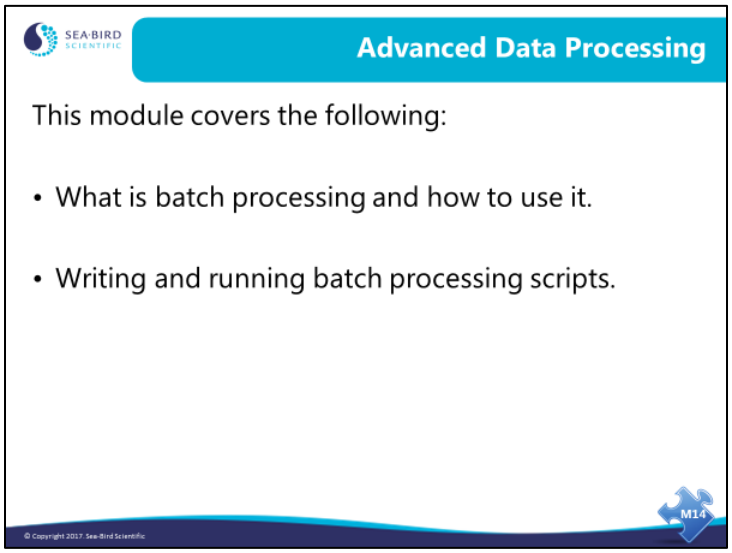
Advanced Data Processing: Batch Processing

Sea-Bird Scientific University Module 14

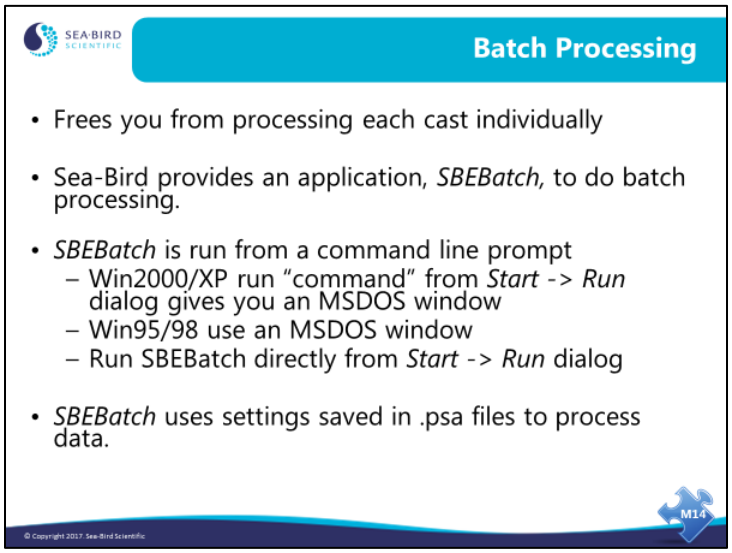
© Copyright 2017. Sea-Bird Scientific

Overview



This section of the course is the final topic in profiling. We will talk batch processing large numbers of files.

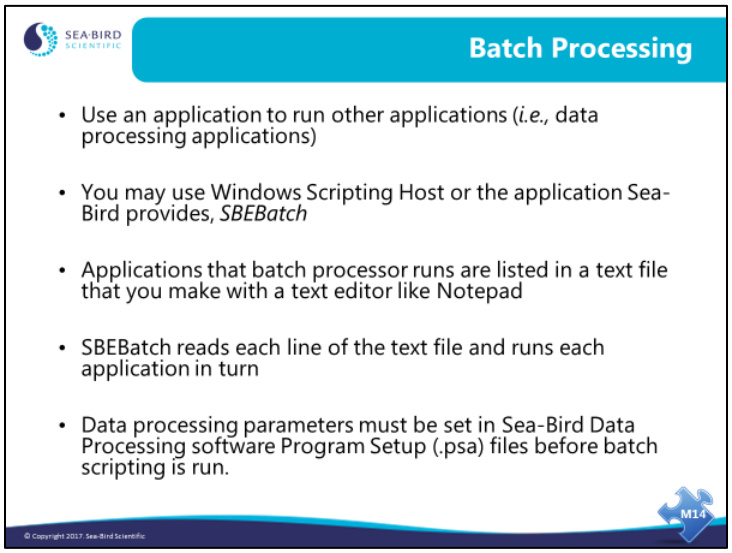
Data Processing Large Numbers of Files



For processing large sets of cast data, batch mode processing automates the job. You can use the windows scripting host or a program provided with SBE Data Processing, SBEBatch.exe. Your batch file can take advantage of command line parameters and wild card characters.

You can run SBEBatch from a DOS window or from the Windows Run dialog (Start -> Run). In all the examples we'll use today (and the examples in the SBE Data Processing manual), we're assuming you are running from the Windows Run dialog box.

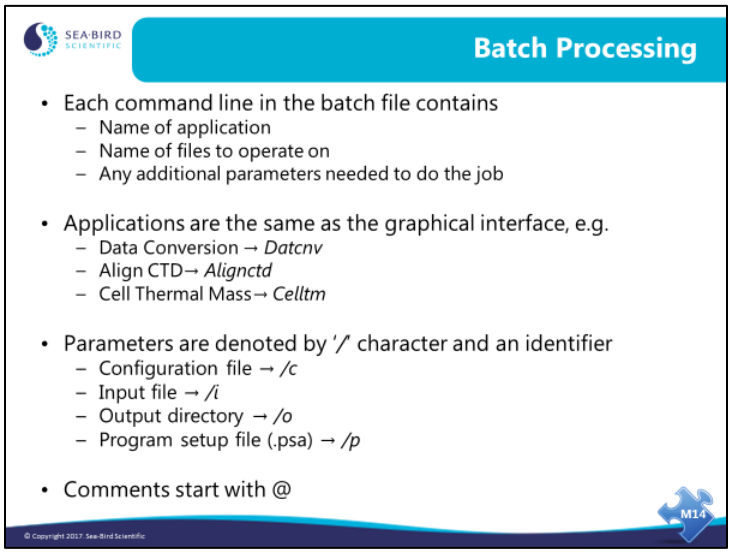
Data Processing Large Numbers of Files *(continued)*



Applications:

Module	Process Name
Align CTD	Alignctd
ASCII In	Asciin
ASCII Out	Asciiout
Bin Average	Binavg
Bottle Summary	Bottlesum
Buoyancy	Buoyancy
Cell Thermal Mass	Celltm
Data Conversion	Datcnv
Derive	Derive
Filter	Filter
Loop Edit	Loopedit
Mark Scan	Markscan
Sea Plot	Seaplot
Section	Section
Split	Split
Strip	Strip
Translate	Trans
Wild Edit	Wildedit
Window Filter	Wfilter

Data Processing Large Numbers of Files *(continued)*

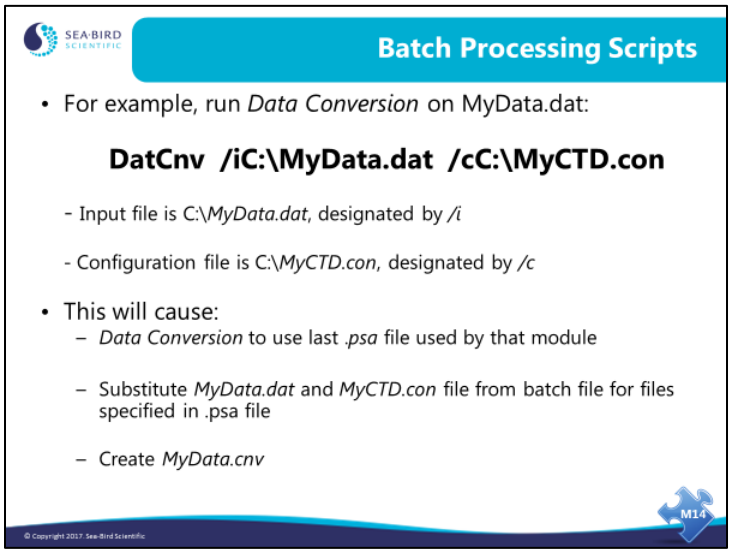


A list of the most commonly used parameters follows; see the SBE Data Processing manual for a complete list:

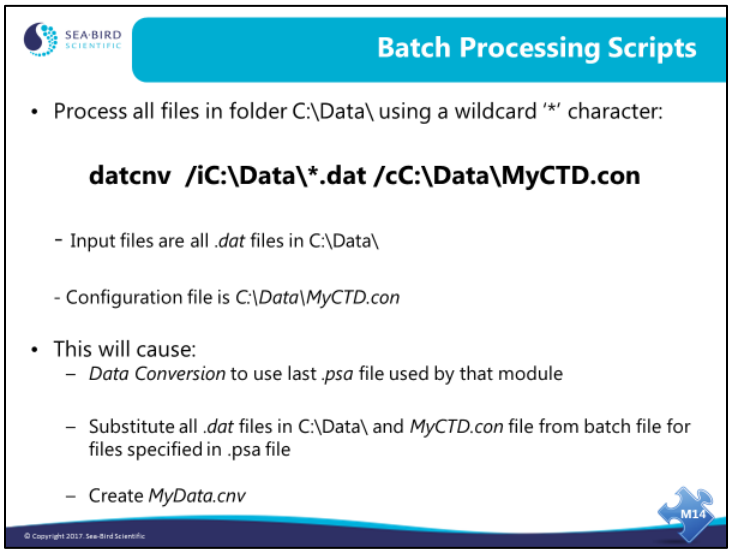
Parameter	Description
/cString	Use *String* as instrument **configuration (.con or .xmlcon) file**. *String* must include full path and file name. Note: If using /cString, must also specify input file name (using /iString).
/iString	Use *String* as **input file** name. *String* must include full path and file name. This parameter supports standard wildcard expansion: • ? matches any single character in specified position within file name or extension • * matches any set of characters starting at specified position within file name or extension and continuing until end of file name or extension or another specified character
/oString	Use *String* as **output directory** (not including file name).
/fString	Use *String* as **output file name** (not including directory).
/aString	**Append** *String* **to output file name** (before extension).
/pString	Use *String* as **Program Setup (.psa)** file. String must include full path and file name.
/xModule: String	Use *String* to define an additional parameter to pass to Module. Not all modules have x parameters; see module descriptions for details. If specifying multiple x parameters, enclose in double quotes and separate with a space. *Example*: Run Data Conversion, telling it to skip first 1000 scans: /xdatcnv:skip1000
#m	Minimize SBE Data Processing window while processing data, allowing you to do other work on computer.

If specifying multiple parameters, insert a space between each parameter in the list.

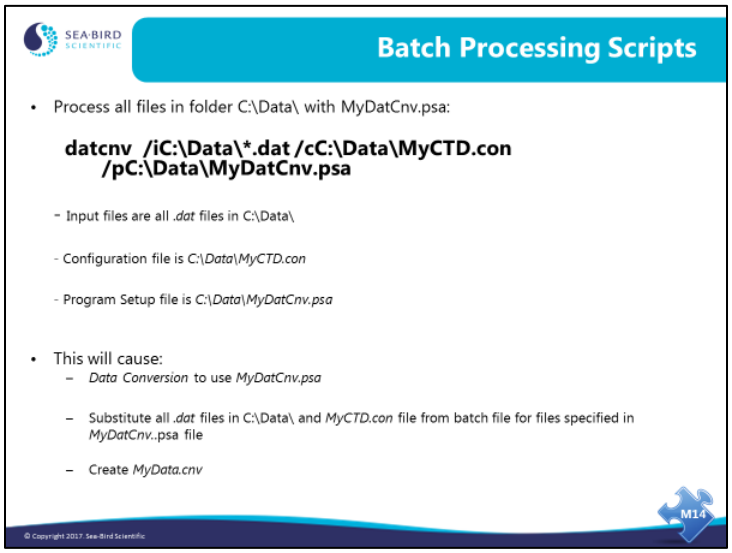
Writing Batch Processing Scripts



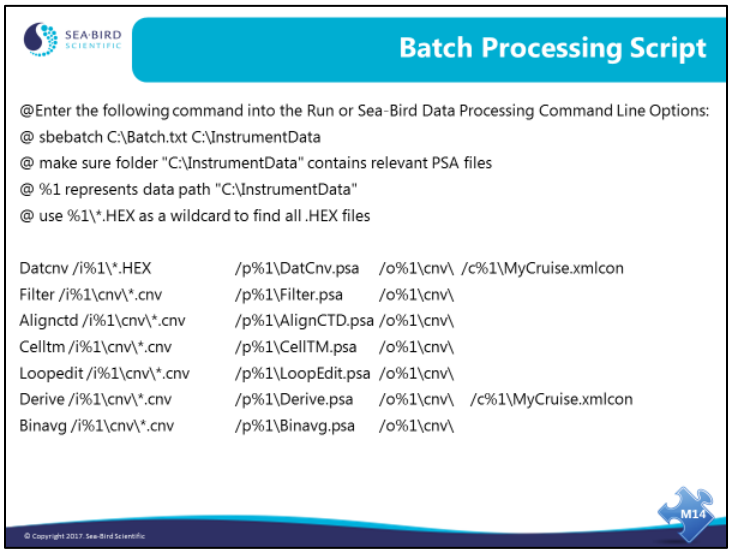
Writing Batch Processing Scripts *(continued)*



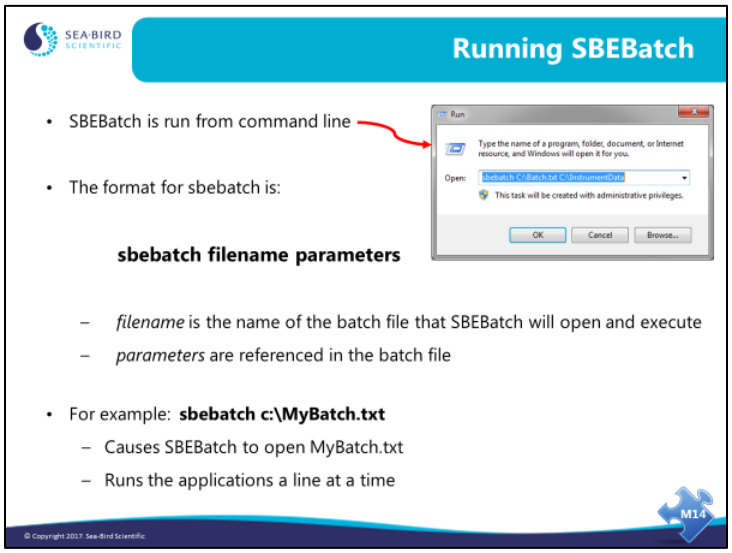
Writing Batch Processing Scripts *(continued)*



Writing Batch Processing Scripts *(continued)*



Running SBEBatch



Here's an example of how to use batch processing to run Data Conversion and Derive to process all the files in C:\Data:

1. Run Data Conversion, entering the desired choices in the File Setup and Data Setup dialog boxes. Upon completing setup, press Save or Save As on the File Setup tab. The configuration is stored in the .psa file.
   Repeat for Derive.

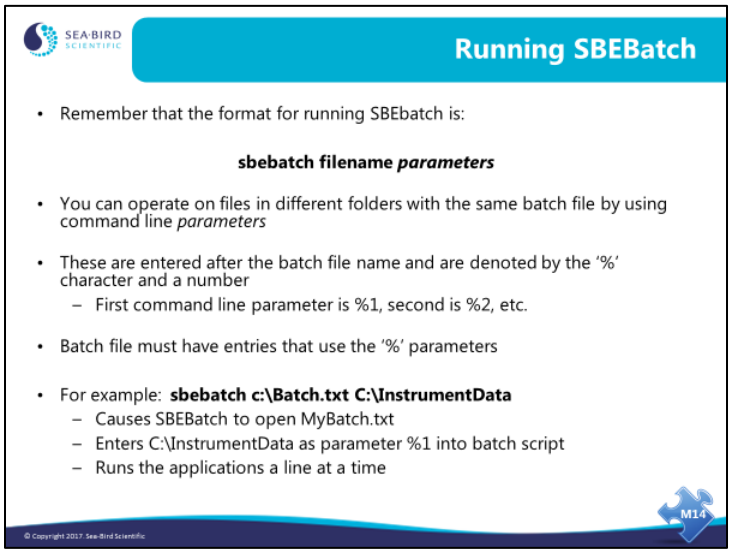
2. Create a batch file named batch.txt in C:\Data, which contains the following lines:
   @Lines starting with @ are comment lines, and have no effect on the result
   @Use these to document what you are doing in the batch file
   @Processing data from February 2006 Cruise
   datcnv /iC:\Data*.dat /cC:\Data\MyCTD.con
   derive /iC:\Data*.cnv /cC:\Data\MyCTD.con

3. Select Run in the Windows Start menu. The Run Dialog box appears.
   Type in the batch processing program name and the .txt file name:
   sbebatch c:\Data\batch.txt

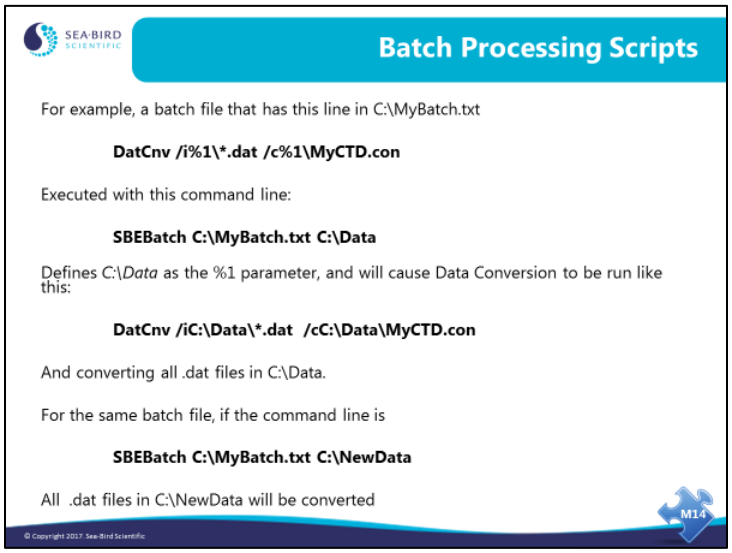
The results:

1. Data Conversion uses its last .psa file, substituting the .con file from the batch file for the .con file specified in the .psa file, and processes **all** .dat files in C:\Data, creating a .cnv file from each .dat file.
2. Derive uses its last .psa file, substituting the .con file from the batch file for the .con file specified in the .psa file, and processes **all** .cnv files in C:\Data (which were just created by Data Conversion), creating a .cnv file from each .cnv file.

Running SBEBatch *(continued)*



Running SBEBatch *(continued)*



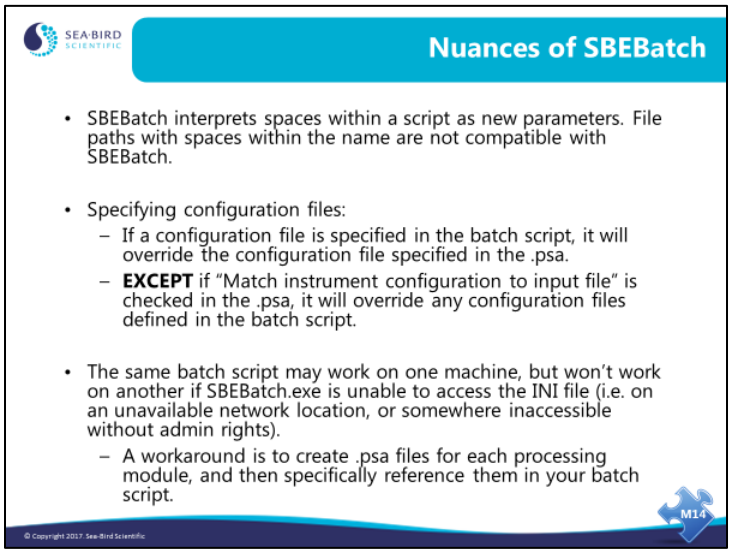
Now let's add a bit more flexibility to the process. Here's an example of how to use batch processing to run Data Conversion and Derive to process all the data files in C:\Data\Leg1, C:\Data\Leg2, and C:\Data\Leg3:

1. Run Data Conversion, entering the desired choices in the File Setup and Data Setup dialog boxes. Select *Match instrument configuration to input file* on the File Setup tab. Upon completing setup, press Save or Save As on the File Setup tab. The configuration is stored in the .psa file.
   Repeat for Derive.
2. Create a batch file named batch.txt in C:\Data, which contains the following lines:
   @Processing data from 3 legs of February 2006 Cruise
   datcnv /i%1*.dat
   derive /i%1*.cnv
3. Select Run in the Windows Start menu. The Run Dialog box appears.
   Type in the batch processing program name, the .txt file name, and the %1 parameter:
   sbebatch C:\Data\batch.txt C:\Data\Leg1
   Repeat for the files in Leg2 and Leg3:
   sbebatch C:\Data\batch.txt C:\Data\Leg2
   sbebatch C:\Data\batch.txt C:\Data\Leg3

The results:

1. Data Conversion uses its last .psa file, substituting the matching .con file for the .con file specified in the .psa file, and processes **all** .dat files in C:\Data\Leg1, creating a .cnv file from each .dat file.
2. Derive uses its last .psa file, substituting the matching .con file for the .con file specified in the .psa file, and processes **all** .cnv files in C:\Data\Leg1 (which were just created by Data Conversion), creating a .cnv file from each .cnv file.
3. Steps 1 and 2 are repeated for the files in C:\Data\Leg2 and C:\Data\Leg3.

Nuances of SBEBatch



Activity: Use Beach Scripting to Process Data

Explanation: We will prepare Data Conversion and Bin Average to operate with your batch file, by setting up the desired output parameters, bin size, etc. The saved .psa files contain all the information on the setup. And then we will create a batch processing script to process a large number of files in the same way.

1. In SBE Data Processing, run *Data Conversion*:
   - Use C:\Data\Module9\Batch\Puget00.hex
   - Downcast only
   - Calculate Pressure, Temperature, and Salinity
   - Save .psa file as C:\Data\Module9\Batch\DatCnv.psa

2. In SBE Data Processing, run *Bin Average* to create 1 decibar bins:
   - Use C:\Data\Module9\Batch\Puget00.cnv
   - Name append *B*
   - Data Setup
     Choose Pressure for *Bin Type*
     Enter *Bin size* of 1
     Process the downcast
     Make other selections as desired

3. Delete Puget00.cnv and Puget00B.cnv.
   - Delete these output files from Data Conversion and Bin Average because they will be recreated by the batch processing script.

4. Create a batch processing script to process **all** files in C:\Data\Module9\Batch (Puget00.hex, Puget01.hex, etc.):
   - Use Notepad to write your batch file, using a %1 parameter for the file locations.
   - Save your batch file as C:\Data\Module9\Batch\MyBatch.txt

5. Launch a command line session: Click Start > Run.

6. Run your batch file with the command line:
   Sbebatch C:\Data\Module9\Batch\MyBatch.txt C:\Data\Module9\Batch

7. Check to see that you were successful with Notepad