A MICROCOMPUTER-BASED WATERSHED-MODELING AND DATA-MANAGEMENT SYSTEM

G.H. Leavesley[1], A.M. Lumb[2], and L.G. Saindon[1]

INTRODUCTION

Program size, computation time, and data-storage requirements historically have limited the use of most watershed models to mainframe computers and minicomputers. However, recent advances in microcomputer technology have minimized or eliminated many of these limitations and have made feasible the application of microcomputers to watershed modeling. The use of microcomputers decreases the costs of model applications and makes models accessible to a much larger group of users in the fields of hydrology and water-resources management.

To take advantage of these technological advances and to provide watershed-modeling capabilities to the microcomputer-user community, the U.S. Geological Survey recently began the development of a microcomputer-based watershed-modeling and data-management system. Some system components are being developed from operational programs currently running on larger computer systems; other components are new programs. The major components of the system are the U. S. Geological Survey's Precipitation-Runoff Modeling System (PRMS) (Leavesley and others, 1983), the U.S. Geological Survey's interactive data management and control program ANNIE (A. M. Lumb, U.S. Geological Survey, written commun., 1987), and a modified version of the National Weather Service's Extended Streamflow Prediction (ESP) program (Day, 1985). Together they form a complete system that enables a user to reduce, analyze, and prepare data for model application, simulate and forecast watershed response, and statistically and graphically analyze model results.

The purpose of this report is to describe the features and capabilities of the watershed-modeling and data-management system. Some of the features discussed currently (1987) are being developed and tested.

SYSTEM DESCRIPTION

The three major components of the system and the data transfers among these components are depicted schematically in Figure 1. ANNIE is a modular-design interactive program that provides the data-management and analysis functions of the system. Data from a variety of sources are reformatted by ANNIE to a system-compatible file structure. Hydrologic and meteorologic data are placed in a direct-access file structure called the Watershed Data Management (WDM) file. Data in the WDM file can be analyzed using a number of statistical, graphical, and data-management functions. Model- and watershed- specific data can be entered interactively into ANNIE to construct the required input files for PRMS.

PRMS is a modular-design, distributed-parameter watershed model. PRMS reads model initialization and watershed-characterization data from files created by ANNIE. Meteorologic and hydrologic data are read directly from the WDM file. During a simulation, time traces of simulated and measured streamflow and selected variables that describe the state of the watershed system can be printed out or written to a data file called the PLOTGEN file. Data in the PLOTGEN file can be analyzed after a simulation using the statistical and graphical functions in ANNIE.

A modified version of the ESP program is coupled to PRMS for forecasting purposes. The ESP component produces a probabilistic forecast of streamflow variables, such as maximum flow, flow volume, and the date flow decreases to less than a selected value. A more detailed description of ESP and the other system components in Figure 1 is presented hereafter.

Presented at the Western Snow Conference, Vancouver, British Columbia
1 Hydrologist, U.S. Geological Survey, Denver, Colo.
2 Hydrologist, U.S. Geological Survey, Reston, Va.

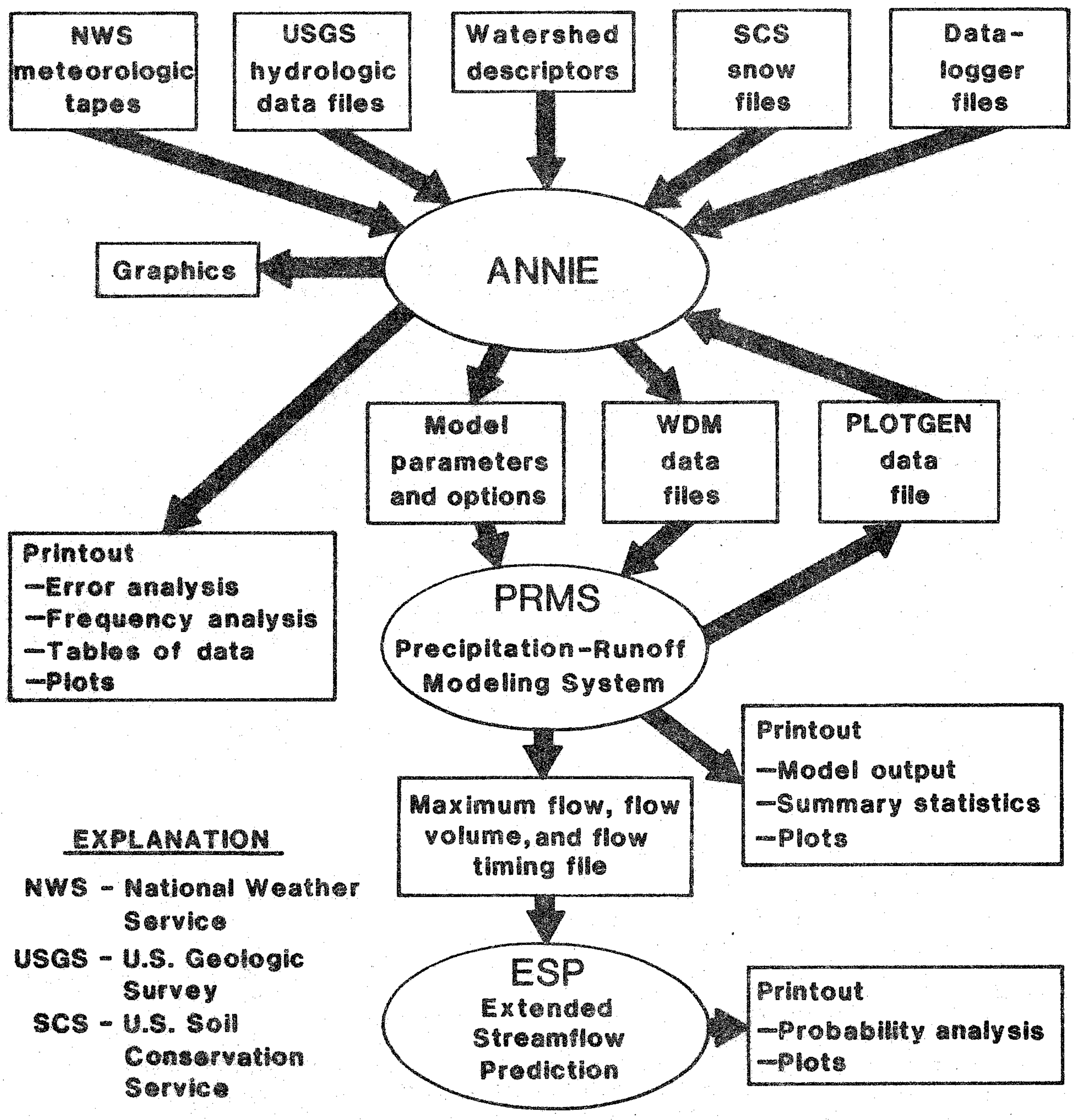

Figure 1. Schematic diagram of data transfers in the watershed-modeling and data-management system.

<u>ANNIE</u>

ANNIE is a system of software modules designed to help a user interactively create, check, and update input to hydrologic models, and to provide statistical and graphical tools to assist in the analysis of model input and output. The design of the user interactions in ANNIE is a hierarchial system of branches. Branches are selected by the use of menus. Each response moves the user farther down the hierarchy until the desired operation is reached. A user response of a question mark to any menu item produces a help message with additional information about the question and the alternative responses. Documentation on the concepts, design, and application of ANNIE is being developed.

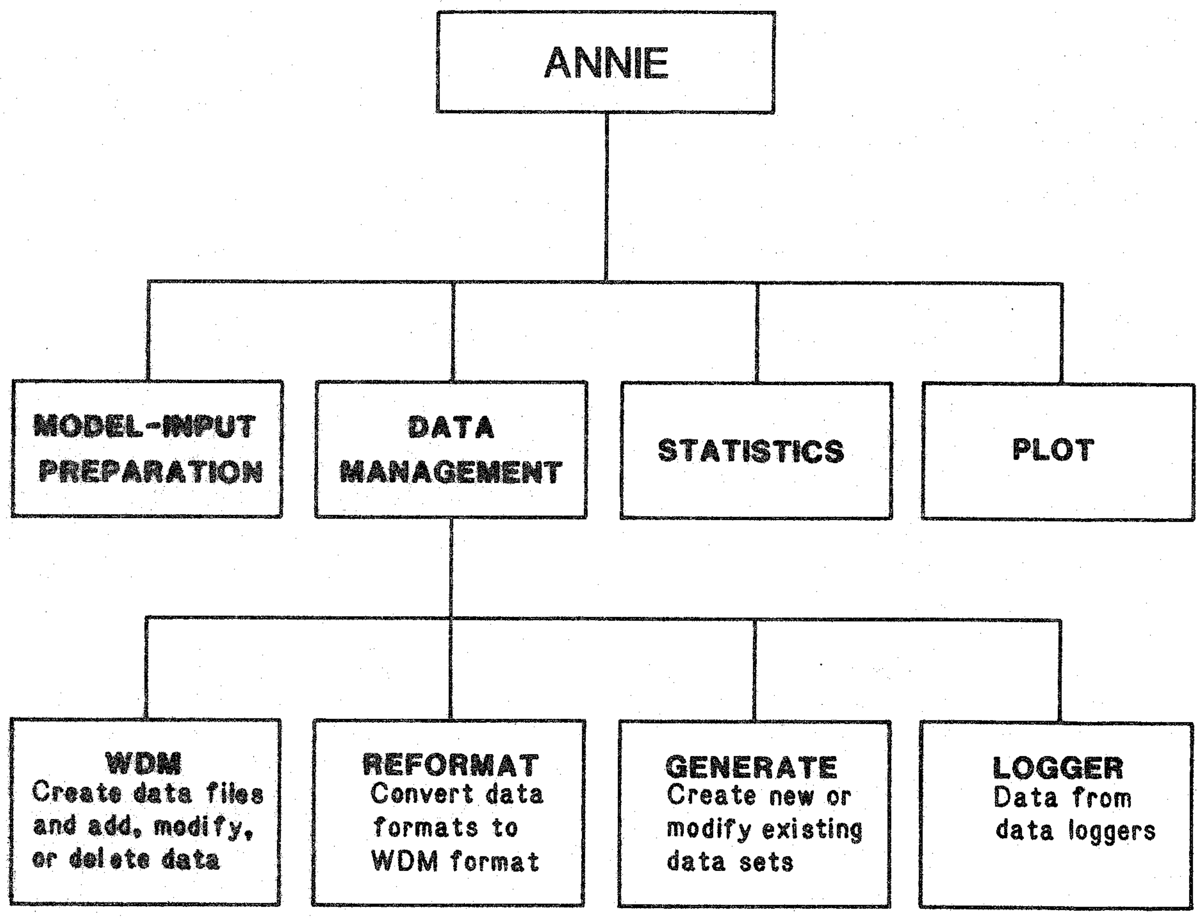

Figure 2. Schematic diagram of the major branches of the data-management and analysis program ANNIE.

The major branches of ANNIE useful to watershed modeling are depicted schematically in Figure 2. A description of these branches is presented hereafter.

Model-Input Preparation

This component assists in the preparation of input files for PRMS. Model parameters and variables, watershed characteristics, and model options are contained in these files. Input to PRMS is divided into eight files. The number of files required for a model run is a function of the model options selected. The user is queried for the data values in each of the files.

ANNIE places user responses in the correct file format, then provides a series of checks on selected responses. Some responses are checked against predefined maximum and minimum values to insure that the response is within an acceptable range. Responses that affect other responses are compared with those responses to insure no incompatibilities.

Data Management

WDM - The WDM component is used to create a WDM file, define the data sets in the WDM file, and perform a variety of data-management functions on the data sets. After the WDM file is created, data sets are added by specifying the data-set number and the attributes associated with each data set. Attributes describe the type, geographic location, and other characteristics of the data set. A selected number of attributes are required by ANNIE for each data set; additional attributes may be supplied by the user. Data values can be entered into the data set using the Reformat option, or they can be entered manually from the keyboard. A third data-entry option accepts data directly from data loggers; it is described later in this report.

WDM management options include functions to add, modify, or delete data in specified data sets. Data also can be copied from one data set to another. A list function enables the listing of as many as six data sets in a tabular format, where rows are time-steps, and each column is a data set. The list function also can be used to produce summaries of time-series data; sums or averages of the data values for periods of 1 day, 1 month, or 1 year can be requested. The tabular file generated by the list function can be used as an input file to several commercially available statistical packages.

Search capabilities, using Boolean algebra, are available to find data sets in the WDM file with attributes that match a set of user-defined attributes. The data-set numbers of matching data sets are placed in a buffer for subsequent use with other ANNIE functions. Functions such as list, modify, plot, reformat, and statistics can be performed on the selected data sets.

Reformat - Reformat converts hydrologic and meteorologic time-series data from a number of different sources into the system-compatible, direct-access WDM file structure. Input from the National Weather Service's climatic-data files, U.S. Geological Survey's National Water Data Storage and Retrieval (WATSTORE) system, and the U.S. Soil Conservation Service's Centralized Forecast System (CFS) currently (1987) are supported. Conversion from a WDM file format to a WATSTORE format also is supported to enable the transfer of data from ANNIE to other applications.

Time-series data can have time steps ranging from 1 minute to 1 year; they may be stored as a uniform time step or as breakpoint or date-tagged data. A quality flag can be set to indicate such conditions as missing data and estimated data. Space for each data set is allocated dynamically and space from deleted data sets is reused.

Generate - Generate provides the capability to create new data sets. One procedure generates a new data set by applying a mathematical operation to one, or selected combinations of two, existing data sets. Twenty-three different mathematical operations are available for use. Mathematical functions include log transformations, power transformations, weighted sums, and the addition or subtraction of, and multiplication or division by, constants.

A second procedure generates a new data set by changing the time step of an existing data set. A disaggregate transformation expands a time series into shorter time steps. Data values for the new time steps can be set equal to the data values for the associated old time steps or they can be set to be an equal proportion of the data values for the associated old time steps. An aggregate transformation condenses a time series into longer time steps. The new time-step values can be specified to be the average, sum, maximum, or minimum of the values in the associated old time steps.

Logger - Hydrologic and meteorologic data, collected using digital data loggers, can be directly entered into a WDM file. Procedures currently (1987) available were developed to work with data collected using Campbell Scientific CR-21 and CR-21x data loggers*. Data-reduction programs supplied with these loggers put all retrieved data into a tabular format. Table columns include a row identifier, julian date, time, and one or more data fields. The row identifier normally denotes the time step or frequency at which the data values in that row were collected. Data for a maximum of 5 different time steps, with a maximum of 12 data fields per time step, can be entered into a WDM file at one time. If more than 12 data fields occur in any of the time steps, the file needs to be read multiple times.

After the data sets are in the WDM file, they can be reviewed, analyzed, or modified using other ANNIE functions. The plot functions permit a rapid visual analysis of the data, and the list functions provide a printed copy of the data. Data recorded in units defined by the sensor can be converted to engineering units using the mathematical operations available in the Generate component of ANNIE.

---

* The use of trade-names is for descriptive purposes only and does not constitute endorsement by the U.S. Geological Survey.

Statistics

Statistical-analysis capabilities are provided by both internal- and external-applications programs. Commonly used statistical procedures such as the computation of means, standard deviations, correlations, and simple linear regressions are done in ANNIE. Frequency analysis and error analysis also are supported in ANNIE. More complex statistical procedures are accommodated by reformatting WDM data sets to a sequential flat file that is compatible with statistical packages external to ANNIE.

Plot

The plot component provides the graphic capabilities in ANNIE. Data sets in the WDM file, results of selected statistical analyses, and PRMS results placed in a PLOTGEN file can be displayed graphically on a monitor or on a pen plotter. Graph formats include time-series plots of one to five data sets, x-y plots of selected data sets, frequency plots, flow-duration plots, and scatter diagrams.

PRMS

PRMS is a modular-design, distributed-parameter physical-process watershed model that was developed to evaluate the effects of various combinations of precipitation, climate, and land use on watershed response. Watershed response to normal and extreme rainfall and snowmelt can be simulated to evaluate changes in water-balance relations, flow regimes, flood peaks and volumes, soil-water relations, sediment yields, and ground-water recharge. Parameter-optimization and sensitivity-analysis capabilities are provided to fit selected model parameters and to evaluate their individual and joint effects on model output.

Watershed response can be simulated at both a daily and a storm time scale. The daily mode simulates hydrologic components as daily average or total values. Streamflow is computed as a mean daily flow. The storm mode simulates selected hydrologic components at time intervals shorter than 1 day. The minimum time interval is 1 minute. The storm mode computes storm hydrographs and sediment yields for selected rainstorms. Sediment-modeling capabilities are only provided in the storm mode.

Distributed-parameter capabilities are provided by partitioning a watershed into units, using characteristics such as slope, aspect, elevation, vegetation type, soil type, and precipitation distribution. Each unit is assumed to be homogeneous with respect to its hydrologic response and to the characteristics listed above; each unit is called a hydrologic response unit (HRU). A water balance and an energy balance are computed daily for each HRU. The sum of the responses of all HRUs, weighted on a unit-area basis, produces the daily watershed response.

A second level of partitioning is used for storm-mode computations. The watershed is conceptualized as a series of interconnected flow-plane and channel segments. Surface runoff is routed over the flow planes into the channel segments using the kinematic wave approximation to overland flow; channel flow is routed through the watershed channel system using the kinematic wave approximation for channel flow. An HRU can be considered the equivalent of a single flow plane, or it can be divided into a number of flow planes.

The conceptualized watershed system and its inputs are schematically depicted in Figure 3. Model inputs are daily precipitation, maximum and minimum air temperature, and solar radiation. Precipitation in the form of rain, snow, or a mixture of both is decreased by interception and it becomes net precipitation delivered to the watershed surface. The energy inputs of air temperature and solar radiation data are used in the computation of evaporation, transpiration, sublimation, and snowmelt. If solar-radiation data are unavailable, estimates of daily solar radiation are made using air temperature and computed potential solar-radiation data.

The watershed system is conceptualized as a series of reservoirs, the responses of which combine to produce the total system response. The soil-zone reservoir represents that part of the soil mantle that can lose water through the processes of evaporation and transpiration. Average rooting depth of the predominant vegetation covering the soil surface defines the depth of this zone. Water storage in the soil zone is increased by infiltration of rainfall and snowmelt and it is decreased by evapotranspiration. The soil zone is

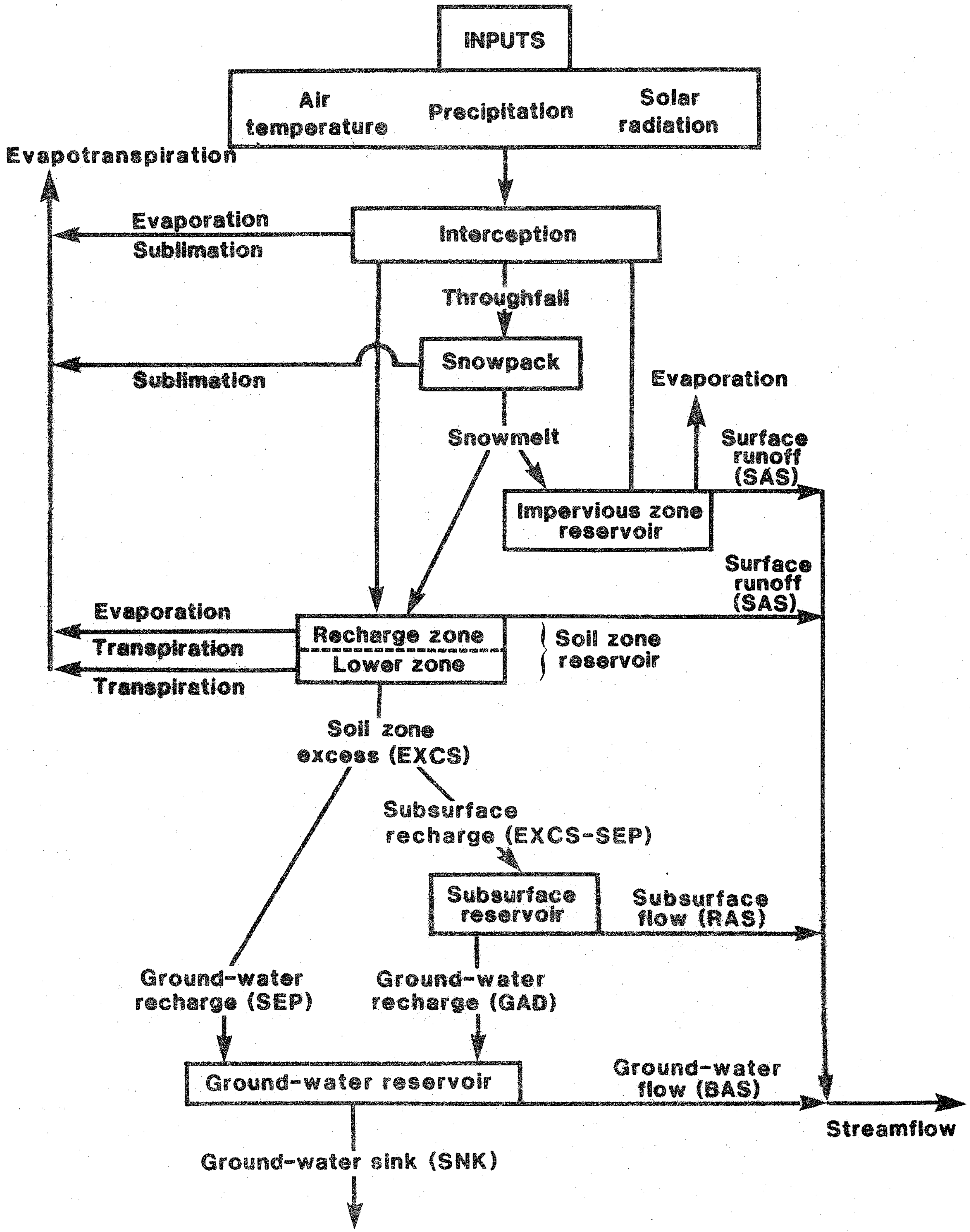

Figure 3. Schematic diagram of Precipitation-Runoff Modeling System (modified from Leavesley and others, 1983).

treated as a two-layered system. Losses from the upper zone, termed the recharge zone, occur from evaporation and transpiration; losses from the lower zone occur only through transpiration.

When the soil reservoir reaches field capacity, infiltrated water in excess of field capacity moves to the subsurface and ground-water reservoirs. Recharge to the ground-water reservoir is removed first from the soil-reservoir excess and is assumed to have a daily upper limit. Water in excess of this limit is delivered to the suburface reservoir. The subsurface reservoir simulates the relatively rapid component of flow that may occur in the saturated-unsaturated and ground-water zones during periods of rainfall and snowmelt. The ground-water reservoir simulates the slower component of flow from the ground-water zone. The subsurface reservoir can be defined as being linear or nonlinear; the ground-water reservoir is assumed to be linear.

The snow components of PRMS simulate the accumulation and depletion of a snowpack on each HRU. A snowpack is maintained and modified both as a water reservoir and as a dynamic heat reservoir. A water balance is computed daily, and an energy balance is computed twice each day. The energy-balance computations consider net shortwave and longwave radiation, approximations of convection and condensation terms, and the heat content of precipitation.

The snowpack is assumed to be a two-layered system. The surface layer is the upper 3-5 centimeters of the snowpack and the lower layer is the remaining snowpack. Heat transfer between layers occurs by conduction when the temperature of the surface layer is less than 0 °C (degrees Celsius). When the surface temperature equals 0 °C, heat transfer occurs as conduction when the net energy balance at the air-snow interface is negative; the heat transfer occurs as mass transfer from surface melting when the net energy balance is positive. Heat transfer from precipitation occurs as a mass-transfer process. Conduction of heat across the soil-snow interface is assumed to be negligible.

Surface-layer melt and rainfall moving into the lower layer are used first to satisfy part or all of the heat deficit in the lower snowpack. When the entire snowpack is isothermal at 0 °C, remaining liquid water is used to satisfy part or all of the free-water holding capacity of the snowpack. After the free-water holding capacity is satisified, the remaining liquid water becomes available for infiltration and surface runoff.

Parameter-optimization routines are included to automatically adjust a user-specified set of parameters to improve the agreement between measured and predicted runoff. Two optimization techniques are available: one is the Rosenbrock technique (Rosenbrock, 1960); the second is a Gauss-Newton technique similar to the linearization method described by Draper and Smith (1966), but with the incorporation of the Box-Kanemasu interpolation modification (Beck and Arnold, 1977). One of four objective functions can be selected for use with either technique. Parameter-sensitivity-analysis routines are available to evaluate the extent to which uncertainty in the parameters results in uncertainty in the predicted runoff, and to assess the magnitude of parameter errors and parameter intercorrelations when an optimization is performed.

Model-printout options range from a table of measured and predicted streamflow values to daily, monthly, annual and storm streamflow summaries. Daily, monthly, or annual summaries also are available for major water-balance components and selected model variables that describe the state of the watershed system. These detailed summaries are available for basin-average values and for each individual HRU.

Daily values of precipitation, potential and simulated evapotranspiration, soil moisture, measured and predicted streamflow, and the surface, subsurface, and ground-water flow components of predicted streamflow can be written to a PLOTGEN file for processing by ANNIE. These values can be output for the total basin or for as many as 10 HRUs. ANNIE then can be used to read the PLOTGEN file, and to produce time-series plots and statistical analyses of some or all of the data in the PLOTGEN file.

Extended Streamflow Prediction (ESP)

A modified version of the ESP program has been coupled to PRMS to provide forecasting capabilities. The ESP procedure uses historic or synthesized meteorologic data to forecast future streamflow, given the simulated hydrologic conditions for a watershed at a specified point in time. When historic data are used, the procedure assumes that past meteorological events are representative of future meteorological events. Alternative assumptions about future meteorological conditions can be made with the use of synthesized meteorological data.

To implement the ESP procedure, the user defines in PRMS a forecast period, the meteorological data set to be used, and the forecast variables of interest. The forecast period can vary from a few days to an entire water year. PRMS simulates to the start of the forecast period then stores all model variables that describe the state of the watershed system. One streamflow trace for the forecast period then is simulated for each year in the meteorological data base. PRMS is reinitialized at the start of each new forecast trace using the previously stored variables. The simulation results obtained from this iterative procedure are termed the conditional simulation.

Streamflow variables of interest are extracted from each forecast trace and stored in an ESP file. The variables that can be extracted are the maximum daily flow value, minimum daily flow value, flow volume, and the date that the flow decreases to less than a selected threshold value. One variable, or any combination of variables, can be selected for a given model run.

The ESP program reads the ESP file, performs a frequency analysis on the variables, and produces a probabilistic forecast for each of the variables in the file. The Log Pearson Type III probability distribution is currently (1987) supported, but other distributions can be added. Frequency analysis results for the conditional simulation are presented in tabular form and in a printer plot.

A frequency analysis also can be performed with ESP using the historic streamflow record or the simulation results obtained from running PRMS using the historic or synthesized meteorologic data without resetting the initial conditions each year. The results of this continuous simulation are termed the historic simulation when historic data are used. A comparison of frequency-analysis results between the measured data and the historic simulation provides a measure of the adequacy of the PRMS calibration. Differences between the analyses reflect the effects of model and data errors on the historic simulation, and they also provide a measure of the bias in the conditional-simulation results.

A continuous simulation using synthesized meteorologic data is termed the synthetic simulation. Differences between the analyses for the historic simulation and the synthetic simulation provide a measure of the effect of the assumed meteorological changes in the synthesized record on streamflow.

The ESP component provides a valuable tool for a variety of hydrologic forecasting and analysis applications. These applications include short-term or seasonal forecasting for water supply and flooding, evaluation of the effects of land-use changes on streamflow characteristics under a variety of meteorologic and watershed conditions, and the effects of climatic changes on streamflow under existing or modified watershed conditions.

SYSTEM STATUS

All capabilities in the WDM and Reformat options of ANNIE are fully operational. Parts of the Model-input, Generate, Logger, and Statistics functions currently (1987) are being downloaded from a minicomputer and tested. The Plot functions are being developed under contract; delivery of these routines is anticipated in August 1987.

The daily-mode components of PRMS are fully operational and the parameter-optimization and sensitivity-analysis routines are being tested. Downloading and testing of the storm-mode components is not anticipated until July 1987. All components of the ESP package are fully operational.

In its present state, the system can create and modify WDM data files, perform daily-mode simulation with PRMS, and forecast selected streamflow characteristics using ESP. Implementation of the full data-management, analysis, and simulation capabilities of the system is anticipated in the near future.

SYSTEM SPECIFICATIONS

Software and Hardware

All components of the system are written in standard ANSI Fortran 77 and they have been compiled and linked for use with the MS DOS or PC DOS operating system. Minimum hardware requirements are 640K bytes of random-access memory (RAM) and a 10-megabyte hard disk. To use the graphics capabilities of ANNIE, a graphics card and compatible monitor also are required.

Performance

The current (1987) executable run file for ANNIE is 267K bytes and it uses a direct-access data file containing the ANNIE dialog that is 350K bytes. Any WDM file created by ANNIE will require additional storage. For example, a WDM file containing 8 water years of daily precipitation, maximum and minimum air temperature, solar radiation, and streamflow data uses 81,920 bytes of disk storage. However, two WDM files containing an equal number of water years and data sets may have slightly different storage requirements. A data-compression technique is used in the WDM when sequences of the same value occur within a data set.

The computational speed of ANNIE is, in part, dependent on the user because the system is interactive. Total time required for any of the procedures will be a function of the users' familiarity with ANNIE and microcomputers. Creation of the example WDM file discussed before took the author about 15 minutes.

The executable run file for PRMS exceeds 640K bytes, but the program is implemented using a series of overlay procedures. Computational speed will vary with the microcomputer used. Test results using a watershed from a region with several months of snowmelt runoff, configured with 15 HRUs, and simulated in the daily mode, showed a computation speed of about 1 minute and 50 seconds per water year on an IBM PC AT using a 6-megahertz clock. This time was decreased to 1 minute and 20 seconds per water year on a Compaq AT compatible using an 8-megahertz clock. Both systems had an 80287 math coprocessor installed. Computational speed using a Toshiba AT compatible with an 8-megahertz clock, but no 80287 math coprocessor, was 4 minutes and 15 seconds per water year.

The executable run file for ESP is about 200K bytes. Neither storage or computational speed are major concerns for this component.

SUMMARY

A watershed-modeling and data-management system that provides a range of capabilities from data reduction and analysis to simulation and prediction of watershed response has been developed for microcomputer application. The system is composed of three major components. Interactive data-management and analysis capabilities are provided by a modular set of programs called ANNIE. Data can be reduced, reformatted, transformed, analyzed statistically and graphically, and placed in a model-compatible format using ANNIE. Watershed-simulation capabilities are provided by PRMS, a modular-design, distributed-parameter watershed model. PRMS can simulate watershed response in both a daily mode and a shorter time-step storm mode. Forecasting capabilities are provided by a modified version of the ESP program. ESP has been coupled with PRMS to produce probabilistic forecasts of a number of user-selected streamflow variables. Together these components form a powerful tool for operational and research applications.

Limitations are imposed on the system by the current capabilities of microcomputer technology and by current knowledge of hydrologic modeling. The speed and capabilities of the system will increase as technological advances in available memory, processing speed, and hardware design are implemented, and as advances are made in the hydrologic sciences. The modular design of the system provides a flexible framework in which advances in both microcomputer technology and hydrology easily can be incorporated.

REFERENCES

Beck, J. V., and Arnold, K. J., 1977, Parameter estimation in engineering and science: New York, John Wiley and Sons, Inc., p. 340-349.

Day, G. N., 1985, Extended streamflow forecasting using NWSRFS: American Society of Civil Engineers, Journal of Water Resources Planning and Management, v. 111, no. 2, p. 157-170.

Draper, N. R., and Smith, Harry, 1966, Applied regression analysis: New York, John Wiley and Sons, Inc., p. 267-270.

Leavesley, G. H., Lichty, R. W., Troutman, B. M., and Saindon, L. G., 1983, Precipitation-runoff modeling system--User's manual: U.S. Geological Survey Water Resources Investigations Report 83-4238, 207 p.

Rosenbrock, H. H., 1960, An automatic method of finding the greatest or least value of a function: Computer Journal, v. 3, p. 175-184.