Report from Dagstuhl Seminar 22331

# Visualization and Decision Making Design Under Uncertainty

**Nadia Boukhelifa*[1], Christopher R. Johnson*[2], and Kristi Potter*[3]**

1 **INRAE – Palaiseau, FR.** nadia.boukhelifa@inrae.fr
2 **University of Utah – Salt Lake City, US.** crj@sci.utah.edu
3 **NREL – Golden, US.** kristi.potter@nrel.gov

**Abstract**

This report documents the program and the outcomes of Dagstuhl Seminar 22331 "Visualization and Decision Making Design Under Uncertainty". The seminar brought together 33 researchers and practitioners from different domains concerned with visualization and decision making under uncertainty including visualization, visual analytics, human-computer interaction, artificial intelligence, climate research, geography and geology. The programme was organized in two parts: In the first part which lasted two days, participants gave short talks where they discussed current practices and the uncertainty visualization challenges they encountered in their own research. At the end of day two, participants brainstormed collectively around the main uncertainty visualization research challenges across domains and applications. In the second part, participants voted for the following three main challenges they wished to discuss for the remainder of the seminar (one and a half days): applications, human-centered uncertainty visualization, a design process for uncertainty visualization. Thus three break-out groups were formed to discuss these challenges. Abstracts for the individual talks and the break-out group activities are included in this report.

**Seminar** August 15–19, 2022 – http://www.dagstuhl.de/22331
**2012 ACM Subject Classification** Human-centered computing → Visualization
**Keywords and phrases** Decision making, Uncertainty visualization, Visual Analytics, Visualization
**Digital Object Identifier** 10.4230/DagRep.12.8.1

## 1 Executive Summary

*Nadia Boukhelifa (INRAE – Palaiseau, FR)*
*Christopher R. Johnson (University of Utah – Salt Lake City, US)*
*Kristi Potter (NREL – Golden, US)*



Uncertainty is an important aspect to data understanding. Without awareness of the variability, error, or reliability of a data set, the ability to make decisions on that data is limited. However, practices around uncertainty visualization remain domain-specific, rooted in convention, and in many instances, absent entirely. Part of the reason for this may be a lack of established guidelines for navigating difficult choices of when uncertainty should be added, how to visualize uncertainty, and how to evaluate its effectiveness. Unsurprisingly, the inclusion of uncertainty into visualizations is a major challenge to visualization [1]. As work concerned with uncertainty visualization grows, it has become clear that simple visual

* Editor / Organizer

additions of uncertainty information to traditional visualization methods do not appropriately convey the meaning of the uncertainty, pose many perceptual challenges, and, in the worst case, can lead a viewer to a completely wrong understanding of the data.

The goal of this Dagstuhl Seminar was to bring together experts with diverse knowledge of uncertainty visualization and comprehension toward building a foundation of accessible, practical knowledge that practitioners and researchers alike can rely on in addressing challenges related to uncertainty. Specifically, this seminar brought together leaders in the field of uncertainty visualization and communication, along with experts on quantification and practitioners and domain experts dealing with uncertainty on a daily basis. Drawing on the knowledge of the participants, the seminar worked toward goals of synthesizing disparate findings and approaches from across computer science and related literature, noting current practices surrounding uncertainty, and identifying unsolved problems in common workflows, and areas needing further study.

As a major result from the seminar, the following challenges and research topics in visualization and decision making under uncertainty have been identified:

- Applications,
- Human-centered uncertainty visualization (including how to support "feeling uncertain"),
- A design process for uncertainty visualization,
- Defining terms related to uncertainty,
- Algorithms and uncertainty quantification,
- Software dissemination,
- User studies,
- Ethics of uncertainty (when to include uncertainty information),
- Surveys of uncertainty-aware visual analytics, and
- Teaching uncertainty visualization.

The top three challenges were discussed in depth during this intensive three and a half days Dagstuhl Seminar as part of the break-out groups, and are further discussed in this report. In particular, the break-out groups examined uncertainty visualization research challenges from three complementary perspectives: from an application viewpoint looking at how uncertainty visualization and assessment are used in many domains; from a human-centered perspective considering the needs and information of the viewer; and from a more theoretical stand focusing on the problem space for designing uncertainty visualization.

The seminar ended with a presentation from each group and discussions on the next steps. Interesting research questions and potential solutions were identified during the discussions, and plans were made to continue the collaboration. Details of the individual talks and break-out group discussions are provided in this report.

## References

1 Chris R. Johnson and Allen R. Sanderson. *A next step: Visualizing errors and uncertainty*. IEEE Computer Graphics and Applications, 2003, vol. 23, no 5, p. 6-10.

## 2 Table of Contents

# 3 Overview of Talks

## 3.1 Statistical Analysis for Uncertainty Quantification and Visualization of Scientific Data

*Tushar Athawale (Oak Ridge National Laboratory, US)*


**Joint work of** Tushar Athawale, Chris R. Johnson

Data visualization has become indispensable for efficient interpretation of complex data generated across diverse scientific domains, such as biomedical imaging and meteorology. Many critical decisions directly rely on the quality of data visualizations. Inaccuracies in visualizations cannot be averted due to uncertainties inherent in underlying data and non-linear transformations of data caused by the stages of the visualization pipeline. The uncertainty in the final visualizations can adversely impact the decision-making process. The accurate quantification of uncertainties in data visualizations has, therefore, been recognized as the top research challenge for minimizing risks associated with scientific decisions.

In this talk, I will present the abstract statistical methods for uncertainty visualization and a few uncertainty visualization applications. My main topics of discussion are as follows: 1) Need for uncertainty visualizations, 2) abstract statistical methods for uncertainty quantification, 3) a few applications of uncertainty visualization to key scientific visualization techniques, such as fiber surfaces and Morse complexes, and domain-specific data, e.g., biomedical imaging, 4) open research challenges in uncertainty visualization. Our experimental results relevant to uncertainty visualizations confirm the significance or need for incorporating statistical error analysis into computational models for visualization applications.

## 3.2 A Tentative List of Uncertainty Visualization Research Challenges

*Nadia Boukhelifa*



Uncertainty visualization research has made considerable progress leading to a variety of techniques, algorithms, systems, frameworks and user studies. The goal of this talk is to provide a preliminary list of open problems and challenges that our visualization community has been focused on in the last 20 years. I present findings from a literature survey of 17 papers from 2002 -2022, covering multiple domains including scientific visualisation, information visualization and visual analytics. I focus on surveys, state-of-the-art reports, viewpoint articles and position papers rather than on papers on specific techniques, algorithms, systems or user studies.

The results of this survey shows eight main areas of open challenges related to conceptualisation, evaluation, formalisation and theory, quantification, representation, training and dissemination, uncertainty-aware tools, and user Interaction. Some of the found challenges may have already been solved, and new ones may not yet have been fully documented. There is a need to review progress of the field of uncertainty visualization across domains, and to highlight success stories, long-standing problems as well as emerging and new ones.

## 3.3 Visualization of Climate Simulation Data and related Uncertainty

*Michael Böttinger (DKRZ Hamburg, DE)*



Climate models simulate the most important processes governing the climate system, i.e. the coupled system of atmosphere, ocean, sea ice, land-biosphere and ocean-biogeochemistry. Simulations result in 3D time-dependent multivariate data sets, characterized by high variability at various time scales. Internal variability of the coupled climate system additionally contributes to this noise. However, the high variability reduces the signal-to-noise ratio, thus makes it hard to detect climate change signals. Analyzing and visualizing climate change in the presence of noise is challenging, but with ensemble simulations, the signal-to-noise ratio can be enhanced and the internal climate variability assessed. I present examples from climate change research that show the visualization of robustness in the presence of a highly variable field. However, with respect to the climate change to 2100, the largest uncertainty is in the range of possible evolutions of the socio-economic system. Furthermore, I show visualizations of the CMIP6 multi model ensemble of simulations conducted globally with regard to the 6th IPCC report that capture this range through a range of scenarios describing different socio-economic development pathways. Finally, I briefly present recent collaborative work with Gerik Scheuermann's group to highlight the challenges in the visualization of uncertain topology-based features for highly variable complex phenomena such as the North Atlantic Oscillation and its evolution in a changing climate.

### References

1 Vietinghoff, D., Heine, C., Böttinger, M., Maher, N., Jungclaus J., and Scheuermann, G. *Visual Analysis of Spatio-Temporal Trends in Time-Dependent Ensemble Data Sets on the Example of the North Atlantic Oscillation.* 2021 IEEE 14th Pacific Visualization Symposium (PacificVis), 2021, pp. 71-80, doi:10.1109/PacificVis52677.2021.00017

2 Vietinghoff, D., Böttinger, M., Scheuermann, G. and Heine, C *Detecting Critical Points in 2D Scalar Field Ensembles Using Bayesian Inference* IEEE 15th Pacific Visualization Symposium (PacificVis), 2022, pp. 1-10, doi:10.1109/PacificVis53943.2022.00009.

## 3.4 The Impossibility of Zero: Effects of Individual Differences in Medical Decision Making

*Remco Chang (Tufts University – Medford, US)*



Making decisions that might affect a person's long-term physical wellbeing can be difficult and stressful. As most medical diagnosis contains some amount of uncertainty (including type I and type II errors), it is often up to a patient to assess their own comfort level with different treatment options. In this talk, I present three challenges relating to medical decision making through the perspective of the patients, namely risk communication, reasoning with conditional probability, and visualization design for decision making.

First, I present a design study of a visualization tool for communicating a patient's prostate cancer risk. After interviewing 6 prostate cancer patients and two urologists, we iteratively designed the visualization based on the participants' feedback. Our takeaways from this design study include: (1) prostate cancer patients (who tend to be older men) have trouble using even basic visualizations (e.g. bar chart, stacked area chart, scatterplot, etc.). Text explanations that accompany the visualizations are a must. (2) Emotion and stress can affect a patient's ability to reason about their diagnosis. After receiving a positive diagnosis, a patient often has limited cognitive capacity to think through the diagnosis rationally. (3) Most Patients' first question after receiving a positive diagnosis is "how much time do I have left," suggesting that there's an order to the presenting of information that can best meet the patients' decision-making needs.

Second, I present an experimental study on people's ability to reason about their diagnosis as conditional probabilities. Most screening tests contain some amount of uncertainty, in particular as type I and type II errors. When a patient is told that they have a positive diagnosis for a disease, it is often up to the patient to reason through these probabilities to assess what their "true risks" are. In our experiment, we tested 6 visualization designs that were accompanied by text explanations. Our initial analysis of the results found no statistical significance between the effectiveness of the 6 visualizations. However, when the participants were stratified based on their spatial ability scores (as measured using the paper-folding test), we found that some of the visualizations are very effective (near 100% accuracy) for the participants with high spatial abilities. Unfortunately, we found no visualization that was helpful for the participants with low spatial abilities.

Lastly, I discuss the challenges in designing visualizations for helping patients make difficult medical decisions. For example, a patient might not perceive any difference between a diagnosis with 30% or 31% of having a disease. However, when the difference is between 0% and 1% chance of having a disease, the same difference of 1% becomes more significant to a patient as it represents "not having a disease" versus "possibly having a disease." A visualization will need to incorporate individuals' risk perception and risk tolerance utility curves to best support their decision making process.

## 3.5 Underthinking Uncertainty Visualization

*Michael Correll (Tableau Software – Seattle, US)*



Uncertainty visualization is viewed as a hard problem. Sources of these difficulties include complexity and disagreement around how uncertainty is modeled or quantified and clashes between between idealized forms of decision-making and the actual behavior of human beings. It is true that these are problems. But we can't wait for statisticians and psychologists to settle all of their internal disputes on these topics; people have decisions to make today. What we can do, however, is find solutions that are likely to be generally good enough for many practical purposes.

In this talk I will introduce a framework for ways to address uncertainty without having to think too hard, specifically around leveraging the ability of people to estimate statistical properties in visualizations without additional scaffolding, and the ability of visualization designers to "nudge" these estimates to align with statistical models of decision-making without being dogmatic or domineering. This is good news for uncertainty visualization as a discipline in that it does not require either designers or viewers of visualizations to be perfectly rational statistical deities to get their work done, but perhaps bad news in that we now have to do much more work as a field to build a deeper understanding of graphical perception for "fuzzier" tasks, take stronger stances around desired behavior from viewers of visualizations, and to better integrate statistical models, models of inference, and rhetorical goals into our design thinking.

## 3.6 Uncertainty in Public Policy Decision Making

*Stephanie Deitrick (Arizona State University – Tempe, US)*



Public policy decision makers leverage both qualitative and quantitative data as part of their decision-making processes. With increased interests in science-based information and leveraging data for their decisions, agencies are often expanding their workforce to include more data scientists and partnering with researcher on a variety of topics. While policy makers understand that data are uncertain at some level, that may not be something they explicitly consider as part of how they currently leverage data.

Since data are often communicated through visualization, such as maps and charts, should uncertainty be part of that communication? Would it produce better or more informed decisions?

## 3.7 Uncertainty-aware Visual Analytics

*Christina Gillmann (Universität Leipzig, DE)*



Visual analytics has been successfully applied to a variety of applications also in terms of uncertainty analysis. Unfortunately, the visual analytics process does not include a mechanism to systematically handle uncertainty. In order to solve this issue, we developed the concept of uncertainty-aware visual analytics. Therefore, an extension of the classic visual analytics cycle is achieved that includes the quantification of uncertainty in each component, the exchange of analysis and visualization approaches in general by uncertainty-aware options and the introduction of provenance to monitor the accumulation and propagation of uncertainty throughout the visual analytics cycle. In order to create uncertainty-aware visual analytics cycles for particular applications, we determined a workflow that consists of 5 steps that constructs an uncertainty-aware visual analytics cycle starting from the classic approach. The procedure is based on a developed taxonomy of uncertainties that allow to understand the nature of different uncertainty events and their effect on the visual analytics cycle.

### 3.8 Visualizing uncertainty in digital geologic map databases

*Amy Gilmer (USGS – Denver, US) and Kathleen Warrell (UCAR – Boulder, US)*



The geologic map remains the primary tool geologists use to model and communicate what we know about Earth's surface. All geologic models contain some level of uncertainty, but this uncertainty is rarely incorporated in traditional geologic maps, potentially limiting application by decision makers. Even as our geological depictions have migrated to digital geologic map databases, our map symbology has largely remained the same as that used on traditional paper maps. While varying dash length for contacts and faults may convey a relative sense of uncertainty to experienced users, it does not convey meaning to the nonexpert user. Cartographic uncertainty visualizations are an effective way to communicate how well we know what and where something is.

The adoption of the Geologic Map Schema (GeMS) standard for geologic maps has enabled geologists to capture feature-level metadata, including location uncertainty, as well as feature identity and existence confidence. To visually communicate the underlying locational uncertainty in the USGS Intermountain West geological framework database, we have developed an ArcGIS Python toolbox that extracts existing location confidence data from feature attributes, and then buffers and aggregates the uncertainty across a tessellation grid. The tessellation grid can then be visualized by any of the statistical fields generated. This toolbox can be applied to any geologic map database adhering to the GeMS format to produce visualizations summarizing uncertainty. While there is still much we can do to refine how we quantify uncertainty in mapping geologic features, this type of visualization, when provided alongside the geologic map data, summarizes the uncertainty without requiring the user to understand the nuances of traditional map cartography. Additionally, this quantitative approach can help identify areas characterized by high levels of uncertainty, potentially a result of low-resolution map data, that can be used for geologic mapping needs assessments and to better inform end users to limit improper use of the map data.

### 3.9 Summarization, Uncertainty, Estimation...: Models as a basis for visualization

*Michael Gleicher (University of Wisconsin-Madison, US)*



Models are part of most (if not all) use of data. However, they are often hidden or implicit. We often expect viewers to figure out what they are, estimate their parameters, and apply them correctly to achieve their goals. I argue that models should be a first class citizen in how we help people work with data. Many problems, including uncertainty, seem to be made worse because the models are hidden. Many concepts, such as summarization, estimation, and uncertainty are often conflated, especially when models are hidden. My conjecture is that by having a better way to include models in our thinking and by de-conflating the key terms, we can better discuss, design, and evaluate tools to help people work with data.

## 3.10 Uncertainty in Definition

*Hans-Christian Hege (Zuse-Institute Berlin, DE)*



Facts about the world are mainly represented linguistically. The building blocks of language are concepts, both concrete and abstract. Concepts help us humans to organize, understand and explain the world. We use them in cognitive processes such as categorization, reasoning, and decision-making, as well as in explanation and communication. An important task of data science/visualization is to connect the world of data and the world of concepts by finding equivalents of the concepts in the data.

However, concepts are only defined in language and often imprecisely. This leads to "uncertainty of definition". Metaphorically speaking, concepts are not points in conceptual space, but rather regions with blurred boundaries. Examples: What exactly is a vortex in a flow? What exactly is the spatial extent of a vortex? Which patients are considered to have died from COVID-19 as opposed to patients who died with COVID-19? What exactly is an epidemic wave and what is not? Which atmospheric phenomenon is a hurricane and which is not? Almost every statistic or visualization is preceded by such questions. Different answers are possible, but they lead to different results: the definition uncertainties propagate into the results. If we take into account the uncertainties in definition, we get ensembles of results.

We should be aware of this type of uncertainty, capture it and its propagation into the results, communicate it, and reduce it. Visualization can help with the latter by showing the variance that results from different definitions of the concepts.

### References

1 Natalia Mikula, Tom Dörffel, Daniel Baum, Hans-Christian Hege. *An Interactive Approach for Identifying Structure Definitions*. Computer Graphics Forum, 41:3, pp. 321-332 (2022), DOI: 10.1111/cgf.14543

## 3.11 Visualization and Analysis of XCT Data – Decision Making under Uncertainty

*Christoph Heinzl (Universität Passau, DE)*



Visualization and analysis of "rich" X-ray computed tomography (XCT) data has become highly attractive for boosting research endeavors in the materials science domain. On the one hand, XCT allows to generate detailed and cumulative data of the specimens under investigation in a non-destructive way. On the other hand, through the conception, the development, and the implementation of novel, tailored analysis and visualization techniques, in-depth investigations of complex material systems turned into reality.

This talk presents contributions to computer science in terms of design studies, methods, and techniques, which are advancing visual analysis and visualization for enabling insights into "rich" XCT data. The introduced methods and techniques focus on three distinct technical areas of visual analysis and visualization of XCT data, which are interactive visualization of spatial and quantitative data, visual parameter space analysis of respective data processing and visualization pipelines, uncertainty and sensitivity analysis. For each area, the problem statements, important research questions to be solved as well as some of the author's contributions thereto are discussed.

**References**

1 A. Reh, C. Gusenbauer, J. Kastner, E. Gröller, C. Heinzl. *MObjects – A Novel Method for the Visualization and Interactive Exploration of Defects in Industrial XCT Data.* IEEE Transactions on Visualization and Computer Graphics, Vol. 19, No. 12, 2013, pp. 2906-2915
2 A. Reh, A. Amirkhanov, J. Kastner, E. Gröller, C. Heinzl. *Fuzzy Feature Tracking: Visual Analysis of Industrial 4D XCT Data*, In Computers & Graphics vol. 53, Part B, 2015, pp. 177–184
3 B. Fröhler, T. Möller, C. Heinzl. *GEMSe: Visualization-Guided Exploration of Multi-channel Segmentation Algorithms.* Computer Graphics Forum, vol. 35, 2016, pp. 191-200
4 B. Fröhler, T. Elberfeld, T. Möller, HC Hege, J. Weissenböck, J. De Beenhouwer, J. Sijbers, J Kastner, C. Heinzl. *A Visual Tool for the Analysis of Algorithms for Tomographic Fiber Reconstruction in Materials Science.* Computers Graphics Forum vol. 38 (3), 2019, pp. 273-283
5 J. Weissenböck, B. Fröhler, E. Gröller, J. Kastner, C. Heinzl. *Dynamic Volume Lines: Visual Comparison of 3D Volumes through Space-filling Curves.* IEEE Transactions on Visualization and Computer Graphics, vol. 25 (1), 2019, pp. 1040-1049
6 A. Amirkhanov, C. Heinzl, C. Kuhn, J. Kastner, E. Gröller. *Fuzzy CT Metrology: Dimensional Measurements on Uncertain Data.* Proceedings of the 29th Spring Conference on Computer Graphics, 81-90, 2013
7 B. Fröhler, T. Elberfeld, T. Möller, H.C. Hege, J. Maurer, C. Heinzl. *Sensitive vPSA– Exploring Sensitivity in Visual Parameter Space Analysis.* arXiv preprint arXiv:2204.01823

## 3.12 Designing, de-“bias”ing, and de-probabilizing uncertainty visualization

*Matthew Kay (Northwestern University – Evanston, US)*



I discuss three challenges in uncertainty visualization: (1) how do we design uncertainty visualizations systematically? (2) how do we (and should we) de-bias uncertainty visualizations? (3) how do we visualize possibilistic and qualitative forms of uncertainty?

## 3.13 Centering Uncertainty on People

*Miriah Meyer (Linköping University, SE)*



From a perspective of data as a situated perspective – one that is inherently partial and incomplete – knowledge about the shortcomings of data is often known by domain experts. In recent work we propose a framing of this knowledge as data hunches, and argue that hunches are a source of qualitative uncertainty. Acknowledging and valuing the hunches people bring to visual analysis opens new opportunities to design visualization tools that support people in externalizing and communicating their hunches.

## 3.14 Actionable Uncertainty Visualization

*Kristi Potter (NREL – Golden, US)*



Ensemble simulations capture the variability present in the predictions of future states by combining multiple runs of a computational model with different parameter settings. Datasets derived from ensemble simulations are often quite large and complex, making it hard to create visualizations that facilitate decisions, particularly for people not intimately involved with the scientific domain or the creation of the dataset. An example comes from the renewable energy space, where improvements to the electrical grid to facilitate the large-scale reduction of carbon emissions involves making decisions on highly complex systems. Traditional methods for uncertainty visualizations primarily focus on the challenge of visually presenting large-scale, high-dimensional datasets in an exploratory manner. However those approaches do not facilitate decision making by non-experts, such as policy-makers, who may not know enough about the computational system to appropriately choose appropriate parameter settings to achieve a desirable outcome. In this talk I will discuss ideas for distilling down the parameter space by importance, annotating contextual information needed for better understanding, and designing a visualization tool that is streamlined for decision-making.

## 3.15 A design theory for uncertainty visualization?

*Maria Riveiro (Jönköping University, SE)*



Despite the large volume of research on uncertainty visualization, we do not fully understand the impact of uncertainty visualization on decision-making. There is evidence of both positive and negative effects of visually depicted uncertainty on decision-making.

This talk presents examples of evaluations carried out with practitioners in various application areas, including autonomous driving, air traffic risk assessment and maritime surveillance. I summarise the effects of the uncertainty visualizations provided on the users and their decision-making processes in these evaluations.

Finally, we discuss the need for a design theory/space of uncertainty visualization and elaborate on the multiple dimensions/variables that such a design space should have.

**References**

1 Helldin, T., Falkman, G., Riveiro, M. and Davidsson, S. (2013) Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. Proc. 5th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 13), Eindhoven, The Netherlands.
2 Riveiro, M., Helldin, T., Falkman, G. and Lebram, M. (2014) Effects of visualizing uncertainty on decision-making in a target identification scenario, Computers & Graphics, Volume 41, Pages 84-98, Elsevier, ISSN 0097-8493.
3 Riveiro, M. (2016). Visually supported reasoning under uncertain conditions: Effects of domain expertise on air traffic risk assessment. Spatial Cognition & Computation: An Interdisciplinary Journal, vol. 16(2), pp. 133-153. Taylor & Francis.

4 Kinkeldey, C., MacEachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations. Cartography and Geographic Information Science, 44(1), 1-21.

## 3.16 Critical Points of an uncertain Scalar field

*Gerik Scheuermann (Universität Leipzig, DE)*



Critical points like extrema or saddles are a well established concept in (deterministic) scalar field visualization. There is strong practical interest, a clear mathematical concept in continuous and discrete settings, and corresponding algorithms, including implementations in commercial systems. Looking at uncertain scalar fields, formally described as smooth stochastic processes, practically often given as ensembles over a common grid, the situation changes. The concept of a "critical point" is not exactly defined. Two major definitions are "critical point of the (deterministic) mean field" or "probability distribution of critical points in a sample from the stochastic process/ensemble". The choice of concept definition has effects on the visualization and its interpretation, like "a maximum being multiple (significant) maxima in one sample and no significant maximum in another case". Also, the sampling quality of the ensemble should somehow be integrated into the visualization. The talk concern these issues. Looking at the distribution definition, I show how to infer critical point distributions from ensembles using Bayesian Inference. Looking at the mean field definition, I will discuss how bootstrapping allows to reason about the sampling quality to derive significant results. Finally, the talk shows how this allows to decide two practical questions regarding the future of the north atlantic oscillation (NAO) depending on Climate Change. (NAO describes they interplay between Iceland Low and Azore High – which is the most dominant factor in European winter weather.) We derive that the centers of action of both pressure systems move substantially depending on the global warming, and that the IceLand Low will most likely see a split into two centers of action in the extreme scenarios. The work was done with Dominik Vietinghoff, Christian Heine, and Michael Böttinger.

### References

1 Vietinghoff, Heine, Böttinger, Maher, Jungclaus, Scheuermann. *Visual analysis of spatio-temporal trends in time-dependent ensemble data sets on the example of the north atlantic oscillation.* IEEE PacificVis 2021 Proceedings, 71-80, 2021
2 Vietinghoff, Heine, Böttinger, Scheuermann. *An Extension of Empirical Orthogonal Functions for the Analysis of Time-Dependent 2D Scalar Field Ensembles.* IEEE PacificVis 2021 Short Papers, 46-50, 2021.
3 Vietinghoff, Böttinger, Scheuermann, Heine. *Detecting Critical Points in 2D Scalar Field Ensembles Using Bayesian Inference.* IEEE PacificVis 2022 Proceedings, 1-10, 2022.

## 3.17 Uncertainty in Time Series and Geographic Data

*Johanna Schmidt (VRVis – Wien, AT)*


**URL** https://www.digi-hydro.com/

Design decisions must be made to make data visual, and modifications to the data are needed. Data modification includes reconstruction, resampling, filtering, and aggregation. In one of our projects, we have to deal with time series data recorded from sensors installed in hydropower machines. The project's purpose is to better understand which sensors can give information about the current state of the hydropower machine. This needs to be done with exploratory data analysis. It is not yet known to the mechanical engineers which sensors will be descriptive for detecting certain stages during machine operation. However, the data is large (approximately 30 TB of data), and it is impossible to analyze the raw data in this case. We, therefore, need to apply resampling and filtering to the data, which introduces uncertainty in the analysis the mechanical engineers should be informed about. In the case of geological data, reconstruction (of point cloud data) and 3D rendering introduce uncertainty in the data representation. When performing analyses, methods like plane fitting are also not wholly accurate. This uncertainty in the data and how it is presented to the users needs to be communicated, as both user groups (mechanical engineers and geologists) highly depend on detailed analysis results.

## 3.18 Quantifying and Visualizing Uncertainty in Medical Image Segmentation

*Thomas Schultz (Universität Bonn, DE)*


**Joint work of** Shekoufeh Gorgi Zadeh, Thomas Schultz
**Main reference** Shekoufeh Gorgi Zadeh, Maximilian W. M. Wintergerst, Thomas Schultz: "Intelligent interaction and uncertainty visualization for efficient drusen and retinal layer segmentation in Optical Coherence Tomography", Comput. Graph., Vol. 83, pp. 51–61, 2019.
**URL** http://dx.doi.org/10.1016/j.cag.2019.07.001

Neural networks have greatly increased the accuracy in many medical image segmentation tasks, and have been successfully deployed for large-scale image analysis. However, fully automated results are still not reliable enough to be trusted blindly in applications where segmentation quality might be critical to the well-being of individuals. Using an application example in ophthalmology, we demonstrate that visualizing the uncertainty in neural network based segmentations, and providing uncertainty-aware tools for segmentation editing, can make it more time efficient to identify and correct remaining segmentation errors. We also discuss the important open question of reliable uncertainty quantification in an out-of-distribution setting, for example when processing images that have been acquired with a different scanner, and we mention strategies for approaching that problem.

## 3.19 Visualizing the Uncertainty in Image Analysis – Previous work and new opportunities

*Brian Summa (Tulane University – New Orleans, US)*



In this talk, I give an overview of my research in the visualization of uncertainty in scientific data, while highlighting new opportunities for uncertainty quantification in topological data analysis (TDA) or in accounting for uncertainty due to human variability.

## 3.20 Uncertainty and Trustworthy AI

*Stefan Hagen Weber (Siemens – München, DE)*


**Joint work of** Daniela Oelke, Stefan H. Weber.
**Main reference** Daniela Oelke, Stefan H. Weber: "Line Density Plots – Visualizing uncertainty in forecast ensembles". Talk on IEEE VIS 2018 VisInPractice event.

Visual representations of density under uncertainty have been explored for geographic data, scatterplots, line charts or parallel coordinates. Ensemble forecasting is a widely known application for uncertainty visualization. Often end users have specific requirements and tasks for the visualization, e.g.

- each forecast (line) should be visible and interactable,
- the resulting chart should not be overcrowded
- the uncertain space between ensembles should be filled by upsampling
- the uncertainty should be made visible
- Identifying outliers is as important as spotting the main trend

All these requirements can be realized by a novel technique for generating density representations for line charts that is visually and computationally scalable with respect to the number of lines that are shown. In contrast to alternative kernel-based density representations, it also keeps the course of the lines visible unless the local density is very high. Points are on top of each other (or crossing lines) represent the (un)certainty (density surface) and are mapped to color. A smoothed representation with an upsampling effect is done by adding a "glow" around the lines. This glow is implemented by decreasing the alpha value with increasing distance from the line. The amount of glow at a certain distance is determined by the shape of the specific kernel function. The kernel width determines the extension of the glow around the line. Some considerable effort was spent to design and implement the visualization and integrate it into a commercial system (linking & brushing), to fulfill the end user's requirements. The result was evaluated together with the end user who wanted to gain more insight into their ensemble forecasts to answer the question "When is the best time to buy oil?". The end user first inspected the overall distribution pattern in time over all ensemble members. They used the median to separate the higher half of the distribution from the lower. They got immediate insights regarding the trend and distribution. The visualization was more explored, and the end user provided very positive feedback. Our expectation was of course that the result will be used from now on. However, the opposite happened. The end user so far was not sure if he can trust the ensemble forecast method. With our visualization he gained trust in the AI after a few hours. From that point on he was fine with a simple KPI: “Just tell me when to buy oil”. Lessons learned: A lot of effort was spent for a visualization that was only used a few times. You might argue that it was not worth the effort. However, it turned out that the initial task of understanding the uncertainty aspect was only the first step. The final effect was that the user increased his trust in the AI. Trustworthy AI is a valuable asset. It might be a frustrating experience but increasing trust in AI is a huge long-term benefit. Even if the visualization was only one short part of the journey. Showing uncertainly in a proper way can increase trust.

## 3.21 Overcoming Uncertainties in Molecular Visualization

*Thomas Wischgoll (Wright State University – Dayton, US)*


**Joint work of** Thomas Wischgoll, Christina Gillmann, Robin Maack, Matthew Marangoni
**URL** https://avida.cs.wright.edu

Uncertainties are difficult if not impossible to avoid. Capturing data from the analog world almost always results in some form of uncertainty. The amount of uncertainty depends on the method of measurement and its accuracy. When visualizing data that has some associated uncertainty, it is essential to properly process and convey such uncertainty and especially the amount of uncertainty keeping in mind that additional processing steps can amplify the uncertainty. There are various sources of uncertainty, such as numerical limitations or limitations of the capture device. However, there are other sources of uncertainty. Some of these uncertainties stem from model assumptions or limitations of how we translate natural specimens to 3D representations. Molecular structures are one example of this. This talk will illustrate this further and point to some of the solutions.

## 3.22 Uncertainty Visualization of Health Data

*Liang Zhou (Peking University, CN)*



Health science relies on a wide range of different types of data. There, uncertainty is ubiquitous and is aware by health science experts. Uncertainty visualization is, therefore, important and could potentially aid decision making. In this talk, I will introduce my own research work on new visualization techniques for representative health data. These examples focus on uncertainty visualization of ensemble medical imaging data, local correlation and subspace visualization for multidimensional data, and perceptual enhancement for visualization images. I will also discuss works on visual analytics of health data with uncertainties from missing data. Finally, I will discuss uncertainty challenges that I identified in the various types of health data.

# 4 Working groups

## 4.1 Applications

*Tushar Athawale (Oak Ridge National Laboratory, US), Michael Böttinger (DKRZ Hamburg, DE), Amy Gilmer (USGS – Denver, US), Hans-Christian Hege (Zuse-Institute Berlin, DE), Christoph Heinzl (Universität Passau, DE), Christopher R. Johnson (University of Utah – Salt Lake City, US), Gerik Scheuermann (Universität Leipzig, DE), Johanna Schmidt (VRVis – Wien, AT), Thomas Schultz (Universität Bonn, DE), Jarke J. van Wijk (TU Eindhoven, NL), Stefan Hagen Weber (Siemens – München, DE), and Xiaoru Yuan (Peking University, CN)*



Uncertainty visualization and assessment are used in many domains, including medical applications, non-destructive testing, industrial AI, geology, renewable energies, and climate research. The way uncertainty is used by users in these domains differs depending on the required tasks and the data used. As an outcome of this working group, we identified success stories of published or successfully applied in seven domains. Based on these success stories, we identified common open challenges and research questions that will be worth working on: (i) Uncertainty could be viewed from a mathematical point of view, looking at stochastic processes, statistics, correlations, and similar. This would also enable the quantification of uncertainty in different domains. (ii) The different sources of uncertainty need to be discussed – whether they are similar in different domains and to which degree they depend on tasks and data. Also, the terminology used in different domains to describe uncertainty differs. (iii) An interesting question is to differentiate between visualization applications where uncertainty visualization is needed and where not. It might depend on the task, and the types of questions users have, whether it makes sense to include uncertainty in a visual representation or not. (iv) Visualizing uncertainty also relates to perceptual issues, describing how well uncertainty can be perceived using different encodings. (v) As a wrap-up, it would be interesting to find out how far uncertainty visualization is already used in commercial software.

## 4.2 What's the point?: Focusing on the human in uncertainty vis

*Nadia Boukhelifa (INRAE – Palaiseau, FR), Michael Correll (Tableau Software – Seattle, US), Stephanie Deitrick (Arizona State University – Tempe, US), Matthew Kay (Northwestern University – Evanston, US), Miriah Meyer (Linköping University, SE), Kristi Potter (NREL – Golden, US), Paul Rosen (University of Utah – Salt Lake City, US), and Regina Maria Veronika Schuster (Universität Wien, AT)*



Current techniques around uncertainty visualization are often oriented around statistical models, with the efficacy of an uncertainty visualization viewed as either how accurately the viewer is able to retrieve or estimate specific model values, or how well the viewer's decision-making aligns with that of some normative model of utility or decision quality. This model-driven rather than human-driven perspective introduces several key limitations when designing or evaluating uncertainty visualizations. For one, it elides many aspects of decision-making under uncertainty that are not amenable to tidy quantification, such as situated or implicit knowledge. For another, it ignores psychological, sociological, rhetorical, or ethical aspects of presenting uncertainty information. We propose a human-centered view of uncertainty visualization in which the needs and information of the viewer, rather than backing statistical or inferential models, are given precedence.

In the human-centered view of uncertainty visualization, viewers are neither rote reciters of p-values, nor conditioned to mimic the actions of a statistical test. Rather, they have many goals, including being able to audit or justify their decisions, build appropriate trust in the data source and designers, integrate their own mental models and domain knowledge with existing data, or even just walk away satisfied that they made a reasonable decision given the information they had to hand. In this paper, we show how existing frames around uncertainty visualization may fail to result in designs that accomplish these goals, and present both existing strategies for better integrating the human in the uncertainty visualization design process as well as open problems in visualization research.

## 4.3 A Problem Space for Designing (Uncertainty) Visualizations

*Maria Riveiro (Jönköping University, SE), Remco Chang (Tufts University – Medford, US), Oliver Deussen (Universität Konstanz, DE), Christina Gillmann (Universität Leipzig, DE), Michael Gleicher (University of Wisconsin-Madison, US), and Tatiana von Landesberger (Universität Köln, DE)*


**Main reference** Hans-Jorg Schulz, Thomas Nocke, Magnus Heitzler, and Heidrun Schumann. 2013. A Design Space of Visualization Tasks. IEEE Transactions on Visualization and Computer Graphics, 19(12):2366–2375.

Visualization researchers seek appropriate abstractions to help us design, analyze, organize, and evaluate the things we create. Information visualization literature has many task structures (taxonomies, typologies, etc.), design spaces, and related frameworks. In this working group, we discussed current frameworks for designing visualizations, and we considered developing a new problem space that complements the existing ones by focusing on the needs that a visualization is meant to solve. Briefly, the proposed problem space is based on the earlier work by [1], considering the 5Ws and H (who, why, what, where, when and how).

We believe that this problem space provides a valuable conceptual tool for designing and discussing visualizations, including uncertainty visualisations.

### References

1 Hans-Jorg Schulz, Thomas Nocke, Magnus Heitzler, and Heidrun Schumann. 2013. A Design Space of Visualization Tasks. IEEE Transactions on Visualization and Computer Graphics, 19(12):2366–2375.

## Participants

- Michael Böttinger
DKRZ Hamburg, DE
- Nadia Boukhelifa
INRAE – Palaiseau, FR
- Remco Chang
Tufts University – Medford, US
- Oliver Deussen
Universität Konstanz, DE
- Christina Gillmann
Universität Leipzig, DE
- Amy Gilmer
USGS – Denver, US
- Michael Gleicher
University of Wisconsin-Madison, US
- Hans-Christian Hege
Zuse-Institute Berlin, DE
- Christoph Heinzl
Universität Passau, DE
- Miriah Meyer
Linköping University, SE
- Kristi Potter
NREL – Golden, US
- Maria Riveiro
Jönköping University, SE
- Gerik Scheuermann
Universität Leipzig, DE
- Johanna Schmidt
VRVis – Wien, AT
- Thomas Schultz
Universität Bonn, DE
- Regina Maria Veronika Schuster
Universität Wien, AT
- Jarke J. van Wijk
TU Eindhoven, NL
- Tatiana von Landesberger
Universität Köln, DE
- Stefan Hagen Weber
Siemens – München, DE

## Remote Participants

- Tushar Athawale
Oak Ridge National Laboratory, US
- Mehdi Chakhchoukh
INRAE Paris, FR
- Michael Correll
Tableau Software – Seattle, US
- Stephanie Deitrick
Arizona State University – Tempe, US
- Jake Hofman
Microsoft – New York, US
- Amit Jena
Indian Institute of Technology Bombay, IN
- Christopher R. Johnson
University of Utah – Salt Lake City, US
- Matthew Kay
Northwestern University – Evanston, US
- Robert Lempert
RAND – Santa Monica, US
- Paul Rosen
University of Utah – Salt Lake City, US
- Han-Wei Shen
Ohio State University – Columbus, US
- Brian Summa
Tulane University – New Orleans, US
- Kathleen Warrell
UCAR – Boulder, US
- Thomas Wischgoll
Wright State University – Dayton, US
- Xiaoru Yuan
Peking University, CN
- Liang Zhou
Peking University, CN