



## The state of the art in visual analysis approaches for ocean and atmospheric datasets

Item Type	Article
Authors	Afzal, Shehzad;Hittawe, M. M.;Ghani, Sohaib;Jamil, Tahira;Knio, Omar;Hadwiger, Markus;Hoteit, Ibrahim
Citation	Afzal, S., Hittawe, M. M., Ghani, S., Jamil, T., Knio, O., Hadwiger, M., & Hoteit, I. (2019). The State of the Art in Visual Analysis Approaches for Ocean and Atmospheric Datasets. Computer Graphics Forum, 38(3), 881–907. doi:10.1111/cgf.13731
Eprint version	Post-print
DOI	<a href="https://doi.org/10.1111/cgf.13731">10.1111/cgf.13731</a>
Publisher	Wiley
Journal	Computer Graphics Forum
Rights	Archived with thanks to Computer Graphics Forum
Download date	2024-03-13 07:10:03
Link to Item	<a href="http://hdl.handle.net/10754/656777">http://hdl.handle.net/10754/656777</a>

# The State of the Art in Visual Analysis Approaches for Ocean and Atmospheric Datasets

S. Afzal<sup>1</sup>, M.M. Hittawe<sup>1</sup>, S. Ghani<sup>1</sup>, T. Jamil<sup>1</sup>, O. Knio<sup>1</sup>, M. Hadwiger<sup>1</sup> and I. Hoteit<sup>1</sup>

<sup>1</sup>King Abdullah University of Science & Technology (KAUST), Saudi Arabia

---

## Abstract

*The analysis of ocean and atmospheric datasets offers a unique set of challenges to scientists working in different application areas. These challenges include dealing with extremely large volumes of multidimensional data, supporting interactive visual analysis, ensembles exploration and visualization, exploring model sensitivities to inputs, mesoscale ocean features analysis, predictive analytics, heterogeneity and complexity of observational data, representing uncertainty, and many more. Researchers across disciplines collaborate to address such challenges, which led to significant research and development advances in ocean and atmospheric sciences, and also in several relevant areas such as visualization and visual analytics, big data analytics, machine learning and statistics. In this report, we perform an extensive survey of research advances in the visual analysis of ocean and atmospheric datasets. First, we survey the task requirements by conducting interviews with researchers, domain experts, and end users working with these datasets on a spectrum of analytics problems in the domain of ocean and atmospheric sciences. We then discuss existing models and frameworks related to data analysis, sense-making, and knowledge discovery for visual analytics applications. We categorize the techniques, systems, and tools presented in the literature based on the taxonomies of task requirements, interaction methods, visualization techniques, machine learning and statistical methods, evaluation methods, data types, data dimensions and size, spatial scale and application areas. We then evaluate the task requirements identified based on our interviews with domain experts in the context of categorized research based on our taxonomies, and existing models and frameworks of visual analytics to determine the extent to which they fulfill these task requirements, and identify the gaps in current research. In the last part of this report, we summarize the trends, challenges, and opportunities for future research in this area.*

(see <http://www.acm.org/about/class/class/2012>)

## CCS Concepts

• **Human Centered Computing** → Visualization; Visual Analytics; • **Physical Sciences and Engineering** → Earth and atmospheric sciences;

---

## 1. Introduction

Ocean and atmospheric data analysis is important for many domains to facilitate forecasting, planning, and decision making. Research in this area is highly interdisciplinary, and scientists from diverse research backgrounds collaborate to address challenges in enabling such analysis. Significant research has been carried out in the visualization and visual analytics domain to develop techniques and systems to facilitate further research in the domain of ocean and atmospheric sciences. Advances in technology and computational resources have enabled scientists to generate large amounts of data with higher density, complexity, and diversity. Researchers aim to run high-resolution ocean and atmospheric models [SMG\*15], ensembles visualization and analysis [BLLS17, FKRW17, WLSL17, KRRW18], study ensemble uncertainties [SZD\*10, WMK13, ME18, LBR\*17, KTB\*18, AZM\*15], examine the sensitivity of parameters [WLSL17,

KBL18, KRRW18], analyze eddies (mesoscale ocean features) and associated uncertainty in high-fidelity simulations [WPS\*16, WHP\*11], interactive querying and filtering in multidimensional data [ME18, KHS\*18b, SSJKF09, LZM14, DPD\*15], and apply predictive analytics [LPCRH18, LMK\*15]. Dealing with data of such volume and complexity for different application requirements presents new challenges for visualization and visual analytics researchers, and has led to the development of many data analysis and visualization techniques. As a result, there is an ever-growing number of publications that target different areas including: big data storage, management, processing and summarization; visualization design for multi-resolution, multi-dimensional and time-varying datasets; support for interactive visual analysis, multidimensional querying and filtering; ensembles visualization; eddies analysis; uncertainty representation and visualization; predictive analytics; and interactive data analysis and steering of simulations [SGL\*16, HMC\*13, WJW\*17, KRRW18, KHS\*18a, KBL18,

FTMKH\*18, WPS\*16]. Keeping in view this rapid growth in the visualization research relevant to the ocean and atmospheric data, there is a need for a structured review of the current state of the art. As well as reviewing the current state of the art, we also need to evaluate whether research advancements in the visualization domain are in agreement with the task and system requirements of ocean and atmospheric scientists. We also need to determine whether existing visual analytics models and frameworks for data analysis, sense making, and knowledge discovery adequately address the task and system requirements of ocean and atmospheric domain experts. We also need to assess to what extent the current visualization research related to ocean and atmospheric datasets supports the components and functionality of these visual analytics models and frameworks.

To achieve these goals, we have conducted interviews with 18 domain experts from academia and industry working with ocean and atmospheric datasets to gather their task and system requirements. Then, we surveyed recent research advances in the visualization community related to the visual analysis of ocean and atmospheric datasets. We then categorized and cataloged this research based on the taxonomies of *task requirements*, *interaction methods*, *visualization techniques*, *machine learning and statistical methods*, *evaluation methods*, *data types*, *data dimensions and size*, *spatial scale*, and *application areas*. The task requirements identified earlier based on our interviews with domain experts are evaluated in the context of categorized research based on our taxonomies. This helped us understand the current research trends and identify any gaps that exist between the task requirements of the domain experts and visual analytics research. In the end, we examined how the task requirements of domain experts relate to visual analytics frameworks and models. We also apply these models and frameworks to the survey results generated by coding survey papers based on ten different taxonomies. This helped us gain insights about the major challenges, current trends, and opportunities for future research to the visualization community.

## 2. Survey Organization

This state-of-the-art report is organized as follows. Section 3 discusses relevant surveys on this and other closely related topics [RBS\*17, KH13, WHLS18], focusing in particular on the visualization and visual analytics research, highlighting the significance and differences of our survey report as compared to others, and what motivated us to write a survey on this topic.

In section 4, we provide details about our survey methodology, including details about how we defined the scope of our work, built a set of keywords to search, and generated a sample set of papers for classification based on our taxonomies. Section 5 provides details about interviews conducted with domain experts to gather task and system requirements for designing visual analytics applications for ocean and atmospheric datasets.

Section 6 provides details about the ten taxonomies used for classification of papers included in this survey and contains the corresponding tables showing these classification or coding results. We also discuss why we chose these categories of taxonomies, and how it compares with the task requirements of domain experts mentioned in section 5.

In Section 7, we discuss existing visual analytics models and frameworks related to data analysis, knowledge generation, and sense-making [SSS\*14, FWR\*17, PC05, EM10, KMH06a, KMH06b, TC05], and evaluate their suitability in the context of supporting the task requirements for visual analysis of ocean and atmospheric datasets, identified during our interviews with domain experts.

Section 8 provides details about the web application to explore papers included in this survey across different taxonomies. Section 9 & 10 discuss details about the current trends and patterns in visualization research relevant to ocean and atmospheric datasets, highlights opportunities and challenges to conduct future research that aligns with the tasks and requirements identified during interviews with domain experts. We also provide details about whether current visual analytics frameworks can address these research challenges.

## 3. Previous Relevant Surveys

There are some recent surveys that are related to our survey topic. Rautenhaus et al. [RBS\*17] conducted a survey on visualization in meteorology focusing on tools and techniques utilized in operational weather forecasting and atmospheric research with a scope limited to data analysis tasks. The survey was conducted from the perspective of meteorological research (not from the perspective of visual analytics) with an objective to review visualization research that can advance meteorological research. The scope of this survey does not extend to other areas such as decision-making, communication, frameworks and models for visual analytics, evaluation methods, interaction techniques, etc. In this report, we attempt to address these shortcomings with the objective of identifying the challenges and limitations that restrict the collaboration and exchange between visualization researchers and ocean/atmospheric researchers.

Tominski et al. [TDN11] performed a study with 76 participants working in climate research to analyze the use of interactive visual analysis methods in their work. They found that the state-of-the-art methods are used very infrequently and integration of existing solutions in their work flows is problematic. There is a need to systematically analyze the visualization research and task requirements of domain experts working on ocean and atmospheric sciences to understand different issues in this interdisciplinary research. In this survey, we utilize certain well-defined taxonomies to classify the surveyed research, and analyze the trends and patterns in light of task requirements of ocean and atmospheric domain experts.

A survey conducted by Wang et al. [WHLS18] examined the state of the art in ensemble visualization and was particularly focused on ensembles from computer simulation models. The survey was structured based on three categories: ensemble data (variable, location, time, member, ensemble), visualization techniques (point, curve surface, volume, non-spatial), and analytic tasks (overview, compare, cluster, trend, feature, parameter).

Nocke et al. [NSBW08] conducted a questionnaire with researchers using visualization tools working in the area of climate impact, and discussed the state of the art in climate data visualization. The questionnaire was focused on determining the most frequently used visualizations, major tasks, important features, and

[illegible]

ble i

OS a  
a lite  
tion  
n-in

used  
iona  
our  
ns

of ke  
ces  
ted

One  
phen  
g in  
en v

ide  
vis  
eva  
me  
mo

su  
'S  
om

Was  
as  
un  
an

4.

We  
for  
the  
sea  
no  
da  
of  
alc  
us





Interaction	BBP08																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
-------------	-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 3: Interaction taxonomy and coded papers.

3D velocity, temperature, sensitivity fields) in their analysis tasks, and observational data (e.g., satellite observational data, rainfall, atmospheric data) was the next most-used data category. In some cases, the experts have to simultaneously analyze data belonging to different categories. Our interviews revealed that almost 60% of the experts work with data at the Terabyte scale, 10% at Petabyte, and the remainder at the Gigabyte scale. Their analysis tasks usually involve dealing with data having 1-10 dimensions at any given time.

More than 75% of the domain experts use high-performance clusters or super computers while working with ocean and atmospheric datasets. It is evident that any visualization tool intended to fulfill the task requirements of scientists working in this area must address these storage and computational challenges. The experts described how they require the capability to analyze data at different spatial (e.g., global, synoptic, meso, micro) and temporal scales (e.g., hourly, daily, monthly, yearly, seasonal, decadal) with features like selection, aggregation, summarization, filtering, sorting, etc., along with the support for interactive exploration where users have an overview of the data first, and can explore more details-on-demand.

Some experts mentioned a lack of interface, interoperability, or difficulty to move back and forth between the data analysis and visualization tools. In some scenarios, they have their custom data analysis code but they cannot easily integrate visualizations in their existing code base due to these issues. Below is a comment from one of the experts:

*"I used CDO (climate data operators) and NCO (netCDF operators) which supports the analysis of big and long-term data for climate research but it doesn't support visualization of big data"*

CDO [CDO] and NCO [NCO] are a collection of command-line programs and operators that can operate with different file formats (e.g. netCDF, GRIB 1/2, HDF), and can be used to manipulate and analyze climate and numerical weather prediction datasets.

Domain experts emphasized the importance of interactive visual analysis coupled with automated statistical and machine learning

techniques in facilitating discovery of trends, patterns, and anomalies in multidimensional datasets. This approach can also enable tasks such as model building, validating existing models, model refinement by attribute space exploration, correlative analysis of multiple variables, clustering, and analyzing simulations. Experts feel that it is inevitable to have these visual analytics features, considering the scale and complexity of ocean and atmospheric datasets. Some of these experts also mentioned that having a toolbox or collection of 2D and 3D visualizations with simple and flexible integration can help in their routine tasks. One of the experts expressed a desire to have 4D data visualization (i.e., two data dimensions, one dimension for time, and one for ensembles). When asked whether having a generic visualization tool that provides lots of toolboxes would be better or if they prefer a custom designed visual analytics tool tailored to solve a particular research problem, most of the analysts preferred having custom tools designed for solving their problems and they should adapt in their data processing pipeline. One of the experts gave the following comment about interaction support in visualization tools for ocean and atmospheric datasets:

*"A visualization tool with a variety of interactive exploration features would be good where you can customize different aspects of the visualizations if needed (a default setting would be good for a quick overview)"*

Furthermore, when the experts were asked about the use of big data tools in their analysis tasks and how well these tools integrate with the visualization tools they intend to use, they mentioned that they face lots of challenges and limitations in such scenarios. They particularly pointed out that they require efficient data access and loading, flexible interfaces that support a variety of data formats and structures, more desirable statistical analysis features, query-based interfaces for multi-dimensional data retrieval and visualization, flexible visual encoding and visualization options, capability to provide an overview and details-on-demand, and support for interactive big-data visualization with interactive querying and filtering. Comments from one of the domain experts are mentioned below:

	JBS08	KLW08	WY08	JG09	LS09	SFJK09	SSJKF09	SWMM09	WS09	AAB10	DIANS10	LSL10	GGA11	LLEJ11	PH11	GXY12	LLBP12	DNN13	GYHZ13	DDW14	KSU14	PDW14b	BJA15	CZC15	DKG15	HMRH15	JDKW15	MN15	WFZ15	BGO16	DPD16	HDS16	BLLS17	FKRW17	WJWP17	WLSL17	ZYL17	KBL18	KTB18	ME18	NR18	VAN18	YIT18	ZW18	BMLC19	Sum	
ML & STATS																																															
Dimension Reduction																																															12
Classification																																															9
Clustering																																															31
Regression/Correlation Analysis																																															9
Others																																															5

**Table 4:** High-level Categories of machine learning and statistics techniques used for coding papers. We classify techniques into 4 categories— Dimensionality Reduction, Classification, Clustering, Regression/Correlation Analysis, and others. For atmospheric analysis, SOM (self-organizing maps) was used for both clustering and dimension-projection.

"The visualization tool is expected to have fast access to the data. If possible we will only need to provide an interface between the visualization tool and the data with different format/structures"

As a consequence of these limitations, the experts often use a combination of different tools to accomplish their tasks. In most scenarios, they have limited visualization choices offering little flexibility, especially considering the scale and complexity of the data. These scientists and analysts working on big datasets often don't have the luxury of utilizing an integrated visual analytics environment to accomplish their tasks. Instead, they visualize the output data produced at different stages of their workflows. In some cases, visualization is performed as a post-process, after they are done with their analysis tasks.

It is therefore necessary to understand the key challenges and requirements to enable extensive use of visual analytics tools by experts in this domain. Thus, it is essential to look into visual analytics research in the domain of ocean and atmospheric sciences based on such taxonomies that may uncover these limitations and challenges, and may serve as a guide for future research opportunities. Our interviews were structured in such a way that we can not only understand their task requirements, but also gain some insights on associated challenges.

Based on these interviews, we summarize detailed task requirements for visual analytics applications from a perspective of a domain expert in the field of ocean and atmospheric sciences.

#### Data requirements

- R1** Support data compression, storage, and efficient retrieval of the data for visualization tasks from high-performance clusters, supercomputers.
- R2** Query-based retrieval, analysis, and visualization across multiple data dimensions and for multiple data attributes.
- R3** Features for data pre-processing, cleaning, transformation, data reshaping, filtering, scale management, and support of native formats.
- R4** Superimpose and analyze data from disparate sources accommodating differences in resolutions and scale. Support for different types of grids.

#### Interactions needed

- R5** Interactive data analysis and exploration with supporting features like overview and details-on-demand, presenting data at

different abstraction levels, multiple linked visualizations, customizable visual encoding options for viewing data from different perspectives, and overlaying contextual information.

#### Features/tasks required

- R6** Identification of trends, patterns, and anomalies in the data.
- R7** Supporting correlative analysis, comparison, clustering, aggregation, and classification.
- R8** Predictive analytics, forecasting, and model validation.
- R9** Visual analysis of ensembles and simulations.

#### Visual Analytics Tool Design

- R10** Custom designed visualizations and visual analytics solutions for a variety of problem scenarios, and customizable tools for presenting results.

These task requirements will not only help us understand the requirements of domain experts but it will also help us understand whether the current research direction in the field of visual analytics related to ocean and atmospheric datasets is in line with the task requirements of ocean and atmospheric sciences, and how successfully it is addressing the needs of this domain. In the next sections, we will look into existing visual analytics models and frameworks, and provide details of survey results.

#### 6. Taxonomies

To systematically review visualization research in ocean and atmospheric sciences, we identified 10 different taxonomies to categorize and code the final sample set of papers collected, as explained in 4. We developed this set of taxonomies after observing the workflows of domain experts working on ocean and atmospheric datasets, analyzing different usage scenarios of visual analytics and visualization on different climatology dataset, receiving feedback from domain experts especially related to challenges in incorporating interactive visual analysis techniques in their workflows, and reviewing visualization literature.

The purpose of utilizing these taxonomies is to analyze different aspects of published research in an attempt to identify the gaps between the visualization requirements of ocean and atmospheric sciences and the focus of research of the visualization community related to this area. This would help address the challenges in adoption and usage of interactive visual analysis by the climate research community [TDN11]. The coding results are presented in the form



**Table 5:** *Papers coded based on a taxonomy of visualization techniques.*

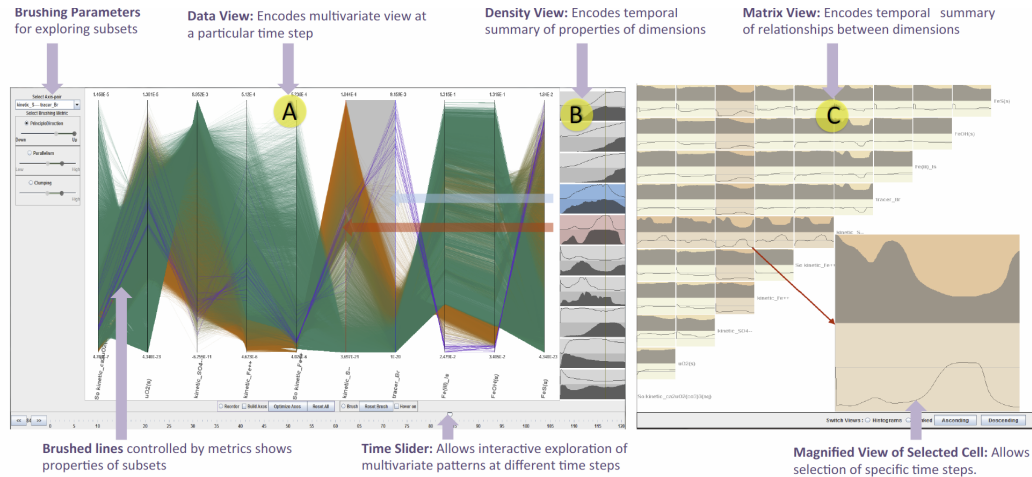
### 6.1. Tasks

Other data-related task categories such as ‘*data cleaning, pre-processing, transformations, finding inconsistencies*’(R3) [DNN13, VMN\*18, TWC\*18] and ‘*scale management from big data perspective*’(R3) [PSR17, GXY12, GYHZ13], ‘*simultaneous analysis of heterogeneous data / overlay*’(R4) [ME18] have limited coverage in the visualization research. Some of these data handling tasks overlap with big data research but they are also important to handle in the visualization domain with the objective to support interactive visual analysis tasks. The term ‘big

Other task categories such as ‘*automatic detection of patterns, trends, and anomalies*’ relates to **R6** [LPCR<sup>H</sup>18], ‘*prediction/forecasting future trends*’ corresponds to **R8** [WFZ<sup>\*</sup>15], and ‘*identifying relationships between different variables*’ partially corresponds to **R7** [WSL<sup>\*</sup>14]. The remaining categories do not have limited coverage in the visualization research but some of them do not correspond to any task requirements of domain experts.

Data types refer to features such as attributes, dimensions, temporal and hierarchical information, etc., based on which we can taxonomize datasets. Ben Shneiderman in 1996 in his famous paper “The Eyes Have It”, proposed a data-type taxonomy with seven types (1-2-3 dimensional, temporal, multidimensional, tree, and network data) [Shn96]. We use this taxonomy to code our surveyed papers (Table 2). We have combined 1-dimensional, tree, and network data types into ‘Others’ category as there were only few papers in these categories. Below we describe this taxonomy in detail. Also, ocean and atmospheric datasets are generally a combination of these data types.

**1-dimensional:** This data type includes scalar data generally organized in a sequential layout. Text data, programming source codes, list of items, etc., are common examples of this data type. Each item in this data type is a line of text constituted of character strings. The main challenge to visualize this data is the design decisions to choose fonts, color, size, etc., along with interaction methods, such as overview, zoom, scroll, etc., based on size of the



**Figure 1:** VIMTEX interface with *Select*, *Explore*, *Reconfigure*, *Abstract*, *Filter*, and *Connect* interactions. [DKG15]

data. Also, this data type is generally unstructured so designing visualization for it is difficult. For ocean and atmospheric datasets, we did not find any example belonging to this type.

**2-dimensional:** This data type includes planar/map data such as geographic maps, floorplans, layouts, etc. Each data entity in this data type has attributes such as name, type, value, etc. Ocean and atmospheric datasets often have this data type representing spatial information [SGL\*16, GSFT16, BPR\*15]. The spatial data is presented on the map as a 2D layer and often visual analytics systems have multiple layers that users can turn on and off based on their requirements.

**3-dimensional:** Daily life items with volume info such as human body, car, buildings, etc., are examples of this data type. The data might also contain relationship information among data items. Various weather and climate models often generate 3D data, such as ensemble data [LLBP12, PWB\*09], hurricane movement [MGA\*08], tornado structures [WPB\*11], etc. Volume rendering (Figure 6) is a common visualization technique for this type of dataset [LLBP12]. In general, this data type is commonly present in ocean and atmospheric datasets.

**Temporal:** Data that varies with time is temporal data. Examples are temperature variations, ocean waves height over time, wind magnitude varying with time, etc. This is one of the common data types often present in ocean and atmospheric datasets that are often temporal, multivariate, and multidimensional datasets. The temporal component is one of the essential parts of oceanic and atmospheric datasets and is generally visualized in the form of 2D charts [DKG15, WS09, MHR\*11], calendar visualization [MMME11], animation [FKRW17], timesliders [DKG15], glyphs [DNN13], etc.

**Multi-dimensional:** Data with multiple dimensions and attributes is associated with this data type. Multivariate data also belongs to this data type. Multidimensional data has multiple dimensions whereas multivariate data has various attributes associated with each dimension or item. Examples include databases, weather and climate simulation data, social network data, etc. Oceanic and

atmospheric data are often multidimensional and multivariate, and thus belong to this data type. Our interviews with domain scientists (Section 5) also show that experts mostly work with multidimensional and multivariate simulation data (Task Requirements R2 and R9).

Multidimensional data is commonly visualized on 2D displays with either data being summarized with fewer dimensions (please refer to section 6.4 for more details) [FBW16, KFH10, SWMW09], using multiple views for different attributes/dimensions [LTJ10, BBP08, LLEJ11], or with parallel coordinates [WLSL17, CZC\*15]. Parallel coordinates is a commonly used visualization technique for multidimensional data (Figure 1 (A) and Figure 4).

**Tree:** They represent hierarchical data with all items having a link to a parent item (except the root) [Shn96]. These datasets are visualized using a tree visualization where the root is present at the top level, its children are on next level extending to the leaf nodes (which don't have any children). This data type is not common for ocean and atmospheric data as we did not find a single example of it in our surveyed papers (Table 2).

**Network:** This data type represents relationship data among items that cannot be captured with a tree datatype as items might have an arbitrary number of relationships with other items. Node-link diagram and matrix representation are common visualizations for the network data. Shu et al. [SGL\*16] derived network data from the time invariant ensemble simulation data. They use node-link diagram to graphically display the data with nodes depicting sub-regions having similar ensemble properties, and edges portraying spatial overlap in neighboring time frames. Poco et al. [PDW\*14b] use matrix representation for finding similarities in climate model structure and outputs. They also derived the network data from simulation models (Figure 5).

### 6.3. Interaction Methods

Interactions are one of the most important tools of the visual analytics system. They allow the user to ask the system questions and



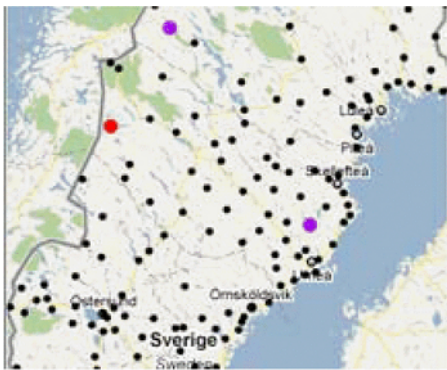
receive an answer. It is also clear from our discussions and interviews with domain scientists (Section 5) that interaction is an essential feature required in visual analytics tools (Task Requirement R5).

Yi et al. [YaKSJ07] provided the taxonomy of interaction by defining seven interaction categories. Lu et al. [LGH\*17], in their predictive visual analytics survey, not only used this interaction taxonomy but also added a Shepherd technique that helps users to tune the modeling process. We also coded our surveyed papers (Table 3) based on the interaction techniques taxonomy used by Lu et al. [LGH\*17]. Below we discuss these interaction techniques with examples.

### 6.3.1. Select

Select interaction allows users to mark particular items of interest so that they remain easily visible to them among other displayed items. This interaction is important when there are lots of items on the viewport and users are interested in a few particular items, and this way they can select those items and keep track of them easily.

Generally, select is the most common interaction technique present in many visual analytics systems designed for ocean and atmospheric datasets. Lundblad et al. [LEH09] used select interaction to allow their users to select interesting ship movement routes. Tao et al. [TWC\*18] used selection to highlight critical nodes in their semantic flow graphs in the form of halos. Lundblad et al. [LLEJ11] enabled users to select a particular weather data station in their visual analytic system to explore multiple weather station data (Figure 2). The multiple coordinated view system would then update other views according to selected data items. This shows that selection is often coupled with other interactions. Yi et al. [YaKSJ07] also mentioned that select is a basic interaction that is generally followed by other interactions such as filter, connect, or abstract. This helps users in the data exploration process and to find insights in the data.



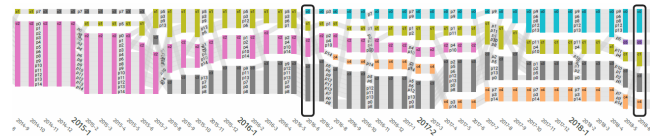
**Figure 2:** Selection interaction example where the user is selecting particular weather stations. [LLEJ11]

### 6.3.2. Explore

Data available for visualization is often large and complex, such that it cannot be visualized on the viewport or the user's display screen. Explore interaction allows users to analyze items which are

currently not viewable and changes the subset of the data to be displayed on the viewport. The most common example of explore is panning, which allows users to change the scene to bring other items into view.

Ocean and atmospheric datasets are generally very large and complex so this interaction is often present in the visual analytics system designed for such datasets [BJA\*15]. Höllt et al. [HMZ\*14] used explore to allow users to pan and analyze their simulated 3D ocean surface in detail. Ma and Entezari [ME18] utilized select and explore to interactively analyze the uncertainty in subsets of ensemble isocontours for ensemble forecasting. Kappe et al. [KBL18] used scrolling to change the timeline—to bring another subset of the data in the view—in analyzing decadal climate predictions (Figure 3).



**Figure 3:** Explore interaction example where timeline can be scrolled to bring other items in the view. Also, couple of items are selected in the timeline using select interaction. [KBL18]

### 6.3.3. Reconfigure

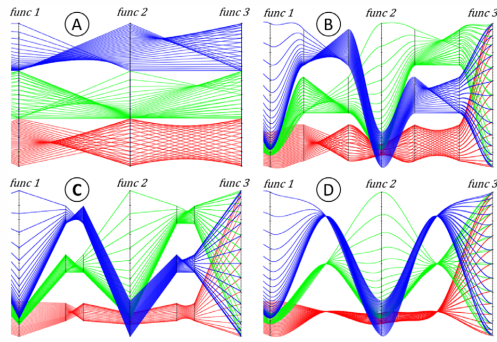
Reconfigure interaction allows users to change the spatial arrangement of the visualization. Sorting, realigning, rearranging, etc. are common types of this interaction. This interaction allows users to see different perspectives of the dataset that might not be available in a given perspective to gain valuable insights. A simple example of this interaction is reordering the parallel coordinates plot axis or changing scatterplot axis attributes.

Ocean and atmospheric datasets are often multivariate and this interaction helps to find patterns and correlations among various attributes present in the data. VIMTEX [DKG15] is a visualization interface comprised of parallel coordinates and two summary views for time-varying multivariate data exploration. Users can reconfigure parallel coordinate axes to reveal different patterns. Users can also select various items of interest and can explore different subsets of data (Figure 1).

### 6.3.4. Encode

Encode interaction enables users to change the representation of the data by changing various visual attributes such as color, size, and shape. With this interaction users can pick and choose different attributes of the multivariate dataset to be encoded by different visual attributes in the visualization. It also allows the user to change the visual representation. For example, a node link representation of the network data could be viewed as a matrix representation to get different perspectives of the data. Overall, this interaction helps users to find insights in the data by analyzing the data using various representations.

Biswas et al. [BLLS17] presented a visualization of time-varying weather ensemble data and allowed users to change the encoding



**Figure 4:** An example of encode interaction in nested parallel coordinate plots. Users can change different function parameters for each axis to update the visual representation. [WLSL17]

by choosing various visual representations, aggregation levels, and color scales. Wang et al. [WLSL17] also allow users to change the encoding of visual representation in their proposed nested parallel coordinates visualization (NPCP) for analyzing parameter correlations in climate ensembles (Figure 4).

### 6.3.5. Abstract/Elaborate

This interaction enables users to see the visual representation at different scales. For example, users can see the overview of the whole representation or can scale down to see details, or can look in between levels as well. Zooming and showing details on the cursor tooltip by hovering over an item of interest are common types of abstract/elaborate interactions, often present in various visual analytics system.

Because ocean and atmospheric datasets are generally complex and large, there are therefore usually various aggregations and summarizations needed to be performed to represent the data. This interaction technique is helpful in such cases as it enables users to see different levels of details as required. A good example of abstract/elaborate is presented by Wang et al. [WLSL17]. They presented an overview+detail technique to analyze spatio-temporal climate ensembles at different resolutions. Figure 1 also shows an example of abstract/elaborate in VIMTEX interface, where users can select a temporal summary cell to see its details [DKG15].

### 6.3.6. Filter

Filter interaction allows users to see a subset of items based on specific conditions. In this interaction, users generally specify the range or condition to see a subset of data meeting that criteria. Filter interaction only hides the data not satisfying the given condition and once the condition is removed the whole data is once again available to be presented. Search and query are common types of filtering interactions.

For complex and multivariate oceanic and atmospheric datasets, this interaction is very useful as it helps users to see only the data of interest [GGA\*11, GHA\*08]. Lundblad et al. [LEH09] used filter sliders to enable users to hide uninteresting ship voyages. Wang

et al. [WLSL17] in their nested parallel coordinate plots representation (Figure 4) also allowed users to specify the range of values for each axis by displaying a range slider on each axis. VIMTEX interface also enabled users to filter the data according to different parameters [DKG15].

### 6.3.7. Connect

Connect interaction allows users to see associations and relationships between data items or to show hidden items relevant to specific displayed items. This interaction technique can apply to single view visualizations—for example, to highlight neighbors of selected node in a node link diagram—or to multiple view visualizations where a selected item is highlighted across multiple views and this connect interaction technique is called brushing.

An ocean and atmospheric data analytics system usually comprised of multiple views as the data to be visualized is often temporal, multidimensional, and multivariate [LSL\*10, USKD12]. The VIMTEX interface [DKG15] shown in Figure 1 has a brushing technique to highlight a selected item in other views. Blaas et al. [BBP08] also used connect interaction technique in their multiple view visualization for multi-time point volumetric dataset (for example, hurricane data) to highlight an item selected in one view in other views.

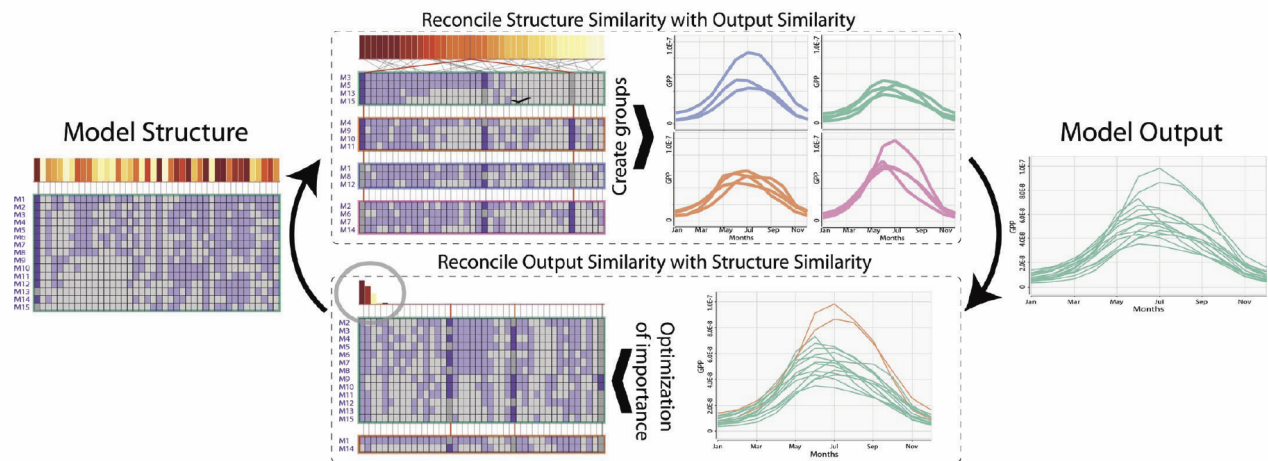
### 6.3.8. Shepherd

The Shepherd [LGH\*17] interaction technique enables users to tune the modelling process either directly or indirectly. Direct shepherding allows users to choose model parameters settings such as updating number of clusters in k-means. Indirect shepherding includes specifying constraints such as defining threshold from the interface. This technique is not that common in visual analytic systems for ocean and atmospheric datasets but could be useful. This particular type of interaction is unique because it is an interaction with the model and it also partially overlaps with the machine learning taxonomy. However, it was more appropriate to put it here as it involves directing the model through user interaction [LGH\*17]. With this interaction technique, users can tune the model parameters according to their needs and can see different perspectives of the data.

We only found a few papers from our surveyed papers using this interaction technique. Ferstl [FKRW17] allowed users to tune the clustering parameters in analyzing the temporal uncertainty in weather forecast ensemble visualization. Poco et al. [PDW\*14b] defined a visual reconciliation paradigm that is an iterative human in a loop strategy for understanding climate model structure similarity and model output similarity (Figure 5). Valsangkar et al. [VMN\*18] also allowed users to tune the parameters for cyclone identification and tracking.

## 6.4. Statistics and Machine Learning Methods

Initially, we adopted a taxonomy of machine learning algorithms adapted from Patgiri [Pat18] but after completing the coding of papers based on this taxonomy, we were left with a very sparse table; there was especially very limited work related to deep learning



**Figure 5:** An example of Shepherd interaction to enable the user to refine the groups and explore different parameters in climate model similarity analysis. [PDW\*14b]

in general. Thus, we decided to merge them into five major categories: dimensionality reduction, classification, clustering, regression/correlation, and Others. Table 4 shows the coding results of this revised taxonomy table. In this table, we have included coding for those papers for which we can easily identify which machine learning or statistics technique was used. That is the reason this table has fewer entries as compared to other taxonomy coding tables.

#### 6.4.1. Clustering

Clustering is known as a prediction task; many applications used clustering in case there was no available labeled data [WFZ\*15]. The models created by clustering cannot be generalized, so, selecting a suitable similarity metric and validation are considered as the main challenges in clustering. Clustering configuration, evaluation, and exploration is a common task used in visual analytics (Task Requirements R5 and R7).

Ma and Entezar [ME18] presented an interactive framework for visualizing uncertainty of ensemble isocontours. They used high-density clustering algorithms and a bandwidth selection method for generating better density function for ensemble representation. They also allow users to interact with the representation and select specific regions of interest. The system then automatically recalculates the clustering and updates the results. Jänicke et al. [JBMS09] used clustering to group similar climate change regions. Köthür et al. [KSU\*14] utilized hierarchical clustering to aggregate the spatial regions associated with time series into a hierarchy for generating visual summaries.

The efficacy of the obtained results by a clustering method depends on the application; this leads to the development of a variety of clustering methods. For atmospheric data visualization, Self Organizing Map(SOM) [AAB\*10] and K-means [WYM08] are frequently used. This was observed from the detailed coding results for clustering subtype in machine learning and statistics taxonomy tables.

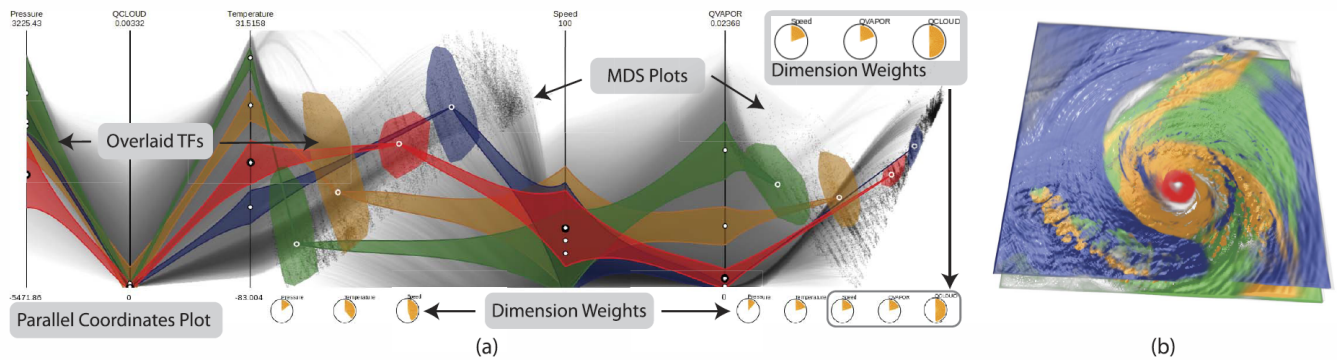
In k-means clustering, the first step is the partitioning of the input data into k initial sets randomly or by using heuristic data. The second step is to find the centroid of each set. A new partition is created by associating each point with its closest centroid. The centroids are regenerated for the new clusters, and the algorithm repeats until some convergence condition is satisfied [WS09, SWMW09, WYM08]. For example, In [WYM08], k-means clustering algorithm was used to cluster importance curves. They presented an importance-driven, time-varying data visualization approach that highlights the essential aspects of such datasets. This approach is based on blockwise-analysis where an importance driven curve is derived for each block based on conditional entropy from information theory that characterizes the local temperature behavior. These importance curves are clustered based on k-means (Lloyd's algorithm), and these clusters then classify underlying data. This can be used to identify any temporal trends in these clusters. They support features like cluster highlighting, abnormality detection, time-budget allocation, and time step selection.

Self Organizing Map (SOM) SOM is an unsupervised learning technique which has expanded in popularity. Firstly, SOM produces topological clusters in a condition where similar clusters are neighboring. Then, SOM uses a projection technique by projecting multidimensional data into a two-dimensional data. For example, in [AAB\*10], a framework is presented based on SOM, combination of clustering and dimensionality reduction, and interactive visual analysis for spatiotemporal datasets. Multiple linked views including a couple of SOM matrix views (one for showing spatial distribution and other for temporal variations), a time-series graph view, and periodic pattern view are presented. (Task Requirement R7)

#### 6.4.2. Dimensionality Reduction

Visualization of high-dimensional data is considered a challenging task, so reducing data dimensionality is a usual scenario to try [JBS08] (Task Requirement R7). In [KFH10], the authors have shown that combining interactive visual analysis of multidimen-





**Figure 6:** Volume visualization of hurricane Isabel multivariate data. (a) Transfer function specification with multidimensional scaling plots embedded in parallel coordinates in the interface for feature selections. (b) Volume rendering visualization based on transfer function specified in (a). [GXY12]

sional data in the form of multiple coordinated views (with brushing and linking) along with statistical techniques can be effective for analysis. They integrate outliers detection methods and four statistical moments (mean, variance, skewness, and kurtosis) in the visual analysis process. A couple of detailed climate studies are presented to show the usefulness of the presented technique. Also, in [FKRW17], the authors presented a visualization method called ‘streamline variability plots’ that can be used to visualize the vector field ensembles passing through certain locations. It takes a set of streamlines and uses PCA to transform this set of streamlines to a low-dimensional Euclidean space. It then clusters those streamlines to identify trends that are approximated by Multivariate Gaussian Distribution, which yields a probabilistic mixture model for the streamline distribution.

Cluster analysis of ensemble weather forecast data is a commonly used technique in meteorology for analyzing forecast uncertainty [KRRW19]. However, the sensitivity or confidence of the clustering results is often not known. In [KTB\*18], an interactive visual interface to analyze the sensitivity of clustering results is presented. They provide multiple views to help users to access a variation in the composition of clusters, find variability in cluster groups, and analyze the uncertainty in spatial areas of the trends. A real-world example of Cyclone Karl is provided to show the effectiveness of the technique.

Multivariate clustering is commonly used in geovisualization applications. The results of the clustering are often displayed using choropleth maps with cluster labels projected on them. However, this representation can be greatly affected by variation in the data or clustering parameters [ZLMM16]. Zhang et al. [ZLMM16] present a geovisual analytics framework to help users understand the variations in visualizations with geographical variations for multivariate data clustering. Various visualization views (scatter-plot, area profiler, rose plot, etc.) are provided.

#### 6.4.3. Classification

Classification is a task of predicting the value of a categorical variable. This is achieved by building a model; this model learns exemplars using the data characteristics from the training set. Then

the model can be applied to unknown examples to predict their classes [AAB\*10, DIANS10, BGOJ16] (Task Requirement R7 and R8). For example, authors in [PDW\*14a] presented a classification of the specific domain intents of climate scientists, and underlying data facets, and then bridged the intents and facets with the visualizations tasks and designs through a classification scheme, and finally presented a tool called ‘SimilarityExplorer’ that implements this classification scheme. This tool also contains coordinated multiple views (CMV) that presents two case-studies and evaluations by three climate scientists.

Another approach in [CWMW11] is focused on the correlation analysis in time-varying multivariate datasets. It creates a static volume classification by utilizing a sampling-based approach for classifying correlation patterns (with respect to various reference locations). The overall process has three steps: 1) Selecting important samples from the volume; 2) priority distance computation for sample pairs; 3) approximate volume-based correlation with sample-based correlation.

#### 6.4.4. Regression

One of the significant challenges in data analysis is to identify the multivariate relationship within data variables, especially when the numbers of the data variables are large. Visual analytic approaches have demonstrated good potential to evaluate to what degree observed relations can be attributed to underlying events and to build causal analyses (Task Requirement R2, R6, R7 and R8). In [SSJKF09], the authors augmented the interactive multivariate representations with statistical model information to better understand the weather data. Regression and correlation analysis are integrated with parallel coordinate plots to design a new technique called multidimensional data explorer (MDX).

#### 6.5. Visualization Techniques

There are various visualization techniques available to visualize ocean and atmospheric datasets. Ocean and atmospheric datasets are generally multidimensional, multivariate, and time-varying. Therefore, they often require multiple visualization techniques for

sta

ove  
sigh  
ot to  
ion  
mor

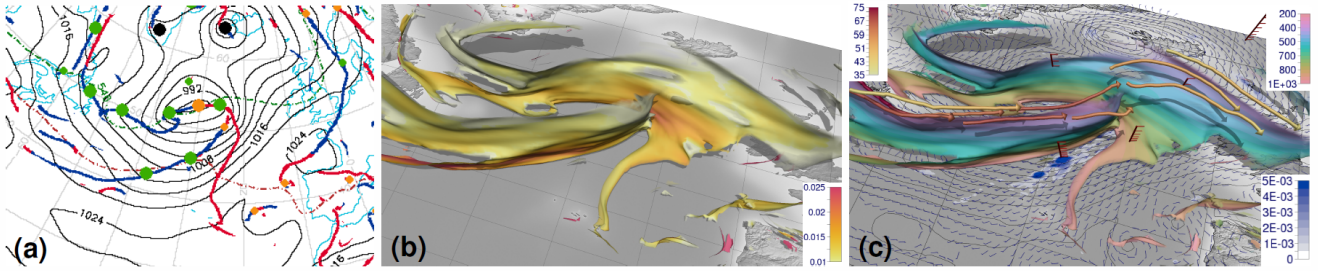
qual

death  
organizat  
ul. [br  
ation  
to u  
on i  
vis  
data

al. [10] are li

users  
i02  
of  
an  
s by  
s. A





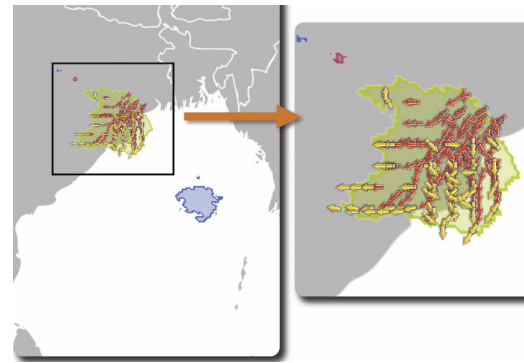
**Figure 7:** 3D atmospheric fronts visualization of cyclone “Vladiana”. (a) shows identified fronts. (b) Visualization of 3D fronts. (c) 3D fronts along with other meteorology features. [KHS\*18b]

Atmospheric data is often multidimensional and multivariate, parallel coordinates is one of the most popular techniques used for visualizing such data. Steed et al. [SFJK\*09], Poco et al. [PDW\*14a], and Jin and Guo [JG09] presented an interactive parallel coordinates based technique for multivariate climate datasets. There are examples where parallel coordinates technique have been extended to multivariate volume datasets. Blaas et al. [BBP08] presented a parallel coordinate based visualization for multi time point volume data. Guo et al. [GXY12] also presented a parallel coordinate and multidimensional scaling plots technique for multivariate volume data visualization (Figure 6).

Scatterplot displays data as points—often encoded in terms of color, shape, size, etc., to represent different data attributes—in 2D cartesian space [LZM14]. Qu et al. [QCX\*07] used scatterplot to show relationships in neighboring parallel coordinate axes while visualizing air pollution data. Bachthaler and Weiskopf [BW08] extended the scatterplot technique as “continuous scatterplots” for spatially continuous input data. They applied the technique on tornado dataset—with a 3D airflow velocity field—to show its efficacy. There are various other proposed techniques of this type, mainly derived from parallel coordinates and scatterplot, being used in visual analytics systems.

### 6.5.3. Iconic Displays

Iconic displays is another visualization technique that maps various attributes of multidimensional data to features such as color, shape of an icon, or glyph. Glyphs are commonly used in weather and ocean data displays for showing wind speed and direction [WP13], and for displaying ocean currents [RSAS17]. Pilar and Ware [PW13] evaluated various glyphs and streamlines and showed that integrating glyphs along with streamlines is useful in understanding patterns in flow visualization. Jarema et al. [JDKW15] presented a glyph based visualization technique to visualize directional probability density function of vector field ensemble data. Doraiswamy et al. [DNN13] used arrows as glyphs on the map to show cloud movement (Figure 8). The color of the arrows is used to show past or present trajectory movement. Maskey and Newman [MN15] proposed a texture based glyph visualization for multivariate weather data visualization. In summary, these techniques are useful for ocean and atmospheric datasets. However, choosing the right glyph for a given attribute is important and can help users to better perceive the visualization.



**Figure 8:** Visualization of clouds movement using arrows as glyphs. Red arrow represents past trajectory and yellow arrow depicts future trajectory. [DNN13]

### 6.5.4. Dense Pixel Displays

Dense pixel display map a data attribute value to a colored pixel on a display. As this technique uses one pixel per data value, it therefore allows the largest amount of data possible to be displayed on the viewport. The main difference in various types of dense pixel displays is generally the arrangement of pixels. Pixels could be arranged in different layouts to help find patterns and correlations in the data. Common examples of different layouts are grid-based layout (the most popular one) [AAB\*10] and radial/circular layout [LZM14].

Andrienko et al. [AAB\*10] used color-coded pixels in the SOM matrix for displaying temporal variations in multivariate spatiotemporal data. Poco et al. [PDW\*14b] utilized color-coded pixels for displaying matrix for performing visual reconciliation of grouping in comparing climate models (Figure 5). Li et al. [LZM14] presented a pixel-based global radial map to analyze changes in climate states in temporal multidimensional climate data (Figure 9).

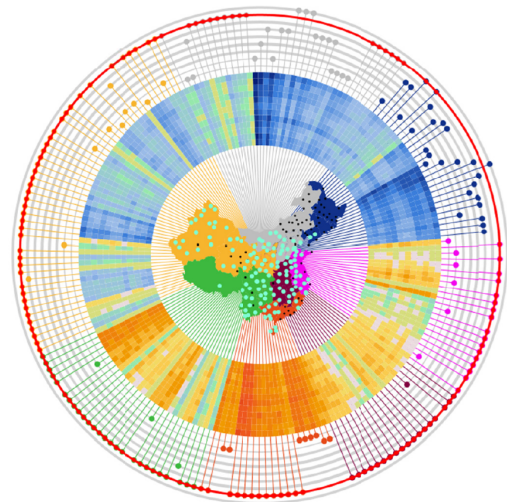
### 6.5.5. Stacked Display

Stacked displays are designed to present hierarchical partitioned data. For multidimensional data, it is important to choose appropriate dimension for partitioning and hierarchy building [Kei02]. Treemaps are a common example of this technique that visu-

alize hierarchy information of multidimensional data. Zhou et al. [ZYL\*17] used treemap in the multiple coordinated view visualization system for air quality data to analyze hierarchy details of specified spatial cluster.

## 6.6. Evaluation Methods

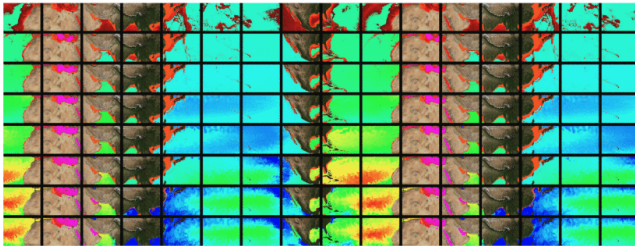
Lam et al. [LBI\*12] argued that our approach to evaluate papers should be based on evaluation scenarios instead of evaluation methods, and they propose two major categorizations of evaluation scenarios: (1) understanding data analysis; (2) understanding visualizations. Understanding data analysis is further categorized into the



following: understanding environments and work practises (UWP); evaluating visual data analysis and reasoning (VDAR); evaluating communication through visualization (CTV); and evaluating collaborative data analysis (CDA). Isenberg et al. [IIC\*13] proposed another evaluation category named the qualitative result inspection (QRI) after extending Lam et al.’s scenario based approach to systematically analyze work published in the IEEE Visualization conference.

Understanding environments and work practises (UWP) is cen-

tered on gathering design requirements for visualization tools and the set of features that should be supported (Task Requirement **R10**). There can be two possible variations of evaluations conducted in UWP: The first is observing the workflows of people without using the tool similar, and the second is observing the workflows while using the tool to understand the adoption and identify features for future revisions. Ellsworth et al. [EHN17] observed domain experts while they worked with the data to identify the features required to visualize Petabyte scale ocean simulation data sets. They designed a hyperwall as shown in Figure 10, comprising 245 million pixels operating at near native resolution.



**Figure 10:** The hyperwall showing the salt concentration levels [EHN17]

Visual data analysis and reasoning (VDAR) evaluation scenario assesses the success of a visualization tool in visual analysis, reasoning, and hypothesis development/testing, based on insights and knowledge gained with the use of the tool. Biswas et al. [BLLS17] places the domain expert in the ensemble data exploration loop and adjusts the output based on the requirements of the domain expert.

Communication through visualization (CTV) explores whether communication (conveying desired information/message to a target audience) is supported in the visualization tool and how this support is provided. Different metrics such as interaction scenarios, information retention, accuracy, learning rate [LBI\*12] can be used to quantify communication quality.

A collaborative data analysis (CDA) evaluation scenario is concerned with how visualization tools support performing certain tasks (taskwork) through collaborative analysis and decision making (teamwork) [PG02, LBI\*12]. This evaluation scenario also questions whether certain features of the system are utilized during collaborative task analysis, and what type of social exchange or communication features are supported.

In User Performance (UP), objectively measurable metrics (e.g., time, error rate, work quality) are calculated for visualizations and can be used to determine associated human visual perception and cognition factors, and compare user performance across visualizations. A User Performance (UP) evaluation method is utilized in these works [LPCR18, LBR\*17, JDKW15, KLM\*08].

User Evaluation (UE) is based on subjective user response on whether a specific visualization supports the desired tasks, and the user response or opinion may be collected in different forms such as written, spoken, or body sensors. Tao et al. [TWC\*18] conducted a user evaluation with a domain expert. They conducted three sessions in order to first familiarize the user with the tool, then gave

the user some test tasks to complete a moderately complex dataset, and then was provided with a complete dataset with full featured visualization support.

Algorithm Performance (AP) is based on a quantitative evaluation of the output of the visualization algorithm measure usually in terms of performance (e.g., computational efficiency) or visualization quality (e.g., image quality measures) (Task Requirement **R1**). Algorithm Performance (AP) is often utilized in scenarios where a computationally expensive algorithm or technique is involved in the visualization generation process [RSG\*13, NR18, ME18, WSL\*14].

Qualitative Result Inspection (QRI) is based on the qualitative analysis and discussion on the resultant visualization and then the viewer is asked to make an assessment on his own. As compared to some other evaluation scenarios, no participants or actual end users are involved in the evaluation. This evaluation method is the most commonly used evaluation scenario, as shown in the Table 6 coding results.

## 6.7. Dimensions, Size, and Scale

In recent years, the increase in data size, dimensions, and scale is a continuous challenge for the visualization community. In this section, we broadly classified the taxonomy into three groups: 1) models/spatial scale, 2) number of dimensions and 3) data scale. This taxonomy is derived from our interviews with domain scientists and form the gathered tasks requirements. Table 7 summarized the results of coded papers for these classifications. A paper is present in the table if it can be coded according to at least one of the taxonomies.

### 6.7.1. Spatial Scale of Models or Data

Oceanic and atmospheric data are usually the output of simulation models; therefore they encompass a wide range of spatial scales ranging from micro to global. The classification of the spatial scale of models or data (micro, meso, synoptic, and global) for our taxonomy is adopted from Rautenhaus et al. [RBS\*17]. Visualization requirements for choosing the appropriate spatial scale depend on the purpose of the analysis; for example, forecasting and ensembles are mainly concerned with the meso and synoptic scales [HMC\*13].

Most of the visualization framework in ocean and atmospheric data domain also featured the meso and/or synoptic spatial scale of models or data [WPH\*12]. The visualization framework Ensemblegraph [SGL\*16] visualized the analysis of spatiotemporal behaviors in ensemble simulation data at the meso-scale. Nooodle [SZD\*10] integrated a numerical weather model at the meso-scale to visualize ensemble uncertainty and Travis [HMRH15] visualized microclimate research data also at the meso-scale. On the other hand, VisMAT [LZM14], Curve Boxplot [MWK14], and Albero [DPD\*16] visualized synoptic data. Similarityexplorer [PDW\*14a] deals with global data.

Almost, 40% of the domain experts we interviewed work with micro-scale data or models. However, in the coded papers no visualization tool is found for visualizing micro-scale models and data. This is contradictory because domain experts whom we interviewed mentioned that they have extensively used micro-scale



	QCX07	BBP08	BVM08	GHA08	JB508	KLM08	MGA08	WYH08	JBMS09	JO9	LEH09	LS09	MGKH09	PMB09	SFK09	SSJKF09	SVMM09	WS09	AAB10	DIA10	LSL10	LTJ10	SZD10	GGA11	LLE11	MHR11	MMME11	PH11	WHP11	WPB11	GXY12	LLBP12	USKD12	VLBA12	WPH12	DNH13	GYHZ13	HMC13	HRB13	PW13	RSG13	WBD13	VMK13	WP13	XYM13	JS14	LZM14	PDW14a	PDW14b	Sum								
Application Areas / Disciplines																																																										13
Weather Forecasting																																																										4
Operational Forecasting																																																										4
Ensemble Forecasting and Parameter Analysis																																																										9
Climate study Atmospheric Research/Meteorology																																																										40
Large Scale Data Simulation/Visualization																																																										9
Exploring Large Multivariate Data																																																										12
Others																																																										12
	DDW14	HMZ14	KSU14	MWK14	WSL14	BJA15	BPR15	CZC15	DKG15	DPD15	EGG15	HMR15	JDKW15	LMK15	MN15	WFZ15	BGO16	DPD16	FKRW16	GHP16	GSFT16	HDS16	QMI16	RSM16	SGL16	WPS16	BLS17	EHM17	FKRW17	LBR17	PSR17	RSAS17	WJW17	WLSL17	ZYL17	KEL18	KHS18a	KHS18b	KRRW18	KTB18	LPCR18	ME18	MPBM18	NK18	TWC18	VNM18	YIT18	ZW18	BMLC19	Sum								
Application Areas / Disciplines																																																										21
Weather Forecasting																																																										0
Operational Forecasting																																																										0
Ensemble Forecasting and Parameter Analysis																																																										10
Climate study Atmospheric Research/Meteorology																																																										30
Large Scale Data Simulation/Visualization																																																										0
Exploring Large Multivariate Data																																																										9
Others																																																										7

Table 8: Papers coded based on ocean and atmospheric sciences application areas.

models along with another scale of models or data (task requirements **R4**). This represents a gap in the visualization domain and domain experts task requirement.

### 6.7.2. Number of Dimensions

The classification within this class is adopted from Sedlmair et al. [SHB\*14]. It classifies the dimensions into two broad categories: (1) multi-dimensional (1-10D, 11-25D, and 26-100D), and (2) complex objects. Different variables associated with multivariate data or uncertainties derived from ensembles are also considered as dimensions in our classification scheme. Furthermore, multi-dimensional data subtypes (1-10D, 11-25D, and 26-100D are considered mutually exclusive (i.e., 1-10D is not included in 11-25D, etc.) during the coding process. The complex objects are 2D/3D images, animations, performance graphs, etc. As most of the data in oceanic and atmospheric sciences is simulated and observational numerical data, nothing from our surveyed papers except Peng et al. [PSR17] falls in the complex category.

In the coded papers, we found that the number of dimensions rarely exceeded 10 dimensions. Domain expert interview responses also correlate with taxonomy results (task requirements **R2**).

### 6.7.3. Data Scale

We classified data scale for ocean and atmospheric data into four categories: MB or Smaller, GB, TB, and PB or greater. In majority of papers, the scale of data was indefinable, therefore the taxonomy for this category is sparse.

Among the coded papers, a large proportion of works deal with GB data where the domain experts we interviewed are mostly using TB data. About 70% of the domain scientists from our interviews are using TB and 10% are using PB or greater. We think that there are gaps in task requirements of domain experts and available visual analytics frameworks for handling data scale of TB, PB, or greater (task requirements **R1**, **R2**, **R3**).

## 6.8. Application Areas

Ocean and atmospheric datasets are encountered in numerous application areas. We included this taxonomy in the survey to under-

stand the distribution of visualization research efforts across different application areas. Initially, we chose more categories of application areas than the ones shown in Table 8, but for some categories the coding table was sparse so we merged those categories into the ‘Others’ category. It can be seen from the coding table (Table 8) that ‘climate study/atmospheric research/meteorology’ has the most number of publications.

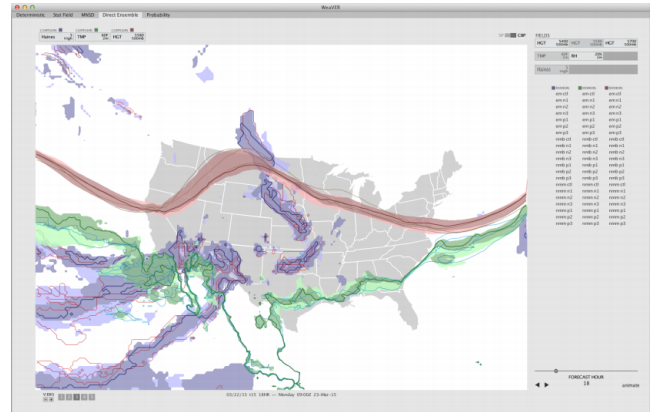
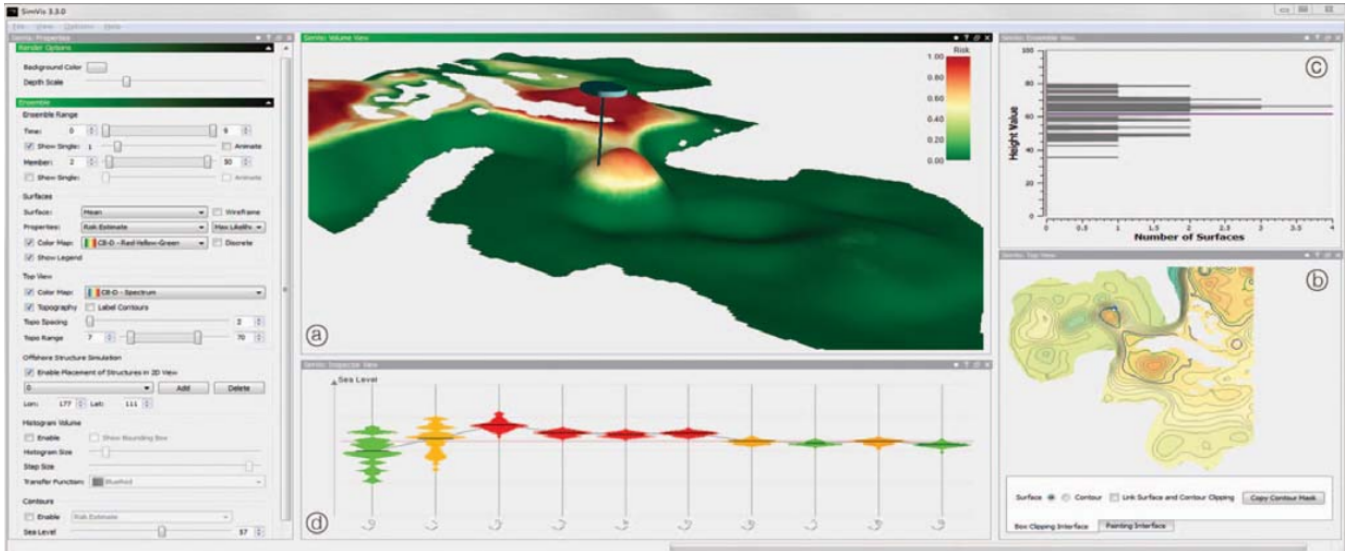


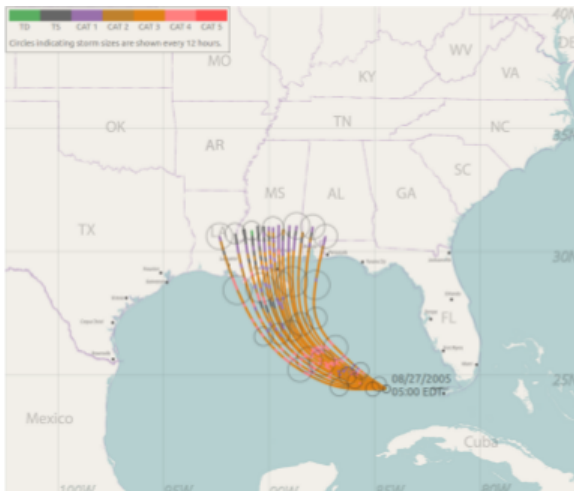
Figure 11: WeaVER visualization tool to study multiple geological features. [QM16]

Features supported by visual analytics environments such as brushing and linking can be used for hypothesis generation and testing for time-varying climate datasets [KLM\*08, RSM\*16]. Trends identified during analysis can be used to select regions that can act as climate change indicators. WeaVER [QM16], an open source visualization tool shown in Figure 11, allows people without visualization knowledge to generate visualizations based on visualization principles and existing meteorological conventions that can support comparative study of multiple meteorological features.

Höllt et al. [HMC\*13] presented a visualization framework for interactive exploration of ocean forecast ensemble data that allowed users to perform exploratory analysis, both in the spatial and temporal domain, to understand the prediction of ocean states in detail (Figure 12). This type of work has lots of applications in the oil



**Figure 12:** Linked views to explore ocean forecast ensembles [HMC\*13]



**Figure 13:** Sampling original sample to produce a smaller representation of ensemble using annotated tracks [LPCRH18]

and gas industry. Liu et al. [LPCRH18] presented a visualization of ensemble forecasts by calculating the representative set of the uncertainty data and representing uncertainty by its spatial distribution. Figure 13 shows the results of this technique by selectively sampling the original ensemble and representing in the form of annotated tracks. Linked views between 3D volume rendering view and multi-charts visualization (consisting of bar and line charts) obtained by linearizing 3D data points from 3D ensembles are utilized by Demir et al. [DDW14] to explore 3D ensemble datasets, and supports features like coordinated brushing, linking, and interactive querying.

Interactive visual exploration of large multivariate ocean and at-

mospheric datasets is a desirable feature (R2, R7, and R5) to support many analysis tasks. There are many extensions to parallel coordinate plots [YIT\*18, ZW18, GXY12, SFJK\*09, LS09], that is a popular technique to analyze multivariate datasets.

We also collected various tools, API's, and programming languages used by the surveyed papers (Table 9). Some of the papers have not mentioned the tools used or are using custom designed visualization tools, so we skip those papers in this table. Overall, from the table, it is clear that researchers in the oceanic and atmospheric domain are mostly using tools such as Met.3D [RKSW15], Matlab, ParaView [AGL05], and GPU based programming languages, Cuda and OpenGL to render large data. There are also few works which are using custom designed tools [QM16, ME18, WLSL17] (not mentioned in the table). This table serves a nice overview of the tools and programming languages commonly used for data analysis in the ocean and atmospheric domain.

## 7. Visual Analytics Models and Framework

In this survey, we have examined the task requirements of domain experts working on ocean and atmospheric datasets, and also performed a systematic study by looking into the published research in the visualization community related to ocean and atmospheric sciences. We also need to ground this discussion into existing visual analytics frameworks and models for sense-making and knowledge generation. This can help us analyze which visual analytics model or framework is most suitable to capture and represent the task and system requirements for visual analysis of ocean and atmospheric datasets, and help us convey the survey results from a visual analytics models perspective.

Pirollo and Card [PC05] presented a sense-making model that captured the iterative process that an analyst may perform to get insights from the data. This model consists of a foraging (seeking relevant information) and a sense-making loop (iterative mental



Application Areas / Disciplines	Tools Used
Weather Forecasting	Met.3D[FKRW16, KHS*18a, KHS*18b, KRRW18, KTB*18], Matlab[LLEJ11, FKRW16], R [LPCR18], (Nektar++ software package/VTK/CGAL)[WMK13,MWK14], (NCL (NCAR Command Language)/Avizo/ParaView)[BPR*15], (Javascript/Jquery)[DPD*16, LTJ10], VISUS framework[WJW*17], (C++/Cuda/OpenGL)[XYM13, GYH213, PSR17], Eprime Software[LBR*17], SREF weather ensemble \SQL [PWB*09]
Operational Forecasting	Qt 4.5[SZD*10], ParaView [WHP*11], Cuda [HMC*13]
Ensemble Forecasting and Parameter Analysis	Met.3D[FKRW16, FKRW17], Matlab[JDKW15,FKRW16, FKRW17], (CUDA/OpenGL/GLSL Shaders)[HMZ*14], (C++/Qt)[GYH213, CZC*15], (Processing/C++)[QM16], Eprime Software[LBR*17], R [LPCR18]
Climate Study/ Atmospheric Research/ Meteorology	(C++/OpenGL/Qt)[MMME11, WSL*14, CZC*15, SGL*16, ZW18, NR18], Met.3D[FKRW16, KHS*18a,KTB*18], Matlab[SWMW09, JDKW15, FKRW16 FKRW17], Paraview [HRB*13, RSAS17, GSFT16, BPR*15], (NCL (NCAR Command Language)/Avizo)[BPR*15], OpenGL's vertex shader[EGG*15], (LULC Map/GDAL)[HMRH15], (FTLE) Finite-Time Lyapunov Exponent tool [GHP*16], ModelforPredictionAcrossScales(MPAS)-Ocean[WPS*16], (C++/Direct3D/Amira)[GSFT16], VISUS framework[WJW*17], (C++/Cuda)[GXY12, PSR17], C++\OpenCV\Farneback [DNN13],
Large Scale Data Simulation/Visualization	Flex/Bison/SQL [GHA*08], C++/Cuda/fortran [GXY12]
Exploring Large Multivariate Data	Matlab[JDKW15], (C++/Qt)[CZC*15, ZW18, NR18], (LULC Map/GDAL)[HMRH15], (C++/Cuda)[JC09, GXY12, HMC*13, PSR17],
Others	(Java/WEKA machine learning suite)[BJA*15], (C++/Direct3D/Paraview/Amira)[GSFT16, HDS16], VISUS framework[WJW*17], Cuda[MPBM*18], Visual C++\MySQL\OpenGL [LEH09, MMME11], ParaView [HRB*13], MatLab\GPU [SWMW09],

**Table 9:** Various tools and programming languages used by surveyed papers in the ocean and atmospheric sciences application areas.

model development). One aspect missing in this model is to capture the reasoning and thinking process of a human analyst [ERT\*17]. Klein et al. [KMH06a, KMH06b] presented a human-centered approach towards sensemaking based on the Data/Frame theory. They represent the current information or understanding of humans in terms of ‘frames’. New frames are created or existing frames are strengthened, weakened, rejected or modified as they acquire, explore, and analyze more data. A more recent model by Sacha et al. [SSS\*14] is a knowledge generation and synthesis model for visual analytics that consists of both computer and human parts (Figure 14), representing a more human-centered approach. The left part of this model represents a visual analytics system and the right part presents a human knowledge generation process through exploration, verification, and sense-making loops. Ribarsky et al. [RF16] proposed an extension to this model based on insights derived from cognitive science principles. There are some other relevant models and frameworks related to the intersection of visual analytics, machine learning and dimensionality reduction [SZS\*17, SSZ\*17]. We believe that Sacha’s model [SSS\*14] more accurately captures the workflows of ocean and atmospheric domain experts that utilize visual analytics tools. We have annotated different parts of this model with the task requirements identified during our interviews with domain experts (Section 5) to highlight components of this model associated with different task requirements.

As shown in Figure 14, task requirements **R1**, **R2**, **R3**, and **R4** correspond to the data component of this model, whereas **R2** partially overlaps with the Visualization, and **R3** partially overlaps with the Model. Task requirement **R5** is associated with the Visualization as it is primarily concerned with providing interactive data analysis and exploration with some added features. **R6**, **R7**, and **R8** are mainly associated with the Model component whereas **R9** is associated with both Visualization and Model component. **R10** is associated with the entire visual analytics system.

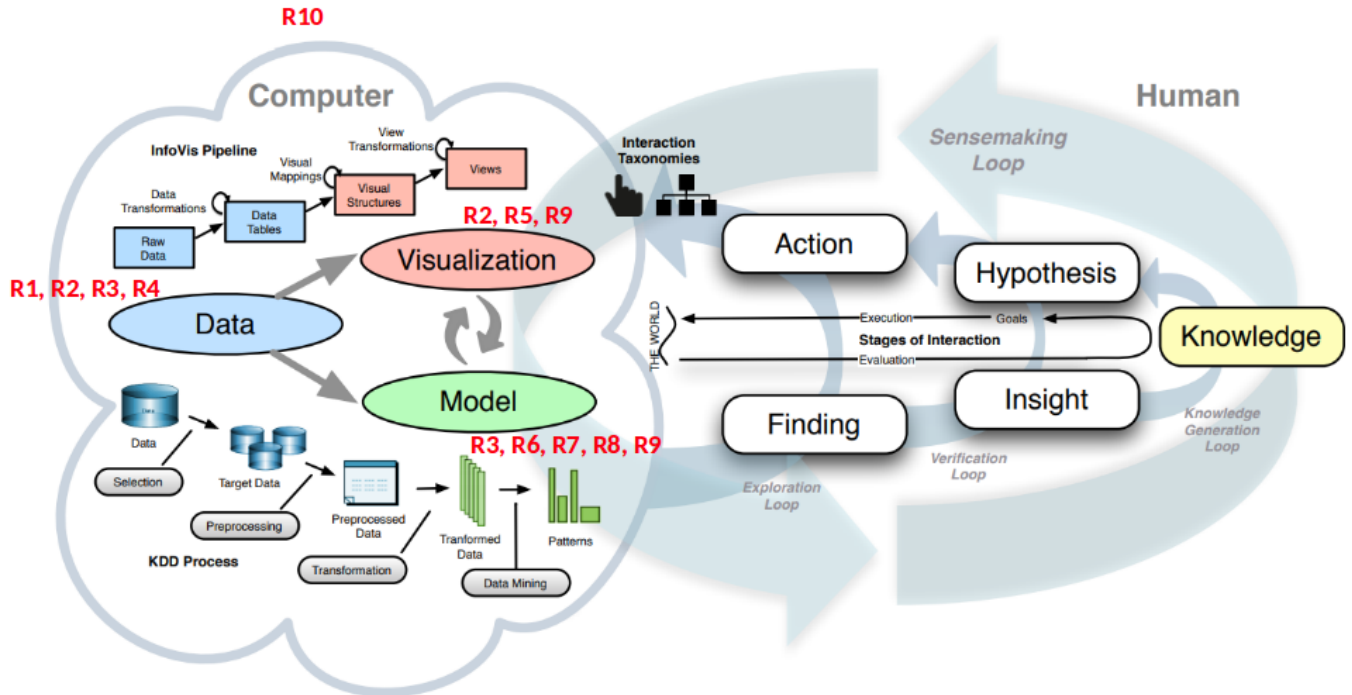
In addition to identifying the relationships between different task requirements and visual analytics components in this knowledge generation model, we apply this model to the coding results of different taxonomies used in this survey report as shown in Figure 15, by adopting an approach similar to Sacha et al. [SSS\*14] for comparative system assessment. The only difference is that Sacha et al. [SSS\*14] applied the model to individual systems whereas we

applied the model to survey coding tables. This will help us identify how strongly different functions or components are being supported by our sample set of survey papers.

We rank each component/functionality of this model on a scale of weak, basic, or strong as shown in the legend (Figure 15) and encode this strength through the weight of the line and color intensity. We have encoded the strength of all components of this visual analytics model, i.e., Data (D), Visualization(V), Model (M), Exploration loop (EL), Validation Loop (VL), and Knowledge Generation Loop (KL) after deriving the strength from the coding tables.

To calculate the strength, we identified the list of taxonomy subcategories in all taxonomy tables that belong to each component of the model. Then we added the paper counts corresponding to this list of taxonomy subcategories to calculate an overall sum for each component. This process is then repeated for all components of the model. The value calculated for each component is normalized based on an aggregated value obtained by adding individual component values. Then we assigned a score of weak, basic or strong to the model components based on this normalized value. Once this score is calculated, we looked at the domain experts tasks requirements that belong to each component of the knowledge generation model (e.g., R1, R2, R3, and R4 belong to the Data component as shown in Figure 14) to determine the extent to which these task requirements are covered by the taxonomy tables (based on paper counts).

Figure 15 shows that, overall, surveyed papers have strong support for the Visualization component of the model derived mainly from the visualization and interaction taxonomy coding tables, and is related to **R5** task and system requirement of domain experts. Since there is strong support for the Visualization category and a basic support for the Model component, the exploration loop as a result has strong support as well. The Data component is assigned weak support based on factors such as: there are 4 task and system requirements identified during our interviews with the domain experts, and there were shortcomings identified in the taxonomy tables related to these task requirements. For example, domain experts also work on the Terabyte and Petabyte scale data, but in our paper coding results of data sizes, the majority of works were at the Gigabyte scale. Based on such factors, we adjusted the support score for the Data component of the model.



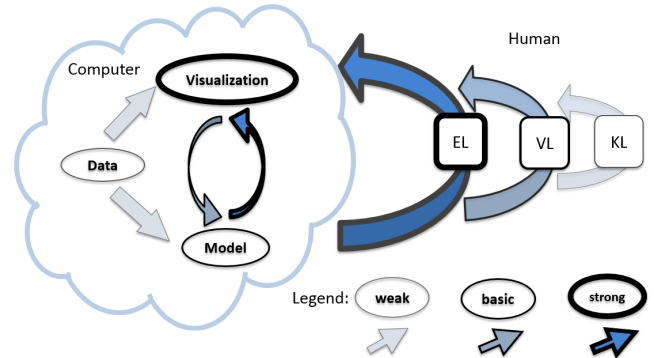
**Figure 14:** Knowledge generation model for visual analytics by Sacha et al. [SSS\*14]. Red text represents task requirements collected from domain experts and is mapped to corresponding components.

The Verification loop was assigned medium support strength due to an overall medium score derived from taxonomy tables and domain experts task requirements. The knowledge generation loop support strength is difficult to evaluate [SSS\*14] as we have to consider the domain knowledge of users as well as other external factors. We assigned a lower score for the knowledge generation loop because the features in surveyed papers are mostly at the lower level (exploration level), as evident from domain experts interviews, and there is limited explicit support for knowledge generation and verification loops in coded papers.

The resultant model helps us connect the information obtained through coded taxonomy tables and domain experts task requirements. A strong component means that it is not only well represented in the taxonomy tables but the domain experts task requirements corresponding to that component are also well represented in the taxonomy tables. This model summarizes and highlights the gaps between the task requirements of domain experts and the current focus of visualization research in the domain of ocean and atmospheric sciences.

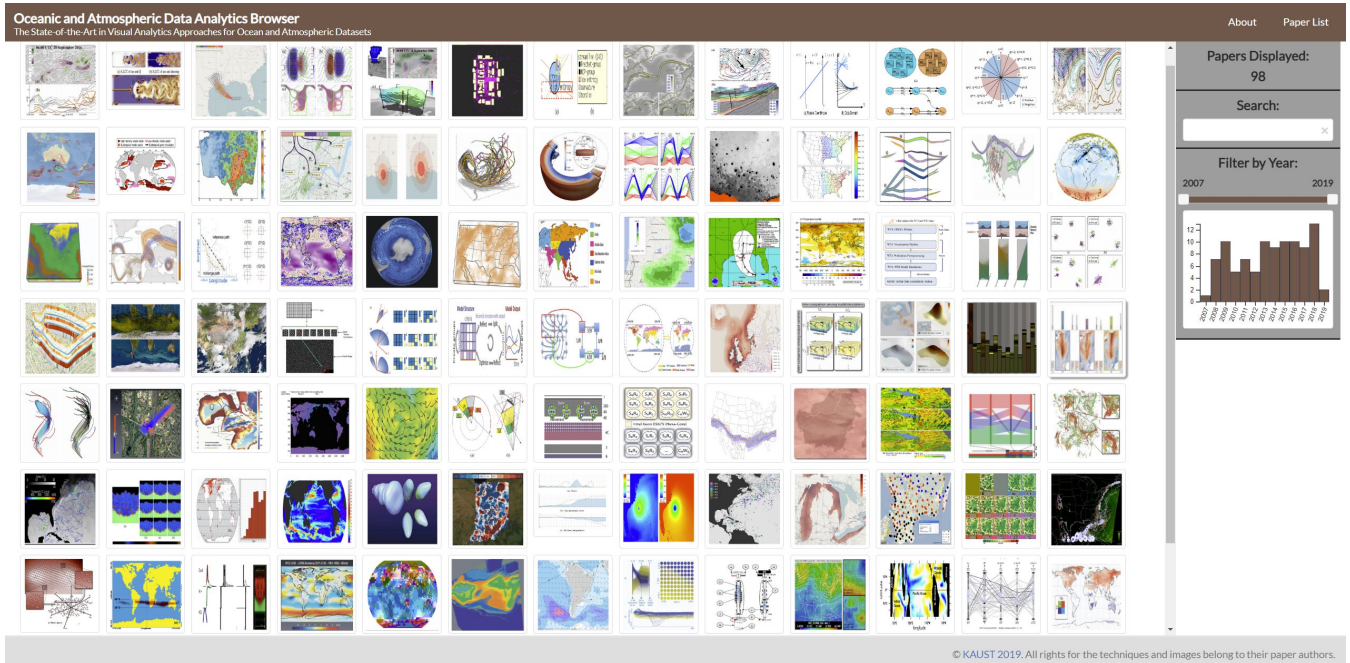
## 8. Web Browser for Exploring Papers Included in the Survey

We have designed and implemented a web browser inspired by the text visualization browser [KK15], displaying all of the filtered papers we discussed in our report. Each paper is represented as a thumbnail chosen from its paper as its representation in the main panel. Users can hover over the thumbnail to see the title of the paper. They can also click on any thumbnail to see



**Figure 15:** Model application to coding results for different taxonomies of our survey. Border intensity of each component (Data, Visualization, and Model) encodes the extent to which that component is represented in the surveyed papers (papers coded based on different taxonomies). Similarly, arrow border and color intensity encodes the extent to which that particular loop or process is covered in the surveyed papers.

the paper details (paper title, authors, and publication venue) in a popup window. On the side bar, there are filtering options available to filter the papers based on different attributes. In the future extension of this web browser, we plan to provide additional filtering options based on all taxonomy categories. Figure 16 shows an image of the browser. The browser is



**Figure 16:** Interactive Ocean and Atmospheric Data Analytics Browser for our surveyed papers.

available online at <https://kaust-vislab.github.io/Ocean-and-Atmospheric-Data-Analytics-Browser/>.

## 9. Discussion

In this paper, we have summarized the state of the art in visual analytics for the ocean and atmospheric sciences, identified task requirements by conducting interviews with domain experts, coded papers based on ten different taxonomies, and grounded the discussion and findings of our survey within the context of visual analytics frameworks and models. In doing so, we have identified areas of potential research where there is a need for collaborative efforts between visualization, and ocean and atmospheric researchers. Rautenhaus et al.'s survey [RBS\*17] is primarily looking at operational forecasting and meteorological visualization research, and their survey is focused on data analysis tasks. In our work, we are looking at current ocean and atmospheric research more from a perspective of visual analytics; thus, why we have coded this research based on the taxonomies of task requirements, interaction methods, visualization techniques, machine learning and statistical methods, evaluation methods, data types, data dimensions and size, spatial scale and application areas to identify barriers to adoption of visual analytics in ocean and atmospheric research and development [TDN11].

While looking at the coding results based on machine learning and statistics (Section 6.4), we found that deep learning methods are very rarely used in the sample set of papers used for coding. The detailed coding tables are provided as supplementary material with this paper. We merged all these categories to create a more compact coding table included in this paper (Table 4). Utilizing these models

can be advantageous in areas like forecasting, anomaly detection, feature detection, predictive analytics, etc.

The complexity of different models and simulations will continue to increase due to advancements in these fields; visual analytics and data mining need to be designed or adapted to handle the output data. Forecasting and prediction are essential analysis tasks (Task requirement **R8**) from a domain experts perspective, and in future, there could be more emphasis on these tasks but at better resolution [BTB15], and machine learning and visual analytics have to provide improved automated methods that can handle this scale and complexity.

We found a significant gap between the scale of data currently being used in the majority of visualization work (Gigabyte scale) and the data requirements of ocean and atmospheric domain experts (Petabyte and Terabyte scale) (Task requirement **R3**). This mismatch is problematic as it makes the adoption and usage of visualization tools difficult for data at that scale. In order to fill this gap, visualization research can incorporate big data architectural support in the implementation, focusing on capabilities like data sampling, summarization, compression, remote rendering and processing (in-situ visualization), and scalable client server visualization frameworks. Figure 15 also shows that there is a weak support for data requirements of domain experts in the current visualization research related to ocean and atmospheric sciences.

During our interviews for task requirements analysis, we found that domain researchers do not commonly use state-of-the-art visualization techniques. This is also highlighted in [TDN11]. There are some significant barriers that exist in successfully integrating their datasets with visualization tools. These barriers or limitations are related to data storage and efficient retrieval, data representa-



tion, preprocessing, reshaping, transformations, interface with data processing tools on high performance clusters and servers, remote data visualizations (such as web-based visualizations linked with processes running on servers), resolving differences in data (resolution, scale, and type), disparate data sources, and superimposing multiple datasets for simultaneous analysis. There is a need to address the scalability issues associated with visual analytics tools that can handle large simulations, ensemble datasets, etc. This is also shown in the results of applying visual analytics model to paper coding results in figure 15. As shown in the figure, the data component is weakly supported in coded papers from the perspective of the ocean and atmospheric sciences.

There are visualization tools designed for expert use in climate data analysis, such as UV-CDAT [WBD\*13], which is an extended version of an earlier tool that was unable to handle large datasets. There are other general purpose tools that are used in ocean and atmospheric domain like Met.3D [RKSW15], Paraview [AGL05], VTK [SLM04], and VisIt [CBW\*12]. There are in-situ analysis techniques as well that are useful to address the computational and storage requirements for large scale datasets [AJO\*14, WPS\*16, LAA\*17]. They generally build on platforms that support remote rendering capabilities. General purpose visualization tools can address common task requirements for analysis in this domain but for researchers working on advanced research problem often require customizations to address their needs that mandate designing custom visualization tools (Task requirement R3).

There is a need to carry out collaborative work between ocean and atmospheric domain experts and visual analytics researchers to design custom visualization tools (Task requirement R10) tailor-made to solve specific problems and address challenges associated with the adoption of visualization techniques, and extend the state of the art on the intersection of these two areas.

If we look at the coding tables of evaluation (Figure 6), we observe that, besides QRI, AP, and UE, there is limited focus on other evaluation methods. Using evaluation scenarios like UWP can benefit visual analytics researchers in general, as it can help evaluate the workflows of the ocean and atmospheric domain experts, and can provide guidelines for visualization tool design.

After looking at the coded tables based on the interaction taxonomy, we observed that Shepherd's method (direct or indirect) [LGH\*17] is relatively underutilized. Incorporating this method can encourage more engagement of users with visualizations and provide them with analysis capabilities at higher abstraction levels, that again also requires more automated methods in visual analytics, machine learning, and data mining.

## 10. Future Directions, Challenges, and Opportunities

Based on the analysis of the results of our survey, we have identified several key challenges and opportunities for future work.

In section 7, we annotated the knowledge generation and synthesis model proposed by Sacha et al. [SSS\*14] with the task requirements gathered from domain experts based on our interviews (Figure 14). If we look at the coding table of the interaction taxonomy (Table 3), there is relatively better support for interaction techniques in visual analytics literature (except for Shepherd's method

[LGH\*17]), such as we mentioned earlier, presents an opportunity for future work. In the visualization techniques taxonomy table, 2D and 3D visualizations have relatively better coverage in visualization literature as compared to other categories included in this taxonomy. It is possible that this is due to the nature and type of datasets utilized in ocean and atmospheric research. Also, if we look at Figure 15, we can see that Visualization component has strong support, and it addresses interaction requirements (Task requirement R5) of the domain experts.

In the machine learning and statistics taxonomy table (Table 4), clustering is the most popular category and there is relatively less support for other techniques in the literature. Figure 15 also shows that Model has basic support. Our initial coding tables included more categories specifically related to machine learning, but the corresponding coding results were very sparse. This indicates an opportunity for future research especially for integrating deep learning models in visual analytics and provide interactive visualizations that facilitate modifications at a higher abstraction level.

Analyzing the coding tables for evaluation taxonomy (Table 6), we can see that QRI, AP, and UE are the most commonly used evaluation scenarios. This also shows that there is an opportunity for research, especially in collaboration, communication, and studying environments and work practices (Task requirement R10) of the ocean and atmospheric researchers. This indicates a need to improve the usability design of visualization tools targeted for the ocean and atmospheric domain.

Based on our interviews with domain experts, we found that there is a need to adapt the visualization tools to the pipelines and environments. They also mentioned the limitations associated with managing the scale, transformations, storage, efficient retrieval, filtering, and other data related issues (Task requirements R1, R2, R3 and R4). Figures 14 and 15 show that there is a strong need to do more collaborative work between the ocean and climate research community and visualization researchers, to address the data related task requirements of domain experts. The coding table related to tasks taxonomy (Table 1) also shows that there is limited work related to data overlays, simultaneous analysis of multiple heterogeneous datasets, incorporating datasets with different resolutions, and scale management. With future advancements in computational technology and improved models, there is a need to provide more automated techniques coupled with interactive visualizations. As mentioned earlier, one of the major gaps we found in the task requirements of domain experts and the current visualization research is the difference in the scale of data used in practice. Domain experts mentioned that they routinely encounter datasets at Petabyte or Terabyte scale, but our survey results show that mostly data at Gigabyte scale or less is used in the visualization research, so visual analytics systems need to scale better, bearing in mind these requirements.

## 11. Conclusion

In this survey, we provide a comprehensive survey of the state of the art in visual analysis approaches related to the ocean and atmospheric datasets. We have conducted interviews with 18 domain experts to identify their task and system requirements for a visual analytics system related to ocean and atmospheric datasets. We have

grouped these task requirements into four major categories: data, interactions, features/tasks, and visual analytics design. We have conducted the survey based on ten different taxonomies, and coded papers based on the categories defined within these taxonomies. We relate the results of our coding with the task requirements, and identified the gaps in domain scientists task requirements, and the visualization and visual analytics research. We also discuss these survey results in the context of visual analytics knowledge generation and sense-making models and frameworks. We apply one of these models to our survey results, and draw important conclusions and recommendations for future research.

## 12. Acknowledgements

This study was supported by the Office of Sponsored Research (OSR) at King Abdullah University of Science and Technology (KAUST) under the “Virtual Red Sea Initiative” (award #REP/1/3268-01-01). We want to acknowledge Dr. Hari Prasad Dasari for his help and contribution in organizing the interviews with domain experts. We also like to thank KAUST Visualization Core Lab for their help and support.

## References

- [AAB\*10] ANDRIENKO G., ANDRIENKO N., BREMM S., SCHRECK T., VON LANDESBERGER T., BAK P., KEIM D.: Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. In *Computer Graphics Forum* (2010), pp. 913–922. 11, 12, 14
- [AGL05] AHRENS J., GEVECI B., LAW C.: Paraview: An end-user tool for large data visualization. *The visualization handbook* 717 (2005). 18, 22
- [AJO\*14] AHRENS J., JOURDAIN S., OLEARY P., PATCHETT J., ROGERS D. H., PETERSEN M.: An image-based approach to extreme scale in situ visualization and analysis. In *SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (Nov 2014), pp. 424–434. 22
- [AZM\*15] ANREDDY S., ZHANG S., MERCER A., DYER J., SWAN J. E.: Visual scalability of spatial ensemble uncertainty. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (Oct 2015), pp. 187–188. 1
- [BAOL12] BRODLIE K., ALLENDES-OSORIO R., LOPES A.: *A Review of Uncertainty in Data Visualization*. Springer, London, 2012, pp. 81–109. 3
- [BBP08] BLAAS J., BOTHA C., POST F.: Extensions of parallel coordinates for interactive exploration of large multi-timepoint data sets. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1436–1451. 8, 10, 13, 14
- [BGOJ16] BENSEMA K., GOSINK L., OBERMAIER H., JOY K. I.: Modality-driven classification and visualization of ensemble variance. *IEEE Transactions on Visualization and Computer Graphics* 22, 10 (Oct 2016), 2289–2299. 12
- [BHJ\*14] BONNEAU G.-P., HEGE H.-C., JOHNSON C. R., OLIVEIRA M. M., POTTER K., RHEINGANS P., SCHULTZ T.: Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization*. Springer, 2014, pp. 3–27. 3
- [BJA\*15] BUCHMÜLLER J., JANETZKO H., ANDRIENKO G., ANDRIENKO N., FUCHS G., KEIM D. A.: Visual Analytics for Exploring Local Impact of Air Traffic. *Computer Graphics Forum* (2015). 9
- [BLS17] BISWAS A., LIN G., LIU X., SHEN H.-W.: Visualization of time-varying weather ensembles across multiple resolutions. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 841–850. 1, 9, 16
- [BMLC19] BERENJKOUB M., MONICO R. O., LARAMEE R. S., CHEN G.: Visual analysis of spatio-temporal relations of pairwise attributes in unsteady flow. *IEEE transactions on visualization and computer graphics* 25, 1 (2019), 1246–1256. 7, 13
- [BPR\*15] BÖTTINGER M., POHLMANN H., RÖBER N., MEIER-FLEISCHER K., SPICKERMANN D.: Visualization of 2D Uncertainty in Decadal Climate Predictions. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)* (2015), pp. 1–5. 8
- [BTB15] BAUER P., THORPE A., BRUNET G.: The quiet revolution of numerical weather prediction. *Nature* 525, 7567 (2015), 47. 21
- [BW08] BACHTHALER S., WEISKOPF D.: Continuous scatterplots. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (Nov 2008), 1428–1435. 14
- [CBW\*12] CHILDS H., BRUGGER E., WHITLOCK B., MEREDITH J., AHERN S., PUGMIRE D., BIAGAS K., MILLER M., HARRISON C., WEBER G. H., KRISHNAN H., FOGAL T., SANDERSON A., GARTH C., BETHEL E. W., CAMP D., RÜBEL O., DURANT M., FAVRE J. M., NAVRÁTIL P.: VisIt: An End-User Tool For Visualizing and Analyzing Very Large Data. In *High Performance Visualization—Enabling Extreme-Scale Scientific Insight*. Oct 2012, pp. 357–372. 22
- [CDO] Climate data operators. <https://code.mpimet.mpg.de/projects/cdo>. Accessed: 2019-01-29. 5
- [CWMW11] CHEN C.-K., WANG C., MA K.-L., WITTENBERG A. T.: Static correlation visualization for large time-varying volume data. In *Visualization Symposium (PacificVis), IEEE Pacific* (2011), pp. 27–34. 12
- [CZC\*15] CHEN H., ZHANG S., CHEN W., MEI H., ZHANG J., MERCER A., LIANG R., QU H.: Uncertainty-aware multidimensional ensemble data visualization and exploration. *IEEE Transactions on Visualization and Computer Graphics* 21, 9 (Sept 2015), 1072–1086. 8
- [DDW14] DEMIR I., DICK C., WESTERMANN R.: Multi-charts for comparative 3d ensemble visualization. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec 2014), 2694–2703. 18
- [DIANS10] DANIELS II J., ANDERSON E. W., NONATO L. G., SILVA C. T.: Interactive vector field feature identification. *IEEE Transactions on Visualization & Computer Graphics*, 6 (2010), 1560–1568. 7, 12
- [DKG15] DASGUPTA A., KOSARA R., GOSINK L.: Vimtex: A visualization interface for multivariate, time-varying, geological data exploration. In *Computer Graphics Forum* (2015), vol. 34, pp. 341–350. 8, 9, 10, 13
- [DNN13] DORAISWAMY H., NATARAJAN V., NANJUNDIAH R. S.: An exploration framework to identify and track movement of cloud systems. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec 2013), 2896–2905. 7, 8, 14
- [DPD\*15] DIEHL A., PELOROSSO L., DELRIEUX C., SAULO C., RUIZ J., GRÖLLER M. E., BRUCKNER S.: Visual analysis of spatio-temporal data: Applications in weather forecasting. In *Computer Graphics Forum* (2015), Wiley Online Library, pp. 381–390. 1
- [DPD\*16] DIEHL A., PELOROSSO L., DELRIEUX C., MATKOVIC K., RUIZ J., GRÖLLER M. E., BRUCKNER S.: Albero: A Visual Analytics Approach for Probabilistic Weather Forecasting. *Computer Graphics Forum* (2016). 16
- [EGG\*15] ELSHEHALY M., GRAČANIN D., GAD M., ELMONGUI H. G., MATKOVIĆ K.: Interactive fusion and tracking for multi-modal spatial data visualization. In *Computer Graphics Forum* (2015), pp. 251–260. 7
- [EHN17] ELLSWORTH D. A., HENZE C. E., NELSON B. C.: Interactive visualization of high-dimensional petascale ocean data. In *2017 IEEE 7th Symposium on Large Data Analysis and Visualization (LDAV)* (Oct 2017), pp. 36–44. 16
- [EM10] ELLIS G., MANSMANN F.: *Mastering the information age solving problems with visual analytics*. Eurographics, 2010. 2



- [ERT\*17] ENDERT A., RIBARSKY W., TURKAY C., WONG B. W., NABNEY I., BLANCO I. D., ROSSI F.: The state of the art in integrating machine learning into visual analytics. *Computer Graphics Forum* (2017). 19
- [ET08] ELMQVIST N., TSIGAS P.: A taxonomy of 3d occlusion management for visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 5 (2008), 1095–1109. 13
- [FBW16] FERSTL F., BÄURGER K., WESTERMANN R.: Streamline variability plots for characterizing the uncertainty in vector field ensembles. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan 2016), 767–776. 8
- [FKRW16] FERSTL F., KANZLER M., RAUTENHAUS M., WESTERMANN R.: Visual analysis of spatial variability and global correlations in ensembles of iso-contours. In *Computer Graphics Forum* (2016), pp. 221–230. 13
- [FKRW17] FERSTL F., KANZLER M., RAUTENHAUS M., WESTERMANN R.: Time-hierarchical clustering and visualization of weather forecast ensembles. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 831–840. 1, 8, 10, 12, 13
- [FTMKH\*18] FRIEDRICI A., TOYE MAHAMADOU KELE H., HOTEIT I., WEINKAUF T., THEISEL H., HADWIGER M.: A lagrangian method for extracting eddy boundaries in the red sea and the gulf of aden. In *IEEE Scientific Visualization 2018 Short Papers* (2018). 1
- [FWR\*17] FEDERICO P., WAGNER M., RIND A., AMOR-AMORÓS A., MIKSCH S., AIGNER W.: The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *Proc. IEEE Conference on Visual Analytics Science and Technology (VAST)* (2017). 2
- [GGA\*11] GOSINK L. J., GARTH C., ANDERSON J. C., BETHEL E. W., JOY K. I.: An application of multivariate statistical analysis for query-driven visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 3 (2011), 264–275. 10
- [GHA\*08] GLATTER M., HUANG J., AHERN S., DANIEL J., LU A.: Visualizing temporal patterns in large multivariate data using textual pattern matching. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (Nov 2008), 1467–1474. 10
- [GHP\*16] GUO H., HE W., PETERKA T., SHEN H.-W., COLLIS S. M., HELMUS J. J.: Finite-time lyapunov exponents and lagrangian coherent structures in uncertain unsteady flows. *IEEE transactions on visualization and computer graphics* 22, 6 (2016), 1672–1682. 13
- [GSFT16] GÜNTHER T., SCHULZE M., FRIEDERICI A., THEISEL H.: Visualizing volcanic clouds in the atmosphere and their impact on air traffic. *IEEE computer graphics and applications* 36, 3 (2016), 36–47. 8
- [GXY12] GUO H., XIAO H., YUAN X.: Scalable multivariate volume visualization and analysis based on dimension projection and parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics* 18, 9 (Sept 2012), 1397–1410. 7, 12, 13, 14, 18
- [GYHZ13] GUO H., YUAN X., HUANG J., ZHU X.: Coupled ensemble flow line advection and analysis. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (Dec 2013), 2733–2742. 7, 13
- [HDS16] HAZARIKA S., DUTTA S., SHEN H.-W.: Visualizing the variations of ensemble of isosurfaces. In *IEEE Pacific Visualization Symposium (PacificVis)* (2016), pp. 209–213. 13
- [HMC\*13] HÖLLT T., MAGDY A., CHEN G., GOPALAKRISHNAN G., HOTEIT I., HANSEN C. D., HADWIGER M.: Visual analysis of uncertainties in ocean forecasts for planning and operation of off-shore structures. In *IEEE Pacific Visualization Symposium (PacificVis)* (2013), pp. 185–192. 1, 16, 17, 18
- [HMRH15] HÄB K., MIDDEL A., RUDELL B. L., HAGEN H.: Travis—a visualization framework for mobile transect data sets in an urban microclimate context. In *IEEE Pacific Visualization Symposium (PacificVis)* (2015), pp. 167–174. 16
- [HMZ\*14] HÖLLT T., MAGDY A., ZHAN P., CHEN G., GOPALAKRISHNAN G., HOTEIT I., HANSEN C. D., HADWIGER M.: Ovis: A framework for visual analysis of ocean forecast ensembles. *IEEE Transactions on Visualization and Computer Graphics* 20, 8 (2014), 1114–1126. 9
- [HRB\*13] HELBIG C., RINK K., BAUER H.-S., WULFMEYER V., FRANK M., KOLDITZ O.: 3D Visualization of Atmospheric Data for Analytical Approaches. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)* (2013), Kolditz O., Rink K., Scheuermann G., (Eds.), The Eurographics Association. 13
- [IIC\*13] ISENBERG T., ISENBERG P., CHEN J., SEDLMAYER M., MÄÜLLER T.: A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics* (2013). 15
- [JBMS09] JANICKE H., BOTTINGER M., MIKOLAJEWICZ U., SCHEUERMANN G.: Visual exploration of climate variability changes using wavelet analysis. *IEEE transactions on visualization and computer graphics* 15, 6 (2009). 11
- [JBS08] JÄNICKE H., BÖTTINGER M., SCHEUERMANN G.: Brushing of attribute clouds for the visualization of multivariate data. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008). 11
- [JDKW15] JAREMA M., DEMIR I., KEHRER J., WESTERMANN R.: Comparative visual analysis of vector field ensembles. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (2015), pp. 81–88. 14, 16
- [JG09] JIN H., GUO D.: Understanding climate change patterns with multivariate geovisualization. In *IEEE International Conference on Data Mining Workshops. ICDMW.* (2009), pp. 217–222. 14
- [JNL10] JOHANSSON J., NESSET T. S., LINNÄL R. B.: Evaluating climate visualization: An information visualization approach. In *14th International Conference Information Visualisation* (July 2010), pp. 156–161. 3
- [JS14] JÄDNICKE S., SCHEUERMANN G.: Utilizing GeoTemCo for Visualizing Environmental Data. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)* (2014). 13
- [KBL18] KAPPE C. P., BOTTINGER M., LEITTE H.: Exploring variability within ensembles of decadal climate predictions. *IEEE Transactions on Visualization & Computer Graphics* (2018), 1499–1512. 1, 9, 13
- [KCA\*16] KO S., CHO I., AFZAL S., YAU C., CHAE J., MALIK A., BECK K., JANG Y., RIBARSKY W., EBERT D. S.: A survey on visual analysis approaches for financial data. *Computer Graphics Forum* 35, 3 (2016), 599–617. 13
- [Kei02] KEIM D.: Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics* 8, 1 (2002), 1–8. 13, 14
- [KFH10] KEHRER J., FILZMOSER P., HAUSER H.: Brushing moments in interactive visual analysis. In *Computer Graphics Forum* (2010), Wiley Online Library, pp. 813–822. 8, 11
- [KH13] KEHRER J., HAUSER H.: Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE transactions on visualization and computer graphics* 19, 3 (2013), 495–513. 2, 3
- [KHS\*18a] KERN M., HEWSON T., SADLO F., WESTERMANN R., RAUTENHAUS M.: Robust detection and visualization of jet-stream core lines in atmospheric flow. *IEEE transactions on visualization and computer graphics* 24, 1 (2018), 893–902. 1, 13
- [KHS\*18b] KERN M., HEWSON T., SCHÄFLER A., WESTERMANN R., RAUTENHAUS M.: Interactive 3d visual analysis of atmospheric fronts. *IEEE transactions on visualization and computer graphics* (2018). 1, 13, 14
- [KK15] KUCHER K., KERREN A.: Text visualization techniques: Taxonomy, visual survey, and community insights. In *IEEE Pacific Visualization Symposium (PacificVis)* (April 2015), pp. 117–121. 20
- [KLM\*08] KEHRER J., LADSTÄDTER F., MUIGG P., DOLEISCH H., STEINER A., HAUSER H.: Hypothesis generation in climate research with interactive visual data exploration. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1579–1586. 16, 17

- [KMH06a] KLEIN G., MOON B., HOFFMAN R. R.: Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems* 21, 4 (July 2006), 70–73. 2, 19
- [KMH06b] KLEIN G., MOON B., HOFFMAN R. R.: Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems* 21, 5 (Sept 2006), 88–92. 2, 19
- [KRRW18] KUMPF A., RAUTENHAUS M., RIEMER M., WESTERMANN R.: Visual analysis of the temporal evolution of ensemble forecast sensitivities. *IEEE transactions on visualization and computer graphics* (2018). 1
- [KRRW19] KUMPF A., RAUTENHAUS M., RIEMER M., WESTERMANN R.: Visual analysis of the temporal evolution of ensemble forecast sensitivities. *IEEE transactions on visualization and computer graphics* 25, 1 (2019), 98–108. 7, 12
- [KSU\*14] KÖTHUR P., SIPS M., UNGER A., KUHLMANN J., DRANSCH D.: Interactive visual summaries for detection and assessment of spatiotemporal patterns in geospatial time series. *Information Visualization* 13, 3 (2014), 283–298. 11
- [KTB\*18] KUMPF A., TOST B., BAUMGART M., RIEMER M., WESTERMANN R., RAUTENHAUS M.: Visualizing confidence in cluster-based ensemble weather forecast analyses. *IEEE transactions on visualization and computer graphics* 24, 1 (2018), 109–119. 1, 7, 12
- [LAA\*17] LARSEN M., AHRENS J., AYACHIT U., BRUGGER E., CHILDS H., GEVECI B., HARRISON C.: The alpine in situ infrastructure: Ascending from the ashes of strawman. In *Proceedings of the In Situ Infrastructures on Enabling Extreme-Scale Analysis and Visualization* (New York, NY, USA, 2017), ISAV'17, ACM, pp. 42–46. 22
- [LBI\*12] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPENDALE S.: Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics* (2012). 15, 16
- [LBR\*17] LIU L., BOONE A. P., RUGINSKI I. T., PADILLA L., HEGARTY M., CREEM-REGEHR S. H., THOMPSON W. B., YUKSEL C., HOUSE D. H.: Uncertainty visualization by representative sampling from prediction ensembles. *IEEE transactions on visualization and computer graphics* 23, 9 (2017), 2165–2178. 1, 16
- [LEH09] LUNDBLAD P., EURENIUS O., HELDRING T.: Interactive visualization of weather and ship data. In *IEEE 13th International Conference Information Visualisation* (2009), pp. 379–386. 9, 10
- [LGH\*17] LU Y., GARCIA R., HANSEN B., GLEICHER M., MACIEJEWSKI R.: The state-of-the-art in predictive visual analytics. *Computer Graphics Forum* 36, 3 (June 2017), 539–562. 3, 9, 10, 22
- [LLBP12] LIU S., LEVINE J. A., BREMER P., PASCUCCI V.: Gaussian mixture model based volume visualization. In *IEEE Symposium on Large Data Analysis and Visualization (LDAV)* (Oct 2012), pp. 73–77. 8, 13
- [LLEJ11] LUNDBLAD P., LÖFVING H., ELOVSSON A., JOHANSSON J.: Exploratory visualization for weather data verification. In *IEEE 15th International Conference on Information Visualisation* (2011), pp. 306–313. 9, 13
- [LMK\*15] LIU L., MIRZARGAR M., KIRBY R. M., WHITAKER R., HOUSE D. H.: Visualizing time-specific hurricane predictions, with uncertainty, from storm path ensembles. In *Computer Graphics Forum* (2015), vol. 34, pp. 371–380. 1
- [LMW\*17] LIU S., MALJOVEC D., WANG B., BREMER P.-T., PASCUCCI V.: Visualizing high-dimensional data: Advances in the past decade. *IEEE Transactions on Visualization and Computer Graphics* 23, 3 (Mar. 2017), 1249–1268. 3
- [LPCR18] LIU L., PADILLA L., CREEM-REGEHR S. H., HOUSE D. H.: Visualizing uncertain tropical cyclone predictions using representative samples from ensembles of forecast tracks. *IEEE transactions on visualization and computer graphics* (2018). 1, 7, 16, 18
- [LS09] LEE T., SHEN H.: Visualization and exploration of temporal trend relationships in multivariate time-varying data. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov 2009), 1359–1366. 18
- [LSL\*10] LADSTÄDTER F., STEINER A. K., LACKNER B. C., PIRSCHER B., KIRCHENGAST G., KEHRER J., HAUSER H., MUIGG P., DOLEISCH H.: Exploration of climate data using interactive visualization. *Journal of Atmospheric and Oceanic Technology* 27, 4 (2010), 667–679. 10
- [LTJ10] LUNDBLAD P., THOURSIE J., JERN M.: Swedish road weather visualization. In *14th International Conference Information Visualisation* (July 2010), pp. 313–321. 8
- [LZM14] LI J., ZHANG K., MENG Z.-P.: Vismate: Interactive visual analysis of station-based observation data on climate changes. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (2014), pp. 133–142. 1, 13, 14, 15, 16
- [ME18] MA B., ENTEZARI A.: An interactive framework for visualization of weather forecast ensembles. *IEEE transactions on visualization and computer graphics* (2018). 1, 7, 9, 11, 13, 16, 18
- [MGA\*08] MEYER M., GOSINK L. J., ANDERSON J. C., BETHEL E. W., JOY K. I.: Query-driven visualization of time-varying adaptive mesh refinement data. *IEEE transactions on visualization and computer graphics* 14, 6 (2008), 1715–1722. 8
- [MGKH09] MATKOVIC K., GRACANIN D., KLARIN B., HAUSER H.: Interactive visual analysis of complex scientific data as families of data surfaces. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (Nov 2009), 1351–1358. 13
- [MHR\*11] MACIEJEWSKI R., HAFEN R., RUDOLPH S., LAREW S. G., MITCHELL M. A., CLEVELAND W. S., EBERT D. S.: Forecasting hotspots—a predictive analytics approach. *IEEE Transactions on Visualization and Computer Graphics* 17, 4 (2011), 440–453. 8, 13
- [MME11] MALIK A., MACIEJEWSKI R., MAULE B., EBERT D. S.: A visual analytics process for maritime resource allocation and risk assessment. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (2011), pp. 221–230. 8, 13
- [MN15] MASKEY M., NEWMAN T. S.: Directional texture for visualization - new technique and application study. In *19th International Conference on Information Visualisation* (July 2015), pp. 1–8. 14
- [MPBM\*18] MUNOZ-PANDIELLA I., BOSCH C., MERILLOU S., MERILLOU N., PATOW G., PUEYO X.: Urban weathering: Interactive rendering of polluted cities. *IEEE Transactions on Visualization and Computer Graphics* (2018). 13
- [MWK14] MIRZARGAR M., WHITAKER R. T., KIRBY R. M.: Curve boxplot: Generalization of boxplot for ensembles of curves. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec 2014), 2654–2663. 13, 16
- [NBD\*15] NOCKE T., BUSCHMANN S., DONGES J. F., MARWAN N., SCHULZ H.-J., TOMINSKI C.: visual analytics of climate networks. *Nonlinear Processes in Geophysics* 22, 5 (2015), 545. 3
- [NCO] Netcdf operators (nco). <http://nco.sourceforge.net/>. Accessed: 2019-01-29. 5
- [NR18] NGUYEN H., ROSEN P.: Dspcp: A data scalable approach for identifying relationships in parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics* 24, 3 (March 2018), 1301–1315. 16
- [NSBW08] NOCKE T., STERZEL T., BÖTTINGER M., WROBEL M.: Visualization of climate and climate change data: An overview. *Digital earth summit on geoinformatics* (2008), 226–232. 2
- [Pat18] PATGIRI R.: Taxonomy of big data: A survey. *preprint arXiv:1808.08474* (2018). 10
- [PC05] PIROLI P., CARD S.: The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis* (2005), vol. 5, McLean, VA, USA, pp. 2–4. 2, 18

- [PDW\*14a] POCO J., DASGUPTA A., WEI Y., HARGROVE W., SCHWALM C., COOK R., BERTINI E., SILVA C.: Similarityexplorer: A visual inter-comparison tool for multifaceted climate data. In *Computer Graphics Forum* (2014). 12, 14, 16
- [PDW\*14b] POCO J., DASGUPTA A., WEI Y., HARGROVE W., SCHWALM C. R., HUNTZINGER D. N., COOK R., BERTINI E., SILVA C. T.: Visual reconciliation of alternative similarity spaces in climate modeling. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 1923–1932. 8, 10, 11, 13, 14
- [PG02] PINELLE D., GUTWIN C.: Groupware walkthrough: Adding context to groupware usability evaluation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2002), ACM. 16
- [PH11] POTHKOW K., HEGE H.: Positional uncertainty of isocontours: Condition analysis and probabilistic measures. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (Oct 2011), 1393–1406. 13
- [PSR17] PENG C., SAHANI S., RUSHING J.: A gpu-accelerated approach for feature tracking in time-varying imagery datasets. *IEEE Transactions on Visualization and Computer Graphics* 23, 10 (Oct 2017), 2262–2274. 7, 17
- [PW13] PILAR D. H. F., WARE C.: Representing flow patterns by using streamlines with glyphs. *IEEE Transactions on Visualization and Computer Graphics* 19, 8 (Aug 2013), 1331–1341. 14
- [PWB\*09] POTTER K., WILSON A., BREMER P.-T., WILLIAMS D., DOUTRIAUX C., PASCUCCI V., JOHNSON C. R.: Ensemble-vis: A framework for the statistical visualization of ensemble data. In *Data Mining Workshops. ICDMW'09*. (2009), pp. 233–240. 8
- [QCX\*07] QU H., CHAN W., XU A., CHUNG K., LAU K., GUO P.: Visual analysis of the air pollution problem in hong kong. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov 2007), 1408–1415. 14
- [QM16] QUINAN P. S., MEYER M.: Visually comparing weather features in forecasts. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 389–398. 17, 18
- [RBS\*17] RAUTENHAUS M., BÖTTINGER M., SIEMEN S., HOFFMAN R., KIRBY R. M., MIRZARGAR M., RÖBER N., WESTERMANN R.: Visualization in meteorology—a survey of techniques and tools for data analysis tasks. *IEEE Transactions on Visualization and Computer Graphics* (2017), 3268–3296. 2, 16, 21
- [RF16] RIBARSKY W., FISHER B.: The human-computer system: Towards an operational model for problem solving. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (Jan 2016), pp. 1446–1455. 19
- [RKS\*15] RAUTENHAUS M., KERN M., SCHÄFLER A., WESTERMANN R.: Three-dimensional visualization of ensemble weather forecasts—part 1: The visualization tool met. 3d (version 1.0). *Geoscientific Model Development* (2015), 2329–2353. 18, 22
- [RSAS17] ROCHA A., SILVA J. D., ALIM U., SOUSA M. C.: Multi-variate Visualization of Oceanography Data Using Decals. In *Workshop on Visualisation in Environmental Sciences (EnvirVis)* (2017), The Eurographics Association. 14
- [RSG\*13] RÖBER N., SALIM M., GIERISCH A., BÖTTINGER M., SCHLÜNZEN H.: Visualization of urban micro-climate simulations. In *Proceedings of workshop on visualisation in environmental sciences (EnvirVis)* (2013), pp. 53–57. 16
- [RSM\*16] RADOŠ S., SPLECHTNA R., MATKOVIĆ K., ĐURAS M., GRÖLLER E., HAUSER H.: Towards quantitative visual analytics with structured brushing and linked statistics. In *Computer Graphics Forum* (2016), vol. 35, pp. 251–260. 7, 17
- [SFJK\*09] STEED C. A., FITZPATRICK P. J., JANKUN-KELLY T., YANCEY A. N., II J. E. S.: An interactive parallel coordinates technique applied to a tropical cyclone climate analysis. *Computers & Geosciences* (2009). 14, 18
- [SGL\*16] SHU Q., GUO H., LIANG J., CHE L., LIU J., YUAN X.: Ensemblegraph: Interactive visual analysis of spatiotemporal behaviors in ensemble simulation data. In *IEEE Pacific Visualization Symposium (PacificVis)* (2016), pp. 56–63. 1, 8, 16
- [SHB\*14] SEDLMAIR M., HEINZL C., BRUCKNER S., PIRINGER H., MÖLLER T.: Visual parameter space analysis: A conceptual framework. *IEEE Transactions on Visualization and Computer Graphics* (2014). 17
- [Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings IEEE Symposium on Visual Languages* (1996), pp. 336–343. 7, 8
- [SLM04] SCHROEDER W. J., LORENSEN B., MARTIN K.: *The visualization toolkit: an object-oriented approach to 3D graphics*. Kitware, 2004. 22
- [SMG\*15] SPLECHTNA R., MATKOVIĆ K., GRAČANIN D., JELOVIĆ M., HAUSER H.: Interactive visual steering of hierarchical simulation ensembles. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (Oct 2015), pp. 89–96. 1
- [SSJKF09] STEED C. A., SWAN J. E., JANKUN-KELLY T., FITZPATRICK P. J.: Guided analysis of hurricane trends using statistical processes integrated with interactive parallel coordinates. In *IEEE Symposium on Visual Analytics Science and Technology (VAST)* (2009), pp. 19–26. 1, 12
- [SSS\*14] SACHA D., STOFFEL A., STOFFEL F., KWON B. C., ELLIS G., KEIM D. A.: Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec 2014), 1604–1613. 2, 19, 20, 22
- [SSZ\*17] SACHA D., SEDLMAIR M., ZHANG L., LEE J. A., PELTONEN J., WEISKOPF D., NORTH S. C., KEIM D. A.: What you see is what you can change: Human-centered machine learning by interactive visualization. *Neurocomputing* (2017). 19
- [SWMW09] SUKHAREV J., WANG C., MA K.-L., WITTENBERG A. T.: Correlation study of time-varying multivariate climate data sets. In *IEEE Pacific Visualization Symposium (PacificVis)* (2009), pp. 161–168. 8, 11
- [SZD\*10] SANYAL J., ZHANG S., DYER J., MERCER A., AMBURN P., MOORHEAD R.: Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization & Computer Graphics*, 6 (2010), 1421–1430. 1, 16
- [SZS\*17] SACHA D., ZHANG L., SEDLMAIR M., LEE J. A., PELTONEN J., WEISKOPF D., NORTH S. C., KEIM D. A.: Visual interaction with dimensionality reduction: A structured literature analysis. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 241–250. 3, 19
- [TC05] THOMAS J. J., COOK K. A. (Eds.): *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Press, 2005. 2
- [TDN11] TOMINSKI C., DONGES J. F., NOCKE T.: Information visualization in climate research. In *15th International Conference on Information Visualization* (July 2011), pp. 298–305. 2, 6, 21
- [TWC\*18] TAO J., WANG C., CHAWLA N. V., SHI L., KIM S. H.: Semantic flow graph: A framework for discovering object relationships in flow fields. *IEEE Transactions on Visualization and Computer Graphics* 24, 12 (Dec 2018), 3200–3213. 7, 9, 16
- [USKD12] UNGER A., SCHULTE S., KLEMMANN V., DRANSCH D.: A visual analysis concept for the validation of geoscientific simulation models. *IEEE Transactions on Visualization & Computer Graphics*, 12 (2012), 2216–2225. 10
- [VLBA\*12] VON LANDESBERGER T., BREMM S., ANDRIENKO N., ANDRIENKO G., TEKUSOVA M.: Visual analytics methods for categorical spatio-temporal data. In *IEEE Conference on Visual Analytics Science and Technology (VAST)* (2012), pp. 183–192. 15
- [VMN\*18] VALSANGKAR A. A., MONTEIRO J. M., NARAYANAN V., HOTZ I., NATARAJAN V.: An exploratory framework for cyclone identification and tracking. *IEEE Transactions on Visualization and Computer Graphics* (2018), 1460–1473. 7, 10



- [WBD\*13] WILLIAMS D. N., BREMER T., DOUTRIAUX C., PATCHETT J., WILLIAMS S., SHIPMAN G., MILLER R., PUGMIRE D. R., SMITH B., STEED C., BETHEL E. W., CHILDS H., KRISHNAN H., PRABHAT P., WEHNER M., SILVA C. T., SANTOS E., KOOP D., ELLQVIST T., POCO J., GEVECI B., CHAUDHARY A., BAUER A., PLETZER A., KINDIG D., POTTER G. L., MAXWELL T. P.: Ultrascale visualization of climate data. *Computer* 46, 9 (September 2013), 68–76. [22](#)
- [WFZ\*15] WANG Y., FAN C., ZHANG J., NIU T., ZHANG S., JIANG J.: Forecast verification and visualization based on gaussian mixture model co-estimation. *Computer Graphics Forum* 34, 6 (2015), 99–110. [7](#), [11](#)
- [WHL18] WANG J., HAZARIKA S., LI C., SHEN H.: Visualization and visual analysis of ensemble data: A survey. *IEEE Transactions on Visualization and Computer Graphics* (2018). [2](#)
- [WHP\*11] WILLIAMS S., HECHT M., PETERSEN M., STRELITZ R., MALTRUD M., AHRENS J., HLAWITSCHKA M., HAMANN B.: Visualization and analysis of eddies in a global ocean simulation. In *Computer Graphics Forum* (2011), pp. 991–1000. [1](#)
- [WJW\*17] WIDANAGAMAACHCHI W., JACQUES A., WANG B., CROSMAN E., BREMER P.-T., PASCUCCI V., HOREL J., ET AL.: Exploring the evolution of pressure-perturbations to understand atmospheric phenomena. In *IEEE Pacific Visualization Symposium (PacificVis)* (2017), pp. 101–110. [1](#)
- [WLSL17] WANG J., LIU X., SHEN H.-W., LIN G.: Multi-resolution climate ensemble parameter analysis with nested parallel coordinates plots. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 81–90. [1](#), [8](#), [10](#), [13](#), [18](#)
- [WMK13] WHITAKER R. T., MIRZARGAR M., KIRBY R. M.: Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2713–2722. [1](#)
- [WP13] WARE C., PLUMLEE M. D.: Designing a better weather display. *Information Visualization* 12, 3–4 (2013), 221–239. [14](#)
- [WPB\*11] WILLIAMS S., PETERSEN M., BREMER P.-T., HECHT M., PASCUCCI V., AHRENS J., HLAWITSCHKA M., HAMANN B.: Adaptive extraction and quantification of geophysical vortices. *IEEE Transactions on Visualization & Computer Graphics*, 12 (2011), 2088–2095. [8](#)
- [WPH\*12] WILLIAMS S., PETERSEN M., HECHT M., MALTRUD M., PATCHETT J., AHRENS J., HAMANN B.: Interface exchange as an indicator for eddy heat transport. In *Computer Graphics Forum* (2012), pp. 1125–1134. [16](#)
- [WPS\*16] WOODRING J., PETERSEN M., SCHMEIßER A., PATCHETT J., AHRENS J., HAGEN H.: In situ eddy analysis in a high-resolution ocean climate model. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 857–866. [1](#), [22](#)
- [WS09] WOODRING J., SHEN H.-W.: Multiscale time activity data exploration via temporal clustering visualization spreadsheet. *IEEE Transactions on Visualization and Computer Graphics* 15, 1 (2009), 123–137. [8](#), [11](#), [13](#)
- [WSL\*14] WONG P. C., SHEN H.-W., LEUNG R., HAGOS S., LEE T., TONG X., LU K.: Visual analytics of large-scale climate model data. In *IEEE 4th Symposium on Large Data Analysis and Visualization (LDAV)* (Nov 2014), pp. 85–92. [7](#), [16](#)
- [WYM08] WANG C., YU H., MA K.: Importance-driven time-varying data visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (Nov 2008), 1547–1554. [11](#)
- [XYM13] XIE J., YU H., MA K.-L.: Interactive ray casting of geodesic grids. In *Computer Graphics Forum* (2013), vol. 32, pp. 481–490. [13](#)
- [YaKSJ07] YI J. S., AH KANG Y., STASKO J. T., JACKO J. A.: Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1224–1231. [9](#)
- [YIT\*18] YANO M., ITOH T., TANAKA Y., MATSUOKA D., ARAKI F.: A comparative 3d visualization tool for observation of mode water. In *IEEE Pacific Visualization Symposium (PacificVis)* (2018), pp. 230–234. [18](#)
- [ZLMM16] ZHANG Y., LUO W., MACK E. A., MACIEJEWSKI R.: Visualizing the impact of geographical variations on multivariate clustering. In *Computer Graphics Forum* (2016), Wiley Online Library, pp. 101–110. [12](#)
- [ZW18] ZHOU L., WEISKOPF D.: Indexed-points parallel coordinates visualization of multivariate correlations. *IEEE Transactions on Visualization and Computer Graphics* 24, 6 (June 2018), 1997–2010. [7](#), [18](#)
- [ZYL\*17] ZHOU Z., YE Z., LIU Y., LIU F., TAO Y., SU W.: Visual analytics for spatial clusters of air-quality data. *IEEE computer graphics and applications* 37, 5 (2017), 98–105. [15](#)