EVALUATION OF THE COLLECTION, ARCHIVING AND PUBLICATION OF DAILY SNOW DATA IN THE UNITED STATES

David A. Robinson
Department of Geography
Rutgers, The State University
New Brunswick, New Jersey 08903

Abstract: Snow is an important component of the earth's environment, and can significantly impact the everyday life of those living in a large portion of the United States. However, accurate and complete information on snowfall and snow cover is presently being collected at only 57% of the official climate observing stations in states where regional snows occur on an annual basis. The quality of the data varies significantly from one state to another, although it tends to be weakest where snow is least common. As a result of this evaluation, the list of stations with daily snow data published in National Oceanic and Atmospheric Administration *Climatological Data* publications has been revised and expanded. [Key words: snow, snow observations, climate data, United States.]

INTRODUCTION

Across much of the United States, the impact of snow on human activity and the environment is considerable. Falling snow or snow lying on the ground affect weather, climate, hydrology, biology, environmental chemistry, earth-surface processes, engineering, agriculture, travel, recreation, commerce and safety, among others. In turn, the presence or state of snow is influenced by weather, climate, topography, proximity to water bodies, human beings and other factors. Examples of literature addressing these often complex associations include: Rooney (1969), Changnon (1969), Kukla (1979), Gray and Male (1981), Goodrich (1982), Namias (1984), Telfer and Kelsall (1984), Barry (1985), Hall and Martinec (1985), Goodison et al. (1987) and Walsh (1987).

For geographers and others to explore further these relationships, it is imperative that accurate information on snowfall and snow cover be available. Satellite sensors provide important information on large-scale snow cover extent and show potential for providing snow depth data (Hall, 1988). However, there remains a great need for gathering snowfall and snow-cover data at ground stations (Barry and Armstrong, 1987). In the United States, this includes data from the several hundred first order synoptic stations recording hourly data, plus the over seven thousand daily cooperative stations. To rely solely on first order reports would be ill advised, considering the high spatial variability of snow and the potential serious biasing of first order data from urbanization effects.

Physical Geography, 1989, **10**, 2, pp. 120–130.

Here, the quality of snow data presently being gathered at stations across the U.S. is evaluated. Project results have been used to revise the list of stations with daily snow data published in *Climatological Data* (CD), a monthly National Oceanic and Atmospheric Administration report with issues for 44 U.S. states or regions. CDs are where most users first turn for snow data. It is important to note that all data reported by first order and cooperative stations are retained in a digital data base at the NOAA National Climatic Data Center (NCDC) in Asheville, NC. However, with this revision, the first in over ten years, users of snow data will know better which stations are keeping reliable records and have these data available in a published format.

This evaluation was conducted while the author was a visiting scientist at NCDC, and was part of a major effort to assemble the first quality-controlled digital set of long-term (50 years or more) daily snow data (Robinson, 1988). This set of over 1000 primarily cooperative stations will be employed in climatic research, including assessing the utility of snow as an indicator of climate change.

MONITORING SNOW OVER CONTINENTS

Snowfall

Direct means of identifying where and how much snow is falling or has fallen include using ground-based radar and station observations. The former distinguishes between snow and liquid precipitation to some extent, but has difficulties identifying the presence of a light snowfall. Station observations remain the primary means of monitoring snowfall, despite the spatial limitations of observing networks. Indirect means of estimating snowfall over land include evaluating changes in snow cover derived from satellite shortwave and microwave data. These presently have little or no utility in quantitative studies of snowfall.

Snow Cover

Among the various means of gathering information on continental snow cover, aerial surveys have high spatial resolution and permit the percent coverage of snow to be determined. However, their overall spatial and temporal coverages are severely restricted.

Snow-cover information from satellites includes that derived from shortwave and passive microwave sensors (cf. Hall and Martinec, 1985). Shortwave data facilitate continental assessment of snow coverage with a relatively high spatial resolution. Information on surface albedo is also gleaned from the data. Shortcomings include the inability to ascertain cover when solar illumination is low or absent or when skies are cloudy; the lack of all but the most general information on pack depth; and the relatively short (two decade) record, which has increased in quality with time (Dewey and Heim, 1982; Wiesnet et al., 1987). Microwave-derived snow products provide continental coverage. The capacity to recognize snow cover is independent of solar illumination or the absence of clouds (Kunzi et al., 1982; Foster et al., 1984; McFarland et al., 1987). Certain information on depth and liquid-water content can be acquired, although its accuracy remains suspect. Spatial resolution is lower than in the shortwave but is generally sufficient for large-scale regional studies, except where snow cover is patchy. It is difficult to identify shallow or wet snow using microwaves and no direct information on surface albedo may be obtained. Microwave snow products are only available for the past ten years.

Station reports of snow cover have the disadvantages of being point measurements and having inadequate spatial coverage outside the lower elevations of middle latitudes (Robinson and Kukla, 1988). Positive aspects of station data include daily measurements of snow depth, observations of snow-pack water equivalence and long periods of record. These advantages, along with a dense network of observing sites, make station snowfall and snow-cover data particularly useful over much of the United States.

EVALUATION PROCEDURE

The quality of snow observations at each official U.S. observing station was appraised by applying a step-wise series of three tests to monthly snow data (Fig. 1) for each of the 7637 mainland U.S. stations in operation as of August 1988. Study data were from December through March for the winters of 1985-86, 1986-87 and 1987-88. To facilitate the analysis, stations were arranged according to the 353 state climatic divisions in the mainland U.S. (including Alaska). In the first test, snowfall totals had to be

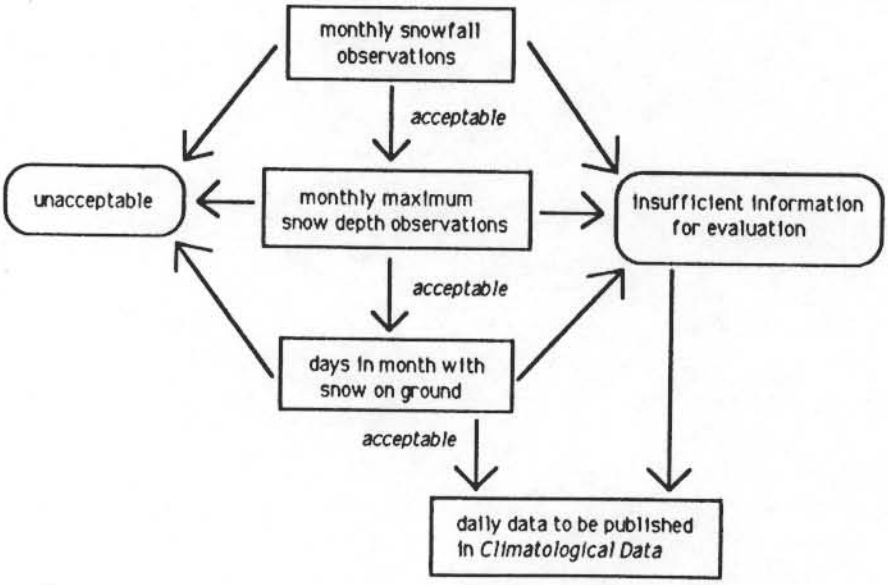

Fig. 1. Step-wise procedure employed in the evaluation of U.S. first order and cooperative station daily snow data.

deemed acceptable in at least nine of the twelve study months. This judgement was made in an educated and objective manner, and gave the benefit of the doubt to any borderline questionable data whenever appropriate. Of the stations remaining after the first examination, the same evaluative criteria were applied to monthly maximum snow depths. Few observers make satisfactory snow depth observations while at the same time making inferior snowfall measurements. Finally, those stations passing the first two tests had to have at least nine of twelve acceptable months of data for number of days with one inch (2.5 cm) or more of snow on the ground. This served to identify those stations which tend to record snow depth only on the day of a snowfall event, ignoring snow cover on subsequent days.

Snow observations at stations passing the three tests are considered to be of a relatively high quality. Stations failing any one of the three tests tended to have an abundance of missing data or data showing zero values when it was apparent from surrounding stations that some snow had fallen or had lain on the ground. An insufficient number of regional snow events in southern and west coast states prohibited assessing whether a given station had improperly recorded snow in at least three months of the study period. In a few cases where one or two regional events occurred, stations were dropped if it was obvious that they were entirely missed by the observer.

Nine acceptable months of snow data were not necessary for the 91 stations which began reporting after December 1985. To receive a positive evaluation, these stations only had to have fewer than three months of missing or suspicious data. Miskeying at NCDC may have led to stations failing the tests, although these cases should be few, given the criterion of needing three or more months of poor data. Finally, the late arrival of data at NCDC may explain a few negative station evaluations.

RESULTS OF THE REVIEW PROCESS

In the 49 mainland states, 5871 or 77% of the stations passed the monthly snowfall test and remained under consideration for the snow-depth examination. Some 4938 stations remained following the depth test and 4847 stations after the days-on-ground evaluation. Of the 4847, some 2026 are from states bordering the Pacific Ocean (excluding Alaska), Mexico, the Gulf of Mexico or the Atlantic Ocean south of 37°N. In at least several climate divisions in these states, an insufficient number of snow episodes in the test data set prohibited a realistic state-wide evaluation of data quality. Given enough cases to study, it is probable that many more stations in these states would be found to have unacceptable snow data. This is suggested by the poor quality of observations in the divisions of these states which were able to be evaluated, as well as in nearby states where snow is somewhat more common.

The remainder of the evaluation will therefore concentrate on results from the 37 snowier states, where a sufficient number of snow events

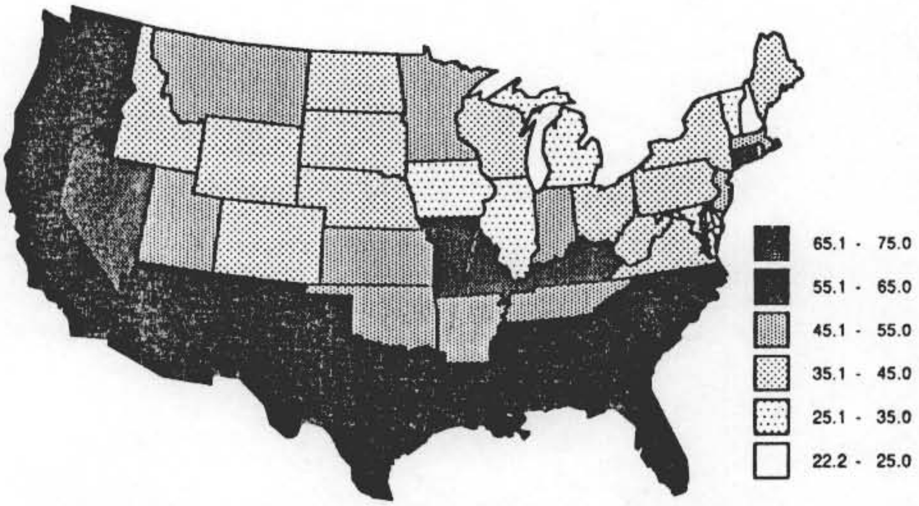

Fig. 2. Percentage of climatic stations failing to pass one of the three snow examinations in the 37 state "snowy" region (Alaska = 20%). Those states with insufficient data for evaluative purposes are shown in black.

occurred to permit confident state-wide assessments. Of the 4960 stations in this region, 2139, or approximately 43%, failed one of the three examinations (Fig. 2). Some 28% failed the snowfall test (Fig. 3), with 13% failing owing to poor maximum snow-depth data, despite having acceptable snowfall observations (Fig. 4), and another 2% failed owing to inadequate days-on-ground data. Table 1 accentuates the positive, presenting state-wide totals for stations passing the tests. All first-order stations in the 37 state group passed the three tests, indicating that the observation problems lie

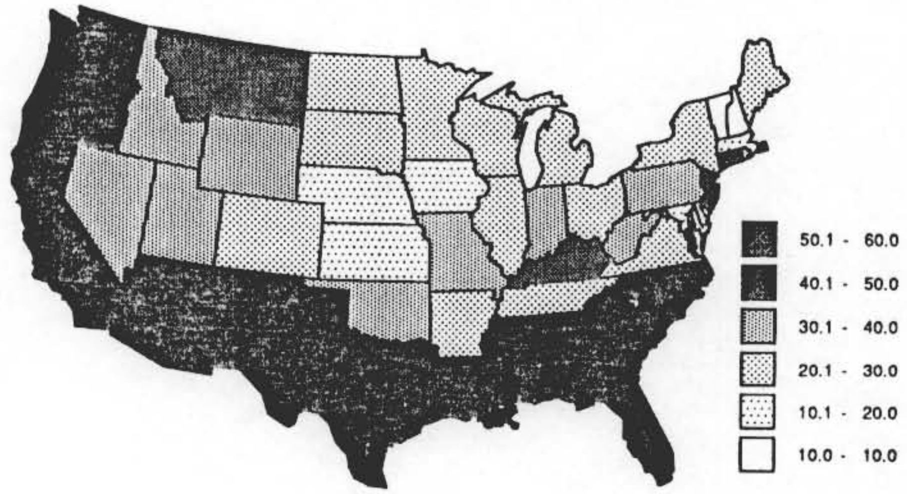

Fig. 3. Percentage of climatic stations failing to meet the criteria established for acceptable snowfall data (Alaska = 20%). State coverage same as in Figure 2.

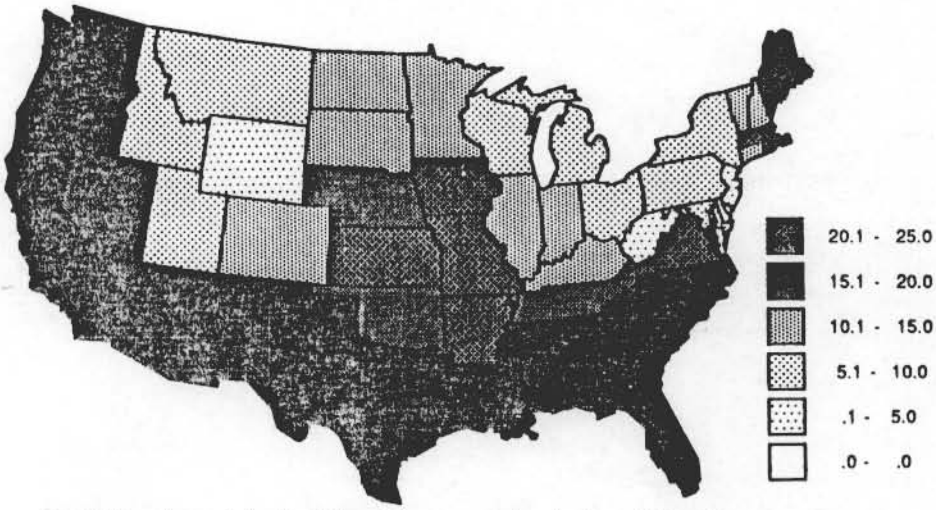

Fig. 4. Percentage of climatic stations failing to meet the criteria established for acceptable maximum snow depth data, after having passed the snowfall test (Alaska = 0%). State coverage is the same as in Figure 2.

with the cooperative stations. In fact, considerably more cooperative stations would have failed had the 9 of 12 months criterion been tightened.

Within the snowier region, there is a tendency for data quality to be best where snow is quite common (e.g., Alaska and states bordering the Great Lakes). Differences in the quality of snowfall data are often substantial in neighboring states (e.g., within New England and between Nebraska, Iowa, Illinois and Indiana). Missouri stations often fail to record satisfactory snowfall and snow-cover data, while states such as Vermont and Alaska are strong on both accounts. Some states, Iowa in particular, have a high percentage of stations recording credible snowfall data but far fewer providing accurate snow cover observations. Wisconsin has relatively good snowfall and snow-depth data, but 10 stations meeting these standards fail to collect accurate days-on-ground information.

Differences between states may be due to variations in the emphasis given snow observations amongst NOAA Cooperative Program Managers and state climate programs. Also contributing to this may be inconsistencies in the manual quality control effort at NCDC, resulting in incomplete totals being identified and flagged as missing in some states (or months) while passing through in others. NCDC employees tend to process different states each month, so testing this hypothesis is difficult.

CLIMATOLOGICAL DATA DAILY SNOW LIST

All stations which passed the three examinations, following the step-wise procedure shown in Figure 1, have been included in the revised *Climatological Data* daily snow list. Stations with insufficient evaluative information will also be included until such time that adequate data are available for their

Table 1. Results of the Three Step Evaluation of Snow Data for the 37 States Where a Confident Assessment of the Quality of Snow Observations Could Be Made

		Stations with Acceptable Data on		
State	Total Stations[a]	SF[b]	SF and MD	SF, MD and DG
Alaska	153	123	123	123
Arkansas	151	118	80	76
Colorado	210	157	131	128
Connecticut	34	16	12	11
Delaware	9	8	8	7
Idaho	140	92	81	80
Illinois	176	139	116	115
Indiana	106	70	54	53
Iowa	161	144	105	105
Kansas	247	199	139	135
Kentucky	176	98	72	69
Maine	59	46	35	35
Maryland	52	42	37	37
Massachusetts	71	57	40	39
Michigan	165	126	114	110
Minnesota	179	127	103	94
Missouri	191	125	81	78
Montana	263	148	123	120
Nebraska	228	184	141	134
Nevada	102	61	44	44
New Hampshire	50	45	39	38
New Jersey	56	33	30	30
New York	250	174	159	156
North Dakota	151	105	87	83
Ohio	143	103	94	92
Oklahoma	202	131	95	94
Pennsylvania	213	147	131	129
Rhode Island	6	4	3	3
South Dakota	145	114	96	94
Tennessee	101	74	55	52
Utah	189	129	107	101
Vermont	39	35	30	28
Virginia	132	104	76	76
West Virginia	107	69	65	65
Wisconsin	175	137	121	111
Wyoming	128	85	78	76
Totals	4960	3569	2905	2821

[a]Total cooperative and first order stations reporting as of August 1988.

[b]SF is snowfall, MD is maximum snow depth, DG is days with snow on the ground.

proper evaluations. This will permit publication of any accurate data coming from these stations. The 4847 stations selected almost doubles the 2574 stations which were formerly published. The need for this revision is clearly illustrated by the fact that 646 (25.1%) of the stations formerly published failed to make the updated daily snow list.

Plans are for NCDC to continue revising the list on an annual or biennial basis. Recommendations as to how the revision might be done in an objective, accurate and efficient manner using monthly data are presented in Appendix 1. They illustrate the detailed and interactive effort required to evaluate climatic data correctly.

IMPROVING THE COLLECTION AND ARCHIVING OF SNOW DATA

The collection and archiving of accurate and complete daily snow data for all stations across the U.S. should not be considered an unrealistic goal. Steps should be taken to encourage the many observers who presently collect inferior snow data to begin gathering better information. The recent publication of snow-observing tips and the anticipated publication of a condensed version of this report in the *National Cooperative Observer Newsletter* is a start. Continued education by Cooperative Program Managers and state climatology programs is also essential. In addition, Cooperative Program Managers must take it upon themselves to scrutinize the data their observers are collecting and correct those who are doing a poor job.

At NCDC, two steps are needed to improve the quality of the snow data being digitized and archived. First, an additional flag needs to be placed in the automated quality control routine which is presently applied to all daily cooperative data. This flag should identify days with measurable precipitation and with a maximum temperature of less than 30°F (−1°C), yet with no snowfall reported. Some freezing rain events will be flagged, but experience has shown that these typically can be identified through ancillary data and notes on observer forms. All other flagged events should be considered incorrect and the snowfall digitized as missing, unless it is determined that it is the report of measurable precipitation which is incorrect, in which case this value should be replaced with a zero.

Unfortunately, no flag can unambiguously resolve the problem of identifying missing snowfall events on days where the maximum temperature exceeds 30°F. Here, a careful pre-screening of the data by NCDC personnel prior to digitization is essential if such omissions are to be caught and flagged as missing before getting into the data set as zeros. This activity needs greater emphasis, particularly until the automated quality control routine is amended and until better data begin to arrive at NCDC. Staff time should be made available for these functions.

CONCLUSION

Whether it snows every other day or every other decade, high-quality station observations of snowfall and snow cover are important. Without them, vital data are being lost and, worst yet, under the present circumstances, incorrect data are being entered into the national climatic archives. Presently, only somewhat more than half of the official observing stations in the U.S. are collecting snow data considered of an acceptable quality, even using relatively liberal evaluative criteria. Observing problems are primarily found at cooperative stations. Through education, technical changes and an improved effort by all of those involved in the data collection and archiving process, accurate and complete snow data can become a reality. Until that time arrives, it is essential that geographers and all others who employ snow data, or for that matter any climate data, in their research be cognizant of the fact that the data must be carefully screened before being used.

Acknowledgments: Thanks to R. Quayle, G. Goodge and A. McNab for their advice and encouragement throughout this project, and to L. Dimmick for assistance with the data reduction. The author's stay at the National Climatic Data Center was funded by grants ATM 87-19865 and ATM 89-96113 from the Office of Climate Dynamics, National Science Foundation.

BIBLIOGRAPHY

Barry, R. G. (1985) The cryosphere and climate change. In *Detecting the Climatic Effects of Increasing Carbon Dioxide,* U.S. Dept. of Energy DOE/ER-0235, 109–148.

_______ and Armstrong, R. L. (1987) Snow cover data management: the role of WDC-A for Glaciology. *Hydrological Sciences Journal,* Vol. 32, 281–295.

Changnon, S. A., Jr. (1969) *Climatology of Severe Winter Storms in Illinois.* Illinois State Water Survey Bulletin 53.

Dewey, K. F., and Heim Jr., R. (1982) A digital archive of Northern Hemisphere snow cover, November 1966 through December 1980. *Bulletin of the American Meteorological Society,* Vol. 63, 1132–1141.

Foster, J., Hall, D. K., Chang, A. T. C., and Rango, A. (1984) An overview of passive microwave snow research and results. *Reviews of Geophysics and Space Physics,* Vol. 22, 195–208.

Goodison, B. E., Barry, R. G., and Dozier, J. (eds.) (1987) *Larger Scale Effects of Seasonal Snow Cover.* International Association of Hydrological Sciences Publication No. 166.

Goodrich, L. E. (1982) The influence of snow cover on the ground thermal regime. *Canadian Geotechnical Journal,* Vol. 19, 421–432.

Gray, D. M. and Male, D. H. (1981) *Handbook of Snow.* Pergamon Press, 776pp.

Hall, D. K. (1988) Assessment of polar climate change using satellite technology. *Reviews of Geophysics,* Vol. 26, 26–39.

Hall, D. K., and Martinec, J. (1985) *Remote Sensing of Ice and Snow.* Chapman and Hall, New York, 189pp.

Kukla, G. (1979) Climatic role of snow covers. In *Sea Level, Ice, and Climatic Change (Proceedings of the Canberra Symposium),* International Association of Hydrological Sciences Publication 131, 79–107.

Kunzi, K. F., Patil, S., and Rott, H. (1982) Snow-cover parameters retrieved from Nimbus-7 SMMR data. *IEEE Transactions in Geoscience and Remote Sensing,* Vol. GE-20, 452–467.

McFarland, M. J., Wilke, G. D., and Harder, II, P. H. (1987) Nimbus 7 SMMR investigation of snowpack properties in the northern Great Plains for the winter of 1978-1979. *IEEE Transactions in Geoscience and Remote Sensing,* Vol. GE-25, 35–46.

Namias, J. (1984) Some empirical evidence for the influence of snow cover on temperature and precipitation. In *Proceedings of the Ninth Annual Climate Diagnostics Workshop,* U. S. Department of Commerce, 32–43.

Robinson, D. A. (1988) Construction of a United States historical snow data base. In *Proceedings of the 1988 Eastern Snow Conference,* 50–59.

_______ and Kukla, G. (1988) Comments on "Comparison of Northern Hemisphere snow cover data sets". *Journal of Climate,* Vol. 1, 435–440.

Rooney, J. F., Jr. (1969) The economic and social implications of snow and ice. In *Water, Earth and Man,* Methuen, London, 389–401.

Telfer, E. S., and Kelsall, J. P. (1984) Adaptation of some large North American mammals for survival in snow. *Ecology,* Vol. 65, 1828–1834.

Walsh, J. E. (1987) Large-scale effects of seasonal snow cover. In *Large Scale Effects of Seasonal Snow Cover,* International Association of Hydrological Sciences Publication 166, 3–14.

Wiesnet, D. R., Ropelewski, D. F., Kukla, G. J., and Robinson, D. A. (1987) A discussion of the accuracy of NOAA satellite-derived global seasonal snow cover measurements. In *Large Scale Effects of Seasonal Snow Cover,* International Association of Hydrological Sciences Publication 166, 291–304.

APPENDIX 1

Recommendations as to how the *Climatological Data* daily snow list might be revised in an objective, accurate and efficient manner using monthly data.

(Step 1) Identify (flag) a station which does not report snowfall in a month when at least 80% of the stations in its climatic division do report snow. Eliminate those stations which are flagged more than a selected number of times in a given period (e.g., the "three and out" criterion used in the initial revision). Special procedures will be required for divisions with very few stations or for stations with significantly different snow climatologies than others in the division to avoid unjustly eliminating stations from the list (e.g., Chatham, MA). This step may be accomplished in an automated manner if special cases are identified and methods for handling them are developed prior to implementation.

(2) Repeat step one using monthly maximum snow-depth data. Do so only for those stations passing the first step.

(3) Repeat again for observations of days in the month with one inch or more of snow on the ground. Do so only for those stations passing the second step.

(4) Flag any remaining station which has monthly snowfall totals some percentage above or below that of an average of the remaining stations in its climatic division. Experimentation will be required to determine the appropriate percentages. Subjective evaluation of those cases which are flagged will be required. This should be performed by someone familiar with the climatology of the region in which the station is situated. If it is determined that the data for a given month are incorrect and the total of unacceptable months for all steps performed exceeds the selected cut off, the station is eliminated.

(5) Repeat step four procedures using monthly maximum snow-depth data. Do so only for those stations remaining after step four.

(6) Repeat step four procedures again using days in the month with one inch or more snow on the ground. Do so only for those stations passing the fifth step.