

Recent developments in the USA in the collection and pre-processing of hydrological data

ROBERT A. CLARK Office of Hydrology, NWS, NOAA, Silver Spring, USA

Abstract. With the implementation of AFOS (Automation of Field Operations and Services) the collection and processing of hydrological data will be altered considerably from those procedures used previously in the USA. In addition, new techniques in remote sensing and automatic data acquisition have provided our hydrological forecasting service with new tools for providing an improved service. Promising new data relay techniques such as the radio event reporting gauge, satellite data relay, and the meteor burst technique offer marked improvement in the relay of real time data. Recent innovations in the pre-processing of such data to ensure high quality are of equal concern.

Developpements récents aux Etats-Unis dans le domaine de la collecte et du traitement préliminaire des données hydrologiques

Résumé. Avec la réalisation du projet AFOS (Automatisation des travaux de terrains et de bureau) les procédés de collecte et de traitement des données hydrologiques tels qu'ils sont encore utilisés aux Etats-Unis seront modifiés considérablement. En plus, de nouvelles techniques dans les domaines de la télédétection et de l'acquisition automatique des données ont pourvu les services de prévisions hydrologiques de nouveaux outils qui permettent d'offrir de meilleurs services. Des techniques nouvelles et prometteuses de transmission de données, telles que l'appareil enregistreur d'événements avec transmission par radio, la télétransmission par satellites et la technique des événements hydrologiques de caractère violent présentent un progrès notable dans la transmission, en temps réel des données. De récentes innovations dans le traitement préliminaire de telles données en vue d'assurer une bonne qualité aux résultats, sont également très intéressantes.

INTRODUCTION

The river and flood forecasting service in the USA produces annually over 400 000 forecasts for 2500 riverside communities. These forecasts are prepared by 13 River Forecast Centers (RFC) who collect and process daily hydrological data from some 7000 river and rainfall observation points. Ninety per cent of the data is manually collected and relayed by telephone to National Weather Service (NWS) offices.

Floods cause nearly 200 deaths and over \$2 billion of damage annually in the USA. In recent years they have become the major natural disaster in our country.

NETWORK AUTOMATION

A river forecasting service is only as good as its data collection system. Data are the life blood of the complex hydrological forecast models and the heart of the system that can simulate river flow conditions upon which good water management decisions can be made.

Data may include rainfall, streamflow, temperature, humidity, soil moisture, snow water equivalent, snow depth, cloud cover, radiation, evaporation or evapotranspiration, and winds. The data requirements are dependent upon the type of forecast relationship utilized, area of concern, and accuracy of required forecasts. The time scale is important also. For example, the data requirements may vary for the forecasting of flash floods, major river flooding, low flows, and water supply for water management.

Most hydrological data collection systems are manual. As a result, these systems are inadequate today to meet real time river forecasting and flood warning services. Thus, automation of data systems is necessary to forecast accurately rapid river fluctuations, to issue timely forecasts for operation of spillways and flood control works, and to ensure prompt forecasting of river flow for power generation, barge traffic, and pollution abatement programs.

AFOS

Certainly one of the most important programs initiated in the NWS in the past few years has been the Automation of Field Operations and Services (AFOS). It is basically a data handling and transfer system. All offices of the NWS will be linked on national circuits using minicomputers to transmit all data and/or information. Our RFCs will also use the minicomputers as remote terminals into the central computer facility (3 IBM 360/195s) of the National Oceanic and Atmospheric Administration (NOAA) located near Washington, DC. A primary advantage of the program is that information available to one office will be available to all. This is particularly important to the river program which cuts across state and regional lines.

AFOS contains many desirable features beyond its data transfer capability. Data can be retained in the system for several days; thus, it will be possible to utilize these data in models that require a relatively long memory, e.g. low flow computations or soil moisture accounting. Also, it will be possible to display graphically products such as hydrographs and precipitation maps. This will alleviate work done by hand in the past and speed up the issuance of watches and warnings in areas affected by severe flooding. The formatting of messages for these forecasts can be done also *a priori*, which speeds up their preparation.

AHOS

Beginning in 1966 an Automatic Hydrologic Observing System (AHOS) was established by the NWS in the Potomac River basin above Washington, DC. Data from telemetered rain and river gauges were transmitted via telephone lines to a central interrogator in Washington. These data were printed on a teletypewriter output on demand. The AHOS/T (telephone) system has been expanded to include about 415 gauges operating throughout the USA. This system is now interrogated by minicomputers which use automatic switching (direct dial) telephone facilities. Where no land lines exist the AHOS/S (satellite) system is used. It utilizes a geostationary satellite (GOES) with communications relay facilities. The Data Collection Platform (DCP) utilizes a UHF radio set for data transmission to the satellite which in turn transponds data to a base station which then relays it to RFCs or other users. The NWS currently has about 46 AHOS/S stations operating. Another 400 are being installed by other agencies.

The AHOS/S stations have proved very reliable. During severe flood situations the telephone systems utilized by the AHOS/T gauges have frequently failed. Our current policy has been to install the AHOS/S stations wherever possible. The difference in installed cost is from approximately \$2000 for an AHOS/T station to \$6000 for an AHOS/S station.

Both tipping bucket and weighing type precipitation gauges are used with AHOS. Although the accuracy of the tipping bucket gauge is reduced at high rainfall rates because of splash-out from the rapidly oscillating tipper, these gauges have proved to be very reliable under adverse conditions. Some difficulties have been encountered with the weighing type gauges relative to the telemetry contacts.

Event reporting gauge

A precipitation gauge which has proved very effective has been the event reporting gauge. These gauges employ a tipping bucket precipitation gauge installed in a stand

pipe type casing and relay the information on each tip (1 mm) via VHF radio in real time to a master station. A number of these stations have been in operation for several years in California. So far, they have proved very reliable. In a recent test conducted by the NWS Test and Evaluation Division more than 22 000 tips in 12 days of the gauge were made to evaluate the drain on the battery which supplies power and the reliability of the gauge. That many tips represents a catch of several years at most sites. The battery drain was less than about 10 per cent and the reliability of transmission of data was over 99 per cent (an erroneous transmission on only 43 tips out of 22 000). This type of gauge is being employed primarily in areas subject to flash floods where continuous response is needed.

Radar

Considerable emphasis has been placed in recent years on radar to estimate precipitation. In particular, the NWS has been experimenting with the operational application of digital radar (currently we have four sets in operation). The overall success has been disappointing. The primary problem has been due to the lack of reliability in converting digital radar return into rainfall estimates. A secondary problem involves moving the large volume of information available from the radar site to the various offices requiring the data. Current activities include the development of a data management and analysis program that performs preliminary processing on-site in the radar mini-computer. These analyses are then transmitted to a central site having large-scale data processing capabilities for merging or compositing information from multiple radar sites and combining it with raingauge and/or satellite data. Although it was hoped that with the implementation of the AFOS communications network that the data transmission problem would be solved, this may not be true due to other hydrological/meteorological data loading requirements.

Satellites

With the launching of the GOES series (Geostationary Operational Environmental Satellite) in 1974, the potential for estimating rainfall from satellite imagery was enhanced markedly. Data are routinely available from these satellites every half hour. Models have been developed which utilize both infrared and visible imagery. Basically, they are oriented toward convective storms and use a spatial enhancement of infrared imagery to pick up thunderstorm cores. The method follows a decision tree approach where, depending on storm development, a numerical estimate of rainfall for the previous half-hour period is made for particular points. The results to date have been particularly encouraging although the frequent lack of ground truth has made it difficult to evaluate fully the worth of this technique.

Meteor burst telemetry (SNOTEL)

The Soil Conservation Service (SCS) in the USA has been automating their snow gauging sites recently using meteor burst trails to relay VHF radio transmitted data. Currently 230 stations have been installed by the SCS in the lower 48 States and several stations in an interagency program in Alaska. Information relayed includes temperature, precipitation, and snow water equivalent data. Results to date have been very encouraging.

DATA MANAGEMENT AND PREPROCESSING

Since 1971 the NWS has been developing a hydrological forecasting system known as NWSRFS (National Weather Service River Forecast System) that includes a comprehensive set of hydrological techniques required by the RFCs to perform their operational functions. In addition to including mathematical models for soil moisture

accounting, river routing, snow accumulation and ablation, and many others, the system also includes procedures for archiving, retrieving, and processing the types of data needed to apply the system.

The NWSRFS has been organized into three separate groups of modules:

- (1) data management (DM) modules,
- (2) preprocessor modules,
- (3) forecast modules.

Only the first two modules will be discussed here.

DM module

This module is the initial interface between the user and the forecast system. It allows the user to: (1) date the observation data, (2) enter time series data (i.e. precipitation, temperature, stage/discharge, and potential evapotranspiration data), (3) display/print time series data, (4) enter parameter data (these are required to execute other preprocessor and forecast models) and (5) execute a forecast run.

Unfortunately, the only real quality control of data in the DM module currently involves visual checks made of data printed out – frequently in mapped form. The NWSRFS operational forecast system is currently undergoing extensive revision. The new version will contain decode, validity check and quality control subroutines for editing of the raw data. These data will then be read into an observed data base read/write system where they will be utilized by the preprocessor routines.

Preprocessing module

The preprocessor module uses the time series data and their corresponding parameters to compute: (1) mean areal precipitation, (2) mean areal temperature, (3) potential evapotranspiration, and (4) discharges. The forecast module then uses these computed values along with carryover values to compute runoff at specified forecast points. This module is crucial to an effective forecast system. A wide variety of observed data are available at any one time. In addition to our official observing network each office has access to many unofficial hydrometeorological gauges which can be introduced as ‘stranger’ reports. ‘Observed’ data can also include certain projected quantities such as Quantitative Precipitation Forecasts (QPF) and projected reservoir outflow (for those reservoirs with gated outlets). The preprocessor also must be able to handle data for a wide range of time intervals from 1 to 24 h.

A brief description of each subroutine currently in the NWSRFS preprocessor module is given below.

Mean areal precipitation (MAP)

This routine is presently written as a relatively simple procedure that estimates precipitation values at selected points. This is done by finding the closest station with an actual report for a given time in each of four quadrants surrounding the point to be estimated. The weight factor applied to each station is the inverse of the distance squared. In mountainous terrain where orographic influences dominate it is necessary to employ a station characteristic for each station. If enough points are computed it is a simple matter to compute the arithmetic mean of all points within a given basin to obtain the basin mean.

Since most observers report only once each 24 h (generally early in the morning), it is necessary to distribute the 24-h MAP values into values for shorter durations. This is done generally by computing the percentage of 24-h precipitation falling during selected durations at automatic and first order stations. The centre of mass of the MAP area is then determined and the 24-h MAP for that basin distributed temporally using the nearest stations having time distributed percentage data in each of four quadrants.

The weight applied to each station equals the inverse of the distance squared from the centre of mass.

Other objective analysis procedures are being examined for inclusion in the MAP preprocessor. Also, techniques for estimating more accurately the precipitation in orographic areas have been developed recently and will be incorporated into this module.

Mean areal temperature (MAT)

MAT calculations are used primarily in the snowmelt routines. In non-mountainous areas the temperature varies linearly with distance. Thus, values at grid points within a given drainage basin can be determined by the quadrant technique described for estimating precipitation data but computing station weights using the reciprocal of the distance rather than the distance squared. If stations are distributed in a reasonably uniform manner, the Thiessen method or an arithmetic average will yield satisfactory results.

In mountainous areas it is frequently necessary to compute temperatures over basins with little, if any, data. In regard to station weights, the two most important factors are probably distance and elevation. If two stations are equidistant from point X , studies have shown that the one closest in terms of elevation is usually the best estimator. Experience has shown also that the differences between station means are a good indication of the typical variations in temperature that exist in mountainous areas. Generally, mean monthly temperatures can be used to estimate this variation.

Temperatures for a given station are normally available as maximum/minimum data. Since snowmelt occurs primarily during the spring, several linear relationships have been developed relating the mean temperature for selected 6-h periods, e.g. midnight to 6 a.m., to these data. Thus it is impossible to compute MAT values over selected basins for durations of variable length. This is particularly important when working with estimating runoff from small basins.

Potential evapotranspiration (PE)

Estimates of PE are obtained in several ways. One of the most straightforward methods is to use pan evaporation data as an estimate. Although this method has been employed by some offices, the results to date have not been encouraging. Most evaporation stations in the USA employ the Class A pan. The variability of the pan coefficient is probably large enough, even on a daily basis, to present problems in daily use. The most common method is to simply use daily mean values of PE demand which vary with day of the year. PE values also are computed using the meteorological factors of air temperature, dew point, daily wind movement and solar radiation (which is generally inferred from per cent sunshine or sky cover). In most cases this latter technique utilizes the observed values from yesterday to estimate the PE for today.

Discharges

This routine simply converts observed stages to discharges for each forecast point and *vice versa*. A 25-point table is used with interpolation between points. If a reported stage is 'out-of-bounds', a log-log or hydraulic extension procedure is used to estimate the discharge. In the current version loop ratings are not allowed. This can cause serious errors on many streams with soft bottoms or highly variable slope during severe floods. Future modifications will attempt to correct this problem.

