EuroVis Workshop on Visual Analytics (2019)
T. von Landesberger and C. Turkay (Editors)

Quantifying Uncertainty in Multivariate Time Series Pre-Processing

C. Bors[1] and J. Bernard[2] and M. Bögl[1] and T. Gschwandtner[1] and J. Kohlhammer[3] and S. Miksch[1]

[1]TU Wien, Vienna, Austria
[2]TU Darmstadt, Darmstadt, Germany
[3]Fraunhofer IGD, Darmstadt, Germany

Abstract

In multivariate time series analysis, pre-processing is integral for enabling analysis, but inevitably introduces uncertainty into the data. Enabling the assessment of the uncertainty and allowing uncertainty-aware analysis, the uncertainty needs to be quantified initially. We address this challenge by formalizing the quantification of uncertainty for multivariate time series pre-processing. To tackle the large design space, we elaborate key considerations for quantifying and aggregating uncertainty. We provide an example how the quantified uncertainty is used in a multivariate time series pre-processing application to assess the effectiveness of pre-processing steps and adjust the pipeline to minimize the introduction of uncertainty.

CCS Concepts

• ***Mathematics of computing*** → ***Time series analysis;*** • ***Information systems*** → ***Uncertainty;*** • ***Human-centered computing*** → ***Visualization theory, concepts and paradigms;*** *Visual analytics;* • ***Computing methodologies*** → ***Uncertainty quantification;***

1. Introduction

In Visual Analytics (VA) research and related fields, the awareness and need to incorporate uncertainty information into the analysis has increased considerably. This holds true for both a methodological, design, and implementation perspective. How uncertainty was introduced into the data can be distinguished by the different sources of uncertainty, including observations inherent to the data, generated by models or simulations, or introduced by the processing or visualization processes [PRJ12, BHJ*14]. Even though pre-processing inevitably introduces uncertainty by altering the original data, these routines are rarely analyzed towards their impact on uncertainty. When analyzing multivariate time series (MVTS), pre-processing is an integral part to enable further analysis. Several approaches analyze uncertainty introduced by pre-processing [CCM09, WYM12], aggregating uncertainty for individual processing steps. When assessing the influence of uncertainty on MVTS, inappropriate aggregation would omit temporal characteristics that can also be affected by processing (e.g., rastering [BBGM17], or sampling).

When pre-processing MVTS, a common processing pipeline would consist of multiple steps: (1) imputing missing values, (2) performing linear interpolation, (3) smoothing the time series by applying a moving average kernel, and (4) sampling the data to reduce the size. This implies that first we need to assess how individual processes influence uncertainty, but also how subsequent steps of the analysis are affected. To allow this, it is necessary to inspect the MVTS and corresponding uncertainties in detail, in order to audit the impact of single pre-processing routines. Another challenge regards concatenation of pre-processing steps: Consecutively executing different data transformation steps also propagates uncertainties throughout the processes. This makes it increasingly difficult to determine how much uncertainty was introduced at which step. Adequately monitoring uncertainty during pre-processing allows identifying individual steps that alter the value or temporal domain inappropriately.

We address the special challenges of uncertainty quantification when pre-processing MVTS. Our contributions are: (1) A formalization of uncertainty quantification and aggregation, addressing the particularity of the temporal domain, relevant for MVTS pre-processing. (2) An elaboration of important considerations for the quantification and aggregation of uncertainty, along with selected examples. (3) A usage scenario that shows how uncertainty can be quantified in a concrete MVTS pre-processing application.

2. Related Work

Uncertainty Quantification. Bonneau et al. [BHJ*14] defined uncertainties to be observable from different sources: (1) uncertainty from sampled data, (2) uncertainty generated by models or simulations, and (3) uncertainty introduced by data processing. Wu et al. [WYM12] quantified and visualized uncertainty in processing pipelines as error ellipsoids for every employed routine and communicated the extent and propagation of uncertainty throughout the pipeline in a flow visualization. Correa et al. [CCM09] stated that uncertainties are propagated and aggregated throughout data transformations in parametric modeling of data distributions, but also stated that these propagation and aggregation steps are applicable on more general data transformation techniques. However, they encountered the difficulty of analyzing the impact of uncertainty on single dimensions or variables, and motivated analyzing uncertainty locally throughout the transformations.

Uncertainty in Visual Analytics. Visualization of uncertainty has disseminated into most application and research domains, like scientific-, information-, geographic (spatio-temporal) [MRH*05], and workflow visualization [WYM12], visual analytics [Mac15], and time series analysis [GBFM16, WBFvL17]. Uncertainty influences users' decision-making, awareness of uncertainty can build trust and reduce user errors [SSK*16] but also affect risk assessment and perception [KMRS17]. Sacha et al. [SSK*16] incorporated the notion of uncertainty in a knowledge generation model for visual analytics to determine how it should be adequately generated and propagated. This underlines the need for appropriate integration and support of uncertainty in visual analytics. Seipp et al. [SOGV16] argued for uncertainty information to be available to the user at any stage in the sensemaking process, including pre-processing. Even though recent approaches aimed at integrating uncertainty into pre-processing and data quality assessment approaches [BBB*18, BBGM17], the inspection of uncertainty information produced by models or processing algorithms along a pipeline remains an open challenge [LFR17].

Visual Analytics and Pre-Processing. Liu et al. [LAW*18] presented a framework for steering data quality, and identified that data pre-processing and analysis can introduce uncertainty. Pre-processing of MVTS can have unforeseeable effects on the data, visual interactive support helps users assess the impact on the time series [BRG*12]. VA approaches have been employed to determine quality issues in time series [GAM*14]. Bors et al. [BBGM17] attempted to derive uncertainty from pre-processing utilizing domain knowledge of the processing routines and temporal domain characteristics, like temporal granularity and temporal deviation. However, this approach lacks multivariate aspects and limits uncertainty quantification to one pre-processing step. We iterate on this approach of integrating uncertainty in pre-processing. Building on the aspects addressed by [SSK*16, BBB*18, LFR17, PRJ12, LAW*18], we condensed a formalization of uncertainty quantification for pre-processing MVTS.

3. Quantifying and Aggregating Uncertainty

We identified two challenges regarding pre-processing of MVTS: How can uncertainty be consistently quantified for this type of data, and how can multiple pre-processing steps be assessed and compared towards their impact on uncertainty. To effectively quantify uncertainty from pre-processing MVTS and ultimately allow visual analysis, the time and variables (also referred to as data dimensions) of the MVTS, and the pre-processing steps span a cube of dimensions (see Figure 1a) that influence uncertainty introduced by MVTS pre-processing. We elaborate how uncertainty can be quantified and aggregated in different ways, and describe why quantification and aggregation depends on the above mentioned dimensions.

3.1. Quantifying Uncertainties

We define the uncertainty quantification for the three dimensions of the cube shown in Figure 1a: time and variables of the MVTS, and pre-processing steps. We refer to a p-dimensional time series data by $\mathbf{X} = \{\mathbf{x}_{(t_1,v)}, \dots, \mathbf{x}_{(t_n,v)}\}$ measured at time point $t_1, \dots, t_n$ with variables $v = 1, \dots, p$ (cf. Figure 1b). A pre-processing pipeline for MVTS consists of m pre-processing steps that modify the MVTS and introduce uncertainty. Each pre-processing step s takes a MVTS $\mathbf{X}_{s-1} = \{\mathbf{x}_{(t_1,v,s-1)}, \dots, \mathbf{x}_{(t_1,v,s-1)}\}$ as input and generates a modified MVTS $\mathbf{X}_s = \{\mathbf{x}_{(t_1,v,s)}, \dots, \mathbf{x}_{(t_1,v,s)}\}$ which is the input of the next step. $\mathbf{X}_0$ is the MVTS as input to the whole pre-processing pipeline, $\mathbf{X}_m$ the resulting MVTS, and $\mathbf{X}_s$ with $s = 1, \dots, m-1$ the MVTS between the single pre-processing steps. The natural atomic representation of uncertainty for such a processing step is determined by the quantification function $u(\mathbf{X}_s, \mathbf{X}_{s-1})$ that computes the uncertainty per timestamp and variable $u(x_{(t,v,s)}, x_{(t,v,s-1)})$. However, depending on the pre-processing operation, the uncertainty quantification can only be done on a specific level of granularity, if the temporal domain or the dimensionality of the MVTS are affected. In the following we discuss the different cube dimensions' dependencies on quantification.

Dependency on Variables. If MVTS variables are individually analyzed, it is sufficient to determine the absolute value difference between the input and output time series of a pre-processing step: $u_{abs}(abs(z_{(t,v)}))$, where $z_{(t,v)} = x_{(t,v,s)} - x_{(t,v,s-1)}$ denotes the value difference. This results in an uncertainty value that is value domain dependent, as it needs to be considered in the context of the respective scale of the value domain. Thus, if uncertainties of variables with different value domains are to be compared or assessed simultaneously, normalized relative differences need to be determined instead: $u_{rel}(z_{(t,v)}) = \frac{z_{(t,v)} - \mu_z}{\sigma_z}$, where μ_z is the mean difference and σ_z the deviation. This way, the influence of multiple variables on the uncertainty at time $x_{(t,s)}$ is comparable for any v. If the uncertainty of each variable cannot be quantified for single time points, the uncertainty needs to be computed for single variables across all time points $u_t(x_{(v,s)}, x_{(v,s-1)})$. This is for example the case, if the temporal space is modified, like temporal sampling or rastering (only u_v is applicable).

Dependency on Time. The quantification of uncertainty over single time points and dimensions $u(x_{(t,v,s)}, x_{(t,v,s-1)})$ allows to identify time points or time ranges that have a high, low, or normal level of uncertainty in the value domain. If the uncertainty of time points cannot be quantified for single variables, the uncertainty needs to be computed for single time points across all variables $u_v(x_{(t,s)}, x_{(t,s-1)})$. This is for example the case, if the time series dimensionality is altered, e.g., by dimensionality reduction routines (only u_t is applicable). In the case of aggregating over time (cf. Section 3.2), e.g., for rastering or sampling a time series to a coarser temporal granularity, the uncertainty introduced in the temporal domain needs to be considered in the quantification. This can be done by computing the relative or absolute temporal differences Δt of all time points that are merged in the raster intervals of the coarser granularity level, similarly to computing relative value differences formalized for variables, but in the temporal domain.

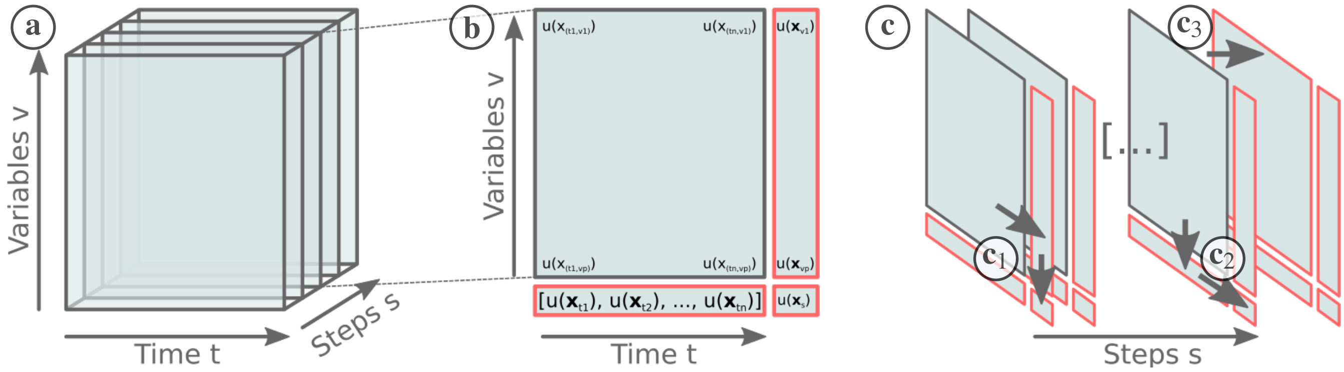

Figure 1: *Illustration of quantification of uncertainties and aggregation on values and uncertainties. (a) shows the three variables time, variables, and processing steps. (b) represents a single processing step slice with the dimensions time and variables together with the uncertainty aggregation, either across time or variables (shown as red boxes). (c) indicates the different aggregation paths within single processing step slices (c_1, c_2) and across all steps (c_3).*

Dependency on Pre-Processing Steps. Each pre-processing method has different effects on the introduced uncertainty. However, these effects can be derived when taking into account the error that is introduced by the specific method and its parametrization. Moreover, this on average introduced error can be estimated (e.g., moving average changes the value domain consistently). We formalize the introduced uncertainty accordingly: $u_{err}(x_{(t,s)}) = f_{err}(x_{(t,s)}, \mathbf{k})$, where f_{err} is an error function for quantifying uncertainty, and $\mathbf{k} = \{k_1, \dots, k_l\}$ is the current parameter vector of the pre-processing method.

3.2. Aggregating Uncertainties

Figure 1c illustrates the different types of aggregation of uncertainties over all processing steps. As with quantifying uncertainty, aggregation can be applied on all of the cube's dimensions: time and variables of the MVTS, and pre-processing steps. Generally it is advisable to quantify uncertainty at the finest granularity level and aggregate to coarser granularities if necessary. In the following we use a general $agg_{i=1}^{n}(\cdot)$ function to indicate that there are multiple different aggregation methods that could be applied. More specifically this can be a simple summarization $\sum_{i=1}^{n}(\cdot)$, a multiplication $\prod_{i=1}^{n}(\cdot)$, or other statistical aggregations of uncertainty, like the mean uncertainty $\mu(u)$, mean squared uncertainty $\mu(u^2)$, or root mean squared uncertainty $\sqrt{\mu(u^2)}$.

Aggregating by Time. Quantifying uncertainty on timestamp granularity is not always beneficial. Analogous to visualization of large MVTS, aggregating uncertainty to a coarser temporal granularity allows maintaining a representative dataset if the scale of the original data is too large. Aggregating uncertainty can be done on different levels of temporal granularity (cf. Fig. 1c_1). To remove the temporal dimension from the quantified uncertainty, we can aggregate over the entire time dimension $u(\mathbf{x}_{(v,s)}) = agg_{t=1}^{n}(u(x_{t,v,s}, x_{t,v,s-1}))$. This allows an abstract representation of uncertainty without time, e.g., a single value of uncertainty for an entire time series variable v, and pre-processing step s.

Aggregating by Variables. Analyzing uncertainty of individual variables allows detailed inspection of effects on the value domain. However, variables can be affected differently by pre-processing. Uncertainty can be aggregated by variables $u(\mathbf{x}_{(t,s)}) = agg_{v=1}^{p}(u(x_{t,v,s}, x_{t,v,s-1}))$ to determine a single value of uncertainty for these variables, e.g., $\mu(u(\mathbf{x}_{(t,s)}))$ (cf. Fig. 1c_2).

Aggregating by Pre-Processing Steps. To obtain an overview of uncertainties for one step s of the pre-processing, we compute the uncertainty of each pre-processing step $u(\mathbf{x}_s)$. Comparison of different steps can be done on different levels of aggregation, by variable $u(\mathbf{x}_{(t,s)}) = agg_{v=1}^{p}(u(x_{t,v,s}, x_{t,v,s-1}))$ or time $u(\mathbf{x}_{(v,s)}) = agg_{t=1}^{n}(u(x_{t,v,s}, x_{t,v,s-1}))$. However, it is also possible to aggregate over a whole pre-processing pipeline, to assess the introduced uncertainty of a sequence of pre-processing steps $u(\mathbf{x}_{(t,v)}) = agg_{s=1}^{m}(u(x_{t,v,s}, x_{t,v,s-1}))$.

To enable more distinct assessment, aggregation can be nested consecutively. Aggregating by variables allows comparison over time $u(\mathbf{x}_{(t)}) = agg_{v=1}^{p} agg_{s=1}^{m}(u(x_{v,t,s}, x_{v,t,s-1})$. This allows more detailed inspection if the time series was affected by pre-processing uniformly. Conversely, aggregating by time allows comparison over variables $u(\mathbf{x}_{(v)}) = agg_{t=1}^{n} agg_{s=1}^{m}(u(x_{v,t,s}, x_{v,t,s-1}))$. Ultimately, aggregating over time, variables, and pre-processing steps produces a single value of uncertainty for the entire pre-processing pipeline $u(\mathbf{x}) = agg_{t=1}^{n} agg_{v=1}^{p} agg_{s=1}^{m}(u(x_{v,t,s}, x_{v,t,s-1}))$ (cf. Fig. 1c_3).

4. Usage Scenario: Cleansing and Reduction of MVTS Data

The MVTS processed in the scenario contains weather experiment data measured in Antarctica [RLKLI12] and used by our collaborator for downstream analysis. We exemplify the use of uncertainty quantification in a visual analytics tool for pre-processing of MVTS presented by Bernard et al. [BHR*19] to support analysis scenarios with uncertainty on different aggregation levels (Please be referred to this work for a detailed description of the interactive VA approach). Among others, it enables the assessment of (a) uncertainty introduced by a pre-processing step (cf. Figure 2), (b) uncertainty

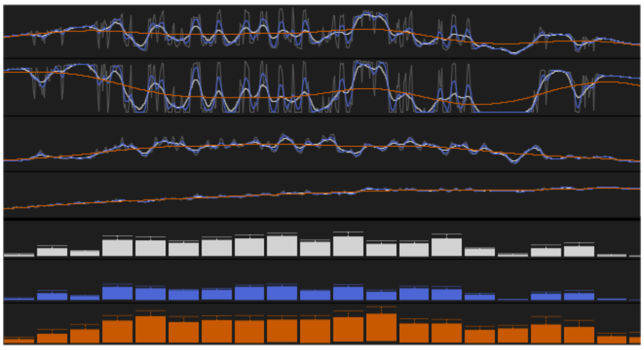

Figure 2: *Analysis of a Moving Average pre-processing step: Multiple MVTS dimensions are visualized with three different parameter settings (top), for each parameter uncertainty is aggregated by dimensions and time to give three boxplots over time (bottom). Reprinted from [BHR*19].*

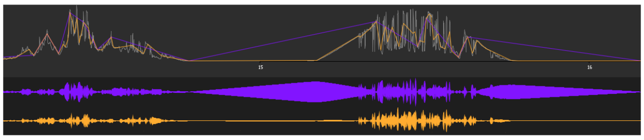

Figure 3: *Assessment of uncertainty introduced by a sampling routine for one dimension, applied with two parameter values (purple, orange). The purple parametrization is too coarse, introducing a considerable amount of uncertainty. Reprinted from [BHR*19].*

influencing individual and multiple variables, and (c) uncertainty influenced by alternative pre-processing parameter values (cf. Figure 3). For all steps and parameters used in the following examples, uncertainty is quantified as the normalized relative difference on a timestamp and individual variable level, $u_{rel}(z_{(t,v)})$.

First, we highlight how the collaborator applies a smoothing routine to remove noise and reduce the effect of outliers, i.e., to improve data quality. Figure 2 shows how the effect of the smoothing routine can be assessed for four dimensions and three different parametrizations (gray, blue, orange linecharts). Using aggregation by variable allows assessment of the average uncertainty across all selected dimensions, aggregation by time allows analysis of the uncertainty introduced for cyclic patterns observed in the first two dimensions. The orange boxplots on the bottom (cf. Fig. 2) indicate a considerably higher uncertainty with this parametrization and removes the cyclic patterns entirely. The collaborator proceeds by adding a sampling routine with two sampling window sizes, aiming for a more compact MVTS. To grasp the effect of the sampling routine at a fine-grained level, the collaborator inspects the sampling results (cf. Fig. 3) of one individual dimension of the MVTS (top purple and orange linecharts) and the corresponding uncertainties (bottom symmetric area charts), meaning we don't apply aggregation in the variable domain. It shows that the purple sampling routine introduces excessive uncertainty, due to a too coarse sampling kernel. Finally, the collaborator wants to validate the pipeline as a whole. Again, an adequate level of aggregation is used to exhibit the uncertainty of several routines. The uncertainties are aggregated over all variables, but shown for every pre-processing and time stamp individually. That way, the collaborator can identify which routines introduced the largest amount of uncertainty in comparison to the others.

With the visual analytics approach building upon our methodology, the collaborator was able to conduct the uncertainty-aware pre-processing of MVTS. She was able to make informed decisions in the creation as well as in the validation phase. Without a visual-interactive approach, selection adequate parameters would have required iterative comparison of intermittent processing results.

5. Discussion & Conclusion

In this paper we presented a formalization for quantifying and aggregating uncertainty that was introduced by pre-processing of MVTS and we identified the dimensions that affect the way uncertainty needs to be quantified and aggregated.

We distinguish uncertainty at the time stamp level, the data variable level, as well as uncertainty introduced at each step of a data pre-processing pipeline. We argue that uncertainty should be quantified for the finest granularity level possible (i.e., for each time stamp, data variable, and pre-processing step), as aggregated uncertainty values are not sufficient for all analysis tasks. If coarser uncertainty information is required to support an effective analysis, this fine-grained uncertainty can subsequently be aggregated. On the other hand, it is not always possible to quantify uncertainty at the finest granularity level. Some pre-processing methods transform the granularity of the MVTS, such as dimensionality reduction or temporal sampling. This change of granularity needs to be considered in the employed uncertainty quantification method, as a simple comparison of input and output values of the pre-processing step is not feasible in such cases. Moreover, we elaborated on the different possibilities for uncertainty aggregation. Finally, we presented a use case of how our formalization can be applied to quantify uncertainty in a visual interactive pre-processing environment and how different uncertainty aggregations support analyzing and fine-tuning of the pre-processing pipeline.

While the visual representation of uncertainty information and the need to include information about the uncertainty of the data that is visualized into VA environments gains awareness, it is often assumed that the uncertainty information is given. Yet, almost any data analysis is preceded by data pre-processing which also introduces considerable uncertainty into the data. Thus, we formalized the quantification and aggregation of uncertainty from MVTS pre-processing. This might be done to evaluate the appropriateness of the pre-processing pipeline as such, but also to include this uncertainty information into the final data representation to foster informed reasoning. Our formalization helps visualization designers to understand and consider relevant aspects in this context.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) and the Austrian Science Fund (FWF), Project No. I 2850 (-N31), Lead Agency Procedure (D-A-CH) "Visual Segmentation and Labeling of Multivariate Time Series (VISSECT)".

References

[BBB*18] BERNARD J., BORS C., BÖGL M., EICHNER C., GSCHWANDTNER T., MIKSCH S., SCHUMANN H., KOHLHAMMER J.: Combining the Automated Segmentation and Visual Analysis of Multivariate Time Series. In *EuroVis Workshop on Visual Analytics (EuroVA)* (2018), The Eurographics Association. doi:10.2312/eurova.20181112. 2

[BBGM17] BORS C., BÖGL M., GSCHWANDTNER T., MIKSCH S.: Visual Support for Rastering of Unequally Spaced Time Series. In *Proceedings of the 10th International Symposium on Visual Information Communication and Interaction* (New York, NY, USA, 2017), VINCI '17, ACM, pp. 53–57. doi:10.1145/3105971.3105984. 1, 2

[BHJ*14] BONNEAU G.-P., HEGE H.-C., JOHNSON C. R., OLIVEIRA M. M., POTTER K., RHEINGANS P., SCHULTZ T.: Overview and State-of-the-Art of Uncertainty Visualization. In *Scientific Visualization*, Mathematics and Visualization. Springer, London, 2014, pp. 3–27. URL: https://link.springer.com/chapter/10.1007/978-1-4471-6497-5_1, doi:10.1007/978-1-4471-6497-5_1. 1

[BHR*19] BERNARD J., HUTTER M., REINEMUTH H., PFEIFER H., BORS C., KOHLHAMMER J.: Visual-Interactive Preprocessing of Multivariate Time Series Data. *Computer Graphics Forum (EuroVis)* (2019). 3, 4

[BRG*12] BERNARD J., RUPPERT T., GOROLL O., MAY T., KOHLHAMMER J.: Visual-Interactive Preprocessing of Time Series Data. In *SIGRAD* (2012), Kerren A., Seipel S., (Eds.), vol. 81 of *Linköping Electronic Conference Proceedings*, Linköping University Electronic Press, pp. 39–48. 2

[CCM09] CORREA C., CHAN Y.-H., MA K.-L.: A framework for uncertainty-aware visual analytics. In *Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on* (2009), IEEE, pp. 51–58. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5332611. 1

[GAM*14] GSCHWANDTNER T., AIGNER W., MIKSCH S., GÄRTNER J., KRIGLSTEIN S., POHL M., SUCHY N.: TimeCleanser: A Visual Analytics Approach for Data Cleansing of Time-oriented Data. In *Proceedings of the 14th International Conference on Knowledge Technologies and Data-driven Business* (New York, NY, USA, 2014), i-KNOW '14, ACM, pp. 18:1–18:8. URL: http://doi.acm.org/10.1145/2637748.2638423, doi:10.1145/2637748.2638423. 2

[GBFM16] GSCHWANDTNER T., BÖGL M., FEDERICO P., MIKSCH S.: Visual Encodings of Temporal Uncertainty: A Comparative User Study. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan. 2016), 539–548. doi:10.1109/TVCG.2015.2467752. 2

[KMRS17] KINKELDEY C., MACEACHREN A. M., RIVEIRO M., SCHIEWE J.: Evaluating the effect of visually represented geodata uncertainty on decision-making: systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science 44*, 1 (Jan. 2017), 1–21. URL: https://doi.org/10.1080/15230406.2015.1089792, doi:10.1080/15230406.2015.1089792. 2

[LAW*18] LIU S., ANDRIENKO G., WU Y., CAO N., JIANG L., SHI C., WANG Y.-S., HONG S.: Steering data quality with visual analytics: The complexity challenge. *Visual Informatics 2*, 4 (Dec. 2018), 191–197. URL: http://www.sciencedirect.com/science/article/pii/S2468502X18300573, doi:10.1016/j.visinf.2018.12.001. 2

[LFR17] LANDESBERGER T. V., FELLNER D. W., RUDDLE R. A.: Visualization System Requirements for Data Processing Pipeline Design and Optimization. *IEEE Transactions on Visualization and Computer Graphics 23*, 8 (Aug. 2017), 2028–2041. doi:10.1109/TVCG.2016.2603178. 2

[Mac15] MACEACHREN A. M.: Visual Analytics and Uncertainty: Its Not About the Data. In *EuroVis Workshop on Visual Analytics (EuroVA)* (2015), The Eurographics Association. doi:10.2312/eurova.20151104. 2

[MRH*05] MACEACHREN A. M., ROBINSON A., HOPPER S., GARDNER S., MURRAY R., GAHEGAN M., HETZLER E.: Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science 32*, 3 (2005), 139–160. URL: http://www.tandfonline.com/doi/abs/10.1559/1523040054738936. 2

[PRJ12] POTTER K., ROSEN P., JOHNSON C. R.: From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches. *IFIP advances in information and communication technology 377* (2012), 226–249. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4319674/, doi:10.1007/978-3-642-32677-6_15. 1, 2

[RLKLI12] RIMBU N., LOHMANN G., KÖNIG-LANGLO G., IONITA M.: 30 years of synoptic observations from Neumayer Station with links to datasets., 2012. 3

[SOGV16] SEIPP K., OCHOA X., GUTIÉRREZ F., VERBERT K.: A Research Agenda for Managing Uncertainty in Visual Analytics. URL: http://dl.gi.de/handle/20.500.12116/316, doi:10.18420/muc2016-ws11-0007. 2

[SSK*16] SACHA D., SENARATNE H., KWON B. C., ELLIS G., KEIM D. A.: The Role of Uncertainty, Awareness, and Trust in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (Jan. 2016), 240–249. doi:10.1109/TVCG.2015.2467591. 2

[WBFvL17] WUNDERLICH M., BALLWEG K., FUCHS G., VON LANDESBERGER T.: Visualization of Delay Uncertainty and its Impact on Train Trip Planning: A Design Study. *Computer Graphics Forum 36*, 3 (June 2017), 317–328. URL: http://doi.wiley.com/10.1111/cgf.13190, doi:10.1111/cgf.13190. 2

[WYM12] WU Y., YUAN G.-X., MA K.-L.: Visualizing flow of uncertainty through analytical processes. *Visualization and Computer Graphics, IEEE Transactions on 18*, 12 (2012), 2526–2535. URL: https://ieeexplore.ieee.org/document/6327258. 1, 2