Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data

Kristin Potter[1], Andrew Wilson[2], Peer-Timo Bremer[3], Dean Williams[3], Charles Doutriaux[3], Valerio Pascucci[1], and Chris R. Johnson[1]

[1] Scientific Computing and Imaging Institute, University of Utah

[2] Sandia National Laboratories

[3] Lawrence Livermore National Laboratory

SCI INSTITUTE

What is ensemble data?

Collection of data sets (*members*) generated by computational simulations.

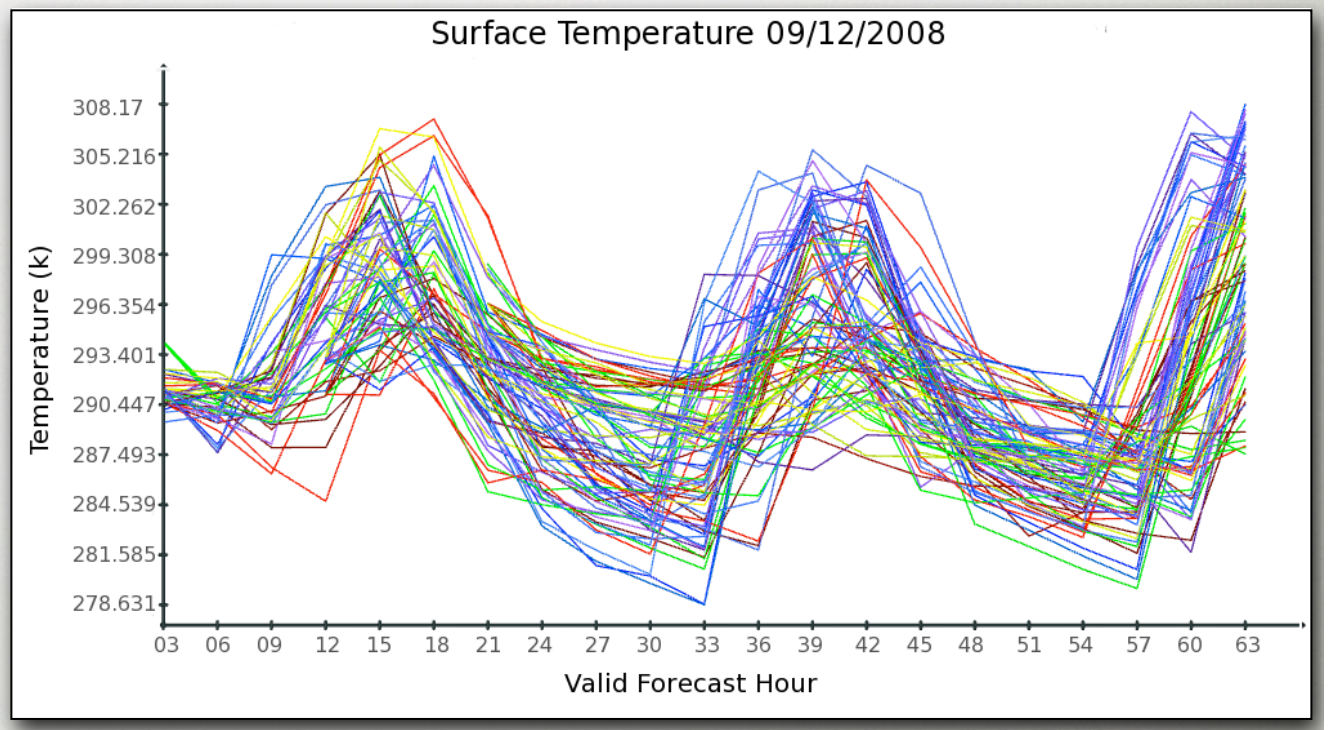
- Multidimensional
  - 2D or 3D spatial domain plus
  - time component
- Multivariate
  - simulations predict for numerous variables (i.e. temperature, humidity, etc)
- Multivalued
  - several values for each variable at each point

Why use ensemble data?

- Simulate complex systems
- Handle unknowns in initial conditions
- Investigate sensitivity to parameters
- Mitigate uncertainty

Ensemble data is complicated!

- Information rich
- Can get very large
- Not clear how or what to visualize
- Need an approach that handles these issues



Single weather station, single variable, all runs, across all valid forecast hours

Driving Applications

Short-Range Ensemble Forecasts (SREF)

- NOAA / NCEP
- Domain across North America
- Forecast weather variables out to ~3.5 days
- Public notification & warnings, aviation forecasts

Driving Applications

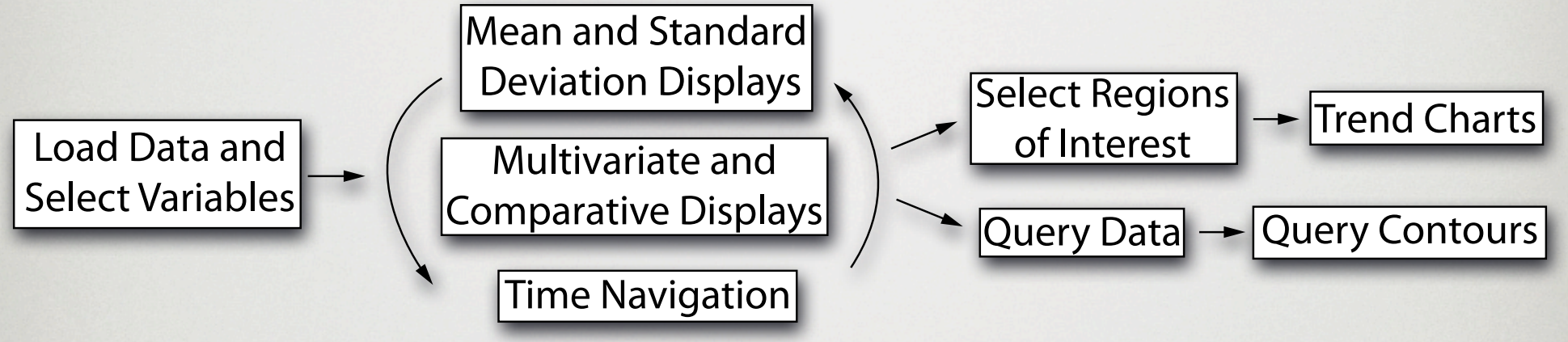
Climate Modeling

- IPCC Climate of the 20th century
- Spatial domain the whole globe
- Evolution over hundreds of years
- Impact of human activity, trends in natural disasters

IPCC
Intergovernamental Panel On Climate Change

SCI INSTITUTE

The Ensemble-Vis Framework

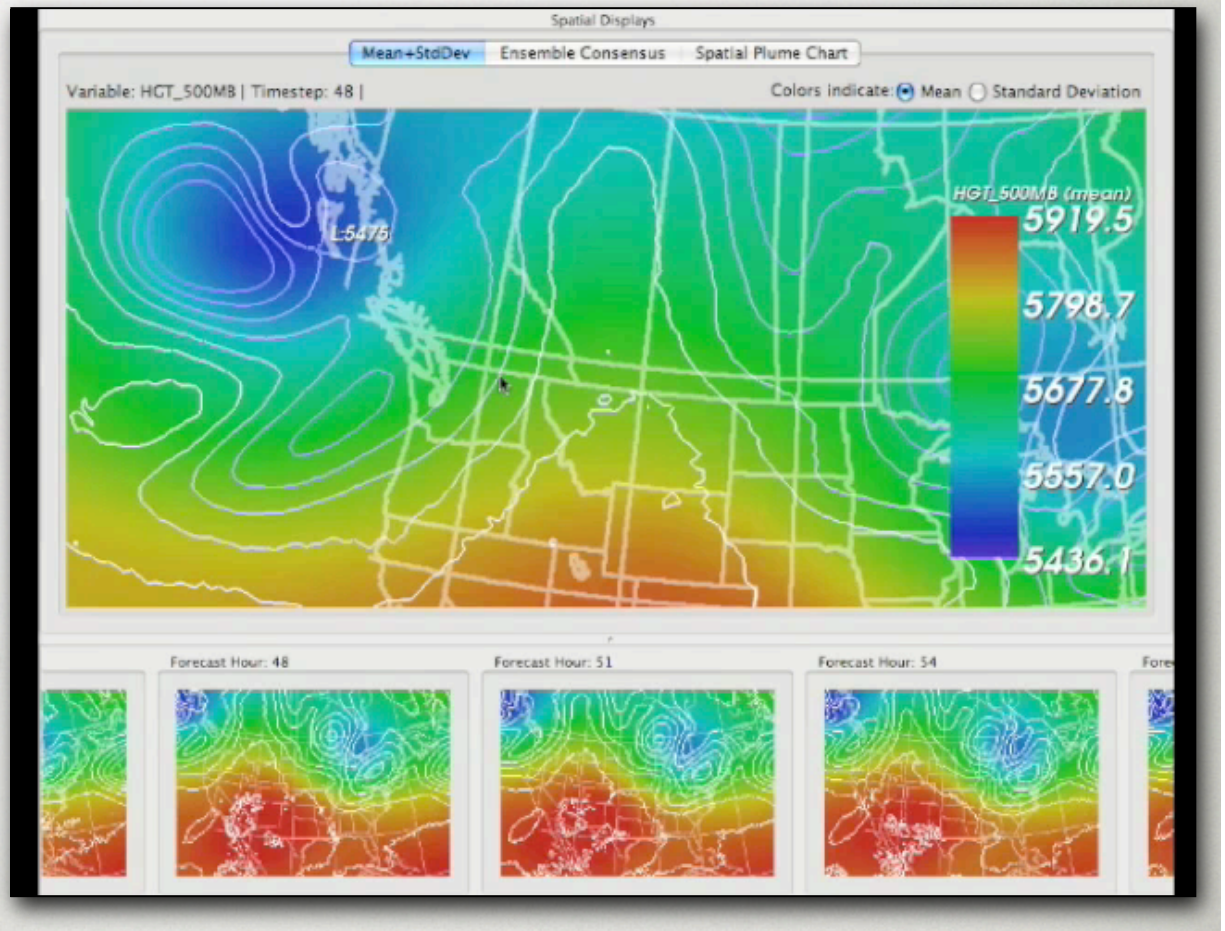


- User-driven, component-based framework
- Explore the range of possible predictions
- Probability of outcomes
- Interrogate the ensemble

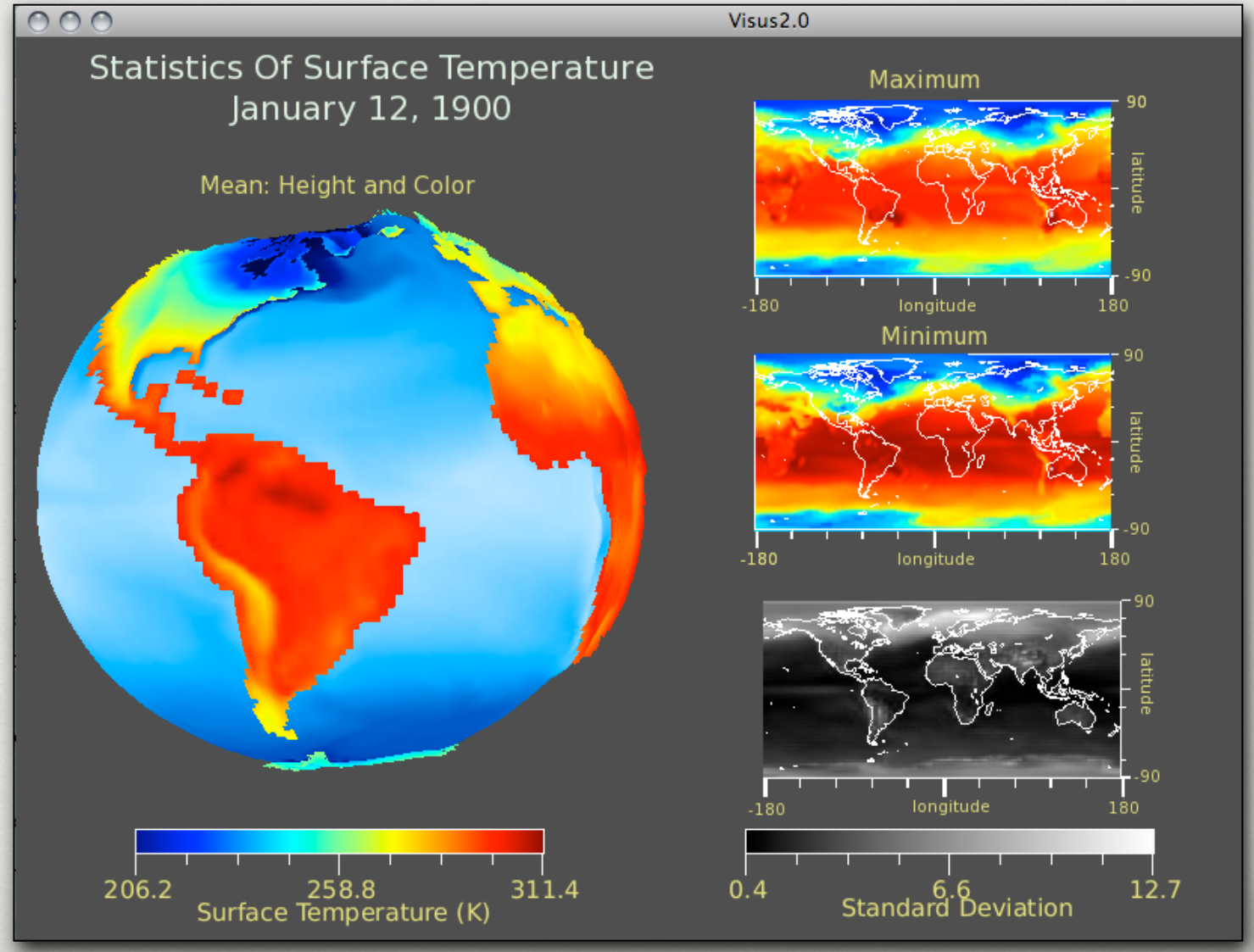
Ensemble Overviews

Mean & standard deviation

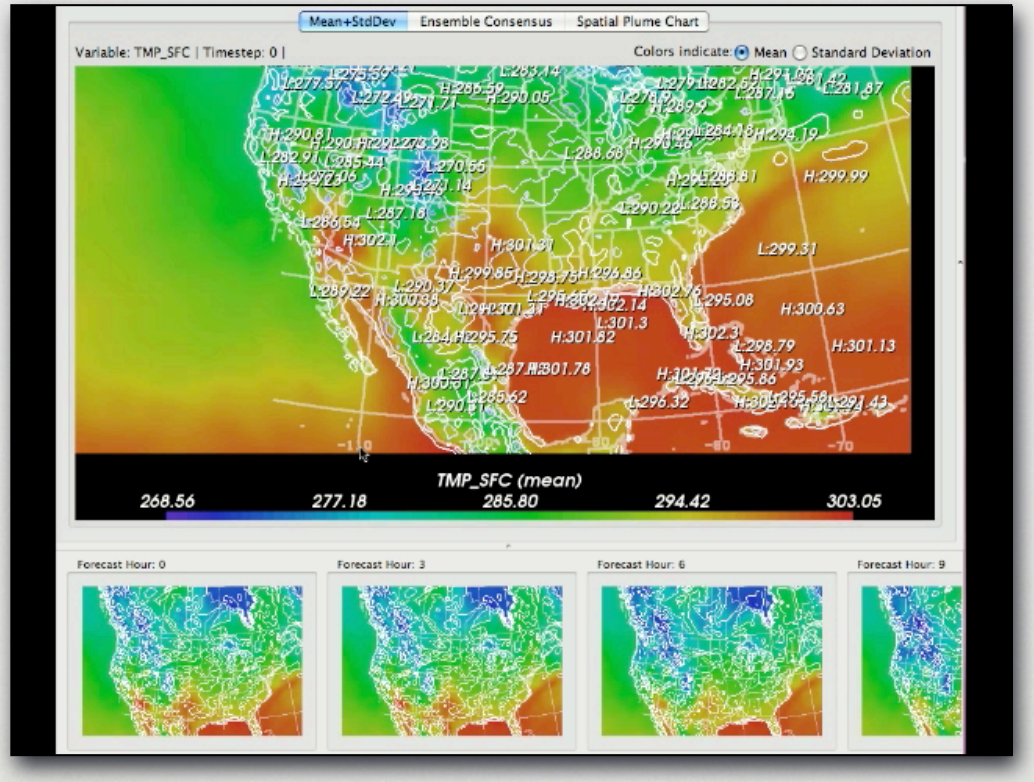
- roughly indicate value
- highlight areas of variation
- single time step across spatial domain



Height and Comparative Displays



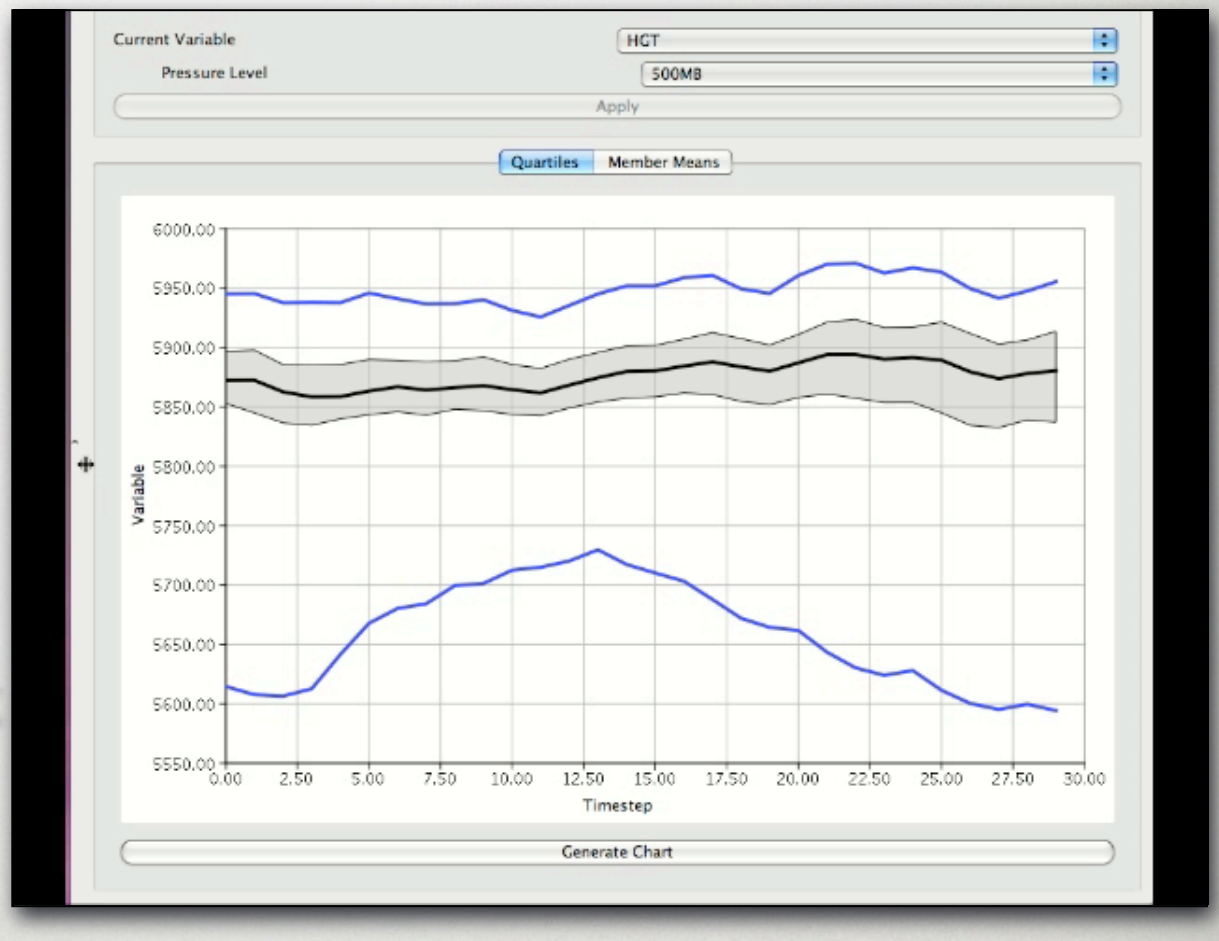
Time Navigation Summaries



- Small multiples showing each time step
- Quickly see evolution across time
- Choose time step of interest

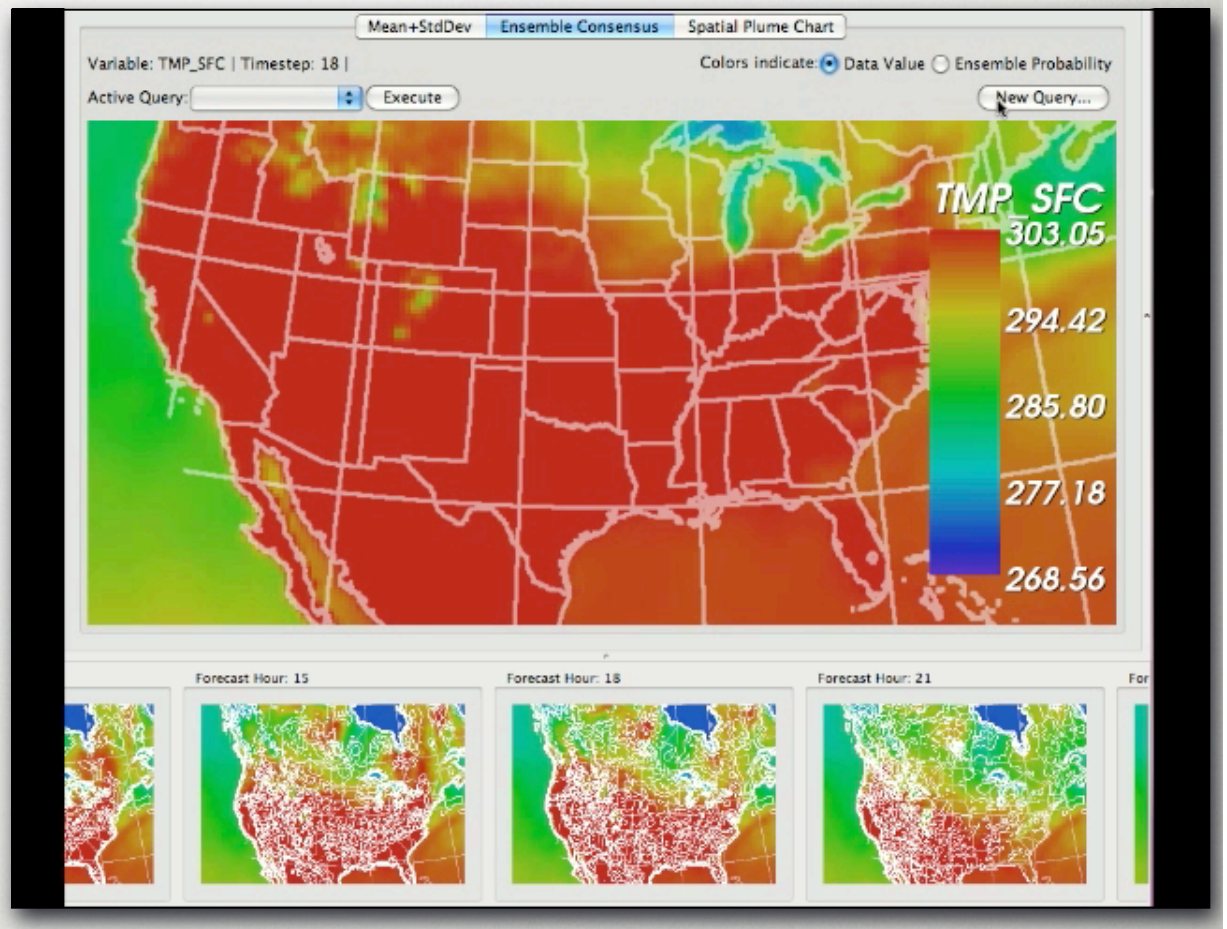
Trend Charts

- Select region of interest
- Show statistics like mean, quartiles, etc
- Drill-down to direct data display



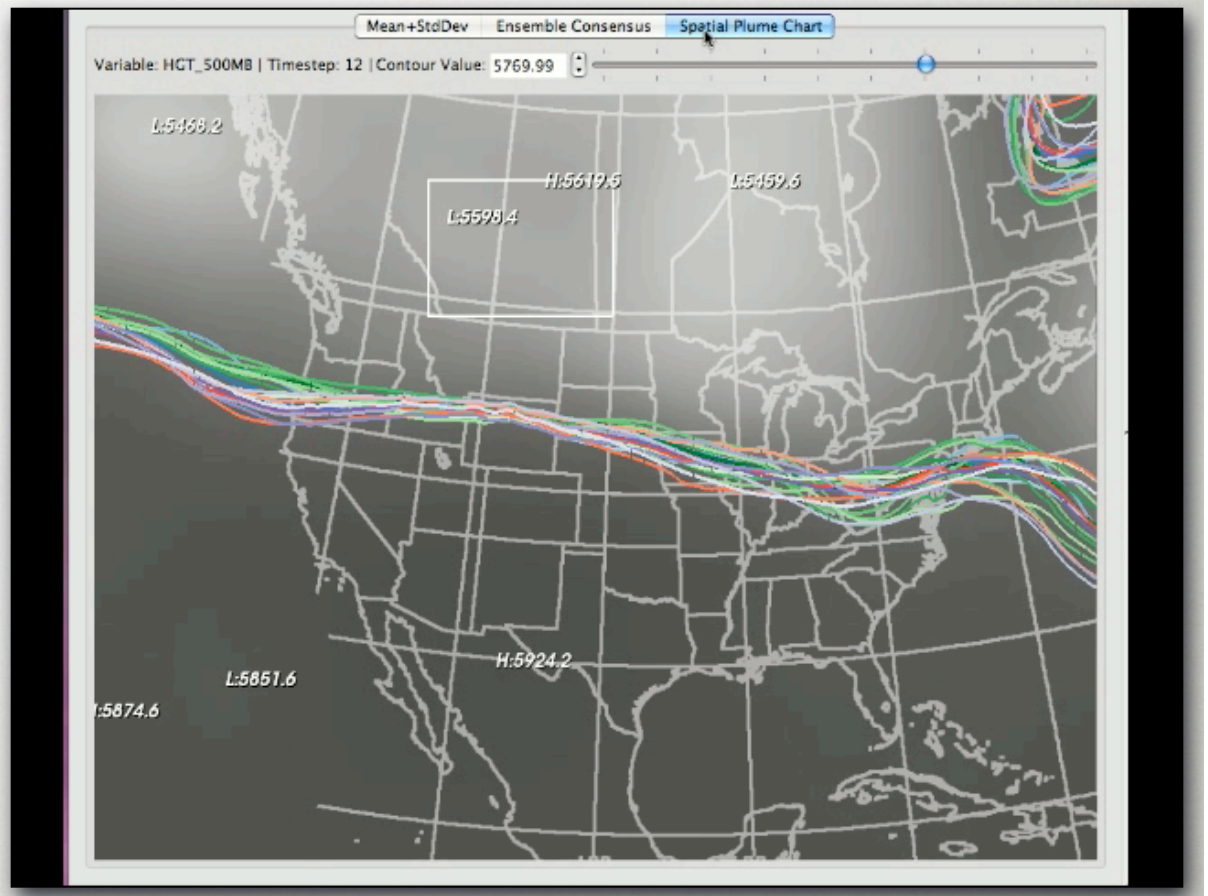
Query Contours

- User-driven query
- Select subset of data
- List of points where conditions are satisfied
- Scalar value at each point indicates number or percentage of satisfying members



Spaghetti Plots

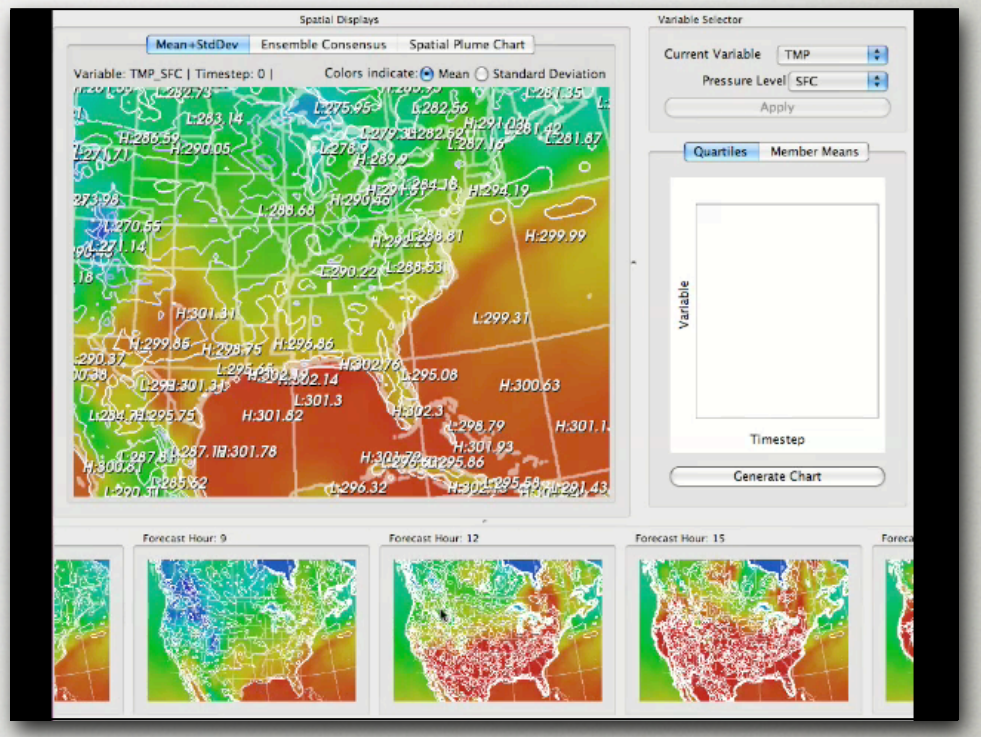
- Show variation across ensemble over space
- User chosen contour value
- Isocontour for each desired member
- Highlights outliers and divergence



Implementation

Two Prototypical Systems:

- SREF Weather Explorer
  - VTK filters, Qt widgets
  - Relational database backend MySQL & parallel Netezza

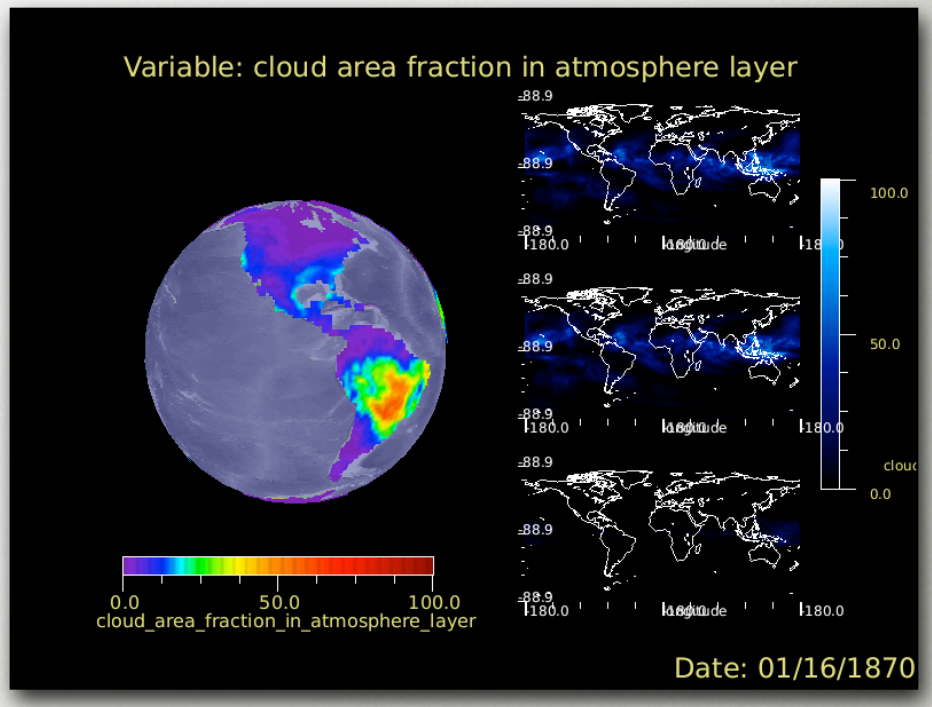


SCI INSTITUTE

Implementation

Two Prototypical Systems:

- ViSUS
  - Climate Data Analysis Tools (CDAT) integration
  - C++, OpenGL, Python, FLTK
  - Out-of-core streaming



SCI INSTITUTE

Conclusion

- Framework to let users drive visualization
- Combine various representations to highlight different aspects of the data
- General approach can be applied to numerous other fields
- Further work includes extension into higher spatial dimensions, feature detection algorithms, handling of non-normal distribs

Thanks!

This work was funded in part by the DOE SciDAC Visualization and Analytics Center for Enabling Technologies (www.vacet.org) and the NIH NCRR Center for Integrative Biomedical Computing(www.sci.utah.edu/cibc), NIH NCRR Grant No. 5P41RR012553-02.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

SCI INSTITUTE