Global Modeling and Assimilation Office

GMAO Office Note No. 1 (Version 2.3)

File Specification for MERRA Products

Release Date: 8/3/2012

Global Modeling and Assimilation Office
Earth Sciences Division
NASA Goddard Space Flight Center
Greenbelt, Maryland 20771

This page intentionally left blank.

File Specification for MERRA Products

Document maintained by Rob Lucchesi (GMAO, SSAI)

This document should be cited as
Lucchesi, R., 2012: File Specification for MERRA Products. GMAO Office Note No. 1 (Version 2.3), 82 pp, available from http://gmao.gsfc.nasa.gov/pubs/office_notes.

Approved by:

____

Michele M. Rienecker Date

Head, Global Modeling and Assimilation Office
Code 610.1, NASA GSFC

REVISION HISTORY

Version Number	Revision Date	Extent of Changes
1.0	03/24/2006	Baseline
2.0	04/19/2007	First version posted for comment on MERRA web site.
2.1	09/02/2008	Collection tables updated to reflect MERRA production. Glossary in progress. Appendix H (Budgets) added.
2.2	09/01/2011	Added documentation for ocean product. Some corrections were made throughout the document, including updates to the descriptions of the SWTNT* and SPEED variables. The detailed descriptions of MERRA variables in Appendix F was updated and moved to a separate document.
2.3	08/03/2012	Appendix C removed. Added documentation for MERRA-Land product (tavg1_2d_mld_Nx) throughout and corrected discussion of land budgets in Appendix F.

TABLE OF CONTENTS

1. Introduction

The Modern-Era Retrospective analysis for Research and Applications (MERRA) is a NASA atmospheric reanalysis for the satellite era using the Goddard Earth Observing System Model, Version 5 (GEOS-5) with its Atmospheric Data Assimilation System (ADAS), version 5.2.0. MERRA focuses on historical analyses of the hydrological cycle on a broad range of weather and climate time scales and places the NASA EOS suite of observations in a climate context. MERRA covers the period 1979-present, continuing as an ongoing climate analysis as resources allow.

The MERRA system is documented in Rienecker et al. (2008). The GEOS-5 system actively assimilates roughly 2×10^6 observations for each analysis, including about 7.5×10^5 AIRS radiance data. The input stream is roughly twice this volume, but because of the large volume, the data are thinned commensurate with the analysis grid to reduce the computational burden. Data are also rejected from the analysis through quality control procedures designed to detect, for example, the presence of cloud. To minimize the spurious periodic perturbations of the analysis, MERRA uses the Incremental Analysis Update (IAU) technique developed by Bloom et al. (1996). More details of this procedure are given in Appendix A.

Observations used for MERRA are summarized in Appendix B. The data streams assimilated by the GEOS-5 DAS for MERRA are shown in Tables B.1 (conventional) and B.2 (satellite); their availability and the source of the data assimilated for MERRA are shown in Tables B.3 (conventional) and B.4 (satellite).

The analysis is performed at a horizontal resolution of 2/3-degree longitude by ½-degree latitude and at 72 levels, extending to 0.01 hPa. Some products, such as the instantaneous analysis fields, are available on the native three-dimensional grid. Hourly two-dimensional diagnostic fields are also available at the native horizontal resolution. Other products are available on a coarser horizontal grid with resolution of 1.25 ×1.25 degrees or 1 × 1.25 degrees, the latter for use by the chemistry transport community. These may be on the model's native vertical grid or at 42 pressure surfaces extending to 0.1 hPa. Surface data, near surface meteorology, selected upper-air levels, and vertically integrated fluxes and budgets are produced at one-hour intervals, which will help the development of offline land and ocean models and data assimilation systems by resolving the diurnal cycle. Monthly mean versions of most of these products are also available.

This document describes the gridded output files produced by the MERRA reanalysis. Further details regarding MERRA can be found in the peer-reviewed literature (e.g., Rienecker et al., 2011 and other papers in the Journal of Climate MERRA Special Collection, http://journals.ametsoc.org/page/MERRA), and at http://gmao.gsfc.nasa.gov/research/merra/. See Reichle (2012) and Reichle et al. (2011) for additional information about *MERRA-Land*, a supplemental MERRA land surface data product.

The MERRA data are available online through the Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (http://disc.sci.gsfc.nasa.gov/mdisc/). Detailed documentation of the data access procedures can be found at http://gmao.gsfc.nasa.gov/MERRA.

2. Format and File Organization

GEOS-5 files are in HDF-EOS format, which is an extension of the Hierarchical Data Format Version 4 (HDF-4), developed at the National Center for Supercomputing Applications http://www.hdfgroup.org/. HDF-EOS provides additional capabilities over HDF-4, but does not prevent the use of the files as standard HDF-4 files (de Witt, 1996; Gross, 1997).

Each GEOS-5 file contains a single HDF-EOS grid, which in turn contains a number of geophysical quantities that we refer to as "fields" or "variables." Some files contain 2-D variables on a longitude-latitude grid, and some files contain 3-D variables, or a mixture of 2-D and 3-D variables on the same longitude-latitude grid, but with a vertical dimension applicable to all the 3D variables.

The variables are created using the **GDdeffield** function from the HDF-EOS GD (GriD) API, which stores them as HDF Scientific Data Set (SDS) arrays, so that they can be read with standard HDF routines. In addition to the geophysical variables, the files have SDS arrays that define dimension scales (or coordinate variables). There are two distinct scales for each dimension, which ensures that a wide variety of graphical display tools can interpret the data. In particular, there are dimension scales that adhere to the CF conventions, as well as ones that adhere to the COARDS conventions.

Due to the size of the MERRA archive, all data are compressed with a GRIB-like method that is invisible to the user. This method does degrade the precision of the data, but every effort has been made to ensure that differences between the product and the original, uncompressed data are not scientifically meaningful. Once the precision has been degraded, the files are written using the standard (internal) gzip deflation available in HDF-4.

EOS Core System (ECS) metadata and other information are stored as global attributes.

2.1 Dimensions

GMAO HDF-EOS files contain two sets of dimension scale (coordinate) information. One set of dimensions is defined using the **SDsetdimscale** function of the standard HDF SD interface. This set of scales has an attribute named "units," set to an appropriate string defined by the CF and COARDS conventions that can be used by applications to identify the dimension. The other set of dimension scales is created using the **GDdeffield/GDwritefield** functions.

Table 2.1-1. Dimension Variables Contained in GMAO HDF-EOS Files

Name	**Description**	**Type**	***units* attribute**
XDim:EOSGRID	Longitude	float32	degrees_east
YDim:EOSGRID	Latitude	float32	degrees_north
Height:EOSGRID (3D only)	pressure or layer index	float32	hPa or layer
TIME:EOSGRID	minutes since first time in file	float32	minutes
XDim	Longitude, in degrees east	float64	N/A
YDim	Latitude, in degrees north	float64	N/A
Height (3D only)	pressure or level indices	float64	N/A
Time	seconds since 00:00Z on 1 January 1993	float64	N/A

The 32-bit dimension variables have a "units" attribute that makes them COARDS-compliant, while the 64-bit dimension variables satisfy ECS requirements.

2.2 Variables

Variables are stored as SDS arrays, even though they are defined with the HDF-EOS **GDdeffield** function. As a result, one can use the SD interface of the HDF library to read any variable from the file. The only thing one must know is the SDS variable name listed in the Tables in Appendix D and the number of dimensions (the rank). One can quickly list the variables in the file by using common utilities such as *ncdump* or *hdp*. Both utilities are distributed by the HDF Group with the HDF-4 library. In Section 8 we present sample code for reading one or more data fields from this file. The *short names* for all variables in all GMAO data products are listed in Appendix D. A glossary with a brief description of each variable is available in the separate GEOS-5 Variable Definition Glossary, available on the GMAO web page.

Each variable has several useful metadata attributes. Many of these attributes are required by the CF and COARDS conventions, while others are specific for GMAO products. The following table lists required attributes. Other attributes may be included for internal GMAO use and can be ignored.

Table 2.2-1 Metadata attributes associated with each SDS.

Name	**Type**	**Description**
_FillValue	float32	Floating-point value used to identify missing data. Normally set to 1e15. Required by CF.
missing_value	float32	Same as _FillValue. Required for COARDS backwards compatibility.
valid_range	float32, array(2)	This attribute defines the valid range of the variable. The first element is the smallest valid value and the second element is the largest valid value. Required by CF. In MERRA files these are set to -/+ _FillValue.
long_name	String	A brief description of the variable contents taken from the *Description* column of the tables in Appendix D.
standard_name	String	An ad hoc description of the variable as required by COARDS. It approximates the standard names as defined in an early version of CF conventions. (See References).
units	String	The units of the variable. Must be a string that can be recognized by UNIDATA's Udunits package.
scale_factor	float32	If variable is packed as 16-bit integers, this is the scale_factor for expanding to floating-point. Currently data are not packed, thus value is 1.0.
add_offset	float32	If variable is packed as 16-bit integers, this is the offset for expanding to floating-point. Currently, data are not packed, thus value is 0.0.

2.3 Global Attributes

In addition to SDS arrays containing variables and dimension scales, global metadata is also stored in GMAO HDF-EOS files. Some metadata are required by the CF/COARDS conventions, some are present to meet ECS requirements, and others as a convenience to users of GMAO products. A summary of global attributes present in all MERRA files is shown in Table 2.3-1.

Table 2.3-1 Global metadata attributes associated with each SDS.

	Type	**Description**
Conventions	character	Identification of the file convention used, currently "CF-1.0"
title	character	Experiment identification: "MERRA" or "MERRA-Land"
history	character	CVS tag of this release. CVS tags are used internally by the GMAO to designate versions of the system.
institution	character	"NASA Global Modeling and Assimilation Office"
source	character	CFIO Version (CFIO is the GMAO's IO layer)
references	character	GMAO website address
comment	character	As required
HDFEOSVersion	character	Version of the HDF-EOS library used to create this file.
StructMetadata.0	character	This is the GridStructure metadata that is created by the HDF-EOS library.
CoreMetadata.0	character	The ECS inventory metadata.
ArchivedMetadata.0	character	The ECS archive metadata

3. Instantaneous vs Time-averaged Products

Each file collection listed in Section 6 contains either instantaneous or time-averaged products, but not both.

All instantaneous collections contain fields at *synoptic times* (00 GMT, 06 GMT, 12 GMT, and 18 GMT). In addition, three-hourly instantaneous collections also include snapshots at *mid-synoptic times* (03 GMT, 09 GMT, 15 GMT, and 21 GMT).

Time-averaged collections contain either hourly, three-hourly, monthly, or seasonal means, but not mixtures of these. (Files with daily mean values are also available through the MDISC subsetting tools.) Each time-averaged collection consists of a continuous sequence of data averaged over the indicated interval and time stamped with the central time of the interval. For hourly data, for example, these times are 00:30 GMT, 01:30 GMT, 02:30 GMT, etc. Only products consisting exclusively of two-dimensional (horizontal) fields are produced hourly. Three-hourly time-averaged files contain averages over time intervals centered and time stamped at 01:30 GMT, 04:30 GMT, 07:30 GMT, and so on. Monthly files represent averages for the calendar months, accounting for leap years. For monthly means, each file contains a single month.

Each hourly, three-hourly, or six-hourly collection, whether instantaneous or time-averaged, consist of a set of daily files, with the date as part of the filename. For collections of monthly or seasonal means each month or season is in a separate file, and file names also include a date in the file name.

4. Grid Structure

4.1 Horizontal Structure

Fields will be produced on three different horizontal grids: the a native grid of the Finite-Volume (FV) dynamical core, with a resolution of 2/3 longitude by 1/2 degree latitude, a coarse version of the FV grid, with a resolution of 1.25 longitude by 1 degree latitude, and a uniform coarse grid, with a resolution of 1.25 by 1.25 degrees.

The GEOS-5 MERRA ***native resolution*** gridded output is on a global horizontal grid, consisting of **IMn=540** points in the longitudinal direction and **JMn=361** points in the latitudinal direction. The horizontal native grid origin, associated with variables indexed (i=1, j=1) represents a grid point located at (180°W, 90°S). Latitude (φ) and longitude (λ) of grid points as a function of their indices (i, j) can be determined by:

$$\lambda_i = -180 + (\Delta\lambda)_n (i-1), \quad i = 1, \mathrm{IMn}$$
$$\varphi_j = -90 + (\Delta\varphi)_n (j-1), \quad j = 1, \mathrm{JMn}$$

Where $(\Delta\lambda)_n$=2/3° and $(\Delta\varphi)_n$ = 1/2°. For example, (i = 271, j = 181) corresponds to a grid point at (λ = 0, φ = 0).

FV ***reduced resolution*** gridded output is used only for the specialized chemistry forcing files (see Appendix ED). It is on a global horizontal grid, consisting of **IMf=288** points in the longitudinal direction and **JMf=181** points in the latitudinal direction. The horizontal native grid origin, associated with variables indexed (i=1,j=1) represents a grid point located at (180°W, 90°S). Latitude and longitude of grid points as a function of their indices (i, j) can be determined by:

$$\lambda_i = -180 + (\Delta\lambda)_f (i-1), \quad i = 1, \mathrm{IMf}$$
$$\varphi_j = -90 + (\Delta\varphi)_f (j-1), \quad j = 1, \mathrm{JMf}$$

where $(\Delta\lambda)_f$ = 1.25° and $(\Delta\varphi)_f$ = 1°.

MERRA ***reduced resolution*** gridded output is on a global horizontal grid consisting of **IMc=288** points in the longitudinal direction and **JMc=144** points in the latitudinal direction. The horizontal coarse grid origin, associated with variables indexed (i=1, j=1), represents a grid point at (179.375°W, 89.375°S), so that the Dateline and the South Pole are "edges" of the first grid box. Latitude and longitude as a function of their indices (i, j) for the coarse grid can be determined by:

$$\lambda_i = -180 + (\Delta\lambda)_c (i - \tfrac{1}{2}), \quad i = 1, \mathrm{IMc}$$
$$\varphi_j = -90 + (\Delta\varphi)_c (j - \tfrac{1}{2}), \quad j = 1, \mathrm{JMc}$$

where $(\Delta\lambda)_c = (\Delta\varphi)_c = 1.25^\circ$.

4.2 Vertical Structure

Gridded products use four different vertical configurations: Horizontal-only (can be vertical averages, single level, or surface values), pressure-level, model-level, or model-edge. Horizontal-only data for a given variable appear as 3-dimensional fields (x, y, time), while pressure-level, model-level, or model-edge data appear as 4-dimensional fields (x, y, z, time). In all cases the time dimension spans multiple files. Pressure-level data is output on the **LMp=42** pressure levels shown in Appendix C. The GEOS-5 model layers used for MERRA products are on a terrain-following hybrid sigma-p coordinate. Model-level data will be output on the **LM=72** layers shown in the second table of Appendix C. The model-edge products contain fields with **LMe = LM + 1** levels representing the layer edges. The pressure at the model top is a fixed constant, **PTOP=0.01 hPa**. Pressures at model edges should be computed by summing the DELP starting at PTOP. A representative pressure for the layer can then be obtained from these. In the GEOS-4 *eta* files, one could compute the pressure on the edges by using the "ak" and "bk" values and the surface pressure. In GEOS-5, the full 3-dimensional pressure variables are explicitly provided through ($\mathrm{DELP}_{\mathbf{ijl}}$) and PTOP. For the MERRA products documented here, all model-level fields are on a hybrid-sigma coordinate, and their vertical location could be obtained from the "ak-bk" relationship as well. But this may change in future GMAO systems. We thus recommend that users rely on the reported 3D pressure distribution, and not use ones computed from the "ak" and "bk".

Note that the indexing for the GEOS-5 vertical coordinate system in the vertical is top to bottom, i.e., layer 1 is the top layer of the atmosphere, while layer LM is adjacent to the earth's surface. The same is true for edge variables, with level 1 being the top of the model's atmosphere (PTOP), and level LM+1 being the surface.

5. File Naming Conventions

Each GEOS-5 product file will have a complete file name identified in the EOSDIS metadata as "LocalGranuleID". EOSDIS also requires eight-character abbreviated naming indices for each Earth Science Data Type (ESDT). In MERRA each file collection has a unique ESDT index. The ESDT index convention is described in section 5.2.

5.1 File Names

The standard full name for the assimilated GEOS-5 MERRA products will consist of five dot-delimited nodes:

runid.runtype.config.collection.timestamp

The node fields, which vary from file to file, are defined as follows:

runid

All MERRA *Mainstream* runs will be identified by **MERRA**SVv , where the numeric qualifiers S and Vv denote the production *Stream* and the *Version* numbers. For MERRA-Land files, the *runid* identifies the MERRA *Stream* and *Version* that was used to construct the surface meteorological forcing data.

MERRA will be run in three production *Streams*, each covering approximately a third of the MERRA period. If the stream number is not applicable (the case for ancillary products, such as forecasts), *S* is set to 0, otherwise it is 1, 2, or 3. Each of the three MERRA streams also had a multi-year spin-up period and these periods are available from the GES DISC for specialized analyses. Filenames from the spin-up periods will be denoted with a *runid* of **SPINUP_MERRA**SVv

The *Version* numbers are non-zero when there is more than one version of the dataset. It is usually zero, denoting the original processing. MERRA will be conducted with a frozen assimilation system, so there should be no updates or patches to the GEOS-5 software. Version changes indicate either that a problem was encountered after product release and a reprocessing was necessary (v), or a period was reprocessed with a modified version of the system for scientific studies (V). For any such reprocessing, we will increment the version number appropriately and this will be documented in the metadata parameter “history.” Information on version differences will also be available at the MERRA web site (http://gmao.gsfc.nasa.gov/merra/).

***runtype*:**

All MERRA *mainstream* and *spin-up* filenames have a ***runtype*** designation **prod.**
prod - Standard product
swep - Sweeper/Scout product
rosb - Reduced observing system product
cers - CERES observing system product

***config*:**

The GEOS-5 analysis and forecast system can run in different configurations.

assim – Assimilation. Uses a combination of atmospheric data analysis and model forecasting to generate a time-series of global atmospheric quantities.

simul – Simulation. Uses a free-running atmospheric model with some prescribed external forcing, such as sea-surface temperatures. MERRA-Land is a land-only "replay" (model integration) of the MERRA land surface model component with prescribed surface meteorological forcing data. For MERRA-Land the configuration is therefore described as a "simulation".

frcst – Forecasts. Uses a free-running atmospheric model initialized from an analyzed state.

The main configuration used to support MERRA is assimilation (**assim**).

***collection*:**

All MERRA data are organized into file *collections* that contain fields with common characteristics. These collections are used to make the data more accessible for specific purposes. Fields may appear in more than one collection. Collection names are of the form *freq_dims_group_HV*, where the four attributes are:

freq: time-independent (**cnst**), instantaneous (**inst***F***),** or time-average (**tavg***F*), where *F* indicates the frequency or averaging interval and can be any of the following:
1 = Hourly
3 = 3-Hourly
6 = 6-Hourly
M = Monthly mean
U = Monthly-Diurnal mean
0 = Not Applicable
A *freq* designation of **M** or **U** can apply to either an **inst** or a **tavg** file depending on whether it is a monthly mean of instantaneous or time-averaged data.

dims: **2d** for collections with only 2-dimensional fields or **3d** for collections with a mix of 2- and 3-dimensional fields.

group: A three-letter mnemonic for the type of fields in the collection. It is a lowercase version of the group designation used in the ESDT name, as listed in the next section.

HV: Horizontal and Vertical grid.

H can be:

N: Native (2/3 x ½) horizontal resolution
C: Reduced (1.25 x 1.25) horizontal resolution
F: Reduced FV (1.25 x 1) horizontal resolution

V can be:

x: horizontal-only data (surface, single level, etc.) ; *dims* must be **2D**
p: pressure-level data (see Appendix D for levels) ; *dims* must be **3D**
v: model layer centers (see Appendix D) *dims* must be **3D**
e: model layer edges (see Appendix D) *dims* must be **3D**

timestamp:

This node defines the date and time associated with the data in the file. It has the form *yyyymmdd* for either instantaneous or time-averaged daily files, *yyyymm* for monthly-mean files.

yyyy - year string (e.g. , "2002")
mm - month string (e.g.., "09" for September)
dd - day of the month string (optional)

EXAMPLE:

MERRA300.prod.assim.tavg3_3d_tdt_Cp.20020915.hdf

This is an example of a MERRA filename from the production ("prod") segment of the original version of the third (most recent) assimilation stream ("MERRA300"). The data are hourly time-averages ("tavg1"), three-dimensional ("3d"), temperature tendency products ("tdt"), at reduced horizontal resolution and interpolated to pressure levels ("Cp"). The file contains all data for 15 September 2002 and is in "hdf" format.

5.2 Earth Science Data Types (ESDT) Name

To accommodate EOSDIS toolkit requirements, all MERRA files are associated with a nine-character ESDT. The ESDT is a short (and rather cryptic) handle for users to access sets of files. In MERRA the ESDT will be used to identify the *Mainstream collections* and consists of a compressed version of the collection name of the form:

M*CTFHVGGG*

where

C: Configuration
A = Assimilation
F = Forecast
S = Simulation

T: Time Description:
I = Instantaneous
T = Time-averaged
C = Time-independent

F: Frequency
1 = Hourly
3 = 3-Hourly
6 = 6-Hourly
M = Monthly mean
U = Monthly-Diurnal mean
0 = Not Applicable

H: Horizontal Resolution
N = Native
F **=** Reduced resolution version of model grid
C = Reduced resolution

V: Vertical Location
X = Two-dimensional
P = Pressure
V = model layer center
E = model layer edge

GGG: Group
ANA = direct analysis products
ASM = assimilated state variables (from the IAU corrector, see Appendix A)
TDT = tendencies of temperature
UDT = tendencies of eastward and northward wind components
QDT = tendencies of specific humidity
ODT = tendencies of ozone
LND = land surface variables
MLD =MERRA-Land
FLX = surface turbulent fluxes and related quantities
MST = moist processes

CLD = clouds
RAD = radiation
TRB = turbulence
SLV = single level
INT = vertical integrals
CHM = chemistry forcing
OCN = ocean

6. Summary of MERRA file collections

The GEOS-5 MERRA product is organized into the 26 collections listed in Tables 6.1 and 6.2. These are described in detail in Appendix D. The 19 collections in Table 6-1 are the “standard” products intended for most diagnostic work. The seven “chemistry” collections in table 6-2 are more specialized products, intended for forcing off-line chemistry transport models (CTMs). The last three of these, in particular, are for CTMs using the GEOS-5 transport cores.

Collection names follow the conventions described in Section 5. Sizes are given in Gbytes/day and Tbytes for the entire collection. All numbers referred to the compressed file sizes.

Table 6-1 List of standard collections

Name	**Description**	**Size Gbytes/day // Tbytes**
const_2d_asm_Nx	Constant fields	
const_2d_mld_Nx	MERRA land surface constants	
inst6_3d_ana_Nv	Analyzed fields on model layers	0.452
inst6_3d_ana_Np	Analyzed fields at pressure levels	0.291
inst3_3d_asm_Cp	Basic assimilated fields from IAU corrector	0.231
tavg3_3d_cld_Cp	Upper-air cloud related diagnostics	0.075
tavg3_3d_mst_Cp	Upper-air diagnostics from moist processes	0.056
tavg3_3d_trb_Cp	Upper-air diagnostics from turbulence	0.147
tavg3_3d_rad_Cp	Upper-air diagnostics from radiation	0.088
tavg3_3d_tdt_Cp	Upper-air temperature tendencies by process	0.191
tavg3_3d_udt_Cp	Upper-air wind tendencies by process	0.224
tavg3_3d_qdt_Cp	Upper-air humidity tendencies by process	0.166
tavg3_3d_odt_Cp	Upper-air ozone tendencies by process	0.083
tavg1_2d_slv_Nx	Single-level atmospheric state variables	0.285
tavg1_2d_flx_Nx	Surface turbulent fluxes and related quantities	0.267
tavg1_2d_rad_Nx	Surface and TOA radiative fluxes	0.189
tavg1_2d_lnd_Nx	Land surface diagnostics (MERRA)	0. 146
tavg1_2d_mld_Nx	Land surface diagnostics (MERRA-Land)	0.118
tavg1_2d_ocn_Nx	Ocean quantities	0.090

Name	Description	Size Gbytes/day // Tbytes
tavg1_2d_int_Nx	Vertical integrals of tendencies	1.500
inst1_2d_int_Nx	Vertical integrals of quantities	0.115
TOTAL		4.714 // 51.8

Table 6-2 List of chemistry collections

Name	Description	Size (Gbytes)
const_2d_chm_Fx	2-D invariants on chemistry grid	
tavg3_3d_chm_Fv	Chemistry related 3-D at model layer centers	0.329
tavg3_3d_chm_Fe	Chemistry related 3-D at model layer edges	0.166
tavg3_2d_chm_Fx	Chemistry related 2-DSingle-level	0.020
tavg3_3d_chm_Nv	Accumulated transport fields at layers	0.915
tavg3_3d_chm_Ne	Accumulated transport fields at edges	0.469
inst3_3d_chm_Ne	Instantaneous fields for off-line transport	0.050
TOTAL CHEM		1.949 // 21.44

7. Metadata

GEOS-5 gridded output files include two types of metadata. When using the HDF-EOS library and tools, the EOSDIS metadata are used (Dopplick, 1997). Other utilities such as GrADS or FERRET use the CF metadata (NOAA, 1995).

7.1 EOSDIS Metadata

The EOSDIS toolkit only uses the EOSDIS metadata. EOSDIS identifies two major types of metadata, collection and granule.

Collection metadata are stored in a separate index file. This file describes an ESDT and is like a card in a library catalog. Each GMAO data product has an ESDT description in the EOS Core System that contains its unique collection attributes.

Granule metadata is the "table of contents" information stored on the data file itself. The EOSDIS granule metadata include:

- File name (local granule ID)
- Grid structure
- Number of time stamps stored in the file
- Number of vertical levels for each variable in this file
- Names of variables in this file
- Variable format (32-bit floating point, 16-bit integer, etc.)
- Variable storage dimensions

    time, latitude and longitude for 2-d fields

    time, latitude, longitude and vertical levels 3-d fields

- "Missing" value for each variable.

7.2 CF Metadata

When GrADS or FERRET are used to view GEOS-5 gridded data sets, the application uses the CF metadata embedded in the data products. These metadata include the following information:

- Space-time grid information (dimension variables)
- Variable names and descriptions
- Variable units
- "Missing" value for each variable.

The grid information and units comply with the CF conventions. The variable names and descriptions are only loosely based on an early version of CF conventions.

8. Sample Software

The following example illustrates the use of the standard HDF library or the ECS HDF-EOS library to read GEOS-5 products. The program shown below accepts as command line arguments a file name and a field name. It opens the file, reads the requested field at the first time, computes an average for this field, and prints the result to standard output. There are two versions of this program. The first version uses the HDF-EOS library to read the file. The second version uses the standard HDF library to read the file. Electronic copies of these programs can be obtained from the Operations section of the GMAO web page.

8.1 HDF-EOS Example

```
/*****************************************************************************/
/*  This program demonstrates how to read a field from a GMAO HDF-EOS        */
/*  product using the HDF-EOS library. It will take a file name and          */
/*  field name on the command line, read the first time of the given         */
/*  field, calculate an average of that time and print the average.          */
/*                                                                           */
/*  usage: avg <file name> <field name>                                      */
/*                                                                           */
/*  Rob Lucchesi                                                             */
/*  rob.lucchesi@nasa.gov                                                    */
/*****************************************************************************/

#include "hdf.h"
#include "mfhdf.h"
#include <stdio.h>

main(int argc,char *argv[]) {

int32 sd_id, sds_id, status;
int32 sds_index;
int32 start[4], edges[4], stride[4];
char *fname, *vname;
float32 *data_array;
float32 avg, sum;
int32 i;
int32 file_id, gd_id;
int32 xdim, ydim, zdim, len;

if (argc != 3) {
        printf("Usage: avg <filename> <field> \n");
        exit (-1);
}
```

```
fname = argv[1];
vname = argv[2];

/* Open the file (read-only) */

if ( (file_id = GDopen (fname, DFACC_RDONLY))< 0) {
        printf ("Could not open %s\n",fname);
        exit(-1);
}

/*  Attach to the EOS grid contained within the file.                          */
/*  The GMAO uses the generic name "EOSGRID" for the grid in all products. */


if ( (gd_id = GDattach (file_id,"EOSGRID"))< 0) {
        printf ("Could not open %s\n",fname);
        exit(-1);
}

status = GDget(file_id,xdim,ydim,zdim),

/*  Set positioning arrays to read the entire field at the first time. */

start[0] = 0;
start[1] = 0;
start[2] = 0;
start[3] = 0;

stride[0] = 1;
stride[1] = 1;
stride[2] = 1;
stride[3] = 1;

edges[0] = 1;
edges[1] = zdim;
edges[2] = ydim;
edges[3] = xdim;

len =  xdim*ydim*zdim;
data_array = (float32 *)malloc(len);

/*  Read the data into data_array */

status = GDreadfield (gd_id, vname, start, stride, edges, data_array);
printf ("Read status=%d\n",status);
```

```
/*  Calculate and print the average */

sum=0.0;
for (i=0; i<len; i++)  sum += data_array[i];
avg = sum/(float32)len;

printf ("Average of %s in 3 dimensions is=%f\n",vname,avg);

/* Close file. */

status = GDdetach (gd_id);
status = GDclose (file_id);

}
```

8.2 HDF (non EOS) Example

```
/****************************************************************************/
/* This program demonstrates how to read a field from a GMAO HDF-EOS       */
/* product using the HDF library (HDF-EOS not required). It will take      */
/* a file name and field name on the command line, read the first time     */
/* of the given field, calculate an average of that time and print the average. */
/*                                                                          */
/* usage: avg <file name> <field name>                                      */
/*                                                                          */
/* Rob Lucchesi                                                             */
/* rob.lucchesi@nasa.gov                                                    */
/*                                                                          */
/****************************************************************************/

#include "hdf.h"
#include "mfhdf.h"
#include <stdio.h>

main(int argc,char *argv[]) {
int32 sd_id, sds_id, status;
int32 sds_index;
int32 start[4], edges[4], stride[4];
char *fname, *vname;
float32 *data_array;
float32 avg, sum;
int32 i, xdim, ydim , zdim, len;

if (argc != 3) {
        printf("Usage: avg <filename> <field> \n");
        exit (-1);
}

fname = argv[1];
vname = argv[2];

/* Open the file (read-only) */

if ( (sd_id = SDstart (fname, DFACC_RDONLY))< 0) {
        printf ("Could not open %s\n",fname);
        exit(-1);
}

/* Find the index and ID of the SDS for the given variable name and get its dimensions. */
```

```
if ( (sds_index = SDnametoindex (sd_id, vname)) < 0) {
        printf ("Could not find %s\n",vname);
        exit(-1);
}

sds_id = SDselect (sd_id,sds_index);
status = GDget(file_id,xdim,ydim,zdim),

/* Set positioning arrays to read the entire field at the first time. */

start[0] = 0;
start[1] = 0;
start[2] = 0;
start[3] = 0;

stride[0] = 1;
stride[1] = 1;
stride[2] = 1;
stride[3] = 1;

edges[0] = 1;
edges[1] = zdim;
edges[2] = ydim;
edges[3] = xdim;

len =  xdim*ydim*zdim;
data_array = (float32 *)malloc(len);

/* Read the data into data_array */

status = SDreaddata (sds_id, start, stride, edges, (VOIDP) data_array);
printf ("read status=%d\n",status);

/* Calculate and print the average */

sum=0.0;
for (i=0; i<len; i++)  sum += data_array[i];
avg = sum/(float32)len;
printf ("Average of %s in 3 dimensions is=%f\n",vname,avg);

/* Close file. */

status = SDendaccess (sds_id);
status = SDend (sd_id);
}
```

Appendix A: The IAU procedure

The implementation of the Incremental Analysis Update (IAU, Bloom et al., 1996) used for MERRA is summarized in Figure A.1. Every six hours, at the synoptic times, an analysis is performed using backgrounds at that time, three hours earlier, and three hours later, which incorporates observations during the six-hour period spanned by the three backgrounds. The results of this analysis are the products included in the two analysis collections. The analysis increments (i.e., the difference between the analysis and the corresponding synoptic background) are then divided by a time scale (in MERRA we used 6 hours) to produce an "analysis tendency." The model is then "backed-up", restarting it from its state three hours before the analysis time, and run for six hours, adding in the time-invariant "analysis tendency" in addition to its normal physics tendencies. At that point a restart is created that will be used next time the model is backed-up, and the first background for the next analysis cycle is saved. We refer to this first 6-hour run as the "corrector" segment of the IAU. The run is then continued without an analysis tendency for another six hours, saving the other two backgrounds needed by the next analysis---one at the next synoptic time and another at the end of the six hours. We refer to this 6-hour run as the "predictor" segment of the IAU. The entire cycle is then repeated for subsequent synoptic times. Note that during each of the four daily analysis cycles the model is run for 12 hours---a 6-hour "corrector" followed by a 6-hour "predictor."

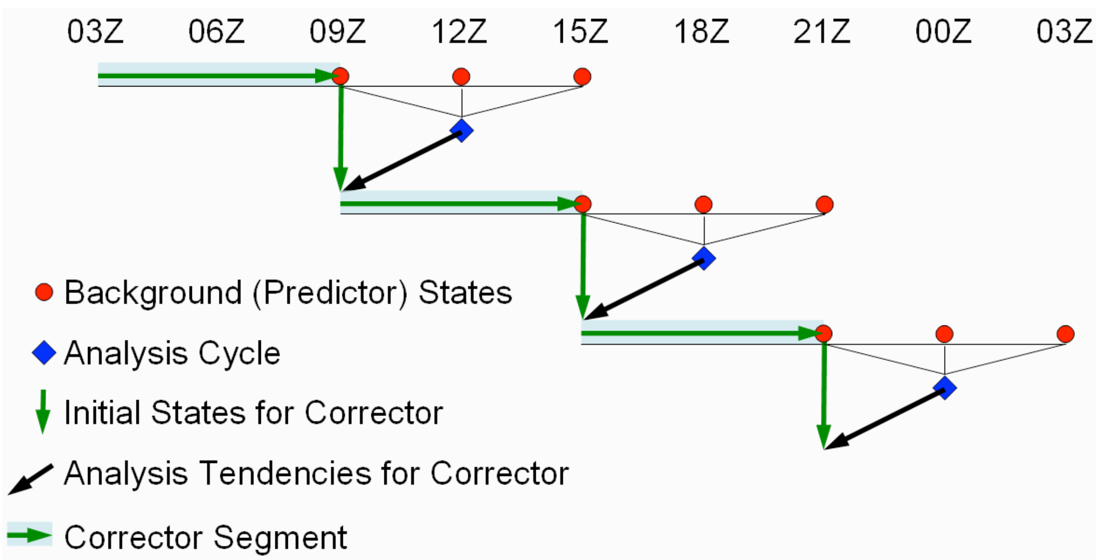


Figure A.1

Except for the analyses themselves, all products from MERRA are produced by the model during the corrector run segment. The sequence of corrector segments (follow the green line in the figure) is a continuous model run, with the extra forcing term from the analysis tendencies. The analysis tendencies do change abruptly every six hours, but state variables are continuous (within the model's time step) solutions of the equations of motion, albeit with the extra forcing term.

Appendix B: Observational Inputs to MERRA

Table B.1 Input conventional observations.

Radiosondes	u, v, T, q, P_s
PIBAL winds	u, v
Wind profiles	u, v
Conventional, ASDAR, and MDCRS aircraft reports	u, v, T
NEXRAD radar winds	u, v
Dropsondes	u, v, T, P_s
PAOB	P_s
GMS, METEOSAT, cloud drift IR and visible winds	u, v
MODIS clear sky and water vapor winds	u, v
GOES cloud drift IR winds	u, v
GOES water vapor cloud top winds	u, v
Surface land observations	P_s
Surface ship and buoy observations	u, v, T, q, P_s
SSM/I	Rain rate, wind speed
TMI	Rain rate
QuikSCAT	Wind speed and direction
SBUV2 ozone (Version 8 retrievals)	Nimbus 7, NOAA 9, 11, 14, 16, 17

Table B.2 Input Satellite Radiance Data

TOVS 1b Radiances	AMSU-A: N15, N16, N18 AMSU-B: N15, N16, N17 MHS: N18 HIRS2: TIROS-N, N6, N7, N8, N9, N10, N11, N12, N14 HIRS3: N16, N17 HIRS4: N18 MSU: TIROS-N, N6, N7, N8, N9, N10, N11, N12, N14 SSU: TIROS-N, N6, N7, N8, N9, N10, N11, N14
EOS/Aqua Level 1b Radiances	AIRS (152 channels), AMSU-A
SSM/I radiances	DSMP-13, DSMP-14, DSMP-15 (7 channels)
GOES sounder T_B	GOES-08, GOES-10, GOES-12 Channels 1-18

Table B.3 Conventional data availability and source.

DATA SOURCE/TYPE	PERIOD	DATA SUPPLIER
Radiosondes	1970 - present	NOAA/NCEP
PIBAL winds	1970 - present	NOAA/NCEP
Wind profiles	1992/5/14 - present	UCAR CDAS
Conventional, ASDAR, and MDCRS aircraft reports	1970 - present	NOAA/NCEP
Dropsondes	1970 - present	NOAA/NCEP
PAOB	1978 – 2010/08	NCEP CDAS
GMS, METEOSAT, cloud drift IR and visible winds	1977 – present	NOAA/NCEP
GOES cloud drift winds	1997 – present	NOAA/NCEP
EOS/Terra/MODIS winds	2002/7/01 - present	NOAA/NCEP
EOS/Aqua/MODIS winds	2003/9/01 - present	NOAA/NCEP
Surface land observations	1970 - present	NOAA/NCEP
Surface ship and buoy observations	1977 - present	NOAA/NCEP
SSM/I rain rate	1987/7 – 2009/09/16	NASA/GSFC/DAAC
SSM/I V6 wind speed	1987/7 – 2011/2/28	RSS
TMI rain rate	1997/12 - present	NASA/GSFC/DAAC
QuikSCAT surface winds	1999/7 – 2009/09	JPL
ERS-1 surface winds	1991/8/5 – 1996/5/21	CERSAT
ERS-2 surface winds	1996/3/19 – 2001/1/17	CERSAT
SBUV2 ozone	1978/10 - present	NASA/GSFC

Table B.4 Satellite radiance data availability and source

DATA SOURCE/TYPE	**PERIOD**	**DATA SUPPLIER**
TOVS/tn (TIROS N)	1978/10/30 – 1980/06/01	NCAR
TOVS/na (NOAA 6)	1979/07/02 – 1983/04/17	NCAR
TOVS/nc (NOAA 7)	1981/07/11 - 1986/06/01	NCAR
TOVS/ne (NOAA 8)	1983/04/26 - 1985/01/01	NCAR
TOVS/nf (NOAA 9)	1985/01/01 - 1988/11/01	NOAA/NESDIS & NCAR
TOVS/ng (NOAA 10)	1986/11/25 - 1991/09/17	NOAA/NESDIS & NCAR
TOVS/nh (NOAA 11)	1988/09/02 - 1994/12/31	NOAA/NESDIS & NCAR
TOVS/nd (NOAA 12)	1991/08/18 - 1997/07/14	NOAA/NESDIS & NCAR
TOVS/nj (NOAA 14)	1995/01/19 -2006/10/10	NOAA/NESDIS
TOVS/nk (NOAA 15)	1998/07/01 - present	NOAA/NESDIS
TOVS/nl (NOAA 16)	2001/03/02 – 2010/04/14	NOAA/NESDIS
TOVS/nm (NOAA 17)	2003/03/01 -2011/12/11	NOAA/NESDIS
TOVS/nn (NOAA 18)	2005/10/05 – present	NOAA/NESDIS
EOS/Aqua	2002/10 - present	NOAA/NESDIS
SSM/I V6 (F08)	1987/7 - 1991/12/04	RSS
SSM/I V6 (F10)	1990/12 – 1997/11/13	RSS
SSM/I V6 (F11)	1991/12 – 1999/12/17	RSS
SSM/I V6 (F13)	1995/5 – 2009/11/19	RSS
SSM/I V6 (F14)	1997/5 – 2008/08/23	RSS
SSM/I V6 (F15)	1999/12 – 2006/07/25	RSS
GOES sounder T_B	2001/01 – 2007/12/04	NOAA/NCEP

Appendix C: Vertical Structure

Pressure-level data will be output on the following 42 pressure levels:

Level	P(hPa)	Level	P(hPa)	Level	P(hPa)	Level	P(hPa)	Level	P(hPa)	Level	P(hPa)
1	1000	8	825	15	600	22	250	29	30	36	2
2	975	9	800	16	550	23	200	30	20	37	1
3	950	10	775	17	500	24	150	31	10	38	0.7
4	925	11	750	18	450	25	100	32	7	39	0.5
5	900	12	725	19	400	26	70	33	5	40	0.4
6	875	13	700	20	350	27	50	34	4	41	0.3
7	850	14	650	21	300	28	40	35	3	42	0.1

Products on the native vertical grid will be output on the following levels. Pressures are nominal for a 1000 hPa surface pressure and refer to the top edge of the layer. Note that the bottom layer has a nominal thickness of 15 hPa.

Lev	P(hPa)	Lev	P(hPa)	Lev	P(hPa)	Lev	P(hPa)	Lev	P(hPa)	Lev	P(hPa)
1	0.0100	13	0.6168	25	9.2929	37	78.5123	49	450.000	61	820.000
2	0.0200	14	0.7951	26	11.2769	38	92.3657	50	487.500	62	835.000
3	0.0327	15	1.0194	27	13.6434	39	108.663	51	525.000	63	850.000
4	0.0476	16	1.3005	28	16.4571	40	127.837	52	562.500	64	865.000
5	0.0660	17	1.6508	29	19.7916	41	150.393	53	600.000	65	880.000
6	0.0893	18	2.0850	30	23.7304	42	176.930	54	637.500	66	895.000
7	0.1197	19	2.6202	31	28.3678	43	208.152	55	675.000	67	910.000
8	0.1595	20	3.2764	32	33.8100	44	244.875	56	700.000	68	925.000
9	0.2113	21	4.0766	33	40.1754	45	288.083	57	725.000	69	940.000
10	0.2785	22	5.0468	34	47.6439	46	337.500	58	750.000	70	955.000
11	0.3650	23	6.2168	35	56.3879	47	375.000	59	775.000	71	970.000
12	0.4758	24	7.6198	36	66.6034	48	412.500	60	800.000	72	985.000

Appendix D: GEOS-5 data collections

This section lists the variables in each data collection. More details on the variable definitions may be found in the GEOS-5 Variable Definition Glossary, available at the GMAO web page. In the tables, variable names refer to HDF names, which are uppercase.

D.1 Time-independent Variables

These are prescribed 2-dimensional fields that do not vary during the reanalysis. They are available as: **MERRA000.prod.assim.const_2d_asm_Nx.00000000.hdf** or **MERRA000.prod.assim.const_2d_mld_Nx.00000000.hdf** at the full resolution.

const_2d_asm_Nx

ECS short name:	**AC0NXASM**
ECS long name:	MERRA DAS 2d constants,
Characteristics:	**Constant at native resolution**
Dimensions:	longitude: **540**, latitude: **361**

Variable Name	Description	Units
PHIS	Surface geopotential	$m^2\ s^{-2}$
SGH	Standard deviation of topography for gravity wave drag	m
FRLAKE	Fraction of lake type in grid box	fraction
FRLAND	Fraction of land type in grid box	fraction
FRLANDICE	Fraction of land ice type in grid box	fraction
FROCEAN	Fraction of ocean in grid box	fraction
AREA	Area of grid box	m^2

const_2d_mld_Nx

ECS short name: **MSC0NXMLD**

ECS long name:	MERRA-Land 2d land surface constants
Name:	const_2d_mld_Nx
Characteristics:	**Constant at native resolution**
Dimensions:	longitude: **540**, latitude: **361**

Variable Name	Description	Units
DZSF	Thickness of soil layer associated with SFMC and GWETTOP	m
DZRZ	Thickness of soil layer associated with RZMC and GWETROOT	m
DZPR	Thickness of soil layer associated with PRMC and GWETPROF (“depth-to-bedrock” in the Catchment model)	m
DZTS	Thickness of soil layer associated with TSAT, TUNST, and TWLT	m
DZGT1	Thickness of soil layer associated with TSOIL1	m
DZGT2	Thickness of soil layer associated with TSOIL2	m
DZGT3	Thickness of soil layer associated with TSOIL3	m
DZGT4	Thickness of soil layer associated with TSOIL4	m
DZGT5	Thickness of soil layer associated with TSOIL5	m
DZGT6	Thickness of soil layer associated with TSOIL6	m
WPWET	Soil wilting point (in degree of saturation units)	Fraction
WPMC	Soil wilting point (in volumetric units)	$m^3\ m^{-3}$
WPEMW	Soil wilting point (in units of equivalent mass of total profile water)	$kg\ m^{-2}$

D.2 Analysis Collections

These are the fields resulting from the GSI analyses performed at the four synoptic times (states labeled ANALYSIS in Figure A.1). They are produced on the native horizontal grid and on both model levels and pressure levels in the vertical. The data on model levels are the values actually analyzed.

inst6_3d_ana_Nv

ECS short name:	**AI6NVANA**
ECS long name:	MERRA DAS 3d analyzed state,
Characteristics:	**Instantaneous, on model levels, at native resolution**
Dimensions:	longitude: **540**, latitude: **361**, levels: **72** (see Appendix C)
Times:	**00, 06, 12, 18 GMT**

Variable Name	Dims	Description	Units
PS	2D	Surface pressure	Pa
DELP	3D	Layer pressure thickness	Pa
T	3D	Air temperature	K
U	3D	Eastward wind component	$m\ s^{-1}$
V	3D	Northward wind component	$m\ s^{-1}$
QV	3D	Specific humidity	$kg\ kg^{-1}$
O3	3D	Ozone mixing ratio	$kg\ kg^{-1}$

inst6_3d_ana_Np

ECS short name: **AI6NPANA**
ECS long name: MERRA DAS 3d analyzed state on pressure,
Characteristics: **Instantaneous, on pressure levels, at native resolution**
Dimensions: longitude: **540**, latitude: **361**, pressure levels: **42** (see Appendix C)
Times: **00, 06, 12, 18 GMT; monthly and seasonal also available.**

Variable Name	Dims	Description	Units
SLP	2D	Sea-level pressure	Pa
PS	2D	Surface pressure	Pa
H	3D	Geopotential height	m
T	3D	Air temperature	K
U	3D	Eastward wind component	$m\ s^{-1}$
V	3D	Northward wind component	$m\ s^{-1}$
QV	3D	Specific humidity	$kg\ kg^{-1}$
O3	3D	Ozone mixing ratio	$kg\ kg^{-1}$

D.3 History Collections

These histories are produced from the GCM during the "corrector" segment of the IAU cycle. All collections in this group are at reduced horizontal resolution and 3-D fields are on pressure levels.

inst3_3d_asm_Cp

ECS short name: **AI3CPASM**
ECS long name: MERRA IAU 3d assimilated state on pressure,
Characteristics: **Instantaneous, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, pressure levels: **42** (see Appendix C)
Times: **00, 03, 06, 09, 12, 15, 18, 21 GMT**

Variable Name	Dims	Description	Units
SLP	2D	Sea-level pressure	Pa
PS	2D	Surface pressure	Pa
PHIS	2D	Surface Geopotential	$m^2 s^{-2}$
H	3D	Geopotential height	m
O3	3D	Ozone mixing ratio	$kg\ kg^{-1}$
QV	3D	Specific humidity	$kg\ kg^{-1}$
QL	3D	Cloud liquid water mixing ratio	$kg\ kg^{-1}$
QI	3D	Cloud ice mixing ratio	$kg\ kg^{-1}$
RH	3D	Relative humidity	fraction
T	3D	Air temperature	K
U	3D	Eastward wind component	$m\ s^{-1}$
V	3D	Northward wind component	$m\ s^{-1}$
EPV	3D	Ertel potential vorticity	$K\ m^2 kg^{-1} s^{-1}$
OMEGA	3D	Vertical pressure velocity	$Pa\ sec^{-1}$

tavg3_3d_cld_Cp

ECS short name: **AT3CPCLD**
ECS long name: MERRA IAU 3d cloud diagnostics,
Characteristics: **Time-averaged, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
RH	Relative humidity	fraction
QLLS	Cloud liquid water mixing ratio – large-scale	kg kg^{-1}
QILS	Cloud ice mixing ratio – large-scale	kg kg^{-1}
QLAN	Cloud liquid water mixing ratio – anvils	kg kg^{-1}
QIAN	Cloud ice mixing ratio – anvils	kg kg^{-1}
QCCU	Cloud condensate mixing ratio – convective updraft	kg kg^{-1}
CFLS	3-D Cloud fraction – large scale	fraction
CFAN	3-D Cloud fraction – anvils	fraction
CFCU	3-D Cloud fraction – convective	fraction

tavg3_3d_mst_Cp

ECS short name: **AT3CVMST**
ECS long name: MERRA IAU 3d moist processes diagnostic,
Characteristics: **Time-averaged, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
CMFMC	Upward moist convective mass flux	kg m^{-2} s^{-1}
DQRCU	Precipitation production rate – convective	kg kg^{-1} s^{-1}
DQRLSAN	Precipitation production rate - large-scale+anvil	kg kg^{-1} s^{-1}
PFLCU	Downward flux of liquid precipitation – convective	kg m^{-2} s^{-1}
PFICU	Downward flux of ice precipitation – convective	kg m^{-2} s^{-1}
PFLLSAN	Downward flux of liquid precip - large-scale+anvil	kg m^{-2} s^{-1}
PFILSAN	Downward flux of ice precip - large-scale+anvil	kg m^{-2} s^{-1}
REEVAPCN	Evaporation of precipitating convective condensate	kg kg^{-1} s^{-1}
REEVAPLSAN	Evaporation of precipitating LS &anvil condensate	kg kg^{-1} s^{-1}

tavg3_3d_rad_Cp

ECS short name:	**AT3CPRAD**
ECS long name:	MERRA IAU 3d radiation diagnostics,
Characteristics:	**Time-averaged, on pressure levels, at reduced resolution**
Dimensions:	longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times:	**1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	**Description**	**Units**
CLOUD	3-D Cloud fraction	fraction
DTDTLWR	T tendency from terrestrial radiation	$K\ s^{-1}$
DTDTLWRCLR	T tendency from terrestrial radiation (clear sky)	$K\ s^{-1}$
DTDTSWR	T tendency from solar radiation	$K\ s^{-1}$
DTDTSWRCLR	T tendency from solar radiation (clear sky)	$K\ s^{-1}$

tavg3_3d_trb_Cp

ECS short name:	**AT3CPTRB**
ECS long name:	MERRA IAU 3d turbulence diagnostics,
Characteristics:	**Time-averaged, on pressure levels, at reduced resolution**
Dimensions:	longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times:	**1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	**Description**	**Units**
KM	Momentum diffusivity	$m^2 s^{-1}$
KMLS	Momentum diffusivity from Louis	$m^2 s^{-1}$
KMLK	Momentum diffusivity from Lock	$m^2 s^{-1}$
KH	Heat (scalar) diffusivity	$m^2 s^{-1}$
KHLS	Heat (scalar) diffusivity from Louis	$m^2 s^{-1}$
KHLK	Heat (scalar) diffusivity from Lock	$m^2 s^{-1}$
KHRAD	Heat (scalar) diffusivity Lock radiative contribution	$m^2 s^{-1}$
KHSFC	Heat (scalar) diffusivity Lock surface contribution	$m^2 s^{-1}$
RI	Richardson Number	Nondimensional

tavg3_3d_tdt_Cp

ECS short name: **AT3CPTDT**
ECS long name: MERRA IAU 3d temperature tendencies,
Characteristics: **Time-averaged, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
DTDTRAD	Temperature tendency from radiation	$K\ s^{-1}$
DTDTMST	Temperature tendency from moist physics	$K\ s^{-1}$
DTDTTRB	Temperature tendency from turbulent mixing	$K\ s^{-1}$
DTDTFRI	Temperature tendency from turbulent friction	$K\ s^{-1}$
DTDTGWD	Temperature tendency from gravity wave friction	$K\ s^{-1}$
DTDTTOT	Temperature tendency from physics	$K\ s^{-1}$
DTDTDYN	Temperature tendency from dynamics	$K\ s^{-1}$
DTDTANA	Temperature tendency from analysis	$K\ s^{-1}$

tavg3_3d_udt_Cp

ECS short name:	**AT3CPUDT**
ECS long name:	MERRA IAU 3d eastward wind tendencies,
Characteristics:	**Time-averaged, on pressure levels, at reduced resolution**
Dimensions:	longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times:	**1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
DUDTMST	U-wind tendency from moist physics	$m\ s^{-2}$
DUDTTRB	U-wind tendency from turbulence	$m\ s^{-2}$
DUDTGWD	U-wind tend from gravity wave drag	$m\ s^{-2}$
DUDTDYN	U-wind tendency from dynamics	$m\ s^{-2}$
DUDTANA	U-wind tendency from analysis	$m\ s^{-2}$
DVDTMST	V-wind tendency from moist physics	$m\ s^{-2}$
DVDTTRB	V-wind tendency from turbulence	$m\ s^{-2}$
DVDTGWD	V-wind tend from gravity wave drag	$m\ s^{-2}$
DVDTDYN	V-wind tendency from dynamics	$m\ s^{-2}$
DVDTANA	V-wind tendency from analysis	$m\ s^{-2}$

tavg3_3d_qdt_Cp

ECS short name: **AT3CPQDT**
ECS long name: MERRA IAU 3d moisture tendencies,
Characteristics: **Time-averaged, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, pressure levels: **42** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
DQVDTMST	Water vapor tendency from moist physics	$kg\ kg^{-1}\ s^{-1}$
DQVDTTRB	Water vapor tendency from turbulence	$kg\ kg^{-1}\ s^{-1}$
DQVDTCHM	Water vapor tendency from chemistry	$kg\ kg^{-1}\ s^{-1}$
DQVDTDYN	Water vapor tendency from dynamics	$kg\ kg^{-1}\ s^{-1}$
DQVDTANA	Water vapor tendency from analysis	$kg\ kg^{-1}\ s^{-1}$
DQIDTMST	Ice tendency from moist physics	$kg\ kg^{-1}\ s^{-1}$
DQIDTTRB	Ice tendency from turbulence	$kg\ kg^{-1}\ s^{-1}$
DQIDTDYN	Ice tendency from dynamics	$kg\ kg^{-1}\ s^{-1}$
DQLDTMST	Liquid water tendency from moist physics	$kg\ kg^{-1}\ s^{-1}$
DQLDTTRB	Liquid tendency from turbulence	$kg\ kg^{-1}\ s^{-1}$
DQLDTDYN	Liquid tendency from dynamics	$kg\ kg^{-1}\ s^{-1}$

tavg3_3d_odt_Cp

ECS short name: **AT3CPODT**
ECS long name: MERRA IAU 3d ozone tendencies,
Characteristics: **Time-averaged, on pressure levels, at reduced resolution**
Dimensions: longitude: **288**, latitude: **144**, levels: **42** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
DOXDTMST	Ozone tendency from moist physics	$kg\ kg^{-1}\ s^{-1}$
DOXDTTRB	Ozone tendency from turbulence	$kg\ kg^{-1}\ s^{-1}$
DOXDTCHM	Ozone tendency from chemistry	$kg\ kg^{-1}\ s^{-1}$
DOXDTDYN	Ozone tendency from dynamics	$kg\ kg^{-1}\ s^{-1}$
DOXDTANA	Ozone tendency from analysis	$kg\ kg^{-1}\ s^{-1}$

tavg1_2d_slv_Nx

ECS short name: **AT1NXSLV**
ECS long name: MERRA IAU 2d atmospheric single-level diagnostics
Characteristics: **Time-averaged, single-level, at native resolution**
Dimensions: longitude: **540**, latitude: **361**
Times: **0:30, 1:30, 2:30, 3:30, 4:30, … GMT**

Variable Name	Description	Units
SLP	Sea level pressure	Pa
PS	Time averaged surface pressure	Pa
U850	Eastward wind at 850 hPa	$m\ s^{-1}$
U500	Eastward wind at 500 hPa	$m\ s^{-1}$
U250	Eastward wind at 250 hPa	$m\ s^{-1}$
V850	Northward wind at 850 hPa	$m\ s^{-1}$
V500	Northward wind at 500 hPa	$m\ s^{-1}$
V250	Northward wind at 250 hPa	$m\ s^{-1}$
T850	Temperature at 850 hPa	K
T500	Temperature at 500 hPa	K
T250	Temperature at 250 hPa	K
Q850	Specific humidity at 850 hPa	$kg\ kg^{-1}$
Q500	Specific humidity at 500 hPa	$kg\ kg^{-1}$
Q250	Specific humidity at 250 hPa	$kg\ kg^{-1}$
H1000	Height at 1000 hPa	m
H850	Height at 850 hPa	m
H500	Height at 500 hPa	m
H250	Height at 250 hPa	m
OMEGA500	Vertical pressure velocity at 500 hPa	$Pa\ s^{-1}$
U10M	Eastward wind at 10 m above displacement height	$m\ s^{-1}$
U2M	Eastward wind at 2 m above the displacement height	$m\ s^{-1}$
U50M	Eastward wind at 50 m above surface	$m\ s^{-1}$
V10M	Northward wind at 10 m above the displacement height	$m\ s^{-1}$
V2M	Northward wind at 2 m above the displacement height	$m\ s^{-1}$
V50M	Northward wind at 50 m above surface	$m\ s^{-1}$
T10M	Temperature at 10 m above the displacement height	K
T2M	Temperature at 2 m above the displacement height	K

Variable Name	Description	Units
QV10M	Specific humidity at 10 m above the displacement height	kg kg^{-1}
QV2M	Specific humidity at 2 m above the displacement height	kg kg^{-1}
TS	Surface skin temperature	K
DISPH	Displacement height	m
TROPPV	PV based tropopause pressure	Pa
TROPPT	T based tropopause pressure	Pa
TROPPB	Blended tropopause pressure	Pa
TROPT	Tropopause temperature	K
TROPQ	Tropopause specific humidity	kg kg^{-1}
CLDPRS	Cloud-top pressure	Pa
CLDTMP	Cloud-top temperature	K

tavg1_2d_flx_Nx

ECS short name: **AT1NXFLX**
ECS long name: MERRA IAU 2d surface turbulent flux diagnostics
Characteristics: **Time-averaged, single level, at native resolution**
Dimensions: longitude: **540**, latitude: **361**
Times: **0:30, 1:30, 2:30, 3:30, 4:30, … GMT**

Variable Name	Description	Units
EFLUX	Latent heat flux (positive upward)	W m^{-2}
EVAP	Surface evaporation	kg m^{-2} s^{-1}
HFLUX	Sensible heat flux (positive upward)	W m^{-2}
TAUX	Eastward surface wind stress	N m^{-2}
TAUY	Northward surface wind stress	N m^{-2}
TAUGWX	Eastward gravity wave surface stress	N m^{-2}
TAUGWY	Northward gravity wave surface stress	N m^{-2}
PBLH	Planetary boundary layer height	m
DISPH	Displacement height	m
BSTAR	Surface buoyancy scale	m s^{-1}
USTAR	Surface velocity scale	m s^{-1}
TSTAR	Surface temperature scale	K
QSTAR	Surface humidity scale	kg kg^{-1}
RI	Surface Richardson number	nondimensional
Z0H	Roughness length, sensible heat	m
Z0M	Roughness length, momentum	m
HLML	Height of center of lowest model layer	m
TLML	Temperature of lowest model layer	K
QLML	Specific humidity of lowest model layer	kg kg^{-1}
ULML	Eastward wind of lowest model layer	m s^{-1}
VLML	Northward wind of lowest model layer	m s^{-1}
RHOA	Surface air density	kg m^{-3}
SPEED	3-dimensional wind speed for surface fluxes	m s^{-1}
CDH	Surface exchange coefficient for heat	kg m^{-2} s^{-1}
CDQ	Surface exchange coefficient for moisture	kg m^{-2} s^{-1}
CDM	Surface exchange coefficient for momentum	kg m^{-2} s^{-1}
CN	Surface neutral drag coefficient	nondimensional

Variable Name	**Description**	**Units**
TSH	Effective turbulence skin temperature	K
QSH	Effective turbulence skin humidity	kg kg^{-1}
FRSEAICE	Fraction of sea-ice	Fraction
PRECANV	Surface precipitation flux from anvils	kg m^{-2} s^{-1}
PRECCON	Surface precipitation flux from convection	kg m^{-2} s^{-1}
PRECLSC	Surface precipitation flux from large-scale	kg m^{-2} s^{-1}
PRECSNO	Surface snowfall flux	kg m^{-2} s^{-1}
PRECTOT	Total surface precipitation flux	kg m^{-2} s^{-1}
PGENTOT	Total generation of precipitation	kg m^{-2} s^{-1}
PREVTOT	Total re-evaporation of precipitation	kg m^{-2} s^{-1}

tavg1_2d_rad_Nx

ECS short name: **AT1NXRAD**
ECS long name: MERRA IAU 2d surface and TOA radiation fluxes
Characteristics: **Time-averaged, single level, at native resolution**
Dimensions : longitude : **540**, latitude : **361**
Times: **0 :30, 1 :30, 2 :30, 3 :30, 4 :30, … GMT**

Variable Name	Description	Units
EMIS	Surface emissivity	fraction
TS	Surface skin temperature	K
ALBEDO	Surface albedo	fraction
ALBNIRDF	Diffuse beam NIR surface albedo	fraction
ALBNIRDR	Direct beam NIR surface albedo	fraction
ALBVISDF	Diffuse beam VIS-UV surface albedo	fraction
ALBVISDR	Direct beam VIS-UV surface albedo	fraction
LWGEM	Emitted longwave at the surface	W m^{-2}
LWGAB	Surface absorbed longwave	W m^{-2}
LWGABCLR	Surface absorbed longwave assuming clear sky	W m^{-2}
LWGABCLRCLN	Surface absorbed longwave assuming clear clean sky	W m^{-2}
LWGNT	Surface net downward longwave flux	W m^{-2}
LWGNTCLR	Surface net downward longwave flux assuming clear sky	W m^{-2}
LWGNTCLRCLN	Surface net downward longwave flux assuming clear clean sky	W m^{-2}
LWTUP	Top of atmosphere (TOA) upward longwave flux	W m^{-2}
LWTUPCLR	TOA upward longwave flux assuming clear sky	W m^{-2}
LWTUPCLRCLN	TOA upward longwave flux assuming clear clean sky	W m^{-2}
SWTDN	TOA incident shortwave flux	W m^{-2}
SWGDN	Surface incident shortwave flux	W m^{-2}
SWGDNCLR	Surface incident shortwave flux assuming clear sky	W m^{-2}
SWGNT	Surface net downward shortwave flux	W m^{-2}
SWGNTCLR	Surface net downward shortwave flux assuming clear sky	W m^{-2}
SWGNTCLN	Surface net downward shortwave flux assuming clean sky	W m^{-2}
SWGNTCLRCLN	Surface net downward shortwave flux assuming clear clean sky	W m^{-2}
SWTNT	TOA net downward shortwave flux	W m^{-2}
SWTNTCLR	TOA net downward shortwave flux assuming clear sky	W m^{-2}

Variable Name	Description	Units
SWTNTCLN	TOA net downward shortwave flux assuming clean sky	W m^{-2}
SWTNTCLRCLN	TOA net downward shortwave flux assuming clear clean sky	W m^{-2}
TAUHGH	Optical thickness of high clouds	dimensionless
TAULOW	Optical thickness of low clouds	dimensionless
TAUMID	Optical thickness of mid-level clouds	dimensionless
TAUTOT	Optical thickness of all clouds	dimensionless
CLDHGH	High-level (above 400 hPa) cloud fraction	fraction
CLDLOW	Low-level (1000-700 hPa) cloud fraction	fraction
CLDMID	Mid-level (700-400 hPa) cloud fraction	fraction
CLDTOT	Total cloud fraction	fraction

tavg1_2d_lnd_Nx

ECS short name: **AT1NXLND**
ECS long name: MERRA IAU 2d land surface diagnostics
Characteristics: **Time-averaged, single level, at native resolution**
Dimensions: longitude: **540**, latitude: **361**
Times: **0:30, 1:30, 2:30, 3:30, 4:30, … GMT**

Variable Name	Description	Units
GRN	Vegetation greenness fraction	Fraction
LAI	Leaf area index	$m^2\ m^{-2}$
GWETROOT	Root zone soil wetness	fraction
GWETTOP	Top soil layer wetness	fraction
TPSNOW	Top snow layer temperature	K
TUNST	Surface temperature of unsaturated zone	K
TSAT	Surface temperature of saturated zone	K
TWLT	Surface temperature of wilted zone	K
PRECSNO	Surface snowfall	$kg\ m^{-2}\ s^{-1}$
PRECTOT	Total surface precipitation	$kg\ m^{-2}\ s^{-1}$
SNOMAS	Snow mass	$kg\ m^{-2}$
SNODP	Snow depth	m
EVPSOIL	Bare soil evaporation	$W\ m^{-2}$
EVPTRNS	Transpiration	$W\ m^{-2}$
EVPINTR	Interception loss	$W\ m^{-2}$
EVPSBLN	Sublimation	$W\ m^{-2}$
RUNOFF	Overland runoff	$kg\ m^{-2}\ s^{-1}$
BASEFLOW	Baseflow	$kg\ m^{-2}\ s^{-1}$
SMLAND	Snowmelt	$kg\ m^{-2}\ s^{-1}$
FRUNST	Fractional unsaturated area	fraction
FRSAT	Fractional saturated area	fraction
FRSNO	Fractional snow-covered area	fraction
FRWLT	Fractional wilting area	fraction
PARDF	Surface downward PAR diffuse flux	$W\ m^{-2}$
PARDR	Surface downward PAR beam flux	$W\ m^{-2}$
SHLAND	Sensible heat flux from land	$W\ m^{-2}$
LHLAND	Latent heat flux from land	$W\ m^{-2}$

Variable Name	Description	Units
EVLAND	Evaporation from land	kg m^{-2} s^{-1}
LWLAND	Net downward longwave flux over land	W m^{-2}
SWLAND	Net downward shortwave flux over land	W m^{-2}
GHLAND	Downward heat flux at base of top soil layer	W m^{-2}
TWLAND	Total water store in land reservoirs	kg m^{-2}
TELAND	Energy store in all land reservoirs	J m^{-2}
WCHANGE	Total land water change per unit time	kg m^{-2} s^{-1}
ECHANGE	Total land energy change per unit time	W m^{-2}
SPLAND	Spurious land energy source	W m^{-2}
SPWATR	Spurious land water source	kg m^{-2} s^{-1}
SPSNOW	Spurious snow source	kg m^{-2} s^{-1}

tavg1_2d_mld_Nx

ECS short name:	**MST1NXMLD**
ECS long name:	MERRA-Land 2d land surface diagnostics
Name:	tavg1_2d_mld_Nx
Characteristics:	**Time-averaged, single level, at native resolution**
Dimensions:	longitude: **540**, latitude: **361**
Times:	**0:30, 1:30, 2:30, 3:30, 4:30, … GMT**

Variable Name	Description	Units
GRN	Vegetation greenness fraction (LAI-weighted)	Fraction
LAI	Leaf area index	$m^2\ m^{-2}$
GWETPROF*	Total profile soil wetness	Fraction
GWETROOT	Root zone soil wetness	Fraction
GWETTOP	Top soil layer wetness	Fraction
PRMC*	Total profile soil moisture content	$m^3\ m^{-3}$
RZMC*	Root zone soil moisture content	$m^3\ m^{-3}$
SFMC*	Top soil layer soil moisture content	$m^3\ m^{-3}$
TSURF*	Mean land surface temperature (incl. snow)	K
TPSNOW	Top snow layer temperature	K
TUNST	Surface temperature of unsaturated (but non-wilting) zone	K
TSAT	Surface temperature of saturated zone	K
TWLT	Surface temperature of wilting zone	K
TSOIL1*	Soil temperature in layer 1	K
TSOIL2*	Soil temperature in layer 2	K
TSOIL3*	Soil temperature in layer 3	K
TSOIL4*	Soil temperature in layer 4	K
TSOIL5*	Soil temperature in layer 5	K
TSOIL6*	Soil temperature in layer 6	K
PRECSNO	Surface snowfall	$kg\ m^{-2}\ s^{-1}$
PRECTOT	Total surface precipitation	$kg\ m^{-2}\ s^{-1}$
SNOMAS	Snow mass	$kg\ m^{-2}$
SNODP	Snow depth	m
EVPSOIL	Bare soil evaporation	$W\ m^{-2}$
EVPTRNS	Transpiration	$W\ m^{-2}$
EVPINTR	Interception loss	$W\ m^{-2}$
EVPSBLN	Sublimation	$W\ m^{-2}$

Variable Name	Description	Units
RUNOFF	Overland runoff	kg m^{-2} s^{-1}
BASEFLOW	Baseflow	kg m^{-2} s^{-1}
SMLAND	Snowmelt over land	kg m^{-2} s^{-1}
QINFIL*	Soil water infiltration rate	kg m^{-2} s^{-1}
FRUNST	Fractional unsaturated (but non-wilting) area	Fraction
FRSAT	Fractional saturated area	Fraction
FRSNO	Fractional snow-covered area	Fraction
FRWLT	Fractional wilting area	Fraction
PARDF	Surface downward PAR** diffuse flux	W m^{-2}
PARDR	Surface downward PAR** beam flux	W m^{-2}
SHLAND	Sensible heat flux from land	W m^{-2}
LHLAND	Latent heat flux from land	W m^{-2}
EVLAND	Evaporation from land	kg m^{-2} s^{-1}
LWLAND	Net downward longwave flux over land	W m^{-2}
SWLAND	Net downward shortwave flux over land	W m^{-2}
GHLAND	Downward heat flux into top soil layer	W m^{-2}
TWLAND	Total water stored in land reservoirs	kg m^{-2}
TELAND	Energy stored in all land reservoirs	J m^{-2}
WCHANGE	Total land water change per unit time	kg m^{-2} s^{-1}
ECHANGE	Total land energy change per unit time	W m^{-2}
SPLAND	Spurious land energy source	W m^{-2}
SPWATR	Spurious land water source	kg m^{-2} s^{-1}
SPSNOW	Spurious snow energy source	W m^{-2}

**Denotes variables that are new in MERRA-Land and not available in the MERRA "lnd" collection.*
***PAR = Photosynthetically Active Radiation*

tavg1_2d_ocn_Nx

ECS short name:	**AT1NXOCN**
ECS long name:	MERRA IAU 2d ocean surface diagnostics
Characteristics:	**Time-averaged, single level, at native resolution**
Dimensions:	longitude: **540**, latitude: **361**
Times:	**0:30, 1:30, 2:30, 3:30, 4:30, … GMT**

Variable Name	Description	Units
U10M	Eastward wind at 10 m above displacement height	$m\ s^{-1}$
V10M	Northward wind at 10 m above the displacement height	$m\ s^{-1}$
T10M	Temperature at 10 m above the displacement height	K
QV10M	Specific humidity at 10 m above the displacement height	$kg\ kg^{-1}$
HFLUXWTR	Open water upward sensible heat flux	$W\ m^{-2}$
HFLUXICE	Sea ice upward sensible heat flux	$W\ m^{-2}$
EFLUXWTR	Open water latent heat (energy) flux	$W\ m^{-2}$
EFLUXICE	Sea ice latent heat (energy) flux	$W\ m^{-2}$
LWGNTWTR	Open water net downward longwave flux	$W\ m^{-2}$
LWGNTICE	Sea ice net downward longwave flux	$W\ m^{-2}$
SWGNTWTR	Open water net downward shortwave flux	$W\ m^{-2}$
SWGNTICE	Sea ice net downward shortwave flux	$W\ m^{-2}$
PRECSNOOCN	Snowfall over ocean	$kg\ m^{-2}\ s^{-1}$
RAINOCN	Rainfall over ocean	$kg\ m^{-2}\ s^{-1}$
TAUXWTR	Eastward component of surface stress over open water	$N\ m^{-2}$
TAUYWTR	Northward component of surface stress over open water	$N\ m^{-2}$
TAUXICE	Eastward component of surface stress over sea ice	$N\ m^{-2}$
TAUYICE	Northward component of surface stress over sea ice	$N\ m^{-2}$
FRSEAICE	Fraction of ocean covered by sea ice	1

tavg1_2d_int_Nx

ECS short name: **AT1NXINT**
ECS long name: MERRA IAU 2d Vertical integrals
Characteristics: **Time-averaged, single level, at native resolution**
Dimensions : longitude : **540**, latitude : **361**
Times: **0 :30, 1 :30, 2 :30, 3 :30, 4 :30, … GMT**

Variable Name	Description	Units
DMDT_ANA	Vertically integrated atmospheric mass tendency for analysis	kg m^{-2} s^{-1}
DMDT_DYN	Vertically integrated atmospheric mass tendency for dynamics	kg m^{-2} s^{-1}
DQVDT_DYN	Vertically integrated water tendency for dynamics	kg m^{-2} s^{-1}
DQVDT_PHY	Vertically integrated water tendency for physics	kg m^{-2} s^{-1}
DQVDT_MST	Vertically integrated water tendency for moist	kg m^{-2} s^{-1}
DQVDT_TRB	Vertically integrated water tendency for turbulence	kg m^{-2} s^{-1}
DQVDT_CHM	Vertically integrated water tendency for chemistry	kg m^{-2} s^{-1}
DQVDT_ANA	Vertically integrated water tendency for analysis	kg m^{-2} s^{-1}
DQLDT_DYN	Vertically integrated liquid water tendency for dynamics	kg m^{-2} s^{-1}
DQLDT_PHY	Vertically integrated liquid water tendency for physics	kg m^{-2} s^{-1}
DQLDT_ANA	Vertically integrated liquid water tendency for analysis	kg m^{-2} s^{-1}
DQLDT_MST	Vertically integrated liquid water tendency for moist	kg m^{-2} s^{-1}
DQIDT_DYN	Vertically integrated ice water tendency for dynamics	kg m^{-2} s^{-1}
DQIDT_PHY	Vertically integrated ice water tendency for physics	kg m^{-2} s^{-1}
DQIDT_ANA	Vertically integrated ice water tendency for analysis	kg m^{-2} s^{-1}
DQIDT_MST	Vertically integrated ice water tendency for moist	kg m^{-2} s^{-1}
DOXDT_DYN	Vertically integrated total ozone tendency for dynamics	kg m^{-2} s^{-1}
DOXDT_PHY	Vertically integrated total ozone tendency for physics	kg m^{-2} s^{-1}
DOXDT_CHM	Vertically integrated total ozone tendency for chemistry	kg m^{-2} s^{-1}
DOXDT_ANA	Vertically integrated total ozone tendency for analysis	kg m^{-2} s^{-1}
DKDT_DYN	Vertically integrated kinetic energy tendency for dynamics	W m^{-2}
DKDT_PHY	Vertically integrated kinetic energy tendency for physics	W m^{-2}

Variable Name	Description	Units
DKDT_ANA	Vertically integrated kinetic energy tendency for analysis	W m^{-2}
DKDT_PHYPHY	Kinetic energy tendency as computed in physics	W m^{-2}
DHDT_DYN	Vertically integrated cpT tendency for dynamics	W m^{-2}
DHDT_PHY	Vertically integrated cpT tendency for physics	W m^{-2}
DHDT_ANA	Vertically integrated cpT tendency for analysis	W m^{-2}
DHDT_RES	Residual cpT tendency	W m^{-2}
DPDT_DYN	Potential energy tendency for dynamics	W m^{-2}
DPDT_PHY	Potential energy tendency for physics	W m^{-2}
DPDT_ANA	Potential energy tendency for analysis	W m^{-2}
UFLXCPT	Vertically integrated eastward flux of dry enthalpy	J m^{-1} s^{-1}
VFLXCPT	Vertically integrated northward flux of dry enthalpy	J m^{-1} s^{-1}
UFLXPHI	Vertically integrated eastward flux of geopotential	J m^{-1} s^{-1}
VFLXPHI	Vertically integrated northward flux of geopotential	J m^{-1} s^{-1}
UFLXKE	Vertically integrated eastward flux of kinetic energy	J m^{-1} s^{-1}
VFLXKE	Vertically integrated northward flux of kinetic energy	J m^{-1} s^{-1}
UFLXQV	Vertically integrated eastward flux of specific humidity	kg m^{-1} s^{-1}
VFLXQV	Vertically integrated northward flux of specific humidity	kg m^{-1} s^{-1}
UFLXQL	Vertically integrated eastward flux of liquid condensate	kg m^{-1} s^{-1}
VFLXQL	Vertically integrated northward flux of liquid condensate	kg m^{-1} s^{-1}
UFLXQI	Vertically integrated eastward flux of ice condensate	kg m^{-1} s^{-1}
VFLXQI	Vertically integrated northward flux of ice condensate	kg m^{-1} s^{-1}
CONVCPT	Vertically integrated convergence of dry enthalpy	W m^{-2}
CONVPHI	Vertically integrated convergence of geopotential	W m^{-2}
CONVKE	Vertically integrated convergence of kinetic energy	W m^{-2}
CONVTHV	Vertically integrated convergence of potential temperature	K
TEFIXER	Total energy added by artificial energy “fixer”	W m^{-2}
DKDT_GEN	Generation of kinetic energy	W m^{-2}
DKDT_PG	Kinetic energy tendency due to pressure gradient force	W m^{-2}
DKDT_REMAP	Kinetic energy tendency due to remapping (spurious)	W m^{-2}
DKDT_INRES	Kinetic energy tendency residual from inertial terms (spurious)	W m^{-2}

Variable Name	Description	Units
DKDT_PGRES	Kinetic energy tendency residual from pressure terms (spurious)	W m^{-2}
DKDT_GWD	Kinetic energy tendency due to gravity wave drag (GWD)	W m^{-2}
DKDT_RAY	Kinetic energy tendency due to Rayleigh friction	W m^{-2}
DKDT_BKG	Kinetic energy tendency due to background GWD	W m^{-2}
DKDT_ORO	Kinetic energy tendency due to orographic GWD	W m^{-2}
DKDT_GWDRES	Kinetic energy residual due to errors in GWD (spurious)	W m^{-2}
BKGERR	Energy residual due to errors in background GWD (spurious)	W m^{-2}
DKDT_TRB	Kinetic energy tendency due to turbulence	W m^{-2}
DKDT_SRF	Kinetic energy tendency due to surface friction	W m^{-2}
DKDT_INT	Kinetic energy tendency due to internal friction	W m^{-2}
DKDT_TOP	Kinetic energy tendency due to topographic low-level drag	W m^{-2}
DKDT_MST	Kinetic energy tendency due to moist processes	W m^{-2}
DHDT_REMAP	Virtual enthalpy change due to remapping	W m^{-2}
DHDT_GWD	Virtual enthalpy change due to all gravity wave drag processes	W m^{-2}
DHDT_RAY	Virtual enthalpy change due to Rayleigh friction	W m^{-2}
DHDT_BKG	Virtual enthalpy change due to background gravity wave drag	W m^{-2}
DHDT_ORO	Virtual enthalpy change due to orographic gravity wave drag	W m^{-2}
DHDT_TRB	Virtual enthalpy change due to all turbulent	W m^{-2}
DHDT_MST	Virtual enthalpy change due to all moist processes	W m^{-2}
DHDT_FRI	Virtual enthalpy change due to all frictional processes	W m^{-2}
DHDT_RAD	Virtual enthalpy change due to radiation	W m^{-2}
DHDT_CUF	Virtual enthalpy change due to cumulus friction	W m^{-2}
DPDT_REMAP	Potential energy change due to remapping (spurious)	W m^{-2}
QTFILL	Artificial “filling” of total water	kg m^{-2} s^{-1}
DQVDT_FIL	Artificial “filling” of water vapor	kg m^{-2} s^{-1}
DQIDT_FIL	Artificial “filling” of frozen water	kg m^{-2} s^{-1}
DQLDT_FIL	Artificial “filling” of liquid water	kg m^{-2} s^{-1}
DOXDT_FIL	Artificial “filling” of odd oxygen	kg m^{-2} s^{-1}

Variable Name	**Description**	**Units**
HFLUX	Upward turbulent flux of sensible heat at the surface	W m^{-2}
EVAP	Upward turbulent flux of water vapor at the surface	kg m^{-2} s^{-1}
PRECCU	Liquid precipitation from convection at the surface	kg m^{-2} s^{-1}
PRECLS	Liquid precipitation from large scale processes at the surface	kg m^{-2} s^{-1}
PRECSN	Frozen precipitation at the surface	kg m^{-2} s^{-1}
DTHDT_ANA	Virtual potential tendency due to analysis	K s^{-1}
DTHDT_PHY	Virtual potential tendency due to physics	K s^{-1}
DTHDT_DYN	Virtual potential tendency due to dynamics	K s^{-1}
DTHDT_REMAP	Virtual potential tendency due to dynamics remapping	K s^{-1}
DTHDT_CONSV	Virtual potential tendency due to dynamics conservation	K s^{-1}
DTHDT_FIL	Virtual potential tendency due to dynamics water filling	K s^{-1}
LWTNET	Net Downward longwave radiation at the top of the atmosphere	W m^{-2}
LWGNET	Net Downward longwave radiation at the surface	W m^{-2}
SWNETTOA	Net Downward shortwave radiation at the top of the atmosphere	W m^{-2}
SWNETSRF	Net Downward shortwave radiation at the surface	W m^{-2}
LSCNVCL	Large-scale conversion of water vapor to cloud liquid	kg m^{-2} s^{-1}
LSCNVCI	Large-scale conversion of water vapor to cloud ice	kg m^{-2} s^{-1}
LSCNVRN	Large-scale conversion of water vapor to liquid precipitation	kg m^{-2} s^{-1}
CUCNVCL	Convective conversion of water vapor to cloud liquid	kg m^{-2} s^{-1}
CUCNVCI	Convective conversion of water vapor to cloud ice	kg m^{-2} s^{-1}
CUCNVRN	Convective conversion of water vapor to liquid precipitation	kg m^{-2} s^{-1}
EVPCL	Evaporation of cloud liquid water	kg m^{-2} s^{-1}
EVPRN	Evaporation of rain liquid water	kg m^{-2} s^{-1}
SUBCI	Sublimation of cloud ice	kg m^{-2} s^{-1}
SUBSN	Sublimation of frozen precipitation	kg m^{-2} s^{-1}
AUTCNVRN	Autoconversion of cloud liquid water to liquid precipitation	kg m^{-2} s^{-1}
SDMCI	Sedimentation of cloud ice	kg m^{-2} s^{-1}
COLCNVRN	Conversion of cloud liquid water to rain through collection	kg m^{-2} s^{-1}
COLCNVSN	Conversion of cloud liquid water to snow through collection	kg m^{-2} s^{-1}
FRZCL	Net freezing of cloud water	kg m^{-2} s^{-1}

Variable Name	Description	Units
FRZRN	Net freezing of rain water	kg m^{-2} s^{-1}

inst1_2d_int_Nx

ECS short name: **AI1NXINT**
ECS long name: MERRA IAU 2d Vertical integrals
Characteristics: **Instantaneous, single level, at native resolution**
Dimensions: longitude: **540**, latitude: **361**
Times: **0, 1, 2, 3, 4, …,23 GMT**

Variable Name	Description	Units
TQV	Total water vapor	kg m^{-2}
TQI	Total cloud ice water	kg m^{-2}
TQL	Total cloud liquid water	kg m^{-2}
TOX	Total column odd oxygen	kg m^{-2}
MASS	Atmospheric mass	kg m^{-2}
KE	Kinetic Energy	J m^{-2}
CPT	Dry enthalpy	J m^{-2}
THV	Virtual potential temperature	K

D.4 Chemistry Forcing Files

These histories are intended for forcing off-line chemistry/aerosol models with the results of the reanalysis. Like other histories, they are produced from the GCM during the "corrector" segment of the IAU cycle. As explained in Appendix A, such results represent a continuous evolution of the model's state, without the shocks of data insertion and initialization. By doing it in this way, we can provide consistent data at intervals finer than the six-hour analysis interval. In order to provide the most accurate forcing for off-line calculations, these forcings are produced at 3-hourly intervals and on the native vertical grid using a reduced horizontal resolution version of the model's horizontal grid. This horizontal grid was chosen for backward compatibility with earlier products. In addition, the last three streams in this section consist of those fields needed to reproduce as closely as possible the model's large-scale transport calculation. These are on the 3-dimensional C-grid used by the model and can reproduce, within time truncation, the model's continuity equation. These native streams are fairly specialized and are intended primarily for users of Goddard's Chemistry Transport Model or its Finite-Volume advection module.

const_2d_chm_Fx

ECS short name:	**AC0FXCNS**
ECS long name:	MERRA CHM 2d constants
Characteristics:	**Constant at reduced FV resolution**
Dimensions:	longitude: **288**, latitude: **181**

Variable Name	Description	Units
PHIS	Surface geopotential	$m^2 s^{-2}$
SGH	Standard deviation of topography for gravity wave drag	m
FRLAKE	Fraction of lake type in grid box	fraction
FRLAND	Fraction of land type in grid box	fraction
FRLANDICE	Fraction of land ice type in grid box	fraction
FROCEAN	Fraction of ocean in grid box	fraction

tavg3_3d_chm_Fv

ECS short name: **AT3FVCHM**
ECS long name: MERRA IAU 3d Chem On Layers
Characteristics: **Time-averaged, 3D model levels, at reduced FV resolution**
Dimensions: longitude: **288**, latitude: **181**, levels: **72** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
DELP	Layer pressure thickness	Pa
T	Air temperature	K
QV	Specific humidity	kg kg^{-1}
QL	Cloud liquid water mixing ratio	kg kg^{-1}
QI	Cloud ice mixing ratio	kg kg^{-1}
U	Eastward component of wind	m s^{-1}
V	Northward component of wind	m s^{-1}
CFLS	Large-scale cloud fraction	fraction
CFAN	Anvil cloud fraction	fraction
CFCU	Convective cloud fraction	fraction
DQRCON	Precipitation production rate – convective	kg kg^{-1} s^{-1}
DQRLSC	Precipitation production rate - large-scale	kg kg^{-1} s^{-1}
DQRANV	Precipitation production rate - anvils	kg kg^{-1} s^{-1}
DTRAIN	Detrainment cloud mass flux	kg m^{-2} s^{-1}
TAUCLI	Layer ice cloud optical thickness	nondimensional
TAUCLW	Layer liquid cloud optical thickness	nondimensional

tavg3_3d_chm_Fe

ECS short name:	**AT3FECHM**
ECS long name:	MERRA IAU 3d Chem On Layer Edges
Characteristics:	**Time-averaged, 3D model levels, at reduced FV resolution**
Dimensions:	longitude: **288**, latitude: **181**, levels: **73** (see Appendix C)
Times:	**1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
MFZ	Upward resolved mass flux	$kg\ m^{-2}\ s^{-1}$
CMFMC	Upward moist convective mass flux	$kg\ m^{-2}\ s^{-1}$
PFLLSAN	Liquid large-scale plus anvil precipitation	$kg\ m^{-2}\ s^{-1}$
PFILSAN	Ice large-scale plus anvil precipitation	$kg\ m^{-2}\ s^{-1}$
PFLCU	Liquid convective precipitation	$kg\ m^{-2}\ s^{-1}$
PFICU	Ice convective precipitation	$kg\ m^{-2}\ s^{-1}$
KH	Total scalar diffusivity	$m^2\ s^{-1}$

tavg3_2d_chm_Fx

ECS short name: **AT3FXCHM**
ECS long name: MERRA IAU 2d Chem
Characteristics: **Time-averaged, single-level, at reduced FV resolution**
Dimensions: longitude: **288**, latitude: **181** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
PRECANV	Surface precipitation flux from anvils	$kg\ m^{-2}\ s^{-1}$
PRECCON	Surface precipitation flux from convection	$kg\ m^{-2}\ s^{-1}$
PRECLSC	Surface precipitation flux from large-scale	$kg\ m^{-2}\ s^{-1}$
FRCLS	Fractional area of large-scale precipitation	fraction
FRCAN	Fractional area of anvil precipitation	fraction
FRCCN	Fractional area of convective precipitation	fraction
PRECSNO	Surface snowfall flux	$kg\ m^{-2}\ s^{-1}$
TS	Surface skin temperature	K
QV2M	Specific humidity 2 m above displacement height	$kg\ kg^{-1}$
T2M	Temperature 2 m above displacement height	K
U10M	Eastward wind 10 m above displacement height	$m\ s^{-1}$
V10M	Northward wind 10 m above displacement height	$m\ s^{-1}$
PARDF	Surface downward PAR diffuse flux	$W\ m^{-2}$
PARDR	Surface downward PAR beam flux	$W\ m^{-2}$
NIRDF	Surface downward NIR diffuse flux	$W\ m^{-2}$
NIRDR	Surface downward NIR beam flux	$W\ m^{-2}$
LWGNET	Surface net downward longwave flux	$W\ m^{-2}$
SWGNET	Net surface downward shortwave flux	$W\ m^{-2}$
LWGDWN	Surface downward longwave flux	$W\ m^{-2}$
SWGDWN	Surface downward shortwave flux	$W\ m^{-2}$
EVAP	Surface evaporation	$kg\ m^{-2}\ s^{-1}$
HFLUX	Sensible heat flux (positive upward)	$W\ m^{-2}$
GWETROOT	Root zone soil wetness	fraction
GWETTOP	Top soil layer wetness	fraction
CLDHGH	High-level (above 400 hPa) cloud fraction	fraction
CLDLOW	Low-level (1000-700 hPa) cloud fraction	fraction
CLDMID	Mid-level (700-400 hPa) cloud fraction	fraction
CLDTOT	Total cloud fraction	fraction

Variable Name	Description	Units
PBLH	Planetary boundary layer height above surface	m
PBLTOP	Pressure at PBL top	Pa
PS	Surface pressure	Pa

tavg3_3d_chm_Nv

ECS short name: **AT3NVCHM**
ECS long name: MERRA IAU 3d Chem On Layers
Characteristics: **Time-averaged, 3D model levels, at native resolution**
Dimensions: longitude: **540**, latitude: **361**, model levels: **72** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
MFXC	Eastward mass flux on C-Grid	Pa m s^{-1}
MFYC	Northward mass flux on C-Grid	Pa m s^{-1}
UC	Eastward component of wind on C-Grid	m s^{-1}
VC	Northward component of wind on C-Grid	m s^{-1}
DTRAIN	Detrainment cloud mass flux	kg m^{-2} s^{-1}

tavg3_3d_chm_Ne

ECS short name: **AT3NECHM**
ECS long name: MERRA IAU 3d Chem On Layer Edges
Characteristics: **Time-averaged, 3D model levels, at native resolution**
Dimensions: longitude: **540**, latitude: **361**, model levels: **73** (see Appendix C)
Times: **1:30, 4:30, 7:30, 10:30, 13:30, 16:30, 19:30, 22:30 GMT**

Variable Name	Description	Units
MFZ	Upward mass flux on C-Grid	kg m^{-2} s^{-1}
CMFMC	Upward moist convective mass flux	kg m^{-2} s^{-1}
KH	Total scalar diffusivity	m^2 s^{-1}

inst3_3d_chm_Ne

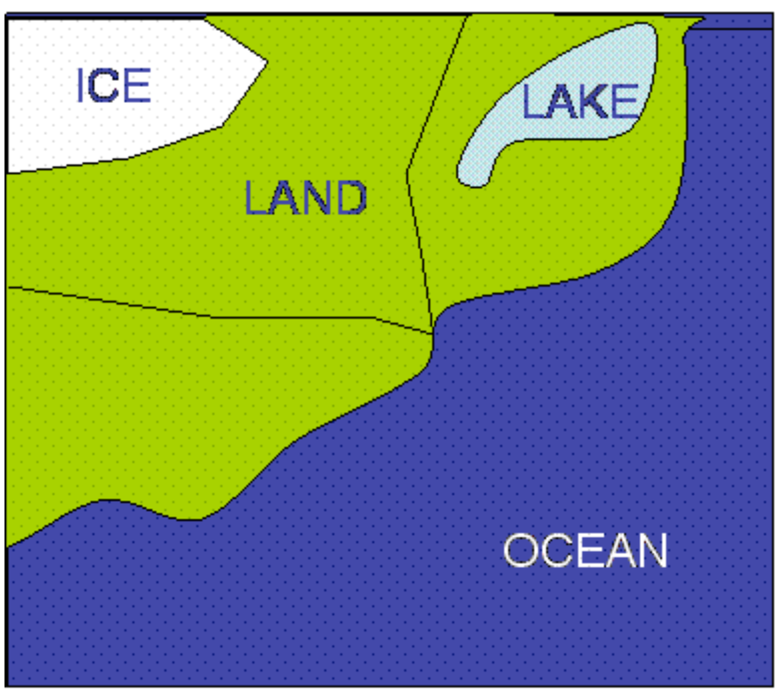
ECS short name: **AI3NECHM**
ECS long name: MERRA IAU 3d Chem On Layer Edges
Characteristics: **Instantaneous, 3D model levels, at native resolution**
Dimensions: longitude: **540**, latitude: **361**, model levels: **73** (see Appendix C)
Times: **0, 3, 6, 9, 12, 15, 18, 21 -GMT**

Variable Name	Description	Units
PLE	Edge pressures	Pa

Appendix E: Surface Representation

In GEOS-5 the surface below each atmospheric column consists of a set of tiles that represent various surface types. Tiles can be of four different types: Ocean, Land, (land) Ice, or Lake, as illustrated in the figure below. In each grid box a single (land) Ice tile represents those areas covered by permanent ice. Similarly a single Lake tile represents continental areas covered permanently by water. Other continental areas (non Lake or Ice) can be further subdivided into tiles that represent parts of the grid box in different hydrological catchments, defined according to the Pfafstetter (1989) system. Each of these is, in turn, divided into sub-tiles (not shown in the figure) that represent the wilting, unsaturated (but non-wilting), saturated, and snow-covered fractions of the tile. These sub-tile fractions vary with time and are predicted by the model based on the hydrological state of the catchment and its fine-scale topographic statistics. Details of the land model, including the partitioning into sub-tiles, can be found in Koster et al. (2000) and Ducharne et al. (2000). The Ocean tile can be divided into two sub-tiles that represent the ice-covered and ice-free parts of the ocean part of the atmospheric grid box. The fractional cover of these sub-tiles also varies with time. In MERRA, the fractional cover of sea ice and the ocean surface temperature are specified from the climatology of Reynolds and Smith (1994).

The area fraction of each surface type within each grid cell can be found in the MERRA 2d constants file collection (MAC0NXASM or “const_2d_asm_Nx”).



Appendix F: Budgets

As described in Appendix A, the concatenation of the IAU corrector segments of each analysis cycle is equivalent to a single, continuous run of the AGCM with an additional term–the analysis tendency–on the right-hand-side of all equations that predict analyzed quantities. Budgets in MERRA are thus identical to those in the free-running AGCM, with the addition of the effects of analysis tendencies, which are constant during the six-hour duration of each corrector segment.
In the MERRA output, we produce a full accounting for the budgets of eight atmospheric quantities: the mass of the atmosphere, the mass of water in vapor, liquid, and ice forms, kinetic energy, the virtual enthalpy, virtual potential temperature, and the total mass of odd oxygen. We also keep exact budgets for total water and energy at the land surface.

Atmospheric Budgets

An exact accounting for each of the eight atmospheric quantities is kept in two file collections: inst1_2d_int_Nx and tavg1_2d_int_Nx. Both collections consist of 2-dimensional data on the model's native horizontal grid, and both contain only vertical integrals. The first contains instantaneous mass-weighted vertical integrals of the eight quantities at the beginning of each hour (i.e., 00:00Z, 01:00Z, etc). The second collection contains various hourly-mean, vertically-integrated contributions to the tendencies of each quantity. This Appendix explains how these contributions can be combined to produce the local change in the quantity implied by values in the inst1_2d_int_Nx collection, how the contributions were computed, and what assumptions enter this accounting.

In the following, we will denote the finite-differenced, mass-weighted, vertical integrals of quantities by an overbar:

$$\overline{X} = \sum_{l=layers} X_l \delta m_l \,,$$

where $\delta m_l = m_d + m_v = \dfrac{\delta p_l}{g}$ is the mass of layer l, interpreted to be the sum of the mass of dry air and water vapor. Here δp_l is the pressure thickness of the layer and $\sum_{l=layers} \delta p_l = p_S - p_T$, where p_S is the surface pressure, and p_T is the pressure at the model's top (1 Pa).

Atmospheric Mass–The model satisfies the following conservation equation:

$$\frac{\partial M}{\partial t} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}}) + \left[\frac{\partial M}{\partial t}\right]_{ANA}$$

The spatial distribution of the total atmospheric mass, $M = (p_S - p_T)/g$, is stored in the variable **MASS** in the inst1_2d_int_Nx collection. In the tavg1_2d_int_Nx collection, the total tendency is separated into contributions from dynamics (**DMDT_DYN)**, and analysis (**DMDT_ANA)**. These satisfy collection-variable equations such as:

$\mathrm{MASS}_{1:00Z}$ – $\mathrm{MASS}_{0:00Z}$ = 3600. (DMDT_DYN + DMDT_ANA) $_{0:30Z}$

The finite-volume dynamics used in GEOS-5 conserves mass to within round-off at 64-bit precision, so the global mean of **DMDT_DYN** is effectively zero. The physics used in the GEOS-5 AGCM does not include changes in column mass due to changes in water content; there is, therefore, no physics contribution. Although the atmospheric physics does not include atmospheric water in its mass budget, realistic changes in atmospheric mass can still occur since the surface pressure is analyzed and surface pressure observations include the effect. Mass changes due to water content thus appear as analysis increments rather than physics increments. It is not clear how much improvement in the mass budget, as measured by, say, the degree of invariance of dry mass, would be achieved if the atmospheric model had a more complete treatment.

Atmospheric Water–The model predicts water vapor (specific humidity) as well as cloud ice and liquid water. Only specific humidity is analyzed, but analysis tendencies of the condensates can occur through the analysis increments of atmospheric mass.

For each of these three quantities we have:

$$\frac{\partial \overline{q_{(v,l,i)}}}{\partial t} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}\, q_{(v,l,i)}}) + \left[\frac{\partial \overline{q_{(v,l,i)}}}{\partial t}\right]_{PHY} + \left[\frac{\partial \overline{q_{(v,l,i)}}}{\partial t}\right]_{ANA}$$

The sum of the contributions obtained from the tavg1_2d_int_Nx collection exactly matches the difference of values in the inst1_2d_int_Nx. For example,

$\mathrm{TQV}_{1:00Z}$ – $\mathrm{TQV}_{0:00Z}$ = 3600 (DQVDT_DYN + DQVDT_PHY + DQVDT_ANA) $_{0:30Z}$
$\mathrm{TQL}_{1:00Z}$ – $\mathrm{TQL}_{0:00Z}$ = 3600 (DQLDT_DYN + DQLDT_PHY + DQLDT_ANA) $_{0:30Z}$
$\mathrm{TQI}_{1:00Z}$ – $\mathrm{TQI}_{0:00Z}$ = 3600 (DQIDT_DYN + DQIDT_PHY + DQIDT_ANA) $_{0:30Z}$

The total atmospheric water, $w = q_v + q_l + q_i$, satisfies:

$$\frac{\partial \overline{w}}{\partial t} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}\, w}) + (E - P) + \left[\frac{\partial \overline{w}}{\partial t}\right]_{FILL} + \left[\frac{\partial \overline{w}}{\partial t}\right]_{ANA}, \qquad (6)$$

where $P = P_l^{cu} + P_l^{ls} + P_s$ is the total precipitation rate at the surface and includes contributions from convective and large-scale rain and snow; E is the upward turbulent water vapor flux at the surface; and the "filling" of spurious negative water is part of the physics contribution. In the model, surface turbulent fluxes of the condensates are assumed to be zero. The contributions to total water from the dynamics, physics, and analysis are simply the sums of the corresponding terms in the three collection-variable equations. As implied by (6), the sum of the physics contributions to tendencies of the three phases of water also satisfies:

DQVDT_PHY + DQLDT_PHY + DQIDT_PHY = EVAP– PRECCU – PRECLS – PRECSN + QTFILL.

The model ignores the storage of falling precipitation, so the production and loss terms of rain and snow must balance.

$$C_r^{cu} + C_r^{ls} + G_r^{coll} + G_r^{auto} - E_r - F_r = P_l^{cu} + P_l^{ls}$$
$$G_s^{coll} + G_s^{sedm} - S_s + F_r = P_s \qquad (7)$$

Here C_r^{cu} is the direct production of rain from vapor by the convection parameterization and C_r^{ls} is the direct production of rain from vapor by the large-scale condensation , $G_{r,s}^{coll}$ are the generation of rain and snow from collection of cloud liquid droplets, G_r^{auto} is the increase of rainwater by auto conversion of cloud liquid water, G_s^{sedm} is the increase in frozen precipitation from the sedimentation of cloud ice, E_r is evaporation of rain and S_s is the sublimation of frozen precipitation (snow), F_r is the net conversion between rain and snow due to freezing and melting, and $P_{r,s}$ are the rain and snow reaching the surface. The corresponding collection-variable equations are:

CUCNVRN + LSCNVRN + COLCNVRN +AUTCNVRN– EVPRN – FRZRN
= PRECCU + PRECLS
COLCNVSN – SDMCI + SUBSN + FRZRN = PRECSN

The sum of the two equations also satisfies

$$G_P - E_P = P,$$

where $G_P = C_r^{cu} + C_r^{ls} + G_r^{coll} + G_r^{auto} - G_s^{coll} + G_s^{sedm}$ is the total production of precipitation, $E_P = E_r + S_s$ is the total evaporation and sublimation of rain and snow, and $P = P_l^{cu} + P_l^{ls} + P_s$ is the total precipitation reaching the surface. The corresponding collection-variable equation is:

PGENTOT – PREVTOT = PRECTOT ,

which are variables in the tavg1_2d_flx_Nx collection. This last balance is subject to greater round-off, since this collection has its precision degraded for compression.

The physics tendency of specific humidity is separated into contributions from moist processes, turbulence, chemistry (a small stratospheric source), and the numerical 'filling" of spurious negative values:

$$\left[\frac{\partial \overline{q_v}}{\partial t}\right]_{PHY} = \left[\frac{\partial \overline{q_v}}{\partial t}\right]_{MST} + \left[\frac{\partial \overline{q_v}}{\partial t}\right]_{TRB} + \left[\frac{\partial \overline{q_v}}{\partial t}\right]_{CHM} + \left[\frac{\partial \overline{q_v}}{\partial t}\right]_{FIL}.$$

The corresponding collection-variable equation is:

DQVDT_PHY = DQVDT_MST + DQVDT_TRB + DQVDT_CHM + DQVDT_FIL

Contributions from the moist and turbulence parameterizations can be written as:

$$\left[\frac{\partial \overline{q_v}}{\partial t}\right]_{MST} = -(C_r^{cu} + C_r^{ls} + C_l^{cu} + C_l^{ls} + C_i^{cu} + C_i^{ls}) + S_i + E_l + S_s + E_r,$$

$$\left[\frac{\partial \overline{q_v}}{\partial t}\right]_{TRB} = E\,,$$

In addition to the precipitation sources and sinks in (7), the contribution from moist processes includes the sources and sinks for cloud condensates. The terms C_r, C_l, and C_i are the direct conversions from water vapor to liquid precipitation, to cloud liquid, and to cloud ice, separated into cumulus and large-scale contributions; S_i and S_s is the vapor source from sublimation of cloud ice and snow; and E_l and E_r is the vapor source from evaporation of cloud liquid water and rain. The turbulence contribution is simply the evaporation from the surface, E. In terms of collection variables these are:

DQVDT_MST = – (CUCNVRN+LSCNVRN+CUCNVCL+LSCNVCL +CUCNVCI+LSCNVCI)
+ SUBCI + EVPCL + SUBSN + EVPRN

DQVDT_TRB = EVAP

The budgets for cloud liquid and cloud ice involve only terms from the moist processes and water filling. Although these are mixed by the turbulence, we ignore surface turbulent fluxes of cloud condensates.

$$\left[\frac{\partial \overline{q_{l,i}}}{\partial t}\right]_{PHY} = \left[\frac{\partial \overline{q_{l,i}}}{\partial t}\right]_{MST} + \left[\frac{\partial \overline{q_{l,i}}}{\partial t}\right]_{FIL},$$

And in equation variable form:

DQLDT_PHY = DQLDT_MST + DQLDT_FIL
DQIDT_PHY = DQIDT_MST + DQIDT_FIL.

The moist contributions can be written as:

$$\left[\frac{\partial \overline{q_l}}{\partial t}\right]_{MST} = C_l^{cu} + C_l^{ls} - (G_r^{coll} + G_r^{auto} + G_s^{coll}) - E_l - F_l\,,$$

$$\left[\frac{\partial \overline{q_i}}{\partial t}\right]_{MST} = C_i^{cu} + C_i^{ls} - G_s^{sedm} - S_i + F_l\,,$$

where F_l is the net freezing (freezing-melting) of cloud condensates. In terms of collection variables these are:

DQLDT_MST = CUCNVCL+ LSCNVCL – (COLCNVRN+AUTCNVRN+COLCNVSN)
– EVPCL –FRZCL
DQIDT_MST = CUCNVCI+ LSCNVCI – SDMCI – SUBCI + FRZCL

Atmospheric Kinetic Energy –The instantaneous total kinetic energy for each grid column,

$K = \overline{\frac{1}{2}|\underset{\rightarrow}{v}|^2}$, is stored in the variable **KE** in the inst1_2d_int_Nx collection. In the tavg1_2d_int_Nx collection, the total tendency is separated into contributions from dynamics (**DKDT_DYN)**, physics (**DKDT_PHY)**, and analysis increments of winds and atmospheric mass (**DKDT_ANA)**.

$$\frac{\partial K}{\partial t} = \left[\frac{\partial K}{\partial t}\right]_{DYN} + \left[\frac{\partial K}{\partial t}\right]_{PHY} + \left[\frac{\partial K}{\partial t}\right]_{ANA} \qquad (1)$$

The sum of the three contributions obtained from the tavg1_2d_int_Nx collection exactly matches the difference of values in the inst1_2d_int_Nx, since the contributions are computed by differencing K before and after each process is included. They thus satisfy collection-variable equations of the form:

$$\mathbf{KE_{1:00Z} - KE_{0:00Z} = 3600\ (DKDT\_DYN + DKDT\_PHY + DKDT\_ANA)_{0:30Z}}.$$

The dynamical contribution can be further divided by considering the vertically integrated kinetic energy equation written in flux form. Ignoring contributions from the analysis and from explicit frictional processes that are included in the physics, this is:

$$\left[\frac{\partial K}{\partial t}\right]_{DYN} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}\tfrac{1}{2}|\underset{\rightarrow}{v}|^2}) - \nabla \bullet (\overline{\underset{\rightarrow}{v}\phi}) - \overline{\omega\alpha} - \frac{\partial}{\partial t}(\phi_S p_S - \phi_T p_T) - Q_{NUM} \qquad (2)$$

The first term on the right hand side of (2) is the contribution from the inertial terms in the momentum equations, which reduces to the vertically integrated convergence of the kinetic energy flux. The next three terms are the total work done by the pressure gradient. The $\overline{\omega\alpha}$ term, which appears with opposite sign in the thermodynamic equation, is the conversion between kinetic and total potential energy. The fourth term is the change in gravitational potential energy associated with raising the mass in the column from sea-level to the local height of the surface topography and the change in gravitational potential energy of the mass above the model top (1 Pa or about 0.1 kg m^{-2}) as the height of the top rises and falls. Finally, we add a term, Q_{NUM}, to account for spurious numerical contributions, which we write as a dissipation.

We can form a collection-variable equation that corresponds to (2):

DKDT_DYN = CONVKE + CONVPHI + DKDT_GEN – DPDT_DYN +
DKDT_REMAP + DKDT_INRES + DKDT_PGRES. (2a)

The last three terms on the r.h.s. of (2a) represent a breakdown of the numerical residual Q_{NUM}. This breakdown deserves more explanation.

The numerics of finite volume dynamical core used in GEOS-5 do not naturally conserve total energy (Lin 200x). This is so for several reasons. First, the advection of the flow field produces numerical dissipation and so does not conserve kinetic energy. Some numerical dissipation is present in all discrete formulations of the dynamics useful for long-term integrations. Even conservative formulations, such as spectral schemes or centered finite-difference schemes like those based on the Arakawa Jacobians, need to be supplemented by explicit dissipation from horizontal diffusion or filters applied to the flow. In the finite volume dynamical core, the dissipation is not added explicitly and so is harder to track. Errors from these sources appear in the **DKDT_INRES** variable, which represents the errors made while the core is in the vertically Lagrangian mode, where all layers are bounded by material surfaces.

Spurious contributions from the inertial terms also occur when the core returns to hybrid sigma-pressure coordinates by doing a vertical coordinate transformation or "vertical remapping". Since this step does not involve any new physics, only a change in vertical coordinate, it should produce no energy transformations. The vertical remapping is indeed done to conserve the sum of internal, gravitational potential, and kinetic energies, but not the three individually. Non-conservation of kinetic energy by remapping is spurious and appears in the **DKDT_REMAP** variable.

Another spurious tendency of total energy comes from the form of the pressure gradient force and the discretization of the thermodynamic equation. In the continuous equations, the vertical integral of the rate of work done by the pressure gradient force is represented by second through fourth terms in (2), with the $\overline{\omega\alpha}$ term being identical to the corresponding term in the thermodynamic equation. This also is not true in the discrete core. For accounting purposes, non-conservation of total energy due to this mismatch of the conversion term is kept as a residual in the **DKDT_PGRES** variable.

Still another spurious contribution to kinetic energy comes from time-truncation errors. This arises because the state variables that are directly updated are the horizontal components of the wind, the surface pressure, and the virtual potential temperature, while the kinetic and total potential energy before and after the time step are non-linear combinations of these. For kinetic energy, these effects appear in both the **DKDT_INRES** and the **DKDT_REMAP** variables.

Finally, there is some error in our ability to estimate the convergence of kinetic energy and geopotential. These are not errors cause by the core's numerics, but by inaccuracies in our attempt to separate the various contributions. These errors implicitly appear in **DKDT_INRES** and the **DKDT_PGRES,** respectively.

To enforce total energy conservation, the core needs to compensate for all of these spurious sources. This is done through the addition of a "total energy fixer" heating, as discussed below with the thermodynamic equation breakdown.

In the global mean, we have:

$$\left[\frac{\partial}{\partial t}\left\langle K+\phi_S p_S-\phi_T p_T\right\rangle\right]_{DYN}=-\left\langle\overline{\omega\alpha}\right\rangle-\left\langle Q_{NUM}\right\rangle,$$

where the angled brackets denote the global mean.

The vertically integrated physics contribution to the kinetic energy tendency consists of frictional effects due to subgrid-scale motions. It can be separated into contributions due to turbulence, including all boundary layer friction, due to vertical mixing of momentum by the convective parameterization, and due to the friction associated with momentum transport by parameterized gravity waves. Since the GEOS-5 physics does not time split the effects of these three processes, there is some time truncation in the separation and this is stored as a residual.

$$\left[\frac{\partial K}{\partial t}\right]_{PHY}=-\overline{D_{GWD}}-\overline{D_{MST}}-\overline{D_{TRB}}-\overline{\Re_{PHY}},$$

In terms of collection variables, this is:

$$\textbf{DKDT\_PHY} = \textbf{DKDT\_GWD} + \textbf{DKDT\_MST} + \textbf{DKDT\_TRB} + \text{Residual},$$

where the small spurious contribution, $\overline{\Re_{PHY}}$, from time and space truncation errors, the latter due to the staggering of the winds, were not stored, but can be computed as a residual.

Finally, contributions from the gravity-wave-drag and turbulence parameterizations are further separated into different physical mechanisms:

DKDT_GWD = DKDT_ORO + DKDT_BKG + DKDT_RAY + DKDT_GWDRES

DKDT_TRB = DKDT_INT + DKDT_SFC + DKDT_TOP

Here **DKDT_GWD** is separated into contributions from orographic gravity waves, background gravity waves, Rayleigh friction in the upper atmosphere, and a residual due to the way the terms are implemented. **DKDT_TRB** is separated into contributions from internal diffusion, surface friction, and low-level topographic drag; no residual was stored.

Virtual Enthalpy– We keep a budget of the virtual enthalpy, $H = \overline{c_p T_v}$, where c_p is the heat capacity of dry air at constant pressure and T_v is the virtual temperature. Instantaneous values of this quantity are stored for each grid column in the variable **CPT** in the inst1_2d_int_Nx collection. In the tavg1_2d_int_Nx collection, its total tendency is first separated into contributions from dynamics (**DHDT_DYN**), physics (**DHDT_PHY**), and analysis (**DHDT_ANA):**

$$\frac{\partial H}{\partial t} = \left[\frac{\partial H}{\partial t}\right]_{DYN} + \left[\frac{\partial H}{\partial t}\right]_{PHY} + \left[\frac{\partial H}{\partial t}\right]_{ANA}.$$

As with K, the sum of the three contributions obtained from the tavg1_2d_int_Nx collection exactly matches the difference of values in the inst1_2d_int_Nx. For example,

$$\mathbf{CPT_{1:00Z} - CPT_{0:00Z} = 3600.(DHDT\_DYN + DHDT\_PHY + DHDT\_ANA)_{0:30Z}}.$$

The dynamical contribution can be rewritten, separating the conversion term. This leaves:

$$\frac{\partial H}{\partial t} = -\nabla \bullet (\overline{\vec{v} c_p T_v}) + \overline{\omega \alpha} + Q_{NUM} + \left[\frac{\partial H}{\partial t}\right]_{PHY} + \left[\frac{\partial H}{\partial t}\right]_{ANA}, \tag{3}$$

where Q_{NUM} is a heating added by the dynamics to conserve total energy in the presence of numerical dissipation–the "energy fixer." The corresponding collection-variable equationfor the dynamics contribution is:

DHDT_DYN = CONVCPT – DKDT_GEN + TEFIXER, (3a)

where

TEFIXER = - DKDT_REMAP -<DKDT_INRES + DKDT_PGRES>, (3b)

as required for total energy conservation (c.f., Eq. (2a-c)). As before, the angle brackets around the last two terms denote a global mean. The horizontal distributions of these terms are simply diagnostic estimates, and only their global means are used by the dynamical core.

The vertically integrated physics contribution to the H tendency can be separated into contributions due to turbulence, including sensible heat fluxes as well as turbulent friction; contributions due moist processes, including condensation and cumulus friction; and contributions due to the friction associated with momentum transport by parameterized gravity waves. Since the GEOS-5 physics does not time split the effects of these three processes, there is some time truncation in the separation and this is stored as a residual, $\overline{\Re_{PHY}}$.

$$\left[\frac{\partial H}{\partial t}\right]_{PHY} = \overline{Q_{TRB}} + \overline{Q_{MST}} + \overline{Q_{GWD}} + \overline{Q_{RAD}} + \overline{\Re_{PHY}}, \tag{4}$$

and in terms of collection variables:

DHDT_PHY = DHDT_GWD + DHDT_MST + DHDT_TRB + DHDT_RAD + RPHYRES

The first four quantities on the r.h.s. can be further separated into contributions from more detailed physical processes:

$$\overline{Q_{TRB}} = SH + \overline{D_{TRB}}$$

$$\overline{Q_{GWD}} = \overline{D_{GWD}}$$

$$\overline{Q_{MST}} = -L_v \left[\frac{\partial \overline{q_v}}{\partial t}\right]_{MST} + L_f \left[\frac{\partial \overline{q_i}}{\partial t}\right]_{MST} + L_f P_s + \overline{D_{MST}}$$

$$= -L_v (E_l + E_r - S_s - C_l^{cu} - C_l^{ls} - C_r) - L_s (S_i - C_i^{cu} - C_i^{ls}) + L_f (F_r + F_l + G_s^{coll} - S_s) + \overline{D_{MST}}$$

$$\overline{Q_{RAD}} = (SW_T - SW_S) - (OLR + LW_S).$$

Here L_v, L_s, L_f are the latent heats of evaporation, sublimation, and fusion of water, SW_T and SW_S are the net downward fluxes of solar radiation at the top of the atmosphere and at the surface, LW_S is the net downward flux of longwave radiation at the surface, OLR is the outgoing longwave radiation at the top of the atmosphere, and *SH* is the sensible heat flux at the surface. The *D* terms are dissipations, as they appeared in the kinetic energy equation, and the remaining terms are water conversions to precipitation or exchanges between the three phases of suspended atmospheric water, as they appeared in the water budgets.

In terms of collection variables these are:

DHDT_TRB = HFLUX - DKDT_TRB (4a)

DHDT_GWD = - DKDT_GWD (4b)

DHDT_MST = $-L_v$ **DQVDT_MST** $-$ L_f **(DQIDT_MST + PRESN) – DKDT_MST** (4c)

DHDT_RAD = (LWTNET-LWGNET) + (SWNETTOA-SWNETSRF). (4d)

Total Energy–

The total energy of the atmospheric column is

$$T = K + H + \phi_S p_S - \phi_T p_T + L_v q_v - L_f q_i$$

Combining equations (1-4):

$$\frac{\partial(K+H+\phi_S p_S-\phi_T p_T)}{\partial t}=-\nabla\bullet(\overline{\underset{\rightarrow}{v}\tfrac{1}{2}|\underset{\rightarrow}{v}|^2}+\overline{\underset{\rightarrow}{v}c_p T_v}+\overline{\underset{\rightarrow}{v}\phi})+\left[\frac{\partial K}{\partial t}+\frac{\partial H}{\partial t}\right]_{PHY,ANA}+\left[\frac{\partial(\phi_S p_S-\phi_T p_T)}{\partial t}\right]_{ANA} \tag{5}$$

where $\frac{\partial(\phi_S p_S-\phi_T p_T)}{\partial t}$ on the l.h.s. includes the dynamical contribution from (2) as well as the contribution from the physics, which is not separated in the diagnostics. This is a small effect, since in MERRA the physics does not alter the surface pressure. The sum of the physics contributions to (5) can be written:

$$\left[\frac{\partial K}{\partial t}+\frac{\partial H}{\partial t}\right]_{PHY}=(SW_T-SW_S)-(OLR+LW_S)-L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{PHY}-L_f\left[\frac{\partial\overline{q_i}}{\partial t}\right]_{PHY}+SH+\overline{\Re_{PHY}} \tag{6}$$

From the atmospheric water vapor and ice budgets we have

$$\frac{\partial(L_v\overline{q_v}-L_f\overline{q_i})}{\partial t}=-\nabla\bullet(L_v\overline{\underset{\rightarrow}{v}q_v}-L_f\overline{\underset{\rightarrow}{v}q_i})+L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{PHY,ANA}-L_f\left[\frac{\partial\overline{q_i}}{\partial t}\right]_{PHY,ANA},$$

Where the physics contributions can be written

$$\left[L_v\frac{\partial\overline{q_v}}{\partial t}-L_f\frac{\partial\overline{q_i}}{\partial t}\right]_{PHY}=-L_v(C_r+C_l-E_l-S_s-E_r-E)-L_s(C_i-S_i)-L_f(F_l-G_s^{sedm})+L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{CHM}+L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{FIL}-L_f\left[\frac{\partial\overline{q_i}}{\partial t}\right]_{FIL} \tag{7}$$

Now adding (6) and (7),

$$\left[\frac{\partial T}{\partial t}\right]_{PHY}=(SW_T-SW_S)-(OLR+LW_S)+SH+L_vE+L_f(F_r+S_s+G_s^{sedm}+G_s^{coll})+L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{CHM}+L_v\left[\frac{\partial\overline{q_v}}{\partial t}\right]_{FIL}-L_f\left[\frac{\partial\overline{q_i}}{\partial t}\right]_{FIL}$$

Potential Temperature– The model's dynamics uses virtual potential temperature, θ_v, as the prognostic variable in its thermodynamic equation, we therefore keep its budget. First, we separate its local vertically integrated tendency into contributions from the dynamics, physics, and analysis:

$$\frac{\partial\overline{\theta}_v}{\partial t}=\left[\frac{\partial\overline{\theta_v}}{\partial t}\right]_{DYN}+\left[\frac{\partial\overline{\theta_v}}{\partial t}\right]_{PHY}+\left[\frac{\partial\overline{\theta_v}}{\partial t}\right]_{ANA}$$

The sum of the three contributions obtained from the tavg1_2d_int_Nx collection exactly matches the difference of values in the inst1_2d_int_Nx. For example, the collection-variable equation

$$\mathbf{THV_{1:00Z}-THV_{0:00Z}=3600.(DTHDT\_DYN+DTHDT\_PHY+DTHDT\_ANA)_{0:30Z}}$$

is satisfied within round-off.

Since virtual potential temperature is the prognostic variable used by the dynamical core, it is within round-off while in vertically Lagrangian mode. Both horizontal and vertical advection (i.e., the vertical remapping from the Lagrangian vertical coordinate to hybrid sigma-p) result in diabatic contributions; these are assumed to conserve total energy, but of course, will not conserve θ_v. The total dynamics contribution includes these diabatic tendencies, in addition to the vertically integrated convergence of virtual potential temperature.

$$\left[\frac{\partial \bar{\theta}_v}{\partial t}\right]_{DYN} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}\,\theta_v}) + \left[\frac{\partial \overline{\theta_v}}{\partial t}\right]_V + \left[\frac{\partial \overline{\theta_v}}{\partial t}\right]_H$$

or

$$\mathbf{DTHDT\_DYN = CONVTHV + DTHDT\_REMAP + DTHDT\_CONSV}$$

No further decomposition is done for this budget

Atmospheric Ozone– The model uses odd oxygen mixing ratio, q_{O_x}, as its prognostic variable. Its vertically integrated tendency is given by:

$$\frac{\partial \overline{q_{O_x}}}{\partial t} = -\nabla \bullet (\overline{\underset{\rightarrow}{v}\,q_{O_x}}) + \left[\frac{\partial \overline{q_{O_x}}}{\partial t}\right]_{PHY} + \left[\frac{\partial \overline{q_{O_x}}}{\partial t}\right]_{ANA}$$

or

DOXDT = DOXDT_DYN + DOXDT_PHY + DOXDT_ANA.

The dynamics contribution is simply the vertically integrated convergence. The physics contributions result from parameterized production and loss terms, mostly in the stratosphere. In MERRA, ozone is analyzed and so q_{O_x} can change due to ozone increments or due to increments in atmospheric dry mass. We report only the total analysis contribution.

Land Budgets

Complete budgets for total land water storage and total land energy storage are accessible in the tavg1_2d_int_Nx collection at the full resolution. More detailed partitioning of budget terms can be achieved from quantities in the tavg1_2d_flx_Nx and tavg1_2d_lnd_Nx (or tavg1_2d_mld_Nx for MERRA-Land) collections, but these are not saved at the full precision.

Total Land Water Budget– The land surface water balance equation can be written

$$\frac{\partial W}{\partial t} = P_l + P_s - E_L - R_L + \Re_W$$

where W is the total water held in all land surface reservoirs (comprising the soil, the interception reservoir, and the snowpack), P_l and P_s are the liquid rain and "snowfall" rates, respectively, E_L is the total evapotranspiration rate, R_L is the total runoff–surface (or overland) plus baseflow, and $\Re_W$ is a spurious water source (nonzero for MERRA due to logistical issues with the land-atmosphere interface but identical to zero for MERRA-Land). The corresponding collection-variable equation, using quantities stored in tavg1_2d_int_Nx and tavg1_2d_lnd_Nx (or tavg1_2d_mld_Nx for MERRA-Land), is

WCHANGE = PRECTOT – EVLAND – RUNOFF – BASEFLOW + SPWATR.

See Appendix E for variable definitions. Note that all of these quantities are values per unit land area, in kg m^{-2} s^{-1} – i.e., not weighted by fractional land area of the grid cell and excluding any contributions from lake, land ice, or ocean tiles. Note also that **WCHANGE, PRECTOT, EVLAND, RUNOFF, BASEFLOW** and **SPWATR** are defined only at grid cells for which the land fraction is non-zero and are set to the undefined value elsewhere. The two precipitation rates, however, are the same as found in other collections, since in MERRA rain and snow fall uniformly across all subgrid surface types and are defined everywhere. (In contrast to EVLAND, the MERRA field EVAP includes evaporation from all surface types, including land, lake, land-ice, and ocean.) Note that the precipitation rate in MERRA-Land differs from that found in MERRA collections due to the correction of the precipitation forcing with a gauge-based precipitation data product (Reichle, 2012).

In the tavg1_2d_lnd_Nx and tavg1_2d_mld_Nx collections, the evapotranspiration is separated into components according to its sources as follows:

$$E_L = E_{tr} + E_{bs} + E_{int} + E_{snow} ,$$

where E_{tr} is the transpiration rate, E_{bs} is the evaporation from bare soil surfaces, E_{int} is the evaporation from the land's interception reservoir (e.g., water droplets sitting on leaves after a rainstorm), and E_{snow} is the water that sublimates from the snowpack. The corresponding collection-variable equation is

$$\mathbf{EVLAND} = \frac{1}{L_v}(\mathbf{EVPTRNS + EVPSOIL + EVPINTR}) + \frac{1}{L_s}\mathbf{EVPSBLN}$$

where L_v and L_s are the latent heat of vaporization and the latent heat of sublimation, respectively. Following GEOS-5.7.2, MERRA-Land uses L_v=2.4665e6 J kg^{-1} (at 15°C) and L_s=2.8002e6 J kg^{-1}, whereas MERRA (GEOS-5.2.0) used L_v=2.4548e6 J kg^{-1} and L_s=2.8368e6 J kg^{-1}.

Total Land Energy Budget –The balance equation for total land surface energy can be written:

$$\frac{\partial \varepsilon}{\partial t} = SW_L + LW_L - SH_L - L_v E_L - L_f \Delta SWE + \Re_L ,$$

where ε is the total heat content (in the soil, canopy, and snowpack) relative to liquid water. SW_L is the net shortwave radiation, LW_L is the net longwave radiation, L_v is the latent heat of vaporization (from liquid), E_L is the total evaporation from the land surface, SH_L is the sensible heat flux from the land surface, L_f is the latent heat of fusion, and ΔSWE is the change in the snow water equivalent (through addition of frozen precipitation falling on the surface or removal of snow through melt or sublimation). The term L_f ΔSWE therefore accounts for the snowmelt energy and the added energy needed to evaporate from solid rather than liquid water. The spurious snow energy source $\Re_L$ is explained below.

In terms of "lnd" and "mld" collection variables the energy budget is:

$$\mathbf{ECHANGE = SWLAND + LWLAND - SHLAND} - L_v\,\mathbf{EVLAND}$$
$$- L_f\mathbf{PRECSNO - SPLAND - SPSNOW.}$$

All of these quantities are computed over land only and are *not* weighted by fractional land area. SPLAND, in analogy to SPWATR, is associated with the coupled land-atmosphere interface in MERRA and is always zero for MERRA-Land. The term SPSNOW contains "spurious" snow-related energy sources and sinks associated with several small accounting inconsistencies across the coupled models. When, for example, the same amount of snow falls at -20°C in one region and at 0°C in a second region, more energy is needed in the first region to melt the snow, because in the first region, energy is needed first to warm the snow up to 0°C. The atmospheric model does not distinguish between the energy content of snow falling at -20°C and that falling at 0°C, whereas the land model does account for this energy difference. To rectify this inconsistency between the land and atmosphere models, the "negative energy" of the colder snow (i.e., the energy deficit relative to 0°C snow) is "invented" and added to the snow's internal energy as soon as it hits the surface. It is therefore implicitly included in **ECHANGE**.

References

Bloom, S., L. Takacs, A. DaSilva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256-1271.

Collins, N., G. Theurich, C. DeLuca, M. Suarez, A. Trayanov, V. Balaji, P. Li, W. Yang, C. Hill, and A. da Silva, 2005: Design and implementation of components in the Earth System Modeling Framework. *Int. J. High Perf. Comput. Appl*., **19**, 341-350, DOI: 10.1177/1094342005056120.

Derber, J. C., R. J. Purser, W.-S. Wu, R. Treadon, M. Pondeca, D. Parrish, and D. Kleist, 2003: Flow-dependent Jb in a global grid-point 3D-Var. *Proc. ECMWF annual seminar on recent developments in data assimilation for atmosphere and ocean.* Reading, UK, 8-12 Sept. 2003.

deWitt, Shaun, 1996: Writing HDF-EOS Grid Products for Optimum Subsetting Services. ECS Technical Paper 170-TP-007-001.

Dopplick, T., 1997: The Role of Metadata in EOSDIS. EOSDIS Technical Paper 160-TP-013-001.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model 2. Parameter estimation and model demonstration. *J. Geophys. Res.*, **105**, 24823–24838, doi:10.1029/2000JD900328.

Gross, C., 1997: B.0 Earth Science Data Model for the ECS Project. EOSDIS Technical Paper 420-TP-015-001.

Koster, R. D., M. J. Suárez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a GCM, Part 1, Model Structure. *J. Geophys. Res.*, **105**, 24809-24822.

NOAA, 1995: Conventions for the Standardization of NetCDF Files (COARDS). http://ferret.wrc.noaa.gov/noaa_coop/coop_cdf_profile.html

Pfafstetter, O., 1989. Classification of hydrographic basins: coding methodology, unpublished manuscript, Departamento Nacional de Obras de Saneamento, August 18, 1989, Rio de Janeiro; available from J.P. Verdin, U.S. Geological Survey, EROS Data Center, Sioux Falls, South Dakota 57198 USA. See, for example:

> Verdin, K.L. and J.P. Verdin, 1999: A topological system for delineation and codification of the Earth's river basins. *J. Hydrology*, **218**, nos. 1-2, pp. 1-12
> or
> http://gis.esri.com/library/userconf/proc01/professional/papers/pap1008/p1008.htm.

Reichle, R. H., 2012: The MERRA-Land Data Product, version 1.1. GMAO Technical Report, NASA Global Modeling and Assimilation Office, Goddard Space Flight Center, Greenbelt, MD, USA. Available at http://gmao.gsfc.nasa.gov/research/merra/file_specifications.php

Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate,* **24**, 6322-6338, doi:10.1175/JCLI-D-10-05033.1.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analysis using optimal interpolation. *J. Climate,* **7,** 929–948.

Rienecker and Coauthors, 2011: MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24**, 3624-3648, doi:10.1175/JCLI-D-11-00015.1.

Rienecker, M.M., M.J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R.D. Koster, R. Gelaro, I. Stajner, and E. Nielsen, 2008: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data Assimilation 104606*, **Vol. 27**.

Wu, W.-S., R.J. Purser and D.F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.*, **130**, 2905-2916.

Web Resources

GMAO web site: http://gmao.gsfc.nasa.gov/

GMAO Operations page: http://gmao.gsfc.nasa.gov/products/

CF Standard Description: http://cf-pcmdi.llnl.gov/

GEOS-5 Variable Definition Glossary:
http://gmao.gsfc.nasa.gov/products/documents/GEOS5_Glossary_11_01_2011.pdf

The HDF Group: HDF-4 information. http://www.hdfgroup.org /

NASA, 1999: HDF-EOS Standards. http://earthdata.nasa.gov/data/references/