Originally published as:

Köthur, P., Sips, M., Dobslaw, H., Dransch, D. (2014): Visual Analytics for Comparison of Ocean Model Output with Reference Data: Detecting and Analyzing Geophysical Processes Using Clustering Ensembles. - *IEEE Transactions on Visualization and Computer Graphics*, *20*, 12, pp. 1893–1902.

DOI: http://doi.org/10.1109/TVCG.2014.2346751

Visual Analytics for Comparison of Ocean Model Output with Reference Data: Detecting and Analyzing Geophysical Processes using Clustering Ensembles

Patrick Köthur[1,2], Mike Sips[1], Henryk Dobslaw[1] and Doris Dransch[1,2]

[1]GFZ German Research Centre for Geosciences, Potsdam, Germany
E-mail: {koethurp, sips, dobslaw, dransch}@gfz-potsdam.de
[2]Humboldt-Universität zu Berlin, Geography Department, Berlin, Germany

Abstract

Researchers assess the quality of an ocean model by comparing its output to that of a previous model version or to observations. One objective of the comparison is to detect and to analyze differences and similarities between both data sets regarding geophysical processes, such as particular ocean currents. This task involves the analysis of thousands or hundreds of thousands of geographically referenced temporal profiles in the data. To cope with the amount of data, modelers combine aggregation of temporal profiles to single statistical values with visual comparison. Although this strategy is based on experience and a well-grounded body of expert knowledge, our discussions with domain experts have shown that it has two limitations: (1) using a single statistical measure results in a rather limited scope of the comparison and in significant loss of information, and (2) the decisions modelers have to make in the process may lead to important aspects being overlooked.

In this article, we propose a visual analytics approach that broadens the scope of the analysis, reduces subjectivity, and facilitates comparison of the two data sets. It comprises three steps: First, it allows modelers to consider many aspects of the temporal behavior of geophysical processes by conducting multiple clusterings of the temporal profiles in each data set. Modelers can choose different features describing the temporal behavior of relevant processes, clustering algorithms, and parameterizations. Second, our approach consolidates the clusterings of one data set into a single clustering via a clustering ensembles approach. The consolidated clustering presents an overview of the geospatial distribution of temporal behavior in a data set. Third, a visual interface allows modelers to compare the two consolidated clusterings. It enables them to detect clusters of temporal profiles that represent geophysical processes and to analyze differences and similarities between two data sets.

This work is the result of a close collaboration with ocean modelers. They employed our concept to find aspects of improvement in a new version of the Ocean Model for Circulation and Tides (OMCT).

Index terms—Ocean modeling, model assessment, geospatial time series, cluster ensembles, visual comparison, visual analytics

1 Introduction

Geoscientific simulation models, in particular ocean and climate system models, consider complex interactions between processes in the Earth system [38]. For example, the salinity of the oceans affects the circulation within oceans, which in turn impacts the energy exchange between oceans and atmosphere. The latter influences air temperature, which may affect continental and sea ice. To come full circle, melting ice leads to freshwater influx into the oceans which influences their salinity.

Modeling such interactions serves three important purposes. First, simulation models provide data about real-world phenomena in geographic regions that are not or only partially covered by monitoring devices. Second, they enable scientists to identify and study causalities between geophysical processes to gain insight into the physics of the Earth system. Last, when researchers have acquired a sound understanding of the mechanisms within the Earth system and their interactions, they can produce reliable predictions; a prominent example being future greenhouse gas concentrations and their effect on, e.g., sea level rise, availability of fresh water, or natural hazards.

In this paper, we specifically focus on ocean models. Intense collaboration with ocean modelers at the German Research Center for Geosciences GFZ enabled us to gain an understanding of the challenges involved in model assessment. The development of ocean models is an iterative process in which the assessment of new model versions is a critical part. After each change to the model, researchers compare the model output to reference data to determine whether the new version improves the simulation. The reference data may either stem from a previous version of the same model or from observations. The latter is usually the case when a model is new and there is little knowledge about its behavior and performance against observations. The former applies, e.g., when scientists want to improve specific aspects of a tried-and-tested model and are already familiar with the observational data and the behavior of previous model versions.

For the comparison, modelers need to locate and compare geophysical processes in both data sets. In our context, a geophysical process is a broad concept. It includes, e.g., the El Niño southern oscillation or particular ocean currents such as western or eastern boundary currents. Every process has particular temporal and geospatial characteristics that manifest to a varying degree in the data.

To detect geophysical processes, our partners try to identify and locate temporal profiles in the data that are characteristic for the processes under study. This is a challenging task due to the volume and complexity of the data. Ocean models depict the ocean as a regularly structured three-dimensional grid. The grid points represent geographic coordinates and have temporal profiles associated with them that describe the temporal behavior at these coordinates. For the analysis, modelers frequently focus on the topmost layer, the sea surface. This is appropriate because most mechanisms within the ocean manifest themselves in changes to sea surface heights. But even with the focus on the sea surface, the data comprise thousands or hundreds of thousands of time series.

To address this challenge, scientists employ data aggregation and visualization in a two-step process. First, modelers aggregate the time dimension by computing a single statistical measure for each temporal profile in the data. Scientists plot this measure in a separate geographic map for each data set to visually analyze and compare its geospatial distribution. Second, modelers choose a small number of geographic coordinates (typically not more than 20) for a more detailed analysis of the temporal behavior. Line charts are used to study and compare the temporal profiles associated with the selected coordinates.

Although scientists aggregate and compare the data based on experience and a well-grounded body of expert knowledge, they know that data aggregation results in loss of information, and that the subjectivity involved in this strategy may have them miss important aspects. The chosen statistical measure focusses the analysis on a particular characteristic of temporal behavior; information about other aspects of temporal behavior is lost. In addition, other important details may not be noticed because modelers focus the comparison of temporal profiles on a limited number of geographic coordinates.

In this article, we introduce a visual analytics approach that presents modelers with a more comprehensive view on the temporal behavior in the data. It allows modelers (a) to create multiple spatial clusterings of the temporal profiles in model output and reference data, (b) to consolidate the various clusterings for each data set with an ensemble approach [34], and (c) to interactively explore and compare the two consolidation results.

We chose clustering of temporal profiles because (1) it allows modelers to systematically identify and locate the predominant types of temporal behavior in the data, and (2) because it allows researchers to base the analysis on many different characteristics of temporal behavior. Our concept enables scientists to compute multiple clusterings with varying user-chosen features of temporal behavior, clustering algorithms, and parameterizations. In the following, a *feature* denotes any representation of a temporal profile that captures a particular aspect of temporal behavior and supports definition of a (dis)similarity metric. To unite the different aspects of temporal behavior reflected in various clusterings, a clustering ensemble combines all clusterings of one data set into a single consolidated clustering. A visual interface facilitates comparison of the two consolidation results for model data and reference data. It allows researchers to identify clusters of temporal profiles that represent geophysical processes as well as to explore differences and similarities between the two data sets.

In particular, the contributions of this article are as follows:

- We closely collaborated with ocean modelers and conducted a thorough task and requirement analysis to identify the key challenges in the comparison of ocean model output with reference data.
- We combine cluster ensembles and interactive visual exploration for a novel approach to supporting the assessment of ocean models.
- We demonstrate how our concept enables modelers to conduct a fast comparison of model data with reference data that complements the existing statistical methods.

2 Related work

Although numerous guidelines and techniques exist for the visual analysis of geospatial data [15, 30], time series data [2], and geospatio-temporal data [5, 7], the complexity and volume of geoscientific data still presents significant challenges [29]. Clustering is an established technique for approaching these challenges. Its aim is to divide data into groups of similar objects (for further reading, please refer to [10, 20]). A number of visual analytics works apply clustering to analyze geospatial time series. Andrienko et al. [4] introduce two perspectives to such an analysis: 'space-in-time' and 'time-in-space'. The former analyzes how the geospatial distribution of data values changes over time; the latter studies how the temporal behavior is distributed in geographic space. To address both perspectives, the before mentioned work uses self-organizing maps (SOM) as a clustering and visualization technique and combines it with multiple linked views. In own previous work [26], we consider the 'space-in-time' perspective and combine hierarchical clustering with visual exploration to support detection of dominant spatial states in geoscientific data. Approaches that apply clustering to analyze multiple temporal profiles ('time-in-space') are more numerous. Guo et al. [13] and Andrienko et al. [3] use SOMs and combine them with geographic maps, small multiples, reorderable matrices, time series charts, or parallel coordinates. Another work by Andrienko et al. [6] combines clustering and interactive visual analysis with the aim of statistical modeling of geospatial time series. Woodring and Shen [40] combine wavelet transform, clustering, and interactive visualization for analysis of trends at varying temporal scales. These works use a single clustering, which captures only a particular aspect in the data. A different clustering leads to different results. In our application, however, it is important to consider multiple aspects of temporal behavior simultaneously.

Clustering ensembles [43] address this issue. They combine multiple clusterings into one clustering solution that shares as much information as possible with the input clusterings. With this approach one is able to cluster the data with varying features. In addition, it eases the burden on the users to find an optimal combination of (dis)similarity measure, algorithm, and parameterization of the clustering. They can select a set of plausible configurations and use cluster ensembles to combine the results. The resulting consolidated clustering is generally more robust and more accurate [18,31,34]. Although clustering ensembles have been shown to improve data analysis in a variety of fields – e.g., cancer research [22,41] or remote sensing [42] – we are, to the best of our knowledge, not aware of any works that apply this approach and its benefits to ocean modeling.

Another important aspect of our concept is the visual comparison of model data with reference data. As noted in a recent survey [29], many approaches for visualization and visual analysis in the Earth sciences have been introduced, but only a few works support the comparison of geoscientific simulation data. Nocke et al. [32] provide a library of comparative visualization techniques tailored to climate modeling. Based on the characteristics of the data and the task at hand, their framework generates an appropriate visualization. Unger et al. [39] address the validation of geoscientific simulation models. They compare many model outputs with sparse and uncertain observations. The focus is to find an appropriate model parameterization that best matches the observations. Ahrens et al. [1] use comparative visualization to support detection of errors in simulation model code. To this end, they conduct a numerical comparison of several output variables. Kehrer et al. [24,25] and Ladtstädter et al. [27] support visual analysis and comparison of different variables of climate model output by multiple linked views. Recent work by Poco et al. [33] focusses on the comparison of output from different climate models. Their approach concentrates on the analysis of correlations between data sets. These works do not base the comparison on geophysical processes, a key requirement of our users.

The concept that is closely related to our application, focusing on processes and applying clustering, was introduced by Frey et al. [19]. They support comparison of two temporal field data sets by combining automated detection of processes with interactive visual exploration. For the detection of processes, they use recurrence analysis and clustering. This approach regards processes as recurring events and places the emphasis on temporal similarity between data. In our application, we do not focus on recurrences but are interested in geographic regions that exhibit similar temporal behavior.

3 Visual analytics approach and requirements

We adopted a user- and task-centered approach [14] in our collaboration with ocean modelers at the German Research Center for Geosciences GFZ. This involved frequent meetings and discussions with our partners to obtain a detailed understanding of the model assessment process and the associated challenges. In this section, we provide an overview of our concept and the associated requirements.

3.1 Objectives for a visual analytics approach

As a result of our analysis we identified three main objectives for a visual analytics approach to facilitate comparison of model data with reference data.

(1) Less temporal aggregation Using a single statistical measure as a feature to describe a time series is a rather drastic approach to temporal aggregation. The ability to employ other types of features, such as the power spectrum of a time series, would reduce the amount of information lost and allow modelers to study more sophisticated characteristics of temporal behavior.

(2) More comprehensive comparison process The current comparison process requires modelers to make two main decisions that are rather subjective and may result in important aspects to remain hidden in the data. First, they have to choose a feature for temporal aggregation. A single feature, however, only describes a particular aspect of the temporal behavior.

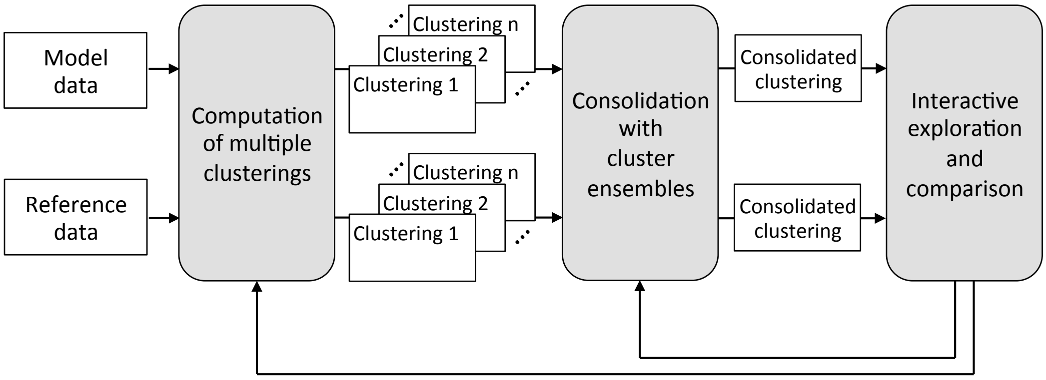


Figure 1: Our visual analytics concept: (1) modelers cluster model data and reference data with multiple configurations, (2) each set of clusterings is combined into one consolidated clustering, and (3) modelers interactively explore and compare the two consolidated clusterings. At any time during the exploration, scientists can either change the configuration of the consolidation or the set of input clusterings. Note that the two data sets are considered independently in the clustering and consolidation process.

Other features may capture different – but equally valid – aspects, and, hence, may yield different results. Being able to consider various features of temporal behavior in the comparison would broaden the scope of the analysis. Second, modelers hand-pick geographic coordinates for detailed comparison of temporal behavior. However, there is no guarantee that all relevant behavior can be observed at the selected coordinates. To reduce the risk of overlooking important aspects, modelers need to take geographic areas into account, not just a few coordinates.

(3) Enhanced visual exploration and comparison When the two objectives above are met, modelers will be able to study the output of a model in comparison to reference data from a broader perspective. To take full advantage of the additional information, scientists need a visual analytics tool that supports the comparison from this broader perspective. The tool should enable modelers to quickly identify and inspect temporal profiles that point to relevant geophysical processes in model data and reference data, and to assess differences and similarities between the data sets.

3.2 Concept

Based on our analysis and the identified objectives, we developed a threefold concept (Figure 1):

1. *Computation of multiple geospatial clusterings.*

   Our concept allows modelers to widen the scope of the comparison by performing multiple clusterings of the temporal profiles in model data and reference data. Since clustering considers the entire geographic domain and systematically detects groups of similar temporal profiles; modelers can compare geographic regions (the clusters) instead of a few geographic coordinates. For the clusterings, scientists can choose from a broad range of features of temporal profiles, instead of just single statistical measures. This provides them with more options for the detection of geophysical processes. The resulting two sets of clusterings represent various perspectives on the temporal behavior in the data.

2. *Consolidation of the clusterings via clustering ensembles.*

   To enable modelers to compare model data and reference data based on the two sets of clusterings, we combine each set into a separate consolidated clustering using clustering ensembles [34]. From a modeler's perspective, the consolidation of clusterings that are based on different features presents a more comprehensive, more robust view on the temporal behavior in a data set.

3. *Interactive visual exploration and comparison of the consolidated clusterings.*

   An interactive visual interface allows modelers to study and compare the two consolidated clusterings of model data and reference data. Scientists can perform visual queries to identify clusters that represent geoscientific processes, and to explore differences and similarities between the two data sets with respect to these clusters. Furthermore, modelers can go back to the previous two steps in the pipeline at any time. They can either choose to explore a differently configured consolidation result or they can change the set of input clusterings to, e.g., consider additional features of temporal behavior.

3.3 Requirements

In discussions with ocean modelers, we were able to identify four analytical requirements (ARs) and three visualization requirements (VRs) for our concept.

3.3.1 Analytical requirements

The analytical requirements comprise two requirements for the computation of multiple clusterings (AR1 and AR2), and two requirements for the consolidation with clustering ensembles (AR3 and AR4).

AR1 *Broad range of features*

In order to detect different geophysical processes or to study various characteristics of a single process, the set of input clusterings must represent different kinds of temporal behavior. Therefore, modelers need to be able to choose from many features to capture a wide variety of characteristics of temporal behavior when computing the geospatial clusterings.

AR2 *Many distinct cluster parameterizations*

A parameterization of a clustering algorithm reflects a specific assumption about the data to be clustered. However, modelers can often only make vague assumptions about the temporal behavior in model and reference data. For example, although the number of distinct types of temporal behavior in a data set can usually be narrowed down to a plausible range, it is difficult to anticipate the exact number. To account for this challenge, modelers need to be able to conduct clusterings with a varying number of clusters.

AR3 *Flexible configuration of the consolidation*

Modelers must consider two important aspects of a consolidated clustering: the amount of information it shares with the set of input clusterings, the quality, and its number of clusters, the complexity. A consolidated clustering that shares little information with the set of input clusterings is difficult to interpret. Likewise, a large number of clusters also complicates interpretation and comparison. To achieve a good balance between quality and complexity, modelers need to be able to configure different consolidations with a varying number of clusters and to study and compare the results.

AR4 *Quantitative measures to support the assessment of consolidated clusterings*

When presented with a consolidated clustering, modelers want to know how much information the individual clusters share with the input clusterings. This enables them to discriminate the clusters whose geographic locations were often considered as similar in the input clusterings, from the clusters where the input clusterings agree less on. This is important since the latter may not allow for a meaningful interpretation. Furthermore, to support an initial assessment of the relationships among clusters, modelers need quantitative measures that describe the geospatial similarity and the feature similarity between consolidation clusters.

3.3.2 Visualization requirements

VR1 *Overview of consolidated clusterings*

Modelers have to be able to interpret the results of the cluster ensembles. Therefore, they need to obtain an overview of the clusters and their relations in the two consolidated clusterings for model data and reference data. This requires visualizations that allow them (a) to filter the clusters that share only very little information with the input clusterings, (b) to detect clusters in each data set that may represent relevant geophysical processes, and (c) to identify potential matches between clusters from different data sets to guide further comparison. This requirement is associated with the following tasks:

- Assess distribution of clusters in geographic space.
- Obtain overview of geospatial similarity among clusters.
- Obtain overview of feature similarity among clusters.
- Compare individual clusters regarding the information they share with the input clusterings.

VR2 *Inspection of cluster properties*

Modelers must understand the properties of a single consolidated cluster to judge whether it is related to a geophysical process. To this end, they need to:

- Assess the temporal variations of input time series associated with a cluster.
- Inspect the distributions of feature values associated with a cluster.
- Inspect the distribution of cluster members in geographic space.
- Explore relations between input time series, feature values, and geographic distribution.

VR3 *Detailed comparison of clusters*

For a judgement of model quality, modelers need to explore and evaluate differences and similarities between consolidated clusters in model data and reference data. This allows them to identify geographic regions where the model performs well and where it needs improvement. The associated tasks are:

- Compare the temporal variations of input time series associated with clusters.
- Compare the distributions of feature values for multiple clusters.
- Compare the distribution of clusters in geographic space.

4 Clustering and consolidation

In this section, we describe the clustering and consolidation part of our concept and how it meets the analytical requirements AR1–AR4.

4.1 Computation of multiple geospatial clusterings

The computation of multiple geospatial clusterings has two aspects. As explained in AR1 (Section 3.3.1), modelers need to be able to cluster the data based on a variety of features. Therefore, our approach provides many features describing aspects of temporal behavior that scientists consider important for detecting geophysical processes; e.g., minimum and maximum value, mean, standard deviation, power spectrum, and logarithmic power spectrum. Note that scientists have the choice to cluster the raw data without prior computation of features. Scientists can also add additional features to focus on other types of geophysical processes.

Secondly, modelers can create multiple clusterings by varying the parameterizations of a clustering algorithm (AR2). The challenge in our application scenario was to identify a clustering method that is appropriate for ocean model output and reference data. We conducted a large number of experiments, clustering well understood observational data with different algorithms, parameterizations, features, and distance or similarity measures. The methods tested in these experiments were hierarchical clustering [23], DBSCAN [16] as a density-based method, a Gaussian mixture model approach [17], and k-means [8]. The clusterings were conducted with the following distance and similarity measures: Euclidean, Manhattan, mutual information [12], normalized compression distance [9], dissimilarity based on cross-correlation [28], and dynamic time warping [11]. We chose k-means and Euclidean distance because this combination yielded meaningful clusters over a broad range of parameterizations and features. Note that our concept is not limited to a specific clustering algorithm or distance measure. If need be, other methods can be included to provide modelers with additional options.

To satisfy AR2, modelers can vary the number of clusters and the number of iterations for the k-means algorithm. In addition, they can choose a distance measure (with Euclidean as default).

4.2 Consolidation

We use the cluster ensembles framework of Strehl and Ghosh [34] to combine the multiple clusterings into a single consolidated clustering. This popular technique aims at finding a consolidated clustering that maximizes the mutual information with a set of input clusterings. Mutual information is especially suited for our application because modelers are interested in the geographic regions where the input clusterings agree most on.

Since maximizing the mutual information is computationally prohibitive, we apply the three heuristics suggested in the cluster ensembles framework: the cluster-based similarity partitioning algorithm (CSPA), the hypergraph partitioning algorithm (HGPA), and the meta-clustering algorithm (MCLA). All heuristics first transform the set of input clusterings into a hypergraph. CSPA uses the relationships between objects expressed through the input clusterings to construct a measure of pairwise similarity. This measure can then be used with any similarity-based clustering algorithm (k-means in our case, see Section 4.1). HGPA performs a minimum cut operation to approximately maximize the mutual information. MCLA uses the hypergraph to identify and consolidate meta-clusters. Out of the consolidations resulting from the three heuristics, the one that has the highest average normalized mutual information (ANMI) with the input clusterings is chosen (see [34] for further details).

While HGPA determines the optimum number of clusters automatically, the other two heuristics allow for controlling the final number of clusters. To provide modelers with the required flexibility in the consolidation process (AR3), our tool enables them to apply MCLA and CSPA for a varying number of final clusters. They can then choose to be presented with the best result in terms of information shared, or select any of the other consolidated clusterings that were created by the three heuristics.

4.3 Quantitative measures

To support the assessment of the consolidated clusterings (AR4) we compute three quantitative measures.

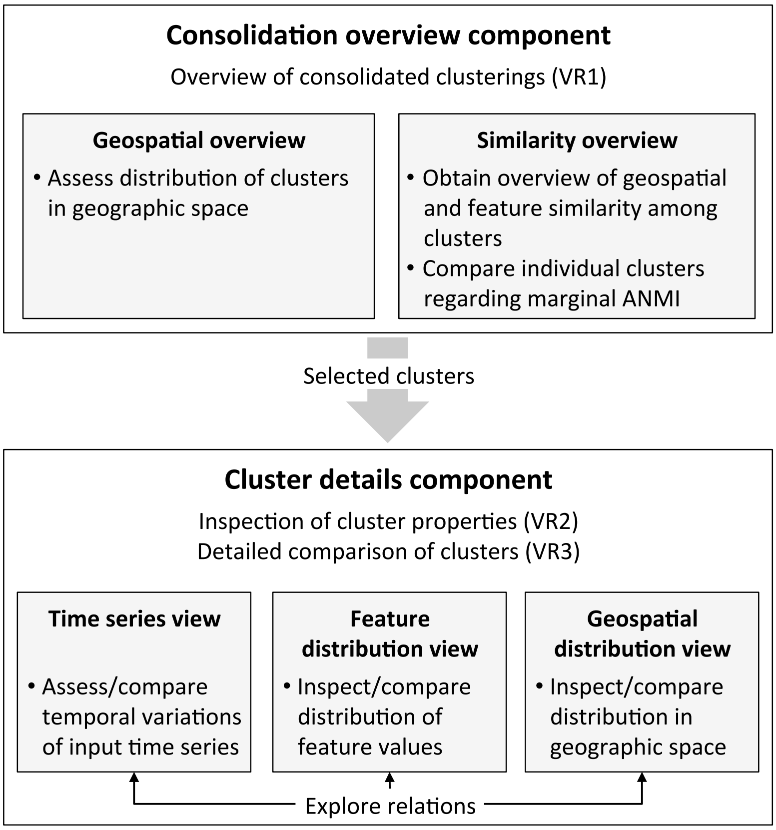


Figure 2: The components and views of our visual interface, and the visualization requirements and tasks they support. The consolidation overview component comprises two views which support modelers to obtain an overview of the consolidated clusterings. Based on this overview, modelers select clusters for detailed inspection and comparison in the cluster details component.

The first captures the information that a particular cluster in a consolidation result shares with the input clusterings. For its computation, we use the ANMI criterion from the cluster ensembles framework [34], but assume that the consolidated clustering only contains this particular cluster. We call this measure *marginal ANMI*. It allows modelers to identify clusters that should not be interpreted as geophysical processes.

We further compute the geospatial similarity as well as the feature similarity between consolidated clusters. These similarity measures are important criteria in the comparison of two data sets because they allow modelers to identify pairs of clusters that represent the same geophysical process. For the pairwise geospatial similarity, we compute the percentage of geographic overlap between clusters.

The pairwise feature similarity is determined as follows (assuming that multiple feature spaces were used to produce the input clusterings): first, we calculate the distance in each feature space between the centroids of two consolidated clusters; second, we normalize the separate distances, weigh and combine them. The weights for each feature space are assigned according to the number of input clusterings that were conducted in the respective feature space. Hence, the weights are implicitly provided by modelers since they create the set of input clusterings for the consolidation. Lastly, we convert this single distance measure into a similarity score.

5 Interactive visual exploration and comparison

To meet the visualization requirements outlined in Section 3.3.2, our visual interface for exploration, interpretation, and comparison of the consolidated clusterings comprises two types of

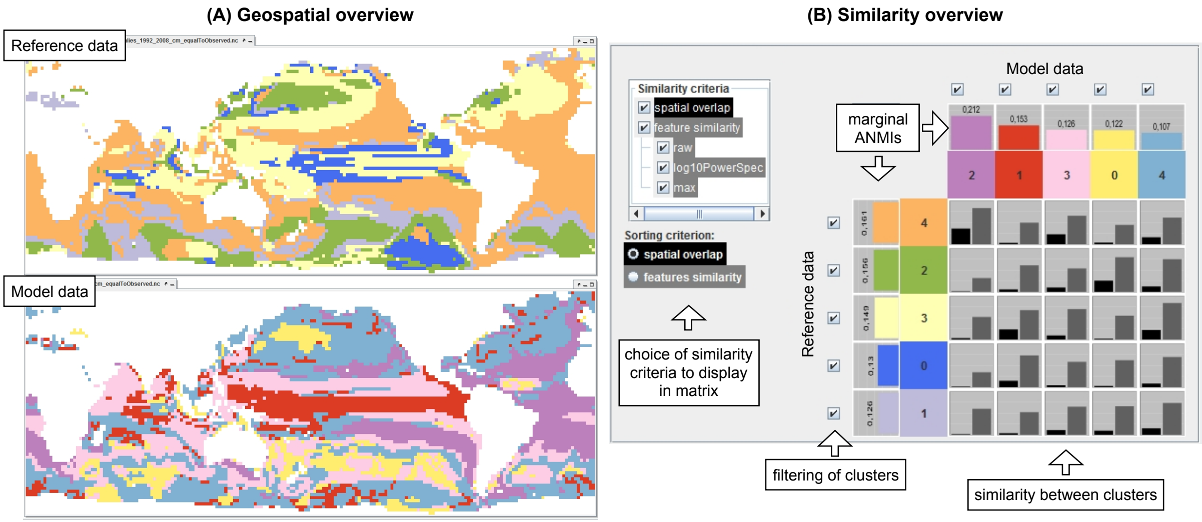


Figure 3: The consolidation overview component. Note that each cluster has its unique color since the cluster ensemble computes different clusters in both data sets. The consolidation overview supports scientists to focus on clusters that allow for a meaningful interpretation and to understand the relationships between clusters by comparing their geospatial and feature similarity.

coupled visualization components (Figure 2). A *consolidation overview component* enables users to gain a basic understanding of relations among clusters in geographic space and in feature space, and to decide on subsequent analysis steps (VR1). Researchers select clusters in this component and pass them to a *cluster details component* where they can inspect the properties of a single cluster (VR2) but also compare multiple clusters in detail (VR3). Both components allow for visual queries to support the identified analysis tasks.

To establish a visual link, all clusters are color-coded consistently across the consolidation overview and cluster detail components. To this end, we use one of ColorBrewer's qualitative color schemes [21] as well as colors sampled from the CIELAB color space (see Guo et al. [13] for a suitable sampling strategy). We chose the ColorBrewer colors to provide users with carefully designed and easily distinguishable colors. If additional colors are required, we use the CIELAB samples. This strategy yields a sufficient number of distinguishable colors. In addition, users can change the colors manually to adjust the color coding according to their preference. Note that we assign a unique color to clusters in both data sets since the consolidation process computes different clusters in both data sets.

The tight integration of the analytical part (clustering and consolidation via cluster ensembles) allows scientists to change the configuration of the consolidation as well as the set of input clusterings. The former is done to improve the balance between quality and complexity of the consolidated clusterings (see AR3 in Section 3.3.1); the latter allows modelers to change the features considered in the analysis. They can make these changes at any time during visual exploration and study the resulting consolidated clusterings.

In the following, we explain the visual encoding and interactive capabilities of each visualization component, and how they contribute to the visualization requirements VR1-VR3.

5.1 Consolidation overview component

This component provides an overview of the two consolidation results (VR1) and, thus, acts as a starting point for the visual comparison process. It includes two views: a *geospatial overview* and a *similarity overview* (Figure 3).

5.1.1 Geospatial overview

This view depicts the geospatial distribution of clusters in two maps; one for each consolidated clustering. The maps are juxtaposed to allow for visual comparison; cluster membership of geographic locations is encoded with a unique color for each cluster.

The geospatial overview allows users to observe a variety of patterns. For example, the maps in Figure 3 depict clusters that form quite coherent geospatial structures – e.g., the orange cluster in the model data – but also clusters with members that are more distributed over geographic space – e.g., the yellow cluster in the reference data. Notice also the different sizes of clusters as well as the horizontal structures along the equator.

Modelers interpret these patterns based on their domain knowledge and identify clusters that may represent geophysical processes. For example, the red cluster along the equator in the reference data probably represents a process in the tropics.

5.1.2 Similarity overview

This view provides quantitative information about the clusters and their pairwise similarity (Figure 3) to support modelers to develop a first understanding of differences and similarities between model data and reference data. For a compact visual overview, the clusters are arranged in a matrix layout. The rows represent the clusters from model data; columns represent the clusters in the reference data. The clusters are ordered with respect to their marginal ANMI score. Each matrix cell contains a bar chart that depicts the geographic similarity and the feature similarity (see Section 4.3) between a pair of clusters. This enables modelers to quickly detect similar clusters.

The similarity overview also allows modelers to assess the clusters regarding the information they share with the input clusterings. For this purpose, we visualize the marginal ANMI (see Section 4.3) for each cluster as colored bars next to the cluster labels in the row and column headers. We arrange the ANMI bars in this way to facilitate comparison of clusters by judging position along a common scale. Since users were primarily interested in a relative comparison of clusters, the bars are scaled to the highest marginal ANMI score among all clusters. In addition, the bars are labeled with their ANMI values to help users judge the absolute amount of information shared with input clusterings.

5.1.3 Linking and interaction

To support modelers in the assessment of relations among clusters, the similarity overview offers several filtering and sorting options.

In order to allow scientists to focus on a particular aspect of cluster similarity, we provide a checkbox tree (Figure 3) where they can select between displaying information on geospatial similarity, feature similarity, or both. They can further select the features to be considered in the computation of the feature similarity.

To obtain a better overview of similarities between clusters, modelers can also change the order of clusters in the matrix by mouse-clicking on a cluster label. When users click on a row label, the clusters in the columns are sorted in order of decreasing geospatial or feature similarity to the selected cluster, and vice versa. Users may switch the sorting criterion any time during analysis.

Lastly, checkboxes adjacent to the marginal ANMI bars allow modelers to visually filter clusters that they consider irrelevant or that share very little information with the input clusterings. Deselected clusters are greyed out in the similarity overview as well as in the geospatial overview.

An important functionality of the consolidation overview component is to enable users to choose and pass clusters to cluster details views for detailed inspection and comparison. To this end, modelers may either use the geospatial overview or the similarity view. In the geospatial overview, they may click on one of the maps to select a cluster and pass it to a cluster details view. In the similarity overview, modelers can click on a matrix cell to compare the two clusters associated with that cell in a new cluster details view.

5.2 Cluster details component

This visualization component allows modelers to inspect the properties of a single cluster (VR2) to determine whether it represents a geophysical process. It also enables them to compare multiple clusters in detail (VR3) to study differences and similarity between clusters in model and reference data.

A cluster details component comprises three visualizations (Figure 4a). The *time series view* depicts the temporal signatures associated with clusters, the *feature distribution view* helps users to inspect and compare the distribution of feature values for each cluster, and a *geospatial distribution view* facilitates detailed inspection and comparison of clusters in geographic space. All three views are tightly-coupled and allow users to conduct visual queries to the consolidation results via linking and brushing.

5.2.1 Time series view

This view presents the time series of the cluster centroid – which is the average time series over all cluster members – in a line chart. To assess the range of temporal variations in clusters, modelers can choose to display a semi-transparent minimum-maximum ribbon around each cluster representative (Figure 4b). The colors of the semi-transparent ribbons from different clusters mix in areas of overlap. This allows users to visually compare the temporal variations of multiple clusters. The time series view also enables users to interactively change the time frame via zooming to focus on particular time periods of interest.

5.2.2 Feature distribution view

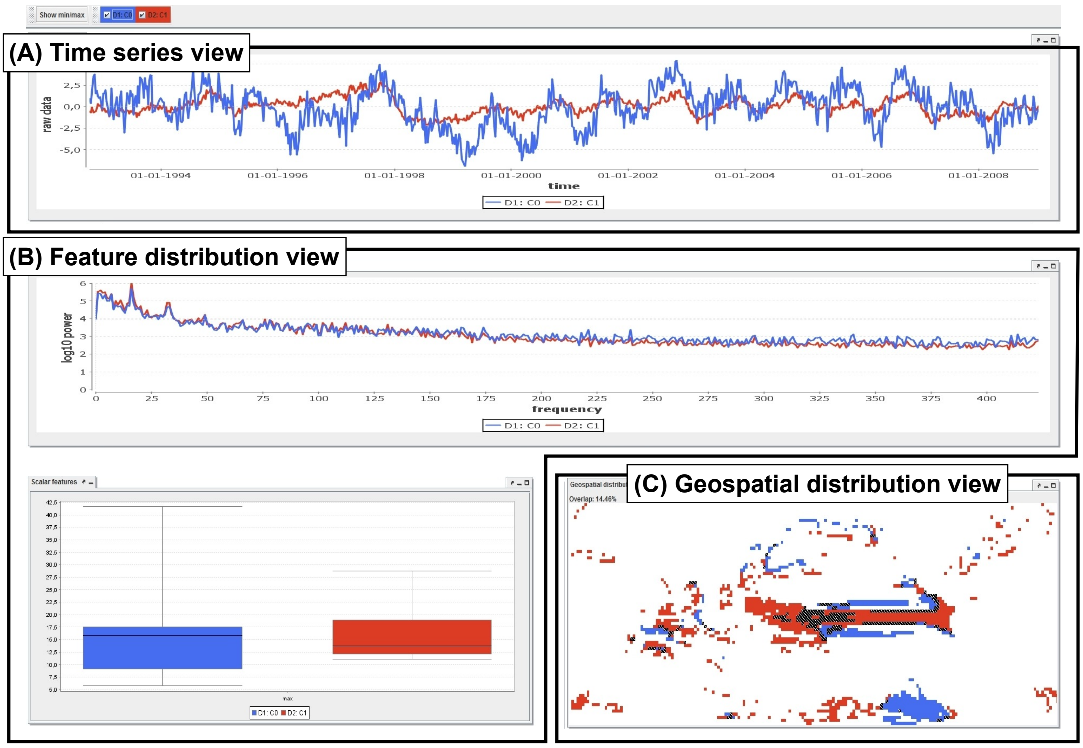
This view supports inspection and comparison of clusters regarding their distributions of feature values. Each feature is depicted in a separate visualization to reduce visual complexity. In accordance with the preference of our collaborators, we use line charts for features like the power spectrum of a temporal profile, and box-and-whisker plots for scalar features such as standard deviation. These two visualization types are appropriate for all features that our partners use for their analyses. To accommodate other types of features, this view can be easily extended to include additional visualizations such as star plots, 3D charts, or scatterplot.

The line charts are constructed in the same way as the time series view; cluster representatives are shown and optionally surrounded by a semi-transparent minimum-maximum ribbon (Figure 4b).

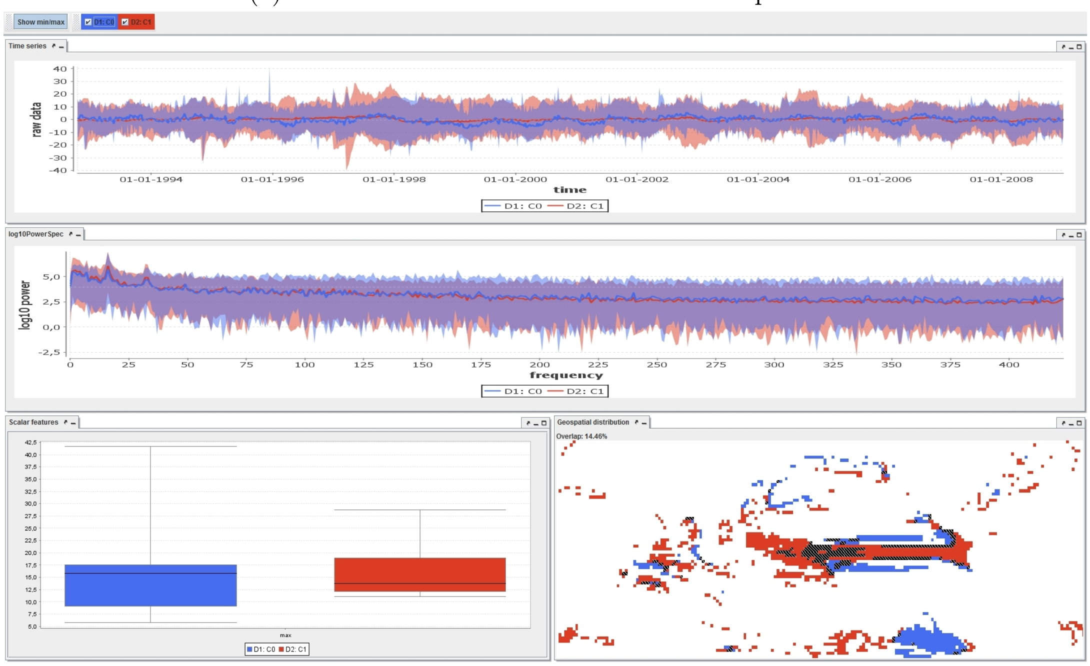
Box-and-whisker plots communicate comprehensive information about the distribution of feature values in clusters. Furthermore, juxtaposing box-and-whisker items from different clusters allows users to judge position along a common scale and, thus, facilitates comparison.

5.2.3 Geospatial distribution view

In this visualization, all clusters in the cluster details component are plotted in the same geographic space to facilitate detailed comparison of clusters from different consolidated clusterings. Areas where clusters overlap are highlighted to point researchers to similarities in geographic space (black areas in geospatial distribution views in Figure 4). This view also shows the total percentage of overlap between all selected clusters as quantitative information above the plotting area.



(a) The three views of a cluster details component.



(b) Cluster details component with semi-transparent minimum-maximum range ribbons in line charts.

Figure 4: The cluster details component.

5.2.4 Linking and interaction

The cluster details component provides several means of interaction to support inspection and comparison of clusters.

Apart from tooltips and zooming functionality, users can filter the clusters in this component using checkboxes at the top. This enables modelers to reduce visual complexity, for example, to focus on varying pairwise comparisons, to revisit the properties of a single cluster, or to assess how much a particular cluster contributes to the total geospatial overlap.

To allow modelers to explore relations between input time series, feature values, and geographic distribution of clusters, this visualization component offers three brushing mechanisms. First, researchers can brush a range of feature values in a box-and-whisker plot. The distributions of all cluster members that fall within the selected range are highlighted in the other plots in the cluster details component. Second, modelers may apply a vertical line brush in a line chart to determine all cluster members with values in the selected y-axis range at the specified x-axis index. Again, the corresponding distributions are highlighted in the other visualizations. Third, modelers can select either all overlapping or all non-overlapping geographic locations in the geospatial distribution view to study the distributions of features for these regions in the other views.

6 Application example: Ocean Model for Circulation and Tides

In this section, we explain how our approach supported the assessment of the Ocean Model for Circulation and Tides (OMCT) [36]. On the domain expert side of our collaboration in this particular example were two ocean modelers, one of them a leading OMCT expert and also co-author of this paper.

The OMCT simulates currents and tides of the global ocean and is used for removal of aliasing artifacts from observational data produced by the Gravity Recovery and Climate Experiment (GRACE) satellite mission [35]. GRACE data are used in the geosciences to gain valuable insight into a number of important processes on Earth, e.g., ice-mass changes, ocean tides, or Earth crust displacements associated with major earthquakes. To ensure a high quality of these widely used data, noise correction with models such as the OMCT is crucial.

The OMCT yields volumetric time series data, representing the ocean as 13 vertical depth layers of regular horizontal grids. Although the data comprise three geographic dimensions, an initial assessment of model output only requires to analyze the topmost layer, the sea surface, since most mechanisms within the ocean manifest themselves in changes to sea surface heights.

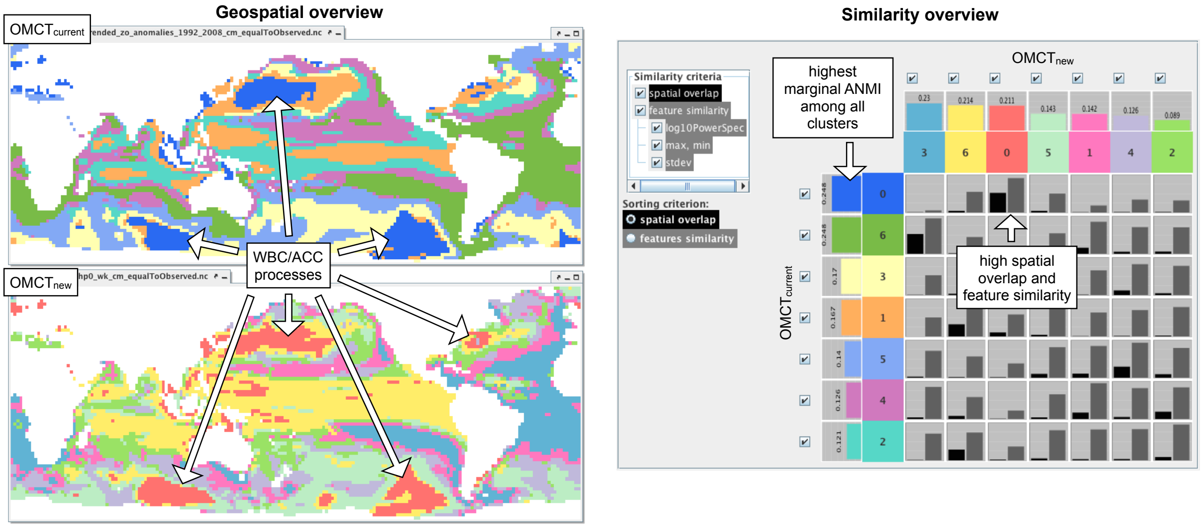
The most important ocean processes to consider during the assessment of OMCT data are the western boundary currents (WBC) and the antarctic circumpolar current (ACC). These ocean currents cause high spatial and temporal variability of sea surface heights around Antarctica, South Africa and on the northwest-side of the Atlantic and Pacific Ocean basins.

The remainder of this section describes how our tool supported the assessment of a new version of the OMCT ($OMCT_{new}$) regarding its ability to depict WBC and ACC processes.

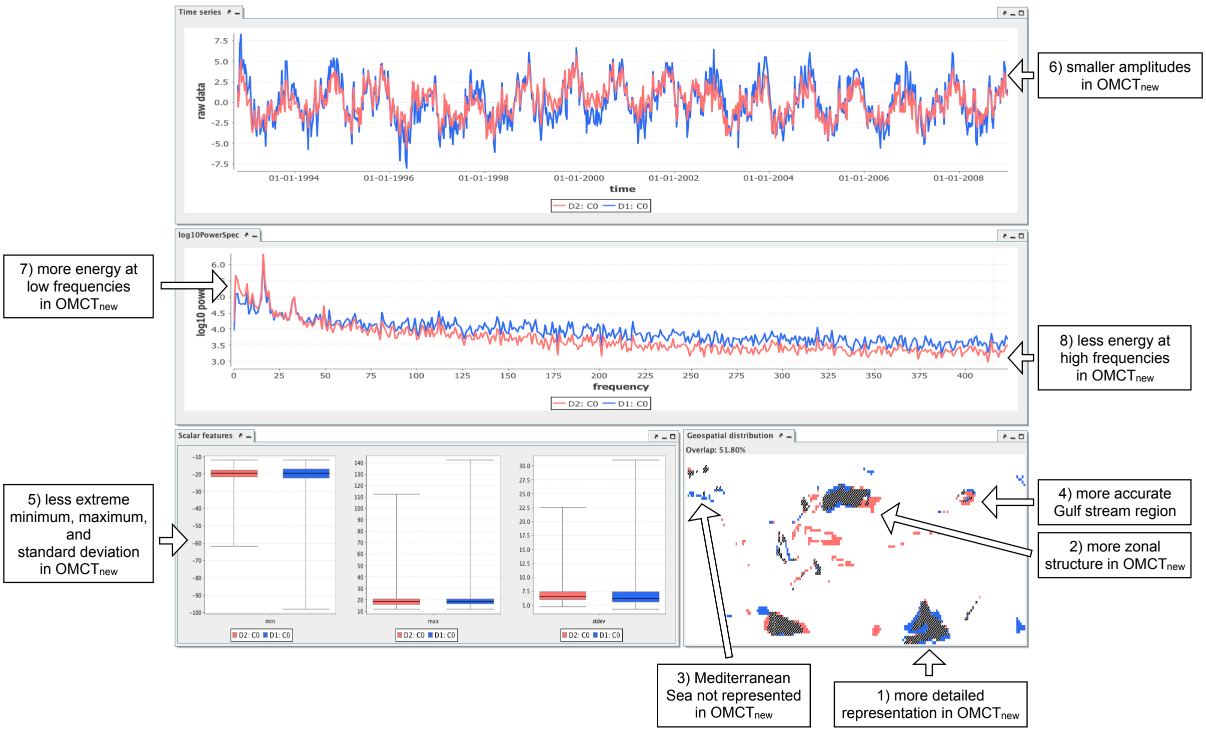
6.1 Results

To asses the $OMCT_{new}$, modelers compared sea-level anomalies simulated with this version with data produced by the current state-of-the-art OMCT version ($OMCT_{current}$). Each data set comprised approximately 9000 time series.

Applying our approach to the comparison of the two OMCT versions allowed for three important accomplishments: (1) the consolidation enabled modelers to capture the WBC and ACC processes in both data sets, (2) the visual interface permitted modelers to readily identify the clusters that represent these processes and to study relations between these clusters across



(a) Consolidation overview showing consolidated clusterings for $OMCT_{current}$ (upper left) and $OMCT_{new}$ (lower left). The dark blue cluster and the red cluster represent western boundary currents and antarctic circumpolar currents. The similarity overview indicates that both clusters share a relatively high amount of information with the input clusterings and have high spatial overlap and feature similarity.



(b) Cluster details component comparing WBC/ACC clusters from $OMCT_{current}$ (dark blue) and $OMCT_{new}$ (red). $OMCT_{new}$ improves the simulation in various aspects. The red $OMCT_{new}$ cluster is a more accurate geographic representation of WBC/ACC processes (1-4), it also exhibits less extreme temporal behavior (5, 6), as well as an improved power spectrum (7, 8).

Figure 5: Comparison of a new version of the Ocean Model for Circulation and Tides (OMCT) – $OMCT_{new}$ – with the current state-of-the-art OMCT version – $OMCT_{current}$.

data sets, and (3) detailed visual comparison helped them two determine that the $OMCT_{new}$ significantly improves the state-of-the-art OMCT version.

In the following, we provide more details about each of these three accomplishments.

(1) Capturing WBC and ACC processes. To describe WBC and ACC, modelers chose standard deviation, minimum and maximum, and the logarithmic power spectrum as features. These features describe the variation in a temporal profile from varying angles and in different granularities. Standard deviation summarizes the variability in a temporal profile in a single scalar value. The minimum-maximum feature is a vector of size two that provides information about the range of values that can be found in a temporal profile. The logarithmic power spectrum is a vector of size $n/2 - 1$ where n is the length of a time series. It provides detailed information about the energy that a temporal profile exhibits at particular frequencies. Out of these three features, modelers previously could only use standard deviation to capture the variability of temporal profiles and to compare them in geographic space. Likewise, the power spectrum could only be used for detailed comparison of temporal profiles at a few geographic coordinates. Our visual analytics approach allows modelers to use all three features to compare the temporal profiles at all geographic coordinates.

Our collaborators concluded that WBC and ACC processes can be differentiated into at least two but not more than twelve different types. Therefore, they chose k-means clustering for each of the three features with k ranging from two to twelve as input for the consolidation. This resulted in 33 input clusterings for each data set. Modelers also set the number of clusters for the consolidated clusterings to range from two to twelve. The consolidated clusterings with twelve clusters shared the most information with the sets of input clusterings. However, modelers determined with our tool that the results with only seven clusters had a much better balance between quality and complexity. In particular, the number of clusters was reduced by 40%, while the information shared with the input clusterings decreased by only 0.2% for $OMCT_{current}$ and 4% for $OMCT_{new}$.

(2) Identifying WBC/ACC clusters in both data sets. First, the geospatial overview in our tool allowed modelers to scan the consolidated clusterings for clusters that represent WBC and ACC processes (Figure 5a left). Based on their knowledge about the geospatial distribution of WBC/ACC and the geospatial shapes of the clusters, modelers quickly identified the dark blue cluster in the $OMCT_{current}$ data and the red cluster in the $OMCT_{new}$ as candidates.

Next, modelers were able to discern from the marginal ANMI scores in the similarity overview (Figure 5a right) that these two clusters share significant information with the input clusterings, and, hence, can be interpreted as geophysical processes. Of all clusters in the two data sets, the dark blue WBC/ACC cluster had the highest marginal ANMI, while the score for the red cluster was also relatively high. Modelers could also tell from the similarity overview that the two clusters have high geospatial and feature similarity and, thus, represent the same geophysical process.

(3) Detailed comparison of model versions. In developing a new model version, ocean modelers wanted to improve a number of aspects of the current state-of-the-art OMCT version ($OMCT_{current}$): (a) more detailed simulation of WBC and ACC regions, (b) improved simulation of WBC in the Gulf stream region, (c) smaller amplitudes in sea-surface heights over time, and (d) more energy for low frequencies and less energy for high frequencies in the power spectra of temporal profiles.

Exploring the two WBC/ACC clusters in a cluster details component (Figure 5b) allowed for assessing all of these aspects. The geospatial distribution view enabled modelers to notice four geographic characteristics that have improved with the new model version: (1) the region west of South Africa (lower right overlap) is represented in much more detail, (2) the WBC region in the northwest Pacific (top middle overlap region) is simulated in a more zonal structure, (3) the Mediterranean Sea is not represented in the red $OMCT_{new}$ cluster, and (4) the Gulf stream region is also represented more accurately in the $OMCT_{new}$ cluster. In sum, modelers concluded from the geospatial distribution view that the new model version provides a more accurate geographic representation of WBC and ACC processes.

In addition, the time series view and the feature distribution view show that the red $OMCT_{new}$ cluster exhibits less extreme temporal behavior with smaller amplitudes. Lastly, a comparison of the logarithmic power spectra (Figure 5b) reveals that the final objective has also been met. In comparison to the $OMCT_{current}$ cluster, the $OMCT_{new}$ cluster has more energy at low frequencies and less energy in higher frequencies.

All the above findings constitute aspects of potential improvement in the new version of the OMCT – aspects that our partners could readily study with the help of our approach. After additional statistical analyses to corroborate the increase in quality, the $OMCT_{new}$ became the new state-of-the-art OMCT version.

6.2 User feedback

Our partners consider our approach a valuable complement to their existing tools and routine for four primary reasons: (1) they are not limited any more to single statistical measures for the detection of geophysical processes, instead, they have access to a range of features of temporal behavior, (2) the combination of multiple clusterings and cluster ensembles improves the detection of geophysical processes because multiple features of temporal behavior are considered simultaneously, (3) our interactive tool enables modelers to obtain a more complete picture about differences and similarities between model and reference data, and (4) it has great potential for speeding up the model development process because it supports a quick initial assessment of new model versions.

Our partners also value the flexibility of our approach. Our tool can be extended to include any feature that describes the temporal behavior represented in a temporal profile and, hence, allows scientists to study any geophysical process that may be of relevance to the assessment. Modelers are also highly flexible regarding the visual exploration. Although the consolidation overview component provides them with important information that guides subsequent analysis steps, they can always decide to make an educated guess and readily study any potentially interesting aspect. Before they had our tool, such an educated guess was not feasible because it involved time-consuming scripting and plotting procedures, a problem that also exists in related domains such as climate research [33, 37]. The means of interaction provided in our tool effectively remove this hurdle.

6.3 Discussion

Although our concept provides significant benefits to modelers in the assessment of ocean models, several issues need to be discussed.

First, the quality of the input clusterings determines the quality of the consolidated clusterings. Therefore, the features chosen for the clusterings must capture geophysical processes. If not, the consolidation will not yield meaningful clusters. However, since our approach was developed with and for expert users, one can assume that the features chosen will be appropriate for the respective analysis task.

Second, the computational complexity of the consolidation process is quadratic in the number of temporal profiles in a data set. Although this can be addressed in future work, it has not been a major issue in our application for two reasons. (1) A single simulation run of an ocean model typically takes several days if not weeks; in this context, the required time for the consolidation process is negligible. (2) In the opinion of modelers, the benefits of our approach outweigh the downside of the high computational complexity. Depending on the hardware available, the automated analysis part of our approach is applicable to models in the range of 100K grid points.

Another point worth mentioning is that the time series view is currently limited to a reasonable number of time steps in the temporal profiles. In practice, however, this was no issue because ocean modelers typically study weekly, monthly, or even seasonal averages. This approach is also applied by climate scientists (as described by Poco et al. [33]) and significantly reduces the number of time steps per year.

7 Conclusion and future work

In this article, we presented a visual analytics concept that addresses a crucial part in ocean modeling: the comparison of model output with reference data. This concept was developed in close collaboration with ocean modelers, which allowed us to identify the primary challenges: the drastic aggregation that had to be performed and the high degree of subjectivity in the comparison process. To address these challenges we integrate clustering ensembles and interactive visual analysis into a tightly-coupled system. This approach is based on a comprehensive task and requirement analysis.

We have shown that the combination of data mining and interactive visual analysis can be of high value to the assessment of ocean models. We could also observe in our collaboration that the promising results of our work have led to increasing acceptance of visual analytics in the ocean modeling community.

To further enhance our approach, we identified several major areas of future work. The next step is to conduct an in-depth user study to further corroborate the promising results we were able to achieve so far. We also want to extend our concept to very high-resolution ocean models, which requires improving the scalability of the visualization components as well as reducing the time and storage complexity of the consolidation. Therefore, we would like to investigate the applicability of distributed computing and GPU processing to our concept. We also plan on working on efficient algorithms to further speed up the consolidation and to better adapt to the characteristics of ocean model data. Currently, our concept supports the analysis of sea surface heights or any other meaningful two-dimensional layer. To apply our approach to all three geospatial dimensions, we will, again, have to address the scalability of the automated analysis, but in addition, identify and meet the visualization requirements that come with the third spatial dimension. We would also like to extend our approach from a two-way comparison to a three-way comparison. This would support an even more comprehensive assessment. A three-way comparison, however, also introduces additional challenges for visual analytics. Finally, our vision is to incorporate other types of geoscientific simulation models, beginning with climate models since their characteristics are somewhat similar to ocean models.

Acknowledgements

The authors thank Joachim Fohringer, Ralf Friedeman, and Alexander Bobach for their help in implementing the prototype. This study was partially supported by the German Federal Ministry for Education and Research (BMBF) via the Potsdam Research Cluster for Georisk Analysis, Environmental Change and Sustainability (PROGRESS) (grant number 03IS2191A).

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(6):16–28, 2010.

[2] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. *Visualization of time-oriented data.* Springer, 2011.

[3] G. Andrienko, N. Andrienko, P. Bak, S. Bremm, D. Keim, T. von Landesberger, C. Pölitz, and T. Schreck. A framework for using self-organising maps to analyse spatio-temporal patterns, exemplified by analysis of mobile phone usage. *Journal of Location Based Services*, 4(3–4):200–221, 2010.

[4] G. Andrienko, N. Andrienko, S. Bremm, T. Schreck, T. Von Landesberger, P. Bak, and D. Keim. Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. *Computer Graphics Forum*, 29(3):913–922, 2010.

[5] N. Andrienko and G. Andrienko. *Exploratory analysis of spatial and temporal data.* Springer, Berlin, Heidelberg, 2006.

[6] N. Andrienko and G. Andrienko. A visual analytics framework for spatio-temporal analysis and modelling. *Data Mining and Knowledge Discovery*, 27(1):55–83, 2013.

[7] N. Andrienko, G. Andrienko, and P. Gatalsky. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing*, 14(6):503–541, 2003.

[8] D. Arthur and S. Vassilvitskii. k-means++: the advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035, 2007.

[9] C. Bennett, P. Gacs, M. Li, P. Vitanyi, and W. Zurek. Information distance. *IEEE Transactions on Information Theory*, 44(4):1407–1423, 1998.

[10] P. Berkhin. A survey of clustering data mining techniques. In J. Kogan, C. Nicholas, and M. Teboulle, editors, *Grouping Multidimensional Data*, pages 25–71. Springer, Berlin, Heidelberg, 2006.

[11] D. J. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *AAAI-94 Workshop on Knowledge Discovery in Databases*, volume 10, pages 359–370. Seattle, WA, 1994.

[12] M. Cover Thomas and A. Thomas Joy. *Elements of information theory.* Wiley, New York, 1991.

[13] Diansheng Guo, Jin Chen, A. M. MacEachren, and Ke Liao. A Visualization System for Space-Time and Multivariate Patterns (VIS-STAMP). *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1461–1474, 2006.

[14] 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.

[15] J. Dykes, A. M. MacEachren, and M.-J. Kraak, editors. *Exploring geovisualization.* Elsevier, Oxford, 2005.

[16] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, volume 96, pages 226–231, 1996.

[17] C. Fraley and A. E. Raftery. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–631, 2002.

[18] A. L. N. Fred and A. Jain. Data clustering using evidence accumulation. In *Proceedings of the 16th International Conference on Pattern Recognition*, volume 4, pages 276–280, 2002.

[19] S. Frey, F. Sadlo, and T. Ertl. Visualization of temporal similarity in field data. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2023–2032, 2012.

[20] J. Han and M. Kamber. *Data mining: Concepts and techniques.* Morgan Kaufmann, San Francisco, 2nd edition, 2006.

[21] M. Harrower and C. A. Brewer. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, 40(1):27–37, 2003.

[22] N. Iam-On, S. Garrett, C. Price, and T. Boongoen. Link-based cluster ensembles for heterogeneous biological data analysis. In *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2010*, pages 573–578, Dec 2010.

[23] L. Kaufman and P. J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis.* John Wiley & Sons, 2009.

[24] 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, Nov 2008.

[25] 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):934–946, July 2011.

[26] P. Köthur, M. Sips, A. Unger, J. Kuhlmann, and D. Dransch. Interactive visual summaries for detection and assessment of spatiotemporal patterns in geospatial time series. *Information Visualization*, 13(3):283–298, 2014.

[27] 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.

[28] T. W. Liao. Clustering of time series data – a survey. *Pattern Recognition*, 38(11):1857 – 1874, 2005.

[29] D. R. Lipşa, R. S. Laramee, S. J. Cox, J. C. Roberts, R. Walker, M. A. Borkin, and H. Pfister. Visualization for the physical sciences. *Computer Graphics Forum*, 31(8):2317–2347, 2012.

[30] A. M. MacEachren. *How maps work: representation, visualization, and design.* Guilford Press, 2004.

[31] M. Nisha, S. Mohanavalli, and R. Swathika. Improving the quality of clustering using cluster ensembles. In *Proceedings of the IEEE Conference on Information Communication Technologies (ICT) 2013*, pages 88–92, 2013.

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

[33] J. Poco, A. Dasgupta, Y. Wei, W. Hargrove, C. Schwalm, R. Cook, E. Bertini, and C. Silva. SimilarityExplorer: A Visual Inter-Comparison Tool for Multifaceted Climate Data. *Computer Graphics Forum*, 33(3):341–350, 2014.

[34] A. Strehl and J. Ghosh. Cluster ensembles - a knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research*, 2002(3):583–617, 2002.

[35] B. D. Tapley, S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins. Grace measurements of mass variability in the earth system. *Science*, 305(5683):503–505, 2004.

[36] M. Thomas, J. Sündermann, and E. Maier-Reimer. Consideration of ocean tides in an ogcm and impacts on subseasonal to decadal polar motion excitation. *Geophysical Research Letters*, 28(12):2457–2460, 2001.

[37] C. Tominski, J. Donges, and T. Nocke. Information visualization in climate research. In *Proceedings of the 15th International Conference on Information Visualization (IV)*, pages 298–305, 2011.

[38] K. E. Trenberth, editor. *Climate System Modeling.* Cambridge Universtity Press, New York, 2009.

[39] A. Unger, S. Schulte, V. Klemann, and D. Dransch. A visual analysis concept for the validation of geoscientific simulation models. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2216–2225, 2012.

[40] J. Woodring and H.-W. Shen. Multiscale Time Activity Data Exploration via Temporal Clustering Visualization Spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 15(1):123–137, 2009.

[41] Z. Yu, H.-S. Wongb, J. You, Q. Yang, and H. Liao. Knowledge based cluster ensemble for cancer discovery from biomolecular data. *IEEE Transactions on NanoBioscience*, 10(2):76–85, 2011.

[42] X. Zhang, L. Jiao, F. Liu, L. Bo, and M. Gong. Spectral clustering ensemble applied to SAR image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7):2126–2136, 2008.

[43] Z.-H. Zhou. *Ensemble methods: foundations and algorithms.* Machine learning & pattern recognition series. Chapman & Hall/CRC, Boca Raton, FL, 2012.