



Delft-FEWS and long-term storage of data:

## > the Deltares OpenArchive

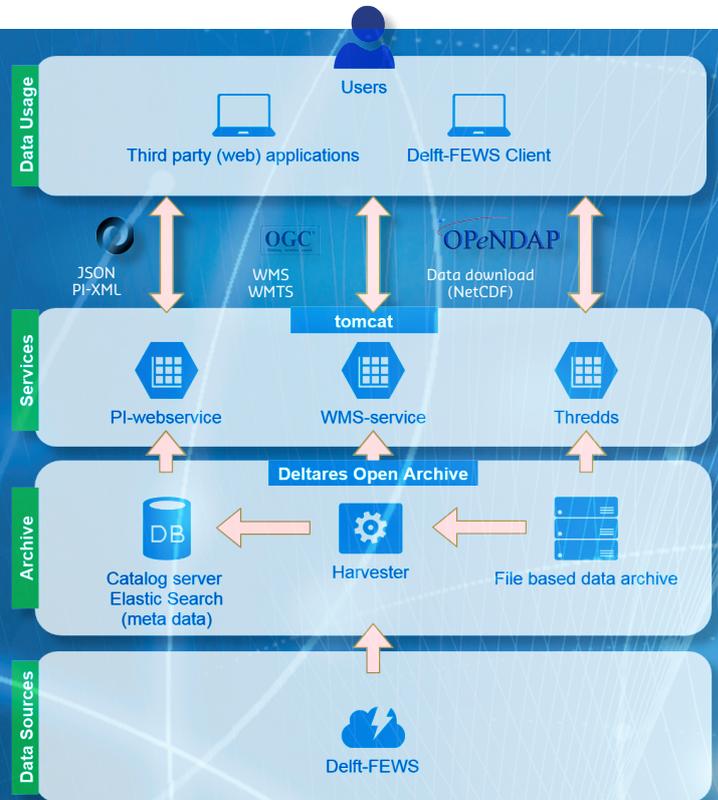


Figure 1: Overview of the Deltares Open Archive components.

### > Long-term storage of data

Storage of data is a basic task in any Water Resources Management activity. Nowadays, data storage requirements are no longer limited to timeseries of measurements, observations and historical model simulations but also include weather forecasts, hydrological/hydrodynamic model forecasts and non-timeseries data such as text reports, communications logs, model initial states and forecasting database snapshots. To allow for long-term storage, management and retrieval of such a variety of data types, Deltares developed the OpenArchive.

The Deltares OpenArchive is a file based storage solution combined with a catalogue for harvesting and indexing the data, and a web service for data access (Figure 1). The OpenArchive intimately integrates with Delft-FEWS applications. At the same time, data can be easily accessed by third party applications using web services, without intervention of a Delft-FEWS application.

### > Example applications

The OpenArchive can contribute to analyses for which there is an inherent requirement to use historic data records. Examples of these are hindcasting (also referred to as retrospective forecasting or reforecasting) and forecast verification. The OpenArchive supports such analyses by allowing external applications to directly source data from the archive.

For example, the US National Weather Services's Ensemble Verification System<sup>1</sup> can perform its verification analyses by directly accessing OpenArchive data through a web service. This eliminates the labour intensive task of making accessible required data in a bespoke location. In a similar fashion, web portals may source data from the archive and visualize this through a web browser.

## The OpenArchive and Delft-FEWS: seamless integration

In the context of Delft-FEWS applications, the OpenArchive is used in combination with a Delft-FEWS operational database. The operational database is designed for near 100% availability and speed of access; this implies that it contains as little data as possible. The OpenArchive, on the other hand, is designed for long-term, high volume storage of data. Delft-FEWS, its database and the OpenArchive integrate seamlessly. By 'seamless' we mean that the Delft-FEWS application can source data from either its operational database or an instance of the OpenArchive, or both, without requiring any changes in the underlying configuration.

Delft-FEWS and the OpenArchive interact in a number of ways. These include export of forecast data to the archive, direct sourcing of data and manual or scheduled download/import of data from the archive.

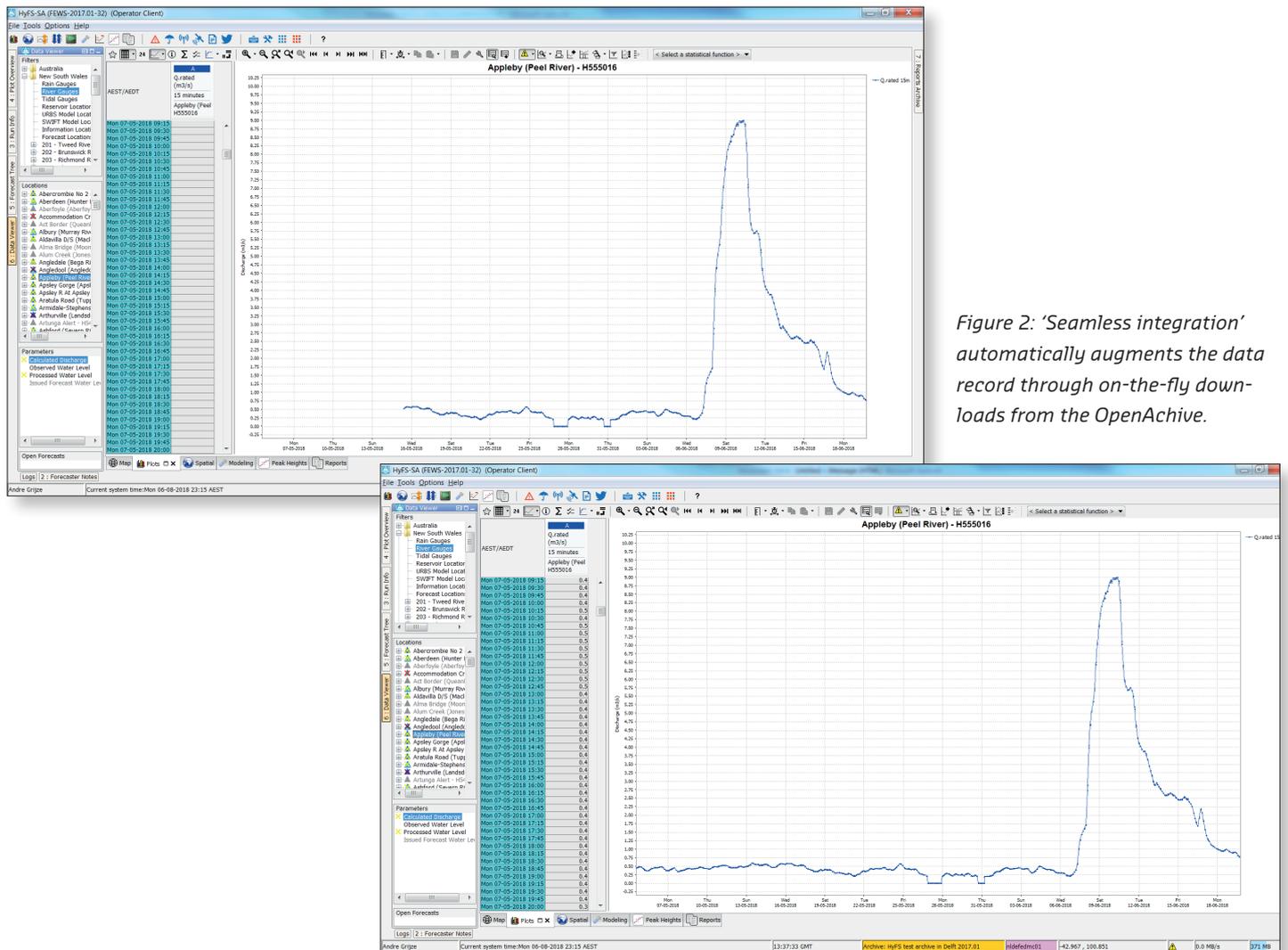


Figure 2: 'Seamless integration' automatically augments the data record through on-the-fly downloads from the OpenArchive.

1 The Ensemble Verification System allows for verifying forecasts against observations. The open source software may be downloaded from <http://amazon.nws.noaa.gov/ohd/evs/evs.html>

First of all, Delft-FEWS can export data to the OpenArchive. Obvious examples include observations from telemetry systems, weather forecasts, and hydrological forecasts and simulations that are produced from within the Delft-FEWS application. In addition, Delft-FEWS can export reports, messages, modifiers, model states and database snapshots. All of this can be re-used in Delft-FEWS at a later time.

The Delft-FEWS timeseries viewer allows for directly showing observations and measurements that are stored in the OpenArchive. This includes the actual variables as well as any metadata that was added when the data resided in the database including quality flags and comments. As of the 2018.02 release, this will be possible with gridded data also.

Other data such as weather forecasts, model states, modifiers as well as model simulations and forecasts can be interactively browsed from within the Delft-FEWS application (Figure 3 and Figure 4). Upon selection, the data is downloaded to the application and can subsequently be imported for use in the application. This mechanism can also be invoked in Delft-FEWS workflows. The workflow then downloads, imports and uses data from the OpenArchive without requiring interactive user intervention.

Note that the OpenArchive integrates with both stand-alone and client-server implementation of Delft-FEWS - often simultaneously.

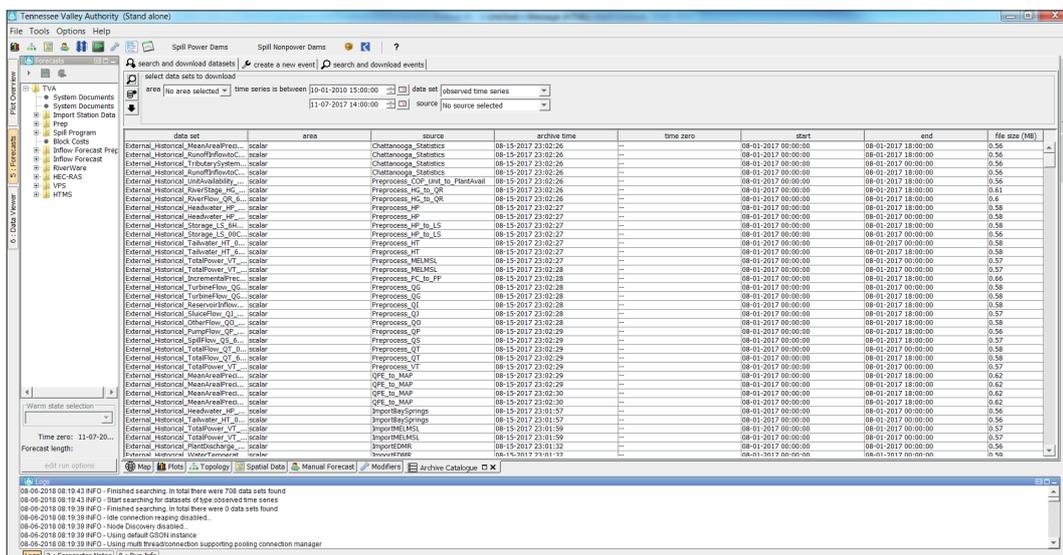


Figure 3: Browsing the OpenArchive from within a Delft-FEWS application: the Archive Catalogue display.

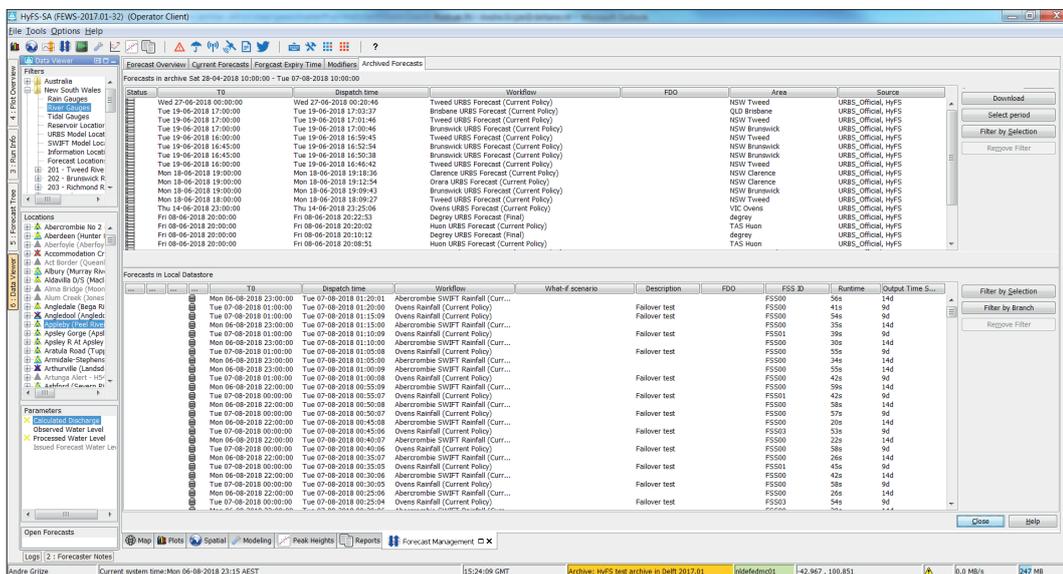


Figure 4: Browsing past forecasts available from the OpenArchive from within a Delft-FEWS application.

## ➤ Third party applications

Data in an OpenArchive can be accessed in a number of ways including through the Delft-FEWS Web Services. These include the PI Web Service which uses either the SOAP or REST protocols and allows for retrieval of scalar timeseries in XML or JSON format.

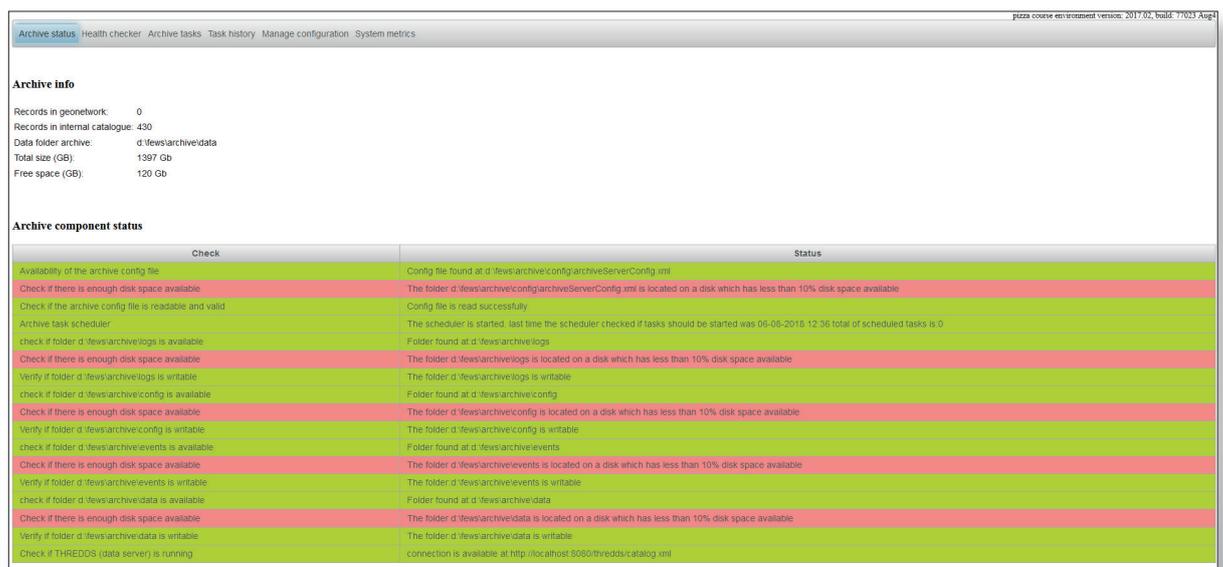
Using this service, third-party applications can use the archived data for pretty much any purpose. For example, web viewers can directly access the data for visualization in a browser. Similar applications include retrieval of data for posting on Twitter as well as disseminating via a Telegram channel.

The web service is instrumental in accessing OpenArchive data in scripting environments such as R, Matlab and Python. Using the PI Web Service, the archive can be searched for the location of requested data. This data can subsequently be accessed by the script using the OpenDAP protocol.

## ➤ Technical details and full documentation

The Deltares Open Archive comprises a number of open source components including Tomcat, Elastic Search and THREDDDS. Schematically, these components combine to form multiple layers including an Archive Layer, a Services Layer and a Data Usage layer (Figure 1). The OpenArchive can be managed through an internet browser (Figure 5).

Additional technical details and full documentation of the OpenArchive are available from the Deltares public WIKI: <https://publicwiki.deltares.nl/display/FEWSDOC/The+Deltares+Open+Archive>



The screenshot shows a web-based dashboard for the OpenArchive. At the top, there are navigation tabs: 'Archive status', 'Health checker', 'Archive tasks', 'Task history', 'Manage configuration', and 'System metrics'. The 'Archive status' tab is active. Below the tabs, there is a section titled 'Archive info' with the following data:

- Records in geonetwork: 0
- Records in internal catalogue: 430
- Data folder archive: d:\fevs\archive\data
- Total size (GB): 1397 Gb
- Free space (GB): 120 Gb

Below this is a section titled 'Archive component status' which contains a table with two columns: 'Check' and 'Status'. The table lists various system checks and their corresponding status messages.

Check	Status
Availability of the archive config file	Config file found at d:\fevs\archive\config\archiveServerConfig.xml
Check if there is enough disk space available	The folder d:\fevs\archive\config\archiveServerConfig.xml is located on a disk which has less than 10% disk space available
Check if the archive config file is readable and valid	Config file is read successfully
Archive task scheduler	The scheduler is started. last time the scheduler checked if tasks should be started was 06-05-2015 12:35 total of scheduled tasks is 0
check if folder d:\fevs\archive\logs is available	Folder found at d:\fevs\archive\logs
Check if there is enough disk space available	The folder d:\fevs\archive\logs is located on a disk which has less than 10% disk space available
Verify if folder d:\fevs\archive\logs is writable	The folder d:\fevs\archive\logs is writable
check if folder d:\fevs\archive\config is available	Folder found at d:\fevs\archive\config
Check if there is enough disk space available	The folder d:\fevs\archive\config is located on a disk which has less than 10% disk space available
Verify if folder d:\fevs\archive\config is writable	The folder d:\fevs\archive\config is writable
check if folder d:\fevs\archive\events is available	Folder found at d:\fevs\archive\events
Check if there is enough disk space available	The folder d:\fevs\archive\events is located on a disk which has less than 10% disk space available
Verify if folder d:\fevs\archive\events is writable	The folder d:\fevs\archive\events is writable
check if folder d:\fevs\archive\data is available	Folder found at d:\fevs\archive\data
Check if there is enough disk space available	The folder d:\fevs\archive\data is located on a disk which has less than 10% disk space available
Verify if folder d:\fevs\archive\data is writable	The folder d:\fevs\archive\data is writable
Check if THREDDDS (data server) is running	connection is available at http://localhost:8080/threads/catalog.xml

Figure 5: The OpenArchive browser-based dashboard gives an overview of the archive's status and allows for changing its settings..