Processing of Observations from Radiosondes and Upper Air Wind Systems and related or unique surface synoptic observations

at the

Canadian Meteorological Centre

1. Introduction

The basic principles of sounding the atmosphere with radiosondes have changed very little since their inception in the middle of the 20th century. The extraction of observations of pressure, temperature, humidity and winds has progressed through manual methods to more automated ones. Even the alphanumeric code and transmission of the reports, containing the observations, have remained relatively stable over the years.

The number of reporting sites, at any one time, is usually about 750 and has probably never exceeded 1000. Although the numbers are miniscule with respect to the voluminous profiles from satellites, their observations still serve as an important source of data at a high vertical resolution and for comparison and calibration of other vertical profiles or single level data. They also play a vital part in the verification of numerical forecasts and determining global upper air climatology, with at least one but often two soundings per day at the main synoptic hours of 00 or 12 UTC.

Sensing of the observed parameters is usually of the order of seconds, so that since WMO has developed criteria for defining significant levels in a sounding, averaging occurs and the transmitted observations are only a small subset of the raw ones. Apart from winds, the other parameters lend themselves very well to the use of the hydrostatic equation in computing geopotential heights and in a preliminary assessment of their quality. It is the generally high quality of the observations that has also enhanced their usefulness. Even if this is not the case, monitoring their quality over a longer time than for any other data and advances in sensing technology have paid dividends in improving the data.

The following sections describe the peculiarities of the reports, decoding, preliminary checks on the quality of the observations, consolidation of a unique profile, comparison with background fields and assimilation. They end with a description of monthly monitoring, which helps to assess the performance of the Canadian network of stations, in particular, and of global ones.

2. Automated Processing

For many reasons, historical and technical, the transmitted observations from one sounding appear in several bulletins. As a typical sounding takes 100-120 minutes to reach termination, the launch of a radiosonde may be as much as 45 minutes before the usual nominal hour of 00 or 12 UTC. These facts and various thermodynamic constraints, including the main one due to the hydrostatic equation, require different processing from that for most other in-situ observations. The general aim is to combine the observations into a unique profile.

2.1 Coded Reports

Global soundings have multiple permutations and combinations of parts to their reports, reporting practices and types of radiosondes. Each part has an alphabetic identifier of A-D and is usually in a specific bulletin for that part from several stations. Depending on the reporting practices, observations may constitute WMO TEMP and/or PILOT reports, with each possibly being qualified by SHIP or MOBIL and TEMP by DROP. While TEMP reports contain all observed parameters, PILOT ones only have winds. In each case, parts A and B contain observations from the surface to 100 hPa and parts C and D include those above 100 hPa.

Generally, parts A and C contain data at standard or mandatory levels, while those at significant levels are in parts B and D that may not always form part of the global exchange of data. While altitudes of significant levels for temperatures are always in units of pressure, those for winds may be pressure or multiples of 300m (~1000 ft) or 500m. Wind speeds may also be in m/s or knots, which demand a factor for conversion into the MKS units of the data bases.

Common identifying information in the reports is only the day of the month, whose code determines the units of wind speed, and the rounded hour of the observation. By exception, this hour may be rounded up to the next hour, if the time of launch is no more than 45 minutes before the rounded hour. Depending on the type of station, its identifier is either a WMO block and station number or the platform’s identifier, with the latter’s associated coordinates and an elevation for a MOBIL station.

2.2 Decoded Data

The multiple parts and their possible incoherent arrival suggest storing the reports initially in a similar fashion, as blocks within a record. The decoder searches an ASCII file of stations, in order to determine the coordinates and elevation of known ones with a block and station number. It takes the rounded hour as the time of the observation and calculates the stored time of delay as the difference between the time of arrival of the first part and this hour. It also stores the time of launch and details of the type of radiosonde, most commonly in TEMP part B, in a separate block from each one containing the observed parameters from an individual part. In order to associate these parameters with their type of level, the decoder adds their vertical sounding significance.

One station's record in the data base contains all parts associated with each WMO type of code, so that TEMP and PILOT reports are separate records. A record also contains, in separate blocks, any corrected or repeated parts. Unlike most other observations, repetitions are common and do not necessarily contain exactly the same information as a previous part and corrections may expand existing observations.

Such an expansion is particularly common with Canadian observations, because the elapsed time of a sounding tends to be in conflict with the demands of obtaining as much data as possible, before beginning the regional analysis. This conflict has always existed, because slower manual extraction of data demanded a separate RADAT report, which commonly only had observations at mandatory levels and no higher than that at 500 hPa. Software still contains remnants of code to treat such a report. Since the time of observation may be the same within an extreme window of 74 minutes after a short sounding, another potential conflict is a possible second release, the observed parameters of which may only differ from the original ones, but which do not qualify as corrections.

Apart from these conflicts and multiple reports, whose resolution awaits later processing, the decoder performs relatively little quality control outside checking for adherence to WMO codes. It checks for physical and climatological limits and flags the parameter, if it fails, and writes the coded report to a log and does not store it, if it fails too many of these limits or violates too much of the WMO code.

2.3 Merging Data

As a preliminary step towards creating a unique profile, there is a gathering of all reports for all stations, which have a unique location or identifier, within the time window of the analysis. This includes any SYNOP ones for similarly identified stations, so that they may form part of that profile. A filter restricts how many reports of a particular WMO type of code for an individual station pass it. The number is usually one, with the preference going to that closest to the centre of the window, although the criterion for SYNOP reports is more restrictive, so that none of these for a certain station may actually pass through it. Another possible result is that the times of the SYNOP and chosen assembly of parts may be different.

Concurrently, the latest correction of any part takes precedence over a previous one or the original, but any repetitions fall by the wayside without any comparison. This means ignoring any one of these that may contain more data, but are still compatible with the rest and may only be the result of a peculiarity in the usual functioning of equipment. At this stage, no resolution of any ambiguity of a second release having the same time of observation as the original one is possible and even if its time does differ, the first filter may exclude a probably more complete sounding, because its time may not be the closest to the centre of the time window. The final result after these filters is an ensemble of reports that serve as input to the next stage, where a more comprehensive check of the quality of the observations occurs.

2.4 Deriving Data

In spite of the previous filters, the ensemble of reports most likely still contains redundant and missing data, which may require, respectively, elimination or replacement in the final unique profile. The first ones are useful in comparing data at coincident levels and keeping the parameters with the best quality, while the missing data provide an opportunity to replace or derive them from other parts, usually by interpolation.

While interpolation may not be either desirable or necessary for all applications, it is possible because the WMO definition of significant levels permits it between them. If there is none at all or they are absent between mandatory levels or, as a precaution, there is more than one mandatory level between two significant levels, no interpolation occurs. In this regard, any reported tropopause or level of the maximum wind serves as a significant level, despite only appearing in parts A and C of the coded reports. Another constraint on interpolation of an observed parameter is the presence of one, which has already failed previous checks and other successful ones at other levels can not compensate for it.

One of the first checks here is whether it is desirable to assimilate the surface wind. This occurs by verifying the status of a bit in a dictionary of upper air stations. If the bit is on, it prevents interpolation between the surface and the first significant level. Bits set by the decoder or set by previous checks on winds also prevent the associated parameters from participating in further checks. These are on lapse rate, wind shear and hydrostatic consistency. They all check the appropriate observed parameters, three levels at a time, with various conditions and thresholds specific to the parameter. Finer checks occur for winds, if there are more levels of them than normally exist in TEMP part A and they undergo a final check, if interpolation has occurred with respect to height.

Before this interpolation can occur, there is a need to express pressure in terms of geopotential height and vice versa. In the first case, the use of the hydrostatic equation, in order to produce reported geopotential heights at mandatory levels, satisfies that need for these levels, but as long as a system of assimilation requires these heights or winds, reported with respect to height, derived at or interpolated to fixed pressure levels, there is a need to compute these heights at these levels. In the second case, a table derived from a consensus of ones of pressure versus height, used mainly by tropical countries, whose stations regularly report winds with respect to pressure at mandatory levels, but no associated geopotential heights, and with respect to heights at significant ones, helps to interpolate these latter winds to the fixed pressure levels.

The derivation of geopotential heights involves a matrix solution to a set of simultaneous equations, using as much as possible the full profile of heights and virtual temperatures. Similar criteria apply in this derivation as the interpolation of other parameters, but virtual temperatures have additional ones. A doubtful temperature or one ≥ -15C and without a dew-point depression prevents the computation of a virtual one and thus an interpolation of heights between two mandatory levels with reported heights, which the preliminary hydrostatic check may not have already found to be doubtful or which are missing. For a colder temperature, without a dew-point depression, the virtual one equals this ambient one.

It is possible that during this derivation or in the earlier hydrostatic check, a distinction between an original and a second release, having the same times of observations, may be feasible. As the surface data in a TEMP report most often take precedence over those in a SYNOP one, with the same time of observation, it may also be desirable to formulate an algorithm to check the hydrostatic consistency of the surface pressure. If it is not consistent, the SYNOP report may have a more consistent one, taking into account a possible difference in elevation, with which to replace it and probably to flag all occurrences of pressure.

Another operation is the correction of geopotential heights for certain types of radiosondes that suffer from undesirable effects on observed ambient temperatures by either solar radiation or infrared cooling, especially in the upper troposphere and lower stratosphere. Since an integration of the hydrostatic equation derives these heights, the variance of their differences from an analysis or background field invariably increases more rapidly and the associated bias is more monotonic, with respect to decreasing pressure, than corresponding ones for temperature. Thus it is more desirable to correct heights rather than temperatures with statistics, gathered from long term monitoring for at least VIZ and the Chinese Shanghai radiosondes still in use. These statistics originally come from the ECMWF for 16 mandatory pressure levels with respect to different sun angles. Like other parameters, they undergo interpolation to any more numerous levels.

There are also stations whose data have systematic discrepancies or ones caused by various malfunctions. Instead of correcting their delinquent parameters, these stations appear in an adjustable black list with an appropriate indicator for the parameter in question, so that it may in turn set a bit for that parameter in the final unique profile.

3. Data Assimilation

The unique sounding now exists as one record, with at least one block of data and an associated one of flags, indicating previous processing and any failures. Except in the rarer cases, when only parts C or D arrive and there is no accompanying SYNOP report, the surface data are in another block. If the chosen SYNOP data have a different time of observation from the sounding, they appear in another record for such a station with their associated block of flags.

The upper air data appear at all available mandatory levels, above the elevation of the station, and are similarly at any intermediate fixed levels, for which there are any original, interpolated or derived data. The data are missing, where constraints of interpolation or derivation prevented it, and have appropriate flags when they successfully satisfy the constraints.

A comparison of these data with equivalent ones, at the closest time step to the time of the observation from a background field that is a forecast by a global numerical model from the previous objective analysis, serves to identify any gross errors in the data and flags them as rejects. A similar, but separate, background check occurs for the SYNOP or surface data. These data and those at lower levels of a sounding undergo further checks with respect to the orography and land-sea mask used by the model. The criteria may flag certain data as being unrepresentative for further assimilation. Irrespective of any flags, this comparison also provides a residual or difference between each element of data and its interpolated value from the background field.

A more severe test of quality follows as variational quality control to give a final weight to the data and reject any whose probability of being a gross error is greater than 0.75. This precedes the calculation of increments at the location of each element. Their algebraic addition to the interpolated values of the background field has as its result an objective analysis. Then, by putting these residuals and increments in separate blocks and adding them to those of the associated data and their flags, the whole ensemble can act as a useful diagnostic tool for the behaviour of observations from stations over an extended time.

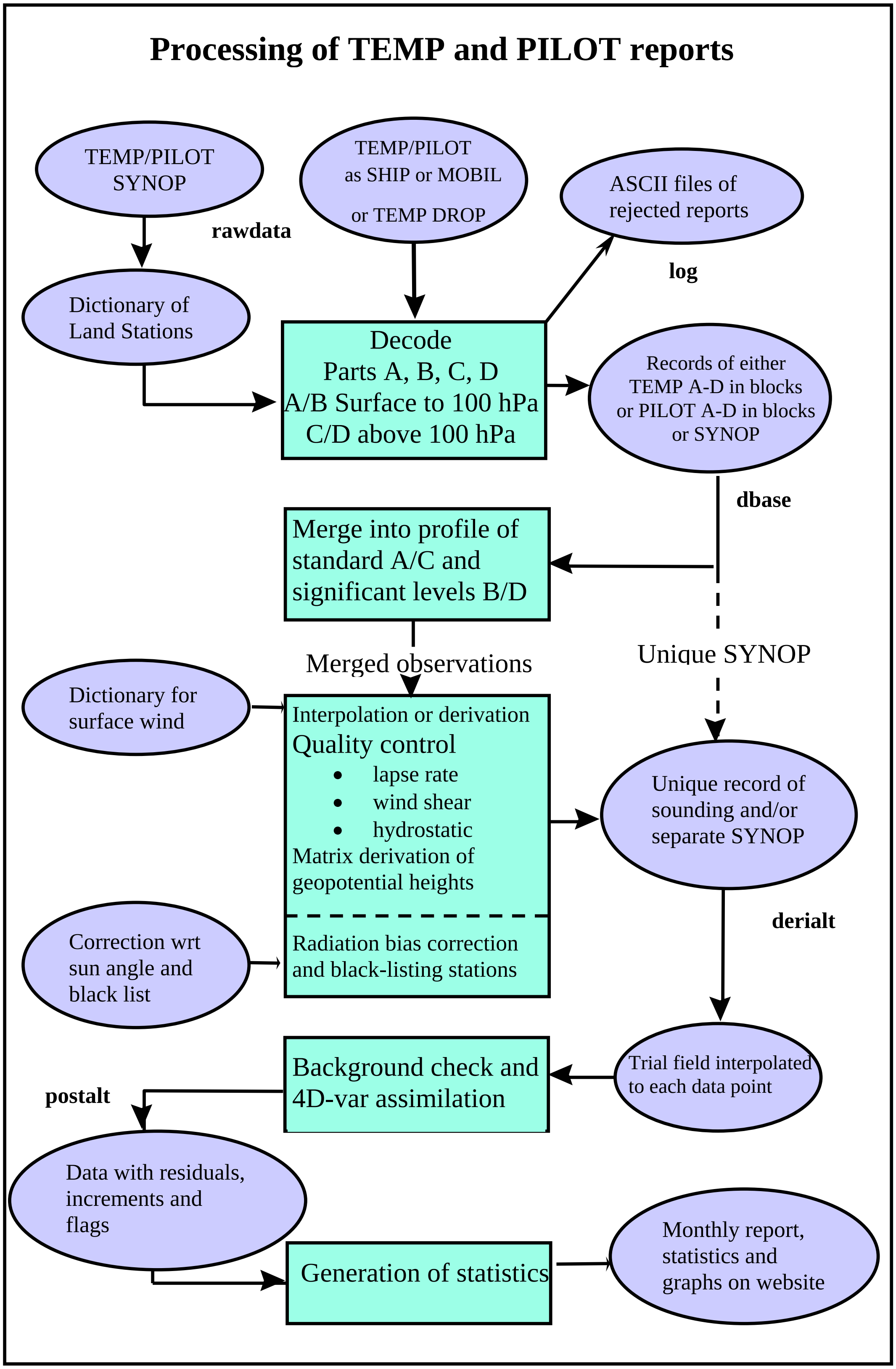
4. Data Monitoring

As numerical weather prediction improved over time, the background fields for all observed parameters, except humidity, in profiles from radiosondes have become reliable enough to assess the overall performance of these observations. Derived statistics from their residuals over a period such as a month may reveal systematic biases or large variances. These may have their origin in problems with equipment or procedures, which the producing country may be able to correct, if so informed.

This information now exists as six-monthly lists of suspect stations, compiled by the ECMWF, which the CBS of the WMO recommended as the lead centre for radiosondes in 1988. The CBS has also formulated criteria for an exchange of monthly lists between major GDPS centres. These criteria include thresholds on the variance of the residuals of geopotential heights, weighted for mandatory levels from 1000-30 hPa, and of the vector wind from 1000-100 hPa and the bias of the wind direction, constrained by thresholds on the standard deviation at all mandatory levels between 500 and 150 hPa.

Graphical representations of the statistics on residuals may include their temporal trend and geographical extent and, with similar ones for the increments, a combination of them in the vertical may also prove to be useful. The observations themselves and associated flags are important in verifying forecasts and averaging the number of soundings over a month and plotting their global distribution also form a part of a monthly exchange of information.

Even daily monitoring of soundings depends on the results of processing and assimilation of the data from radiosondes. Within the Canadian network, observations of wind are particularly vulnerable to external influences on whichever navigational aid they use. Their results from this process may determine that a change to an alternative navigational aid is necessary.


Processing of TEMP and PILOT reports
TEMP/PILOT SYNOP
TEMP/PILOT as SHIP or MOBIL or TEMP DROP
ASCII files of rejected reports
rawdata
log
Dictionary of Land Stations
Decode Parts A, B, C, D A/B Surface to 100 hPa C/D above 100 hPa
Records of either TEMP A-D in blocks or PILOT A-D in blocks or SYNOP
dbase
Merge into profile of standard A/C and significant levels B/D
Merged observations
Unique SYNOP
Dictionary for surface wind
Interpolation or derivation
Quality control
lapse rate
wind shear
hydrostatic
Matrix derivation of geopotential heights
Radiation bias correction and black-listing stations
Unique record of sounding and/or separate SYNOP
Correction wrt sun angle and black list
derialt
Background check and 4D-var assimilation
Trial field interpolated to each data point
postalt
Data with residuals, increments and flags
Generation of statistics
Monthly report, statistics and graphs on website