AUTOMATED OPERATIONAL VALIDATION OF METEOROLOGICAL OBSERVATIONS IN THE NETHERLANDS

Wiel M.F. Wauben

Royal Netherlands Meteorological Institute (KNMI)

P.O. Box 201, 3730 AE De Bilt, The Netherlands

Tel. +31-30-2206 482, Fax +31-30-2210 407, E-mail: Wiel.Wauben@knmi.nl

ABSTRACT

The meteorological measurement network of the Netherlands is fully automated. The network provides near real-time meteorological data to both internal and external users. Hence a strong emphasis needs to be placed on the data quality of the automated observations. The KNMI measurement network employs several automated data validation steps. In this presentation an overview of the validation steps at different points in the measurement network will be given. The performance of these validation steps will be discussed. Furthermore, recent developments on the automation validation of data will be presented.

1. Introduction

Since the introduction of the new meteorological measurement network of KNMI (including 2 airbases of the Dutch Royal Navy) in November 2002 all synoptic and climatological reports are generated fully automatically[1]. In 2005 the 9 airbases of the Dutch Royal Air Force where upgraded and connected to the central system so that for these stations too the synoptic and climatological reports are generated fully automatically. Currently KNMI in collaboration with the Netherlands Oil and Gas Exploration and Production Association is implementing 8 platforms in the Dutch sector of the North Sea. KNMI and its partners employ observers only at airports and airbases, but only for the aeronautical reports. Since mid 2005 fully automated aeronautical reports are used operationally during closing hours of regional airports. Currently KNMI is working towards the full automation of the regional airports and airbases. The aeronautical reports for the North Sea platforms are generated fully automatically at the central location in De Bilt. All meteorological data and status information in the meteorological network are acquired centrally at least every 10 minutes and are distributed near real-time to internal and external users. In addition hourly meteorological reports are produced and disseminated automatically. At airports real-time 12-second sensor data, particularly runway visible range (RVR) and wind, is provided to users as well as half hourly routine and special aeronautical reports.

This paper reflects the preliminary work on a project that is starting up at KNMI. The aim of the project is to make an inventory of the currently used and available automated validation steps in order to move towards a uniform, optimized and improved automation of validation of meteorological observations. The focus here is on the automated meteorological measurement network in the Netherlands, but the project includes all meteorological measurement, including data received internationally. Currently validation checks of the data are performed at various places in the measurement network, but still it can happen that obviously faulty data is delivered to users. On the other hand known shortcomings are corrected before data enters the climatological database. External data sources are generally available on an hourly bases and are often encoded of entered manually. This data goes though coding and consistency checks before or while the data enters data bases, models or other applications. In this project the usage and location of the automated validation steps will be (re)considered as will the handling of the validation information.

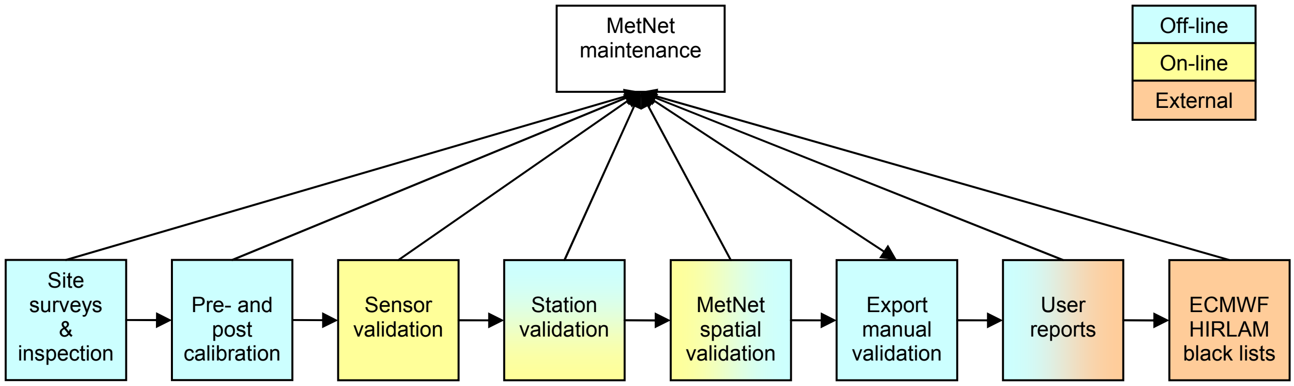

Figure 1. A schematic representation of the process chain of the quality control of MetNet data.

2. Quality control in the automated meteorological measurement network

The maintenance of the measurement network is distributed over different departments within KNMI. In Figure 1 an overview is given of the processes involved in the quality control of the measurement chain. The figure shows the steps involved in the internal data quality control and not the monitoring of the quality assurance of the process itself. The latter is monitored routinely via timeliness, availability and MTBF/MTTR by the instrumental department, which is ISO 9001/2000 certified. The upcoming reorganization within KNMI will put the maintenance staff involved in MetNet and other operational systems closer together. If not within the same department, then at least in the same sector of KNMI. Furthermore a person will be appointed that is responsible for the overall system so that issues at the interfaces and over more departments can be handled more effectively. The embedding of the maintenance of the observation network within an organization with clear arrangements, procedures, monitoring, documentation and evaluation is crucial. The guidelines provided by e.g. KNMI and WMO give useful information on (the organization) of the validation and the quality aspect[2-4].

In the rest of this paper the focus will be more on the physical and technical aspects involved. The individual components involved in the current quality control scheme at KNMI are:

- Site surveys and inspection which are performed manually in order to insure that the conditions under which the observations are performed are up to agreed standards. This includes the selection of the site for performing the observations and a regular (half yearly) inspection of the site and its surroundings. This also includes a visual inspection of the sensors. Any discrepancies are documented and stored in the station archive, and solved when possible.
- Sensors, after extensive evaluations and tests prior to selection for usage in the network, are maintained and calibrated at KNMI following fixed procedures. Each sensor is calibrated before usage, during which it should be within certain limits. If necessary a sensor is adjusted before calibration, since KNMI uses a fixed calibration constant for each type of instrument so that no calibration factor needs to be taken into account and each sensor can easily be replaced by another. Before the calibration period is passed the sensor is exchanged and subjected to a post calibration. Is the sensor is outside specific limits, which are broader than the original calibration limits, the users are notified. The limits and calibration periods are based on user accuracy requirements and field experiences. Depending on the results of the post calibrations and the number of sensors malfunctions the calibration period can be reconsidered.
- Each sensor is connected to a KNMI sensor interface (SIAM) that transforms the sensor output into a fixed format with meteorological units and status information. The sensor interface, and sometimes also the sensor - e.g. smart sensors like cloud base ceilometers and present weather sensors that internally check the correct operation of components and the validity of the results, perform the first quality check. This check is restricted to the sensor itself, although the sensor interface cross checks ambient and dew point temperature, or wind speed and direction. These checks include the well known range checks, consistency checks (e.g. dew

point ≤ ambient temperature, minimum ≤ average ≤ maximum temperature etc.) and temporal checks (e.g. sudden jumps between successive measurements or fixation of a value). The sensor interface uses fixed warning and error limits for these tests. In case of an error no value is reported.

- KNMI operates automatic weather stations and airport systems at a station level. An automatic weather station only stores the data temporarily before it is transferred to the central site for further processing. At airports the sensor data is processed and forwarded to local users. Furthermore an observer/forecaster is generally present that can monitor sensor and derived data, enter aeronautical reports and discard or overrule sensor values. The automated validation at a station level presently only included corrections for light precipitation (rain gauge versus present weather sensor), correction of precipitation type (present weather sensor precipitation type and visibility, temperature and relative humidity), backup of faulty sensors and coding checks.
- At the next level all data is available in the central database. Here the same checks as on an airport system are available and when appropriate they are applied. It should however be noted that generally the sensors are not redundant. Hence backup sensors are only applicable at airports and some platforms. Spatial information is available in the central database, but currently no spatial validation is performed. Data can be visualized in the central system, but its main purpose is to maintain and monitor sensors, stations and connections; and acquire, process and make the observations available. Alarms are generated in case of malfunctions. Daily status reports are generated or are made upon request.
- Data is exported from the central site database to other applications for further usage and validation. On a daily basis data is validated whereby interrelations between parameters and spatial information is used. The validation is mainly performed on hourly data. Consistency checks like averaged wind speed ≤ wind gust are performed automatically, but interrelation and spatial check are generally performed manually.
- Several applications are available to visualize 10-minute MetNet data. Due to the easy access of the data, the feedback of users increased after realization of the measurement network. MetNet data is also used by models and other applications that perform a validity check and give feedback on the quality.

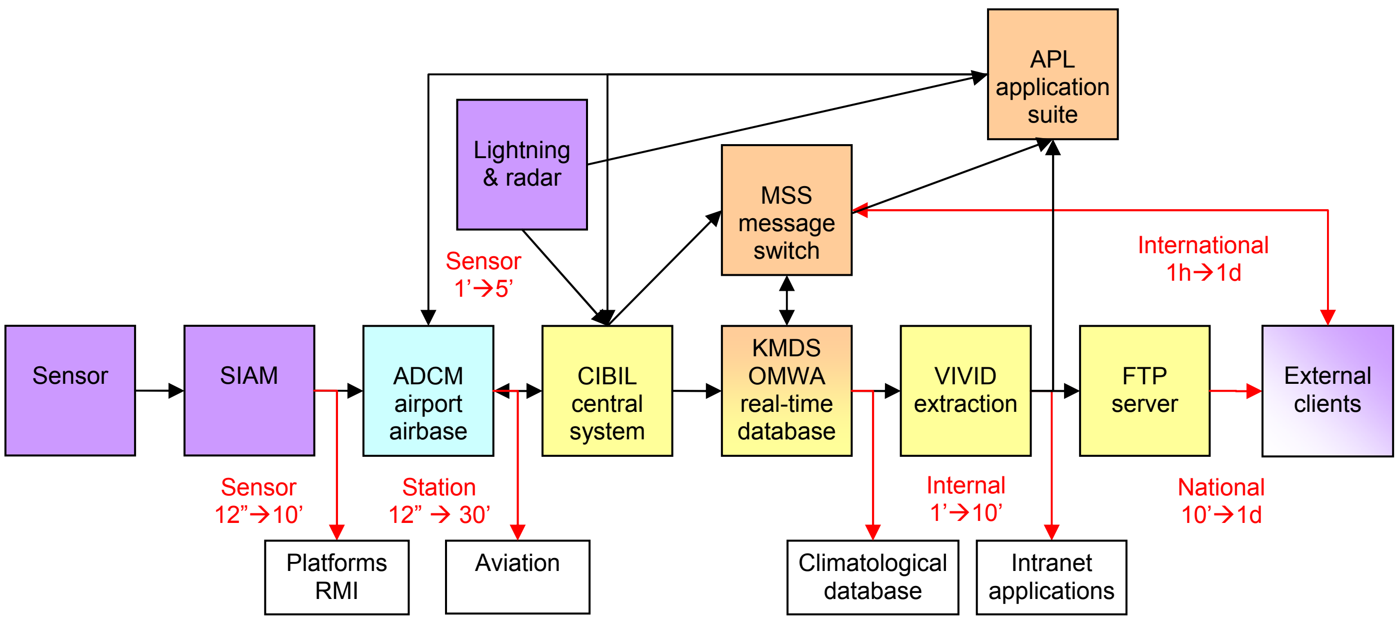

Figure 2. Schematic overview of the data flow between sensors/systems of MetNet.

3. Data flow in the automated meteorological measurement network

The data flow of the MetNet system is schematically shown in Figure 2. The location of data clients is also indicated. Although all users agree the validation of data is important, it may not have any significant impact on the timeliness of the data. This puts constraints on the extend of the validation available at a specific location. Since the KNMI sensors and sensor interfaces are also used on platforms in the North Sea and within sister organizations in the ministry of transport and water works responsible for the coastal and inland waters, the most basic user level is that of the output of the sensor interfaces. The validation at that point is real-time since data is produced every 12 seconds over a time interval ranging from sample values to 10-minute averages for 'individual' sensors. It seems that the current range, consistency and temporal check are optimal at that user level.

The next phase at which users receive data is at the station level. This particularly applies to aeronautical usage where 12-second (processed) sensor data is made available together with meteorological reports. At this level the interrelations between meteorological parameters can be applied and sometimes limited spatial information is available. A first quick scan showed that several relations are used in the process of the manual validation. The interrelations are, e.g. between rain gauge precipitation versus present weather, precipitation radar, humidity and cloud cover; or cloud cover versus global radiation and radar. However, generally these interrelations cannot guarantee the validity of instantaneous data because differences between various parameters can occur instantaneously and only deviations over longer time periods suggest a possible problem. The situation regarding backup sensors is also not clear at present, since often the backup sensor is not closely co-located with the operational sensor since it serves another purpose as well, e.g. at airports the wind sensor set along a runway might serve as the backup of the wind along another runway if the local conditions permit this. The validity of a backup for wind is determined by the gust factor (which is related to the local surface roughness) and is determined by values per wind sector averaged over a long time period. The behavior of instantaneous differences needs to be studied in more detail.

The users generally get the data from the central database at 10-minute intervals. The maximum acceptable delay is about 6 minutes, which includes the acquisition, processing and dissemination of all data. At this level spatial information over the Netherlands is available. The current off-line spatial validation is performed manually. However, it is expected that for pressure and wind a simple model can be used for validation. For the other parameters further investigations are required in order to find out whether a simple model (possible assisted with statistical information) can be used as an automated validation tool, or a reduced real-time meso-scale model. The situation on the international level is generally the same as on the national level although the frequency of the received data (generally hourly), density and the uniformity of the observations differ. The external observations are presently subjected to range and consistency tests before usage.

4. Basic assumptions regarding validation

The improved automated operational validation at KNMI kicked of recently with a brainstorm session involving both experts and users. The basic assumption agreed upon include:

- Automated operational validation is considered useful. It will improve the real-time data quality 24*7 and will reduce manual labor.
- Validation should not lead to a delay in the real-time data flow.
- The result of validation should be indicated by a quality number. The results for each test can be recorded separately in a binary format. A quality relevant for each user group resulting in e.g. an easy distinction between e.g. good, suspect, bad data, can be generated by applying a binary filter to the quality number.
- The validation process will not change or discard values. If a correction is applied this should be done on another variable so that the raw data is still available.
- The user should always use value AND quality of the data. As a first step an export filter masking suspect and/or bad data can be used.

- The results of a validation step cannot be expressed in a (reduced) uncertainty range since validation is often related to malfunctions and/or specific situations.
- A better way to monitor the validation steps is to look at statistics for each validation step (applied on-line as well as off-line). This gives not only information on the usefulness of the validation steps, but it also points to specific problems regarding observations that should if possible be traced and solved.
- Validation should be embedded in the maintenance process. The results of the validation should be monitored and evaluated and adequate actions should be taken to solve if possible the cause of the discrepancies.
- The validation steps should be consistent and modular so that they can be applied at various locations.
- The automated validation process should allow manual input of e.g. environmental and technical disturbances that serves as a warning for reduced quality and as a link to a detailed description in station/sensor archives.
- The usage of quality indicators should be accommodated and treated consistently/traceable between the various systems/formats, e.g. between the KNMI sensor interface and the disseminated BUFR report.
- The usage/meaning of quality indicators should be transparent and suitable for users, so that e.g. the climatological data base can make use of the results of the automated validation steps.
- It should be possible to generate alarms for failures of specific validation steps.

It should be noted that the points mentioned above sometimes might conflict with KNMI or WMO practices/recommendation. The purpose was first to agree on general assumptions, next find out what validation methods are available, and finally to decide which validation steps should be implemented where in the KNMI infrastructure and how. At that stage the usefulness will be considered as well as the costs involved, i.e. success rate and importance of a validation step will be weighed against implementation costs and saving in operational costs.

5. Outlook

The common validation steps regarding range, consistency and temporal checks are performed at various places within KNMI. The criteria of these tests need to be clarified and homogenized if possible. For the correct validation of international data the local conditions/ experiences should be taken into account. Therefore detailed information on the criteria used in their checks is important. The next step forward in automated real-time validation, specifically usage of interrelations and spatial information, is not straightforward. Before these validation steps can be performed automatically, studies need to be performed into the available validation checks and the optimal thresholds for the situation in the Netherlands. It is expected that KNMI will investigated this in the near future. Since the work is just starting up it is a good opportunity to take assumptions, methods and experiences of other meteorological institutes into account. Therefore any information and suggestions regarding (automated real-time) validation is much appreciated.

At the TECO 2006 I expect to present further details on the validation steps used and considered by KNMI together with an indication of their impact.

6. References

[1] Wauben, W. M. F., Hart, L.M.: The new meteorological observation network in the Netherlands; Status and operational experience, WMO-TECO, Bucharest, Rumania, 2005.

[2] Handboek waarnemingen (Handbook on observations), KNMI, De Bilt, The Netherlands, 2005

[3] Guide to Meteorological Instruments and Methods of Observation, Sixth edition, WMO No. 8, Geneva, Switzerland, 1996.

[4] Guide on Global Data Processing Systems, WMO No. 305, Geneva, Switzerland, 1993.