WREG Shell Program Version 1.0
March 2011

Modified April 2013 for slight changes in installation.

By Gregory E. Granato, Julie Kiang and Ken Eng

The WREG shell program is designed to facilitate use of the Weighted-Multiple-Linear Regression Program (WREG) by Eng and others (2009). The WREG shell program was written in Visual Basic 6.0 at the behest of the U.S. Geological Survey Office of Surface Water in response to user comments during WREG training sessions. The purpose of the program is to provide a graphical-user interface to augment file-handling features. The WREG shell program is designed to improve the file-handling features of WREG, by doing simple checks on the content and structure of input files, and by providing lists of errors and warnings for each input file. The WREG shell program also augments the file-handling features by copying the WREG runtime files to a user-defined directory and running WREG using the shell feature of Visual Basic (Siler and Spots, 1998). Use of the WREG shell program allows the main WREG program to be installed in a single file location, which can be separate from the location of input files.

The scope of the program is to sit outside the peer-reviewed and USGS approved version of WREG and run the program rather than being incorporated into the MATLAB runtime interface. The WREG shell program provides features that are not as readily included in the MATLAB runtime interface, which was selected for its ability to process the complex matrix algebra necessary for the various regression models implemented in WREG (Eng and others, 2009).

PROGRAM INSTALLATION

The WREG shell program is installed using a standard Windows installation package. It must be installed after WREG is installed. To install the package copy the "setup.exe", "Setup.Lst", and "WREGshellv01.CAB" files to a temporary folder on your computer. Double left-click the "setup.exe" file and follow the instructions provided by the installation wizard. You must have administrative rights to install this program on your computer.

The WREG shell installation package comes with a copy of the main WREG executable, and requires this file to run WREG. The WREG executable file must be located in the directory where the WREG shell was installed. If a new version of WREG is available, the new executable may be placed in this directory. In addition, appropriate changes must be made to the WREGshell.ini file, so that the program knows to use the updated WREG executable. These edits can be made with administrative rights by simply swapping the name of the old executable file for the new executable file.

To uninstall the WREG shell program, standard uninstall program functions found in the Windows control panel should be used.

FORM CONTROLS

The WREG shell program is written in Visual Basic 6.0 and uses a graphical user interface with controls that should be somewhat familiar to a Windows User (Fig. 1). The controls on the WREG shell interface read from left to right and from top to bottom. The four primary steps for running the shell are numbered and the program highlights progress from command button to command button as you work through the process. The four steps are:

Step 1: Specify Input Filenames

Step 2: Check Input Files

Step 3: Load WREG Files and Run WREG

Step 4: Open the Output Directory to See Results

All controls on the form are activated by left-clicking once on the control. All controls are enabled throughout the process, but the WREG shell will not let you skip a step.Clicking a button out of sequence will produce an error message prompting you to select the proper command button. Most controls will be described in the context of their use when stepping through execution of the WREG shell program, but the more general controls will first be discussed here. The form has 8 help buttons, each denoted by a question mark ("?"), that launch message boxes with basic information about the adjacent control. The form has 6 check boxes, which indicate the current status of available selections. If a box is checked the WREG shell will use that option as the default for file processing. If you change an option during processing, you will need to restart the process to insure that the shell completes all pre-processing tasks correctly. The form has four boxes that are used to visually group related controls.

A.

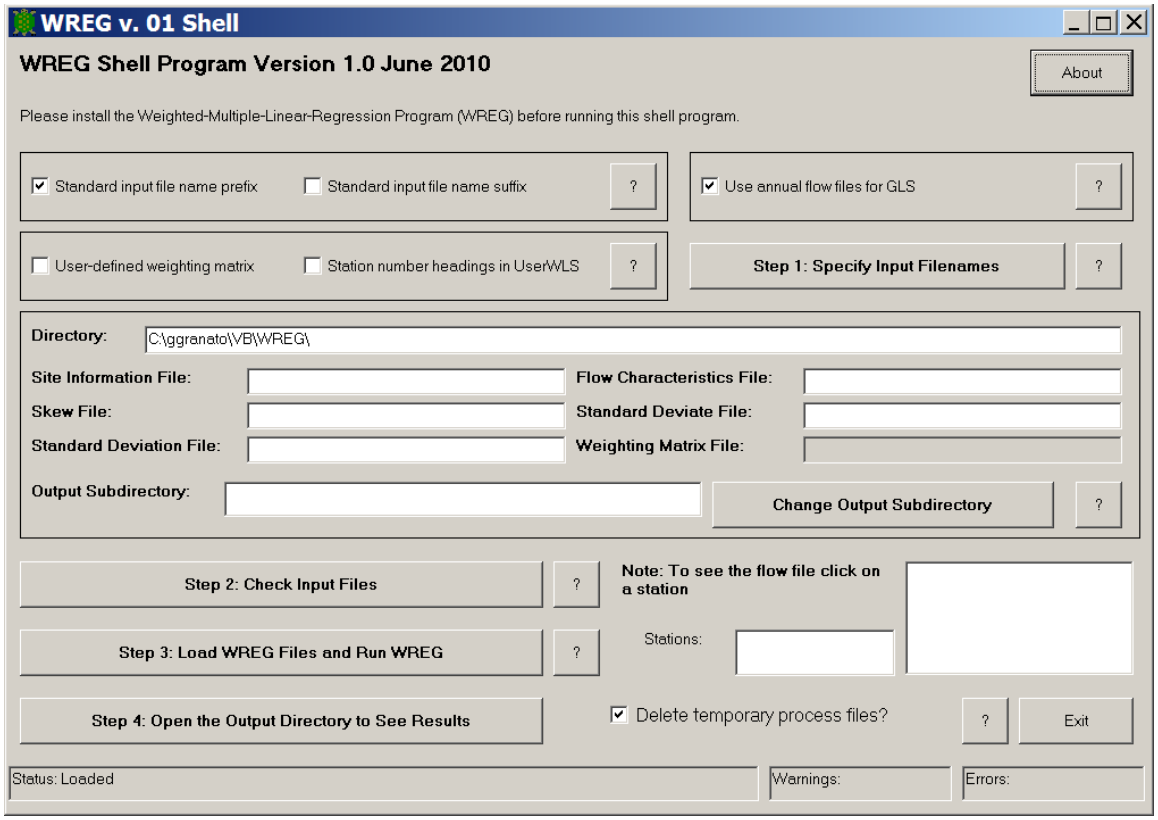


Figure 1. Examples of the WREG shell main-menu form ***A*** upon launching the program and ***B*** after specifying input-filenames.

B.

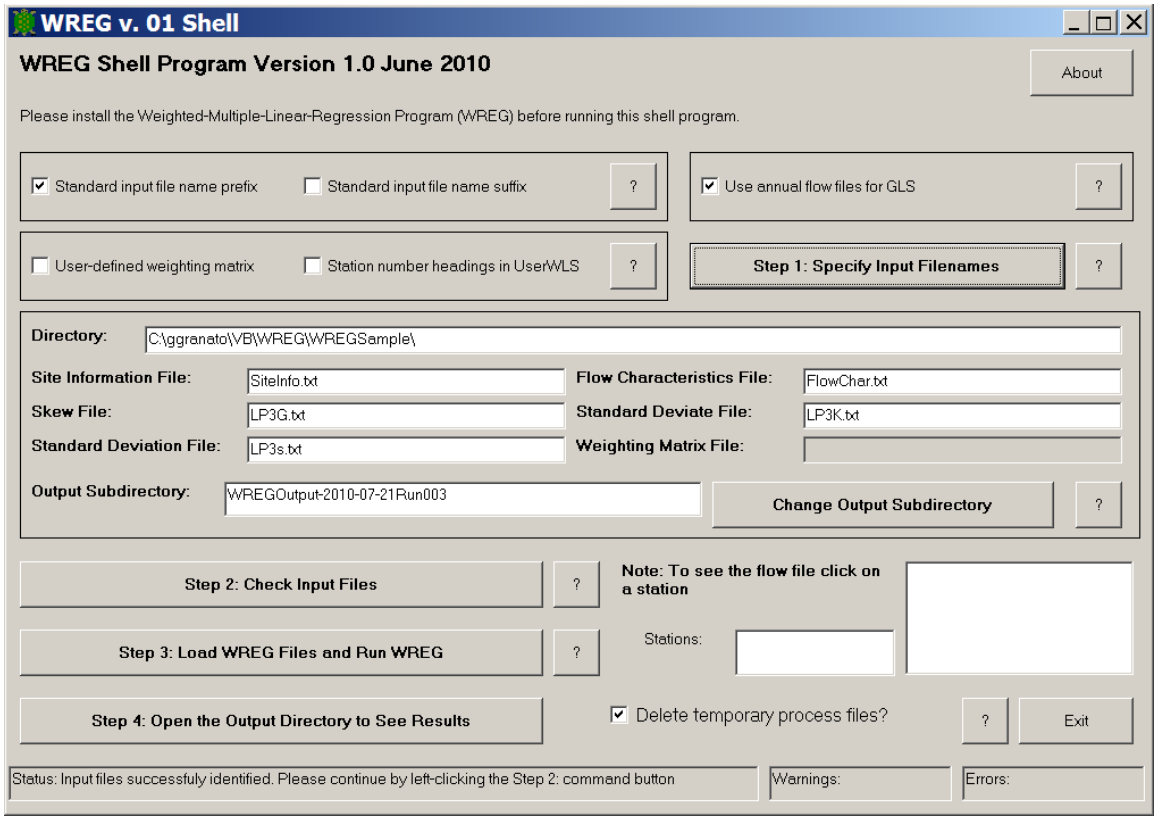


Figure 1. Examples of the WREG shell main-menu form ***A*** upon launching the program and ***B*** after specifying input-filenames.

There are three status boxes on the bottom of the form. The largest status box in the lower left will indicate the progress within the WREG preprocessing sequence. The status box in the bottom center indicates the number of warnings encountered during preprocessing and the status box in the lower right indicates the number of errors encountered during preprocessing. Additional warnings and error alerts are given in popup windows when the WREG shell checks input files. In general, warnings are less serious than errors, but may impact the WREG run. Errors are problems that are likely to impact the accuracy of WREG results, or to prevent the program from running at all. If the shell identifies problems requiring correction, you will need exit the WREG shell, edit your input files, and try again.

INPUT-FILE SELECTION

The shell improves the file-handling features of WREG because it allows you to select an input-file directory, select input files with various names, and select an output subdirectory. These file-handling features help you do multiple WREG runs by allowing you to create different input files (with different file names) in the same working directory(for example, files that include a different set of streamgages).

Figure 1A shows the main interface form as it appears when the program is launched from the programs menu. The top-left input box lets you select either a standard input file name prefix (for example, MyAnalysis01SiteInfo.txt) or a standard input file name suffix (for example, SiteInfoMyAnalysis01.txt). The next input box lets you include a user-defined weighting matrix as described in the WREG manual (Eng and others, 2009) by clicking the check box. If you click the "Station number headings in USERWLS" check box it loads a matrix with station numbers in the first column and first row. The input box in the upper right lets you use the annual flow files for generalized least squares regression (Eng and others, 2009). The first command button in the upper right "Step 1: Specify Input Filenames" runs the file-specification process. Once the file-specification process is complete the input directory and filenames will be displayed in the input-file textboxes (fig. 1B).

The first step in the file-specification process is to navigate to your data-file directory and select the site information text file. WREG shell prompts you for the directory and file using the common dialog file specification form (fig. 2) that is standard for most windows applications. The default name is "SiteInfo.txt" but you may select a file with a standard prefix, a standard suffix, or with any random name. Automatic file-name selections for other input files (FlowChar.txt, LP3G.txt, LP3K.txt, LP3s.txt, and UserWLS.txt) will be based on this initial filename. For example, if the siteinfo file is named "MyAnalysis01SiteInfo.txt" and the standard input file prefix checkbox is selected (fig. 1A) the program will look for a flow characteristics file named "MyAnalysis01FlowChar.txt". Similarly, if the file is named "SiteInfoMyAnalysis01.txt" and the standard input file suffix checkbox is selected (fig. 1A) the program will look for a flow characteristics file named "FlowCharMyAnalysis01.txt ". If either checkbox is selected and the standard "SiteInfo.txt" filename is used WREG shell will look for the standard input filenames. If all the filenames follow the user-specified format the rest of the input files will be identified automatically for you. If any one of these files (or all of the files) is not found using the specified prefix or suffix then the program will prompt you for each unidentified file using the common dialog file specification form on a case-by-case basis. The common dialog file specification form in WREG shell has been programmed so that the type of file is printed on the title bar and the default WREG filename is in the filename window of the form. If you cause a file-specification error or you refuse to identify one of the input files WREG shell will not allow you to proceed to the next step.

The files that contain the annual-time series of interest for each streamflow-gaging station (the "USGS########.txt files) are treated differently than the other files. These files are not identified with the Step 1 command button, but are identified using the "Step 2: Check input files" command button. These files may have any name as long as the text "USGS" and the station number is included somewhere in the filename. The WREG shell was developed with the assumption that you will only have one annual-time series file for each station in the specified input directory.

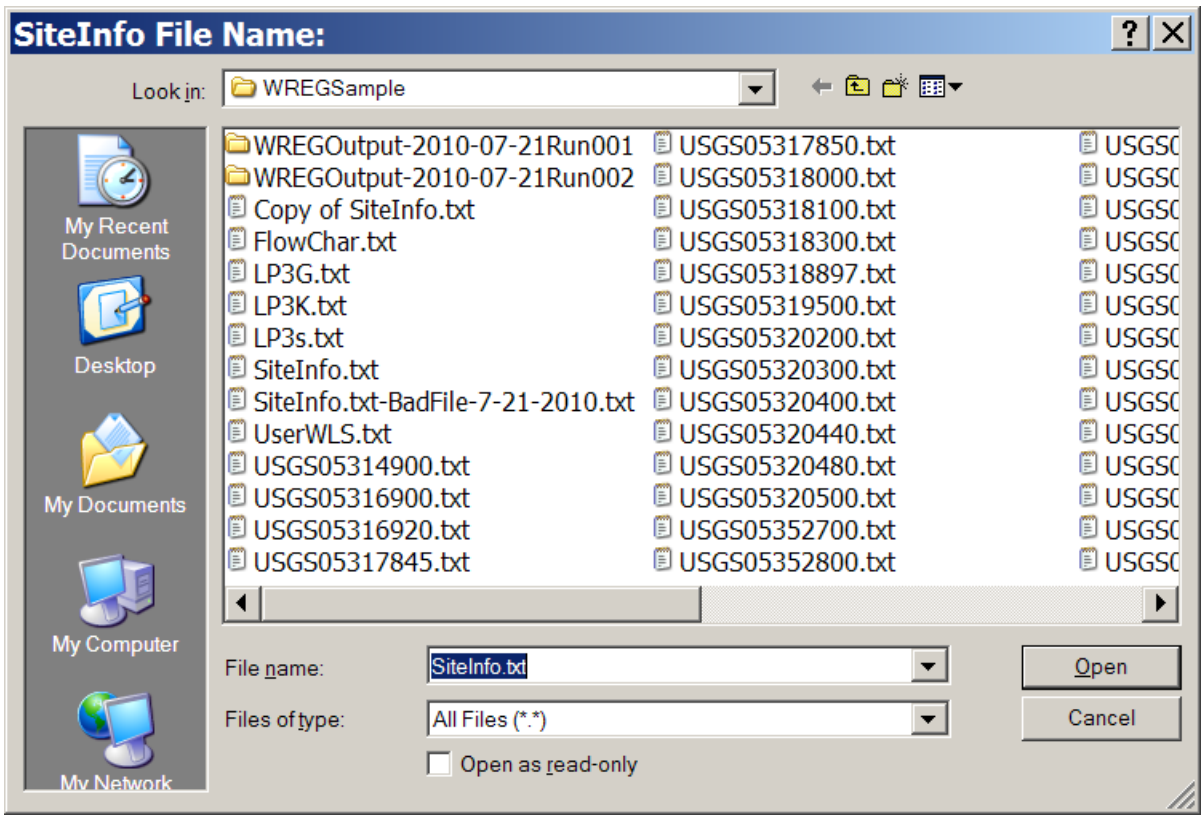


Figure 2. Example common dialog file specification form with the title bar and default file name indicating that the shell is looking for the site information file (SiteInfo.txt).

OUTPUT SUBDIRECTORY

The "Change Output Subdirectory" command button in the file-name box (fig. 1) lets you change the name of the output subdirectory. The default is to place output in a subdirectory of the directory in which input files are located. This default name for this subdirectory is "WREGOutput-" the date in YYYY-MM-DD format, "Run" and the run sequence number for a given day (for example: WREGOutput-2010-07-21Run001 is the first run on July 21, 2010). A directory name can be up to 100 characters long. The WREG shell prompts you for a subdirectory name using an input box (fig. 3). You can specify a name using any keyboard characters that are allowed by the Windows operating system, however, the program will trim any leading or trailing spaces and replace any internal spaces with underscores. The WREG shell will not allow you to overwrite an existing subdirectory. The WREG shell checks the target directory and reverts to the standard folder name if the subdirectory already exists. The WREG shell also reverts to the standard subdirectory name if you select cancel. You can rename the output subdirectory at any time after specifying the input filenames and before running the "Step 3: Load WREG Files and Run WREG" command button.

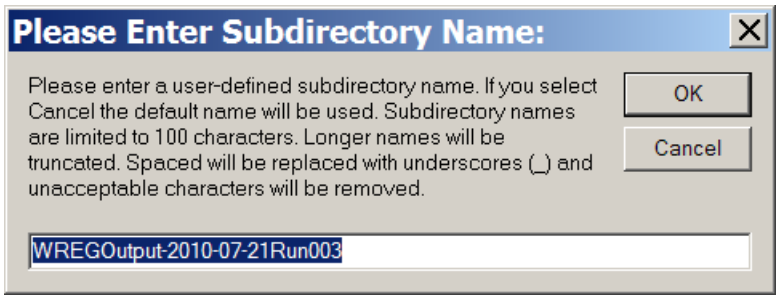


Figure 3. Example of an output-subdirectory input-box.

CHECKING INPUT FILES

The next step in the WREG shell process is checking input files. In this step the program does rudimentary checks on the content and/or format of each of the WREG input files. The quality assurance-quality control (QA/QC) process is designed to identify typographical errors rather than hydrological errors on the assumption that you have done the prerequisite analysis correctly. The WREG shell generates an error for any problem that is likely to have an adverse effect on a WREG analysis and a warning when something is found that does not meet input specifications. It also generates a warning for critical items that may be optional such as the regional skew value. Issues identified as either warnings or errors can cause WREG to produce unexpected results, so both should be checked.

All the WREG input files are tab delimited so the shell program reads each line and parses the line using the tab character. If an input file has two tabs in a row the shell will identify this as an empty column. The shell program uses the Visual Basic Trim function to delete leading and trailing spaces from input values so a column with one or more spaces without any data also will be identified as an empty column. WREG requires an input value for each tab delimited column so this condition will cause an identified error in the QA/QC output.

The shell program launches a message box (fig 4.) warning you that input files may be overwritten in the quality assurance-quality control (QA/QC) process. A copy of the original input file is saved in the input directory. Nevertheless, the message box provides the opportunity to backup input files. The default selection is "Yes" to backup input files, which will exit the QA/QC process. You can either do the backup while the shell program is waiting for a response or select yes to exit the QA/QC process and make the backups.

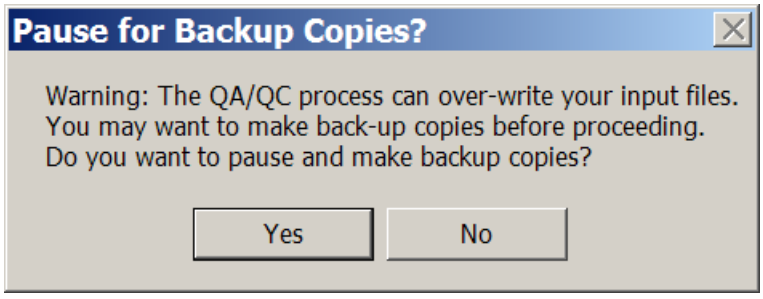


Figure 4. The backup message box.

Site Information File

The shell program checks the site information file first because many of the QA/QC tests are based on the content of this file. The site information file has a heading line and then one line of data for each streamgage included in the analysis (Eng and others, 2009). The site information file has eight standard columns, which are followed by one or more columns containing the value of each explanatory variable associated with each streamgage. The shell program checks to ensure that there are 9 or more heading columns in the input file. It checks that the heading name of the 8 standard columns match exactly the names given in the user manual. It counts the number of stations in the input file and it checks that the data line for each station has the same number of columns as the heading line. The input file must include data for two or more streamgages (because it takes at least two points to define a line). The program checks the number of columns in comparison to the heading row and generates an error if there are too many or too few. The program checks to ensure that the station number is an integer and counts the number of digits in the number; it generates an error if the value is not a number. The program checks the latitude and longitude to ensure that they are decimal numbers; it generates a warning for each problem it detects. The program checks the annual series number to ensure that it is an integer; and generates an error if the value is not a number. The program checks the

"Zero-1;NonZero-2" value to ensure that it is either 1 or 2 and generates a warning if it is not (note: this should be converted to an error if WREG is changed to incorporate this variable). The program checks the FreqZero value to ensure that it is a decimal number greater than zero and generates a warning if the value does not meet these criteria. The program checks the regional skew value to ensure that it is a decimal number and generates a warning if the value does not meet this criterion. The program checks the "Cont-1:PR-2" values and generates an error if the values are not integers equal to 1 or 2.

The station numbers in the site information file are used as the basis for identifying all data in the other input files and for identifying the annual-time series data files. The WREG shell identifies four potential problems that result in non-standard station numbers and spawns a station number correction form to address these problems (fig. 5). Spreadsheet programs, such as Microsoft Excel, commonly are used to create and edit tab-delimited text files. However, such programs will strip leading zeroes from input values that are recognized as numeric data. You can avoid this problem by formatting the column as text before entering station numbers in a spreadsheet. Data-processing programs (or persons compiling data) also may add a single quote (') or other non-numeric character to preserve leading zeros in station numbers.

The WREG shell identifies non-numeric leading characters and flags station numbers with this problem for automatic correction. The number of station numbers with non-numeric leading characters is displayed in the upper-left textbox of the station number correction form (fig. 5) and the "Strip Leading Characters From Station Number" command button will fix these problems and reset the number to zero. If this problem does not exist in the site information file the number will be zero and the command button will be disabled as indicated by the grayed out appearance of the text on the command button in figure 5. Nonnumeric characters also may inadvertently be added to station numbers in the data compilation process. The WREG shell also looks for these problems, flags station numbers with this problem, and facilitates automatic correction. The number of stations with this problem is displayed in the second text box in the upper right and these problems may be fixed using the "Strip Internal Characters From Station Number" command button. This button also will be disabled if station numbers do not have internal non-numeric characters or if these problems have been fixed.

Most USGS streamgages have leading zeros in the station numbers and are eight numbers long. The WREG shell identifies potential problems with leading zeroes by identifying station numbers that are seven digits long. The number of these suspect station numbers is listed in the third textbox in the upper-right corner of the form. These problems will be corrected automatically using the "Add Leading Zero to Seven Digit Stations" command button (fig. 5). If one or more seven-digit station numbers are detected in the site information file then all the 8 to 14-digit station numbers will be displayed in the listbox in the lower left of the form (fig. 5) so that the user can determine whether or not to add leading zeros to these stations as well. All 8-14 digit station numbers are displayed because some numbers may be correct (they should not start with a leading zero). For example, some stations are identified by the 14 digit ground-water site

inventory (GWSI) number, which identifies the station by latitude, longitude, and sequence numbers. Also, some studies may have valid surface-water station number with and without leading zeros. You may select one station numbers (by clicking on it), a block of station numbers (by clicking on the first one, holding the shift key, and clicking on the last one) or multiple station numbers that are not in a group (by holding the Ctrl key and clicking on individual station numbers). In any case, clicking on the "Fix It" button will add a leading zero to the station number(s). These numbers will be corrected in the list box when you click the "Fix It" button. You can select numbers and correct them until you decide to close this form. The WREG shell program, however, will not let you add more than one leading zero to a site number.

If errors in the site information file are not resolved the WREG shell program will launch an information form with a list of errors and warnings (fig. 6).

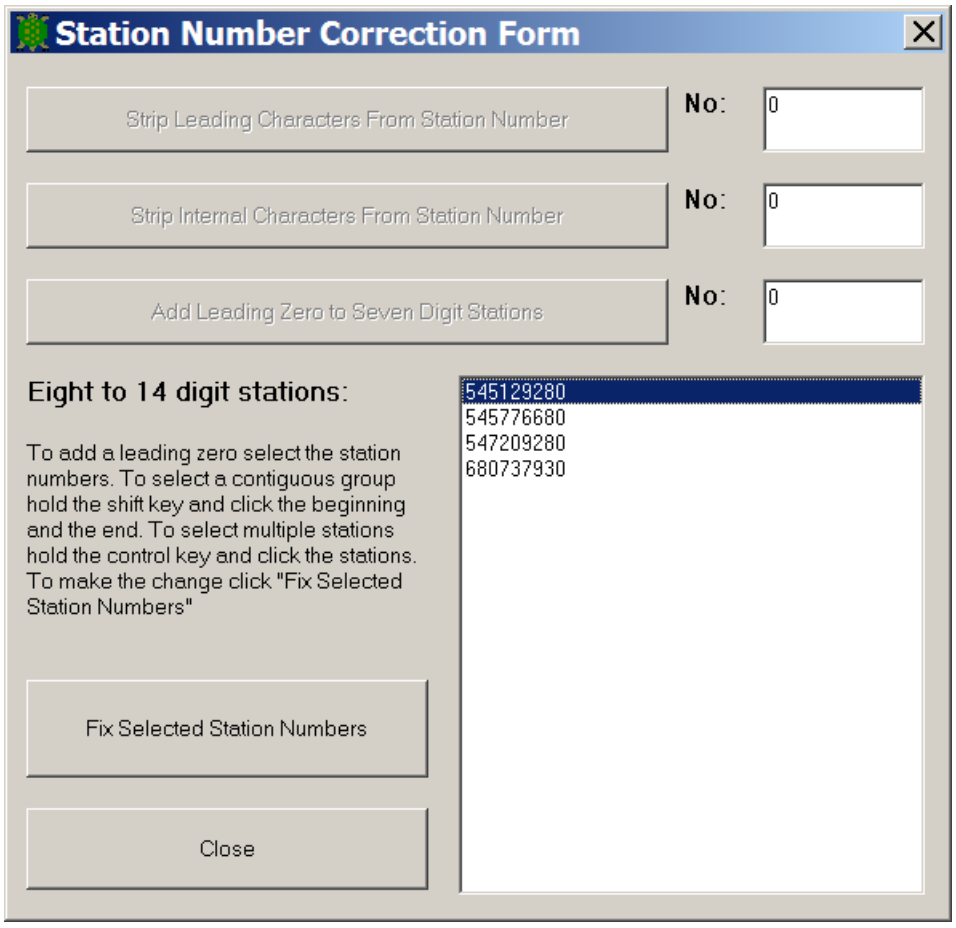


Figure 5. Example of the station number correction form.

Flow Characteristics File

The shell program checks the flow characteristics file (FlowChar.txt in Eng and others, 2009) to ensure that it has the same number of lines as the site information file, that the station numbers are the same as in the site information file, that it has more than one column, and that each line has the same number of columns as in the heading line. The flow characteristics file is the second file to be checked because this file defines the dependent variables that are used within the Log-Pearson Type III input files (LP3G.txt, LPGK.txt, and LP3s.txt). If the number of stations or the station numbers do not match this causes an error condition that precludes the use of WREG using the shell program. When checking the stations, the shell program examines the station numbers and will ignore differences caused by adding the leading zero. However, if all the station numbers exist, but are out of order the shell will give you the option of saving the file with the proper order and proceeding to use WREG. If you do not save the file the shell will not let you proceed. All errors and warnings will be displayed in the output information form (fig. 6).

Log-Pearson Type III Input Files

The shell program checks the to ensure that each of the Log-Pearson Type III input files (LP3G.txt, LPGK.txt, and LP3s.txt) has (1) the same number of lines as the site information file, (2) has the same station numbers as the site information file, (3) has the same number of columns as the flow characteristics file, and (4) has the same column headings as the flow characteristics file. When checking station numbers the WREG shell accounts for the possibility that one file has leading zeros in the station number, while the other may not so there will be no identified mismatches based on this factor. Error conditions 1-3 will create an error that will preclude use of WREG during the current run. Condition 4 creates a warning. You are advised to fix the input files and rerun the shell program to run WREG. All errors and warnings will be displayed in the output information form (fig. 6).

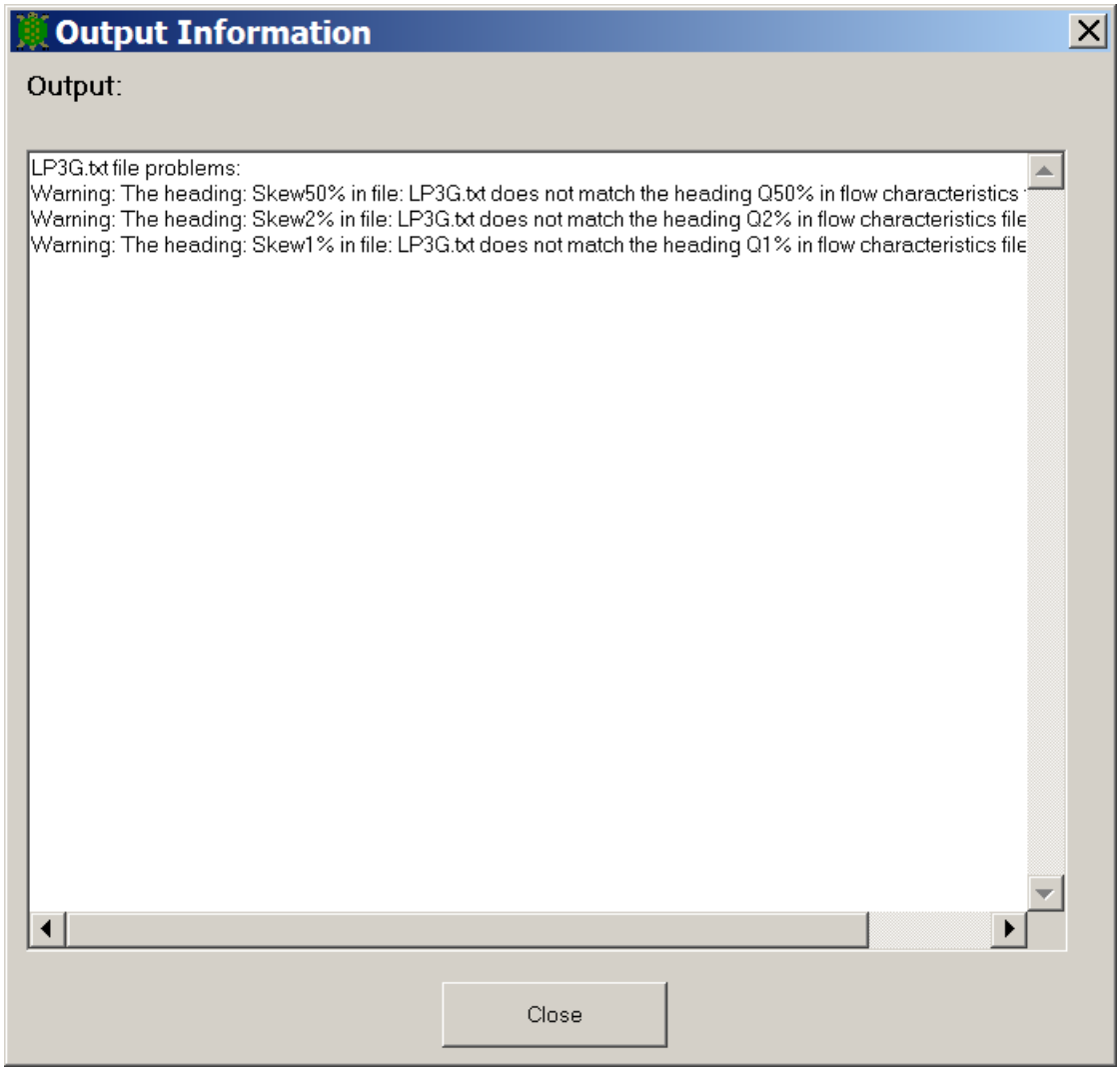


Figure 6. Example of the output information form with LP3G file-content warnings. The textbox with the output information is implemented so that the user can copy and paste the information into a text file to facilitate input-file corrections.

User-Defined Weighting Matrix File

The user defined weighting matrix file is an optional file that can be used to manually weight the stations in the analysis (Eng and others, 2009). WREG only reads a file named "USERWLS.txt" that does not contain header information (Eng and others, 2009). The WREG shell, however, lets you create a user defined weighting matrix file that does have station numbers in the first row and first column, which greatly facilitates manual assembly and review of the matrix in the file. WREG will only load and check the user defined weighting matrix file if the checkbox for this file (in the upper-left part of the form) is checked. It will identify heading lines as errors unless the "Station number headings in UserWLS" check box is checked.

The shell has to delete station numbers in the first row and first column and save the array to the standard file name ("USERWLS.txt") for use with WREG. Therefore it is necessary to use a different filename to preserve the file with headings. Using a standard prefix or suffix for all the input files will achieve this objective.

The WREG shell reads the user-defined weighting matrix file and checks (1) the number of rows, (2) the number of columns, (3) if there are station number headings, (4) the order of stations in the columns (5) that the station numbers in the columns match the station numbers in the first row, and (6) that the sum of diagonal values equals the number of stations once the sum is rounded to an integer value. All errors and warnings will be displayed in the output information form (fig. 6).

Annual-time series "USGS########.txt" Files

The "USGS########.txt" files contain the annual-time series of interest for each streamflow gaging station used in the analysis, which are used only for generalized least squares (GLS) analysis (Eng and others, 2009). The WREG shell concatenates the content of all these files into one master file named "USGSAnnualTimeSeries.txt" for use with WREG. The data files are concatenated in the same order as the list of station numbers in the site information file. The WREG shell will only process these files if the "Use annual flow files for GLS" checkbox is checked. The default is to not to check these files so you need to check this box if you are using these files for GLS regression. The WREG shell checks to see that an annual-time series file exists for each station number in the SiteInfo.txt file. The expected format of an annual-time series file is a tab-delimited file with the station name, the year, and a flow value. The WREG shell tests the number of columns in each file generating an error if there are less than three columns and a warning if there are more than three columns (WREG will only read the first three columns) in any line.

The shell does six tests on the year values:

(1) tests that the value is a number,

(2) tests that the year is not -9999, which is an error code,

(3) tests that the year is greater than zero (an error test),

(4) tests that the year is in the range from 1879 to the current water year,

(5) tests that the years progress in a logical order without any gaps, and

(6) tests that there are not multiple values for any given year in the annual time series.

The shell also tests the third column (containing the streamflow statistic) to see if it is a number.

As it does these tests, the shell program compiles a list of warnings and a list of errors that are identified. All errors and warnings will be displayed in the output information form (fig. 6). The number of annual-time series files and the length of each file increase the chance that there may be a large number of warnings and errors. A Visual Basic Textbox can hold only 64,000 characters so the information form is designed to break long error strings into blocks and to display them using two command buttons, "See Previous Part of Output" and "See Next Part of Output", to scroll backwards and forwards in the warning or error message array, respectively (fig. 7).

The WREG shell provides an interface to see the annual-time series input files. The "Stations" text box to the right of the command button (fig. 8) lists the number of stations in the site information file. You can view the contents of a site information file in the textbox on the output information form by clicking on the station number in the list box on the lower-right quadrant of the form above the "Exit" button.

The WREG shell rewrites all of the time series into a single file, called USGSAnnualTimeSeries.txt

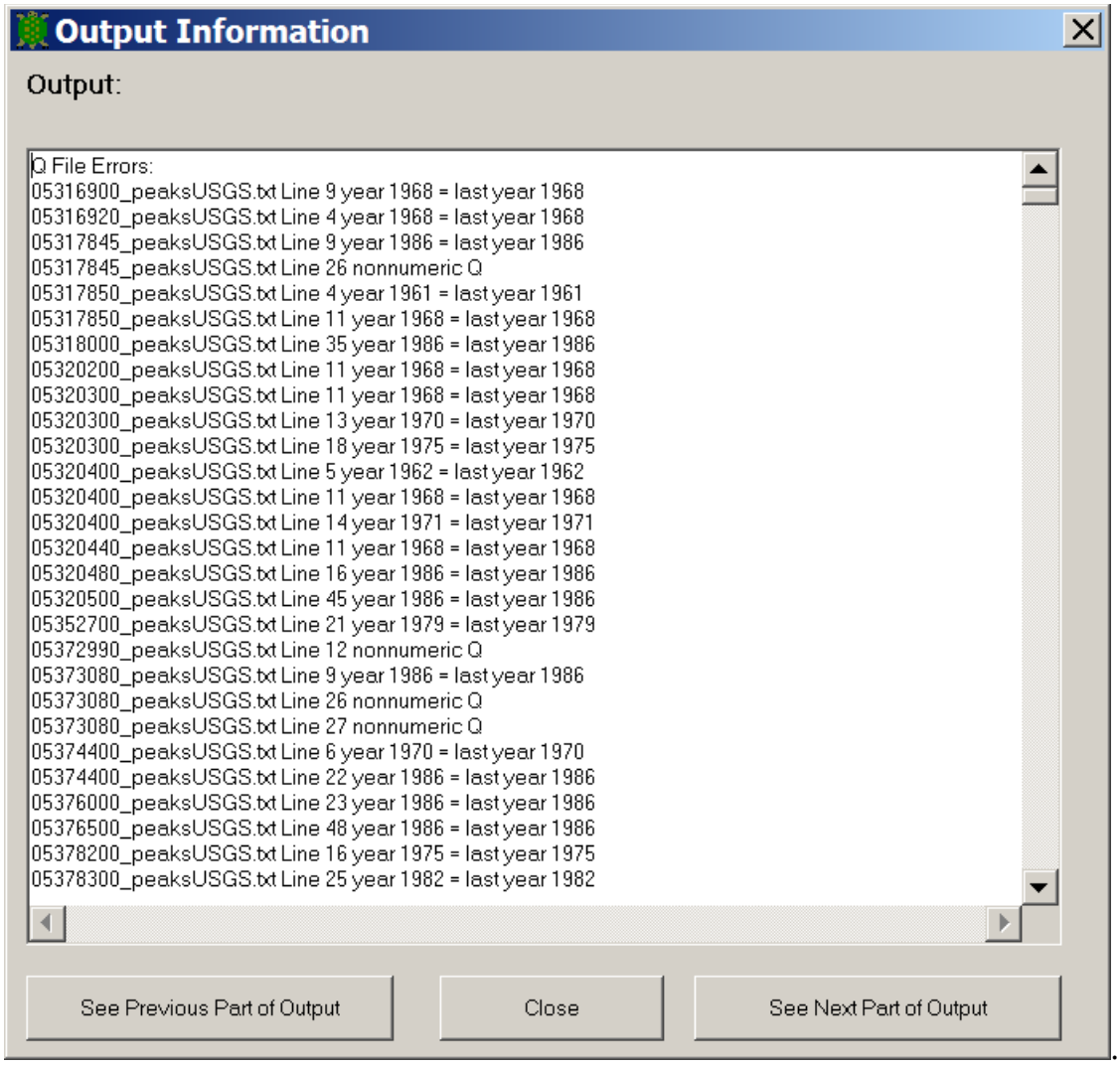


Figure 7. Example of the output information form with a large number of annual-time series errors. The information form will display the "See Previous Part of Output" and "See Next Part of Output" command buttons so you can scroll through a large set of error or warning messages. The textbox with the output information is implemented so that the user can copy and paste the information from each subsequent block of messages into a text file to facilitate input-file corrections.

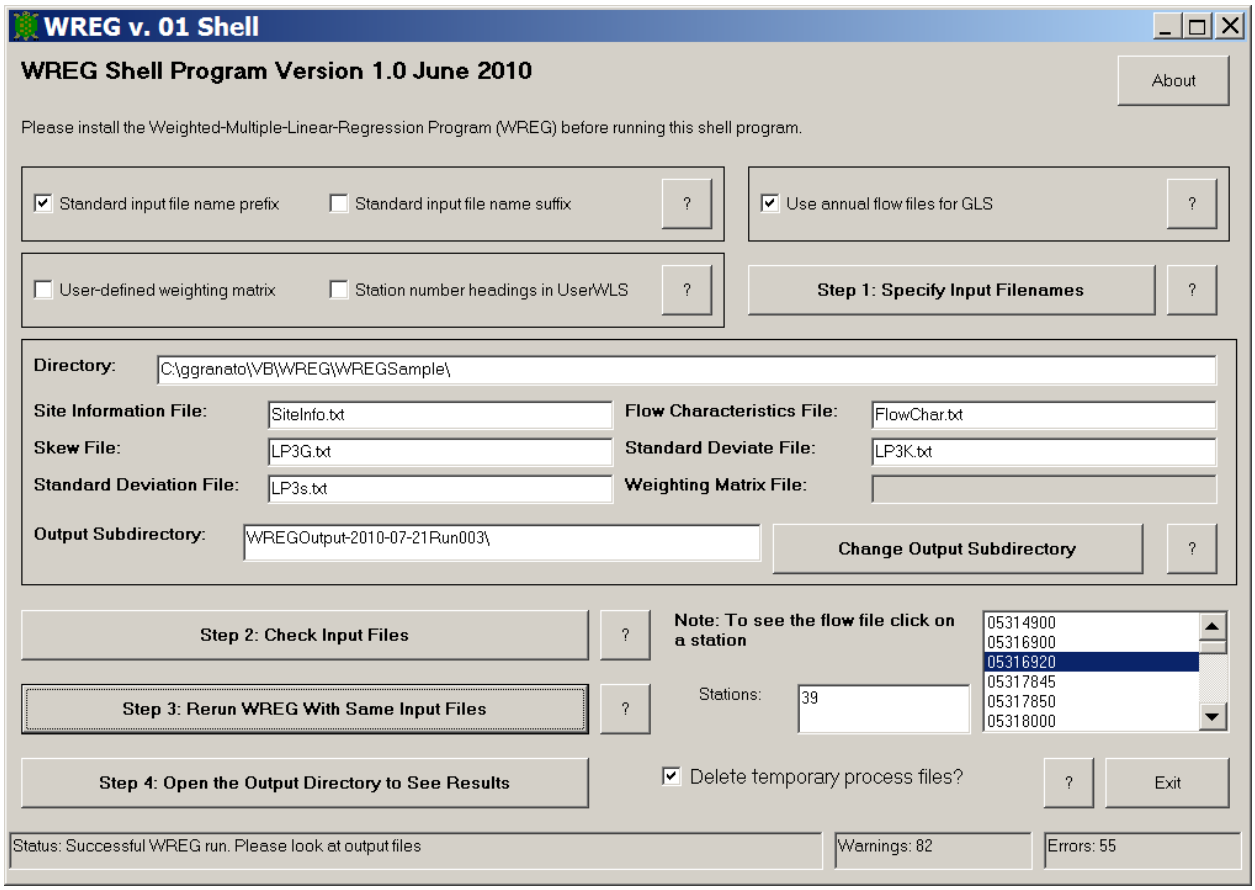


Figure 8. Example of the WREG shell main-menu form with the focus on station 05316902 in the flow-file list box. Left-clicking the list box will launch the information form, which will display the contents of the annual-time series input file.

RUNNING WREG, RERUNNING WREG, AND OPENING THE OUTPUT SUBDIRECTORY

After the input files are processed, click on the "Step 3: Load WREG Files and Run WREG" command button. If any errors were detected when processing the input files, the shell provides one alert (fig 9). Most errors will cause a problem with the main WREG program so these errors should be fixed.

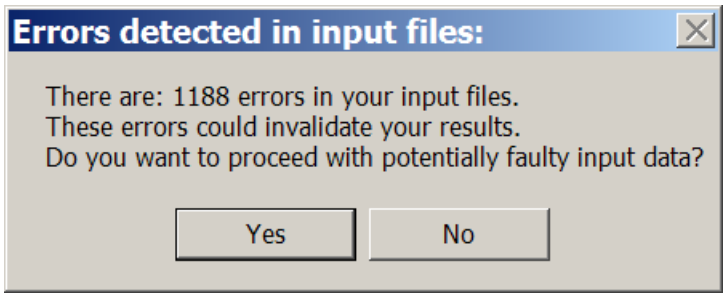
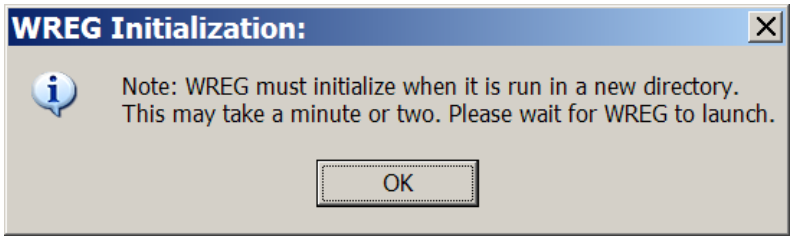


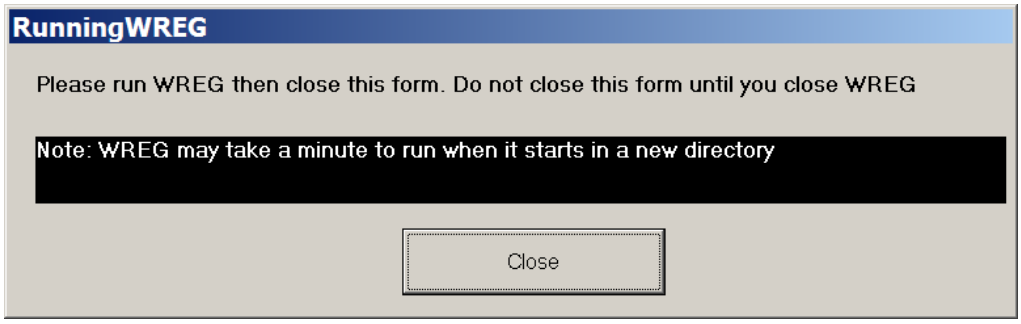
Figure 9. Example of final alert regarding errors in input files.

The program then copies WREG to the working directory and launches the program. At this point, the shell provides a message box to remind you that WREG initialization may take a moment (fig 10A). The WREG shell them minimizes the main input form to provide space for the WREG interface and launches a modal form that pauses operation of the WREG shell until you complete the WREG analysis (fig 10 B). If you try to close this form the shell will prompt you to ensure that you have completed your analysis before proceeding with the post-analysis cleanup (fig 10 C). Once you have completed the analysis the WREG shell will prompt you to save the input files with the output (fig 10 D). This option helps document input if you change variables or reselect stations for subsequent analysis. This option will copy all the standard input files except the individual annual-time series files. It does, however, save the concatenated "USGSAnnualTimeSeries.txt" file if this option is selected.

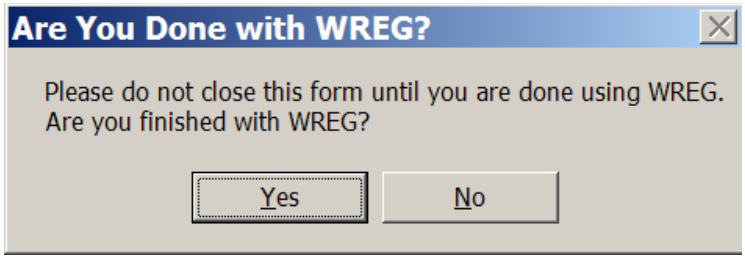
A.



B.



C.



D.

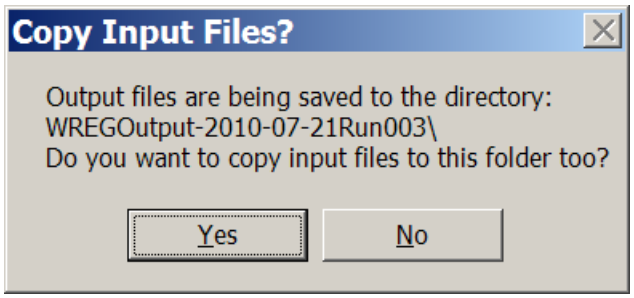


Figure 10. Examples of the message boxes and the form for running or rerunning WREG including *A.* the message box with the WREG initialization notice, *B.* the modal form that is used to pause the WREG shell for WREG analysis, *C.* the message box to confirm that you have completed the analysis, and *D.* the message box to prompt you to save the WREG input files with the output files.

The WREG shell is designed to help you try alternate regression models using the same input files. Once you complete the initial run, the caption on the Step 3 command button changes to "Step:3: Rerun WREG With Same Input Files" (fig. 8) and the program shifts into rerun mode. When you rerun the WREG shell you are prompted with the default name for the new output subdirectory and given the option to change this name (fig. 11). As with the first output subdirectory the name is comprised of the text "WREGOutput" the date in "YYYY-MM-DD" format and the run number (for example, "Run002"). If you choose to rename the output subdirectory, the subdirectory input box will appear (fig. 3) for you to input a new name. If the selected name exists, however, the WREG shell will use the default directory and you can change the name using the operating system. If you rerun WREG, the shell program uses the existing copy of the WREG program and support files in the target directory. If you have deleted these files the Shell will recreate them, but you will have to wait for WREG to reinitialize.

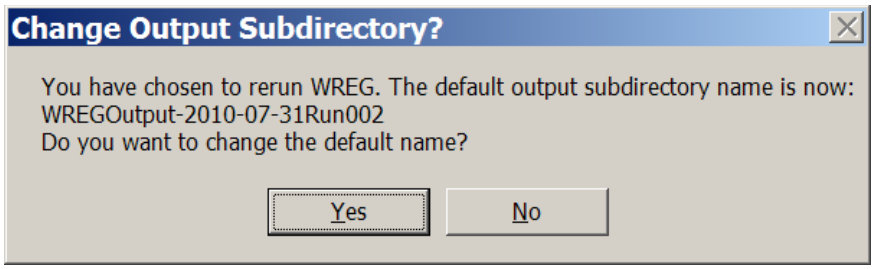


Figure 11. Example of the message box used to list the new output subdirectory name and prompt the user for a user-defined output subdirectory name.

After each run the WREG shell will allow you to open the output subdirectory using the "Step 4: Open the Output Directory to See Results" command button (fig. 8). This feature is useful to ensure that you are looking at the most current directory. It also facilitates the process of navigating through your directory structure. This command button opens a new window each time it is used for subsequent runs of WREG, as such it facilitates comparison of results from different WREG runs.

DELETE TEMPORARY PROCESS FILES AND EXIT WREG

The status of "Delete temporary process files" checkbox determines the actions of the WREG shell when you exit the program. If you exit the program before running WREG the program closes without deleting any files. If you have run WREG using the shell and the "Delete temporary process files" checkbox is selected, then the program will check to see if the WREG execution files exist and if they do the shell will delete them. However, if your working directory is also your WREG shell directory then the shell will not delete the WREG execution files because these are the master copies for use with the shell. Deleting the temporary files is the default selection so that multiple copies (and perhaps versions) of WREG will not be distributed throughout your working directories.

DISCLAIMER

The WREG shell software and documentation are made available by the U.S. Geological Survey (USGS) to be used in the public interest and in the advancement of science. This USGS software is free and in the public domain. The WREG shell is provided "as is," the authors, the USGS, and the United States Government have made no warranty, express or implied, as to accuracy or completeness and are not obligated to provide the user with any support, consulting, training or assistance of any kind with regard to the use, operation, and performance of this software nor to provide the user with any updates, revisions, new versions or "bug fixes". The user assumes all risk for any damages whatsoever resulting from loss of use, data, or profits arising in connection with the access, use, quality, or performance of this software (U.S. Geological Survey, 2008).

REFERENCES

Eng, Ken, Chen, Yin-Yu, and Kiang, J.E., 2009, User's guide to the weighted-multiple-linear-regression program (WREG version 1.0): U.S. Geological Survey Techniques and Methods, book 4, chap. A8, 21 p. (Also available at http://pubs.usgs.gov/tm/4A8.)

Microsoft Corporation, 2010, Support Statement for Visual Basic 6.0 on Windows Vista, Windows Server 2008 and Windows 7: http://msdn.microsoft.com/en-us/vbasic/ms788708.aspx

Siler, Brian, and Spotts, Jeff, 1998, Using Visual Basic 6. Special Edition: Que Corporation, Indianapolis, Indiana 860 p.