

Preparing global DEM data for QSWAT

Chris George

March 2015

Version 2

1 Introduction

You may have your own digital elevation map (DEM) for your watershed, but if you don't this document explains how to obtain and prepare DEM data obtainable free from the web. The techniques for clipping and reprojecting DEMs may also be useful for DEMs from other sources.

There are also notes about masking DEMs and about dealing with large DEMs in the final two sections that should be of general interest.

2 Obtaining SRTM DEM data

SRTM (Shuttle Radar Tomography Mission) is a good source of DEM data for almost anywhere in the world. It is available at 90 metre and, from the end of 2014, 30 metre resolutions.

The first thing to do is to make a note of the extent of the watershed you want to model: the maximum and minimum latitudes and longitudes. Decimal longitudes that are west of the Greenwich meridian should be expressed as negative numbers: if you have a longitude above 180°, subtract 360. Decimal degrees, rather than degrees, minutes and seconds, will be required for clipping later, so convert yours if necessary (just divide the minutes by 60 and round to 1 decimal place) and make a note of them.

Choose or create a folder for the maps to be stored. We will refer to this as the *map folder*.

2.1 90m data

You might like to use the less detailed 90m data if your watershed is large or if you are new to QSWAT. Go to <http://srtm.csi.cgiar.org/>, and select the *SRTM Data Search and Download* menu item. Then select *GeoTiff* as the file format. You can now select tiles by selecting *Multiple Selection* and clicking on them, or by selecting *Input Coordinates* and entering the extent. Then click *Begin Search* and on the next page select the *Data Download (FTP)* or *(HTTP)* for each of your tiles. You will not need the mask. Each tile comes in a zipped archive file containing a text file with information about SRTM (including citation data) and a .tif file, plus .hdr and .tfw files. Unzip each archive and put the files in the map folder.

2.2 30m data

Go to <http://earthexplorer.usgs.gov/> and register if it is your first visit, or log on. You won't be able to download anything until you are logged on.

Under *Coordinates*, select *Decimal* and click *Add Coordinate*. Enter, the latitude and longitude of one of the 4 corners of the extent. Add the other 3 corners similarly. Then above *Enter Search Criteria* select the second tab, *Data Sets*.

Find *Digital Elevation*, click on the plus sign to expand and find *SRTM*. Expand that and check *SRTM 1 Arc-Second Global*. Then go straight to the fourth tab, *Results*.

After a few seconds you should see the tiles you need listed. For each one. Click on *Download options* (the 5th symbol) and then click *Download* beside *GeoTIFF*. Save the file, close the *Download Options* form, and repeat for each of your tiles. Note that the save form says the file is a zipped archive, but its extension is .tif and it is not zipped. Move the .tif files to the map folder.

3 Clipping the grids

Start QGIS and load the first grid (using *Layer* → *Add Raster Layer*, or the corresponding button on the left of the QGIS window). Select *Raster* → *Extraction* → *Clipper* to bring up the *Clipper* form (Figure 1). Set the grid to the *Input file*, type `<grid_name>_clip_ll` as the *Output file*, check *No data value*, and enter the appropriate value, which is -32768 for 90m data and -32767 for 30m data¹. You can find the appropriate value by right clicking on the entry in the *Layers* panel, selecting *Properties*, and selecting the *General* tab. The no data value is included in the *Layer info* box.

Use *Extent* as the *Clipping mode*, and enter the extent you used to download the grids. It is a good idea to widen these somewhat, say by 0.2 of a degree, to make sure there are no problems. *x* here means longitude and *y* is latitude. Make a note of the values you use as you may need them again. The extent will be shown as a rectangle on the map canvas, so you can see it is where you expect it to be. Check *Load into canvas when finished* and click *OK*. You may get a warning about going outside the raster, which you can ignore if this is what you expected. Don't close the *Clipper* form, so that you can keep the no data value and extent. Right click on the original, unclipped grid in the *Layers* panel, select *Remove* and confirm. You will be left with the clipped portion. Load each of the other grids, and clip in the same way. Remember to change the output file name each time or you will overwrite a previous clip. Close the *Clipper* form when you have clipped all your grids and checked that the pieces abut to form a rectangle as expected.

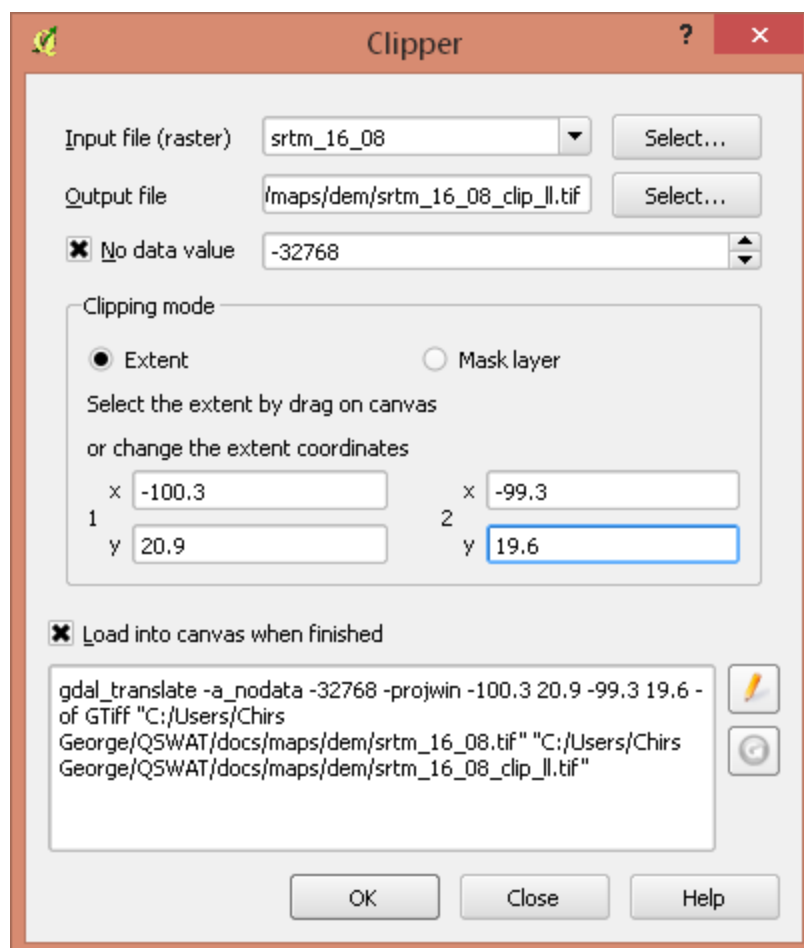


Figure 1: *Clipper form*

¹ The value should not matter, as long as it is less than any occurring in the actual data, but if it is different from the existing no data value it can cause problems with the existing value being accepted as an actual data value.

4 Merging the grids

You should have all the clipped grids loaded. Use *Raster* → *Miscellaneous* → *Merge* to open the *Merge* form (Figure 2). Click the *Select* button next to *Input files* and in the next form navigate to the map folder. Hold *Ctrl* and click on each of the files ending in *_clip_ll.tif*. Click *Open*. You will be back in the *Merge* form. Click *Select* by *Output file*, which should return you to the map folder, else navigate to it, choose a suitable name, perhaps ending in *dem_ll* to remind you what it is, and click *Save*. Back in the *Merge* form, check *No data value* and enter the value you used earlier, e.g. -32768 for 90m, -32767 for 30m. If necessary, check *Load into canvas when finished*, and click *OK*. Check the next two messages don't tell you of errors, and click *OK* to close each of them.

Check your merged DEM is OK by removing the visibility of the other layers. If all is well you can *Close* the *Merge* form.

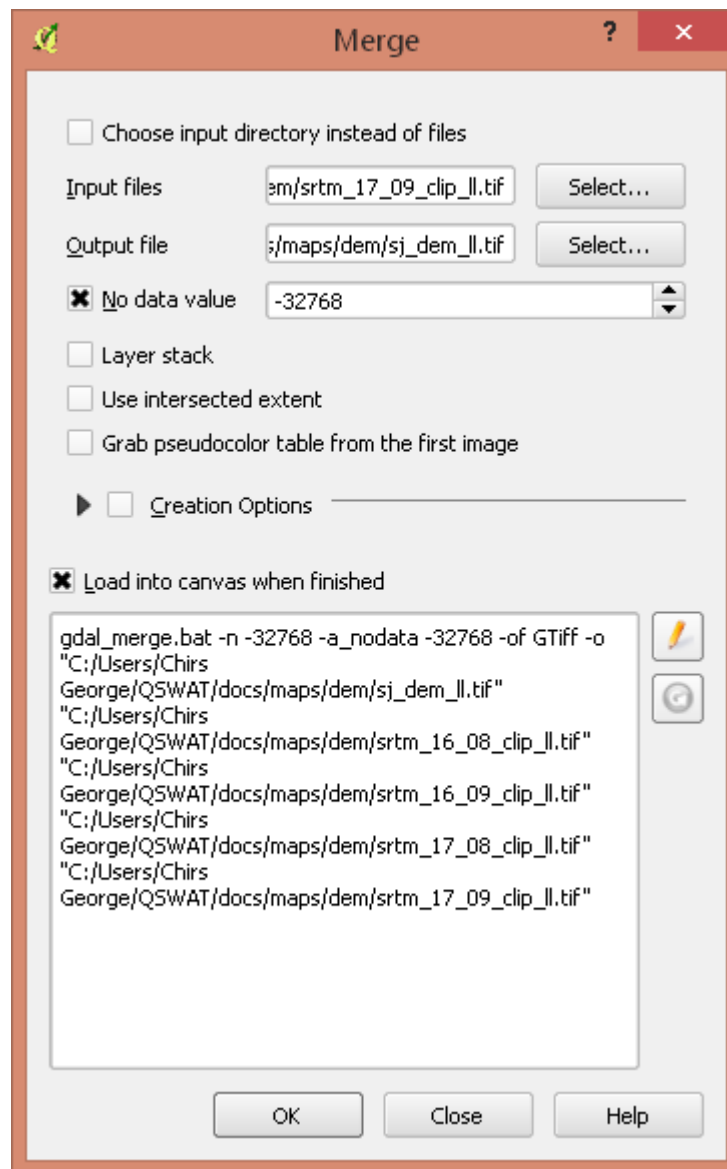


Figure 2: Merge form

5 Reprojecting the grid

Our final task is to reproject the merged grid into an “equal area” projection, i.e. one in which each grid cell has the same area. The current lat-long projection is not appropriate because at different latitudes the same amounts of longitude will represent different distances. Universal Transverse Mercator (UTM) is a good general choice for the new projection. It is available almost everywhere, and while not exactly an equal area projection is sufficiently close for most SWAT modelling. We need a suitable UTM zone. We can find the zone from a map like <http://www.dmap.co.uk/utmworld.htm>, or use the formula $(x + 180)/6$ rounded up to the next integer, for a reasonably central longitude x in our area. So, for example, if x is -100 the UTM zone will be $80/6$, or 13.3, rounded up, to give 14. Then we note whether we are in the northern or southern hemisphere.

Use *Raster* → *Projections* → *Warp (Reproject)* to open the *Warp (Reproject)* form (Figure 4). Select the merged grid as the *Input file* from the pull-down menu. Click *Select* next to *Output file*, navigate if necessary to the map folder, and type in a suitable name, perhaps ending in *dem_utm* to remind you what it is.

Check the *Target SRS* box and click the *Select button* beside it to bring up the *Select the target SRS* form (Figure 3). For zone 14 in the northern hemisphere, for example, we need to find WGS84 / UTM zone 14N. If we have used it recently it will appear in the upper *Recently used* box. Otherwise we have to find it in the lower box. It will appear under ***Projected Coordinate Systems*** rather than ***Geographic*** and can be found much closer to the bottom than the top. Note that it has an EPSG number of 32614, while the corresponding zone in the southern hemisphere has an EPSG number of 32714. So if you know the zone and the north or south hemisphere you can calculate the EPSG number. QGIS does not support finding the CRS by EPSG number, however. On the other hand, once you have used it it will appear in the recently used list, making it much quicker next time. Make sure the right zone appears in the *Selected CRS* box, and click *OK*.

Check *No Data values* and put in the same no data value, -32768 for 90m, -32767 for 30m. Uncheck *Load into canvas when finished*: the projection is different from the current lat-long projection and should be viewed in a new QGIS window.

TauDEM recommends that DEMs have square cells. Gdalwarp, the tool that we are using to do the reprojection, has a parameter to generate an output grid with a particular resolution, but the QGIS form does not offer this option. We can achieve it, however, by editing the command in the bottom window of the *Warp* form. To do this, click on the *Edit* button (the yellow pencil symbol) and include the string *-tr 90 90*. This is the parameter for 90 metre square cells. If you were using the 30 metre SRTM data you would use the string *-tr 30 30*. You can put this parameter string before the *-of* parameter, for example. Don't put it at the end as the final parameter is the output file name. See Figure 4. Do not click the *Edit* button or it will remove the extra parameter; just click *OK* to perform the reprojection. You should see a message saying that processing is complete (click *OK*), and a second one showing the output is created, finishing with 0...10...20...30...40...50...60...70...80...90...100 – done. Click *OK* to close this.

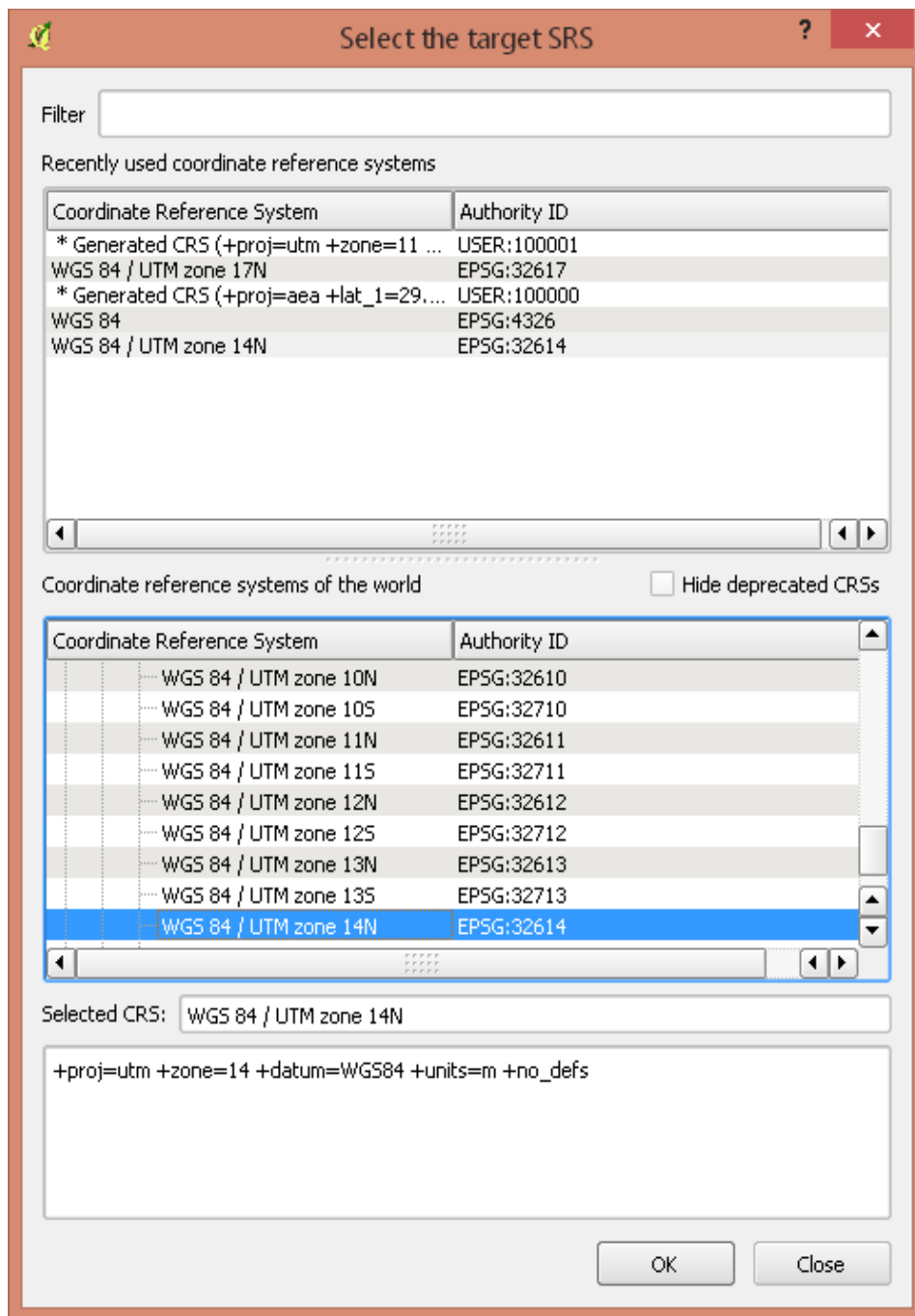


Figure 3: Select the target SRS form

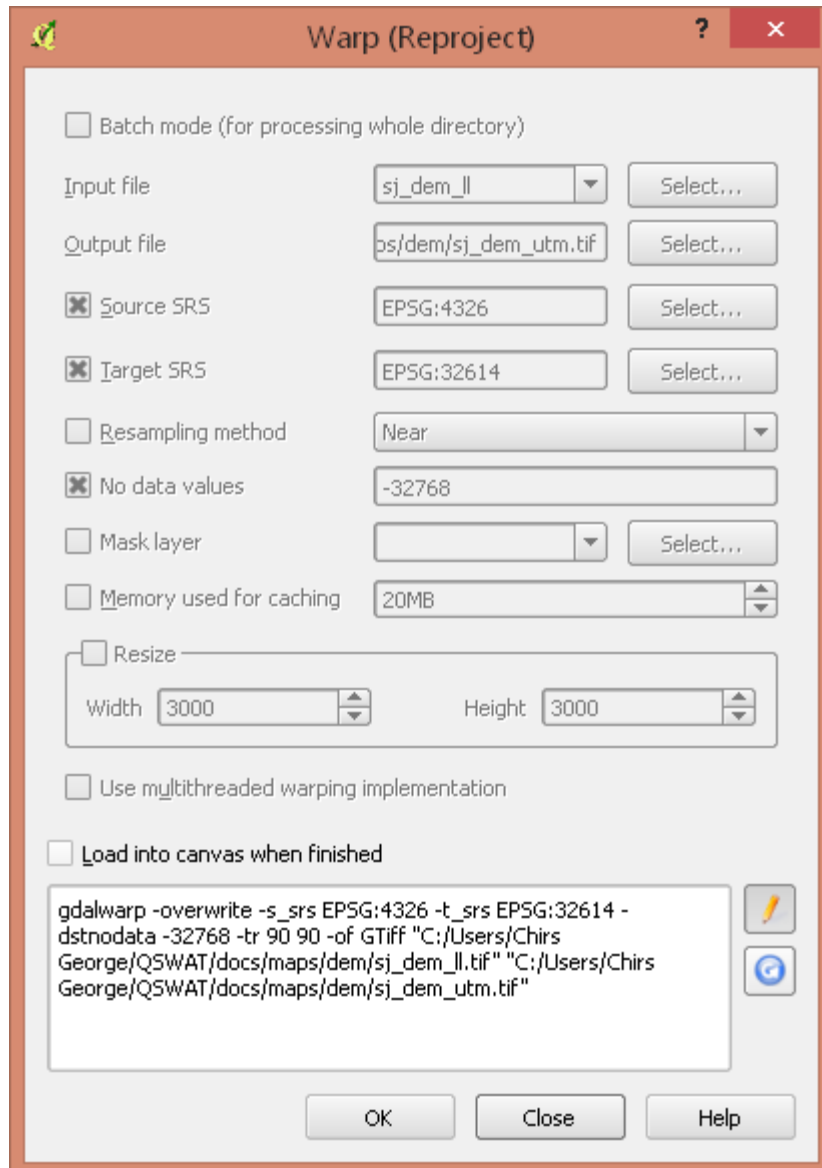


Figure 4: Warp (Reproject) form

Start a new QGIS instance (leaving the old one in case you find you made an error and need to redo something). Load dem_utm file and check it looks OK. Note that coordinates in the box at the bottom are now shown as metres instead of degrees (you won't recognise the numbers but you will see the difference from the previous instance of QGIS)². If all looks well close the previous instance of QGIS. Choose *Discard* when asked if you want to save the project.

6 Masking a DEM

QSWAT does not itself provide a facility for masking a DEM, because the capability is already in QGIS. Before showing how to mask, a note about when to do it. Masking, or reducing the area of the

² If you see degrees instead of metres as the coordinates you should open *Project* → *Project Properties* and select the *CRS* tab. If *Enable 'on the fly' CRS transformation* is checked you are strongly advised to uncheck it. If it is checked then you will not be informed about loading layers with different projections, which can lead to errors in QSWAT.

DEM to be delineated, is done for two reasons:

1. To reduce the amount of time that delineation takes.
2. To improve the accuracy of delineation.

The first reason is much less true now than it was some years ago because of the improved speed of computers. Unless your DEM is very large and detailed it will typically take a few minutes at most to delineate, and as you don't often repeat the process (and QSWAT tries to reduce the amount of recalculation when it is repeated, as well as supporting multi-core processing) reducing the area more than the rough clipping described in this document is not likely to be worth doing. See section 7 for more information about large DEMs.

Masking is also dangerous. If after delineation you find that part of your watershed boundary is on the edge of your DEM then it is quite possible that you have excluded some of the watershed, and it will be a good idea to reclip the DEM more generously and run again. So the general advice on masking your DEM is don't do it!

The second reason is much more compelling. DEMs are far from perfect representations of the real terrain, and can easily include in the watershed areas which are known to be outside it. In this case, if you have a reliable map of the actual watershed boundary, you can use it to clip the DEM precisely.

So we assume you have a good map of the watershed boundary, which we assume is a shapefile with a single shape in it. We further assume it is in the projection of your final DEM: if it is in some other projection right click on the entry in the *Layers* panel, select *Save as...*, and in the next form change the CRS from *Layer CRS* to *Selected CRS*, use *Browse* to find the projection, enter a new file name and click *OK* (Figure 5).

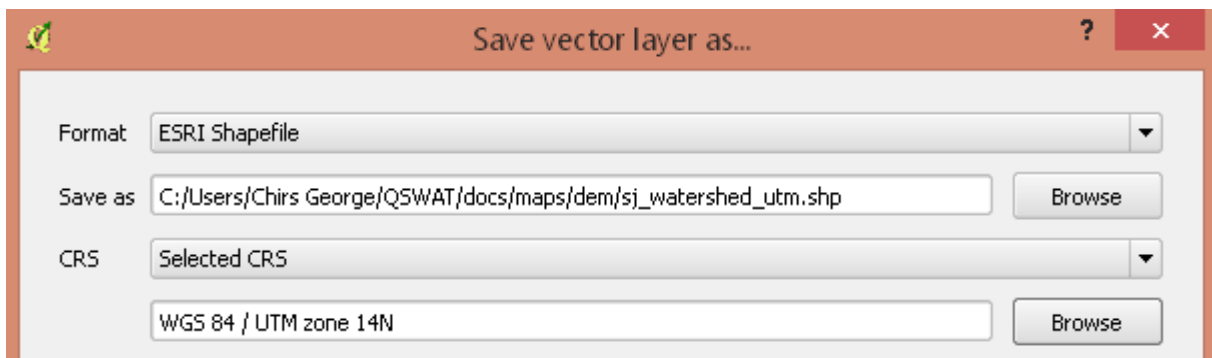


Figure 5: Reprojecting a shapefile

We can now use this boundary shapefile to clip our DEM. Start QGIS, load the projected DEM, and then load the projected watershed shapefile (Figure 6) Check the shape lies within the DEM, or we made a mistake somewhere!

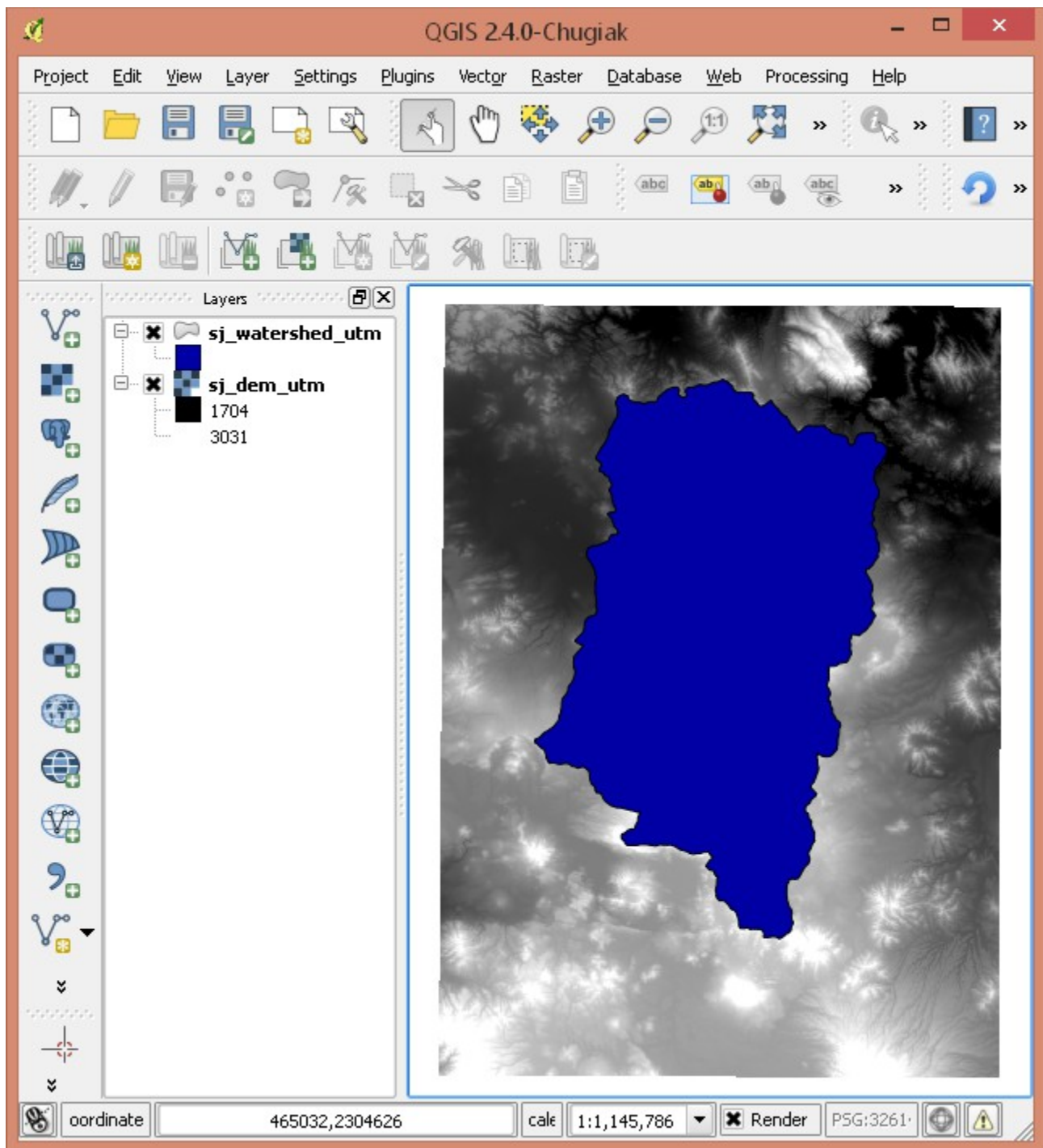


Figure 6: Ready to clip DEM with watershed boundary shape

Then use *Raster* → *Extraction* → *Clipper* to open the *Clipper* form (Figure 7). If not already selected, select the dem_utm from the pull-down menu as the input, Click the output *Select* button, navigate to the map folder, and use perhaps sj_dem_clip_utm as the file name. Set the no data value to that of the DEM, -32768 for the 90m, -32767 for the 30m. Set the *Clipping mode* to *Mask layer* and make sure the boundary shapefile is selected. Click *OK* to reproject, and again to clear the completion message.

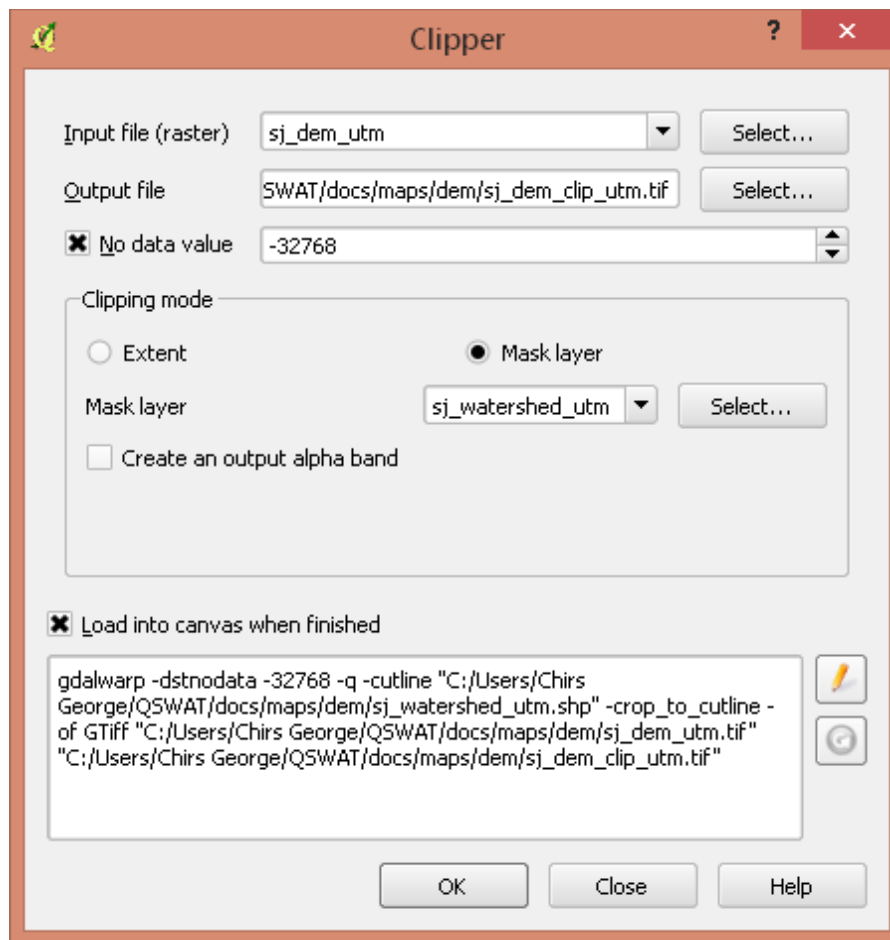


Figure 7: Clipping the DEM by mask

Close the clipper form and remove the dem_utm to check that we have a DEM clipped to the shape.

7 Large DEMs

Some processing times for large DEMs are shown in Table 1. These are not intended to be anything more than indicative: while the number of cells in a DEM is probably the main factor influencing processing time, it is not the only one. DEM sizes are in millions of cells: for example the 100 million DEM has 10000 rows and 10000 columns. You can see the size of your DEM by loading it in QGIS, right clicking the entry in the *Layers* panel and selecting *Properties*. The *Layer info* section under the *General* tab shows the number of columns and rows.

The times in Table 1 are for running from after loading the DEM to the end of stream delineation: this includes the two most long-running TauDEM functions (D8FlowDir and DinffFlowDir) Note that these two functions only need to be rerun if the DEM changes. If you want to change the delineation threshold, or the inlets/outlets, the rerun will be much faster than these times. Where there is a second time in brackets, this is for running without using the Message Passing Interface (MPI), i.e. with a single process. The machines used to obtain these figures are:

- A. A 32-bit laptop with an Intel T9400 2.53GHz processor with 2 cores and 2 GB of memory. MPI

was used with 4 processes.

- B. A 64-bit PC with an Intel i3 3.3GHz processor with 2 cores, 4 logical processors and 6GB of memory. MPI was used with 8 processes.
- C. A 64-bit PC with an Intel i7 3.4GHz processor with 4 cores, 8 logical processors and 8GB of memory. MPI was used with 16 processes.

Machine/ DEM size	A	B	C
100	2100(*)	255(560)	104(450)
25	*(480)	64(140)	28(111)
12	159(205)	28(61)	12(48)
6	42(57)	8 (16)	4 (13)
4	13(22)	3(5)	1(3)
2	7(9)	1(2)	1(2)
1	3(4)	1(1)	1(1)

Table 1: Times in minutes for processing large DEMs (sizes in millions of cells)

* - run failed

There are two conclusions to be drawn from these figures:

- In this range of DEM sizes the times increase initially linearly with the size of the DEM, but later much more sharply. A DEM of 369 million cells processed on machine C took over 15 hours, much worse than the rough '1 minute per million cells' that you would expect for this machine from the table. Typically 'Not Responding' will appear at the top of the main QGIS window and/or the delineation form. This is a result of the TauDEM processes running in the background but using almost all the CPU resources. Ignore the lack of responsiveness and be patient!
- MPI only makes a substantial difference when the DEMs are above a few million cells.