

Integrated National In-Situ Observational Network

Dr Igor Zahumensky

Slovak Hydrometeorological Institute
Jeseniova 17, 833 15 Bratislava, Slovakia
Tel./Fax.: +421 541 36 24, Igor.Zahumensky@shmu.sk

Abstract

In order to maximize the potential value and benefit of observations from different observation networks, there is a need to integrate the planning, operation, data management, system monitoring, and life cycle support {operation, maintenance, repair} to existing and future observations. It is a challenge how to improve efficiency and effectiveness in order to optimize the value of the investments in the observing systems for existing and future observations and observing systems.

Defining and validating requirements is the starting point for an integration activity. It is not always suitable, feasible or acceptable to integrate all aspects of observing systems (existing or future ones). The quality of an observing system can be assessed by comparison of users' requirements and the ability of the system to fulfil them.

The goal is to minimize development and operating costs and maximize the volume, quality, and utility of observations.

An Integrated In-Situ observational network includes also an associated Integrated Data and Monitoring System.

Introduction

In National Meteorological Services (NMSs), many of the present users' needs have been usually met by different observational networks, belonging to various departments or divisions, set up in the past. Those have not been developed in an integrated manner and have been funded and operated separately to meet their own purposes and goals. Some of those networks are in danger of disappearing, even though their needs remain, although other useful elements could remain, but there is no clear course of operational running and funding.

A key issue for the national observing system is the integration of separate observational networks and data sets to maximize their utility for different users and purposes. Profit of the investment in individual observational networks, which is substantial, can be increased by more cost-efficient integration of data sets, by shared exploitation of observational sites, by an integrated data and network management, and by an information delivery system.

The integration is an end-to-end process of examination, planning and implementation. It includes: the variables to be measured and/or phenomena observed; the site and sensors installed for the measurements; associated data and metadata management, such as collection, processing, control, archiving and monitoring; network management; monitoring of performance and functionality of all system components; data, metadata and products dissemination; and other elements and processes.

Defining and validating of users' requirements for the data is the starting point for the integration. It is not always suitable, feasible or acceptable to integrate all aspects of existing or planned observing systems. The aim is to minimize the development and operating costs and to maximize the quality, value, utility and availability of observations, providing easy and timely access and long term archive of observational data and related metadata.

The Reasons for and the Goals of the Integration

Generally, the national surface observational network managed by the National Meteorological or Hydrometeorological Service consists of several separated networks of from tens to hundreds of observing stations each of them with several sensors installed. Each separate network/system has been managed separately from the others and this has been the main shortcoming of this management style.

To overcome this and save the financial sources as well as maximize the potential value and benefit of observations, it is required to think through a need to integrate the planning, operation, data management, system monitoring, and life support cycle to existing and future observations.

Therefore, it is necessary for the NMS to prepare its "Strategy and Policy" including:

1. Assessment of existing observing systems (an analysis of their extent and characteristics, capabilities to meet present and planned requirements and defining gaps and shortcomings of the systems);
2. Integration existing observing systems assets into an integrated network through integrated data and network management and communications;
3. Development of end products based on users' requirements;
4. Improvement of infrastructure and technology of the new observing system;
5. Modernization of data and information systems;
6. Periodic reviews and assessment of the integrated network, and modification as necessary.

Three dominant goals of the desired integrated observational network will be:

1. To sustain existing observations;
2. To integrate existing and new observations; and
3. To adapt the system to meet short-term and long-term users' requirements, to make sustained observations more complete, more effective, and more affordable in real time and, of course, in higher quality.

Features of an Integrated Observational Network

There are several requirements that should be met by the observing program of the NMS:

- The observing program has to keep pace with growing demands and needs of users and with current development of technology.
- Data processing has to be done in real time and should be automated as much as possible.
- Quality Assurance has to be implemented in all stages of the data acquisition process.
- Data Quality Control has to be efficient; data quality shall be known at any stage of data processing.

In combination with those, the integrated network (as an upgraded multilevel network) for monitoring the weather, water, climate variability, and other environmental components should (*or may be shall*):

- Maintain a network size and station density that satisfies all major needs:
 - the station spacing and interval between observations should correspond with the desired spatial and temporal resolution of the variables to be measured or observed;
 - the total number of stations should, for reasons of economy, be as small as possible but large enough to meet the various users' requirements;
- Be expandable and adaptable to meet future users' needs and requirements on observations;
- Provide coordinated data collection to minimize duplication, reduce costs and maximize data availability;
- Ensure high-resolution and accurate real-time data for all required applications (including operational forecasts, warnings, and their verification);
- Provide easy and rapid access to data and information, and associated metadata;
- Ensure adequate uniformity and standardization in the observing practices, data processing, and QC procedures;
- Ensure high and known quality of the data;
- Provide opportunities to develop many different applications and value-added products; and, of course
- Meet the established WMO standards for siting and exposure, sensors/network performance and maintenance, data availability, data quality, and required metadata.

Regardless the integration and automation, a human observer will continue to be a valued component of the integrated network, though his function will change with time. Until the subjective and visual observation or manual measurement of some variables can be automated, human observers are indispensable.

The integrated network and its observing systems have to be implemented in a way that allows flexibility as both requirements and technology have been evolving permanently. The formulation of data requirements is an evolving process based on experiences with the observing systems and improvements in data assimilation techniques. The process balances demands of users with the technical feasibility of data accuracy and resolution. To ensure that, an effective integrated observational network has to:

- Include ongoing strategic planning to set evolving requirements for the system based on evolving users' needs and requirements;
- Evaluate technological development for potential improvement and recommend required technical developments;
- Provide data and information that are standards based and interoperable;
- Address integrated data archival and management needs of users.

And finally, an integrated surface observation network would maximize the utility and cost effectiveness of the national investment in surface observations by:

- Creating a modern network around which all surface monitoring networks are integrated to provide the highest possible quality in real-time weather, water, and climate data and information;
- Establishing a high-quality infrastructure that integrates surface observations from a wide range of sites, quality assures those data, and makes it available within minutes of the observation time.

Planning of the Integration

There are several aspects that have to be taken into account during the phase of planning a new observing station or network. During this phase, the management of the NMS responsible of the development should answer several questions that can be as follows:

- What is the representativeness of meteorological observations of an area in accordance with application for which they are used?
- What are the general requirements of a station/network for siting and exposure, for inspection and maintenance, for monitoring of system performance, availability and quality of data?
- What are the standards and definitions for measurements?
- Which system to choose for the required observation?
- What are the required accuracy and achievable uncertainties?
- What are requirements for instrument?
- How to carry out meteorological observations?
- How to establish effective liaison procedures between those responsible for monitoring and for maintenance and calibration to facilitate rapid response to reports from the monitoring system and remedial actions if needed?
- What kind of metadata related to meteorological observations is needed?
- How to optimize cost/benefit and cost/performance ratios?

All disciplines involved with meteorological observations (synoptic meteorology, climatology, aeronautical meteorology, agricultural meteorology, and hydrology) formulate their own functional requirements on observations to satisfy individual functional requirements for observations to satisfying specific service needs. All disciplines, however, state that it is beneficial to apply universal rules or standard methods of observation to avoid unnecessary confusion and to achieve data compatibility. In line with this policy, standardization and integration will be beneficial if network design fulfils the requirements of various disciplines and applications.

A plan for the sustained integrated observational network must have the following characteristics:

- Coordinated data collection efforts to minimize duplication, reduce costs, and maximize data availability;
- Development of an integrated data management to ensure continuous data streams, timely delivery of data, and adequate data quality control;
- Ability to meet the requirements of multiple users by integrating observations collected for different purposes.

The process of integration is very complex and begins with answering of these key questions:

- What are the users' needs?
- What level of integration is desired and cost-effective (Which observing systems functions are related to the identified societal benefits)?
- What are existing resources and future possibilities of them?
- Which observing subsystem functions to integrate?
- Which observations need to be integrated?
- What tools and methods will be used to accomplish the integration of the observing subsystem functions?
- What plans need to be developed and implemented?

Based on the previous analysis, some important steps that are needed to be undertaken are as follows:

- To formulate an initial design of the integrated observational network based on identified users' needs and on the existing monitoring networks. This means to translate users' requirements into functional specifications and then to technical specifications for the network, which are very complex tasks requiring a sound knowledge of measuring technology, instruments and methods of observation, and WMO regulations and standards;
- To establish and utilize criteria (e.g., cost/benefit ratios, operational constraints, etc.) for prioritizing the observing subsystems. These will provide funding sources mechanism for evaluating the present networks and the intended future integrated system. Observations that are critical for the decision-making require specific prioritization;
- To identify deficiencies and shortcomings of existing observing components based on users' needs;
- To analyse ongoing data collection systems and identify how well they meet the users' needs;
- To propose augmentations to the existing subsystems and a new network with estimated benefits, costs, and schedules;
- To estimate new users' needs and adjust observing system elements to meet changing requirements, technologies, and available resources;
- To evaluate present data management activities and capabilities, and the way how to improve them. It is most likely that the greatest benefits from the integrated observational network will be realized through the development of an effective data and network management methodology; and
- To improve and expand the management and communication facilities necessary for routine monitoring, analysis, warning and prediction on required time scales.

As because there are several gaps but limited resources, comprehensive analysis and planning are needed to determine the location and attributes of existing and planned observing systems, and where to fill gaps primarily for the greatest benefit. In the interim, maintaining the existing systems and integrating data available will provide users with rapid and easy data access. Developing new products that combine available data sets is also high priority action to consider.

During the process of planning, observing system priorities for operations should be kept in mind. It is assumed that the most important are following: reliability, efficiency, cost effectiveness, and continuity.

The development of the integrated network should be based on the following key architecture principles:

- To support a broad range of implementation options driven by users' needs and incorporate new technology and methods;
- To address operational and planned observing systems required by users to make products, forecasts, warnings and related decisions.

Technical Systems Integration

Technical systems integration is an inherent part of network integration. It addresses the coordination of observing system technology and data management systems that enable operational and research applications. Across many of the societal areas, there are common

challenges that would benefit from an investment in integrated solutions. The most important and noticeable are as follows:

- Information technology integration: This integration includes control of the data flow and processing as well as the integration necessary for distribution of data, related metadata and products to a variety of users' applications.
- Observing site integration: This integration could include the physical integration of observing systems at the same ground site, such as sharing space and observing masts or towers on the ground for various observations; and maintaining and upgrading sub-networks in an orderly way.
- Multifunctional integration: This integration means using observations for all possible applications not only for the main purpose the observing site was built up.
- Standards and procedures integration: This integration means implementation and the use of unified standards and procedures in all processes of the system, including metadata. These standards will address data and metadata access, discovery, interoperability, processing, dissemination and archiving.

Integration considerations must also address observing system evolution. Observing system upgrades, new systems, system replacements, etc., they all need to be considered in the context of a structured system integration approach. The significance and role assigned to models and other decision supportive tools in the integration process must also be considered.

Network Management Unit

To reach effectiveness and efficiency of the integrated observational network, it is necessary to specify precisely the separate processes of the network quality management system, and criteria for the control of its quality including the procedures for monitoring. Special attention shall be given to guidance dealing with network management and all its components.

The operational tasks of the Network Management Unit (NMU) are based on the activities and performance of individual stations. Amongst the functions performed by the NMU are:

- To maintain functions of the network;
- To exercise functional control and inspection of the network;
- To monitor the performance of the network, recommend and implement improvements;
- To monitor and review the efficiency and effectiveness of the network;
- To develop and define performance and calibration standards, procedures, and functional requirements for observations, instruments and equipment, and relevant issue instructions;
- To produce and maintain observation specifications which detail installation specifications for network observations;
- To advise on technical training for all those involved in the network;
- To provide consumables for the network measurements;
- To provide effective liaison between users of meteorological observations and suppliers of data and equipment;
- To formulate plans and policy for the development, maintenance and operation of the network;
- To advise on long term re-equipment plans.

Management and Information System

Multifunctional Management and Information System (MIS) as a supervisor system for data and network management, for dissemination of data, metadata, products and other information is an integral part of an integrated observational network that:

- Provides integrated, quality-controlled real-time datasets available to hydrometeorological community as well as other agencies upon agreement;
- Performs the comprehensive observational data and alert/warning collection and distribution of operational data, information and service messages;
- Provides system responses in case of detection of alert/alarm condition(s);
- Provides monitoring of:
 - the availability of observational data and monitoring statistics on the frequency and magnitude of observational errors encountered at each station location;
 - data remedial actions;

- data from point of view of alert/alarm threshold/rate of change, specified by the users;
- alerts/alarms received from the integrated network and corresponding responses of the system;
- Provides sufficient control and monitoring of performance and functionality of all system components;
- Provides remote control and diagnostic of network automatic stations;
- Provides effective real-time interaction with the NMU, e.g.;
 - Issues alerts to warn the NMU of poor data quality, malfunctions or failures.
 - Issues automatic “trouble tickets” to the NMU so that problems can be fixed and sensors replaced or repaired if needed as soon as possible;
- Compiles information from laboratory calibration and testing, on-site testing, station inspection and maintenance, automated QC of data, and network and data monitoring, each of them producing valuable information on the network’s performance;
- Requires real-time access to the observations as well as related metadata;

The key features of MIS are (at least) as follows:

- It has to provide observations and information that are essential for decision-making;
- It has to be fully automatic and operational 24 hours/day, 7 days/week.
- It has to be a flexible and modular system; modular architecture enables that most modules can be updated transparently or new procedures, modules, units etc. can be added without interrupting the performance of the system and service to users, without any problems and necessity to rebuild the system;
- It has to be reliable and designed to operate during the most severe weather; the greatest benefit from the effective system occurs when weather is extreme, widespread, and/or sudden.
- There has to be the two-way communication between the separate units of the integrated network and the system as itself, as decisive importance is their cooperation, i.e. there has to be the feedback between the corresponding units.

Quality Assurance

The Quality Assurance system and its standardisation is becoming more important with the introduction of an integrated observational network. The provision of good quality hydrometeorological data is not a simple matter, and it is impossible without quality assurance implemented by a quality management (QM) system. The effective QM system operates continuously at all points of the whole observing system, from network planning and training, through installation and station operations to data transmission and archiving, with continuous interaction and feedback among all parts involved in the observing system.

There is a strong need to develop and implement the comprehensive guidelines for a sophisticated quality management. The QA system represents a compilation of the processes that range from simple testing to sophisticated control and monitoring procedures. It is a chain that starts far before the installation of the sensors and ends providing data of high and known quality accompanied by corresponding metadata. To support the significance of QA the following statement, as an example, can be used: There are no methods that can correct data from poorly installed, exposed or maintained sensors.

Data from different observing sub-systems/sub-networks are typically managed by individual systems differently. The integrated system applies the same QA procedures for all parts of the new network; it produces a more uniform, comprehensive data set suitable for a wider range of applications and improves the confidence level of the QC processed data combined with associated metadata.

Quality of Data

It is better to prevent errors than to cure them later and it is by far the cheaper option. Making corrections retrospectively can also mean that the incorrect data may have already been used in a number of analyses before being corrected. Therefore, data validation and cleaning is an important part of the data quality process which is one of the most important elements of QA system implemented inside the observing system.

NMS as a data producer should ensure that:

- Data and the data quality are adequately and accurately documented;
- Data quality is maintained at the highest possible level at all times;
- Data are available in a timely and accurate manner with documentation (metadata) that allows users to determine “fitness for use”;
- QC procedures are implemented and exercised during data acquisition;
- Validation checks are routinely carried out on all observational data;
- Validation checks carried out are fully documented;
- All known errors are fully documented and made known to users;
- Feedback from users on the data quality is dealt with in a timely manner.

It is of the utmost importance that NMS should make adequate provision for QC of data to ensure that they are as free from error as possible and the quality of data is known in every level of the data obtaining process.

The QC System implemented by NMS should include:

- Data Validation,
- Data Cleaning,
- QC Monitoring.

Data Validation is a process used to determine how accurate data are, complete, consistent or reasonable. The process may include completeness checks, plausible value checks, time consistency checks, and internal consistency checks. These processes usually result in flagging, documenting and subsequent checking and correction of suspect records. Validation checks may also involve checking for compliance against applicable standards, rules, and conventions. A key stage in data validation is to identify the causes of the errors detected and to focus on preventing those errors from re-occurring. The quality of data should be known at any point of the validation process and can change through time as more information becomes available. QC flags should be used for that purpose.

Data Cleaning (remedial actions) refers to the process of “fixing” errors in the data sets that have been identified during the validation process. It is reasonable to retain both the original data and the corrected data side by side in the database so that if mistakes are made in the cleaning process, the original information can be recovered. The general framework for data cleaning is:

- To define and determine error types (missing, inconsistent, doubtful, erroneous);
- To search and identify error instances;
- To activate appropriate remedial procedures;
- To correct the errors, fill the gaps;
- To document error instances and error types;
- To modify data entry procedures to reduce errors in future.

The system should include procedures for returning to the source of data (original data) to verify them and to prevent recurrence of the errors. It also should include immediate feedback to the origin of the erroneous data providing a record for use by the NMU.

Some assessments and corrections can be and should be performed in real time, whereas other evaluations can only be accomplished after gathering of sufficient data over a longer period. No appreciable delay should be caused by these procedures because the data must be transmitted in real time for operational use. On the other side, the real-time QC at the observing point is, however, of paramount importance since many of the errors introduced during the observation process cannot be eliminated later.

QC Monitoring at the network level should be implemented and performed by MIS as real time quality control procedures have their limitations and some errors can go undetected, such as sensor drift or bias, as well as errors in data transmission. All possibilities for automatic monitoring of errors should be used to recognise errors in advance before they could affect the processed values.

Effective real time QC monitoring as an integral part of a QC system has to include checks of the following items:

- Completeness of observations at the observing station;
- Quality of data transmitted from the station;
- Completeness and timeliness of the data collection from the whole integrated network

at the MIS.

QC monitoring requires the preparation of summaries and various statistics on observational errors of individual meteorological variables and for individual observing point, through a series of flags indicating the results of each check, and generates hourly, daily, weekly, monthly and yearly summaries of, e.g.:

- The total number of observations scheduled and available for each variable (completeness of data);
- The total number of observations which failed the QC checks for each variable (quality of data) from the reason of:
 - Plausible inconsistency,
 - Time inconsistency,
 - Internal inconsistency;
- The percentage of failed observations (quality of data);
- The error and threshold values for each failed observation (reason of failure)
- The frequency and magnitude of observation errors encountered at each station.

Stations with large percentages of failed observations are probably experiencing hardware or software failures or inappropriate maintenance. These should be referred to the NMO as the statistics provide information for the purpose of:

- Monitoring quality of station performance,
- Locating persistent biases or failures in observations,
- Evaluating improvement of quality of observation data, performance and maintenance of station/network.

The Role of Metadata

Metadata are critical to the success of any integrated data management system.

The NMS is responsible for providing adequate metadata on all parts of the observing system. Metadata describing the circumstances of the measurement/observation are particularly important for users of data. They are absolutely essential to ensure that the final data user has no doubt about the conditions in which data have been recorded, gathered, transmitted and processed.

Metadata help organize, maintain, and insure the investment in data. They should describe data from all aspects of the whole process but mainly provide information on:

- Identification (name of the data and data set; responsibility for the data; geographic coverage, etc.);
- Description from where and how the observations were collected, procedures related to the observation and data collection (stations, instruments);
- Data Quality (how good data are including any QC processing used to identify any changes to original observational data, as well as any retrospective reprocessing);
- Data access (location, access, restrictions).

Metadata coupled with observations significantly increase the value and the use of data and information. Observations without relevant metadata are subject to higher levels of uncertainty and risk, and lower probability of use. The standards, such as ISO 19115, address aspects of metadata documentation. Each observational network has to specify the particulars of the metadata required.

The development of a metadata system requires considerable interdisciplinary organization and its operation requires constant attention. Standard procedures has to be established for collecting all significant changes made in instrumentation, observing and processing practices, sensor exposure and siting, sensor failures, etc.

Conclusion

There is no doubt about the profit that the integration of separate networks into an integrated system will bring the NMS. The integration of all components and processes is a challenging task how to improve efficiency and effectiveness and optimize the value of the investments in the observational networks.

The benefit of the integrated observing system will be:

- Improved information (data and end products) for the hydrological, meteorological, climatological and other environmental monitoring;
- Improved data management system (data quality monitoring, performance monitoring necessary for network management);
- Improved observing system performance with uniform monitoring tools and evaluation standards;
- Improved data accessibility by all users;
- Improved metadata management;
- More cost-effective observing system;

It is likely that the greatest benefit from an integrated observational network will be realized through the development of an effective data and information management methodology, i.e. the development of effective approaches for the acquisition, processing, dissemination and utilization of above mentioned information.

The success of the integrated network will depend on several design elements:

- Provision of the right set of measured variables to meet users' requirements;
- Implementation of new technologies to improve all aspects of the system, especially the timeliness and accuracy of providing information, and
- Usefulness of its end user tailored products.

Reference

- (1) Guide on the GOS, WMO-No. 488
- (2) Guide to Meteorological Instruments and Method of Observation, WMO-No. 8
- (3) Strategic Plan for the U.S. Integrated Earth Observation System,
http://usgeo.gov/docs/EOCStrategic_Plan.pdf
- (4) Integrating Surface Observing Systems. (ISOS, Implementation Plan, A NOAA First Step, Toward an Integrated National Surface Observing Network,
www.ncdc.noaa.gov/oa/isosmtg/isos-ex-plan.pdf
- (5) Toward a U.S. Plan for an Integrated, Sustained Ocean Observing System,
<http://www.nopp.org/resources/nopp/towardintegrated.pdf>
- (6) COOP Modernization: Building the National Cooperative Mesonet,
http://ams.confex.com/ams/84Annual/techprogram/paper_72214.htm