JOURNAL OF HYDROLOGY (N.Z.) Vol. 14, No. 1, 1975

DATA PROCESSING: A REVIEW OF THE ROLE OF THE MINISTRY OF WORKS AND DEVELOPMENT

R. P. Ibbitt*

ABSTRACT

When instrumentation is chosen for collecting data, decisions are made – often implicitly – about the purposes for which the data are to be collected and about the accuracy needed to fulfil these purposes. Frequently, unforeseen circumstances affect the type of instrument required or the objective(s) for which the data are being collected. Some of the consequences of the use of inappropriate instruments are examined.

A description is then given of the different types of data currently being processed by the Ministry of Works and Development and of how the type of data affects the processing. The relative merits of different forms of publication based on the data are discussed, and the conclusions lead to a description of the ways in which users can obtain the data.

INTRODUCTION

The publication of the Dunford report (Dunford, 1973) and the discussions that led up to it made people aware that the management of the nation's water resources could be roughly divided into two parts. Short-term and day-to-day activities have come to be referred to as operational activities, while the word 'planning' has been attached to longer-term activities.

While the preceding statements are broad generalizations and there are many 'grey' areas between planning and operational activities, they will be used as a background for putting into perspective the hydrological data-processing work carried out in New Zealand.

INSTRUMENTATION

Any discussion of data processing must begin with a consideration of the purpose for which a phenomenon is being measured since this controls the type of instrument to be used.

* Water and Soil Division, Ministry of Works and Development, Wellington.

In operational work, the value of a phenomenon at the present moment in time and its value in the immediately preceding few hours are the data that are usually required for making decisions. For example, when deciding whether or not to issue a flood warning the following data are ideally needed:

(a) the current water level;
(b) whether the level is rising, falling or static;
(c) the catchment's time of concentration, t_c hours;
(d) how much rainfall has occurred in the preceding t_c hours.

The main characteristic of these data, if they are to be useful, is that they must be measured and communicated to the decision maker in the minimum possible time. Implicit in this assertion is the fact that the decision maker is unlikely to be interested in what happened outside the immediate past, and that he will require data for one purpose only. Thus there is no point in making elaborate provision for permanently recording the data, especially if this delays communication of the data to the decision maker.

Planning work, on the other hand, instead of using a small amount of data immediately, uses large quantities of data over long lengths of time. Speedy communication of data to the planner is rarely needed. The amount of data required, however, is considerable and must be recorded permanently for future (possibly multiple) use. Further, the quantity of data usually outstrips the human ability to digest and manipulate it. Consequently, not only must the data be filed in an orderly manner, but also retrieval from the file must enable the user to carry out any desired calculations rapidly and easily. Thus, computers arrive on the scene. If – as is almost inevitable – a computer is to be used to file the data and permit subsequent retrieval and manipulation it makes good sense to use field recording instruments giving output which can be fed into the computer-based filing system with the minimum of human interference. The smaller the amount of manual work required to file the data, the more computer compatible it is said to be. In summary, therefore, instrumentation for planning purposes should make permanent records that are as computer compatible as possible.

PERMANENTLY RECORDING INSTRUMENTS USED IN NEW ZEALAND

Since the vast majority of hydrological data automatically and permanently recorded in New Zealand come from either chart recorders or punched-paper-tape recorders, only recordings from these two types of instrument will be discussed.

The older of the two forms of recording instrument are the chart recorders. In one form or another they have been in use for over 50 years, and have been developed to record a wide range of both time scales and phenomenon values. They come in many different physical forms, ranging from linear scales for both variables through to circular scales for both variables. With the exception of long strip charts, e.g. the Lambrecht, they all have one feature in common: they give an immediate visual record of what has been happening from the moment the chart was put on up to the moment the instrument was visited. With the strip charts, a sufficient amount of the past is normally displayed for operational purposes if the recorder is being used in this capacity. Because the immediate visual record from a chart recorder gives the type of information needed for operational work, they were frequently used for this purpose before more sophisticated techniques removed the delays caused by actually having to visit an instrument.

In spite of being superseded in the operational domain, chart recorders have remained popular as planning instruments because they produce a permanent record, they are simple, and they are reliable. Unfortunately, the records from chart recorders are not amenable to calculations, so they tend to get stockpiled, with selected records being withdrawn and manually processed as and when information on a particular event or type of event is required. This does not make efficient use of the data because almost invariably some form of interpretation is placed on the data; for example, daily totals may be extracted, all information at finer time scales being ignored, so that if someone wants hourly values the job of reading the charts has to be repeated. To improve efficiency the charts need to be processed in a non-interpretative manner, i.e. one that will allow the basic document to be reconstructed if necessary. To do this manually is impractical in many cases. Methods and machines have been developed to aid the non-interpretative processing of the data but the fact remains that a considerable amount of human effort has to be expended to make a chart computer compatible.

This last feature has brought about, over the last 10 years, the gradual replacement of many chart recorders with punched-tape recorders. Instead of giving a continuous trace of the behaviour of a phenomenon, punched-tape recorders sample the magnitude of the phenomenon at times predetermined by a set of rules. These recorders give no visual record, so they are of little or no use for operational purposes. The record from them is coded on to paper tape that is much more computer compatible than charts in spite of being unsuitable for direct reading by the computer.

Punched-tape recorders are more sophisticated instruments than chart recorders, and usually record data to a greater number of figures than the phenomenon can be measured. Herein lies the trap for those who equate recording-instrument precision to the accuracy of measurement. This situation has been brought about by the punched-tape recorder being a digital instrument, i.e. one that can only record whole numbers. The chart recorder, by contrast, has a continuous trace the thickness of which is often comparable with the tolerances that would be placed on the measurement being recorded. For instance, on water-level charts it is not uncommon to find that the thickness of the trace limits the precision of water-level measurement to ± 50 mm, a figure that is commensurate with the accuracy and meaning that can be placed on a water-level measurement on some of New Zealand's rivers. Punched-tape recorders would, however, record to at least ± 3 mm for the same site.

The warning here is that the use of sophisticated instruments to obtain computer-compatible data does not imply a high accuracy of data measurement, and that where possible the processing of the data should be cognisant of this fact.

PROCESSING OF DATA RECORDS

The first step in the processing of any data is to get them into a computer-compatible form.

The chief problem with punched paper tape is that it cannot be read at the speed necessary for efficient computer use. Consequently, it is necessary to convert it into a medium suitable for high-speed reading by the computer. This conversion can be automatically done using a machine that can be operated by staff with no hydrological knowledge.

Conversion of charts is far more difficult. To convert the trace into a computer-compatible form, X–Y coordinates are taken off using a machine designed for the purpose. The digitizing of the trace, however, requires manpower. The work tends to be tedious (although not as tedious as manually reading the charts) and inefficient with inexperienced staff, since far more X–Y coordinates have to be taken off to guarantee that significant parts of the trace are not missed than are really necessary for an adequate representation. Hydrologically experienced staff do the job much more efficiently.

The record on either the paper tape or on the chart is basically only an index of what happened at the site. To give the index physical significance requires it to be put through computer programs that adjust the index to agree with the field measurements taken at the site when the chart or paper tape was changed. If these field measurements are not taken properly then the data collected cannot be given an accurate physical interpretation. This process is referred to as translation.

The results of the translation are stored on the computer's magnetic discs, and evidence of the translation – usually in plot form – is sent to the appropriate Ministry of Works and Development district office for checking. When sufficient data have been accumulated they are merged with data already processed and resident on magnetic tape, to produce a new and updated magnetic tape.

With the tapes of river and lake water-level data, additional information in the form of rating curves is necessary if flows are to be computed. Rating data are prepared manually and merged with existing data in much the same way as new water-level data.

With all merging operations, existing data are taken from one magnetic tape and written to another magnetic tape, but with the addition of new information. Each time this process is repeated a new magnetic tape is produced. After an arbitrary number of new tapes have been written (four in the case of the Ministry of Works and Development) the first of the magnetic tapes is overwritten. This results in a limited number of magnetic tapes being used and provides several editions of the data, the latest being the most up-to-date.

This method of handling the processed data has proved to be reliable and easy to operate, and has been applied to storing various types of time-series data. The largest sets of data held by the Ministry of Works and Development are the river and lake level records. The number of these records held is sufficiently large for each Works district to have at least one set of four tapes on which to store the data. In three districts a further subdivision has been found necessary.

Besides the surface water-level data, two other types of data are being stored. The more advanced of these are the rainfall data. This file does not duplicate data available from the Meteorological Service but supplements it. The data stored are from automatic raingauges (event raingauges and a limited number of Lambrecht recorders) operated by the Ministry of Works and Development on its experimental and representative basins. The data are stored in a non-interpretative manner so that intensity data over variable time periods and cumulative volumes suitable for input into catchment models can be recovered.

The latest of the data types being filed is groundwater data. Three aspects of this type are being processed: (i) water-level data; (ii) water-quality data; (iii) aquifer properties and well characteristics, including well location and construction details. The groundwater data file is only just emerging from the feasibility stage, and its processing is still a long way behind that of the rainfall and surface water-level files.

WHAT DATA ARE AVAILABLE?

There are two aspects to this topic:

(i) at which locations are data available?

(ii) what are available at selected locations?

Groundwater Data

At the time of writing (October 1974), unedited groundwater levels are available for two stations – both in the Hutt Valley. Data for these two stations have been processed to find out what snags were likely to be encountered in the processing of this type of data. Instructions have now been drafted for the general preparation of groundwater data, and it is expected that a substantial amount of new data will be added to the file in 1975.

Rainfall Data

Data from 22 stations are available on the file in an unedited form. Most of the data came from Fischer and Porter event rain-gauges sited in experimental or representative basins although a limited amount of data from stations equipped with Lambrecht or Dines recorders have been processed. The relatively recent introduction of the Fischer and Porter event raingauge means that most of the records are of less than two years duration. The exceptions are two chart-recorder sites at the head of Lake Wanaka, for which data are available from 1965–1970 in one case and 1955–1970 in the other.

Surface-Water Data

Surface-water data are the most prolific, information being available from approximately 330 stations distributed throughout New Zealand. Data available for selected sites vary from the Lake Rotoiti record which runs from 1905 through to the present date, to new stations with only a few months of data. A significant proportion of the sites have over 10 years of records. Details of the records available are given in the *Filed Data List* produced by the Ministry of Works and Development.

The surface-water files are in a more advanced stage of processing than the other two types of file. Besides water-level data, most sites have rating data with which to compute flows, and comments that relate to the quality of the data. Much editing work has been carried out to correct errors that creep in during the recording and processing operations. The district offices are responsible for the hydrological quality of the data, and when they are satisfied that a given piece of data has been edited to the point where no further improvements can be made with the resources they have available they 'certify' the data. It must be emphasized that certification does not imply that the data are of high quality; it does imply that the data have been checked and found to be consistent with the quality of the recording instrument and the conditions prevailing at the station. Should any errors be detected they should be communicated immediately to the appropriate district office for correction. The fact that data are certified does not remove the responsibility of users to check that the data are of adequate quality for their purposes. To check this users should read the comments associated with the data and, if necessary, contact the local district office for more information.

The principle of certification that is applied to the surface-water data will in due course be applied to the other types of data file.

PUBLICATIONS AND ACCESS TO THE DATA

Historically, access to hydrological data has been obtained by going to the appropriate publication, usually some form of yearbook. The use of this type of approach is rapidly declining because:

(a) it is only feasible to publish a limited amount of data;
(b) to limit the amount of data published some form of interpretation has to take place, and this tends to constrain the scope of work;
(c) techniques for using data are becoming increasingly computer orientated, and printed data are incompatible with computers;
(d) the exponentially increasing amount of data is causing greater delays in publication;
(e) amendments to data have to be delayed until the next publication.

A recent survey of the New Zealand data yearbook, the *Hydrology Annual*, has revealed that:

(a) only a limited number of data are presented;
(b) the data receive much interpretation;
(c) the information in the *Annual* is not computer orientated;

(d) there are considerable delays in publication;
(e) amendments to the data usually pass unnoticed.

Besides these factors, it requires approximately 5 man-years of labour to produce each volume – and it is used by fewer than 10 percent of the recipients.

In view of the recent comments about one of the American yearbooks (Ruhe, 1973), the purpose of yearbooks has been reviewed in the light of modern developments. The basic failing with yearbooks is that they attempt, by offering a summary of the data, to serve two frequently conflicting purposes: (i) to provide a non-specific catalogue of the data; (ii) to provide specific data for design purposes. It is the growth of the amount of data, coupled with the increasingly wide range of specific data required, that has caused yearbooks to become unwieldy. They are being replaced by two different types of publication – a catalogue of the data, known as the *Filed Data List*, and design manuals. Only the former publication is the concern of the data-processing section, and Volume 6 has recently been issued.

The *Filed Data List* gives a list of the sites for which data are being processed and what has been processed to the time of publication. This list is based on the data available on the magnetic tape files, and since these are usually updated at intervals of at least a week, the publication rapidly dates itself. As new issues supersede old ones, the old issues should be destroyed. Consequently, the document has been deliberately produced in a flimsy form and has been circulated only to known users of the data. Up-to-the-minute amendments can be obtained by contacting the appropriate district office who will issue a revised listing directly from the computer.

Finding out what data are available is only one aspect of actually being able to use them. Until recently access to the data has been difficult, but with the introduction of computer terminals into district offices any bona fide request for data will be actioned in district office. Data can be provided in three basic forms: (i) as a computer printout; (ii) as a deck of punched cards; (iii) as a computer plot. Under normal terminal operations a printout can be obtained in a matter of hours. The other two forms of output have to be posted from Wellington, and this introduces several days' delay.

When asking for data the user must be specific about what he wants. He must state:

(a) the site(s) for which data are required;
(b) the period(s) of time over which data are required;

(c) the type of data required, e.g. mean daily water level, mean daily flows, instantaneous peak flows, or flow-duration figures – the district office should be consulted to get the most suitable type of output;
(d) the form of output, i.e. whether cards, printout or plots are required.

Note that in all cases the users will have been expected to determine that what they request has been processed. If it is not yet available on file, a request may be made for it to be processed – but this may take some time, depending upon the amount of work involved.

The system of contacting a district office is expected to work satisfactorily where requests do not place too heavy a burden on district staff. However, for users who do not have easy access to a district office, or who will have large and frequent requirements, a new development is taking place. The Ministry of Works and Development is willing, under certain conditions, to permit organizations with suitable facilities to acquire a subset of the data-handling programs available on its own computer, together with copies of data tapes which will be updated on request. Installation of the system is at present under way at the University of Canterbury, and the success of this operation will determine the feasibility of extending the system to other institutions.

It is expected that, with the distribution of both the data-handling system and the data tapes, hydrological design will improve and there will be a greater awareness of the processing problems and design limitations created by the choice of an unsuitable recording instrument in the field.

ACKNOWLEDGMENT

Permission to publish the paper was given by the Commissioner of Works.

REFERENCES

Dunford, E. G. 1973: *Water and Soil Conservation Research in New Zealand.* A Report to the Minister of Works and Development. Government Printer, Wellington.

Ministry of Works and Development (n.d.): *Filed Data List of the Water and Soil Division.* MWD, Wellington.

Ruhe, R. V. 1973: Hydrologic data – what to do with it. *EOS 54* (7): 718–719.