

Operator Manual: Streamflow and Weather Database

From:



To:



1 July 2012

Table of Contents

1. Introduction.....	3
2. Entities and Relationships.....	3
2.1.1.1. Entity Relationship Diagram.....	3
2.1.1.2. Entity Relationship Table	5
3. System Administration	6
3.1.1.1. Default Modules	7
3.1.1.2. Roles	7
3.1.1.3. Users	7
4. Master Data Management.....	7
4.1.1.1. Generic Master Data Management.....	8
4.1.1. Organisations	10
4.1.1.1. Project/Sites	10
4.1.1. Station	10
4.1.1.1. Unit of measure	11
4.1.1.2. Offering.....	12
4.1.1.3. Data Schema.....	13
4.1.1.4. File Split Condition	15
4.1.1.5. Field Configuration	16
4.1.1.5.1. Type of Fields.....	17
4.1.1.6. Data Source	23
4.1.1.7. Universal Data Schema	24
4.1.1.8. Sensor Specific Data Schema	25
4.1.1.1. Datasource Transformations.....	27
4.1.1.1. Datasource Roles.....	28
4.1.1.2. Phenomenon	29
4.1.1.3. SensorProcedure	31
4.1.1.4. Data source Processing (Import Batches)	32
4.1.1.1. Import Data File Form	33
5. The Data Processor	34

1. Introduction

This document provides information and guidance to the end user of the system.

2. Entities and Relationships

2.1.1.1. Entity Relationship Diagram

2.1.1.2. Entity Relationship Table

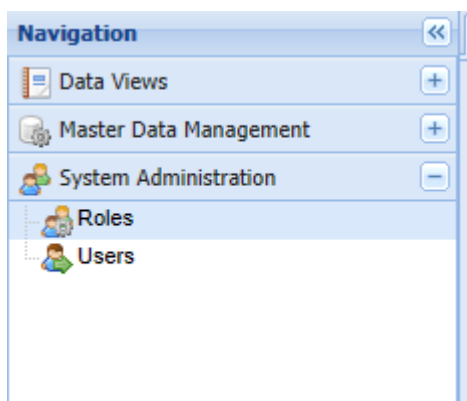
Entity	Relationships
organisation	<ul style="list-style-type: none"> Has many projectsite
projectsite	<ul style="list-style-type: none"> Has one organisation Has many station
Station	<ul style="list-style-type: none"> Has one projectsite Has many sensorprocedure
dataschema	<ul style="list-style-type: none"> Has one datasourcetype Has <u>none/many</u> datasource Has <u>none/many</u> sensorprocedure
datasource	<ul style="list-style-type: none"> Has none/one dataschema Has <u>none/many</u> transformation Has many role
sensorprocedure	<ul style="list-style-type: none"> Has one station Has one phenomenon Has one datasource Has none/one dataschema.
phenomenon	<ul style="list-style-type: none"> Has many unitofmeasure Has many offering Has many transformation Has many sensorprocedure
offering	<ul style="list-style-type: none"> Has many phenomenon
unitofmeasure	<ul style="list-style-type: none"> Has many phenomenon
dataschema	<ul style="list-style-type: none"> Has <u>none/many</u> datasource Has <u>none/many</u> sensorprocedure Has one datasource type

aspnet_users	<ul style="list-style-type: none"> • Has many roles
aspnet_role	<ul style="list-style-type: none"> • Has many user • Has many module • Has many datasource
module	<ul style="list-style-type: none"> • Has many role
datalog	<ul style="list-style-type: none"> • Has none/one sensorprocedure • Has none/one phenomnonoffering • Has none/one phenomnonunitofmeasure • Has one importbatch • Has one status
observation	<ul style="list-style-type: none"> • Has one sensorprocedure • Has one phenomnonoffering • Has one phenomnonunitofmeasure • Has one importbatch

3. System Administration

This section consists of the different modules that are responsible for managing user access to specific system functions and data sets.

The System Administration is located within the Navigation panel of the User Interface.



3.1.1.1. Default Modules

The following modules exist in the system by default and are not managed through the front end.

- Data Sources
- Sensors
- Unit of Measure
- System Administration
- Phenomenon
- Master Data Management
- Roles
- Data Views
- Stations
- Observations
- Projects/Sites
- Offerings
- Data Schemas
- Import Batches
- Organisations
- Data Query Display
- Users

3.1.1.2. Roles

Roles are implemented to group together access to specific system functionality and datasource feeds.

A role can be linked to multiple modules. When a user is linked to the role he/she will have access to all the modules associated with the role.

A role can be linked to multiple datasources. The association includes a time interval in which the role (Users of the role) are granted access to a specific data range of the datasource feed.

3.1.1.3. Users

A user can be assigned to multiple roles. The user is granted access to the functionality and data feeds that have been configured with the Role.

4. Master Data Management

This section consists of the capturing and configuration module for the system entities.

Each entity set has a link in the Navigation panel that activates a separate tab in content section containing the relevant user layouts for the specific entity.

4.1.1.1. Generic Master Data Management

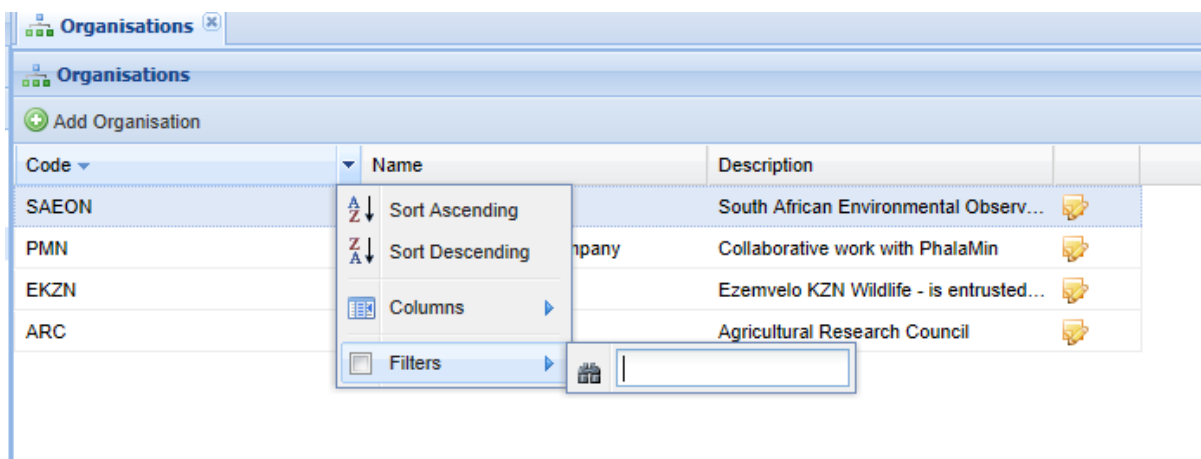
Each entity set consists of at least 2 layouts

- List
 - Find
 - Filter
 - Sort
 - Page
 - Delete
 - Export

- Detail
 - Add
 - Edit

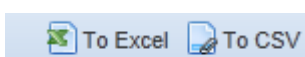
The List layout presents the user with an interactive list that displays a column for each of the most important attributes for the specific entity.

The list can be sorted on any of the columns. A dynamic filter for each attribute type is available and accessible by hovering over the relevant column header.

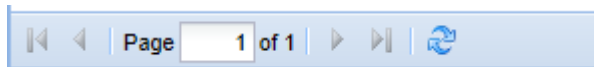


The toolbar at the top of the list will at least contain an action button which enables the user to add a new instance of a specific entity type.

The toolbar also contains two export buttons, To Excel, To CSV. The export function will output the data set to the relevant format on each of the available pages with the applied filters.



The last column on each row consist of at least an action button which enables the user to activate the Edit entity layout form which will allow amendments to the specific entity instance.



The bottom of the list layout form has a paging mechanism which displays the current page and the number of pages available in the entity (data) set and dynamically changes as filters are applied.

The detail form of each entity is presented in the form of a modal popup input form. The detail form for each entity will consist of at least the following fields.

Code

A unique code for the specific instance of the entity set.

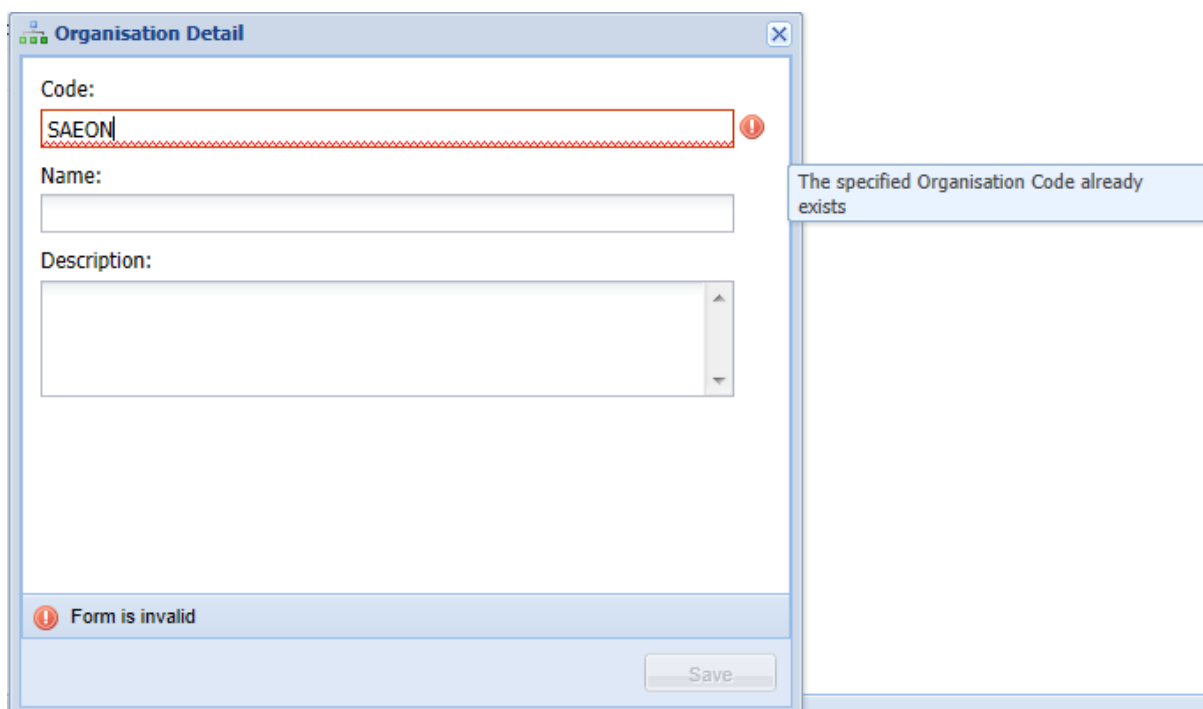
Name

A unique name for the specific instance of the entity set.

Description

Descriptive information on the specific instance of the entity set.

The Save button at the bottom of the form activates only once all required fields have been provided and validation on specialized field succeeds. The validity of the input is also indicated by the label at the bottom of the Detail form.



Each input field includes a validity indicator to the right of the field which are activated when the input was invalid. Descriptive information of the input validation error is presented when the mouse is hovered over the validity indicator.

When the save button is clicked the relevant entity with configuration are stored and the user is taken back to the relevant Entity Set list layout form.

4.1.1.Organisations

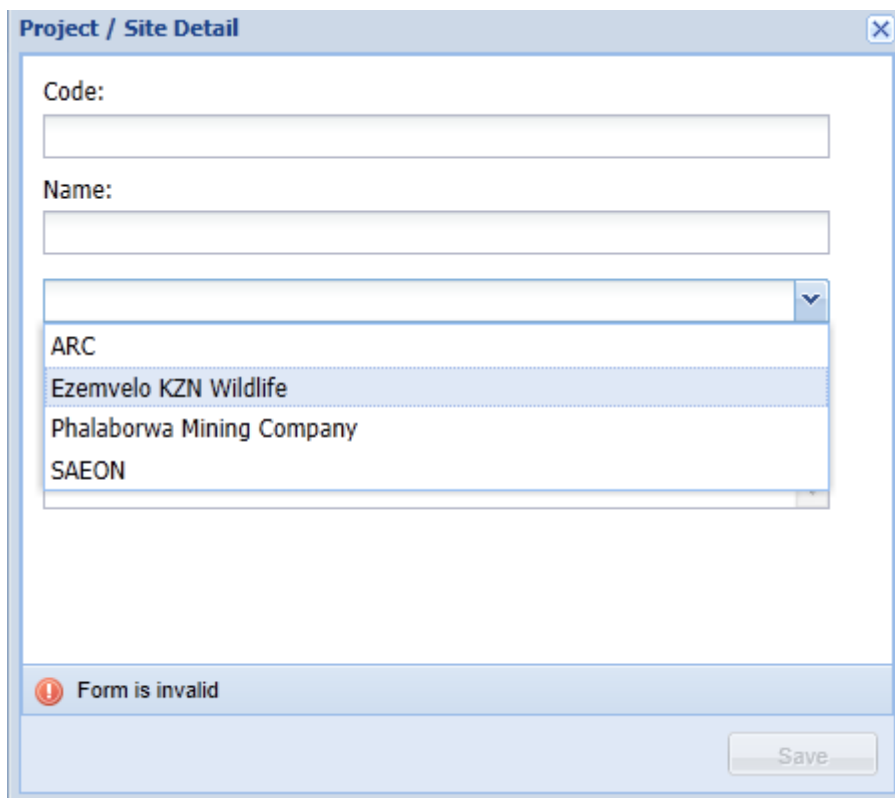
No Dependency

See section 4.1.1 General Data Management

4.1.1.1. Project/Sites

Depends on Organisation

See section 4.1.1 General Data Management



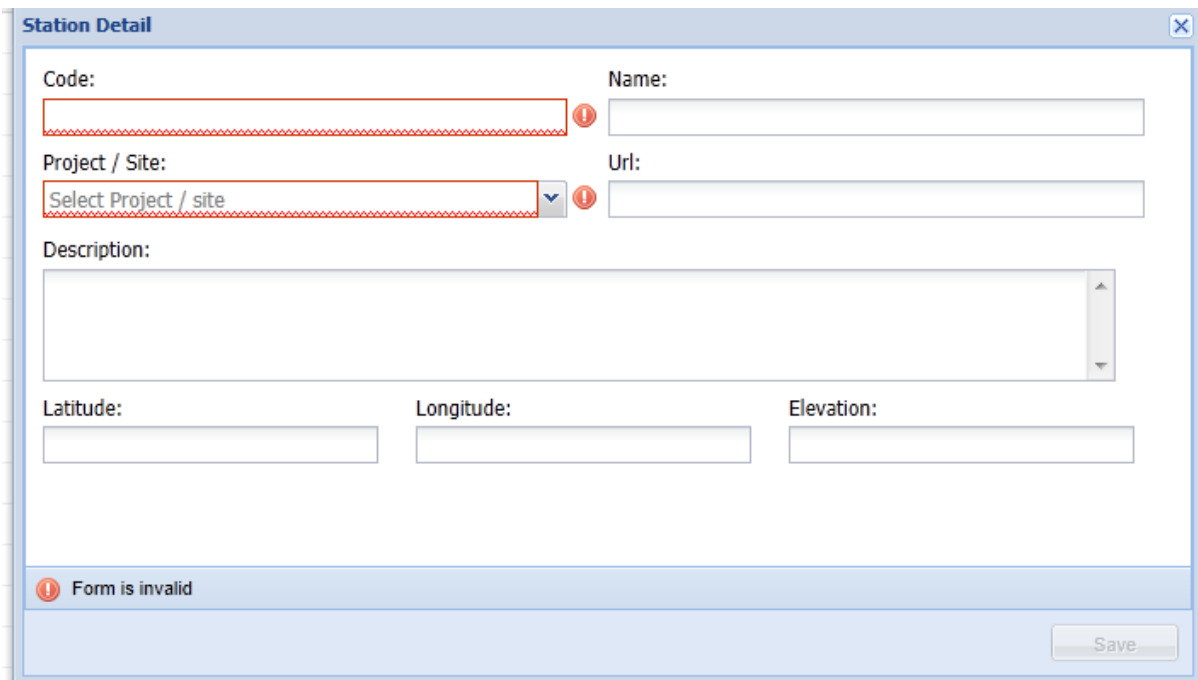
The screenshot shows a web form titled "Project / Site Detail". It contains several input fields: "Code:" (text box), "Name:" (text box), and a dropdown menu. The dropdown menu is open, showing a list of organisations: "ARC", "Ezemvelo KZN Wildlife", "Phalaborwa Mining Company", and "SAEON". At the bottom of the form, there is a red error icon and the text "Form is invalid". A "Save" button is located at the bottom right of the form.

Lookup Data (Organisations) are presented in the form of a Drop down list and only one item can be selected.


4.1.1.Station

Depends on Projects/Sites



See section 4.1.1 General Data Management



Station Detail

Code: 


Name:

Project / Site:  

Url:

Description:

Latitude: Longitude: Elevation:

 Form is invalid

Save

Lookup Data (Projects/Sites) are presented in the form of a Drop down list and only one item can be selected.

4.1.1.1. Unit of measure

No Dependency

See section 4.1.1 General Data Management

The image shows a software window titled "Unit Detail" with a close button in the top right corner. The window contains three text input fields: "Code:", "Unit:", and "Symbol:". The "Unit:" field is highlighted with a red border and a red exclamation mark icon, indicating a validation error. At the bottom of the window, there is a status bar with a red exclamation mark icon and the text "Form is invalid". A "Save" button is located in the bottom right corner of the window.

4.1.1.2. Offering

No Dependency

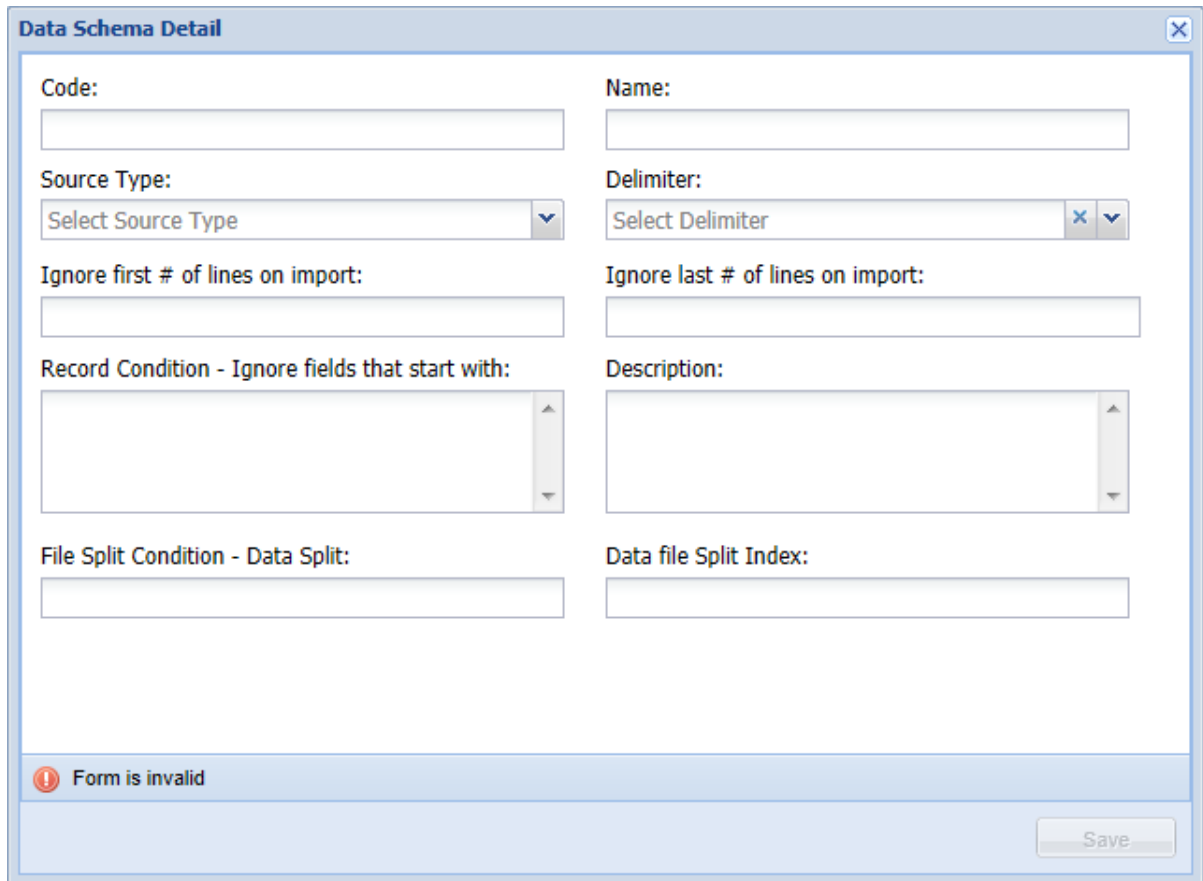
See section 4.1.1 General Data Management

The screenshot shows a web form titled "Offering Detail" with a close button in the top right corner. The form contains three input fields: "Code:" with the value "act", "Name:", and "Description:". Each field has a red dashed border and a red exclamation mark icon to its right, indicating a validation error. At the bottom of the form, there is a blue bar with a red exclamation mark icon and the text "Form is invalid". A "Save" button is located at the bottom right of the form.

4.1.1.3. Data Schema

No Dependency

See section 4.1.1 General Data Management



The dataschema is a mechanism used to configure the type and layout of the data that will be used to process data from a specific feed. The following tasks form part of the data schema definition process.

Task	Description
Specify the physical Data Format	The format in which the data are provided is configured. Options are Delimited Fixed Length Flat file SMS or USSD Messages.
Specify the delimiter (if the type is delimited)	Options are Comma Delimited (,) Pipe Delimited (\) Tab Delimited (\t) Semi Comma Delimited(,;)

<p>Specify where in the file(feed) to look for the data and the condition for a valid data record</p>	<p>The following settings are available to filter the feed for the actual data</p> <p># of lines to ignore at the start of the file</p> <p># of lines to ignore at the end of the file</p> <p>Record condition – Ignore fields that start with</p> <p>File Split Condition – Data Split</p> <p>Data file Split Index</p>
<p>Field Configuration</p>	<p>Each column in the data feed is individually configured according to the following variables</p> <p>Field Type</p> <p>Date of observation</p> <p>Time of observation</p> <p>Observation value (Offering)</p> <p>The data in the field should be ignored.</p> <p>The value in the field is a comment for the actual values,</p> <p>The data in the field is always a fixed value.</p>

4.1.1.4. File Split Condition

One data feed can contain multiple sets of data in a single file. A data schema should be configured for each of the different sets in the file. The File Split Condition and Data Split Index parameters should be set accordingly to apply a data schema to a specific data set in the file.

In the scenario below 2 different sets of data are displayed. The data processor should be configured to start at row number 2 (# of lines to ignore at the start of the file) with the File Split Condition

Parameter set to “<STATION>”. This means the file will be split into two separate data sets, it will split whenever “<STATION>” is encountered.

The data file will be split into an individual set for Waterlevel and Temperature respectively for this scenario.

To define the schema for retrieving Waterlevel observations in this scenario the Data File Split Index should be set 1. If the Temperature observations is required the Data File Split Index should be set to

2. In Addition the “Record condition – Ignore fields that start with” parameter should be set to “<INTEXTVAL”. This will indicate to the processor that the Record in this instance that starts with “<INTEXTVAL” should be ignored as it is not a valid data record.

```
<STATION>0TIERKLOOF</STATION><SENSOR>WLEV</SENSOR><DATEFORMAT>YYYYMMDD</DATEFORMAT>
<INTEXTVAL DATE="20111027" MINVAL="0.175" MINTIME="170000" MAXVAL="0.184" MAXTIME="070000" />
20111027;220000;0.177
20111027;230000;0.178
20111028;000000;0.178
20111028;010000;0.178
20111028;020000;0.178
20111028;030000;0.178
20111028;040000;0.179
20111028;050000;0.179
20111028;060000;0.178
20111028;070000;0.179
20111028;080000;0.179
20111028;090000;0.179
<STATION>0TIERKLOOF</STATION><SENSOR>TEMP</SENSOR><DATEFORMAT>YYYYMMDD</DATEFORMAT>
<INTEXTVAL DATE="20111027" MINVAL="12.1" MINTIME="073000" MAXVAL="14.0" MAXTIME="163000" />
20111027;220000;12.8
20111027;230000;12.6
20111028;000000;12.3
20111028;010000;12.1
20111028;020000;11.9
20111028;030000;11.9
20111028;040000;11.8
20111028;050000;11.8
20111028;060000;11.8
20111028;070000;11.7
20111028;080000;11.8
20111028;090000;12.1
```

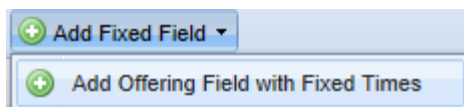
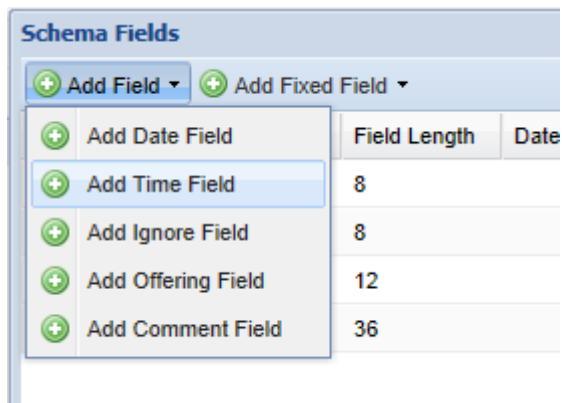
4.1.1.5. Field Configuration

This activity involves the configuration of each data field (column) in the data feed. The data in each field should be assessed and classified according to the following variables.

- Is the Date of observation
- Is the Time of observation
- Is the Observation value (Offering)
- The data in the field should be ignored.
- The data in the field is always a fixed value.
- If the type of feed is Fixed Length Flat file what is the length of the data field.

The data processor needs to know what the value of each field will hold, what the type of that value will be, how to identify the valid value of the field value based on a predefined format and whether the data processor should try to process a field or discard the information.























4.1.1.5.1. Type of Fields



Field Type	Description	Additional Parameters
Date Field	The actual date of the observation	<ul style="list-style-type: none"> ○ Field Name ○ The format to used to read the value into a system date. I E "dd/MM/yyyy" ○ A field length for fix length files.
Time Value	The actual time of the observation	<ul style="list-style-type: none"> ○ Field Name ○ The format to use to read the value into a system time. I E "hh:mm:ss:" ○ A field length for fix length files.
Ignore field	The field should be ignored by the data processor	<ul style="list-style-type: none"> ○ Field Name ○ A field length for fix length files.
Offering Field	The observation data value	<ul style="list-style-type: none"> ○ Field Name ○ Phenomenon ○ Phenomenon Offering ○ Unit of measure ○ Empty value ○ If this value is encountered while processing data value will be interpreted to System NULL which means that the observation was conducted but the actual value is incomplete ○ A field length for fix length files.

Comment Field	A comment for each observation value in the same data row	<ul style="list-style-type: none"> ○ Field Name ○ A field length for fix length files.
Offering Field with fix times	The observation data value accompanied with a fixed time for the observation reading	<ul style="list-style-type: none"> ○ Field Name ○ Phenomenon ○ The fixed Time of the Observation. ○ Phenomenon Offering ○ Unit of measure ○ Empty value ○ If this value is encountered while processing data value will be interpreted to System NULL which means that the observation was conducted but the actual value is incomplete

Below a typical configuration for an Automated whether station data feed which has a Date field, Time field ,multiple Offering(Phenomenon) data fields and multiple ignore fields.

Schema Fields						
+ Add Field ▾ + Add Fixed Field ▾						
Name	Date Field	Time Field	Ignore Field	Offering Field	Fixed Field	
Date	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	 
Time	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	 
Temp_Out	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Temp_Hi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Temp_Low	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Humidity	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Dew_Point	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Wind_Speed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Wind_Dir	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	 
Wind_Run	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 
Hi_Speed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	 

In the sample above(Automated whether station) the fields are clearly defined with column headers.

Temp	Hi	Low	Out	Dew	Wind	Wind	Wind	Hi	Hi	W:	
Date	Time	Out	Temp	Temp	Hum	Pt.	Speed	Dir	Run	Speed	
11/06/18	3:30	a	7.8	7.9	7.5	91	6.5	12.9	E	6.44	25.7
11/06/18	4:00	a	7.8	7.9	7.8	90	6.3	14.5	E	7.24	27.4
11/06/18	4:30	a	7.8	7.9	7.8	91	6.5	14.5	E	7.24	25.7
11/06/18	5:00	a	7.6	7.8	7.6	91	6.2	14.5	E	7.24	25.7
11/06/18	5:30	a	7.7	7.7	7.6	92	6.4	14.5	E	7.24	30.6
11/06/18	6:00	a	7.6	7.7	7.5	92	6.4	12.9	E	6.44	24.1
11/06/18	6:30	a	7.5	7.6	7.5	93	6.4	12.9	E	6.44	25.7
11/06/18	7:00	a	7.5	7.6	7.4	93	6.4	11.3	E	5.63	20.9
11/06/18	7:30	a	7.8	7.8	7.5	92	6.6	11.3	E	5.63	22.5
11/06/18	8:00	a	8.5	8.5	7.7	91	7.1	11.3	E	5.63	20.9
11/06/18	8:30	a	8.4	8.8	8.4	89	6.7	11.3	E	5.63	22.5
11/06/18	9:00	a	8.6	8.7	8.4	87	6.6	12.9	E	6.44	25.7
11/06/18	9:30	a	9.1	9.3	8.6	85	6.7	12.9	E	6.44	24.1
11/06/18	10:00	a	9.0	9.0	8.5	80	6.5	12.9	E	6.44	20.9

Below the configuration for each field

Date

Date field Detail ✕

Name:

Date Format:

Form is valid

Time

Time field Detail [X]

Name:

Time Format:

Form is valid

Temp Out

Offering Field Detail

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:

Temp High

Offering Field Detail

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:

Temp Low

Offering Field Detail

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:

Humidity


Offering Field Detail

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:


Dew Point

Offering Field Detail ✕

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:

Wind Speed

Offering Field Detail ✕

Name:

Phenomenon:

Offering:

Unit of Measure:

Empty Value:

Wind Direction

Comment Field Detail

Name:
Wind_Dir

Form is valid

Save

Wind Run

Offering Field Detail

Name:
Wind_Run

Phenomenon:
Wind

Offering:
Wind run

Unit of Measure:
Meters per second

Empty Value:

4.1.1.6. Data Source

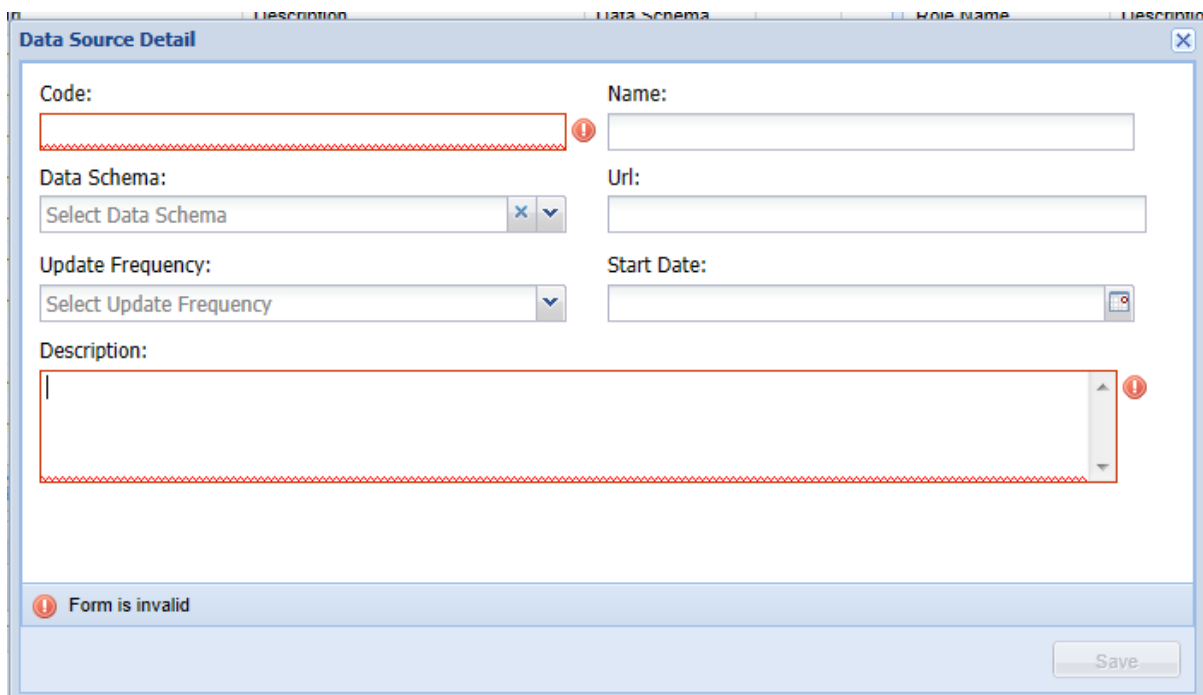
Partially Depends On Data Schema

See section 4.1.1 General Data Management

Below the configuration for each field

The data source setup involves the following configuration

- The virtual location from where the same data feed will be available from online.
- The update frequency in the event where the feed will be pulled automatically.
- The start date of the automated pull process.
- Specify a universal Data Schema if multiple sensors consume the same set of data (data feed, data file) while the data processor reads a single feed in a tabular fashion.



Based on the datasource file layout the decision should be made whether a “Universal Data Schema” or a “Sensor Specific Data Schema” is applicable.

When the “Universal Data Schema” method is applicable a Data Schema “SHOULD” be selected on the Data Source Master Detail.

4.1.1.7. Universal Data Schema

A universal data schema should be specified for the datasource when multiple sensors consume the same data feed (datasource) while the data processor reads a single feed in a tabular fashion.

That is each column in the data feed can be associated with a single phenomenon and offering except for the date and time fields. The Automated whether station is a typical example of a universal data schema scenario.

Date	Time	Out	Temp	Temp	Hum	Pt.	Speed	Dir	Run	Speed	
11/06/18	3:30	a	7.8	7.9	7.5	91	6.5	12.9	E	6.44	25.7
11/06/18	4:00	a	7.8	7.9	7.8	90	6.3	14.5	E	7.24	27.4
11/06/18	4:30	a	7.8	7.9	7.8	91	6.5	14.5	E	7.24	25.7
11/06/18	5:00	a	7.6	7.8	7.6	91	6.2	14.5	E	7.24	25.7
11/06/18	5:30	a	7.7	7.7	7.6	92	6.4	14.5	E	7.24	30.6
11/06/18	6:00	a	7.6	7.7	7.5	92	6.4	12.9	E	6.44	24.1
11/06/18	6:30	a	7.5	7.6	7.5	93	6.4	12.9	E	6.44	25.7
11/06/18	7:00	a	7.5	7.6	7.4	93	6.4	11.3	E	5.63	20.9
11/06/18	7:30	a	7.8	7.8	7.5	92	6.6	11.3	E	5.63	22.5
11/06/18	8:00	a	8.5	8.5	7.7	91	7.1	11.3	E	5.63	20.9
11/06/18	8:30	a	8.4	8.8	8.4	89	6.7	11.3	E	5.63	22.5
11/06/18	9:00	a	8.6	8.7	8.4	87	6.6	12.9	E	6.44	25.7
11/06/18	9:30	a	9.1	9.3	8.6	85	6.7	12.9	E	6.44	24.1

In this scenario multiple phenomenon (sensors) readings are available per date/time frequency/

One data schema can be setup and linked to the datasource. While this data feed is processed the processor can dynamically resolve the sensor that should be associated with the reading as each of the field will be associated with a specific phenomenon and offering.

4.1.1.8. Sensor Specific Data Schema

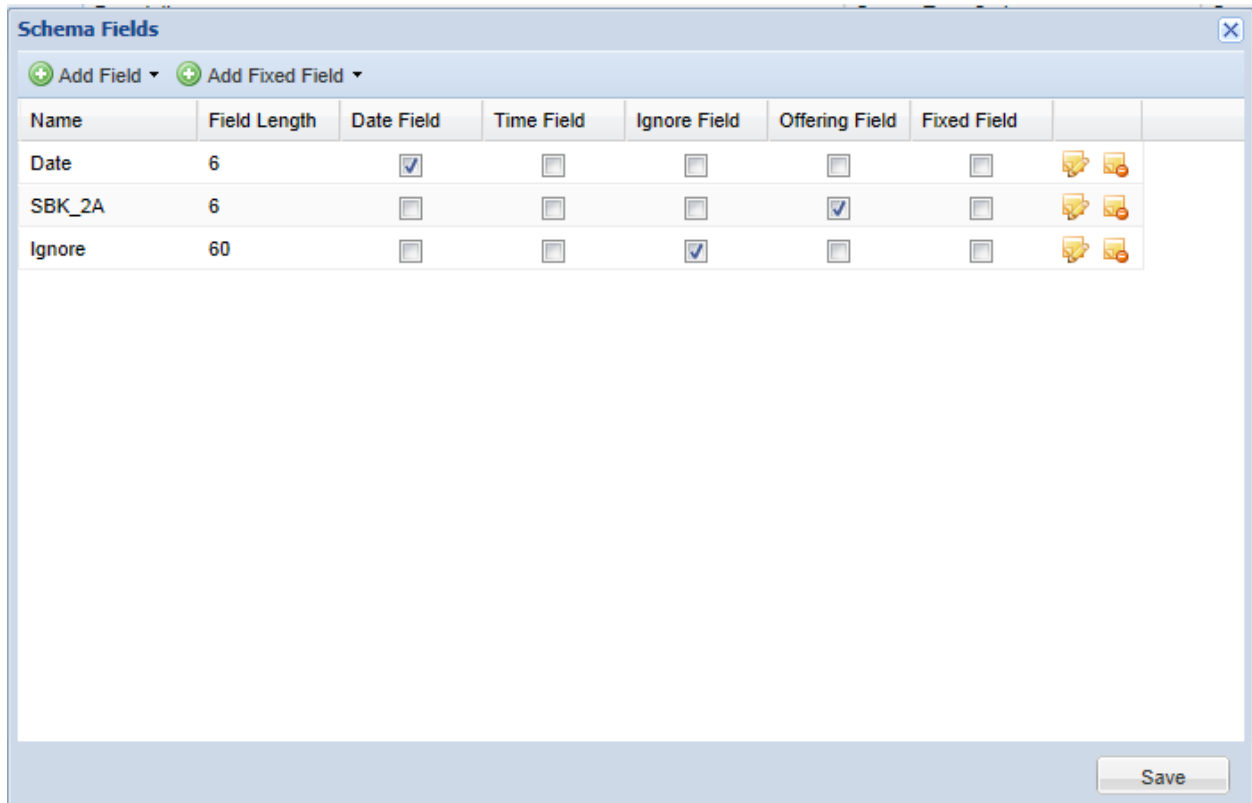
When multiple readings of the same phenomenon are present in the datasource file, a data schema should be provided for each sensor that consumes the data source.

The manually captured rainfall readings file for Jonkerhoek is an example of when the Sensor specific data schema is applicable.

DATE	2A	2B	7A	7B	8A	8B	9A	9B	19B	20A	20M
500120	0.3	0.3	0.0	0.0	0.5	0.5	0.3	0.3	.	.	.
500123	0.5	0.5	0.0	0.0	1.0	1.0	0.3	0.3	.	.	.
500124	1.5	1.5	1.3	1.3	1.0	1.0	1.5	1.8	.	.	.
500131	3.8	3.8	3.8	3.8	4.1	4.1	4.1	4.1	.	.	.
500209	7.1	7.4	0.3	0.3	3.3	3.6	1.0	1.5	.	.	.
500214	1.3	1.3	0.8	0.8	1.5	1.5	0.8	0.8	.	.	.
500223	1.8	1.8	1.5	1.5	1.8	1.8	1.8	1.8	.	.	.
500301	2.3	2.5	1.8	1.8	1.8	2.1	2.0	2.0	.	.	.
500321	7.9	8.4	6.1	6.6	6.9	7.7	6.6	7.4	.	.	.
500326	0.3	0.3	0.3	0.3	0.3	0.3	0.5	0.5	.	.	.
500327	8.1	8.6	8.6	9.2	8.6	9.5	8.6	9.0	.	.	.
500329	2.3	2.5	1.8	1.8	1.8	2.1	1.5	1.5	.	.	.
500401	11.9	12.5	11.4	12.0	9.7	10.5	11.2	11.8	.	.	.
500403	4.1	4.1	1.0	1.0	2.8	3.1	1.5	1.5	.	.	.
500404	2.3	2.3	1.0	1.0	0.5	0.5	0.5	0.5	.	.	.
500409	23.1	24.1	21.6	22.8	24.4	25.9	22.6	23.5	.	.	.
500416	2.3	2.5	4.3	4.9	4.3	5.1	2.5	3.1	.	.	.
500417	36.1	36.1	25.7	26.3	36.3	36.7	28.7	28.9	.	.	.
500418	2.0	2.0	8.9	8.7	5.8	6.2	7.4	7.2	.	.	.
500423	9.7	9.4	6.6	6.6	6.1	6.4	7.6	7.7	.	.	.
500424	17.5	18.0	6.1	6.6	9.4	9.5	9.1	9.5	.	.	.

Each column of the file, except Date, defines a rainfall reading (and offering) specific to a SensorProcedure which should be defined for the data source and assigned a unique Data Schema

The first data schema to be configured will be for SensorProcedure “2A”. This configuration will only be concerned with the date of the reading and the actual reading values in “2A” column of the file, hence the configuration as follows.



Since this is a fixed length flat file, the length of each column is defined as part of the configuration. The first 6 character of each row defines the date value of each reading.

The second field in the data schema specifies the offset and length of the column where the rainfall readings for SensorProcedure “2A” should be found when the file is processed. (Offset being the end of the Date column in this instance).

To complete the configuration the third field is configured to ignore the rest of the contents of each row. In this case, 60 characters.

In order to process the readings SensorProcedure “2A” should be configure to use the created schema when the datasource is processed. This is done under the Master Data -> SensorProcedure section. See section 0

Sensor Procedure Detail ✕

Code: Name:

URL:

Description:

Station:

Phenomenon:

DataSource:

Data Schema: ✕ ▼

Form is valid

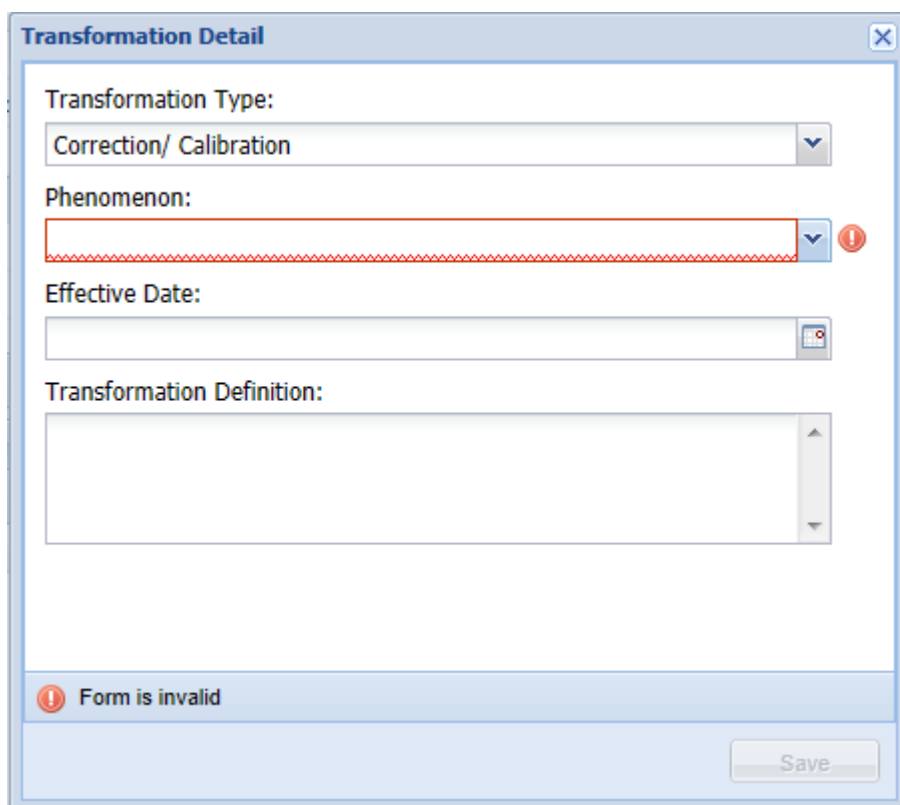
4.1.1.1. Datasource Transformations

See section 4.1.1 General Data Management

The following Transformation Types exist.

Transformation	Description
Rating Table	The actual observation value are compared to a list(table) of chronological numerical values. The closest match is used to override the Data value of the observation

Quality Control on Values	These values indicate the valid range of the transformed data value. If the value falls outside the range, the import batch is moved to the datalog.
Correction/ Calibration	A mathematical expression is applied to the observation and the data value is overwritten with the result

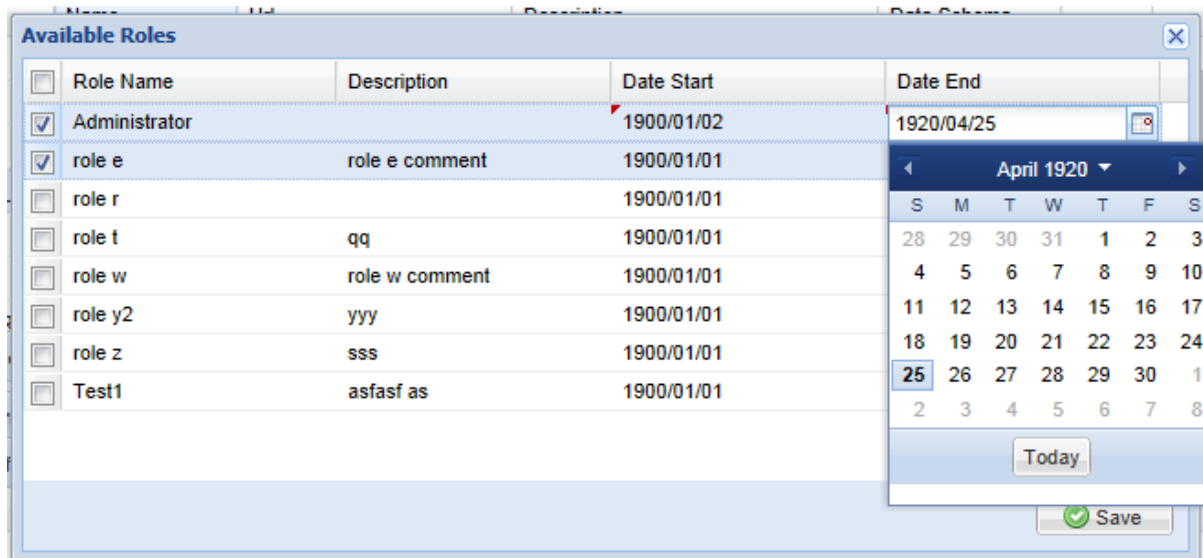


The Rating Table values, Calibration expression and Control values are captured within the Transformation Definition Input field.

The Transformation applies to a specific phenomenon for a specific data feed (datasource).

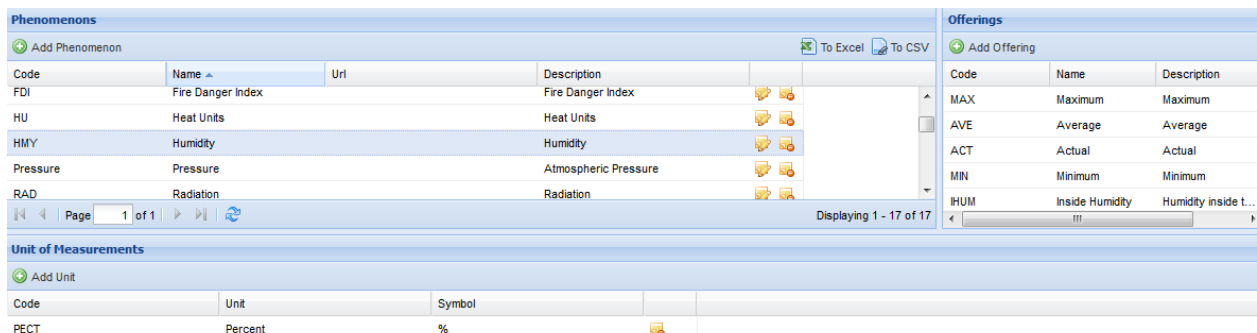
4.1.1.1. Datasource Roles

Access to the data collected by a specific datasource can also be limited to specific user roles. When a role is assigned to the datasource a date range should also be specified which will give users of the specific role access to data that has collected within the selected date period.



4.1.1.2. Phenomenon

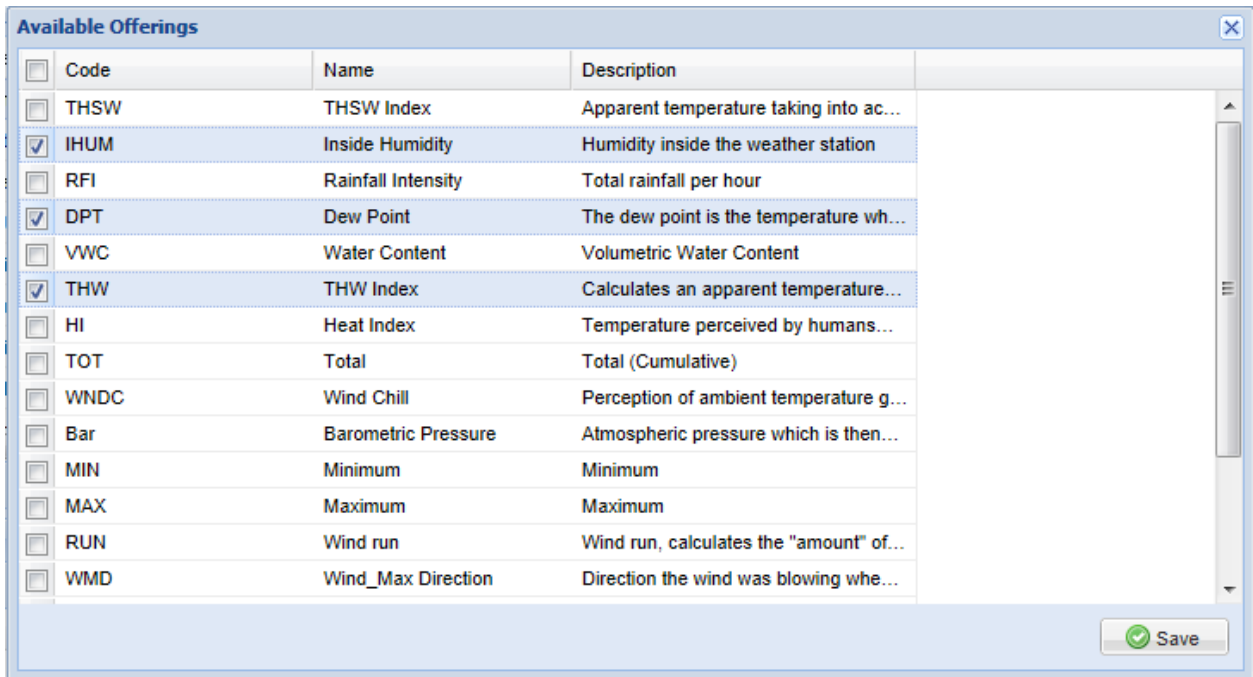
The phenomenon page is used to setup the phenomenons and how they are linked to offerings and units of measure.



The right hand pane displays all the offering that has been assigned to the phenomenon selected in the topleft pane.

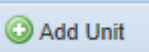
Offerings can be removed from the selected phenomenon or new ones assigned from this section.

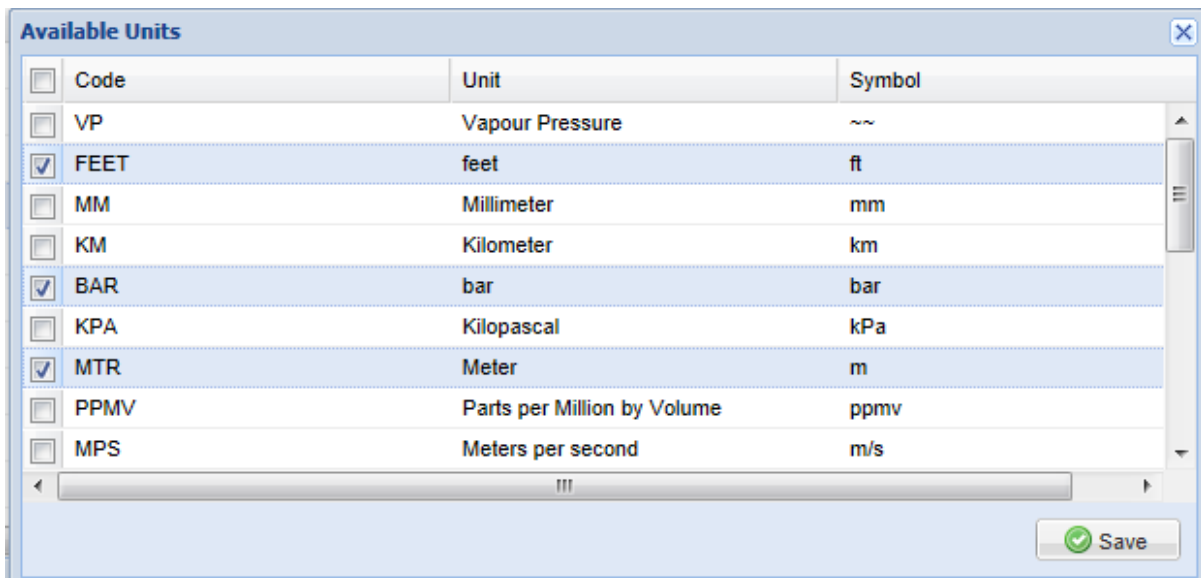
Click the **Add Offering** button to bring up the available offerings window. Select (Checkbox on the left hand side) the relevant offerings and hit the save button. The selected offering will now be associated with the selected phenomenon.



The bottom pane displays all the UOM that has been assigned to the phenomenon selected in the topleft pane.

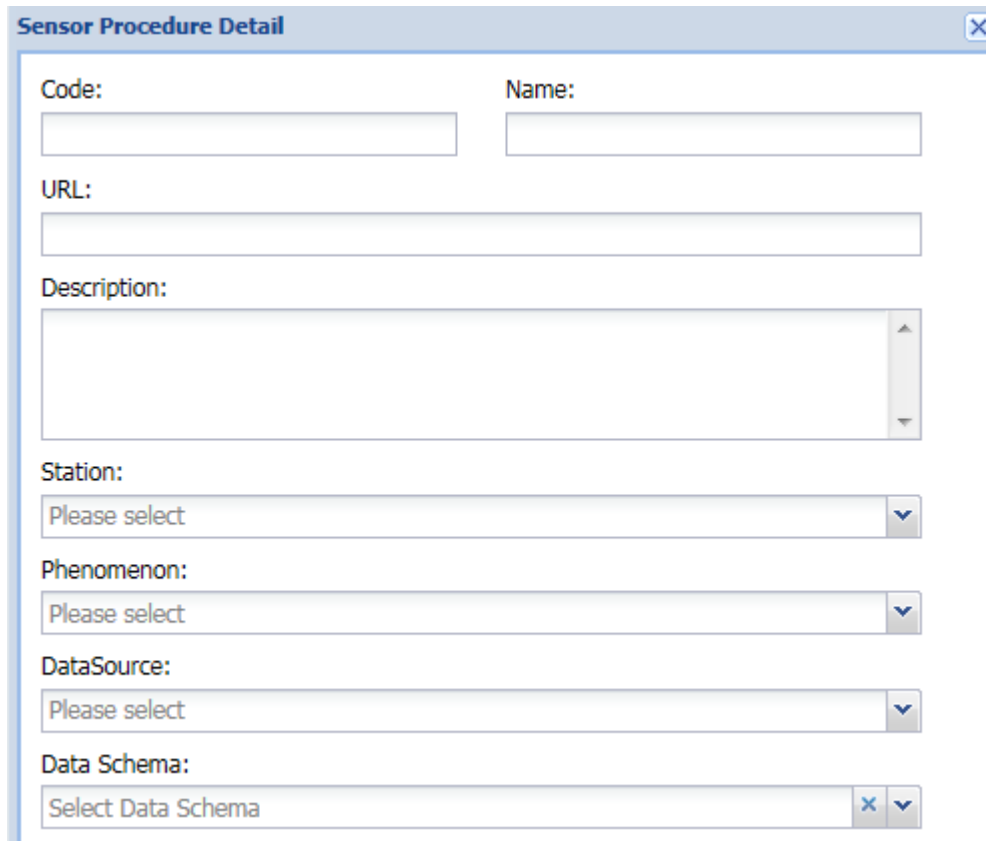
UOM can be removed from the selected phenomenon or new ones assigned from this section.

Click the  button to bring up the available UOM window. Select(Checkbox on the left hand side) the relevant offerings and hit the save button. The selected UOM will now be associated with the selected phenomenon.



4.1.1.3. SensorProcedure

A sensor consists of the following elements:



Sensor Procedure Detail

Code: Name:

URL:

Description:

Station:

Phenomenon:

DataSource:

Data Schema:

Sensor details:

- Belongs to a station
- Has one phenomenon
- Has one data source
- Has an optional data schema

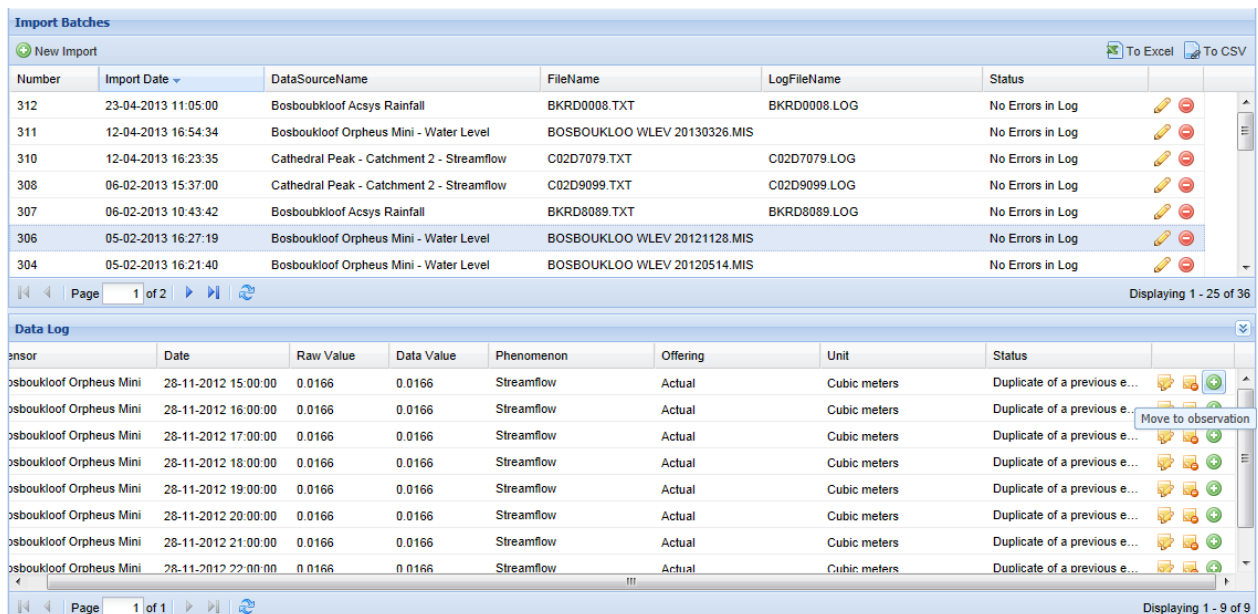
Based on the datasource file layout the decision should be made whether a “Universal Data Schema” or a “Sensor Specific Data Schema” is applicable.

When the “Sensor Specific Data Schema” is applicable a Data Schema "SHOULD" be selected on the SensorProcedure Master Detail.

The data schema has to be selected when the selected data source does not have a data schema selected. You will not be able to select a data schema on the SensorProcedure when the selected data source already has a data schema selected and vice versa.

4.1.1.4. Data source Processing (Import Batches)

Here is where all Master data setup and configuration are used to process the contents of a Raw file into observations (information) and persist them for analysis and further processing (the observations database).



Number	Import Date	DataSourceName	FileName	LogFileName	Status
312	23-04-2013 11:05:00	Bosboukloof Acsys Rainfall	BKRD0008.TXT	BKRD0008.LOG	No Errors in Log
311	12-04-2013 16:54:34	Bosboukloof Orpheus Mini - Water Level	BOSBOUKLOO WLEV 20130326.MIS		No Errors in Log
310	12-04-2013 16:23:35	Cathedral Peak - Catchment 2 - Streamflow	C02D7079.TXT	C02D7079.LOG	No Errors in Log
308	06-02-2013 15:37:00	Cathedral Peak - Catchment 2 - Streamflow	C02D9099.TXT	C02D9099.LOG	No Errors in Log
307	06-02-2013 10:43:42	Bosboukloof Acsys Rainfall	BKRD8089.TXT	BKRD8089.LOG	No Errors in Log
306	05-02-2013 16:27:19	Bosboukloof Orpheus Mini - Water Level	BOSBOUKLOO WLEV 20121128.MIS		No Errors in Log
304	05-02-2013 16:21:40	Bosboukloof Orpheus Mini - Water Level	BOSBOUKLOO WLEV 20120514.MIS		No Errors in Log


Sensor	Date	Raw Value	Data Value	Phenomenon	Offering	Unit	Status
bsboukloof Orpheus Mini	28-11-2012 15:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 16:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 17:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 18:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 19:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 20:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 21:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...
bsboukloof Orpheus Mini	28-11-2012 22:00:00	0.0166	0.0166	Streamflow	Actual	Cubic meters	Duplicate of a previous e...

The top pane displays a list of import events with the following detail

- Unique system number assigned for each import.
- The import event date and time
- The datasource used for the import.
- The Name of the file which was imported and processed.
- If applicable (Acsys), the Name of the Log file used as part of the import.
- The status of the Import.

The bottom panel displays the entries that exist in the data log for the selected Import event in the top panel. If entries exist in the Data log it means that the data processed as the import happened was not consistent with a set of processing (Business Rules).

The consistencies in each entry that exist in the data log for an import event should be resolved before an import batch can be processed to the observation store.

Once all consistencies are fixed the import batch can be processed to the observation store by using the  action button.

Alternatively if inconsistencies/discrepancies exist because of invalid schema configuration or when an incorrect file is imported, the import entry can be deleted and processed later when schema configuration is updated.

The following detail is displayed in the Data log detail view.

- The sensor for which the inconsistency is relevant
- The Date of the observation.

- Raw value from the data source file.
- The process data value.
- The phenomenon for which the inconsistency is relevant.
- The offering of the phenomenon for which the inconsistency is relevant.
- The UOM of the phenomenon for which the inconsistency is relevant.
- Status. The system generated message, detail of the consistency.

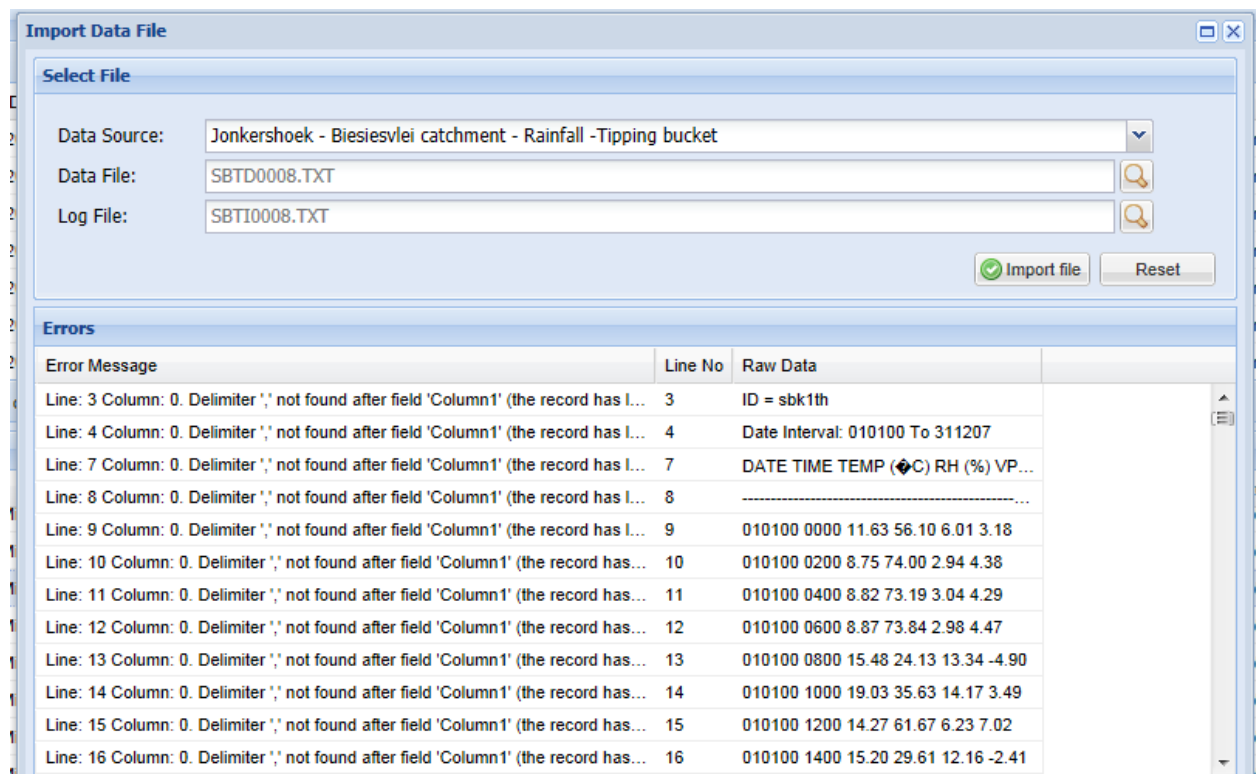
The status column of the Datalog display (bottom pane) together with the action buttons should be used to resolve the inconsistency with the relevant observation (Selected entry in the bottom panel) in order to process the record into the observations data store.

The edit button should be used to view the full detail of the selected record and to correct inconsistencies/discrepancies in the date or the value of the observation. An example of such an inconsistency is where an Empty/"System NULL" value was found while processing the data source file or where the date/time format was different to the format configured for the data schema field when the file is processed.

The delete button can be used to delete an entry from the data log. Note that this entry will be removed from the data log store and will be lost.

The move to observation button is enabled when a duplicate entry already exist in the observation store. This option moves the entry from the Data log and overwrites the data in the observations store.

4.1.1.1. Import Data File Form



Error Message	Line No	Raw Data
Line: 3 Column: 0. Delimiter ',' not found after field 'Column1' (the record has l...	3	ID = sbk11h
Line: 4 Column: 0. Delimiter ',' not found after field 'Column1' (the record has l...	4	Date Interval: 010100 To 311207
Line: 7 Column: 0. Delimiter ',' not found after field 'Column1' (the record has l...	7	DATE TIME TEMP (◆) RH (%) VP...
Line: 8 Column: 0. Delimiter ',' not found after field 'Column1' (the record has l...	8
Line: 9 Column: 0. Delimiter ',' not found after field 'Column1' (the record has l...	9	010100 0000 11.63 56.10 6.01 3.18
Line: 10 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	10	010100 0200 8.75 74.00 2.94 4.38
Line: 11 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	11	010100 0400 8.82 73.19 3.04 4.29
Line: 12 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	12	010100 0600 8.87 73.84 2.98 4.47
Line: 13 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	13	010100 0800 15.48 24.13 13.34 -4.90
Line: 14 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	14	010100 1000 19.03 35.63 14.17 3.49
Line: 15 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	15	010100 1200 14.27 61.67 6.23 7.02
Line: 16 Column: 0. Delimiter ',' not found after field 'Column1' (the record has...	16	010100 1400 15.20 29.61 12.16 -2.41

The form required of the following input

- The Data Source to which the import is relevant.
- Data file. The file containing the raw observation values.
- Log file. Only applicable to ACSYS historic data files where the log files will indicate data gaps.

When a critical error encountered during the file processing, the Error Pane will display the detail on each error. Errors shown in this error pane would have caused processing to be aborted and all processing done to the point when the error was encountered, rolled back.

5. The Data Processor

The processor starts up by gathering the schema configuration from the selected data source. From the configuration data the processor determines whether the “**Universal Schema Data Schema**” or the “**Sensor Specific Data Schema**” method is applicable. It does this by checking executing the following logic

- If there is a Data Schema defined on the **Data Source** the processor determines that the “**Universal Schema Data Schema**” is applicable. This indicates to the processor that each column of the raw data file defines a unique phenomenon offering. It validates that none of the sensors defined for this data source has got a data schema defined and that each SensorProcedure defined as part of the Schema configuration can be resolved based on the Phenomenon defined for each field in the schema. This also indicates to the processor that the import file needs be iterated (row by row) only once in order to satisfy the schema, that is, all observations(Phenomenon values) in a single row, can be processed by iterating through the file once
- If there is no Data Schema defined on the **Data Source** the processor determines that the “**Sensor Specific Data Schema**” is applicable. This indicates to the processor that each SensorProcedure defined to the data source should have its own data schema associated to it. This indicates to the processor that the processing will be done for each individual SensorProcedure. While processing the import file will be iterated as many times as too the amount of SensorProcedure defined for the datasource, that is, all observations (Phenomenon values) in a single row are gathered by processing each SensorProcedure(Schema) individually and then combining the result.

The Process then parses the raw input file contents to the contract defined as part of the Data Schema/s. If the contents of the raw import file satisfy the contract the processor continues, otherwise the Errors Pane of the “Import Data File” form is populated and processing aborted.

The processor then gathers the transformation process defined on the data source. Each value of the parsed data set is send through a data transformation process whereby the processor determines to which of the transformations the value is applicable and then the transformation of the value.

The final result is the transformed value to each record in the parsed data set.

The processor then split the parsed data set into two groups, valid observations and observations with discrepancies.

The processor then starts writing the parsed data to the relevant store (Data log / Observations). This process involves further validation just before the data is persisted. If a duplicate record exists in the observation store/data log the record will always be moved into the Data log for user intervention. The

status of the Data log will be updated with the relevant error information which should assist the user in resolving the conflict and correcting the data before processing it back into the Observations store.