CERES Data Management Status

Presented to
CERES Phase-2 Science Team

Erika Geier

Mike Little

NASA Langley Research Center

November 2, 2004

michael.m.little@nasa.gov

erika.b.geier@nasa.gov

Agenda

- Systems Evolution
- Data Products
  - Documentation Overview
  - CERES File Naming Conventions
  - Questions about Data Set
  - CERES Data Product Summary Tables
    - New Data Sets since March 2004
    - Data Sets expected before next Science Team Meeting
    - Older CERES Data Sets

SCF and ASDC System Evolution

- **Plan ahead for future computation needs: measurements not missions**
  - Limitations of future funding
    - Reduce cost
    - Permit flexibility to pay-as-you-go
  - NPOESS-like computing environment: IBM 970
  - Flexibility to run individual codes for special studies
  - As instrument life continues, collecting large data volumes
- **Transform Computational Environment by 12/06**
  - Replace SGI environment with commodity platforms
    - ASDC: IBM-970 cluster
    - SCF: Mac G5 workstations and cluster
  - Increase Storage using commodity solutions accessible by all platforms
    - 25TB at SCF using Apple Xserve RAID (RAID-5)
    - 20TB at SCF using LaCie firewire disks (non-RAID)
  - Replace production management tools
    - Apply LINUX tools at ASDC to manage mass-production
    - S4P production management tool from GSFC at ASDC and SCF
    - Apply OS X & grid-computing tools at SCF for special studies
  - Migrate codes to open-source-like development environment

CERES Code Development Environment

- **Future computation and user environment will change**
  - Need to run codes as part of special studies on non-ASDC platforms
  - Need to prove scientific equivalency of data products regardless of platform
    - Configuration management
    - Transfer of validation from one version to another
  - More directive guidance on software engineering by NASA (NPR 7150.2)
  - Grid computing yields opportunities for decentralized production model
    - Need to re-write production control files
  - Preparation for NPOESS computational environment
- **Convert Code to Open-Source Model**
  - Current effort to build/test conversion process to be complete by 02/05
    - Code conversion to open source environment vs. re-hosting to new platform
    - Data product validation as scientifically equivalent
    - Production tools to run special studies at SCF
  - Code conversion of production to run at multiple facilities
    - IBM-970 LINUX at ASDC
    - G-5 Cluster at SCF
    - IBM-970 AIX at NPOESS computation facilities

Documentation Overview

- Data Quality Summaries
  - Detailed information about a particular data set
  - Always consult Quality Summary prior to using data or publishing research

- Data Products Catalog
  - Parameter lists for each data product
  - Version of pages that apply to data set included with order

- Collection Guides
  - User Guide for data product

- Description/Abstract
  - Record of differences between data sets and configuration codes

Questions about Data Set

- Look over Data Products Catalog pages
- Reread Data Quality Summary
- Consult Collection Guide, if available
- Specific science questions may be sent to Contact Scientist listed in Section 2.2 of Collection Guide or in Description/Abstract
- All other questions should be sent to User Services

    larc@eos.nasa.gov

- For data products for which no Collection Guide or Description/Abstract is available, send all questions to User Services

CERES Science Data Sets
New Since March 2004

Data Set	Comments
CER_SSF_Terra-FMx-MODIS_Edition2B	• Terra Edition2 Clouds • Final Terra ADMs • Processing 3/00 – 12/03
CER_SFC_Terra-FMx-MODIS_Edition2B	• Edition2B SSF are input
CER_CRS_Terra-FMx-MODIS_Edition2B	• Edition2B SSF are input • Latest version of SARB algorithms • Processing crosstrack data for 3/00 – 6/03 • Not yet publicly available
CER_FSW_Terra-FMx-MODIS_Edition2B	• Edition2B CRS are input • Only crosstrack data processed • Not yet publicly available
CER_SSF_Aqua-FMx-MODIS_Edition1A	• Processing 2 years of data (7/02 – 6/04) • Not yet publicly available
CER_SFC_Aqua-FMx-MODIS_Edition1A	• Edition1A SSF are input • Processing up to 2 years of data (7/02 – 6/04) • Not yet publicly available
CER_SYN_Terra-FMx-MODIS_Beta2 CER_AVG_Terra-FMx-MODIS_Beta2 CER_ZAVG_Terra-FMx-MODIS_Beta2	• First Synoptic SARB and TISA averaged data sets for Terra • Inputs from Edition2A data stream • Processing 5 months of data (3/00, 1/01, 4/01, 7/01, 10/01) • Not yet publicly available

CERES Science Data Sets
Expected Before Next Science Team Meeting

Data Set	Comments	When?
CER_SFC_Terra-FMx-MODIS_Edition2C	• Correct region number problem • 3 years of data to be processed (3/00 – 2/03)	Expected to start processing in December'04
CER_SRBAVG_Terra-FMx-MODIS_Edition2C	• Edition2C SFC are input • 3 years of data to be processed (3/00 – 2/03)	Expected to start processing in December'04
CER_SFC_Aqua-FMx-MODIS_Edition1B	• Correct region number problem • 1 year of data: 7/02 – 6/03	Expected to start processing in early 2005
CER_SRBAVG_Aqua-FMx-MODIS_Edition1B	• Edition1B SFC are input • Up to 1 year of data	Expected to start processing in early 2005
CER_CRS_Aqua-FMx-MODIS_Beta1	• First Aqua SARB data set • 1 year of crosstrack data to be processed • Aqua Edition1A SSF are input	Expected to start processing in Spring 2005
CER_FSW_Aqua-FMx-MODIS_Beta1	• First Aqua FSW data set • 1 year of crosstrack data to be processed • Aqua Beta1 CRS are input	Expected to start processing in Spring 2005

Older CERES Science Data Sets
Still of Interest (Part 1)

Data Set	Comments
CER_BDS_TRMM-PFM_Edition1	• TRMM period Jan'98 – Aug'98; Mar'00
CER_BDS_TRMM-PFM_Transient-Ops2	• Transient Operations period Jan'99 – Jul'99 • Not publicly released
CER_ES8_TRMM-PFM_Edition2	• TRMM period Jan'98 – Aug'98; Mar'00
CER_ES8_TRMM-PFM_Transient-Ops2	• Transient Operations period Jan'99 – Jul'99
CER_ES4_TRMM-PFM_Edition2 CER_ES9_TRMM-PFM_Edition2	• TRMM period Jan'98 – Aug'98; Mar'00
CER_ES4_PFM+FM1+FM2_Edition2 CER_ES9_PFM+FM1+FM2_Edition2 CER_ES4_PFM+FM1_Edition2 CER_ES9_PFM+FM1_Edition2 CER_ES4_PFM+FM2_Edition2 CER_ES9_PFM+FM2_Edition2	• Mar'00 only month of TRMM, Terra overlap
CER_SSF_TRMM-PFM-VIRS_Edition2B	• TRMM period Jan'98 – Aug'98; Mar'00
CER_SSF_TRMM-PFM-VIRS_Edition2B-TransOps	• Transient Operations period Jan'99 – Jul'99
CER_SSF_TRMM-SIM-VIRS_Edition2-VIRSonly	• Processed only for months when CERES was off • Currently available through Jul'01
CER_SFC_TRMM-PFM-VIRS_Edition2B	• TRMM period Jan'98 – Aug'98; Mar'00
CER_SRBAVG_TRMM-PFM-VIRS_Edition2B	• TRMM period Jan'98 – Aug'98; Mar'00

Older CERES Science Data Sets
Still of Interest (Part 2)

Data Set	Comments
CER_CRS_ TRMM-PFM-VIRS_Edition2C	• TRMM period Jan'98 – Aug'98; Mar'00
CER_FSW_ TRMM-PFM-VIRS_Edition2C	• TRMM period Jan'98 – Aug'98; Mar'00
CER_BDS_Terra-FMx_Edition2 CER_ES8_Terra-FMx_Edition2 CER_ES4_Terra-FMx_Edition2 CER_ES9_Terra-FMx_Edition2	• Terra Edition2 BDS and ERBElike available Mar'00 through Dec'03
CER_ES4_FM1+FM2_Edition2 CER_ES9_FM1+FM2_Edition2	• Available Mar'00 through Dec'02 • Production of these data sets halted • Replaced by Terra/Aqua combined crosstrack data sets (FM1+FM3, FM1+FM4 Edition2)
CER_BDS_Aqua-FMx_Edition2 CER_ES8_Aqua-FMx_Edition2 CER_ES4_Aqua-FMx_Edition2 CER_ES9_Aqua-FMx_Edition2	• Instrument drifts removed • 6/18/02 – 2/29/04 currently available
CER_ES4_FM1+FM3_Edition2 CER_ES4_FM1+FM4_Edition2 CER_ES9_FM1+FM3_Edition2 CER_ES9_FM1+FM4_Edition2	• Combine only crosstrack instruments • Jul'02 – Dec'03 currently available • uses Terra Edition2 and Aqua Edition2 inputs

Science Data Product URLs and Contacts

- Ordering Data
  - http://eosweb.larc.nasa.gov/HBDOCS/langley_web_tool.html
  - Can also order data from EOS Data Gateway
- Subsets of SSF, CRS, and ES8 are available
  - Order data using Java version of Langley Ordering Tool
  - Can subset by parameters or latitude/longitude box
- Contact Points
  - All questions regarding production data products and their use
    - E-mail: larc@eos.nasa.gov
    - Langley ASDC Customer Service
  - Contents of this presentation: michael.m.little@nasa.gov
- CERES News (e-mail)
  - Subscribe from CERES datasets webpage
  - All new public datasets are announced as soon as available
  - Mechanism for distributing important CERES information