Information about NCAR Data Support (DSS), 1981 – 87

Roy Jenne
8 Nov 2000

This set of papers about DSS has 4 sets of slides and one short paper.

1. Data Support Section, Jan 1981, 12 slides

2. Long-Term Climate Data Sets at NCAR, May 1982, 34 slides

3. Data Support Section, 1 Sep 1982, 7 pages

4. Data Support Section, March 1987, 31 slides
   › We had 237 datasets then, slide on data centers in USA. Slide on costs of computing, etc.

5. Data Support Section, Dec 1987, 38 slides
   › For a user conference. Formats, avoid traps, etc.
DATA SUPPORT SECTION

Data status and plans
TN-111, WMO data text, National and WMO plans

Data set plans
- UA
- Surface
- Hydrological
- Ocean
- Satellite
- Paleoclimatic
- Geographical and Economic

Data volume
- At NCAR
- ETAC Asheville
- Navy
- Satellite data
  NOAA, NOSS Satellite

TBM data activity

Archive costs

Roy Jenne
Jan 1981
UPPER AIR DATA

- Continue the sets of NMC basic data and analyses
- Prepare major subsets

- Obtain FGGE data

- Obtain national data back to 1950
  India—Brazil—USSR—etc.

- Make available the 1940-60 data at USAF—Asheville
SURFACE DATA

- Continue to obtain NMC and Navy data on tape
- Obtain more NMC data in real time as the need develops
- Make the global 1900-60 USAF-ETAC data available
- NCAR will soon receive the US daily co-op data from NCC
- Improve the data set of year-month surface data
- Obtain selected additional daily time series of precipitation and temperature--global
Hydrological Data

- Obtain selected river flow data
- Encourage the preparation of water use, data
- Obtain better global precipitation data
  - Long series at 300 to 70 km spacing
  - Don't get the 5-10 km data
  - Get selected compact radar data
  - Get compact satellite data
Ocean Data

- NCAR now has permanent ship surface data
- Obtain a good set of ship data
  - NCC-UK-Navy-USSR
- Obtain the set of BT - XBT data
- Obtain sets of drifting buoy data
- Get a better set of ocean bathymetry data
- Obtain selected ocean satellite data
- Don't get Navy spectral wave grids
SATELLITE DATA

- Continue work to define intermediate sized data sets

- Obtain selected data, especially for clouds—ocean studies—convective precipitation

- Plan and prepare for Ocean Data User Support—obtain selected Seasat data
PALEOClimatic Data

- Work with NSF to prepare TN (inventory and status)
- Obtain data to support modeling

Geographical and Economic Data

- Identify selected available data
  Soil types
  Energy use
  Crop acreage
### Selected Large DSS Data Sets

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Tapes</th>
<th>Bits (10^9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMC anal. 1963-72 (n. hemisphere)</td>
<td>27</td>
<td>6.8</td>
</tr>
<tr>
<td>NMC Tropical anal. 1968-72</td>
<td>7</td>
<td>1.75</td>
</tr>
<tr>
<td>NMC Global and 1976-79</td>
<td>95</td>
<td>25.</td>
</tr>
<tr>
<td>NMC LFM grids (1971-79)</td>
<td>80</td>
<td>21.6</td>
</tr>
<tr>
<td>Australian Anal. 1972-79</td>
<td>19</td>
<td>5.3</td>
</tr>
<tr>
<td>Other sets</td>
<td>50</td>
<td>12.</td>
</tr>
</tbody>
</table>

### Observations

<table>
<thead>
<tr>
<th>Observations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NMC UA obs. 1962-79</td>
<td>259</td>
<td>60.</td>
</tr>
<tr>
<td>Time series UA (USA, UK, AUSTR., Etc.)</td>
<td>90</td>
<td>23.</td>
</tr>
<tr>
<td>NMC Sfc obs 1976-79</td>
<td>182</td>
<td>40.</td>
</tr>
</tbody>
</table>

### Aircraft and Satellite

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GATE Aircraft</td>
<td>72</td>
<td>16.</td>
</tr>
<tr>
<td>GOES satl for GATE area</td>
<td>83</td>
<td>20.8</td>
</tr>
</tbody>
</table>

**Total ~ 1164 283.**

### Getting Soon:

- **Daily US surface (was 650 tapes)**: 170, 51
- **Foreign synop (now 1414)**: 180, 68
- **Ship obs. (now 550)**: 170, 62
ETAC--Asheville

<table>
<thead>
<tr>
<th>Observations</th>
<th>(10^9 bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface 1965-79</td>
<td>324</td>
</tr>
<tr>
<td>Foreign SFC. 1901-71</td>
<td>52</td>
</tr>
<tr>
<td>SFC. Airways 1931-74</td>
<td>79</td>
</tr>
<tr>
<td>UA 1968-79</td>
<td>530</td>
</tr>
<tr>
<td>Foreign UA 1943-67</td>
<td>12</td>
</tr>
<tr>
<td>Foreign Winds 1930-71</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyses</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyses 1971-79</td>
<td>120</td>
</tr>
<tr>
<td>3-D Neph 1971-79</td>
<td>1441</td>
</tr>
</tbody>
</table>

NAVY

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>62</td>
</tr>
</tbody>
</table>
## Satellite Data

<table>
<thead>
<tr>
<th></th>
<th>Now</th>
<th>All Data</th>
<th>Select</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NOAA SR Grids (20 km) (1973-78)</strong></td>
<td>2600</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td><strong>NOAA VTPR (Nov 72-Jan 79)</strong></td>
<td>1130</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td><strong>TIROS 4 km (Oct 78-On)</strong></td>
<td>5800/yr</td>
<td></td>
<td>95/yr</td>
</tr>
<tr>
<td><strong>Each GOES</strong></td>
<td></td>
<td>9530/yr</td>
<td>88/yr</td>
</tr>
</tbody>
</table>

### NOSS (Only Basic Data)

<table>
<thead>
<tr>
<th></th>
<th>Tapes Per Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Altimeter</strong></td>
<td>638</td>
</tr>
<tr>
<td><strong>SCAT</strong></td>
<td>924</td>
</tr>
<tr>
<td><strong>LAMRR (SMRR)</strong></td>
<td>2200</td>
</tr>
<tr>
<td><strong>CZCS</strong></td>
<td>35406</td>
</tr>
</tbody>
</table>

- Tapes per Yr
  - 70 (4 km)
  - 150 (48 km)
  - 350
  - 180 (16 km)
<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBM ARCHIVE</td>
<td>$10^9$ BITS</td>
</tr>
<tr>
<td>TOTAL ARCHIVE (TIMES 2)</td>
<td>4000</td>
</tr>
<tr>
<td>USER BITS MOVED/MO</td>
<td>1139</td>
</tr>
<tr>
<td>TOTAL BITS MOVED/MO</td>
<td>5234</td>
</tr>
<tr>
<td>ONE MEGABIT RATE</td>
<td>2635/MO</td>
</tr>
</tbody>
</table>
Table 1. Data From Geostation Satellites. Note that saving all data from one GOES satellite would take 9149 tapes/year (6250 BPI) for visible data but only 23.8 tapes when the data is sampled each 8 Km.

<table>
<thead>
<tr>
<th></th>
<th>ESA Meteosat</th>
<th>GMS Japan</th>
<th>GOES USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visible resolution</td>
<td>2.5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>IR resolution</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Water vapor resolution</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vis samples in scan</td>
<td>E 6116</td>
<td>E 7645</td>
<td>15290</td>
</tr>
<tr>
<td>Vis scan lines in pix</td>
<td>E 5828</td>
<td>E 7284</td>
<td>14568</td>
</tr>
<tr>
<td>IR spots in scan</td>
<td>E 3058</td>
<td>E 2548</td>
<td>3823</td>
</tr>
<tr>
<td>IR scan lines</td>
<td>E 2914</td>
<td>E 2428</td>
<td>1821</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2.96 x 10^8</th>
<th>4.62 x 10^8</th>
<th>1.85 x 10^9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis Bits/picture (8 bit spot)</td>
<td>7.41 x 10^7</td>
<td>5.15 x 10^7</td>
<td>5.79 x 10^7</td>
</tr>
<tr>
<td>IR Bits/picture (8 bit spot)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data from Satellite

<table>
<thead>
<tr>
<th></th>
<th>14.2 (48p)</th>
<th>3.70 (8p)*</th>
<th>89.0 (48p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis Bits/day (8 bit) (10^9)</td>
<td>3.56(48p)</td>
<td>.721 (14p)</td>
<td>2.78(48p)</td>
</tr>
<tr>
<td>IR Bits/day (8 bit) (10^9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cut vis at dark line, drop IR overlap

<table>
<thead>
<tr>
<th></th>
<th>5.33 (48p)</th>
<th>1.39 (8p)</th>
<th>33.4 (48p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis Bits/day (6 bit) (10^9)</td>
<td>3.56 (48p)</td>
<td>.721 (14p)</td>
<td>1.39 (48p)</td>
</tr>
<tr>
<td>IR Bits/day (8 bit) (10^9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Cut most space data (25%)

<table>
<thead>
<tr>
<th></th>
<th>1460 (48p)</th>
<th>383 (8p)</th>
<th>9149 (48p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis Bits/year (6 bit) (10^9)</td>
<td>957 (48p)</td>
<td>199 (14p)</td>
<td>381 (48p)</td>
</tr>
<tr>
<td>IR Bit/year (8 bit) (10^9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>15.2 (8p)</th>
<th>15.3 (8p)</th>
<th>23.8 (8p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vis Bits/year (6 bit) (10^9)</td>
<td>39.8 (8p)</td>
<td>28.0 (8p)</td>
<td>63.5 (8p)</td>
</tr>
<tr>
<td>IR Bits/year (8 bit) (10^9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: 7000 tapes from GMS

*(8p) means 8 pictures/day

Note: A 1600 BPI tape holds .3 x 10^9 bits
A 6250 BPI tape holds 1.0 x 10^9 bits
A U. Wisc. recorder tape holds 22.3 x 10^9 bits
LONG TERM CLIMATE DATA SETS
AT
NCAR

FUNCTIONS NECESSARY FOR DATA ANALYSIS AND DISPLAY
SELECTED CLIMATE STUDIES NEEDING DATA
MAJOR PROJECTS TO PREPARE DATA
SELECTED DATA AT NCAR
DAILY ANALYSES-GRIDS
UPPER AIR OBSERVATIONS
SATELLITE DATA
SURFACE DATA
OCEAN DATA
ELEVATION AND DEPTH DATA
PALEO-CLIMATIC INFORMATION
DATA INFORMATION SOURCES

Roy Jenne
May 1982
FUNCTIONS NECESSARY FOR DATA ANALYSIS AND DISPLAY

1. OBTAIN AND PREPARE DATA SETS

2. MOVE DATA SETS WITHIN THE HARDWARE SYSTEM

3. MAINTAIN GENERAL AIDS FOR UNBLOCKING AND UNPACKING DATA

4. PREPARE ACCESS Routines FOR EACH DATA SET

5. PREPARE AIDS FOR CALCULATIONS

6. MAKE CALCULATIONS

7. DISPLAY THE OUTPUT

8. PROVIDE DATA INVENTORY INFORMATION
SELECTED CLIMATE STUDIES NEEDING DATA

1. CIRCULATION STATISTICS
2. ATMOSPHERIC TEMPERATURE TRENDS
3. SURFACE TEMPERATURE TRENDS
4. SURFACE AND UPPER AIR CLIMATOLOGY
5. LARGE SCALE CLIMATIC RELATIONSHIPS
6. RAINFALL COMPARED TO OCEAN CONDITIONS
7. PACIFIC EQUATORIAL STUDIES
8. TROPICAL STORMS
9. BLOCKING PATTERNS/CLIMATE PREDICTION
10. OCEAN STUDIES
11. HYDROLOGY
12. PREDICTABILITY
13. PALEOClimATIC STUDIES
14. LOCAL EFFECTS MODELS
15. SOLAR AND VOLCANIC EFFECTS
MAJOR PROJECTS TO PREPARE DATA

1. SHIP MARINE DATA

   A WORKING DATA SET 1850 - 1980 IN 12 MONTHS

   - WITH 71 MILLION OBSERVATIONS
   - GUESS 5 MILLION STILL OUT
   - 16 MILLION OBSERVATIONS NEVER PUNCHED

2. SATELLITE DATA

   INTERNATIONAL SATELLITE CLOUD PROJECT -- DATA FOR CLOUDS

   NOAA SOUNDER DATA 1972 - PRESENT

   NOAA IR/VIS SCANNERS 1972 - PRESENT

3. MORE LONG PERIOD SURFACE BASED DATA

   DAILY DATA FOR USA (HAVE THIS)

   MORE MONTHLY DATA (DSS AND SHEA)

      ------------------------------------

0 DO YOU NEED A TYPE OF DATA THAT WE DO NOT HAVE?
### SELECTED DATA AT NCAR

#### ANALYSES

<table>
<thead>
<tr>
<th>Period</th>
<th>Description</th>
<th>Grids</th>
<th>Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1899-81</td>
<td>DAILY SFC, N. H.</td>
<td>65,000</td>
<td>4</td>
</tr>
<tr>
<td>1962-81</td>
<td>SFC TO 10 MB, N. H.</td>
<td>920,000</td>
<td>88</td>
</tr>
<tr>
<td>1971-81</td>
<td>FINE MESH, N. AMER.</td>
<td>1,500,000</td>
<td>80</td>
</tr>
<tr>
<td>1972-81</td>
<td>DAILY S. HEM. AUSTRALIA</td>
<td>225,000</td>
<td>20</td>
</tr>
<tr>
<td>1976-81</td>
<td>GLOBAL</td>
<td>290,000</td>
<td>120</td>
</tr>
</tbody>
</table>

#### PRECIPITATION

<table>
<thead>
<tr>
<th>Description</th>
<th>Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MONTHLY FOR 1800 GLOBAL STNS, LONG RECORDS</td>
<td>4</td>
</tr>
<tr>
<td>MONTHLY FOR 1000 AFRICAN STNS, LONG RECORDS</td>
<td>1</td>
</tr>
<tr>
<td>DAILY FROM 9000 U. S. STNS, 40 YEARS</td>
<td>55</td>
</tr>
<tr>
<td>HOURLY FROM 3000 U. S. STNS, 30 YEARS</td>
<td>40</td>
</tr>
</tbody>
</table>

#### RAWINSONDES

<table>
<thead>
<tr>
<th>Description</th>
<th>Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAILY FROM 1948 FOR U. S. STNS (4,200,000 OBS)</td>
<td>49</td>
</tr>
<tr>
<td>FROM U. K., AUSTRALIA, N. Z., FRANCE, INDIA</td>
<td>12</td>
</tr>
</tbody>
</table>

#### OBSERVATIONS FROM TELECOMMUNICATIONS

<table>
<thead>
<tr>
<th>Description</th>
<th>Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAILY UPPER AIR OBS FROM 1962</td>
<td>179</td>
</tr>
<tr>
<td>DAILY SURFACE OBS 1972-78</td>
<td>98</td>
</tr>
</tbody>
</table>

#### OTHER

<table>
<thead>
<tr>
<th>Description</th>
<th>Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATELLITE HEAT BUDGET AVERAGES</td>
<td>1</td>
</tr>
<tr>
<td>GLOBAL CLIMATOLOGY TO 100 MB (10 FOR N. H.)</td>
<td>3</td>
</tr>
<tr>
<td>LAND ELEVATION, OCEAN DEPTH DATA</td>
<td>5</td>
</tr>
<tr>
<td>MANY OTHER DATA SETS</td>
<td></td>
</tr>
</tbody>
</table>

**Roy L. Jenne**

**September 1981**
### Daily Analyses

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>US</strong> daily</td>
<td>sfc-500 mb</td>
<td>sfc-300 mb</td>
<td>N. Hem</td>
</tr>
</tbody>
</table>

#### N. Hem
- Daily N. Hem (USSR) 500 mb
- Daily N. Hem (USN) sfc-100 mb
- Daily N. Hem (German) Mar 66 sfc-100 mb
- Daily Strato H,T (German) 50-10 mb 50-30 mb
- Mo heights (German) 100 - 10 mb 50-30 mb
- Weekly analyses (US and German) 2 and .4 mb add 5 mb 1968 1973

#### Global
- Daily global (NMC) Dec 72
- Monthly global (GFDL) 1958 1973
- Daily SLP and SST (USN) Aug 74 Poor Improved Jun 74
- Daily SST (NESS) Sep 73
- Mo SST (from ship data - several sources, for many years, still problems)

#### Tropical
- Daily Tropical (NMC) 700-200 mb Jan 68 sfc-200 mb Jan 73
- Daily Tropical (USN)

#### S. Hemisphere
- Daily S. Hem (Australia) 500 mb sfc-100 mb 1968 Apr 72 sfc-50 mb Aug 77 sfc-100 mb Jul 73
- Daily S. Hem (S. Africa)
- Daily S. Hem (USN)

#### N. Hem SST
- Daily (USN) Sea Surface Temperature 1961

**Note:** This is not a complete listing of available analyses

**Figure 6** Daily Analyses. Most sets of analyses have both strong and weak points that must be further determined and considered when using them.
Fig. 1-1. Northern hemisphere analyses of sea level pressures, heights, and temperatures. The daily sea level pressure analyses start in 1899. The southern limit for most grids is about 15 or 20°N. The resolution of the more recent grids is about 400 km and the grids are usually two per day. Hashed lines show time periods when analyses are not available for every day of every month. Some levels, such as 250, 150, and 70 mb, are available but are not listed in the figure.
Fig. 12-1. Shows the year-month grids and the earlier long-term mean grids that were available for use in this climatological study. The indicated data are year-month grids except where noted. Dashed lines show year-month grids that were more recently acquired. The first 18 months of data in the 50 mb height mean were from Muench (from Jenne et al., in preparation).
AUSTRALIAN GRIDS

APRIL 72 | NOV 76 | AUG 78 | JULY 1980

- 45 TAPES RECEIVED
- 19 TAPES IN ARCHIVE
  - FGGE HAS JAN 78 - NOV 79 ON 23 TAPES
  - WE HAVE IT ON 5 TAPES

PROCESSING PROGRAMS
- CHECK DATA
- CONVERT FORMAT
- SORT IT
- ELIM DUPE GRIDS
- INVENTORY
- MAKE MONTH MEANS

USER SOFTWARE
- READ AND UNPACK TAPE
- DERIVE DESIRED LAT-LON GRID
- CONTOUR ALL OR PART OF GRID

USERS

47 SEPARATE REQUESTS - 32 USERS

11 IN 1978
13 IN 1979
21 IN 1980
2 IN 1981

USERS FROM
- NCAR
- UNIVERSITIES (10)
- UK
- BRAZIL
- TAIWAN
Upper Air Observations

UA observed data from GTS (rawinsonde, etc.)

<table>
<thead>
<tr>
<th>Year</th>
<th>N Hem</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

US - NMC
- Mar 62
- Jun 66

Germany
- Aug 66

Australia

UA observed data, not GTS
- From MIT: 1958-1963
- Also see Fig. 2

Special Aircraft Data
- Winds: 1960-1972
- Tropics

Air Quality

AIDS

ASDAR

Balloons
- S. Hem. TWERLE (winds, H): Jul 75
- S. Hem. EOLE (winds): Aug 71-Dec 72
- FGGE Tropical (winds): Jan 79

Rockets
- 6 stations needs work: 1968
- 33 stns

Satellite
- Cloud winds
- I R sounders
- TIROS-N (IR and microwave)

Figure 4 Upper Air Observations

10
Fig. 1

Typical data coverage and types from the FGGE composite observing system for a 6-hour period during SOP-I (see text).
Rawinsonde Data
Not Collected in Real Time

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>USA</strong> Americas</td>
<td>Prepared</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Includes Alaska, Pacific Islands, Caribbean, Mexico, Central America.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chile, <strong>not</strong> Argentina <strong>or</strong> Brazil.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>Prepared</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 stns UK, 3 Atlantic ships</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1948</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19 stns overseas, completed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45 stns including islands</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1943-1955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13 stns including islands</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 stns completed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1951</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USSR</td>
<td>Prepared</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Many other countries have similar data sets.

Figure 5 Rawinsonde data

1982
Roy Jenne
UPPER AIR DATA

- Continue the sets of NMC basic data and analyses
- Prepare major subsets

- Obtain FGGE data

- Obtain national data back to 1950
  India—Brazil—USSR—etc.

- Make available the 1940-60 data at USAF—Asheville
Figure 2. Satellite digital data. Data are also available from synchronous satellites.
Figure 3. Selected sets of older climate data for global and regional problems. Some of these data will be in global archives. Other data will remain in regional archives. Global data for a given data type will therefore be obtained by contacting one to several data archives.
Fig. 1 - 15 (1) - Spatial correlation of precipitation totals for various time intervals (Vilnius, USSR, summer periods) from data supplied by M. V. Golchin and others (1967).

1 = Total for 24 hr period
2 = Total for 24 hr period
3 = Total for 10 day period
4 = Total for month
5 = Total for season

Fig. 1 - 15 (2) - Spatial correlation of precipitation totals for various time intervals (Illinois, USA) from data by F. A. Bee and J. W. Skipp (1969).

1 = Total for rain
2 = Total for month
3 = Total for season
Figure 5. A summary of primary surface data archives for the World Climate Program, prepared in delayed time. About eight times as many (8000) surface synoptic stations will be available in real time. The insert shows the station spacing and the number of stations for the world and for the contiguous states of the USA. An outline of portions of Europe is on the same map.
Table 1. Daily surface synoptic data at Asheville, prepared by the USAF. Many stations do not cover the entire stated period of record (POR). The number of station-decades of data is included to indicate data coverage. The file is global. A sample of the data is listed below. The data are on 9 track, 1600 BPI tapes, about 75,000 observations per tape at Asheville. The data are in time series order on the tapes. (From Jenne, 1975).

<table>
<thead>
<tr>
<th>Area</th>
<th># stns</th>
<th>POR</th>
<th># YR-MO</th>
<th>Sta-Decades</th>
<th># obs (Thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific Ocean Islands</td>
<td>180</td>
<td>39-68</td>
<td>18,144</td>
<td>151</td>
<td>3,576</td>
</tr>
<tr>
<td>S.E. Asia</td>
<td>352</td>
<td>43-66</td>
<td>39,988</td>
<td>333</td>
<td>6,600</td>
</tr>
<tr>
<td>Indonesia</td>
<td>80</td>
<td>42-66</td>
<td>10,740</td>
<td>134</td>
<td>712</td>
</tr>
<tr>
<td>Central and S. America</td>
<td>739</td>
<td>30-71</td>
<td>65,741</td>
<td>548</td>
<td>4,308</td>
</tr>
<tr>
<td>Philippines, Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formosa, Japan, New Zealand</td>
<td>305</td>
<td>45-69</td>
<td>39,958</td>
<td>333</td>
<td>11,168</td>
</tr>
<tr>
<td>Europe, Br. Isles</td>
<td>1600</td>
<td>26-72</td>
<td>127,960</td>
<td>1066</td>
<td>18,464</td>
</tr>
<tr>
<td>Middle East</td>
<td>602</td>
<td>35-68</td>
<td>47,149</td>
<td>393</td>
<td>4,155</td>
</tr>
<tr>
<td>Africa</td>
<td>757</td>
<td>41-68</td>
<td>63,431</td>
<td>529</td>
<td>5,015</td>
</tr>
</tbody>
</table>
AFRICA

MAP 4.1(b)

DENSITY OF PRECIPITATION STATIONS

<table>
<thead>
<tr>
<th>Category</th>
<th>km²/station</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 100</td>
<td></td>
</tr>
<tr>
<td>101 - 600</td>
<td></td>
</tr>
<tr>
<td>601 - 1500</td>
<td></td>
</tr>
<tr>
<td>1501 - 5000</td>
<td></td>
</tr>
<tr>
<td>&gt; 5000</td>
<td></td>
</tr>
<tr>
<td>Data not received</td>
<td></td>
</tr>
</tbody>
</table>
Table 10-1

Monthly surface climatological station data. Number of available monthly averages of sea level pressure, station pressure, temperature, and precipitation for a global network of surface stations. Data are included through 1980. The explanation of the symbols follows:

- RECS = No. of logical records in the decade (station/years)
- S = No. of stations in the decade
- SSLP = No. of stations in the decade with sea level pressures
- MLSP = No. of months of SLP in the decade
- SP = No. of stations in the decade with pressures
- MP = No. of months of pressure in the decade
- ST = No. of stations in the decade with temperatures
- MT = No. of months of temperature
- SR = No. of stations in the decade with precipitation
- MR = No. of months of precipitation

<table>
<thead>
<tr>
<th>DECADE</th>
<th>RECS</th>
<th>S</th>
<th>SSLP</th>
<th>MLSP</th>
<th>SP</th>
<th>MP</th>
<th>ST</th>
<th>TEMP</th>
<th>MT</th>
<th>Rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971-1980</td>
<td>14358</td>
<td>1928</td>
<td>1917</td>
<td>139700</td>
<td>1443</td>
<td>140043</td>
<td>1920</td>
<td>149140</td>
<td>1905</td>
<td>146630</td>
</tr>
<tr>
<td>1961-1970</td>
<td>14107</td>
<td>1728</td>
<td>1648</td>
<td>144334</td>
<td>144</td>
<td>1722</td>
<td>152848</td>
<td>1701</td>
<td>149046</td>
<td></td>
</tr>
<tr>
<td>1951-1960</td>
<td>18046</td>
<td>1877</td>
<td>1311</td>
<td>145307</td>
<td>1140</td>
<td>123816</td>
<td>1807</td>
<td>201484</td>
<td>1827</td>
<td>207835</td>
</tr>
<tr>
<td>1941-1950</td>
<td>8032</td>
<td>902</td>
<td>583</td>
<td>584221</td>
<td>576</td>
<td>59409</td>
<td>859</td>
<td>88637</td>
<td>882</td>
<td>93003</td>
</tr>
<tr>
<td>1931-1940</td>
<td>6664</td>
<td>732</td>
<td>309</td>
<td>31517</td>
<td>429</td>
<td>46055</td>
<td>640</td>
<td>69633</td>
<td>725</td>
<td>80312</td>
</tr>
<tr>
<td>1921-1930</td>
<td>612</td>
<td>648</td>
<td>222</td>
<td>25375</td>
<td>375</td>
<td>42347</td>
<td>538</td>
<td>60151</td>
<td>640</td>
<td>72225</td>
</tr>
<tr>
<td>1911-1920</td>
<td>5591</td>
<td>585</td>
<td>134</td>
<td>15351</td>
<td>341</td>
<td>38482</td>
<td>476</td>
<td>53554</td>
<td>551</td>
<td>62341</td>
</tr>
<tr>
<td>1901-1910</td>
<td>4965</td>
<td>538</td>
<td>120</td>
<td>12707</td>
<td>322</td>
<td>35345</td>
<td>437</td>
<td>47075</td>
<td>502</td>
<td>54827</td>
</tr>
<tr>
<td>1891-1900</td>
<td>3948</td>
<td>439</td>
<td>86</td>
<td>9093</td>
<td>263</td>
<td>28438</td>
<td>348</td>
<td>37026</td>
<td>407</td>
<td>43136</td>
</tr>
<tr>
<td>1881-1890</td>
<td>3040</td>
<td>342</td>
<td>58</td>
<td>5828</td>
<td>205</td>
<td>20709</td>
<td>274</td>
<td>28135</td>
<td>313</td>
<td>33241</td>
</tr>
<tr>
<td>1871-1880</td>
<td>1949</td>
<td>245</td>
<td>40</td>
<td>3620</td>
<td>122</td>
<td>10194</td>
<td>173</td>
<td>15176</td>
<td>228</td>
<td>21429</td>
</tr>
<tr>
<td>1861-1870</td>
<td>1058</td>
<td>138</td>
<td>17</td>
<td>1713</td>
<td>36</td>
<td>3289</td>
<td>75</td>
<td>7029</td>
<td>123</td>
<td>11089</td>
</tr>
<tr>
<td>1851-1860</td>
<td>718</td>
<td>84</td>
<td>10</td>
<td>864</td>
<td>20</td>
<td>1663</td>
<td>49</td>
<td>4550</td>
<td>71</td>
<td>6972</td>
</tr>
<tr>
<td>1841-1850</td>
<td>456</td>
<td>55</td>
<td>6</td>
<td>549</td>
<td>7</td>
<td>766</td>
<td>29</td>
<td>3054</td>
<td>43</td>
<td>3945</td>
</tr>
<tr>
<td>1831-1840</td>
<td>296</td>
<td>36</td>
<td>2</td>
<td>240</td>
<td>5</td>
<td>600</td>
<td>22</td>
<td>2307</td>
<td>25</td>
<td>2021</td>
</tr>
<tr>
<td>1821-1830</td>
<td>231</td>
<td>26</td>
<td>2</td>
<td>240</td>
<td>5</td>
<td>600</td>
<td>17</td>
<td>1896</td>
<td>15</td>
<td>1295</td>
</tr>
<tr>
<td>1811-1820</td>
<td>158</td>
<td>21</td>
<td>2</td>
<td>240</td>
<td>5</td>
<td>479</td>
<td>14</td>
<td>1344</td>
<td>9</td>
<td>764</td>
</tr>
<tr>
<td>1801-1810</td>
<td>125</td>
<td>13</td>
<td>2</td>
<td>240</td>
<td>3</td>
<td>264</td>
<td>10</td>
<td>1200</td>
<td>5</td>
<td>503</td>
</tr>
<tr>
<td>1791-1800</td>
<td>113</td>
<td>12</td>
<td>2</td>
<td>240</td>
<td>2</td>
<td>240</td>
<td>10</td>
<td>1108</td>
<td>3</td>
<td>360</td>
</tr>
<tr>
<td>1781-1790</td>
<td>115</td>
<td>12</td>
<td>2</td>
<td>240</td>
<td>2</td>
<td>240</td>
<td>9</td>
<td>1079</td>
<td>4</td>
<td>372</td>
</tr>
<tr>
<td>1771-1780</td>
<td>57</td>
<td>7</td>
<td>2</td>
<td>240</td>
<td>2</td>
<td>240</td>
<td>6</td>
<td>564</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>1761-1770</td>
<td>34</td>
<td>6</td>
<td>2</td>
<td>132</td>
<td>2</td>
<td>156</td>
<td>4</td>
<td>264</td>
<td>2</td>
<td>142</td>
</tr>
<tr>
<td>1751-1760</td>
<td>16</td>
<td>2</td>
<td>1</td>
<td>72</td>
<td>1</td>
<td>72</td>
<td>1</td>
<td>72</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>1741-1750</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td>1731-1740</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>36</td>
</tr>
</tbody>
</table>

TOTAL DATA COVERAGE: 90452 2451 2417 595703 2317 556729 2771 927326 2814 991884
SURFACE DATA

- Continue to obtain NMC and NAVY data on tape
- Obtain more NMC data in real time as the need develops
- Make the global 1900-50 USAF-ETAC data available
- NCAR will soon receive the US daily co-op data from NCC
- Improve the data set of year-month surface data
- Obtain selected additional daily time series of precipitation and temperature—global
DATA EXCHANGE

USSR

- Monthly station temperature and precipitation, Northern Hemisphere
- Daily data (temperature, precipitation, other), 120 USSR stations from 1880
- Ship data
- Solar data
- RAOBS 1960-1970

INDIA

- Have RAOBS
- Negotiating for daily rainfall

23

1982
R Penne
Figure 2-1e.
Ship Observations September 1 - 15, 1980
AUTUMNS PRECEDING SEESAW WINTERS
SEA LEVEL DIFFERENCE (mm) GB - GA

AMERICAN CITY
PORTSMOUTH
HAMPTON
WILMINGTON
CHARLESTON
FT. PULASKI
MAYPORT
MIAMI

BERMUDA

-63.07
-12.47
-60.41
-27.73
Ocean Data

At Depth

- Station data 560,000 reports, 1900 to June 1975
  - Select from 200,000 adnl ($10 each)
- MBT, XBT 1,020,000 observations now 1955

Tide and Sea Level

- Prepare about 200 global stations
  - Yr-Mo USA - 100 Stations 1800

Deep Seat Tides and Bottom Pressure

- AOML: 10 Stations (Pressure and Currents)
- Check for Other Global Data

Fixed Buoy and Platform Data

- Prepare data for a selection of platforms

Long Time Series of Daily Data

- About each 200 km along coasts prepare time series of daily atmospheric and ocean data

Surface Currents

- Have 4,200,000 ship observations
- Prepare monthly summary
- Prepare drifting buoy data

Sub-Surface Currents

- Give special priority to data sets in areas with major current systems

Fig. 2-13 Ocean data

27
HISTORICAL AND PALEOCLIMATIC DATA

HISTORICAL RECORDS
  • MAJOR EFFORTS UK, SWITZERLAND, CHINA

TREE RINGS
  • ABOUT 1,000 TREE SITES IN WORLD
  • SEASONAL RECONSTRUCTIONS WESTERN NORTH AMERICA

COHMAP (POLEN)
  • NORTHERN HEMISPHERE TEMPERATURE AT 3,000, 6,000, 9,0000, AND 12,000 YEARS BP
  • DECADE POLLEN DATA TO 1,000 BC (90 SITES)
  • CENTURY DATA TO 8,000 BC (90 SITES)

ICE CORES
  • GREENLAND
    • 1 TO 100,000 YEARS
    • 5 TO 1,000-2,000 YEARS
  • ANTARCTICA
    • ABOUT 4 CORES

CLIMAP
  • RECONSTRUCTIONS FOR 18,000 YEARS BP

CONTINENTAL DRIFT
  • RECONSTRUCTIONS TO ABOUT 300 MILLION YEARS BP
RADIATION

- USSR ARCHIVE FOR 600 SURFACE STATIONS FROM ABOUT 1965

- MINUTES OF SUNSHINE: MANY STATIONS

- SATELLITE MONTHLY RADIATION GRIDS FROM CSU
  INCIDENT RADIATION
  REFLECTED
  EMITTED
  ALBEDO
  NET FLUX IS IMPLIED
  MINIMUM ALBEDO

- NOAA DAILY AND MONTHLY RADIATION
  1974 - CURRENT (GAP IN 1978)

- SOLAR CONSTANT
  PRELIMINARY DATA AVAILABLE FROM 1975

29
ELEVATION DATA

ONE DEGREE ELEVATION DATA

TEN MINUTE GLOBAL ELEVATION
   - WITH LAND USE

FIVE MINUTE ELEVATION
   NORTH AMERICA AND EUROPE

ONE KM ELEVATION (USA)

DEPTH

ONE DEGREE DEPTH DATA

FIVE MINUTE DEPTH
   NORTHERN HEMISPHERE

30
DATA INFORMATION SOURCES

1975 (JENNE) DATA SETS FOR METEOROLOGICAL RESEARCH

1977 (WMO) STATISTICAL INFORMATION ON ACTIVITIES IN OPERATIONAL HYDROLOGY. WMO 464

1978 (WMO) CATALOGUE OF METEOROLOGICAL DATA FOR RESEARCH

1979 (POSEY, NASA) CANDIDATE NASA DATA SETS APPLICABLE TO THE CLIMATE PROGRAM

1979 (DREYFUS, ED—CODATA—ICSU). CODATA DIRECTORY OF DATA SOURCES FOR SCIENCE AND TECHNOLOGY. CHAPTER 2, HYDROLOGY

1981 (JENNE) THE GLOBAL DATA BASE FOR CLIMATIC RESEARCH (IN PREPARATION)

1980 (PREDOEHL, NOAA) THE INTERIM CLIMATE DATA INVENTORY
DATA INFORMATION

- USSR
- CHINA
- AUSTRALIA
- MALAYSIA
- INDIA
- CANADA
- Kenya
DATA EXCHANGE

USSR

- Monthly station temperature and precipitation, northern hemisphere
- Daily data (temperature, precipitation, other), 120 USSR stations from 1880
- Ship data
- Solar data
- RAOBS 1960-1970

INDIA

- Have RAOBS
- Negotiating for daily rainfall

33
Attachment 1

Selected projects in the World Climate Research Programme. Taken from the report of the fourth session of the JSC, Venice, March 1983. Several US scientists were members of the drafting group.


This 5-year project started data collection in July 1983. The six satellites send down about 62 terabits ($62 \times 10^{12}$) of data per year. This is sampled so that the two major archives for the experiment have 400 gigabits and 50 gigabits per year respectively. The archives will be used to prepare cloud and radiation statistics, and to improve the specification of cloud-radiation feedback processes in climate models.

2. International Precipitation Climatology Project (page 16).

Year-month averages of precipitation for $2^0$ squares are desired for land and ocean areas. For other projects, much additional daily precipitation data is needed.

3. Two projects include climate trends and interannual variability.

For this research, long records of land and ocean data are necessary.


This includes radiative processes in the lower troposphere, structure of cloud-topped boundary layers, role of boundary layer in mesoscale processes, fluxes over complex terrain including the role played by vegetation, and moisture budgets over limited tropical areas.

5. Joint study of hydrological and land-atmosphere processes (page 24).

A pilot study is being started to initiate work in this area.

6. Ocean heat flux studies (page 31).

Measure and model the global pattern of energy fluxes at the ocean-atmosphere interface. This involves satellite data and surface oceanic and atmospheric data.

7. Tropical Ocean and Global Atmosphere (TOGA) Program (page 29).

This program is aimed at studying the interannual variability of the tropical ocean and global atmosphere considered as a coupled system. The data will include XBT sections, surface ship data, and oceanographic satellites.
Roy Jenne
1 Sept. 1982

DATA SUPPORT SECTION

This text was prepared to help answer questions about how the Data Support Section (DSS) allocates its time to different tasks, and how much of the service that we provide results in support for the university or a wider community of users. We will include information about the sale of data (at cost), and about the amount of archive data being used on-line. Other information about services and costs is also included.

A. Data Support Staff Allocation of Time

Table 1 shows the breakdown of staff time by major functions. Note that most of our time is spent in data archive activities which include archive preparation, data distribution, documentation, and necessary program development. The last column shows the percentage of the activity that is estimated to represent a service to the research community outside of NCAR.

B. Data Requests Filled by Mailing Data

Paul Mulder surveyed the job orders for 1981 and found that for data sent out, the distribution of effort is given below. He gave some weight to the effort needed for each job, but noted that the average job size doesn't vary much between the different categories.

<table>
<thead>
<tr>
<th>Request Category</th>
<th>Portion of Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. U.S. Universities</td>
<td>41%</td>
</tr>
<tr>
<td>2. U.S. Government (major part) and private companies</td>
<td>49%</td>
</tr>
<tr>
<td>3. Foreign Governments</td>
<td>6%</td>
</tr>
<tr>
<td>4. Foreign Universities</td>
<td>4%</td>
</tr>
</tbody>
</table>

C. Charges for Data Requests from Outside NCAR

Note that selling prices were increased on 9 November 81:

1. From 9 Nov 81 - 26 Aug 82 we billed $31,927 for data.
2. The full year equivalent is $40,046.
3. In addition we gave away data on the basis of data exchange. The amount of "free" data for a full year was:

    Doc 83
The equivalent cost of this was about $8,000. Thus, the total value of data sent was about $48,050. This doesn't include the value of publications and movies. Many data sets go to other archives where they are used for more than one purpose. Thus there is also a multiplier effect. Also note that this "value" is measured only by minimal tape copy costs.

D. On-line and off-line use of data.

Much data is used on-line at NCAR and not charged-for by us. However, the computer time needed is charged to the project using the data. The following table shows the number of tape mounts and equivalent mass store accesses made that create or use DSS archives. The second column is included in the first total. The yearly access counts are for 1981. See Table 2.

Times per year that a tape or mass store data volume is accessed:

<table>
<thead>
<tr>
<th>Item</th>
<th>Jobs</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tapes</td>
<td>39</td>
<td>112 Tapes</td>
</tr>
<tr>
<td>Fiche &amp; DD80</td>
<td>14</td>
<td>6051 Pages</td>
</tr>
<tr>
<td>Printer</td>
<td>10</td>
<td>2644 Pages</td>
</tr>
</tbody>
</table>

Users of our archive volumes sometimes hit a basic archive tape several times within a week or two (of some concern). We estimate that there is an 80% rehit rate; thus only 56% of the user on-line hits are equivalent to mailing a tape. Using this rehit rate, and the above table, we estimate that the on-line use of data (by other than DSS) is equivalent to 1.35 times the amount mailed. The relative amounts of machine use are estimated to be as follows:

<table>
<thead>
<tr>
<th></th>
<th>Accesses by DSS</th>
<th>Accesses Not by DSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Accesses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for DSS</td>
<td>Accesses</td>
<td>On-Line accesses</td>
</tr>
<tr>
<td>to DSS</td>
<td></td>
<td>to DSS archives</td>
</tr>
<tr>
<td>send data *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tape Access</td>
<td>10549</td>
<td>(1579)</td>
</tr>
<tr>
<td>MS Access</td>
<td>3793</td>
<td>( )</td>
</tr>
<tr>
<td>Tapes Sent to Users</td>
<td>739</td>
<td>3053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2782</td>
</tr>
</tbody>
</table>

Users of our archive volumes sometimes hit a basic archive tape several times within a week or two (of some concern). We estimate that there is an 80% rehit rate; thus only 56% of the user on-line hits are equivalent to mailing a tape. Using this rehit rate, and the above table, we estimate that the on-line use of data (by other than DSS) is equivalent to 1.35 times the amount mailed. The relative amounts of machine use are estimated to be as follows:

* These are actual counts of data volumes accessed and tapes sent in 1981. A computer count would be 5% higher due to job failures.
1. Archive maintenance, update, and calculations by DSS. 4.9
2. All data sent 1.0
3. Used to prepare calculations and displays for users. .2
4. Used on-line by NCAR staff & visitors. .35
5. Used on-line by universities. 1.0
6. Thus the equivalent copy-cost of all data provided is about 2.55 times the value given in C above, or $122,530 per year.

These statistics ignore the data processing that we do on the 4341. For 1981, this assumption shouldn't alter the results very much. The effect is that the archive maintenance number would be somewhat larger than indicated.

E. Preparation of Resource Materials and Scientific Results.

The DSS outputs a variety of written material including data inventories, information about procedures used when data was prepared at NMC and elsewhere, and selected scientific results. In addition, several motion pictures have been produced for education and research uses. This includes movies for:

- climatology of S. Hemisphere
- daily 500 mb patterns, and 15, 30 and 60 day filtered patterns
- El Nino: Temperature and winds over the tropical Pacific Ocean
- the Stratosphere
- others

About 200 copies of the movie films have been sent out at cost.

F. Some data is provided to NCAR free of charge by NMC (NOAA):

- 4 tapes per week (208 tapes/year, value about $10,000) (each week there is 1 tape for grids, 1 U/A observed, 1 sfc synop, 1 sfc ship-3 hrly).
- other tapes are also received.

G. CRU Use By The Data Support Section, NCAR.
The increase in machine use in 1981 (Table 3) primarily was caused by the addition of three major data sets for climate and other research. There has also been an increase in the amount of data sent. The cost of a CDC-7600 CRU unit is about 35 cents, and is used to obtain the cost in Table 3. This cost seems a bit puzzling because the machine has the capability of outputting over 12 million of these units a year and that would give a return of $4.2 million dollars. Actual use is no longer as high.
<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>R. Jenne</th>
<th>D. Joseph</th>
<th>P. Mulder</th>
<th>W. Spangler</th>
<th>G. Walters</th>
<th>Total FTE's</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I. Data Archive Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Acquiring, preparing, and maintaining</td>
<td>8</td>
<td>25</td>
<td>39</td>
<td>20</td>
<td>20</td>
<td>1.12</td>
</tr>
<tr>
<td>data sets.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Distributing Data to Users</td>
<td>2</td>
<td>10</td>
<td>36</td>
<td>20</td>
<td>15</td>
<td>.83</td>
</tr>
<tr>
<td>3. Consulting on data sets</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>.31</td>
</tr>
<tr>
<td>4. Documenting and Cataloging</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>15</td>
<td>.40</td>
</tr>
<tr>
<td>5. Developing Programs</td>
<td>-</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>30</td>
<td>.58</td>
</tr>
<tr>
<td><strong>II. General User Support</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Consulting</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>.15</td>
</tr>
<tr>
<td>2. Developing General Purpose Codes</td>
<td>-</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>.10</td>
</tr>
<tr>
<td>**III. Direct Project Support and Research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programming</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>13</td>
<td>2</td>
<td>.23</td>
</tr>
<tr>
<td><strong>IV. Data Planning Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. National and International Planning</td>
<td>20</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>.26</td>
</tr>
<tr>
<td>2. Software, Hardware, and Data Planning</td>
<td>8</td>
<td>15</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>.24</td>
</tr>
<tr>
<td>for Selected Functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**V. Publications, Talks, and Classes on</td>
<td>20</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>.25</td>
</tr>
<tr>
<td>Data and Data Methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>VI. Staff Education and Training</strong></td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>.18</td>
</tr>
<tr>
<td><strong>VII. General Administrative Functions</strong></td>
<td>20</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>.35</td>
</tr>
</tbody>
</table>

Total FTE's by Function: 5.00

Percentage of Staff Member Time
(10% is approximately 23 working days/year)
All staff members are fulltime.
Table 2. Access to DSS data archives by DSS and by others. The number of times that tapes or TBM archive volumes are read or written by the Data Support Section is on the left. The number of times other people read DSS archive data is also given. The count of jobs includes times when an archive volume is accessed that has already been staged to a TLIB disk.

<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Jobs</th>
<th>Tapes</th>
<th>TBM</th>
<th>Jobs</th>
<th>Tapes</th>
<th>TBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>07</td>
<td>907</td>
<td>1152</td>
<td>360</td>
<td>779</td>
<td>342</td>
<td>336</td>
</tr>
<tr>
<td></td>
<td>08</td>
<td>734</td>
<td>761</td>
<td>427</td>
<td>886</td>
<td>437</td>
<td>302</td>
</tr>
<tr>
<td></td>
<td>09</td>
<td>675</td>
<td>939</td>
<td>129</td>
<td>591</td>
<td>386</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>811</td>
<td>1117</td>
<td>94</td>
<td>780</td>
<td>473</td>
<td>184</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>769</td>
<td>1087</td>
<td>193</td>
<td>865</td>
<td>342</td>
<td>305</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>555</td>
<td>757</td>
<td>231</td>
<td>443</td>
<td>190</td>
<td>179</td>
</tr>
<tr>
<td>1981</td>
<td>1</td>
<td>640</td>
<td>830</td>
<td>162</td>
<td>628</td>
<td>258</td>
<td>348</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>650</td>
<td>795</td>
<td>283</td>
<td>594</td>
<td>215</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>693</td>
<td>849</td>
<td>263</td>
<td>674</td>
<td>243</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>704</td>
<td>837</td>
<td>186</td>
<td>446</td>
<td>179</td>
<td>242</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>784</td>
<td>992</td>
<td>229</td>
<td>317</td>
<td>190</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>662</td>
<td>902</td>
<td>145</td>
<td>458</td>
<td>267</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>577</td>
<td>816</td>
<td>257</td>
<td>432</td>
<td>289</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>513</td>
<td>570</td>
<td>200</td>
<td>526</td>
<td>155</td>
<td>315</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>910</td>
<td>1140</td>
<td>465</td>
<td>698</td>
<td>301</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>884</td>
<td>1171</td>
<td>596</td>
<td>637</td>
<td>387</td>
<td>190</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>826</td>
<td>971</td>
<td>396</td>
<td>525</td>
<td>354</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>768</td>
<td>677</td>
<td>611</td>
<td>560</td>
<td>215</td>
<td>262</td>
</tr>
</tbody>
</table>

1981 Total | 10549 | 3793 | 3053 | 2782
Table 3. CRU Use on CDC-7600 by Data Support Section

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>7091</td>
<td>12912</td>
<td>14182</td>
<td>23946</td>
<td>25514</td>
<td>24975</td>
<td>25470</td>
</tr>
<tr>
<td>Feb</td>
<td>9397</td>
<td>12635</td>
<td>13335</td>
<td>43451</td>
<td>36757</td>
<td>33050</td>
<td>45243</td>
</tr>
<tr>
<td>Mar</td>
<td>13124</td>
<td>14923</td>
<td>18186</td>
<td>37761</td>
<td>27139</td>
<td>41930</td>
<td>48627</td>
</tr>
<tr>
<td>Apr</td>
<td>12874</td>
<td>12368</td>
<td>13524</td>
<td>31892</td>
<td>46759</td>
<td>35484</td>
<td>44776</td>
</tr>
<tr>
<td>May</td>
<td>10366</td>
<td>10235</td>
<td>32356</td>
<td>22423</td>
<td>25971</td>
<td>56918</td>
<td>26625</td>
</tr>
<tr>
<td>June</td>
<td>9775</td>
<td>12233</td>
<td>38046</td>
<td>24839</td>
<td>28427</td>
<td>39825</td>
<td>47970</td>
</tr>
<tr>
<td>Jul</td>
<td>7768</td>
<td>5582</td>
<td>23932</td>
<td>28162</td>
<td>24216</td>
<td>46731</td>
<td>27705</td>
</tr>
<tr>
<td>Aug</td>
<td>5902</td>
<td>8188</td>
<td>30527</td>
<td>18613</td>
<td>26500</td>
<td>35494</td>
<td>37194</td>
</tr>
<tr>
<td>Sept</td>
<td>13804</td>
<td>7784</td>
<td>25824</td>
<td>28847</td>
<td>26657</td>
<td>70710</td>
<td>60863</td>
</tr>
<tr>
<td>Oct</td>
<td>17078</td>
<td>9921</td>
<td>30031</td>
<td>30386</td>
<td>25254</td>
<td>69727</td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td>16334</td>
<td>16333</td>
<td>33323</td>
<td>21351</td>
<td>27408</td>
<td>54270</td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td>16499</td>
<td>16411</td>
<td>26178</td>
<td>15356</td>
<td>23840</td>
<td>40081</td>
<td></td>
</tr>
</tbody>
</table>

|       | 140012 | 139525 | 388285* | 327027 | 344512 | 549195 |

| TIMES 2.5 | 350030 | 348810 | 115400 | 114460 | 120580 | 192220 |

*Jan thru April 1978, CRUs were multiplied by 2.5 before adding for the 1978 total. A new charging algorithm was started in May 1978 that gave much higher charges to I/O processes. This increased the charges for our average job by a factor of 2.5.
DATA SUPPORT SECTION

- Staff, Budget, Time Use

- Archive Tasks
  - Keep track of data sets
  - Update tasks
  - Size of data sets

- Selected Data Tasks
  - Ship project
  - U.S. surface hourly data
  - World surface synoptic data
  - Daily precipitation data from India
  - Miscellaneous data sets

- Computing
  - Analysis of strategies
  - Mass storage, supercomputing

- Data Plans and Availability
  - Selected publications
  - Selected involvements for data
  - Project involvement, TOGA, WOCE, etc.

- The Future
  - Thrusts in world climate research
  - Bilateral data exchanges, etc.

Roy Jenne
March 1986

Some data centers in USA
Some Questions

• What new discipline areas (if any) do we need to cover?

• Data delivery
  - How far to go on generating special disks, etc.
  - Prepare to generate subsets for user to get via communications

• Display aids for users
  - Unidata should help
  - Ease of making movies, VCR

• Need for more data description and inventory work

• More satellite data
  - A satellite data facility is being talked about
  - Suggest we do not try to get every high resolution bit
  - Costs are important
  - Save NOAA TOVS data from destruction

• How many people
  - Especially need to replace one who will retire
  - NSF would pay for an Indian visitor
  - A number of projects are backed up from when we had a short term visitor

• Cost
  - How control data flow costs & storage costs
FUNCTIONS NECESSARY FOR DATA ANALYSIS AND DISPLAY

1. Obtain and prepare data sets

2. Accomplish archive functions

3. Move datasets within the hardware system

4. Maintain general aids for unblocking and unpacking data

5. Prepare access routines for each data set
   - Ability to provide small subsets of data via communications

6. Prepare aids for calculations

7. Make calculations (user function)

8. Display the output (mostly a user function)

9. Provide data inventory information
<table>
<thead>
<tr>
<th>Description</th>
<th>FY 86</th>
<th>FTE</th>
<th>Salary</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Data Archive Activities</td>
<td></td>
<td>2.40</td>
<td>$118</td>
<td></td>
</tr>
<tr>
<td>1. Acquiring, Preparing and Maintaining Datasets.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Documenting and Cataloging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Developing Programs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. General User Support</td>
<td></td>
<td>1.58</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>1. Distributing Data to Users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Consulting on Use of Datasets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Consulting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Developing General Purpose Codes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Direct Project Support and Research Programming</td>
<td></td>
<td>.55</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>D. Support NSF Radar Project</td>
<td></td>
<td>.85</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>E. Data Planning Activities</td>
<td></td>
<td>.59</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>1. National International Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Software, Hardware, and Data Planning for Selected Functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. Public, Talks, Classes</td>
<td></td>
<td>.33</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>G. Staff Education and Training</td>
<td></td>
<td>.36</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>H. General Adm. and Internal Plans</td>
<td></td>
<td>.34</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>7.00</td>
<td>$341</td>
<td></td>
</tr>
</tbody>
</table>

* This is in a separate budget.
TABLE I -- Data Support Section Staff Time by Major Function

Percentage of Staff Member Time
(10% is approximately 23 working days/year)
All staff members are fulltime.

<table>
<thead>
<tr>
<th>FUNCTION</th>
<th>R. Jenne</th>
<th>D. Joseph</th>
<th>P. Mulder</th>
<th>W. Spangler</th>
<th>G. Walters</th>
<th>R. Barnes</th>
<th>New Hire</th>
<th>Total FTE's by Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Data Archive Activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Acquiring, preparing, and maintaining datasets.</td>
<td>8</td>
<td>25</td>
<td>38</td>
<td>25</td>
<td>20</td>
<td>5</td>
<td>31</td>
<td>1.21</td>
</tr>
<tr>
<td>2. Documenting and Cataloging</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>12</td>
<td>2</td>
<td></td>
<td>.39</td>
</tr>
<tr>
<td>3. Developing Programs</td>
<td></td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>20</td>
<td>-</td>
<td></td>
<td>.49</td>
</tr>
<tr>
<td>B. General User Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Distributing Data to Users</td>
<td>2</td>
<td>12</td>
<td>38</td>
<td>20</td>
<td>20</td>
<td>8</td>
<td></td>
<td>.08</td>
</tr>
<tr>
<td>2. Consulting on Use of Datasets</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>-</td>
<td></td>
<td>.28</td>
</tr>
<tr>
<td>3. Consulting</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>-</td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td>4. Developing General Purpose Codes</td>
<td></td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td></td>
<td>.07</td>
</tr>
<tr>
<td>C. Direct Project Support and Research Programming</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. Support NSF Incoherent Scatter Radar Data Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Data Planning Activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. National and International Planning</td>
<td>24</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
<td>.15</td>
</tr>
<tr>
<td>2. Software, Hardware, and Data Planning for Selected Functions</td>
<td>8</td>
<td>5</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td></td>
<td>.14</td>
</tr>
<tr>
<td>F. Publications, Talks, and Classes on Data and Data Methods</td>
<td>18</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>10</td>
<td>.33</td>
</tr>
<tr>
<td>G. Staff Education and Training</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>-</td>
<td>20</td>
<td>.36</td>
</tr>
<tr>
<td>H. General Administrative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal Plans</td>
<td>18</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>.34</td>
</tr>
</tbody>
</table>

100% 100 100 100 100 100 100 7.00 FTE
## NCAR DATA SUPPORT

### Data Requests

<table>
<thead>
<tr>
<th>Year</th>
<th>Outside Requests</th>
<th>Input Tapes</th>
<th>Output Tapes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan - Dec 1977</td>
<td>140</td>
<td>1,237</td>
<td>511</td>
</tr>
<tr>
<td>Jan - Dec 1978</td>
<td>190</td>
<td>1,400</td>
<td>780</td>
</tr>
<tr>
<td>Jan - Dec 1979</td>
<td>200</td>
<td>1,237</td>
<td>511</td>
</tr>
<tr>
<td>Jan - Dec 1980</td>
<td>240</td>
<td>1,350</td>
<td>600</td>
</tr>
<tr>
<td>Jan - Dec 1981</td>
<td>186</td>
<td>1,045</td>
<td>550</td>
</tr>
<tr>
<td>Oct 1981 - Sept 1982</td>
<td>230</td>
<td>1,654</td>
<td>708</td>
</tr>
<tr>
<td>Oct 1982 - Sept 1983</td>
<td>190</td>
<td>979</td>
<td>452</td>
</tr>
<tr>
<td>Oct 1983 - Sept 1984</td>
<td>220</td>
<td>1,680</td>
<td>750</td>
</tr>
<tr>
<td>Oct 1984 - Sept 1985</td>
<td>246</td>
<td>1,697</td>
<td>784</td>
</tr>
</tbody>
</table>

Info about the 720 tapes mailed in 1984:
- Significant use of 6250 BPI tapes started late 1983 (some from 1981)
- We sent 30 to 40% of the tapes at 6250 BPI in 1984
- 95 tapes sent to India at 800 BPI, early 1984
- Rest of tapes sent at 1600 BPI in 1984

### REQUESTS FOR DATA INFORMATION AND PROGRAMMING HELP:

Walk in and telephone requests about data (estimate, Jan 1984):
- About 420 per month (5000 per year)
Ways to extend what we can do

- Cooperate with ERL, CIRES, NCDC on ships
- Gather the best of community efforts
- NSF would pay for an Indian visitor a year
- Perhaps use some part time help,

Roy Jenne
March 1986
TIME USED FOR DATA SENT (NOV 81-OCT 82)

US UNIVERSITIES 41%
US GOVT (MOST) & PRIVATE 49%
FOREIGN UNIVERSITIES 4%
FOREIGN GOVERNMENTS 6%

CHARGES FOR DATA SENT

- BILLED $40,050
- SENT FREE ON EXCHANGE 8,000

Total $48,050

DDS ACCESS

<table>
<thead>
<tr>
<th>Mass Store</th>
<th>Prepare</th>
<th>To Send</th>
<th>Access</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tape, MS Accesses</td>
<td>Archives</td>
<td>Data</td>
<td>NOT DSS</td>
</tr>
<tr>
<td>Equivalent Accesses</td>
<td>12648</td>
<td>1658</td>
<td>5835</td>
</tr>
<tr>
<td></td>
<td>12648</td>
<td>1658</td>
<td>3268</td>
</tr>
</tbody>
</table>

Direct on-line access to data volumes.
**SPENT IN FY 83**

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salaries &amp; Benefits</td>
<td>$240,000</td>
</tr>
<tr>
<td>Supplies</td>
<td>650</td>
</tr>
<tr>
<td>Data Purchase</td>
<td>11,400</td>
</tr>
<tr>
<td>Domestic Travel</td>
<td>3,000</td>
</tr>
<tr>
<td>Foreign Travel</td>
<td>2,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>$257,050</strong></td>
</tr>
</tbody>
</table>

CRAY use (108 CCU * $1500) *(Plus extra use for ships)* $162,000

**FY 1985**

Non-computing budget FY 85, about $300,000
Cray use in FY 85 approx 380 GAU, about $280,000
AVOID TRAPS

- Tightly Coupled Systems
  — The data must be in system format to use it.

- All data are very valuable and need equal care.
  — Little used or less critical data should be handled differently.

- A full McIdas type display should be available for all data.
  — This can be built up with user participation, not as a big system.

- Use a formal DBMS, even for large datasets.
  — Timing is lousy.

- Offer unlimited format flexibility to users.
  — Takes too much time.
  — A compromise will work well.

- The communications era adds a new set of problems.
  \textit{yes}

- AI must be developed to have a decent data system.
  \textit{no}

- We don't have to gather and prepare datasets, because communications can get the pieces when needed.
  \textit{no}

R. Jenne
March 1986
<table>
<thead>
<tr>
<th>Function</th>
<th>SDSD</th>
<th>PCDS</th>
<th>PODS</th>
<th>NASA Project</th>
<th>NCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prepare Data</td>
<td>22%</td>
<td>3</td>
<td>12</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td>Do Calculations</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>User Support</td>
<td>47</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>Data Calculations Aids</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>User Display Output</td>
<td>6</td>
<td>44</td>
<td>29</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Support Interactive</td>
<td>1</td>
<td>25</td>
<td>24</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Communications Systems</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Publish, Plan, etc.</td>
<td>12</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>18</td>
</tr>
</tbody>
</table>

| Percent                   | 100  | 100  | 100  | 100          | 100  |
| Approx. Budget $10^6      | $2.2 | 3.5  | 2.0  | --           | 0.6  |
| Total People              | 50   | 50   | 21   | --           | 6    |
Data Support Section

Data Support for "Other" Disciplines

1. Inventory of datasets. Provide system hooks so that a discipline inventory can be browsed on-line as well as on-paper. The discipline would do most of the work of preparing and updating the inventory.

2. Provide consulting advice on data handling on the NCAR systems, such as how to load datasets onto the mass-store.

3. Loading of portions of a discipline archive onto NCAR Mass Store, or tape archive.

If a discipline has a large archive that needs to reside at NCAR, we will work with the discipline (consultive) to load the data and data information onto the mass-store.

- Users then can use NCAR for general data, consultive help

- Users use the discipline archive for technical information about the archives, and for some consultive help.

- When data access programs are written, they could be logged into the system for use by other users. The persons placing these and other programs in the archive would be responsible for any necessary support.

- NCAR will provide a facility so that users can contribute various programs and display routines that may be of general interest (up to a space limit). The discipline would need to periodically help classify, index and purge this set so that it could be used easily. NCAR might keep subroutine use statistics on the routines.

4. Make a facility for users to contribute small datasets along with comments that give format and summary information to facilitate browsing.

5. Make our data handling tools available to the community and provide consultive help in their use.
# DATA SETS

**TOTAL SETS IN DSS COLLECTION**: 237

- **Gridded Data**: 77
  - Daily: 46
  - Means and Climatology: 21
  - Oceanographic: 10

- **Station Data**: 60
  - Upper Air: 15
  - Surface: 36
  - Surface Means: 9

- **Data from Experiments (GATE, MONEX, FGGE, etc.)**: 22

- **Satellite Observed and Derived Products**: 22

- **Geophysical (Mostly Elevation Data)**: 11

- **Other**: 45
## SELECTED DATA SET SIZES

<table>
<thead>
<tr>
<th></th>
<th>TAPES IN PRIMARY SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. NMC Global Analyses (July 1976 – Present)</td>
<td>170</td>
</tr>
<tr>
<td>2. NMC Upper Air Obs (1962 – Present)</td>
<td>260</td>
</tr>
<tr>
<td>3. Time Series Raob Data</td>
<td>170</td>
</tr>
<tr>
<td>4. LFM Analyses and Forecasts (Sept 1972 – Present)</td>
<td>155</td>
</tr>
<tr>
<td>5. FGGE Analyses by ECMWF - Full Resolution</td>
<td></td>
</tr>
<tr>
<td>(Half Resolution)</td>
<td>86</td>
</tr>
<tr>
<td>6. United States Co-op Data</td>
<td>48</td>
</tr>
<tr>
<td>7. Various Grid Time Series Sets</td>
<td>30</td>
</tr>
<tr>
<td>8. World Monthly Surface Climatology (Binary)</td>
<td>1</td>
</tr>
<tr>
<td>(Character)</td>
<td>3</td>
</tr>
<tr>
<td>9. Navy 10’ Elevation Data</td>
<td>1</td>
</tr>
<tr>
<td>10. Nicholson African Rainfall</td>
<td>1</td>
</tr>
</tbody>
</table>

### MAIN PRIMARY DATA VOLUMES:

- **Tapes**: 2813
- **TBM**: 1565

**Total Data Volumes - 13,900**

v4 1984
DAILY PRECIPITATION DATA FROM INDIA

- Daily Precipitation Data for Jan 1901 - Dec 1970
- 1632 stations in 1901
  2195 stations in 1930
  2402 stations in 1950
  3203 stations in 1957
  2536 stations in 1970
- 1473 stations have 67 or more years of data
- On 4 tapes (6250 BPI) at NCAR

- 20918 station-decades of data
  - 76.4 million daily rainfall observations

This data was received from India about 1982. We provided analysis, etc. in return.

- R. Jenne
MISCELLANEOUS DATA SETS

- GISS Vegetation, Land Use, and Albedo Data
- Palmer Drought Index by State and District
- Circulation Statistics By Lau and White
- Various Ozone Sets
- Rocket Sounding Data
- Aircraft Data from GATE, MONEX, and ALPEX
- Various Station Library Sets
- Manually Digitized Radar Set
- Selected Oceanographic Data
COMPUTING METHODS

- Strategies to develop and access large sets of atmospheric data (Jan 81)
- Talk to NAS about data handling and processing (Jun 81)
- Talk to CDC about file and data management (Sep 81)
- Cost and efficiency of (processing) in computing systems (Jan 82)
- NAS panel “Solar-Terrestrial Data Access, Distribution, and Archiving” (1982-83)
- NCAR AVC document ”Scientific Data Processing” (Dec 82)
- Mass storage analysis and projections (Jun 83)
- Problems and costs of preparing and accessing databases (Mar 83)
- Computing/data information for countries (for WMO, 1983)
- AVC data flow information (1984)
- Data packing, data flow and storage costs for different memory levels (Oct 84)
- Supercomputer output (July 1985)
- Mass storage and dataflow issues in preparation for NCAR X-MP (Sep 85)
  - Also a related larger text

March 1986
SELECTED INVOLVEMENTS FOR DATA

- International Satellite Cloud Climate Program
  — Several meetings, papers, talks to define program (over 3 years)
  — Data started July 1983

- Academy Panel on Data, CSSP/CSTR
  — Space Sciences & Solar Terrestrial (1982-83)

- New NAS Panel for Climate Data Problems
  — 1983-1986

- Problems of Saving NOAA Satellite Data

- NSF Program: Incoherent Scatter Radar
  — NCAR has central database
OTHER ACTIVITIES

• **On-line Documentation Maintenance**
  
  - Volume & summary information for each set is available to users with CRAY access

• **General Consulting**
  
  - Data set access (programs and tape numbers)
  - Finding appropriate data sets
  - Tape and other I/O problems
  - Handling data transfer to/from NCAR’s System
  - Data formatting problems
SELECTED PUBLICATIONS ON DATA PLANS & AVAILABILITY

- Planning Guidance for World Climate Data System
  - Jan, Feb 1980 in Geneva to write text
  - Published by WMO in 1982

- New Handbook of Applied Meteorology
  - 163 page data chapter (print 1985)

- Data Availability at NCAR
  - Updated each 6 to 15 months

- Selected Data Initiatives (for NCPO, Oct. 82)

- Data for Antarctic Research
  - SCAR meeting in Hamburg, Aug 1983
  - Prepared inventory information and helped on plan

- Global Data for Climate Diagnostic Research
  - Paper for Climate Diagnostic Workshop, Oct 1982

- List of Satellite Data for Monsoon Research, 1984

- Data for Monsoon Research (Oct 1984)

- Data Support Section, 5-Year Plan (July 1985)
  - Ocean Data Tasks in Data Support (Feb 1986)
  - Data for TOGA (May 1985)
INFORMATION ABOUT DATA

Sea Surface Temperature Data (Oct 82 - Dec 83)

NMC Analyses & Satellite Soundings (Feb 82)

Status of FGGE Data (Feb 82)

Daily Rainfall Data from India (Dec 83)

Inventory of Japanese Data (1983)

Compare Analyses

NMC Stratospheric Analyses (E 1982)

Papers About Ship Data, Analysis Methods, Navy Analyses (1983-86)

India Trip Notes (Sep 1984)

Kenya Trip Notes (Apr 1985)
THRUSTS IN WORLD CLIMATE RESEARCH

(A Challenge for Data Support)

- Satellite Cloud Climatology
- Precipitation Climatology
- Climate Trends and Interannual Variability
- Atmospheric Boundary Layer Processes
- Hydrological and Land-Atmosphere Processes
- Ocean Heat Flux Studies
- Tropical Ocean and Global Atmosphere (TOGA)
SELECTED CLIMATE STUDIES NEEDING DATA

1. CIRCULATION STATISTICS
2. ATMOSPHERIC TEMPERATURE TRENDS
3. SURFACE TEMPERATURE TRENDS
4. SURFACE AND UPPER AIR CLIMATOLOGY
5. LARGE SCALE CLIMATIC RELATIONSHIPS
6. RAINFALL COMPARED TO OCEAN CONDITIONS
7. PACIFIC EQUATORIAL STUDIES
8. TROPICAL STORMS
9. BLOCKING PATTERNS/CLIMATE PREDICTION
10. OCEAN STUDIES
11. HYDROLOGY
12. PREDICTABILITY
13. PALEOCLIMATIC STUDIES
14. LOCAL EFFECTS MODELS
15. SOLAR AND VOLCANIC EFFECTS
DATA FOR DISCIPLINES

- Core Disciplines
  - Meteorology and Oceanography, some hydrology

- NCAR/NSF Incoherent Radar Data
  - Scientist HAO, programmer in Data Support

- NCAR Archives for Other Disciplines
  - We could work with other disciplines to help them establish archives at NCAR.

- Must get ready for multidiscipline projects.
  - Monitor status of other data.
  - Will other necessary archives survive?
  - Get a better set of paleoclimatic data.
  - Firor and others want some help with socio-economic data.

- Field Project Data
  - Our role is to provide the "background" data.
  - Others should prepare the project data and answer questions for some time.
  - We archive parts of it.
SATELLITE DATA

- Usually a professor has $5000 in annual budget for data, etc.

- Sums like $15,000 and more are often removed from proposals.

- Processing satellite data is often data cost intensive and compute cost intensive.

- One plan is to set up a satellite data facility that gives out free data.
  - Some say up to 10 people.
  - A data purchase budget of perhaps $100K/yr.
  - Absorb the cost of giving data free.
  - But, should it give free data?
  - I think the facility should start small and be monitored.

- We need efforts to reduce real data handling costs.
  - But some want to discourage data flow and archives by making high costs.
  - Hardware configurations to be chosen to reduce costs.
  - Technology needs to be pushed to reduce costs.
Cost comparisons of computing

The following shows that the cost of computing has come down a lot. The cost of IO capability has also come down, but the charges have not.

<table>
<thead>
<tr>
<th>Compute charges</th>
<th>NCAR CDC-7600</th>
<th>Cray 1 As NCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU, Task</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day</td>
<td>$864/Task</td>
<td>$864/Task</td>
</tr>
<tr>
<td>Night</td>
<td>$86</td>
<td>$86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use</th>
<th>NCAR</th>
<th>Cray 1 As NCAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>All memory</td>
<td>$648/comp.hr</td>
<td>$323/comp.hr</td>
</tr>
<tr>
<td>Little memory</td>
<td>404/202</td>
<td>$750/250</td>
</tr>
<tr>
<td>Non-technical</td>
<td>$250/30</td>
<td></td>
</tr>
</tbody>
</table>

valid 1981-83

Therefore compute charges have come down $864 \rightarrow $323 \rightarrow $83

Data flow is less costly but the charges have not decreased.
New CRAY Charges, Especially Tape and TBM

Tape Copy Costs at NOAA-Boulder

For comparison, the costs to copy tapes at NOAA, Boulder will be given. These are substantially less than costs from the NCAR algorithm. I think that some of their low density costs may be too low. However, I think that our NCAR costs are too high.

The following costs to copy data are given for reasonably full 6250, 1600, and 800 BPI tapes at NOAA. The charge is based on the amount of data moved; the record size or tape density doesn't enter. The following information is from Scott Woodruff, CIRES:

<table>
<thead>
<tr>
<th>Density</th>
<th>Bits (10^6)</th>
<th>Bit Movement</th>
<th>Mounts</th>
<th>CPU &amp; Print, etc</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>6250</td>
<td>1000</td>
<td>$20.08</td>
<td>$.50</td>
<td>$2.</td>
<td>$22.50</td>
</tr>
<tr>
<td>1600</td>
<td>300</td>
<td>6.00</td>
<td>.50</td>
<td>1.</td>
<td>7.50</td>
</tr>
<tr>
<td>800</td>
<td>100</td>
<td>2.00</td>
<td>.50</td>
<td>.50</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Tape to Tape Cost at NOAA

<table>
<thead>
<tr>
<th>Density</th>
<th>Bits (10^6)</th>
<th>Bit Movement</th>
<th>Mounts</th>
<th>CPU &amp; Print, etc</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>6250</td>
<td>1000</td>
<td>$40.39</td>
<td>$1.50*</td>
<td>$2.</td>
<td>$43.90</td>
</tr>
<tr>
<td>1600</td>
<td>300</td>
<td>12.12</td>
<td>1.50</td>
<td>1.</td>
<td>14.60</td>
</tr>
<tr>
<td>800</td>
<td>100</td>
<td>4.04</td>
<td>1.50</td>
<td>.50</td>
<td>6.00</td>
</tr>
</tbody>
</table>

*The first mount costs $.50 and succeeding ones $1.00

Note: The cost of a disk-to-tape operation is thus about $1.44 per megaword (64 bit).

Feb 1986

Note: The CRAY charges to copy a 6250 BPI tape at NCAR are now $76. (1984-86 period)

- Cost to copy a 6250 tape now $76.
- On the 7600 the cost was $42.

U. Work rates 1985

They charge $60 per I/O hour for tapes at about $10.55 per 10^9 bits, even on a disk drive (1.55 mbit/sec).

A 6250 BPI tape copy would then be about $25.
TABLE 2: Data archive cost. The annual archive cost is based on 2.2 times the annual investment cost for media.

<table>
<thead>
<tr>
<th>Data on Offline Media</th>
<th>Media Cost Per Megabyte</th>
<th>Media Cost Per (10^9) Bits</th>
<th>Annual Archive Cost Per (10^9) Bits</th>
<th>Annual Archive Cost Per (10^{12}) Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use 556 bpi tapes with 80 megabits (cost $14, 1960-1970)</td>
<td>$1.40</td>
<td>$175</td>
<td>$57.75</td>
<td>$57,750</td>
</tr>
<tr>
<td>Use 6250 bpi tapes with 1000 megabits (cost $13, 1981-)</td>
<td>.104</td>
<td>13</td>
<td>4.29</td>
<td>4,290</td>
</tr>
<tr>
<td>A cartridge tape with 1350 megabits (cost $14, 1985-)</td>
<td>.083</td>
<td>10.4</td>
<td>3.43</td>
<td>3,430</td>
</tr>
<tr>
<td>An optical disk plate, 2000 megabytes (16,000 mgbits) cost $200 in 1986</td>
<td>.10</td>
<td>12.5</td>
<td>4.12</td>
<td>4,120</td>
</tr>
<tr>
<td>A cartridge tape in 1990 with 10,000 megabits cost $15</td>
<td>.012</td>
<td>1.50</td>
<td>.50</td>
<td>496</td>
</tr>
<tr>
<td>On-line disks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use a hard disk 300 megabytes (cost, $30,000)</td>
<td>$100</td>
<td>$12,500</td>
<td>$5,500</td>
<td>$5,500,000</td>
</tr>
<tr>
<td>New hard disk 5000 megabytes (cost, $100,000)</td>
<td>20</td>
<td>$2,500</td>
<td>1,100</td>
<td>$1,100,000</td>
</tr>
</tbody>
</table>

Examples of Data Volume:

- Daily high, low temperature and precipitation for 10 stations for 50 years (about 2.2 megabytes or 17.6 megabits).

- Five years of satellite sounder data and 4 km/5-channel global scanner data from one NOAA satellite (about \(10^{13}\) bits).

- Each (of two) US GOES satellite, data for one year is \(25.5 \times 10^{12}\) bits.
<table>
<thead>
<tr>
<th>Function</th>
<th>Effective Rate</th>
<th>Data Moved</th>
<th>Cost Foreground</th>
<th>Cost Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 min of TBM channel time</td>
<td>3.0 mbit/sec</td>
<td>1.8 Gbit</td>
<td>$48.60</td>
<td>$48.60</td>
</tr>
<tr>
<td>10 min of tape cartridge channel</td>
<td>7.0 mbit/sec</td>
<td>4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 min (later) tape cartridge</td>
<td>17 mbit/sec</td>
<td>10.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 min of disk channel time</td>
<td>20 mbit/sec</td>
<td>12.0</td>
<td>9.50</td>
<td>3.20</td>
</tr>
<tr>
<td>10 min CRAY-1 CPU time</td>
<td>-</td>
<td>-</td>
<td>$125</td>
<td>$41.70</td>
</tr>
</tbody>
</table>
Channel Time Needed to Move Data to/from Cartridges

We will now consider the time needed to move data over two channels to the IBM 3480 cartridge devices. In 1987 the channel rate will probably be up to an effective 17 mbit/sec., as seen below. Most of the data flow should actually be to the disk farm. The flow to cartridges should only be about 25% of that given below; thus, about 1.1 hours per day of channel time will be needed. Tape mounting and searching takes much more time than the channel time.

<table>
<thead>
<tr>
<th>Gbit/Day</th>
<th>Assumed 3480 Rate</th>
<th>Time to Move Data</th>
<th>Portion of Available Channel Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984-85, 2 Crays</td>
<td>99.3</td>
<td>7 mbit/sec.</td>
<td>3.94 channel-hr.</td>
</tr>
<tr>
<td>1987, X-MP, 1 CRAY-1A</td>
<td>258</td>
<td>7</td>
<td>10.24 channel-hr.</td>
</tr>
<tr>
<td>1987, X-MP, 1 CRAY-1A</td>
<td>258</td>
<td>17</td>
<td>4.22 channel-hr.</td>
</tr>
</tbody>
</table>

Data Flow History and Forecast to 1990

<table>
<thead>
<tr>
<th>Type of Storage</th>
<th>Bits/Tape ((10^3))</th>
<th>Speed ((\text{mbit/sec}))</th>
<th>Read Time for (10^5) Bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960 7 track, 556 bpi tapes</td>
<td>0.09</td>
<td>.50</td>
<td>.40</td>
</tr>
<tr>
<td>1972 9 track, 1600 bpi (150 IPS)</td>
<td>0.3</td>
<td>1.92</td>
<td>1.7</td>
</tr>
<tr>
<td>1980 9 track, 6250 (200 IPS)</td>
<td>1.0</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>1986 IBM Cartridge (1.6 Gbits)</td>
<td>1.45</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>1987 IBM Cartridge (1.6 Gbits)</td>
<td>1.45</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>1990 IBM Cartridge</td>
<td>8</td>
<td>?</td>
<td>35</td>
</tr>
</tbody>
</table>
PRESENT MAIN COMMITTEES, ETC.

- NAS Panel on Climate Related Data, 1983-86 (nearly done)
- NAS Committee on Geophysical Data About 1980 to June 1987
- JOI SODSSWG Ocean Science Data Working Group
  — From about early 1984
- NAS ( engr) Panel on Satellite Data (title), 1986
- WOCE/TOGA Data Management Working Group
  — From about early 1984
- NAS Ocean Climate Research Committee of the Ocean Studies Board
  (Feb 86 - Dec 87)
- Negotiations with ECMWF for their analysis data
  — Especially from March 1984
  — Involved from 1976
- US Focal Point for US-India Data Exchange (from Mar 83)
  — Many datasets to prepare and exchange for monsoon research
  — Many equipment issues
  — Several side program tasks
- US-China Data Exchange
  — Involved from Dec 80 (This will now have to wait.)
- SCD Computing
AVAILABILITY AND DELIVERY OF DATA  
(National and NCAR)

Questions We Hear

Progress in Data Availability

National Data Systems
   NASA, NOAA, USAF, (ETAC), DOE (CO₂), U Wisc, NCAR

Selected Data at NCAR

Aspect of Data Management
   Data set problems
   Avoid traps
   Trends
   Need exchange media
   Problems discussed, problems ignored
   NASA archive rates
   How often is a dataset used
   Archive costs

NCAR Data Plans

Dec BS

Roy Jenne
Dec 1987
DISCIPLINE DATASETS ARE NEEDED, READILY AVAILABLE

- Biggest problem often that data not prepared:
  - No surface wind data for US dustbowl in 1930s;
  - Upper air data before 1946 not digital.

MAIN THRUST

- For primary discipline areas we need to:
  - Prepare the necessary datasets;
    --Gather fruits of other efforts;
  - Copy data and extract subsets for users;
  - Handle user interface problems;
  - Provide associated information systems;
  - Computing, data flow, storage, exchange media, communications.
Selected Questions About Data

• I want to study relationships between Atlantic Ocean water temperature and rainfall over Africa.
  *Do you have the necessary data?*

• We are launching stratospheric balloons over Brazil. We need to know how they will drift.
  *Do you have wind information near 60,000 feet?*

• I need daily analyses of wind and temperature over the N. Hemisphere after 1960.
  *Do you have this?*

• I would like ozone data as derived from satellites.
  *Do you have these data?*

• My model needs the reflectance (albedo) of the world’s surface and the roughness.
  *Is there data to help?*

• I want to calculate the changes in climate and its variability over the last 100 years.
  *What data does NCAR have that would help?*

• We brought this tape from computer X, but can’t read it.
  *Here’s what happened. Do you know what might be wrong?*
PROGRESS

IN 1967 NCDC, ASHEVILLE HAD 600,000,000 CARDS:
- They filled the building;
- NCAR paid $4.50/1000 cards to buy data (1967-70);
- By 1978 could buy 500,000 cards on a 1600 BPI tape (for $60).
  (Cost $0.12 per 1000 cards)

RAOBS FROM NCDC:
- Some raobs purchased on cards during 1966-70;
- All NCDC USA raobs 1948-1986 at NCAR.
  (On 21 tapes at NCAR in binary (6250)

USAF TD-13 WORLD SFC SYNOP:
- Variable coverage 1930-65;

ANALYSES (Hemispheric and Global):
- NCAR went through over 2000 tapes for 1958-73;
- Also read thousands of cards;
- Now on 50 tapes (hemispheric).
PROGRESS

SHIP DATA:
- To 1982 The World’s Ship Data Was on Several Hundred Tapes;
  (Some was not released by a nation)
- Project: NCAR, CIRES, NCDC 1982-84;
- Now, 1850-1979 Data Is On 45 Tapes, Cleaned Up;
  (plus statistics)
- Needs Update.

US COOP DATA (10,000 STATIONS):
- To 1978 Was On About 700 Tapes At NCDC;
- Sent To NCAR On 50 Tapes In 1978 (1600 BPI);
- We Have It On 20 Tapes;
- Still Lots Of Work To Do.

NMC GLOBAL OBSERVED DATA (1962-72):
- 1971-74: NCAR Got Data From 2000 Tapes;
- First On 300 Tapes (7tr, 800 BPI);
- Now On 116 Tapes (1600 BPI).

NOAA VTPR SOUNDER DATA (1972-79):
- In 1979 On 1194 Tapes At NOAA;
- Tried 1979-83 To Get $25,000 To Copy Tapes;
  (Couldn’t obtain cash)
- In 1986 NCAR And NOAA/SDSD Cooperated;
- Now On 48 Tapes At NCAR.
NATIONAL DATA ACTIVITIES
NASA DATA

PLANETARY:
- 700-800 people in planetary sciences in US;
- Flight projects $300 million/year;
- Deep space com net about $125 million/year;
- NASA research budgets are about $15 million/year;
- Planetary data group $3 million/year.

PROCESS DATA FROM NIMBUS-7 (1978-1987):
- Seven major data streams;
- Cost $6 million/year.

COST TO DOWNLINK NEAR-EARTH DATA (1986):
- $175 million/year.

NASA INFO SYSTEMS OFFICE:
- Budget $20 million/year;
- Hq staff is 5 people;
- 86 civil service plus 200 contract FTE at Goddard;
  — includes NSSDC, pilot climate, pilot land;
  — includes supercomputer center (Cyber 205).
Dec 1987

NASA OCEAN DATA SYSTEM (JPL)

<table>
<thead>
<tr>
<th>FY</th>
<th>STAFF(FTE)</th>
<th>Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>12</td>
<td>$1350K</td>
</tr>
<tr>
<td>1982</td>
<td>15</td>
<td>1600</td>
</tr>
<tr>
<td>1983</td>
<td>20</td>
<td>2300</td>
</tr>
<tr>
<td>1984</td>
<td>23</td>
<td>2275</td>
</tr>
</tbody>
</table>

STAFF USE IN FY 1984

<table>
<thead>
<tr>
<th>FTE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer operations</td>
<td>3.7</td>
</tr>
<tr>
<td>Archive system</td>
<td>5.0 (keep it going)</td>
</tr>
<tr>
<td>Data workstation</td>
<td>4.8 (improve it)</td>
</tr>
<tr>
<td>Work on datasets</td>
<td>4.5</td>
</tr>
<tr>
<td>Management, workshop</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Total 23
NASA OSSA DATA SYSTEMS

COMMENT JUNE 86:
- $200 million/year being spent on OSSA data systems;
- Some think we aren’t getting our money’s worth.

CONCERN OF BERT EDELSON (June 86), Head of OSSA
- EOS data systems talk of needing $300 to $1000M;
- Is there a better way of doing business?
- Should an info system be 1%, 10%, 50% of budget?
- Wanted our committee to reflect on these questions.
COSTS

1. PLANETARY SYSTEMS
   Flight projects                           $300
   Deep space communications                125
   Planetary science research               15
   Planetary data system                    3

2. UARS
   One satellite, no instruments             $192.2
   Build 9 instruments (Scientist teams)     206
     —teams analyze data (Science)           14
     —associated data processing              4
   Theoretical research (11 teams)           13.7
   Central data support                      38.3

   $468.3

3. NOAA RAWINSONDE NETWORK
   (about 100 stations)
   Cost to operate network                  $20,000,000
   Cost to gather data & prepare dataset     200,000
   Cost for NCAR to buy copy                 300
   Cost to archive 40 years of data          500
     (40 to 100 tapes)

4. THOUGHT CASE FOR DATA MANAGEMENT
   COST/DATASET
   $2,000,000 for 20 datasets                 $200,000/yr
NATIONAL OCEANOGRAPHIC DATA CENTER

- Gathers Ocean Data for the USA
- Started About 1962
- Staff

<table>
<thead>
<tr>
<th>Year</th>
<th>FTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960s</td>
<td>120-145</td>
</tr>
<tr>
<td>About</td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>110-120</td>
</tr>
<tr>
<td>Sep 1985</td>
<td>96</td>
</tr>
<tr>
<td>May 1986</td>
<td>83</td>
</tr>
<tr>
<td>Aug 1987</td>
<td>85</td>
</tr>
</tbody>
</table>

(incl 15 in computing)
(computing moved)

- Budget

<table>
<thead>
<tr>
<th>FY</th>
<th>Millions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>$4.5</td>
</tr>
<tr>
<td>1985</td>
<td>4.1</td>
</tr>
<tr>
<td>1987</td>
<td>3.6</td>
</tr>
</tbody>
</table>
NATIONAL GEOPHYSICAL DATA CENTER

FY 1980 $3.1 million
  1983  3.0
  1986  2.6

- Has about 70 FTE
- Plus 11 FTE in R. Barry’s Ice and Snow group

USAF ETAC OL-A, Asheville

\begin{array}{ccc}
1960s & 190 \text{ people} & \text{ (much key entry)} \\
1985-86 & 67 & \\
Aug 1987 & 85 & \\
\end{array}

- Archives of early (1930-1965) world data are largely based on these efforts.
NATIONAL CLIMATIC DATA CENTER (1984)

- About 80,000 User Transactions/Yr
- 30,103 Non-digital Requests/FY83 $607,300
- 1,223 Digital Requests $673,100
- Send out 100,000 Subscriptions $450,000
  (each 13 mailings/yr)

Satellite Data:
- Copy 4,100 Tapes $600,000
- Satellite Photos, 1,600 jobs/18,000 Photos $171,700

Total Budget ~$12 million

Total Staff 361 FTE
OAK RIDGE CO$_2$ INFORMATION CENTER

An Organization For CO$_2$ Research And Data:
- Carbon Cycle Research Center;
  10 people
- Carbon Dioxide Information Center.
  10 people


### SATELLITE DATA AT THE UNIVERSITY OF WISCONSIN

<table>
<thead>
<tr>
<th>Tapes</th>
<th>Comments</th>
<th>Bits Each (10^9)</th>
<th>Total (10^{12} bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9747</td>
<td>GOES low density tapes</td>
<td>21</td>
<td>204.7</td>
</tr>
<tr>
<td>5760</td>
<td>GOES high density tapes through Mar 85</td>
<td>42</td>
<td>241.9</td>
</tr>
<tr>
<td>1460</td>
<td>GOES high density tapes (Apr 85 - Mar 87)</td>
<td>42</td>
<td>61.3</td>
</tr>
<tr>
<td>1215</td>
<td>Meteosat (FGGE, 6250 BPI)</td>
<td>1</td>
<td>1.22</td>
</tr>
<tr>
<td>5200</td>
<td>GMS (FGGE, 6250 BPI)</td>
<td>1</td>
<td>5.20</td>
</tr>
</tbody>
</table>

**TOTAL Archive** 514.32 Tbits

error rate 10^{-5} to 10^{-6}

### COST OF THE ARCHIVE: (Valid 1987 & for the previous few yrs.)

The costs of the Wisconsin archives are for media, student help to change tapes, the recorders, and maintenance:

- **Media Cost for One Satellite**
  
  $20,500/yr

  $2 * 365 days * $28/tape

- **People To Change Tapes**
  
  7,100

  3.5 hrs weekdays, more on other days

  1415 hr/yr, $5 each

- **Read Units (about $23K each, they last 7 years)**
  
  18,000

  About $6000 capital cost each unit per year, includes maintenance. About 3 units.

- **Assume About $10K Per Year For Other Costs and storage**
  
  10,000

**TOTAL For One Satellite** $55,600/yr

**TOTAL If Two Satellites, About**

(About 60 * 10^{12} bits of data)

$80,000/yr

---

15
DATA SUPPORT STAFF

FTE

Archive Activities 2.25
(acquire, prepare, update, catalog data)
General User Support 1.50
(send data, consult, etc.)
Direct Project Support and 0.40
Research Programming
Support NSF Radar Project 0.85
(special NSF grant)
Other Planning, Training, Research Tasks 1.00

Total FTE 6.00

------------- FTE -------------

DISCIPLINE USE OF FTE

<table>
<thead>
<tr>
<th>Discipline</th>
<th>1985,86</th>
<th>Late 87,88</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorology, Climate, etc</td>
<td>3.85</td>
<td>3.2</td>
</tr>
<tr>
<td>Both Meteorology, Oceanography</td>
<td>1.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Pure Ocean</td>
<td>0.2</td>
<td>1.1</td>
</tr>
<tr>
<td>NSF Radar</td>
<td>0.85</td>
<td>0.80</td>
</tr>
<tr>
<td>Total</td>
<td>6.00</td>
<td>6.00</td>
</tr>
</tbody>
</table>
DATA SUPPORT ARCHIVES AT NCAR

- One Of The Most Comprehensive Archives For Climate, Etc. In The World.
- Maintains, Adds, Updates Datasets. Now Over 300 Sets.
- Data From NOAA, NASA, Navy, USAF, ECMWF, Australia, Germany, India, Etc.

SELECTED DATASETS (10 sets of the 300+):

<table>
<thead>
<tr>
<th>Dataset Description</th>
<th>Volume (10^9 bits)</th>
<th>added each year (10^9 bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All global and hemispheric analyses and N. American fine-mesh (start 1963)</td>
<td>263</td>
<td>23</td>
</tr>
<tr>
<td>Global upper air observations starting 1963, surface starting 1976</td>
<td>296</td>
<td>24.3</td>
</tr>
<tr>
<td>E 70 million older global surface (4900 station decades), was 1000 tapes</td>
<td>93</td>
<td>none</td>
</tr>
<tr>
<td>World ocean ship data, 1854-1979 (72 million observations)</td>
<td>39.5</td>
<td>--</td>
</tr>
</tbody>
</table>

SAVED NOAA SATELLITE DATA:

- TOVS sounders (6.5 years) | 2094 |
- GAC 4 km global data, 5 chan | 6557 |
- NOAA 1 km data (ocean and land studies) | 1643 |

SUPPORT USERS: Send About 750 Tapes/Year; Most Use Is On-Line:

- Provide Access Programs;
- Help Them With Input/Output Problems;
- Provide Catalog Information;
- Provide Software Tools;
- Obtain New Data.
Fig. 1  Network of stations and location of regional divisions for rainfall analyses. Inset map: location of regions averaged to form rainfall departure series for Sahelo-Saharan, Sahel, Soudan and Soudano-Guinean zones and for northern and southern Kalahari.

1080 stns 18 from Sharon Nicholson usually 40-50 yrs. length
Figure 5. Number of days in a 7-day period in 1982 (18-24 April) where stations report precipitation (either the amount or "none"). All numbers should be 7 if adequate precipitation information is available.
NUMBER OF DAYS WITH RAOBS OR PIBALLS IN A WEEK

Figure 6. Coverage of NMC raobs and piballs for a week in 1982. NCAR archives of these data start March 1962 and became global in June 1966. Data receipt has gradually improved compared with the early years.
TRENDS

DATA MANAGEMENT

EXCHANGE MEDIA
TRENDS

• DATA WILL BE PREPARED OR GATHERED CENTRALLY:
  — But Also Distributed More to User Locations;
  — On Tapes, CDrom, Etc.

• EVEN MORE DATA USE WILL OCCUR AT THE CENTRAL COMPUTER.

• PEOPLE WILL OBTAIN SMALL DATA SUBSETS BY COMMUNICATIONS.

• SO NEED TO PREPARE SMALL DATA SUBSETS EASILY:
  — We Do This Now For Many Datasets;
  — But This Needs To Be More Automated For Some Uses.
DATA SET PROBLEMS

- NOBODY PUT SET INTO REASONABLE FORM.

- NOBODY GATHERED SEVERAL SIMILAR DATA SETS TOGETHER.

- SOMEBODY MERGED DATA TOO MUCH.

- NOBODY MADE SUBSETS OF HIGH VOLUME SETS.

ACCESS PROBLEMS

- USE AT HOME - WRITE UNPACK PROGRAM.

- USE ELSEWHERE:
  — Learn System Software;
  — Learn Rules for Data Language;
  — Should Have Simple Rules for Record and Fortran Interface.
AVOID TRAPS

- TIGHTLY COUPLED SYSTEMS:
  — The Data Must Be In System Format To Use It.

- ALL DATA ARE VERY VALUABLE AND NEED EQUAL CARE:
  — Little Used Or Less Critical Data Should Be Handled Differently.

- A FULL McIDAS TYPE DISPLAY SHOULD BE AVAILABLE FOR ALL DATA:

- USE A FORMAL DBMS, EVEN FOR LARGE DATASETS:
  — Timing Is Lousy.

- OFFER UNLIMITED FORMAT FLEXIBILITY TO USERS:
  — Takes Too Much Time;
  — A Compromise Will Work Well.

- WE DON’T HAVE TO GATHER AND PREPARE DATASETS, BECAUSE COMMUNICATION CAN GET THE PIECES WHEN NEEDED. (No)

- OTHER
  — The Communications Era Adds a New Set of Problems. (Yes)
  — AI Must be Developed to Have a Decent Data System. (No)
DATA MANAGEMENT

PROBLEMS THAT ARE OFTEN DISCUSSED (& Funded):

- Set Up Communications;
- Set Up Data Catalog Systems;
- Set Up Data Display Systems.
- Set up mass storage systems

PROBLEMS OFTEN IGNORED (The Data):

- Gathering Global Observations, So That 40+ Years Of Data Can Be Used;
  - such as ship data
  - gathering is also needed for raobs, satellite winds, aircraft, etc.
- Data Are Not Prepared To Study The US Dust Bowl Days;
- Copy Good Data From Old Tapes Or Lose It;
- Paper Is Crumbling, Should Microfilm Some Of It.

NOTE: R. Corell at NSF has noted this problem of balance.
WE NEED STANDARD EXCHANGE MEDIA

- The whole world can read half inch tapes;
- 3480 cartridges may become a standard;
- Optical media do not have standards;

- We must have 2 or 3 types of standard media, but:
  - The market for half inch technology is decreasing
  - Another standard has not emerged

- With a $5K micro, people don’t pay $15K for I/O.
FOR DATA
WHAT IS NEEDED FROM COMPANIES

- Read a variable length record and obtain status;
- Subroutine Gbytes to handle binary data.

DATA PACKAGING

WE NEED A FEW STANDARDS:
- Handle variable length records;
- Have a checksum on data blocks;
- Handle a "volume" that can have several logical files.

NOTES:
- IBM handles variable length;
  — No checksum
  — No imbedded file marks
- CRAY-block is close to what is needed.
ARCHIVING

COST AND STRATEGY
<table>
<thead>
<tr>
<th>Year</th>
<th>Effective Cap. (Gbit)</th>
<th>----- Media Costs -----</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Per Gbyte</td>
<td>Per Tbit</td>
</tr>
<tr>
<td>1960</td>
<td>556 BPI tape ($14), 10 mBytes</td>
<td>0.08</td>
<td>$1,400</td>
</tr>
<tr>
<td>1972</td>
<td>1600 BPI tape ($14), 42 mBytes, holds 37.5 mBytes</td>
<td>0.30</td>
<td>374</td>
</tr>
<tr>
<td>1980</td>
<td>6250 BPI tape ($13), 150 mBytes, holds 125 mBytes</td>
<td>1.00</td>
<td>104</td>
</tr>
<tr>
<td>Feb</td>
<td>IBM 3480 tape cart ($14), 200 mBytes, holds 174 mBytes</td>
<td>1.39</td>
<td>81</td>
</tr>
<tr>
<td>Feb</td>
<td>3480, $10.50 each</td>
<td>1.39</td>
<td>60.5</td>
</tr>
<tr>
<td>June</td>
<td>3480, $7.50 each</td>
<td>1.39</td>
<td>43.2</td>
</tr>
<tr>
<td>Late</td>
<td>Honeywell VHS Cassette holds 5200 mBytes, costs $8.50</td>
<td>40.5</td>
<td>1.7</td>
</tr>
</tbody>
</table>
Figure 3: This figure was prepared by the Office of Space Tracking and Data Systems in July 1986.

DAILY DATA VOLUME TIME PROFILE

July 1986 Plans

20 Tbits/DAY
(= 231 Mbits/SEC)

YEAR

1998
1997
1996
1995
1994
1993
1992

TELEMETRY

0.16
0.15
0.14
0.13
0.12
0.11
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00

TOTAL DAILY DATA VOL (TBYTES)
NCAR OCEAN MODELS

- ATLANTIC, 15° S - 65° N
- 120 ARCHIVE STEPS PER YEAR

<table>
<thead>
<tr>
<th>Levels</th>
<th>Track</th>
<th>Each Step (10^6)</th>
<th>Each 5 Yrs (10^9)</th>
<th>Bits per 5 Yrs (10^9)</th>
<th>XMP CPU Hr Per 5 Yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>T, Salin</td>
<td>5.922</td>
<td>3.55</td>
<td>227</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>T, Salin</td>
<td>8.583</td>
<td>5.15</td>
<td>330</td>
<td>300</td>
</tr>
<tr>
<td>30</td>
<td>T, Salin</td>
<td>11.782</td>
<td>7.07</td>
<td>452</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+2 Tracers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 1/3 Degree Resolution
NASA ARCHIVES

THERE ARE 100,000 TAPES AT NSSDC (1987):

- Estimate that 20 terabits of data on these;
- These have most NASA data saved to now;
- Not enough resources to save it all.

IN 1995:

- NASA plans to save 15 tbits/day.
# ACCESS RATE TO DATASETS

<table>
<thead>
<tr>
<th></th>
<th>Total Datasets</th>
<th>Total Accesses Per Day</th>
<th>Accesses Per Day to 97% of Archive</th>
<th>Average Datasets Read Per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984-85(2C-1)</td>
<td>105,000</td>
<td>950 (570 read)</td>
<td>163 (87)</td>
<td>30%</td>
</tr>
<tr>
<td>1987-88(XMP,C-1)</td>
<td>260,000</td>
<td>1900 (1140 read)</td>
<td>315 (163)</td>
<td>23%</td>
</tr>
</tbody>
</table>

## RATE OF USE OF CLIMATE MODEL DATASETS (Estimate):
- Use data 1.5 times during year one;
- Use another 1.5 times during years 2-7;
- Assume the archive is kept 7 years;
- Part that is used per year: 43%.

## NOTE: RATE OF USE OF AN OLD BOOK:
- Once each 100 to 200 years.
INVESTMENT COST FOR DATA STORAGE

<table>
<thead>
<tr>
<th></th>
<th>Media Cost/Tbit</th>
<th>Media Plus Robotics Cost/Tbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM 3480 Cartridge Tapes</td>
<td>$5,400</td>
<td>$88,000</td>
</tr>
<tr>
<td>(175 mbytes each [of 200])</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VHS Tapes</td>
<td>$210</td>
<td>$18,500</td>
</tr>
<tr>
<td>(6750 mbytes [of 7000])</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

EXAMPLE:
- Investment Cost Is $88 Per $10^9$ Bits (7-year life);
- It Is Used Once Each Three Years;
- Then The Storage Cost Per Access Is $50; If Data Isn’t Used For 3 years, No Harm To Wait 5 Minutes.

ASSUME FOR ROBOTICS:
- 3480: $700,000 For 6,000 Cartridge Unit, Holds 8.4 Tbits;
- VHS: $600,000 For 600 Cartridge Unit, Holds 32.4 Tbits.

NOTE: Optical Media Usually Costs 2 to 5 Times The Cost Of 3480 Cartridges.
CATALOG EFFORTS

ENCOURAGE OTHER CATALOG EFFORTS:
• Are Working With NASA;
• Have Worked With WMO to Define System.

WE HAVE PREPARED INFORMATION ABOUT CATALOG STRATEGY.

MAINTAIN A DATA CATALOG
• Update It Each Year;
• Refer to Other Catalogs.

CONCENTRATE ON CATALOG INFORMATION RATHER THAN ON PARTICULAR ACCESS SYSTEMS:
• Give More Emphasis to Scientific Aspects of the Data Than Others Do.

PREPARE INFORMATION ABOUT:
• Sources of Real-Time Data (done);
• Data for Applications (book is done);
• Other NCAR Data;
• Boulder Local Area Data.
NCAR PLANS

- MAINTAIN THE ARCHIVE AND SUPPORT USERS:
  — Update Many Of 300 Datasets;
  — Fill Many Requests For Data;
  — Takes Nearly Four FTE.

- PROVIDE MORE DATA SUPPORT FOR TOGA.

- PROVIDE DATA SUPPORT FOR OCEAN MODELING:
  — Make Ocean Satellite Data Easy For Model Use;
  — Distribute Model Data.

- ADD DATA FOR BOUNDARY-LAYER RESEARCH AND SURFACE PROCESSES.

- IMPROVE THE ARCHIVES OF SATELLITE DATA AT NCAR.

- ARCHIVE NSF INCOHERENT SCATTER RADAR DATA AND CEDAR DATA.

- GIVE DATA SUPPORT FOR CLIMATE ASSESSMENTS AND MODEL DEVELOPMENT.
SELECTED TASKS

- Have Delayed Update On African Precip For 18 Months.

- Ships Are Thru 1979.
  — Badly Need An Update
  — Have Delayed 18 mo To Save Satellite Data

- People Need SFC And Upper Air Data For Africa.

- Need a Better Set Of Raob Time Series Data.

- Need Time Series Of Daily Solar & Streamflow For Dickinson.
  — Delayed 18 mo

- Make Aircraft Dataset For US-India And TOGA, Etc.

- Need To Keep The US-USSR Exchange Going.

- People Need More Info As In Our 1975 Tech Note re Datasets.

- Need US Data For The 1930s.