Spatio-Temporal Aggregation of European Air Quality Observations in the Sensor Web

Christoph Stasch[a,*], Theodor Foerster[a], Christian Autermann[a], Edzer Pebesma[a,b]

[a] *Institute for Geoinformatics, University of Muenster, Weseler Strasse 253, 48153 Muenster, Germany*
[b] *52 North Initiative for Geospatial Open Source Software GmbH, Martin-Luther-King-Weg 24, 48151 Muenster, Germany*

Abstract

An increasing amount of observations from different applications such as long-term environmental monitoring or disaster management is published in the Web using Sensor Web technologies. The standardization of these technologies eases the integration of heterogeneous observations into several applications. However, as observations differ in spatio-temporal coverage and resolution, aggregation of observations in space and time is needed. We present an approach for spatio-temporal aggregation in the Sensor Web using the Geoprocessing Web. In particular, we define a tailored observation model for different aggregation levels, a process model for aggregation processes and a Spatio-Temporal Aggregation Service. The presented approach

*Corresponding author. Current affiliation: Institute for Geoinformatics, University of Muenster, Weseler Strasse 253, 48151 Muenster, Germany. Tel.: +49 251 8331965. Fax.: +49 251 8339763

Email addresses: `staschc@uni-muenster.de` (Christoph Stasch), `foerster@uni-muenster.de` (Theodor Foerster), `autermann@uni-muenster.de` (Christian Autermann), `e.pebesma@uni-muenster.de` (Edzer Pebesma), `pebesma@52north.org` (Edzer Pebesma)

is demonstrated by a case study of delivering aggregated air quality observations on-demand in the Sensor Web.
Keywords: Geoprocessing Web, Sensor Web, Spatiotemporal Aggregation, Air Quality Monitoring

1. Introduction

An increasing amount of observations gathered by geosensor networks is published via standardized Sensor Web technologies to enable an ad-hoc integration of heterogeneous observations in different applications (Broering et al., 2011). As observations usually differ in their spatio-temporal coverage and resolution, aggregation of observations in space and time is needed. Moreover, due to the heterogeneity of these observations, aggregating them is also not trivial. The aggregations need to be performed by dedicated geoprocessing facilities. The Geoprocessing Web with its aim to provide common analysis and transformation of geospatial data into geospatial information is promising to realize spatio-temporal aggregation in the Sensor Web. Currently, data coming from the Sensor Web and Geoprocessing facilities are tightly coupled and only realized for specific scenarios. Though aggregated observations are already available on the Web through for instance weather portals (WetterOnline, 2011) or public observation portals (EEA, 2011), these observations are only aggregated in space *or* in time. An integrative approach for spatio-temporal aggregation is missing. Moreover, these aggregates cannot be calculated on-demand nor are they accessible on the web in standardized formats. In addition, metadata about provenance or aggregation methods is currently not available.

A comprehensive approach for spatio-temporal aggregation in the Sensor Web allowing a flexible integration of observations at a required aggregation level needs to be investigated. The approach has to be flexible to enable easy reuse, integration, and composition of existing aggregation methods. Also, it needs to allow for an on-demand aggregation. To allow retracing aggregated observations to original observations, the approach needs to provide machine readable metadata about the original observations and the aggregation processes. The main contributions of the paper regarding these requirements are:

1. A data model for observations that can be used across different aggregation levels. This model also incorporates metadata about provenance and aggregation method (Section 3).
2. A process model for spatio-temporal aggregation (Section 4).
3. A web service architecture for aggregation of observations including the definition of the Spatio-Temporal Aggregation Service (STAS) (Section 5).

In our approach, we propose a Service-Oriented Architecture (SOA) for spatio-temporal aggregation of observations. As the Open Geospatial Consortium (OGC) provides well-defined encodings and service interfaces for both, the Sensor Web and the Geoprocessing Web, we are utilizing these standards in our approach. As a basis for our SOA, we define a tailored observation model and process model for spatio-temporal aggregation. The proposed SOA consists of Sensor Observation Services (SOS), the standard service for providing observations in the Sensor Web (Na and Priest, 2007), and the Spatio-Temporal Aggregation Service (STAS), which is defined as a profile of the Web Processing Service (WPS). The WPS provides the basic service interface for the Geoprocessing Web (Schut, 2007). In a case study, we demonstrate how our approach meets the requirements identified above by *temporally* aggregating hourly measurements of Ozone to daily maxima and by *spatially* averaging these maxima for each federal state in Germany.

The remainder of the paper is structured as follows: At first, we provide a brief overview on related work and background information (Section 2). We then describe the tailored observation model that can be used across different aggregation levels (Section 3). Afterwards, we present the process model for spatio-temporal aggregation (Section 4). In the next section, we describe how we provide these processes in the Sensor Web (Section 5). The implementation of the approach for an aggregation of air quality observations is presented afterwards (Section 6). Finally, we discuss our results and identify further research (Section 7).

2. Background

This section provides the related work. At first, we give an overview about spatio-temporal aggregation, which forms the framework for this work. Afterwards we provide background information about Geospatial Web Services including Sensor Web technology, the Geoprocessing Web and the Model Web. Geospatial Web Services have been identified as a foundation of this work to enable interoperability of spatio-temporal aggregation on the Web. They provide common means to build interoperable geospatial applications in the Web (Zhao and Di, 2010).

An *aggregation process* computes a single value, an *aggregate*, for a group of attribute values by means of an *aggregation function* (Jeong et al., 2004). The attribute values are grouped by a *partitioning predicate*. In our work, spatio-temporal aggregation combines objects in space and time and provides means to compute aggregates for certain attribute values attached to these objects. Most of the research on spatio-temporal aggregation during the last years has focused on improving aggregation operations in spatio-temporal databases. For example, Vega Lopez et al. (2005) give a comprehensive survey on spatio-temporal aggregation methods in databases. Others develop general models for space and time that can be used as a basis for spatio-temporal aggregation: Worboys (1994) defines a unified model for space and time and Camossi et al. (2003) introduce a multi-granular spatio-temporal data model. Jeong et al. (2004) define a generic algorithmic framework for spatio-temporal aggregation processes in databases. Related research regarding sensor observations deals with the aggregation of low-level sensor data to reduce the communication load from sensors to databases and clients. For example, Madden et al. (2002) introduce a tiny aggregation service for in-network aggregation of observations. However, in the case of low-level sensor data aggregation, observations with a higher resolution are usually lost. This is in contrast to our approach which provides flexible spatio-temporal aggregation of sensor observations to different aggregation levels in the Web.

Geosensor networks are interconnected sensors for monitoring environmental phenomena or geographic processes (Nittel and Stefanidis, 2005). The Sensor Web thereby abstracts from low-level interfaces and protocols in geosensor networks by adding an additional application layer in the Web (Broering et al., 2011). The Sensor Web Enablement (SWE) initiative of the

Open Geospatial Consortium (OGC) aims to standardize the Sensor Web with a suite of standardized interfaces (Botts et al., 2007). The goal of the SWE initiative is to improve the interoperability of discovery, access and tasking of sensors in the Web. The Sensor Observation Service (SOS) forms the web-based interface for accessing observations and sensor metadata in the Sensor Web (Na and Priest, 2007). It allows client applications to query different kinds of observations through standardized operations and filters and retrieve the observations in a common format. The available observation data in the SOS can be retrieved in the Observations&Measurements (O&M) format, which is a model and encoding for observations (Cox, 2007a). Metadata about sensors that are registered at a SOS interface is provided in the Sensor Model Language (SensorML) (Botts and Robin, 2007). The observations can be queried in a flexible way from a SOS interface regarding space, time or thematic attributes. Though Havlik et al. (2009) introduce a system of cascading SOS instances, which is able to aggregate observations in time, an (spatio-temporal) aggregation functionality is currently not supported by the SOS interface. Following seperation of concerns, aggregation *functionality* should be rather provided by other processing services and the aggregated observations should be accessible via the SOS interface.

In the past, most Geoprocessing functionality has been provided by monolithic Geographic Information Systems (GIS). By standardizing the interface for geoprocessing on the Web such as the Web Processing Service (WPS) (Schut, 2007), geoprocessing functionality has been integrated into various applications (Brauner et al., 2009) and the Geoprocessing Web evolved. The Geoprocessing Web makes geoprocessing functionality available on the web, which can be used interchangeably. To ensure interoperability of this functionality, profiles have been proposed to be used in the Geoprocessing Web. A profile consists of unique identifiers for its processes implemented as Unified Resource Names (URN), and of process descriptions including the definition of input and output parameters. An example of a profile related to aggregation is described by Foerster (2010) in the context of generalization. Related to processing of observations, Chen et al. (2010) describe a standards based processing system for wildfire detection in an Sensor Web environment. The use of standardized geoprocessing in wildfire analysis, smoke data analysis, and forecast has also been described and evaluated by Falke et al. (2008). As a possibility for a web-based aggregation, Pebesma et al. (2011) introduce a web service for the automated spatial interpolation of observations. However, the service does not provide spatio-temporal interpolation methods.

When processing sensor data in the Web, provenance information is crucial to determine the quality of the information derived. Recently, several initiatives have developed models for providing provenance information in the Web. The Open Provenance Model (OPM)[1] defines a model for provenance graphs enabling to retrace an information item in the Web back to its origin. Similarly, a Provenance Vocabulary has been defined that can be used, for example, in Linked Open Data (Hartig and Zhao, 2010). Related to sensors, Liu et al. (2010) propose a provenance-aware virtual sensor using the OPM. The virtual sensor provides continuous observations estimated from values gathered by surrounding physical sensors. We are also conceptualizing the

[1]http://openprovenance.org/

aggregation process as a virtual sensor, but rather in the sense of a software sensor like described by Kabadayi et al. (2006) to integrate the aggregation process in our observation data. Instead of adding additional provenance metadata like described by Park and Heidemann (2008), the provenance information is directly provided in our model. Thus, following the final report of the W3C Provenance Incubator Group[2], our approach is providing provenance information passed by value and embedded in the representations. It allows to retrieve relevant provenance information similar as described by Patni et al. (2010) for Linked Open Data.

3. A Data Model for Observations across Different Aggregation Levels

This section describes the tailored observation model that can be used for observations at different aggregation levels. The model is based on the O&M model (Section 2) and is shown in Figure 1. The `result` of an observation can either be a numerical value with information about the unit of measurement (uom) or a value coverage. While the single value can be used to represent non-aggregated as well as aggregated observations, the value coverage can be used for interpolation results. The `procedure` has created an observation. In order to represent aggregation processes, we have added an `AggregationProcess` as a procedure (see also Section 4). In case of non-aggregated observations, we are using a `SensorSystem` that can represent single sensors as well as sensor systems like air quality monitoring stations.

[2] http://www.w3.org/2005/Incubator/prov/XGR-prov-20101214/

The spatial geometry of an observation is part of the `featureOfInterest`. Environmental phenomena are usually fields in geographic space. Thus, an in-situ measurement is a sampling at a certain location of the field-based phenomenon. We have restricted the `featureOfInterest` either to be a `SF_SamplingPoint` for a point as geometry or to be a `SF_SamplingSurface` for a polygon. Both features are defined in the sampling feature specification (Cox, 2007b) and thus provide a reference to a superior feature, the `sampledFeature`. In case of an aggregation, the featureOfInterest usually changes from `SF_SamplingPoint` to `SF_SamplingPolygon`. An aggregation from smaller to larger polygons is also possible. The `observedProperty` represents the phenomenon that is observed (e.g. ozone concentration). Usually, it is a property of the sampled feature (e.g. atmosphere). No specific adjustments of the `observedProperty` are necessary in our model. The time when an observation applies is kept in the `samplingTime`. In our model, this can either be an instant in time or a time period. While the time instant can only be used for non-aggregated observations, the time period is usually used for aggregated observations.

When aggregating observations, information about the quality of the aggregation result is crucial. Therefore, the `resultQuality` can contain statistical information like standard deviation (e.g. in addition to averaged values). In addition, provenance information is needed. The `AggregationProcess`

Figure 1: UML diagram of the observation model that can be used at different aggregation levels. The blue-colored classes stem from the O&M model while the yellow-colored classes have been added to represent different aggregation levels.

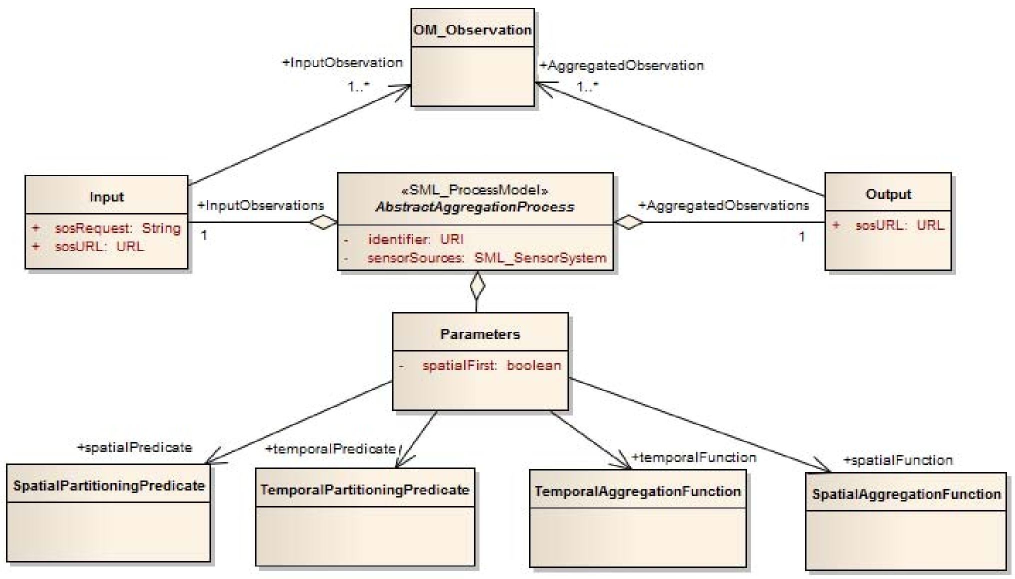
contains associations to the `SensorSystems` which have gathered the original observations. This is not applicable for providing sufficient provenance information, as not all observations produced by a sensor are aggregated. Thus, we have introduced the `isAggregateOf` association between observations for referencing the original observations through identifiers. Each observation which has been aggregated can be retraced by following the `isAggregateOf` association. The original observations again provide information about the time when they have been produced, the spatial location and the initial measurement process as described above. Thus, our model contains provenance information about the original observations as well as the aggregation process.

4. A Process Model for Spatio-temporal Aggregation Processes

Based on the tailored observation model for different aggregation levels described in the previous section, this section defines the aggregation process model. In an aggregation process, observations are grouped by a partitioning predicate before applying a certain aggregation function to its values (Section 2). We currently consider the partitioning predicates to be spatial and/or temporal. Thus, observations are grouped spatially and/or temporally and aggregation functions are then applied to the result values of the observations in one group to calculate a new aggregated observation. Figure 2 shows the UML diagram of our basic aggregation process model.

The main class is the `AbstractAggregationProcess`. All aggregation processes are `ProcessModels` as defined in SensorML and thus are having an `Input`, an `Output`, and `Parameters`. Additionally, they have a name, a de-

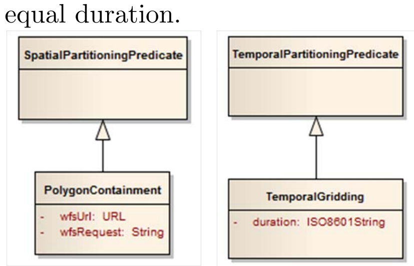
Figure 2: UML diagram of the model for aggregation processes. An aggregation process has a set of observations as inputs and outputs. It is defined by the parameters containing the grouping predicates and the aggregation functions.



scription and metadata properties including a global identifier. For simplicity, we only show the `identifier` of the aggregation process in our diagram. The `Input` of an `AbstractAggregationProcess` acts as a container for several input observations. In the same way, the `Output` of an `AggregationProcess` is a container for several aggregated observations. Both, `Input` and `Output`, can either directly contain the observations or reference SOS instances. The `Parameters` class contains the spatial and temporal predicates as well as the aggregation functions that are used to aggregate the result values of the observations. The subtypes of the `AbstractAggregationProcess` are defined by the subtypes of the predicates and aggregation functions.

Several subtypes of the predicates can be defined. Two examples are shown in Figure 3. For example, the spatial predicate `PolygonContainment` is defined to aggregate point measurements to polygons that are retrieved from a Web Feature Service (WFS), the main web service for retrieving spatial vector data. Thus, the type has two additional parameters, namely an URL pointing to a WFS and a request which selects certain polygon features. The temporal predicate `TemporalGridding` groups the temporal extent of all observations into time intervals of equal duration. In the same manner, aggregation functions can be defined, that are applied to the result values after spatial or temporal grouping. Depending on the chosen aggregation function, the order of grouping and the order of applying an aggregation function (spatial first or temporal first) can be of importance. For example, computing the daily maxima first and then averaging them spatially is different than calculating spatial averages first and then applying the max function to the spatial averages. Thus, an additional parameter indicating the order is introduced (`spatialFirst`). Together with our tailored observation model, we now have the two models that allow us to define the aggregation processes that shall be provided by our spatio-temporal aggregation service.

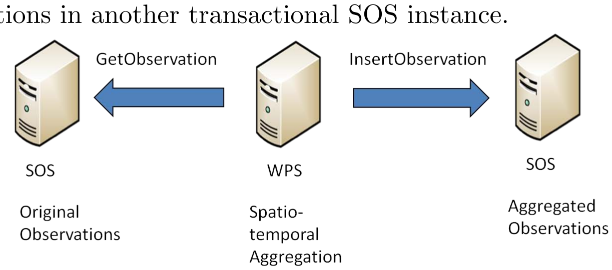
Figure 3: UML diagram showing two subtypes of spatial and temporal partitioning predicates. The PolygonContainment predicate is defined to group points to polygons that are provided by a WFS. The TemporalGridding predicate is used for paritioning a temporal extent in time intervals of equal duration.



5. Spatio-Temporal Aggregation Service

To provide aggregation functionality in the Sensor Web in an interoperable way, we define the Spatio-Temporal Aggregation Service (STAS) as a profile of the OGC WPS. Figure 4 illustrates the basic service architecture for spatio-temporal aggregation in the Sensor Web consisting of SOS instances and the STAS. The STAS can be linked dynamically with SOS instances to retrieve input observations and publish the aggregated observations.

Figure 4: Basic service architecture for spatio-temporal aggregation in the Sensor Web. The STAS queries non-aggregated observations from a SOS instance and inserts the aggregated observations in another transactional SOS instance.



According to the definition of WPS profiles (Section 2), we now describe the two parts of our STAS profile: An URN scheme for defining the identifiers of the processes, and the implementation of process descriptions according to our aggregation process model (Section 4). Following the common URN scheme of the OGC as defined in Whiteside (2009), a basic URN has the form `urn:ogc:def:objectType`. Thus, the basic URN for aggregation processes is defined as `urn:ogc:def:aggregationProcess`. The spatial grouping predicate (sgp) and spatial aggregation function (saf) is appended to the basic URN followed by the temporal grouping predicate (tgp) and temporal aggregation function (taf). Thus, the resulting generic URN is `urn:ogc:def:aggregationProcess:sgp:saf:tgp:taf`. For example, the URN `urn:ogc:def:aggregationProcess:polygonContainment:spatial-Mean:temporalGridding:temporalMax` identifies the process which averages observations that are contained in polygons and calculates the maximum value of these aggregates for temporal intervals. Depending on the value of the `spatialFirst` parameter (see Section 4), the temporal aggregation can also be applied first. The additional parameters needed for the predicates and aggregation functions are defined in the process descriptions as defined in the next paragraph.

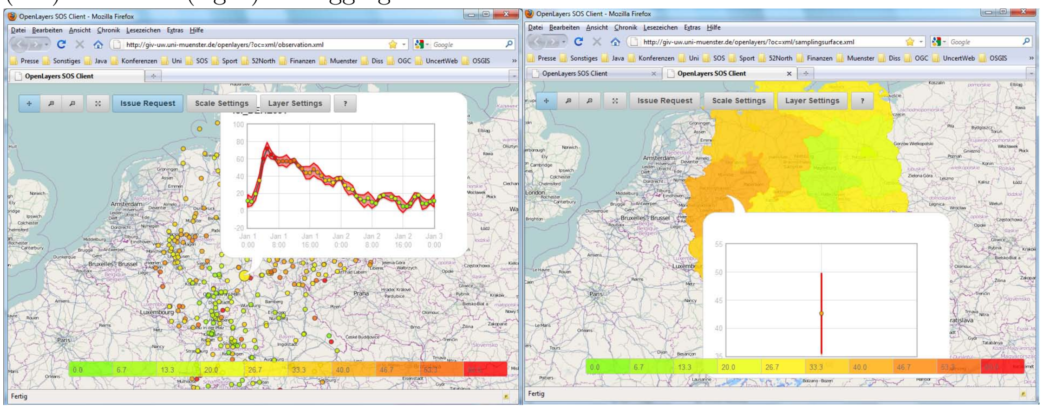
To perform a spatio-temporal aggregation, the Execute operation of the STAS has to be invoked. The parameters of the Execute request for a specific aggregation process are described in its process description, which can be retrieved through the DescribeProcess operation. We now describe the parameters of the aggregation processes. The URN of the process has to be passed in every Execute request to identify the aggregation process. Additionally, each Execute request contains an URL of the SOS instance providing the input observations and a SOS request to identify the observations which should be aggregated. As the output observations are also published via the SOS interface, another input parameter is the URL of the SOS instance where the aggregated observations should be published. Following the aggregation process model (Section 4), each aggregation process is defined by the subtypes of the spatial and temporal grouping predicates and aggregation functions. Depending on these subtypes, each aggregation process has additional parameters. In case of the examples shown in Figure 3, the additional parameters are an URL pointing to a WFS and a request identifying polygonal features in WFS. For the temporal gridding, a duration has to be defined and passed to the STAS. Finally, the aggregation functions can define additional parameters. For example, if the spatial aggregation method is a Kriging interpolation method, there will be additional Kriging parameters like a variogram. After aggregating the observations, every aggregation process returns a reference to the SOS instance offering the set of aggregated observations. If requested by the client, the aggregated observations can also be included in the Execute response.

6. Case Study - Spatio-Temporal Aggregation of European Air Quality Observations

This section presents the spatio-temporal aggregation of European air quality observations, as provided by the EEA, to demonstrate the developed approach (Section 3-5). In particular, the approach is implemented to aggregate Ozone observations collected in Germany to mean and maximum values in temporal intervals and to averages of these temporal aggregates in space. Figure 5 shows the OpenLayers client visualizing observations before the aggregation (left side) and after the aggregation (right side). The red bar behind the aggregated value is showing the confidence interval calculated from the standard deviation of the values. The statistics are encoded in O&M using the Uncertainty Markup Language (UncertML) (Williams et al., 2009).

The service deployment for the case study is shown in Figure 6. One SOS instance serves approximately 30 million observations extracted from the AirBase database files of the European Environmental Agency (EEA, 2011).

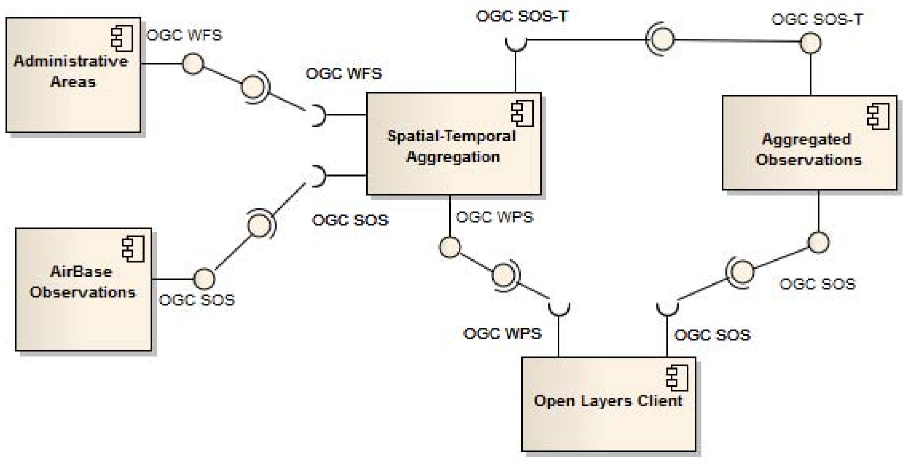
Figure 5: Screenshots of the OpenLayers client visualizing air quality observations before (left) and after (right) the aggregation.



Administrative areas of Germany are provided as vector-based features by a WFS. The spatio-temporal aggregation is implemented in a prototype of the STAS. Currently, it provides means to aggregate points to polygons in space and to partition temporal extents of observations into time intervals. Supported aggregation functions for space and time are MIN, MAX, SUM and MEAN. Another SOS instance provides the aggregated observations. As a client, the OpenLayers SOS client (Eijnden, 2010) has been extended and a simple user interface has been developed for sending requests to SOS and WPS.

Figure 7 shows the workflow of an exemplary spatio-temporal aggregation of ozone observations. As an aggregation process is executed on-demand, a client sends an Execute request to the STAS. Then, the STAS retrieves the observations from a SOS instance and, in parallel, the administrative areas from a WFS instance. Now, the observations are grouped temporally, then aggregated, and afterwards these temporal aggregates are grouped and aggregated spatially. The aggregated observations are then inserted into another

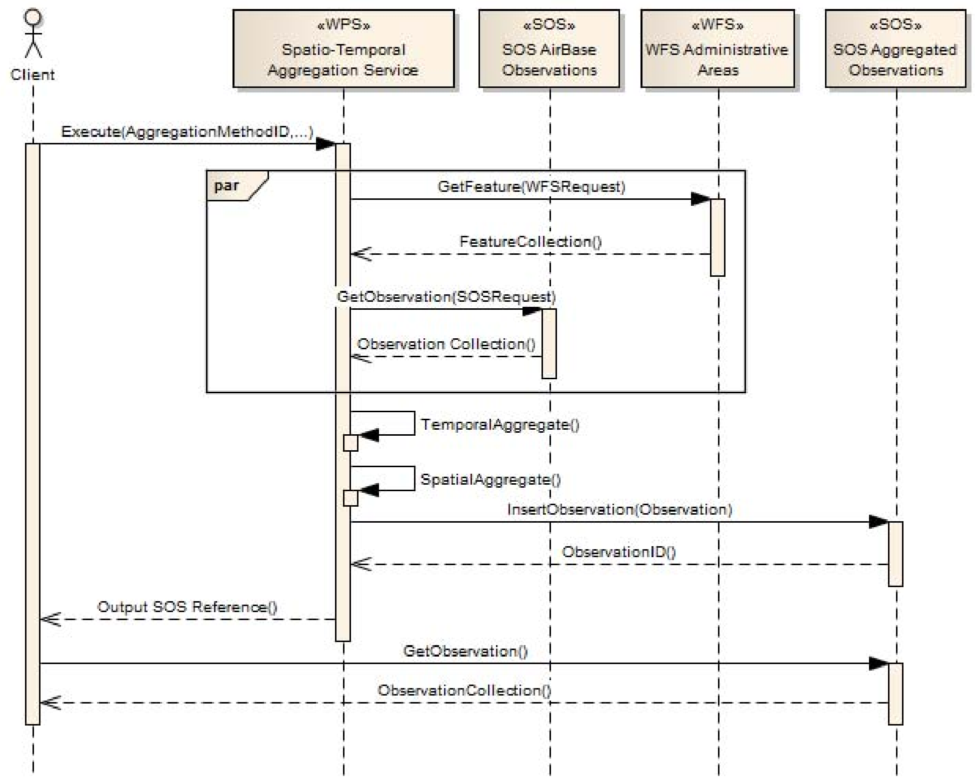
Figure 6: Service deployment for spatio-temporal aggregation of European air quality observations. The architecture consists of one SOS instance offering the basic air quality observations, a WPS instance implementing the STAS, a SOS instance that provides the aggregated observations, and a WFS providing administrative areas of Germany. The OpenLayers client has been customized to interact with WPS and SOS and to visualize observations.



SOS instance and the reference to the aggregated observations is returned by the STAS to the client. The client can now retrieve the aggregated observations from a SOS. As an option, the aggregated observations can also be directly contained in the aggregation service response. The aggregation functions applied can be easily exchanged by invoking another aggregation process.

Provenance information is provided directly in the aggregated observations. The isAggregateOf element points to the original ozone observations provided by SOSs. Retrieving this observations allows to retrieve similar provenance information like described by Patni et al. (2010) for Linked Data:

Figure 7: Workflow of spatio-temporal aggregation of European air quality observations.



the location and time of the observation as well as the description of the measurement procedure that has produced the observation result provided as SensorML description. The procedure element of the aggregated observations points to the SensorML description of the aggregation process. Thus, information about the applied aggregation function and predicates can be retrieved. We do net yet include information about the users of the aggregated or source observations like described in Park and Heidemann (2008) and consider this to be future work.

7. Discussion & Conclusions

In this paper, we describe an approach for spatio-temporal aggregation of observations in the Sensor Web. As a foundation for our approach, we define a tailored observation model (Section 3) that can be used at different aggregation levels. Additionally, a process model for spatio-temporal aggregation processes is developed (Section 4). To realize the spatio-temporal aggregation in the Sensor Web, a service-oriented architecture using Geospatial Web Services is proposed. The central component of this architecture is the Spatio-Temporal Aggregation Service (STAS) that is defined as a profile of the OGC Web Processing Service (Section 5). The STAS can be dynamically linked with Sensor Observation Service (SOS) instances that offer non-aggregated observations and with SOS instances that allow to insert the aggregated observations. Hence, by relying on these standardized service interfaces the aggregation methods can be easily re-used in the Web and the observations are always provided in the same format at different aggregation levels. Additionally, the tailored observation model allows to retrace non-aggregated observations from aggregated observations. Furthermore, the reference from observations to the aggregation process allows to retrieve information about the aggregation process in a standardized format. The presented approach is applied to a case study of aggregating air quality data (Section 6).

Our case study shows that the STAS allows for a flexible integration of aggregation processes within SOSs and, thus, in the Sensor Web and Geoprocessing Web. Aggregation methods can be exchanged in a flexible way, while the common service operations remain the same for both, the execution of the aggregation as well the access to the aggregated observations. This eases the integration in clients and other applications. Other approaches (Liu et al., 2008; Kabadayi et al., 2006; Patni et al., 2010) usually focus on aggregating raw measurements provided in a proprietary format and then providing the aggregates in standardized formats. Furthermore, the STAS aggregates a set of source observations to a set of aggregated observations whereas the other approaches usually focus on producing a single aggregate out of several observations.

The process model is general enough for the definition of different aggregation processes, as the spatial and temporal grouping predicates and the aggregation functions allow for additional parameters depending on the method chosen. However, we do not yet consider a thematic aggregation as described by Patni et al. (2010) where the grouping predicate might be thematic (e.g. high windspeed, low temperature, heavy snowfall) in order to aggregate the observations to a higher level event (e.g. Blizzard). It has to be explored, whether the STAS can also be used for such a thematic aggregation of observations.

As the STAS is implemented as a WPS profile, it allows for an on-demand aggregation. Most approaches for aggregating observations through virtual sensors use predefined aggregation processes and the execution time of the aggregation is predefined as well, see for example Liu et al. (2008). However, the STAS might be used in these approaches to provide the aggregation processing, for example the estimation of values for a new virtual sensor where no physical sensors are available. Encapsulating the aggregation processes through an WPS interface allows to integrate aggregation in Geoprocessing workflows. The STAS is able to mediate between different observation and model services providing or requiring the data in different resolutions. In order to realize complex aggregation workflows that require several aggregation steps, a WPS-T interface as proposed by Schaeffer (2008) might be used. It allows to create complex workflows out of the basic aggregation processes which then can be published as complex WPS processes again. This enables the STAS to provide complex aggregation chains in the same manner as the simple aggregation processes. While we propose a model for aggregation of observation, we did not yet provide disaggregation methods like e.g. disaggregating grid cells to point observations. Their integration needs to be explored in future work.

While the STAS allows for a flexible and seamless integration of aggregation processes in the Sensor Web, the communication overhead for transferring large datasets between the services has been identified as a possible drawback of the approach. In that case, one approach might be to tightly couple the STAS with a SOS instance and thus reduce the communication overhead: Instead of passing the URLs of the SOS instances, the observations can automatically be fetched from a SOS running on the same machine. Also, the aggregated observations can directly be inserted in this SOS instance again. Another approach to cope with large datasets might be efficient caching and indexing strategies as described by Sivasubramanian et al. (2007).

To enable the discovery of already deployed aggregation processes in the Web and to automate observation aggregation workflows, semantic descriptions of observations and aggregation processes are needed. In a next step, we plan to semantically enable the STAS by using already existing approaches for integrating semantics in web-based spatial data infrastructures (Janowicz et al., 2010) and by utilizing observation ontologies like defined by the W3C (Janowicz and Compton, 2010). Finally, we consider to extend the concept of the STAS from aggregation processes for observations to aggregation processes for general spatio-temporal data in SDIs.

Acknowledgments

The presented work is developed within the 52° North statistics community, and partly funded by the European project *UncertWeb* (FP7-248488, see http://www.uncertweb.org). We are thankful to the reviewers for the valuable comments provided to improve the paper.

References

Botts, M., Percivall, G., Reed, C., Davidson, J. (Eds.), 2007. OGC Sensor Web Enablement: Overview And High Level Architecture. OGC 07-165. http://portal.opengeospatial.org/files/?artifact_id=25562 [[accessed 20 March 2011.

Botts, M., Robin, A. (Eds.), 2007. OpenGIS Sensor Model Language (SensorML) Implementation Specification. OGC 07-000. Open Geospatial Consortium, Inc., 180 pp, http://portal.opengeospatial.org/files/?artifact_id=21273, [accessed 22 March, 2011].

Brauner, J., Foerster, T., Schaeffer, B., Baranski, B., 2009. Towards a Research Agenda for Geoprocessing Services, in: 12th AGILE International Conference on Geographic Information Science.

Broering, A., Echterhoff, J., Jirka, S., Simonis, I., Everding, T., Stasch, C., Liang, S., Lemmens, R., 2011. New Generation Sensor Web Enablement. Sensors 11, 2652–2699.

Camossi, E., Bertolotto, M., Bertino, E., Guerrini, G., 2003. A Multigranular Spatiotemporal Data Model, in: GIS '03: Proceedings of the 11th ACM international symposium on Advances in geographic information systems, ACM, New York, NY, USA. pp. 94–101.

Chen, N., Di, L., Yu, G., Gong, J., 2010. Geo-processing Workflow Driven Wildfire Hot Pixel Detection under Sensor Web Environment. Computers & Geosciences 36, 362–372.

Cox, S. (Ed.), 2007a. Observation and Measurements - Part 1 - Observation Schema. OGC 07-022r1. Open Geospatial Consortium, Inc., 86 pp, http://portal.opengeospatial.org/files/?artifact_id=22466, [accessed 22 March, 2011].

Cox, S. (Ed.), 2007b. Observation and Measurements - Part 2 - Sampling Features. OGC 07-002r3. Open Geospatial Consortium, Inc., 46 pp,http://portal.opengeospatial.org/files/?artifact_id=22467, [accessed 22 March, 2011].

EEA, 2011. AirBase - The European air quality database - datasets. http://www.eea.europa.eu/data-and-maps/data/airbase-the-european-air-quality-database-2 [accessed 22 March, 2011].

Eijnden, B.v. (Ed.), 2010. OpenLayers: SOS and Inspire. Bart van Eijnden. Presentation at FOSS4G. http://2010.foss4g.org/presentations/2891.pdf [accessed 22 March, 2011].

Falke, S., Dorner, E., Dunn, B., Sullivan, D., 2008. Processing Services in Earth Observation Sensor Web Information Architectures, in: Earth Science Technology Conference, University of Maryland, June 24-26.

Foerster, T., 2010. Web-based Architecture for On-demand Maps - Integrating Meaningful Generalization Processing. Ph.D. thesis. University of Twente.

Hartig, O., Zhao, J., 2010. Publishing and consuming provenance metadata on the web of linked data, in: In: Proc. of 3rd Int. Provenance and Annotation Workshop.

Havlik, D., Bleier, T., Schimak, G., 2009. Sharing Sensor Data with SensorSA and Cascading Sensor Observation Service. Sensors 9, 5493–5502.

Janowicz, K., Compton, M., 2010. The Stimulus-Sensor-Observation Ontology Design Pattern and its Integration into the Semantic Sensor Network Ontology, in: Kerry Taylor, Arun Ayyagari, D.D.R. (Ed.), Proceedings of the 3rd International workshop on Semantic Sensor Networks 2010 (SSN10) in conjunction with the 9th International Semantic Web Conference (ISWC 2010), CEUR-WS.

Janowicz, K., Schade, S., Broering, A., Kessler, C., Maue, P., Stasch, C., 2010. Semantic Enablement for Spatial Data Infrastructures. Transactions in GIS 14, 111–129.

Jeong, S.H., Fernandes, A.A.A., Paton, N.W., Griffiths, T., 2004. A Generic Algorithmic Framework for Aggregation of Spatio-Temporal Data, in: SS-DBM '04: Proceedings of the 16th International Conference on Scientific and Statistical Database Management, IEEE Computer Society, Washington, DC, USA. p. 245.

Kabadayi, S., Pridgen, A., Julien, C., 2006. Virtual Sensors: Abstracting Data from Physical Sensors, in: 4th International Workshop on Mobile Distributed Computing.

Liu, Y., Abdelzaher, T., Heo, J., Choi, J., Minsker, B., Fazio, D., 2008. Virtual sensor-powered spatiotemporal aggregation and transformation: A case study analyzing near-real-time nexrad and precipitation gage data in a digital watershed, in: Environmental Information Management Conference 2008, September 10 - 11.

Liu, Y., Futrelle, J., Myers, J., Rodriguez, A., Kooper, R., 2010. A provenance-aware virtual sensor system using the open provenance model, in: Collaborative Technologies and Systems (CTS), 2010 International Symposium on, pp. 330 –339.

Madden, S., Franklin, M.J., Hellerstein, J.M., Hong, W., 2002. TAG: a Tiny AGgregation service for ad-hoc sensor networks. SIGOPS Oper. Syst. Rev. 36, 131–146.

Na, A., Priest, M. (Eds.), 2007. Sensor Observation Service. OGC 06-009r6. Open Geospatial Consortium, Inc., 104pp. [accessed 22 March 2011].

Nittel, S., Stefanidis, A., 2005. GeoSensors Networks. CRC. chapter GeoSensor Networks and Virtual GeoReality.

Park, U., Heidemann, J., 2008. Provenance in sensornet republishing, in: Proceedings of the 2nd International Provenance and Annotation Workshop, Springer-Verlag, Salt Lake City, Utah, USA. pp. 208–292.

Patni, P., Sahoo, S., Henson, C., Sheth, A., 2010. Provenance aware linked sensor data, in: Kärger, P., Olmedilla, D., Passant, A., Polleres, A. (Eds.), Proceedings of the Second Workshop on Trust and Privacy on the Social and Semantic Web.

Pebesma, E., Cornford, D., Dubois, G., Heuvelink, G.B., Hristopulos, D., Pilz, J., Sthlker, U., Morin, G., Skien, J.O., 2011. INTAMAP: The design and implementation of an interoperable automated interpolation web service. Computers & Geosciences 37, 343 – 352.

Schaeffer, B., 2008. Towards a Transactional Web Processing Service, in: Pebesma, E., Bishr, M., Bartoscheck, T. (Eds.), 6th Geographic Information Days, Ifgi Prints. pp. 91–116.

Schut, P. (Ed.), 2007. OpenGIS Web Processing Service. OGC 05-007r7. Open Geospatial Consortium, Inc., 87pp. http://portal.opengeospatial.org/files/?artifact_id=24151 [accessed 22 March, 2011].

Sivasubramanian, S., Pierre, G., Steen, M., Alonso, G., 2007. Analysis of caching and replication strategies for Web applications. IEEE Internet Computing 11, 60–66.

Vega Lopez, I.F., Snodgrass, R.T., Moon, B., 2005. Spatiotemporal Aggregate Computation: A Survey. IEEE Trans. on Knowl. and Data Eng. 17, 271–286.

WetterOnline (Ed.), 2011. Wetter WetterOnline Wettervorhersage Aktuelles Wetter Regenradar Reisewetter http://www.wetteronline.de/, [accessed 22 March, 2011].

Whiteside, A. (Ed.), 2009. Definition identifier URNs in OGC namespace. OGC 07-092r3. Open Geospatial Consortium, Inc., 46pp. http://portal.opengeospatial.org/files/?artifact_id=30575 [accessed 22 March, 2011].

Williams, M., Cornford, D., Bastin, L., Pebesma, E. (Eds.), 2009. Uncertainty Markup Language. OGC 08-122r2. Open Geospatial Consortium, Inc., 61pp. http://portal.opengeospatial.org/files/?artifact_id=33234 [accessed 22 March, 2011].

Worboys, M., 1994. A Unified Model for Time and Space. The Computer Journal 37, 26–34.

Zhao, P., Di, L., 2010. Geospatial Web Services: Advances in Information Interoperability. Information Science Publishing.