Climate Model Diagnostic Analyzer

Seungwon Lee, Lei Pan, Chengxing Zhai, Benyang Tang,
Jonathan Jiang, Terry Kubar, Juilin Li, Joao Teixeira,
Zhangfan Xing, and Hook Hua

at

Jet Propulsion Laboratory, Caltech

Jia Zhang and her team at Carnegie Mellon University

Objectives

- Develop a novel methodology to diagnose model biases in contemporary climate models, to identify the physical processes responsible for creating model biases, and to incorporate the understanding into new model presentation that reduce the model biases.

- Implement the methodology as a web-service based, cloud-enabled, provenance-supported climate-model evaluation system named Climate Model Diagnostic Analyzer (CMDA) for the Earth science modeling and model analysis community.
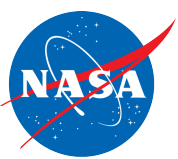
CMDA

- Enables multi-aspect, physics-based climate data analyses.
- Facilitates comprehensive and synergistic use of observational data, reanalysis data, and model outputs.
- Is a web-service oriented system.
- Does not require local software/library installation.
- Provides all the input data needed for analysis.
- Runs on the Amazon cloud system.
- Supports interactive visualization.
- Collects provenance and supports provenance-based search.

CMDA Data Sets

Model Outputs from CMIP5 project

To be Evaluated

- Experiments:
  - Historical, AMIPs, Forecast
- Models:
  - CCCMA/canesm2, GFDL/esm2g, GISS/e2-h, GISS/e2-r, NCAR/cam5, NCC/noresm, UKMO/hadgem2-es, CCCMA/canam4, CSIRO/mk3.6, GFDL/cm3, IPSL/cm5a-lr, MIROC/miroc5, UKMO/hadgem

Observation Data from Obs4MIPs

References

- AMSR-E surface temperature
- AIRS and MLS air temperature & water vapor content
- MODIS total cloud fraction, leaf area index
- GPCP and TRMM precipitation
- AVISO sea surface height
- CERES radiation fluxes

Reanalysis Data from ECMWF and Merra

References

- Vertical Wind
- Relative Humidity

CMDA Analysis Tools

Single Variable Analysis

Multiple Variable Analysis

2-D Variable Analysis

3-D Variable Analysis

- temporal averaging and plotting
- time series generation and plotting
- zonal mean calculation and plotting
- horizontal regridding

- vertical profile averaging and plotting
- pressure-level slicing and plotting
- zonal mean calculation and plotting
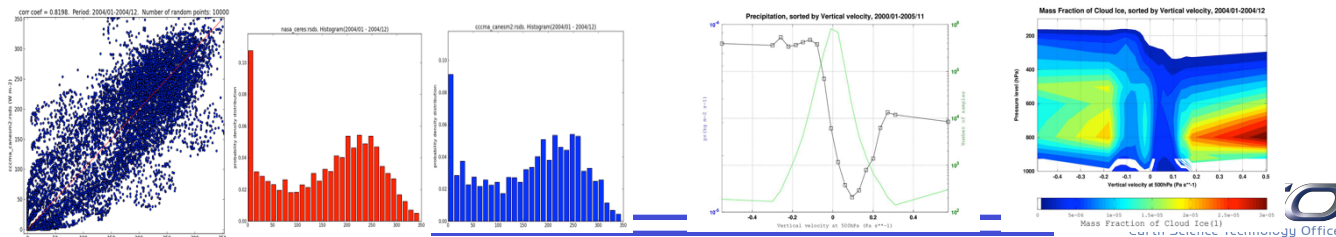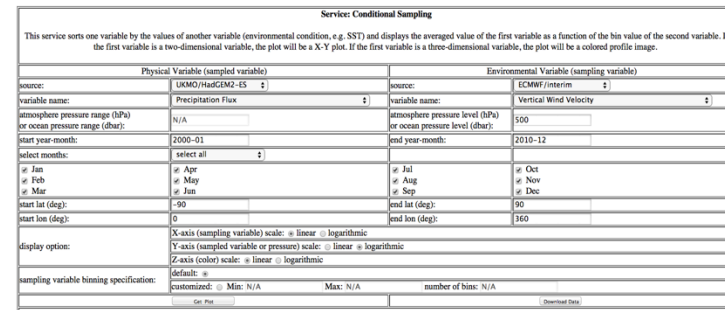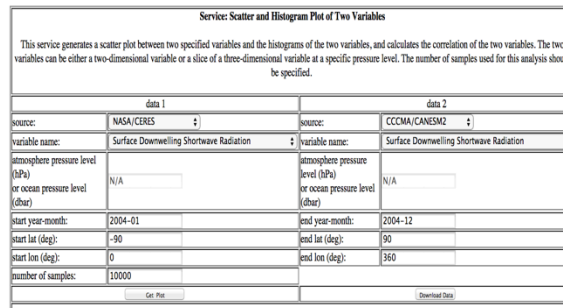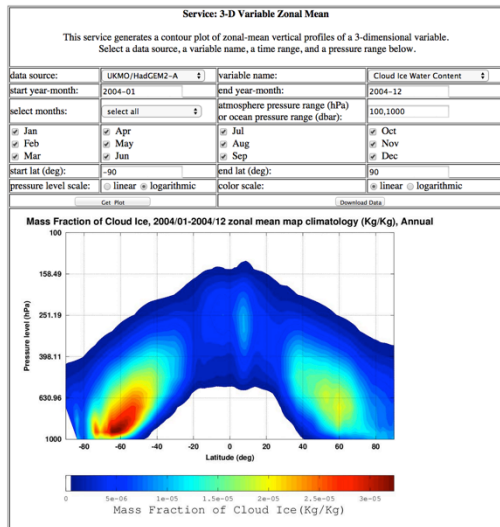- Horizontal and vertical regridding
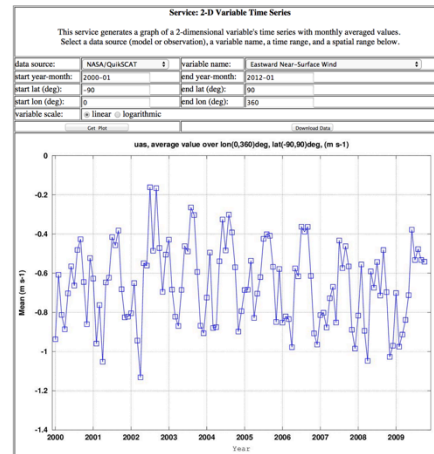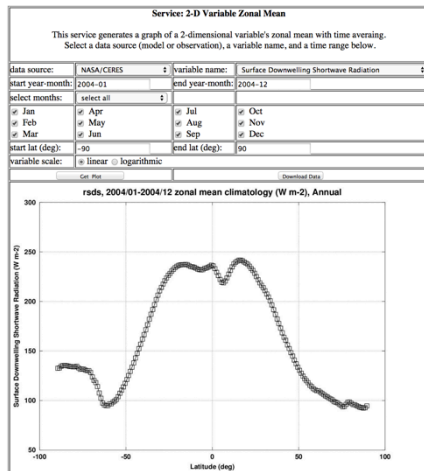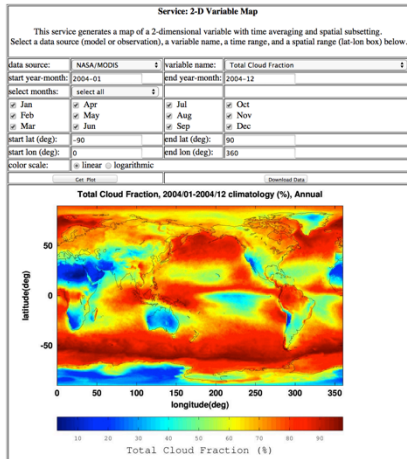
- time-lagged correlation calculation
- scatter plotting
- histogram
- difference calculation and plotting
- conditional sampling
- random forest
- conditional probability density

Analysis Tool Screenshots

http://cmacws2.jpl.nasa.gov:8080/cmac/
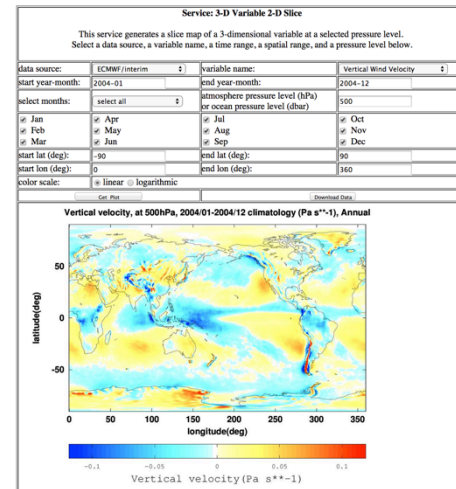
Previous Educational Use

- **CMDA was used as an educational tool for the Summer School** organized by JPL's Center for Climate Science in September 22-26, 2014.

- **The theme** of the Summer School in 2014 was **"Using Satellite Observations to Advance Climate Models"**.

- **23 students attended** the summer school: graduate students or postdocs from around the world.

- **Five research topics were prepared** by the JPL Earth scientists.

- **The students worked on their research topic for total 5 hours** of group research time.

- **Each student was assigned to a virtual machine** in Amazon Cloud to use CMDA.

- **An one-hour session of the CMDA introduction was given** and the students were able to use the tool immediately to do their research.

- **The students presented their result** on the last day of the school.
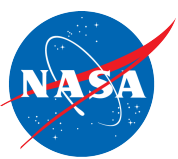
2014 JPL Summer School
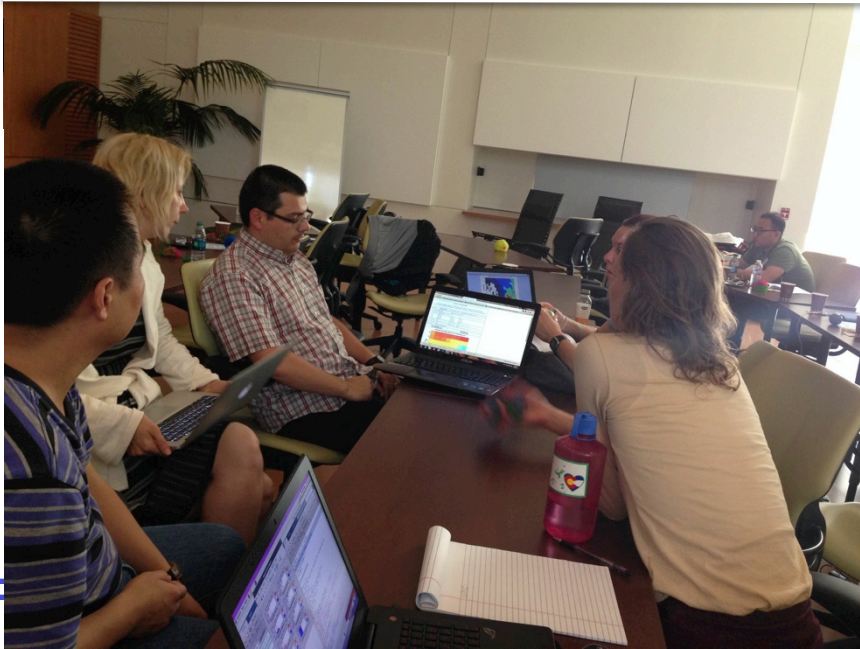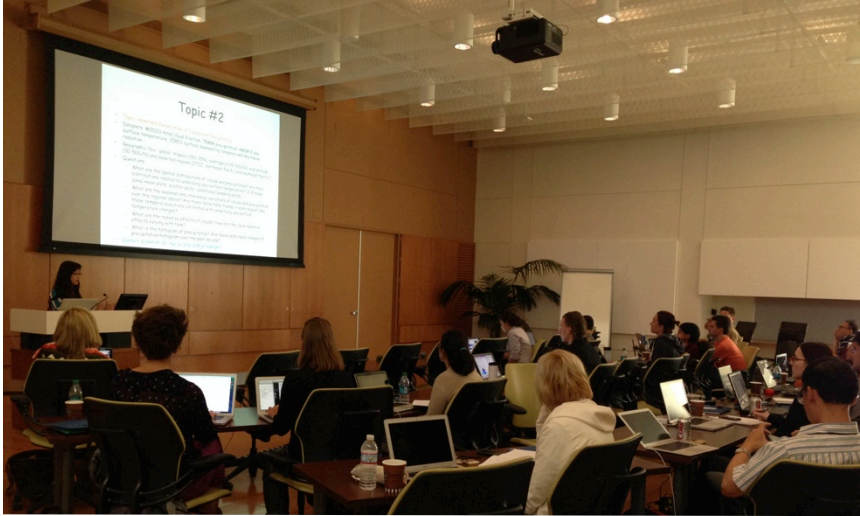
Student Feedback

- A tool survey was given to students, and students comments were collected.

- **Comments from students for appreciation:**
  - "I found the tool to be very user-friendly."
  - "It is really straightforward to use."
  - "It is not so difficult to use, seems intuitive."
  - "It was very useful for the exploratory analysis"
  - "The web tool was pretty instrumental in the completion of the research project."
  - "I could see this tool bing easily used in classroom settings."

- **Comments from students for improvement:**
  - "Provide a comprehensive description of the datasets available."
  - "Provide the time ranges over which the variables are available."
  - "Provide a tool to download data with a user specified condition (time, space, grid resolution for more detailed analyses in other programs."
  - "Provide interactive visualization capabilities: interactive space selection, color range selection, etc."
  - "Provide an exemplary picture of each analysis tool's output"
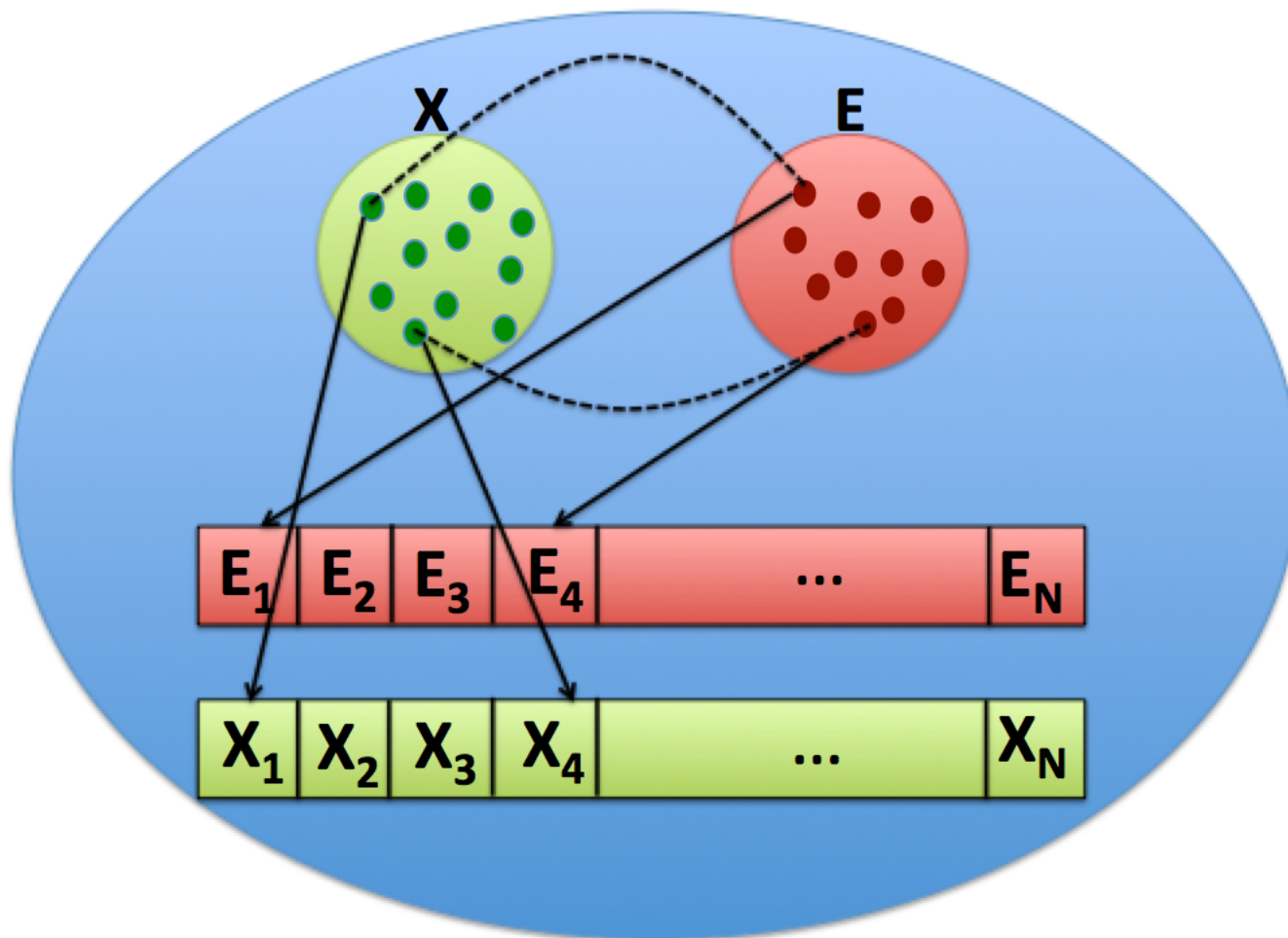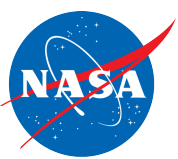
Diagnostic Methodology

- Conditional Sampling

- Random Forest Feature Ranking

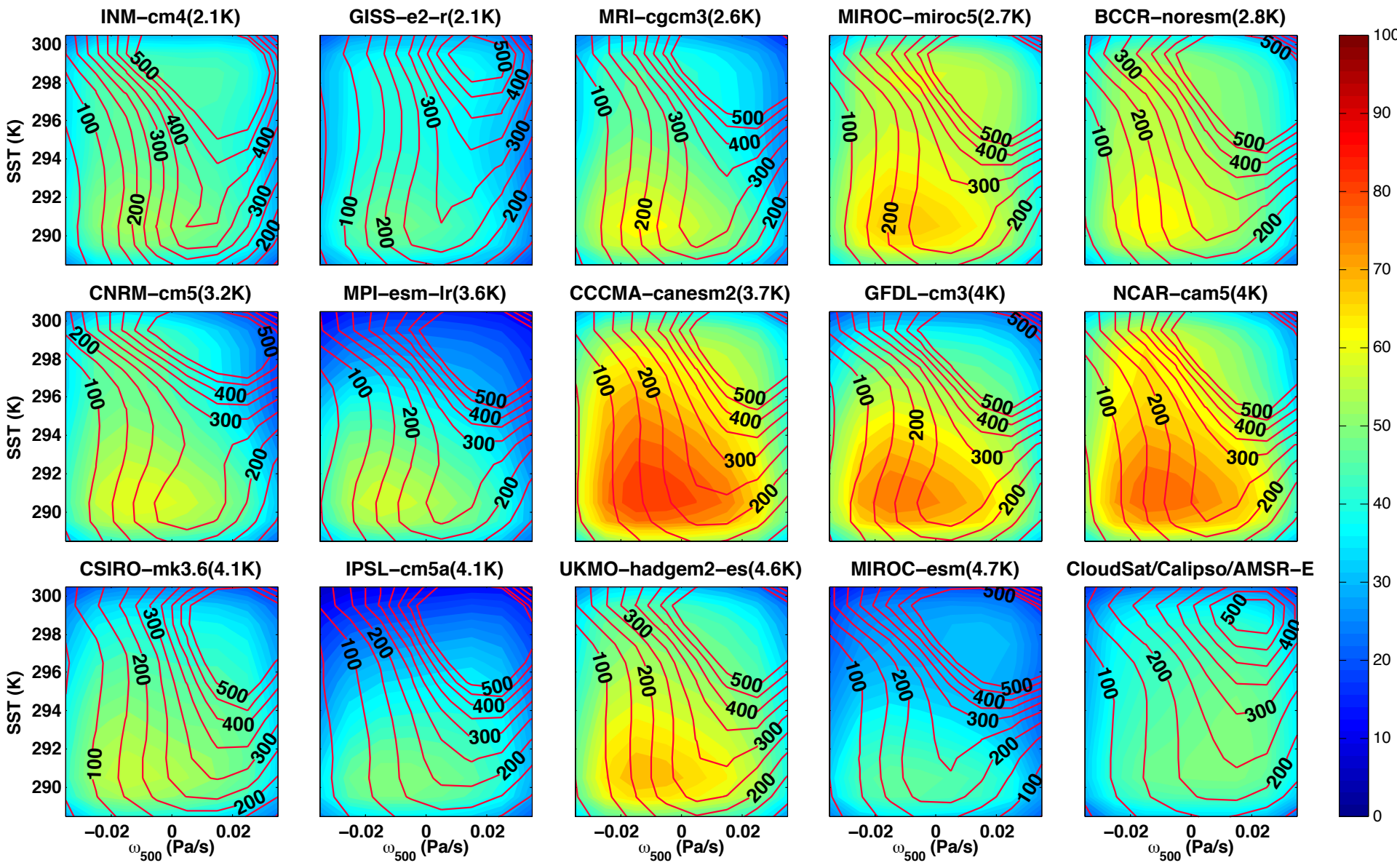- Conditional Probability Density

- Time Lagged Correlation

Conditional Sampling

Displays a physical quantity **X** according to the values of another physical quantity **E**, which is related to **X** via physical processes.
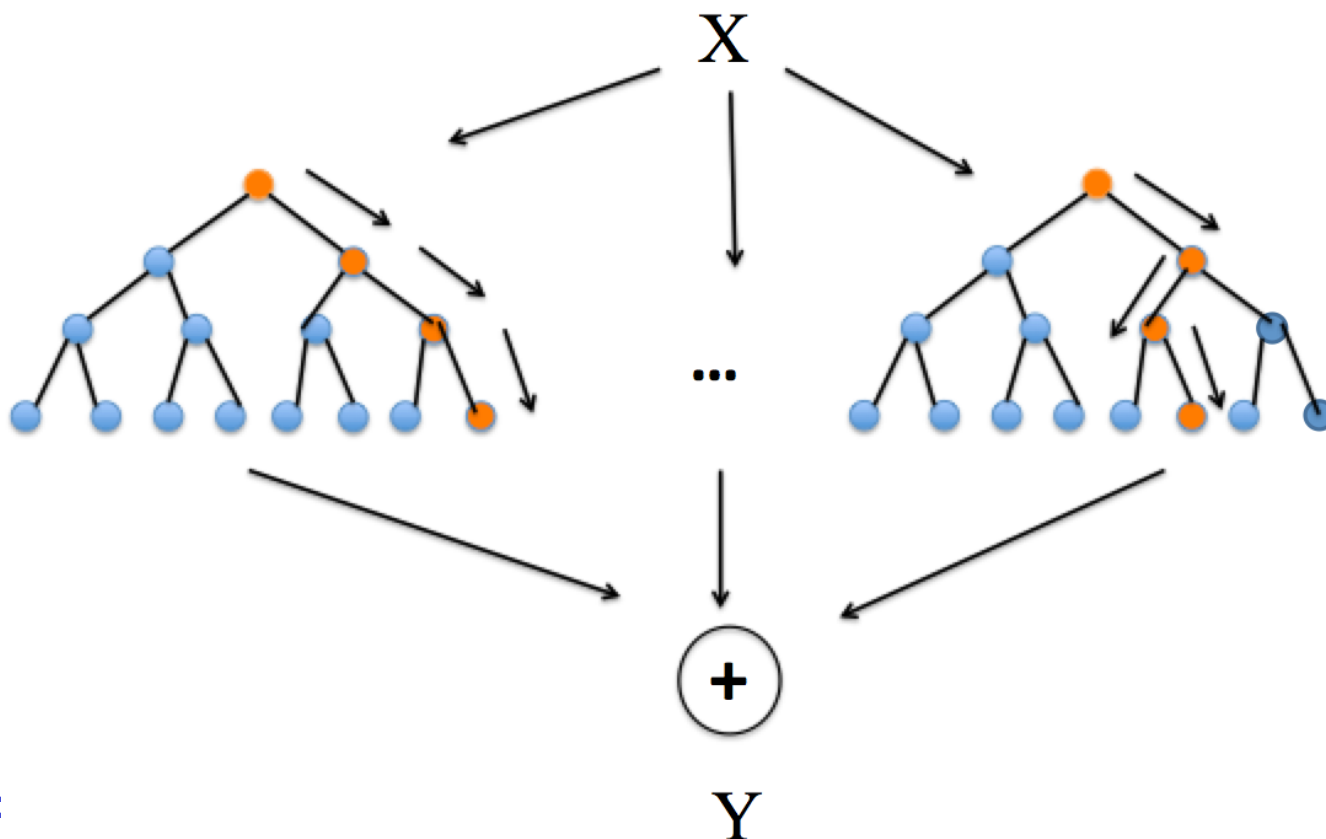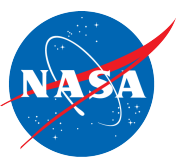
Conditional Sampling
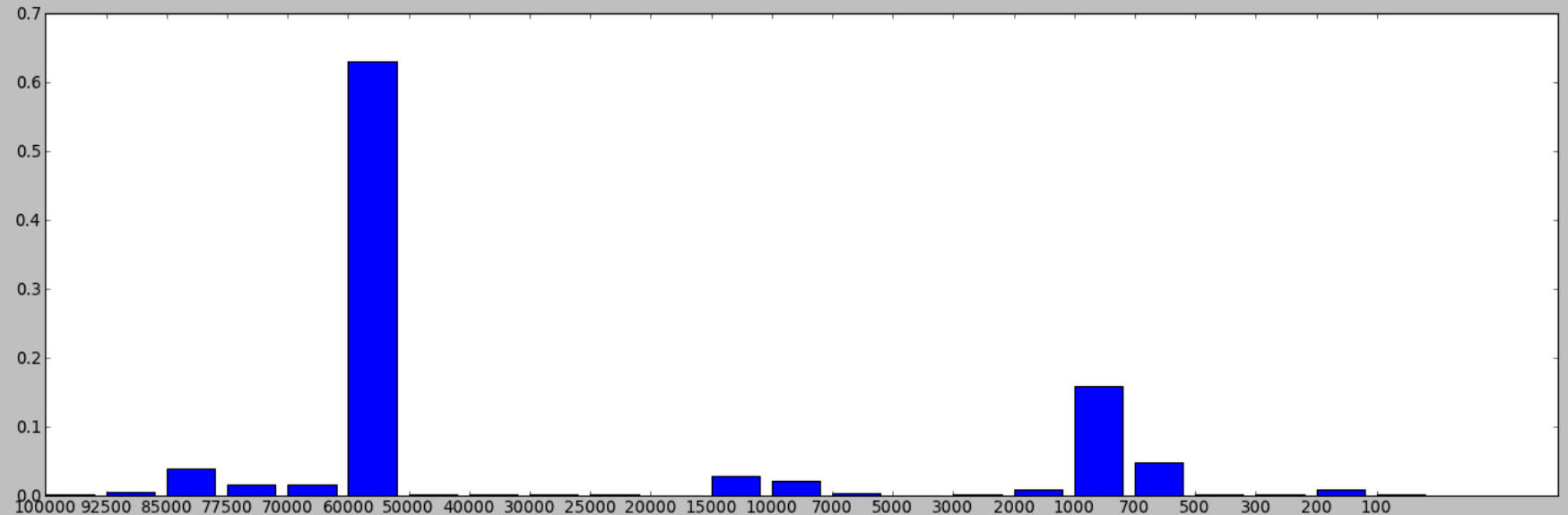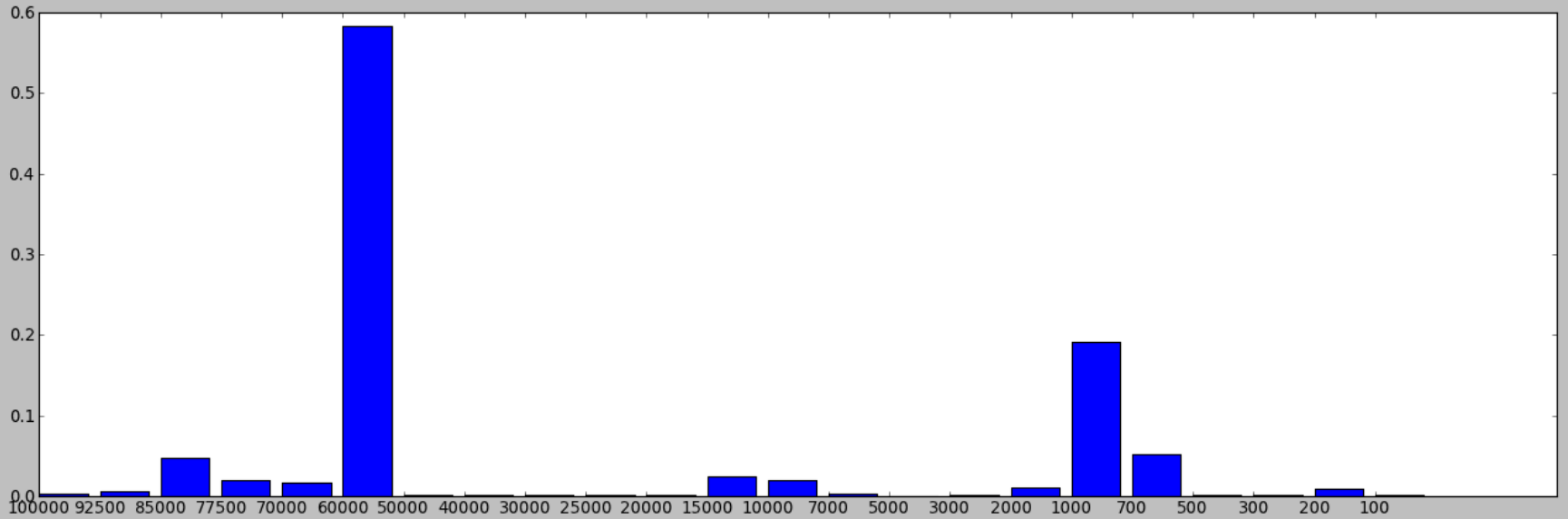
Random Forest Feature Ranking

Measures how important each input variable X_i is in predicting the regression target variable Y without assuming a particular relationship. Build an ensemble of decision trees.
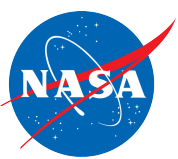
Random Forest
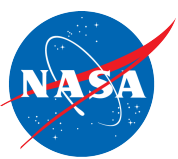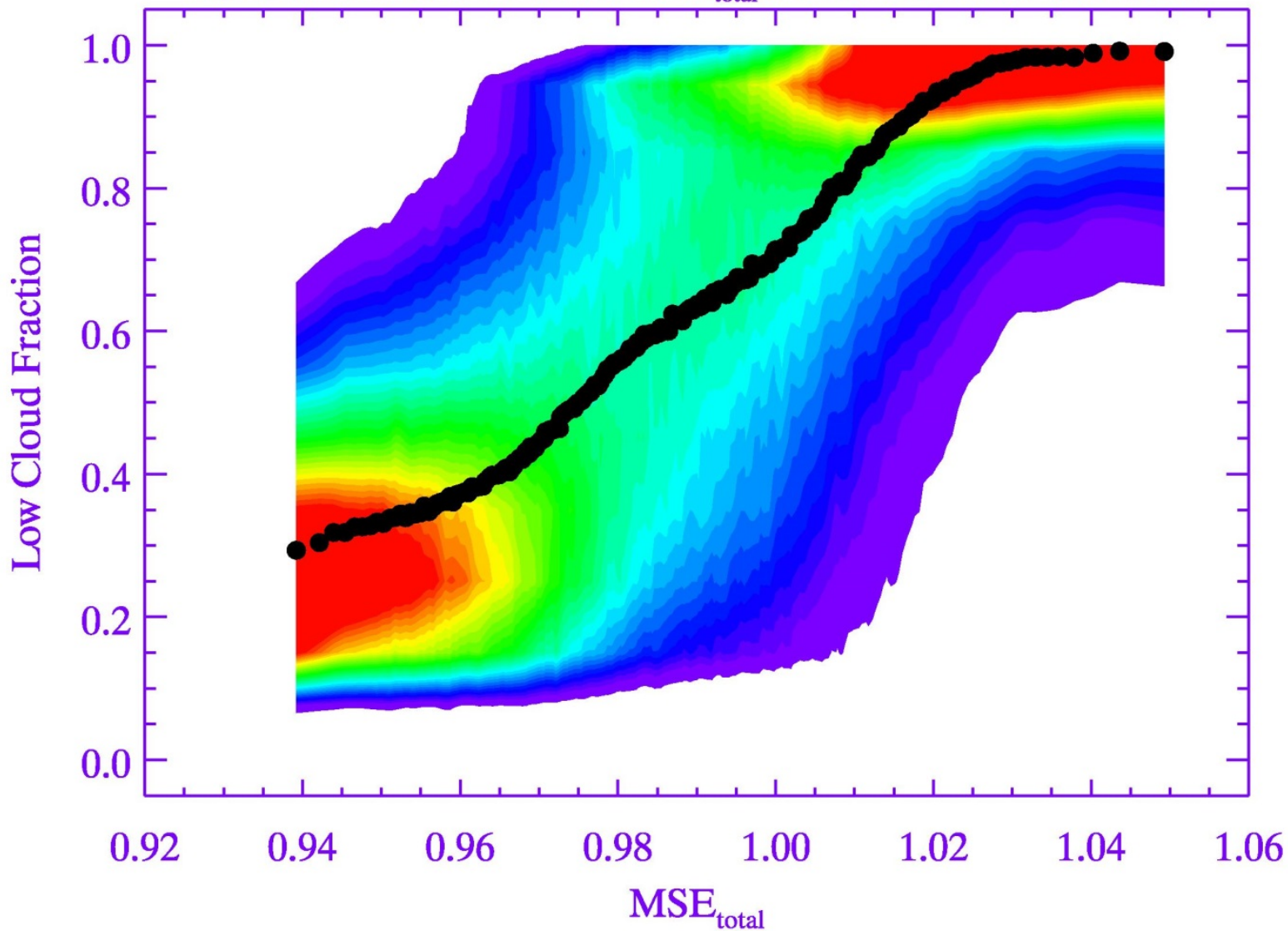
Conditional Probability Density

Displays a probability density function of the sampled parameter **X** for a given sampling parameter bin **E**. Carries considerably more information about the relationship between parameter **X** and **E** than the mean and variance of the parameter **X** for a given **E**.

MODIS Low Cloud PDFs vs MSE_{total} over Entire NE & SE Pacific

● MODIS Median Low CF for CF>0 Low Cloud Scenes

Distribution of Low Cloud Fraction Within Each MSE Bin

0.05 0.07 0.09 0.11 0.13 0.15 0.17 0.19

Provenance System

- A pilot study with Carnegie Mellon University Prof Jia Zhang's Group.
- Building a provenance-driven recommendation engine for CMDA

http://einstein.sv.cmu.edu:9019/climate

- Service configuration/execution provenance
- Data usage provenance
- Service usage provenance
- Provenance-based search
- Data recommendation
- Service recommendation
- Bug report
- Web service publication and interface design

Summary

- CMDA will provide the climate modeling and model analysis community with a diagnostic tool to evaluate climate models.

- CMDA will help the scientists identify the physical processes responsible for creating model biases.

- CMDA will facilitate community-wide use and relatively effortless adoption of the novel diagnostic methodology through web-service and cloud technology.

- CMDA will collect processing history and allow provenance-based search.