

IVP Batch Program User's Manual For Verification

By

Henry Herr
Office of Hydrologic Development
National Weather Service

Table of Contents

1.0	<i>Overview</i>	4
1.1	<i>Using This Manual</i>	4
2.0	<i>Notation and Definitions</i>	4
3.0	<i>Execution</i>	5
4.0	<i>Apps-defaults Tokens</i>	5
5.0	<i>Files and Directory Structure</i>	6
6.0	<i>File Types</i>	6
7.0	<i>Batch File Format</i>	6
8.0	<i>Instructions</i>	7
9.0	<i>Commands and Actions Summary Table</i>	8
10.0	<i>Batch Commands</i>	10
	ACTIVE_STATUS = "<ACTIVE/INACTIVE/BOTH>"	10
	ANALYSIS_INTERVAL = "<quantity> <unit>"	10
	BREAKDOWN_BY_FCSTTS = <ON/OFF>.....	11
	BREAKDOWN_BY_LID = <ON/OFF>.....	11
	COMP_VARIABLE = "<variable>"	11
	DUR = "<DUR1>,<DUR2>,...,<DURn>"	11
	END_TIME = "<DATE>"	11
	EXTREMUM = "<EXT1>,<EXT2>,...,<EXTn>"	12
	FCST_CAT = "<CAT1>,<CAT2>,...,<CATn>"	12
	FCST_CAT_USED = <category>.....	12
	FCST_TS = "<TS1>,<TS2>,...,<TSn>"	13
	GRAPH_TEMPLATE = <filename>.....	13
	ISSUANCE_END = "<number of hours>" -or- "<quantity> <unit>"	13
	ISSUANCE_START = "<number of hours>" -or- "<quantity> <unit>"	14
	ISSUANCE_STEP = "<number of hours>" -or- "<quantity> <unit>"	14
	LEADTIME_END = "<number of hours>" -or- "<quantity> <unit>"	14
	LEADTIME_START = "<number of hours>" -or- "<quantity> <unit>"	14
	LEADTIME_STEP = "<number of hours>" -or- "<quantity> <unit>"	14
	OBS_CAT = "<CAT1>,<CAT2>,...,<CATn>"	15
	OBS_CAT_USED = <category>.....	15
	OBS_TYPE = <RAW/PROCESSED>.....	15
	OUTPUT_FILE = "<filename>,<c/a>"	15
	PAIRS_FILE = "<filename>,<c/a>"	16
	PDF_BINS = "<positive integer>"	16
	PE = "<PE1>,<PE2>,...,<PEn>"	16
	PRIMARY_STATS = "<batch stat>,<batch stat>,..."	16
	PRIMARY_PLOT_TYPE = <type>.....	16
	QUANTILES = "<prob1>,<prob2>,..."	17
	RIVERRESPONSE = "<RESP1>,<RESP2>,..."	17
	ROC_PTS = "<positive integer>"	17
	SECONDARY_STATS = "<batch stat>,<batch stat>,..."	17
	SECONDARY_PLOT_TYPE = <type>.....	17
	START_TIME = "<DATE>"	17
	XAXIS_VARIABLE = "<variable>"	18
11.0	<i>Batch Actions</i>	18
	CALCSTATS = "<category>,<STAT1>,<STAT2>,...,<STATn>"	18
	CLEAR_GROUPS = <anything>.....	18
	DEF_GRP = "<LID1>,<LID2>,...,<LIDn>"	18
	DEF_LOC = "<LID1>,<LID2>,...,<LIDn>" or "NATIONAL"	19

GEN_GRAPH = <image filename>[,<data filename>]" 19
NATLSTATS = "<LID1>,<LID2>,...,<LIDn>" 20
12.0 Special Tokens.....20
@FILE = <filename>..... 21
@+ = <line continuation>..... 21
13.0 The Verification Data Gathering Algorithm21
14.0 Output Files22
 14.1 Output Statistics File..... 23
 14.2 Pairs File 24
 14.3 Output Image File..... 24
 14.4 Image Data File 24
15.0 Examples.....24

1.0 Overview

The IVP Batch Program serves three functions:

- (1) constructing forecast-observed data pairs,
- (2) calculating verification statistics, and
- (3) generating graphical products.

This manual provides instructions for the second and third functions, calculating verification statistics using data stored in the vfympairs table of the archive database, and generating graphics based on those statistics.

1.1 Using This Manual

This manual provides a description of the format of an IVP Batch Program input batch file. If producing raw statistics (numbers), it is recommended that the user review the commands in this manual and the examples prior to using the software. Also, read Section 8.0 for instructions on how to put together a batch input file. If producing graphics, it is recommended that the user run the IVP program (see the *Interactive Verification Program User's Manual*) to create the batch file. If editing is needed, this manual can be used to determine how to edit the batch file, and the IVP Batch Builder can be used to do the editing.

Section 13.0 describes the formats of the output files. Refer to this section in order to understand the output produced in the output files.

2.0 Notation and Definitions

This section provides definitions used throughout this section and the remaining sections. Additional definitions to those in this section will be provided as needed. The definitions are as follows:

- *input token*: The <token> portion of a batch file input line. All tokens are displayed in this font.
- *token value*: The <value> portion of a batch file input line. All token values are displayed in this font and in quotes "".
- *batch command*: An input token that is used to specify a parameter. Its token value is stored by the batch file processor. A batch command does not result in any calculations or output (except logging output written to the terminal). Batch commands are displayed in **bold**.
- *batch action*: An input token that triggers a particular action, such as querying the archive database for pairs or performing calculation of verification statistics. The token value specifies parameters of that action. Batch actions are displayed in **bold**.
- *data pair*: A forecast-observed pair, defined in the vfympairs table of the archive database.
- *verification group*: A collection of verification locations that are to be lumped together to produce one set of statistics, and parameters that define how those statistics are to be computed.
- *verification location*: A forecast point and data type for which verification is to be performed. A verification location is uniquely defined by its location id and SHEF pedtsep code.

3.0 Execution

Before executing the IVP Batch Program, be sure that the tables vfyruninfo, rivercrit, and location are all populated correctly for each location for which verification is to be done. The vfyruninfo table is populated using the Vfyruninfo Editor. The rivercrit table must be populated in order for the flood stages and flows to be found. The location table must be populated in order for the rfc to be identified for a location, given by the field rfc. If no rfc is found, then the rfc is assumed to be "NONE".

To execute the Verification Batch Program, enter:

```
cd $(get_apps_defaults verify_dir)/scripts
ivpbatch [-c] <batch file name>
```

where <batch_file_name> is the name of the batch file. If the first letter of the name of the batch file is either '.' or '/', then the file name is assumed to be fully specified, relative to the current directory. Otherwise, it is assumed to be specified relative to the directory given by the apps-defaults token "vsys_input". Use the -c option only if executed within a cron.

4.0 Apps-defaults Tokens

The following apps-defaults tokens are used by the IVP Batch Program:

- adb_name : <depends upon RFC>
- util_rls : \$(util_dir)/bin/RELEASE
- verify_dir : /rfc_arc/verify
- vsys_dir : \$(verify_dir)
- vsys_input : \$(vsys_dir)/input
- vsys_output : \$(vsys_dir)/output
- vsys_files : \$(vsys_dir)/files
- rax_pghost : ax <optional; if not present, it uses the PGHOST environment variable>
- pguser : pguser <optional; if not present, it uses the PGUSER environment variable>
- pgport : 5432 <optional; if not present, it uses the PGPORT environment variable>

Each of the above directories must exist for the IVP Batch Program to run properly. It is also recommended that the following directory be created for each user:

```
$(vsys_input)/$LOGNAME
```

If this recommendation is followed, then the apps-defaults site file should override the setting of vsys_input as follows:

- vsys_input : \$(vsys_dir)/input/\$(LOGNAME)

All of these directories should be constructed prior to running the IVP Batch Program.

5.0 Files and Directory Structure

All of the directories corresponding to apps-defaults tokens in Section 4.0 must exist. In addition, the following files must be in place for the IVP to execute properly:

```
$(util_rls)/rfc.ohd.jar
$(sys_java_dir)/lib/ChartDirector.jar
/usr/share/java/pg74.215.jdbc3.jar
$(vsys_dir)/scripts/ivpbatch
```

6.0 File Types

The following types of files are used within the IVP (recommended file extensions are given in parentheses; for the output image files, the extensions are required)

- **Batch Input Files (.bat):** These files are input files to the IVP Batch Program, and can be loaded by the GUI in order to specify parameters of the data used and plots generated. These files are, by default, assumed to be in the `$(vsys_input)` directory.
- **Graph Template Files (.txt):** These files specify properties of the charts to create, including labels, fonts, colors, sizes, axis limits, and others. These files are, by default, assumed to be in the `$(vsys_files)/$LOGNAME/templates` directory.
- **Output Image Files (.png, .jpg, .jpeg):** These files are images of the charts generated by the IVP or IVP Batch Program. These files are, by default, assumed to be in the `$(vsys_output)/$LOGNAME` directory. Any such file **MUST** have one of the extensions listed.
- **Output Data Files (.dat):** These files are ASCII format files that provide the data plotted to a chart in a tabular format. These files are, by default, assumed to be in the `$(vsys_output)/$LOGNAME` directory.
- **System Settings File:** This file is an ASCII format file used to specify defaults for the appearance of the IVP and its charts. The file must be in the directory `$(vsys_dir)/app-defaults` and have the name `IVP_SYSTEM_FILE.txt`. See Appendix A for more details.

7.0 Batch File Format

Each line of the batch file corresponds to a command or a parameter setting. All lines must be of the format:

```
<token> = <value>
```

with the following restrictions:

1. The token is not case sensitive.
2. Any number of spaces or tabs may be placed before or after the token and before or after the value.
3. A space, new-line (carriage return), tab, or pound ('#') marks the end of a value or token.
4. Double quotes must be placed around the value if it is to contain tabs, spaces, or pounds, but the value may never contain a new-line. For example,

```
my_name = john doe
```

has a token of "my_name" and a value of "john", whereas

```
my_name = "john doe"
```

has a token of "my_name" and a value of "john doe". If a new-line is encountered, it is treated as the closing double-quote.

5. The character '#', unless it is within double quotes, is used to indicate a comment. All characters after a '#' are ignored.
6. The equal sign ('=') must not be used as part of a value.

If a line is found which does not follow this format or specifies an unrecognized token, an error message will be generated. Blank lines are ignored.

8.0 Instructions

The following steps can be used to setup a batch input file to calculate verification statistics and output those statistics to an ASCII tabular output file:

1. Decide on the locations for which verification is to be done. A location is defined in the vfyruninfo table by a location id (lid), physical element (pe), duration (dur), extremum (ext), and forecast type source (ts). For each location, if you wish to define forecast or observed value categories, then do so by using the FCST_CAT and OBS_CAT. Also, specify the OBS_TYPE appropriately for if you wish to draw the pairs from the vfypairs or vfyprocpairs table. Each location is then made available, or defined, for verification via the DEF_LOC command. Categories can be set in either absolute terms or relative to the location's critical values, as defined in the rivercrit table of the archive database.
2. Decide on the data time interval for which verification is to be done. Set the time frame in the batch file using the START_TIME and END_TIME commands.
 - a. If you wish to break down the overall time interval into subintervals, then determine the width of each subinterval and set it within the batch file using the ANALYSIS_INTERVAL command.
3. If you wish to only calculate statistics for forecasts with a basis time hour-of-day within a particular range of values, then determine the overall issuance time-of-day interval, and set it within the batch file using the ISSUANCE_START and ISSUANCE_END commands.
 - a. If you wish to further break down the issuance time interval into subintervals, each with separately calculated statistics, then determine the width of each subinterval and set it within the batch file using the ISSUANCE_STEP command.
4. If you wish to only calculate statistics for forecasts with a lead time within a particular range of values, then determine the total lead time interval, and set it within the batch file using the LEADTIME_START and LEADTIME_END commands.
 - a. If you wish to further break down the total lead time interval into subintervals, each with separately calculated statistics, then determine the width of each subinterval and set it within the batch file using the LEADTIME_STEP command.
5. Given the defined locations, if you wish to only calculate statistics for locations with specific river response times, then determine the response times you wish to use (slow, medium, or fast), and set it within the batch file via the RIVERRESPONSE command.

6. Given the defined locations, if you wish to limit the analysis to data values with particular physical elements, durations, extremums, or forecast type sources, then determine what physical elements, durations, extremums, or forecast type sources you wish to use and set them in the batch file using the PE, DUR, EXTREMUM and FCST_TS commands.
7. If you want to include inactive locations or both active and inactive locations in your analysis, specify the ACTIVE_STATUS command accordingly.
8. If you want to compute statistics independently for each location, the set the BREAKDOWN_BY_LID token to 'ON'.
9. If you want to compute statistics independently for each forecast type source, the set the BREAKDOWN_BY_FCSTTS token to 'ON'.
10. Determine if you want to generate a pairs file, in addition to an output file. Determine the file names for both files and, if they are not the default names, specify them with the PAIRS_FILE and OUTPUT_FILE command. Note that both file names can use replacement strings, as described in Appendix C of the *IVP User's Manual*, which allows for more generic and robust file names.
11. Determine the verification groups you wish to construct and define the groups using the DEF_GRP command. A verification group defines a collection of locations for which one set of verification statistics are to be produced. A location will only be added to a group if all of these conditions hold:
 - a. It has a river response time within those response times given in Step 5.
 - b. It satisfies all of the constraints of Step 6.
 - c. It has the same number of forecast categories and observed categories, defined in Step 1, as all other locations added to the group.
 - d. Its physical element has a unit that is identical to all other physical elements of the group.
12. Determine the statistics you wish to produce and call the CALCSTATS action in the batch file passing in those statistics.
13. Use the created batch file as input to the IVP Batch Program. The desired output file will be generated.

If the user desires to produce graphics via the IVP Batch Program, the user should use the IVP to generate such a batch file. See the *IVP User's Manual* for more details.

9.0 Commands and Actions Summary Table

The following table summarizes all the batch commands and actions available used in verification within the IVP Batch Program. Specific details for each command and action are provided in the following sections.

Command/Action	Purpose
COMMANDS	
ACTIVE_STATUS	Sets the required active status for a location to be included in the group.
ANALYSIS_INTERVAL	Sets the time interval breaking down the overall analysis period into equal sized subintervals.
BREAKDOWN_BY_FCSTTS	Results in statistics being computed independently for each forecast type source.
BREAKDOWN_BY_LID	Results in statistics being computed independently for each location id.
COMP_VARIABLE	Sets the comparison variable to use when building a plot.
DUR	Sets the required SHEF duration code for a location to be included in the group.

END_TIME	Sets the end time of the overall analysis period.
EXTREMUM	Sets the required SHEF extremum code for a location to be included in the group.
FCST_CAT	Sets the forecast categories for a verification location.
FCST_CAT_USED	Sets the required forecast category for a pair to be included in statistic computations for a plot.
FCST_TS	Sets the required SHEF forecast type source code for a location to be included in the group.
GRAPH_TEMPLATE	Sets the template file to use for setting the properties (labels, fonts, etc.) for a plot.
ISSUANCE_END	Sets the overall end time for the issuance time-of-day interval.
ISSUANCE_START	Sets the overall start time for the issuance time-of-day interval.
ISSUANCE_STEP	Sets the time interval breaking down the overall issuance time-of-day interval into equal sized subintervals.
LEADTIME_END	Sets the overall end time for the lead time interval.
LEADTIME_START	Sets the overall start time for the lead time interval.
LEADTIME_STEP	Sets the time interval breaking down the overall lead time interval into equal sized subintervals.
OBS_CAT	Sets the observed categories for a verification location.
OBS_CAT_USED	Sets the required observed category for a pair to be included in statistic computations for a plot.
OBS_TYPE	Sets the table (vfy pairs or vfypropairs) from which pairs will be drawn when acquiring data for a verification location
OUTPUT_FILE	Specifies the file name to where the output will be written and whether or not the file should be created or appended to.
PAIRS_FILE	Specifies the file name to where the pairs will be written and whether or not the file should be created or appended to.
PE	Sets the required SHEF physical element code for a location to be included in the group.
PRIMARY_STATS	Specifies the statistics to produce and display against the left hand y-axis on a plot.
PRIMARY_PLOT_TYPE	Specifies the plot type, SCATTER, LINE, or BAR, for the primary statistics.
RIVERRESPONSE	Sets the required river response time for a location to be included in the group.
SECONDARY_STATS	Specifies the statistics to produce and display against the right hand y-axis on a plot.
SECONDARY_PLOT_TYPE	Specifies the plot type, SCATTER, LINE, or BAR, for the secondary statistics.
START_TIME	Sets the start time of the overall analysis period.
XAXIS_VARIABLE	Sets the x-axis variable to use when building a plot.
ACTIONS	
DEF_GRP	Defines a group containing the specified locations and for the parameters currently defined in the batch file.
DEF_LOC	Defines verification locations for the specified location ids and the parameters currently defined in the batch file.
GEN_GRAPH	Produces an output image file, and optionally an image data file, for the current defined groups and plot definition parameters.
NATL_STATS	Computes national statistics.

10.0 Batch Commands

A batch command sets a parameter that is used by a batch action (see Section 8). This section provides an alphabetical listing of all of the available batch commands. Commands associated with verification groups (see DEF_GRP action) are in blue; those associated with verification locations (see DEF_LOC action) are in red; those associated with both are in purple; and those associated with statistics and/or graphics (see GEN_GRAPH action) are in green. Acceptable values will be listed for each command, as well as the default if the command is not specified or if the command's token value is "<DEFAULT>". If the passed in token value is not acceptable, then the batch program will print an error message and stop. The following are batch commands:

ACTIVE_STATUS = "<ACTIVE/INACTIVE/BOTH>"

Description: Restricts the locations included in the verification group based on the active status of the location. The active status is defined for a location in the vfyruninfo table via the Vfyruninfo Editor.

Acceptable Values: "ACTIVE" if only those locations for which the record in the vfyruninfo table has an act column value of 'Y'. "INACTIVE" for records with a column value of 'N'. "BOTH" for both.

Default Value: "ACTIVE".

ANALYSIS_INTERVAL = "<quantity> <unit>"

Description: Defines the analysis interval for the verification calculations. This interval breaks down the total [START_TIME, END_TIME] interval into subintervals, each of width equal to that amount given by the token's value. When statistics are calculated, they will be calculated independently for each interval. A value of "NONE" can be used to specify that no subintervals are to be created.

Acceptable Values:

- "<quantity> <unit>": If specified, then the overall interval will be broken down into subintervals of the sized specified. The <quantity> must be a positive integer and the <unit> must be either "WEEKS" ("WEEK" or "WK"), "DAYS" ("DAY" or "DY"), or "HOURS" ("HOURL" or "HR").
- "MONTHLY": If specified, then the overall interval will be broken down into months, with statistics being computed for each month. If the overall period from START_TIME to END_TIME covers all 12 months in the year, then the statistics are computed for each month from January to December, in order. If not, then the statistics are computed starting from the month of the START_TIME and ending at the month of the END_TIME.
- "YEARLY": If specified, then the year components of the START_TIME and END_TIME are assumed to specify the years over which analysis is to be performed, while the rest of the dates specify the interval within each year for which analysis is to be performed. Year wrapping will occur if the END_TIME month/day/time is before the START_TIME month/day/time. For example, a yearly interval for START_TIME "2000-12-01" and END_TIME "2002-02-14" would create two intervals: 2000-12-01 through 2001-02-14 and 2001-12-01 through 2002-02-14.
- "YEARLY_READONLY": If specified, read in the data that would normally be read with a setting of "YEARLY", but compute one set of statistics for all of the pairs lumped together over all of the years. For example, a YEARLY_READONLY interval for START_TIME 2000-12-01 and END_TIME 2002-02-14 would create only one interval containing all of the data from 2000-12-01 – 2001-02-14 and 2001-12-01 – 2002-02-14.

Default Value: "NONE".

BREAKDOWN_BY_FCSTTS = <ON/OFF>

Description: If "ON", then statistics will be computed independently for each forecast type source.

Acceptable Values: Either "ON" or "OFF".

Default Value: "OFF".

BREAKDOWN_BY_LID = <ON/OFF>

Description: If "ON", then statistics will be computed for each verification location lid independently of the others. For example, to compute statistics independently for all location lid's available in the vfyruninfo table, define all of the locations, set this command to "ON" and call DEF_GRP with a value of "ALL".

Acceptable Values: Either "ON" or "OFF".

Default Value: "OFF".

COMP_VARIABLE = "<variable>"

Description: Defines the comparison variable, which is displayed within the legend of the generated plot.

Acceptable Values: Must be "NONE", "LOCATION", "ANALYSIS INTERVAL", "ISSUANCE INTERVAL", "LEADTIME INTERVAL", "OBSERVED CATEGORY", "FORECAST CATEGORY", or "FORECAST TYPE SOURCE".

Default Value: "NONE"

NOTE: The comparison variable and x-axis variable cannot be the same.

DUR = "<DUR1>,<DUR2>,...,<DURn>"

Description: Defines a list of durations. When a location is defined via a DEF_LOC action, this restricts the locations defined to be those which have a duration code from within this list. When a group is defined via a DEF_GRP action, this restricts the locations added to the group to be those with a duration code from within this list.

Acceptable Values: A list of valid 1-character SHEF duration codes or "ALL" to allow for any duration. The list must be comma separated and, if you have spaces within the list, must be within double quotes.

Default Value: "ALL".

END_TIME = "<DATE>"

Description: Defines the end date/time for the pairing run. The end time can be absolute or can be relative to the current system time. Any data pair included in statistics calculation must have a valid time prior or equal to this date/time.

Acceptable Values: If the date is absolute, then it must be of one of these formats:

- "CCYY-MM-DD" (assumes time of 23:59:59 that day),
- "CCYY-MM-DD hh:mm:ss",
- "CCYY-MM-DD hh:mm:ss TZC",
- "MMDDCCYY:hh".

If the date is set relative to the current system time, then it must of the following format:

"* [\leftarrow or \rightarrow <quantity> <unit> <quantity> <unit> ...]".

Everything in '[']' is optional. The <quantity> must be a positive integer and the <unit> must be either "WEEKS" ("WEEK" or "WK"), "DAYS" ("DAY" or "DY"), or "HOURS" ("HOUR" or "HR").

Default Value: "*" (the current system time).

EXTREMUM = "<EXT1>,<EXT2>,...,<EXTn>"

Description: Defines a list of extremums. When a location is defined via a DEF_LOC action, this restricts the locations defined to be those which have an extremum code from within this list. When a group is defined via a DEF_GRP action, this restricts the locations added to the group to be those with an extremum code from within this list.

Acceptable Values: A list of valid 1-character SHEF extremum codes or "ALL" to allow for any extremum. The list must be comma separated and, if you have spaces within the list, must be within double quotes.

Default Value: "ALL".

FCST_CAT = "<CAT1>,<CAT2>,...,<CATn>"

Description: Defines categories used to break down data pairs by the forecast value. The categories are defined as follows: ($CAT1 \leq x < CAT2$), ..., ($CATn-1 \leq x \leq CATn$). Any data pair included in statistics calculation must have a forecast value within at least one of the defined categories. The categories can be location dependent if defined relative to the flood stage.

Acceptable Values: A list of categories or "NONE" to not categorize data. The list must be comma separated and, if you have spaces within the list, must be within double quotes. Each "<CATn>" value must one of the following:

- a number (decimal or otherwise; for stage data, unit is assumed to be feet),
- "MIN" to denote no lower bound,
- "MAX" to denote no upper bound,
- "<scalar>*<critical value>".

In the last case, the scalar is a positive integer value, and the critical value is either "AS", "FS", "ModFS", "MajFS", "RS" for height data; or "AF", "FF", "ModFF", "MajFF", or "RF" for discharge data. These correspond to fields of the rivercrit table: action, flood, modflood, majflood, record, actionf, floodf, modfloodf, majfloodf, and recordf, respectively. If the desired value cannot be found for a location, a message will be generated stating that no critical value was found and the batch processor will print an error message and stop. The list will be sorted into ascending order prior to use. See Section 15.0 for examples.

Default Value: "NONE", which is equivalent to "MIN, MAX".

FCST_CAT_USED = <category>

Description: Sets the forecast category to use. See Section 17.6 of the *Interactive Verification Program User's Manual* for a description of how the forecast category is specified for a graphic within the GUI. Either this command or OBS_CAT_USED must be something other than "NONE". If this command is set to "NONE" and OBS_CAT_USED is already "NONE", then OBS_CAT_USED will be set to "ALL". If this is set to a value other than "NONE", then OBS_CAT_USED will be set to "NONE".

Acceptable Values: Must be "NONE" (equivalent to "Do Not Use" in IVP), "ALL" ("All Categories Combined" or "Use Only Category" in IVP), or "CAT#" ("Category #" in IVP) where '#' is the category number (the categories are sorted into ascending order and numbered 1, 2,...).

Default Value: "NONE"

FCST_TS = "<TS1>,<TS2>,...,<TSn>"

Description: Defines a list of forecast type sources. When a location is defined via a DEF_LOC action, this restricts the locations defined to be those which have a forecast type source code from within this list. When a group is defined via a DEF_GRP action, this restricts the locations added to the group to be those with a forecast type source code from within this list.

Acceptable Values: A list of valid forecast type sources, "ALL" to allow for any valid type source, or "ALL_BUT_PERSIST" to allow for any valid type source, other than 'FR', which is associated with persistence. The list must be comma separated and, if there are spaces within the list, must be within double quotes.

Default Value: "ALL".

NOTE: To include persistence forecasts in the analysis, 'FR' must be a part of the FCST_TS value or the FCST_TS value must be "ALL" for both the DEF_LOC and DEF_GRP commands.

GRAPH_TEMPLATE = <filename>

Description: Defines the name of a template file that contains chart properties for the chart to create (via GEN_GRAPH). See Appendix D of the *Interactive Verification Program User's Manual* for a description of template files.

Acceptable Values: A filename of a file that can be read on the system, or "NONE". If the filename does not begin with a '.' or '/', then it is assumed to be located within the directory corresponding to \$(vsys_files)/\$(LOGNAME)/templates. If "NONE" is specified, then default chart properties are used.

Default Value: "NONE".

NOTE: Though the file is in ASCII format, it is highly recommended that only the IVP **Chart Property Manager** window be used to edit the properties contained in a template file.

NOTE: Template file names can use replacement strings, as described in Appendix C of the *Interactive Verification Program User's Manual*, which allows for more generic and robust file names.

ISSUANCE_END = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the largest issuance hour-of-day to be used in verification. Any data pair included in statistics calculation must have a basistime hour smaller than or equal to this upper bound. A value of "NONE" can be used to specify no upper bound.

Acceptable Values: "NONE", a positive integer (assumed to be hours), or "<quantity> <unit>" (see ANALYSIS_INTERVAL). Note that the <quantity> must be within 0-24, and the <unit> can only be "hours" ("hour" or "hr"). The value of this token, in hours, must be larger than ISSUANCE_START.

Default Value: "NONE".

ISSUANCE_START = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the smallest issuance hour-of-day to be used in verification. Any data pair included in statistics calculation must have a basistime hour larger than or equal to this lower bound. A value of "NONE" can be used to specify no lower bound.

Acceptable Values: See ISSUANCE_END. The value of this token, in hours, must be smaller than ISSUANCE_END.

Default Value: "NONE".

NOTE: If either ISSUANCE_END or ISSUANCE_START is not "NONE", then neither can be.

ISSUANCE_STEP = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the issuance time interval for the verification calculations. This interval breaks down the total [ISSUANCE_START, ISSUANCE_END] interval into subintervals, each with a width equal to the token's value. When statistics are calculated, they will be calculated independently for each issuance time interval. Note that each subinterval interval excludes the lower bound and includes the upper bound, except for the first interval which includes both.

Acceptable Values: See ISSUANCE_END. "NONE" specifies use of the entire issuance time interval.

Default Value: "NONE".

LEADTIME_END = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the largest lead time to be used in verification. Any data pair included in statistics calculation must have a (validtime – basistime) smaller than or equal to this upper bound. A value of "NONE" can be used to specify no upper bound.

Acceptable Values: "NONE", a positive integer (assumed to be hours), or "<quantity> <unit>" (see ANALYSIS_INTERVAL). The value of this token, in hours, must be larger than LEADTIME_START.

Default Value: "NONE".

LEADTIME_START = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the smallest lead time to be used in verification. Any data pair included in statistics calculation must have a (validtime – basistime) greater than this lower bound. A value of "NONE" can be used to specify no lower bound (this is equivalent to a LEADTIME_START of "0").

Acceptable Values: "NONE", a positive integer (assumed to be hours), or "<quantity> <unit>" (see ANALYSIS_INTERVAL).

Default Value: "NONE".

NOTE: If the if either LEADTIME_END or LEADTIME_START is not "NONE", the neither can be.

LEADTIME_STEP = "<number of hours>" -or- "<quantity> <unit>"

Description: Defines the lead time interval for the verification calculations. This interval breaks down the total (LEADTIME_START, LEADTIME_END] interval into subintervals, each with a width equal to the token's value. When statistics are calculated, they will be calculated independently for each lead time interval. Note that each interval excludes the lower bound and includes the upper bound. For example, if LEADTIME_STEP is 6 and LEADTIME_START is 0, then the first interval will be (0,6],

the next will be (6, 12], and so on. A value of "NONE" can be used to specify that no subintervals are to be created.

Acceptable Values: "NONE", a positive integer, or see ANALYSIS_INTERVAL if "<quantity> <unit>" is used.

Default Value: "NONE".

OBS_CAT = "<CAT1>,<CAT2>,...,<CATn>"

Description: Defines categories used to break down data pairs by the observed value. The categories are defined as follows: ($CAT1 \leq x < CAT2$), ..., ($CATn-1 \leq x \leq CATn$). Any data pair included in statistics calculation must have an observed value within at least one of the defined categories. The categories can be location dependent if defined relative to the flood stage.

Acceptable Values: See FCST_CAT command above, except that the categories bound the observed value, not the forecast value.

Default Value: "NONE".

OBS_CAT_USED = <category>

Description: Sets the observed category to use. See Section 17.6 of the *Interactive Verification Program User's Manual* for a description of how the observed category is specified for a graphic within the GUI. Either this command or FCST_CAT_USED must be something other than "NONE". If this command is set to "NONE" and FCST_CAT_USED is already "NONE", then FCST_CAT_USED will be set to "ALL". If this is set to a value other than "NONE", then FCST_CAT_USED will be set to "NONE".

Acceptable Values: See FCST_CAT_USED command above.

Default Value: "NONE"

OBS_TYPE = <RAW/PROCESSED>

Description: Defines if the vfy pairs (setting of "RAW") or vfyprocpairs (setting of "PROCESSED") table will be searched for pairs associated with the verification location.

Acceptable Values: Either "RAW" or "PROCESSED".

Default Value: "RAW"

OUTPUT_FILE = "<filename>,<c/a>"

Description: Defines the name of the one output file to use for the output statistics of the verification group and whether to open it for creation or append. When CALCSTATS is executed for multiple groups that have the same OUTPUT_FILE parameter, the output file will be opened according to the <c/a> flag for the first group that requires the file, but will always append to the file for any other groups that require the file (i.e., CALCSTATS will open each output file only one time).

Acceptable Values: Either "NONE" (to generate no output file), "<DEFAULT>" (for the default, below), or "<filename>,<c/a>", where:

- <filename>: The name of the file to open or "<DEFAULT>" to use the default name. If the first character is not a '/' or '.', then the file will be opened relative to the directory pointed to by apps-defaults token vsys_output. The file name can include directory components, which will be created if needed by the IVP Batch Program.
- <c/a>: Either 'c' or 'a'. If 'c', then the file specified by the <filename> will be created or overwritten when first created. If 'a', then the file will be opened for appending.

Defaults: "vfyoutput.txt,c"

NOTE: Output file names can use replacement strings, as described in Appendix C of the Interactive Verification Program User's Manual, which allows for more generic and robust file names.

PAIRS_FILE = "<filename>,<c/a>"

Description: Defines the name of the one pairs file to use for the pairs used in computing the statistics for the verification group and whether to open it for creation or append.

Acceptable Values: See OUTPUT_FILE.

Default Value: "NONE".

NOTE: Pairs file names can use replacement strings, as described in Appendix C of the Interactive Verification Program User's Manual, which allows for more generic and robust file names.

PDF_BINS = "<positive integer>"

Description: Defines the number of "bins" used to estimate the probability density functions produced in the PDF special plot. See Appendix B of the *IVP User's Manual* for more details.

Acceptable Values: Positive integer.

Default Value: "10".

PE = "<PE1>,<PE2>,...,<PEn>"

Description: Defines a list of physical elements. When a location is defined via a DEF_LOC action, this restricts the locations defined to be those which have a physical elements code from within this list. When a group is defined via a DEF_GRP action, this restricts the locations added to the group to be those with a physical elements code from within this list.

Acceptable Values: A list of valid 2-character SHEF physical element codes or "ALL" to allow for any valid physical element. The list must be comma separated and, if you have spaces within the list, must be within double quotes.

Default Value: "ALL".

PRIMARY_STATS = "<batch stat>,<batch stat>,..."

Description: Sets the primary statistics, or those statistics displayed against the left-hand y-axis of the generated plot.

Acceptable Values: Each <batch stat> must correspond to an output statistics column header. All of the statistics must have the same scale, which means they must be in the same group within the **IVP Statistic Chooser Manager** of the IVP. See Appendix B.1 in *Interactive Verification Program User's Manual* for a listing of the output column headers, and Section 18.0 for a listing of the groups.

Default Value: No default value. This command must have a value.

PRIMARY_PLOT_TYPE = <type>

Description: Sets the plot type to use for the primary statistics in the generated plot.

Acceptable Values: Must be "BAR", "LINE", or "SCATTER".

Default Value: "SCATTER"

QUANTILES = "<prob1>,<prob2>,..."

Description: Defines the quantiles computed when the statistics to calculate in a CALCSTATS action are defined as QUANTILES (see the next Section). It also controls what quantiles are available to plot and can be listed as a <batch stat> for the PRIMARY_STATS and SECONDARY_STATS commands, described in this section. See Appendix B of the *IVP User's Manual* for more details.

Acceptable Values: Each <prob#> must be a probability (between 0 and 1).

Default Value: "0.25, 0.50, 0.75".

RIVERRESPONSE = "<RESP1>,<RESP2>,..."

Description: Defines a list of river response times. The response time is stored in the vfyruninfo table of the archive database as the resptime field. Any location included in statistics calculation must have a response time in this list. If no resptime field can be found for a given location in the vfyruninfo table, then that location will be assigned a response time of "NONE".

Acceptable Values: A list of valid response times ("SLOW", "MEDIUM", "FAST") or "ALL" to allow for any response time, including "NONE". The list must be comma separated and, if you have spaces within the list, must be within double quotes.

Default Value: "ALL".

ROC_PTS = "<positive integer>"

Description: Defines the number of points computed on the relative operating characteristic (ROC) curve for the ROC special plot. See Appendix B of the *IVP User's Manual* for more details.

Acceptable Values: Positive integer.

Default Value: "100".

SECONDARY_STATS = "<batch stat>,<batch stat>,..."

Description: Sets the secondary statistics, or those statistics displayed against the right-hand y-axis of the generated plot.

Acceptable Values: See the PRIMARY_STATS command, above.

Default Value: "NONE"

SECONDARY_PLOT_TYPE = <type>

Description: Sets the plot type to use for the secondary statistics in the generated plot.

Acceptable Values: Must be "BAR", "LINE", or "SCATTER".

Default Value: "SCATTER"

START_TIME = "<DATE>"

Description: Defines the start date/time for the pairing run. The start date/time can be absolute or can be relative to the current system time. Any data pair included in statistics calculation must have a valid time after or equal to this date/time.

Acceptable Values: See END_TIME above, except that for the format of "CCYY-MM-DD" a time of 00:00:00 is assumed.

Default Value: "* - 14 DAYS" (two weeks prior to current system time).

XAXIS_VARIABLE = "<variable>"

Description: Defines the variable to display along the x-axis, against which the statistics are to be plotted in the generated plot.

Acceptable Values: Must be "LOCATION", "ANALYSIS INTERVAL", "ISSUANCE INTERVAL", "LEADTIME INTERVAL", "OBSERVED CATEGORY", "FORECAST CATEGORY", or "FORECAST TYPE SOURCE".

Default Value: "LOCATION"

NOTE: The comparison variable and x-axis variable cannot be the same.

11.0 Batch Actions

Actions instruct the verification program to do something, such as open a file or calculate statistics. The nature of what is done depends on the action given. Acceptable values will be listed for each token value. If the value is not acceptable, then the batch program will print an error message and stop.

The following are valid actions within the verification system:

CALCSTATS = "<category>,<STAT1>,<STAT2>,...,<STATn>"

Description: This action causes statistics to be generated given the current command settings provided prior to this line in the batch file. Statistics will be produced for each verification group. All output is sent to the output file specified by the OUTPUT_FILE associated with the group.

Acceptable Values:

- <category>: Specifies if categories are to be constructed relative to observed values, forecast values, or both observed and then forecast. The value must be either "OBS" for observed categories, "FCST" for forecast categories, or "BOTH" for both categories. If you have not defined any categories and just want to compute overall statistics, then put either "OBS" or "FCST" as the category, and ignore the category information in the output file.
- <STATn>: Specifies a group of statistics to compute. Values for <STATn> are provided as the CALCSTATS identifier in Appendix B.1 of the *Interactive Verification Program User's Manual*.

Default Value: Does not apply. The value must be acceptable.

CLEAR_GROUPS = <anything>

Description: Clears all of the groups that have been defined up until this point in the batch file.

Acceptable Values: The value is ignored.

Default Value: Does not apply. The value is ignored.

DEF_GRP = "<LID1>,<LID2>,...,<LIDn>"

Description: Defines a group for a collection of location ids, each one of which having at least one verification location defined at least via a DEF_LOC action previously within the batch file. The locations provided will be added to the group only if they pass the restrictions placed on them by the commands PE, DUR, EXTREMUM, FCST_TS, ACTIVE_STATUS, and RIVERRESPONSE. In addition to a list of locations, associated with each group are parameters. The parameters are specified by these batch commands: START_TIME, END_TIME, ANALYSIS_INTERVAL, ISSUANCE_START, ISSUANCE_END, ISSUANCE_STEP, LEADTIME_START,

LEADTIME_END, LEADTIME_STEP, OUTPUT_FILE, PAIRS_FILE,
BREAKDOWN_BY_FCSTTS, and BREAKDOWN_BY_LID.

Acceptable Values:

- <LIDn>: A defined location (via DEF_LOC, described below); "ALL" to include every location currently defined via DEF_LOC in one group; or ONE_PER_LOCATION to create one group per location. Note that:
 - All locations must be defined at least once via a DEF_LOC action.
 - All locations within a group must have an identical number of forecast categories and an identical number of observed categories.

Default Value: None.

DEF_LOC = "<LID1>,<LID2>,...,<LIDn>" or "NATIONAL"

Description: Defines the locations given in the list using parameters as set via commands in the batch file prior to this action. This action uses the PE, DUR, EXTREMUM, and FCST_TS commands to determine the specific verification locations being defined. Associated with each verification location are parameters specified by these batch commands: OBS_CAT, FCST_CAT, and OBS_TYPE. A location (defined by a location id, physical element, duration, extremum, and forecast type source) can only be defined once in memory, so any previous instance of the defined location, possibly generated with a previous DEF_LOC call, will be overwritten.

Acceptable Values:

- <LIDn>: A location to define for which a record must exist within the vfyruninfo table, "ALL" to define all locations given within vfyruninfo, or
- "NATIONAL" to define all locations in the vfyruninfo table with a national flag value of 'Y'. This flag is set for a location within the Vfyruninfo Editor.

Default Value: None.

GEN_GRAPH = <image filename>[,<data filename>]"

Description: Generates a graphic, using the values of commands PRIMARY_STATS, SECONDARY_STATS, PRIMARY_PLOT_TYPE, SECONDARY_PLOT_TYPE, XAXIS_VARIABLE, COMP_VARIABLE, FCST_CAT_USED, OBS_CAT_USED to define the parameters of a plot, exactly as the corresponding fields in the **Verification Plot Definition Manager** defines the plot for the IVP (see Section 17.0 of the *Interactive Verification Program User's Manual*). The value of command GRAPH_TEMPLATE defines a template file to be read which contains properties for the displayed chart. These properties are those defined in the **Chart Property Manager** of the IVP. One graphic will be generated per defined group, so that it is imperative that replacement strings (see Appendix C in the *Interactive Verification Program User's Manual*) are used in the image and data file names, or else the last defined group will overwrite the earlier groups.

Acceptable Values:

- <image filename>: The output image file to create. The filename must have a ".jpeg", ".jpg", or ".png" extension, so that the IVP Batch Program knows which type of image to create.
- <data filename>: Optional. The name of the image data file to create. This file is an ASCII file, equivalent to clicking on "Save to File" in the **IVP Statistics Data Viewer** of the IVP (see Section 16.0 of the *Interactive Verification Program User's Manual*).

In both cases above, the file name can include directory components, which will be created if needed by the IVP Batch Program.

Default Value: None. The <image filename> must be specified, though the <data filename> is optional.

NOTE: See Section 14.0 for a description of the output files generated by a GEN_GRAPH action.

NATLSTATS = "<LID1>,<LID2>,...,<LIDn>"

Description: This action causes national statistics to be generated for each of the passed in locations. All previously defined commands are used as with CALCSTATS, except that the following are assumed:

- ANALYSIS_INTERVAL = NONE
- FCST_CAT = NONE
- ISSUANCE_END = NONE
- ISSUANCE_START = NONE
- ISSUANCE_STEP = NONE
- LEADTIME_END = 72
- LEADTIME_START = 0
- LEADTIME_STEP = 6
- OBS_CAT = "*0.0,*1.0,*10.0"

It is assumed that every verification location defined by the passed in location ids is to have its statistics calculated independently. Therefore, any previously defined verification locations and groups are ignored. Locations can still be restricted by using the PE, DUR, EXTREMUM and FCST_TS commands (see the NOTE, below, however). The OUTPUT_FILE is ignored: the output statistics produced are organized into files by location RFC, leadtime interval, and category. The names of the output files are:

"<rfc>_<ccyymm>_<above/below>.stat_<lead time interval end>hr_tab".

The file is placed in the directory \$(vsys_output). <rfc> is the lower case five letter rfc abbreviation; <ccyymm> is the four digit year and two digit month of the date specified by the batch command START_TIME; <above/below> is either "above" for data in the upper category or "below" for data in the lower category; and <lead time interval end> is the upper bound of the lead time interval of the statistics contained in the file.

Acceptable Values: See the DEF_LOC action; the "NATIONAL" value is particularly useful, as it will include any location for which the national flag is set to 'Y'; persistence forecasts are always assumed to have a national flag of 'N'. Using "NATIONAL" is the recommended approach, but be sure to specify the national flag via the Vfyruninfo Editor for all locations to include.

Default Value: None.

NOTE: If the FCST_TS command is set ot ALL, this action will set it to ALL_BUT_PERSIST. Persistence forecasts are assumed to never be part of the national program.

12.0 Special Tokens

A special token is a batch language token that is not related to specifying parameters or performing actions related to the software program.

This section provides an alphabetical listing of all special tokens. The following are special tokens:

@FILE = <filename>

Description: Forces batch processor to process the file specified by <filename>. The file is treated just as if its contents were included within the original batch file.

Acceptable Values: Any valid file that will not force the batch processor to enter into an infinite, recursive loop (i.e. if file 1 references file 2, then file 2 can never reference file 1).

Default Value: Does not apply. The file name must be valid.

@+ = <line continuation>

Description: Line continuation token. The batch processor will take the value of this line and append it to the value on the previous line.

Acceptable Values: Depends upon previous lines batch command or action.

Default Value: Depends upon previous lines batch command or action.

13.0 The Verification Data Gathering Algorithm

The algorithm used to collect verification data pairs from the vfpairs table of the archive database is as follows:

1. Every time the action DEF_LOC is called, a collection of locations is defined for the specified location ids and user specified list of pes (via PE command), durations (DUR command), extremums (EXTREMUM command), and forecast type sources (FCST_TS command). Associated with each defined location in the collection are the following:
 - List of forecast value categories given by the FCST_CAT command. This is stored in its unprocessed (comma separated list) form.
 - List of observed value categories given by the OBS_CAT command. This is stored in its unprocessed (comma separated list) form.
 - The table from which pairs are to be drawn, based on the OBS_TYPE: vfpairs if "RAW" or vfyprocpairs if "PROCESSED".
 - Other information, including RFC, response time, and critical values, each of which is loaded from the archive database (tables location, vfyruninfo, and rivercrit, respectively).

2. Every time the action DEF_GRP is called, a verification group is defined. Associated with each defined group are the following:
 - List of locations in the group, generated by cross referencing the defined locations (via the DEF_LOC action above) with the values of the DEF_GRP action (a list of locations; see Section 11.0), and the commands PE, DUR, EXTREMUM, FCST_TS, ACTIVE_STATUS, and RIVERRESPONSE. A verification location must satisfy all of these constraints to be included in a verification group.
 - Analysis window start and end times (START_TIME, END_TIME commands) and analysis interval (ANALYSIS_INTERVAL commands).
 - Issuance time-of-day interval start and end (ISSUANCE_START, ISSUANCE_END commands), and issuance time step (ISSUANCE_STEP commands).
 - Lead time interval start and end (LEADTIME_START, LEADTIME_END commands), and lead time step (LEADTIME_STEP commands).

- Flags specifying if the statistics are to be computed independently for each location id (BREAKDOWN_BY_LID) or each forecast type source (BREAKDOWN_BY_FCSTTS).
 - An output file name and whether to create it from scratch or append to the existing file.
 - A pairs file name and whether to create it from scratch or append to the existing file.
3. When the CALCSTATS algorithm is called, for each group in the list of groups and for each analysis interval in the list of analysis intervals for that group (based on START_TIME, END_TIME, and ANALYSIS_INTERVAL), the following is done:
- a. All pairs records that have:
 - a corresponding verification location for the current group,
 - a forecast validtime within the current analysis interval,
 - a forecast basistime hour-of-day that is within the overall issuance time interval (based on ISSUANCE_START and ISSUANCE_END),
 - a forecast (validtime – basistime) within the overall lead time interval (based on LEADTIME_START and LEADTIME_END) for the current group, and
 - a quality_code field value that is larger than 1073741824 (i.e. not considered 'BAD')

will be loaded from the database and stored as raw forecast-observed pairs.
 - b. For each issuance time-of-day subinterval (based on ISSUANCE_STEP), a subset of the current pairs records is created.
 - c. For each lead time subinterval (based on LEADTIME_STEP), a subset of the current pairs records is created.
 - d. If the first argument of the CALCSTATS command is either "BOTH" or "FCST", then for each subset from Step 4b and forecast category, the following is done:
 - i. All vfypairs records are extracted from the list gathered in Step 3c for which the forecast value is in the current category.
 - ii. All statistics are produced for the current records.
 - iii. All statistics are output to the group's output file in the appropriate format.
 - e. If the first argument of the CALCSTATS command is either "BOTH" or "OBS", for each observed category, Steps 3d-i, 3d-ii, and 3d-iii, are each performed, except that in step 3d-i, the observed value is checked in place of the forecast value.

NOTE: The forecast category statistics and observed category statistics are computed independently. It is not possible to acquire statistics for when the pairs are within a specific forecast category and a specific observed category, simultaneously.

14.0 Output Files

The IVP Batch Program can generate the following types of output files (recommended file extension is given in parentheses):

File Type	Description	Generating Action
Output Statistics File (.txt or .dat)	Contains rows of output statistics sectioned by verification group. There is one row per combination	CALCSTATS (defined with

	of location (if broken down by location), analysis interval, issuance time-of-day interval, lead time interval, forecast type source (if broken down by type source), and forecast/observed category.	OUTPUT_FILE command)
Pairs File (.txt or .dat)	Contains output pairs used in computing statistics for each group within the output statistics file.	CALCSTATS (defined with PAIRS_FILE command)
Output Image File (.png, .jpeg, or .jpg; required)	Plot image in either PNG or JPG format. If none of the extensions are specified for a file name, .png will be added.	GEN_GRAPH
Image Data File (.dat)	Data corresponding to plot image. If .dat is not specified as the extension file name, it will be added.	GEN_GRAPH

A description of each output file type is provided below.

14.1 Output Statistics File

An output statistics file stores the statistics computed when a CALCSTATS action is called. The output file has sections devoted to each group for which statistics are computed. The sections have two parts:

- Header: Provides a listing of the parameters that make up the definition of the verification locations and verification group. Also provides a line listing the column headers for the data.
- Data: Pipe (|) delimited tabular data of this format (column name is given in parentheses):
 - Location id (LOCATION): 8 character id; same as the lid in the archive database. "ALL" if all locations are lumped together.
 - Analysis subinterval start time (START): In the format "CCYY-MM-DD hh:mm:ss" (GMT is assumed).
 - Analysis subinterval end time (END): In the format "CCYY-MM-DD hh:mm:ss" (GMT is assumed).
 - Analysis month (MONTH): The numerical month (1-12) indicating what month is being analyzed if the ANALYSIS_INTERVAL command is set to MONTHLY. It is -999 if not.
 - Forecast type source (FCST TS): "ALL" if all type sources are lumped together.
 - Issuance time-of-day subinterval start hour (ISSUANCE START): In the format "#hr"; "-999hr" if no issuance intervals were defined for this group.
 - Issuance time-of-day subinterval end hour (ISSUANCE END): In the format "#hr"; "-999hr" if no issuance intervals were defined for this group.
 - Lead time subinterval start hour (LEAD START): In the format "#hr"; "-999hr" if no lead time intervals were defined for this group.
 - Lead time subinterval end hour (LEAD END): In the format "#hr"; "-999hr" if no lead time intervals were defined for this group.
 - Category variable (CAT VAR): "fcst" if the statistics for the output line were calculated for a forecast category, "obs" for observed.
 - Category number (CATEGORY): The number identifying the category. The categories for each variable are numbered 1, 2, etc.
 - Number of samples (NUM SAMPLES): The number of forecast-observed data pairs found satisfying the conditions given by the preceding columns.
 - Statistics: There is one column per statistic computed. See Appendix B of the *Interactive Verification Program User's Manual* for a description.

14.2 Pairs File

A pairs file provides the records of the vfypairs and vfyprocpairs tables used in the computing the statistic for a group. The pairs file has sections devoted to each group for which pairs exist with the file. The sections have two parts:

- Header: Provides a listing of the parameters that make up the definition of the verification locations and verification group. Also provides a line listing the column headers for the data.
- Data: Pipe (|) delimited tabular data of this format:
 - Location id: 8 character id; same as the lid in the archive database.
 - Forecast integer duration: integer duration code.
 - Forecast SHEF type source: 2 character type source code.
 - Forecast SHEF extremum: 1 character extremum code.
 - Forecast SHEF probability: 1 character probability code.
 - Forecast basistime: In the format "CCYY-MM-DD hh:mm:ss" (GMT is assumed).
 - Forecast validtime: In the format "CCYY-MM-DD hh:mm:ss" (GMT is assumed).
 - Forecast value: number.
 - SHEF Physical Element: identical for both forecast and observed value; 2 character physical element code.
 - Observed SHEF duration: 1 character duration code.
 - Observed SHEF type source: 2 character type source code.
 - Observed SHEF extremum: 1 character extremum code.
 - Observed obstime: In the format "CCYY-MM-DD hh:mm:ss" (GMT is assumed).
 - Observed value: number.
 - Observed SHEF quality code: Hardwired to 'Z' for backward compatibility with pairs files from early verification software.
 - Pairs quality code: The quality bit-code for the pair: 1879048191 if good; 0 if bad. Other values are possible, but if above 1610612736, it is assumed to be good; if between 1073741824 and 1610612736, it is assumed to be questionable; if below 1073741824 it is assumed to be bad.

14.3 Output Image File

An output image file is a PNG or JPEG format file that displays the plot specified by the IVP Batch Program parameters, exactly as it would appear in the IVP GUI. The extension of the file MUST be either .png, .jpg, or .jpeg. It is the extension which IVP uses to determine the format of the file.

14.4 Image Data File

An image data file provides the statistics displayed in an output image file, exactly as it would appear in the **IVP Statistics Data Viewer** of the IVP GUI (see the *Interactive Verification Program User's Manual*).

15.0 Examples

The following examples illustrate how to construct batch files to compute statistics for specific locations. Refer to the *IVP Examples* document for examples that involve generating graphics.

Example 1: A batch file to produce verification statistics for October 2003 for the national verification project. The locations used have lids of AAAAA,BBBBB,CCCCC, and DDDDD. The files to create will have prefix "stats_OCT_2003" and will be in the default output directory \$(vsys_output)/\$(LOGNAME).

```

START_TIME = "20031001 00:00:00"
END_TIME   = "20031031 23:59:59"

# Do the statistics...
NATLSTATS  = "AAAAA,BBBBB,CCCCC,DDDDD"

# To do national statistic for every location with a national flag of
# 'Y' in the vfyruninfo table (set via the Vfyruninfo Editor) do this:
# NATLSTATS = NATIONAL

```

Example 2: A batch file to produce verification statistics for October 2003 and for all locations defined within the vfyruninfo table with slow or medium response times. Only overall statistics are needed; categories are not needed. The statistics to produce are the error statistics ("ERRORS") and will be written to the default file. Persistence and normal forecasts are included in the verification.

```

DEF_LOC      = ALL

START_TIME   = "20031001 00:00:00"
END_TIME     = "20031031 23:59:59"
RIVERRESPONSE = "SLOW, MEDIUM"
DEF_GRP      = ALL

# Do the statistics...
# For overall statistics, I can pass any of the three
# acceptable values as the first argument.
CALCSTATS    = obs,ERRORS

```

Example 3: As example 2, except that for location AAAAA (with forecast type source FF), which is a slow response time river, create two categories for the observed value: $(2.0 < x \leq 15.0)$, and (above 15.0). For all other locations, use the categories below flood stage and above flood stage. Also, add the categorical statistics.

```

# First, define all locations.
OBS_CAT      = "MIN, 1.0*FS ,MAX"
DEF_LOC      = ALL

# Override the categories for lid AAAAA by calling
# DEF_LOC again for it only.
OBS_CAT      = 2.0,15.0,MAX
FCST_TS      = FF

```

```

DEF_LOC      = AAAAA

FCST_TS      = <default> # This is to undo the FCST_TS setting above.
                # This token is used within the DEF_GRP
                # below.

# Now define the general group parameters and build the group.
START_TIME   = "20031001 00:00:00"
END_TIME     = "20031031 23:59:59"
RIVERRESPONSE = "SLOW, MEDIUM"
DEF_GRP      = ALL

# Do the statistics...
CALCSTATS    = "obs, errors, catstats"

```

Example 4: Same as example 3, except produce statistics for lead times between 0 and 72 hours and at 24 hour time steps and only for locations AAAAA (with forecast type source FF), BBBBB (with forecast type source PE), and CCCCC (with forecast type source FE). Also, send the output to a file with name "outfile.txt" relative to the default output directory \$(vsys_output)/\$(LOGNAME). Finally, do not include persistence forecasts in verification. The file should be appended to as it may already contain previously produced statistics.

```

# First, define the BBBBB and CCCCC locations for FCST_TS of FE. Do
# not define the persistence forecasts.
OBS_CAT      = "MIN, 1.0*FS,MAX"
FCST_TS      = FE
DEF_LOC      = BBBBB,CCCCC

# Now define the AAAAA location. Do not define the persistence
# forecasts.
OBS_CAT      = 2.0,15.0,MAX
FCST_TS      = FF
DEF_LOC      = AAAAA

# Now define the general group parameters.
START_TIME   = "20031001 00:00:00"
END_TIME     = "20031031 23:59:59"
RIVERRESPONSE = "SLOW, MEDIUM"
LEADTIME_START = 0
LEADTIME_END   = 72
LEADTIME_STEP  = 24
FCST_TS       = ALL # Persistence forecast locations are not defined
                # and are therefore not included.

# Do the statistics...
DEF_GRP      = AAAAA,BBBBB,CCCCC
OUTPUT_FILE   = "outfile.txt,a"
CALCSTATS    = "obs, errors, catstats"

```

Example 5: Same as example 4, except include persistence forecasts in the computations and force the statistics to be computed separately for the non-persistence and persistence forecasts.

```
# First, define the BBBB and CCCC locations.
OBS_CAT      = "MIN, 1.0*FS,MAX"
FCST_TS      = FE,FR
DEF_LOC      = BBBB,CCCC

# Now define the AAAAA location.
OBS_CAT      = 2.0,15.0,MAX
FCST_TS      = FF,FR
DEF_LOC      = AAAAA

# Now define the general group parameters.
START_TIME   = "20031001 00:00:00"
END_TIME     = "20031031 23:59:59"
RIVERRESPONSE = "SLOW, MEDIUM"
LEADTIME_START = 0
LEADTIME_END   = 72
LEADTIME_STEP  = 24
FCST_TS       = ALL
BREAKDOWN_BY_FCSTTS = ON

# Do the statistics...
DEF_GRP      = AAAAA,BBBB,CCCC
OUTPUT_FILE  = "outfile.txt,a"
CALCSTATS   = "obs, errors, catstats"
```
