

Converting US Topo GeoPDF[®] Layers to GeoTIFF

Larry Moore
U.S. Geological Survey
May, 2016
Document Version 2.5

Based on Global Mapper v17.1 and GDAL v1.11.3

Readers are cautioned that this reformat is complex and tedious, with significant learning curves for most people. This paper is offered in the spirit of sharing specialized technical information with a niche audience. These procedures are at best interim solutions to the “PDF in GIS” problem.

This paper may be freely shared, but please use references to <http://www.usgs.gov/faq/categories/9797/3704> instead of rehosting copies. Outdated versions of this document that remain online tend to float to the top of search engine results.

Table of Contents

1	Introduction	2
1.1	Background	2
1.2	Scope.....	2
2	Global Mapper Commercial Software	3
3	GDAL Open Source Software	5
3.1	Environments.....	5
3.2	Download software.....	5
3.3	Example data.....	5
3.4	Basic reformat.....	6
3.4.1	Use gdalinfo to get the names of layers in the PDF.....	6
3.4.2	Use gdal_translate to convert selected layers to GeoTIFFs.....	7
3.5	Reducing output file size.....	8
4	More Information and Important URLs	9
	Appendix A. Global Mapper script for batch GeoPDF-to-GeoTIFF reformat.....	10
	Appendix B. Perl Script for GDAL GeoPDF-to-GeoTIFF reformat.....	12
	Appendix C. List of US Topo layers.....	13
	Appendix D. GeoTIFF file sizes with various GDAL options.....	15
	Appendix E. Finding USGS GIS data.....	16

1 Introduction

1.1 Background

In 2009 the U.S. Geological Survey (USGS) began publishing digital quadrangle maps modeled on the legacy 7.5-minute topographic series of 1947-1992. The new maps, branded US Topo in 2010, are published in GeoPDF[®] format – Portable Document Format (PDF) with geospatial extensions. PDF was chosen as the physical format because it is the only format in common use that

- Can carry vector, raster, and text data.
- Can display complex structures on a typical office computer without specialized software or expertise.
- Allows a map to be printed at correct scale without specialized software or expertise.

PDF geospatial extensions were implemented in US Topo because they add value “for free” – no cost to end users, and do not conflict with base PDF¹.

US Topo was intended to be a map product, not a GIS product. US Topo maps are fundamentally an output of, not an input to, GIS. They are derived directly from national GIS databases and represent a repackaging of existing data, not creation of new data. Nevertheless, there is demand for the ability to load these symbolized maps into GIS software. The basic benefit of this capability is to supply the GIS user with a pre-built, symbolized background map.

There are several potential solutions to this GIS user problem. The USGS is working on ways to distribute GIS vector data and symbol libraries corresponding to US Topo². It is also likely that the major GIS vendors will eventually implement geospatial PDF import functions. Web services (Appendix E) also provide partial solutions. Converting US Topo PDFs to some other format is also an option, and is the subject of this paper.

A common question from GIS users is “how can I convert a US Topo to GeoTIFF?” GeoTIFF is a raster image, not an intelligent GIS format, but provides a simple mechanism to display map symbols and annotation. GeoPDF-to-GeoTIFF conversion would seem like a relatively easy way to import US Topos into GIS. Unfortunately, this reformat is not especially simple, mostly due to the lack of powerful software for manipulating layered PDF files and for preserving georeferencing through reformat processes.

1.2 Scope

This paper presents software and procedures for separating the layers of a US Topo and converting the separated layers or combinations of layers to GeoTIFFs. Two methods are discussed:

1. Global Mapper commercial software. This is the easier of the two methods by a wide margin, but it depends completely on proprietary software.
2. Utilities of the open source Geospatial Data Abstraction Library (GDAL). This process is more powerful but more complex. It requires familiarity with Unix-like software setup and command-line processing.

Neither process has been fully automated, though both can be partially automated with scripts. Both produce results of high visual quality.

¹ These extensions are currently added with TerraGo software, so the resulting files can be referred to with the trademarked term GeoPDF[®]. “Geospatial PDF” is a generic term for georeferenced PDF documents.

² For more information (as of 5/16), and sample products, about a new product of GIS vectors with symbol templates packaged in 7.5-minute cells, see <http://viewer.nationalmap.gov/tools/topotemplate/>

Software to manipulate geospatial PDFs is evolving relatively rapidly, and this document may have a short useful life. Check for later versions of the software described here and read the documentation distributed with the software.

Earlier versions of this paper contained a section on the USGS [Historical Topographic Map Collection](#) (HTMC). In 2015 the USGS [National Geologic Map Database](#) project created GeoTIFF versions of the HTMC maps. These are distributed through the TopoView application at <http://ngmdb.usgs.gov/maps/TopoView/>. Several other formats are also served through TopoView. The GeoTIFF versions of HTMC maps are not yet available for bulk delivery or through other USGS interfaces. USGS does not intend to create standard GeoTIFF versions of US Topo, in part because the US Topo maps contain too many data layers to include in a single, simple image file.

2 Global Mapper Commercial Software

See <http://www.blumarblegeo.com/products/global-mapper.php> for technical and purchasing information about Global Mapper.³

This description is based on Global Mapper v17.1. Global Mapper documentation indicates that similar procedures should work for all versions beginning with v11.

Global Mapper can load layered geospatial PDFs and preserve the georeferencing. However, Global Mapper PDF display behavior is different than Adobe Acrobat behavior. Global Mapper rasterizes the entire dataset, flattening all layers to a single raster image plane. Therefore, the Global Mapper display of a geospatial PDF does not look like an Acrobat display of vectors, but rather like a file that has already been converted to a raster format such as TIFF or JPEG.

As part of the file load process, Global Mapper gives the user the option to select layers (Figure 1). This layer selection must be done when the file is loaded; once the map is displayed in Global Mapper, it has been converted to one raster image and the layer characteristics of the original PDF are discarded.

When Global Mapper rasterizes a US Topo, it sets the resolution at about two ground meters per pixel, or about 300 lines per inch at 1:24,000 scale. If the data are exported to a raster file, this resolution is maintained by default, though it can be changed. Decreasing the resolution will reduce the file size and lower the image quality. Increasing the resolution will increase the file size, but will not noticeably improve image quality.

The displayed image can be exported to a GeoTIFF (or other raster formats) using the standard Global Mapper export functions.

Global Mapper includes a powerful scripting language, and this process can be scripted to convert any number of maps at one time. A listing of such a script is included in Appendix A of this document. The script and a map cell Shapefile (needed for clipping collars; see the inline documentation of the script listing) are also included in a tar file at ftp://ftpext.usgs.gov/pub/cr/co/lakewood/Larry_Moore/ustopo2tiff/

³ USGS has no business relationship with Global Mapper or Blue Marble Geographic, Inc., and does not endorse or sponsor this commercial product. Global Mapper procedures are described here because we believe it provides unique capabilities for separating US Topo layers and converting them to GeoTIFF.

As explained in the inline documentation, changing the behavior of this script (to alter the layers selected for export, for example), requires changes to the code. More complex versions of this script that would present run-time options to the user could be written.

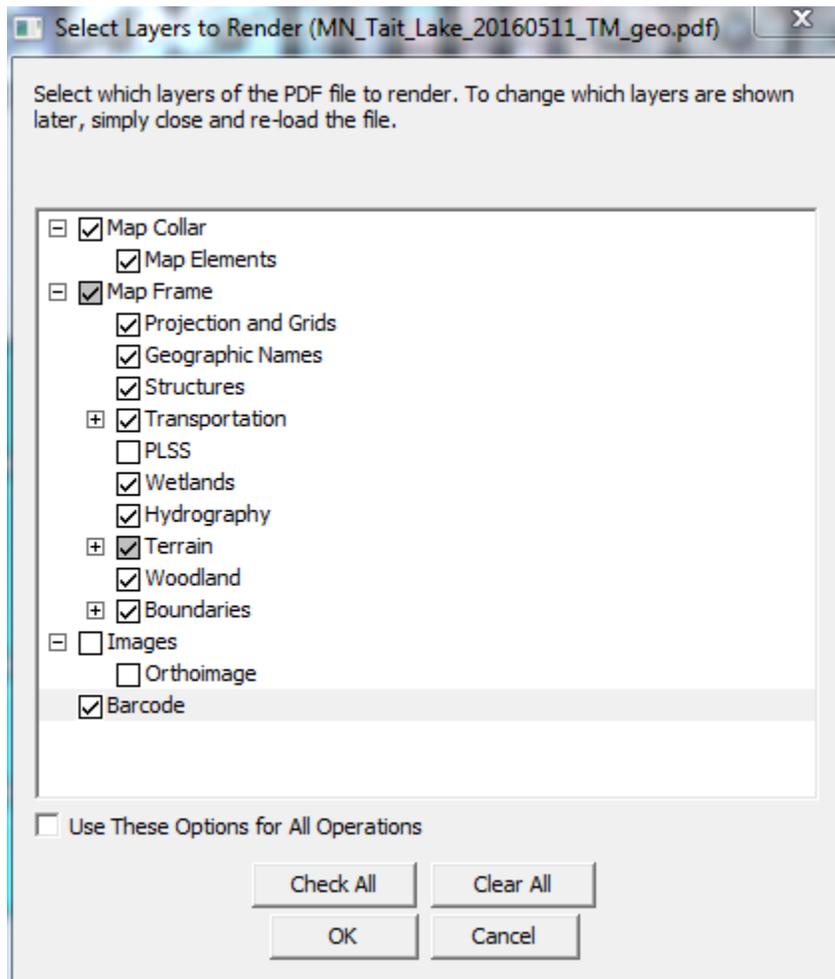


Figure 1. Screen shot of Global Mapper layer selection box, with all folders expanded and “on by default” layers checked. Unlike previous Global Mapper versions, the layer list is not completely expanded.

Layers may be clicked on or off when a US Topo geospatial PDF is opened. Global Mapper recognizes the GeoPDF file instructions regarding which layers should be “on” when the file is opened. Layer selection must be done as part of the file open process; layers cannot be selected later. This example is from a US Topo map published in 2015. The exact layer list will vary with the vintage of the US Topo; see Appendix C for more information.

The bar code layer was added to US Topo in early 2016. Unlike other new layers, bar codes were added retroactively to all 2014 and 2015 maps. Global Mapper appears unable to exclude this layer; it is not known at this writing if this is a problem with the data or is a bug in Global Mapper.

3 GDAL Open Source Software

The Geospatial Data Abstraction Library (GDAL) has several advantages over Global Mapper: the software is no-cost and open source, gives the user more options and control, and allows more advanced processing. On the other hand, using GDAL utilities is more difficult. Software setup is relatively complicated. Familiarity with command-line processing is required. Commands are long and tedious. Familiarity with scripting techniques is very helpful.

This paper explores only a small subset of GDAL capabilities.

3.1 Environments

The GDAL utilities discussed below can be used in several different environments:

1. **Simplest, but least powerful:** The critical utilities can be downloaded as pre-compiled Windows binaries that can be run from the Command Prompt application that is included with Windows.
2. **More powerful, more difficult:** A Unix emulation environment within Windows provides more power and flexibility, especially for scripting. It also allows software to be built from source (though precompiled Windows binaries also work). Examples of such environments are MinGW/MSYS and Cygwin.
3. **Even more powerful:** use a Unix or Linux computer. This is not as difficult as it might sound. Modern Linux distributions are easy to install, and can be run on older computers that are too small to run current Windows versions. Software installation and management for the procedures described below are easier on Linux than on Windows. The OSGeo-Live distribution is a particularly convenient way to access GDAL and other open source GIS software on Linux.
4. **Apple computers:** The GDAL download pages build for Mac OS X. USGS has not tested these programs or procedures on Macs, but a GIS user who independently investigated these reformat reports that the GDAL utilities work well on a MacBook Pro with OS X.

3.2 Download software

The Geospatial Data Abstraction Library (GDAL) home page is <http://gdal.org>. The software download page is <http://trac.osgeo.org/gdal/wiki/DownloadingGdalBinaries>. A subordinate page, <http://www.gisinternals.com/sdk/> serves precompiled Windows binaries. To see the GDAL distribution version, click on the **Package Info** column, and scroll down to **GDAL Version Information**. This paper is based on Windows binaries of GDAL version 1.11.3 running in MinGW/MSYS.

See the GDAL documentation for installation instructions. Installation procedures depend partly on the environment you are using.

The GDAL software library and associated utilities are very general and powerful. This paper documents only a few selected GDAL capabilities.

3.3 Example data

Examples in this paper use the Grand Marais, MN, 2016 US Topo map (MN_Grand_Marais_20160511_TM_geo.pdf), which can be downloaded from

http://ims.er.usgs.gov/gda_services/download?item_id=8056633

3.4 Basic reformat

The basics of separating US Topo layers and converting them to GeoTIFF require only two GDAL programs: `gdalinfo` and `gdal_translate`.

3.4.1 Use `gdalinfo` to get the names of layers in the PDF

In an appropriate command-line environment, run the command

```
1) gdalinfo MN_Grand_Marais_20160511_TM_geo.pdf -mdd LAYERS
```

The output will contain metadata and coordinate system information, and a section listing the layers in the PDF:

```
Metadata (LAYERS):
  LAYER_00_NAME=Map_Collar
  LAYER_01_NAME=Map_Collar.Map_Elements
  LAYER_02_NAME=Map_Frame
  LAYER_03_NAME=Map_Frame.Projection_and_Grids
  LAYER_04_NAME=Map_Frame.Geographic_Names
  LAYER_05_NAME=Map_Frame.Structures
  LAYER_06_NAME=Map_Frame.Transportation
  LAYER_07_NAME=Map_Frame.Transportation.Road_Names_and_Shields
  LAYER_08_NAME=Map_Frame.Transportation.Road_Features
  LAYER_09_NAME=Map_Frame.Transportation.Trails
  LAYER_10_NAME=Map_Frame.Transportation.Railroads
  LAYER_11_NAME=Map_Frame.Transportation.Airports
  LAYER_12_NAME=Map_Frame.PLSS
  LAYER_13_NAME=Map_Frame.Wetlands
  LAYER_14_NAME=Map_Frame.Hydrography
  LAYER_15_NAME=Map_Frame.Terrain
  LAYER_16_NAME=Map_Frame.Terrain.Contours
  LAYER_17_NAME=Map_Frame.Terrain.Shaded_Relief
  LAYER_18_NAME=Map_Frame.Woodland
  LAYER_19_NAME=Map_Frame.Boundaries
  LAYER_20_NAME=Map_Frame.Boundaries.Jurisdictional_Boundaries
  LAYER_21_NAME=Map_Frame.Boundaries.Jurisdictional_Boundaries.International
  LAYER_22_NAME=Map_Frame.Boundaries.Jurisdictional_Boundaries.State_or_Territory
  LAYER_23_NAME=Map_Frame.Boundaries.Jurisdictional_Boundaries.County_or_Equivalent
  LAYER_24_NAME=Map_Frame.Boundaries.Federal_Administered_Lands
  LAYER_25_NAME=Map_Frame.Boundaries.Federal_Administered_Lands.National_Cemetery
  LAYER_26_NAME=Map_Frame.Boundaries.Federal_Administered_Lands.National_Park_Service
  LAYER_27_NAME=Map_Frame.Boundaries.Federal_Administered_Lands.Department_of_Defense
  LAYER_28_NAME=Map_Frame.Boundaries.Federal_Administered_Lands.Forest_Service
  LAYER_29_NAME=Images
  LAYER_30_NAME=Images.Orthoimage
  LAYER_31_NAME=Barcode
```

The layer names to the right of the equal sign can be used as input to `gdal_translate` to extract and rasterize layers or combinations of layers.

The number and names of the layers will vary with the US Topo vintage. A complete list of layer names for US Topo maps for the time period 2014-2016 is given in Appendix C.

3.4.2 Use `gdal_translate` to convert selected layers to GeoTIFFs⁴

To create a GeoTIFF of just the contour layer:

```
2) gdal_translate MN_Grand_Marais_20160511_TM_geo.pdf marais_contours.tif \  
   --config GDAL_PDF_LAYERS "Map_Frame.Terrain.Contours" --config GDAL_PDF_BANDS 3 \  
   --config GDAL_PDF_DPI 400
```

GeoTIFF is the default output format; this can be made explicit with the option `-of GTiff`

The backslash at the end of the first line is a Bash shell continuation character. `^^` is the comparable DOS character. The continuation character is not needed when typing these commands in a command line interface, but is included here to present correct syntax for multi-line commands.

The output also defaults to RGB color model with no compression; this can be made explicit with the options

```
-co PHOTOMETRIC=RGB -co COMPRESS=NONE
```

The option `--config GDAL_PDF_BANDS 3` removes the alpha channel and creates 3-band RGB output. Omitting this option and preserving the alpha channel may be useful for some specialized applications.

The argument to the `GDAL_PDF_LAYERS` parameter can be multiple layers, contained in double quotes and separated by commas (with no spaces or other added characters). For example,

```
"Map_Frame.Hydrography,Map_Frame.Terrain.Contours"
```

`GDAL_PDF_LAYERS` specifies layers to turn on. `GDAL_PDF_LAYERS_OFF` specifies layers to turn off, which is sometimes more convenient. However, the relationship between these two parameters is somewhat complicated due to US Topo data characteristics.

- GDAL respects the PDF characteristics of “off by default” and “on by default,” but takes this information from the layer level. In a US Topo file, the orthoimage **folder** is off by default, but the one layer in this folder is **on** by default. In the absence of other instructions, GDAL will therefore convert the orthoimage layer. To advanced users reading this paper, this may look like a US Topo design flaw, but it is actually an intentional feature that benefits non-specialist users.
- If neither `GDAL_PDF_LAYERS` nor `GDAL_PDF_LAYERS_OFF` are used, the utility will convert all the normal on-by-default layers, including the orthoimage layer as per the previous bullet.
- Specifying `--config GDAL_PDF_LAYERS "ALL"` will convert all layers, which is normally not a very useful thing to do. However, the next bullet describes a variation that is useful.
- The two parameters can be used together. For example, `--config GDAL_PDF_LAYERS "ALL" --config GDAL_PDF_LAYERS_OFF "Images"` will convert all layers except the orthoimage.
- When writing software to automate these conversions, the safest thing to do is probably to build a list of all layers to be included as input to `GDAL_PDF_LAYERS`.

⁴ There is one significant caveat to this section: these procedures do not work correctly on the great majority of US Topo maps produced earlier than about July, 2011. There are several variations of the problem, but the most common effect is that GDAL will not create a GeoTIFF that excludes the orthoimage layer. The cause is not known, and the bug is not likely to be either found or fixed. Global Mapper **will** convert these files correctly. Any reader who discovers any helpful information about this issue is encouraged to send it to lmoore@usgs.gov. US Topo maps produced this early in the project have other deficiencies, and have all been replaced at least once, so USGS does not consider this problem a high priority.

This command will create a GeoTIFF that includes all layers except the orthoimage, even those layers that are not turned on by default by Adobe Reader:

```
3) gdal_translate MN_Grand_Marais_20160511_TM_geo.pdf marais_map.tif \  
  --config GDAL_PDF_LAYERS "all" --config GDAL_PDF_LAYERS_OFF "Images" \  
  --config GDAL_PDF_BANDS 3 --config GDAL_PDF_DPI 400
```

3.5 Reducing output file size

The output of commands 2 and 3 above are uncompressed RGB 24-bit TIFFs. For a typical US Topo, this file will be roughly 500 MB, a very large file for this kind of image.

File size can be dramatically reduced by various combinations of JPEG compression, different color models, and reduced resolution. This appears to be both faster and more reliable when done as a two-step conversion. First, use commands such as examples 2 and 3 above to create a 24-bit RGB TIFF, and then run `gdal_translate` a second time on the output. The following command, for example, will reduce the file size of the color composite Grand Marais map (example 3 above) from about 300 MB to about 8 MB:

```
4) gdal_translate -co COMPRESS=JPEG -co PHOTOMETRIC=YCBCR marais_map.tif \  
  marais_map_ycbr.tif
```

The compression and YCBCR color model slightly reduce the sharpness of the map linework, but not enough to be noticeable for most applications. Reducing the resolution further is also an option for many applications. See Appendix D for more information about photometric, compression, and resolution effects on file size.

Since most US Topo layers are vectors, the overall resolution of a US Topo PDF is ambiguous. The orthoimage layer has a ground resolution of 1 or 1.5 meter per pixel (depending on map date), which are about 600 DPI and 400 DPI respectively at map scale. 600 DPI is the declared resolution of the PDF and is the default output of `gdal_translate` operations. This output resolution is higher than needed for most applications. Reducing it with the `--config GDAL_PDF_DPI` as in the above examples reduces the output file size. For most applications, 300 DPI is probably an adequate resolution. Resolutions above 500 DPI provide very little improvement in image quality while rapidly increasing file sizes.

See the GDAL pages referenced in section 4 for more information about compression and color model options.

Appendix B of this paper describes a Perl script that partially automates selected GeoTIFF conversions of US Topo files with GDAL utilities.

4 More Information and Important URLs

The Grand Marais, MN map used in examples, the Global Mapper, script, a Perl script that uses the GDAL utilities, a Shapefile of 7.5-minute cells, and a copy of this paper are in a .zip file at:

ftp://ftpext.usgs.gov/pub/cr/co/lakewood/Larry_Moore/ustopo2tiff/, file name `ustopo2gtif_v2.5.zip`.

Global Mapper: <http://www.bluemarblegeo.com/products/global-mapper.php>

GDAL (Geospatial Data Abstraction Library): <http://www.gdal.org>

MinGW/MSYS: <http://www.mingw.org> (home) <http://sourceforge.net/projects/mingw/files> (download)

GDAL utility documentation pages:

- gdalinfo: <http://www.gdal.org/gdalinfo.html>
- gdal_translate: http://www.gdal.org/gdal_translate.html
- TIFF creation options: http://www.gdal.org/frmt_gtiff.html
- Geospatial PDF configuration options: http://www.gdal.org/frmt_pdf.html

Additional GDAL documentation for advanced users and programmers:

http://trac.osgeo.org/gdal/wiki/USGS_PDF_Topo

OSGeo: <http://www.osgeo.org/>

OSGeo-Live (open source GIS software bundled with a Linux distribution):

<http://live.osgeo.org/en/index.html>

The Historical Topographic Map Collection: <http://nationalmap.gov/historical/>

HTMC FAQs: <http://www.usgs.gov/faq/categories/9796>

US Topo: <http://nationalmap.gov/ustopo/>

US Topo FAQs: <http://www.usgs.gov/faq/categories/9797>

A CSV database dump of all US Topo and HTMC maps is posted at <http://geonames.usgs.gov/apex/f?p=262>. This list is refreshed nightly. It includes database identifiers, basic metadata, and download URLs for all distributable US Topo and HTMC product instances.

Appendix A. Global Mapper script for batch GeoPDF-to-GeoTIFF reformat

This script is included with the bundle of sample data and software at

ftp://ftpext.usgs.gov/pub/cr/co/lakewood/Larry_Moore/ustopo2tiff/

```
// This Global Mapper script converts USGS GeoPDF US Topo maps to Geotiffs.
// The script supports two outputs: maps with collars, or with collars clipped.
//   Read inline documentation below.
//
// Author: Steve Aichele, saichele@usgs.gov
// Modified by Larry Moore, lmoore@usgs.gov, last changes 5/2016
//
// Lots of stuff is hardcoded that may have to be modified depending on
//   1) vintage of US Topo and
//   2) desired content of output
//
// RUN INSTRUCTIONS:
// Place this script, and GeoPDF USTopos in a directory.
// Open GlobalMapper 13.1 or higher and run this script.
//
// If you want to clip map collars, the shapefile "All_7.5min_cells.shp"
//   must be in the same directory as the script and PDFs.
//
// USTopos with collars will have the prefix "WC" appended to the quad name

GLOBAL_MAPPER_SCRIPT VERSION="1.00"
UNLOAD_ALL

// set starting directory
DIR_LOOP_START DIRECTORY="." FILENAME_MASKS="*.pdf" RECURSE_DIR=NO

// Layer list follows.
// Comment out lines for layers you don't want in the output TIFF
// Notes:
//   There is variation in the actual content of different vintages of US Topos...
//   ...and also variation in layer names due to bugs and editorial design changes.
//   Having "extra" layers in this list does not hurt anything,
//   but having a layer missing means it cannot be exported to GeoTIFF
//   The list below is of most, but not all, layers that occur in any US Topo
//   in the period 2014-16
//   The presence of a layer in this list does *not* necessarily mean that layer is
//   present in all US Topo maps. The product is evolving, both in content and design.
//
// IMPORTANT: Previous versions of this script broke the layer parameters of the LOAD_FLAGS option
//   into multiple lines. Global Mapper 17.x appears to not support that syntax, and the
//   layer names must be strung together on one line.
//   The layer list here includes all US Topo layers *except* orthoimage, shaded relief,
//   PLSS, and bar code. The list must be edited to create other output.
//
// [If this listing is in a Word or PDF document, it shows wrapped text, but an operational
//   script must have all layers on one line. Be careful cutting and pasting from here
//   into a text document; better to download an actual copy of the script.]

IMPORT FILENAME="%FNAME_W_DIR%" \
  TYPE="PDF" RECTIFY="Automatic" \
  LOAD_FLAGS="0,600,Map Collar->Map Elements<SEP>Map Frame->Projection and Grids<SEP>Map Frame-
>Geographic Names<SEP>Map Frame->Structures<SEP>Map Frame->Transportation<SEP>Map Frame-
>Transportation->Road Names and Shields<SEP>Map Frame->Transportation->Road Features<SEP> Map Frame-
>Transportation->Trails<SEP>Map Frame->Transportation->Railroads<SEP>Map Frame->Transportation-
>Airports<SEP>Map Frame->Wetlands<SEP>Map Frame->Hydrography<SEP>Map Frame->Terrain<SEP>Map Frame-
>Terrain->Contours<SEP>Map Frame->Woodland<SEP>Map Frame->Boundaries<SEP>Map Frame->Boundaries-
>Jurisdictional Boundaries<SEP>Map Frame->Boundaries->Jurisdictional Boundaries-
>International<SEP>Map Frame->Boundaries->Jurisdictional Boundaries->State or Territory<SEP>Map
Frame->Boundaries->Jurisdictional Boundaries->County or Equivalent<SEP>Map Frame->Boundaries->Federal
Administered Lands<SEP>Map Frame->Boundaries->Federal Administered Lands->National Cemetery<SEP>Map
Frame->Boundaries->Federal Administered Lands->National Park Service<SEP>Map Frame->Boundaries-
>Federal Administered Lands->Department of Defense<SEP>Map Frame->Boundaries->Federal Administered
Lands->Forest Service,1" \
  SAMPLING_METHOD="NEAREST_NEIGHBOR" AUTO_CONTRAST="NO" CONTRAST_SHARED="YES"
CONTRAST_MODE="NONE" \
  CLIP_COLLAR="NONE" TEXTURE_MAP="NO"
```

```
// This command, if uncommented, creates GeoTiff's with Collars
EXPORT_RASTER FILENAME="%DIR%WC_%FNAME_WO_EXT%.tif" TYPE=GEOTIFF PALETTE=HALFTONE

// This command, if uncommented, outputs GeoTIFFs without map collars -- everything outside the
projection line is deleted
// Clipping requires a shapefile of the standard 7.5-minute cells to define where the clip is done
// EXPORT_RASTER FILENAME="%DIR%FNAME_WO_EXT%.tif" TYPE=GEOTIFF PALETTE=HALFTONE
POLYGON_CROP_FILE="All_7.5min_cells.shp"

UNLOAD_ALL

DIR_LOOP_END
```

Appendix B. Perl Script for GDAL GeoPDF-to-GeoTIFF reformat

Included in the zip file that goes with this paper

(ftp://ftpext.usgs.gov/pub/cr/co/lakewood/Larry_Moore/ustopo2tiff/) is a Perl script named “ustopo2gtif.pl” that automates selected options for GDAL GeoPDF-to-GeoTIFF conversion. This script was written by the author for personal use, and is not especially robust -- no guarantees, no support, use at your own risk, etc. The capabilities and behavior of this script are summarized here.

```
$ ustopo2gtif.pl -?
```

```
use: ustopo2tif [options] [input_pdf]
    Converts USGS US Topo GeoPDFs to a collection of GeoTIFFs.
    One PDF is an optional input; if no file specified, all US Topo maps in the
        current directory will be processed.
    Input and output locations are the current directory
```

```
options:
```

```
--lines
    Output a GeoTIFF file for each linework layer.
--images
    Output GeoTIFF files for shaded relief and orthimage layers.
--map
    Make a color composite; layers to include are hard-coded at the top of the script.
--zip
    Put all outputs into a .zip file.
--all
    Select all the above options; that is, make all GeoTIFFs and zip them.
--help or -?
    Print this information
```

The script has three basic behavior options:

- Each layer in the GeoPDF is converted to a GeoTIFF file. Layers that contain no data are detected and named distinctly.
- A color composite that simulates a traditional topographic line map is created. The layers that go into this composite are hard-coded in the script, though they are not difficult to change by modifying the code. (Several other layer combinations, for image and shaded relief maps, are in the script in commented-out lines.)
- All output files can be bundled in one zip file.

Command-line options specify which of these behaviors to activate. If a single PDF file is named on the command line, the script converts just this file. If no input file is specified, the script converts all US Topo files in the current directory. Neither option depends on any particular file name convention; the script checks selected header information to judge whether a PDF is a USGS US Topo file.

Outputs use the base name of the input file, with additional characters to describe the layer. Outputs are created in the current directory.

In addition to the layers that compose the color composite, several other output parameters are hard-coded. These include values for photogrammetric type, compression type, and output resolution. All hardcoded parameters are defined near the top of the script.

Appendix C. List of US Topo layers

The layers present in US Topo maps vary due to addition of new content and product design changes. It is likely the layer table of contents will never be completely stable. The following table shows layers that occur in the set of “current maps” as of May, 2016 – mostly maps published from mid-2013 through mid-2016. Older product instances will contain other layers not included in this list.

Scripts for either Global Mapper or `gdal_translate` can include “extra” layers, which are simply ignored by the programs (`gdal_translate` will display warning messages). Fairly general scripts can therefore be written by including all the layer names that refer to a feature type of interest.

The table uses GDAL syntax. To convert to Global Mapper syntax, replace underbars (`_`) with spaces, and periods (`.`) with `->` Example:

- [GDAL syntax] `Map_Frame.Terrain.Contours`
- [Global Mapper syntax] `Map Frame->Terrain->Contours`

Table 1. Layers present in the set of “current maps” as of May, 2016. If the “last use” is 2016 and the map count is relatively large, the layer is part of the current product design. The layer design has grown more stable with time, and this list will probably to shrink as 2013 maps are replaced, though it may never be completely fixed and stable. Shaded rows are folders that never contain data of their own.

<i>Layer name</i>	<i>Number of maps</i>	<i>Last use</i>
Map_Collar	55,690	May-16
Map_Collar.Map_Elements	55,690	May-16
Map_Frame	55,690	May-16
Map_Frame.Projection_and_Grids	52,101	May-16
Map_Frame.Projection_and_Grids.Geographic_and_Grid_Tics	971	Mar-13
Map_Frame.Projection_and_Grids.Grid_Lines	971	Mar-13
Map_Frame.Projection_and_Grids.Projection_Coordinate_Values	971	Mar-13
Map_Frame.Projection_and_Grids.Projection_Line_Mask	971	Mar-13
Map_Frame.Projections_and_Grids	3,589	Jan-14
Map_Frame.Geographic_Names	55,690	May-16
Map_Frame.Geographic_Names.Geographic_Names	971	Mar-13
Map_Frame.Structures	52,467	May-16
Map_Frame.Structure_Features_and_Names	3,223	Sep-13
Map_Frame.Structures.Structures	971	Mar-13
Map_Frame.Structures.Structures_Names	971	Mar-13
Map_Frame.Transportation	55,690	May-16
Map_Frame.Transportation.Roads	971	Mar-13
Map_Frame.Transportation.Road_Names_and_Shields	55,690	May-16
Map_Frame.Transportation.Road_Features	54,719	May-16
Map_Frame.Transportation.Trails	47,928	May-16
Map_Frame.Transportation.Railroads	52,467	May-16
Map_Frame.Transportation.Railroad_Names	971	Mar-13

Map_Frame.Transportation.Railroad_Features_and_Names	3,223	Sep-13
Map_Frame.Transportation.Airports	52,467	May-16
Map_Frame.Transportation.Airport_Features_and_Names	3,223	Sep-13
Map_Frame.Transportation.Airport_Names	971	Mar-13
Map_Frame.Hydrography	52,467	May-16
Map_Frame.Wetlands	10,579	May-16
Map_Frame.Hydrography.Hydrography_Features	971	Mar-13
Map_Frame.Hydrography.Hydrography_Names	971	Mar-13
Map_Frame.Hydrography_Features_and_Names	3,223	Sep-13
Map_Frame.PLSS	36,705	May-16
Map_Frame.Terrain	54,719	May-16
Map_Frame.Terrain.Contours	51,496	May-16
Map_Frame.Terrain.Contour_Features_and_Names	3,223	Sep-13
Map_Frame.Terrain.Shaded_Relief	54,719	May-16
Map_Frame.Contours	971	Mar-13
Map_Frame.Contours.Contour_Features	971	Mar-13
Map_Frame.Contours.Contour_Names	971	Mar-13
Map_Frame.Woodland	51,496	May-16
Map_Frame.Woodland_Features	3,223	Sep-13
Map_Frame.Land_Cover	971	Mar-13
Map_Frame.Land_Cover.Woodland	971	Mar-13
Map_Frame.Boundaries	52,467	May-16
Map_Frame.Boundaries.Boundaries	971	Mar-13
Map_Frame.Boundaries.Boundary_Names	971	Mar-13
Map_Frame.Boundary_Features_and_Names	3,223	Sep-13
Map_Frame.Boundaries.Jurisdictional_Boundaries	51,496	May-16
Map_Frame.Boundaries.Jurisdictional_Boundaries.International	51,496	May-16
Map_Frame.Boundaries.Jurisdictional_Boundaries.State_or_Territory	51,496	May-16
Map_Frame.Boundaries.Jurisdictional_Boundaries.County_or_Equivalent	51,496	May-16
Map_Frame.Boundaries.Federal_Administered_Lands	51,412	May-16
Map_Frame.Boundaries.Federal_Administered_Lands.National_Park_Service	50,101	May-16
Map_Frame.Boundaries.Federal_Administered_Lands.Department_of_Defense	51,412	May-16
Map_Frame.Boundaries.Federal_Administered_Lands.Forest_Service	51,412	May-16
Map_Frame.Boundaries.Federal_Administered_Lands.Fish_and_Wildlife_Service	3,441	May-16
Map_Frame.Boundaries.Federal_Administered_Lands.Bureau_of_Land_Management	398	Jun-15
Map_Frame.Boundaries.Federal_Administered_Lands.National_Cemetery	24,812	May-16
Images	55,635	May-16
Images.Orthoimage	55,635	May-16
Barcode	48,281	May-16

Appendix D. GeoTIFF file sizes with various GDAL options

The following table shows GeoTIFF file sizes for various types of data content converted using `gdal_translate` and with several combinations of photometric, compression, and resolution. There is little variation in visual quality between any of these combinations, but a wide range of file sizes. Particular applications may have reasons for preferring certain combinations. These data are from tests on only a few US Topo instances, and the amount of variation that can be expected between maps has not been studied.

Even so, it is clear that for applications that do not have specialized requirements for format or individual pixel color value preservation, the YCBCR color model with JPEG compression and relatively low resolution is the best choice.

The “traditional line map” column is a color composite that includes all vector layers, the map collar and grids, and the green timber layer, but omits the orthoimage, shaded relief, and any land boundary layers represented with transparent area fill.

Photometric/ Compression	Output Resolution (dpi)	GeoTIFF file size in MB		
		Contours and hydrography	Traditional line map	Orthoimage
RGB/LZW	600	30	50	514
	400	16	30	232
RGB/JPEG	600	85	63	92
	400	27	36	49
YCBCR/JPEG	600	18	25	34
	400	10	14	18
RGB/Packbits	600	82	222	460
	400	32	103	205

Appendix E. Finding USGS GIS data

As noted in the introduction to this paper, US Topo maps are GIS outputs -- a repackaging of GIS data for non-specialist map users that were not intended to be GIS inputs. Before going to the work of reversing this process and making GeoTIFFs for GIS uses, consider USGS products designed for GIS use. Web services (point 2 below) can be especially convenient for accessing GIS-friendly base maps.

Following is a summary of distribution interfaces for all products of the USGS National Geospatial Program (NGP), and gateways to other USGS programs and products.

1. [The National Map Download Client](http://nationalmap.gov) is the primary interface for all data of *The National Map* (<http://nationalmap.gov>). Many NGP data products can be downloaded through this viewer. [Data download FAQs](#).
2. Data from many of *The National Map* databases are also available through [GIS services](#).
3. Current US Topo topographic maps, and legacy maps in the Historical Topographic Map Collection, can be downloaded through *The National Map* Download Client (point 1 above), or [through other interfaces](#) dedicated to topographic quadrangle maps.
4. Individual data themes (e.g., elevation, hydrography, geographic names...) also maintain interfaces to specific products. Go to *The National Map* homepage (<http://nationalmap.gov>) and select a data theme.
5. More orthoimage, elevation, and land cover data are served from the Earth Resources Observation and Science (EROS) Center, including satellite data not served through other interfaces. See <http://eros.usgs.gov/find-data> for general information. The Earth Explorer viewer (<http://earthexplorer.usgs.gov>) is the primary download application. Some archived datasets are also served through Earth Explorer; see the Long Term Archive page at <https://lta.cr.usgs.gov/>.
6. [Bulk data deliveries and downloads](#) are also available for some products

For paper maps, books and scientific reports, educational literature, and other hardcopy products, visit the USGS Store (<http://store.usgs.gov/>) or the USGS Publications Warehouse (<http://pubs.er.usgs.gov/>). To search for other types of earth science data from USGS scientific disciplines, start at the USGS home page (<http://usgs.gov>). You can also contact USGS by email or phone (<http://www.usgs.gov/ask>).