Uncertainty-aware Visual Analytics - Scope, Opportunities and Challenges

Robin Georg Claus Maack

University of Kaiserslautern: Technische Universität Kaiserslautern

Gerik Scheuermann

Leipzig University: Universität Leipzig

Hans Hagen

TU Kaiserslautern: Technische Universität Kaiserslautern

Jose Tiberio Hernández Peñaloza

Universidad de Los Andes

Christina Gillmann (✉ gillmann@informatik.uni-leipzig.de)

Leipzig University: Universität Leipzig

Research Article

Keywords: Visual Analytics, Uncertainty Analysis, Uncertainty-aware Visualization

Posted Date: December 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1177485/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Uncertainty-aware Visual Analytics

Scope, Opportunities, and Challenges

Robin G. C. Maack^{1*}, Gerik Scheuermann², Hans Hagen¹, Jose Tiberio Hernández Peñaloza³ and Christina Gillmann²

¹Computer Graphics and HCI Group, University of Kaiserslautern, Erwin-Schrödinger-Straße 52, Kaiserslautern, 67663, Rhineland-Palatinate, Germany.

²Image and Signal Processing Group, Leipzig University, Augustusplatz 10, Leipzig, 04109, Saxony, Germany.

³IMAGINE Group, Universidad de los Andes, Cra. 1 18a-12, Bogotá, Cundinamarca, Colombia.

*Corresponding author. E-mail: maack@rhrk.uni-kl.de;

Contributing authors. E-mails: scheuermann@informatik.uni-leipzig.de;
hagen@cs.uni-kl.de; jhernand@uniandes.edu.co; gillmann@informatik.uni-leipzig.de;

Abstract

In many applications, Visual Analytics(VA) has developed into a standard tool to ease data access and knowledge generation. Unfortunately, many data sources, used in the VA process, are affected by uncertainty. In addition, the VA cycle itself can introduce uncertainty to the knowledge generation process. The classic VA cycle does not provide a mechanism to handle these sources of uncertainty. In this manuscript, we aim to provide an extended VA cycle that is capable of handling uncertainty by quantification, propagation, and visualization. Different data types and application scenarios that can be handled by such a cycle, examples, and a list of open challenges in the area of uncertainty-aware VA are provided.

Keywords: Visual Analytics, Uncertainty Analysis, Uncertainty-aware Visualization

1 Introduction

Huge amounts of data are created every day that need to be properly analyzed. This need drove the development of a new data processing concept, called Visual Analytics(VA) [17]. It states that analytic reasoning should be supported by interactive visual interfaces that allow users to explore datasets according to their needs and perform decision-making tasks.

Keim et al. [53] described the VA process as a graph consisting of four major components (dataset, hypothesis, visualization, and insight). These components are connected by functions that allow to transform and analyze given input datasets while creating new insights, as shown in Figure 1. In many applications, VA is applied as a standard tool to find novel insights in datasets and perform decision-making [57].

The role of uncertainty in the VA process has been described by Sacha et al. [82]. It mainly

Fig. 1: Visual analytics cycle introduced by Keim et al. [53]. The cycle consists of four major components (Dataset, Hypothesis, Visualization, and Insight) that are connected, including a feedback loop that inserts insights back into the VA process.

states that uncertainty has to be properly communicated to allow decision-makers to perform their tasks properly. Keim et al. [52] stated that the integration of uncertainty is one of the major challenges in VA. By now, many applications have started incorporating uncertainty analysis in their VA tools. Unfortunately, an overall description of uncertainty-aware VA that assists visualization researchers in identifying the necessary steps that need to be fulfilled, is not available yet. Therefore such a description of uncertainty-aware VA is provided in Section 2.

The data and models used in real-world applications are often affected by uncertainty due to a variety of effects such as data incompleteness, imprecise measurements, reconstruction artifacts, or model imprecision [35], as shown in Section 3. Each component in the VA cycle can be affected by uncertainty that needs to be quantified, propagated, and communicated throughout the VA cycle. Although a variety of approaches in different applications for uncertainty-aware VA exist, there is a lack of a unified description that defines the necessary steps to achieve this goal.

This forms the motivation of the presented work. We aim to provide a general description of an uncertainty-aware visual analytics cycle (see Section 4) by revisiting the VA definition of Keim et al. [53], extending it to provide uncertainty-aware quantification and transformation approaches along the VA cycle. Furthermore, we added novel connections and steps in the VA

process, when required, to achieve uncertainty-awareness. In its entirety, this formulates an uncertainty-aware description of the VA process.

We show the applicability of the presented approach by providing scenarios for potential use-cases, utilizing various datasets and application scenarios (see Section 5). Based on the presented description of an uncertainty-aware visual analytics cycle, we will identify components and connections that require further research to be properly defined (see Section 6).

In this work, we contribute:

- A quick guide to uncertainty analysis
- Refinement of the VA cycle to achieve uncertainty-awareness
- A summary of potential data sources and application scenarios for uncertainty-aware VA
- A summary of open problems in uncertainty-aware VA

This work explicitly does not aim to contribute the following:

- A rating of uncertainty-aware visual analytics approaches - we list examples
- A state of the art analysis of uncertainty-aware visual analytics approaches - we define necessary components
- A description of the construction of such a cycle - we explain it

2 Related Work

In the following, we aim to summarize approaches that target the challenge of including uncertainty-aware visual analytics for specific types of data or scenarios (see Section 2.1). In addition, there exist a variety of works in related disciplines such as sensitivity analysis (see Section 2.2). In this section, we aim to summarize the most related work in the context of the given approach.

2.1 Uncertainty-aware Visual Analytics

Bertini and Roberts [69] stated that a classic visualization approach is not sufficient to deal with uncertainty. They proposed that uncertainty-aware VA is required as it provides the user with an approach to tackle data not restricted to visualization. We use this statement as a starting point

for the presented description of uncertainty-aware VA.

Based on the classic definition of the VA cycle by Keim et al. [53], a massive amount of VA applications have been developed and successfully applied. Still, the field of VA holds a set of open problems. One is the proper quantification, communication, and visualization of uncertainty in the VA cycle [53].

Sacha et al. [82] formulated requirements that need to be fulfilled to obtain an uncertainty-aware visualization. Their suggestions include: uncertainty quantification, uncertainty propagation, visualization of uncertainty in each component, and suitable interaction with uncertainty-aware visualization. These requirements will be used to adapt the classic VA cycle in this work.

Correa et al. [15] showed how the requirements by Sacha et al. could be described mathematically. Although this gives first hints on the requirements needed to implement an uncertainty-aware VA cycle, it does not clearly state where this information comes from and how it can be applied. In contrast to the presented work, we aim to provide an adapted and extended VA cycle that incorporates the suggestion by Sacha et al.

Karami et al. [50] provided an uncertainty-aware VA cycle that allows the processing of big datasets. Their work includes precise descriptions of each component in the VA cycle when considering big data. This limitation neglects further flavors of data that will be targeted in the presented work by a description of uncertainty-aware VA.

Senaratne [85] presented an uncertainty-aware VA cycle for Spatio-temporal data. Here, each component and connection of the cycle is explained in terms of uncertainty incorporation. We use this work as a starting point, while not being restricted to Spatio-temporal data.

Although the problem of uncertainty in VA is quite known, a generalized uncertainty-aware description of the VA cycle does not exist. As parts of the VA cycle highly depend on the underlying dataset, we aim to include uncertainty-aware descriptions of different datatypes.

2.2 Uncertainty-awareness in Related Disciplines

Uncertainty-awareness is highly related to a set of other disciplines including sensitivity analysis or VA of ensembles. We aim to shed light on these approaches and define starting points for our research.

VA in the context of ensemble datasets is highly related to the presented topic, as ensemble data can be transformed into uncertainty data (including loss of knowledge) and vice versa. Wang et al. [98] provided a state-of-the-art analysis for VA of ensembles and showed that a suitable communication of variability in an ensemble can be achieved by VA approaches. We would like to derive important knowledge from this work to achieve an uncertainty-aware VA cycle.

Liu et al. [65] showed that the quality of data is an important aspect that needs to be monitored in the VA cycle. In their work, they provided a mechanism that extends the VA cycle to enhance data quality and create awareness of data flaws. The quality of data is highly affected by data uncertainty. Resulting from this, we will include the data quality defined by Liu et al. in this proposed approach.

Sensitivity analysis [83] is highly related to uncertainty analysis as this discipline examines the effect of changing input variables to the output variable(s). Especially in machine learning, VA approaches are derived to conduct sensitivity analysis [89]. This also highly relates to uncertainty analysis as uncertainty expresses the variability of parameters in a system. We will include sensitivity analysis in the presented work if it is applicable. Although we found several related disciplines to uncertainty analysis and visualization, these sciences cannot build an uncertainty-aware VA cycle right away. This is based on two reasons. First, the related disciplines are themselves not solved and second uncertainty cannot be transformed into another problem without loss. Therefore, this work aims to provide an uncertainty-aware VA cycle.

3 Definitions

This section defines the mathematical basics of uncertainty. Here, we describe how to define,

quantify, propagate, and accumulate uncertainty as a reference for the remaining manuscript.

3.1 Definition of Uncertainty

Independent from the data source, task, and user, datasets are usually acquired by measuring or simulating a phenomenon creating data points. These measurements can be distorted by a variety of effects leading to measurement errors and uncertainty. Error and uncertainty are referring to two different aspects when considering measurements.

Let $a \in (-\infty, \infty)$ be a measurand and a^* be the true value of this measurand. When performing the measurement, the result will be a' . a^* and a' may be the same value, but in reality, their values differ due to a variety of effects. The error e of the performed measurement can be defined as the difference between the measured value and the true value of the measurand [9]. This means: $e = \|a^* - a'\|$. As a consequence, the quantification of an error requires a ground truth that clearly shows the difference between the actual value and the measured value.

The *uncertainty* of a measurement is a quantification of the doubt about the measurement result [39]. If this uncertainty is known, the measurand is defined to be *uncertainty-aware*. In contrast, if this uncertainty is unknown, a measurand is called *uncertain*.

Unfortunately, there is no unique definition of how to compute uncertainty. Arbitrary functions can be considered to achieve uncertainty quantification. In many cases, uncertainty is described as a *boundary* around the measurand [72]. It defines an interval around the measurand that can be defined as: $u_B(a) = [a' - u, a' + u]$. This description of uncertainty is chosen when the distribution of the occurrences is not important. Instead, it is important to know the limits in this variation [7].

Another popular definition of uncertainty utilizes *probabilistic* distribution functions [66] $u_{\text{PDB}}(a)$. These functions allow describing the probability density of a measurand to be located at an arbitrary point in some space. Here, the measurand usually defines the most probable location of the true value that was captured. The most prominent choice of probabilistic distribution functions are Gaussian distribution functions, but in general, any distribution can be used

to express uncertainty, including generalized linear models, Poisson distribution, and count-based models [40].

3.2 Quantification of Uncertainty

To achieve an uncertainty definition, proper uncertainty quantification is required. Uncertainty quantification expresses or measures the doubt about the data measurement. Its approaches are manifold. The most important categories can be roughly separated into four categories: *forward uncertainty quantification*, *sensitivity analysis methods*, *response surface methods*, and *dimension reduction methods* [63]. We will explain each category briefly in the following.

Most of the uncertainty forward propagation techniques aim to assign a statistical distribution for each of the model parameters considered to be uncertain. Most of these techniques are data-based methods that include standard statistical techniques such as maximum likelihood estimation, minimum distance estimation, method of moment estimation, and Bayesian inference. A summary of these techniques can be found in [60].

Sensitivity analysis methods can be used for uncertainty quantification. Here, the idea is to provide a measure of the variability of input parameters in a system. As a result, the effect of variability of input parameters to the output of a system can be described [4].

Response surface method approximation techniques aim to approximate a mathematical model by providing a simplified meta-model mostly using linear or quadratic functions [33]. These methods are used to reduce the computational effort in large and complex systems.

As parameter analysis can be computationally expensive, dimension reduction methods can be utilized for uncertainty quantification [13]. These techniques aim to reduce the set of input parameters to facilitate uncertainty quantification. A summary of dimension reduction approaches can be found in [25].

3.3 Propagation of Uncertainty

The propagation of uncertainty is an important issue when data (including their uncertainty) is transformed.

Data is mostly propagated through mathematical operations O . These operations do not solely

Fig. 2: Pathways of uncertainty propagation. When an operation O is applied to an attribute a , an adapted function $\bar{O}$ needs to be applied to the uncertainty of the attribute. The uncertainty of the attribute influences the function O . In addition, uncertain values need to be adapted by a damping factor d .

affect the data, but also the attached uncertainty. Besides, mathematical operations are affected by the uncertainty of their operands. This results in the need to adjust mathematical operations to be able to handle uncertainty, as shown in Figure 2.

In order to extend mathematical operations, an operation O is modified to $\bar{O}$, where $\bar{O} : \bar{a} \rightarrow \bar{a}^*$. This includes three computational paths: first, the manipulation of the attribute itself. Second, the manipulation of the uncertainty quantification of the attribute, and third a damping factor that manipulates the influence of an attribute according to its uncertainty.

An uncertainty-aware formulation of O can be achieved by:

$$O^*(\bar{a}) = (O(a) * d(u(a)), \bar{O}(u(a))) \quad (1)$$

where $d(a)$ is the damping factor of each attribute. $d(a)$ can be defined as:

$$d(u(a)) = \begin{cases} \frac{1}{u(a)} & u(a) > 0 \\ 1 & \text{else} \end{cases} \quad (2)$$

This means that every time an attribute is utilized in a mathematical operation, the attribute value will be damped when the respective uncertainty is high. When uncertainty is zero, the attribute value will be fully considered. Furthermore, all mathematical operations that are applied to an attribute will be applied to the uncertainty quantification of this attribute. Here, the function $\bar{O}$ is dependent on the mathematical function O and can be derived considering

the uncertainty propagation rules summarized by Gillmann et al. [31].

3.4 Accumulation of Uncertainty

As shown above, uncertainty can be introduced into the VA cycle at all components, or multiple sources of uncertainty can affect one component. This results in the need for a mechanism that allows the accumulation of uncertainty.

The accumulation of uncertainty can, in principle, be achieved by arbitrary accumulation functions. Cai et al. [12] presented a survey of aggregation functions. In the VA process, a proper aggregation function needs to be able to aggregate all sources of uncertainty in the VA cycle in an orderly manner, also allowing the user to adjust the importance of all sources of uncertainty in the VA cycle. This is required, as users may need to determine which sources of uncertainty are more important than others or even discard specific sources.

3.5 The Role of Uncertainty in Visual Analytics

Keim et al. [52] proposed that the inclusion of uncertainty into the VA cycle is a non-trivial task. This is due to a variety of sources of uncertainty in this cycle. This section aims to summarize these sources to create a basis for the required adaptations in the VA cycle to make it uncertainty-aware.

In fact, each main component of the VA cycle can introduce uncertainty. The sources of uncertainty can have different origins: uncertainty based on the underlying model (epistemic uncertainty) or statistical uncertainty resulting from variations in the measurement result when running an experiment multiple times (aleatoric uncertainty). In most cases, aleatoric uncertainty is usually the type of uncertainty that is requested to be visualized to enhance a decision-making process in a given application [76].

Starting from the input dataset, uncertainty can be introduced into the VA cycle by data incompleteness, finite instrument resolution, non-representative sampling, variations in observations, and incomplete knowledge about the measurand [8].

When considering Hypothesis, uncertainty can be introduced by parameter uncertainty. This

means that computational models often require parameters, which can be hard to find in many cases or it is hard to determine if a chosen parameter is optimal [70]. Furthermore, the computational model itself introduces uncertainty into the VA cycle. Models are incomplete or approximate physical behavior by definition. As our knowledge of the world and computational power is limited, hypothesis forming is affected by uncertainty [26].

In terms of visualization, uncertainty can be introduced by the mapping of visual variables of the visualization algorithm, as well as the resolution of the display device [71]. Also, users reviewing the shown visualization can introduce uncertainty into the VA process that stems from perceptual uncertainty, memory uncertainty, and thinking uncertainty [23].

At last, uncertainty can be introduced into the VA cycle while creating a hypothesis. Here, users can introduce uncertainty through a decision-making bias. This means that users may tend to ignore VA results, as they might be biased by previous results. Besides, the experience and knowledge of domain experts can also introduce uncertainty into the VA cycle [95].

Please note that not all mentioned sources of uncertainty are present in each scenario where VA is applied. Also, there exist cases where a specific source of uncertainty may be present, but is neglected as its influence is too small. This decision is highly dependent on the use case, data source, and user. Still, one or even multiple sources of uncertainty are likely introduced into a specific implementation of the VA cycle. As the VA cycle is a chain of operations that is connected while visiting different components in the cycle, uncertainty needs to be *propagated* and *accumulated*.

4 Uncertainty-aware Visual Analytics

In this work, we aim to provide a description of uncertainty-aware VA that allows visualization researchers to get a quick overview of the necessary steps that need to be accomplished when being confronted with a dataset affected by uncertainty. This includes two important adaptations to the traditional VA cycle. First, all existing components and connections in the VA cycle need to be *extended* or *adapted* to incorporate uncertainty

information. Second, the existing traditional VA cycle does not hold mechanisms to insert uncertainty knowledge into the VA cycle and keep track of them, which means that there are missing components and connections in the classic VA cycle that need to be *added*.

At first, we will follow the definition of the VA cycle by Keim et al. [53], as shown in Figure 1. In this work, the VA cycle is composed of four components:

- Dataset (■)
- Hypothesis (■)
- Visualization (■)
- Insight (■)

The components are connected by operations required in the VA process. These operations are encoded as connections between components and are defined as functions that transform one component into another. We sort these operations into the four main components according to where they fit best.

Please note that all connections originating from the classic VA process will be marked by a box (■) in the respective color of the category they belong to.

To describe a complete uncertainty-aware VA cycle, we need to introduce two novel components and several connections to already existing components. Namely, the novel components are:

- U-Dataset (▲)
- Provenance (▲)

The novel connections include uncertainty quantification and provenance generation concerning the existing components and connections of the VA cycle. Please note, that all novel components will be marked by a triangle (▲) in the respective color throughout the entire manuscript for smooth reading. The presented description will be structured along with the six *components* we defined.

4.1 Dataset *S*

A dataset *S* is a very general concept that consists of *n* records ($r_1, r_2, \dots, r_n$), where each record r_i , consists of *m* observations, variables or attributes ($a_1, a_2, \dots, a_n$). An attribute a_i is a single entity such as a number or symbol. Dataset holds a structure that can be syntactic or semantic [99]. They can be

Fig. 3: Uncertainty-aware visual analytics. Starting from an input dataset, a U-Dataset can be created. This U-Dataset can be utilized to create an uncertainty-aware hypothesis or an uncertainty-aware visualization. Based on these analysis techniques, an uncertainty-aware insight can be generated through user interaction and provenance creation. The uncertainty-aware insight can be fed back into the Dataset or U-Dataset component.

generally defined as a function t . These relations are normally used to differentiate various types of data, e.g. attributes that are aligned on a grid are usually referred to as image data.

Based on the respective problem description, a dataset S is generated to be analyzed in the VA cycle. In contrast to the classic definition of the VA cycle, an uncertainty-aware VA cycle requires mechanisms that allow extending the dataset into an uncertainty-aware U-Dataset. The required steps in this process will be explained in the following.

4.1.1 Preprocessing D_W ■

The classic VA cycle allows processing the input dataset by four different operations: Data Transformation D_T , Data Cleaning D_C , Data Selection D_S , and Data Integration D_I . Up to the point, where no uncertainty definition or quantification has been performed, these operations can be applied as defined in the classic VA cycle. Although data preprocessing is an important step and in many cases indispensable step in the VA cycle, it is not recommended before an uncertainty definition and quantification has been achieved [11]. This results from the problem that usually all input data, independent from their quality is required to define a suitable uncertainty quantification of a dataset. Applying data preprocessing algorithms before an uncertainty

quantification can fade important characteristics in the input dataset which are required for proper uncertainty quantification. Still, the uncertainty-aware VA cycle contains these connections for completeness.

4.1.2 Uncertainty Quantification Q_S ▲

Depending on the data format, application, and task that the user needs to fulfill, proper uncertainty quantification is required. This holds for each record (and its attributes) in a dataset, as well as for the relations defined in the dataset. There exist a variety of datasets that are acquired in conjunction with an uncertainty quantification such as molecular data. In this case, uncertainty quantification of the input dataset can be neglected if the provided uncertainty quantification expresses the uncertainty of the input dataset well enough.

As mentioned in Section 3.2, uncertainty quantification approaches are manifold. In the case of input datasets, the uncertainty can mostly be quantified by using forward propagation techniques, as this group of measures results in a statistical description of the input parameters. Still, the presented uncertainty-aware VA cycle can handle all types of uncertainty descriptions.

4.2 U-Dataset $\tilde{S}$

Resulting from the input dataset S in conjunction with the achieved uncertainty quantification D_Q we aim to achieve an uncertainty-aware dataset (U-Dataset) $\tilde{S}$.

As a first definition, we require the uncertainty of an attribute. Let a be an attribute, and A be the set of all possible values for a , then $\bar{a} = (a, u(a))$ is the uncertainty-aware description of the attribute a . Note that $\bar{A}$ holds all possible uncertainty-aware descriptions of the set of attributes a . Attributes can be single measurands, but in the following, they can also contain entire datasets (large and complex data). This means, that dataset combinations such as multi-field data or ensemble datasets are explicitly possible.

The uncertainty quantification of a dataset can also affect the function t , expressing the relation within the dataset. Resulting from this, uncertainty quantification can result in a novel function $\bar{t} = (s, u(t))$ that allows to express uncertainty within the relation function. One example is the connection between points within a graph. Here, the function that defines the relationship between data points can be adapted to capture the degree of certainty that the respective points are connected.

4.2.1 Uncertainty-aware Data Preprocessing D_W ▲

Once an uncertainty-aware dataset is achieved, preprocessing operations can be applied to transform the dataset into a format that allows the creation of hypotheses or apply visualization approaches. Here, data transformation $\bar{D}_T$, data cleaning $\bar{D}_C$, data selection $\bar{D}_{SL}$ and data integration $\bar{D}_I$ are available, as defined in the original VA cycle. Still, they need to be adapted to be uncertainty-aware.

The transformation of data is concerned with an application of mathematical functions to describe the transformation. As we consider U-Datasets in the uncertainty-aware VA cycle in the form $\tilde{S} = (S, u(s))$, we require mathematical operations that can be applied in this setting. Here, three different pathways have to be followed, as shown in Figure 2.

In the classic VA cycle as well as in most other data analysis scenarios, datasets are cleaned,

selected, and integrated into each other to provide a stable dataset that can be processed. When considering data cleaning, we propose two important adaptations in this process: *Do not eliminate* any captured data point and *Merging* data points (including their uncertainty).

When eliminating a data point, the information (no matter how uncertain it is) is neglected in the VA cycle. No matter how well selected these points are, the selection is based on a hypothesis or metric that could be wrong or incomplete. To avoid this, we propose to find a suitable uncertainty quantification that assigns a very high uncertainty to the selected data point. In this case, data points are not sorted out completely, but if a data point has very high uncertainty, the data point will not be considered as important as other values in the VA cycle.

The merging of data points arises when a phenomenon is captured in the data multiple times. Here, data points are merged to avoid multiple occurrences of the same phenomena in the dataset. In this case, one must not only merge the data points. In addition, the uncertainty of the data points needs to be merged as well, resulting in an accumulation of uncertainty. This accumulation can be computed based on the suggestion in Section 3.4.

4.3 Hypothesis $\tilde{H}$

A hypothesis is a supposition or proposed explanation made based on limited evidence as a starting point for further investigation. To achieve this, the null hypothesis is usually utilized. In this case, a hypothesis is formed and tested. Then the hypothesis can be either rejected or fail to be rejected. Greco [34] explained in his work that a hypothesis can never be accepted or confirmed as it does not capture a confidence value of the rejection criterion of a hypothesis.

In the classic VA cycle, the component hypothesis H is described as a general tool to create insight or knowledge based on statistical analysis. When considering hypotheses that are based on uncertainty-aware datasets, we need to define an uncertainty-aware Hypothesis $\tilde{H} = (H, u(H))$. Here, $u(H)$ describes a confidence value for the formulated Hypothesis. This means that whatever the output of a statistical analysis method

is, a result is composed of the derived Hypothesis H and an uncertainty quantification $u(H)$ of the generated hypothesis. The generation of an uncertainty-aware hypothesis and possible interaction methods will be shown in the following.

4.3.1 Uncertainty Quantification in Hypothesis $Q_{\bar{H}}$ ■

As shown in Section 3.5, uncertainty can be introduced by a Hypothesis itself, namely through parameter uncertainty, incompleteness, and approximation of models. For input parameter uncertainty we suggest utilizing sensitivity analysis uncertainty quantification approaches, or for a high number of input parameters dimension reduction uncertainty quantification approaches should be used.

The incompleteness and approximation approach in a model can be described using model reliability approaches. A summary including evaluation of these approaches can be found in [80].

The quantified uncertainties need to be combined with the uncertainty quantification that is attached to the input dataset or the input visualization using an uncertainty accumulation approach as described in Section 3.2.

4.3.2 Generation ■

The generation of an uncertainty-aware hypothesis $\bar{H}$ can be described by a function starting from two sources: an uncertainty-aware dataset ($\bar{H}_S : \bar{S} \rightarrow \bar{H}$) and an uncertainty-aware visualization ($\bar{H}_V : \bar{V} \rightarrow \bar{H}$).

Generation from U-Datasets

In the classic VA cycle, the generation of hypothesis H can be based on a dataset utilizing a set of statistical analysis tools $\{f_{S1}, f_{S2}, \dots, f_{Sq}\}$. These statistical operations need to be redefined to provide an uncertainty-aware creation of a hypothesis. Fortunately, physicians and engineers are concerned with this issue for decades and massive literature is available that summarizes the hypothesis generation based on statistical analysis. Devore [19] summarized uncertainty-aware descriptions of all standard statistical tests for uncertainty-aware datasets. It includes average, variance, standard deviation, the sum of

squares, root sum of squares, pooled variance, linear interpolation, linear regression, sensitivity coefficient, covariance, and correlation. For statistical approaches that have not been described yet, we suggest the uncertainty propagation rules described in Section 3.3. As statistical analysis is based on mathematical functions, the application of uncertainty propagation rules works right away.

During the last decades, machine learning approaches become increasingly important in the generation of hypotheses and are a standard tool by now. In this context, clustering approaches are a popular form of machine learning. A survey on uncertainty-aware clustering approaches was presented by Aggarwal et al. [1]. These algorithms are capable of transforming uncertainty throughout their computational model and provide an uncertainty-aware hypothesis forming. Neural networks are increasingly popular in providing hypotheses, as well. Here Gal provided a state of the art analysis of uncertainty-aware approaches [27]. Most popular in this context are deep learning approaches that utilize Bayesian theory [97] to output an uncertainty-aware hypothesis.

Generation from Visualization ■

Building uncertainty-aware hypothesis from uncertainty-aware visualizations is defined as the function $\bar{H}_V : \bar{V} \rightarrow \bar{H}$. Unfortunately, this process cannot be determined analytically in its entirety, as it involves the subjective impression of a user to refine a hypothesis when regarding the available visualization. What can be determined is the user input that leads to a hypothesis. Here, we suggest letting the user quantify how certain his selections are to express the uncertainty of the hypothesis generation at least partially.

4.3.3 User Interaction with Hypothesis ■

User interaction with a hypothesis usually concerns operations such as selecting a proper hypothesis generation algorithm or adapting previous choices. In this context, the interaction is only allowed to select uncertainty-aware hypothesis forming operations. In addition, a user may be enabled to adapt input parameters required for computing uncertainty-aware hypothesis forming. Here, users need to be able to not solely set the

input parameter J . In addition, the input parameter needs to be expressed with an uncertainty quantification $u(J)$ as well. Thus, the user should be enabled to manipulate this uncertainty quantification. The resulting uncertainty-aware input parameter $\bar{J} = (J, u(J))$ needs to be considered in the uncertainty-aware computation based on the propagation rules defined in Section 3.3. Here sensitivity analysis can be utilized to quantify this uncertainty.

4.4 Visualization $\bar{V}$

The visualization of uncertainty-aware datasets is a key component in the VA cycle. It allows users to gain valuable insight into the dataset and provide a natural understanding of the underlying uncertainty [46]. In the uncertainty-aware VA cycle, an uncertainty-aware visualization is defined as $\bar{V} = (V, u(V))$. This means that a suitable visualization of uncertainty generates a visualization that finds suitable visual metaphors for the captured uncertainty in the underlying dataset. The creation of an uncertainty-aware visualization approach can be achieved considering U-datasets or uncertainty-aware hypotheses.

4.4.1 Uncertainty Quantification $Q_{\bar{V}}$ ■

Section 3.5 shows that the visualization process itself introduces uncertainty into the VA process, namely mapping uncertainty, perceptual uncertainty, and memory and thinking uncertainty. Dasgupta and Kosara [18] summarized the need for quality metrics in visualization that can quantify uncertainty such as mapping uncertainty. Diamond [20] provided a survey on perceptual uncertainty and how it can be expressed. Coutinho [16] described the role of memory and thinking uncertainty when reviewing a visualization. They propose that the description of these uncertainties is hard to achieve as human cognition is very complex and parts of its functionality are unknown.

Uncertainty quantification of the visualization process needs to be accumulated with already existing uncertainty quantification approaches that arise from visualization generation by hypothesis or uncertainty-aware datasets as shown in Section 3.2.

4.4.2 Generation ■

The generation of an uncertainty-aware hypothesis $\bar{V}$ can be described by a function starting from two sources: an uncertainty-aware dataset ($\bar{V}_S : \bar{S} \rightarrow \bar{V}$) and an uncertainty-aware visualization ($\bar{V}_H : \bar{H} \rightarrow \bar{V}$).

Generation from U-Datasets

Uncertainty-aware visualization is a very active field that has been researched for decades resulting in a variety of visualization approaches. Still, it only represents one component of the VA process. Therefore, visualization can be seen as one computational step in the pipeline.

In general, the utilized visual variables that are considered to express uncertainty in visualization can be listed as follows: comparison techniques, attribute modification, glyphs, and image discontinuity [76]. The choice of uncertainty visualization and the visual variable expressing the uncertainty is highly dependent on the underlying dataset and the use case, the VA cycle is designed for. As a hint, the selected visualization approach needs to be able to visually encode all types of uncertainty that are encoded in the uncertainty-aware dataset and the accumulation of uncertainty that has been performed so far [82].

Generation from Hypothesis

The process of generating an uncertainty-aware visualization based on an uncertainty-aware hypothesis can be described as $\bar{V}_H : \bar{H} \rightarrow \bar{V}$. Here, we assume that an uncertainty-aware statistical analysis has been conducted requiring a proper visualization. Depending on the output of the statistical analysis, a U-dataset can be created. The specific data type depends on the underlying statistical analysis approach and requires a sophisticated visualization approach. Here, the same rules apply as in Section 4.4.2.

4.4.3 User Interaction with Visualization ■

User interaction with visualizations can be quite manifold. A summary of available interaction techniques was given by Brodbeck et al. [10]. In terms of interaction with uncertainty visualization, Sacha et al. [82] proposed a suitable user interaction with uncertainty-aware visualization

approaches as a fundamental requirement to provide a suitable uncertainty-aware VA cycle. Still, a summary of all necessary interaction metaphors is not available. In this context, we would like to suggest the following considerations when designing uncertainty-aware interactions for visualization.

First, there needs to be specific selection or zooming operations that are based on the data uncertainty, not on the data itself. Second, the result of the current interaction methodology needs to provide information about the currently shown uncertainty about the overall uncertainty captured in the dataset.

4.5 Insight $\bar{I}$

The term insight I can be defined as knowledge that is gained during analysis and has to be internalized, synthesized, and related to prior knowledge [82]. In terms of uncertainty, an uncertainty-aware insight $\bar{I} = (I, u(I))$ is composed by the insight generated from the uncertainty-aware VA cycle and quantification of the credibility of the derived result $u(I)$. In reality, insight cannot be defined mathematically in many cases, as it is a subjective impression of the user, often affected by personal bias that runs the VA cycle. Based on this problem, it might not be possible to describe the respective uncertainty quantification.

4.5.1 Uncertainty Quantification $Q_{\bar{I}}$ ▲

Insight generated in the VA cycle can be affected by uncertainty due to decision-making bias or experience and knowledge that may keep a user from accepting novel findings. Lewandowsky et al. [61] stated that knowledge is always affected by uncertainty. Unfortunately, Insights are very subjective, such that uncertainty quantification is hard to achieve. Most considerations are philosophical rather than computational [3]. For evaluation purposes, benchmark tasks have shown to be useful for identifying and assessing analytic findings but are not sufficient in most cases [75]. Here, a clear strategy of uncertainty quantification is missing.

4.5.2 User Interaction to create Insights from Hypothesis U_{CH}^- ■

Uncertainty-aware insight generated from an uncertainty-aware Hypothesis can luckily be quantified mathematically (at least up to the point where analysis results are interpreted). Here, the uncertainty-aware hypothesis directly implies the uncertainty of the derived uncertainty-aware insight. In fact, they are identical, which means $u(I) = u(H)$.

4.5.3 User Interaction to create Insights from Visualization U_{CV}^- ■

Throughout the interaction of the user with uncertainty-aware visualization, insight is generated. In comparison to the insight based on hypothesis, the considered insight can usually not be described mathematically as it is depending on a subjective user experience. Here, visualization evaluation approaches come into play, as they offer metrics and approaches to quantify the amount of insight generated by a visualization.

In terms of uncertainty visualization, Hullmann et al. [45] presented a state-of-the-art report that summarizes uncertainty visualization evaluation approaches. These approaches can be used to at least approximate the insight generated by an uncertainty-aware visualization approach.

4.5.4 Feedback Loop $F(S)$ ■ in Difference to uncertainty-aware Feedback Loop $\bar{F}(S)$ ▲

As indicated by the classic VA approach, VA is designed to be a cycle $F(S)$. When generating new knowledge, this knowledge can act as further data input. As already shown, generated insight from the uncertainty-aware VA cycle can be of two types: insight, with uncertainty quantification, and insight without uncertainty quantification. These types of insight need to be treated differently. Insights without uncertainty quantification that need to be reinserted in the VA cycle need to be fed back into the dataset component. This is the reason why an uncertainty-aware VA cycle still requires the component S . Starting from here, a suitable uncertainty quantification needs to be found according to the data structure of

the insight. As the uncertainty of insight cannot be computed directly in many cases, insight can be modeled as a normal dataset and then be transferred into a U-Dataset through a suitable uncertainty quantification as described in Section 3.2.

On the other hand, insights that have an uncertainty quantification need to be inserted in the U-Dataset component, as there is no uncertainty definition or quantification required.

4.6 Provenance Generation ▲

When running an uncertainty-aware VA cycle, uncertainty will be propagated and accumulated along with the performed operations of the VA cycle. The importance of provenance analysis and visualization has been described by Varga et al. [92]. This implies the tracking of uncertainty throughout each computational step of the VA cycle, referred to as provenance. Therefore, each time an uncertainty-aware dataset, a hypothesis, or a visualization is created, the current uncertainty quantification and the respective operation need to be stored and are subject to further analysis.

In fact, we encourage to provide a visualization and interaction tool to let users follow the development of uncertainty throughout the VA process. This can give users important hints on which operations caused a drastic increase of uncertainty or at which point the accumulated uncertainty exceeds a threshold that is known to be the highest amount of uncertainty that still allows for interpretation. Herschel et al. [42] provided a survey on provenance creation. Still, the selection of a suitable tool is highly dependent on the underlying dataset and the application scenario. Ragan et al. [78] characterized provenance in visualization and data analysis approaches. We will use this work as a basis for the following descriptions of uncertainty provenance that are required in the VA cycle.

4.6.1 Provenance Generation for U-Datasets $P_{\bar{S}}$ ▲ and uncertainty-aware Hypothesis $P_{\bar{H}}$ ▲

The provenance of data focuses on the history of changes and movement of data. Data provenance

is often heavily emphasized in computational simulations and scientific visualization, in which significant data processing is conducted. The history of data changes can include subsetting, data merging, formatting, transformations, or execution of a simulation to ingest or generate new data [78]. This can be directly transferred to the uncertainty of a U-Dataset and the uncertainty of a hypothesis.

4.6.2 Provenance Generation for uncertainty-aware Visualization $P_{\bar{V}}$ ▲

As Ragan et al. [78] stated visualization provenance is concerned with the history of graphical views and visualization states. This process is tightly coupled with data transformation and the interactions used to produce the visualization. These concepts need to be adapted to provide a provenance generation for the uncertainty in uncertainty-aware visualization. A survey on available methods in provenance visualization and user interaction was conducted by Xu et al. [102].

4.6.3 Provenance Generation for uncertainty-aware Insight $P_{\bar{I}}$ ▲

The provenance of uncertainty-aware insights needs to include the component of uncertainty as well. Unlike data computations, insights are not directly observable in all cases and so their uncertainty is not observable as well directly, as shown in Section 4.5. Here, solely quantifiable insights can be included in the provenance generation of uncertainty.

5 Opportunities of uncertainty-aware Visual Analytics

Based on the proposed uncertainty-aware Visual Analytics(VA) cycle, its implementation to various dataset types and applications fields is provided here.

5.1 Data Types

To apply the provided definition of uncertainty to different data types, the characteristics of each data type have to be considered. Table 1 shows the

Original DataType	Description	Uncertainties	U-DataType
Geospatial Data	Locations or trajectories that have attributes assigned to	Spatial Uncertainty $\bar{L}$ Attribute Uncertainty $\bar{A}$	$\bar{S}_1 = (\bar{L}, \bar{A}, f : \bar{L} \rightarrow \bar{A})$
Graph/Network Data	Nodes that are connected by links that can have attributes assigned to	Node Uncertainty $\bar{V}$ Link Uncertainty $\bar{E}$ Attribute Uncertainty $\bar{A}$	$\bar{S}_2 = (\bar{V}, \bar{E}, \bar{A}, f_V = \bar{V} \rightarrow \bar{A}, f : \bar{E} \rightarrow \bar{A})$
Field Data	Field with specific size and spacing that have attributes assigned to each field point	Grid Uncertainty $\bar{P}$ Attribute Uncertainty $\bar{A}$	$S_3 = (\bar{P}, f : \bar{P} \rightarrow \bar{A})$
High-dimensional Data	Tuples of attributes	Attribute Uncertainty A	$\bar{S}_4 = (A^N)$
Temporal Data	Attributes that are sorted along a timeline	Time Uncertainty T Attribute Uncertainty A	$S_5 = (\bar{T}, \bar{A}, f : \bar{T} \rightarrow \bar{A})$
Text/ Document Data	Positions holding different attributes	Attribute Uncertainty A	$\bar{S}_6 = (\bar{P}, A, f : \bar{P} \rightarrow A)$

Table 1: Uncertainty-aware data types and their definition. For each datatype, an informal description of contained uncertainties, and an uncertainty-aware definition of the datatype is provided.

most prominent data types occurring in the context of visual analytics. It holds a short description of the dataset characteristics as well as a list of different types of uncertainties occurring in specific data types.

Geospatial data S_1 uses geospatial locations or trajectories L . Here, various attributes A are assigned to such a domain L by a function $f : L \rightarrow A$. Therefore, two types of uncertainty, namely spatial uncertainty and attribute uncertainty [62], are found in such datasets. Spatial uncertainty originates from the underlying areas or trajectories that can be displaced or shifted in shape, deviating from the stored data. Attribute uncertainty, on the other hand, describes the uncertainty of data attributes themselves. Figure 4a illustrates both types of uncertainty by showing positional and attribute uncertainty. Li et al. [62] described how analytic models can be utilized to achieve uncertainty quantification.

Visualizations of uncertainty-aware spatial data are manifold and include earth, space, and environmental sciences [100], urban science [87, 88], terrain visualization [91] and geographic/geospatial visualization [68]. An example is shown in Figure 5a, providing the uncertainty in predicting wildfires, color-coding a map of terrain at risk.

Graph data S_2 connects a set of nodes V via links E creating a network called graph. These nodes and links can hold various attributes, provided by functions $f : V \rightarrow A$ and $g : E \rightarrow A$. Graph data can hold three different types of uncertainty [51]. First, the presence of a node can be uncertain. Second, a link between nodes can be uncertain, and third, the attributes contained in nodes or links can be uncertain. It should be noted that the position of nodes in a visualization is not a fundamental uncertainty, as it is derived from the graph description or some graph drawing algorithm. Engel et al. [22] provided an uncertainty quantification for graph data. A visual indication of these types of uncertainty can be found in Figure 4b.

Uncertainty-aware graph-based data, occurring in applications like business and finance [36, 94], social and information sciences [6], sensor networks [21, 86], bioinformatics [93] and cybersecurity [2], can be visualized by a variety of approaches. These approaches are usually based on uncertainty-aware graph-drawing algorithms.

Fig. 4: The average and standard deviation of critical parameters: Region R4

Figure 5b shows an example where edge and node attributes that contain uncertainty are visually encoded by areas of varying sizes.

Field data S_3 can contain scalars, vectors, and tensors (attributes A), often arranged on some grid. This grid is defined by a set of positions and neighborhood relations on those given positions. The result is cells or positions with neighborhood information about their adjoined cells or positions, while each cell holds its attribute L . They are connected by a function $f : P \rightarrow A$ where P are the set of positions or set of cells. Here, two types of uncertainty can occur, as depicted in Figure 4c. Both positions, as well as the attributes defined over P can be uncertain [37]. It is important to note that each attribute value may be affected by uncertainty to differing extents. This means, for example, that vector entries can have varying

uncertainty depending on their dimension. Potter et al. [76] provided a summary on uncertainty quantification for field data.

Uncertainty-aware Field data visualizations can be found in mathematics, physical sciences and engineering [64], multimedia(image/video/music) [101], biomedical and medical [38, 56, 81] applications. Here, the visualization highly depends on the attributes that are encoded in the respective field. Figure 5c provides an example of an uncertainty-aware visualization using diffusion tensors. The surrounding transparent surfaces indicate the varying visual appearance of the visualized tensor.

High-dimensional Data S_4 is defined by a dimension N that determines the number of attributes A contained in one entry. N is a larger number, usually higher than 10, even though some authors talk about high-dimensional data if $N >$

Fig. 5: The average and standard deviation of critical parameters: Region R4

3. Here, only attribute uncertainty needs to be considered, as shown in Figure 4d.

Uncertainty-aware High-dimensional data can be found in a variety of applications. Hoffmann et al. [43] provided a survey of potential visualization approaches. Figure 5d shows an uncertainty-aware parallel coordinates visualization. Instead of visualizing lines between axes, the images visually indicate areas with varying occurrences of connecting lines utilizing a color table.

Temporal Data S_5 contains attributes A that are sorted along a time-line T utilizing a function $f : T \rightarrow A$. These attributes can be manifold and may be of any type of data that was mentioned before. Here, two types of uncertainty arise: time uncertainty and attribute uncertainty [14], as shown in Figure 4e. Each point in time can be affected by uncertainty as well as the attribute attached to this point in time. Zhen et al. [44] demonstrated the quantification of uncertainty in temporal data.

Uncertainty-aware Time-dependent data often occurs in digital humanities [90], as well as

robotics [74]. Figure 5e shows a timeline visualization utilizing different glyphs to indicate the uncertainty of specific time steps.

Text/Document Data S_6 is data in the form of text or documents that hold attributes A at a specific character position P . This connection is given by the function $f : P \rightarrow A$. Here, two types of uncertainty can arise, as shown in Figure 4f: Document uncertainty and attribute uncertainty [54]. Each document can have an overall uncertainty and all of its entries can be affected by uncertainty. Quantification of uncertainty in textual data was given by Kerdjoudj et al. [54].

Uncertainty-aware Text/Document Data can occur in nearly all kinds of applications. Prominent examples are digital humanities [90] and software visualization [5]. The visualization strongly depends on the underlying text that is visualized. Figure 5f shows the visualization of a tag cloud that is adapted according to the uncertainty of the underlying words. Uncertain words are shown with a lower opacity compared to certain words.

For each of these data types a proper description, including the quantification of uncertainty, needs to be found. As mentioned before, the quantification highly depends on the underlying dataset and use case and cannot be generalized. Although examples were given on how to find an uncertainty quantification algorithm, they may not be considered to be complete or appropriate. This manuscript does not aim to provide a state-of-the-art analysis of uncertainty quantification strategies. Instead, the necessary steps that need to be accomplished to run an uncertainty-aware visual analytics cycle are provided. The given references are mentioned to deliver a starting point for visualization researchers that wish to learn more about uncertainty quantification.

5.2 Applications

In the following, we would like to summarize potential applications that can benefit from the proposed uncertainty-aware visual analytics cycle or that are already using uncertainty-aware visual analytics cycles.

Medicine

Medical data, independent from their origin, is always affected by uncertainty. Especially when considering medical imaging, which is an important data source for medical applications, a variety of sources of uncertainty are introduced, as shown by Gillmann et al. [28]. In general, the need for uncertainty-aware visual analytics has been highlighted as one of the ten recent challenges in medical visualization [30].

Many medical applications are affected by uncertainty mainly originating from missing data, ambiguous data, or data that is affected by error sources in the data acquisition process. Examples are lesion examination [28], surgery assistance [29] and vascular analysis [73]. This leads to a decision-making process that is highly affected by uncertainty. Especially in medicine, where the patients' health can be highly affected by a decision made, uncertainty-aware visual analytics can be of great benefit.

Biochemistry

Biochemistry is a largely developing field where simulations are used to predict the behavior of molecules and their interaction. Regarding drug

development and pandemic research, this research field gained massive interest. In this discipline, researchers aim to predict the resulting molecular structure. These structures are based on atomic positions in space, which are by default not certain.

Maack et al. [67] provided an uncertainty-aware visual analytics approach that allows understanding the positional uncertainty within atoms and molecules to understand the potential effectiveness of a molecule. In their work, they provided visual hints that allow understanding how molecules may change and how their stability is affected by this. This example shows the effectiveness of uncertainty-aware visual analytics in this discipline and how its use can ease drug development and examination of molecules.

Environmental Sciences

Environmental sciences usually deal with massive amounts of data resulting from large simulations. Alternatively, they use remote sensing images to capture current states of weather conditions. All these types of data contain uncertainty and resulting from this, the computations made are also affected by uncertainty.

Many applications in environmental sciences deal with uncertainty such as weather forecasts, climate change simulations, or sea examinations, which need to be analyzed and understood by visual analytics approaches.

Raith et al. [79] proposed a visual analytics approach to understand different sources of uncertainty inherent in Eddy detection of large oceans. Eddies are water currents that play a dramatic role in the exchange of minerals in an ocean and are therefore of great interest. Here, visual analytics provides a great opportunity to understand where eddies may occur and what size they have.

Urban Planning

Urban planning is concerned with the development of cities and rural areas. Here, many scenarios need to be considered to plan sustainable cities that are capable of providing all necessary structures to their cities [24]. This results in a variety of uncertainties that need to be communicated.

An example of an urban planning task is the control and planning of traffic in cities. Here, Lee et al. [59] provided a visual analytic approach that

assists in predicting road traffic congestion. Here, uncertainty can be included to show how traffic may vary leading to differing scenarios.

Mechanical Engineering

In mechanical engineering, the proper construction of workpieces and the detection of bottlenecks in a production line is a typical problem that holds a high amount of uncertainty. To achieve this, many simulations are run to examine problems in the production process. Kretzschmar et al. [58] provided an uncertainty-aware visual analytics system that allows the examination of different simulation scenarios and sees which construction is the most stable.

Digital Humanities

Digital humanities are concerned with the systematic examination of historical or sociological data. Here, a variety of sources of uncertainty can be contained in the data ranging from ambiguity over missing values to uncertain sources of data. The computational analysis of data originating from the digital humanities community has become very popular in recent years, where uncertainty inherent in the data has been named as one major challenge [55].

Jänicke et al. [49] provided an uncertainty-aware visual analytics approach to examine missing or ambiguous data of musicians. Here, users are enabled to understand the trustworthiness of the data they review.

6 Open Challenges

Although we successfully described an uncertainty-aware visual analytics cycle and show how it can be applied to a variety of cases, there remain open problems that need further investigation. The problems separate into two groups: open problems that result from the visual analytics cycle (see Section 6.1) itself and open problems that result from the inclusion of uncertainty (see Section 6.2).

6.1 Open Problems that result from the visual analytics cycle

Generalization

In this paper, we showed that the visual analytics cycle can be extended to include uncertainty.

Although this is a suitable extension for many real-world problems, there exist further cases that cannot be treated with the classic visual analytics cycle. These cases include ensemble datasets or multi-modal datasets. Here, proper extensions of the visual analytics cycles are required.

Proper description of the Insight

As shown in this manuscript, the insight that can be generated using a visual analytics cycle, regardless of whether it incorporated uncertainty or not, cannot be quantified properly. This is because the insight is mainly depending on the user of the provided cycle. Here, proper quantification approaches of the insight are required to drive the development of visual analytics cycles in general.

Approximation of the Amount of Knowledge that is generated by a Visualization.

As shown in Section 4, the amount of insight that can be generated based on visualization cannot be quantified so far. Based on this problem, the uncertainty of the insight also lacks proper quantification. Although the amount of visualization that can be created by visualization is a highly subjective process depending on the user, at least an approximation of the knowledge would be beneficial. This would contribute to classic VA as well as uncertainty-aware VA.

6.2 Open Problems that result from the inclusion of uncertainty

Selection of proper scenarios

A further open problem is when uncertainty-aware visual analytics is required. Naturally, the extension of the classic visual analytics cycle requires further resources. There might exist cases where the effect of uncertainty can be neglected or where the effort in extending an uncertainty-aware visual analytics cycle might be too big in comparison to the insight that is generated.

Survey of existing Techniques

We showed that there exists a variety of work that deals with uncertainty-aware visual analytics in many applications and for many data types. Still, a holistic state-of-the-art report in this area is missing. Such a report may be a good starting

point for researchers that start in the field and need to understand what possibilities they have. In addition, further open problems in the field could be identified.

Construction of Uncertainty-aware visual analytics cycles

In this work, we showed that a visual analytics cycle can be described. A logical next step would be to determine a standardized way to construct such a cycle. A good starting point might be the use of a classic visual analytics cycle, then deriving rules on how to provide uncertainty-awareness. There exist several approaches to construct a visual analytics cycle that may assist as a starting point.

Frameworks/Libraries with ready-to-use uncertainty-aware Visual Analytics Approaches.

In this work, we identified multiple steps in the uncertainty-aware VA pipeline that can be accomplished by existing methodologies. Examples are the determination and description of uncertainty-aware datasets, adaptation of preprocessing and hypothesis generation approaches, and provenance generation. In this context, frameworks or libraries that provide at least the uncertainty-aware visual analytic steps that can be standardized would be a massive contribution to the VA community. Gillmann et al. [32] provided a survey on uncertainty-awareness in open-source visualization solutions, which can be a great starting point for the creation of an uncertainty-aware VA framework. Still, the implementation of such a framework was not conducted so far.

Teaching of uncertainty-aware Principles.

Although uncertainty is an effect that is occurring in nearly all data acquisition processes, the application of uncertainty-aware analysis techniques, in general, is often a neglected point. This can be due to a variety of reasons. One major reason is that uncertainty-aware analysis principles are rarely taught to students. Here, lectures on uncertainty-aware VA would help new visualization researchers to understand the problems of data that is affected by uncertainty, giving them the awareness of principles that have to be kept in mind when dealing with uncertainty in datasets.

Approximation of Knowledge Uncertainty.

As mentioned before, the amount of uncertainty in insight can only be quantified in the case that the extracted knowledge is based on an uncertainty-aware hypothesis. This is an important open problem for uncertainty-aware VA as this distorts the feedback loop in the analysis cycle. Although we proposed two feedback cycle connections, the right one has to be picked. Here, suitable approaches to quantify insight and its uncertainty are highly requested.

The Missing Link between Ensemble Visualization and Uncertainty Visualization.

In contrast to uncertainty visualization, an ensemble visualization is concerned with visualizing multiple datasets representing the same captured scenario. Still, these disciplines are closely related. There are approaches available, where uncertainties can be generated from ensembles or ensembles that can be generated from an uncertainty distribution. Ensemble visualization is a highly active research field [98], providing a massive amount of VA solutions. Unfortunately, the link between these two disciplines is not defined properly. If one could arbitrarily transform ensemble datasets into uncertainty datasets, both disciplines could benefit from each other.

7 Conclusion

In this work, we described an uncertainty-aware visual analytics cycle. Here, the original visual analytics cycle is extended such that uncertainty can be quantified, propagated, and communicated in each component of the visual analytics cycle. This results in a holistic mechanism to tackle uncertainty originating from data, models, and humans in visual analytics approaches. We showed how to use this concept to tackle different types of input data as well as various use cases. As a result, we were able to formulate a variety of open problems originating from the visual analytics cycle and the incorporation of uncertainty.

8 Conflicts of Interest

The authors have no conflicts of interests to declare.

References

- [1] Charu C. Aggarwal and Chandan K. Reddy, *Data clustering: Algorithms and applications*, 1st ed., Chapman & Hall/CRC, 2013.
- [2] Marco Angelini and Giuseppe Santucci, *Visual cyber situational awareness for critical infrastructures*, Proceedings of the 8th International Symposium on Visual Information Communication and Interaction (New York, NY, USA), VINCI '15, ACM, 2015, pp. 83–92.
- [3] Marquis D' Argens, *Philosophical dissertations on the uncertainty of human knowledge. with some remarks on the theology of the grecian philosophers*, Gale Ecco, 04 2018.
- [4] Leon Arriola and James M Hyman, *Sensitivity analysis for uncertainty quantification in mathematical models*, Mathematical and statistical estimation approaches in epidemiology, Springer, 2009, pp. 195–247.
- [5] Sarita Bassil and Rudolf K Keller, *Software visualization tools: Survey and analysis*, Proceedings 9th International Workshop on Program Comprehension. IWPC 2001, IEEE, 2001, pp. 7–17.
- [6] Fabian Beck, Michael Burch, Stephan Diehl, and Daniel Weiskopf, *The State of the Art in Visualizing Dynamic Graphs*, EuroVis - STARs (R. Borgo, R. Maciejewski, and I. Viola, eds.), The Eurographics Association, 2014.
- [7] G. Belforte, B. Bona, and V. Cerone, *Bounded measurement error estimates: Their properties and their use for small sets of data*, Measurement **5** (1987), no. 4, 167 – 175.
- [8] Marcel Boumans, Giora Hon, and Arthur C Petersen, *Error and uncertainty in scientific practice*, Routledge, 2015.
- [9] Ajay Kumar Boyat and Brijendra Kumar Joshi, *A review paper: Noise models in digital image processing*, CoRR (2015), 63–75.
- [10] Dominique Brodbeck, Riccardo Mazza, and Denis Lalanne, *Interactive visualization-a survey*, Human machine interaction, Springer, 2009, pp. 27–46.
- [11] Guowei Cai and Sankaran Mahadevan, *Big data analytics in uncertainty quantification: Application to structural diagnosis and prognosis*, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering **4** (2018), no. 1, 04018003.
- [12] Simin Cai, Barbara Gallina, Dag Nyström, and Cristina Seceseanu, *Data aggregation processes: a survey, a taxonomy, and design guidelines*, Computing **101** (2019), no. 10, 1397–1429.
- [13] Peng Chen and Alfio Quarteroni, *A new algorithm for high-dimensional uncertainty quantification based on dimension-adaptive sparse grid approximation and reduced basis methods*, J. Comput. Phys. **298** (2015), no. C, 176–193.
- [14] R. Cheng, T. Emrich, H. Kriegel, N. Mamoulis, M. Renz, G. Trajcevski, and A. Züfle, *Managing uncertainty in spatial and spatio-temporal data*, 2014 IEEE 30th International Conference on Data Engineering, 2014, pp. 1302–1305.
- [15] C. D. Correa, Y. Chan, and K. Ma, *A framework for uncertainty-aware visual analytics*, 2009 IEEE Symposium on Visual Analytics Science and Technology, 2009, pp. 51–58.
- [16] M. Coutinho, Joshua S. Redford, B. Church, Alexandria C Zakrzewski, Justin J. Couchman, and J. D. Smith, *The interplay between uncertainty monitoring and working memory: Can metacognition become automatic?*, Memory & Cognition **43** (2015), 990–1006.
- [17] Wenqiang Cui, *Visual analytics: A comprehensive overview*, IEEE Access (2019), 1–1.
- [18] Aritra Dasgupta and Robert Kosara, *The need for information loss metrics in visualization*, Workshop on The Role of Theory in

Information Visualization, 2010.

- [19] Jay L. Devore, *Probability and statistics for engineering and the sciences*, 8th ed., Brooks/Cole, January 2011.
- [20] Mathew E Diamond, *Perceptual uncertainty*, PLoS biology **17** (2019), no. 8, e3000430.
- [21] Gulustan Dogan and Ted Brown, *Uncertainty modeling in wireless sensor networks*, Proceedings of the International Conference on Big Data and Internet of Thing (New York, NY, USA), BDIOT2017, ACM, 2017, pp. 200–204.
- [22] David W Engel, Kenneth D Jarman, Zhi-jie Xu, Bin Zheng, Alexandre M Tartakovsky, X Yang, Ramakrishna Tipireddy, Huan Lei, and Jian Yin, *Uq methods for hpda and cybersecurity models, data, and use cases*, Tech. report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2015.
- [23] Benjamin Enke and Thomas Graeber, *Cognitive uncertainty*, Tech. report, National Bureau of Economic Research, 2019.
- [24] José Miguel Fernández Güell, *How to approach urban complexity, diversity and uncertainty when involving stakeholders into the planning process*, (2017).
- [25] I K Fodor, *A survey of dimension reduction techniques*, Tech. report, Lawrence Livermore National Lab., CA (US), 2002.
- [26] Andrew U Frank, *Incompleteness, error, approximation, and uncertainty: An ontological approach to data quality*, Geographic Uncertainty in Environmental Security, Springer, 2007, pp. 107–131.
- [27] Yarin Gal, *Uncertainty in deep learning*, Ph.D. thesis, University of Cambridge, 2016.
- [28] C. Gillmann, Lucas Hempel, Carlo Schmidt, Dorothee Saur, and Gerik Scheuermann, *Visualizing Multi-Modal Deep Learning for Lesion Prediction*, IEEE Computer Graphics and Applications, 2021.
- [29] C. Gillmann, Maack, R.G.C., Post, T., Wischgoll, T., and Hagen, H., *An uncertainty-aware workflow for keyhole surgery planning using hierarchical image semantics*, Visual Informatics **2** (2018), no. 1, 26 – 36, Proceedings of PacificVAST 2018.
- [30] C. Gillmann, Noeska N. Smit, Eduard Gröller, Bernhard Preim, Anna Vilanova, and Thomas Wischgoll, *10 Challenges in Medical Visualization*, IEEE Computer Graphics and Applications, 2021.
- [31] Christina Gillmann, Pablo Arbelaez, Jose Hernandez, Hans Hagen, and Thomas Wischgoll, *An uncertainty-aware visual system for image pre-processing*, Journal of Imaging **4** (2018), no. 9, 109.
- [32] Christina Gillmann, Thomas Wischgoll, and Hans Hagen, *Uncertainty-awareness in open source visualization solutions*, 2016.
- [33] AA Giunta, MS Eldred, and JP Castro, *Uncertainty quantification using response surface approximations*, 9th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, Citeseer, 2004, pp. 26–28.
- [34] Daniel Greco, *Significance Testing in Theory and Practice*, The British Journal for the Philosophy of Science **62** (2011), no. 3, 607–637.
- [35] Henning Griethe, Heidrun Schumann, et al., *The visualization of uncertain data: Methods and problems.*, SimVis, 2006, pp. 143–156.
- [36] Shunan Guo, Fan Du, Sana Malik, Eun-yeo Koh, Sungchul Kim, Zhicheng Liu, Donghyun Kim, Hongyuan Zha, and Nan Cao, *Visualizing uncertainty and alternatives in event sequence predictions*, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA), CHI '19, ACM, 2019, pp. 573:1–573:12.

- [37] Charles D. Hansen, Min Chen, Christopher R. Johnson, Arie E. Kaufman, and Hans Hagen, *Scientific visualization: Uncertainty, multifield, biomedical, and scalable visualization*, Springer Publishing Company, Incorporated, 2014.
- [38] ———, *Scientific visualization: Uncertainty, multifield, biomedical, and scalable visualization*, Springer Publishing Company, Incorporated, 2014.
- [39] S. W. Hasinoff, F. Durand, and W. T. Freeman, *Noise-optimal capture for high dynamic range photography*, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 2010, pp. 553–560.
- [40] Troy M Hegel, Samuel A Cushman, Jeffrey Evans, and Falk Huettmann, *Current state of the art for statistical modelling of species distributions*, Spatial complexity, informatics, and wildlife conservation, Springer, 2010, pp. 273–311.
- [41] Julian Heinrich and Daniel Weiskopf, *State of the Art of Parallel Coordinates*, Eurographics 2013 - State of the Art Reports (M. Sbert and L. Szirmay-Kalos, eds.), The Eurographics Association, 2013.
- [42] Melanie Herschel, Ralf Diestelkämper, and Housseem Ben Lahmar, *A survey on provenance: What for? what form? what from?*, The VLDB Journal **26** (2017), no. 6, 881–906.
- [43] Patrick E Hoffman and Georges G Grinstein, *A survey of visualizations for high-dimensional data mining*, Information visualization in data mining and knowledge discovery, 2001, pp. 47–82.
- [44] Zhen Hu, Sankaran Mahadevan, and Xiaoping Du, *Uncertainty quantification in time-dependent reliability analysis*, International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 57083, American Society of Mechanical Engineers, 2015, p. V02BT03A062.
- [45] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay, *In pursuit of error: A survey of uncertainty visualization evaluation*, IEEE Transactions on Visualization and Computer Graphics **25** (2019), no. 1, 903–913.
- [46] Jessica Hullman, *Why authors don't visualize uncertainty*, 2019.
- [47] Fangxiang Jiao, Jeff M. Phillips, Jeroen Stinstra, Jens Krger, Raj Varma, Edward Hsu, Julie Korenberg, and Chris R. Johnson, *Metrics for uncertainty analysis and visualization of diffusion tensor images*, Medical Imaging and Augmented Reality (Berlin, Heidelberg) (Hongen Liao, P. J. “Eddie” Edwards, Xiaochuan Pan, Yong Fan, and Guang-Zhong Yang, eds.), Springer Berlin Heidelberg, 2010, pp. 179–190.
- [48] S. Jänicke, J. Focht, and G. Scheuermann, *Interactive visual profiling of musicians*, IEEE Transactions on Visualization and Computer Graphics **22** (2016), no. 1, 200–209.
- [49] Stefan Jänicke, Annette Geßner, Marco Büchler, and Grik Scheuermann, *Visualizations for text re-use*, 2014 International Conference on Information Visualization Theory and Applications (IVAPP), 2014, pp. 59–70.
- [50] Amin Karami, *A framework for uncertainty-aware visual analytics in big data*, CEUR Workshop Proceedings, vol. 1510, CEUR Workshop Proceedings, 2015, pp. 146–155.
- [51] Vasileios Kassiano, Anastasios Gounaris, Apostolos N Papadopoulos, and Kostas Tsichlas, *Mining uncertain graphs: An overview*, International Workshop of Algorithmic Aspects of Cloud Computing, Springer, 2016, pp. 87–116.
- [52] Daniel Keim and Leishi Zhang, *Solving problems with visual analytics: challenges and*

- applications*, Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies, 2011, pp. 1–4.
- [53] Daniel A Keim, Florian Mansmann, Jörn Schneidewind, Jim Thomas, and Hartmut Ziegler, *Visual analytics: Scope and challenges*, Visual data mining, Springer, 2008, pp. 76–90.
 - [54] Fadhela Kerdjoudj and Olivier Curé, *Evaluating Uncertainty in Textual Document*, URSW at ISWC (Bethlehem, United States), 2015.
 - [55] R. Khulusi, J. Kusnick, C. Meinecke, C. Gillmann, J. Focht, and S. Jänicke, *A survey on visualizations for musical data*, Computer Graphics Forum **39** (2020), no. 6, 82–110.
 - [56] Joe Michael Kniss, *Managing uncertainty in visualization and analysis of medical data*, 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2008, pp. 832–835.
 - [57] Jörn Kohlhammer, Thorsten May, and Marcus Hoffmann, *Visual analytics for the strategic decision making process*, GeoSpatial Visual Analytics (Dordrecht) (Raffaele De Amicis, Radovan Stojanovic, and Giuseppe Conti, eds.), Springer Netherlands, 2009, pp. 299–310.
 - [58] Kretzschmar, V., C. Gillmann, Günther, F., Stommel, M., and Scheuermann, G., *Statistically Informed Visualization Framework for Assisting Interface Optimization of Hybrid Component Design*, VMV, 25th International Symposium on Vision, Modeling, and Visualization, 2020.
 - [59] Chunggi Lee, Yeon-Rye Kim, Seungmin Jin, Dongmin Kim, R. Maciejewski, D. Ebert, and Sungahn Ko, *A visual analytics system for exploring, monitoring, and forecasting road traffic congestion*, IEEE Transactions on Visualization and Computer Graphics **26** (2020), 3133–3146.
 - [60] Guesuk Lee, W. Kim, H. Oh, B. D. Youn, and N. Kim, *Review of statistical model calibration and validation—from the perspective of uncertainty structures*, Structural and Multidisciplinary Optimization (2019), 1–26.
 - [61] Stephan Lewandowsky, Timothy Ballard, and Richard D. Pancost, *Uncertainty as knowledge*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **373** (2015), 1–11.
 - [62] Linna Li, Hyowon Ban, Suzanne Wechsler, and Bo Xu, *Spatial data uncertainty*, vol. 1, pp. 313–340, Elsevier, 12 2017.
 - [63] Guang Lin, David W Engel, and Paul W Eslinger, *Survey and evaluate uncertainty quantification methodologies*, Tech. report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2012.
 - [64] Dan R. Lipşa, Robert S. Laramée, Simon J. Cox, Jonathan C. Roberts, Rick Walker, Michelle A. Borkin, and Hanspeter Pfister, *Visualization for the physical sciences*, Comput. Graph. Forum **31** (2012), no. 8, 2317–2347.
 - [65] Shixia Liu, Gennady Andrienko, Yingcai Wu, Nan Cao, Liu Jiang, Conglei Shi, Yu-Shuen Wang, and Seokhee Hong, *Steering data quality with visual analytics: The complexity challenge*, Visual Informatics **2** (2018), no. 4, 191 – 197.
 - [66] Daniel P. Loucks and Eelco van Beek, *An introduction to probability, statistics, and uncertainty*, pp. 213–300, Springer International Publishing, Cham, 2017.
 - [67] R. G. C. Maack, M. Raymer, T. Wischgoll, H. Hagen, and Gillmann C., *A framework for uncertainty-aware visual analytics of proteins*, Special Section on Molecular Graphics and Visual Analysis of Molecular Data, Computers & Graphics, 2021.
 - [68] Alan MacEachren, Anthony Robinson, Susan Hopper, Steven Gardner, Robert Murray, Mark Gahegan, and Elisabeth

- Hetzler, *Visualizing geospatial information uncertainty: What we know and what we need to know*, Cartography and Geographic Information Science **32** (2005), 139–160.
- [69] Alan M. MacEachren, *Visual Analytics and Uncertainty: Its Not About the Data*, EuroVis Workshop on Visual Analytics (EuroVA) (E. Bertini and J. C. Roberts, eds.), The Eurographics Association, 2015.
- [70] H.R. Maier and B.A. Tolson, *Sensitivity and uncertainty*, Encyclopedia of Ecology (Sven Erik Jørgensen and Brian D. Fath, eds.), Academic Press, Oxford, 2008, pp. 3214 – 3221.
- [71] Michael Mastrandrea, Katharine Mach, Gian-Kasper Plattner, Ottmar Edenhofer, Thomas Stocker, Christopher Field, Kristie Ebi, and Patrick Matschoss, *The ipcc ar5 guidance note on consistent treatment of uncertainties: A common approach across the working groups*, Climatic Change **108** (2011), 675–691.
- [72] Chris Olston and Jock D. Mackinlay, *Visualizing data with bounded uncertainty*, Proceedings of the IEEE Symposium on Information Visualization (InfoVis’02) (USA), INFOVIS ’02, IEEE Computer Society, 2002, p. 37.
- [73] Pras Pathmanathan, Jonathan M. Cordeiro, and Richard A. Gray, *Comprehensive uncertainty quantification and sensitivity analysis for cardiac action potential models*, Frontiers in Physiology **10** (2019), 721.
- [74] Joseph J. Pfeiffer, *Using brightness and saturation to visualize belief and uncertainty*, Diagrammatic Representation and Inference (Berlin, Heidelberg) (Mary Hegarty, Bernd Meyer, and N. Hari Narayanan, eds.), Springer Berlin Heidelberg, 2002, pp. 279–289.
- [75] Catherine Plaisant, *The challenge of information visualization evaluation*, Proceedings of the Working Conference on Advanced Visual Interfaces (New York, NY, USA), AVI ’04, Association for Computing Machinery, 2004, p. 109–116.
- [76] Kristin Potter, Paul Rosen, and Chris R. Johnson, *From quantification to visualization: A taxonomy of uncertainty visualization approaches*, Uncertainty Quantification in Scientific Computing (Berlin, Heidelberg) (Andrew M. Dienstfrey and Ronald F. Boisvert, eds.), Springer Berlin Heidelberg, 2012, pp. 226–249.
- [77] Annie Preston, Maksim Gomov, and Kwan-Liu Ma, *Uncertainty-aware visualization for analyzing heterogeneous wildfire detections*, IEEE Computer Graphics and Applications **39** (2019), no. 5, 72–82.
- [78] Eric D Ragan, Alex Endert, Jibonananda Sanyal, and Jian Chen, *Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes*, IEEE transactions on visualization and computer graphics **22** (2015), no. 1, 31–40.
- [79] F. Raith, G. Scheuermann, and C. Gillmann, *Uncertainty-aware Detection and Visualization of Ocean Eddies in Ensemble Flow Fields - A Case Study of the Red Sea*, Proceedings of the Workshop on Visualisation in Environmental Sciences, 2021.
- [80] Ramesh Rebba and Sankaran Mahadevan, *Computational methods for model reliability assessment*, Reliability Engineering & System Safety **93** (2008), no. 8, 1197–1207.
- [81] Gordan Ristovski, Tobias Preußner, Horst K. Hahn, and Lars Linsen, *Uncertainty in medical visualization: Towards a taxonomy*, Computers & Graphics **39** (2014), 60–73.
- [82] Dominik Sacha, Hansi Senaratne, Bum Chul Kwon, Geoffrey Ellis, and Daniel A. Keim, *The role of uncertainty, awareness, and trust in visual analytics*, IEEE Transactions on Visualization and Computer Graphics **22** (2016), no. 1, 240–249.
- [83] Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto, *Sensitivity analysis in practice: A guide to*

assessing scientific models, Halsted Press, USA, 2004.

- [84] Christoph Schulz, Arlind Nocaj, Jochen Goertler, Oliver Deussen, Ulrik Brandes, and Daniel Weiskopf, *Probabilistic graph layout for uncertain network visualization*, IEEE Transactions on Visualization and Computer Graphics **23** (2017), no. 1, 531–540.
- [85] Hansi Vihara Senaratne, *Uncertainty-aware visual analytics for spatio-temporal data exploration*, Ph.D. thesis, Universität Konstanz, Konstanz, 2017.
- [86] Tim Sodergren, Jessica Hair, Jeff M Phillips, and Bei Wang, *Visualizing sensor network coverage with location uncertainty*, 2017 IEEE Visualization in Data Science (VDS), IEEE, 2017, pp. 52–59.
- [87] Xing Su, Sanat Talmaki, Hubo Cai, and Vineet R. Kamat, *Uncertainty-aware visualization and proximity monitoring in urban excavation: a geospatial augmented reality approach*, Visualization in Engineering **1** (2013), no. 1, 2.
- [88] Peter J Taylor and Ben Derudder, *World city network: a global urban analysis*, Routledge, 2016 (eng).
- [89] Roberto Therón and Juan Francisco De Paz, *Visual sensitivity analysis for artificial neural networks*, International Conference on Intelligent Data Engineering and Automated Learning, Springer, 2006, pp. 191–198.
- [90] Roberto Therón, Alejandro Benito Santos, Rodrigo Vicente, and Antonio Gómez, *Towards an uncertainty-aware visualization in the digital humanities*, Informatics **6** (2019), 31.
- [91] Sabine Timpf and Patrick Laube, *Advances in spatial data handling: Geospatial dynamics, geosimulation and exploratory visualization*, Springer Science & Business Media, 2012.
- [92] Margaret Varga and Caroline Varga, *Visual analytics: Data, analytical and reasoning provenance*, Building Trust in Information (Cham) (Victoria L. Lemieux, ed.), Springer International Publishing, 2016, pp. 141–150.
- [93] Corinna Vehlow, Jan Hasenauer, Andrei Kramer, Andreas Raue, Sabine Hug, Jens Timmer, Nicole Radde, Fabian Theis, and Daniel Weiskopf, *ivun: interactive visualization of uncertain biochemical reaction networks*, BMC bioinformatics **14 Suppl 19** (2013), S2.
- [94] Zana Vosough, Dietrich Kammer, Mandy Keck, and Rainer Groh, *Visualizing uncertainty in flow diagrams: A case study in product costing*, Proceedings of the 10th International Symposium on Visual Information Communication and Interaction (New York, NY, USA), VINCI '17, ACM, 2017, pp. 1–8.
- [95] Emily Wall, Leslie M Blaha, Celeste Lyn Paul, Kristin Cook, and Alex Endert, *Four perspectives on human bias in visual analytics*, Cognitive biases in visualizations, Springer, 2018, pp. 29–42.
- [96] M. Wallace and N. Platis, *The uncertain tag cloud*, 2015 10th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), 2015, pp. 1–5.
- [97] Hao Wang and Dit-Yan Yeung, *A survey on bayesian deep learning*, 2016.
- [98] Junpeng Wang, Subhashis Hazarika, Cheng Li, and Han-Wei Shen, *Visualization and visual analysis of ensemble data: A survey*, IEEE Transactions on Visualization and Computer Graphics **PP** (2018), 1–1.
- [99] Matthew Ward, Georges Grinstein, and Daniel Keim, *Interactive data visualization: Foundations, techniques, and applications*, A. K. Peters, Ltd., USA, 2010.
- [100] Norihiro Watanabe, Wenqing Wang, Christopher I. McDermott, Takeo Taniguchi, and Olaf Kolditz, *Uncertainty*

analysis of thermo-hydro-mechanical coupled processes in heterogeneous porous media, Computational Mechanics **45** (2009), no. 4, 263.

- [101] R. Wilson and G. H. Granlund, *The uncertainty principle in image processing*, IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-6** (1984), no. 6, 758–767.
- [102] Kai Xu, Alvitta Ottley, Conny Walchshofer, Marc Streit, Remco Chang, and John Wenskovich, *Survey on the analysis of user interactions and visualization provenance*, Computer Graphics Forum **39** (2020), no. 3, 757–783.