Understanding Analysis Uncertainty

David Myrick
Western Region Headquarters
Scientific Services Division

2nd RFC Verification Workshop - 11/20/08

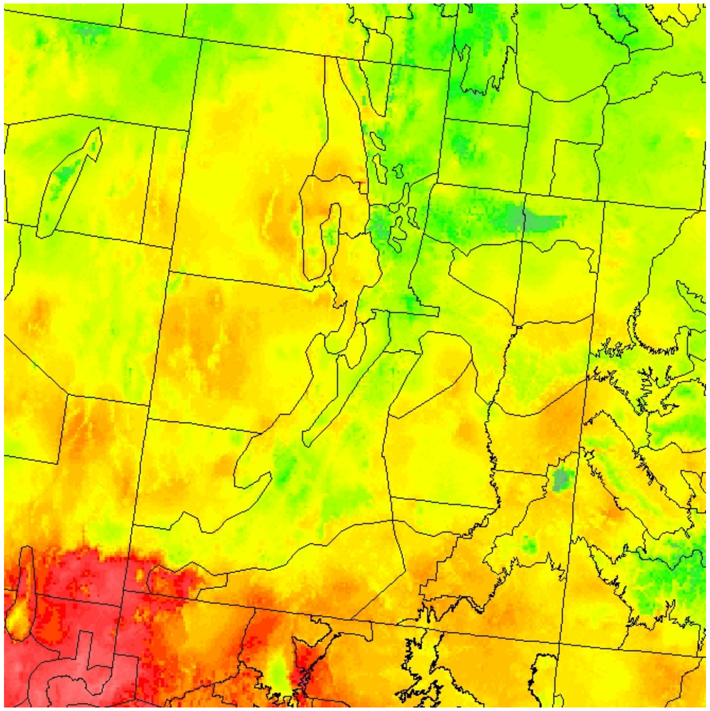
Acknowledgements

- John Horel & Dan Tyndall (Univ. of Utah)
- Manuel Pondeca (NCEP/EMC)

Outline

- Sources of analysis uncertainty
- RTMA/AOR Project
- Using analysis uncertainty estimates in forecast verification

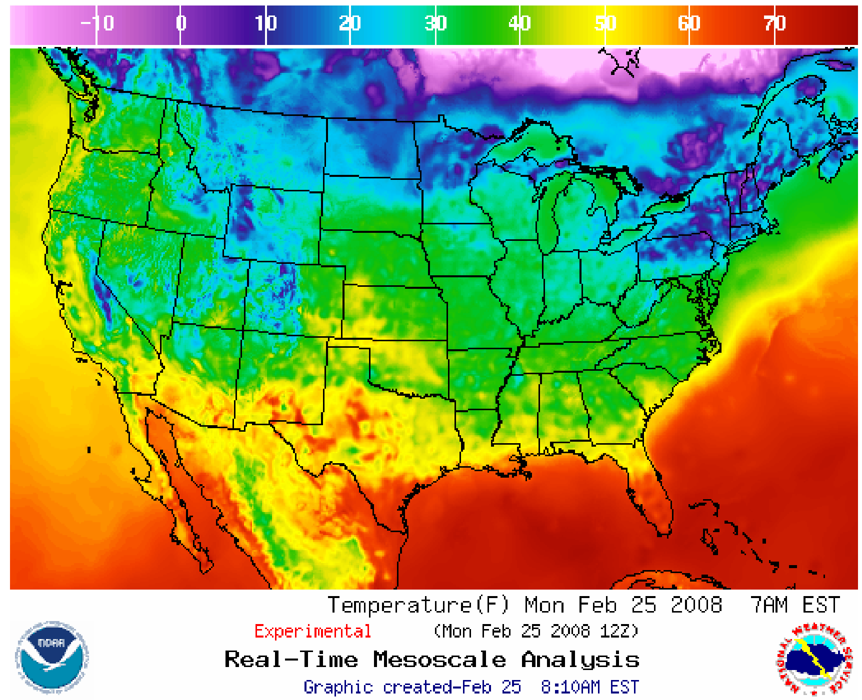
Objective Analysis



- A map or picture of a meteorological field
- Relies on:
  - observations
  - background field
- Used for:
  - Initialization for a model forecast
  - Situational awareness
  - Verification grid

Objective Analyses

- Hand drawn analysis
- Model initialization panel
- LAPS
- MSAS
- MatchObsAll
- NCEP Reanalysis
- NARR
- RTMA
- AOR (future)



ABC's of Objective Analysis

- In the simplest of terms:

Analysis Value =

Background Value +

Observation **C**orrection

Background Values

- Obtained from an analysis:
  - Climatology
  - An objective analysis at a coarser resolution
  - Short term forecast
  - Analysis from previous hour

Observations

- Observations are not perfect…
  - Gross errors
  - Local siting errors
  - Instrument errors
  - Representativeness errors

A good analysis requires…

- A good background field
- Observations with sufficient density to resolve critical weather & climate features
- Information on the error characteristics (uncertainty) of the obs & background
- Analysis scheme that takes into account that obs & background contain errors

Source of Confusion

- Analysis systems like MatchObsAll suggest that the analysis should exactly match every ob
- Objective analysis values usually don't match surface observations
  - Analysis schemes are intended to develop the "best fit" to the differences between the observations and the background taking into account observational and background errors when evaluated over a large sample of cases

Incorporating Errors

- Basic example:

$$T_a = T_b + W(T_o - T_b) \qquad W = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}$$

σ_b = background error variance

σ_o = observation error variance

- So – the analysis won't always match an ob

Objective Analysis Schemes

- Successive Corrections
- Optimal Interpolation
- Variational (3DVar, 4DVar)
- Kalman Filtering

simple

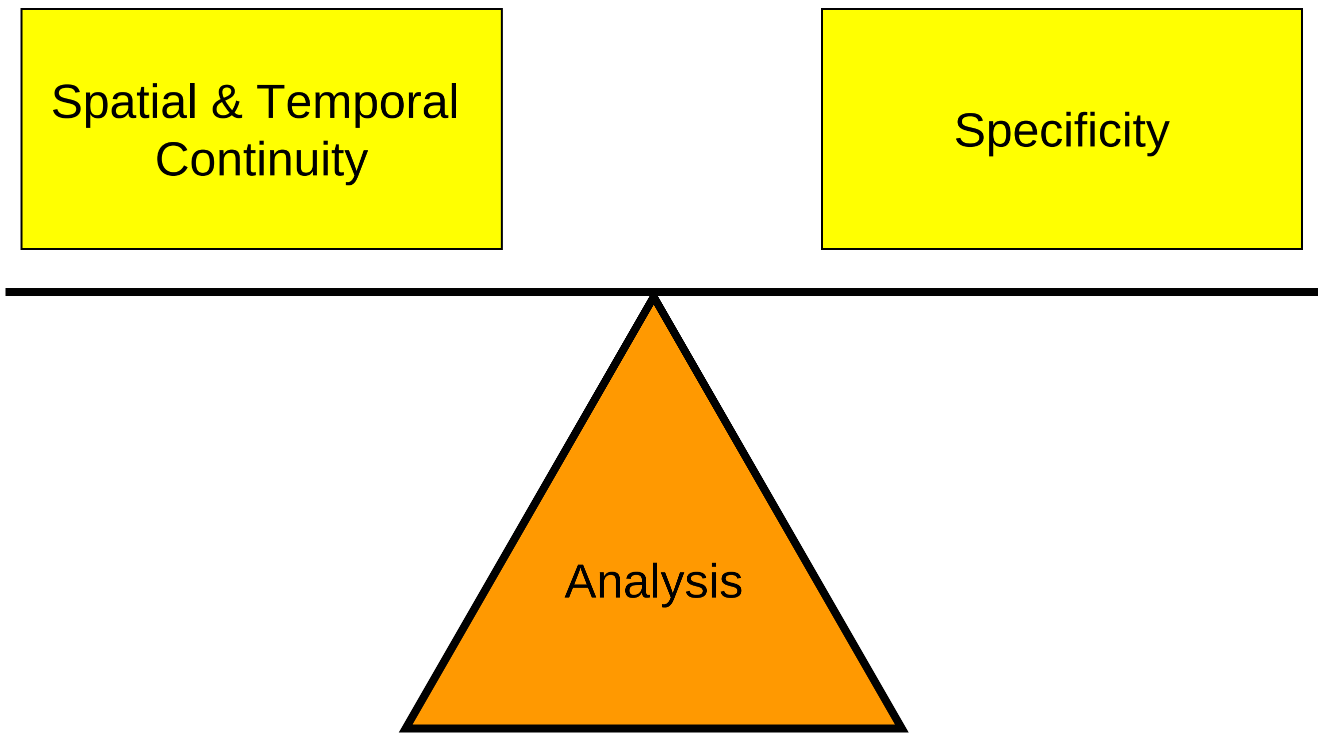
↕

complex

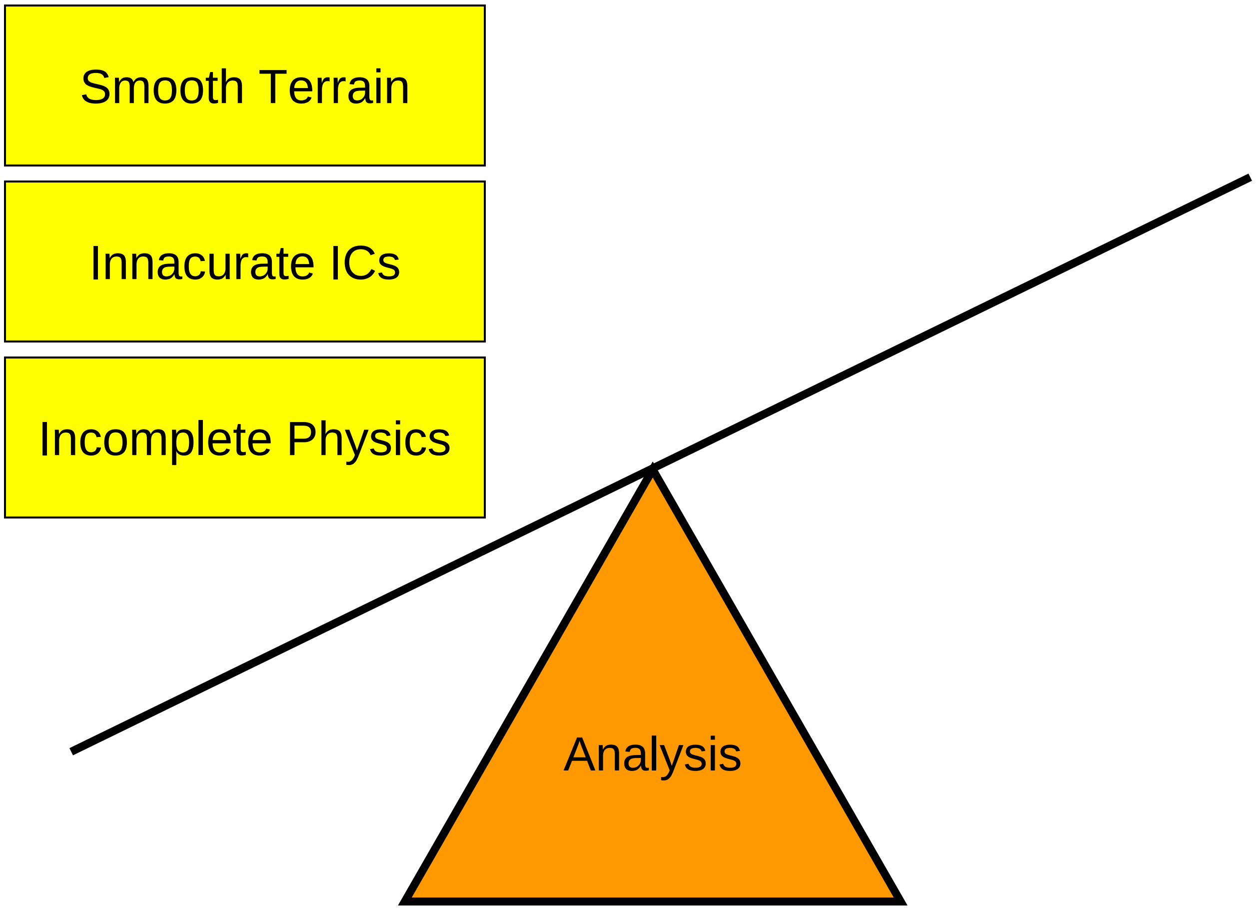
- Kalnay (2003) Chapter 5 – good overview of different schemes

Need for a balance…

Models or observations cannot independently define weather and climate processes effectively

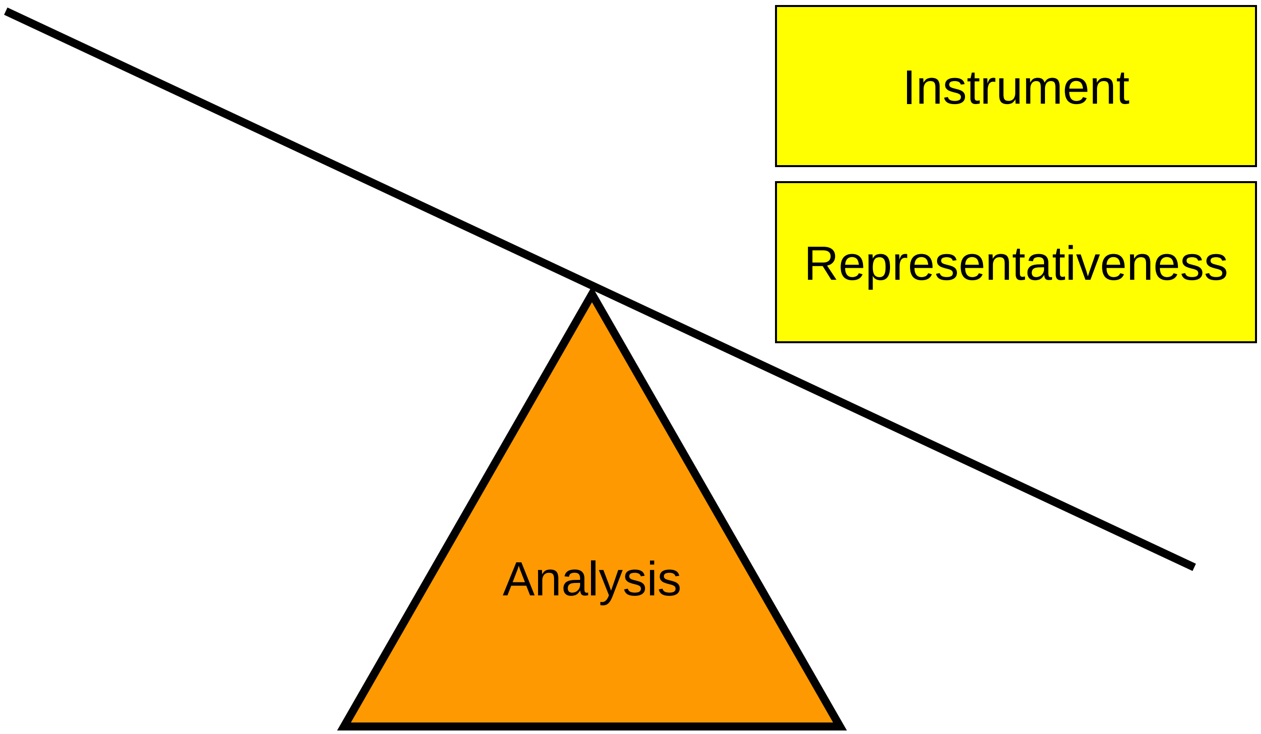


Sources of Error
NWP Model Errors



Sources of Error

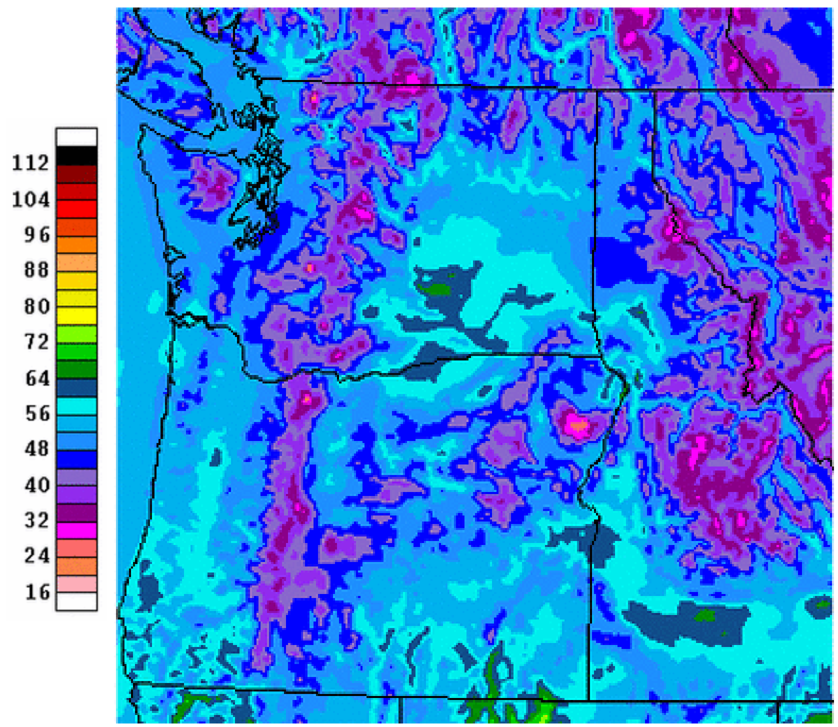
Observations



RTMA Project

(De Pondeca et al. 2007)

- 1st step in a multi-year project to build an "Analysis of Record" (Horel and Colman 2005)
- Collaborative effort
  - NCEP/EMC
  - ESRL/GSD
  - NESDIS
  - And lots of input from NWS & Universities



0000 UTC 4 October 2007
RTMA Temperature (^{o}F)

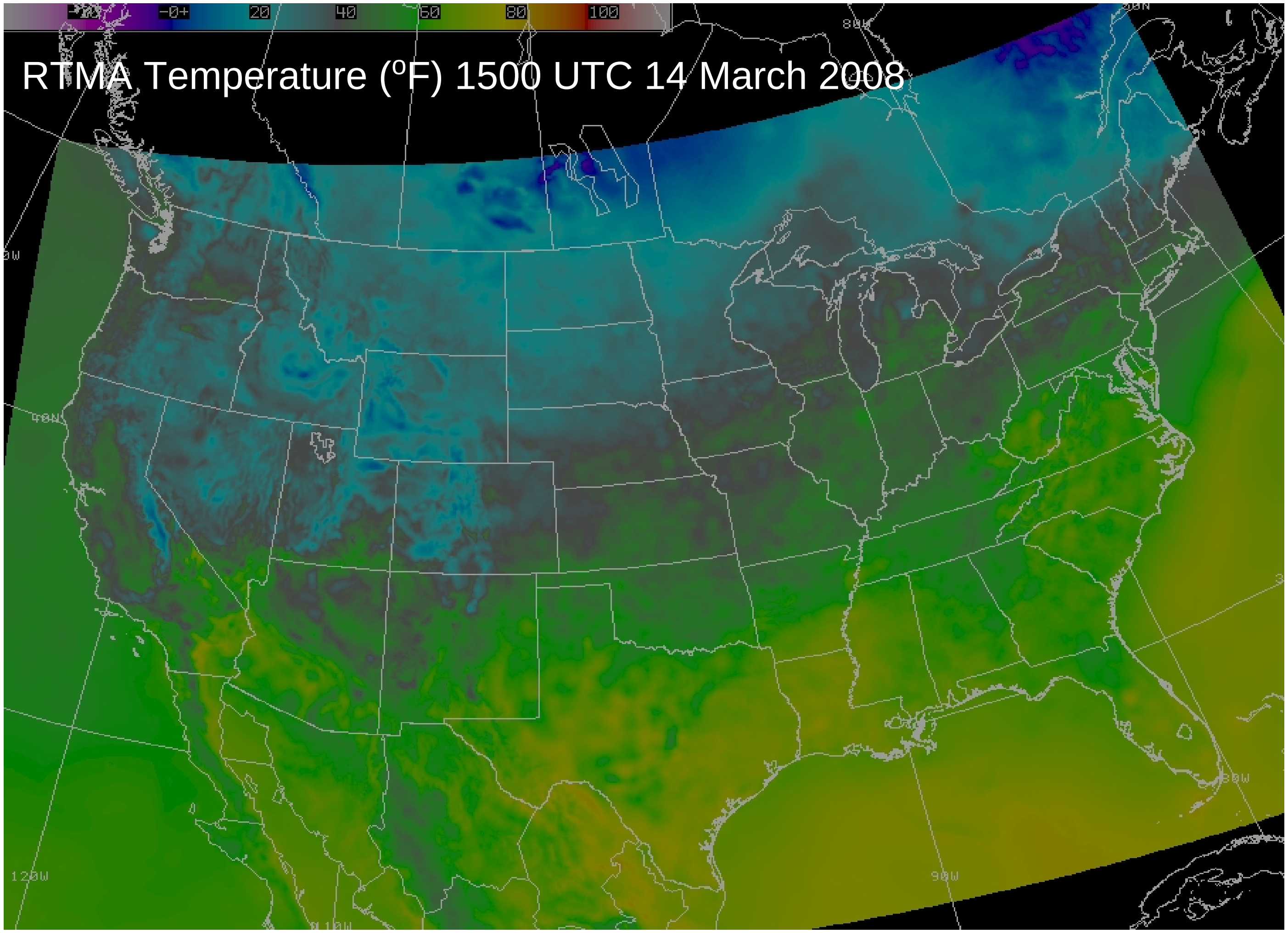
RTMA Program Components

1. **Real-time Mesoscale Analysis (RTMA) – Current Effort**
   - Hourly within ~45 minutes of nominal observation time
   - Initially a prototype, or proof-of-concept, for AOR

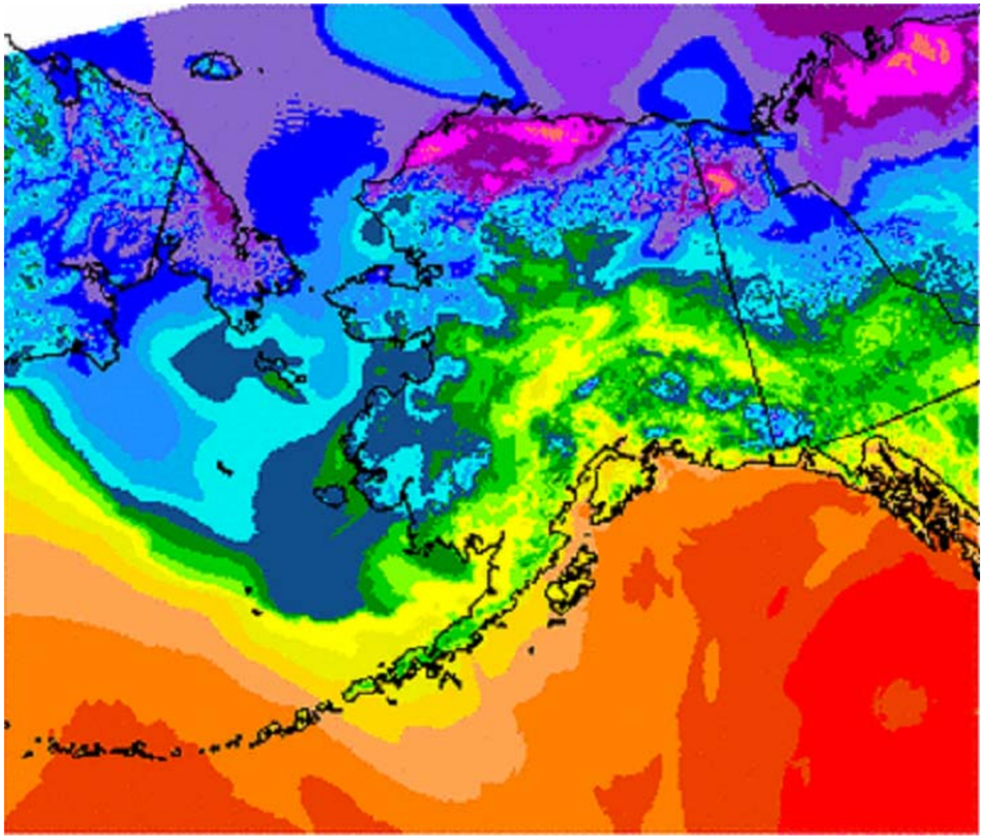
2. **Analysis of Record (AOR) – Future Effort**
   - State-of-the-science analysis (best possible)
   - Delayed for late arriving data assets
   - Methodology to be determined (likely community effort)
   - Accepted ‘truth’ for use in studies and verification
   - Not yet funded (funding proposed)

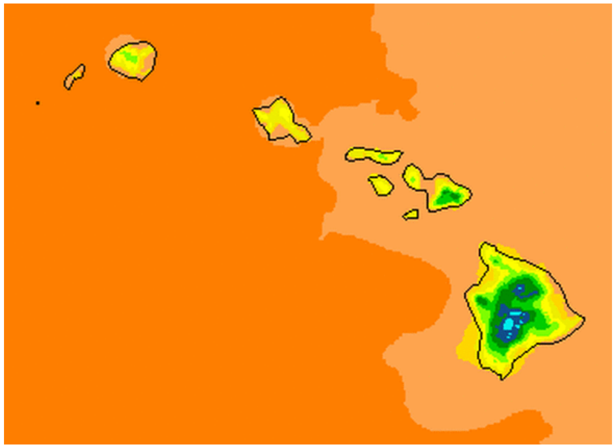
3. **One-time-only reanalysis – Future Effort**
   - Apply mature AOR methodology retrospectively
   - 30 year time history of AORs
   - Not yet funded (funding proposed)

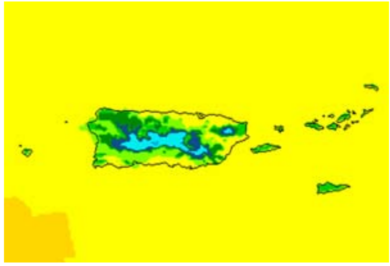
CONUS RTMA


-0+
20
40
60
80
100
RTMA Temperature (°F) 1500 UTC 14 March 2008
120W
90W
80W

OCONUS RTMA Regions







RTMA Variables (available hourly)

- Temperature, Dewpoint, & Wind
  - Analysis "uncertainty" grids for each variable
  - More analysis variables planned (not funded yet)
- Precip field = NCEP stage II
- Sky cover = GOES effective cloud amount from NESDIS

RTMA Specs

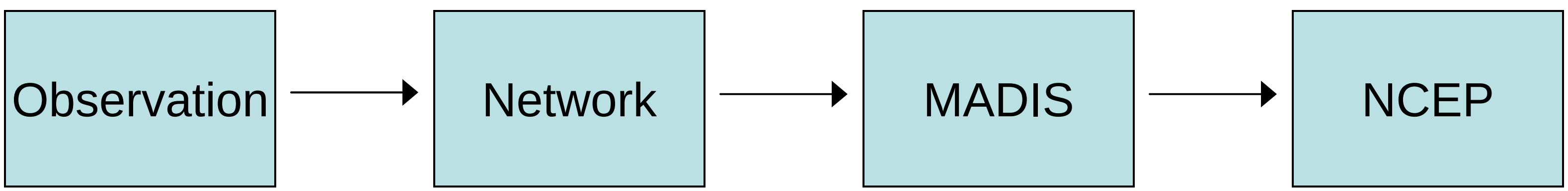
	CONUS	Alaska	Hawaii	Puerto Rico	Guam (Planned)
Resolution	5-km	~6-km	2.5-km	2.5-km	2.5-km
Background (downscaled)	RUC	NAM	NAM	NAM	GFS
Precip?	NCEP Stage II analysis	No	No	No	No
Effective Cloud Amount?	GOES sounder	No	No	No	No

Analysis Methodology

- NCEP Gridpoint Statistical Interpolation (GSI) was modified from 3-D to 2-D to create surface analyses
- Background fields are downscaled
  - CONUS 13-km RUC to 5-km NDFD grid
- Surface observations from MADIS
- Iterative analysis
- Anisotropic background error covariances
  - Based on terrain (future plans include info from prevailing wind direction, potential temp, etc.)

Surface Data Issues

- Real-time RTMA analysis begins ~30 min past the hour
- Getting the data in time is a challenge
  - RTMA uses obs taken (+/-12 min from top of hour)
    - RAWS (-30 min to +12 min)



- Many remote obs still don't make it in time, as well as some obs with weird data dumps (Snotel = every 3 hrs)

RTMA Quality Control

- MADIS QC flags – obs withheld if MADIS flag is bad >25% of the time, list updated quarterly (future: to updated monthly)
- WFO blacklists – submitted through your regional SSD, updated quarterly
- Gross checks
- RTMA dynamic reject list – based on stats from (ob – guess)

RTMA

- “operational” version in AWIPS
  - Updates implemented by NCO ~1x/year
- “parallel” version available by ftp
  - Used for testing changes
  - A “work in progress” with little to no funding
  - Field input/feedback important!
- RTMA Training Module Available from COMET (S. Jascourt)

Uncertainty Verification

- Motivating Question: Is there a way we can define what constitutes a "good enough" forecast?
- Problem: Much hesitation when it comes to grid-based verification
- Reality: Forecasters need feedback across the entire grid

The dreaded *verifying* analysis

- “The objective analysis never draws for the cold air that pools in the X valley!”
- “I’ll be penalized for adding detail to my grids!”
- “The analysis never matches the observations in my northern mountain zones!”

Forecaster Concerns

- Quality of the verifying analysis
- Penalty for adding mesoscale detail to grids in areas unresolved by analysis
- Bad (mesonet) observations influencing analysis
- Analysis in remote areas – driven mostly by the background model

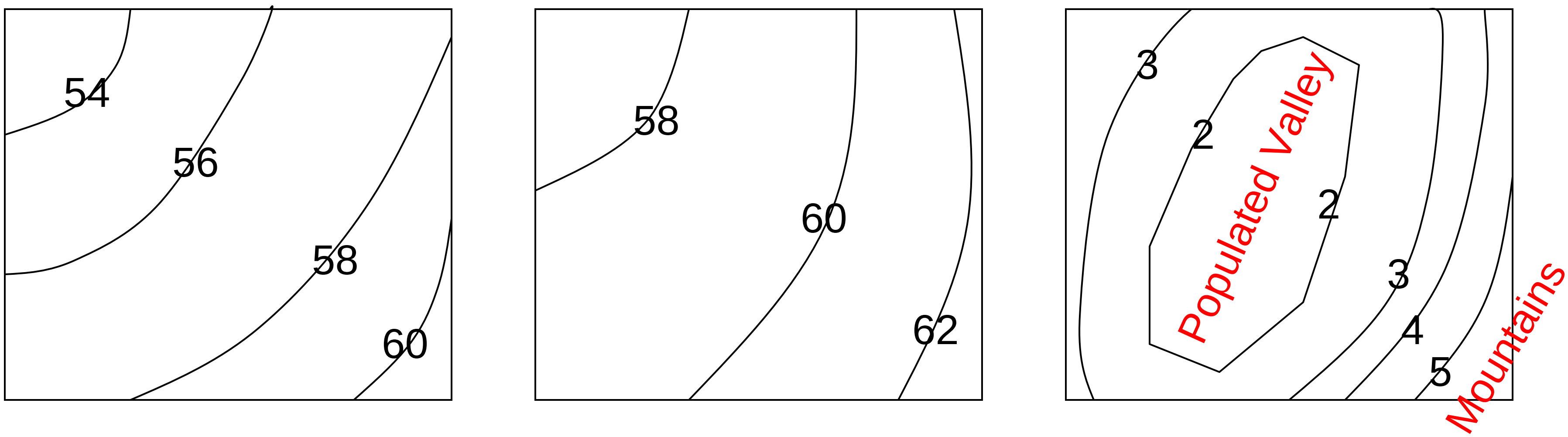
RTMA Uncertainty Estimates

- Experimental product (under development)
  - Error is computed using Lanzcos method in conjunction with the conjugate gradient method of the GSI minimization procedure
- Goal:
  - Higher uncertainty in data sparse areas
  - Lower uncertainty in data dense areas
- Currently available for T, T_d, Wind

Idea

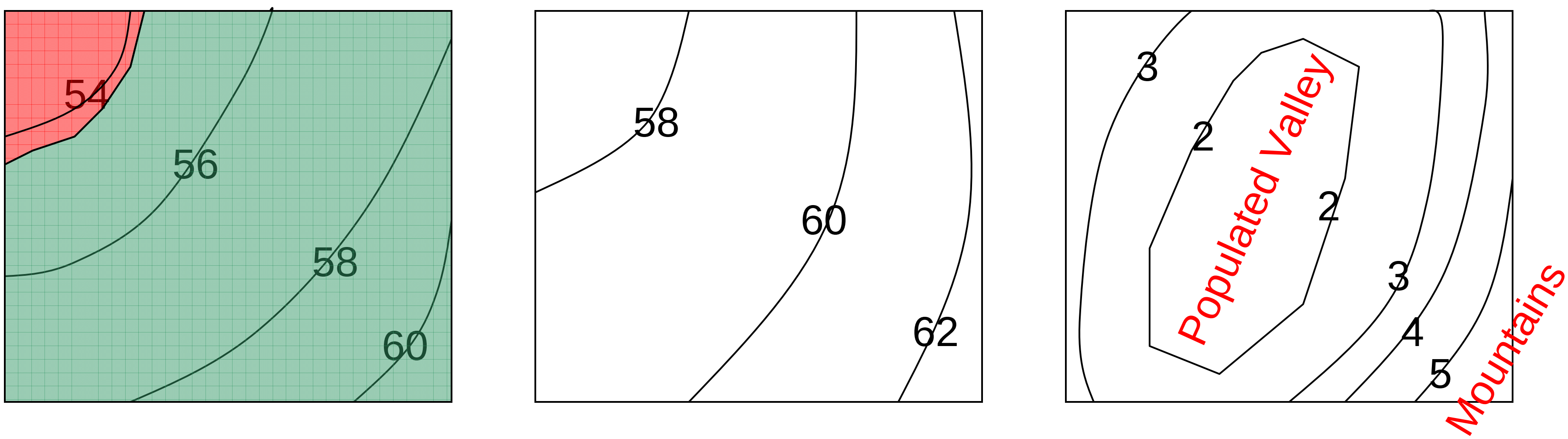
- Can we use the RTMA analysis uncertainty estimate as a proxy for a good forecast?
  - Lower margin of error in areas with obs
  - Forecasters would not be penalized as much in areas where the analysis struggles

Temperature (°F) Forecast Example



Forecast RTMA RTMA Uncertainty

Temperature (°F) Forecast Example



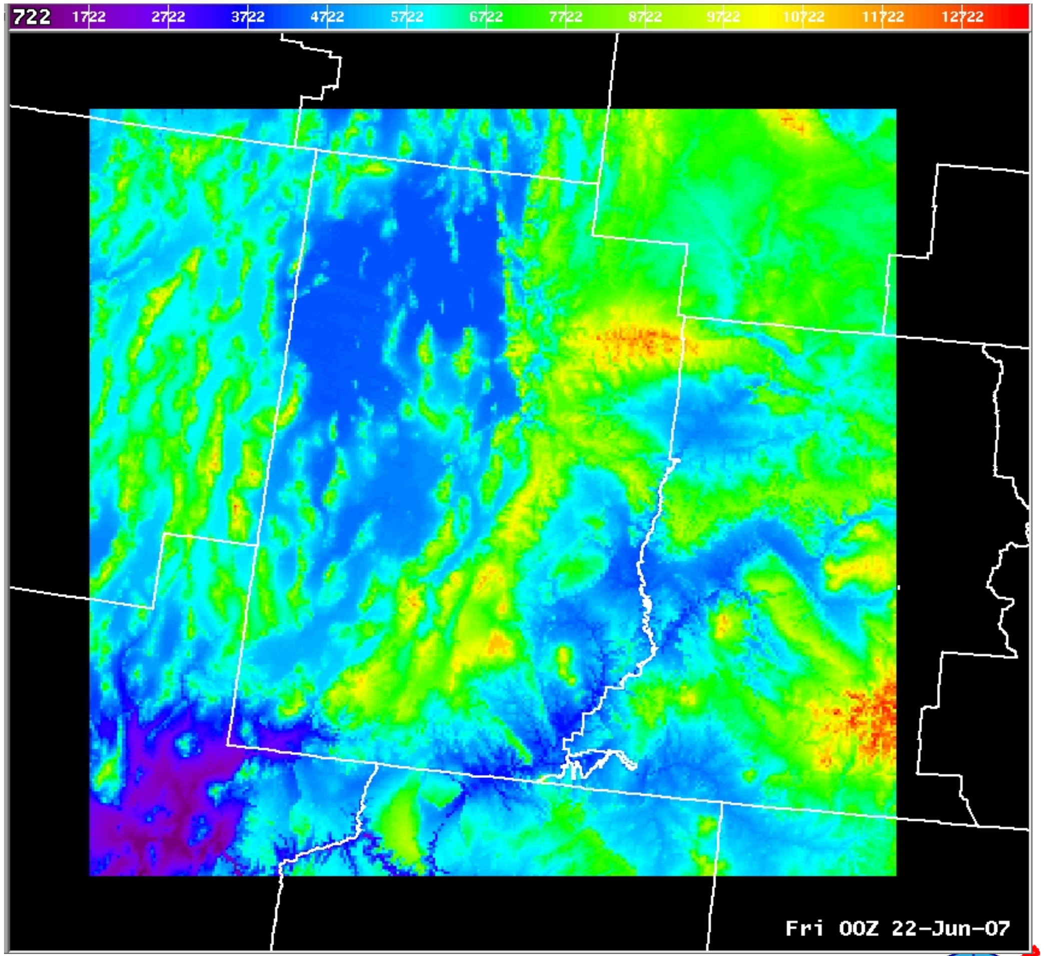
Forecast RTMA RTMA Uncertainty

Green = Forecasts are within the bounds of the analysis uncertainty

Red = abs(RTMA – Forecast) > Uncertainty

Utah Example

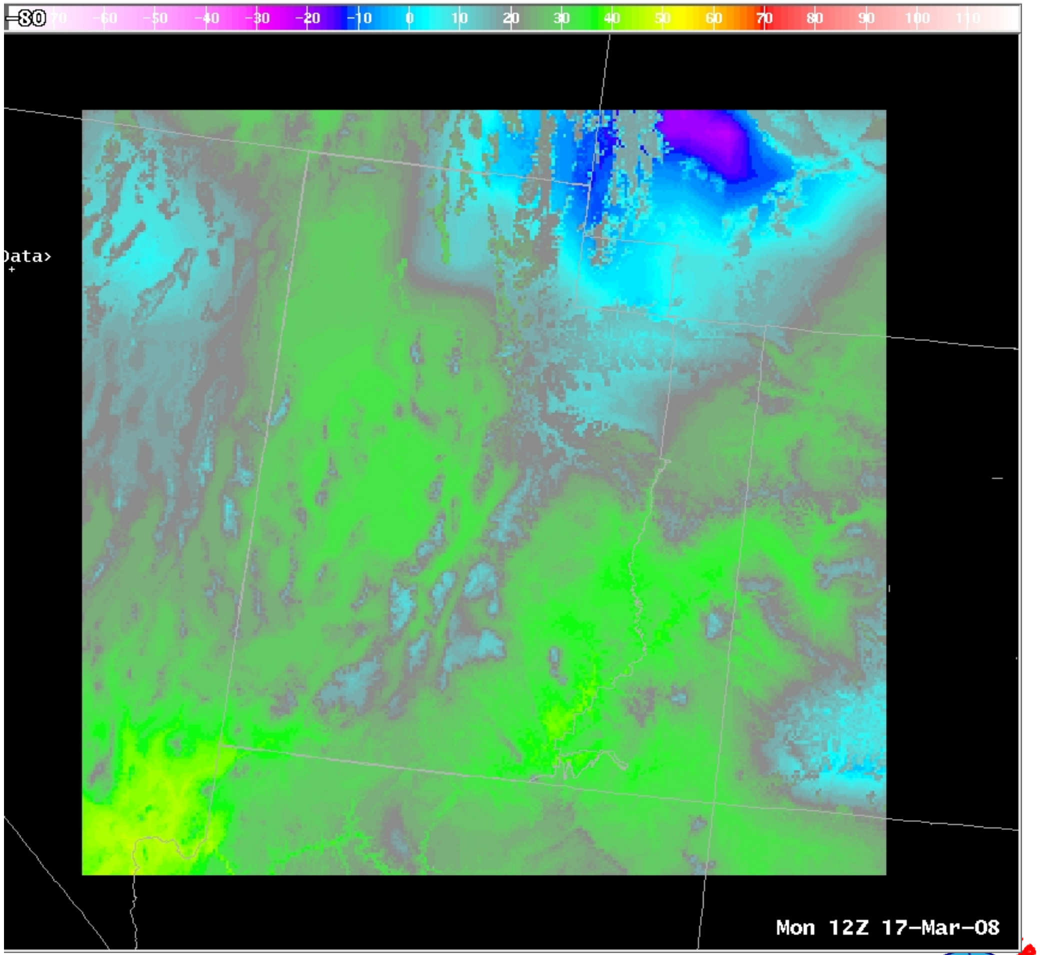
- NDFD terrain (used by the RTMA) captures the complex mountain/valley topography of the Great Basin



GFS40 T (^{o}F) Initialization

1200 UTC 17 March 2008

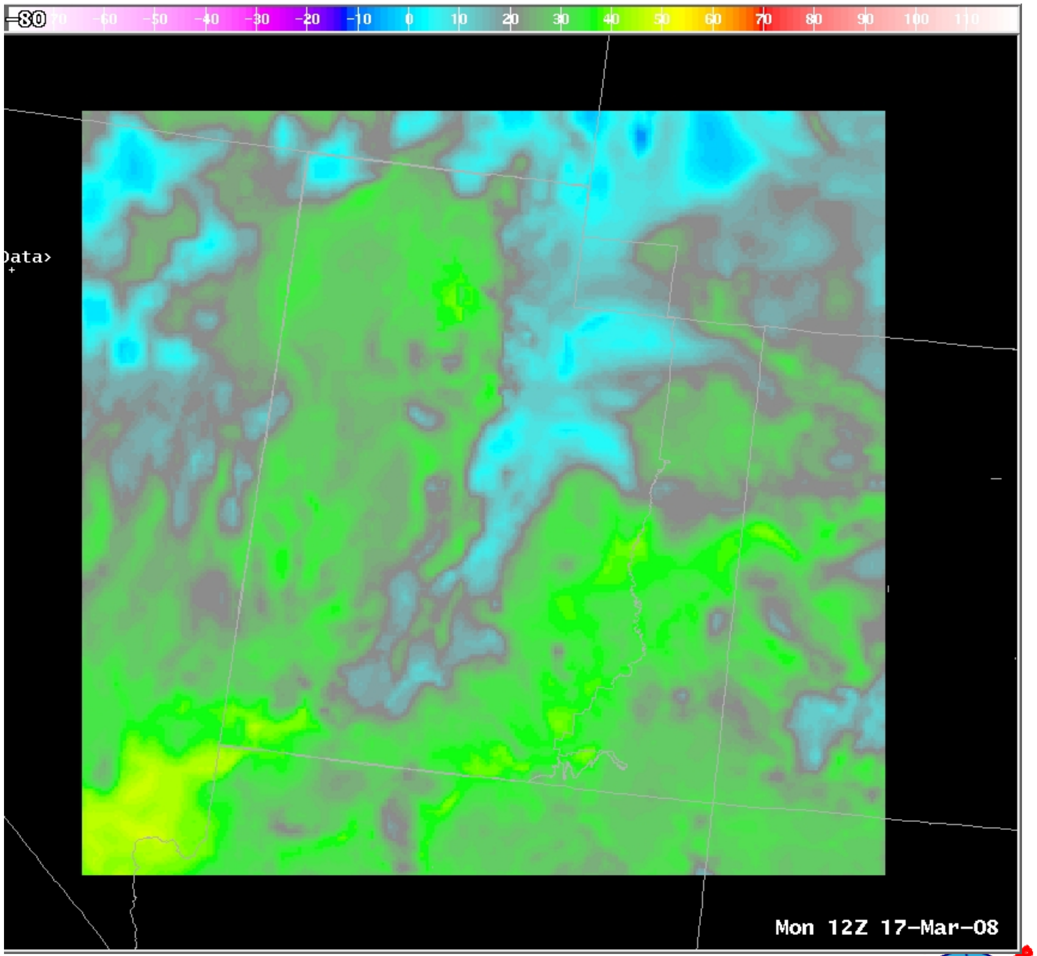
- Using model data to test technique
- GFS40 smartinit does a fairly good job downscaling to the terrain
- GFS40 contains sharp boundary in temps just north of Salt Lake City



Parallel RTMA T (^{o}F)

1200 UTC 17 March 2008

- Cold boundary not evident in parallel RTMA



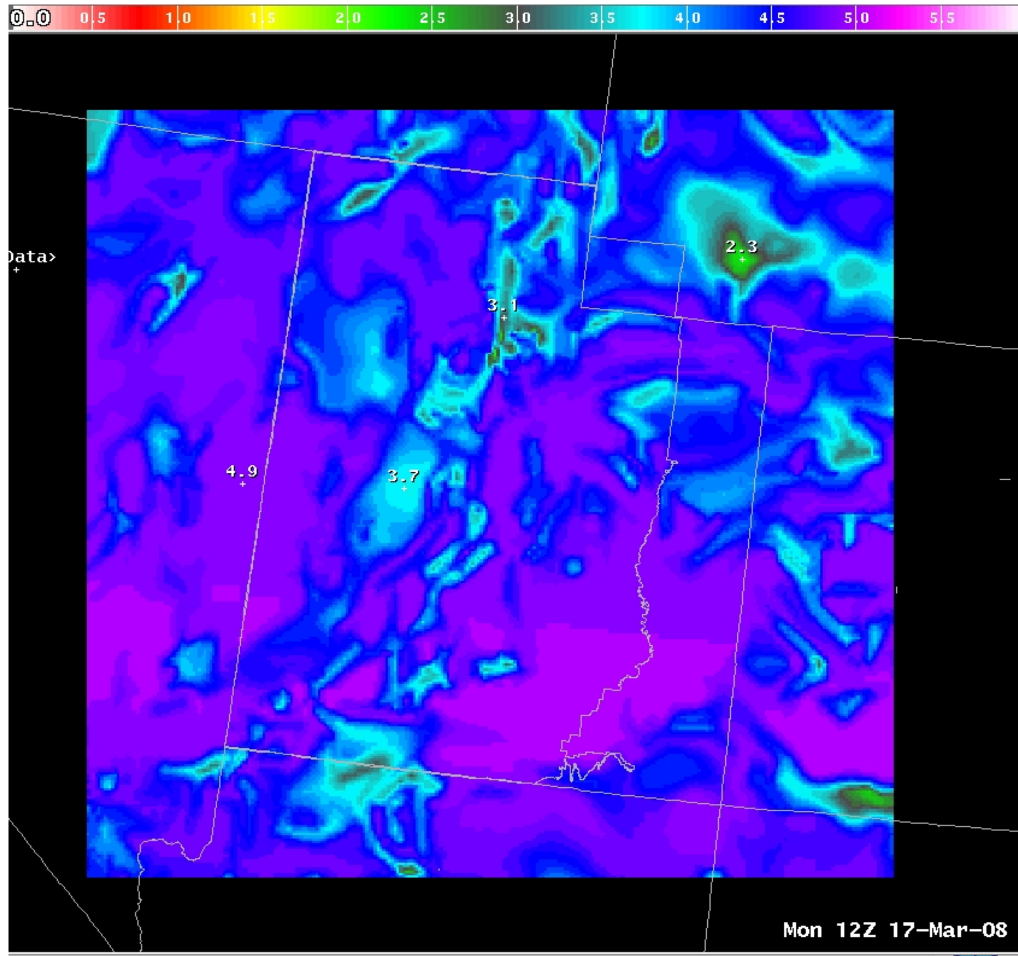
RTMA T (^{o}F) Uncertainty

1200 UTC 17 March 2008

- Uncertainty fields are a work in progress

- Formal cross-validation study needed to determine the magnitude of the analysis errors

- Goal: lower values in data dense areas/valleys, higher values in data sparse areas/mountains



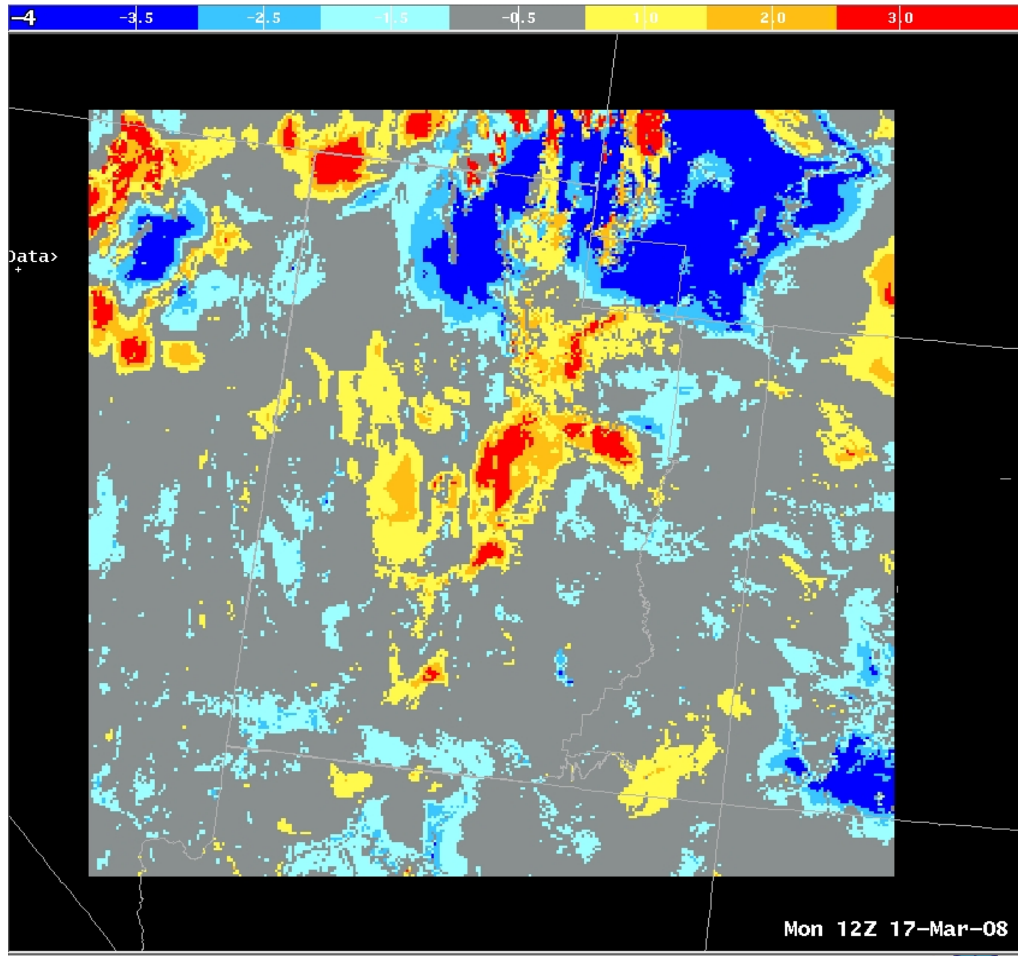
Uncertainty Verification T (^oF)

1200 UTC 17 March 2008

- Gray areas = good forecasts (forecast is within bounds of analysis uncertainty)

- GFS40 too cold north of Salt Lake City and in SW Wyoming

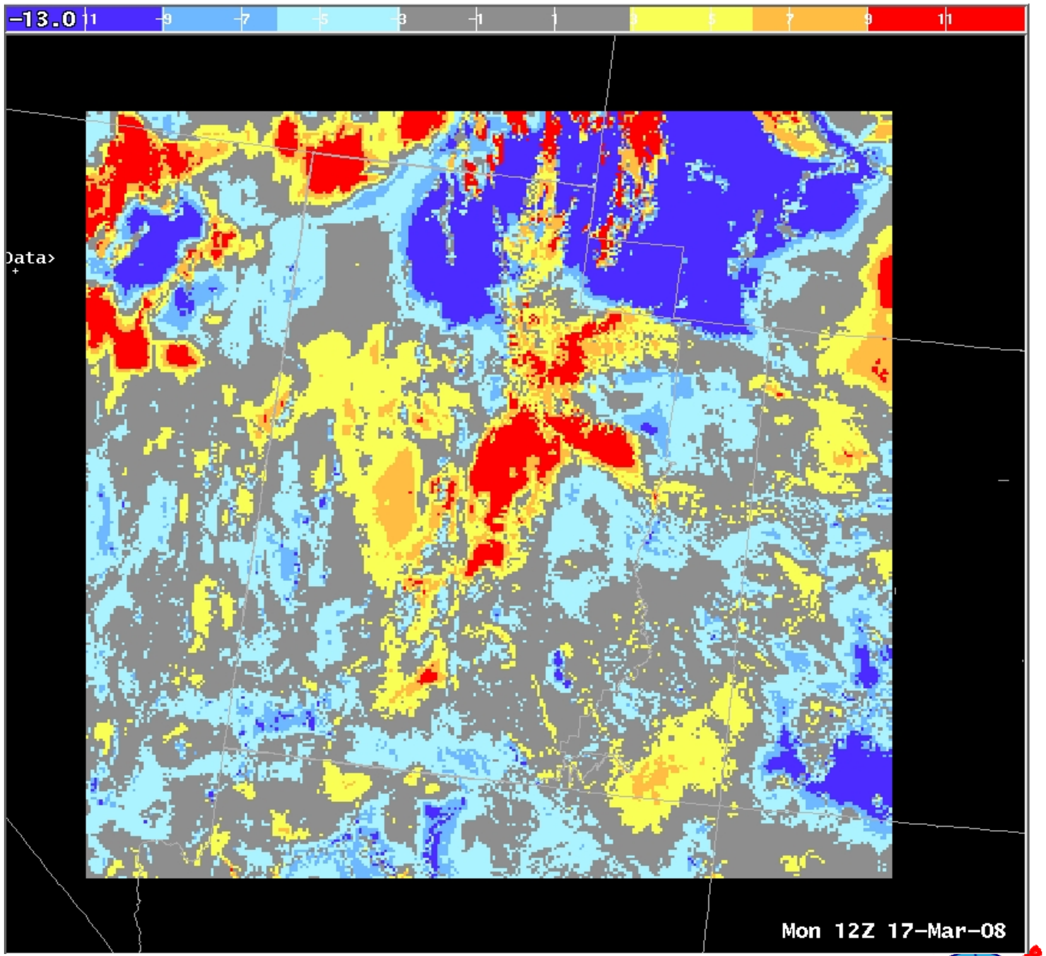
- GFS too warm in other regions (shaded in red)



Bias T (^oF)

1200 UTC 17 March 2008

- Similar error pattern
- No dependence on analysis uncertainty
- Forecasters penalized more in areas that the analysis struggles



Summary

- Objective analyses = balancing act:
  - Specificity of observations
  - Spatial and temporal continuity
- It is important that we understand ob & background errors
- RTMA effort underway…
- Uncertainty verification – an idea for grid-based verification studies – needs work

References

Benjamin, S., J. M. Brown, G. Manikin, and G. Mann, 2007: The RTMA background – hourly downscaling of RUC data to 5-km detail. Preprints, *22nd Conf. on WAF/18th Conf. on NWP*, Park City, UT, Amer. Meteor. Soc., 4A.6.

De Pondeca, M., and Coauthors, 2007: The status of the Real Time Mesoscale Analysis at NCEP. Preprints, *22nd Conf. on WAF/18th Conf. on NWP*, Park City, UT, Amer. Meteor. Soc., 4A.5.

Horel, J., and B. Colman, 2005: Real-time and retrospective mesoscale objective analyses. *Bull. Amer. Meteor. Soc*., **86**, 1477-1480.

Kahler, C. M., and D. T. Myrick, 2007: Evaluation of the Real-Time Mesoscale Analysis (RTMA) over complex terrain. Preprints, *22nd Conf. on WAF/18th Conf. on NWP*, Park City, UT., Amer. Meteor. Soc., P1.36.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge, 341 pp.

Background Downscaling

Benjamin et al. (2007)

- CONUS RTMA background = 1-h forecast from the NCEP-operational 13-km RUC downscaled to the 5-km NDFD terrain

    1. Horizontal - bilinear interpolation
    2. Vertical interpolation – varies by variable, for temperature it is based on near-surface stability and moisture from the RUC native data used to adjust to the RTMA 5-km terrain
        – If RTMA terrain lower than RUC, then local RUC lapse rate used (between dry adiabatic and isothermal)
        – If RTMA terrain higher than RUC, interpolated between native RUC vertical levels, but shallow, surface-based inversions maintained
    3. Coastline sharpening