

22-1 Datasets of the Deltares Open Archive

- [Observations](#)
 - [Description of the Observations dataset](#)
 - [Time Series Metadata](#)
- [Messages](#)
 - [Description of Messages dataset](#)
 - [Messages Metadata](#)
- [External forecasts](#)
 - [Description of external forecasts dataset](#)
 - [Time Series Metadata](#)
- [Simulations](#)
 - [Data set description](#)
 - [Simulations Metadata](#)
- [Configuration](#)
 - [Description of Configuration dataset](#)
 - [Configuration metadata](#)
- [Rating curves](#)
 - [Description of Rating Curves dataset](#)
 - [Rating Curves metadata](#)
- [Snap shot](#)
 - [Description of Snap shot dataset](#)
 - [Snap shot metadata](#)

Observations

Description of the Observations dataset

An Observations dataset holds one data type: time series, stored in NetCDF files.

Observations are historical time series data, continuous over time, with a regular or irregular time step. Typically these time series are scalar (0-dimensional in space). At each moment in time only one numerical value applies per station and variable. For scalar time series, a quality flag and string comment can be included for individual time stamps

This type of dataset consists of one or more NetCDF files and a 'metaData.xml'. The NetCDF files contain the observation timeseries and the 'metaData.xml' file contains meta-data about the content of the NetCDF files.

Observation timeseries are stored in NetCDF (1.4) using the data structure and metadata definitions as prescribed in the CF metadata convention (<http://cfconventions.org/>). Multiple stations and variables can be included in one netcdf file.

Time Series Metadata

The metadata file for this data set follows the schema definition as defined in:
<http://fewswildelft.nl/schemas/version1.0/archive-schemas/netcdfMetaData.xsd>.
 See Figure 3.1 and Table 3.1 for the detailed explanation.

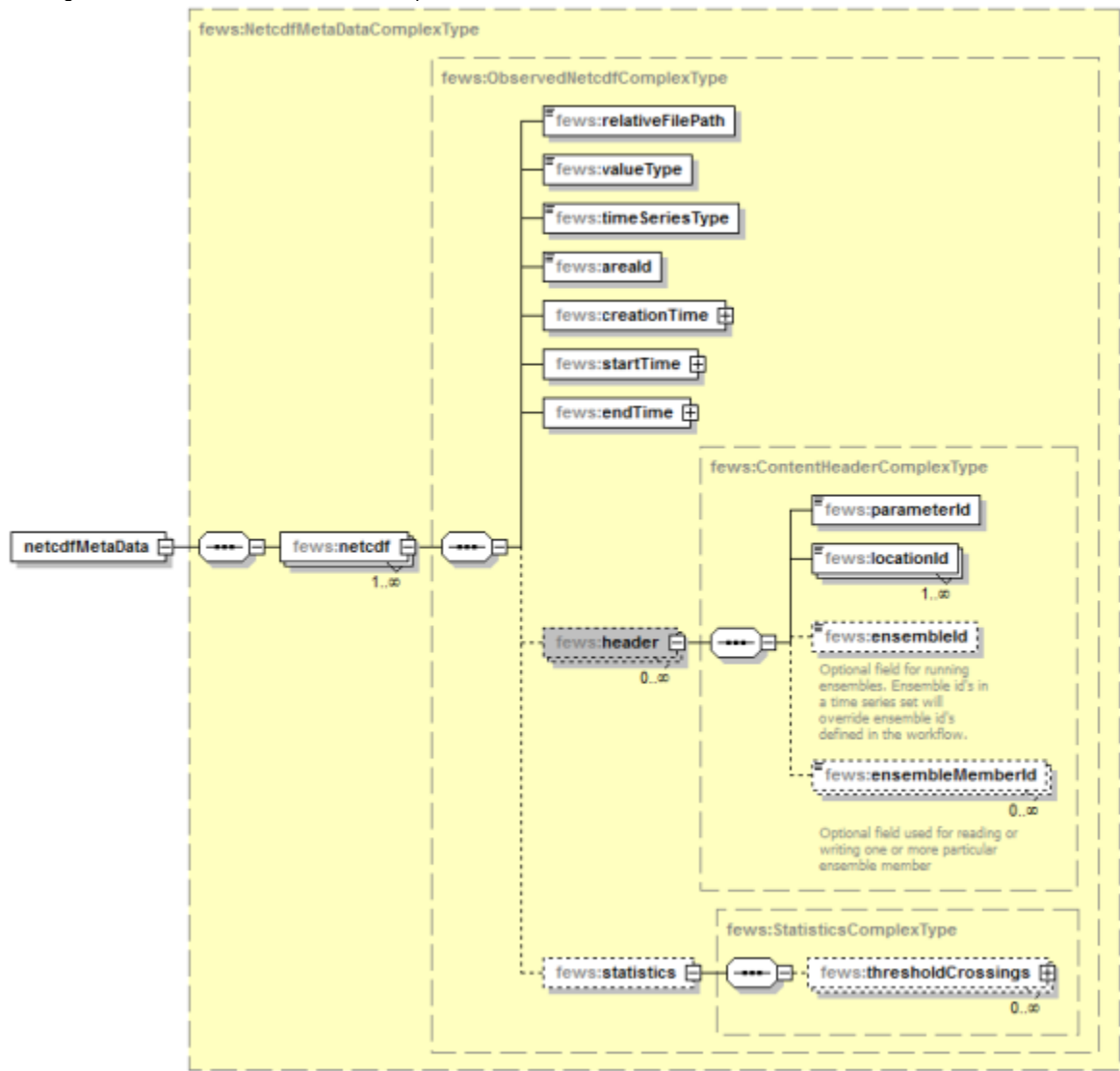


Figure 3.1 Metadata describing the contents of a netcdf file

Table 3.1Metadata content of time series (Netcdf ComplexType)

Element Name	Format	Description
relativeFilePath	string	file path relative from metadata file
valuetype	enum(1): grid scalar	Determines appropriate netcdf parser
timeSeriesType	enum (1): observed simulated externalForecast	Determines directory structure
areaid	string	area reference, Use no space, will become part of directory path
creationTime	attributes date time	Creation time of data set as posted to archive
startTime	attributes date time	Start time of time series held in file
endTime	attributes date time	End time of time series held in file
header	ContentHeader ComplexType	
header-parameterid	string, use underscore (_) no dot (.)	variable name in NetCDF header

header-locationId	string, use underscore (_) no dot (.)	station name in NetCDF header
header-ensembleId	string	ensemble/realization identifier
header-ensembleMemberId	string	ensemble/realization member identifier
statistics	Statistics ComplexType	
statistics-thresholdCrossings	ThresholdsCrossings ComplexType	
thresholdCrossings-header	StatisticsHeader ComplexType	just one location, rest same as normal ContentHeader
thresholdCrossings-thresholdCrossing	ThresholdsCrossing ComplexType	
thresholdCrossing-thresholdId	string	identifier of threshold crossed
value	double	threshold value crossed

Messages

Description of Messages dataset

The Messages dataset holds one data type: text, stored in text files.

Typically, all text from the same origin (e.g. system logs, model logs or Forecaster Notes) are collected together and stored in one file. Generally, each message in such file has a time stamp associated with it. All messages are collated by day and stored by day. This type of dataset consists of one or more text files and a metaData.xml file.

Messages Metadata

The metadata file for this data set follows the schema definition as defined in:

<http://fews.wdelft.nl/schemas/version1.0/archive-schemas/messagesMetaData.xsd>.

The dataset may hold multiple message files, where each file is described by its metadata (see Figure 3.2 and Table 3.2).

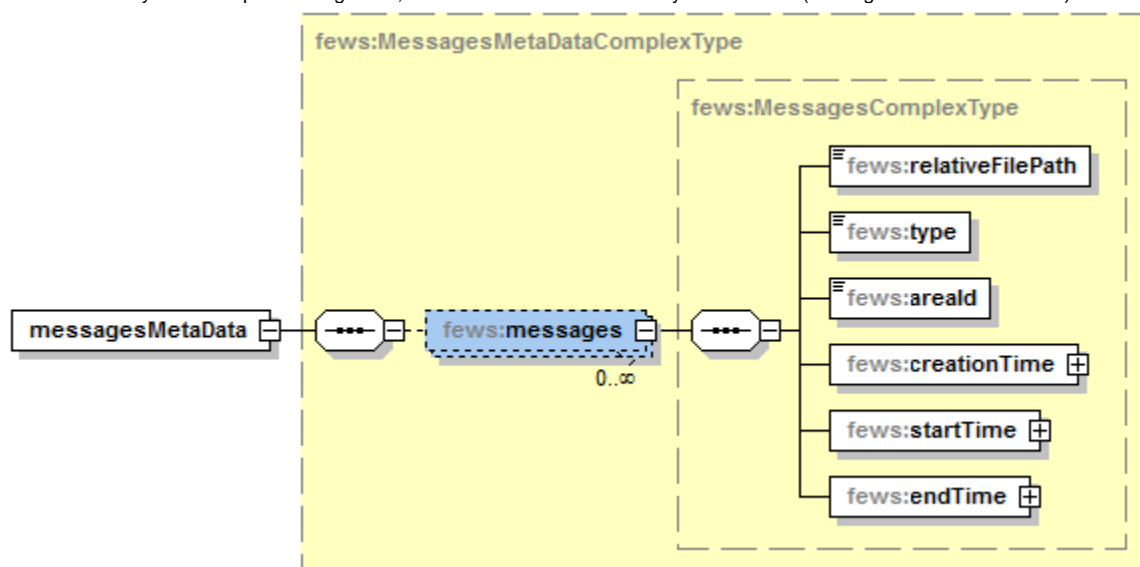


Figure 3.2 Metadata for messages according to messagesMetaData.xsd

Table 3.2 Metadata content of messages

Element Name	Format	Description
relativeFilePath	string	file path relative from metadata file
type	string	as understood by software
areald	string	area reference for which message applies Use no space, will become part of directory path
creationTime	attributes date time	Creation time of dataset as posted to archive

startTime	attributes date time	Start time of messages held in file
endTime	attributes date time	End time of messages held in file

External forecasts

Description of external forecasts dataset

The external forecasts dataset holds one data type: time series, stored in NetCDF files.

External forecasts refer to forecast time series from external sources such as UK MetOffice, KNMI, NCEP, ECMWF etc. Characteristic of a forecast time series is the fact that multiple versions may exist at a particular point in time. Each version is identified by its reference or forecast time, e.g. the 11am forecast. The time stamp when the value applies is typically called valid time. Given that hydrological operational forecasting systems may use multiple NWP products, which all share the same reference time, separate datasets need to be archived by NWP source.

External Forecast timeseries are stored in NetCDF (1.4) using the data structure and metadata definitions as prescribed in the CF metadata convention (<http://cfconventions.org/>). For gridded timeseries, each file contains one grid and one or more variables. For scalar time series, multiple stations and variables can be included in one netcdf file.

This type of dataset is accompanied with a metaData.xml file.

Time Series Metadata

The metadata file for this data set follows the schema definition as defined in:

<http://fews.wdelft.nl/schemas/version1.0/archive-schemas/netcdfMetaData.xsd>.

This is the same metadata definition as discussed for the observations (see Figure 2.1).

The content of the metadata file differs in the choice of timeSeriesType (externalForecast), while the valueType will be 'grid' in most situations. Currently, no statistics can be stored in the metadata for external forecasts.

Simulations

Data set description

General

The Simulations data set may hold a variety of data types, e.g.

- time series, stored in NetCDF files
- model states, stored in native model/application format
- run info in xml format
- model run settings stored in native application format (e.g. FEWS modifiers in PI-xml)
- reports stored in document formats (e.g. HTML, PDF)

Characteristic to this dataset is its origin: a simulation. This simulation could be a State updating run or a Forecast run in a forecasting application (e.g. Delft-FEWS), or it could be a reference or scenario run by any other simulator (e.g. a Python script or model application like DeltaShell).

The dataset generally holds every data item which is needed for full traceability of the result. This includes the time series as generated by simulation, possibly the time series that are the boundary conditions for the simulation, the model state (if deviating from the default), model run settings, including any deviations from the default, etc. If the dataset holds the 'default' or reference, it may even include the entire model setup.

Timeseries are stored in NetCDF (1.4) using the data structure and metadata definitions as prescribed in the CF metadata convention (<http://cfconventions.org/>). For gridded timeseries, each file contains one grid and one or more variables. For scalar time series, multiple stations and variables can be included in one netcdf file.

Model run settings are stored in native application format.

Reports are stored in HTML, ZIP or PDF format.

Delft-FEWS specific files in the simulations datasets

For Delft-FEWS the run settings are defined in the runinfo.xml file, which meets the schema as defined in <http://fews.wldelft.nl/schemas/version1.0/archive-schemas/runinfo.xsd>.

As can be seen in Figure 3.3, this runinfo file includes information on the configuration version used to create the results.

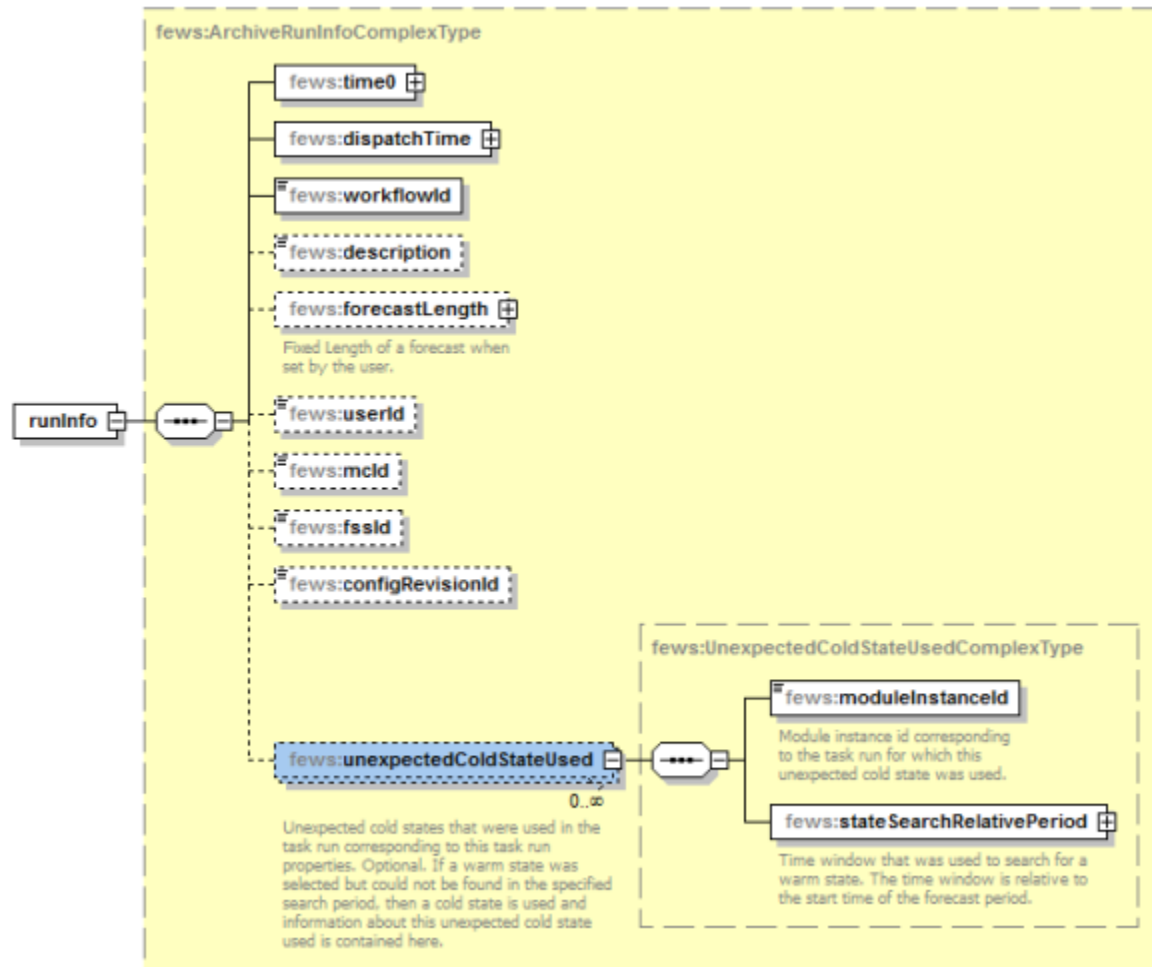
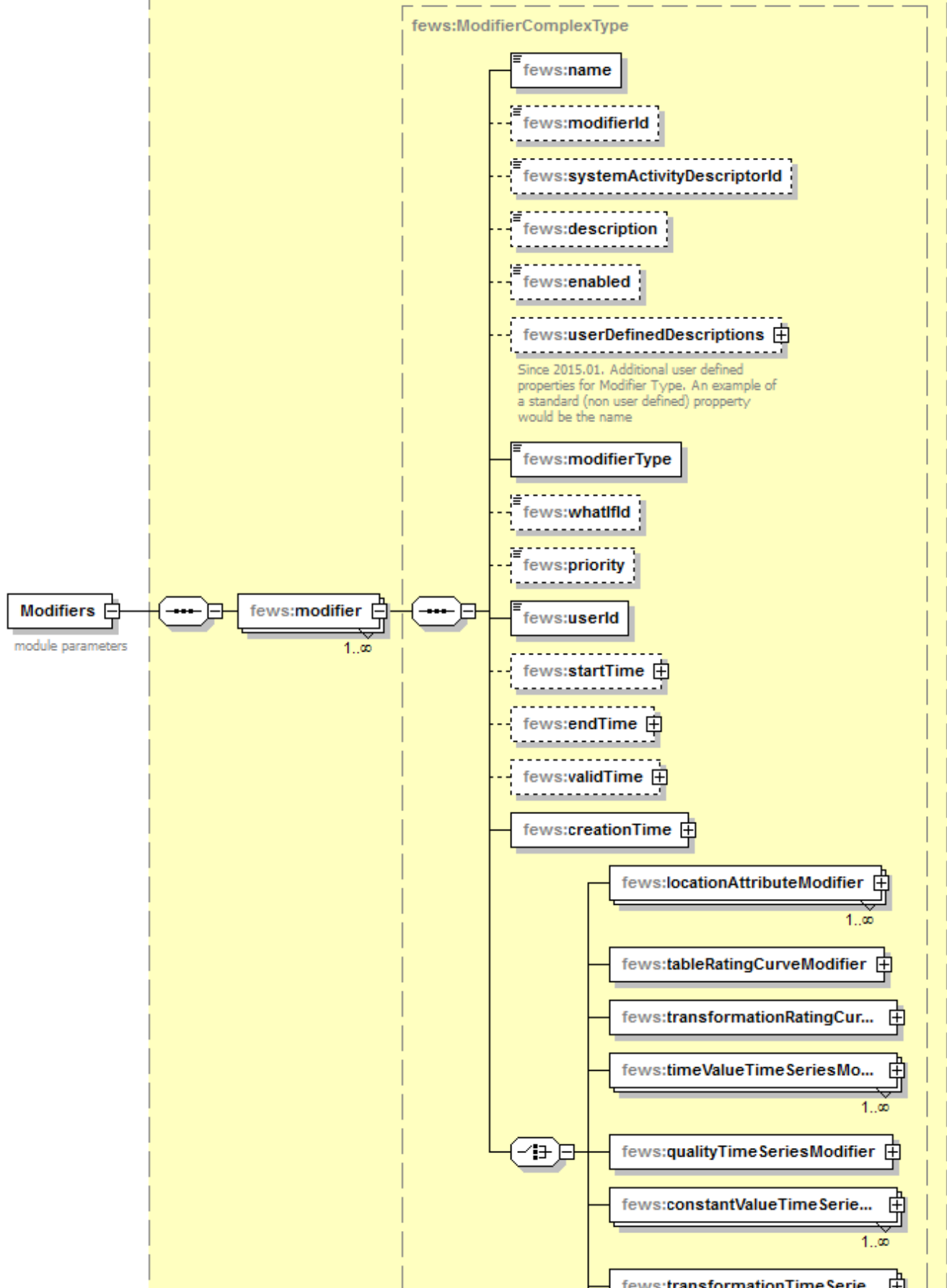


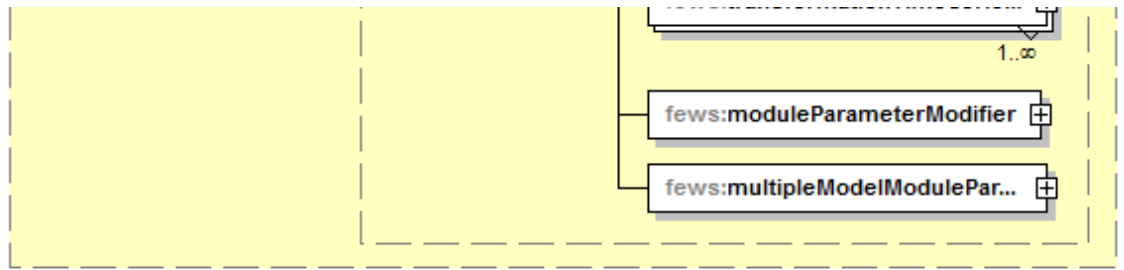
Figure 3.3 Delft-FEWS runinfo.xsd

Modifiers used in this run are stored in the modifiers.xml, which follows a FEW PI-schema: http://fews.wldelft.nl/schemas/version1.0/pi-schemas/pi_modifiers.xsd.

At this moment all modifiers types are supported.

fews:ModifiersComplexType





Generated by XMLSpy

www.altova.com

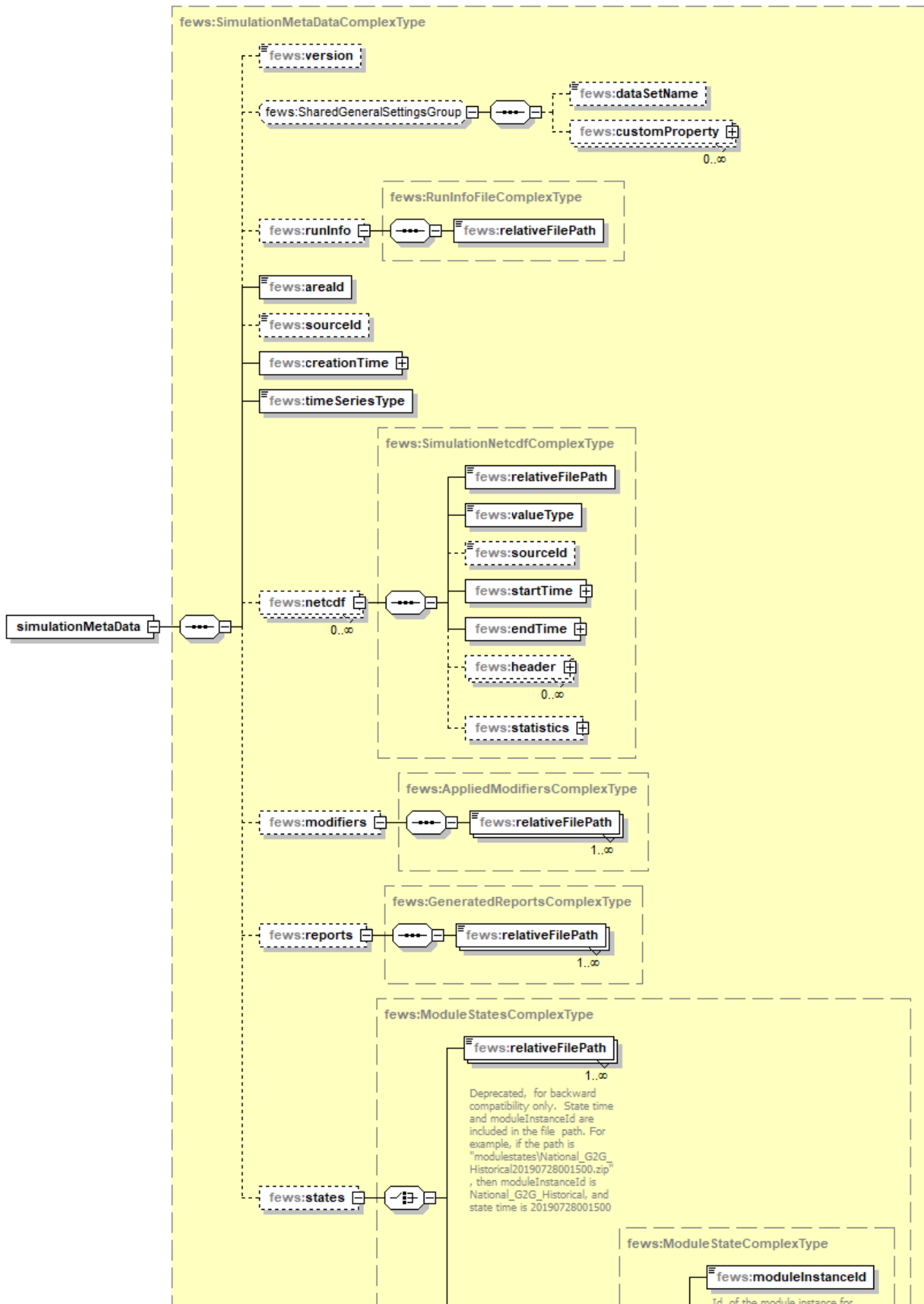
Figure 3.4 Delft-FEWS PI-schema modifiers.xsd

Simulations Metadata

The metadata file for this data set follows the schema definition as defined in:

<http://fews.wldelft.nl/schemas/version1.0/archive-schemas/simulationMetaData.xsd>.

As can be seen in Figure 3.5, the associated metadata file holds references to each individual portion of the dataset



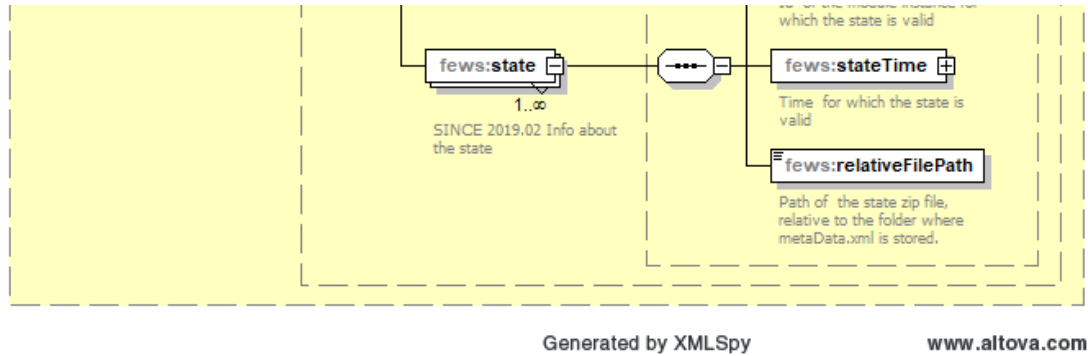


Figure 3.5 Metadata for simulation datasets according to simulationMetaData.xsd

Table 3.3 Metadata content for simulation

Element Name	Format	Description
runInfo	RunInfoFile ComplexType	
runInfo-relativeFilePath	string	path to runInfo file relative from metadata file
areald	string	area reference, Use no space, will become part of directory path
creationTime	attributes date time	Creation time of data set as posted to archive
timeSeriesType	enum (1): observed simulated externalForecast	Determines directory structure. Typically simulated
netcdf	simulationNetCDF ComplexType	subset of the normal netCDF ComplexType. For details see section 3.1.2
modifiers	AppliedModifiers Complextype	
modifiers-relativeFilePath	string	path(s) to modifier data file(s) relative from metadata file
reports	GeneratedReports ComplexType	
reports-relativeFilePath	string	path(s) to report zip file(s) relative from metadata file
states	ModuleStates ComplexType	
states-relativeFilePath	string	path(s) to module state file(s) relative from metadata file (Deprecated)
states-state	ModuleStatesComplexType	list of module state file descriptions
state-moduleInstanceld	string	Id of the module instance for which the state is valid
state-stateTime	attributes date time	Time for which the state is valid
state-relativeFilePath	string	Path of the state zip file, relative to the folder where metaData.xml is stored

Configuration

Description of Configuration dataset

The Configuration dataset is a zip file holding a Delft-FEWS configuration or complete model setup. The Configuration is identified by a revision identifier. For Delft-FEWS, the revision identifier is the Master Controller revision identifier. For non-FEWS configurations, any other unique identifier could be used, e.g. an SVN-revision number.

The dataset is as accompanied by a metadata.xml

Configuration metadata

The metadata file for this data set follows the schema definition as defined in:
<http://fews.wldelft.nl/schemas/version1.0/archive-schemas/configMetaData.xsd>.
The dataset may hold one zip file, described by its metadata (see Figure 3.6 and Table 3.4).

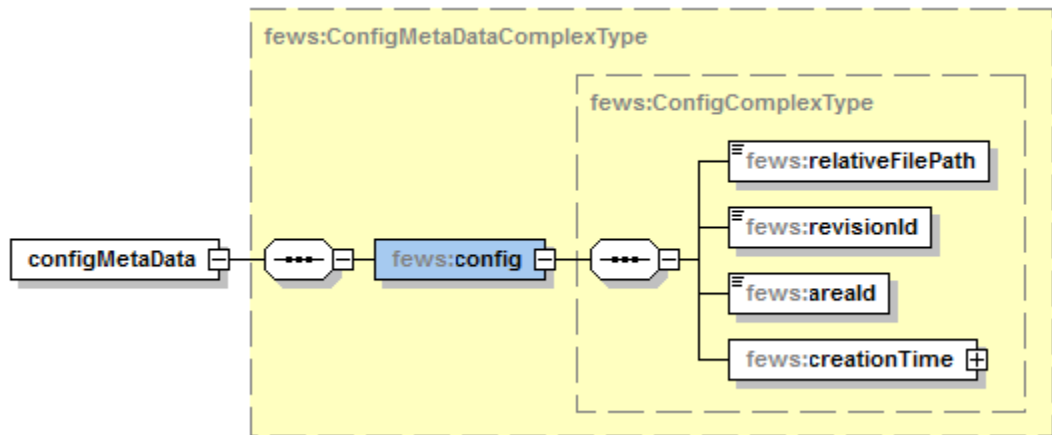


Figure 3.6 Metadata for configurations according to configMetaData.xsd

Table 3.4 Metadata content for configuration

Element Name	Format	Description
relativeFilePath	string	path to configuration zip file relative from metadata file
revisionId	string	Identifier/version number
arealD	string	area reference for which message applies Use no space, will become part of directory path
creationTime	attributes date time	Creation time of dataset as posted to archive

Rating curves

Description of Rating Curves dataset

The rating curves dataset holds a set of rating curve files accompanied by a metaData.xml file. Any format can be chosen but Deltares software supports the Delft-FEWS PI_ratingcurves.xsd schema definition using a rating table. For details check: http://fews.wldelft.nl/schemas/version1.0/pi-schemas/pi_ratingcurves.xsd.

Rating Curves metadata

The metadata file for this data set follows the schema definition as defined in:
<http://fews.wldelft.nl/schemas/version1.0/archive-schemas/ratingCurvesMetaData.xsd>.
The dataset may hold multiple rating curve file, described by its metadata (see Figure 3.7 and Table 3.5).

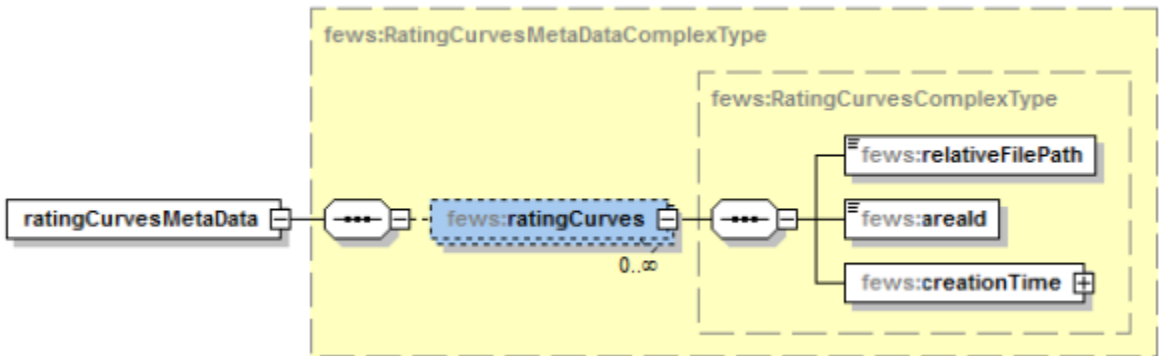


Figure 3.7 Metadata for rating curves according to ratingCurvesMetaData.xsd

Table 3.5 Metadata content for rating curves

Element Name	Format	Description
relativeFilePath	string	file path relative from metadata file
areald	string	area reference in which rating curves apply Use no space, will become part of directory path
creationTime	attributes date time	Creation time of dataset as posted to archive

Snap shot

Description of Snap shot dataset

The snap shot dataset holds a complete snap shot of the model database with all its data. Within Delft-FEWS this is equal to a local datastore in Firebird format as available on a Forecasting Shell Server. This dataset ensures that ALL information including its history, as held in the system for a particular run, is archived. Snap shots are particular important when reviews or legal inquiries are of concern to the operator. When snapshots are made at a high frequency, duplication of data is likely to occur.

Snap shot metadata

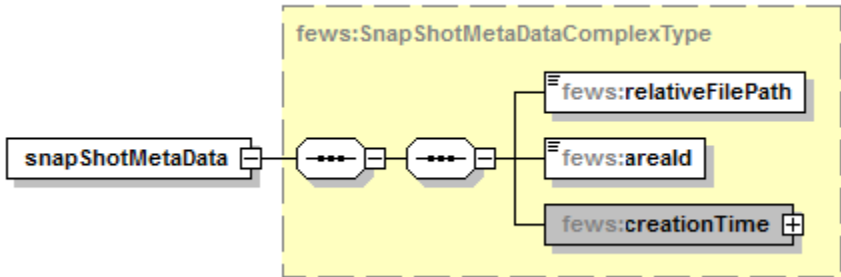


Figure 3.8 Metadata for snapshots according to snapshotMetaData.xsd

Table 3.6 Metadata content for snapshot

Element Name	Format	Description
relativeFilePath	string	file path relative from metadata file
areald	string	area reference in which rating curves apply Use no space, will become part of directory path
creationTime	attributes date time	Creation time of dataset as posted to archive