ICOADS: Update Status and Data Distribution

Steven J. Worley

Scott D. Woodruff

Sandra J. Lubker

Ziahua Ji

J. Eric Freeman

NCAR, NOAA/ESRL, NOAA/NCDC

CLIMAR-III, 6-9 May 2008, Gdynia, Poland

NCAR

CISL
Computational and Information Systems Laboratory

Topic Outline

- Data Distribution
  - *How well are we doing? Metrics*
- Release 2.5 (new)
  - *Introduction and Status*
  - *Next Presentation, "Data Characteristics and Future Directions", Woodruff et al.*

Data Distribution

- All available options @ http://icoads.noaa.gov/

International COADS

Jump to: ICOADS

Search for: Go!

Search options... Contact us Privacy policy

ETMC

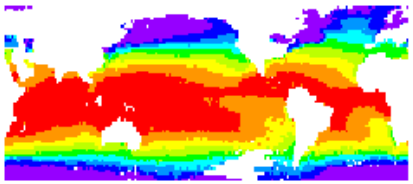
RECLAIM

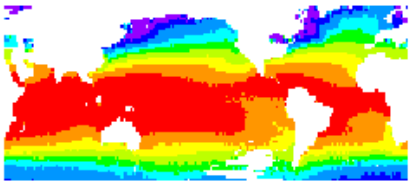
Events:

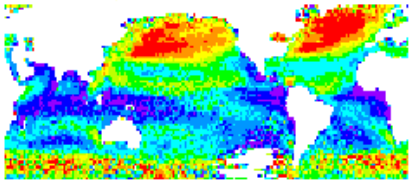
CLIMAR-III
6-9 May 2008,
Gdynia, Poland

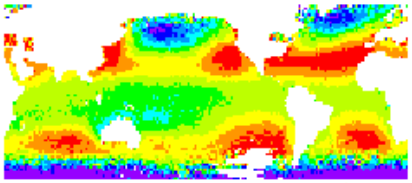
MARCDAT-II
17-20 Oct. 2005,
Exeter, UK

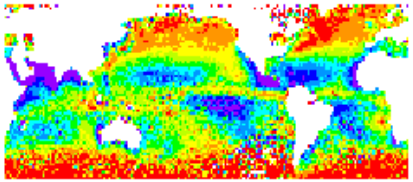
International Comprehensive Ocean-Atmosphere Data (ICOADS)


sea surface temperature


air temperature


wind


sea level pressure


cloudiness

Chronology and News
Project Status
Related Data and Resources

Data and Metadata
Publications
Contact Points

CISL
Computational and Information Systems Laboratory

Data Distribution, NCAR

CAR

Data

- All Observations
- All Monthly Summary Statistics (1°, 2°)

- Observations as Example (ditto for Stats)
  - *Direct File download*

- Fortran language code **software**.
- IMMA **documentation** .

View/Download Selected Files | Perl Download Script | Csh Download Script

INDEX	File Name	Description	Size	Data Format	Archive Format	Date Archived
1	IMMA_2.4_1784_1799.tar	178402-179912	567.3K	IMMA	Z.TAR	07/19/2007
2	IMMA_2.4_1800_1849.tar	180001-184912	42.1M	IMMA	Z.TAR	07/19/2007
3	IMMA_2.4_1850_1899.tar	185001-189912	275.8M	IMMA	Z.TAR	07/19/2007
4	IMMA_2.4_1900_1929.tar	190001-192912	740.5M	IMMA	Z.TAR	07/19/2007
5	IMMA_2.4_1930_1949.tar	193001-194912	624.6M	IMMA	Z.TAR	07/19/2007
6	IMMA_2.4_1950_1959.tar	195001-195912	607.4M	IMMA	Z.TAR	07/19/2007
7	IMMA_2.4_1960_1969.tar	196001-196912	1.1G	IMMA	Z.TAR	07/19/2007
8	IMMA_2.4_1970_1974.tar	197001-197412	658.3M	IMMA	Z.TAR	07/19/2007
9	IMMA_2.4_1975_1979.tar	197501-197912	682.9M	IMMA	Z.TAR	07/19/2007

CISL Computational Systems Laboratory

Data Distribution, NCAR

CAR

Sub-selection Interface IMMA observation example

- Temporal and spatial selection

CISL
Computational
Systems Labora

Data Distribution, NCAR

NCAR

- QA/QC, trimming flag, data mixture choices
  - *Seeded with 'Enhanced Filtering'*

- ○ Standard Filtering
- ◉ Enhanced Filtering
- ○ Make your own Filtering Selection

day night options	☑ day night obs.	☐ night obs. only	☐ day obs. only	
platform type options	☐ ships obs. only	☑ ships + buoys + others		
source exclusion flags	☑ used	☐ ignored		
composite QC flags	☑ used	☐ ignored		
outlier trimming level	☐ 2.8 sigma	☐ 3.5 sigma	☑ 4.5 sigma	☐ untrimmed

CISL
Computational and Information Systems Laboratory

Data Distribution, NCAR

- Pick variable from IMMA Core section

The IMMA Core

Reset Core Selection

Click imma_short.pdf for detail desciption of the IMMA Core Fields.

☑ YR	year UTC	☑ MO	month UTC	☑ DY	day UTC
☑ HR	hour UTC	☑ LAT	latitude	☑ LON	longitude
☐ IM	IMMA version	☐ ATTC	attm count	☐ TI	time indicator
☐ LI	latitude/long. indic.	☐ DS	ship course	☐ VS	ship speed
☐ NID	national source indic.	☐ II	ID indicator	☐ ID	identification/call sign
☐ C1	country code	☐ DI	wind direction indic.	☐ D*	wind direction
☐ WI	wind speed indicator	☐ W*	wind speed	☐ VI	VV indic.
☐ VV*	visibility	☐ WW	present weather	☐ W1*	past weather
☐ SLP*	sea level pressure	☐ A*	characteristic of PPP	☐ PPP*	amt. pressure tend.
☐ IT	indic. for temperatures	☐ AT*	air temperature	☐ WBTI	indic. for WBT
☐ WBT*	web-bulb temperature	☐ DPTI	DPT indic.	☐ DPT*	dew-point temp.
☐ SI	SST meas. method	☐ SST*	sea surface temp.	☐ N*	total cloud amount

CISL Computational Systems Laboratory

Data Distribution, NCAR

- Pick record ID information from ICOADS Attachment

ICOADS attm

Clear&Close ICOADS attm Selection | Reset ICOADS attm Selection

Click imma_short.pdf for detail description of ICOADS attm Fields.

☐ B10	10 degree box number	☐ B1	1 degree box number	☑ DCK	deck
☑ SID	source ID	☑ PT	platform type	☐ DUPS	dup status
☐ DUPC	dup check	☐ TC	track check	☐ PB	pressure bias
☐ WX	wave period indicator	☐ SX	swell period indicator	☐ C2	2nd country code

CISL
Computational and Information Systems Laboratory

Data Distribution, NCAR

- Other IMMA attachments
  - *Ship Metadata (Kent and Berry, 1973-2006)*

CISL
Computational and Information Systems Laboratory

Data Distribution, NCAR

- Select Auxiliary Data to supplement ICOADS
  - *Data sources not in ICOADS, any IMMA format data*
  - *Believed to be:*
    - Convenient data service for collaborators
    - Early access to new data for users

ICOADS Auxiliary Data Selection (Optional)

DS Name	DS Type	DCK	SID	Date Range*	Description
☑ ICOADS	ICOADS	1-999	1-120	None	data in ICOADS archive
☐ RussRV	New	735	64-65	1936-2000	Russian research vessel obs.
☐ JapanWhaling	New	761	115-116	1946-1984	Japanese Whaling Ship Data
☐ PMELJAMSTEC	Replacement	144	117-120	1990-1997	PMEL/JAMSTEC buoy data
☐ COAPSRV	New	740	130	1990-1998	COAPS research vessel obs.
☐ USMMJ	New	704	125	1878-1894	US Marine Meteorological Journals
☐ RNWW2	New	245	126	1936-1948	UK Royal Navy WW II Logs
☐ CLI21	New	730	124	1662-1855	Climatological Database for the World's Oceans

* Available date ranges of the listed data sources that are NOT included in ICOADS yet

Data Metrics, NCAR

From the Observation Interface

NCAR ICOADS Observation Interface Requests (IMMA Core + ICOADS Attm)

Year	Enh. Filter	Stand. Filter	Self-def. Filter	Total Users	Other Attms	Aux. Data
2007	**317**	49	33	**399**	26	**5**

Percentage of Variables Requested Through Observation Interface (11/05-03/08)

	SST	W	D	AT	SLP	DPT	WH	WD	SH	WW	SD
%	**75**	**67**	**65**	**60**	**52**	35	**30**	**27**	**25**	25	**25**

- 70-80% Users choose Enhanced Filtering (recommended choice)
- About one request per day (399)
- Lower than anticipated use of Auxiliary Data (5)
- Top % variable, expected (SST, W, D, AT, SLP)
- Interesting 25% request wave data (WH, WD, SH, SD)

Data Metrics, NCAR

From the Observation Archive, File Download

NCAR ICOADS Observation Archive File Downloads		
Year	**Unique Users**	**Data Volume (GB)**
2007	**192**	619

- Surprising number of Unique Users download files (192)

CISL
Computational and Information Systems Laboratory

Data Metrics, NCAR

From the Monthly Summary Statistics Interface

NCAR ICOADS Monthly Summary Statistics Interface			
Year	**Requests**	**Unique Users**	**Data volume (GB)**
2007	**251**	**101**	13

Percentage of Variable Requested Through Statistics Interface
(11/05-03/08)

	SST	AT	W	V	U	SLP	R	C
%	**83**	36	32	27	27	19	16	15

- Requests (251), Users (101) => 2 to 3 requests per user
- SST dominates as the preferred variable, over 80%
- About 90% choose Enhanced Statistics (not shown)

Data Metrics, NCAR

From the Monthly Summary Statistics Archive, File Downloads

NCAR ICOADS Monthly Summary Statistics File Downloads		
Year	**Unique Users**	**Data Volume (GB)**
2007	**68**	41

- Not many Unique Users for archive files (68)

CISL
Computational and Information Systems Laboratory

Data Metrics, NCAR

Top 10 International Rankings

NCAR

Countries Using the Interfaces
(Obs. + Statistics, 11/05-03/08)
Ordered by Number of Requests

Rank	Country	Unique Users	Requests	Data Volume (GB)
1	U.S.A.	163	768	185
2	U.K.	30	107	14
3	INDIA	11	84	5
4	CANADA	20	69	20
5	BRAZIL	7	57	5
6	JAPAN	11	37	40
7	SPAIN	13	32	3
8	P.R.CHINA	19	29	8
9	PORTUGAL	9	28	.3
10	GERMANY	9	26	10
Total		**385**	**1448**	**376**

Countries Downloading Archive Files
(Obs. + Statistics, 11/05-03/08)
Ordered by Number of Files Downloaded

Rank	Country	Unique Users	Files	Data Volume (GB)
1	U.S.A.	158	2101	422
2	JAPAN	27	747	175
3	P.R.CHINA	38	455	127
4	FRANCE	15	434	29
5	U.K.	29	261	95
6	TAIWAN	8	144	61
7	GERMANY	14	141	66
8	INDIA	19	135	78
9	AUSTRALIA	7	118	35
10	CUBA	1	104	24
Total		**443**	**5213**	**1268**

- Rankings vary between preference, interfaces .vs. file downloads
  - U.S. is ranked #1 in both - good for U.S. budget justifications
- Many Unique Users in both cases (385 and 443)
  - Both services are important

CISL
Computational and Information Systems Laboratory

Data Metrics, NCAR

NCAR

Summing it all up

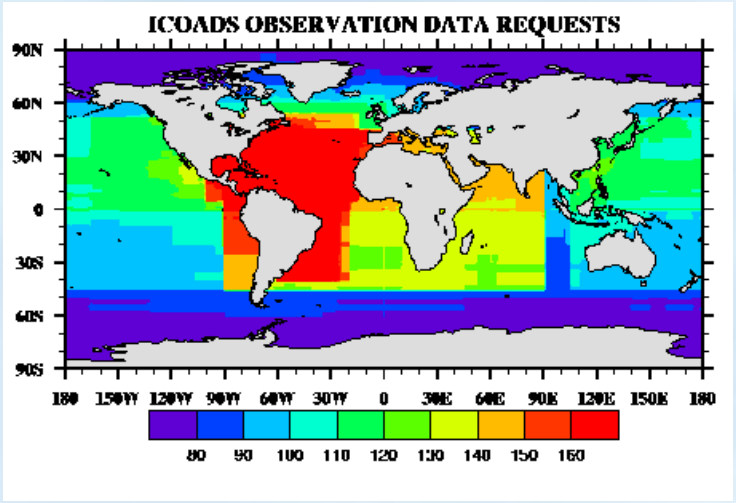
Grand Totals All Products (11/05-03/08)		
	Unique Users	Data Volume (GB)
TOTAL	**723**	1693
% Monthly Summary Statistics	**38%**	10%
% Observations	**62%**	90%

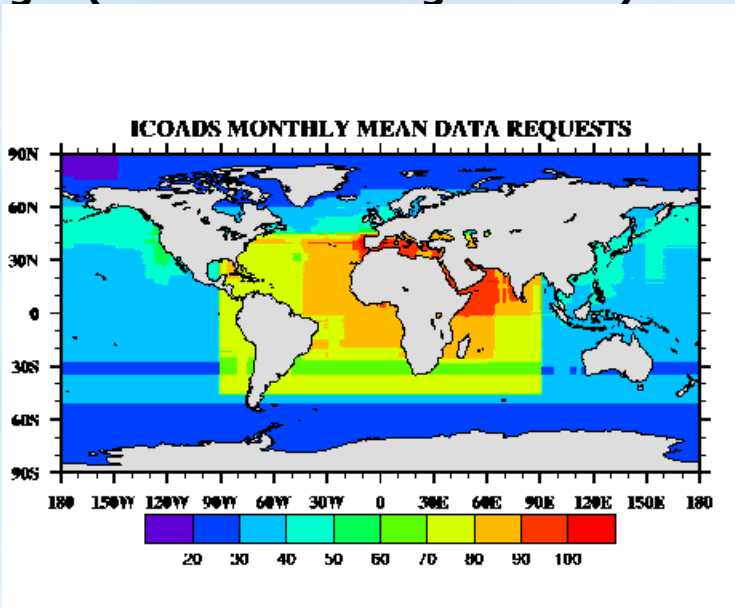
- Over about 2.5 years
  - 700+ Unique Users
- ~ 40% use MSS, ~ 60% use Observations

CISL
Computational and Information Systems Laboratory

Data Metrics, NCAR

What areas of data are most interesting? (From subsetting metrics)





Features:

- Observations - Atlantic Ocean
- Monthly Summary Statistics - Mediterranean Sea, Northwest Indian Ocean

CISL
Computational and Information Systems Laboratory

Data Distribution, NOAA/ESRL

NCAR

Products:

- Monthly Summary Statistics Files (1°,2°)
  - *netCDF format*
- Long-term Means (LTM) (2°)
  - *netCDF format*
- Near-Real Time (NRT) ICOADS-like extensions
  - *Based on GTS, abbreviated (limited statistics and variables) format*
  - *Observations,*
    - ASCII format
  - *Monthly Summary Statistics*
    - ASCII and netCDF format

CISL
Computational and Information Systems Laboratory

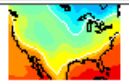
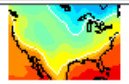
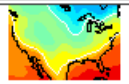
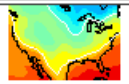
Data Distribution, NOAA/ESRL

NCAR

Illustration of Web Access

ICOADS 2-degree Enhanced

Sea Surface Temperature

Create a plot or subset.	Statistic	Start Date	End Date	Level	Link to files
 Make plot or subset	Monthly Fifth Sextile	1800/1	2007/5	Surface	sst.sextile5.nc
 Make plot or subset	Monthly First Sextile	1800/1	2007/5	Surface	sst.sextile1.nc
 Make plot or subset	Monthly Fraction of Observations in Daylight	1800/1	2007/5	Surface	sst.day_fraction.nc
 Make plot or subset	Monthly Mean	1800/1	2007/5	Surface	sst.mean.nc

- Point and click with one statistic (e.g. mean) per file
- Full record length (1800-5/2007) in one file
- Very handy for climate studies and analysis tools (NCL, GRADS, MatLab, NCO, etc)

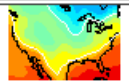
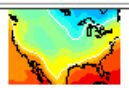
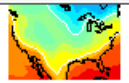
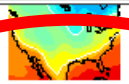
CISL
Computational and Information Systems Laboratory

Data Distribution, NOAA/ESRL

NCAR

ICOADS 2-degree Enhanced

Sea Surface Temperature

Create a plot or subset.	Statistic	Start Date	End Date	Level	Link to files
 Make plot or subset	Monthly Fifth Sextile	1800/1	2007/5	Surface	sst.sextile5.nc
 Make plot or subset	Monthly First Sextile	1800/1	2007/5	Surface	sst.sextile1.nc
 Make plot or subset	Monthly Fraction of Observations in Daylight	1800/1	2007/5	Surface	sst.day_fraction.nc
 Make plot or subset	Monthly Mean	1800/1	2007/5	Surface	sst.mean.nc

Features:

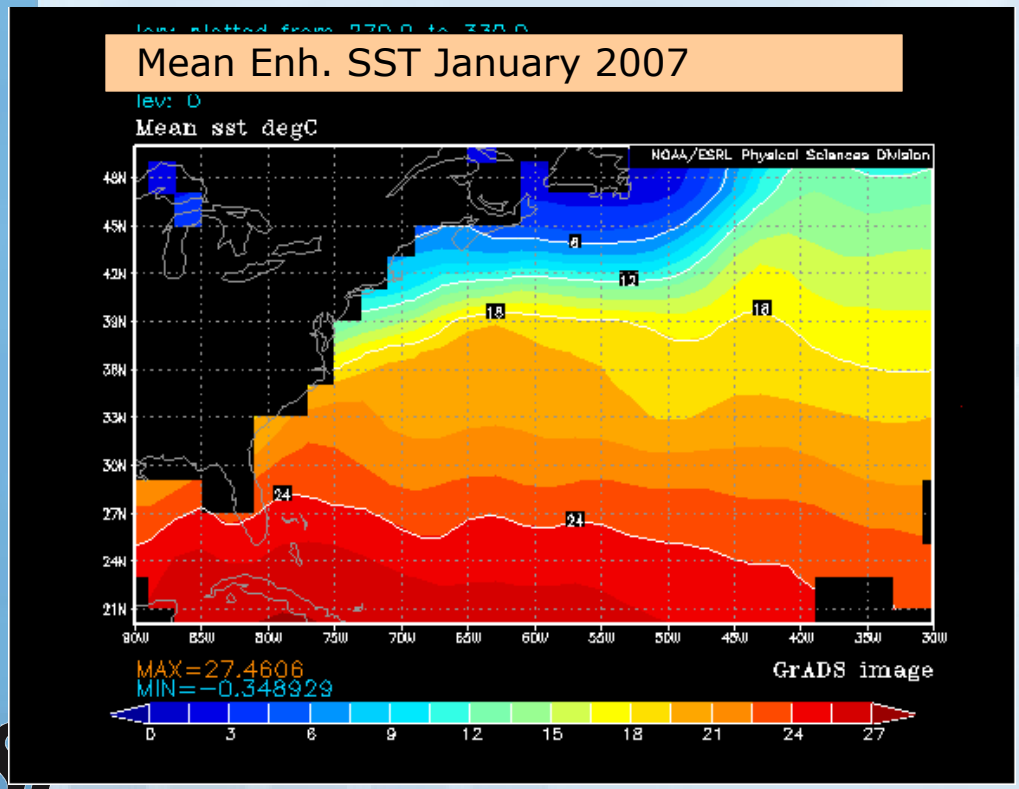
- Make plot or create subset file

CISL
Computational and Information Systems Laboratory

Data Distribution, NOAA/ESRL

NCAR

- Select region and time
- Create plot(s), and/or output dataset in netCDF



Computational and Information Systems Laboratory

Data Distribution, NOAA/ESRL

NCAR

Web Access @ NOAA/ESRL
Percentage of 2007 Access
Total = **560**K Accesses (files, images, subsets)
Estimates from Web Logs

	Monthly Summary Statistics	ΛTM∏c	NRT	Total
HTTP	**12**	1	**65**	78
OPeNDAP	**16**	< 1	< 1	16
OPeNDAP Server	< 1	< 1	< 1	< 1
LAS		**6**		6

Web Highlights

- NRT, most desired (65%)
- MSS; HTTP(12%), OPeNDAP (16%)
- LAS, 6% across products

FTP Highlights

- 25K Transfers
- MSS and NRT Monthly, most desired (10K & 14K)

FTP Access @ NOAA/ESRL
2007 (scaling based on 10 month record)
Estimates from Web Logs

	Monthly Summary Statistics	ΛTM∏c	NRT Monthly	NRT Obs.	Total
Transfers (K)	**10**	.8	**14**	.8	**25**
Volume (GB)	357	< 1	30	8	395

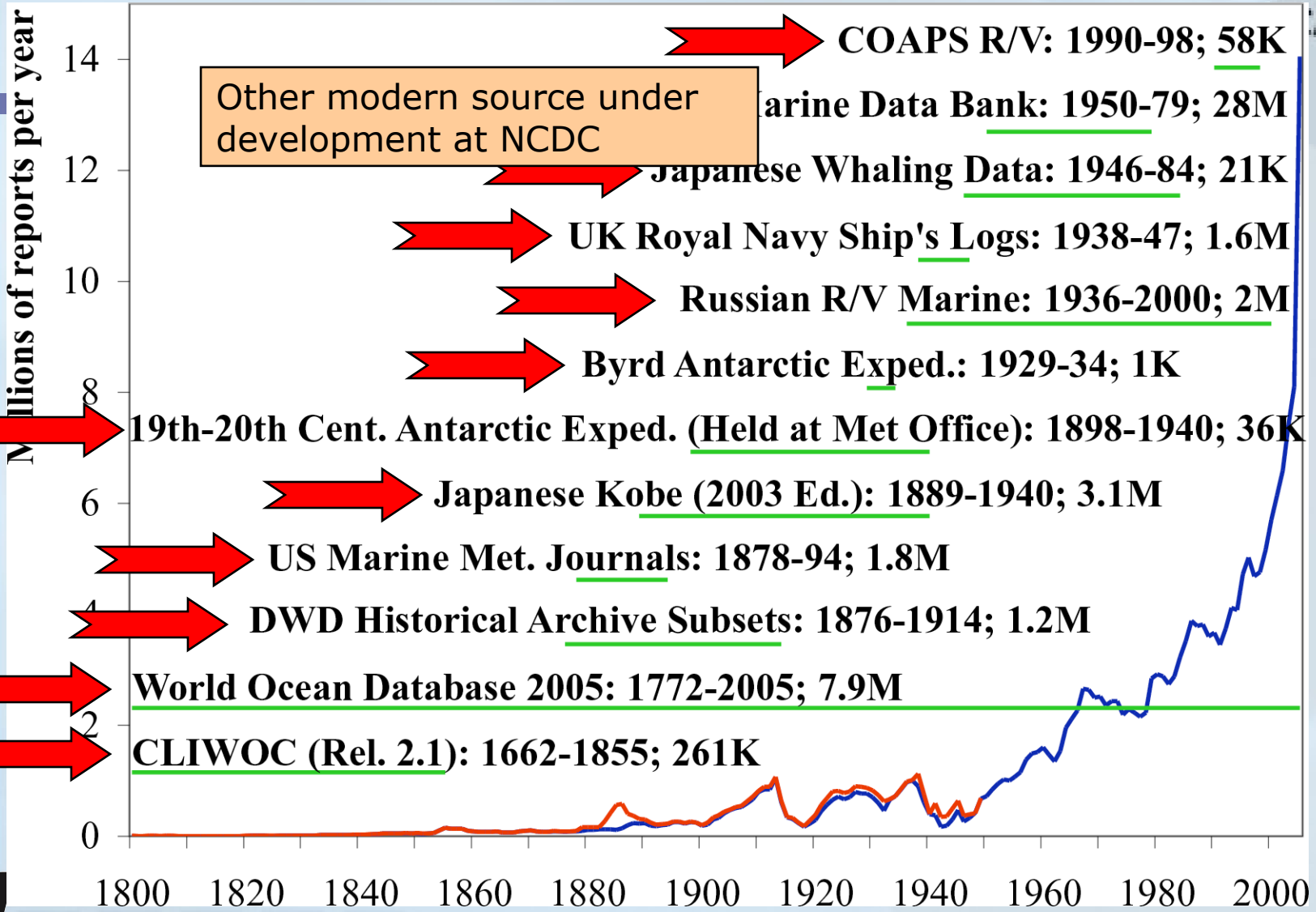
CISL
Computational and Information Systems Laboratory

Release 2.5 Status, New Release '08

Release 2.5, Long-period reprocessing 1662 -> 2000

- Progress to date
  - *Period of record completed, 1662 - 1949 (obs. and summaries)*

Release 2.5 Status, New Release '08



CISL
Computational and Information Systems Laboratory

Future for Data Distribution and Release 2.5

NCAR

Data Distribution

- Find better ways to promote Auxiliary Data usage
  - *More obvious presence on the interface?*
  - *Automatic inclusion with proper documentation?*
- Work toward a routine ICOADS near-real time update
  - *Possibly monthly*
  - *Retire 'light weight' NRT products currently available*
  - *Create compliant MSS and Observation Products*
- Provide observational data from NCDC

CISL
Computational and Information Systems Laboratory

Future for Data Distribution and Release 2.5

NCAR

Release 2.5

- Complete Release 2.5 data processing
- Refresh archives at the partners sites
- Notify Users of Release 2.5 availability
  - *Automated from 'registration' data base at NCAR*
- NOAA/ESRL has proposed a press release for ICOADS Release 2.5
- BAMS Article
  - *Probably good idea, expanded POR, new data sources*

CISL
Computational and Information Systems Laboratory

Summary

- ICOADS data distribution is going well
  - *Reaching the International community*
  - *Find MSS and Observations archives to both be important*
  - *Find file download and subsetting both to be important*

- Completing Release 2.5 is still a big task, but it is underway

International COADS

- Global community has/is adding great value ICOADS

NCAR

END

worley@ucar.edu

http://icoads.noaa.gov/

CISL
Computational and Information Systems Laboratory

Roles at the US Partners

NOAA/ESRL

- ICOADS Project Coordination and Home Website
- Documentation and IMMA format development
- User software, Release data processing
- Data Access

NCAR

- Long-term data preservation
- Data Access

NOAA/NCDC

- Long-term data preservation
- Real-time data collation
- Data rescue, i.e. CDMP

All

- Source data preparation

Acknowledge International Support

International COADS

- **Global community has/is adding great value ICOADS**

CISL
Computational and Information Systems Laboratory

Data Distribution, NCAR

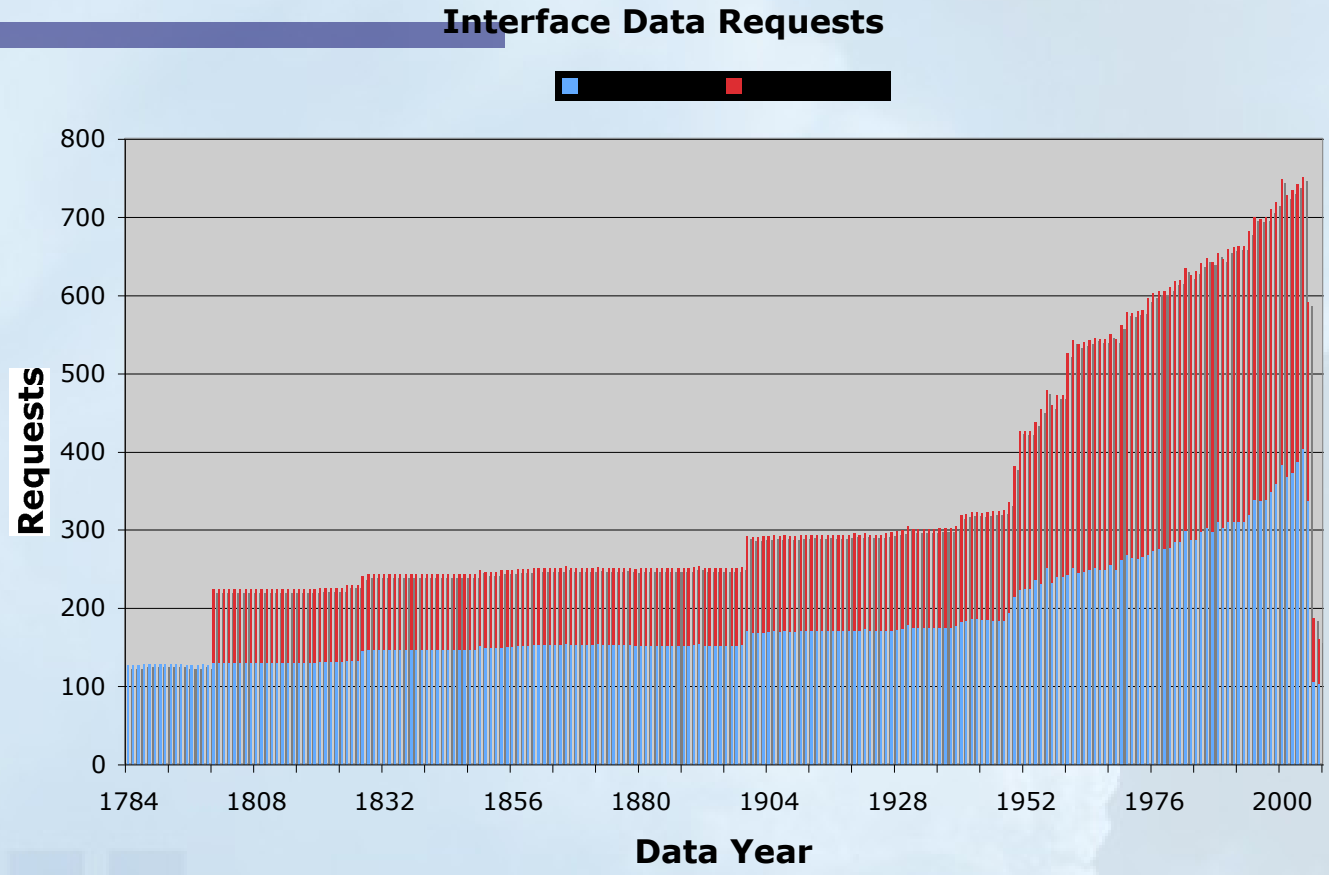
NCAR

- Sub-set Process
  - *Delayed processing - run in ~ 5 minute*
  - *Email notification*
  - *Users download*
  - *ASCII data*

CISL
Computational and Information Systems Laboratory

Data Metrics, NCAR

What time periods of data are most interesting? (11/05-03/08)



- Number of requests that touched specific data years
- 128 Requests go back to 1784 - climate interests are strong
- More interest in the modern period