



# 2nd NASA Data Mining Workshop



Issues & Applications in  
Earth Science

May 23-24, 2006  
Pasadena, CA

Elaine Dobinson

# Acknowledgements

- To our sponsors from NASA
  - Martha Maiden and Frank Lindsay
- To my co-organizers
  - Amy Braverman and Sara Graves
- To my program committee
  - Jeanne Behnke, Mike Burl, Becky Castano, Tom Hinke, Chris Lynnes, Bernard Minster, Rahul Ramachandran
- To the Westin Hotel
  - Janice Glosup
- To my secretary, web team, and volunteers
  - Stephanie Chong, Lynne Carver, Danny Hardin, Stephanie Granger, Brian Wilson

# Background

- The 1st workshop was held in October 1999, in Huntsville, AL
  - Smaller, less structured, getting together of folks interested in pursuing data mining for Earth science
  - EOS satellites were on the brink of sending data
  - Recommendations were made to NASA to promote the use of data mining (see report on line)
- This workshop was conceived to follow-up the previous one
  - To see what has been done in the intervening 6 years; hence the call for papers
  - To identify areas where data mining could potentially yield significant scientific advances in Earth science in the near and medium term
  - To move forward with the previous recommendations and make some new ones

# Workshop Objectives

- To bring together researchers from the Earth science community with researchers and technologists from the data mining and statistics communities
  - To identify ways in which data mining and statistics can help solve problems in Earth science
  - To better match the needs of the scientific community to the capabilities of the technologists, and to form working collaborations
  - To suggest future research directions the technologists may pursue to advance Earth science research & analysis
  - To identify ways that these technologies can be better infused into Earth science research & analysis

# Workshop Format

- Talks
  - Invited speakers from NASA and Interface
  - 16 oral presentations in 5 sessions selected in response to workshop call
- Posters
  - 15 poster presentations at Tuesday evening reception selected from the same call
- Guided discussions at the end of each session
  - As well as informal discussions throughout the day
- Panel & Discussion on Wednesday afternoon
  - *How to promote the infusion of data mining and statistical technologies into Earth science*
- Workshop Report
  - Electronic version of workshop highlights & recommendations
  - Appendices containing abstracts, talks, and posters

# Workshop Agenda

(full version in handout)

## **Tuesday, May 23rd - Fountain II**

8:00 - 8:30 Registration & Continental Breakfast

8:30 - 10:00 Session 1: Opening Talks, Session Chair: Elaine Dobinson; Session Recorder: Jeanne Behnke

10:00 - 10:30 Break

10:30 - 12:00 Session 2: Atmosphere 1, Session Chair: Chris Lynnes; Session Recorder: Stephanie Granger

12:00 - 1:30 Lunch (at the Westin)

1:30 - 3:00 Session 3: Climate, Session Chair: Becky Castano; Session Recorder: Chris Lynnes

3:00 - 3:30 Break

3:30 - 5:00 Session 4: Surfaces, Session Chair: Amy Walton; Session Recorder: Rebecca Castano

5:30 - 7:00 Poster Session & Reception - **Madera**

## **Wednesday, May 24th - Fountain II**

8:00 - 8:30 Continental Breakfast

8:30 - 10:00 Session 5: Land Cover, Session Chair: Mike Burl; Session Recorder: Rahul Ramachandran

10:00 - 10:30 Break

10:30 - 12:00 Session 6: Atmosphere 2, Session Chair: Rahul Ramachandran; Session Recorder: Brian Wilson

12:00 - 1:30 Lunch (on your own)

1:30 - 3:00 Session 7: Application of Data Mining and Statistics to Earth Science Research,

Session Chair & Panel Moderator: Amy Braverman; Session Recorders: Chris Lynnes, Stephanie Granger

3:00 Workshop Adjourns – Report Writing Begins (Program Committee, Panel Members, & Volunteers)

# Discussion Questions for Each Session

1. What data mining or statistical methods were used?
2. How were the techniques developed?
  - a) from scratch in response to a specific Earth science question, or
  - b) modified from another application or research project?
3. What is the importance of the science question?
4. What scientific results were obtained that would have been difficult or impossible without data mining and/or statistics?
5. What were the obstacles?
6. Did this work result in a geoscience publication? If not, why not?