

9.A

Command Line Data Aggregation Program

Robert A. McDougall

This chapter describes the command line data aggregation program , also known as FLEXAgg, which is included in the GTAP 5 data package. It is based on user documentation shipped with the package.

9.A.1 System Requirements

The package has been tested on Windows 98. It will likely work with other WIN32 operating systems. It may even work with Windows 3.x, if you unzip the included DPMI server archive 'csdpmi4b.zip' (see section 4, "Contents").

9.A.2 Preparing Aggregations

To prepare an aggregation, you need to provide the data base to be aggregated, prepare an aggregation mapping file and run the aggregation program.

Providing the Data Base to be Aggregated

The program aggregates a data base consisting of four files:

- GSDSET.HAR : sets
- GSDPAR.HAR : parameters
- GSDDAT.HAR : data
- GSDVOLE.HAR : energy volumes

These files should be present in the directory where you install this package. With the full data package, they come with the package; with the aggregation program package, it is up to you to provide them. With the full data package, the files comprise the standard GTAP 5 data base; with the aggregation program package, you may use any data base with the GTAP 5 structure (so, for example, you may use the aggregation program to aggregate an already-aggregated data base).

Preparing an Aggregation Mapping File

To prepare an aggregation mapping file, make a copy of the file INPUT.TEM, and edit it to suit your requirements. For details on layout, see the instructions included in INPUT.TEM itself. For an example of a completed mapping file, see the file `tstin\TEST335.TXT`. Save the edited file with a name ending in the extension .TXT, for example, MY_AGG.TXT.

Running the Aggregation Program

The aggregation program is DATA-AGG.BAT. To run it:

Make the directory with the GTAP data package your current directory. For example, if you have installed the GTAP data package in the directory C:\GTAP, you might enter the command

```
cd C:\GTAP
```

Invoke the program, using either:

— syntax 1:

```
DATA-AGG map-file
```

or

— syntax 2:

```
DATA-AGG [map-dir\]map-file [output-dir] [work-dir]
```

Note:

- For *map-file* substitute the name of your aggregation mapping file, omitting the extension .TXT. For example, if your mapping file is MY_AGG.TXT, substitute MY_AGG for *map-file*.

To use syntax 1, your mapping file must be in the same directory as the GTAP data package. With syntax 2 you can specify the location of your mapping file, the aggregated data files and intermediate files. You exercise this control through the optional arguments shown in square brackets. (Don't type the brackets in your command line; they're there only to indicate that the arguments are optional.) Substitute as follows:

`map-dir` the directory containing your mapping file
`output-dir` the directory to contain the aggregated data files
`work-dir` the directory to contain all intermediate working files.

For example, to use a mapping file `test335.txt` in the local subdirectory `tstin`, writing the aggregated data files to a local subdirectory `tstout`, and the working files to `tstwrk`, enter

```
data-agg tstin\test335 tstout tstwrk
```

This example is ready to run as a test case.

- Besides the aggregated data files, the program will create report files `gview.har` and `gtax.har` containing summary information on the data base. To inspect these files we recommend the GEMPACK utility `ViewHAR`.

Further details:

- With syntax 1, the program behaves as in GTAP release 2.
 - The program will look in the GTAP data package directory for a mapping file with the extension `.TXT`.
 - It will create a subdirectory with the same base name as your mapping file. For example, if your mapping file is `my_agg.txt`, it will create a subdirectory `my_agg`.
 - It will write the aggregated data files `gset.har`, `gpar.dat`, `gdat.har`, `gvole.har`, `gtax.har` and `gview.har` to the new subdirectory.
 - It will leave intermediate working files in the current directory.
- After running an aggregation with syntax 1, you can remove intermediate files from your GTAP directory with the command:


```
aggclean
```
- With syntax 2, the program behaves as follows:
 - It looks for the mapping file in the directory indicated by the argument *map-dir*. If you do not supply a value for *map-dir*, it looks in the GTAP data package directory. If you supply a file name but the file does not exist, it stops with an error.

- It writes the aggregated data files `gset.har` to the directory indicated by *output-dir*. If you do not supply a directory for *output-dir*, it writes the files to the GTAP data package directory. If you supply a directory name but the directory does not exist, it stops with an error.
- It writes intermediate working files to the directory indicated by *work-dir*. If you do not supply a value for *work-dir*, it writes the working files to the same directory as the aggregated data files. If you supply a directory name but the directory does not exist, it stops with an error.

Trouble Shooting

error message "environment space insufficient"

- In MS Windows 9x or NT, you can increase your environment space as follows:
- Open a DOS window.
- Press Alt + Spacebar
- A drop-down menu appears. Press P (for "Properties").
- A dialogue box COMMAND Properties appears.
- Press Ctrl+Tab twice to select the *Memory* page.
- Press Alt+V to select the *Initial environment* text box.
- Enter a new environment size (2048 should be adequate for most users).
- Press Enter to exit [*sic*].

error message "error in map file"

This message can be generated by an error in your mapping file. An error describing the error in the mapping file may be found in the intermediate working file `gmap.txt`. This is in the main GTAP directory, if you used syntax 1, or in your selected work directory, if you used syntax 2. See also below under the heading "Mapping file errors".

error message "user-specified directory or file does not exist"

This message is written if the program cannot find the mapping file (syntax 1 and 2), or if it cannot find the output or work directories (syntax 2 only).

One way to generate this error is to include the extension `.TXT` in the *map-file* parameter. For example, the command

```
data-agg MY_AGG.TXT          WRONG!
```

would generate this error.

Mapping File Errors

(This is taken from earlier documentation by Randy Wigle.)

The following error checks are made on the input map (the .TXT file.)

- If an aggregated commodity is referred to in section 2 of the mapping, but it has not been defined in section 1, an error is noted on the first occurrence and the program exits
- If an aggregated region is referred to in section 4 of the mapping, but it has not been defined in section 3, an error is noted on the first occurrence and the program exits.
- If a line or section of the file is not set up properly, an error is noted and the current line is printed. NOTE that this will not always be the incorrect line! (In particular, if you leave out one of the " = = " lines, the line where the error prints will usually be later in the file.
- If an aggregated commodity is duplicated in section 1, an error will be generated, but the program usually chokes in section 3 (it is finding two matching aggregated commodities for one disaggregated commodity.)
- If an aggregated commodity or region name violates GEMPACK rules (e.g. if it is longer than 12 characters, or if it includes characters other than letters, digits and underscores, or if it doesn't start with a letter), then an error will be generated and the program terminates. See the discussion in the file INPUT . TEM.

9.A.3 Legalese

As always, to the extent permitted by applicable law, THE GTAP PACKAGE COMES WITH ABSOLUTELY NO WARRANTY.

9.A.4 Contents

Unpacked, the distribution contains the following files:

User Information

README this file

GTAP Standard Data Base

- gsddat.har
- gsdpar.dat
- gsdset.har
- gsdvole.har

These files are not included in the command line aggregation package, but are included in the full database package.

To limit the size of the distribution, we have not included the corresponding GTAPView files.

Bundled Freeware

A number of programs from outside sources are needed for unpacking and aggregating the data base:

<code>awk.exe</code>	from DJGPP, stub for gawk
<code>gawk.exe</code>	from DJGPP, text processing language
<code>modhar.exe</code>	from COPS/IMPACT, modifies HA files
<code>rd1hmd.exe</code>	from COPS/IMPACT, reads header from GEMPACK-formatted text file
<code>seehar.exe</code>	from COPS/IMPACT, converts HA files to text
<code>csdpmi4b.zip</code>	from Charles Sandmann, DPMI server, manages memory for gawk under Windows 3.x or DOS

In accordance with the licence for awk and gawk (see COPYING), we offer program source code on request to purchasers of the data package, within three years of the original purchase, charging for distribution only. More conveniently, however, users seeking documentation and source code can obtain it through the site from which we obtained the executables:

`ftp://ftp.simtel.net/pub/simtelnet/gnu/djgpp/v2gnu/`

These programs should run with all reasonably current versions of MS-DOS and MS-Windows (and some less-than-reasonably current); if nevertheless you do encounter problems, please advise us (see *Contact* below).

Aggregation Mapping Files

Some files are included to assist you in specifying mappings for data base aggregations

<code>.\INPUT.TEM</code>	template for aggregation mapping files
<code>.\tstin\TEST335.TXT</code>	sample aggregation mapping file

Aggregation Program Files

The aggregation program lets you reduce the size of the data base by aggregating across sectors and regions. For user instructions, see section 9.A.2.

AGGMAP . AWK
 DATA-AGG . BAT
 ENVSPCER . MSG
 GBARUN . BAT
 GSET . PRE
 G_USAGE . MSG
 MAPERROR . MSG
 NODIR . MSG
 NOTBASE . AWK
 PP . AWK
 PREDIM . AWK
 PREEND . AWK
 PREETR . AWK
 PREREG . AWK
 PRESEC . AWK
 README
 aggclean.bat
 aggdatt.*
 aggparr.*
 gtapview.*
 premar.awk
 preslg.awk

Miscellaneous Files

. \AGGCLEAN . BAT cleans up after an aggregation run

Additional Documentation

COPYING copyright notice for awk

Source Files for Aggregation Program

src\aggdat.tab
 src\aggpar.tab
 src\aggvole.tab
 src\gtapview.tab

9.A.5 Contact

If you experience problems with this package that you cannot resolve by reference to the documentation, please contact us by email at gtap@ftp.purdue.edu.