Deltares
Enabling Delta Life

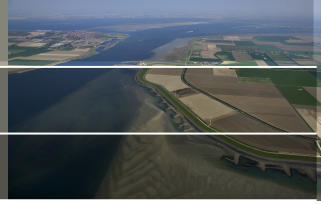
Comparison of time series ingest performance of various standardized file formats using DelftFEWS

Peter Gijsbers, Erik de Rooij

Open Water Symposium and Workshops
April 19, 2011

UNESCO-IHE
Institute for Water Education

Topics



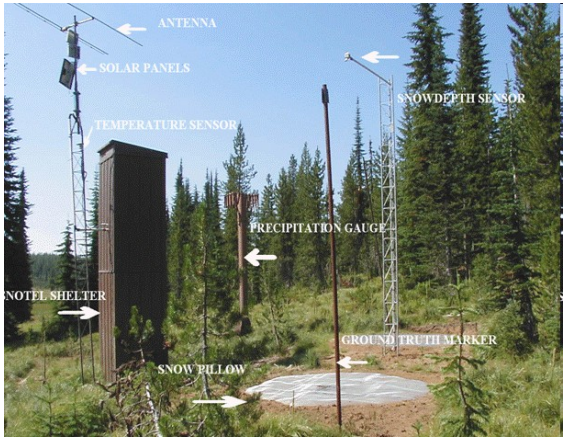
- Context
  - Standardization to share hydrological data
  - Flood Forecasting
- Objective of the experiment
- Results
- Lessons learned

Context (i): Data sharing

- DWG Hydro

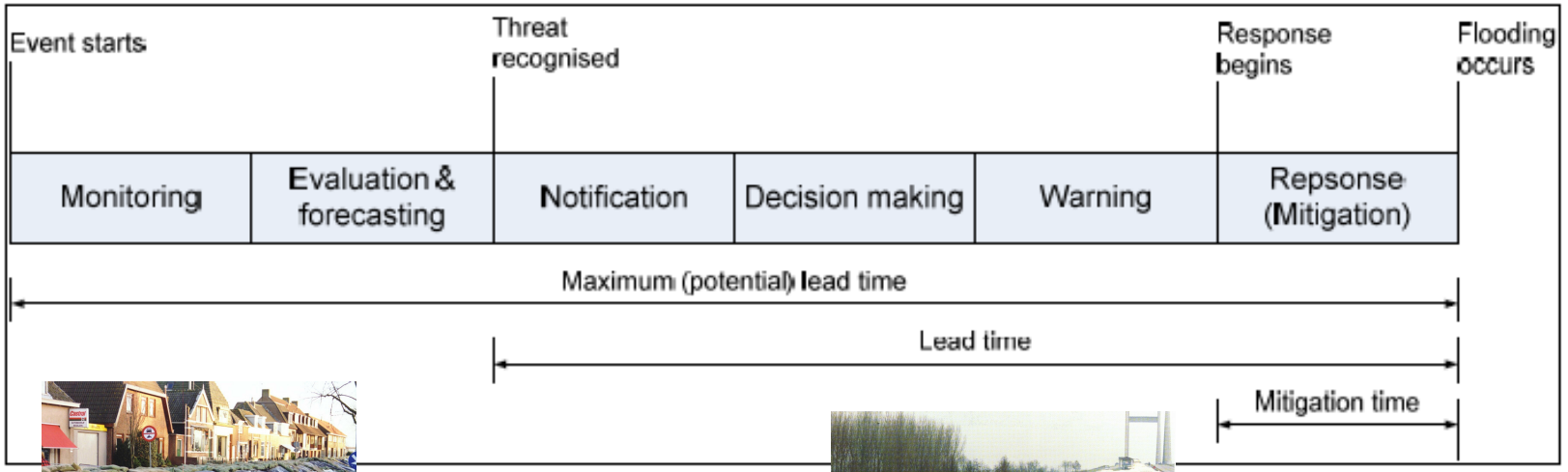
- Sensor web enablement
- Hydrological time series data exchange
- WaterML 2.0

- Interoperability Experiments



Deltares

Context (ii): Flood Forecasting



Forecasts are only useful when there is time left to warn and act

Deltares

Surface Water Interoperability Experiment

- Part of Surface Water IE by DWG Hydro
- Use Case 2: data ingest

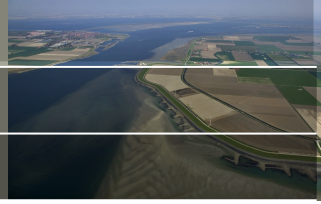
- Objective:
- Asses suitability of OGC standards for incremental data ingest in real time to support forecasting systems
  - SOS2.0: web-service (call method)
  - WaterML 2.0: data encoding (format)

Preparation for Hydrologic Forecasting IE

SWIE-UC2 (OHI0) – Operational Hydrologic incremental data Ingest
HFIE-OHF1 – Operational Hydrologic flood Forecasting
HFIE-HIO2 – Hydrologic Infrastructure Operation
HFIE-MHOF3 – joint Meteo-Hydro Operational Forecasting

Planning:
- Kickoff sep 2011

Methodology



Compare performance of data ingest of OGC encodings with other standardized data formats:

- NWS-SHEF
- FEWS-PI-xml
- O&M1 (SWE-common)
- WaterML 2.0

Comparison criteria:

- Lines of source code in client: → proxy for code complexity
- File size (un/compressed) → proxy for network load
- Ingest time (net) required by DelftFEWS

Deltares

File formats – NOAA/NWS-SHEF

Standard Hydrologic Exchange Format

Fortran aged encoding style (semi fixed formatting)

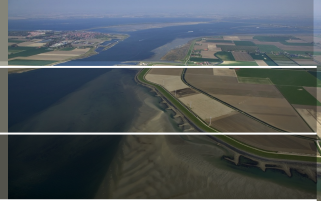
Various flavours (.A, .AR, .B., .E, .ER)

```
: SHEF file
: Date/time forecast: 20110413

.ER TAPM4 20110403 Z DH06/DC201104151718/QIN    /DIH01
.E1    780.00/  780.00/  785.00/  785.00/  785.00/  790.00/  795.00/  801.00/
.E2    801.00/  806.00/  811.00/  816.00/  821.00/  821.00/  826.00/  832.00/
```

Deltares

File formats – FEWS PI-xml



DelftFEWS Published Interface

xml-encoding

```
<TimeSeries xmlns="http://www.wldelft.nl/fews/PI" ....">
    <timeZone>0.0</timeZone>
    <series>
        <header>
            <type>instantaneous</type>
            <locationId>TAPM4</locationId>
            <parameterId>QIN</parameterId>
            <timeStep unit="nonequidistant"/>
            <startDate date="2011-04-13" time="06:00:00"/>
            <endDate date="2011-04-15" time="06:00:00"/>
            <missVal>-999.0</missVal>
            <stationName>Paradise 12W</stationName>
            <units>CMS</units>
        </header>
        <event date="2011-04-13" time="06:00:00" value="2329.9998" flag="0"/>
        <event date="2011-04-13" time="07:00:00" value="2329.9998" flag="0"/>
        <event date="2011-04-13" time="08:00:00" value="2329.9998" flag="0"/>
     ....
    </series>
</TimeSeries>
```

Deltares

File formats – OGC-O&M1- SWE common

OGC-Observatoins & Measurements using SWE Common encoding

```
<om:ObservationCollection gml:id="oc_oc1302859889906" ...>
  <gml:boundedBy>
    <gml:Envelope srsName="urn:ogc:def:crs:EPSG:4326">
      <gml:lowerCorner>51.58278906918545 11.794808643213615</gml:lowerCorner>
      <gml:upperCorner>51.58278906918545 11.794808643213615</gml:upperCorner>
    </gml:Envelope>
  </gml:boundedBy>
  <om:member>
    <om:Observation>
      <om:samplingTime>
        <gml:TimePeriod xsi:type="gml:TimePeriodType">
          <gml:beginPosition>2011-04-14T08:00:00.000+02:00</gml:beginPosition>
          <gml:endPosition>2011-04-15T05:00:00.000+02:00</gml:endPosition>
        </gml:TimePeriod>
      </om:samplingTime>
      <om:procedure xlink:href="Wasserstand-Wettin_Up_570840"/>
      <om:observedProperty>
        <swe:CompositePhenomenon gml:id="cpid0" dimension="1">
          <gml:name>resultComponents</gml:name>
          <swe:component xlink:href="urn:ogc:data:time:iso8601"/>
          <swe:component xlink:href="Wasserstand"/>
        </swe:CompositePhenomenon>
      </om:observedProperty>
      <om:featureOfInterest>
        <gml:FeatureCollection>
          <gml:featureMember>
            <sa:SamplingPoint gml:id="Wettin_Up_570840" xsi:schemaLocation=" http://www.opengis.net/sampling/1.0
http://schemas.opengis.net/sampling/1.0.0/sampling.xsd">
              <gml:name>Wettin Up - Kilometer: 70</gml:name>
              <sa:sampledFeature xlink:role="urn:x-ogc:def:property:river" xlink:href="SAALE"/>
              <sa:position>
                <gml:Point>
                  <gml:pos srsName="urn:ogc:def:crs:EPSG:4326">51.58278906918545 11.794808643213615</gml:pos>
                </gml:Point>
              </sa:position>
            </sa:SamplingPoint>
          </gml:featureMember>
        </gml:FeatureCollection>
      </om:featureOfInterest>
      <om:result>
        <swe:DataArray>
          <swe:elementCount>
            <swe:Count>
              <swe:value>57</swe:value>
            </swe:Count>
          </swe:elementCount>
          <swe:elementType name="Components">
            <swe:DataRecord>
              <swe:field name="Time">
                <swe:Time definition="urn:ogc:data:time:iso8601"/>
              </swe:field>
              <swe:field name="feature">
                <swe:Text definition="urn:ogc:data:feature"/>
              </swe:field>
              <swe:field name="Wasserstand">
                <swe:Quantity definition="Wasserstand">
                  <swe:uom code="cm"/>
                </swe:Quantity>
              </swe:field>
            </swe:DataRecord>
          </swe:elementType>
          <swe:encoding>
            <swe:TextBlock decimalSeparator="." tokenSeparator="," blockSeparator=";"/>
          </swe:encoding>
          <swe:values>2011-04-14T08:00:00.000+02:00,Wettin_Up_570840,366.0;2011-04-14T08:15:00.000+02:00,Wettin_Up_570840,366.0;</swe:values>
        </swe:DataArray>
      </om:result>
```

Deltares

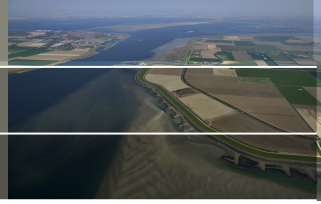
File formats – OGC-WaterML 2.0

OGC-Observatoins & Measurements using WaterML2 encoding

```
<wml2:TimeseriesObservation ../waterml2.xsd">
<gml:identifier codeSpace="http://nwis.waterdata.usgs.gov/WI/nwis">USGS.04071765</gml:identifier>
<gml:name codeSpace="http://nwis.waterdata.usgs.gov/WI/nwis">OCONTO RIVER NEAR OCONTO, WI</gml:name>
<om:metadata>
        <wml2:ObservationMetadata>
                <gmd:contact xlink:href="mailto:h2o@usgs.gov"/>
                <gmd:dateStamp>
                <gco:Date>2011-04-13</gco:Date>
                </gmd:dateStamp>
                <gmd:identificationInfo xlink:href="urn:OGC:unknown"/>
                <wml2:status xlink:href="http://waterdata.usgs.gov/WI/nwis/help/?provisional"/>
        </wml2:ObservationMetadata>
</om:metadata>
<om:phenomenonTime>
        <gml:TimePeriod gml:id="ts_period">
                <gml:beginPosition>2011-04-03T00:00:00-06:00</gml:beginPosition>
                <gml:endPosition>2011-04-13T07:00:00-06:00</gml:endPosition>
        </gml:TimePeriod>
</om:phenomenonTime>
<om:resultTime>
        <gml:TimeInstant gml:id="result_time">
                <gml:timePosition>2011-04-13T09:43:31</gml:timePosition>
        </gml:TimeInstant>
</om:resultTime>
<om:procedure xlink:href="http://www.nemi.gov" xlink:title="Discharge"/>
<om:observedProperty xlink:href="urn:ogc:def:phenomenon:OGC:Discharge" xlink:title="Discharge"/>
<om:featureOfInterest>
        <wml2:MonitoringPoint gml:id="USGS.WMP.04071765">
                <sf:sampledFeature xlink:href="http://nwisvaws02.er.usgs.gov/ogc-swie/wfs?request=GetFeature&featureId=04071765"/>
                <sf:parameter>
                        <om:NamedValue>
                                <om:name xlink:title="Watershed"/>
                                <om:value>Oconto</om:value>
                        </om:NamedValue>
                </sf:parameter>
                <sams:shape>
                        <gml:Point gml:id="USGS.P.04071765">
                                <gml:pos srsName="urn:ogc:def:crs:EPSG:4269">44.86054589 -87.98399270</gml:pos>
                        </gml:Point>
                </sams:shape>
                <wml2:descriptionReference xlink:href="http://external.opengis.org/twiki_public/bin/view/HydrologyDWG/" xlink:title="IE"/>
                <wml2:timeZone>
                        <wml2:TimeZone>
                                <wml2:zoneOffset>-05:00</wml2:zoneOffset>
                                <wml2:zoneAbbreviation>CST</wml2:zoneAbbreviation>
                        </wml2:TimeZone>
                </wml2:timeZone>
        </wml2:MonitoringPoint>
</om:featureOfInterest>
<om:result>
        <wml2:owner>
                <gmd:organisationName>
                        <gmd:CharacterString>Wisconsin Water Science Center</gmd:CharacterString>
                </gmd:organisationName>
        </wml2:owner>
        <wml2:Timeseries gml:id="time_series_loc_0">
                <wml2:domainExtent xlink:href="ts_period">
                        <gml:TimePeriod gml:id="USGS.TP.04071765">
                                <gml:beginPosition>2011-04-03T00:00:00-06:00</gml:beginPosition>
                                <gml:endPosition>2011-04-13T07:00:00-06:00</gml:endPosition>
                        </gml:TimePeriod>
                </wml2:domainExtent>
                <wml2:defaultTimeValuePair>
                        <wml2:TimeValuePair>
                                <wml2:unitOfMeasure uom="cfs"/>
                                <wml2:dataType xlink:href="http://www.opengis.net/def/timeseriesType/WaterML/2.0/Continuous"
xlink:title="Continuous/Instantaneous"/>
                                <wml2:qualifier xlink:href="http://waterdata.usgs.gov/WI/nwis/help/?provisional" xlink:title="Provisional data subject to
revision."/>
                        </wml2:TimeValuePair>
                </wml2:defaultTimeValuePair>
                <wml2:point>
                        <wml2:TimeValuePair>
                                <wml2:time>2011-04-13T07:00:00-06:00</wml2:time>
                                <wml2:value>3240</wml2:value>
```

Deltares

Service details



USGS (USA):

- SOS2 service with WaterML2.0 data encoding
  - Hourly data: 1 station
  - 15 min. data: 9 stations

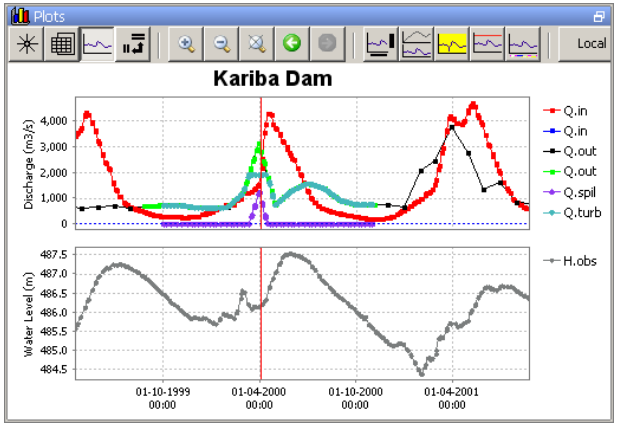
Pegelonline (Germany):

- SOS1 service with O&M1 (SWE common) data encoding
  - 1 min data: 1 station
  - 15 min. data: 7 stations

  - Data retrieval for ten days
  - Data retrieval for one day

Delft FEWS

Open Shell Forecasting System



System for operational forecasting (resilience!)

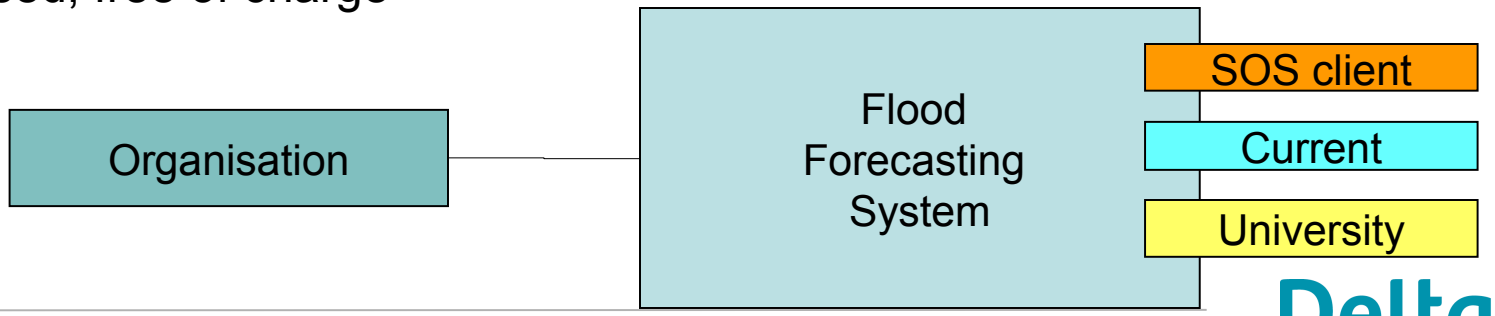
- Floods & droughts
- Reservoir operation
- Storm surge
- Water quality

Organise data from ingest to forecast

Open interface for variety of models

Fully configurable

Licensed, free of charge

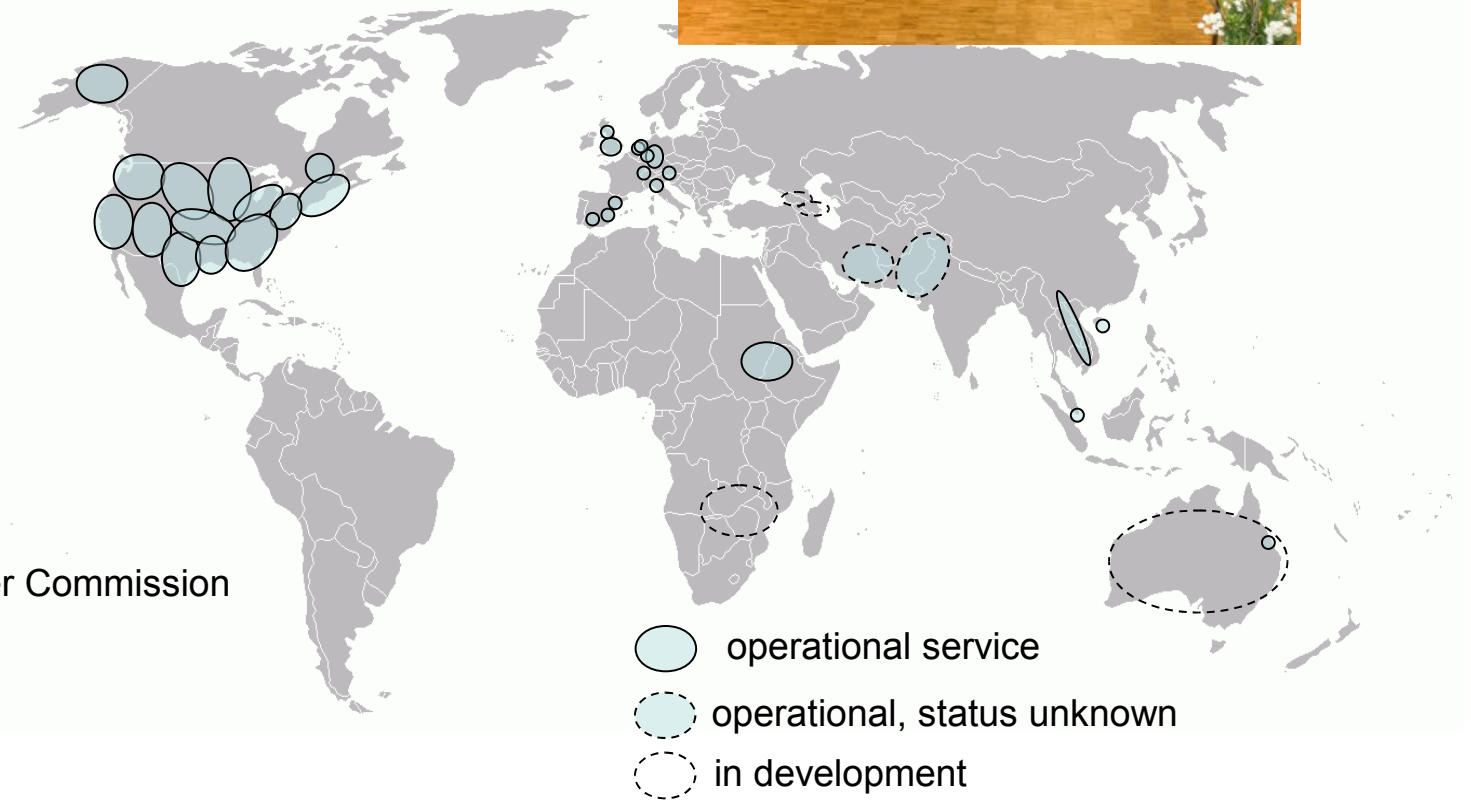


Deltares

Delft FEWS

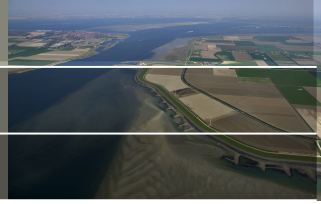
User Community

- USA (NWS)
- Canada
- UK
- Netherlands
- Germany
- Suisse
- Italy
- Austria
- Spain
- Singapore
- Taiwan
- Mekong River Commission
- Australia
- Sudan
- Georgia
- Azerbaijan
- Zambezi



Deltares

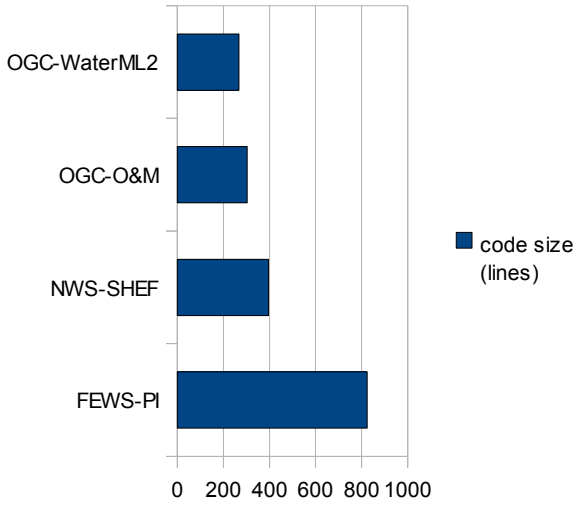
Preliminary results

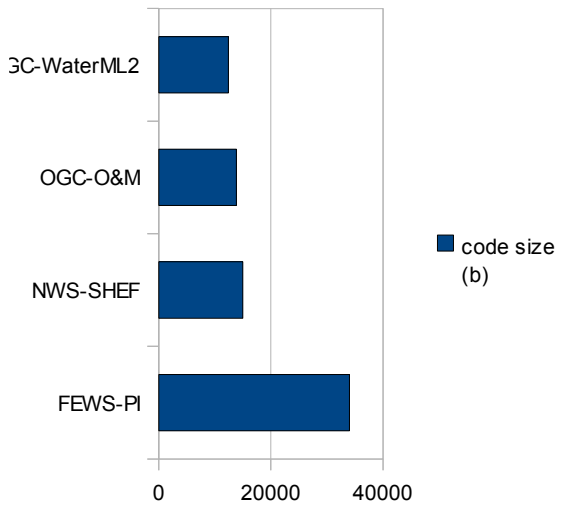


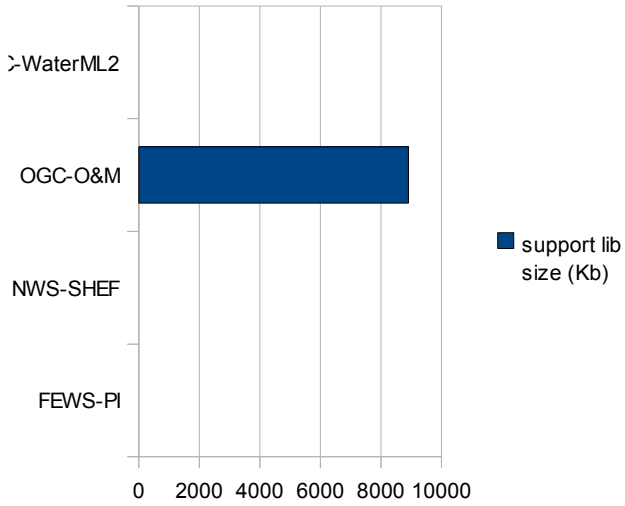
Comparison criteria:

- Lines of source code in client (→ proxy for code complexity)
- File size (→ proxy for network load):
  - > Uncompressed, compressed
- Ingest time (net) required by DelftFEWS

Preliminary results – code complexity of the client



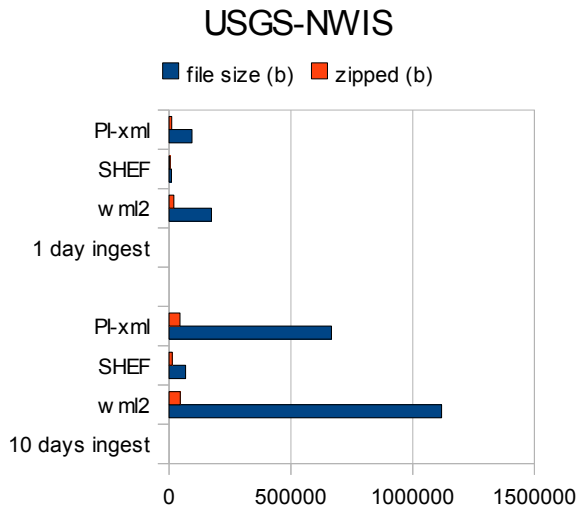


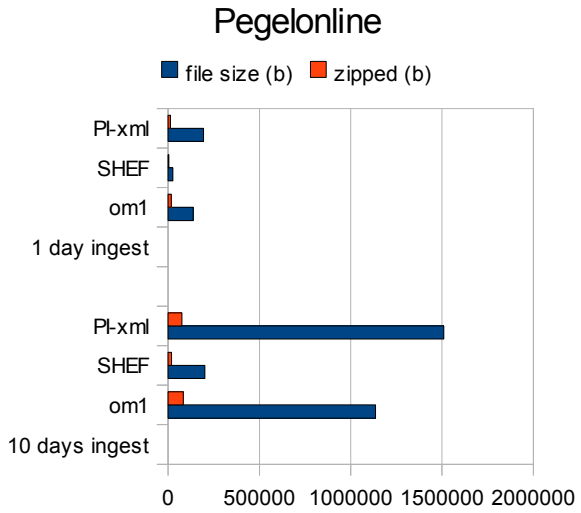


	FEWS-PI	NWS-SHEF	OGC-O&M1	OGC-WaterML2
code size (lines)	824	396	303	267
code size (b)	34025	15005	13871	12420
support lib size (Kb)	0	0	8911	0

Deltares

Preliminary results – File sizes





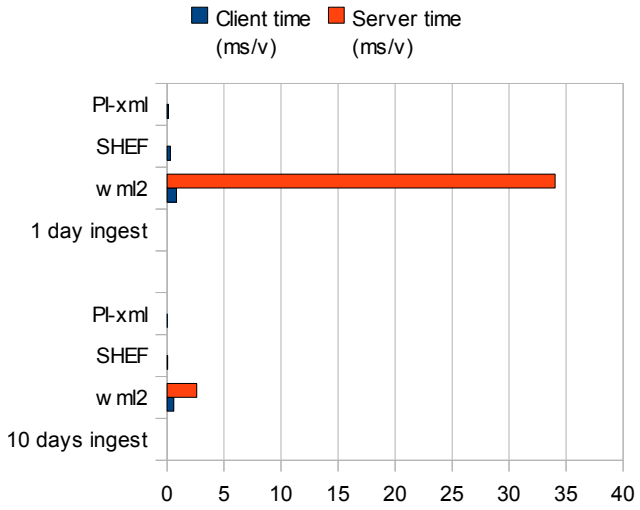
	File size (b)	Zipped (b)	nr.values	nr.series
10 days				
wml2	1119461	46245	8788	10
SHEF	68016	13387	7719	
PI-xml	666820	44145	8656	
1 day				
wml2	174294	19369	885	10
SHEF	9917	4114	875	
PI-xml	93541	10867	1117	

	File size (b)	Zipped (b)	nr.values	Nr series
10 days				
om1	1135933	82822	20787	8
SHEF	202350	19378	20309	
PI-xml	1509931	76389	20320	
1 day				
om1	137378	17129	2058	8
SHEF	26052	4409	2534	
PI-xml	193847	13618	2534	

Deltares

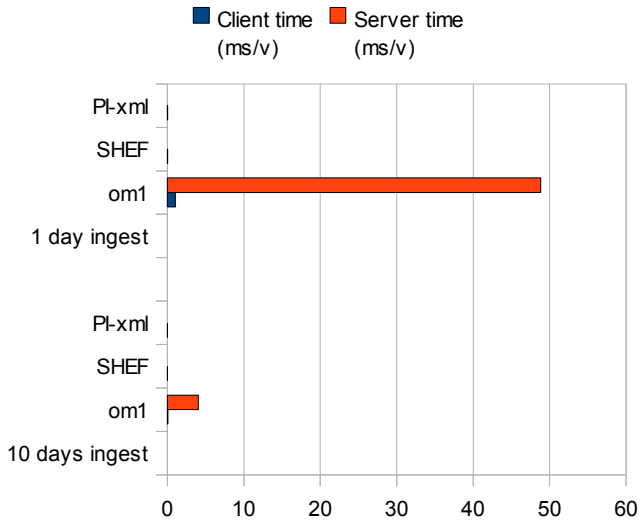
Preliminary results – Ingest performance

USGS-NWIS



10 days	Client (ms/value)	Server (ms/value)	Server (ms)	Client (ms)
wml2	0.594	2.62	23229	5273
SHEF	0.051		0	390
PI-xml	0.023		0	203
1 day				
wml2	0.828	34.04	30124	733
SHEF	0.285		0	249
PI-xml	0.126		0	141

Pegelonline



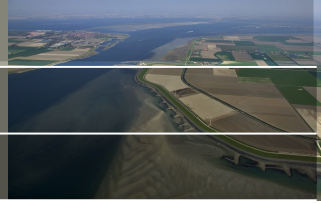
10 days	Client (ms/value)	Server (ms/value)	Server (ms)	Client (ms)
om1	0.071	4.06	84476	1466
SHEF	0.005		0	109
PI-xml	0.008		0	172
1 day				
om1	1.046	48.82	100481	2153
SHEF	0.037		0	93
PI-xml	0.043		0	110

Deltares

Lessons learned

Deltares

Lessons learned



File size

- WaterML2, O&M1-SWEcommon are a bit more volumious than FEWS-PI, especially with short timeseries

Performance

- WaterML2 and O&M1-SWEcommon are significantly slower than SHEF or FEWS-PI
- Efficiency with WaterML2 and O&M1-SWEcommon improves a little with longer timeseries

Binary representations might improve performance,

But ...

Still room for improvement

Deltares