Conception and Implementation of an OGC-compliant Sensor Observation Service for a Standardized Access to Raster Data

Juergen Sorg and Ralf Kunkel
Research Center Juelich - Institute of Bio- and Geosciences (IBG-3), Leo Brandt Strasse, Juelich, Germany
{j.sorg, r.kunkel}@fz-juelich.de

Keywords: TERENO, Sensor Observation Service, SOS, SWE, OGC, Raster Data.

Abstract: The target of the Open Geospatial Consortium (OGC) is interoperability of Geographic Information Systems (GIS), which means creating opportunities to access geodata in a consistent, standardized way. In the domain of sensor data, the target will be picked up within the OGC Sensor Web Enablement Initiative and especially reached through the Sensor Observation Service (SOS) specification. This one defines a service for a standardized access to time series data and is usually used for in-situ sensors (like discharge gauges, climate stations). Although the specification considers raster data, no implementation of the standard for raster data exists presently. In this paper an OGC-compliant Sensor Observation Service for a standardized access to raster data is described. A data model was developed, which enables an effective storage of the raster data with the corresponding metadata in a database, reading this data in an efficient way and encoding it with result formats that the SOS-specification provides.

1 INTRODUCTION

Long term changes of temperature, precipitation and other climate parameters have direct and indirect influences to the terrestrial systems soil, air and water and result in social, economical and political effects to the society.

To understand and predict these interconnected and continuously evolving processes of the earth system integrated models to quantify theses effects need to be developed. However, these models require the observation and analysis of long-term data sets from different scientific topics, e.g. from physics, chemistry, meteorology, geology or anthropology. In this context the development and implementation of Spatial Data Infrastructures (SDI) for terrestrial research is gaining importance.

TERENO (Terrestrial Environment Observatories) is an interdisciplinary long term research project initialized by the Helmholtz association, which gathers long term ecological, social and economical results of global change on a regionally scale (Zacharias, Bogena et al. 2011). Four terrestrial observatories are established, which are coordinated by five Helmholtz research centers. Installation of equipment has started in 2007 and will be finished in 2013. Data collection, however, is planned to be performed for at least 30 years. Within these observatories sensor networks for intensive measurement of soil moisture, soil temperature, water gauge as well energy and fluid fluxes are deployed. In addition, four weather radar and rain scanner devices to remotely measure precipitation rates quantities have been installed.

Observed data from each observatory are stored and published in decentral infrastructures, each operated by the individual centers responsible for an observatory. The data portal TEODOOR (TEReno Online Data repOsitORy: http://www.tereno.net) provides access to the data for scientists and stakeholders and allows to search in metadata catalogs, visualize the data as well as to download it. Heterogeneity of observed data, but also different measurement technics and usage of different database systems require the application of metadata standards to describe data and measurement technics as well as the access to distributed data and metadata using standardized interfaces (Botts, Percivall et al. 2008). The Open Geospatial Consortium (OGC) defines these standards and provides specifications of interfaces for an interoperable access to geo data, which services have to implement.

In-Situ measurement stations for observing physical phenomenons (e.g. temperature, soil moisture, etc.) always related to a single geographic point. In contrast, remote sensing stations deliver
area differentiated data, related to a certain geographic area. The common approach for standardized access to raster data is the OGC Web Coverage Service (WCS) specification (OGC 2003, OGC 2010b). However, a WCS cannot be used for time series, since the datasets stored within a WCS have no temporal relation. Therefore, the OGC (OGC 2011b) defines a specialization of the WCS specification for earth observations (WCS-EO), which requires to describe each dataset by an additional metadata set. This additional metadata set allows to identify the temporal relation for each raster data set and enables on the other hand the possibility to subsume datasets to identifiable and queryable sets (Dataset Series). A description of each data set can be retrieved by the new operation (DescribeEOCoverageSet), whereas datasets itself can be selected by temporal filters. In a second step data are requested with the GetCoverage operation using the unique identifiers.

A new approach to manage raster time series data supports the OGC sensor observation service (SOS) specification. It comprises methods for a standardized access to all kinds of time series data with spatial relation to the earth. The advantage of using a SOS instead of a WCS-EO is the inherence of the temporal relation of each dataset, since a SOS is particularly designed for managing time series. Temporal selection of datasets is performed in a direct way rather than using an additional operation with an afterwards extraction of the required identifiers. Furthermore, a SOS supports to apply thematic filters to extract thematic attributes of raster data sets. In this paper we describe the conception and implementation of an OGC compliant SOS for a standardized access to raster time series data, which allows to select raster data sets using temporal, spatial and thematic filters and to deliver it in a standardized way. In addition, a solution is described, which prepares the data for a fast perfunctory verification in a time critical visualization on a web page and gives a standardized access.

2 OGC SENSOR OBSERVATION SERVICE

The Open Geospatial Consortium (OGC) is a confederation of leading GIS manufacturer, data producers, authorities, organizations, research facilities and universities. It was founded in 1994 and since it develops metadata standards and standardized interfaces for an interoperable access to geographic data. The Interoperability of sensor and sensor networks are covered by the Web Enablement Initiative of the OGC (Bröring, Echterhoff et al. 2011). Within these initiative there are several to each other aligned standards and interfaces defined:

- Sensor Observation Service (SOS) specification defines a standardized web service to search, filter and retrieval sensor data and sensor information (OGC 2006a, OGC 2012a).
- Observation & Measurement is a standard to describe and deal with observational data (OGC 2007a, OGC 2007b, OGC 2010a, OGC 2011a). It is majorly used to encode geo data, which a SOS delivers
- Sensor Model Language (SensorML) is a information model to describe sensors and the observed properties (OGC 2007d).
- Sensor Alert Service (SAS) gives users over a standardized interface the possibility to receive alert messages from sensors (OGC 2007c).

Presently there is a discussion about extending the SAS to a Sensor Event Service (SES), which is able to handle all kinds of event messages (OGC 2008).

Beside these standards the SWE also comprises the standards Transducer Model Language (TML), Sensor Planing Service (SPS) and Web Notification Service (WNS), which are not within the scope of the paper.

The SOS specification comprises eleven operations to access observation data, but only the operations GetCapabilities, DescribeSensor and GetObservation are mandatory. The GetCapabilities Operation yields general information about the service and all information necessary to call the supported operations. The DescribeSensor Operation yields a description of a sensor encoded by the SensorML language (OGC 2007d) and contains among others identifiers for the observed properties, coordinates of the station and a time domain for which data is available. Finally, data can be requested by means of a GetObservation operation. The XML fragment within Example 1 displays an example of a GetObservation request with all available parameters (OGC 2006a).

```xml
<GetObservation>
  <offering>Reflectivity</offering>
  <eventTime>
    <o:TM_During>
      <o:PropertyName>
        om:samplingTime
      </o:PropertyName>
      <gml:TimePeriod>
        <gml:beginPosition>
          2014-01-01T00:00:00.000Z
        </gml:beginPosition>
        <gml:endPosition>
          2014-12-31T23:59:59.999Z
        </gml:endPosition>
      </gml:TimePeriod>
    </o:TM_During>
  </eventTime>
</GetObservation>
```
Example 1: GetObservation request with all available parameters.

The unique and mandatory Offering identifies a free selectable, use case dependent logical combination of sensors (e.g. by observed properties) (OGC 2006a). The parameters for temporal, spatial and thematic filters (eventTime Line 3-17, featureOfInterest Line 22-34 and result Line 35-42) and the unique identifiers of the station (procedure Line 18) and the observed properties (observedProperty Line 19) are optional. Figure 1 shows some examples of spatial filters that are supported by the SOS specification. On the left side a line segment filter is depicted, which extracts all values located on a street course. In the center a rectangle filter and on the right hand side a not orthogonal polygonal filter is depicted. Both filters selecting objects located within the filter face.

Beside the spatial filters, the SOS specification also provides thematic compare filters, which selects objects not by their location but by means of values of a variable. In lines 35-42 of Example 1 a thematic filter is specified, which selects all meshes of a raster dataset, that attribute “Reflectivity” has a value greater than 36.

The SOS processes requests under consideration of the specified filter conditions and returns the results as O&M documents (OGC 2007a, OGC 2007b, OGC 2010a, OGC 2011a). Thereby the O&M standard facilitates to encode vector data as well raster data by use of XML documents. Crucial thereby is, that despite of the very different structure the O&M standard provides both, the encoding of raster and vector data (OGC 2007b). On the one hand the generalized coverage model of ISO19123 is used to encode raster data. On the other hand vector data are covered with several for this purpose specifically developed models. Within a GetObservation request the desired result model is specified by the Parameter resultModel (line 46 in example 1) (OGC 2006a). Therefore, it is possible to request raster as well as vector data from a SOS in a standardized way. But certainly, raster data are not supported from the common SOS implementations.

3 IMPLEMENTATION

There exist a huge number of implementations of the SOS standard in any kind of programming languages (Nengcheng, Liping et al. 2009). Because the SOS implementation of the 52°North company is easily modifiable it was used to be extended to give access to raster data in a standardized way. Following steps were necessary to achieve this, which are explained in following sections:

- A data model, based on the 52°North SOS data model, to store the raster data and its describing metadata in an efficient way was developed (see Figure 2)
- Efficient algorithms to apply filters were implemented.
- Several O&M models to return the raster data in a standardized and flexible manner were realized.
3.1 Data Model

The raster data are stored in a relational database management system (RDBMS), whereas PostgreSQL (Pfeiffer and Wenk 2010) with the PostGIS (Obe and Hsu 2011) extension for spatial data are used. Although the new PostGIS version (2.0) supports raster data, it cannot be used here, because PostGIS intends to use one table for each raster data set (Holl 2012). Within the TERENO project a huge number of raster datasets (up to 1.5 Mio. weather radar data sets within the project period) will be created, making the usage of PostGIS raster data tables not practicable. Therefore, a data model was developed, which supports to store all raster data in one single data table. This can be accomplished in two different ways: as a binary large object (BLOB) or row based. A BLOB is a database type for large, not nearer specified binary objects. Because there is no additional information about the stored object in a BLOB, the amount of space is minimized on the field of table layer. Certainly storage in a BLOB is inconvenient for a thematic filtering of the data, because in this case the entire binary object must be read and for each raster mash the filter condition must be proofed. Keeping each raster mesh in a row of a data table has sure the advantage that searching can be done in an efficient manner, but the considerable disadvantage is the large amount of inserts and rows that accrue for each raster dataset (e.g. for a relative small raster with 800x800 pixels 640000 inserts and rows must be handled).

To achieve an efficient thematic search within the raster data a new coarser grid was used, which is a compromise between a small storage space and an efficient access. In this manner database indices can be used to have a more efficient access then it can be achieved with a sequential method (Kemper and Eickler 2006). The coarser grid is in its resolution free selectable and keeps for each of its raster meshes the maximum and minimum of the underlying original raster meshes. The resolution is selected in this way, as on one hand inserts can be done in an adequate time frame and on the other hand not too much original pixel have to be read and proofed.

Figure 2 shows the developed data model. The time series describing meta data is distributed stored over the normalized tables `coverage_structure`, `phenomenons`, `stations`, `offerings` and `coverage_out_of_band`. Moreover the coarser grid is also normalized in the tables `coverage_coarse_raster` and `coverage_coarse_geometry`. And finally the raster data itself is stored as a BLOB field in the table `weatherradar_coverage` with a foreign key link to its time series in table `coverage_structure`.

3.2 Filter

The existing filters of the 52°North implementation are coupled to the database intrinsic functions, they only support vector based requests. However, for our developed SOS for raster data, it is required to have filter operations, which can be applied also to raster data. Our implementation supports thematic as well as spatial filters. The former selects raster cells by means of values of a variable. This is realized with the aforementioned coarse grid, which has the advantage, as only relevant pieces of the original dataset must be read. Because the selected pixel within the coarse raster always represents an orthogonal polygon, an efficient sweep algorithm (Güting and Dieker 2004) was implemented to read these orthogonal faces. Figure 3 shows a clip of such an orthogonal polygon. The sweep algorithm scans
Figure 3: Sweep algorithm to scan orthogonal polygons.

the polygon in vertical direction from top to bottom and holds line segments in a temporary cache, which mark domains that can be read in blocks from the original raster dataset. Only for these pixels the filter condition must be proofed and if they pass then the raster mesh is appended to the result set.

The sweep algorithm is also usable for spatial filters, which are dealing directly with the original raster dataset (not over the coarse grid). This is only possible, if the spatial filter is orthogonal (see Figure 1). If the spatial filter is not orthogonal, the sweep algorithm is not usable. In this case a vectorized representation of the raster is used, which holds for each raster mesh from the original raster the geometry and the index of the cell. With this vectorized raster it is possible to use the PostGIS intrinsic functions like intersects, contains, etc. (Obe and Hsu 2011). Moreover, our implementation supports any spatial filter geometry, which is encodeable with GML within the GetObservation request. This is because the GML encoded filter can easily transformed to the WKT (Well Known Text) format (OGC 2011c), which is the encoding format for filter geometries of the aforementioned PostGIS functions.

3.3 Result Models

Two different kinds of methods were implemented to return the result sets. First there are used O&M models for the direct encoding of the raster data. These are the DiscreteCoverageObservation, the TimeSeriesObservation and a generic O&M model (OGC 2007b) (Broering and Meyer 2008), which are used to encode spatial and thematic filtered data. Second, there are entire raster datasets provided as georeferenced images over an external OGC compliant Web Map Service.

For the direct transfer, specified by the parameter INLINE of the GetObservation request (see example 1), two different models were implemented. Both models are included from the ISO19123 standard (ISO 2005). The DiscreteCoverageObservation O&M model can be used for the encoding of spatial and thematic filtered data, while the TimeSeriesObservation O&M model is suitable for the encoding of a time series of a selected raster mesh. In this case the requests use a point filter. Example 2 shows the geometry value pair of a raster mesh encoded with the DiscreteCoverageObservation model.

```
<ns:element>
  <ns:CV_GeometryValuePair>
    <ns:geometry>
      <ns:spatialElement>
        <gml:Polygon srsName="EPSG::31466"
          xsi:type="gml:PolygonType">
          <gml:exterior>
            <gml:LinearRing
              xsi:type="gml:LinearRingType">
              <gml:coordinates>
                5640414.72905 2468620.92429,
                5640664.72905 2468620.92429,
                5640664.72905 2468870.92429,
                5640414.72905 2468870.92429,
                5640414.72905 2468620.92429
              </gml:coordinates>
            </gml:LinearRing>
          </gml:exterior>
        </gml:Polygon>
      </ns:spatialElement>
    </ns:geometry>
    <ns:value
      uom="dbz"
      xsi:type="gml:MeasureType">
      8.944881889763703
    </ns:value>
  </ns:CV_GeometryValuePair>
</ns:element>
```

Example 2: Geometry value pair of a raster mesh encoded with the DiscreteCoverageObservation-model.

The geometry is encoded as a polygon in the Geography Markup Language (GML), while the thematic value is encoded as a decimal value with its unit of measure (UOM). In a similar manner the dataset is encoded in the TimeSeriesObservation model, but here the geometry is replaced by a timestamp. Because both models are xml based and very character intensive, a third generic O&M model (OGC 2007b) was used to encode the data with comma separated values. For each raster cell the index or the geometry and the value are given.

For the indirect transfer of data the SOS specification provides the possibility to answer to a GetObservation request with a reference to an external source (OGC 2007b). We use this option to provide the raster data as a georeferenced image.
from an OGC compliant Web Map Service (WMS) (OGC 2006b). References are encoded in a generic O&M model. Within the GetObservation request the response mode is given by the value OUT-OF-BAND for the parameter ResponseMode. Example 3 shows a reference to a raster image provided by a WMS and encoded with a generic O&M model.

<om:member>
  <om:Observation>
    <om:samplingTime>
      <gml:TimeInstant xsi:type="gml:TimeInstantType">
        <gml:timePosition>2012-10-10T06:00:00.000+02:00</gml:timePosition>
      </gml:TimeInstant>
    </om:samplingTime>
    <om:procedure xlink:href="WRradar"/>
    <om:observedProperty xlink:href="Reflectivity"/>
    <om:featureOfInterest>
      <gml:GridCoverage gml:id="WRadar">
        <gml:name>Weatherradar</gml:name>
        <gml:boundedBy>
          <gml:Envelope>
            <gml:lowerCorner srsName="EPSG::31466">
              5543664.729 2432120.924
            </gml:lowerCorner>
            <gml:upperCorner srsName="EPSG::31466">
              5743664.729 2632120.924
            </gml:upperCorner>
          </gml:Envelope>
        </gml:boundedBy>
        <gml:gridDomain/>
        <gml:rangeSet>
          <gml:ValueArray/>
        </gml:rangeSet>
      </gml:GridCoverage>
    </om:featureOfInterest>
    </om:result>
  </om:Observation>
</om:member>

Example 3: Reference to a raster image provided by a WMS and encoded with a generic O&M model.

4 CONCLUSION AND OUTLOOK

In this paper we explained the concept and an implementation of an OGC compliant SOS for an interoperable access to area related raster time series. To give applications the opportunity to request entire raster data sets for a visualization of the data, a system was implemented, that uses the raster SOS to select the desired datasets and an OGC WMS to provide the datasets as georeferenced raster images. The selection of the datasets is efficient, because database indices can be used. The interoperability and the use of standard web software like OpenLayers, but also the retrieval of the datasets in an acceptable time slot is given by using an OGC compliant WMS. A data model to store raster time series data in an efficient manner but also software components for an efficient access to this raster data was implemented. The implemented algorithm allows the retrieval of raster datasets by thematic, spatial and temporal filters. With the algorithms extracted data is encoded with different O&M-models, to support suitable models for different use cases.

Because of the verbosity of O&M coverage models, in particular of the DiscreteCoverage model, this models are not feasible to deliver raster data in a reasonable time. Therefore a generic O&M model with a comma separated encoding of the data was implemented.

The approach to use a SOS to manage raster data has several advantages mentioned already in the introduction. But to deliver data in a well known scientific owned data format like NetCDF is an advantage the WCS approach has. To adapt this a mapping between O&M and NetCDF must be defined. In (OGC 2012b) this is done for WaterML, which is a profile to the O&M Standard for hydrological data. In further work this can also accomplished for an O&M coverage model.

The implemented OUT-OF-BAND method is not limited for external resource hosted on a WMS. It is also usable in conjunction with a WCS. An interesting question for this is how the filters defined in a GetObservation request can be transformed to filters in a WCS GetCoverage request.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support by TERENO (Terrestrial Environmental Observatories) funded by the Helmholtz Association.

REFERENCES

Visualisierung hydrologischer Zeitreihen mit Hilfe
Symposium und Fachmesse., Salzburg, Österreich.
Bröning, A., J. Echterhoff, S. Jirka, I. Simonis, T.
Everding, C. Stasch, S. Liang and R. Lemmens
Algorithmen, Teubner Verlag.
2012, Dessau-Rosslau, FOSSGIS.
Funktionsschema (DIN EN ISO 19123), DIN
Deutsches Institut für Normung e.V.: 71.
Kemper, A. and A. Eickler (2006). Datenbanksysteme -
eine Einführung, Oldenbourg Wissenschaftsverlag
GmbH.
"A flexible geospatial sensor observation service for
diverse sensor data based on Web service." ISPRS
Journal of Photogrammetry and Remote Sensing
Obe, R. O. and L. S. Hsu (2011). PostGIS in Action,
Manning Publications.
OGC (2003). Web Coverage Service (WCS) Wayland,
MA, USA, Open Geospatial Consortium Inc.: 67.
Implementation Specification. Wayland, MA, USA,
Open Geospatial Consortium Inc.: 91.
OGC (2006b). OpenGIS® Web Map Server
Implementation Specification, Open Geospatial
Consortium Inc.
OGC (2007a). Observations and Measurements – Part 2 -
Sampling Features, Open Geospatial Consortium Inc.: 36.
Observation schema Wayland, MA, USA, Open
Geospatial Consortium Inc.: 85.
OGC (2007c). OGC® Sensor Alert Service
Implementation Specification, Open Geospatial
Consortium Inc: 128.
(SensorML). Implementation Specification. Wayland,
MA, USA, Open Geospatial Consortium. 1.0.0: 180.
Specification (proposed), Open Geospatial Consortium
Inc.: 88.
OGC (2010a). Geographic Information: Observations and
Measurements. OGC Abstract Specification Topic 20,
Open Geospatial Consortium Inc. 2.0.0: 57.
OGC (2010b). OGC® WCS 2.0 Interface Standard - Core,
Open Geospatial Consortium Inc.: 45.
OGC (2011a). Observations and Measurements. XML
Implementation. Wayland, MA, USA, Open
Geospatial Consortium. 2.0: 76.
OGC (2011b). OGC® Web Coverage Service 2.0
Interface Standard - Earth Observation Application
Profile, Open GeoSpatial Consortium Inc.: 48.
OGC (2011c). OpenGIS® Implementation Standard for
Geographic information - Simple feature access - Part
I: Common architecture, Open Geospatial
Consortium, Inc.: 92.
OGC (2012a). OGC® Sensor Observation Service
OGC (2012b). WaterML 2.0 – Timeseries – NetCDF
Pfeiffer, T. and A. Wenk (2010). PostgreSQL : das
Praxisbuch. Bonn, Galileo-Press.
Zacharias, S., H. Bogena, L. Samaniego, M. Mauder, R.
Fuß, T. Pütz, M. Frenzel, M. Schwank, C. Baessler, K.
Butterbach-Bahl, O. Bens, E. Borg, A. Brauer, P.
Dietrich, I. Hajnsek, G. Helle, R. Kiese, H.
Kunstmann, S. Klozas, J. C. Munch, H. Papen, E.
Priesack, H. P. Schmidt, R. Steinbrecher, U.
Rosenbaum, G. Teutsch and H. Vereecken (2011). "A
Network of Terrestrial Environmental Observatories
in Germany." Vadose Zone Journal 10(3): 955-973.