"Building an Atmospheric Rivers Catalog Using Big Data Approach"

CHRS - CONNECT and PERSIANN CDR

Soroosh Sorooshian

Phu Nguyen

Hao Liu

Center for Hydrometeorology and Remote Sensing
University of California Irvine

"Improving Sub-Seasonal & Seasonal Precipitation Forecast for drought Preparedness"

DWR-WSWC Workshop, San Diego CA May 28th, 2015

University of California Irvine (UCI)
Research Team: Present and Recent Past

and many more ...

Prediction Requirements for Water Resources

Short Range - · - · - · - · - · - · - · - · - · - · - · - · - · - · - · - · - · - · -> Long Range

hours - - - - -> days - - - - -> weeks - - - -> months - - -> seasons - - -> years - - - - - - -> decades

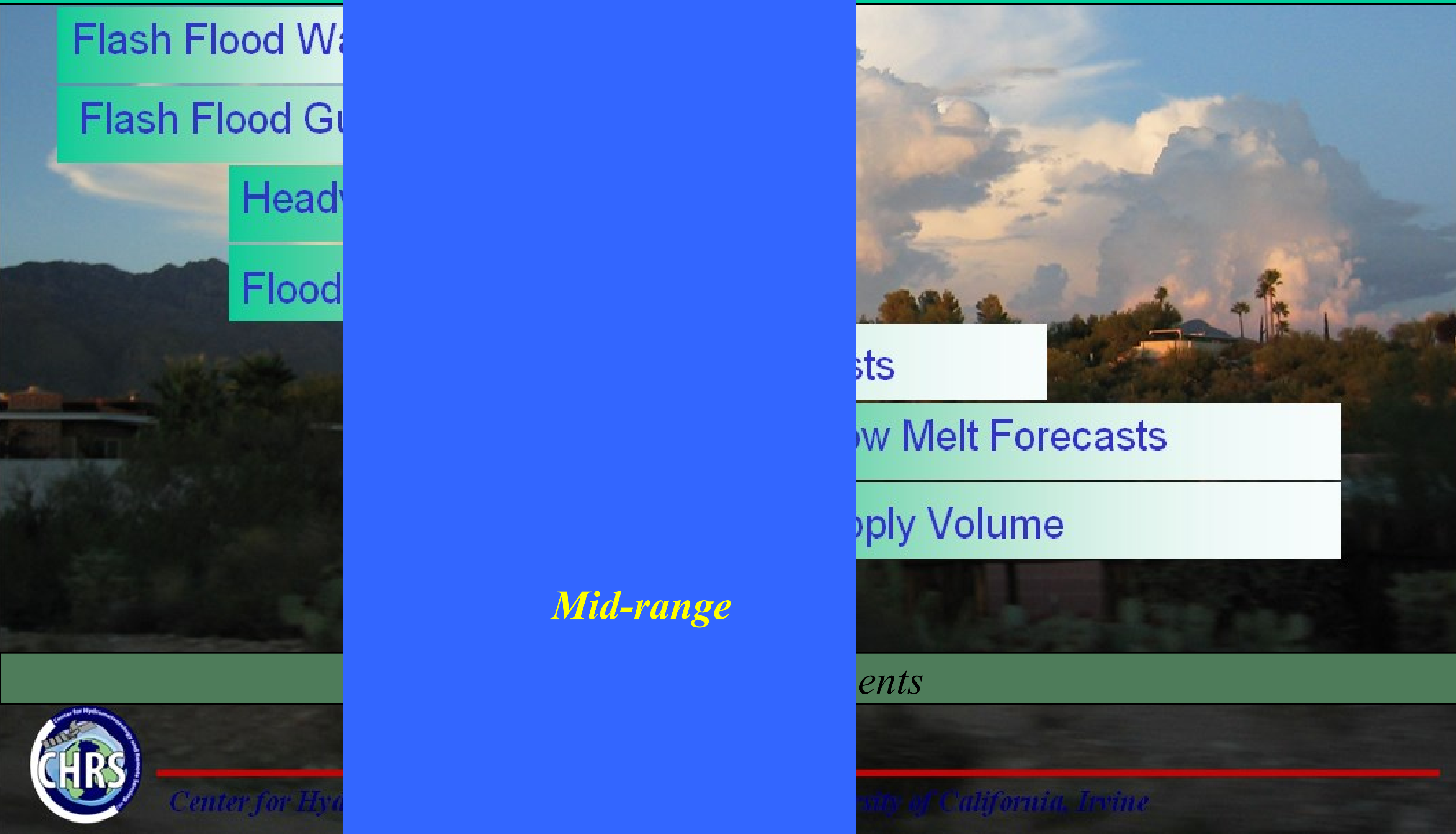
Mid-range

Long-range

Prediction Requirements for Water Resources

Short Range – · – · – · – · – · – · – · – · – · – · – · – · – · – · – · –> Long Range

hours - - - - -> days - - - - -> weeks - - - -> months - - -> seasons - - -> years - - - - - - -> decades



Flash Flood Wa

Flash Flood Gu

Head

Flood

sts

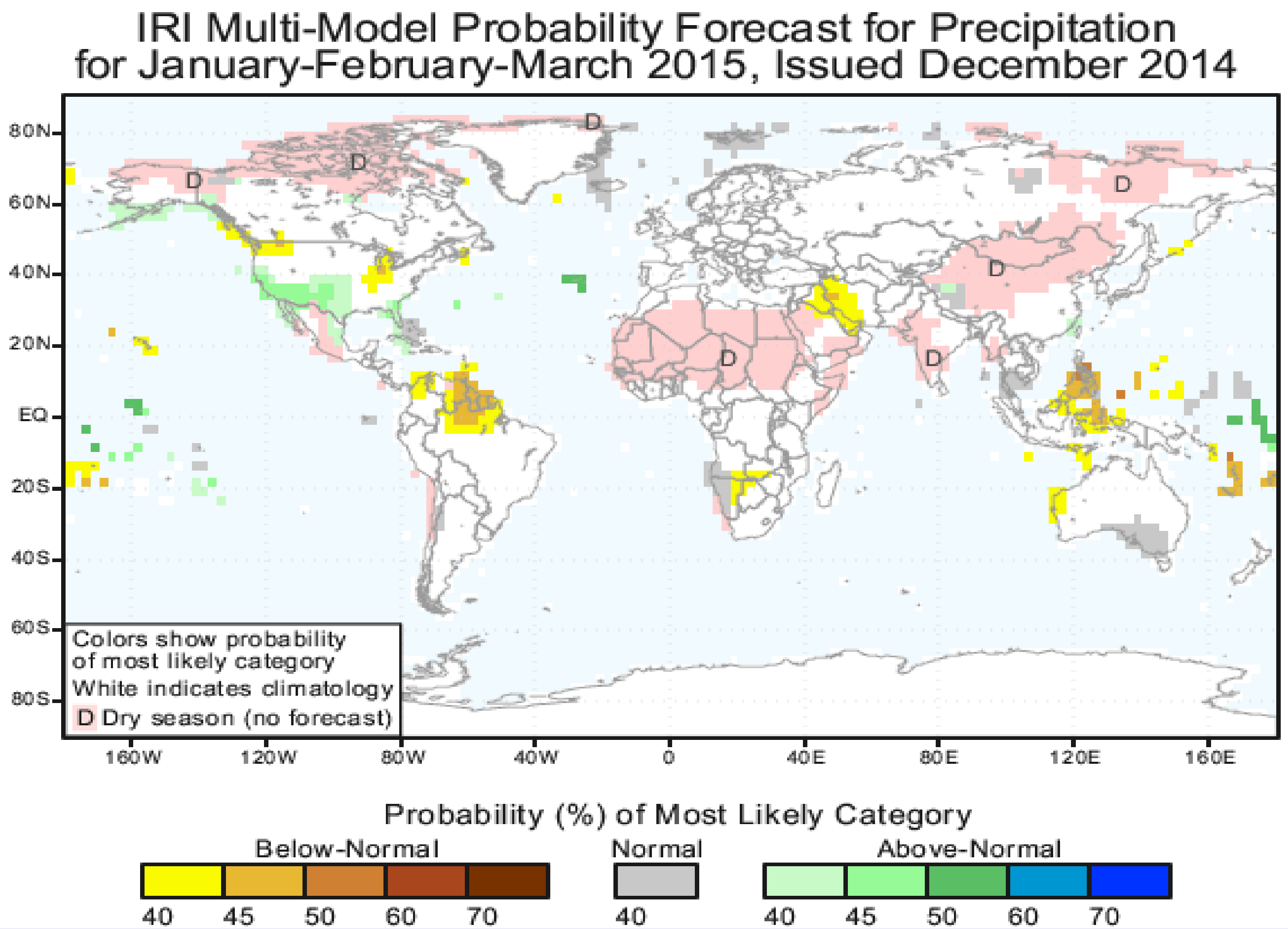
w Melt Forecasts

pply Volume

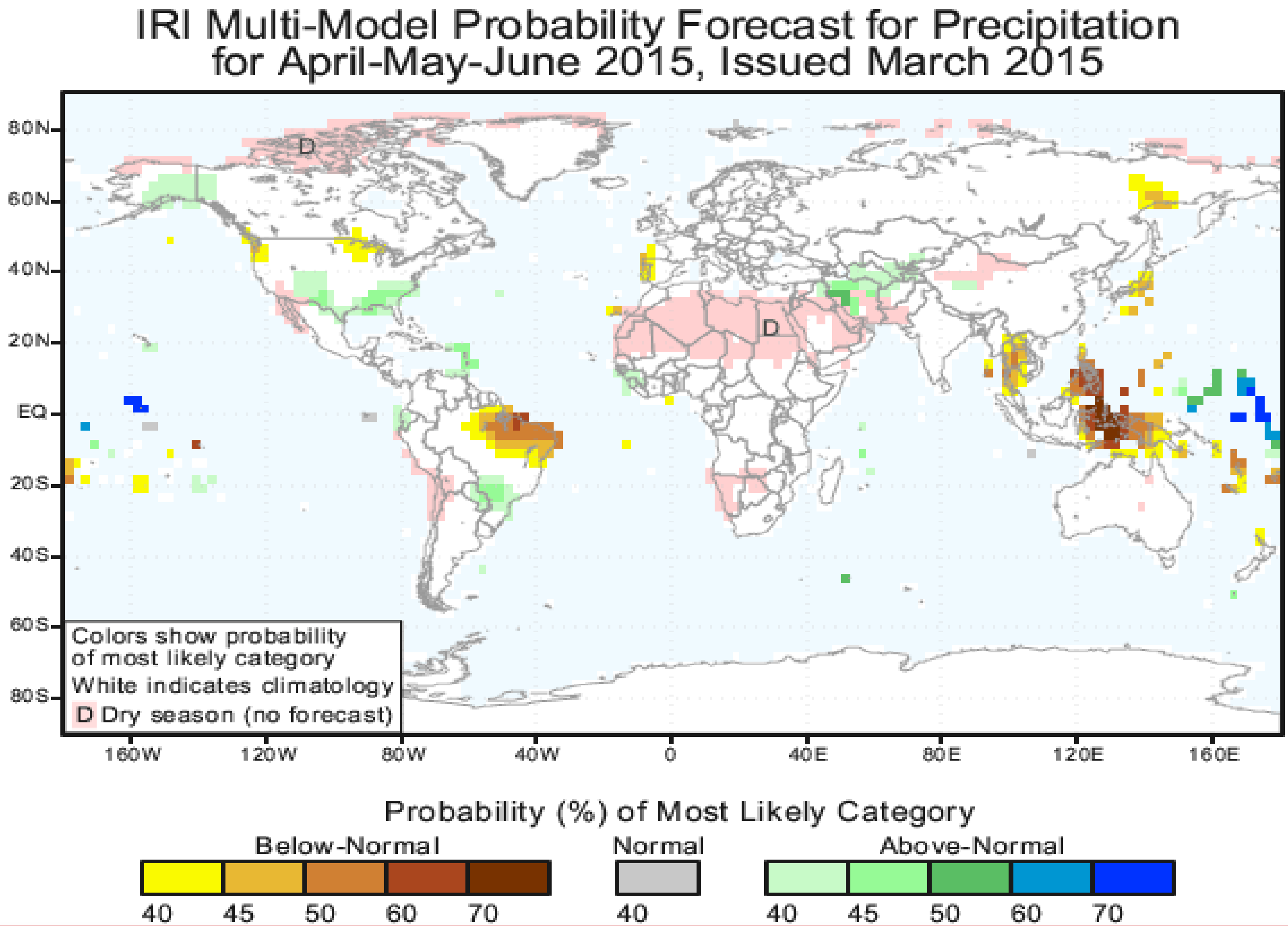
Mid-range

ents

IRI 3-Month Multi-Model Probability Precipitation Forecast



IRI 3-Month Multi-Model Probability Precipitation Forecast



Climate-Scale approaches to addressing hydrologic extremes

Short Range – · – · – · – · – · – · – · – · – · – · – · – · – · – · – · –> Long Range

hours ----> days ----> weeks ---> months --> seasons --> years ------> decades

Flash Flood Warning

Flash Flood Guidance

Headwater Guidance

Flood Forecast Guidance

Reservoir Inflow Forecas

Spring Sno

Water Sup

Long-range

Forecast Requirem

Hydrologically-Relevant Climate Variables

Climate Models Projections of Future Precipitation Patterns and Amounts?

Western U.S. future model projections

Dr. Chiyuan Miao - BNU

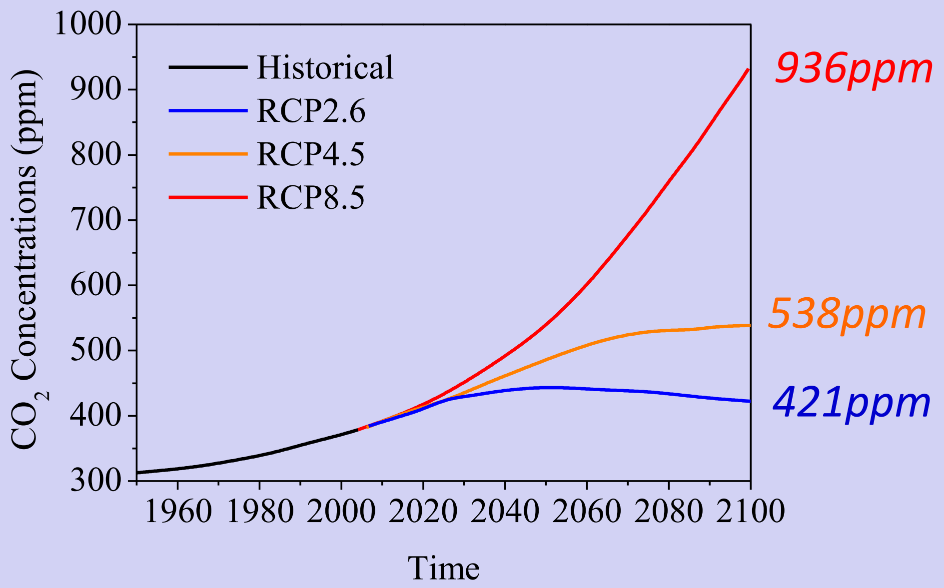
Future Modeling Scenarios – IPCC AR5

Representative Concentration Pathways (RCP) Scenarios:

RCP2.6: represent 'low' scenarios featured by the radiative forcing of 2.6 W/m^2 by 2100, the resulting CO_2-equivalent concentrations is 421 ppm in the year 2100 .

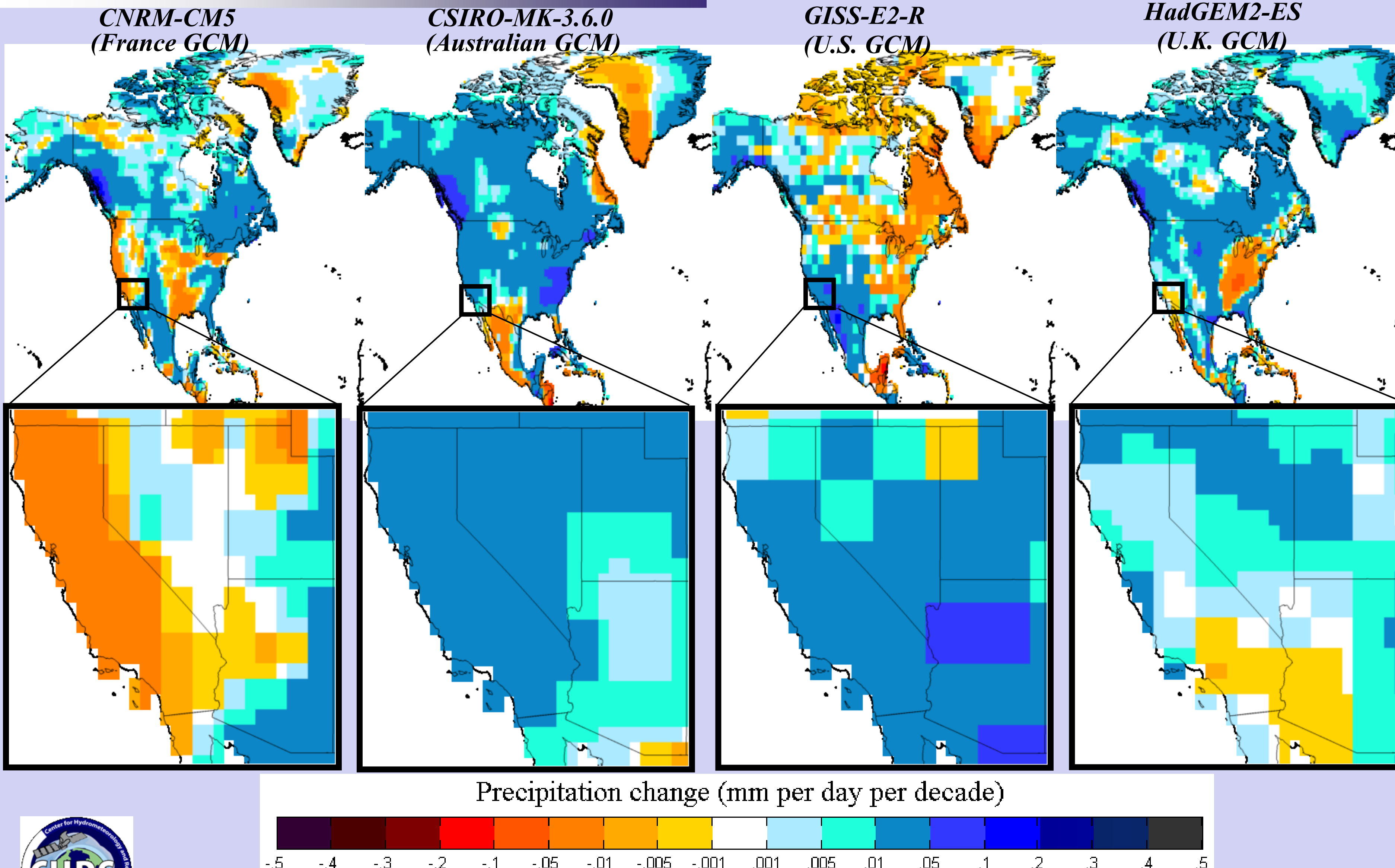
RCP4.5: represent 'medium' scenarios featured by the radiative forcing of 4.5 W/m^2 by 2100, the resulting CO_2-equivalent concentrations is 538 ppm in the year 2100 .

RCP8.5: represent 'high' scenarios featured by the radiative forcing of 8.5 W/m^2 by 2100, the resulting CO_2-equivalent concentrations is 936 ppm in the year 2100 .



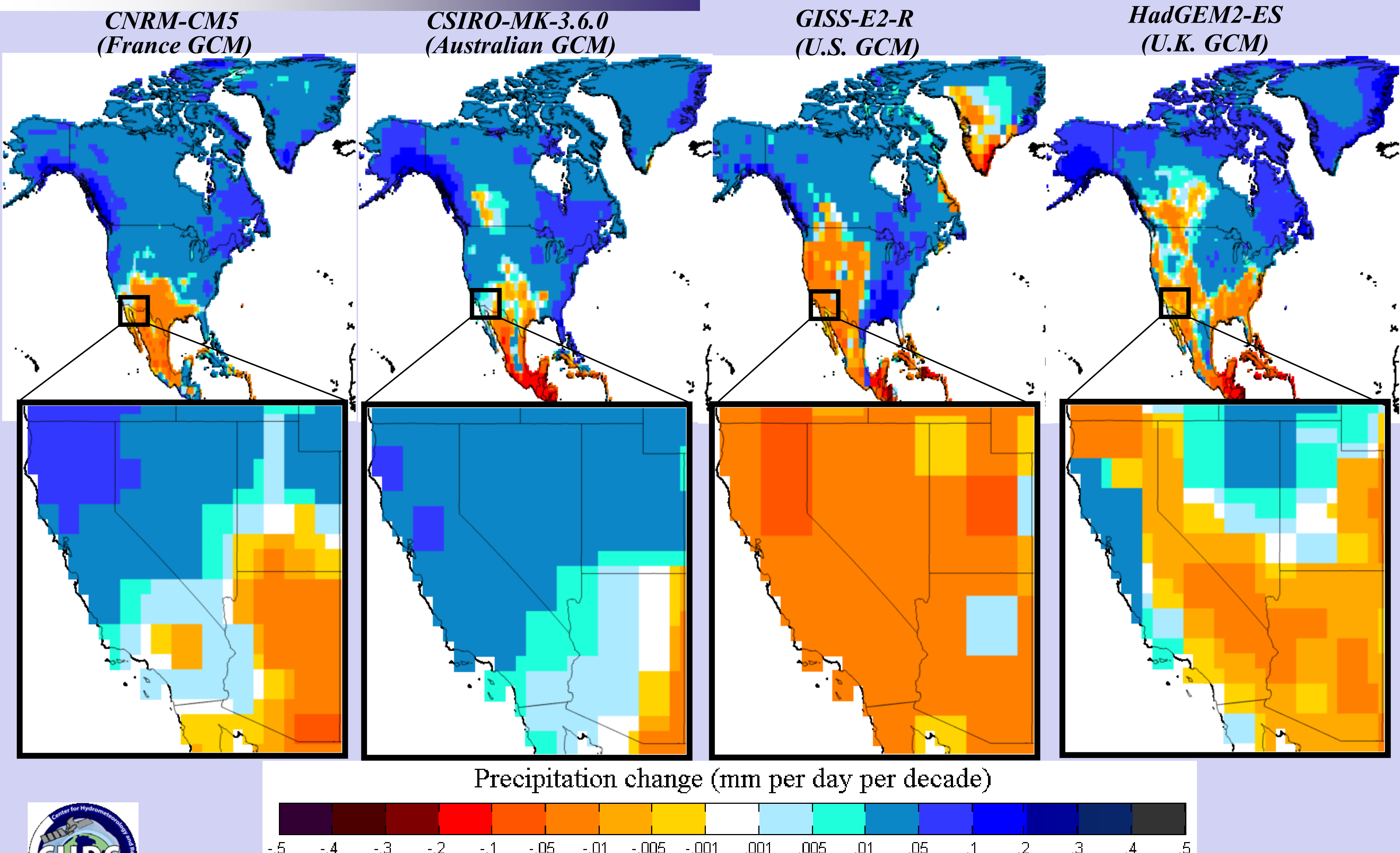
RCP2.6

Time period: 2006-2099



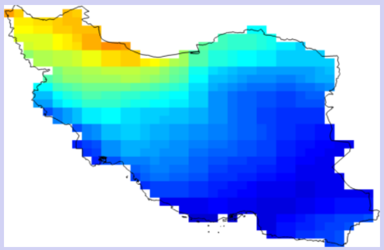
RCP8.5

Time period: 2006-2099

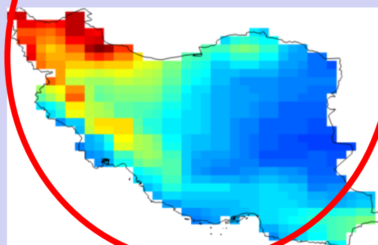


Model historical simulation (1983-2005)

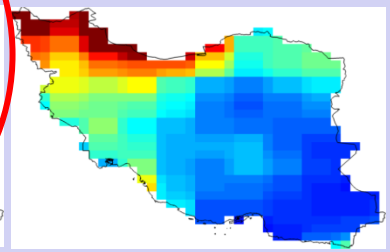
bcc_csm1_1_m
(Chinese GCM)



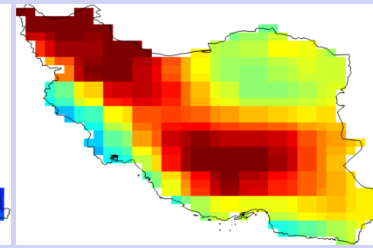
CCSM4
(NCAR, USA GCM)



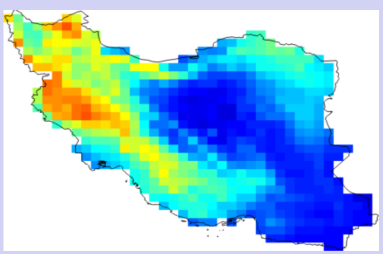
HadGEM2-ES
(U.K GCM)



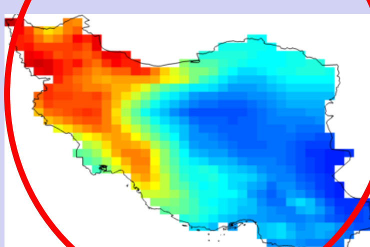
MIROC5
(Japan GCM)

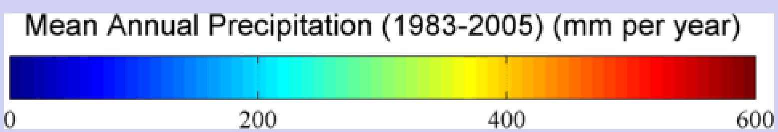


Observation
(CRU Dataset)

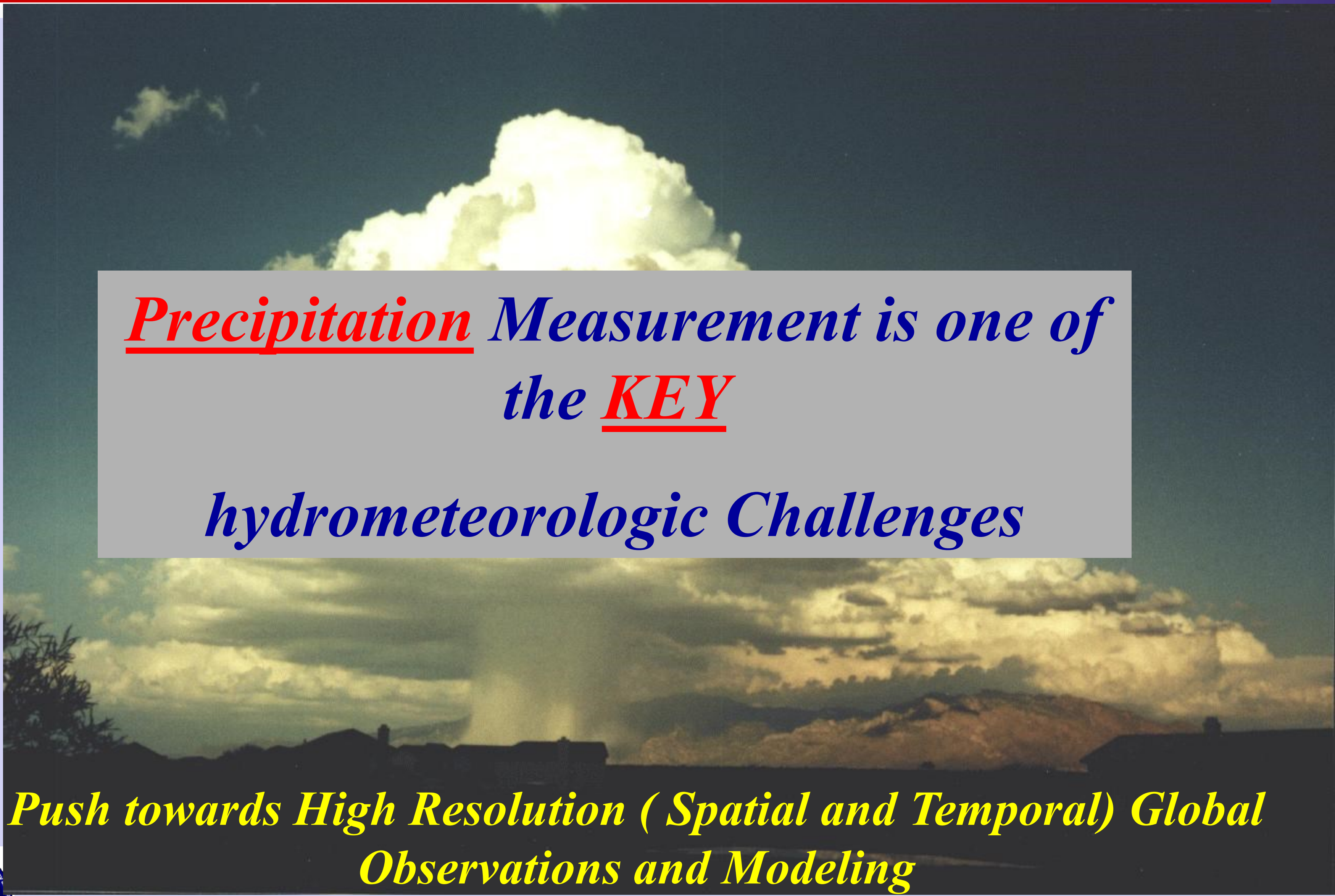


Remotely Sensed Estimates
(PERSIANN-CDR)





A Key Requirement!



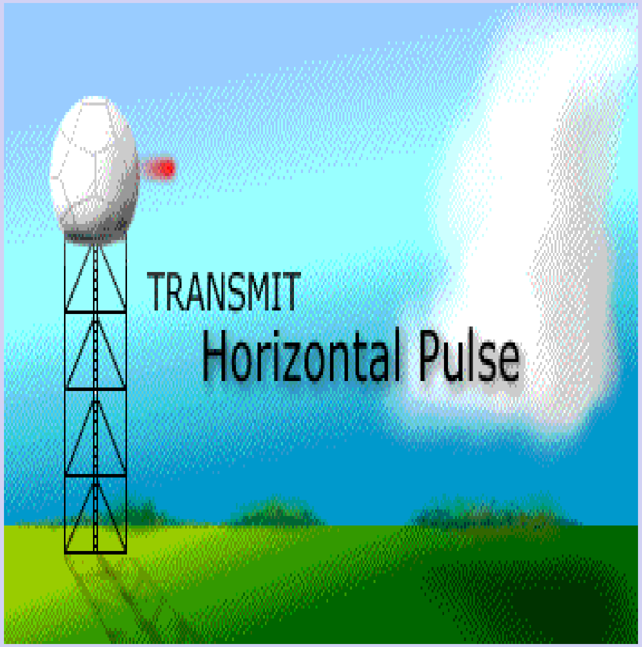
Precipitation Measurement is one of the KEY

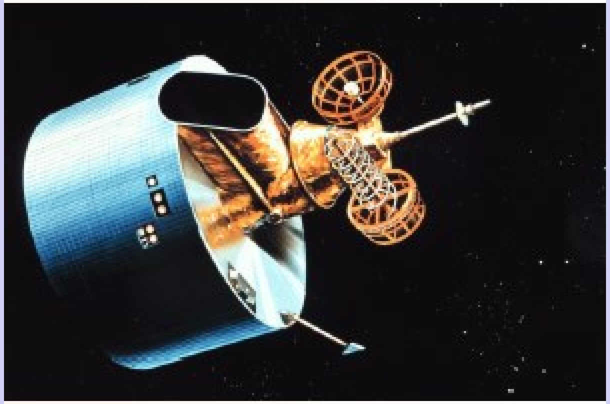
hydrometeorologic Challenges

Push towards High Resolution (Spatial and Temporal) Global Observations and Modeling

Precipitation Observations: Which to trust??

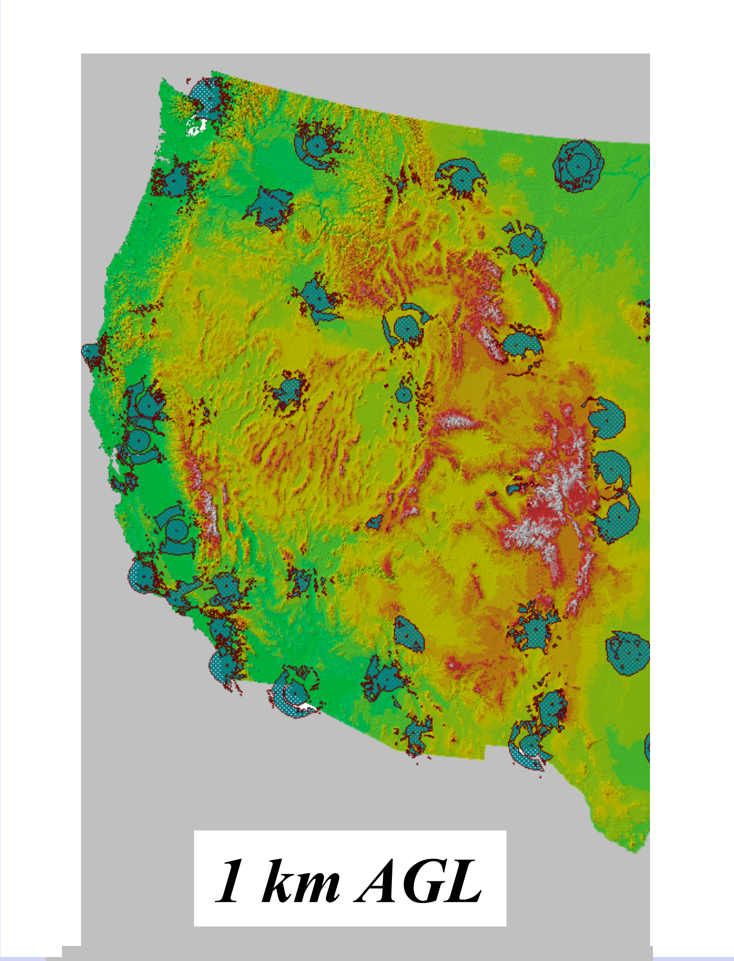
Rain Gauges



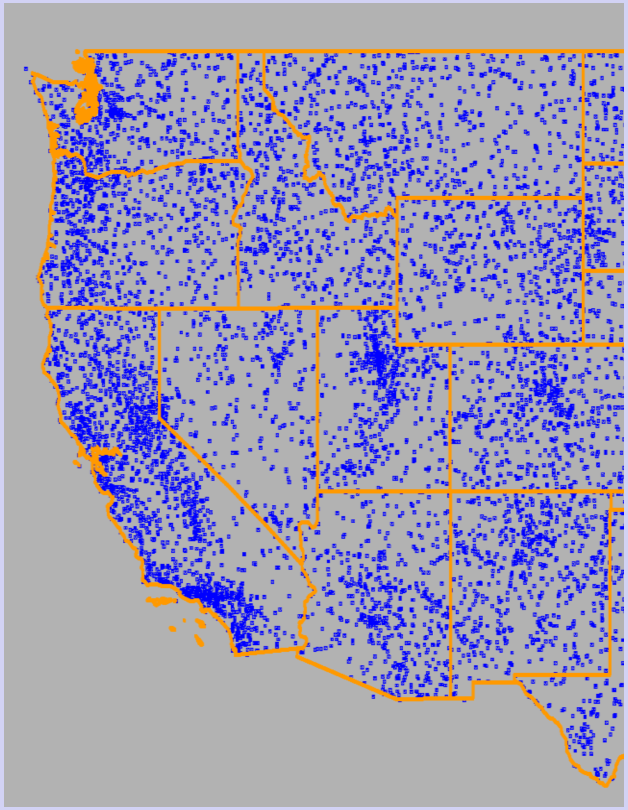


Satellite

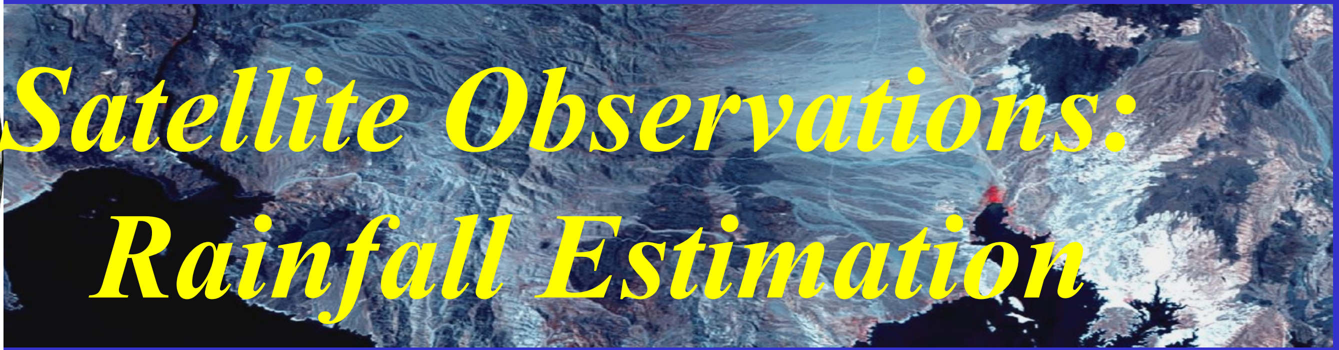
Coverage of the WSR-88D and gauge networks



Maddox, et al., 2002

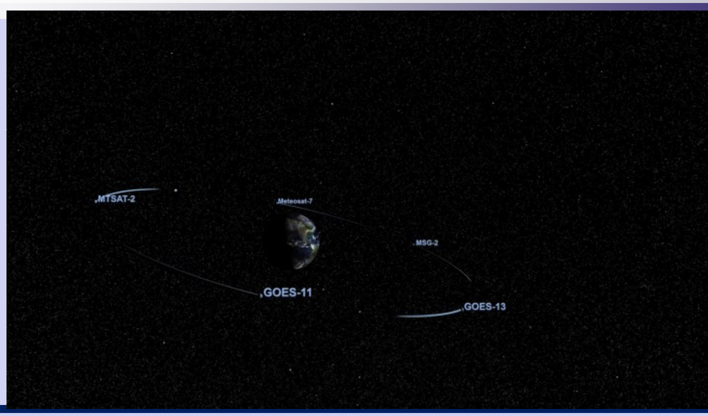


Daily precipitation gages (1 station per 600 km^2 for Colorado River basin) hourly coverage even more sparse

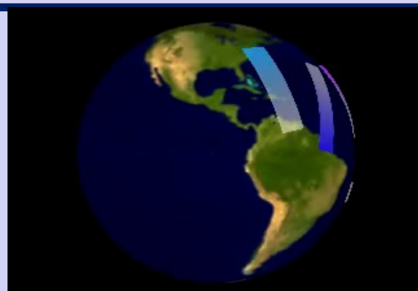


Satellite Observations: Rainfall Estimation

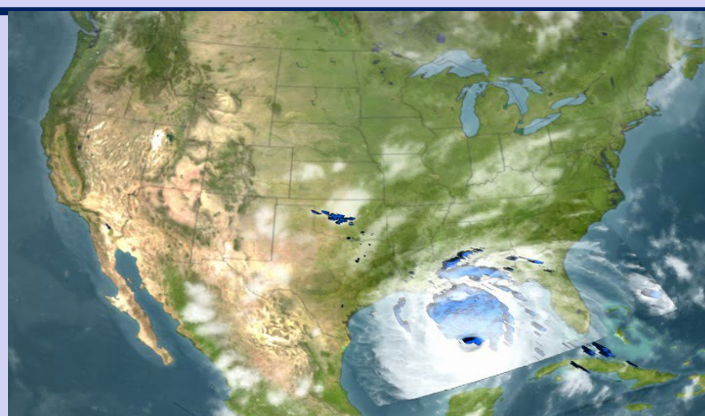
Satellite Data for Precipitation estimation



Geostationary IR
Cloud top data
15-30 minute temporal resolution

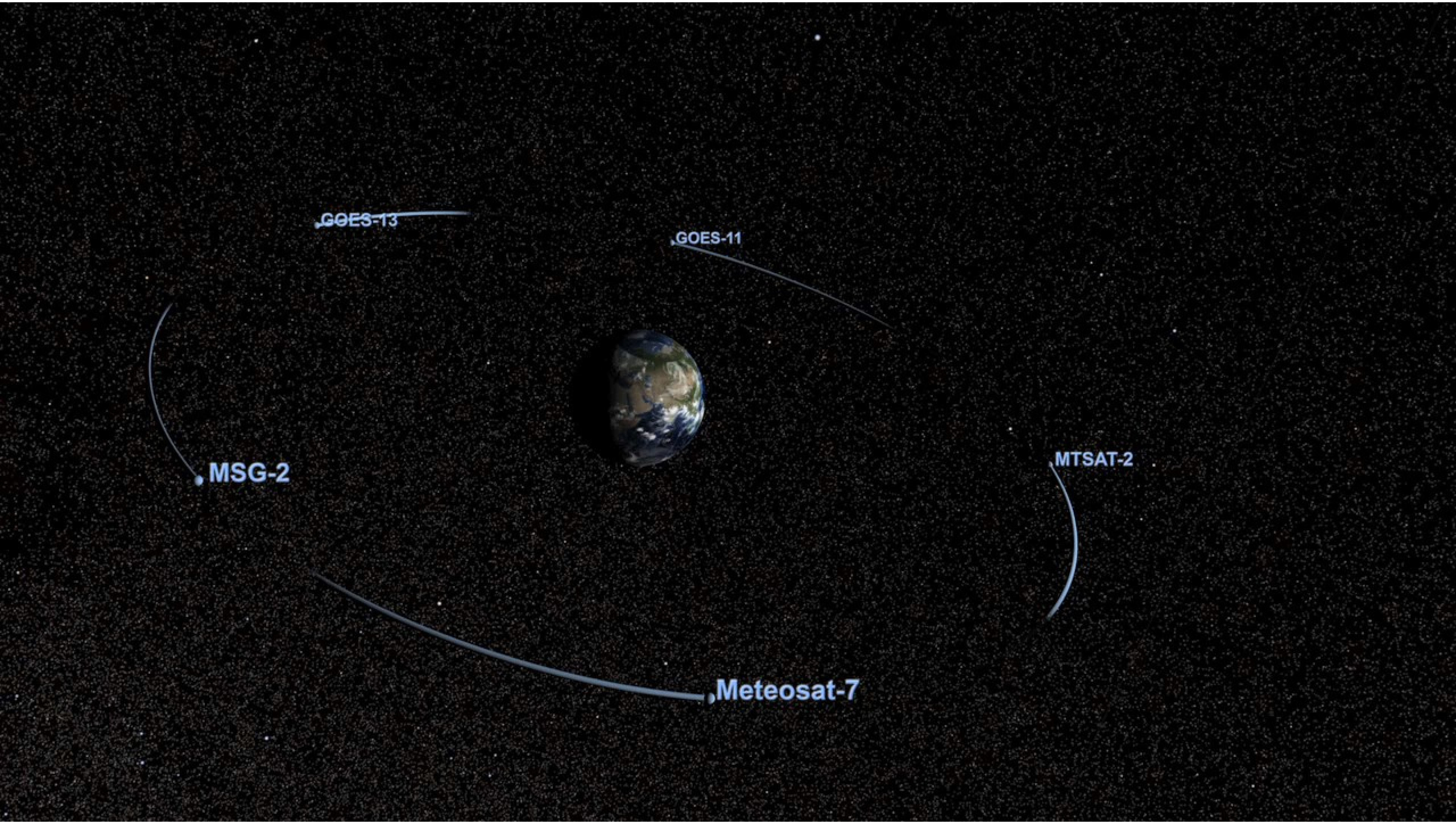


Passive Microwave (SSM/I)
Some characterisation of rainfall
~2 overpasses per day per spacecraft, moving to 3-hour return time (GPM)



TRMM precipitation RADAR
3D imaging of rainfall
1-2 days between overpasses
(S-35° N-35 °)

Geostationary Satellites Infrared (IR) Channel Tb



Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN)

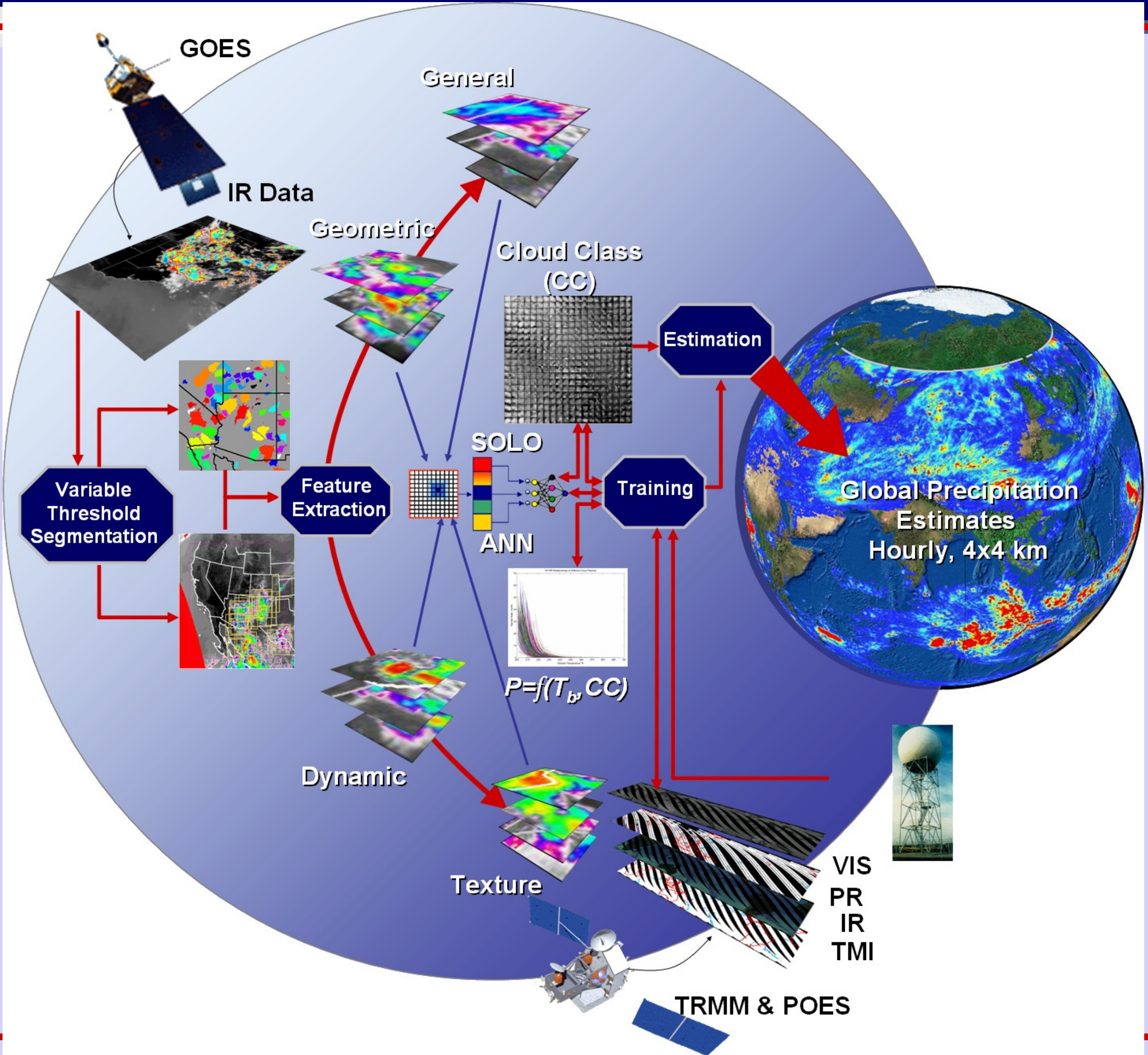
PERSIANN System

Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks

Kuolin Hsu
Algorithm Development

Bisher Imam
G-WADI site development

PERSIANN-CCS (Real-time 4 km)



PERSIANN Extensions: Climate-Related

- ***PERSIANN-CONNECT***
- ***PERSIANN-CDR***

PERSIANN-CONNECT

PERSIANN-CONNected precipitation objECT

EOS®

EOS, TRANSACTIONS, AMERICAN GEOPHYSICAL UNION

VOLUME 94 NUMBER 32 6 AUGUST 2013

Computational Earth Science: Big Data Transformed Into Insight

More than ever in the history of science, researchers have at their fingertips an unprecedented wealth of data from continuously orbiting satellites, weather monitoring instruments, ecological observatories, seismic stations, moored buoys, floats, and even model simulations and forecasts. With just an internet connection, scientists and engineers can access atmospheric and oceanic gridded data and time series observations, seismographs from around the world, minute-by-minute conditions of the near-Earth space environment, and other data streams that provide information on events across local, regional, and global scales. These data sets have become essential for monitoring and understanding the associated impacts of geological and environmental phenomena on society.

This increasing amount of data has led us

If such algorithms are run in a computer environment designed to home in on characteristics of objects or events of interest, then the data can be crunched even more efficiently, allowing insights from big data to be revealed at a quicker pace. Such machine learning evolved from artificial intelligence research and focuses on developing models that are based on the behaviors and characteristics of empirical data. Capturing the behaviors and characteristics from data and determining their underlying probability distributions can provide new knowledge regarding the object or characteristic of interest. Typically, the properties or "true" underlying probability distributions of the observed variable of interest are not explicitly known. However, by seeking to define or describe these underlying probability distributions, data mining can help scientists

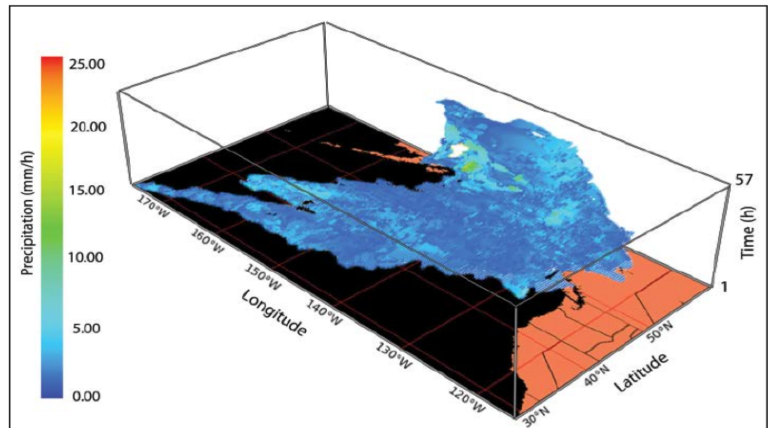
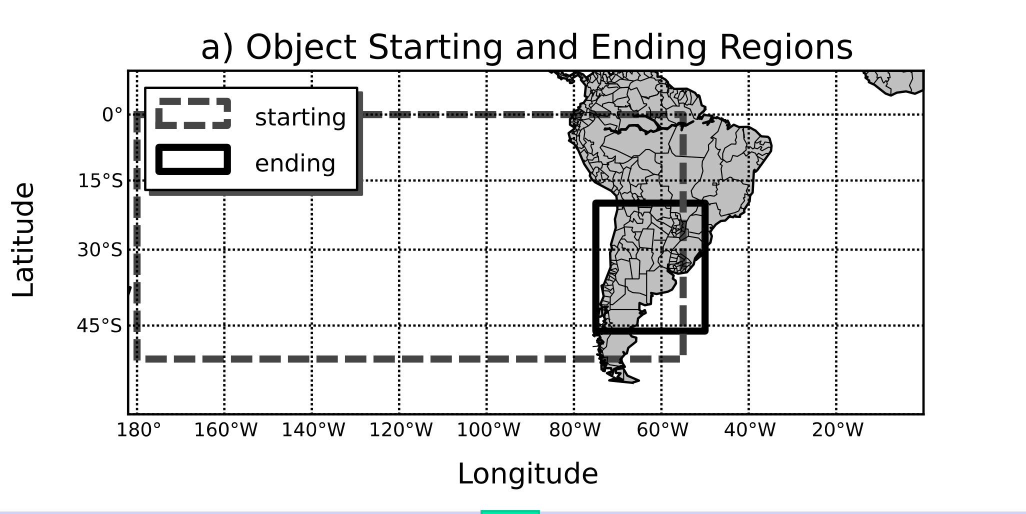
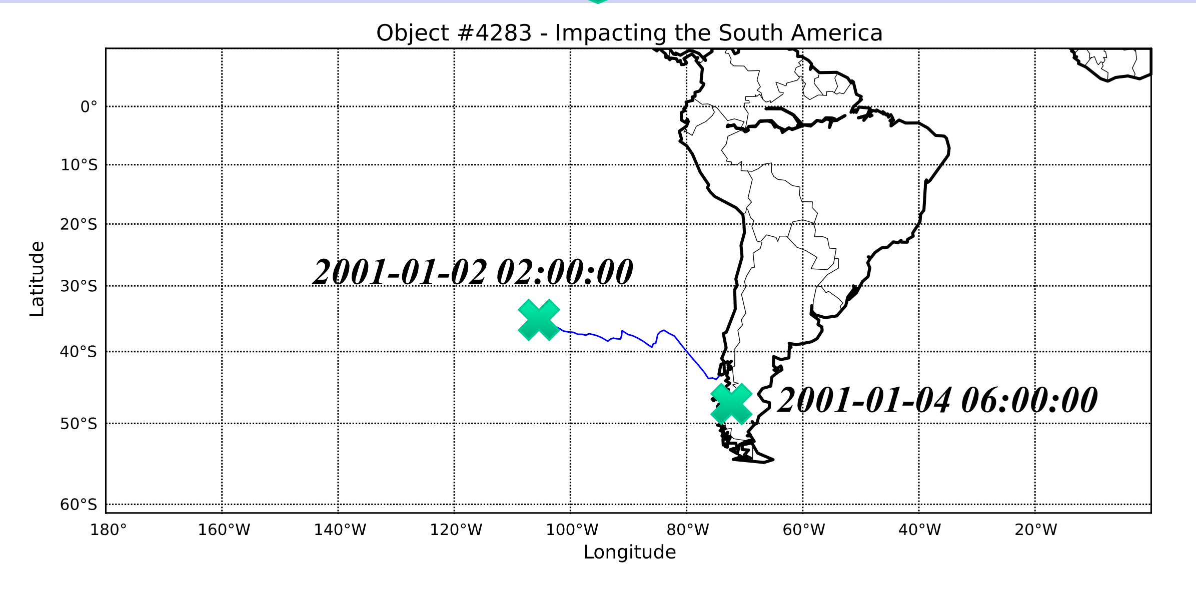


Fig. 1. A connected four-dimensional atmospheric river, or "precipitation object," extracted from the PostgreSQL database. The atmospheric river originated in the eastern Pacific and affected the western United States from 28 to 30 December 2005.

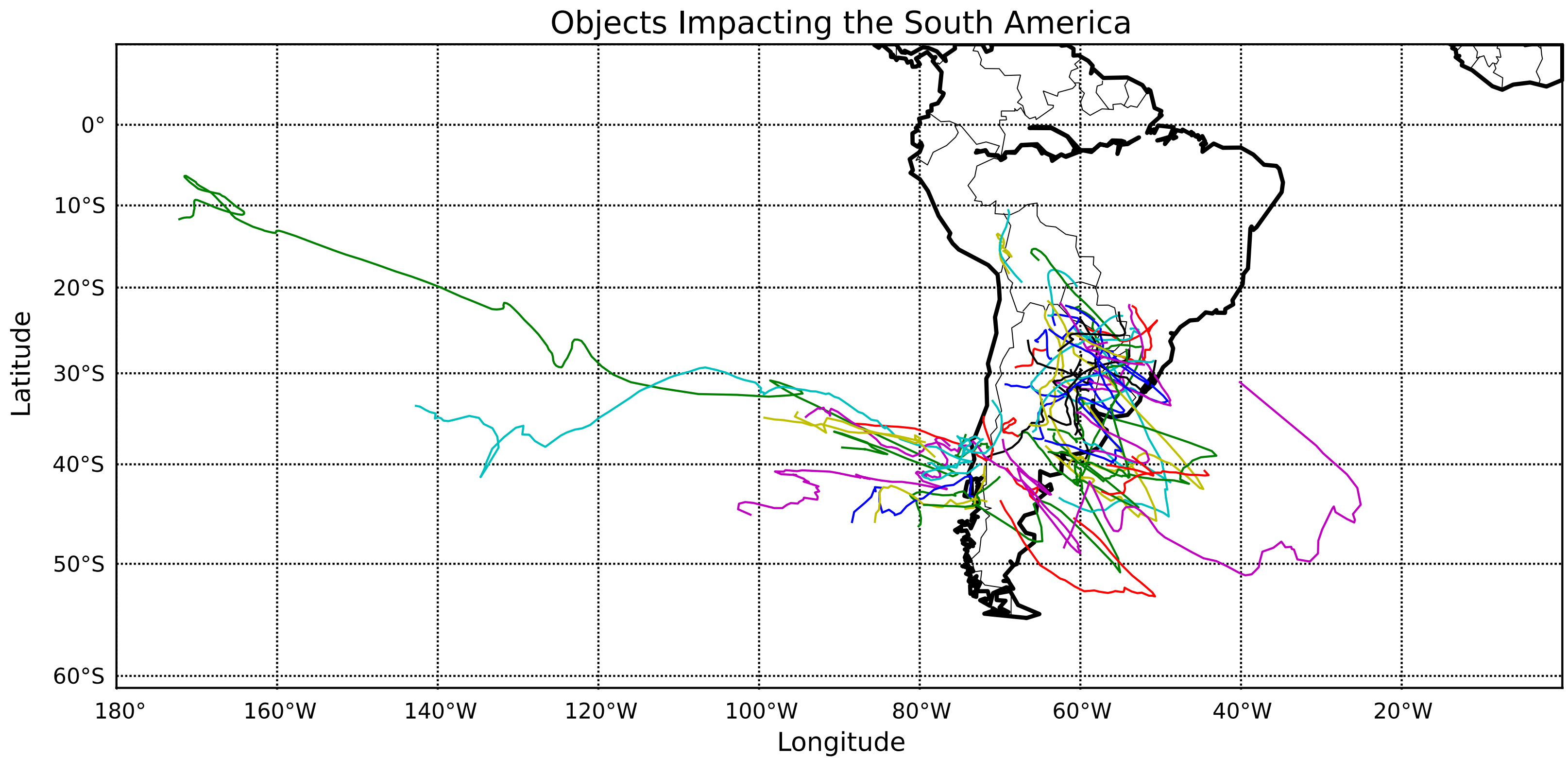
***Sellars, S., P. Nguyen, W. Chu, X. Gao, K. Hsu, and S. Sorooshian (2013), Computational Earth Science: Big Data Transformed Into Insight, EOS Trans. AGU, 94(32),277**

Regional Search: Specific Storm

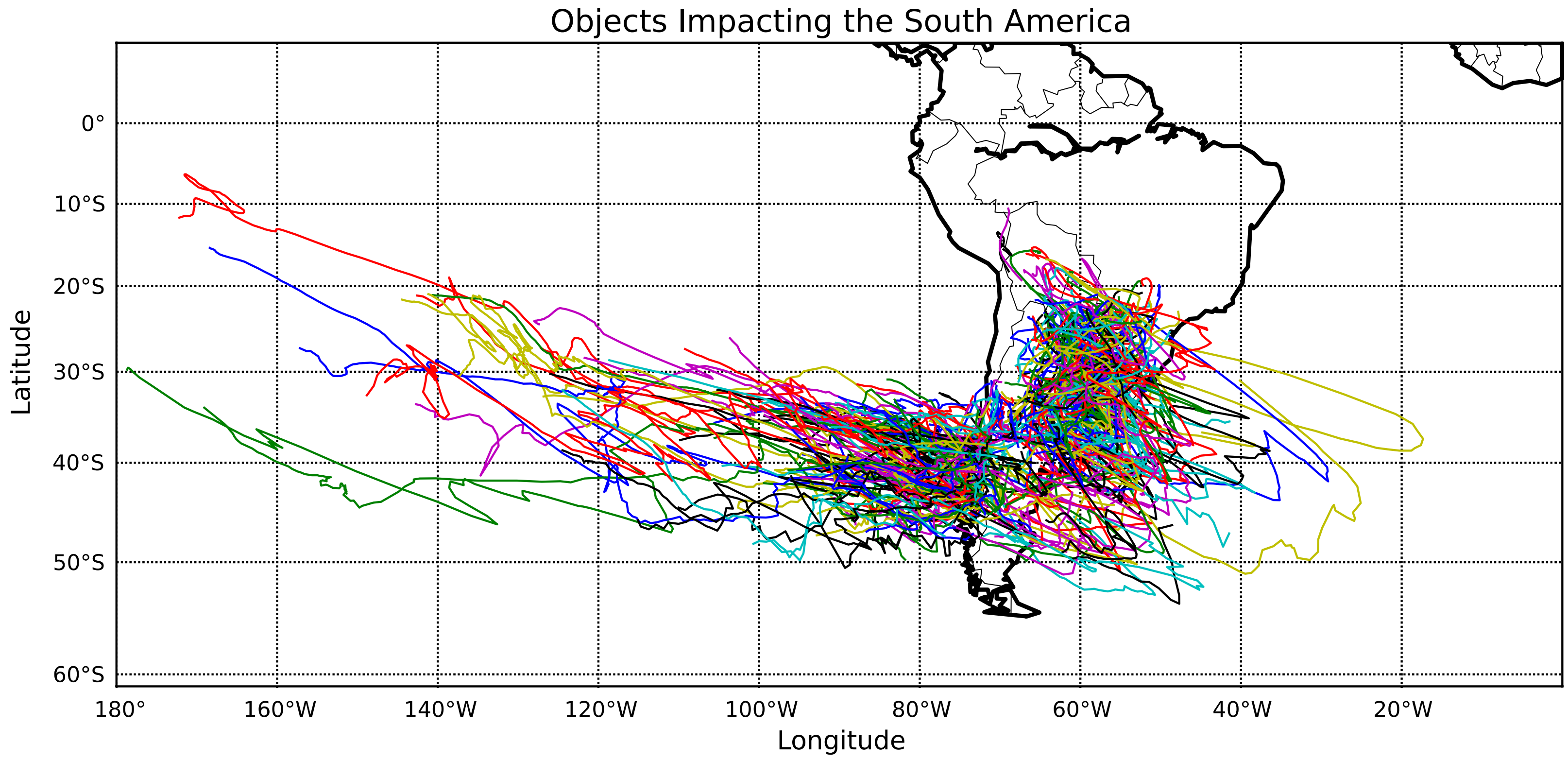




Regional Search: 2001 Storms



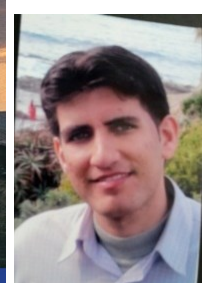
Regional Search: All Storms (2000-2010)

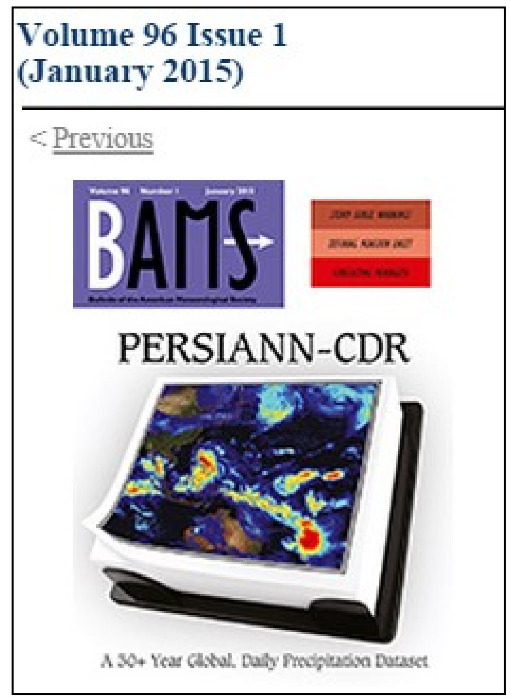


PERSIANN -CDR

http://www.ncdc.noaa.gov/cdr/operationalcdrs.html

- *Daily Precipitation Data*
- *Data Period: 1983~2014*
- *Coverage: 60^{o}S ~ 60^{o}N*
- *Spatial Resolution: 0.25^{o}x0.25^o*





Ashouri, Hsu et al., BAMS, 2015.

PERSIANN-CDR: PERSIANN Climate Data Record

http://www.ncdc.noaa.gov/cdr/operationalcdrs.html

NOAA NATIONAL CLIMATIC DATA CENTER

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Home | Operational CDRs | Developmental CDRs | Opportunities | Library | SEARCH NCDC

CLIMATE DATA RECORD PROGRAM

- Serving the Public
- Data
- Development Guidelines
- Contact Us

News

Climate Data and Applications Workshop - A Focus on Precipitation - Dec 3-4, 2013

Congratulations Cheng-Zhi Zou

2013 CDR Annual Meetings Presentations now available

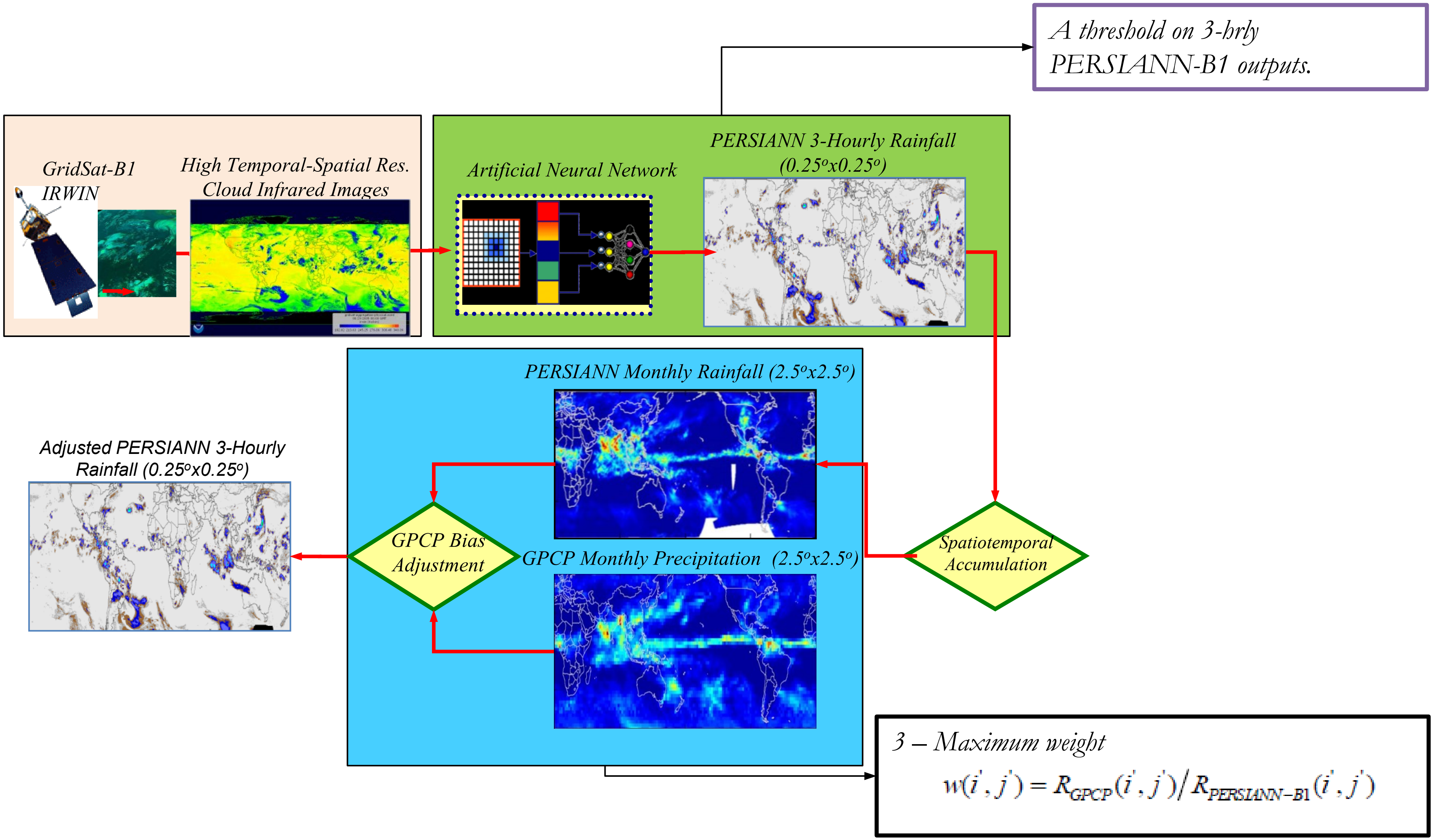
Operational Climate Data Records (CDRs)

In addition to embracing the National Research Council CDR definition (Climate Data Records from Environmental Satellites: Interim Report 2004) , NOAA operational CDRs are routinely assessed for quality and systematically generated. The first step in establishing an operational CDR includes public posting of the source code that generated the CDR dataset, the dataset itself, and supporting documentation through a six-phase Research-to-Operations process that is described in the Developers Guidelines.

Once posted to the NCDC webpage, the CDRs are grouped by Fundamental CDRs and Thematic (Atmospheric, Oceanic, and Terrestrial) CDRs. Fundamental CDRs are sensor data (e.g. calibrated radiances, brightness temperatures) that have been improved and quality controlled over time, together with the ancillary data used to calibrate them. Thematic CDRs are geophysical variables derived from the FCDRs, such as sea surface temperature and sea ice concentration, and they are specific to various disciplines. Thematic CDRs are often generated by blending satellite observations, in-situ data, and/or model output.

Atmospheric CDRs	Information	Serving Public	Available Data Access	Source Code	Documentation
Aerosol Optical Thickness	Overview Contact Us Registration		Use Agreement THREDDS FTP		Algorithm Description Data Flow Diagram Maturity Matrix
Mean Layer Temperatures - RSS	Overview Contact Us Registration		Use Agreement THREDDS FTP		Algorithm Description Data Flow Diagram Maturity Matrix
Mean Layer Temperatures - UAH	Overview Contact Us Registration		Use Agreement THREDDS FTP		Algorithm Description Data Flow Diagram Maturity Matrix
Outgoing Longwave Radiation	Overview Contact Us Registration		Use Agreement THREDDS FTP		Algorithm Description Data Flow Diagram Maturity Matrix
Precipitation - PERSIANN-CDR	Overview Contact Us Registration		Use Agreement THREDDS FTP		Algorithm Description Data Flow Diagram Maturity Matrix

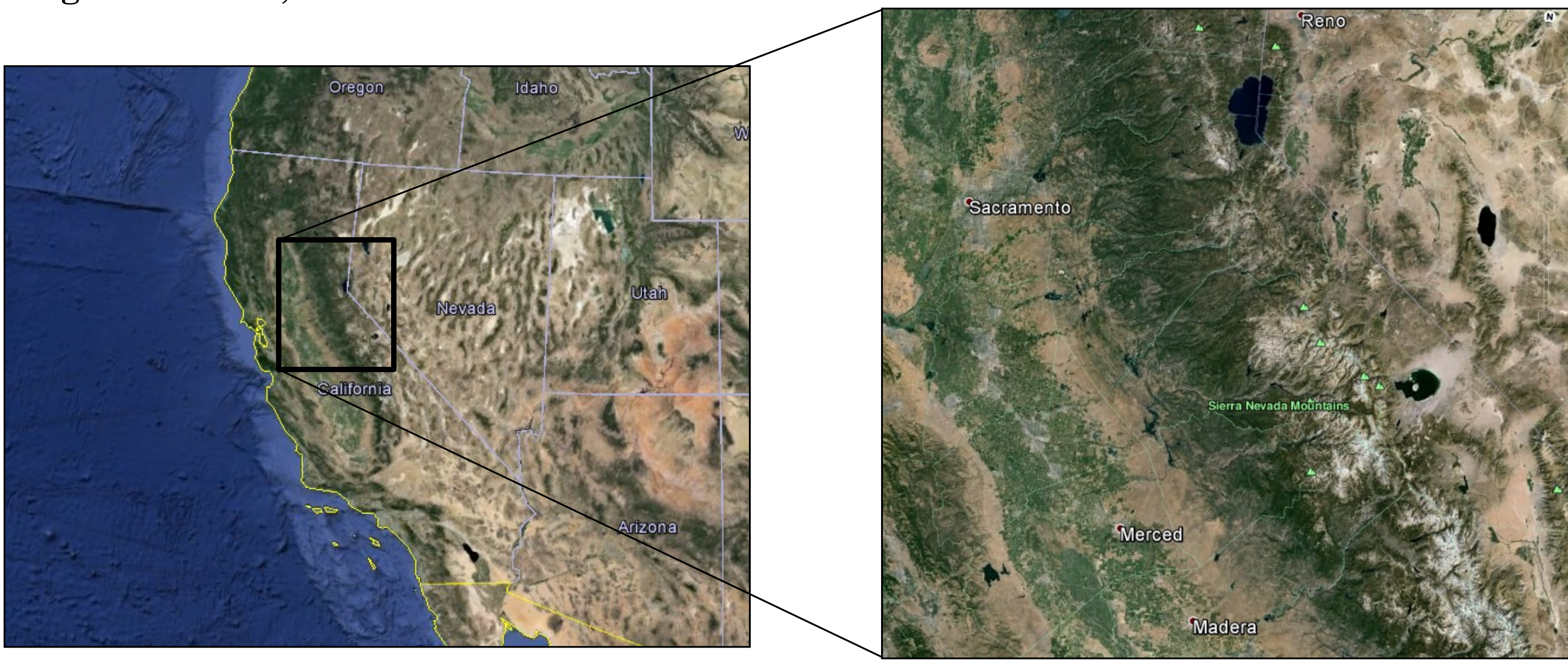
PERSIANN-CDR Algorithm



Sierra-Nevada Mountain Region

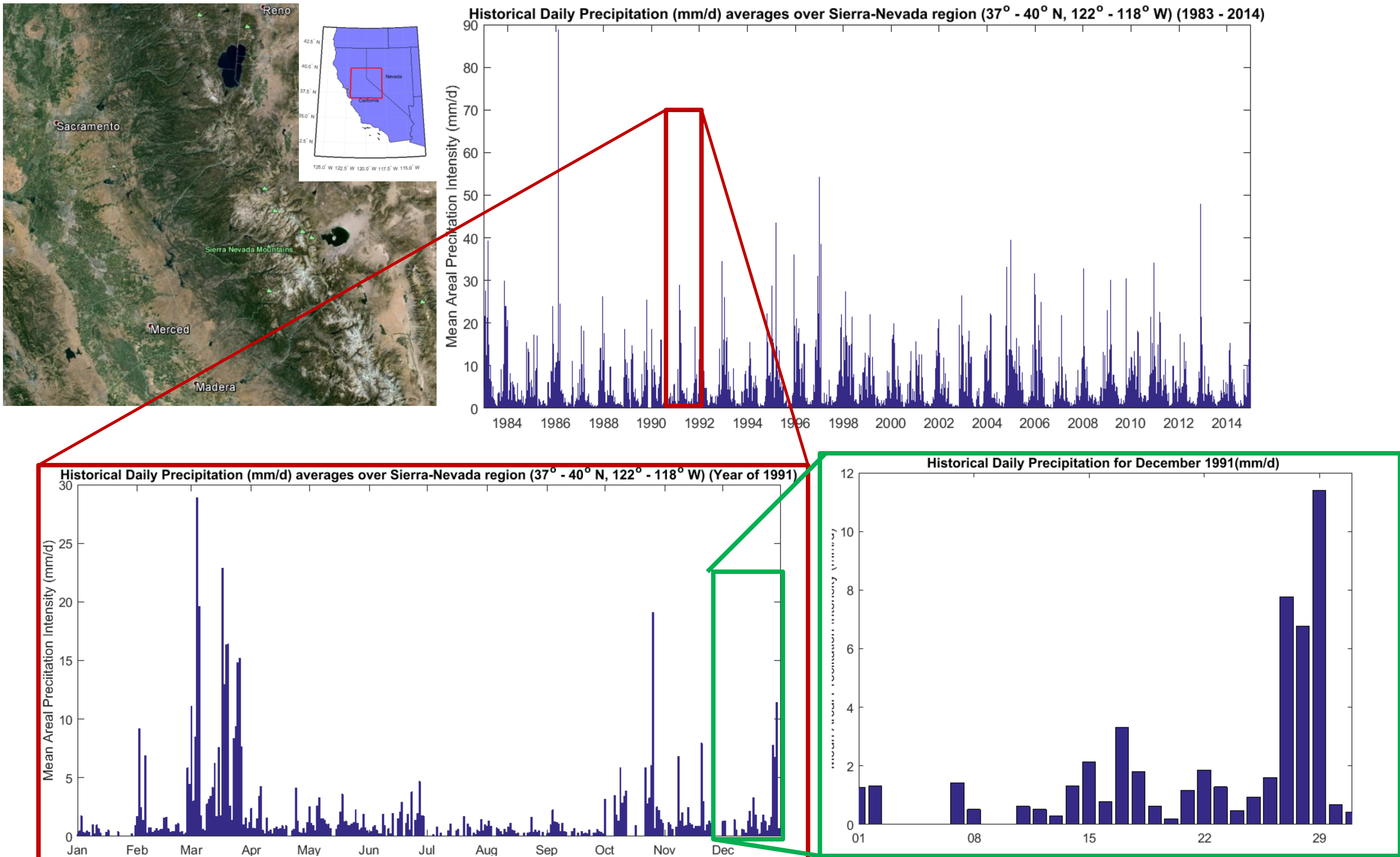
Area: 63,100 square kilometers (24,370 sq mi)

Length: 400 mile, Width: 64 mile.



Source: Google Earth

Sierra-Nevada Mountain (California and Nevada)



DEMO

CHRS CONNECT Database

Center for Hydrometeorology & Remote Sensing
University of California, Irvine

About | Research Areas | Research Projects | Resources | GWADI

CHRS Home > Research Areas > Voxel

PERSIANN - CONNected precipitation objECT (CONNECT) Dataset

Location

North

60

West

0

East

360

-60

South

Time

Please define the start and end period (YYYY-MM-DD HH:MM:SS):

Start Date:

End Date:

Intensity

Please set minimum and maximum precipitation intensity criteria (mm/hr):

Min Intensity 1

Max Intensity 200

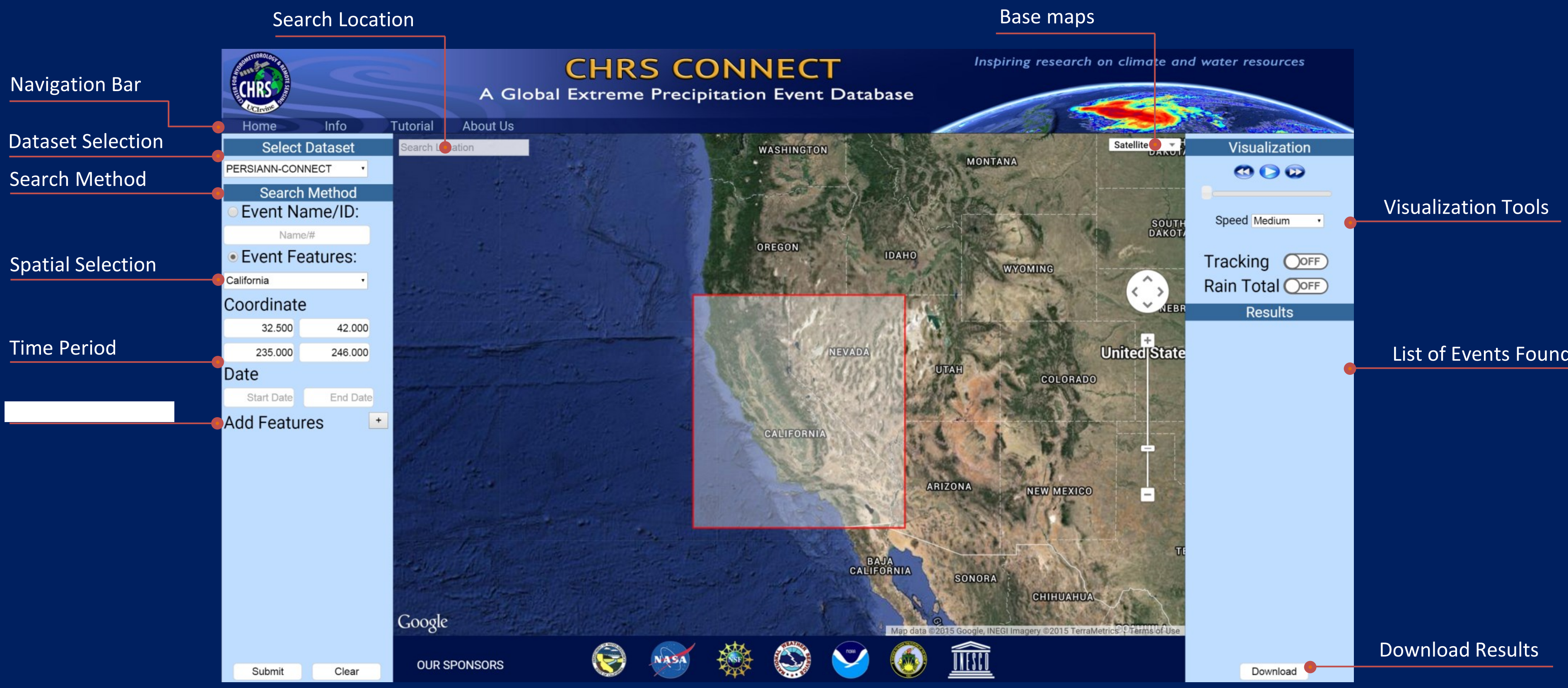
Download Options

Download Options:

- [x] Download raw object data
- [x] Download statistics data

Submit

CHRS CONNECT Interface



Search Extreme Precipitation Events

Extreme precipitation events can be searched using event IDs or features. Event IDs are the unique numbers indicating each event in the database.

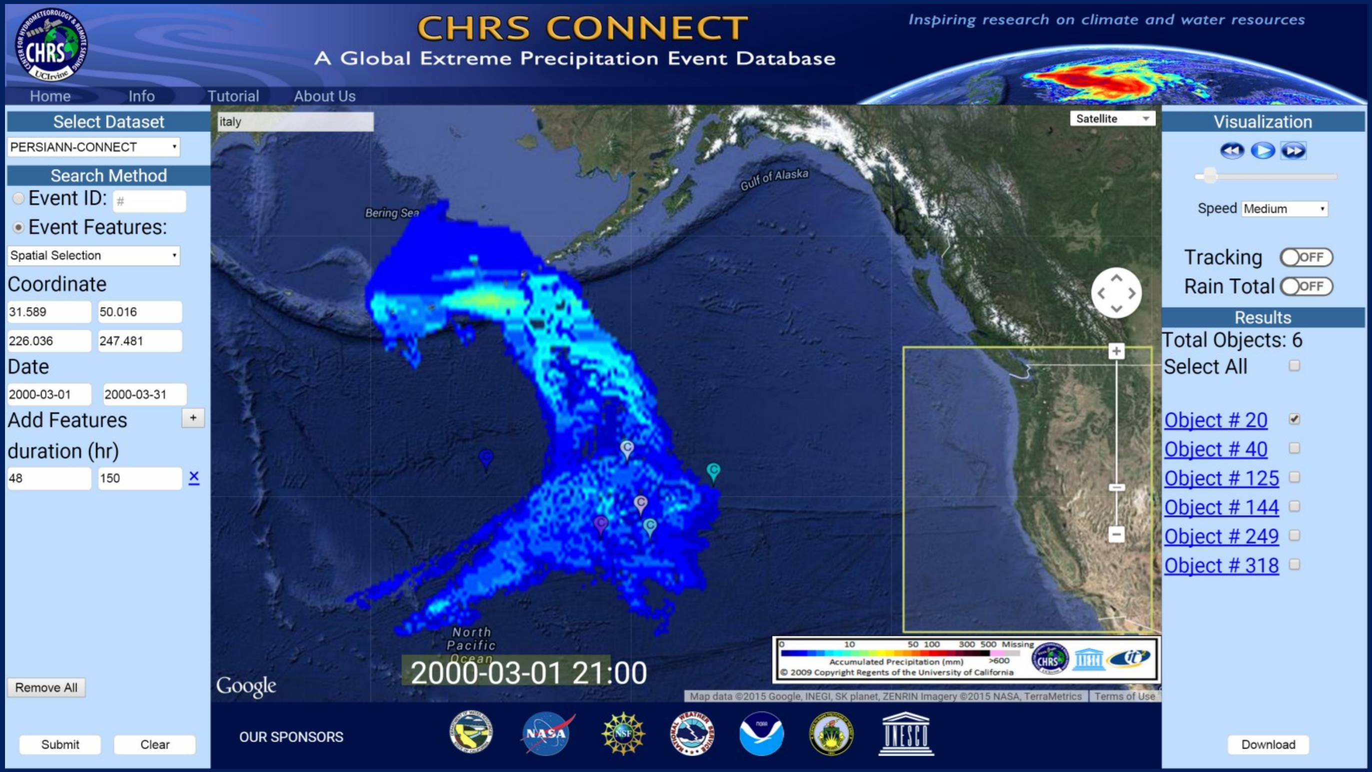
Steps to search events using event features:

- Select a dataset (PERSIANN-CONNECT or PERSIANN-CDR-CONNECT)
- Select a region by using Spatial Selection → Select One Region/Start-End Region → Draw a region (if One Region selected) or 2 regions (if Start-End Region selected).
- Select Start and End dates
- Add event features for searching (optional)
- Submit
- Extreme precipitation events will appear as markers on the map and on the result tab

Interactive Functionalities for Events

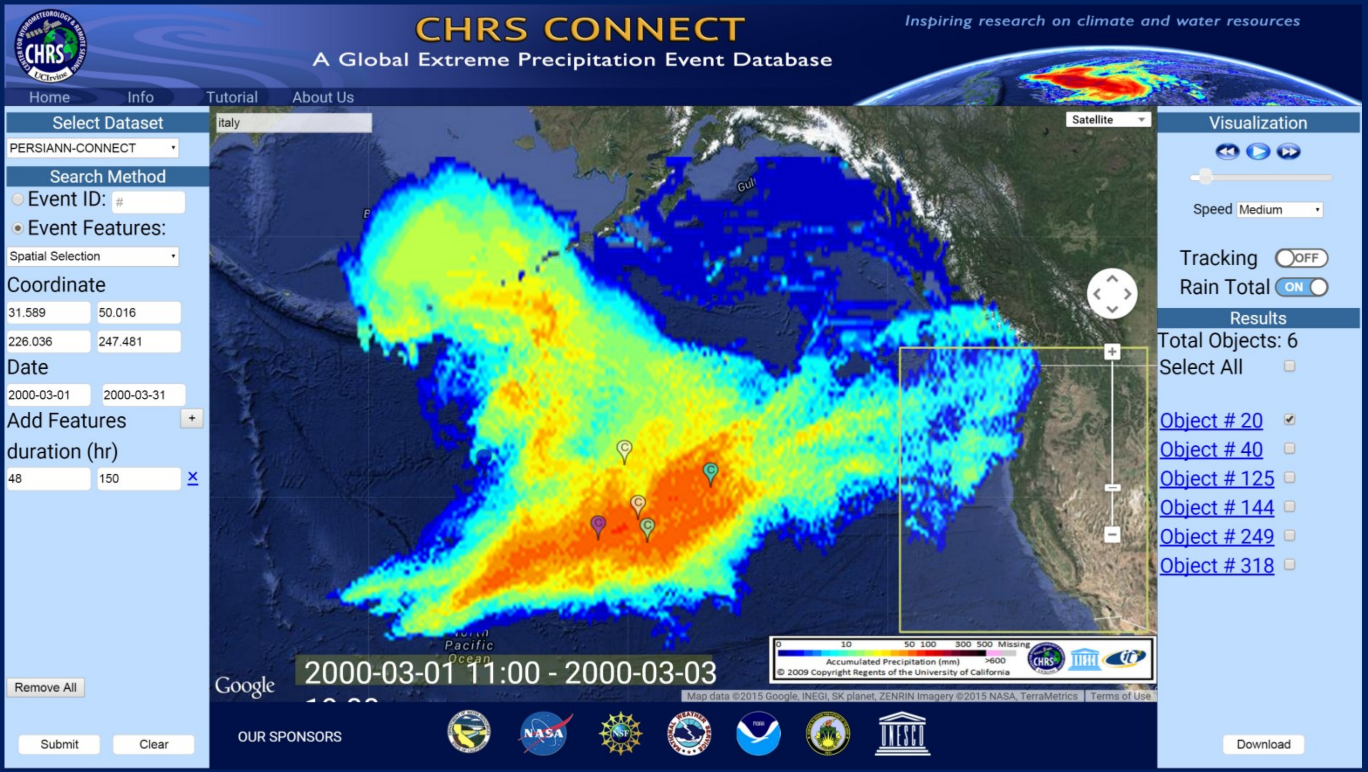
- ➢ **Select an event by clicking the checkbox next to the object**
- ➢ **Users can show either animation, tracking, or rain total of an event**
- ➢ **Hover over an event marker and the same event on the result tab will be highlighted, vice-versa**

Animation



- Select an event to animate (only 1 event can be animated at a time)
- Click on the Play button to animate the event
- Event information and a legend will appear while the event is animating
- Change speed of an event animation by selecting a speed mode (slow, medium, fast)

Rain Total



- ➢ Select an event to show rain total (only 1 event can be shown at a time)
- ➢ Turn Rain Total button on to show the rain total
- ➢ Event rain total will appear on the map along with a legend

Tracking

➢ Select event(s) to show tracking (multiple event can be selected)

➢ Turn Tracking button on to show tracking of selected events

➢ Event tracks will be displayed on the map

Statistics

- Users can click on a marker to display statistic table of an event
- Users can also click on an event on the Result tab to display statistics
- Even statistics can be downloaded by selecting events and clicking on the Download button

Download

- Download dialog box will appear after clicking on Download button
- Type in an email address and click Submit button to receive event statistics
- Users can just click on Submit button without an email address to download event statistics straight to the local computer

Search Events by Names

Search Location

- ➢ **Type in a location**
- ➢ **The map will center to that location**

UCI PERSIANN-CONNECT on Studying the Atmospheric River Events

WSWC/CDWR Workshop on Sub-Seasonal and Seasonal Precipitation Forecasting

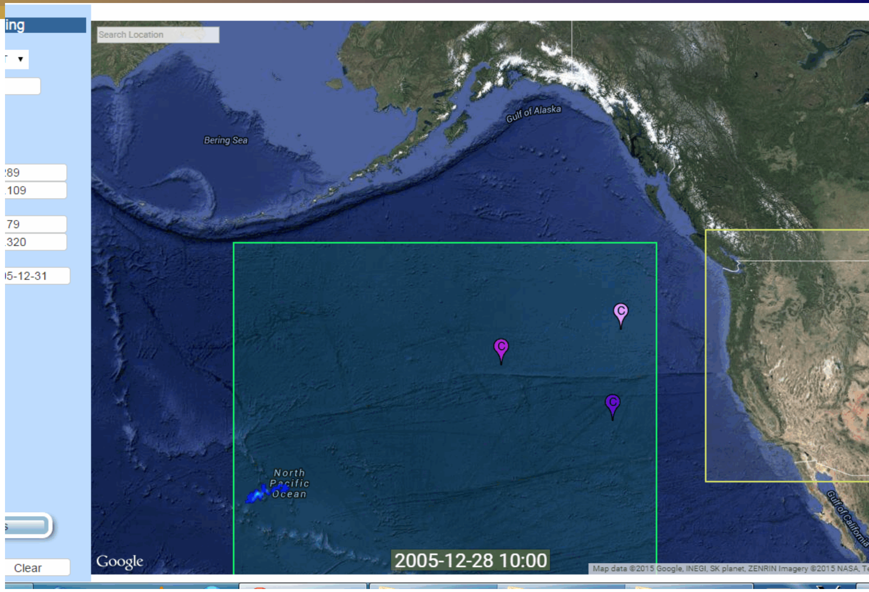
May 28th, 2015

Soroosh Sorooshian, Phu Nguyen, Hao Liu

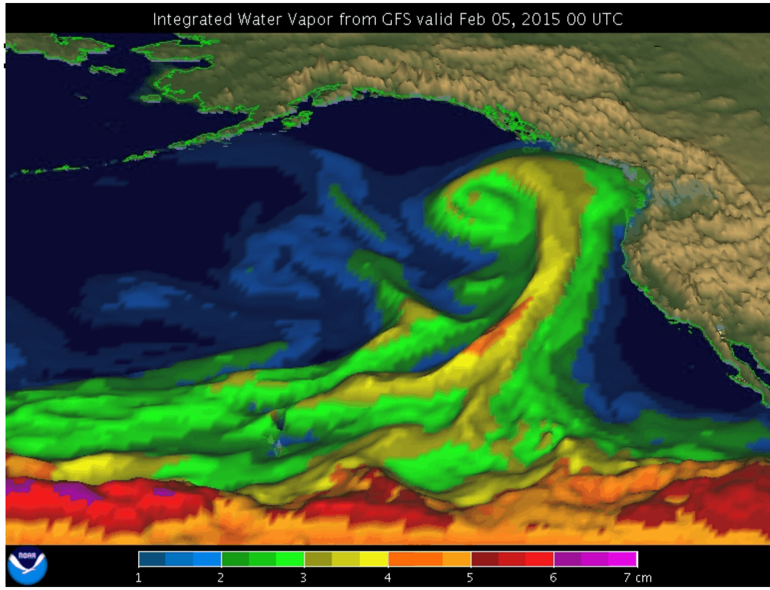
Department of Civil and Environmental Engineering
University of California, Irvine

Background: PERSIANN-CONNECT and Atmospheric Rivers

- PERSIANN-CONNECT's advantages:
  - High temporal-resolution
  - Object-oriented database
  - Track-based and regional search function
  - Unique event ID
- PERSIANN-CONNECT's functionalities:
  - Archive and retrieve the historic events
  - Visualization
  - Scientific research



- Atmospheric River (AR, Dettinger et al, 2011):
  - Water vapor flux (using polar-orbit satellite)
  - Water vapor band that is above a threshold
  - 1000 km wide, 2000 km long
  - Contribute 30 ~50% of California's precipitation



Building a Complete AR Database

- AR event database: PERSIANN-CONNECT and other available resources
- UCI PRESIANN-CONNECT:
  - Precipitation- based AR database (under-developing)
- UCSD-SCRIPPS (Dettinger et al. 2011):
  - SSM/I-based AR events dates
  - Flux of integreted water vapor flux (1000km wide, 2000 km long,)
  - 162 days in CA (1998-2008)
  - RS daily-average water vapor flux
- NOAA ESRL PSD (http://www.esrl.noaa.gov/psd/atmrivers/events/):
  - Research publication –based AR database
  - 8 AR events in CA (1986-2010)
- CA-DWR & U.S. NWS historical flooding events
  - For cross-reference purposes
  - 1193 flooding records in CA (1806-2013)

AR events retrieved from the PERSIANN-CONNECT

NOAA-ESRL-PSD AR Events	PERSIANN-CONNECT	Dettinger et al, 2011	CA-DWR & NWS /Flooding Record
• N/A (No published paper)	• Object id: 13661	• 2002-Dec-13 • 2002-Dec-14 • 2002-Dec-16	• 2002, Dec. 13-21 • Santa Clara, Riverside, Napa,
• 2004 Feb. 16-18	• Object id: 19404	• 2004-Feb-16 • 2004-Feb-17	• 2004, Feb. 18 • Riverside, Flash
• 2005 Jan. 7-11	• Object id: 23945 • Object id: 23916	• 2005-Jan-09 • 2005-Jan-10	• 2005, Jan. • Orange, 127
• 2005 Dec. 28-29	• Object id: 28708	• 2005-Dec-22 • 2005-Dec-28	• New Year's Eve Flood
• 2009 Oct. 13-14	• Object id: 48945	• N/A (out of date range)	• 2009 Oct. • Monterey, 13
• 2010 Feb. 5-6	• Object id:50394	• N/A (out of date range)	• 2010, Feb. • Los Angeles, Orange, Riverside, San

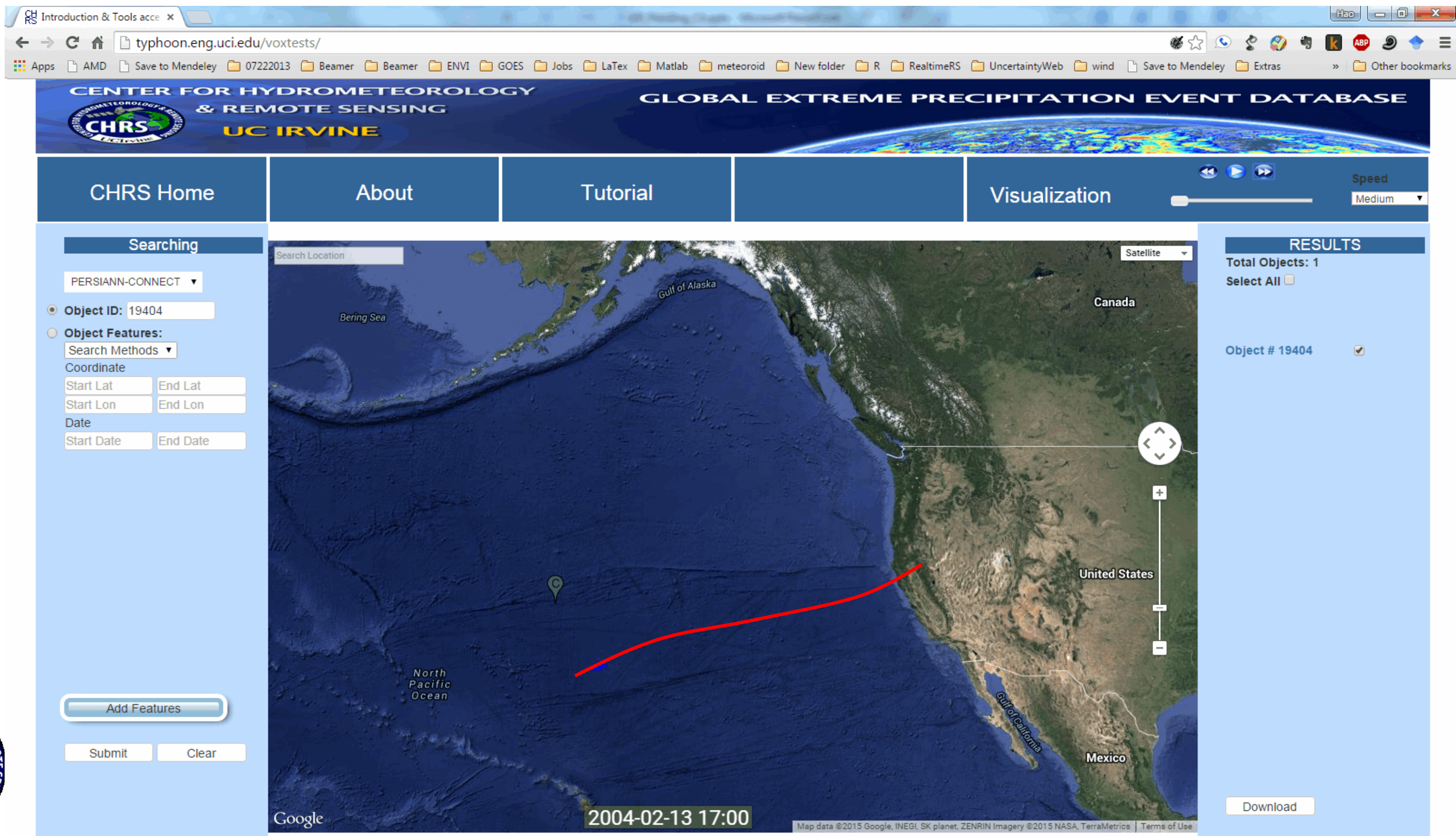
AR events visualization using PERSIANN-CONNECT: 2004 Feb.

NOAA-ESRL-PSD AR Events
- 2004 Feb. 16-18

PERSIANN-CONNECT
- Object id: 19404

Dettinger et al, 2011
- 2004-Feb-16
- 2004-Feb-17

CA-DWR Flooding Record
- 2004, Feb. 18
- Riverside, Flash



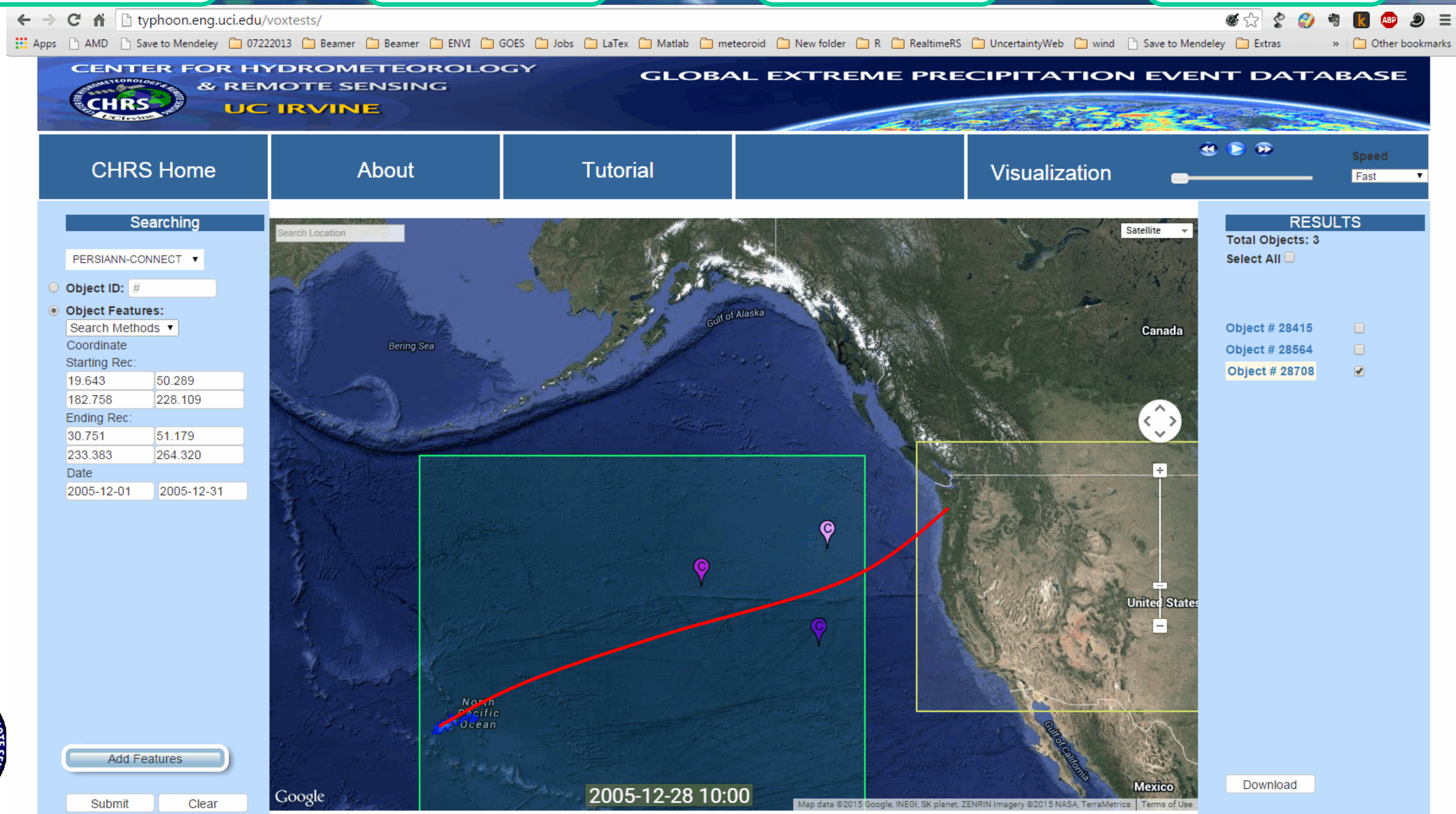
AR event visualization using PESIANN-CONNECT: 2005 Dec.

NOAA-ESRL-PSD AR Events
- 2005 Dec. 28-29

PERSIANN-CONNECT
- Object id: 28708

Dettinger et al, 2011
- 2005-Dec-22
- 2005-Dec-28
- 2005-Dec-30
- 2005-Dec-31

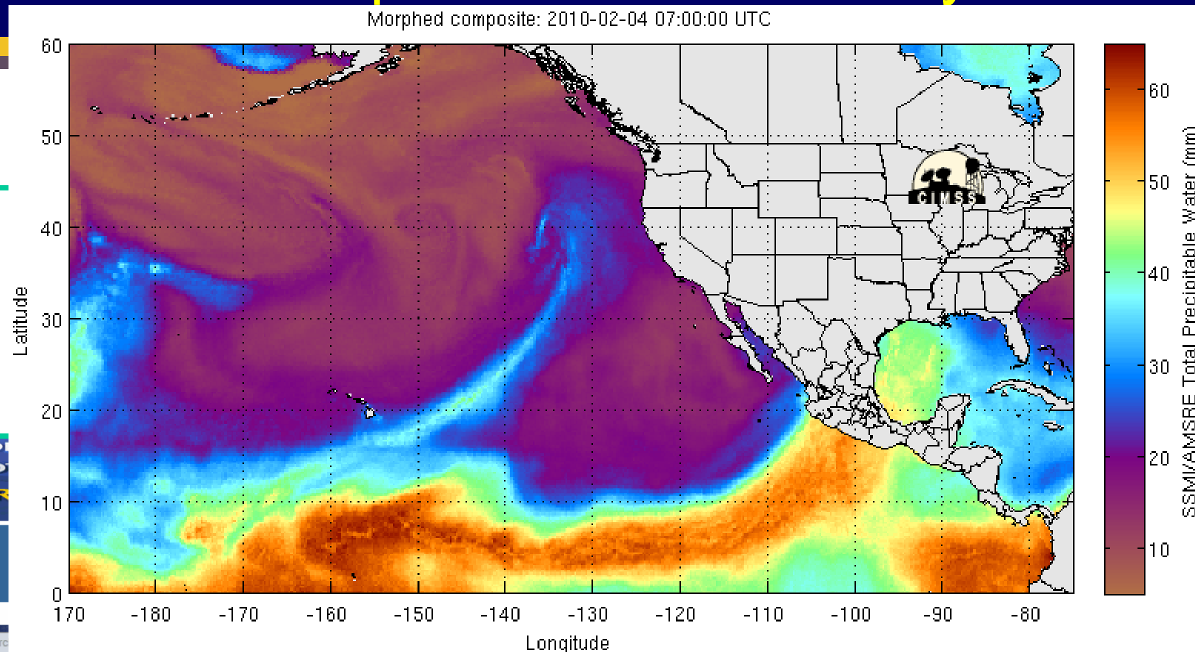
CA-DWR Flooding Record
- New Year's Eve Flood
- 300 million statewide

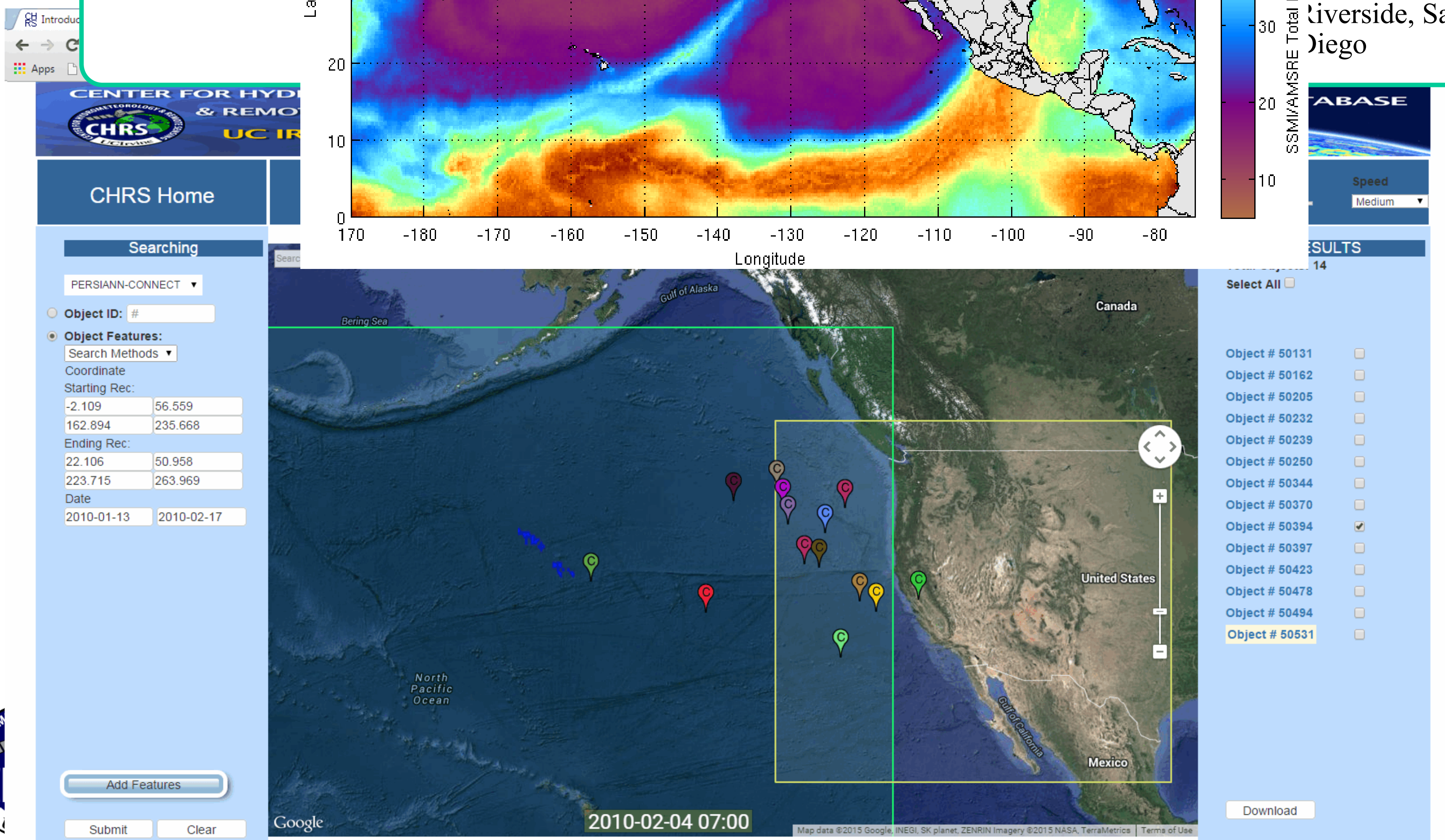


Understand the rainfall processes caused by the AR events

NOAA-ESRL-PSD AR Events

- 2010 Feb. 5-6





...R
...g Record

...010, Feb.
...os Angeles,
...range,
...iverside, San
...iego

Conclusions

- **PERSIANN-CONNECT records the flooding ARs at early stage**
  - **High temporal resolution**
  - **Good tracking capability**

Developing an AR-classification algorithms based on PERSIANN-CONNECT:

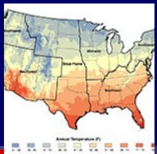
- **Useful attributes:**
  - **Moving directions and moving speeds**
  - **Total rainfall amount**

Thank you

Thank You for Your Presence and Listening

Somewhere in New Mexico, USA - Photo: J. Sorooshian

Back up slides



Precipitation

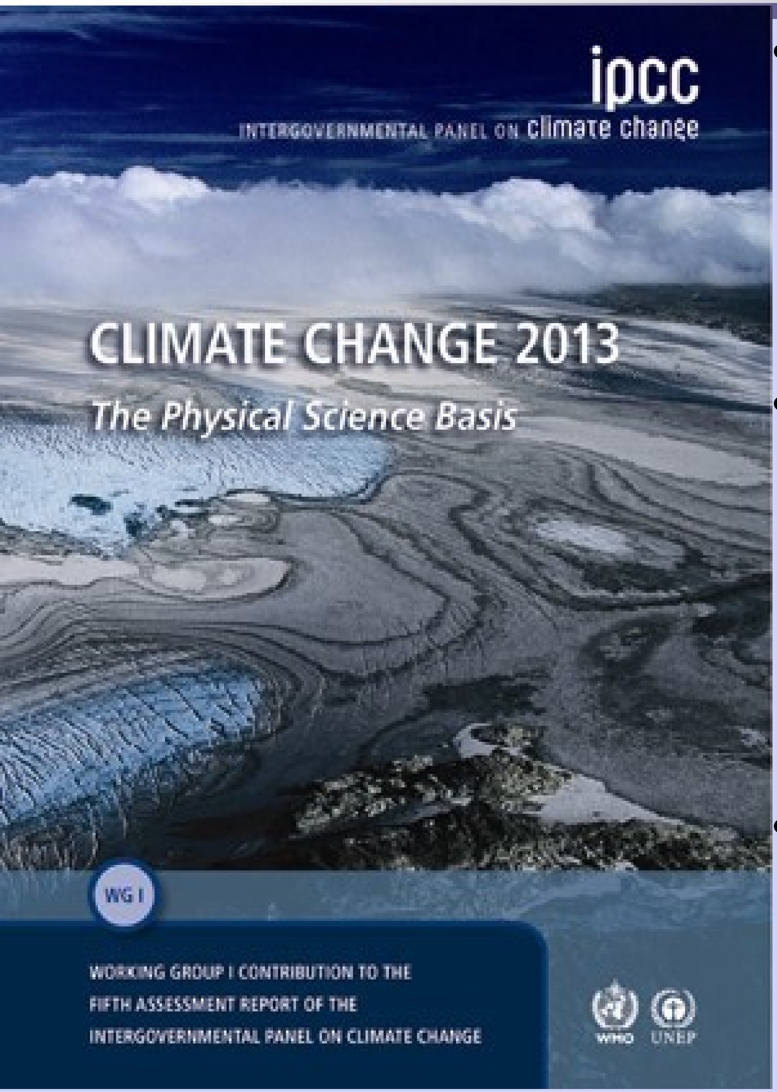
- ***"Precipitation does not exhibit any obvious long-term trends for the Southwest U.S., except for Fall, which shows a slight upward trend. Trends are not statistically significant for any season.***

- ***The region experienced its wettest conditions in the 1980s and 1990s (coinciding with a shift in Pacific climate in 1976, after which El Niño became much more frequent), but has dried in the last decade.***

Extremes

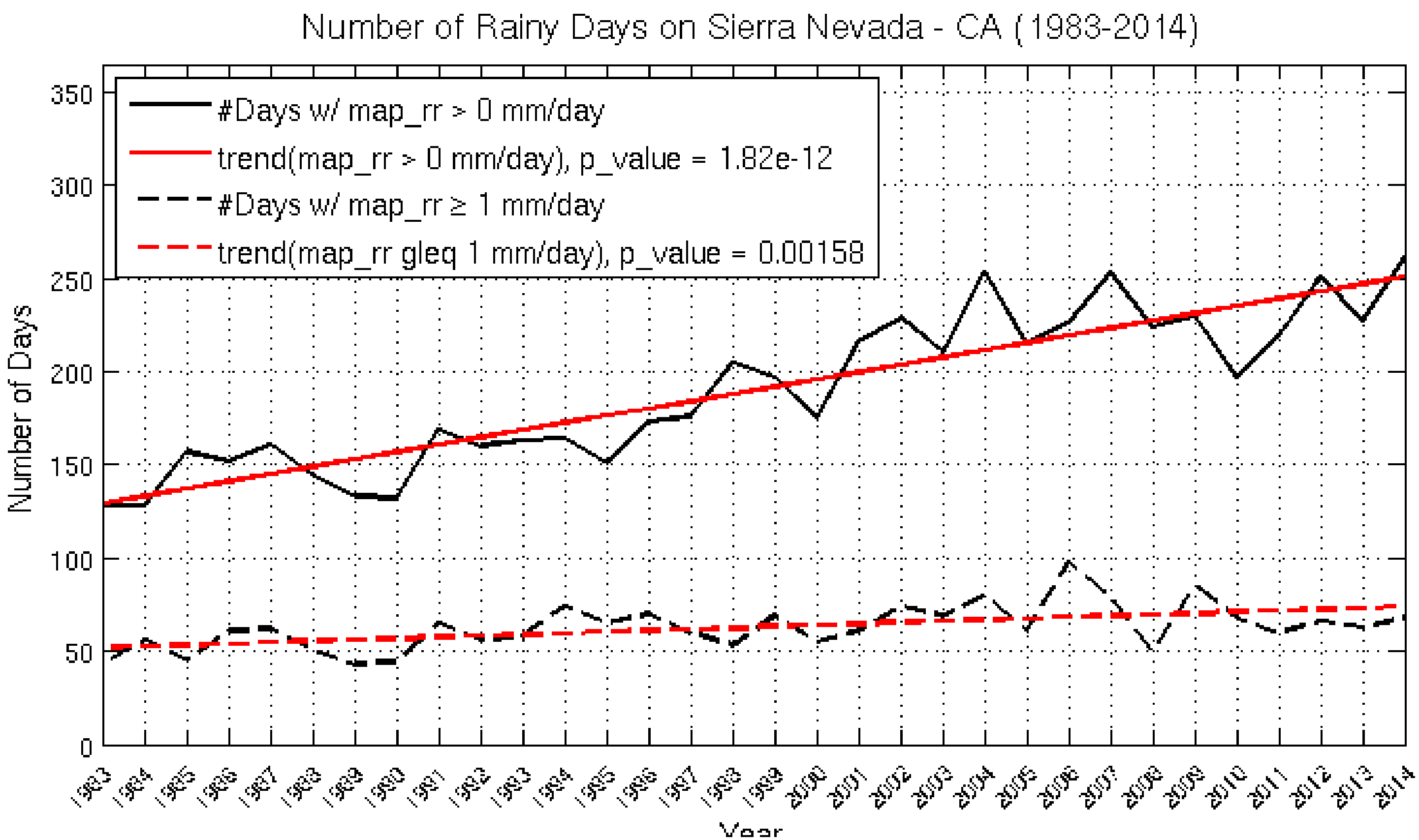
- ***There is not a statistically significant trend in the occurrence of extreme precipitation events in the Southwest."***

Recently Released IPCC Report (AR5) - Sept. 2013



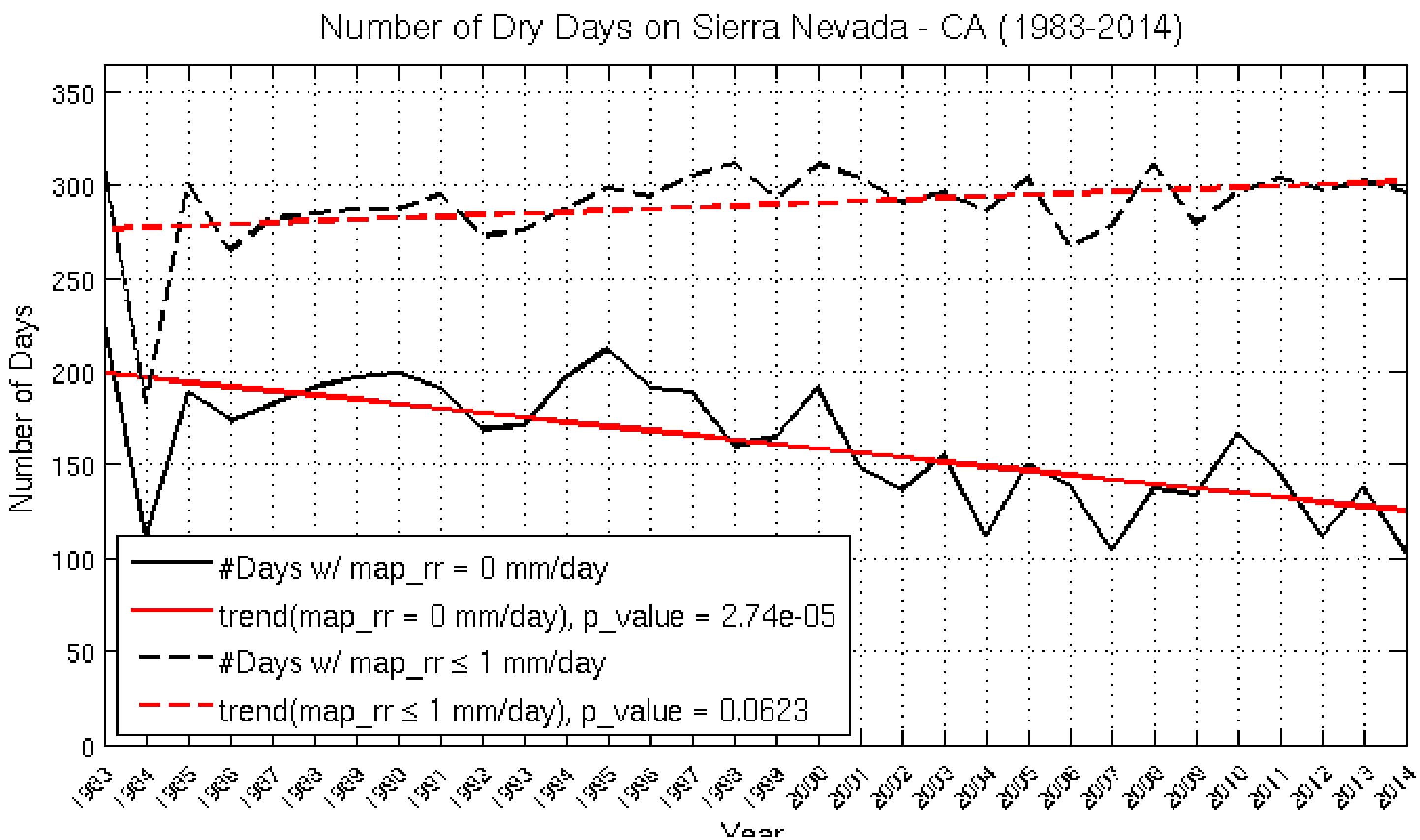
- ***"It is likely that since 1950 the number of heavy precipitation events over land has increased in more regions than it has decreased."***
- ***"There is medium confidence that human influence on climate has affected stream flow and evapotranspiration in limited regions of middle and high latitudes of the Northern Hemisphere."***
- ***"In summary, there continues to be a lack of evidence and thus low confidence regarding the sign of trend in the magnitude and/or frequency of floods on a global scale"***

Trend in the number of Rainy days



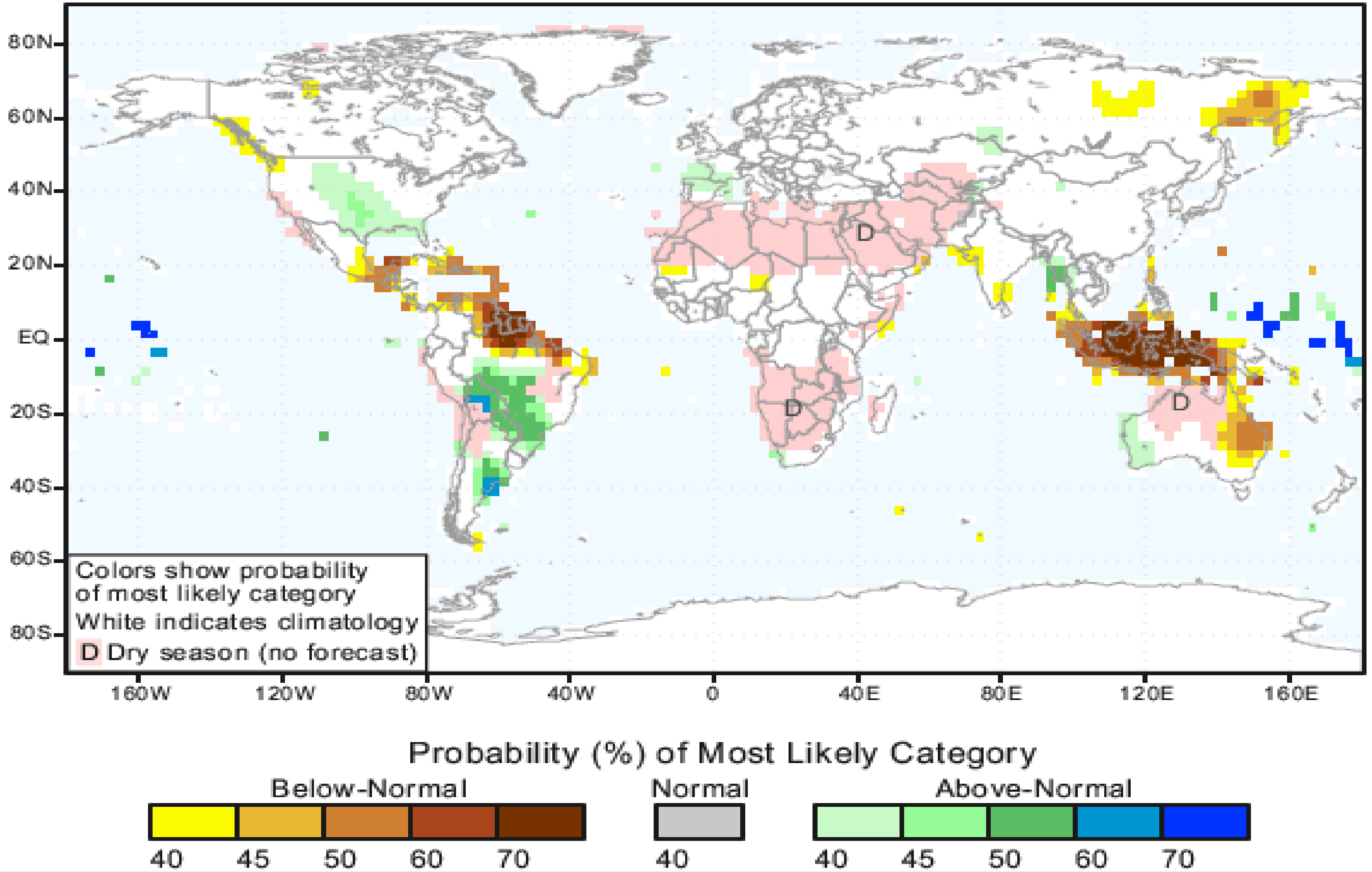
At a 5% significance level (95% CI), the positive trends in the number of rainy days are statistically significant.

Trend in the number of dry days



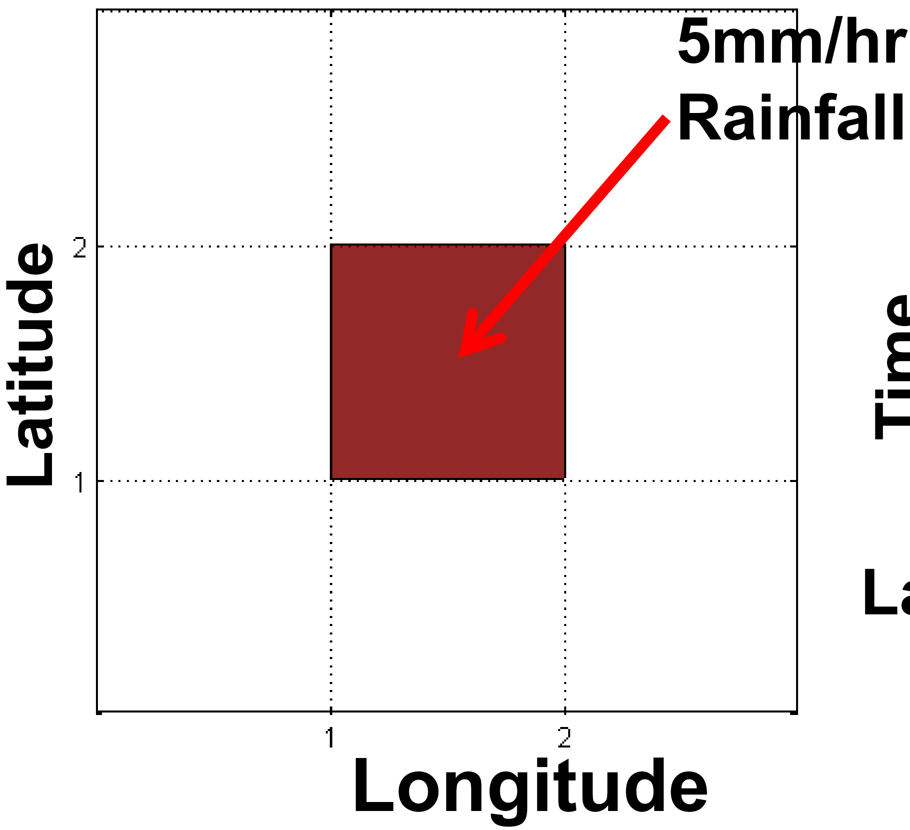
At a 5% significance level (95% CI), the negative trend in the number of dry days (mean areal precip. = 0 mm/day) is statistically significant.

IRI Multi-Model Probability Forecast for Precipitation for June-July-August 2015, Issued May 2015

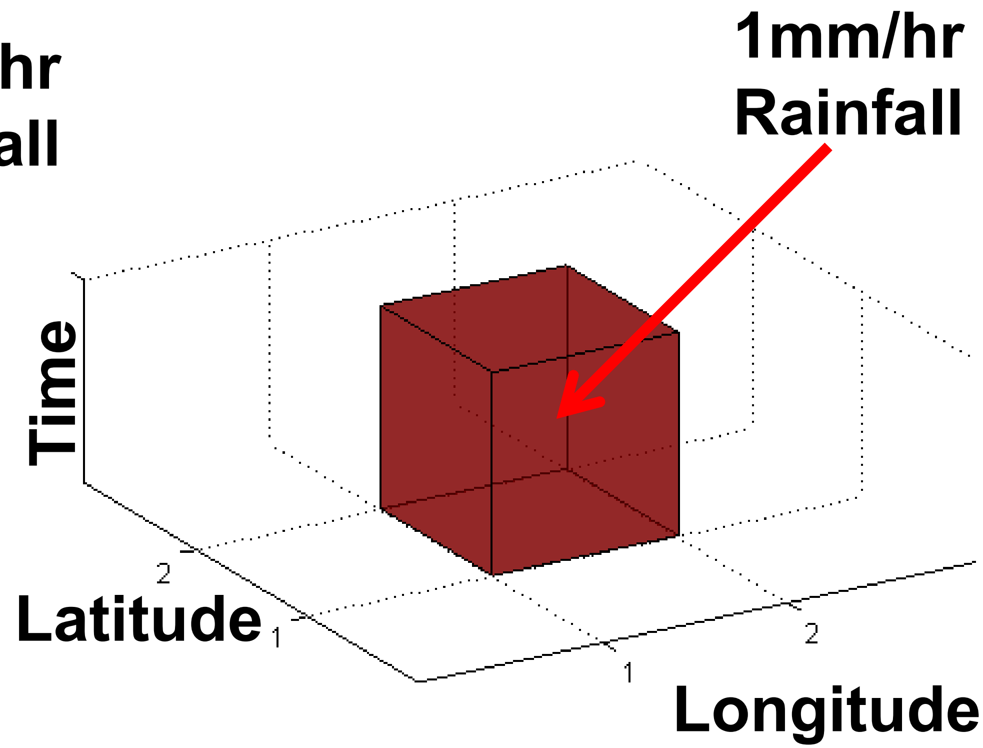


Object Based Approach: 4-D

Traditional

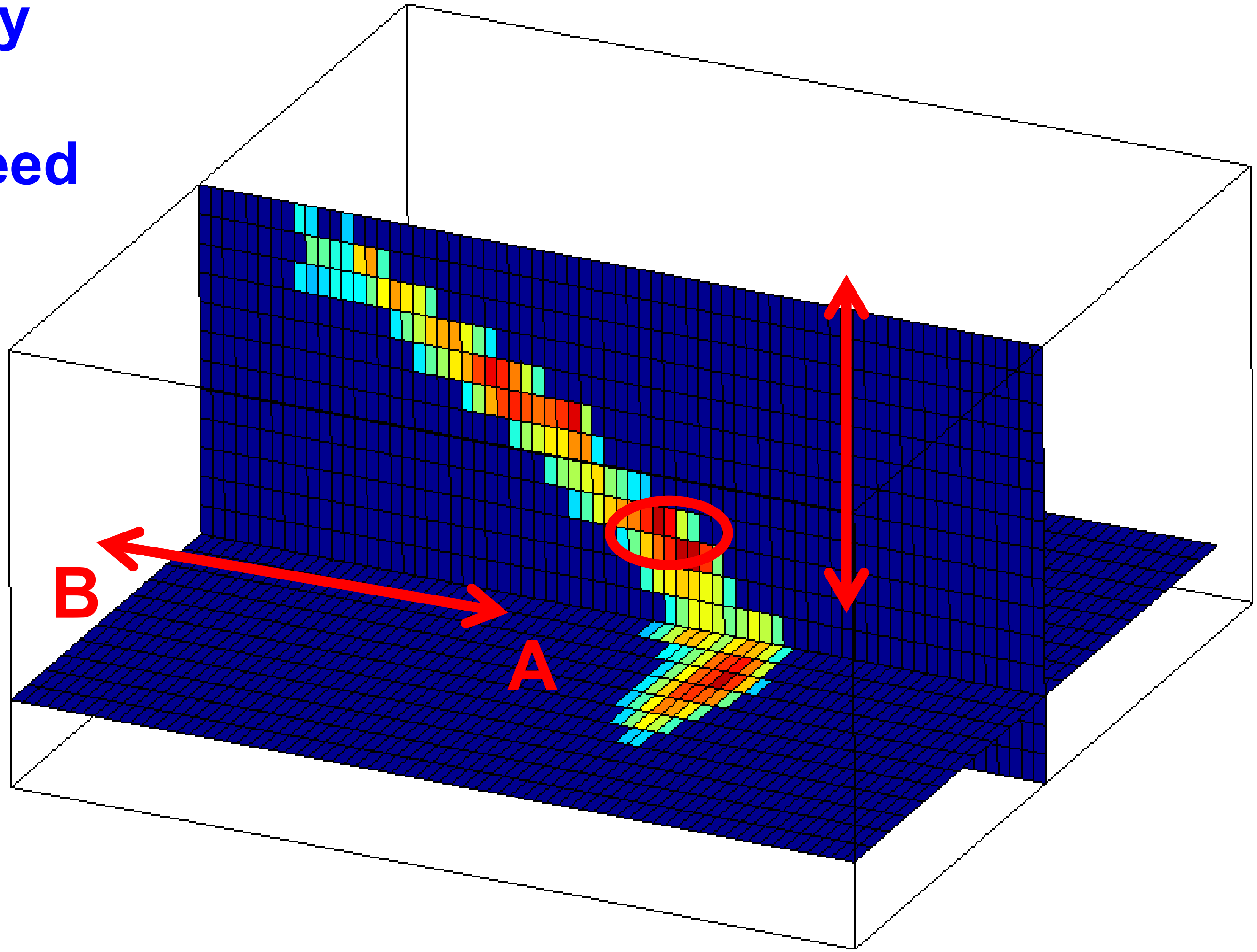


4 Dimensional Approach



4-D Objects Characteristics

1) Max Intensity
2) Duration
3) Average Speed



4-D Objects Characteristics: Nxd

1) Volume
2) Max Intensity
3) Average Intensity
4) Duration
5) Average Speed
6) Centroid Lat
7) Centroid Lon

…

…

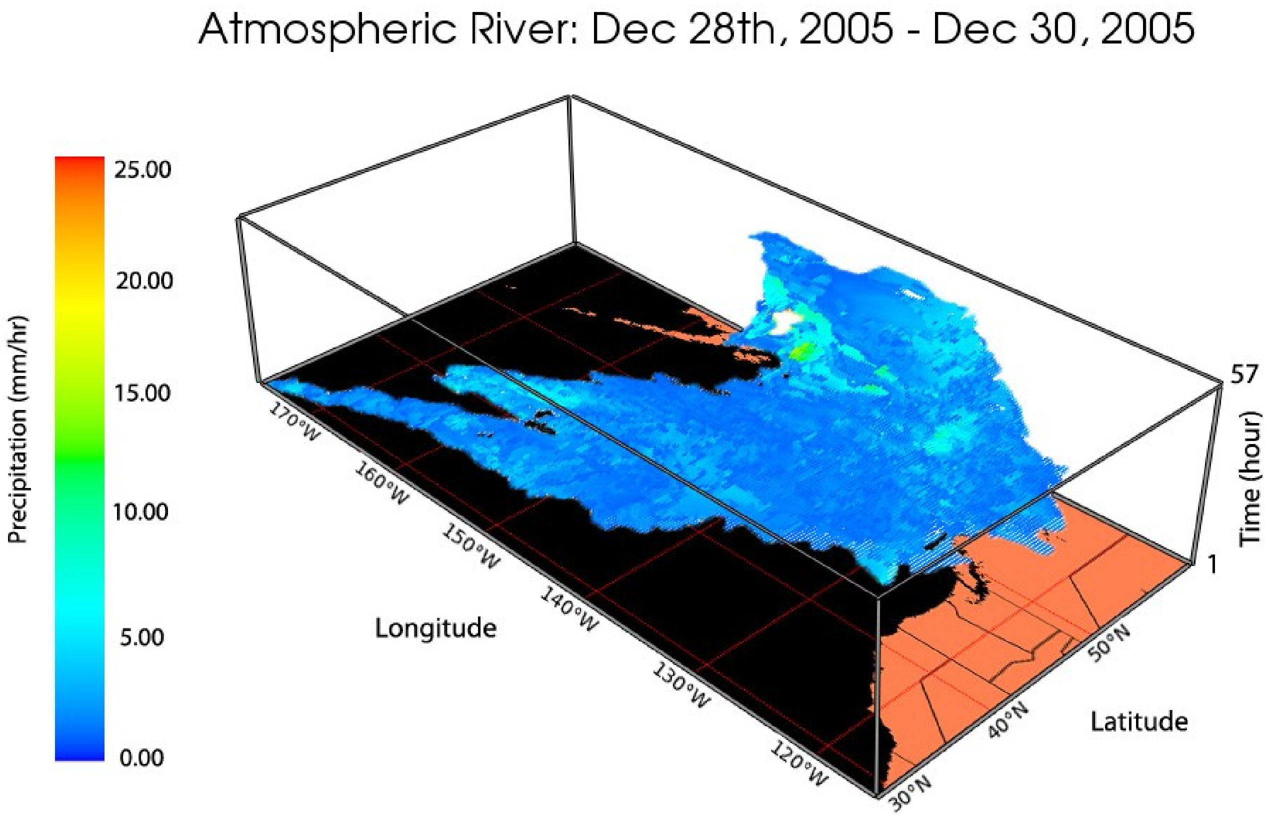
58) NINO3+4,
59) NINO4
60) Start Centroid Latitude

N = # of events
d = # of dimensions (60 in this case)

Nxd Matrix

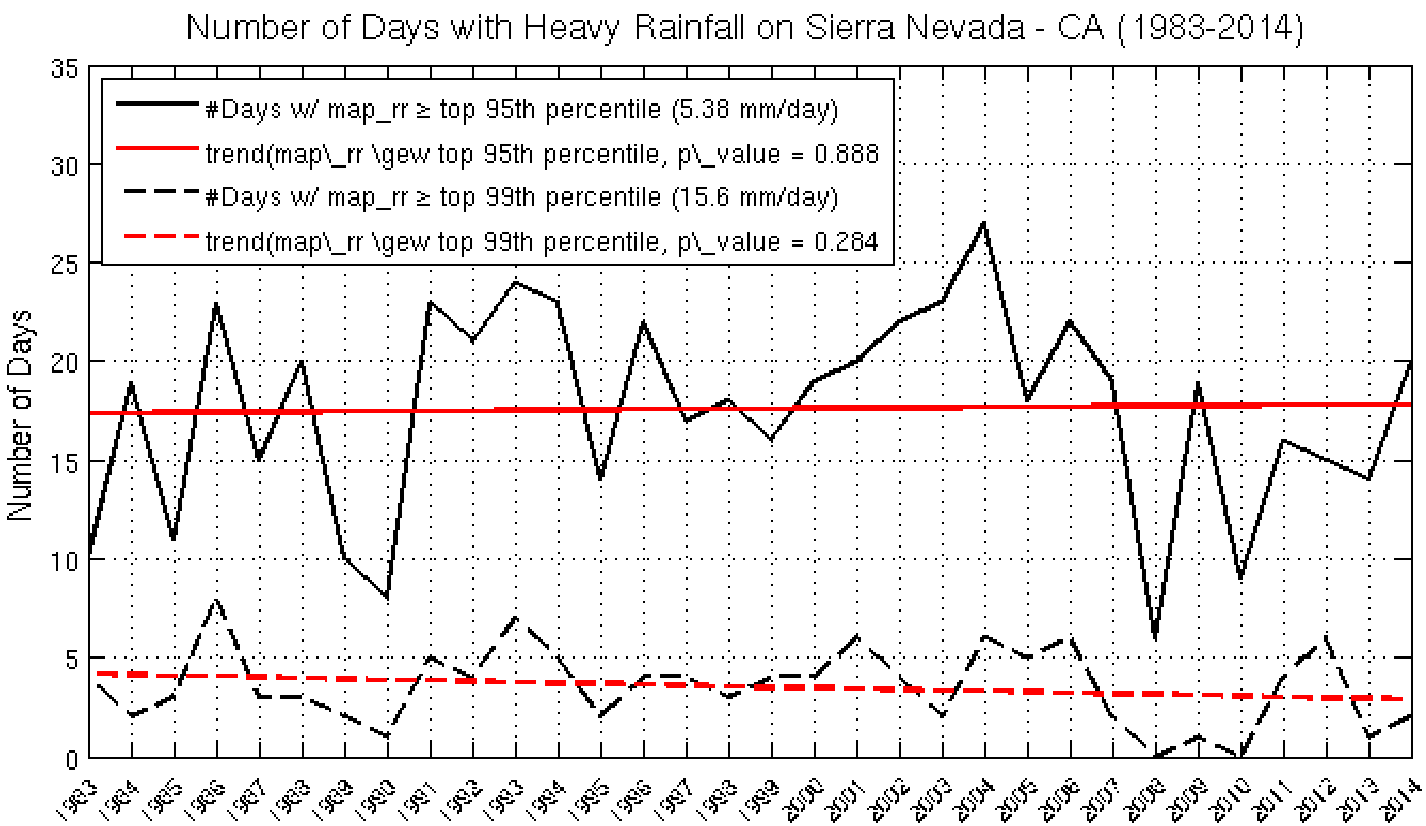
$$\mathbf{X} = \begin{bmatrix} X_1^1 & X_2^1 & \cdots & X_d^1 \\ X_1^2 & X_2^2 & \cdots & X_d^2 \\ \vdots & & & \\ X_1^N & X_2^N & \cdots & X_d^N \end{bmatrix}$$

4-D Object: Atmospheric River Example



Sellars, S., P. Nguyen, W. Chu, X. Gao, K. Hsu, and S. Sorooshian (2013), Computational Earth Science: Big Data Transformed Into Insight, EOS Trans. AGU, 94(32),277

Trend in the number of heavy rainy days



Though a reduction in extreme precipitation events is identified in the recent years (after 2007), at a 5% significance level (95% CI), none of the above trends is statistically significant.