Status of Reanalysis

Roy Jenne
National Center for Atmospheric Research

October 1995

• The text is mostly about the NMC/NCAR project to reanalyze global data for 40 years.

• Some information is also given about the status of projects at ECMWF, Goddard, and COLA.

• A partial description of similarities between data inputs at NMC and ECMWF is given.

• The status of sending data from NCAR to NMC is given. Many more details are in other texts.

• The volume of data in several popular sets of reanalysis output data is shown.

• The first CD-ROM for reanalysis will arrive about Dec 1995.
# Table of Contents

(A Selected List)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Status of the Project</td>
<td>1</td>
</tr>
<tr>
<td>2. Availability of Data on CD-ROMs and Tape</td>
<td>1</td>
</tr>
<tr>
<td>6. Surface Temperature Too Hot Locally?</td>
<td>3</td>
</tr>
<tr>
<td>11. TOGA COARE Raobs</td>
<td>4</td>
</tr>
<tr>
<td>12. Sea Ice Data for Reanalysis</td>
<td>5</td>
</tr>
<tr>
<td>14. Sea Surface Temperature</td>
<td>6</td>
</tr>
<tr>
<td>17. The TOVS Sounder Data</td>
<td>7</td>
</tr>
<tr>
<td>18. VTPR Satellite Sounder Data, 1972-79</td>
<td>7</td>
</tr>
<tr>
<td>19. Cloud Drift Winds from Satellite Data</td>
<td>7</td>
</tr>
<tr>
<td>20. How Does the NMC/NCAR Reanalysis Compare with NMC Operations</td>
<td>8</td>
</tr>
<tr>
<td>29. Status of Reanalysis at ECMWF</td>
<td>10</td>
</tr>
<tr>
<td>31. Volume of Data from ECMWF Reanalysis</td>
<td>12</td>
</tr>
<tr>
<td>32. The NASA Goddard Reanalysis Project</td>
<td>13</td>
</tr>
<tr>
<td>33. A Pilot Reanalysis Project at COLA</td>
<td>13</td>
</tr>
</tbody>
</table>

Attachment 1 - Previous Status Documents
Attachment 2 - Loss of Some Aircraft Data
Attachment 3 - The Volume of Output from NMC/NCAR Reanalysis
Attachment 4 - Analyses from NMC
Attachment 5 - Status of Sending Data to NMC
Attachment 6 - Inventories of Input Data for Reanalysis
Attachment 7 - Russian Ship Upper Air Data
Attachment 8 - Does a Surface Algorithm Change Affect the Upper Air?
Status of Reanalysis (Aug 1995)

This text gives the status of the NMC/NCAR reanalysis project on 2 Aug 1995. A little information about the ECMWF project is included, plus some information about Goddard and COLA.

1. Status of the Project

   The NMC/NCAR reanalysis is going well. However, people keep running into some new problems to solve. This information is from Bob Kistler at NMC on 2 Aug 1995.

                 — All data for 1982 and Jan 1985 - Oct 1993 are done (total of 9.8 years).
   7 Sep 1995:  — Data for 1982 through 1993 are done, but not 1984 (11 years are done)
   Oct 1995:  — 12 years are done.

   • Now (Aug 2) 9.8 years are done
   • On Sep 7, 11 years are done

   • NMC needs to finish developing code to be fully operational on the new computers. These big tasks are progressing well. The Cray Y-MP being used now will leave, probably on Sep 28, 1995.

   • Then NMC has to re-do data from CDAS for 1994 and 1995. Also they will do 1984. So before long 1982-94 (13 years) will be done.

   • NMC needs SST grids from the UK in order to do reanalysis for 1981 and earlier. (The global weekly, 1° grids by Dick Reynolds start Nov 1981.)

Note: There was a small glitch in aircraft QC for the first half of 1982 that NMC did earlier, so they did reanalysis for that again (took a week).

2. Availability of Data on CD-ROMs and Tape

   NMC is preparing CD-ROMs that have a popular subset of the reanalysis data. There will be one CD-ROM for each year of data. The first one (for 1985) will be ready about Dec 1995. For information, the practical amount of data for a CD-ROM is about 660 MB. By about 1998, we expect a new standard technology that will hold 6 to 10 times this much data. The standards are still being argued. The CD-ROMs with reanalysis data will be available from NCAR and NCDC. The cost will probably be in the range of $15-$20. There may also be an initial mailing from NMC.
3. **Availability of Data from NCAR**

- Data for 1982-93 has been analyzed, 12 years are done (Oct 1995).
- NCAR has data for 1985-93 (9 years)
  - Data for 1985 will all be replaced (model spinup was slower than we hoped)
  - 1985-93: Pressure coordinate data and statistics that were hurt by the polar wind calculation will be replaced.

These data have been defined as a bulk transfer set, where we charge only $50 per Gbyte. (Our standard price, with volume discount, is about $450 per Gbyte.) The cost will probably come down a little more when we are set up to handle bulk copies in a better way. NCAR also has data from the "with SSMI winds" analyses for Jul 1987 - Dec 1991. It is expected that most users will not use these data. We do not advertise them.

We discuss the problem of polar wind calculations that will be fixed (in the pressure stack) and the problem of surface temperature (too hot) for calm winds. The resolution of the wind issue will cause some changes in the present archives.

4. **NMC Update on 16 Oct 1995**

NMC re-ran data for all of 1985 because of the time that it takes to spin up the soil moisture and the stratospheric water. The data at NCAR will be replaced.

Data for 1982-93 (12 years) is ready, but several datasets already at NCAR are affected by the problem with the derivation of polar winds. These pressure level fields and associated monthly statistics will have to be replaced. We think that this data replacement will start about mid Nov 1995, and it will take some time.

We will discuss the effects of calm winds at the surface on surface temperature and heat (and radiation) fluxes. The program will be changed to fix this problem. The years 1994 and 1995 will be run with the fix, and 1957-81 will have the fix.

Especially during Jun-Oct 1995, NMC has been doing much extra work because of the change in computers. The Y-MP was shut down 12 Oct 1995. There have been some data flow problems that are frustrating and serious. The computer periodically lost a block of data (when tapes are spanned). More recently, some of the data reads are wrong. The data has been altered and the engineers do not understand the problem yet.

SST: In mid Oct 1995, NMC received the new sea surface temperature data from the British Met Office.

5. **Status of Sending Data to NMC**

Much information about the preparation of many datasets for NMC is given in Attachment 5.
6. Surface Maximum Temperature Too Hot Locally?

In mid-Aug 1995, NMC found that some of the surface daytime temperatures may become too hot over land when the wind speed is very close to zero and it is sunny. In this case, the surface parameterization formulation doesn’t have good enough methods to carry heat away. This can produce some temperatures that are much too hot in a few tropical points. This can occasionally cause periodic local problems such as 50°C over the Amazon and 80°C over Saudi Arabia. NMC is trying to decide what to do. The first information was that in some places, such as the Amazon, the very high surface temperatures could appear day after day in some seasons. Fortunately, this is not the case. However, people say that there is a small climate jump at the surface if the algorithm is changed. Attachment 8 discusses some of the comparisons and tests that have been made. The good news is that the statistics of the upper air data and precipitation do not change when the surface algorithm is changed.

Status on Sep 26, 1995: NMC will make some new reanalysis tests for a few months using a weaker version of the change in the surface algorithm than the one used in the first tests. This weaker version is the one that will soon go into operational production.

- 12 years of reanalysis have been completed.

This problem will also affect the surface heat, moisture, and radiation fluxes. The algorithm will be changed in reanalysis to avoid these surface temperature and flux problems.

7. Winds in the Analyses (pressure coordinates) and in Statistics

There was a problem in the calculation of winds near the poles. This affected the winds from about 82° or 85° latitude to the poles, in the analyses on pressure coordinates. All of these grids will be replaced by corrected grids. This also affected the time series of analyses that are about to go on a CD-ROM for 1985. The archive of statistics will also have to be replaced. Also the pressure stack of grids from the forecast will be replaced. Fixing this will delay the CD-ROM to about December 1995.

Another time delay:

- In mid-Sep 1995, NMC found that the mass storage gear had dropped some blocks of data from files. The problem happened 12 times. The problem stopped but it wasn’t fixed.

- When will the files with bad grids be fixed? The earliest that NCAR will start receiving new files is mid Nov 1995.

8. Soil Moisture

NMC started reanalysis with Jan 1985 and came forward in time. They started with December of the previous year to spin up the models. Kanamitsu said (19 Sep 1995) that
they are noticing that the soil moisture was not spun up very well, and shows a jump in the statistics from 1984 to 1985. It would probably take 3 or 4 months to spin up properly. Also it takes more time than a month to spin up the stratospheric moisture. Another start was made with Dec 1981 data. This was a "warm" start so 1982 data should be okay.

The reanalysis procedure nudges soil moisture toward climatology. The Mintz and Serafini climatology was used.

9. **Comment on the Recent Problems (19 Oct 1995)**

The issues about the problem in surface fluxes have been resolved so that the project can move forward again. A considerable amount of data will be moved to fix the wind problem (the primary outputs from reanalysis do not have to be changed). The problems associated with the change of computers at NMC have been more severe than anyone anticipated. We are hopeful that the computer hardware problems will soon be fixed so that the reanalysis can move forward quickly.

10. **Precision of Vorticity and Divergent Fields**

NMC is considering whether to add more precision to the vorticity and divergence fields in the pressure stack of grids. Trenberth's groups was using the sigma analyses when they brought up the question of precision. Should there be enough precision in vorticity and divergence to transform back to U and V within a precision of 0.1 m/sec? The U and V are also in the archive. The following message was from Kevin Trenberth's group on 1 Sep 1995:

> Kevin and I are concerned about the precision of the vorticity and divergent fields from the NMC reanalysis grbsanl file. (For reference I was looking at the .3125 level in grbsanl85010100 from A00016.) It appears that all the values have been rounded to 6 decimal places. As a result, values on the order of 10E-6 have only one significant digit (and 10E-5, 2 significant digits ...; maximum values of vorticity are slightly over 10E-4 and have 3 significant digits.)

\[ r_{\text{max}} = 2.83000 \times 10^{-4} \]

\[ r_{\text{min}} = -1.91000 \times 10^{-4} \]

Doug Lindholm
Climate Analysis Section

*My comment:* The larger values now have three significant digits. I doubt that it is practical to add enough precision to derive winds from vorticity and divergence. However, we note that the archives do include a dataset with 32-bit precision. NMC has increased the precision of the vorticity.


NMC used the TOGA raobs. They had trouble with too many levels in these raobs, so they quickly thinned the levels to get to a practical number. Most people think of the TOGA
COARE period as 1 Nov 1992 - 28 Feb 1993, the intensive observing period (data used in reanalysis). There were enhanced observations for the 1-year period 1 Jul 1992 - 30 Jun 1993.

12. Sea Ice Data for Reanalysis

The sea ice that has been used for reanalysis is the daily product calculated directly from SMMR (starts Nov 1978) and SSMI (starts July 1987) microwave data. The alternative weekly option from the Joint Ice Center has not been used. Summary of ice data used:

- Daily ice grids for Dec 1978 - 26 Dec 1991 are the same for both NMC and ECMWF.
- Both NMC and ECMWF have used this same ice product. After 1991, the sources will have only minor differences between NMC and ECMWF.

The Snow and Ice Data Center (SNIDC) has worked with Goddard to prepare daily grids of ice concentration, about 25-km resolution. These data have been used in reanalysis. Monthly ice grids have also been calculated. The daily ice concentration grids for Nov 1978 to near present are available on two CD-ROMs. The reanalysis actually uses ice cover instead of concentration. The concentration is used to derive ice cover. ECMWF also uses ice cover.

For SMMR, the instrument was on one day and off the next. Per Gloersen (Goddard) took the data for the active day and calculated the sea ice concentration grid. So the CD-ROM has sea ice concentration grids every other day. It isn't entirely a regular schedule; sometimes there are extra days and sometimes fewer days.

ECMWF processed the CD-ROM ice data for Dec 1978 through 26 Dec 1991, and prepared daily, 1° grids of sea ice concentration. These are being used by both ECMWF and NMC for reanalysis. To obtain data past Dec 1991, NMC started with basic SSMI data from other sources and calculated similar 1° grids of sea ice concentration. They made calculations for a month of overlap period, to verify that their calculations were very close to the ECMWF data.

ECMWF compared sea ice data derived directly from microwave data with the weekly product from joint Navy/NOAA sea ice center. They strongly favor the direct calculation of sea ice from microwave. See (Nomura, A.) for a description.

The present sea ice data is good, but Goddard is now (Aug 1995) working with the radiances to take out some small inconsistencies over the 16-year period. A new daily ice concentration product may become available in the next year or two. (This information is from Ron Weaver, NSIDC, Aug 1995).


- Sea ice concentration, late 1978-on, 25-km resolution, on CD-ROM. See above text.
- Daily grids of concentration, 1° resolution, Dec 1978-on, prepared by ECMWF.

Earlier years: There is microwave data (ESMR) from the Nimbus 5 and Nimbus 6 satellites for at least Dec 1972 through Dec 1976 and there is some data for the gap period until Nimbus 7 starts in late 1978.

14. Sea Surface Temperature

The NOAA AVHRR (GAC) 4-km data exists from Nov 1978. The split window channel starts about Nov 1981; it permits a better SST to be derived. Dick Reynolds (NMC) obtained COADS surface marine data from NCAR in 1992. He used COADS (or operational GTS data) plus satellite SST (from GAC) to prepare weekly 1° fields of global SST. These are global fields that go to the edge of the sea ice. The sea ice fields used are the digital weekly, 2° fields from the Navy/NOAA Joint Ice Center. Note that the source of the sea ice data for SST is different from the sea ice used for boundary conditions in reanalysis.

The weekly 1° SST by Reynolds (Nov 1981-on) has been used for reanalysis at both NMC and ECMWF. This weekly SST data from Reynolds (for reanalysis) used observations from COADS through 1989. For 1990-on, it uses the operational marine data on GTS, as obtained by NMC.

15. SST Data Before Nov 1981

Before Nov 1981, a better set of SST grids are needed for reanalysis than are now available. The UK Met office is working on grids, but a direct analysis using data that are usually sparse loses much of the Pacific equatorial SST pattern that is associated with El Niño changes. Reynolds has been working on methods to get the patterns right. Now Reynolds and the UK are working together. SST grids for the period 1957-81 will be ready by about Jul 1995, for use in reanalysis.

- Reynolds used the new methods to produce monthly SST data (2x2°) for 1950-92 using COADS. The output is for 66°N to 45°S (not global). More information is in the April 1995 status report.

- The UK will use these methods, and blend with ice edges to produce a monthly, global SST product.

- NMC expects to obtain these grids from the UK by 1 Oct 1995. Then they will start to analyze data for 1979 through 1982.

- ECMWF used the older SST data from the UK Meteorology Office (with problems in the tropics) for 1979-82. NMC will use the new version.
16. The Use of Surface Marine Data in Reanalysis

Both NMC and ECMWF are using COADS data for reanalysis, for 1980 through 1993. NMC has received the data from Boulder, put it into BUFR, and sent it to ECMWF. For 1994, NMC is only using data from the original NMC decode.

The FGGE dataset is for Dec 1978 through Nov 1979. ECMWF put the last version of the surface marine data (made about 1984) into the BUFR format and sent a copy of the data to NMC.

The FGGE year did not include data for Dec 1979. In Feb 1995, NCAR sent NMC COADS data for Dec 1979. For Dec 1979, ECMWF used COADS plus data from their own archive.

17. The TOVS Sounder Data

For reanalysis, both NMC and ECMWF are using the same set of 2.5° TOVS data produced by NOAA/NESDIS. This dataset has cloud-cleared radiances and derived soundings. The version being used is the set from the research archives at NESDIS/NMC, because it has the fewest gaps and the best date order. NCAR got this research archive from NMC, and also got a copy of the archives at Asheville. We made extensive inventories and used the NCDC set to fill a few gaps in the research set. Goddard obtained some more gap fillers from NMC daily dump tapes that they have. These were sent to NMC. There are still a few gaps. ECMWF has prepared a program to start from Level 1B TOVS from NCAR and calculate cloud-cleared radiance data, etc., to fill most of the remaining gaps. This work is now going on. ECMWF helped to pay for the copy of the NCDC archive. They got a copy of the research archive and the gap fillers via NCAR. I think they also got the gap fillers from NASA via NMC.

18. VTPR Satellite Sounder Data, 1972-79

VTPR sounding data is available for 15 Nov 1972 to 28 Feb 1979. The primary record has original radiances, cloud-cleared data, and soundings. There are gaps in the record, but nearly all of these gaps can be filled by using soundings on tapes from NMC that are at NCAR.

VTPR overlaps with TOVS during Nov 1978 through Feb 1979. This will permit NMC to run tests to compare analyses made by using TOVS to analyses using VTPR. On 31 May 1995, NCAR sent NMC samples of TOVS and can quickly send all of it when NMC wants it.

19. Cloud Drift Winds from Satellite Data

These data are used from the original NMC tapes at NCAR. Also, JMA provided a copy of GMS cloud winds from their archive for the period 1978-91. The JMA data were sent directly to NMC.
Rex Gibson says that they also received a copy of the complete archive of GMS cloud winds from JMA. A message from Rex Gibson (ECMWF, Apr 1995) says that they are using the GMS cloud winds from JMA from Apr 1981-on. These are the original GMS winds as calculated in Japan. They have not been revised or recalculated. The data copy received by ECMWF started by Jan 1979 or earlier, but they did not start inputting the data for reanalysis until data for April 1981. Since these are data from the source, ECMWF uses them in preference to any GMS data received earlier via GTS.

20. **How Does the NMC/NCAR Reanalysis Compare with NMC Operations**

The methods used for reanalysis are the same as the methods that went into production at NMC in Jan 1995. The production methods are still the same in mid-Aug 1995. At the end of Sep 1995, NMC hopes to put some changes into the production model. There will be some small changes in the physics. The big change is that for satellite TOVS sounders they will use radiances, not the inverted soundings currently used. The resolution for TOVS will be 80 km instead of 250 km now. NMC sees big changes in forecast skills, especially in the S. Hemisphere with these new methods. At the same time, NMC will fix the problem of bad surface land temperatures on sunny days when winds are calm.

The resolution for the NMC production runs is T126 and 28 levels. NMC would like to move to T170 and 42 levels, but this certainly won't happen before around Feb 1996 and even then it is questionable (the factors are tests, machine time, etc.). The resolution for reanalysis is T62 and 28 levels.


NMC got a new Cray C-90 (16 processors) in Dec 1992. Reanalysis has been using the old Cray Y-MP (8 processors) for production. The Y-MP will probably leave near Sep 20, 1995.

The two new Cray Jedi computers (each has 16 processors) got to NMC at the end of June 1995. Each processor is about 0.6 of a Y-MP processor. Each Jedi has a lot of memory—256 Mwords (64 bits each). This is 4 times the memory that is on the Y-MP (64 MW). The mass store links are working. NMC is testing code on these computers now (Aug 2).

Sep 7: Disks have been moved from the Y-MP to the Jedi, but Kistler will run a little longer on the Y-MP. The Y-MP will leave about Sep 28, 1995.

22. **The Budget for NMC/NCAR Reanalysis**

NMC is taking some hits on budgets, but it appears that reanalysis may go on roughly okay. However, some of the funds come from Mike Hall's budget (Climate and Global Change). Hall's total budget was $71M for FY95. In July 1995, the bill that removes $14M for FY95 was signed by Clinton. This leaves about $57M for this year (FY95). People guess that the FY96 amount may be around $53M. There seem to be special worries about FY97.
NMC met with Mike Hall's office on Aug 2 to obtain some information. The office recognizes that reanalysis is an important program. Money is assured through 1996. They will want to see a more extensive proposal text by summer 1996, concerning future work.

23. **Government Shutdown on Oct 1?; No, but maybe on Nov 15**

There is a big worry (now Sep 7) that the government may shut down on Oct 1. Clinton told employees to prepare for a 60-day shutdown. NMC has just been declared as essential, so all of them would keep working.

24. **Previous Status Documents about Reanalysis**

A list of previous documents about the status of reanalysis is given in Attachment 1. A little information about the content of the documents is included. A few related documents are also listed. There are many other documents about various types of data that are not listed here.

25. **Aircraft Data from Japan (1978-94)**

NMC received Japan's complete file of global aircraft data for mid-1978 through Dec 1994. Both the NMC data and all of Japan's data were input. For duplicates, NMC gave preference to keeping the NMC record. The Japanese data helped to fill in the coverage, especially over the Western Pacific. The combined data will be in the output files.

A message from ECMWF said that they also used aircraft data from Japan. I assume that it was for these same dates.

26. **Loss of Some Aircraft Data**

For reanalysis, about 3% to 5% of the aircraft reports on the original NMC tapes were lost for the period July 1987 - March 1991 (Attachment 2). This means that the input data (and not the reanalysis output observations) should be the starting point for any future reanalysis. The loss was for all synoptic times, but it was worse at 12 and 18 GMT. Probably the Japanese aircraft archives helped to bring back much of the "lost" data.

The description of the problem (in Attachment 2) also indicates that the problem came from a "problem waiting to happen," a complexity in the data format. Other users of the original data tapes should be aware of this.

27. **Archived Grids from NMC, History and Reanalysis**

The history of model resolution used for production at NMC is summarized in Attachment 4. This is compared with the resolution used in reanalysis. A little information about the data archives from operations and from reanalysis is also given.
28. **Inventories of Input Data for Reanalysis**

We have a series of online inventories of various datasets that will give the reader a feeling for what was used in reanalysis. All of the NMC surface and upper air archives at NCAR was fed back into reanalysis. Therefore these inventories can be used directly. However, NCAR also added data from many other sources, so the inventories of NMC data will not be everything used for reanalysis.

See Attachment 6 for information about how to obtain the inventories for datasets.

29. **Status of Reanalysis at ECMWF**

*Summary of the status at ECMWF in Aug 1995:*

- In early May 1995, 3.0 years were done (1979, 1980, 1981)
- On Aug 3, 5.6 years were done
- Sep 22, 1995: ECMWF recently finished a total of 7 years.
- Oct 19, 1995: Doing Dec 1986 (have almost finished 8 years)
- They plan to finish 15 years (1979-93) by July 1996
- In early Oct 1995, the ECMWF science advisory panel encouraged them to also do 1994 and 1995.
- The reanalysis programs are almost exactly like what was first put into operational production on 4 Apr 1995 and which is still being run in production in Aug 1995. However, operations has T213 resolution (63 km).

In Apr 1995, ECMWF noted that "reanalysis production is running smoothly, and we will complete 1981 early in May. We are averaging about 11 days per day." The system has 31 levels (sigma coordinates) T106 resolution (1.125° or 125 km). The T106 Gaussian grids have dimensions of 320x160 points, and T213 is 640x320 points.

As of Aug 3, 1995 (another 3 months), ECMWF has finished another 2.8 years, into Oct 1983, so the rate of doing the new analysis is staying about the same. So a total of 5.6 years are done. They have been doing about 11 months of production each month.

*Analysis methods and data (ECMWF, Apr 1995)*:

After careful consideration, ECMWF eventually decided to go ahead with re-analysis production using the OI system augmented by 1D-VAR for the cloud cleared radiance data, so we no longer intend to use 3D-VAR. The final production scheme corresponds almost exactly to the operational changes made at ECMWF on 4 April 1995 - moisture is handled in grid point space, dynamics spectral, fully interpolating Lagrangian, 31 levels, T106, with reduced N80 Gaussian grid for grid point computations. Also used is the new 31 level ECMWF cloud scheme, and full 31 level cloud data are archived.

The augmented data include COADS and 250 km CCR (cloud-cleared radiances, TOVS) as you mention. They also included additional TEMP and AIREP data from JMA, PAOB data
from Melbourne, and (from Apr 1981 onward) the complete archives of GMS cloud winds from JMA, Japan.

They have just produced the first locally regenerated CCR data from full NESDIS 1b TOVS data, using the system they hope to use for gaps in the 250 km CCR data. They have also, with help from NESDIS, been successful in recovering some of the mislocated satellite data by recomputing positions from orbital information.

ECMWF noted in Apr 1995 that the hard work of setting up stable production was complete, and that there might be more time to keep people better informed on the status of the project.

More information (Aug 3, 1995): The TOVS sounder 1b data (basic radiances) has been useful—they used it successfully for the October 1983 gap in 2.5° TOVS data, and plan to use it for the remaining gaps. But they are still learning, and will re-do the October case several times in order to refine the methods.

On Aug 3, ECMWF stated: "Our re-analysis production is currently running August 1984, so we have completed 5 years and 7 months. We believe our initial delay and period of experimentation has been justified, in that our system is more or less exactly that being run in daily operations today. Thus we expect to have the 15 years (1979-93) complete by July 1986, all with an invariant system, compatible with that used in operations in what will be by then the very recent past. ECMWF will change computers in Sep 1996, and the reanalysis system is not likely to work on the new computers." Note: ECMWF would like to do 1994 and 1995 also, but there may not be enough time. They hope also to rerun some months where the science was right but some things in post processing were different.

"We can now easily sustain 10 months of production per month. Our peak is 12 months in a month, and this is our target for the remainder of the production. We are producing a number of monitoring diagnostics as routine production validation—some early results drawn from these will be presented at the AMS next January."

How TOVS sounder data is used at ECMWF:

The use of satellite data is pretty complex at ECMWF, but is based 95% on radiances. Below 100 Hpa we use radiances only. Above 100 Hpa soundings are used, but only as a "first guess"; they are subsequently modified using the radiance data. ECMWF will send a paper at some time giving full details of how it is done.

Use of computers at ECMWF (Aug 1995):

- ECMWF currently uses about 10% of the Cray C-90 (with 16 processors) for ERA (reanalysis) to do the modeling—remaining work is done using 2 SGI crimsons with lots of disk space (~35 Gigabytes).

- The tenders for the C-90 replacement are currently (Aug 1995) being evaluated. The replacement machine is due June 1996, with a 3-month overlap.
• ECMWF will obtain a new computer with about 5 times the power of the C-90 (with 16 processors) in June 1996, and over 20 times the power of the C-90 in 1998. The formal decisions and announcement have not been made yet.

• So for the ERA (reanalysis) project the heat is really on - we MUST complete before the C-90 is taken away in September 1996!!

Two comments: ECMWF hopes to do 1979 and 1980 again if they have time. They will use the SST analysis chosen by NMC (the new one).

Note: Much of this information was from Rex Gibson, who is the project manager for reanalysis at ECMWF.

Will ECMWF analyze data for earlier years too? They still have these goals, but the main thrust is to complete 1979-on. To do earlier years, they would need to convince their council. Also, TOVS satellite data is a significant help in the analyses, and they hate to give that up (not available in early years).

30. Reasons To Analyze Data for 1979 Again

There are three main reasons that ECMWF would like to do 1979 again. They are using analysis statistics to prepare bias corrections for raobs. These are applied above about 200 mb. They use this procedure for whole block numbers (such as N. America) where a significant area has similar raob types, so that they can get good statistics, and it is used only if there is a clear signal. These methods are used in operations. In reanalysis, they were used for 1980-on. So they hope to use the methods for 1979 also.

Second, in 1979 there was a small difference in the way that vertical advection was done compared with 1980-on. This should not make much difference. Third, ECMWF and NMC both wanted the data and the boundary conditions to be the same for 1979. ECMWF used an early version of SST from the UK Met Office. NMC will use the revised version that is better, especially in the tropics. This information was from Per Kallberg.

31. Volume of Data from ECMWF Reanalysis

The data volume of reanalysis from ECMWF will be approximately as follows:

1. Model level data - 13 Gbytes per year (31 levels) (includes surface fields)
2. Pressure level data - 7.5 Gbytes per year (15 levels) (includes surface fields)
3. 2.5-degree data (15 pressure levels) - 4.5 Gbytes per year (includes 2.5-degree surface fields)
4. Surface fields - 3 Gbytes per year model grid, 0.85 Gbytes per year on 2.5-deg grid
5. Climate and stats - 0.75 Gbytes per year

ECMWF plans to support 3480, 3490 (cartridge tapes), and Exabyte tapes.
ECMWF will have a few papers in the AMS next January (1996), including one from Saki Uppala making the point that they see better performance from a two satellite system than from one.

32. The NASA Goddard Reanalysis Project

The results of the Goddard reanalysis for Mar 1985 - Feb 1990 (5 years) are presented in ASA Tech Memo 104606, Vol 6, Apr 1995, 183 pp. The methods are described in Schubert, et al., *Bul AMS*, Dec 1993. The analysis is done on a 2° lat x 2.5° lon grid with 14 pressure levels (1000-20 mb) plus sea level. The forecast model has a resolution of 2° x 2.5° with 20 sigma levels. More information is on their WWW home page (http://hera.gsfc.nasa.gov/dao.home_page.html). If you email to data@dao.gsfc.nasa.gov, you will get a message that tells about the data and how to get it.

Output is each 6 hours for the upper air fields and each 3 hours for the surface (precipitation, OLR, etc.). There is an extracted marine set of fields such as wind, temperature, stress, etc. There is a set for chemistry that has all sigma levels, each 6 hours, but a reduced set of variables, U, V, T, etc. There are monthly means online.


33. A Pilot Reanalysis Project at COLA

COLA made a reanalysis for a 19-month period (May 1982 - Nov 1983). The resolution used was R40 and 18 sigma levels. The project and output is described in the May 1995 *Bul AMS*.

Reference

Nomura, Atsushi: Global Sea Ice Concentration Data Set for use with the ECMWF Reanalysis System. ECMWF TR No. 76.
Table 1. Output Data from Reanalysis
This table gives the name of the data files, the volume for 1985, and the type of data. See the text for more details.

<table>
<thead>
<tr>
<th>No.</th>
<th>File</th>
<th>Mbytes (1985)</th>
<th>Type of Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>grbsanl</td>
<td>4909</td>
<td>The GRIB analyses for users (most don’t use guess)</td>
</tr>
<tr>
<td>1.1</td>
<td>grbsf06</td>
<td>4895</td>
<td>Sigma analyses, Gaussian grid</td>
</tr>
<tr>
<td>1.2</td>
<td>pgb.f00</td>
<td>2533</td>
<td>Sigma guess, Gaussian grid</td>
</tr>
<tr>
<td>1.3</td>
<td>pgb.f06</td>
<td>2530</td>
<td>Analysis stack, 2.5° pressure</td>
</tr>
<tr>
<td>1.4</td>
<td>prs</td>
<td>2152</td>
<td>Guess stack, 2.5° pressure</td>
</tr>
<tr>
<td>1.5</td>
<td>prs</td>
<td>2152</td>
<td>Time sort analyses, 2.5° pressure</td>
</tr>
<tr>
<td>2.</td>
<td>grb2d</td>
<td>1989</td>
<td>The flux files (GRIB format) for users</td>
</tr>
<tr>
<td>2.1</td>
<td>2D</td>
<td>2606</td>
<td>Synop sort fluxes, Gaussian grid, GRIB format</td>
</tr>
<tr>
<td>2.2</td>
<td>2D</td>
<td>2606</td>
<td>Time sort, all of grb2d plus some 2.5° data from pgb.f00</td>
</tr>
<tr>
<td>3.</td>
<td>ipvanl</td>
<td>1462</td>
<td>Data for isentropic surfaces (GRIB format)</td>
</tr>
<tr>
<td>3.1</td>
<td>theta</td>
<td>1971</td>
<td>Analyses on theta derived from sanl (synop sort)</td>
</tr>
<tr>
<td>3.2</td>
<td>theta</td>
<td>1971</td>
<td>Data for 11 theta surfaces, 11 variables each (time)</td>
</tr>
<tr>
<td>4.</td>
<td>grib</td>
<td>269</td>
<td>Means and statistics</td>
</tr>
<tr>
<td>4.1</td>
<td>grib</td>
<td>269</td>
<td>Monthly means, variance of pgb.f00, grb2d and ipvanl</td>
</tr>
<tr>
<td>4.2</td>
<td>znl.f00</td>
<td>52</td>
<td>Zonal band averages of analyses (each 6 hours)</td>
</tr>
<tr>
<td>4.3</td>
<td>znl.f06</td>
<td>52</td>
<td>Zonal band averages of guess (each 6 hours)</td>
</tr>
<tr>
<td>4.4</td>
<td>optavg</td>
<td>30</td>
<td>Optimal averages (each 6 hours)</td>
</tr>
<tr>
<td>4.5</td>
<td>3D</td>
<td>192</td>
<td>3D heating, etc., monthly statistics from grb3d</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td>2972</td>
<td>Data for CD-ROM (6-hour data and statistics)</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td>2972</td>
<td>Forecast data. An 8-day forecast each 5 days</td>
</tr>
<tr>
<td>7.</td>
<td>preqnm</td>
<td>7276</td>
<td>Observed data with all metadata (BUFR format)</td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td>Four basic (32-bit binary) files follow</td>
</tr>
<tr>
<td>8.1</td>
<td>sanl</td>
<td>2697</td>
<td>All basic analysis data (sigma), spectral</td>
</tr>
<tr>
<td>8.2</td>
<td>sfcasnl</td>
<td>1692</td>
<td>All other fields for forecast run (SST, flux, etc.)</td>
</tr>
<tr>
<td>8.3</td>
<td>sges</td>
<td>2799</td>
<td>Like sanl, only from end of 6-hour forecast</td>
</tr>
<tr>
<td>8.4</td>
<td>bges</td>
<td>1549</td>
<td>Has all flux fields, etc., from 6-hour forecast</td>
</tr>
<tr>
<td>9.</td>
<td>grb3d</td>
<td>9381</td>
<td>3D heating, momentum, and moistening follows</td>
</tr>
<tr>
<td>9.1</td>
<td>grb3d</td>
<td>9381</td>
<td>Model diagnostics each 6 hours on sigma levels, GRIB format, Gaussian grid</td>
</tr>
</tbody>
</table>

Total: 54.01 Gbytes for 1 year
Attachment 1

Previous Status Documents

This lists the main previous status documents that were prepared by Jenne. There are many more documents about broad classes of data that are not listed here.

1. Observations for Global Reanalysis, 3 pp (23 Mar 1995)
   A short summary of the input data for reanalysis.

2. Some Reanalysis Data Issues, 5 pp (11 Nov 1993)
   Summarizes information about different types of data.

3. Arctic Raobs (12 Nov 1993)
   A summary of information about data that are available in the Arctic. This is important.

4. Considerations for Reanalysis, 4 pp (7 Dec 1994)
   Includes information about data and about early hand analyses that might be used to help S. Hemisphere reanalysis if the models get lost because of no satellite data in early years.

   This is still needed for its information about types of data available. This needs to be brought up to date. It also includes a section about the distribution of high-volume datasets.

   This includes a history of events. It has a short piece that compares data formats. Information about COADS update for 1992-93 is given. Has a helpful table about output products from reanalysis. Has lots of useful information about the output files from reanalysis.

   Has information about the low number of raob significant winds on original NMC tapes. The 26-month period affected was Aug 1989 to Sep 1991. NMC got raob data from NMC for this period and combined the two sources.

   Has information about data for reanalysis and data coverage. This is a set of briefing slides for a talk about reanalysis at Goddard (Mar 1995).
9. **Data Status from NMC/NCAR Reanalysis, 9 pp (27 Apr 1995)**

Gives the status of the project. Has information about the SSMI wind problems. Data for 1985 through 1991 (7 years) were done but data for July 1987-on will have to be done again. Changes in satellite sounder data, an overlap of sounders during Oct 1978 to Feb 1979. Bad locations in 2.5° TOVS. Information about early SST grids that are needed. Information about Australian surface bogus.

The original NMC tapes lost some important S. Hemisphere raob data during Jan 1992 - June 1993. ECMWF provided raob and pibal data for blocks 97 and 98 so that NMC could restore these data. There is raob filler data for 1979-93. This is described.

10. **Considerations for reanalysis (7 Dec 1994)**

This text reviews the coverage of different types of data versus time. One question is whether the analyses will get lost in the S. Hemisphere in early years. This reviews some of the early grid analyses at NCAR that might be used to help (4 pp).

11. **Another big paper (written winter-spring 1995)**

A paper about reanalysis (Kalnay, et al.) will be in *Bul AMS* in Feb 1996.
Attachment 2

Loss of Some Aircraft Data

For reanalysis, about 3% to 5% of the aircraft reports on the original NMC tapes were lost for the period July 1987 - March 1991. This means that the input data (and not the reanalysis output observations) should be the starting point for any future reanalysis.

The description of the problem below also indicates that the problem came from a "problem waiting to happen," a complexity in the data format. Other users of the original data tapes should be aware of this. The following is from Bob Kistler, NMC:

I called Jack. He quickly determined that the problem was not in oiqc but arose in the aircraft QC program prepacqc. We then reran the prepacqc compiled with -Rbc (array bounds check) and saw a multitude of instances of 0 indices. However, Jack also noted that most of the data had been excluded via the "receipt time" check - i.e. if the time of data receipt is prior to ob time—the data is excluded. This was instituted by Paul Julian to exclude data that came in as two day old data. I asked Jack to see how extensive the problem was. Jack then make plots of the aircraft counts before and after all qc. It revealed that when the ACAR data came on line (April 1991) the counts of accepted data dropped at 18Z. I then reprogrammed the count program that runs as a monthly summary of daily counts to do synoptic counts. This quickly confirmed the 1800 GMT plot jack had made and also noted that a loss of aircraft at 1200 GMT but not at 0000 and 0600!

It took 3 days of confused detective work trying to reconcile all the available evidence. We were delayed by Jack and Dennis having to attend to a failure of the operational data preprocessor. We then tried to reconcile the question of why the problem apparently began with the onset of the ACAR data. After laboring fruitlessly with this question, an observation by Dennis Kaiser's that this problem almost never occurs in operations finally pointed me toward the answer. If the problem is non-existant in operations it should be non-existant in CDAS. I rerun the synoptic count plots for April 1995 and saw the same behavior as in the reanl - large drop in accepted aircraft at 1200 and 1800 GMT!
This then focused attention on difference between operations and CDAS: the conversion of ON29 to bufr (common to reanl and CDAS). I called Jack and asked him to compare the receipt times for a June aircraft file before and after the conversion to bufr. He saw that they differed!

The aircraft receipt time caused problems

Jack and Dennis finally discovered that the receipt time conversion from ON29 to bufr was corrupted by an accident waiting to happen! ON29 does not have a place for receipt time—it had to be encoded into a "reserve word". Fine. However, the encoding is done by splitting the time into two non-contiguous
pieces! Presumably, the reserve word chosen was already being used for something else. When Jack was doing the conversion to bufr, he was working from an earlier version of ON29 that did not indicate that the receipt time had been broken apart!

Lost some aircraft reports

What are the consequences for reanl and CDAS of the receipt time error? The receipt time error has been dropping data via prepacqc since it was introduced into the processing for July 1987. It amounts to a 3-5% loss of aircraft reports at all synoptic times.

It is my opinion that this does not constitute a reason to rerun the period from July 1987 - March 1991. However, the loss of data from the onset of ACAR processing—April 1991—is unacceptable. Therefore, decided to rerun from that point. Similarly, the CDAS has to be rerun.

The rerun starting April 1, 1991 began at 1300 Friday, June 3, 1995.

A series of severe thunderstorms occurred on Sat Jun 4. Since then the assimilation system has failed three times, and the monthly avg program for Jun 91 also failed. All the failures seem to relate to disk I/O, and all attempts to rerun were successful. We lost approximately 24 hours due to these transient failures.

- Bob Kistler
Attachment 3

The Volume of Output from NMC/NCAR Reanalysis

The total output from reanalysis is about 54 Gbytes for each year (same as 430 tapes if 6250 BPI). For a description, see my "status" notes for Dec 1994, and the Bul AMS paper (fall 1995). The volume for some of the most popular grids is given below. The data are generally for four times each day.

<table>
<thead>
<tr>
<th>File</th>
<th>MB/year</th>
<th>Data Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>grbsanl</td>
<td>4930</td>
<td>Sigma analyses, Gaussian grid</td>
</tr>
<tr>
<td>pgb.foo</td>
<td>2536</td>
<td>Analysis stack, 2.5* pressure</td>
</tr>
<tr>
<td>grb2d</td>
<td>1990</td>
<td>Flux data, Gaussian, synop sort</td>
</tr>
<tr>
<td>grib</td>
<td>269</td>
<td>Monthly means, variance</td>
</tr>
<tr>
<td>znl (etc.)</td>
<td>104</td>
<td>Zonal means</td>
</tr>
<tr>
<td>--</td>
<td>3875</td>
<td>An 8-day forecast each 5 days</td>
</tr>
</tbody>
</table>

Notes:  
1. Files similar to Nos. 2 and 3 are also available in time sort.  
2. One CD-ROM for each year is being prepared by NMC. It has a popular sample of the data. The first one (for 1985) will be available about Nov 1995. NCAR will have copies to distribute at a rather low cost.

1. How can I access the data?

For those who can use the NCAR computers, the data are available online. For people who want to use the data at home, we plan to make the data available on tapes (Exabyte), with a low price if no data selection is needed. We assume the users have workstations, mainframes, or high-end PCs (not including Macs now).

2. Forecasts

NMC is running an 8-day forecast for reanalysis, done once each 5 days. NCAR has data for 1985 through 1990, but two cartridges for 1985 have to be replaced. The data volume is about 3875 MB/year (1920 MB for atmospheric state and 1055 MB for flux fields).
Attachment 4

Analyses from NMC

NCAR has analyses from NMC from the 1950s-on for the N. Hemisphere. The main stack of grids starts in 1962. Daily grids are on a CD-ROM. The archived global analyses start Jul 1976, but the tropics were very poor in these until Sep 1978. We have received the global data at 2.5° resolution, twice a day. From 1990-on, we also have archives of model resolution data, each 6 hours. The following shows selected changes in the NMC model.

<table>
<thead>
<tr>
<th>Change</th>
<th>Model</th>
<th>Resolution</th>
<th>Gaussian Points</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>14 Aug 1982</td>
<td>R30, L12</td>
<td>2.22° x 3.75°</td>
<td>96 x 80</td>
<td>Like T42 resolution</td>
</tr>
<tr>
<td>28 May 1986</td>
<td>R40, L18</td>
<td></td>
<td></td>
<td>Like T56 resolution</td>
</tr>
<tr>
<td>12 Aug 1987</td>
<td>T80, L18</td>
<td>1.48°, 165 km</td>
<td>243 x 122</td>
<td></td>
</tr>
<tr>
<td>6 Mar 1991</td>
<td>T126, L18</td>
<td>0.94°, 104 km</td>
<td>384 x 190</td>
<td></td>
</tr>
<tr>
<td>25 Jun 1991</td>
<td>T126, L18</td>
<td>0.94°, 104 km</td>
<td>384 x 190</td>
<td>Start SSI methods</td>
</tr>
<tr>
<td>11 Aug 1993</td>
<td>T126, L28</td>
<td>0.94°, 104 km</td>
<td>384 x 190</td>
<td>Convection is better</td>
</tr>
<tr>
<td>--</td>
<td>T170, L42</td>
<td>0.70°, 78 km</td>
<td>512 x 256</td>
<td>On wish list for 1996</td>
</tr>
<tr>
<td>Jun 1994</td>
<td>T62, L28</td>
<td>1.88°, 208 km</td>
<td>192 x 94</td>
<td>For reanalysis</td>
</tr>
</tbody>
</table>

- A diurnal cycle was added in Aug 1987.
- The SSI methods were first developed at T80 resolution.
- Aug 1993: The NMC production computer was still the Cray Y-MP (8 processors).
- Dec 1993: A Cray C-90 (16 processors) was delivered to NMC.
- 1 Apr 1994: Production was moved from the Cray Y-MP8 to the C90-16.

1. **The NMC pressure grids from operations**

From NMC operations, NCAR has received data for the pressure levels 1000-50 mb (12 levels) for many years, each 12 hours, plus a tropopause level and plus surface information. In the early 1980s data at 50 mb wasn't good from this set. Now we should obtain data for higher levels also.

The data are on 2.5° grids for the period from 1978-on. The dimensions are 145x37 points for each hemisphere.

2. **The NMC pressure grids from reanalysis**

These reanalysis grids at 2.5° resolution are for 17 levels, 1000, 925, 850, ...—10 mb, and are given each 6 hours (u-wind, v-wind, height, temperature, vorticity, vertical velocity (1000-100 mb), humidity (1000-300 mb). These are global grids (144x73 points). Model sigma grids and flux grids are also available from NCAR.
Attachment 5

Status of Sending Data to NMC

We will describe part of the status of preparing data at NCAR and sending it to NMC.

1. **NMC has all of the original NMC data for 1978-present**
   - Surface and upper air
   - A reanalysis header is on each report
   - Locations of raobs have been through time series checks and corrections made for 1973 through 1994.

2. **Special upper air rawinsonde data and some winds-only data (1973-95)**

   These data are from various national archives.
   - NMC has this data for 1981-on.
   - Working on data for 1973-80. There are some station number things and location things still to correct (this will be ready by about 15 Sep 1995).
   - More information about this is given later.

3. **Surface land synoptic data, 1947-95**

   NCAR will use the following sources:
   1. TD 13 from USAF for 1947 - Dec 1966. Still 2 months more work for this.
   2. USAF GTS decode for Jan 1967 - Jan 1977. We use the "Dick Davis" version from NCDC (still 1 month of work to do on this to solve some station number and location things).

   In addition, we will add U.S. 3-hour airways data for 1948 - Dec 1966, taken from the hourly dataset from NCDC (DS 470.0 at NCAR). The text "Surface Land Synoptic Data," 29 Mar 1994, has more information.
   - We have NMC surface data from Jul 1976-on, but we will use the Air Force data through Jan 1977 because the coverage is better in that period.

4. **NMC upper air data, Mar 1962 - Dec 1972**

   NCAR has the original version of upper air data from NMC for 4 Mar 1962 - Dec 1972. The data are in two separate basic data streams, one for 00Z and one for 12Z. The data have raobs and piballs, aircraft, cloud winds, and satellite sounders (from NASA SIRS). At first
the data was just for the N. Hemisphere; in Jun 1966 the data became global. Prior to 1 Nov 1969, only the 00Z and 12Z soundings were saved; since then 06Z and 18Z data are also available. Also see NCAR TN/IA-111, 1975, page 63, for more information.

- NCAR has accomplished a number of cleaning steps on the data (there is a text).
- We have already separated the aircraft data. As usual, we will also put the cloud winds and satellite sounders into separate data streams when the data are sent to NMC.
- We are working on time series checks of raob location data (Joey). This will fix the location errors that have been in all of these datasets (probably around 2 weeks to do this). Then more time will be needed to fix some of the worst elevations.
- We haven't added the reanalysis headers to each report yet. This will go fast (5 to 10 days) once all correction information is available.
- One problem is that, for example, during this period the raobs from Canada changed their block number from 72 (like USA) to 71. But their old Canadian data from Canadian archives has the new numbers. So we need to handle the differences in numbers.

5. The MIT raob dataset (May 1958 - Apr 1963)

The 5-year MIT dataset has global coverage. It was used to calculate atmospheric circulation statistics (Oort and Rasmusson, 1971). The data was gathered by Professor V. Starr's project at MIT. The data tapes that got to NCAR came via Kung at the University of Missouri. In 1994, Professor Kung did more cleaning on the dataset, and NCAR has the results of this work. It will be used for reanalysis.

The NMC decode (N. Hemisphere) starts Mar 1962 and became global in Jun 1966. It will also be used. The earlier dataset (TD 54) prepared by the USAF will also be used. We note that the MIT data has only a few stations for Australia. Data from Australia will completely fill in that region.

- About a month of staff time will be needed to put on reanalysis headers, etc., and get this ready to send.


Sherhag's group at the Frei University of Berlin has drawn daily stratospheric maps (N. Hemisphere) for many years (now Karin Labitzke leads the group). They made a special collection of raob data. Dick Davis, NCDC, sent us a copy in Aug 1995. It is supposed to have data for 1954-62. This probably only has the higher levels of raob data. We are still trying to gather some format information.

For 1958-63, I think that most of this same data should also be in the MIT dataset, and in TD54.
7. **COADS surface marine data**

   a. Data for 1980-92 was sent to NMC

      Jul 1993: Release 1a (1980-92) was sent to NMC. Used by Dick Reynolds for SST, and used for reanalysis.

   b. Data for 1992, 93 was sent to NMC


   c. Data for Dec 1979 was sent to NMC in Feb 1995: To fill gap between FGGE and 1980.

   d. Data for 1947-79 is being prepared

      The COADS project is working on this period. An old version is available, but it needs new inputs and a few fixes.

      - We are doing the 1970s first
      - All of the inputs have been converted
      - FGGE IIb has been added to check for duplication
      - Data for the 1970s has been sorted together
      - We think we know the best rules for eliminating duplication
      - Data for 1970 through 1977 will be ready by late Nov 1995.
      - Data for the 1960s will be ready in Mar 1996.


Constant level balloons near 150 mb gave rather good data coverage of the southern hemisphere during this period. There was an altimeter that gave the distance to the ocean (or land) with an accuracy of about 10 meters. The reports over water can be used like one known level (near 150 mb) in a rawinsonde ascent.

Status Aug 9: Bob Dattore has been processing this data. So far it looks in good order, but I am worried that winds are not given for one general case where they could be calculated.

9. **Aircraft and reconnaissance data**

The data from NMC decode for aircraft and reconnaissance has the following status:

   - 1978-on has been sent to NMC
   - 1973-77 could be sent within days, if needed
   - 1962-72: The NCAR work to prepare this data from older NMC tapes has been completed, but reanalysis headers need to be added (1 week).
Additional data have been gathered from several sources. There is data from New Zealand from Dec 1978 to Jun 1988, some of which never did get onto GTS. Kidson in New Zealand thinks that most of their data did get onto GTS after this time. Data from GATE (Jun-Sep 1974) and FGGE (1979) will be used.

Other sets of aircraft and reconnaissance data have mostly been processed. I will try to prepare an updated text about aircraft data. There's no time now.

10. **Russian ship UA data**

   The attached sheet (Attachment 7) has counts of ship upper air data from Russia. All of set 1 (54,000 reports) has been going into reanalysis. Set 3 wasn’t ready in time to include the latest years. But data for either prior to 1982 or prior to 1978 will be used. Something happened and set 2 can’t be located at NCDC. Russia will send another copy. This is an impressive set of ship UA data. The reports with year equals zero can’t be recovered because there wasn’t any apparent order to the data on the tapes. These data do not include raobs from Russian military ships.

   - Data for both set 1 and set 3 are essentially ready as of Sep 26, 1995. Only a few minor things remain to do.

11. **Raobs and pibals (the extra data for NMC)**

   As we go further back in time, NCAR has been providing NMC with increasing amounts of data besides the NMC decode of GTS. There are some texts and figures that give more information about the extra data. Some of these need updates.

   1985-on: Additional raob data were stations in Antarctica, and isolated stations—see the list (about 20 stations). Unfortunately, some of the Antarctica stations—S. Pole, McMurdo, Japan, and UK special stations—didn’t get into this time period, but they did get into earlier years. However, much of the data was still available from GTS.

   1981-84: Has an expanded list that includes Brazil (raobs and winds), Argentina, S. Africa, Australia (raobs and winds), UK overseas, UK and Japan, and Antarctica.

   1973-80: Added in about everything we have. Included Caribbean stations, etc., plus all of the above.

   More data for Antarctica: In the bilateral exchange, we met with Russian colleagues in July 1995. They plan to send long records of their Antarctic raob data during 1996. I think that most of this is already available for reanalysis from other sources, but not all.
12. GATE raob data, summer 1974

The GATE experiment was for Jun-Sep 1974, off the coast of N. Africa. By about Aug 21, 1995, Bob Dattore will have these UA data ready for reanalysis. They include rawinsonde data from the ship array, some land stations, and dropsondes. Dennis Joseph spent the summer in Dakar for this experiment.

13. Use of TOGA COARE raob data

In late Jan 1995, I talked with Jim Moore and Steve Williams at NCAR about the status of raob data preparation for TOGA COARE. They now have the high-resolution data for 19 sites (2-second to 10-second data). Winds have been properly filtered to remove noise. The problems of low humidity and high humidity in instruments have been corrected as best they can.

Most of the 19 stations probably did come in at least part of the time on GTS, but this is better data. It included three ships that probably did not get on GTS (1 ship, 90 days; 2 other ships, 2 weeks each). Steve Lord at NMC probably has programs that could go from the high-resolution soundings to data with a reasonable number of levels for reanalysis. If the data are used, they say that the new files now online at NCAR should be used.

Status of TOGA COARE raobs:

Jan 1995: The dataset has 5,000 soundings. When they are finished with everything, there will be 13,000 soundings.

5 Jun 1995: Now about 85% of the soundings are completed and online at NCAR.

Note: When NMC picked up the raob data for use in reanalysis, about 85% to 90% of the soundings were probably available.

14. Fake raob stations

Sometimes a raob has a wrong station number (probably a garble), and by chance it looks like a surface station number. The NMC procedures accept that as a valid raob site. NCAR handles these so that they will not be used for reanalysis. If there are steady raobs at a site that is indicated as only surface, we do accept them. I am sure that there are some garbles in raob numbers that send a raob to a wrong true upper air location. More tests should be made on the wrong number data. This happens more often than I am comfortable with. A good report at a wrong location can cause damage.
15. Some raob data from other sources

The following are described in more detail in the status report dated 27 Apr 1995:

- The count of significant wind reports on the original NMC tapes was low from Aug 1989 to Sep 1991 (26 months). ECMWF sent all of their raob data for this period and NMC did a merge.

- Some missing raobs during Jan 1992 - Jun 1993 (18 months). ECMWF provided data for blocks 97 and 98 to restore these data.

- Some Pacific upper air data from Japan. This JMA upper air filler data (raobs and pibal) is for a selection of Pacific and China stations, and covers the period 1979-93. NMC found that most of the data was also in other sources, but about 10 to 20% of it filled gaps and was used.

16. Location and elevation of upper air stations

It is difficult to derive correct information about the location and elevation of stations. All station library sources have errors, and the locations on the NMC tapes have errors. NCAR (Joey Comeaux) pulled the location and elevation from every UA report on the NMC tapes, 1962-present. For raobs, we also hydrostatically calculated an elevation from every report. We have graphics (time plots) that show the relationship between the reported elevation and the calculated elevation. There are many errors. For the location of stations, we found that the best way to discover errors is to do a time series analysis of the reported position that tells us when the location changes by more than 50 km. Sometimes there are huge persistent errors. One day on the NMC tapes, all S. Hemisphere stations were in the N. Hemisphere. These problems are then inspected and compared with other library tapes to determine what is correct. This step is done by manual inspection and is quite time consuming since the correct values are often difficult to determine. (Later on, we will also prepare a library tape that has more updates.) Summary of what has been done:

Station locations: These checks and corrections for location have been done for 1962-95, and we are confident in the results within the 50-km error limit. Actually it will be Oct 15, 1995, before the 1962-72 part is finished; and 100-im error limits may be used to speed this process. For early data (1962-63) the reports have no elevations, so later NMC data and other sources will be used to fill these in. This covers the raobs and the wind reports at raob locations. During 1973-95, the locations of pibal-only reports have not been through these checks. For 1962-72, the pibals are also checked.

Raob station elevation: These corrections have not yet been on the reports sent to NMC. For 1973-95 we had about 60% of the corrections done, but there was not time to put them in the tapes for 1978 to 1994 already sent to NMC. Calculated elevations show obvious elevation errors quickly, but the noise in these calculated values makes determining the date of the change quite difficult. Calculated elevations were used to resolve discrepancies for elevations of special raobs.
17. Location and elevation of surface stations

We have also done a lot of work to improve the reliability of the location of surface land reports. On a typical day there are reports from about 7500 stations. This will need continued work after the first reanalysis, and statistics from the analysis could help determine and correct problems.

18. Ice cap buoy data

NCAR has 3-hourly data from Antarctica and Greenland from Chuck Stearn's group at the University of Wisconsin. The data includes stations around the edge of Antarctica. By Apr 1994 the data through 1992 was available. In mid-1995 NCAR obtained data for 1993. Checks (by the University of Wisconsin) aren't completed for 1994 data as yet.

I think that both NMC and ECMWF picked up the data for 1980 - Dec 1992 about May 1994. Please verify. In more recent years, most of this data is also on GTS. The number of station months for sample years are:

Station months of ice cap surface data

<table>
<thead>
<tr>
<th>Year</th>
<th>Antarctica</th>
<th>Greenland</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>65</td>
<td>0</td>
</tr>
<tr>
<td>1984</td>
<td>190</td>
<td>0</td>
</tr>
<tr>
<td>1987</td>
<td>279</td>
<td>8</td>
</tr>
<tr>
<td>1992</td>
<td>320</td>
<td>64</td>
</tr>
</tbody>
</table>

Note: See the text "Ice Cap Buoy Update," 5 Apr 1994.


The next two items give the data status. It is important to use LIMS, Version 5, for Oct 1978 - May 1979. The final FGGE dataset still had Version 4. Better limb sounding data could be calculated for these periods, but it has not been done (as of Sep 1995).

20. Other Stratospheric Data (Limb sounding data)

John Gille (NCAR) had instruments on Nimbus-6 and -7 that measured stratospheric temperature and gas content by measuring IR in vertical profiles, while looking at the limb of the earth. The data have a line of profiles along the earth's limb, that are parallel to the path of each orbit. The sensor could only work for 7 months because it required cooling, and there was a limited supply of coolant.
• LRIR - Limb sounder on Nimbus-6
  — The data are for 19 Jun 1975 to 7 Jan 1976 (7 months)
• LIMS - Limb sounder on Nimbus-7
  — The data is for 25 Oct 1978 to 29 May 1979 (7.1 months). These data are in FGGE IIb data, but not the last version. John Gille (NCAR) prepared these soundings. In 1993, Gille said that they now have new methods that would improve the accuracy of the LIMS soundings. The work has not been carried out. John thinks it may be possible to do it by sometime in 1996 (written in Aug 1994).
  -- Note: If we wait too long to achieve some of these improvements, the expertise will disappear.

21. Use of LIMS Satellite Stratospheric Data

In Jan 1987 I wrote a summary of the status of "FGGE reanalysis and FGGE observed data" (Roy Jenne). The portion of that text about LIMS follows:

The earth limb scanning data was taken during the first 7 months of FGGE. The data is for the period Oct 25, 1978 to May 27, 1979. Data were calculated for 64S-84N in steps of 4° latitude.

ECMWF found that the quality of the early LIMS used for the 1981 FGGE analyses was fairly poor and seemed to vary from day to day. They saw some variations of even 5° that didn't look correct. Sometimes there was a bad satellite swath between two good ones. LIMS didn't get much analysis weight because of the quality.

For the reanalyses, ECMWF used the version of LIMS that was in the final IIb data (1985) from Sweden. This was LIMS Version 4. Paul Bailey and John Gille (NCAR) note that most of the problems in the earlier LIMS were taken care of by Version 4. It still had perhaps 100 cases where a whole derived sounding was bad and a few episodes of problems associated with stratospheric warmings. In Version 5 (about fall 1983) there are still about 10 bad soundings, but the warming cases are handled properly. There are major improvements in the derivations of stratospheric gas concentrations in Version 5.

The new LIMS (Version 5) will be used by GFDL and GLAS. Thomas Miles (NASA Langley), says that the new LIMS looks very good; much better than the TOVS retrievals. LIMS has a vertical resolution of about 3 km and can resolve only about six horizontal waves. SSU (in TOVS) has a vertical resolution of 15 km (it has 3 channels) and can resolve about 12 zonal waves. There was about a year when the top channel of SSU did not work. Comparisons of the new derivation of LIMS (FGGE) with rockets look very good.

The LIMS data for FGGE only gave temperature (not gas concentration - gas data is on other tapes) for the pressure levels: 100 mb, 70, 50, 30, 16, 10, 7, 5, 3, 2, 1.5, 1. The satellite could scan the earth's limb both on the ascending and descending side of each orbit. It views the emission of infrared from the limb. It is 3500 km from the orbit location to the
tangent point where a sounding is located. The time period was Oct 1978 to May 27, 1979. Data could be calculated for 64S to 84N in steps of 4° latitude. In the FGGE format for a sounding, 37 characters are used for the ID and for each pressure level. If we disregard 10 to 15% of missing data, the volume is about (215 days) (2 times/day) (14 orbits/day) (38 latitudes) (12 levels + 1) (37 bytes) equals 110 megabytes.

Version 5 of LIMS did not get into the last (1985) IIb, as prepared by Sweden (Version 4 did get in). Asheville replaced LIMS 4 with LIMS 5 in their set. NCDC also made subsets; they have the data (all are 1600 BPI tapes; these can be put three tapes to one 6250 BPI tape).
Attachment 6

Inventories of Input Data for Reanalysis

NCAR has a lot of inventories of the data for reanalysis, at different levels of detail. We will talk about inventories that show summarized information about daily data. In many cases, there is one print line for each day that shows the number of observations for each main synoptic hour.

1. Text with inventory information

The text "Data for Reanalysis, Inventories," Nov 1992, was prepared to give a sense of the data coverage that will be available for reanalysis. It includes maps of data coverage. It mainly gives information for about 1948-on.

2. Satellite sounders

The text (Aug 1994) about satellite sounders has the inventory information (down to the day level) that is needed. The inventory of the 2.5° TOVS sounders follows:

FTP: datasets/ds686.0/inventories/nmc

The inventory of gap fillers is also in this directory.

3. Inventories of daily upper air data from NMC, 1973-94

These inventories show the count of global raobs, pibals, aircraft, satellite winds, etc., for each day. There is one print line for each synop hour. Note that the numbering system for NCAR datasets has a form such as ds353.4

FTP: datasets/ds353.4/inventories/adpupa.[year].sum

4. NMC upper air data for 1962-72

This text "NMC Upper Air Data, 1962-72" has information about seven types of upper air data (raobs, aircraft, cloud winds, sounders, etc.) on these tapes. It includes a variety of inventory information. There is usually one print line for each synoptic hour.

FTP: datasets/ds353.0/inventories/analysis_00z&12z
5. **Raobs per month for each station on NMC tapes**

The monthly counts of raobs here are based on daily counts in the above two items. There is one print line for each station-year.

- 1973-94
- Period: 1962-72 (not available yet)

FTP: datasets/ds353.4/inventories/advupaa.rec.w[mmnn]

where mm is the first block number included in the file and nn is the last.

6. **Time series rawinsonde list**

One of the lists at NCAR shows data availability over time for each rawinsonde station. This list shows data from sources, not including GTS. We will make available the list that combines various sources. One print line shows up to 60 years of coverage for a station. One symbol for each year shows the months of daily data coverage in each year.

FTP: datasets/ds390.2/inventories/stations

7. **COADS surface ocean data**

There are papers that have plots of data counts, and figures that show data coverage versus time. The data includes ships, fixed buoys, drifting buoys, arctic ice buoys and surface data from XBTs, etc.

FTP: pub/coads/release_1a

8. **Surface land synoptic data**

A text ("Surface Land Synoptic Data," May 1994) summarizes the availability of land synoptic data. It describes the datasets that will be used for reanalysis. It has data coverage plots for somewhat arbitrary years such as 1956, 1963, and 1982. Portions of daily listings show how the counts of data each day (and each hour) have changed with time.

8.1 **Surface land reports**

These counts of surface land reports are for each 3 hours, each day (1 line per day)

- NMC (Jul 1976 - 1994): See 9a below
- TD13 (1901-1973): See 9.1
9. **How to find inventories of surface data from NMC, 1976-94**

   a. Reports per synop hour
      FTP: datasets/ds464.0/inventories/adpsfc.[year].sum

      Monthly station receipts
      FTP: datasets/ds464.0/inventories/adpsfc[3/6].rec.w[mmnn]

   b. FTP: ftp://ncardata.ucar.edu/datasets/ds463.0/inventories/by_date/inv[yr]

9.1 **Inventories for TD13 data**

   Total reports for each year-month are broken down by synoptic hour.

   FTP: datasets/ds467.0/inventories/nobmo

10. **How to find general online information about NCAR data**

    The general e-mail address which may be used for initial contact with DSS is:

    datahelp@ncar.ucar.edu

    There is a wide variety of additional information in the DSS information area. There are
    often additional types of inventories in the same directories as those noted above. Check out
    the README file in these directories. Additional information is available by email, anonymous
    FTP, or over the web.

    FTP: ncardata.ucar.edu

    WWW: http://www.ucar.edu/dss

11. **How to quickly connect to find these inventories**

    ftp ncardata.ucar.edu

    *or type:*

    ftp 128.117.8.111

    < connected... >
    < name: >

    anonymous

    < password: >
your email address

< ok >
< ftp >

cd datasets/dsxxx.x ...

Note: Credit goes to Dennis Joseph, Will Spangler and other members of our group who prepared most of this information.
1) Received at NCAR - Feb. 1991
   - Russian tape CCCPA1
   - USSR upper air data taken from ships 1961-1987
     - actually first date found was 1913, probably a mistake
     - 54059 records

2) Received at NCDC - approx. Dec. 1992
   - Russian tape RUA2
   - Russian Upper Air Data 1961-1990
     - 27094 reports

3) Received at NCAR - Feb. 1993
   - two data files on Russian tape RUS64
     1) Marine ship data - added to Russian ship data archive
     2) Upper air observations taken from ships
        12968 records

<table>
<thead>
<tr>
<th>YEAR</th>
<th>COUNT</th>
<th>YEAR</th>
<th>COUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1913</td>
<td>1</td>
<td>1947</td>
<td>4</td>
</tr>
<tr>
<td>1948</td>
<td>59</td>
<td>1949</td>
<td>46</td>
</tr>
<tr>
<td>1952</td>
<td>2</td>
<td>1961</td>
<td>209</td>
</tr>
<tr>
<td>1962</td>
<td>349</td>
<td>1963</td>
<td>336</td>
</tr>
<tr>
<td>1964</td>
<td>722</td>
<td>1965</td>
<td>134</td>
</tr>
<tr>
<td>1966</td>
<td>1166</td>
<td>1967</td>
<td>1</td>
</tr>
<tr>
<td>1968</td>
<td>307</td>
<td>1969</td>
<td>11</td>
</tr>
<tr>
<td>1970</td>
<td>1526</td>
<td>1971</td>
<td>137</td>
</tr>
<tr>
<td>1972</td>
<td>322</td>
<td>1973</td>
<td>2</td>
</tr>
<tr>
<td>1974</td>
<td>357</td>
<td>1975</td>
<td>1</td>
</tr>
<tr>
<td>1976</td>
<td>464</td>
<td>1977</td>
<td>116</td>
</tr>
<tr>
<td>1978</td>
<td>360</td>
<td>1979</td>
<td>3884</td>
</tr>
<tr>
<td>1980</td>
<td>1260</td>
<td>1981</td>
<td>2953</td>
</tr>
<tr>
<td>1982</td>
<td>555</td>
<td>1983</td>
<td>1438</td>
</tr>
<tr>
<td>1984</td>
<td>567</td>
<td>1985</td>
<td>1896</td>
</tr>
<tr>
<td>1986</td>
<td>630</td>
<td>1987</td>
<td>2862</td>
</tr>
<tr>
<td>1988</td>
<td>88</td>
<td>1989</td>
<td>7650</td>
</tr>
</tbody>
</table>

YEARS OUT OF RANGE: 625

SUMMARY:
TOTAL RECORDS READ: 12968
TOTAL RECORDS READ: 54059
OUT OF RANGE: 0
Attachment 8

Does a Surface Algorithm Change Affect the Upper Air?

Is it possible to change the surface algorithm over land in order to control the surface maximum temperature problem and still not change the climate of the upper air, or the precipitation? We will give some information from tests made at NMC so far. It appears that the answer is that the algorithm could be changed without affecting these aspects of the climate.

1. Butterfly effect

Whenever any small change is made in an assimilation system, it creates changes due to the butterfly effect. The small changes can evolve to slightly different states during the forecast (first guess) run. When different runs are compared, one has to expect these small changes, even if the runs are almost the same. Therefore, changes seen have to be measured against the butterfly effect, and not against zero change.

2. For a test, a few months were rerun with a new algorithm (done Sep 1995)

Kistler re-ran a few months where the surface wind speed was forced to be at least 1 m/sec over both land and ocean. If the algorithm is changed, this rule will only be used over land. This test is therefore more extreme than any changes that might be made.

Did the Hadley cell change? Glenn White compared monthly zonal plots of temperature, u wind and v wind. There were essentially no differences. One temperature contour had a very minor difference near 500 mb.

Compare monthly maps. U component wind: There were only a few spots near the poles and one small region near the equator where the differences were as big as 0.5 m/sec (this was Jan 1983).

Did the precipitation change? He looked at zonal mean precipitation over land. There were some places with a slight reduction over land, but not enough to worry about.

Glen looked at divergence of the 10m surface wind, monthly means. This was very close for the two cases. This is another indication that this change in the surface algorithm does not change the Hadley cell.

Kistler prepared statistics on RMS differences of daily values of Z, T, U, V. The local differences in heights were under 5 meters overall and within 2 to 3 meters at 850m. Temperature differences were under 0.3°K overall and 0.4°K at 850 mb in the tropics. RMS changes in u and v were under 1 m/sec for any place on the global map.

If we look at certain percent changes in winds between runs, these can get as high as 25%. With the other small changes described above, and with the Hadley cell okay, I do not think we have to worry about this. I think it just means that some of the overall winds are weak enough that percent changes can be big. At 850 and 700 mb we can see these changes of
20% to 30%. At 200 mb they are not bigger than about 15%, probably because the winds are a little stronger.

3. **How many land points have had high temperatures?**

Consider the frequency of low winds at land grid points. Only land points with low speeds were counted, but the percentage is given with reference to all global grid points.

- land points under 0.75 m/sec: 2.3% of global points
- land points under 1.0 m/sec: 5% of global points

Wesley inspected the bad points. He thinks that only 1 in 5000 of these really have a problem. He now thinks that there may only be one bad point somewhere in the world each 3 days. But people also say that the climate of some of the surface variables does change significantly, if the algorithm changes. Perhaps this is because there are smaller effects at many grid points, but only a few look very wrong. Wesley Ebisuzaki is in CAC. He is spear heading the effort to prepare the CD-ROM.

4. **Some small changes made starting with Jan 1990 data in reanalysis**

Some small changes were made in reanalysis starting with data for Jan 1990 (only data for 1985-89 do not have these changes).

- A different Cray compiler was used.

- The convective parameterization could create some negative precipitation. The result of this wasn't allowed to affect precipitation, but it did affect the amount of water in the column. Then the global water balance was not quite right. When it was fixed, the evaporation increased a little bit.

- Sometimes the Richardson number was misbehaved and a small change was made to fix that.

These small changes should not worry us much, but they do create some small differences.

NMC has tested a more "surgical" approach to solving this problem. With a low-speed surface wind, the wind (for the algorithm) was increased to 1 m/sec only over land and only when the stratification was unstable. This led to essentially no noticeable changes except over some equatorial regions. Even here we expect that the change in surface climate will be quite small.
Some Reanalysis Notes

These notes are mostly based on a conversation with Bob Kistler (NMC, now NCEP), Nov 27, 1995.

1. Status of Reanalysis

NMC has done reanalysis for 1982 through Jul 1995 (13.6 years done). Bob Kistler also had to redo Dec 1985 because of bad tapes. And a month in 1994 was redone. They will soon start to do 1979, 1980, and 1981.

2. Use of TOVS for Dec 1978 and 1979-on

There are 5° TOVS in the FGGE data for Dec 1978. NMC will use it for spinup, and then they will use the 2.5° TOVS for 1979-on. They have been using 2.5° TOVS data for 1982-on for the analyses that are done.

3. Test of VTPR data

We should get more VTPR data to NMC so that they can run a test for the overlap period (TOVS/VTPR) of Jan, Feb 1979. VTPR data goes from late 1972 to Feb 1979.

4. Cloud drift winds

The SDSD archive starts with data for 1974. This is the delayed data archive from GOES data. For some years it has more data than are on the NMC tapes. Send NMC the necessary data.

5. COADS data

Doing reanalysis for 1979-80-81 will be about 6 or 7 weeks of work. However, NMC may be ready to start loading COADS for the 1970s in about 2 weeks.

6. Day-night differences in raob data

Bob would like to have a lot of the early raob data (1957-on) so that they can analyze day-night differences in the stratosphere to see what corrections should be made to the data.

7. A change on NMC data tapes, 1995

The automation division made a change in the handling of recco data starting Aug 1995. In NMC, they have figured out what they need to do, as of Nov 27. Then they will reanalyze more of 1995.
8. The fix of surface temperatures

Data for 1994, 1995 and pre 1982 will have the fix. The 1982-93 data will not have the fix for occasional high daytime surface temperatures with calm wind speeds. Low speed winds are boosted to 1 m/sec for the calculation of heat flux from the surface over land. This does not change the archived wind.

9. A special CD-ROM for Bulletin AMS

Over the holiday weekend, Kistler selected a lot of data files needed for a CD-ROM for AMS (with 13 years of data). We hope that this CD-ROM can be made in time for the Feb 1996 issue of the Bulletin AMS.

The CD-ROM will have year-monthly means for pressure levels (17 levels) for 1982-94. It will also have some monthly climatologies. There will still be room for a few fields of daily data.

10. Received third UA ship raob tape from Russia

NCAR has processed two tapes of Russian ship raob data. The third tape has been missing. Russia sent another copy on Aug 15, 1995, and it arrived at NCAR on Nov 30, 1995. It should have data for the years 1958-91.

11. Problems with the new computers at NMC

The machine conversion was difficult (from Cray Y-MP to Cray Jedi's). First NMC had to change their software to achieve efficiency. From Aug-Nov 1995 there have been machine problems: 1 hardware and 3 software problems. The last of the system software bugs will soon be fixed. They are through the bad period.

Another tape storage silo (third one) became operational on Nov 29, 1995. It has eight tape drives and they are a faster version.

12. Data status (29 Nov 1995)

a. NCAR has received data for 1985 through 1993, but 1985 will be sent again, and polar winds in pressure files will be corrected.

b. Data for the year 1984 was sent to NCAR about Nov 29, 1995.

c. Data for the year 1985 was recalculated. It will all be sent again to NCAR.
d. Problem of bad polar winds in pressure grids:
   — Have corrected data for all years (1982-93) at the first level.
   — Have corrected other affected files for 1982 and 1983 (monthly statistics, etc.).
     This takes 2 days per year. By the end of the year, NMC will have all file
     corrections done. Then they will start sending the corrections.
   — Data for 1994 and 1995 will all be okay when we get it.

e. The 10m winds are all okay. They were not affected by the polar wind problem.

f. The forecast data (a long forecast each 5 days)
   — All the flux fields are okay.
   — The winds on pressure levels have the polar wind problem. The forecasts would
     have to be run again to fix this. We don't plan these reruns.
   — The forecast winds for 1984, 1985, 1994, and 1995 will be okay. Other years
     during 1982-95 will have the polar wind problem.
Reanalysis Update, Dec 13

The following is from a talk with Bob Kistler, NCEP

1. 13 Dec 1995

Kistler will finish data for Nov 1995 today.

— So 1982 through Nov 1995 (13.9 years) are done.
— Some data for Jan 1995 were lost so that has to be fixed.

2. Analyses of 1979-81

Jack is now working on the data for FGGE. The TOVS data for Dec 1978 in FGGE is not identified well as to what it is, so they are not going to use it.

For FGGE, NCEP will do the same as ECMWF did. Use VTPR for Dec 1978, then use 2.5° TOVS for Jan 1979-on, as planned. I hope to see a parallel test using VTPR for Jan-Feb 1979 to help evaluate VTPR vs. TOVS.

NCEP will start FGGE about next week and will analyze data for 1979-81. This will go fast because they have optimized for the new computer. It is now faster than the Y-MP. (But the optimization would have made the Y-MP faster too). These years may go as fast as a week to do a year.

3. The computer

The new computer at NCEP still is not entirely stable.

4. Sending data to NCAR

a. NCEP previously sent us data for 1984.
b. They copied 1982 for us last week.
c. They will copy 1983 for us next, probably next week.
d. They haven’t sent the new data for 1985 yet. It still needs repair work and a verification that all files are okay. The computer was really up and down during that period.

Notes: (1) The new tape drives are faster. Today Kistler copied 20 tapes to 20 tapes in an hour.
(2) Some data for Jan 1995 were lost so that has to be fixed.

5. Radiation corrections

Kistler found copies of the radiation correction tables that were derived from a 19-month period during 1974-76. He has the Tech Note by Fred Finger.
6. **Data for the 1970s**

   We should send COADS data and all the other types of data by the first part of Jan 1996.

7. **Surface stress**

   Kistler says that the stress grids (on Gaussian) in the grib2d file should be used.

8. **Summaries of data bias, etc.**

   Bill Collins (x8161) has been preparing some monthly summaries of the biases.
Status of Observations for Reanalyses

Observations for 1978-95 are already at NCEP. This is a list of data going into reanalyses. NCEP is now 20% of the way through the 1979-82 period. Then they will do reanalyses for 1972-78 (the VTFR era) and need the data. In this list we are concentrating on data for 1968-79, but most other data are also listed. NMC will later have data for the whole 1948-95 period (48 years). Item 6 discusses the batches of raob (and pibal) data that will be sent to NCEP (NCEP readers can ignore the other raob data source summaries if they want to).

- At NCAR, we will concentrate on moving data for 1968-77 to NCEP fast enough that the data ingest by Jack Woollen is fully busy.

- At any given time, NCAR is working on a lot of the older datasets so that they have a good chance to be ready to send to NCEP when needed.

- NCAR merges many sources of data. The blocks of data that will be sent to NCEP are listed in Attachment 1. This is the main summary information that NCEP needs to read.

- For most observations used in reanalysis, NCAR puts a short header on each observation that gives the best information about time, location, and source.

1. How many years of reanalysis are done?

The period 1979-81 is now being worked on. The next period will probably be for the early 1970s thru 1978.

<table>
<thead>
<tr>
<th>Date</th>
<th>Years of reanalysis done</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 Sep 1995</td>
<td>11.0 (1982-93, but not 1984)</td>
</tr>
<tr>
<td>5 Jan 1996</td>
<td>14.6 (1982-95 plus part of 1979), thru Jul 1979</td>
</tr>
<tr>
<td>11 Jan 1996</td>
<td>15 years are done. The snow stopped things for a few days</td>
</tr>
</tbody>
</table>

Notes:
1. On Jan 7, 8, 1996 there was a huge snow storm in Washington, DC (20-24 inches), the most snow in 70 years. On Jan 9, another 3 to 6 inches of snow fell on Annapolis and Baltimore. Jan 12, snowing at NCEP, another 6 in. has fallen.
2. During the period late Aug 1995 - Nov 1995, the NCEP project was slowed down a lot by system and hardware problems in moving to a new computing system.

2. Doing reanalysis for the next batch of years; which years?

Reanalysis for 1979-95 (17 years) will soon be finished. Which batch of years should be next? NCAR is now in the process of sending data for 1968-78. Figure 1 shows the availability of satellite sounding data. Bob Kistler and I assume that no one will have the
time to do the research necessary to use the SIRS raw data. I note that SIRS soundings on the early NCEP tapes at NCAR start May 1969; VTPR starts roughly 20 Nov 1972. Please see Table 1 for SIRS sounder coverage; there are some gaps.

For the next batch of reanalyses, NCEP would avoid the problems (in the S. Hemisphere) of facing a world without satellite sounders by starting with Nov 1972. However, it would also make sense to start the next batch with May 1969, when the sounder data is first available. I suspect that this may be a better choice.

From Figure 2, please note that NCAR has 500 mb analyses of the S. Hemisphere (starting Aug 1968) from Australia. If necessary, these could be used as a guess.

3. COADS data (1970-79), ocean surface marine

The data are ready but NCAR needs to put it into a synoptic sort. The volume is about 17 to 20 MB per month.
— The data was FTP'd to NCEP on Jan 11, 1996.
— We can send COADS data for 1969 as a special case, if that is needed soon.

4. Some main raob sources for reanalysis

4.1 NCEP decode of GTS
— the 1973-95 dataset, all in one format
— the 1962-72 dataset, another format
(The station library problems have been a bear for 1962-72, etc.)

4.2 MIT raob dataset (May 1958 - Apr 1963)

This is the near global collection that was used for circulation statistics (e.g., Oort and Rasmusson, 1971).

4.3 TD54 raob data from USAF (mostly 1945-65)
— Some data goes to 1971
— Sequence 1-19 is from work done on TD54 at NCAR in the 1970s

4.4 Data for Antarctica

This is a special collection at NCAR. It includes data from tapes from the U.S. Japan, Australia, UK, and S. Africa. We may obtain more old data from Russian stations during 1995. We have a short text and inventory about this collection, and use it to make sure that all the pieces of data are included.

4.5 Arctic rawinsonde data

Jonathan Kahl (University of Wisconsin) has been in the process of assembling a set of UA data for the Arctic. He has digitized much data including Ptarmigan dropsondes. He needs money to digitize more. See the separate text.
• About Jan 16, 1996, Jon Kahl will send NCAR North Pole Russian raobs for 1954-87 (only 0-3 km now). About April 1996, we will get data for the upper levels.

• NCAR will prepare the 0-3 km data for 1968-78 by about Feb 10, 1996.

4.6 Sherhag raob collection (1954-62) is mostly for Europe and Asia

There are about 50 stations in this dataset, mostly for Europe and Asia.

5. The NCEP GTS decode of UA data for 1962-72

NCAR has a text about this dataset (which is also listed above). NCAR will send the following files back to NCEP:

• Raob and pibal data
• Aircraft reports
• SIRS soundings
• Early satellite cloud wind data

6. Batches of raobs for NCEP

These raobs are all sent in synoptic sort. The "all-else raobs" pick up raobs from many datasets listed in this text. See Attachment 1 for dates of future shipments

• NCEP decode (raobs, winds) » they have 1978-on
  » 1973-77 will be ready Jan 26, 1996
• All-else raobs batch » 1981-84, 22 Mar 1995
  » 1973-80, was sent 6 Oct 1995
• All-else winds-only batch » 1981-84, 22 Mar 1995
  » 1973-80, was sent 6 Oct 1995

• Three Russian ship batches
  — Part 1 is in raobs and in winds-only above
  — Part 2 » was ready for NCEP on 3 Nov 1995.
  — Part 3 (biggest set) » estimate ready for NCEP on 12 Jan 1996. The data are mostly for 1961-91. They were sent from Russia on 18 Aug 1995, and arrived at NCAR on Nov 30, 1995.

Note: For the data period 1973-80, NCAR sent NCEP a complete set of the "all-else" raob data that we had available. For later years we sent a more selective set. Our plan for earlier years is to send almost everything.

• Plan: NCAR will make similar batches of data for 1968-72 next, and they will be sent to NCEP about Jan 19, 1996.
7. Raob data, general ship collections

7.1 Data from GTS (this is in the NCEP data, 1962-95)
7.2 The three tapes from Russia, see above
7.3 U.S. permanent ship data (on old U.S. control tapes)—most were collected onto tape sequences 103 and 104 at NCAR.
7.4 Canadian ship P (with the Canadian data)
7.5 European permanent ships
7.6 Collection of U.S. ships from NCDC (TD6210), includes Navy carrier data, 1946-93
   — NCAR received 4 cartridges from NCDC about 18 Dec 1997. Will has copied the data.
7.7 Germany has a research ship collection (we don't have it)
7.8 Ship data collection by S. Africa (NCAR does not have this).

Note: The last big tapes with Russian ship data arrived. On 4 Jan 1996, Bob Dattore is halfway through data from Russian ships (tape 3). This data will be ready Jan 12, 1996.

8. Raob and wind data from countries

Most of the following datasets have been put into one format at NCAR. The data are grouped into "all else" files and sent to NCEP. Figure 3 summarizes much of this information about the time coverage of raobs from countries. NCAR sends other datasets with global coverage that are part of the data merge at NCEP.

8.1 Canada, including early data, 1940s-on. We have a text about Canadian raobs that includes inventories and information about data that have never been digitized.

8.2 U.S. raob stations (1948-on), see other text

8.3 U.S. control for Mexico, Caribbean, Pacific Islands, West Coast of S. America, some in SE Asia, etc.

8.4 Brazil
   — raobs (1961-86), some earlier
   — winds only (1951-86), 30 stations

8.5 Argentina (1958-91)

8.6 Ascension Island (1953-69), from USAF

8.7 China (1954-62). E. Kung processed data for 1958-62. NCAR will put it into a standard format.

8.8 India (1950-78), 55 MB

8.9 Singapore (1957-90)
8.10 Australia (1943-87), 94 MB soundings and 183 MB wind only
8.11 New Zealand (1955-75), 12 MB soundings and 29 MB wind only
8.12 French Polynesia (1950s-1973), 16 MB, from France
8.13 French Africa (1951-65), data for seven African stations
8.14 S. Africa (1968-89)
8.15 UK overseas (about 1950-70), from UK

9. **Raob and dropsonde data from projects**

Data from the following projects are being used for reanalysis:

9.1 Raob data for IGY (Jul 1957 - Dec 1958)

   I think that NCDC prepared data on punched cards for this project, and that the data got into one of several datasets that we have.

9.2 GATE ships, drops, some land raobs (summer 1974)

9.3 FGGE ships and drops (Dec 1978 - Nov 1979)

9.4 Line Islands (Feb-Apr 1967)

9.5 TOGA COARE (Nov 1992 thru Feb 1993). UCAR had these raobs online. About 90% of them were available when NCEP got the data about Jun 1995. This includes data from three ships.

10. **Constant level balloon data for the S. Hemisphere**

    There are texts at NCAR about TWERLE data and EOLE data.

10.1 TWERLE balloons (Jul 1975 - Aug 1976)

    Bob Dattore can convert TWERLE data in 1 or 2 days (will send it 13 Jan 1996).

10.2 EOLE balloons (Aug 1971 - Dec 1972)

    Use the data for 27 Aug 1971 - 5 Jul 1972 if we have time to calculate winds.
Schedule to send these balloon data

- TWERLE balloon data will be ready to send about Jan 17, 1996. There will be lots of wind data for mid-latitude balloons. We will not have time to calculate additional winds in the tropics.

- Send EOLE balloon data (Aug 1971 - Jul 1972) about early Feb 1996??

11. Aircraft data

There is a temporary text "Status of Aircraft Data for Reanalysis," 3 Jan 1996. Figure 4 summarizes the availability of aircraft data.

11.1 On NCEP tapes
   - the 1973-on dataset
   - the 1962-72 dataset

11.2 Sadler data (1960-73), mostly for the tropical region.

11.3 FGGE data (Dec 1978 - Nov 1979). Data is in the FGGE dataset and it has been used. This includes data for FGGE, and summer and winter monex.

11.4 GATE data (1974), research and civil data. This has been used.

11.5 NASA GASPS (Mar 1975 - Jul 1979), tropical routes

11.6 New Zealand (Oct 1978 - Jun 1988) - this has been used.

11.7 USAF decode (1976-85)

11.8 Australian GTS (Dec 1971-89) - not decoded.

11.9 Navy decode (1971-on) - this has not been used.

12. Cloud drift winds

12.1 On NCEP tapes
   - the 1973-on dataset
   - the 1962-72 dataset

NCAR has data from SDSD starting Jan 1974 for several years. We will send the data for 1974-78 to NCEP. Much of the data will overlap with what is on the NCEP tapes.

12.2 Basic archive of GMS winds (1978-91)

Japan sent tapes to NCEP with the original archive of cloud wind calculated from GMS. This is original data. Kanamitsu verified that JMA did not make new wind calculations.

12.3 Original archive of Meteosat data

This archive wasn't available, but most data should be on the NCEP tapes.

12.4 Cloud drift winds for FGGE.

These are in the FGGE dataset and have been used for reanalysis. The data includes winds calculated from a U.S. GOES satellite over the Indian ocean.

13. Satellite sounder data

Figure 1 summarizes much of the availability of satellite sounder data. In the 1970s, there are stratospheric channels from some NASA satellites that are not in the figure.

13.1 TOVS Level 1 data (29 Oct 1978-96)
— TOVS 2.5° data (Jan 1979-96) was used for reanalysis

13.2 VTPR, Level 1, CCR, and soundings
— VTPR data is for 20 Nov 1972 - 28 Feb 1979 (overlaps with TOVS)
— Use this data for reanalysis
— NCEP has looked at a data sample. They will have the whole set of VTPR data by 17 Jan 1996.
— There are 23 gaps of 2 days or longer. The longest four are 21, 18, 7, and 5 days. All gaps except for one or two shorter ones can be filled using the 1973-79 dataset below.

13.3 Nimbus SIRS (Apr 1969 - Sep 1972)
— Level 1 SIRS data is for Apr 1969 - Apr 1971 and is at NCAR.
— Soundings on NCEP tapes are for May 1969 - Sep 1972, with some gaps (Table 1).

— Ignore this data for this reanalyses effort. Use VTPR data during the period.

13.5 LIMS strato sounder (Oct 25, 1978 - May 27, 1979)
— Data for Dec 1978 - May 1979 is in the FGGE dataset (Version 4).
— Version 5 should really be used.
13.6 VTPR satellite soundings on NCEP tapes (1973-79)

During Jan 1973 thru Jan 6, 1979, satellite soundings were on the NCEP tapes that were sent to NCAR. The volume of all this is 1.3 Gbytes. Then NOAA kept these data only in the separate archives that are being used for reanalyses.

These data will be used to fill gaps in the separate archive of soundings of VTPR data described above.

13.7 Data from UARS satellite (Sep 1991-95)

UARS includes stratospheric data for temperature and winds. It is not being used in reanalysis. A CD-ROM is available that has data for Sep 1991 to Sep 1994. A text about UARS data is almost ready.

13.8 Data for stratospheric channels during the 1970s

The SSU instrument in TOVS started Oct 1978, and is being used in reanalysis. Earlier strato data from NASA satellites is not being used.

Schedule to send sounder data

- NCEP has 2.5° TOVS data (1979-95)
- Send all VTPR (SDSD set) by 12 Jan 1996
- Send VTPR gap fillers from NCEP tapes by 26 Jan 1996

14. Surface land synoptic data (1940s-on)

14.1 Surface synoptic data (3 hourly) from NCEP decode (Jul 1976 - Dec 1995)

14.2 USAF decode (3 hourly) (1967 to 1980)

This data is the Dick Davis version. Use this version for Jan 1967 - Dec 1976.

14.3 USAF TD13, land synoptic data (early 1900s to 1971)

Prepare this data for 1940s-on, in synoptic sort

14.4 Icecap buoy data from the University of Wisconsin (1980-93)

There are 3-hourly data for stations on and around the Antarctic ice cap (1980-93) and for Greenland (1987-93). Both NCEP and ECMWF got the data for 1980-92, and I assume that it was used. See "Status of Reanalysis," Oct 1995, for more information.
Notes:  
(1) Both NCEP and USAF data will be used during Jul-Dec 1976.
(3) NCAR has Navy surface synop data (from GTS) starting Oct 1966. It isn't being used for reanalysis because the other sources should be enough.

Schedule to send synoptic land data

- NCEP (1976-77) - send by 19 Jan 1996
- USAF (1971-76) - send by 19 Jan 1996
- TD13 (~1947-71) - send by mid Feb 1996

15. Backup bogus data for S. Hemisphere analysis in early years

If the Reanalysis Project has analysis and first guess problems south of 30°S in early years, NCAR has some analyses for 1958-on that could be used as an aid. The trouble is that they do not cover the whole area for the whole period. More information is available.

16. More information

Attachment 1

Summary of Plans to Send Data to NCEP in Jan 1996

1. Raob and wind data

   - NCEP data (1978-on) - NCAR already sent this data to NCEP.
   - NCEP data (1968-72) - NCAR will send this by Feb 9.
   - NCAR all-else data (1973-on) - NCEP already has this.
   - NCAR all-else data (1968-72) - NCAR will send this by Jan 18.
     - The new Russian ship data will be in
     - Leave out delayed Canada and U.S. data (receipt is okay in GTS)
     - Leave out U.S. ship set TD6210? (time)
     - Include N. Pole Russian ice island raobs by Feb 15.

2. Aircraft data

   - NCEP data (1978-on) - NCAR already sent this to NCEP.
   - NCEP data (1973-77) - NCAR will send this by Jan 26, maybe earlier.
   - NCEP data (1968-72) - NCAR will send this by Feb 9.
   - NCAR all-else data - NCAR will send this by Jan 22.

   Notes:  (1) Balloon data (TWERLE and EOLE) will be included. See the text for dates when available.
           (2) Includes data from projects and countries
                — NCAR will put in USAF set for 1976-78
                — leave out the Australia set (there's no time, it is rather well covered)

3. Cloud drift winds data

   - NCEP data (1978-on) - NCAR already sent this to NCEP.
   - NCEP data (1973-77) - NCAR will send this by Jan 26 or sooner.
   - NCEP data (1968-72) - NCAR will send this by Feb 9.
   - U.S. GOES SDSD set for 1974-77 (ask Woollen) - NCAR will send this by Jan 26.

4. COADS data

   - NCEP has FGGE ship data.
   - NCEP has COADS data for Dec 1979 and 1980-93.
   - NCAR will send COADS data for 1970-79 by Jan 12.
5. Satellite sounders

- NCEP has 2.5" TOVS data for Jan 1979-on.
- NCAR sent VTPR data for 1972-79 on Jan 10 (fast mail). This was NOAA SDSD data (it has Level 1, CCR, soundings).
- NCAR will send VTPR gap fillers from NCEP tapes for 1972-78 (NCAR will send the whole set by Jan 26).
- NCAR will send the SIRS file on NCEP tapes for 1969-72 by about Jan 25.

6. Surface land synop data

See the text in the first section.
### Table 1

**SIRS satellite sounder data on original NMC tapes, Apr69-Dec72**

(00Z and 12Z data are shown separately and combined)

#### 00Z hour, year/month summary

<table>
<thead>
<tr>
<th>SIRS (type 51, Apr69- )</th>
<th>year</th>
<th>jan</th>
<th>feb</th>
<th>mar</th>
<th>apr</th>
<th>may</th>
<th>jun</th>
<th>jul</th>
<th>aug</th>
<th>sep</th>
<th>oct</th>
<th>nov</th>
<th>dec</th>
<th>ann</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1969</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1683</td>
<td>4165</td>
<td>4848</td>
</tr>
<tr>
<td></td>
<td>1970</td>
<td>3387</td>
<td>3417</td>
<td>1575</td>
<td>1000</td>
<td>1223</td>
<td>886</td>
<td>1595</td>
<td>1886</td>
<td>1833</td>
<td>1920</td>
<td>2854</td>
<td>2609</td>
<td>24185</td>
</tr>
<tr>
<td></td>
<td>1971</td>
<td>1788</td>
<td>1994</td>
<td>2255</td>
<td>528</td>
<td>656</td>
<td>1021</td>
<td>777</td>
<td>644</td>
<td>0</td>
<td>0</td>
<td>1710</td>
<td>1334</td>
<td>12707</td>
</tr>
<tr>
<td></td>
<td>1972</td>
<td>1487</td>
<td>868</td>
<td>1560</td>
<td>1078</td>
<td>1714</td>
<td>644</td>
<td>207</td>
<td>0</td>
<td>354</td>
<td>0</td>
<td>0</td>
<td>7912</td>
<td></td>
</tr>
<tr>
<td><strong>total</strong></td>
<td></td>
<td>6662</td>
<td>6279</td>
<td>5390</td>
<td>2606</td>
<td>3593</td>
<td>2551</td>
<td>2579</td>
<td>2530</td>
<td>2187</td>
<td>1920</td>
<td>6247</td>
<td>7108</td>
<td>49652</td>
</tr>
</tbody>
</table>

#### 12Z hour, year/month summary

<table>
<thead>
<tr>
<th>SIRS (type 51, Apr69- )</th>
<th>year</th>
<th>jan</th>
<th>feb</th>
<th>mar</th>
<th>apr</th>
<th>may</th>
<th>jun</th>
<th>jul</th>
<th>aug</th>
<th>sep</th>
<th>oct</th>
<th>nov</th>
<th>dec</th>
<th>ann</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1969</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>229</td>
<td>3877</td>
<td>7695</td>
<td>9632</td>
<td>7735</td>
<td>5563</td>
<td>2056</td>
<td>2977</td>
<td>39764</td>
</tr>
<tr>
<td></td>
<td>1970</td>
<td>4080</td>
<td>3732</td>
<td>2152</td>
<td>2283</td>
<td>2750</td>
<td>2144</td>
<td>2834</td>
<td>2564</td>
<td>2611</td>
<td>2813</td>
<td>4012</td>
<td>4112</td>
<td>36087</td>
</tr>
<tr>
<td></td>
<td>1971</td>
<td>2066</td>
<td>3037</td>
<td>2934</td>
<td>547</td>
<td>1648</td>
<td>1074</td>
<td>1420</td>
<td>979</td>
<td>0</td>
<td>0</td>
<td>1896</td>
<td>2407</td>
<td>18008</td>
</tr>
<tr>
<td></td>
<td>1972</td>
<td>2722</td>
<td>1400</td>
<td>2348</td>
<td>1824</td>
<td>2007</td>
<td>622</td>
<td>247</td>
<td>252</td>
<td>0</td>
<td>57</td>
<td>0</td>
<td>0</td>
<td>11179</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td></td>
<td>8868</td>
<td>8169</td>
<td>7434</td>
<td>4454</td>
<td>6634</td>
<td>7717</td>
<td>12196</td>
<td>13327</td>
<td>10346</td>
<td>8433</td>
<td>7964</td>
<td>9496</td>
<td>105038</td>
</tr>
</tbody>
</table>

#### Both 00Z and 12Z hours, year/month summary

<table>
<thead>
<tr>
<th>SIRS (type 51, Apr69- )</th>
<th>year</th>
<th>jan</th>
<th>feb</th>
<th>mar</th>
<th>apr</th>
<th>may</th>
<th>jun</th>
<th>jul</th>
<th>aug</th>
<th>sep</th>
<th>oct</th>
<th>nov</th>
<th>dec</th>
<th>ann</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1969</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>229</td>
<td>3877</td>
<td>7695</td>
<td>9632</td>
<td>7735</td>
<td>5563</td>
<td>3739</td>
<td>6142</td>
<td>44612</td>
</tr>
<tr>
<td></td>
<td>1970</td>
<td>7467</td>
<td>7149</td>
<td>3729</td>
<td>3283</td>
<td>3973</td>
<td>3030</td>
<td>4429</td>
<td>4450</td>
<td>4444</td>
<td>4733</td>
<td>6866</td>
<td>6721</td>
<td>60272</td>
</tr>
<tr>
<td></td>
<td>1971</td>
<td>3854</td>
<td>5031</td>
<td>5189</td>
<td>1075</td>
<td>2304</td>
<td>2095</td>
<td>2197</td>
<td>1623</td>
<td>0</td>
<td>0</td>
<td>3606</td>
<td>3741</td>
<td>30715</td>
</tr>
<tr>
<td></td>
<td>1972</td>
<td>4209</td>
<td>2268</td>
<td>3908</td>
<td>2702</td>
<td>3721</td>
<td>1266</td>
<td>454</td>
<td>152</td>
<td>354</td>
<td>57</td>
<td>0</td>
<td>19091</td>
<td></td>
</tr>
<tr>
<td><strong>total</strong></td>
<td></td>
<td>15530</td>
<td>14448</td>
<td>12824</td>
<td>7060</td>
<td>10227</td>
<td>10268</td>
<td>14775</td>
<td>15857</td>
<td>12533</td>
<td>10353</td>
<td>14211</td>
<td>16604</td>
<td>154690</td>
</tr>
</tbody>
</table>
Note 1: We will not use HIRS data (1975-76) for reanalysis
Note 2: UARS data (for stratosphere) started Sep 1991. Not used.

Figure 1. Satellite Sounders
Available daily analyses for the S. Hemisphere at NCAR

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>Early sfc analyses</td>
</tr>
<tr>
<td></td>
<td>1968 S. Hemisphere 500 mb</td>
</tr>
<tr>
<td>Jun</td>
<td>1957 40% of Hemisphere</td>
</tr>
<tr>
<td></td>
<td>Dec 1957, IGY SLP and 500 mb</td>
</tr>
<tr>
<td></td>
<td>S. Africa drew maps</td>
</tr>
<tr>
<td>1970</td>
<td>Apr 1972 Australian analyses - stack</td>
</tr>
<tr>
<td>1980</td>
<td>SLP from New Zealand</td>
</tr>
</tbody>
</table>

Figure 2

R. Jenny
Mar 1995
### Rawinsonde Data from Countries

<table>
<thead>
<tr>
<th>Year</th>
<th>Data Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1950</td>
<td>'55 Canada</td>
</tr>
<tr>
<td>1960</td>
<td>U.S. Control</td>
</tr>
<tr>
<td>1970</td>
<td>West Coast of S. America</td>
</tr>
<tr>
<td>1980</td>
<td>'61 Brazil '86</td>
</tr>
<tr>
<td>1990</td>
<td>'58 Argentina '91</td>
</tr>
<tr>
<td></td>
<td>'54 '62 China</td>
</tr>
<tr>
<td></td>
<td>India '78 55 MB</td>
</tr>
<tr>
<td></td>
<td>'57 Singapore</td>
</tr>
<tr>
<td></td>
<td>Australia '87</td>
</tr>
<tr>
<td></td>
<td>New Zealand '75 12 MB, winds only</td>
</tr>
<tr>
<td></td>
<td>French Polynesia '73 16 MB</td>
</tr>
<tr>
<td></td>
<td>'51 French Africa '65</td>
</tr>
<tr>
<td></td>
<td>S. Africa '89</td>
</tr>
<tr>
<td></td>
<td>UK overseas '84 MB</td>
</tr>
<tr>
<td></td>
<td>Ascension Island</td>
</tr>
<tr>
<td></td>
<td>Antarctica (different stations, not all yet)</td>
</tr>
<tr>
<td></td>
<td>'61 Russian ship rawinsondes '87</td>
</tr>
</tbody>
</table>

**Experiments:**
- GATE ships, drops and some land (summer 1974, 50% complete)
- Line Islands (Feb-Apr 1967), 50% done

---

**Figure 3**

Roy Jenne
Mar 1995
Aircraft Data
Aircraft and balloon data at NCAR.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1947</td>
<td>AF Recon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dec</td>
<td></td>
<td>62 MB</td>
<td>133 MB</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1959</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sadler</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2690 MB</td>
<td></td>
<td>Mar</td>
<td>1016 MB</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1962</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NMC Aircraft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>1976 USAF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>1985 New Zealand</td>
<td></td>
<td></td>
<td></td>
<td>1985</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oct '78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jun '88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1971 Australia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tropical widebody aircraft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>331 MB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Research and civil aircraft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>72 MB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Balloons covered S. Hem near 150 mb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GASP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Mar Jul '75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Jun-Aug '74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FGGE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1979</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TWERLE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Jun Aug '75 '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Options:
Navy aircraft, 1971-on (NCAR has)
Japan aircraft, ~1979-on (don't have), NMC has this

Roy Jenne
Mar 1995

Figure 4
5.0 PLANS FOR ISOLATED STATIONS AND ANTARCTICA

5.1 Data for Permanent Ships and Antarctica

NCAR has data for a number of these ships now, but not all. We will compare inventories with NCDC. Then NCDC will take a lead role in obtaining additional ship data. We hope to obtain data inputs from the USSR.

Data for Antarctica will be handled in a similar way.

5.2 Arctic Rawinsonde Data

Jonathan Kahl has been in the process (since about September 1988) of assembling a set of upper-air data for the Arctic (all available) since about 1949. He has digitized much data including Ptarmigan dropsondes. He has obtained many Arctic station raobs from NCAR. A list of a few types of data (not the land stations) follows:

<table>
<thead>
<tr>
<th>Station</th>
<th>Date</th>
<th>Approximate Number Soundings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ice island alpha</td>
<td>July 57 - December 58</td>
<td>1000</td>
</tr>
<tr>
<td>Ice island T3</td>
<td>June 52 - March 53</td>
<td>350</td>
</tr>
<tr>
<td>T3</td>
<td>June 66 - May 68</td>
<td>2000</td>
</tr>
<tr>
<td>T3</td>
<td>January 68 - April 71</td>
<td>2500 (see overlap)</td>
</tr>
<tr>
<td>Tarmigan drops</td>
<td>1949 - 62</td>
<td>3000</td>
</tr>
<tr>
<td>Ice-fast ship soundings</td>
<td>--</td>
<td>100,000</td>
</tr>
<tr>
<td>Cearex ship, ice island</td>
<td>1988 - 89</td>
<td>500</td>
</tr>
<tr>
<td>AGasp drops</td>
<td>1983, 86, 89</td>
<td>200</td>
</tr>
<tr>
<td>Two USSR ice island</td>
<td>Prob from late 1940s</td>
<td>not yet</td>
</tr>
</tbody>
</table>

NCAR will work with Jon Kahl to obtain data we don’t already have.

5.3 Arctic Buoy and Other Surface Data

A separate memo is available that traces the data available (or almost available) from about 1950 on. The University of Washington Polar Science Center (Roger Colony) is very active in pulling these data together.

6.0 DATA FROM PROJECTS

We will include data from several of the past projects, as time permits.

6.1 IGY Data (July 1957-Dec. 1958)

I believe that NCDC prepared surface and upper-air data on punched cards (global) for this period. Did the data get onto tape? See the discussion for card deck 510.

6.2 Data for GATE (June-Sept. 1974)

If time is available, we will try hard to include the GATE ship data, aircraft and cloud winds. Good radar precipitation data is available for this tropical Atlantic Ocean ship array. If the input data is included in the reanalysis data, some valuable checks of model output could be made (such as model precipitation compared with radar precipitation).
The NCEP/NCAR Project to Reanalyze the Atmosphere

In this paper we will discuss the Reanalysis Project. The main subjects covered will be:

- Why do a reanalysis project (why make new atmospheric analyses of old data)?
- What is the present status of the NCEP/NCAR Reanalysis Project?
- Other reanalysis projects.

It is important to have good information about the state of the whole atmosphere for as many years as possible (temperature, winds, moisture, etc.). In the Reanalysis Project, we are preparing new analyses of the state of the whole atmosphere, for each 6-hour period, for 40 years. Data for about 16 years are already done. If a reasonably consistent time series of global analyses are available, they can be used for studies such as:

- What are the changes in atmospheric circulation (and associated precipitation patterns) that are associated with several El Niño episodes and other long-term changes in the atmosphere?
- How have the overall circulation and energy exchanges of the atmosphere changed with time?
- To study the interconnections between circulation patterns (and weather events) in different parts of the world. These and El Niño studies form the basis of efforts to prepare forecast outlooks for one season and one year in advance.
- To provide better circulation statistics of the real atmosphere that can be compared with climate model simulations of the present atmosphere.
- The transport of pollution, trace gases, and aerosols is of interest. The reanalysis winds will help these studies.
- The ocean surface winds and surface energy budgets can be used to drive ocean models.
- To help enable studies of land surface processes. Land surface studies use information about wind, temperature, evaporation, precipitation, snow, radiation, hydrology, etc.
- The reanalysis helps the development of forecast models. All of the comparisons and checks on reanalysis output helps to improve these forecast models.

1. The goal of reanalysis

The goal of the NCEP/NCAR Reanalyses Project is to prepare new analyses of the atmosphere (winds, temperature, etc.) for each 6-hour period, for 40 years (1957-96). Also, we hope that it may be possible to extend this period back to 1946. Modern methods to prepare analyses combine the best of analytical methods with the capability of the best
forecast models. These are called data assimilation methods. The forecast models permit us to obtain rather good analyses over places like the N. Pacific Ocean where there are not a lot of observations. The models also give us a lot of diagnostic information about precipitation, radiation, clouds, etc. To help evaluate the output from reanalysis, these data are compared directly with other surface measurements (e.g., precipitation, snow, radiation) and satellite measurements (the earth’s outgoing longwave and shortwave radiation, ocean precipitation, cloud data, etc.).

The ability to prepare global analyses has improved a lot with time. This means that more information can now be extracted from the observations, and more observations can be provided than were available in real time. Therefore, the new analyses will be much better than those already available for 1978-on; they will be more consistent in time, and global analyses for earlier years will also become available from this project.

How does this reanalysis project relate to global change issues? Almost everyone sees plots that show estimates of how the world’s surface temperature has changed during the past 130 years, or how the global upper air temperature has changed in the past 17 to 35 years. To help understand what any of the trends mean, much more information about the whole structure of the atmosphere is needed—including the energy transports, etc. Reanalysis will supply much of this data.

Projects to prepare reanalyses will be seen as some of the big projects for the atmosphere and ocean sciences during the 1990s.

2. **Summary of project timing**

A summary of the timing of the Reanalyses Project follows:

- **1987-90:** The desirability and possibility of preparing a 10-year (or more) reanalysis was beginning to occur to some people. The movement spread.
- **1991-94:** Both NCEP and NCAR worked on projects to prepare data, implement operational methods, and improve analysis methods, using many years of previous work and experience.
- **Jun 1994:** The production of analyses started.
- **7 Sep 1995:** 11.0 years were completed (1982-93, but not 1984)
- **5 Jan 1996:** 14.6 years (1982-95 and Jan-Jul 1979)
- **16 Jan 1996:** 15.5 years were completed
- **11 Feb 1996:** 17.0 years were completed (1979-95)
- **Jan - Feb 1996:** NCAR is sending NCEP the rest of the observations for 1968-77. Then NCEP will have input data for 1968-95 (28 years).
3. **Doing a reanalysis project is very complex**

A project to prepare new analyses for many years is very difficult. The improvement of the models, the operational procedures (including automated checks for various data and analysis problems), and the preparation of many datasets of the world's observations are all very difficult tasks. NCEP considers the reanalysis project to be the largest and most challenging ever attempted at NCEP. Everything took much longer than their most pessimistic estimates. The difficulties of dealing with so many different data (observations and output) at NCEP is huge. The tasks to collect, check, and organize these many sets of observations at NCAR is very challenging. It is only possible because of past work at NCAR, joint projects (COADS), real-time data from NCEP starting in 1962, work done in many countries (giving permanent archives), and global data flow over the GTS.

A number of years have had to be analyzed twice to correct some problems (such as some bad satellite SSMI wind derivations, and some bad surface temperatures under selected conditions). These problems (as well as some of the hardware problems) are documented in periodic status reports that are available. A list of reports that describe project status, the coverage of data, and information about different types of data is available. Some of these reports are still being extended.

4. **What types of data are used for reanalysis?**

Reanalysis projects use many types of observations. The beauty of modern assimilation's methods is that the observations can have irregular spacing, and the types of observations can vary with time and space. We only need enough observations to define the system, given the ability of models to fill gaps. Some of the main types of observations are given in Table 1.

5. **Handling the data output**

The diversity and volume of output data is both an opportunity and a challenge. The documentation of procedures and algorithms takes time. The data volume means that we have to develop new methods to deliver data that reduce the time for a user to handle data, and which sharply reduces the cost. Technologies such as CD-ROMs (for parts of the data), and bulk delivery methods (on tape) will be used.

6. **A new paper; a new CD-ROM**

A major paper that describes the Reanalyses Project and its methods will be in the April 1996 issue of the *Bulletin of the American Meteorological Society*. For the first time, a CD-ROM will be included with the mailing of this issue of the Bulletin. It will have a selection of output from the Reanalysis Project that covers a 13-year period (1982-94). This mailing goes to about 13,000 people. A description of the contents of this CD-ROM is available.
In addition, NCEP is preparing one CD-ROM for each year of reanalysis. It will have a larger sample of the output. NCAR will help to distribute these CD-ROMs—at a rather low price. The bigger archives of output will be at NCAR, NCDC, and NOAA ERL (Boulder).

7. Plans to make another reanalysis

We do not view reanalysis as a project that should only be done once. Just as the community found it very beneficial to analyze FGGE data more than once, the same will be true of reanalysis. Why should this be true? The methods used for reanalysis are based on the early 1995 state of the art of data assimilation. This was a big improvement compared with 1990 or 1980 capability, but we expect a significant set of advances by the 1998 time frame. The process of making a reanalysis creates a good product, and it is also a learning experience on how to improve the methods and the models. We would also like to increase the resolution for the next reanalysis. The reanalysis gives us information about the observations (biases, errors) that can be used in the next reanalysis. Also, we expect that some additional observations can be made available for later reanalyses (e.g., UARS, SCR channels in 1970s, and more old raobs) that were not used for the first project. SCR means "Selective Chopper Radiometer."

8. Are there other big reanalysis projects?

There are two other large reanalysis projects. NASA Goddard has a data assimilation office. By Dec 1995 they had completed about 10 years of reanalysis (1985-on). This project gives them experience to develop systems to assimilate the coming data from EOS satellites.

ECMWF (European Centre for Medium Range Weather Forecasting) in England also has a major reanalysis project. In early May 1995 they had completed 3 years (1979-81). They completed 8 years by 20 Oct 1995, and plan to have 15 years (1979-93) by July 1996.

9. Home pages with reanalysis information

NCEP URL


NCAR URL

http://www.ucar.edu:80/dss/pub/reanalysis.html

NOAA CDC URL

http://www.cdc.noaa.gov/cdc/reanalysis/reanalysis.html
Table 1. Data Inputs for Reanalysis

1. Upper air raob and pibal data
   - Data from NCEP for 1962-95
   - Data from many countries
2. Aircraft data
   - Several sources
3. Satellite cloud drift winds
   - NCEP decode
   - Plus the original Japanese archives
4. Satellite sounders (cloud-cleared radiances)
   - TOVS (1979-95)
   - VTPR (1972-79)
   - SIRS (1969-72)
5. Surface marine data — COADS
   - Ship and buoy data
   - Arctic ice buoys
6. Land synoptic (3-hour observations)
   - For 1947-95
   - From NCEP, USAF
   - From many countries for early years
7. Boundary conditions
   - SST: sea surface temperature; weekly and monthly analyses
   - Sea ice: daily ice cover calculations from satellite microwave
What is the Status of Reanalysis?

We will give a short status report on the NCEP/NCAR reanalysis project. The new analyses are completed for 17 years (1979-95). Also, a little information will be given about the projects at ECMWF and at NASA/Goddard.

1. A Short History of the NCEP/NCAR Reanalysis Project

Following is a short history:

<table>
<thead>
<tr>
<th>DATE</th>
<th>COMMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>Some funding started to work on Data operational methods, and models.</td>
</tr>
<tr>
<td>Jun 1994</td>
<td>The production of analysis started.</td>
</tr>
<tr>
<td>7 Sep 1995</td>
<td>11.0 years were completed</td>
</tr>
<tr>
<td>5 Jan 1996</td>
<td>14.6 years (1982-95, some 1979)</td>
</tr>
<tr>
<td>16 Jan 1996</td>
<td>15.5 years</td>
</tr>
<tr>
<td>11 Feb 1996</td>
<td>17.0 years (1979-95)</td>
</tr>
</tbody>
</table>

*NOTE:* A number of years had to be analyzed twice because of various problems.

2. Computer Disruptions at NCEP

In summer 1995, NCEP prepared software so that they could move reanalysis from the Cray YMP-8 to the new Cray Jedi with 16 processors.

- The Cray YMP departed Oct 1995
- There were many hardware and software problems Aug-Nov 1995. Several of these corrupted data files which then had to be fixed. It was a very trying period.
- The Cray YMP (with 8 processors) could deliver a real speed of 1.08 Gflops on a climate model (at NCAR). With a job mix, it delivers almost one Gflop. The Cray Jedi (with 16 processors) allowed NCEP to do reanalysis a little bit faster.

3. Sending Observations from NCAR to NCEP

The main types of observations for the world that are being used for reanalysis are:

- Upper air Data from rawinsonde balloons
- Pibals and other winds only data
- Aircraft reports (mainly winds)
- Satellite Sounder Data (thermal structure of atmosphere)
- Satellite cloud drift winds
- Surface land synoptic reports (from 7500 stations)
- COADS surface ocean marine data. Includes ships, drifting buoys, fixed buoys, and ice buoys.
The status of moving data to NCEP is as follows. We note that this data preparation is a huge task.

- NCEP has the data for 1979 - 1995 (17 years)
- They now (Mar 27) have 98% of the data for 1968 - 1978 (11 years)
- The next batch that NCAR will send is data for 1957 - 1967 (11 years). Preparation work for this period has been underway for a few years, but several more months are needed for more diagnostic checks, fixes, station location work, and data packaging for NMC.

4. Reanalysis Output Data

The main archives from reanalysis are given below. The volume of data for each year of reanalysis is about 54 Gbytes

a. Data in model coordinates (28 levels, on Gaussian grid, 208 Km resolution)
b. Flux fields: Lots of model generated data: precip, evaporation, surface and top radiation components, clouds, 2m temp, 10m wind, etc. (on Gaussian grid)
c. Global analyses in pressure coordinates, 2.5 degree resolution, 17 levels
d. Monthly means and statistics (2.5 degree)
e. Data in isentropic coordinates
f. All the observations
g. Several other files

5. Data from Reanalysis now at NCAR

NCAR receives 144 cartridges (Double density) with data for each year of reanalysis (about 54 Gbytes). We now have data for the 17 years (1979-95) that have been completed. However, some of the present data has to be replaced because of some recent reruns, and because of a bad wind transformation that affected winds in pressure coordinates poleward of about 80 degrees. The replacements still needed (Mar 27) are:

- Data for 1985 (This year was rerun)
- Data for Oct-Dec 1995 (a recent rerun)
- 1986-93: The data affected by bad winds at the poles (in pressure coordinates) needs to be replaced. The monthly statistics are also affected by this problem.

NOTE: Data on the CD-ROM for the bulletin AMS are all based on the corrected data.

6. CD-ROMs for NCEP/NCAR Reanalysis

For the first time ever, the Bulletin of the AMS for March 1996 will be sent out with a free CD-ROM. This issue started to be delivered to subscribers about Mar 26, 1996.

6.1 The March 1996 issue of Bulletin AMS
- Has a paper about the NCEP/NCAR reanalysis
- A CD-ROM with selected data (mostly monthly reanalysis statistics) for 13 years (1982 - 94). (AMS sent a CD with each copy of the Bulletin.) It also has a subset of daily data for one year (1993). A text describes the contents of this CD-ROM.
- The circulation of the Bulletin is about 13,000 copies.
6.2 A CD-ROM for each year
   - The CD-ROM with a selection of data for the first year (1985) may be ready about May 1996. (An update: It was later than this.)

7. Tests Being Run at NCEP

Some of the projects being worked on at NCEP are:

a. Does the use of VTPR sounder data (for 1973-79) give similar results to the use of TOVS (in the overlap period)? I think the answer will be "yes." Then the years 1973-78 can be analyzed.

b. When there is no satellite data in the west wind belt of the S. Hemisphere, what is the effect on the quality of the analyses in that region?

c. An 8-day forecast is run each 5 days during reanalysis. Do the verification scores change over the 17 years that have been completed?

8. Data to Help the Analysis, 30 -70 South, if Needed

It may turn out that we need to help the analyses in the S. Hemisphere during the years prior to satellite soundings. Sources of manual analyses (to use to create bogus reports) are being documented by NCAR. NCEP is thinking about how such data would be used, if necessary. There are manual analyses for the whole period 1957 - 1972 (and later), but not all of them are in digital form.

9. Status of Reanalysis at ECMWF

ECMWF officially started the production phase of their reanalysis in June 1994, but the runs led to more model development until early 1995. Following is a short history of their progress:

<table>
<thead>
<tr>
<th>DATE</th>
<th>YEARS DONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early May 1995</td>
<td>3.0 (1979-81)</td>
</tr>
<tr>
<td>3 Aug 1995</td>
<td>5.6</td>
</tr>
<tr>
<td>22 Sep 1995</td>
<td>7.0</td>
</tr>
<tr>
<td>20 Oct 1995</td>
<td>8.0</td>
</tr>
<tr>
<td>11 Mar 1996</td>
<td>12.3 (1979-Mar 91)</td>
</tr>
</tbody>
</table>

The model is T106, 31 levels, sigma coordinates. It has the new ECMWF 31-level cloud scheme. ECMWF will probably analyze Jan 79 - Aug 80 again to make the data consistent with later years, and to use the same boundary conditions in 1979 that NCEP used.

They will try hard to complete data through 1994 (16 years), and may be able to do 1995. They can’t run their present reanalysis system after the Cray C-90 leaves on 1 Oct 1996.
10. The NASA Reanalysis at Goddard

The Data Assimilation Office (DAO) at Goddard is preparing a reanalysis, using their GEOS-1 model. The forecast model is a 2° Latitude x 2.5° Longitude model with 20 sigma levels. The actual data assimilation is done at sea level and on 14 upper air pressure levels (resolution also 2° x 2.5°). DAO is gaining experience so that their model can assimilate data from EOS satellites. The progress of their reanalysis has been as follows:

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 1993</td>
<td>Start production</td>
</tr>
<tr>
<td>May 1994</td>
<td>5 years were finished (1985-89)</td>
</tr>
<tr>
<td>20 Jan 1996</td>
<td>9 years were completed (1985-93)</td>
</tr>
<tr>
<td>27 Mar 1996</td>
<td>10 years (also did 1980)</td>
</tr>
</tbody>
</table>

NCEP has been sending DAO a copy of the observations for the 1980-1984 period.

The GEOS-2 model will probably be ready by about June 1996. It will have the same horizontal resolution, but it will have 70 levels. There will be about double the vertical resolution in the lower atmosphere. The top level will be near the mesopause (About 0.01mb, or 70 Km).
Status Reanalysis, June 12, 1996
From talk with Kistler

1. Reanalysis on Dec 1972 was run yesterday.

   The production of 1973-78 (6 years) starts today. Dec took under 12 hours. It was such that Bob could do a year in 8 days; now it may go as fast as 6 days.


   Only one more year to fix. They will be at NCAR soon.


   The basic tapes are corrupted. He will have to rerun the analyses for these months.


   During a period of time, NCEP did not use cloud winds from the US Pacific Satellite, even when the data was available. In the past day, they found that the data was there in the first stage of data, but the Pacific Satellite was not used in the analysis. They found that there was code to exclude this satellite (because there had once been a problem). They will determine how long the error existed (not using data).

   - Goes 9 was used in production (GDAS) starting May 1996
   - It was used in CDAS starting June 1996

5. Data processing is on the old IBM machines.

   NCEP needs to get off the old machines. They have been used about 23 years. NCEP will start a bulletin heading archive (but I should check). A lot of Jack’s time is going to this project.

6. Problem of air temp in deck 555 (COADS).

   Often this deck is biased 0.5 degrees warm in air temperature compared with the same data from ship logs. The last digit is always 0.5 when we looked in Boulder. Deck 555 is GTS Navy and is in COADS. It stopped in mid 1973. It appears that for a period of time the original air temp was rounded to whole degrees and that for later periods truncation was used. In both cases NCDC added 0.5° C which made the first case worse and the 2nd case better. Scott says that GTS Navy dew point is always whole degrees (0.5 is not added). SST is given to tenths. We note
that both air-sea temperature difference, and dew-point spread are important variables that should be as accurate as possible.

7. **How is the tropopause level calculated in the analysis?**

   The model level analyses are used to calculate tropopause data. (A post processor is used). The procedure at NCEP looks for the lowest level above 450 mb where the lapse rate is under 0.2° K per km. A more exact location for the tropopause between two model levels is determined, and the pressure and temperature are calculated for that spot in the vertical. The trop is not allowed to be above 85 mb. This information was from Masao Kanamitsu, NCEP, June 1996. NCEP has a listing of the program used.

   In spite of changing observations, the analyses should be quite stable over time. Therefore the derived tropopause data should be stable over time.
Reanalysis
(Talk with Bob Kistler)

The reanalysis of 1973-78 still has not started. Have been looking more at the input of data (It started 19 June 1996).

1. Aircraft

The character of the aircraft data archive is quite different in 1972 and 1973. The aircraft ID starts in 1973. NMC does not use one isolated aircraft report. This helps them to protect against bad locations.

2. Use of cloud drift winds from the Pacific satellite

When running data for Aug 95, some checks were needed on the geosynchronous satellites that said don’t use winds from the Pacific satellite. Then this code got into the reanalysis processor. Probably it affected 1979-81 (don’t use Pacific winds during these years). But the FGGE data for all satellites were properly used (Dec 78-Nov 79). We should check the month Dec 1979, which is after the FGGE year.

3. Inconsistencies between ice and SST

UK is sending new ice and SST for earlier years. This affects 79-80-81 some too (years already done).

4. How long to do 1957-72 (16 years)

Bob says this period would take a minimum of 16 weeks (year a week). I think that various issues will almost certainly make reanalysis slower than this. Bob thinks most trouble is probably past. Computer changes, etc.

5. Discussion at NCEP over rerun questions

Some people wanted to rerun everything soon because of the perception problem. The reruns go fast. Most of the data flow of output is simultaneous with the calculations.

6. Bill Large at NCAR--user of reanalysis for ocean work

On June 14, he told me that he had just heard about the Southern Hemisphere bogus problem (location of Australian bogus wrong). He said he doesn’t think he could use the affected years in journal papers. The reviewers would probably yell.
7. Tapes to send one year of reanalysis data
   
   • 150 tapes are sent to Boulder
   • 360 tapes are sent to NCDC

   It is quite an effort just to label tapes, put them in the NCEP system, take them out, send them. The cost of buying blank tapes hurts!

8. The 6-hour vs. the month time scale

   On their checks, they have focused mostly on the monthly scale.

9. The NCEP HDF computer (GTS ingest)

   This IBM system is 23 years old. It was supposed to go to rest at the end of June 1996. Maintenance is costing too much. This computer handles a lot of tasks:
   
   • All GTS observation decoding
   • All product generation

   This is a huge job to prepare new programs. Jack Woollen is helping. I am arguing for an archive of bulletins before decode. I think that is in the plans.

10. The calculation of 2 m temp and 10 m wind

   The calculation uses surface physics. The temperature and wind will be in part, a function of surface heat flux and latent flux (a function of mixing). To get these physics terms, it is necessary to run at least one time step in the model.

   For reanalysis, the 2 m temp and 10 m winds are derived from the model levels at the end of a 6-hour forecast. For operations, NCEP recently decided to derive these variables from the model levels at analysis time. This seems better because possible errors in the forecast are not added to the normal uncertainty. The derivation of these variables at analysis time went into operational production in Spring 1996 (but not into reanalysis).

11. Start periods in reanalysis

   We think that there are no significant problems remaining, due to starting reanalysis on selected years. NCEP started reanalysis with 1985. At that time, a good
climatology of some variables (like soil moisture) was not available. The month of
spinup (Dec 1984) was not enough time when the climatology was poor. Now they
have good climate data.

- The year 1985 has been rerun, starting with 1984 reanalysis data (so no spinup
  problem remains).
- 1982 was a start (but it was a warm start on 1 Dec 1981 with a good climate).
- 1979 was a warm start. ($\Delta \omega \Delta 1979-81$)
- 1973 was a warm start (to do 1973-78).
Status Reanalysis Data, July 24

This is a progress report on the preparation of data for reanalysis at NCAR for 1946-1967. NCAR has already sent data for 1968-1995 to NCEP. Also see the reports:

- **Send data for Reanalysis 1957-1967** (7 pp), 9 May 1996
- **Sending 1968-78 data to NCEP** (12 pp), 21 Mar 1996
- **Possible manual map help for reanalysis 1957-1972** (13 pp), 25 Mar 1996
- **Older Rawinsonde data** (31 pp), 1 July 1996
- **Data for Reanalysis; Inventories** (120 pp), Nov 1992
  —has data coverage information

1. **When will NCEP start years 1973-1978**

   Basic sea ice and SST datasets have had some problems that they needed to resolve. This has been going on about six weeks. Now (July 24) they are nearly ready to start production on 1973-78 (in next day or two).

2. **Russian ship raob data (Bob)**

   Doing all the ship track checks is slow work. Sometimes 2 or 3 ships have the same ID so that the tracks have to be separated. The dates for some reports are a month ahead or a month behind (such as June 31, 12Z) and these can be fixed. There are a number of cases where there is a wrong sign on latitude or longitude. There will be a list of raobs where changes have been made.

   - This will still take a few more days (about 3/4 done on July 24).
   - 1968-78 done (ship tracks)
   - Now doing all the rest—take a week


   - No additional work yet
   - The original version has more levels. The new version makes some small changes in temperature that are not desirable. We will use the original, but will make more comparisons with the new version.

4. **Sherhag raobs (Bob)**

   - Fairly far along on this job
   - It will be possible to do without documentation
5. The CD-ROM of North America raobs

- Will has memo about it from Chi-Fan
- Data on it only goes to 100mb
- Some extra data on it, not on our tape (25 AF stations)

6. N. American Rawinsonde data

There is a document that describes the relationship between what data are on the CD-ROM (US and Canada) and what data are on the tapes. It outlines what we need to do to obtain the best data for the 50 states (N. American Rawinsonde Data for Reanalysis, 4 Mar 1996).


There is a document that describes the status of TD54 and the work that we need to do (TD54 World Rawinsondes, 2 pages, 7 Mar 1996).

8. Some stations with raobs

- Singapore is in this set
- Ascension island 1957-71 (on US control tape), 8° S, 14° W. Also Ascension Island is available for 1943-1946.

Ascension Island from Patrick AFB (Jan 68-1981). See DS400. NCAR received tapes with raobs for this period from the USAF.

9. Three sets of US control raob tapes from NCDC

a. Data for US raobs for 1946-47 is on a few separate tapes.

b. Data only for mandatory levels
   In early years, NCAR was able to obtain tapes with raob data on mandatory levels. This was the rich set of levels—data each 50 mbs. We still have these tapes. The years of data are 1961-1970.

c. Raob data with both mandatory levels and sigs.
   In the early 1970's, NCDC prepared tapes with the US raobs that had both mandatory level raob data and the significant levels. Since then we keep updating these tapes. The data are now for 1948-1995.

10. Our set of country raobs (15 July 1996)

We have prepared data for each country that sent raobs and wind data. These will be combined into one dataset to send to NCEP

Will is still making some improvements. Raob data Seq 78 has 3 stations in Europe that start about 1954. Over 400 raobs have only the high levels (lowest is 70-
90mb). One of these may be called a surface level. Other wrong things happen. In this data sequence, most raobs do also include lower levels.

He ran the new hydro check program on an old set of US control raobs. It flagged a lot more data than checks a few years ago and his first reaction was "HELP!" But it actually flagged less data. Most excess flags were because the surface elevation information did not look reasonable.

- We will make a dataset of these raobs from countries.
- This will include many winds-only reports too.
- We will not include the 4 months of early Easter Island (elev. problem).
- We will have the full set of Antarctica raobs included.
- S. Africa, French Islands, etc. will be in.
- It will have the permanent ships, including Canadian P, the US ships, and the 3 other UK permanent ships.

11. The story of Boise raobs

There are problems with the Boise, Idaho raobs during 1948-59 that need to be resolved. On the merge tape from NCDC, the raobs for Boise are bad for 1948-59. This is because the mandatory levels for one time were merged with sig levels from a different time--offset by 12 hours. There are 8847 soundings in this period in our set and 8842 of these are flagged bad by the hydrostatic check.

Consider the inventory of Boise raobs on our merge tape (from NCDC) and from Boise on the CD-ROM. The merge tape shows a solid 2/day raobs for the whole period 1948-1960 and on. The data from the CD-ROM has a fix of the bad-merge problem. But it has far fewer raobs. Starting July 1952, the CD-ROM has a solid 2/day of raobs, but in earlier years, many raobs were tossed out by the cure. It often only has 1 raob per day. In the whole 1948-59 period, there are 1027 fewer Boise raobs on the CD-ROM than are on the merge tape (The CD-ROM has a total of 7820 raobs--compared with 8847--and there are 462 of these that are flagged with height/temp problems).

**Question:** Why should so many Boise raobs be tossed? The implication of a bad merge of sig levels is that we could just leave out the sig levels (if necessary) and still retain most of the information from all of the soundings.

We had hoped that we could use Boise raobs from the CD-ROM for reanalysis. This is not possible for two reasons. First, 1027 raobs were dropped; and second, all of them were cut off at 100mb.

12. UA data for 1962-72, NMC was the source.

This set is the original capture of upper air data from GTS, done by NCEP. It starts March 1962. NCAR got it on a lot of tapes (in about 1973). The data includes rawinsondes, UA winds, aircraft data, cloud drift winds, and satellite soundings.
This dataset was for the N. Hemisphere. It became global in June 1966. There is a text about these data.

- These data have already been sent to NCEP for 1968-1972.
- The data for 1962-67 will be sent in about 2 weeks (written 16 July 96).

Joey has finished most of the work for the UA station library. Some elevations had to be extrapolated back in time. Another check will be made of the library info vs. library data on the tapes.

13. TD52 World Wind aloft data (data for 1922-1971)

This dataset of winds aloft data (pibals, etc.) was gathered together by the USAF section at Asheville during about 1960-73. In mid 1995, the Airforce decided that they no longer wanted this dataset, so they gave it to NCDC in boxes. Dick Davis sent the tapes to NCAR. We received 33 boxes of tapes on April 17-19, 1996 with 211 tapes. The inventories are very encouraging. The old data for the 1930's and 1940's is still on the tapes. The data coverage for Africa is quite good for 1949-1966 (especially for 1952-1965). This will help to compensate for the low counts of raobs in southern Africa during 1963-66.

In July 1996, Debby and Joey key entered library information for over 1000 stations. Gregg is working on the assignment of proper height information to the wind level data.

- We think that this dataset can be ready by Sep 1996.
- There are a total of 3,786,205 reports in this dataset. Each report has data at several levels aloft.
- Several inventories show counts each year, counts in regions, and data coverage for each station.

14. Data from projects

a. Line Islands data (Mar and Apr 1967)
   - Include raob and aircraft data

15. Aircraft Data

This is a brief summary of the aircraft data that will be used for the period up to 1967. Figure 2 summarizes the availability of aircraft data for 1947 - 1995.

- From NCEP decode, Mar 1962 - 1995
- Sadler tropical aircraft, 1960 - 1973
- USAF recon, 1947 - Dec 1959
- Aircraft data from projects (FGGE, Gate, Line Islands, etc.)
- Other sources for 1971 - recent

The following datasets are being used. They cover the period 1946-1995. This section is a summary of information about these datasets.

a. USAF TD13, a global collection of land 3-hour synop data for early years, mostly for 1946-1972. Does not have US data
   - We will use 1946-1966 for reanalysis.
   - The Dick Davis set was thin for coverage of Australia during 1967-76. Would TD13 help? Dennis Joseph prepared plots in July 1996 that show problems in surface synop coverage for Australia and New Zealand from the Dick Davis set. Plots were made from the NCEP set for comparison.

b. Use hourly surface data from about 300 US stations, 1948-1966 (from NCDC)

c. USAF surface land data from GTS, Jan 1967-Dec 1980. NCAR gets this in the form of the "Dick Davis" dataset.
   - We are using Jan 1967-Dec 1976 for reanalysis.
   - Data for 1968-Dec 1976 was sent to NCEP 1 Mar 1996.
   - We will send 1967 to NCEP by about July 26.

d. NCEP decode of data from GTS
   - Data for July 1976-1995 has been sent to NCEP

e. U. Wisconsin Antarctic and Greenland data
   Ice buoys were established with data for about 1980-on.

f. Land surface synoptic data for Russia and the former USSR
Russia provided the US with 3 or 6-hourly surface synoptic data for 223 stations. NCAR is working on this dataset. It is needed for at least the late 1940's and early 1950's.
   - Total period 1936-1985
   - 6-hour data is for 1936-1965
   - 3-hour data is for 1966-1985


The Air Force at Offutt AFB decodes real time GTS data and prepares a dataset called Datsav. Dick Davis at NCDC took the AF data and put it into a character format that was a lot like the original GTS data. At NCAR we often call this the Dick Davis dataset. NCAR obtained a copy of these tapes.
• NCEP has a copy of these tapes for 1968-1976
• NCAR will send data for 1967 by July 26.


This data set was prepared by the USAF from many sources. It is in a format for reanalysis. The supplementary fields are still carried along. NCAR key entered library location data to include with the reports. All of the checks, etc. are essentially done for the whole period.

• Data for 1957-1967 will be sent to NCEP around 6 Aug 1996.
• The earlier data could be sent quickly when needed.

19. COADS; ocean surface marine data

These data are prepared in a joint project between NCAR, NOAA/ERL and NCDC.

• Data for 1970-79 was sent Jan 1996
• Data for 1960-69 was sent to NCEP July 11, 1996
• Data for 1950-59 will be ready about Sept 1996
• For the 34 year period (1960-1993) there are 87,027,014 observations in COADS

20. NCEP needs part of the older data soon

NCEP would like part of the data soon so that they can start preparing it. Therefore when major datasets are ready, we should send them rather soon instead of sending almost everything at a later time.

• COADS data for 1960s was sent July 11.
• The set of country raobs will be sent before some others.
• The UA GTS data (from NMC) for 1962-67 will be sent by Aug 5.
• The TD13 surface synop for 1957-76 will be sent by Aug 5.
• See this text for other comments about dates.

21. Russian North Pole Raobs

The University of Wisconsin (Jon Kahl) previously sent the N. Pole raobs with lower level raob data (under about 700mb). About July 8 they sent more raob data for North Pole ice stations 4-17 which have data for the Period Jan 1957-Oct 1969,

We expected that this data would be only the upper part of raobs that match the lower parts that we got before. Instead, this set includes full raobs that give the other time of day not given before. And it has the upper parts to match our earlier data. Between the upper and lower parts, we thought that data for 3 Km would be
repeated. That data is not repeated but it appears possible to get a good match of the
date/time data so that we can merge soundings.

The new reports have data time in both GMT and local time. The conversion
between GMT and local time does not seem to match the way we converted from
local date/time to GMT in the previous batch. This will need to be resolved.

22. Sea ice and SST from UK met (info July 17)

The issues of monthly sea ice and SST from the UK met office are gradually
getting settled at NCEP. The UK prepares global, monthly, one degree grids. UK has
the ice concentration in percent but only the 100% cutoff coverage got to NCEP. The
NCEP cutoff is at 50% coverage. This is now fixed. the SST near the ice edge had
troubles. The UK fixed this. Now Bob finds that there is some open water (not real)
near the coasts between his land and where the sea ice starts. This will soon be fixed.
There are also seven big bodies of inland seas: Great Lakes, Caspian, Canadian
Lakes, etc. Bob is going to put a satellite based ice climatology into them. There were
some problems with the ice data for these inland areas. The data for 1972 had no
Great Lakes ice. For 1973 data from John Walsh was supposed to be there, but the
numbers didn't appear correct. It is best now just to have a good ice climate for
these regions.

• The monthly global, 1° UK set goes back to 1903.
• Bob Kistler now has the data from 1949-on.
• Kistler hopes to start production of analyses for 1973-1978 in the next few days.
The date now is July 17 1996.

23. Workshop about the ECMWF reanalysis, July 8-10, 1996, England

ECMWF has 15 years done (1979-1993), but 4 years need to be rerun (Jan 79-
Aug 80, Jun 90-Oct 92). The C-90 computer leaves Sep 30; they must be done by that
time. The meeting was mostly about the ECMWF reanalysis, but some information
about the other big projects was given (NCEP/NCAR and Goddard).

24. Reanalysis needs for CLIVAR (climate variability)

There is a need to understand the variability of climate. Lennart Bengtsson
spoke about this issue. Long time series of paleoclimate data need to be gathered
and processed. Lennart encouraged ECMWF and NCEP to prepare reanalyses for as
long a period as possible.

25. Reanalysis plans at NCEP
(By Eugenia Kalnay, July 1996)

• Finish 40 years in 1997 (1957-1996)
• Extend analyses back to 1946-1956
• Generate 1/year CD-ROMs
• Continue active model development. Most present problems have been attached and solved (e.g., convection in S.E. USA, high albedo, spectral rain bands in high lats)
• In 1999-2000, start a second reanalysis with a state-of-the-art system
• Data input and QC will be much easier then

26. ECMWF reanalysis plans
(by David Burridge, July 1996)

1. Complete production (1979-93 or 94) by 30 Sep 1996
2. Do clean-up work (documents, etc.) 1996/97
3. 1997: adapt variational analysis assimilation system to new computer system.
4. 1997/1998: Do observing system experiments (OSE)
6. ????: Do new reanalysis using 4D-Var
7. Maybe get persuaded by Lennart Bengtsson to do a reanalysis starting with data in the late 1940's.
Fixing Old Raobs

1. Deck TD54
   - Find sheet showing hydro failures in each deck of TD54.
   - a number of reports have the "80 mb problem: several strato levels at the wrong pressures.
   - Get an inventory of the number of reports in each station-year-month that have problems with the same raob level (for any level).
   - Fix the strato data and then see if the hydrostatic check is OK.

2. Canada raob data from Canada.
   - It has duplicate sfc levels, and some sig levels are within 1 mb of mandatory levels and both have equal data.
   - We have rules to fix these
   - The station numbers on the tape from Canada were 7 or 8 digit Canadian numbers. Will put them into WBAN numbers.
   - We need a bulk data compare between this Canada raob set and the same data on the US control tapes (NCDC).

   Note: Probably prefer the Canada in US control over the other. The tests will clarify the answer.

3. The C-Cards (1949-1966)
   - This is Northern Hemisphere data tabs.
   - Will is making an inventory showing raobs in (a) TD54 (b) US control (c) all sources not C-Cards (d) C-Cards.

4. Some comments about when Rawinsonde observations started.
   a. US National Stations: The data on these tapes starts in 1948. Other tapes have data starting 1946, but these have not been converted as yet. The early data are in complicated card formats with overpunches.
   b. Mexico: Three stations started in 1948, about two in the 1950's and more later.
   d. Panama Canal Zone: The rawinsonde data are for Nov 1946-on.

5. TD53: US winds aloft data.

   During 10-11 February, 1997, 62 boxes of tapes arrived from Dick Davis at NCDC. Also, a bag has two copies of documentation, each 0.75 inches thick. Each box has five tapes (these look like original round tapes from about ten years ago). My guess is that these are 1600 BPI.
NCAR TN IA/III (1975) has inventory information that the AF gave me in 1974. Region for TD52 is North America and US stations abroad. Years 1920-1967, 85,331 station year months of data (312 tape reels, 9 track, 800 BPI).


Roy Barnes has sent the 1957-67 TD13 data to NCEP. He will send 1946-56 data by June 1997.
Data for Reanalysis, March 1997
An Update on Reanalysis, March 1997

Section A: Talk with Bob Kistler on March 11.

1. Our status at NCAR on sending observations (March 11, 1997).

I briefed Bob on where we stand on the data (March 11). The data preparation is coming along, but it is slower than we hoped.

NCAR is finding:

a. There are a few stations in US control each with a few years of record, where the data time is off by 12 or more hours (stations such as San Diego, Kwajalein Island, etc.). In addition, in these cases, mandatory levels and sigs were merged that were not for the same time. A mess. So far, the data for these same stations (mandatory levels only) is OK in TD54. So we will toss the data from US control and use it from TD54.

   • March 13: We have now found ten stations with this type of problem.

b. A few stations (these seem to be military sites) in US control have constant data during day after day, for several years. It is interesting that the number of levels of the constant data does change from day to day. This data error was the same in US control and in TD54 so the massive data compares did not identify this error. We will eliminate this bad data. These observations are probably lost forever, as we do not know another source. I do not think that this will hurt reanalysis very much because there are often other stations in the vicinity.

c. I told Bob Kistler that we will keep him informed about where we stand. I noted that the work done to fix data for 1957-67 will help a lot with the earlier years 1946-56.

Bob noted that in summer 1996, he had to spend a good deal of time to figure out ways to handle SST and sea ice for the earlier years. The data that he received had many problems to overcome. But his solutions solved the problems for all of the early years. He thinks that we at NCAR are in the same situation now in solving some of the long standing problems of the early datasets. This seems to be true.
2. Status of reanalysis at NCEP.

There have been a series of problems to slow down operations at NCEP.

a. On March 3, Bob told me that Eugenia Kalnay had wanted statistics for 1979 about how well the first guess and the analysis fit the Antarctic stations during FGGE. It turned out that the Antarctic raobs were not used. The FGGE data was received from ECMWF in BUFR format. The raobs were in the data, but the NCEP BUFR ingester does not find them. This is a vice of several of the more complicated data format systems. The present NCEP BUFR ingester does find the Antarctic raob stations in the same dataset.

b. The NCEP Cray C90 has been having trouble, so operations keeps stealing Bob’s Cray Jedi (16 proc) to run operational weather forecasts for the US.

c. September - December 1972 was completed, but the data got wiped out, so it had to be done again. This was rerun, which completed 1968-1972, but 1968, 69 needs to be rerun because of some bad ship locations, and other reasons.

d. Bob is now doing April 1968 (on March 11, 1997). He will finish 1968-69 about March 25.

e. FGGE (1979) will be rerun because of the problem of no Antarctic raobs, and to use the proper snow cover boundary conditions. It should be done about April 10. But the FGGE (with satellite data) will be run first. After April 10, a FGGE (nosat) will probably also be rerun. This will take about eight days.

f. Consider the original NCEP data from GTS for 1962-72. Jack Woollen just found that the raob tropopause data had not been decoded. This tends to leave out the coldest temperature data in the raob soundings. Some rerun tests will be made to evaluate this. I do not think it will make much difference, because the analysis procedure seems to do OK, even if the sig levels in soundings are not available. However, it is a concern, because the methods of analysis use temperature and winds, but they do not use the rawinsonde heights.

g. When will NCEP need observed data for 1957-1967? They need much of it soon to start getting ready to do these years. Two weeks ago, it appeared that they would start operations on 1957-67 about March 20, 1997. Now, on March 11, it appears that they will start operations on this period about April 20.
Section B: The status of sending several datasets.

1. NCEP original GTS decode (1962-72).
   - This dataset has rawinsondes, pibals, aircraft, cloud winds.
   - NCEP already has 1968-72.
   - This dataset will be ready March 13.

The last of the problems we need to take care of now are nearly resolved. There are still troubles where data for a wrong hour is in some of the 12-hour files. Sometimes there is a mix of good and bad data in such original files. These problems can’t all be resolved, but we can warn people about when the dangerous data/times occur in the record.

2. Set of country raobs (mostly 1950’s on).

These are from various countries. Has rawinsondes, and some countries also sent us pibals. During March 6-11, there have been various runs to compare station libraries. Some library data has been at wrong longitudes, and even the wrong hemisphere. Some stations were not in the master library.

Four raob stations will be excluded because they have hit and miss reporting and under 50 raobs each (looks like they came from unknown errors). The pibal stations were all OK in this regard.

There were 36 raob stations that had much data, which were not in the master library. These are now added to the library.
   - This dataset (for 1957-67) will be ready about March 13.

3. TD52; World pibal data 1920’s-1971. NCAR has done much work on this in the last ten days. Gregg has sorted out the last of the questions of what altitude levels to assign the data to. Most decks are GMT time, but a few are local time, so these will be converted to GMT. Some heights of data are meters and some are in feet. These will all be put into miles. A few speed units are Km/hour.

Bob has half of TD52 converted and will finish about March 12.

Most wind data are all given as height above station ground level (AGL), as has long been the standard practice in the world. Some of these stations (one deck) report as height above ground level (AGL) in the lowest levels and then change to height above MSL.

Then various compares between stations libraries will have to be made; plus compares of data. These will probably take a week. Perhaps ready by March 19.
4. Scherhag raobs (31 stations).

Data comparisons with MIT data indicate that the wind units are usually m/sec. But one stations is knots. This assumes that the wind units in MIT are all m/sec.


This dataset has been worked on heavily for many months at NCAR. Many compares were made with the US control data to help verify the MIT station lists (and to correct problems). At the 12Z time, there are fewer stations and most of them are US control stations. Comparisons with US control were made to identify the MIT stations at 12Z. There was no valid list of stations for the MIT 12Z data. Everyone thought that the 00Z stations list would be the proper list to use at 12Z also (this was not true).

The MIT dataset has already been compared with the NCEP decode data for January and April 1963. The comparisons of the actual reported daily data look good and gives us confidence that the MIT dataset is now in rather good shape.

In the past few days the MIT library has been compared with the master library. On this basis a few stations were added to the master library.

- The MIT dataset will be ready about March 13.

6. US control raob data.

NCAR has been detecting the problems in a number of stations as noted above. The US control raobs usually use WBAN stations numbers. It will take a few more days to get the library cross references between WBAN and WMO numbers correct.

As we make more tests on the actual data, we keep finding more things wrong. This dataset (for 1957-67) will not be ready, at least before March 21.

7. South Africa raobs; no more work done yet. Should not take too long.


9. Other datasets.

a. Big TD54 raobs. Still problems to fix.
b. Line Islands raobs.
c. Early USAF aircraft (recon)
e. Etc.
10. Surface data coverage over Russia and Asia.

We will summarize the coverage of surface data in the TD13 dataset from the USAF.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mid-Russia</th>
<th>Hi Lat Russia</th>
<th>China</th>
<th>India</th>
</tr>
</thead>
<tbody>
<tr>
<td>1946</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>dense</td>
</tr>
<tr>
<td>1948</td>
<td>good</td>
<td>poor</td>
<td>none</td>
<td>dense</td>
</tr>
<tr>
<td>1954</td>
<td>dense</td>
<td>poor</td>
<td>dense</td>
<td>dense</td>
</tr>
<tr>
<td>1956</td>
<td>dense</td>
<td>good</td>
<td>dense</td>
<td>dense</td>
</tr>
<tr>
<td>1963</td>
<td>dense</td>
<td>good</td>
<td>dense</td>
<td>dense</td>
</tr>
<tr>
<td>1967</td>
<td>good</td>
<td>good</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

NCAR has a dataset of USSR surface synoptic data from Russia for 1936-on. We may use this for 1946-1956 when the coverage of stations in northern Russia is not very good.

11. Check the data in component datasets before merges.

Our philosophy has been to make a lot of checks on the data within datsets before any data merges are done. If this is not done, various problems in the data are merged together in a way that will be impossible to understand later on. One powerful method to discover systematic errors in the datasets is to make comparisons of the reported data from the same stations in different datasets. The checks for systematic errors in the data between datasets, are also powerful tools to discover errors and provide confidence that the data are OK. Since there will always be some data differences on some days, (due to minor errors in the data), it is important to set up the tests so that a person does not have to analyze each small problem, but can still readily see the larger systematic problems that need to be fixed.

We note that hydrostatic checks could help detect mis-identified stations (because of wrong elevations). However, hydrostatic checks can not find problems of systematic offsets in date/time. If rawinsonde data has been assigned to the wrong pressure levels (as for some stations in TD54), a hydrostatic check program might well try to fix the data, but it would not make the proper fixes. The big systematic errors have to be discovered and fixed first, if at all possible.

The cleaner versions of the big component datasets will be the most important resource for future world use. If other problems are discovered in the future, these datasets should allow people to understand the problems, and cope with them.
To: Bob Kistler  
From: Roy Jenne  

Status Report on Data for Reanalysis  
(14 April 1997)  

A. The following report is from Bob Dattore  

Monday, April 14, 1997  

1. MIT Raobs (1958-63): ---(almost done)  

   Joey has made changes to the library and I am now in the process of putting  
   the headers on the MIT data. If there are no more conflicts with the station  
   library, the data could be ready as early as this afternoon.  

2. TD52 Pibals (1930's-1971):  

   I am in the process of putting the 1946-1967 reports out in the ftp area. So  
   far, I have put out 1946-1957 (done April 11). The last 10 years (58-67) should  
   be out there by this afternoon. (Rerun will be required to correct a few station  
   IDs. Ready by April 17. --Dennis Joseph).  


   I am still working on correcting bad soundings. This could be completed by  
   tomorrow morning, April 15. Then I will put reanalysis headers on the data  
   and do some track-checking to look for locations that are obviously bad. I  
   would set a ready date as Friday 4/19. \( \text{[that Apr 22]} \)  


   These are in the ftp area and I have told Dennis Joseph about them. They  
   were put there on 4/8.  


   These are in the ftp area and I have told Dennis about them. They were put  
   there on April 2.  


   No new progress. (This should be used for reanalysis. --Jenne).  

   \( \text{gave 57-59 on May 8, and earlier data bits due to too high} \)
7. Line Islands raobs (1967):

   No new progress (this can be delayed some—Jenne).


   No new progress (there is already good coverage for mid 1956-on, in TD54).

B. Selected datasets.

1. TD54 (1946-71)

   TD54 is needed for reanalysis. The coverage would be too thin in both the 1950's and 1960's without it. The things that we must still fix can be done before too long. Will Spangler had to miss four days last week due to a very bad cold.

   Seven stations had a bad stratosphere for several years (raob data placed at wrong levels). We now know how to resolve this problem without losing data. There are other systematic problems that we are still trying to cope with. Present estimate: ready by April 23. This dataset has been difficult but I think that this time estimate is believable.


   The list of bad stations (time shift) is 95% complete, but it needs some checking. The library work for US control is within two days of being done. Then we need to put reanalysis headers on and do last checks. It should be ready for NCEP on April 22.


   The library is done. Nearly everything else is done. Ready April 16.


   This can wait awhile. Not a big job. Too busy on other things.

C. Other datasets

A few datasets are not listed here. See the April 7 text.
Data for Reanalysis

de some problems we are finding, etc.

We will discuss some of the methods being used to remove data errors that would cause serious trouble for the reanalysis. In the second part, we will outline the progress in sending data to NCEP.

Part A: Preparing data for reanalysis.

1. Some of the data problems we are finding.

The data preparation is coming along rather well, but it is slower than we hoped. This is as of March 13.

NCAR is finding:

a. A few stations in US control each with a few years of record, where the data time is off by 12 or more hours (stations such as San Diego, Kwajalein Island, etc.). In addition, in these cases, mandatory levels and sigs were merged that were not for the same time. A mess. So far, the data for these same stations (mandatory levels only) is OK in TD54. So we will toss the data from US control and use it from TD54.

- March 13: We have now found ten stations with this type of problem.

b. A few stations (these seem to be military sites) in US control have constant data during day after day, for several years. It is interesting that the number of levels of the constant data does change from day to day. This data error was the same in US control and in TD54 so the massive data compares did not identify this error. We will eliminate this bad data. These observations are probably lost forever, as we do not know another source. I do not think that this will hurt reanalysis very much because there are often other stations in the vicinity.

c. I told Bob Kistler that we will keep him informed about where we stand. I noted that the work done to fix data for 1957-67 will help a lot with the earlier years 1946-56.

Bob noted that in summer 1996, he had to spend a good deal of time to figure out ways to handle SST and sea ice for the earlier years. The data that he received had many problems to overcome. But his solutions solved the problems for all of the early years. He thinks that we at NCAR are in the same situation now in solving some of the long standing problems of the early datasets. This seems to be true.

copy of this went to Halley
2. Our NCAR methods of finding errors in raobs.

In general we try to concentrate on fixing data problems that will cause a lot of problems for the reanalysis operations at NCEP. For example, if the data are hydrostatically correct, but the date-time or the location is wrong, it will badly affect the analyses, at least part of the time. If stations A is incorrectly called station B, this will also put good data at a wrong location. Also, if rawinsonde data was systematically assigned to the wrong pressure levels (as for some stations in one big dataset, TD54) the NCEP hydrostatic check programs will try to fix it, but they will not be able to do the correct repairs.

In some cases we find that the data are systematically wrong in one dataset and correct in another. Once enough tests have been done to gain this knowledge, we know how to get rid of the wrong data, and use correct data when possible.

3. The identification of stations.

When NCAR obtains data from national archives, the typical practice is that there is a station number and a date/time on each observation. Some countries (including US, NCDC) use their own station numbering methods rather than WMO numbers. Usually the relationship between the station number and location comes on pieces of paper; sometimes in a digital file. The location information seems to have some trouble in most sources, and there is no perfect source. By checking several sources against each other, we are becoming quite confident in the locations we assign to the data.

The data from NCEP have locations on each report. NCAR extracts this from the dataset and runs time series checks on the locations for each station for all years (1962-95). There were some stations that were off location by even 1000 Km, some 10,000 Km, for long periods of time. These have been fixed, so that the reanalysis headers have good station location data.

Part B:

1. Selected big datasets now being sent.

1.1 NCEP data from GTS, Mar 62-Dec 1972.
This has data from the original NCEP decode of GTS data (March 1962-Dec 1972). This was hemispheric. It became global in June 1996. But the coverage was not as good as possible in early years. So we need other data sources too.
- The 1968-72 portion of this data was sent on March 1, 1996.
- Data for bad off-hour times should be thrown out on ingest.
- Please see the daily inventories (on Internet server at NCAR). These show some 12-hour periods that are likely to have questionable data.
- This dataset has the following types of data.
<table>
<thead>
<tr>
<th>Report type</th>
<th>Reports for 1962-72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raobs, dropsondes, ship UA</td>
<td>4,509,454</td>
</tr>
<tr>
<td>Pibals (wind aloft)</td>
<td>2,065,987</td>
</tr>
<tr>
<td>Aircraft and recco obs.</td>
<td>3,642,307</td>
</tr>
<tr>
<td>Satellite Cloud Winds (1967-on, some earlier)</td>
<td>313,075</td>
</tr>
<tr>
<td>Satellite Soundings (Sirs data: May ‘69 - Sep ‘72)</td>
<td>154,690</td>
</tr>
<tr>
<td>Other data</td>
<td></td>
</tr>
</tbody>
</table>

- These counts will change some.
- The period 1968-72 was sent March 20, 1996. This replaces an earlier version sent March 1. It was found at NCAR that a few reports had a bad longitude in the reanalysis headers, so we corrected the file and sent it again.

1.2 The countries dataset (mostly 1950's -1985)
This is a collection of data from various countries, not US and not Canada.
- This actual data was compared with NCEP decode data for January 1965.
  The compares look good. But there were many stations in the countries data that did not find any matching data on the NCEP tapes.
- This data was ready March 13, 1997.

1.3 TD52 Foreign Winds Aloft (pibals) 1922-1971
The reanalysis headers are being put on. NCAR keyed in the station library information. We are using the keyed location data, except that a few stations had to be fixed, based on other comparisons.
- This dataset will be ready about March 17, 1997.

2. Other raob datasets in preparation.

2.1 The MIT raob dataset (May 58-Apr 1963)
Bob Dattore is done with a million checks and fixes. There were many, many problems with station identification. Joey Comeaux is now working on more station library issues. There were 50 MIT raob stations that were not in the master library. These can be quickly put into the library. The are a few stations that may still be at the wrong longitude or in the wrong hemisphere. These will be fixed.
- This will be ready about March 24.

2.2 South Africa raobs 1961-67
This is a critical dataset. It fills a gap. On March 13, it is all converted and through the hydrostatic check. In general, it looks very good.
- The year 1961 for station 68032 failed hydro checks all the time. It looks like the data was assigned to wrong pressure levels for that year. There is no comparison data in any other dataset.
• Probably ready March 20. We may have to toss the one station year. We are contacting S. Africa for more information.

2.3 Scherhag raob dataset (1954-62).
This has 31 raob stations. We had no information about the units of wind speed. Comparisons with MIT look good. They also show that all but one station has units of m/sec. The one station is knots. Bob Dattore will also compare data with the countries dataset.
• Ready March 18.

2.4 China raobs (1954-57).
We have good China coverage from mid-1956 — on. This dataset is not critical for the 1957—67 period.

2.5 Arctic raobs.
a. North Pole Russian Ice Islands raobs
• You have 1968-72
• NCAR is still working on 1954-1967
b. Arctic dropsonde data (Apr 1950-Dec 1961)
• Ptarmigan recon: send by March 26.

3. Three large datasets still giving problems.
a. US control data (~1947-1967)
This dataset has US data, Mexican, Caribbean, Pacific Islands, and some others. It is a necessary dataset to have.
• We have found about 10 stations that have periods of several years (different periods for different stations) when the soundings all have the wrong date/time.

---Wrong date/time---

• On March 12-13, we are doing massive comparisons to make certain that we found all of these cases of bad data.
• We will knock the bad data out of US control. Usually there will be good data for the same station in TD54.
• Some RH data (low values) are set negative. They should be set positive and used in a normal way.
• Probably ready by March 21.

b. TD54, 1946-1971
Will Spangler has done a lot of work on this large dataset. It still has some systematic problems. We have isolated most of them. Work is going on to isolate the rest of them. This set has data that does not exist in other sets.
• This may be ready by March 28.
c. The Northern Hemisphere C-Cards (1949-1966)
   We obtained this raob data last fall. We inventoried it. There has not been
time to process this data. We know that it has some stations that are not in
other datasets. We may have time to include part of the data.

4. Land surface data
   4.2 TD13 world synop (not US), 1920s-1971.
     • We sent 1957-67 on August 12, 1996.
     • We will send 1946-56 by June 1997.
   4.3 US hourly (1948-67), 300 stations. Send by March 25.
   4.4 Russia synop from Russia, 1936-on. Not decided yet.

5. Surface marine data.
   • COADS; 1950-93; you have this.
   • We are doing an update on 1990-95 (Done about March 25, 1997).

6. Aircraft datasets.
   a. In NCEP decode, 1962-72
      — Ready March 13, 1997
   b. Sadler aircraft for tropics, 1960-73
      — Sent to NCEP 14 March 1996. You should have it.
   c. USAF recon data (several routes), 1947-Dec 1959.
      • NCAR still needs to process this.
      • It has 717,019 aircraft reports (volume is 132.65 MB).
      • We hope to send this by March 27, but raobs have a higher priority.
NCEP/NCAR Reanalysis
(Talk with Kistler, May 30)

1. Data from NCAR for 1957-1967.
   NCAR sent it May 23 (Friday); Kistler got it May 28.
   a. S. Africa raobs for 1960 still not available.
   b. Remember the raob time change during April & May 1957
   c. NCEP will probably use June 57 for spinup and start with July 1957.

2. Line Island raobs (in 1967)
   This is still being prepared at NCAR
   a. The wind data are for each 30 sec. Most runs about 90 min (180 levels). One
      has about 315 levels.
   b. Kistler says the NCEP limit on levels is 255.

3. FGGE has been rerun.
   a. Bob made sat and no-sat reruns.
   b. All of the Antarctic raob data did get in.
   c. They will send the new output.

4. Some raob data not used (a problem).
   a. About Fall 1996, they found that during FGGE, very few Antarctic raobs had
      been used. It turned out that after the raob data was in Bufr code, the first
      batch (or two) of raobs for each time was not used (an IO error). The cause
      of the problems were only found recently.
      The FGGE raobs were sorted from South Pole to North Pole, so the data lost
      was Antarctic.
      Note: The input tapes do have the observations
   b. This was really a more general problem (it happened many years)
      (1) A change in a program was made in Aug 1995 that caused this loss of the
          first batches of data each time. So it affected 1968-1978, probably CDAS (late
          years) etc.
      (2) Most years the stations with low numbers were first in the data file, so
          Scandinavia, etc. was hit. NCEP will map the years hit.
      (3) At 6 & 18 Z, there is less data, so a bigger fraction was lost.

5. TOGA COARE

Some TOGA COARE raobs had too many levels to handle well (over 255). Some
data was split into 2 parts at NCEP. The complex QC did not calculate virtual
temperatures. The SSI (model level) picked up temps and thought they were
virtual, but they were real. So air temp too low over TOGA and this gave bigger ocean to air fluxes, etc. (this is how it was found).

NCEP will rerun the four months of TOGA COARE and sent us the data. The problems of bad bogus land temps will also be fixed.

(1) They may make new CD-ROMs for the two years (no decision yet).

6. Limit raob observations to save money.

This plan was stopped. The observations will continue.

7. Texts about datasets.

Bob would like to see some of the texts about the 1957-67 data even if they are not in final form.

8. Corrected daily precip from reanalysis
   Jun 10, 1997; from Wesley Ebisuzaki at NCEP

From: Wesley Ebisuzaki
Date: Tue Jun 10 07:43:41 1997
To: joseph@ncar.ucar.edu
Subject: corrected precip

Dennis,

The corrected daily precipitation is currently not on-line.
Bob K. has a new linux box with anonymous FTP facilities and may make it available at a later date.

Wesley

There is model precip each 6-hours. The correction uses 6-hr inputs and produces a corrected daily precip.
Status of Data for Reanalysis
(Roy Jenne, Oct 21, 1998)

Information is given about several datasets that will be ready to use in a few days.

1. THE B-3 NCEP DATA (MAR 1962 – DEC 1972)

   This is the NCEP archive of data from GTS and other sources in real time. It has raobs, pibals, aircraft, some early cloud winds, etc.

   NCAR obtained the original NCEP archive tapes around 1974-75. There was one stream of tapes for 00Z data and one set of tapes for the other half day. There were 17 cases where the data for one half day was duplicated exactly at the next half day, but labeled with the date/time of the next half day. These cases were identified and removed before NCEP got the data.

   There were about 50 other time periods with more complicated problems. One of these bad cases for a given date/time often had data for the right date/time, but it also had other files of data for a wrong data time. We used time series of rawinsondes (for 6 stations) from other datasets (with proper date/time control) to use to match observed data from the 6 stations with data from the same stations in the B-3 data. This allowed us to determine which were good files and which were bad. This second set of fix-up work has been done during May-Oct 98 (part-time work). It is done now.

   a. A lot of Type 4 reports in the B-3 data turned out to be good raobs and pibals. The data in them matched data in the 6 marker raobs for 54,810 reports that were excluded earlier. Of these, we tossed 9 because they were garbage.

   There were a lot more type 4 reports that we did not recover. Most were too short to be a raob and some were too long.

   b. Internal hour wrong. A bunch of raobs had previously been excluded because they had an internal hour of 24 or 31. Will Spangler found that if he assumed that these really had the same day/hour as the file header, then the data in this B-3 file matched the data for the 6 marker stations. So this data was recovered—a total of 90,675 reports. (6681 reports with hour 24 and 83,994 reports with hour 31.) Note that hour equals 31 was “all bits on.” Maybe it was a missing code.

   c. Note: The report types were not set in the original NCEP tapes during 1962, 63. NCAR wrote programs to look at each report and determine what type it was. This was done before NCEP got the data, and it was done again on the Oct 1998 version.

2. C-CARDS RAOBS

   We did more clean-up on C-Cards in the first 5 months of 1998. This dataset required a lot of cleanup work. By looking at the statistics of failures in hystatic checks, we could identify and fix systematic problems. The units of the data were different for different stations. For blocks of stations, the units were supposed to have changed on certain dates, but often the real dates of change were different. We took care of a lot of these problems before NCEP got the data, but we were able to do a lot more station-by-station cleaning later on. I'm pretty happy with the dataset now.

   In the past few months we have been addressing the remaining station library issues (from lots of cross-comparies between different sources of library data). This work is within 4 days of being done.

   This was sent to ECHOF
3. **MIT RAOBS (5 YEARS, MAY 1958 – APR 1963)**

This has been run against the new (Oct 98) station library. The run gives diagnostics of old vs. new library information. After a few library fixes, it has been rerun.

4. **TD54 RAOBS (~1946-71)**

The data has been in final form for a year. The last round of station library comparisons is 70% done. The TD54 dataset is in 5 different pieces. GFDL did the initial processing on 90% of it (Part A). NCAR did all of the processing on Part B. Each part is divided into data with WBAN station numbers, and ones with WMO numbers.

There is also a Part 5 with stations from India and Pakistan. The reports are all flagged as to type of station number. The reasons for having 5 files instead of 1 are past. There is no reason not to merge them.

5. **THE COUNTRIES SET OF RAOBS (1950’S TO ~MID 80’S)**

This data stream has rawinsondes and pibals that were received directly from about 20 sources, mostly national archives.

We had prepared the data in batches to send to NCEP: 1946-67; 1968-72; 1973-80; 1981-84.

We are still determining how the present files in this dataset compare with what was sent to NCEP. Data was arriving at NCAR at different points in time. Some of it was too late to meet NCEP deadlines, but most of it did get into the 50-year reanalysis from one source or another. Also, for some recent years of data, we only added data for the most important isolated stations rather than sending all data.

6. **NEW WORK ON UPPER AIR STATION LIBRARIES**

We built station libraries for each station from all of the daily UA reports on the NCEP tapes for 1962-72 (the “B-3” tapes), and from NCEP data for 1973-95. The operational tapes had various changes, most good and some not good. By looking at time series of locations and elevations we have gradually been able to remove problems. We calculated elevations from the rawinsonde information, and have used that to improve the elevations when the calculations can be made. In some cases, we see large enough elevation changes that we would expect a location change at the same time. But there may be no information. We have done enough work that we think the remaining errors are constrained to fairly small limits.

Each report carries its original library information (from NCEP, from a list from a country, from the corrected MIT list, from our key entered tables for C-Cards, etc.) That information stays with the report, plus the best data we have from our latest library efforts. We have several main sources of library data and have done many cross comparisons. All of this has helped to make decisions about what to believe. When we insert the best library information into the reports to be sent, diagnostics are made of the comparisons with original data. This allows us to keep tight control of any new key entry problems.

It is a relief that the library is finally at the stage of development that exists now.

The final raob library is now being inserted into B-3 raobs, C-Cards, MIT, TD54, Countries, TD90, etc.

7. **SUMMARY OF DATA THAT NCAR WILL SOON SEND**

About Oct 26-30, NCAR will be sending several batches of data:

- b. C-CARDS raobs (1949 – 65), 1.414m obs
- c. MIT raobs (5 years, 1958 – 1963), 1.031m obs
d. TD54 raobs (~1946 – 71), 5.065m obs
e. Countries raobs (1946 – 84), 2.17m obs
f. US control raobs (1946 to ~1980) Send somewhat later
g. TD90 raobs (1943 – 62; about 22 stations per day during 1950-59) Has 57 stations.

Note: Data counts in NCEP B-3 set (10.8 years, 1962-72):

- Rawinsondes: 4.45m
- Pibals: 2.16m
- Aircraft: 3.65m
- Satellite wind: 0.311m
- Satellite soundings: 0.152m

Note: Counts of data in the countries dataset:
a. 1946-47 about 2800 obs/year
b. 20k in 1948, increase to 61k in 1967
c. 101k in 1968-70
d. 89k in 1971, decrease to 34k in 1984

8. SOME OTHER DATASETS

Some datasets such as China raobs and French raobs are still in small side files rather than being sorted into the countries dataset. We will tell you about these later.

9. SELECTED CREDITS.

Bob Dattore did a lot of cleanup work on several main datasets (C-Cards, MIT, TD90, several countries, etc.) Will Spangler has done the cleanup on NCEP B-3, much on TD54, much on US control, etc. Joey Comeaux has done most of the overall library work: many comparisons to find the conflicts and problems in data sources, checks on continuity, etc. Dennis Joseph worked on North Pole raobs, various calculations, and has helped to coordinate many projects.

10. COADS DATA

This will be sent about the end of November 98. More information will be sent about this.

Note: We will be sending about two more messages.

Best wishes,
Roy Jenne
Comments About Reanalysis Data  
(Roy Jenne, 9 Nov 1998)

This has selected information about several datasets. It also has a discussion of station library numbers, especially for upper air data. One section notes that to bring duplicate raob and pibal data together, one must use the location information.

1. NCEP B-3 Data (1962-72)

   The original data were in two separate tape sequences, one for 00Z and one for 12Z. The 00Z tape sequence later included data for 06Z and the 12Z tapes included 18Z data. Each time period had files for all of the different types of data: raobs, aircraft, etc. In the processing done during May-Oct 1998, Will Spangler put the 00Z and 12Z data streams together into their proper date/time order.

   In Oct 1998, the most recent station library information has been added into the header block of each rawinsonde and pibal report. The data that you will get is in the following data sequences for the period. Each data sequence has separate files for each year.

   a. Rawinsonde and pibal reports  
   b. Aircraft  
   c. Satellite cloud winds  
   d. Sirs (satellite soundings in general)  
   e. Miscellaneous  
   f. Other  

   Note: The last two are tiny files of some type of data that we could not recognize. These files probably are not useful to you.

   SMALL NUMBERS OF RANDOM WRONG RAOB AND PIBAL REPORTS.  
   Because of errors in WMO station numbers due to garbling, etc., the B-3 tapes had a few reports for stations that NCEP thought might be a valid station, but the data was really valid for some different station number. We took steps to get rid of this random trash (good reports at a wrong location). If the count of raobs plus pibals for a station was under 3 per month, the data was tossed. If the data kept was under 26 reports in the whole period of record (1962-72) then we tossed all of the reports for that station.

2. The NCEP GTS tapes for 1973-1998

   The NCEP GTS tapes for 1973-1998 use C7A, C7B, etc., for the ships that still existed. We understand that C7 was considered a proper WMO designation, whereas 4Y was set up for international aviation. Other groups have often defined a set of numbers for the ships such as 99001 (4YA), 99002, etc. But these schemes are not all consistent.

   We also had some trouble with the location of permanent ships. In some data sources, the location of some ships was very wrong, or it was not set at all, so that one would get the location of 0,0 if this was not fixed. But a ship actually moves a little within a one degree box while it is on location. We ended up just setting the location of the ships to the standard location. Each report on the reanalysis tapes has an internal location and an external location in the header section of the report. It would be somewhat better, if one checked the internal location versus the external. If the internal one was within one degree or 0.5 degree (?) then the internal location could be accepted and this might often restore the tenths of degree (almost) precision in the location of the ship. NCAR has not done this in the data. Later, we will note that some ships have used a 4Y designation while steaming to the standard location. In this case, we would throw out good data if we demanded that a ship location has to be close to the permanent position.
SOME COMMENTS

3. C-Cards (1949-65), and add 1966

This data was from the big Northern Hemisphere Data Tabulations project of NCDC. We did a lot more cleanup work in the first 5 months of 1998. When NCEP got 1949-59 by March 1998, most of the cleanup for those years had been done. But they should get a new version, both because of more special case cleanup, and the new libraries that were ready in Oct 1998.

The original project prepared plotted charts at one time of day, and prepared digital data at the other time of day. The digital data is nearly all for 03Z or 00Z (not for the other half day).

Now (Nov 98) the data for 1949-65 are all ready, and all with new library information.

Note: The C-Cards document says there is data through Oct 1966, and an inventory in the documents says that there is one station that goes into 1967. There are two stations (with 335 reports) in 1966 (10 months for 03953 and 1 month for 06610). These will be available by FTP. There is nothing in 1967.

4. TD90 has Data for 1943-52. No data for 1947

This dataset does not have as many stations as sets like C-Cards, TD54, MIT, and Countries.

5. MIT (1958-63, 5 years)

There are separate files for 00Z and 12Z. There is a lot more data for 00Z (about 3X more).

The original MIT station numbers were just a sequential list of numbers, about 1 to 600. This original number is now gone from the reports. They are all WMO numbers. It took some time to get this right, because one number was missing from the list we got and this had the effect of changing the WMO numbers on half of the stations. Anyway, you do not have to worry about these typical troubles any more.

6. TD54

This had 5 files of data as we have described. In Oct 98, Bob Dattore combined 2 pairs of files. Now there are 3 files. There is one file for WMO station numbers (Parts A and B together) and one for WBAN (Parts A and B) and a miscellaneous part. The station numbers in the miscellaneous deck looked like they may have been WMO numbers, but WMO would have been inconsistent with the locations given in the printed book. We called the type of number integer-other. The stations have the locations given in the book.

7. The Countries Tapes

NCAR put the data received directly from countries into this dataset. The reports have source codes so that you can still tell where they came from. A few similar datasets were processed later (early China raobs, French national, etc.) These are now still separate small datasets.

The Countries tapes also received some “filler raobs” that we put in from various sources that were not in the other main upper air sources for reanalysis.

8. Volume of Data in Three Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Volume</th>
<th>Size</th>
<th>Largest Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-Cards (1949-65)</td>
<td>148 MB</td>
<td>27 MB</td>
<td></td>
</tr>
<tr>
<td>MIT (1958-63)</td>
<td>337 MB</td>
<td>53 MB</td>
<td></td>
</tr>
<tr>
<td>TD90</td>
<td>17 MB</td>
<td>2 MB</td>
<td></td>
</tr>
</tbody>
</table>
THE STATION NUMBERS FOR REANALYSIS OBSERVATIONS

9. Comments About Station Numbers (raobs and pibals).

It would be nice if there were one consistent station numbering scheme for rawinsonde and pibal (wind aloft) stations but that has not been possible. Following are some comments about station numbers.

Both US and Canadian stations had block 72 WMO numbers for many years. Then on 1 July 1977 Canada all changed to block 71. We do not know if the numbers actually all changed on this date, or whether it took time for some stations to do the new thing. In the NCEP decode data for 1962-1977, we did set the Canadian stations to block 71 in the report headers. In other datasets, the block numbers for Canada were not changed in early years. We have kept the reported numbers in the record. Canada may be a clean enough change that one could just backdate all of those numbers in the other datasets too, if desired.

Alaska has always had block 70 WMO numbers.

Germany used blocks 9 (East Germany) and 10 for West Germany. Then they went to one block number with reunification after 1990. During April 1992, the NCEP decode starts having block 10 for all German stations. We have also seen this pattern in the surface data. We have not checked on what numbers were used when Germany was one country in the 1940s.

The station numbers for permanent ships have been somewhat messy. The NCEP GTS decode tapes for 1962-72 use the 4YA, 4YB, etc., call letters for ships.

The US Air Force has used a 6-digit station numbering scheme for many years. The first 5 digits are the normal WMO numbers and the last digit allows one to handle such things as dual stations on the local scale. NCAR has a sample of their station library tapes over a period of many years. The surface synoptic data for 1967-1980 (we sent 1967-1976) has this kind of station numbers. The original source of that data was the decode of GTS by the USAF at the Global Weather Center near Omaha.

There was a radical change of surface station numbers in Scandinavia (maybe only Sweden) that started 1 July 1977—like Canada. Numbers were switched between stations, so that a lot of the same WMO numbers existed, but they were for different stations. We kept the original numbers (first old, then new), but the location given for the station will reflect the change. For a given actual station, the station number will change, but the location stayed the same.

The Swedish upper air stations changed numbers at the same time (01 July 1977), but the numbers assigned were all new unique numbers. We did not change either the new numbers or the old numbers. The locations will be correct.

In TD54, there is Part A that has WMO and WBAN numbers for the rawinsondes, there is a similar Part B. Then there is the miscellaneous part with stations for India and Pakistan. We first thought that these were WMO numbers, but that is not possible. We accepted the numbers as is, and used the location information printed in the books.

10. What are WBAN numbers?

WBAN numbers stand for Weather Bureau, Army, Navy. The joint scheme was set up in the 1940s or 1950 (I think) in the US. The idea was to have a unique number for a station that had a major move, even though the WMO number might remain the same. Also, if two local stations had an overlap period in time, it gave a way to keep track of both of them. For one WMO number there may be 2 or 3 WBAN numbers over a period
of time. The result is that a number of datasets include stations with WBAN numbers. It is important to know whether a number is WMO or WBAN because the two schemes have some numbers in common.

All of the US control data for early years has WBAN numbers. This includes US stations, Canadian stations, Pacific Island, etc. We still get updates for US stations each year from Asheville. The data has WBAN numbers. The annual update has rawinsonde data for the Caribbean, Mexico, the west coast of South America, and Pacific Islands, all with WBAN numbers.

At times, we thought of converting the eternal station numbers to a common scheme (perhaps WMO), but it was never possible to get a 1 to 1 mapping for all stations. We do have several conversion routines that can convert most stations. In any case, it is important to keep original numbers so that any problems can be traced back. One is bound to keep receiving data in a variety of national and other numbering methods. There will be some changes in time and the older history will keep coming up in some datasets.

11. Numbers for Permanent Ships

All of our so-called time series sources like data from various countries have the permanent ships with a number scheme of 99001 (Ship A), 99002 (Ship B), etc. So Ship K from France, and the South African ship will have these kinds of numbers set by us. The location will be at the reference location. I do not think that any of these sources provided a more precise hour-by-hour location.

Consider the ID of the ships on the NCEP GTS tapes for 1962-1998. These still have call letters (either 4YA or 7CA as noted). The ship location given on the ship report from GTS is just carried along. We did not check it for accuracy. I would guess that NCEP did check it for accuracy. A ship is supposed to call itself a permanent ship (4YA, etc.) only when they are near the standard permanent location. However, we all think that we have seen cases where a ship may be called 4YP but it is clearly steaming to or from port. This means that we would throw out some good data if we demanded that a permanent ship observation always be close to the standard location.

Many of the permanent ships were also in early datasets that we got from Asheville. Here the different ships had numbers from 1 to the 90s. Ship T was #91. Ship N had 4 numbers at different periods of time: 15 (10/49-12/50), 20 (12/50-12/53), 24 (12/53-12/71), and 14 (1972-74). In the present data, all of these are called 14 (N is the 14th letter), and the header has 99014. All of our evidence indicates that Ship N stayed at the same standard location (30N, 140W) for the whole period (1949-74). The proof is quite good that Ship N was always at one location, but some sources disagree. We did make spot checks of data plotted on maps. That helped to confirm the standard location. We note that the surface data for Ship N, all says it is at the same location (30N, 140W).

On a dataset that we got directly from Canada, all of the continental stations plus Ship P came in with 7-digit Canadian numbers. In the NCAR processing, WMO numbers were assigned to most of the stations, but a WBAN number was assigned to Ship P. The WBAN list of numbers does include Ship P. WMO never assigned the usual WMO numbers to permanent ships.

12. Summary of Station Numbers in Several Upper Air Datasets

a. C-Cards: All WMO numbers
b. TD-54: A mix. See the comments.
c. Countries tape: Now it is all WMO; it used to have a few WBAN that have been changed or else moved to another dataset. This tape also has permanent ship data for 1946-80.
d. TD-90: All WMO
e. France and China raobs: All WMO
f. Scherhag: WMO
g. NCEP GTS (62-72): WMO, but see comments about ships
h. NCEP GTS (73-98): WMO, but see ship comments
i. US Control: All WBAN
j. TD52: World UA winds; almost all WMO
   Some IDs are under 1000; some non-standard India numbers for prior to about 1948. We added a flag.
k. TD53: US control, UA winds; WBAN

13. Using the Upper Data with Changes in Station Numbers.

   The numbers of some stations change with time. Also, in different datasets there may be duplicate data for a
   station, one with a WMO number and the other with a WBAN number. The locations for the two duplicates
   will be very close, but often not identical. At present, it is the location information that allows one to bring
   the duplicates all together. This is what NCEP used.

   The TD13 surface synoptic dataset for early years has a mix of WMO numbers and local numbers.

14. The People at NCEP who Handled the Station IDs.

   The key people are Jack Woollen, Bill Collins, and Bob Kistler. They handled the station numbers and
   brought identical (or near identical) reports together.

NOTE: More information about the libraries will be sent a little later.

- Roy Jenne