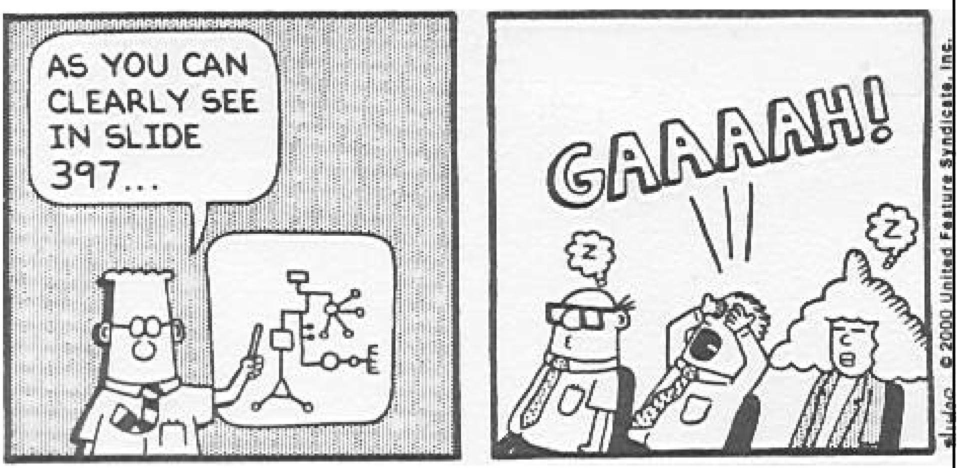
The North American Regional Reanalysis: Results Obtained and Lessons Learned

Fedor Mesinger[1], Geoff DiMego[2], Eugenia Kalnay[3], Perry C. Shafran[4], Wesley Ebisuzaki[5], Dusan Jovic[4], Jack Woollen[4], Kenneth Mitchell[2], Eric Rogers[2], Ernesto H. Berbery[3], Michael B. Ek[1], Yun Fan[6], Robert Grumbine[2], Wayne Higgins[5], Hong Li[3], Ying Lin[2], Geoff Manikin[2], David Parrish[2], and Wei Shi[6]

[1]NCEP/EMC and UCAR, [2]NCEP/EMC, [3]Univ. of MD, [4]NCEP/EMC and SAIC/GSO, [5]NCEP/CPC, [6]NCEP/CPC and RSIS

Potential European Regional Reanalysis Project Workshop, ECMWF, 21-22 November 2005



Hope to avoid achieving this result

In 1997, during the final stages of the production of NCEP-NCAR Global Reanalysis ("GR"), exploration of a regional reanalysis project was suggested,

"particularly if the RDAS
[Regional Data Assimilation System]
is significantly better than the global reanalysis
at capturing the regional hydrological cycle, the diurnal cycle and other important features of weather and climate variability."

(Report of the NCEP-NCAR GR Panel, Randy Dole, Chair)

Goal, Motivation

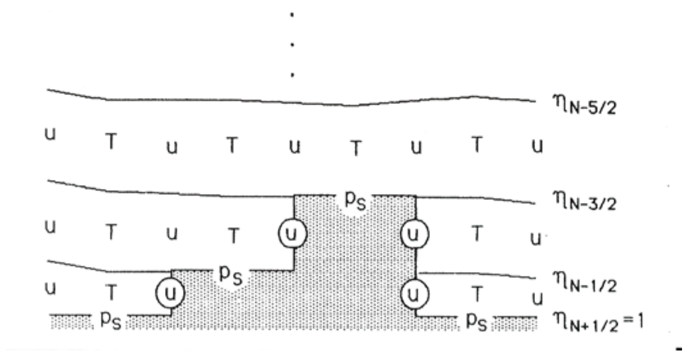
- Create long-term set of consistent climate data on a regional scale for the North American domain
- Use of the GR to drive the Regional Reanalysis (RR) system
- Superior to NCEP/NCAR Global Reanalysis due to:
  - Use of a regional model (the Eta Model);
  - Advances in modeling and data assimilation since 1995, especially:
    - Precipitation assimilation
    - Direct assimilation of radiances
    - Land-surface model updates

NARR Support

- **NOAA Office of Global Programs (OGP)**
  - GAPP: GEWEX America Prediction Project:

  Total actual funding: ~$4M/ 6 years; perhaps ~20 person/years
- **NCEP (EMC and CPC)**
  - Computational resources / IT Staff
  - Staff scientific and technical expertise
- **Science Steering Committee**
  - J. Roads, Chair
- **NCDC, UCAR**
  - public access to NARR database (archiving)

Some of the points re the system used:



A two-dimensional illustration of the eta coordinate step-topography

With model numerics (Arakawa style) conservation enforcing and flux-like, the vertical sides of model grid-boxes being about the same results in the model being very nearly finite-volume

Why should finite-volume be a good idea?

We do "physics" on box-averages,
and for individual model grid boxes,
and thus work against smoothness of model fields

Traditional finite-differencing/ spectral methods
assume smoothness

Eta: a mature model, complex physics, e.g.:

Eta / Noah LAND-SURFACE MODEL UPGRADES: 24 July 2001

- **assimilation of hourly precipitation**
  - -- hourly 4-km radar/gage analysis (Stage IV)
- **cold season processes**(*Koren et al 1999*)
  - -- patchy snow cover
  - -- frozen soil (new state variable)
  - -- snow density (new state variable)
- **bare soil evaporation refinements**
  - -- parameterize upper sfc crust cap on evap
- **soil heat flux**
  - -- new soil thermal conductivity (*Peters-Lidard et al 1998*)
  - -- under snowpack (*Lunardini, 1981*)
  - -- vegetation reduction of thermal cond. (*Peters-Lidard et al 1997*)
- **surface characterization**
  - -- maximum snow albedo database (*Robinson & Kukla 1985*)
  - -- dynamic thermal roughness length refinements
- **vegetation**
  - -- deeper rooting depth in forests
  - -- canopy resistance refinements

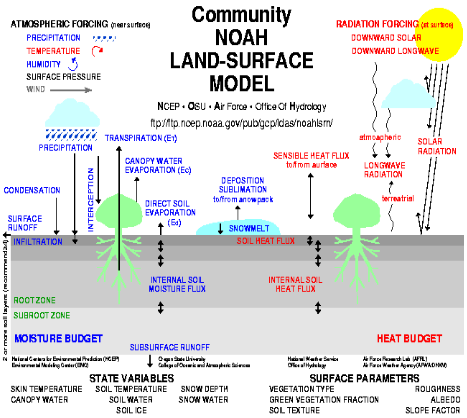


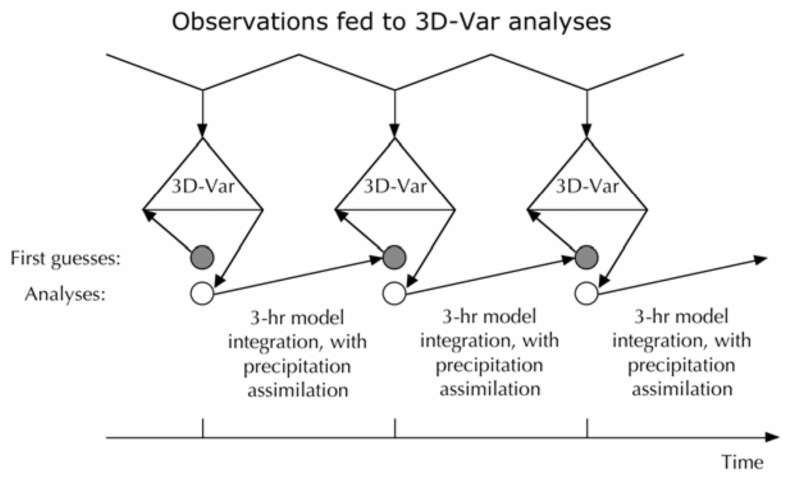
Figure 1. Schematic of the NCEP community NOAH land-surface model of multi-layer soil, vegetation, and snowpack.

NOAH LSM tested in various land-model intercomparison projects, e.g., GSWP, PILPS 2a, 2c, 2d, 2e, Rhone, and (near-future) DMIP.

System Design

- Fully cycled 3 hr EDAS (Eta Data Assimilation System)
- Lateral boundary conditions supplied by Global Reanalysis 2
- Resolution: 32 km, 45 layers
- RR time period: 1979-2003 (continued later in near-real time, as in CDAS: R-CDAS)
- Free forecasts out to 72 hr every 2.5 days, using GR2 forecast boundary conditions ("reforecasts")

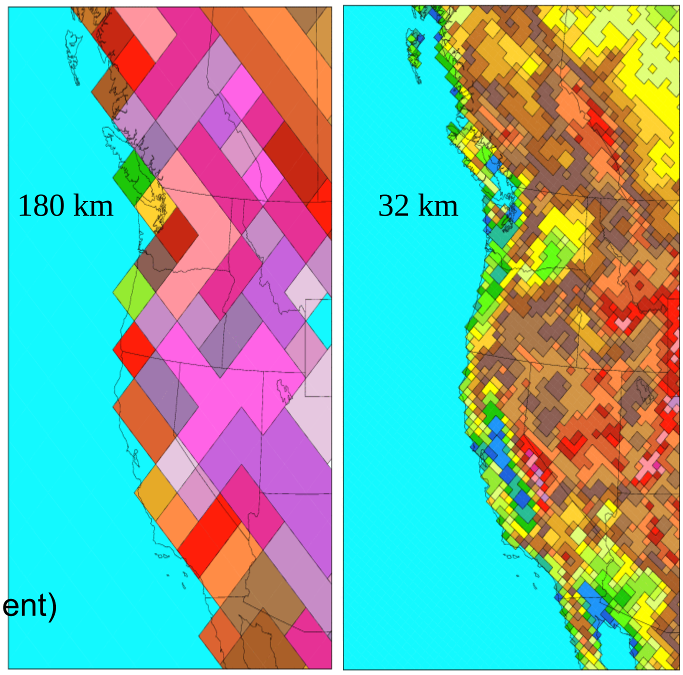
The NARR system:



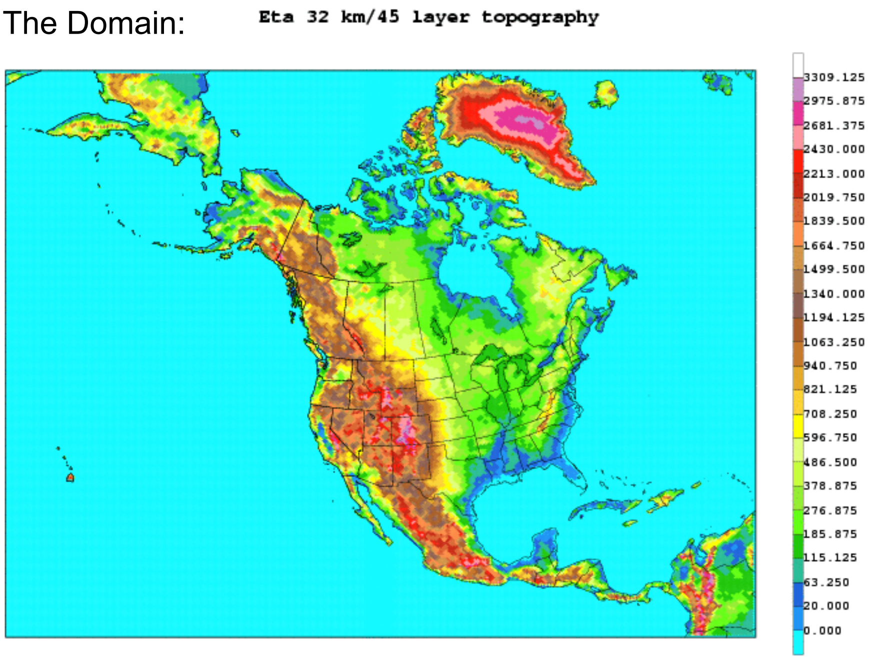
Challenge:

In designing a data assimilation + model system, not that much attention tends to be given to the results at 3 h fcst time

Resolution:



The Domain:



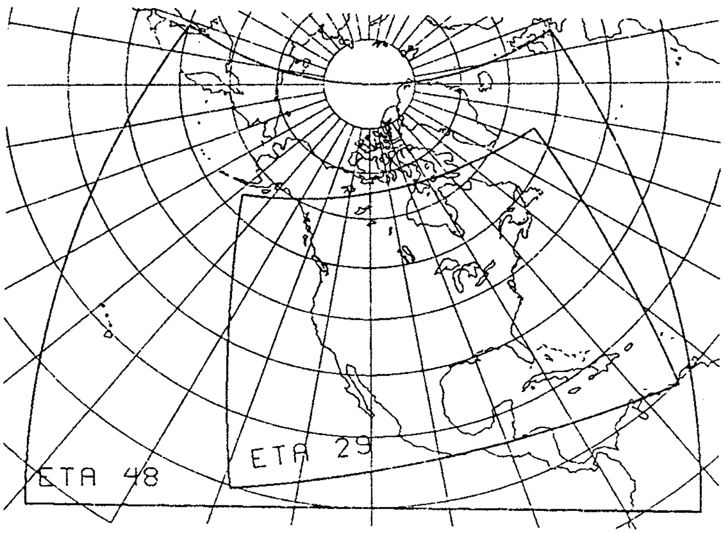
Why is the domain this large ?

The "Early" vs the "Meso" Eta (an inadvertent experiment)

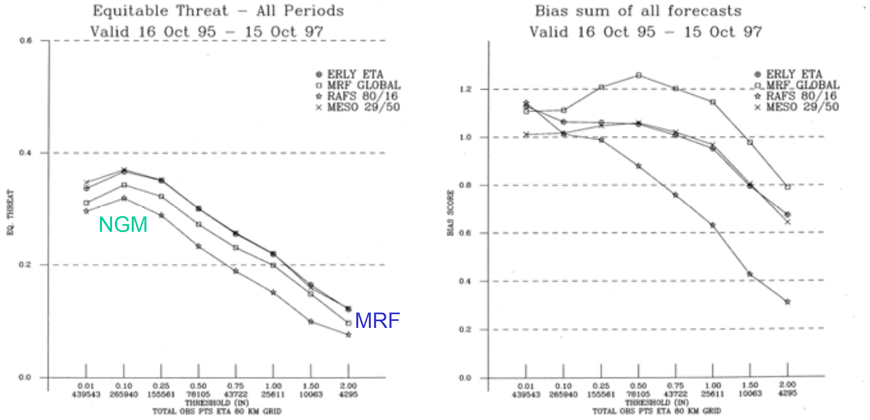
"Early": 48 km, 12 h old Avn LBCs,

"Meso": 29 km, current Avn LBCs;

Domains:



2 years of scores:



Scores of the the “Early” and the “Meso” Eta about the same!

The benefit of the large domain compensates the combined benefit of more accurate LBCs and higher resolution !!

Data Used in Global Reanalysis and Regional Reanalysis

Dataset	Observed variable	Source
Rawinsondes	Temperature, wind, moisture	NCEP/DOE Global Reanalysis (GR2)
Dropsondes	Same as above	GR2
Pibals	Wind	GR2
Aircraft	Temperature and wind	GR2
Surface	Pressure	GR2
Geostationary satellites	Cloud drift wind	GR2

Data Added or Improved Upon for Regional Reanalysis

Dataset	Details	Source
Precipitation, disaggregated into hours	CONUS (with PRISM), Mexico, Canada, CMAP over oceans (<42.5°N)	NCEP/CPC, Canada, Mexico
TOVS-1B radiances	Temperature, precipitable water over ocean	NESDIS
NCEP Surface	Wind, moisture	GR2
MDL Surface	Pressure, wind, moisture	NCAR
COADS	Ship and buoy data	NCEP/EMC
Air Force Snow	Snow depth	Air Force Weather Agency
SST	1-degree Reynolds, with Great Lakes SSTs	NCEP/EMC, GLERL
Sea and lake ice	Contains data on Canadian lakes, Great Lakes	NCEP/EMC, GLERL, Ice Services Canada
Tropical cyclones	Locations used for blocking of CMAP precipitation	Lawrence Livermore National Laboratory

Climatologies

Dataset	Used for, details	Source
Green vegetation fraction	Specification of vegetation cover extent, monthly interpolated to daily	NESDIS
Baseline snow-free albedo	Specification of land albedo, quarterly interpolated to daily	NASA

Fixed Fields

Land mask (land=1; sea=0)
Vegetation type [index, 1-13]
Soil type [index, 1-9]
Surface slope type [index]
Snow-free albedo [%]
Maximum snow albedo [%]
Surface roughness [m]
Soil column bottom temp. [K]
Number of root zone soil layers [non-dim]

It was a lot of work /
There were problems to solve:

- Acquire, or create (improved reprocessing) a variety of datasets, e.g.,
  - TOVS-1B data for 1979-2002 time period;
  - Air Force snow depth dataset;
  - High (RR) resolution sea and lake ice, and sea and lake surface temperatures;
  - Precipitation data (U.S., Mexico, Canada, CMAP) collection and processing (PRISM/"Mountain Mapper", disaggregation to hourly);

It was a lot of work / There were problems to solve, cont'd

- Improvements/ refinements of various components of EDAS, e.g.,
  - Precipitation assimilation (additional regions, blending, …)
  - Upgrade 3D-Var to run using the satellite bias correction for numerous additional satellites;

… a lot of work (cont'd)

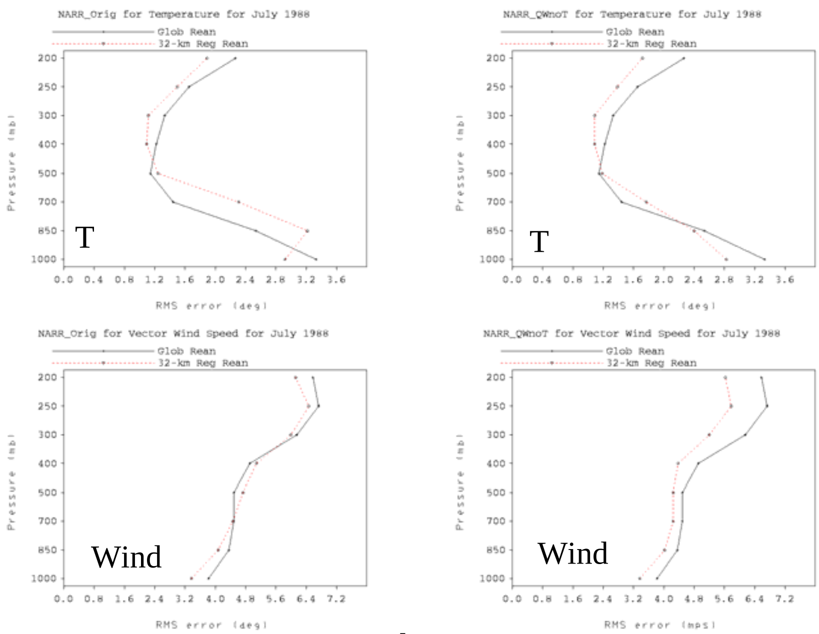
- Presented/ discussed RR-related work at numerous conferences
  - AMS Annual Meeting in Orlando, FL, in January 2002
  - GEWEX Mississippi River Hydrology & Climate Conference in New Orleans, LA, in May 2002
  - AMS Annual Meeting in Long Beach, CA, in February 2003
  - Climate Diagnostics Workshop in Reno, NV, in October 2003
  - AMS Annual Meeting in Seattle, WA, in January 2004
  - NARR Users Workshop, San Diego, January 2005
  - AGU Spring Meeting, New Orleans, May 2005
- Numerous contacts with potential users;
- Completed 24 years of RR production in just over 3 months, running 4 streams, on all of the previously mainframe NCEP IBM ASP

We've learned a few things along the way, e.g.,

- Assimilation of 2 m land surface station air temperatures, with the system we had:
  - Harmful in the sense of making the first guess considerably worse, throughout the troposphere (!)

"Full" Assimilation

No 2 m T assimilation



July 1998

A 2D-Var 2 m temperature analysis recently implemented in the NCEP operational EDAS

More “things learned” (disappointments ?):

- No overall advantage identified from the use of the Eta 4D-Var (“mixed resolution”) compared to the Eta 3D-Var (“fine resolution”);

More “things learned” [disappointments (?) cont’d]:

- While RR’s 3 vs GR’s 6 hr analysis frequency resulted in higher RR time resolution, it did not increase accuracy (3 hr first guess not closer to observations than the 6 hr first guess would have been)

More “things learned” [disappointments (?) cont’d]:

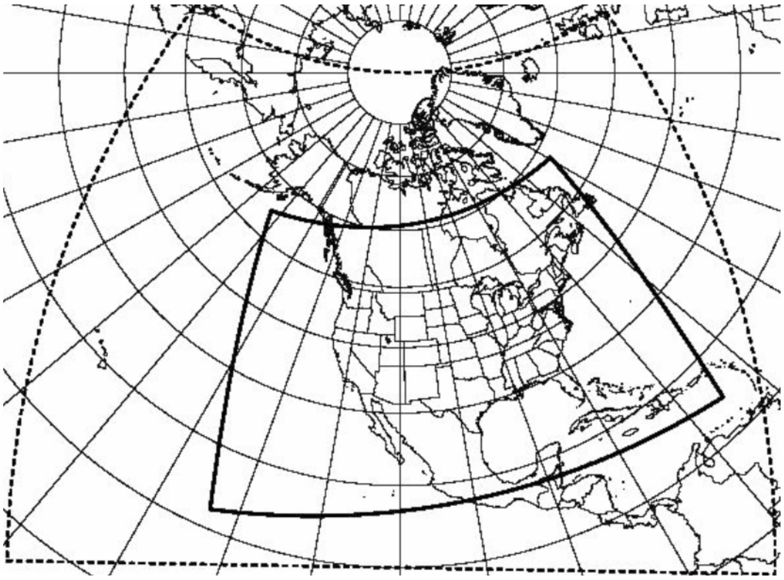
- No benefit identified from direct assimilation of radiances (the Eta Model top, 25 mb, too low ?);

But what matters most is the **result**, thus

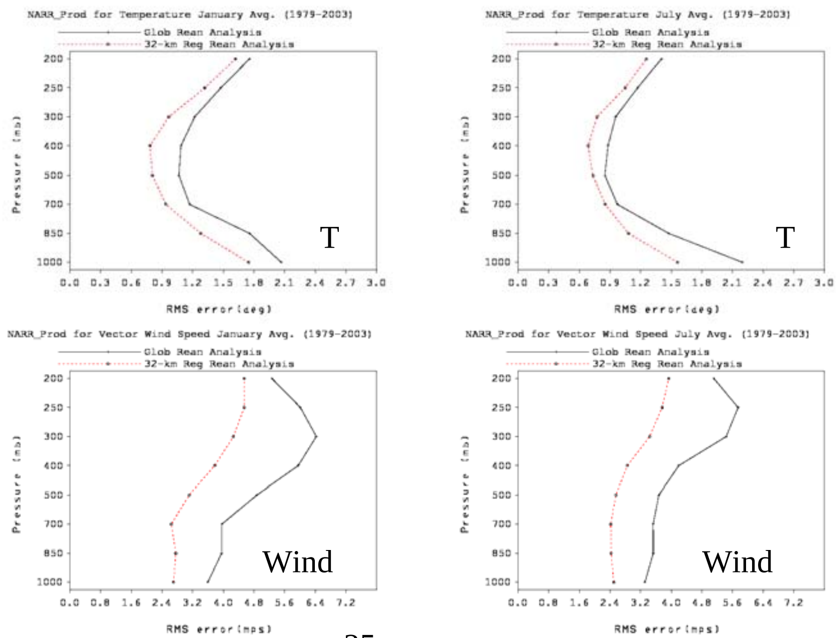
Upper-air (T, wind), and near-surface verifications

(“Near-surface”:
2 m temperature and 10 m wind)

Verification domain (heavy solid line) for upper-air and near-surface:

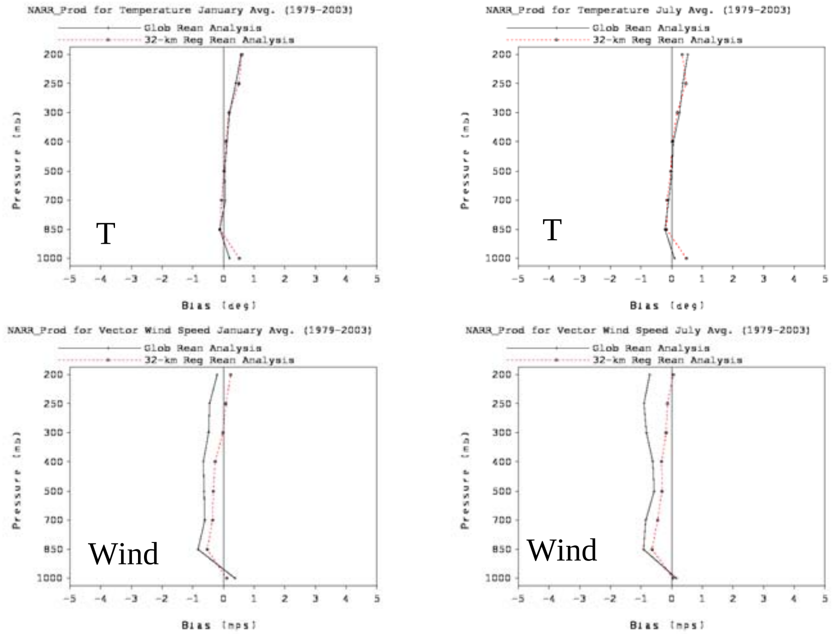


January Analysis, RMS July

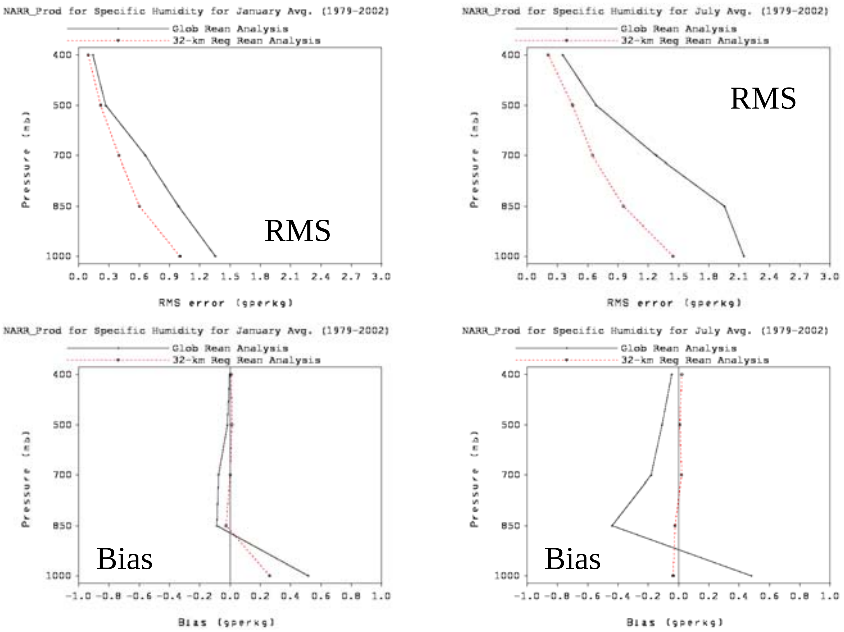


25 yr averages

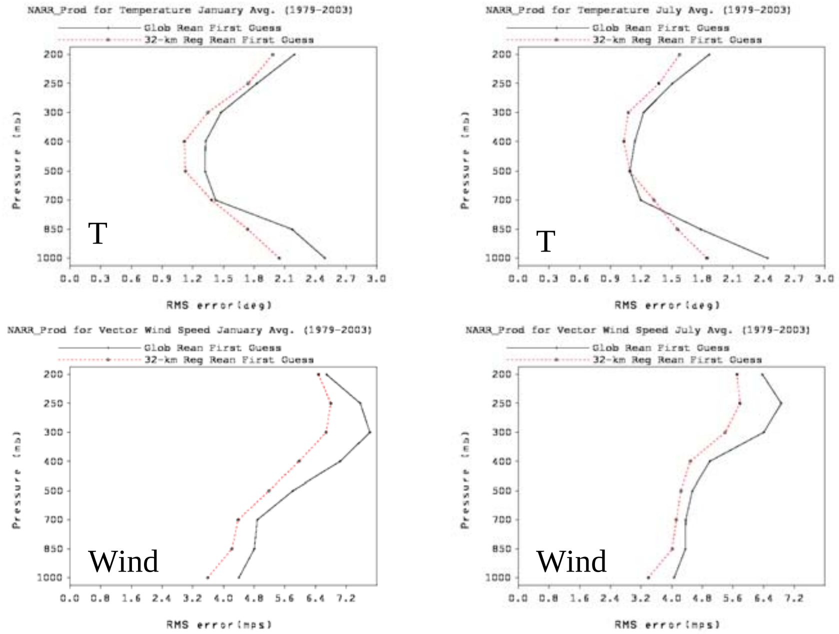
January Analysis, Bias July



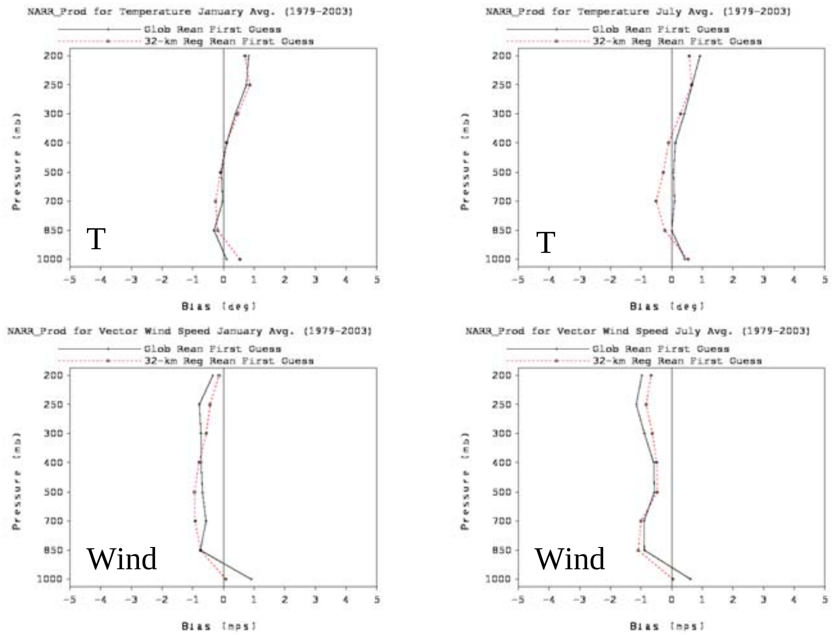
January Q, Analysis July



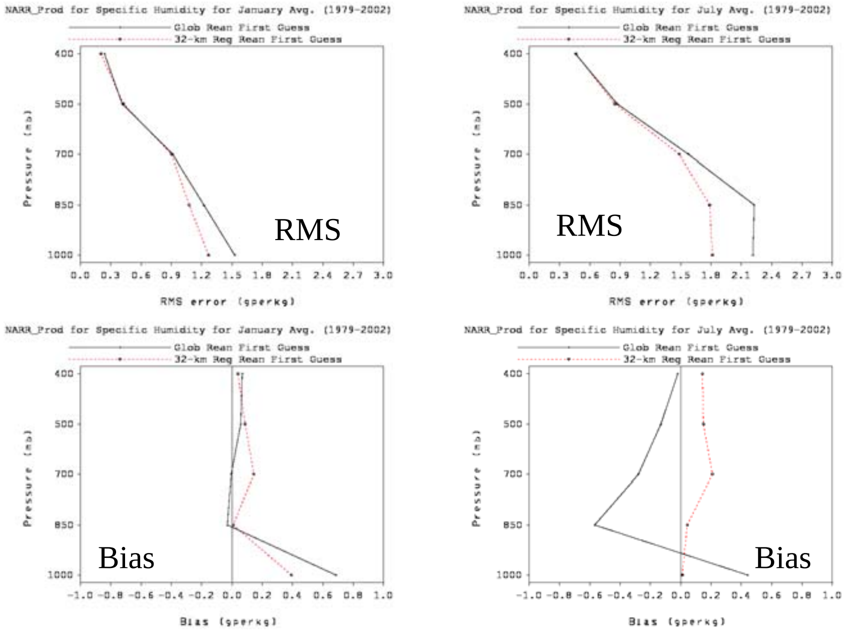
January First Guess, RMS July



January First Guess, Bias July



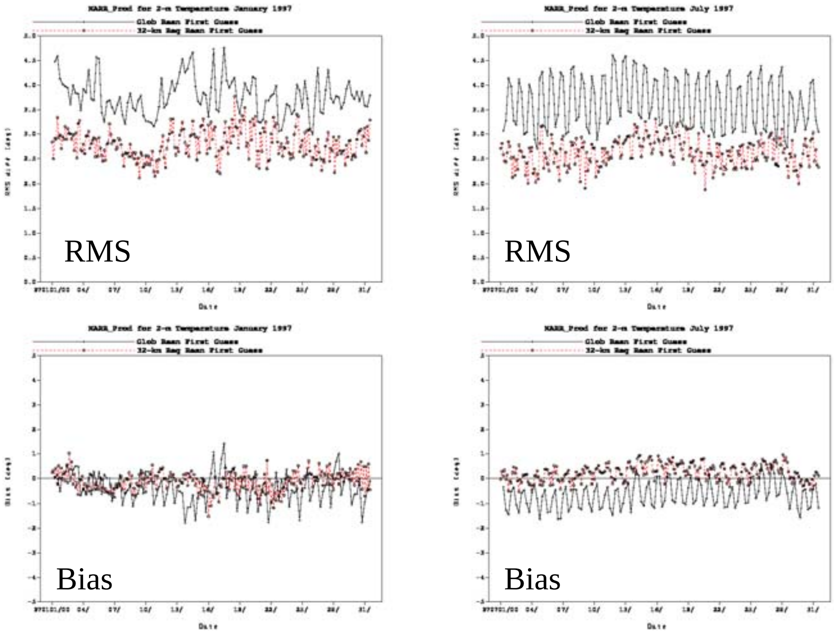
January Q, First Guess July



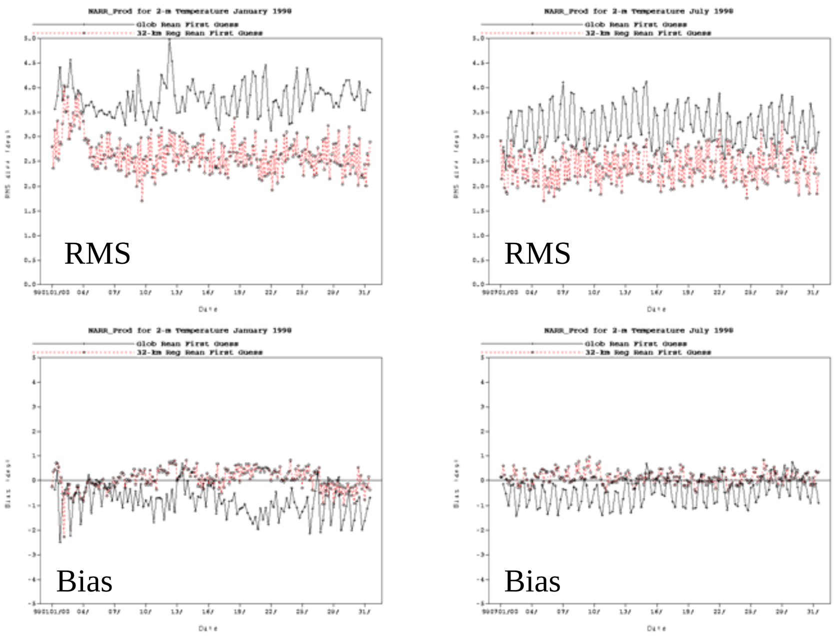
Near Surface

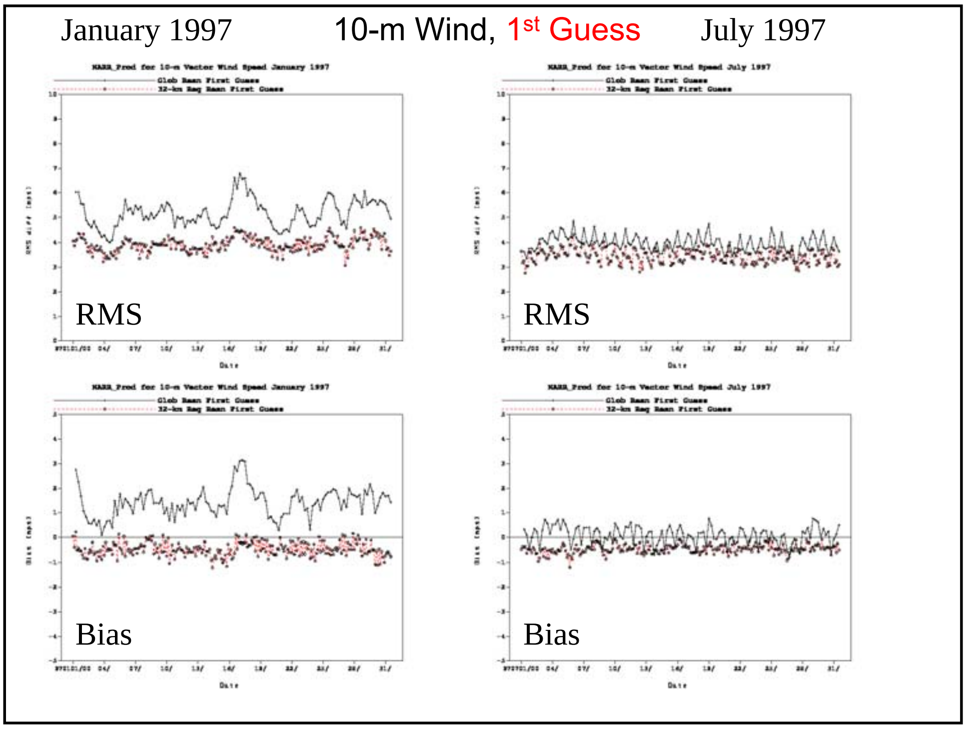
- Analysis (not shown): Comparison against the GR2 not possible. RMSs and biases look reasonable;

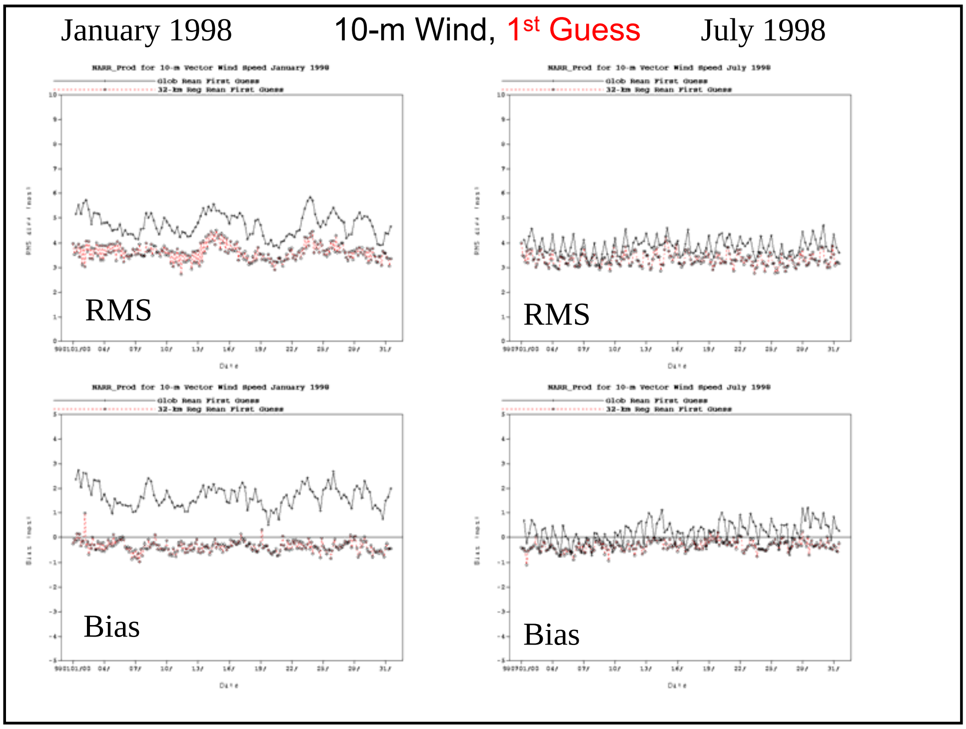
January 1997 2 m Temperature, 1st July 1997



January 1998 2 m Temperature, 1st July 1998




January 1997
10-m Wind, 1st Guess
July 1997
RMS
RMS
Bias
Bias


January 1998
10-m Wind, 1st Guess
July 1998
RMS
RMS
Bias
Bias

Summary, upper air and near surface:

- Just about all variables improved compared to GR, either considerably, or somewhat;
- Improvements generally greater in winter;
- Upper air: greatest improvements in winds at the jet stream level !
- Surface temperature: improved considerably both winter and summer;
- 10 m winds: improved considerably in winter, little bit in summer

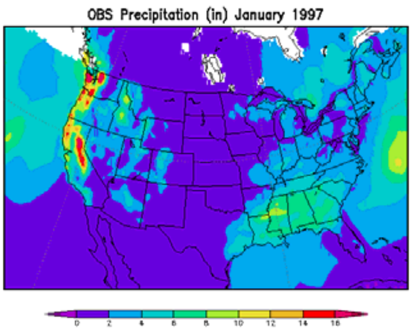
Results: Precipitation

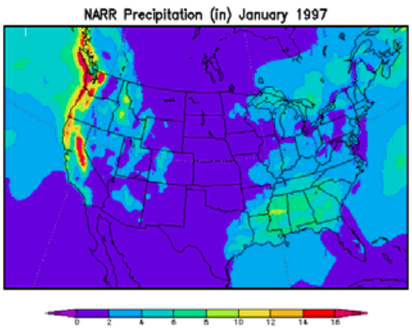
- Several sources of precipitation
  - ConUS data with PRISM (Mountain Mapper) to improve orographic effects
  - Canada
  - Mexico
  - CMAP (combination of satellite and gauge data) over oceans; CMAP is blocked:
    - Near central areas of hurricanes (7.5 by 7.5 deg)
    - Observed precipitation > 100 mm/day
    - A 15-degree "blending belt" between 27.5 and 42.5 N, with no CMAP north of 42.5 N

Precipitation (cont'd)

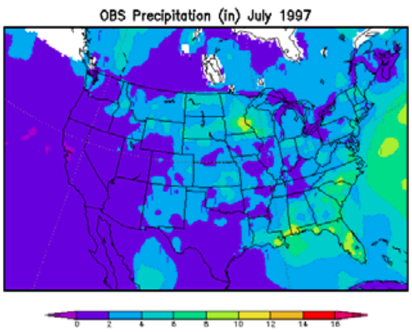
- Precipitation observations used to prescribe the latent heat profile in the model;
- Model uses given latent heat profile to simulate precipitation;
- Moisture adjusted not to be in conflict with the precip analysis used

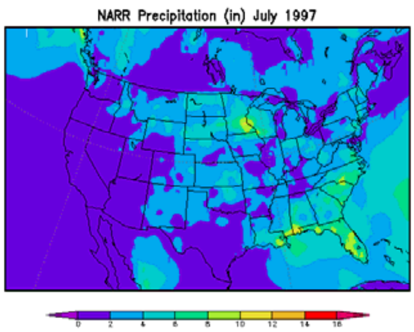
January 1997 Precipitation Results



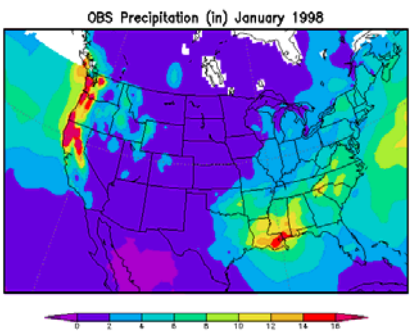


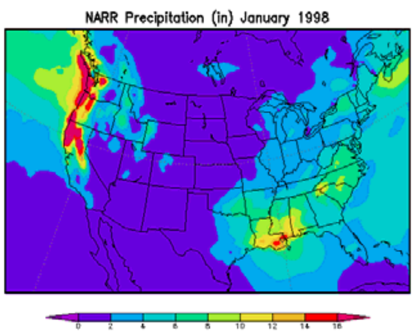
July 1997 Precipitation Results





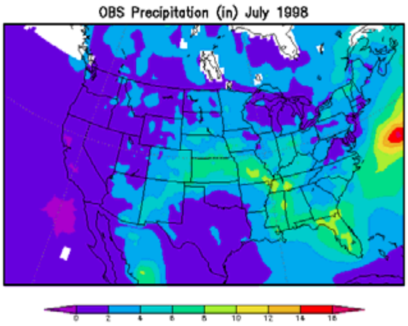
January 1998 Precipitation Results

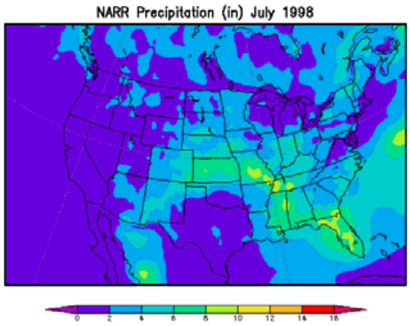




(A strong El Niño)

July 1998 Precipitation Results





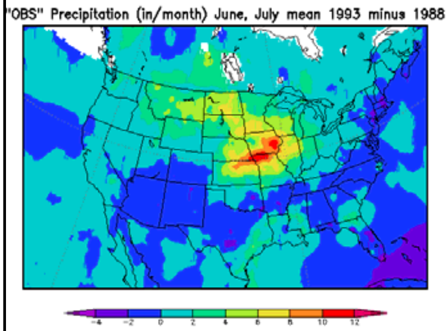
Summary (precip)

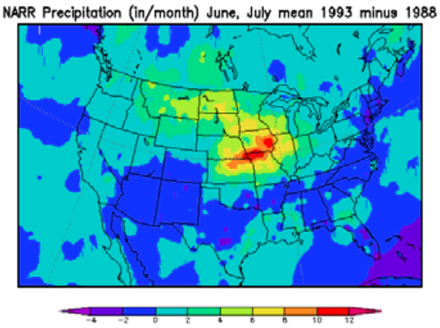
- Precip over land very near analyzed, bodes well for land surface hydrology over the ConUS area, were the precip analyses are most accurate

Flood of 1993 vs Drought of 1988

- Shows impact of summer precipitation;
- Extreme years compared;
- Monthly mean June/July precip of 1993 minus monthly mean June/July precip of 1988;
- Shows the success of precipitation assimilation even for drought and flood years

Flood of 1993 vs Drought of 1988



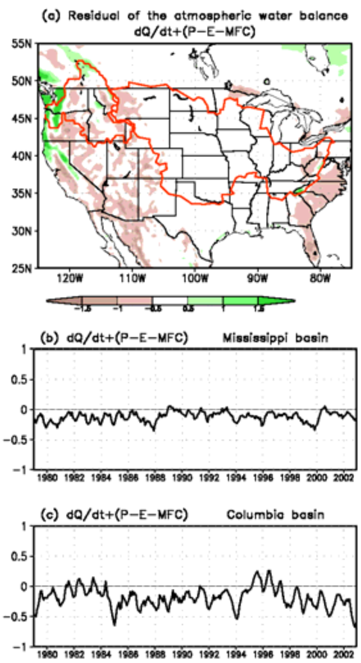


Moisture budget

Roads et al. (JGR, 2003):
comparison of a number of models and analyses;
need to better "close the budgets"

$$\frac{\partial Q}{\partial t} + P - E - MFC = R$$

Roads et al.: 4 years of Eta operational analyses;
residuals significantly smaller than GR1 and GR2



The typical magnitude
(after removing the mean):
about half of those of GR2

The mean:
about 1/5 of that of GR2

mm/day; (b) and (c): 13 month running means

R-CDAS

- Running in near real-time (similar to the current CDAS that runs with the Global Reanalysis system),
- Precipitation:
  - CMORPH, a 1/2 hr high-resolution satellite-based dataset, replaces CMAP over oceans, used for 2003 and beyond;
  - No Canadian precipitation; using EDAS precipitation instead
  - Mexico and US (N-LDAS) precipitation in real-time;
- Data available after ~ 21 hr delay
- System, including data ingest script, created and tested
- Currently running ~ 4 days behind real time, responsibility of NCEP/CPC

Our output files produced (about 75 Tb total):

Six sets of reanalysis files, at 3 hr intervals each:

- "Restart files": on "native" (model) grid, all variables needed to restart the model;
- On the "native" (E) grid, but interpolated to p levels:
  Analysis files, First guess files;
- Interpolated in addition to Grid 221 (Lambert, "AWIPS"):
  Analysis files, First guess files;
- "Merged" files (archived): Grid 221, a selection of (mostly) analysis variables, enhanced by, e.g., fluxes ("fcst", from 1st guess), 3-h averages, and 3-hr accumulations (4 types of time validity)

Output files produced, cont'd

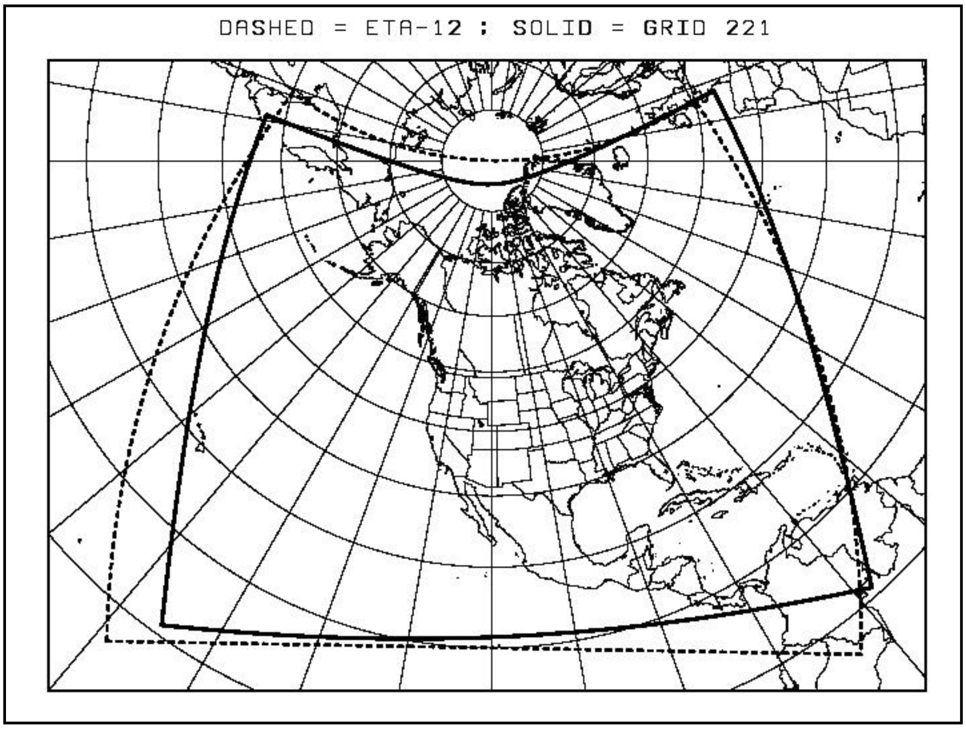
- Input observations;
- Climatological and fixed fields used;
- Various plots;

- Special subsets:
  - Land-surface subset - fields related to land-surface/hydrology modeling
  - Climatological subset (CPC)

- "Free forecasts" (reforecasts): 3-day forecast every 2.5 days

Output files produced, cont'd; more on the merged files

- Two main sets of the merged files produced; so-called merged-a and merged-b;
- Temporal frequency 3 hour (8 files per day, archived in a single tar file)
- Merged-a daily tar file is about 450 Mb, merged-b about 75 Mb;
- Horizontal grid standard NCEP grid 221 (Lambert conformal, 32 km). Note that a sliver of 221 is sticking out of the RR domain/ variables undefined;
- In vertical: pressure levels, "hybrid levels" (lowest eta mid-layers), soil layers, boundary layer etc.;
- Fields in merged-a files mostly analysis; some 3-hr accumulations, 3-hr averages, and 3-hr forecasts;
- When the same filed is output as analysis and 3-hr forecast, 3-hr forecast is placed in merged-b file (GrADS problem);
- Some fields are interpolated to 221 using nearest-neighbor algorithm (labeled with asterisk in column three in the list)



Archiving

- Two (or more) archiving centers/ two have the "merged fields" + some of the special smaller sets:
  - National Climatic Data Center (NCDC);
  - National Center for Atmospheric Research (NCAR);
  - San Diego Supercomputing Center (SDSC): contacts in progress, maybe;

NARR Archive at NCAR

- All NARR data files are on-line at
http://dss.ucar.edu/pub/narr

- ftp download is available

- Contact:
chifan@ucar.edu
303-497-1833

Chi-Fan Shih

NARR on the Web

http://wwwt.emc.ncep.noaa.gov/mmb/rreanl/index.htm
l

Summary (results)

- Long-term, consistent, high-resolution climate dataset for North America a significant improvement over earlier global reanalyses;
- Near-surface temperatures and winds closer to observations;
- Winds in the upper troposphere, in addition to land-surface hydrology over the ConUS, maybe the strongest areas of improvement;
- Improvements greater in winter than in summer;

But also a few problems/ weaknesses, that we mostly became aware of after the fact:

- Excessive strength of the summertime Gulf of California low-level jet, :-(NAME. Not understood at this point;
- Precipitation over Canada: the number of gauge observations available may have been insufficient to do better than the model might have done;
- Precipitation over northern Atlantic not simulated well;
- Atlantic hurricanes apparently not an RR strong feature;
- . . . (TBD)

confident:

objectives set out at the outset of the project:

to create a long-term, consistent, high-resolution climate dataset … as a major improvement … fully met;

and that

the NARR will for some years to come be the best/ most accurate North American weather and climate dataset, for numerous research and application purposes

Early Usage

In contact with 20-30 groups. What are people doing? Some of the groups outside NCEP:

NARR validation

West, Steenburgh, Univ. Utah

Water and energy budget, hydrologic cycle studies (predictability, climate variability, …)

Dery, Wood, Kerr, Princeton Univ.;

Gochis, NCAR;

Korolevich et al., Nat'l Resources Canada, Ottawa;

Luo, Berbery, Mitchell, Univ. Md, College Park;

NA monsoon studies/ moisture fluxes

Salstein, Cady-Pereira, Atmos. Environm. Res., Lexington, MA

Dynamical and precip structures/ systems, 1979-2003, or specific episodes, interannual variability

Caetano, Mendez, Magaña, Nat'l Univ. Mexico;

Milrad, McGill Univ., Montreal;

Ruiz-Barada, Nigam, Univ. Md, College Park;

Model validation, severe weather predictors

Durnford, Gyakum, Atallah, McGill Univ., Montreal;

Jaye, Tripoli, Univ. Wisconsin, WI;

Impact in simulation of spec. systems, in driving regional climate simulations

Nunes, Roads, Kanamitsu, Arkin, Scripps ECPC, La Jolla, CA

Vasić, Xue, UCLA

Wind energy assessment and air pollution transport

Moon, WindLogics Inc.

Water management engineering

P. Trimble, Southern Florida Water Management Directorate

…