I Leave You .... The Archive

- Jacquie Blackburn

When making a career change in the mid 50's, job security was not a serious consideration. Looking back from 1987, the year of my retirement, I feel the move to Climatological Data Processing was indeed a wise choice. Not only was the job 'secure', it gave me an opportunity to make a contribution to the Climatological Archive. In the future you will be able to attribute some 'good' or 'bad' feature of this archive to me. There are things I have experienced in the last 32 years which I shall always remember. I would like to share some of them with you.

The steady growth of the archive has convinced me that, regardless of the state of the economy, uneffected by union demands or government ceilings, production continues across the nation, 24 hours a day. But in spite of years of experience, Field Services/ Weather Services still has trouble with scheduling. With a certain common liquid, it's either a drenching or a drought. When the quantites are right, the timing is wrong. Nobody wants it during July or August, except farmers and gardeners, and they want it only between 11pm and 5am---thank you very much. Then there is that white fluffy stuff; they seem to produce copious amounts, during morning and evening rush hours, even though we have been told repeatly it's only a week-end seller and only in rural areas, which are hilly.

All the while I have been subjected to continous pressure, both sea level and station, now expressed in those new units, kilopascals. Mercury, it's had it's ups and downs, usually averaging out over a period of.....is it seven years or fifteen? Is it really getting colder? Weather observations and I have become good friends. In recent years it has been my job to look after them in their old age.

Transcribing of weather data to punched cards began in 1950, before I started, with five IBM Type 001 machines in use by October of that year. (I have a copy of the manual for this machine prepared for the U.S. Weather Bureau by IBM, dated 1947) This was a table top model requiring cards to be hand fed, one by one. It did have an eject key, which, when depressed, would toss the card, usually within reach of anyone with long arms. An experienced operator could punch 250 aviation weather records a day using this table top model. Production increased to 1200 cards/dy with the introduction of the IBM Type 016 Card Punch a few years later. This, our first record type, aptly named Type #1, contained such information as height of the ceiling, visibility, type of weather, temperature, pressures, wind, as well as cloud layer information. All airports reported these data each hour; 744 times in a 31 day month, 8760 times in a year, more in a leap year.

It soon became obvious that punching all stations would require a large increase in staff; person years was a consideration even then. Two years later, in November, 1952 a field punching program began which was to last until June 1976. An IBM Type 001 was installed in each of the selected stations and for the technicians at these sites one of the

'other related duties' was to key a card each hour, place same face down in a special shipping box and dispatch to headquarters monthly. In spite of the controversy over this duty the program persisted -- with a total of 73 stations involved, as many as 60 at one time. When the program terminated in 1976, 54 stations 'turned in their machines'. The stations which participated for the full period contributed 207,500 hourly weather observations; that is 207,500 wind speeds, 207,500 dry bulb temperatures, 207,500....., or 1200 pounds of cards. Almost from the beginning the key-punching machines, and to a lesser extent, the punchers, were hard to maintain with perhaps hundreds of memoranda exchanged between stations, regional offices and headquarters. In later years it was impossible to get parts for the machines, necessitating much cannabalizing of good parts from machines beyond repair to fix other machines. At least one station, Calgary, used the same machine from 1952-1976. One of these machines was given to Ernie Greckol, a part of Climat Data Processing from the beginning, on the occassion of his retirement in December, 1981.

All observations were and still are subjected to some form of quality control. For the early data, this involved using a sorter with pocket counters. In the mid 50's we acquired several Electronic Statistical Machines, IBM Type 101. These were really glorified sorters controlled by a wired panel board and had the added feature of 60 unit counters and 2 accumulating counters, which could be printed in fixed positions. The panels enhanced quality control, but did not reduce the effort to impose the checks. Operators wore out the carpeting moving gracefully back and forth as they emptied the pockets and refilled the hopper. Several passes using different boards and settings were required. Rejected cards, identified by the run on which they failed and the pocket they fell in, were passed to technicians for review and possible correction. Since a card was not subjected to further controls after it was rejected, it was necessary to fully validate the rejected card.

I can remember how comforted I was when my supervisor, explaining the procedure to be followed with the cards that fell into certain pockets, told me with calm assurance 'Tornadoes are not allowed in Canada'. Each time the wired control panel #6 was in place and the selector switch was set at #2, I watched pocket 0, the one to receive the offending Tornado hour should it ever occur, with baited breath. Thankfully pocket 0 remained empty.

With the use of modern day computers the quality controls have been greatly enhanced with logical and mathematical checks over space and time as well as against extensive stored information. Enhancements will continue to be made using new methods which will blossom under the Climate Data Management System (CDMS).

Card storage rapidly became a problem as more and different kinds of stations were added. In addition projects were established to key the historical records to give us a greater depth of record. The most significant of these was established in 1960 under the Agricultural Rehabilitation & Developement Act, to key the historical Daily Climatological data, to the #4 card format. The fields captured were

maximum temperature, minimum temperatures, rainfall, snowfall, total precipitation and last day of the month snow cover. This project was fondly referred to as ARDA.

One aviation weather station taking observations each hour, resulted in 87,600 #1 type cards or 40 card storage boxes, in just 10 years. A simple request to provide a wind frequency of speed classes by direction took many passes of cards through the statistical machines, taking many hours of machine and people time.

In 1966 we took delivery of a IBM 360 Machine (we had a IBM 1440 for a year prior to this) with no tape drives but with a generous compliment of disk drives and packs. Hourly data from 20 of the most frequently requested airports were loaded to these packs as 34 character mini-records. A similar format was available for the #4 Daily Climat card. Utility programs, the earlier versions of the current General Report Programs, used these formats as input. However, cards remained as the official archive. Tape drives were added a year later presenting an opportunity to 'do something' about the archive which had expanded to 75 million cards.

Card image, which had become almost a 'sacred cow', was to remain. Each station was checked for completeness, and to some extent, quality controlled during the taping process. Progress was slow with the number of errors detected greater than anticipated. More cards were being punched all the while, bringing the total to a whopping 100 million or 45,500 boxes of them. Finally by 1977 we had won. The official archive was now proudly sitting on tape, still as card images.

Weather is still 'being produced' Canada wide. However, there is only a finite number of sites equipped to record it. A stations is established as part of a Network, with each Network equipped to record selected weather elements.

350 stations report aviation weather, our first record type.
300 record the amount of bright sunshine each hour.
2500 log the highest and lowest temperature and the amount of rain or snow which falls each day. Many of these are operated by volunteers and their contribution is significant. There is probably such a station in your neighbourhood.

Yet other networks observe such things as radiation, soil temperature rainfall intenstities, evaporation and even the upper airs are probed. Beyond our shores there is coverage, with ships at sea regularly logging the conditions that they are exposed to. Many stations are part of several networks. They open .... they close. Historically reports have been received from 9000 different sites. Understandably the coverage is sparse in some areas but generally it is adequate to describe the Climate of Canada.

Probably the greatest change to the Canadian Climatological Archive that I have seen resulted from Canada's committement to 'GO METRIC' in 1977. This decision presented an opportunity to finally turn our back on the 'sacred cow', and to reduce the number of record formats.

1987.07.03

Each network had a different card format. Constrained by the 80 columns, complex coding structures had been introduced, involving multiple entries in each column. Even simple numeric fields had become difficult to read with hi-level languages. I can remember asking a researcher if he was having any difficulty using the archival data which he required. His reply was 'Things went well after I broke the code'. Not a good indication of the useability of the archive.

Although each Network observed many kinds of data, they did so at regular intervals, usually hourly or daily, which we summarized by month. Recognizing this, three basic formats were adopted and a unique element number was assigned to each type of weather data.

Each format consists of the Station identification date and element number, followed by the element values.

-The Daily Record of Hourly Values provides for all types of weather reported at hourly intervals. The indicative portion consists if station ID, year, month, day and an element number. The value is then repeated 24 times, once for each hour of the day,

-The Monthly Record of Daily Values has 31 repeats of the value. The day is not required in the indicative portion of the record, but is implied by the location within the record.

-Lastly, the Annual Record of Monthly values follows the same philosophy with 12 repeats of the value, once for each month.

These formats have become a vehicle for entering derived or non-standard data for processing through existing software. The element number implies the content and could be used to record such diverse information as, say, the number of Monarch butterflies sited each hour heading south. These counts could them be analyzed with other archival elements using existing General Report Programs.

Only a few of the records are still keyed at Headquarters. The aviation weather is collected from the operational teletype circuits which span the country. Yet other data are received on magnetic tape from the regional centres. Some networks have been automated with output to diskettes. Continuously, new data are received, quality controlled and added to the now 'friendly' archive.

At regular intervals, data are summarized for distribution to a large subscription list and to the Regional Centres. Microfiche is now a popular media for this type of output, with increased use each year. Canada, and all member countries of the World Meteorological Organization, prepare 30 year normals every 10 years. You are exposed to these on a daily basis, with to-days temperature expressed as being higher or lower for this date than is normal, or exceeding the record set back in -----.

The Canadian Climatological Archive is maintained to support all climate sensitive areas of the economy. Such a wealth of data would be of little use without the means to extract that portion which will help

users solve their particular problem.

There is an analogy between access to the archive and the Toronto telephone book. If a boss asks her secretary for a telephone number, and he hands her the book .... that's data. If he looks the number up for her .... that's information. SERVICE is when he dials the number for her.

Just as there is more than one way to get a telephone number there is more than one way to have access to the Climate Archive. Users can arrange for a tape or diskette copy of the data (the telephone book approach) to explore on their own computing facility. Authorized users of the Downsview computing facility have access to the data and the support software. We maintain a comprehensive library of computer programs to prepare reports reflecting user limiting criteria. There is often a need to know the frequency or persistance of combinations of weather elements. For example ---- durations of more than 6 hours when the temperture is less than -15.0 degs C and the wind speed is more than 30 km/hr, from the north.

The combinations are endless, the results, ---- information. Many programs are available to reduce voluminous basic data to meaningful reports.

Remembering the analogy to a good secretary, we will do it for you. Hundreds of requests are serviced annually, usually with an average queue time of ten days. A project number was and still is assigned to each request. One of the first projects was #00452, the 4th project assigned in 1952. This project provided frequency of closed conditions at Seven Islands, using October 1949 through September 1951, just two years of data. The breakdown of hours spent follows:

105 hours for punching
24 hours for checking
12 hours for sorting
24 hours for grouping & typing

I have a copy of the covering memoranda written on September 2, 1952, signed by Pat D MacTaggard-Cowan for Andrew Thomas. The wrap-up paragraph in this letter says in part, 'The attached report is for your information only, and it is not intended as an encouragement to forward requests for work of this kind. The punch card unit is extremely hard pressed at the present time putting the basic data on punched cards, and machine facilities are not yet available to enable us to handle any more projects of this kind.'

The first machine produced Psychrometric Tables, issued to the stations for implementation in January, 1960, were prepared using an IBM Type 514 Reproducing Punch, 077 Collator, 602 Calculator and 402 Tabulator under Project #01358. I wish I could demonstrate for you the techniques used to perform these calculations, with all input and output on cards, in all colours!

What changes do I predict for you? Certainly more automation at the

source, with networks sampling almost continuously. Satellite imagery will provide area coverage never before possible, as well as creating archiving problems ... never before possible. Introduction of new communications facilities and the use of modern data management and access techniques will make even to-days weather available to-day.

Looking back to pre '77 when the archive was card images on tape ... and even further back when it was actually cards, tons of them, I recognize that where we are to-day is NOT A BAD PLACE TO BE. Where you will be to-morrow must be a Data Managers heaven.

I wish all of you 'Happy Data Processing'. When you find an error in the archive, a bug in one of the programs or some other problem apparently left to annoy you, please remember ..... I DIDN'T MEAN TO DO IT !

1987.07.03