Technical Memo

European Centre for Medium-Range Weather Forecasts

850

WIGOS Data Quality Monitoring System at ECMWF

Cristina Prates, Erik Andersson,
Thomas Haiden (Research Department)

July 2019

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our website under:
http://www.ecmwf.int/en/research/publications

Contact: library@ecmwf.int

Abstract

This document summarises ECMWF's participation in the World Meteorological Organisation (WMO) effort to modernize the monitoring of the global observing system (GOS) focusing on the benefits that our engagement can bring to the monitoring of in situ observations at ECMWF. WMO is developing a modern system for the quality management of the surface-based component of the WMO Integrated Global Observing System (WIGOS). ECMWF, the German Weather Service (DWD), the US National Centers for Environmental Prediction (NCEP) and the Japan Meteorological Agency (JMA) have contributed to the pilot development phase of the project. The proposed system, WDQMS itself has the potential to bring far-reaching benefits in terms of improvement of WMO management of in-situ components of WIGOS concerning the quality of the observations and of the associated station metadata (available in the OSCAR/Surface database). The NWP community benefits not only from a higher-quality network, but also from near-real-time access to comparable quality monitoring data from several global NWP centres. This has already proved to be beneficial for ECMWF, helping to detect (and subsequently resolve) differences in data reception and differences in data usage (e.g. differences in quality control, station height and other metadata) as shown in this document.

1 Introduction

The World Meteorological Organisation (WMO) has launched an initiative to modernise the monitoring of the surface-based component of WMO Integrated Global Observational System (WIGOS). Hitherto, WMO monitoring of conventional observations has been based on monthly reports produced by Lead Centres following the recommendations in Attachment II.9 of WMO Manual GDPFS (WMO, Manual on the GDPFS 2010). ECMWF has been the WMO Lead Centre for upper-air observations since 1988[1] and is still producing a Global Data Monitoring report[2] monthly in which information on availability of land surface observations is also included. The goal is to move towards a near-real-time (e.g. daily) monitoring of the status of the Global Observation System (GOS) in terms of availability and data quality, which would help WMO to take actions, namely reporting back to data providers to have the problem fixed in a timely manner. This activity under the umbrella of WIGOS is key for monitoring the actual performance of the observational capabilities recorded in the surface-based component of the Observation Systems Capability Analysis and Review Tool (OSCAR), OSCAR/Surface, database. This database, the global repository of WIGOS metadata for all surface-based observations, will be used in the monitoring as a source of observational metadata.

[1] The 9th session of the WMO CBS (Geneva 1988) recommended that lead centres should be appointed for monitoring the quality of each main type of observation. They should liaise with participating centres and coordinate all the results, inform the WMO Secretariat immediately of obvious problems, and produce every six months a consolidated list of observations believed to be of low quality. ECMWF was subsequently nominated as the lead centre for radiosonde and pilot observations.

[2] available on https://www.ecmwf.int/en/forecasts/quality-our-forecasts/monitoring-observing-system/ecmwf-global-data-monitoring-report-archive

The two WIGOS workshops on quality monitoring and incident management, held in December 2014 and December 2015, reviewed the monitoring of the in situ component of GOS and developed the plans for the WIGOS Data Quality Monitoring System (WDQMS) which was designed to provide the near-real time monitoring and identification of the observational data quality issues and, if needed, follow-up actions on a station by station basis. It consists of three main functions: the WIGOS Quality Monitoring (QM) Function; the WIGOS Evaluation (Ev) Function and the WIGOS Incident Management (IM) Function. The QM Function will receive quality monitoring information daily from Global NWP centres. This information - a by-product of the centres' assimilation systems - is provided for each land station and must be complemented with the associated station metadata extracted from OSCAR/Surface. It is the role of the QM function "to generate reports of the results of comparisons of the received data with the expected availability, timeliness and observational quality criteria" (Top Level Description of WDQMS). The Ev Function will take the outputs of the QM Function and analyse all the observational issues highlighted in the QM reports and determine if they justify being formally raised as incidents with the observational data providers, taking into account the expectation of typical performance and other contextual information (e.g. geo-political, environmental). Any issue considered as incident by the Ev Function will be undertaken by the IM Function, which will request the data provider to investigate and to resolve the incident within a reasonable time. It is the role of the IM function to record, communicate and follow-up on the incidents with the data suppliers as well as data users to ensure they take suitable precaution with the given source. The Regional WIGOS Centres (RWCs) are the regional component of WIGOS responsible for implementing the Ev and IM functions as well as supporting their members in updating, maintaining and quality controlling WIGOS station metadata in the OSCAR/Surface database.

A Task Team on the WIGOS Data Quality Monitoring System (TT-WDQMS) was created in May 2016 to further develop and extend the concept of the WDQMS and oversee the implementation of the pilot WDQMS. In particular, a pilot project on quality monitoring was established to implement the QM Component of WDQMS which relies on quality monitoring information provided by global NWP centres. Additionally, a prototype of a webtool for displaying the monitoring outputs from NWP centres was developed at Secretariat/WIGOS PO, which includes interactive graphic displays (maps and time series). ECMWF has taken an active role in these pilot studies and has been asked to take on the responsibility of developing and running the future operational webtool on behalf of WMO.

The proposed system, WDQMS itself has the potential to bring far-reaching benefits in terms of improvement of WMO management of in-situ component of WIGOS concerning the quality of the observations and of the associated station metadata (recorded in the OSCAR/Surface database). NWP community benefits not only from a better-quality network, but also from near-real-time access to comparable quality monitoring data from multiple global NWP centres. The potential benefit of exchanging regularly the monitoring results from different NWP centres is highlighted in Hollingsworth et al. (1986) as a diagnostic tool to support the investigations of statistic anomalies, particularly to disentangle observation errors from model errors. It is really an important collective achievement being able to have this exchange in near-real time that allows for a quick and more efficient response towards the observation providers. Here, we show some of the benefits that our participation in this project has brought already to ECMWF's observation monitoring activities, such as helping to detect differences in data reception and in data usage between participating NWP centres, as well as issues in data quality.

This report is organized as follows. In section 2, a general overview of the pilot project on data quality monitoring is given, while in section 3 its practical implementation in near-real time monitoring is presented focusing on the description of the developed web-based graphical User Interface (GUI) prototype and of the initial capabilities of the operational webtool tool (under development), mainly regarding the quality performance. The utilisation of quality monitoring information from other global NWP centres to strengthen ECMWF monitoring capabilities is addressed in section 4. The illustration of the benefit of having Observation-minus-Background (O-B) departures from other models to help us identify and investigate some of the flagged quality issues is illustrated in section 5. Finally, some conclusion as well as future developments are given in section 6.

2 Pilot Project on the Quality Monitoring (QM) function

The development of a pilot project to exchange global observational data quality information was initiated in 2015 (Prates and Richardson, 2016) aiming to explore possible designs for a future implementation of the QM component of WDQMS. First, a template defining the data exchange format for land surface observations (mainly SYNOP) was agreed amongst the NWP global centres taking part in the project. In March 2015, an FTP user account for the pilot project was created by ECMWF to upload the quality monitoring report files and have them available to the WMO secretariat. Also, ECMWF created and maintains a WIKI page[3] dedicated to the project in which all technical details are provided. Additionally, ECMWF made available on the WIKI page some near-real time products to support the Demonstration Project in Africa that ran successfully for nine months (July 2016-March 2017).

Later, it was decided that the monitoring should be extended to the upper-air land observations (December 2016). After agreeing on the template for the exchange reports, ECMWF and the Japan Meteorological Agency (JMA) initiated the generation of near-real-time monitoring reports of upper-air observations over land (e.g. radiosondes).

The structure and format of the files for exchanging monitoring information have evolved significantly since the beginning of the pilot project for both surface and upper-air observations, therefore they required a versioning control which is now applied (information is available on the ECMWF WIKI page under item “4. Template versioning”).

When completed, WDQMS aims at monitoring observational availability, observational quality and observational timeliness for data produced by all WIGOS observing components: the Global Observing System (GOS), the observing component of Global Atmospheric Watch (GAW), the WMO Hydrological Observations (WHO) and the Observing component of Global Cryosphere watch (GCW); and also the co-sponsored observing systems, in particular the Global Climate Observing System (WMO, 2019). The ultimate goal of WDQMS is to monitor the performance of all observing platforms and stations documented in OSCAR/Surface either in near-real time in the case of weather observations or in delayed mode for climate observations. The prototype has successfully integrated the QM of observations from both surface and upper-air stations of GOS located on land. The extension to cover other type of observations has been considered in the prototype development and recently the template

[3] https://software.ecmwf.int/wiki/display/WIGOS/WIGOS+pilot+project+on+data+quality+monitoring

for exchanging the monitoring information of airborne (ABO) observations has been drafted and approved.

2.1 Land surface observations

ECMWF, the German Weather Service (DWD), the US National Centers for Environmental Prediction (NCEP) and JMA are providing quality monitoring reports of land surface observations based on feedback from their data assimilation (DA) systems on a daily basis. These reports include qualitative (quality flag) and quantitative (Observation-minus-Background, O-B, departures) information covering the following observed physical quantities: surface pressure, 2-metre temperature, 2-metre relative humidity and 10-metre wind. From these reports (4 daily, centred at the 4 main synoptic hours, 00, 06, 12 and 18UTC), it is possible to infer the performance of the land surface network both in terms of availability and quality.

2.2 Upper-air land stations

Up to the time of writing, only ECMWF and JMA have been generating quality monitoring reports for upper-air observations. Similar to the surface reports, these include qualitative as well as quantitative information. However, the quantitative information provided is obtained by aggregating the O-B departures into three main categories: surface (Surf); layer between the first pressure level up 100hPa inclusive (Trop); and the layer from 100hPa up to the last reported level (Stra). The quantitative information for the two aforementioned layers consists of both average and standard deviation of O-B departures over the layer for the following observed physical quantities: upper-air temperature, upper-air humidity and upper-air wind. Provision of these data reports (4 daily, centred at the 4 main synoptic hours, 00, 06, 12 and 18UTC), typically happens 24 hours after the actual observation. The availability, quality and completeness of these conventional profiling observations can be easily assessed based on the information provided by these monitoring reports.

2.3 Aircraft based observations

Very recently (October 2018) it was agreed to extend the monitoring capability of WDQMS to airborne observations as part of the strategy to integrate all the WIGOS observing components, particularly the in situ observing systems, into WDQMS. The template has been agreed and includes both qualitative (status flag) and quantitative information of the following observed physical quantities: upper air temperature, upper air humidity, upper air wind, presence of airframe icing, turbulence index and mean turbulence intensity (eddy dissipation rate). The quantitative information includes the O-B departures as well as the observed value itself. In this case, the provision of the data reports (4 daily, centred at the 4 main synoptic hours, 00, 06, 12 and 18UTC), should happen 48 hours after the actual observation, to comply with the airlines data policy.

2.4 Other observation types

The goal of the WDQMS is to integrate all the WIGOS observing components particularly the surface-based systems. Therefore, monitoring activities are planned to be extended to cover other components of the WIGOS that have not been considered such as marine, climate and hydrological observations.

3 Near-real time monitoring of the performance of the Global Observing System

The data quality monitoring practices will be based on the assessment of the performance of observational systems against a set of targets defined for the three performance measures -availability, timeliness and quality - that are proposed in the Technical Guidelines for Regional WIGOS Centres on the WIGOS Data Quality Monitoring System (WMO, 2018) hereafter called WDQMS Guidance Document. The provision of a web-based, interactive Graphical User Interface (GUI) providing access to the monitoring data and presenting it graphically in charts and diagrams is one of the main goals of WDQMS. This WDQMS webtool will be the front end of the QM Function designed to support the Ev Function main activities; all three measures of performance will need to be implemented. WMO developed a prototype of the GUI in which some functionalities have been implemented, mainly related to data availability. Figure 1 shows the diagram of QM data flow for the WDQMS pilot project, in which the database fed by the NWP QM reports and the web-based GUI constitute the back and front end of the QM system, respectively.

3.1 Data Availability

The monitoring of data availability of the surface-based network will be based on performance figures obtained from comparing the observations received from the network to those required and expected to be ingested to the WMO Information System (WIS) according to the schedule determined from OSCAR/Surface metadata. Performance targets can refer to daily, monthly or even annual figures. Currently, availability is based on a 6-hourly temporal aggregation centred on the main synoptic hours and has been implemented for both land surface and upper-air observations. Two extra levels of temporal aggregation should be included to fulfil the WMO requirement in the WDQMS Guidance Document: daily - already partially done in the station time series - and monthly.

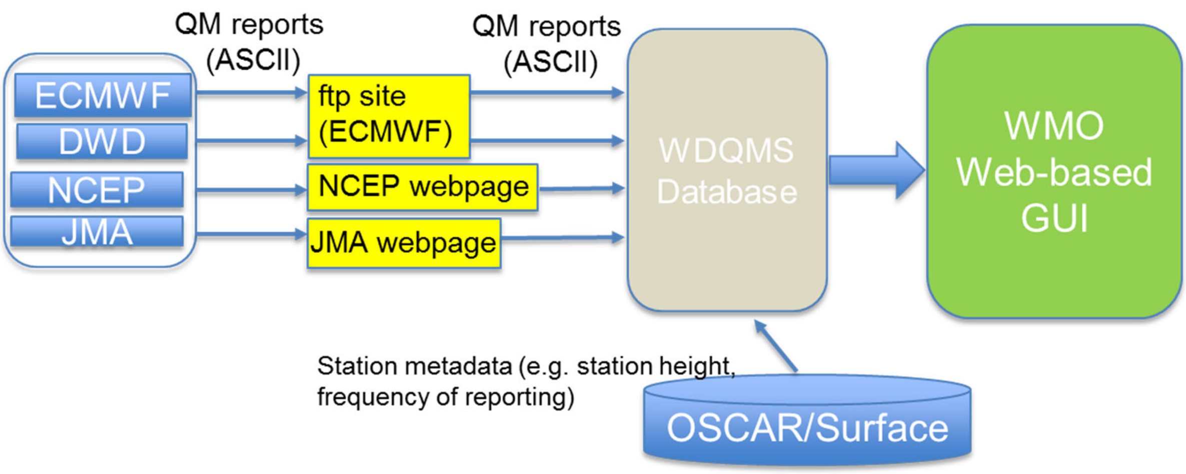


Figure 1. Quality Monitoring (QM) data flow in the WDQMS pilot project.

Land surface observations

Figure 2 shows a snapshot of the WMO web-based GUI for the pilot project in which the global availability of land surface observations is displayed by combining metadata information from OSCAR/Surface with quality monitoring information provided by the four NWP participating centres (i.e. ECMWF, NCEP, JMA and DWD). This tool gives near-real time information about the status of the observational network in terms of availability, highlighting the stations with observational issues, e.g., from not reporting at all to reporting less frequently than expected, and even showing stations that are not included in OSCAR/Surface. Over Europe, the land surface network is generally in good condition with most of stations displayed as green dots (meaning station "reporting as expected"). Noteworthy is the large amount of stations in Iceland and Spain that are reporting, but do not have a WMO-ID attributed (yellow dots, meaning not included in OSCAR/Surface). Those stations are reporting in the new format (Binary Universal Format for the Representation of Meteorological Data - BUFR). This pattern is shown by most of the NWP centres except NCEP (not shown), which at the time of writing is not yet assimilating SYNOP observations in the new BUFR format, therefore they are not made available to the NCEP DA system.

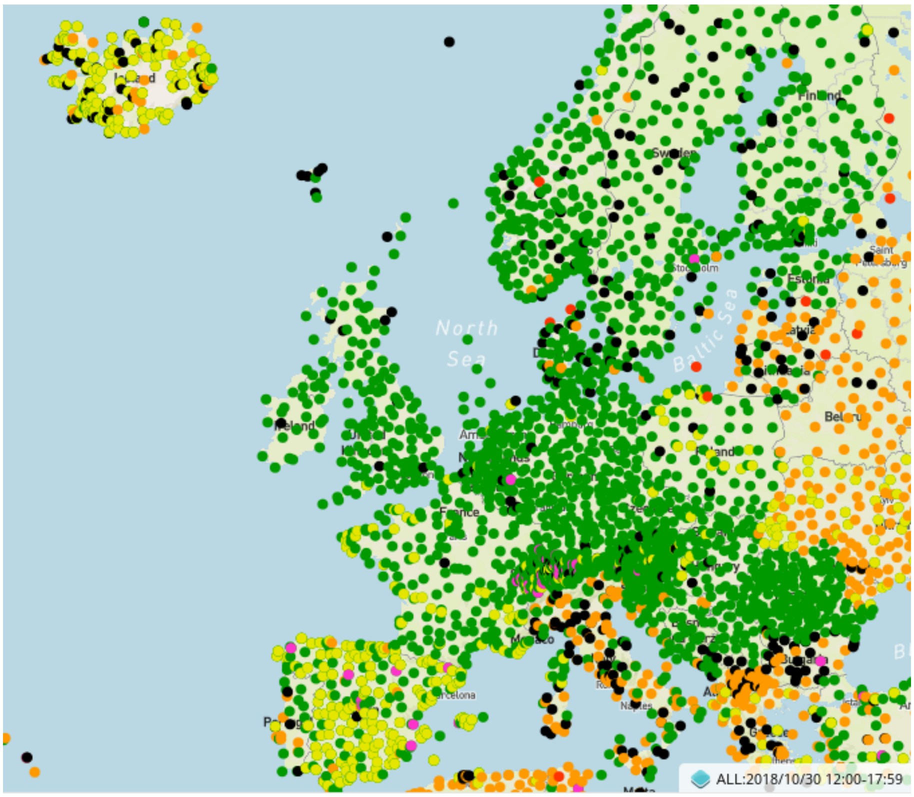


Figure 2. WMO prototype web-tool displaying the status of land surface observational network for 30 of October 2018 at 12UTC, zoomed over Europe, showing the combination of monitoring results from the four NWP centres providing QM data (i.e. ECMWF, NCEP, JMA and DWD). Markers show stations with the number of observations 80% or more of the expected value for the period (green), between 30 and 80% of the expected value (orange), below 30% of the expected (red), above expected (pink), totally missing (black) and station not listed in WMO OSCAR/Surface (yellow).

Upper-air land observations

ECMWF and JMA are providing daily quality monitoring information on upper-air observations (mainly radiosondes over land). The WMO web interface displays in near-real time the map of availability and completeness of these observations. In the old alphanumeric format (TAC) these observations were split into four different reports called parts A (for mandatory level data from surface up to 100 hPa), B (for significant level data from surface up to 100 hPa), C (for mandatory level data above 100 hPa) and D (for significant level data above 100 hPa). With the new high resolution BUFR format, a report is sent with observations up to 100hPa first and after the balloon burst a second report is sent with the full radiosonde (up to the level of balloon burst). Since October 2014 - deadline to migrate to the new format - the old TAC bulletins have been gradually replaced by the new BUFR bulletins. However, there are regions of the world where this has not happened yet. ECMWF has been able to assimilate the new observations as soon as they were deemed of good quality, at least of similar quality as the TAC ones. Figure 3 shows two snapshots of the WMO web-based GUI for the pilot project in which the global availability of upper-air land observations is displayed based on the quality information provided by ECMWF (left) and JMA (right). This tool provides an important description of the status of the upper-air land observation network in terms of availability: it distinguishes the stations that provide a complete report (i.e. all the variables are reported up to and above 100 hPa - in green dots) from the others in which variables are missing (in yellow) or did not report above 100 hPa (orange dots) or did not report at all (black dots). Noteworthy is the common pattern in both ECMWF and JMA results for Chinese stations (yellow), as they do not send humidity observations above 100 hPa. On the other hand, ECMWF and JMA exhibit different results for most of the Japanese stations: ECMWF showing complete reports whereas JMA is missing humidity above 100 hPa. The reason for this discrepancy is the fact that ECMWF was already actively using the BUFR whereas JMA was still assimilating the TAC version, for which humidity observations above 100 hPa are not included in the reports (part B and D). Since then, JMA has started processing high-resolution BUFR observations, but it is still using less than ECMWF.

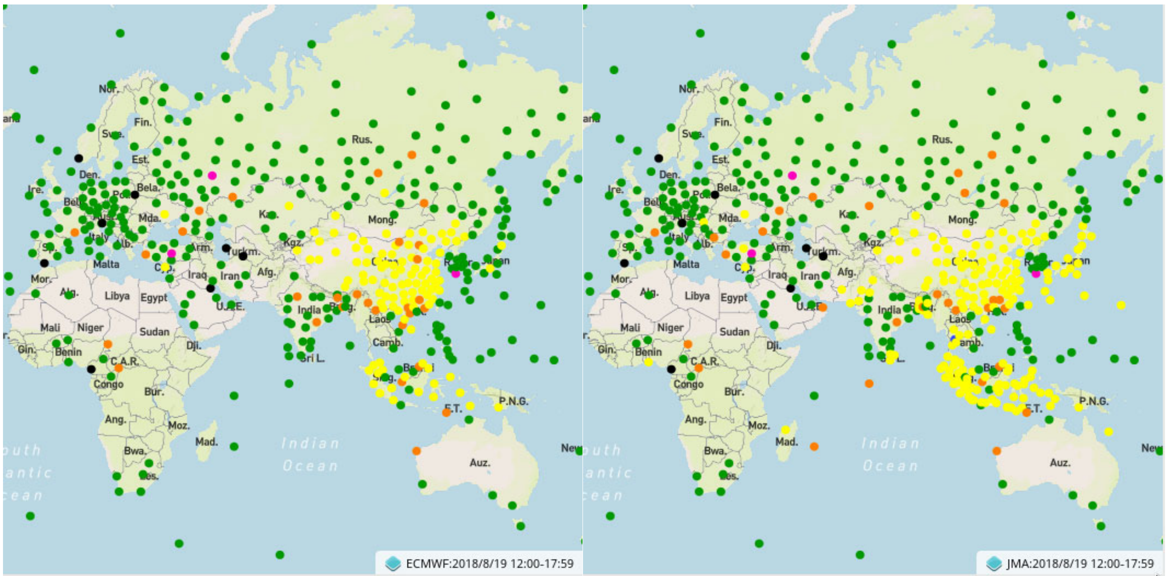


Figure 3. WM O prototype web-tool displaying the status of upper-air land observational network for 19 August 2018 at 12UTC for WMO regions I, II, V and VI, showing JMA (left) and ECMWF (right) monitoring results. Markers show complete launches (green), missing variables (yellow), missing layers (orange), station not reporting (black) and station not listed in WMO Volume A (pink).

3.2 Data Quality

This important measure of performance has not been implemented in the webtool GUI prototype yet. The main quality indicators to be considered are trueness, precision and gross error (WDQMS Guidance Document). The quality indicators will be applied only to the measured quantities whose O-B departures are available in the NWP monitoring reports, i.e. the ones whose model equivalent is available from the NWP assimilation system. Therefore, the surface physical quantities that will be checked are the following: surface pressure, 2-metre temperature, 2-metre relative humidity and 10-metre wind (meridional and zonal components). For the upper-air land observations, the quantities are: air temperature, relative humidity and wind, both meridional and zonal components.

An important distinction needs to be made between land stations in mountainous areas (high elevation stations) and the rest of the land surface stations, because in general NWP centres assess surface pressure, while in the case of stations in mountainous areas they assess geopotential height.

The bias, or systematic error, is used as a measure of trueness and is calculated as the average of O-B departures over a certain period. The targets regarding trueness are stated such that the bias should be close to zero for all measured variables. The trueness will be assessed daily and monthly as recommended in the WDQMS Guidance document. Also, a 5-day running mean of the absolute value of daily calculated O-B departures needs to be calculated daily for all observed variables and compared against the prescribed thresholds. This will be used as one of the main performance indicators on the daily monitoring activities. The standard deviation - an estimate of random error - is the quantitative measure of precision. The targets for precision are applied to the standard deviation of O-B departures over a certain period for each of the observed variables. Like trueness, precision will be assessed daily and monthly. Also, the 5-day moving average of daily calculated standard deviation of O-B will be calculated for all variables and compared to the respective prescribed threshold. This together with the performance indicator for trueness will be used by the Ev function in their daily monitoring activities to determine the level of priority for stations showing accuracy/measurement uncertainty issues (see table in Annex2 of WDQMS Guidance Document).

The number of gross errors in a month (number of single observations whose O-B departures exceed the prescribed threshold) will be computed for each physical quantity at each land station. It will be flagged as a problem when the percentage of gross error per physical quantity is larger than 15% of the total observations of that quantity in the month. For different physical quantities, one must apply different thresholds. The thresholds proposed for land surface observations are the following (from WDQMS Guiding Document): 10hPa for surface pressure; 100 m for geopotential height; 10 K for 2-metre temperature; 15 m/s for wind vector; and 0.30 for relative humidity.

4 Data inter-comparison

ECMWF's participation in the pilot project provides a wealth of information on NWP observational data availability and quality that is not yet being fully exploited. The opportunity to better understand the availability and usage of surface-base data in our data assimilation system is apparent and two potential applications are the detection of differences in data reception (e.g. missing observations) and in data usage (e.g. check if blacklisted observations are used by others) in our data assimilation system. In general, ECMWF data monitoring capabilities can be enhanced by extracting the data quality information provided by the other NWP Centers and comparing it against our own results. Some tools based on WDQMS quality files have been developed in-house to assist the data monitoring activities at ECMWF, namely the blacklisting of in situ observations.

4.1 Data usage

The most basic evaluation is to compare the number of observations each NWP has available in their assimilation system. Having access to land surface quality monitoring information (based on SYNOP observations) from the four NWP centres participating in the Pilot project allows us to compute routinely the average number of observations available at each NWP centre daily. An example plotted in figure 4 shows the average number of land surface pressure/geopotential height observations in a 24-hour period for June 2018. ECMWF shows the largest amount of observations available followed by JMA, NCEP and DWD. However, the discrepancies in the total amount of observations seen by the different NWP centres can be partly explained by the Data Assimilation (DA) characteristics of each of the different NWP systems. For example, NCEP shows only the data that pass a quality control step prior to DA, therefore part of the data available that is deemed to be of poor quality or duplicate is filtered out and is not available to the DA, whereas ECMWF shows even the data that is deemed to be of poor quality and is rejected and/or blacklisted. That explains some of the difference seen in totals from the two centres.

On the other hand, a fairer comparison is to look at the number of assimilated surface pressure/geopotential observations from the land surface reports, and this shows that ECMWF assimilates by far the largest number, followed by NCEP, DWD and JMA. The major differences seen in percentage of used observations are partly due to the frequency of data assimilated (i.e. 1-hourly at ECMWF versus 3-hourly at JMA). Also, DWD's current global DA system can only use actively one SYNOP observation per 3-hour window at a given site.

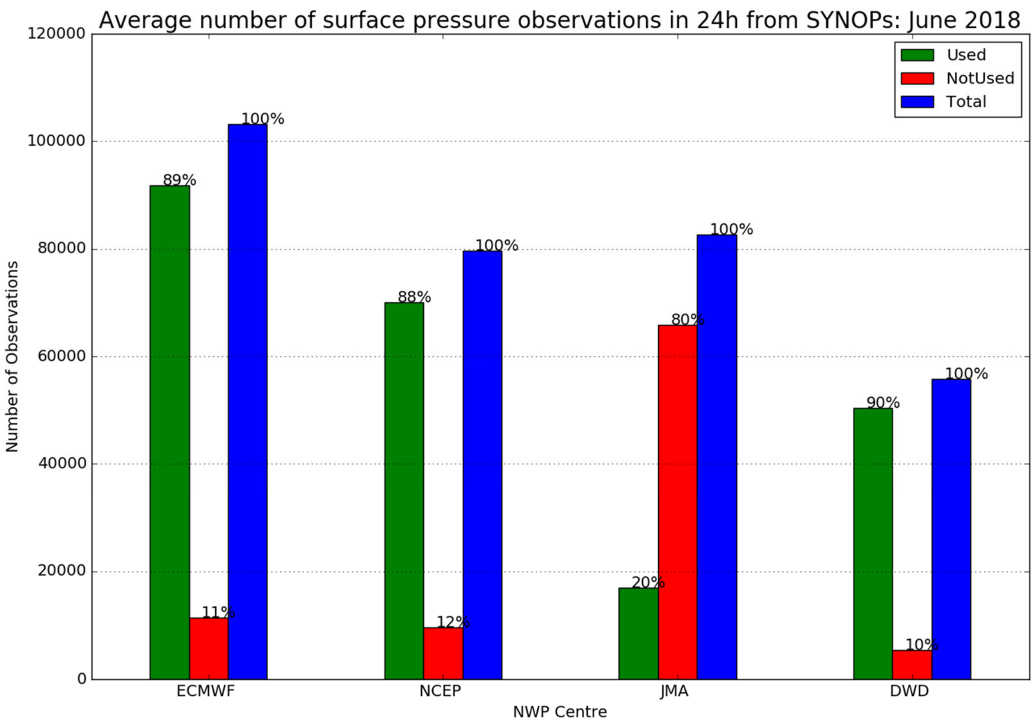
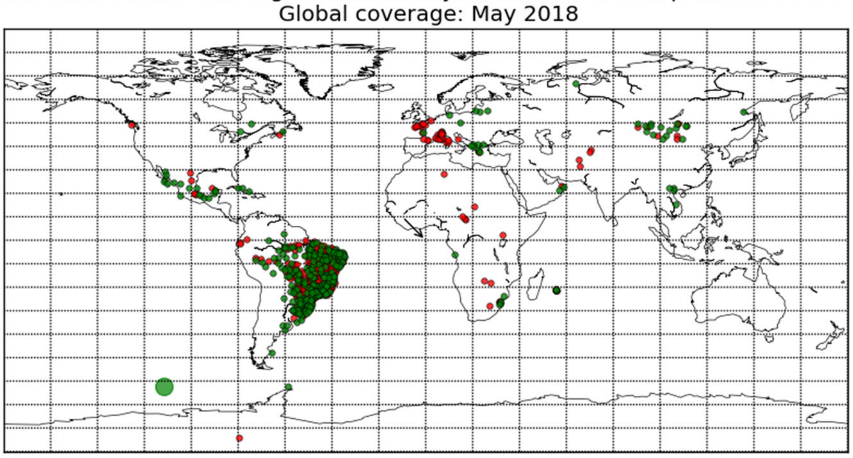


Figure 4. Monthly average of the number of land surface observations in 24 hours available for Data Assimilation in four NWP global centres: ECMWF, NCEP, JMA and DWD. The total (blue) is breakdown in used in the assimilation (green) and not used in the assimilation (red).

Despite having in general more observations than other centres, ECMWF can still miss observations that are available to other NWP DA systems. Regularly checking the data availability against the other three centres can prove to be useful in detecting inconsistencies in the global data coverage. For example, Figure 5 shows the global coverage of stations that were available to the DWD DA system but missing in ECMWF's at least 10 days in a month. The coverage for May 2018 shows many stations over Brazil, some used (green dots) some not used (red dots) by DWD. These stations are 1-hourly BUFR land SYNOP reports disseminated via GTS, which at the time were not included in ECMWF's Observation Database (ODB). After some quality checks, they were included in ECMWF DA to be passively monitored as we can confirm from the coverage for November 2018.



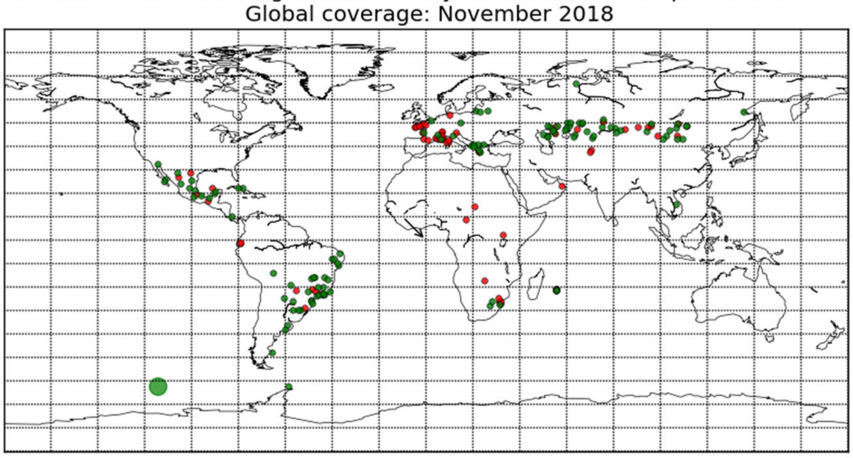
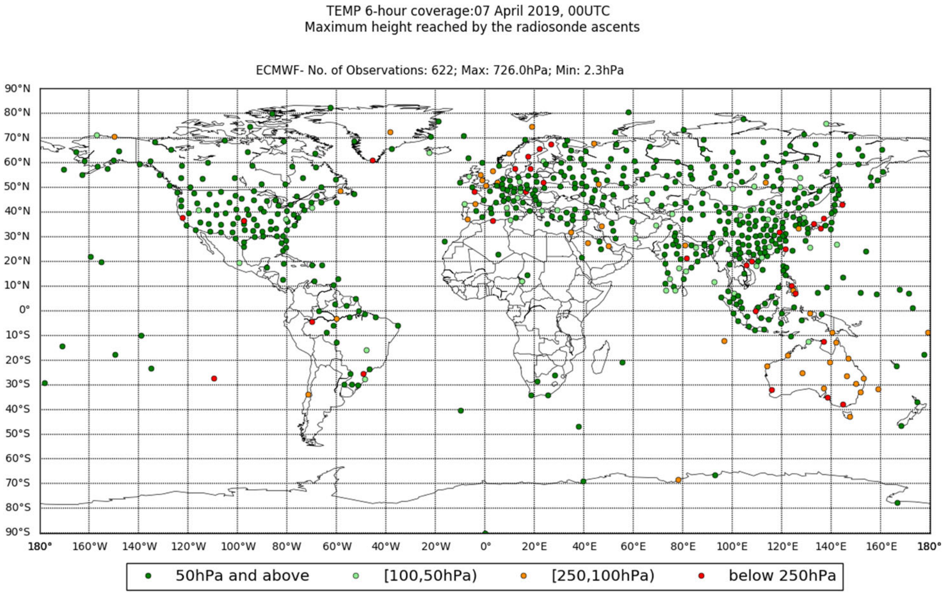


Figure 5. Surface synoptic land stations with observation available in DWD DA system but missing in ECMWF for at least 10 days within a month: global coverage for May 2018 (top) and November 2018 (bottom). The green dots refer to stations whose observations (mainly surface pressure or geopotential height) have been used by DWD and the red dots pertain to stations with observations that have not been used by DWD data assimilation system.

Another important aspect of having information available in near-real time from other global NWP centres, is that it allows for an immediate cross-check in case of any detected anomaly. For example, on the 7th of April 2019, the automatic alarm system flagged missing radiosonde ascents in some WMO regions (mainly region VI and V). The inspection of 00UTC maps of radiosonde maximum height on 7th of April showed many radiosonde ascents achieving low burst heights over Europe and Australia; and a cross-check with JMA monitoring data revealed a similar pattern (see Figure 6). Furthermore, on the 12UTC maps of 7th of April the observations from those stations were completely missing in both ECMWF and JMA Data Assimilation systems. The problem was later identified to be related to the GPS Epoch[4] on the morning of 6th of April 2019. The GPS Epoch event affected some of the Vaisala systems globally. The disruption depended on the firmware in which the systems are operating. For example, all UK autosonde systems failed, but the manual sondes were not affected. Vaisala promptly provided a firmware update which needed to be manually installed at the autosonde sites and gradually the network went back to normal (48 hours after the event most of the issues had been sorted).

[4] GPS signals from satellites include a timestamp, needed in part to calculate one's location, that stores the week number using ten binary bits. That means the week number can have 210 or 1,024 integer values, counting from zero to 1,023 in this case. Every 1,024 weeks, or roughly every 20 years, the counter rolls over from 1,023 to zero. The first Saturday in April 2019 (6th) marked the end of the 1,024th week, after which the counter will spill over from 1,023 to zero. The last time the week number overflowed like this was in 1999, nearly two decades on from the first epoch in January 1980. If devices in use today are not designed or patched to handle this latest rollover, they will revert to an earlier year after that 1,024th week in April, causing attempts to calculate position to potentially fail.



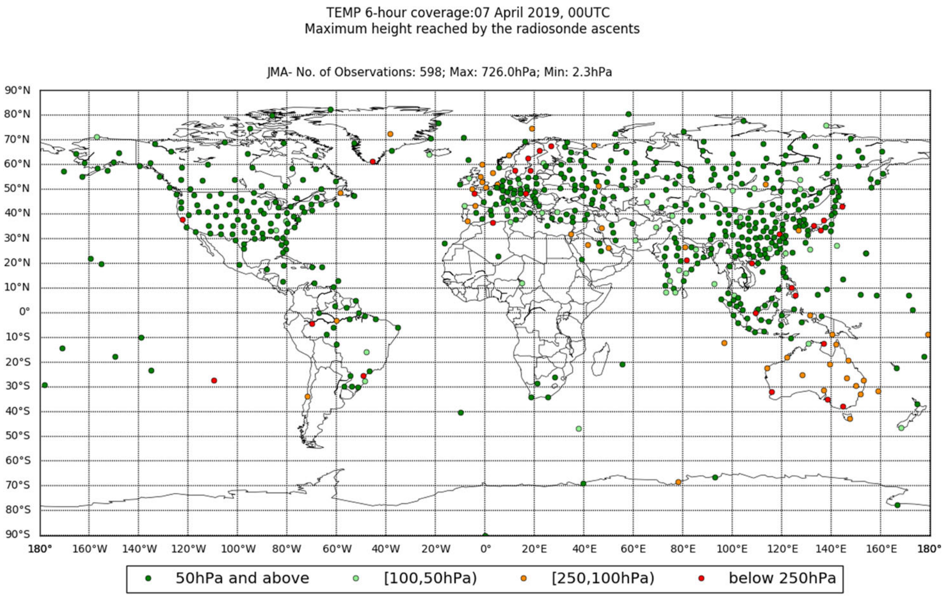
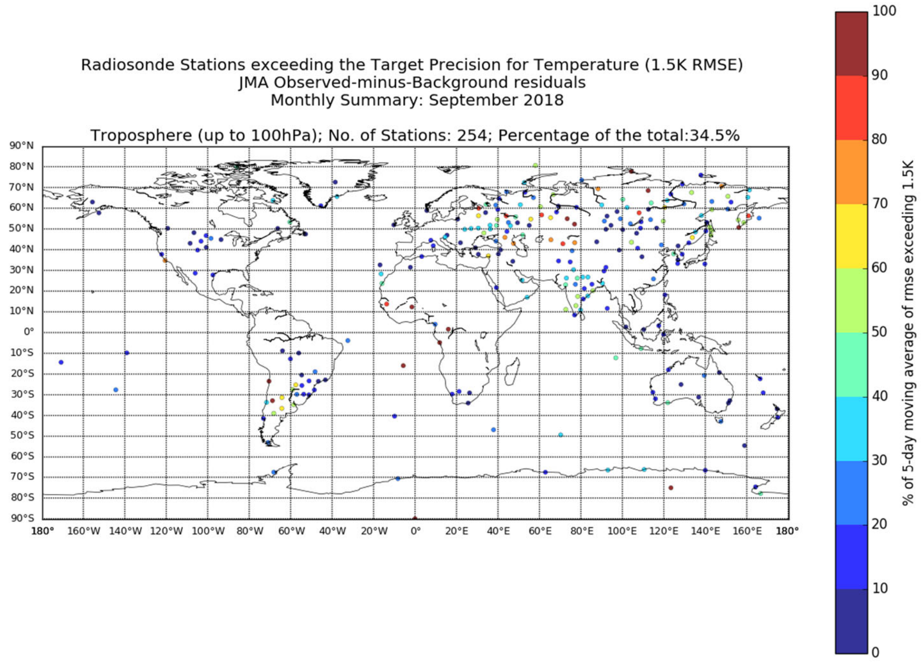


Figure 6. The radiosonde last reported pressure level for all upper-air land stations reporting air temperature within 6-hour interval centred at 00UTC of 7th of April 2019. The top plot refers to the monitoring information provided by ECMWF and the bottom one to the JMA monitoring results.

4.2 Data accuracy

One important aspect of monitoring the health of the GOS is to assess the quality of each individual observation, so that one can assess the average quality of a station for example. All data assimilation systems perform a quality check on the observations (Kalnay 2003, section 5.8) before assimilation and observations may not be used because they are blacklisted (e.g. due to known poor quality) or simply because they are rejected during the assimilation cycle. The evaluation of the quality of the observations on a station basis is paramount to the monitoring component of WDQMS. The Guidelines published in the WDQMS Guidance document regarding the performance quality indicators required by the QM component will need to be implemented in the web graphical interface (as described in section 3.2). At ECMWF, some preliminary studies have been done to develop products to assess the quality of surface-based stations of the GOS. For example, the guidelines for upper-air observations recommend a 1.5 K threshold to be applied to the rmse of temperature O-B profile; and, if the station persists at least for five consecutive days with daily rmse values exceeding the threshold, the Ev component must raise the issue to IM component so that an action is taken towards the observation provider. Figure 7 displays an example based on temperature observations up to 100 hPa in September 2018. The percentage of 5-day moving average of rmse exceeding the 1.5 K threshold for ECMWF and JMA is plotted for all upper-air stations providing temperature observation in that month. The differences between the two global NWP centres are significant. ECMWF only flags 20.8% of the total number of stations whereas Japan flags 34.5% of the total network. It is worth mentioning that in the case of upper-air land observations the total number of observations available to both ECMWF and JMA are similar. These plots illustrate well the difficulty on defining aggregation rules to decide if a particular station is producing good quality observations because the answer will depend on the NWP model background against which observations are compared. Obviously, if both centres agree on a particular station, this indicates that the chances of the observations from this station being of poor quality are very high. Furthermore, the overall quality of the temperature measured by upper-air stations since January 2017 (when the monitoring reports became available) displays a seasonal behaviour. This is readily apparent in Figure 8 where the time series with the percentage of observations being flagged monthly by applying the threshold to the 5-day moving average of temperature O-B rmse is shown for the two centres. Both centres exhibit a larger percentage of stations with quality issues in winter (January is the worst month) than in summer (September exhibits the best scores).



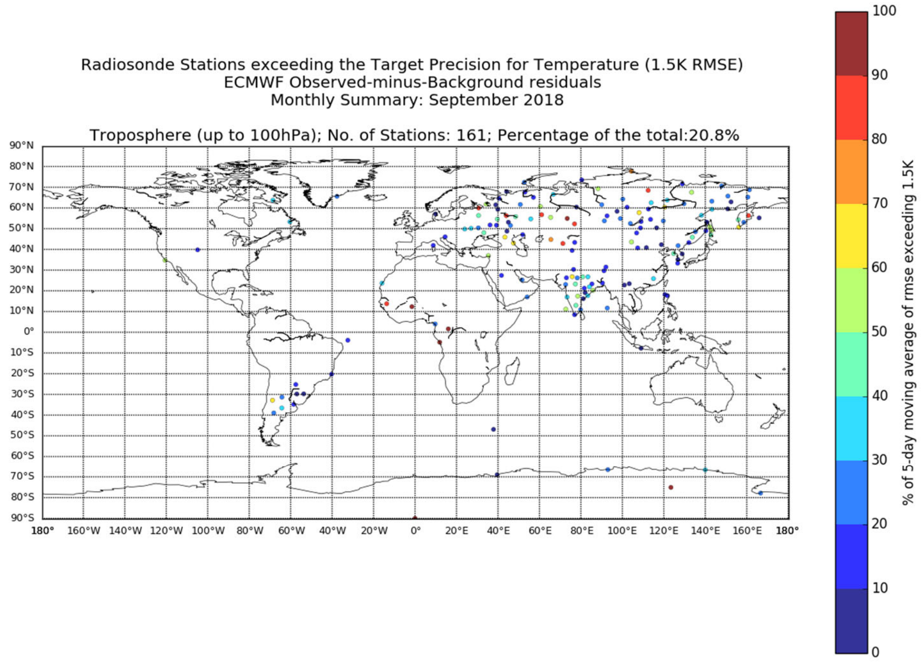


Figure 7. Flagged stations that fail the quality target of 5-day moving average of rmse exceeding 1.5 K.

Apart from the seasonal effect on the O-B fit to the observations, there is also a noticeable trend towards smaller number of stations with quality issues. in both ECMWF and JMA systems. There are several possible reasons for this trend, the observations are becoming more accurate, or the models are getting better or a combination of both. We will try to answer this question below. As mentioned in section 3.1, the migration to high-resolution BUFR has been gradually taking place and an increasing proportion of BUFR reports are being used by the NWP models. At ECMWF, for example, in January 2017 only a small percentage of observations from BUFR reports was used in the assimilation (around 12%); the steepest increase occurred between August and November 2017, reaching at that time around 31% of the total of used observations in the system; since then a gradual increase occurred and the percentage has plateaued around 38% since October 2018. The high-resolution BUFR reports provide a more accurate representation of the measured profile, which allows for a better use in NWP, namely the suboptimal assimilation of significant levels (typical in TAC messages) can be replaced by levels randomly selected by the thinning process (Ingleby, Pauley, et al. 2016). The most significant improvements in the use of radiosonde data have been introduced in IFS Cycle 45r1 which became operational in June 2018. Importantly, the extra information about the balloon drift provided by the new BUFR messages has started to be taken into account in the assimilation of radiosonde measurements for the stations that are reporting in BUFR. This might explain the significant improvement seen in the monthly values compared with the same months in the previous year. According to Ingleby et al. (2018) the improvement can be up to 10% in the temperature O-B standard deviation values for those reporting the drift positions. In Figure 8 we also plot the time series of monthly percentage of stations that failed the temperature criteria in the ERA5 reanalysis (Hersbach and Dee 2016). Note that ERA5 is based on 4D-Var data assimilation using cycle 41r2 of IFS, which was operational at ECMWF in 2016. As expected the larger differences between ECMWF and ERA5 are seen in the months after ECMWF last model upgrade which already incorporates the drift from BUFR radiosondes (June 2018).

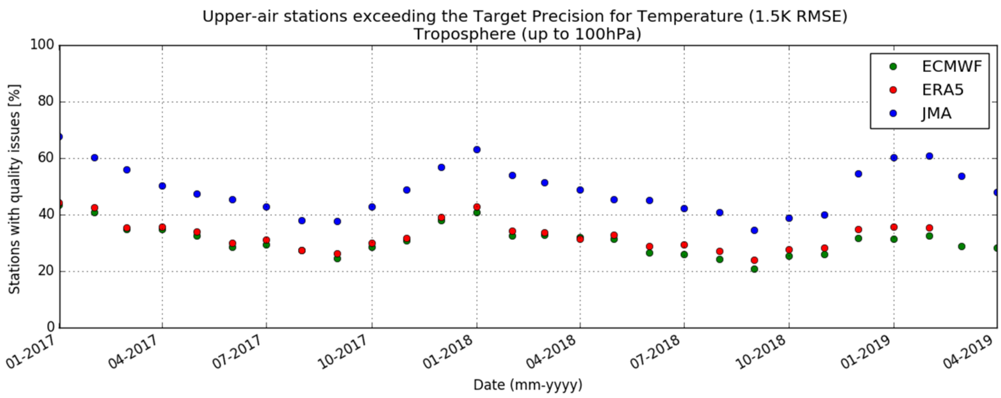


Figure 8. Time series of the monthly percentage of stations that failed the temperature quality target in ECMWF (green), JMA (blue).and ERA5 (red) DA systems.

4.3 Blacklisting

At ECMWF, the quality control (QC) used for the in situ observations is the so-called variational quality control approach (Andersson and Jarvinen 1999) performed within 4D-Var, and thus it is part of the analysis itself. However, there is still some quality control prior to assimilation that is done monthly and it involves not only identifying stations which are providing systematically poor quality observations of the assimilated physical quantities, but also the ones which were not used but for which the quality of observations has improved (Haiden, et al. 2018).The former list of station/physical quantity is called blacklist and the latter whitelist. These lists are based on the monthly statistics of O-B departures for each of the relevant physical quantities.

The process of building these lists relies heavily on a dedicated observation events database populated daily by the automatic alarm system (Dahoui, Bormann and Isaksen 2014) with information of anomalous and improved in situ observations. Although highly automated, the blacklisting/whitelisting process still requires a final human evaluation and intervention. There is a web tool designed to manage all the processes that display the automatic proposed blacklist/whitelist generated from the observation events database, and also some ad-hoc proposals. This webtool allows the evaluator to accept or reject the proposals and add comments. The idea is to have a documented justification for each of the choices made.

The quality monitoring information provided by the participating NWP centres is already being used to support the blacklisting activities at ECMWF. First, observations from all the land stations/physical quantities included in the automatic proposed list are compared against the statistics from all the four NWP centres, if available, and displayed on the web-tool under the column "Multi Centre comparison". In this way, it is easy to assess if it is an issue common to all centres or specific to ECMWF DA system. Second, an automatic quality check is performed against other NWP results to assess if any of the stations assimilated by other centres but blacklisted by ECMWF are of good enough quality to be whitelisted. Initially, the number of proposed stations was huge and showed that the alarm system was missing some surface pressure observations. This helped to identify and correct the issue in the alarm system (i.e. use the generic surface pressure variable to check for the pressure at surface, which can be either station level pressure or pressure reduce to the sea level depending on the quantity that has been used in assimilation). Currently, it complements the alarm system that sometimes does not spot a quality improvement, adding on average a few more stations to the whitelist every month. For example, the number of land stations sugested by the Multi Centre comparion based on July 2018 statistics was seven and only one of them was also suggested by the automatic system[5] A plot with the time series of O-B surface pressure is always produced and included in the web-tool interface to support the evaluator.

5 Data Quality Issues

The diagnosis of an anomaly in the O-B statistics requires investigations to determine the origin of the issue. The mismatch between observations and model short range forecasts can have different causes, including inaccurate metadata, poor quality observations, and model errors. It can be difficult to differentiate between some of the causes, mainly between observations and model errors. Having regular exchange of monitoring results from other NWP centres can help to identify the problem and even differentiate between possible causes. For example, if some correct observation deviates from the background because the NWP model is not perfect, the fact that we can cross-check this with other model forecasts helps to discriminate whether errors are related to observation or to the NWP model

[5] https://confluence.ecmwf.int/display/EVAL/Conventional+Observation+Blacklist+Proposal%3A+2018080700

itself. Some examples are shown in this section, illustrating the benefit of intercomparing monitoring results from different analysis centres.

5.1 Metadata (Sensor) issues

The list of stations that should contribute towards the Regional Basic Synoptic Network (RBSN) can be obtained from the OSCAR/Surface database (which was deployed operationally on 2 May 2016 to replace WMO No. 9, Volume A). This list, together with the associated station metadata, is extracted to be used in the assimilation of surface observations, particularly surface pressure.

Surface pressure is the most important *in situ* observed quantity for global NWP forecasting, and in some cases the only observed surface quantity over land used in the global atmospheric data assimilation, e.g. in the JMA global atmospheric data assimilation system (JMA 2017). ECMWF's atmopsheric 4D-Var also assimilates relative humidity over land, but only at nighttime. Both station level pressure and pressure reduced to sea level should be reported in SYNOP messages (for high elevation stations the height of a standard pressure level replaces the pressure reduced to sea level). ECMWF DA system tends to give preference to station level pressure - rather than pressure reduced to sea level - when both are available in the SYNOP messages. Hence the barometer height is essential for the observation operator to generate a model counterpart for the observation quantity – pressure in this case - in observation space. The DA system relies solely on Volume A metadata in the case of the old format (traditional alphanumeric code -TAC) reports, whereas in the new format (Binary Universal Format for the Representation of Meteorological Data -BUFR) this information is also included in the report. However, for the RBSN stations that are reporting in the new format the system checks the metadata provided in the report against the volume A legacy information. A good illustration of the importance of metadata in the usability/quality of the pressure measurements is given in the example below. Furthermore, it also shows how the near-real time access to O-B departures from multiple NWP centres can help to diagnose this type of quality issue.

After a short outage between 11 and 18 of February 2019, a large positive surface pressure bias was noticed in ECMWF O-B statistics for the surface land station Trollenhagen, in Germany (WMO-ID 10281). A similar bias was also detected in the statistics based on the other NWP centres (Figure 9), which indicated that something had changed after the outage and was causing these biases in the surface pressure measurement reported in the SYNOP bulletins from this station. It was discovered that the origin was a metadata problem: the barometer originally at 93m had been moved to a new position on the ground floor of the observation tower, being the new barometer height 69.4m. The OSCAR/Surface had been updated accordingly, however none of the NWP centres made the correction to reflect the change, except DWD (not shown). After a prompt correction in the metadata to reflect the changes, ECMWF statistics for that station went back to the previous values.

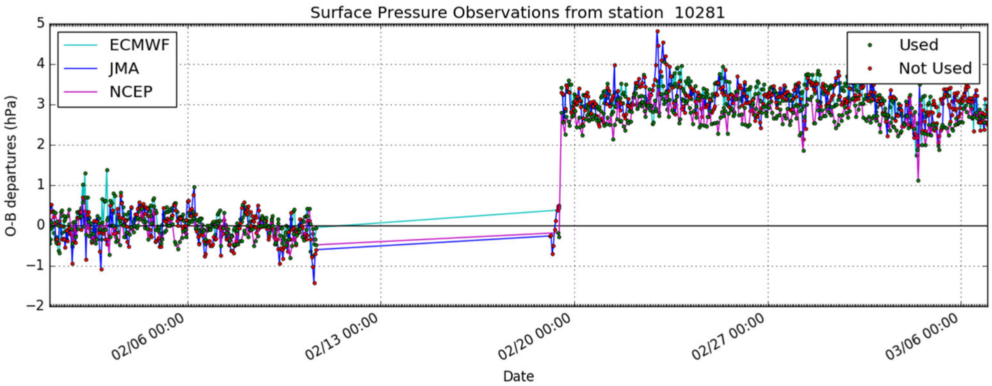


Figure 9.Time series of surface pressure Observations minus Background (O-B) departures for the land surface station Trollenhagen in Germany (WMO-ID 10281). The different lines correspond to O-B from three NWP centres: ECMWF (light blue), JMA (dark blue) and NCEP (pink). The green and red circles correspond to observations that were used and not used, respectively, by each of the individual data assimilation systems.

5.2 Observation errors

The Russian station Pirovskoe (WMO-ID 29363) was flagged by ECMWF's automatic alarm system due to the deterioration of the monthly surface pressure O-B statistics for December 2018. Before accepting the automatic proposal to blacklist the surface pressure observations from this station, some checks were made to understand the possible nature of this deterioration. By cross-checking the ECMWF O-B departures with the departures from the other NWP centres (Figure 10) it became apparent that an observation problem may have been the cause of this, as all the three NWP centres (ECMWF, NCEP and JMA) exhibited a similar pattern. During December there were two periods of a few consecutive days (3-9 and 24-29) in which large departures were obtained in all the NWP monitoring statistics available. Careful examination of the observations revealed that during these two periods in December, the reported surface pressure observations were constant throughout the periods (1013.4 hPa and 1014.7 hPa for the first and second period, respectively). Therefore, the variations detected in the O-B time series in those periods were the result of surface pressure fluctuations in the background, and obviously in some cases the departures became very large because the observation values were completely wrong and did not represent the real surface pressure.

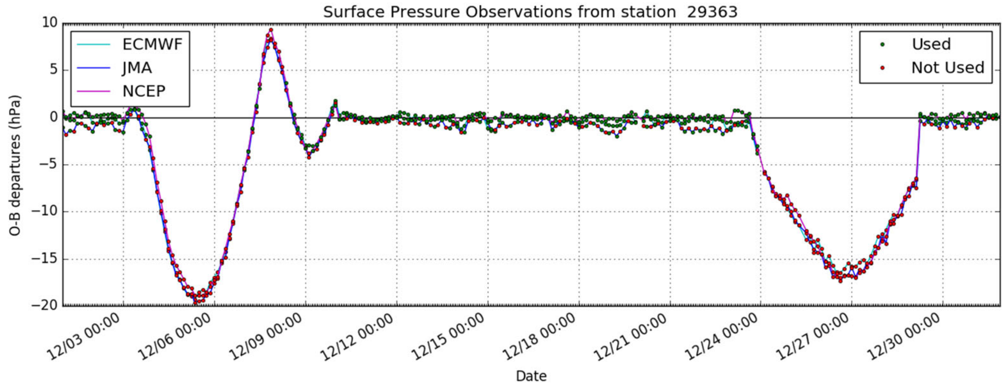


Figure 10.Time series of surface pressure Observation minus Background (O-B) departures for the synoptic station Pirovskoe, Russian Federation (WMO-ID 29363). The light blue, dark blue and pink pertain to ECMWF, JMA and NCEP O-B departures, respectively. The green and red circles correspond to observations that were used and not used, respectively, by each of the individual data assimilation systems.

5.3 Model Issues

A strong signal in the Extreme Forecasts Index (Lalaurette 2003; Zsótér 2006) for 2-metre temperature was detected in the Tibetan Plateau region in Asia in October 2018 and discussed as part of daily report activities at ECMWF (David Levers, personal communication, October 26, 2018). The 2-metre O-B departures from some stations in the region were investigated, and in some cases very large positive biases were found, namely in the stations Nagqu (WMO-ID 55299) and Lhasa (WMO-ID 55591). These large biases started at the beginning of October and persisted throughout the month. The initial hypothesis of observation errors being the cause of such large bias was dismissed when ECMWF's O-B time series for October 2018 was compared with JMA's. Figure 11 shows this comparison for Nagqu station. It is apparent that the large positive bias (10K on average) seen in ECMWF O-B departures during October, is not mirrored in JMA statistics at all. Actually, the JMA statistics suggest that the observations are of good quality, occasionally with some negative departures of the order of -5K.

The large bias is definitively due to errors in ECMWF short-range forecasts of 2-meter temperature, which are systematically lower than the observed values. Figure 12 displays the observed 2-metre temperature and the model background derived from ECMWF and JMA. The ECMWF model exhibits 2-metre temperatures well below freezing, whereas the observed 2-metre temperatures oscillate around 0°C. JMA short-range forecasts for 2-metre temperature are mirroring the observation fluctuations. Some further investigations revealed that a snowfall event took place in the 27^{th} of September and lead to snow accumulation on the ground in both the model and the real world. However, the snow in the real world melted afterwards (observed temperatures were above zero in late September), whereas in the model the snow did not melt completely. As a result, the snow that remained on the ground cooled the air and probably a surface-based inversion arose as a consequence and more radiation was reflected reinforcing the temperature drop in the model. This problem persisted for all of October 2018. The fact that fresh surface snow often tends to take too long to melt although ground temperatures are forecast above 0°C is a known problem in the ECMWF model (Owens and Hewson 2018, section 9.2).

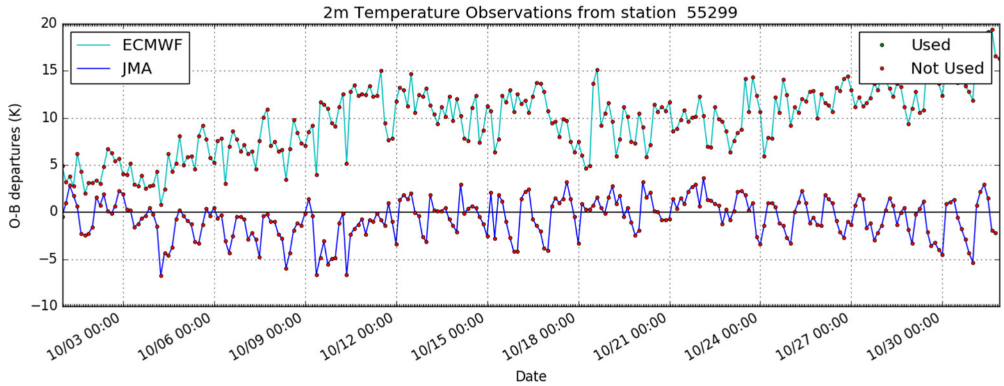


Figure 11. Time series of 2-metre temperature Observation minus Background (O-B) departures for the synoptic station Nagqu in the Tibetan Plateau (WMO-ID 55299). The light and dark blue pertain to ECMWF and JMA O-B departures respectively. Note that 2-metre temperature over land is not assimilated by both NWP, therefore the observations are represented by red circles.

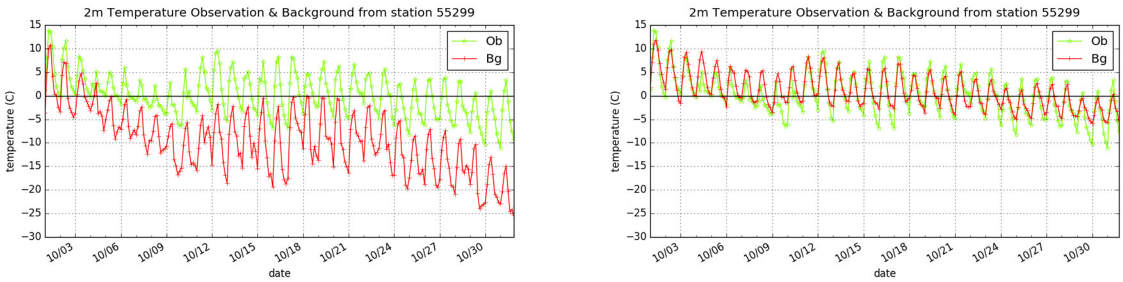


Figure 12. Time series (October 2018) of observed 2-metre temperature (Ob, green) for synoptic station Nagqu in the Tibetan Plateau (WMO-ID 55299) and the model background (Bg, red) derived from ECMWF (left) and JMA model forecasts (right).

6 Conclusion and future developments

The WIGOS pilot project on data quality monitoring, including the GUI prototype developed by WMO, has shown that the WDQMS is a viable concept. Once fully implemented, WDQMS will allow the management in near-real time of the in-situ component of WIGOS. With its three-main functions - QM, Ev and IM-, the system aims at monitoring, evaluating and triggering corrective procedures when anomalies are found in the comparison between actual WIGOS observational data and the user requirements for these observational data. These requirements include availability, timeliness of delivery and observational data quality, including completeness. The incident management process when in place should ensure that issues with individual stations are detected and acted upon.

WDQMS is considered a fundamental WIGOS technical tool to help WMO members with network evaluation and design, as well as trouble-shooting, and is potentially a transformational activity for WMO itself. The pilot project, which has involved a substantial effort from many NWP centres, namely ECMWF, NCEP, JMA and DWD, is now sufficiently mature to transform into an operational implementation. Despite covering only data availability, the WDQMS GUI prototype has already proven to be a useful web tool, most notably in revealing poor data coverage and inadequate reporting of observational data over some areas of the globe (WMO, 2019); this web tool is seen as the front end of the QM function designed to support Ev function's main activities. ECMWF has agreed to develop the future operational GUI and expand it to include also the monitoring of data quality. Furthermore, ECMWF will become the lead WDQMS Quality Monitoring Centre with the responsibility of running and maintain the future WDQMS web-based GUI. A MoU between ECMWF and WMO setting ECMWF's future role has been signed in Septembre 2018.

Both OSCAR/Surface and WDQMS are vital for WIGOS. OSCAR/Surface is intended to document the capabilities of surfaced-based WIGOS observing systems and to support the WIGOS Rolling Review of Requirements process. WDQMS, on the other hand, is intended to monitor the actual performance of the surface-based WIGOS components, which is essential for any meaningful optimization or redesign activity. As shown, OSCAR/Surface will provide WDQMS with the relevant observational metadata that informs the expected performance of observational data under consideration. On the other hand, WDQMS will inform about the status of observational systems regarding availability, quality and timeliness and this information should be recorded in OSCAR/Surface. The need to make WDQMS interoperable with OSCAR/Surface, meaning that essential information should flow between the two systems to ensure consistency, has been stressed by both task teams, TT-WDQMS and the task team on OSCAR Development. Finally, the establishment of RWCs is critical for advancing operation in WIGOS. These centres will be responsible for the regional WIGOS metadata management (OSCAR/Surface) and the regional WIGOS performance monitoring and incident management (Ev and IM components of WDQMS). In practice, they will have to work closely with data providers to facilitate collecting, updating and quality control of WIGOS metadata in OSCAR/Surface. Also, they will have the mandate to evaluate and raise incidents when appropriate and follow-up with data providers in case of data availability and quality issues. It is worth noting the impact that the timely updating of OSCAR/Surface metadata by Members to reflect the latest status of their observing networks has in minimising issues being identified with observing network performance against OSCAR/Surface declared intent.

The benefits of WDQMS itself are wide-ranging. Many application areas, most notably NWP, will benefit from improved performance of in situ observing network/systems and improved and documented data quality. The benefits to NWP are obvious in terms of improved observation quality, and shorter periods of blacklisting of problematic stations, which will contribute to improved forecast quality. Data providers - mainly National Meteorological Services (NMSs) - will see their observations being used due to improved quality particularly by NWP centres which can extract the full benefit from the investment made in observations. Furthermore, the open access to WDQMS GUI will allow all NMSs to get feedback in near-real time on the usage and quality of their observations. Also, the access to the WDQMS GUI will allow the different global NWP centres to enhance their monitoring capabilities by comparing the data coverage and data quality with other NWP centres.

At ECMWF, we have seen that our participation in this project has already been beneficial for the daily activities of monitoring the availability and quality of the surface-based network. The extension to other observation types will enhance further our monitoring capabilities, like the exchange of quality monitoring information of ABO, which will be extremely useful to understand some of the issues that are detected daily in the monitoring of the rejections. Given the importance of aircraft observations in the ECMWF assimilation system (Bormann, Lawrence and Farnan 2019) and the need for their quality control prior to assimilation, the cross-checking of the monitoring results from other NWP centres is of particular interest and would definitively improve the blacklist activity.

In situ observations are also widely used in forecast verification. At ECMWF, the list of radiosonde stations to be used in the standardised verification of NWP products is generated annually - as part of our tasks as WMO Lead Centre - based on ECMWF DA feedback files. This quality control has the drawback of creating an observational dataset too dependent on our own model. Therefore, a multi-centre approach in which monitoring information from different centres is aggregated to generate a list of good quality stations is a better alternative. In this way, the model errors from a single model will potentially have less impact on the assessment of the quality of observations and potentially less observations would be wrongly removed from the list. Furthermore, there is also the opportunity to implement the new guidelines on the quality of observations (WMO, 2018) potentially leading to a more accurate best quality station list.

Acknowledgements

We would like to thank Lars Peter Riishojgaard, Luís Nunes and Timo Pröscholdt from WMO Secretariat as well as all the members of TT-WDQMS for their contribution to the conception, design and prototype implementation of WDQMS.

References

Andersson, E., and H. Jarvinen. 1999. "Variational quality control." *Q.J.R. Meteorol. Soc.* 697-722.

Bormann, N., H. Lawrence, and J. Farnan. 2019. "Global observing system experiments in the ECMWF assimilation system." *ECMWF Technical Memorandum* No. 839.

Dahoui, M., N. Bormann, and L. Isaksen. 2014. "Automatic checking of observations at ECMWF." *ECMWF Newsletter* No. 140: 21-24.

Haiden, T., M. Dahoui, B. Ingleby, P. de Rosnay, C. Prates, E. Kuscu, T. Hewson, et al. 2018. "Use of in situ surface observations at ECMWF." *ECMWF Technical Memorandum* No. 834.

Hersbach, H., and D. Dee. 2016. "ERA5 reanalysis is in production." *ECMWF Newsletter* 147.

Hollingsworth, A., D. Shaw, P. Lönnberg, L. Illari, K. Arpe, and A. Simmons. 1986. "Monitoring of observation and analysis quality by data assimilation system." *Mon. Wea. Rev.* 114: 861-879.

Ingleby, B., L. Isaksen, T. Kral, T. Haiden, and M. Dahoui. 2018. "Improved use of atmospheric in situ data." *ECMWF Newsletter* No. 155: 20-25.

Ingleby, B., P. Pauley, A. Kats, J. Ator, D. Keyser, A. Doerenbecher, E. Fucile, et al. 2016. "Progress towards high-resolution, real-time radiosonde reports." *Bull. Amer. Meteor. Soc.* 97: 2149-2161.

JMA. 2017. "Join WMO Technical progress report on the global data processing and forecasting system and numerical weather prediction research activities for 2017." https://www.jma.go.jp/jma/jma-eng/jma-center/nwp/report/2017_Japan.pdf.

Kalnay, E. 2003. "Atmospheric modelling, Data assimilation and Predictability." *Cambridge University Press.*

Lalaurette, F. 2003. "Early detection of abnormal weather conditions using probabilistic extreme forecast index." *Q. J. R. Meteorol. Soc.* 3037-3057.

Owens, R., and T. Hewson. 2018. "ECMWF Forecast User Guide." *Reading:ECMWF.* doi:10.21957/m1cs7h.

Prates, C., and D. Richardson. 2016. "ECMWF takes part in WMO data monitoring project." *ECMWF Newsletter* No. 148,19.

WMO. 2010. "Manual on the GDPFS." *WMO-No. 485* Volume 1. https://www.wmo.int/pages/prog/www/DPFS/documents/485_Vol_I_en.pdf.

WMO. 2018. "Technical Guidance for Regional WIGOS Centres on the WIGOS Data Quality Monitoring System." *WMO-No. 1224.* https://library.wmo.int/doc_num.php?explnum_id=5681.

WMO. 2019. “WDQMS: NWP Pilot Project and Preliminary Resulst.” http://www.wmo.int/pages/prog/www/wigos/tools.html.

Zsótér, E. 2006. “Recent development in extreme weather forecasting.” *ECMWF Newsletter* No. 107: 8-17.