U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANOGRAPHIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION

OFFICE NOTE 423

COMPLEX QUALITY CONTROL OF RAWINSONDE HEIGHTS AND TEMPERATURES AT THE NATIONAL METEOROLOGICAL CENTER: QUALITY CONTROL OF MANDATORY LEVEL DATA

William G. Collins
Lev S. Gandin
NATIONAL METEOROLOGICAL CENTER
(NOW NCEP/ENVIRONMENTAL MODELING CENTER)

JUNE 1994
(as Office Note MAY 1998)

NOTE

This paper was originally written for submission to one of the AMS journals, but it was realized that it was much too long for that purpose. Much time lapsed and much thought, primarily by Lev, went into how to rewrite the paper to make it publishable. In the end, Lev became ill and no paper was produced. Also, in the interim, the complex quality control code was completely rewritten, and so this paper became grossly out of date. However, it is felt that this paper is valuable to make available to those interested, because it contains useful historical information and also contains much of the views of Lev's, generally regarding quality control. (Note also that at the time, NCEP did not exist.)

ABSTRACT

The Complex Quality Control of rawinsonde data on mandatory level Heights and Temperatures (CQCHT), consisting of the hydrostatic check, the baseline check, statistical checks, and an advanced Decision Making Algorithm (DMA), was designed at the National Meteorological Center in Washington and implemented into operational quality control system in 1991. The principles of the operational quality control in general and of the CQCHT approach in particular are discussed in the article. The DMA reaction to errors of various origin is described in detail and illustrated by numerous examples. Some statistics on the CQCHT performance are presented and compared with those for the previously applied CHQC algorithm (Collins, Gandin, 1990).

1. Introduction

As is well known, some meteorological data received at prognostic centers are distorted by so-called rough errors. Such errors may originate in the course of measuring, processing or communicating the data. Although comparatively rare, rough errors may lead, particularly in data-poor regions, to substantial errors in analyzed fields and, therefore, in predicted ones. That is why some special procedures are performed, both manually and automatically, at every prognostic center in order to get rid of rough errors. These procedures are usually referred to as the quality control (QC) of operational meteorological information.

The necessity of an automatic QC performed by computer was recognized at the beginning of the numerical weather prediction era (Gilchrist and Cressman, 1954), and the first such methods were proposed and applied soon after that (Bergthorsson and Döös, 1955, Bedient and Cressman, 1957, Staff Members, Joint Numerical Weather Prediction Unit, 1957). There was, however, little improvement of QC methods during several following decades, because the most important task was to improve existing numerical weather prediction models and data assimilation systems, and also because the QC design was considered by many specialists as a purely technical task having nothing to do with science. As a result, the QC systems in operational use at major prognostic centers, including NMC, were due to tradition rather than to logical reasons.

It was recognized as recently as 1988 that the operational QC system at NMC needed substantial improvement. An important decision was made by W. Bonner, then the NMC Director, to begin the design of the new NMC QC system from scratch rather than to try to improve the existing system. Leading principles for the new system were agreed upon after thorough discussions (Julian, 1989; see also Collins and Gandin, 1990). Only a few of them, essential for the further discussion, will be mentioned here.

The main principle was (and continues to be) that the new QC system should be as much automated as possible. Experience shows that severe time limitations under operational conditions make a subjective QC by human specialists very difficult even if they limit themselves to quality control of traditional data, like those from rawinsondes, and only over a limited area, like North America. At the same time, the improved performance of the NMC medium-range forecast (MRF) model made it necessary to perform the QC worldwide, particularly because the frequency of rough errors over many regions of the globe was (and continues to be) much higher, and/or the station network much sparser, than over United States and Canada. Even more importantly, new kinds of measurements, most notably the satellite soundings, provide us with large amounts of data, and one cannot even think about controlling the quality of all these data manually. The automatic QC by a computer remains the only practical alternative if we want to exclude rough errors from all kinds of operationally available information.

This does not mean, however, that human specialists will not be involved in the operational QC. Just the opposite is true. The more sensitive is an automatic QC method, i.e., the smaller are rough errors it is capable to detect, the higher is the probability that it will sometimes be unable to decide what to do with one or another suspected datum. It may even happen, though rather rarely, that an automatic QC will make a wrong decision concerning rejection or even correction of one or another suspected datum. It is necessary therefore to make the QC algorithm capable of diagnosing all cases when its decision was questionable. The information on each such case should be automatically transmitted to a specialist for subjective consideration together with the information on cases when no definite decision could have been made automatically. It is up to the specialist to decide what should be done in each such case. However, every result of a specialist's action should be immediately subjected to the automatic QC and accepted only if the modified report passes the QC test. Otherwise, an alternative action can be proposed and tested, etc.

The outlined procedure is known as the interactive QC (IAQC). It is clear that the number of reports needing to undergo the IAQC is very small as compared not only with the overall number of reports, but also with the number of confident rejections and corrections performed by the QC algorithm. It is possible therefore for human specialists to get involved in all such cases. It also happens sometimes that a human specialist suspects some information not caught by the automatic QC, or wants to keep the information as it has been reported despite a confident rejection or correction by the QC algorithm. This can be done within the IAQC framework as well.

Another principle agreed upon in 1988 was that QC algorithms should be observation system dependent. In other words, the sets of QC checks for one or another parameter, as well as decisions resulting from the checks, should be, generally speaking, different not only for different meteorological parameters, but also for the same parameter observed by different systems. There are two reasons for this dependence. First, some specific kinds of errors may exist for one observation system and not exist for another one. Secondly, some checks, sensitive enough for one observation system, may be less sensitive or just impossible for another system.

For example, the upper-air winds may be those observed by rawinsonde, by aircraft, by satellite, or by profiler. The aircraft winds are often distorted by position errors, which never occur to, say, profiler winds. As to checking methods, the vertical statistical interpolation check proved to be sensitive for rawinsonde and profiler winds, while not being applicable to aircraft or satellite winds. These examples illustrate the fact that separate QC algorithms are needed for data from different observation systems, although some general principles, like that of maximal automation, should be universal.

The last of the principles to be discussed here is that the QC algorithms should follow the so-called complex quality control (CQC) approach (Gandin, 1969, 1988), as opposite to sequential (or hierarchical) approach, which was traditionally used in automatic QC methods at that time.

According to the sequential approach, the least sensitive check, capable of detecting only very large errors, is applied first, and data suspected by this check are flagged as wrong information not to be used by the data assimilation system (we will refer to such decision as a rejection, although "physically" every such datum is not excluded from the data set). Remaining information is then subjected to another, more sensitive check, which additionally rejects some less erroneous data, and so on.

In contrast to this, no information is rejected (or corrected) by its CQC until it undergoes all checks. The CQC algorithm thus consists of two major parts: the checks and the decision making algorithm (DMA). Results of each check are expressed not by flags, but quantitatively, by so-called residuals. The DMA then analyses all residuals. If none of them is large (by absolute value), as is the case with an overwhelming majority of reports, then the DMA concludes that there is no reason to suspect any error. If, however, at least one residual is large, then the DMA analyses the pattern of various residuals, trying to locate the error (or errors), to explain its origin and, if possible, to correct erroneous datum. Corrections made by the DMA are directed towards the restoration of correct values which were distorted by errors made while processing the data or originated on communication channels. It may be mentioned in this respect that the larger an error, the higher is the probability that it was not caused by the measurement itself, but originated later, in the course of processing or communicating the data, and the datum in error may be corrected instead of being rejected.

The CQC approach was applied in the former USSR for QC of rawinsonde height and temperature for comparatively long time (Antsipovich, 1980, Aldukhov, 1982) showing its substantial advantages over a sequential approach. It was natural therefore to apply the CQC approach in the new NMC QC system.

Fig. 1.1 presents schematically the transformation of the QC system at NMC during recent years as a result of the work on design, testing, implementation and monitoring of new QC methods, performed at the NMC Development Division (DD) under general supervision by E. Kalnay (the scheme also contains some "extrapolation" in time reflecting DD plans for a near future). As may be seen from the scheme, the first of these methods was the Comprehensive Hydrostatic Quality Control (CHQC) of rawinsonde data on mandatory level heights and temperatures (Collins and Gandin, 1990, hereafter referred to as CG90). Although applying only one, hydrostatic, check, the CHQC algorithm analyzed residuals of this check for as many as three layers (i.e., four levels) before making any decision. Correspondingly, the CHQC included a comparatively advanced DMA, the first one ever designed at NMC.

Our initial intention was first to enrich the CHQC by adding some other checks to the hydrostatic one and only then to implement the resulting algorithm. However, the results of CHQC testing were so encouraging, and the need to improve the QC at NMC so demanding, that the decision was made to first implement the CHQC algorithm as it was. It was soon complemented by a QC of significant level temperatures performed by a complex of hydrostatic and vertical interpolation checks using already quality-controlled mandatory level heights and temperatures (Collins, 1990).

The CHQC was in operational use at NMC for about three years. It proved to be very productive not only in its operational mode. Quasi-operational monitoring of the CHQC outputs, which was performed by DD specialists, allowed us to discover many problems with operational data arriving at NMC and to resolve some of these problems (Gandin, Morone, Collins, 1993). The CHQC algorithm is well documented, and the code is now used at many centers both in this country and abroad.

As mentioned in CG90 (see also Table 8.1), the CHQC DMA was able to automatically correct about 50% of errors suspected by it. It also submitted all its information on remaining suspicions to the Senior Duty Meteorologist (SDM) or to other specialists at the NMC Meteorological Operations Division (MOD). That marked the beginning of a scheduled interaction between automatic and manual QC at NMC, which later resulted, as shown in Fig. 1.1, in the IAQC system designed by J. Woollen (Collins and Woollen, 1993).

As our monitoring of CHQC outputs demonstrated, it would not be difficult for a specialist to make a proper decision in an overwhelming majority (more than 80%) of cases only suspected but not decided upon by the CHQC DMA. Nevertheless, the MOD treatment of these outputs proved to be not very successful, and the absence of IAQC equipment was not the only obstacle. A deeper understanding of the CHQC DMA was needed, as well as a deeper interest in data outside USA, where rough errors occur much more often. Still, we believe that an involvement with the CHQC outputs was useful for MOD specialists as a preparation to much more complicated operations with the IAQC.

At the same time, the results of our monitoring of the CHQC performance demonstrated again that a more advanced QC, containing other checks in addition to the hydrostatic one and thus capable of automatically correcting much more hydrostatically detected errors, should be designed and implemented as soon as possible. The design of the new QC, called the Complex Quality Control of rawinsonde Height and Temperature (CQCHT) began in 1990. After extensive testing and improvements, the CQCHT became operational in November 1991, replacing the CHQC.

There were several other improvements of the NMC QC system during recent years. The most important of them was the design and implementation by J. Woollen of the Optimum Interpolation Quality Control (OIQC) (Woollen, 1992). As shown in Fig. 1.1, the OIQC replaced the so-called gross check and buddy check which were components of the former, sequential QC system. As a result, none of the components of the former NMC system, except the manual non-interactive QC, continues to operate now, all components are new. One may add that the OIQC is, like the CQCHT, a kind of CQC.

The CQCHT DMA is much more productive than the CHQC DMA was because every CQCHT suspicion and DMA decision is based on a variety of checks. As a result, the CQCHT is capable not only of detecting a larger number of errors, but also of correcting, entirely automatically, a much larger proportion of correctable errors than was the case with CHQC. In particular, the CQCHT DMA automatically corrects (or decides not to correct) an overwhelming majority of those suspected errors, information on which would otherwise only be transferred by the CHQC to the NMC MOD specialists for their help.

Even more importantly, the CQCHT, unlike the CHQC, reacts not only to errors originating in the course of processing the observation results or afterwards, in the course of communicating them, but also to so-called observational errors that originated before the processing began. Naturally, the CQCHT DMA is unable to correct observational errors (except those in surface-air pressure measurement), deciding instead to reject such erroneous data or to assimilate them with diminished weights.

The CQCHT DMA produces a special file for interaction with human specialists, the so-called SDM file, just as the CHQC DMA did. Superficially, these two kinds of SDM files look analogous, but the essence is quite different. The CHQC SDM files contained only those cases when the DMA did not make its decisions and requested human help. In contrast to this, an overwhelming majority of cases included into the CQCHT SDM files are those for which the CQCHT DMA did make all its decisions, but it concluded that a specialist may wish to change some of them, particularly if there exists some additional information that was not available to the CQCHT.

The involvement of human specialists in the interaction with the CQCHT is therefore much less "automatic" and more challenging than it was the case with the CHQC. It requires detailed knowledge and good understanding of the CQCHT DMA by the NMC senior duty meteorologists (SDMs) and by other specialists of the NMC Meteorological Operations Division (MOD) interacting with it. To achieve such knowledge and understanding was (and continues to be) a rather difficult task, because the CQCHT is a very advanced and complicated algorithm, and also because problems like this never appeared before.

Every effort has been therefore undertaken in order to provide the MOD personnel with necessary training and consultations. An extensive NMC Office Note on the CQCHT (Collins, Gandin, 1992) has been written especially for this purpose. It could also be used, and actually was already used, by those of our colleagues at various institutes in this country, as well as abroad, who wanted to design and implement analogous QC algorithms.

This article is also devoted to the CQCHT, but its aim is quite different. Since the CQCHT began to operate at NMC and was briefly described by us and by E. Kalnay at several national and international meetings, many meteorologists, oceanographers and scientists in related fields, including those not directly involved in the QC of operational information, expressed their desire to learn more about the CQCHT methodology and, generally, about the CQC approach and Decision Making Algorithms. This article has been written in order to meet this demand.

2. Some Definitions

This article, like any scientific paper, uses, so to say, its own terminology: many terms in the article are new, and the meaning of many others is different, to one or another extent, from that assumed in other publications. The latter is particularly true with terms that have no commonly accepted meaning and are often given different meanings by different specialists. It is worthwhile therefore to give exact definitions of basic terms as they will be understood in this article.

Perhaps no other term, used in this article, may be, and actually is, understood in such variety of meanings, as is the case with the term *quality control.* This is partly due to different meanings of the word control in different languages. To control means something like to manage, or to rule, in English, while in other languages the meaning is rather to check the reliability of something, or to verify it. Consequently, the term quality control is often used, at least in this country, in a very wide sense, as including all actions connected, directly or indirectly, with the quality of some objects and/or operations. From that point of view, for example, detection and correction of computer failures at a weather prediction center is a part of the quality control operations at this center, as is the software for collection and storage of information used in its data assimilation system.

Opposite to this, the term quality control in this paper, like in most other publications about the automatic quality control of meteorological data, has a rather narrow and concrete meaning as a set of procedures used in order to detect and, if possible, correct the so-called *rough errors* in meteorological data.

This leads us to the term *error*, which is also used, particularly in meteorology, in a variety of meanings. A difference between objectively analyzed value and observed one is often called the *analysis error*, and the difference between predicted and observed values is referred to as the *forecast error.* One can even find a term *climatology error* understood as the difference between the climatological norm and observed value (in other words, as minus anomaly). Not discussing here this kind of wording, it is necessary to stress that in this paper, an *error* is understood in quite different, more natural way as a difference between the *reported* value of a meteorological parameter and its *actual* value.

It is important to distinguish between random, rough and systematic errors in meteorological data. *Random errors* are inherent in all data and caused by a variety of factors, like (non-systematic) measurement errors or small-scale turbulence. Being more or less independent from each other at different points and times, they form what is called a random noise in the data. It is, of course, impossible to correct random errors, but it is important to properly take into account the noise level, usually characterized by the root mean square (RMS) random error, when performing many operations with the data, including their quality control.

Unlike the random errors, the so-called *rough errors* in meteorological reports occur comparatively seldom; the majority of reports don't contain any rough errors. Each rough error has its definite cause which may happen in the course of measurement, processing, or communicating the data. It is the task of the quality control to detect each rough error in arriving reports and, if possible, to correct erroneous data. Otherwise, it must mark the data for rejection from the operational data assimilation system or for assimilating them with smaller weights. Certainly, some errors of this kind may be rather small. It is, however, impossible even to recognize any such error, unless its absolute value substantially exceeds the noise level. That is why the errors dealt with by QC have been given the name rough errors.

As to the errors of the third category, the *systematic errors*, they are usually small but, unlike rough and random errors, they persist in time. Such errors may result from some insufficiencies either in measurement devices or in procedures designed to take care of these insufficiencies. Substantial averaging in time, e.g., over a month, is needed in order to detect systematic errors. This process, known as the *data quality monitoring* (DQM), as distinct from the (operational) data quality control, will be not dealt with in this paper devoted to operational QC, which deals exclusively with rough errors It should be mentioned, however, that the application of a CQC approach, like that described in this paper, would result in substantially more productive DQM methods than those applied nowadays.

Depending on their origin, rough errors may be divided into three categories: observational, computational and communication-related errors (see Fig. 2.1). *Computational errors* are those originating in the course of processing of the sounding data, particularly in the computation of mandatory surface heights at the station (or elsewhere). All rough errors made before this processing began are called, in this paper, *observational errors*, and all rough errors committed after the processing ended are called communication-related errors (or, simply, *communication errors*). The category of observational errors thus contains not only measurement errors, but also those made at the station when the rawinsonde signals were received and put into the processing. As to the communication errors, they include all rough errors made when coding reports for their transmission and putting them into communication lines, when the reports follow the communication lines and when they are received, sorted and decoded at the NMC. Rough errors made in the course of subjective QC may be also included into this category.

It should be mentioned in this respect that what we call *reported values* in this paper are actually the values entering the quality control algorithm. Due to the influence of various communication errors, the reported values, as they are understood here, may differ from values actually reported by the station.

As already mentioned, no decisions concerning any reported datum are made by the CQC algorithm before the quality-control by a series of more or less independent methods, called the *checks*, or the components of this CQC, is applied to the report. Each check results in its *residual*, quantitatively reflecting the degree (and sign) of the inconsistency in this datum discovered by this particular check.

Most of the CQCHT checks, namely, those called in this article *statistical checks*, deal not with reported values themselves, but with their deviations from the so-called *forecast first guess*, usually a 6-hour numerical forecast for the time under consideration. There exists no commonly accepted name for these deviations, some authors call them innovations (e.g., Daley, 1991), but most often, they (with the opposite sign) are referred to as first guess errors. We will call these deviations the *increments*, as proposed by Thiebaux and Pedder (1987).

The simplest among statistical checks is simply based on the value of the increment and is called the *incremental check*. The residual of this check is just the increment itself. Other statistical checks include the interpolation of increments either from neighboring levels (*vertical check*) or from neighboring stations (*horizontal check*). The residual of each such check is the difference between the increment at the point under check and its interpolated value.

Other components of the CQCHT, the *hydrostatic check* and the *baseline check*, may be called quasi-functional checks (as opposite to statistical ones). The hydrostatic check uses the hydrostatic equation (more exactly, the so-called barometric equation) for each layer between two adjacent mandatory surfaces to check the consistency between heights and temperatures at this pair of surfaces. The residual of the hydrostatic check is the difference between the layer thickness computed from two heights and the same thickness computed from two temperatures. The baseline check is also based on the hydrostatic equation, but in this case the equation is applied to another pair of levels: the station level (more exactly, the level of its surface observations) and the middle of the layer between two lowest mandatory surfaces. The *baseline residual* is the difference between the station elevation above the mean sea level (taken from the NMC dictionary of stations) and the same elevation computed by the baseline check.

Due to their quasi-functional nature, both hydrostatic and baseline check are, generally speaking, substantially more sensitive in detecting rough errors than statistical checks are. That is why all CQCHT suspicions of computational and communication errors are made on the basis of hydrostatic and baseline residuals. At the same time, statistical checks play a very important role when they are performed in a complex with quasi-functional checks, as is the case in the CQCHT.

The CQCHT algorithm thus consists of two major parts: the checks and the Decision-Making Algorithm (DMA). The CQCHT DMA is an advanced, logically complicated code which analyses residuals of all checks and makes all decisions. Programs like this are often referred to as expert systems, or even as artificial intelligence. No serious objections can be made against such terminology, although it sounds like an advertisement rather than science. The point is, however, that this wording might lead to an impression that the DMA design is a prerogative of specialists in expert systems, rather than meteorologists (or oceanographers). Numerous experience, both positive and negative, in the QC design at NMC proves quite convincingly that a good meteorological background and understanding is absolutely necessary for success of this design.

3. CQCHT Checks

Table 3.1 contains basic information about each check used in the CQCHT algorithm.

The hydrostatic check is applied to each layer between two neighboring mandatory isobaric surfaces with *complete* information, i.e., with none of heights and temperatures at the two levels missing in the report. It is based on so-called *hydrostatic redundancy* in rawinsonde reports, that is, by the fact that both temperature T_i and height z_i of each mandatory level p_i are reported, while the barometric equation (also called the hypsometric equation)

$$z_i - z_{i-1} = A_{i,i-1} + B_{i,i-1}(T_{i-1} + T_i) \tag{3.1}$$

is, in the absence of rough errors, obeyed with high accuracy for each pair of mandatory levels. In (3.1),

$$A_{i,i-1} = \frac{R}{g} T_{00} \ln\left(\frac{p_{i-1}}{p_i}\right); \quad B_{i,i-1} = \frac{A_{i,i-1}}{2T_{00}} \tag{3.2}$$

where g is the acceleration of gravity, R is the gas constant for the air, and T_{00} is the Kelvin temperature of 0°C, so that T in (3.1) is in °C. The equation (3.1) follows from the hydrostatic equation under the assumption that the temperature varies linearly with ln(p) within the layer; its left hand side is the layer thickness computed from the heights of its boundaries, while the right hand side is the same thickness computed from temperatures at the boundaries.

It should be mentioned that the applicability of the barometric equation (3.1) to reported temperatures and heights has nothing to do with the approximate nature of this equation, i.e., with the fact that the hydrostatic equation is an approximate form of the equation of motion in projection on the vertical. Equations similar to (3.1) are used to compute the mandatory heights while processing the rawinsonde data at stations (or elsewhere). The redundancy expressed by equation (3.1) is therefore of communicational rather than physical nature. An analogous redundancy might exist, e.g., in rawinsonde wind data, if, say, the zonal wind component were computed at stations and included into reports along with wind speed and direction. That would substantially increase the possibilities of the rawinsonde wind quality control because it will be able to detect all rough communication errors.

Due to the communicational nature of hydrostatic redundancy in rawinsonde reports, the hydrostatic check does not react at all to observational errors; the equation (3.1) holds if there were no communication or computation errors. As discussed in some details below (see Section 7), this fact is extensively used by the CQCHT DMA for detection of observational errors.

Each of the coefficients A and B (3.2) in equations (3.1) depends only on two pressures and is therefore constant for a given layer. Table 3.2 contains these coefficients for all *elementary layers*, i.e., layers between neighboring mandatory surfaces. According to (2), both A and B are *additive*: if a layer consists, say, of two elementary ones, then each of the coefficients A and B for this layer is just the sum of two values for the elementary layers, and so on.

There are several effects causing violations of equations (3.1) or, in other words, leading to their residuals

$$s_{i,i-1} = (z_i - z_{i-1}) - \left[A_{i,i-1} + B_{i,i-1}(T_{i-1} + T_i)\right] \quad (3.3)$$

even in the absence of rough (computational or communication related) errors, namely, non-linearities of the temperature profile (with respect to ln(p)) in the layers, various random errors, and differences between the temperatures in (3.1) (and (3.3)) and corresponding virtual temperatures. In order to be detectable on the background of the noise caused by these effects, a rough error should result in residuals (3.3) substantially exceeding the noise level. This leads to so-called *magnitude conditions*: a report is suspected for *hydrostatic error(s)*, i.e., errors detected by the hydrostatic check, only if at least one of hydrostatic residuals exceeds by absolute value the *admissible residual* for the corresponding layer.

Table 3.3 contains the admissible residuals, used in the CQCHT, for all elementary layers, expressed both in terms of height (3.3) and in terms of temperature

$$x_{i,i-1} = \frac{s_{i,i-1}}{B_{i,i-1}} \quad (3.4)$$

They were specified on the basis of routinely collected statistics on the residual frequency distributions among all hydrostatically not suspected reports.

To achieve a better understanding, a series of what may be called numerical experiments with admissible residuals was also performed. The hydrostatic check was applied to the same set of reports with varying, gradually decreasing admissible residuals. An explosion-like growth in the number of hydrostatically suspected reports took place when the admissible residuals became comparatively small, which clearly demonstrated that many error-free reports began to be suspected as containing hydrostatic errors.

As to the admissible residuals for non-elementary layers, they are computed from those in Table 3.3 using the simplest hypothesis of statistical independence between hydrostatic residuals for neighboring elementary levels, so that, e.g., the admissible residual $s^{adm}{}_{(i+2,i)}$ for a layer consisting of two elementary ones is

$$s^{adm}{}_{i+2,i} = \sqrt{\left(s^{adm}{}_{i+1,i}\right)^2 + \left(s^{adm}{}_{i+2,i+1}\right)^2} \quad (3.5)$$

and so on. If, however, two or more mandatory levels in a row are missing or incomplete in a report, so that the layer between neighboring complete levels becomes rather thick, then the non-linearity of the temperature profile in such a layer can, by itself, cause a large hydrostatic residual (particularly, if the tropopause level is within this layer). This is why the hydrostatic residuals over the *data holes*, containing two or more mandatory levels, are not used by the CQCHT algorithm as means for hydrostatic suspicions; they are just ignored. The CQCHT DMA analyses each report with such a data hole as if it were two separate reports, one below the hole, another above it, treating the level before the hole as the upper level of the "first" report, and the level after the hole as the lowest level of the "second" report

The *baseline check* also uses the hydrostatic equation, but applies it not to layers between mandatory surfaces but to the layer between the station level z_s and the middle

$$\bar{z} = (z_1 + z_2)/2 \quad (3.6)$$

between the heights of two lowest reported mandatory isobaric surfaces p1 and p2 (used independently on whether the temperature at any of them is reported). A linear temperature profile

$$T(z) = T_s - \gamma z \tag{3.7}$$

is assumed for the layer between z_s and $\bar{z}$ with the standard lapse rate

$$\gamma = 6.5\text{x}10^{-3}\ \text{Km}^{-1} \tag{3.8}$$

Under the assumption (3.7), known as that of polytropic atmosphere, the pressure p decreases with height z proportional to $z^{1/c}$ where

$$c = \frac{R\gamma}{g}, \tag{3.9}$$

so that

$$\frac{(z - z_1)}{(z_2 - z_1)} = \frac{(p_1^{\ c} - p^c)}{(p_1^{\ c} - p_2^{\ c})}. \tag{3.10}$$

With the standard lapse rate (3.8), the non-dimensional parameter c (3.9) is equal to 0.190.

We define the *baseline residual* (in terms of the station elevation) b_{zs} as the difference

$$b_{zs} = z_s - z_{sc} \tag{3.11}$$

between the station elevation above the mean sea level, z_s, known from the NMC Upper-Air Station Dictionary, and its value

$$z_{sc} = z_1 + (z_2 - z_1)\left(\frac{b}{a}\right), \tag{3.12}$$

computed by the use of (3.10) from reported values of z_1, z_2 and surface air pressure p_s. In (3.12),

$$a = 1 - \left(\frac{p_s}{p_1}\right)^c \quad \text{and} \quad b = 1 - \left(\frac{p_2}{p_1}\right)^c. \tag{3.13}$$

Alternatively, one may express the baseline residual in terms of z_1, or z_2, or p_s, using equations analogous to (3.11):

$$b_{z1} = z_1 - z_{1c}, \quad b_{z2} = z_2 - z_{2c}, \quad b_{ps} = p_s - p_{sc}, \tag{3.14}$$

where

$$z_{1c} = \left(\frac{bz_s - az_2}{b - a}\right), \tag{3.15}$$

$$z_{2c} = (z_1 - z_s)\left(\frac{b}{a}\right) \tag{3.16}$$

and

$$p_{sc} = p_1\left[1 + b\left(\frac{z_1 - z_s}{z_2 - z_1}\right)\right]^{\frac{1}{c}}. \tag{3.17}$$

Each of the residuals (3.11), (3.14) with an opposite sign is equal to the correction that should be added to reported value of the corresponding parameter (and only to it) in order to make the baseline residual equal to zero.

There are two reasons why it is preferable to use reported heights, rather than temperatures, for the baseline check. First, the temperature at a given level near the ground may be distorted by a small scale disturbance, while such distortions are smoothed to some degree in the course of height computations. Secondly, the mandatory surface heights are, according to existing rules, computed and *reported* even for (some) mandatory isobaric surfaces which actually are under the ground (while their temperature is missing in reports). The polytropic hypothesis (3.7) with the standard lapse rate (3.8) is applied to compute such underground heights. Their use for the baseline check often results in the fact that the extrapolation downwards to the station level turns out to be an interpolation, or even an upward extrapolation.

The admissible residual of the baseline check (in terms of the station elevation) used in the CQCHT, $(b_{zs})^{adm}$, is equal to 40 m. If, however, some other check, or checks, of the same information, particularly statistical checks of (reduced) mean-sea-level pressure, also results in some suspicions, then a half of this value, $(1/2)(b_{zs})^{adm} = 20$ m, is considered sufficient in order to suspect an error.

Having in mind the specific role which the incremental and horizontal checks of the mean-sea-level pressure play in conjunction with the baseline check, it is convenient to call this set of three checks the *baseline checks.*

From the formal point of view, the *incremental check* does not differ from what was called the gross check, it just compares the absolute value of an increment with its admissible value. However, if applied in conjunction with other, quasi-functional checks, as is always the case in the CQCHT, the incremental check is much more sensitive than if it were used solely, as a gross check. Consequently, the admissible residuals of the incremental check, presented in Table 3.4, are rather small as compared to those used in a gross check. Moreover, even halves of values in Table 3.4 are used as admissible residuals if the incremental check is applied to confirm (or deny) suspicions resulting from other checks. This "two-margin" approach is analogous to that used in the baseline check. It is applied to other statistical checks as well.

Both the *horizontal check* and the *vertical check* are optimum interpolation checks of increments. The increment i_o at the point under check is compared with the result of optimum interpolation

$$\left(i_0\right)^{\text{int}} = \sum_{k=1}^{m} w_k i_k \tag{3.18}$$

from its values i_k at m surrounding stations (horizontal check) or at m surrounding mandatory levels (vertical check). Here, w_k are the optimum interpolation weights computed from the system of linear equations

$$\sum_{k=1}^{m} r_{kl} w_k + \eta^2 w_l = r_{l0} \qquad (l = 1,2,\ldots,m), \tag{3.19}$$

where r_{kl} is the correlation coefficient between the values of the increment at points k and l, and η^2 is the "relative" variance of random observational errors, i.e., its ratio to the variance of increments. The relative RMS difference ϵ between interpolated and observed values (the so-called RMS comparison error) is also computed, as a by-product of the optimum interpolation, for each check, using the equation

$$\varepsilon^2 = 1 + \eta^2 - \sum_{k=1}^{m} r_{k0} w_k . \tag{3.20}$$

Increments at four (or less) closest surrounding stations situated in different quadrants around the station under check are used for the horizontal check, provided that the distance between each of them and the station under check does not exceed 1000 km and that the increment at none of them is too large by absolute value.. For the vertical check of an increment at an intermediate level, the interpolation is performed from two closest levels, one below the level under check, another above it. As to the lowest reported level and the highest one, the vertical interpolation to them reduces itself to the optimum extrapolation from one closest level.

The correlation coefficients in equations (3.19) and (3.20) are produced by correlation functions of increments similar to those used in the NMC Regional Data Assimilation System (DiMego, 1988), namely

$$r_{kl} = \exp\left(-\alpha d_{kl}{}^{2}\right) \tag{3.21}$$

for the horizontal correlation as a function of the distance d, and

$$r_{kl} = \frac{1}{1+\beta\left|\ln\left(\frac{p_k}{p_l}\right)^{1.2}\right|} \tag{3.22}$$

for the vertical correlation as a function of the ratio of pressures. The coefficient α is taken equal to 3.5x10-6 m^{-2} for both height and temperature, while the coefficient β is assumed equal to 1.1 for height and 9.0 for temperature.

One may argue that more accurate interpolation results can be achieved by improving the correlation function approximations and by using larger numbers of influencing points for the interpolation. This may be rather important when using the optimum interpolation for objective analysis but not for the quality control, just because it deals only with rough errors. Moreover, the more surrounding points are used in an interpolation check, the higher is the danger that some of neighboring values would be distorted by rough errors as well. That is why the interpolation for the quality control should be performed using much smaller numbers of influencing points than the numbers used in the data assimilation.

The application of incremental and horizontal checks to the surface pressure is slightly more complicated, as compared with those for mandatory level heights and temperatures, because the model topography is used to compute the first guess and because the elevations of neighboring stations may be quite different from each other even if the distance between the stations is small. The surface pressure is first reduced to the mean sea level using equations analogous to those applied in the baseline check, and first guess pressure at the mean sea level is computed. This means that the incremental and horizontal checks are actually applied to the mean sea level pressure increments.

Admissible residuals for height and temperature for the horizontal and vertical checks are presented in Table 3.4 along with those for the incremental check. They are approximately seven times the standard deviation of the residuals in the absence of errors.

4. Decision Making Algorithm (DMA)

This section contains a general description of the CQCHT DMA in comparison with the CHQC DMA. The CQCHT DMA actions concerning various kinds of suspected errors will be discussed in more detail and illustrated by examples in Sections 5-7.

As described in CG90, the CHQC algorithm included only two kinds of checks, the hydrostatic check for all layers between neighboring mandatory surfaces and the baseline check. The CHQC DMA successively analyzed each set of hydrostatic residuals for three layers (four levels), moving upwards and using the magnitude conditions (Table 3.3) to detect *large* hydrostatic *residuals*. If there were none, the DMA moved to the next set of layers and continued this scanning until it either found a set containing large hydrostatic residual(s) or reached the upper level of the report having found no large residuals. In the latter case the CHQC DMA concluded that the report did not contain hydrostatically detectable rough errors (that happened, of course, with an overwhelming majority of reports) and went to the next report. As to the former case, when the DMA did find a set containing at least one large residual, it then applied another group of conditions, the so-called *existence conditions*, separating the patterns of hydrostatic residuals caused by rough errors of various types and thus allowing the DMA to conclude what was the (most probable) cause of the error(s) and to locate it (them).

In many cases, the CHQC DMA was able to go further by computing the errors and thus correcting them. It even tried to find the so-called *simple corrections* resulting in changing only one digit, or only sign, or in transposition of digits. That is justified by the fact that a majority of these errors are introduced in the course of manual operations and are therefore, most probably, simple ones.

There were many other cases, when the CHQC DMA just assigned the error type but did not perform any correction, because, based only on hydrostatic residuals, it would be either risky to make corresponding corrections, or even impossible to univaluedly determine them. The DMA included all its information about every such case into a special file, called the SDM file, and sent this file to the SDM, who decided what to do in each such case. This happened to about a half of all errors detected by the CHQC, so that its DMA was able to automatically correct only about 50% of all hydrostatically detected errors (CG90).

As to the baseline check, the CHQC DMA did not even try to recognize the origin of any large residual and, thus, did not perform any corrections based on the baseline check results. There was no other option because, as it will be discussed in detail in Section 6, errors of quite different origin, e.g., a communication error in the surface pressure and an error in computing the height of the lowest mandatory surface, may result in the same baseline residual (producing no hydrostatic residuals). Quite different actions should be undertaken to correct these two types of errors, a change of reported surface pressure in the first case or the change of all reported heights in the second case. It was impossible for the CHQC algorithm to distinguish between these (and other) types of errors resulting in large baseline residuals just because it did not contain other, statistical checks. One can say that, although the CHQC algorithm did include the baseline check, its DMA did not pay much attention to its results.

Even so, it was important to have the baseline check within the CHQC algorithm. Particularly, our monitoring of its results, presented in the CHQC Monthly Summaries, allowed us to detect erroneous elevation of some stations in the NMC upper-air station dictionary caused, most probably, by the station movement, and to correct the wrong elevations (Gandin, Morone, Collins, 1993). It was clear, however, that the inclusion of statistical checks would substantially improve the use of the baseline check information.

Generally, we realized at the beginning of the CQCHT design that it would be a much more advanced and complicated algorithm than the CHQC. At the same time, it was important to develop it on the base of the CHQC algorithm, rather than to design a completely new one, so that if, by one or another reason, there is no first guess available, then the CQCHT will just work us a CHQC. Our present experience with the CQCHT does show that it happens, though very rarely, that the proper first guess is not available. Moreover, the NMC NWP models were unable, for many years, to produce reliable first guess fields above the 50-hPa level because they did not have sufficient vertical resolution in the stratosphere. For several years, the operational CQCHT at the NMC was actually a symbiosis of CQCHT up to 50 hPa and CHQC above. That caused much inconvenience, but it was better than just giving up any attempts to detect and correct rough errors above 50 hPa.

Superficially, it may seem that the CQCHT DMA reaction on the hydrostatic check results does not differ much from that of the CHQC DMA: the latter performs the same vertical scanning of each report and uses the same sets of existence and magnitude conditions as the former does in order to locate the possible hydrostatic errors and to determine their types (listed in Table 5.1) The essence of the CQCHT DMA actions is, however, different. Strictly speaking, there do not exist such things as hydrostatic errors, i.e., errors dealt with the hydrostatic check without any influence of other checks in the CQCHT DMA. Instead, the described actions result in hydrostatically *suspected* errors, or *hydrostatic suspicions*. Additional set of conditions, called the acceptance conditions and based on the residuals of statistical checks, is then applied to each hydrostatic suspicion before the DMA makes its decision.

The idea of acceptance conditions is that none of the residuals of statistical (and baseline) checks should remain or become large after the correction of suspected error(s). The DMA just computes all these residuals. If at least one of them is large, then this correction is not performed, and the DMA examines other alternatives, like the rehabilitation of the suspected value (i.e., retaining it as it was) or postponement of the final decision until other parameters in the report are checked.

The acceptance conditions are particularly useful for *multivalued hydrostatic suspicions*, in other words, when the suspicion may be caused by various kinds of errors. If, for example, only the hydrostatic residual for the highest layer was large, then it was impossible to decide without other checks whether the height or the temperature of the highest level is wrong (or, maybe, both are). The CHQC DMA just passed its information about all multivalued suspicions to the SDM. As to the CQCHT DMA, it applies the acceptance conditions to decide what was wrong and to automatically correct such errors.

Even more important is the role of acceptance conditions for the errors suspected by the baseline checks, because, as it was mentioned above, every *baseline suspicion* is multivalued. Strictly speaking, the term acceptance conditions does not adequately describe the role of these conditions for the baseline suspicions, as well as for any multivariate suspicions: they are used, first of all, to clarify the type of error and then to investigate whether its correction should be accepted.

One of the error types recognizable by the aid of the baseline suspicions (they are listed in Table 6.1) is an observation error in the surface-air pressure p_s, denoted as a Type 106 error. Its correction would result in changing not only p_s, but also heights of all surfaces, and the CQCHT DMA computes all these corrections. It was decided, however, after consultations with W. Whitmore and other MOD specialists, that the DMA should not automatically correct Type 106 errors. Instead, it provides MOD with all its information about each such error. The reason for this unique decision (all other correctable errors are automatically corrected with or without passing the information to MOD) is the fact that it can actually be an error not in p_s observation, but in its forecast first guess, also resulting in large statistical check residuals for heights. This actually happens sometimes in data-poor regions and/or in situations with rapidly changing p_s. It is safer therefore to pass all information about each Type 106 error to the MOD specialists so that they can make their final decisions.

Type 106 errors are exceptional ones also in the sense that they are the only observational errors that can be corrected. Errors of observational origin in temperatures and height errors caused by them do happen comparatively often, but they can never be corrected because the reported values of T and z are not observed values but those computed at stations and because the hydrostatic check does not react to observation errors.

At the same time, the absence of hydrostatic suspicions proves to be a rather powerful means for the automatic detection of observational errors by the CQCHT DMA. If there are large residuals of statistical checks and no hydrostatic suspicions, then it is highly probable that the errors are of observational origin. There may be two other causes of such situation: errors in the forecast first guess or an error in the station position caused by improper station identifier (the latter happens more often for ship observations). If the residuals of both incremental and horizontal checks are available, then it may be possible to recognize such causes. In any case, our experience shows that an overwhelming majority of errors, detected by the DMA as observational ones, are really of observational origin.

Having detected a report with observational errors, the DMA decides to reject some reported data from the DAS and/or to assimilate some data with smaller weights than if they were error-free. The DMA also includes all such reports, together with its decisions, into its SDM File, so that a MOD specialist can modify the DMA decisions.

Due to the possibility to deal with observational errors, the CQCHT algorithm performs by two successive scans denoted scan 1 and scan 2. Every report undergoes scan 1, but only those that were suspected, and perhaps corrected, by scan 1 are subjected to scan 2. Again, only if there were any suspicions (or corrections) by scan 2, then this information is stored. Such organization assures that, as a rule, any correction or retention is made by Scan 1, while any decision about observational errors is performed by scan 2. Only for reports with multiple non-observational errors, does it sometimes happen that scan 1 is unable to perform all corrections, so that a part of them is made by scan 2

Summarizing what has been said about various decisions made by the CQCHT DMA, one can see that there are as many as five decision types, listed in Table 4.1, as compared with only two decision types, Nos. 1 and 5, (not denoted this way) in the CHQC DMA. Types of all decisions made by the CQCHT DMA for each report are stored together with corresponding scan numbers.

A distinctive property of the CQCHT algorithm is that it automatically creates numerous files reflecting, with various degree of detail, each DMA action and used for various purposes. The Action Motivation File is the most detailed one, containing all information that is necessary in order to understand why each particular action has been undertaken. These files were extensively used by us in the course of CQCHT design and improvement and are still used occasionally to consider further possible improvements. The most condensed, least detailed file is called the Events File. It presents each DMA action by one line, containing all information necessary to understand what the DMA did, but not always sufficient to understand why it did so. Unlike other CQCHT files, stored for only several days, the *Events File* information is accumulated during more than a month and forms a basis for the CQCHT *Monthly Summaries*.

The most widely used CQCHT files, the *Operational Output Files* (OOFs), are intermediate in detail between the two files described above and ordained for human inspection of the CQCHT actions. There are usually several such files, including the SDM File and the Monitoring File(s), all presented in the same format, which is the easiest for understanding. Like any other CQCHT file, the OOFs contains information only about reports suspected by the DMA. As illustrated by Fig. 4.1., the operational output for each suspected report (and each scan) consists of four parts: the heading, containing information on the station position and observation time, the quick recognition table, the main body, and the final part reporting the DMA actions (if any). The main body of the output contains all information, level by level (beginning by the station level) on reported heights and temperatures, as well as residuals of all checks. The same information is reflected in the quick recognition table, preceding the main body, but it is presented in a quasi-qualitative way, which facilitates the recognition of the problem by a specialist. Numbers in the IHSC column are types of suspected hydrostatic (or baseline) errors, while digits 0 (no suspicion), 1 (suspicion) and 2 (strong suspicion) in other columns reflect the residuals of statistical checks. It may be immediately seen from this table in Fig. 4.1, that the hydrostatic suspicion of Type 2 (communication error in temperature) at 500 hPa was supported by all other checks. The DMA diagnosed a simple error, that in one digit and sign, and corrected it.

All examples presented in Sections 5-7 are taken from the CQCHT operational outputs. To save space, their quick recognition tables are not shown, and only a part of the main body, essential for the DMA actions, is shown in each example.

5. Hydrostatically Suspected Errors

Analyzing the pattern of large hydrostatic residuals, the DMA not only finds which values of temperature and/or height (if any) should be suspected, but also assigns one or another *hydrostatically suspected error type* to each suspicion. These types, listed in Table 5.1, are the same as they were for the CHQC, but further actions of the CQCHT DMA are quite different. The CHQC DMA just made corrections of all suspected large isolated errors at intermediate levels (types 1 and 2) and those at two neighboring levels (types 7-10) and included information about other suspected errors, not trying to correct them, into the SDM file. As to the CQCHT DMA, it first uses the *acceptance conditions* for each hydrostatically proposed correction, and if at least one of the acceptance conditions is violated, then the correction is not made.

The acceptance conditions for temperature corrections are straightforward: none of the statistical check residuals of the corrected value (including its increment) should exceed the value indicated by the magnitude condition. One can easily examine that just looking at initial statistical residuals: their values should be close to those of the hydrostatic residuals in terms of temperature, as was the case in the example in Fig. 4.1.

The use or, better to say, formulation of acceptance conditions for height corrections is slightly more complicated. It is necessary to take into account that the mandatory level heights are not measured but computed at the station from the temperature profile by using the hydrostatic equation. If by one or another reason the temperature measurement errors persist (say, are of the same sign) along the vertical, then the accumulated influence of such errors, even of small ones, results in comparatively large height increments and horizontal check residuals of the same sign, forming a kind of background. The incremental and horizontal check residuals of corrected height should be thus compared not with zero but with this background, and that is what the acceptance conditions for the height corrections do, comparing each such residual with the mean between its values for neighboring levels above and below. In other words, it is not the residual R_i at the level i but the value

$$R'_i = R_i - \frac{1}{2}\left(R_{i-1} + R_{i+1}\right) \qquad (5.1)$$

that is compared with the admissible residual when applying the acceptance conditions to the height increments and horizontal residuals. It is also not difficult to examine these conditions just by looking at the CQCHT output because the values R_{i-1} and R_{i+1} are usually close to each other.

The example in Fig. 5.1 is, in a sense, an extreme example of such a situation. Residuals of incremental and horizontal checks of Z_{700} were both small, and it might seem therefore that no correction was needed. However, both residuals were different from their background, and this difference was close to the hydrostatic residuals (in terms of height). That was exactly what the DMA needed in order to accept and perform the correction.

One can also see that the corrections in both examples presented on Figs. 4.1 and 5.1 were simple: one digit plus sign in Fig. 4.1 and single digit in Fig. 5.1. Attempting to find a simple correction, the DMA examines their slightly modified values, both with the same and the opposite sign, as was described in some detail in CG90. The only difference is that, if the modified, simple, correction does not satisfy the acceptance conditions, then the CQCHT DMA returns to the initially suggested correction to check the acceptance conditions for it.

Although human mistakes most often result in simple errors correctable by the CQCHT (and CHQC) DMA, not every human error, however simple, in a broad sense, it is, leads to a single digit, transposition of digits, or/and a sign error. For example, one digit may be missing, which is a simple error, but does not belong to the set of possible simple errors examined by the DMA. Another example is a "repetition" error, when a reported temperature or even height of some level is erroneously repeated for the next level. No simple corrections of errors like that are needed, but the DMA "does not know" this and still tries, sometimes "successfully", to make a simple correction. It would be possible to include special provisions to the DMA making it capable of recognizing repetition errors and other events like that, but that would result in only slight, if any, improvement of the DMA performance on the expense of further complication of an already complicated algorithm. This is just an example illustrating the general point of the DMA design: the inclusion of additional provisions into it should be limited not by possibility, or complexity, of an addition but by its *desirability*. As will be shown later, such dilemmas often emerge when the DMA treatment of reports with multiple errors is considered.

As to the isolated errors considered above, the CQCHT corrections to a majority of them coincide with those made by the CHQC. There are, however, some exceptions. First, the CQCHT DMA corrects many small hydrostatically suspected errors in either height (type 11) or temperature (type 22), while the CHQC DMA only passed its information on them to SDM. In fact, the CQCHT treatment of types 11 and 22 does not differ from that of types 1 and 2, and the only reason to preserve this distinction is to provide a protection for rare situations when there is no first guess.

Second, it happens sometimes that the hydrostatic suspicion of a small isolated error (type 11 or 22), or even of a large one (type 1 or 2) is not confirmed by the acceptance conditions, as was in the case shown in Fig. 5.2. There were two large hydrostatic residuals in a row close to each other (in terms of temperature) which made the DMA to suspect a type 2 error. However, the acceptance conditions did not support this suspicion, and the DMA concluded that there was no error (decision 2). It is easy to explain, considering Fig. 5.2., what actually happened in this case: the assumption of a linear (with respect to log(p)) temperature profile was strongly violated in the two layers because they are close to the tropopause, and that led to fictitious hydrostatic suspicion of a communication error in the temperature.

The statistical residuals (more exactly, the height increments and horizontal residuals) play a crucial role in diagnosing a computational error, i.e., an error in computing (or writing down) a thickness when processing the sounding at a station. While such an error results in a hydrostatic residual for a single layer, its correction necessitates the subtraction of this error from all mandatory level heights beginning with that of the upper boundary of this layer. It would be rather risky to make these multiple corrections based only on the hydrostatic residual, particularly because such isolated hydrostatic residual might be, as illustrated by some further examples, of quite different origin. That is why the CHQC DMA did not correct suspected computational (type 6) errors, just passing its information about them to the SDM. The situation with the CQCHT is quite different, as may be seen from an example in Fig. 5.3. Both increments and horizontal residuals above the layer strongly confirm the hydrostatic suspicion of a computation error, and the CQCHT DMA corrected all erroneous heights.

In fact, the heights in this example were erroneous only up to the 100-hPa level, while those at 70 hPa and above were correct. This happens comparatively often and is caused by the fact that parts A (up to 100 hPa) and C (above) of rawinsonde reports are transmitted separately. At the time when the part C was prepared for transmission, the computation error reflected in part A had been discovered and corrected, but the station did not retransmit the corrected part A (or, at least, NMC did not receive it). Although the CQCHT DMA does not include explicit provisions for cases like that, it treats them quite successfully, first correcting all heights including those in part C and then "re-correcting back" the part C heights. The DMA even does not need its two scans for doing so: both operations are performed by scan 1.

Another type of hydrostatically suspected errors that might be, in principle, corrected by the hydrostatic check alone, but were not corrected by the CHQC DMA, is the type 3 hydrostatic suspicion: suspected communication errors in both height and temperature of the same level. The two equations for hydrostatic residuals for layers below and above the level in question form a system of two linear equations for two unknowns, the corrected height and temperature at this level. However, solutions of such systems are often not stable enough, a small variation of residuals may result in a large change in corrections. Even more important is the fact that type 3 error hydrostatic suspicions are often caused by errors of quite different origin, just as for type 6 suspicions. The availability of statistical residuals allows the CQCHT DMA to automatically correct type 3 errors, as it did in example presented in Fig. 5.4. This output form does not contain the values of hydrostatic residuals after the correction, but one can check that they are rather small. At the same time, both the height and temperature corrections agree quite well with results of statistical checks.

Fig. 5.5 illustrates another situation. Although the hydrostatic check suspected both height and temperature of the 400-hPa level, the DMA conclusion, based on statistical residuals, was that there was no error in the temperature, and it corrected only the height. In other words, the DMA "transformed" the type 3 suspicion into type 1 correction. It is convenient to denote such action by an arrow directed from suspected type to corrected one; in this case, it was a type 3 → 1 correction. Using these notations, we can say that the CQCHT DMA is able to perform corrections of types 3 → 3, 3 → 1, 3 → 2 and 3 → 0 (the last meaning no correction).

The situation with two hydrostatically suspected errors at neighboring levels (types 7-10) is quite analogous to that. For example, 7ype 7 suspicions (those of height errors at both levels) may result in corrections of types 7 → 7, 7 → 1+0 (only the lower height), 7 → 0+1 and 7 → 0+0.

The next example (Fig. 5.6) illustrates a quite different situation. It was the hydrostatic suspicion of type 6 error at 500-hPa level, i.e., of a computational error in the 700--500-hPa layer thickness. However, statistical residuals did not confirm this suspicion, and the DMA did not make any. What actually happened in this case, however, was a type 3 error with what we call a *compensation effect*. The contributions of height and temperature errors to the hydrostatic residual for the 500-hPa layer were of opposite signs and close by absolute values. That is why this residual was, as seen on Fig. 5.6, small, and this prevented the DMA from type 3 hydrostatic suspicion, making it suspect a type 6 error instead. Information about this case was included, along with that on all cases with DMA decisions different from 1, into the SDM file, so that an MOD specialist could make a proper decision.

One may ask why the DMA rejected correctable data instead of correcting them? The answer is simple: the DMA, in its present version, does not contain special provisions necessary for diagnosing the compensation effects. Such provisions could be, of course, introduced, but we decided not to do it, because such effects take place extremely seldom and because there would still remain even more complicated situations (e.g., combinations of type 3 and type 6 errors) requiring further complication of the DMA. This is another illustration of above mentioned dilemmas concerning the possibility and desirability of the DMA extensions.

Unlike the hydrostatic suspicions considered so far, those of type 4 (an error at the lowest reported level) and 5 (at the highest one) require the use of statistical residuals not only in acceptance conditions but, first of all, in what may be called the *selection conditions*. A type 4 or 5 error results in only one large hydrostatic residual which just signals that something is, most probably, wrong, and it is impossible without statistical residuals to decide *what* (if anything) is wrong. For example, in the case presented in Fig. 5.7, the type 5 suspicion indicates that there is, most probably, *either* an error of about 200 m in the 200-hPa height, or an error of approximately 60 K in its temperature (or, maybe, both are wrong). It is clear in this case, that only the height was wrong (a type 5 → 1 error), and that was the DMA conclusion based on hydrostatic and statistical residuals.

Most of the type 5 suspicions result in either 5 → 1 or 5 → 2 correction, but it happens sometimes that both height and temperature of the upper level are wrong. The role of statistical residuals in such, type 5 → 3, cases, like that illustrated by Fig. 5.8, is to *partition* the hydrostatic residual into those caused by errors in temperature and in height.

The list of possible DMA reactions on type 5 suspicion, 5 → 1, 5 → 2, 5 → 3 and 5 → 0 (no correction), looks analogously to that for type 3 suspicions but, unlike type 3 → 1 and 3 → 2 errors, which happen as exceptions, type 5 → 1 and 5 → 2 errors are most common among those suspected as type 5 errors.

The DMA reaction to a type 4 hydrostatic suspicion, that of an error at the lowest level, is to a large extent analogous to its reaction on a type 5 suspicion, as may be seen in Fig. 5.9. illustrating a 4 → 2 correction. There is, however, an important difference: the list of possible errors resulting in type 4 suspicions includes additionally a computational error in the thickness of the lowest layer (type 4 → 6 error). As illustrated by Fig. 5.10, the DMA corrects such an error by modifying all heights, except that of the lowest level, by the same quantity, just as it acts with "ordinary" type 6 errors, those at intermediate layers. The list of possible DMA reactions to a type 4 hydrostatic suspicion thus includes 5 options: 4 → 1, 4 → 2, 4 → 3, 4 → 0 and 4 → 6. (Strictly speaking, the same is true for type 5 suspicions, but type 5 → 6 errors cannot, and should not, be distinguished from type 5 → 1 errors).

The so-called compensation effect, discussed above in connection with type 3 errors, is even more destructive with type 4 and 5 errors. While in the case of a type 3 error, a compensation effect results in disappearance of one of two hydrostatic residuals, and the remaining one signals that something may be wrong, an analogous effect for a type 4 or 5 error leads to the disappearance of the only hydrostatic residual which would exist otherwise. That means that type 4 → 3 or 5 → 3 errors with compensation are not suspected at all by the hydrostatic check. Fortunately, the situation with such errors is not so bad as it may seem. As described in the next Section, the baseline check makes it possible to diagnose and correct type 4 → 3 errors with compensation in spite of the absence of hydrostatic suspicions. As to type 5 → 3 errors with compensation, they cannot be distinguished from observation errors (considered in Section 7), and it is much safer to reject these erroneous data than to try to correct them.

The last kind of hydrostatic suspicions to be considered here is that of a so-called data hole (types 13 and 14), i.e., two or more levels in a row with missing data followed by at least one level with complete information. It may seem that the CQCHT treatment of data holes does not differ from that by CHQC because, in a majority of such cases, the CQCHT DMA just includes its information about the data hole into the SDM file like the CHQC DMA did. In fact, however, the CQCHT DMA does much more investigating, for each hole, whether there are any errors at the lower or/and upper boundary of the hole and, if so, trys to correct the errors.

When doing so, the DMA does not pay any attention to the hydrostatic residual within the hole. This residual may be, and very often is, rather large not because of any error but simply because the hole occupies a thick layer, so that the hydrostatic check assumption of linear (with respect to log(*p*)) temperature profile is strongly violated. (The same may happen if the hole includes the tropopause; that is why a type 13 hole is diagnosed even if it consists of only one upper level of the part A, the 100-hPa level).

Ignoring the hydrostatic residual within the hole, the CQCHT DMA thus treats each hole-containing report as if it were two independent reports, one below the hole and another above it. This means that the DMA may hydrostatically suspect and, if possible, correct error(s) at the hole's lower boundary just like it does so for the upper level of the whole report (type 5 hydrostatically suspected errors). Analogously, the DMA hydrostatically suspects and corrects errors at the holes' upper boundaries as if they were errors at the lowest level (type 4).

Quite naturally, there usually are no errors at holes' boundaries, and it may seem that the CQCHT DMA actions in such cases, illustrated by Fig. 5.11, do not differ from what the CHQC could do. In fact, however, the DMA has checked if there were any errors at 400 and/or 70 hPa and concluded that there were none (decisions 2). Looking at Fig. 5.11, one can see that all increments and horizontal residuals for these levels are really small. At the same time, the hydrostatic residual between these two levels is rather large. There is no doubt that it was entirely caused by the non-linearity of temperature profile between these levels, and the only communication-related error in this report was just the presence of the data hole.

Fig. 5.12 illustrates the DMA actions concerning the errors at a data hole boundary. The height at the upper boundary of the hole was found wrong and corrected by the DMA (type 14 → 4' → 1 corrections). The correction is close to that indicated by corresponding statistical residuals.

What has been said so far about the CQCHT DMA reaction to data holes should not leave a false impression that the data holes have no adverse influence on the quality control. Both detection and correction of errors at the lowest and the highest levels of reports are more difficult and therefore less productive than at intermediate levels, and so are the error detection and correction at the holes' boundaries. In addition, there is a type of errors that cannot be confidently diagnosed because of the data holes, namely, computational errors (those of type 6) within the holes or just below them. As mentioned above, the hydrostatic residuals within the holes cannot be believed in, and that makes the diagnosis and subsequent correction of computational errors within the holes practically impossible.

6. Baseline-type Errors

As mentioned above, the baseline check is essentially a hydrostatic check but applied not for layers between two mandatory levels (as is the case with what is called the hydrostatic check in this paper) but for the layer between the station level and the lowest mandatory surfaces. Large residuals of the baseline check can be created by errors of various origins, and statistical residuals play a crucial role in distinguishing between these origins. This is particularly true for the residuals of incremental and horizontal checks of the surface-air pressure. It is convenient therefore to consider these two checks together with the baseline check as a separate, baseline-related group of checks and call the errors which influence either the statistical residuals of the surface pressure, or the baseline residual, or both the *baseline errors*.

Table 6.1. contains the list of various baseline error types. There was nothing like this classification in CG90 or in any other publication on the quality control. We shall therefore consider the classification of baseline errors in some more detail than was done in the previous section for hydrostatically suspected errors.

If the surface-air pressure was measured correctly and correctly used for the computation of the first mandatory level height, but distorted afterwards, in the course of communication, then it is a type 100 error, a communication-related error in the surface pressure p_s. Such an error influences the baseline check residual, as well as the surface pressure increment and horizontal residual, and it does not influence anything else. To recognize a type 100 error, the DMA considers the baseline residual in terms of p_s, the p_s increment and its horizontal residual (if the latter is available). If all of them are large and close to each other, then the DMA diagnoses the type 100 error, computes its value by averaging the three (or two) estimates and introduces the correction, as it did in the case shown in Fig. 6.1. In fact the DMA does slightly more than that. Type 100 errors are, most probably, human errors, and the DMA tries therefore to find a simple correction of each type 100 error like it does for every communication error.

In this case, the DMA did not find a simple correction, and we know why it did not: actually, it was an error not in the surface pressure, but in the station elevation. From a "theoretical" point of view, each type 100 error may be a communication error either in the surface pressure p_s or in the station elevation z_s. However, the station elevations are never communicated, they are taken from the stored file called the NMC upper-air station dictionary. However, it happens sometimes, though very seldom, that a station moves, changing its elevation as well, and no information about the relocation is passed to NMC (or to any other NWP center). That was exactly what happened in the case illustrated by Fig. 6.1. We knew it because this error, with about the same baseline residual, occurred with each report from this station, before we discovered this by analyzing a CQCHT monthly summary, and the station elevation in the dictionary was corrected. It is interesting to mention that, although the elevation of this station changed by as much as -250 m, its horizontal position did not change substantially, as witnessed by small statistical residuals at mandatory surfaces.

Because of the persistence in the reaction of the baseline check to a station elevation error, such errors may be discovered based on the baseline check alone, and we did it several years ago, before the CQCHT was designed and implemented (Gandin, Morone, Collins, 1993). Under such conditions, however, a permanent baseline residual might be, at least in principle, caused not by wrong elevation of the station but by systematic errors of its barometer. This alternative does not exist for the CQCHT DMA, because an observational error in the surface pressure p_s, denoted type 106 error and illustrated by Fig. 6.2, leads to a quite different pattern of the CQCHT residuals as compared with that for a communication error in p_s, or for an error in z_s. Unlike a type 100 error, that of type 106 does not create any baseline residual, it results only in a large increment and horizontal residual of the surface pressure. Even more important, a type 106 error, unlike a type 100 one, does lead to errors in all mandatory level heights, reflected by their large statistical residuals (while those for mandatory level temperatures remain small). All height increments and horizontal residuals are close to each other and to the product of the surface pressure measurement error Dp_s and the so-called barometric step

$$\frac{Dz}{Dp_s} = \frac{RT_s}{gp_s} \quad (6.1)$$

approximately equal to 8 m/hPa.

That is exactly what happened in the case presented in Fig. 6.2. The pattern of height statistical residuals is reminiscent of a thickness computation error (type 6, thus the type 106 notation), but *all* heights are erroneous and should be corrected and, additionally, the surface pressure is erroneous and should be corrected as well.

Type 106 errors are exceptional from several points of view. First, no other error of observational origin than that in surface pressure can be corrected, simply because reported heights and temperatures of mandatory levels were not measured but computed from the measurement results. Second, it may happen to a type 106 error, as for any observational error diagnosed by the CQCHT (or by any other method using the forecast first guess), that it actually was an error in the first guess, not in the observation. As for suspected observational errors at mandatory surfaces, this is not very dangerous because, at worst, some temperatures and/or heights will be mistakenly assimilated with smaller weights or even rejected. The situation with type 106 errors is much more serious because each wrong type 106 diagnosis would lead to wrong corrections not only of p_s but of all mandatory level heights as well.

Type 106 errors occur seldom, 2 or 3 times, in the mean, per a main observation time worldwide. We monitored all such errors for more than a year. Only about 70% of them appeared to be actually type 106 errors, others were, most probably, caused by other effects. We decided therefore, following W. Whitmore's suggestion, not to make automatically type 106 corrections, but to pass, instead, all CQCHT information on each such error, including computed corrections, to the SDM file, so that it is up to a MOD specialist to make the final decision.

The corrections in Table 6.2. are thus not performed but only proposed, as indicated by decisions 5. This is also an exception; in all other situations with decisions different from 1, the corrections, independently on whether any was tried, are put equal to 0.

The situation in the next example (Fig. 6.3) looks quite similar to that in Fig. 6.2, but the error origin is different: the error has been made not while measuring p_s but while computing the height z1 of the lowest mandatory surface (type 116 error). Such an error also results in height errors of all mandatory surfaces. These errors are close to each other and to the baseline residual (in terms of the station elevation) with the opposite sign. In order to correct a type 116 error, it is necessary to subtract it from all heights but, unlike type 106 errors, that is all: the surface pressure does not need any correction.

We believe that the situation with the error types 106 and 116 is a good demonstration of the CQC approach, i.e., its attempts to recognize the origin of any suspected error, its, so to say, mechanism, before deciding to correct (or reject) any erroneous data. It is not difficult for the CQCHT DMA to distinguish between type 106 and 116 errors, although both show the same pattern of statistical residuals for heights (and no large residuals for temperature or hydrostatic residuals). The baseline residuals show different patterns. A type 106 error is recognized as having *large* p_s increment and horizontal residual, which are close to each other, to the baseline residual in terms of p_s

and to 1/8 of the height increments and horizontal residuals. For type 116 error, the DMA recognizes it as having *small* p_s increment and horizontal residual and large baseline residual in terms of the station elevation, close by absolute value and opposite by sign to the height increments and horizontal residuals.

An important difference between type 106 and 116 errors is that a residual pattern typical for a type 116 error can not be a result of observation errors or, to be more exact, there is no simple mechanism for observation errors to look like a type 116 error, as is the case for type 106 errors. All type 116 errors are therefore automatically corrected by the DMA.

The same is true with the remaining type of errors that can be recognized with the aid of the baseline check, type 101 errors. The type 101 error is a communication-related error in the height, z1, of the lowest mandatory level with missing temperature of this level in the report. According to existing rules, the 1000-hPa height should be computed, while processing the rawinsonde data, at all stations (except highly elevated ones), even if this surface is under the ground. The extrapolation applied in order to compute such underground heights is analogous to that used in our baseline check and described in Section 3. The extrapolated heights are included into reports, while the underground temperatures, also computed in the course of extrapolation, are not. It happens therefore rather often that the temperature of the lowest mandatory level (or even of several such levels) is missing in a report while its height is there.

If there was an error in communicating such height, a type 101 error, then the baseline check is used by the DMA, instead of the missing hydrostatic check for this, incomplete level, in order to diagnose and correct the error, as it did in the example presented at Fig. 6.4. As illustrated by this case, the DMA tries to find simple type 101 errors. In this case, it was a sign error with transposition of digits.

An analogous procedure could be designed to identify and correct a height communication error at the second reported level if its temperature is also missing. We decided, however, not to do it, because such cases happen extremely seldom, and also because the computation of the baseline residual in terms of z_2 is often not stable enough: small variations in z_1, z_2 and p_s may result in large changes of this residual.

There also exist situations when the baseline check is used by the DMA as an auxiliary means to confirm (or deny) a decision made on the basis of other checks. For example, the partition of the hydrostatic residual in the type 4 → 3 correction at Fig. 6.5. was made by using the statistical residuals of height and temperature and supported by the baseline check, whose residual (in terms of z_1) agreed quite well with statistical residuals of the height.

As mentioned before, the DMA behavior with the errors diagnosed with the baseline check's aid, i.e., with those of types 100, 101, 106 and 116, substantially differs from that with respect to hydrostatic errors. While the DMA first assigned an error type to each hydrostatically suspected error and then used results of other checks to make its final decision, the DMA uses the baseline checks in conjunction with other checks in order to assign any type listed above, and that leaves no ambiguity: all corrections of type 100, 101 and 116 errors are automatically performed by the DMA, and all type 106 corrections are computed but not performed.

The question remains therefore what should the DMA do in situations when the baseline residuals or, at least, some of them are large, but none of the condition sets for the four listed types is satisfied? Unfortunately, this happens comparatively often, as one could easily foresee. The extrapolation to the mean sea level, involved in baseline checks, is rather approximate, particularly over elevated terrain. The vertical temperature profile near the ground may strongly differ from a standard one assumed in the checks. And the forecast first guess is less reliable near the earth surface than anywhere else (except, maybe, for the upper stratosphere).

There is therefore a special error type, 102, a non-identified baseline error always accompanied by decision 5, a request for human help. It happens sometimes that the DMA itself solves the problem. For example, a type 1 or 3 correction at the second level may lead to what we call the annihilation of the type 102 suspicion, so that the scan 2 does not make it. Fig. 6.6 presents a slightly more complicated example of this kind, when the type 3→1 correction of the 925-hPa height transformed a type 102 error diagnosis by the first scan to the type 100 diagnosis and correction performed by the second scan.

In most cases, however, the type 102 diagnosis remains intact, and it is difficult to understand without additional information what actually happened in each such case, like that in Fig. 6.7. That is why the information about all unresolved baseline suspicions (except those for highly elevated stations) is included by the DMA into the SDM file.

7. Observational Errors. Second Scan.

As was mentioned above, the term *observational error* is applied in this article in a wide sense, encompassing all errors made before the processing of the report at a station began. It includes not only measurement errors but those committed when the measurement results were sent to the station, received there and prepared for undergoing the processing. It was also mentioned that, as long as, in the course of this processing, the mandatory level heights are hydrostatically computed from the temperature profile, the hydrostatic check does not react at all to the observational errors. That is why no observational errors were detected by the CHQC.

The situation with the CQCHT is quite different. All statistical checks react to observational errors, and the hydrostatic check plays an important role: if large errors detected by statistical checks did not cause large hydrostatic residuals, then, if they really are errors, they are, most probably, not of computational or communication-related nature, but of observational origin.

It is necessary to realize, however, that a similar configuration of residuals may be caused by errors in the forecast first guess; such errors also result in large statistical residuals and do not influence the hydrostatic ones. To provide some protection against misdiagnosing first guess errors as those of observational origin, the CQCHT DMA requires that the error should be comparatively large in order to be diagnosed as observational one. As may be seen by comparing Tables 3.3 and 3.4, the conditions are generally more severe than the existence conditions for hydrostatic suspicions. In other words, the smallest observational errors detectable by the CQCHT are larger by absolute value than the smallest detectable communication-related and computational errors. Still, the CQCHT sensitivity to observational errors is substantially higher than that of the gross check and the buddy check applied before.

Unlike the errors considered so far, the observational errors are very seldom isolated vertically, they usually occur at several levels. That is easy to explain by the fact that reported temperatures are those computed from measured ones, so that even a single measurement error may influence several computed temperatures. Another reason for the vertical persistence of the observation errors is that, once having happened, such an error is unlikely to disappear in the course of the rawinsonde ascent.

Even more pronounced is the vertical persistence of the height statistical residuals caused by observational errors in temperatures. The mandatory level heights are computed, upwards, from the temperature profile using the hydrostatic equation. A single observation error leads therefore to vertically persistent height errors. As mentioned in Section 5, this effect is accounted for by the CQCHT DMA when formulating the acceptance conditions for corrections of height communication errors.

As to vertically persistent observation errors, they result therefore, as illustrated by Fig. 7.1, in vertically increasing statistical residuals of height. This increases the DMA sensitivity to observational errors. Moreover, the DMA often decides in such situations to reject (decision 4) some heights, while retaining all temperatures, as it did in this case, or to artificially assign a higher RMS random observation error for these heights in order to assimilate them with smaller weights (decision 3).

Errors of this kind happen comparatively often. Their most probable cause is the so-called calibration error, i.e. an error in adjusting the rawinsonde sensor(s) to conditions at the launch site. That leads to a shift in temperature or/and pressure scale which, in its turn, results in vertically persisting errors. This cause of observational errors would not exist if the calibration of rawinsondes were performed automatically.

Fig. 7.2. illustrates another typical pattern of observational errors which occurs more often in reports from automatically processing stations. In this case, the temperature errors are much larger by absolute value, but they occupy only several mandatory levels in a row, while those below and above are error free. L. Morone, another NMC specialist in the quality control, came across such cases some time ago, and she succeeded in discovering what actually happened in these cases (Gandin, Morone, Collins, 1993). The receiving antenna at the station was for some time erroneously fixed on one of side lobes of the sonde signal, before the operator realized this and redirected the antenna to re-fix to the main lobe. It is easy to understand why such situations happen more often at stations with automatic processing of rawinsonde reports than at those where it is performed manually. Although it is still the operator's responsibility to ensure the proper antenna directions, he (or she) may pay, consciously or subconsciously, less attention to that under the relaxing environment of a computer-equipped station.

The fact that the automation of processing results sometimes in increased numbers of human errors was detected not long ago (Schwartz and Doswell, 1990). It was even proposed to return back to manual processing in order to avoid adverse influence of such errors on analyses and forecasts. It is, of course, bad to lose some information as a result of these errors. Our example demonstrates, however, that the erroneous data will never enter the assimilation system if there is an automatic quality control. As to the errors illustrated by this example, they will never occur after the antenna directing becomes automated as well.

The situation in Fig. 7.3 looks similar to that in Fig. 7.1, there also are persistent large statistical residuals of temperature which caused growing height residuals which, in their turn, resulted in the DMA decision to assimilate the 50-hPa height with a smaller weight. However, our monitoring of these errors revealed that they often persist in time over this region (Alaska), and the most reasonable explanation of this fact (if not the only possible one) is that these are not the observation errors but errors in the forecast first guess. This assumption is supported by the fact that the number of such events has recently decreased due, as we believe, to improvements in the NMC DASes and, particularly, to increased vertical resolution of the NMC global NWP model in the upper stratosphere.

There were no suspected communication or computation errors in the examples considered so far in this Section, so that the DMA performed its actions only at the second scan. The next example (Fig. 7.4) is typical for most situations when reports contain errors of both observational and non-observational origin: the DMA makes all its corrections of communication and computation errors at scan 1, while all its decisions 3 and 4, concerning observational errors, are made at scan 2. It even happens sometimes, though as exceptions rather than the rule, that some data corrected by the first scan are then rejected by the second one. That was the case in the example above; the 300-hPa height was confidently corrected by scan 1, but scan 2 rejected the corrected value because this part of report was also distorted by observational errors. Both these actions by the DMA were well motivated, there actually was a communication error in this height committed additionally to the observation errors.

That was not the worst that can happen with reports distorted by both observational and non-observational errors. The presence of observational error(s) can prevent the DMA from proper reaction to communication and/or computation error(s) and finally result in rejection of correctable data. Fig. 7.5 presents an example of such situation. The hydrostatic check quite correctly (as we can see) suspected type 6 error at 500 hPa, but acceptance conditions were not satisfied because the statistical residuals were influenced by observational errors. The DMA was therefore unable to correct the suspected value, and it was rejected by the second scan.

This example also demonstrates that the DMA decisions 3, typical for scan 2, are sometimes made by scan 1. The opposite situation, when scan 2 performs some correction(s) (i.e., decisions 1) in addition to those made by scan 1, also takes place sometimes with reports distorted by multiple communication and/or computational errors. The DMA usually succeeds to correct all such errors at its scan 1 if they are isolated from each other, i.e., if the hydrostatically suspected levels are separated from each other by suspicion-free levels. As to non-isolated errors, the DMA is not always capable, as illustrated by Fig. 7.6, to correct all of them at scan 1, and remaining corrections are made at scan 2. Another example of this kind, dealing with the baseline error suspicion, was demonstrated in the previous Section (Fig. 6.6).

All information about the observational errors is included in the SDM file, so that the MOD specialists can modify the DMA decisions by rejecting either more or less data than it has been done by the DMA. They can also distinguish, based on their experience, between the observational and first-guess errors and thus preserve some data that otherwise would be lost. Even more important is the possibility to perform the monitoring of observational errors made at various stations. The feed-back of this information to stations involved helps them to detect their problems and to improve their performance.

8. Some Statistics of the CQCHT Performance

The CQCHT actions undergo regular monitoring of various kinds by specialists belonging to several NMC divisions. The operational monitoring is performed by the SDMs and/or other MOD specialists in the interactive quality control (IAQC) mode, allowing them to make their subjective decisions, as well as to modify those made automatically by the DMA, in operationally acceptable time. The CQCHT designers perform, on a regular basis, what we call quasi-operational monitoring of CQCHT. Only the most complete data sets, those for the final runs (twice a day) are analyzed with a delay not exceeding several days. The main aim of the quasi-operational monitoring is the improvement of an existing algorithm and the design of new algorithms. The operational implementation of any new algorithm or of any new version of an existing algorithm is not done until it is thoroughly investigated, often in parallel with the operational version, in the quasi-operational monitoring mode. Such investigations usually require much time because only a small number of reports contain rough errors, and because the most complicated cases needed for the investigation happen particularly seldom.

The CQCHT quasi-operational monitoring is also a powerful means for detecting problems with the rawinsonde data that occur sometimes at one or another station or somewhere else outside or inside the NMC. The feed-back connections with those responsible for the problem is usually provided by the NMC QAG specialists and results, as a rule, in solution of appearing problems.

Both operational and quasi-operational monitoring use the CQCHT operational outputs described above (Section 4) and illustrated by Fig. 4.1. As to the third kind of the CQCHT monitoring, the monthly monitoring, also performed by us, it is based on the CQCHT monthly summaries produced at the end of each month by a special code. These summaries are regularly disseminated among various specialists at NMC and also sent to other specialists in this country and abroad who expressed their interest in this information.

Each monthly summary consists of three major parts. The first, most voluminous part is a list of all DMA actions for a month as they are exposed in the CQCHT Events Files (see Section 4) for major observation times. Information presented in this part is the only source of all statistics on the rough error distribution by various types and by various regions computed by the Monthly Summary Code and presented in the second part of the summary. Its third part contains statistics summarized for each station that committed at least one rough error during the month. The information collected in the CQCHT monthly summaries it stored for a long time. This archive is used in order to obtain various statistics averaged over larger periods of time.

The main applications of the monthly summaries are the analysis of rawinsonde data quality over various regions of the Earth and the detection and investigation of specific problems with the data which appeared repeatedly or persistently over some regions or at some stations. Some results of such analysis have been reported elsewhere (Morone, Gandin, Collins, 1992).

The information collected in CQCHT monthly summaries may also be used for the evaluation of the overall CQCHT performance. The corresponding statistics of monthly mean numbers of errors per observation time, averaged over the whole globe, are presented in Tables 8.1 and 8.2.

Table 8.1 contains the error numbers averaged over 18 months, as well as their standard deviations, for four major categories: hydrostatically suspected errors (except the data holes), errors suspected by the baseline checks, data holes, and observational errors. The table also includes corresponding numbers for the CHQC allowing comparison of the productivity of the two methods.

This table illustrates two major advantages of CQCHT over CHQC. First, suspecting hydrostatically errors in the same reports, the CQCHT DMA corrects a much higher number of these errors, more than 75% as compared with less than 50% of them (it also concludes that the remaining hydrostatically suspected errors should not be corrected). Second, the CQCHT, unlike the CHQC, detects observational errors and decides either to reject such erroneous data or to assimilate them with smaller weights.

Other CQCHT advantages are expressed by smaller numbers just because the errors of involved categories occur more seldom. The CQCHT baseline checks detect more errors than the single baseline check of the CHQC did and, much more importantly, the CQCHT DMA corrects a majority of these errors while the CHQC corrected none of them.

Some additional statistics of the CQCHT performance concerning the hydrostatically suspected errors are given in Table 8.2. They are based on a smaller sample than that in Table 8.1 and presented by relative numbers (in percent) rather than by absolute numbers of errors per observation time.

All types of hydrostatically suspected errors presented in Table 8.2. may be divided into two categories: those which would be automatically corrected by the CHQC (category I, the CHQC-correctable errors, types 1, 2 and 7-10) and those for which the CHQC DMA would rely on human help (category II, the CHQC-non-correctable errors, types 3-6, 11 and 22). Table 8.2 shows that the overall frequency of the category II suspected errors is even slightly higher than that for category I. At the same time, the number of corrections made by the CQCHT DMA is smaller for category II than for category I because the percent of the CQCHT corrections (decisions 1) of suspected errors is substantially smaller for types belonging to category II. This fact may be considered as a justification of the decision made several years ago when the CHQC DMA was designed, to make it automatically correct only the category I errors.

There are several reasons why the correction percent is comparatively small for suspected errors of various types belonging to category II. Type 11 and 22 suspected errors are, by definition, small, and the DMA often decides that there was no error at all and rehabilitates the suspected datum (decision 2). The same happens, although much more seldom, with type 4 suspicions which may be caused by the non-linearity of the temperature profile near the ground and/or by the influence of humidity on the virtual temperature. On the other hand, there were very few rehabilitation decisions for type 5 and 6 suspicions, and none for type 3. For these types, substantial fractions of the DMA non-correction decisions led to rejection of data (decision 4), to their assimilation with diminished weights (decision 3), or to requests for a human help (decision 5).

As mentioned above, the forecast first guess information was for several years of the CQCHT operation available only up to 50 hPa, and the CQCHT algorithm acted as the CHQC for 30-, 20- and 10-hPa levels. Being unable therefore to decide what to do with the hydrostatically suspected errors of category II at those levels, the CQCHT assigned decision 5 to any such suspicion. Naturally, this happened most often to suspected errors at the highest reported level (type 5). The number of such decisions would be much less if the forecast first guess were available everywhere.

As to the hydrostatically suspected errors of category I, those which would be automatically corrected by the CHQC DMA, the numbers in Table 8.2. show that an overwhelming majority of these corrections are performed by the CQCHT DMA as well. This conclusion confirms that, in rare situations when there is no first guess available, the CQCHT can still produce good results working as a CHQC.

9. Some Further Developments

The main purpose of the CQCHT design and operational implementation was, of course, the improvement of the NMC DASes and NWP results. At the same time, it has formed a basis for further development of the QC algorithms at the NMC DD. All these algorithms, briefly described in this Section, were already used at NMC to one or another extent.

Complex quality control of significant level temperatures. A new CQC of significant level temperatures was designed and implemented by one of us (W.C.). It includes their hydrostatic and incremental checks, as well as vertical checks by interpolation from mandatory levels only and from neighboring significant levels. The mandatory level temperatures applied in this algorithm are those that already underwent CQCHT. However, the algorithm checks them again, using significant level temperatures. Sometimes, though very seldom, its result is contradictory to the CQCHT correction, and this correction may be modified or even rejected. The new algorithm proved to be much better than the previous one (Collins, 1990) that did not use the forecast first guess.

CQCHT for the NMC Climate Data Assimilation System (CDAS). CDAS is the NMC global data assimilation system designed for climatological studies and using therefore a much longer cut-off time, than the operational GDAS does, in an effort to assimilate as much information as possible. Although the CQCHT might be applied to CDAS information as it is, an even more productive CQCHT version was designed for CDAS. It additionally includes the temporal check whose residual is the difference between a reported value and that linearly interpolated in time using one observation before and one after the time in question (provided that none of them is more than 24 hours apart). This additional check applied to height and to temperature is most important for isolated stations, because the horizontal check for them is not productive or may be even impossible. The CDAS CQCHT was also applied for the NMC-NCAR Reanalysis project (Kalnay, Jenne, 1991).

Preliminary quality control for Reanalysis. Initial data for the reanalysis are collected from various sources worldwide, and they should be subjected to some quality control before entering the reanalysis, when the first guess is not available. The CQCHT version with the temporal check (and without first guess) was used for this purpose. The role of the temporal checks is particularly crucial in the absence of other statistical checks. Our experiments with the reanalysis data have shown that, whenever the temporal checks of height and temperature are available, their complex with the hydrostatic and baseline checks is only slightly less efficient in detecting and correcting errors than is the complete CQCHT algorithm including incremental, horizontal and vertical checks.

The temporal check may be, and already has been, used also for the preliminary quality control of other reanalysis data. The most sensitive, under such circumstances, would be, perhaps, its complex with horizontal and vertical checks of anomalies.

Acknowledgements

We wish to thank Dr. Ronald McPherson for his permanent interest in our work, to Dr. Eugenia Kalnay for her leadership and personal involvement in the quality control design at the NMC DD, to our colleagues at other NMC Divisions, P. Julian, B. Balish, R. Hirano, J. Hoke and W.Whitmore for their cooperation and constructive criticism, and to L. Morone for her help in editing the article.

References

Aldukhov, O.A., 1982: Combined quality control of height and temperature values for isobaric surfaces in FGGE upper-air reports. *GARP International Conf. on Scientific Results of the Monsoon Experiments. Extended Abstracts and Panel Sesion*. Denpasar, Bali, GARP, WMO, 8-15.

Antsipovich, V.A., 1980: Complex quality control of height and temperature at mandatory isobaric surfaces. Hydrometeorological Center, *Trudy*, No 217 (in Russian).

Bedient, H.A. and G.P. Cressman, 1957: An experiment in automatic data processing. *Mon. Wea. Rev.*, 85 (10), 333-340.

Bergthorsson, P. and B.R. Doos, 1955: Numerical weather map analysis. *Tellus*, 7 (3), 329-340.

Collins, W.G., 1990: Quality control of significant level rawinsonde temperatures and pressures. NMC Office Note 373. [Available from NCEP, 5200 Auth Road, Washington, D.C., 20233]

Collins, W.G., L.S. Gandin, 1990: Comprehensive hydrostatic quality control at the National Meteorological Center. *Mon. Wea. Rev.*, 118 (12), 2752-2767.

Collins, W.G., L.S. Gandin, 1992: Complex quality control of rawinsonde heights and temperatures (CQCHT) at the National Meteorological Center. NMC Office Note 390. [Available from NCEP, 5200 Auth Road, Washington, D.C., 20233]

Collins, W.G., Woollen, J., 1993: An interactive system for the quality control of upper air observations at the National Meteorological Center. *Preprints, Ninth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology*, Annaheim, CA, 17-22 January 1993.

Daley, R., 1991: *Atmospheric data analysis*. Cambridge University Press, 457 pp.

DiMego, G.J., 1988: The National Meteorological Center regional analysis system. *Mon. Wea. Rev.*, 116, (5) 977-1000.

Gandin, L.S., 1969: On automatic quality control of current meteorological information. *Meteorol. Gidrol.*, No 3, 3-13 (in Russian).

Gandin, L.S., 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, 116 (5), 1137-1156.

Gandin, L.S., L.L. Morone, W.G. Collins, 1993: Two years of comprehensive hydrostatic quality control at the National Meteorological Center. *Wea. Forecasting.*, 8 (1), 57-72.

Gilcrist, B. and G.P. Cressman, 1954: An experiment in objective analysis. *Tellus*, 6 (4), 309-318.

Julian, P.R., 1989: Data quality control procedures at NMC Washington. *WMO/ECMWF Workshop. Data Quality Control Procedures*, 69-73.

Kalnay, E., R. Jenne, 1991: Summary of the NMC/NCAR reanalysis workshop of April 1991. *Bull. Amer. Meteorol. Soc.*, 72 (12), 1897-1904.

Morone, L.L., L.S. Gandin, W.C. Collins, 1992: Quasi-operational monitoring of the NMC complex quality control performance. *12th Conference on Probability and Statistics in the Atmosperic Sciences*, American Meteorological Society, Toronto, Ont., Canada, June 22-26, 1992, 305-309.

Schwartz, B.E, C.A. Doswell III, 1991: North American rawinsonde observations: problems, concerns, and call for actions. *Bull. Amer. Meteorol. Soc.*, 72, (12), 1885-1896.

Staff Members, Joint Numerical Weather Prediction Unit, 1957: One year of operational numerical weather prediction. *Bull. Amer. Meteorol. Soc.*, 38, (5), 364-369.

Thiebaux, H.J., M.A. Pedder, 1987: *Spatial objective analysis: with applications in atmospheric sciences.* Academic Press, 299 pp.

Woollen, J., 1991: New NMC operational OI quality control. *Preprints, Ninth Conference on Numerical Weather Prediction,* American Meteorological Society, Denver, CO, 10-14 October 1991, 24-27.

Table 3.1 CQCHT checks

Name	Applied to	Residual
Hydrostatic	Each layer between neighboring "complete" mandatory surfaces i.e., surfaces with neither height (z) nor temperature (T) missing	Difference between the layer thickness computed from heights of its boundaries and that hydrostatically computed from their temperatures. Also applied in terms of temperature
Baseline	Layer between the station level (z_s) and lowest reported mandatory surfaces	Difference between z_s in station dictionary and z_s hydrostatically computed from surface pressure p_s and heights z_1 and z_2 of lowest reported surfaces. Also applied in terms of p_s, z_1, and z_2
Incremental	Reduced mean sea level pressure, temperature and height of all mandatory surfaces, if (and where) the forecast first guess is not missing	Difference between the reported value (or reduced mean sea level pressure) and its first guess (called the increment)
Horizontal	Reduced mean sea level pressure, temperature and height of all mandatory surfaces, if (and where) the forecast first guess is not missing	Difference between the increment at the station and its value interpolated from four (or fewer) surrounding stations situated in different quadrants
Vertical	Temperature and height of all mandatory surfaces, if (and where) the forecast first guess is not missing	Difference between the increment at the level and its value interpolated from two surrounding levels (or, for the first and the last level, extrapolated from the neighboring level)

Table 3.2 Coefficients A and B

PRESSURE RANGE	A	B
1000-925	623.3	1.141
925-850	676.1	1.238
850-700	1552.3	2.842
700-500	2690.2	4.924
500-400	1784.1	3.266
400-300	2300.1	4.210
300-250	1457.7	2.668
250-200	1784.1	3.266
200-150	2300.1	4.210
150-100	3241.8	5.934
100-70	2851.7	5.220
70-50	2690.2	4.924
50-30	4084.2	7.476
30-20	3241.8	5.934
20-10	5542.0	10.145

Table 3.3 Admissible hydrostatic residuals

PRESSURE RANGE	admissable residual (m)	admissable residual in terms of temperature (K)
1000-925	65	57.0
925-850	65	52.5
850-700	35	12.3
700-500	50	10.2
500-400	35	10.7
400-300	40	9.5
300-250	35	13.1
250-200	40	12.3
200-150	50	11.9
150-100	85	14.3
100-70	70	13.4
70-50	70	14.2
50-30	80	10.7
30-20	70	11.8
20-10	100	9.9

Table 3.4 Admissible residual magnitudes (7 standard deviations)

	Height (meter)			Temperature (degree K)		
Pressure	increment	horizontal	vertical	increment	horizontal	vertical
1000	130	104	82	19	17	17
925	101	85	33	17	15	17
850	90	73	35	13	13	14
700	90	77	42	11	10	11
500	113	104	51	11	11	11
400	136	123	52	12	12	11
300	169	153	59	13	12	12
250	190	168	62	15	12	15
200	217	183	73	17	11	16
150	273	220	117	17	14	17
100	303	247	135	17	15	17
70	313	264	133	17	17	17
50	319	279	137	17	17	17
30	353	319	174	17	17	17
20	400	379	247	17	17	17
10	439	405	345	17	17	17

Table 4.1 Decision types.

Decision	**Meaning**
0	No error suspected
1	Data automatically corrected
2	Data suspected, found correct
3	Data questionable, not corrected
4	Data examined and found bad, not corrected
5	Baseline error of undetermined type

Table 5.1 Types of hydrostatically suspected errors.

Type	Suspicion
1	Communication, in Z_k $(2 \le k \le N\text{-}1)$
2	Communication, in T_k $(2 \le k \le N\text{-}1)$
3	Communication, in T_k and Z_k $(2 \le k \le N\text{-}1)$
4	Communication, in T_1 and/or Z_1, or computation of Z_2-Z_1
5	Communication, in T_N and/or Z_N
6	Computation of Z_{k+1} - Z_k $(2 \le k \le N\text{-}2)$
7	Communication, in Z_k and Z_{k+1} $(2 \le k \le N\text{-}2)$
8	Communication, in T_k and T_{k+1} $(2 \le k \le N\text{-}2)$
9	Communication, in Z_k and T_{k+1} $(2 \le k \le N\text{-}2)$
10	Communication, in T_k and Z_{k+1} $(2 \le k \le N\text{-}2)$
11	Like Type 1, but small
12	Hydrostatically proposed correction would lead to substantial super-adiabatic lapse rate
13	Data hole including 100 HPa surface
14	Data hole not including 100 HPa surface
22	Like Type 2, but small
99	Hydrostatically proposed corrections of Type 8, 9, or 10 would lead to substantial super-adiabatic lapse rate

Table 6.1 Types of baseline suspicions

Type	Suspicion
100	Surface pressure communication error
101	Height at bottom level when temperature is missing at this level
102	Undetermined baseline error
106	Surface pressure measurement error
116	Error in height of the lowest mandatory level

Table 8.1 Accumulated statistics-- July 1992 - December 1993 - - monthly mean numbers per observation time.

		CHQC		CQCHT	
		average	**std dev**	**average**	**std dev**
hydrostatic--	suspected	67.9	10.9	67.9	10.9
	corrected	28.4	4.2	51.5	11.0
baseline--	suspected	5.6	1.4	7.6	1.7
	corrected	-	-	5.0	0.7
holes--	detected	10.2	4.8	10.2	4.8
	corrected	-	-	2.0	0.6
observation--		-	-	47.9	5.9

Table 8.2 Relative statistics (in percent) on the overall numbers of various hydrostatic suspicions and on the DMA decisions for each suspicion type. The statistics are averaged over 3 randomly selected months.

Hydrostatic suspicion type		**Decisions**				
	Overall	**1**	**2**	**3**	**4**	**5**
1	17.1	96.0	0.9	2.2	0.9	0.0
2	18.2	86.5	4.6	5.8	3.1	0.0
3	17.2	55.7	0.0	22.6	8.5	13.2
4	9.0	72.9	13.0	9.9	4.2	0.0
5	20.2	47.5	1.8	6.4	4.7	39.6
6	3.8	63.7	5.3	3.8	7.6	19.6
7	1.9	59.9	14.6	19.4	5.4	0.7
8	1.2	90.6	5.7	2.9	0.8	0.0
9	0.9	81.7	11.5	4.5	2.3	0.0
10	1.0	68.0	20.2	11.1	6.9	0.0
11	5.5	26.5	55.7	12.8	5.0	0.0
22	4.0	43.5	25.9	25.2	5.4	0.0

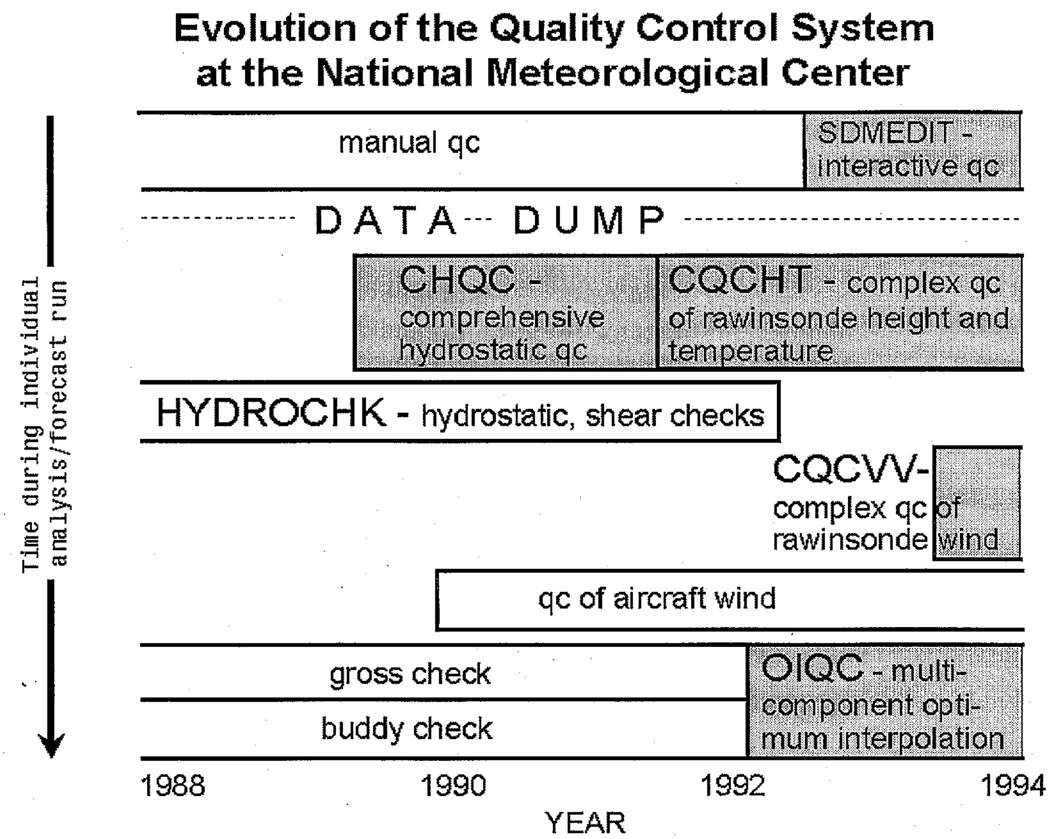

Fig. 1.1 Transformation of the NMC QC system.

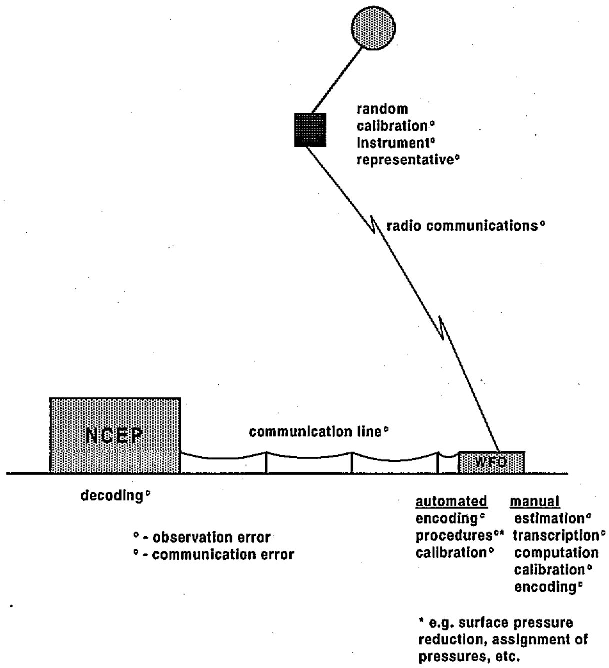

Fig. 2.1 Errors associated with rawinsonde observations and processing.

Figure 4.1 Operational output

heading

STN ID: 29282 **LAT**: 58.38 **LON**: 97.48 **EAST** **TIME**: 94/05/01/00
SCAN: 1

quick recognition table

	IINC		**IVOI**		**IHOI**		**ITMP**					
PRES	**Z**	**T**	**Z**	**T**	**Z**	**T**	**Z**	**T**	**IHSC**	**IBAS**	**IIPL**	**IHPL**
1000	0	0	0	0	0	0	0	0	0	0	0	0
925	0	0	0	0	0	0	0	0	0			
850	0	0	0	0	0	0	0	0	0			
700	0	0	0	0	0	0	0	0	0			
500	0	0	0	**2**	0	0	0	0	0			
400	0	**2**	0	**2**	0	**2**	0	0	**2**			
300	0	0	0	**2**	0	0	0	0	0			
250	0	0	0	0	0	0	0	0	0			
200	0	0	0	0	0	0	0	0	0			
150	0	0	0	0	0	0	0	0	0			
100	0	0	0	0	0	0	0	0	0			
70	0	0	0	0	0	0	0	0	0			
50	0	0	0	0	0	0	0	0	0			
30	0	0	0	0	0	0	0	0	0			
20	0	0	0	0	0	0	0	0	0			
10	0	0	0	0	0	0	0	0	0			

main body

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1003.	131.	-1.5	-2.5	-0.5	-3.9	3.7

	OBSERVATION		**INCREMENT**		**HYRES**	**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
1000	159.	2.2	-5.	3.8	----	----	-9.	2.6	-2.	1.9
925	782.	-1.3	7.	2.4	-1.	-1.2	7.	1.4	-5.	-0.5
850	1450.	-6.3	3.	-1.1	1.	1.1	1.	-1.9	-6.	-1.3
700	2952.	-12.3	-4.	-0.7	3.	0.9	-5.	-0.3	-16.	-1.2
500	5450.	-28.3	-1.	-0.6	8.	1.6	2.	-19.6	-25.	-1.0
400	7010.	**33.0**	-7.	**72.3**	-239.	**-73.3**	3.	**72.9**	-36.	**72.2**
300	8900.	-56.1	-25.	-2.0	-313.	**-74.3**	-13.	-16.1	-51.	-1.8
250	10050.	-60.5	-23.	0.9	3.	1.3	-7.	1.6	-43.	1.2
200	11450.	-58.1	-17.	-0.5	3.	1.0	8.	-0.3	-34.	-0.9
150	13280.	-53.7	-45.	-1.8	1.	0.1	-27.	-2.0	-46.	-0.5
100	15890.	-52.9	-34.	1.6	1.	0.1	-9.	1.9	-44.	0.6
70	18200.	-51.7	-27.	0.0	4.	0.8	-7.	-0.2	-37.	0.1
50	20390.	-51.3	-20.	-0.6	7.	1.4	-6.	-0.6	-37.	-1.0
30	23730.	-49.3	----	----	8.	1.1	----	----	----	----
20	26400.	-47.9	----	----	5.	0.8	----	----	----	----
10	31000.	-42.9	----	----	-21.	-2.1	----	----	----	----

DMA actions

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	400	T	2	**1**	33.0	**-72.0**	-39.0

Figure 5.1 Type 1 correction with a "small" increment.

STN ID: 35394 **LAT**: 49.80 **LON**: 73.13 **EAST** **TIME**: 92/04/31/12
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
300	8950.	-46.3	**91.**	3.3		64.	**47.**
					-104.		
250	10050.	-48.9	**8.**	3.3		**-69.**	**-32.**
					101.		
200	11610.	-50.7	**110.**	-0.5		63.	**74.**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	250	Z	1	**1**	10050.	**100.**	10150.

Figure 5.2 Type 1 suspicion not confirmed by statistical checks.

STN ID: 48698 **LAT**: 1.37 **LON**: 103.98 **EAST** **TIME**: 94/05/04/12
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**TEMP**	**TEMP**	**TEMP**
150	14250.	-67.5	-48.	-2.3		-2.6	-1.0
					-9.4		
100	16590.	-75.1	-67.	5.0		4.9	3.0
					-6.2		
70	18620.	-76.1	-44.	2.2		1.8	1.0

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	100	T	2	**3**	-75.1	**0.0**	-75.1

Figure 5.3 Type 6 error in 400--300-hPa layer. No errors in Part C.

STN ID: 58238 **LAT**: 32.00 **LON**: 118.80 **EAST** **TIME**: 92/12/29/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
500	5650.	-21.1	14.	-0.8		9.	5.
					6.		
400	7270.	-30.9	13.	0.9		-72.	8.
					202.		
300	9460.	-43.3	**223.**	-1.8		125.	**205.**
					7.		
250	10690.	-44.5	**221.**	-1.1		45.	**203.**
					14.		
200	12190.	-46.7	**232.**	2.6		50.	**205.**
					-3.		
150	14060.	-54.7	**261.**	3.5		76.	**209.**
					-7.		
100	16580.	-65.7	**278.**	-3.1		175.	**201.**
					-195.		
70	18530.	-69.7	35.	-6.5		-71.	-13.
					6.		
50	20580.	-61.5	15.	0.5		-3.	-51.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	300	Z	6	**1**	9460.	**-200.**	9260.
1	250	Z	0	**1**	10690.	**-200.**	10490.
1	200	Z	0	**1**	12190.	**-200.**	11990.
1	150	Z	0	**1**	14060.	**-200.**	13860.
1	100	Z	0	**1**	16580.	**-200.**	16380.
1	70	Z	6	**1**	18530.	**-200.**	18330.
1	50	Z	0	**1**	20580.	**-200.**	20380.
1	30	Z	0	**1**	23770.	**-200.**	23570.
1	70	Z	6	**1**	18530.	**0.0**	18530.
1	50	Z	0	**1**	20580.	**0.0**	20580.
1	30	Z	0	**1**	23770.	**0.0**	23770.

Figure 5.4 Type 3 corrections.

STN ID: 98223 **LAT**: 18.18 **LON**: 120.53 **EAST** **TIME**: 92/04/11/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
500	5860.	-5.9	3.	-0.5			94.	10.5	8.	0.1
					-63.	**-19.3**				
400	7380.	-55.7	**-196.**	**-39.0**			**-202.**	**-38.9**	**-186.**	**-39.**
					371.	**88.2**				
300	9680.	-32.5	15.	0.5			75.	8.5	37.	0.7

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	400	T	3	**1**	-55.7	**40.2**	-15.5
1	400	Z	3	**1**	7380.	**200.**	7580.

Figure 5.5 Type 3⇒1 correction.

STN ID: 42868 **LAT**: 21.10 **LON**: 79.05 **EAST** **TIME**: 94/05/10/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
925	728.	29.2	16.	3.0		126.	4.5	2.	1.7
					-395.				
850	1074.	23.4	**-381.**	-3.2		**-380.**	-4.2	**-398.**	-2.5
					413.				
700	3137.	11.0	19.	-1.3		178.	-0.2	-3.	-1.0

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	850	Z	3	**1**	1074.	**400.**	1474.

Figure 5.6 Type 3 errors with compensation.

STN ID: 44292 **LAT**: 47.93 **LON**: 106.98 **EAST** **TIME**: 93/07/15/12
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
850	1367.	17.2	4.	-3.5		-3.	-3.0	-9.	-1.9
					8.				
700	2990.	5.0	-1.	-1.6		15.	-3.8	-13.	-1.6
					-149.				
500	5600.	9.0	**-55.**	**19.0**		**-58.**	**19.4**	**-72.**	**19.0**
					3.				
400	7350.	-20.5	6.	-0.3		29.	-5.5	-11.	-0.6

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	500	Z	6	**5**	5600.	0.0	5600.

Figure 5.7 Type 5⇒1 correction.

STN ID: 24266 **LAT**: 67.55 **LON**: 133.38 **EAST** **TIME**: 94/05/04/12
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**
300	8720.	-52.5	15.	-0.2			0.	18.
					-2.	-0.9		
250	9900.	-50.7	**15.**	-0.2			76.	**22.**
					-193.	**-59.0**		
200	11170.	-47.7	**-183.**	-0.3			**-192.**	**-175.**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	200	Z	5	**1**	11170.	**200.**	11370.

Figure 5.8 Type 5⇒3 correction.

STN ID: 94996 **LAT**: -29.03 **LON**: 167.93 **EAST** **TIME**: 94/05/12/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
150	14070.	-65.3	-31.	-2.3		28.	-13.4	-22.	-1.9
					-577.				
100	16340.	-1.3	**-164.**	**72.4**		**-149.**	**72.8**	**155.**	**72.2**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	100	T	5	**1**	-1.3	**-70.1**	-71.4
1	100	Z	5	**1**	16340.	**142.**	16482.

Figure 5.9 Type 4⇒2 correction.

STN ID: 20744 **LAT**: 72.38 **LON**: 52.73 **EAST** **TIME**: 94/05/02/12
SCAN: 1

FULL VALUES		**SURFACE PRESSURE**		**BASELINE CHECK RESIDUALS**		
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1011.	19.	-0.6	-1.0	0.3	2.5	-2.2

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**TEMP**	**TEMP**	**TEMP**
1000	100.	-71.9	-4.	**-58.3**		**-58.4**	**-58.6**
					57.8		
925	690.	-15.1	12.	0.2		21.8	0.4

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	T	4	**1**	-71.9	**60.0**	-11.9

Figure 5.10 Type 4⇒6 correction.

STN ID: 46734 **LAT**: 23.57 **LON**: 119.62 **EAST** **TIME**: 92/04/29/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	107.	24.2	11.	-2.0		-53.	7.
925	---	---	---	---	**106.**	---	---
850	1615.	19.0	**108.**	-2.5		68.	**101.**
					-1.0		
700	3250.	10.2	**96.**	-1.0		16.	**82.**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	T	4	**2**	24.2	**0.0**	24.1
1	1000	Z	4	**2**	107.	**0.**	107.
1	850	Z	0	**1**	1615.	**-100.**	1515.
1	700	Z	0	**1**	3250.	**-100.**	3150.
1	500	Z	0	**1**	5970.	**-100.**	5870.
1	400	Z	0	**1**	7680.	**-100.**	7580.
1	300	Z	0	**1**	9760.	**-100.**	9660.
1	250	Z	0	**1**	11020.	**-100.**	10920.
1	200	Z	0	**1**	12480.	**-100.**	12380.
1	150	Z	0	**1**	14270.	**-100.**	14170.
1	100	Z	0	**1**	16670.	**-100.**	16570.
1	70	Z	0	**1**	18740.	**-100.**	18640.
1	50	Z	0	**1**	20750.	**-100.**	20650.
1	30	Z	0	**1**	23950.	**-100.**	23850.
1	20	Z	0	**1**	26570.	**-100.**	26470.
1	10	Z	0	**1**	31200.	**-100.**	31100.

Figure 5.11 Data hole, no errors.

STN ID: 71722 **LAT**: 46.38 **LON**: 284.03 **EAST** **TIME**: 92/04/13/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
400	6800.	-41.9	-44.	-1.6	↑		-35.	-0.6	-60.	-1.2

300	----	----	----	----			----	----	----	----
:	:	:	:	:	-122.	:	:	:	:	:
100	----	----	----	----			----	----	----	----
						-4.8				
70	18170.	-53.9	-30.	-1.1	↓		-13.	-0.2	-30.	-0.2

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	400	T	0	2	-41.9	0.0	-41.9
1	400	Z	0	2	6800.	0.0	6800.
1	70	T	13	2	-53.9	0.0	-53.9
1	70	Z	13	2	18170.	0.0	18170.

Figure 5.12 After-the-hole correction: type 13⇒4′⇒1 correction.

STN ID: 43346 **LAT**: 10.92 **LON**: 79.83 **EAST** **TIME**: 94/05/18/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
250	10990.	-42.7	10.	-4.1		-6.	17.
200	---	---	---	---		---	---
150	---	---	---	---	**-3135.**	---	---
100	---	---	---	---		---	---
70	15860.	-73.9	**-2804.**	6.0		**-2796.**	**-2819.**
					2723.		
50	20590.	-64.9	-16.	5.1		1424.	16.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	70	Z	13	**1**	15860.	**2700.**	18560.
2	70	T	13	**2**	-73.9	**0.0**	-73.9
2	70	Z	13	**2**	18560.	**0.**	18560.

Figure 6.1 Type 100 error.

STN ID: 47827 **LAT**: 31.63 **LON**: 130.60 **EAST** **TIME**: 94/04/15/00
SCAN: 1

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1020	283	27.1	28.5	30.3	254.2	-202.3

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	196.	11.8	-11.	-3.1		0.	-3.
					3.		
925	848.	10.4	-18.	0.7		-10.	-4.

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
1	1000	PS	100	**1**	1020.0	**-28.7**	991.3

Figure 6.2 Type 106 error.

STN ID: 26477 **LAT**: 56.38 **LON**: 30.60 **EAST** **TIME**: 94/05/01/00
SCAN: 1

FULL	VALUES	SURFACE	PRESSURE	BASELINE	CHECK	RESIDUALS
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1007.	98.	**10.4**	**10.8**	-0.1	-1.0	0.9

	OBSERVATION		INCREMENT		HYRES	VERTICAL	HORIZONTAL
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	158.	12.0	**101.**	-1.1		39.	**92.**
					5.		
925	812.	10.2	**98.**	-0.6		20.	**96.**
					-7.		
850	1500.	5.4	**95.**	-0.1		18.	**86.**

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
1	1000	Z	**106**	**5**	158.	**-89.**	69.
1	925	Z	0	**5**	812.	**-89.**	723.
1	850	Z	0	**5**	1500.	**-89.**	1411.
1	700	Z	0	**5**	3065.	**-89.**	2976.
1	500	Z	0	**5**	5630.	**-89.**	5541.
1	400	Z	0	**5**	7250.	**-89.**	7161.
1	300	Z	0	**5**	9220.	**-89.**	9131.
1	250	Z	0	**5**	10390.	**-89.**	10301.
1	200	Z	0	**5**	11790.	**-89.**	11701.
1	150	Z	0	**5**	13610.	**-89.**	13521.

Figure 6.3 Type 116 correction.

STN ID: 28952 **LAT**: 53.22 **LON**: 63.62 **EAST** **TIME**: 92/05/08/00
SCAN: 1

FULL PS	**VALUES ZS**	**SURFACE INCR**	**PRESSURE HOR-RES**	**BASELINE in Ps**	**CHECK in Zs**	**RESIDUALS in Z1**
990.0	171.	-1.5	-2.0	-9.0	**-76.**	81.

PRESS	**OBSERVATION HEIGHT**	**TEMP**	**INCREMENT HEIGHT**	**TEMP**	**HYRES HEIGHT**	**VERTICAL HEIGHT**	**HORIZONTAL HEIGHT**
1000	163	8.6	**65.**	4.0		14.	**66.**
925	---	---	**---**	---	8.	---	**---**
850	1500	-6.3	**86.**	3.8		21.	**81.**
					-5		
700	3040	-21.3	**104.**	3.9		18.	**101.**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	Z	116	**1**	163.	**-76**	87.
1	850	Z	0	**1**	1500.	**-76**	1424.
1	700	Z	0	**1**	3040.	**-76**	2964.
1	500	Z	0	**1**	5600.	**-76**	5524.
1	400	Z	0	**1**	7210.	**-76**	7134.
1	300	Z	0	**1**	9170.	**-76**	9084.
1	250	Z	0	**1**	10350.	**-76**	10274.
1	200	Z	0	**1**	11740.	**-76**	11664.
1	150	Z	0	**1**	13540.	**-76**	13464.
1	100	Z	0	**1**	16020.	**-76**	15944.

Figure 6.4 Type 101 error.

STN ID: 48820 **LAT**: 21.02 **LON**: 105.80 **EAST** **TIME**: 94/05/02/12
SCAN: 1

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
997.	9.	-1.9	-1.1	-153.4	-442.2	**460.0**

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	442.	----	**425.**	----		**467.**	**445.**

925	678.	33.0	-66.	6.0		-227.	-44.
					0.		
850	1430.	28.0	-26.	0.9		12.	-21.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	Z	101	**1**	442.	**-484.**	-42.

Figure 6.5 Type 4⇒3 correction.

STN ID: 97014 **LAT**: 1.53 **LON**: 124.92 **EAST** **TIME**: 92/05/07/00
SCAN: 1

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1002.0	80.	0.5	0.4	-10.5	-99.	**98.**

	OBSERVATION		**INCREMENT**		**HYRES**	**HYRES**	**VERTICAL**		**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**
1000	**0.**	**0.0**	**-89.**	**-30.4**			**-89.**	**-29.9**	**-89.**	**-30.**
925	---	---	---	---	**162.**	**68.2**	---	---	---	---
850	1504	17.8	1.	-1.5			**38.**	**8.0**	1.	-1.2

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	T	4	1	0.0	30.3	30.3
1	1000	Z	4	1	0.	90.	90.

Figure 6.6 Type 102 problem resolved after the z_{925} correction.

STN ID: 58150 **LAT**: 33.77 **LON**: 120.35 **EAST** **TIME**: 94/04/19/12
SCAN: 1

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1099.	7.	**89.0**	**89.8**	**78.9**	283.2	-126.9

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	82.	11.6	-21.	-2.9		233.	-6.
					-359.		
925	373.	11.6	**-401.**	-2.3		**-385.**	**-369.**
					371.		
850	1450.	12.6	-18.	-0.1		166.	-8.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	Z	102	**5**	82.	0.0	82.
1	925	Z	3	**1**	373.	**376.0**	749.

SCAN: 2

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
1099.	7.	**87.9**	**88.7**	**90.3**	746.	-334.4

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	82.	11.6	-21.	-2.9		-5.	-6.
					17.		
925	749.	11.6	-25.	-2.3		-9.	7.
					-5.		
850	1450.	12.6	-18.	-0.1		-4.	-8.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	Z	102	**5**	82.	0.0	82.
1	925	Z	3	**1**	373.	376.0	749.
2	1000	PS	**100**	**1**	1099.	**-89.**	1010.

Figure 6.7 Type 102 problem.

STN ID: 42875 **LAT**: 21.33 **LON**: 81.65 **EAST** **TIME**: 94/05/03/12
SCAN: 1

FULL	**VALUES**	**SURFACE**	**PRESSURE**	**BASELINE**	**CHECK**	**RESIDUALS**
PS	**ZS**	**INCR**	**HOR-RES**	**in Ps**	**in Zs**	**in Z1**
961.	298.	**-18.8**	**-16.7**	**-7.4**	**-70.2**	**143.9**

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
1000	1.	----	**-31.**	----		3.	-22.

925	718.	35.2	**-53.**	**6.0**		-38.	-11.
					5.		
850	1479.	29.0	-6.	0.2		16.	-8.

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	1000	Z	102	**5**	1.	0.0	1.

Figure 7.1 Vertically persistent observation errors.

STN ID: 62306 **LAT**: 31.33 **LON**: 27.22 **EAST** **TIME**: 94/04/19/12
SCAN: 1

	OBSERVATION		INCREMENT		HYRES	HORIZONTAL	
PRESS	HEIGHT	TEMP	HEIGHT	TEMP	HEIGHT	HEIGHT	TEMP
1000	-9.	----	12.	----		-20.	----

925	720.	36.6	-4.	**7.7**		-12.	**6.9**
					-12.		
850	1464.	28.2	27.	**5.3**		16.	**5.0**
					3.		
700	3136.	12.8	51.	2.5		38.	2.3
					4.		
500	5850.	-8.7	**102.**	**6.2**		**83.**	**5.8**
					-4.		
400	7530.	-21.9	**131.**	**4.7**		**107.**	**3.9**
					7.		
300	9600.	-34.3	**188.**	**6.7**		**156.**	**5.7**
					1.		
250	10850.	-43.9	**226.**	**6.5**		**186.**	**5.4**

200	12310.	----	**257.**	**----**		**215.**	**----**
					-40.		
150	14150.	-55.7	**306.**	**7.2**		**257.**	**4.9**
					3.		
100	16720.	-58.1	**406.**	**8.7**		**391.**	**5.2**

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
2	300	Z	0	**4**	9600.	0.0	9600.
2	250	Z	0	**4**	10850.	0.0	10850.
2	200	Z	0	**4**	12310.	0.0	12310.
2	150	Z	0	**4**	14150.	0.0	14150.
2	100	Z	0	**4**	16720.	0.0	16720.

Figure 7.2 Observation errors in a part of the sounding.

STN ID: 72747 **LAT**: 48.57 **LON**: 266.62 **EAST** **TIME**: 93/02/08/12
SCAN: 1

	OBSERVATION		INCREMENT		HYRES	HORIZONTAL	
PRESS	HEIGHT	TEMP	HEIGHT	TEMP	HEIGHT	HEIGHT	TEMP
500	5550.	-19.3	24.	-0.9		13.	0.0
					-4.		
400	7190.	-23.5	34.	**7.9**		33.	**8.5**
					16.		
300	9270.	-32.5	**153.**	**16.6**		**156.**	**16.8**
					0.		
250	10530.	-41.5	**244.**	**16.6**		**246.**	**17.7**
					4.		
200	12010.	-52.7	**347.**	**7.8**		**347.**	**6.4**
					-11.		
150	13810.	-63.5	**321.**	**-7.8**		**323.**	**-8.9**
					-9.		
100	16350.	-53.3	**274.**	**3.6**		**276.**	**4.2**
					11.		
70	18650.	-54.5	**343.**	**5.3**		**341.**	**6.0**
					-1.		
50	20780.	-59.1	**371.**	-0.1		**374.**	0.5

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
2	400	T	0	**3**	-23.5	0.0	-23.5
2	300	T	0	**3**	-32.5	0.0	-32.5
2	250	T	0	**3**	-41.5	0.0	-41.5
2	250	Z	0	**4**	10530.	0.0	10530.
2	200	T	0	**3**	-52.7	0.0	-52.7
2	200	Z	0	**4**	12010.	0.0	12010.
2	150	T	0	**3**	-63.5	0.0	-63.5
2	150	Z	0	**4**	13810.	0.0	13810.
2	100	Z	0	**4**	16350.	0.0	16350.
2	70	Z	0	**3**	18650.	0.0	18650.
2	50	Z	0	**3**	20780.	0.0	20780.

Figure 7.3 Probable first guess errors.

STN ID: 70219 **LAT**: 60.78 **LON**: 161.80 **WEST** **TIME**: 94/04/25/00
SCAN: 2

	OBSERVATION		**INCREMENT**		**HYRES**	**VERTICAL**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**TEMP**	**TEMP**	**TEMP**
300	8860.	-54.5	56.	0.1		-1.3	-0.1
					1.7		
250	10030.	-48.7	52.	2.9		2.7	1.3
					-4.6		
200	11510.	-46.3	63.	0.7		-0.6	0.3
					1.9		
150	13430.	-44.7	81.	3.0		2.1	1.9
					0.7		
100	16140.	-43.7	**118.**	**4.8**		3.3	2.6
					-1.0		
70	18520.	-46.5	**186.**	**5.8**		4.2	2.6
					0.0		
50	20760.	-46.5	**223.**	**3.9**		2.6	2.8

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
2	50	Z	0	**3**	20760.	0.0	20760.

Figure 7.4 Communication and observation errors. Corrected T_{300} rejected by scan 2.

STN ID: 42314 **LAT**: 27.48 **LON**: 95.02 **EAST** **TIME**: 94/05/11/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**TEMP**	**HEIGHT**	**TEMP**
400	7550.	-11.9	70.	4.2		48.	3.4
					-42.4		
300	9710.	21.0	**114.**	**49.0**		**89.**	**49.0**
					-39.6		
250	11040.	-29.3	**159.**	7.8		**136.**	8.5
					-1.1		
200	12590.	-41.3	**204.**	7.0		**179.**	7.4

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	300	T	2	**1**	21.0	**-42.0**	-21.0

SCAN: 2

	OBSERVATION		**INCREMENT**		**HYRES**	**HORIZONTAL**	
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**TEMP**	**HEIGHT**	**TEMP**
400	7550.	-11.9	70.	4.2		48.	3.4
					-0.4		
300	9710.	-21.0	**114.**	**7.0**		**89.**	**7.0**
					2.4		
250	11040.	-29.3	**159.**	**7.8**		**136.**	8.5
					-1.1		
200	12590.	-41.3	**204.**	**7.0**		**179.**	7.4

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	300	T	2	**1**	21.0	**-42.0**	-21.0
2	300	T	0	**4**	-21.0	0.0	-21.0
2	250	T	0	**4**	-29.3	0.0	-29.3
2	200	Z	0	**4**	12590.	0.0	12590.

Figure 7.5 Corrections not made because of observational errors.

STN ID: 42701 **LAT**: 23.32 **LON**: 85.32 **EAST** **TIME**: 94/05/16/00
SCAN: 1

	OBSERVATION		**INCREMENT**		**HYRES**	**HORIZONTAL**
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**
700	3134.	9.8	35.	0.2		29.
					89.	
500	5940.	-4.3	**151.**	**3.4**		**146.**
					14.	
400	7670.	-16.5	**167.**	**-1.9**		**172.**

DMA RESULTS

SCAN	**PRESS**	**VARIABLE**	**IHSC**	**DECISION**	**OLD VALUE**	**CORRECTION**	**NEW VALUE**
1	500	Z	6	**5**	5940.	**0.**	5940.
2	500	Z	6	**4**	5940.	**0.**	5940.
2	400	Z	0	**4**	7670.	**0.**	7670.

Figure 7.6 Multiple errors corrected by two scans.

STN ID: 91610 **LAT**: 1.35 **LON**: 172.92 **EAST** **TIME**: 93/02/14/12
SCAN: 1

	OBSERVATION		INCREMENT		HYRES	VERTICAL	HORIZONTAL
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
300	9630.	-33.9	-30.	-3.3		-92.	-37.
					208.		
250	11090.	-43.1	**160.**	-3.1		604.	**153.**
					-1321.		
200	11240.	-52.7	**-1176.**	-0.7		-255.	**-1180.**
					-1609.		
150	11420.	-68.7	**-2792.**	-0.8		-720.	**-2791.**
					-2081.		
100	11650.	-88.1	**-4875.**	-11.4		-3519.	**-4874.**

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
1	250	Z	3	**1**	11090.	**-157.**	10933.
1	200	Z	3	**1**	11240.	**1178.**	12418.
1	150	Z	3	**1**	11420.	**2781.**	14201.

SCAN: 2

	OBSERVATION		INCREMENT		HYRES	VERTICAL	HORIZONTAL
PRESS	**HEIGHT**	**TEMP**	**HEIGHT**	**TEMP**	**HEIGHT**	**HEIGHT**	**HEIGHT**
300	9630.	-33.9	-30.	-3.3		-25.	-37.
					5.		
250	10933.	-43.1	3.	-3.1		14.	-4.
					51.		
200	12418.	-52.7	2.	-0.7		5.	2.
					14.		
150	14201.	-68.7	-11.	-0.8		1582.	-10.
					-6.		
100	11650.	-88.1	**-4875.**	**-11.4**		-4870.	**-4874.**

DMA RESULTS

SCAN	PRESS	VARIABLE	IHSC	DECISION	OLD VALUE	CORRECTION	NEW VALUE
1	250	Z	3	1	11090.	-157.	10933.
1	200	Z	3	1	11240.	1178.	12418.
1	150	Z	3	1	11420.	2781.	14201.
2	100	T	5	**1**	-88.1	**5.9**	-82.2
2	100	Z	5	**1**	11650.	**4897.**	16547.