Internship Experiences
University of Illinois Urbana-Champaign
Department of Mathematics
Graduate Program

2021

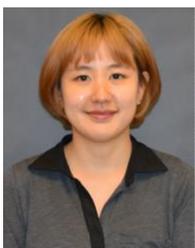Eion Blanchard
Microsoft Research

Wooli Chae
USAA Life Insurance

Ran Ji
Oak Ridge National Laboratory (MSGI program)

Derek Kielty
Sandia National Laboratories

Hanna Kim
Los Alamos National Laboratory (MSGI program)

Sophie Le
AbbVie (Inmas program)

Adriana Morales Miranda
Sandia National Laboratories

Souktik Roy
Schneider Transportation (Inmas program)

Emily Shinkle
Los Alamos National Laboratory (MSGI program)

Sarah Simpson
Ameren (Inmas program)

Nikolas Wojtalewicz
Vision Systems Inc.

Yong Xie
IBM

2020

Neer Bhardwaj
Argonne National Lab (MSGI program)

Eion Blanchard
Sandia National Laboratories

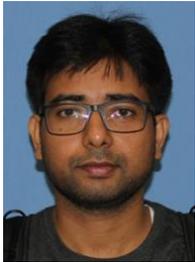Adriana Morales Miranda
Sandia National Laboratories

Shinhae Park
Corteva Agriscience (PI4 program)

2019

Eion Blanchard
Applied Research Laboratories at UT Austin (PI4 program)
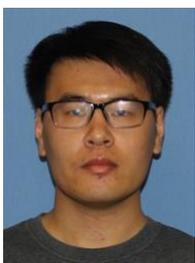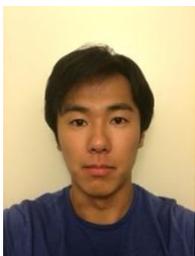

Ankush Hore
Ameren

Ran Ji
Wolfram Research (PI4 program)

Xujun Liu
AbbVie (PI4 program)

Shufan Mao
Psychology Department (PI4 program)

Tsutomu Okano
Industrial and Systems Engineering Department (PI4 program)

Shinhae Park
AbbVie (PI4 program)

Henry Solberg
NASA Kennedy Space Center

2018

Daniel Carmody worked with Professor Richard Sowers studying traffic in New York. One of the central goals was to determine the tradeoff between travel time and safety on various routes through Manhattan. This involved coming up with an efficient method to solve a bi-objective optimization problem (time and number of accidents) and writing python code on a computing cluster to perform the optimization. The project culminated in a paper, "Tradeoffs between Safety and Time: A Scale-Free Routing Approach," which is currently submitted.

Mary Angelica Gramcko-Tursi began an internship with the aim of rigorously quantifying uncertainty in object detection. The first step to this goal was to create and train an object detector for cars in clear satellite imagery. The point, however, is not to craft the best detector for all possible environmental circumstances, but to test and mathematically describe how a well-performing detector trained on clear images will fail when noise is introduced to the data. The noise could be internal to the satellite, or some environmental obstruction like clouds or fog. The summer internship was the start of a year-long project in which she is still involved, but over the summer she tested multiple approaches to detection and examined how they failed by looking at how they dealt with before settling on one approach, which I can now use for measuring uncertainty. Since her ultimate work is more theoretical, he also focused on the mathematics behind the algorithms used to train the different models.

Lina Li worked as an intern at Ameren innovation center. Her first project aimed to study demographic data of Ameren customers and determine the likelihood that a given customer will enroll in the special energy saving programs. This project involved using multiple approaches to perform feature selection and building predictive model with random forest classifier. Her second project focused on analyzing accelerometer data of power poles and proposing an innovative statistical approach to quantify pole health from existing pole sensor accelerometer data.

William Linz worked as an intern at Sandia National Laboratories. He worked on two projects related to optimization in optimal power flow problems. His research investigated linear and convex relaxations of the optimal power flow model. The first project concerned the scheduling of a power grid; the second project studied interdiction of a power grid. As part of these projects, he worked on optimal power flow models using the Pyomo modeling environment. William was also affiliated with the MARTIANS group at Sandia, where he was able to network with other math and statistics interns at Sandia, and through which he improved his knowledge of machine learning theory.

Michael Livesay's internship in summer 2018 was with Sandia National Laboratories. He worked on two projects one relating to social modeling and the other relating to security systems and their reliability. For the social modeling project, he applied methods of measuring the complexity of different aspects of the models and worked to justify or improve these measurements. For the reliability of security systems, we confined ourselves to the problem of the steady state case. Part of what he did was write a program which calculated conditional events determined by the structure of any directed graph.

Sarah Mousley Mackay interned at the Naval Research Laboratory in the Center for High Assurance Computer Systems branch. The goal of the project was to design network protocols that allow two parties communicate covertly without being detected by a warden watching their communication channel. She quantified the asymptotic behavior of a particular detector in network covert channels given the asymptotic behavior of embedding method. She also studied the maximum size of message that can be sent while ensuring unlikely detection. She also implemented various embedders and detectors in the software.

Dileep Menon worked as an Analytics Intern at John Deere World Headquarters in Moline, IL. He worked on a number of data science projects throughout his internship. In one project he used clustering algorithms to discover patterns in tractor machine data which he communicated with engineering and marketing teams. In another, the goal was to build a live Tableau dashboard that monitors data quality at the enterprise level. He also had the chance to participate in a company-wide data science competition in which he placed second. For the competition, he developed a model using XGBoost that predicts the rejection probability of a warranty claim for tractors and combines.


Joel Villatoro collaborated at the Ameren innovation Center with fellow graduate student Lina Li. See Lina's internship summary above.

Lan Wang worked as a researcher in the Data Science group at Open Data Group in Chicago. Her job was to research and develop a recommender system using machine learning models, intended to provide accurate and value-enabling recommendations on Mortgage-Backed Securitization (MBS) that their clients may be interested in. First, she designed, implemented and fine-tuned a Restricted Boltzmann Machine (RBM) with TensorFlow to realize a proposed collaborative filtering concept, only to discover that the prediction accuracy of clients' ratings on the products is suboptimal. Subsequently, she applied ensemble techniques on a collection of the Autoencoder model, content-based model and linear regression model, which resulted in a more sophisticated hybrid recommender system, leveraging three data sources in the computation: clients' ratings on the products, products' features and clients' characteristics. The eventual model accuracy is higher than 80%. In between Wang's own project agenda, she has also learned many other interesting projects from my colleagues such as natural language processing and time series modeling. "All in all, the experience of this internship was eye-opening and fruitful," says Wang.

Xiao Wang's second project with BudLab was to optimize the company's turnover model in a pilot zone and to support the scale-up of the model building process in other zones. She extracted and cleaned the data of the pilot zone from the database and performed exploratory data analysis. By applying natural language processing techniques to analyze the feedback data from HR, she designed and automated a ranking system for all the questions that could be provided to HR, to make sure that the effective questions are more likely to be selected in the future. Later she helped with the model scale-up by creating data anonymization process and building standard modules in the whole workflow. The project involved working with people from different departments in different zones, during which she learned a lot about how to collaborate and communicate with coworkers.
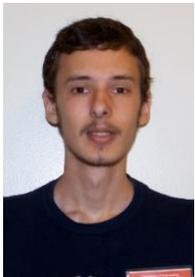
Benjamin Wright's internship at the National Center for Supercomputing Applications was supervised by Dr. Vlad Kindratenko. Benjamin researched existing methods for training neural networks, in order to figure out which methods scale well to large numbers of GPUs.
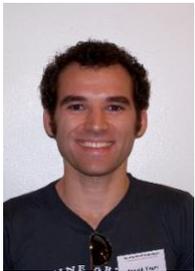
Derrek Yager was an intern at Sandia National Laboratories. He worked in the Sensor Exploitation Applications Department under the Autonomy for Hypersonics Mission Campaign. Satellites and other aircraft collect images using Synthetic Aperture Radar (SAR). This is expensive and results in a limited dataset. However, synthetic SAR images can be generated using a target's CAD model. Nevertheless, this process is not perfect, and we need a way to measure the quality of this synthetic image. Moreover, we would like to use machine learning and Generative Adversarial Networks (GANs) to transform these synthetic images into more realistic ones. Derrek created metrics for measuring the similarities of the real and synthetic images, and he developed the software to do it. He also ran the math for different loss functions to be used in the GANs.
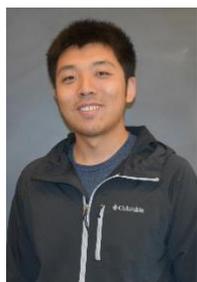
Wentao Zhang worked as a software engineering intern in the Compose Team of Alation on two projects related to prototyping a new search engine in Alation data catalog for metadata of sheets and dashboards of Tableau a popular business intelligence platform. The first project was indexing source tables of sheets and dashboards of Tableau. Each sheet and dashboard of Tableau may use source tables that are stored in some data servers. The goal of the project is to index these source tables so that the search engine can return correct sheets and dashboards whose source tables match query keywords. The second project was about data field peek to show business glossary definition. In Alation data catalog, users can write and store articles about definitions of business glossary (BG) terms. In the search result page, there are facets that show the data fields used in the sheets and dashboards. When users hover over a data field in the browser, if the name of the data field matches with certain BG definition, we want to show this matched BG definition in a peek to help users learn BG terms.

Dara Zirlin classified families of spider toxins using hidden Markov models, in a scientific internship with Dr. Brenda Wilson in the Department of Microbiology.

2017

During the summer, Stacey Butler worked with health data at the Champaign-Urbana Public Health District. Specifically, she looked for patterns in sanitary (restaurant) inspections for Champaign area food establishments. Several years of data were analyzed using Pandas in Python, checking for inspector bias regarding different cuisine types using OLS regression and variation among inspectors' average scores, and examining the effects of policy changes on inspection scores.

This summer Sneha Chaubey worked as an intern in the Scientific Content group at Wolfram Research, in Champaign. She continued her work from last year building on the Wolfram function site in Wolfram Alpha, collecting formulas, properties, theorems, etc. for both elementary and complex mathematical functions and presenting the information in Mathematica notebooks. Functions include the Riemann zeta function, Dirichlet L-function, gamma function, exponential, logarithmic, sine and cosine integrals, and Jacobi and Elliptic integrals. Another project involved writing programs to construct general differential equations whose solutions in special cases are the Heun functions or Legendre functions.

Yongwoon Escobar worked as an intern at Sandia National Laboratories in the Center for Cybersecurity Defenders program. His first project involved studying cryptographic protocols, gaining experience on implementing a post-quantum public-key protocol on hardware, and establishing a pipeline of tools to process documents of interest relevant to cryptography. His second project implemented additional features and solutions for a graph partitioning algorithm and library in C/C++ for large graphs, on the scale of billions of vertices and trillions of edges, for systems running in parallel. Near the conclusion of his internship, he met and interviewed with various Sandia teams and organizations for potential future employment.

Benjamin Fulan
AbbVie (Research Park)

The internship this summer on the tinnitus project in the Speech and Hearing Science Department (supported by the PI4 grant) continued Mary Angelica Gramcko-Tursi's work from last year. The overall aim of the project is to find biomarkers for Tinnitus that can be used for diagnosis. The project involves lead matrices, which are used to recover the cyclic order of activity levels in specified regions of interest in the brain, based on fMRI data. Mary Angelica examined differences in lead matrices over each session as well as across different sessions, by projecting them onto subspaces of high variation to determine whether there was any discernible difference between the two subject groups (with and without tinnitus). Her work showed that the data being extracted from lead matrices is in fact meaningful, rather than being a result of noise. These methods may also be useful in other diagnostic settings too.

During this summer, Tigran Hakobyan was a PhD software engineering intern at Facebook. He worked on designing a Traveling Salesman heuristic algorithm for nonlinear path weight and finite traversal duration constraints with arbitrary spatial graph partitioning. His work provided distributional statistics on the number, costs, and values of salesman routes under different constraints.

David Hannasch
Department of Defense
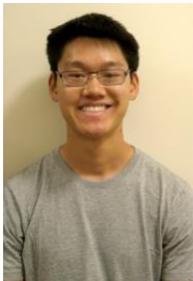
Derek Jung worked at the North Las Vegas facility of National Security Technologies, in an internship arranged through the NSF Mathematical Sciences Graduate Internship program. He worked on deblurring images, by investigating a mathematical model of the blurred image of an opaque edge. That problem serves as a model for understanding how all images are blurred. At a technical level, his work showed a certain integral operator equation is ill-posed while a Tikhonov regularized problem is well-posed, by using techniques from measure theory, functional analysis and harmonic analysis.

William Karr worked at the Caterpillar Data Innovation Lab in the Research Park during the summer of 2017. The Data Innovation Lab hosts a diverse set of projects related to data and technology, proposed by dealers and clients of Caterpillar. Bill worked on developing a mobile web application called vSite for use by workers and supervisors in mining sites. The application would be used to mark problem areas and keep track of work that had been done around the site.
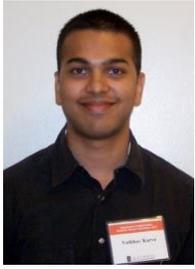
Vaibhav Karve worked with Professor Richard Sowers in the Department of Industrial and Enterprise Systems Engineering this summer. He continued his research into taxi traffic patterns in New York City by helping to develop a new algorithm for Non-negative Matrix Factors, a technique that is gaining traction for finding low-rank structures in large datasets and for predictive estimation of missing data values. Vaibhav is exploring also the underlying topological and geometrical structure of the New York City road network. The insights from this work could find use in the planning and maintenance of large cities.

As an Intern at the New York office of Ernst & Young's Quantitative Advisory Services (QAS), Artur Kirkoryan consulted on projects at one of America's big banks. QAS validates financial models, examining their underlying theory, implementation, and performance. The work requires a good grasp of stochastic processes and statistics, as well as some basic coding skills: one project used linear regression techniques for analyzing data, and a second used dynamic programming for estimating the values of certain options. Artur appreciated gaining valuable industry experience and a wider perspective on applications of mathematics.

Nicholas Kosar
Personify (Research Park)

Paulina Koutsaki spent the summer of 2017 as a software development intern at the AbbVie Innovation Center in the University of Illinois Research Park. As part of a server migration for the Statistical Analysis Environment, there was a need to replace custom Unix-only utilities that were developed in the 1990s. Her job focused on analyzing off-the-shelf tools and giving a detailed description of their features and how they could replace the functionality of the utilities. Moreover, during her internship she participated in a hackathon where she had the opportunity to collaborate with corporate employees and other interns and win the Innovation Award.

Michael Livesay internship at the Champaign-Urbana Public Health District was supported by the PI4 grant. There he was given death certificate data to analyze. By using the two-sample Kolmogorov-Smirnov test was able to find a dependency between a person's life span and their month of brith, even when looking at only people who survived to adulthood. Further, using the statistical Chi-squared test he found confirmable trends between the day of the week one dies and location of death.

Amita Malik continued her theorem curation project from last year at Wolfram Research, in the Scientific Department. This summer her main focus was to format concepts from complex analysis, written in Mathematica last year, and create a database of definitions and theorems in complex analysis. The long term goal is to be able to verify results by computer, based on this database.

During the summer, Shufan Mao worked in the Language Acquisition Lab of the Psychology Department, running a study on children and doing eye-coding work. He also attended weekly lab meetings for reading papers and discussing the experiments conducted in the Lab. Meanwhile, he participated in group meetings for an ongoing project called "Baby SRL", led by Professors Cynthia Fisher (head of Language Acquisition Lab) and Daniel Roth (Computer Science Department). They are working on a computational model for children's language acquisition, especially syntax acquisition. The model is built on an unsupervised parser, and Shufan worked with Dr. Lori Moon (Linguistics Department) to understand the algorithm of the parser and determine what the mechanism could contribute to the Baby SRL model. The project is ongoing, and he will continue working with the group in the coming year.

Cara Monical
Sandia National Labs

Sarah Mousley was an intern at Sandia National Laboratories, funded through the NSF Mathematical Sciences Graduate Internship (MSGI) program. She worked in the Center for Computing Research (CCR), optimizing an energy function used by the meshing community, with the goal of measuring and improving mesh quality. The energy of a mesh measures its fitness for usage in Discrete Exterior Calculus, DEC, a method for numerically solving Partial Differential Equations (PDEs). More specifically, the energy bounds the discretization error of the Hodge-star operator that can be used for numerical formulations of many PDEs. She explored the landscape of the energy function to better understand how to optimize it, using software she developed.

During Joseph Rennie's summer internship at the Champaign-Urbana Public Health District with funding from the NSF PI4 grant, he examined birth records for Champaign County from 2014. Given hundreds of variables, he searched the literature to create summary variables for whether or not a birth should be considered unusual. Then, grouping by zipcode, he searched for deviations from uniformity in the spatial distribution of unusual births. He found a potentially significant trend, but needs more data before reaching a conclusion with a meaningful degree of certainty.

This summer Vanessa Rivera-Quiñones interned at the lab of Professor Carla Cáceres from the Department of Animal Biology. Her work focused on studying the changes in the phototactic behavior (i.e. movement towards or away from light) of the zooplankter *Daphnia Dentifera* as a response to infection pressures. As part of her internship, she received lab training and collaborated on the experimental design process, a new experience for her. Vanessa will continue gathering data through the fall semester with the ultimate goal of building a mathematical model that describes this phenomenon.

Nishant Rodrigues worked with the Formal Systems Laboratory in the Computer Science Department to help write an executable semantics for the Ethereum Virtual Machine in the K Framework. The group submitted a paper for review. The K framework has it's foundations in Rewriting and Matching Logic. It allows creating mathematically well-defined and executable specifications for programming languages. With these semantics, K can use Reachability Logic for program verification, that is, to automatically generate proofs of correctness for properties of programs written in these languages. Rewriting Logic, Matching Logic and Reachability have solid mathematical foundations. Many members of the Formal Methods sub-department are interested in working more closely with the Mathematics Department.

Hao Sun's project at the Ameren Innovation Center in the Research Park aimed to describe the relationship between the temperature and Ameren customers' usage, and hence to predict future usage based on historical temperature data. Hao combined smart-meter customer data and NOAA's temperature data in several different mathematical models to predict daily customer usage. Model errors were analyzed using linear regression, multiple linear regression, and modified multiple linear regression. Based on the multiple linear regression model, he used machine learning techniques to construct the modified multiple linear regression model, which was found to give the best predictions.

This summer Albert Tamazyan worked as a software engineering intern with the Communication Products team at Yahoo. His projects in the Android Mail group were about developing machine learning algorithms for categorizing emails in mailboxes and for generating message filters for user accounts. He learned relevant tools and technologies, and gained valuable programming experience.

During the summer of 2017, Corbin Tucker worked as an intern on the Predictive Analytics team at Blue Cross Blue Shield of Tennessee in Chattanooga. His responsibilities included researching possible analysis methods from peer-reviewed articles, assisting in identifying predictive variables, and cleaning text data for a machine learning algorithm. He also had a project of his own predicting patients who are most likely to suffer from a first time heart attack or stroke. As part of the internship, Corbin was introduced to new programming languages, analytics tools, and statistical methods which helped him to build his data analysis skills.

Venkata Sravani Vadali worked at the Ameren Innovation Center (Research Park) for the summer on two separate projects. In one project, her team had to model a new billing scheme for Ameren that analyzes a demand-based charge as opposed to traditional billing. They analyzed terabytes of customers' data and came up with mathematical equations that related the energy charge to the demand charge to keep Ameren's revenue unchanged. In her second project, Sravani and the team determined the environmental factors that would affect the movement of utility poles. Various modeling and machine learning techniques were implemented in the project. She will continue working on these projects in the Fall. Her analyses was done in R, Python and SQL / Hive.

During the summer, Lan Wang worked as a researcher in the Data Science group at John Deere, in the Research Park. Her task was to improve the accuracy of position data and smooth out the driving paths using Kalman Filter techniques, and then to analyze the paths using machine learning techniques, visualize the mower stripes, and finally deploy all Python scripts on AWS (Amazon Web Services) Lambda platform. During this work, she had the opportunity to combine her Math knowledge with real industrial problems. When applying the Kalman Filter, she built both linear and non-linear models to make the filter work better. When doing visualization, she took advantage of geometric insights to find efficient and effective methods to clean and parallelize mower paths. Meanwhile, she improved her programming capabilities and learned how to do operations in the cloud, taking advantage of Amazon Web Services, Amazon Turk, and the Google Compute Engine. She really enjoyed working at John Deere!

Xiao Wang's internship at the Bud Analytics Lab in the Research Park centered on customer data sets from the eCommerce department in China (especially through Tmall.com and Wechat stores). Her job included: analyzing customers' behavior, especially different promotions and how they affect member conversion rate; training random forests to classify golden/platinum members; clustering customers from their behavior using Kmodes, ROCK (RObust Clustering using linKs) and Correlation Explanation algorithms; applying time series analysis to forecast eCommerce transaction volume.
She enjoyed learning how to connect technical results to business insights.

This past summer at Dow AgroSciences, Argen West worked on weed resistance modeling. First he learned about the different genetic mechanisms that cause weeds to build up resistance to herbicides and the methods that are used to model those mechanisms. Then he wrote Matlab code to predict the onset of herbicide resistance depending on different input variables. His internship culminated in a co-authored paper with his adviser that is set to be submitted for publication.
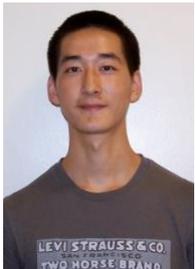
During the summer, Qiang Wu worked on the Networked Stackelberg Model under the supervision of Professor Subhonmesh Bose (ECE Department). The specific project involved simulating the model and estimating the efficiency loss in a general network. Qiang began by learning the Classical Cournot Model, then coded up a simulation of the general Stackelberg Network competition. Besides that, he also revised the proof of some important theorems in a theoretical analysis of the model. During the process, he gained experience on using several optimization software packages.
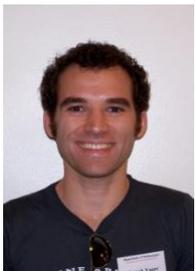
During the summer of 2017, Derrek Yager was partially supported by the National Science Foundation to work for the New York City Department of Transportation on taxi path inference problems. The NYC Taxi & Limousine Commission gathers GPS location trails, or "breadcrumbs," from taxis at 1-2 minute intervals. A given month can produce over 300 million breadcrumbs. Due to GPS noise as a result of the urban canyon effect, the breadcrumbs rarely fall on the road. Using pgrouting, python, and parallel processing, Derrek separated the breadcrumbs into trips and then modeled potential paths connecting the breadcrumbs. Using Markov chains, he chose the most probable path so that he can now analyze the data using his own modification of Sparse Non-negative Matrix Factorization. This work extends a PI4 project from the previous summer with Vaibhav Karve, and ties in with an Illinois Geometry Lab project he has been co-supervising during the academic year.

Instagram is a rapidly growing platform with over 500 million users sharing and consuming content every day. The Instagram Ads team in the Menlo Park, where Mingyu Zhao spent her summer internship, connects businesses to people by building a platform to show ads to the users. The team is focused on using state-of-the-art machine learning techniques to understand users' interests and increase the relevance of business content shown on Instagram. In particular, Mingyu worked on ad auctions, which include selection and pricing mechanisms for choosing the best ad based on the predicted performance. This project relied on game theory to make the auction incentive-compatible.

During the summer, Dara Zirlin attended the IMA Math-to-Industry Boot Camp at the University of Minnesota. This was a six week program, where the first three weeks focused on courses in programming, data analysis, and mathematical modeling. The next three weeks were spent on group projects. The first project involved predicting the outcome of men's singles tennis tournaments. The second involved optimizing oil well placement to maximize profit.

2016

Yingjie Bai
AXIS Capital (Research Park)

During summer 2016, Stacey Butler was supported by the PI4 program to work in the O'Dwyer lab in the Plant Biology Department, on questions from theoretical ecology involving systems of interacting species. The evolutionary dynamics of a network of species containing mutualistic and competitive interactions was modeled using a Lotka-Volterra equation. The first question was an investigation into a pattern noticed in the outcome of an optimization algorithm performed on networks of interacting species. The species interactions were altered to optimize species abundances. It was observed that the number of real eigenvalues of the linearized system at the end of the optimization was significantly increased over the initial situation. An explanation for this pattern and its ecological consequences could be useful for understanding ecological systems in nature. The second question involved searching for patterns in the resistance of a network of interacting species to perturbations in species-intrinsic growth rates. This perturbation can be thought of as representing environmental fluctuations. The structure of the network and the interaction types can impact this resistance to perturbation. These questions were explored with simulations written in Python.
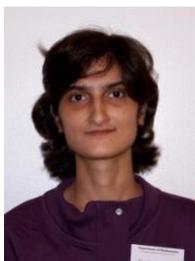
This summer Sneha Chaubey worked as an intern in the Scientific Content group at Wolfram Research, in Champaign. Her main job was to build on the Wolfram function site in Wolfram Alpha. She collected formulas, properties and theorems for both elementary and complex mathematical functions, including the Riemann zeta function, Dirichlet L-function, gamma function, exponential, logarithmic, sine and cosine integrals, and Jacobi and Elliptic integrals. The other project she worked on involved writing programs to construct general differential equations whose solutions in special cases are the Heun functions, Legendre functions and other associated functions.

Jed Chou worked at Ab InBev in the summer of 2016. He developed and integrated an application on Teradata to generate weekly growth forecasts of all Ab InBev craft beer brands. For a second project, he applied machine learning to optimize online auctions hosted by ABI and was able to provide specific recommendations on which auction type to run and which suppliers to invite for different auctions. Jed learned about time series forecasting, auction theory, and machine learning. He also enjoyed many tasty beers.

Erin Compaan worked at CERT, a division of the Software Engineering Institute (SEI) at Carnegie Mellon University. SEI is a federally funded research corporation focused on Internet security. An SEI employee contacted the department in search of mathematics interns and corresponding with her led to the internship offer. The project was open-ended, and interns were encouraged to find problems to pursue which accorded with their skills and interests. Erin's work centered on analyzing Internet connectivity data, with the object of understanding the growth and vulnerabilities in the network. Day-to-day, this involved a lot of data parsing, programming, and visualization. However, she was also encouraged to spend time learning relevant theory to inform her investigations, such as the mechanics of Internet traffic, statistical concepts, and random network theory.

Lin Cong worked with Wolfram Alpha this summer, for his second summer and the fourth semester internship there. After more than one year's work, he helped finish the Function Space project. His duty there was to summarize properties of function spaces, so that users can access them on the wolframalpha.com website and manipulate the connections between different spaces. He was also participated in the Semantic Math project, which is an on-going and ambitious new development. The first goal is for computers to "understand" pure mathematics theorems, and eventually, to let the computer search for and even give proofs for particular theorems. The group has so far done excellent work towards the first goal by implementing mathematical definitions and structures.

Eliana Duarte - Wolfram Research
During the summer internship, Eliana tested algorithms to compute greatest common divisors (GCDs) of multivariate polynomials over finite fields and over algebraic extensions. The main algorithm she tested uses a Gröbner basis to compute GCDs using syzygies, and together with her supervisor, she wrote the paper "Polynomial GCDs by syzygies" which has been accepted for publication at the conference SYNASC 2016. A second project involved writing a function to factorize multivariate polynomials over finite fields. The main ingredient of this algorithm is to reduce to the univariate case and do Hensel lifting to recover multivariate factors.

For his 2016 summer PI4 internship, Matthew Ellis worked with Professors Andrew Ferguson (Materials Science) and Lee DeVille (Mathematics). The project developed from Ferguson's research into the reconstruction of single molecule free energy surfaces (FES) from experimentally measurable time series. The FES can be conceived of as a low-dimensional manifold in a high-dimensional space of atomic coordinates. However, these coordinates are experimentally unavailable, meaning the FES cannot be recovered precisely. If a coarse-grained time series is accessible (e.g., the molecular size as a function of time), then with Takens' Theorem from differential geometry coupled with nonlinear manifold learning techniques, a delay map can be used to recover the FES manifold up to diffeomorphism. Matthew worked with these delay maps, deriving bounds for the Jacobian of the diffeomorphism and modeling simple systems with MATLAB to verify the bounds. The implications of this work provide a means to place rigorous bounds on the quality of the reconstructed FES. This project introduced him to new concepts in differential geometry and dynamical systems, and he plans to continue working with Ferguson's group into the Fall semester.

Ian Ford
Wolfram Research

Benjamin Fulan worked on data analytics projects at the Ameren Innovation Center at the Research Park, during an internship supported by the PI4 grant. He collaborated with other interns on customer usage clustering and detecting anomalies using Ameren's new cloud-based data. Based on usage patterns, they identified six main clusters containing more than 90% of customers. To detect anomalies, he trained a machine learning model to predict customers' energy usage patterns for the first three months of 2016, using Ameren data collected by smart meters over the year of 2015. Both projects required a strong knowledge of linear algebra, statistics and programming in Python.

During her PI4 internship, Mary Angelica Gramcko-Tursi worked with the Speech and Hearing Department under the supervision of Professors Fatima Hussein and Yuliy Baryshnikov. Her work was part of a larger project aimed at finding a biological basis for Tinnitus. As of now, fMRI scans of people with normal hearing and those with Tinnitus are virtually indistinguishable. If a difference could be found, the discovery would help establish a proper diagnosis and open doors for new treatment. Unfortunately fMRI images have the limitation of detecting brain activity in a single instant, while it is possible that different regions work together over an extended period of time and are activated in patterns or cycles with time delays between each region, rather than activating instantaneously. In order to detect differences in this delayed behavior, Mary Angelica computed "lead matrices," matrices whose elements quantify time delays between levels of activity in certain regions of interest (ROIs) under study. In order to isolate possible differences, she used non-linear programming and optimization techniques to find optimal separating hyperplanes of various pairs of sets of lead matrices, this time interpreted as points in a Euclidean space, differentiated either by subject category or choice of time interval. Finally, she used statistical tools to detect levels of variation between subject groups and time intervals. By testing for separability and using simple analyses on the lead matrices themselves, she was able to isolate certain ROIs which may point a way forward for further research on the project. Throughout the entire process, Mary Angelica learned about the wide range of mathematical tools that can be used in analyzing biological processes.

During the summer of 2016, Byron Heersink held an internship at Sandia National Laboratories in Albuquerque. There, he joined an ongoing project in developing and analyzing certain types of game-theoretic models, and in particular, to contribute rigorous mathematical analysis of the models. In addition to simplifying and improving proofs in some of Sandia's earlier work, he developed algorithms for performing analytical calculations of certain outcomes of the models, and implemented them in Python. He also assisted in the validation of model simulations developed by other interns. Another component of the work was investigating the possibility of applying recent work in control theory to complex systems research at Sandia. Lastly, Byron took part in a program giving interns the opportunity to learn about Sandia's energy research, which included many interesting tours of various Sandia facilities.

Elliot Kaplan worked at Wolfram research during the summer of 2016, in an internship supported by the PI4 program. He worked with a team to develop a way to talk about theorems and definitions in pure mathematics using the Wolfram language. This included numerous discussions about the best way to encode basic ideas in set theory (such as products of sets) as well as abstract concepts (uniqueness up to an equivalence relation) and concrete concepts (Cauchy sequences). In doing this, Elliot learned how pure mathematics, even foundations of mathematics, can be useful in an industry setting. Throughout the summer, Elliot worked with another intern (Ian Ford) to encode the first six chapters of Munkres' classic text "Topology".

William Karr worked at the Caterpillar Data Innovation Lab in Research Park during the summer of 2016. The Data Innovation Lab has a diverse set of data and technology related projects proposed by dealers and clients of Caterpillar. Some of the projects involve data mining and machine learning. Bill worked on developing a mobile application called Track & Trace to be used by municipalities to track and manage fleets of weather-related vehicles. The application allows operators to tag problem areas and allows county engineers and supervisors to monitor progress. This application could eventually be used to analyze historical records to determine how to more efficiently allocate resources for fleets of such vehicles.

As part of the PI4 program, Vaibhav Karve studied the traffic patterns in New York city taxi data. Under the guidance of Professor Richard Sowers in the Department of Industrial and Systems Engineering, he analyzed the post-processing traffic data of nearly ten thousand taxis running on the 260,000 links in the New York city road grid. This data spanned every hour of the day, every day of the year, for four years. Vaibhav used algorithms of Non-negative Matrix Factorization (NMF) to produce good low-rank approximations to the taxi data, and then isolated patterns in the traffic with the aim of improving our understanding of traffic flow and traffic dynamics. To achieve these goals, he used an algorithm known as Sparse-NMF (a combination of NMF and k-means clustering). Sparse-NMF allowed him to compress data while simultaneously clustering road links based on their traffic patterns. This clustering could be used in the future to understand individual components that contribute in the overall traffic dynamics. For this project, Vaibhav wrote code in Python and ran it on the Illinois Campus Computing Cluster. He will continue the visualization aspects of this project with a larger team as part of the Illinois Geometry Lab in Fall 2016.

Nicholas Kosar spent the summer of 2016 as an intern at Personify, on an internship set up through the PI4 program. Nick worked on two projects: exploring the feasibility of augmented reality with a standard smart phone and using a webcam with a depth sensor to create virtual meetings. Through this internship, Nick gained exposure to current research that is being done in augmented and virtual reality; he also gained experience working in computer vision.

Instagram is a rapidly growing platform with over 400 million users sharing and consuming content everyday. The Instagram Data Products team in the Bay Area, where Shiya Liu worked for the summer, helps users to find content of interest by building search, content discovery and recommendation products on Instagram. The team is focused on using state of the art Machine Learning techniques to understand users' interests and increase the relevance of content shown on Instagram. In particular, Shiya worked in the Media Ranking team. Interesting challenges for the coming year include mapping users' interests to content consumed on Instagram, discovering and identifying events happening on Instagram, and connecting users to the most relevant accounts.

As an intern in the Content Development Division at Wolfram|Alpha Scientific Content, Amita Malik's job was to build data related to number theory and complex analysis, by investigating results equivalent to the Riemann Hypothesis (RH). Since, RH arises in many seemingly different fields of Mathematics, it would be useful to find all this data together, and would save a great deal of time and help mathematicians with their research. The second part of Amita's project was to curate theorems in complex analysis, as a first step towards automatic theorem proving in this area using *Mathematica*.

During the summer of 2016, Andrew McConvey worked as a summer intern in the Securities division at Goldman Sachs in New York. Andrew rotated between two different "Strats" teams, working first with the Foreign Exchange Electronic Trading desk before moving to the Structured Credit Trading team. His responsibilities included analyzing methods of streaming data to clients, specifically how to balance the need for real time data with the cost of sending new information. He also developed a tool to help evaluate and easily display the cost and risk of a trade. As part of the internship, Andrew was introduced to new programming languages and packages, which helped him to improve his coding and data analysis skills.

Anna Mitchell worked this summer at RLI Insurance Company, a specialty product insurance company headquartered in Peoria. She worked with the reserving team in the risk services department, whose main responsibilities include determining the amount of money to set aside, or "reserve", for future claims and liabilities. Anna primarily provided assistance with quarterly reserve reviews, and reconciled data for reserve studies, but also had a project of her own analyzing the effectiveness of five methods used by the reserving team to predict future loss development. She had the opportunity to present this information to all the top executives of the firm. This internship experience enhanced Anna's professional skills and introduced her to several successful practitioners in actuarial science.

During the summer of 2016, Sarka Petrickova worked on data analysis and model fitting in the biological sciences, in an internship arranged through the PI4 program. Under the guidance of David LeBauer (Energy Biosciences Institute) and Yuliy Baryshnikov (Mathematics), she worked together with another intern, Benjamin Wright, on two projects. In the first project they studied light interception in canopies. The second project concerned root systems of canopies like sorghum or maize. Here, the main goal was to compare the precision of various measuring tools that are, or could be, used in the field to estimate important parameters of the root system. Sarka reports that this internship was a great opportunity to learn about crop sciences, improve her programming skills, and get a feel for what it is like to work in an applied setting.

Michael Raftery
Centers for Medicare and Medicaid Services

Sepideh Rezvani worked on data analytics projects at the Ameren Innovation Center at the Research Park. She collaborated with other interns on customer usage clustering, detecting anomalies, and sentiment analysis projects using Ameren's new cloud-based data. Based on usage patterns, they identified six main clusters containing more than 90% of customers. To detect anomalies, she trained a machine learning model to predict customers' energy usage patterns for the first three months of 2016, using Ameren data collected by smart meters over the year of 2015. This project helps detect fraud, and the goal for the Fall is to improve the model using neural networks. In the sentiment analysis project, Sepideh analyzed Twitter data in order to alert Ameren if the general sentiment is "very negative" in a specific location, for instance due to an outage. She reports that the internship provided a great opportunity to learn cloud-based tools for cleaning, manipulating, analyzing and visualizing data, and also presentation skills using Tableau dashboards. All the projects required a strong knowledge of linear algebra, statistics and programming in Python.

This summer, Albert Tamazyan rejoined the "Scientific Content" group of WolframAlpha and continued a project from last year's internship about adding Entities for new families of mathematical functions. He collected formulas and theorems from the literature, studied the underlying functions and derived new formulas. Relevant properties include integral representations, functional equations, and asymptotic expansions.

Hongfei Tian spent the summer as one of twelve full-time residents in the DS12 data science residency program embedded in DataScience, Inc., a Culver City, CA startup. Hongfei was one of the 2% of applicants selected to participate in the highly competitive program. DS12's elite and intensive curriculum goes beyond the fundamentals of data science to dive into terabyte scale data sets while working with purely functional Scala and the Apache Spark ecosystem. During the 3-month program, residents mine and model data from real DataScience, Inc. data sets to develop actionable insights for clients. During the final capstone project, Hongfei worked with a small team of other DS12 residents and DataScience employees to engineer a data pipeline that supports training and testing graphical models of Waze event data for the city of Los Angeles.

Bolor Turmunkh worked as a Data Science intern for the Bud Analytics Lab, the data analysis research lab of ABinBev, located at the Research Park. She investigated an exciting data set consisting of consumers' basket items, and developed methodology and models to predict beer purchase behaviors based on consumers' purchase history. Besides getting hands-on experience and learning what it means to deliver value to a company, she learned data analysis topics such as Discrete Choice modeling and Bayesian Hierarchical modeling and visualization tools that contributed to her development as a Data Scientist.

Argen West worked as a student researcher at the John Deere Technology and Innovation Center in the Research Park in summer 2016. He collected and analyzed crop information to better understand causes of yield variability, which involved in both outdoor collection of field data and subsequent statistical analysis and report development. He also worked on a visualization project using soil moisture data and another project focused on reducing construction equipment input cost. Argen gained considerable experience applying his knowledge to solve quantitative problems for a variety of business applications.
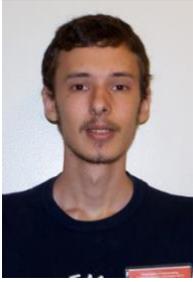
Benjamin Wright
David LeBauer (Energy Biosciences Institute)
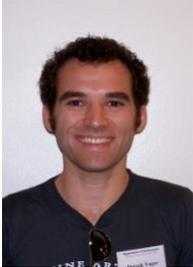
Derrek Yager worked with Professor Richard Sowers in the Department of Industrial and Enterprise Engineering in the summer of 2016, on an internship funded by the PI4 program. He developed a low-rank approximation of a large set of New York City taxi traffic data. From this analysis, they will extrapolate trends to characterize the general traffic behavior. Due to the large size of the data set, Derrek worked with Python and the University of Illinois Campus Cluster to perform the factorization and analyze the data. Derrek will continue this project in the Fall with students in the Illinois Geometry Lab, teaching undergraduates the basics of Python and developing their visualization skills.

Bingji Yi - Goldman Sachs (New York; quantitative analyst in their controllers team)
During the summer at Goldman Sachs, Bingji worked on price verification in the controllers modeling team. Price verification is an independent modeling process that uses external market data to verify the internal pricing models. The specific project involved price verification of certain exotic interest rate products. Bingji began by learning the Libor Market Model, which is a continuous-time diffusion model, and then developed an approximation method based on linear regression. He ran tests on the method and implemented it into the internal programming environment at Goldman Sachs.

At the Research Park this summer, Shibo Zhu worked at AXIS Reinsurance on selecting and capturing data for professional liability from historical files. Then he used the captured data for pricing analysis.

During her summer 2016 PI4 Internship, Dara Zirlin worked in the Pathobiology Department of the College of Veterinary Medicine with Professor Rebecca Smith. She modeled the spread of bovine tuberculosis in a population of deer in Michigan, with the aim being to study how decreases in hunting will impact its prevalence in the deer population. During this experience she acquired useful programming skills and learned to assess and interpret scientific literature to solve problems.

Nerses Aramyan: During his internship at Wolfram Research in Summer 2015, Nerses Aramyan investigated and developed identities for classes of special functions such as the Wigner D, Siegel Theta, and Bell Y functions. These identities will now be included on the Wolfram Alpha website and in the knowledge base of the Mathematica software. Nerses converted abstract mathematical expressions into a machine-readable form, and tested the formulas extensively both symbolically and numerically. He reports that "one of the most satisfying theoretical topics I learned in the course of this work was the connection between Lie group theory, representation theory, harmonic analysis, and the theory of special functions."

Hannah Burson worked with Professor Rebecca Smith in the Pathobiology Department of the College of Veterinary Medicine in the summer of 2015, an internship arranged through the PI4 program. She worked on models of the spread of Orf, a poxvirus common in sheep, which persists even in closed populations. As little is known about the transmission dynamics of Orf, Hannah used the Python programming language to implement several different possible models. Collaborators have designed a study to collect data that will be compared with the output of the code, in order to better estimate the dynamics Orf transmission.

For the summer 2015 PI4 internship, Stacey Butler worked on a project in James O'Dwyer's lab in the Plant Biology Department. The project explored properties of communities of interacting species, with competitive, mutualistic and trophic interactions. She studied a variety of dynamic evolutionary models, mostly involving ordinary differential equations of Lotka-Volterra type, taken from a variety of ecology papers searching for consistent patterns in the structure of the communities. In particular, the project examined the spectrum of the interaction matrix and of the Jacobian of the system, and the eigenvector associated with the leading eigenvalue, and the models were analyzed via simulations written in Python.

In the summer of 2015, Jed Chou worked in Professor Tandy Warnow's lab where he and his teammates empirically evaluated a new method, SVDquartets, for inferring evolutionary phylogenies from DNA sequences generated under the Multispecies Coalescent (a statistical model in which gene trees can be topologically different from the overall species tree). A contentious topic in this area is whether so-called "summary methods", which infer tree topologies on individual genes and then combine these gene trees into an overall species trees, should be used to infer species trees. SVDquartets was designed to bypass this issue of gene tree estimation error, but had not yet been compared to other coalescent-based species tree estimation methods. Jed and his team compared it with several other leading methods for phylogenetic inference on simulated datasets with very short gene alignments. One interesting outcome was that leading summary methods can in fact be highly accurate relative to the other methods on some datasets with short sequences. Through this internship, Jed learned about using simulations to understand algorithms, improved his Python and UNIX scripting skills, and he also helped write a paper that will be published in BMC Genomics.

Erin Compaan spent the summer interning at a communications consulting firm in San Diego. She had met the director of the company at a lecture the previous year, and arranged the internship herself. The projects involved using machine learning algorithms to sift through large volumes of data and identify salient characteristics. While programming was a big part of the day-to-day job - mostly in C and Python - the best part of the summer was the opportunity to collaborate with senior research scientists and other interns to make progress on a relevant and challenging problem.

Lin Cong worked with Wolfram Alpha this summer as an intern, in the scientific content group. His duty there was to summarize properties of function spaces, so that users can access them on the wolframalpha.com website and manipulate the connections between different spaces. To achieve this goal, Lin consulted academic articles, compared between different authors and references, determined the desired properties, and finally encoded the properties into the Data file. The last step involves combining mathematical expressions with the Wolfram language (Mathematica software). He used knowledge from Functional Analysis and Harmonic Analysis repeatedly, along with learning lots of new mathematics. The team environment in the scientific content group was friendly and helpful, and Lin enjoyed the internship tremendously.

Elizabeth Field, along with fellow mathematics student Goran Tomic, held an internship at the Institute for Genomic Biology in Summer 2015, organized through PI4 and supervised by Dr. David LeBauer. They joined the Predictive Ecosystem Analyzer (PEcAn) project, which is a toolbox consisting of two main parts: 1) a scientific workflow which manages environmental data and which is compatible with various ecosystem models, and 2) a Bayesian data assimilation system which synthesizes environmental data with the ecological models. Their project was aimed at developing a method of assimilating observations of biomass in order to provide better estimates for initial parameter values for an ecosystem model. To do this, they wrote an R package which implements a particle filter method. Through this internship, she learned about Baysian data analysis and gained valuable programming skills.

For the second year in a row, Meghan Galiardi held a summer internship at Sandia National Laboratories in New Mexico. She mainly worked on her thesis research while at Sandia, but took several steps to connect her research with projects at Sandia. First, Meghan worked on a team trying to identify Sandia models in which complex systems features play an integral role. Meghan's role was add more mathematical/analytical capabilities to the project. Second, Meghan worked closely with two staff members, explaining to them the details of her thesis and brainstorming possible applications to Sandia projects. Additionally, Meghan learned skills in project development, presentation-giving, and interviewing. Sandia has a very welcoming environment and is a great place to work. Meghan advises anyone who is considering an internship at Sandia to meet with representatives and learn about opportunities at the Labs. Meghan also thanks the PI4 grant for travel support.

Alessandro Gondolo worked as a full-time intern with Agrible Inc., a Research Park start-up providing actionable information to growers. Alessandro's internship was mediated by the department's PI4 initiative. The projects he worked on bridged the gap between mathematics and computer science—taking his knowledge of numerical analysis and PDEs and implementing physical models in Python. The experience pushed Alessandro to better understand both broad computer science concepts such as object-oriented programming and design patterns and day-to-day coding etiquette such as data scraping and version control system, Git. He gained experience with code reviews and ticketing systems, JIRA and Trac. At the end of the summer, Agrible Inc. offered Alessandro a fall internship; he has accepted the position and extends his thanks to the PI4 initiative for the opportunities PI4 provided.

Byron Heersink worked under the supervision of Dr. Samuel Beshers in the Entomology Department on modeling ant colonies, in a Summer 2015 internship arranged through the PI4 program. He designed and wrote a program in Python emulating the behavior of an ant colony, using a stimulus-threshold response model to control how workers perceive and respond to various tasks. Hence Byron could investigate how the level of need for a task or the sensitivity of workers to perform a task would affect the patterns of division of labor and the colony's ability to manage its task needs.

The Ameren Illinois Innovation Center hosted Sushma Kini during Summer 2015, for an internship arranged through the PI4 program in the Department of Mathematics. During her internship period in the customer analytics project, she analyzed the customers of Ameren Illinois and their energy usage patterns by means of statistical inferences and Python programming. By the end of the summer, the teams could segment the customers and identify which energy efficiency programs should be offered to them. The work culture at Innovation Center provides a sound technical learning experience along with a focus on opportunities for developing one's leadership qualities. "Big data" analysts are in high demand in the job market right now, and Sushma highly recommends working as an intern to transition smoothly into a full time career in the field.

Nicholas Kosar spent the summer of 2015 as an intern at the Ameren iCenter in the University of Illinois Research Park, on an internship set up through the PI4 program. His two main goals were to segment customers based on electricity and gas consumption patterns and identify customers that could benefit from energy efficiency offerings. In addition, he also helped out on a project that was analyzing the reliability of Ameren's electrical grid. During this internship, Nicholas gained experience using Python to look at large data sets.

During the Summer of 2015, Andrew McConvey worked at Akuna Capital, a proprietary trading firm based in Chicago. He undertook a month-long crash course in options theory and basic trading strategies, and then got to apply this theory in a simulated environment before trading actual products in the market. Throughout the summer, Andrew also worked with a small team of other interns to develop a model for the volatility of option markets and identify trading opportunities. This project deepened his knowledge of Python and introduced him to several packages for data analysis. Andrew met representatives from Akuna Capital at the Mathematics Corporate Day last October, and was offered the internship in January.

During her internship at Dow AgroSciences in Summer 2015, which was arranged by the PI4 program, Tara Negron Santiago worked on the problem of weed resistance resulting from repeated use of a herbicide. Weeds evolve and build up resistance to specific herbicides such as glyphosate (Roundup) through repeated use. Academic herbicide resistance modeling tools have been developed to integrate knowledge about weed biology (species specific), genetics, weed management (herbicide dose response, herbicide rotation), and a variety of other interacting factors such as seed bank density, cropping system, and the initial resistant allele frequency in large weed populations. Tara developed a graphical user interface and post-processing routines in VBA coupled with Matlab, to enable weed scientists and field managers to easily use and interpret results from an existing but rather complicated weed resistance research tool. Additional functionality included the ability to explore various Best Management Practices to try to avoid future issues analogous to what is observed today with glyphosate resistant weeds.
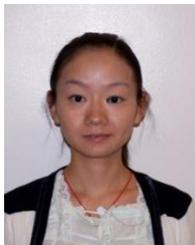
This summer Wei Qin worked as a Back End Developer intern at Agrible. This data science company at the UI Research Park provides business-critical information to select the agricultural communities. The goal for the internship was to learn the agricultural models and implement them in Python. These models typically involve second order partial differential equations which cannot be solved explicitly, and so numerical methods are essential. Thanks to her prior coursework and research in the Mathematics department, Wei was able to construct a finite volume method to solve the models, yielding solutions that are stable and consistent. In addition, she employed statistical methods to analyze and calibrate the model results, using the statistical package R to analyze the sensitivity of the model with respect to parameters. Wei says "the ability to learn new things fast will always be important, no matter which jobs we are doing!"

As part of the PI4 program, Vanessa Rivera-Quiñones interned at the lab of Professor Carla Cáceres in the Department of Animal Biology. Her work focused on the zooplankter Daphnia dentifera, commonly known as the "water flea", which experiences epidemics of the fungus Metschnikowia bicuspidata. To integrate the three roles of Daphnia as consumers, competitors and hosts, Vanessa studied an ordinary differential equations system and analyzed its dynamics. Specifically, she focused on the case were two strains are present, the wild host and the invader. She used techniques from the theory of adaptive dynamics to describe the long term evolution of the population and determined conditions that allow polymorphisms to occur in the system.

Emily Schlafly spent the summer of 2015 working on tinnitus detection and diagnosis with a team in the Speech and Hearing group on campus. Although tinnitus is relatively common, there is no physiological identifier with which to classify and diagnose the condition, which makes treatment difficult. The Speech and Hearing group aims to identify biomarkers associated with tinnitus and hence contribute to the development of a diagnostic tool. Emily and fellow Mathematics student Benjamin Wright worked on developing a method that will establish the activation sequence of given brain regions using functional MRI time series data, so that functional neural connectivity can be compared in sufferers versus non-sufferers. The algorithm was coded in Matlab and its performance analyzed with noisy data to give context to the results. Emily enjoyed working with Professor Fatima Husain, and reports that she was enthusiastic about working with the students from Mathematics, clear about what she was looking for, and fantastic at giving feedback.

Sean Shahkarami worked at Argonne National Lab this summer on developing an implementation of the WASH123d watershed model on top of the PETSc framework. The goal of the project was to develop flexible and scalable tools to help other users define interconnected systems of rivers and ponds, as well as groundwater and atmospheric sources, and then simulate their description as a system of interconnected differential-algebraic equations. Most of his time was split between writing tools in Python, doing numerical work in C on top of PETSc, experimenting with numerical methods for hyperbolic partial differential equations and doing some fluid mechanics. Overall, he enjoyed working on a variety of mathematical and physical problems and being put in charge of a project. He even had a chance to contribute a simple network monitoring tool to the PETSc library, which was then used during some of the testing.

The Wolfram Language (Mathematica) includes "Entities", which carry general types of information. Albert Tamazyan worked at Wolfram Research during the summer of 2015, adding Entities for new families of mathematical functions. The relevant properties include "integral representations", "functional equations", "asymptotic expansions" and many others. During his internship he collected formulas and theorems from the literature, studying the underlying functions and deriving new formulas. Some of these functions are related to his PhD topic, and so the summer work contributed also to his research. Along with a wonderful work experience in the "Scientific Content" department at Wolfram, Albert enjoyed many valuable conversations with experts on mathematics and Mathematica.

During his internship at the Energy Biosciences Institute (which is funded by BP and housed in the Institute for Genomic Biology), Goran Tomic worked with fellow PI4 intern Elizabeth Fields and investigated various data assimilation packages, under the supervision of Dr. David LeBauer. Data assimilation is one step of a larger process that includes model calibration, validation, and application. The interns worked on the Predictive Ecosystem Analyzer (PEcAn) project, which has two components: a scientific workflow and a Bayesian data assimilation system. They contributed to the second component by implementing a particle filter algorithm to be used in the data assimilation step. The algorithm was implemented based on an academic article using particle filters for parameter estimation in weather forecasting. The algorithm was tested on a real dataset in which biomass observations were used to improve parameter estimates.

Bolor Turmunkh spent the summer of 2015 as a Data Science intern at the text analytics startup company Idibon, in San Francisco. Idibon provides cloud-based natural language processing services that enable organizations to efficiently structure and derive business insights from their language data. During this internship, Bolor worked with millions of twitter data, used a crowd-sourced annotation process to give structure to the data, and built predictive models. Bolor improved her skills in Python and R programming languages, implemented Machine Learning algorithms, and generally learned a great deal about good practices in Data Science. She found this opportunity independently, through an extensive online search.

Nuoya Wang held a summer 2015 internship at Dow AgroSciences at the University of Illinois Research Park. Nuoya's main task in the summer project was to build a predictive model for selecting good corn seeds using genome data. She learned a lot in this project, especially about data transformation, predictive modeling and validating the model. Nuoya got the internship through an on-campus career fair.

Xiao Wang worked at Dow AgroSciences on a summer internship, arranged through the PI4 program. The work focused on a deterministic equivalent modeling method (DEMM), which allows for direct incorporation of uncertainty in model predictions, based explicitly upon uncertainty in inputs. This project was initially performed by the company over a decade ago and needed to be updated using the most recent version of the mathematical software package Maple. The initial several weeks of the project dealt with a literature review covering the existing applied mathematics techniques and current approaches being proposed. The company's existing Maple libraries were refined and novel functions (e.g., arbitrary polynomial chaos expansions) were added. The project culminated in the use of DEMM with a model which estimated the evaporation of a water droplet under variable meteorological conditions, and the associated uncertainty in the solution. Through the internship, Xiao learned how to program in Maple, performed a literature survey of relevant applied mathematics papers, incorporated uncertainty into mechanistic and black-box models using traditional Monte Carlo techniques and DEMM, and was exposed to a mechanistic model describing droplet evaporation. She became a mathematical scientist!

Argen West: This past summer, Argen West was a student researcher at John Deere Technology and Innovation Center, in the UI Research Park. He worked on incorporating site data to improve modeling of crop yield, and on evaluating planter performance using statistical methods. While working on these projects, he picked up new software skills (Tableau for data visualization and ArcGIS for geographic and spatial data) and gained experience working with large data sets.
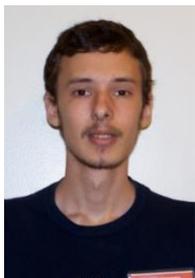
Benjamin Wright
See joint project with Emily Schlafly, above.
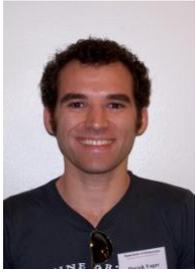
Derrek Yager: Derrek Yager worked as graduate coordinator for the Valparaiso "Experience in Research by Undergraduate Mathematicians" program over the summer of 2015. His duties included liaising between the undergraduate students and the faculty supervisors, and preparing a series of presentations on how to create papers, talks, and posters using LaTeX. He also gave a seminar talk sharing his own research. He aided the Combinatorics group as a research mentor, and arranged trips, lunches, and activities to keep the summer experience lively.

Bingji Yi: At Deutsche Bank during Summer 2015, Bingji Yi found a method to extrapolate the Parametric Implied Volatility (PIV) surface to keep it arbitrage-free and twice differentiable. This surface is a fundamental building block in quantitative finance, for it parameterizes the Black-Scholes implied volatility as a function of strike and expiration; hence modelers can still use Black-Scholes formulas for pricing and hedging efficiently. The problem with PIV is the lack of any guarantee that it will lead to arbitrage-free option prices. In practice, arbitrage is usually found in PIV with extreme strikes, which leads to the idea of extrapolating the PIV outside a certain strike range. During this internship, Bingji developed a new method of extrapolation, and proved the method will work for any (arbitrage-free) market data. He also found how to select optimal model parameters, using numerical algorithms.

2014

Stephen Berning held a Summer 2014 position at the John Deere Technology and Innovation Center in the University of Illinois Research Park as an Algorithm Scientist. Stephen learned about this opportunity at the 2013 Mathematics Corporate Forum. His main duty was working on a variation of the Vehicle Routing Problem for Deere customers. This included the mathematical development of the problem and the algorithm development and implementation for data extraction, for data analysis and for the problem itself. His duties also included aiding in the development of mathematical solutions to various problems being tackled by Deere.

Stacey Butler worked with James O'Dwyer in the Plant Biology Department on modeling ecological systems. This internship was set up through the PI4 program. She worked on models of mutualistic networks, such as plant-pollinator networks.
Using the Python programming language Stacey wrote code to implement the models. Finding equilibrium states for the network involved finding fixed points of a system of ODE's. Different parameter values were tested to determine the effect on the network structure. Through the course of the internship, she learned about theoretical ecology and programming.

Jed Chou spent the summer of 2014 as an intern at Personify, a startup in Champaign developing an image-segmentation app for video conferences. His primary duties were to understand the theoretical aspects of the algorithms implemented in the software and to try and optimize some parts of the software. To this extent, he and his partner wrote some code in Matlab, ran experiments on a large number of images, and reported the results to one of the developers at Personfiy. Jed found this internship through the NSF-funded program PI4.

Erin Compaan held an internship at the Department of Defense in the summer of 2014, granted as part of the National Physical Sciences Consortium fellowship program. She worked on a network behavioral analytics problem using machine learning techniques. There was extensive collaboration with researchers, analysts, and other interns. Much of her time was spent programming in Python and R, which proved a challenge because of her lack of training. However, the experience was excellent, and she found her supervisors valued perseverance and willingness to learn as much as prior knowledge. In addition to programming and modelling skills, she took away a new appreciation for the challenges of working with real (and messy!) data to meet clients' often unclear goals.

Matthew Ellis held a summer 2014 internship at the Institute of Genomic Biology (IGB) at the University of Illinois, working as part of a group from the Energy Biosciences Institute that focuses on modeling biofuel crops, such as sugarcane and agave. He found this internship through the PI4 program, which connected him with the research group after a "Computational Bootcamp," where he learned Python and several applications, including image processing and triangulation. Working with another Math graduate student, Nouya Wang, he developed a model for an agave plant, and wrote a ray-tracing algorithm to calculate the amount of light an agave plant receives, which ultimately determines the plant's rate of photosynthesis. Through his work, Matthew learned about 3D imaging, Bayesian data analysis, and ray-tracing.

Elizabeth Field had an internship in the biology lab of Professor Karen Sears, during the summer of 2014, through the Program for Interdisciplinary and Industrial Internships at Illinois (PI4). She worked alongside fellow student Daniel Carmody and was also advised by Prof. Zoi Rapti, both of whom are also in the mathematics department. Her project focused on using a reaction-diffusion system of partial differential equations to model the limb development of tetrapods, such as mice and bats. The model was then used to conduct a sensitivity analysis to determine which parameters are most sensitive to small changes. The results of this project will be used by biologists studying the morphological development of particular tetrapods to help them better understand the various biological factors influencing the morphological development of such species.

Benjamin Fulan held an internship in Prof. Andrew Ferguson's lab in the UIUC Materials Science Department during the summer of 2014. He was matched with the internship as part of the Mathematics Department's PI4 program. The primary goal of the project was to develop an algorithm to aid in the creation of a vaccine for HIV, a process complicated by the virus's ability to defend itself via mutation. Given data describing the virus's replicative fitness as a function of its amino acid sequence, the algorithm is designed to find the set of amino acids which will most impair the virus when targeted by the immune system. As a result of the internship, Ben gained an increased knowledge of both mathematical optimization techniques (such as genetic algorithms) and computer programming.

Meghan Galiardi held an internship at Sandia National Laboratories during the summer of 2014, in Albuquerque. She connected with Sandia after meeting their representatives on their regular recruiting visits to the Mathematics Department. The representatives passed her CV on to the department at Sandia that eventually offered her an internship. Meghan worked with the complex systems group doing population health modeling. Most of her time was spent programming models in Matlab. Meghan learned the complexities that go into a model when it needs to be delivered to a customer. She advises anyone who is considering an internship at Sandia to meet with representatives and learn about opportunities at the Lab. Sandia has a very welcoming environment and is a great place to work. Meg also thanks the PI4 grant for travel support.

Lisa Hickok held an internship with the Department of Defense during the summer of 2014. She first found out about the internship possibilities when a visitor from the DoD came to the department to give a talk. After applying online and going through a formal background check and several interviews, Lisa was offered the summer position. During the internship, Lisa created a recommender system for linking users to the tools that would most likely benefit them, by using latent factors uncovered through statistical methods and linear algebra techniques. She gained experience developing algorithms in Python and using Hadoop Map/Reduce. Lisa found previous programming experience very valuable. Lisa recommends looking early for an internship with the DoD, as the application and security clearance process can take 6-8 months.

Nicholas Kosar spent the summer of 2014 as an intern in Dr. Alejandro Dominguez-Garcia's research group at the University of Illinois. Nicholas got this internship through the department's Program for Interdisciplinary and Industrial Internships at Illinois (PI4). His work involved writing a computer program to study a feedback system for regulating frequency in islanded microgrids. Nicholas then used this program to collect data and make a conjecture about how to minimize the time it takes for the frequency to be regulated. Through this internship, Nicholas learned about the goals in power systems and how these are achieved in practice. He also gained valuable programming experience.

Lanius   Fulan   Terry

Melinda Lanius, Benjamin Fulan, Alonza Terry; along with Candice Lanius (Rensselaer Polytechnic Institute) participated as a team in the non-MBA Case Competition hosted by Cornell and Rockefeller Universities (http://cr-casecomp.org), in August 2014. Sixty teams applied to participate, and our team was among the top thirty chosen for the final presentation day in New York City. The distinctive feature of this case study event was designed to "foster the next generation of consultants with advanced degrees" and so it focused on students from PhD programs (rather than MBA programs). Case competitions help students hone their analytical and presentation skills while working on consultant-type problems, which can range from implementing a new type of technology to suggesting potential business acquisitions and mergers. The Illinois team worked for a week to develop a 3-5 year plan for Yahoo!, analyzing Yahoo CEO Marissa Mayer's current strategy for the company and suggesting how to deploy the $10 billion windfall Yahoo stands to gain from an initial public offering of one of its subsidiaries (the e-commerce company Alibaba). Then the team traveled to New York for a day of presentations and networking, with travel support from the PI4 grant. Interesting opportunities came out of the trip, including the chance to apply for "mini-internships" next summer, which are 1 week job-shadowing programs hosted by some of the major consulting companies.

During the summer of 2014, Andrew McConvey held an internship at Cantor Fitzgerald & Co., a brokerage firm in New York. He joined a team of quantitative analysts and software developers at Cantor who were in the process of developing an electronic trading platform. His responsibilities included implementing models to evaluate cost and risk of prospective trades, gathering and processing necessary data, and using these models to implement a portfolio optimizer. As part of the internship, he gained valuable industry experience and improved his knowledge of languages such as Python, SQL, and Java. Andrew was first introduced to the team in early 2014 while on a networking trip to New York. After interviewing for the position in the firm's internship program and a second meeting with the team, he was offered the summer job.

This summer Daniel McDonald worked as the Special Functions Intern at Wolfram Research in historic Champaign, after learning about the position at the Research Park Career Fair in the spring and going through a quick interview process. He says the various career fairs and engineering expos throughout the year seem quite useful in finding both temporary and permanent employment. At Wolfram, his main job function was to read through mathematical literature and verify and input interesting formulas involving special functions for use in the Wolfram Functions site, as well as for possible Mathematica implementation. While working with partial Bell polynomials and their generalizations, which are used in computing arbitrary derivatives of compositions of functions, he devised a novel functional recurrence that suggested a quick recursive algorithm for computing generalized Bell polynomials. This algorithm ran much faster than Mathematica's existing method, and so the new algorithm was implemented into the newly-released Mathematica version 10.0.1.
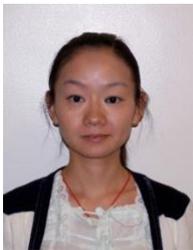
Wei Qin held a summer 2014 internship at Akuna Capital, which is a trading firm founded in Chicago. Her work was to build a quantitative model that will be used in trading. The first task she was assigned was to use cubic splines to fit discrete data so that the fitted curve is smooth and preserves some good properties of the data. Later, she used the component analysis method to analyze the spline result so that the spline nodes could satisfy the requirements of the trading process. An important skill Wei learned during the internship was how to compile the code in a team, since all team members shared the same code. As part of her work, she became more proficient in Python.

Han Wang held a summer 2014 internship at John Deere Technology Innovation Center, where he worked as a data and GIS analyst. He thanks the PI4 faculty for arranging the internship. His primary duty was to analyze the crop yield data collected from the sensors of harvesting combines. He and his teammates developed several signal processing strategies to create the yield map from raw data. As a benefit of this internship, Han gained a new understanding of precision agriculture and the modern practices in agricultural industry. He also learned about robotics and sensors used at John Deere for various purposes. He mainly used R, Matlab and Python as the working programming languages.

Nuoya Wang held a summer 2014 internship at Institute of Genomic Engineering at the University of Illinois, Urbana-Champaign. Nuoya's main task in the summer project was to develop a 3D canopy model and analyze the photosynthesis effect for plants. By analyzing 3D triangular meshed data, Nuoya developed an algorithm to extract plant features, such as canopy height, leaf length, leaf width and leaf curvature from given format files. With this information, statistical strategies can be used to generate new plant models. From this internship Nuoya gained a better understanding of applied geometry and graph theory. She also learnt a lot about data cleaning, data analysis and statistical methods. Nuoya learned about this internship from Professor Yuliy Baryshnikov, associated with the PI4 Program.

During the summer of 2014, Anna Weigandt interned with Personify, at the UI Research Park. Anna studied an image segmentation algorithm called GrabCut. This algorithm analyzes an image and separates the pixels into foreground and background components. GrabCut works by setting up an energy minimization problem. Pixels of very different colors are easier to separate into different components than pixels with similar color data. Anna worked on the 'finger' problem. When segmenting images of people, often fingers are incorrectly marked as background pixels. She implemented various modifications to the base algorithm in Matlab, to attempt to address this issue. Anna connected with Personify through the PI4 program.
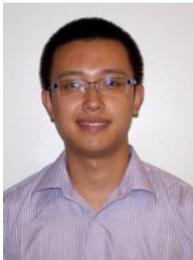
Sishen Zhou learned about this opportunity at John Deere (UI Research Park) from the Mathematics Department grad-careers mailing list. The position was described as a full-time contract job at the beginning. Sishen contacted the manager, showed his interest and explained that he is not going to graduate soon. After a discussion, the manager turned the position into a summer 2013 internship. The main topic of the work is to create improved maps of yield distribution by analyzing data sets collected by several sensors on a combine harvester. As a result of the internship, Sishen developed his skills in signal processing, image processing and optimization, and got plenty of experience in handling large, real data sets.
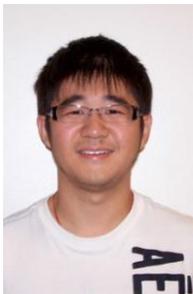
2013

Chris Bonnell, a 2013 graduate of our PhD program, is now an Associate Predictive Modeler at Allstate in Evanston, Illinois. Chris held a 2012 summer internship with Travelers. The internship focused on (1) the building and testing of predictive models built on data, and (2) the design of tools to build new and better models. Chris originally found this internship through networking with graduate school contacts.

Susannah Johnson worked at Sandia National Lab during the summer of 2013. She first connected with representatives from Sandia when they visited the Mathematics Department in December 2012. One of the visitors was looking for an intern with knowledge of theorem provers, one of the primary tools in Susannah's research. After some further conversation and a formal background review, Susannah was offered the internship. At Sandia, Susannah worked with others to develop a theorem prover related to computer security. She learned about computer security, a new area for her, and developed her skills in creating mathematical models. Susannah says that her work was mostly theoretical in nature, while some of the other interns were working in very applied areas of mathematics at Sandia.

During the summer of 2013, Ziying (Cedar) Pan interned at the State Farm Research & Development Center in the University of Illinois Research Park. Cedar found this opportunity through an on-campus career fair. During the internship, he worked with four other interns to develop a prototype of an onboard device for driving risk evaluation. His role in development focused on data clustering component in the system. He designed project's SQL database, implemented its SQL interface, and also implemented a machine learning algorithm for data clustering. As a result of his experience with State Farm, Cedar gained knowledge of databases, of software development in C++ and of machine learning.

Austin Rochford held a summer 2013 internship at Monetate, an e-commerce analytics contractor. One of his primary duties was to find ways to improve the accuracy of the statistical models used for evaluating the success or failure of experiments run on e-commerce websites. There were several obvious points were the assumptions of the model did not correspond to reality. Austin developed strategies for countering assumptions of the model which did not correspond to reality and thereby increased the accuracy of the reports by approximately 40%. He also collaborated with user interface and user experience designers to develop an interface for reporting the outcomes of experiments to marketers. From his internship, Austin gained an understanding of the methodology of professional software development and large-scale data analysis. Austin learned about the Monetate internship by doing an online search for data-oriented software developer internships in the Philadelphia area. He also recommends Hacker News , a software developer news/discussion site that has monthly "Who's Hiring" posts. Austin's coding experience, including publicly available projects (on GitHub, for instance), was valuable in landing this position.
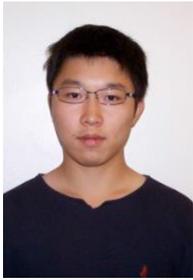
Rui Song held a 2013 summer internship at the insurance company Plymouth Rock. His primary task was to determine the variables which are significant in predicting some key response variables, such as loss ratio, pure premium, counts of claims. He also did theoretical work such as comparing different predictive models in several aspects. As a result of this internship, Rui gained an understanding of the gap between real data with lots of irregular input, and ideal data, and learned about normalizing real data. Rui learned about the Plymouth Rock internship through a personal connection, and was offered the position after interviewing.

Yat Sen Wong spent the summer of 2013 as an intern at Neustar, a real-time cloud-based information services and analytics company. Yat Sen learned about this internship opportunity through the Engineering Career Fair and the Research Park Career Fair, both at the University of Illinois. His main duty was to implement scalable machine learning algorithms. This means not only building a mathematical model, but also ensuring that the model can give a satisfactory running time performance. when dealing with large data sets. As part of his work, he learned several programming languages including Hadoop, Hive, Python, Java and some elementary shell script.