Deltares

enabling delta life

Deltares

enabling delta life

Deltares

Delft-FEWS

Basic Configuration Course

Module: Importing Data

Module Motivation

- We are learning FEWS in order to manage data! (And maybe run some models)
- Importing data is therefore a basic and essential task.
- Many different data types can be imported into Delft-FEWS (Google: Delft-FEWS Import)
- A few elements are needed to set-up a data import, including id mapping.
- Data once imported can be viewed in a variety of ways.

Learning Objectives

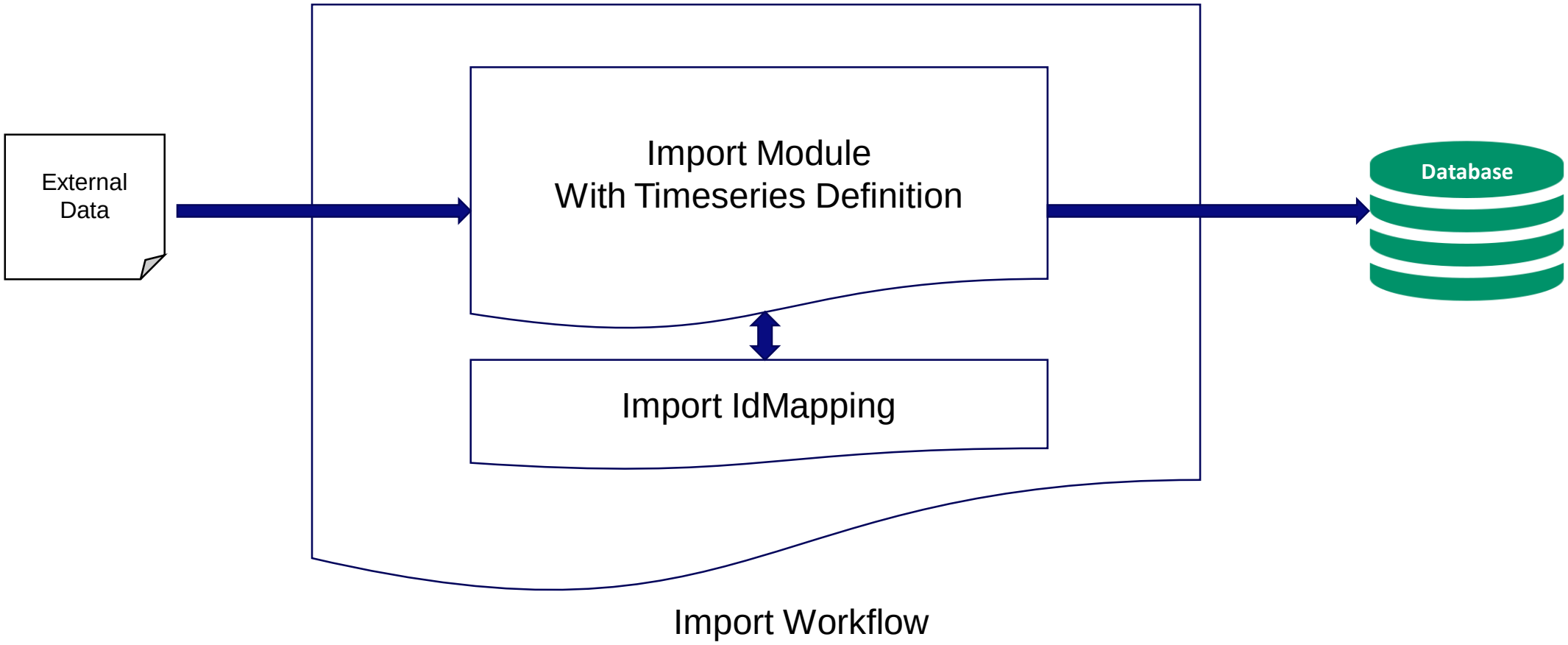
By the end of this module, you will have met the following learning objectives:

1) Have an overview of scalar data types that can be imported into Delft-FEWS

2) Be exposed to the required elements for importing data, including the import module and id mapping.

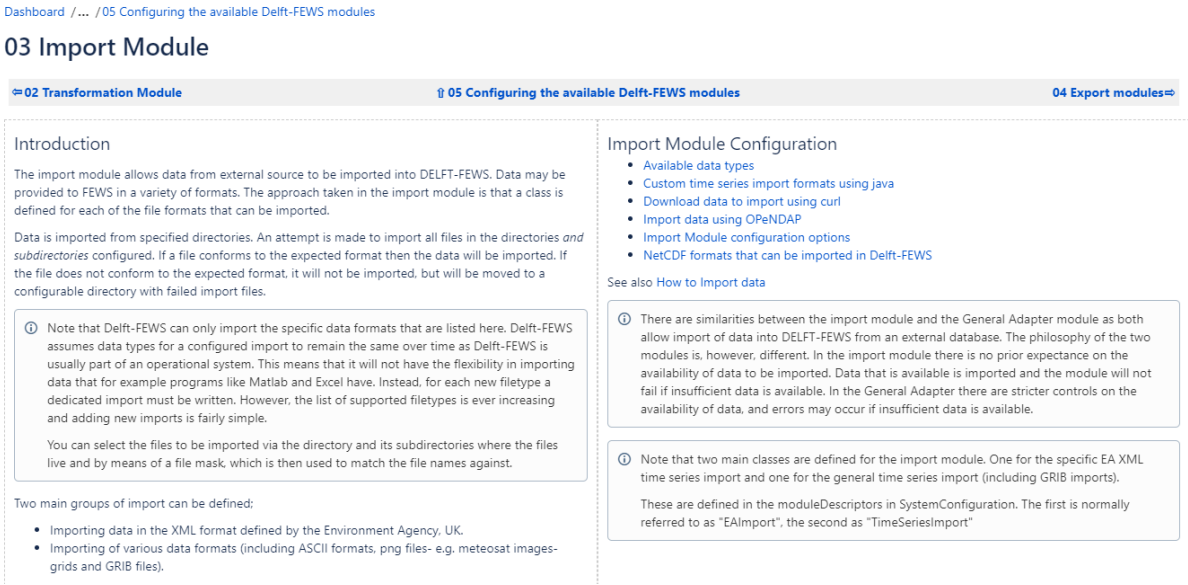
3) Have basic experience setting up an import workflow for scalar data, running it and seeing the results.

Import External Data



Delft-FEWS: Import of external time series

- Delft-FEWS can import large variety of data sources
- Supports standards in data exchange formats: GRIB, NetCDF, csv, xml, etc.
- Emerging standards: WaterML – OpenGIS standard for exchange of hydrological data (USGS, NWS, CUAHSI), OpenDAP, ..
- Plugin-technology to extend integration of data formats
- Google “Delft-FEWS Import” for all available data types



Dashboard / ... / 05 Configuring the available Delft-FEWS modules

03 Import Module

⇦02 Transformation Module | ⇧05 Configuring the available Delft-FEWS modules | 04 Export modules⇨

Introduction

The import module allows data from external source to be imported into DELFT-FEWS. Data may be provided to FEWS in a variety of formats. The approach taken in the import module is that a class is defined for each of the file formats that can be imported.

Data is imported from specified directories. An attempt is made to import all files in the directories *and subdirectories* configured. If a file conforms to the expected format then the data will be imported. If the file does not conform to the expected format, it will not be imported, but will be moved to a configurable directory with failed import files.

Note that Delft-FEWS can only import the specific data formats that are listed here. Delft-FEWS assumes data types for a configured import to remain the same over time as Delft-FEWS is usually part of an operational system. This means that it will not have the flexibility in importing data that for example programs like Matlab and Excel have. Instead, for each new filetype a dedicated import must be written. However, the list of supported filetypes is ever increasing and adding new imports is fairly simple.

You can select the files to be imported via the directory and its subdirectories where the files live and by means of a file mask, which is then used to match the file names against.

Two main groups of import can be defined;

- Importing data in the XML format defined by the Environment Agency, UK.
- Importing of various data formats (including ASCII formats, png files- e.g. meteosat images-grids and GRIB files).

Import Module Configuration

- Available data types
- Custom time series import formats using java
- Download data to import using curl
- Import data using OPeNDAP
- Import Module configuration options
- NetCDF formats that can be imported in Delft-FEWS

See also How to Import data

There are similarities between the import module and the General Adapter module as both allow import of data into DELFT-FEWS from an external database. The philosophy of the two modules is, however, different. In the import module there is no prior expectance on the availability of data to be imported. Data that is available is imported and the module will not fail if insufficient data is available. In the General Adapter there are stricter controls on the availability of data, and errors may occur if insufficient data is available.

Note that two main classes are defined for the import module. One for the specific EA XML time series import and one for the general time series import (including GRIB imports).

These are defined in the moduleDescriptors in SystemConfiguration. The first is normally referred to as "EAImport", the second as "TimeSeriesImport"

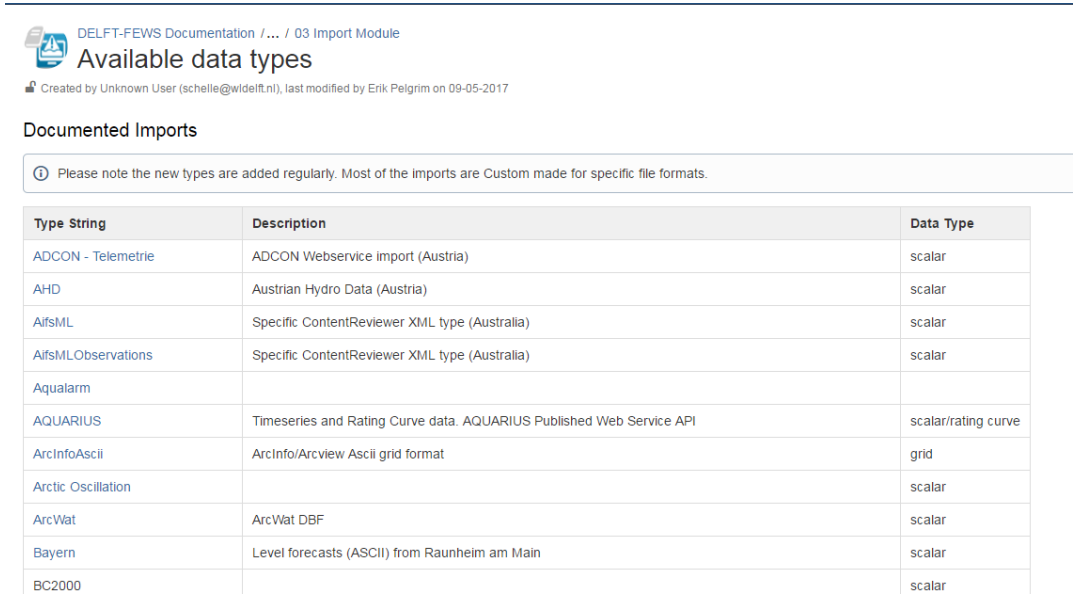
Deltares

Delft-FEWS: Import of external time series

- The FEWS WIKI has documentation on all Import functions in Delft-FEWS
  https://publicwiki.deltares.nl/display/FEWSDOC/Available+data+types

- Import module can import data from
  - Files from import folders
  - Files from ftp servers
  - Files from HTTP
  - Files from Webservices
  - Files from OpenDAP servers
  - Series from databases

- Procedure for data import is same for all data types



DELFT-FEWS Documentation / ... / 03 Import Module

Available data types

Created by Unknown User (schelle@wldelft.nl), last modified by Erik Pelgrim on 09-05-2017

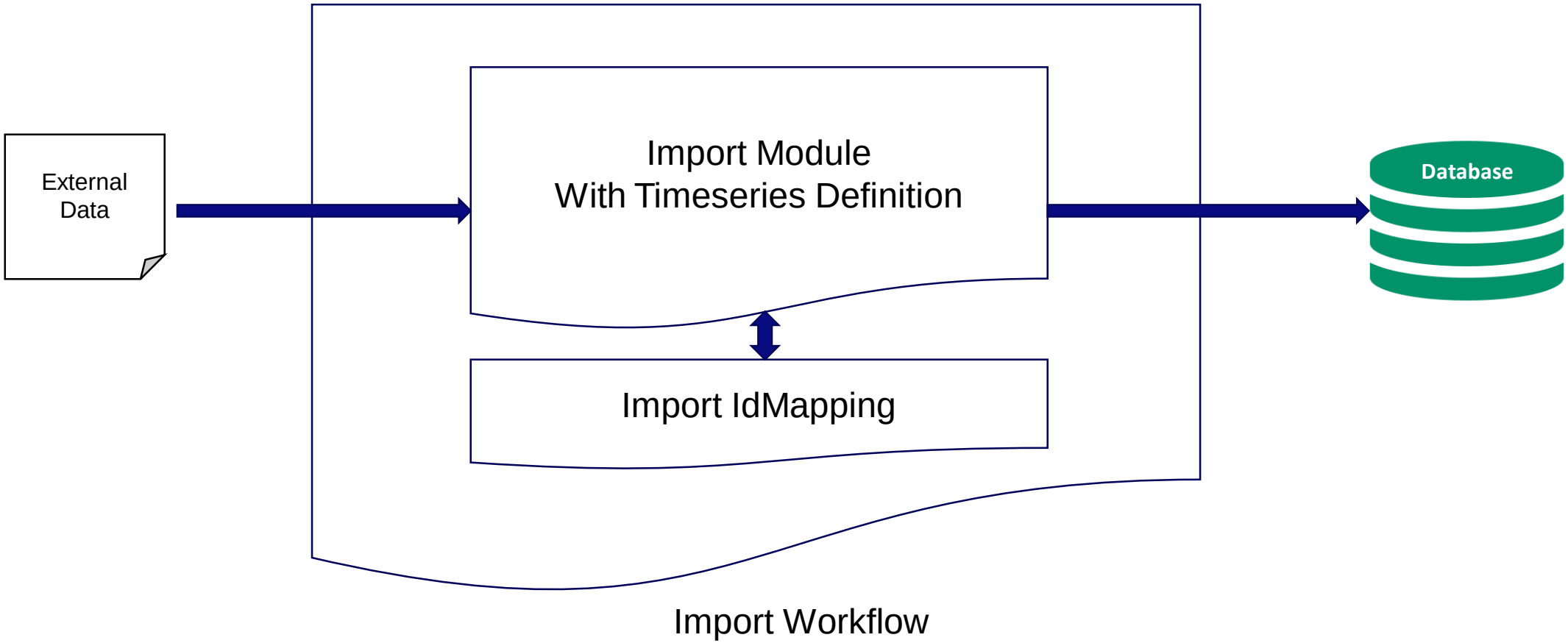
Documented Imports

Please note the new types are added regularly. Most of the imports are Custom made for specific file formats.

Type String	Description	Data Type
ADCON - Telemetrie	ADCON Webservice import (Austria)	scalar
AHD	Austrian Hydro Data (Austria)	scalar
AifsML	Specific ContentReviewer XML type (Australia)	scalar
AifsMLObservations	Specific ContentReviewer XML type (Australia)	scalar
Aqualarm		
AQUARIUS	Timeseries and Rating Curve data. AQUARIUS Published Web Service API	scalar/rating curve
ArcInfoAscii	ArcInfo/Arcview Ascii grid format	grid
Arctic Oscillation		scalar
ArcWat	ArcWat DBF	scalar
Bayern	Level forecasts (ASCII) from Raunheim am Main	scalar
BC2000		scalar

Deltares

Import External Data



Delft-FEWS: Data import module

- A configured Data Import module instance is required for each file format
- A workflow calls the module instances to import the data

- To import data, configuration is defined in:
  - \ModuleConfigFiles\Import\ImportForecast.xml (Import instructions)
  - \IdMapFiles\IdImport<TYPE>.xml (Id mapping instructions)
  - \RegionConfigFiles\ModuleInstanceDescriptors.xml (Registers import module instance)
  - \WorkflowFiles\ImportForecastGrids.xml (list with import module instances activities)

- Also more generally
  - %REGION_HOME%\global.properties (location of the import folder)
  - \RegionConfigFiles\LocationSets.xml (locations that contain time series)

Delft-FEWS: Data import module

- Data Import Module has special functions to read the specific file formats
- Data Import Module configuration tells the software:
  - What file format to import
  - Where to find the data files
  - How to translate the information in the files to internal Ids
  - How to store the imported time series in the database

```
<timeSeriesImportRun xmlns="http://www.wldelft.nl/fews" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.wldelft.nl/fews http://fews.wldelft.nl/schemas/version1.0/timeSeriesImportRun.xsd">
  <import>
    <general>
      <importType>csv</importType>
      <folder>$IMPORT_FOLDER$/WSC</folder>
      <failedFolder>$IMPORT_FAILED_FOLDER$/WSC</failedFolder>
      <backupFolder>$IMPORT_BACKUP_FOLDER$/WSC</backupFolder>
      <idMapId>IdImportWSC</idMapId>
      <unitConversionsId>ImportUnitConversions</unitConversionsId>
    </general>
    <timeSeriesSet>
      <moduleInstanceId>ImportWSC</moduleInstanceId>
      <valueType>scalar</valueType>
      <parameterId>QR.obs</parameterId>
      <locationSetId>$MODELNAME1$HydroStations</locationSetId>
      <timeSeriesType>external historical</timeSeriesType>
      <timeStep unit="day" multiplier="1"/>
      <readWriteMode>read only</readWriteMode>
    </timeSeriesSet>
  </import>
</timeSeriesImportRun>
```

IdMap Files

- IdMaps are defined to map internal location and parameter ID's to external location and parameter ID's. The configuration of these can be done in two ways.

1) Separate mappings can be defined for the locations and for the parameters. Although this is the most efficient method, it is not suitable in all cases, as specific locations may require a different mapping.

```
<map internalLocation="RDPA" internalParameter="PC.sim" externalLocation="none" externalParameter="RDPA.6F_PR"/>
```

2) A definition can be created where the mapping is done on the basis of the unique combination of location/parameter.

```
<parameter internal="TA.obs" internalQualifier="max" external="air_temperature_yesterday_high"/>
<parameter internal="TA.obs" internalQualifier="min" external="air_temperature_yesterday_low"/>
```

Each IdMap configuration may only use one method of defining mappings to avoid ambiguity.

Module Summary

- A variety of data types can be directly imported into Delft-FEWS, either directly from your local disk or from online sources.

- For standardized formats (NetCDF, WaterML2, grib2), pre-defined imports have been built that can import data directly.

- Data import can requires configuration of several files. Learning what files need to be changed is learning Delft-FEWS configuration

Next Steps

- Once data is imported, it often needs to be validated and processed.
- Delft-FEWS has very rich functionality for both data validation and processing.
- In this course, we will look at processing with transformations.
- Once learned, even complicated transformations can be implemented quite simply.
- It helps to think sequentially!

Additional Resources

- Google “Delft-FEWS WIKI”
- Google “Delft-FEWS Configuration Guide”
- Google “Delft-FEWS Forum”
- Email fews-pm@Deltares.nl