VI.2.5 EXTENDED STREAMFLOW PREDICTION SYSTEM (ESP) APPLICATION INFORMATION

This Section contains information about some of the items that should be considered before using the Extended Streamflow Prediction (ESP) system.

The ESP system includes the ESP Initialization (ESPINIT) program and the ESP Function in program FCST.

Introduction

The ESP system is that portion of the NWSRFS which provides the capability of making long-range probabilistic forecasts of streamflow and streamflow-related variables. ESP uses conceptual models to forecast future streamflow using the current snow, soil moisture, river and reservoir conditions with historical meteorological data. The ESP procedure assumes that meteorological events that occurred in the past are representative of events that may occur in the future. Each year of historical meteorological data is assumed to be a possible representation of the future and is used to simulate a streamflow trace. The simulated streamflow traces can be scanned for maximum flow, minimum flow, volume of flow, reservoir stage, etc., for any period in the future. ESP produces a probabilistic forecast for each streamflow variable and period of interest.

Much of the information needed for ESP is available from the Forecast Component of the Operational Forecast System. The order of Segment execution, the Operations Table for each Segment, the parameters for each Operation, Rating Curve information and the carryover are all obtained from Forecast Component files. Historical data are obtained from the ESP historical data files (calibration type files) and future data are obtained from the processed data base. Additional ESP type information is available from the ESP parameter file. The ESPINIT program is used to load ESP information on the ESP parameter file.

Implementation

A number of steps must be accomplished if ESP is to be successfully implemented. The first step should be the development of an implementation plan. It is important to develop a plan to ensure that all the steps have been identified and will be completed by the time they are needed. It may be helpful to include a test area as part of the plan. The test area could consist of a limited number of basins with important ESP applications. Experience with the test area would enable the user to become familiar with the capabilities and constraints of the ESP system. The knowledge gained from the test area will be very valuable when ESP is set up on large portions of a user's area.

Some of the steps that should be considered in the implementation plan are outlined briefly here. One of the most time-consuming tasks in the implementation process is the data preparation step. ESP requires historical time series of MAP and MAT. MAT is only needed in areas where there is significant snow accumulation. For ESP purposes it is desirable to have as long a historical record as possible. It is also essential that all the historical time series needed to simulate a river system be developed for the same period, since the output from one basin is needed as input to another. 20 years are probably needed to define a reasonable frequency distribution, but ESP will work with as many or as few years as are available. If a long period of MAP and MAT is not available, then it must be developed. This involves loading station data on the calibration data files from the historical data tapes, checking the data for consistency, calculating station weights, and creating mean areal time series.

In the past the historical data tapes have been extremely difficult to work with. In order to obtain the entire period of record for a given station it was necessary to access several tapes. It also took several years for the most recent data to become available. This procedure will be streamlined so that fewer tapes will have to be accessed to obtain a station's record and the lag time before data becomes available will be reduced.

The consistency of each station's data should be checked to ensure that the mean areal time series created are also consistent. Even stations that are only used for estimating missing data should be checked. Calculating station weights may or may not be a serious problem. In non-mountainous areas grid point or Thiessen weights can be used to calculate station weights, and it is a relatively straight forward procedure. However, in mountainous areas the effect of elevation on precipitation and temperature must be accounted for. This requires hydrometeorological expertise and manual computations outside the program. An isohyetal map is invaluable when generating MAP time series in mountainous areas.

After the mean areal time series have been created, some basins may still require calibration, because: they were not included as part of the daily forecast program, a different model should be used for ESP than is currently being used in the daily forecast program, or the current calibration is not considered adequate for ESP purposes. All of the models used to simulate a basin must be calibrated. The difficulty in calibrating a model varies with the number of unobservable parameters, the sensitivity of the parameters, the complexity of the model, and the experience of the user with the model. Research is being done to make the calibration of the Sacramento Soil Moisture Accounting model more automated. Calibration can also be a problem because of the difficulty in obtaining a natural un-depleted streamflow time series. This is a serious problem in the West, because of the many diversions that occur and the difficulty in quantifying them. It is important to develop a streamflow time series that is accurate, so that the model parameters will not be distorted.

Once the models have been calibrated and the historical time series prepared, the basin should be added to the Forecast Component if it does not already exist there. See Section VI.2.2D on FCINIT for a discussion of the steps required for defining information in the Forecast Component. Some specific considerations are given below that relate to the ESP usage of the Forecast Component information. After the information for the basin has been defined, the FCST program can be run to enter data, process time series, execute forecast runs, and save carryover. Some attempt should be made to evaluate forecast system accuracy. If the models constantly need updating to stay on track, the system may not support long-term forecasting with the current models and calibrations. However, it may only be that the operational network is not adequate to keep the models on track. In this case as long as the models are updated, ESP forecasts may not be effected since ESP relies on the climatological data network to make long-term forecasts. These are questions that must be addressed as part of the implementation process.

ESPINIT is used to define ESP information for a Segment after it has been defined in the Forecast Component. The ESP information which must be defined includes information describing the time series to be used and the analysis to be performed. In ESP, input time series are read from the historical data files and output time series are written to the ESP time series files. New external location information must be specified for both input and output time series. The analysis information defined with ESPINIT includes which output variables are of interest for which time series, and with what displays. ESP provides the capability to define, delete, print, and redefine ESP information for a Segment. Once the ESP information has been defined for a Segment, the Segment can be executed with the ESP Function in the FCST program.

FCINIT Considerations

Section VI.2.2D described the FCINIT program and listed some factors to be considered when using FCINIT. This section specifically addresses factors that can potentially impact the ESP system.

- o If the time series needed for a Segment already exist on an ESP historical file before the Segment is defined with FCINIT, the time series should be entered in the DEF-TS section of the SEGDEF command in the same order they are stored on the ESP historical file. This will ensure that the minimum number of disk accesses are made when reading the data.

- o In some cases it may be desirable to analyze time series in ESP that are not needed for daily forecasting. Any time series used in ESP must be defined first in the Forecast Component.

- o ESP will accept any Operations Table defined with FCINIT, but the order of the Operations can be very important in ESP. The number of disk accesses and the number of intermediate time series that must be stored when proceeding downstream with ESP computations can be minimized if all of the Operations needed for each forecast point are executed in succession. This means that all the Operations for a given forecast point should be put in a single Segment or at least in successive Segments.

- o ESP simulates weeks or even months into the future without the benefits of updating, so it is important that the rainfall/runoff Operation used be reliable and have a long memory. ESP can use any continuous procedure, but the procedure should be capable of modeling the entire range of flows needed for forecasting. An API Operation such as API-MKC may be adequate for generating long-range peak forecasts using ESP, but because of its constant baseflow assumption it is not suitable for low flow or volume type forecasts. If low flow and/or volume forecasts are needed, it may be necessary to switch to a different rainfall/runoff Operation when implementing ESP. In other cases the model may be suitable for low flow or volume forecasts, but it was not calibrated with these forecasts in mind. It may be necessary to re-calibrate some basins before using ESP.

- o Special Forecast Groups can be used to group Segments that have a common ESP application and should be executed together.

<u>ESPINIT Considerations</u>

The purpose and capabilities of the ESPINIT program were described briefly above. Detailed input summary documentation for the program is given in Section VI.3.5. Specific considerations for use of the program are given below.

- o ESPINIT has an option to copy time series from the calibration data files to the ESP historical data files. The ESP program has the ability to use time series directly from the calibration data files, but the ESP historical files are more efficient. Time series are stored in the ESP historical files with a mixed data type, such that all the time series needed for a Segment are stored together month by month in the order they are needed in the program. The time series are read in the order in which they were specified in the DEF-TS section of FCINIT. Storing the time series in this order in the historical data files will keep the disk accesses to a minimum.

- o When a Segment is defined with ESPINIT careful thought should be given to each time series. If a time series that has been defined in the Forecast Component is not specified in the DEF-TS section of the DEFSEG command in ESPINIT, it will be assigned defaults. In most cases input and output time series must be specified so that new external location information can be defined. See Section VI.3.5-DEFSEG-DEF-TS for detailed information on how to define time series with ESPINIT.

- o Input time series will usually be found on the ESP historical files and output time series will usually be written to the ESP time series files. If a time series has been specified with an external location on the ESP time series files at initialization time, it can be written to either the ESP scratch time series file or an ESP permanent time series file using a run-time option. For the special case of update time series, ESPINIT allows the output location to be different from the input location, since the ESP historical files cannot be overwritten.

- If potential evaporation data are being used in the daily forecast program, then it is also required in ESP. If the data are not available historically, ESP can generate a time series from average monthly values. See Section VI.3.5-DEFSEG-DEF-TS.

- If future MAP and MAT time series are available on the processed data base, they can be blended with the historical data before they are input to ESP. The user specifies a weight to be assigned to the future values at the beginning of the weighting period, a weight to be assigned to the future values at the end of the weighting period, a weighting period length, and a blending period length. See Section VI.3.5-DEFSEG-DEF-TS.

- ESP analyzes observed streamflow data so that it can be compared to an analysis of the simulated streamflow. In general, the observed data available operationally are instantaneous, while the observed data available historically are mean daily. Since no new time series can be added when the ESP Segment is defined, ESPINIT has the capability to replace an instantaneous flow time series needed for the daily forecast program with a mean daily flow time series needed for ESP. This substitution has no effect on the daily forecast program. See Section VI.3.5-DEFSEG-DEF-TS.

- ESP output time series are usually written to the ESP scratch file, so that they can be used as input time series to a downstream Segment. In large ESP runs, the ESP scratch file could become full. When a time series is defined in ESPINIT as input from the ESP time series file, it can be marked for deletion if it is not needed in a later Segment. When ESP is executed, the space occupied by the time series on the ESP scratch time series file is freed after the time series is read for the Segment. See Section VI.3.5-DEFSEG-DEF-TS.

- It is important to define all of the desired output variables when an ESP Segment is defined. Output variables can be turned off at execution time, but no new output variables can be added without redefining the ESP Segment. See Section VI.3.5-DEFSEG-ANALYSIS.

- All of the display options should also be specified when the ESP Segment is defined. Plots and displays can be turned OFF, but they cannot be defined with the execution program.

- The user can select from three frequency distributions: empirical, normal, and log-normal. The empirical distribution ranks the values and calculates the probabilities as m/(n+1), where m is the rank and n is the number of values. If the proper distribution for a particular output variable is not known, either the empirical or the normal distribution should be used.

- o Streamflow is the most commonly analyzed type of time series in ESP, but ESP has the ability to analyze other types of data. Other data types which might be of interest in an ESP analysis include: river stage, soil moisture, snow water equivalent, etc. Any time series specified in the ANALYSIS section must be defined first in the DEF-TS section.

<u>ESP Considerations</u>

The execution portion of ESP is a Function contained in the FCST program. Section VI.2.3B describes the FCST program and gives considerations for its use. Specific input summary type documentation for the ESP Function can be found in Section VI.5.3C-ESP. The ESP Techniques are documented with the other FCST program Techniques in Section VI.5.3D.

Some general considerations for ESP execution are:

- o Special Forecast Groups can be used to group Segments in a manner that is appropriate for ESP execution. This grouping can be done on the basis of application, e.g., headwaters for water supply, or it may be done on the basis of required computer resources. It is probably best to break extremely large ESP runs into several pieces, so that if the system went down the entire run would not be lost. Intermediate time series that are needed for a downstream Segment in a later run could be written to the ESP permanent time series files. The ESP permanent time series files can be used as input to downstream Segments or as input to external programs for further analysis.

- o The historical simulation option is controlled with the HISTSIM Technique. When this Technique is ON, the ESP program simulates the entire historical period of record continuously without resetting the initial conditions. The analysis of the historical simulation can be compared to the analysis of the observed flows to assess any bias that might exist in the system. This information can be used to subjectively adjust the conditional simulation before a forecast is made.

- o The adjust simulation option is controlled with the ADJSIM Technique. The adjust simulation option works similarly to the historical simulation option, except that the simulated flow is adjusted to match the observed flow whenever there is an adjust Operation. The purpose of this option is to produce a time series as close to an observed time series as possible where an observed time series does not exist. This option was inserted on a test basis. It should be used with caution since it is very costly in terms of run-time, disk space, and array sizes. It can also produce a lot of warnings indicating that the adjust routine did not converge.

- o The base period option was included so that an analysis could be done on the observed data for a historical period which is different from the run period. The purpose of this option is to make it convenient for forecasts that must be expressed as a percent of normal for a given base period. The base period option is controlled with the Technique BASEPER.

- The output from the Summary table display is relatively straightforward. A -999.0 is used to indicate missing values. A 9999.0 is used with the NDTO output variable to indicate that a time series never got above/below the criteria.

- The output from the Frequency display will depend partially on how the Segment was defined with ESPINIT. The probability levels to be displayed, the option to produce an exceedance probability plot, and the option to include sample points on the plot are defined at initialization time.

- The End of the Run Summary table was included to summarize the results of a run in a format similar to the one used in the publication, 'Water Supply Outlook for the Western United States'. Only the output variables designated when the Segment was defined with ESPINIT are shown in the table.

- ESP has the capability to weight historical years differently in the analysis. The simulations proceed as normal, but weights are applied to the output variable values accumulated for each year. Initially, this option will probably be limited to excluding historical years that are considered to be totally unrepresentative of the current year, since no objective procedure is presently available for determining the weights. Historical years can be excluded from the analysis by assigning them zero weights. The mean and standard deviation shown in the Summary Table display will reflect the assigned weights. The values calculated in the Frequency display for the different exceedance probability levels will also reflect the assigned weights if the normal or log-normal distribution was selected. The plotting position of the sample points will not be effected by the weights. The year weighting option will not have any effect on the Frequency display if the empirical distribution was selected. The year weighting option is controlled with the YRWEIGHT Technique. This Technique allows the user to specify the weights at run time. See Section VI.5.3D-TECH-YRWEIGHT.

- The blending Techniques for precipitation and temperature are used to override the blending parameters specified at initialization time. The changes apply to all temperature or precipitation time series within a given Segment. See Sections VI.5.3D-TECH-BLENPREC and VI.5.3D-TECH-BLENTEMP.

- The criteria Technique is used to change the criteria level for the output variables NDTO and NDIS. The change applies to all output variables with the specified type and KODE within a given Segment. See Section VI.5.3D-TECH-CRITERIA.

Most of the Techniques for the ESP Function are related to turning options, output variables or displays on or off. See Section VI.5.3C-ESP for a complete list of the available Techniques for ESP and Section VI.5.3D for a description of these Techniques.