Monthly Reports, DSS, 1992 – 94

- Year 1992 (reports for 11 months here)
- Year 1993 (reports for only 6 months here)
- Year 1994 (all 12 months here)
- Year 1995 (all of 1995 is in another text, 3 mos are here)

Some comments:
- Tough times in Russian data lab.
- NRC study about “National Archives.”
- Feb 1994: Work to prepare observations for reanalysis is “hot and heavy” in DSS.
- Nov 1994: Japan will probably need observations for reanalysis.

- Ready for scan Mar 4, 2003, 76 p, RJ0273

Roy Jenne
03/03/03
(These reports are online in RJ0273, 76 p)

Roy Jenne
Mar 03, 2003

1992 (11 months here), July 1992 missing ......................................................... p 1
* Nov 1992: Feelings are running high at NASA-Goddard about formats, EOSDIS, etc.
  • “Who will stop them now? Who will stop EOSDIS? It is like a black hole out of control.”

1993 (6 months here), only Jan, Feb, Apr, May, Jul, Aug 1993 are here ....... p 32
* Jan 1993: It is hard to even pay the electricity bills at the data lab in Russia.
* Tony Hollingsworth, ECMWF, visited NCAR to talk about reanalysis and the data that they need (the ERA-15 project).
* Apr 1993: NAS study about national archives.
* Apr 1993: Status of reanalysis observations (huge tasks).
* Apr 1993: Connects to DSS online system.
* Aug 1993: Trip to Uruguay
* Aug 1993: Tough changes in data lab near Moscow.

1994 (all 12 months here) ................................................................................. p 47
* Feb 1994: The work to prepare observations for reanalysis is “hot and heavy” in our data support unit. Gives a partial list of tasks.

  NOTE: Production of the NCEP/NCAR Reanalysis started June 1994 at NCEP.

* Mar 1994: Some data policy issues. I do not want the observations for reanalysis to be restricted. Europe invited me and one other from WMO to their meeting. They wanted to help with data, but then restrict the data.
* Nov 1994: Japan probably will do a reanalysis and will need data from us.
* Dec 1994: NRC archives panel. Difficult. I will write a minority report. It is good to use PI teams for data, but we can’t run a big data system with only PI teams. I am in the middle; I do not like to support monster data systems either.
DATA SUPPORT SECTION

Monthly Status Report
January 1992

Dataset Updates, Additions, and Requests (all)

Data sets updated during January and the most current dates available are:

NMC surface and upper air station data, 4 January 1992.
Navy surface and upper air station data, 31 December 1991.
World monthly upper air climatology through 1990 (ds430.0).
TDL U.S. and Canada surface hourlies (ds472.0), July, 1991
The CI data set (ds742.0), several gaps filled.
The ERBE data set (ds732.0) has been updated to May 1987.
NMC Medium Range Forecast 10 day forecasts, 31 December 1991.
NASA Global Tropo Temp Analyses from MSU data(ds701.0), Dec 1991.

An update of ds067.1 (Gelman’s CAC 65x65 N. and S. Hemisphere Analyses) was begun, which would have updated the dataset through 5 October 91, but several checksum errors were found as well as two grids labeled with the wrong year. A new tape has been ordered.

The Navy surface and upper air station data (SPOT data) for the period Oct 1985 through Dec 1991 was split out to surface, upper air, and satellite (plus miscellaneous) subsets. The entire period of record (1971-1991) was inventoried and the results saved on the MSS. Copied data (backups and originals) on 173 C-tapes to Y-volumes on the MSS.

A new data set has been created from the Diaz monthly precipitation data. The data set number is ds238.0. The data set contains 4 files which contain the monthly precipitation amounts for 1951-1989, monthly precipitation anomalies for 1851-1989, DOE precipitation station data and the monthly precipitation climatology. The data is on a 4 degree latitude by 5 degree longitude grid. To create the gridded data 5,562 stations were utilized. There have already been 2 requests for this data set.

There were 29 orders completed during January.

General Archive Activity (all)

Discussions and testing of improvements and new options for our internal database are ongoing. Modified the code for the new charging algorithm to deal with unanticipated problems. Updated the datahelp version of the Air Force Dictionary.
New codes were written to provide character version for data in TD14 packed binary format (ds470.0), and to select/compress data records in TD3280 format (ds470.0).

NMC is planning to ship the MRF data to us in GRIB format, starting sometime early this year. A few of the November grids in the flux archive (DS084.5) are already in this format. A subroutine to decode the NMC GRIB format (which is slightly different from the ECMWF GRIB format) was received via FTP from NMC.

The access code for DS084.5 was modified to check record lengths to determine whether a record is in GRIB or the old format, and to use the appropriate decoding subroutine. A new document for this dataset was written, which describes both the old and the new formats. This document will be sent out with the dataset, and is also available in the new DSS FTP area.

NMC also sent routines to convert from spherical harmonic coordinates to grid point space for the 10-day forecast dataset. Work continues on adapting these routines for use with the sigma level analyses.

Current copies of the access programs for the ECMWF TOGA datasets (DS111.0, .1, .2, and .3) were put into the FTP area. It was decided to make a different version of the main program for each dataset, with a single copy of the decoding subroutines which was then linked to a directory entry in each dataset's directory in the FTP area. Other documentation for these datasets was also placed into this area, along with some sample data for the most popular of these (DS111.2).

A short article on the DSS FTP area was written and distributed to the Internet newsgroup sci.geo.meteorology and to the CLIMLIST mailing list. These are electronic forums devoted to the discussion of meteorological and climatological research. We have already received some positive comments, and a few requests for data. This has helped spread the word about NCAR's data support facilities to scientists who might not otherwise know of them.

Scientific Support

Time series of coop data for period of record for Phoenix (ds510.0) was derived for Mary Downton. A character version for Australian monthly means of sea level pressure and 500mb heights and temperatures (ds195.1) was set up for Harry van Loon to take to New Zealand. For other NCAR scientists we derived long term means and harmonic analyses, and monthly means of height, temperature, and u-wind component for all stations south of 35 degrees south in the world monthly upper air climatology (ds430.0); and sea level pressure for all stations south of 35 degrees south in the world monthly surface climatology (ds570.0).

Unidata System Support

Considerable extra effort was put into the Unidata system in support of STORM-FEST. Arrangements were made for backup systems and backup data sources, and problems with system reliability were debugged.

DSS is also supporting the ACD Hawaii Field Experiment, January 15 to February 15, 1992. They are using the weather maps generated by wetterhorn every day. The maps are faxed to Hawaii daily. Peter Hess is using the data stored on MSS for 10-day trajectory analysis.
Anonymous ftp Area

The information system available via anonymous ftp is growing rapidly as is the delivery of data files by ftp. All master data set documents are now available and much of the software and supporting documentation is being put on the system.

Incoherent Scatter Radar Project (Barnes)

As of 11 Feb, the programs (print, time series plot, height profile plot, height versus time averaging and plot, and antenna scan plot) were installed on Shavano and Bierstadt. A common library of subroutines was established for subroutines. The work is not quite complete though. During the course of the install on the Cray, I uncovered a cft77 compiler bug, but it turns out the bug has been fixed in the new compiler. One request for database and access information from Brazil was answered.

Problem Areas

As more users move to Unix based systems, they are having more difficulty handling the basic I/O functions and external media. Interestingly, many of these problems have come from Cray users. A small code and test data set which read on Shavano would not run on the San Diego Cray.

The Exabyte tape drive is becoming a serious bottleneck in data delivery. Many users are requesting data on Exabyte, but turn around on copying data from the MSS to Exabyte can often be several days and failures are common. DSS is encouraging users for moderate sized orders to accept 1/2" tape when this is an option for them. Writing data to multiple Exabyte tapes even when the data will fit on one is sometimes more successful.

Meetings

Roy Jenne, Dennis Joseph, Joey Comeaux, and Chi-Fan Shih attended the annual AMS meeting in Atlanta. Chi-Fan presented a paper at the meeting. The special CDROM meetings were held at this time.

Comments by R. Jenne

1. CD-ROM MEETING IN ATLANTA (5-8 January): The first night of UCAR-sponsored CD-ROM Meetings (6 Jan. 1992, 2.5 hours) there were ten people around the table and twenty more in the room. The second night the total population was about eight. The first night was fairly "heavy" since a number of issues are involved. Cliff Mass put off the thorniest of the subjects, which was "formats" until day two, and that was more peaceful than I expected. The main decisions:

   • We will start with three main data series.
   • We talked about prices of about $50 per disk. This is the NOAA charge, NASA usually charges about $10 to $20 for one.
   • The data on the CD-ROM will be in a compact format; probably we will use GRIB for recent analyses. There will be an option to convert data to NetCDF on the output.
In late November and December, I gathered more information about the format question. There are many large sets of data that NetCDF does not handle well. Sandy McDonald, NOAA-Boulder, was there and noted the same trouble with the handling of rawinsonde data that we saw. Unidata may work on ways to handle compact formats, which must include the ability to handle variable length data.

2. COADS WORKSHOP IN BOULDER (13-15 January):
   - Gave talk on the reanalysis project, including COADS. People seemed very interested in this subject.
   - Chair a session with eight speakers.
   - Wrote an abstract, most of a paper, and the final session summary.
   - Steve Worley gave a talk on various types of data in COADS.

3. WORLD TEMPERATURE WORKSHOP (16 January): Got good information about several projects that created datasets, and the results of scientific comparisons between datasets.

4. MEETING REGARDING NCAR SPEC (16 January)

5. REANALYSIS: NMC called - need budget information and sample data.

6. EXECUTIVE MEETING (Friday, 17 January): Three secretarial interviews, visitors from UK and Australia, phone calls, documents (a 12-hour day). Last day for Karen; she has been very helpful.

7. NAS DATA FORUM: A US interagency data plan has been in preparation for three or four years. At least two attempts failed. In the last year George Ludwig brought together the present version of the plan which will probably pass the tests, after a number of changes. The disciplines of "Human Science" are apparently coming into the plan with full member standing. If a good bandwagon like global change starts, everyone tries to get on board. I am on the Academy panel for Geophysical Data that has been talking about the data problems for some years. This group oversees the US planning process. About ninety people were at the meeting (Washinton, DC).

8. NATIONAL ARCHIVES PLANNING MEETING: On January 29-31 I attended a NAS National Archives Planning Meeting. The Archives (NARA) has responsibility for long term (50-year scale) national archives. They have been getting hit to worry more about digital data. The talks included NASA, NOAA, DOE, NIST (materials data), DOD, LBL, myself, etc. We devised a strategy to make a study of the problem. Our committee had about ten members and there were about 45 people at the meeting (Washington, DC).
DATA SUPPORT SECTION

Monthly Status Report
February 1992

Dataset Updates, Additions, and Requests

Data sets updated during February and the most current dates available are:
- ds018.0 (U.S. Navy N. Hem Sea-Level Press & 500mb Ht, daily) --
  navy sea-level pressure was updated thru December, 1991
- ds066.0 (NMC 65x65 N. Hem Tropo Analys) --
  lists A0 (all 47x51 grids), A1 (47x51 heights, temps, sea-level
  pressure and surface temp), A2 (47x51 grids not in A1 or A3
  including tropical grids), and A3 (47x51 forecast grids)
  were updated thru December, 1991
- ds067.1 (Gelman’s CAC 65x65 N. & S. Hem Analys, daily) --
  update from Gelman was added for the period 1991Mar24-1991Oct05
- ds085.0 (N. Hem 47x51 Time Series Tropo Analys, monthly) --
  navy sea-level pressure was updated thru December, 1991
- ds085.1 (N. Hem 72x19 Time Series Tropo Analys, monthly) --
  navy sea-level pressure was updated thru December, 1991
- ds195.0 (N. Hem 47x51 Time Series Tropo Analys) --
  navy sea-level pressure was updated thru December, 1991
  500mb heights and temperatures were updated thru December, 1991
- ds195.1 (N. Hem 72x19 Time Series Tropo Analys) --
  navy sea-level pressure was updated thru December, 1991
  500mb heights and temperatures were updated thru December, 1991
- NMC limited fine mesh model outputs updated through 91/11/30
- NMC nested grid model outputs updated through 92/01/18
- U.S. first order summary of day through December 1990 (ds509.0).
- Canadian time series raobs through April 1991 (ds390.0).
- World monthly surface station climatology through December 1990
- Daily lat-lon time series for 700mb and 500mb u and v through
  June 1991 (195.1).
- The C1 data set (ds742.0) for the ISCCP, several gaps filled.
- Several data gaps of the ERBE data set (ds732.0) were filled.
- NMC Medium Range Forecast 10 day forecasts, 30 January 1991.
- The Drifting Buoy Data Set prepared by MEDS. Add the 1991 data.

A new data set has been created from the NOAA GPCP pentad precipitation estimates. The data set contains 3 files which contain the pentad precipitation amounts for 1986-1991, Psuedo estimates for each month and Changs monthly estimates. The Chang's monthly estimates file has some problems with the record length and the total chronological length of this data is also
shorter than they reported. Both of these problems are being worked on. The first 2 files of data appear to be fine. The GPCP data is on a 2.5 degree grid for the domain 38.75N to 38.75S and 1.25E to 1.25 W. Chang’s estimates are on a 5 degree grid for the domain 50N to 50S and 0 to 5 W.

There were 35 data requests completed during February. The World Monthly Surface Climatology and the time series raob files were the most popular sets.

General Archive Activity

The codes provided with the ISCCP data were improved and made more portable. These new codes were tested and stored under the new anonymous ftp. The Navy SPOT and Navy grid applications programs were converted to UNICOS.

Documentation for the NMC upper air data (NMC Office Note 29, plus collected notes) was transcribed and put in an on-line file.

The task of backing up the primary files in ds698.0 in an effort to remove 1/2 inch tapes from the archive was begun. Thus far, 30 primaries have been backed up.

The South African raob data from the latest update (1978-1989) has now been put into packed-binary format. The next steps are to sort the raobs and then inventory them.

Reanalysis Project

As part of data preparation for reanalysis, DSS has begun making more extensive inventories of the NMC ADP data and examining aircraft data from various sources.

A new summary of NMC ADP upper air data is being made and nearly 400 of an eventual 726 inventories have been made. New codes to summarize these reports has been written.

Data were dumped from the USAF DATSAV tapes for the period 1976-1985, as per Roy’s request. This data has now been inventoried and sample printouts also exist.

The aircraft data tapes from GASP (NASA Global Atmospheric Sampling Program) for the period 1975-1983 were imported, and inventories and sample printouts were done for these files. Available data fields include temperature, dewpoint, pressure, altitude, wind direction and speed. For the period 1975-1978, 397668 observations were found in 5645 flights of five aircraft.

DSS ftp Information System

Files are continuous added to the collection in the ftp area and users are making extensive use of this service.

A small subset of ds111.2, the most popular ECMWF dataset, was extracted from the January 1990 data. The binary data, the output resulting from running the 'access.f' read program, and the metafile containing the resulting plots were placed in a subdirectory in the FTP area. Some users have already reported that this has been very useful in helping them modify the read code for use on their machines.

More successes and satisfaction with the DSS FTP facility:

---
>From Roy Mendelsohn at the Pacific Fisheries Environmental Group.

"we got 1978 with no problem.... This is really a fantastic and mind-boggling experience, having worked on computers since 1967. First, the amount of data being transferred in this manner, and second that we will be able to access it on a desktop PC.

I just want to say that I think your guys approach of making the data available in a decentralized manner is the right one, and conflicts with much of the trend in NOAA.

Thanks again, Roy"

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

Our anonymous ftp directory system on huron is not only proving useful to scientific researchers, but also for education of science students. One of the satisfied clients comments are given below.

Mark A. Horrell Illinois Mathematics and Science Academy 1500 West Sullivan Road Aurora, IL 60506-1039

"I retrieved ds270.2 and ds750.1. The latter looks terrific on Image and Spyglass. My students are playing with changing sea level and will use it to look at large-scale features of the ocean floor. ...."

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

UNIDATA System and STORM-FEST

During February STORM-FEST access to data from the UNIDATA system went fairly smoothly. There were a few glitches but nothing serious according to Phil Haagenson. The backup system was available but was not used. Backup data sources from University of Wyoming and Purdue were not needed.

ACD Hawaii Field Experiment ended on 15 February and the map generation was stopped at the end of the month.

SCD Computing News

An article was written describing the Data Support Section's archive of terrain and vegetation datasets. This is one of a series of articles highlighting some frequently-used datasets. We have gotten a lot of positive feedback from these articles.

Kuwait Fires Project

Code was written and tested to read the NMC sigma level dataset, convert it to gridded form, extract a subgrid over the Persian Gulf area, and write it out in NMC Office Note 84 format. Upper Air and surface data were extracted for a window around the Persian Gulf. These files
were sorted into time series order and both synoptic and time series order were made available to
the Kuwait Archive Group.

Comments by R. Jenne

DATA FOR ARCTIC ICE ISLANDS

We need to obtain as much of the arctic upper-air data as is possible for reanalysis. Jon
Kahl, University of Wisconsin, is helping to prepare data. We have also given him data. Talked
with Jon on 3 February.

PHILOSOPHY OF DATA MANAGEMENT

Dick Davis at Asheville is on the interagency format standards committee. I gave him a
copy of the "Belgium Paper" (May 1988, in 1989 book) that traces data management history and
popular (not necessarily sensible) ideas over a 20-year period. He liked the report and will pass
it around the center.

METHODS TO DESCRIBE DATA FORMATS

Don Mock and Jim Welch (NOAA, Boulder) visited about ways to describe formats so that
computer systems can automatically use the data regardless of the format. Jim wants me to give
him more time on this issue but I don’t have more time now.

DEMOnSTRATION OF INO DATA SYSTEM (Feb. 26)

The core part of this system is the Neons display system developed in Monterey. The internal
formats are the WMO standards, Bufr and Grib. It was a fairly interesting demonstration.

FORMATS

I sent NASA some information about the format questions. This issue always brews.

OCEAN CLIMATE DATA WORKSHOP (February 18-21)

An ocean data workshop was held at Goddard with about 100 people attending. A mix of
science, and data people attended. Top officials from agencies, and international organizations
spoke. The purpose was to help the planning for data systems for ocean work. This seemed to
have useful results. Several display systems were demonstrated.

THE US SOLAR NETWORK

The US solar stations were planned to close down 1 July 92 because of data problems. I
attended a February 25 NOAA meeting in Silver Springs to try to keep the net going. Our inabili-
ity to keep these stations going, (and with good data) is a disgrace to the US and an embarras-
sement in the US-Russian exchange. About 15% of the data is lost before it gets to an archive.
I have the feeling, "If Canada can do it, why can’t we?" There are ways to mostly fix these prob-
lems, at low-cost. I hope it happens.

PROBLEMS WITH LEVELS IN CANADIAN RAOBS
There are probably 50,000 or more cases where two levels separated by 0.5 mb have the same height, which is impossible. We are working with Canada to resolve this problem and one other problem in the surface levels. Credit to Will Spangler and his good diagnostic checks.

REANALYSIS PROJECT (Feb 28)

Kanamitsu visited about the reanalysis. They need detailed inventories (each 6-hours) for the NMC upper-air data since 1973. We sent the last 6-years within a week.

NMC would really like three months of all data in 1985 ready by the last of March. I do not think that this is possible; data for ten years is almost as easy as data for three months.

ARCHIVE OF ANALYSES FROM GODDARD

Mike Seablom visited from Goddard on February 14. They archived some older model runs here and are interested in sending a lot of newer data. We will probably agree on a subset that will have tolerable volume. Also, they will probably put it into Grib format for us to help control the volume.
DATA SUPPORT SECTION
Monthly Status Report
March 1992

Dataset Updates, Additions, and Requests

Data sets updated during March and the most current dates available are:
- U.S. first order summary of day through December 1990 (ds509.0).
- Canadian time series raobs through April 1991 (ds390.0).
- World monthly surface station climatology through December 1990 (ds570.0).
- Daily lat-lon time series for 700mb and 500mb u and v through June 1991 (195.1).
- NCAR N. Hemisphere Daily/Monthly Sea-Level Pressure, 31 December 1991
- Trenberth's N. Hemisphere Daily Sea-Level Pressure, 31 December 1991
- Gelman's 65x65 Grid Daily Analyses, 15 February 1992
- 47x51 N. Hemisphere Monthly Analyses, 31 December 1991
- 47x51 N. Hemisphere Daily Analyses, 31 December 1991
- 72x19 N. Hemisphere Daily Analyses, 31 December 1991
- U.S. Navy N. Hemisphere Daily SST Analyses, 31 December 1991
- U.S. and Canadian Hourly Station data, November, 1991
- NOAA GPCP Pentad Estimates of precip, December, 1991
- Satellite Cloud Climatology Data (ISCCP), December 31, 1989.

The ECMWF FGGE archive was cleaned and tapes removed. The permanent weather ship archive was augmented with UNIX ASCII files for easier access.

Several problems with recent NMC updates have delayed the update of this archive. NMC will try to fix the problems when we identify the missing dates.

General Archive Activity

The task of backing up the primary files in ds698.0 in an effort to remove 1/2 inch tapes from the archive continues. This month, 424 primaries were backed up for a total of 454 to date.

More South African raob data were put into packed-binary format. The period of record is now 1968-1989 (1968-1977 were added). The data are currently being run through a hydrostatic check. Problems have been encountered, but at least initially, it looks like the majority of the trouble is due to data formats and documentation that do not agree, and this can be dealt with.

The NMC flux archive is now being received in GRIB format. Because NMC uses extremely long blockizes, code was written to split the GRIB records into several smaller records. This increases the portability of this dataset, since many computers can't handle large blockizes. Code to read this format using routines supplied by NMC was written. Code to select specific grids and write them to disk was written.
A routine to output data records in GRIB format was written. This will be useful in filling requests for subgrids of global datasets in GRIB format, such as the ECMWF/TOGA data and the NMC flux data.

A program to extract particular parameters from Kevin Trenberth’s ECMWF/WMO long-term means and statistics dataset (ds110.3) and write them to disk in ASCII was written. We have had several requests for small subsets of this dataset.

Work continued on the global surface station set inventory (TD13). Edits to the master inventory were completed and it was combined into one file. An inventory plot by year program was built and used to produce continental/political outline background maps of average number of observations per day for each year of data. Some station location information remains to be key entered.

UNIDATA

STORM-FEST project ended on March 15, 1992. There may be another experiment similar to the scale of STORMFEST next year for WISP program according to Phil Haagenson.

Reanalysis Project

A new summary of NMC ADP upper air data is being prepared. The first pass through 726 volumes to put inventories on-line is completed. The inventory program produces additional diagnostic information, including "trouble" reports that flag potential problems.

COADS Ship Project

The DSS in conjunction with the Environmental Research Laboratory of NOAA has processed a limited amount of data for the COADS update project. In order to meet the needs of reanalysis projects that are beginning at the University of Maryland data from 1982-1983 are required. To meet this need 26.6 million ship and buoy reports were translated into COADS binary format from which 9.5 million reports were extracted for the 1982-1983 period. This subset was sorted by geographical location and prepared for further duplicate report elimination data processing. Following duplicate elimination 4.2 million reports remained. This small dataset will be used by the University of Maryland with another key element they require to start their global reanalysis project.

GrADS Demonstration

The DSS hosted a NCAR Lecture/Demonstration of a data display system. Jim Kinter and Brian Doty from the University of Maryland, Center for Ocean-Land-Atmosphere Interactions presented, "The Grid Analysis and Display System (GrADS): A System for Data Manipulation and Data Display". The demonstration was run a Sparc workstation with two slave monitors illustrating the system to the audience. The system was impressive and the attendees showed enough interest so that the DSS decided to make it available to the public. Execution of GrADS is now available to all persons connected to the SCD file server (crestone.scd.ucar.edu). Furthermore, documentation, and sample files for running GrADS are available on the DSS anonymous ftp machine (ncardata.ucar.edu). Brian Doty has volunteered to help other NCAR divisions, or any other interested parties for that matter, install GrADS on their own computing systems. Doty’s offer is noted in the documentation on the DSS anonymous ftp machine.
Kuwait Fires Project

Code was written and tested to read the NMC flux dataset, convert it to gridded form, extract a subgrid over the Persian Gulf area, and write it out in NMC Office Note 84 format. A similar program was prepared to do the same to the NMC globals. The output grids of flux and sigma level parameters were rotated to match the orientation of the global grids.

Incoherent Scatter Radar Project (Barnes)

One request went to UK, another for information and potential login from Mexico. The annual inventory update was started in preparation for generating the catalogue. One contributor will eventually use Exabyte, but has ftp'd up to 40MB files onto "upload", no more than one file per day for some 15 days to get data to me in time to include in the catalogue. Electronic transfer has been more inconvenient than traditional so far. Comments by R. Jenne

DAILY WORLD DATA (7500 STATIONS)

We have a project with CAC and two other centers to prepare a better set of world daily data. For two months we have needed an update with the past two years of data. CAC has been struggling with their VAX and Exabyte system to get us a copy of the data. They have too many competing priorities. It is now too late to get the update into a CGD Tech. Note. Frustrating!!

MONTHLY OCEAN PRECIPITATION DATA FOR 1979-91

People need precipitation data. By end March, the NASA Marshall group (Spencer and Christie) hopes to have monthly 10-year means for 1982-91, and anomalies for 1979-91. This is calculated from NOAA satellite microwave data. NCAR sent them a lot of basic data; now they get updates from NOAA.

ARCTIC ICE BUOY DATA

We had a meeting in Boulder to coordinate activities across the country to produce the best dataset of Arctic Ice Buoy data. NCAR, NOAA/ERL, NOAA/(Snow & Ice), and the University of Washington were represented. There is an array of automatic stations sitting on pack ice that are read out by satellites. There have been manned Russian ice islands for many years. We prepared notes about what data are available, and plans for new datasets. This data will be a part of COADS. It mainly describes pressure and temperature over the Arctic Ocean.

CREDIT CARDS TO PURCHASE DATASETS

A PI from the NOAA Air Resources Laboratory in Washington called (March 10) to find where he could find data about world soils. We have two datasets that should solve his problem - cost $50. He asked if he could use a government credit card to pay the bill. He could do that directly and without paperwork hassle. NCAR can't handle this yet - so NOAA has to make out a purchase order. I have the impression that this procedure used about two hours of their time, and two extra telephone calls to me. I called George Meneghini. NCAR has been thinking about allowing credit card purchases but it did not sound as if anything could happen soon. Interest in the idea seemed to increase during our conversation. He said that the paperwork for NCAR would be reduced if we permitted credit card purchases. The paperwork would certainly be reduced for the purchaser.
GAO REPORT ON EOSDIS

This report was released about February 25, and was written about in Science (6 March 92) issue. I finally read it on March 8. It seems to me that NASA is criticized for not building a super-fancy system. NASA scientists have been trying to achieve a good basic system and one that will work. I vote with them. I wrote a review of the GAO report.

DEMONSTRATION OF DISPLAY SYSTEM

Jim Kinter and Brian Doty (University of Maryland), demonstrated their Grid Analysis and Display System (GrADS) at NCAR on March 25. Many functions and displays are already built into the system. During the next four months, they plan to add a GUI interface. The lecture and demo was attended by 38 people. I suggested that they prepare a video tape. Kinter likes the idea; they have the facilities to do it.

COADS SHIP PROJECT

The University of Maryland wanted to start a 1982-1983 reanalysis about April 1. NMC has needed test data for 1985 reanalysis for three months. During March we have been working very hard to overcome a lot of ship and buoy problems, and to get conversions and sorts into production. We hope to have it ready by about April 7.

— Three people are full-time on this in NOAA/Boulder.

— In March we used the equivalent of 1.5 FTE at NCAR.

— One person half-time in Asheville.

Gregg has been doing a lot of work to handle special cases and non-standard events in NMC upper-air (UA) data in order to make new UA inventories.

CD-ROMS

This project is coming along, but certainly not fast enough yet. Part of the problem are all the other demands on my time. There are a few data inputs for which I need to do some background work.

WORLD 1-KM ELEVATION DATA

We have elevation data spaced five minutes or ten minutes over the world-mostly ten minutes (about 18 km). We have 1 km data over the US. People need 1 km data over the world. I made several phone calls to get the status of possible new datasets from several sources. A four-page text is now available. (The particular recent request was for a project in Mexico).
April 1992

TAUGHT A CLASS AT CSU

CSU (Department of Meteorology) has a course that includes a variety of activities such as a tour of Weather Service forecast offices. This year they wanted to add a lecture about data—what is available and how to obtain it. I talked to this class of 25 people on April 13.

STATUS OF REANALYSIS PROJECTS

NOAA taxed several groups (GFDL and NMC) to raise $50K to support data help by Ernie Kung, University of Missouri. This has been added to Jenne's reanalysis grant. Pete and others (NCAR and UCAR) have figured out a way to send the $50K to the University of Missouri without NCAR overheads (still with $2700 of UCAR overheads).

Status Report:

A document was needed to define the status of the various reanalysis data projects. The main people who need the info are NMC, GFDL, University of Missouri, and DSS staff at NCAR.

GLOBAL PRECIPITATION FROM MSU

We hosted a talk on this subject by Roy Spencer from NASA Marshall on April 8 (about 25 people from NCAR and NOAA attended). This group has also prepared global atmospheric trends from MSU. There is no increase in temperature during 1979-1992.

WORK WITH ARCHIVE GROUPS

During the past 2 years, interfaces with National Library and Archive groups have become more common. In January 1992, I attended a NSF panel meeting about the National Archives—and their entry into the area of digital data. On April 9 Gertrude Long visited NCAR; she is archivist for the International Monetary Fund. She wanted to know about long-term digital storage and how to approach the various problems.

COADS SHIP PROJECT

The COADS ship project has been in high gear for a couple of months. The following major outputs have been delivered:

- Ships for 1982-1983 (for University of Maryland) became ready for use on April 3.
- Ships for 1985 (for NMC). These data were ready on April 14.
AMELIA EHRHART

She was lost in the Pacific about 175°W, 5°S on July 2, 1937. Jacobson of the Office of Naval Research is helping to research the case. He called on April 10 to ask about old code forms and weather data.

NOAA SATELLITE DATA PLANS

NOAA wants to organize their satellite data facility into more of a research data facility and couple it strongly with the research section of NESDIS. Steven Ambrose (SDSD) and Bob Money (NCDC) visited here about these questions on April 28. I gave them copies of "Readings About Data Formats" and "Computing Power from PCs to Supers."

CD-ROM PROJECT

Rick Anthes will be with the University Relations Committee on May 4. We sent him a memo about the status of the CD-ROM project (April 30).

A LARGE VOLUME OF WORLD GAC DATA WILL ARRIVE

A large quantity (up to 9000 tapes) of NOAA 4-km global GAC satellite data will start to arrive by late May (for the 5 years, 1987-1991). A memo is available. The volume is 255 Gbytes per year.

DATA FROM NEXRAD RADARS

Tim Crum visited on Apr. 27. He is from the Nexrad group in Oklahoma. Bob Serafin chairs one of the main Nexrad committees and asked that Tim visit us.

There will finally be 159 radar sites. If nearly all data were saved from each radar, they expect:

- Basic data: 500 GB per year per site
- Derived map data: 6 GB per year per site

The data will go onto Exabyte tapes (4.6 GB each), which will be sent to Asheville. NCDC (Asheville) says that it would cost them $600,000 per year to keep all of the data. This includes making a copy of the data. This is based on receiving a tape each 3 days from each site (19,345 tapes/year, almost 4.6 GB each) (total volume of 80.4 Tbytes per year.) NCDC estimates that it will take them 5.6 hours to catalog and copy each Exabyte tape. The NCDC cost estimate is very reasonable for handling this amount of data.

I do not think that the U.S. should save all of the basic data all of the time. Tim indicated that the agencies also have doubts about the wisdom of saving all of the data, all the time.

I think that NCAR should obtain a complete set of derived map data (159 radar times 6 GB per year each), thus about 954 GB per year.
VISIT TO NMC (Reanalysis) (May 29)

Five of us talked for two hours about the reanalysis project. Then I saw individuals until 6:20 pm. They got our test file of data inputs on May 27, but had not tried it yet. We talked about the different datasets and how we will share the work. At NCAR, we will not handle all of the necessary work to eliminate duplicates, but we will try to bring duplicates together. NMC is gearing up to handle the rest of the work on duplicates. It is most important that we have time to assemble the data. I have notes that cover many more subjects from the trip.

ASOS MEETING (May 28)

Our group looks at the NWS observational program from a climate point of view. We met at the NOAA instrument test facility at Sterling, Virginia. NOAA is making more of an effort to handle climate concerns, but there are still problems to work on--like measuring total precip correctly. Notes from the trip are available.

CD-ROMs

The University Relations Committee met at NCAR on May 4. One subject for discussion was CD-ROMs. I prepared a memo for Rick Anthes, "Status of the Preparation of CD-ROMs."

NRC PANEL ON EOSDIS

The NRC panel put out an interim report on EOSDIS, and asked me to review it. On 13 May 1992 I sent them a 6-page review. Overall, I think that the panel needs to be a little tougher.

BUDGETS TO HANDLE SATELLITE DATA IN EUROPE

The budget for satellite operations and to handle satellite data during 1992 through 1998 in Europe is 180 million ECU (about $220 million). This is a cost of about $32 million per year. This cost includes ground operations for the satellites; bringing the data to the ground; putting the data into archives, inventories, and user support. It includes some QC on the data. About 200 people are involved in this effort.

Under this program, Europe handles ERS-1 data. Europe will handle Landsat-6, Radarsat, Sea wifs, and other birds. They figure that their budgets are about one-fifth of those in the U.S. for similar work. These budgets do not include other spending to obtain basic data from the above archive money to do research, or to generate data products within individual countries in Europe. This information is from Gerhard Treibnig. I gave him a sheet of questions about the handling of satellite data in Europe, and budgets (6 May 1992).
SAR DATA FROM ERS-1

Europe obtains data from four readout centers. From Aug. 1991 through Apr. 1992, they obtained about 300,000 scenes of SAR, each about 100 km square. Each scene has about 120 MB of data. This gives a SAR volume for 9 months of about 36 Tbytes of data. The processing includes making a second copy of the data.

OTHER DATA FROM ERS-1

All of the other data from ERS-1 probably only is about 10% of the SAR volume.

FORMATS: TALK WITH JOE KING OF NASA (5 May 1992)

In June 1992, Joe will bring together about eight format groups to help describe the properties of their format systems.

Then in the fall, he will have a bigger format meeting to discuss the result. Joe will ask me to this meeting. I gave him a sheet with some of the information that I would like about each format.

DEVELOPMENTS IN HDF AND Net-CDF

Barbara Mihalas (NCSA) told me that there will be a co-development effort between the two groups to try to take care of some of the problems that I documented last fall.

WRITINGS ON FORMATS

During April-May, Olivia and I bundled together a set of papers about data format issues. I gave a copy of this bundle of papers to Joe King (NASA) and to Sara at NASA.

Sara said that the main format person at Goddard will be Don Sawyer. She will pass the set of papers on to him.

I should send a copy of the format papers to Elaine Dobinson at JPL. Elaine has been involved in interoperability issues for catalog and data issues. I think that she has a better feel for the subject than most people. A few more details are still being cleaned up in the format text.

GLOBAL GAC SATELLITE DATA

NASA will send five years of 4-km (5-channel) GAC data to NCAR. It is on about 1800 round tapes each year. The volume is about 255 Gbytes per year. This data is used to calculate variables such as clouds, sea surface temperature, surface albedo, and vegetation indices. Coakley (OSU) and others need access to the data. The data for 1989 arrived at NCAR about May 26. A 3-page memo is available, dated 30 Apr. 1992. This is a very large project.
TOVS SOUNDER DATA


NEXRAD RADAR DATA

There will be about 159 radars, and the cost for the radars is $3.5 million each. About 56 radars will be installed by the end of 1993. If all of the data were saved, it would be 80.5 Tbytes per year.

Asheville has estimated that they could inventory the data and make a copy of the Exabyte tapes for $600,000 per year. This is an attractive estimate for this quantity of data (80 TB). The data probably will not all be saved.

The archives need better definition to give us what we really need. I have a 6-page memo that is almost ready for Bob Serafin and others to read.

WORRY ABOUT TIME

The GAC and TOVS tapes are data that we need, but these large projects are hitting us at a bad time. We cannot afford to delay the other large projects (CD-ROMs, COADS, Reanalysis, and tasks to get NMC/ECMWF data).

May 1992

VISIT TO NMC (Reanalysis) (May 29)

Five of us talked for two hours about the reanalysis project. Then I saw individuals until 6:20 pm. They got our test file of data inputs on May 27, but had not tried it yet. We talked about the different datasets and how we will share the work. At NCAR, we will not handle all of the necessary work to eliminate duplicates, but we will try to bring duplicates together. NMC is gearing up to handle the rest of the work on duplicates. It is most important that we have time to assemble the data. I have notes that cover many more subjects from the trip.

ASOS MEETING (May 28)

Our group looks at the NWS observational program from a climate point of view. We met at the NOAA instrument test facility at Stirling, Virginia. NOAA is making more of an effort to handle climate concerns, but there are still problems to work on—like measuring total precip correctly.
The archives need better definition to give us what we really need. I have a 6-page memo that is almost ready for Bob Serafin and others to read.

WORRY ABOUT TIME

The GAC and TOVS tapes are data that we need, but these large projects are hitting us at a bad time. We cannot afford to delay the other large projects (CD-ROMs, COADS, Reanalysis, and tasks to get NMC/ECMWF data).

NAS MEETING OF DATA CENTER DIRECTORS (June 9)

The second meeting of the NAS Data Centers Directors was held at NCAR (one meeting was held 2 years ago). Data centers at NOAA, NASA, USGS, and DOE were represented plus academy people and others (about 25 representatives). I gave a short talk about format issues around the country. Several of the right people, such as Jim Green at NASA and Jerry Bartrom at NOAA, were very interested.

PACIFIC BOMB TESTS, 1960s, U2 FILM

We have had 1 or 2 tons of 70mm U2 film from Pacific bomb tests for 15 years, with beautiful pix of clouds, etc. MMM does not need it, and on June 10 NCDC (Asheville) said they wanted to keep it. We will ship this stuff to Asheville.

NATIONAL OBSERVING NET OF NWS

I prepared notes about the status of observations and plans for the new NWS observing net. We still do not have all of the obs that we need for climate, but the committee on climate has helped.

WEATHER RADAR - NEXRAD

I had been refining notes and recommendations about the radars for 3 weeks. On June 10, these papers were sent to Serafin and to others. The volume of data could be about 80.5 Tbytes per year. The archive plans need attention badly. NCDC estimates that they can copy, inventory, and archive this amount of new data each year for $600,000 per year. This is an attractive price.

A BIG NASA, NOAA, NCAR PROJECT - SATELLITE SOUNDERS

NASA and NOAA have a "pathfinder" project to do good science using NOAA TOVS sounder data, Oct 1978-present. NASA wants us to help provide the data to the PIs. NCAR has worked up mechanical and fiscal details, and I sent the memo out on June 11. Under the plans, NASA pays NOAA $190,000 to send us an update for Oct 1989 - Mar 1992. Then we send old and new
data to three places (including back to NOAA). NASA pays NCAR about $325,000. NASA headquarters says that the three organizations should take credit for this good work. True. This project also will help the reanalysis project via NASA inputs to NOAA.

MET WITH NCDC (ASHEVILLE) CHIEFS

On June 10 I met with Ken Hadeen (Director of NCDC) and Gus Shumbera to discuss data swaps, troubles with U.S. raobs, the USSR bilateral and other subjects. The meeting was very useful. There is a need to prepare one version of the U.S. daily cooperative data (8000 stations, 70 years) in a reverse sort from now. NCDC likes the idea, so we may help them by doing the sort.

WORLD RIVER RUNOFF DATA CENTER, GERMANY

On June 15-17 there was a meeting in Koblenz, Germany, to discuss hydrological data and programs, and to help formulate plans for future activities of the runoff center. I told them about NCAR data and reanalysis. They said that many centers would be jealous of the data holdings at NCAR. I argued for more emphasis to obtain the long records of runoff data, not just for data for the 1978-on period. About 23 people attended this invited meeting.

Text: Computing Power from PCs to Supers

A new version of this text became available on Jun 11. It now has most of the comparisons between computers, including history, that I wanted to include. The section with the history of PCs is now stronger also.

COADS PAPERS

We had an international COADS meeting in Boulder in Jan 1992. I now have two of my texts done for the report; one more short one is needed. I reviewed two of the papers that others have written, but the natives are restless about waiting.

NSF BUDGET

Congress is worried about the deficit. Pam Stephens, NSF, said on June 30 that Congress says that NSF will get a zero increase in funds for FY93. This is a far cry from the 26.5% increase recommended by the President a few months ago (and no one believed it would happen). It may also be that Congress was irritated over the bad NSF forecasts of the need for engineers (see Science, 10 Apr 1992 and Nature, 16 Apr 1992).

-end of June 1992 report-

20
VISIT FROM SOUTH AFRICA (Aug. 13-14, 1992)

Chris Koch visited from South Africa; Chris is the head of the climate section in Pretoria. They sent us South African raob data last year. I now have several pages of notes about many good datasets that they have. This includes about 20 years of upper air data over southern oceans taken from their research ship.

CLASSIFIED MILITARY DATA

On 14 Aug 1992, I sent in paperwork for an updated security clearance. Senator Gore has asked the CIA to brief several of us about what military data exists that would be useful for global change studies.

HIGH DATA PRICES AT LOS ALAMOS

Los Alamos set the data storage and movement charges about 13 years ago. Bill Collins thinks the charges were more or less reasonable then. Since then they have only come down a little, yet the price of computations has come down a lot. He said that the high prices are causing them a lot of problems.

I said that as a user I would consider setting up my own storage system. He said that is what people are doing.

Management is reluctant to lower their prices, because the present prices generate a lot of real money. I think that this is normal behavior for monopolies, but a little competition is starting.

GERMAN MODEL CENTER, HAMBURG

On Aug 26, 1992, Hans von Storch visited from the Max Plank Center, headed by Bengtsson. He needs a lot of data to carry out his work, and got a lot of updated information about the data at NCAR.

I asked if we could get back some of the long series of raob data for Germany. He said that would be very hard and that I would have an easier time of getting it than they would. He said that the data situation in Germany is a real mess, and bureaucratic. The data center is under the Weather Service, which does not respond to the needs of science. Peter Spaeth used ECMWF analyses, with the permission of the Weather Service. He got a huge bill for 4,000,000 DM, which had to be paid within 6 weeks unless he gave a full report of the research they did.

European scientists use our help a lot. They should also band together and beat up on their own governments.
INFORMATION ABOUT FORMATS

In late August we revised a new version of "Readings About Data Formats." This includes some information about the many common formats and how the users' life is getting worse. There seems to be very few people who know or who care.

September 1992

Comments by Janne

ADVISORY PANEL FOR NCDC-ASHEVILLE

On 1-3 Sep 1992 I went to Asheville to attend a meeting of their advisory panel. This should help to get an inventory of their punched cards that are still on 16 mm film (called FOSDIC).

I learned that Sony has a new and better quality tape for Exabyte, that should avoid some of the problems of tape life.

MEETING ABOUT NOAA SATELLITE DATA PRODUCTS (9-11 Sep 1992)

I attended a workshop about the resolution and handling of products from NOAA satellite data. About 35 people were there. I gave an invited talk; there will be a report.

TERABYTE WORKSHOP (16-18 Sep 1992)

I attended the Terabyte Workshop sponsored by NOAA and MITRE Corp., and gave an invited talk. I included information about the experiences and costs of several groups who are actually handling very large datasets, or are planning to handle them. A paper is available entitled "Handling Large Datasets," Sep 1992. I took 30 to 40 copies of this paper and they were gone quickly. About 60 people attended.

By chance I met David Sebring (IBM) at the meeting, and had a good chat with him over lunch. David lives in Washington, DC, but is the head of the Houston group.

ACCESS TO CLIMATE DATA AND BROWSE PRODUCTS

I attended a workshop (with about 60 people) at the University of Maryland on 25 Sep 1992 and gave an invited talk. The meeting was sponsored by CESDIS (Center of Excellence in Space Data and Information Sciences). The center is run by NASA, University Space Research Association, and the University of Maryland. I guess that 35% of the people there were from computer science departments. Milt Halem and Jim Green were also there. I was pleased that the group seemed more sensitive to ease-of-use issues than I have noticed with similar groups in the past.
SECURITY CLEARANCE

The security clearance activities are going on to prepare for a briefing about what classified data is held by the CIA and military that might be useful for global change research.

CLOUD RESEARCH

I went to a meeting on 28 Sep 1992 at the University of Colorado to discuss strategies to continue with necessary cloud research, and the surface observations needed for the work. I don't want us to do much work on this project until after the reanalysis work. I think that we can meet this need. Julius London, Bill Rossow (GISS), Judy Curry and Peter Webster were there.

Bill Rossow said that his cloud data was put into the HDF format by NASA Langley, which increased the volume by a factor of five. This is crazy.

REQUESTS FOR THE CLOUD TAPE

A few years ago, we were involved in a project that prepared world cloud statistics on a tape. The tape has been obtained by about 146 groups; it has been popular.

October 1992

CD-ROM OF NMC GLOBAL 2.5 GRIDS

On Oct 12 we received copies of this CD-ROM. On Oct 15, I talked to the UCAR department chairman about it. Fortunately, top officials in NOAA were also there.

UPDATE OF TOVS

On Oct 23, 1992, Bud Booth called from NOAA to say that they are ready to send TOVS data for Oct 1989 through Mar 1992. There are 2064 tape cartridges in 70 boxes.

CLASSIFIED DATA FOR CLIMATE RESEARCH?

I attended a meeting of the Environmental Task Force, set up by MITRE Corp. and started by Senator Gore (Oct 19-20). There were 10 panels to talk about data needed. In attendance were about 50 scientists, plus about 100 people from the Intelligence community and MITRE. This meeting was unclassified.

NEW MOUNTAIN OF GAC DATA

Bob Rank called from NASA Goddard on Oct 28; he guessed that we will have two more pallets of data tapes in 3 or 4 weeks.
DAILY WEATHER DATA FOR THE WORLD

On Oct 13, Paul Dyke visited from Texas A&M. He has a big project in agricultural research for which they need daily temperature and precip data for thousands of world stations. We have been working with Paul, CAC and NCDC since Oct 1991 to help make this happen. NCAR is also using the data we have for 8000 stations. Paul wants access to the older data that we have.

HISTORY OF SATELLITE CLOUD PROGRAM

The initial planning for the International Satellite Cloud Climate (ISCCP) was during about 1976-1979. There was a U.S. meeting about it in 1980 and an international meeting in Hungary in mid-1981. I was mainly involved during 1976-1985. Joel Genuth, from the Center for History of Physics, New York, interviewed me for 3 hours (on Oct 14) about the history of ISCCP.

DATA FORMATS

On the national scene, formats are a hot topic. In Jun 1992, I told the data center directors that they should become more interested in the subject because troubles are brewing. NOAA acted on that issue and has set up a NOAA format workshop for Nov. 19. They asked me to talk at the workshop, but I will be at another meeting that week. I hope to have good feedback from the meeting.

We made copies on Oct 29 of the new version of the formats text entitled "Readings About Data Formats."

OCEAN CLIMATE FOR 400,000 YEARS

John Imbrie called from Lamont. He is ready to send us data for 16 long sediment cores in the world ocean. Each core has the usual isotope ratio and other counts that give the temperature of surface or middle water, and some give the frequency of dust storms. He says he gets good reviews on the format, which is a character string with many comments. It has only 2 MB, but that is still a lot of data.

SEA LEVEL PRESSURE DATA

Sea level pressure data has always been quite popular, but for some reason, we have had several requests for long series of data in the last few weeks. For example, the tide gage people would really like global monthly grids of pressure back to 1900 (1 millibar of pressure is about the same as 1 cm of water depth). I wrote an overview of the status and plans for SLP datasets to help answer questions.

end of Oct 1992
November 1992

EOSDIS

Feelings are running high at NASA-Goddard about formats and related ideas. One person there gave me back comments on the formats text. One statement was "Who will stop them now? Who will stop EOSDIS? It is like a black hole out of control."

On Nov 11, Ted Meyers and R. Furesh from the EOSDIS project office visited here to talk about formats and related data strategies. It was somewhat an HDF sales pitch, but in fact we had a good basic technical and philosophical discussion.

In mid-October, I saw Jim Green at another meeting and asked if he could talk to some of his CDF people about the practicality of taking a somewhat more general approach to formats. On Nov 13 we got a message that they had done this and would soon be sending comments.

CO₂ RELEASE FROM WORLD CITIES

There is a UN program to compare energy use and CO₂ release by a dozen world cities. Denver is one of these cities, and the state and city are involved. Peter Gilman asked me if I would attend their meeting in Denver on Nov 3. I put together some information about energy, CO₂ release, and other environmental-related information. These notes were rather popular, and didn't have enough copies with me.

COADS PROJECT

This project is coming along fairly well, but there are still many problems in data to fix. There have been too many problems at Asheville with bad format conversions and inability to track tapes. A meeting was held in July. People keep wanting to change the text. I worked with Scott Woodruff, NOAA-Boulder, on these issues.

We held an International COADS Workshop in Boulder Jan 13-15, 1992. The proceedings (2.2 cm thick) became available in mid-Nov. I'm glad that the work on the texts is done. I wrote a paper entitled "The Importance of COADS for Reanalysis" and a session summary. Steve Worley wrote about the status of component datasets for COADS. The report has 39 papers plus a summary.

We had hoped to finish the COADS update for the 1980s by Jan 1993. It appears that Feb 1993 is now a better estimate.

SMALL CLIMATE SATELLITES BY NASA

Jim Hansen, NASA-GISS, asked me to review their text from a Feb 1992 workshop. It gives a lot of good information about climate forcings and feedbacks; it proposed a series of small satellites. Both the technical and the cost scaling arguments are good. Given the new ideas in NASA, I am hopeful that these small satellites will be built.
WORLD SURFACE RADIATION DATA

There is a need for better archives of surface radiation data for modeling of evaporation, crops, etc.; for checking climate; and for trends. On Saturday, Nov 14, Anatoly Tsvetkov from St. Petersburg was here. Deluisi (NOAA), Wells (DOE) and I met with him. We will enhance activities under the bilateral WG-VIII and probably put in a proposal to the western nation’s institute in Moscow to help them survive.

Anatoly brought 6 months of 1991 world stations radiation data on floppy disks.

DATA INVENTORIES FOR REANALYSIS

In mid-Oct we sent a series of data inventories to ECMWF. By Nov 7 we added information about three old datasets, and packaged all of the information in a better way. It has 106 pages and covers many types of data; copies were sent out.

TWO-PART TRIP TO WASHINGTON, DC, NOV 15-24

I attended a meeting of the Environmental Task Force at MITRE Nov 16-20. I signed in at 7:30 am and the meetings were from 8:00 am to 6:30 pm daily.

I attended a meeting of the Academy Committee for Geophysical and Environmental Data (CGED), Nov 23-24. I gave a 25-minute talk on formats. Dick Reed said that this was the first time he could really understand some format issues. He asked about more technical format issues at the airport. Reed and Guy Brasseur also are members now.

December 1992

VISIT TO NMC (Dec 16)

I had very useful 5-hour discussions at NMC about the reanalysis project and about the state of model developments at NMC. They are coming along very well. In 12 to 24 months, they may have useful versions of revolutionary analysis procedures. In these procedures, data out to about plus or minus 12 hours will be used at once to prepare an analysis for the middle time. Observations can be inserted at their proper point in time, and dynamical consistency in time is enforced within the coupled forecast model. This is a neat idea. John Derber thinks a plus or minus 12-hour system would take about 10 times the computing as running the forecast model by itself for 24 hours.

The way that satellite sounding data is used in the analyses will get much better over the next year.

Asheville is sensitive about NMC coming to NCAR for input data, rather than getting the data from their own NOAA data center in Asheville. At a recent meeting in Washington, DC, NMC
people were asked why this was the case. I've known for a year that some of these politics were going on. This issue may heat up more. I think that because of this contract involving NCAR, NMC will have more data, a better chance to start on time, and less chance of trouble.

NMC also feels somewhat nervous because we have not already dumped a of data on their doorstep. I understand this feeling, but I don't like to give them too many half-finished datasets or too many problems that we can resolve.

We have had very useful ties with NMC, especially with the Development Division, and especially for the last 5 years.

MEETINGS AT NASA

Some NASA people ask me what the secret of our data success is; I've been trying to figure this out too. In many meetings about data problems, the big problems are not solved quickly and the opportunities are not spotted. In meetings, including the recent Goddard DAAC meeting (about 10 scientists and 8 systems people in the room), there is a presentation by a science group about their data problems. The data people in the science group giving the presentation are quite good, but they need some guidance on how to handle the data. There is another whole group of data software people who are more focused on building their own systems whether or not it solves the problem. Fortunately, this group doesn't have much influence in the Goddard DAAC at this time. We might expect that the combination of the science advisory group, managers, and data people in the room would help to home in on the problem and to help solve it. It could be that the group doesn't understand enough of the particular science to help solve it; it could be that they don't have the blended skills needed in science and technology. But in a real sense the problem is often ignored or deferred by the group. I have seen this happen in a number of meetings. If there is time, I often start off-line dialog with such a group to obtain a set of the technical numbers by which one can understand the problem well enough to develop solutions with reasonable tradeoffs. It is often possible to come to a solution in less time than five committees can spend walking around it. The science groups seem to welcome this sort of interaction because it helps with their problem, doesn't waste a lot of their time, and keeps things as simple as possible. Most of the conversation can be short and very much on their wavelength.

NASA GODDARD DAAC MEETING (Dec 14-15)

This was our second meeting; the first meeting was held in Dec 1991. Some good things are happening, but NASA Headquarters is feeling frustrated with the rate of progress. The NASA Goddard Land Data System and Climate Data System were merged into the DAAC. They had developed respect and trust with their users. The success of these land and climate data systems has been strongly undermined by management decisions, national politics, and the need to prepare for new DAAC projects. I ended up feeling frustrated and sad. It almost seemed like it was everybody's fault and nobody's fault. Anyway, as some of the science PIs have privately told me, "they have been screwed." Could this have been avoided? Yes, in large part, but not entirely.

In the case of the land data system, about 75% of the customers are Goddard. About 1987-1988 the data support was poor, and the question was whether continued support should be given
to the data group for land data. About 1989, Greg Hunolt (NASA) told me "they are doing better; give them a chance." This was good advice. Now their data support is good and the NASA scientists trust them. So what happens? It was decided that Oak Ridge must have a DAAC, and that the new DAAC will give service for this aspect of land surface data. This is being called the Gore DAAC. The morale in this Goddard unit is extremely low. The Goddard land scientists see their data support being messed up for administrative reasons. The best solution now is to transfer at least 1 to 3 land data people into the science division to give continued support. The Goddard scientists will need some local support regardless of whether they may finally start getting good support from Oak Ridge.

The Climate Data group at Goddard has had 7-9 people since about 1980. It almost expired in the mid 1980s because of too much emphasis on systems. During about 1987-1989, Lola Olsen did a good job of reviving it. She has a science background and the respect of the science groups. She has provided a lot of good support on a budget of $350,000 per year. Now, although the DAAC has a budget of about $3 million per year, this climate support cannot be continued as well as needed--because of all of the shuffles, plus the need to prepare for new projects.

In Nov 1992, I received a message from a respected source at Goddard. This was in reference to data activities:

It is bound to be that emotions will run high in the present situation, and the message should partly be understood in this context. However, it is frustrating that it has too much of the ring of reality. Can we see a light at the end of the tunnel? Maybe.

We do not want a data system that is by systems people and for systems people. Bob Schiffer and others emphasized this point again. Schiffer's comment was "This is not a galactic game for computer weenies. It must provide a useful service for the science community." I agree entirely. We still need systems, but they should be constrained to be useful and costs should be managed. Also, they need to be easy to use.

In the DAAC meeting, I ran into several discussions where I thought that there was a very good attitude toward constraining costs and making good technical judgments.

PATHFINDER PROGRAMS

Schiffer noted that if the Pathfinder programs cannot be done well, then there is little faith that all of EOS can be done well. (Schiffer is from NASA Headquarters.) I got the impression that these efforts are coming along rather well. I also think they are a good idea. At NCAR we are helping to supply science groups with the TOVS data, to support one of the Pathfinder programs. NASA is helping us by providing 5 years of 4-km GAC data; in turn, NCAR will give access to some of the PIs (e.g., Coakley, Ramanathan) that need it for NASA work and NSF work.

NEW STORAGE SYSTEMS FOR THE DAAC

The Goddard DAAC will not use the computer division's storage systems for EOS. The DAAC is setting up hardware storage systems and seems to be doing a good job.
**Metrum storage units:** They are getting a Metrum silo (600 cartridges).

Each VHS tape holds 14.4 GB; the data rates are 2 MB/sec, with read after write. I also think that this is the best choice they had. They will have four tape units on the system, each with a channel. The data goes to a small front end. This may be a bottleneck. The FE is a silicon graphics 440 server with 8 GB of disk space. There are 8.7 TB of data in a silo unit, and they will finally obtain two units. The carousel ($260K) and four drives ($120K) cost about $350K (or $380K?).

They will also have two silo units to mount optical disks. Each silo holds 1.2 TB. Much of their ocean color data is now on optical disks. Their file system for all of this will be Unitree.

**Summary:** The DAAC will have 19.8 TB of on-line data storage.

A person at Goddard told me in Oct 1992 that the timing in HDF is rather poor. Mary James is getting me some of the timing comparisons. She also said that they sent us 2 years of GAC data about Dec 10. We received the data about Dec 20, on 4 pallets (about 2750 tapes), which was "a Christmas present." Mary is one of the very good people around Goddard. She has an uncommon amount of expertise and good practical horse sense.

**DATA COMPRESSION WORKSHOP (Dec 17)**

I stayed in Washington, DC, an extra day to attend this workshop on data compression at Goddard; it was useful. There needs to be more emphasis on simpler methods that are practical now and on ways to integrate the methods with the flow of data in computer memory. Ken Salem (University of Maryland) gave an interesting talk where he had put compression options on the inside of the Goddard CDF format system. People gave some of the timing information that is needed, but not enough.

**STATUS OF FORMAT DEBATES**

Our text "Readings About Data Formats" is gradually getting out to the people who can contribute to solutions. A partial update on the status of this technical issue follows:

- EOSDIS Project Office decided on the HDF format for EOSDIS products. This was done about Feb 1992. Two key people from this NASA office visited NCAR on Nov 11, 1992 (T. Meyers and R. Furesh). We had a good, basic technical and philosophical discussion on formats and large-scale data handling (they have 2 copies of "Readings About Data Formats").

- The next day, Meyers and Furesh dropped by to get 2 more copies of the format text. One of these copies went to Fulk at the University of Illinois. He is now the main developer of the HDF format.

- 8 Dec 1992: Jim Green at NASA Goddard asked for 4 more copies of the formats text, which were sent to him.
• Steve Unger is in the Land Surface Science Division at Goddard. I have known him for 2 years, through other NASA committees. In Dec 1992 he said that he had heard about the formats text and would like a copy.

• Al Arking had been the head of the Climate Radiation Division at Goddard. He is now on the DAAC panel. He had read the formats text and thought it was a very good review of the problems. He came away with the idea that we should worry more about the user interface than the exact format. This is correct, but efficiency is also a constraint.

• At the data compression meeting (Dec 17), a person from the Naval Research Labs (Mitchell Grunes) had budding ideas about formats that seemed useful. He said that things are getting too complex. We will send him a copy of the format text.

• Ken Salem (University of Maryland): On Dec 17, 1992, Ken gave a good talk at Goddard about how he put choices of data compression on the inside of the CDF format, and left the user view the same as before. He even included the ability to decompress part of a grid but not a of it.

• A copy of the formats text was sent to Warren Wiscombe, Goddard DAAC advisor. Mary James saw the text and made 5 or 6 copies of parts of it for people that work with data at Goddard. She also told me in Dec 1992 that they are leaving the big datasets (like GAC) in simple flat formats as long as possible.

• Another section at Goddard told me the same thing. They are not (now) being pushed to convert data to HDF.

• Patterns of data use: Our experience, and Goddard's, has been that people will resist the use of data in complex formats that come with thick manuals. The head of NASA Ocean Research was at the DAAC meeting. He uses the Goddard data system (that has data in CDF format) but he always uses the system to convert the data to an easy format to bring it home for use. This is consistent with how Goddard says that the systems are being used (PIs ask for data on tapes in native format, and when they use the on-line system, they take small amounts of data home in ASCII format. Therefore they do not have to learn the common format).

• Jim Kinter at the University of Maryland sent me a fax (Dec 3, 1992) saying that he read the formats text and found it to be very useful. He gave it to Brian Doty to read (I see Brian as one of the very good developers in the U.S. who can contribute to solutions).

• Time needed to use a common format: Steve Worley got some of his ocean data back from IBM/Houston in netCDF format. The netCDF software was already loaded on his computer. It took 1.5 days to figure out how to read the data. When it takes this long for an expert in the data and programming areas, it means that there would be trouble in forcing the format onto all users.
— Many groups around the U.S. are designing such common formats, but they are not paying enough attention to the usability of them by users.

- Trouble: Lots of people design data systems that blow up the volume of big datasets. We need to avoid this and to make it possible to use easy ways to compress data. I mentioned this at the Dec. 1992 meeting on data compression.

- On Nov 24, 1992, I talked to the NAS Geophysical Data Panel about data formats. I told them that at the present time it was best to realize that there are problems with what is happening in the U.S. with formats, but that it is best to let it "muddle along" for awhile to clarify options for better designs.

- People at Goddard have told me that the timing for data flow on HDF is not good. They will send me some numbers.

- I met Bob Evans again (University of Miami - Oceans) at the AGU meeting in San Francisco. He is getting very frustrated with the number of complicated common formats--and the fads keep changing.

- I have a set of briefing notes valid 3 Dec 1992 by Mona Kessel, Goddard: "What is CDF? Why use CDF?"

REANALYSIS AT ECMWF

Tony Hollingsworth at ECMWF sent a letter to me on Dec 18, 1992, saying that their European Council approved a 15-year reanalysis project to be done by ECMWF. The reanalysis will start in the fall of 1993, and they hope to complete it by fall 1995. They need part of the data from us. Tony will visit NCAR on Jan 29, 1993.
DATA SUPPORT SECTION

Monthly Status Report
January 1993

Comments by Roy Jenne

NOAA/NASA VISIT

On Jan 15, 1993, I had three visitors here for another meeting. Concerning NOAA, I was asked whether a review committee about data flow problems at Asheville was reasonable; the answer was yes. It surprised me when Greg Hunolt from NASA Headquarters said that the EOS DIS decision for the HDF format was an interim decision, mostly for image data. Now they realize that there are many other types of data, and they have people checking on other formats such as E-BUFR. At the moment, the data center situation in NASA is not good.

MEETING IN WASHINGTON, DC

During Jan 10-14, 1993, I attended another meeting of the Environmental Task Force. The report is coming along.

U.S.-RUSSIA BILATERAL

Perhaps I should have attended a bilateral meeting in Washington, DC, and gone to Russia Jan 16-24, but I didn’t. I did take some time to do U.S. coordination about the agreements. Also I have written reviews to try to help two U.S.-Russia data proposals.

RUSSIA

Even the electricity bills for the data labs in Obninsk, Russia, are difficult to cope with. The bill for Dec 1991 was 700,000 rubles; by Dec 1992 the bill increased to 9,000,000 rubles (more than inflation). In Jul 1992, the exchange rate was 142 Rb to $1; on Jan 15, 1993, it was 475 Rb, and then about Jan 25 it was 550 Rb to $1. A person, like an NCAR division director, made about 30,000 Rb/year as of Jan 1993 (about $60/year). Key entry staff are paid 1200 Rb/month now. People say that a policeman makes 3 times this amount; it used to be that scientists were treated well in Russia.

A train ride from Moscow to St. Petersburg cost 700 Rb (mid-Jan 1993). A poor hotel in Moscow cost $75 per night, and a good hotel in St. Petersburg cost $90. People from Russia would have obtained a much better price in the same hotel.

Some of us here (and in Russia) are worried about whether the data center (and the data) will survive. Some people don’t expect it to last beyond 18 months.

MASTER DIRECTORY (FOR USA)

NASA administers the U.S. Master Directory that has dataset listings from all agencies (and other countries). They have several discipline coordinators who help obtain information for the different sciences. John Scialdone visited us on Jan 26, and we gave him much more information, mostly for weather, climate, and oceans. I was worried that we didn’t have time to give him enough information; he started feeling that he was receiving too much info to absorb.
TIGER TEAM AT GODDARD

Dave Halpern called on Jan 22 from JPL (ocean science) (and he did run the data system at JPL). He was at a meeting at Goddard about the ocean satellite Seawifs program (there are about 10 people in Seawifs at Goddard). The head of the Goddard DAAC said that a Goddard tiger team had been set up to try to improve communications between Seawifs and the DAAC. Dave could hardly believe what he was hearing, but he knew the program was in trouble when this is going on.

INFORMATION ABOUT ECMWF AND REANALYSIS

Tony Hollingsworth visited NCAR on Jan 29 to talk about reanalysis and the data that they need.

REANALYSIS AT ECMWF

ECMWF recently obtained the approval from their council to do a 15-year reanalysis (1979 through 1993 and perhaps 1994).

The project starts Feb 1, 1993; they hope to start production in Oct or Nov 1993. ECMWF is funding Rex Gibson, who will head the project. Two more people will be funded by the U.S. (DOE and NSF). Larry Gates preurchased the analyses for the DOE part; some help comes from Shukla. The Project Manager is David Burridge, head of ECMWF. The project is "definitely the director's baby."

The reanalysis will take 8-10% of the C-90. They plan that it will take 3 years (Nov 1993 - Nov 1996). They hope to do 10 days of data per day; this would mean 1.5 years to finish but with reruns and days with no runs, they plan for 3 years. ECMWF has not decided on which years to do first. It depends somewhat on data availability; they don’t want to wait too long to do FGGE again.

WHAT DATA DOES ECMWF NEED?

They will use FGGE data and their own decoded data for 1980-199X. The added data that ECMWF especially wants are:

1. SST grids. I gave Tony a text about the available grids and the plans.
2. TOVS Soundings (clear radiances). It appears that we have a deal with NCDC (as of Jan 29) to obtain the set at NCAR for $30,000 (regular price about $62,000). NCAR will return inventory information to NCDC; ECMWF will help pay. There is "no way that ECMWF could afford $68,000."
3. COADS: 1979-1993. NCAR will send this data to ECMWF when it is developed.
4. Rawinsondes for remote stations. NCAR will create a set, but it won’t be as complete as we would like.

I think that the following datasets should also be considered for use, since they fill gaps in remote areas:
5. Some added aircraft data. I think that at least the data for the New Zealand region, tropical GASP aircraft, and maybe data from Australia should be used.

6. Ice cap buoy data during the 1980s (for surface pressure).

WHAT DOES NCAR NEED?

I said that NCAR can help ECMWF to obtain the data, but what we need is some agreement about access to the data.

SEMINAR BY HOLLINGSWORTH

After Tony got here, we arranged for him to give a seminar. He gave us a rundown on model developments at ECMWF. He especially wanted feedback about possible causes for the behavior of winds in the deep boundary layer. Sixteen people attended, and the seminar was very productive.

One side note: ECMWF runs an ocean wave model that gives very good results. But it does not feed back into the atmosphere. Their work is helping to produce much better algorithms to work up satellite altimeter and scatterometer data.

In dry, big subsidence areas\(^1\), the model has about 60\% too much precipitable water; in the moist, narrow tropical strip (areas of uplift), the model is 10-15\% too dry. This is probably because the large-scale vertical motion is not quite right.

METEOROLOGICAL SERVICES IN EUROPE AND THE U.S.

Tony noted that in the weather services (observing and forecasting, not all of the climate) there are 25,000 staff members in Europe and 5000 in the U.S. In Europe there is a lot of pressure on budgets for these functions. This causes problems with the cost of data and access to it.

THE C-90 AT ECMWF

ECMWF signed a 4-year contract with Cray Research that included increases in computing power. It was for Jan 1991 through Dec 1994. The first 2 years of the contract gave them a Y-MP-8. Since fall 1992, the C-90 has been running well. However, he said that they are only able to get 50\% utilization, which is low; we need to check. It has 16 processors, 128 million words, and an SSD with 512 mwords. The cost has been $4.8 mil pounds/year, with some increase each year; the 4-year cost is (4 * 4.8L plus 15\%). The plan was to obtain a 12-processor system, but it did not quite meet the specs.

\(^1\) such as northwest of South Africa.
DATA SUPPORT SECTION

Monthly Status Report
February 1993

Dataset Updates, Additions, and Requests

Datasets updated during the month and the most current dates available are:

- NMC Medium Range Forecast 10 day forecasts, 31 Jan 1993.
- NMC Medium Range Forecast sigma level analyses, 3 Feb 1993.
- USSR Marine Ship Data, current to 1990
- MEDS Drifting Buoy Data, current 1992
- ISCCP C1 (ds742.0) and ISCCP C2 (ds742.1) to June 30, 1991.
- NESS Visual and IR Brightness Anal, August, 1992
- NMC surface and upper air station data, 23 January 1993.
- NMC upper air split, 2 January 1993
- NMC 65x65 and 47x51 Grids, all 47x51 grids thru December 1992
- 47x51 Daily Grids, thru December 1992
- 72x19 Daily Grids, thru December 1992
- 47x51 Monthly Grids, thru December 1992
- 72x19 Monthly Grids, thru December 1992
- Global MSU Temperatures, thru January 1993
- CAC daily summary of the day put into time series order. (ds512.0)
- World Monthly Surface Station Climatology (ds570.0), 1991.
- UNESCO river data, added German WMO data.

DSS completed 33 orders in February. There were three orders from Japan, two from Italy, and one each from France, Mexico, Korea, and Canada. Approximately 25 datahelp/email users were assisted in getting data questions answered and/or data requests filled (four of these requests were passed on to other DSS members, and one request was passed on to SCD consultation).

General Archive Activity

Code was written to convert the older NMC flux data (in quasi-ON84 format) to the newer GRIB format. Users who want to use data that spans the format change period can now convert the older data and then use the same read code for the whole period.

FTP Logging

We have compiled the following statistics from the FTP logger on huron.scd.ucar.edu (ncardata) for February:

34 - 2
• DSS had 2930 total connections, including 2886 anonymous connections (up from 2656 in January), from 584 different sites. (Non-anonymous connections originate from DSS staff, and are not included in the statistics.) 2427 of these connections resulted in file retrieval.

• 645 different files were retrieved, including "MASTER" files for 100 datasets, and 14 accesses to retrieve information for data orders (usually multiple files per order) from the "download" directory. 180 users retrieved the general "README" file, and 54 users retrieved the file describing the DSS pricing policy.

CD-ROM

Programs to sort out records of AFOS dump tape have been completed. We began to investigate SAO, synoptics, upper-air, etc., reports. It may be necessary to merge sources to get the best possible set of hourly obs, which will complicate preparing data for the CD-ROM.

Satellite Data

The 1991 update of GAC data has been completed. All tapes from the 1990 and 1991 shipments, a total of 2804 tapes, have been archived since Dec 22, 1992. Catalogs are available for all of these files, and access programs are being worked on.

ATD has begun extracting NOAA/GAC data from the NCAR/GAC archive, dataset ds703.0, for the Kuwait oil fires study. In order to use the GAC binary data from the NCAR archive with their Terascan software, Sea Space has modified the ingest routine of Terascan software for ATD. Currently ATD is processing the GAC data in a production mode. With interactive verification of satellite aerial coverage, ATD can process one year of GAC data on their workstation in 4-5 hours. Once regional images of interest are extracted, two copies of these images are saved using the UNIX tar command on two separate storage media, DAT and Exabyte tapes. As for the missing data, ATD has found about 5-7 missing passes out of a total of about 500 passes over the Persian Gulf in 1991.

All 128 ISCCP B3 tapes are ready to be imported.

The import of TOVS cartridges for the update through Mar 1992 has been completed, and the primary archive preparation is done. Some work remains to prepare the MSU and SSU subset archives.

Unidata

Wetterhorn has been running very stable recently.

MMM WISPI began on Feb 16, 1993. The UNIDATA system support for WISPI includes several data streams. Upper-air, SAO netCDF files are routinely generated using raw data from wetterhorn, Purdue University, and the University of Wyoming. Dave Gill has run front-end tests to examine these files. All netCDF files are usable for MMS runs. NMC NPS MRF model grids are routinely generated using raw data from wetterhorn, FSL, and Purdue University. The University of Wyoming files are available from MSS directly. All netCDF files have been examined, and they are all usable. FSL MAPS data files are not available consistently. Shell scripts to check and generate MAPS netCDF files are routinely executed. If MAPS files are incomplete, an e-mail message will be sent to NCAR operators to ask them to
check the availability of MAPS data by calling FSL operators. To make sure the wetterhorn files are available on MSS for Scripps Institute, the files they need are checked regularly.

**COADS Ship Project**

Final data arrived at NCAR and data format translations necessary for Release 1a COADS were completed. This included data on (1) about 30 cartridges from NMC, (2) two tapes from Russia, (3) three tapes with the 1992 drifting buoy data from MEDS, and (4) one tape containing 1992 data from the Inter-American Tropical Tuna Commission. Processing of Release 1a has begun and will proceed by operating stepwise on 2-year subsets for the 1980-1992 period. All totaled, there are 130 million records (10.2 GBytes) in 428 data files for initial input. Following the data selection and duplicate elimination processing steps there will be roughly a 50% reduction in the total number of records.

In support of ongoing work in Russia, the 1980-1991 interim MSTG COADS product was sent to the WDC-B in Obninsk.

Various sorts and merges were done in support of COADS processing. The 88,89 COADS test compress/sort/merge in statistical order and the 80,81 COADS compress/sort/merge in B10 order were competed.

The IATTC Tuna and Porpoise log data has been converted to the LMR6 format. The data covers the period 1971-1991. At this time, an update to this set is being converted.

**Incoherent Scatter Radar Project (Barnes)**

The backlog of data requests was processed and a detailed study of the Millstone Hill Madrigal package was begun, in preparation for writing the User Guide chapter. Since the local man pages were removed (when cedar was reconfigured) and the Madrigal man pages were moved, a script which identifies and prints one-line descriptions of each database command (mancedar) had to be made smart enough to conditionally find the Madrigal man pages.

**Reanalysis Project**

The two GASP tapes received from Gregory Nastrom at St. Cloud State University did not contain the data to fill the data gap in our GASP archive. It may be that there are no data to fill this gap, though inventories lead one to believe otherwise. Anyhow, a short note stating the situation concerning the gap will be included in the dataset, and the GASP archive will now be considered complete.

Work is progressing on the GATE research aircraft to get the data ready for inclusion into Reanalysis. Time-series of the 1-minute averaged data are being assembled at approximately 3-minute intervals, and these are being "tweaked" some to give time-series with the minimum amount of missing data.

Work has begun on matching stations in the NCDC-reformatted USAF synoptic dataset ds463.0 to their locations in station dictionaries. Inventories by station are being made. It's evident that many "stations" only show one or two reports per year, and are probably typos or other errors.
Comments by Roy Jenne

DATA FOR INTERNATIONAL GEWEX PROJECT (WORLD WATER CYCLE)

Paul Try called from the International Office near Washington, DC. He needed information about some of our analyses that will help this project.

In their planning, they are also running into the format problem. Some scientists at Langley say that they put data into an easy format; it was changed by the DAAC into HDF and then no one could read it. In general, people in the GEWEX project think that formats such as HDF are too complicated for general use. So it sounds like Schiffer (NASA headquarters) will give the go-ahead to use simpler "native" formats.

NCAR - AVAILABILITY OF ATMOSPHERIC MODEL OF T, P, AND MOIST SL (30 km)

We have a good deal of information about the mean atmosphere that people can use in models, but not enough to answer all questions.

HANDLING LARGE DATASETS

The first version of this 25-page text was written in Aug 1992, to prepare for a Terabyte conference in Sep 1992. It gives information about several projects around the U.S. and Europe that process and store huge amounts of data. Such comparisons should help all of us to plan large data-handling projects, and to talk with groups who have gained experience. The text was updated in Feb 1993 to reflect new technology. It will be an appendix in another government document.

DAILY WORLD SURFACE DATA FROM 1979-ON

We have a set of the daily data (precipitation, temperature, etc.) for which we update for Apr-Dec 1992. Data for Jun and Sep 1986 are still missing. Texas can help us fill this gap. Trenberth’s group and ECMWF in England need the data (3400 MB) in station sort. Spangler just accomplished this sort, but it needs the updates.

Internet seems to go at speeds of 21 KB/sec (1.25 MB/min) between NMC and NCAR. We can probably use this to update the files (140 MB each 3 months).
DATA SUPPORT SECTION

Monthly Status Report
April 1993

Comments by Roy Jenne

DATA FOR GEWEX

GEWEX is a 10-year global energy and water experiment that includes surface process, hydrology, radiation, etc. Data for this experiment will be collected for 5 years. They need archives from global models (NMC, ECMWF, Canada); we have the first two and prefer not to add Canada, except for a subset.

Archives from the NMC 40-km regional model (ETA) are needed; the volume is about 1 GB per day. John Leese (formerly with INO but now with GEWEX) visited NCAR on Apr 15. He said some people argue that the data should go to Asheville; others argue that it should definitely go to NCAR, "where they know about model data." I have had similar feedback from three sources in the last 4 months. Leese also asked whether our data services, represented by Jenne, was a person or an institution. I told him that there was some cause for concern, but that the group has the necessary institutional skills.

We should have this ETA model archive for other reasons. Some focused subsets will be needed, and we will need to invent better ways for low-cost data transfer. This can be done.

TOGA-COARE DATA (Apr 15 meeting)

I attended 3 hours of TOGA COARE national meetings on Apr 15, and gave a short talk on how we could help with some of the data problems. Some argued for a totally distributed data system, where the data is not gathered. I said that it was good to distribute the data preparation, but the results must be gathered or the data would evaporate on the 5- to 20-year timeframe.

VISIT TO SCD FROM NSF

Dick Greenfield from NSF visited SCD management on Apr 15. He is the new head of Meteorology at NSF and wanted to introduce himself, hear about different computing issues, and develop stronger NSF-NCAR cooperation in the planning process.

NAS MEETING ABOUT THE NATIONAL ARCHIVES

The National Archives and Congress want to do more to save digital data that should be saved for the 20- to 50-year time scale for physical sciences, lab sciences, and technology. A NAS study is being funded to help define what should be saved either in agencies, the Archives, or both. One preliminary meeting was held in Jan 1992, and the first study meeting was held Apr 5-6. The lady who is now the acting head of the Archives told me out of the blue that she read my "Belgium paper" about data management issues and liked it. I was pleased, because some of the principles covered will foster long-term retention, lower costs, and data integrity.
MEETING AT NMC ABOUT REANALYSIS

I attended a monthly meeting at NMC about reanalysis, and gave a 15-minute talk about the data status. The 2.5° TOVS data is still a problem (gaps and quality). Three of us talked with Larry McMillin (satellite research) about a new TOVS algorithm that could be used (the one being developed for the TOVS pathfinder).

For related reasons we need to send Susskind at Goddard more basic TOVS data, but we need the funding to do it. On Apr 16, I found out that Martha Maiden (NASA) will soon send the rest of our TOVS money (we received $112K before and will receive the rest of the $325K (approximate total) this year.

REANALYSIS STATUS OVERALL

- COADS is coming along but the timeline is very tight. Also, we need to check whether the data from the 1979 FGGE ships is still the old version (as of 1982) that had some bad wind units.

- Older raobs: There are still a huge number of issues to iron out. Coordination with European countries will take time (Finland work has started). I have to iron out some options with NCDC. The GFDL work on TD54 is coming along well.

- Aircraft: The work on many datasets is progressing okay.

- Surface land data is coming along, but timing is a worry.

- TOVS issues: NCAR, NMC, NESDIS, and ECMWF still have not sorted out how to fill gaps in the 2.5° data.

- Getting 1962-92 data to NMC: This is coming along but slower than needed.

- Overall: Making timing goals will be difficult, but there is still a chance to be in an acceptable range.

NOAA MEETING ON DATA CONTINUITY

Data continuity means that we need to accomplish our observing, data management, and science functions well enough that our 5-year to 100-year descriptions of the behavior of the earth and its changes (temperature, ozone, CO₂, aerosols, etc.) are accurate. I was asked to attend NOAA’s second meeting on this subject, Apr 19-20. There were reports on the new NWS observing system, COADS, reanalysis, satellite data, fisheries data, etc.; the meeting was useful.

MORE CLIMATE MODEL WORK; ASSESSMENT STUDIES

Before the global Rio meeting (June 1992), President Bush promised countries that the U.S. would assist them in analyzing the effects of environmental change, such as CO2 warming. We would help support emissions inventories also. Official requests for support are coming in to EPA from countries. This year they have funds for studies in about 20 countries. Each country will set up a more general office for various assessment studies than before. About 10 U.S. Federal agencies are involved.

Ron Benioff from EPA called on Apr 23, 1993. He asked if we could help with the climate model data again. I told him yes, with some support. We can make it easy to access the data,
but we do not want to do the interface work to various models, such as crop or forest models. He said that people such as Cynthia Rosenzweig (GISS) will also be involved to help with some of the other tasks. Our data support involvement and the studies will probably start about Oct 1993.

USE OF THE NCAR DATA SUPPORT ON-LINE SYSTEM

This system has catalogs of our datasets, documentation about them, selected reports and software, and some smaller amounts of data, all for quick access on-line. It was useful as soon as it became available for users in Jan 1992. By Jul 1992 another main step in adding information had been attained, and more is being added all the time. Ilana Stern was able to gather usage statistics starting in mid-Dec 1992; 3 months of summary statistics follow:

<table>
<thead>
<tr>
<th></th>
<th>Jan 1993</th>
<th>Feb 1993</th>
<th>Mar 1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total connects (not DSS)</td>
<td>2656</td>
<td>2886</td>
<td>3235</td>
</tr>
<tr>
<td>Total sites</td>
<td>584</td>
<td>613</td>
<td>865</td>
</tr>
<tr>
<td>- education sites</td>
<td>271</td>
<td>293</td>
<td>432</td>
</tr>
<tr>
<td>- government sites</td>
<td>95</td>
<td>97</td>
<td>103</td>
</tr>
<tr>
<td>- sites outside the U.S.</td>
<td>70</td>
<td>76</td>
<td>128</td>
</tr>
<tr>
<td>- other sites</td>
<td>148</td>
<td>147</td>
<td>202</td>
</tr>
<tr>
<td>Number of connects with file retrievals</td>
<td>2138</td>
<td>2427</td>
<td>2169</td>
</tr>
<tr>
<td>Total retrievals</td>
<td>7016</td>
<td>7427</td>
<td>7141</td>
</tr>
<tr>
<td>- for weather data</td>
<td>4807</td>
<td>5511</td>
<td>4961</td>
</tr>
<tr>
<td>- for dataset info and other</td>
<td>2209</td>
<td>1916</td>
<td>2180</td>
</tr>
</tbody>
</table>

Note: The weather data available on-line (and free) consists of only a few small datasets, some limited data and forecasts for cities, and ski slope information. The system is accessed by anonymous ftp to ncardata.ucar.edu

DATA CATALOG AND COMMUNICATION, APR 25 - MAY 1, SAN DIEGO

I attended a meeting of the advisory panel for the Master Directory for multidiscipline datasets as administered by NASA. Links with European and Japanese activities are going well. I think that people need higher level ways to find useful data (such as fact sheets, summaries, etc.). This time I made that point again, and others seemed to pick up on it for the first time. There are questions about how these master directory activities will relate to growing "global change" directory efforts. NASA may make a separate information system for the global change program; in any case, one system should have all of the high-level information.

There were many presentations given on world communications. I have some briefing slides on catalogs and communications, and more will be sent to me.
DATA SUPPORT SECTION

Monthly Status Report
May 1993

Comments by Roy Jenne

OZONE DATA

Julius London, University of Colorado, had a request for ozone data from China. They noted that I had said that some ozone data could be part of an exchange agreement. We have London's early global grids and other data. I made up a list that summarizes the ozone data that are available in the U.S. and the world. London and I will talk again before he visits the Chinese in July.

RATE OF ICE ACCUMULATION ON MT. FUJI

Sam Hensef (Electronic Space Systems Corp.) is putting some instruments on the top of Mt. Fuji (11,000 feet). They need to know temperature, winds, and rate of precipitation in order to determine the heating requirements needed to melt the snow and ice. He only has part of the data and strategy needed; they have some data from Mt. Washington to help. His phone number in Massachusetts is 508-369-7200. This is an example of various questions we get; Gayle Gray in the NCAR Library received the call on May 6, 1993.

OLD RAOBS

There are still many problems in the datasets of old raobs (1940-70). There is difficulty in knowing about the existence of certain data, and the overall history of what was done to prepare the data. I talked with Dick Davis at NCDC/Asheville and then put together a few pages of notes to describe the big picture, as well as we know it now. This compilation of notes was sent to Davis on May 7.

TOVS DATA FOR REANALYSIS

In the 2.5° dataset, there are many time gaps. These gaps can be filled using the basic data. I found that a NOAA/ERL group in Boulder is also interested in this problem; they want to help fill the gaps. I put together a 20-page package of notes to describe the TOVS issues and what we know about them.

STATUS OF PC COMPUTING

Intel recently announced the Pentium processing chip (the 586 chip that people have waited for). It is a hot chip! Even floating point processing is quite fast. Table 1 shows the history of Intel chips from 1974-on. Table 2 compares the speed of the Pentium chip with a number of other fast chips that are being used.

OTHER EVENTS IN MAY

On May 3 I saw Mark Thorley from the Antarctic section in the UK. (They will give us raob data.) On May 9-13, I was at a NOAA ice and snow meeting at Penn State. Peter Finger visited from the Global Precipitation Center in Germany.
Table 1. The Evolution of Intel PC Chips, 1974-1993

<table>
<thead>
<tr>
<th>PC chip (in MHz)</th>
<th>MIPS</th>
<th>Transistors</th>
<th>Internal Bus (in bits)</th>
<th>Date Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.64</td>
<td>6,000</td>
<td>8</td>
<td>Apr 1974</td>
</tr>
<tr>
<td>8080 (2)</td>
<td>0.33</td>
<td>29,000</td>
<td>16</td>
<td>Jun 1978</td>
</tr>
<tr>
<td>8086 (5)</td>
<td>0.33</td>
<td>29,000</td>
<td>16</td>
<td>Jun 1979</td>
</tr>
<tr>
<td>80286 (8)</td>
<td>1.2</td>
<td>134,000</td>
<td>16</td>
<td>Feb 1982</td>
</tr>
<tr>
<td>80386DX (16)</td>
<td>6</td>
<td>275,000</td>
<td>32</td>
<td>Oct 1985</td>
</tr>
<tr>
<td>80386SX (16)</td>
<td>2.5</td>
<td>275,000</td>
<td>32</td>
<td>Jun 1988</td>
</tr>
<tr>
<td>486SX (20)</td>
<td>16.5</td>
<td>1.185M</td>
<td>32</td>
<td>Apr 1991</td>
</tr>
<tr>
<td>486DX2 (25/50)</td>
<td>40</td>
<td>1.2M</td>
<td>32</td>
<td>Mar 1992</td>
</tr>
<tr>
<td>586DX, Pentium (66)</td>
<td>112</td>
<td>3.1M</td>
<td>64</td>
<td>May 1993</td>
</tr>
</tbody>
</table>

- 1974 - The 8080 chip was used in the Altair, the first personal computer.
- 1981 - IBM PC introduced; this changed the world of computing.
- 1984 - IBM PC-AT introduced with the 286 chip.
- 1986 - COMPAQ DeskPro was the first computer with the 386 chip.

Notes: (1) This information was mostly from PC Magazine (Apr 27, 1993)

(2) The Pentium has two computing pipes, not just one pipe
    as with the others listed here. It uses 13 watts under normal conditions, 16 watts at
    peak power. The DX2 chips have a fast internal speed and a slower speed to talk
    with the outer world. The next Intel chip family, the 686, is expected to be released
    about late 1995, with 10 million transistors.

Table 2. Compare Integer and Floating Point Speeds

<table>
<thead>
<tr>
<th>PC Chip (in MHz)</th>
<th>MIPS</th>
<th>Spec Integer '92</th>
<th>Spec Floating Point '92</th>
</tr>
</thead>
<tbody>
<tr>
<td>486DX2 (33/66)</td>
<td>54</td>
<td>32</td>
<td>16</td>
</tr>
<tr>
<td>586/Pentium (66)</td>
<td>112</td>
<td>65</td>
<td>57</td>
</tr>
<tr>
<td>DEC Alpha/21064 (150)</td>
<td>74</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>MIPS/R4000 (50/100)</td>
<td>62</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>MIPS/R4400 (75/150)</td>
<td>82</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>PA - RISC/PA 7100 (99)</td>
<td>80</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td>IBM Power PC/601 (66)</td>
<td>e 50</td>
<td>e 80</td>
<td></td>
</tr>
<tr>
<td>SPARC/Super SPARC (40)</td>
<td>53</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

(based on PC Magazine, Apr 27, 1993) Megahertz times for the Pentium are expected to range from 60 to 99 MHz)
Two missing months of daily world data for the end of 1986 have been missing from our 14-year record of daily world data from 7500 stations. We have been trying for two years to fill the gap; the data was also missing in original CAC files and in Asheville files. But they were in the data that Texas had obtained from CAC. Three cheers, the record is now complete.

Meeting on ice and snow at Penn State, May 11-12.

I went to Penn State U. on May 10-13 to discuss ice and snow datasets that NOAA needs to take care of. The meeting was sponsored by NOAA. I also found information about the YMP-2 that Eric Barton's group uses for climate modeling (I have notes). Wilson said that SCD had helped them a lot, and specifically mentioned Dick Weisbrod's help with some of the software routines they needed.
**NCAR**  
*Extension Summary Report*  
*Period: 07/01/93 - 07/31/93*  
*Date: 08/01/93*

**Department: 3971**

<table>
<thead>
<tr>
<th>Ext. Identification</th>
<th>Calls</th>
<th>Duration</th>
<th>Cost</th>
<th>Equip. Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1214 STERN, ILANA</td>
<td>5</td>
<td>0:26.1</td>
<td>3.48</td>
<td>3.48</td>
</tr>
<tr>
<td>1215 JENNE, ROY</td>
<td>57</td>
<td>9:37.5</td>
<td>76.72</td>
<td>76.72</td>
</tr>
<tr>
<td>1216 JOSEPH, DENNIS</td>
<td>24</td>
<td>1:27.4</td>
<td>10.84</td>
<td>10.84</td>
</tr>
<tr>
<td>1217 SPANGLER, WILBUR</td>
<td>2</td>
<td>0:07.7</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td>1218 WALTERS, GREGG</td>
<td>2</td>
<td>0:15.8</td>
<td>2.11</td>
<td>2.11</td>
</tr>
<tr>
<td>1219 JENNE, ROY</td>
<td>2</td>
<td>0:00.0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1230 BARNES, ROY</td>
<td>2</td>
<td>0:37.9</td>
<td>4.33</td>
<td>4.33</td>
</tr>
<tr>
<td>1248 WORLEY, STEVE</td>
<td>23</td>
<td>4:34.6</td>
<td>33.45</td>
<td>33.45</td>
</tr>
<tr>
<td>1256 COMEAUX, JOSEPH</td>
<td>13</td>
<td>1:11.0</td>
<td>8.26</td>
<td>8.26</td>
</tr>
<tr>
<td>1825 DATORRE, ROBERT</td>
<td>8</td>
<td>0:17.0</td>
<td>2.12</td>
<td>2.12</td>
</tr>
<tr>
<td>1833 SHIH, CHI FAN</td>
<td>6</td>
<td>0:38.7</td>
<td>24.15</td>
<td>24.15</td>
</tr>
</tbody>
</table>

**Department: 3971**

<table>
<thead>
<tr>
<th>Calls</th>
<th>Duration</th>
<th>Cost</th>
<th>Equipment Total</th>
<th>Cost Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>142</td>
<td>19:13.7</td>
<td>166.49</td>
<td>0.00</td>
<td>166.49</td>
</tr>
</tbody>
</table>

---

41
DATA SUPPORT SECTION

Monthly Status Report
July 1993

Comments by Roy Jenne

COADS

The big COADS update for 1980-92 (13 years) was recently completed. (This is a big achievement.) There are 40.9 million reports with a total volume of 3.76 Gbytes. Others (here and at NOAA) have been involved, but this is a primary achievement of Steve Worley and Wilbur Spangler in our group. Will put it all into synoptic sort by July 17, and Steve sent it to NMC on July 20. It will first be used there for new calculations of sea surface temperature which are needed for reanalysis. We have been trying to find time for this major update since 1986.

On July 23, we met to talk about the next steps needed to get the data for 1947-79 ready for reanalysis.

NATIONAL ARCHIVES MEETING, IRVINE, CALIF. (July 6-9)

I am on a NAS panel and subcommittee (atmospheric sciences) to advise the National Archives and the agencies on issues relating to the long-term retention of digital data (50 years and more). Both the main panel (and all of the discipline subcommittees) met in Irvine, Calif.

TOVS 2.5° SATELLITE DATA

On July 22, Dave Burridge (head of ECMWF) called. He is ready to pay his half of the TOVS bill and needs an invoice. He wants to know when he can get the data. We have it for 1979 - Jul 1985 and 1992. NCDC-Ashville will try to be faster on the tape copies. Joey is making us a detailed inventory. We will work with ECMWF to try to figure out how to fill gaps. There are big gaps in the NOAA archives.

REANALYSIS

This project will be reviewed again in October. I prepared a text to describe the status of our data project.

TEXT ABOUT COMPUTING OPTIONS

A new version of the text "Selected Fast Computers and Computer Options at NCAR" was completed, but there is still more to do.

PROJECT TO ARCHIVE NMC ETA MODEL OUTPUT

John Leese (NOAA GEWEX Project Office) talked with me about our archiving the NMC ETA model output. These are high-resolution grids for N. America and are needed by the GEWEX Project (global water and energy project). The test period is scheduled for summer 1994 and continues in 1995.
EOSDIS MEETING, JULY 28-29

I attended the 2nd meeting of the Academy EOSDIS panel in Washington, DC. We are making some headway in understanding what is going on. It does have about $2.5B overall. The Hughes contract is about $811M, or $81M a year. All of the DAACs together get $19M. The Hughes contract includes some flight operations work. Too much planning reminds me of the syndrome of big bucks (and systems) plans as in DoD and NASA.
DATA SUPPORT SECTION

Monthly Status Report
August 1993

Comments by Roy Jenne

TRIP TO URUGUAY (July 30 - August 6)

I attended a meeting of the Inter American Institute (IAI) for global change in Uruguay. Bob Corell of NSF is chair of the overall activity and wrote to encourage me to go. I gave a paper that included information on greenhouse gases, our climate model data for assessment studies, other data that we have, and information about data that we have. A lot of our reports disappeared from the table. I wrote sections on data for the workshop report.

Since the 1930s and 1940s, we have had good estimates of the biomass of fish species in the California current. There is a natural variability that is very large. Fishing should not always be blamed for changes in the population.

Off the coast of Chili and Peru is 15% of the world’s fisheries: sardines, etc. About 1950 the catch was 0.1 million tons per year; from 1965-70 the catch was 1 million tons, and from 1985-on it was 6 million tons. Chile ships a lot of fish meal, but the low price (about $500 per ton) makes it hard to make much money. Other fish bring $3000 to $4000 per ton. Chili wants to become the second biggest exporter of salmon. Most species are now being over exploited. They want to be able to determine whether a country is really growing or whether they are using up their natural resources.

About 100 people attended the workshop, which dwindled to 60.

VISIT TO THE SNOW AND ICE DATA CENTER (Aug 11-12)

As an Academy task for the EOSDIS Panel, Elaine Hansen (University of Colorado) and I visited SNIDC in Boulder. I have thought of them as a nicely focused data center with a staff of 10 to 12 people doing good work. With the help of NASA they have a staff of 35; they are planning for a staff of 68 by 1996 and 87 by the year 2000. Parkinson is at work. Is this what we really want in an era of "reinventing government"?

I have a text about this visit and associated pages and figures from their reports. I also have more summary notes on EOSDIS. These papers will go back to the Academy panel.

NATIONAL DATA ISSUES

The EOSDIS issues are eating up my time. The Academy also wants me at the 3rd National Data Forum at the end of September to consider the Agency plans again. At least I’m getting more definite information on what is happening in NASA. Bill Bonner wants me in on the review of NOAA plans.

A NEW VERSION OF THE TEXT ABOUT FORMAT ISSUES

A new version of the text "Readings About Data Formats" has been completed and has been sent to about 8 people.
DATA FOR HIGH SCHOOL USE

On August 18, a science teacher from Denver (Richard Irwin) was passed to me by Outreach at UCAR. He has an advanced class of juniors and seniors. They talk about atmospheric science, climate models, evidence for and against climate change, population, energy, etc. We will get him some of the digital data for Denver weather that he wants. Sometime I hope to find time to do a little focused work that is aimed at needs for education.

TRIP TO RUSSIA (Aug 21-28)

I visited the Hydro Meteorological Institute of Russia, at Obninsk, 106 km west of Moscow. Five years ago their staff consisted of 1250 people; now they have 750 people, of which 550 are actively working. Conditions were so bad last year in Russia that they could not invite us. If we came, they said that they could not talk about accomplishing anything during the year. This year we did go, and they even were able to pay for our hotel. Slack from USGS (about rivers) and Shumbera (from NCDC) went with me. I think that we will get a lot of river data exchanged during the next year, but those negotiations were difficult.

On Wednesday we visited the U.S. Embassy for an hour. There is a Science and Technology Institute in Moscow (funded by the West). We have had proposals there to try to obtain funds to help Russian PIs survive and put datasets together. This funding won’t work—it is designed only to employ nuclear scientists so they won’t go to other countries. There are other possible sources of funding. A U.S. ocean scientist got into a flap 2 year ago. We talked about this. The embassy was trying to resolve it before Vice President Gore visits there soon.

We have a set of written plans for the next year’s data activities. These will be fed into the Washington, DC, office of the bilateral exchange. It was a successful trip. Conditions are bad in Russia, but people are surviving. I have notes. Inflation is 18% each month. Apartments seem cheap, but food and clothing are expensive.
MONTHLY REPORTS

JAN 1994 – MAR 1995

BY ROY JENNE

DATA SUPPORT

SCD – NCAR
Monthly Status Report
Comments by Roy Jenne

January 1994

PREPARATION OF CLIMATE MODEL DATA AND TEXT

I have been very busy preparing for the Country Studies Workshop being held from Jan 31 to Feb 4. There are lots of issues and lots of text to prepare. Data for the UK 1989 run arrived on Jan 25. Dennis Joseph prepared data from 3 model runs on 7 floppy disks, plus access programs, plus text. Joey had to make 25 copies of the 7 disks of data that were needed at the workshop (175 disks).

REANALYSIS

I prepared another text to give the status of datasets for reanalysis. It has been a hectic month on this issue. On Jan 27, NMC received the 6 cartridges from us with upper air data for 1985. More will follow soon. USAF at Asheville is copying 2 tapes for us with raob data from Brazil. The aircraft data is flowing from New Zealand.

PANEL ON LONG-TERM DATA STORAGE (National Archives)

On Jan 18-21 I went to a meeting of the NAS archives panel. We are developing advice about national policy on saving digital data for a long time, e.g., a 100-year archive. Cost control is critical, or else data will be needlessly lost. They asked me to write a section on how to accomplish the necessary archive functions and still control costs. I have almost finished these writing assignments for the report.

AMS MEETING IN NASHVILLE, Jan 23-25

I could only attend meetings for 1.5 days, but I found out many useful things. Bob Cess and Jeff Kiehl (NCAR) gave talks where they outlined things that have been wrong in climate model radiation codes for many years. If this is fixed, it will help upper troposphere heating rates, and it will improve the amount of energy getting to the surface.

On Monday evening, I talked with Ken Hadeen of NCDC for 2 hours--until midnight--about reanalysis data, etc. On Sunday evening, they gave me a certificate and declared me a Fellow in the AMS. Also I bought a very good satellite book and found a good paleoclimate book (18K years) to order.
TALK WITH A VISITOR FROM ECMWF ABOUT REANALYSIS

Rex Gibson visited from ECMWF on Jan 27-28. We talked for 9 hours about many data issues—status of various datasets and policy issues—charging, etc. ECMWF is helping to fill gaps in the TOVS 2.5-degree archives. NMC mailed us their archive of 150 tapes of TOVS data about Jan 24. ECMWF now hopes to start reanalysis production in March 1994.

February 1994

TRIP TO WASHINGTON, D.C. (3 meetings), Jan 30 to Feb 4

- Country Studies Meeting. Dennis and I were very busy getting documents and data ready for this trip. About 1 to 3 people from each of 17 countries attended this workshop. Each country got copies of climate model data on 8 floppy disks from us. Both Dennis and I gave talks. We also talked separately with people from various countries. I attended 2.4 out of 6 days. This was a useful meeting.

- Meeting at NMC about their mesoscale model data. NMC will start saving data (analyses and forecasts) on May 1 and send it to NCAR. It includes diagnostic data such as precip, radiation, and clouds. The volume, gulp, is about 6 Gbytes per month. At least this is down from a first proposal of 94 Gbytes per month. This project is associated with GEWEX and includes NMC people and John Lease of GEWEX.

- Meeting of the Committee on Earth Studies (CES), Feb 3-4. We had briefings by NASA and NOAA about their plans. I asked whether there was hope to obtain some laser wind measurements. Kennel's assistant seemed interested and talked with me about it afterwards.

VISITOR FROM POLAND, Feb 8

Professor Kaczmarek is from the Academy of Sciences in Poland. They have used our climate model archives for assessment studies. He was at the Country Studies meeting in early Feb. UCAR invited Poland to join the UCAR affiliates program, so he visited here. I talked with him for 4 hours and have 4 pages of notes about data in Poland. I think that he will send us some daily and monthly river data.

VISITOR FROM TAIWAN

The head of the Taiwan weather service visited NCAR on Feb 22. He was in the university system there before this job. Rick Anthes knows him well and introduced us. He remembered Dennis and I from 15 years ago. He is interested in some of our data. We could use some of their observations.
His assistant took notes.

SATellite DATA Handling IN JAPAN

On Feb 10-11, I met with about 6 people from Japan's NASA. Our mass store people were there too. NASA is setting up a DAAC in Japan to handle satellite data products from their satellites. They later sent me some information about expected data rates. It was very useful, because I now have better information about their satellite plans for the next 5 to 8 years.

Reanalysis

This work is hot and heavy for the group. We are making progress on many fronts.

. Two-thirds of the New Zealand aircraft data is here.

. More TOVS sounder problems are understood. The complete set of 2.5-degree data from NMC is now here (in addition to the NCDC set). The NMC set came on 167 cartridges.

. We got all data from Antarctic and Greenland ice cap buoys for 1980 through 1992 from the University of Wisconsin. This includes the update for 1992 and older corrections. (We now have various inventories.) The data is almost ready to hand over to NMC and ECMWF.

. Other sets of raob and aircraft data are coming along.

. NMC has hurricane data.

. NMC has nearly all data for 1985, their starting year.

. We need to fix the data from ships for 1947-79 by Aug 1994—a big job.

MeETING ABOUT GODDARD Reanalysis

On Feb 17-18, I made notes about Goddard reanalysis for 1985-89. A visitor was here from Goddard (Sigfried Schubert) and gave a seminar too. There are a few bad gaps in TOVS 2.5-degree data that they might help fill from a different set of 1-a-day NMC tapes that only they have.

Options For Computing

I did some more work describing the options for computing equipment and obtaining the latest information about timing tests.
March 1994

A WIND SATELLITE

It is looking possible that we may get a satellite to measure global winds, using lasers, and achieve this in 3.5 years. In 1990-92, NASA Marshall and others were designing the $500M to $600M satellite (often called the $500M to $1B satellite) that was looking more and more doubtful to achieve, even by the year 2004. In late 1992 the likelihood of sharp increases in NASA budgets was poor; they started working on a low-cost way to measure this important data.

On March 1, 1994, Greg Wilson called from NASA Marshall. He thinks that they can build a wind satellite and launch it within 3 years from the official start. The total mission cost is now down to $30M to $50M, including some data handling. The NASA wind satellite planning was officially stopped about Jan 1994. Also in Jan 1994, I was in an Academy panel meeting with the new head of NASA earth science, Charles Kennel, and his deputy. I asked if there was a possibility of obtaining a wind satellite and said that we didn't need the high-cost version. Kennel visited NASA Marshall in Feb 1994. Wilson says that with Academy backing, there is a chance that a miracle will happen and we can build this satellite. Kennel was very interested in it.

JOINT PROJECT WITH EUROPE AND DATA POLICY

Europe wants to work with us on a joint project to improve rawinsonde data. I like this idea, but then they want to classify the data (we would have to track users and deny commercial access). On March 1, their project manager called me from Spain. I'll write another position paper within a week about what rules we could agree with.

Europe wants me to agree to work with Europe on preparing data and then keep the data restricted (only for research). I have been resisting this policy of restricted data for a year, because it would help to set precedents and be a pain. It would become very hard for us to serve the general education community. I prepared a text on the subject of data policy to argue open data, and sent it to the coordinator in Spain.

I called Bretherton. This is now an even hotter issue in the U.S. than I realized. A WMO drafting team has almost completed a policy--designed by Europe. The U.S. and Australia are trying to delay this move. Bretherton wanted some examples. I gave him a copy of the paper on March 16. One case is that we will not use Canadian surface observations in reanalysis because they are too costly, and they are restricted.
COUNTRY STUDIES PROGRAM

This program with climate model data has taken a lot of time to organize more data and to prepare new text about the models, and about model variability, etc. The draft was sent in on March 17.

ISCCP 1995 FOR WASHINGTON, DC

We are preparing data for this program to help assessment modelers obtain what they need. The U.K. is preparing something, but it isn’t what the models need, so the program is delayed and 200 PIs are scattered around the world. We have to prepare 1000 to 1800 copies of floppy disks. Yuks! I sure wish we had time to make a CD-ROM and that the world’s PIs had CD-ROM readers or Internet.

April 1994

MULTIPART TRIP TO WASHINGTON, DC, Apr 6-14

Attended the following meetings:

- Meeting of the Committee on Earth Studies (satellites) (Apr 7-8), sponsored by NAS (Apr 7-8)
- Visited NMC about the reanalysis project (Apr 11)
- Visited the World Bank about their need for climate data and forest info (Apr 11)
- Attended the NASA Catalog Interoperability meeting and workshop (Apr 12-13)

May 1994

MEETING ON GLOBAL CHANGE MODELING, May 1-4

A meeting was held at NSF, Washington, DC, to help plan future directions for modeling efforts in the U.S. About 80 people attended including 6 or 8 agency reps. On Sunday evening and for two days thereafter there were concentrated talks about what is happening across the country about models, and on the future outlook. On the last day, the direction and changes for U.S. plans were discussed.

A nice joke was told in a summary: "The university modeling groups claimed that more money should be given to the big labs, but the labs said 'No, no. We don't need anymore.'"
About computing, there seemed to be a strong sentiment to restrict the use of big climate models on big computers. Jim O'Brien (FSU) stated this gently by saying that any computer center director who tried to enlist the 1001st customer should be fired. Jim Hansen, GISS, argued against supercomputer centers because it is easier and less expensive to compute locally.

SATELLITE LIDAR WINDS

Jeff Rothermel visited from NASA-Huntsville about the design of a satellite to measure winds, using backscatter from lidar beams. He gave me slides that give technical details about the satellite and about the scattering properties of the atmosphere. They will develop more tradeoff and cost information and present it to our CES panel in July. This will include possible cost sharing with France or Russia (for launch, and maybe for the spacecraft bus, which is the satellite itself without the instruments).

CLIMATE MODEL ARCHIVES, May 16-17

I attended a meeting (Dickinson, chair) in Tucson, Ariz, to discuss more archives from climate models. The agencies (NSF and DOE) seem to be interested in more effort in this area. It is also the case that most people did not realize what has been going on here at NCAR since 1987. I gave 'a talk about our archives of model data (to about 50 people). The meeting finished without making good enough use of the expertise in the room. Several people were frustrated by this. Four of us got together afterwards and discussed a strawman plan. I will write it up. Update June 9: The strawman draft is almost done.

The person from DOE said they would probably invite me and a few others to Washington and get more action going. I want the new project to make the daily data available.

DICKINSON EOS PROJECT

There was a meeting for the PIs of the Dickinson project at NCAR on May 19. The PIs are from NCAR, JPL, CU, Wyoming, Arizona. The project has been criticized for not having enough glue to hold it together. There are certain themes that tie it together, reanalysis is one of these. There will probably be a longer meeting in August. Our data part of the project has gotten good reviews.

DATA FOR GCIP, May 23-24

A meeting about GCIP data was held in Boulder on May 23, 24. About 15 people attended from around the U.S. For this experiment there are three main "modules" for data. We have the one for model data.
I gave a talk on the status of planning to have the Eta model data (NMC new mesoscale model). The new model is run at 30-km resolution. The archive will have a resolution of 40 km. This procedure matches a standard NWS product for forecasters, it reduces the volume a lot, and it avoids the staggered grid complexity of the basic model output. We will invent a new system to deliver large amounts of data at a low cost. I also provided information about our information systems.

The volume of data that NCAR gets from Eta will be about 6 Gbytes each month. This is a lot of data.

Attachment

GCDIS implementation plan for NCAR

3.x

The National Center for Atmospheric Research (NCAR) in Boulder, Colorado, through the Data Support Section (DSS) of the Scientific Computing Division (SCD), archives and distributes a wide range of meteorological, oceanographic, and climatology data. Although this is not a formal activity assigned to this NSF facility it is supported as a service to the NCAR and university research community. The University Corporation for Atmospheric Research (UCAR) oversees the operations at NCAR and in particular and SCD Advisory Panel, comprised of UCAR representatives, meets annually to review the status and future plans of SCD including the data management activities.

The 10 person DSS staff maintains, updates, distributes, and develop data products from over 400 different data sets that use four Terabytes storage on the NCAR Mass Storage System (MSS). The archive represents an irreplaceable store of data that are used for major national and international atmospheric and oceanic research projects. The data can be roughly divided into the categories of atmospheric analyses and observations, oceanic analyses and observations, model output, satellite data, and supporting geophysical data sets. The NCAR archive is best known for the data in the atmospheric and oceanic categories. The complete data collection is accessible to approximately 400 NCAR and 650 university personnel who use the NCAR computing facilities. Access to the data is also extended to U.S. Government agencies, the private research sector, and research groups worldwide by distribution on various media and through network transfers. In brief summary, the DSS services over 400 offsite data requests, requiring about 150 Gigabytes of data to be shipped annuallly. Data transfers for those using the NCAR facilities are much higher.
The efficient MSS and the connected computing systems provide the DSS with the capability to routinely manipulate large amounts of data. This capability contributes important support for research projects at NCAR and throughout the research community. For example, the DSS has and is preparing data for Global Atmospheric Reanalysis, the Comprehensive Ocean-Atmosphere Data Set (COADS) is produced in cooperation with NOAA, and output from climate models is collected and distributed worldwide for global impact assessment studies. NCAR is also a participant in bilateral data exchange agreements with Russia and China with goals to foster international cooperation and more efficient data resources worldwide. Maintenance and growth of the NCAR archive and continued work on these and many other large data manipulation efforts will remain a primary focus for the DSS in the years to come.

DSS provides access to data and metadata in a variety of forms and methods. Some older printed atlases, technical notes, and documentation still contain important data and metadata. These hard copies are distributed in original form. The DSS strives to convert these documents to digital format. This will enhance the access for many users, but support for hard copies will remain to service requests from users who are not as technologically advanced. A DSS staff specialist is assigned to each data set in the archive. The specialists have experience with the data sets and are capable of answering most questions about format, access procedures, distribution, and general documentation. The DSS also maintains Internet accessible online information systems of data and metadata for the archive. At this time the system is populated with metadata for all data sets in a quasi-standard format along with some small or sample data files. At the most elementary level, persons with Internet connectivity can access and browse the online system with simple ftp login and file transfer methods. This system, while very successful, has been enhanced to support World Wide Web information searches through the National Center for Supercomputing Applications Mosaic interface. This interface nicely provides a GUI (using hypertext and simple ASCII files) for users with Mosaic capability and yet does is gauged by the fact that it is accessed in about 3000 sessions during which 7000 files are transferred per month. The DSS will continue to strengthen this system by adding more data, metadata, data set inventories, browse data, and by employing Mosaic capabilities to handle query and key word searches.

Not only does NCAR provide data support and distribution on a per request basis, but it has also begun a project to produce CD-ROMs for mass distribution. One CD-ROM containing atmospheric analyses has been completed and is available for distribution. A series of other CD-ROMs is planned. These will contain more atmospheric analyses, atmospheric observations, and possibly long term data for climate studies. Presently, NCAR has several data sets that need to be distributed, but are too large for CD-ROMs or the existing network communications. NCAR intends to distribute
these data on existing and new low-cost media as they are
developed.

NCAR has begun to reach beyond the science research community and
is now, through the SCD Users Services Section, distributing
information about computing, science, and general activities to
undergraduate universities, high schools, and K-12 students.
Announcements of opportunities at NCAR can be posted to an e-mail
group specifically working in K-12 education. NCAR has also
hosted the Colorado Computational Science Fair, 1994, in
cooperations with Colorado State University. Activities like this
will become more prevalent as we recognize and organize
information about and at NCAR that can serve as stimulating
material for students.

The NCAR library resources are also available to anyone with
Internet connectivity. Using an online library catalog and key
word search capability information about the library holdings
(books, technical notes, journals, atlases, maps, etc) can be
reviewed. Users with authorization can then directly request that
the hard copy material be delivered to them. Unauthorized persons
must use an inter-library loan procedures to obtain copies of NCAR
library holdings.

The NCAR archives are diverse in type and sometimes large in
volume, so far, the data support focus has been on data content,
data information, and easy user data accesses. NCAR will remain
committed to these principles. We will continue to take advantage
of the most practical and publicly available information
distribution tools. Many of the NCAR data sets are described in
the GCMD. The NCAR will continue to provide information to that
directory, as well as support automatic connection to the NCAR
data information system. This system will meet the level four
contribution criteria for GCDIS directory, guide, and inventory
functionality.

As the transmission capacity of networks increases, larger amounts
of data will be directly transferred to and from NCAR.
Accompanying the growth of network capability will be the
development of network tools that efficiently search for data,
browse data, and examine and display information files. The
resident NCAR data information system will remain flexible and
capable of adapting to the new network features and, in doing so,
will meet the higher levels of interoperability and connec-
tivity required for the GCDIS.

Since NCAR has long been a data provider to the university
community and worldwide, it presently offers level-3 and -2 GCDIS
data order and distribution functionality, respectively. As more
specific GCDIS guidelines are defined, NCAR will make procedural
adjustments to accommodate these standards as considered
appropriate.
MEETING ON IPCC ASSESSMENT STUDIES (June 6-7)

The lead authors for world assessment studies met in Washington, D.C., on June 6-11. About 125 people were there. I attended the workshop on June 6-7 and gave a talk about the climate models. It was fortunate that I attended, because there are a number of issues connected with the models (How do you assign real years to model decades? Which data on the disks should people use?). NCAR prepared 200 copies of PC disks with model data (transient runs) that the PIs received earlier.

NCAR has been working with the IPCC office since Nov 1993. The PIs have what they need, for now. I am glad that our involvement can be somewhat less for awhile. The model data has also been available online at NCAR starting April 15.

I got a copy of the draft chapter on world energy policy and hope to help review it. For energy and for transportation, I made a plea to give projections of the growth in use, plus the efficiency of delivering the service. Too often the terms for use were not there.

LIDAR WIND SOUNDING SATELLITE

The main missing observations from satellites are global winds. On the CES panel of NAS, I am the person trying to help this technology along--with controlled costs. The team at NASA Marshall has been working very hard to design a Lidar sounding satellite at a cost much better than the $750 million that it was 2 years ago. I met with the team (a dozen people) at Marshall on June 16. They were talking about a duty cycle of only 10%. We can use a lower laser shot rate, but we do need a near 100% duty cycle to see all the world.

On the plane home I made up some suggested viewgraphs for their CES briefing in July, and sent them by fax. They would need to do a lot of engineering work to fill in the numbers. Wayman Baker (NOAA) suggested that they follow this script (which they did).

REANALYSIS

A lot of work is going on to prepare data for NMC. NMC is getting a good start on reanalysis. Their present data from us starts with Dec 1984 (a warmup for 1985).

. NMC had received the data for Dec 1984 to Dec 1985 before June 14.
. Dennis sent cartridges for 1986 by overnight mail on June 15.
. Dennis will send cartridges for 1987 on June 16.
. Then he'll send about a year of data each week.

NMC has TOVS data for the whole Jan 1979 to Jan 1984 period. They are getting 2 months of gap fillers from Goddard.

VACATION

I was gone to Washington State on vacation from June 19-28.

July 1994

FORECASTING CONFERENCE IN PORTLAND, OR (July 18-22)

I attended meetings about forecasting models in Portland. There was one session about reanalysis, where I gave a talk. People from NMC, ECMWF, NASA and lots more were there. Topics covered mesoscale models, global models, clusters of forecast runs for longer range forecasts, etc. Status of reanalysis about July 18:

. NMC had done Dec 1984 thru Jun 1986 (1.5 years)
. ECMWF has completed 6 months in 1979

In side meetings I obtained more information about the proposed data archives. NMC needs to work up some more technical information. If sigma data has really been transformed to 2.5 degrees, it isn’t okay.

SATELLITE CES MEETING IN WASHINGTON, D.C. (July 6-7)

I attended this 2-day meeting at NAS. We had briefings on NASA and NOAA satellite programs. About eight people came up from Huntsville to brief the CES panel on the planning for a Lidar wind satellite. They have found a way to get a duty cycle of 100% Cheers! (they first talked about only 10% in Huntsville). There are 10 laser shots per second. They outlined costs for three different options, as I suggested. Their briefing went over well.

Three of us are still working on the findings and recommendations for our chapter about models in the CES text.

The data chapter (for EOSDIS) and an education chapter will be deleted. I am glad because I think that EOSDIS focused too much on big programs and lots of money, rather than getting the job done and constrained budgets. The latter could produce a better outcome for less cash.

. I like to see groups struggle with tradeoffs, learning from
experience, and cost control. I have not seen this in EOSDIS. The budget for EOSDIS is being cut. There are review teams at Goddard. I prepared some information and sent it to them.

COADS DATA

On July 12 we (ERL, NCAR) had another meeting to discuss the projects needed to complete a new COADS. A 3-page summary memo is available. The reanalysis project is moving along. We will soon need the older COADS data.

SATELLITE SOUNDER DATA FOR REANALYSIS

I wrote most of a text about the status of the satellite data for reanalyses. Copies will be sent to NMC and ECMWF.

August 1994

MODEL DATA: FIVE VISITORS FROM NGDC ON AUG 15

NOAA NGDC in Boulder has been thinking of getting into the model data business. They have been talking with others at NCAR, GFDL, etc. What do people want? Their strengths are paleo work and solid earth. I talked with them for 2 hours. They said that "more than one person" has told them that they should model their approach after our Data Support group. I asked them how people define what we do. He answered that it is the close interface with PIs.

My note: Most planners think that you get the best scientific input by having more data meetings and more advisory groups. This is about 90% wrong.

TEXT ABOUT SATELLITE SOUNDER DATA FOR 25 YEARS

There is a need to show the reanalysis groups the inventories of the satellite sounder data that we have, and other information about the data. We prepared a much more extensive text about sounders than has been available so far. It has considerable information about four main sounder systems, including helpful things to know about data use.

STATUS OF REANALYSIS

The NMC/NCAR production of reanalyses started in mid-June. By July 19, they had completed 19 months (Dec 1984 - Jul 1986). In late July they found that July and August heated up too much at the surface in the model. This problem is related to the way that surface vegetation was described and handled in the model (growing
wheat became dry stubble, etc.). They fixed this and started the reanalysis again about Aug 15.

I prepared a text about the status of reanalysis. It has a lot of information needed to help plan to handle the data. About 50 copies were sent out 30 Aug to 2 Sep.

The International TOGA group met at UCAR and got 10 copies of the text.

TESTING OF NetCDF BY CGD

CGD will probably change the data format that they use for model output data. I had a sidewalk talk with Jim Hack. They did a lot of testing of NetCDF. They found that it was very inefficient. I heard some strong words about the situation. It is interesting to me that many people push hard for these solutions, yet they do not make the necessary tests.

NEW CALCULATIONS OF CLOUD DATA

On June 30, Bill Rossow from GISS in New York visited. He has been doing the calculations of clouds for ISCCP. New procedures are almost developed. By about April 1996, we should have new clouds for Jul 1983 to Jun 1995 (12 years). This will also extend the present record by 4 years.

Bill is also going to get me a present--the calculation of radiation at the earth's surface, based on the satellite data. We have been talking about this for about 3 years.

TRIPS

I have not needed to go on a trip from July 25 through August. This has helped.

SORT PAPERS

In late August I spent a few days to get some better control on the papers that flow in at a fast rate.

September 1994

EOS PROJECT MEETING (Sep 7-9), TUCSON, AZ

Dickinson (University of Arizona) leads this interdisciplinary climate study for EOS. It has about 20 PIs, most from NCAR, some from JPL, and some from universities. We met for 2 days to get our various science and data plans better formulated, and linked better into a true cooperative whole. (Jenne is one of the PIs.) The daytime temperatures were around 13
103 degrees F. The summer in the western U.S. has been very hot. A text is available that describes the EOS project.

COADS STATUS

The ship data project is coming along well, but it sure is a lot of work. Woodruff, Worley and Jenne met Sep 13 to review the status and discuss the issues. In 2 or 3 months we will need part of the new data for reanalysis.

ASSESSMENT STUDIES, IPCC, TRIP TO ENGLAND, SEP 17-22

The IPCC-2 group for climate assessment studies has about 150-200 PIs around the world working on text for the 1995 report. There was trouble with some of the activities of a group in England last winter (the data was late and not what people needed). The IPCC people in Washington, DC, encouraged me to go this meeting Sep 19-21 in Norwich, England. It was useful to go. There were several good talks about models, and data "downscaling." There was a separate meeting about the IPCC issues. They were not resolved.

CD-ROM UPDATE WORK WITH THE UNIVERSITY OF WASHINGTON

The CD-ROM with long (40-year) N. Hemisphere analyses has been very popular. We are starting the work to update it and to add some more types of data.

One issue is whether the record of NMC analyses has real or virtual temsips in different periods. Dennis has done many diagnostics to almost answer the questions. Jenne has prepared text to describe this issue.

STATUS OF REANALYSIS

NMC has completed the analyses for 1985 and 1986. On Sep 26 they were nearly ready to start analyses for 1987.

NMC had a computer crisis--no money to pay for the Y-MP8 they are using. It will go away in May 1995, and they will get two Cray ELS (Jedi's), with 16 processors each. This will give them even more computing power than now. A Jedi processor is 0.6 of a Y-MP processor.
GEWEX (GCIP) AFFAIRS

NCAR will handle the regional mesoscale model (Eta) for GCIP--on 40-km grids. It is a lot of data. I'm afraid because of too many big projects. John Leese called from the project office in Washington, DC. He will meet with me the week of Oct 31, and I will go to a meeting in Washington, DC, on Dec 14.

October 1994

GODDARD EOS DAAC USER MEETING (Oct 3-4)

The DAAC is now in much better condition than a year ago. They have a number of datasets in the system. The staff is getting fairly large; 8 civil service (CS) and 51 FTEs from contractor. The budget for FY95 is $4.48 million and this does not include the salaries of CS people. It includes computer and storage hardware and maintenance. Only about 100K of this budget is for pass through funding for other organizations.

There was some information from the main EOS data contract shop. There were a couple of encouraging things (that indicated knowledge), but I did not get a good feeling that they know how to deal with the coming EOS data problem.

BOB CURRAN LEAVES HUGHES EOSDIS PROJECT

Bob Curran was chief scientist for Hughes. He left recently for at least a year. I was told by well placed sources that what is being said is that he was asked to leave. His job was to get input from scientists (and many of them have been unhappy with EOSDIS). The source said that Hughes needed a scape goat for the troubles. This interpretation seems plausible, but is not verified.

DATA MIGRATION IN NASA

NASA has been designing data systems that always require a data migration, whether or not the data is already on a local mass storage system with a reasonable file structure.

During 1980-91, when the climate data system was operational, the data from PIs had to be copied into the NCDS data system and usually the format was changed. During 1993-94 the DAAC copies data into its system and usually the format is changed. The version 1 data system is designed so that all data has to be copied into their system (again!). It isn't good enough that it is on an accessible mass storage system. People do not design adaptive data systems. This is not that hard to do!
VISITED JIM GREEN, GODDARD NSSDC DATA CENTER

Their center is for space and astronomy data. They did have the earth science data too, but it will be sent away. A number of datasets have been sent to the DAACs. More of the datasets will go to the DAACs. The rest will go to NOAA in 2 years.

JUPITER COMET CRASH

Three places, including JPL and Goddard, captured comet crash images from many sources. Jim Green thinks that their archives are the most complete. When the news was hot, Goddard NSSDC was getting 6000 accesses per hour for comet information.

RADAR DATA FOR THE SUMMER OLYMPICS

Dan Smith called from the NWS southern region office in NWS. There is a need to make a study of the climatology in Georgia for the olympics in summer 1996. They need to do work on a PC. We have the data.

THE VALIDITY OF CLIMATE MODELS

On Oct 12-13 I attended a meeting in Washington about the validity of climate models. It was held because Congressman Dingle from Michigan asked the general accounting office to make such a study. So the climate science office set up a meeting of about 15 people to discuss (and write) on this issue such that 3 or 4 people from GAO could listen and get ideas for their study. I found it interesting. The first day there were fireworks. Two of the strongest anti-greenhouse scientists (Richard Lindzen of MIT and Pat Michaels of the University of Virginia) in the country were there. There were some strong debates. The chairman from Penn State University (Eric Barron) was about to give up after 3 hours, but it finally settled down. It was agreed that people with contrary opinions will be able to write them into the report.

A person there (Bob Watson) who is very close to the science advisor to the President said: "This administration is absolutely committed to having the year 2000 emissions the same as 1990." I doubt that this is possible unless the U.S. economy really slows down. It is very likely that these arguments will heat up more.

Dingle tries to protect the automotive industry and they are often in the line of fire on this CO2 issue. In fact, we do need cars as gas efficient as we can make them.

I have been wondering if there are quasi reliable ways to estimate the amount of CO2 millions of years ago. Eric Barron says there are four ways: estimate weathering rates, soil column, ocean sediments and one more.
STATUS OF EOSDIS

It does not seem that NASA is getting a good grip on the EOS data problem. The talks seem to emphasize the whole system without a sense that people have determined the critical components, how to manage them, and how to integrate them.

Mahlman (GFDL) and Barron (Penn State) are now on a NASA committee about the data subject. I talked with them Oct 12 at another meeting. Their worries are similar to mine. Mahlman also commented on some of the huge data rates being planned and shares concerns about whether people have a sensible plan to handle it. People are also concerned about the cost. I suspect that this whole thing will heat up more.

November 1994

GEWEX AND GCIP MEETING, NOV 1-4

About 70 people were at NCAR for these meetings. I attended 60% of the sessions. On Nov 1 from 7-9 pm, four of us discussed the mesoscale Eta model data that will soon come to NCAR. We have to use it to support GCIP. The volume is about 5.5 GB per month. I also had a meeting about GCIP with John Leese on Oct 31.

REANALYSIS AND COMPUTING IN JAPAN

On Nov 3 I met with Nobuo Sato from Japan. He is chief of the numerical modeling and analysis branch of JMA. He said that it is 60% sure that they will want to do a reanalysis. He would need data from us. He now needs our dataset of CAC daily precip for 7000 stations.

In Aug 1995, JMA will obtain a new Hitachi computer (4 processors) that is twice as fast as a C-90 with 16 processors. It will become the operational computer in Mar 1996. The contract costs 95 million yen each month for 5 years (about $1 million per month). This includes networks around Tokyo, workstations, etc., too.

RUSSIA AND RAOBS; TRIP TO ASHEVILLE, NOV 6-11

On Nov 7-8 I met with Russian counterparts to negotiate the next year of activities in the U.S.-Russian data exchange. A document is available. They have a new chief and he is good to work with.

On Nov 9-10 I met with an advisory panel about Asheville's CARDS project (raob data). A year ago they had not accomplished much. This year they have made progress. NCAR has a similar project to prepare data for reanalysis. I would not trade our
situation on raob data with them. We are preparing many more types of data and our costs are much lower.

WORKSHOP "ADVOCACY FOR SCIENCE," NOV 14

This workshop was organized by John McCarthy and others in UCAR. There were speakers from Washington, D.C. The main themes were (1) what is going on in the federal budget and (2) how do people effectively talk with congress (and government).

On Nov 8 the U.S. voted in a landslide for Republicans. This will change congress. There was some information about likely policy changes, but it is too early.

Note that 14 inches of snow fell between Sunday evening and Monday morning (Nov 14).

TRIP TO GREENBELT, MD, NOV 21-22 (University-NASA)

I attended a panel meeting of USRA that manages the University-NASA set of programs for earth sciences. Part of the news is very tough: the Goddard science divisions are being asked to slim down by 20% this year (FY 95) and probably another 20% next year.

NAS PANEL ON ARCHIVES; MINORITY REPORT

I haven't been too happy with some key recommendations in the National Archives panel report—the report asks for a new top organization to set policy for all national science and technology data. On Sunday I called John Perry at NAS. On Monday noon (Nov 21) I called Hadeen at NCDC. On Monday morning I wrote a minority report, and on Monday evening I talked by phone with NAS staff (11:20 to 11:59 pm near Washington, DC). The key staff is almost glad I did it. They may rewrite the report. One or two other panel people had worries. Several reviewers worried a lot about the recommendation.

REVIEW TEXT ON RAOBS

I was asked to review a paper about the CARDS raob program ($1M per year) at NCDC. I first looked at it on Nov 23. The paper is difficult to review, because it has a bunch of problems.

December 1994

TRIP TO WASHINGTON, D.C., (Reanalysis, GWEX, NRC)

For this trip I needed to prepare slides for reanalysis, slides for a GCIP talk, and to study a lot on the NRC text on national archives.
Our annual panel on reanalysis met on Dec 12-13 at NMC. I talked about data inputs and later about data outputs. The program has some problems, but it is going well. A person from Livermore labeled it MOAR (mother of all reanalyses). One present problem is that the NMC decode did not include 80% of the GTS sig level winds for 2 years. We need to quickly see if they are in the Navy decode of GTS (we have it at NCAR).

GCIP EXPERIMENT

NCAR has a model data role. I met with 10 people in Washington, D.C., (Silver Spring, MD) about this problem on Dec 14. I'm worried that this will be a fair amount more work, and there is little time.

NATIONAL ARCHIVES PANEL

On the evening of Dec 13 I met with NRC staff from 6:00 to 10:00 pm about the panel report. Earlier I said I would have to write a minority report. Paul Uhlir (NRC) has been working hard to revise the report. At the end of this discussion, Paul said that he agrees with me, but is not sure that he can persuade the committee. I feel more comfortable now, regardless of how it comes out. It has been difficult to think intensely about these national issues, handle several huge science projects, and also be affected by NCAR politics.

During the trip there was no free time:

MAIN ACTIVE BIG PROJECTS

+ Reanalysis input data - intense amount of work
+ COADS - fairly intense
+ Reanalysis output - intense
+ GCIP project (Eta model output) - getting worse
+ Climate model data for U.S. Country Studies program - becoming intense
+ Satellite data issues - no time, so these stack up.

WIND SATELLITES

On Dec 13 I talked with Wayman Baker of NMC. He said that the USAF may be willing to provide the satellite launch and a satellite bus (all but instrument) for this. This is not yet a sure thing, but it's very hopeful.
MEETING ABOUT COADS

Woodruff, Worley and I met about the present set of COADS issues.

STATUS OF REANALYSIS

I prepared a main text about reanalysis about Dec 8. It tells the users a lot about the datasets that are included. At this time, 2 years of reanalysis were on the mass store. Each year has about 51 Gbytes of data, but it does not yet include 8-day forecasts. NMC has also (by Dec 13) prepared a draft paper about the reanalysis project.

REANALYSIS DATA ON THE MSS

As of Jan 3, 1995, 4 years of data (1985-88) are on the MSS. By Jan 4, inventories will be complete. One file is bad and will have to be replaced. NCAR receives 144 cartridges (2x density) per year of reanalysis. One cartridge will be replaced.

VACATION

I finally took some vacation, Dec 18-31.

January 1995

MEETING OF THE SATELLITE PANEL (SAR data), Jan 9-10

I attended a meeting of the NRC CES panel in Irvine, Calif. The purpose of the meeting was to start a study to advise NASA on what sort of SAR radar missions to plan for. SAR can be used to measure sea ice drift (done now), derive motion of ice cap ice, obtain ocean currents, obtain biomass, etc. People at JPL will prepare some cost trade-off information for me. I prepared some information about data issues that need to be resolved.

JPL DATA PLANS, Jan 9

Frank Carney from JPL said that they had planned for a DAAC (data system) size of about 20 people, which would then decrease some in size after it was set up. The EOSDIS plan was that the size of JPL would grow to 90 people. Now the plan is for the JPL DAAC to grow to 30 or 35 people. A lot of effort will be required just to talk with the EOSDIS contractor. He said that "you (NCAR Data Support) are always used as the example of someone who can get a lot done with only 2 people." (True, but there's some exaggeration).
MEETING OF EPA AND ARGONNE COUNTRY STUDIES TEAM, Jan 12

About 15 of us met in Washington, DC, on Jan 12, to help coordinate the program and prepare for the workshop with 30 new countries in Hawaii in early Feb. The Country Studies program is popular in the government, but they are worried about future decisions by Congress. A few people do not like the idea of funding global change activities.

HOW GOOD ARE CLIMATE MODELS?

Congress asked for a GAO study to evaluate how good climate models are. In Oct 1994 a forum was held, chaired by Eric Barron. A text is now complete. It will soon be published, probably by OSTP.

TROUBLE IN THE NEW NWS OBSERVATIONS

A meeting was held at AMS-Dallas on Jan 15 to discuss the data problems. The measurement of Precipitation is the main problem, but there are others. In good weather conditions, the measured precip with the new equipment is about 5% low; in snow conditions it is about 30 to 40% low. All the top NWS brass were at the meeting in Dallas.

COUNTRY STUDIES

Jenne prepared a paper for the Feb 1995 workshop. It has updated information about climate model data, new plans for climate model runs and information about climate science to help users understand the data. About 30 new countries are attending this workshop to help start programs. In Feb 1994 the first 18 countries started.

February 1995

COUNTRY STUDIES

People from 30 countries attended this workshop in Hawaii to learn about how to use models for crop growth, forests, rivers, rangelands, etc. They also learned some about the climate models, and how to obtain model output for any part of the world. Dennis Joseph and I attended Jan 28 through Feb 6. I gave a talk about the climate models and led one of the workshops (with E. European countries) on how to extract desired data using software by D. Joseph. Dennis gave a talk and led the workshop of African countries.

INTER AMERICAN PROGRAM

On Feb 9 I gave a talk about data policy and data exchange possibilities for interamerican work. It was good to see some of
the people who I met earlier in Uruguay. What we need are easy ways to exchange data. I fear that the people who like systems will define difficult methods that will hamper any data exchange.

GCIP-GEWEX

I went to meetings on Feb 22-23 in Boulder about data for GCIP. I gave a talk about the status of the Eta model data and used about 3.5 days on this topic. I'm trying to sort out the status of Eta and the plans. Not done yet.

DATASET OF WORLD RAOB DATA

On Feb 28 I was part of a conference call about raobs. We are preparing them for reanalysis. NCDC has a CARDS project. They have been getting about $1M a year for the project, from DOE. Money runs freely and people use expensive methods. Their software has problems. DOE will quit funding them. They want me to be part of a big project to help justify NOAA funds ($0.5M available). I will try to duck. They have not been good to work with.

OTHER

I attended the ARG meeting on Feb 21.

March 1995

TRIP TO GODDARD AND NMC, MARCH 5-9

NASA had a workshop to discuss their reanalysis (1985-90) on March 6-8. There were a lot of papers to compare reanalyses with each other and with other observations. I gave a talk about data inputs, and the status of the output files. A short 3-page version will be printed. I also visited NMC for a day to talk about reanalysis and about the mesoscale Eta model.

PROBLEMS WITH SSMI WINDS IN NMC/NCAR REANALYSIS

We have known for some time that when SSMI ocean surface winds became available starting July 1987, they seemed to cause a "climate jump" in some reanalysis variables. On March 7, NMC found that the speed seemed to be biased about 2.0 m/sec high because of a confusion between an antenna temperature and a brightness temperature. It will take 2 months to sort out a cure; in the meantime, NMC will do 1982-84 and NCAR will quickly send the data for those years. Later, Jul 1987 to Dec 1991 4.5 years) will be done again. On March 24, I contacted the main expert on SSMI (Wentz) and made two pages of updated notes on what data are available.
GETTING READY FOR MESOSCALE ETA ARCHIVES

On Feb 17 I wrote a draft paper about ETA with updated model and data volume information. The volumes are getting large and some of the NMC plans are changing. NMC requested that I visit to sort out the issues and firm up plans. (I did visit NMC on Mar 9.) NCAR has ETA data for summer 1994, and the new flood of data will start May 1995. A text is available. In March I spent about 3 days on the new text.

CAN NASA SAVE MONEY ON INFRASTRUCTURE AND FACILITIES?

Joe King (Goddard) told me (on March 8) that NASA is trying to evaluate whether they can save enough money on the cost of infrastructure that they can keep the main projects going, even though budgets are reduced. What should be contracted out? How can money be saved? The facilities include functions like data centers, computing, libraries, communications, etc. Joe is head of the Space Science Data Center. He asked what I thought the core functions of a data center should be. I think it is possible to define core functions and roughly how to constrain costs. Cost comparisons are needed in the U.S. and Europe.

SEND NMC THE DATA FOR 1982-84 QUICKLY

On Friday 10 March, four of us at NCAR met to more precisely describe what data to send NMC. By 23 March, NMC had a total of 14 years of data (1981-94) from us, plus a few parts of the earlier data.

THE U.S.-RUSSIA BILATERAL

Gus Shumbera visited (from NCDC) on March 27 to talk about this program. Two of us will visit Russia about July 8-15 to plan next year's program. The USGS is preparing river flow data for us.

NEW ISCCP PRODUCTS (clouds and radiation)

I got new information from Rossow at Goddard Institute in New York and wrote text on the status of these (clouds and radiation) data (dated March 13).

USE DRONE AIRCRAFT (March 28)

Kuettner called about a possible project to measure things using drone aircraft at 60,000 feet, at 100 locations around the world. But they can't stand winds over 150 knots. He wanted probability statistics about high winds. We have atlases and datasets that would help, but some processing would be needed.
OUR EOS RESEARCH PROJECT (Dickinson project)

I have prepared several items for our project (has about 20 PIs). These were sent to Arizona.

+ List of data that would help our pilot research project
+ Update of my science plans to March 1995
+ Update of timeline for the project
+ Brief answers to their 25 questions about plans and progress
+ Plans on data flows are poor. Wrote a text "The Big Problems of Using Huge Datasets." Now members of our project say we want 1000 Gbytes of data per day, and will give back 300 Gbytes of products. The planning has been handwaving so far. I showed the text to Sloan and Harano.

+ I prepared a list of steps that are needed to help figure out how to handle the EOS data.

+ Our needs for this project are to (1) constrain data volume where possible (formats, sampling, compression) and (2) to handle the necessary data flow at good costs.

TRENDS THAT MAKE DATA ACCESS HARDER

In the U.S. (and in UCAR) there are some trends that will make data access harder. Planners do not seem to know or care. I wrote two texts to help clarify these issues. These issues are a drain on my time that impacts reanalysis, GEWEX, and climate assessment studies.