

DAO Office Note 96-01

## Office Note Series on Global Modeling and Data Assimilation

Richard B. Rood, Head  
*Data Assimilation Office*  
*Goddard Space Flight Center*  
*Greenbelt, Maryland*

## Documentation of South American Time-series Files for the Multi-year GEOS-1 Assimilation

Arlindo da Silva

*Data Assimilation Office, Goddard Laboratory for Atmospheres*  
*Goddard Space Flight Center, Greenbelt, Maryland*



Goddard Space Flight Center  
Greenbelt, Maryland 20771  
February 1996

### **Abstract**

This document describes a subset of the pressure level data sets in time-series format obtained from the multi-year assimilation with the GEOS-1 DAS (currently March 1985 to November 1993). The data sets include selected prognostic and diagnostic quantities (both surface and upper air) from the full assimilated data set, restricted to the SOUTH AMERICAN region. The data are given at a variety of temporal resolutions: 3-hourly, 6-hourly, daily and monthly means. The files are available by anonymous ftp from [dao.gsfc.nasa.gov](ftp://dao.gsfc.nasa.gov). An on-line version of this document with the relevant links can be accessed from

<ftp://niteroi.gsfc.nasa.gov/www/on9601/sa.html>

Consult also the Data Assimilation Office's home page at

<http://dao.gsfc.nasa.gov/>

## Contents

<b>Abstract</b>	<b>iii</b>
<b>1 Overview</b>	<b>1</b>
<b>2 Contents of the data sets</b>	<b>2</b>
<b>3 Description of 3- and 6-hourly South American data sets</b>	<b>2</b>
3.1 Contents of the GrADS descriptor (control) file . . . . .	3
3.2 List of available quantities . . . . .	4
<b>4 Description of daily South American data sets</b>	<b>7</b>
4.1 Contents of the GrADS descriptor (control) file . . . . .	7
4.2 List of available quantities . . . . .	8
<b>5 Description of monthly South American data sets</b>	<b>9</b>
5.1 Contents of the GrADS descriptor (control) file . . . . .	9
5.2 List of available quantities . . . . .	10
<b>6 Electronic access/ordering information</b>	<b>12</b>
<b>7 Files available on NASA/GSFC Unitree System</b>	<b>14</b>
<b>8 Acknowledgments</b>	<b>14</b>
<b>Appendix: GrADS data sets</b>	<b>15</b>
<b>References</b>	<b>24</b>

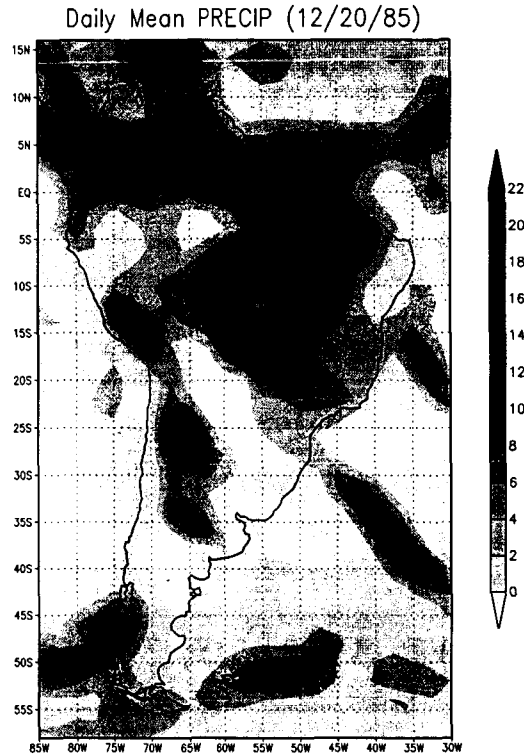


Figure 1: Daily mean precipitation rate (mm/day) for December 20, 1985. This figure shows the geographic domain of the South American subset described in this document.

## 1 Overview

The Data Assimilation Office (DAO) at Goddard Space Flight Center is currently producing a multi-year control assimilation (currently 1985-93) with version 1 of the Goddard EOS Data Assimilation System (GEOS-1/DAS). One of the main goals of this project, in addition to bench-marking our system, is to produce a research quality data set suitable for Earth Science applications. By making the data available to the general scientific community, we hope to learn more about the deficiencies in our data assimilation system, the usefulness of various diagnostic quantities, and the required resolution (in both space and time) which will satisfy the majority of researchers. We realize that certain problems will require specialized data sets, and we have built our system in a way that allows a large degree of flexibility in how the data are saved.

An overview of the GEOS-1 Multi-year Assimilation project can be found in Schubert *et al.* (1993). The Atmospheric General Circulation Model used in the assimilation system is documented in Takacs *et al.* (1994) and Suarez and Takacs (1995). The statistical analysis portion of GEOS-1/DAS is described in Pfaendtner *et al.* (1995), while Schubert *et al.* (1995a) provide an overview of the main results. The proceedings of the *Workshop on the GEOS-1 Five-year Assimilation* (Schubert *et al.* 1995) provides a good summary of the strengths and weaknesses of the data set. Most of these documents are available on-line from DAO's home page (<http://dao.gsfc.nasa.gov>).

This document describes a subset of the pressure level data sets in time-series format obtained from the multi-year assimilation with the GEOS-1 DAS (currently March 1985 to

November 1993). The data sets include selected prognostic and diagnostic quantities (both surface and upper air) from the full assimilated data set, restricted to the South American region (see Figure 1). The data are given at a variety of temporal resolutions: 3-hourly, 6-hourly, daily and monthly means, as described below.

## 2 Contents of the data sets

The data are stored in a GrADS compatible format. These are IEEE compliant, 32 bit *big endian*<sup>1</sup> floating point files WITHOUT FORTRAN control words at the end of each record. (Therefore, if using FORTRAN you must open these files as direct access.) See the Appendix for additional information on the file format.

Each data set consists of a series of records. Each record contains one level of a field. The field at one level is read, then the next, and so on. Upper air fields (both prognostic and diagnostic) are available at 18 or 6 levels; see the GrADS control files in subsequent sections for specific information about the pressure levels, the resolution, the date and time stamp, etc. See the Appendix for additional information on how to read a standard GrADS file.

The upper air data sets were obtained by interpolating assimilated data from 20 sigma levels to 18 pressure levels (14 mandatory levels given by 1000 hPa, 850 hPa, 700 hPa, 500 hPa, 400 hPa, 300 hPa, 250 hPa, 200 hPa, 150 hPa, 100 hPa, 70 hPa, 50 hPa, 30 hPa, and 20 hPa plus 4 additional levels given by 950 hPa, 900 hPa, 800 hPa and 600 hPa).

The upper air fields on pressure levels are read with the bottom level (highest pressure) first, i.e.

```
u wind at 1000 hPa
u wind at 950 hPa
.
.
.
u wind at 30 hPa
u wind at 20 hPa
```

## 3 Description of 3- and 6-hourly South American data sets

These data sets (as well as this section) are derived from the global time-series files prepared by Wayne Higgins. The file name convention is as follows:

---

<sup>1</sup>Since the IEEE standard does not specify the particular byte ordering of float point data, there exists variation among vendors. Most Unix platforms, including Sun, SGI, HP and Cray, use the *big endian* byte ordering. Platforms with *little endian* byte ordering include x86 systems (MS-DOS, OS/2, etc.) and Digital Equipment Corporation (DEC) workstations. When using GrADS (version 1.4 or later) the user does not need to worry about byte ordering as the software handles it internally. However, when reading these files on a *little endian* platform with a high level language such as FORTRAN or C the user will need to either 1) instruct the compiler to automatically perform *big endian* I/O (recent versions of DEC compilers allow this), or 2) manually perform a byte swap.

```
NAME.sa.prs.byymmdd (data file)
NAME.sa.ctl.byymmdd (GrADS control file)
```

where

**NAME:** name of the desired upper level field (see Tables 1-6, subsection 3.2) or name given to the surface prognostic/ diagnostic data set. The surface prognostic and diagnostic data sets have multiple fields. The naming convention for the surface prognostics is *sfcprog*. The naming convention for the diagnostics is *diag $n$* , where  $n=1,2,3$  or  $4$ ; see Tables 3-6 for a list of available quantities.

**sa:** stands for South America.

**byymmdd:** beginning year, month and date (e.g. 850301)

For example, the file with the zonal wind for March of 1985 are named:

```
uwnd.sa.prs.b850301 (data file)
uwnd.sa.ctl.b850301 (GrADS control file)
```

### 3.1 Contents of the GrADS descriptor (control) file

The GrADS control file contains all information necessary to read a given data set (see Appendix). Each data set has an associated GrADS control file. The following is an example of a control file for a single upper air prognostic quantity (control file name: *tmpu.sa.ctl.b920201*):

```
DSET tmpu.sa.prs.b920201
OPTIONS big_endian
TITLE OI1.3.2/VC5.2 Geos-1.0 Version
UNDEF 0.100000E+16
XDEF 23 LINEAR -85.0 2.500
YDEF 38 LINEAR -58.0 2.000
ZDEF 18 LEVELS 1000 950 900 850 800 700 600 500 400 300 250
      200 150 100 70 50 30 20
TDEF 116 LINEAR 00:00Z01FEB92 6hr
VARS 1
TMPU 18 0 TEMPERATURE (K)
ENDVARS
```

In this example, the first line contains the name of the GrADS data file, while the second line indicates that this is an IEEE *big endian* binary file (see section 2). Next, the title of the experiment is given followed by the undefined, or missing, value  $1.0E+15$ . There are 23 grid points in the zonal direction (i.e. longitude) with the first grid point at 85 West and with a grid spacing of 2.5 degrees. There are 38 grid points in the meridional direction (i.e. latitude) with the first grid point at 58 South and with a grid spacing of 2 degrees. There are 18 pressure levels with the bottom level (highest pressure) first. Data were written 116

times starting at 00Z on February 1, 1992 and every six hours thereafter. There is 1 variable in this data set, namely the temperature field (in degrees Kelvin).

The upper-air time series given in Table 2b are derived from similar time series prepared by the Goddard DAAC (<http://daac.gsfc.nasa.gov>) and contain only 6 vertical levels. A typical control file for these data is

```
DSET tmpu.sa.prs.b920201
OPTIONS big_endian
TITLE OI1.3.2/VC5.2 Geos-1.0 Version
UNDEF 0.100000E+16
XDEF 23 LINEAR -85.0 2.500
YDEF 38 LINEAR -58.0 2.000
ZDEF 6 LEVELS 1000.0 850.0 700.0 500.0 300.0 200.0
TDEF 116 LINEAR 00:00Z01FEB92 6hr
VARS 1
TMPU 6 0 TEMPERATURE (K)
ENDVARS
```

### 3.2 List of available quantities

Tables 1-2 list the surface and upper prognostic quantities which are given every 6 hours; these correspond to instantaneous values of the prognostic fields at the synoptic time. The surface (and top of the atmosphere) diagnostic quantities (Tables 3-6) are given every 3 hours over the South American continent. The 3-hourly diagnostics are averages from the previous 3 hours. For example, the 3-hour average evaporation valid at 12Z corresponds to a 3-hour average from 9Z to 12Z.

Table 1: Surface prognostic quantities (sfcprog), every 6 hours.

```
----- File names:
          sfcprog.sa.ctl.byymmdd (GrADS control file)
          sfcprog.sa.prs.byymmdd (data file)
Data file size: ~4 Mbytes/month
```

NAME	LEVELS	DESCRIPTION
phis	1	SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
albd	1	SURFACE ALBEDO (0-1)
gwet	1	SURFACE GROUND WETNESS (0-1)
ps-ptop	1	SURFACE PRESSURE - P TOP (hPa)
gtmp	1	SURFACE GROUND TEMPERATURE (K)
slp	1	SEA LEVEL PRESSURE (hPa)
lwi	1	WATER_(1) LAND_(2) PERMANENT ICE_(3) SEA ICE_(4) FLAGS
ubar	1	VERTICALLY INTEGRATED (BAROTROPIC) UWND (M/S)
vbar	1	VERTICALLY INTEGRATED (BAROTROPIC) VWND (M/S)

Table 2a: Upper air prognostic quantities, every 6 hours. File names are:

----- NAME.sa.ctl.byymmdd (GrADS control file)  
 NAME.sa.prs.byymmdd (data file)  
 where NAME is given in the first column.  
 Data file size: ~8 Mbytes/month

NAME	LEVELS	DESCRIPTION
uwnd	18	U-WIND (M/S)
vwnd	18	V-WIND (M/S)
hght	18	GEOPOTENTIAL HEIGHT (M)
tmpu	18	TEMPERATURE (K)
sphu	18	SPECIFIC HUMIDITY (G/KG)
omega(*)	18	VERTICAL VELOCITY (hPa/DAY)

(\*) Actually a diagnostic quantity, included here for convenience.  
 Notice that the vertical velocity fields correspond to 6-hour averages centered at the synoptic time, while the other prognostic quantities are instantaneous snapshots at the synoptic time.

Table 2b: Upper air prognostic quantities, every 6 hours. File names are:

----- NAME.sa.ctl.yymmdd (GrADS control file)  
 NAME.sa.prs.yymmdd (data file)  
 where NAME is given in the first column. These files are identical to the files in Table 2a, except for the reduced number of vertical levels. These files were derived from similar time series prepared by the Goddard DAAC (<http://daac.gsfc.nasa.gov>).  
 Data file size: ~3 Mbytes/month

NAME	LEVELS	DESCRIPTION
uwnd	6	U-WIND (M/S)
vwnd	6	V-WIND (M/S)
hght	6	GEOPOTENTIAL HEIGHT (M)
tmpu	6	TEMPERATURE (K)
sphu	6	SPECIFIC HUMIDITY (G/KG)

Table 3: Surface diagnostic quantities associated with the hydrological cycle (diag1), every 3 hours.

----- File names:  
 diag1.sa.ctl.byymmdd (GrADS control file)  
 diag1.sa.prs.byymmdd (data file)  
 Data file size: ~8 Mbytes/month

NAME	LEVELS	DESCRIPTION
------	--------	-------------



ps-ptop	1	SURFACE PRESSURE - PTOP (hPa)
preacc	1	TOTAL PRECIPITATION (MM/DAY)
precon	1	CONVECTIVE PRECIPITATION (MM/DAY)
evap	1	SURFACE EVAPORATION (MM/DAY)
vintuq	1	VERTICALLY INTEGRATED U*Q (M/S G/KG)
vintvq	1	VERTICALLY INTEGRATED V*Q (M/S G/KG)
vintut	1	VERTICALLY INTEGRATED U*T (M/S DEG)
vintvt	1	VERTICALLY INTEGRATED V*T (M/S DEG)
qint	1	PRECIPITABLE WATER (GM/CM**2)

Table 4: Surface diagnostic quantities associated with surface momentum and heat fluxes (diag2), every 3 hours. File names:  
diag2.sactl.byymdd (GrADS control file)  
diag2.sa.prs.byymdd (data file)  
Data file size: ~8 Mbytes/month

NAME	LEVELS	DESCRIPTION
ps-ptop	1	SURFACE PRESSURE - PTOP (hPa)
uflux	1	U-MOMENTUM SURFACE STRESS (N/M**2)
vflux	1	V-MOMENTUM SURFACE STRESS (N/M**2)
hflux	1	SURFACE FLUX OF SENSIBLE HEAT (W/M**2)
ct	1	SURFACE DRAG COEFFICIENT FOR T AND Q (M/S)
cu	1	SURFACE DRAG COEFFICIENT FOR U AND V (M/S)
winds	1	SURFACE WIND SPEED (M/S)
ustar	1	USTAR (M/S)
z0	1	SURFACE ROUGHNESS (M)
pbl	1	PBL DEPTH (hPa)

Table 5: Diagnostic quantities associated with surface and top of the atmosphere radiational fluxes (diag3), every 3 hours. File names:  
diag3.sactl.byymdd (GrADS control file)  
diag3.sa.prs.byymdd (data file)  
Data file size: ~9 Mbytes/month

NAME	LEVELS	DESCRIPTION
ps-ptop	1	SURFACE PRESSURE - PTOP (hPa)
radlwg	1	NET UPWARD LW RADIATION AT GROUND (W/M**2)
radswg	1	NET DOWNWARD SW RADIATION AT GROUND (W/M**2)
olr	1	OUTGOING LONGWAVE RADIATION (W/M**2)
olrclr	1	OUTGOING LONGWAVE RADIATION CLEAR SKY (W/M**2)
lwgclr	1	SURFACE LONGWAVE FLUX CLEAR SKY (W/M**2)
radswt	1	INCIDENT SW RADIATION AT TOP OF ATMOS (W/M**2)

```

osr          1          OUTGOING SW RADIATION (W/M**2)
osrclr       1          OUTGOING SHORTWAVE RADIATION CLEAR SKY (W/M**2)
swgclr       1          SURFACE SHORTWAVE FLUX CLEAR SKY (W/M**2)
cldfrc       1          2-DIMENSIONAL TOTAL CLOUD FRACTION (0-1)

```

---

Table 6: Miscellaneous surface diagnostic quantities (diag4),  
----- every 3 hours. File names  
          diag4.sa.ctl.byymmdd (GrADS control file)  
          diag4.sa.prs.byymmdd (data file)  
Data file size: ~11 Mbytes/month

---

NAME	LEVELS	DESCRIPTION
ps-ptop	1	SURFACE PRESSURE - PTOP (hPa)
tg	1	GROUND TEMPERATURE (DEG)
ts	1	SURFACE AIR TEMPERATURE (DEG)
qs	1	SATURATION SURFACE SPECIFIC HUMIDITY (G/KG)
dpdt	1	SURFACE PRESSURE TENDENCY (hPa/DAY)
u2m	1	U AT 2 METERS (M/S)
v2m	1	V AT 2 METERS (M/S)
t2m	1	T AT 2 METERS (DEG)
q2m	1	Q AT 2 METERS (KG/KG)
u10m	1	U AT 10 METERS (M/S)
v10m	1	V AT 10 METERS (M/S)
t10m	1	T AT 10 METERS (DEG)
q10m	1	Q AT 10 METERS (KG/KG)

---

## 4 Description of daily South American data sets

These time series files from the Multi-year GEOS-1 Assimilation consist of daily means for the full length of the assimilation (currently 9 years, from 1985 to 1993) for the South American region. The data offered here represent a selection of popular quantities and levels. The quantities available are given in Table 7.

### 4.1 Contents of the GrADS descriptor (control) file

The GrADS control file contains all information necessary to read a given data set (see Appendix). Each data set has an associated GrADS control file. The following is an example of a control file for the daily mean time series files for the South American region (z500.sa.ctl):

```

DSET ^z500.sa.dat
OPTIONS big_endian
TITLE Daily mean 500 hPa height from e0054A

```

```

UNDEF 1.0E+15
XDEF 23 LINEAR -85 2.5
YDEF 38 LINEAR -58 2
ZDEF 1 LEVELS 1 500
TDEF 3197 LINEAR 0z01MAR1985 24hr
VARS 1
hght 0 0 500 hPa HEIGHT (M)
ENDVARS

```

In this example, the first line contains the name of the GrADS data file, while the second line indicates that this is an IEEE *big endian* binary file (see section 2). Next, the title of the experiment is given followed by the undefined, or missing, value 1.0E+15. There are 23 grid points in the zonal direction (i.e. longitude) with the first grid point at 85 West and with a grid spacing of 2.5 degrees. There are 38 grid points in the meridional direction (i.e. latitude) with the first grid point at 58 South and with a grid spacing of 2 degrees. All daily mean files have a single vertical level, in this example 500 hPa. Data were written 3197 times starting at 00Z on March 1, 1985 and every 24 hours thereafter. There is 1 variable in this file, namely the 500 hPa geopotential heights (in meters).

## 4.2 List of available quantities

The quantities currently available are listed in Table 7, along with the GrADS data file name.

Table 7: Daily mean time series at selected levels for South America.  
----- The GrADS control files have extension "ctl", e.g., the data file "evap.sa.dat" has a control file named "evap.sa.ctl".  
Data file size: ~11 Mbytes/9 years.

NAME	FILE NAME	DESCRIPTION
evap	evap.sa.dat	surface evaporation
preacc	preacc.sa.dat	total precipitation
precon	precon.sa.dat	convective precipitation
qint	qint.sa.dat	total precipitable water
slp	slp.sa.dat	sea level pressure
stf200	stf200.sa.dat	stream function at 200 hPa
stf850	stf850.sa.dat	stream function at 850 hPa
t700	t700.sa.dat	temperature at 700 hPa
u200	u200.sa.dat	u-wind at 200 hPa
u850	u850.sa.dat	u-wind at 850 hPa
v200	v200.sa.dat	v-wind at 200 hPa
v850	v850.sa.dat	v-wind at 850 hPa
velp200	velp200.sa.dat	velocity potential at 200 hPa
velp850	velp850.sa.dat	velocity potential at 850 hPa
vintuq	vintuq.sa.dat	vertically integrated u*q
vintvq	vintvq.sa.dat	vertically integrated v*q
z500	z500.sa.dat	500 hPa height

-----  
These data are in GrADS format, one level per file.

## 5 Description of monthly South American data sets

There are 3 groups of files with monthly means for the South American region. Their contents are listed in Tables 9-11. The monthly mean prognostic file (Table 9) also contains the (diagnostic) vertical velocity and several second-order fluxes<sup>2</sup> (*e.g.*,  $uv$ ,  $uT$ , etc.). The surface and top of the atmosphere diagnostic file (Table 10) contains several fluxes of interest. Finally, the monthly mean analysis increments (Table 11) are essential quantities for any budget study from data assimilation systems (see section 2 of Schubert et al., 1995a).

### 5.1 Contents of the GrADS descriptor (control) file

The GrADS control file contains all information necessary to read a given data set (see Appendix). Each data set has an associated GrADS control file. The following is an example of a control file for the monthly mean time series files for the South American region (`tmpu.sa.ct1.dat`):

```
DSET ^tmpu.sa.mm.dat
options big_endian
TITLE Reassimilation of 1985-1989 --- OI1.3/VC5B20 Geos-0.1 Version
UNDEF 0.100000E+16
XDEF 23 LINEAR -85.0 2.500
YDEF 38 LINEAR -58.0 2.000
ZDEF 18 LEVELS 1000.00 950.00 900.00 850.00 800.00
          700.00 600.00 500.00 400.00 300.00
          250.00 200.00 150.00 100.00 70.00
          50.00 30.00 20.00
TDEF 105 LINEAR 18:00Z01MAR85 1mo
VARS 1
TMPU 18 0 TEMPERATURE (K)
ENDVARS
```

In this example, the first line contains the name of the GrADS data file, while the second line indicates that this is an IEEE *big endian* binary file (see section 2). Next, the title of the experiment is given followed by the undefined, or missing, value  $1.0E+15$ . There are 23 grid points in the zonal direction (*i.e.* longitude) with the first grid point at 85 West and

<sup>2</sup>Notice that the second order fluxes are *not* eddy fluxes. For example, if  $\bar{u}$  and  $\bar{v}$  represent the monthly mean zonal and meridional wind components, and  $u'$  and  $v'$  represent departures from the monthly mean, then the quantity  $uv$  listed in Table 9 is the monthly mean

$$\overline{uv} = \bar{u} \cdot \bar{v} + \overline{u'v'}$$

In order to obtain the eddy fluxes  $\overline{u'v'}$  you will need to subtract the quantity  $\bar{u} \cdot \bar{v}$  from  $\overline{uv}$ .

with a grid spacing of 2.5 degrees. There are 38 grid points in the meridional direction (i.e. latitude) with the first grid point at 58 South and with a grid spacing of 2 degrees. There are 38 grid points in the meridional direction (i.e. latitude) with the first grid point at 58 South and with a grid spacing of 2 degrees. Data were written 105 times starting at March 1985 and every month thereafter. There is 1 variable in this file, namely the temperature field (in degrees Kelvin).

## 5.2 List of available quantities

The quantities currently available are listed in Tables 9-11. The corresponding file names and sizes are indicated in the caption of each table. The monthly mean prognostic quantities are stored with 1 variable per file (Table 9), while the surface/top of the atmosphere diagnostics are stored with several variables per file (see Tables 10-11).

Table 9: Monthly mean prognostic (and some diagnostic) for the  
 ----- South American region. The file names are:  
           NAME.sa.mm.ct1 (GrADS control file)  
           NAME.sa.mm.data (GrADS data file)  
 where NAME is given in the first column.  
 The upper air data are given in 18 pressure levels.  
 Data file size: ~7 Mbytes/9 years.

NAME	LEVELS	DESCRIPTION
phis	1	SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
albd	1	SURFACE ALBEDO (0-1)
gwet	1	SURFACE GROUND WETNESS (0-1)
psptop	1	SURFACE PRESSURE - P <sub>TOP</sub> (hPa)
gtmp	1	SURFACE GROUND TEMPERATURE (K)
slp	1	SEA LEVEL PRESSURE (hPa)
lwi	1	LAND_(2), WATER_(1), ICE_(3) FLAGS
ubar	1	VERTICALLY INTEGRATED (BAROTROPIC) UWND (M/S)
vbar	1	VERTICALLY INTEGRATED (BAROTROPIC) VWND (M/S)
uwnd	18	U-WIND (M/S)
vwnd	18	V-WIND (M/S)
hght	18	GEOPOTENTIAL HEIGHT (VIRTUAL) (M)
tmpu	18	TEMPERATURE (K)
sphu	18	SPECIFIC HUMIDITY (G/KG)
qq	18	TURBULENT KINETIC ENERGY (M/S)**2
omega	18	OMEGA VERTICAL VELOCITY (hPa/DAY)
uu	18	UWND * UWND (M/S)**2
vv	18	VWND * VWND (M/S)**2
tt	18	TMPU * TMPU (K**2)
zz	18	HGHT * HGHT (M**2)
mm	18	SPHU * SPHU (G/KG)**2
uv	18	UWND * VWND (M/S)**2
ut	18	UWND * TMPU (M*K)
vt	18	VWND * TMPU (M*K)
uz	18	UWND * HGHT (M**2/S)

vz	18	VWND * HGHT	(M**2/S)
uq	18	UWND * SPHU	(M/S)*(G/KG)
vq	18	VWND * SPHU	(M/S)*(G/KG)
ww	18	OMEGA*OMEGA	(hPa/DAY)**2
wt	18	OMEGA*TMPU	(hPa/DAY)*K
wq	18	OMEGA*SPHU	(hPa/DAY)*(G/KG)

---

Table 10: Monthly mean surface and top of the atmosphere diagnostics  
 ----- for the South American region. The file names are:

diag.sfc.sa.mm.ct1 (GrADS control file)  
 diag.sfc.sa.mm.dat (GrADS data file)  
 Data file size: ~15 Mbytes/9 years.

---

NAME	LEVELS	DESCRIPTION
phis	1	SURFACE GEOPOTENTIAL HEIGHTS (M/S)**2
psptop	1	SURFACE PRESSURE - P TOP (hPa)
uflux	1	U-MOMENTUM SURFACE STRESS (N/M**2)
vflux	1	V-MOMENTUM SURFACE STRESS (N/M**2)
hflux	1	SRFCE FLX OF SENSIBLE HEAT (W/M**2)
radlwg	1	NET UPWARD LW RAD. AT GRND (W/M**2)
radswg	1	NET DOWNWARD SW RAD. AT GRND (W/M**2)
ct	1	SURFACE DRAG COEF. FOR T AND Q (DIMENSIONLESS)
cu	1	SURFACE DRAG COEF. FOR U AND V (DIMENSIONLESS)
preacc	1	TOTAL PRECIPITATION (MM/DAY)
precon	1	CONVECTIVE PRECIPITATION (MM/DAY)
winds	1	SURFACE WIND SPEED (M/S)
tg	1	GROUND TEMPERATURE (DEG)
ts	1	SURFACE AIR TEMPERATURE (DEG)
qs	1	SATURATION SURFACE SPEC. HUM. (G/KG)
olr	1	OUTGOING LONGWAVE RADIATION (W/M**2)
olrclr	1	OUTGOING LONGWAVE RAD CLEAR SKY (W/M**2)
lwgclr	1	SURFACE LONGWAVE FLUX CLEAR SKY (W/M**2)
radswt	1	INCIDENT SW RAD. AT TOP OF ATM. (W/M**2)
evap	1	SURFACE EVAPORATION (MM/DAY)
dpdt	1	SURFACE PRESSURE TENDENCY (hPa/DAY)
ustar	1	USTAR (M/SEC)
z0	1	SURFACE ROUGHNESS Z0 (M)
pbl	1	PBL DEPTH (hPa)
osr	1	OUTGOING SHORTWAVE RADIATION (W/M**2)
osrclr	1	OUTGOING SHORTWAVE RAD CLEAR (W/M**2)
swgclr	1	SURF. SHORTWAVE FLUX CLEAR SKY (W/M**2)
vintuq	1	VERTICALLY INTEGRATED U*Q (M/SEC G/KG)
vintvq	1	VERTICALLY INTEGRATED V*Q (M/SEC G/KG)
vintut	1	VERTICALLY INTEGRATED U*T (M/SEC DEG)
vintvt	1	VERTICALLY INTEGRATED V*T (M/SEC DEG)
cldfrc	1	2-DIMENSIONAL TOTAL CLOUD FRACTION (0-1)
qint	1	PRECIPITABLE WATER (GM/CM**2)
u2m	1	U AT 2 METERS (M/SEC)
v2m	1	V AT 2 METERS (M/SEC)
t2m	1	T AT 2 METERS (DEG)

q2m	1	Q AT 2 METERS (KG/KG)
u10m	1	U AT 10 METERS (M/SEC)
v10m	1	V AT 10 METERS (M/SEC)
t10m	1	T AT 10 METERS (DEG)
q10m	1	Q AT 10 METERS (KG/KG)

---

Table 11: Monthly mean analysis increments for the South American region given at 20 sigma levels. The file names are:  
 ----- NAME.sa.mm.ct1 (GrADS control file)  
 NAME.sa.mm.dat (GrADS data file)  
 where NAME is listed in the first column. See NOTES below.  
 Data file size: ~7 Mbytes/9 years.

---

NAME	LEVELS	DESCRIPTION
analp	1	ANALYSIS INCREMENT OF PSURF (hPa/S)
analu	20	ANALYSIS INCREMENT OF U-WIND (M/S/S)
analv	20	ANALYSIS INCREMENT OF V-WIND (M/S/S)
analt	20	ANALYSIS INCREMENT OF THETA (hPa*K)/(hPa**Kappa)/S
analq	20	ANALYSIS INCREMENT OF SPHU (hPa*KG/KG)/S

---

#### NOTES:

1. The GEOS GCM “temperature” variable  $\Theta$  is defined by

$$\Theta \equiv (p_s - p_{top})T p^{-\kappa}$$

where  $T$  is the temperature in Kelvin, and  $p$  stands for pressure in hPa with  $p_s$  and  $p_{top}$  being its value at the surface and at the top of the model (equal to 10 hPa in the GEOS-1 Assimilation run);  $\kappa = R/C_p \approx 0.287$ .

2. Similarly, the “moisture” variable  $Q$  in the GEOS GCM is

$$Q \equiv (p_s - p_{top})q$$

where  $q$  is the specific humidity in kg/kg. See page 47 of Takacs *et al.* (1994) for details.

3. Generally, all quantities are saved on an Arakawa A-grid. The exceptions are the analysis increments in Table 11 which are saved on the GEOS-1 GCM’s C-grid (Takacs *et al.* 1994). Therefore, the `analv` and `analq` fields are not on the same grid as the monthly mean `uwnd` and `vwnd` fields listed in Table 9.

## 6 Electronic access/ordering information

Several files reside on an anonymous ftp account on a local workstation in the DAO to allow outside users to get information about 1) selected monthly means and time series, 2)

statistics about the observations going into the assimilation, 3) the latest documentation of the data sets (this document), 4) a detailed description of the diagnostics, and 5) known problems with the data. There is also a special file to allow users of the data to document any problems they have found with the data.

The anonymous ftp site is on `dao.gsfc.nasa.gov`, and some of the South American files can be found at

```
/pub/assimilation/e0054A/data/time_series/south_america/3-hourly/  
6-hourly/  
daily/  
monthly/
```

The daily and monthly data sets will reside on this workstation. However, the 3- and 6-hourly files will be made available upon request. Please send your request by e-mail to `data@dao.gsfc.nasa.gov`.

Other information concerning the full GEOS-1 data set can be found at the directories below.

```
/pub/assimilation/e0054A/documentation/
```

```
/pub/assimilation/e0054A/messages
```

This is a place for you to document any problems you have discovered with the data. We are very interested in feedback from the community. You can also send e-mail to `data@dao.gsfc.nasa.gov` (this assures a quicker response).

```
/pub/assimilation/e0054A/observations/
```

```
bar_obscountyymm.ps
```

These are postscript files (bar graphs of data counts)  
There are also text files summarizing the data going  
into the assimilation

```
/pub/assimilation/e0054A/data
```

```
/monthly_means  
/time_series
```

These directories contain selected GLOBAL monthly  
means and time series in GrADS format available  
via anonymous ftp

To get further information about the data please direct your questions to:



E-mail: [data@dao.gsfc.nasa.gov](mailto:data@dao.gsfc.nasa.gov)

You can also visit DAO's home page at

WWW: <http://dao.gsfc.nasa.gov/>

or the Goddard DAAC home page

WWW: <http://daac.gsfc.nasa.gov/>

For specific questions regarding the South American subset write to Arlindo da Silva ([dasilva@gsfc.nasa.gov](mailto:dasilva@gsfc.nasa.gov)).

## 7 Files available on NASA/GSFC Unitree System

Users with access to NASA/GSFC Unitree System ([dirac.gsfc.nasa.gov](http://dirac.gsfc.nasa.gov)) can find the South American subset in the following directories:

```
/u4/yfamd/geos/south_america/3-hourly/  
    .6-hourly/1-level  
        /6-levels  
        /18-levels  
    daily/  
    monthly/
```

In case of difficulties please contact [data@dao.gsfc.nasa.gov](mailto:data@dao.gsfc.nasa.gov).

## 8 Acknowledgments

A formal technical review of this document was conducted on February 13, 1996 at the data Assimilation Office. I would like to thank Ken Ekers (review leader), Siegfried Schubert and Chung Kyu Park (reviewers) for valuable suggestions. I would like also to thank Dr. Carlos Nobre from the Center for Weather Prediction and Climate Studies (CPTEC/INPE) in Brazil for encouraging the preparation of this data subset.

## Appendix: GrADS data sets

The Grid Analysis and Display System (GrADS) is an interactive desktop tool that is currently in use worldwide for the analysis and display of earth science data. GrADS is implemented on all commonly available UNIX workstations and DOS based PCs, and is freely distributed over the Internet via anonymous ftp from either of these sites

```
ftp://grads.iges.org/  
ftp://sprite.llnl.gov/pub/fiorino/grads/
```

For information on GrADS visit the GrADS home page at

```
http://grads.iges.org/grads/head.html
```

This appendix is reproduced from an earlier version of the GrADS manual by Brian Doty, with permission. It is intended to provide a detailed documentation of the format used to create the data sets described in this document.

### Gridded Data Sets

The GrADS gridded data set is a direct access binary data set. It may contain any number of variables at specified longitude, latitude, vertical, and time intervals.

GrADS views this data set as a giant array -- with X (longitude) varying the fastest, then Y (latitude), then Z (vertical level), then the variable type, then T (time).

It is easier for us to think of the data set in terms of a sequence of horizontal grids, where longitude and latitude vary. Each horizontal grid represents a particular variable at a particular height and time. Each horizontal grid is the same size in any particular GrADS data set (if you have grids of different sizes, you must create separate data sets).

These grids are written to the data set in the following order: starting with a particular variable, grids for each vertical level (at a particular time) are written out in ascending order. Then the grids for the next variable are written out. When all the grids at a particular time have been written, grids for the next time are written.

The format of this data set is thus exactly the same as the COLA Pressure History format, except: there are no date/time records, and latitude varies from south to north (not north to south as in the pressure history data).

Each binary gridded data set is described by a data descriptor file, essentially a table of contents for the binary data set. Following is

an example of such a file:

```
DSET      ua.dat
TITLE    Upper Air Data
UNDEF    -9.99E33
XDEF     80 LINEAR  -140.0  1.0
YDEF     50 LINEAR   20.0  1.0
ZDEF     10 LEVELS  1000 850 700 500 400 300 250 200 150 100
TDEF     4 LINEAR  OZ10apr1991 12hr
VARS     5
slp      0      0 sea level pressure
z        10     0 heights
t        10     0 temps
td       6      0 dewpoints
u        10     0 u winds
v        10     0 v winds
ENDVARS
```

The data descriptor file is 'free format', ie each entry is blank delimited and may appear in any column. Comment records start with an asterisk ('\*') in column 1. Comments may not appear in the list of variable records (between the vars and endvars records). Records may not be more than 80 characters long.

In this example, the binary data set is named ua.dat, the undefined, or missing, data value is -9.99e33, there are 80 grid points in the X direction, 50 in the Y direction, 10 levels, 4 times, and 5 variables. The variables z, t, u, and v have 10 levels, the variable td has 6 levels, and the variable slp has one level (see below for a more specific description of each entry).

Think in terms of the X and Y data points at one level for one variable at one time being a horizontal grid. This grid is exactly in the same storage order as a FORTRAN array, in this case an array DIMENSION A(80,50). The first dimension always varies from west to east, the second from south to north.

In the above example the horizontal grids would be written in the following order:

```
Time 1, Level  ?, Variable slp
Time 1, Level 1000, Variable z
Time 1, Level 850, Variable z
      then levels 700, 500, 400, 300, 250, 200, then
Time 1, Level 150, Variable z
Time 1, Level 100, Variable z
Time 1, Level 1000, Variable t
Time 1, Level 850, Variable t
      then levels 700, 500, 400, 300, 250, 200, then
Time 1, Level 150, Variable t
Time 1, Level 100, Variable t
Time 1, Level 1000, Variable td
Time 1, Level 850, Variable td
```

```

Time 1, Level 700, Variable td
Time 1, Level 500, Variable td
Time 1, Level 400, Variable td
Time 1, Level 300, Variable td
Time 1, Level 1000, Variable u
Time 1, Level 850, Variable u
      then levels 700, 500, 400, 300, 250, 200, then
Time 1, Level 150, Variable u
Time 1, Level 100, Variable u
Time 1, Level 1000, Variable v
Time 1, Level 850, Variable v
      then levels 700, 500, 400, 300, 250, 200, then
Time 1, Level 150, Variable v
Time 1, Level 100, Variable v
Time 2, Level  ?, Variable slp
Time 2, Level 1000, Variable z
Time 2, Level 850, Variable z
Time 2, Level 700, Variable z
Time 2, Level 500, Variable z
Time 2, Level 400, Variable z
.
.
.
etc

```

A description of each record in the GrADS data descriptor file follows:

**DSET data-set-name**

This entry specifies the name of the binary data set. It may be entered in mixed case.

If the binary data set is in the same directory as the data descriptor file, you may enter the filename in the data descriptor file without a full path name by prefixing it with a ^ character. For example, if the data descriptor file is:

```
/data/wx/grads/sa.ct1
```

and the binary data file is:

```
/data/wx/grads/sa.dat
```

you could use the following file name in the data descriptor file:

```
DSET ^sa.dat
```

instead of:

```
DSET /data/wx/grads/sa.dat
```

As long as you keep the two files together, you may move them to any directory without changing the entries in the data descriptor

file.

#### TITLE string

A brief description of the contents of the data set. This will be displayed during a QUERY command, so it is helpful to put meaningful information here.

#### UNDEF value

The undefined, or missing, data value. GrADS operations and graphics routines will ignore data with this value from this data set.

#### BYTESWAPPED

Indicates the binary data file is in reverse byte order from the normal byte order of the machine. This would happen if you sent a file in binary format from, for example, a Sun to a PC. Putting this keyword in the descriptor file tells GrADS to swap the byte order as the data is being read.

XDEF number <LINEAR start increment>  
<LEVELS value-list>

Defines the mapping between grid values and longitude. Specifically:

number -- the number of grid values in the X direction, specified as an integer number.  
Must be  $\geq 1$ .

LINEAR or LEVELS -- Indicates the grid mapping type.

For LINEAR:

start -- the starting longitude, or the longitude for  $X = 1$ . Specified as a floating point value, where negative indicates degrees west.

increment -- the spacing between grid value in the X direction. It is assumed that the X dimension values go from west to east. Specified as a positive floating value.

For LEVELS:

value-list -- List of 'number' values representing the longitude of each X dimension. May start and continue on the next record in the descriptor file (records may not be  $> 80$  characters). There must be at least 2 levels (otherwise use LINEAR mapping).

YDEF number mapping start <increment>  
<LEVELS value-list>

Defines the mapping between grid values and latitude.  
Specifically:

number -- the number of grid values in the X direction,  
specified as an integer number.  
mapping -- mapping type, specified as a keyword.  
Valid are:  
    LINEAR -- Linear mapping  
    GAUSR15 -- Gaussian R15 latitudes  
    GAUSR20 -- Gaussian R20 latitudes  
    GAUSR30 -- Gaussian R30 latitudes  
    GAUSR40 -- Gaussian R40 latitudes  
start -- For LINEAR mapping, the starting latitude, ie the  
latitude for Y = 1, and is specified as a floating  
point value, with negative indicating degrees south.  
For GAUSRxx mapping, the start value indicates  
the first gaussian grid number, where 1 would be the  
southernmost gaussian grid latitude.  
increment -- the spacing between grid values in the Y direction.  
It is assumed that the Y dimension values go from  
south to north. Specified as a positive floating  
point value. Used only for LINEAR mapping.

For LEVELS:

value-list -- List of 'number' values representing the  
latitude of each X dimension. May start and  
continue on the next record in the descriptor  
file (records may not be > 80 characters).  
There must be at least 2 levels (otherwise  
use LINEAR mapping).

Examples of specifying GAUSRxx mapping:

YDEF 20 GAUSR40 15

Indicates that there are 20 Y dimension values which start  
at Gaussian Latitude 15 (64.10 south) on the Gaussian R40 grid.  
Thus the 20 values would correspond to Latitudes:

-64.10, -62.34, -60.58, -58.83, -57.07, -55.32, -53.56,  
-51.80, -50.05, -48.29, -46.54, -44.78, -43.02, -41.27,  
-39.51, -37.76, -36.00, -34.24, -32.49, -30.73

YDEF 102 GAUSR40 1

The entire gaussian grid is present, starting at the southernmost  
latitude (-88.66).

ZDEF number mapping <start increment>  
          <value-list>

Defines the mapping between grid values and pressure level.  
Specifically:

number -- the number of grid values in the X direction, specified as an integer number.

mapping -- mapping type, specified as a keyword.  
Valid are:  
    LINEAR -- Linear mapping  
    LEVELS -- Arbitrary pressure levels

start -- when mapping is LINEAR, this is the starting value, or the value when Z=1.

increment -- when mapping is LINEAR, the increment in the Z direction, or from lower to higher. This may be a negative value, for example:

```
ZDEF 10 LINEAR 1000 -100
```

indicating that the data is for levels 1000, 900, 800, 700, etc.

value-list -- when the mapping is LEVELS, the specific levels are simply listed in ascending order. If there is only one level, use LINEAR, since LEVELS implies at least two levels.

TDEF number LINEAR start-time increment

Defines the mapping between grid values and time.  
Specifically:

number -- the number of times in the data set.  
Specified as an integer number.

start-time -- The starting date/time value, specified in GrADS absolute date/time format. This is the value when T=1. The date/time format is:

```
hh:mmZddmmmyyyy
```

where:

hh = hour (two digit integer)  
mm = minutes (two digit integer)  
dd = day (one or two digit integer)  
mmm = month (jan, feb, mar, apr, may, jun, jul, aug, sep, oct, nov, dec)  
yyyy = year (two or four digit integer.  
two digits implies a year between 1950 and 2049).

If not specified, hh defaults to 00, mm defaults to 00, and dd defaults to 1. The month and year must be specified. No intervening blanks are allowed in a GrADS absolute date/time.

Examples:

12Z1JAN1990  
14:20Z22JAN1987  
JUN1960

increment -- time increment. Specified in GrADS time increment  
format:

vvkk

where:

vv = an integer number, 1 or 2 digits  
kk = an increment keyword,  
mn = minutes  
hr = hours  
dy = days  
mo = months  
yr = year

Examples:

20mn -- increment is 20 minutes  
1mo -- increment is 1 month  
2dy -- increment is 2 days

Some examples of a TDEF statement:

TDEF 24 LINEAR 00Z01JUN1987 1HR

The data set has 24 times, starting at 00Z on 1 Jun, 1987,  
with an increment of 1 hour.

TDEF 30 LINEAR 2JUN1988 1DY

The data set has 30 times, starting at 00Z on 2 Jun, 1988,  
with an increment of 1 day.

VARS number

Indicates the start of the records describing the variables  
in the data set.

number -- the number of variable records

Each variable record is in the following format:

abrev levs units description

abrev -- a 1 to 8 character abbreviation for this variable.  
This abbreviation must start with an alphabetic  
character (a-z) and be composed of alphabetic  
characters and numbers. This abbreviation will be  
the "name" the variable is accessed by from within  
GrADS.



levs            -- an integer value specifying the number of levels  
                   this variable has in the data set. It may not  
                   exceed the number of levels in the ZDEF statement.  
                   A lev's value of 0 indicates this variable has one  
                   "level" that does not correspond to a vertical level.  
                   An example would be a surface variable.

units           -- Reserved for future use. Put a value of 99 here.

description - A text description of the variable, max 40 characters.

After the last variable record comes the ENDDVARS statement. This ends  
 the GrADS data descriptor file.

#### Examples of Creating a Gridded Data Set

On a workstation, the binary GrADS data sets need to be  
 created as a 'stream' data set, ie, it should not have the  
 normal FORTRAN record descriptor words imbedded in it.  
 This can be done from FORTRAN using direct access I/O:

```

REAL Z(72,46,16)
.
.
OPEN (8,FILE='grads.dat',FORM='UNFORMATTED',
& ACCESS='DIRECT',RECL=72*46)
.
IREC=1
DO 10 I=1,16
  WRITE (8,REC=IREC) ((Z(J,K,I),J=1,72),K=1,46)
  IREC=IREC+1
10 CONTINUE

```

This example writes out 16 levels of one variable to a file  
 in direct access format. We are really writing the data out  
 sequentially, and using direct access to avoid having the  
 record descriptor words written. There may be options in  
 your compiler to do this more directly, or you may wish to  
 write the data using a C program.

Another simple sample might be:

```

REAL X(100)
DO 10 I=1,100
  X(I)=I
10 CONTINUE
OPEN (8,FILE='samp.dat',FORM='UNFORMATTED',ACCESS='DIRECT',
& RECL=100)
WRITE (8,REC=1) X
STOP
END

```

The associated descriptor file:

```
DSET samp.dat
TITLE Sample Data Set
UNDEF -9.99E33
XDEF 100 LINEAR 1 1
YDEF 1 LINEAR 1 1
ZDEF 1 LINEAR 1 1
TDEF 1 LINEAR 1JAN2000 1DY
VARS 1
x 0 99 100 Data Points
ENDVARS
```

Once created, you can use this data set to experiment with  
GRADS data functions, such as:

```
display sin(x/50)
```

## References

- Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobie, A. da Silva, 1995: Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System—Version 1. NASA Tech. Memo. No. 104606, volume 4, Goddard Space Flight Center, Greenbelt, MD 20771.
- Schubert, S.D., R. B. Rood, and J. Pfaendtner, 1993: An assimilated data set for earth science applications. *Bul. Amer. Meteor. Soc.*, 74, 2331-2342.
- Schubert, S., C.-K. Park, Chung-Yu Wu, W. Higgins, Y. Kondratyeva, A. Molod, L. Tkacs, M. Seablom, and R. Rood, 1995a: A Multiyear Assimilation with the GEOS-1 System: Overview and Results. NASA Tech. Memo. 104606, Vol. 6. Goddard Space Flight Center, Greenbelt, MD 20771.
- Schubert, S. D. and R. Rood, 1995b: Proceedings of the Workshop on the GEOS-1 Five-Year Assimilation. NASA Tech. Memo. 104606, Vol. 7, Goddard Space Flight Center, Greenbelt, MD 20771.
- Suarez, M.J. and Takacs, L.L., 1995: Documentation of the Aries-GEOS Dynamical Core: Version 2. NASA Tech. Memo. No. 104606, volume 5, Goddard Space Flight Center, Greenbelt, MD 20771.
- Takacs, L.L., A. Molod, and T. Wang, 1994: Documentation of the Goddard Earth Observing System (GEOS) General Circulation Model—Version 1. NASA Tech. Memo. No. 104606, volume 1, Goddard Space Flight Center, Greenbelt, MD 20771.

**NOTE:** The NASA Tech Memoranda listed above are available by anonymous ftp from

`ftp://dao.gsfc.nasa.gov/pub/tech_memos.`