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

OFFICE NOTE 358

QUALITY CONTROL OF THE AIRCRAFT FILE AT THE NMC
Part I

Paul R. Julian

June 1989

Office Note 358

QUALITY CONTROL OF THE AIRCRAFT FILE AT THE NMC
Part 1

Paul R. Julian

June 1989

Introduction

Since the GATE in 1975, observations taken aboard commercial aircraft, supplemented by observations from military and private aircraft, have become a very valuable source of upper-air data [e.g.Julian & Steinberg, 1975]. The transition from the old 'course-made-good' type of navigation to inertial and omega-type navigation instrumentation resulted in much improved wind vector data. Not only are INS-NAVAID winds much more accurate than the older computed winds, they also are essentially instantaneous, i.e.they are not averaged over legs of an aircraft's course. The development of avionics for jet aircraft also made possible the automatic recording and transmission of the meteorological information from the in-flight aircraft to a collecting site. The first such system, used extensively in GATE, recorded raw instrument output data on-board the aircraft: those data were processed on the ground after the aircraft had landed to produce meteorological reports. The quality and usefulness of data collected in this fashion was forcefully demonstrated in GATE, and led to the design and deployment of the ASDAR system during the Global Experiment. The ASDAR system, which used the data collection system on the synchronous satellites to relay the meteorological data to the collecting site, was even more successful in terms of the quantity and quality of the reports. The current follow-on to the ASDAR system and the even more complex and comprehensive ACARS represent the culmination of a highly reliable, automatic, and high quality data system.

For purposes of the Note, it is necessary to point out that evaluation of meteorological data from aircraft, when all reports are taken together, makes clear that the main sources of error in an operational center's file of aircraft reports are not instrumental but originate from the cumbersome, manually-based system of gathering and transmitting the reports. The conventional AIREP and the PIREP reporting systems depend upon a series of manual actions, data hand-offs, and clerical procedures that do not utilize the automated systems referred to above and result in a major source of error in any present file of aircraft reports.

The analysis that is included in this Note will attempt to document this statement by examining the types and frequency of errors that are termed 'gross' or 'rough' errors- those that cannot be ascribed to any instrument failure of the equipment involved in making the observation. A report is defined as being in error if there is any discrepancy between the quantities being reported (position, time,

T5	Apparent correction to bad report
T6	Inconsistent track (not any of T1-T5)

Sorting by collocation. The second restructuring of the production file is to produce a set of collocated data. All reports within a one degree latitude diameter are placed together and sorted, then, by time of report. Single, isolated reports are not included in this file. Such a sort allows rapid comparison of reports at a single navigation check point, and the detection of outliers is expedited. Of course, some errors/problems that are evident in this type of sort are also easily detectable in the sort by flight identification. However, some are not.

The following Table gives the definition of error that was used in surveying the collocation-sort files. Note is made of the error type(s) that are equivalent to those defined for the flight-id sort.

Table Two
Definition of Errors detectable from Collocated Files

C1	Speed Units (*2 or *1/2)
C2	Initial Direction digit
C3a	Missing Direction digit
C3b	Missing speed digit
C4a	Extra Direction digit
C4b	Extra speed digit
C5	Incorrect date inferred
C6	Unclassified/unknown
C7	Calm wind for missing wind
C8	Same report but for lat/lon (<0.5 deg) [=T4]
C9	Same report but for lat/lon (>0.5 deg) [=T1]
C10	Same report but for time (<0.10 hrs) [=T4]
C11	Same report but for time (>0.10 hrs) [=T2]
C12	Same report but a/c id one char different[=T4]
C13	Apparent correction to bad report [=T5]

Examples of all the error types defined above are given in Figures 1 to 15. All the examples were taken from the restructured production files- only the date is not given.

Frequency of Error

From the sample of seven sort-by-aircraft-id files, a count was made of the Tx occurrences of error. In relatively few cases was it necessary to exercise some subjective judgment as to the proper error class. Nearly all of these involved a decision as to whether an inconsistent report had been corrected (T5,C13): the example shown (Figure 14) is one is which the decision was that the error in longitude was detected and a correction sent. An alternative explanation is that the two reports came from different sources. Since

Table 4

Frequency of error in Collocation sort (18 files in sample)

	Mode(#with)	High(#with)	Low(#with)
C1	3(5)	5(3)	0(2)
C2	4(3)	9(1)	1(2)
C3a	1(3)	1(3)	0(15)
C3b	1(7)	3(3)	0(5)
C4a	1(3)	1(3)	0(15)
C4b	1(5)	1(5)	0(13)
C5	2(5)	7(1)	0(6)
C6	12(3)	20(2)	4(1)

	Number
C7	38
C8	11
C9	7
C10	10
C11	5
C12	2
C13	7

Interpretation of the statistics

From Tables 1 & 3, it is certain that the overwhelming source of error is due to the same report time being appended to reports that are in reality not for the same time and are otherwise consistent with an expected aircraft track. This type of error has been recognized for some years. (I believe the first time I was aware of it was in some WMO/WWW publication: but I cannot locate it or recall exactly the reference.) This error must have its source in either the aircrew or the ground receiving unit attaching a time to the report which is not necessarily the time of the observation, but the time the report is either sent or recorded.

The second most frequent error results, apparently, from the fact that there is more than one source for the observations. There are actually two sub-types here, which I did not bother to distinguish: the first, most common one, is when the entire report is identical save the position, which differs by a fraction of one degree. My interpretation of this error is that the various sources of reports are using slightly different dictionaries for converting named navigation check-points to latitude/longitude coordinates. The second type has identical reports except for times which differ by a few minutes. This problem could arise from different sources using different algorithms and/or units in handling times in hours/minutes/seconds or in decimal form.

Error type T3a has a very probable explanation stemming from the way the AIREP reports are collected. On overseas flights an aircraft may send a sequence of reports to a ground collection site when coming within radio range. Frequently, this sequence will contain reports for

Inferences for NMC Quality Control

There are a number of steps that can be taken to alleviate the error rates that are being observed. We may classify these as internal and external to the NMC. As already mentioned, a new quality control and super-observation program has been written, debugged, and calibrated so as to detect most of the errors that are of concern here. This code uses only the collocation sort, and does not make use of the flight id sort. This decision was a conscious one based upon the experience that most of the gross errors can be detected without the extra computation time (and program development time) to go through a complete track consistency check. The quantity most often in error, the time of report, is used in the new program, but is not a critical quantity: some calibration of adjustable constants can be used to compensate for an uncertain time of observation. (This and all matters pertaining to the new program are to be found in Part 2 of this Office Note)

To protect against the 'calm' for missing wind error, the current (June 1989) data preprocessor deletes any aircraft wind report of 360 degrees/000 knots(in ON29 notation). While this almost certainly deletes some actual calms, the rationale for these deletions was that there are more incorrect than correct calms in the file, plus the fact that with a good assimilating forecast model the report of an actual calm is not contributing much information. This ad hoc procedure will be abandoned when the new quality control program becomes operational- that program is quite adept at detecting false calm wind reports. External to the NMC, certain of the procedures used in reporting aircraft observations can be checked. For example, we have instituted a cross-check of navigation check-point dictionaries at Carswell AFB and within the NWS to insure that the converted positions agree to a hundredth of a degree. (The accuracy of those positions is not nearly as important as their precision!). The problem of what to do about the attaching of identical times to different observations is not easy to solve. We would need to be able to determine exactly where in the chain of reporting procedure these errors are taking place.
As the new ASDAR systems take over a greater portion of the aircraft report volumes, the errors attributable to clerical and multiple sources will decline. The analysis given in this Part 1, however, is a good lead-in to the quality control and super-observation algorithm described in Part 2.

Figure 3
Example of Error Type C3a
Missing Direction Digit
All reports at 48.83N, 180.00W

Aircraft Id	Time	Altitude	Wind Direction	Wind Speed
KAL012	1103	10668	255	85
DAL58	1223	10058	263	91
CDN8	1227	10668	265	95
NWA24	1237	8839	070	100**
DAL52	1245	9449	264	89
JAL2	1292	10668	270	95
THA740	1432	10668	270	90
NWA8	1455	11278	270	95

** Almost certainly 270 degrees

FDX716	1500	11278	270	88

Figure 4
Example of Error Type C3b
Missing Speed Digit
All reports at 24.50N, 153.88W

Aircraft Id	Time	Altitude	Wind Direction	Wind Speed
JAL811	1055	11887	250	140
SIA1	1158	11887	250	14**
PAL107	1255	11887	255	135

** Almost certainly >140 knots

Figure 5
Example of Error Type C4a
Extra Direction Digit
All reports at 24.00N, 180.00W

Aircraft Id	Time	Altitude	Wind Direction	Wind Speed
NWA9	2172	10668	040	30
NWA51	2363	10668	025	30
NWA15	2403	10668	030	30
NWA21	2645	10058	160	30**

** Almost certainly 060 degrees

Figure 8
Example of Error Type C6
Unknown or Unclassified Error
All reports at 53.17N, 145.00W

Aircraft Id	Time	Altitude	Wind Direction	Wind Speed
KAL001	2188	9449	110	25
UAL819	2208	10668	115	10
UAL53	2255	10058	090	10
UAL809	2258	11278	130	10
KAL011	2288	9449	100	20
NWA23	2362	10058	090	25
JAL61	2375	9449	090	57**
UAL97	2457	9449	100	15
CAL003	2585	10668	115	20
CAL005	2650	10668	020	100***

*** This report was also rejected - apparently direction and speed were transposed.

Figure 9
Example of Error Type C7
Calm for Missing Wind
All reports at 36.17N, 86.67W

Aircraft Id	Time	Altitude	Wind Direction	Wind Speed
XX999	2327	10058	310	90
AAL5	2560	10668	287	74
XX999	2710	11278	360	00**

** XX999 is designator given by NMC to all converted PIREPs

Figure 10
Example of Error Type T1
All reports are from flight TWA880

Aircraft Pos		Time	Altitude	Wind Dir.	Wind Speed
46.00	50.00	0123	11278	335	100
46.00	30.00	0205	11278	325	65**
46.00	40.00	0205	11278	325	65**
46.00	30.00	0283	11278	240	50***
46.00	35.00	0283	11278	260	45***

** Two reports differing only in position greater then 0.5 deg (T1)
*** Two different reports with same report time (T3a)

Figure 14
Example of Error Type T5
All reports are from flight QFA11

Aircraft Pos	Time	Altitude	Wind Dir.	Wind Speed
-17.67 175.00	0980	11278	250	45
-13.17 169.50	1060	11278	230	35
- 3.50 200.00	1215	11278	225	25**
- 3.50 160.00	1215	11278	225	25
7.17 150.00	1382	11278	260	50
17.00 140.00	1538	11278	280	80

** Apparent bad report, subsequently corrected.(?)

Figure 15
Example of Error Type T6
All reports are from flight PAA46

Aircraft Pos	Time	Altitude	Wind Dir.	Wind Speed
7.00 20.00	0628	11278	220	20
65.50 20.00	0628	11278	220	20

Both reports cannot be correct- unclassified error

Figure 16
Duplication of Reports, Different Locations
Separation greater than one degree

Aircraft	Position		Time	Altitude	Wind Dir/Spd
UAL819	53.50	169.17	2442	10363	200/85
NWA1	53.50	169.17	2553	10363	200/90
NWA27	53.50	169.17	2678	10973	200/35
UAL819	55.50	169.17	2442	10363	200/85
NWA1	55.50	169.17	2553	10363	200/90
NWA27	55.50	169.17	2678	10973	200/35