

# Samuel Araya, Ph.D.

Data Science / Environmental Soil Physics / Python, R, SQL

[SamuelNA.netlify.app](http://SamuelNA.netlify.app)

## EXPERIENCE

---

### Data Scientist

**Corteva Agriscience, Indianapolis, IN (Remote)** | MARCH 2023 – PRESENT

Contracted by Insight Global to perform crop yield and environment related data aggregation, data analysis and modeling across multiple environments and scales .

- \* Performed agricultural data analysis and model building across multiple environments and scales.
- \* Conducted data aggregation from on farm records, in-house crop models, soil and weather APIs into cohesive databases for product selection and model training.
- \* Worked with in-house data science tools built in Python environments.

### Soil (Data) Scientist

**US Department of Agriculture, NRCS, Texas** | OCTOBER 2021 – MARCH 2023

Performed data-driven analysis and interpretations. Developed and maintained soil-related geospatial data. Utilized soil-landscape relationships to map and describe soils.

- \* Performed data wrangling and quality analysis of large state-wide soil and environmental geospatial databases using R, Python, and SQL.
- \* Performed data interpretation and created insights using data visualization and statistical methods including regression, PCA, and cluster analysis.
- \* Performed data fusion from multi-source spatial and non-spatial data to create new databases.
- \* Published analysis results in markdown reports and company ArcGIS Online geodatabases.
- \* Performed large-scale terrain analysis in Google Earth Engine and ArcGIS environments.
- \* Retrieved satellite imagery using APIs to investigate fire and natural disaster.
- \* Conducted field soil and ecological data collections.

### Postdoctoral Researcher

**Stanford University, Palo Alto** | SEPTEMBER 2019 – SEPTEMBER 2021

Researched remote sensing and GIS methods to detect soil heavy metal contamination in rice paddies and natural landscapes.

- \* Performed multi-year time series analysis of rice paddies using Sentinel-2 and PlanetScope imagery.
- \* Performed time series analysis of soil temperature and moisture data.
- \* Developed and implemented predictive models using machine learning algorithms such as random forest, feedforward neural network, support vector machines, k-means clustering, and PCA.
- \* Performed fusion of data with temporal and spatial dimensions.
- \* Utilized Google Earth Engine environment to retrieve and process large-scale satellite and gridded environmental data.
- \* Planned and conducted a year-long field soil data collection and monitoring project.

## SKILLS

---

### Data analysis & programming

- \* Python (pandas, GeoPandas, scikit-learn, TensorFlow, etc.)
- \* R ( tidyverse, caret, XGBoost, sf, terra, etc.)
- \* JavaScript
- \* SQL
- \* Git

### GIS & remote sensing

- \* Google Earth Engine
- \* ArcGIS
- \* QGIS
- \* ENVI

## CONTACT

---

### LinkedIn

[linkedin.com/in/saraya209/](https://www.linkedin.com/in/saraya209/)

### GitHub

[github.com/saraya209](https://github.com/saraya209)

## EXPERIENCE

---

### Teaching Assistant

University of California, Merced | JANUARY 2011 – AUGUST 2019

- \* Taught graduate- and undergraduate-level laboratory and discussion sessions for courses: Stable Isotope Ecology (BIO 174/ESS 174); Fundamentals of Soil Science (ES 170); Contemporary Biology (BIOL 01); Earth Resources and Society (ESS 010); Sustainability Science (ESS 002).

Prepared teaching plans and study materials. Graded and assessed student learning.

### Academic Programming Director

Hamelmallo Agricultural College, Eritrea | MARCH 2007 – AUGUST 2010

- \* Developed and managed schoolwide courses database.
- \* Prepared class and exam timetables using MS Access and FET Timetabling software.
- \* Prepared semester reports and recommendations.

### Teaching Assistant

Hamelmallo Agricultural College, Eritrea | SEPTEMBER 2006 – MAY 2010

- \* Taught laboratory and discussion sessions for multiple courses: Fundamentals of Remote Sensing and GIS (LREN 241); Fundamentals of Soil Science (LREN 212D); Land Use Planning (LREN 202).
- \* Prepared teaching plans and study materials. Graded and assessed student learning.

## EDUCATION

---

### Ph.D. in Environmental Systems (Soil Physics)

University of California, Merced | AUGUST 2014 – AUGUST 2019

Dissertation: Soil Structure and Land Surface Controls on Soil Hydraulic Properties and Processes  
*Teamrat A. Ghezzehei (Chair), YangQuan Chen, Martha Conklin, Jason Sexton, and Joshua Viers*

### M.S. in Environmental Systems (Soil Biogeochemistry)

University of California, Merced | AUGUST 2010 – AUGUST 2014

Thesis: Effect of Combustion Temperature on Soil and Soil Organic Matter Properties  
*Asmeret A. Berhe (Chair), Randy Dahlgren, and Samuel Traina*

### B.S. in Land Resources and Environment

University of Asmara, Eritrea | SEPTEMBER 2002 – MAY 2006

Thesis: Application of GIS to Derive Surface Parameters for Rainfall-Runoff Modeling at Afdeyu Watershed

## PEER-REVIEWED PUBLICATIONS

---

- \* S. N. Araya, Mitchell, J. P., Hopmans, J. W., Ghezzehei, T. A. (2022). Long-term impact of cover crop and reduced disturbance tillage on soil pore size distribution and soil water storage. *SOIL*.
- \* Mitchell, J. P., Shrestha, A., Epstein, L., Dahlberg, J. A., Ghezzehei, T. A., Araya, S. N., Richter, B., Kaur S., Henry, P., Munk, D. S., Light, S., Bottens, M., Zaccaria, D. (2022). No-tillage sorghum and garbanzo yields match or exceed standard tillage yields. *California Agriculture*.
- \* S. N. Araya, Fryjoff-Hung, A., Anderson, A., Viers, J. H., Ghezzehei, T. A. (2021). Advances in soil moisture retrieval from multispectral remote sensing using unoccupied aircraft systems and machine learning techniques. *Hydrology and Earth System Sciences*.

- \* S. N. Araya, Fryjoff-Hung, A., Anderson, A., Viers, J. H., Teamrat A Ghezzehei (2020). Machine learning based soil moisture retrieval from unmanned aircraft system Multispectral remote sensing. *IEEE IGARSS 2020*.
- \* S. N. Araya, Ghezzehei, T. A. (2019). Using machine learning for prediction of saturated hydraulic conductivity and its sensitivity to soil structural perturbations. *Water Resources Research*.
- \* S. N. Araya, Fogel, M. L., Berhe, A. A. (2017). Thermal alteration of soil organic matter properties: a systematic study to infer response of Sierra Nevada climosequence soils to forest fires. *SOIL*.
- \* S. N. Araya, Meding, M., Berhe, A. A. (2016). Thermal alteration of soil physico-chemical properties: a systematic study to infer response of Sierra Nevada climosequence soils to forest fires. *SOIL*.

## PROFESSIONAL SERVICE

---

- \* **Reviewer board member:** “Soil Systems” journal (ISSN 2571-8789).
- \* **Session Co-Chair:** “Soils and Hydrology Session”, IEEE IGARSS 2020.
- \* **Ad hoc reviewer:** Journal papers (approximately 12 per year).