CANADA

DEPARTMENT OF TRANSPORT
METEOROLOGICAL BRANCH

QUALITY CONTROL OF SURFACE METEOROLOGICAL DATA FOR CLIMATOLOGICAL PURPOSES

J. G. POTTER

CLI—5—69
JULY 15, 1969

CANADA

DEPARTMENT OF TRANSPORT

METEOROLOGICAL BRANCH

QUALITY CONTROL OF SURFACE METEOROLOGICAL DATA FOR CLIMATOLOGICAL PURPOSES

J. G. POTTER

CLI—5—69
JULY 15, 1969

U.D.C.: 551.506.1.004.1

QUALITY CONTROL OF SURFACE METEOROLOGICAL DATA FOR CLIMATOLOGICAL PURPOSES

TABLE OF CONTENTS

1. INTRODUCTION

Traditionally climatologists are vitally concerned with the completeness and accuracy of climatological data. Data published in official data publications and accumulating in national archives must be as complete and error-free as possible. Until fairly recently data were processed by clerical methods, and at various points in this processing professional staff passed judgment on the acceptability of the data. In the last two decades however, with the advent of modern methods of data processing and with the availability of experienced technical staff, it has been possible to use more sophisticated procedures. During this period the term "quality control" came into use to describe clerical, technical, machine and electronics procedures.

There are at least six basic steps in a good meteorological data quality control system:

Step 1 – The data are scanned and those whose accuracy is in doubt are selected for further action.

Step 2 – The selected data are examined in an effort to justify their accuracy.

Step 3 – When data cannot be justified an attempt is made to determine some logical way the observer made a mistake and to correct the data accordingly.

Step 4 – When the data cannot be justified (Step 1), or the type of mistake determined and a correction made (Step 2), then the recorded value must be rejected and not included in the transfer to punched cards or further processing.

Step 5 – At this time estimated data are entered for rejected values or missing data depending on the type of data and the requirement for completeness.

Step 6 – Action is initiated with the observational programme managers for correction of faulty instruments or observer instruction and training.

It is of course desirable to automate such quality control systems as much as possible, and in processing meteorological data in a forecast system there is normally insufficient time for human intervention. At present it is virtually impossible to automate Steps 2, 3 and 6 listed above. In a fully automated system the extent to which observer errors may be accepted as unusual weather conditions, or the extent to which unusual occurrences of concern to the climatologists are rejected by the system, depends entirely on the knowledge of the personnel stating the requirements and the ingenuity of those writing the computer programmes. In such a system the usual practice of detecting suspect data by comparing them against the limits of their occurrence in the past may be employed. Under these conditions the system will reject all data outside limits of past occurrences or, if broader limits are used, accept new extremes which are actually observer or transmission errors, and thus may give no new knowledge on the occurrence of extreme conditions. Similarly, it is possible that any estimation for mission or rejected values may be biased towards mean conditions.

In processing meteorological data for climatological purposes when more time is available, it is possible to employ computing facilities to a large extent in scanning the data for selection of suspect values (Step 1) and in estimation for missing or rejected values (Step 5). Technical staff, having access to professional advice, are then used exclusively in carrying out Steps 2, 3, 4 and 6. They may also carry out some of the scanning of data in Step 1, and in confirming estimates made by computer or doing the actual estimating required in Step 5.

This circular describes the quality control system currently in use in the Climatology Division in checking the two main types of meteorological data used for climatological purposes, i.e. data in the hourly weather observations and the observations of daily extremes of temperature and totals of precipitation. Sections 1-5 contain a general description of the system. Appendices A and B contain detailed specifications of all checks imposed on the data. These Appendices are included for possible references by others interested in developing similar systems. They will also provide a permanent documentation of the methods currently employed in the Climatology Division in checking data as may be required by future users of the data.

2. QUALITY CONTROL STEP 1 – USE OF COMPUTER FACILITIES

Cudbird (1961, 1968) has described the various types of unit record equipment used in the Climatology Division since 1950 and the current IBM 360 system installed in September 1966 (for a period August 1965 – August 1966 an IBM 1440 computer was used on a temporary basis). Most of the development work in the use of unit record equipment in quality control was carried out prior to the beginning of the International Geophysical Year (IGY) in July 1957. The introduction of the IBM 360 computer in 1966 made it necessary to write a completely new set of programmes for scanning data for quality control. These new programmes included all the basic concepts used with the unit record equipment, but many refinements were possible with the more powerful computer.

In developing requirements for programming the IBM 360 computer system for selection of suspect data in Step 1 of the quality control system it was found useful to organize these requirements under 5 different types of checks. As a matter of fact, these types are not peculiar to computer checking in particular, but are equally useful in instructing technicians in visual scanning of data. The 5 types of checks imposed on data for selection of suspect data are:

Type A – A check that the reported value of an element in an observation is within limits. These limits may be either arbitrary limits or actual limits of past occurrences;

Type B – A check that the various elements or an observation are consistent with each other;

Type B – A check when applicable on mathematical computations;

Type D – A check of the change in an element of an observation against the same element in an earlier observation from the same station against arbitrary limits established for a logical change in time;

Type E – A check against imposed limits for the variation in the reported values for an element for station to station.

In describing some of the ways in which each type of check is used by the computer in selecting suspect data, references will be made mainly to the scanning of hourly weather data (#1 punched card) and daily climatological data (#4 punched card). Detailed listings of all requirements for computer programmes to carry out these checks are included in Appendix A and Appendix B. For illustration purposes references will be made to only a portion of these checks.

TYPE A – CHECK AGAINST LIMITS

Checking data against limits to select suspect data is one of the more common checks used in the quality control of all types of data. For example, in the hourly weather observations all values of mean sea level data outside the limits 960-1060 mb. may be selected as suspect data. Similarly, acceptable limits may be set for temperature, dew point, etc. In the same manner provision may be made to select very high wind speeds (70 miles or over) or very low relative humidities (less than 10%) for verification. In addition, in this type of check on the hourly weather observations it is possible to restrict the values punched in the card to the reportable values defined in the instructions to the observers and select all others for verification. This applies to the reportable values assigned to ceiling, visibility, weather, obstructions to vision, wind direction and clouds. With the IBM 360 computer configuration it is also possible to carry a

table of monthly extreme values (maximum, minimum temperatures, 24 hour rainfall, 24 hour precipitation, etc.) and check values reported by a station against those in the table. These tables may be actually based on values reported at the station in the past, or may be composed of estimated values such that any data reported beyond these limits in the future will be checked and verified.

TYPE B – CHECK FOR CONSISTANCY

In some types of observations transferred to punched cards there must be some relationship in the data arising from definitions or from normal occurrences in nature. As examples of the first it will be noted that in the hourly weather observations the ceiling value must be reported as unlimited if the total opacity is not more than 5 tenths, and a ceiling other than unlimited reported when the opacity is greater than 5. Provision is made in the computer programme for selecting any occurrences other than these two conditions. When a ceiling value other than unlimited is reported it is further checked for agreement with the reported heights in the cloud layer columns. If the same value cannot be found for one of the cloud layers, the observation is suspect. Similarly, a check may be imposed on the reported visibility and the reported intensity of those types of precipitation whose intensity is defined in terms of visibility. Examples of combinations of events in the hourly weather observations that are selected for further examination before acceptance as true occurrences in nature include: occurrences of freezing rain outside the the temperature range 20-33°F, occurrences of liquid precipitation not indicated as freezing at temperatures below 30°F, occurrences of all types of solid precipitation, except hail, at temperatures above 42°F. In applying this type of check to climatological data it is usual to include a simple check that the maximum temperature for the day is equal or greater than the minimum temperature.

TYPE C – MATHEMATICAL CHECK

The extent to which computations performed by the observer may be later checked by the computer will depend on the amount and precision of the data transferred to punched cards. For example, in checking the psychrometric data where the original computations were carried out using temperatures to the tenth of a degree, and these temperatures were rounded off to the nearest degree when transferred to punched cards, an exact reproduction of the calculations is impossible. In this case, tabular data containing the upper and lower limits of both dew point and relative humidity possible with each pair of dry and wet bulb temperatures rounded off to the nearest degree have been prepared and are available in the storage of the computer system. The psychrometric data in each punched card containing hourly weather observations are checked against this tabular data. If any reported value of dew point or relative humidity lies outside the limits, the value is listed as suspect. This check is interesting in that while it will select errors of 1 degree or more in dew point and over 3% in relative humidity at higher temperatures, it becomes a much coarser screen at low temperatures. At temperatures below zero where difficulty in maintenance of the ice bulb and inability to read thermometers with the precision required result in errors which make all psychrometric data suspect, there is practically no check possible on the computations.

TYPE D – CHECK FOR CHANGE IN TIME

In nature the change in many elements with time is in an ordered manner and unusual variations from the normal may be selected as suspect. One example of this is the check imposed on the station and sea level pressures reported by stations in the hourly weather observations. The computer compares the station and sea level pressure at each observation with the

value reported for the preceding hour. If the difference is 4 mb. or more, the data are listed as suspect. Similarly, in the hourly weather observations, limits for change from hour to hour in dry bulb temperature, dew point, wind direction, wind speed, etc. may be established for selecting data for further examination. When the time interval between observations is greater than 1 hour this type of check becomes less valuable. In checking pressures with longer intervals between observations it has been found more useful to calculate the reduction to sea level pressure for each observation and list differences of these for scanning by technical staff. Setting limits for the day to day variation has also proved useful in computer scanning of daily extremes of temperature at both principal and ordinary climatological stations, even though these limits are quite large. Currently a limit of 20°F or more in the maximum temperature and 30°F or more in the minimum temperature is used in checking day to day variations at the same station. It has proved useful in directing attention to stations with poor data or to stations recording data on the wrong day.

TYPE E – CHECK FOR CHANGE IN SPACE

Although very little has been done in the Climatology Division in utilizing checks between reports of the same data from two stations at the same time for most types of data, it has proved very useful in scanning the reports of daily extremes of temperature and precipitation from climatological observers. The computer is used in this check to prepare 4 separate listings of (1) daily maximum temperatures, (2) daily minimum temperatures, (3) daily precipitation and, (4) resultant monthly totals, means and extremes from the processed data (fig. 2-5). In each of the first 3 listings the daily values for each station in a climatological district (fig. 1) are printed on a single line. The computer sorts the stations for listing in a designated geographical sequence. Instances are flagged where there is a difference greater than 10°F in the maximum temperatures, and greater than 15°F in the minimum temperatures, reported by adjacent stations. The technical staff then scan the listings of data, using flagged values as guides to problem stations, and select and take action on suspect values.

There are many areas in which the computer is similarly used simply to list observed data, or summarize them, and technical staff then follow through all six steps of the quality control procedures. The most extensive scanning by technical staff is done in connection with the listings of daily temperature extremes and daily total precipitation from climatological stations as described above. While the flags put on the temperature data by the computer are helpful to indicate variations from day to day at a station, or between stations on the same day, the technical staff normally do a more exact comparison in time and space whenever it is known that data reported by a particular station are likely to contain errors. In quality control of reported daily totals of precipitation, the principal use made of the computer is to produce the listings of data. Technical staff then scan the data for selection of super values and carry out other steps as described in the next section.

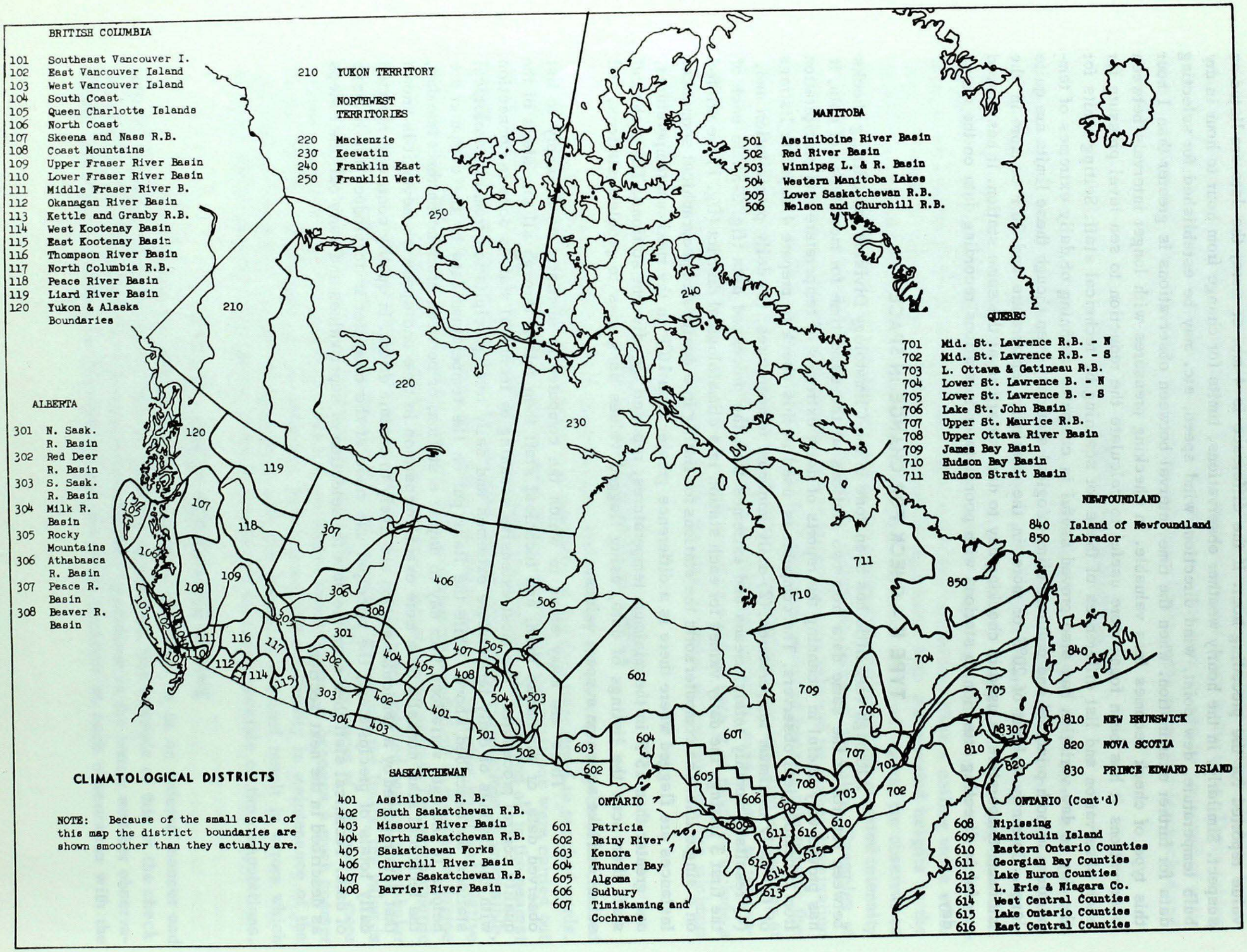


FIGURE 1 – CLIMATOLOGICAL DISTRICTS

604 THUNDER BAY 68 04 MAX. ARRAY

GEO	NO	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	MEAN	DIFF
140*	0325	35	34	25	29	35	46	41	36	36	45	54	60	41	37	41	51	60	50	53	55	43	43	35	29	33	47	51	58	58	56		43.9	.8
215	9096	36	39	28	28	38	51	42	35	41	45	69 T	58	46	32	52	63	59	49	55	48	45	55	33 T	28	36	48	53	57	57	60		46.2	.8
260	3930	42	43	38	31	43	50	44	40	48	47	67	71 S	50 T	35	59 T	71	66	54	52	47	48	57	43	31	36	52	59	63	63	65		50.5	2.7
270*	2500	46	42	37	29	37	48	42	40	49	45	54 S	73	47 T	38	59 T	56 S	56	50	46	45	49	53	47	35	36	51	59	63	62	60		48.5	2.2
271	6595	45	38	37	30	37	46	41	40	50	44	50	73 T	45 T	36	57 T	55	52	47	43	44	50	52	48	32	36	52	54	59	61	56		47.0	
325	1109	42	36	35	27	37	45	42	39	45	47	58	66	49	34	52	58	48	47	54	55	46	48	45	30	30	46	52	58	59	55		46.2	2.6
*	4735	36	30	28	22	32	M	44	36	39	44	51	~~M~~	M	M	~~63~~ S	55	51	42	54	~~M~~	~~54~~	44	42	27	28	M	M	56	54	47		42.6	
*	7810	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	~~M~~	~~45~~	55	41 S	47	~~58~~ S	36 T	32	41	43	41 S	44	49			
370	7627	38	37	34	26	36	40	41	40	41	40	45	63	48	45	47	50	41	47	56 S	50	42	49	48	32	35	46	49	50	50	58		44.1	2.5
374	0081	38	36	40	30	32	38	44	42	39	42	47	59	50	41	43	52	42	48	52	58	47	49	56	35 T	33	49	54	48	55	57		45.2	1.6
463*	2715	35	37	28	23	31	45	44	37	40	42	56	67	46 T	36	48	56	45	47	54	54	47	45	44	34	29	44	50	58	54	51		44.2	
468	2723	36	38	27	20	34	M	M	35	40	42	56	M	M	M	M	55	42	45	~~45~~	M	~~M~~	46	40	26	29	43	M	M	54	54			
*	4560	34	38	27	22	34	45	46	36	38	42	55	64	47	~~25~~ T	47 T	57	~~57~~ S	48	58 S	55	47	47	44	29	30	43	50	58	52	51		44.2	
520	4903	40	41	34	19	33	44	48	39	43	45	56	65	47	47	47	62	47	42	57	57	55	45	63	32 T	27	45	52	58	52	49		46.4	
550	4959	37	37	41	23	36	42	50	38	39	42	49	51 S	48	48	45	57	47	45	57	58	48	47	66	35 T	32	43	45	45 S	43	53		44.9	3.6
*	1221	M	M	M	M	M	M	M	M	32	40	41	47	39	41	40	~~44~~ S	49	40	47	~~44~~ S	43	41	~~45~~ S	38	33	38	42	49	44	42 S			

FIGURE 2 – Example of a computer listing of daily maximum temperatures for quality control with corrections indicated.

This listing is for stations in the Thunder Bay District in Ontario for April 1968. In the first line the digits 01 to 31 are headings for the days of the month. Daily maximum temperatures for each station in the district are listed on a line under these headings. On the left of each line is the 3-digit number assigned to each station for geographical sorting and the 4-digit number which, combined with the district number, identifies the station. The mean maximum temperature and difference from normal are listed to the right. (refer also to para. 7.3.1.)

604 THUNDER BAY 68 04 MIN. ARRAY ✓

GEO	NO	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	MEAN	DIFF
140*	0325	12	05	14	04	-09	11	28	30	18	13	32	28	19	27	11	32	31	22	17	21	37	34	28	23	20	19	12	15	28	25		20.2	.8
215	9096	18	14	18	13	-09	08	28	33	23	21	27	38	24	27	14	22	37	28	18	32	38	37	32	23	22	19	16	20	35	28		23.5	.8
260	3930	23	26	27	16	00	13	28	34	26	26	28	28	28	32	23	32	39	32	21	28	39	40	35	28	26	26	22	23	36	29		27.1	2.7
270*	2500	23	28	22	14	05	14	31	35	28	29	31	28	30	29	25	37	35	28	21	36	39	38	34	28	26	26	23	23	31	28		27.5	2.2
271	6595	24	28	30	17	08	19	30	35	29	30	30	30	30	33	25	36	38	34	23	34	38	37	38	29	28	29	27	25	31	30		29.2	
* 325	1109	14	13	25	15	-06	~~-12~~ +12/ I	29	34	26	22	32	29	27	28	21	32	38	31	20	26	40	38	35	28	23	26	20	23	37	28		~~24.1~~ 25.5	(2.6)
* 335	4735	21	14	23	14	-07	~~11~~ 22/ T	M	~~22~~ M/	26	~~08~~ M/	30	M	M	M	M	~~19~~ M/	32	29	22/	M	M	30	36	25	22	22	M	~~11~~ 24 M/	30			
* 368	7810	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	~~11~~ 31/ M	35	34	34	37	27	22	28	30	30	34	30			
	370	7627	12	19	25	12	01	14	30	35	28	24	25	35	27	31	24	34	37	32	29	34	34	34	39	27	20	26	25	25	30	29		26.6	2.5
	374	0081	12	19	28	15	-02	18/	24	36	30	25	28	34	30	34	24	35	38	34	22	32	35	35	38	28	22	28	27	26	30	28		27.1	1.6
	463*	2715	14	09	15	-12	-18	17/ T	30	27	18	11	29	29	22	19	23	34	33	21	16	24	37	35	34	23	17	24	17	19	29	24		20.7	
*	468	2723	16	13	20	10/ S	-07	M	M	20	25	22	29	M	~~11~~ 22/	M	M	~~22~~ M/	39	28	20	~~11~~ 29/	M	~~29~~ M/	35	24	19	24	M	M	23	28			
	475	4560	17	09	18	12	-13	05	30	34	25	20	23	34	22	24	28	28	41	29	19	27	39	37	36	28	19	20	18	20	33	27		23.6	
	520	4903	18	15	23	12	-10	10	32	36	26	27	26	24	25	32	25	35	43	33	23	26	45	38	40	26	19	24	25	28	36	30		26.4	
	550	4959	24	22	28	15	03	32/ S	35	34	31	28/	28	42/	30	31	28	38	39	36	30	31	45	35	39	29	23	27	24	25	35	28		29.8	3.6
*	599	1221	M	M	M	M	M	M	M	M	~~11~~ 32/ M	~~32~~ /	30	37 S	34	31	34	36	35	35	33	35	36	35	34	33	28	29	32	31	37	33			

FIGURE 3 – Example of a computer listing of daily minimum temperatures for stations in the Thunder Bay District for April 1968, for quality control. The headings and arrangement of data in this listing are the same as that used in Fig. 2. Corrections to the data are indicated. (refer also to para. 7.4.1.)

604 THUNDER BAY								68 04			MAX. ARRAY																					MEAN	DIFF
GEO	NO	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31		
140	*0325	35	34	25	29	35	46	41	36	36	45	54	60	41	37	41	51	60	50	53	55	43	43	35	29	33	47	51	58	58	56		43.9	.8
215	9096	36	39	28	28	38	51	42	35	41	45	69 T	58	46	32	52	63	59	49	55	48	45	55	33 T	28	36	48	53	57	57	60		46.2	.8
260	3930	42	43	38	31	43	50	44	40	48	47	67	71 S	50 T	35	59 T	71	66	54	52	47	48	57	43	31	36	52	59	63	63	65		50.5	2.7
270	*2500	46	42	37	29	37	48	42	40	49	45	54 S	73	47 T	38	59 T	56 S	56	50	46	45	49	53	47	35	36	51	59	63	62	60		48.5	2.2
271	6595	45	38	37	30	37	46	41	40	50	44	50	73 T	45 T	36	57 T	55	52	47	43	44	50	52	48	32	36	52	54	59	61	56		47.0	
325	1109	42	36	35	27	37	45	42	39	45	47	58	66	49	34	52	58	48	47	54	55	46	48	45	30	30	46	52	58	59	55		46.2	2.6
*335	4735	36	30	28	22	32	M	44	36	39	44	51	~~M~~ 63	M	M	~~63~~ M S	55	51	42	54	~~M~~ 54	~~54~~ M	44	42	27	28	M	M	56	54	47		42.6	
*368	7810	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	~~M~~ 45	~~45~~ M	55	41 S	47	~~58~~ 48 S	36 T	32	41	43	41 S	44	49			
370	7627	38	37	34	26	36	40	41	40	41	40	45	63	48	45	47	50	41	47	56 S	50	42	49	48	32	35	46	49	50	50	58		44.1	2.5
374	0081	38	36	40	30	32	38	44	42	39	42	47	59	50	41	43	52	42	48	52 S	58	47	49	56	35 T	33	49	54	48	55	57		45.2	1.6
463	*2715	35	37	28	23	31	45	44	37	40	42	56	67	46 T	36	48	56	45	47	54	54	47	45	44	34	29	44	50	58	54	51		44.2	
468	2723	36	38	27	20	34	M	M	35	40	42	56	M	M	M	M	55	42	45	~~45~~ M	M	~~M~~ 43.	46	40	26	29	43	M	M	54	54			
*475	4560	34	38	27	22	34	45	46	36	38	42	55	64	47	~~25~~ 45 T	47 T	57	~~57~~ 47 S	48	58 S	55	47	47	44	29	30	43	50	58	52	51		44.2 5	
520	4903	40	41	34	19	33	44	48	39	43	45	56	65	47	47	47	62	47	42	57	57	55	45	63	32	27	45	52	58	52	49		46.4	
550	4959	37	37	41	23	36	42	50	38	39	42	49	51 S	48	48	45	57 M	47	45	57	58 M	48	47	66 M	35 T	32	43	45	45 S	43	53		44.9	3.6
*599	1221	M	M	M	M	M	M	M	M	32	40	41	47	39	41	40	~~44~~ M S	49	40	47	~~44~~ M S	43	41	~~45~~ M S	38	33	38	42	49	44	42 S			

FIGURE 2 – Example of a computer listing of daily maximum temperatures for quality control with corrections indicated.

This listing is for stations in the Thunder Bay District in Ontario for April 1968. In the first line the digits 01 to 31 are headings for the days of the month. Daily maximum temperatures for each station in the district are listed on a line under these headings. On the left of each line is the 3-digit number assigned to each station for geographical sorting and the 4-digit number which, combined with the district number, identifies the station. The mean maximum temperature and difference from normal are listed to the right. (refer also to para. 7.3.1.)

604 THUNDER BAY 68 04 MIN. ARRAY ✓

GEO	NO	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	MEAN	DIFF
140*	0325	12	05	14	04	-09	11	28	30	18	13	32	28	19	27	11	32	31	22	17	21	37	34	28	23	20	19	12	15	28	25		20.2	.8
215	9096	18	14	18	13	-09	08	28	33	23	21	27	38	24	27	14	22	37	28	18	32	38	37	32	23	22	19	16	20	35	28		23.5	.8
260	3930	23	26	27	16	00	13	28	34	26	26	28	28	28	32	23	32	39	32	21	28	39	40	35	28	26	26	22	23	36	29		27.1	2.7
270*	2500	23	28	22	14	05	14	31	35	28	29	31	28	30	29	25	37	35	28	21	36	39	38	34	28	26	26	23	23	31	28		27.5	2.2
271	6595	24	28	30	17	08	19	30	35	29	30	30	30	30	33	25	36	38	34	23	34	38	37	38	29	28	29	27	25	31	30		29.2	
*	325	1109	14	13	25	15	-06	~~-12~~ +12	29	34	26	22	32	29	27	28	21	32	38	31	20	26	40	38	35	28	23	26	20	23	37	28		~~25.5~~ 24.1	2.6
*	335	4735	21	14	23	14	-07	~~22~~ T	M	~~22~~	26	~~08~~	30	M	M	M	M	~~19~~	32	29	22	M	M	30	36	25	22	22	M	~~24~~	24	30		
*	368	7810	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	~~31~~	31	35	34	34	37	27	22	28	30	30	34	30		
	370	7627	12	19	25	12	01	14	30	35	28	24	25	35	27	31	24	34	37	32	29	34	34	34	39	27	20	26	25	25	30	29		26.6	2.5
	374	0081	12	19	28	15	-02	18	24	36	30	25	28	34	30	34	24	35	38	34	22	32	35	35	38	28	22	28	27	26	30	28		27.1	1.6
	463*	2715	14	09	15	-12	-18	17	30	27	18	11	29	29	22	19	23	34	33	21	16	24	37	35	34	23	17	24	17	19	29	24		20.7
*	468	2723	16	13	20	10	-07	M	M	20	25	22	29	M	~~22~~	M	M	~~M~~	39	28	20	~~29~~	M	~~M~~	35	24	19	24	M	M	23	28			
	475	4560	17	09	18	12	-13	05	30	34	25	20	23	34	22	24	28	28	41	29	19	27	39	37	36	28	19	20	18	20	33	27		23.6	
	520	4903	18	15	23	12	-10	10	32	36	26	27	26	24	25	32	25	35	43	33	23	26	45	38	40	26	19	24	25	28	36	30		26.4	
	550	4959	24	22	28	15	03	32	35	34	31	28	28	42	30	31	28	38	39	36	30	31	45	35	39	29	23	27	24	25	35	28		29.8	3.6
*	599	1221	M	M	M	M	M	M	M	M	~~M~~	~~32~~	30	37	34	31	34	36	35	35	33	35	36	35	34	33	28	29	32	31	37	33			

FIGURE 3 – Example of a computer listing of daily minimum temperatures for stations in the Thunder Bay District for April 1968, for quality control. The headings and arrangement of data in this listing are the same as that used in Fig. 2. Corrections to the data are indicated. (refer also to para. 7.4.1.)

604 THUNDER BAY 68 04 PREC. ARRAY ✓

GEO	NO	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	DEP	TOTAL	DIFF
140*	0325	T		28	04			09	14	T					83	03	T				T	39	01	139	73	09				T			0	4.02	2.21
215	9096	T		30				36	23	05				~~40~~	S 40		02				01	03	T	110	05				02				0	2.57	.76 →
260	3930			68			35	33	06	30				11	11		T	T			80	07		49	15				01	03			0	3.49	1.56
270*	2500	T		21	18			06	15	07	03	T	T	T	46			07			64	13		61	20	T			T	10			0	2.91	.94
271	6595			40	04		04	19	02	08				35	18		04	02			77	04		75	14	03		~~10~~	~~02~~ ✓10	02			0	3.21 →	
325	1109			45	05			24	06	26	02			10	50		10	19			28	21		104	10	05			T	16			0	3.81	2.20
330	6549			60	05			46	23	37				01	35	10	T	09			12	01		25	45	05			T	15			0	3.29	1.50
335	4735		T	30	10			45	12	25				-	40	11	T	09			15	-		60	15	05				37			0	3.14	1.83
368	7810	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M		32	11		60	43	03	T			14			T		
370	7627			45	T			T	RE 10/					37	45		55	T			45	48	T	RE 70/	T	T				15			0	~~2.90~~ 3.70E	(.60) →
374	0081			25	20			15	10					05	126		20	36			50	31		81	10	07				18			0	4.54	2.14
463*	2715	01		31	32			17	17	08		T		T	74	T		T	18		T	51	T	19	37	15	01			13			0	3.34	
465	2755	T	T	62	10			*20	15	12	T				89	T	22	08				54		42	20	10				18			0	3.82	
468	2723	T	T	94	02			R/J C	38	21	T				~~M~~ RS	~~M~~ RS	109	10				45		20	42	01				22			0	4.04 →	
475	4560	T		85	12			23	14	16				35	50	02	12	10	02		10	27	T	44	26	06				31			0	4.05	
520	4903			20	25			49	29	-				65	50	.	-	27			-	27		37	25	15				22			0	3.91	
550	4959			26	24			59	80					21	91		10	32			29	13		59	22	05	06			18			0	4.95	
599	1221	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M		0		

FIGURE 4 – Example of a computer listing of daily precipitation for stations in the Thunder Bay District for April 1968 for quality control.

The headings and arrangement of data are the same as that used in Fig. 2. Corrections are included with the type of precipitation in the correction indicated by the use of R for rain, S for snow and E for an estimated value. (refer also to para. 7.6.5.)

604		THUNDER BAY					68	04																	
140	0325	ARMSTRONG A																							
	1065		43.9	20.2	32.1	.8	60	12	- 9	05	28	.57	5	37.8	9	4.02	11	2.21	.39	21	13.7	23	1.39	23	0
215	9096	UPSALA																							
	1587		46.2	23.5	34.9	.8	69	11	- 9	05	24	.52	6	20.5	6	2.57	11	.76	.36	07	11.0	23	1.10	23	0
260	3930	KAKABEKA FAL																							
	912		50.5	27.1	38.8	2.7	71	12	0	05	24	2.02	9	14.7	6	3.49	13	1.56	.80	20	6.8	03	.80	20	0
270	2500	FORT WILLIAM																							
	644		48.5	27.5	38.0	2.2	73	12	5	05	23	2.23	10	7.8	6	2.91	13	.94	.64	20	2.7	03	.64	20	0
271	6595	PORT ARTHUR																							
	602		47.0	29.2	38.1		73	12	8	05	21	2.45	12	7.6	6	3.21	17		.77	20	4.0	03	.77	20	0
325	1109	CAMERON FALL																							
	750		46.2	~~24.7~~ 25.5	35.~~8~~ 9	2.6	66	12	~~-12~~ -6	~~06~~ 05	24	2.76	12	10.5	6	3.81	16	2.20	.84	23	4.5	03	1.04	23	0
330	6549	PINE PORTAGE																							
	765	30 30										1.34	9	19.5	8	3.29	15	1.50	.46	07	6.0	03	.60	03	0
335	4735	MACDIARMID																							
	1025	7 10	42.6						-7	05		1.64	7	15.0	8	3.14	13	1.83	.45	07	4.0	14	.60	23	0
368	7810	SLATE ISLAND																							
	617	18 18 18																					.60	23	T
370	7627	SCHREIBER																							
	990		44.1	26.6	35.4	2.5	63	12	1	05	22	~~2.45~~ 3.25E	~~6~~ 8	4.5	1	~~2.90~~ 3.70E	~~7~~ 9	.60	~~.55~~	~~16~~	4.5	03	~~.55~~	~~16~~	0
374	0081	AGUASABON																							
	625		45.2	27.1	36.2	1.6	59	12	- 2	05	21	3.94	11	6.0	4	4.54	14	2.14	1.26	14	2.5	03	1.26	14	0
463	2715	GERALDTON																							
	1085		44.2	20.7	32.5		67	12	-18	05	25	1.45	8	16.7	9	3.34	14		.51	21	4.1	14	.74	14	0
465	2755	GERALDTON ON																							
	1110	30 30										2.06	8	17.6	8	3.82	13		.54	21	6.2	03	.89	14	0
468	2723	GERALDTON FO																							
	1110	10 10 2' 2							- 7	05		2.12	—	19.2	—	4.04	—		.45	21	9.4	03	.94	05	0
475	4560	LONGLAC P AN																							
	1124		44.~~2~~ 5	23.6	~~33.9~~ 34.1		64	12	-13	05	23	2.18	12	18.7	8	4.05	17		.32	14	8.5	03	.85	03	0
520	4903	MANITOUWADGE																							
	1090		46.4	26.4	36.4		65	12	-10	05	22	2.91	8	10.0	5	3.91	12		.65	13	2.5	04	.65	13	0
550	4959	MARATHON																							
	602		44.9	29.8	37.4	3.6	66	23	3	05	20	4.18	12	7.7	5	4.95	15		.86	14	2.6	03	.91	14	0
599	1221	CARIBOU ISLA																							
	612	8 9 30					49	17						.0	0						.0				0

FIGURE 5 – Example of a computer listing of summarized data for stations in the Thunder Bay District for April 1968 for quality control.

Data for each station are on 2 lines. The first line contains the identifying numbers and name. On the second line will be found the station elevation, the number of missing daily values of maximum and minimum temperature, and precipitation (see Slate Island), if any. These are followed by monthly mean maximum, mean minimum, and mean temperature and difference from normal. The highest and lowest temperature for the month with their date of occurrence and number of days with frost are given. The next data included are the totals of rain, snow and precipitation with the number of days with each, the difference from normal precipitation, the greatest daily amount of rain, snow and precipitation with the date of occurrence. The final entry on the line is the depth of snow cover on the last day of the month. (refer also to para. 7.6.5.)

```
25346 6808  SANDSPIT A            BC     25 FT    P  BOOK NO. 1

0119 --- 150           0199 053   C 00 10199 065 068 065    1 1   1 1 CU   030                                  YYY
0322 --- 150           0193 055 WNW22 10193 061 052 060    1 1   1 1 CI   300                                  YYY
0400 --- 150           0196 047  NW16 10193 059 052 065    0 0   0 0   0  ---                                  YYYYY
1115 080 100   R-      0178 049   C 00 10179 062 054 062 1010  1010 AS   080                                  YYYYY
1414 --- 150           0104 055  SE07 10104 060 057 082 1010   1 1 CU   010 66 SC  030  7 33 AC  080 10       YYY
1418 010 040   RW- F   0104 056 WSW17 10104 059 056 090 1010   6 6 SF   010 44 NS  050 10                     YYY
1712 120 150   RW-     0001 053  SE20 10001 056 055 089    9 9   3 3 CU   010 22 SC  025  5 55 AS  120 10       YYY
1811 070 150           0013 054  SE20 10013 060 056 081    8 9   1 1 CF   005 22 CU+ 015  3 33 AC  070  6 N  AS  160YYY
1816 --- 150           0041 054 ESE22 10041 061 057 079    4 6   1 1 CF   015 22 CB  020  3 13 CI  300          YYYYY
 #
2108 030 080   RW-     0086 051 SSE03 10086 055 059 087 1010   1 1 CF   010 66 SC  030  7 33 AS  080 10       YYY
2112 --- 150           0072 056  SE18 10072 059 057 088 1010   2 2 CF   010 77 SC  040  9 11 AS  080 10       YYY
2306 --- 150           0082 046   C 00 10082 050 048 086    6 8   1 1 CF   007 56 SC  045  7 01 CI  300  8       YYY
2404 050 150           0098 050 WSW06 10098 052 050 094 1010   4 4 SC   013 66 SC  050 10                     YYY
2406 070 15-           0099 049 WSW04 10099 051 050 094  710   1 1 CF   002 11 CU  010  2 44 AC  060  6 02 CI 290YYYYY
2906 030 030   L- F    0217 052 SSE08 10217 055 054 091 1010   5 5 ST   010 55 SC  030                        YYY
 #
2907 030 080   L-      0220 053  SE08 10220 055 054 093 1010   5 5 SF   025 55 SC  030                        YYY
 #
3008 030 150           0254 051 SSE04 10254 057 054 081    8 8   3 3 SC   050 44 AC  090  7 11 CI  300  8       YYYYY

  17        CARD(S)
```

FIGURE 6 – Computer listing of hourly weather data from #1 punched card with suspect data underlined and corrections indicated.

The listing is a print out of col. 10-80 of the #1 punched cards with numerical punching for weather, obstructions to vision, wind direction and cloud type interpreted.

The three columns of "Y" on the right are used to indicate, when circled, the reason for an error as follows:

First Y – indicates data acceptable

Second Y – indicates an error in punching

Third Y – indicates an error in the original observation

3. QUALITY CONTROL STEPS 2 TO 6 – ACTION BY TECHNICAL STAFF

Technical staff engaged in the quality control programme for surface data are technicians who have taken a basic course in Meteorology and observing procedures and have a minimum of 5 years experience in surface weather observations. Preferably this experience should include assignment to observing duties in various climatic regimes, while experience in a Weather Office or in supervisory duties proves useful. While it may be impossible to select individual staff members with a wide variety of experience, the selection of staff with observing experience in various parts of Canada will provide an alternate way of broadening the outlook of individual members by an interchange of knowledge gained through experience.

Step 2: The justification of suspect data

Unless the technical staff is completely familiar with computers and the programmes written for the quality control there is always a tendency to view the data selected by computers for scanning as real errors. To guard against this they are instructed to first try to justify the data as correct even when they do not meet the conditions written in the programme. In other words, they are encouraged to try to prove that the original observer was correct. This is particularly important in dealing with climatological data of extremes of temperature and totals of precipitation as reported by observers at ordinary climatological stations. In making such decisions the technical staff have not only the computer listings, but have on hand the complete records for the month from the station in question, and similar records for other nearby stations are available. Other useful aids are a knowledge, or availability of a card index file, of past performance of particular stations and observers, and daily weather maps.

It is particularly important that any computer produced listing of suspect data for use by the technicians be on a form designed to facilitate their work. For example, in the quality control of the hourly weather observations as transferred to the #1 card, three methods have been used. In the beginning, a card containing suspect data in any field was rejected by the unit processing equipment in use at that time. On receipt of this card the technician was required to examine each punch in the card as there was no indication of the reason for the card rejection. The work of the technician became simpler later when marks were put on the edge of the card to indicate general areas, i.e. cloud field or psychrometric data, which were suspect. In the third phase the output of the computer consists of a listing of the observations in a form very similar to that used in recording the same data on the document and the element or elements suspect in the observations are underlined (fig 6). By creating this listing designed to aid the technicians the amount of data they can check has been almost doubled and eye strain has been reduced.

Step 3: Correction of Data

If the data control technician decides that the data cannot be justified, he is next directed to try to determine some logical way in which the observer made a mistake. In this context observer mistakes includes transposition of digits, misplaced decimals, errors in calculation (sea level pressure, psychrometric data, etc.) reading thermometers in error by 5, 10, 15° etc. too high or too low, resetting maximum and minimum thermometers between observations or not resetting properly at time of observation, reading the wrong end of the index in the minimum thermometer, omitting minus signs with below zero temperatures, recording data on the wrong day, or in the wrong space, or not in the exact form required for transfer to punched cards etc. The greater portion of suspect data selected by the computer in the climatological quality control system are simple mistakes of this nature. Once the quality control technician has determined the likely mistake a correction can be made which retrieves the original observed value.

Step 4: Rejection of Data

Only after decisions have been made that suspect values cannot be justified and are not the result of a mistake on the part of the observer in the original observation and in its recording, is action taken to remove the data from further processing. This is accomplished simply by correcting the punched card or any other machine sensible form in which the data may appear to indicate the values as missing or substitute estimated values. However, it should be noted that on the original record as completed by the observer the data which are rejected at this step should not be obliterated. By simply drawing a line through the original entry it will not be made illegible. It is possible that sometime in the future the observation may be examined again and, on the basis of further information, a decision be made that the data were correct as originally recorded.

Step 5: Estimation for Rejected or Missing Data

Much of the work in Climatology involves the statistical analysis of time series of data for a point. Such analysis may be biased or of little value if there are gaps in the series due to missing observations or faulty data. During the past few years there has been increasing awareness of this problem and two methods are currently in use to overcome them.

In computing the standard normals based on the 30 years 1921-50 missing monthly values for temperature and precipitation were estimated when not more than 5 of these for a particular month were missing. Average monthly values for the data available were computed and these values were used with the departure from normal for the climatological district to estimate a value for each particular month when data were not available. Similar procedures were used in computing the standard normals based on the years 1931-60 and are still used for estimating climatological data when all observations for a station for a month are missing or faulty. It will be noted that this type of estimating is done only by the supervisory staff in quality control during the final check of summaries prepared from the data, or later by supervisory staff at the time normal values are being prepared.

The second method of estimating for missing values has been adopted more recently and involves estimation of hourly or daily data rather than monthly mean values. It was the first used to estimate occasional missing hourly values of wind, pressure, temperature, psychrometric data etc. in otherwise complete records of hourly weather observations. In climatological data, particularly from ordinary climatological and precipitation stations, where the total of precipitation for a month is considered incomplete when it is judged that an observer missed the measurement of one or more occurrences of precipitation which gives an error of 0.25 in. or 5% of the monthly total, which ever is greater, it has been judged preferable to estimate for these one or two missing values rather than estimate for the complete month at a later date using the method described above. It has been found also that when means of the daily extremes of temperature for a month are calculated whenever not more than 5 values are missing, the resultant mean may be considerably in error if the missing temperatures are grouped in extremely warm or cold periods during the month. The use of estimated daily values during these unusual periods will improve the mean values. Pressure has also been exerted to have the records as complete as possible in machine sensible form. When these data are used later for project purposes, the users, especially those outside the Meteorological Branch, may not be accustomed to using data with missing values and encounter considerable difficulty in their programming and interpretation.

Since January 1, 1968 the estimated values for climatological data and any values derived from them are flagged in the punched cards and are listed in standard data publications followed by the letter "E".

Estimation of hourly and daily values is incorporated as part of the regular data control system with the estimated values checked by the supervisory staff. Up to this point the methods used for estimating these values are in the development stage. Briefly, most estimation is based on interpolation, both in time and in space, with a further adjustment for differences in sites and/or weather types.

Step 6: Action to Improve Data

The quality of meteorological data depends mainly on the original training and continuing supervision of the observers, the accuracy of the instruments in use, the representativeness of the observing site, etc. The quality control programme as carried out in the Climatology Division may be only a small part of the overall control of the data gathering system, but it serves a useful purpose in that for some type of data it is the only consistent checking carried out. The results of this scrutiny will not only correct observer errors made in the past but, more important, it may serve as a means of indicating when action is required to prevent similar errors occurring in the data in the future.

When the first hourly weather data were transferred to punched cards in the early 1950's the need for the quality control of the data, and further training of many observers, became evident. At the time the first quality control procedures were developed they included a method of indicating on each punched card containing an error the correction required and then returning the card through administrative channels to the original observer. This procedure of routing all observer errors in hourly weather data to the original observing station has resulted in a marked improvement of data throughout the years. The improvement includes not only a reduction in the number of errors corrected in the quality control programme, but also improvement in those features of the observation which are not checked in this programme. In connection with this latter point the official statement of requirements for accuracy in records of hourly weather observations for climatological purposes is as follows:

> "To meet the Climatology Division requirements for accuracy, records of data from the hourly weather observations which are transferred to machine sensible form (#1 punched card) must be as free from error as possible when the data are to be used in statistical analyses as a basis for climatological advice for decision making. In interpreting this requirement it should be noted:
>
> (a) that an error count of greater than 2% in the records of the observations from a station for a month as determined from the quality control programme at Meteorological Headquarters, is an indicator of a requirement for special supervisory action in improving the observing and recording of data at the station;
>
> (b) when the error count for a station reaches 4%, all data from such a station are suspect and the entire observational programme should be reviewed;
>
> (c) for purposes of interpretation the error count is the percentage of observations requiring one or more corrections".

The advantages gained in preventing errors before their occurrence rather than correcting them after they have occurred is obvious. While the system outlined above for implementing this procedure for the hourly weather observations has been satisfactory in most respects, similar systems have not been established for all types of data. There are two main difficulties to overcome in the effective implementation of this step. In the first place, this step is not required in the current operation of the quality control unit for the Climatology Division, and is often overlooked due to pressure of meeting immediate requirements. Secondly, since the training and supervision of observers required to eliminate current problems and prevent future occurrences of errors in the responsibility of others who are far removed in the organization from those responsible for the quality control, it is very difficult to organize and sustain the co-operation required.

4. CONCLUSION

Based on experience in the use of the quality control system for climatological purposes the main features of the system may be summarized as follows:

(a) The unit record equipment used at the beginning, and currently the computer facilities, provide efficient means of consistently scanning data, or arranging and listing data for scanning by technical staff, to select suspect values.

(b) The efficiency with which the computer selects suspect values depends on the specifications written for data control programmes. These programmes should be kept under constant review and modified as additional computer capabilities and new meteorological knowledge becomes available. The development of these programmes requires knowledge in depth of general meteorology, climatology, observing instruments and procedures and the computer system in use.

(c) By employing the computer only to select suspect data, or to list observed data and summarized data for scanning by technical staff where suitable programmes are not developed for scanning by computer, and using experienced technical staff to make decisions in all cases as to whether data should be accepted, corrected, or rejected and an estimated value used, the completeness of data may be retained.

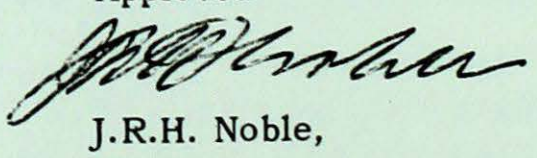
(d) In quality control of data from a continuing programme it is equally or more important to prevent errors in the future than to correct errors which have occurred in the past. This step in the quality control system may be overlooked entirely, or not carried out efficiently through lack of understanding and cooperation by observational programme managers.

(e) Users of the data and observational programme managers should be familiar with the limitations of the quality control system. From the illustrations used in sections 1 and 3, and from the specifications for computer programmes listed in the appendices, it will be evident that the efficiency of quality control system varies from fairly exact controls to relatively little control. As an example of the latter, in the hourly weather data there is particularly no check imposed on the observers' estimations of visibility, cloud heights, etc.

4.1 ACKNOWLEDGEMENTS

Many individuals in the Machine Processing and Operations Sections of the Climatology Division have contributed to the planning and development of the quality control system described in this circular. It must also be acknowledged that when machine methods were first used in the quality control procedures the system then in use at the National Weather Records Center, Environmental Science Services Administration, United States Department of Commerce, served as a model. In the intervening years there have been many fruitful interchanges of ideas on further development of the system with the staff at the National Weather Records Center.

Approved



J.R.H. Noble,
Director,
Meteorological Branch

5. REFERENCES

1961 Cudbird, B.S.V.

Modern Techniques in Canadian Climatic Data Processing.
Dept. of Transport, Meteorological Branch,
CIR-3558, CLI-28, Nov. 1, 1961. 21 p.

1968 Cudbird, B.S.V.

Modern Techniques in Canadian Climatic Data Processing.
Dept. of Transport, Meteorological Branch,
CLI 2-68, June 30, 1968. 17 p.

1964 Card Type #1, Card Format Documentation

Department of Transport, Meteorological Branch,
Climatology Division, June 10, 1964. 10 p.

1965 Card Type #4, Card Format Documentation

Department of Transport, Meteorological Branch,
Climatology Division, March 15, 1965. 12 p.

6. APPENDIX A

Specifications for computer programmes for selection of suspect data from the hourly weather observations (#1 punched card)

6.1. Specifications for the programme for selecting suspect data are listed for each element of the observation and then in relation to the 5 types of checks described in Section 2. For convenience reference will be made to columns of the # 1 punched card (Fig. 7) to which the data are transferred.

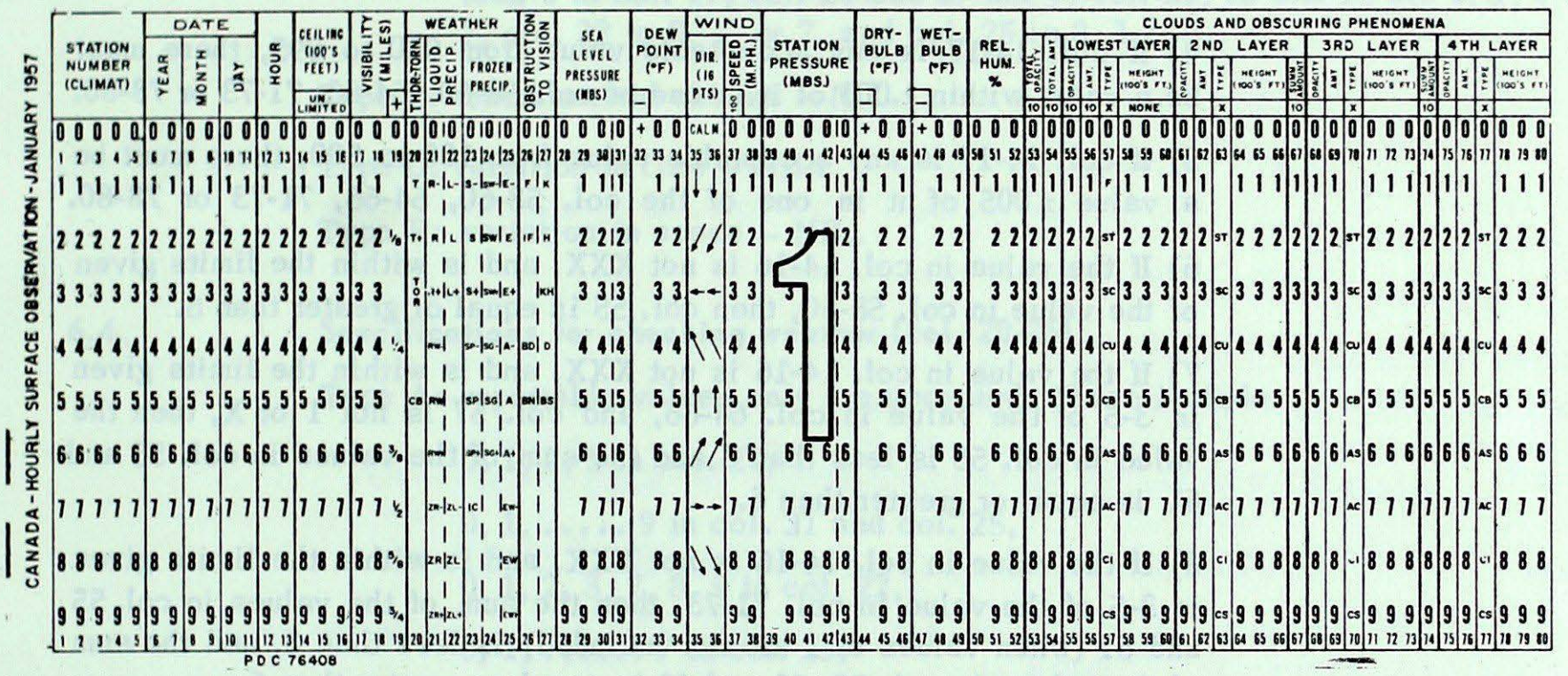


FIGURE 7 – Meteorological Service of Canada Punched card #1 for hourly weather data

In these specifications statements are made of conditions that must exist in a # 1 punched card or, if not met, the observation is listed with the element which did not meet the specification underlined as suspect data (Fig. 6).

6.1.1. # 1 punched cards are normally single punched with reliance placed on the computer scanning to pick out punching errors as well as other types of errors. The only exception to this is a spot check where col. 10-13, 20-27, 35-38 and 67 for a few stations are verified each month.

6.2. Specifications for checking ceiling (col. 14-16)

Type A: Reportable values only are acceptable. These include:

000, 001, 002,..... 050,

055, 060, 065,..... 100,

110, 120, 130,..... 500,

XXX, or there may be no entry when col. 17-80 also have no entry.

Type B: Internal consistency check:

1) When the value in col. 14-16 is XXX, then col. 53 is 0-5, or if col. 53 is 6-9 when col. 14-16 is XXX, then col. 57 is 1 or X, and col. 61, 68 and 75 are 0, 1, or blank.

2) When col. 14-17 is any value except XXX, then the punching in col. 53. is 6-10.

3) If col. 14-17 is any acceptable value from 000 to 049, there must be the same value in one of col. 58-60, 64-66, 71-73 or 78-80.

4) If col. 14-17 is any acceptable value from 050 to 095, there must be a value within ± 003 of it in one of col. 58-60. 64-66, 71-73 or 78-80.

5) If col. 14-17 is any acceptable value from 100 to 500, there must be a value ± 005 of it in one of the col. 58-60, 64-66, 71-73 or 78-80.

6) If the value in col. 14-16 is not XXX, and is within the limits given of the value in col. 58-60, then col. 55 is equal or greater than 6.

7) If the value in col. 14-16 is not XXX, and is within the limits given in 3-5 of the value in col. 64-66, and col. 57 is not 1 or X, then the value in col. 55 is less than 6 and the sum of the values in col. 55 and 61 is equal or greater than 6.

8) If the value in col. 14-16 is not XXX, and is within the limits given in 3-5 of the value in col. 71-73, then the sum of the values in col. 55 and 61 (when values of 1 are not included) is less than 6, and the sum of the values in col. 55, 61 and 68 is equal or greater than 6.

9) If the value in col. 14-16 is not XXX, and is within the limits given in 3-5 of the value in col. 78-80, then the sum of the values in col. 55, 61 and 75 (when values of 1 are not included) is less than 6, and the sum of the values in col. 55, 61 68 and 75 is equal or greater than 6.

Type C: mathematical checks – NIL.

Type D: variation in time – NIL.

Type E: variation in space – NIL.

6.3. Specifications for checking visibility (col. 17-19)

Type A; reportable values only acceptable. These include:

0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 1, 1 1/4, 1 1/2, 1 3/4, 2, 2 1/4, 2 1/2 3, 4, 5, ...15, 15+, 20, 25, 30 ....95, 99. When there is no entry in col. 17-19 there are also no entries in col. 14-16 and 20-80.

Type B: Internal consistency check:

1) When the value in col. 19 is 0, 2 or 4, then there must be a value other than 0 in one of col. 26-27, or a value of 2, 3, 5, 6, 8 or 9 in col. 21 or 25, or a value of 3 or 9 in col. 22, or a value of 3 or 6 in col. 23 or col. 24.

2) When the value in col. 19 is 6 or 7, there must be a value other than 0 in one of col. 26- 7, or a value of 2, 3, 5, 6, 8 or 9 in col. 21 or 25, or a value of 2 or 8 in col. 22, or a value of 2 or 4 in col. 23 or 24.

3) When col. 19 is 8 or 9, or col. 18 is 1, 2, 3, 4, 5 or 6, there is a value other than zero in one of col. 26-27, and a value other than 0 in col. 21 or 25, or a value of 1 or 7 in col. 22 or 23, or a value of 1 in col. 24.

4) When the value in col. 18 is 7 or greater, or there is a value other than 0 in col. 17, col. 26 and 27 are 0, col. 21, 23 and 25 are 0, 1, 4 or 7, col. 22 is 0, 1, or 7, and col. 25 is 0, 1 or 4.

Type C: mathematical checks – NIL.

Type D: variation in time – NIL.

Type E: variation in space – NIL.

6.4. Specifications for checking weather (col. 20-25)

Type A: reportable values only are accepted. These include:

0, 1, 2, 5 in col. 20,

0, 1, ... 9 in col. 21 and col. 25,

0, 1, 2, 3, 7, 8, 9 in col. 22,

0, 1, ... 7 in col. 23,

0, 1, ... 6 in col. 24.

Entries in col. 20-25 are acceptable only when there are entries in some or all of the col. 14-19 and 26-80.

Type B: Internal consistency check:

1) With values of 1, 2, 3, 4, 5, 6 in col. 21, or of 1, 2, 3 in col. 22, the entry in col. 44-46 must be for a value equal or greater than 290.

2) With values of 7, 8 or 9 in col. 22 the entry in col. 44-46 must be for a value in the range of 10-33°

3) With values of 7, 8 or 9 in col. 21. the entry in col. 44-46 must be for a value in the range of 20-30°.

4) With values of 1, 2 .....7 in col. 23, or 1, 2, ... 6 in col. 24, or 1, 2, 3, 7, 8, or 9 in col. 25, the entry in col. 44-46 must be for a value equal or less than 42°.

5) With values of 4, 5 or 6 in col. 25, the entry in col. 44-46 must be for a value equal or greater than 25°.

6) With values of 1 or 2 in col. 20, the entry in col. 44-46 must be for value equal or greater than 20°.

Type C: mathematical check – NIL.

Type D: variation in time – NIL.

Type E: variation in space – NIL.

6.4.1. Refer to 6.3 for internal consistency checks against visibility.

6.5. Specifications for checking obstructions to vision, (col. 26-27)

Type A: Reportable values only are accepted. These include:

0, 1, 2, 4, 5 in col. 26 and 0, 1, 2, 3, 4, 5 in col. 27

Type B: Internal consistency check:

1) With a value of 4 or 5 in col. 26 the entry in col. 44-46 must be for a value equal or greater than 40°.

2) With a value of 5 in col. 27 the entry in col. 44-46 must be for a value equal or less than 39°.

3) With a value of 4 or 5 in col. 26, or 5 in col. 27, the entry in col. 37-38 must be for a value equal or greater than 12 mph.

Type C: mathematical check – NIL.

Type D: variation in time – NIL.

Type E: variation in space –NIL.

6.5.1. Refer to 6.3 for internal consistency check against visibility.

6.6. Specifications for checking pressure (col. 28-31 and 39-43)

Type A: 1) Entries in col. 28-31 are to be for values in the range 960.0 to 1059.9 mb.

2) Entries in col. 39-43 are to be for values in the range 700.0 to 1059.9 mb.

Type B: Internal consistency check:

1) When a value is entered in either col. 28-31 or col. 39-43, the other must contain a value.

2) If the established elevation of the station is less than 50 feet, the value in col. 28-31 equals the value in col. 39-43.

3) If the established elevation of the station is equal or greater than 50 feet, the value in col. 28-31 must be greater than that in col. 39-43.

Type C: mathematical check – NIL.

Type D: Variation in time:

1) Differences in the values in successive observations for both col. 28-31 and 39-43 are equal or less than the limits in the following table

Interval in Hours Between Observations	Limit of Change (mb.)
1	3.9
2	6.9
3	9.9
4	12.9
5	14.9
6	16.9
Greater than 6	No limit

2) When the values in col. 39-43 are subtracted from those in col. 28-31 and differences taken of the results in successive observations, the differences, regardless of the time interval between observations, must be equal or less than the values in the following table:

Station Elevation (feet)	Limit (mb.)
0 – 999	1.9
1000 – 1999	2.9
greater than 1999	3.9

3) If col. 28-31 for the first observation of a month does not contain a value, then these columns do not contain a value for all other observations in the month. The same condition holds for col. 39-43.

Type E: variation in space – NIL.

6.7. Specifications for checking wind (col. 35-38)

Type A: Acceptable combination of entries for col. 35-36 are: 00, 11, 12, 22, 32, 33, 34, 44, 54, 55, 56, 66, 76, 77, 78, 88, 18.

Acceptable entries for col. 37-38 are for all values less than 70 mph., including 00. Missing values are only acceptable for those stations reporting less than 24 hourlies per day.

Type B: Internal consistency check:

If either the entry in col. 35-36 is 00, or the entry in col. 37-38 is 00, then both are 00.

Type C: mathematical check – NIL.

Type D: variation in time – NIL.

Type E: variation in space – NIL.

6.8 Specifications for checking temperature (col. 44-46) and psychrometric data (col. 32-34 and 47-52)

Type A: Allowable ranges for entries are:

For col. 44-46 the range of values is from -59° to 105° and for col. 50-52 the range of values is from 10 to 100%

Type B: Internal consistency check:

1) If col. 44-46 are blank, then col. 32-34 and 47-52 are also blank.

2) If col. 44-46 contain a value less than -39° then col. 47-49 do not contain a value.

3) If col. 44-46 contain a value less than -35°F and col. 47-49 also contain a value, the absolute value of their difference is one or less.

4) If col. 44-46 contain a value equal or greater than -30°, then there are also values in col. 32-34, 47-49 and 50-52.

5) If col. 44-46 contains a value greater than -30°, then that value is equal or greater than the value in col. 47-49.

6) If col. 44-46 contains a value greater than 32°F and col. 50-52 a value of 100%, then these values in col. 44-46, 47-49 and 32-34 are equal.

7) If col. 47-49 do not contain a value, then col. 32-34 and 50-52 do not contain values.

8) If col. 47-49 contain an entry, then col. 32-34, 44-46 and 50-51 contain values.

Type C: Mathematical Check:

1) For all observations with data complete in col. 32-34 and 44-52 a table lookup is required to confirm that for each observation the values in col. 32-34 and 50-52 are within the limits of possible values computed originally from dry bulb and wet bulb depression in tenths of degrees according to the following:

Station elevation (feet)	Type of Psychrometer	Source of data for limits	
		Lower Limits	Upper Limits
0 – 999	ventilated	Form 2240-1	Form 2240-1
1000 and above	ventilated	Form 2240-2	Form 2240-3
All elevations	unventilated	Form 2240-4	Form 2240-6

Type D: Change in time:

Hourly changes in values are equal or less than the following:

For col. 32-34 the limit is 15°, for col. 44-46 the limit is 29°, and for col. 47-49 the limit is 19°.

Type E: Change in space – NIL.

6.9 Specifications for checking cloud data (col. 53-80)

Type A: Acceptable entries in col. 58-60, 64-66, 71-73, 78-80 are only for values from 000-500, XXX or blank.

Type B: Internal consistency check:

1) The sum of the entries in col. 56, 62, 69 and 76 is equal or greater than the value in col. 54.

2) When col. 53 and 54 are both 0, then col. 53-57 are 00000, col. 58-60 are XXX and col. 61-80 are blank.

3) When col. 61-80 are blank, the entries in col. 53 and 55 are equal, and also the entry in col. 54 must be the same as that in 56.

4) When col. 68-80 are blank the entry in col. 54 is the same as that in col. 67.

5) When col. 75-80 are blank the entry in col. 54 must be the same as that in col. 74.

6) When the entry in col. 54 does not equal any of the entries in col. 56, 67 or 74, then col. 75-80 are not blank.

7) When the entry in col. 57 is 1, and the value in both 55 and 56 is 10, then col. 58-60 contain a value in the range 000 to 020.

8) When the value in col. 57 is 1, and those in col. 55 and 56 are both less than 10 and equal, then the entry in col. 58-60 must be XXX.

9) No two of the following 4 groups of col. may have the same value: col. 58-60, 64-66, 71-73. 78-80.

Type C: Mathematical check – NIL.

Type D: Variation in time – NIL.

Type E: Variation in space – NIL.

6.9.1. Refer to 6.2 for checks of ceiling against opacity and layer heights.

7. APPENDIX B

Specifications for computer programmes for selection of suspect data and listing observed and summarized data for scanning from daily climatological data (#4 punched card)

7.1. Specifications for the programme for selecting suspect data are listed for each element under the 5 types of checks described in section 2. For convenience reference will be made at times to columns of the #4 punched card (Fig. 8) to which the data are transferred.

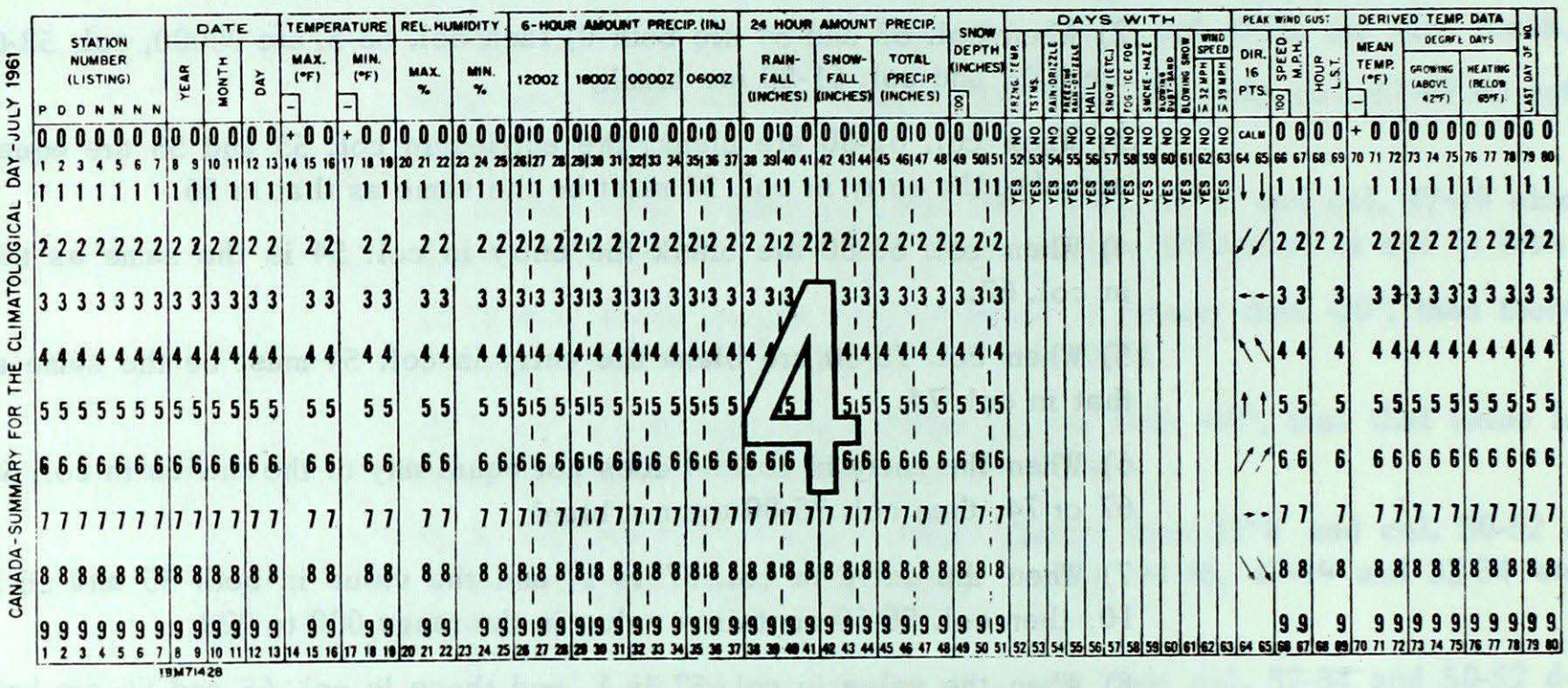


FIGURE 8 – Meteorological Service of Canada Punched card #4 for daily climatological data

7.1.1. For precipitation stations col. 1-13 and 38-48 are completed for all days, and col. 49-51 for the last day of the month only.

7.1.2. For ordinary climatological stations col. 1-19 and 38-48 are completed for all days, and col. 49-51 for the last day of the month, except col. 49-51 are completed for each day for agrometeorological and climatological reference stations.

7.1.3. For principal stations col. 1-69 are completed for all stations taking 24 hourly observations per day and equipped with an anemometer measuring gusts. When observations are taken for less than 24 hours per day col. 20-25 are not completed, and col. 64-69 are not completed if the station is not equipped with an anemometer recorder which records gusts.

7.2. Elimination of Punching Errors

7.1.2. All # 4 cards from principal stations are punched and verified.

7.2.2. In punching # 4 cards for precipitation and ordinary climatological stations an additional card containing the monthly sums of the daily values of maximum, minimum temperatures, daily rain, snow and precipitation as recorded on the original record is included with the daily cards for the month. Differences between these sums on the additional card and the corresponding sums obtained by adding the values punched in the cards are listed by the IBM 108 card proving machine. Zero balances are taken as indicators of correct punching, while differences other than zero are resolved by first checking the monthly sums entered on the original documents and, if these are correct, the punching of the daily values in the card is checked for elimination of the error.

7.3. Specifications for checking maximum temperature (col. 14-16)

Type A: When a daily maximum but no daily minimum temperature is available the daily maximum temperature is in the range 99 to -49.

Type B: The maximum temperature is greater than the minimum temperature for the same day and the difference between the maximum and minimum is equal or less than 50°.

Type C: Mathematical Check – NIL.

Type D: The difference between the maximum temperature on day N and that on day N-1 is equal or less than 20°.

Type E: The difference between the maximum temperatures reported for the same day at station A and station B is equal or less than 10° where station B follows station A in the geographical arrangement within a sub-district (see 7.3.1.)

7.3.1. As an aid for further scanning of the data by technical staff a listing is required for each month for each climatological district (Fig. 1) of the daily values of maximum temperature and their mean for each station (Fig. 2). The values for each station appear on a single line with the stations arranged in numerical order of the 3-figure geographical number. Changes between sub-districts are indicated by a change in the hundreds digit of the geographical number. On this listing the letter T is printed below values which do not meet the condition in the type D check, and S for data which do not meet the conditions in Type E check.

7.3.2. See para. 7.4.2. for futher requirement for listing of summarized data.

7.4. Specifications for checking minimum temperature (col. 17-19)

Type A: When daily minimum but no daily maximum temperature is available the daily minimum temperature is in the range 79 to -69°.

Type B: See para. 7.3. Type B

Type C: Mathematical Check – NIL.

Type D: The difference between the minimum temperature on day N and that on day N-1 is equal or less than 30°.

Type E: The difference between the minimum temperatures reported for the same day at stations A and B is equal or less than 15°, where station A follows station B in the Geographical arrangement within a subdistrict (see 7.3.1.)

7.4.1. For use in scanning of the data by technical staff a listing is required for minimum temperature data (Fig. 3) to the same specifications as that required for maximum temperature data in para. 7.3.1.

7.4.2. As an aid in scanning observed values and derived means and extremes a further listing is required. Stations are listed in order of the geographical numbers and the following data are included: The last 4 digits of the climatological listing number, station name, elevation, mean monthly maximum, minimum and mean temperatures, difference of the mean monthly temperatures for the month and dates of occurrence, number of days with frost and the number of missing maximum and minimum temperatures in the month (Fig. 5).

7.5. Specifications for checking maximum (col. 20-22) and minimum relative humidity (col. 23-25)

Type A: Acceptable entries for both maximum and minimum relative humidity are in the range 10-100%.

Type B: 1) The maximum relative humidity is greater than the minimum except they may be equal when both are 100%.

2) When either the maximum or minimum relative humidity is missing the other must also be missing.

3) When data are available for day N, there will also be data for day N + 1, and if they are missing on day N, they will also be missing on day N + 1.

Type C: Mathematical Check – NIL.

Type D: Variation in time – NIL.

Type E: Variation in space –NIL.

7.6. Specifications for checking precipitation (col. 26-48)

7.6.1. Six-hourly amounts (col. 26-37)

Type A: Acceptable entries for six-hourly amounts are for values in the range 0-5.00 in., or T, or blank.

Type B: 1) When the total precipitation for the day (col. 45-48) is zero, then the six-hourly amounts are zero or blank.

2) When the total precipitation is T, then at least one six-hourly amount is T and the others are zero, T, or blank.

3) When the total precipitation is missing, then one or more of the six-hourly values are missing.

4) When the total precipitation is a value other than zero or T, and not all the six-hourly values are blank, then the sum of the entries for six-hourly amounts equals the total precipitation.

Type C, D and E – NIL.

7.6.2. Daily rainfall (col. 38-41)

Type A: 1) Acceptable entries are values in the range 0 – 15.00 in., C, L, T or blank.

2) When the entry in col. 38 is zero, then col. 39-41 are zero.

3) When col. 38-40 are blank, then col. 41 is blank or contains an entry of T, L, or C.

4) When col. 38 and 39 are blank and col. 40 is not blank, then col. 40-41 contain values in the range 00 – 99 in.

5) When col. 38 is blank and col. 39 is not blank, then the value in col. 39-41 is in the range 1.00 – 9.99 in.

6) When col. 38 is not blank or zero then the value in col. 38-41 is in the range 10.00 – 15.00 in.

Type B: When the total rain is not an accumulated amount (C.L.) and the snowfall is zero, then the total rain equals the precipitation. See also para 7.6.4.

Type C: Mathematical Check – NIL.

Types D and E: refer to checks D and E in para. 7.6.4.

7.6.3. Daily snowfall (col. 42-44)

Type A: 1) Acceptable entries are values in the range 0 – 40.0 in., C. L, T or blank.

2) When Col. 42 is zero, then col. 39-41 are zero.

3) When col. 42-43 are blank, then col. 44 is any digit or T, L or C.

4) when col. 42 is blank and col. 43 is not blank, then col. 43-44 contain values in the range 1.0 – 9.9 in.

5) When col. 42 is not blank and col. 42 is not zero then col. 42-43 contain values in the range 10.0 to 40.0 in.

Types B, C,: internal consistency and mathematical check – NIL.

Types D and E: refer to checks D and E in para. 7.6.4.

7.6.4. Daily precipitation (col. 45-48)

Type A: 1) Acceptable entries are values in the range 0 – 15.00 in., C, T, L or blank.

2) When col. 45 is zero the values in col. 46-48 are also zero.

3) When col. 45-47 are blank, the entry in col. 48 must be T, C, L or blank.

4) When col. 45-46 are blank and col. 47 contains a value, the value in col. 47-48 must be in the range .01 – 99 in.

5) When col. 45 is blank and col. 46 contains a value, the value in col. 46-48 are in the range 1.00 – 9.99 in.

6) When col. 45 is not blank or zero the value in col. 45-48 are in the range 10.00 to 15.00 in.

Type B: 1) When the total precipitation is zero the total rain and total snow are both zero.

2) When the total precipitation is T, then either one or both of the totals of rain and snow are T.

3) When the total precipitation is missing then one or both of the totals of rain and snow are missing.

4) If the total precipitation is C or L then one or both of the totals for rain and snow must have the same value.

Type C: Mathematical Check – NIL.

Types D and E: Variations in time and space:

7.6.5. As an aid in scanning of observed daily values of precipitation and monthly totals of rain, snow and precipitation and extreme 24-hourly amounts of each, there are 2 requirements for listing data for each climatological region. The first listing (Fig. 4) is to include daily precipitation and monthly total of precipitation for each station and be similar in format and geographical method of arrangement of stations to that specified for daily maximum temperatures in para. 7.3.1. The second requirement is for inclusion with the summary listing specified in para. 7.4.2. the following information for each station: the monthly totals of rain, snow and precipitation and the number of "days with" for each; the difference from normal of the monthly precipitation; the 24-hourly extremes for the month of rain, snow and precipitation and the data of occurrence of each; the depth of snow cover on the last day of the month; the number of days with missing precipitation amount, and the number of days for which entries of C and L are used (Fig. 5).

7.7. Snow cover depth (col. 49-51)

Type A: 1) Acceptable entries are for values from 0 to 199 in and a trace.

2) When col. 49 and 50 are blank col. 51 has the entry for a trace.

3) When col. 49 is blank and col. 50 is not blank, then col. 50 contains a digit in the range 1-9 and col. 51 is blank.

4) When col. 49 is a digit in the range 1-9, or an X over-punch over any digit 0-9, then col. 50 is not blank and col. 51 is blank.

5) When col. 49 is 0 both col. 50 and 51 are zero.

Type B: Internal consistency check – NIL.

Type D: mathematical Check – NIL.

Type D: 1) The decrease in snow depth in successive days is not greater than 5 in.

2) The increase in snow depth in successive days does not exceed by more than 2 in. the sum of the depth of snow cover of the first day and the snowfall in the following 24 hours.

3) For any day with reported snow cover depth of zero, the depth on the previous day is 0 or T.

7.8. **"Days with" (col. 52-63)**

Type A: 1) Acceptable entries are 0, 1 or blank.

2) Col. 54 and 57 must be blank (data are no longer punched in these col. as the information can be derived from the entries in col. 38-48 by the computer).

3) For principal stations the entries in col. 53, 55, 56 and 58-61 are 0 or 1.

Type B: 1) When col. 53 is 1, then the maximum temperature (col. 14-16) is in the range 30-99.

2) When col. 55 is 1 and col. 38 is not 0, then the maximum temperature (col. 17-19) is equal or be less than 32°.

3) When col. 55 is 0 and col. 38 is not 0, then the maximum temperature (col. 14-16) is equal or greater than 32°.

4) When col. 60 is 1 the maximum temperature (col. 14-17) is equal or greater than 30°.

5) When col. 63 is 0 then col. 62 is 1.

6) When col. 63 is blank, col. 62 is also blank.

7) When col. 63 is 1, then col. 62 is also 1.

Types B, C, D, E, – NIL.

7.9. Peak wind (col. 64-69)

Type A: 1) Acceptable entries in col. 64-65 are standard direction to 16 points (see para. 6.7.) or blank.

2) Acceptable entries in col. 66-67 are values in the range 19-99 or blank.

3) Acceptable entries in col. 66-67 are values in the range 19-99 or blank.

Type B: 1) When col. 64 is blank then col. 65-69 are also blank.

2) When col. 64 is not blank then col. 65-69 are also not blank.

Types C, D, E – NIL.