

David **Caleb** Robinson

Phone: (662) 380-6690

Email: [dcrobins@gatech.edu](mailto:dcrobins@gatech.edu)

Website: <http://www.calebrob.com/>

US Citizen

---

## INTEREST STATEMENT

---

I am interested in using machine learning, optimization, and simulation techniques to create models that give better insights into the interactions of complex systems. An example of this is modeling the migration patterns of populations under different scenarios of sea level rise in order to understand and better plan for accelerated urban growth.

---

## EDUCATION

---

- Georgia Institute of Technology, 2015-  
Ph.D. Computational Science and Engineering
- University of Mississippi, 2011-2015  
Major B.S. Computer and Information Science  
Minor in Mathematics for Engineers  
GPA - 3.92

---

## SKILLS

---

- Familiar with Python; including Python data science, visualization, and geospatial libraries
- Occasionally use Java, C#, HTML/CSS/Javascript, and PHP
- Familiar working with geospatial datasets using QGIS
- Familiar with formulating/solving mathematical programs using CLPEX or Gurobi
- Familiar with Linux and Windows, use Ubuntu Linux as main OS
- Graduate classes: Algorithms, Machine Learning, Computational Sustainability, Modeling and Simulation, Numerical Linear Algebra, Deep Learning, Network Science, Development Economics, Econometric Analysis

---

## FELLOWSHIPS AND AWARDS

---

- UN Data for Climate Action Award (Fall 2017)
- Serve Learn Sustain Fellowship (Fall 2017)
- Microsoft Azure Research Award (Fall 2017)
- NSF Graduate Research Fellowship Program (GRFP) Honorable mention (Spring 2017)

---

## PUBLICATIONS

---

### Journal Articles

W. Zhang, C. Robinson, S. Guhathakurta, V. M. Garikapati, B. Dilkina, M. A. Brown, and R. M. Pendyala, "Estimating residential energy consumption in metropolitan areas: A microsimulation approach," *Energy*, 2018. (In review)

C. Robinson, B. Dilkina, J. Hubbs, W. Zhang, S. Guhathakurta, M. A. Brown, and R. M. Pendyala, "Machine learning approaches for estimating commercial building energy consumption," *Applied Energy*, 2017

### Conference Proceedings

C. Robinson and R. Fujimoto, "Toward a common object model for integrated transportation and land use models," 2018. (In review)

C. Robinson, F. Hohman, and B. Dilkina, "A deep learning approach for population estimation from satellite imagery," in *Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities, GeoHumanities'17*, ACM, 2017

C. Robinson, A. Shirazi, M. Liu, and B. Dilkina, “Network optimization of food flows in the U.S.,” in *Big Data (Big Data)*, 2016 *IEEE International Conference on*, IEEE, 2016

A. Jain, C. Robinson, B. Dilkina, and R. Fujimoto, “An approach to integrate inter-dependent simulations using HLA with applications to sustainable urban development,” in *2016 Winter Simulation Conference (WSC)*, IEEE, 2016

C. Robinson and J. Xue, “Sparse local binary pattern histograms for face recognition with limited training samples,” in *Proceedings of the 2014 ACM Southeast Regional Conference*, ACM, 2014

## Other

A. Gupta, C. Robinson, and B. Dilkina, “Predicting and alleviating road flooding for climate mitigation,” *UN Data for Climate Action Challenge*, 2017

C. Robinson and B. Dilkina, “A machine learning approach to modeling human migration,” *arXiv preprint arXiv:1711.05462*, 2017

C. Robinson, “Modelling global climate variables with cellular automata networks.” May 2015 (Honors Thesis)

## RESEARCH PROJECTS AND COURSEWORK

---

### **Human migration patterns under different scenarios of sea level rise**

Ongoing

with Bistra Dilkina and Juan Moreno-Cruz

- We couple sea level rise and human migration models to create a framework for studying the effects of sea level rise on human population distributions.
- We implement the framework using the radiation model of human migration and NOAA Digital Coast sea level rise estimates.
- Our results show how the indirect effects of sea level rise (people living in areas that will experience large migrant influxes) can be much larger than the direct effects.

### **A machine learning approach to modeling human migration**

Ongoing

with Bistra Dilkina

- We show how machine learning models of human migration can outperform traditional physics based models of human migration by using historical training data.
- Using socioeconomic data further improves the models.
- We test our models on migrations between counties in the US and migrations between countries globally.

### **A deep learning approach for population estimation from satellite imagery**

Ongoing

with Fred Hohman and Bistra Dilkina

- We train convolutional neural networks to estimate population counts from satellite imagery. More specifically, we train CNNs that take  $1\text{km}^2$  patches of Landsat 7 satellite images as input, then directly regress how many people live in the area covered by the input image.
- We train our models using disaggregated Census tract data for 2000, then validate the models both by quantitatively by making predictions for
- We validate our models quantitatively by aggregating their predictions

### **Predicting and alleviating road flooding for climate mitigation**

2017

with Amrita Gupta and Bistra Dilkina

- We created a computational framework to determine which roads in Senegal should be fortified against flooding in order to maximize accessibility under a fixed budget.
- We combined road network data, data about different flooding scenarios, and human mobility data to estimate how accessibility over the road network in Senegal would be affected under different flooding scenarios.
- We found that optimizing over different accessibility measures give nonuniform improvements to mobility in different parts of the country, and demonstrate the tradeoffs between the available budget for road network repairs and the overall benefit of the repairs.

## Machine learning approaches for estimating commercial building energy consumption 2017

with Bistra Dilkina, Jeffrey Hubbs, Wenwen Zhang, Subhrajit Guhathakurta, Marilyn A. Brown, and Ram M. Pendyala

- We build machine learning models for estimating commercial building energy consumption, based on data from the Commercial Building Energy Consumption Survey.
- We show that machine learning models which only rely on square footage, number of floors, principle building activity, number of heating degree days, and number of cooling degree days, can achieve good results for predicting the commercial building energy consumption.
- We validate our models on energy consumption data from New York City.

## Network optimization of food flows in the U.S. 2016

with Arezoo Shirazi, Mengmeng Liu, and Bistra Dilkina

- We used data from the Commodity Flow Survey and linear optimization models to determine how flows of food in the United States could be restructured more efficiently.
- We minimize the number of food ton miles in the network of flows of tons of food between states. In this problem we constrain the amount of outgoing and incoming commodities to be the same for each state, but to more efficient locations.
- We formulate the problem as a multi-objective optimization problem to find solutions that will be efficient but resilient.

## Optimization with integrated transportation and land use models 2016

with Ajitesh Jain, Bistra Dilkina, and Richard Fujimoto

- The goal of this project is to see if it is possible to influence where people live in an urban environment by changing the transportation networks with the purpose of achieving sustainability goals.
- We have coupled the recently released version of UrbanSim with MATSim to create a modeling framework in which to study this problem.
- I created several tools to visualize geographic and road network data to test the models.

## Triangle densest k-subgraph problem with integer linear programming 2015

with Bistra Dilkina

- Finding the Triangle Densest Subgraph of size  $k$  is a NP-hard problem that is useful for finding quasi-cliques in a graph.
- We are investigating finding and approximating hard instances of this problem with an Integer Linear Programming approach and comparing the performance against greedy heuristic based algorithms.

## Cellular automata networks for predicting weather 2015

with Dawn Wilkins

- This was my undergraduate honors thesis project. I examined simulating climate variables with cellular automata models.
- I was interested if adding in long range connections to the cellular automata model could improve the accuracy of the model by learning the influences certain climate indicators (like El Nino) have on local weather.
- During the project I automated the training of over 10,000 neural networks on the Mississippi Center for Supercomputer Research's cluster.

## Automating measurements of space plants 2015

with Josh Vandenbrink

- This was my Senior year capstone project. I was tasked with creating a framework that automated the data collection process from images of seedlings grown on the International Space Station.
- A lab in the UM Department of Biology received groups of 80+ images, showing the growth of up to 10 seedlings per image over time, then had to measure each seedling in each image by hand with a graphics program.

- I created a Python program that facilitated faster manual measurements and automatically performed OCR, perspective transformations, and image registration on these image groups to standardize the measurements as much as possible.

### Face recognition with limited training samples

2014

with Jianxia Xue

- This project was focused around the problems involved with automating classroom attendance using face recognition with a single training sample for each individual.
- I created a web application with a Python backend that performed online facial detection and recognition via a HTML5 webcam access.
- We examined improving the standard Local Binary Pattern Histogram approach for face recognition and using active learning to improve recognition accuracy.

## JOBS AND INTERNSHIPS

---

### Research Intern, Oak Ridge National Laboratory.

May 2016 - Aug 2016

- Created a web based visualizer that showed active hurricane tracks, and the probable power outages associated with the hurricane.
- Working on modeling human migration as a regression problem with deep learning techniques.

### Software Development Intern, FNC Inc.

June 2015 - Aug 2015

- Created address validation software using machine learning techniques
- Created an automatic XML posting tool to assist testing API's

### Software Development Intern, CSpire Wireless

June 2014 - Aug 2014

- Used Java with the Swing Framework and other enterprise level Java technologies
- Developed middleware REST services that facilitated communication between applications in different areas of the company

### Software Development Intern, FNC Inc

June 2013 - Aug 2013

- Worked on a development team using Agile Scrum methodology with 2 week sprints
- Used Team Foundation Server and Visual Studio with C# and ASP.NET
- Was able to integrate with the team quickly and wrote code that made it into production within the first week

### Software Development Intern, Wallace Community College, Selma, AL

Dec 2012 - Jan 2013

- Worked on starting development of an air traffic controller simulator program
- Used Java to implement a multithreaded server as well as a full screen client application that could interact with multiple industry standard flight simulator programs

### Assistant Network Administrator, University of Mississippi Computer Science Department

Jan 2012 Aug 2012

- Helped manage 79 laboratory computers spread over three different labs
- Created PHP based web application that ran on top of the Drupal CMS and interacted with departments active directory to manage student user accounts
- Recreated the department's Linux mailserver

Last updated: January 23, 2018

<http://calebrob.com/cv.pdf>