

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.

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

- 30 1. A data model for observations that can be used across different aggre-
31 gation levels. This model also incorporates metadata about provenance
32 and aggregation method (Section 3).
- 33 2. A process model for spatio-temporal aggregation (Section 4).
- 34 3. A web service architecture for aggregation of observations including the
35 definition of the Spatio-Temporal Aggregation Service (STAS) (Section
36 5).

37 In our approach, we propose a Service-Oriented Architecture (SOA) for
38 spatio-temporal aggregation of observations. As the Open Geospatial Con-
39 sortium (OGC) provides well-defined encodings and service interfaces for
40 both, the Sensor Web and the Geoprocessing Web, we are utilizing these
41 standards in our approach. As a basis for our SOA, we define a tailored
42 observation model and process model for spatio-temporal aggregation. The
43 proposed SOA consists of Sensor Observation Services (SOS), the standard
44 service for providing observations in the Sensor Web (Na and Priest, 2007),
45 and the Spatio-Temporal Aggregation Service (STAS), which is defined as a

46 profile of the Web Processing Service (WPS). The WPS provides the basic
47 service interface for the Geoprocessing Web (Schut, 2007). In a case study,
48 we demonstrate how our approach meets the requirements identified above
49 by *temporally* aggregating hourly measurements of Ozone to daily maxima
50 and by *spatially* averaging these maxima for each federal state in Germany.

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

60 **2. Background**

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

69 An *aggregation process* computes a single value, an *aggregate*, for a group

70 of attribute values by means of an *aggregation function* (Jeong et al., 2004).
71 The attribute values are grouped by a *partitioning predicate*. In our work,
72 spatio-temporal aggregation combines objects in space and time and provides
73 means to compute aggregates for certain attribute values attached to these
74 objects. Most of the research on spatio-temporal aggregation during the last
75 years has focused on improving aggregation operations in spatio-temporal
76 databases. For example, Vega Lopez et al. (2005) give a comprehensive sur-
77 vey on spatio-temporal aggregation methods in databases. Others develop
78 general models for space and time that can be used as a basis for spatio-
79 temporal aggregation: Worboys (1994) defines a unified model for space and
80 time and Camossi et al. (2003) introduce a multi-granular spatio-temporal
81 data model. Jeong et al. (2004) define a generic algorithmic framework for
82 spatio-temporal aggregation processes in databases. Related research regard-
83 ing sensor observations deals with the aggregation of low-level sensor data
84 to reduce the communication load from sensors to databases and clients.
85 For example, Madden et al. (2002) introduce a tiny aggregation service for
86 in-network aggregation of observations. However, in the case of low-level sen-
87 sor data aggregation, observations with a higher resolution are usually lost.
88 This is in contrast to our approach which provides flexible spatio-temporal
89 aggregation of sensor observations to different aggregation levels in the Web.

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

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

114 In the past, most Geoprocessing functionality has been provided by mono-
115 lithic Geographic Information Systems (GIS). By standardizing the interface
116 for geoprocessing on the Web such as the Web Processing Service (WPS)
117 (Schut, 2007), geoprocessing functionality has been integrated into various
118 applications (Brauner et al., 2009) and the Geoprocessing Web evolved. The
119 Geoprocessing Web makes geoprocessing functionality available on the web,

120 which can be used interchangeably. To ensure interoperability of this func-
121 tionality, profiles have been proposed to be used in the Geoprocessing Web. A
122 profile consists of unique identifiers for its processes implemented as Unified
123 Resource Names (URN), and of process descriptions including the definition
124 of input and output parameters. An example of a profile related to aggrega-
125 tion is described by Foerster (2010) in the context of generalization. Related
126 to processing of observations, Chen et al. (2010) describe a standards based
127 processing system for wildfire detection in an Sensor Web environment. The
128 use of standardized geoprocessing in wildfire analysis, smoke data analysis,
129 and forecast has also been described and evaluated by Falke et al. (2008). As
130 a possibility for a web-based aggregation, Pebesma et al. (2011) introduce a
131 web service for the automated spatial interpolation of observations. However,
132 the service does not provide spatio-temporal interpolation methods.

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

¹<http://openprovenance.org/>

143 aggregation process as a virtual sensor, but rather in the sense of a software
144 sensor like described by Kabadayi et al. (2006) to integrate the aggregation
145 process in our observation data. Instead of adding additional provenance
146 metadata like described by Park and Heidemann (2008), the provenance in-
147 formation is directly provided in our model. Thus, following the final report
148 of the W3C Provenance Incubator Group², our approach is providing prove-
149 nance information passed by value and embedded in the representations. It
150 allows to retrieve relevant provenance information similar as described by
151 Patni et al. (2010) for Linked Open Data.

152 **3. A Data Model for Observations across Different Aggregation** 153 **Levels**

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

²<http://www.w3.org/2005/Incubator/prov/XGR-prov-20101214/>

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

182 When aggregating observations, information about the quality of the ag-
183 gregation result is crucial. Therefore, the `resultQuality` can contain statis-
184 tical information like standard deviation (e.g. in addition to averaged values).
185 In addition, provenance information is needed. The `AggregationProcess`
186 contains associations to the `SensorSystems` which have gathered the original
187 observations. This is not applicable for providing sufficient provenance infor-
188 mation, as not all observations produced by a sensor are aggregated. Thus,
189 we have introduced the `isAggregateOf` association between observations for

190 referencing the original observations through identifiers. Each observation
191 which has been aggregated can be retraced by following the `isAggregateOf`
192 association. The original observations again provide information about the
193 time when they have been produced, the spatial location and the initial mea-
194 surement process as described above. Thus, our model contains provenance
195 information about the original observations as well as the aggregation pro-
196 cess.

197 4. A Process Model for Spatio-temporal Aggregation Processes

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

207 The main class is the `AbstractAggregationProcess`. All aggregation
208 processes are `ProcessModels` as defined in SensorML and thus are having an
209 `Input`, an `Output`, and `Parameters`. Additionally, they have a name, a de-
210 scription and metadata properties including a global identifier. For simplicity,
211 we only show the `identifier` of the aggregation process in our diagram. The
212 `Input` of an `AbstractAggregationProcess` acts as a container for several
213 input observations. In the same way, the `Output` of an `AggregationProcess`

214 is a container for several aggregated observations. Both, `Input` and `Output`,
215 can either directly contain the observations or reference SOS instances. The
216 `Parameters` class contains the spatial and temporal predicates as well as the
217 aggregation functions that are used to aggregate the result values of the ob-
218 servations. The subtypes of the `AbstractAggregationProcess` are defined
219 by the subtypes of the predicates and aggregation functions.

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

238 5. Spatio-Temporal Aggregation Service

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

245 According to the definition of WPS profiles (Section 2), we now describe
246 the two parts of our STAS profile: An URN scheme for defining the iden-
247 tifiers of the processes, and the implementation of process descriptions ac-
248 cording to our aggregation process model (Section 4). Following the com-
249 mon URN scheme of the OGC as defined in Whiteside (2009), a basic URN
250 has the form `urn:ogc:def:objectType`. Thus, the basic URN for aggrega-
251 tion processes is defined as `urn:ogc:def:aggregationProcess`. The spa-
252 tial grouping predicate (`sgp`) and spatial aggregation function (`saf`) is ap-
253 pended to the basic URN followed by the temporal grouping predicate (`tgp`)
254 and temporal aggregation function (`taf`). Thus, the resulting generic URN
255 is `urn:ogc:def:aggregationProcess:sgp:saf:tgp:taf`. For example, the
256 URN `urn:ogc:def:aggregationProcess:polygonContainment:spatial-`
257 `Mean:temporalGridding:temporalMax` identifies the process which averages
258 observations that are contained in polygons and calculates the maximum
259 value of these aggregates for temporal intervals. Depending on the value of
260 the `spatialFirst` parameter (see Section 4), the temporal aggregation can
261 also be applied first. The additional parameters needed for the predicates
262 and aggregation functions are defined in the process descriptions as defined

263 in the next paragraph.

264 To perform a spatio-temporal aggregation, the Execute operation of the
265 STAS has to be invoked. The parameters of the Execute request for a spe-
266 cific aggregation process are described in its process description, which can
267 be retrieved through the DescribeProcess operation. We now describe the
268 parameters of the aggregation processes. The URN of the process has to be
269 passed in every Execute request to identify the aggregation process. Addi-
270 tionally, each Execute request contains an URL of the SOS instance providing
271 the input observations and a SOS request to identify the observations which
272 should be aggregated. As the output observations are also published via
273 the SOS interface, another input parameter is the URL of the SOS instance
274 where the aggregated observations should be published. Following the ag-
275 gregation process model (Section 4), each aggregation process is defined by
276 the subtypes of the spatial and temporal grouping predicates and aggrega-
277 tion functions. Depending on these subtypes, each aggregation process has
278 additional parameters. In case of the examples shown in Figure 3, the addi-
279 tional parameters are an URL pointing to a WFS and a request identifying
280 polygonal features in WFS. For the temporal gridding, a duration has to be
281 defined and passed to the STAS. Finally, the aggregation functions can define
282 additional parameters. For example, if the spatial aggregation method is a
283 Kriging interpolation method, there will be additional Kriging parameters
284 like a variogram. After aggregating the observations, every aggregation pro-
285 cess returns a reference to the SOS instance offering the set of aggregated
286 observations. If requested by the client, the aggregated observations can also
287 be included in the Execute response.

288 **6. Case Study - Spatio-Temporal Aggregation of European Air** 289 **Quality Observations**

290 This section presents the spatio-temporal aggregation of European air
291 quality observations, as provided by the EEA, to demonstrate the developed
292 approach (Section 3-5). In particular, the approach is implemented to ag-
293 gregate Ozone observations collected in Germany to mean and maximum
294 values in temporal intervals and to averages of these temporal aggregates
295 in space. Figure 5 shows the OpenLayers client visualizing observations be-
296 fore the aggregation (left side) and after the aggregation (right side). The
297 red bar behind the aggregated value is showing the confidence interval calcu-
298 lated from the standard deviation of the values. The statistics are encoded in
299 O&M using the Uncertainty Markup Language (UncertML) (Williams et al.,
300 2009).

301 The service deployment for the case study is shown in Figure 6. One SOS
302 instance serves approximately 30 million observations extracted from the
303 AirBase database files of the European Environmental Agency (EEA, 2011).
304 Administrative areas of Germany are provided as vector-based features by
305 a WFS. The spatio-temporal aggregation is implemented in a prototype of
306 the STAS. Currently, it provides means to aggregate points to polygons in
307 space and to partition temporal extents of observations into time intervals.
308 Supported aggregation functions for space and time are MIN, MAX, SUM
309 and MEAN. Another SOS instance provides the aggregated observations. As
310 a client, the OpenLayers SOS client (Eijnden, 2010) has been extended and
311 a simple user interface has been developed for sending requests to SOS and
312 WPS.

313 Figure 7 shows the workflow of an exemplary spatio-temporal aggregation
314 of ozone observations. As an aggregation process is executed on-demand, a
315 client sends an Execute request to the STAS. Then, the STAS retrieves the
316 observations from a SOS instance and, in parallel, the administrative areas
317 from a WFS instance. Now, the observations are grouped temporally, then
318 aggregated, and afterwards these temporal aggregates are grouped and aggre-
319 gated spatially. The aggregated observations are then inserted into another
320 SOS instance and the reference to the aggregated observations is returned
321 by the STAS to the client. The client can now retrieve the aggregated ob-
322 servations from a SOS. As an option, the aggregated observations can also
323 be directly contained in the aggregation service response. The aggregation
324 functions applied can be easily exchanged by invoking another aggregation
325 process.

326 Provenance information is provided directly in the aggregated observa-
327 tions. The `isAggregateOf` element points to the original ozone observations
328 provided by SOSs. Retrieving this observations allows to retrieve similar
329 provenance information like described by Patni et al. (2010) for Linked Data:
330 the location and time of the observation as well as the description of the
331 measurement procedure that has produced the observation result provided
332 as SensorML description. The procedure element of the aggregated observa-
333 tions points to the SensorML description of the aggregation process. Thus,
334 information about the applied aggregation function and predicates can be
335 retrieved. We do not yet include information about the users of the aggre-
336 gated or source observations like described in Park and Heidemann (2008)
337 and consider this to be future work.

338 7. Discussion & Conclusions

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

358 Our case study shows that the STAS allows for a flexible integration of
359 aggregation processes within SOSs and, thus, in the Sensor Web and Geo-
360 processing Web. Aggregation methods can be exchanged in a flexible way,
361 while the common service operations remain the same for both, the execution
362 of the aggregation as well the access to the aggregated observations. This

363 eases the integration in clients and other applications. Other approaches
364 (Liu et al., 2008; Kabadayi et al., 2006; Patni et al., 2010) usually focus on
365 aggregating raw measurements provided in a proprietary format and then
366 providing the aggregates in standardized formats. Furthermore, the STAS
367 aggregates a set of source observations to a set of aggregated observations
368 whereas the other approaches usually focus on producing a single aggregate
369 out of several observations.

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

379 As the STAS is implemented as a WPS profile, it allows for an on-demand
380 aggregation. Most approaches for aggregating observations through virtual
381 sensors use predefined aggregation processes and the execution time of the
382 aggregation is predefined as well, see for example Liu et al. (2008). However,
383 the STAS might be used in these approaches to provide the aggregation pro-
384 cessing, for example the estimation of values for a new virtual sensor where
385 no physical sensors are available. Encapsulating the aggregation processes
386 through an WPS interface allows to integrate aggregation in Geoprocessing
387 workflows. The STAS is able to mediate between different observation and

388 model services providing or requiring the data in different resolutions. In or-
389 der to realize complex aggregation workflows that require several aggregation
390 steps, a WPS-T interface as proposed by Schaeffer (2008) might be used. It
391 allows to create complex workflows out of the basic aggregation processes
392 which then can be published as complex WPS processes again. This enables
393 the STAS to provide complex aggregation chains in the same manner as the
394 simple aggregation processes. While we propose a model for aggregation of
395 observation, we did not yet provide disaggregation methods like e.g. dis-
396 aggregating grid cells to point observations. Their integration needs to be
397 explored in future work.

398 While the STAS allows for a flexible and seamless integration of aggre-
399 gation processes in the Sensor Web, the communication overhead for trans-
400 ferring large datasets between the services has been identified as a possible
401 drawback of the approach. In that case, one approach might be to tightly
402 couple the STAS with a SOS instance and thus reduce the communication
403 overhead: Instead of passing the URLs of the SOS instances, the observa-
404 tions can automatically be fetched from a SOS running on the same ma-
405 chine. Also, the aggregated observations can directly be inserted in this SOS
406 instance again. Another approach to cope with large datasets might be effi-
407 cient caching and indexing strategies as described by Sivasubramanian et al.
408 (2007).

409 To enable the discovery of already deployed aggregation processes in the
410 Web and to automate observation aggregation workflows, semantic descrip-
411 tions of observations and aggregation processes are needed. In a next step, we
412 plan to semantically enable the STAS by using already existing approaches

413 for integrating semantics in web-based spatial data infrastructures (Janow-
414 icz et al., 2010) and by utilizing observation ontologies like defined by the
415 W3C (Janowicz and Compton, 2010). Finally, we consider to extend the con-
416 cept of the STAS from aggregation processes for observations to aggregation
417 processes for general spatio-temporal data in SDIs.

418 **Acknowledgments**

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

423 **References**

424 Botts, M., Percivall, G., Reed, C., Davidson, J. (Eds.), 2007. OGC Sen-
425 sor Web Enablement: Overview And High Level Architecture. OGC 07-
426 165. http://portal.opengeospatial.org/files/?artifact_id=25562 [[accessed
427 20 March 2011].

428 Botts, M., Robin, A. (Eds.), 2007. OpenGIS Sensor
429 Model Language (SensorML) Implementation Specification.
430 OGC 07-000. Open Geospatial Consortium, Inc., 180 pp,
431 http://portal.opengeospatial.org/files/?artifact_id=21273, [accessed
432 22 March, 2011].

433 Brauner, J., Foerster, T., Schaeffer, B., Baranski, B., 2009. Towards a Re-

434 search Agenda for Geoprocessing Services, in: 12th AGILE International
435 Conference on Geographic Information Science.

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

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

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

446 Cox, S. (Ed.), 2007a. Observation and Measurements - Part 1 - Obser-
447 vation Schema. OGC 07-022r1. Open Geospatial Consortium, Inc., 86
448 pp, http://portal.opengeospatial.org/files/?artifact_id=22466, [accessed
449 22 March, 2011].

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

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

- 457 Eijnden, B.v. (Ed.), 2010. OpenLayers: SOS and Inspire. Bart van Eijnden.
458 Presentation at FOSS4G. <http://2010.foss4g.org/presentations/2891.pdf>
459 [accessed 22 March, 2011].
- 460 Falke, S., Dorner, E., Dunn, B., Sullivan, D., 2008. Processing Services
461 in Earth Observation Sensor Web Information Architectures, in: Earth
462 Science Technology Conference, University of Maryland, June 24-26.
- 463 Foerster, T., 2010. Web-based Architecture for On-demand Maps -
464 Integrating Meaningful Generalization Processing. Ph.D. thesis. University
465 of Twente.
- 466 Hartig, O., Zhao, J., 2010. Publishing and consuming provenance meta-
467 data on the web of linked data, in: In: Proc. of 3rd Int. Provenance and
468 Annotation Workshop.
- 469 Havlik, D., Bleier, T., Schimak, G., 2009. Sharing Sensor Data with SensorSA
470 and Cascading Sensor Observation Service. *Sensors* 9, 5493–5502.
- 471 Janowicz, K., Compton, M., 2010. The Stimulus-Sensor-Observation Ontol-
472 ogy Design Pattern and its Integration into the Semantic Sensor Network
473 Ontology, in: Kerry Taylor, Arun Ayyagari, D.D.R. (Ed.), Proceedings of
474 the 3rd International workshop on Semantic Sensor Networks 2010 (SSN10)
475 in conjunction with the 9th International Semantic Web Conference (ISWC
476 2010), CEUR-WS.
- 477 Janowicz, K., Schade, S., Broering, A., Kessler, C., Maue, P., Stasch, C.,
478 2010. Semantic Enablement for Spatial Data Infrastructures. *Transactions*
479 in GIS 14, 111–129.

- 480 Jeong, S.H., Fernandes, A.A.A., Paton, N.W., Griffiths, T., 2004. A Generic
481 Algorithmic Framework for Aggregation of Spatio-Temporal Data, in: SS-
482 DBM '04: Proceedings of the 16th International Conference on Scientific
483 and Statistical Database Management, IEEE Computer Society, Washing-
484 ton, DC, USA. p. 245.
- 485 Kabadayi, S., Pridgen, A., Julien, C., 2006. Virtual Sensors: Abstracting
486 Data from Physical Sensors, in: 4th International Workshop on Mobile
487 Distributed Computing.
- 488 Liu, Y., Abdelzaher, T., Heo, J., Choi, J., Minsker, B., Fazio, D., 2008.
489 Virtual sensor-powered spatiotemporal aggregation and transformation: A
490 case study analyzing near-real-time nexrad and precipitation gage data in a
491 digital watershed, in: Environmental Information Management Conference
492 2008, September 10 - 11.
- 493 Liu, Y., Futrelle, J., Myers, J., Rodriguez, A., Kooper, R., 2010. A
494 provenance-aware virtual sensor system using the open provenance model,
495 in: Collaborative Technologies and Systems (CTS), 2010 International
496 Symposium on, pp. 330 –339.
- 497 Madden, S., Franklin, M.J., Hellerstein, J.M., Hong, W., 2002. TAG: a Tiny
498 AGgregation service for ad-hoc sensor networks. SIGOPS Oper. Syst. Rev.
499 36, 131–146.
- 500 Na, A., Priest, M. (Eds.), 2007. Sensor Observation Service. OGC 06-009r6.
501 Open Geospatial Consortium, Inc., 104pp. [accessed 22 March 2011].

- 502 Nittel, S., Stefanidis, A., 2005. GeoSensors Networks. CRC. chapter GeoSen-
503 sor Networks and Virtual GeoReality.
- 504 Park, U., Heidemann, J., 2008. Provenance in sensornet republishing, in:
505 Proceedings of the 2nd International Provenance and Annotation Work-
506 shop, Springer-Verlag, Salt Lake City, Utah, USA. pp. 208–292.
- 507 Patni, P., Sahoo, S., Henson, C., Sheth, A., 2010. Provenance aware linked
508 sensor data, in: Kärger, P., Olmedilla, D., Passant, A., Polleres, A. (Eds.),
509 Proceedings of the Second Workshop on Trust and Privacy on the Social
510 and Semantic Web.
- 511 Pebesma, E., Cornford, D., Dubois, G., Heuvelink, G.B., Hristopulos, D.,
512 Pilz, J., Sthlker, U., Morin, G., Skien, J.O., 2011. INTAMAP: The de-
513 sign and implementation of an interoperable automated interpolation web
514 service. *Computers & Geosciences* 37, 343 – 352.
- 515 Schaeffer, B., 2008. Towards a Transactional Web Processing Service, in:
516 Pebesma, E., Bishr, M., Bartoscheck, T. (Eds.), 6th Geographic Informa-
517 tion Days, Ifgi Prints. pp. 91–116.
- 518 Schut, P. (Ed.), 2007. OpenGIS Web Processing Service.
519 OGC 05-007r7. Open Geospatial Consortium, Inc., 87pp.
520 http://portal.opengeospatial.org/files/?artifact_id=24151 [accessed
521 22 March, 2011].
- 522 Sivasubramanian, S., Pierre, G., Steen, M., Alonso, G., 2007. Analysis of
523 caching and replication strategies for Web applications. *IEEE Internet*
524 *Computing* 11, 60–66.

- 525 Vega Lopez, I.F., Snodgrass, R.T., Moon, B., 2005. Spatiotemporal Aggre-
526 gate Computation: A Survey. *IEEE Trans. on Knowl. and Data Eng.* 17,
527 271–286.
- 528 WetterOnline (Ed.), 2011. Wetter WetterOnline Wettervorhersage Aktuelles
529 Wetter Regenradar Reisewetter <http://www.wetteronline.de/>, [accessed 22
530 March, 2011].
- 531 Whiteside, A. (Ed.), 2009. Definition identifier URNs in OGC names-
532 pace. OGC 07-092r3. Open Geospatial Consortium, Inc., 46pp.
533 http://portal.opengeospatial.org/files/?artifact_id=30575 [accessed
534 22 March, 2011].
- 535 Williams, M., Cornford, D., Bastin, L., Pebesma, E. (Eds.), 2009. Uncer-
536 tainty Markup Language. OGC 08-122r2. Open Geospatial Consortium,
537 Inc., 61pp. http://portal.opengeospatial.org/files/?artifact_id=33234 [ac-
538 cessed 22 March, 2011].
- 539 Worboys, M., 1994. A Unified Model for Time and Space. *The Computer*
540 *Journal* 37, 26–34.
- 541 Zhao, P., Di, L., 2010. *Geospatial Web Services: Advances in Information*
542 *Interoperability*. Information Science Publishing.

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.

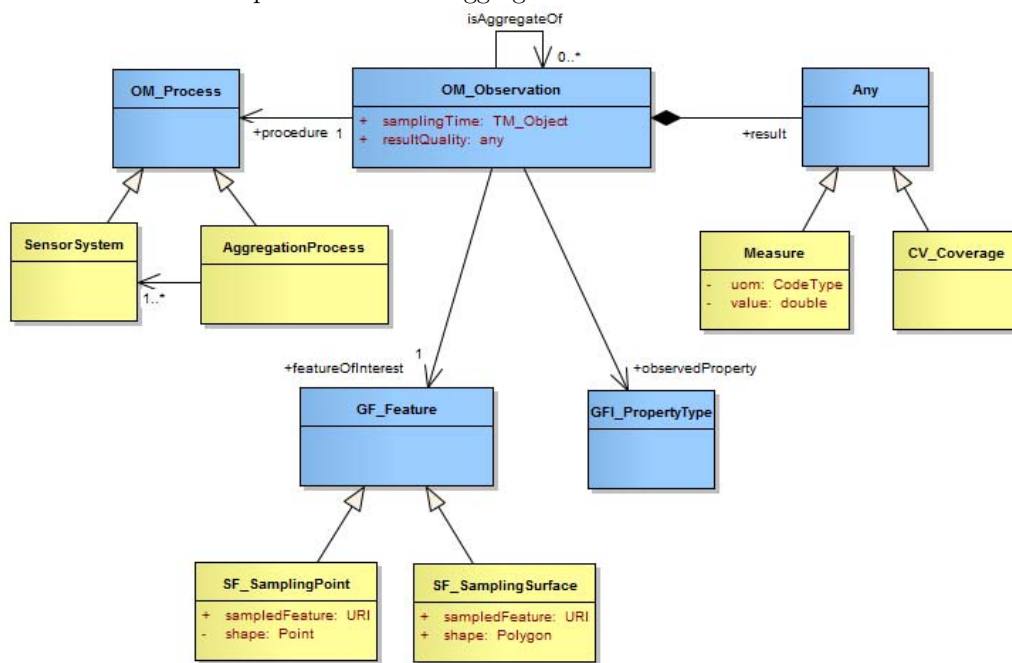


Figure 2: UML diagram of the model for aggregation processes. An aggregation process has a set of observations as inputs and outputs. It is defined by the parameters containing the grouping predicates and the aggregation functions.

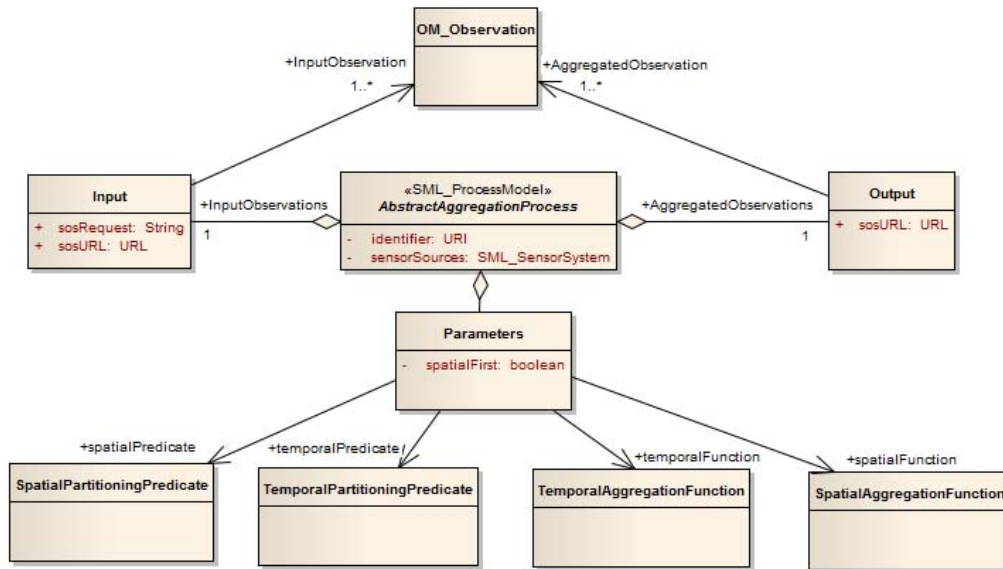


Figure 3: UML diagram showing two subtypes of spatial and temporal partitioning predicates. The PolygonContainment predicate is defined to group points to polygons that are provided by a WFS. The TemporalGridding predicate is used for partitioning a temporal extent in time intervals of equal duration.

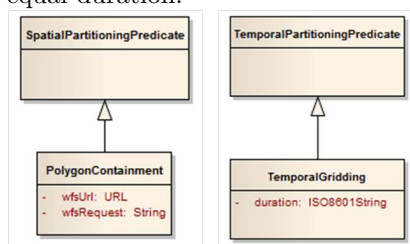


Figure 4: Basic service architecture for spatio-temporal aggregation in the Sensor Web. The STAS queries non-aggregated observations from a SOS instance and inserts the aggregated observations in another transactional SOS instance.

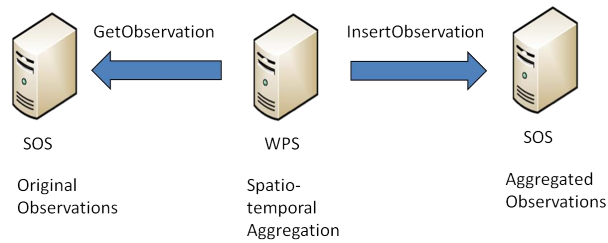


Figure 5: Screenshots of the OpenLayers client visualizing air quality observations before (left) and after (right) the aggregation.

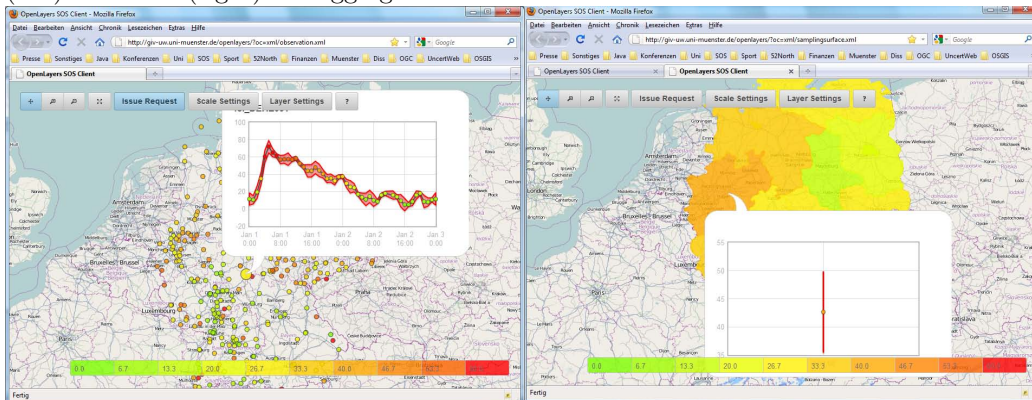


Figure 6: Service deployment for spatio-temporal aggregation of European air quality observations. The architecture consists of one SOS instance offering the basic air quality observations, a WPS instance implementing the STAS, a SOS instance that provides the aggregated observations, and a WFS providing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

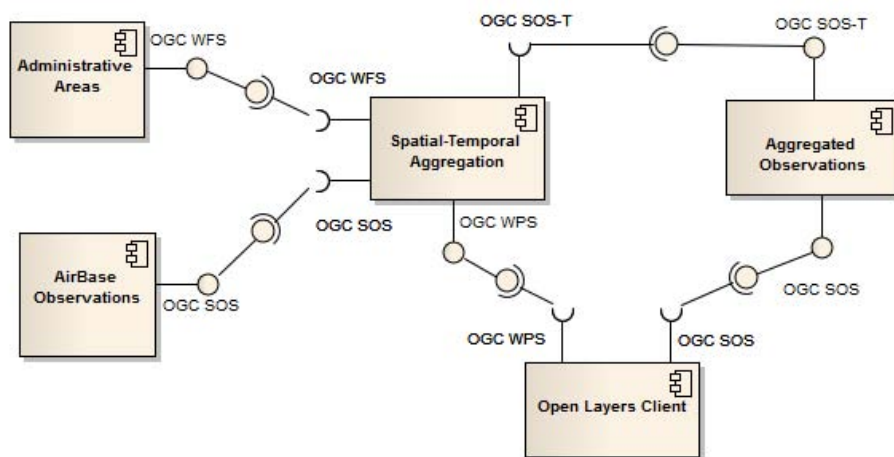


Figure 7: Workflow of spatio-temporal aggregation of European air quality observations.

