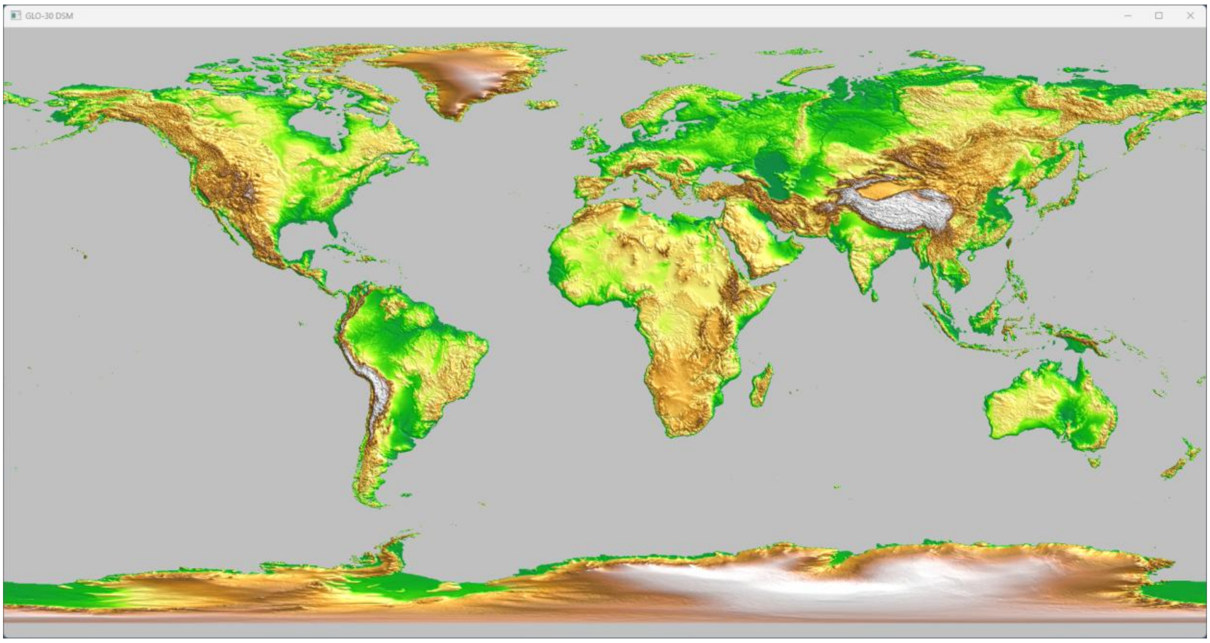
GLO-30 in MRR Format: A single, global, high-resolution Terrain DSM

Sam Roberts, Roberts Geospatial Engineering, 28 November 2024



GLO-30 DSM MRR rendered in ProRaster Scientific

Introduction

GLO-30 is formally known as the "Copernicus DEM Global Elevation Model GLO-30" dataset. It can be acquired from the European Space Agency Copernicus portal as 1x1 degree tiles. The tiles are arranged by country and then by region. Many tiles are duplicated as they overlap more than one country. You can find a description of the data, the terms of use, and a downloadable product handbook at the link below.

https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model

You can download the tiles from the Copernicus Data Service Portal (CDSPortal). The following YouTube video provides a useful guide to downloading these tiles using FTP.

https://www.youtube.com/watch?v=4piMj4DbCow

GLO-30 DEM is a Digital Surface Model (DSM) that represents the surface of the Earth including buildings, infrastructure and vegetation. GLO-30 offers global coverage of the Earth at a resolution of 30 metres with total surface coverage of approximately 149 million square kilometres. It can be acquired for free and used in accordance with the terms and conditions (see the web page listed above). The raster tiles are in geographic coordinates (WGS84), EPSG 4326. Grid spacing in the latitude direction is 1.0 arcsecond (about 31m at the equator and 20m at 50 degrees latitude). Grid spacing in longitude direction is variable and based on latitude. For more information, please consult the product handbook. For latitudes 0 – 50 degrees, grid spacing is 1arcsecond. This steps up to 1.5, 2.0, 3.0, 5.0 and then 10.0 arcseconds as you approach the poles. The terrain data is stored in decimal metres.

The data is derived from the WorldDEM terrain raster product and is predominantly data acquired by the TanDEM-X satellite mission between 2011 and 2015. This data is infilled on a local basis with the following DEMs: ASTER, SRTM90, SRTM30, SRTM30plus, GMTED2010, TerraSAR-X Radargrammetric DEM, ALOS World 3D-30m, Norway National DEM and Spanish National DEM.

At time of writing, WorldDEM is available for purchase through Airbus and grid spacing resolution has improved from the original 12 metre resolution to 5 metre resolution. It is now known as WorldDEM Neo and is derived from TanDEM-X between 2017 and 2021. It is available as a DSM and as an improved DTM with vegetation and man-made objects removed.

Compared to WorldDEM Neo, GLO-30 is lower resolution, older, and poorer quality. However, at the present time GLO-30 is the best, free, and highest resolution global DSM available. Statistically and visually, GLO-30 is clearly higher quality than other comparable datasets like SRTM, ALOS and ASTER. However, it still has some quality issues (which I will explore) and the fact that it is a DSM incorporating vegetation and man-made infrastructure (and not a clean DTM) remains a blocker for some applications.

Acquisition

I downloaded release number 2023-1 between September - October 2024 using the FTP site for bulk downloads. I acquired the 32-bit GeoTIFF DGED tiles which include the quality masks. The entire data set is split into thousands of tiles, each of which are downloaded as a compressed archive. It is designed to downloaded by country and is quite onerous and inefficient to download in bulk. The total download volume ran to multiple terabytes.

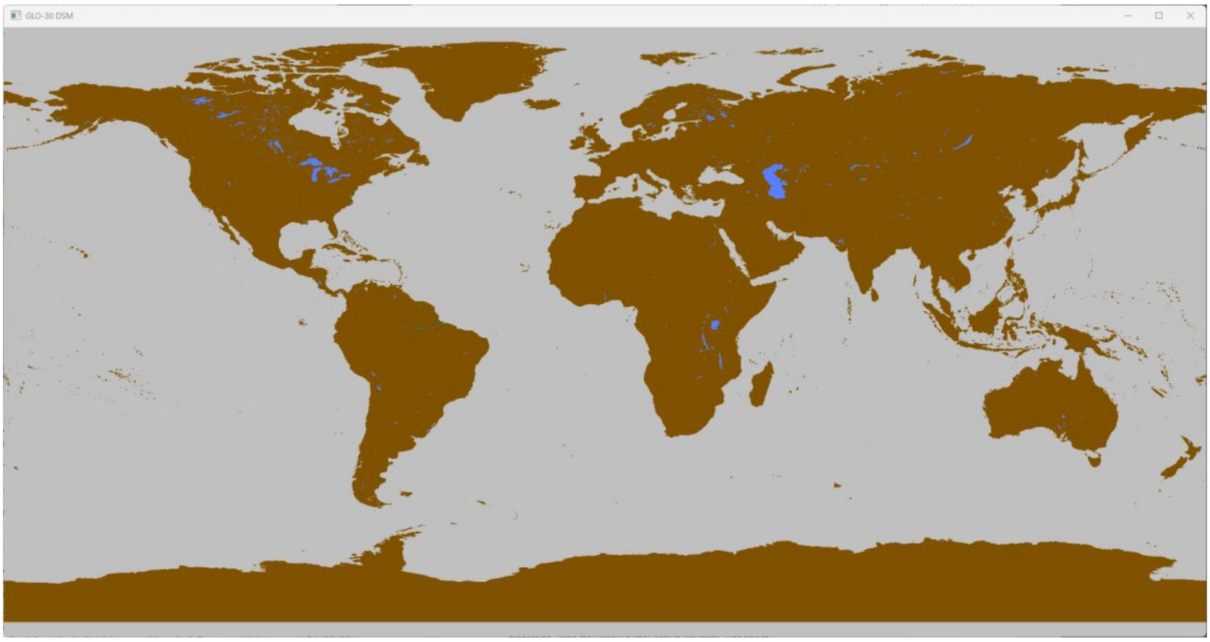
The supplied TIFF files are standard TIFF format. They are not in COG format and no OVR files are supplied. In addition to the DSM TIFF file, several quality layers are supplied as classified rasters in TIFF format. Consult the handbook for details.

- The Editing Mask (EDM) records the last editing process applied to each pixel, during the terrain and hydro editing process. For example, it may record Infill, Interpolated pixels, Smoothed pixels, Flattened pixels etc.
- The Filling Mask (FLM) records the data source for each pixel. For example, it may be sourced from ASTER, SRTM90 or SRTM30 etc.
- The Water Body Mask (WBM) classifies all pixels as Land, Ocean, Lake or River.
- The Height Error Mask (HEM) records an estimated height error measure for each pixel.

Each tile is in a TAR archive. To process the data, I decompressed each tile TAR. To reduce storage requirements, I deleted all the ancillary files and folders. I also deleted the HEM rasters as I made a judgement that the HEM data was not sufficiently valuable to warrant retention or processing.

Processing Masks

I produced four rasters in MRR format - three classified rasters for the WBM, FLM, and EDM masks, and one DSM terrain raster. In total there were 29314 input TIFF files for each raster, although some are duplicates and could be excluded. There were 26474 unique input files per raster.



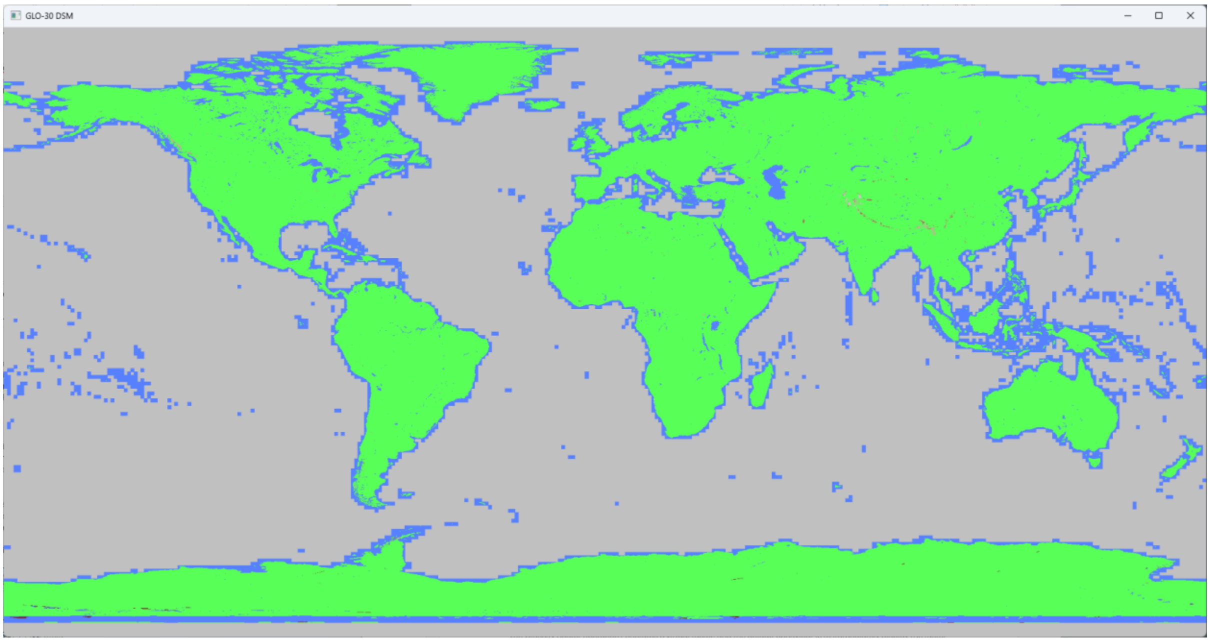
GLO-30 WBM Water Body Mask MRR rendered in ProRaster Scientific

I created the three classified mask rasters using ProRaster Scientific running on a 6-core medium specification laptop. In each case the processing was a three-step process.

1. Create a raster source that links to all the input TIFF files.
2. Execute a Classify Raster operation (the output is a virtual raster).
3. Execute an Export Raster operation to export the classified virtual raster to a crystallised raster in MRR format.

In ProRaster, a “raster source” is used to associate driver preferences with a raster that modify the treatment of the raster by the raster engine when rendering or processing. It is also used to group multiple rasters with matching properties together so that users can deal with a single “source” object for many rasters when rendering or processing. Using the Raster Source Editor in ProRaster, I was able to –

- Quickly and easily select almost 30000 rasters from a complicated folder structure using a filename wildcard like “*WBM.tif”.
- Prevent processing of duplicate rasters (matching by file name).
- Prevent generation of overviews by opening in base support mode only.
- Ensure the TIFF files are loaded as a continuous field rasters.



GLO-30 FLM Fill Mask MRR rendered in ProRaster Scientific

The Classify Raster operation requires the specification of a Classification table and a data mapping from the original TIFF data to the final class index. The classification table information was taken from the handbook, although colors were not specified in the handbook. The content of the CSV files for WBM, FLM, and EDM rasters is reproduced below. For information on how to format these CSV files, see the ProRaster User Guide. Note that in the WBM raster I discarded the Ocean pixels as these contain no useful information.

```
Mask,OriginalClass,ClassColour,ClassLabel
Data,Class,Colour,Label
UNSIGNED_INT8,UNSIGNED_INT8,RGBA,STRING
1,0,"RGBA(128,82,0,255)",Land
1,2,"RGBA(87,128,255,255)",Lake
1,3,"RGBA(0,105,128,255)",River
```

```
OriginalClass,ClassLabel,ClassColour
Class,Label,Colour
UNSIGNED_INT8,STRING,RGBA
0,"Void (no data)","RGBA(120,120,120,255)"
1,"Edited (except filled pixels)","RGBA(87,128,255,255)"
2,"Not edited / not filled","RGBA(87,255,87,255)"
3,"ASTER","RGBA(169,129,94,255)"
4,"SRTM90","RGBA(152,77,80,255)"
```

```
5,"SRTM30","RGBA(190,183,148,255)"
6,"GMTED2010","RGBA(119,40,45,255)"
7,"SRTM30plus","RGBA(162,106,89,255)"
8,"TerraSAR-X Radargrammetric DEM","RGBA(175,153,101,255)"
9,"AW3D30","RGBA(204,202,195,255)"
100,"Norway DEM","RGBA(159,93,86,255)"
101,"DSM05 Spain","RGBA(138,58,65,255)"
102,"Norway DEM v2","RGBA(182,170,119,255)"
```

```
OriginalClass,ClassLabel,ClassColour
Class,Label,Colour
UNSIGNED_INT8,STRING,RGBA
0,"Void (no data)","RGBA(120,120,120,255)"
1,"Not edited","RGBA(87,255,87,255)"
2,"Infill of external elevation data","RGBA(169,129,94,255)"
3,"Interpolated pixels","RGBA(152,77,80,255)"
4,"Smoothed pixels","RGBA(190,183,148,255)"
5,"Airport editing","RGBA(119,40,45,255)"
6,"Raised negative elevation pixels","RGBA(162,106,89,255)"
7,"Flattened pixels","RGBA(175,153,101,255)"
8,"Ocean pixels","RGBA(0,61,128,255)"
9,"Lake pixels","RGBA(87,128,255,255)"
10,"River pixels","RGBA(0,105,128,255)"
11,"Shoreline pixels","RGBA(255,220,114,255)"
12,"Morphed pixels (series of pixels manually set)","RGBA(204,202,195,255)"
13,"Shifted pixels","RGBA(159,93,86,255)"
```

The data transform CSV files for WBM, FLM, and EDM are reproduced below. For information on how to format these CSV files, see the ProRaster User Guide.

```
NUMERIC
0,0
2,1
3,2
```

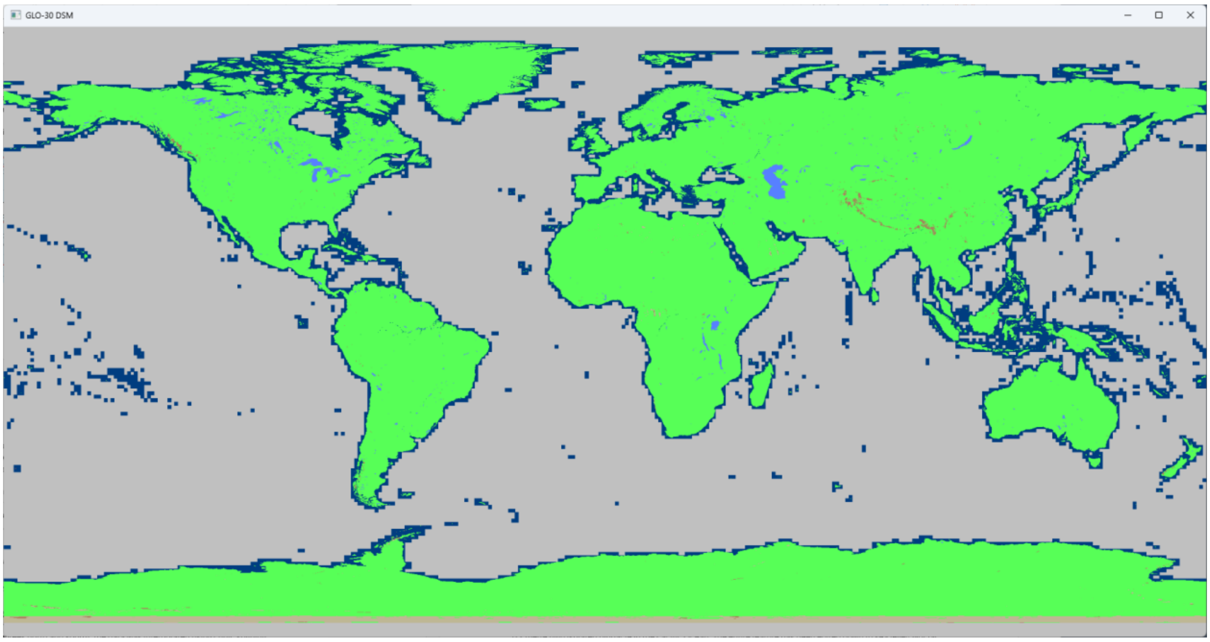
```
NUMERIC
0,0
1,1
2,2
3,3
4,4
5,5
6,6
7,7
8,8
9,9
100,10
101,11
102,12
```

```
NUMERIC
0,0
1,1
2,2
3,3
4,4
5,5
6,6
7,7
8,8
9,9
10,10
11,11
12,12
13,13
```

The Classify Raster operation generates a virtual raster, and the default behaviour is to immediately display the raster after it is created. However, as these rasters have no overviews and the raster source was explicitly used to prevent creation of overviews, the virtual raster is not suitable for rendering. It is only used as a link in the processing chain. In ProRaster, you can elect to prevent rendering of the output of an operation when you execute the operation. I elected to use an 8-bit unsigned integer as the class index in the classified rasters (for improved processing performance).

The Export Raster operation converts a raster of any kind into an MRR format raster. In particular, the operation is designed to convert MVR virtual rasters (produced by processing operations in ProRaster) into MRR or TIFF format rasters. It provides advanced settings for MRR format rasters that allow you to customise the storage data type and data compression, the overview levels, and other options. For large high-resolution classified rasters, I use a large tile size, run-length encoding, and highest quality LZMA lossless compression. I used the “Any” overview coverage rule to ensure that any populated cell in the base level will generate a populated cell coincident with it right up the overview pyramid. When visualising the data, this makes it easy to find small islands isolated data when zooming in from broad coverage to local scale.

The final raster was 1x1 arcsecond resolution. Where the source data was at a lower resolution (above 50 degrees north and south), the data was interpolated using nearest neighbour.



GLO-30 EDM Edit Mask MRR rendered in ProRaster Scientific

Processing Terrain DSM

The terrain raster was generated using the WBM mask to mask out all ocean pixels. Only pixels classified as land, lake, or river were retained. The final raster was 1x1 arc second resolution. Where the source data was at a lower resolution (above 50 degrees north and south), the data was interpolated using cubic splining.

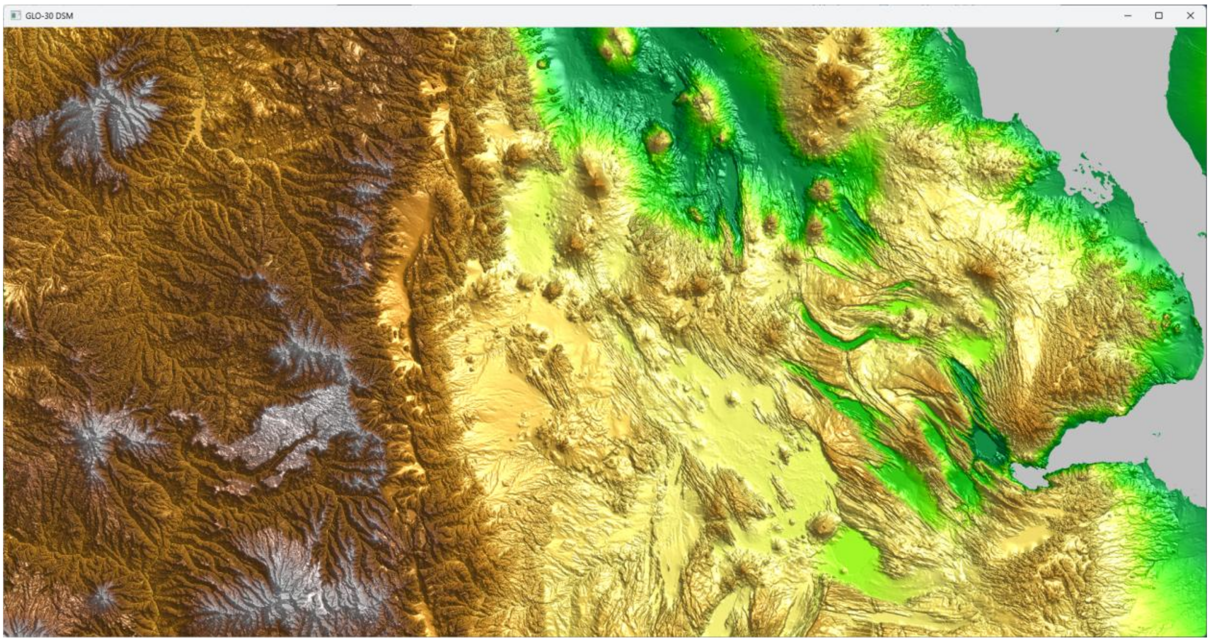
Using the DSM TIFF files and the WBM classified MRR raster (see above), you can create the raster in ProRaster Scientific using a three-step process.

1. Create a raster source that links to all the source TIFFs.
2. Execute a Mask Raster operation to remove ocean pixels (the output is a virtual raster).
3. Execute an Export Raster operation to export the masked virtual raster to a crystallised raster in MRR format.

Whilst this works well for subsets of the dataset (e.g. for a country), I found that this method was too slow to process the entire raster. Instead, I wrote a custom Merge processing operation that applied the mask and exported the raster in one step, working through the source rasters sequentially and in a spatially optimised pattern. On a 4-core older NUC PC this took about 30 hours to complete.

To better reflect the vertical data quality, and to maximise data compression, I rounded the elevations to the nearest 0.2 metres vertically. The data is stored as an unsigned 16-bit integer to which a transform (offset and scale) is applied to convert to floating point. Predictive encoding was performed prior to high-quality LZMA lossless compression.

The overview pyramid for the DSM was generated in the usual way where cells are only propagated upwards where half or more of the area of the overview level cell is populated.



Ethiopia: GLO-30 DSM MRR rendered in ProRaster Scientific

Outcomes

Starting from 105,896 unique TIFF files (excluding overviews which were not supplied), generated four MRR format rasters. The Water Body Mask (WBM) classified raster is 466 MB on disk. The Filling Mask (FLM) classified raster is 2.14 GB on disk. The Editing Mask (EDM) classified raster is 2.29 GB on disk. The Terrain (DSM) continuous raster is 173 GB on disk.

Each raster is 1x1 arcsecond resolution with global coverage from -180 to 180 degrees longitude and -90 to 90 degrees latitude. Each raster is 1,296,001 x 648,001 cells in extent. The DSM contains approximately 382 billion valid grid cells and so it requires approximately 3.9 bits of storage per cell.

By converting to MRR I achieved the following outcomes.

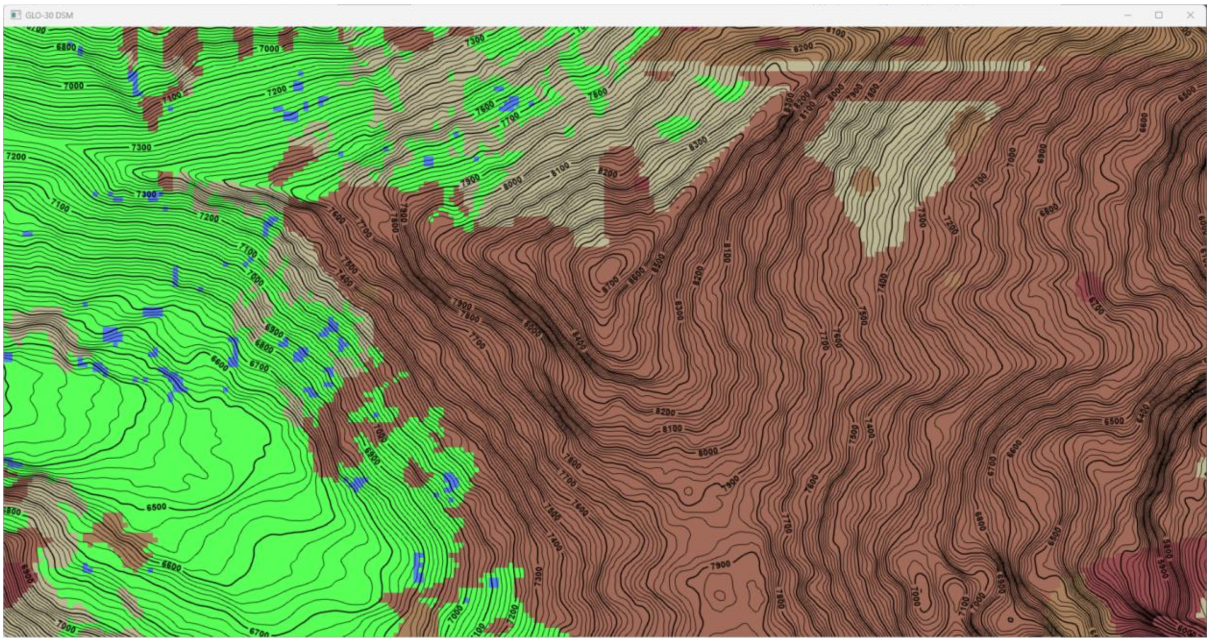
- Ease of access. For each raster, 26,474 TIFF files were condensed to one file.
- Ease of storage. The disk space requirement has been reduced by an order of magnitude.
- Quality rendering. Overviews are stored in the MRR.
- Analysis. Statistics are stored in the MRR.
- Efficiency. Ocean cells have been discarded (partial coverage, no value).
- Correctness. Unsupportable vertical decimal resolution has been discarded.
- Joy! The whole Earth accessible from a single raster for visualisation and processing.

Issues

The DSM data has some quality issues, but I would preface these comments with a reiteration that this the highest quality and resolution, free terrain raster with global coverage. The satellites that gathered the data were not free, so let's not complain too loudly about quality issues with a free gift.

Firstly, this dataset is predominantly but not completely sourced from the TanDEM-X mission data. You can spatially browse the FLM fill mask raster to see where data has been sourced from different datasets. Often, the quality of the data is similar, but usually it isn't. At the boundaries between regions sourced from different datasets there may be visible transitions.

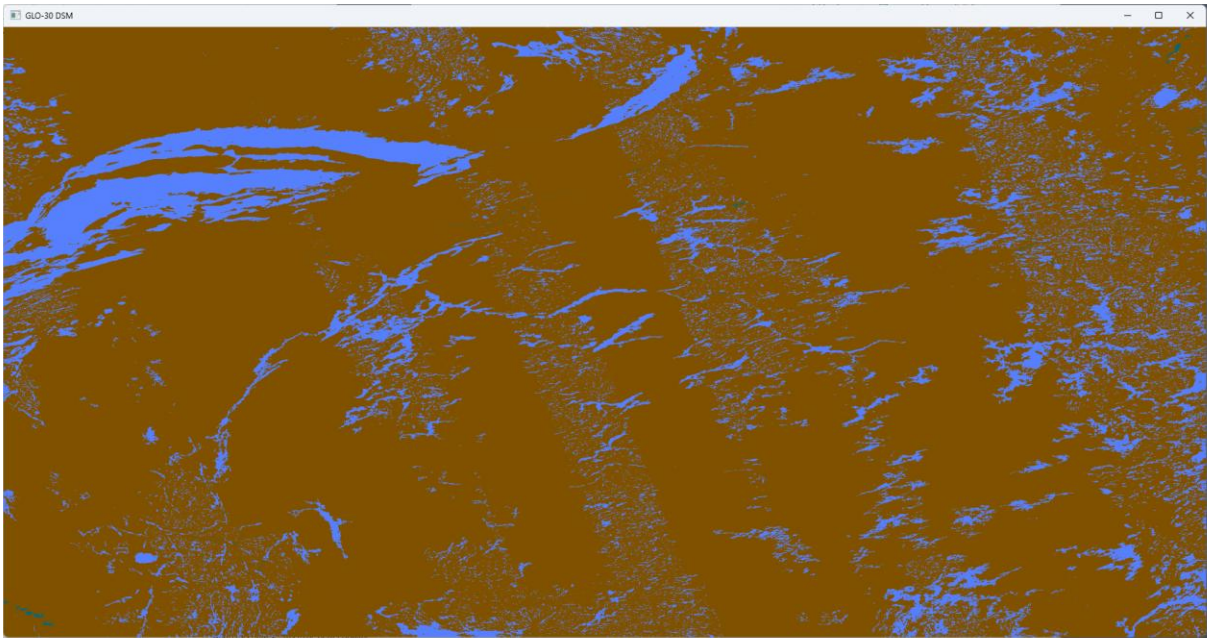
I have found some examples where terrain heights are substantially incorrect. One example is Mt Everest which, at 8737 metres, is 110 metres lower than its known height. In mountainous terrain, and particularly in the Himalaya, fill from other datasets is very common. Mount Everest is SRTM30+ data. A less important example is Balls Pyramid – a 572 metre high isolated pinnacle in the Pacific Ocean. The entire feature has been edited down to sea level! Not far away is Lord Howe Island – another rugged Pacific Island. In this case the entire Southwest corner of the island has been edited out of existence.



Mt Everest: DSM Contours 20m/100m and FLM Fill, rendered in ProRaster Scientific. Green pixels are TanDEM-X derived, dark brown is SRTM30+.

On the other hand, the DSM can reveal errors in other maps. For example, Google Maps places Montagu Island (the largest of the South Sandwich Islands, located in the Scotia Sea off the coast of Antarctica) and several islands near it in the wrong location. Microsoft and OpenStreetMaps have the islands in the right place, coincident with the GLO-30 DSM.

Finally, the WBM Water Body Mask raster seems to be compromised in Canada. The thousands of lakes in northern Canada are only partially represented. Large swathes, following satellite paths, are missing. Elsewhere around the world this data seems to be better quality.



Canada: WBM Water Mask, rendered in ProRaster Scientific. NW-SE striping is evident.