

## **Fisseha Berhane, PhD**

Data Scientist at [Aurotech](#)

Phone: 443-970-2353

Website: <http://datascience-enthusiast.com/>

Email: [fisseha@jhu.edu](mailto:fisseha@jhu.edu)

Regular writer for [DataScience+](#) and [R-bloggers](#)

### **Current Employment**

Data Scientist at [Aurotech](#)

Sep 2015-

Solving various problems using data analytics and machine learning with Spark, R, Python, Hadoop ecosystem, and Tableau.

Projects:

- Hadoop Data Lake for analytics and machine learning with big data
  - Created a Hadoop cluster on AWS EC2
  - Ingested disparate data from various sources and in different formats to the lake
  - Cleaned and transformed the data for downstream analytics pipeline
  - Developed machine learning applications using Spark's MLlib library
  - Connected Tableau with the data lake and created visualizations using Spark SQL with ODBC connector.
- Predicting drug recall potential using various machine learning techniques and various data sources  
Architecture diagram available [here](#)
- Interactive drug adverse event knowledge discovery with R and Shiny using unsupervised machine learning techniques  
Cleaned and merged lots of adverse event datasets and stored them in a database.  
Developed an R-shiny application that clusters (using optics and hierarchical clustering) drug adverse events to discover new insights interactively  
Architecture diagram available [here](#)
- Real-time tracking of disease outbreaks using social media with R and Tableau  
Created a complete pipeline that automates social media data collection, cleaning and processing, sentiment analysis, trend analysis and creates a Tableau dashboard.  
Architecture diagram available [here](#)
- R-Shiny dashboard API that helps to download the FDA adverse events data  
Created an API that helps users to download data based on search query from the FDA adverse events database
- Social media mining to track natural hazards at real-time  
Created a Tableau dashboard that helps to track flooding
- Google Trends Analytics with R-Shiny  
Created an R shiny application that closely listens to google search trends and identifies anomalies in disease related google searches.

## Education

Johns Hopkins University, Baltimore, MD

Ph.D., Earth and Planetary Sciences, 2016

Thesis: Intraseasonal precipitation variability over tropical Africa

Advisor: Benjamin F. Zaitchik

M.A., Earth and Planetary Sciences May 2013

University of Connecticut

M.S., Natural Resources and the Environment, May 2011

Thesis: Model based assessment of potential impacts of climate change on the flow of the main headwaters of the Nile River: Equatorial Lakes Region and Blue Nile Basins

Advisor: Richard Anyah

Mekelle University, Ethiopia

B.Sc., Civil Engineering, June 2006

## Research Positions

Graduate Research Assistant, Department of Earth and Planetary Science, Johns Hopkins

University, Baltimore, Maryland.

August 2011 – 2015

- Built semi-automated rainfall prediction models, with various machine learning techniques such as Tree-based ensemble models (**Random Forest** and **Boosting**), **Support vector Machines** and **Artificial Neural Network**, with **R**, HTML, JavaScript, and CSS.
- Employed various statistical analysis and data mining techniques using **Python** and **R** to understand interactions of atmospheric waves and their impacts on rainfall using large volume climate data.
- Analyzed large volume climate data, using **Python** and **R**, to investigate future climate conditions
- Completed many side-projects on big data using **Spark** (e.g., movie recommendation, web server log analysis, text mining and entity resolution and click-through prediction; available on my [website](#))
- Worked on many other side-projects using **R** (available on my [website](#))
- In addition to the data science courses I have done in grad school, I have taken more than 20 edx, coursera and Udacity data science courses with **R**, **Spark**, **Python**, **Matlab**, and **Hadoop** and **MapReduce** (certificates on my [website](#))

Graduate Research Assistant, Department of Natural Resources and the Environment,

University of Connecticut, Storrs, CT

2009 – May 2011

- Built and evaluated a model that predicts Nile River flow. Further, examined possible impacts of climate change on river flow using different climate scenarios.
- The main tools I used in this study: **R**, **Python** and GIS.

## **Awards**

Research Assistantship, Department of Earth and Planetary Sciences, Johns Hopkins University,  
Baltimore, Maryland 2012-2015

Morton K. Blaustein Fellowship, Department of Earth and Planetary Sciences, Johns Hopkins  
University, Baltimore, Maryland 2011-2012

Research Assistantship, Department of Natural Resources and the Environment, University of  
Connecticut, Storrs, CT 2009-2011

## **Teaching Experience**

Teaching assistant (TA), Department of Earth and Planetary Science, The Johns Hopkins  
University, Baltimore, Maryland. Spring 2013

Assistant Lecturer, Department of Civil Engineering, Mekelle University, Ethiopia 2006-2009

## **Peer-Reviewed publications**

**Berhane F** and BF Zaitchik: An MJO-mediated mechanism to explain ENSO and IOD impacts on  
East African short rains. in prep.

**Berhane F**, BF Zaitchik and HS Badr, 2015: The Madden-Julian Oscillation's influence on Spring  
Precipitation over Equatorial West Africa. J. Climate. doi: <http://dx.doi.org/10.1175/JCLI-D-14-00510.1>.

**Berhane F** and BF Zaitchik, 2014: Modulation of Daily Precipitation over East Africa by the  
Madden-Julian Oscillation. J. Climate, 27(15): 6016-6034. doi:  
<http://dx.doi.org/10.1175/JCLI-D-13-00693.1>.

**Berhane F**, BF Zaitchik and A Dezfuli, 2013: Sub-seasonal analysis of precipitation variability in  
the Blue Nile River basin. J. Climate, 27(1): 325-344. doi: <http://dx.doi.org/10.1175/JCLI-D-13-00094.1>.

## **Data Science related courses I have done**

### **Graduate**

- Time Series Analysis
- Statistical Computing
- Data Analytics for Engineering, Policy Analysis and Management
- Inversion Modeling & Data Assimilation
- Spatial Statistics and Modelling
- Environmental Quantitative Methods
- Python Scripting for GIS

### **Undergraduate**

- Probability and Statistics
- Computer Programming (C++)
- Applied Mathematics I
- Applied Mathematics II
- Numerical Methods

### **Online (Coursera, edX, Udacity; certificates linked)**

- [Machine Learning](#)
- [BerkeleyX: CS100.1x Intro to Big Data with Apache Spark](#)
- [MITx - 6.00.1x Intro to Computer Science and Programming Using Python](#)
- [Practical Machine Learning](#)
- [BerkeleyX: CS190.1x Scalable Machine Learning](#)
- [Developing Data products](#)
- [Intro to Data Science](#)
- [DAT201x: Querying with Transact-SQL](#)
- [R Programming](#)
- [Reproducible Research](#)
- [The Data Scientist's Toolbox](#)
- [Getting and Cleaning Data](#)
- [Regression Models](#)
- [MITx: 15.071x The Analytics Edge](#)
- [W3C-HTML5](#)
- [Statistical Inference](#)
- [Exploratory Data Analysis](#)
- [Intro to Hadoop and MapReduce](#)
- [Text Mining, Scarping and Sentiment Analysis with R](#)
- [Tableau 9 For Data Science: Real-Life Data Science Exercise](#)
- [Tableau 9 Advanced Training: Master Tableau for Data Science](#)

### **Other Skills**

- ❖ **Operating Systems:** Windows , Unix and Linux
- ❖ **Software:** Python, R, Apache Spark, Tableau, Hadoop, SQL, Matlab, C++, Octave, GRADS, Ferret, NCL, WRF, ArcGIS, SWAT, ERDAS IMAGINE, ENVI, RegCM, Fortran, HTML5, JavaScript, CSS, Git

### **Selected Presentations**

Berhane F and BF Zaitchik, 2015: The influence of the MJO on Spring Equatorial West African convection. 95<sup>th</sup> AMS Annual Meeting 2015, Sixth Conference on Weather, Climate, and the New Energy Economy, Phoenix, AZ.

- Berhane F and BF Zaitchik, 2014: Intraseasonal variability of the impacts of the Madden-Julian Oscillation on East African long and short rains. 94<sup>th</sup> AMS Annual Meeting 2014, Second Symposium on Prediction of the Madden-Julian Oscillation: Impacts on Weather and Climate Extremes, Atlanta, GA.
- Berhane F and BF Zaitchik, 2014: Intraseasonal variability of the impacts of the Madden-Julian Oscillation in the Gulf of Guinea. 94<sup>th</sup> AMS Annual Meeting 2014, Fifth Conference on Weather, Climate, and the New Energy Economy, Atlanta, GA.
- Berhane F, BF Zaitchik and A Dezfuli, 2013: Evolution of intraseasonal precipitation variability in the Blue Nile River basin. 93<sup>rd</sup> AMS Annual Meeting 2013, 25th Conference on Climate Variability and Change, Austin, Texas, USA.
- Berhane F, 2013: Modulation of daily rainfall over Africa by the Madden-Julian oscillation. 5<sup>th</sup> annual Atmosphere-Ocean Science Days seminar, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland
- Berhane F, 2013: Intraseasonal variability of the modulation of daily rainfall over Africa by the Madden-Julian oscillation. Atmosphere-Ocean Seminar. Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland
- Berhane F, 2013: Modulation of daily rainfall over Africa by the Madden-Julian oscillation. Journal Club, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland
- Berhane F, 2012: Intraseasonal variability of precipitation in the Blue Nile River Basin. Climate Dynamics of Tropical Africa: Present Understanding and Future Directions, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA.
- Berhane F, 2012: Rainfall anomalies in the Blue Nile basin and their teleconnections with the Indian Summer Monsoon. Journal Club, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland
- Berhane F, 2012: Evolution of drivers and mechanisms of precipitation variability in the Blue Nile River Basin. Eastern Nile Technical Regional Office- Nile Basin Initiative. Addis Ababa, August 2012.
- Berhane F, 2012: Model based assessment of potential impacts of climate change on the flow of the Blue Nile Basin. Eastern Nile Technical Regional Office- Nile Basin Initiative. Addis Ababa, August 2012.
- Berhane F, Anyah R.O., 2010: Hydrological Response to Climate Change over the Blue Nile Basin Distributed hydrological modeling based on surrogate climate change scenarios. American Geophysical Union Fall Meeting 20140, San Francisco, California, USA.

### **Professional Memberships**

- Member of American Meteorological Society  
Member of American Geophysical Union

