

08 Error Correction Module (ARMA)

What	nameofinstance.xml
Description	Configuration o the ARMA module
schema location	https://fewdocs.deltares.nl/schemas/version1.0/errorModelSets.xsd

- [AR Module Background information](#)
- [Error correction module configuration](#)
 - [errorModelSet](#)
 - [inputVariable](#)
 - [TimeSeriesSet](#)
 - [autoOrderMethod](#)
 - [orderSelection](#)
 - [order_ar](#)
 - [order_ma](#)
 - [parameters](#)
 - [subtractMean](#)
 - [boxcoxTransformation](#)
 - [lambda](#)
 - [ObservedTimeSeriesId](#)
 - [SimulatedTimeSeriesId](#)
 - [OutputTimeSeriesId](#)
 - [fixedOrderMethod](#)
 - [correctionModel](#)
 - [order_ar](#)
 - [order_ma](#)
 - [ObservedTimeSeriesId](#)
 - [SimulatedTimeSeriesId](#)
 - [OutputTimeSeriesId](#)
 - [interpolationOptions](#)
 - [interpolationType](#)
 - [gapLength](#)
 - [defaultValue](#)
 - [maxObserved](#)
 - [minObserved](#)
 - [maxResult](#)
 - [minResult](#)
 - [ignoreDoubtful](#)
 - [ignoreTrailingMissingsInSimulatedTimeSeries](#)
 - [loopOverMultipleTimeSeries](#)
 - [logLevelNoObservedValues](#)
 - [outputVariable](#)
 - [logLevelCoefficientsInfo](#)

Error correction module configuration

The error modelling module is a generic forecasting module. The module is used to improve the reliability of forecast by attempting to identify the structure of the error a forecasting module makes during the modelling phase where both the simulated and observed values are available, and then applying this structure to the forecast values. This is under the assumption that the structure of the error remains unchanged. A description of the background of this module can be found at [AR Module Background information](#). In defining the error model three time series will need to be defined;

- Merged input time series of simulated model output for the historical period and of forecasted model output for the forecast period. The time series in the historical period will be used for establishing error model through comparison with the observed time series. The error forecast will be applied to the time series in the forecast period.
- Input time series for the observed data.
- Output time series for the updated simulated data for the historical period and the updated forecast data for the forecast period.

Two methods of establishing an error model are available. The first uses an AR (Auto Regressive) model only, but allows the order of the model to be determined automatically. The second method uses an ARMA model, but the order of both the AR and the MA (Moving Average) model must be defined. In both cases various transformations may be applied to normalise the residuals prior to establishing the error model.

When available as configuration on the file system, the name of the XML file for configuring an instance of the error module called for example GreatCorby_ErrorModel_Forecast may be:

GreatCorby_ErrorModel_Forecast 1.00 default.xml

GreatCorby_ErrorModel_Forecast	File name for the GreatCorby_ErrorModel_Forecast configuration.
--------------------------------	---

1.00	Version number
default	Flag to indicate the version is the default configuration (otherwise omitted).

The observed input time series may contain missing values. If this is the case, then the error model will be applied iteratively, with each section of missing values considered a forecast. The simulated series, both historical and forecast may not contain missing values. The error module will not be applied if these are found and appropriate message generated. If missing values may occur in these series, then these will need to be filled prior to applying the error module using a suitably configured interpolation module.

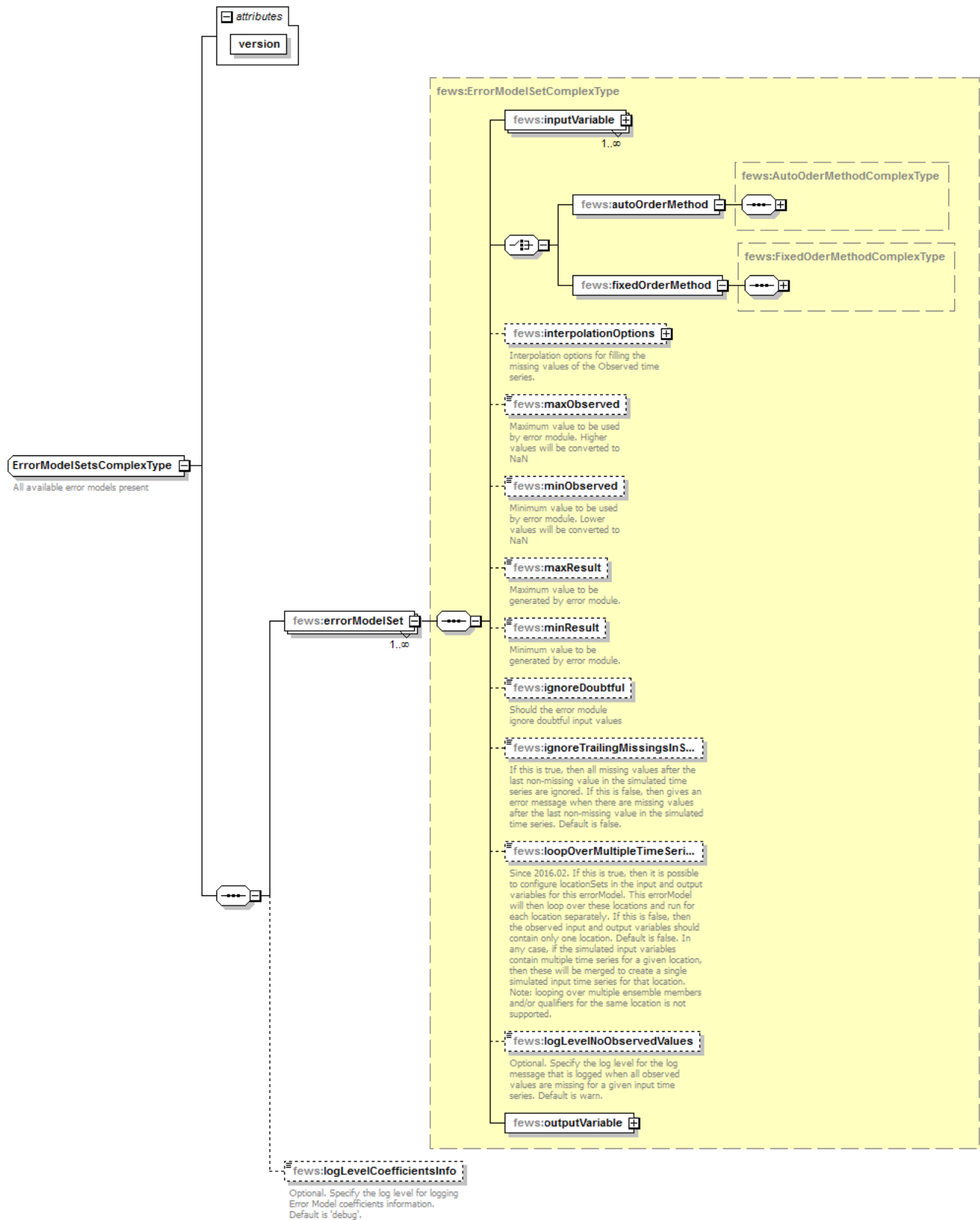


Figure 86 Elements of the error module configuration.

errorModelSet

Root element for definition of an error model set.

inputVariable

Definition of input variable to be used in the error correction model. At least two entries are required in the error model, one for observed time series and one for simulated time series. For each entry an input variable will need to be identified. The variableId is used to refer to the time series. See Transformation Module for definition of inputVariable configuration. It is not possible to use LocationSets in the TimeSeriesSet definition, so it is required to define an ErrorModelSet per location.

TimeSeriesSet

ConvertDatum is not supported in the ARMA module. The input timeseries should be in the same datum.

autoOrderMethod

Root element for defining an error model using the AR structure.

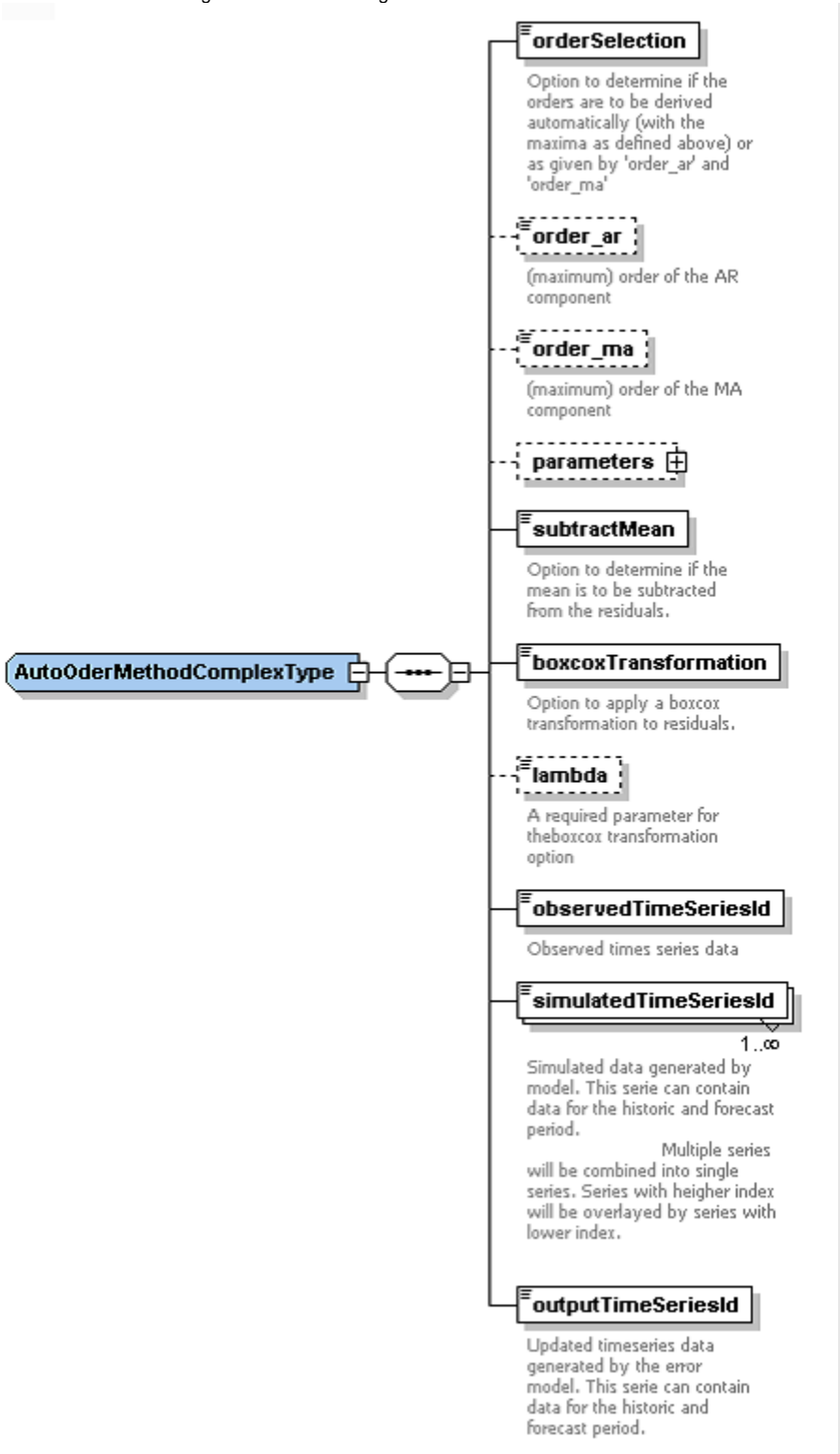


Figure 87 Elements of the autoOrderMethod configuration.

orderSelection

Boolean to indicate if order of AR components should be established automatically or if the given order should be used.

Since 2018.02 an attribute tag ("@ATTRIBUTE_ID@") can be given to use an attribute value of the output location instead. The used attribute must be a <boolean> attribute.

order_ar

Order of the AR model. If the orderSelection is true, then this value is the maximum order (may not exceed 50). In literature mostly an value of the AR order up to 3 is chosen, higher values are possible, but will have a smaller contribution to the overall result of the error correction.

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the order_ar using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

order_ma

Order of the MA model. If the orderSelection is true, then this value is used as the maximum order. In this case, for a given ar order, the corresponding ma order will always be calculated as follows: maOrder = arOrder - (maxArOrder - maxMaOrder). During the order optimization process, each ar order for which this results in a valid ma order will be tried.

For example, if the configured <order_ar> is 5 and the configured <order_ma> is 2, then the order optimization process will attempt an arOrder of 0 (special case), 3, 4, and 5. The ar orders of 1 and 2 will be omitted since they would result in a negative ma order.

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the order_ma using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

parameters

This optional setting can be used to exactly specify the values for all the parameters (multipliers, powers, dividers, etc) used in the error correction model. An example is shown below. Please note that you will need to establish these parameters first. One way to do this is to run a long historical run with auto-parameters on. The log file will show the parameters determined by the model. These parameters can be used to fix the parameters for the forecast.

1.1

at

inputVariable (3)

autoOrderMethod

orderSelection

false

order_ar

3

order_ma

0

parameters

parameter (3)

	= type	= or...	Abc Text
1	ar	1	1.844
2	ar	2	-1.142
3	ar	3	0.27249

subtractMean

false

boxcoxTransformation

false

lambda

0

observedTimeSeriesId

Observed

simulatedTimeSeriesId (2)

outputTimeSeriesId

Updated

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the parameters using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

subtractMean

Boolean to indicate if mean of residuals should be subtracted prior to establishing error model.

Since 2018.02 an attribute tag ("@ATTRIBUTE_ID@") can be given to use an attribute value of the output location instead. The used attribute must be a <boolean> attribute.

boxcoxTransformation

Boolean to indicate if the residuals should be transformed using Box-Cox transformation prior to establishing error model.

Since 2018.02 an attribute tag ("@ATTRIBUTE_ID@") can be given to use an attribute value of the output location instead. The used attribute must be a <boolean> attribute.

lambda

Lambda parameter to use in Box-Cox transformation (note: value of 0 means the transformation is a natural logarithm). Values ranging from 0 to 0.5 are often used.

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the lambda using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

ObservedTimeSeriesId

Input time series set to be defined as the observed data to compare simulated model output to.

SimulatedTimeSeriesId

Input time series set to be defined as the simulated model output for both the historic and the forecast period. Multiple series will be combined into single series. Series with higher index will be overlayed by series with lower index.

OutputTimeSeriesId

Updated timeseries data generated by the error model. This serie can contain data for the historic and the forecast period.

fixedOrderMethod

Root element for defining an error model using the ARMA structure.

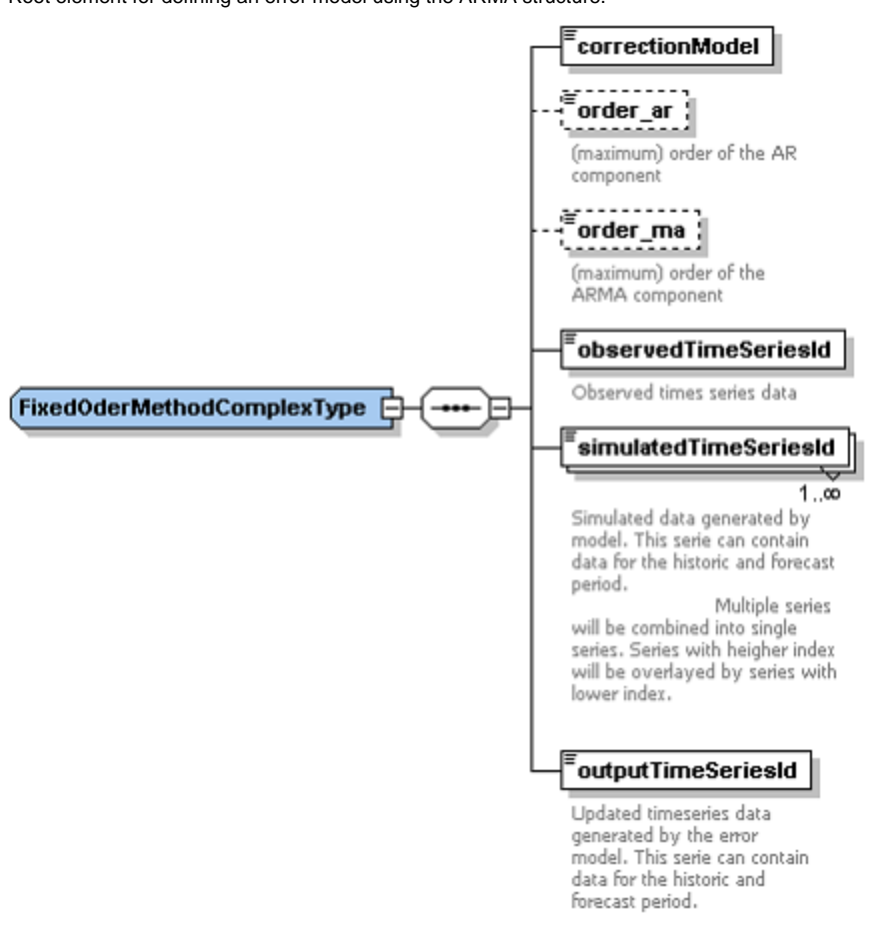


Figure 88 Elements of the fixedOrderMethod configuration.

correctionModel

Structure of the error model to be used. The model selection includes the selection of initial transformations. Enumeration of options included;

- *none*
- *ARMA+ systematic*
- *systematic*
- *ARMA*
- *ARMA+ log transformation*
- *ARMA+ systematic+ log transformation*

order_ar

Order of the AR part of the model. In literature mostly an value of the AR order up to 3 is chosen, higher values are possible, but will have a smaller contribution to the overall result of the error correction.

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the order_ar using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

order_ma

Order of the MA part of the model. The order you specify determines the length of the period effected by the moving average function. The higher the order, the longer the effected period. The moving average model is not operational yet.

Since 2018.02 attribute tags ("@ATTRIBUTE_ID@") can be used to determine the order_ma using attribute values of the output location. The used attributes must be a <number> or <boolean> attribute (in the case of <boolean> the "true" value will be interpreted as 1 and the "false" value as 0).

ObservedTimeSeriesId

Input time series set to be defined as the observed data to compare simulated model output to.

SimulatedTimeSeriesId

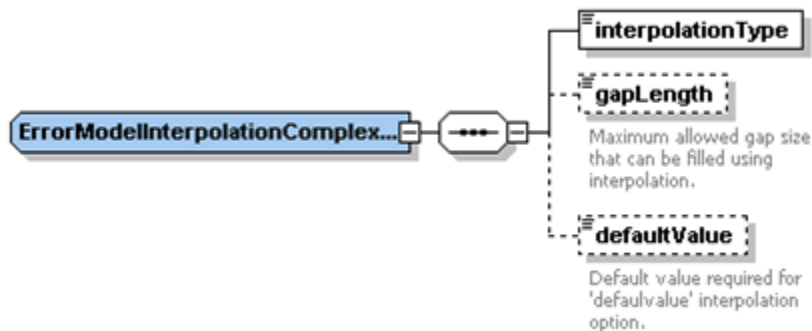
Input time series set to be defined as the simulated model output for both the historic and the forecast period. Multiple series will be combined into single series. Series with higher index will be overlayed by series with lower index.

OutputTimeSeriesId

Updated timeseries data generated by the error model. This serie can contain data for the historic and the forecast period.

interpolationOptions

- Interpolation options for filling the missing values of the observed time series. This parameter is optional.



interpolationType

You can make a selection of a type of interpolation. Enumeration of available options is;

- *linear* ; for linear interpolation between available values
- *block* ; for block interpolation (note: the last available value is then used until a new value available).
- *default* ; for replacing unreliable values with a default.

gapLength

Maximum allowed gap size that can be filled using interpolation.

defaultValue

Default value required for 'defaultvalue' interpolation option.

maxObserved

Maximum value to be used by the error module. Higher values will be converted to NaN and not used as input for error correction. This parameter is optional.

minObserved

Minimum value to be used by the error module. Lower values will be converted to NaN and not used as input for error correction. This parameter is optional.

maxResult

Maximum value to be generated by the error module. This setting can be used to specify an upper limit of the generated output timeseries. This parameter is optional.

minResult

Minimum value to be generated by the error module. This setting can be used to specify a lower limit of the generated output timeseries. This parameter is optional.

ignoreDoubtful

Should the error module ignore doubtful input values. This parameter is optional.

ignoreTrailingMissingsInSimulatedTimeSeries

If this is true, then all missing values after the last non-missing value in the simulated time series are ignored. If this is false, then gives an error message when there are missing values after the last non-missing value in the simulated time series. Default is false.

loopOverMultipleTimeSeries

Since 2016.02. If this is true, then it is possible to configure locationSets in the input and output variables for this errorModel. This errorModel will then loop over these locations and run for each location separately. If this is false, then the observed input and output variables should contain only one location. Default is false. In any case, if the simulated input variables contain multiple time series for a given location, then these will be merged to create a single simulated input time series for that location. Note: looping over multiple ensemble members and/or qualifiers for the same location is not supported.

logLevelNoObservedValues

Optional. Specify the log level for the log message that is logged when all observed values are missing for a given input time series. Can be error, warn or info. Default is warn.

outputVariable

Definition of output variable as a result of the error model. A single timeSeriesSet for one location output variable error model is defined.

logLevelCoefficientsInfo

Option 'logLevelCoefficientsInfo' is set in ErrorModelSetsComplexType directly under element 'errorModelSet'. The value defaults to 'debug', such that the computed coefficients are only logged when the module is run in debug mode. Set the value to 'info' to show the coefficients in the standard logging.