U.S. DEPARTMENT OF COMMERCE NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION NATIONAL WEATHER SERVICE NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 212

A Review of Developments for Monitoring and Controlling the Timeliness, Accuracy, and Completeness of Data for Meteorological Analysis

> James Kerlin* Lt. Colonel, USAF

Arthur Thomas Climate Analysis Center

MARCH 1980

*Air Force Liaison to the National Meteorological Center, NWS/NOAA

This is an unreviewed manuscript, primarily intended for informal exchange of information among NMC staff members.

Abstract

Acquisition of timely, high quality data from reasonably dense evenly distributed observing points is essential to production of accurate sequential analyses. Establishment and continuous operation of such networks are a demanding task in light of national interests, economics, technology, and geopolitical factors.

Apart from the important aspects of international cooperation and comprehensive network design, increased automation at major world weather analysis centers has introduced a unique set of problems associated with total procurement, utilization, and onward dissemination of data. Due to computer sensitivity and a rigid requirement for inter-system compatibility, practically every meteorological analysis produced is degraded to some extent by the absence of essential data. At times the losses of these data have direct implications on the accurate depiction of the state of the atmsophere.

Detection, investigation, and resolution of data deficiencies is essential to optimum system operation but is, at best, a complex and involved undertaking. Cause/effect relationships of data deficiencies must be determined and then followed by decision processes appropriate in dealing with the characteristics of the deficiency. Objectivity is essential in determining the reasons for data deficiencies in order to avoid initiating inappropriate action which may be resented.

A review is presented of the latest procedures for evaluating data in terms of timeliness, accuracy, and completeness. Recent developments of procedures and techniques for systematically appraising the distribution and quality of global data is discussed. New involvements of the WMO in comparing the availability and quality of data at each major meteorological center throughout the world are taken into account. Procedures developed and implemented by the National Meteorological Center and Air Force Global Weather Central for minimizing data discrepancies and losses are presented.

INTRODUCTION

Quality control in meteorology has traditionally emphasized the taking of accurate observations, proper encoding of the observations and reliability in transmitting the observations. The U.S. and most other meteorological services throughout the world expend a lot of effort on the checking and validation of observational records. This form of quality control has, despite its limitations, accomplished a great deal in ensuring that archived data are made about as correct as possible. This fills an important need by creating consistently accurate information for future climatological studies and other research needs. The process of performing comprehensive post-facto data validation also carries with it the ability to feed back to the data sources information concerning inaccurate observations and incorrect procedures.

Despite their value, such methods of quality control have serious shortcomings in bringing about significant continuing improvements in data for the purpose of operational analysis and subsequent forecasting (prognostication). Perhaps the biggest drawback is the extensive delay involved in informing data sources of observational deficiencies- generally on the order of several months. These delays do little to impress observers about the potential impact on operations of "accidental" errors which are by far the most prevalent. Moreover, record checking fails by and large to take into account most errors that occur during the transmission of data. Also, due to budgetary restraints in recent years, these quality

control programs have become increasingly more limited. Whereas some data centers were formerly heavily staffed and quality-controlled every observation, in recent years the trend has been toward spot checking and using statistical techniques to evaluate performance. More complete quality control of full data sets is then performed by data users.

3

In the mid-60's the entire concept of data quality control began to change. A system was proposed which would result in permanently improving data for operational use and also for archiving data for subsequent applications. This program was named the World Weather Watch. Of course, the World Weather Watch is more than simply the quality control of data. "Its primary purpose is to ensure that all World Meteorological Organization Members obtain the meteorological information they require for operational and research purposes. The conversion of this theoretical concept into practical proposals required a thorough re-appraisal of the world weather system and preparation of new plans incorporating all modern scientific and technological developments." (WMO, 1968).

Monitoring and quality control are an integral part of this concept. From the early stages of the World Weather Watch came a series of Planning Reports one of which was No. 26 - Quality Control Procedures for Meteorological Data. This comprehensive document formed the basis for much of what was to evolve in World Meteorological Organization quality control activities. WWW Planning Report No. 26, like its companion planning reports, became predecessors of later written guides and manuals on the Global Observing, Telecommunications, and Data Processing Systems. The U.S., as a member country, has played an important role in the development, extension, and implementation of these concepts. This discussion centers on the purpose and application of these developments.

REQUIREMENTS

While meteorological interests are broadening geographically, they are at the same time becoming more precise scale-wise. This type of expansion causes the requirements for data which are needed to describe the state of the atmosphere to increase exponentially. Thus, deficiencies increase proportionally.

The most urgent requirement for high quality data exists within the National, Regional, and World Meteorological Centers, as defined by the World Meteorological Organization. Within the United States, Organizations such as the National Meteorological Center, the Fleet Numerical Oceanography Center, the Air Force Global Weather Central, and other comparable centers have equally important needs for the procurement of accurate data. These organizations can ill-afford to allow erroneous data to pass into their data bases where they are apt to mislead an objective analysis and deteriorate a forecast. More importantly the centers must be judicious about depriving their analyses of too much "suspect" data. This in itself "waters down" an analysis and destroys resolution. Also, the cost of taking and communicating the data necessitates optimum use of the data.

A study prepared by the Bureau of Census (1977) lists primary activities supported by the National Weather Service's products in the organizations survey. These activities include Agriculture, Airlines, Airport Operations, Air Services, Automobile Clubs, General Aviation, Commerce and Finance, Consultants, Federal Government, Local Governments, State Governments, Manufacturing, Marine, Mineral Processing, Museums, Press Services,

Private Meteorologists, Railroads, Research and Development, Retailing, Schools, Research Institutes, Service Companies, Television and Radio, Trucking Companies, and Utilities. Many of these data recipients have to either forfeit data they are unable to control or endeavor to control data without the necessary tools or skills.

The World Weather Watch's concept for quality control employs the total systems concept. It has provisions for monitoring and quality control at the observing point, within the telecommunications phase, and at the meteorological centers. It includes provisions for real-time and non-real-time activities. It includes both the monitoring and the quality control functions. Logical objectives of the monitoring and quality control effort are to improve the performance of the World Weather Watch on a national, regional, and global level.

Cooperative non-duplicative quality control programs are carried out by both the National Meteorological Center and the Air Force Global Weather Central (AFGWC, 1979). Because both centers exchange considerable amounts of data, each benefits from the activities of the other. These activities are among the most advanced in the world and are directed toward:

A. Salvaging data which would otherwise be lost.

B. Developing statistical information for management analysis.

C. Informing data sources in real-time of data deficiencies (e.g. missing data) and discrepancies (e.g. incorrect values).

Obtaining high quality data from the worldwide meteorological observing network is necessary for the production of accurate analyses of the atmosphere worldwide.

Research into improving numerical weather prediction is being conducted at national centers. These activities center on improving not only the physics of the models but also the horizontal and vertical scales of motion being dealt with. The keys to improvement lie not only with better NWP forecast models but with a better (i.e. more accurate) depiction of the initial state of the atmosphere. As more sophisticated techniques develop to use the data, the data in turn must be more accurate or an incorrect interpretation of the state of the atmosphere may result.

In addition to dealing with the traditional twice-a-day 00Z and 12Z observing cycles there are now, and have been for some time, 6-hourly data analysis and forecast cycles at the AFGWC and the NMC. As the trend to deal with smaller scale features continues, more emphasis is being put on synoptic and asynoptic data both from conventional and remotely placed sensors.

As an example of the volume of data we are talking about, the worldwide rocketsonde network provides 25 observations each week. The radiosonde network reports every 12 hours for 650 stations, and every 6 hours for 150 stations. The surface observing network has at least 8000 stations reporting every 3 hours. In addition there are during a 24 hour period a considerable number of aircraft reports, ASDAR reports, constant level balloon reports, and satellite soundings.

All of these types of observing systems have their own characteristics and methods of observing data and encoding reports. How these data are now dealt with and what to look for in the future will greatly impact on the progress that will be made in weather forecasting in the future.

The Air Force and the Navy utilize radio broadcast weather intercept sites to collect data broadcast by other nations that is neither available

from other sources nor as timely. This intercept system collects 60 to 70% of the Eastern Hemisphere data used by the automated weather network. Granted that some of these data may be available through the GTS, it is still necessary to intercept them by radio since any GTS data usually go through a number of processing or switching centers before reaching the NMC or the AFGWC.

In addition to the intercept system other unique methods of obtaining data include use of civil contractors such as Aeronautical Radio Inc. (AIRINC) who supply AIREP data and Braniff Airlines who provide about half of the data from South America.

PROCEDURES

The quality control of data starts with the instrument at the observing point and ends when the information is delivered to a user. The stages at which errors can and do occur include instrument manufacture, instrument installation, instrument deteoriation, visual observing, reading of instrument, reading of observation, instrument corrections, processing of an instrument's signal, coding of the observation, and transmission of the observation (WMO, 1968).

Since these types of causes of error are normally not apparent to the user of the data, these errors have to be kept in mind when processing the data. Some help is provided by processing centers such as the NMC and the AFGWC by their ability to summarize the errors of individual stations and feed these results back to the station (Figures 3, 4, 5, 6, and 7). The transmission of data through communications channels remains as a large source of errors.

The checking of data as to its acceptability at a processing center usually is derived from experience. This checking deals with first

identifying the data, then decoding it, and then checking its meteorological goodness.

According to the WMO "the purpose of quality control of observational data is error detection, possible error correction, and therefore error prevention in order to ensure the highest possible standard of accuracy for the optimum use of these data by all possible users" (WMO 1977). The WMO Minimum Standards for Quality Control of Incoming Data described in the WMO Manual of the GDPS (1979) include real and non-real time procedures.

Frequent exchange of information between WMO centers is used to improve the quality of data. This exchange of information is accomplished by the use of monitoring reports. The reports are important because the primary responsibility for quality control, with the aim of error detection and correction of all observational data, should rest with the national meteorological services from which the observations originated (WMO, 1977).

Some general methods for quality control recommended by the WMO and used in some form by providing centers include internal consistency checks, qualitative logical inferences, time and space consistency, physical limitations, statistical checks, and empirical checks.

Internal consistency might include looking at cloud types compared to cloud height or high relative humidity compared to the weather reported. A logical inference could compare a visibility with an obstruction to vision or wind speed. Time and space consistency can check past and present aircraft location reports, or pressures with pressure tendency. Physical limits are checks against a standard such as climatological extremes. Statistical and empirical checks deal with such relationships.

PROGRAMS

Quality control is carried out in basically two ways. The first involves interception and identification of discrepancies and either correction or rejection of data prior to their use within the centers. This is primarily real-time. The second form is in the issuance of information to data sources concerning discrepancies that have been encountered. This procedure is carried out in real-time, and in non-real-time on a weekly, monthly, and semi-annual basis. Figure 1 is a generalized diagram of the process.

The procedures used to perform quality control of data vary with the type of data. In general, the data flow and quality control influences are basically the same. Figure 2 is a generalized scheme of the typical data flow in a major processing center. A number of communication lines connect to a high speed processor. As the meteorological bulletins flow into the computer the bulletin headings are compared to those in a master directory. If a match-up is found, a check is made between the circuit the bulletin was received on and the one it was supposed to be received on. If they are the same and the bulletin is one scheduled to be relayed elsewhere, it is retransmitted onto the circuits designated in the directory. If the circuit on which the bulletin was received was not the one designated, the bulletin is not relayed. However, its contents are stored for analysis at the center. If the bulletin heading agrees with none other in the directory, the bulletin is deviated for electronic display and manual inspection.

If the inspector can determine that the bulletin heading is in error and can resolve the discrepancy based on the bulletin contents or clues in the heading itself, the editor corrects the heading. The bulletin is then re-entered into the computer and re-cycled through







¢.,

Figure 2. Data Control Real-Time



the system as though it was entering from an external circuit. The success rate of this quality control procedure is quite good in salvaging bulletins with discrepant identifications. The remaining ones, generally with completely indistinguishable identifications and contents, are discarded.

A record is kept of unknown bulletins. If they are received repeatedly and frequently they are identified and added to the master directory. This important aspect of quality control accounts for preventing the loss of thousands of pieces of data daily. However, it is not completely fool-proof. There is nothing to prevent a communications source from assigning a legitimate yet incompatible bulletin heading to a message. For example, a collection of upper air data could be transmitted with a heading reserved for terminal forecasts. Although this doesn't appear to happen very often, it does happen. When it occurs, the bulletin flows into and through the system automatically and the deficiency will remain undetected. As a result, some clients will not receive their data, others will get something they probably don't need, and the data will be lost from the analysis. The solution to this problem lies in automatically comparing bulletin headings with their contents to ensure agreement; something presently being done by the AFGWC.

The bulletin heading identifies the contents, originating country or region, originating communications point, and the date and time of either the transmission or when the contents of the bulletins were assimilated. As the various bulletins flow into the system, they are directed to specified files based on their heading.

For the sake of brevity and to simplify the discussion we will trace two major types of data, TEMP (upper air) and SYNOP (synoptic) through the system and describe the process of internal and external quality control.

It is important to note that TEMP data include a time of observation in each individual report but SYNOP data have the observation time only in the heading. Also, data collections often come in from far distant countries late in synoptic periods and sometimes during an ensuing period. This can compound problems more with SYNOP data which have synoptic times separated by only six hours than with TEMP data which are usually analyzed on a 12-hour cycle. Late receipts of SYNOP data are more apt to spill into the next collection period than are late receipts of TEMP data. This precludes most automated tests of SYNOP data for synopticity. They must be taken at face value initially. TEMP data on the other hand have considerable redundancy between the bulletin headings and individual reports.

General Characteristics of TEMP and SYNOP Data

			SYNOP	TEMP
Number of Parts	to a Report		1	4
Similarity with	Intermediate	Reports	yes	yes

The first four characters of each TEMP report can be matched against the first two characters of the bulletin heading for consistency. The time included in each report can be compared to the bulletin heading for comparability. Each level in each TEMP report has a fixed and identified location which must be accounted for. If an extra data group is included or one is omitted, the sequence of levels is interrupted and this can be automatically detected and dealt with.

When discussing quality control of meteorological data it is important to note that the style and format of several types of telecommunicated

reports have remained essentially unchanged since their introduction during the 1940's and 50's. The code forms were designed for efficiency in operations other than computerized ones. Brevity of communications, speed in manual interpretation and plotting, and simplicity for extracting and archiving of data were some of the paramount considerations in the early development. Even when evaluations of coding practices began to take place during the 60's and 70's, automated processing could not always be the principal consideration. The world is still made up of dual operations, automated and manual, even in so-called developed or technologically advanced countries. Greater emphasis is now being placed on designing code forms more suitable for automated processing and analysis.

A typical example of the conflict involved in designing data formats which can easily be processed automatically with a low incidence of misinterpretation by a computer program and yet be easily read and understood by a person are the hourly aviation data. The user community, aviation interests in this case, want frequent and expeditious availability of accurate fine-scale analyses and forecasts. Yet they also want the individual reports for quick personal updates. Thus the data must be readable by a person and processable by a computer with a low failure rate in both regards. Existing hourly data, although reasonably simple to read and translate, are not easily encoded, are highly susceptible to error either in preparation or transmission, and cannot be easily programmed for computer processing. The reports are not in a fixed format, they are a blend of alphanumeric characters, and elements are not uniquely identified.

Despite these shortcomings, considerable progress has been achieved through extensive and complex programming to automatically process hourly

reports. Unfortunately, the error rate in hourly data remains relatively high thereby increasing the risk factor in misinterpreting parameters.

SYNOP data, although essentially in a fixed format, have few distinctive identifiers for specific groups. Indeed, the data groups most essential for operations; those which are compulsory for foreign exchange, have none. This means that if a data group is inadvertently or purposely added to or omitted from the report, all ensuing data are apt to be misinterpreted. SYNOP data cannot stand alone and readily be identified as to source, time or type.

TEMP and pilot data, although not perfectly formatted for computer operations, are the least difficult to handle with automatic processors. Each report is a self-contained entity, independently identifiable as to type, time, and origin.

Though not completely infallible, upper air data lend themselves most easily to automated processing. They are in a fixed format. They are almost completely numeric except for their initial identifier group. Data sets within the reports have double-digit identifiers throughout. If a processor "stumbles" on a discrepancy, it can key on a subsequent identifier in order to prevent loss of a full data set.

Incidently, when a discrepancy does occur the computer can be programmed to register and store it. Moreover, upper air data can be tested for hydrostatic and vertical consistency with a high degree of accuracy. Surface data can be exposed to rudimentary and limited internal consistency and time consistency checks but these techniques impose heavy storage and processing demands on the on-line operational systems.

The National Weather Service and the Air Weather Service are proceeding

on ambitious and progressive data monitoring and quality control programs. Three major aspects of data gathering and utilization are being monitored and quality controlled: observing, telecommunications, and processing.

The Air Force Global Weather Central and National Meteorological Center both perform in-house man-machine quality control. They disseminate and exchange real time notifications of data deficiences. Jointly or separately they disseminate weekly, monthly, and semi-annually produced appraisals of data receipts, losses, and quality on a local, regional, national, and international basis. They also perform special studies to investigate and diagnose chronic or persistent problem areas and initiate action to get such problems corrected.

Program developments are shared between the centers in order to provide mutual benefits and minimize duplicated activities. Parallel analysis of data acquisition successes and failures which are essential for isolating major system problems are exchanged between the centers. The quality control activities of these centers result in many data subscribers obtaining data which they would otherwise not receive.

The AFGWC communications hub at Carswell AFB, Texas automatically screens incoming data. Garbled or improperly formatted reports are directed to a manual monitoring position and are repaired and re-entered into the data base. The data base is periodically scanned for missing reports. The scan frequency is a function of data type. Hourlies are scanned at least twice each hour; three- and six-hourly synoptics less frequently, and six- or twelve-hourly upper-air data least frequently.

NMC scans its six-hourly synoptic data base four times a day, three hours after observation time. NMC also interrogates its upper air data base four times during each 12-hour cycle.

Each center transmits real time advisory messages informing data sources if their report has not been properly received. The purpose of these messages is to advise data sources that necessary data are unavailable at the center and a retransmission from the originator is desirable. They are not intended to discredit a station for poor performance because the data loss is often beyond the originator's control. Unfortunately, when chronic or persistent problems prevail, this unavoidably and understandably becomes the net result. Figures 3 thru 7 are examples of the advisory bulletins issued by the centers.

The major difference between the quality control method used by the NWS and the AWS as compared to those of other centers is the emphasis on action in real-time to minimize discrepancies.

The NMC and the AFGWC Detachment 7 at Carswell AFB, Texas, both monitor incoming traffic for discrepancies. Recognition codes are programmed into the communications computer to identify bulletin types. If recognition is not accomplished or, at Carswell, if an inconsistency exists between the bulletin heading and its contents, the bulletin is directed to a human monitor/quality control/edit function for correction and is re-entered into the system.

Bulletins or data which are so badly mangled that they cannot be confidently corrected are discarded. Bulletins or data which are "clean" but cannot be identified are flagged and stored. Repetitive receipts of the same unknown message or contents triggers an effort to determine legitimacy and to identify the source. When these things have been verified, the item becomes part of the recognition code.

Applications

Figure 3 shows an NMC advisory of unreceived SYNOP reports. It is transmitted at about 2 1/2 to 3 hours after each main synoptic observation time. These advisories are addressed to the country of data origin for subsequent procurement and retransmission of the report by the data source or the individual national meteorological center. NMC disseminates these message to all countries throughout the Western Hemisphere.

Figure 3

2C,2C NBC942 AANX60 KWBC 270844

FOLLOWING AURFACE REPORTS FROM HE INDICATED **HND** STRIIONS HAVE NOTE BEEN SHE SSFUL $V \subseteq O$ F 6 PROCESSED IN THE UNC **URGHINGTON** RE TRANSMIT THEN AS SOON AS POSSIBLE.

270600

76151 76323 76471 76525 76625 76632 76644 76647 76675 76685 76692 76695 76737 76743 76775 76833

MASHINGTON MONITOR

Figure 4 is a composite example of a series of advisories transmitted by the NMC Washington to upper air units throughout North America. The NMC advisory identifies the nature of the data discrepancy by:

a. Data source.

b. Date/time deficiency was noted.

c. Time of observation.

d. Circuit.

e. Non-receipt by PART A, B, C or D.

f. I.D. discrepancy by PART.

g. Incomplete PART A and identification of the first missing isobaric surface.

h. Text error by PART, level, and parameter.

These messages are transmitted at least four times every twelve hours.

Figure 4

	ZCZC WB AAXX48	C466 KWBC 2	10921		ODV (0.07				· · ·	Ļ
	UPPER A	TR REC	ETLL.	TDA TC	SORI (JUZ .					
	AAAA CTUA TD	NTD	тъ	τci	תעיבית	п					1
1	DIA ID	NR.	ΤD	τu	.T. T.	Ľ					11
	70086	CD	A	15	99W	OOH	OOD				
	72528	B D	1 - E								
1	74486		А		99P	OOH	85H				
į	76256			20							
	76612	ABCD									
	78016				B11	B12	B1 3	C22	C23		
	78367		в		50W	40H	30H	07T	07W	05H	i
	78367		В		50W	812 40H	30H	07T	07W	05H	

Figure 5 is an example of an AFGWC real-time message notifying synoptic reporting stations of unreceived or unidentified SYNOP reports. These are updated and retransmitted several times during each synoptic period.

Figure 5

ZCZC W8C773 AXXX15 KAWN 211840 STATNS NOT RCVD KAWN 71081 71094 71911 71919 71925 71939 99021 STAINS SURVYD 48 REPORT MISG 7 INT ZDK

Figure 6 is an example of an AFGWC real-time message notifying upper air stations of unreceived or unidentified PART A TEMP reports. Similar advisories are transmitted for PARTS B, C, and D of the TEMP reports and for all PARTS of PILOT (winds aloft) reports. As with SYNOP advisories, these are updated and retransmitted periodically during each 12-hour period. This advisory is for stations in North, South, and Central America, the Caribbean, and Antartica.

Figure 6

ZCZC WBC985 AXXX61 KAWN 220320 220000 TTAA RAOB DATA NOT RCVD KAWN 70026 71926 72385 76723 78016 78367 78501 78954 80222 84628 87047 87418 87715 89664 99008 STATNS SURVYD 156 REPORT MISG 15 Figure 7 is an example of an AFGWC teletype message providing summary statistics of the receipt of PART A TEMP data. This message is updated and transmitted several times each cycle. The column showing "Reports Expected" is based on the recent history of data availability from each block, not on the published WMO schedule of operations. In this regard it is useful in planning meteorological operations such as objective analysis schedules.

Fi	g	u	r	e	7
----	---	---	---	---	---

					· · · · · · · · · · · · · · · · · · ·		
7676	MBC902	e di seria a					
RULI	YANN 22	0150					
1 2200	007 7793	2010 00	UN 1999	1.0			
SIV.	9010 9010	PETC	PCNT FYPETD	81.8	RPTS	RPTS	PONT EXPOTE
100	CUDENS	2045	PEND	82	EXPCTO	RCVD	RCVD
04	- 002	000	000	62	0.05	697	\$77
6.5	666	000	100	0.4	007	003	042
60	605	0.05	083	87	0.04	006	150
, VO 60	000	667	666	69	0.03	0.02	066
- 100 - 4 a	000	005	470	43	607	0.03	100
: <u>L</u> V :`	005	664	140	43	602	682	100
11	603	002	640	- 45	005	666	120
10	0.04	000	000	20.	0.00	0.02	025
27	003	000	9.6.V 4.6.6	20	0.05	004	055
61	200 U	0.00	000	51	010	667	67.6
23	. 008	600	100	· 24 67	614	607	053
25	V1 V	00Z	020	20	007	6.05	071
27	963	683	NSI	60	001	000	14
29	003	004	844	. 60	610	600	400
31	014	007	150	32	003	000	100
33	011	883	027	34	000	882	600
35	009	.005	.055		603	004	000
37.	885	\$\$\$	000	- 38	015	0.03	020 -
40	014	0.01	007	42	885	000	000
43	004	000	000	44	000	000	000
45	001	000	000	46	603	- 800	000
47	023	014	050	48	008	601	012
5.0	004	001	025	51	011	003	027
52	-012	0.03	066	- 53	010	0.05	858
54	.014	005	035	55	0.02	0.01	050
56	017	010	053	57	017	003.	V17-
58	012	000	000	59	012	000	
60	002	001	050	61	601	000	000
62	002	001	050	.63	041	- 002	200
67	000	003	999	68	800	003	637
78	014	015	107	71	\$30	014	046
72	065	068	103	74	002	003	-150
76	009	004	044	73	610	005	050
80	001	000	000	83	606	000	(1)
84	601	000	000	85	002	001	Fedgel: 050 - edi
87	000	004	• 999	39	0.03	000	000
94	022	811	050	. 93	006	600	100
96	029	622	075	96	003	000	000
97	0.03	000	0.0.0	98	803	001	023
	606	644	016		1.1		
Tot	01 654	759	652		· · · · · · · · · · · · · · · · · · ·	1999 - 1999 -	

MANAGEMENT QUALITY CONTROL

The final step in centralized quality control of meteorological data destined for operational analysis deals with the follow-up action which takes place to minimize future discrepancies and losses. This phase of quality control closes the loop which began with on-station monitoring and control. It has many applications. To name a few:

a. It provides information on comparative performance both spatially and time-wise.

b. It aids in isolating system and procedural deficiencies.

c. It enables centralized analysis and forecasting centers to gear objective analysis schedules to optimum times of data availability.

d. It allows major centers to do statistical comparisons of data quality and availability and to isolate internal or exchange problems.

A number of programs have been developed by both AFGWC and NMC to produce detailed viable summary statistics on data receipt and quality. Most of these can easily be modified to provide specific information depending on the application. In this manner special problem areas can be isolated for more detailed evaluation. Often the statistical summaries provide the necessary clues concerning the general nature of a problem. This can be instrumental in significantly minimizing investigative and corrective processes.

Figure 8 shows a typical NMC summary of receipt of radiosonde data for a select group of National Weather Service stations. This type of summary can be produced for any combination of stations in the world and any TEMP or PILOT data parameter and standard level. It can identify data by PART and time of availability. Total daily receipts, by level

and parameter for the stations chosen, are produced along with period averages. Station groupings can be made on the basis of region, nation, communications circuit or many other predetermined choices. The high degree of flexibility in this program readily lends it to versatile applications of monitoring and quality control.

Figure 8

				DEFINITION O	FP	PARAMETER	S
	÷		IF LOUT				
	. 2		TE LGH I	I - TEMPERATURE	P	- PRE55	UKE N WIND
	00)' -	1000MB			AAR -	PART A BASED ON TEST THAT 1000 MB HEIGHT PRESENT
	99);	SURFACE			BBH -	BASED ON TEST THAT THERE ARE ONE OR MORE WINDS
	88		TROPOPAUSE				PRESENT WITH HEIGHTS LESS THAN OR EQUAL TO
	85	; ;-:	850MB				60,000 FEET (EXCLUDING SURFACE WIND)
	77		MAXIMUM WIND			BBR -	PART & BASED ON TEST THAT ONE OR MORE SIGNIFICANT
	70) -	700MB	and the second			LEVELS PRESENT WITH PRESSURE GREATER THAN CR
	50)	500MB				EQUAL TO 100 MB (EXCLUDING SURFACE LEVEL)
	44	÷	VERTICAL WIND	SHEAP		DDH -	BASED ON TEST THAT THERE ARE ONE OR MORE WINDS
	40) -	400MB			21	PRESENT WITH HEIGHTS GREATER THAN 60,000 FEET
	30) -	300MB			CCR -	PART C BASED ON TEST THAT 70 MB HEIGHT PRESENT
	25	- :	250MB			BBP -	BASED ON TEST THERE ARE ONE OR MORE WINDS
	20		200MB		• .		AT PRESSURE LEVELS GREATER THAN OR EQUAL TO
	15	-	150MB				100 MB (EXCLUDING SURFACE WIND)
	10	- 1	100MB			DDR -	PART D BASED ON TEST THAT ONE OR MORE SIGNIFICANT
	07	-	70MB				LEVELS PRESENT WITH PRESSURE LESS THAN 100 MB
	05	-	50MB	•		DDP -	BASED ON TEST THAT THERE ARE ONE UR MORE WINDS
	03	-	30MB			• 1	AT PRESSURE LEVELS LESS THAN 100 MB
	0 <i>2</i>	-	20 MB			MIU -	DATA CORRECTED OR INSERTED BY MANUAL
	01	-	10MB				ASSISTANCE VIA SANDERS CRT TUBE
	+7	-	7MB			SOT -	NUMBER OF TIMES STATION REPORTED SOMETHING
	÷5	-	SMB			PAC -	STATION REPORTED TWO OUT OF THREE HEIGHTS(700.
•	*3		3M8				400, AND 200MB)
	۶	-	2M8				
	¥1	: -	1M8				
• •				STATIONS USED IN TH	ts_f	RUN ARE:	
				72425 72429 72518	7	2402 7 2520 7	2405 2528
				12606 12/12 74486	- 74	4494	



Figure 9 is an example of an AFGWC real-time message intended to notify U.S. and Canadian aviation hourly reporting stations in realtime of unreceived or unrecognized reports. These are also updated and transmitted more than once an hour.

Figure 9

2020	: W8	6785							
AXXX	(30	KANN	2120	30					
STAT	INS	NOT F	2CAD	KANN					
BCE	BLU	CID	ELY	EVV	EVW	EYW	FTY	HHR	MDT
MIV	PUC	RDG	RTN	SZN	WAE	Мах	MBD	WBJ	WBS
MC.A	MDH	MDS	WEO	MEU	WEY	WGN	WGR	WHJ	WHN
WHT	WJC	WKE	MLG	WNN	WRP	NSQ	WTC	WTW	MUR
MUU	WUW	WUW	WUZ	WVK	WVM	WVT	WWV	WZC	WZW
YAJ	YCB	934	YFR	YKF	YLT	YMD	YOC	Ϋ́ΡΥ	YSR
YÜA	YUJ	YUK	YUQ	YUR	YUS	YUX	YVN	420	ZUE
STAT	INS	SURVY	10 68	15 RE	POR	I MIS	16 7(1	
INT	ZDK								

As with the statistical summary for TEMP and PILOT data, the equivalent program for global synoptic data has corresponding good versatility. See Figure 10. At the moment individual parameters cannot be surveyed but development to do this at NMC is underway. Individual stations and groupings of stations can be selected. Availability of data as a function of elapsed time after the observation time is included. Subtotals of the amount of data processed according to observation time is included. An indicator of the frequency receipt for off-time reports is given. Summary totals for the station grouping and period are appended. A separate list of "zero-receipt stations" is given. An inventory of dates and observation times surveyed is prefixed.

Figure 10 LAND STATION RECEIPT TIME SUMMARY

THE FOUR CHARACTERS TO THE LEFT OF EACH BLOCK AND STATION NUMBER ARE INDICATORS OF THE NUMBER OF RECEIPTS WITH OBJERVATION TIMES OTHER THAN ONE OF THE NOMINAL SYNOPTIC TIMES. THE ONLY OFF-TIME RECEIPTS SUMMARIZED ARE THOSE WITH DES TIMES ONE OR TWO HOURS EARLIER THAN A NOMINAL SYNOPTIC TIME. RECEIPTS WITH OBS TIMES OF 222 OR 232 ARE GROUPED UNDER 002, FOR EXAMPLE. THE FIRST CHARACTER GIVES THE COUNT FOR 002, THE SECOND FOR 062, THE THIRD FOR 122 AND THE FOURTH FOR 182. A SLASH (/) INDICATES ZERO RECEIPTS. AN ASTERISF INDICATES ONE OR MORE RECEIPTS, NONE OFF-TIME. AN A, B, OR C INDICATES ONE, TWO, OR THREE OR MORE OFF-TIME RECEIPTS, RESPECTIVELY.

THE SUMMARY IS CATEGORIZED ACCORDING TO THE TIME INTERVAL BETWEEN THE RECEIPT TIME AND THE NOMINAL SYNOPTIC TIME. A REPORT WITH AN OFF-TIME OBS TIME IS PLACED IN THE 0-1 HOUR CATEGORY IF ITS RECEIPT TIME IS EARLIER THAN THE NOMINAL SYNOPTIC TIME.

	STATION		0 1	HOUF	۶ [:]		1-2	HOURS	6		2-3	HOUR	IS .			3-6 F	IOUR	s .	6н	OURS	0R	MORE			TO	TAL			
× + + + + + + + + + + + + + + + + + + +	72405 72406 72407 72408 72412 72414 72425 72503 72504 72507	002 29 28 27 28 29 29 29 29 29 29 29 28 29	2 062 31 31 28 30 30 30 29 29	2 122 30 27 30 31 31 31 31 29 29 31	2 182 26 29 25 28 30 31 31 29 29	00 1 1 0 1 1 0 1 1 0 1	Z 06Z 0 0 0 0 0 0 0 0 0	122 0 0 0 0 0 0 0 0 0 0 0	182 2 0 1 2 3 0 0 0 0 2	00 0 0 0 0 0 0 0 0 0 0	Z 062 0 0 0 0 0 0 0 0 0 0 0 0	2 122 0 0 0 0 0 0 0 0 0 0 0 0	182 0 0 0 0 0 0 0 0 0 0		002 0 0 0 0 0 0 0 0 0 0	062 0 0 0 0 0 0 0 0 0 0 0 0	122 0 0 0 0 0 0 0 0 0 0 0 0	182 0 0 0 0 0 0 0 0 0 0 0		Z 06 0 1 0 0 0 0 0 0	Z 12 0 0 0 0 0 0 0 0 0 0 0 0	Z 18 C C C C C C C C C C C C C C C C C C C	3Z))))))	002 30 297 29 30 29 30 29 30 29 30 28 30 28 30	2 06Z 31 31 30 30 30 30 29 29	122 30 27 30 31 31 31 31 29 29 31	182 28 29 26 30 31 30 31 31 29 31	nation of the term	119 116 112 121 122 120 122 120 115 121
2342 3235 4235 4235 4235 4235 4235 2335 23	72508 72509 72513 72515 72517 72518 72519 72520 72523 72528	27 29 28 29 28 29 28 29 28 29 28 29 24 28	26 31 30 30 31 29 31 0 31	28 31 30 31 30 29 31 30 31	24 30 31 30 27 29 29 31 27 26	1 1 0 1 0 1 1 0	0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 1 0 0 0 0 0	1 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0		0 0 0 0 0 0 0 0 0 0			0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0		0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0			28 30 29 29 29 28 28 30 25 28	26 31 30 30 31 29 31 0 31	28 31 30 31 30 31 29 31 30 31	25 30 31 30 27 29 31 27 27	u u tra a han a na a a s	107 122 120 116 119 115 123 82 117
*A** **** **** **** *** *** *** ***	72529 72605 72606 72608 72612 72613 72613 72617 72618 72712 72712	27 26 28 27 27 29 28 27 29 28 27 27	30 28 29 0 30 29 30 28 0 0	30 30 31 29 30 31 30 30 30 28	289 2231 2251 2251 230 300 0		1 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 1	0 0 0 1 0 1 0 0 0	0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0		0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	0 0 0 1 0 1 0 0 0		000000000000000000000000000000000000000	0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0		28 27 29 28 28 30 29 29 29 28 28 28	31 29 29 0 30 29 30 29 30 28 0 0	30 30 31 29 30 31 30 30 30 29	28 29 31 28 27 31 31 30 30 30	if the restriction of the second s	117 115 120 85 115 121 120 117 88 57
1 4 4 4 1 3 4 4 4 4 4 4 1 4 4 4 1 4 4 4 1 4 4 4 1 4 4 4 1 4 1	74486 74494 78016	28 29 22	31 30 29	28 30 25	30 27 27	2 1 0	0 0 0	0 0 0	010	C 0 1	0 0 0	0 0 0	0 1 0	· · · · · ·	0 0 0	0 0 0	0 1 0	0 0 0	0 0 4	0 0 2	0 0 3	0 0 3		30 30 27	31 30 31	28 31 28	30 29 30	= =	119 120 116

The AFGWC produces a variety of similar shared information pertaining to the availability of data at Carswell AFB. Space permits discussion of only a couple of typical AFGWC monitoring and quality control products. Normally, AFGWC statistics are made available on a restricted one-time basis only to specific requestors in other Federal agencies with reasonable justification. Extensive use of the products is made by the AFGWC for special studies and applications. The net result is smoother uninterrupted flow of data among the centers.

Figure 11 is a copy of one page of an AFGWC reference library. It provides vital information about all known weather stations including their location, observing program, and availability of data. It also shows certain types of equipment used and variations from normal operations. It is a valuable supplement to the NMC production because it provides information on availability of additional types of data like hourlies and AIREPS.

for many second s				and the second
MASTER STATION LINNARY - 1108 FORM	AT - VERSION D	38 JULIAN DATE 017 1960	DATE 020380	PAGE 340
702710 GRN + GULKANA AIRPORT	W L	AK US 6209N 14527W 0481 11 4		N362502 800207WHO 0
102715 STO ++ TYONE LAKE	· ·	AK US 6237N 14643# 0720 10 4		N362627 720331ETAC 0
702715 ANN 5	אא א	K US 6154N 147198 0899 10 4		N362502 780420USVA 0
	PR W	W PADRSNNK		N362531 7103 USWA 0
702718 218 +* LAXE LOUISE	P	-X KS0+2+X	A .	N362505 74061450PP 0
-702718x4WN E-	pR	H UC 6133N 145134 8518 10 4		N362502 700503K NW 0
-702719KAWA 5	PR	W BE CITER TROPP OF LETER	<u> </u>	1 N3(1702 701000 W0 0
	нас РЯ	#/K DS BILSN 14948W 0064 101408		xx2Z L 4888 JAN 1
702723 EDF +* ELMENDORF AF8	P	7K US 6115N 14948W 0059 101468		H361702 /70505#M0 0
TUZTES LHO +* ANCHORAGE/LAKE HOOD	P	AK US 6111N 14957# 0022 10 4		N361702 750418U5WB 0
702730 ANC + ANCHORAGE INTL M	WRPL	7X US 6110N 15001W 0038 11 408 6	110N 15001W 0045 1711	N361710 750307WHO D
702733 PANC +* ANCHORAGE INTL	A 1+2+	AK US 6110N 15001W 0038 11 408		N361702 750214840 0
702735 MRI * ANCHORAGE/MERRILL F	PR N	H AK US 6113N 14950W 0041 10 4	A	N361700 710331USWR 0
	PRR	-W-XZZYYZZZXYXZZYZZXZXYZXXY-D	A.	1 N361702 7711105UPP 0
704730KAWI 5-	PR	-W-TVTVUUWW2	A.1	190 E
-702746KAWN-5	p	-W-UVSFGGGF.9		
-762750 VNS + VALDEZ M.	ρ	-M- AAAASAAAAXXXAXXAASAMASXXXASX-D		N362507 V80622W40 0
702755 4XY + #CCARTHY	W P	4K US 6126N 14255W 0455 10 4 W-P08r+8-T+0xFTTU000PL:0	A	N362518 7412135UPP 0
702755 VD2 + VALDEZ 2	W 9	/K US 6108N 14615W 0034 10 4		N362500 75041150Pp 0
702757 5WT +4 WHITTIER	n	4X US 6046N 14814W 0047 10 4		N362502 790920N044 0
702770 540 + SEWARD		AR US 6007N 14927W 0023 10 4		N361718 740913460 0
702790 5HN * CAPE HINCHINBROOK	W	rk US 6019N 14639W 0056 1024		C NEC2535 790726440 0
	P	1K US 6024N 14605W 0110 10 4		N362511 741213511Pp 0
	PR	LK US 5953N 14745W 0012 10 4	A state of the sta	N362500 741108USWA 0
-702796%AXH 5-	P	H US 6257N 14156W 0523 10 4	A-	N362601 741108ww0 0
- 702910 XARN 5	eR	-W. XXYWXXYXWXXZ22ZZYZYZYYY 0	A.	N362604 74082350PP 0
702915 552 + SLAMAZOOFFTS TAVEN	PR	-W.LLOMMPML	32200 183214 0685 1711	N362602 7910098064 0
702920 TSG -* TANACROSS 	к Р	H	POECA TADETA DADA TIAT	
702924 Z12 +* TOX (USCG)	PR	/K US 6319N 14243W 0509 15 4	A_	NOD UZ (98505KAWN 10
702950 CDV + CORDOVA/MILE 13	₩ L	/K US 6030N 14530W 0012 10 4	A	N362518 740913840 0
TU2965 JOH ++ JOHNSTONE POINT		14 US 6029N 14636W 0010 10 4		N362502 7905035UPP 0
	W N	1 × US 6005N 14230W 0004 10 4	States 1	N362518 740913WMO 0
/U2980XAXN_S	PR	UGUUTXSUR7+8+6+++RUTTTX5-0	R-1	······································

Figure 11



A program similar to the one producing summary statistics for TEMP and PILOT data received is the NMC "Error Check Program." This program provides companion information to the one producing TEMP and PILOT data receipt statistics.

However, the Error Check Program specifies greater detail about the quality of upper air data processed by the NMC. Figure 12 is an example of the information provided by this program. It categorizes information according to:

a. Data source.

b. Heights, temperatures, or winds reported.

c. Number of times the parameter was available during the period for levels from 1000 to 100 mb, or 70 to 10 mb inclusively.

d. Frequency with which the parameter was automatically processed and used in analysis without interference.

e. Frequency with which the parameter failed vertical consistency checks and what the disposition was.

f. How often the data were forcibly held in or discarded from the analysis.

g. How frequently a manual monitor/edit function intervened to manipulate the data for improvement of the analysis.

h. Other discrepancies such as extraneous characters within the raw data.

In addition to specific information about individual cases, the program further summarizes the statistics into totals and percentages by class for the period and station grouping evaluated.

Figure 12 illustrates the performance of a broad grouping of stations in the Western European sector. Standard isobaric height between 1000





-------..... CODE FISURE ******* MEANING PASSED VERTICAL CONSISTENCY CHECK WITH TIGHT LIMITS FAILED VERTICAL CONSISTENCY CHECK AND HAS NOT BEEN RECOMPUTED 8 c FAILED VERIFICAL CONSISTENCY CHECK AND RECOMPUTED F HAS BEEN CHECKED BUT DID NOT PASS VERTICAL CONSISTENCY CHECK WITH LOOSE LIMITS P PURSE UNDER ALL CIRCUMSTANCES н HOLD UNDER ALL CIRCUMSTANCES MAN MANUAL INTERVENTIONS OF ALL TYPES OTHER CODE FIGURES(SUCH AS BLANK-NOT SPECIFIED) OTHER FINA DATA SUMMARY FOR PERIND 1 FEE 1960 TO 29 FEB 1960 1 FEB 1950 TO 29 FEB 1980 ¥ 30 BASED ON RECEIPT BY H+1000-75 DH + 0 KINUTES - BASED ON RECEIPT BY G+1000-74 OR + 0 MINUTES TOTAL 007 HUNS 29 TWPS-TOT WURS ACCESS 1000-100WR HET. VA KS ONLY TOT HETS-TOT TWPS-TOT WUDS ACCESS 1000-100WB HET. TOTAL POSSIBLE 290 TOT HETS-TOT MARKS ONL Y± TRPS . ۵. ST 1705 TMPN 8 с F 2465 ĩş 236 100010000000 000509 000000000 NAUNUNNNNN - NNG KANN NNG AGN ANAANNA KAN A MAGAAAN I KAAANA Disebaadaa - far salaa nna dinarasi fila disebaadaan Naununnnnn - nng kana dinarasi fila disebaadaan i kaanaa a Q NONHOCONT 7 1 10 000010000 រុ 100004 8 00 10000000000000000 20 00000000000 202 21222222222 Ď 0000 10 0 0,0 000 ×1342745 21191705 0000000 00000 000000 00000 100 103

CONCLUSION

In the past decade revolutionary changes in the philosophy and applications of quality control of meteorological data have taken place. This evolution has came about both from necessity and foresight. Quality control has been necessitated by increasingly automated operations at most meteorological centers where ingestion of erroneous data can result in serious degradation of analyses, forecasts and other products because of a lack of human safeguards. Early in the transition from manual to computerized communications, data processing, and objective analysis of data, it became obvious that a high degree of monitoring and control of these data was both necessary and desirable.

Foresight in improving quality control was applied by people with the vision to perceive both the tangible and intangible benefits to be gained by introducing objective and statistical methods to get at the root of data problems and to correct them there.

The World Meteorological Organization developed the World Weather Watch Plan and made data monitoring and quality control an integral part of it. They assigned responsibilities within the Global Observing System, the Global Telecommunication System, and the Global Data Processing System and provided recommended procedures and guidelines for effective programs. Federal meteorological agencies, civil and military alike, adopted and implemented these World Weather Watch precepts and concepts both domestically and internationally.

The Air Force Global Weather Central and the National Meteorological Center are bilaterally developing complementary quality control procedures. These are a delicately balanced blend of automated and manual procedures aimed at controlling the availability and accuracy of present and

future data for both current operations and subsequent use. These Centers' procedures address data difficulties associated with observing practices, measurements, data reduction, coding communications, and processing. They use real-time, non-real-time and management concepts to bring about a full measure of comprehensive inspection and control.

OUTLOOK

The outlook for the 1980's suggest that there will be broader applications of meteorological principles to more fundamental preactical considerations such as weather influences on ecology, world food production, defense, energy use, and basic human comfort and safety.

A greater dependency on technology will be essential to further the capabilities of the discipline. Fiscal restraints are likely to remain high as the climate of government activity becomes increasingly austere. This signals a need for applications of automated systems in playing an ever-increasing role in carrying out work heretofore performed by people. Data monitoring and quality control are prime candidates in such applications.

REFERENCES

AFGWC, 1979: Data Base Quality Control Handbook; AFGWC Pamphlet 105-2. Bureau of the Census, 1977: Weather Data Needs Survey 1977; Conducted for the National Oceanic and Atmospheric Administration, National

Weather Service, 88 pp.

- WMO, 1966: The Global Data-Processing System and Meteorological Service to Aviation, 40 pp.
- WMO, 1968: Quality Control Procedures for Meteorological Data; World Weather Watch Planning Report No. 26, 38 pp.
- WMO, 1974: Manual on the Global Telecommunications System; Publication No. 386, Vol. I, 119 pp.
- WMO, 1977: Guide on the Global Observing System; Publication No. 488, 100 pp.
- WMO, 1979: Manual on the Global Data-Processing System, Vol. I, Supplement No. 2; Publication No. 485, 9 pp.