

DESCRIPTION OF THE QUALITY CONTROL PROCEDURES UTILIZED BY
THE SPECIAL AIRCRAFT DATA CENTRE

H.M. de Jong
Koninklijk Nederlands Meteorologisch Instituut,
De Bilt, The Netherlands

CONTENTS

I. INTRODUCTION

- 1.1 Data to be checked
- 1.2 General approach to quality control
- 1.3 Data consistency checks
- 1.4 Error characteristics
- 1.5 Quality control programme

II. QUALITY CONTROL PROCEDURES

- 2.1 Principles
- 2.2 Error criteria and error bounds
- 2.3 Moving flagging method
- 2.4 Circuit check
- 2.5 Circularity
- 2.6 Back-tracking
- 2.7 Gross error check
- 2.8 Spatial-temporal data consistency checks
- 2.9 Duplicate and near-identical reports
- 2.10 Validity check of the date
- 2.11 Formal and code checks

III. STRATEGIES

- 3.1 AIDS data (ASDAR ad hoc)
- 3.2 Post flight AIREPs
- 3.3 Concorde SST/AIDS data
- 3.4 Examples

IV. CONCLUSION

REFERENCES

APPENDIX I

APPENDIX II

I. INTRODUCTION

In the FGGE basic data set the subset of aircraft meteorological observations consists of the following groups:

<u>Level II-a</u>	<u>Level II-b</u>
conventional AIREPs	AIDS
ASDAR	Post Flight AIREPs

Note: AIREP is an abbreviation and contraction of Air Report.
AIDS stands for Aircraft Integrated Data System,
ASDAR for Aircraft to Satellite Data Relay.

During the FGGE operational year Level II-a reports were collected by the four FGGE Level II-b Area Sub Centres.

The collection and processing of the Level II-b reports was the responsibility of the Special Aircraft Data Centre (SADC).

The subsets of Level II-a reports constitute basic data from the World Weather Watch Global Observing System (GOS) Surface Based and Satellite system. These reports are distributed via GTS and other channels.

It is noticed that in the WWV/GOS there is also a provision for the non-real time distribution of Post Flight AIREPs. These reports are delivered by pilots when debriefing in compliance with an arrangement between WMO and ICAO [4]. During the global experiment an extra collection of Post Flight AIREPs in a delayed mode was one of the tasks of the SADC.

1.1 Data to be checked

This report describes the procedures and methods, some of them innovatory, which were used by the SADC to perform suitable validity and quality checks prior to the transcription of the reports in the FGGE international exchange format and subsequent delivery to the Level II-b Space Based and Special Observing System Data Centre, Norrköping, Sweden.

Aircraft meteorological reports comprise the following elements:

day	}	date]	identification elements
month				
year				
hours	}	observation time		
minutes				
airline identification				
flight number				
latitude (degr., min.)				
longitude (degr., min.)				
pressure altitude/flight level (ft)*]	meteorological data content		
ambient air temperature (°C)				
wind direction (degr.)				
wind speed (knots)				

* Atmospheric pressure is derivable from this element using the ICAO standard atmosphere specifications.

For formats and codes used, reference is made to the Implementation/Operations Plan [1], [2] and [3]. On page 3 a section of a print-out is reproduced for series of AIDS data records.

The quality control methods to be described apply not only to suites of AIDS data and Post Flight AIREPs but also to their Level II-a counterparts: conventional (in-flight) AIREPs and ASDAR data.

1.2 General approach to quality control

The objective of quality control is the detection and possible correction of the information content in data records.

The quality control, to be efficient, should also comprise the tracing of errors which are attributable to what may be called "defects in data manipulation". Such defects manifest themselves when

- duplicate or near-identical reports are received;
- an invalid date (day, month, year) is entered;
- departures from agreed formats and codes are inevitable.

The data records are divided into two sections, one containing identification elements and one the physical elements (see Section 1.1).

The checking of the physical elements is considered to be the main task, but it should be realized that validity checks of the identification elements are equally or even more important. When for example in a record a faulty position coordinate is found, not being recoverable or restorable, then the report has no intrinsic value any more for research and should be eliminated accordingly from the data set.

In the Implementation/Operations Plan [3] it is stated that incorrect physical parameter values should not be changed but flagged instead. However, when defective identification elements are traced, these are allowed to be restored in case of high data redundancy.

Owing to a shortcoming in the Plan no provision has been made to include error marks in the formats and codes for flagging uncoverable and unrestorable identification elements. To warrant the integrity of the data base for research it has been decided in the SADC that such reports be removed from the files.

The quality marks used for flagging the meteorological data (air pressure, temperature and wind) are given in Table 1.

Table 1. Quality control marks used

<u>code figure</u>	<u>description</u>
0	no definite control
1	value correct
2	value suspect
3	value erroneous
9	value missing

In recent years numerical quality control has made good progress, but few methods have become firmly established.

This forced the SADC to try out special service programs of its own design for quality control.

As the information is mostly available in arrays of data points along one-dimensional flight trajectories embedded in

3-space, some powerful new checking algorithms could be developed.

Ultimately, the Centre has introduced the following methods:

- (a) tests against absolute, empirical and climatological limits applied to single observations (gross error check);
- (b) tests for spatial and/or temporal consistency applied to arrays of data points.

Note: As wind, pressure and temperature are very weakly correlated in a single data point, no checks were made of the internal data consistency.

Gross error checks are used for the detection of erroneous data (mark = 3); the more powerful data consistency checks serve to identify and flag suspect values (mark = 2).

1.3 Data consistency checks

Consistency checks require the availability of error criteria formulated for data in two or more proximate data points, proximate in the sense of a fairly high spatial or temporal data resolution. In series of air reports there exists a coupling between spatial and temporal resolution. This is typical of all observational series made on board moving platform stations (balloons, satellites, aircraft). This coupling is stronger with greater speed of the platform and smaller period of sampling. In Table 2 estimates are given of the spacing of data points for various groups of Air reports.

Along-track consistency checks are the more versatile when the data resolution increases and the tolerance limits in the error criteria can be reduced accordingly.

This implies, as Table 2 shows, that such consistency checks are most profitable when utilized in ASDAR data series and the least in series of in-flight and Post Flight AIREPs.

Table 2. Space-time resolution in data arrays of aircraft meteorological observations

data source	sampling period	corresponding horizontal spacing (km)	remarks
conventional AIREPs	~ 1 hour	~ 1000	reporting when crossing 10 degr. meridians over oceans.
Post Flight AIREPs	variable (< 1 hour)	less than 1000	in general more frequent reporting than with AIREPs.
AIDS			
jet aircraft	800 s	200-250	
Concorde SST	800 s subsonic 400 s supersonic	200-250 200-250	change-over near Mach = 1.5.
ASDAR	400 s (200 s)	100-125 50	commercial airlines military.

The entries in the table suggest that the resolution is even marginal when the consistency check is used for the groups of conventional (in-flight) and Post Flight AIREPs.

As to ASDAR, when the prototype units were tried out in research flights, the resolution was taken even higher than indicated in the table: 16 observations per hour, see also Section 3.4.

Methods of along-track data consistency checks are described in Section 2.8.

It is noticed that the error criteria used for checking the physical data content do not explicitly rest on both the spatial and temporal resolution. Most of these exclude the time variability of the quantities to be checked.

Fig. 1 depicts the data coverage of operational sets of conventional AIREPs and ASDAR data series in a 24-hour period over the northern hemisphere on one day in February 1980 as received at the European Centre for Medium Range Weather Forecasts (Reading, U.K.).

ASDAR data series are present in a fair amount both over oceans and over continents. The data points of conventional AIREPs clearly show a clustering along 10-degree "reporting meridians". The spacing of the data points is 1 degree or more, in compliance with ICAO/WMO reporting procedures. Noticeable is that this spacing is of the same order as for the ASDAR data series.

Apparently, the reporting meridians are also candidates for the application of consistency checks in data arrays, albeit that the error characteristics are somewhat different from those encountered along the flight track of an individual flight.

1.4 Error characteristics

When discussing which methods should be taken for checking purposes, the choice will depend on the special observing features and known error characteristics typical of each group of air reports.

AIDS and ASDAR systems are almost fully automatic data acquisition systems, but in the ICAO/WMO air reporting system numerous operations are still performed manually.

In a fully automatic mode, reading-, administrative- and coding errors are absent; however, these are dominant in the error structure of conventional AIREPs.

Communication errors are absent in AIDS data and Post Flight AIREPs during FGGE, because no radio link was involved in the collection of these reports.

In fully automatic systems, however, mutilations in sensor values may arise owing to malfunctioning of the equipment.



FIG.1

AIREP/ASDAR DATA COVERAGE
00-24Z 80/02/26

Experience gained when checking AIDS data revealed that the following recurrent error sources may be found:

- substitution of faulty dates;
- wrong time setting; instabilities in time recording;
- faulty recording of latitude and/or longitude (alpha-numerically);
- breaks in operational series;
- duplicate reports and near-identical reports, and a few others due to technical causes.

A priori knowledge on certain flight and performance characteristics such as altitude profiles, nominal cruise speeds, course changes en route etc. may also be indispensable. The characteristics of the Concorde SST are so different from those of normal carriers that the SADC had to prepare a special quality control mechanism for the Concorde SST (see Section 3.3).

1.5 Quality control programme

The numerical error checking methods as proposed in Section 1.2 and to be described in later sections have been elaborated in the SADC for use in three data files:

- AIDS data file provided by Douglas DC-10, Boeing 747 and 747 SP;
- AIDS data file provided by Concorde SST, and
- file for Post Flight AIREP's during the Special Observing Periods.

The methods have been utilized also on an ad-hoc basis in some ASDAR data series.

The files are subjected not to one single checking operation only, but to a whole repertoire of different checking algorithms.

This may evoke some interference and priority problems.

When, for example, two checking algorithms are put in action on one and the same given array of data points, this may necessitate the matching of error bounds used in the error criteria.

In the quality control programme gross error checks generally take the lead. When they are followed by data consistency checks, priority rules need to be established, taking the most powerful check last. Absolute priority is reserved, however, to algorithms which trace the defects in data manipulation.

II. QUALITY CONTROL PROCEDURES APPLICABLE TO DATA ARRAYS

Numerical quality control is not intended to search for random instrumental and processing errors of a small magnitude because such errors escape discovery. Checks may be developed for the detection of large errors and outliers which are still random in nature and few in number.

In general, the checking procedures to be described have not any grip on systematic or highly correlated errors.

In an operational environment the best strategy is first to apply validity checks of identification elements and then to continue by checking the physical elements.

When carried out in these two stages, the control of the identification part of the records should include a mechanism for error correction by substituting for estimated values if practicable.

2.1 Principles

When whole data arrays are available, supplied by moving platform stations, it is obvious that for quality control purposes some kind of sequence test should be developed. When activated, such a test runs through all data points of the data array in succession, starting in the entry point and terminating in the exit. Attributes indispensable for a sequence test are: suitable error criteria, error bounds and an effective flagging procedure.

The error criteria need to be evaluated for data in one or more proximate data points ($N = 1, 2, \dots$). For $N = 1$ the check is simply a gross error check. Should it appear that, when computing the expression for error control, the result exceeds a pre-set upper or lower bound, this will be an indication for the presence of erroneous or suspicious parameter values somewhere in the N data points. Without additional clues the identification of the parameter(s), which are really defective, and their precise location, is still not possible. An exemption is a gross error check with only one parameter involved.

An essential tool to single out the real source of error is the flagging of parameters during the sequence test. Here a leading principle is: when an N-point criterion is not violated by the observed values, it is assumed that in the N-points in question all parameter values are correct, because it is very unlikely that large errors, when they are present in more than one parameter (which is already exceptional in itself), will have a compensating effect.

In the present context flagging may pursue two objectives: either to indicate the error status of identification elements or to indicate the error status of physical elements.

In the first case flagging is said to "label" defective identification elements and to indicate how to proceed in subsequent operations, e.g. to ignore the data, to determine substitute values, to eliminate the data, etc.

In the second case flagging is said to "mark" the error status of the physical elements for inclusion in the code formats later.

In an operational environment the complexity of the data quality control increases rapidly with growing number of data points and variables. This suggests that one should search for error criteria with as few data points and as few variables involved as possible.

But it should not be overlooked that when the above proposed basic principle is accepted, more complex criteria, when not violated, may easily discern whole series of correct data !

2.2 Error criteria and error bounds

To develop a sequence test for checking the spatial-temporal consistency within a data array it is required to look for a suitable error criterion in the form

$$\delta_1 \leq F(p_{k;n}, q_{m;n}) \leq \delta_2, \quad (\text{II.1})$$

where F is an analytical expression involving M identification elements q_m ($m = 1 \dots M$) and K physical elements p_k ($k = 1 \dots K$), both evaluated in $n = 1 \dots N$ proximate data points.

δ_1 and δ_2 denote appropriately chosen tolerance limits, such that when the constraint is violated one or more elements must be considered suspect in the sense of their error status.

When F is within the range $[\delta_1, \delta_2]$, then according to the leading principle proposed in 2.1 it is assumed that all parameter values $p_{k;n}$, $q_{k;n}$ are correct, apart from non-detectable measuring- and processing errors (noise). However, when out of range, the conclusion must be that an inconsistency exists in the sampled data.

If more than one parameter is involved in the constraint, the criterion is not selective in the sense that the identity of the parameter(s), which really cause its violation, can be fixed, let alone beyond which error bounds the parameter value(s) are really suspect.

In a series of observations, therefore, the strength of the criterion lies predominantly in the discrimination of error-free parameters.

Crucial in the approach is the choice of suitable threshold or aperture values for δ_1 and δ_2 .

The data consistency criterion should be well posed, so that in general it will not suffice to attach fixed constant values to the tolerance limits.

δ_1 and δ_2 should at least depend on nominal values of $p_{k;n}$ and $q_{m;n}$. But this would be against a basic principle of quality control when both the $p_{k;n}$ and $q_{m;n}$ were not subjected in advance already to a separate process of error checking.

Then obviously there is no other choice than to assign fixed constant values to δ_1 and δ_2 at the risk of meeting with a more weakly defined data consistency criterion.

Things get better when e.g. the independent variables q_m were checked beforehand and corrected independently in a previous stage of the quality control processing using an N -point criterion of the form:

$$\delta_1 \leq F(q_{m;n}) \leq \delta_2 \quad (m = 1 \dots M, n = 1 \dots N). \quad (\text{II.2})$$

Then the threshold values are again allowed to become a function of the q_m improving the effectiveness of control considerably.

Examples are presented in Section 2.7, Tables 4 and 5.

A logical extension is the use of more than one error criterion in a sequence test e.g. when checking the error status of vector quantities (see Section 2.8, B.3).

2.3 Moving flagging method

When checking a data consistency in a data array, the algorithm of a sequence test proceeds stepwise. This actually means that each data point will take part in N consecutive steps when an N -points error criterion is used, except in $N-1$ points near the entry and exit points. In order to detect the error(s), each step should involve a flagging of all parameters in the data points used in that step to indicate whether the parameter values are (possibly) correct or not.

This marking or labelling can be done as follows: when after substitution of the observed values into the checking formula the resulting discrepancy does not surpass a pre-set tolerance limit, the parameter value(s) are flagged as being correct. However, when exceeded, the parameter values are flagged as being suspect, except when they had already been flagged as being correct in a previous step.

The effect of this moving flagging is that the data points are discovered when one (or more) parameters are corrupt.

There are two possibilities now: first, when the checking criterion contains only one variable, the algorithm enables precisely to identify both the data point and corrupt parameter value; second, when more parameters are involved, the moving flagging algorithm only identifies the data point(s) where an inconsistency exists but still fails to single out the parameter which actually is the cause of the data inconsistency. In this situation additional control mechanisms need to be used to discover the wrongdoer. Such additional checks may anew incorporate a moving flagging algorithm.

To illustrate how the method operates, assume that in a data array of N points (Fig. 2, $N = 7$) one data point is defective (noise spike, outlier), then a 2-point flagging operation

tests in each data point, beginning in point 1 and ending in point N-1, whether or not the chosen error detection criterion is violated. Then the data points are marked (labelled) at each step, as indicated. The marking which results after termination of this run indeed registers point 4 (mark = 2) as the suspect data point (step 6).

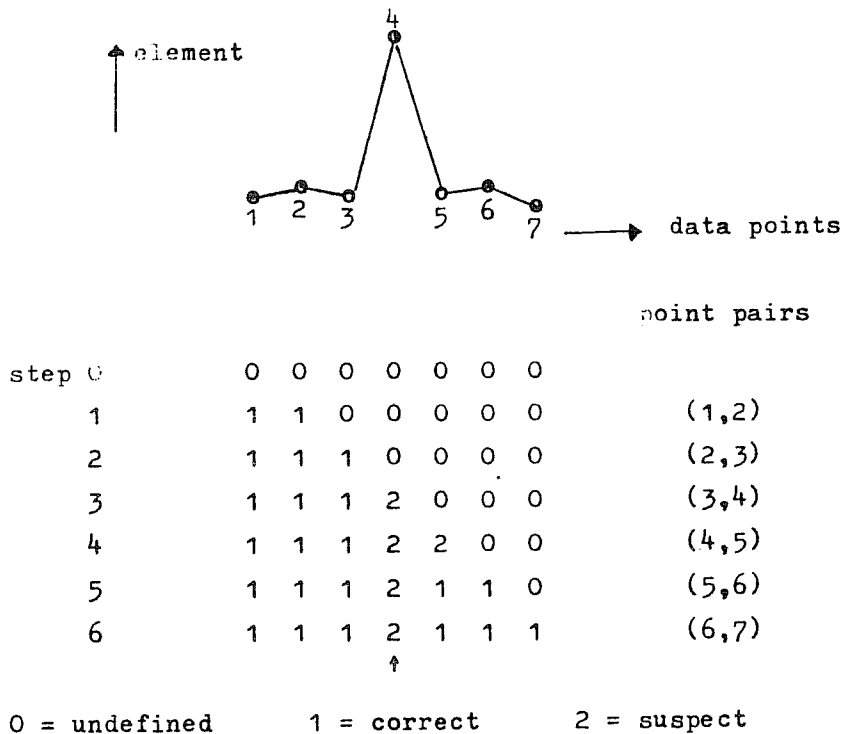


Fig. 2 2-point flagging method identifying a corrupt data point.

A striking feature of the moving flagging algorithm is that when two regimes of a specific quantity meet, the algorithm rightly does not flag the corresponding parameter along the interface as being suspect. Examples are a material course change en route, strong wind shears when crossing a trough line, a sharp temperature rise when passing a frontal surface or tropopause level, etc.

2.4 Circuit check

When examining a sequence test the discrimination of faulty data by using an N-point flagging algorithm is bound to fail in the N-1 data points near the ends of a data array.

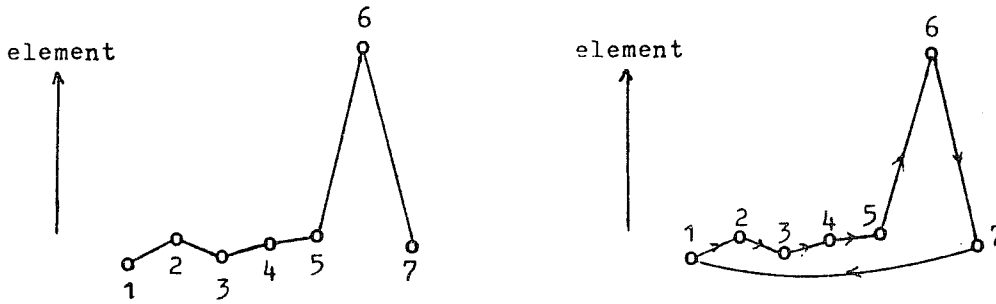
In order to express the fact that the error checking in those points cannot be completed and therefore the error status is still not known there, the error marks (> 1), see Table 1, are changed into mark = 0.

In certain situations, however, the sequence test may be carried on beyond the end points, as if the data array were a "closed" array. The feasibility of such a circuit check depends on the behaviour of the error criterion used.

The circuit check offers the opportunity to assess the error status unambiguously in each data point, entry point and exit point included. An example is a 2-point flagging operation based on the speed control of a moving platform station. Taking as a criterion that the distance flown between two adjacent data points should correspond to the product of the time difference and a nominal or average speed, then the criterion is also applicable between entry- and exit point, with a proviso that the departure of the flown track from the greatcircle route between the end points is not too great (for more details cf. Section 2.8, A1). The circuit check which results is one of the most useful tools for position/time control.

A checking algorithm, which is designed to test the physical data content of a data array, cannot be extended to change over into a circuit check, simply for the reason that in the atmosphere the meteorological data are too weakly correlated over large distances of the order of 1000 km and more. That is why in the FGGE subsets of air reports 2 to 3 per cent. of the meteorological data bears mark = 0. to indicate the undecidability on their error status.

Fig. 3 shows the effect of a 2-point flagging operation, applied, on the left, under normal conditions and, on the right, when generated in a circuit. Whereas on the left the error status in the exit point is undefined (mark = 0), it becomes fully specified on the right.



step	sequence check	point pairs	circuit check
0	0 0 0 0 0 0 0	-	0 0 0 0 0 0 0 -
1	1 1 0 0 0 0 0	(1,2)	1 1 0 0 0 0 0 (1,2)
2	1 1 1 0 0 0 0	(2,3)	1 1 1 0 0 0 0 (2,3)
3	1 1 1 1 0 0 0	(3,4)	1 1 1 1 0 0 0 (3,4)
4	1 1 1 1 1 0 0	(4,5)	1 1 1 1 1 0 0 (4,5)
5	1 1 1 1 1 2 0	(5,6)	1 1 1 1 1 2 0 (5,6)
6	1 1 1 1 1 2 2	(6,7)	1 1 1 1 1 2 2 (6,7)
7	1 1 1 1 1 2 0	(*)	1 1 1 1 1 2 1 (7,1)

(*) change mark = 2 into
mark = 0 in view of
undecidibility.

0 = undefined
1 = correct
2 = suspect

Fig. 3. Moving flagging operation in a circuit check.

2.5 Circularity

It was noticed earlier, Section 2.3, that when more than one, previously unchecked, elements are involved in a data consistency sequence test, a moving flagging method is not capable to identify precisely which element(s) actually cause the inconsistency.

Notable is the case that the method is applied to elements which, along track, show a mutual coupling in one way or another.

An example, treated already in Section 2.4, is the dynamical coupling existing between positions and observation times which are provided by a moving platform station.

Here the problem of circularity may arise: to find a substitute value for a faulty position coordinate using (correct) observation

times or vice versa to find a substitute value for a faulty observation time using (correct) position coordinates. Only in the very exceptional case that both elements are defective the report ought to be discarded.

In practice the circularity problem can be tackled as follows: let by flagging a data point be discovered where both position and observation time are under suspicion; then first derive from proximate (correct) data points estimates for

- a. position by interpolation using the observation times in the coefficients;
- b. observation time by interpolation using positions in the coefficients.

Next a test is made to see if both these estimates fit in the data array. If they do not fit, the report must be discarded. If one fits, the defective element shows up and can then be replaced by its substitute a or b.

A considerable number of AIDS and Post Flight Air reports could be saved by solving this circularity problem.

2.6 Back-tracking

A sequence test proceeds along the track between entry point and exit point. The test may also be performed in reverse direction starting in the exit point and terminating in the entry point.

This back-tracking should of course lead to the detection of the same faulty parameter values. Under certain circumstances, however, some obstacle may cause the algorithm to finish its job before the exit point has been reached, for example when a series of observations shows gaps or when various consecutive data are garbled ("bursts"). Then back-tracking is used to check the data points which escaped control in forward tracking and vice versa.

2.7 Gross error checks

In a 1-point sequence test all elements, except for airline identification/flight number, were subjected to a gross error

check against absolute, physical and empirical limits.

In view of the diversity in flight characteristics of various types of carriers the limits had to be adjusted to special features of wide bodies, SST carriers and other aircraft.

Fig. 4 depicts typical altitude profiles going with wide bodies and the Concorde SST. The Concorde altitude profile is indicative of the cruise climb mode when during supersonic flight the Mach number is held steady at Mach 2.1 ± 0.1 .

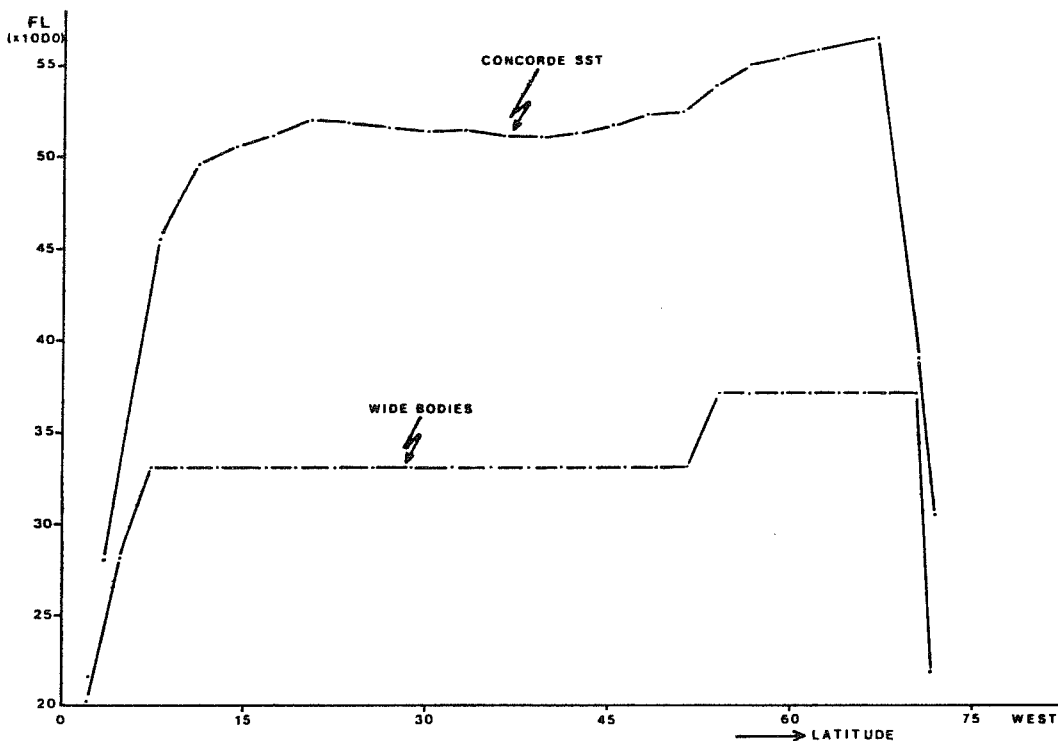


Fig. 4. Altitude profiles typical of wide bodies and Concorde SST.

The profile is strongly temperature-dependent: in the tropics ($< -80^{\circ}\text{C}$) the maximum altitude attained is close to Fl. 61, in the extratropics the altitude reached is sometimes lower than Fl. 55. The altitude profile of wide bodies reflects flights on single levels with a few stepwise changes dictated by Air Traffic Control regulations and flight performance (fuel economy).

Table 3. FIXED GROSS ERROR CHECK LIMITS

used for

a: Post Flight AIREPs

b: AIDS wide bodies

c: Concorde SST AIDS

code parameter	lower limit	upper limit	remarks
A. IDENTIFICATION ELEMENTS:			
day	1	28, 30 or 31	FGGE operational year
<u>date</u> month	1	12	
year	78	79	
<u>obs. time</u> hour	0	23	
minutes	0	59	
<u>Airline identification/ flight number</u>	-	-	no check performed
latitude degrees	90 S	90 N	
minutes	0	59	
longitude degrees	180 W	180 E	
minutes	0	59	
pressure altitude/ flight level	Fl. 1 Fl. 20 Fl. 20	Fl. 45 Fl. 45 Fl. 61	a b c
B. METEOROLOGICAL ELEMENTS:			
<u>air temperature</u>			limits not fixed, see Table 5.
<u>wind direction</u>	0	360	
<u>wind speed</u>	0	225	a and b; for c see Table 4.

The limits used for gross error checks are presented in Table 3.

When gross errors are revealed in the identification part of individual reports, an attempt is made to enter adjusted values into the reports which may be derived from checked information in the data series. However, when gross errors are found in the meteorological data content, the meteorological parameter values are marked as being erroneous (mark = 3) without correction.

A motivation not to correct the physical elements is also the extreme variability of the data in the airspace where AIDS data and Post Flight AIREPs are obtained (tropopause and jet-core crossings, strong wind shears).

Table 3 presents fixed limits only. The use of these limits is straightforward in case of gross error checks of identification elements, but in case of checking wind and temperature the limits should be taken in dependency of the pressure altitude up to Fl. 61. This implies that gross error checks to be performed on these elements need to be preceded by a separate control of the pressure altitude (cf. Section 2.8, B.2).

The variable limits used for wind and temperature gross error checking are reproduced in Tables 4 and 5 together with a presentation of the analytical expressions used.

Table 4. Gross error check limit for wind speed used for the Concorde SST

Flight level x 1000 (Z_p)	Speed (knots)
39	225
40	222
42	215
44	209
46	202
48	195
50	189
52	182
54	175
56	169
58	162
60	155

Analytical expression:

$$Z_p \leq 39 : 225$$

$$39 < Z_p \leq 61 : 225 - 10 (Z_p - 39)/3$$

Table 5. GROSS ERROR CHECK LIMITS FOR UPPER AIR TEMPERATURES AS A FUNCTION OF FLIGHT LEVEL (x 1000 ft).

Flight Level x 1000 ft (Z_p)	Low value °C	High value °C	Remarks
2	-65	54	
4	-64	48	
6	-64	42	
8	-64	36	
10	-64	30	
12	-63	24	
14	-63	18	
16	-63	12	AIDS and
18	-62	6	Post Flight AIREPs
20	-62	0	
22	-64	- 3	
24	-66	- 6	
26	-68	- 9	
28	-70	-13	
30	-72	-16	
32	-74	-19	
34	-76	-22	
36	-78	-26	
38	-80	-29	
40	-82	-30	
42	-84	-30	
44	-86	-30	Boeing 747 - SP
46	-88	-30	
48	-90	-30	
50	-92	-30	
52	-94	-29	
54	-96	-27	Concorde SST AIDS
56	-96	-25	
58	-96	-23	
60	-96	-21	

Z_p denotes pressure altitude/flight level (x 1000 ft):

Analytical expressions:

lower bound

$$Z_p < 20 \quad : \quad -65 + .15 Z_p$$

$$20 \leq Z_p < 54: \quad -42 - Z_p$$

$$54 \leq Z_p \leq 61: \quad -96$$

upper bound

$$Z_p < 20 \quad : \quad 60 - 3 Z_p$$

$$20 \leq Z_p < 40: \quad 32 - 1.6 Z_p$$

$$40 \leq Z_p < 51: \quad -30$$

$$51 \leq Z_p \leq 61: \quad -81 + Z_p$$

2.8 Spatial-temporal data consistency checks

Presently there is no great potential of error detection (and correction) algorithms to use successfully in data arrays supplied by moving platform stations.

For checking purposes the Centre could fall back on the following 2- and 3-point error detection (and correction) methods supported by a moving flagging algorithm:

A. Control of identification elements

- A.1 2-point circuit check based on a nominal cruise speed.
- A.2 predictor-corrector method based on cruise control.
- A.3 3-point sequence test based on cruise control.

B. Control of the meteorological information content

- B.1 3-point data consistency check based on a line integral theorem with specializations.
- B.2 3-point error detection method based on interpolation.
- B.3 2-point data consistency check based on along-track gradients or rate of change of a quantity.

Note: Some of the error criteria do not respond to the asynopticity of the data. Such criteria however remain useful when somewhat coarser error limits are taken in order to incorporate extra time-induced effects.

In the processing scheme all these checking algorithms were activated after gross error checks had been made, with one exception: a gross error check of temperature and wind data is generated not earlier than after finishing a consistency check of en route pressure altitude/flight level data.

ad A. All three algorithms to be described aim at checking the validity of positions and observation time.

The error criteria all are inferred from the most obvious observing features such as sampling interval and the vehicle's speed and maximally possible course changes.

The flight characteristics of the Concorde SST are so special that the error bounds needed a separate treatment. Characteristic features of Concorde SST operations are summarized in Appendix I.

2-point circuit check (A.1)

Let a data array of M data points be specified by:

$$x(i), t(i), \quad i=1\dots M,$$

then a sequence test is applicable using a criterion which is inferred from the requirement that along track the ground speed varies within a given range $[\delta_1, \delta_2]$:

$$\delta_1 \leq \left| \frac{\vec{x}(j) - \vec{x}(i)}{t(j) - t(i)} \right| \leq \delta_2 \quad (\text{II.3})$$

Usually $j=i+1$ in the forward tracking mode, $j=i-1$ when backtracking. δ_1 and δ_2 denote the limits for the ground speed.

As enunciated in Section 2.4, this check may be continued beyond the end points ($i=1, j=M$ or $i=M, j=1$), so that the sequence check changes over into a circuit check. This also means that the supporting 2-point moving flagging method will label all corruptive data points, the end points included.

The check is rather a coarse one for the reason that the ground speed may vary in a broad range under the influence of wind. Therefore the check has been introduced by the Centre primarily to disclose corruptive end points which, when present, should not be used at the start of succeeding control algorithms.

For wide bodies the limits δ_1 and δ_2 are (in knots):

$$\delta_1 = .7 \times 500 = 375$$

$$\delta_2 = 1.35 \times 500 = 675$$

The speed control in SST operations is so variable that the choice of the limits is treated separately in Section 3.3.

Predictor-corrector method (A.2)

Malfunctioning of the equipment may cause interruptions in the recording of data points and the recording of faulty positions or clock times.

At first sight faulty position/time elements may easily be detected by using some kind of "predictor-corrector" method. Close inspection reveals that the problem is complicated in view of insufficient continuity conditions along the flight track. The continuity of the flight track may break down due to considerable course changes which are dictated by the flight plan under control of the navigator (changes in heading of more than 30°). This calls for special measures to be taken in the development of a suitable position/time control mechanism.

The method used by the SADC is as follows:
 Let $\{\vec{x}(i), t(i)\}$, $i=1\dots N$ denote the data points along the flight track. When during the processing the points $\{\vec{x}(i), t(i)\}$ and $\{\vec{x}(j), t(j)\}$ represent the latest points already checked and proven to be correct, then the next point $\{\vec{x}(k), t(k)\}$ should, in view of the continuity of the ground speed, be found within a circle ring bounded by the circles about $\vec{x}(j)$ with radii $r_{1,2}$ (see Fig. 5):

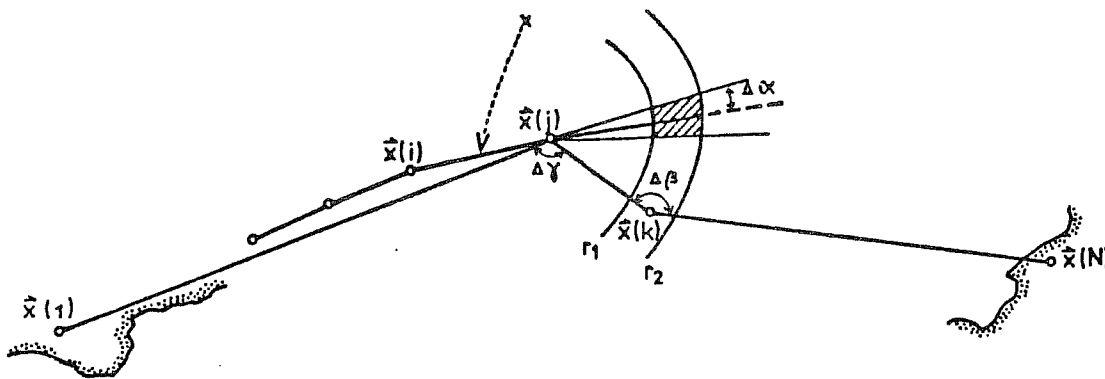


Fig. 5. For explanation see text.

$$r_{1,2} = \frac{t(k)-t(j)}{t(j)-t(i)} (1 \pm \delta) d(i,j) \quad (\text{II.4})$$

where $d(i,j)$ denotes the arc distance between $\vec{x}(i)$ and $\vec{x}(j)$ and δ is some aperture value (e.g. $\delta = 0.2$). When the point $\{\vec{x}(k), t(k)\}$ is not found in the ring, this is an

indication that either $\vec{x}(k)$ is in error or $t(k)$ or both. As the criterion is of the form (II.2) the problem arises of circularity (cf. Section 2.5).

In the case of small course changes, the point $\vec{x}(k)$ should also be within an arc subtended by an angle $\pm \Delta\alpha$ (e.g. $\Delta\alpha < 15^\circ$) as measured with respect to the great circle arc extension through $\vec{x}(i)$ and $\vec{x}(j)$.

However, in the case of an appreciable course change in point $\vec{x}(j)$ the point $\vec{x}(k)$ may be displaced with respect to the great circle arc through $\vec{x}(i)$ and $\vec{x}(j)$ to such an extent that the position $\vec{x}(k)$ may erroneously be considered being in error. Taking a greater limit for the "viewing angle" $\Delta\alpha$ (e.g. $\Delta\alpha = 40^\circ$) is no remedy, for then, as shown experimentally, faulty positions may be misinterpreted as being correct.

This expresses the need to look for another criterion, in order to be sure that the algorithm will not label some faulty element as being correct and/or correct elements as being faulty. The criterion which has ultimately been chosen is based on the observation that along the flight track the "side-viewing angles" $\Delta\phi$ and $\Delta\gamma$ (see Fig. 5) to the end-points $\vec{x}(1)$ and $\vec{x}(N)$ should remain within narrow bounds. This "homing criterion" requires that angle $(\vec{x}(j), \vec{x}(k), \vec{x}(N))$ and angle $(\vec{x}(1), \vec{x}(j), \vec{x}(k))$ remain above a prescribed limit (e.g. $> 155^\circ$).

The algorithm proceeds as follows:

- step 1 Put $i=j, j=k, k=j$
- step 2 Put $k=k+1$. If $k-j > \Delta i$ (e.g. $\Delta i=4$) then label $\vec{x}(n), t(n), n=k \dots N$ as being suspect and jump to step 6. Also if $k > N$, then go to step 6.
- step 3 Compute the radii $r_{1,2}$ with formula (II.4) and compute the arc distance $d(j,k)$.
If $r_1 \leq d(j,k) \leq r_2$, then go to step 4, otherwise label $\vec{x}(k), t(k)$ as being faulty and jump to step 2.

- step 4 If angle $(\vec{x}(i), \vec{x}(j), \vec{x}(k)) > \Delta\alpha$ (e.g. $\Delta\alpha = 170^\circ$), then label $\vec{x}(k)$, $t(k)$ as being correct and go to step 1.
- step 5 If max. (angle $(\vec{x}(1), \vec{x}(j), \vec{x}(k))$, angle $(\vec{x}(j), \vec{x}(k), \vec{x}(N))$) $> \Delta\beta$ (e.g. $\Delta\beta = 160^\circ$), then label $\vec{x}(k)$, $t(k)$ as being correct and go to step 1, otherwise label $\vec{x}(k)$, $t(k)$ as being suspect and go to step 2.
- step 6 If $k < N$, then apply the algorithm in the back-tracking mode.
- step 7 Correct the faulty positions or observation times through interpolation as explained in Section 2.5 for the circularity problem. Exit.

Properties:

- (1) The algorithm has been described as operating in the forward-tracking mode. But as step 2 indicates, it may occur that the process fails to proceed if more than Δi points (e.g. $\Delta i = 4$) in succession are garbled, or when gaps of more than Δi points are present in the observational series. Then a second run is made in the back-tracking mode to check the points which hitherto had escaped control.
- (2) This "search-light" method is also capable to detect systematic departures of positions or observation times arising in portions of the data array.
- (3) Special precautions should be taken in the end-points where the climb phase or descent may be involved (δ_1 and δ_2 to be adjusted). A prerequisite is the correctness of the end-points. This may be established by a forerunning circuit check (A.1).

Experiments have shown that this position/time control mechanism works extremely well in practice. Fig. 6 shows the result of a restoration of a badly recorded flight track in an AIDS data series.

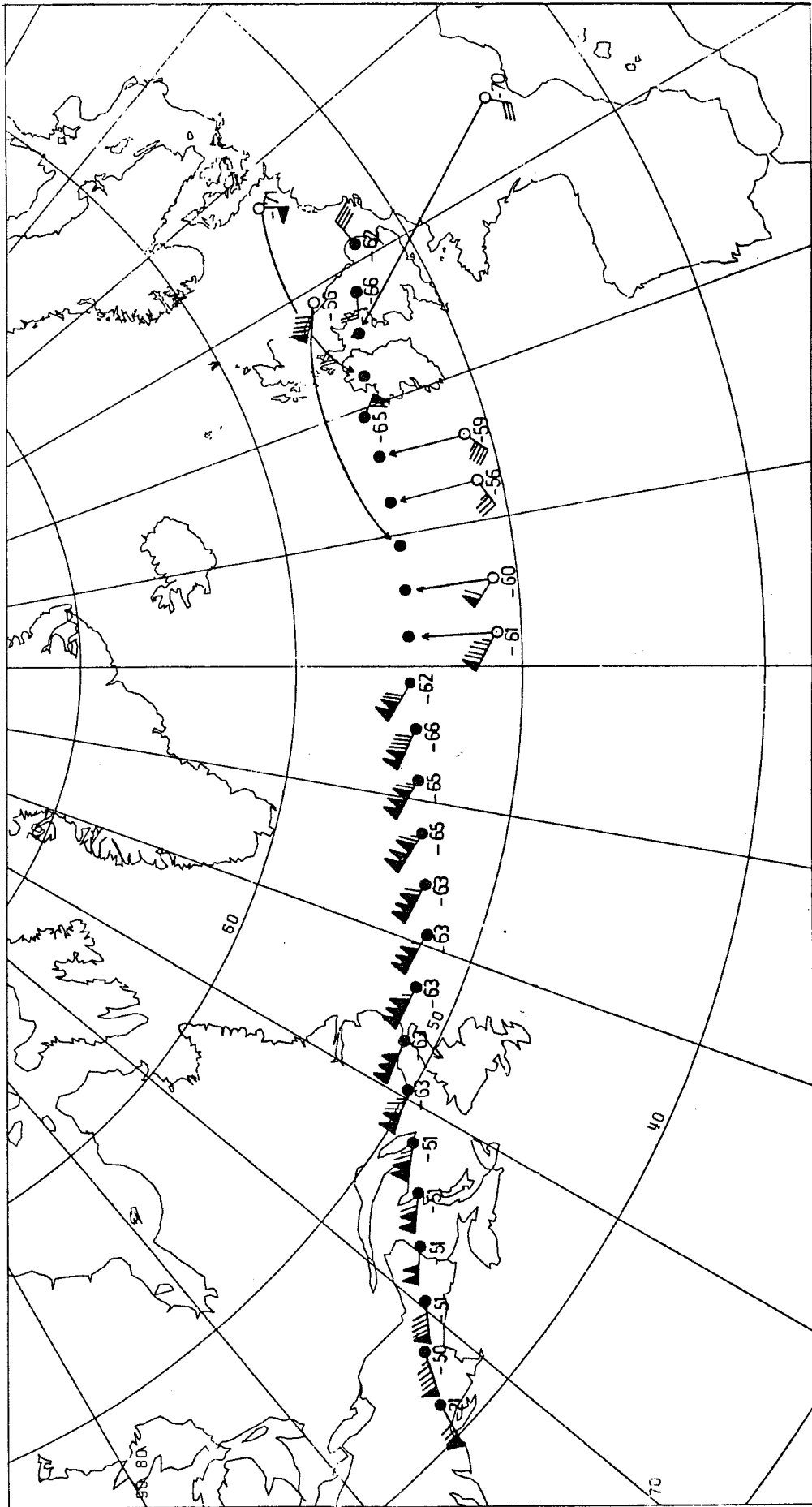


FIG.6 RESTORATION OF FAULTY AIDS DATA POINTS

3-point cruise control check (A.3)

The criterion used in this cruise control check is a generalization of that used in method A.1. It mainly expresses how far en route course changes can go during cruise control.

Consider three data points $\{\vec{x}(i), t(i)\}$, $\{\vec{x}(j), t(j)\}$ and $\{\vec{x}(k), t(k)\}$ (Fig. 7).

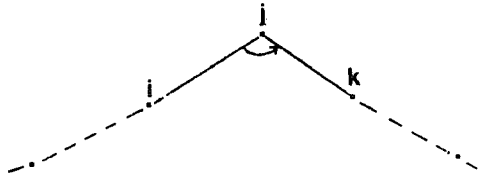


Fig. 7. For explanation see text.

The criterion is made up of three components: the criterion (II.3) for the data points (i,j) and (j,k) and a constraint expressing the maximally permitted en route course change:

$$\delta_1 \leq \left| \frac{\vec{x}(j) - \vec{x}(i)}{t(j) - t(i)} \right| \leq \delta_2 \quad (\text{II.5})$$

$$\delta_1' \leq \left| \frac{\vec{x}(k) - \vec{x}(j)}{t(k) - t(j)} \right| \leq \delta_2' \quad (\text{II.6})$$

$$\delta_1'' \leq \frac{(\vec{x}(k) - \vec{x}(j)) \cdot (\vec{x}(j) - \vec{x}(i))}{|\vec{x}(k) - \vec{x}(j)| \cdot |\vec{x}(j) - \vec{x}(i)|} \leq \delta_2'' \quad (\text{II.7})$$

This compound criterion is workable when considering a brute-force approach to check the "continuity" aspect of a data array. The criterion has been put into practice to check the identification part of Post Flight AIREPs because series of these reports show a marginal resolution not allowing a more powerful sequence test.

The following error bounds were used:

$$\begin{aligned}\delta_1 &= \delta'_1 = 250 \text{ kt} \\ \delta_2 &= \delta'_2 = 800 \text{ kt} \\ \delta''_1 &= -1, \delta''_2 = -.76 \quad ,\end{aligned}$$

corresponding to a course change of 40 degrees maximally.

ad B. Deficiencies in the meteorological elements are checked, using one of the methods to be described below.

Any uncertainty in a parameter is indicated by affixing mark = 2 to the element in question and encoding this in the international exchange or other code format.

None of the methods gives rise to the phenomenon of circularity and none of them turns into a circuit check. Moreover, no dynamics induced by the vehicle's motion play a part in the selected error criteria. On the contrary, the criteria have been inferred from a few, quite elementary diagnostic constraints.

3-point data consistency check (B.1)

The stepped altitude profile, cruise climb in SST operations and course changes in the horizontal, typical of long range flights, and the fairly high resolution of data points do suggest that series of air reports are particularly suited to test a 3-space consistency check, provided that there exist a proper diagnostic equation which links the observations in a set of adjacent data points.

An innovative 3-space consistency check, developed in the SADC, is centred about the application of an integral theorem obtainable from the equation of motion.

This integral theorem states that along an arbitrary closed curve in the atmosphere the following line integral should vanish:

$$\oint \left(\frac{f}{m} (v_g dx - u_g dy) - RT_v d \ln p \right) = 0 \quad (\text{II.8})$$

Here p denotes pressure, T_v the adjusted virtual temperature, u_g and v_g are the components of the geostrophic wind approximation, f is the Coriolis parameter, R is the gas constant of dry air, and m a map factor. The integration is performed in a local (xyz) coordinate system.

The derivation of this equation requires a somewhat lengthy calculation in which use is made of the geostrophic wind equation, the hydrostatic equation and the equation of state for moist air (ref. to Appendix II).

Given a circuit through n adjacent data points, the "circuit integral" practically vanishes after error-free values of pressure, wind and temperature have been inserted. The vanishing of the circuit integral will be violated if one or more of the observed values contain a larger error. Conversely, when the circuit integral does not vanish and exceeds some prefixed limit, this will indicate that the observations are not error-free. Therefore this integral theorem appears to be a powerful tool for checking purposes. The more so, as it can be demonstrated that this checking formula covers as special cases the hydrostatic check and checks based on the thermal wind equation and the quasi non-divergence of the wind.

It needs no arguing that this circuit integral check is worth to be tested within a series of Air reports (AIREPs, AIDS, ASDAR). To apply this checking formula in practice it is necessary to use a very simple circuit configuration. The most elementary circuit is a 3-point circuit (triangle). Let a (sloping) triangle be given by the vertices (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) . The data sets consist of temperatures T_1, T_2, T_3 , the wind velocities $(u_1, v_1, 0)$, $(u_2, v_2, 0)$ and $(u_3, v_3, 0)$ and pressures p_1, p_2, p_3 (See Fig. 8).

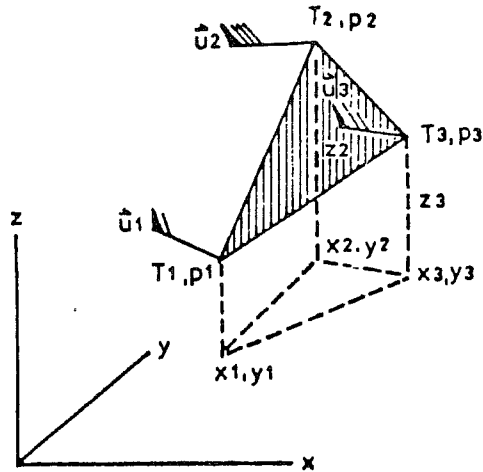


Fig. 8. Circuit integral check in a 3-point circuit.

Note: The adjusted virtual temperature is replaced by the recorded ambient air temperature, because at the levels considered in the data collection the virtual temperature increment amounts to a few tenths of a degree only.

Then (II.8) may be expressed in the following linearized form:

$$\begin{aligned}
 F = & \left(\frac{\bar{f}}{m} \right) [y_1(u_3 - u_2) + y_2(u_1 - u_3) + y_3(u_2 - u_1)] \\
 & + x_1(v_2 - v_3) + x_2(v_3 - v_1) + x_3(v_1 - v_2) \\
 & + R \left[T_1 \ln \frac{p_2}{p_3} + T_2 \ln \frac{p_3}{p_1} + T_3 \ln \frac{p_1}{p_2} \right] \approx 0 \quad (\text{II.9})
 \end{aligned}$$

The circuit integral formula (II.9) is computed for all successive 3-tuples of data points when running a sequence test, whereas simultaneously the moving flagging method affixes the right marks to the meteorological elements. There are two extremes in which the circuit integral method operates, viz.

- (1) 3-tuple of colinear data points in the horizontal (at constant pressure altitude);
- (2) 3-tuple of colinear data points in the vertical.

ad (1) Let the x-axis be taken through all three data points. Then $y_1 = y_2 = y_3 = 0$. As $p_1 = p_2 = p_3$, F reduces to:

$$v_1(x_3-x_2) + v_2(x_1-x_3) + v_3(x_2-x_1) = 0 \quad (\text{II.10})$$

which simply expresses that the shear of the wind component normal to the track should be constant. This special case dominates the control of a wind component in AIDS observational series for straight horizontal portions of the flight. It can also be said that the circuit integral check in this case reduces to a horizontal check of the cross wind, using interpolation.

ad (2) In this case F reduces to

$$T_1 \ln \frac{p_2}{p_3} + T_2 \ln \frac{p_3}{p_1} + T_3 \ln \frac{p_1}{p_2} \cong 0 \quad (\text{II.11})$$

stating that the control of temperature in the vertical is governed by a constant temperature lapse rate. The circuit integral check reduces here to a control of temperature in the vertical, using interpolation. In AIDS data samples this specialization (II.11) does occur very approximatively in the climb and descent phases of the flight.

Of paramount importance is the selection of a suitable tolerance limit for δ_1 and δ_2 in the error criterion:

$$\delta_1 \leq F \leq \delta_2$$

Here $\delta_1 = \delta_2 = \delta$ is taken. It is assumed that the location of data points (x_i, y_i, z_i) , $i=1,2,3$, has been subjected to a consistency control process already in advance. Then δ is made to depend on x_i, y_i and z_i in such a way that the consistency of temperature and wind data is well defined. In the choice of δ account should be taken of the fact that the circuit integral holds for a geostrophic approximation of the upper wind, which is known to be very crude in the

airspace considered here and which breaks down in tropical areas, say, between 20° N and 20° S. It is also supposed that the time variability of the circulation may be neglected within the velocity régime of the air traffic in short flight portions. Besides, the error criterion should also apply in the two extreme cases described above.

This has led to the selection of the following threshold for δ :

$$\delta = 12 h_{\max} + 0.07 d_{\max} |\sin \lambda| \quad (\text{II.12})$$

where h_{\max} and d_{\max} (in km) denote the maximum vertical and horizontal spacing in the 3-tuple of adjacent data points. For $|\lambda| < 20^{\circ}$ the second term on the right is made to vanish, so that this check is rendered inoperative for level flight in the tropics. In data arrays the use of this 3-space consistency check is most profitable in portions of the flight track where the 3-tuples of data points are not colinear, i.e. in the ascent and descent, cruise climb SST, steps and course changes. As already stated in Section 2.7, the checking method is most effective in the discrimination of error-free data. To clear up the detected data inconsistencies in the data arrays more rigorously, it is required to look for additional methods of error checking like those to be described below. These will also be operative in the tropics.

This 3-point consistency check supported by a 3-point flagging algorithm has proven to be invaluable for control purposes along space-embedded trajectories. Its effectiveness in data series other than provided by equipment on board aircraft will still require a closer investigation.

Consistency check using interpolation (B.2)

Along track use can be made of a sequence test involving interpolation in three consecutive data points.

Let $\psi(\vec{x})$ be some physical quantity observed in point \vec{x} and let $\psi(i)$, $\psi(j)$ and $\psi(k)$ denote the observed values in three points $\{\vec{x}(i), t(i)\}$, $\{\vec{x}(j), t(j)\}$ and $\{\vec{x}(k), t(k)\}$, which points had been subjected already to a position/time control check and labelled as being correct.

Then compute in data point $\{\vec{x}(j), t(j)\}$ an estimate for ψ by interpolation and compare the result with the observed value $\psi(j)$.

The criterion is then

$$\delta_1 \leq \psi(i) + \frac{t(k)-t(j)}{t(k)-t(i)} (\psi(k)-\psi(i)) - \psi(j) \leq \delta_2, \quad (\text{II.13})$$

and when interpolated in terms of positions:

$$\delta_1 < \psi(i) + \left| \frac{\vec{x}(k)-\vec{x}(j)}{\vec{x}(k)-\vec{x}(i)} \right| (\psi(k)-\psi(i)) - \psi(j) \leq \delta_2 \quad (\text{II.14})$$

The method, supported by a 3-point flagging operation, has been introduced to resolve the deficiencies in en route temperatures and pressure altitude,

threshold values ($\delta_1 = \delta_2 = \delta$):

air temperature : $\delta = .005 |\vec{x}(k)-\vec{x}(i)|$ °C,

where the distance is expressed in km.

pressure altitude: $\delta = 50$ m.

Consistency check using along-track gradients or rates of change (B.3)

Along track, especially in level flight, use can be made of a 2-point sequence test using the rate of change or gradient of some physical quantity derived from data in two successive and proximate data points as error criterion.

Let $\psi(i)$ and $\psi(j)$ be given in two adjacent data points $\{x(i), t(i)\}$ and $\{x(j), t(j)\}$, then, in general, criteria may be formulated in the form:

$$\delta_1 \leq \frac{\psi(j) - \psi(i)}{|x(j) - x(i)|} \leq \delta_2 \quad (\text{II.15})$$

$$\delta_1 \leq \frac{\psi(j) - \psi(i)}{t(j) - t(i)} \leq \delta_2 \quad (\text{II.16})$$

$$\delta_1 \leq F(\psi(j), \psi(i)) \leq \delta_2 \quad (\text{II.17})$$

etc.

The method has been applied to check speed and direction of the wind, denoted by \vec{v} .

Criteria:

change in wind direction, cos. rule (II.17):

$$.76 \leq \frac{\vec{v}(j) \cdot \vec{v}(i)}{|\vec{v}(j)| \cdot |\vec{v}(i)|} \leq 1 ,$$

corresponding to a wind change of 40 degrees.

wind shear (II.16):

$$0 \leq \frac{||\vec{v}(j)| - |\vec{v}(i)||}{|\vec{x}(j) - \vec{x}(i)|} \leq 0.2 ,$$

$|\vec{v}|$ in knots, distance in km. This corresponds to a shear of approximately 40 knots per 100 nm.

2.9 Duplicates and near-identical reports

The problem of duplicate and near-identical reports is becoming of concern internationally.

There is a strong evidence that occurrence of such reports is very pronounced in the class of Aircraft Meteorological Reports.

To mention a few causes:

1. duplication involved in normal data distribution via the Main Trunk Circuit and its branches;
2. distribution of the same reports via separate channels of communication with inherently a risk of mutilations of the reports;

3. Level II-a reports which are received also as Level II-b reports;
4. Reports of observations which are made very proximate in space-time dimensions when using different sampling systems ("sampling interference");
5. duplication occurring during data tape handling.

Examples:

- o Conventional AIREPs received air-to-ground by more than one Aeronautical Meteorological Office and onward transmission via GTS (ex. 2).
- o ASDAR data relayed via two geostationary satellites to ground stations (ex. 2).
- o Post Flight AIREPs received in a delayed mode, identical to conventional AIREPs (ex. 3).
- o ASDAR data points very close to those of conventional AIREPs, ditto for interfering AIDS data and AIREPs, ASDAR- and AIDS data (ex. 4).
- o Transcription of AIDS data cassette tapes more than once to a master tape (ex. 5).

The detection and rejection of the reports in question need to be done at all levels and at all stages in the information flow. Instructions on how to identify and how to deal with these reports are missing in the Plan.

The SADC tackled this data problem as follows: each record (report), see page 3, was split up into 16 elements, using the Fortran format:

(5I2, A6, 2I2, A1, I3, I2, A1, I3, I4, 2I3)

When in a (monthly) file of records during reading, a record had 14 elements or more in common with a previously scanned record, the record was rejected and removed from the file.

2.10 Validity check of the date

The most critical elements in the records to be checked thoroughly are the elements day, month, year. An invalid date makes the reports virtually useless for research.

The SADC put considerably resources in checking these identification elements both numerically and by visual inspection.

In the AIDS instrumental set-up the date is the only parameter to be inserted manually, prior to take-off.

One of the most recurrent date errors is attributable to dialling in the local calendar date instead of date GMT.

To check the date in sets of AIDS reports, the "777..." and "888..." indicator groups, see page 3, play an important part. 888... groups act as interface between series of observations made in successive flights; a 777... group announces the end of reading of a cassette tape when unloading in the Airline's processing centre.

A sequence of 888... groups together with the enumeration of observation times pertaining to the first and last record, in general suffice to draw a conclusion on what the date really should be, to correct it and put it back into the records.

In a man-machine interactive approach the computer produces lists of date/time groups for end-points of successive flights recorded in one individual A/C AIDS cassette tape plus a first indication of which dates are suspect and what they should be. Then the manual interaction comes into play either to confirm the machine's finding or to reject it, to amend it if necessary and put the restored dates back into the data set.

As to the date-control for Post Flight AIREPs, received in a delayed mode, such a control had to be abandoned simply for the reason that information on the actual time schedules was missing altogether.

It is plausible that the "date problem" is an acute one in many FGGE Level II-b data sets.

It manifested itself clearly in AIDS data sets and most probably it will arise in Post Flight AIREPs also. When planning for future regional and global experiments, the date problem for Level II-b data sets should be studied carefully.

2.11 Format and code checks

Mutilations in the agreed formats did not show up in the automatically encoded AIDS data records. In the relatively small file of Post Flight Air reports format errors were exceptional owing to the verification of the data being punched on magtape. Punching errors which had been overlooked added to mutilations in the data content, undergoing quality tests at a later stage of the control programme.

Code errors were more frequent. They could be ascribed to a misinterpretation of coding instructions. They also could be a consequence of the use of non-standard units or could stem from a bilaterally agreed departure from code regulations.

The a priori knowledge of departures from agreed formats, codes and code tables was sufficient to subject the records in advance to a special sub-programme of format and code corrections and -substitutions.

III. STRATEGIES

The Centre's data processing programme was set up to cover quality control operations for series of

- Post Flight AIREPs
- Wide bodies AIDS data
- Concorde SST/AIDS data
- ASDAR data (ad hoc)

The strategies to be planned had to be somewhat different for each category of reports. In what follows, the main strategy is described for the AIDS data series provided by the Boeing 747 and 747 Special Performance, and Douglas DC-10 (see Section 3.1).

The strategy planned for Concorde AIDS is described in Section 3.3.

Where the methods A.1, A.2, A.3 and B.1, B.2 and B.3 (Section 2.8) make an important contribution in the strategies to be described, it may come about that some parameters experience a multiple checking.

The marking or labelling of these parameters performed by successive flagging operations may then interfere; however, it is so arranged that on command one marking operation overrides another, dependent on the effectiveness of the checking algorithm considered.

3.1 AIDS data (ASDAR ad hoc)

Strategy:

- (a) format and code check; correction, substitution, flagging (mark = 0 or 9).
- (b) search for duplicate and near-identical reports; eliminations.
- (c) date validity check; restoration or elimination.
- (d) gross error check position/time; elimination.
- (e) gross error check pressure altitude/flight level; flagging (mark = 0 or 3).
- (f) consistency check pressure altitude using method B.2; flagging (mark = 1 or 2).
- (g) gross error check air temperature and wind; flagging (mark = 0 or 3).
- (h) position/time consistency check using methods A.1 and A.2; correction c.q. elimination, flagging (labels).
- (i) temperature and wind consistency check using methods B.1, B.2 and B.3; flagging (mark = 1 or 2).

- Notes:
- (i) For gross error check limits see Tables 3, 4 and 5.
 - (ii) When both the circuit check A.1 and search-light method A.2 are generated, the phenomenon of circularity is observed. Restoration of position/time takes place using the method as outlined in Section 2.5.
 - (iii) The error bounds for the various tests have been given earlier, when these tools for error checking were described. In ASDAR data arrays the sampling period is usually 400 s, half that for AIDS. Consequently, a few error bounds had to be re-scaled with a factor 2.

3.2 Post Flight AIREPs

The strategy planned for this group of reports is the same as traced out for AIDS data except that a few methods were made inoperative.

Strategy:

- (a) as AIDS
- (b) as AIDS
- (c) non-operative
- (d) as AIDS
- (e) as AIDS
- (f) non-operative
- (g) as AIDS
- (h) position/time consistency check using method A.3; elimination.
- (i) as AIDS

Notes: (i) Correction of corruptive positions or observation times has been abandoned here, owing to the marginal spacing of adjacent data points.

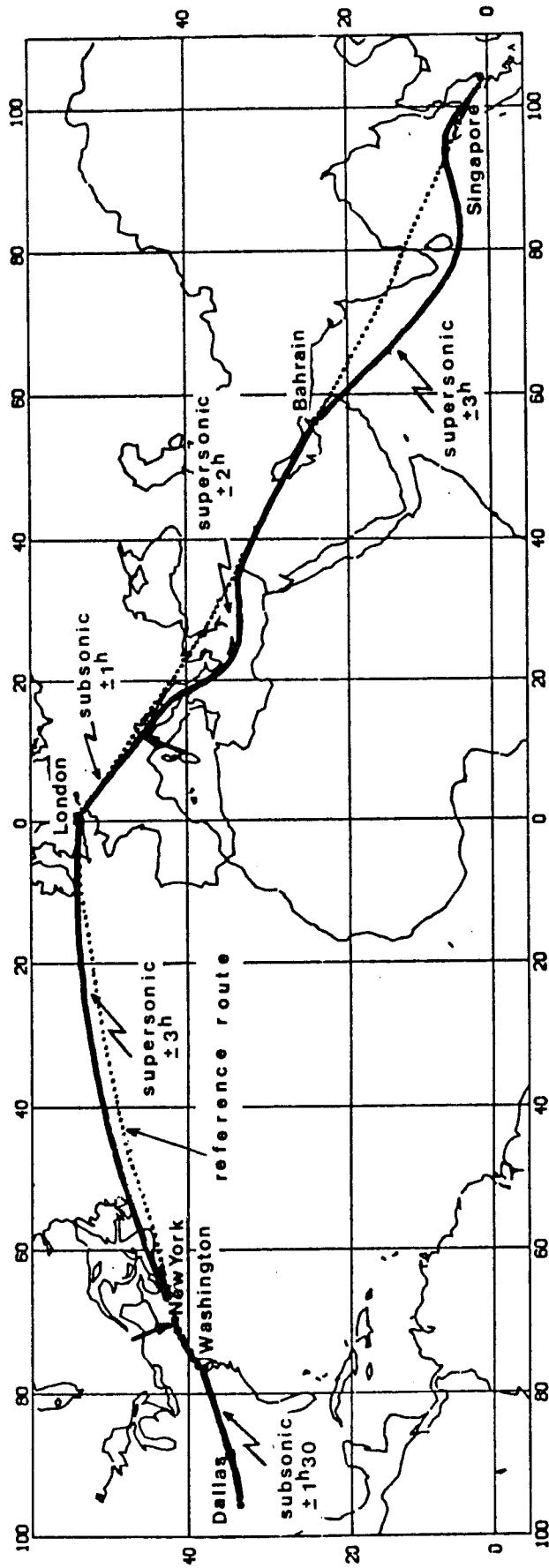
(ii) To carry out a date validity check was out of the question. Therefore a word of caution: the FGGE basic data set may contain Level II-b Post Flight AIREPs with a corruptive coded date (day, month, year).

3.3 Concorde SST/AIDS data

During the FGGE operational year British Airways delivered separate reel tapes to the SADC loaded with Concorde AIDS data series. For specific Concorde flight-, performance and observing characteristics reference is made to Appendix I and Fig. 4.

Fig. 9 presents the flight routes involved, together with an indication of the subsonic (most over land) and supersonic (most over water) flight portions and corresponding flight duration in hours.

With these features in mind the following course of actions was mapped out for quality control purposes:



CONCORDE SST FLIGHT ROUTES

FIG. 9

Strategy:

- (a) as AIDS
- (b) as AIDS
- (c) as AIDS
- (d) gross error check position/time using an additional route check, cf. note (i); elimination.
- (e) as AIDS
- (f) consistency check of pressure altitude using method B.3, see note (ii); flagging (mark = 1 or 2).
- (g) as AIDS
- (h) as AIDS, see note (iii).
- (i) as AIDS

Notes: (i) To strengthen the gross error check for position/time an additional check is used, based on the off-track distance relative to a reference route. For simplicity straight line connections are taken in a linear latitude-longitude projection with 1-degree scaling as reference routes (cf. Fig. 9). Offset limits in degrees:

London-New York (Washington)	5
New York (Washington)-Dallas	3
London-Bahrain	6
Bahrain-Singapore	9

(ii) Method B.3 is used to check the altitude profile. This profile is temperature-dependent (cf. Fig. 4), so that the altitude change per unit time may vary within a wide range. As an estimate for the maximum rate of climbing, when in the cruise climb mode, is taken 2.5 m s^{-1} . The limits δ_1 and δ_2 in expression (II.16) are:

$$\delta_1 = 0, \quad \delta_2 = 2.5$$

(iii) In the important circuit check (method A.1) the upper and lower bound for the cruise speed need to be carefully specified, because this speed is highly variable for the Concorde (cf. Appendix I). Let Δ denote the sampling period. In expression (II.3) the following limits δ_1 and δ_2 have been taken (knots):

$\Delta < 400 \text{ s}$ or ($\Delta > 400 \text{ s}$ and $\text{Fl.} > 40$):

$$\begin{aligned} \delta_1 &= .8 \times 1150 \\ \delta_2 &= 1.2 \times 1150 \end{aligned}$$

$\Delta > 400 \text{ s}$ and $\text{Fl.} \leq 40$:

$$\begin{aligned} \delta_1 &= .4 \times 850 \\ \delta_2 &= 1.2 \times 850 \end{aligned}$$

3.4 Examples

A count of records, which involved mark = 2, 3 or 9 for suspect, erroneous and missing data, yielded as a result only a few tenths of a per cent. of the bulk volume.

In this section two examples are worked out, one referring to an AIDS data series and one to an ASDAR data series.

(1) AIDS data series

Fig. 10 presents a print-out of a series of AIDS data records collected during a flight Bangkok-Karachi. Here the format is the AIDS data record format (see page 3 and [2]), plus three characters: one for flagging pressure (pressure altitude Z_p), one for flagging air temperature T and one for flagging wind \vec{v} . Part of the data content is plotted in a map.

In the three columns reserved for flagging there is one mark (mark = 2) indicating a suspect wind value.

The plotted data reveal that the suspect value really is a mutilated value in the spotted location and is not related to some exceptional feature in the wind field.

The somewhat erratic winds recorded near the fourth line remain undetected, because winds are weak there.

(2) ASDAR data series

During 1978 the performance of a prototype ASDAR unit, which was installed in a NASA research aircraft, was tested.

Using a magtape containing ASDAR data, an experiment was arranged to subject ASDAR data to the same quality program as developed for AIDS, but upgraded to withstand the experimental fourfold higher frequency in ASDAR data.

The ASDAR data are grouped in half-hourly frames consisting of eight records each. The sampling period is therefore 225 s, corresponding to a horizontal spacing of 30 to 35 nm.

Fig. 11 shows a map with plotted data for a flight made on 26 August 1978 with the prototype ASDAR unit. Not all data could be plotted because of lack of space. Some frames are missing.

A copy of part of the quality-controlled ASDAR file is shown, written in the same format as used in Example 1.

When examining the last three flagging columns, it appears that wind and pressure altitude are flagged as being correct (= 1). However, almost all temperature data are marked as being suspect (= 2). A quick glance at the column for temperature reveals that the en route temperatures look very erratic with errors up to 7° C.

The root of this trouble had been discovered already by NOAA. In one of the ASDAR STATUS REPORTS (August 1978) it reads: "Logic in the initial ASDAR contained a flaw that permitted errors (of up to 6° to 8° C) at two points along the temperature scale".

IV. CONCLUSION

Information on the ultimate results of the quality control effort on material committed to the care of the SADC is given in the final report on the participation of the Centre in the first global experiment. [6].

The overall integrity of the FGGE basic data set stands or falls with the quality of all its subsets. This implies that a thorough analysis is desired of the effectiveness of the control as reflected in the marks affixed to the parameters, in other words to investigate the "quality of the quality control mechanism" itself.

In such an effort the support is needed of other processing centres. A valuable tool is the use of graphical displays as used in man-computer interactive data access systems.

An important point in the whole set-up of a control mechanism is the choice of proper error- and tolerance limits. In the Plan there is no guidance on what limits should be taken.

The Centre established its own limits, consulting literature and looking for those used by other FGGE data flow centres. The Centre also interpreted the results of some experimental studies carried out on AIDS data, in order to obtain a first impression of empirical limits in physical elements sec and their along-track gradients and rate of change.

The tolerance- and error limits had to be quite coarse. Probably most criteria assisting in the quality control of FGGE data subsets suffer from this defect. On the other hand, when choosing too sharp limits, it will happen that the control mechanism signalizes data as being suspect where in reality these refer to outstanding atmospheric features.

It is recommended that in the planning of future regional and global experiments one be prepared to strive after an internationally agreed standardization of methods and techniques for quality control. In the meantime the experience gained during the first global experiment has contributed much to an a posteriori knowledge of the error-characteristics in the various groups of Aircraft Meteorological Reports.

For example, some ad hoc quality control experiments performed in ASDAR data series have proven that from the point of view of data quality the ASDAR system is superior to the current WMO/ICAO Air-reporting system and is also, in other respects, most versatile.

REFERENCES

- [1] Implementation/Operations Plan Vol. No. 3. FGGE Data Management Plan.
- [2] Implementation/Operations Plan Vol. No. 5. The FGGE Special Observing Systems, Part E: Aircraft Integrated Data System.
- [3] Implementation/Operations Plan Vol. No. 4. The World Weather Watch in FGGE.
- [4] ICAO-Annex 3 to ICAO convention - Meteorology, Chapter 3.
- [5] de Jong, H.M., 1976, Study on the system mix of radiosonde, aircraft and satellite observations in the North Atlantic region. Scientific Report W.R. 76-5, Royal Netherlands Meteorological Institute, De Bilt.
- [6] Report on the Participation of the Special Aircraft Data Centre in the First GARP Global Experiment. WGDM-5/Doc. 9, ADD. 1.

SOME FLIGHT PERFORMANCE- and OBSERVING CHARACTERISTICS
of importance for the control of Concorde AIDS data.

1. Route pattern.

The flight pattern of AIDS providing Concorde's consisted during the FGGE operational year of two fixed routes:

London - New York (Washington) - Dallas v.v.
 London - Bahrain - Singapore v.v.

Fig. 9 depicts the route pattern together with an indication of flight duration in subsonic and supersonic flight portions. The dotted curves represent control reference routes between the cities mentioned above. The lateral departure of the actual routes from these reference routes has been used to formulate a position/time gross error criterion.

2. Altitude profile.

Concorde altitude profiles are complicated by international agreements prohibiting or allowing supersonic flight over various countries. When Concorde flies subsonically, the flight is straight and level, unless an altitude change is authorised by Air Traffic Control, at a height of between 25000 and 40000 ft. When Concorde flies supersonically, after the initial climb to approximately 50000 ft, it maintains a constant Mach number, so that variations in temperature cause the aircraft to rise and fall slightly, with an overall tendency to rise as fuel is expended. This is called a "cruise climb". Because of the up-and-down movement, Concorde's are separated by a lateral criterion of approximately 3 miles, rather than a vertical criterion.

3. Descent.

On those routes where a considerable proportion is subsonic, there will be two stages of descent: a) from supersonic to subsonic, b) from subsonic to landing. At the tops of those descents the altitudes will be in the ranges a) 50000 to 60000 ft, b) 25000 to 40000 ft. On mostly supersonic flights range a) applies. The altitude ought never to exceed 60000 ft, but very occasionally a Concorde reaches 60500 ft.

4. Course changes.

Currently lateral course changes above 20000 ft do not exceed 45° .

5. Subsonic flight.

Currently, above 20000 ft 19% of time is spent in subsonic flight on the North Atlantic route, and 40% on the Bahrain route. Over all flights this averages 22%.

6. Speed control.

During subsonic flight, the ground speed is approximately 600 knots. The actual groundspeed depends on ambient wind conditions. Stratosphere winds in general are weak.

During supersonic flight, the Mach number is held steady at 2.1 ± 0.1 , which corresponds to a ground speed of approximately 1150 ± 50 knots, depending on wind and temperature conditions.

7. Sampling of AIDS observations.

The sampling period is 800 sec. when subsonic and every 400 sec. when supersonic with the change over at Mach = 1.5 (ground speed approximately 900 knots).

This avoids giving unnecessary resolution at subsonic speeds.

During subsonic climb above 20000 ft, Concorde has a ground speed of between 550 and 590 knots, corresponding to 226 to 242 km per 800 seconds. The aircraft passes through Mach 1.0 between 30000 and 40000 ft and is in supersonic climb between 630 knots and 1150 knots, corresponding to 130 to 237 km per 400 seconds.

Note: Since a large roll-angle, experienced during a course change affects the reliability of AIDS data no data are supplied during a roll, when the roll-angle exceeds 3° .

DERIVATION OF A LINE INTEGRAL THEOREM.

In an elementary application of the equation of horizontal motion a special integral theorem may be formulated, which can be utilized as a powerful diagnostic tool for various purposes especially to check the meteorological elements in an array of data points.

In mid- and high latitude synoptic scale systems the wind and pressure fields are in approximated geostrophic balance. In vector form this balance may be expressed as

$$f \vec{v}_g = \vec{k} \times \frac{1}{\rho} \nabla_z p \quad (1)$$

where \vec{v}_g is the geostrophic velocity and $\nabla_z p$ the horizontal pressure gradient.

$f = 2\Omega \sin \phi$ is the Coriolis parameter;
 Ω = the angular speed of rotation of the earth;
 ϕ = latitude;
 ρ = density;
 \vec{k} = a unit vector pointing to zenith.

In isobaric coördinates the vectorial form of the geostrophic relationship is:

$$f \vec{v}_g = g \vec{k} \times \nabla_p z \equiv \vec{k} \times \nabla_p \phi$$

g = the acceleration of gravity;
 ϕ = the geopotential defined as the work required to raise unit mass from the surface to the height z ;
 $\phi = \int_0^z g dz$

To represent the atmospheric flow in arbitrary surfaces, which are not level, we shall have to derive the horizontal pressure gradient along the surface.

Let Σ be an arbitrarily defined sloping surface in the atmosphere. The surface Σ is specified by some scalar quantity $S = \text{const}$. The surface Σ is a physical surface if S represents a physical parameter.

APPENDIX II-2

Consider a cross section in the (xz) plane. Then the following holds:

$$\left(\frac{\partial p}{\partial x}\right)_z = \left(\frac{\partial p}{\partial x}\right)_S - \left(\frac{\partial p}{\partial z}\right) \left(\frac{\partial z}{\partial x}\right)_S$$

The subscript S means "holding S constant".

If we substitute on the right from the hydrostatic equation and introduce the geopotential, we obtain:

$$\left(\frac{\partial p}{\partial x}\right)_z = \left(\frac{\partial p}{\partial x}\right)_S + \rho \left(\frac{\partial \Phi}{\partial x}\right)_S$$

A similar equation may be written for a cross section in the (yz) plane.

These equations relate the horizontal pressure gradient force per unit mass to the pressure gradient force per unit mass on the surface.

In vectorial form:

$$\nabla_z p = \nabla_S p + \rho \nabla_S \Phi \quad (2)$$

where $\nabla \equiv i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y}$ denotes the horizontal gradient operator.

When Eq. (1) and Eq. (2) are taken together, then the result is the appropriate form of the geostrophic relationship in the Σ surface:

$$f \vec{v}_g = \vec{k} \times \nabla_S \Phi + \frac{1}{\rho} \vec{k} \times \nabla_S p \quad (3)$$

Let Σ be the support of an arbitrarily chosen curve Γ connecting the points P and Q located on the surface. Then we consider the line integral:

$$\int_P^Q \Gamma f [\vec{k} \times d\vec{s}] \vec{v}_g \quad (4)$$

where \vec{s} defines the integration path.

Substitution of Eq. (3) in Eq. (4) yields:

$$\begin{aligned} \Gamma \int_P^Q \mathbf{r} [\vec{k} \times d\vec{s}] \vec{v}_g &= \Gamma \int_P^Q \mathbf{r} [d\vec{s} \times \vec{v}_g] \vec{k} \\ &= \Gamma \int_P^Q (d\vec{s} \times [\vec{k} \times \nabla_S \phi]) \vec{k} + \Gamma \int_P^Q \frac{1}{\rho} (d\vec{s} \times [\vec{k} \times \nabla_S p]) \vec{k} \end{aligned}$$

In view of the vector rule

$$\vec{a} \times [\vec{b} \times \vec{c}] = (\vec{a} \cdot \vec{c}) \vec{b} - (\vec{a} \cdot \vec{b}) \vec{c} \quad \text{we have}$$

$$\begin{aligned} \Gamma \int_P^Q \mathbf{r} [d\vec{s} \times \vec{v}_g] \vec{k} &= \Gamma \int_P^Q (d\vec{s} \cdot \nabla_S \phi) (\vec{k} \cdot \vec{k}) - \Gamma \int_P^Q (d\vec{s} \cdot \vec{k}) (\nabla_S \phi \cdot \vec{k}) \\ &+ \Gamma \int_P^Q \frac{1}{\rho} (d\vec{s} \cdot \nabla_S p) (\vec{k} \cdot \vec{k}) - \Gamma \int_P^Q \frac{1}{\rho} (d\vec{s} \cdot \vec{k}) (\nabla_S p \cdot \vec{k}) \end{aligned}$$

Observing that

$$\begin{aligned} \vec{k} \cdot \vec{k} &= 1 \\ \text{and } \vec{k} \cdot \nabla_S \phi &= 0 \\ \vec{k} \cdot \nabla_S p &= 0 \end{aligned}$$

the expression on the right can be written:

$$\Gamma \int_P^Q \nabla_S \phi \cdot d\vec{s} + \Gamma \int_P^Q \frac{1}{\rho} \nabla_S p \cdot d\vec{s}$$

$$\text{but } \Gamma \int_P^Q \nabla_S \phi \cdot d\vec{s} = \Gamma \int_P^Q d_\Gamma \phi = \phi_Q - \phi_P$$

$$\text{and } \Gamma \int_P^Q \frac{1}{\rho} \nabla_S p \cdot d\vec{s} = \Gamma \int_P^Q \frac{d_\Gamma p}{\rho}$$

The index Γ indicates here that the increments in Φ and p have to be taken along the curve Γ . In a local (xyz) coordinate system the integrand $f [\text{d}\vec{s} \times \vec{v}_g] \cdot \vec{k}$ takes the form $f (v_g dx - u_g dy)$. The integral (4) can best be expressed in a fixed Cartesian coordinate system referring to a conformal map projection. The integral then becomes

$$\Gamma \int_P^Q \frac{f}{m} (v_g dx - u_g dy)$$

where m is the map factor. For instance, in a conformal polar stereographic projection $m = \frac{2}{1 + \sin \Phi}$.

Furthermore, by substituting from the equation of state for moist air:

$$p = \rho RT_v,$$

where R is the gas constant for dry air and T_v the adjusted virtual temperature, we obtain the following expression:

$$\Phi_Q - \Phi_P = \Gamma \int_P^Q \frac{f}{m} (v_g dx - u_g dy) - \Gamma \int_P^Q RT_v d \ln p$$

Finally, when integrated along a closed curve:

$$\oint \left(\frac{f}{m} (v_g dx - u_g dy) - RT_v d \ln p \right) = 0.$$

This theorem holds within a geostrophically approximated air motion at a fixed time. The theorem admits various interpretations, dependent on its use in selected applications. Specializations of the theorem offer the possibility to derive well-known diagnostic relations in dynamical meteorology, e.g. the thermal wind equation and the Montgomery stream function valid for isentropic surfaces.

Obviously, the theorem is particularly suited to be used for checking purposes in series of consecutive observations, provided by upper air soundings, dropsondes, constant level balloon flights, long range flights of commercial aircraft.