MET and MesoVICT - Tools and Data for the Application and Testing of Established and New Spatial Verification Methods
Part 1: MET

Barbara Brown (bgb@ucar.edu)

National Center for Atmospheric Research

Boulder, Colorado USA

International Verification Workshop Online

16 November 2020

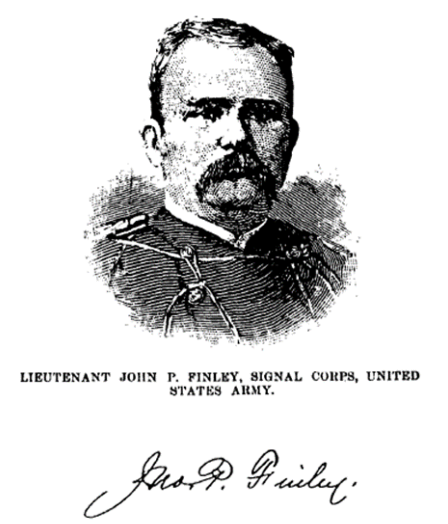
Goals

- Historical basis for modern verification (and MET): *We are part of an evolution that is accelerating*
  - Late 19th and early 20th century
  - Later 20th century
  - Development and testing of spatial methods
- Brief intro to MET and METplus
  - MET concept, history, and development
  - Spatial methods in MET
- Short intro to the ICP, MesoVICT, and Manfred's presentation

Early history: Finley, Gilbert, Pierce, Heidke

Finley tornado verification (1884)

Subjected to scrutiny by mathematicians and scientists due to use of percent correct to verify tornado predictions: *The Famous Finley affair*[1]



Finley

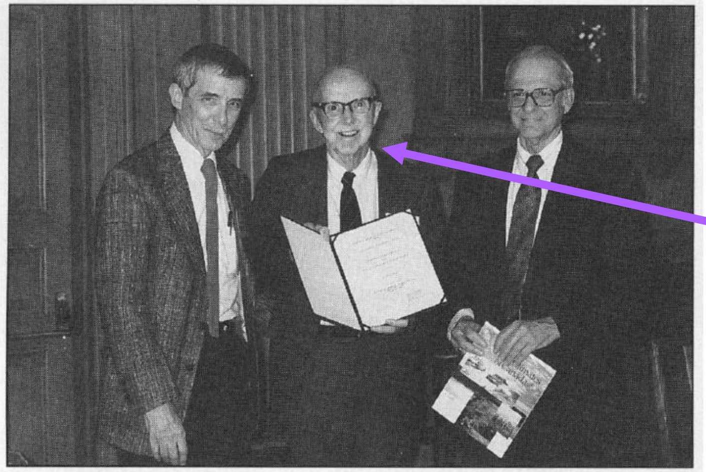
- *GK Gilbert, CS Pierce, and others suggested better approaches (e.g., Gilbert score, aka ETS; Pierce score, aka Hanssen-Kuipers)*
- *Doolittle (1888), Heidke (1926) and others suggested additional contingency table scores*

Gilbert

Pierce

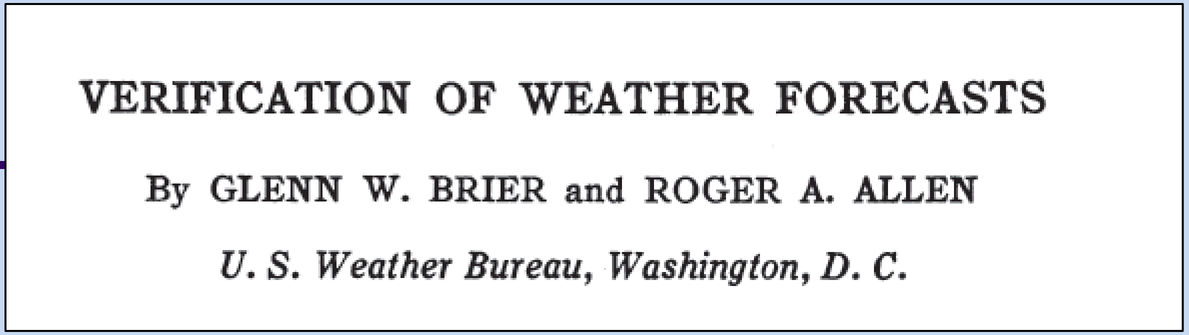
[1]See Murphy 1996 ("The Finley Affair: A Signal Event in the History of Forecast Verification" *Weather and Forecasting*, **11**)

Glenn Brier and others



Left to right: Allan Murphy (chair, SC/IMSC); Glenn Brier (recipient, IMSC outstanding achievement award); and Harry Glahn (speaker, IMSC award).

Glenn Brier



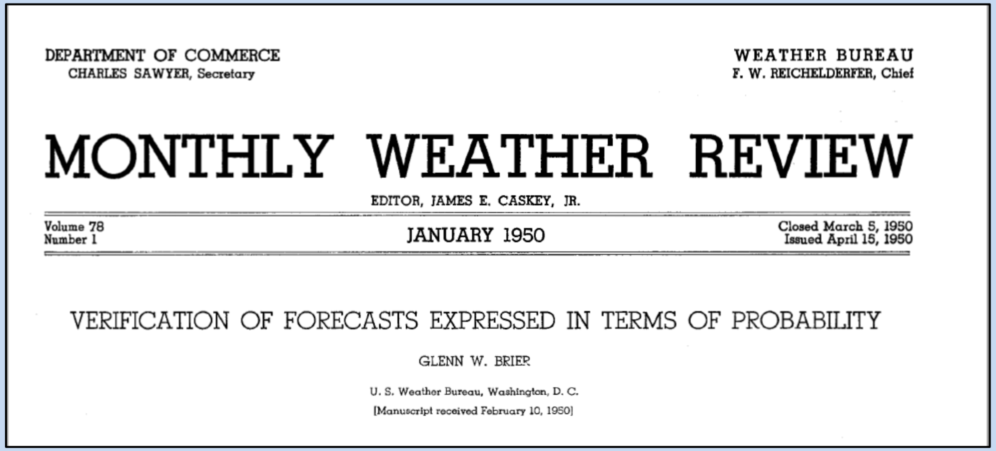
VERIFICATION OF WEATHER FORECASTS

By GLENN W. BRIER and ROGER A. ALLEN

U. S. Weather Bureau, Washington, D. C.

Originated purposes of verification (economic, administrative, scientific)

In "*Compendium of Meteorology*", 1951, American Meteorological Society, Boston, MA



DEPARTMENT OF COMMERCE
CHARLES SAWYER, Secretary

WEATHER BUREAU
F. W. REICHELDERFER, Chief

MONTHLY WEATHER REVIEW

EDITOR, JAMES E. CASKEY, JR.

Volume 78
Number 1

JANUARY 1950

Closed March 5, 1950
Issued April 15, 1950

VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY

GLENN W. BRIER

U. S. Weather Bureau, Washington, D. C.

[Manuscript received February 10, 1950]

Origination of Brier Score (1950)

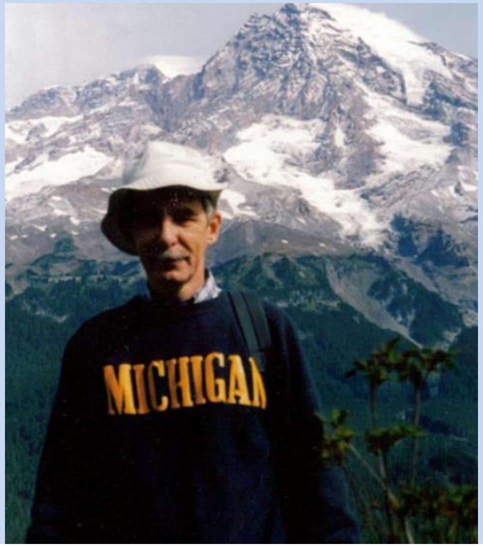
<u>Muller (1944): BAMS, verification of short-range forecasts</u>

Summarizes the international (English, German, French, Dutch, Russian, Danish) literature on verification, with much of it dating back to the Finley-affair period (55 articles total).

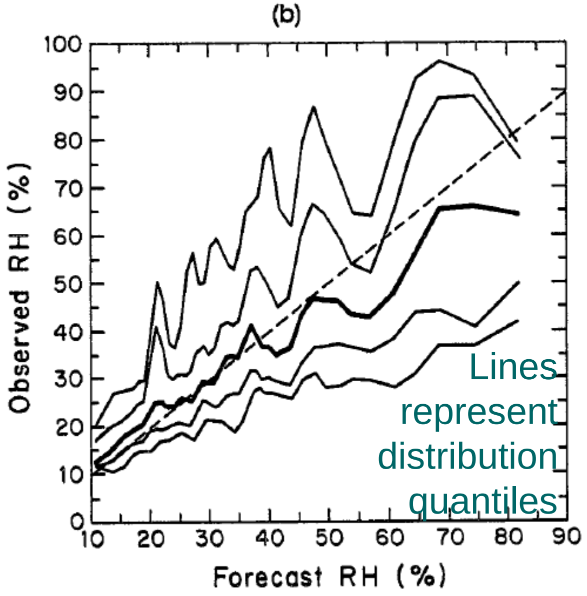
Two major groups of methods:

- Evaluation of hits from comparison to obs
- Comparisons to random or climatic forecasts

Allan Murphy



"***Forecast quality is inherently multifaceted in nature… however, forecast verification has tended to focus on one or two aspects of overall forecasting performance such as accuracy and skill***."



Example: Distribution-Oriented Approach (Fire weather RH forecasts)

Brown et al. 1987 (*Weather and Fcstg*)

Timeline of Murphy's work on Forecast Verification Concepts/Methods:

• Brier score and decompositions	1965-86
• Ranked Probability Score and decompositions	1969-72
• ***Generalizations***: "Quadratic" scoring rules; MSE, correlation, coefficient of determination	1978-96
• Geometrical interpretations	1975, 1986
• "Proper" skill scores	1972-73
• "Objective" and "Subjective" forecasts (comparisons, relationships)	1984-88
• Trends in verification statistics	1986
• ***Framework*** for forecast verification, with extensions	1987-95
• Relationships between ***quality***, ***value***, and ***consistency***	1988, 1993
• Sufficiency	1988-97
• ***Dimensionality*** of verification	1991
• ***Distributions-oriented*** (diagnostic) verification	1989-97
• Equitable skill scores	1992
• History	1996
• Use of statistical models	1997

Other contributions/activities in the 1990s

- Int'l Statistical Climatology and AMS Prob and Stat meetings, with focus on verification
- Tutorial on forecast verification
  - 1998, AMS Annual Meeting
- Methods for evaluating contingency tables (Doswell, Brooks, Marzban)
- Thoughts about uncertainty associated with verification measures (e.g., Hamill)
- Canadian verification methods document (Stanski et al)
- 1st edition of Wilks book
- **Initial ideas around spatial verification, including**
  - P. Neilley paper on concepts for object-based verification (1993, 13th AMS WAF Conference)
  - Verification of microburst "objects"

And Much More!

The 21st century: Motivation to develop spatial methods

- *The conundrum*: As models moved to higher resolution, forecast performance did not appear to improve
- Example:

  Trough position predictions located east of observed

  MAE values

  - 4.19 for 36-km model
  - 4.82 for 12-km
  - 5.25 for 4-km

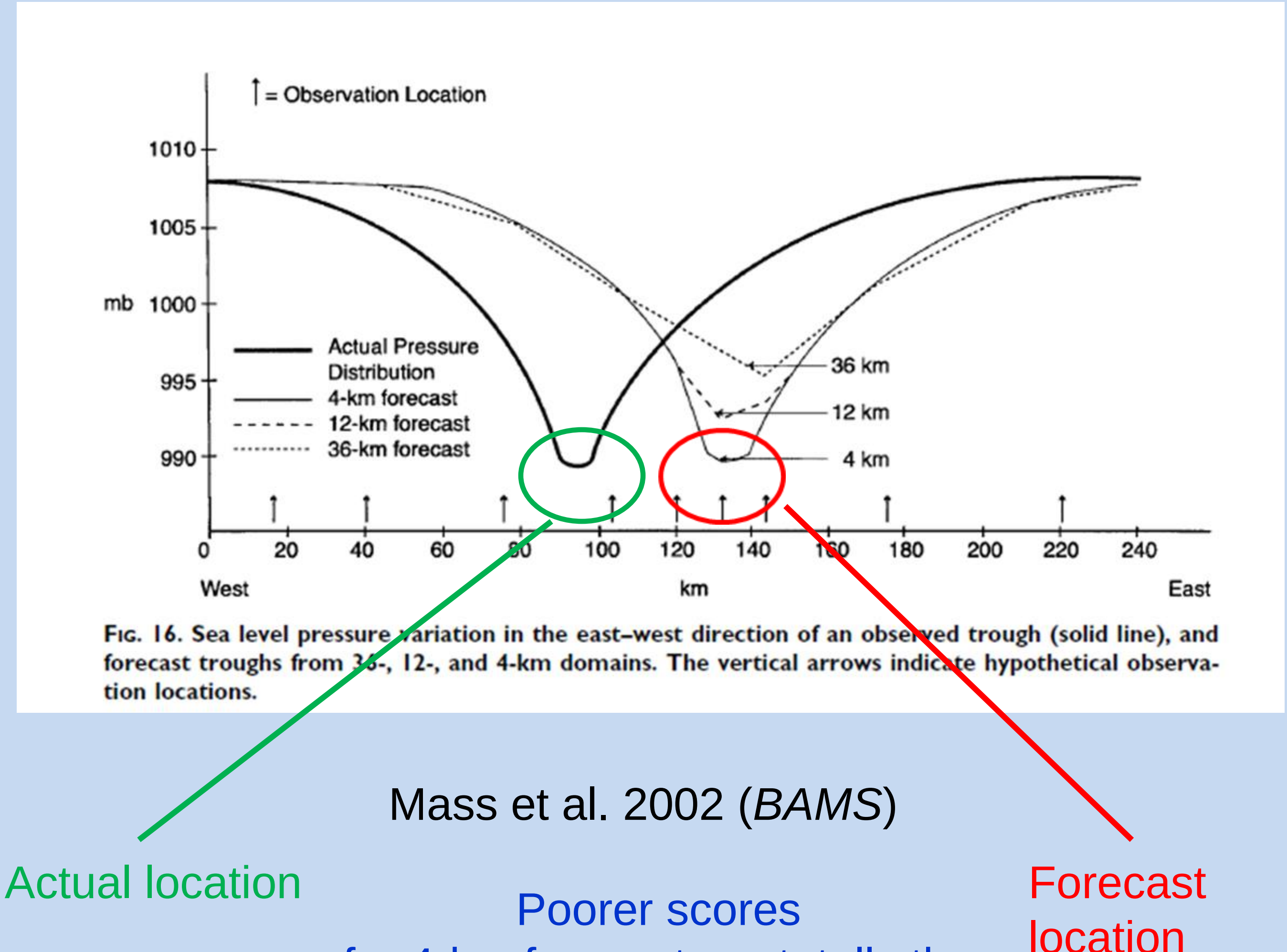


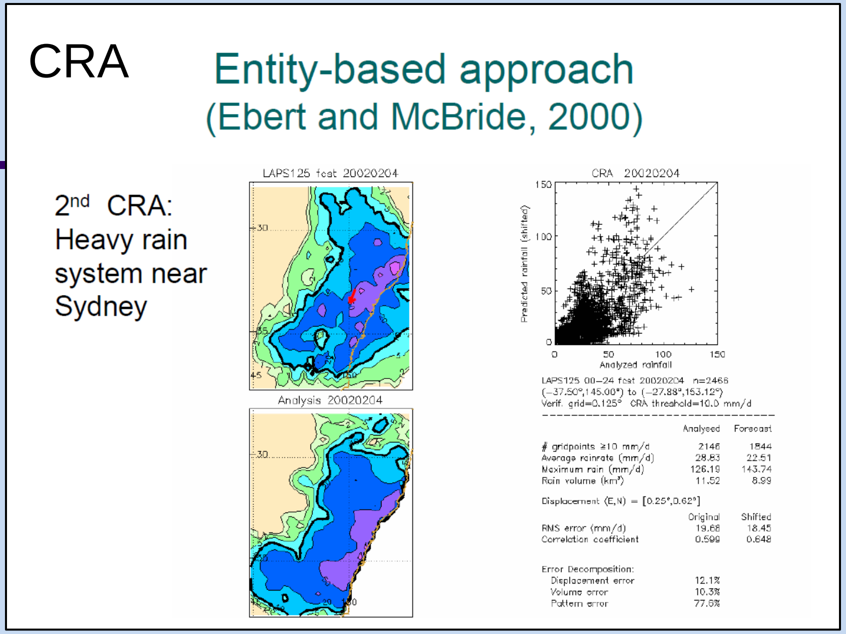
FIG. 16. Sea level pressure variation in the east–west direction of an observed trough (solid line), and forecast troughs from 36-, 12-, and 4-km domains. The vertical arrows indicate hypothetical observation locations.

Mass et al. 2002 (*BAMS*)

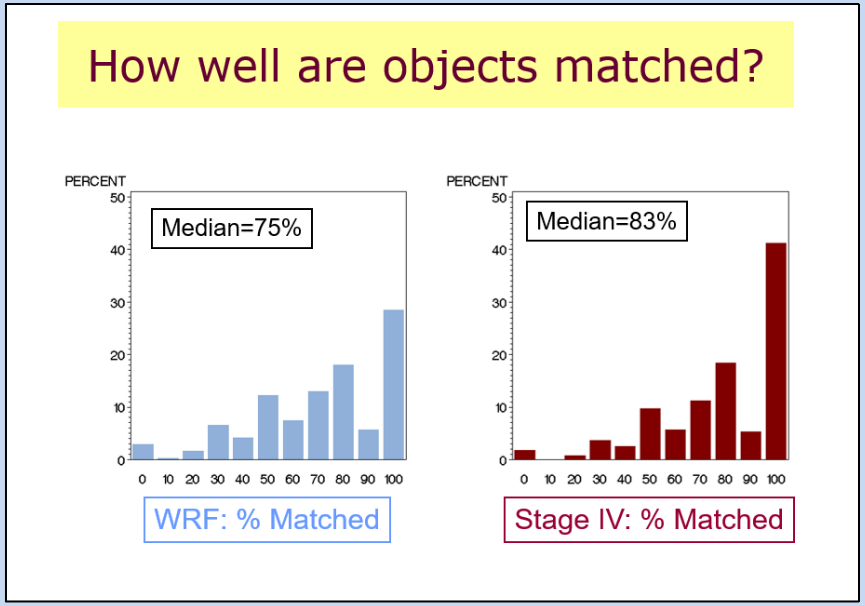
Poorer scores for 4-km forecast are totally the result of location error

Many spatial verification methods were developed in this period, including…

- Feature-based approaches
  - Contiguous rain area (CRA; Ebert and McBride 2000)
  - Method for Object-based Diagnostic Evaluation (MODE; Davis et al. 2006a,b)
  - Object-based approaches developed by Baldwin and others
- Scale separation
  - Ex: Intensity/Scale approach (Casati et al. 2004)
- Distance metrics (Gilleland and others)
- Neighborhood
  - Ex: Fractions Skill Score (FSS)
- Field deformation



MODE



Note: Bibliography available at https://ral.ucar.edu/projects/icp/references.html

Initiation of Model Evaluation Tools (MET)

Developmental Testbed Center (DTC; https://www.dtcenter.org/) created in 2003 to serve as a bridge between research and operations to facilitate NWP development objectives

- Partners included NCAR, NOAA (National Oceanic and Atmospheric Administration), and US Air Force (USAF)
- Activities include *model testing and evaluation*

In 2006, USAF requested development of *"A world class, state of the art verification system for evaluating high-resolution forecast systems… The package will be made available to all WRF [model] users."*

Why 2006?

- Maturing field of verification – with an emphasis on *spatial methods*!
- Increasing user needs: (1) High-resolution models, (2) Need for more accountability, (3) Lack of adequate tools for the whole community

MET applications and generality have grown considerably since its initiation – to include new NWP models and applications and to additional users around the world

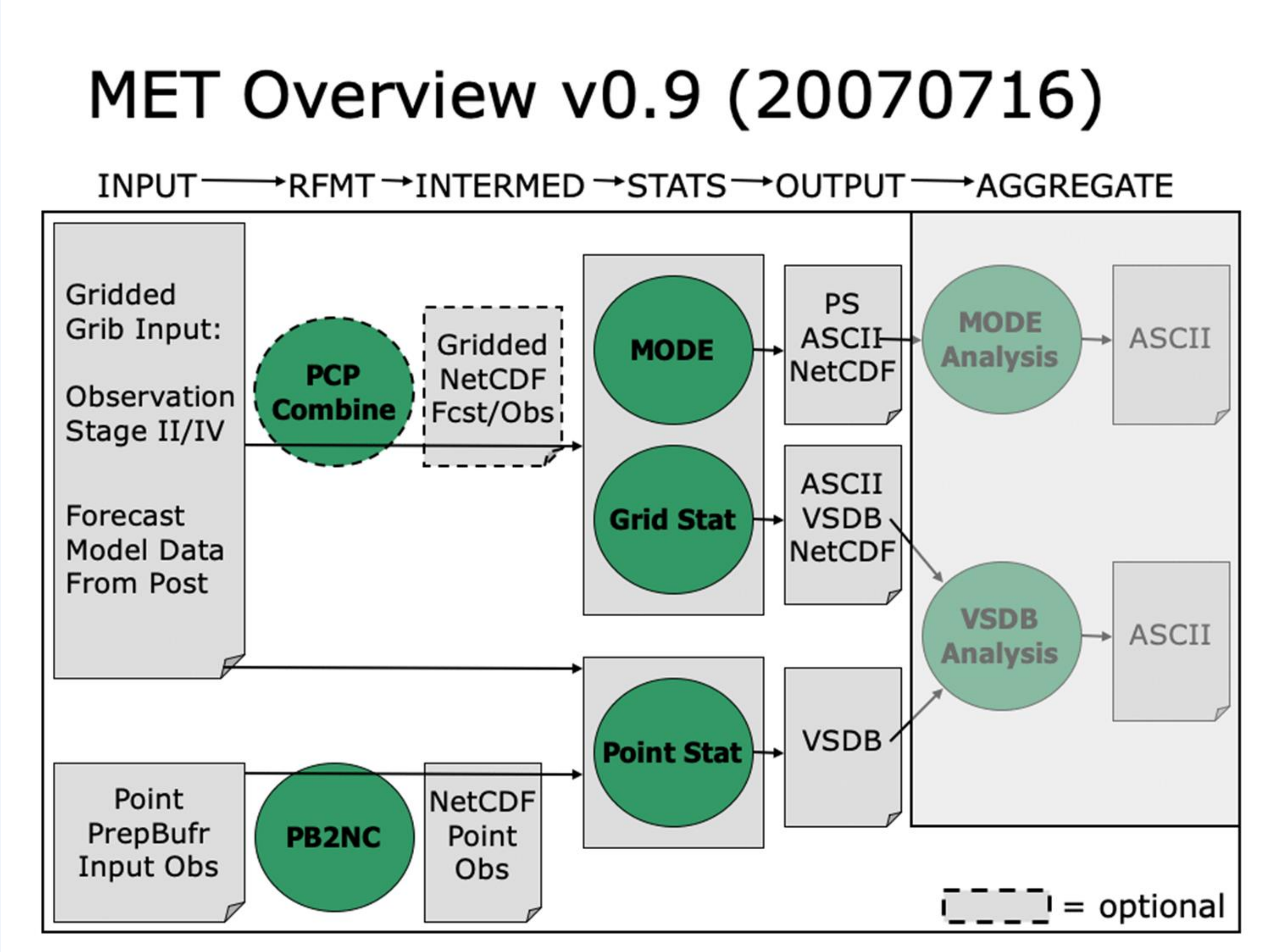
Note: WRF is the community Weather Research and Forecasting model

MET initiation and development

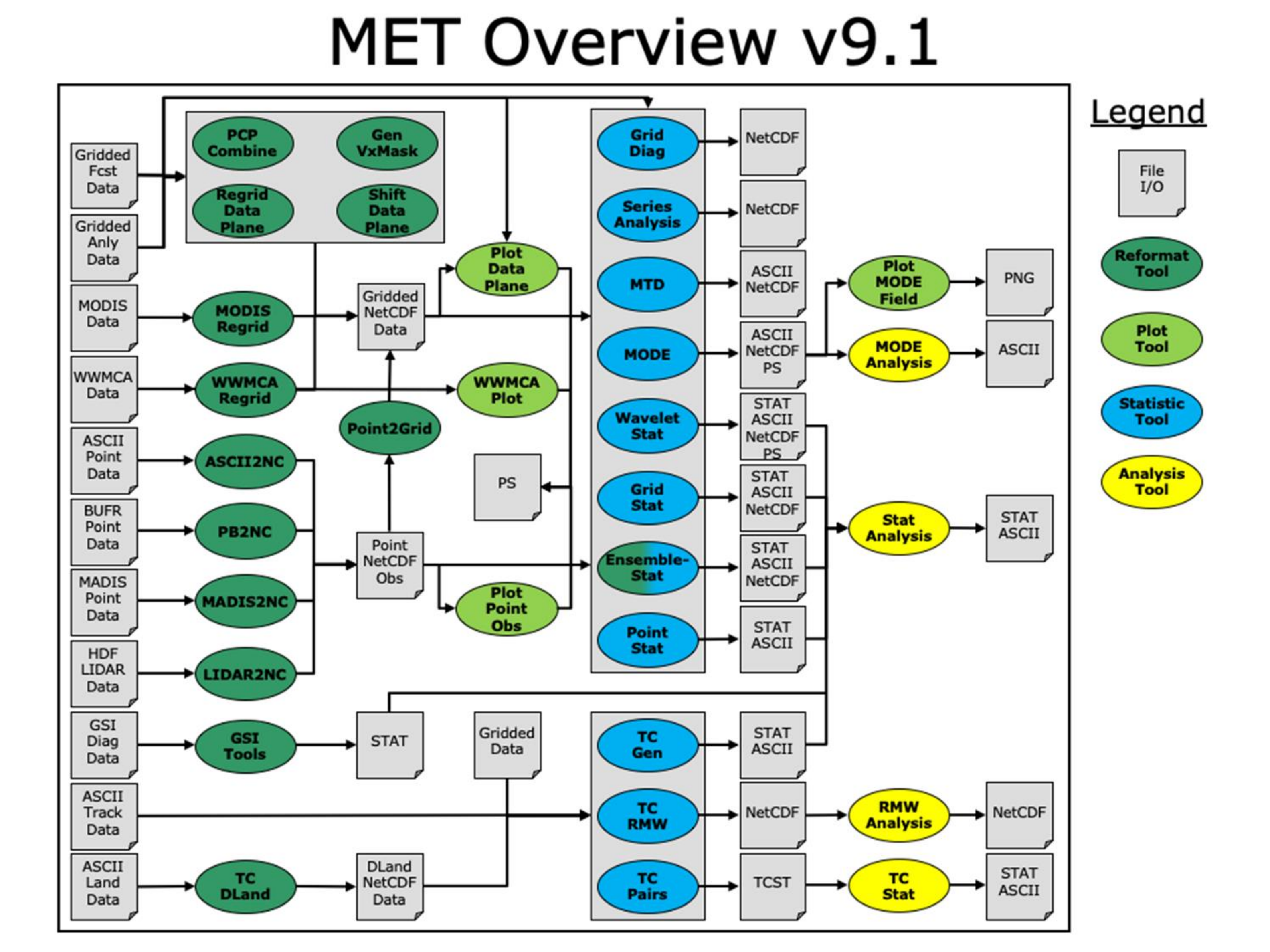
Contributing factor: **Community “creation” and engagement**

- Sydney Olympics!
- Creation of WMO’s Joint Working Group on Forecast Verification Research (JWGFVR)
- Workshops on verification
  - WMO-Sponsored
  - Sponsored by NCAR and the DTC
- *Very* active international verification community

MET then…



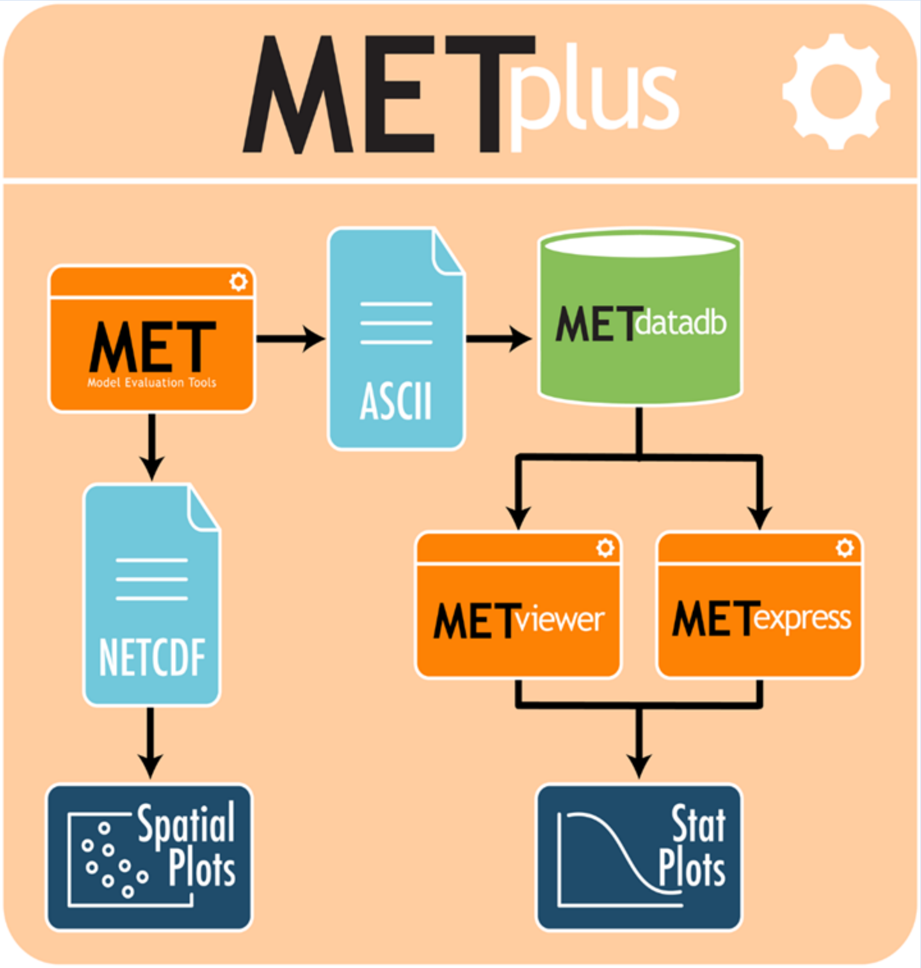
- Limited options for
  - Observations
  - Models
  - Reformatting
- Verification measures include
  - Basic grid- and point-based methods
  - One spatial method (MODE)
- Freely available and supported to all users



MET includes:

- Extensive data input and re-formatting options
- Large number of verification methods
  - Additional spatial methods
  - Ensemble methods
  - Methods for tropical cyclones, ensembles etc.
- Freely available and supported to all users
- Expanding international user-base and collaborations

METplus

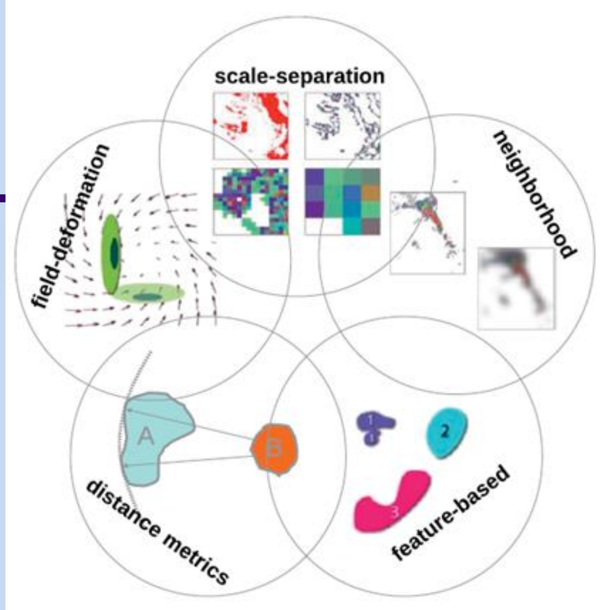


METplus includes MET as well as

- Database system
- METviewer
  - View and analyze MET output (extensive graphical options)
  - Create plots

See talk by Tara Jensen (17 November, 1500 UTC)

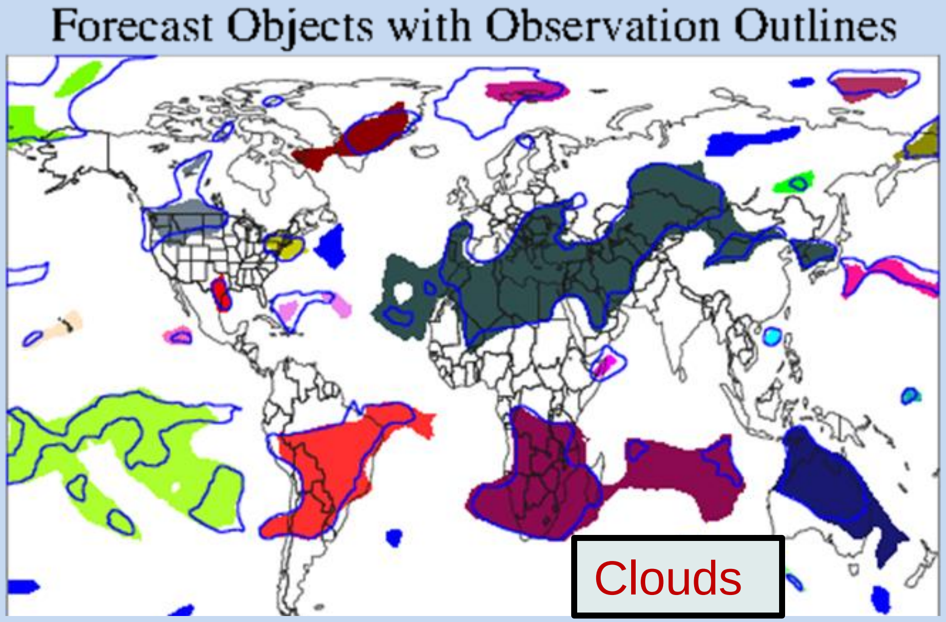
Spatial methods in MET



- MODE and MODE-TD (Feature-based)
- FSS (Neighborhood)
- HiRA (Neighborhood)
- Wavelet-stat (Scale separation)
- Distance metrics (newest)

These methods/tools

- Include 4 out of 5 categories of spatial approaches
- Have been applied to many different phenomena (weather, space-weather, climate, algae, sea ice, etc.)



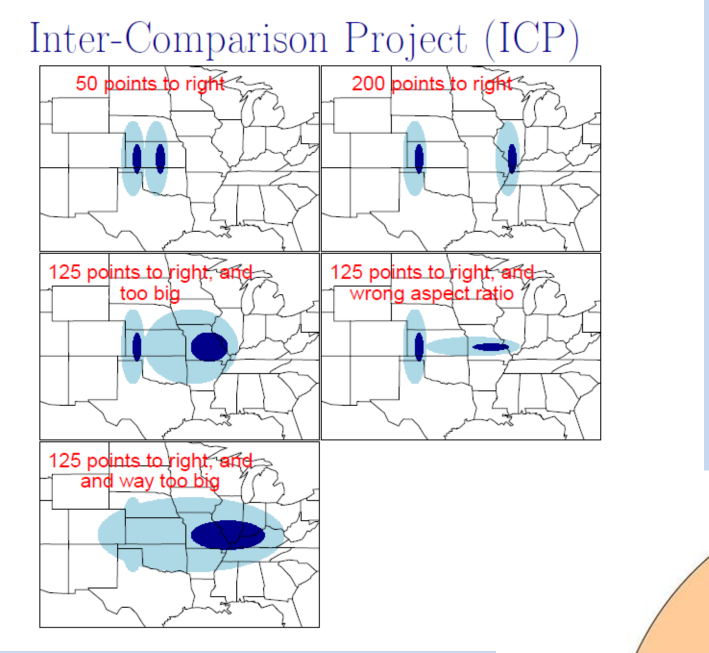
Many methods implemented through the DTC visitor program (https://www.dtcenter.org/visitor-program)

Verification method intercomparison projects: ICP and MesoVICT (Mesoscale Verification in Complex Terrain)

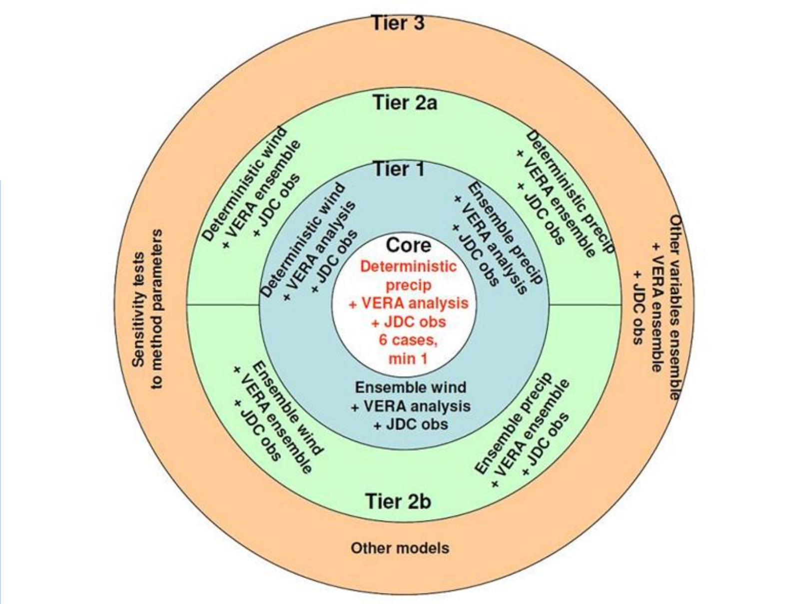
Many of MET's spatial methods included in MET were evaluated in ICP and/or MesoVICT

- **ICP** (Intercomparison Project) focused on forecasts in the central US plains
  - Forecasts compared to gridded observations
  - Resulted in multiple publications
- **MesoVICT** (MesoScale Verification in Complex Terrain) has focused on mountainous region in Europe
- Overarching goals: Provide insights about capabilities of new spatial methods

More about these projects momentarily from Manfred!



Artificial cases examined in first ICP



MesoVICT concept diagram

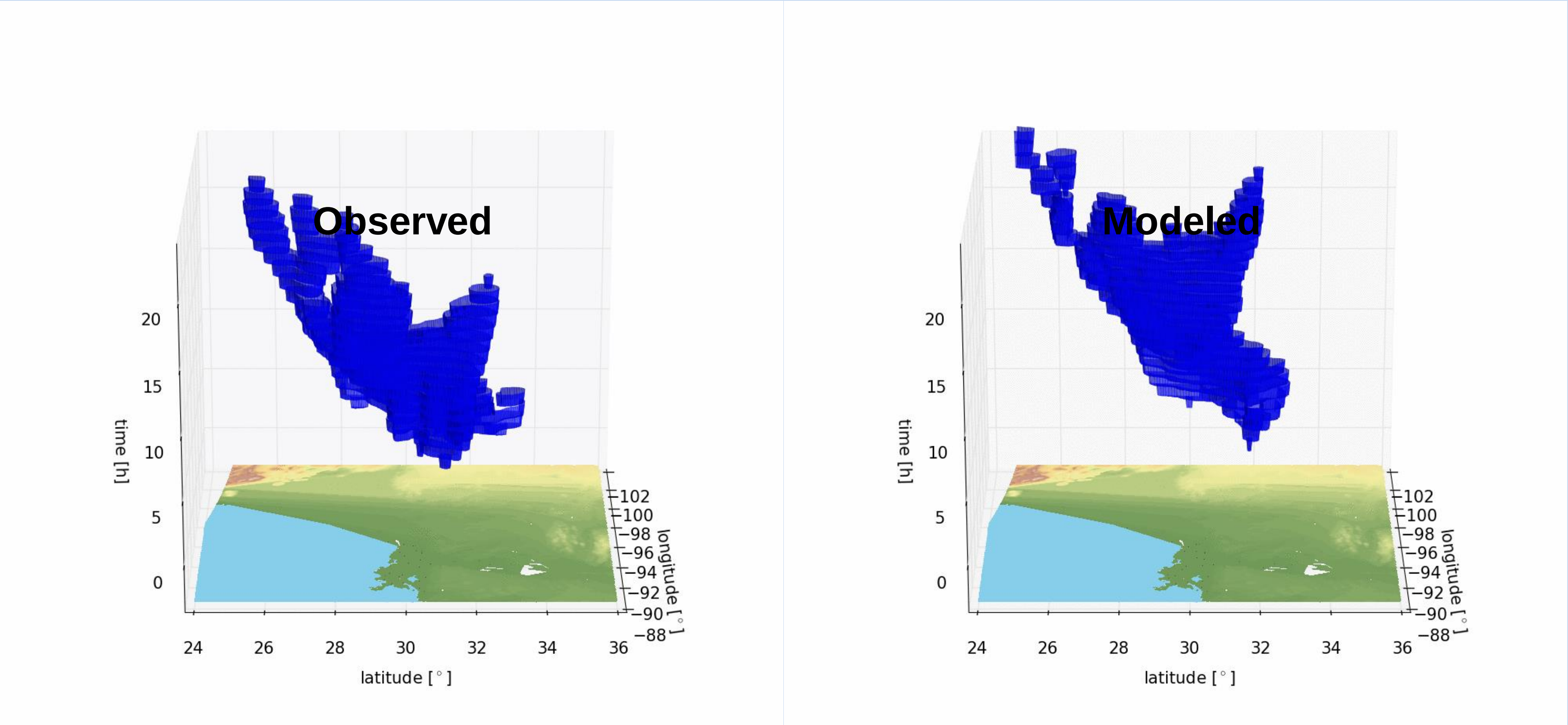
Conclusions

- MET initiation was the culmination of multiple factors, including
  - Historical development of methods and tools
  - Modern capabilities (’90s and early 2000s) and recognition of importance of verification for informing users (including scientists) and improving forecasts
  - Development, testing, and implementation of spatial methods
- MET and Spatial Method development resulted from
  - Maturity of verification knowledge and expertise
  - Needs of modeling and user communities

Brown, B., T. Jensen, and Co-authors, 2020: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society*, in press, DOI 10.1175/BAMS-D-19-0093.1

With gratitude to…

- The WMO and in particular the World Weather Research Project (WWRP) for encouragement and support of verification research efforts for the last two decades;
- NCAR and the DTC for their support of MET development;
- Tara Jensen and the entire MET team and MET contributors for their incredible efforts in developing, testing, and supporting MET;
- All of the participants in the ICP and MesoVICT for their willingness to actively engage in these efforts!



Application of MODE-TD to WRF prediction of an MCS in 2007
(Credit: A. Prein, NCAR)

MODE-TD allows evaluation of storm initiation, movement, velocity, timing errors, storm volume, storm velocity, initiation, decay, etc.