Forecast Consistency Verification for Climate Models

Tressa L. Fowler

Barbara G. Brown, Eric Gilleland, Randy Bullock, John Halley Gotway, Caspar Amman, Tara Jensen, Paul Kucera

Research Applications Laboratory, NCAR

tressa@ucar.edu

Two separate talks

- Evaluation of forecast consistency via forecast revisions.
  - What is a revision? Why is it interesting?
  - Simple – Wind speed
  - Complex - Tropical Cyclones

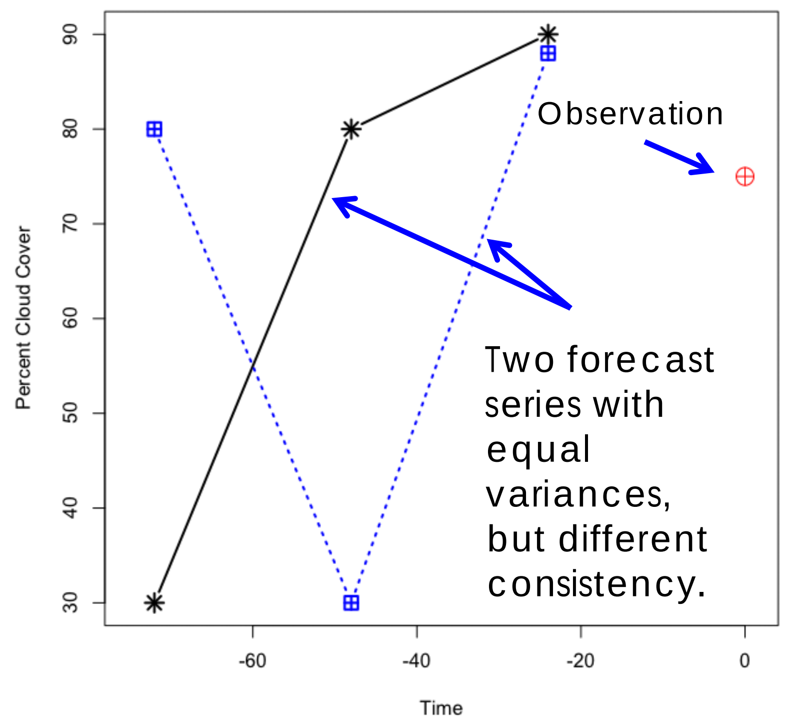
- Adaptation of verification tools and metrics to climate models with applications to hydrologic decision making.
  - Understanding our collaborator's needs.
  - Preliminary tests of verification tools on climate models.

Consistency of Updating Forecasts

Forecasts of a single event (same valid time) with decreasing lead times.

Friday Evening reservation at Flagstaff House
Will we see the stars?

72 h lead (Tuesday) forecast for Friday : 30% cloud cover
48 h lead (Wed): 80%
24 h lead (Thurs): 90%



Revisions

- Revisions are the changes (or updates) in the forecast for the same event.
- In other words, the valid time is the same, but the lead time decreases.
- Two important questions about revisions:
  - Are they large?
  - Are they consistent or random?
- Note: These two questions do not involve the observation!

Magnitude of Revisions

- Simple to use standard statistics like mean, median, standard deviation, box plots.
- Indices for specific weather variables:
  - Ehret, U., 2010: Convergence Index: a new performance measure for the temporal stability of operational rainfall forecasts. *Meteorologische Zeitschrift* **19**, pp. 441-451.
  - Lashley, S., A. Lammers, L. Fisher, R. Simpson, J. Taylor, S. Weisser, and J. Logsdon, 2008: Observing verification trends and applying a methodology to probabilistic precipitation forecasts at a National Weather Service forecast office. *19th conference on Probability and Statistics*, New Orleans, LA. American Meteorological Society.
  - Ruth, D. P., B. Glahn, V. Dagostaro, and K. Gilbert, 2009: The Performance of MOS in the Digital Age. *Weather and Forecasting*, **24**, 504-519.

Is forecast behavior through time **random** or **related**? Forecast ‘consistency’

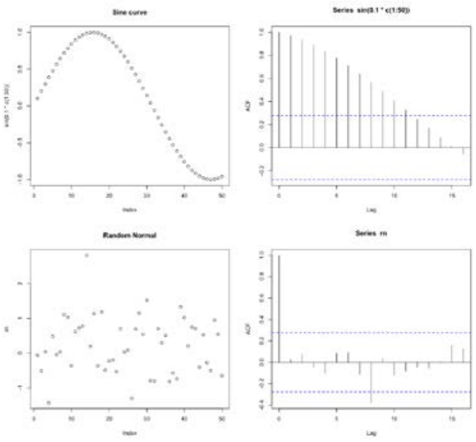
- A property of the forecasts only.
- Also called ‘jumpiness’ or ‘lack of rationality’.
- Consistency can be bad or good, depending on the user.
- Regardless, it should be measured.

- In economics, consistent forecasts are not rational.
  - Information comes in all at once, so a new forecast should incorporate all available info.
- In weather, information trickles in.
  - Maybe forecasts should change gradually, reflecting the continual update of information.
  - For numerical modeling, this may not hold.

Two tests of consistency in a series

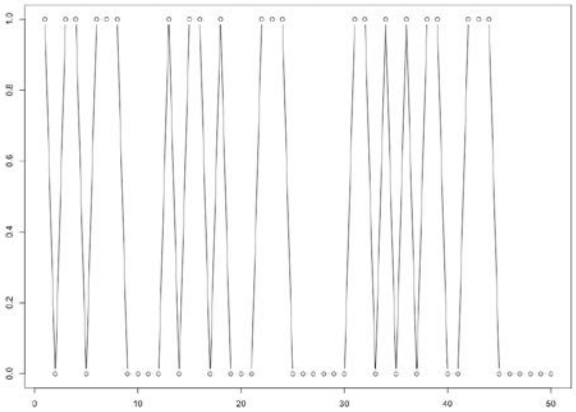
Autocorrelation

- Measures relationship of numbers separated by a specific distance in time.
- ‘Significant’ autocorrelation indicates a relationship in time.



Wald – Wolfowitz

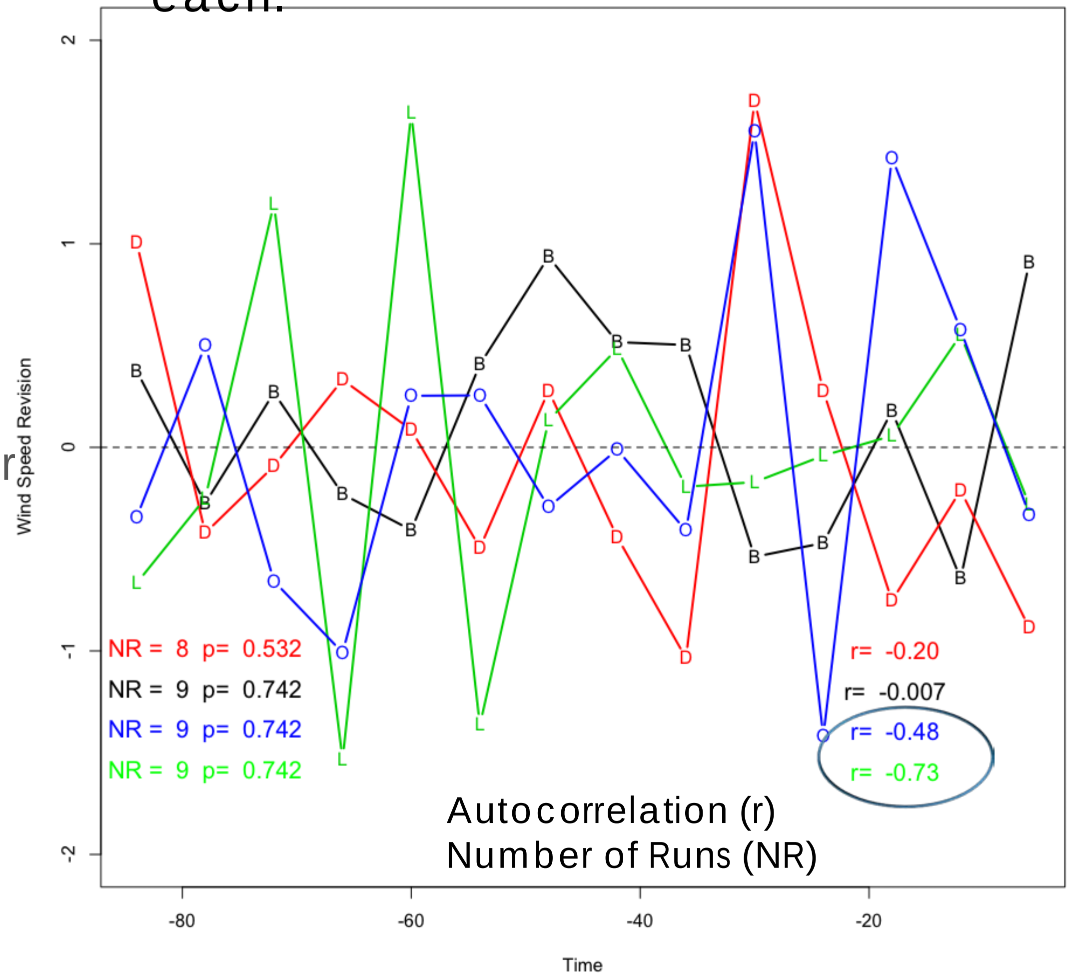
- Measures ‘runs’ above and below some reference.
- More runs than are expected by chance indicates too consistent, e.g. non-random, behavior.



Simple example

- Revision series of wind forecast for 4 locations.
- Blue and green negatively autocorrelated (switch too often).
- Others not differentiable from random, i.e. number of runs for all series and autocorrelation of black and red are not consistent.

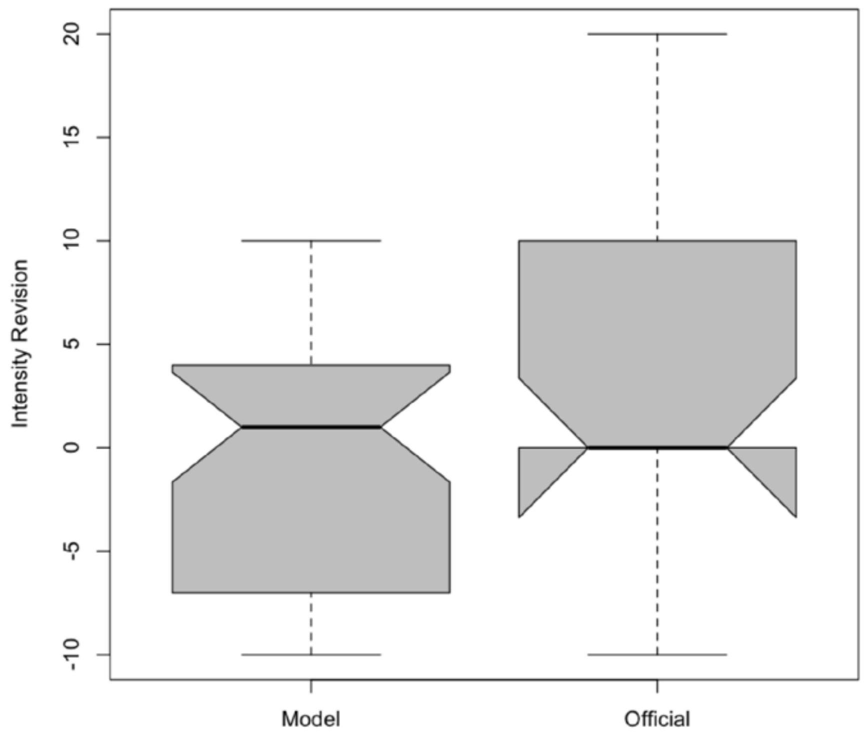
Four revision time series of 14 points each.



Consistency for tropical cyclones (TC)

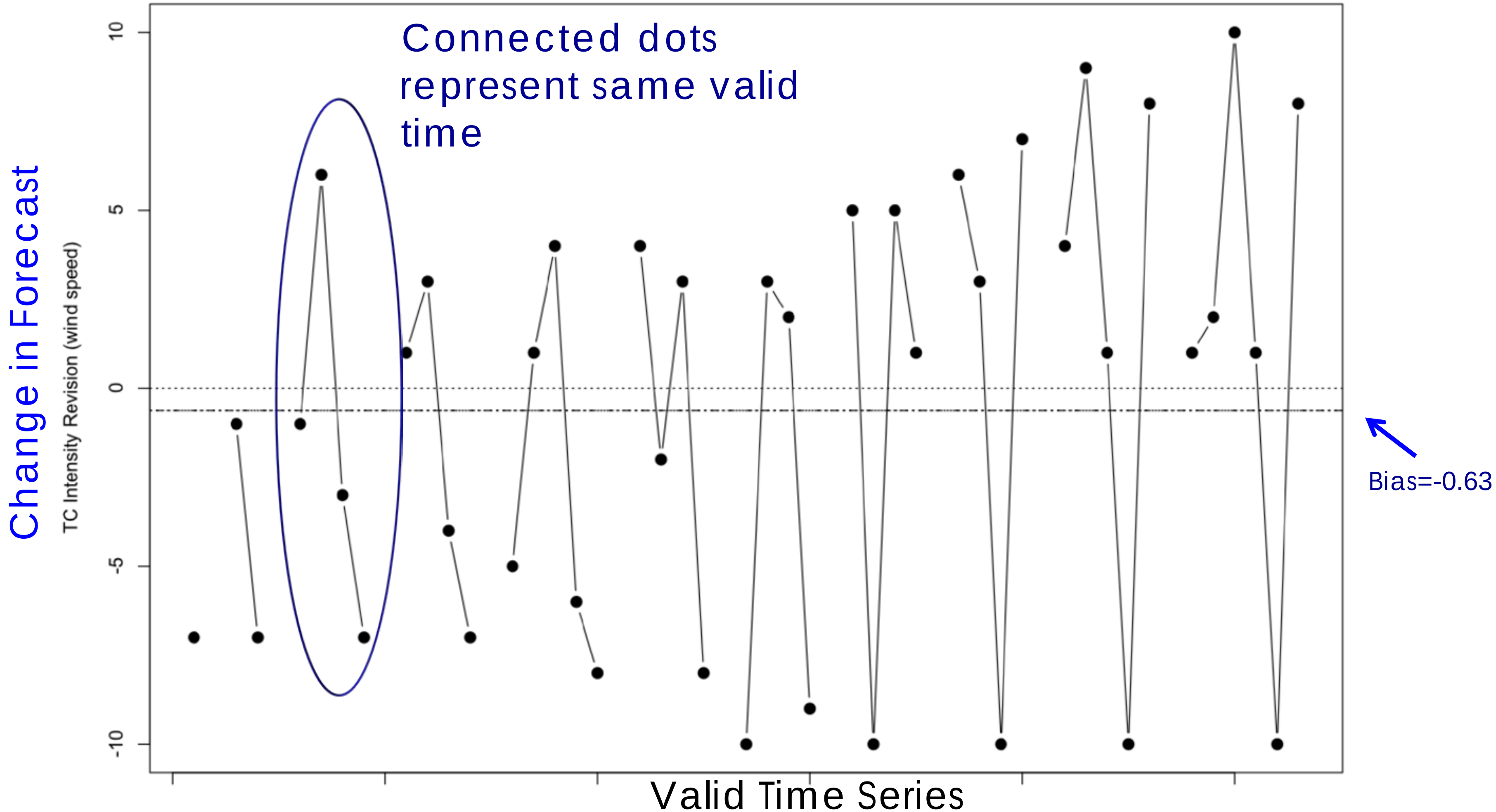
- Intensity (wind speed)
- Track (two dimensional)
  - Along track error
  - Cross track error
- We have many valid times for each storm, how to we summarize?
- Revision series can be very short and of different lengths.
- Track is a two dimensional measure, so there is not a nice time series of revision values.
- Is there a windshield wiper effect in the forecast?
- How do model consistency values compare?
- Is this a measure of (relative) uncertainty?

Magnitude of revisions



- Both model and official revisions center near 0.
- Official forecasts have a wider range of revisions.
- Model more likely to revise to lower intensities (wind speeds), while official more likely to increase forecast intensity.

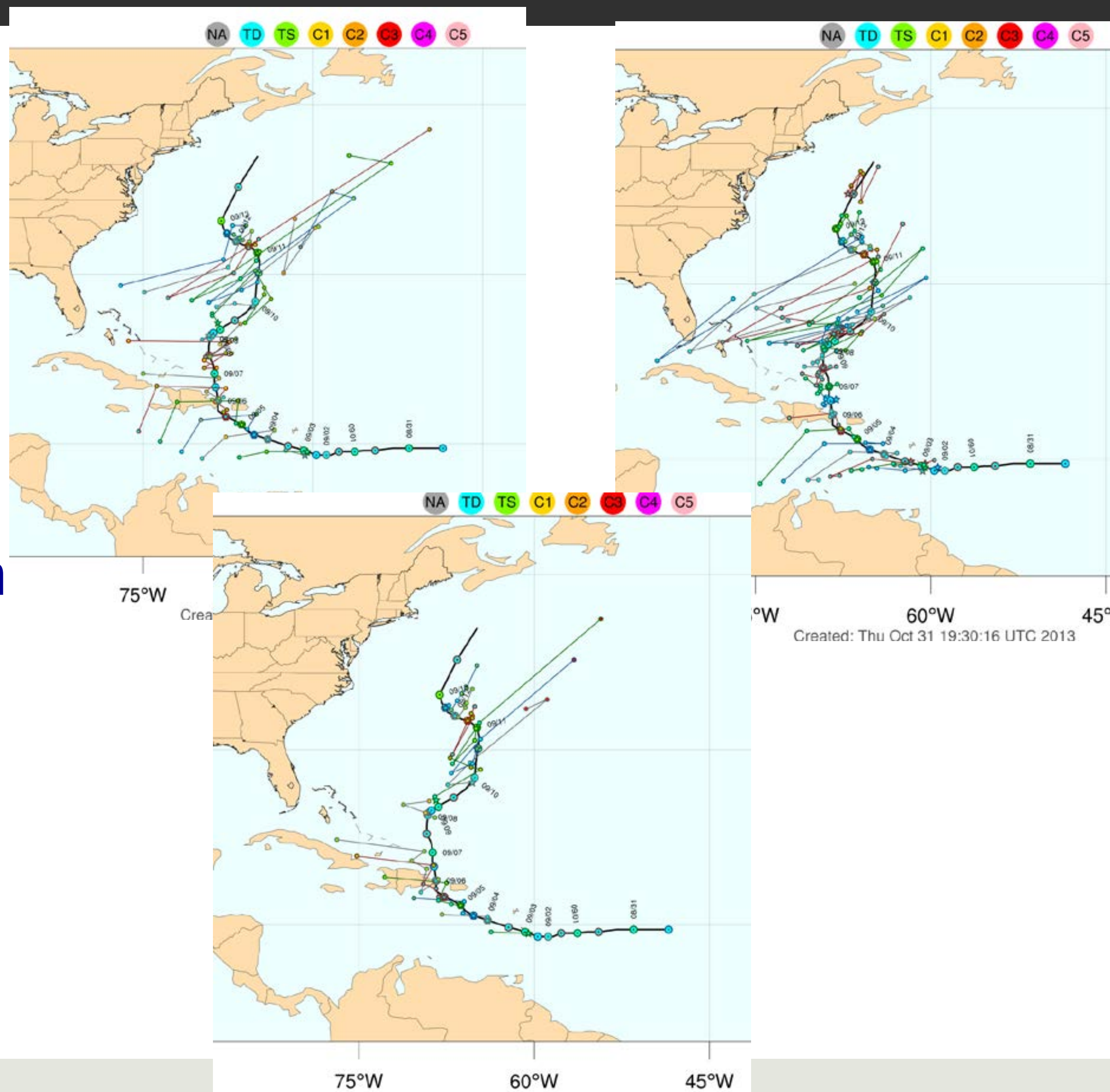
Consistency of several wind speed revision series from TC forecasts

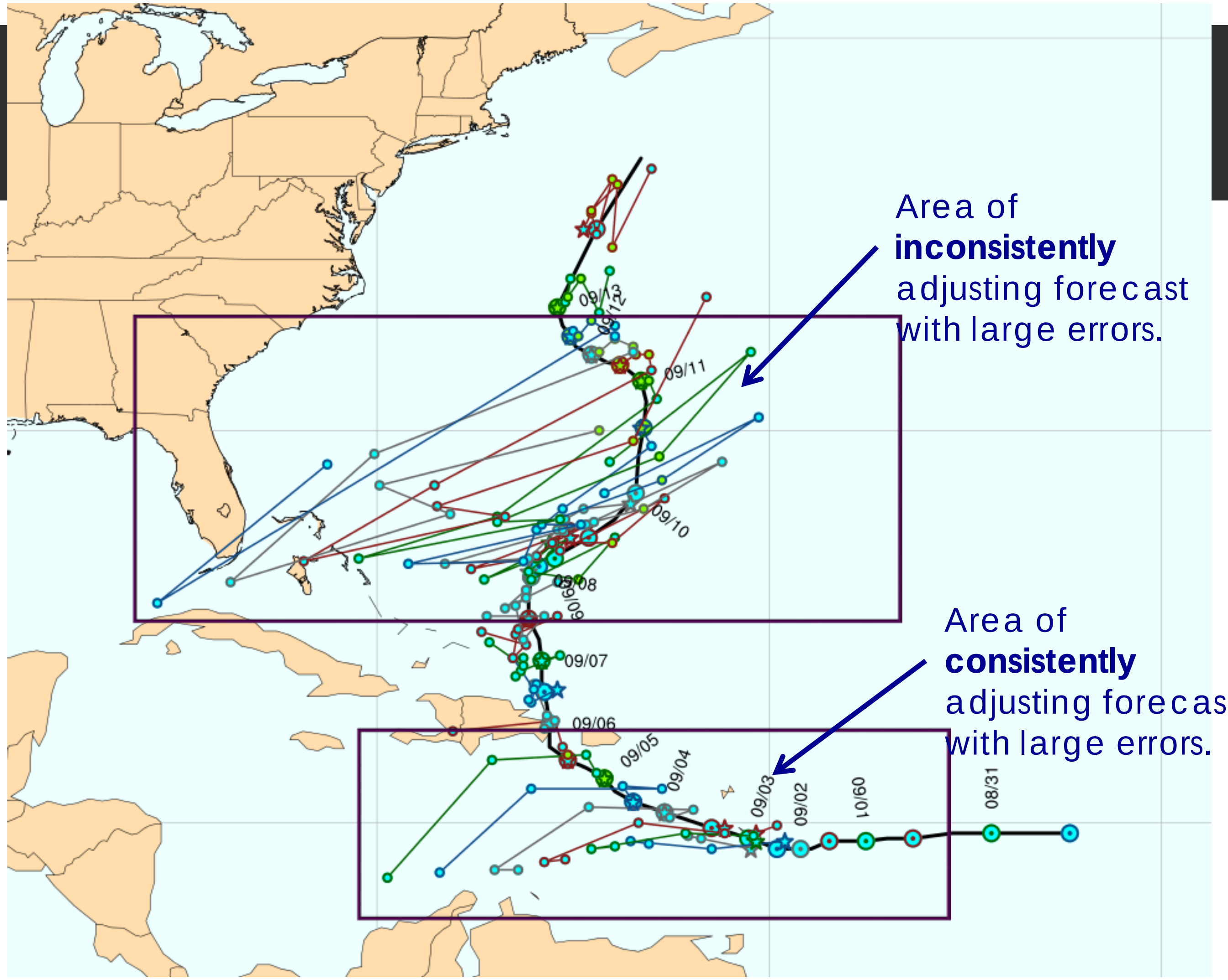


Forecast revision series - GABRIELLE

Other ways to quantify “randomness”.

- Area of revisions.
- Average path length of revisions.
- Number of ‘crossovers’.




Area of **inconsistently** adjusting forecast with large errors.
Area of **consistently** adjusting forecast with large errors.
09/13
09/12
09/11
09/10
09/09
09/08
09/07
09/06
09/05
09/04
09/03
09/02
09/01
08/31

Summary

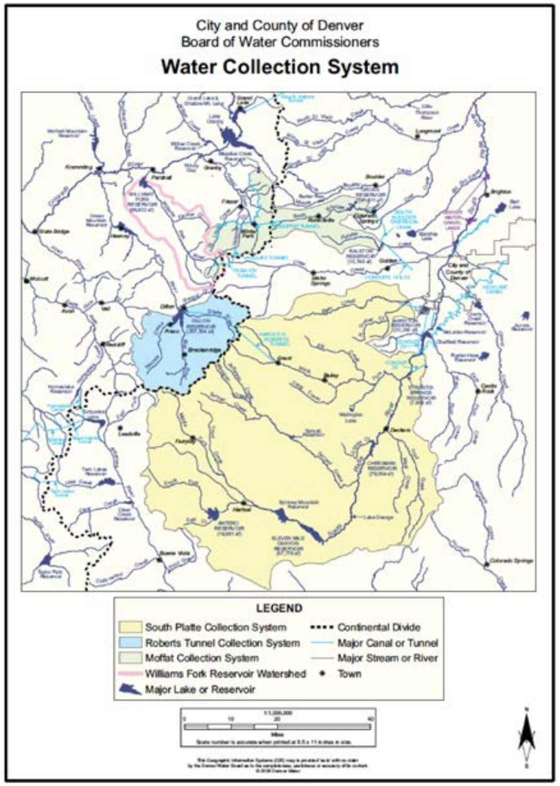
- Examination of revision series gives additional information to forecast users.
- TC forecasts are complex in format, making measuring consistency somewhat difficult. None of these measures is without issues, and there may be pitfalls that are not yet obvious.
- Comparisons between models seems more straightforward than statistical tests of random behavior.
- These measures need to be tested and refined according to users' needs.

Climate Verification

- Advanced Climate and Regional Model Validation for Societal Applications
- Collaboration between climate modelers, hydrologists, NWP verification experts and software engineers.

Climate Verification

- **1. Identify the variables and indices, based on water resource management needs, that threaten or otherwise influence decision-making, applying understanding of key processes and their spatial and temporal scales.**
- **2. Adapt and convert established quantitative weather-forecast verification tools for climate-model metrics. Accessible and transparent metrics will be the cornerstone for establishing "best practice" uses.**
- 3. Characterize changes seen in future climate projections, using the new tools to link the changes and their uncertainties to specific climate change impacts and needs.
- 4. Implement the new validation tools in the CESM diagnostics framework, where they can inform model development and enrich the model assessment through user-developed benchmarks.



Working with Denver Water

Planning for climate change, which might involve new infrastructure.

Particularly interested in 3 year or longer droughts in and near their water collection system.

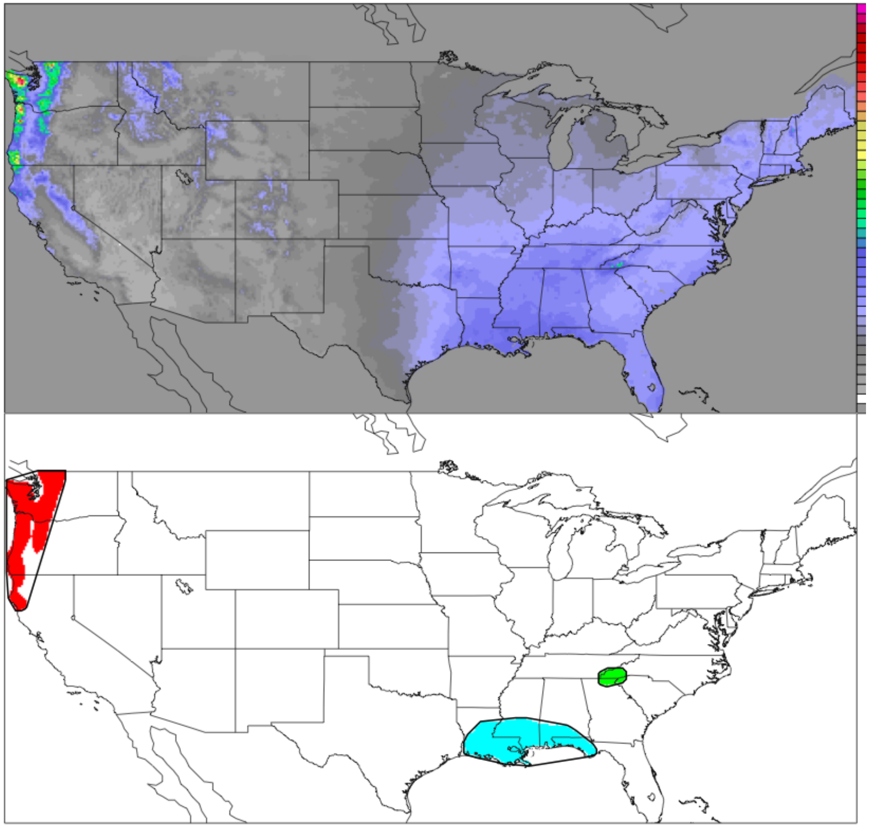
First Efforts

- Examine Drought Index: Standardized Precipitation Index (SPI)
- 36 month periods ending in December.
- Use existing spatial and spatio-temporal verification methods and tools.
- Determine what we can learn about climate model hydrologic processes using these tools.
- Enhance tools to provide additional information.

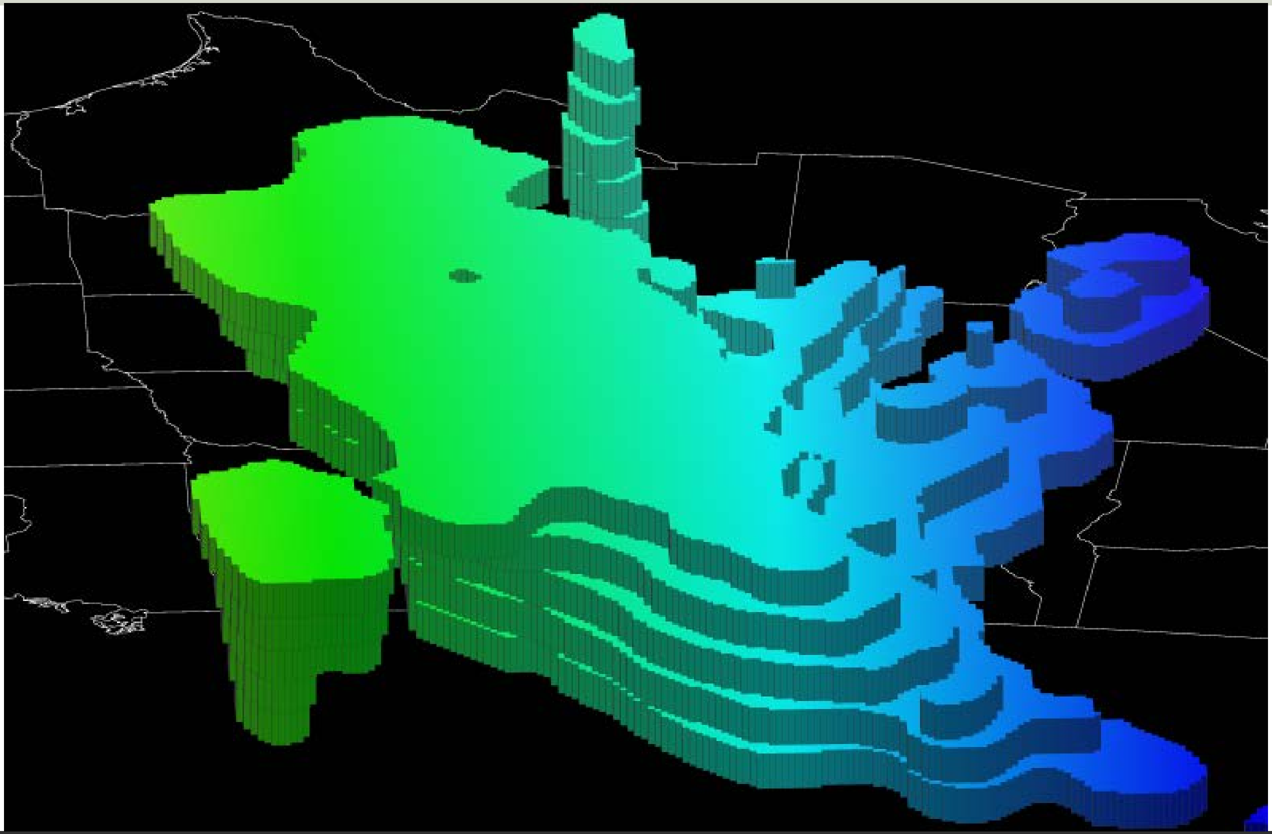
Objects in space

Identify events of interest in two dimensions.

Quantify and compare events with geometric and statistical measures.



Time is "up".



Objects in space and time

Take spatial objects and track them through time.

Answer questions like:

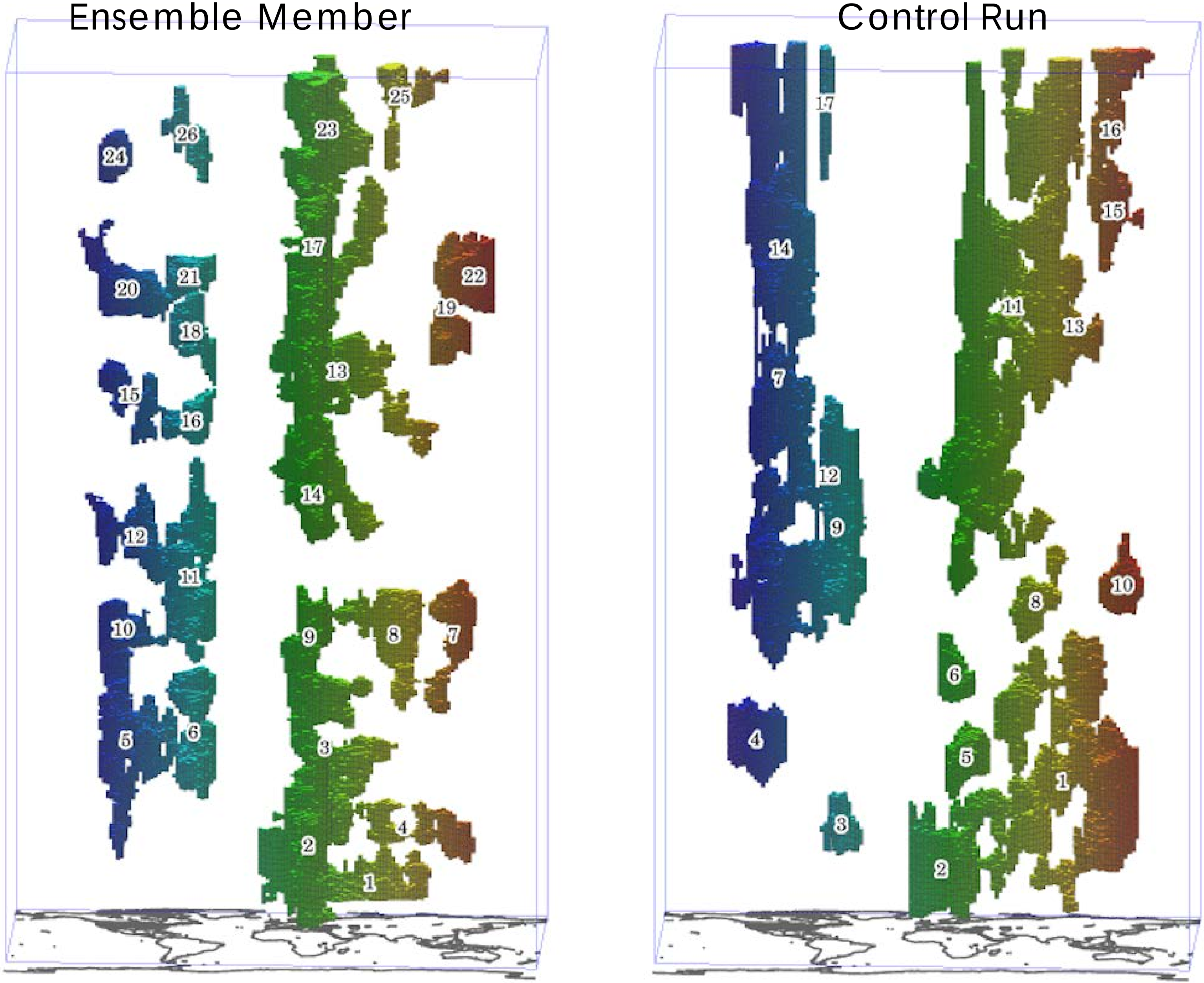
Does the drought move? Endure? Grow?

How does it compare with observations or other models? (Size, duration, location, intensity, etc.)

Time is "up".

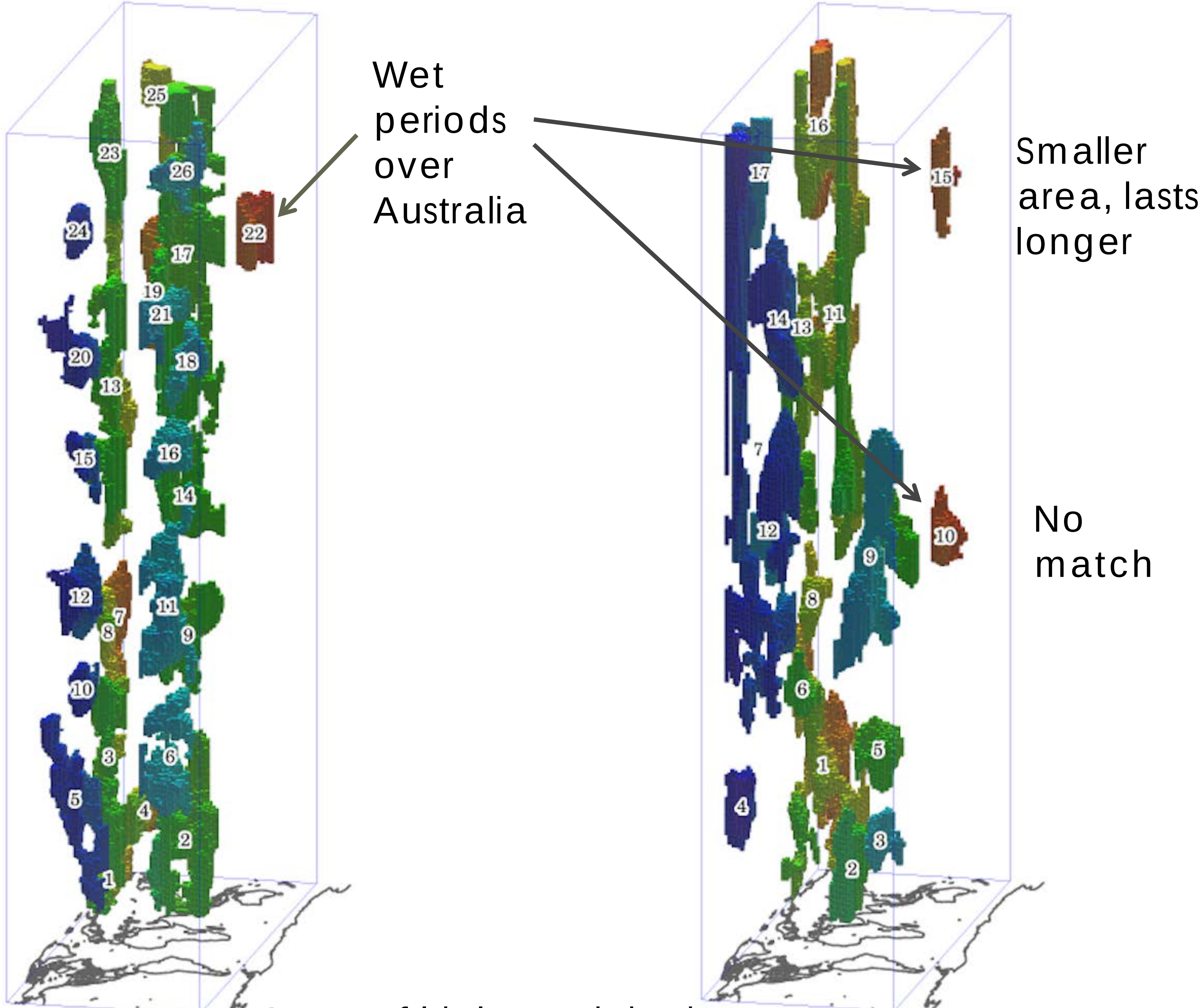
Ensemble Member

Control Run



Areas of high precipitation

Time is “up”.



Areas of high precipitation

Climate information made more user relevant

- Identifying events in time and space facilitates comparison of ensemble members and observations.
- Specific locations, durations, scenarios, etc. can all be examined.
- Facilitates planning and decision making for a variety of users.