GI_Forum 2018, Issue.1
Page: 3 - 15
Full Paper
Corresponding Author:
francisomondi.oloo@sbg.ac.at
DOI: 10.1553/giscience2018_01_s3

Multi-Dimensionality of Uncertainty in big Geospatial Sensor Data

Francis Oloo[1], Raphaela Pagany[1,2] and Gyula Kothencz[1]
[1]University of Salzburg, Austria
[2]Deggendorf Institute of Technology, Germany

Abstract

Developments in sensor technology have contributed immensely to the growth of big geospatial sensor data. Moreover, advances in telecommunications have made it possible to use in-situ sensors to capture and transfer data about the environment in near real-time. The combination of various sensor types within a predefined geographic space and the possibility of making measurements at high temporal resolution contributes to a better understanding of our environment while also generating big geospatial sensor data. Big data from multi-sensor networks, particularly those that capture dynamic characteristics of the environment, have not been spared the challenges that face other types of big data. Specifically, the quality of data and the propagation of uncertainty through the multi-sensor data processing workflows have remained a major concern in the big geospatial sensor data research community. Attempts to document, quantify and communicate the uncertainty associated with sensor data and related sensor network outputs have been made mainly in the context of individual projects. This paper aims to document the state-of-art in defining uncertainty with regard to multi-sensor geospatial data. In particular, we analyse the current literature to outline different types of uncertainty, and document methods for handling uncertainty in the different stages of multi-sensor geospatial data collection, processing and delivery.

Keywords:

big data, data quality, sensors, uncertainty

1 Background

Advances in sensor technology and in telecommunications have made sensors fundamental sources of geospatial data. It is now possible to deploy sensors in extensive, remote and previously inaccessible geographic terrains (Li, Andrew, Foh, Zukerman, & Chen, 2009). Furthermore, with the advent of sensor webs, standardized tools have been developed to facilitate sensor data sharing, discovery, access and visualization (Botts, Percivall, Reed, & Davidson, 2007) over the internet. Consequently, it is now possible to discover and use data from an immense range of sensors from different manufacturers, designed for different types of use, deployed and operated by different entities. The various types of sensors deployed in geographic space and equipped with different communication capabilities can be harnessed into geospatial sensor networks. Geospatial sensor networks enable the capturing of a wide range of environmental variables and include in-situ pre-processing. Additionally, movement trajectories, and health and physiological data of humans and animals, including heart rate, body temperature and muscle tension, can also be recorded. As a result, sensor data have become an integral part of complex environmental monitoring systems. The extraordinary volume of environmental data, captured at high velocities, consequently leads to big data (Zaslavsky, Perera, & Georgakopoulos, 2013).

The heterogeneous characteristics of the hardware, software and computational capabilities of the sensors that compose geospatial sensor networks, coupled with the diversity of the environment in which the sensors are deployed, introduce data uncertainties (Reis et al., 2015). Additionally, individual research objectives determine the choice of sensors to be deployed and the data quality requirements (Dasarathy, 1997). These objectives and requirements further influence the quality of data from sensor nodes that feed geospatial sensor networks. Uncertainty is defined as the quantified description of doubt in a measurement (Bleier et al., 2009). Epistemologically, uncertainty exists in data because any measurement is a simplification of reality. The uncertainty associated with sensor data, particularly data from multi-sensor networks in extensive geographic scales, has been a major concern of many applications (Feyen & Caers, 2006; Geza, Poeter, & McCray, 2009). This is particularly the case when sensor data are used to inform policy decisions (Maxim & van der Sluijs, 2011).

The quality of geospatial data has been an ongoing pursuit in the geographic information science community (Duckham, Mason, Stell, & Worboys, 2001; Frank, 1998; Goodchild & Jeansoulin, 1998). Miniaturization and advances in computational capabilities have contributed to the ubiquity of sensors (Nittel, Labrinidis, & Stefanidis, 2008), resulting in the growth in number of possible sensor data sources and in the multi-dimensionality of the uncertainty of sensor data. Consequently, efforts to define, communicate and handle uncertainty in sensor-driven environment-monitoring systems have become fundamental (Matott, Babendreier, & Purucker, 2009; Refsgaard, Van der Sluijs, Brown, & Van der Keur, 2006; Refsgaard, van der Sluijs, Højberg, & Vanrolleghem, 2007).

Practical solutions for handling uncertainty in geospatial sensor data have been implemented in the context of specific projects or to serve specific application domains. Consequently, there is no up-to-date and widely accepted definition of uncertainty, or description of methods for documenting and handling uncertainty in the context of big data obtained from multi-sensor networks. This paper aims to document the state-of-art in efforts to define uncertainty and outlines the methods for handling uncertainty emanating from big geospatial multi-sensor data. In this endeavour, we aim to answer the following guiding questions: (a) What are the dimensions of uncertainty in big geospatial data emerging from multi-sensor sources? (b) What methods are available for defining, documenting and handling uncertainty at different stages of multi-sensor workflows? Based on these questions, the objectives of this study are: (i) to define the various dimensions of uncertainty in big geospatial multi-sensor data; (ii) to synthesize the methods of handling uncertainty at different levels of multi-sensor environmental monitoring systems.

The remainder of this paper is organized as follows. In Section 2, we provide a definition of uncertainty in the context of big multi-sensor data. In Section 3, we document and synthesize the main methods that are available in the relevant literature. In Section 4, we provide a conclusion regarding the critical elements of uncertainty in the age of big data from geospatial multi-sensor data.

2 Definition of uncertainty

Geospatial sensor data observations and big data analytics

Traditionally, geographic knowledge emerged from the desktop-based spatial analysis and geovisualization of conventional vector and raster data models (Kitchin, 2014; Miller & Goodchild, 2015). However, the proliferation of location-aware sensors has contributed to the growth of geospatial big data. Consequently, advanced methods of analysis are required to address challenges associated with big data (Goodchild, 2009; Kitchin, 2013, 2014; Sui & DeLyser, 2012). Typically, sensor-driven methods for environment monitoring encompass sensing, data processing and visualization (Figure 1). In geospatial sensor data processing systems, space describes the sensed locations on the surface of the earth and provides a reference upon which different dimensions of big data can be aggregated, analysed and interpreted. The integration of space and time during the collection of sensor data facilitates big geospatial sensor data analysis (Lee & Kang, 2015) and opens up further horizons in the understanding of the physical and social environment.

Figure 1: Typical sensing, processing and visualization paradigm of sensor-driven monitoring systems

Uncertainty in sensor-driven environmental management systems

Addressing uncertainty associated with multi-sensor networks requires the understanding of critical elements of multi-sensor data management systems. There are three basic layers in a sensor-web-enabled workflow, namely the observation, management and discovery layers (Figure 2). The observation layer consists of the different sensor nodes in a network. Each sensor node is wirelessly connected to a gateway, hence streamlining the communication between a sensor and the data management system. This design also allows for an in-place computation, which improves the efficiency of sensor resources by ensuring that only pre-processed data are transferred from a sensor node to the centralized data management system.

Due to the technical characteristics of sensors and the variation in the characteristics of the geographic locations at which sensors are deployed, uncertainties may arise from the observation layer. These uncertainties should be documented, qualified and communicated to the centralized data management system. Information on data quality from the observation layer should be transferred along with the observed data.

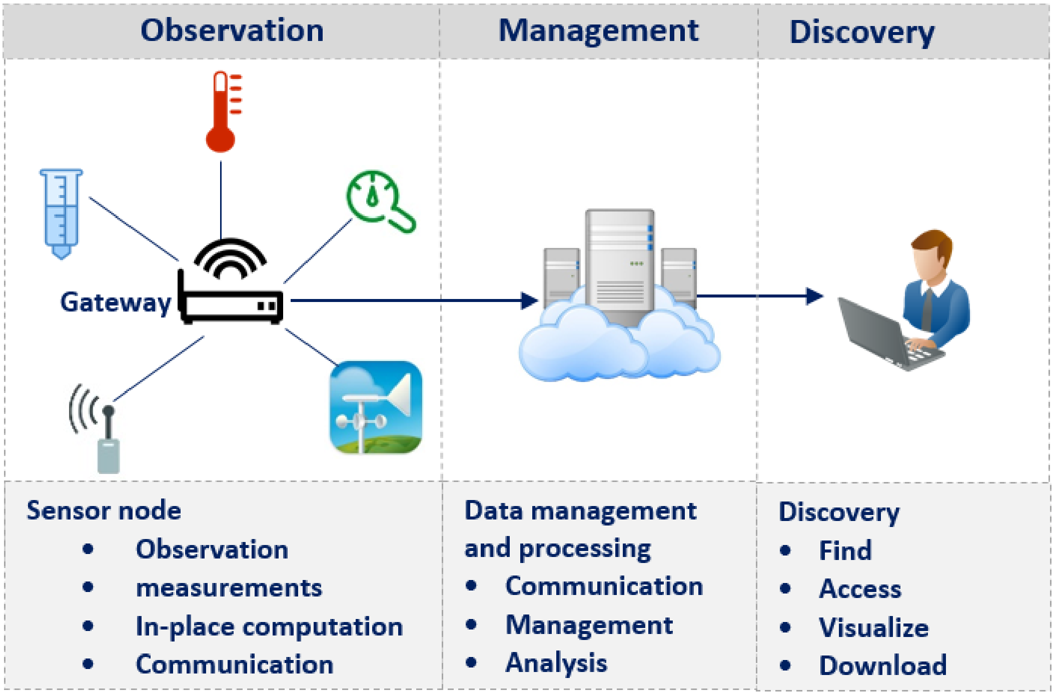

Figure 2: Critical layers of a sensor-driven environmental management system

The second layer is the management (or processing) layer, which provides cloud-based data storage and is accessible via the internet. The management layer also handles communication between the sensor nodes and the end-users. In line with the aims of the Open Geospatial Consortium (OGC), sensor resources are managed as discoverable and consumable resources. The data management layer therefore provides open standards for data processing via Web Processing Services (WPS). Uncertainties associated with the management layer must be documented and communicated as part of the metadata.

The third layer entails discovery and visualization. This is where remote users can discover, access, visualize and download resources from the internet. The process of discovery is actualized through standardized communication protocols. The results of the discovery process can only support policy decisions if the end-users understand both the data and the uncertainties associated with the entire chain of data collection, management, processing and visualization.

There are two categories or types of uncertainty that are relevant for sensor data. The first is the unintentional uncertainty that originates from random effects in the measuring system. Unintentional uncertainty can be evaluated by statistical methods. The second category is the intentional type, which originates from malicious physical or digital attacks on a sensor system (Ni et al., 2009; Shi & Perrig, 2004). In this work, we focus on the unintentional errors that if not properly quantified and documented can propagate through the sensor network and impede the quality of data and information obtained from sensor data management systems.

Research on data quality has been an ongoing endeavour within geographic information science research (Devillers, Bédard, Jeansoulin, & Moulin, 2007; Duckham et al., 2001; Veregin, 1999). Techniques for assessing and improving data quality have also been documented (Batini, Cappiello, Francalanci, & Maurino, 2009). Rodriguez and Servigne (2012) and Rodríguez and Servigne (2013) classified the data quality dimensions in each of the layers as follows: observation layer (accuracy, reliability, spatial precision, completeness and communication reliability); management layer (consistency, currency and volatility); discovery layer (timeliness, availability and adequacy).

Description of uncertainty in the sensor web

The OGC's Sensor Web Enablement (SWE) framework specifies standardized service-oriented interfaces for describing sensor resources (Botts, Percivall, Reed, & Davidson, 2008; Bröring et al., 2011). Noteworthy standards for documenting data quality in the sensor web include: (a) Sensor Model Language (SensorML), (b) Observation & Measurement (O&M), (c) Sensor Observation Service (SOS), (d) Sensor Planning Service (SPS), and (e) SWE Common Data and Services.

SensorML defines models and XML schema for describing processes associated with the measurement and post-measurement transformation of sensor observations (Botts & Robin, 2007; Botts, Robin, Greenwood, & Wesloh, 2014). Specific aims of SensorML that are relevant for documenting sensor data quality include: (i) to provide performance and quality of measurement characteristics (e.g., accuracy, threshold); (ii) to provide an explicit description of the process by which an observation was obtained (i.e., its lineage); (iii) to archive fundamental properties and assumptions regarding sensor systems and computational processes. These aims intended not only to document the processes involved in sensor data measurements, but also to describe and document the quality characteristics associated with the sensor data.

Observation & Measurement (O&M) provides a standard framework for representing observations, measurements, procedures and metadata in a sensor system (Reed, Botts, & Davidson, 2007). Within the O&M specification, an observation is an event aimed at measuring or determining the value of a property by using known procedures. The O&M document can capture information about data quality with respect to the measurement process. An uncertainty-enabled observation and measurement profile has been developed as an extension of the original O&M document (Stasch et al., 2012).

Additionally, in order to bridge the gap between sensor observations and environmental modelling, Uncertainty Markup Language (UncertML) was developed to enable uncertainty-handling in sensor-driven modelling environments (Williams, Cornford, Bastin, & Pebesma, 2009). Consequently, a web-based uncertainty-handling framework that is built around UncertML documentation is now possible via UncertWeb (Bastin et al., 2013). More specifically, the QualityML was defined in order to provide a more holistic description of sensor data quality (Ninyerola et al., 2014).

3 General framework for handling multiple dimensions of uncertainty in multi-sensor networks

Overall dimensions of handling uncertainty

In multi-sensor networks, there are multiple phases, including collection (in the observation layer), processing (management and processing layer), and discovery. In each phase, there may be multiple actors (users and stakeholders) or nodes (sensor nodes). The interaction between the many phases, actors (with different needs and inputs) and different processes involved in the multi-sensor framework contribute to multiple dimensions of uncertainty in multi-sensor data systems. Consequently, various strategies for identifying, managing and eliminating uncertainty in sensor data during acquisition, processing and utilization have been proposed. Specifically, Walker et al. (2003) define dimensions of uncertainty under three categories: location (context, input, parameter, model), levels (statistical, scenario, ignorance), and nature (epistemic and variability). Further taxonomies in the definition and classification of uncertainty are provided by Funtowicz & Ravetz (1990), Klauer & Brown (2004), and Sigel, Klauer, & Pahl-Wostl (2010). Here, we synthesize the dimensions into different categories and further link the dimensions to a specific sensor data management layer to facilitate uncertainty-handling in sensor data from multiple sources.

Methods of handling sensor-based uncertainty are dependent primarily on the needs and preferences of users in a particular application domain. The purpose of the categorization in this study is to provide an efficient guideline that can be adopted by users from multiple disciplines that use sensor data in their analyses. There are three general steps to consider when handling uncertainty of geospatial sensor data: (1) identification, (2) documentation, and (3) clarification.

In the identification phase, a classification scheme or an uncertainty matrix is used to highlight uncertainty associated with specific sensor data and processes (Warmink, Janssen, Booij, & Krol, 2010). The classification scheme provides a systematic means of representing and quantifying the specific aspects of the uncertainty at hand (Refsgaard et al., 2007; Walker et al., 2003). Identification enables a user to frame the uncertainty concisely and guides subsequent phases of handling the uncertainty.

In the description phase, metadata describing the uncertainty and the associated data is created. The description of uncertainty may be supplemented by a description of the dimensions of data quality, including accuracy, completeness, consistency, timeliness, interpretability and accessibility (Batini et al., 2009; Scannapieco & Catarci, 2002). Metadata enables subsequent data-users to recognize and consider the deficits in the dataset. A practical solution to describe uncertain information within metadata is provided by UncertML (Williams, Cornford, Bastin, & Ingram, 2008). When documenting uncertainty information in metadata, a bottom-up or a top-down approach can be adopted (Devillers et al., 2007).

In the bottom-up approach, information on uncertainty from data producers, sensor calibration parameters and systematic deviations are noted in the metadata and parsed to the upper levels in the sensor data management and analysis workflow. At the bottom level, information on uncertainty can be detailed and linked to individual records and processes. However, in higher levels of the data-handling workflow, aggregated measures can be used to describe uncertainty. This introduces a level of complexity in understanding uncertainty, particularly if the data originates from an interconnected network of multiple sensor nodes.

In a top-down approach, quality information is not explicitly available for individual elements. Instead, globally aggregated statements about processes or data are used to communicate uncertainty. Furthermore, in the top-down approach, expert analysis and judgements on the impact of uncertainty on the output from a sensor-monitoring system are made. The subjective nature of human perception introduces the risk of misinterpretation or wrong estimation by the expert, while on the other hand it can allow the identification of errors or inconsistencies which are unforeseeable in sensitivity or statistical analysis.

The final step in data handling is the clarification phase. Here, uncertain records are flagged, which implies adding a quality tag to the potentially erroneous data with the intention of correcting for the influence of uncertainty in the data. This constitutes the highest level of uncertainty-handling. Data and processing quality is not only assessed but also improved by alterations in the process chain or in the data records.

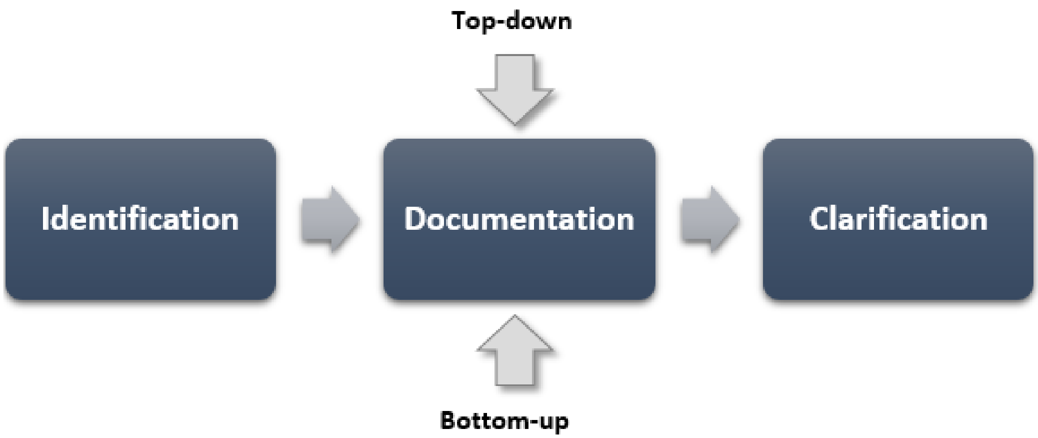

Figure 3: Phases of handling uncertainty in geospatial sensor environmental management systems

Further, strategies for handling uncertainty aim at improving the quality of the data and can be separated into data-driven and process-driven methods (Batini, Cappiello, Francalanci, & Maurino, 2009). Data-driven techniques focus on improving the quality of data by modifying the data value in order to increase their trustworthiness. On the other hand, process-driven methods involve redesigning those components of a sensor data management system that are identified as contributing to the uncertainty in the data. Consequently, in process-driven methods it may be necessary to collect new data after redesigning the specific components of a sensor management system. As a result, process-driven methods tend to be relatively expensive.

Synthesized categories of methods for handling uncertainty in multi-sensor frameworks

Strategies for handling uncertainty (Bastin et al., 2013; Batini et al., 2009; Sigel et al., 2008) can be grouped to specific layers of a sensor network framework. Strategies which can be applied to the observation layer include:

(1) **Merging:** New data, collected under the same acquisition design, can be merged and aggregated to remove outliers or the influence of random errors in a few observations.

(2) **Replacement:** Entails a fresh data collection under the same acquisition design in order to replace uncertain data from an earlier acquisition step. An advantage of this data-driven process is that it is cost-efficient because only the measuring step is repeated. For instance, uncertainty caused by unusual weather phenomena can be compensated by replacement.

(3) **Process-redesign:** Entails collecting new data under a redesigned data-acquisition framework. Arithmetic overflows, unreliability or wrong datasets (incorrect sample) emerging due to poor research design can be accounted for.

(4) **Metadata description:** Generating a complete metadata set that includes calibration parameters of sensors used in observation. This ensures that subsequent data-handling processes benefit from the technical specification of the measuring equipment and the associated influence on precision as part of the descriptors of uncertainty. Additionally, uncertainty flags can also be included in the metadata description (Devaraju, Jirka, Kunkel, & Sorg, 2015).

In the management layer, the following methods can be used:

(1) **Interpolation:** Interpolation methods can be used to integrate data to a unified reference framework. For instance, data from heterogeneous sources or those with different granularity can be synchronized at a uniform temporal resolution. Gaps and errors in the data can be calculated using an interpolation approach and documented in an UncertML document. An example was tested in the INTAMAP project (Williams, Cornford, Bastin, Jones, & Parker, 2011)

(2) **Statistical methods:** Statistical analysis of the data can be executed to flag records that do not fall within expected statistical ranges. The error localization can be corrected by identifying values which do not satisfy certain rules (Batini et al., 2009).

(3) **Probability calculations** are a common approach for describing or quantifying the uncertainty in terms of probability functions (Sigel et al., 2010).

(4) **Sensitivity analysis** methods such as the Monte Carlo method are often used for simulating the impacts of uncertain data within the sensor-data process model. Bastin et al. (2013) give a selection of software packages for handling uncertainty in this manner.

(5) **Ontology-driven models:** These methods use descriptive ontology web languages (OWL) to create standardized semantics for communication between different nodes and users of the geospatial sensor resources (Yi & Calmet, 2005).

(6) **General likelihood uncertainty estimation (GLUE):** These methods identify the most likely source of uncertainty in data by comparing multiple models, including fuzzy set theory and Bayesian models (Beven & Freer, 2001).

Methods which are related to the discovery layer are:

(1) **Scientific judgement:** Experts can use their knowledge to detect uncertainty in the last step of the workflow. While the uncertainty may be identified in this manner, another re-examination of the process workflow would be necessary to improve the data and outcomes of the study. This may include a fresh collection of data under the same research design, or a redesign of the data acquisition process.

(2) **Geovisualization** approaches exist for illustrating the uncertain results graphically or in maps by using transparency, capacity or colour bleaching (Li, Kraak, & Ma, 2007).

A summary of the methods for handling uncertainty in different sensor network layers is provided in Table 1.

Table 1: Categories of methods for handling uncertainty by phase, paradigm of implementation, approach and applicable geospatial sensor network layer.

Method	Phase	Paradigm	Approach	Sensor network layer
Merging	Clarification	Bottom-up	Data-driven	Observation
Replacement	Clarification	Bottom-up	Data-driven	Observation
Process redesign	Clarification	Bottom-up	Process-driven	Observation
Metadata	Description	Bottom-up	Process-driven	Observation
Interpolation	Clarification	Bottom-up	Data-driven	Management
Statistical analysis	Clarification	Bottom-up	Data-driven	Management
Probability	Description	Top-down	Data-driven	Management

Ontology-driven models	Description	Bottom-up	Process-driven	Management
GLUE	Identification	Top-down	Data-driven	Management
Sensitivity analysis	Identification	Bottom-up	Process-driven	Management
Scientific judgement	Identification	Top-down	Process-driven	Discovery
Visualization	Description	Top-down	Process-driven	Discovery

4 Conclusion

In this study, we have provided a concise documentation of methods for handling uncertainty in different dimensions of big geospatial data from multi-sensor data frameworks. The development of a Sensor Web Enablement (SWE) framework and the formulation of OGC standards have contributed immensely to streamlining the quality of sensor data by providing standard schemas for defining and documenting uncertainty in the SWE. Of particular relevance to the research on uncertainty are the Observation & Measurement schemas that standardize the documentation of observations, and SensorML, UncertML and QualityML. Furthermore, UncertWeb now makes it possible not only to document the uncertainties at different points in the modelling framework but also to trace the propagation of the uncertainty through the modelling workflow.

In spite of the efforts to build standardized methods for documenting uncertainty, the developments and implementations of the methods are still domain-specific. This eventuates in the risk of duplicating effort and results in redundancies. In this paper, we have provided a generalized outline that can be adopted by researchers from across thematic domains. Furthermore, we have described common dimensions of uncertainty in geospatial sensor-data management, which may support researchers in identifying the crucial aspects to consider when dealing with big data obtained from multi-sensor sources.

One limitation of this work is that we did not include a practical case to demonstrate the methods and dimensions outlined. However, adequate examples can be found in the literature (Bastin et al., 2013; Devaraju et al., 2015; Erickson, Cline, Tirpankar, & Henderson, 2015). From a usability perspective, the methods described in the literature and the case studies require users to have good computational skills. Future implementations should consider the development of simple user interfaces. In addition, while attempts have been made to document error propagation in sensor-driven environmental modelling systems, further work needs to be done to support a feedback mechanism whereby model results can be used to improve the quality of data (Yue, Zhang, & Tan, 2015).

References

Bastin, L., Cornford, D., Jones, R., Heuvelink, G. B. M., Pebesma, E., Stasch, C., Nativi, S., Mazzetti, P., & Williams, M. (2013). Managing uncertainty in integrated environmental modelling: The UncertWeb framework. Environmental Modelling & Software, 39, 116-134. doi:https://doi.org/10.1016/j.envsoft.2012.02.008

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for data quality assessment and improvement. ACM computing surveys (CSUR), 41(3), 16.

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of hydrology, 249(1), 11-29. doi:https://doi.org/10.1016/S0022-1694(01)00421-8

Bleier, T., Bozic, B., Bumerl-Lexa, R., da Costa, A., Costes, S., Iosifescu, I., Martin, O., Frysinger, S., Havlik, D., & Hilbring, D. (2009). SANY: An open service architecture for sensor networks: SANY consortium.

Botts, M., Percivall, G., Reed, C., & Davidson, J. (2007). OGC White Paper-OGC Sensor Web Enablement: Overview and High Level Architecture (OGC Document Number: 07-165). Wayland, MA: OG.

Botts, M., Percivall, G., Reed, C., & Davidson, J. (2008). OGC® sensor web enablement: Overview and high level architecture. GeoSensor networks, 175-190.

Botts, M., & Robin, A. (2007). OpenGIS sensor model language (SensorML) implementation specification. OpenGIS Implementation Specification OGC, 7(000).

Botts, M., Robin, A., Greenwood, J., & Wesloh, D. (2014). OGC® SensorML: Model and XML Encoding Standard. Technical Standard, 2.

Bröring, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., Liang, S., & Lemmens, R. (2011). New generation sensor web enablement. Sensors, 11(3), 2652-2699.

Dasarathy, B. V. (1997). Sensor fusion potential exploitation-innovative architectures and illustrative applications. Proceedings of the IEEE, 85(1), 24-38. doi:10.1109/5.554206

Devaraju, A., Jirka, S., Kunkel, R., & Sorg, J. (2015). Q-SOS—A sensor observation service for accessing quality descriptions of environmental data. ISPRS International Journal of Geo-Information, 4(3), 1346-1365.

Devillers, R., Bédard, Y., Jeansoulin, R., & Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. International Journal of Geographical Information Science, 21(3), 261-282. doi:10.1080/13658810600911879

Duckham, M., Mason, K., Stell, J., & Worboys, M. (2001). A formal approach to imperfection in geographic information. Computers, Environment and Urban Systems, 25(1), 89-103. doi:https://doi.org/10.1016/S0198-9715(00)00040-5

Erickson, P., Cline, M., Tirpankar, N., & Henderson, T. (2015). Gaussian processes for multi-sensor environmental monitoring. Paper presented at the Multisensor Fusion and Integration for Intelligent Systems (MFI), 2015 IEEE International Conference on.

Feyen, L., & Caers, J. (2006). Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations. Advances in Water Resources, 29(6), 912-929.

Frank, A. U. (1998). Metamodels for data quality description. Data Quality in Geographic Information-From Error to Uncertainty, 192.

Funtowicz, S. O., & Ravetz, J. R. (1990). Uncertainty and quality in science for policy (Vol. 15): Springer Science & Business Media.

Geza, M., Poeter, E. P., & McCray, J. E. (2009). Quantifying predictive uncertainty for a mountain-watershed model. Journal of hydrology, 376(1), 170-181.

Goodchild, M. (2009). NeoGeography and the nature of geographic expertise. Journal of Location Based Services, 3(2), 82-96. doi:10.1080/17489720902950374

Goodchild, M., & Jeansoulin, R. (1998). Data quality in geographic information: from error to uncertainty: Hermes Paris.
Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. Dialogues in human geography, 3(3), 262-267.
Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. Big Data & Society, 1(1), 2053951714528481.
Klauer, B., & Brown, J. (2004). Conceptualising imperfect knowledge in public decision-making: ignorance, uncertainty, error and risk situations. Environmental Research, Engineering and Management, 1.
Lee, J.-G., & Kang, M. (2015). Geospatial big data: challenges and opportunities. Big Data Research, 2(2), 74-81.
Li, J., Andrew, L. L., Foh, C. H., Zukerman, M., & Chen, H.-H. (2009). Connectivity, coverage and placement in wireless sensor networks. Sensors, 9(10), 7664-7693.
Li, X., Kraak, M.-J., & Ma, Z. (2007). Towards visual representations to express uncertainty in temporal geodata. Analysis, 6(9), 3.
Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: a review of concepts and tools. Water Resources Research, 45(6).
Maxim, L., & van der Sluijs, J. P. (2011). Quality in environmental science for policy: Assessing uncertainty as a component of policy analysis. Environmental Science & Policy, 14(4), 482-492.
Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. GeoJournal, 80(4), 449-461.
Ni, K., Ramanathan, N., Chehade, M. N. H., Balzano, L., Nair, S., Zahedi, S., Kohler, E., Pottie, G., Hansen, M., & Srivastava, M. (2009). Sensor network data fault types. ACM Trans. Sen. Netw., 5(3), 1-29. doi:10.1145/1525856.1525863
Ninyerola, M., Sevillano, E., Serral, I., Pons, X., Zabala, A., Bastin, L., & Masó, J. (2014). QualityML: A dictionary for quality metadata encoding. Paper presented at the EGU General Assembly Conference Abstracts.
Nittel, S., Labrinidis, A., & Stefanidis, A. (2008). Introduction to advances in geosensor networks. GeoSensor networks, 1-6.
Reed, C., Botts, M., & Davidson, J. (2007, 17-20 Sept. 2007). Ogc® sensor web enablement:overview and high level achhitecture. Paper presented at the 2007 IEEE Autotestcon.
Refsgaard, J. C., Van der Sluijs, J. P., Brown, J., & Van der Keur, P. (2006). A framework for dealing with uncertainty due to model structure error. Advances in Water Resources, 29(11), 1586-1597.
Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process–a framework and guidance. Environmental Modelling & Software, 22(11), 1543-1556.
Reis, S., Seto, E., Northcross, A., Quinn, N. W., Convertino, M., Jones, R. L., Maier, H. R., Schlink, U., Steinle, S., & Vieno, M. (2015). Integrating modelling and smart sensors for environmental and human health. Environmental Modelling & Software, 74, 238-246.
Rodriguez, C. C. G., & Servigne, S. (2012). Sensor Data Quality for Geospatial Monitoring Applications. Paper presented at the AGILE 2012, 15th Internationale Conference on Geographic Information Science.
Rodriguez, C. C. G., & Servigne, S. (2013). Managing Sensor Data Uncertainty: a data quality approach. International Journal of Agricultural and Environmental Information Systems (IJAEIS), 4(1), 35-54.
Scannapieco, M., & Catarci, T. (2002). Data quality under a computer science perspective. Archivi & Computer, 2, 1-15.
Shi, E., & Perrig, A. (2004). Designing secure sensor networks. IEEE Wireless Communications, 11(6), 38-43. doi:10.1109/MWC.2004.1368895
Sigel, K., Klauer, B., & Pahl-Wostl, C. (2010). Conceptualising uncertainty in environmental decision-making: the example of the EU water framework directive. ecological Economics, 69(3), 502-510.

Stasch, C., Jones, R., Cornford, D., Kiesow, M., Williams, M., & Pebesma, E. (2012). Representing uncertainties in the sensor web. Paper presented at the Workshop sensing a changing world.

Sui, D., & DeLyser, D. (2012). Crossing the qualitative-quantitative chasm I: Hybrid geographies, the spatial turn, and volunteered geographic information (VGI). Progress in Human Geography, 36(1), 111-124.

Veregin, H. (1999). Data quality parameters Geographical information systems (pp. 177-189).

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. Integrated assessment, 4(1), 5-17.

Warmink, J. J., Janssen, J. A. E. B., Booij, M. J., & Krol, M. S. (2010). Identification and classification of uncertainties in the application of environmental models. Environmental Modelling & Software, 25(12), 1518-1527. doi:https://doi.org/10.1016/j.envsoft.2010.04.011

Williams, M., Cornford, D., Bastin, L., & Ingram, B. (2008). UncertML: an XML schema for exchanging uncertainty. Proceedings of GISRUK, Manchester, UK, 44, 0-3.

Williams, M., Cornford, D., Bastin, L., Jones, R., & Parker, S. (2011). Automatic processing, quality assurance and serving of real-time weather data. Computers & Geosciences, 37(3), 353-362. doi:https://doi.org/10.1016/j.cageo.2010.05.010

Williams, M., Cornford, D., Bastin, L., & Pebesma, E. (2009). Uncertainty markup language (UncertML).

Yi, Y., & Calmet, J. (2005, 28-30 Nov. 2005). OntoBayes: An Ontology-Driven Uncertainty Model. Paper presented at the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06).

Yue, P., Zhang, M., & Tan, Z. (2015). A geoprocessing workflow system for environmental monitoring and integrated modelling. Environmental Modelling & Software, 69, 128-140. doi:https://doi.org/10.1016/j.envsoft.2015.03.017

Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2013). Sensing as a service and big data. arXiv preprint arXiv:1301.0159.