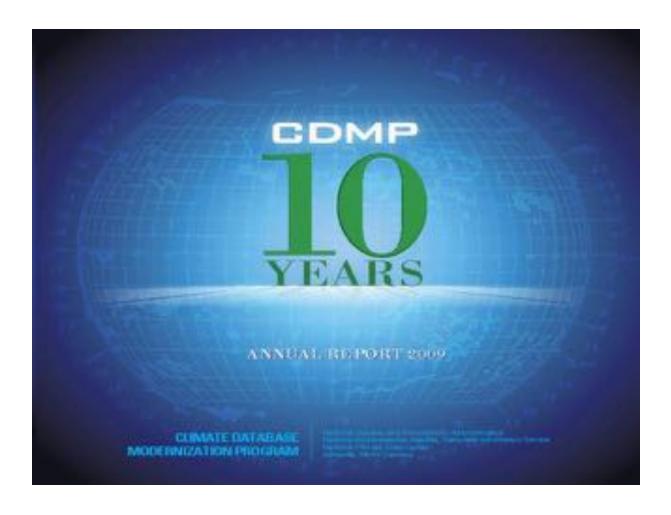


NOAA's Climate Database Modernization Program- More than 10 Years of Data Rescue Activities







ACRE Meeting 3-5 November 2010 -- Tom Ross, CDMP Program Manager



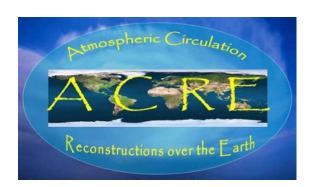






International Partners

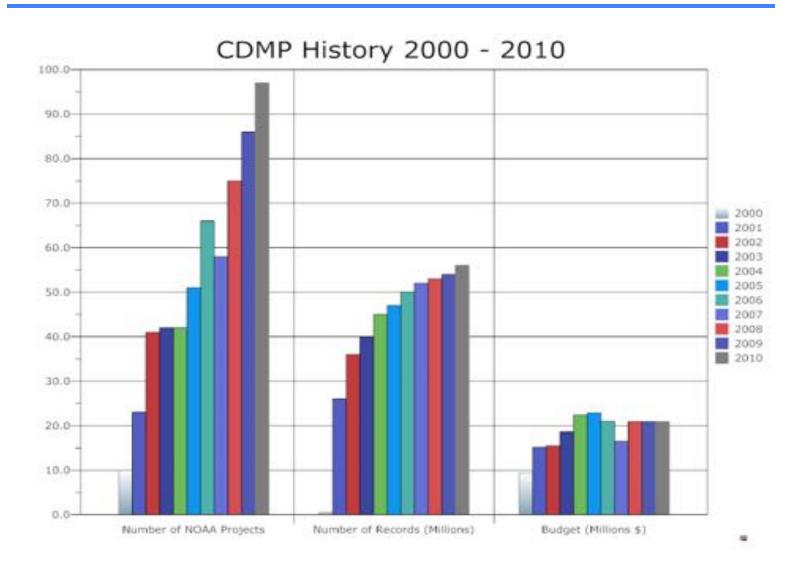




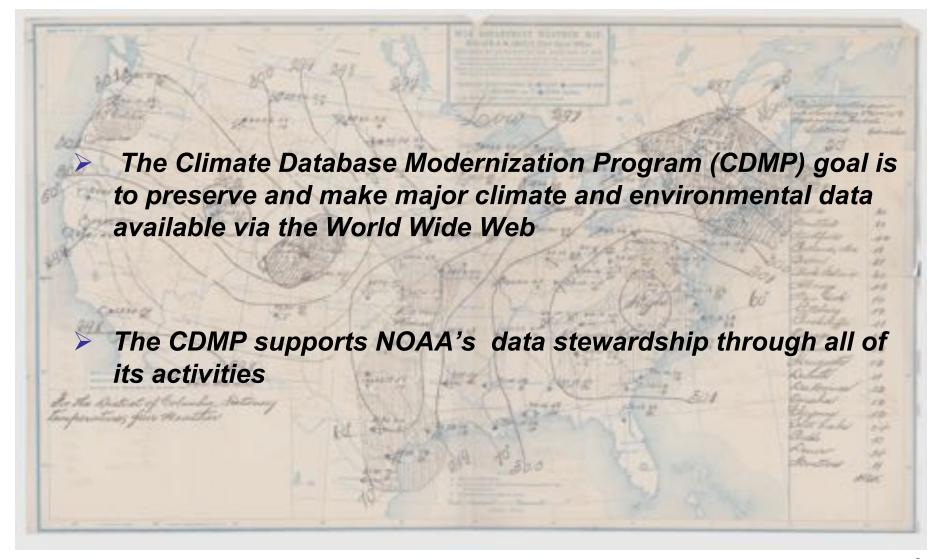








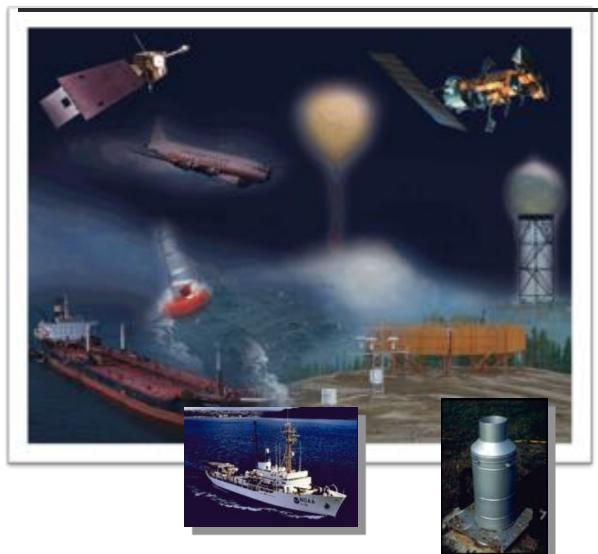






Data Received from Many Sources





Forecast Warning Analysis

Voluntary U.S. Observers

Global Weather Reports

NCEP Weather Charts & Models



Ship, Buoy Reports
Rocketsonde
Weather Balloons
Storm Data

Doppler Radar

(GOES, POES, NPOESS, and many other Satellites

Aircraft Observations

Wind Profiler

Airport Weather Reports (ASOS)

U.S. Climate Reference Network

Climate Models





National Ocean Service - 16

Oceanic and

National Weather Service - 5

NESDIS Geophysical Data Center - 13

Atmospheric Research - 4

National Marine Fisheries Service - 15

Over 14 terabytes of climate data now digitized

- > 54 million weather and environmental images online
- Hundreds of millions of records digitized now online
- International data access and rescue activities
- ▶ 97 current NOAA climate/environmental rescue projects



July 1st, 1842 hourly weather data from Washington, DC, imaged and digitized through the CDMP Program



Imaged Records Example:
Glacial Pairs – Muir Glacier, Alaska

NESDIS National Oceanographic Data Center - 5

Other Government .

Agency Tasks - 4

Climatic Data Center - 35

NESDIS National .





2010 Climate Data Modernization Tasks Across NOAA



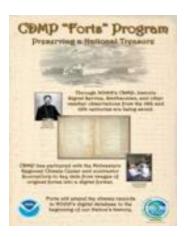




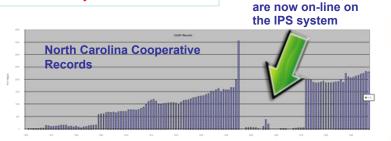
2010 Major NOAA CDMP Data Recovery Tasks

35 NCDC Tasks





Over 410 million surface observations will be added to NCDC's historic synoptic database by the end of 2010





ACAMA Company Continues to Minimum parties of Program (COSM)
Allowed Continues to Minimum to Minimu

•Rescue collections of surface, upper air data, cooperative station data, and historical diaries and journals. Additional projects include "Forts" (mid 19th century) data, NCDC historic publications, various national and international daily and monthly data for inclusion into research datasets (GHCN, IGRA, ICOADS, etc.)

67,022 missing

images 1950-1980







2010 Major NOAA CDMP Data Recovery Tasks

5 NODC Tasks







- Beach Watch Program
- Common Murre Restoration Project
- California Kelp Resources Project
- Scan ~110K survey slides and log sheets

Highlighted tasks include: California marine ecosystem surveys, NODC metadata, World Ocean Database project, NOAA library and film transfer tasks.







2010 Major NOAA CDMP Data Recovery Tasks



Highlighted tasks: Image nautical charts & historical coast pilots, vectorize and geo-reference shoreline charts and photographs, image and key water level gauge and tidal benchmark records, environmental sensitivity maps, water and tide level data, fishery management and catch tracking tasks

Same area on a NOAA nautical chart. Comparing these sources, created at different times, provides information on the rate of change in the coastal zone, which aids in the design of coastal zone mapping projects.





Highlights/Past Successes

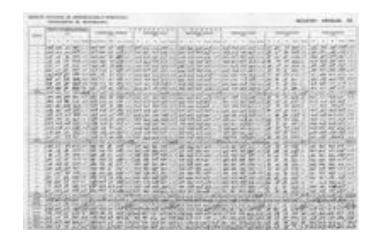
2010 Major NOAA CDMP Data Recovery Tasks

4 OAR/AOML Tasks





The East Indiaman *Warley* (1795), as depicted by Robert Salmon . Courtesy of the National Maritime Museum, London



•Surface observations have been taken at Puerto Baquerizo Moreno (.9°S, 89.6°W) on Isla San Cristobal in Ecuador's Galapagos Islands since 1967. There are no stations within 1,000 km which regularly report surface data for this period. This station is located in the heart of the East Pacific Cold Tongue.

Image and key historic European ship logbooks for incorporation into ICOADS, hurricane reconnaissance information, San Cristobal keying and various research and metadata publications.

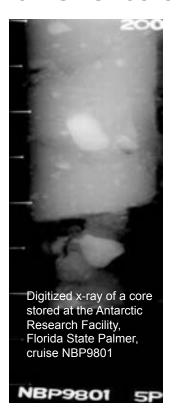






2010 Major NOAA CDMP Data Recovery Tasks

13 NGDC Tasks





•Rescue collections of ionospheric data, various marine geophysical rescue datasets, glacier and tsunami tide data, and solar data rescue projects, cosmic ray and DMSP film data rescue projects. Additional tasks include marine and Lacustrine records of Climate change, glass plate data rescue, solar observations.





Highlights/Past Successes

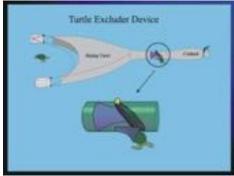
2010 Major NOAA CDMP Data Recovery Tasks

15 NMFS Tasks









Bycatch Reduction Engineering Program to "provide information and outreach"...that will encourage adoption and use of technologies

Sample tasks include: Imaging and keying data on cetaceans, fish eggs/population, coral reefs, plankton data recovery, ecosystem surveys, and Turtle Exclusion data

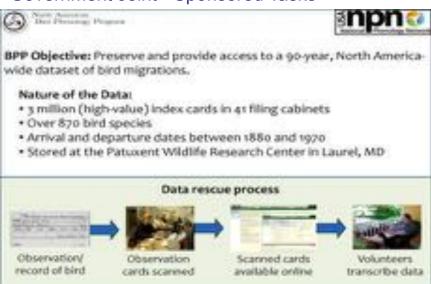


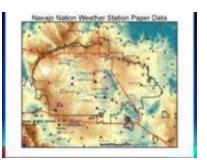


Highlights/Past Successes

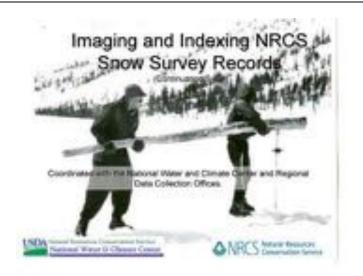
2010 Major NOAA CDMP Data Recovery Tasks

Government Joint - Sponsored Tasks









•Major rescue projects include: USDA Bird Phenology dataset of 1-3 million cards documenting bird migrations; over 250,000 historic snow survey data records from many U.S. western states; keying data from 23 Automated Weather Stations (varying periods from 2-13 years) in the Navajo Nation in Arizona (Hourly and daily) from 1986-1999 - original computer data were lost; Hawaiian island precipitation data project.



NCDC Non-Digital Data Archive



Manuscript / Autograph*

103 Million Pages stored in 120,000 boxes

* Located at Asheville; additional paper records located at the Federal Records Center in Georgia that will be inventoried and prioritized for digitization



Percent digitized (Keyed or imaged)

72% (78 million) 1,500 forms/day

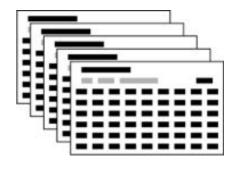


35mm & 16mm Film

125,129 Rolls

Microfiche

860K fiche containing 51 million pages



1.7% (2,105 reels) sporadic

10.0% (86,000 fiche) 3,372 forms/day



Long-term Archive Tasks



Satellite Data

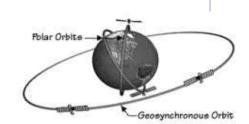
TIROS, NIMBUS, ITOS, ATS, ESSA, GOES, more Mainly 1964-1990

± 1,300 boxes of raw images in many film formats

Film ranges 16mm-70mm

Black & white images, negatives

The many formats makes this a challenging conversion task



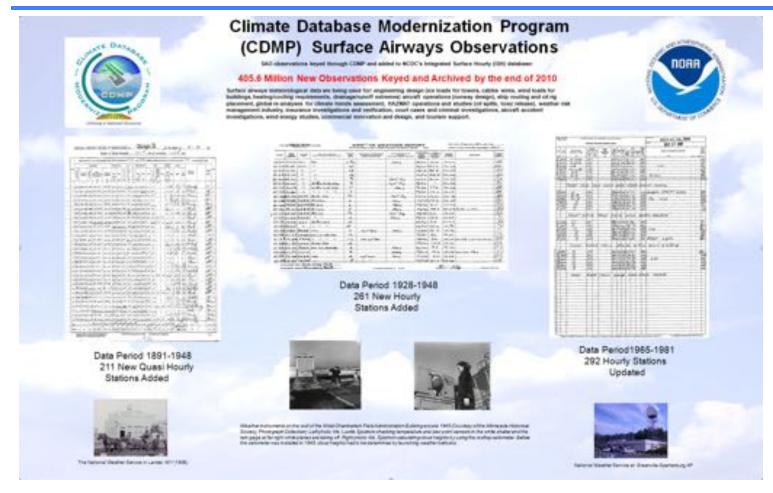
Scientists in RSAD need to access these holdings for use







Data Integration Projects

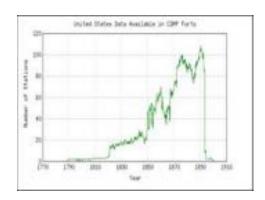


NCDC's digital data for hourly surface observations generally began around 1948. CDMP has imaged over 2 million forms and keyed over 400 million records as part of this keying project. This has added over 50 years of hourly/ synoptic data for over 225 stations.

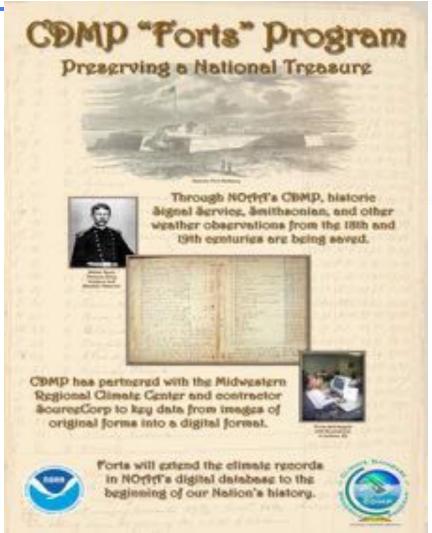




Data Integration Projects



The CDMP "Forts" project is extending daily data records from the beginning of the Weather Bureau era (circa 1892) as far back as the 1780's. Over 390 stations' data have already been keyed. CDMP's Forts team will prepare approximately 75 more stations for keying in the upcoming year.





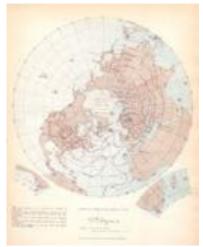
Temperature, precipitation and snowfall tests

Keyed data for all the CDMP Forts stations are available online through the Midwestern Regional Climate Center (MRCC). In addition, close to 200 stations' data have passed Quality Control tests applied at MRCC. Over 175 of these stations are available from NCDC and MRCC via an FTP download.

International Simultaneous Marine ObservationsMulti year Task - Start 2011

- Created by the International Meteorological Congress, Vienna 1873
 - Published by the U.S. National Weather Service from 1875-1887
 - At peak ~ 450 US land stations and ~ 600 vessels reported
 - ~ 30 countries included, from Algeria to the U.S.
 - •~ 5,000,000 daily *land* and marine obs.
- •Maritime Observations POR 1/1877-6/1884
 - Beginning with a few Naval reports to ending with over 600 vessels reporting, average of 75/day
 - $\sim 205,000$ daily observations
- •Elements
 - Locations, pressures, temperatures, winds, clouds, precipitation, sea and weather conditions.
- Caveats
 - Various types of elements, corrections, reductions, scales, etc.



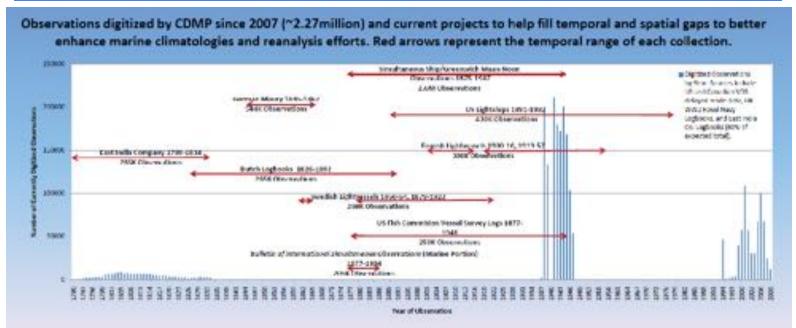




Climate Database Modernization Program Sailing Ahead with Historical Marine Observations



Data Integration Projects







The East Indiaman *Warley* (1795), as depicted by Robert Salmon . Courtesy of the National Maritime Museum, London.



English East India Company Logbooks

- Period of Record: 1790-1834
- 285K Observations
- Digitization goal: 2009



Diamond Shoals Lightship LV 71. US Coast Guard

US Lightship Observations

- Period of Record: 1891-1982
- 430K Observations
- Digitization goal: 2009



Pollock Rip Lightship LV 114, US Coast Guard



R.M.S. Laconia (1912), courtesy of the Steamship Historical Society of America, Inc.

Greenwich Mean Noon / Simultaneous Ship Observations

- Period of Record: 1874-1947
- 2.6 Million Observations
- Digitization goal: 2011



Climate Database Modernization Program (CDMP) Major Historic International Data Rescue Projects







Climate Database Modernization Program Data Integration Projects









58 of 95 Observatories keyed (29 different formats)

-35 coop stations yet to be keyed. Total keying of 13.1 million daily observations (430,000 forms) should be completed by end of 2012 - Data period roughly 1876-1972

- approximately 55% of the data has been keyed. Data will used in international North American Drought monitoring activities with Canada and Mexico and incorporated into Global Historical Climatological

Network (GHCN) at NCDC





Another source of data 1903-1978 Mexican Daily Weather Maps- Daily pressure data in tables Daily data keyed for 30 covering 1890's -1920's -6092 station months. Data incorporated into GHCN

stations total of will be

Hourly data for 16 stations in the 1950-1954 period. Data will be incorporated into the (GHCN) and Integrated Surface Hourly (ISH) database Network at NCDC

http://docs.lib.noaa.gov/rescue/ Foreign_climate/FCD_006_025.pdf/ Mexico carta del tiempo/





Data Integration Projects



Kenya – 5 sites



Senegal – 3 sites



Malawi - 7 sites

- Over 150,000 images of pibal (upper air wind) records from the 1940's to 2003 received to date from 7 African countries (Kenya, Malawi, Mozambique, Niger, Senegal, Tanzania and Zambia)
- Most data have been keyed and after passing quality control checks will be integrated into NOAA's Integrated Global Radiosonde Archive Database (IGRA)
- Digital data files were provided to the host countries that imaged the data.
 Keyed data files also hyperlinked to the actual images providing an easy access to the original record
- Ongoing pibal tasks with Tanzania and Mozambique may expand to surface data in 2011
- NWS and IEDRO (non-profit) provides some technical, administrative and logistical support
- Possible Namibia and ACMAD Microfiche Recovery Project



Mozambique – 2 sites



Tanzania – 2 sites



Zambia – 2 sites



Niger – 2 sites



Making Data User Friendly and Accessible



The index on the CD-ROM will be the main point for searching for related data.

The index page provides links to all available data types (both images of recorded data and actual keyed data) associated with the CD-ROM



Station Number	Date	Time	Table	Dir/Spd	Raw Data	Image
67423	11/07/1970	0400	Table	Dir/Spd	Raw Data	Image
67423	12/21/1979	0400	Table	Dir/Spd	Raw Data	Image
67423	12/22/1979	0400	Table	Dir/Spd	Raw Data	Image
67423	12/24/1979	0400	Table	Dir/Spd	Raw Data	Image
67423	12/27/1979	0400	Table	Dir/Spd	Raw Data	Image
67423	12/29/1979	0400	Table	Dir/Spd	Raw Data	Image
67423	12/30/1979	0400	Table	Dir/Spd	Raw Data	Image

	Station:67423 Day/Mo/Year Time:07/11/1970 04Z										
•	Keyed Image ID: CDMP06MA\MA0001\0009.jpg										
•	Sourc	ce Image	ID: .\MA0	001\1\F	261600)06.jpg					
•											
•			CO	MPUTE	D	OBSEF	RVED				
•	TIME	HGHT	Azimuth	Elev	DIR	SPD	DIR S	PD			
	MINS	S FEE1	DEG	DEG	DEG	KTS	DEG	KTS			
•	> 0	0					130.0	5.0			
•	1	500	327.2	21.2	147	13	0	0			
•	2	1000	337.9	22.4	170	12	0	0			
•	3	1500	239.7	24.9	165	8	0	0			
•	4	2000	342.7	26.1	174	9	0	0			
•	5	2500	337 4	25.5	140	12	0	0			
,	Note:		the Begin		each I	Record	in the Lau	unch			



The highlighted .jpg has a link connecting it to the original image from which the data was keyed

-		C 16.3			HERE TA	100 ft./s		Tat !	- Was	Mar a		
	Promo Commanda		The	1767	MATURE	VYLOCI	VYLOCITY, Rause		SERVICE AND			
min.	Bearing.	Alternation	A/8	Ann.	1 House		3-44	Characters.			cos	
			N 30	-					- Valuette	Fars.A.	Part 1	
2.			10/ 30		120	1		275				
3.11			187.30			1		25%				
			30/ 10			1						
2			28/ 50		1 1 1 1	1		TAMID)				
-		2002	21/ 40		22.0	-	15150	777			100	
		2200	357.79			-	2.15	151/14				
-		222	40/40			100						
- 1		230	45/ 44			170	715					
10		250	30/190			- 2	3.5		12-			
		200	85/188				25		-			
- 12			AR/128		2164	-	150	123	66			
- 13			43,130				-					
10			707/140				-					
			75/150				_					



The data above is an upper air pibal, due to the limited wind data available a keying format was developed to key the azimuth/elevation data along with the surface wind data, found in the header section. The wind data, to the middle right of the image, we found, was often determined by the average of multiple levels of wind data.



The Foreign Data Library (FDL)



The best index we have for our "vast unknown" of foreign data:

Table 16
Detailed Listing of Unfilmed Foreign Data Publications Archived in ASS for Published Dates
Prior to Mid-1975 and Costs for Producing Archival Quality Film from These Publications

Unprocessed/Noncatalogued	Varies	750,000	83.3	37,500	22,500
Polar Regions	1873-1972	154,400	21.6	9,720	5,833
Polynesia/Micronesia	1911-1975	223,500	24.8	11.175	6,705
New Guinea	1914-1964	5,700	0.6	285	171
Australia	1770-1974	351,000	39.0	17,550	10,530
New Zealand	1853-1975	400,200	44.5	20,010	12,006
General Australiasia	1900-1969	7,200	0.8	360	216
Sunda Islands	1866-1974	117,600	13.1	5,880	3,528
Philippines	1348-1975	924,900	102.8	46,245	27,747
Germany	1750-1975	1,382,700	153.6	69,135	41,48
British lules	1845-1975	939,300	101.0	46,965	28,175
Europe Alsa	1797-1973	159,600	17.7	7,980	4,781
Oceans & Water Bodies	1648-1974	563,400	62.6	28,170	16,900
Henisphere	1743-1974	172,800	19.2	8,640	5,18
Tropic	1937-1970	14,400	1.6	720	433
World	1841-1973	219,000	24.3	20,950	6,570
Main Geographic Area Covered by Publications	Period	Pages	Young	BSHR& PC	E
				1100	



2071 boxes of data in the FDL

CDMP multi-year project to inventory data suitable for keying in GHCN daily and monthly, IGRA upper air and ISD surface

Web interface inventory will be developed after inventory is complete





http://docs.lib.noaa.gov/rescue/data_rescue_home.html

Links to climatologica data arranged by Country or Region* Africa Angola ameroon Dibouti dypt Guinea-Bissau Kenya Byde Madagascar Mozambique Namibia South Africa Tunisia Jganda Afghanistan india Pakistan Thailand Turkey Europe Sermany Hungary Norway Poland

Russia

Cuba

rugoslavia

Costa Rica

North America

Dominican Republic

Time Period, Frequency, and Parameters

The time period of coverage ranges from the 1830s through the 1970s with most data from the period prior to 1960. Each series typically includes observations for a number of meteorological and other geophysical parameters. These may include daily and monthly observations, as well as monthly and annual means, for

- · surface temperature
- · precipitation
- · atmospheric pressure
- · wind speed and direction
- · soil temperature
- · radiation
- · sunshine
- · cloudiness
- · upper air measurements
- · geomagnetism

How to Find the Data

To locate the desired data set, select a country or region from the list on the left, then select a title and year.

Credit

Funding for the digitization of these publications was provided by the Climate Data Modernization Project, National Climatic Data Center.

Please credit the NOAA Central Library when using these data. Share the results of your research with any of the staff members below.

Contact Information

you have questions or comments concerning the images, or you need further assistance, please contact

<u>Library Reference@noae.gov</u>, 301-713-2600 ext 124, or

<u>Diana L. Abney@noae.gov</u>, 301-713-2600 ext. 121, or

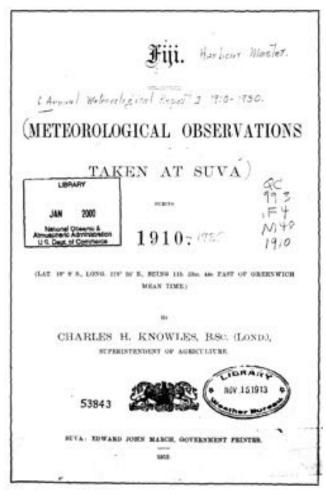
Stanley Elswick@noae.gov, 301-713-2600 ext. 138.

The images are provided in PDF and multi-page TIFF formats.



CDMP-Data Rescue Project NOAA Central Library











CDMP Project Partners











Resources Solutions Results."