

VI.2.2C OPERATIONAL FORECAST SYSTEM PREPROCESSOR DEFINITIONS USING PROGRAM PPINIT

Introduction

The Preprocessor Initialization (PPINIT) program (Section VI.3.3) is used to provide all preprocessor parametric input. The primary functions of PPINIT are:

- o Define all general user parameters for the preprocessors (Section VI.3.3B-DEFINE-USER).
- o Define the entire station network including allocating space on the PPDB for each station, making station entries into the data entry control files, checking and storing all parametric data for each station in the PPPDB and allocating time series space on the PDB for all stations with RRS data (Section VI.3.3B-DEFINE-STATION).
- o Compute estimator stations and weights and alphabetically order the stations after the entire network is defined (Section VI.3.3B-NETWORK).
- o Define basin boundaries by latitude and longitude points if this option is selected by the user (Section VI.3.3B-DEFINE-BASIN).
- o Define FMAP, MAP, MAPX, MAT and MAPE areas including computation of station weights if not specified by the user, checking and storing parametric data on the PPPDB and allocating time series space on the PDB (Section VI.3.3B-DEFINE-AREA).
- o Compute the computational order for MAP and FMAP areas based on the upstream to downstream order that those time series are used in the Forecast Component (Section VI.3.3B-ORDER).
- o Perform a variety of maintenance items after the initial definitions are complete such as:
 - define new stations, make changes to existing stations or delete stations
 - re-alphabetize the station network due to changes
 - re-compute estimators for all stations and station weights for all areas where the weights are not user specified to incorporate new stations and changes to existing stations into the preprocessor computations
 - change basin boundaries if needed, define new basin boundaries or delete basins
 - define new areas, change existing areas or delete areas
 - re-compute station weights for all areas where the weights are not user specified to incorporate new basins or areas or changes to existing ones
 - re-compute the computational order for MAP, MAPX and FMAP based on the addition or removal of areas or changes to how the areas are used in the Forecast Component

- o Display all information that can be defined by PPINIT including the current status of all preprocessor related data bases and files.

Inclusion Rule

The basic rule that is followed in PPINIT for adding new stations or incorporating changes to existing stations or basin boundaries is that the addition or change does not affect the preprocessors until the NETWORK command is run. Thus, the user can change several stations or basin boundaries or add several new stations over some period of time and still control when the changes go into effect by when the NETWORK command is run. The exception to this rule is that when stations are deleted or certain data types are deleted from a station definition, the NETWORK command is run automatically right after the deletions. All previous changes or additions will also be incorporated during this automatic execution of the NETWORK command.

Table 1 in Section VI.3.3B-NETWORK documents the user actions that will trigger certain parts of the NETWORK command to be executed the next time the command is run. The table also specifies which actions cause an automatic execution of NETWORK.

PPINIT Considerations

Following is a list explaining what occurs in the various PPINIT commands and some factors the user should consider when using PPINIT.

- o DEFINE USER - Defines various general user parameters for the preprocessors, some limits for checking values, MDR display information and default PPDB values for RRS data.
 - Even if OFS is only being applied to a part of the user area, it is recommended that the MDR subset be defined to include the entire area. Currently there is no provision to redefine the MDR subset and adjust the existing data on the PPDB.
 - The default minimum days of observed data to be retained and the typical number of observations to be stored for each RRS data type on the PPDB should be selected carefully. Remember these are default values for typical stations. The values can be increased (or decreased) for individual stations. The typical number of observations is the expected total to be stored. For data types that can contain future (projected) data, this is the total of the typical number of observations during the minimum observed days to be retained plus the typical number of future data values. It is probably wise to be a little conservative in specifying the number of observations to be stored to minimize the use of the free pool, but too much extra space will result in increased disk accesses and disk storage requirements.

- o DEFINE STATION - Defines general station parameters, specifies what types of data are available for each station and defines parameters needed for each data type.
 - Remember to use the same identifier for the station as will be used in SHEF reports, normally the NWS Handbook No. 5 identifier.
 - Stations somewhat beyond the user borders should be included to provide adequate estimates of stations near the edge of the user area that are used in MAP and MAT areas.
 - There is no checking of automatically transferred data from the SASM or GOES systems. If there are questions about the quality of these data for a given station, it is better to monitor the station first before having the data transferred to the PPDB. The HADS system can be used to monitor GOES stations, while a second user identifier (non NWSRFS) could be established with the synoptic utility program to monitor synoptic data by obtaining printer reports without transferring the data to the PPDB.
 - Precipitation characteristics are assigned on a station-by-station basis. When used in MAP the following rule applies: if a station with or without characteristics is being estimated by one that has the opposite, the two stations are assumed to have the same characteristic value. Warnings are generated in PPINIT by the NETWORK command when this situation could occur. In general, all stations in areas where the station means cannot be considered equal should use characteristics plus all stations that can be used to estimate missing values at these stations or can have missing values estimated from these stations should have characteristics provided.
 - Significance weights should be used in mountainous areas or at other locations where distance alone does not provide the best long-term estimate of missing data.
- o NETWORK - The explanation and associated table in Section VI.3.3B-NETWORK describes how the NETWORK command works.
- o DEFINE BASIN - Defines basin boundaries for use with MAP, MAPX and/or MAT areas. The latitude and longitude points used to describe the basin boundary are converted to the NWSRFS/HRAP grid system for use in determining station weights (Section VI.3.3C).
 - It is recommended that basin boundaries be used to compute station weights in all areas where long-term station means are approximately equal. In mountainous areas, predetermined station weights should be used. The biggest advantage of using basin boundaries wherever they can logically be used, is that station weights can be automatically computed when the MAP or MAT areas are defined and will be automatically recomputed when there is a change to the station network.

Another advantage is that various utility uses of basin boundaries, such as preparing map backgrounds, can be developed once the basin boundary values are available. Also, basin definitions are used to provide the mechanism for integrating HRAP data. It is recognized that it may take some time for a user to obtain the input data needed to define basin boundaries if they are not currently available, thus, an interim method of automatically generating station weights is provided. This is the $1/d^{**power}$ procedure that computes weights based on the distance between a station and the centroid of an area.

- o DEFINE AREA - Defines areas and associated parameters for Future MAP, MAP, MAPX, MAT and MAPE areas.
 - Future MAP areas must be defined before MAP areas. As discussed in Section VI.2.1, an FMAP area should probably encompass several MAP areas since the data are manually input to the FMAP Preprocessor.
 - Predetermined weights should generally be used in mountainous areas, whereas weights that can be computed automatically should generally be used everywhere else. If basin boundaries are available, grid point or in the case of MAP, Thiessen weights are used. If basin boundaries are not provided, $1/d^{**power}$ weights can be used and still provide the capability to automatically compute weights.
 - As with stations, run-time options control the actual use of MDR data. However, for areas the run-time options are non-universal techniques, i.e., MDR data can be used for some areas and not for others. If basin boundaries are provided, they are used to determine the MDR boxes assigned to the area as discussed in Section VI.3.3C. If actual basin boundaries are not provided, a crude four-point definition of the area is used to assign the MDR boxes.
- o ORDER - Determines the computational order of MAP, MAPX and FMAP areas based on the order they are used in the Forecast Component. The computational order information is needed so that MAP can be run for just parts of the user area, for use in specifying run-time options for areas and for specifying input to FMAP. See Section VI.3.3B-ORDER for more information.
- o RE-DEFINITIONS - PPINIT allows some station, basin boundary, area and general user information to be redefined using the appropriate DEFINE command with a disposition of OLD. Checks are made to make sure that all the necessary values in the files are changed and NETWORK flags are set if needed. Certain changes will cause the NETWORK command to be automatically run as specified in Table 1 of Section VI.3.3B-NETWORK.
- o DELETE - This command is used to remove stations that no longer exist or are no longer used station data types that are no

longer reported or used or unneeded area and basin boundary definitions. All user specified references to an item must be removed before the item can be deleted. The NETWORK command is automatically run after the delete to remove all other references to the item in the OFS files.

- o DUMP - This command is used to display the information that has been defined by PPINIT. The PUNCH option can be useful when doing a redefine by retrieving the current definition, which then can be modified as needed. The DUMP command is described in Section VI.3.3B-DUMP.
- o STATUS - This command prints the current status of all data bases and files used by PPINIT. The STATUS command as described in Section VI.3.3B-STATUS should be used frequently during the initialization process to monitor file usage.