VI.3.3B-DEFINE-STATION PROGRAM PPINIT COMMAND DEFINE STATION

Purpose

Command DEFINE STATION is used to define or redefine station parameters.

The following parameter groups can be defined for a station:

- o General (STAN)
- o Precipitation (PCPN)
- o Temperature (TEMP)
- o Potential evaporation (PE)
- o River, Reservoir and Snow (RRS)

A station will always have general parameters defined and must have at least one of the other parameter groups.

When a station is redefined:

- any information in the old definition that is not specified in the new definition will be deleted; information that is deleted includes:
 - o Data Entry Source Code entries in control files
 - o data groups (PCPN, TEMP, PE and RRS)
 - o precipitation characteristics
 - o RRS time series
- any time series with a data time interval that is compatible with the old definition will copied to a new time series
- any time series with a data time interval that is not compatible with the old definition will be deleted and recreated

Input Data

The general format of the DEFINE command is described on page VI.3.3B-DEFINE. The input described here immediately follows the DEFINE command.

Card	<u>Field</u>	Required/ Optional 1/	Format	Description
1	1	R	A or A(A)	'STAN' or 'STAN(units)' <u>2</u> /
	2	R	A8	Station identifier (same identifier as used for the station in SHEF messages) <u>3</u> /
	3	R #	A20	Descriptive information
	4	R #	A2	State designator (Postal Service 2-character identifier)
	5	R #	R, I or	Latitude <u>4</u> /

Card	Field	Required/ Optional 1/	Format	Description
			I-I	
	6	R #	R, I or I-I	Longitude $\underline{4}/$
	7	R #	R	Elevation in FT or M $5/$
	7+	0 #	I	Station number: <u>6</u> / 'NUMBER(number)'
	7+	0 #	A({A4,A8})	Data Entry Source Code and external identifier: <u>7</u> / 'SOURCE({code,identifier})'
	7+	0 #	A({A4})	Data groups or data types not to be transferred: <u>7</u> / 'NOGOESTR({code})'
	7+	0 #	A({A4})	CDAS data groups or data types not to be transferred: <u>7</u> / 'NOCDASTR({code}'
	7+	0 #	A(I)	Grid point address: <u>8</u> / 'GPA(address)'

The following input is used to define PCPN parameters for the station:

2	1	R	A	'PCPN'
	2	R #	I	Reporting time interval <u>9</u> /
	3	R #	A	<pre>Processing code: 10/ 'NORM' = process normally (default) 'ZERO' = set precipitation to zero if observation is missing 'SYN' = synthetic station</pre>
	4	R #	A	MDR usage option: <u>11</u> / 'MDR' = can use MDR (default) 'NMDR' = can not use MDR
	5	R #	A or A({A8,R})	<pre>Type of station weights: 12/ 'D2' = 1/d**2 (default) 'SIG({id,wt})'</pre>
	5+	0 #	A({R})	Correction factors: $13/$

Card	Field	Required/ Optional 1/	<u>Format</u>	Description
				'CF(n*R)'
	5+	0 #	A(12*R)	Station characteristics: <u>14</u> / 'CHAR(12*R)'
	5+	0 #	A	<pre>Indicator whether station is to be used only to estimate missing data and should not receive any weight in an MAP area: 'NWGT' = do not use as weighting station in MAP areas 'WGT' = use as weighting station in MAP</pre>
				areas (default)
mbo f	allouing	, innut in une	d to dofta	a TEMP parameters for the

The following input is used to define TEMP parameters for the station:

3	1	R	A	'TEMP'
	2	R #	A(12*R)	Mean monthly maximum tempera- tures in DEGF or DEGC: <u>15</u> / 'MX(12*R)'
	3	r#	A(12*R)	Mean monthly minimum tempera- ture in DEGF or DEGC: <u>15</u> / 'MN(12*R)'
	4	r #	A	<pre>Type of data observed: <u>16</u>/ 'MXMN' = maximum/minimum</pre>
	5	R #	I	Time interval of instantaneous data (default is 6 hours; specified only if type of data observed is 'INST' or 'BOTH') <u>17</u> /
	5+	0 #	A(2*R)	Correction factors in DEGF or DEGC: <u>18</u> / 'CF(2*R)'
	5+	0 #	A	Indicator whether forecast maximum/minimum data are available: 'FMM' = data available

Card	Field	Required/ Optional 1/	Format	Description
				'NFMM' = data not available (default)
	5+	O #	A (R)	Elevation weighting factor (Fe - see note for description of units): <u>19</u> / 'FE(R)'

The following input is used to define PE parameters for the station:

4	1	R	A	'PE'
	2	R #	R	Anemometer height in FT or M <u>20</u> /
	3	R #	R	B3 parameter <u>21</u> /
	4	R #	A	<pre>Primary type of radiation data to be used: <u>22</u>/ 'SKY' = sky cover</pre>
	4+	0 #	A(R)	Correction factor: <u>23</u> / 'CF(R)'

The following input is used to define RRS parameters for the station:

5	1	R	A	'RRS'
	2	R	A4	Data type code $24/$
	3	R #	I	Time series data time interval (if negative then a time series will not be created) <u>25</u> /
	4	R #	I	Minimum days of observed data to be retained on the Preprocessor Data Base <u>26</u> /
	5	r #	I	Typical number of observations to be held on the Preprocessor Data Base <u>26</u> /
	Fields 6 t allowed. <u>2</u>	chrough 8 are <u>26</u> /	e require	ed only if missing data is not
	6	0 #	A(A)	Missing data option: <u>27</u> /

Card	Field	Required/ <u>Optional 1/</u>	Format	Description
				'MSNG(YES)' or 'MSNG(NO)' (default is MSNG(YES))
	7	r#	A	<pre>Interpolation option: 28/ 'LIN' = use linear</pre>
	8	R #	R	Extrapolation recession constant <u>29</u> /
	Fields discha	9 and 10 are rge data. <u>30</u> /	allowed	only if the station has mean
	9	0 #	A (R)	Minimum discharge below which fraction of flow typically occurring during each hour of the hydrologic day will be used to distribute flows input using the following format: MIN(min)
	10	0 #	A(24*R)	Fraction of flow typically occurring during each hour of the hydrologic day (starting with hour 1) input using the following format: DIST(24*R)

Repeat fields 2 through 10 to define additional RRS data types.

Repeat cards 1 through 5 to process additional stations.

Notes:

- $\underline{1}/$ A $^{\prime}\#^{\prime}$ indicates that the parameter value can be changed when a station is being redefined.
- $\underline{2}/$ The units in which the data are being input for the entire station can be specified in this field.

If the parameters for the station are being input in English units then the unit code ENGL can be entered. If the parameters for the station are being input in metric units then the unit code METR must be entered.

The default unit code is ENGL.

 $\underline{3}$ / The identifier is used to identify the station throughout the Preprocessor Component. It is also the identifier that will be used

in most cases to identify the station when writing to or reading from the Preprocessor Data Base (station number is also allowed for writing). This identifier should be the identifier that will be used for the station in any SHEF (Standard Hydrologic Exchange Format) messages. If this is not true then observed or edited data in SHEF format cannot be input for the station or SHEF message data could be written for the wrong station. It is important that the NWSRFS identifier for a station be the same as the identifier that will be used for that station in SHEF messages.

The station identifier cannot begin with the letters W, X, Y or Z and be followed by all numbers. This is the format used by SHEF for stranger reports.

<u>4</u>/ The station location can be input in decimal degrees or degrees and minutes. For example specifying 50.5 indicates decimal degrees and is equivalent to 50-30 or 5030 which indicates degrees and minutes. Valid values for minutes are 0 to 59.

If degrees and minutes are input and the optional '-' is omitted then two characters must always be used to specify the minutes. For example 5003 would be a valid input field to specify 50 degrees and 3 minutes.

- <u>5</u>/ If 'STAN(ENGL)' has been specified then the units of the elevation must be feet (FT). If 'STAN(METR) has been specified then the units of the elevation must be meters (M).
- <u>6</u>/ The station number can be used for manual data entry. Data can be written to the Preprocessor Data Base by specifying either the station identifier or the station number. The station number is never used to identify the station for Preprocessor parametric input.

The station number can be any integer number greater than zero and must be unique.

<u>7</u>/ The Data Entry Source Code is used to make entries in control files which are used to transfer observed and forecast data from other data bases. These control files are used by the transfer programs to determine which data are to be transferred to the user's Preprocessor Data Base.

The current valid data entry source codes and the corresponding data source are:

SA	-	Synoptic Airways data from NMC files
SM	_	Synoptic Meteorological data from NMC files
GHB5	_	Geostationary Orbit Earth Satellite data specified by
		NWS Communications Handbook 5 identifier
GPLT	-	Geostationary Orbit Earth Satellite data specified by
		platform identifier
CDAS	_	Centralized Automatic Data Acquisition System (CADAS)

If a code of 'SA' or 'SM' is entered then the SA or SM station identifier is entered as the next field in the parenthesis. For example,

SOURCE (SA, SAID)

would define 'SAID' as the SA identifier for the station.

If a code of 'GPLT' is entered then the 8 character platform identifier is the next field in the parenthesis. For example,

SOURCE (GPLT, PLTFRMID)

would define 'PLTFRMID' as the GOES platform identifier for the station.

If the code 'GHB5' is entered then the 5 character Handbook 5 station identifier is the next field in the parenthesis. For example,

SOURCE (GHB5, HB5ID)

would define 'HB5ID' as the GOES Handbook 5 identifier for the station.

If the code 'CDAS' is entered then the 5 character CDAS station identifier is the next field in the parenthesis. For example,

SOURCE (CDAS, CDSID)

would define 'CDSID' as the CADAS identifier for the station.

If 'SAME' is specified as the data entry source code then the station identifier will be used as the source code identifier.

More than one code can be entered for a station. For example,

SOURCE (SA, SAID, SM, SMID)

would specify that the station has both SA and SM data.

The GHB5 and GPLT Source codes can be specified more than once, each with different identifiers, for a single station. If the other source codes are entered more than once then only the last is used and a warning is printed. A reason for using multiple GHB5 or GPLT source codes for a single station is that multiple DCP's may be used to collect data for what is considered in hydrologic terms to be a single station. An example is a reservoir where one DCP may be used to record pool elevation and another to record tailwater elevation below the dam.

The 'NOGOESTR' and 'NOCDASTR' options are used to specify which data groups (PCPN, TEMP, PE or RRS) or RRS data types are not to be automatically transferred to the Preprocessor Data Base. This option is useful for stations that report data from more than one

source (GOES, CADAS, etc.) and the user does not want certain data to be automatically transferred. For example,

NOGOESTR (RRS)

would specify that GOES RRS data are not to be transferred for the station.

Selected RRS data types can be specified not to be transferred by inputting the data type code(s). For example,

NOGOESTR (PELV, TWEL)

would specify that the CDAS RRS data types PELV and TWEL are not to be transferred.

 $\underline{8}/$ The grid point address is needed when the WGRFC MARO preprocessor function is to be used.

The address is needed for all stations that fall within a defined grid box (see DEFINE GRID). The address is a three or four digit number where the first one or two digits are the degree box number (one degree latitude by one degree longitude is a degree box), the next digit indicates tenths of a degree north of the SE corner of the box and the last digit indicates tenths of a degree west of the SE corner.

<u>9</u>/ The reporting time interval is how often the station reports observed data values.

Valid values are 1, 3, 6 or 24 hours. Space is always allocated for the 24 hour accumulated value. The default is 24 hours.

Even though the MAP Preprocessor only works with 6 hour time intervals, 1 and 3 hour data can be stored in the Preprocessor Data Base. The MAP Preprocessor converts any 1 and 3 hour data into 6 hour summations before processing the values.

 $\underline{10}/$ The processing code to used by the MAP Preprocessor to indicate how the data observed by the station are to be processed.

If 'NORM' is specified then any missing observations will be estimated from the surrounding stations or from MDR data.

If 'ZERO' is specified then any missing observation will be set to zero.

If 'SYN' is specified then it indicates a synthetic station that has no observed data and the data time interval must be 24 hours.

The default is 'NORM'.

 $\underline{11}/$ The MDR usage indicator specifies whether MDR data can ever be used to estimate missing data values. The option for actually using MDR data to estimate missing data is specified by the user

when the MAP Preprocessor is run. The default is 'MDR'.

<u>12</u>/ The type of station weights specifies whether 1/d**2 or significance weights are to be used to estimate missing 24 hours precipitation data.

Significance weights are mainly used in mountainous areas. The method used is based on the Colorado River Basin River Forecast Center's 'high elevation precipitation estimation' technique.

If 'SIG' is specified then the identifiers and weights of the stations to be used must be input. The list of identifiers and weights must be enclosed in parenthesis. The weights can be any non-zero positive number. The maximum number of stations with significance weights is 10.

The default is $1/d^{**2}$.

 $\underline{13}$ / The observed precipitation is multiplied by the PCPN correction factor to increase or decrease the observed precipitation.

One or two correction factors can be specified. If one correction factor is specified it is used when processing data for any month of the year. If there are two factors then the first is used when processing winter season data and the second is used when processing summer season data.

The months that begin the summer and winter season can be specified when the USER parameter group is defined using the DEFINE command.

The default beginning month for the summer season is 5 (May) and for the winter season is 10 (October).

Correction factors are rounded to the nearest .05 when used. The allowable range is 0.5-4.95.

The default correction factor is 1.0.

<u>14</u>/ Station characteristics are used to determine the long term relationships between precipitation stations and are defined as mean monthly precipitation values.

When estimating missing data the ratio of the characteristics between the station being estimated and each of the estimators is used to account for the long term difference in precipitation at each site.

The use of characteristics can result in an estimated value at a station being larger or smaller than the precipitation at any of the stations being used as estimators. Characteristics are used in the estimation of missing station data whenever the station being estimated has characteristics defined. If some of the surrounding stations that are being used as estimators do not have characteristics defined then a characteristic value equal to the characteristic of the station being estimated is assigned. In this case the two stations are assumed to have the same long term precipitation amounts.

Characteristics should be used for all stations in areas where long term means vary from site to site and for all stations that can be used to estimate missing data in these areas.

A warning message will be printed during execution of the NETWORK command to indicate when a station with characteristics defined can be estimated from stations that do not have characteristics defined and vice-versa.

There is one characteristic value for each month starting with January.

If 'STAN(ENGL)' has been specified then the units of the station characteristics must be inches. If 'STAN(METR)' has been specified then the units of the station characteristics must be millimeters.

- <u>15</u>/ One value must be specified for each month starting with January. If 'STAN(ENGL)' has been specified then the units of the mean temperatures must be degrees Fahrenheit (DEGF). If 'STAN(METR)' has been specified then the units of the mean temperatures must be degrees Centigrade (DEGC).
- 16/ The type of data observed by the station specifies the data that will be stored in the PPDB and processed by the MAT Preprocessor.

If 'SYN' is specified then it indicates a synthetic station that has no observed data and the station cannot have Forecast maximum/minimum data available.

The default is 'MXMN'.

17/ The reporting time interval of a station with instantaneous temperature can be 1, 3 or 6 hours. The default is 6 hours.

Even though the MAT Preprocessor generates only 6 hour time series it can use 1 and 3 hour instantaneous temperature data.

 $\underline{18}/$ The correction factor is added to the observed data to increase or decrease the value.

Two correction factors must be specified. The first is used to correct the values of maximum temperature data and the second to correct values of minimum temperature data.

The default value for both correction factors is 0.

If 'STAN(ENGL)' has been specified then the units of the correction factors must be degrees Fahrenheit (DEGF). If 'STAN(METR)' has been specified then the units of the correction factors must be degrees Centigrade (DEGC).

<u>19</u>/ The elevation weighting factor (Fe) is used to account for the differences in elevation in addition to the distance between stations when estimating missing temperature data.

A larger Fe will give more weight to stations with the smallest difference in elevation and less weight to stations with the largest difference in elevation. The Fe factor is normally used in mountainous areas. The default is 0.

If 'STAN(ENGL)' has been specified then the units for Fe are:

MI between stations 1000 FT change in elevation

If 'STAN(METR)' has been specified then the units for Fe are:

KM between station 1000 M change in elevation

A reasonable value for Fe is 20 when English units are used and 100 when using metric units.

- <u>20</u>/ If 'STAN(ENGL)' has been specified then the units of the anemometer height must be feet (FT). If 'STAN(METR)' has been specified then the units of the anemometer height must be meters (M).
- <u>21</u>/ The B3 constant can be obtained from Figure 1. This map was published in Water Resources Research Volume 12, Number 5 (October 1976, page 860).
- <u>22</u>/ The primary type of radiation data specifies the type of data to be used to compute PE data for the station. If 'RAD' or 'SUN' are specified as primary and are not available then 'SKY' is used as the estimate of radiation.
- $\underline{23}$ / The computed PE data is multiplied by the PE correction factor to increase or decrease the PE value. The default is 1.
- <u>24</u>/ The following is a list of the available RRS data types with the compatible SHEF physical element code (if one exists) and description:

RRS	SHEF	
Code	PE Code	Description
Instantar	neous Height	:
STG	HG	River stage observed
PELV	HP	Pool elevation observed
LELV	HL	Lake elevation observed
LAKH	HK	Lake height above specified datum
TWEL	HT	Tailwater stage observed
TWSW	HW	Spillway tailwater stage observed
FBEL	HF	Forebay elevation observed

RRS Code	SHEF PE Code	Description
Insta	ntaneous Discha	arge:
QIN	QR	River discharge observed
RQOT	QT	Reservoir outflow observed and projected
DQIN	QD	Diversion discharge observed and projected
RQIN	QI	Reservoir inflow observed
RQSW	QS	Reservoir discharge spillway observed
PCFD		Diverted flow from channel percent
Mean	Discharge:	
QME	QR	River discharge observed
DQME	QD	Diversion discharge observed and projected
RQME	QT	Reservoir outflow observed and projected
RQIM	QI	Reservoir inflow observed
RQGM	QG	Reservoir discharge power generation observed
Insta	ntaneous Other	Reservoir Data:
RSTO	LS	Reservoir lake storage observed
NFBD		Flash boards down observed and projected
GATE		Gate openings observed and projected
GTCS		Gate control switch
Insta	ntaneous Snow a	and Miscellaneous Data:
AESC	SA	Snow cover areal extent observed
FGDP	GD	Frost depth observed
ICET	IT	Ice thickness observed
SNWE	SW	Snow cover water-equivalent observed
SNOG	SD	Snow cover depth observed
ZELV	ΗZ	Freezing level
Mean	Snow Data:	
PTPS		Fraction snowfall

25/ The time series data time interval specifies the number of hours between the time series values written to the Processed Data Base. Valid values are 1, 2, 3, 4, 6, 8, 12 and 24.

If a negative time series data time interval is specified then a time series will not be created in the Processed Data Base and the RRS Preprocessor will not create a time series from the observed data in the Preprocessor Data Base. This allows the data type to be defined in the Preprocessor Data Base so that data values can be stored and displayed with the Preprocessor Data Base Utility Program (see section VI.6.3) but no space is allocated in the Processed Data Base.

When a station is redefined any time series with a data time interval that is different than the old definition will be deleted and recreated.

<u>26</u>/ The minimum number of days of observed data to be held on the Preprocessor Data Base (PPDB) defines the observed period to be

retained even if more than the typical number of observations specified is exceeded.

The typical number of observations is used to size the primary storage area. For types that can contain future data the typical number of both observed and future data values should be used. If this space is too small to store all the observations for the minimum number of days specified then an overflow area will be used to store the data. It is important to pick a reasonable estimate for the typical number of values to be observed during the period to be retained in order to minimize the use of the overflow area but also not allocate a lot of space that will never be used.

The minimum days to be retained and the typical number of observations must be specified if the default values have not been specified with the DEFINE USER command.

27/ The optional field MSNG specifies whether the time series generated from the RRS input data can contain missing values.

MSNG(YES) indicates missing values are allowed in the time series and is the default value.

MSNG(NO) indicates no missing values are allowed and can only be specified for certain data types.

In almost all cases the MSNG(YES) option should be used since the normal place to fill in missing values and produce an updated or adjusted time series is the Forecast Component and not the RRS Preprocessor. In a few cases, such as an upstream boundary point, a time series with no missing values is needed. In these cases the MSNG(NO) option is used and the output data type will be different from the input data type (all input types allow missing data).

The input RRS data types for which MSNG(NO) is allowed and the data type assigned to the generated time series for each are as follows:

RRS input	RRS output data
<u>data type</u>	type if MSNG=NO
DQIN	DQIE
RQOT	RQIE
DQME	DQMP
RQME	RQMP

<u>28</u>/ The interpolation option is used when the RRS Preprocessor is processing a data type for which missing data is not allowed.

If the 'LIN' option is chosen then linear interpolation will be used to estimate missing values.

If the 'RET' option is chosen then the preceding observed value will be retained until another observed value is found. There is

no default value.

29/ The extrapolation recession constant is used when the RRS Preprocessor is processing a data type for which missing data are not allowed.

The recession constant is applied to the last observed value to estimate missing values for the rest of the period. The value of the recession constant must be greater than 0 and less than or equal to 1. A value of 1 will retain last observed value for the rest of the period to be processed. There is no default value.

The recession amount specified is per 24 hour period. Extrapolated values are computed according to the following equation:

 $X_{t+1} = X_t * R**dt/24$

where X_t = data value at time interval dt = data time interval R = recession constant

<u>30</u>/ The option to distribute flows can only be used for mean discharge data types.

If this option is specified then the time interval of the time series is less than the time interval of the observed data and the flow is less than the minimum specified then the flows will be distributed using the specified distribution when the time series is created. For example the option would be used for a station that reports reservoir outflows every 24 hours, has daily fluctuations due to power generation and will have the data written to a 6 hour time series. The distribution specified would be based on the typical average hourly flow observed when mean discharges are below the specified minimum value. The sum of the fraction of flow typically occurring during each hour of the day should be 1.0. If not then the values will be normalized.

Sample Input

To define a new station that has the following parameters:

General Parameters (STAN): Identifier = DCA Description = WASHINGTON NATL APT State = DC Latitude = 38.75 Longitude = 77.0 Elevation = 200. feet Station number = 1 Data Entry Source Codes: CODE=SA, EXTERNAL ID=DCA Precipitation Parameters (PCPN): Reporting time interval = 6 hours

Processing code = NORM (normal) MDR usage option = MDR (can use MDR) Type of station weights = D2 $(1/d^{*}2)$ Correction factor = 1.1Temperature Parameters (TEMP): Elevation weighting factor = null-field (non-mountainous) Mean monthly temperature (Fahrenheit): JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC 75. MX 35. 45. 55. 65. 85. 85. 75. 65. 55. 45. 35. 65. 75. MN 25. 35. 45. 55. 75. 65. 55. 45. 35. 25. Correction factors: MAX = 0.5 MIN = -1. Forecast maximum/minimum indicator = NFMM (none available) Potential Evaporation Parameters (PE): Anemometer height = 50. feet B3 parameter = .25Primary type of radiation data to be used = SUN (percent sunshine) Correction factor = 1.05River Reservoir and Snow Parameters (RRS) - define 2 data types: Data Time Interp. Extrap. Min Typical Type Interval Option Constant Days Obs DQIN 6 LIN 0.9 10 40 QME 24 (missing allowed) (obtain from URRS default) The following input would be used: **@DEFINE STATION** \$ GENERAL PARAMETERS STAN DCA 'WASHINGTON NAT APT' DC 38.75 77.0 200. NUMBER(1) SOURCE(SA, DCA) \$ PRECIPITATION PARAMETERS PCPN 6 NORM MDR D2 CF(1.1) **\$** TEMPERATURE PARAMETERS TEMP MX(35. 45. 55. 65. 75. 85. 85. 75. 65. 55. 45. 35.) MN(25. 35. 45. 55. 65. 75. 75. 65. 55. 45. 35. 25.) BOTH 6 CF(0.5, 1.0) \$ POTENTIAL EVAPORATION PARAMETERS PE 50. .25 SUN CF(1.05) \$ RIVER, RESERVOIR AND SNOW PARAMETERS RRS QME 24 DQIN 6 10 40 MSNG(NO) LIN .9

