schindler, G. Bloschl, and E. Gröller. Nodes on Ropes: A Comprehensive Data and Control Flow for Steering Ensemble Simulations. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1872–1881, 2011.

[40] A. B. Watts. *Isostasy and Flexure of the Lithosphere*. Cambridge University Press, Cambridge, 2001. V.

[41] X. Yuan, H. Xiao, H. Guo, P. Guo, W. Kendall, J. Huang, and Y. Zhang. Scalable Multivariate Analytics of Seismic and Satellite-based Observational Data. *IEEE Transactions on Visualization and Computer Graphics*, 16:1413–1420, 2010.