5-6-2016

Toward Mining Massive and Multi-dimensional Data for Extreme Hydrometeorological and Climate Event Analyses

Ling Qiu
University of Nebraska-Lincoln, lingqiu33@gmail.com

Francisco Munoz-Arriola
University of Nebraska-Lincoln, fmunoz@unl.edu

Carlos M. Carrilo
University of Nebraska-Lincoln, carlos.cc@unl.edu

Toward Mining Massive and Multi-dimensional Data for Extreme Hydrometeorological and Climate Event Analyses

Ling Qiu[1,2,3], Carlos M. Carrillo[3], and Francisco Munoz-Arriola[3,4,5]

[1]Electrical and Computer Engineering Department, [2]UCARE Fellow; [3]Biological Systems Engineering; [4]School of Natural Resources, [5]Affiliated to Robert Daugherty Water for Food Institute University of Nebraska-Lincoln

Introduction

Missouri River Basin(MRB)

- MRB has 117 million acres in cropland.
- Produce 46%, 22% and 34% US' wheat, corn, and cattle, respectively.

- Economic contribution of agricultural activity can be jeopardized by extreme hydrometeorological and climate events (EHCEs).

Why do we study precipitation and temperature?

- Precipitation and temperature are two of the most important variables to evaluate extreme events.
- It is unclear if precipitation and Minimum and Maximum Temperatures' patterns present a historical persistence in their occurrence.

Research Question – Objectives - Hypothe

What are the main modes of spatial distribution and temporal variability of extreme events (extreme warm/cold, dry/wet) in MRB?

- Extract spatial patterns and temporal variability EHCEs.
- Develop and implement data mining techniques for EHCEs analyses using Python.

MRB modes of spatiotemporal variability can be characterized as areas of historical water deficits/surpluses and warm/cold conditions at the basin scale.

Empirical Orthogonal Functions (EOFs)

- Climate is characterized by nonlinearity and high dimensionality.
- EOFs analyses are widely used to extract important spatiotemporal patterns of variability in atmospheric sciences (Wilks, 2006).
- Python uses a method based on singular value decomposition (SVD) in EOF analysis.
- The input to EOF analysis is a spatial-temporal anomaly matrix of a variable.
- EOFs is our basic analyses for Data Mining of multidimensional data.

DATASETS

- A spatially comprehensive hydrometeorological dataset for Mexico, the U.S., and southern Canada for the period of 1950-2013 (Livneh et al., 2015).
- Includes observed and gridded daily precipitation, maximum and minimum temperatures.
- The resolution of the dataset is 1/16 degree(~6km)
- ~40K grids in the MRB and 20K time steps by grid.

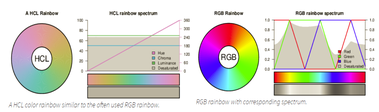
Variable Name	Units	Source
Precipitation	mm	station
Maximum Temperature	℃	station
Minimum Temperature	℃	station

PROGRAMING TOOLS

Matplotlib

- A python 2D plotting library.
- Make plotting easier in python.
- Capable of plotting various kinds of common plots.
- Users are able to control most of the details in a figure.

HCL Color



- HCL(Hue-Chroma-Luminance) is based to how the human perception works.
- Users can directly control the three indexes in HCL.

Conclusions

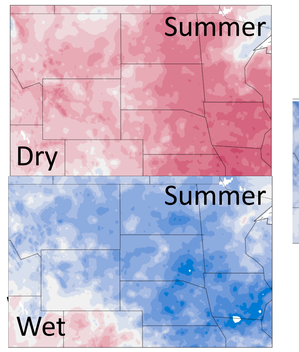
- We found that precipitation and temperature are the ideal meteorological variable to test spatiotemporal variability of extreme events using simple (composite analysis) and sophisticated (EOF) statistical techniques.
- The EOF data mining technique is able to portray the dry and wet events of the MRB climatic variability of the last 50 years. Wet and dry events in the interannual scale of variability are captured as the statistically dominant mode of EOF for summer and winter.
- Python codes for plotting and analysis used in this study were found to be very efficient to work on massive multidimensional climate datasets.

References

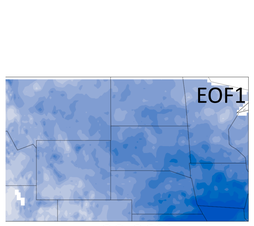
- Wilks, D. S. (2006). Statistical Methods in the Atomspheric Sciences (2nd ed., Vol. 91, International Geophysics Series). Burlington, MA: Elsevier.
- Livneh, B., Bohn, T. J., Pierce, D. W., Munoz-Arriola, F., Nijssen, B., Vose, R., . . . Brekke, L. (2015). A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013. *Scientific Data*. Retrieved March 04, 2016.
- Why HCL. (n.d.). Retrieved April 06, 2016, from http://hclwizard.org/why-hcl/

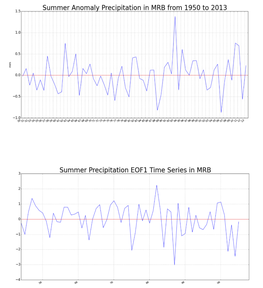
Precipitation Analysis

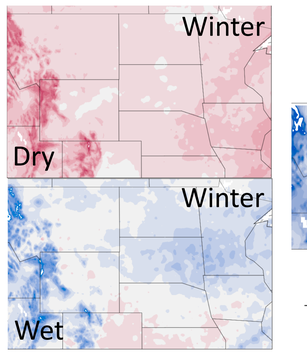
Anomalies

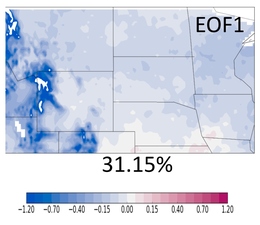


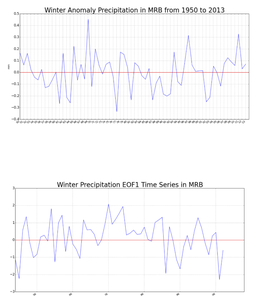
EOFs





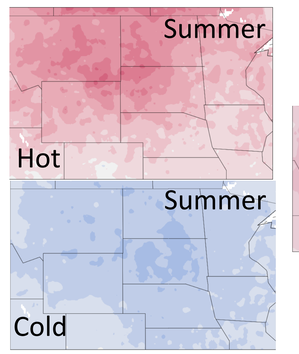






Temperature Analysis

Anomalies



EOFs