CD-ROM BOOKLET

GLOBAL MONTHLY AVHRR CLIMATOLOGY OVER LAND: Clear-sky top-of-the-atmosphere variables

Garik Gutman, Dan Tarpley, Aleksandr Ignatov,
Satellite Research Laboratory, NOAA/NESDIS, Washington, D.C.

Steve Olson
Research and Data Systems Corporation, Greenbelt, MD

1. INTRODUCTION

These CD-ROMs contain a global 5-year top-of-the-atmosphere (TOA) clear-sky monthly climatology over land produced from the Advanced Very High Resolution Radiometer (AVHRR). This was done within the Global Vegetation Index (GVI) project in the Satellite Research Laboratory of the Office of Research and Applications, National Environmental Satellite Data and Information Service (NESDIS), National Oceanic and Atmospheric Administration (NOAA). The GVI project was sponsored by the NOAA Climate and Global Change Program for the purpose of developing improved operational GVI products that are useful for global climate studies.

The data on these CD-ROMs are the third generation GVI D-level products described in the GVI User's Guide (Kidwell, 1994) and Gutman et al. (1995), the latter included on this CD. The products are given at 0.15^o resolution in a rectangular latitude/longitude (equal angle) projection. They include 5-year means and standard deviations for each month of visible (0.63 ?m) and near-infrared (0.85 ?m) reflectances ($@_1$, $@_2$), normalized difference vegetation index (NDVI=($@_2$-$@_1$)/($@_1$+$@_2$)), thermal IR (11 and 12 ?m) brightness temperatures (T_4, T_5), precipitable water index (PWI=T_4-T_5), and observational solar zenith angle.

A number of auxiliary files are included to facilitate data analysis. The satellite based files are the Quality Description (QD) files, an Auxiliary Mask (AM) file, and a 10'-resolution remapped version of NDVI. The non-satellite datasets contain soil and vegetation maps, topography, precipitation, etc. (see the content for details).

The QD files are monthly and include 5-year averages of the number of cloud-free observations (weeks), the biases of satellite scan angle, and the snow masks based on the visible reflectance and 11 ?m temperature data. The AM files are stationary (independent of the month) and contain land/sea tag, political boundaries and two surface type masks based on maximum (V_{max}) and minimum (V_{min}) annual NDVI values characterizing low annual variability in NDVI (V_{max}-V_{min}<0.2) for evergreen (V_{max}>0.45) and deserts (V_{max}<0.2). Descriptions of the QD and AM files are given in Appendix 1.

The data processing is briefly described below. Its full description is given in Gutman et al. (1995) and is provided on this CD-ROM in the directory "/document/avhrrcli". Programs to read the data on several platforms are provided in Appendix 2.

2. BRIEF DATA PROCESSING DESCRIPTION

The processing comprises three stages: B (weekly), C (monthly), and D (climatology) (see Gutman et al. 1995 for the nomenclature and data products generation).

The processing of weekly data (B3.0/W15 ? B3.2/W15) was based on the GVI weekly composite raw data (counts) -- the second generation of GVI. The processing steps included: 1) conversion of counts in AVHRR thermal channels to T_4 and T_5 and correction for calibration non-linearity; 2) calculation of reflectances in the solar channels, $@_1$ and $@_2$, using the updated (Pathfinder) calibration, which accounted for sensor degradation, correction for Sun-Earth distance and normalization to overhead sun; 3) cloud/clear identification using T_4-thresholds dependent upon month, region, and viewing angle. The NDVI=$(@_2-@_1)/(@_2+@_1)$ and PWI=(T_4-T_5) were calculated at the B-level to retain accuracy. All B-product values were packed into 8 bits. Note that no atmospheric or angular corrections were applied to these data, hence no *surface* geophysical variables have been derived. This is yet to be done (see details in Gutman et al. 1995).

The monthly processing (B3.2/W15 ? C3.2/M15) flow included the following steps: 1) QC flags were applied on a weekly basis, resulting in data gaps, 2) each quantity was averaged over one month (five weekly composites) for each map cell to partially fill in the data gaps and reduce some of the angular variability (the weeks that spill over into the next month were used for both the current and the previous months); 3) bi-linear spatial interpolation was applied to the missing data areas with persistent cloudiness in monthly averaged images; 4) 3x3 map cell running average smoothing was done to partially account for the imperfection of cloud screening, to filter out atmospheric and angular variabilities, and to compensate for different number of observations used in monthly averaging and for random spatial sampling from the original Global Area Coverage (GAC) data into the GVI map cells. The monthly time series (C3.2/M15) of all the aforementioned variables for each individual year of the period from April 1985 - September 1994 are available at NOAA National Climatic Data Center (NCDC; e-mail: satorder@ncdc.noaa.gov).

The 5-year TOA AVHRR climatology (D3.2/M15) on this CD has been developed from the C3.2/M15 monthly fields of TOA $@_1$, $@_2$, NDVI, T_4, T_5, and PWI using data from April 1985 to December 1987 (NOAA-9), and from January 1989 to March 1991 (NOAA-11).

3. DISC ORGANIZATION, CONTENT, NAMING CONVENTIONS

There are two volumes to this set of CD-ROMs.

Volume 1 contains the core climatologies and some ancillary vector overlay data, documentation for these data, and some ancillary software tools.

Volume 2 contains selected regriddings of the core climatologies, designed to register with the National Geophysical Data Center’s Global Change Data Base. In addition, data from that data base are included. Scientists wishing to perform integrated global environmental studies may find these ancillary data valuable. Documentation for the ancillary data is also included in Volume 2.

The CD-ROMs have three directories discussed below. The directory structure of each CD-ROM is given in the root directory the respective CD-ROM.

3.1 Directory "data"

This directory contains has five main subdirectories: "average", "qualflag", "standev", "ncillary", and "vector".

The subdirectories "average" and "standev" contain respectively the 5-year monthly means and standard deviations of AVHRR reflectances ($@_1$ and $@_2$) and temperatures (T_4 and T_5), Normalized Difference Vegetation Index (NDVI), Precipitable Water Index (PWI), sensor scan angle (sca), and the solar zenith angle (sza). The 8-bit files can be converted to floating point values using:

Conversion	Naming convention
$@_1(i) = 45.0 * (i) / 255.0 + 5.0$	ch1
$@_2(i) = 35.0 * (i) / 255.0 + 15.0$	ch2
$T_4(i) = 76.0 * (i) / 255.0 + 250.0$	ch4
$T_5(i) = 76.0 * (i) / 255.0 + 250.0$	ch5
$ndvi(i) = 0.8 * (i) / 255.0 - 0.1$	ndvi
$pwi(i) = 7.0 * (i) / 255.0 - 2.0$	pwi
$sca(i) = 110.0 * (i) / 255.0 - 55$	sca
$sza(i) = 50.0 * (i) / 255.0 + 20.0$	sza
$\tilde{O}_1(i) = 4.0 * (i) / 255.0$	ch1
$\tilde{O}_2(i) = 4.0 * (i) / 255.0$	ch2
$\tilde{O}_4(i) = 3.0 * (i) / 255.0$	ch4
$\tilde{O}_5(i) = 3.0 * (i) / 255.0$	ch5
$\tilde{O}_{ndvi}(i) = 0.1 * (i) / 255.0$	ndvi
$\tilde{O}_{pwi}(i) = 0.5 * (i) / 255.0$	pwi
$\tilde{O}_{sca}(i) = 26.0 * (i) / 255.0$	sca
$\tilde{O}_{sza}(i) = 8.0 * (i) / 255.0$	sza

where i is a count value from 0 to 255.

The "qualflag" directory contains 12 monthly QD files monqd.img, where "mon" is as above. The "qualflag" directory also contains an auxiliary mask (AM) file maskam.img which represents fields that do not change with month (see Appendix 1 for details).

The abovementioned subdirectories are in Volume 1.

The subdirectory "ncillary" contains global raster data sets on soils, vegetation, etc. from the Global Ecosystems Database and TerrainBase produced by NGDC. This directory is in Volume 2.

The subdirectory "vector" contains ancillary vector data (coastlines, national boundaries, etc.) in GRASS and IDRISI format. This directory is in Volume 1.

3.2 Directory "document"

This directory contains the documentation for the data. Its subdirectories describe all the databases used to assemble this CD-ROM. A description of the contents for each subdirectory can be found in the file aareadme.1st in the root directory. Portions of this directory are on Volumes 1 and 2, describing the data residing on those two CD-ROMs, respectively.

3.3 Directory "software"

This directory contains three subdirectories: "freeform", "geovu", and "html".

The directory "freeform" contains FreeForm utilities, which the user may find useful for format conversion and documentation for several types of data. Documentation is available in the \geovu\freeform\document directory. Current documentation is updated regularly at the Website of NOAA's National Geophysical Data Center. Currently, that website is http://www.ngdc.noaa.gov/seg/freeform/freeform.html.

The directory "geovu" contains GeoVu browse and visualization software, which has been adapted for use on these CD-ROMs. GeoVu is provided for PCs, Macintoshes, and for Sun and Silicon Graphics UNIX workstations. Documentation and sample data are also provided. Under the GeoVu directory for each hardware platform (for example, PC), there are installation instructions. Make sure that the gut_gvi1.men, gut_gvi2.men, and avhrrcli.lst files are copied from the CD to the directory you create on your computer for running GeoVu. For example, if you create a directory called c:\geovu on your PC to contain GeoVu software, and your CD-ROM is drive e, enter the commands:

```
copy e:\software\geovu\pc\gut_gvi1.men c:\geovu
copy e:\software\geovu\pc\gut_gvi2.men c:\geovu
copy e:\software\geovu\pc\avhrrcli.lst c:\geovu
```

to copy the files from the CD to your GeoVu software directory on your computer.

The geovu package is a general data access tool developed by NGDC to display grids and images, display and plot tabular data sets, and examine related documentation during the data selection process.

The directory "html" contains much of the "plastic Website" on the CD-ROMs. You may ignore this directory, unless you want to "sleuth" the design of the CD.

3.4 Naming conventions

For the averages and standard deviations the naming convention is: varmon.img, where "var" is the variable, defined above (var=ch1, ch2, ...; e.g., pwi is the precipitable water index), and "mon" is the first three letters of the month (e.g., mon=jan, feb, ...)

ACKNOWLEDGMENTS

The GVI project has been supported by the NOAA Climate and Global Change Program. Without its continuous funding, the generation of the 5-year global AVHRR climatology over land would not have been possible. Critical reviewing by the NOAA program manager Dr. A. Gruber motivated this work. are acknowledged for the development of the processing and visualization system. Contributions in the project by J. Powers (NOAA/NESDIS), M. Halpert and C. Ropelewski (NOAA Climate Prediction Center), and P. Schultz, R. Hucek and L. Rukhovetz, D. Sullivan of Research Data Systems Corporation (RDC) are acknowledged. This work was accomplished when one of the authors (A.I.) held National Research Council Associateship at the Satellite Research Laboratory, NOAA/NESDIS, on leave from the Marine Hydrophysics Institute, Sevastopol, Ukraine. David Hastings of the National Geophysical Data Center managed the adaptation of these data to CD-ROM; David Schoolcraft assisted in the adaptation.

The requested form of acknowledgment for use of the data on this CD-ROM is as follows:

Gutman, G., D. Tarpley, A. Ignatov, and S. Olson, 1995: Global Monthly AVHRR climatology over land: Clear-sky top-of-the-atmosphere variables. Ed. D. Hastings. United States Dept. of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Geophysical Data Center, Boulder, Co 80303, USA.

References

Gutman, G., D. Tarpley, A. Ignatov, and S. Olson, 1995: The enhanced NOAA Global Land Dataset from the Advanced Very High Resolution Radiometer. Bulletin of the American Meteorological Society, **76**, 1141-1156,

Kidwell, K., 1994: Global Vegetation Index User’s Guide. U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Climatic Data Center, Satellite Data Services Division, 126 pp.

Fig. 1: The organizational structure of CD-ROM.

APPENDIX 1

QUALITY DESCRIPTION (QD) FILES

The 8-bit QD files were generated for each month to describe some of the inherent statistical characteristics. Each of the 8 bits is a 5-year average of the corresponding bit for each month of each individual year. The following table describes how each bit was assigned:

BIT NUMBER	CRITERIA/COMMENT	BIT VALUE
1	nobs = 0 or 1 (mostly cloudy)	1, else 0
2	nobs = 2 or 3 (moderately cloudy)	1, else 0
3	nobs = 4 or 5 (mostly clear)	1, else 0
4	|sca| < 20 deg (near nadir)	1, else 0
5	sca > 20 deg (forward scatter bias)	1, else 0
6	sca < -20 (back scatter bias)	1, else 0
7	T_4 < 270K and $@_1$>20% (stable snow)	1, else 0
8	270K3$T_4$3280K and $@_1$> 20% (unstable snow)	1, else 0

where "nobs" is number of clear observations (weekly composites after cloud screening) per month per pixel.

AUXILIARY MASKS (AM) FILE

An 8-bit AM file was generated for delineating land, political boundaries, and masking two surface types with low annual variability: evergreen vegetation and desert. The following table describes each bit:

BIT NUMBER	CRITERIA/COMMENT	BIT VALUE
1	land location (no coast used)	1, else 0
2	state borders,inland water	1, else 0
3	evergreen (V_{max}-V_{min}<0.2 and V_{max}>0.45)	1, else 0
4	desert (V_{max}-V_{min}<0.2 and V_{max}<0.2)	1, else 0
5	blank	
6	blank	
7	blank	
8	blank	

where V_{max} and V_{min} are annual maximum and minimum NDVI,respectively.

APPENDIX 2

READING THIRD GENERATION GVI IMAGES

I. INTRODUCTION

This CD-ROM contains global Top-of-the Atmosphere (TOA) images and products derived from Advanced Very High Resolution Radiometer (AVHRR). These products were developed on a Silicon Graphics Inc. (SGI) workstation using the C language as part of the Global Vegetation Index (GVI) project by the Climate Research and Applications Division (CRAD) Land Surface Team of the Office of Research and Applications (ORA). The data on this CD-ROM consists of D-Level products, which are described in the GVI Users Guide (Kidwell, 1994). In order to ensure easy portability across all platforms, the data was scaled to 8-bits or 1-byte. Each uncompressed image is made up of 2500 columns and 904 rows of data, with 0.144 degrees (nominally 0.15 degrees) latitude/longitude resolution, giving a total size of 2260000 bytes/image. The purpose of this section is to read the 2500 x 904 image.

Each image is read by rows, starting with row 1 at 75 N, and ending with row 904 at 55 S. Every row begins at 180 W longitude, and each successive map cell would be 0.144 degrees east of the previous map cell. Therefore, the upper left corner of the image starts at 75 N, 180 W, and ends in the lower right corner of the image, at 55 S, 180 E. The physical range of the data is between 0 and 255, with all ocean values containing the value 0.

Since the binary imagery is character data, you cannot browse through it. However, the character data can be easily converted into integer values. The following is a list of FORTRAN and C programs that can be used as a guide for reading the files. Each of these programs will convert the character data into integer values. Reading the 8-bit imagery is somewhat platform specific. The coding is broken down by language and platform.

NOTE: THE OUTPUT DATA IS BEING WRITTEN TO UNIT=6, THE SCREEN.

II. USING FORTRAN

The following is a list of programs that should be used based on your platform.

A. Silicon Graphics (SGI), SUN Workstations, VAX AND DEC ALPHA WORKSTATIONS

One word of caution with the VAX and DEC ALPHA. Please make sure that the data read from the CD-ROM or whatever media type you are using blocks the data in 2500 byte blocks. Particularly with the VAX and DEC ALPHA machines, there is a tendency to default to 512 byte blocks, which means you will NOT be able to read the data with these programs.

To mount the CD-ROM on a VAX or DEC ALPHA machine, use the following as an example:

```
mount devicename /media=cd /und=fixed:none:2500)
```

```
      PROGRAM CONVRT
C
```

```
C    THIS PROGRAM IS DESIGNED TO READ 8-BIT IMAGE DATA AND
C    OUTPUT THE DATA AS ASCII TEXT WITH 2260000 RECORDS.  THIS
C    WORKS OUT TO 1 RECORD PER PIXEL.  UNIT 10 IS THE INPUT FILE
C    YOU ARE TRYING TO READ, AND IN THE EXAMPLE, UNIT 6 IS WHERE
C    THE OUTPUT IS GOING TO.
C
C ***********************************************************
C   COMMENTS ABOUT THE FORTRAN READING PROGRAM
C ***********************************************************
C
C   THE NUMBER OF COLUMNS IS 2500 AND THE NUMBER OF ROWS IS 904.
C   THE BLOCKING FACTOR IS 2500 BYTES.  YOU'LL NOTICE THAT THE
C   RECL (RECORD LENGTH) IS 625.  625 REPRESENTS THE NUMBER OF 4-
C   BYTE WORDS PER RECORD.  THE DATA WILL BE OUTPUT TO UNIT=6 ONE
C   PIXEL AT A TIME.  THEREFORE, THE OUTPUT FILE WILL CONTAIN
C   2260000 LINES OR RECORDS IN IT.  THE DATA STARTS IN THE UPPER
C   LEFT HAND CORNER AT 75N, 180W AND ENDS IN THE BOTTOM RIGHT
C   HAND CORNER AT 55S, 180E.  EVERY ROW BEGINS AT 180 W
C   LONGITUDE, AND EACH SUCCESSIVE MAP CELL IS 0.144x EAST OF THE
C   PREVIOUS CELL.  THE OUTPUT DATA RANGE WILL BE BETWEEN 0 AND
C   255.  THE OCEAN MAP CELLS ARE ALL ZEROS.
C
C **************************
C     DECLARE ALL VARIABLES
C **************************
C
        CHARACTER*1 IMAGE(2500)
        CHARACTER*8 HEX(2500)
        INTEGER DATA(2500)
C
C ********************************
C         OPEN FILES TO READ
C ********************************
C
       OPEN(UNIT=10,FILE='GVI.DATA',ACCESS='DIRECT',
     $RECL=625,FORM='UNFORMATTED',STATUS='UNKNOWN')

         DO 10 I=1,904
          READ(10,REC=I) IMAGE

           DO 15 J=1,2500
             WRITE(HEX(J),'(Z2)') IMAGE(J)
             READ(HEX(J),'(Z2)') DATA(J)
 15        CONTINUE

           DO 20 J=1,2500
             WRITE(6,'(I4)') DATA(J)
 20        CONTINUE
 10      CONTINUE
         STOP
         END

*****************************************************************
```

B. IBM MAINFRAMES

```
      PROGRAM CONVRT
C
C    THIS PROGRAM IS DESIGNED TO READ 8-BIT IMAGE DATA AND
C    OUTPUT THE DATA AS ASCII TEXT.
C
C ***********************************************************
C   COMMENTS ABOUT THE FORTRAN READING PROGRAM
C ***********************************************************
C
C   THE NUMBER OF COLUMNS IS 2500 AND THE NUMBER OF ROWS IS 904.
C   THE BLOCKING FACTOR IS 2500 BYTES.  THE DATA WILL BE OUTPUT
C   TO UNIT=6 ONE PIXEL AT A TIME.  THEREFORE, THE OUTPUT FILE
C   WILL CONTAIN 2260000 LINES OR RECORDS IN IT.  THE DATA STARTS
C   IN THE UPPER LEFT HAND CORNER AT 75N, 180W AND ENDS IN THE
C   BOTTOM RIGHT HAND CORNER AT 55S, 180E.  EVERY ROW BEGINS AT
C   180 W LONGITUDE, AND EACH SUCCESSIVE MAP CELL IS 0.144x EAST
C   OF THE PREVIOUS CELL.  THE OUTPUT DATA RANGE WILL BE BETWEEN
C   0 AND 255.  THE OCEAN MAP CELLS ARE ALL ZEROS.
C ********************************
C    DECLARE ALL VARIABLES
C ********************************
C
      CHARACTER*1 IMAGE(2500)
      INTEGER*2 DATA(2500)
      EQUIVALENCE(IMAGE,DATA)
C
C********   OPEN FILES TO READ   ********************************
C
      OPEN(UNIT=10,FILE='GVI.DATA')

     DO 10 I=1,904
       READ(10,'(125(20A1))') IMAGE
        DO 20 J=1,2500
         WRITE(6,'(I4)') data(J)
 20     CONTINUE
 10  CONTINUE
     STOP
     END
```

****

C.   CRAY

With the CRAY, the following script (cray.sh) should be used used with the compiled FORTRAN code.

```
      PROGRAM CONVRT
C
C    THIS PROGRAM IS DESIGNED TO READ 8-BIT IMAGE DATA AND
C    OUTPUT THE DATA AS ASCII TEXT.
C
C ***********************************************************
C   COMMENTS ABOUT THE FORTRAN READING PROGRAM
C ***********************************************************
```

```
C
C   THE NUMBER OF COLUMNS IS 2500 AND THE NUMBER OF ROWS IS 904.
C   THE BLOCKING FACTOR IS 2500 BYTES.  YOU'LL NOTICE THAT THE
C   RECL (RECORD LENGTH) IS 625.  625 REPRESENTS THE NUMBER OF 4-
C   BYTE WORDS PER RECORD.  THE DATA WILL BE OUTPUT TO UNIT=6 ONE
C   PIXEL AT A TIME.  THEREFORE, THE OUTPUT FILE WILL CONTAIN
C   2260000 LINES OR RECORDS IN IT.  THE DATA STARTS IN THE UPPER
C   LEFT HAND CORNER AT 75N, 180W AND ENDS IN THE BOTTOM RIGHT
C   HAND CORNER AT 55S, 180E.  EVERY ROW BEGINS AT 180 W
C   LONGITUDE, AND EACH SUCCESSIVE MAP CELL IS 0.144x EAST OF THE
C   PREVIOUS CELL.  THE OUTPUT DATA RANGE WILL BE BETWEEN 0 AND
C   255.  THE OCEAN MAP CELLS ARE ALL ZEROS.
C
C ********************************
C    DECLARE ALL VARIABLES
C ********************************
C
        CHARACTER IMAGE
C
C
         Do J = 1,904
          Do I = 1 , 2500
1          READ(10) IMAGE
           write(6,*) ichar(IMAGE)
          enddo
         enddo
         STOP
         END
```

The shell script to use with the FORTRAN code is the following:

```
#!/bin/sh

assign -R
assign -a gvi.data -sunblocked u:10

cf77 filename

 ./a.out

exit
```

where,

assign -R = Reset all assign attributes.

assign -a filename -sunblocked u:10 = data file "filename" has structure unblocked and has been assigned to unit 10.

cf77 filename = Use the cray FORTRAN 77 compiler on input file "filename" containing program CONVRT code.

```
./a.out  = The executable
```

****

III. USING C Language

The following program can be universally used on all platforms which have a C compiler.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>

/*
   The purpose of this program is to read 1-byte plate carree
   images (C-level or D-level) and output them as ascii text one
   byte at a time.  One word of caution, if you are using a VAX
   DEC ALPHA machine, then you need to make sure that the data
   has been stored in 2500 byte blocks, NOT 512 byte blocks.
*/

static FILE *fp1, *fp2;

void main(int argc, char *argv[])
{
 char image[2260000];
 int xsize=2500, ysize=904;
 register int i;
 short data;

  if (argc != 3)
  {
    printf("usage -- image_read inputfile outputfile\n");
    exit(EXIT_FAILURE);
  }

  if (!(fp1=fopen(argv[1],"rb")))
  {
   printf("Cannot open file %s to read\n",argv[1]);
   exit(EXIT_FAILURE);
  }


  if (!(fp2=fopen(argv[2],"w")))
  {
   printf("Cannot open file %s to write\n",argv[2]);
   exit(EXIT_FAILURE);
  }

  fread(&image[0],1,xsize*ysize,fp1);

  for(i=0;i<2260000;i++)
  {
```

```
    data=(short) image[i];
    fprintf(fp2,"%d\n",data);
  }

/*  Write output to user-defined file  */


}
```

To compile the program, use the following command:

```
cc read_image.c -o read_image
```

Then to run the program, use the following:

```
read_image inputfile outputfile
```

where,

read_image = The executable

inputfile = The input 8-bit binary image

outputfile = The output ASCII image

The output ASCII image will be a sequential file, with each pixel constituting a new record. Therefore, the entire image will have 2260000 records/lines.