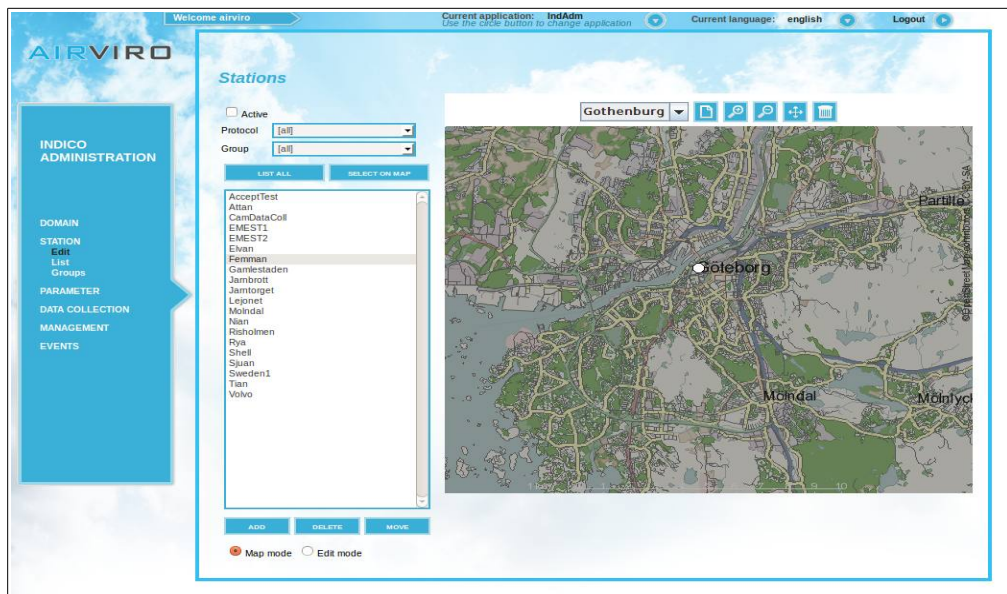# Apertum

Airviro User's Reference

Working with Indico Administration

How to fetch data automatically from remote stations

Amendments

Version	Date changed	Cause of change	Signature
3.11	May 2007	Upgrade to Airviro version 3.11	GS
3.12	January 2009	Upgrade	GS
3.13	January 2009	Upgrade	GS
3.20	August 2010	Upgrade	GS
3.21	Dec 2010	Upgrade	GS
3.21	June2012	Review	GS
3.22	April 2012	Upgrade	GS
3.23	Nov 2014	Upgrade	GS
4.00	April 2015	New Version	GS
4.00	Aug 2018	Review	GS
4.01	May 2019	Upgrade	GS
5.00	Nov 2019	Review	GS
5.00	Dec 2020	Review	DC

Contents

6.1 Introduction

Automatic data collection is impressive once it is working, but setting up a system to handle it is not a trivial task. The Indico Administration module has been built to help with this task so that the whole process is user friendly, but nevertheless, a good understanding of how each part of the data collection works is important if you want to be able to sort out problems.

Data is usually gathered using modems. The Airviro server first runs an **external protocol**, which contacts a specified remote **station**. The station answers and waits for instructions. The external protocol then sends instructions to the remote station including a request for the latest data. The station responds to the instructions, the computer receives the new data and the conversation is complete. The new data is then loaded into the **time series database** so that other programs (such as the Indico Presentation Client) can access it.

Automatic data collection is carried out via external protocol programs. Every external protocol is different, because each type of logger is different. This means that the documentation cannot really become specific enough to be able to lead step by step through the external protocol editor where all the external protocol information is given.

The Indico Administration module is one of the three main tools in Airviro along with Indico Presentation, and Indico Report and Indico Validation, intended for acquiring, storing, editing, presenting, analysing, reporting and exporting time series data. All Airviro modules version 5.00 or higher that require or use geoposition information or geotracking, use OpenStreetMap (OSM), which is an open-content free licensed map of the whole world. The standard OpenStreetMaplayer provide information such as street names, facilities, shops, etc

Swedish Meteorological and Hydrological Institute, SE-601 76 Norrköping, Sweden
Phone: +46-11 495 8000, Fax: +46-11 495 8001

Page 5 (32)

6.2 Getting Started

Once Airviro has been properly installed on the *server*, you can start using it by typing the correct URL in your web browser over the Intranet/Internet. (*Figure 6.2.1. Input User and password*).

 After logging in to Airviro with your user-ID and password, a domain must first be selected, and then **Indico Administration** should be chosen from the available modules.

 At the top of the window, in Current application the name of the module actually selected is shown . Clicking on the [down arrow] button beside the name other modules can be selected.

Clicking on the Current language [down arrow] button it is possible to select the language to work with. This functionality is not yet implemented.

By clicking on the **Logout [**down arrow] button, the current module is closed and the Airviro login page is displayed instead.

Figure 6.2.1. Input User and password.

The main window is split up into different sections (*Figure 6.2.2*):

Main Menu

DOMAIN   **For** selecting the domain to work with.

STATION   Allows you to access the station interface, where it is possible to create, modify or delete existing stations. Automatic data collection is configured here as well.

PARAMETER   Similar to STATION. It allows you to create, edit and delete parameters.

 DATA COLLECTION Used to control the data collection. Here the state of the automatic data collection can be changed .

MANAGEMENT   This enables you to alter the state of the Airviro database manager.

EVENTS   Allows you to define events for time series and stations.
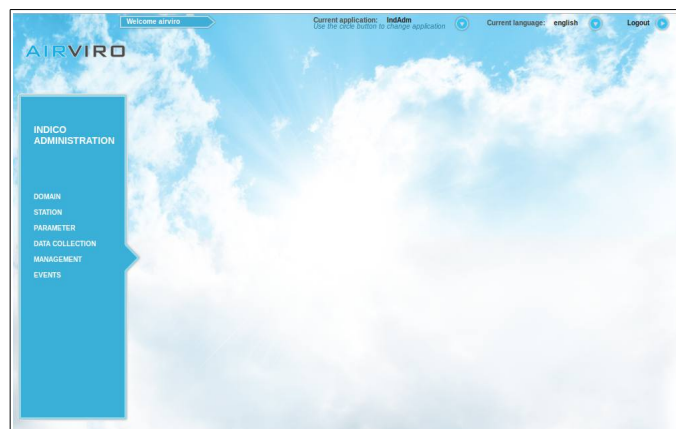
Figure 6.2.2. Main window.

6.3 Viewing and Editing the Parameter Database

The parameter database contains information about the different substances measured. This includes gaseous pollutants such as sulphur dioxide, carbon monoxide, etc. and also meteorological parameters such as temperature and wind speed. A parameter can be anything that is measured as a variation against time.

In the **PARAMETER** menu**,** clicking on the option **List**, a new sub-window will pop up. (see *Figure 6.3.1*)

```
# Key  Descr      Unit    Alim  Gmin  Gmax   Imin    Imax Maxeq  Eps    Grad
# ------------------------------------------------------------------------
  0001 SO2        ug/m3    100    0    100   0e+00   1e+07   6 0.0e+00 -1e+00
  0002 NO         ug/m3   5000    0   1000   0e+00   1e+04   6 0.0e+00 -1e+00
  0003 NO2        ug/m3    100    0    150   0e+00   1e+07   0 0.0e+00 -1e+00
  0004 CO         mg/m3    100    0     15   0e+00   1e+05   6 0.0e+00 -1e+00
  0007 HCl        ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0008 O3         ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  000C Kol        ug/m3    100    0    100  -5e+00   1e+04   0 0.0e+00 -1e+00
  0017 Ammoniak   ug/m3    100    0    100   0e+00   1e+04   6 0.0e+00 -1e+00
  0018 Hg         ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0021 Bensen     ug/m3    100    0     10   0e+00   1e+04  24 0.0e+00 -1e+00
  0022 Toluen     ug/m3    100    0     30   0e+00   1e+04  24 0.0e+00 -1e+00
  0023 P-Xylen    ug/m3    100    0     10   0e+00   1e+04  24 0.0e+00 -1e+00
  0027 Formaldeh  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0040 Klor       ug/m3    100    0    100   0e+00   1e+04   6 0.0e+00 -1e+00
  0042 CO2        ug/m3    100    0    200   0e+00   1e+05   6 0.0e+00 -1e+00
  00EQ Ekoreture  st       500    0    500   0e+00   1e+05  24 0.0e+00 -1e+00
  00HC HC         ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  00hc simHC      mg/m3    100    0    200   0e+00   1e+04   0 0.0e+00 -1e+00
  0820 ethane     ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0830 propane    ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  084i i-butane   ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  084n n-butane   ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  085c cy-pentan  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  085i i-pentane  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  085n n-pentane  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0910 ethene     ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0920 propene    ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0931 1-butene   ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  093c ci2-buten  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  093i i-butene   ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  093t tr2-buten  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0941 1-pentene  ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
  0951 1-hexene   ug/m3    100    0    180   0e+00   1e+04   6 0.0e+00 -1e+00
```

Figure 6.3.1 List of the parameters

 The parameter key is shown in the first column, this is a unique 4 letters identifier which identifies each parameter. The second column contains a parameter description (ie, such as a name), so that you know to which parameter the key is refering to. The fourth column was initiallyused to define alarm limits, but   it is no longer in use (concentration level alarms can now be implemented using Indico Alarms).

The information in the third, fifth and sixth column   is intended to be used by the Indico Presentation Module. They show the measuring units for each   parameter , and   the default minimum and maximum scale values for each parameter,   so that correct scales are used to generate the graphs.

The information shown in the last five columns is used to check against any incoming data associated with a particular parameter, unless the data has already failed a check by an external protocol. **Imin** and **Imax** are the minimum and maximum acceptable values, and **MaxEq** is the maximum number of equal consecutive values that will be accepted, plus or minus the value **Eps**. **Grad** is the maximum difference allowed between two data values that differby one hour. If incoming data fails one of these checks, it will be assigned a status indicating that the check has failed.

6.3.1 Creating and Editing Parameters

On the main menu, click on **Edit** under **PARAMETERS** in order to open the editing window

Figure 6.3.2. Edit parameters

On the left side there is a list with all the parameters currently defined. To view the details stored for a given parameter, just click on the parameter name. All the information that was displayed in the parameter database summary (see *Figure 6.3.1*), is now presented in editing mode, so you can modify it. Click on the [**ADD]** button at the bottom of the list if you want to create a new parameter.

The most important piece of information for each parameter is its key, and once a

parameter has been assigned a key to it, it cannot be changed. The parameter also requires a meaningful name or description so that you can recognise it in the parameter list. Fill in these information in the textboxes beside to **Key** and **Name** respectively.

Then fill in the **Presentation information** section. The **Unit, Graph min**. and **Graph max**. are required by the Indico Presentation module for generating graphs involving parameters. It is always important for a person looking at a graph to know which measuring units are being graphed. For most pollutants this will be μg/m3. You can specify up to 4 new units. **Offset** and **Factor** are the conversion factors used for converting   the default unit of the parameter into another measuring unit. Offset is needed for conversions between measuring units, like Celsius and Fahrenheit (*Figure 6.3.2.*)

Fill in the likely minimum and maximum values that would be required for the graph axis scale. These can then be changed in the Indico Presentation module but default values must be entered when the parameter is created, as these are required when a graph is first generated.

6.3.1.1 Checking the Incoming Data

Now move onto the **Data checks** section. Here it is possible to define various limits to check the incoming data. All incoming data is checked against the limits defined here and a status code is assigned to them,   indicating if data is OK or otherwise the reason why it has not passed the check. The data is ,however, still stored in the time series database whether they are checked OK or not (see Volume 5 -Indico Validation). Users of the Indico Presentation module can choose only to use data with certain status codes.

The **Alarm limit** is no longer used. A consecutive number of equal values often indicates an instrument failure. Set the **Max number of equal values** to what you consider would indicate such a failure. The **Allowed variation** is related to this - if the difference between two adjacent values is less than the allowed variation then they are considered to be equal. Instruments usually have a measurement range outside of which their readings cannot be relied upon. Enter these limits in the textboxes labeled **Absolute min** and

Absolute max. **Max gradient** is the maximum difference allowed between two hourly values of incoming data.

The error messages generated by the data collection are stored in the   /usr/airviro/log/ directory, these files are named avlog.YYMMDD (where YY is the year, MM the month and DD the day). These files are usually kept for a month before being deleted by the system.

If the **Log Error Status** box is checked, a log entry is added in the avlog file for both parameter and station for a time series data with a bad status code.

Once you have completed all the required information to define the parameter click on the [**APPLY]** button to save your data.

6.4 Understanding the Station Database and Setting Up a Station

6.4.1 Viewing the Station Database

A site where data is measured is called a 'station'. In the **STATION** menu**,** clicking on the option **List**, a new sub-window will pop up. (see *Figure 6.4.1*)

```
Station              Op   Bad  Update       Next
-----------------------------------------------------------
265   Sweden1        NO    0s  700101 01:00  700101 01:00
GM1   Shell          NO    0   900122 12:00  700101 01:00
GM2   Lejonet        NO    0   910823 12:00  910823 13:35
GM3   Jarnbrott      NO    0   910116 15:00  910701 13:10
GM4   Risholmen      NO    0   910116 15:00  910701 13:10
GM5   Femman         NO    0   900101 01:00  700101 01:00
GM7   Sjuan          NO    0   900122 12:00  700101 01:00
GM8   Attan          NO    0   070911 02:57  070912 02:57
GM9   Nian           NO    0   070911 02:58  070912 02:58
GMA   Tian           NO    0   070911 02:58  070912 02:58
GMB   Elvan          NO    0   070911 02:59  070912 02:59
GO1   Gamlestaden    NO    0   900122 12:00  700101 01:00
GO2   Molndal        NO    0   900122 12:00  700101 01:00
GO3   Rya            NO    0   900122 12:00  700101 01:00
GO4   Volvo          NO    0   900122 12:00  700101 01:00
GO5   Jarntorget     NO    0   900122 12:00  930407 00:00
GOX   EMEST1         NO    0   900122 12:00  700101 01:00
GOY   EMEST2         NO    0   900122 12:00  700101 01:00
TST   AcceptTest     NO    0   700101 01:00  700101 01:00
XY1   CamDataColl    NO   12s  150304 07:05  150305 11:40
```

Figure 6.4.1. Station List

The first column shows the station key. A unique 3-letter identifier that refers to each station. The second column shows the station name.

The third column shows if the station is in operation or not, i.e. if it is due to be called automatically at some specific time. The next column shows the number of bad (unsuccessful) calls and can also show whether if any alarm limits have been reached. The last two columns show the date and time of the last successfully fetched data, and the next call time (this is only relevant for operational stations).

6.4.2 Creating / Editing Stations

On the main menu, click on **Edit** under **STATIONS** in order to open the editing window

On the left side there is a list with all the stations currently defined. In order to fill in the list click on the **[LIST ALL]** button or select an area on the map by clicking on **[SELECT ON MAP]**.

When you click on [**SELECT ON MAP]** a cross mouse pointer will allow you to select the area on the map that you are interested in. Also you can select a particular area using the ctrl+left mouse button. A white hair cross will appear when you move the mouse pointer onto the map. Move the hair cross to a geographical point that you want to define as a corner in the zoomed area. Click on the left mouse button and drag the mouse to stretch a rectangle. Release the mouse button when you are satisfied with the marked area. Select an area with the mouse by pressing and holding the left mouse button, dragging it over the area and then releasing it.

By checking the **Active** check box above the stations list, only active stations will be shown. (*Figure 6.4.2.*). You can also filter the stations shown in the list according to the protocol used by the stations or by belonging to a certain station group.

Protocols require a good deal of information about the data logger set up which will be contacted, We will see more on protocols later.

StnGroups can be created in Indico Administration. A station can then be assigned to one or several station groups. This is useful if you have a large number of stations.

To view the location of a given station on the map, just click on the station name (make sure the option **Map Mode** is selected), or double click on it for editing the station.

Double click on a station, e.g. Molndal. The **Station Edit** sub-window appears. Once a station was selected on the list, you can also use the toggle button **Edit mode** below the station list to open this sub-window . If you want to create a new station, make sure the option **Map Mode** is selected and then click on the [**ADD]** button below the stations list. When the cross pointer appears, you can hover over the map, and once you found the spot where the new station must be placed, click on it . Then the station edit sub-window will then reappear.

6.4.2.1. Name Information

Start by filling in the station name in the text box labeled "**Name"** and enter a meaningful and recognisable name to identify the station. Then fill in the **Key**. This must be a unique 3-letter identifier, and conventionally follows a naming pattern. The first letter represents the location of the station, for example all stations in Göteborg have G as the first letter of the key. The second letter usually describes the type of data measured at the station. The most important types are O for Opsis analysers and M for Campbell loggers. The last letter distinguishes between different stations of the same type. You do not have to follow this convention - as long as the key is unique.

Figure 6.4.2. Edit Station.

6.4.2.2. Main information & Limits

Next comes **Main information & Limits** , though this is actually the last step in the configuration. Once the station is fully configured for automatic data collection, fill in the date and time for the next data collection in the text box **Time to call** and tick on the check box **Operate** to make it active. Once data collection is running the **Latest update** field automatically updates itself.

If the **Log Error status** box is checked, a log message will be written to avlog for both parameter and station for a time series, and the incoming data will be stored in the time series database with a bad status code,

 Argument: . Each protocol requires a specific argument that indicates the starting date and time of the data to be collected. This has to be wrtitten using the correct format required by the station that is being called

This takes different formats for different protocols - for example, Opsis loggers take a normal date and time format but Campbell loggers require a pointer which does not seem to have anything to do with time at all.

If you are using the "ops" protocol then the argument is the date and time e.g. 000508 12:00. If the OPSIS station has several paths then there must be an argument for each path. Multiple arguments should be separated by a space.

If you are using the "cam" protocol then you must give the pointer position e.g. 1046. If you give a pointer position 0 then all data in the logger will be fetched.

Write a suitable argument in the text box **Argument**. This argument is automatically updated after data has successfully been collected from the station. In general, a resource is also required as a means for contacting the station, and this is usually a modem.

You can define your own **Alarm limit** and **Stop limit**, which typically take values of 5 and 10 respectively. When the number of consecutive bad calls reaches the alarm limit, a warning message can be sent, and when the stop limit is reached another message can be sent and the station ceases to be operational. When this happens you must first find out and solve the problem, then reset the number beside **Bad calls** to 0 and make the station operational again.

Note: When creating a new station, it is usually much easier to copy an existing station and then

Swedish Meteorological and Hydrological Institute, SE-601 76 Norrköping, Sweden
Phone: +46-11 495 8000, Fax: +46-11 495 8001

Page 15 (32)

modify it as required for the new station. First select the station that you want to copy and then click on the [**Add]** button. All the information will be copied to the new station

Airviro systems have been successfully used with loggers from the following manufacturers:

- Monitor Labs/Monitor Europe

- Thermo (42i, 43i, 48i, 49i)

- Met1

- BAM1020Campbell (always used for Airviro principal masts)

- Opsis Analyzers using ComVision

- Opsis loggers

- Horiba

- Envidas

- Philips / DMS / A30

- Odessa

- Nilu

- Aanderaa

- Dasibi

- Marksman

- API

- ESC

- Ecotech

- TEOM

- All loggers / instrument with support for the Extended Bavarian protocol

6.4.2.3. Configuration

- **Restricting the Call Times**

Click on the [**Time restrictions**] button in the Configuration section to open this sub-window, . Here you can restrict the times and the days of the week when the system will try contact the station. Under **Hour,** click on the hours that the station may be contacted (click on the **[All]/[None]** buttons to select / deselect them all). The hours that will be used are those that are highlighted, the same for Weekdays.

Figure 6.4.3.Time Restriction

Under the **Misc. Information** section we have:

Call int. (call interval, in minutes). This indicates for how long the system waits between each call to the station. Of course this depends on the restrictions that you have made under **Hours** and **Week days** - if you have restricted calls to be only between the hours of 6 and 7 then the station will only be contacted between the hours of 6 and 7, even if the call interval is 180 minutes. Try to call fairly frequently to avoid loss of data - it is best to call at least once a day to collect hourly data, and if you want to have fairly recent data during the day you will need to call 3 or 4 times a day, or even every hour.

 Retry int. is the time to wait (retry interval, in minutes) after a bad call before retrying. The optimum time to wait varies between systems, but 3 minutes is quite common.

Start min (minute) and **Stop min** (minute) restrict the parts of the hour to use when calling a station. For example, if **Start min.** is 10 and **Stop min.** is 30, then calls to the station may only be made between 10 past and half past the hour. This can be very useful in helping to avoid more than one process trying to use the modem at the same time.

Max factor is the maximum number of seconds or minutes per hour that the station is allowed to gather data. For each call the maximum allowable call time is calculated, which depends on the number of hours of data to be collected, plus a default connection time that is added because it sometimes takes a while to establish contact with a station. If a call exceeds its maximum call time then the process is terminated by the collection daemon. A value of 0 here means no restrictions.

Click on [**OK]** to save you choices and to return to the **Stations** sub-window.

- **Station group**

You can associate a station to a **Station group** . These groups can be created clicking on the option **Groups** under the **Station** menu on the main menu . A station can then be assigned to one or several station groups. Then when you look at the time series data in Indico Presentation you can choose to only look at stations for a particular station group. This is useful if you have a large number of stations.

Station groups can also be used for dealing with mobile stations. Every time a mobile station moves it becomes a new station with a new station key. However the mobile station can have its own station group so that it is easy to use the Indico Presentation module to select all stations that are actually the same, mobile station.

- **Defining the Protocol to Use**

 From the drop down list select which data collection protocol to use, Then click on the

[**PROTOCOL SETTINGS]** buttonin order to complete all the information required to set up the data collection protocol selected. (for setting up protocols see section 6.4.3)

- **Miscellaneous information**

Click on [**MISCELLANEOUS]** and sub-window will appear.

Figure 6.4.4. Miscellaneous information

In **Geographic info** you can move the station and define its height above the ground.

Time Information: The **birth time** of a station is automatically stored here when a new station is created. If a station ceases permanently to be operational then you may want to call it "dead" and you enter the **death time**. In the future this might be more relevant and mostwill be used with mobile stations.

Under **Protocol,** Resources can be the name of the resource (device file) such as

`/dev/ttyS0` or otherwise to specific from a set of modems it can be specified by:

      type (T)   Indicates the type of resource (ex TModem, TDirect).

      speed (S)   Specifies which speed the resource supports (ex S2400 or S1200).

      property (P) Specifies a certain type of modem (ex PHST, PDiscovery)

Most external protocols require a resource to be used (such as a modem). For systems that have several modems with different capabilities, it is possible to just specify a resource type to use, so that any available resource matching that type can be used.

Alternatively, you may instead just specify a specific resource, but it is better to designate the resource type, as this is unaffected if the actual resource changes its name or if more resources become available when the system is expanded.

If you have got 3 modems, two of them working at 2400 bauds and one at 1200, then you would prefer to use a 2400 modem, therefore, in the text box labeled "resources" you should write:

<div align="center">Tmodem, B2400 Tmodem, B1200</div>

This means "use type modem at 2400 bauds, but if none is available use type modem with 1200 bauds".

 Statistics shows a summary of good and bad calls to the station

Now fill in the **Attributes.** Here you can specify three additional features related to the data that is collected. Choose **Logging** if you want to keep a log of all transactions with the station (you usually will use this option with live stations). Choose **Dummy station** if

the station is not actually operational but data appears at regular intervals anyway (for example if it is distributed from another computer system). Setting the **Dummy station** flag means that you can set the **Operate** flag under **Main information** but it will be ignored by the data collection daemon.

Choose **Save raw data** if you want to keep a copy of the incoming data in the original format as well as in standard format in the time series database. Sometimes this is needed to sort out problems with data collection but in normal cases is not necessary.

Now fill in the **Contact information**. Here you can specify additional information about the contact, including her/his email.

Click on **OK** to save your changes.

6.4.2.4. Error Message

 This text box, is intended to visualise error messages generated by the data collection daemon.

6.4.2.5. Apply your choices

Click on [**APPLY]** to save your changes.

6.4.3 Setting Up the External Protocol Information

This is the most complicated part of setting up an operational station and it is intended   to

be used only by experimented users who have had system administration training. It requires detailed information about how the logger has been configured and which parameter is measured on each channel.

Choose a protocol from the protocol drop down list that appears above the stations list, in order to filter the stations displayed according to your selection.

Now choose a station from the list, and double click on it or otherwise click on **Edit Mode** (below the station list) in order to open the editing window.

Once you have loaded a station your external protocol editor may look like the one shown in *Figure 6.4.5.* depending on the protocol selected (a different input windows is available for some other uncommon protocols, this editor is not explained here). The section **Level** displays a list of properties. Each property specifies some specific information about the logger such as the way the channels are configured or the way in which it can be contacted.

Each property (Level) must therefore be given the correct value or values (Number). A property will enable you to edit it.

With the button **[Root]** you have the possibility to return to the root level. There you will be able to add properties and to enter values to each one of them.

You can clone (copy) these complex properties until you have the required number of them.

Activating the **filter level** allows you to display only the current level.
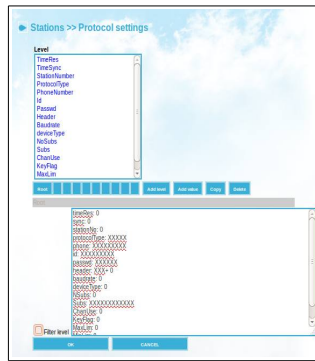
Figure 6.4.5. Protocol Settings

Note: you can request to the System Administrator to configure the "Import" protocol used to extract existent data from a station.

Once you have completed all the required the data for the external protocol, save it by cliking on [**Ok]** button of the protocol editor

6.4.3.4 Starting the Data Collection

It is now time to test the automatic data collection. Return to the Stations Edit window, and select your station from the stations list. Select **Edit Mode** and activate your station by ticking on the check box **Operate** . Fill in **Time to call** with a time 5 minutes from now. Fill in **Latest update time** as some time a few hours ago. It is not a good idea to try and collect too much data at once, especially with a new station, which might not work properly at first, as it is easy to forget to define something important. Save the station and wait for the call time. Check that the station appears in the data collection idle queue (click on **DATA COLLECTION** on the frame). Watch the modem and make sure it makes the calls at the correct time.

When it has finished the call, look in the data collection queue again to see if the call was successful (i.e. 0 bad calls). If this is the case then you have successfully set up automatic data collection, otherwise you will have to go through all your settings very carefully and to work out what is wrong.

6.5 Examining the Data Collection Processes

Click on **DATA COLLECTION** menu. This window displays the processes that are under the control of the data collection daemon (**cold**), together with their current status. The state of **Collection Mode**, in most common cases should be MULTIPLE, i.e. it can deal with more that one data collection process at any one time. The database(s) that currently have operational stations are displayed. In most cases there will be just one database although in some cases it could be more. The other important things here are three queues, called Idle, Wait and Running.

The **Idle** queue contains all the stations that are operational. These stations are to be contacted at the time shown under **Next**. You can also see the number of bad calls and the resources that may be used.

The **Wait** queue is for stations that are due to be contacted but have to wait until a resource is available, for example if the modem that is to be used to contact the station is taken up by another process.

The **Running** queue shows any current activities - all stations from which data is currently being fetched. The length of time that the current collection has taken so far is also shown.
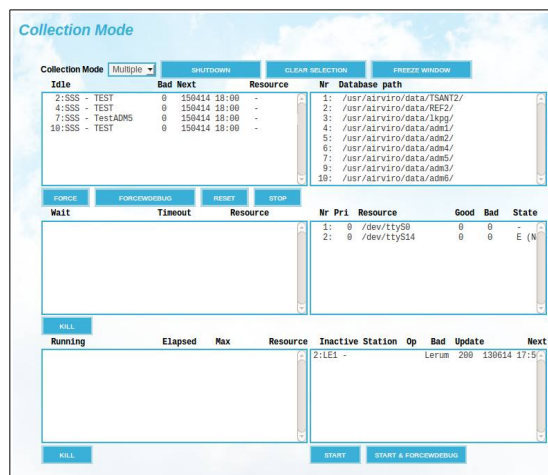

Figure 6.5.1. Data Collection

At the left is a list of all the resources that are available to be used by the data collection daemon. These resources keep a count of good and bad calls that they have made since the collection daemon was started.

This summary of the data collection processes only describes what was happening at all the moment.

6.6 Examining the Database Management Processes

6.6.1 Management

Click on **MANAGEMENT** in the menu, and select **Data Base**. This brings up a sub page where you can view or modify the state of the Airviro database manager. The database manager is storing incoming time series data in the time series database. The manager can run in different modes:
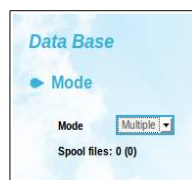
Figure 6.6.1. DataBase-Management

- Off: the database manager is not running. No data will be written to the time series database. The data is stored in a spool archive and when the mode is changed to Multi the data is written to the database.

- Idle: the database manager is running but does nothing. No data is written to the time series database.

- Single: the database manager is running but it only takes care of the first domain that is specified in the configuration file for the time series database manager.

- Multi: the database manager is running and writes data to all domains.

With **Spool files** you can identify the number of spool files and invalid files (0), in the spool directory of the current Domain. If the database manager is running in mode multi, the number of spool files should be 0. Invalid files should always be 0.

HINT: For a normal system running data collection, the state of both of them should be **Multi**. This means that the database manager can deal with several domains.

6.2.3. Distribution

Click on **MANAGEMENT** in the main menu, and select the option **Distribution** In order to save time series data for distribution. Airviro stores this data in a file that is located in /usr/airviro/data/<domain>/dist/ directory. Using the Dist, Fdist and Hdist protocols the data can be sent to other Airviro systems.
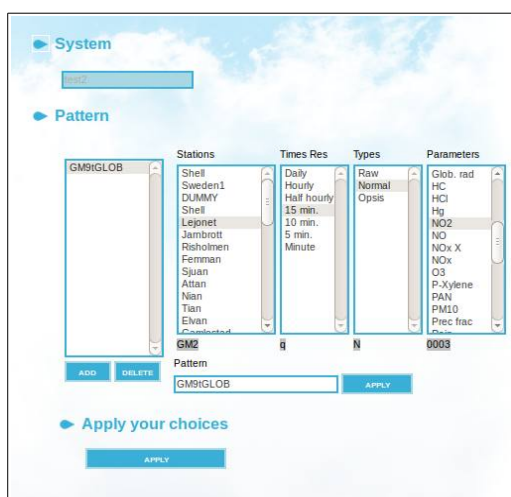
Figure 6.6.2. Distribution Systems

To decide which time series should be saved in distribution files a system based rule is used. The rules are based on the five first letters of the time series key: **SSSRT**, where the three **S:**s are the station key, **R** the time resolution and **T** the type of data. It is possible to define rules using pattern matching. A **?** means any character, a ***** means one or more of any characters.

A rule that is only a * would save all incoming data in the distribution file. A rule that is GM5+? would save all data from the station GM5 with time resolution + (hourly) with any type. A rule that is ???+M would save all normal hourly data in the distribution file.

The information displayed in the lists Stations, Time res, Types and Parameters is only to help you with the time series key letters. By clicking on the names in the lists, the corresponding codes are displayed below. Using these lists it is possible to define a time series pattern.

Airviro can distribute data to one or more systems. These are shown in the leftmost list. Each system can have one or more distribution rules defined. These rules are listed in the **Pattern** list.

A distribution directory with the same name as specified in **System,** is created under */usr/airviro/data/<domain>/dist/* directory.

6.6.3 Time Series Post Processor

Incoming data stored in the time series database can be re-used or post processed to calculate other values A typical post processing is to calculate averages and sums but also to scale raw values to scaled values (See Indico validation user's reference).

To decide which time series should be post processed a system based rule is used. The rules are based on the 12 letters of the time series key: **SSSRTPPPPIII**, where the three

S:s are the station key, **R** the time resolution, **T** the type of data, the four P:s the parameter key and the three I:s the instance. It is possible to define rules using pattern matching. A **?** means any character, a ***** means one or more of any characters.

For each rule an action must be selected and that action then is executed on the data that fulfilled the rule. Available actions are: Calculation of 30 min averages, Calc hourly avg, Calc daily avg, Scale, Hourly sum, Daily sum, Calc NO2, Adjust. TEOM.

A rule that is only a * would perform the action specified on all incoming data. A rule that is GM5+* would apply for all time series with station GM5, time resolution + (hourly) and with any type, parameter and instance. A rule that is ???+MTEMP??? would apply to all time series with hourly resolution, normal values and with TEMP as parameter.

The information displayed in the lists Stations, Time res, Types and Parameters are only to help you with the time series key letters. By clicking on the names in the lists, the corresponding codes are displayed below. Using these lists it is possible to define a time series pattern.

The rules are applied to incoming data in sequential order starting from the topmost rule in the list down until the rule that applies to that data is found and then executed ending the process. This process is then repeated with the next data received, and so on.
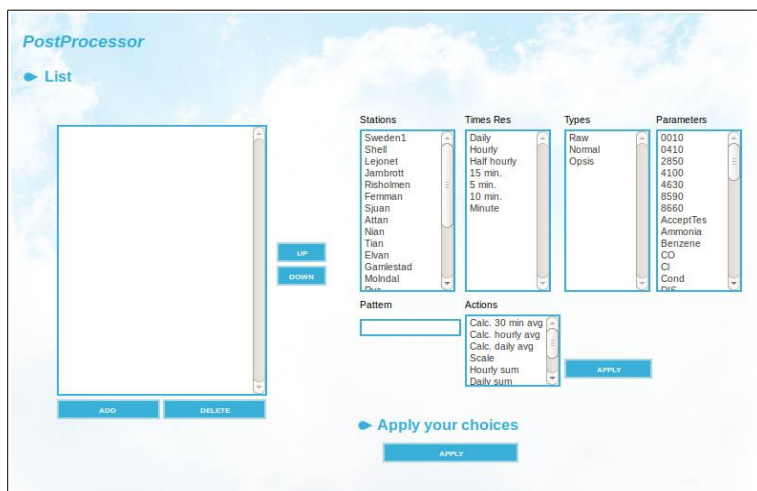

Figure 6.6.3. Postprocessor

6.6.4 Data Checks

Data checks are usually performed for individual parameters but here it is possible to add checks based on time series. The rules defined here will apply to all incoming time series data.

In order to decide which time series should be checked a system based rule is used. The rules are based on the 12 letters of the time series key: **SSSRTPPPPIII**, where the three **S:**s are the station key, **R** the time resolution, **T** the type of data, the four P:s the parameter key and the three I:s the instance. It is possible to define rules using pattern matching. A **?** means any character, a ***** means one or more of any characters.

The rule GM5+MTEMP???   would apply for all time series with station GM5, time resolution + (hourly), type normal, parameter TEMP and for any instance.

Performed checks for each rule are: **Minimum**, **Maximum** and **Suspect.** These checks change the status code and writes a message in to the avlog file. **Alarm** writes an alarm message to the avlog file that contains the **Comment**.

The information displayed in the lists Stations, Time res, Types and Parameters are only to help you with the time series key letters. By clicking on the names in the lists, the corresponding codes are displayed below. Using the lists it is possible to define a time series pattern.

The Indico Real Time module reads the avlog file periodically and when a new alarm appears in the log file the Alarm button starts to flash red and an alarm sound can be heard.
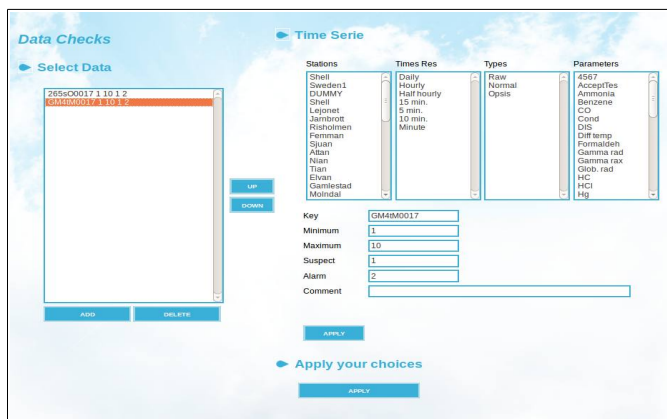
Figure 6.6.4. Data Checks

6.7 Events

These events can be used in Indico Validation. They are used to record events that have happened for a station or a time series as well as adding comments to the validated time series.
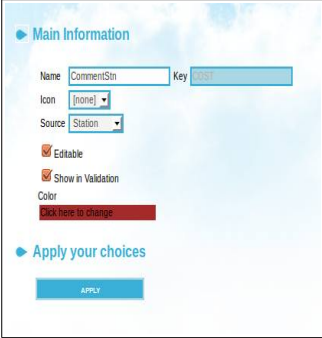
6.7.1 Type

You can define different Event Types. For example:

- Events for a station such as when somebody visited the station or a power failure occurred.

- Comments for a time series, for example when a data is deleted or modified.

- Event messages coming from data collection process.

For each type, you must define a name, an Event Key (4 characters) an icon type and a source (time series or station). Event Types can then be selected in Indico Validation.

You can change the colour assigned to an event.

Figure 6.7.1. Data Checks.

6.7.2 Icon

You can input an icon (upload an image file) and associate it to an event type.

6.7.3 Edit

You can select an event from the list and edit it.

For station, the information shown is: name, date and the event text.

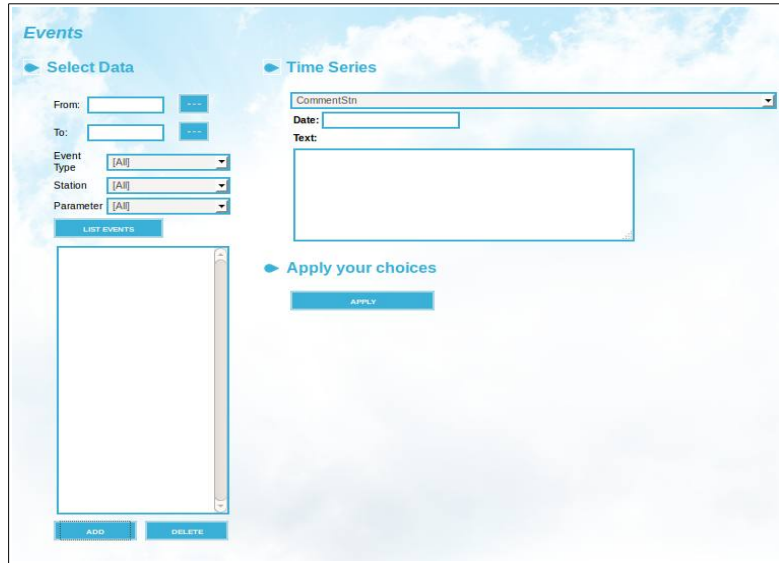For time series, the name, resolution, parameter, instance, date and event text are displayed.

Figure 6.7.2 Edit.