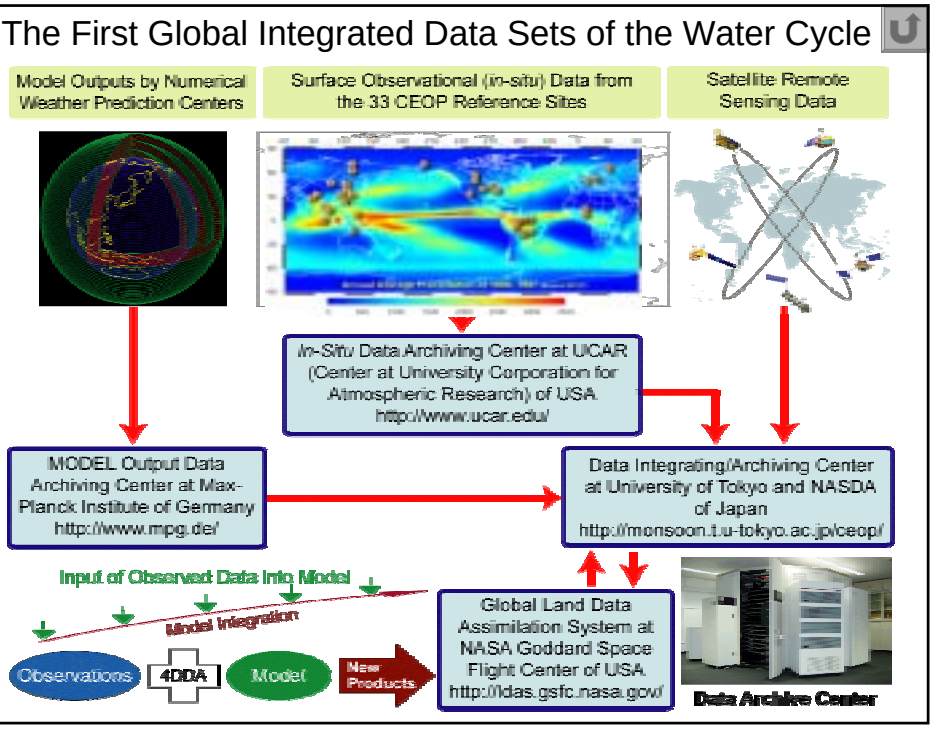
DATA MANAGEMENT UPDATE

Steven Williams, Scot Loehrer, and Linda Cully
UCAR/Joint Office for Science Support (JOSS)
Boulder, Colorado, USA

Hans-Joerg Isemer
GKSS-Forschungszentrum Geesthacht GmbH
Head of the International BALTEX Secretariat
Geesthacht, Germany

3rd International CEOP Implementation & Planning Meeting
University of California-Irvine, CA, USA
10-12 March 2004

The First Global Integrated Data Sets of the Water Cycle



CEOP DATA MANAGEMENT WWW PAGE

http://www.joss.ucar.edu/ghp/ceopdm/

In-Situ Reference Site Data Sets and Information

Data Sets

- CEOP EOP-2 Reference Site Data Sets
- CEOP EOP-1 Reference Site Data Sets
- NASA/GMAO CEOP EOP-1 Reference Site Data Sets in GrADS Format
- Sample Reference Site Data Sets
- CEOP In-Situ Data Source Agency Sites
- Baseline Surface Radiation Network (BSRN)
- GEWEX Land Processes Distributed Data Server

Information

- CEOP Reference Site Data Set Procedures Report (Approved by the SSC 29 July 2003)
- CEOP Reference Site Station Characteristics
- CEOP Reference Site Atlas
- Reference site data management / update (GEWEX SSG Meeting, 20-24 January 2003)
- CEOP Reference Site Investigator Station Questionnaire
  - Responses

Satellite Data and Information

Data Sets

- EOP-1 Satellite Data Sets
- NASA/GMAO GRADS/IDEB CEOP Surface T and Cloud Amount for CEOP EOP1

Information

- CEOP Satellite Data Source Agency Sites

Model Output and Information

Data Policies

- Final DRAFT CEOP Reference Site Data Release Guidelines
- BALTEX
- CAMP
- CATCH
- GAPP
- LBA
- MAGS

Data Standards Information

- Assistance for land-surface modelling activities (ALMA)
- Atmospheric Model Intercomparison Project (AMIP)
- GCIP/LDAS

Documents

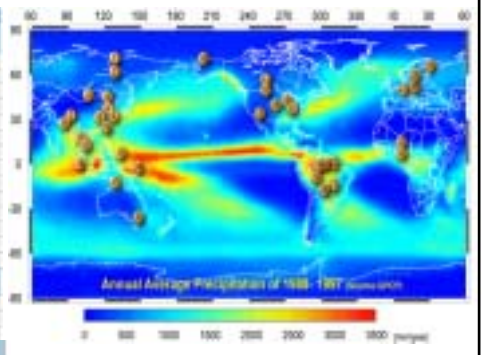
- CEOP Implementation Plan
- Report from the 2nd CEOP Implementation Planning Meeting (DRAFT - 8 July 2003)
- WEST Major Activities Plan (1 June 2003)
- CEOP Reference Site Station Characteristics Questionnaire
- Establishment of a Global Hydrological Observation Network for Climate: GCOS/GTOS/WMP Meeting Report (June 2003)

Other Links

- CEOP Home Page
- WCRP Home Page
- GEWEX Home Page
- CLIVAR Home Page
- CSC Home Page
- ACCP Home Page
- Global Modeling and Assimilation Office (NASA/GSFC)
- Land Information System (NASA/GSFC)
- Model Parameter Estimation Experiment (MOPEX)
- NASA/Goddard Institute for Space Studies (GISS) Data

REFERENCE SITE LOCATIONS

Number	Reference Site Name	Satellite Scene Center Latitude	Satellite Scene Center Longitude
	CAMP		
	Other		
	MAGS		
	GAPP		
	LBA		



Number	Reference Site Name	Satellite Scene Center Latitude	Satellite Scene Center Longitude
	BALTEX		
	CATCH		
	MDB		

CEOP Reference Sites

Data Release Guidelines

Reinforcements in the 11 December 2003 version:

•New section 2.6: Cooperation between Site PIs and *Data Users*

•Executive Summary ("Golden Rules") added

•Action for CDA: Automatic delivery of policy guidelines to *Data Users*

Hans-Jörg Isemer, International BALTEX Secretariat, *isemer@gkss.de*

Executive Summary (1):

1. No financial implications are involved for the CEOP reference site data exchange. *Section 2.1.*

2. Commercial use and exploitation of CEOP reference site data is prohibited. *Section 2.2.*

3. Any re-export or transfer of the original data received from the CDA archive to a third party is prohibited. *Section 2.3.*

4. The origin of CEOP reference site data being used for publication of scientific results must be acknowledged and referenced in the publication. *Section 2.5.*

Hans-Jörg Isemer, International BALTEX Secretariat, *isemer@gkss.de*

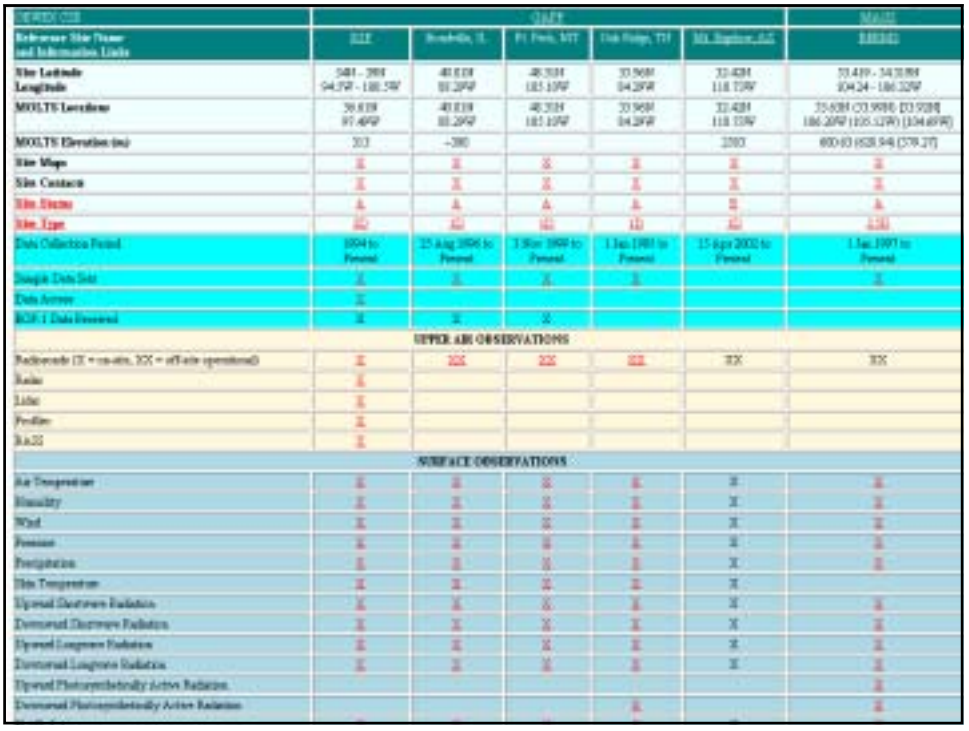
Executive Summary (2):

5. CEOP reference site *data users* are strongly encouraged to establish direct contact with *data providers* for complete interpretation and analysis of data for publication purposes. *Section 2.6.*

6. Co-authorship of *data users* and CEOP reference site Principle Investigators on papers making extensive use of CEOP data is justifiable and highly recommended.

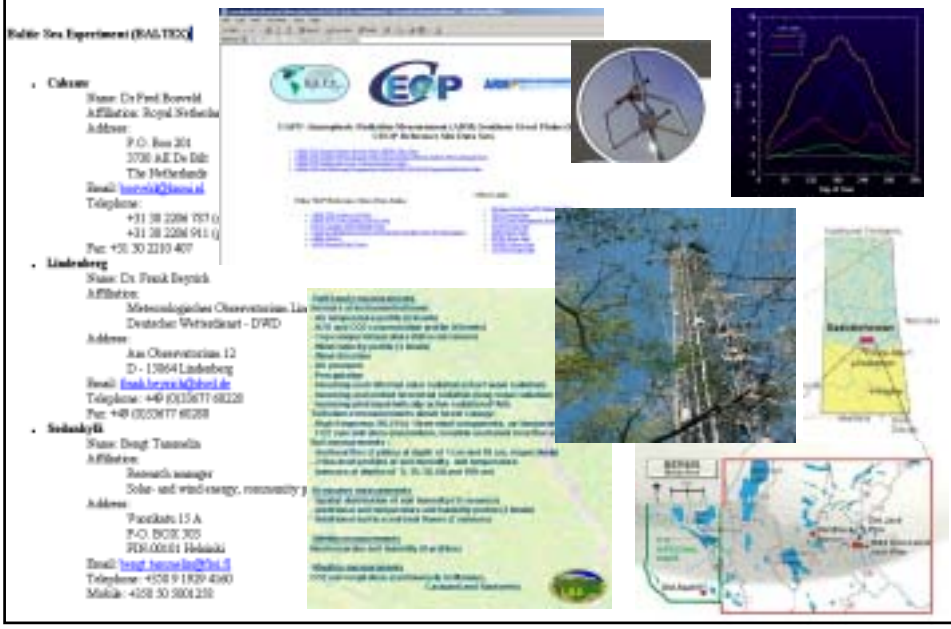
Section 2.7.

Hans-Jörg Isemer, International BALTEX Secretariat, *isemer@gkss.de*

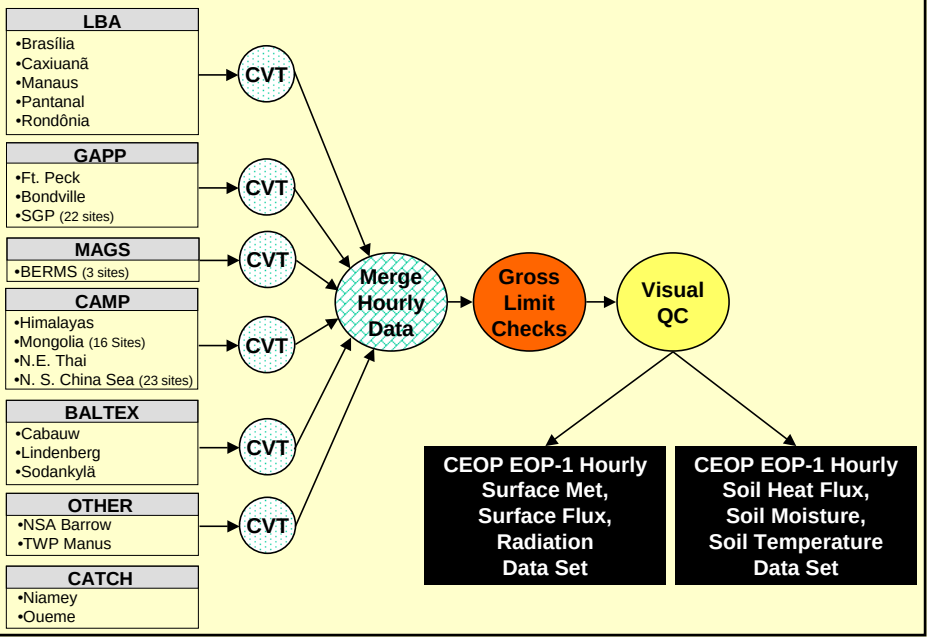


CEOP Reference Site Name and Information Links	[illegible]	[illegible], IL	Pt. Reyes, MT	Oak Ridge, TN	Mt. Bigelow, AZ	[illegible]
Site Latitude Longitude	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
MOLTS Locations	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
MOLTS Elevation (m)	[illegible]	[illegible]			[illegible]	[illegible]
Site Maps	X	X	X	X	X	X
Site Contacts	X	X	X	X	X	X
Site Status	A	A	A	A	A	A
Site Type	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Data Collection Period	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Single Data Set	X	X	X	X		X
Data Access	X					
RSF-1 Data Received	X	X	X			
UPPER AIR OBSERVATIONS						
Radiosonde (X = on-site, XX = off-site operational)	X	XX	XX	XX	XX	XX
Radar	X					
Lidar	X					
Profiler	X					
RASS	X					
SURFACE OBSERVATIONS						
Air Temperature	X	X	X	X	X	X
Humidity	X	X	X	X	X	X
Wind	X	X	X	X	X	X
Pressure	X	X	X	X	X	X
Precipitation	X	X	X	X	X	X
Skin Temperature	X	X	X	X	X	
Upward Shortwave Radiation	X	X	X	X	X	X
Downward Shortwave Radiation	X	X	X	X	X	X
Upward Longwave Radiation	X	X	X	X	X	X
Downward Longwave Radiation	X	X	X	X	X	X
Upward Photosynthetically Active Radiation						X
Downward Photosynthetically Active Radiation				X		X

REFERENCE SITE INVENTORY



EOP-1 "Composite" Data Set Development

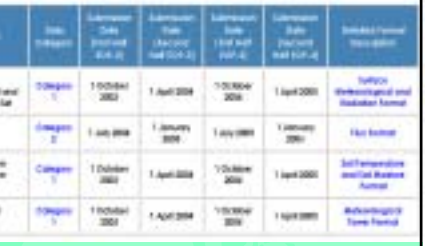


"Composite" Reference Site Dataset Issues from Berlin Meeting (April, 2003)

- Prepare cold weather precipitation questionnaire to document measurement/collection techniques, instrumentation, and data processing
- Update Reference Site documentation and metadata
- Finalize data parameters and file organization
- Finalize temporal resolution, averaging interval, and data formats
- Document data quality assurance/control procedures
- Document data collection and submission for CEOP Annual Cycle Datasets (schedule)

Reference Site Data Set Procedures Report (Approved by the SSC on 29 July 2003)

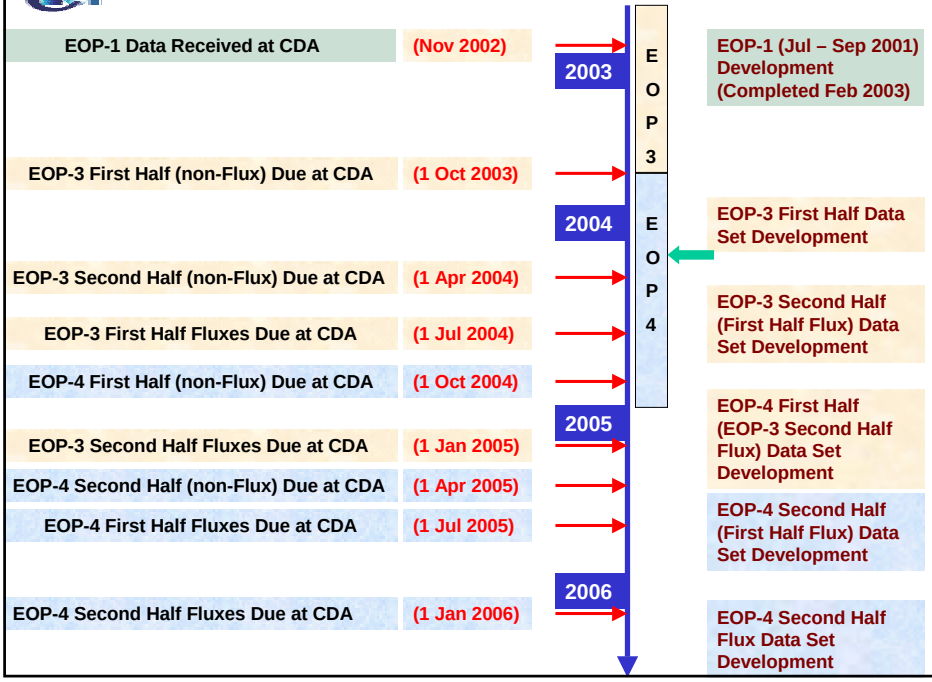
- Introduction
- Data Set Description, submission schedule, formats
- Metadata
- Gap filling
- Quality Control/Quality Assurance
- File naming convention
- Submission instructions to CDA
- CDA Composite formation
- CDA Quality Assurance
- Rawinsonde Data Sets
- Ancillary Data Sets



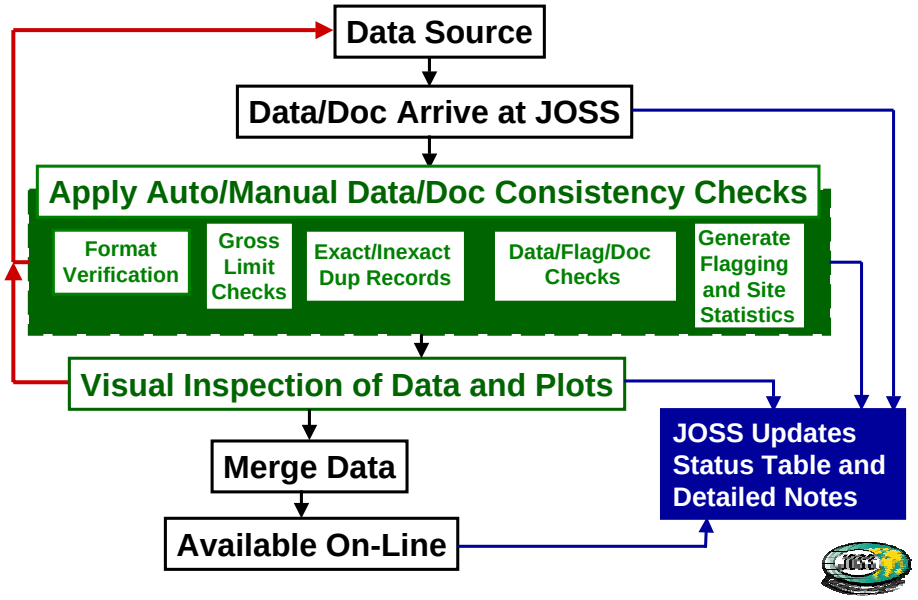
REFERENCE SITE FLUX DATA SET FORMAT

Parameter	C Format	Missing Value	Final Units/Equations/Notes
UTC Nominal Date/Time	16 chars	N/A	yyyy/mm/dd HH:MM, where MM is 00 or 30, only
UTC Actual Date/Time	16 chars	N/A	yyyy/mm/dd HH:MM
CSE Identifier	10 chars	N/A	Fill name with underscores, not spaces.
Reference Site Identifier	15 chars	N/A	Fill name with underscores, not spaces.
Station Identifier	15 chars	N/A	Fill name with underscores, not spaces.
Latitude	f10.5	-99.99999	decimal degrees. South is negative.
Longitude	f11.5	-999.99999	decimal degrees. West is negative.
Elevation	f7.2	-999.99	meters
Sensor Height	f7.2	-999.99	meters; Height of sensor. Positive above ground level. Negative below ground.
Sensible Heat Flux	f8.2	-999.99	W/m^2
Sensible Heat Flux Flag	1 char	M	See Flag values.
Latent Heat Flux	f8.2	-999.99	W/m^2
Latent Heat Flux Flag	1 char	M	See Flag values.
CO2 Flux	f8.2	-999.99	$\mu mol/m^2/s$
CO2 Flux Flag	1 char	M	See Flag values.
Soil Heat Flux	f8.2	-999.99	W/m^2
Soil Heat Flux Flag	1 char	M	See Flag values.

Reference Site Data Set Development Timeline



Reference Site EOP-3 Data Flow



UCAR/JOSS Manual Data and Documentation Consistency Checks

- Examine documentation and data files for completeness.
- Examine in-depth results from UCAR/JOSS automated checking software.
- Consistent CSE, Reference Site, and Station between data file name, data file metadata, and documentation file.
- Consistent station location information between doc and data.
- Consistent sensor heights between doc and data.
- Examine data flag usage.
- Verify data flags properly applied.
- Verify proper units used.
- Examine time series plots of every parameter at every station.

UCAR/JOSS Automated Data Consistency Checks and Statistics for CEOP

Consistency Checks Applied per the 29 July 2003 CEOP SSC Approved Formats:

- File Name in Proper Form (i.e.,CSE_RefSite_StnID_BeginDate_EndDate.suffix)
- File is Sorted Correctly
- File does Not Contain Exact Duplicate Records
- File does Not Contain In-Exact Duplicate Records (i.e., dates/times/names/lat/lons match but not data)
- Every Record is Correct Length
- No Control Chars in Any Record
- Verify Meta Data and Data Field Location and Justification (i.e., locations of slashes, colons, decimal points, spaces, etc.)
- Consistency between File Name and Meta Data and Data in File
- Consistency between Nominal and Actual Date/Times within Each Record
- Gross Limit Check each Meta Data and Data Value
- Check for Data Values of -9.00
- Sensor Height are Valid (Soils <= 0.00, Tower >= 0.00)
- Flag Values are Valid (U,G,D,B,C,M,or I only)
- Missing Data Value Has Missing Flag (i.e., C or M only)
- Identify Completely Missing Records
- Constant Station Location (lat/lon/elev)

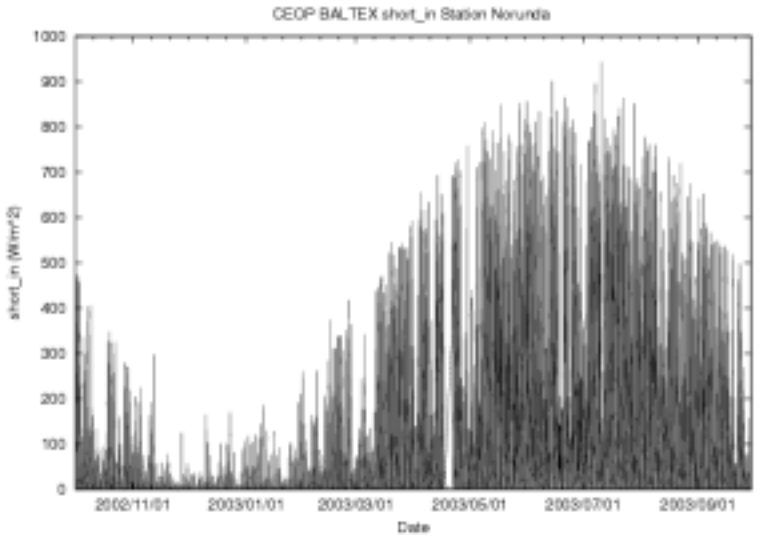
The following statistics are Produced:

- Extreme Values for Each Parameter
- Count & Percent of Each Flag Type for Each Parameter
- List of All Station/Lat/Lon Sets Found
- List of All Sensor Heights Found
- Tracks/Prints Station Information for Every Station
- Summarizes Total Error Counts for Data and Meta Data

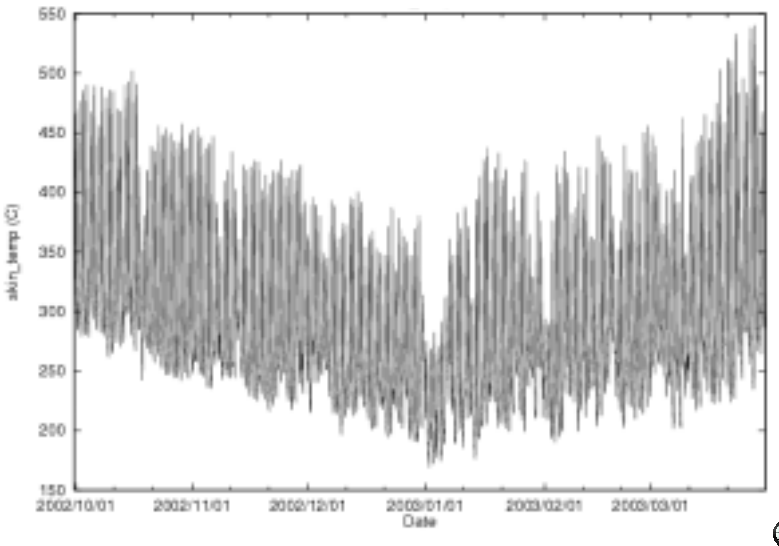
CEOP First Half EOP-3 Data and Documentation Issues

- Incomplete documentation and/or data files.
- Information in documentation file does not match data file (e.g. station locations, heights and depths of measurements).
- Incorrect data file format. Incorrect date format.
- Reporting 00:00 UTC observation as 24:00 UTC.
- Incorrect flag applied to missing data. Incorrect missing value used.
- Inexact duplicates present in the data (e.g. two records of data at the same reporting time but with different parameter values).
- Using missing values when deriving parameters.
- Large amounts of data flagged as Bad or Dubious without explanation in documentation.
- Incorrect units used. Incorrect parameters included in data.
- Sensor heights reported with incorrect sign.
- Reporting the surface pressure at all heights of the tower data.
- Having multiple measurements of the same parameter at the same height or depth without a way to differentiate.

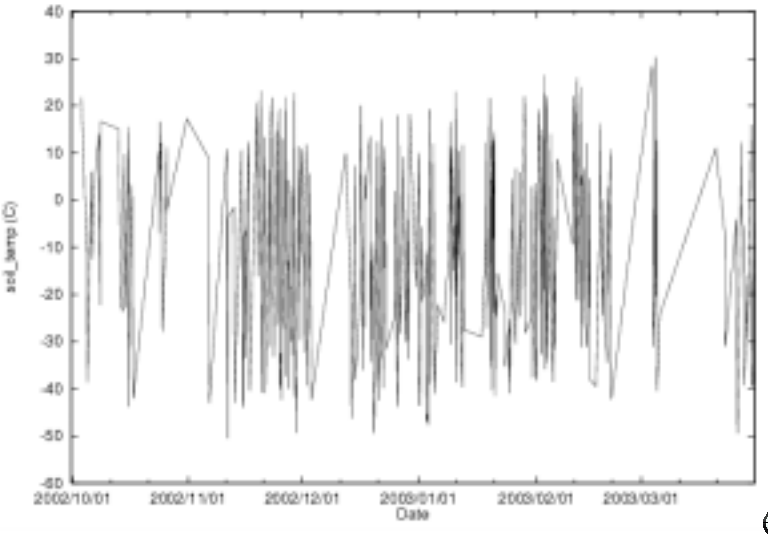
INCOMING SHORTWAVE



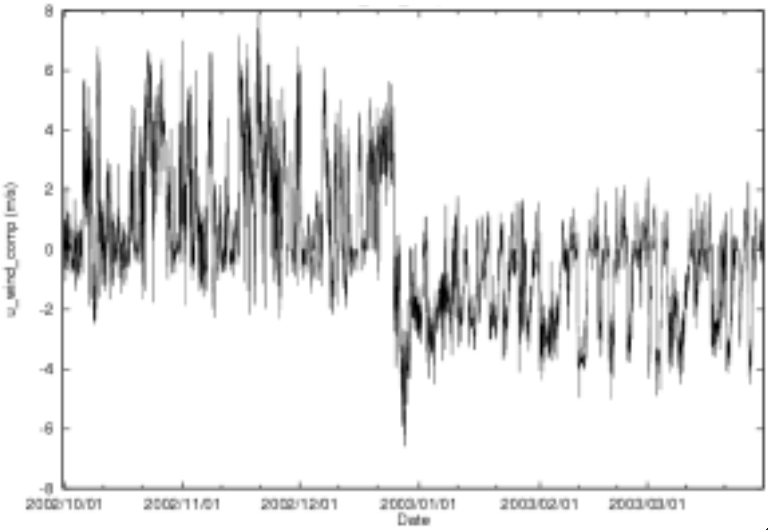
SKIN TEMPERATURE



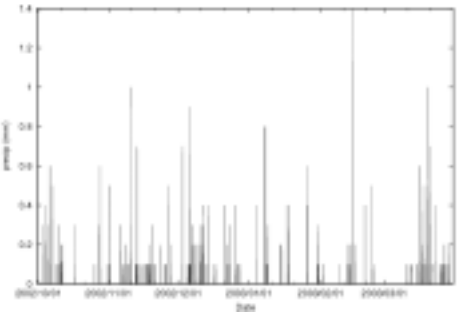
SOIL TEMPERATURE



U-WIND COMPONENT

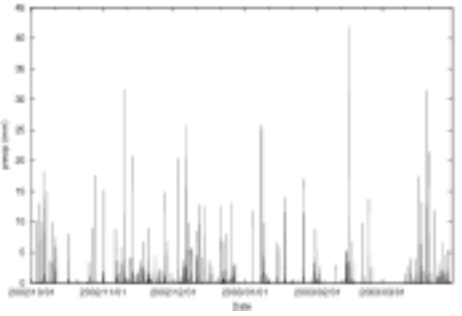


PRECIPITATION DATA CHANGES



BEFORE…

AFTER….



CEOP First Half EOP-3 Reference Site Data Set Status (as of 5 March 2004)

CAMP								
Reference Site	SFC		TWR		STM		FLX	
	Doc	Data	Doc	Data	Doc	Data	Doc	Data
Chao-Phraya								
Equatorial Island								
Himalayas								
Inner Mongolia								
Korean Haenam								
Korea Peninsula								
Mongolia								
NE Thailand								
NSCS-SJ								
Siberia Taiga								
Siberia Tundra								
Tibet								
W. Pacific Ocean								
Yangtze River								

CEOP First Half EOP-3 Reference Site Data Set Status (as of 5 March 2004)

Reference Site	SFC		TWR		STM		FLX	
	Doc	Data	Doc	Data	Doc	Data	Doc	Data
LBA								
Brasilia								
Caxiuana								
Manaus								
Pantanal								
Rondonia								
Santarem								
GAPP								
Bondville								
Fort Peck								
Mt. Bigelow								
Oak Ridge								
ARM SGP								

CEOP First Half EOP-3 Reference Site Data Set Status (as of 5 March 2004)

Reference Site	SFC		TWR		STM		FLX	
	Doc	Data	Doc	Data	Doc	Data	Doc	Data
BALTEX								
Cabauw								
Lindenberg								
Norunda								
Sodankyla								
MAGS								
BERMS								
MDB								
Tumbarumba								
CATCH								
Niamey								
Oueme								
OTHER/ARM								
NSA								
TWP								

Baseline Surface Radiation Network

Archiving

Provisional &Proposed

Goal:

To acquire the highest possible quality, globally-diverse, surface-based radiation measurements for climate research.

Measurements

- **Direct & diffuse solar***
- **Downward IR ***
- **Upwelling irrad.**
- **PAR & UV**
- **Upper air met.**
- **Aerosol optical depth**
- **Surface meteorology***

*all sites

Network Status

- 2561 station-months of data
- 34 archiving sites + 15 potential
- Zurich/ETHZ archive extended
- New Brazilian network
- New site – SIRTA - France
- 2 new Canadian sites proposed
- Progress on a China site(s)
- Potential eventual Siberian site
- GCOS invitation
- July 2004 Mtg. in Exeter U.K.

http://BSRN.ETHZ.CH

BSRN Data

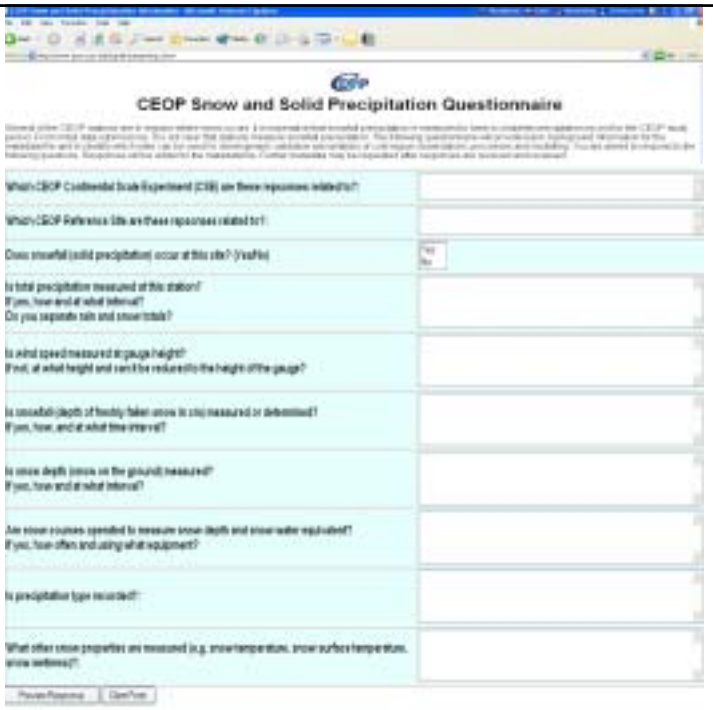
Parameters vary by station.

Base data includes:

- **metadata (location, PI, station characteristics, Instrumentation details)**
- **Radiation data at 1-minute resolution (global, direct, diffuse, downward longwave, air temperature, RH, pressure)**

Other data might include:

- **spectral shortwave, upward shortwave, upward longwave**
- **net radiation**
- **UV**
- **SYNOP**
- **radiosonde**
- **ozone**
- **cloud amount and/or heights**
- **tower measurements**



CEOP Snow and Solid Precipitation Questionnaire

Which CEOP Continental Scale Experiment (CSE) are these repsonses related to?

Which CEOP Reference Site are these repsonses related to?

Does snowfall (solid precipitation) occur at this site? (Yes/No)

Is total precipitation measured at this station?
If yes, how and at what interval?
Do you separate rain and snow totals?

Is wind speed measured at gauge height?
If not, at what height and can it be reduced to the height of the gauge?

Is snowfall (depth of freshly fallen snow in cm) measured or determined?
If yes, how, and at what time interval?

Is snow depth (snow on the ground) measured?
If yes, how and at what interval?

Are snow courses operated to measure snow depth and snow water equivalent?
If yes, how often and using what equipment?

Is precipitation type recorded?

What other snow properties are measured (e.g. snow temperature, snow surface temperature, snow wetness)?

REFERENCE SITE DATA ISSUES (1)

- Metadata and Documentation lacking → need complete specific site descriptions (e.g. land cover, topography, soils, etc.)
- Quality Control and Assurance → most problems are now not in formatting
- Inclusion of BSRN data to CEOP "composite" format → What sites are needed?
- Upper Air Sounding data available in "native" formats → Do we need a common format?
- Data Source turn-around time → Reduce delays and improve data availability

REFERENCE SITE DATA ISSUES (2)

- Prepare Individual Reference Site Reports to be provided to Data Sources
- Organization of Data → Sorting Preference (e.g. dataset type, individual stations, all), CD-ROM?
- Data Integration Plans → Determine alternative data format(s), delivery mechanisms (e.g. DODS/GRaDS servers)
- Data Policy Issues → Acknowledgements and Citations, Do we need a "broader" CEOP Data Policy?
- Need for hydrology data at Reference Sites → Define new hydrology Reference Sites?

Coordinated Enhanced Observing Period (CEOP) Model Output and Information

CEOP Model Data Sets and Information

Model Output Data Sets

- CEOP Model Data Gateway
- NASA/DSMC Data Assimilation Office CEOP Data
- NCEP GFS (AVN/MRF) MOLTS output
- GLDAS CEOP EOP-1 MOLTS output
- CEOP Model Output Source Agency Links

CEOP Model Output Teleconference Notes

- Tenth Formal Telecon (3 December 2003) (DRAFT)
- Ninth Formal Telecon (29 October 2003) (DRAFT)
- Eighth Formal Telecon (17 September 2003)
- Seventh Formal Telecon (30 July 2003)
- Sixth Formal Telecon (2 June 2003)
- Fifth Formal Telecon (28 April 2003)
- Second (17 Dec 2002), Third (3 Feb 2003), and Fourth (11 Mar 2003) Formal Telecons
- First Formal Telecon (15 November 2002)

General Information

- Guidance for CEOP Model Output Generation at NWP Centers, Met Agencies and the DAO (10 Dec 2002)
- Model Output Variables Requested by CEOP (28 Mar 2003)
- MP13 fulfillment (1 Nov 2002)
- Listing of CEOP proposed MOLTS locations
- Map of CEOP proposed MOLTS locations

CEOP Model Center Documentation

BoM

- None

CPTEC

- CPTEC Contribution to CEOP (10 Dec 2002)

ECMWF

- ECMWF Contribution to CEOP (13 Dec 2002)
- ECMWF CEOP MOLTS locations (13 Dec 2002)

ECPC

- ECPC CEOP Contributions (30 May 2003)
- ECPC Model Characteristics (30 May 2003)
- ECPC Model Output Times (30 May 2003)
- ECPC CEOP Variables and Processes (30 May 2003)
- ECPC (RSM/R) MOLTS Characteristics (30 May 2003)

JMA

- JMA Contribution to the CEOP Dataset (17 Dec 2002)
- Additional JMA Comments and Questions on the CEOP Dataset (17 Dec 2002)
- JMA CEOP MOLTS locations (16 Dec 2002)
- Vertical Levels of JMA CEOP Output Data (16 Dec 2002)
- Elements on GRIB2 from JMA Operational 3-DVAR Global Analysis (16 Dec 2002)
- Sample of JMA MOLTS Output (3 Feb 2003)

NASA Global Modeling and Assimilation Office (GMAO; formerly DAO)

- File Specification for the GEOS-DAS Gridded Output (Version 4.5) (4 June 2003)

NASA GLDAS

- None

NCEP Operational

- Output for CEOP from NCEP Global Data Assimilation and Forecast Model (14 Mar 2003)
- NCEP Global Forecast System Implementation (29 Oct 2002)
- NCEP GRIB Table 2 (14 Mar 2003)
- NCEP Output at International CEOP MOLTS Sites (14 Mar 2003)
- NCEP Global Model Characteristics at CEOP MOLTS Reference Sites (14 Mar 2003)
- Vertical Sigma Levels of NCEP Global Model MOLTS Output (14 Mar 2003)
- NCEP data at MP1 (28 Mar 2003)
- NCEP CEOP Data GEWEX Storage (28 Mar 2003)
- Map of proposed NCEP ETA MOLTS locations for NAME
- Map of current NCEP ETA MOLTS locations around the ARM SGP site

NCMRWF

- NCMRWF Notes on Gridded Data (19 Nov 2003)
- NCMRWF CEOP Core Codes (19 Nov 2003)

UKMO

- Met Office Contribution to CEOP (12 Nov 2002)
- MMO GRIB Table 2 for CEOP Data (28 Mar 2003)
- Model Locations of MOLTS Reference Sites (28 Mar 2003)
- NetAir for CEOP MOLTS and Gridded data from UKMO (12 Jun 2003)
- Met Office Scientific Advisory Committee Presentations (28 Mar 2003)

CEOP- Model Output Access

Web-Interface

MODEL OUTPUT ARCHIVE STATUS

(24 February 2004)

CEOP EOP-1 Satellite Datasets

Data Policy

Our data policy is following CEOP Data Release Guidelines.
Please look the **CEOP Data Release Guidelines in detail**.

Link to CEOP EOP-1 Satellite Dataset (Updated Aug. 11 2003)

Dataset Documentation for

- DMSP/ SSM/I (L1B) Low Frequency data (Updated Aug. 11 2003)
- DMSP/ SSM/I (L1B) High Frequency data (Updated Aug. 11 2003)
- TRMM TMI (L1B11) Low Frequency data (Updated Aug. 11 2003)
- TRMM TMI (L1B11) High Frequency data (Updated Aug. 11 2003)
- TRMM PR(L2A25) data (ADDED Mar. 15 2003) (Updated Aug. 11 2003)
- GMS/ 5/ VISSR Visible (Updated Aug. 11 2003)
- GMS/ 5/ VISSR Infrared (Updated Aug. 11 2003)
- NOAA/ AVHRR Visible (Updated Aug. 11 2003)
- NOAA/ AVHRR Infrared (Updated Aug. 11 2003)

- Header Format (2048 Bytes of the data) (Updated Aug. 11 2003)

Reference Site Information

- Latitude, Longitude and Site Code Ver.03 (Updated May. 09 2003)
- Map of reference site (gif)