New concept for Belgian AWS network

Luis González Sotelino, Nicolas De Coster, Peter Beirinckx, Pieter Peeters

Royal Meteorological Institute of Belgium, 3 avenue Circulaire 1180 Uccle, Belgium
Tel: +32 (0) 23730699, Fax: +32 (0) 23746788, E-mail: Luis.Gonzalez@meteo.be

Abstract

The Royal Meteorological Institute of Belgium (RMIB) has a network of 18 automatic weather stations (AWS) recording meteorological parameters such as pressure, temperature at different levels, relative humidity, wind speed and direction, radiation and precipitation. The first RMIB station is now nearly twenty years old and a renovation is needed to meet new needs. Strengths and weaknesses of the actual AWS network will be sketched and requirements for a new AWS concept to address the actual weaknesses described. These requirements are based on the World Meteorological Organisation guide and AWS users' needs.

The status of this new concept, which is already partially implemented for a new AWS at Stabroek, will be presented and a road map for the future will be sketched.

1 Introduction

RMIB is responsible for the Belgian civil meteorology. Civil and military aviation AWS are managed by Belgocontrol [3] and Meteo Wing [14], which also is responsible for the road safety warnings. Environmental measurements and monitoring are the responsibilities of the regions (Brussels-Capital region [4], Walloon region [6] and Flemish region [12]).

The first AWS has been installed in the RMIB headquarters at Uccle in 1995. Since then additional stations have been added all over the country. Over the years, the AWS network has shown several weaknesses and several devices have become obsolete. This was an opportunity to build a new AWS based on experience gained during the last 20 years.

Besides this AWS network, RMIB collects data from a voluntary network of precipitation and temperature measurements of over 200 stations.

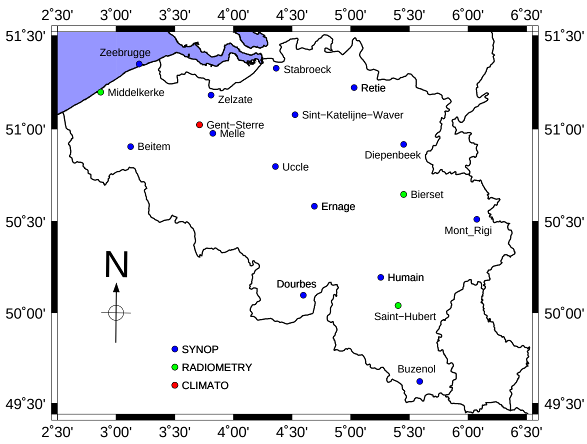

Figure 1.1: RMIB AWS Network (courtesy of Cédric Bertrand)

2 Description of Actual AWS Network

RMIB AWS stations are built from off-the-shelf components. All the electrical and mechanical engineering is done by RMIB except the masts higher than 3m. The instruments are all linked to a datalogger (`CR10` or `CR10X` from `Campbell Scientific`) and a phone connection is available to transfer the data using analog modem connection. Stations are polled every hour from RMIB via modems and the data are transferred to database tables (Oracle). Data collection software (`Loggernet` from `Campbell Scientific`) is run on two redundant `Windows` servers. Meteorological values (mean, minimum, maximum, standard deviation, gusts, ...) are computed on each datalogger for several time steps (10 minutes, 1 hour, 3 hours) and intervals (1 minute, 10 minutes, 1 hour, 3 hours, ...).

2.1 Strengths and weaknesses

Despite their good quality and intensive usage, the data suffer from incomplete and scattered information about the acquisition process, associated quality, uncertainty and processing. Although great efforts have been made in maintaining high standard of measurement quality in automatic weather stations, some more structural problems have been identified. The following paragraphs describe not only the technical and scientific requirements, but also management issues. AWS network management should take into account all aspects able to improve the final data quality, within a defined budget and limited human resources.

2.1.1 Weaknesses

Metrological Uncertainty Michel Leroy from MétéoFrance has defined a maintained performance classification [8]. This classification can be seen as the metrological consideration of its environmental classification. The latter has been added to the WMO CIMO guide in 2010 [16]. This is not the case of the maintained performance classification. This classification is from A (WMO requirement met) to D (very bad or unknown), B being the MétéoFrance goal for their network. The classification of RMIB AWS based on this scale has been realised for temperature (C), pressure (D), and humidity (D) [5]. The reason of this poor classification is different for each meteorological parameter: calibration procedure, lack of routine calibration and insufficient device quality. In particular, the calibration factors of the PT100 have not been taken into account in the data processing.

Poor scalability of stations The connection between datalogger and instruments is analog. Adding new instruments is only possible if free wires are available. The number of additional wires needed to add a new instrument or to replace it by a new type is usually high: in addition to measurement wires, heating wires and instrument status wires are needed.

Datalogger Code

Redundant Code The development of new stations of the network has been carried out at the datalogger program level by copy/paste of previous programs. So there is one program for each installed datalogger. This method has shown its limitation: updates are very time-consuming since each datalogger program has to be changed and are error-prone. Redundant code is strongly discouraged as a bad coding practice [11].

Hard Coded Parameters Calibration parameters are hard-coded in the datalogger code. A new calibration implies a new datalogger program. Hard-coding of parameters is strongly discouraged as a bad coding practice [11]. Parameters should be maintained in a separate file or a data base.

Assembler-Style Programming Language The programming language of CR10 dataloggers is hard to master: a training is necessary and team knowledge redundancy is difficult to maintain.

Lack of Quality Control on Raw Data While statistical calculations (averages, max, min, etc.) are made by the dataloggers, it is impossible to correct or sometimes even to detect different classes of problems without quality control on the raw data. Abnormal noise in data measurements, for example due to a defective multiplexer, will be mitigated by statistical operations that will return plausible, though inaccurate, results. Wrong calculations due to error in the electrical to physical transformation law or improper calibration data are difficult to correct afterwards and generally lead to data loss.

Raw Data not Recorded Statistical calculations (averages, max, min, etc.) are made at the datalogger level. Raw data is not saved. It is difficult to correct any data when new information becomes available, like calibration interpolation. On-site calculations had originally been chosen because telecommunications were judged too slow to transfer large amounts of raw data.

Acquisition Servers Redundancy Switch from one acquisition server to the other one in case of problems is done manually by operators. Parallel connections to acquisition servers are hard to manage, as the data acquisition software does not run as a service.

2.1.2 Strengths

Good Data Availability Better than 99% for the last year (2013).

Low Power Very low power consumption allows 2 days of AWS operations powered by a 12V 7Ah battery (except for heating of the instruments).

Data Acquisition Servers Synchronisation Synchronisation between servers is done using simple synchronisation software (`Allway Sync` [13]) which has shown to be a good trade-off between reliability and maintenance.

2.1.3 Additional Motivations

Datalogger End of Life The dataloggers used in the AWS have reached their end of life and are not available to purchase any more.

Acquisition Servers End of Life The maintenance contract of the servers has expired and the usual lifespan of the servers has been exceeded.

2.2 New Requirements

In addition, during the last years, new needs have arisen from the user side and from the AWS team.

Higher Data Retrieval Scheduling Most of the stations are polled every hour. The weather office and present weather application have requested 10-minute scheduling. This is not possible with the actual system: modems are too slow to allow a 10-minute scheduling for every station.

New Meteorological Measurement Users have required additional measurements like snow height, cloud height and evapotranspiration.

Use of Data for Climatology The data collected at the AWS stations were not intended for climatological studies. Several studies at RMIB have shown that the uncertainty of the measurements was not sufficiently characterised to allow proper climatological studies. The reason for insufficient quality for climate study is different for every parameter. A review of climate monitoring principles edicted by WMO [15] have shown several weaknesses in the data processing chain. Those weaknesses are different depending on the measurement type. In table 1, the principles have been split on 6 different requirements ranging from the quality of the instrument itself to the meta data management and splitted by meteorological measured parameters.

Requirements	Pressure	Relative Humidity	Temperature	Wind	Radiation	Rain
instrument quality	N	N	Y	Y	Y/N	Y/N
instrument siting	Y	Y	Y	Y	Y	Y
instrument maintenance	Y	Y	Y	Y	Y	N
instrument calibration	N	N	N	N	N	N
data processing	N	N	N	N	N	N
meta data	N	N	N	N	N	N

Table 1: Requirements for climate study and status of actual network (Y=meets WMO requirements, N=does not meet WMO requirements)

New Stations It can be seen from figure 1.1 that geographical coverage is poor in the south east part of the country (Hainaut region).

Better Reliability of Data It is not possible to stop acquisition during instrument maintenance. As a result, erroneous measurements could be stored in the data base. Correction is done afterwards, and this procedure can lead to bad data being delivered to clients, depending on scheduling of the maintenance. This situation should be avoided.

Improved Documentation, Retrieval and Sharing of Information Most of the metadata are saved in Excel files but not on the main RMIB server. Excel files have shown not to be well-suited to maintain metadata: concurrent access is limited, search possibilities are limited, automation is nearly impossible, ... Sharing of information inside the AWS team and with other teams, like the calibration laboratory, is done through files and e-mails. Calibration factors are copied to datalogger code, introducing an additional risk of error. Providing access to the metadata to third parties like the climatology department, requires the intervention of an AWS team member. Reports are disseminated through text or PDF files. There is a clear need of better knowledge management.

Automation and Simplification of Calibration, Maintenance, Instrument Updates Pressure control for calibration is still done manually. Automation will save time. Maintenance reports are produced manually. Automatic templates will improve metadata recording and will save time. Changing an instrument in a station involves several persons because of the lack of tools to automate repetitive tasks (datalogger code generation, calibration parameters update).

Improved Warnings of Failure The instruments and electric/electronic devices are not monitored. Failure is usually detected only after quality control has been applied on the data, which could be after several months. Warnings of failure should be raised as soon as possible when feasible.

3 Implementation

Due to the amount of work to be done to comply to all requirements, an implementation plan has been followed. The highest priority task is keeping the network alive: RMIB and outside clients depend strongly on the availability of the AWS network data. An extensive study has been conducted to determine the work load to maintain the actual AWS network [10], concluding that there is insufficient man power to address all issues.

Improved Knowledge Management A directory on the RMIB server has been set up to share information related to automatic weather stations. Templates have been created to ease creation of documentation. Procedures and information related to the stations have been written and shared. Access to information has been granted to RMIB team. Knowledge management has been improved following the Belgian guidelines for public organisations [7].

New Datalogger The dataloggers used in the AWS network are not available any more: a study has been carried out [2] to compare the available solutions. The main criteria were: analog input capacity, accuracy, price, ease of installation, power consumption and remote troubleshooting. `Campbell Scientific` dataloggers have shown to be the best solution according to those criteria for meteorological application. It has also been decided to use internet connections for data communication between dataloggers and RMIB acquisition servers. The speed of internet connections is sufficient to satisfy scheduling requirements, even for raw data collection.

Metadata Database WMO has defined a set of useful data not dealing directly with meteorological or physical values, but with measurement principles, description of environment near the station and other data related to the measurement itself. These form so-called "metadata" [1]. Climate monitoring principles have two points related to these metadata and require that:

- The details and history of local conditions, instruments, operating procedures, data processing algorithms and other factors pertinent to interpreting data (i.e., metadata) should be documented and treated with the same care as the data themselves.

- Data management systems that facilitate access, use and interpretation of data and products should be included as essential elements of climate monitoring systems.

A new and complex structure has been implemented at RMIB to allow metadata to be stored and followed in a complete and automatic way. This new structure would provide researchers and other customers with easy access to all information about the measurement process, from calibration to actual measurements in our automatic weather stations. Moreover, it has been structured in a way to facilitate daily work of our technicians and scientists to maintain this set of information and automate some time-consuming tasks such as dataloggers software maintenance.

The database is now implemented in such a way that the existing table structure remains stable and that all additional information is linked to measurements, from calibration to computation of electrical to physical values. This structure also allows software responsible for calculation to connect to the database and collect automatically all the necessary information to compute properly the conversion from electrical to physical values. When a problem has been detected and identified, if any of this information must be corrected, no information is lost and the data can be easily recomputed.

In the next paragraph, the word "device" refers to any "hardware part", from power transformer to actual meteorological sensors.

The new structure organizes different tables connected to each other to make a coherent set of information. The database has been structured in a as much restrictive as possible way to avoid incoherence.

A complete descriptions of all available tables is out of the scope of this article and would be daunting for the reader. We thus delineate briefly classes of available information.

These tables generally describe one of the following classes of information:

- **Types** Tables containing large classes of devices types such as "Max Thermometer", "Power transformer", "Barometer", etc.
- **Catalogs** Contain actual instantiations of the devices in "Types" tables. The most important tables are:
  - Device: They describe a device : brand, model, version, electrical characteristics and transformation law, electrical limitations in normal use (to detect the most obvious failures automatically), etc.
  - Locations: The different stations, stocks and laboratories, available locations in a given station, etc.
  - Maintenance: Maintenance operations for devices/stations.
  - Others: vendors, manufacturers, etc.
- **Calibrations** Contain the calibration corrections. These may be given in different forms : transformation on electrical or on physical values. A special convention has been adopted to chain transformations (from electrical to physical) and calibration corrections (on electrical or physical values) properly.
- **Events** Keep track of events that can occur. For devices, they describe usual operations such as calibration, displacement, maintenance, repair but also anomaly detection and other less preventable events.
- **Records** Contain measured values coming from devices. This also comprises hardware control parameters (fan speed, internal temperature of dataloggers, battery voltages, etc.).
- **Linking** Structure-related tables to handle n:n relations in database (classical database structure), linking table for external software connections to the structure (see on the following page Data Processing), etc.

A HMTL-PHP/AJAX web interface is being developed to allow technicians and scientists working on automatic weather stations to update and insert data in an easy way. This will allow users to modify the database without compromising its integrity, to generate configuration files that will allow automatic generation of datalogger programs, etc.

This interface will also allow different teams working on different tasks (such as calibration, maintenance, software development) to contribute to the same database and thus keep the complete measurement chain both coherent and efficient.

New Laboratory Calibration Procedures Our calibration laboratory is working on automation of the main procedures. Pressure calibration was adapted a few years ago to automate the different cycles and measurements. More recently, temperature calibration was adapted. Temperature calibration had previously been measured in the calibration laboratory by the dataloggers of the same type as in the stations. This introduces the datalogger measurement uncertainty twice in the global measurement uncertainty. Moreover, the uncertainty on datalogger measurement is difficult to estimate and clear information couldn't be obtained from the manufacturer himself, probably leading to an overestimation on the related uncertainty. Additionally, only three Pt100 could be calibrated at the same time.

The temperature calibration has now been replaced by a much more efficient and reliable one. It is composed of a Fluke "superthermometer" and a dedicated multiplexer allowing measurements of ten probes plus the Pt25 reference per cycle. In addition to providing more measurement channels, the corresponding measurement uncertainty is lower. As a result, the global uncertainty is now better characterised and lowered.

Data Processing To control the uncertainty on the global measurement chain, from calibration to actual measurement, and limit missing data, only raw electric/electronic values of instruments are transmitted by dataloggers in the remote stations, while all calculations are performed at RMIB. The raw values are archived, allowing one to recompute the meteorological data at any time. The expected storage size of all the raw data from our network was estimated at 40GB/year. This storage size can be managed by our database. Unfortunately, once stored in the database, the actual storage size is ten times this amount. A solution has to be found, as this size can not be managed by the RMIB database.

Raw data values are measured at a rate of 4 samples per second for the wind values and at 1 sample per 5 seconds for the others. These data are not relevant on the meteorological scale, and a fortiori on the climatological scale, but allow statistical calculations and detection of some usual electrical problems. Direct access to the raw data values is also highly valuable to diagnose problems remotely.

These short time-scale raw values are used to generate a first 1-minute meteorological table that subsequently allows generation of other database tables available for customers (10-minute values, 1-hour values, synoptic tables, etc.). This separation between electrical and physical values and the short time-scale physical database leads to more convenient management of the customer-oriented databases. It will allow reprocessing of the data and the possibility to create new products from past data.

The processing will be done in one place (acquisition servers) for all stations: maintenance and updates will be simplified in comparison to one program per station.

In order to maintain consistency, the data will be processed at RMIB in exactly the same way as formerly done in the old dataloggers.

New Acquisition servers New acquisition servers have been set up using the CentOS 6.4 Linux distribution and Campbell Scientific Loggernet Linux version running as a daemon. The Linux server can be connected to by Windows clients for remote operation as data transfer supervision. The servers are synchronised using gluster which can be used as a simple synchronisation solution [9]. During the implementation phase of new stations, we have to deal with network and modem connections. It has been decided to purchase multi-RS232 cards to connect from new servers to stations modems. Unfortunately bought cards are incompatible with last `Linux` distributions despite tests on previous versions and compatibility notice. In the meantime it has been decided to use the old servers for modem connection and the new servers for network connection. New RS232 cards have been bought and tests will be carried out in the next months before migration of all stations to new servers.

Datalogger Code Datalogger programming and maintenance were found to be time-consuming. Moreover, they use an inconvenient and specific programming language difficult to learn and manipulate for technicians working on site. This has led us to create an easier way to produce datalogger programs.

Using the database structure, one can generate a simple comma-separated value file (.csv) containing a list of available instruments in a given station with their positions and serial numbers. The datalogger channels and multiplexer configurations can be specified. A dedicated program (written in C++) is then used to transform this .csv file to a complete and ready-to-use program for dataloggers. This allows any of our collaborators to manipulate and adapt programs in our stations, simplifying maintenance greatly and saving time spent on programming. It is also possible to stop acquisition for instruments being maintained, using datalogger flags, and then to avoid filling the database with suspicious data.

Device Status It has been decided to record the status of all devices and to monitor fans using an electronic card developed in-house.

Issue/Requirement	Actions
Metrological Uncertainty	Meta Data Database, New Laboratory Calibration Procedures
Poor scalability of stations	-
Datalogger Code	
Redundant Code	Data Processing: one electrical to physical data code
Hard-Coded Parameters	Meta Data Database
Assembler-Style Programming	Data Processing: C program and automatic generation of datalogger code
Lack of Quality Control on Raw Data	Data Processing: QC on raw data
Raw Data not Recorded	Data Processing: raw data archived
Acquisition Servers Redundancy	New Acquisition servers
Datalogger End of Life	New Datalogger
Acquisition Servers End of Life	New Acquisition servers
Higher Data Retrieval Scheduling	Internet Connection
New Meteorological Measurement	-
Use of Data for Climatology	Meta Data Database
New Stations	-
Better Reliability of Data	Datalogger Code: stopping data acquisition during maintenance
Improved Documentation, Retrieval and Sharing of Information	Improved Knowledge Management, Meta Data Database
Automation and Simplification of Calibration, Maintenance, Instrument Updates	Datalogger Code Generation, Improved Knowledge Management, Meta Data Database
Improved Warnings of Failure	Device Status Monitoring

Table 2: Conformance table

To validate the new concept, a station has been built at Stabroek (see figure 1.1 in the North of Belgium, close to Antwerpen). It has been operating since August 2012 and the whole data processing chain was started in October 2013. A validation using data from nearby stations (including stations from the Netherlands) has been realised. In table 2, the requirements and issues of the actual network are compared to the new implementation of RMIB stations. It has been decided to postpone the building of new stations and the addition of new instruments to free enough man power to implement the other requirements. Before addressing the poor scalability of stations, we are waiting to gain enough experience to switch to digital communication between instruments and datalogger.

4 Future

We are expecting a full update of all stations within the next 4 years. As stations are renovated, routine calibration according to vendor requirements will be realised. The class evaluation of the sites according to WMO environmental classification is foreseen for the next year. The user interface to the metadata database will be finalised this year. Uncertainty of measurements will be estimated for the new stations, and a new maintained performance class assigned. Meteo Wing has also renovated their AWS and there is a project ongoing to ingest data from their stations in our data processing chain.

5 Conclusion

Renovation of the AWS network has been a big challenge and not all issues have been solved. We have developed a new concept for the Belgian AWS network and validated it on the new station at Stabroek, which is operating since august 2012.

References

[1] Enric Aguilar, Inge Auer, Manola Brunet, Peterson Thomas C., and Jon Wieringa. Guidelines on climate metadata and homogenization. TD 1186, WMO, 2003.

[2] Peter Beirinckx. New datalogger study. Technical Report RMIB-AWS-TN-006, Royal Meteorological Institute of Belgium, 2011.

[3] Belgocontrol. `http://www.belgocontrol.be/`.

[4] Bruxelles Environnnement Leefmilieu Brussel. `www.ibgebim.be/`.

[5] Nicolas De Coster and Luis González Sotelino. Estimation of uncertainty on Belgian AWS measurements: temperature, pressure, humidity. *WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation*, (WMO), 2012.

[6] Agence Wallonne de l'air et du climat. `http://airclimat.wallonie.be`.

[7] Knowledge Management for Belgian Public Organisation. `http://www.fedweb.belgium.be/fr/a_propos_de_l_organisation/gestion_des_connaissances/index.jsp`.

[8] Meteo France. Guide technique: classification de performance maintenue. Technical Report DSO/GT/OBS-ClassPerf, Meteo France, 2008.

[9] Gluster. `http://www.gluster.org/`.

[10] Luis González Sotelino. Automatic weather station work load. Technical Report RMIB-AWS-TN-004, Royal Meteorological Institute of Belgium, 2011.

[11] Brian W. Kernighan and P.J. Plauger. *The Elements of Programming Style*. McGraw-Hill, 1978.

[12] Vlaamse Milieumaatschappij. `http://www.vmm.be/`.

[13] Allway Sync. `http://allwaysync.com`.

[14] Meteo Wing. http://www.mil.be/fr/unites/meteo-wing.

[15] WMO. http://www.wmo.int/pages/prog/gcos/documents/GCOS_Climate_Monitoring_Principles.pdf.

[16] WMO. Guide to meteorological instruments and methods of observation. Technical Report 8, World Meteorological Organisation, 2010. 2008 edition, Updated in 2010.