Open Geospatial Consortium

Submission Date: 2014-03-05

Approval Date: 2014-03-30

Publication Date: 2014-04-15

External identifier of this OGC® document: http://www.opengis.net/doc/DP/SOS-hydrology-profile

Internal reference number of this OGC® document: 14-004

Category: OGC® Discussion Paper

Editor: GEOWOW Consortium

OGC Sensor Observation Service 2.0 Hydrology Profile

Warning

This document is not an OGC Standard. This document is an OGC Discussion Paper and is therefore not an official position of the OGC membership. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an OGC Standard. Further, an OGC Discussion Paper should not be referenced as required or mandatory technology in procurements.

Document type:	OGC® Discussion Paper
Document subtype:	not applicable
Document stage:	Approved for public release
Document language:	English

License Agreement

Contents

i. Abstract

This document an interoperable hydrology profile for OGC Sensor Observation Service (SOS) 2.0 implementations serving OGC WaterML 2.0. This development is based on previous activities and results (i.e. Hydrology Interoperability Experiments as well as the European FP7 project GEOWOW). It is guided by the need to overcome mostly semantic issues between different SOS instances serving hydrological data and the according clients. Therefore, this profile focuses on how to use the entities and requests of the standards and covers the necessary technical details.

ii. Keywords

The following are keywords to be used by search engines and document catalogues.

Hydrology, Sensor Web Enablement, Sensor Web, Sensor Observation Service, Profile

iii. Preface

Within work package 5 (Water) of the GEOWOW project a central work item is the creation of a fully interoperable hydrology profile for the usage of the OGC Sensor Observation Service (SOS) serving WaterML 2.0 which is described in this document. Following the development of this profile several members of the GEOWOW consortium implement services and clients supporting this project as a proof of concept and realize one or more use cases as demonstration.

The work on this document has been performed within the European GEOWOW (GEOSS interoperability for Weather, Ocean and Water) project. GEOWOW is co-funded by the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 282915 in response to call ENV.2011.4.1.3-1 “Interoperable integration of Shared Earth Observations in the Global Context”.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The Open Geospatial Consortium shall not be held responsible for identifying any or all such patent rights.

Recipients of this document are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of the standard set forth in this document, and to provide supporting documentation.

iv. Submitting organizations

The following organizations submitted this Document to the Open Geospatial Consortium (OGC):

- 52°North Initiative for Geospatial Open Source Software GmbH
- KISTERS AG

v. Submitters

All questions regarding this submission should be directed to the editor or the submitters:

Name	Affiliation
Andres, Volker	**German Federal Institute of Hydrology (BfG) / Global Runoff Data Centre (GRDC)**
Bredel, Henning	**52°North Initiative for Geospatial Open Source Software GmbH**
Fuest, Stefan	**KISTERS AG**
Hollmann, Carsten	**52°North Initiative for Geospatial Open Source Software GmbH**
Jirka, Simon	**52°North Initiative for Geospatial Open Source Software GmbH**
Schlummer, Manuela	**University of Bonn**
Strauch, Adrian	**University of Bonn**
Utech, Michael	**KISTERS AG**

1. Scope

Based on previous projects and experiences by KISTERS and 52°North, the following use cases are seen as common cases regarding data discovery and data download functionalities in hydrological applications. Although these use cases are described from a client side perspective, of course the server side should allow requests in a corresponding way:

- DISCOVERY: List measuring locations by a geographical extent (bounding box).
- DISCOVERY: Provide a list of measuring locations with metadata such as station id, location, name, type and others.
- DISCOVERY: Provide information about available variables/parameters for a certain measuring location.
- DISCOVERY: Provide information about available time series for a certain measuring location and parameter including data coverage information.
- DATA DOWNLOAD: Retrieve data for certain measuring locations and parameters (or for a defined list of time-series) and a defined time-range.
- DATA DOWNLOAD: Harvest data for a set of defined measuring locations and parameters (or for a defined list of time-series) on a regular base.
- DATA DOWNLOAD: Retrieve a 'layer' of single values for certain measuring locations and parameters (or for a defined list of time-series) at a given date.

This document contains an advanced draft of the SOS 2.0 Hydrology Profile. Each definition is described accurately not only in its meaning but also in the way it has been created and what alternatives have been discussed. Where there are open points they are listed at the end of each subchapter.

The next two chapters briefly describe the current state in hydrological standards and the necessity of an interoperable SOS 2.0 profile for hydrology. A listing of relevant existing standards and experiments is provided in addition to the actual profile content. Chapter 6 is about the basic terms, entities, and identification mechanisms in the used standards while chapter 7 treats the actual service commands. Technical aspects for service implementation and usage are covered in chapter 8. The document finishes with an outlook to remaining work and future options.

This document is intended to serve as basis for further discussion within the relevant OGC Working Groups. Taking into account these discussions, the authors intend to advance this Discussion Paper to an OGC Best Practice Paper.

2. References

The following normative documents contain provisions that, through reference in this text, constitute provisions of this document. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. For undated references, the latest edition of the normative document referred to applies.

- ISO TC 211 (2011). ISO 19156:2011 - Geographic information -- Observations and measurements - International Standard. Geneva, Switzerland, International Organization for Standardization.

- Botts, Mike and Alexandre Robin (2007). OGC Implementation Specification: Sensor Model Language (SensorML) 1.0.0 (07-000). Wayland, MA, USA, Open Geospatial Consortium Inc.

- Bröring, Arne, Christoph Stasch and Johannes Echterhoff (2012). OGC Implementation Specification: Sensor Observation Service (SOS) 2.0 (12-006). Wayland, MA, USA, Open Geospatial Consortium Inc.

- Taylor, Peter (2012). OGC Implementation Specification: WaterML 2.0: Part 1 - Timeseries (0-126r3). Wayland, MA, USA, Open Geospatial Consortium Inc.

3. Conventions

Within this document, the following abbreviations are used:

GEOWOW	GEOSS Interoperability for Weather, Ocean and Water
GML	Geography Markup Language
Hydrology DWG	Hydrology Domain Working Group
O&M	Observations & Measurements
OGC	Open Geospatial Consortium
SOAP	Simple Object Access Protocol
SOS	Sensor Observation Service
WFS	Web Feature Service
WML	WaterML 2.0
WSDL	Web Service Description Language

4. Purpose of this Profile

4.1 Technical vs. semantic interoperability

The technical interoperability of a web service is given by its specification. There it is defined e.g. what protocols to use, how to name certain functions and what parameters and entities exist. On the other hand it does not necessarily say much about how to use these parameters and entities. If a specification provides a higher degree of freedom on these points (e.g. the SOS specification) it can easily lead to semantic problems in communication. This is where it becomes necessary to ensure semantic interoperability.

The experience with the existing SOS services shows that there is a range of ways to apply the semantics of the intentionally generic SOS terminology. Discussions have taken place in several groups and it appears that different approaches of SOS implementations were taken in different communities as well as within similar communities. However, different SOS implementations which use different semantics of the SOS concepts cause ambiguous client/server interaction with technically proper but semantically wrong request/response patterns. Interaction between clients and servers from different vendors and/or domains often requires an adjustment process to ensure common semantics and full compatibility.

The SOS approach is an intentional approach. It has been developed to support a broad range of use cases ranging from fixed in-situ sensors to tracking applications or even complex remote sensing systems. Thus, the semantic flexibility is absolutely necessary to accomplish this. It is up to the community of actors in a certain domain to limit this flexibility to get a reasonable set of options (e.g. defining how the terms “feature of interest”, “offering” and “procedure” shall be used to ensure domain dependent functionality and logic).

4.2 Discovery functions and filtering

The general target for discovery in a hydrological time series service is the discovery up to the actual time series level with appropriate listings and filtering options. Previously this turned out to be difficult because the SOS does not feature anything like a time series list. For the future however a GetDataAvailability command shall solve this problem including provision of coverage information. The contributions made by the GEOWOW project aim at advancing the specification of this extension of the SOS standard.

Basic discovery functions are available in the SOS 2.0 standard with a filterable GetFeatureOfInterest command for retrieval of measuring locations. However, there are no similar options for procedures or observed properties. For filtering in general it may be helpful to add certain mandatory filters or requests to this profile to ensure a better usability.

4.3 Profile Structure

This profile needs to ensure full interoperability of any hydrology SOS with a profile compliant client. First of all it is necessary to publicly announce the profile in the

GetCapabilities request which can be done with the ‘profile’ attribute in the ServiceIdentification section.

Wherever possible, technical definitions will also be expressed in a suitable format such as xsd schemas or extensions. All crucial points will be specified as mandatory definitions, while options not necessary for flawless interoperability may also be added as recommendations to this document.

5. Related Standards and Activities

5.1 GML, O&M and WaterML 2.0

5.1.1 GML

The Geography Markup Language (GML) [1] is a standard of the Open Geospatial Consortium (OGC). It offers a data model as well as an XML encoding for geographic features (= "abstractions of real world phenomena").

GML has been developed to enable the standardized and interoperable exchange of geospatial objects (features) together with their attributes, their relationships to other objects as well as their geometries.

Main driving factors during the development of GML were the following aspects:

- Provision of a framework for describing geospatial application schemas to enable the transport and storage of geospatial information.
- Designed as an open and vendor independent standard that is based on XML
- Allow the creation of GML profiles that comprise subsets of GML (e.g. O&M as an application profile for encoding observation data)
- Support the definition of application schemas that customize GML for example to a specific application domain

For the SOS profile for hydrology, GML is relevant because of two reasons: On the one hand, the Observations and Measurements standard (which is again the foundation for WaterML 2.0) has been defined as an application profile of GML. On the other hand, GML is the standard for modeling and encoding geometries such as the features of interest, to which hydrological observations are related.

5.1.2 O&M

Observations and Measurements (O&M) belongs to the framework of Sensor Web Enablement (SWE) standards of the OGC. It defines a data model as well as an encoding for observation data (e.g. sensor data).

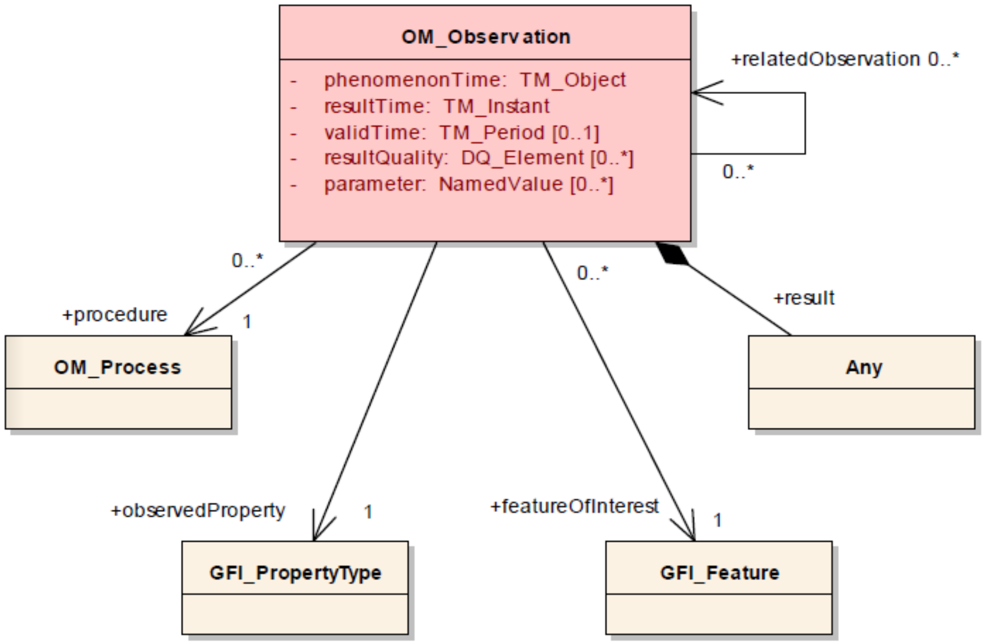

Figure 1: Overview of the Basic O&M Concepts (Source [7])

The figure above shows the general structure of an O&M Observation. Generally an observation contains a set of time stamps:

- The phenomenonTime describes the time for which the result applies to the observed property of the feature-of-interest (e.g. the time at which a certain water level occurred).

- The resultTime describes the time instant at which the result of the observation became available (e.g. when the measurement process was finished; if a 10 minute average is measured the result time is the point in time at which the average was calculated). However, often, the resultTime is identical to the phenomenonTime.

- The optional validTime specifies a time period during which the result can be used (outside this period the result is considered invalid).

Furthermore, every observation contains a description of the observed property (the phenomenon that was observed), the geometric feature to which the observation belongs (feature of interest) and information about the process/sensor (procedure) which has been used for obtaining the result. Finally, an observation may contain optional information about its quality as well as parameters of the measurement process.

The latest version of this specification is O&M 2.0 which was published at the end of 2011. Whereas the 1.0 version of the O&M standard [2], [3] was purely an OGC specification, O&M 2.0 has been adopted by the International Organisation for

Standardisation (ISO). The O&M 2.0 data model has been published as an ISO standard [4], while the XML encoding of this data model is available as an OGC standard [5].

Generally, the differences between O&M 1.0 and 2.0 are of evolutionary nature, so that the underlying concepts of both standard versions are the same. However, O&M 2.0 provides several advancements such as

- Better separation of conceptual model (ISO standard) and its implementation (XML encoding as an OGC standard)
- Addition of a spatial profile
- Additional observation properties (e.g. for relating observations to each other)
- Changes in terminology
- Removal of the observation collection type

As O&M 2.0 is also the basis for the specification of WaterML 2.0 (WaterML 2.0 is an application profile of O&M 2.0), O&M 1.0 has less relevance for the GEOWOW project. Thus, O&M 2.0 will play a more important role.

5.1.3 OGC WaterML 2.0

WaterML 2.0 is a standard information model for the representation of in-situ water observations data, with the intent of allowing the exchange of such data sets across information systems. It is based on Observations and Measurements 2.0 (O&M) and implemented as an application schema according to the rules of the Geography Markup Language version 3.2 (GML). Through the use of existing OGC standards, it aims at being an interoperable exchange format that may be re-used to address a range of hydrological data exchange requirements.

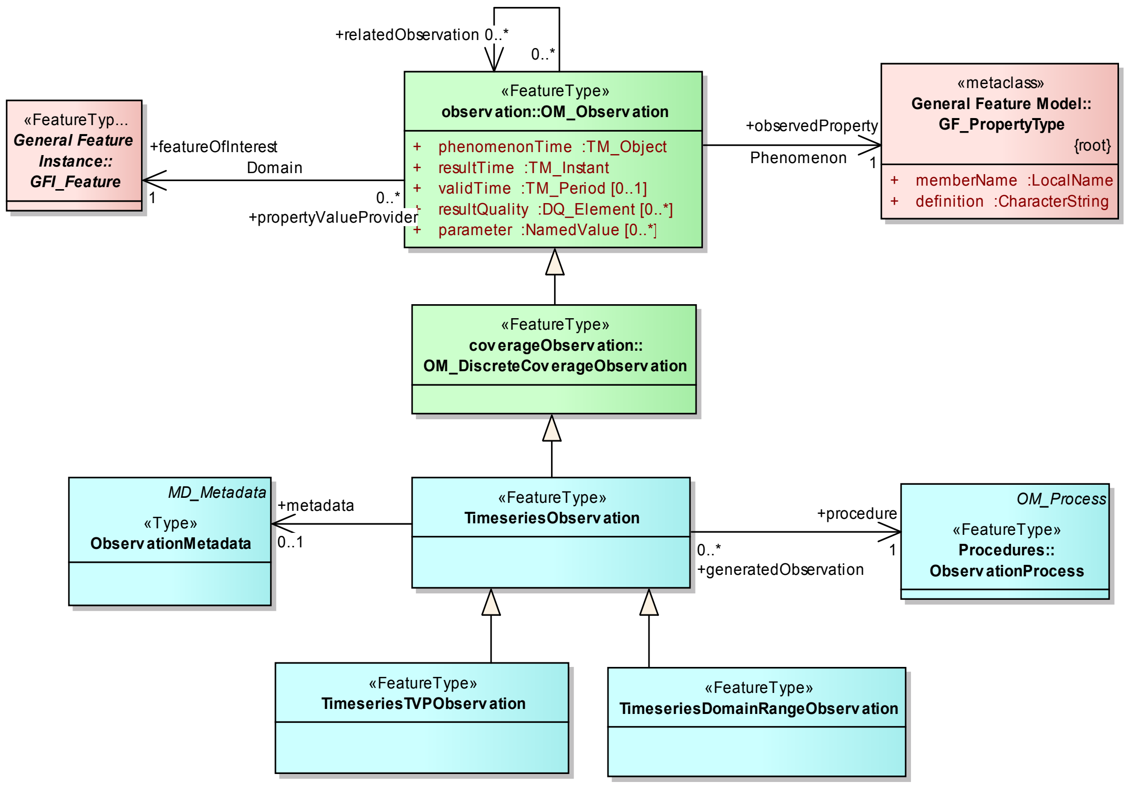

Figure 2: WaterML 2.0 core model as provided in the specification document (Source [9])

The core aspect of the model is in the correct and precise description of hydrological time series. Interpretation of time series relies on understanding the nature of the process that generated them. Thus, WaterML 2.0 customizes the more generic O&M 2.0 specification to the concrete requirements of the hydrology domain.

The information in this chapter is taken from the WaterML 2.0 specification document [9] which has been adopted by the OGC as an official standard in 2012.

Basic element of WaterML 2.0 documents is the wml2:Collection which features one or more wml2:observationMember elements. These members each equal an actual time series and include information about the measuring site and location as well as the parameter, process, and the actual result.

The featureOfInterest is defined as a samplingFeature which is equivalent to a station or site that provides the data. Metadata include location and time zone information as well as gauge datum and monitoring type. Several other parameters are available with the specification and more can be added by self-definable optional parameters.

The observation data can be stored either as categorical or measurement time series in the om:result element. The short description here refers to the more common measurement time series. Metadata is divided into time series based metadata which is valid for the whole time series (e.g. if the series is cumulative, equidistant etc.) and metadata which is valid on a per value basis. This allows storing the unit, quality and interpolation information as well as several other fields for each value individually (or defining a default value). The last part of the result is the list of values with according timestamps as sequence of wml2:point elements.

5.2 OGC Sensor Observation Service (SOS)

The Sensor Observation Service (SOS) is an interface specification included in the Sensor Web Enablement (SWE) framework of the Open Geospatial Consortium (OGC). It defines an interface comprising operations for accessing sensor data as well as metadata.

5.2.1 SOS 1.0

In the SOS 1.0 specification [7] the interface is divided into three groups:

- Core Profile: The main operations for retrieving sensor data and metadata
- Transactional Profile: Operations for inserting new sensors and observations into a SOS server
- Enhanced Profile: Additional operations for specific functionality such as requesting geometries of observed objects/features, determining the time periods for which data is available, etc.

For GEOWOW especially the following operations of SOS 1.0 were expected to be relevant:

- GetCapabilities: Accessing a description of a SOS server, its contents and the supported operations
- GetObservation: Accessing sensor data stored in a SOS server by filter criteria such as temporal extent, observed property, spatial extent, values etc.
- DescribeSensor: Accessing metadata about a sensor
- RegisterSensor: Inserting a new sensor into a SOS server
- InsertObservation: Inserting new observation data for a previously registered sensor into a SOS server
- GetFeatureOfInterest: Accessing the geometries of features of interest (e.g. sensor stations)
- GetFeatureOfInterestTime: Determining for which time periods data is available at a specific feature of interest (please note: in SOS 1.0 this operation has only a limited functionality and is usually not implemented).

However, due to the advantages of the SOS 2.0 standards, in GEOWOW the decision was taken to base the development of the SOS 2.0 Profile for Hydrology solely on the SOS 2.0 standard (see below).

5.2.2 SOS 2.0

The SOS 2.0 standard [8] is an evolutionary advancement of the SOS 1.0 standard. Experiences with the SOS 1.0 specification were used to improve the existing operations, to remove functionality which was not used in practice and to increase interoperability.

The main changes between SOS 1.0 and SOS 2.0 are the following:

- Restructuring of the specification by separating into core and extensions: this means that the distinction between the different profiles of SOS 1.0 is replaced by functional modules (conformance classes) that comprise certain types of functionality

- Introduction of a key value pair (KVP) binding for the SOS interface: although the SOS 1.0 standard mentions a key value pair binding it has not been included in the specification; this deficit is addressed by SOS 2.0

- Increased interoperability by a set of changes in the specification

- Definition of a mandatory set of operators and operands for temporal and spatial filters (clients can rely on the implementation of these operators and operands).

- Introduction of a Spatial Filtering Profile which defines interoperable access to spatial observations (allowing more advanced spatial queries).

- Definition of O&M as default and mandatory response format for observations. Thus, clients can rely on the availability of O&M support in SOS servers, while further data formats may be implemented in a SOS server in addition.

- Redesign of the GetCapabilities responses:

  - Limitation to one sensor per observation offering (please note: a sensor network could also be considered as one sensor). This allows determining the capabilities of a sensor directly within the Capabilities document

  - Listing of related features instead of all features of interest. This change is intended to reduce the length of GetCapabilities responses. As a replacement the functionality of the GetFeatureOfInterest operation has been extended.

- Result handling which allows to store templates for insertion as well as retrieval operations so that repetitive overhead in similar requests is avoided:

    - New operations for result insertion (InsertResult and InsertResultTemplate)

    - New operations for result retrieval (GetResult and GetResultTemplate)

- Removal of operations that were not used in practice

    - DescribeObservationType

    - DescribeResultModel

    - DescribeFeatureType

An additional extension for the retrieval of metadata about available data is in the specification process. This extension offers the so called GetDataAvailability operation which provides more comprehensive functionality than the GetFeatureOfInterestTime operation in the SOS 1.0 specification.

5.3 OGC Hydrology Domain Working Group and Interoperability Experiments

5.3.1 OGC Hydrology Domain Working Group

The following text is the official description of the OGC Hydrology Domain Working Group (DWG)[1].

"The Hydrology Domain Working Group is a domain working group in the Open Geospatial Consortium (OGC). It brings together interested parties to develop and promote the technology for greatly improving the way in which water information is described and shared. This working group is to be hosted by the OGC and co-chaired by a representative from the World Meteorological Organisation's (WMO) Commission for Hydrology (CHy). The Hydrology Domain Working Group will coordinate efforts with other earth science DWGs (Meteorology, Oceans, etc) through the Earth System Science DWG.

The purpose of the Hydro DWG is to provide a venue and mechanism for seeking technical and institutional solutions to the challenge of describing and exchanging data describing the state and location of water resources, both above and below the ground surface. The path to adoption will be through OGC papers and standards, advanced to ISO where appropriate, and also through the World Meteorological Organization's (WMO) and its Commission for Hydrology (CHy) and Information Systems (WIS) activities. While CHy has the recognized mandate to publish and promote standards in this area, OGC contributes to the process with its resources and experience in guiding collaborative development among disparate participants in a rapidly evolving technological environment. It is proposed that the OGC Hydrology DWG will provide a means of developing candidate standards for submission to ISO and for adoption by CHy as appropriate."

[1] Taken from http://external.opengis.org/twiki_public/HydrologyDWG/WebHome

5.3.2 Support by the OGC Hydrology DWG

The content of the SOS 2.0 Hydrology Profile has been developed with the invaluable help of the Hydrology Domain Working Group (DWG). Several important points were presented at meetings of the OGC Hydrology DWG in Reading in June 2012 and in Quebec City in 2013. These presentations lead to vivid discussions during which the participants confirmed the importance of this profile as well as the definitions and added ideas for further topics and solutions.

This document has been and will be passed on to the OGC Hydrology DWG members for review. The intention is to advance this document to an adopted OGC Best Practice Paper.

5.3.3 GroundwaterIE

The Groundwater Interoperability Experiment (GroundwaterIE) was initiated in 2009 and finished in spring 2011. Main purpose was the analysis of WaterML 2.0 development and the OGC service standards in respect to their usage in groundwater scenarios and use cases. Within the scope of this experiment the first indications of technical and semantic problems with the usage of SOS 1.0 in the hydrology domain were discovered and discussed. The focus there was mainly on discovery functionality via the GetCapabilities request that was unsuitable for some of the use cases.

The final report of the GroundwaterIE [10] contains a first set of definitions for a SOS and WFS profile that is dedicated to the usage in the more specialized groundwater domain.

5.3.4 SurfaceWaterIE

The Surface Water Interoperability Experiment (SurfaceWaterIE) was initiated in 2010 and finished in summer 2011. Similar to the GroundwaterIE the development of WaterML 2.0 was supported and based on some previous results and extensive analysis of problems with the SOS was performed. The experiment had three use cases, namely ‘Cross Border Data Exchange’, ‘Forecasting’ from which the current ForecastingIE was developed and ‘Global Runoff’. All use cases focused on data exchange relying on WaterML 2.0 and OGC services.

During this experiment several semantic types of SOS have been identified based on existing setups. Similar to the GroundwaterIE a set of recommendations has been created on how to use SOS and WaterML 2.0 in surface water use cases [11]. These more general findings also are the basis for the hydrology profile development on-going within the frame of GEOWOW.

5.3.5 ForecastingIE

The Forecasting Interoperability Experiment (ForecastingIE) was initiated in autumn 2011. It was also planned to look again at the profiling options for SOS in respect to hydrological forecasting data. The experiment utilized previous results and was planned to possibly adjust with the GEOWOW profile to not have separate developments. However, since mid-2012 the experiment is on hold and it is unsure if it will continue.

5.3.6 GroundwaterIE2

The Groundwater Interoperability Experiment 2 (GroundwaterIE2) started in September 2012 and focuses on further development of GroundWaterML2 as a dedicated language but also makes use of WaterML 2.0 for more generic purposes. Unlike previous experiments the focus is not directly on how to set up services. Therefore, no influence on the Hydrology Profile is expected. In case of a SOS 2.0 server being used within the experiment it is likely that the group will adopt the current profile version for it.

6. Terms and Definitions

This document uses the terms defined in Sub-clause 5.3 of [OGC 06-121r8], which is based on the ISO/IEC Directives, Part 2, Rules for the structure and drafting of International Standards. In particular, the word “shall” (not “must”) is the verb form used to indicate a requirement to be strictly followed to conform to this standard.

For the purposes of this document, the following additional terms and definitions apply.

Terminology from other well established Glossaries such as the OGC Glossary (http://www.opengeospatial.org/ogc/glossary), the ISO TC-211 Glossary (http://www.isotc211.org/Terminology.htm), the OGSA Glossary (http://www.ogf.org/documents/GFD.120.pdf), the OASIS SAML/XACML Glossary (http://www.oasis-open.org/committees/security/docs/draft-sstc-glossary-02.pdf), have been re-used as much as possible and complemented with additional definitions when needed.

6.1 Profile Identification

Any SOS server supporting the SOS 2.0 Hydrology Profile shall announce it within the GetCapabilities request for clients.

Definition 1
The profile shall announce itself in the ServiceIdentification section as attribute profile with the value ‘http://www.opengis.net/spec/SOS_application-profile_hydrology/1.0/ req/hydrosos’.

6.2 FeatureOfInterest / WaterMonitoringPoint

The feature of interest refers to a feature that is the target of an observation. Concerning the feature of interest it is necessary to distinguish between sampled features and sampling features.

Within this profile sampled features shall be considered as objects of study at a domain-level (domain feature), e.g. a country, a river, a river section, a lake, etc. There are relationships between these sampled features, and the in-situ locations where observations are performed.

Sampling features in this profile refer to in-situ locations where observations are performed. While technically it is possible to develop more subtypes for sampling features, the only one currently defined (within this scope) is the wml2:MonitoringPoint for WaterML 2.0. However, for future compatibility the usage of other subtypes has to be possible which leads to a recommendation to use wml2:MonitoringPoint with the option to adopt other types later.

Definition 2
All sampling features within the scope of this profile are recommended to use the type wml2:MonitoringPoint. This applies for filters and responses of the GetFeatureOfInterest and GetObservation operations and satisfies the WaterML 2.0 requirements class http://www.opengis.net/spec/waterml/2.0/req/uml-monitoring-point-feature-of-interest. Each complex wml2:MonitoringPoint element shall comprise at least the following sub elements: - gml:identifier as an identifier for the feature - gml:name as a label for the feature

6.3 ObservedProperty

The observed property is the phenomenon for which an observation result provides a value. It is a property that is associated to the feature of interest of the observation. Examples for an observed property are discharge or water level.

This property is normally implemented using an identifier for the property that is being observed. If possible this identifier shall reference a controlled vocabulary. Currently, there is no agreement upon standard vocabularies for observed phenomena within the hydrology domain, but there is on-going work in the area to address this issue.

6.4 Procedure / ObservationProcess

A procedure is defined as a measurement process, analysis, or processing algorithm that is used to obtain an observation result. Within this profile a process is considered as an algorithm, sensor type, or time series type, but not as an individual, physical device (sensor instance).

Definition 3
All procedures within the scope of this profile shall be restricted to the type wml2:ObservationProcess. This applies for filters of the GetFeatureOfInterest and GetDataAvailability operations as well as all filters and responses of the GetObservation operation. This definition shall satisfy the WaterML 2.0 requirements class http://www.opengis.net/spec/waterml/2.0/req/uml-observervation-process.

6.5 Identification

Identifying an object in the context of this profile means to have exactly one URI pointer for an object and to have at least one human readable/understandable name which serves for labeling that object.

Identification generally should not be achieved using xlink:href and xlink:title from gml:ReferenceType wherever explicit identification is possible. Identification is crucial and therefore shall be done explicitly by using gml:identifier and gml:name from gml:StandardObjectProperties, as well as swes:identifier and swes:name from swes:AbstractSWESType respectively.

Definition 4
Identification of entities shall refer to the gml:identifier field of the entity wherever possible, while the gml:name field shall hold a label name for it.

Where gml:ReferenceType is the only possible identification type (this refers to a 'simple' element), xlink:href shall be used for unique identification whereas xlink:title shall be used to declare a human readable name.

Definition 5
If the entity is not listed as a complex element and therefore does not feature a gml:identifier, the xlink:href attribute shall be used as identifier, while the xlink:title field shall hold a label name for it. This makes the WML2 recommendation http://www.opengis.net/spec/waterml/2.0/req/xsd-xml-rules/rec/xsd-xml-rules/xlink-title mandatory.

Which variant has to be used depends on the underlying specification (e.g. SOS, WaterML 2.0).

Additionally to the identification of individual elements, it also is of general interest to identify a single time series either by an ID or by metadata. Since the concepts of SOS2 and WaterML 2.0 do not provide a real unique time series ID, the combination of the metadata fields feature of interest, observed property and procedure can be used. While it is technically possible to have multiple time series with the same metadata this behavior should be discouraged for obvious reasons – since the metadata fields are used to find and retrieve time series data, a client would not be able to distinguish two or more time series with identical so-called identifying fields.

Definition 6
Any existing combination of the identifying metadata fields feature of interest, observed property and procedure shall be unique within the service and identify either none or only a single time series.

6.6 Open Points

- SOS 2.0 does not provide human readable names in the capabilities. This makes it impossible for clients displaying other than identification URIs which are generally not intended for human identification. To resolve this issue, it is planned to submit a change request to the relevant OGC Standard Working Groups.

- List exactly which elements become mandatory with the definitions e.g. in a MonitoringPoint (these are normally optional elements like gml:identifier). This is partly done and will be finished later.

7. Requests and Responses

7.1 GetCapabilities

The GetCapabilities request is supposed to be the main discovery function that provides a rough overview of a SOS server's functionality and content. It is not meant to expose the complete metadata structure and just leave the actual time series data out.

7.1.1 Avoiding falsely indicated homogeneous distribution of time series in a SOS server instance

Based on the listings of procedures and properties in the GetCapabilities document and the similar list of sampling features in the GetFeatureOfInterest command the service suggests that any possible combination of these entities leads to exactly one time series. In a practical setup however databases tend to be inhomogeneous, e.g. one station only measures precipitation, while another one only measures river stage.

The pattern for structuring the data offered by a SOS server in GetCapabilities responses is the offering concept. However, the offering concept is constrained so that one offering may contain only one procedure. This way it is possible to express the relationship between procedures and observed properties but further structuring is not offered. A filter system for further relationships is only available in the (optional) GetFeatureOfInterest request, where sampling features can be searched based on their related procedures and observed properties.

An explicit listing of existing time series is missing although it can be technically constructed with the filter mechanisms of GetFeatureOfInterest as described above (as long as the used filters are combined with the resulting feature list). However, this would require a rather high degree of complexity in the business logic of clients. To overcome this problem there is an ongoing discussion about a GetDataAvailability extension for the SOS specification that will be looked at closer in section 7.4. This has successfully been implemented in the forecasting IEs and advanced in the GEOWOW project to allow granular definition of the available time extents of individual time series.

7.1.2 Listing of related features

Listing all sampling features of interest (measurement stations) may lead to very large capabilities documents that can affect client performance for parsing. To avoid listing all sampling features available in a SOS server, they have to be available via the getFeatureOfInterest request only. Sampled features may be added as swes:FeatureRelationship.

As the OGC SWE Service Model defines the default feature relationship as http://www.opengis.net/def/nil/OGC/0/unknown the sampled feature's role shall explicitly be marked with http://www.opengis.net/def/featureType/domainFeature.

Definition 7
The GetCapabilities content section shall list sampled features only and those as related features, no sampling features shall be exposed. If no sampled features are available no features at all shall be listed. The OperationsMetadata shall not expose any features at all.

7.2 GetFeatureOfInterest

The GetFeatureOfInterest request is an additional operation of the SOS interface that lists and is capable of filtering features of interest. It is not specified in any more detail what subtypes of features should be contained.

7.2.1 Usage of GetFeatureOfInterest

In the specification this request is optional but since it is crucial for the discovery of sampling features (unless an additional WFS server is available) it is mandatory within the SOS 2.0 Hydrology Profile.

Definition 8
The GetFeatureOfInterest request shall be mandatory.

7.2.2 GetFeatureOfInterest content

Additionally the generic feature of interest entity shall be limited to sampling features with recommendation for wml2:MonitoringPoint to have a direct connection with the according WaterML 2.0 restriction. This way it is ensured that only monitoring points in the sense of sampling features are returned which simplifies the understanding and usability of this operation.

Definition 9
The GetFeatureOfInterest request shall return sampling features only, it is additionally recommended that these be entities of the type wml2:MonitoringPoint.

7.3 DescribeSensor

The DescribeSensor operation is mandatory. According to the SOS specification, it is intended to deliver sensor/procedure descriptions, usually encoded in SensorML.

Within WaterML 2.0 the wml2:observationProcess type is used for describing procedures. Thus, it either must be used as format or it would be necessary to map the contents of a wml2:observationProcess to SensorML documents, which does not provide any additional information. Therefore the wml2:observationProcess shall be used as procedure description format with SensorML being optional.

Definition 10
The procedure description format http://www.opengis.net/waterml/2.0/observationProcess shall be mandatory and shall be announced in the GetCapabilities response.

Definition 11
The DescribeSensor request shall return procedures as http://www.opengis.net/waterml/2.0/observationProcess

7.4 GetDataAvailability

The SOS specification does not provide standardized means for a client to construct valid parameter constellations (i.e. combinations of the query parameters "offering", "procedure", "feature of interest" and "observed property") that refer to an existing time series. Even though best practices are available to let clients construct such constellations, clients still have to deal with SOS dialects to request an existing time series. Within this profile, a mandatoryGetDataAvailability operation is added to the SOS 2.0 interface that provides a filterable time series list and is a result from previous understandings and results of interoperability experiments.

The request enables clients to filter for existing time series by their identifying metadata fields and the offering they may belong to. The resulting time series list contains all time series matching the filter criteria with their identifying metadata fields and their data coverage in form of a phenomenon time element. An optional field for the count of values in a time series is also available.

A detailed specification of this operation will be included in the next iteration of this document.

Definition 12
The support of the GetDataAvailability operation by servers shall be mandatory.

7.4.1 Open Points

- Schema development
- SOAP actions
- Profile entry, probably http://www.opengis.net/spec/SOS/2.0/conf/gda

7.5 GetObservation

The GetObservation operation is the way for data retrieval from SOS servers. It allows several types of entity based filters as well as temporal filters to limit the response.

7.5.1 Response structure if no values are in the time range

Currently compliant SOS implementations do respond with different messages when no observations are present for a given request. While some services still return the metadata (like sampling feature and procedure etc.) of the time series, other services deliver either an error message or a completely empty GetObservation response. For a client however it is crucial to know that the time series exists and just does not contain data for the requested time range. Therefore it is important for the service to return the time series metadata anyway and only respond with an error message if an explicitly requested time series does not exist (e.g. a combination of feature of interest, observed property and procedure has no result).

Definition 13
The GetObservation request shall return metadata of all matching timeseries in form of the OM_Observation element even if they do not contain any values for the requested time range.

7.5.2 Handling of incomplete responses

An incomplete response for example is, if the client sends the request for one procedure, one property and three features. Technically this would lead to a response comprising three time series. But it is relatively common that maybe only two time series exist which leaves the service with the two options to either return everything that was found or to return an error message. Since the SOS specification also allows loose requests (like observed property and procedure only) it makes the most sense to generally deliver everything that matched the criteria.

Definition 14
The GetObservation request shall return all existing time series that match the request parameters.

7.5.3 Handling of requests without temporal filter

Since the GetObservation request does not define a temporal filter as mandatory, a request without this filter technically is supposed to return all data for all time series that match the query parameters. For time series with many values (e.g. 100.000 is quite common) this typically leads to the problem of putting a lot of load on the server. In extreme cases it might even kill the server or equal an involuntary denial-of-service attack. While client applications usually use well-defined requests, direct users tend to start with a simple call without limiting the time range just to ‘see what you get’.

To protect their servers and also the clients several service providers have already changed the response behavior to return the last value of a time series only if no temporal filter is specified. This not only protects the server from unintended large requests but also adds a new feature to the SOS interface. The latest value of a time series often is of special interest for a user and otherwise would have to be requested based on the data coverage of a time series which is not even available with the basic set of requests. It is obvious that this behavior is a double advantage and the full data content of a time series still can be retrieved by specifying a suitable temporal filter.

Definition 15
If a GetObservation request does not contain a temporal filter the service shall only return the latest value of each time series instead of the full content.

7.5.4 Handling of empty responses

Depending on the requested entities it is possible that no time series are found to match the criteria. There are two ways to handle such a case: the service can either respond with an empty GetObservation response or return an error message stating that no results could be found. Since it is more consistent with the previous definitions and also allows easier handling in clients this is defined as an empty response.

Definition 16
If the GetObservation request does not lead to any resulting time series an empty sos:GetObservationResponse document shall be returned.

7.6 Profile appliance to SOS 1.0

Due to the specific requirements of the hydrology domain, which can be covered best through the enhancements offered by the SOS 2.0 standard, the SOS Hydrology Profile has only been defined for SOS 2.0.

8. Technical Aspects

8.1 Simple versus complex elements

For the basic entities the standards underlying the SOS 2.0 Hydrology Profile (in most places) allow to either use a simple element in the form of a one-line reference, e.g.

```
<featureOfInterest href=”http://myservice/features/WaterWell1” title=”WaterWell1”>
```

or to specify a complex version that does not feature the ‘href’ and ‘title’ attributes but has more detailed information about the entity.

This means that the complex versions potentially have different information fields compared to the simple ones (if not resolved based on the reference) which can lead to problems in discovery and uniqueness.

8.2 Bindings

8.2.1 SOAP Binding

While SOS 1.0 does not define a SOAP binding at all, the SOS 2.0 specification provides a basic binding for each request. This however is nothing more than a regular XML based POST request wrapped into a SOAP envelope. It allows usage of generic SOAP features that require SOAP headers to be existent (e.g. authentication services) but does not provide any help with building a client since the necessary complete WSDL file is missing.

With the help of a WSDL file a user could easily just use a generic SOAP client to construct a simple SOS 2.0 client and access any services. On the other hand it has not yet been discussed in depth if it is possible to map the more complex features like filtering to an actual WSDL file. Even a simple WSDL file with basic query options could prove helpful for users interested in building a SOS 2.0 client or just trying out the service. In a future iteration of this document an according WSDL file will be included.

8.2.2 KVP Binding

Neither SOAP nor KVP bindings are actually mandatory in the SOS 2.0 specification. Therefore it is essential for a client to know that at least one of the two bindings is available.

Definition 17
The KVP binding in the SOS 2.0 shall be supported by all SOS servers implementing this profile.

8.2.3 Open Points

- Discuss creation of a useful WSDL for the hydrology profile SOS.

8.3 Response limits

An issue with both the SOS 1.0 and 2.0 specifications is the relative ‘openness’ to extensive queries either by mistake or on purpose. For instance, a request with no time range and just one observed property and offering could easily return some hundred time series with maybe millions of values.

The first step is taken with Definition 15 that limits the response to one value per time series if no temporal filter has been set. But it is still possible to explicitly request the full range of many time series by setting a wide temporal filter in GetObservation requests.

To prevent overloading SOS servers there are two ways. The first one is to make more restricted filtering mandatory which still might be insufficient if large requests are made on purpose. The second option is to allow any request, but simply limit the response size or content. The new SOS 2.0 standard actually specifies an error message indicating that a request would lead to a too large response. The specification does not say anything about what is considered too much though and the message also does not contain any additional information.

Since it is already agreed that a security limitation via response restrictions is needed to protect the server (as well as the client), it has to be decided how a response size limit can be realized. Again there are two options: The simple solution is to limit the number of overall values that the service will return before it responds with an error message that also includes this limit (so that it can be parsed by a client for instance). At the OGC Hydrology DWG meeting in Reading it was criticized that this might make it hard to harvest data from a service because the client needs to be relatively intelligent to determine how to create suitable requests after such an error message.

Because of the complex request options of the SOS interface it is obvious that the service itself cannot deliver a ‘follow-up’ request to indicate some kind of paging mechanism. Therefore it was suggested to think about a limitation on time spans instead of values that would simplify client logic for harvesting. This however introduces new problems, for instance it makes a big difference if the requested time series contain yearly or five minute data. Also there are often time series that are not equidistant. So a time dependent limitation like 500 years of data for the whole request may be very limiting for yearly time series (500 values) while it is already large for five minute data (52560000 values). At the same time it is not possible to distinguish between those cases. Additionally, if more than one time series is requested the client still has no indication which time series are more or less problematic (this is the same as with the count limitation of course).

Because of the collected discussion points it has been decided to go with a count limitation. For facilitating the handling of limits by the client the SOS needs to announce any limit values within the AccessConstraints section of the GetCapabilities document. Additionally, it is recommended that the limit for time series values should be larger than the largest time series in the underlying database.

Definition 18
If a SOS server has a limit on how many time series values can be returned within one request, this limit shall be announced in the AccessConstraints section of the GetCapabilities document. It is recommended that this limit be higher than the value count of the largest time series in the underlying database if possible. The following string shall be included in an AccessConstraints element for indicating the limit of the number of values that can be returned (replace XX with an integer value representing the limit): “maxNumberOfReturnedValues=XX”

Definition 19
If a SOS server has a limit on how many time series can be returned within one request, this limit shall be announced in the AccessConstraints section of the GetCapabilities document. The following string shall be included in an AccessConstraints element for indicating the limit of the number of time series that can be returned (replace XX with an integer value representing the limit): “maxNumberOfReturnedTimeSeries=XX”

9. Outlook

This version of the SOS 2.0 Hydrology Profile relies on a standalone SOS for hydrological data. For additional discovery and search options it would be very helpful to have an additional OGC WFS with basic functionality as support for the SOS (i.e. for managing related sampled features). This is not a new idea and especially large (and therefore unhandy) services would profit from a combined setup.

It is intended that each SOS instance still should be able to work as standalone server and therefore provide the GetFeatureOfInterest request and the GetDataAvailability extension, but the combination with an additional WFS would no doubt be very powerful and fully in the sense of the OGC service architecture.

Furthermore it is planned for the future to advance this paper in close cooperation with the OGC Hydrology Domain Working Group to the status of an OGC Best Practice Paper.

Annex A: Revision History

Date	Release	Author	Paragraph modified	Description
03/03/2014	0.3	All authors listed as contributors	All	Initial version

Annex B: Bibliography

[1] Portele, Clemens (2012). OGC Implementation Specification: Geography Markup Language (GML) - Extended Schemas and Encoding Rules (OGC 10-129r1). Wayland, MA, USA, Open Geospatial Consortium Inc.

[2] Cox, Simon (2007). OGC Implementation Specification: Observations and Measurements (O&M) - Part 1 - Observation Schema 1.0.0 (07-022r1). Wayland, MA, USA, Open Geospatial Consortium Inc.

[3] Cox, Simon (2007). OGC Implementation Specification: Observations and Measurements (O&M) - Part 2 - Sampling Features 1.0.0 (07-002r3). Wayland, MA, USA, Open Geospatial Consortium Inc.

[4] ISO TC 211 (2011). ISO 19156:2011 - Geographic information -- Observations and measurements - International Standard. Geneva, Switzerland, International Organization for Standardization.

[5] Cox, Simon (2011). OGC Implementation Specification: Observations and Measurements (O&M) - XML Implementation 2.0 (10-025r1). Wayland, MA, USA, Open Geospatial Consortium Inc.

[6] Botts, Mike and Alexandre Robin (2007). OGC Implementation Specification: Sensor Model Language (SensorML) 1.0.0 (07-000). Wayland, MA, USA, Open Geospatial Consortium Inc.

[7] Na, Arthur and Mark Priest (2007). OGC Implementation Specification: Sensor Observation Service (SOS) 1.0.0 (06-009r6). Wayland, MA, USA, Open Geospatial Consortium Inc.

[8] Bröring, Arne, Christoph Stasch and Johannes Echterhoff (2012). OGC Implementation Specification: Sensor Observation Service (SOS) 2.0 (12-006). Wayland, MA, USA, Open Geospatial Consortium Inc.

[9] Taylor, Peter (2012). OGC Implementation Specification: WaterML 2.0: Part 1 - Timeseries (0-126r3). Wayland, MA, USA, Open Geospatial Consortium Inc.

[10] Brodaric, Boyan and Booth, Nate (2011). OGC Interoperability Experiment Report: OGC Groundwater Interoperability Experiment – Final Report (10-194r3). Wayland, MA, USA, Open Geospatial Consortium Inc.

[11] Fitch, Peter (2011). OGC Interoperability Experiment Report: OGC Surfacewater Interoperability Experiment – Final Report (12-018). Wayland, MA, USA, Open Geospatial Consortium Inc.