DSM2 Bay-Delta Tutorial 4: Batch Preprocessing

Purpose: This tutorial will demonstrate how to preprocess a number of CalSim output files, each of which represents a different alternative – we will look at three alternatives, but the techniques apply to large numbers of alternatives just as well. In the process of this tutorial, you should become more familiar with how DSM2 and CalSim label their simulations and scenarios and a learn a little bit about batch files

CalSim Files: A typical situation with planning studies is that the input scenarios are represented by different CalSim output files. Sometimes these files reside in a directory structure that follows a pattern, for instance the first two alternatives might look like this:

```
C:/calsim
   /altname1
     /dss
        /d1641
           2020d09edvsa.dss
  /altname2
     /dss
        /d1641
           2020d09edvsa.dss
```

Note that this scheme CalSim uses directory structure to differentiate its output – the files and pathnames are identical.

Another system you may encounter is one where the CalSim files themselves are named after the scenario:

```
C:/calsim
   /altname1_2020d09edvsa.dss
   /altname2_2020d09edvsa.dss
```

Preprocessor requirements:

The DSM2 preprocessing scheme requires three pieces of information for each scenario:
   1. The DSM2 name we want to give the scenario (will become DSM2MODIFIER).
   2. The directory in which the CalSim output is found (will become CALSIMDIR)
   3. The name of the CalSim file (minus the .dss part – will become CALSIMNAME)

So for the first example above
 DSM2MODIFIER=altname1
CALSIMNAME=2020d09edvsa
CALSIMDIR= `c:/calsim/altname1/dss/d1641`

How you will get this information into the preprocessing system depends on approach. We will look at two, but if you are an experienced script writer you will immediately see lots of possibilities.

Two approaches for batch jobs:

For larger studies, you have some choices as to how to set things up. We'll look at a few that may help you get started, while experienced script writers are likely to come up with numerous interesting variations. These exercises will guide you in setting up modest batch processing and familiarize you a bit more with the concept of environmental variables at the command line and in windows "batch" scripts (files with a *.bat extension that list commands for the command line).

1. You can create configuration files for each alternative, e.g. *config_alt1.inp*, *config_alt2.inp*. In each configuration file you hard-wire the information that is required is hardwired for that scenario. This method record of each scenario for people who inherit your study. It is a good choice when the number of alternatives is small. It is also a good choice when things other than CalSim vary between alternatives.
2. Alternatively, you can create a single configuration file that points the three scenario-related variables to generic values. Then you use a batch_prepro.bat script to loop through the scenarios. When the number of simulations is very large (say 100 climate change scenarios) and the only difference in the inputs is CalSim, this method is efficient.

Now let's go through the exercises and check out the details.

Method 1: Using separate configuration files:

1. Create the configuration files:

a. In windows, navigate to *{DSM2_home}**tutorial**ocap_sdip*. The alternatives we are using have generic sounding names, but they are compatible with OCAP assumptions.

b. Copy the configuration file config_sdip_ocap_71.inp to config_alt1.inp

c. Make sure the study dates cover the full 1974-1991 period for planning runs. It is usually a good idea to preprocess the whole period, even if you are going to do run dsm2 on a subset of the simulation period.

d. Replace the three variables indicated below. The three lines may not be next to one another.

```
<file config_alt1.inp>
ENVVAR
NAME            VALUE
[other definitions…]            # NOTE: LINES SHOWN MAY NOT
                                # BE TOGETHER

CALSIMNAME      2005a01edv      # File name, minus .dss
DSM2MODIFIER    alt1            # DSM2 name for alternative
CALSIMDIR       ..data/calsim/alt1 # CalSim output directory
END
```

e. Copy the file config_alt1.inp to config_alt2.inp. Repeat step (d) using alt2 as the DSM2MODIFIER.

f. Prepare *hydro.inp* and *qual.inp* to handle a generic configuration file by making the name of the configuration file at the top of each an ENVVAR. We will be providing this from the command line or batch file – as an operating system environmental variable.

```
<file hydro.inp>
CONFIGURATION
${CONFIGFILE}       # Changed
END
… [other data]
```

g. Prepare a batch file for preprocessing. It will have one line per alternative. Notice the "call" statement – this is the best way to call a succession of other batch files (prepro is itself a batch file called prepro.bat).

```
<file study_prepro.bat>
call prepro config_alt1.inp
call prepro config_alt2.inp
```

h. At the command prompt, launch the preprocessing by typing:

```
> study_prepro.bat
```

i. Now create a batch file that launches QUAL and HYDRO for every alternative in the study. For each alternative, you must set the environment variable CONFIGFILE, then launch the models.

```
<file study.bat>
SET CONFIGFILE=config_alt1.inp
hydro hydro.inp
qual qual_ec.inp
SET CONFIGFILE=config_alt2.inp
hydro hydro.inp
qual qual_ec.inp
```

j. Launch the study batch file by typing at the command prompt:

> study.bat

Method 2: Batch file that loops

1. Create a generic configuration file:

a. In the looping method, we are going to describe the alternatives in a text file and loop through the text file. First we need a configuration file that is generic. Let's begin by copying *config_sdip_ocap_71.inp* one more time to a file called *config_study.inp.* Change the 3 variables (DSM2MODIFIER, CALSIMNAME and CALSIMDIR) as follows.

```
<file config_study.inp>
ENVVAR
NAME            VALUE
[other definitions…]
CALSIMNAME     ${BATCH_CALSIMNAME} # File name, minus .
DSM2MODIFIER   ${BATCH_DSM2MODIFIER}
CALSIMDIR      ${BATCH_CALSIMDIR}  # CalSim output directory
dss
# DSM2 name for alternative
[other definitions…]
END
```

2. Create the scenarios.txt file

a. In the study folder, create a file called scenarios.txt

b. On each line of the file, put the scenario name (DSM2MODIFIER), directory (CALSIMDIR) and file name (CALSIMNAME) minus the ".dss" extension.

```
<file scenarios.txt>
alt1,../data/calsim/alt1,2005a01edv
alt2,../data/calsim/alt2,2005a01edv
```

3. Launch batch_prepro.bat

a. In the study directory, obtain a command prompt and type:

```
> batch_prepro config_study.inp scenarios.txt
```

b. Note: if the batch_prepro script fails for a particular scenario after running others successfully, first fix the problem and eliminate the failed (half-processed) scenario. Then avoid re-running the successful scenarios by adding the "resume" tag, for example:

```
> batch_prepro config_study.inp scenarios.txt resume
```

If you type this command now, batch_prepro.bat will harmlessly do nothing.

4. Examine and use the preprocessing products

a. The preprocessing product is a HEC-DSS file for each scenario in the local time series directory. You should have one file per scenario.

b. If you are doing this tutorial on your own, you may choose to launch dsm2 on each alternative. To do this, change the configuration file in *hydro.inp* and *qual_ec.inp* to the generic one:

```
<file hydro.inp>
CONFIGURATION
config_study.inp
END
```

c. Use batch_run.bat with the same syntax as you did for batch_prepro::

```
> batch_run config_study.inp
```

Note that you may need to modify this script if you use it for something other than qual_ec. We may not be able to run the simulations in class because of the time required – but if you have extra time, change the dates to a one year (1991) and try it out.