Talk for UCAR University Relations Committee, July 1999

- Show how much data is being delivered
- Outline different types of data we have
- State 5 possible data projects done with universities
- There are 14 pages here

Roy Jenne
A Talk in July 1999
(now June 2002)
The Big Projects We Have Now

For Data Support at NCAR

1. Prepare observations for reanalysis (1946-98)
   a. NCEP/NCAR project (1948-97) – 50 years
   b. NCEP-DOE project (1979-97) – 19 years
   c. New ECMWF project (1957-1999) – 43 years

2. Handle the archives from reanalysis
   a. Now about 3.5 Tbytes

3. New Mesoscale Model Archives (1995-on)
   a. Three models for North America
   b. Lots of complexity
   c. About 152 Gbytes per year

4. Help US Climate Assessment Studies
   a. Several climate models; a few runs each
   b. Calculate climate change statistics
   c. Interface with application models

5. Add more data of other types
   a. Rivers, radiation, other

Roy Jenne
Jul 1999
Agenda

Meeting date: July 14, 1999
Meeting time: 2:00-3:30 p.m.
Location: Mesa Lab, Chapman Room
Attendees: Kelvin Droegemeier
           Pete Peterson
           Ginger Caldwell
           and others

Order of Business:

2:00   Ethan Alpert, Group Head, Graphics & Data Analysis Group

2:30   Marla Meehl, Manager, Network Engineering & Technology

3:00   Roy Jenne, Manager, Data Support Section

3:20   Don Middleton, Group Head, Visualization Group

3:30   Adjourn
Subject: Re: University Collaborations (fwd)
Date: Wed, 14 Jul 1999 10:57:51 -0600 (MDT)
From: Ginger Caldwell <cal@scd.ucar.edu>
To: joseph@ucar.edu
CC: cal@ucar.edu (Ginger Caldwell)

Dennis,

July 16, 1999

> > > Wednesday, July 16, 2:00 meeting with Kelvin Droegemeier.
> > > Optional!
> > >
> > > Kelvin Droegemeier, the University Relations Committee’s liaison to the
> > > SCD and also a member of the SCD Advisory Panel, is interested in an
> > > informal meeting (no formal presentations) with SCD staff about projects,
> > > especially how universities are working with the SCD -- not only what
> > > the SCD is providing. I have reserved the Chapman room on Wednesday
> > > from 2-3:30 for this meeting.
> > >
> > > Each section is invited to have a representative (but this is not
> > > required!). In addition, staff who are working on a collaboration
> > > with a university are especially invited.
> > >
> > > Kelvin’s goal is to see if universities are doing their part to help
> > > maintain a strong SCD and to see what "cool things are going on" in SCD.
> > > Kelvin was also adamant that he didn’t want to take a lot of people’s
> > > time with this - so feel free to pass.
> > >
> > > Please let me know if you or someone in your group/section will be
> > > attending and if you have a university collaboration ongoing. Attending
> > > for just part of the meeting is also an option.
> > >
> > > Let me know if you have questions.
> > >
> > > Ginger

At 2:50 - 3:10

in Chapman
Possible Data Projects With Universities

1. Make GOES 3-hour data more readily available.
   - Work with University of Wisconsin.
   - Use some loss-less compression on CD-ROMs.
   - Let others add more software.

2. Make North America hourly observations readily available.
   - For 20 to 50 years of record.
   - Put onto CD-ROMs in synop sort.
   - Prepare basic access software.
   - Let other groups help with more software.

   - With the diurnal cycle included.
   - Perhaps a version with diurnal cycle removed.
   - Resolution about 50 to 100 Km.

4. The gathering of observations.
   - If universities help gather selected observations,
   - NCAR can help archive and merge them.

5. Prepare 12-hour atmospheric grid maps for Northern Hemisphere.
   - 1940s – recent
   - This has been a project with University of Washington.
   - The CD-ROMs are also popular in NWS forecast offices.

Roy Jenne
July 1999
What Types of Data are in DSS Archives?
(By Roy Jenne and Dennis Joseph, NCAR, Oct 1998)

Our Data Support Section at NCAR started work in 1965 and has been working on various large projects, building the data archives and helping users ever since. The Data Support Section (DSS) maintains a large, organized archive of computer-accessible research data that is made available to scientists around the world. The archive represents an irreplaceable store of observed data and analyses and is used for major national and international atmospheric and oceanic research projects.

There are now over 500 distinct datasets in the archive, ranging in size from less than 1MB to over 1TB. The total volume of data in the DSS archive was 2.4 terabytes (TB) in August 1990 and 10 TB in October 1998. We have been adding a lot of reanalysis data and other analyses.

A broad summary of our data holdings follows:

<table>
<thead>
<tr>
<th>Count of Main Datasets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Output data from NCEP/NCAR reanalysis, 50 years (4X/day)</strong></td>
<td>21</td>
</tr>
<tr>
<td>a. Volume about 3.5 TB for NCEP plus ECMWF</td>
<td></td>
</tr>
<tr>
<td><strong>2. Separate CD-ROMs from reanalysis</strong></td>
<td>27</td>
</tr>
<tr>
<td>a. Volume on 27 unique CDs is 18 GB</td>
<td></td>
</tr>
<tr>
<td><strong>3. Observations for reanalysis (in 1997, surface and upper air data)</strong></td>
<td>38</td>
</tr>
<tr>
<td>a. About 180 GB, not level 1 satellite data</td>
<td></td>
</tr>
<tr>
<td><strong>4. Mesoscale model data (North America)</strong></td>
<td>16</td>
</tr>
<tr>
<td><strong>5. Datasets of surface observations and related data</strong></td>
<td>~60</td>
</tr>
<tr>
<td><strong>6. Datasets about the earth’s surface</strong></td>
<td>~20</td>
</tr>
<tr>
<td><strong>7. COADS world ship and buoy data</strong></td>
<td>6</td>
</tr>
<tr>
<td><strong>8. Other datasets (not COADS) for ocean work</strong></td>
<td>~70</td>
</tr>
<tr>
<td><strong>9. Various analysis grids</strong></td>
<td>~20</td>
</tr>
<tr>
<td><strong>10. Main operational analyses from NCEP (was NMC)</strong></td>
<td>~15</td>
</tr>
<tr>
<td><strong>11. Main operational analyses from ECMWF</strong></td>
<td>~7</td>
</tr>
<tr>
<td><strong>12. Climate model data for assessment studies</strong></td>
<td>25</td>
</tr>
<tr>
<td><strong>13. Climate trends datasets</strong></td>
<td>~7</td>
</tr>
<tr>
<td><strong>14. Climatology and circulation statistics</strong></td>
<td>~17</td>
</tr>
<tr>
<td><strong>15. Cloud data</strong></td>
<td>~15</td>
</tr>
<tr>
<td><strong>16. Stratospheric datasets, mostly gridded</strong></td>
<td>~10</td>
</tr>
<tr>
<td><strong>17. Datasets for the very high atmosphere (70-1000 Km)</strong></td>
<td>57</td>
</tr>
<tr>
<td><strong>18. Main radiance data from satellites</strong></td>
<td>~25</td>
</tr>
<tr>
<td><strong>19. Data for the FGGE year (1979)</strong></td>
<td>8 datasets</td>
</tr>
</tbody>
</table>

Comment: The above list has about 464 datasets. Our actual list of datasets has over 500 items, but sometimes several logically different files of data are held within one dataset folder.
GCIP Model Data at NCAR

Roy Jenne
Feb 1999
Observations Over North America

1. Hourly observations with pressure, temp, wind, etc.
   a. North America
      (from TDL, NCEP)
      Dec 1975 1000 stations
      1999
   b. USA
      (from NCDC, Asheville)
      1948 300 stations

2. US and global data from USAF (surface)
   1973
   This data is coming for 1973 – 80.

3. US Co-op stations
   1895 8000 stations
   These have daily precip, temp, snow

Roy Jenne
Feb 1999
### Some Main Reanalysis Projects

<table>
<thead>
<tr>
<th>Project</th>
<th>No. of Years</th>
<th>Span</th>
<th>Finished</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCEP/NCAR*</td>
<td>50</td>
<td>1948 – 1997</td>
<td>Jul 1998 (took 4.1 yrs)</td>
</tr>
<tr>
<td>ECMWF</td>
<td>15</td>
<td>1979 – 1993</td>
<td>Oct 1996 (took 2.3 yrs)</td>
</tr>
<tr>
<td>NCEP</td>
<td>19</td>
<td>1979 – 1997</td>
<td>14 yrs done (at 06/99)</td>
</tr>
<tr>
<td>ECMWF (ERA-40)</td>
<td>40+</td>
<td>1957 – 2001</td>
<td>Start: Dec 99?</td>
</tr>
</tbody>
</table>

* See *Bull. AMS*, March 1996

Roy Jenne  
8 Jul 1999
Figure 1: Data from Data Support archives that are read into user programs (not DSS), which are run on main computers at NCAR. The Gbytes read during each year are shown. A large portion of the use is by the universities. Most of the “other users” are NCAR users.
Figure 3: The total volume of data sent by the Data Support Section of SCD/NCAR. Data sent on tape or CD-ROM is probably used 2 or 3 times on average, so it is actually a greater part of total data use than is indicated here.
CD-ROMs from NCEP/NCAR Reanalysis

- Have been helpful in university education.
- 4500 copies have been sent
  - Cost about $10 each

1. SALES OF REANALYSIS ANNUAL CD-ROMs

See the Dec 1998 Monthly Report for more listings on these sales.

<table>
<thead>
<tr>
<th>Date</th>
<th>Unique CD</th>
<th>Orders</th>
<th>CD-ROMs</th>
<th>Total $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 21, 1997</td>
<td>8</td>
<td>14</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Nov 24, 1997</td>
<td>15</td>
<td>136</td>
<td>1,041</td>
<td>$11,31</td>
</tr>
<tr>
<td>Feb 28, 1998</td>
<td>18</td>
<td>(1979 - 96)</td>
<td>185</td>
<td>1,563</td>
</tr>
<tr>
<td>Jan 4, 1999</td>
<td>28</td>
<td>(1970 - 97)</td>
<td>310</td>
<td>3,278</td>
</tr>
<tr>
<td>May 10, 1999</td>
<td>31</td>
<td>(1967 - 97)</td>
<td>352</td>
<td>4,020</td>
</tr>
<tr>
<td>June 30, 1999</td>
<td>38</td>
<td>(1961 - 98)</td>
<td>380</td>
<td>4,491</td>
</tr>
</tbody>
</table>


2. ONE CD WITH MONTHLY DATA FOR 50-YEAR REANALYSIS

Glen White called from NCEP. About 13,000 copies will be needed for the Bull. AMS. We do not know who will pay. Now, only 300 to 500 copies will be obtained for the 2nd Reanalysis Conference near ECMWF in Aug 1999.

a. Get 500 copies (plastic bags) for $1250.
b. Get 300 copies (plastic boxes) for $1500.
c. Do option a because it’s cheaper. Note: The master costs $600 of this price.

Roy Jenne
July 1999
Precipitation Data for USA

1. Hourly precip data (USA)

- 2500 stations (1948-1996) — NCDC
- Hourly US Precip grids by CPC (1963-on)
- NCEP:
  a. Hourly real time precip, stn archives, started Jan 1996
  b. Hourly US precip grid, stations only, start Jan 1996
  c. Hourly US precip, radar only, started Jan 1996
  d. Hourly US precip, radar plus stations, start Apr 1996

2. Daily precip data (USA)

- 7500 co-op stations (1900-on) — NCDC
- NCEP
  a. Daily precip obs (real time) started Apr 1995
  b. Daily US grids started Apr 1995

Roy Jenne
Dec 1997
SNOW COURSES

Snow courses consist of a number of sample points (usually about 10) along a transect at a permanently marked location. At each point, snow depth and water equivalent observations are made. Densities are also computed to verify each sample’s accuracy. After all the sample points are measured, the depth and water equivalent values are averaged.

The typical snow course is measured monthly, just prior to the first day of the month. Most snow courses are measured February 1 through May 1. However as the seasonal snowpack varies across the Western US, variations to that schedule occur.

While some of the earliest snow courses were established between 1905 and the 1920's, there were less than 100 of these in existence. The greatest number of sites occurs in the mid 1980's when approximately 1600 snow courses existed and nearly 6,500 measurements were taken annually. Since then, with the improved data collection methods, the number of snow