STORET Import Module (SIM)
Version 2.0.2
Training Manual

U.S. Environmental Protection Agency
Washington, DC
August 15, 2006

EPA Technical Support
1-800-424-9067
storet@epa.gov

Electronic copies of this document, along with many others, are available at
http://www.epa.gov/storet/

[This page intentionally left blank.]

Table of Contents

[This page intentionally left blank.]

Introduction

Welcome to the STORET Import Module (SIM) Version 2.0.2 Training Manual. SIM is a software program that helps users quickly and conveniently load data into STORET. SIM allows you to describe the format of the data being imported, ensures that the data are consistent with STORET's requirements, and migrates the data into STORET.

This document contains four tutorials that will help you learn how to do the following:

- **Tutorial 1**—Prepare your data for uploading to SIM and STORET by creating data files and import configurations

- **Tutorial 2**—Deal with SIM import errors

- **Tutorial 3**—Replace and back data out of SIM and STORET

- **Tutorial 4**—Import results data into STORET.

Each tutorial focuses on different aspects of SIM. There are descriptions of the functions and associated concepts, as well as step-by-step instructions for using SIM. You can follow these instructions on your PC because they are called out in text boxes, which are separate from the text. Typically, these boxes will also contain screen captures to illustrate what your screen should look like.

> This is an example of a text box that contains instructions for using SIM with the sample data. Buttons to click are in **bold face**. Menu items and other items to select are in quotation marks.

Included with this training manual is a "SIM_Data_Files" folder, which contains the sample files necessary to complete these tutorials. Take a moment to find it or place this directory on your computer using Microsoft (MS) Windows Explorer or a similar navigation program. Within the "SIM_Data_Files" folder are several MS Excel and text files, plus two folders named "User Created" and "Provided Configuration." Whenever you modify or create a file in these tutorials, be sure to save it to the "User Created" folder. The "Provided Configuration" folder contains files to assist you in completing Tutorial 4, if needed. Because you will be frequently navigating to the "SIM_Data_Files" folder location in these tutorials, you should write down the full path location on a handy piece of paper or change your default settings in SIM so that the program automatically navigates to this directory.

> To change your default settings, open "SIM202."
>
> From the menu bar select "Advanced," then "System Settings."
>
> Locate the DOCUMENT_IMPORT_PATH under the "Item" column in the table.

The "Value**"** column has a file location displayed beside the DOCUMENT_IMPORT_PATH item. Change this location to the pathway that routes to the "SIM_Data_Files" folder.

Another document that will help you understand SIM is the SIM User Guide. To find the guide, go to the "Start" menu on your computer, select "Programs," navigate to "SIM202," and select the "User Guide." This document assumes that you have reviewed that guide and the STORET Data Entry Module User Guide.

If you are a beginning user of STORET and SIM, it should take you about 4 to 5 hours to complete the five sections of this manual. If you are a more advanced user, it should take about 2 to 3 hours.

If you have questions about this tutorial, contact the U.S. Environmental Protection Agency's (EPA's) STORET assistance hotline at 1-800-424-9067 or by e-mail at STORET@epa.gov. You may also refer to EPA's STORET Web site at http://www.epa.gov/storet/.

Tutorial 1. Data Preparation

This tutorial will help you become familiar with the methods for creating STORET Import Module (SIM) import configuration files tailored to your data format. It will also help you prepare your data so that they can be imported into SIM. By the end of this tutorial you should be able to

- Export data from a Microsoft (MS) Excel spreadsheet to create a data file
- Create an Import Configuration file in which you specify the data delimiters
- Specify the fields in your data file and the order in which they appear
- Specify default values for fields
- Specify a data format for a field that requires formatted data in STORET.

In this tutorial, you will specify a data format for importing data for projects.

STORET Data Structure

To understand the way that SIM imports your data, it is important that you understand how data is structured and stored in STORET. The STORET data structure follows the way organizations report water quality monitoring. An organization has projects that require monitoring at specified sampling stations. Individuals at the organization go on monitoring trips to visit stations. At stations, the individuals conduct monitoring activities, which generate results. For the tutorial, the organization is called "DEMOTEST." We will be working on the data files for projects and results of this organization. The figure below illustrates how STORET is structured to follow the business of monitoring.
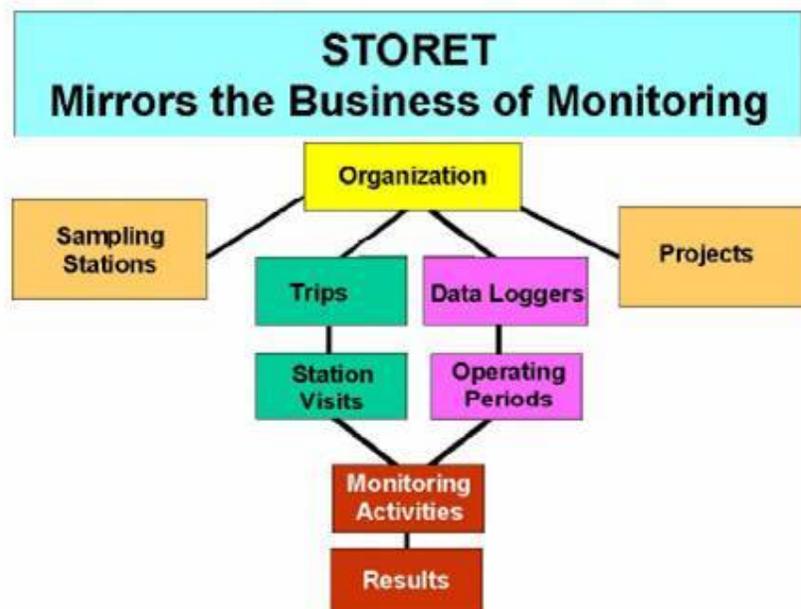
Organization: Organizations participate in Trips to visit Stations/Sites to conduct Monitoring Activities (in support of Projects), which generate Results.

Getting Started with SIM

Let's start the SIM import module by going to the "Start" menu on your computer, selecting "Programs," and then choosing "SIM202." You will see the following SIM202 Welcome screen:

On the lower left corner of the welcome screen you can go directly to different tasks such as "Import a File," "View Pending Migrations," and "Manage Configurations." In addition, the top menu bar contains additional selections. For any file you plan to import, you will need to perform the following three steps:

1. Create a data file for projects, stations, or results
2. Create a corresponding configuration
3. Import/migrate your data into STORET.

Here, you will complete the following three steps for a project:

Step 1: Create a Data File

Now you will create a project file that you want to migrate into STORET. The data file you will migrate will contain project data with the following specifications:

- **Fields**—The project data file will contain the following fields in this order: ProjectID, Name, Contact, Purpose, StartDate, Duration

- **Delimiter**—The file will be pipe (|) delimited. Note that a comma-delimited format would not work for this data because the Contact data element contains a comma (see table below).

- **File name**—The data file will be named "Tutorial1_1 - Projects.txt."

- **Data records**—The data will contain three records as defined in the following table:

ProjectID	Name	Contact	Purpose	StartDate	Duration
TUT1-01	SIM Data Creation Exercise Proj1	John Smith, State Contact	Learning	05/20/2003	Ongoing
TUT1-02	SIM Data Creation Exercise Proj2		Learning	05/20/2003	Ongoing
TUT1-03	SIM Data Creation Exercise Proj3	John Smith, State Contact	Learning	05/15/2000	Complete

Using Notepad, create a new file named "Tutorial1_1 - Projects.txt."

Enter the following text (the pipe symbol "|" is made using the shift key with the back-slash "\" key; note the double pipes on the second line where a data element is blank):

ProjectID|Name|Contact|Purpose|StartDate|Duration|Comments
TUT1-01|SIM Data Creation Exercise Proj1|John Smith, State Contact|Learning|05/20/2003|Ongoing|Not needed information|
TUT1-02|SIM Data Creation Exercise Proj2||Learning|05/20/2003|Ongoing|Not needed information|
TUT1-03|SIM Data Creation Exercise Proj3|John Smith, State Contact|Learning|05/15/2000|Ongoing|Not needed information|

Save the file in the same location as your other sample files. Be sure to keep files that you create or modify in this folder. It is important to remember where you saved the file.

Close the file.

### Step 2: Create an Import Configuration File in SIM

In SIM, data files are separate from configurations, a configuration can be re-used by several data file imports once the data elements are the same and in the same order.

Now you will create an Import Configuration file in SIM to match your data. The Import Configuration file defines the structure that you intend your data to have and provides the basis for how SIM will validate the data in the file that you are importing.

Open SIM.
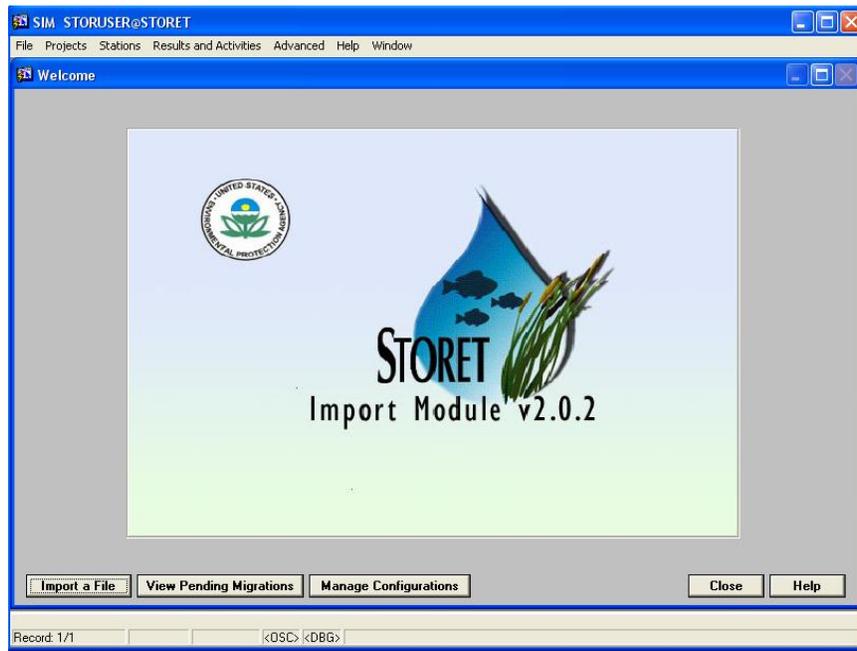
From the Welcome window click the **Manage Configuration** button.
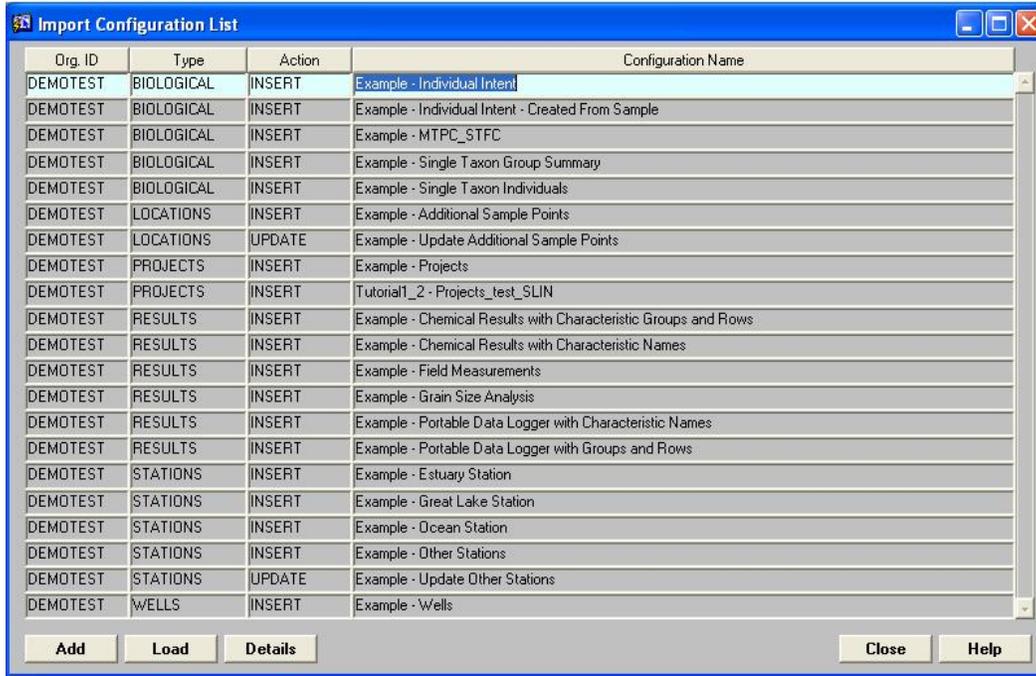
SIM includes a list of configurations for your use, but we will add a new configuration for practice. In the Import Configuration window, click the **Add** button from the lower left corner, click the **Insert** button in the Import Action screen, click the **Projects** button in the Import Type screen to bring up the Import Configuration screen below.
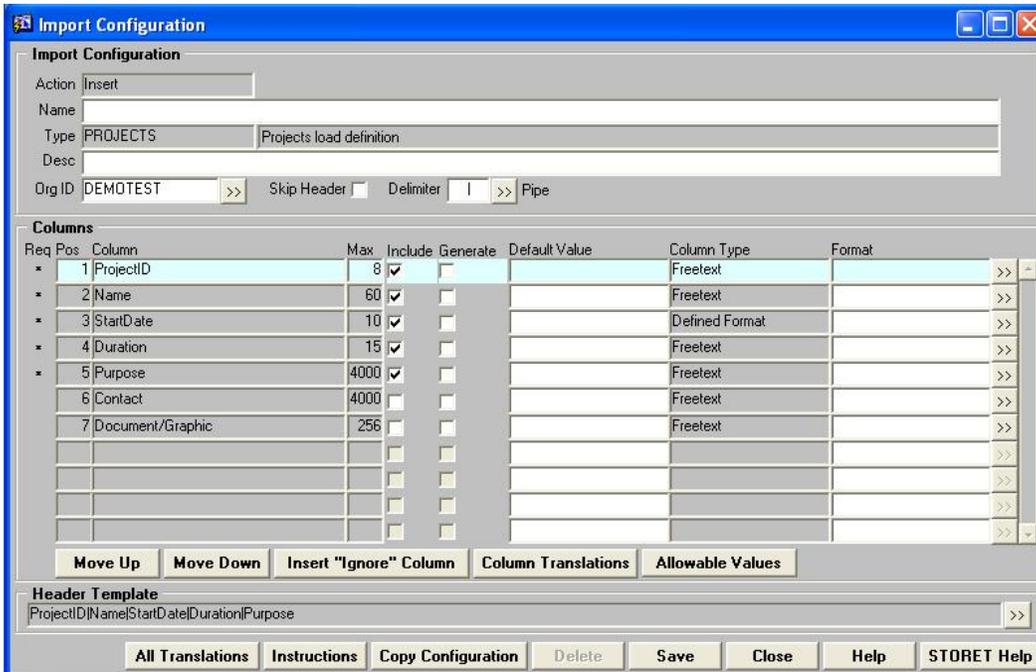
This window contains the following elements:

- **Action**—This explains whether you want to Insert or Add new data or Update data that has already been previously added.

- **Name**—This is where you enter the name of the Import Configuration.

- **Type**—This explains whether you are creating a configuration for projects, stations, results, or other types of data.

- **Desc**—In description, you can enter any information you would like to identify the configuration.

- **Org ID**—All STORET data are connected with an organization. SIM may complete this field for you, but you can get a list of the existing organizations by clicking the downward arrow button on the right of the text field.

- **Skip Header**—Header rows are not accepted in STORET. This box allows you to skip the header row if you have one in your file.

- **Delimiter**—In this box, you can select the type of delimiter you used in your file.

Other parts of the Import Configuration window will be explained later.

Enter "Tutorial1_1 - Projects" in the Name field.

Enter "File for demonstrating project data migration" in the Desc field.

Click **Save** to be able to use other buttons to set up the configuration.

In the data file, the first rows are simply headers and you don't want to import this row as a data record. Check the Skip Header box to indicate the first row to be skipped.

Click the **Include** box next to "Contact" so that the column will be imported from the data file. Note that the check boxes for the field required by STORET cannot be deselected.

For SIM to know the order of your data elements in your data file, you need to designate the "Pos" in the configuration to reflect the position of each column in the data file. You may change the column order to suit your needs by renumbering them or by using the **Move Up** and **Move Down** buttons. For example, in the data file, the "Contact" and "Purpose" columns are right after the "Name" column. To show this is the configuration file, click on the "Contact" and click the **Move Up** button several times to move the Contact below the Name. Repeat this action to move Purpose to the appropriate position. You can also use **Move Down** button to adjust the position to the lower order.

In your data file, the 7th column is "Comments." Because you don't want to include comments in this exercise, click on **Insert "Ignore" Column** and move it up to Position 7.

Click the **Save** button.

In your data set, we have left the Contact field blank in the second record, so here will specify a default value here. Enter the text "John Doe" in the Default Value cell for the Contact field.

Click the "**>>**" button to the right of the Format cell to select the format you want to use from a list of the possible formats. Do this for the StartDate.

Select the "MM/DD/YYYY" format and click the **OK** button to enter that value into the **Format** column.

Click **Save** to save the configuration you have created. The final screen should look like the following figure:
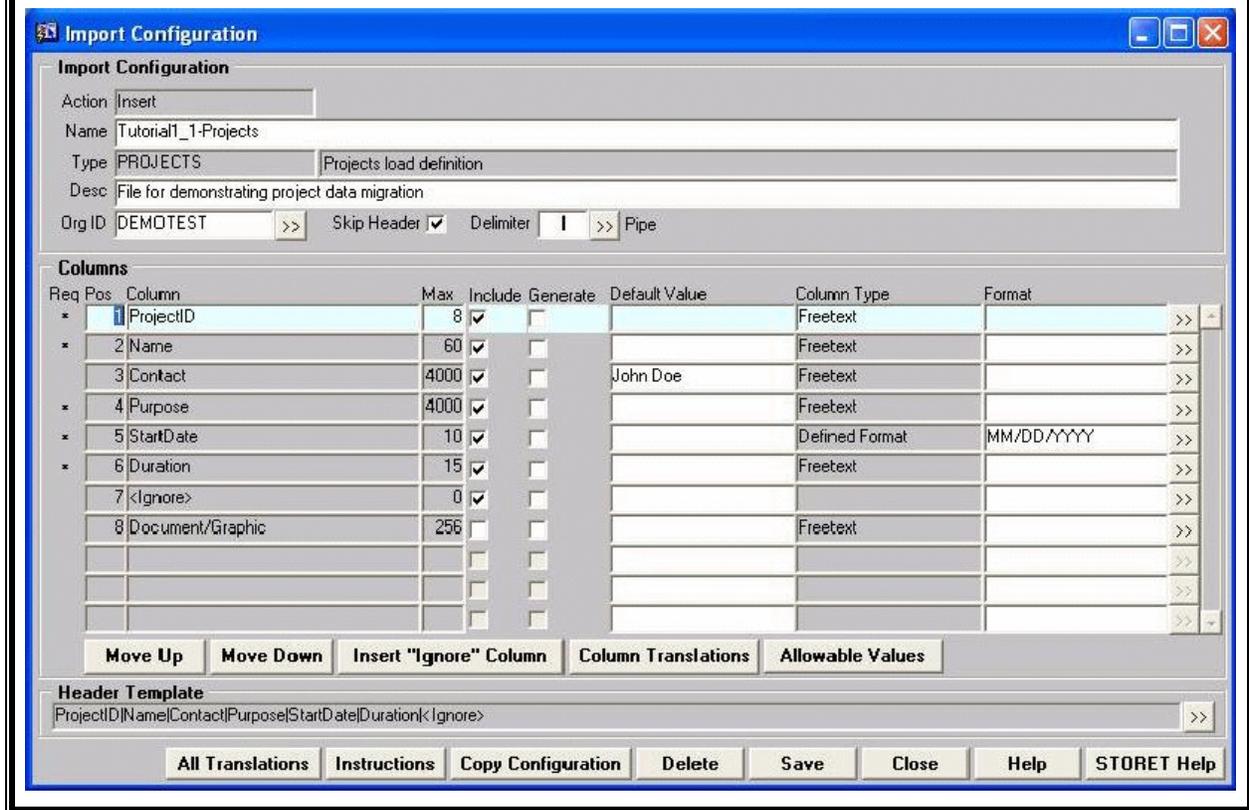
You have just created the configuration settings for your import file. Review the data fields of the Import Configuration window:

- **Name**—This is the name of the configuration file you are creating.

- **Type**—This specifies the type of data you will be importing.

- **Desc**—This is a field where you can enter a description for the configuration file.

- **Org ID**—This allows you to specify an organization that exists in STORET. Clicking the ">>" button to the right of the text field gives you the list of organizations in STORET.

- **Skip Header**—This allows you to skip the first line of the data file.

- **Delimiter**—This is the character that will separate the data elements in your data file.

Under the **Columns** part of the main section of the window, there are several columns of information that include the following:

- **Req**—This column indicates whether a field is required for importing data through SIM.

- **Pos**—This defines the column position or the order in which your columns are to appear in the data file.

- **Column**—This column lists each data element in the data file you plan to import.

- **Max**—This column specifies the maximum length of the field as dictated by STORET. For example, the ProjectID field can have a maximum of eight characters.

- **Include**—This column indicates the columns you have elected to include in your data file.

- **Generate—**This column allows you to generate a value for a field that is not in your data file.

- **Default Value**—This allows you to specify a value that will be placed in the given field by default if the field is empty in your data set.

- **Column Type**—This field describes the type of data that can be entered in the data field as dictated by STORET. Note the "Defined Format" value for this entry is in the StartDate row. This means there are predefined data formats for the date that you can choose from in the **Format** column.

- **Format**—This field contains the format of the data in your import file. For example, a date could be in one of several different formats (e.g., MM/DD/YYYY, MM-DD-YYYY, MM/DD/YY).

Using Help with Configuration Files

The Help function is window-specific. For example, if you click on **Help** from the Import Action window, it will only shows the items of the Import Action window (Insert, Update, and Remove).

During the process of setting up the configuration, you can click **Help** from the lower buttons in the Import Configuration window for information about specific data elements.
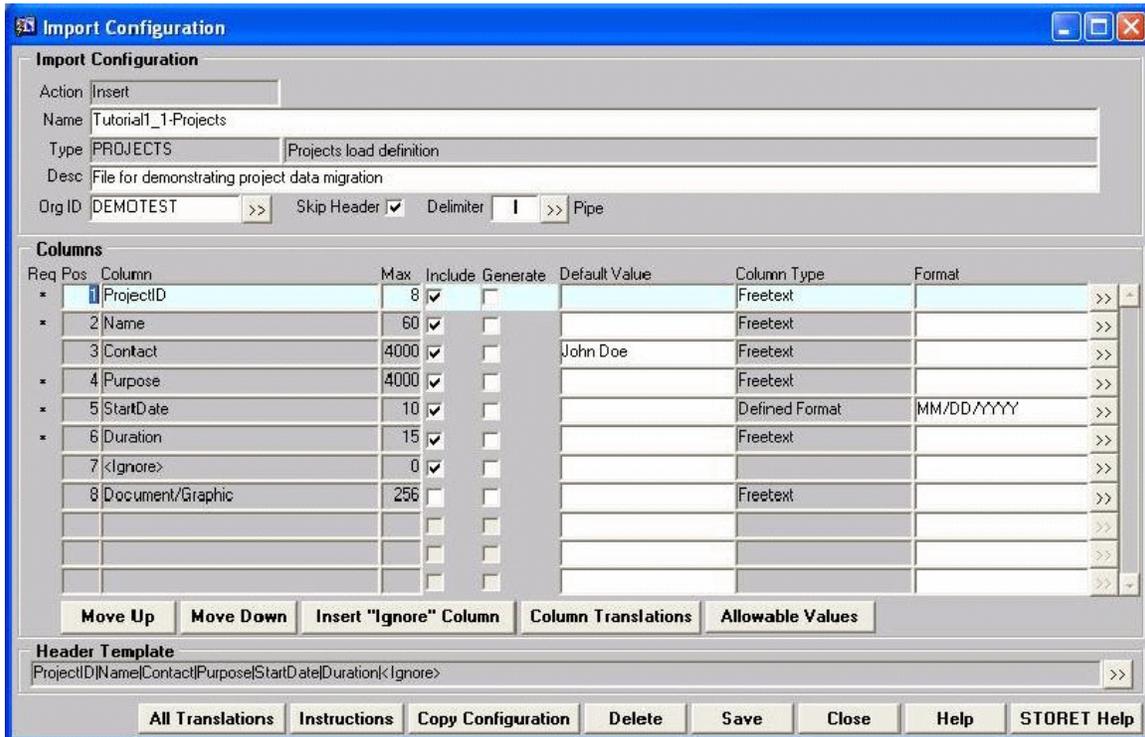
For example, if you find Generate confusing in the Import Configuration window, you can find it in the **Item Help** sections in the User Help windows. As you scroll down to Generate, the text in the light blue section will show you details for Generate.
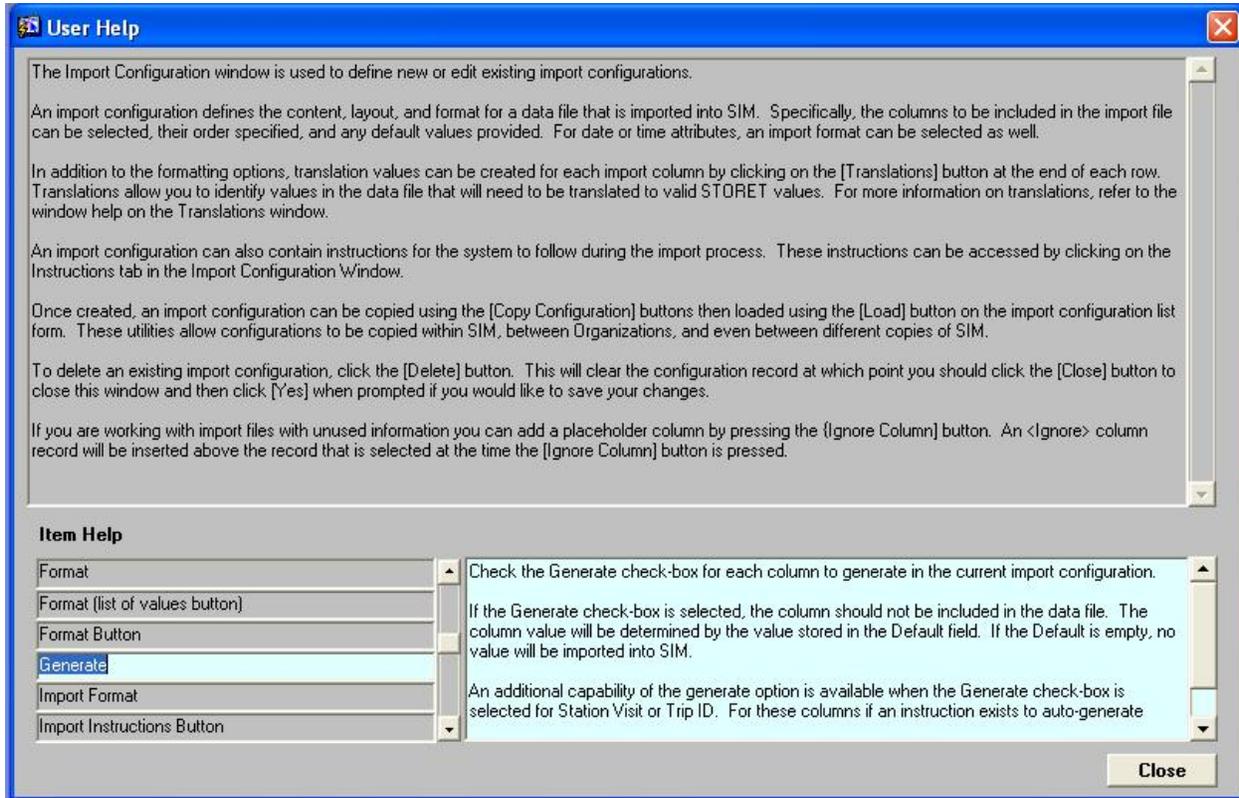
As you read the text in light blue box, you will see that it is not necessary to use the Generate function for the project data file we created because your data has all the data fields required for project data structure to complete the data configuration.

Click **Close** until you return to the Welcome window.

Step 3: Import Your Data File into SIM

From the Welcome window, click the **Import a File** button, click the **Insert** button in the Import Action screen, and click the **Projects** button in the Import Type screen to bring up the Import a File window.

Make sure that the correct Import Configuration is selected, or click the ">>" button to navigate to the correct one.

In the Import a File window, change the name of the data file you want to import to the file you created earlier in this tutorial. Click the ">>" button, navigate to the location of "Tutorial1_1 - Projects.txt," and select it.
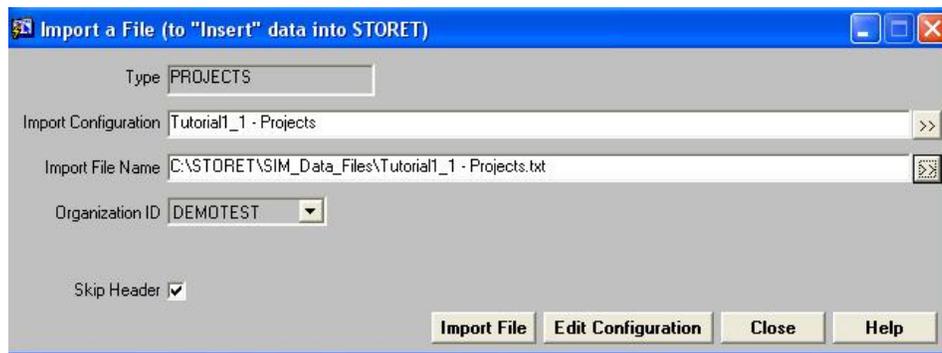
Click the **Import File** button to import the data into SIM.

SIM warns that the first line will be skipped because we checked the **Skip Header** box. It is what we expected, so click the **Continue** button.

When the Import Complete message pops up, click **OK** to get to the Import Status window.
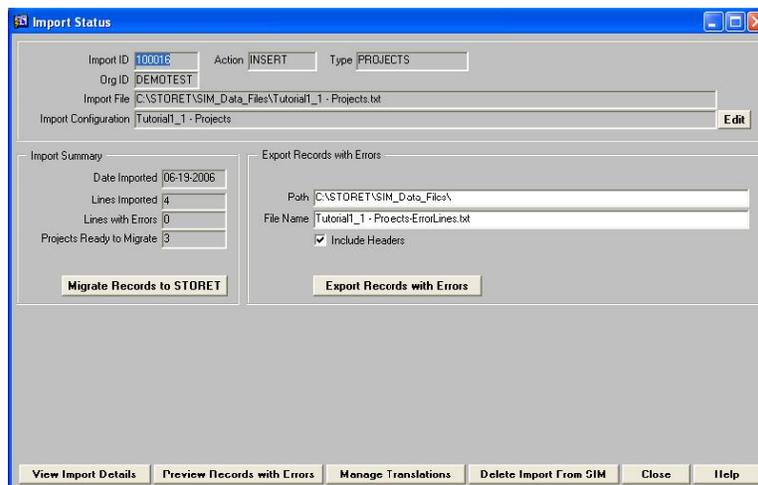
View Imported Data in SIM

If you would like to view your imported data in SIM before it is migrated to STORET, you can use the "Projects" menu bar. You can also view stations and activities/results data before migration by using the menu bar.

Using the menu bar is convenient because you can view the projects without closing the Import Status window. After you finish viewing the projects from the menu bar, you can close the Projects window and keep working on migration in the Import Status window.

Once the data are imported without errors into SIM, migrate the data to STORET by clicking the **Migrate Records to STORET** button.

Click **OK** on the Migration Complete message.
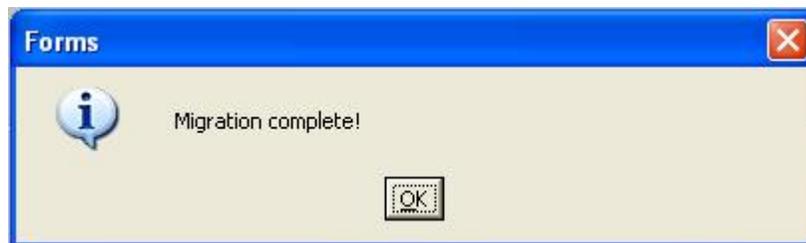
Close all windows, and **Exit** SIM.

The Import Status window should show that four records were read and three records are ready for migration (the four records included the header row that we do not want to migrate). If the window indicates that there are errors, click the **Delete Import from SIM** button and review your file to make sure that it matches with the file "Tutorial1_1 - Projects Correct.txt." If it does, there may be an error in your configuration file. Try importing the "Tutorial1_1 - Projects Correct.txt" file. If you still get errors, click the **Edit Import Configuration** button and make sure the configuration definitions are correct.

Note that once a file has been imported into SIM, but has not yet been migrated, it is added to the Pending Migration List. Once an Import is migrated, it is added to Migration List.

Exporting Data from a Spreadsheet to Create the Data File

Your original data that you want to migrate to STORET may be stored in a spreadsheet or database. These applications generally provide export tools that make it easy for you to generate a text file that can be read by SIM. Now you will create a data file in the same format as the file you created by hand earlier in this tutorial, except that the file will be tab-delimited. You will need to change your configuration file to accommodate this. This example uses data in an MS Excel spreadsheet.

Open MS Excel, and open the spreadsheet "Tutorial1_2 - Projects.xls."

From the menu bar, select "File, Save As."

In the **Save as Type** combo box, select "Text (Tab delimited) (*.txt)," change the File Name to "Tutorial1_2 - Projects.txt," and **Save** the file to the User Created directory.

MS Excel will give you a warning that some data formatting may be lost. Click **Yes** to this message.

Close MS Excel.

Copying and Loading a Configuration

You will create a new import configuration file that recognizes the tab delimiter by making a copy of the configuration you created earlier and editing it to create the new configuration.

Reopen SIM.

From the Welcome window click the **Manage Configurations** button and double click on "Tutorial1_1 - Projects" from Import Configuration List window to select the configuration and view the details.
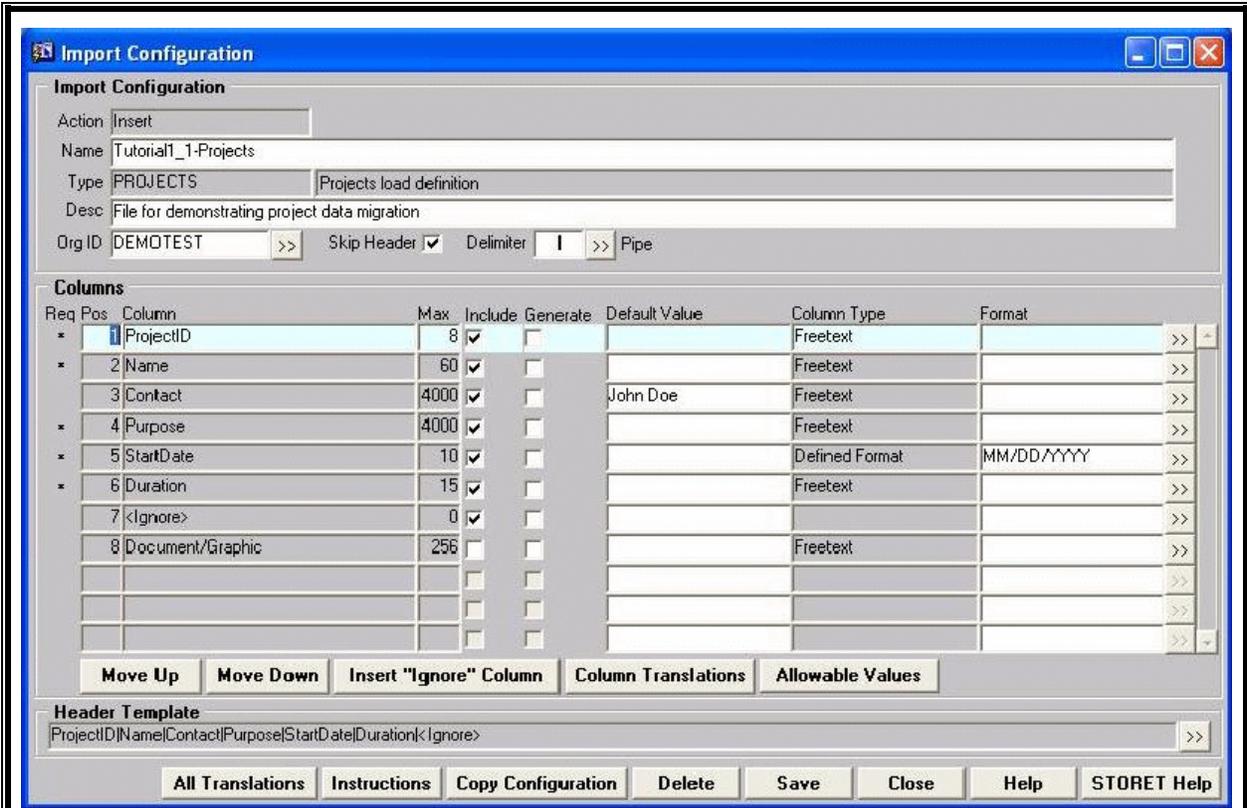
Click on the **Copy Configuration** button**,** and specify the "Tutorial1_2-Projects Config.cfg" in Path and File Name in the Copy Import Configuration window to export your configuration.
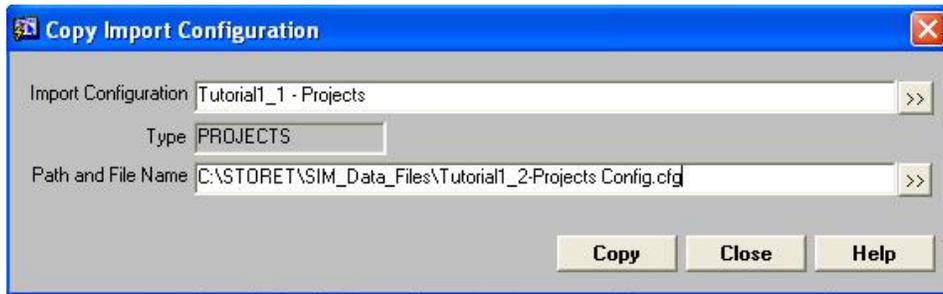
Click the **Copy** button.
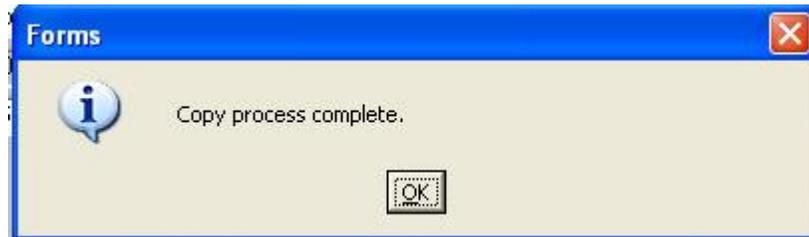
The Copy Process Complete message will appear. Click the **OK** button and close the windows until you return to the Import Configuration List window.

In the Import Configuration List window. Click the **Load** button, and then click the ">>" button to the right of the Path and File Name field.

Navigate to the "Tutorial1_2 - Projects Config.cfg" file you just created, select it, and click the **Open** button. In the Configuration Name field, edit the entry so it reads "Tutorial1_2 - Projects Config," and then click **Load**.
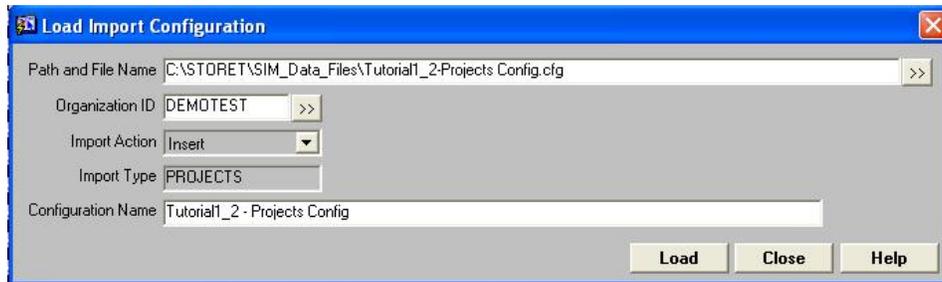
Click **OK** on the Load Process Complete message. The format should be identical to the format you created earlier.

To change the delimiter to a tab, click the ">>" button to the right of the Delimiter field, select "Tab" from the list, and click **OK.** The word "Tab" should appear in the field to the right of the downward arrow button.

Click **Save** to save the file and then close everything to return to the Welcome window.

From the Welcome window, click the **Import a File** button, click the **Insert** button in the Import Action screen, click the **Projects** button in the Import Type screen to bring up the Import a File screen.

Click the ">>" button to select "Tutorial1_2 - Projects Config" as Import Configuration, click the ">>" button to select the "Tutorial1_2 - Projects.txt" file you created from MS Excel, then click **Import File**.

Click **Continue.**

Click **OK** at the Import Complete message to go to the Import Status window.

All three records should have been read with no errors. If you have errors, double check your import configuration and the input file "Tutorial1_2 - Projects.txt" to confirm that they are correctly formatted. After correcting any errors, click the **Delete Import from SIM** button and repeat the steps listed above to import the new file.

You can now migrate the records to STORET.

Click on the **Migrate Records** button to STORET.

Click **OK** in the Migration Complete message.

Press the **Close** button to exit SIM.

Tutorial 2. Dealing with Import Errors

If you spend much time importing data into SIM and migrating it to STORET, you will eventually encounter some errors, so it is important for you to understand how to interpret the error messages and take corrective actions. This tutorial is designed to expose you to some of the errors you may encounter and teach you how to address those errors. During this tutorial you will

- Import data into SIM with errors
- Review and interpret SIM import error messages
- Export records with errors
- Correct errors in your data and re-import them into SIM
- Set up translations.

In this tutorial, you will be working with Station data. To understand some of the errors that you may encounter with these data, you need to first understand the structure of the data in STORET. You can gain some clues about the STORET data structure by looking at the data definitions in the SIM configuration file.

Example SIM Configuration File for Great Lake Station Data

You will look at the "Example - Great Lake Station" import configuration file.

Reopen SIM.

From the Welcome window, click the **Manage Configurations** button to bring up the Import Configuration List window.
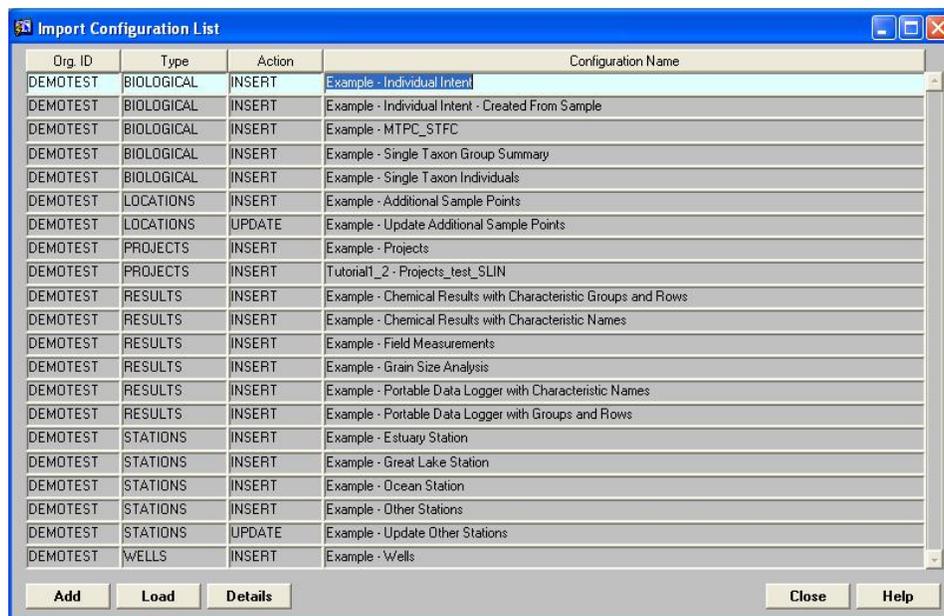
Choose the configuration "Example - Great Lake Station" to bring up the Import Configuration window.
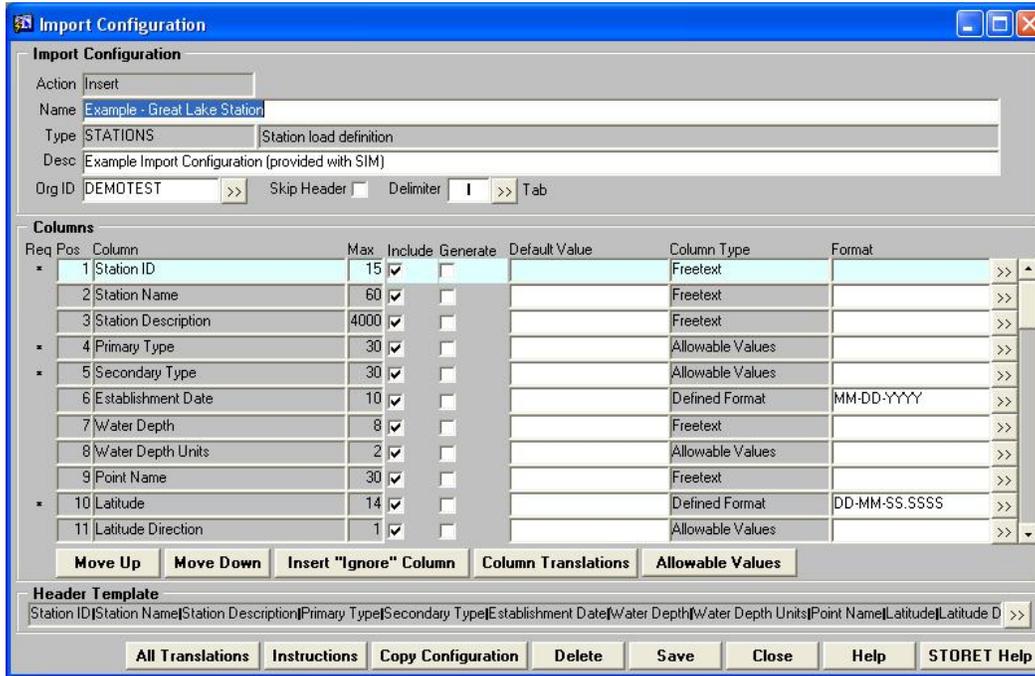
Please note that the configuration file specifies a tab-delimited data file. Also note the list of data elements and the order in which they are expected to appear in the data file.

Review the list of data elements included in the configuration file (i.e., the elements whose "Include" column box is checked). Note the maximum length for each field (Max), the type of data allowed in the field ("Column Type"), and the format defined for the data ("Format").

Click the **Close** button to return to the Import Configuration List window.

Click **Close** to return to the Welcome window.

Correcting Import Errors

Now you will learn how to correct import errors in SIM by importing two data files, one that is correct and one that contains errors. Next, you will review the error log in SIM so that you can become familiar with SIM's error messages. After correcting all of these problems, you will re-import a corrected data file and migrate the two datasets into STORET.

From the Welcome window, click the **Import a File** button.

Click the **Insert** button in the Import Action screen.

Click the **Stations** button in the Import Type screen to bring up the Import a File window.

In the Import File window, click on the ">>" button to the right of the Import Configuration text box.

Select the "Example – Great Lake Station" configuration from the list and click **OK.**

Click on the ">>" button beside the Import File Name text box and locate the directory that contains all the sample files for this tutorial.

Select "Tutorial2 - Great Lake.txt" and click **Open**.

In the Import File window, click on **Import File**.

Click **OK** in the Import Complete message box.

The Import Status window should show an import of one record into SIM. Do **not** migrate the record to STORET.

Close the Import Status window and return to the Welcome window.

Repeat the steps in the last box. This time, keep the Import Configuration the same, but browse to the data file "Tutorial2 - Great Lake with Errors.txt."

Click the **Import File** button**.**

Click **OK** in the Import Complete message box.

Click **View Import Details** to view the error messages in the Event Log Detail window.

ID	Type	Message	Context
100340	Message	Start: 19:16:37	
100341	Error	Station TUT2-GLK-1 already exists in SIM.	Line 1
100342	Warning	Station ID is longer than 15 characters and will be truncated.	Line 2
100343	Error	User-Defined Exception Line 2: Latitude does not match chosen format.	Line 2
100344	Error	Primary Type, GREAT LAKE is not a valid value.	Line 4
100345	Error	State is required.	Line 4
100346	Error	State is required.	Line 4
100347	Error	Latitude Direction, E is not an allowable value.	Line 6
100348	Error	E is not an allowable Latitude Direction.	Line 6
100349	Error	Great Lake Dist to Shore Units, miles is not an allowable value.	Line 7
100350	Error	Great Lake Dist to Shore Units, miles is not a valid value.	Line 7
100351	Error	ORA-06502: PL/SQL: numeric or value error: character string buffer too small - in check storet tran	Line 8
100352	Error	Great Lake, 04020103 is not an allowable value.	Line 8
100353	Error	Great Lake, 04020103 is not a valid value.	Line 8
100354	Error	Great Lake is required when Station type is Great Lake.	Line 8
100355	Error	Station TUT2-GLK-1 already exists in SIM.	Line 10
100356	Message	10 lines processed / 3 stations loaded	
100357	Message	Finished: 19:16:37	

Suppress Warnings   Close   Help

To view the list of data files currently imported into SIM, select **Advanced** from the SIM menu bar and choose **Pending Migration List**.

The Pending Migrations window will show the two station imports. It may also show other imports that you have previously made. To fix the first error, you will need to delete the first import from SIM.

Click on the record to import of "Tutorial2 - Great Lake.txt."

Click on the **Status** button at the lower part of the Pending Migrations window.

In the Import Status window, click the **Delete Import from SIM** button.

Click **Yes** to confirm the deletion, which should return you to the Pending Migrations window.

Click **OK**.

Close the Pending Migrations window.

You will review the messages associated with each of the records and correct these problems. In general, each data record takes one line in your data file. Consequently, the line number in the "Context" column will correspond to the row number of the record with the problem.

- The first error message for Line 1: Station TUT2-GLK-1 already exists in SIM. This indicates that a station record with the same Station ID has already been imported to SIM. In this case, the record in the file "Tutorial2 - Great Lake.txt," which you imported

into SIM first, has the same Station ID as the first record in "Tutorial2 - Great Lake with Errors.txt."

Open the "Tutorial2 - Great Lake with Errors.txt" and "Tutorial2 - Great Lake.txt" files and compare them to confirm these duplicate Station IDs. Close both data files.

▪ The second message is "Warning" for Line 2: Station ID is longer than 15 characters and will be truncated, which is a warning message for the second record relating to the length of the Station ID field. The Station ID should not be longer than 15 characters.

Open the "Tutorial2 - Great Lake with Errors.txt" data file, and you will see that the Station ID for the second record is "TUT2-GLK-1-TOO_LONG," which is longer than 15 characters. SIM will still import this record, but the ID will be truncated to the maximum 15 characters.

To fix the problem, change the ID to "TUT-GLK-1B" in the data file.

Save the data file as "Tutorial2 - Great Lake Corrected.txt" in your "User Created" folder.

▪ The third message is "Error" for Line 2: User-Defined Exception Line 2: Latitude does not match chosen format. Note that you are able to define formats in your Import Configuration file. It is important that your data file format matches the configuration file format.

If you scan the second record in "Tutorial2 - Great Lake Corrected.txt," you will note that the latitude value is "47.4," which is in decimal degrees format rather than degrees-minutes-seconds, the format specified for this data field in the configuration file.

Change the data value to "47-40-20."

Save the file again.

▪ The fourth message is "Error" for Line 4: "Primary Type, GREAT LAKE is not a valid value. At first glance this message may not make sense. Other records in this dataset have a Primary Type value of "Great Lake" and do not generate an error message. The problem here is that the Primary Type value must match the allowed values in STORET, and the case of the text must match as well. In other words, STORET is case sensitive, so "Great Lake" is valid, but "GREAT LAKE" and "Great lake" are not.

Using Translations

Instead of further correcting the "Tutorial2 - Great Lake Corrected" file, you will correct this error using a translation. A translation transforms values in your dataset to new values that are valid in STORET. Instead of manually changing the data in an import file to match STORET's requirements, translations allow you to preserve your data in its original format and correct the conversion in SIM. In this case, the translation will convert the value "GREAT LAKE" in your dataset to "Great Lake," which matches the case for the Primary Type data element in STORET.

Close the Event Log Detail window to return to the Import Status window. **Close** the Import Status window.

Click on the **Manage Configurations** button and open the "Example Great Lake Station" Import Configuration.

In the Import Configuration window under the "Column" section, go to the row for the "Primary Type" column (row 4) and click the **Column Translation** button in the lower part of the window.

Select "Primary Type" as "Column Name" from the dropdown box.

In the first row under the "Translate From," type "GREAT LAKE" in all capital letters.

Now click the ">>" button next to the "Translate To**"** column in the first row. This opens a Translation Allowable Values window, which contains a list of acceptable values in STORET.

Note that for longer lists, such as characteristics, you can either click the value you want from the list or you can use the Find text box at the top to narrow the list of values. The "%" symbol is a wild card for the search.

To narrow the list to all entries starting with the letter "G," click in the Find text box in front of the "%" symbol, and type "G."

Click the **Find** button. The list of values should only include those that begin with the letter "G."

Select "Great Lake" from the list and click **OK**.

Note that the search is not case sensitive, so typing a lower case g would produce the same result. To search for a list containing a certain substring anywhere in the string, type the substring after the "%" symbol. For example, to get a list of all entries with the string "Lake" in them, type "%Lake" in the Find text box and click the **Find** button.

Click **Save** and then **Close** to return to the Import Configuration window.

Note that to delete a Translation, you navigate to the Translations screen, click on the line of the translation you want to change, and click the **Delete** button.

These are all the errors you will correct now.

Close the Import Configuration window to return to the Import Status window, by selecting **"**Advanced**"** from the SIM menu bar and choosing "Pending Migration List." Then click **Status**.

Delete the "Tutorial2 - Great Lakes with Errors.txt" import from SIM by clicking **Delete Import from SIM** and clicking the **Yes** button to confirm the deletion.

Return to the Import File window and make sure "Example - Great Lake Station" is still selected as the Import Configuration. Click **Browse** to select the "Tutorial2 - Great Lake Corrected.txt" data file.

Click the **Import File** button, and then click **OK** to get to the Import Status window. There should now be four rows with errors.

Exporting Records with Errors

At this point, you will export the records with errors, correct those errors, and import the corrected records. These steps are extremely helpful when you have a large number of records that need to be changed.

Notice the **Export Records with Errors** box on the right of the Import Status window. Ensure that the Export Path text box maps to the directory in which you placed the SIM Data Files (the ...SIM Data Files\User Created\directory). Type in the correct file path if necessary.

Name the Export File "Tutorial2 - Export File.txt."

Click the **Export Records with Errors** button to export the records with errors and click **OK** in the Export Complete message box.

Note that the Import Status window data has now changed so that it shows only the status of the records that were imported into SIM without errors.

You will now proceed to migrate error-free records to STORET and correct the errors in the remaining records.

Click on **Migrate Records to STORET** to migrate these error-free rows to STORET.

Click **OK** in the Migration Complete message box and return to the SIM Welcome Screen.

Open the "Tutorial2 - Export File.txt" data file in Notepad.

Let's keep reviewing the rest of errors from the earlier screen shot of **View Import Details** in the tutorial against the "Tutorial2 - Export File.txt."

- For Line 6, which refers to the TUT2-GLK-5 record (the first record in the "Tutorial2 - Export File"), the messages are "Latitude Direction, E is not an allowable value" and "E is not an allowable Latitude Direction." This entry should have been "N" instead of "E."

You can manually correct this error in "Tutorial2 - Export File.txt" by changing the "E" after the latitude value of 47-40-20 to "N" in the first row of the file.

- The error messages for Line 7, which refers to the TUT2-GLK-6 record, the second record in "Tutorial2 - Export File.txt," are "Great Lake Dist to Shore Units, miles is not an allowable value" and "Great Lake Dist to Shore Units, miles is not a valid value." In this case STORET does not recognize the units "miles" spelled out in full, but instead expects the abbreviation "mi" for miles.

Correct this error by using a translation.

Select "Manage Configurations."

Select "Example - Great Lake Station" from the Import Configuration List.

Click on **Details.**

Select "Great Lake Dist to Shore Units" under the "Column" section. In the row under the "Translate From," type "miles." In the row under "Translate To," type "mi."

Click **Save**, and then **Close**.

You can review all the translations by clicking on the **Column Translations** button.

Save and **Close** the Import Configuration window.

- ▪ The error messages for Line 8, which refers to the TUT2-GLK-7 record, are "ORA-06502: PL/SQL: numeric or value error: character string buffer too small - in check storet Great Lake, 04020103 is not an allowable value," "Great Lake, 04020103 is not a valid value," and "Great Lake is required when Station type is Great Lake." In this case, the error messages refer to two fields, HUC and Great Lake, because the field values have been switched in the data file.

Correct this manually in "Tutorial2 - Export File.txt."

- ▪ The error message for Line 10, which refers to the TUT2-GLK-1 record, is "Station TUT2-GLK-1 already exists in SIM." As you may recall, a record with the Station ID of TUT2-GLK-1 was already present earlier in the original input data file.

Correct this by changing the Station ID value to TUT2-GLK-10 in "Tutorial2 - Export File.txt." Save the file.

You will now import the corrected records.

Close the Import Configuration window and the Import Configuration List windows in SIM to return to the Welcome window.

Ensure you have deleted the previous import files with errors from SIM by double-checking **Pending Migrations**.

Choose "Import a File**"** and make sure that "Example - Great Lake Station" is selected as the Import Configuration and click the "**>>**" button to select "Tutorial2 - Export File.txt."

Click the **Import File** button.

Four rows should be ready for migration, and there should be no errors.

Leave the Import Status window open and use the "Stations" menu bar (**->View Stations**) you can view the four stations in the SIM. View the Stations and **Close** the Stations window.

Go back to the Import Status window left open. Go ahead and **Migrate Records** to STORET and click **OK**.

Go to **Advanced, Migration List** from the menu bar of SIM.

To delete the stations migrated to STORET in this tutorial, click on the "Tutorial2 - Great Lake Corrected.txt" in the Migration List window and click on **Delete Migration from STORET** in the lower part of the Migration List window.

Click on **Continue** and **OK** in the message boxes.

Delete the migration "Tutorial2 - Export File.txt" from Migration List window the same way as mentioned above.

Click on the **Close** button to exit SIM.

Tutorial 3. Delete, Update, or Remove Data from STORET

When you work with SIM, you will probably find at some point that you need to remove old data or data with errors from SIM. You may also occasionally need to remove or update data that you have migrated to STORET.

In this tutorial, you will learn how to use the following functions:

- Delete an import from SIM
- Remove data from STORET by deleting a whole migration
- Update data in STORET by using an "Update" action
- Remove data from STORET by using a "Remove" action.

These last two functions ("Update" and "Remove" actions) provide great flexibility because your data corrections are not limited to correcting a full migration: you can specify any amount of data (partial records of the migrated data or all records of multiple migrations) to be corrected. These two functions also allow you to correct data for your organization in STORET that are not in your current copy of SIM.

For the purposes of this tutorial, suppose you have a set of stations with station names beginning with "Example" for which you have already migrated data to STORET. Later, the names of these stations are changed to begin with "Great Lake" to match a new data format adopted by your agency. You want to update the data already imported to STORET to reflect these new station names. You can use one of two basic strategies to accomplish this

- Remove the full migration of the stations with the old station names and re-import the data using a whole new file of stations with the new names
- Import a file using the "Update" action to update just the "Station Name" column.

In this tutorial, we will also practice deleting imports from SIM and removing data from STORET using a "Remove" action. You will also have to import data several times to practice these update and removal techniques.

It may be helpful to view the data files first to become familiar with the data types and formats that SIM accepts.

Open the data files "Tutorial3 - Great Lake Old Names.txt" and "Tutorial3 - Great Lake New Names.txt" using Notepad.

When you have familiarized yourself with the two data files, close them. Note that the data in the second column (the "Station Name") begin with "Example" in the "Old Names" file and with "Great Lake" in the "New Names" file.

Deleting an Import from SIM

You deleted an import from SIM in Tutorial 2, but we will review that here as you migrate a data file with the old station names to STORET to set up for this tutorial.

Open SIM and open the Pending Migrations window by selecting "Advanced, Pending Migration List" from the menu.

If there are station imports in SIM from a previous tutorial, delete them from SIM by clicking the **Status** button to go to the Import Status window for each import and then clicking the **Delete Import from SIM** button.

Close the Pending Migration List window to return to the SIM Welcome window, and click **Import a File**.

Click on the **Insert** button in the Import Action window.

Click on the **Stations** button in the Import Type window.

You are now at the Import a File window. Click the ">>" to the right of the Import Configuration field and select "Example - Great Lake Station." Then click the ">>" button and select "Tutorial3 - Great Lake Old Names.txt."

Click **Import File** and **OK**. This should import five rows into SIM.

Migrate these records to STORET by clicking the **Migrate Records to STORET** button in the Import Status window. When the Migration Complete message appears, click **OK**.

Please note that if you now simply try to re-import a file containing the new station names, you will get an error. We will try this now so you can practice deleting an import from SIM if there were no previous tutorial imports to delete above.

Return to the SIM Welcome window, and click on **Import a File.** Click on the **Insert** button in the Import Action window and click on the **Stations** button in the Import Type window to get to the Import a File window.

Select the "Example - Great Lake Station" configuration and click on the ">>" button to select the file "Tutorial3 - Great Lake New Names.txt."

Click on **Import File** and then **OK** to get to the Import Status window.

Click on the **View Import Details** button to review the import error messages. This should confirm that there are errors in all five records due to duplication of Station IDs. Close the Messages window.

Delete the import from SIM by clicking **Delete Import from SIM**, **Yes**, and then **OK.**

You will be returned to the SIM Welcome window.

Removing Data from STORET by Deleting a Whole Migration

As you saw in the previous step, if you choose to update your data by re-importing it, you need to remove the old data from STORET first. In this part of the tutorial, you will update the station names by first removing from STORET the entire migration you just did containing the old station names and then re-importing the file with the new station names.

Select "Advanced, Migrations List" from the menu to bring up the Migrations List window. This window will list any migrations you have done, including the one you just did as part of this tutorial.

Find the data you just migrated; this should be the last entry with today's date and a migration type of "S" (for stations). Select the row for this migration by clicking on it, then click the **Delete Migration from STORET** button at the bottom of the screen.

Click **Continue** to proceed with the removal, and click **OK** to get back to Migration List window.

Now that the old data have been removed from STORET (and SIM), you are ready to import the new, correct data from "Tutorial3 - Great Lake New Names.txt" and migrate it to STORET.

Return to the SIM Welcome window, and click on **Import a File.** Click on the **Insert** button in the Import Action window and click on the **Stations** button in the Import Type window to get to the Import a File window.

Select the "Example - Great Lake Station" configuration and click on the ">>" button to select the file "Tutorial3 - Great Lake New Names.txt."

Click **Import File,** and then **OK** to get to the Import Status window. The Import Status window should show that five records were read and are ready for migration to STORET. If you have errors, double-check the "Pending Migrations List" and "Migration List" to ensure that you removed the old data from STORET and SIM. Retry importing the new file.

Once your import is ready with no errors, click on **Migrate Records to STORET**; in the Migration message box, click on **OK**.

Before we practice other functions of SIM, we need to clean up the data in SIM and STORET from the first part of this tutorial.

Select "Advanced, Migration Lists" from the menu to bring up the Migrations List window.

Delete all migrations that you have done for these SIM tutorials from STORET, and then close the Migration List window.

Select "Advanced, Pending Migrations List" from the menu.

If there are imports in SIM from these tutorials, delete them by clicking the **Status** button to go to the Import Status window for each import and then clicking the **Delete Import from SIM** button.

Close the windows to return to the Welcome window.

Updating Data in STORET Using the Update Function

As previously noted, there are two basic strategies for updating data already migrated to STORET. In the first part of this tutorial, you accomplished this for a change in station names by removing the whole migration with the old station names from STORET and importing a new file with the new station names. In this part of the tutorial, you will update the station names by using the Update function.

So far, you have only used "Insert" as an Import Action; however, the Import Action window has two other Import Actions: "Update" and "Remove."
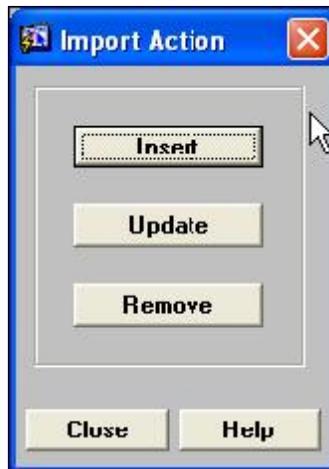
In this version of SIM, you can use the **Import a File** button to add new data to STORET ("Insert"), to update only specified data in STORET ("Update"), or to remove data from STORET ("Remove"). You will need to ensure your import configuration specifies the same Import Action as you specify here (except for "Remove" actions, which do not require a configuration file, see Removing Data from STORET Using the "Remove" Function).

To practice using the "Update" function, you are going to re-import (using the "Insert" function) the station file with the old station names and then update these stations by importing another file with the new names using the "Update" Import Action and an update configuration.

First, re-import the file with the old station names.

Return to the SIM Welcome window, and click on **Import a File.** Click on the **Insert** button in the Import Action window and click on the **Stations** button in the Import Type window to get to the Import a File window.

Select the "Example - Great Lake Station" configuration and click on the ">>" button to select the file "Tutorial3 - Great Lake Old Names.txt."

Click **Import File** and **OK**. This should import five rows into SIM.

Migrate these records to STORET by clicking the **Migrate Records to STORET** button in the Import Status window. When the Migration Complete message appears, click **OK**.

Next, load a configuration with update actions.

From the Welcome window, click on **Manage Configurations.**

In the Import Configuration List, click **Load** to open the Load Import Configuration window.

Browse the path and file name to "...\SIM data files\Provided Configuration\Tutorial3_1-Great Lake Update Config.cfg."
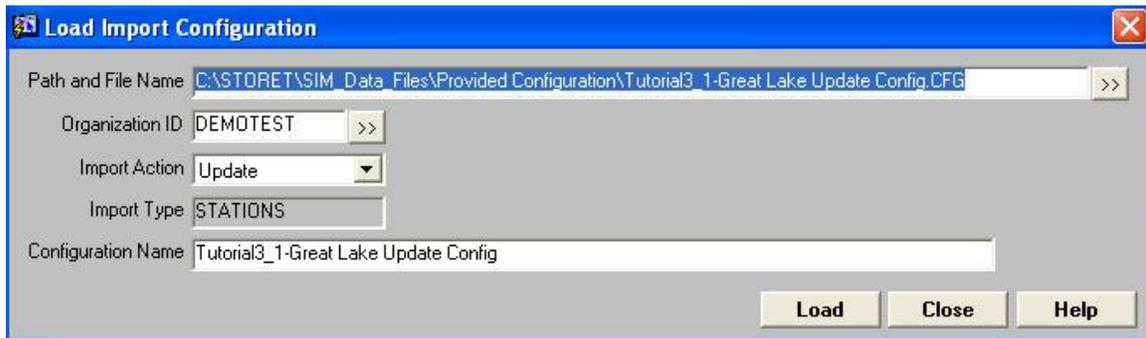
Click **Load**. When the Load Process Complete message appears, click **OK**.

SIM opens the Import Configuration window for you to review. First note that the action is "Update" and also make sure that there are only two columns included, "Station ID" and "Station Name." If more columns are included, click the box under the "Include" column so the additional fields are not included. Close windows to return to the Welcome window.

Next, make sure the update file is ready and then import it.

Open the "Tutorial3 - Great Lake New Names Update.txt" file under the SIM Data File folder. Note that there are only two columns of data, "Station ID" and "Station Name."

Return to the SIM Welcome window, and click on **Import a File.** Select "Update" as the Import Action, and click on the **Stations** button in the Import Type window to get to the Import a File window. Please note that the window title includes the information that this is an import to update data.


Select "Tutorial3_1-Great Lake Update Config" as the configuration and click the ">>" button to select the file "Tutorial3 - Great Lake New Names Update.txt."

Click **Import File**. When the Import Complete message appears, click **OK**.

Review the Import Status window to confirm that five stations are ready to migrate.

Select "Station" and then "View Stations" from the menu to confirm that the new names in the update file have been imported to SIM.


Close the Stations window and return to the Import Status window. Click on **Migrate the Records to STORET**. When the Migration Complete message appears, click **OK**.

Exit SIM.

You have now updated the station names in STORET using the Update function.

Removing Data from STORET Using the "Remove" Function

Previously you have practiced using the "Insert" and "Update" Import Actions to add data to STORET or update data already in STORET. Now you will use the "Remove" Import Action to remove data from STORET.
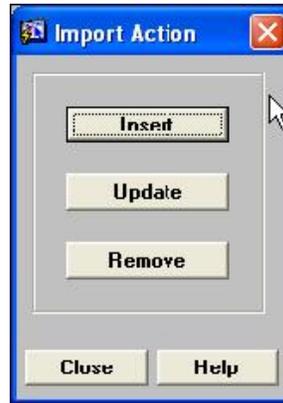
You will use the "Remove" function to remove the stations we added and updated in this tutorial. Importing a file with a "Remove" action is simpler than importing one with an "Import" or "Update" action because you do not need to a configuration, you only need a data file. For stations, SIM will look up the stations to remove from STORET based on the Station IDs in the import file.

First, you will prepare the data file.

Open a new text file using Notepad and type in the following five Station IDs:

    TUT3-GLK-1
    TUT3-GLK-1B
    TUT3-GLK-2
    TUT3-GLK-3
    TUT3-GLK-4

Save as the file as a text file named "Tutorial3 - Great Lake New Names Remove.txt" under the SIM Data File folder.

Now you will use SIM to remove the stations.

Open SIM.

At the Welcome window, click on **Import a File.**

Select "Remove" as the Import Action.

Click on the **Stations** button in the Import Type window.


Note the Import a File window title includes that this import is to remove data from STORET.

For the Import File Name, type in "Tutorial3 - Great Lake New Names Remove.txt" or browse to the "SIM Data Files" folder for the file.

Select "DEMOTEST" as the Organization ID.

Click on **Import File.** When the Remove message box appears, click on **Continue**.

When the Completed Successfully message appears, click **OK** to go back to the Welcome window.

Close SIM to complete this tutorial.

Tutorial 4. Importing Results Data into STORET

This tutorial covers importing results data into STORET using biological data as an example. Its intent is to give you experience in the following:

- Organizing data from a sampling event
- Preparing a data file for import into SIM
- Creating a SIM configuration file for data import
- Importing the data into SIM
- Using the "Instructions" tool to generate station or visit IDs.

To learn these skills, you will work through one example related to taxon abundance.

This tutorial will walk you through the process of migrating one set of hardcopy data from a fictitious sampling trip into SIM and then STORET, including creating the configuration and electronic data files required to accomplish this. It will also cover how to use the "Instructions" tool to generate station or visit IDs.

Although a biological measurement may be something as simple as the weight and length of a fish, storing the information in STORET is complicated by the additional information associated with each measurement (e.g., the method for catching the fish; the location, time, and date of the sampling event). Many individual data elements must be entered for each biological result data record. Breaking these pieces of data into manageable, logically related groups will not only help you think about and structure your data but also make it easier to identify which data are constant across groups of results and which data change from result to result.

Example Taxon Abundance Sampling Event

This example uses a fictitious routine sample collected to describe the fish community in the Blackwater National Wildlife Refuge. The example includes a multi-taxon population census and a single-taxon frequency class count. Information about the sampling event is summarized below.

In general, a sampling event involves a **trip** to a **station** for a particular **project**, during which a sampling **activity** takes place. This activity involves the use of a **sampling procedure** that generates **results**.

In this example, Dr. Lee Manning has recently completed a sampling event in support of a project entitled "Water Quality and Biological Health of the Chesapeake Bay," which has been given the identification number CBCP-001. This ongoing project, which began in 1991, has three specific objectives

- Determine the sources of pollution causing or contributing to existing or anticipated pollution problems in the Bay and its estuaries
- Evaluate the effectiveness of efforts to reduce or eliminate those sources of pollution

▪ Evaluate progress toward achieving and maintaining water quality standards and toward protecting and restoring the fisheries, shell-fisheries, and other living resources of the Bay.

This sampling event occurred at the Blackwater National Wildlife Refuge station on January 2, 2000. During this routine sampling event, Dr. Manning counted and measured fish to assess taxon abundance.

As an experienced STORET user, Dr. Manning knew that STORET requires an identification number for each trip, so he assigned his trip the identifier 01-2000-2. This is his first visit to this particular station.

Dr. Manning used sampling procedures and gear that were already entered into STORET. These include the otter trawling sampling procedure (SP-009), which uses net/horizontal tow gear (NTOT) and the Chesapeake Bay Otter Trawl-1 gear configuration (CBG-014), a highly modified otter trawl custom designed by the Manning Scientific and Correctional Facility Supply. Dr. Manning elected to store his fish samples in amber bottles containing a 70 percent formalin solution (STS-009).

Details of the sampling event are included in the field data sheet on page 4-3, which provides an example of both multi-taxon population census and single-taxon frequency class results (counts for a single species according to size ranges). For the multi-taxon population census, Dr. Manning recorded total counts for each species identified during the sampling event. In this example, there were 54 largemouth bass (*Micropterus salmoides*), 12 yellow bullhead (*Ictalurus natalis*), and 67 brown bullhead (*Ictalurus nebulosus*).

In a single-taxon frequency class, STORET allows you to separate subject taxon individuals into classes based on physical parameters (e.g., length, weight). Counts may be entered for each range of the selected physical parameter. In this example, Dr. Manning used fish fork length to separate individuals in a single taxon (largemouth bass) into three groups (0–20 cm, 21–40 cm, and 41–60 cm).

FISH SAMPLING FIELD DATA SHEET (FRONT)

page __1__ of __1__

STREAM NAME	LOCATION Blackwater National Wildlife Refuge
STATION #CBC-008 RIVERMILE	STREAM CLASS
LAT LONG	RIVER BASIN
STORET #	AGENCY Chesapeake Bay Commission
GEAR otter trawl	INVESTIGATORS
FORM COMPLETED BY Dr. Lee Manning	DATE 1-2-2000 TIME 09 00 (AM) PM REASON FOR SURVEY Routine sample

SAMPLE COLLECTION	How were the fish captured? ☐ back pack ☐ tote barge ☑ other Net/Horiz. tow Block nets used? ☐ YES ☑ NO Sampling Duration Start time 09 00 End time 09 30 Duration 30 min Stream width (in meters) Max_____ Mean_____
HABITAT TYPES	Indicate the percentage of each habitat type present ☐ Riffles____% ☐ Pools____% ☐ Runs____% ☐ Snags____% ☐ Submerged Macrophytes____% ☑ Other (Wetland) 100 %
GENERAL COMMENTS	Sampled at midwater, 23 ft. from surface. Distance fished = 2 nmi.

SPECIES	TOTAL (COUNT)	OPTIONAL: LENGTH (mm)/WEIGHT (g) (25 SPECIMEN MAX SUBSAMPLE)					ANOMALIES							
							D	E	F	L	M	S	T	Z
Largemouth Bass (Micropterus Salmoides)	54						0-20 cm 12 21-40 cm 35 41-60 cm 7							
Yellow Bullhead (Ictalurus natalis)	12													
Brown Bullhead (Ictalurus nebulosus)	67													

Field data sheet for taxon abundance.

Organizing Biological Data for STORET

To enter the data from this example sampling event into STORET, you will need to provide not just the results, but also the associated data relating to the project, station, trip, and sampling procedures used. These categories of information are discussed in the following sections.

Project Information. Detailed information about this project has already been entered into STORET; therefore, all you need to do to associate the sampling event with the project is provide the project ID number. In this case, the project ID number is CBCP-001.

Station Information. The Station number for the Blackwater National Wildlife Refuge station is CBC-008, as indicated in the upper left corner of the field data sheet. STORET already contains the detailed information about this station.

Trip Information. Each trip must have a trip identification number, which in this case has been specified as 01-2000-2. Other data about the trip that will be included in the data file are as follows:

Data Element	Value
Trip ID	01-2000-2
Trip Start Date	1/2/2000
Station Visit Number	1
Station Visit Arrival Date	1/2/2000
Visit Comments	First Visit

Activity Information. Information that you will need to enter into STORET about this sampling event include identifiers for the sampling activities, information about the type of activity, and the time and date the activity began and ended. The activity data elements you will enter into STORET from this sampling event are summarized in the following table.

Data Element	Value
Activity ID	TUT4EX1
Medium	Biological
Activity Type	Sample
Activity Category	Routine Sample
Intent	Taxon Abundance
Community	Fish/Nekton
Activity Start Date	1/2/2000

Sampling Procedure and Sampling Information. The details about the standard sampling procedures used during a sampling event are already contained in STORET. You will need to know the appropriate ID numbers to be able to reference them. In this example, the sample collection procedure, gear ID, gear configuration ID, and sample preservation transport and storage methods have already been entered into STORET. Site-specific details of the sampling event that need to be recorded in STORET are also noted, such as details relating to where in the waterbody the sample was taken and the duration of the sampling event. The following table shows the sampling procedure and sampling data elements that you will enter into STORET from this sampling event.

Data Element	Value
Sample Collection Procedure ID	SP-009
Gear ID	NTOT
Gear Configuration ID	CBG-014
Sample Preservation Transport and Storage ID	STS-009

Results Information. Now that you have worked through all the data that support the results, you are ready to identify the results data. A group ID number and a biological results type are required in STORET if the intent of the sampling event is to measure taxon abundance. In this example, you have two types of biological sampling results, multi-taxon population census and single-taxon frequency class. You can simply assign group IDs of 1 and 2 to these results. For the multi-taxon population census, you have a characteristic (the species name) and a result value (the total number of the species found). The result is an actual count of the number of each species, so the value unit is "count" and the value type is "actual" (because it is not an estimated or calculated value). A summary of the multi-taxon population census data record for largemouth bass is shown in the following results table.

Data Element	Value
Bio Results Type	Multi-Taxon Population Census
Bio Results Group ID	1
Characteristic Name	*micropterus salmoides*
Result Value	54
Result Value Units	Count
Value Type	Actual

For single-taxon frequency class results, you must enter the frequency analysis type (P for physical measures, such as fork length, or B for biological conditions, such as lifestage), the upper and lower bounds for each range, and the associated units. For the single-taxon frequency class data, you have results for largemouth bass for each of three fork length ranges (0–20 cm, 21–40 cm, and 41–60 cm). A summary of the data for the 0–20 cm size range is shown in the following table:

Data Element	Value
Subject Taxon	*micropterus salmoides*
Frequency Analysis Type	P
Common Class Descriptor	Length, Fork (Fish)
Common Class Descriptor Units	Cm
Frequency Class Count	12
Bio Group Count Type	Actual
Lower Class Bound	0
Upper Class Bound	20

Data Summary. The following table provides a complete list of each of the data elements from the sampling event that will be loaded into STORET. Note that there are a handful of additional data elements that have been added that were not included in the information provided in the example. Generating this table from the unformatted information provided by the example completes a major portion of the work required to load the biological sampling data into STORET.

Data Type	Data Element	Column Order	Record 1	Record 2	Record 3	Record 4	Record 5	Record 6
Project	*Project ID*	1	CBCP-001	CBCP-001	CBCP-001	CBCP-001	CBCP-001	CBCP-001
Station	*Station ID*	2	CBC-008	CBC-008	CBC-008	CBC-008	CBC-008	CBC-008
Trip	*Trip ID*	3	01-2000-2	01-2000-2	01-2000-2	01-2000-2	01-2000-2	01-2000-2
	Trip Start Date	4	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000
	Station Visit Number	5	1	1	1	1	1	1
	Station Visit Arrival Date	6	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000
	Visit Comments	7	First Visit	First Visit	First Visit	First Visit	First Visit	First Visit
Activity	*Activity ID*	8	Tut4Ex1	Tut4Ex1	Tut4Ex1	Tut4Ex1	Tut4Ex1	Tut4Ex1
	Medium	9	Biological	Biological	Biological	Biological	Biological	Biological
	Activity Type	10	Sample	Sample	Sample	Sample	Sample	Sample
	Activity Category	11	Routine Sample	Routine Sample	Routine Sample	Routine Sample	Routine Sample	Routine Sample
	Intent	12	Taxon Abundance	Taxon Abundance	Taxon Abundance	Taxon Abundance	Taxon Abundance	Taxon Abundance
	Community	13	Fish/Nekton	Fish/Nekton	Fish/Nekton	Fish/Nekton	Fish/Nekton	Fish/Nekton
	Activity Start Date	14	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000	1/2/2000
Sampling Procedure	*Sample Collection Procedure ID*	15	SP-009	SP-009	SP-009	SP-009	SP-009	SP-009
	Gear ID	16	NTOT	NTOT	NTOT	NTOT	NTOT	NTOT
	Gear Configuration ID	17	CBG-014	CBG-014	CBG-014	CBG-014	CBG-014	CBG-014
	Sample Preservation Transport and Storage ID	18	STS-009	STS-009	STS-009	STS-009	STS-009	STS-009

(continued)

Data Type	Data Element	Column Order	Record 1	Record 2	Record 3	Record 4	Record 5	Record 6
Results	*Bio Results Type*	19	Multi-Taxon Population Census	Multi-Taxon Population Census	Multi-Taxon Population Census	Single-Taxon Frequency Classes	Single-Taxon Frequency Classes	Single-Taxon Frequency Classes
	Bio Results Group ID	20	1	1	1	2	2	2
	Subject Taxon	21				*micropterus salmoides*	*micropterus salmoides*	*micropterus salmoides*
	Frequency Analysis Type	22				P	P	P
	Common Class Descriptor	23				Length, Fork (Fish)	Length, Fork (Fish)	Length, Fork (Fish)
	Common Class Descriptor Units	24				cm	cm	cm
	Characteristic Name	25	*micropterus salmoides*	*ictalurus natalis*	*ictalurus nebulosus*			
	Result Value	26	54	12	67			
	Result Value Units	27	Count	Count	Count			
	Value Type	28	Actual	Actual	Actual			
	Frequency Class Count	29				12	35	7
	Bio Group Count Type	30				Actual	Actual	Actual
	Lower Class Bound	31				0	21	41
	Upper Class Bound	32				20	40	60

Preparing the Data File

Although the tabular form of the data presented above is well structured, it is not in a format that can be imported into SIM. SIM expects each data record to be in a row, with each data element separated by the delimiter specified in the configuration file (in this example, a tab character). This requires transposing the data from the table format (in which each record is a column). You could create the data file from the data summary table by going down the column for each record, starting with the Record 1 column, and typing the data value in each cell followed by a tab character. (Please note that there are other ways to create data files with rows, such as creating database queries or using tools provided from the EPA STORET downloads page). However, a MS Excel data file in the correct format is provided with this tutorial ("Tutorial4_1 - MTPC_STFC.xls"). You will now open and save this file in the needed tab-delimited format.

Open the "Tutorial4_1 - MTPC_STFC.xls" file in MS Excel. The rows of data should match the data table on pages 4-6 and 4-7.

Look at the first row and remember that there is a row of column headers in this file; it will require you to check the "Skip Header" box later when you set up a configuration.

Save this file as a tab-delimited text file named "Tutorial4_1 - MTPC_STFC.txt." Click the **Yes** button when MS Excel warns you that formats will be lost.

View your new file in Notepad. It should look like the following screen shot:

If the lines are wrapping, select "Format" from the menu and click on the **WordWrap** button to uncheck it.

Creating the SIM Configuration File

Now that you have a well-structured data file in the proper format, you will need to create a configuration file that will allow you to import this data into SIM. The import configuration tells SIM what data fields you will import in which order.

Open SIM.

From the Welcome window, click the **Manage Configurations** button. This displays the Import Configuration List window. SIM includes a list of configurations for your use. We will modify a configuration for this example.
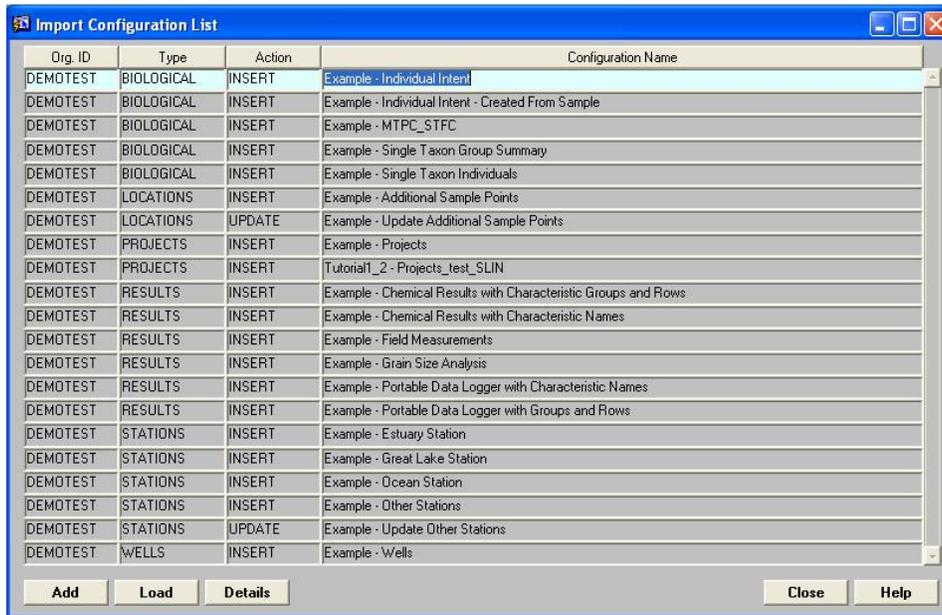
Choose the Example – MTPC_STFC configuration by double clicking on it.

You should now see the Import Configuration window.

Change the Name to "Tutorial4 - MTPC_STFC."

Enter "Modified config for fewer data elements" in the Desc field.

Click the **Save** button at the bottom of the window; this will enable the other buttons needed to finish setting up the configuration.

Note that the **Skip Header** box is checked (to the right of the Org ID in the upper section of the Import Configuration window) to skip the first row of the data file (which you will recall contained headers, not data). SIM will not read this row as data when you check this option.

Click the ">>" button next to the Delimiter field (next to the **Skip Header** box) and ensure that "Tab" is selected from the list.

Now you will select data elements, identify the order in which the data elements are listed in the data file, and select any formats you want to use.

For each item in the "Data Element" column of the example data table on pages 4-6 and 4-7, find the corresponding entry in the "Column" section of the Import Configuration window and ensure that the check box in the "Include" column is checked. You will also need to uncheck the "Include" column for items you do not want to include. Use the scroll bar on the left side of the window to scroll down through the list of data elements. You should be able to find an entry for each Data Element item, and you should have 32 data elements selected when you are done.

For each data element, click the **Move Up** and **Move Down** buttons to adjust the order (position) of the columns until the order matches the column order of the data elements in the example data table on pages 4-6 and 4-7. The number in the "Pos" column changes as you make these adjustments.

For the Trip Start Date and Station Visit Arrival Date elements, click the ">>" button next to the Format field and select the "MM/DD/YYYY" format from the list.

Click on the **Save** button to save the configuration file so far.
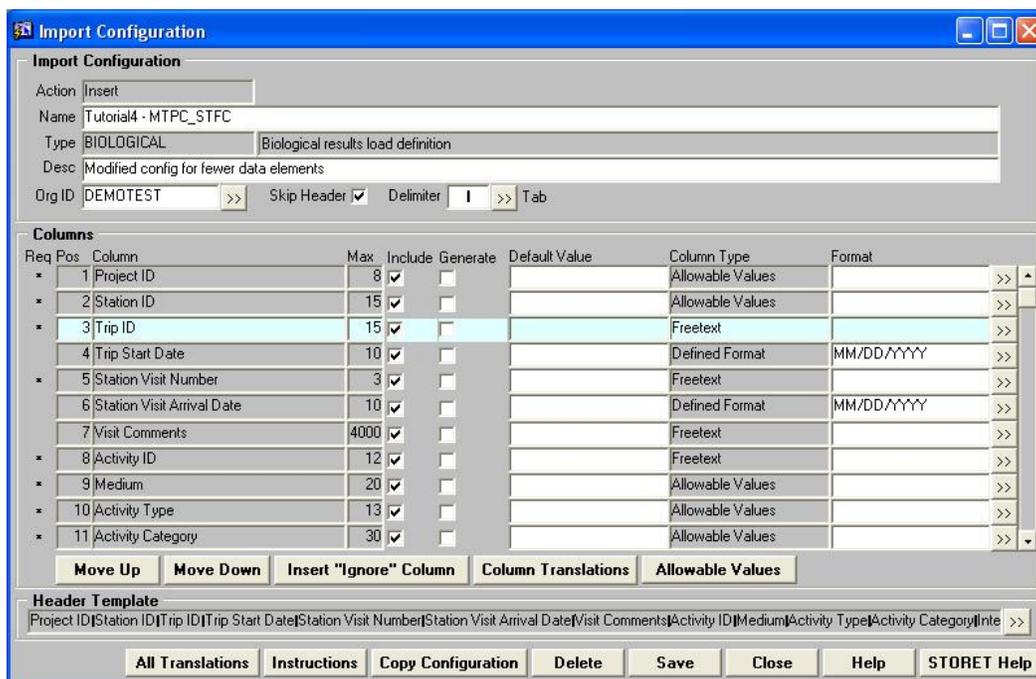
Finally, you will specify any translations you want to use. Using translations in your import configuration can save you a great deal of data processing time. For example, your original data file uses "MTPC" as an abbreviation for multi-taxon population census, but STORET requires that "Multi-Taxon Population Census" be spelled out. Rather than replacing all the occurrences

of MTPC in your data file with the spelled-out phrase, you can tell SIM to translate "MTPC" to "Multi-Taxon Population Census."

In the "Column" section of the Import Configuration Window, select the row for the Medium data element.

Click the **Column Translations** button at the bottom of the "Columns" section. SIM automatically selects Medium as the Column Name, because you had already selected it.

Type Bio in the "Translate From" column.

Type Biological in the "Translate To" column, or click the "**>>**" button and select "Biological."

Click the **Save** button**,** and then **Add** to add another translation.

In the next row under Column Name, select "Activity Type" and translate from "S" to "Sample." Click **Save,** and then **Add**. (Please note that each translation will be saved, even though each one disappears when you click **Save**. This is because the **Column Filter** is set to Medium. Only the translation for the Medium field will appear.)

In the next row under "Column Name," select "Bio Results Type" and translate from "MTPC" to "Multi-Taxon Population Census." Click **Save,** and then the **Add** button.

In the next row under "Column Name," select "Bio Results Type" and translate from "STFC" to "Single Taxon Frequency Classes." Click **Save**. (This is the last translation, so do not click **Add** again.)

To view all the translations, click on the ">>" button next to the **Column Filter** box in the title section of the Translations window and select "*All Columns."
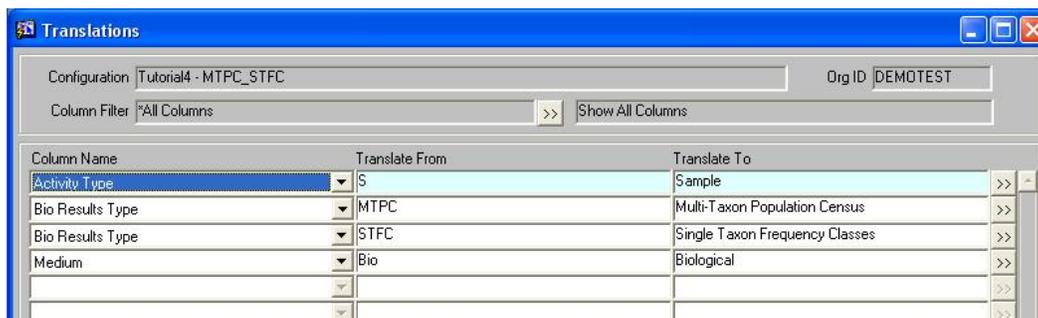
If you have inadvertently entered a duplicate translation or need to delete a translation for any reason, select the one you want to delete and click on the **Delete** button below the translation list.

Close the Translations window. If you have not saved all your changes in the Translations window, a "Save?" message box will appear. Click the **Yes** button to save all the translations and return to the Import Configuration window.

Your configuration file is now complete.

Importing the Data into SIM

You are now ready to load the data into SIM using the data file and configuration file you created. This example uses the configuration you created, "Tutorial4 - MTPC_STFC," which includes data elements for both multi-taxon population census and single-taxon frequency class.

Close all the windows to return to the Welcome window.

From the Welcome window, click **Import a File.**

Click the **Insert** button in the Import Action screen.

Click the **Biological** button in the Import Type screen to bring up the Import a File window.

Select the configuration "Tutorial4 - MTPC_STFC."

Change the name of the data file you want to import to the file you created earlier in this tutorial by clicking the ">>" button, navigating to the location of "Tutorial4 - MTPC_STFC.txt," and selecting it.
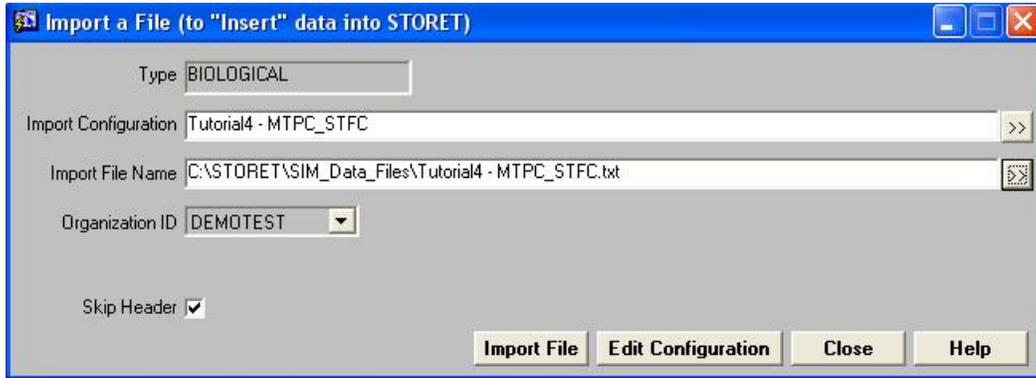
Click the **Import File** button.

When the Verify Skip Header message appears, click the **Continue** button. When the Import Complete message appears, click **OK**.

When the import is complete, the Import Status window will appear. This example should contain no errors.
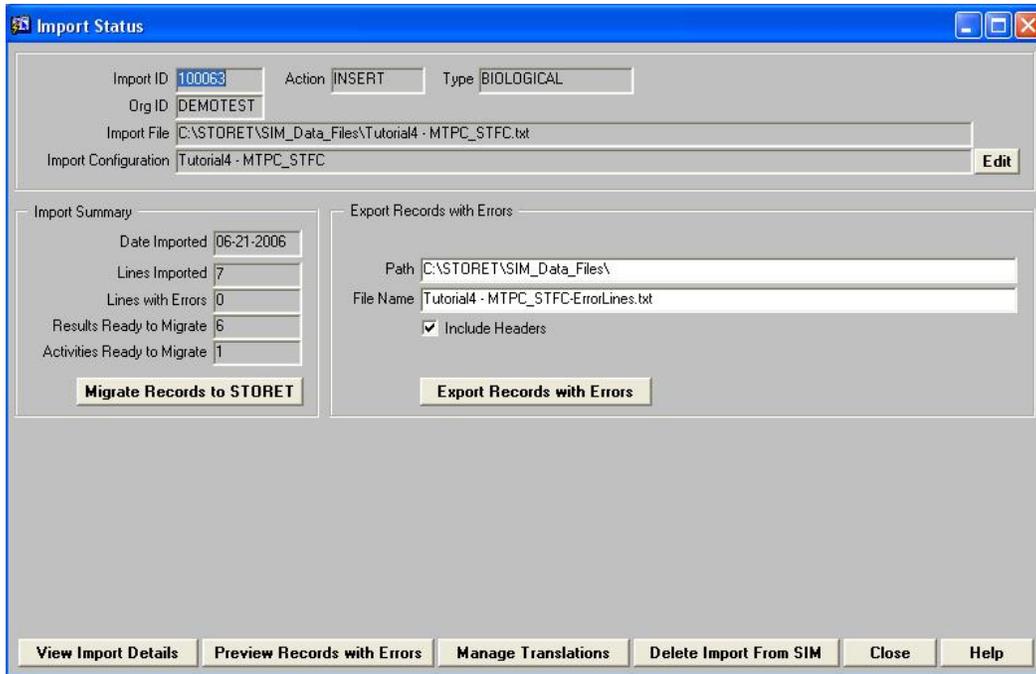
If you get import errors, try the following:

- Click on the **View Import Details** button in the lower left corner of the Import Status window to read SIM's error messages.

- Click on the **Preview Records with Errors** button in the Import Status window to examine the data with errors.

- Double-check your Import Configuration to confirm that you included all necessary data elements and sequenced them in the same order as they appear in the data file.

If you still cannot import your file, the correct configuration file for Tutorial 4 is provided under the "Provided Configuration" folder; however, you should only use this file if you cannot create the configuration yourself.

Even if you do not have import errors, you can check the data you just imported to SIM before you migrate it to STORET.

Leaving the Import Status window open, select "Results and Activities, View Activities" from the menu at the top of the SIM screen.
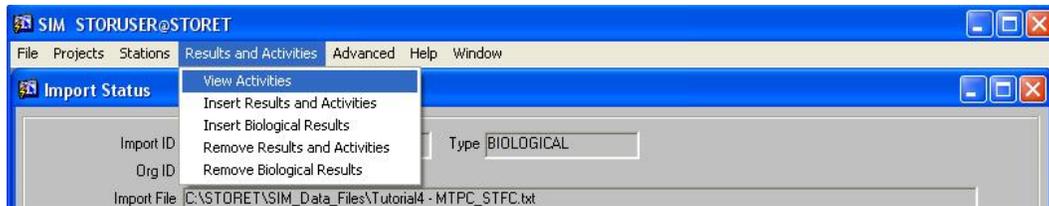
Select Trip ID "01-2000-2" and click on **Detail** to view the Activity Details window.

Click on all the buttons on the lower left corner of the Activity Details window (circled in the above screen capture) to view the imported (not yet migrated) data. When you are done, close the Activity Details window and return to the Import Status window.

In the Import Status window, click on the **Migrate Records to STORET** button.

Using the "Instructions" Tool

Sometimes in preparing data for STORET it is difficult to define a trip or identify the station visit number. Trip lengths are highly variable and can include multiple visits to a particular station. If these data are not collected as part of the standard data collection process for a state, adding this information after the fact can be very time consuming, especially for hundreds or thousands of data records. The "Instructions" tool allows SIM to automatically generate this information for you. The SIM "Instructions" tool is useful for

- ▪ Finding data elements that require a translation

- ▪ Autogenerating Trip IDs

- ▪ Autogenerating Station Visit Numbers.

During the years 2000 and 2001, Dr. Manning has gone every month to the Patuxent River mouth and performed counts of fish in different length ranges. The data file for the biological sampling events is provided in the file "Tutorial4b - Instructions.xls." The file does not contain trip ID numbers or station visit numbers. If you try to import this data into SIM using the configuration file you just created for this tutorial, you will generate a long list of error messages. This part of the tutorial demonstrates how to use the "Instructions" tool to avoid this.

First, save the provided MS Excel file as a tab-delimited file.

Open "Tutorial4b - Instructions.xls" in MS Excel. You should see blank columns of data under Trip ID and Station Visit Number.

Save this as a tab-delimited text file named "Tutorial4b - Instructions.txt" in your User Created directory.

Close MS Excel.

You will now modify the configuration file to use the "Instructions" tool to automatically generate the missing data so that the dataset can be imported successfully into SIM and migrated into STORET.

From the SIM Welcome window, click the **Manage Configurations** button.

From the Import Configuration List window, locate "Tutorial4 - MTPC_STFC."

Click on **Details** to open the Import Configuration window.
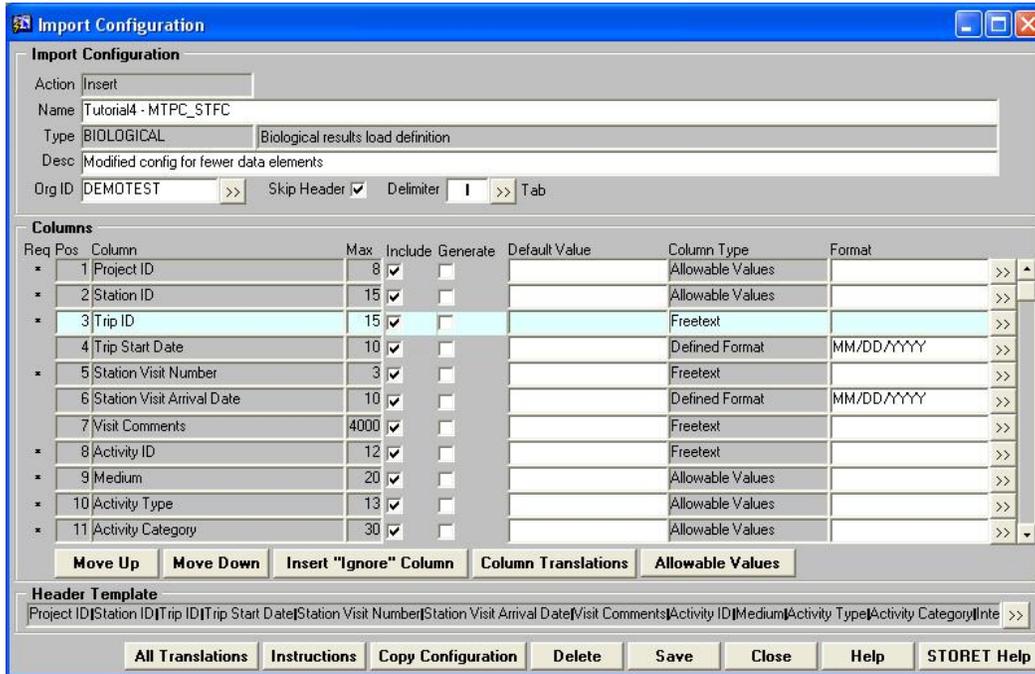
Click on the **Instructions** button at the bottom of the window to open the Import Instructions window.
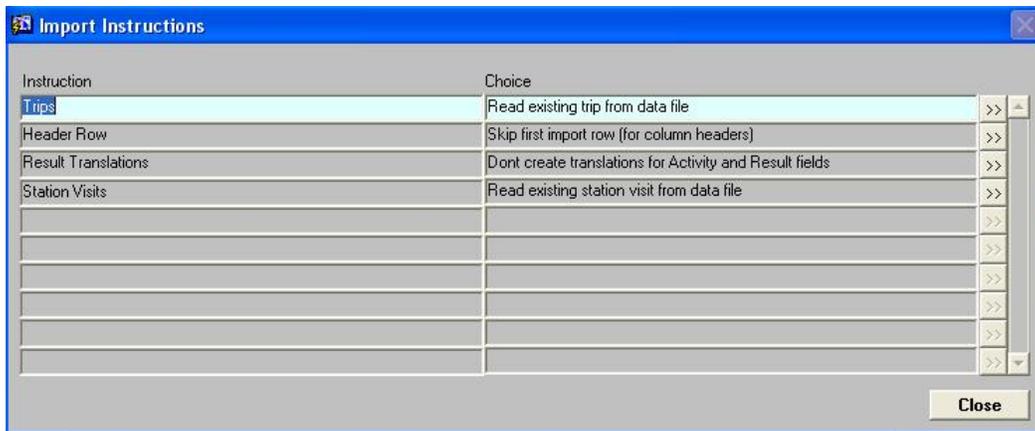
The "Instruction" column indicates the field or topic that the instruction relates to, and the "Choice" column shows the instruction choice.

Click in the Choice text box next to Station Visits (where it says "Read existing station visit from data file"). Click on the ">>" button next to the Choice field and select "Create one station visit per day." Click **OK**.
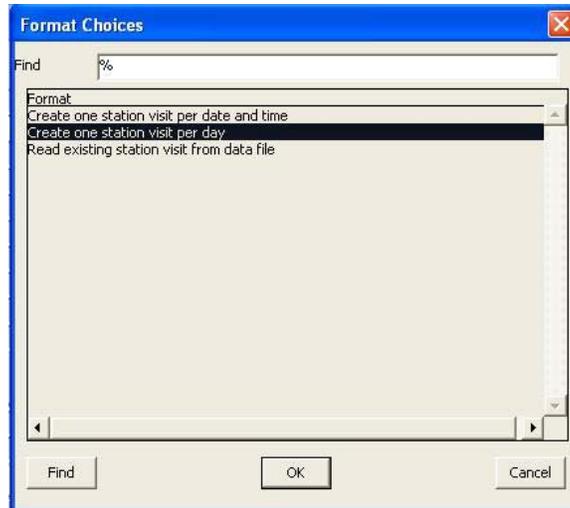
Similarly, for "Trips" select "Create one trip per year."

Keep the default options for "Results Translations" ("Don't create translations for Activity and Result fields") and "Header Row" ("Skip first import row" [for column headers]).

Close the Instructions window to return to the Import Configuration window.

Click on **Save** to save your configuration file, then close the Import Configuration window.

Based on these instructions, SIM will break the data into two trips, one for the year 2000 and one for the year 2001. The 12 sampling events in 2000 will have Station Visit numbers sequentially numbered from 1 to12, and the 12 sampling visits in 2001 will also be sequentially numbered from 1 to 12.

Now you will import the file, using this revised configuration file with instructions.

Return to the Welcome window.

Click **Import a File**, click the **Insert** button in the Import Action window, and click the **Biological** button in the Import Type window to bring up the "Import a File" window.

Choose the "Tutorial4 - MTPC_STFC" import configuration.

For the **Import File Name**, select the file you created earlier in this tutorial by clicking the ">>" button, navigating to the location of "Tutorial4b - Instructions.txt," and selecting it.

Click the **Import File** button to import the data into SIM. When the Skip Header message appears, click **Continue**. When the Import Complete message appears, click **OK**.
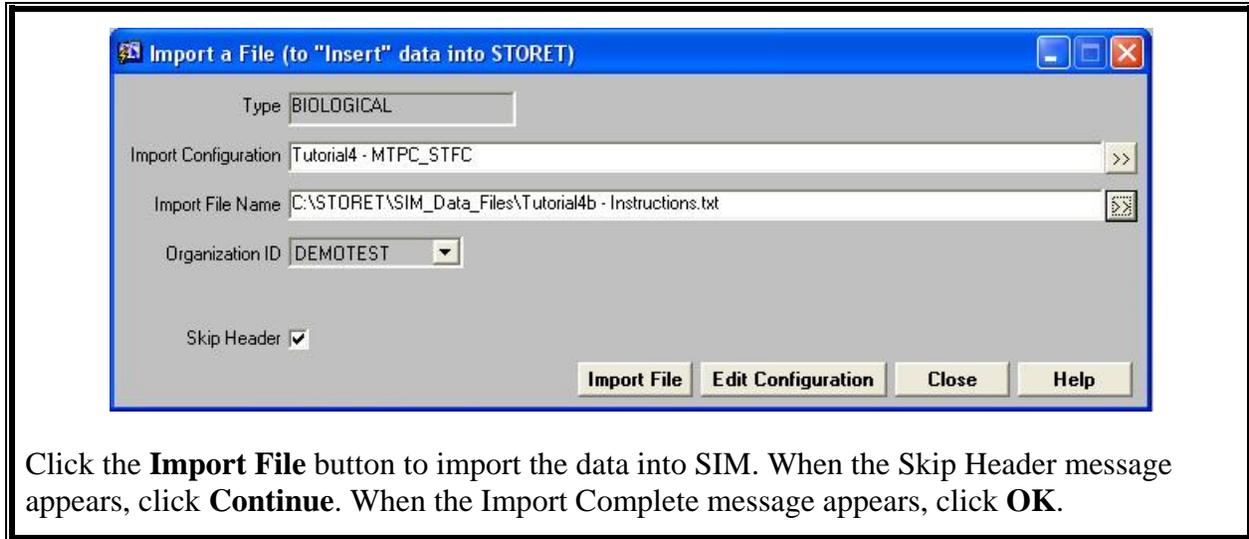
You should have 72 results covering 24 activities (2 trips of 12 visits each) successfully imported. Now you will migrate them to STORET.

Click on the **Migrate Records to STORET** button.

Exit SIM.
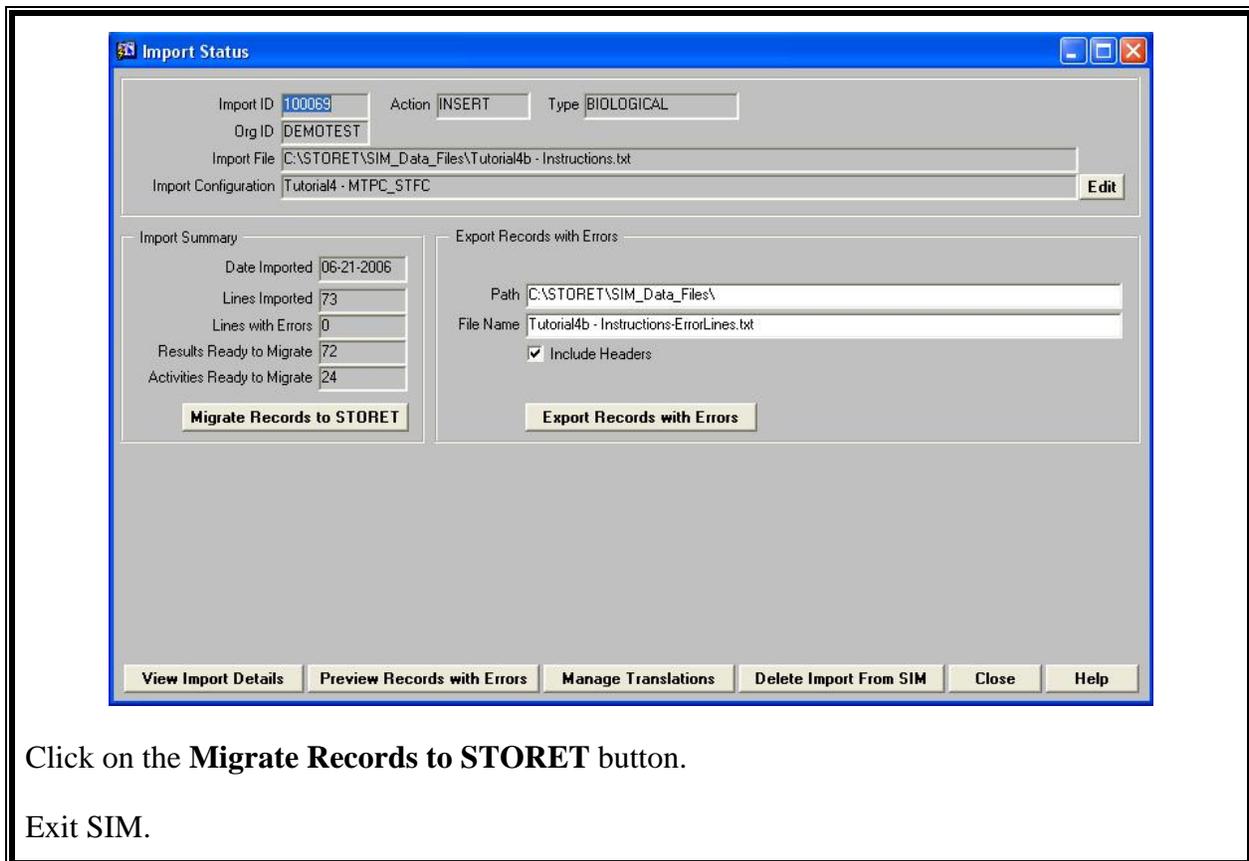
You have now learned how to import results data into STORET with SIM. Although initial preparation of configuration files may require a time investment, you will be able to use the same configurations to add similar data thereafter.