Managing Uncertainty in Evolving Geo-Spatial Data

Andreas Züfle[1], Goce Trajcevski[2], Dieter Pfoser[1], and Joon-Seok Kim[1]

[1]Dep. of Geography and Geoinformation Science, George Mason University, {azufle,dpfoser,jkim258}@gmu.edu

[2]Department of Electrical and Computer Engineering, Iowa State University, gocet25@iastate.edu

***Abstract*—Our ability to extract knowledge from evolving spatial phenomena and make it actionable is often impaired by unreliable, erroneous, obsolete, imprecise, sparse, and noisy data. Integrating the impact of this uncertainty is a paramount when estimating the reliability/confidence of any time-varying query result from the underlying input data. The goal of this advanced seminar is to survey solutions for managing, querying and mining uncertain spatial and spatio-temporal data. We survey different models and show examples of how to efficiently enrich query results with reliability information. We discuss both analytical solutions as well as approximate solutions based on geosimulation.**

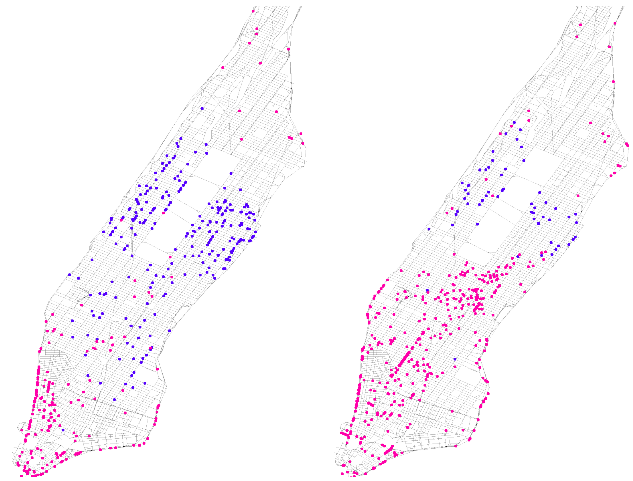
I. Introduction

As the volume, variety and velocity (in terms of the rate of recorded values) of evolving spatial data has increased sharply over the last decades, the impact of uncertainty on all the aspects of data management – i.e., representation, storage, query processing – has increased as well. Until the late 20^{th} century, spatial data available for geographic information science (GIS) was mainly collected, curated, standardized [15], and published by authoritative sources such as the United States Geological Survey (USGS) [64], and data was collected by sensors with known locations (e.g., roadside sensors, sensor stations, etc.) [44, 65]. More recently, data is often obtained from sources of volunteered geographic information (VGI) [16, 57, 46]. Consequentially, our ability to unearth valuable knowledge from large sets of such spatial data is often impaired by the uncertainty of the data which geography has named the "the Achilles heel of GIS" [18] for many reasons:

- Imprecision is caused by physical limitations of sensing devices and connection errors, e.g., using cell-phone GPS [10].
- Data records may be obsolete. In geo-social networks and microblogging platforms such as Twitter, users may update their location infrequently, yielding uncertain location information in-between data records [37].
- Data can be obtained from unreliable sources, such as volunteered geographic information like data in Open-Street-Map obtained from individual users, which may incur inaccurate or plain wrong data, either deliberately or due to human error [21, 16].
- Data sets pertaining to specific questions may be too small to answer questions reliably. Proper statistical inference is required to draw significant conclusions from the data and to avoid basing decisions upon spurious mining results [23, 6].

The main objective of this advanced seminar is to present to the audience the spectrum of effective and efficient solutions to various problems related to modeling, querying and mining of uncertain evolving geo-spatial data, which have been published in the recent years – and, more importantly, place them into proper context in terms of their applicability, existing tools, and various trade-offs.

After a discussion of motivational scenarios, the first part of this advanced seminar will surveys discrete and continuous uncertainty models, as well as the most common approach to interpret an uncertain database using *possible worlds semantics*. We will describe computational challenges of possible worlds semantics and show examples of efficient algorithm to deal with these challenges. This part will also show how query and mining results on possible worlds can be aggregated into representative results and how these results can be enriched with probabilistic guarantees and solutions for probabilistic graph will be surveyed. From a complementary, geostatistical view, we will introduce kriging approaches for spatial interpolation, as well as the impact of evolution of the spatial phenomena, in the context of handling attribute uncertainty.



(a) Undersupply at 10:20 am (b) Oversupply at 11:00 am

Fig. 1: Ride-hailing Geosimulation in New York City

The second part of the tutorial will address uncertainty in spatio-temporal data, centered around three basic themes: (i) infrequently sampled trajectory data for moving point-objects; (ii) discrete sampling and impact on representing deformable continuous phenomena (e.g., shapes of flood); and (iii) combining data from different sources (e.g., location sensors and social networks). We will show what is the impact of relevant parameters changing over time, and how to predict and interpolate such parameters through time and space. We will discuss solutions based on different paradigms – e.g., using space-time geometric models, and fitting stochastic processes on training data of the past, to obtain a model describing the current and the future values of the attributes of interest and queries' answers.

In the third part of the tutorial we will survey geosimulation approaches to approximately query and mine uncertain data. We will show how geosimulation approaches such as Monte-Carlo sampling and agent-based modeling can be leveraged to obtain possible worlds, and how results on possible worlds can be aggregated to provide probabilistic guarantees. As a case study, we will explain the geosimulation approach for improving ride-hailing systems which won ACM SIGSPATIAL

Cup 2019 [31]. Figure 1 shows a screenshot of this system that will be demonstrated to explain the uncertainty in ride-hailing systems. This system shows drivers as pink dots and unassigned passengers as blue dots in New York City. This system will be used to demosntrate how geosimulation can be used to calibrate a ride-hailing system to achieve better assignments between drivers and users. For all of the presented state-of-the-art solutions, we will discuss both the challenges of: (i) *effectiveness in uncertain data* – which is, to correctly determine the set of possible results, each associated with the correct probability of being a result, in order to give a user a confidence about the returned results; and (ii) *efficiency* – which is, to enable fast computations for these results and corresponding probabilities, allowing reasonable querying times, even for large uncertain databases. The main objectives of this tutorial are:

- Provide an introduction to research issues and solutions addressing uncertainty in spatial and spatio-temporal data. This overview is aimed at students with no prior experience in the field, as well as at attendants with some background.
- Extend this overview from spatial data to spatio-temporal data, which is particular for the MDM community often using volunteered or crowd-sourced mobility data.
- Provide examples of techniques in the field of efficient management of geo-spatial data, catering to a broad audience. The aim is to teach the background necessary for researchers to contribute to the field of uncertain data management.
- Provide practitioners with examples and use-cases showing how to leverage uncertain data management for better decision making.

Several tutorials on managing, querying and mining uncertain data have been presented in the recent past [53, 48, 72], touching some of the topics of this tutorial (some of them by the proposers). However, the large parts of this tutorial (constituting at least 40% of the content) have not been presented at any previous tutorial, including:

- Novel solutions for handling uncertain spatial data published since 2014, including the generating functions technique for efficiently answering spatial queries on uncertain data.
- a focus towards approximate solutions that allow to quickly derive confidence intervals to bound the true but unknown query or mining result. Previous tutorials focused on exact solutions, which will be surveyed in less details and constitute no more than 25% of the total tutorial material;
- Solutions for handling uncertain spatial using geosimulation, including recent approaches using Monte-Carlo simulation [55, 71, 56];
- Solutions based on geosimulation [29] to quantify uncertainty in complex spatial systems, including a new case-study leveraging geosimulation for optimizing ride-sharing systems as used by the winner of SIGSPATIAL GIS Cup 2019 [31].

While not mentioned explicitly in the MDM 2020 Call for Research Papers, the management of uncertainty in spatio-temporal data is an aspect paramount to researchers working with personal location data. The techniques presented in this tutorial will provide MDM attendees with an overview of the techniques and tools to understand and explain the uncertainty in their data, and to leverage this uncertainty to provide their results with probabilistic guarantees, i.e., measures to estimate the reliability of these results.

II. Tutorial Outline

The tutorial will be presented jointly by all four authors having complementary backgrounds in handling uncertainty in spatial data. Dr. Trajcevski is an expert in managing uncertainty in moving object databases [62] introduced solutions to spatio-temporal uncertainty using geometric approximations [63]. Dr. Pfoser is an expert in handling uncertainty in spatio-temporal data for trajectory modeling [49] and map-construction [2]. Dr. Züfle has an extensive publication record in querying and mining uncertain data [13, 12, 56] including recent work on Monte-Carlo sampling based approaches to extract representative query results from uncertain data[71, 55]. Dr. Kim is an expert in using geosimulation to model and predict complex spatio-temporal systems [29, 27, 30]. To keep the tutorial vivid an interactive, the four presenters will take turns in presenting sections of the tutorial based on their expertise. We propose a 1.5 hours duration for this advanced seminar. Sections and their presenters are described in the following.

A. Motivation and Application Settings

Living in a world of data-driven science, Dr. Pfoser will kick off the tutorial by introducing examples of uncertainty in modern sources of spatial data, such as Open Street Map[22, 17] and location-based social network data [8]. We briefly introduce spatial and spatio-temporal data and give an overview of existing work that has been done on managing such data, ignoring uncertainty.

B. Part I: Geo-Spatial Uncertainty

Uncertainty Models and Possible World Semantics – In the first main part of the tutorial, Dr. Pfoser will continue to introduce the formal categorization of models for uncertain geo-spatial data: discrete uncertainty models [32] and continuous ones [58, 47], along with attribute [9, 41, 24] and existential [67] uncertainty. The concept of Possible World Semantics, widely used by the data-science community, will be discussed as a mathematically sound and intuitive interpretation of uncertain spatial databases. Additionally, a survey of the Equivalent Worlds Paradigm will be given, to tame the exponential number of possible worlds [33, 1, 54] and #P hard query processing [40]. This paradigm allows to answer a large number of spatial query predicates efficiently. To illustrate this paradigm, Dr. Züfle will take over and survey the technique of using generating functions [41] to efficiently answer range queries and k-NN queries on uncertain spatial data [4].

Representative Querying and Mining of Uncertain Spatial Data using Monte-Carlo Sampling Dr. Züfle will proceed to give a tutorial to efficiently gain approximation of spatial query and mining results using the concept of representative worlds [55, 71]. For this purpose, techniques to draw unbiased samples from the uncertainty models presented in Section II-B will be given, and solutions will be discussed to aggregate query and mining results from sampled worlds into representative results. Solutions to derive probabilistic guarantees such as using Chernoff bounds will be touched upon. This part will also survey computational challenges and approximate solutions for probabilistic networks, having each edge associated with an existential probability [52, 14].

C. Part II: Uncertainty and Evolving Spatial Data

Interpolation, Uncertainty and Evolution: Approaching uncertainty from the perspective of geoinformation-science, Dr. Trajcevski will discuss kriging interpolation methods [45] that explicitly measure uncertainty in the interpolation output. Contrary to previous deterministic methods, Kriging's regression-based methodology includes elements of uncertainty in both the prediction of the final surface as well as in the estimated error surface of the predictions [42, 59, 19]. An implementation with Esri's geostatistical analyst ([36, 26]) will be demonstrated and a multi-variogram approach called Empirical Bayesian Kriging ([20]) will be shown that differs in accounting for the error introduced by estimating the underlying semivariogram and selecting the best fit ([34]). The impact of the temporal dimension and the consequences of using discrete-location samples to represent continuous phenomena will also be discussed [3, 70].

Motion and Sensing: As we will have discussed, the complexity of querying uncertain geo-spatial data is exponential. Considering time, this problem becomes even more complex. Consequently, the main scope of this part of the tutorial is to explore approximate solutions for handing uncertain spatio-temporal data such as trajectory data. Dr. Trajcevski will introduce this part of the tutorial by providing use-cases and examples of uncertain trajectory data in indoor settings and map matching settings. Traditional models to cope with these new challenges will be reviewed (e.g., bounding the possible locations of objects over time by spatio-temporal cylinders [61, 62], diamonds [43] or beads [35, 60]). Based on these models/types, corresponding algorithms for processing certain query categories have been proposed, e.g. range queries [50, 62], kNN queries [61, 7], etc. which will be overviewed. Additionally, some foundational works on location dependency will be reviewed [7, 61, 43, 66] in order to motivate the use of more advanced models. State-of-the-art approaches for querying traffic network data will also be discussed, along with the issues related to spatio-temporal data compression [5, 39, 51], fusion of heterogeneous uncertain location data [68], and the impact of uncertainty on the quality of learning/prediction [69].

D. Part III: Geospatial Simulation

One of the popular techniques to understand and predict uncertain complex phenomena such as social networks, natural disaster, epidemic, etc. – is geospatial modeling and simulation (M&S). Unlike black-box machine learning algorithms, since well-defined geospatial models are an abstractions of the real world, geospatial simulation is understandable and enables to experience diverse what-if scenarios such as spreading of seasonal flu [29]. Plausible scenarios with deliberate policies by domain experts have aided policy makers' decision for uncertain future events such as disaster response (i.e., hurricane and earthquake) [11]. Dr. Kim will introduce a fundamental concept of M&S and exhibit examples of geospatial simulations including epidemic simulations [30], ride-haling [31], collision prediction at intersection [38], sensor network routing, city network generation [28], synchronized world [25], location-based social network [27] and urban patterns of life [29]. The examples will be used to show how repeated simulations can be leverage to estimate the likelihood of events, and how parameters of a system can be adjusted to achieve a desired result such as optimizing ride-hailing algorithms.

III. Presenter Bios

Andreas Züfle received his PhD in Computer Science from LMU, Munich in 2013 under the direction of Dr. Hans-Peter Kriegel. Since January 2016, he is an Assistant Professor at the Department of Geography and Geoinformation Science at George Mason University. His research interests include managing, querying and mining of uncertain geo-spatial data and geosimulation. Since 2016, he has been awarded more than 2M USD in research funding from the National Science Foundation and the Defense Advanced Research Project Agency.

Dieter Pfoser received his PhD in Computer Science from Aalborg University, Denmark in 2000. He is the chair of the Department of Geography and Geoinformation Science, George Mason University, USA. His research focuses on spatiotemporal databases, shortest-path algorithms and data mining methods for user-generated geospatial content such as map matching and map construction. He has co-authored more than 100 papers and his research has been supported by various funding agencies in Europe and the US.

Goce Trajcevski received his PhD from the University of Illinois at Chicago in 2002 and his main research interests are in the areas of Mobile Data Management, Uncertainty Management and Sensor Networks. His research has been funded by the NSF, ONR as well as industry (Nortrhop Grumman and BEA). He has co-authored over 140 publications in refereed conferences and journals, served as a General Co-Chair of ICDE 2014 and ACM SIGSPATIAL 2019, PC Co-Chair of ADBIS 2014 and ACM GIS 2016.

Joon-Seok Kim received his PhD in Computer Science from Pusan National University in 2016. He is currently a post-doctoral research fellow at the Department of Geography and Geoinformation Science, George Mason University. His research interests include spatiotemporal databases, data mining, geospatial simulation, and location privacy. His SmartAgent won the first place at ACM GIS Cup 2019. He was a PC co-chair of the GeoSim in 2018 and 2019, and the chair of AI and Simulation track of SpringSim in 2020.

IV. Acknowledgements

Dr. Pfoser and Dr. Züfle are supported by National Science Foundation AitF grant CCF-1637541. Dr. Trajcevski is supported by National Science Foundation grant CNS 1646107.

References

[1] S. Abiteboul, P. C. Kanellakis, and G. Grahne. On the representation and querying of sets of possible worlds. *Theoretical Computer Science*, 78(1):158–187, 1991.

[2] M. Ahmed, S. Karagiorgou, D. Pfoser, and C. Wenk. A comparison and evaluation of map construction algorithms using vehicle tracking data. *GeoInformatica*, 19(3):601–632, 2015.

[3] B. Avci, G. Trajcevski, and P. Scheuermann. Tracking uncertain shapes with probabilistic bounds in sensor networks. In *ADBIS*, 2016.

[4] T. Bernecker, T. Emrich, H.-P. Kriegel, N. Mamoulis, M. Renz, and A. Züfle. A novel probabilistic pruning approach to speed up similarity queries in uncertain databases. In *Proc. ICDE*, pages 339–350, 2011.

[5] H. Cao, O. Wolfson, and G. Trajcevski. Spatio-temporal data reduction with deterministic error bounds. *VLDB J.*, 15(3):211–228, 2006.

[6] G. Casella and R. L. Berger. *Statistical Inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

[7] R. Cheng, D. Kalashnikov, and S. Prabhakar. Querying imprecise data in moving object environments. In *TKDE*, pages 1112–1127, 2004.

[8] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *SIGKDD*, pages 1082–1090, 2011.

[9] G. Cormode, F. Li, and K. Yi. Semantics of ranking queries for probabilistic data and expected results. In *ICDE*, pages 305–316, 2009.

[10] H. Couclelis. The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge. *Transactions in GIS*, 7(2):165–175, 2003.
[11] A. T. Crooks and S. Wise. Gis and agent-based models for humanitarian assistance. *Computers, Environment & Urban Sys.*, 41:100–111, 2013.
[12] T. Emrich, H.-P. Kriegel, N. Mamoulis, M. Renz, and A. Züfle. Indexing uncertain spatio-temporal data. In *CIKM*, pages 395–404, 2012.
[13] T. Emrich, H.-P. Kriegel, N. Mamoulis, M. Renz, and A. Zufle. Querying uncertain spatio-temporal data. In *2012 IEEE 28th international conference on data engineering*, pages 354–365. IEEE, 2012.
[14] T. Emrich, H.-P. Kriegel, J. Niedermayer, M. Renz, A. Suhartha, and A. Züfle. Exploration of monte-carlo based probabilistic query processing in uncertain graphs. In *CIKM*, pages 2728–2730, 2012.
[15] R. G. Fegeas, J. L. Cascio, and R. A. Lazar. An Overview of FIPS 173, The Spatial Data Transfer Standard. *Cartography and Geographic Information Systems*, 19(5):278–293, 1992.
[16] H. Gao, C. H. Liu, W. Wang, J. R. Zhao, Z. Song, X. Su, J. Crowcroft, and K. K. Leung. A survey of incentive mechanisms for participatory sensing. *IEEE Communications Surveys and Tutorials*, 17(2), 2015.
[17] J.-F. Girres and G. Touya. Quality assessment of the french openstreetmap dataset. *Transactions in GIS*, 14(4):435–459, 2010.
[18] M. F. Goodchild. Uncertainty: The Achilles Heel of GIS. *Geo Info Systems*, 8(11):50–52, 1998.
[19] P. Goovaerts. Geostatistical modelling of uncertainty in soil science. *Geoderma*, 103(1):3–26, 2001.
[20] A. Gribov and K. Krivoruchko. New flexible non-parametric data transformation for trans-gaussian kriging. In *Geostatistics Oslo*. 2012.
[21] J. Grira, Y. Bédard, and S. Roche. Spatial Data Uncertainty in the VGI World: Going from Consumer to Producer. *Geomatica*, 64(1), 2010.
[22] M. Haklay. How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. *Env. and Planning B: Planning and Design*, 37(4):682–703, 2010.
[23] J. Hsu. *Multiple Comparisons: Theory and Methods*. Chapman and Hall/CRC, 1996.
[24] M. Hua, J. Pei, W. Zhang, and X. Lin. Ranking queries on uncertain data: a probabilistic threshold approach. In *SIGMOD*, 2008.
[25] D. Jang, J.-S. Kim, K.-J. Li, and C.-H. Joo. Overlapping and synchronizing two worlds. In *ACM SIGSPATIAL*, pages 493–496, 2011.
[26] K. Johnston, J. M. Ver Hoef, K. Krivoruchko, and N. Lucas. *Using ArcGIS geostatistical analyst*, volume 380. Esri Redlands, 2001.
[27] H. Kavak, J.-S. Kim, A. Crooks, D. Pfoser, C. Wenk, and A. Züfle. Location-based social simulation. In *SSTD*, pages 218–221, 2019.
[28] J.-S. Kim, H. Kavak, and A. Crooks. Procedural city generation beyond game development. *SIGSPATIAL Special*, 10(2):34–41, 2018.
[29] J.-S. Kim, H. Kavak, U. Manzoor, A. Crooks, D. Pfoser, C. Wenk, and A. Züfle. Simulating urban patterns of life: A geo-social data generation framework. In *ACM SIGSPATIAL*, pages 576–579, 2019.
[30] J.-S. Kim, H. Kavak, U. Manzoor, and A. Züfle. Advancing simulation experimentation capabilities with runtime interventions. In *2019 Spring Simulation Conference (SpringSim)*, pages 1–11. IEEE, 2019.
[31] J.-S. Kim, D. Pfoser, and A. Züfle. Distance-aware competitive spatiotemporal searching using spatiotemporal resource matrix factorization (gis cup). In *ACM SIGSPATIAL*, pages 624–627, 2019.
[32] H.-P. Kriegel, P. Kunath, and M. Renz. Probabilistic nearest-neighbor query on uncertain objects. In *Proc. DASFAA*, pages 337–348, 2007.
[33] S. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24(1):1–14, 1959.
[34] K. Krivoruchko. *Spatial statistical data analysis for GIS users*. Esri Press Redlands, 2011.
[35] B. Kuijpers and W. Othman. Trajectory databases: Data models, uncertainty and complete query languages. *Journal of Computer and System Sciences*, 76(7):538–560, 2010.
[36] A. Kumar, S. Maroju, and A. Bhat. Application of arcgis geostatistical analyst for interpolating environmental data from observations. *Environmental Progress*, 26(3):220–225, 2007.
[37] S. Kumar, F. Morstatter, and H. Liu. *Twitter data analytics*. Springer, 2014.
[38] O. Kwon, S.-H. Lee, J.-S. Kim, M.-S. Kim, and K.-J. Li. Collision prediction at intersection in sensor network environment. In *IEEE Intelligent Transportation Systems*, pages 982–987. IEEE, 2006.
[39] R. Lange, F. Dürr, and K. Rothermel. Efficient real-time trajectory tracking. *VLDB J.*, 20(5):671–694, 2011.
[40] J. Li and A. Deshpande. Consensus answers for queries over probabilistic databases. In *Proc. PODS*, pages 259–268, 2009.
[41] J. Li, B. Saha, and A. Deshpande. A unified approach to ranking in probabilistic databases. *PVLDB*, 2(1):502–513, 2009.
[42] C. Lloyd and P. M. Atkinson. Assessing uncertainty in estimates with ordinary and indicator kriging. *Computers & Geosciences*, 27(8):929–937, 2001.
[43] H. Mokhtar and J. Su. Universal trajectory queries for moving object databases. In *MDM*, 2004.
[44] K. Nellore and G. P. Hancke. A survey on urban traffic management system using wireless sensor networks. *Sensors*, 2, 2016.
[45] M. A. Oliver and R. Webster. Kriging: a method of interpolation for geographical information systems. *IJGIS*, 4(3):313–332, 1990.
[46] Open street map. http://www.openstreetmap.org.
[47] K. Patroumpas, M. Papamichalis, and T. K. Sellis. Probabilistic range monitoring of streaming uncertain positions in geosocial networks. In *Proc. SSDBM*, pages 20–37, 2012.
[48] J. Pei, M. Hua, Y. Tao, and X. Lin. Query answering techniques on uncertain and probabilistic data. In *SIGMOD*, 2008.
[49] D. Pfoser and C. S. Jensen. Capturing the uncertainty of moving-object representations. In *SSD*, pages 111–131. Springer, 1999.
[50] D. Pfoser and C. S. Jensen. Capturing the uncertainty of moving-object representations. In *Proc. SSD*, 1999.
[51] I. S. Popa, K. Zeitouni, V. Oria, and A. Kharrat. Spatio-temporal compression of trajectories in road networks. *GeoInformatica*, 19(1):117–145, 2015.
[52] M. Potamias, F. Bonchi, A. Gionis, and G. Kollios. K-nearest neighbors in uncertain graphs. *VLDB*, 3(1-2):997–1008, 2010.
[53] M. Renz, R. Cheng, H.-P. Kriegel, A. Züfle, and T. Bernecker. Similarity search and mining in uncertain databases. *PVLDB*, 3(1-2):1653–1654, 2010.
[54] A. D. Sarma, O. Benjelloun, A. Y. Halevy, and J. Widom. Working models for uncertain data. In *Proc. ICDE*, 2006.
[55] K. A. Schmid and A. Züfle. Representative query answers on uncertain data. In *SSTD*, pages 140–149, 2019.
[56] E. Schubert, A. Koos, T. Emrich, A. Züfle, K. A. Schmid, and A. Zimek. A framework for clustering uncertain data. *VLDB*, 8(12):1976–1979, 2015.
[57] D. Sui, S. Elwood, and M. Goodchild. *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. Springer Science & Business Media, 2012.
[58] Y. Tao, X. Xiao, and R. Cheng. Range search on multidimensional uncertain data. *ACM TODS*, 32(3):15, 2007.
[59] E. Todini and M. Ferraresi. Influence of parameter estimation uncertainty in kriging. *Journal of Hydrology*, 175(1):555–566, 1996.
[60] G. Trajcevski, A. N. Choudhary, O. Wolfson, L. Ye, and G. Li. Uncertain range queries for necklaces. In *MDM*, pages 199–208, 2010.
[61] G. Trajcevski, R. Tamassia, H. Ding, P. Scheuermann, and I. F. Cruz. Continuous probabilistic nearest-neighbor queries for uncertain trajectories. In *Proc. EDBT*, pages 874–885, 2009.
[62] G. Trajcevski, O. Wolfson, K. Hinrichs, and S. Chamberlain. Managing uncertainty in moving objects databases. *ACM TODS*, 29(3), 2004.
[63] G. Trajcevski, O. Wolfson, F. Zhang, and S. Chamberlain. The geometry of uncertainty in moving objects databases. In *EDBT*, pages 233–250. Springer, 2002.
[64] United States Geological Survey. USGS Science Data Catalog. https://data.usgs.gov/datacatalog/.
[65] J. Yang, J. Zhou, Z. Lv, W. Wei, and H. Song. A real-time monitoring system of industry carbon monoxide based on wireless sensor networks. *Sensors*, 11, 2015.
[66] M.-Y. Yeh, K.-L. Wu, P. S. Yu, and M. Chen. PROUD: a probabilistic approach to processing similarity queries over uncertain data streams. In *Proc. EDBT*, pages 684–695, 2009.
[67] M. L. Yiu, N. Mamoulis, X. Dai, Y. Tao, and M. Vaitis. Efficient evaluation of probabilistic advanced spatial queries on existentially uncertain data. *IEEE TKDE*, 21(1):108–122, 2009.
[68] B. Zhang and G. Trajcevski. Towards fusing uncertain location data from heterogeneous sources. *GeoInformatica*, 2016. (to appear).
[69] F. Zhou, Q. Gao, G. Trajcevski, K. Zhang, T. Zhong, and F. Zhang. Trajectory-User Linking via Variational AutoEncoder. . In *IJCAI*, 2018.
[70] F. Zhou, G. Trajcevski, R. Tamassia, B. Avci, A. A. Khokhar, and P. Scheuermann. Bypassing holes in sensor networks: Load-balance vs. latency. *Ad Hoc Networks*, 61:16–32, 2017.
[71] A. Züfle, T. Emrich, K. A. Schmid, N. Mamoulis, A. Zimek, and M. Renz. Representative clustering of uncertain data. In *ACM KDD*, pages 243–252, 2014.
[72] A. Züfle, G. Trajcevski, D. Pfoser, M. Renz, M. T. Rice, T. Leslie, P. Delamater, and T. Emrich. Handling uncertainty in geo-spatial data. In *IEEE ICDE*, pages 1467–1470, 2017.