Lecture #2 PRACTICAL WORK ON QUALITY CONTROL AT THE NMC

In this lecture I plan to expand upon some of the observing system specific quality assessment algorithms mentioned in my previous lecture.

Upper Air Observations. In this category are placed radiosonde, rawinsonde, and all report types containing wind profile observations. Work by W. Collins and L. Gandin (1988,89) has resulted in a new algorithm recently made operational, which makes hydrostatic consistency checks on all reported geopotential heights and temperatures. Using the reported 'full-field' values, the height profile is reconstructed using the reported temperatures; and the temperature profile using the reported heights. A logic tree then examines possibilities for inconsistency arising from encoding or transmission errors. These are; transposition of digits, missing digits, differences of 10 or 100, and incorrect algebraic sign. If the inconsistency can be confidently assigned to one of these errors, the offending datum is corrected before entry into NMC's data base. Table 1 presents summary statistics for one month on the frequency of these hydrostatic errors.

Aircraft Observations.

The new algorithm follows two principles adopted in the previous lecture; viz., the quality control is done by considering unique features of the aircraft wind reporting system, and the algorithm uses full-field values. It has as its purpose:

- o the application of consensus-type quality decisions at those navigation check-points where 3 or more reports are collected within a six-hour time block.

- o to create superobservations when appropriate, preserving information on wind variations with time and altitude.

- o to append a series of hierarchical quality marks depending upon the degree of consensus information.

If only a single report occurs at a latitude-longitude point, it is not considered in the algorithm. If two reports exist co-located or within the superob radius, and the altitude, time, and vector difference are within prescribed limits, the two are superobed: otherwise they pass unchecked.

If 3,4 or 5 reports are co-located, a complicated series of checks occurs based upon observed vector differences between all combinations of data at the same and different levels. These differences are compared with a statistical distribution of vector differences compiled from a large sample of similar aircraft reports and parameterized by

TABLE 1
SUMMARY OF HYDROSTATIC ERROR-CHECKING ALGORITHM AT THE NMC
January 1989

Explanation of error types:

Type	Description
1	confident height correction
2	confident temperature corr.
3	correction pair
4	error at lowest level
5	error at top level
6	isolated large residual

etc,etc

Note: only error types 1 & 2 cause the report to be corrected.

ERROR COUNTS BY ERROR TYPE AND PRESSURE:

PRESS	1	2	3	4	5	6	7	8	9	10	11	12	22
100C	0	0	0	83	0	0	0	0	0	0	C	0	0
85C	30	35	7	77	0	1	2	2	2	1	0	5	0
70C	83	89	14	16	0	11	2	1	5	1	1	12	7
50C	75	67	19	4	22	24	6	2	5	2	32	4	8
40C	84	68	16	0	8	21	14	1	1	2	46	6	4
30C	64	62	23	1	13	20	9	1	4	1	46	5	3
25C	75	58	14	5	9	11	3	2	2	2	57	2	4
20C	58	28	20	1	16	12	3	3	1	2	49	6	7
150	41	25	7	0	12	10	7	3	1	2	27	5	5
10C	29	23	8	1	29	4	8	1	2	1	9	6	0
70	37	16	22	4	60	143	5	3	1	1	9	0	5
5C	40	13	4	0	19	7	9	1	2	2	1C	2	2
3C	46	15	3	0	32	8	6	0	1	1	23	1	2
2C	28	2	2	0	30	1	2	1	1	0	7	1	0
1C	0	C	2	0	38	0	0	0	0	0	C	0	0
TOTAL	690	501	161	192	288	273	76	21	28	18	316	55	47

ERROR COUNTS BY STATION BLOCK (ALL ERROR TYPES):

	0	1	2	3	4	5	6	7	8	9
C	0	3	8	2	5	0	1	1	21	0
1C	11	3	16	8	0	29	10	44	0	0
2C	14	17	11	30	60	32	16	6	25	32
30	49	52	17	10	11	18	22	29	67	0
40	30	35	175	158	121	0	27	30	104	0
50	17	76	38	26	77	18	72	65	65	32
6C	66	29	14	14	3	5	0	16	12	0
7C	9	142	21	0	2	0	56	0	40	0
8C	12	1	22	40	2	5	1	32	0	15
90	0	55	C	50	78	0	86	59	0	0

ERROR COUNTS BY STATION BLOCK (ERROR TYPE = 1):

	0	1	2	3	4	5	6	7	8	9
0	0	1	3	0	1	0	1	0	5	0
10	2	1	7	5	0	3	3	13	0	0
2C	4	4	4	6	20	14	3	2	10	8
30	9	23	2	4	4	6	8	8	19	0
4C	11	6	37	25	27	0	4	11	21	0
5C	8	15	12	5	21	3	10	16	7	9
6C	20	10	3	1	0	0	0	8	4	C
7C	5	44	6	0	1	0	13	0	6	0
8C	2	0	3	10	0	2	0	9	0	5
9C	0	20	0	19	13	0	22	8	0	0

altitude and time differences. If the difference from any pair exceeds a selectable number of standard deviations of that distribution, both observations receive an integer mark. When all pairs have been checked, these marks are summed, and a logic tree is entered to mark bad observations. The remaining observations are either averaged to create a superob, or marked as passing a quality check. This portion of the code has required considerable tuning and the best procedure and amount of checking desirable may require further tuning.

If 6 or more observations are available, quality control is done by calculating means and standard deviations(vector statistics) without each of the observations in turn. Comparison of the statistics with and without each of the reports in turn is used to discard bad or suspect observations. Then, two-dimensional interpolation is used in time and altitude to generate superob winds at the analysis time and level(s).

Some statistics on the distribution of the number of navigation check points having various numbers of co-located reports are: at 00UTC, about 380 points have a single report; 96 have two; 77 have 3,4 or 5; 15 have 6 to 10; and 12 points have more than 10. The latter mostly occur on the U.S. west coast to Hawaii route.

Emphasis should be made here as to the need to monitor the aircraft reporting system to identify those points in the system that are introducing error. (Almost all of the errors occurring in the aircraft report file have a source in the reporting system, not in the wind measuring system itself.) To pursue this monitoring, and to promote the usefulness and quality of the system, NMC intends to establish additional contacts with ATA,ARINC, and ICAO.

Satellite Cloud-motion Vectors

At present SATOB reports, cloud motion vectors, are known to have certain deficiencies which require special treatment for assimilation into numerical analyses. Special monitoring efforts at numerical weather prediction centres (to be described in Lecture #3) have shown convincing evidence that high level SATOB reports are biased in wind speed, with respect, of course, to the assimilating forecast. The exact reason for this bias is not known, and is the subject of present research. At the NMC various procedures have been discussed for accommodating the SATOB data to mitigate the bias. Testing of various alternatives has not been completed, and the final solution will not be reported here. This speed bias problem is a good example of the validation process I defined in the first lecture. Tracking cloud imagery to calculate displacements is the basic measurement; the validation of those measurements as wind vectors, however, is the important, crucial, step not satisfactorily completed.

Satellite Temperature Profiles (SATEMS)

The problem of quality assessment of satellite-based temperature or thickness soundings remains, after a decade and a half a thorny problem. The unique characteristics of these observations and their uncertainties or errors make assessment in an operational context difficult. At the NMC, work at present is toward an interactive retrieval scheme-one in which the information used in converting measured radiances to thicknesses comes from the assimilating model itself. Such interactive schemes are thought to have promise (Eyre and Lorenc (1989)), but will require some thorough investigation because of possible correlation of measurement error and assimilating model error.

Surface Marine Observations

The treatment of surface marine data, both from ships and buoys, is discussed by G. Hamilton elsewhere in this Workshop.

Land Surface Observations

Surface synoptic reports from land stations do not receive any thorough quality assessment from nwp assimilation schemes at NMC. First of all, only the sea-level pressure is used in the analysis process. This information is converted to 1000hPa (or 850hPa) geopotential height information for use at the appropriate analysis level. Second, only a sub-set of the surface synoptic observations are used: the sub-set is created either by thinning or by created super-observations. Quality assessment, and utility, of these observation type is difficult because of differences in the actual as opposed to the model terrain configuration. This area of quality assessment, however, is surely one which demands attention.