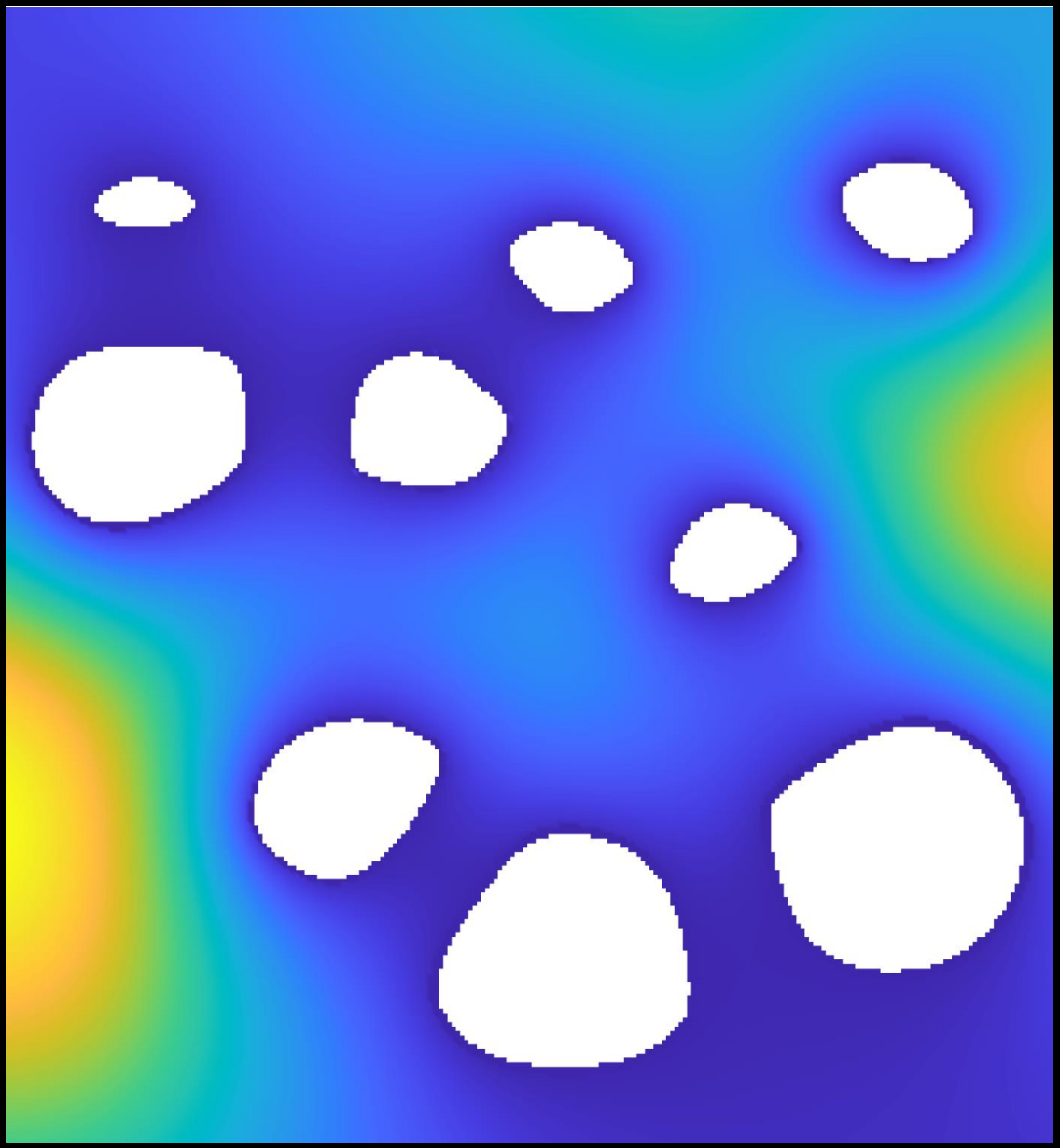
April 14th & 15th



2022 Computational Exposition

Welcome to Computational Exposition 2022, a yearly event where students in the Department of Scientific Computing (DSC) showcase the results of their research in the last year. Higher education was confronted with singularly difficult circumstances this year due to a deadly pandemic, with most students and professors working from home, attending classes online, and having to conduct research and communicate remotely. This research covers a broad spectrum of disciplines, but shares a common thread: they concentrate on algorithm development and blend the computational and the mathematical to solve problems in the applied sciences. The innovation displayed is broad and remarkable. Our students make us proud!

The student posters reflect the breadth and depth of the research carried out in the DSC. They are the direct result of a fulfillment of our two most important missions: providing world-class interdisciplinary research and training in computational science.

Our graduate degree programs and the success of our students bolsters our confidence that we are a premier institution for the training of the next generation of computational scientists. Indeed, looking at what our current students have achieved over the past several years attests to this claim!

So, enjoy the presentation, interact with the students, challenge them, learn from them, and reflect on the fruits of their intelligence, skills, dedication and labor, and join us in thanking them for their contributions to the DSC, to FSU, and to science.

Gordon Erlebacher, Chair
Department of Scientific Computing

Front Cover: Diffusion of a point source initial condition in a multiply-connected domain representing nine eroding grains
courtesy of Jake Cherry.
Back Cover: Data graphs for classification and prediction of Alzheimers,
courtesy of Kevin Mueller.

Presenting Researchers

R E S E A R C H

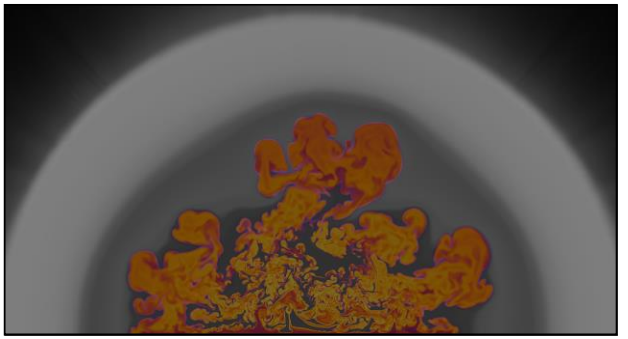
Abstracts & Graphics

Ezra S. Brooker
Ph.D. in Computational Science
Advisor: *Tomasz Plewa*

Turbulence Generated by Thermonuclear Deflagrations in Type Ia Supernovae

The deflagration-to-detonation transition (DDT) mechanism remains one of the major unsolved problems of combustion physics. Astrophysicists have suspected for almost 40 years that it is also directly responsible for a subclass of white dwarf (WD) explosions powering Type Ia supernovae (SN Ia). Much of the research on DDT in SN Ia has focused on the interactions of deflagration fronts with turbulence generated by the flame itself. In our previous studies, we describe a potential DDT mechanism observed in so-called turbulent combustion box models. However, is the turbulence used in these models relevant to conditions in white dwarf stars?

In prior work, we constructed two- and three-dimensional (2D/3D) models of turbulent combustion boxes (TCBs) of white dwarf plasma that experience DDT-like events with nuclear burning, but do not model the actual flame. Using the initial conditions from our TCBs, we further constructed 2D deflagration boxes (DFBs) that model the Rayleigh-Taylor unstable flame itself. Analysis of the turbulence generated in these DFBs bears similarities to our DDT-producing TCBs. We are actively constructing 2D high fidelity models of deflagrating white dwarf (half) stars using realistic progenitor models, gravitational potential field relaxation, and advanced flame front tracking and energetics. Our next steps are to analyze the turbulence that generates from the buoyant flame in these supernova models.



Figure, above: Snapshot of half of an exploding white dwarf star at 800 ms after explosion begins.

Jesse Cherry
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

A time-independent solver for diffusive processes in exterior domains

To describe transport of quantities such as heat or chemical concentrations, the diffusion equation must be solved in complex unbounded geometries. We propose a numerical method recasts the time-dependent heat equation into an elliptic PDE by applying the Laplace transform. By using the Laplace transform, the solution of the PDE can be found at any time without the need to time step. The transformed elliptic PDE is solved with high-order integral equation methods that accurately capture the far-field conditions. After solving the elliptic PDE, an inverse Laplace transform must be performed. This is done by carefully choosing a Bromwich contour that results in an integrand that is non-oscillatory and rapidly decays to zero. Combining these techniques, high-order accuracy in both space and time are achieved. Additionally, we extend the method to compute the flux through the boundary of the domain at any time, and compare our method with a Monte Carlo approach.

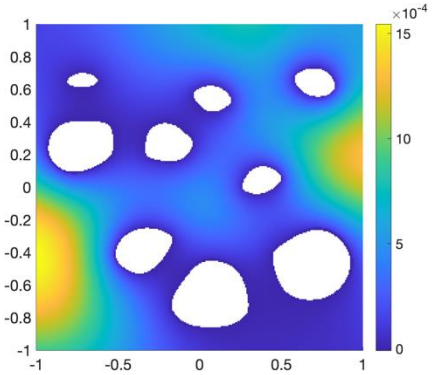


Figure: Diffusion of a point source initial condition in a multiply-connected domain representing nine eroding grains. Solution computed only at time t = 0.25 without timestepping.

Yu-Chieh Chi
Ph.D. in Computational Science
Advisor: *Chen Huang*

High-quality Local Pseudopotential for Metals

Orbital-free density functional theory (OF-DFT) is a very promising tool for large-scale metal simulations. However, the accuracy of OF-DFT is determined by the accuracy of the local pseudopotentials (LPSs) and the approximate kinetic energy density functionals (KEDFs). It is very challenging to build accuracy LPSs and KEDFs for transition metals, due to the localized d electrons. Recently, reliable KEDFs for simple and transition metals have been developed, and the remaining bottleneck is the lack of accurate LPSs for transition metals. In this work, we build high-quality LPSs by approximately reproducing the atomic eigenvalues and atomic orbitals (beyond cutoff radii). For magnetic system, Fe, Co, and Ni, addition constraints on the energy difference between high-spin and non-spin states of these atoms are added. Technically, the LPSs are expanded using the Legendre polynomials, and the expanding coefficients are optimized to achieve such matching. We have built the LPSs for both main group and transition metals and tested them with the Kohn-sham DFT (KS-DFT) calculations of the lattice constants, bulk moduli, and energies of various phases of metals. The LPSs results agree well with the results from the project augmented wave (PAW) method. Our new LPSs open the door for future large-scale, accurate OF-DFT simulations of metals alloys.

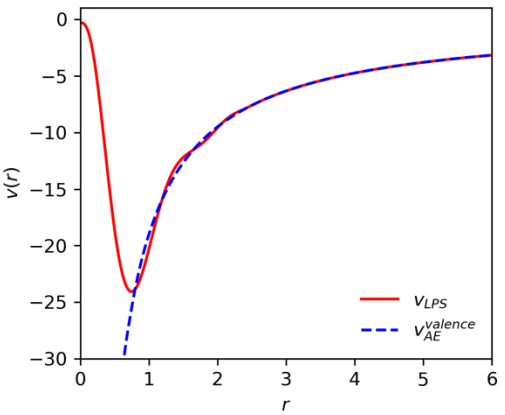
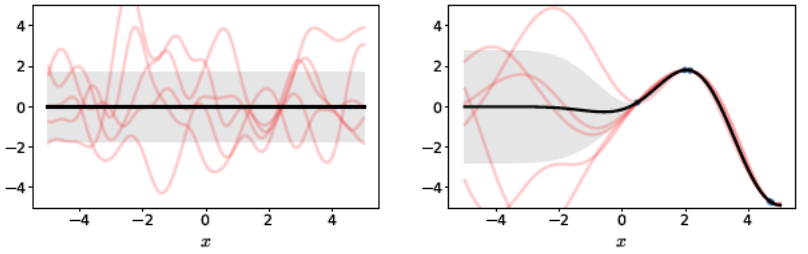


Figure Left: $vAEvalence$ and $vLPS$ for Silver. The valence electrons include silver $4s$, $4p$, $4d$, $5s$, and $5p$ electrons. The cutoff radius is $rcut$=2.5 bohr.

Pankaj Chouhan
Ph.D. in Computational Science
Advisor: *Sachin Shanbhag*

Surrogate Modeling with Gaussian Processes for an Inverse Problem in Polymer Dynamics

In the polymer industry, molecular simulations are used to optimize product quality, increase revenue, and improve manufacturing processes. Molecular simulation models of polymer rheology take the molecular structure of a polymer mixture as input and yield the linear or nonlinear viscoelasticity of the mixture as output. Interestingly, due to the extreme sensitivity of rheology to molecular structure, these models can alternatively be used inversely to characterize the molecular structure, given polymer rheology and molecular model. Different molecular accuracy and speed. For some applications the computational cost of molecular simulations can be prohibitive; this is where surrogate models can help. Surrogate models mimic the input-output behavior of the underlying molecular simulation to reasonable precision at a fraction of the computational cost. This work explores surrogate models based on Gaussian processes for solving inverse problems in polymer dynamics. I will discuss an approach that uses a separable kernel to build a surrogate model. I will also outline the approaches based on Karhunen-Loeve expansion.



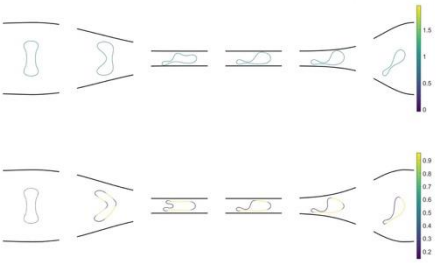
Figures above: Solid black line depicts the mean and the shaded region shows an interval of ±σ around this mean. The red lines depict 5 samples drawn from this Gaussian process. In (a) the kernel is given by the RBF with σ2 = 3 and l = 0.5. In (b), the blue circles denote 5 observations. The optimal hyper-parameters after maximizing the marginal likelihood of observing the data are σ^2 = 7.73 and ^l = 1.57. In posterior distribution using these MLE hyper-parameters and observed data are shown in subfigure (b).

Ashley Gannon
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

Multicomponent vesicles under confinement

The biophysical lipid bilayer membrane (such as a vesicle) often consists of multiple species of macromolecules such as cholesterol, surface proteins, surfactants, and different lipids. I use the phase-field model for a multicomponent lipid membrane and implement the hydrodynamics with a boundary integral code. The simulations covered in this poster show that the multicomponent structure results in fundamentally different behaviors than a single component vesicle.

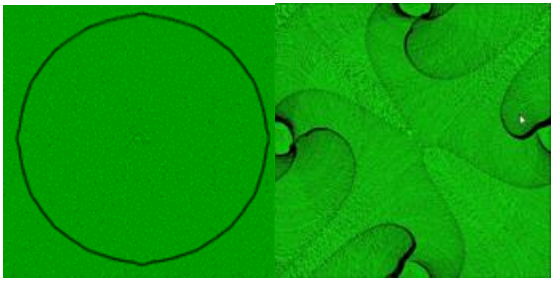
Figure, below: Comparing the shapes of a single-component vesicle and a multi-component vesicle moving through a constriction. Here, the reduced area is 0.4 and the background flow velocity is 5 μm/s.



Jhameika Greenwood
Masters in Computational Science
Advisor: *Bryan Quaife*

A GPU-Accelerated Hydrodynamics Solver For Atmosphere-Fire Interaction

A fundamental process to understand fire spread is the atmospheric flow. Building computational tools to simulate this complex flow has several challenges including boundary layer effects, resolving vegetation and the forest canopies, conserving fluid mass, and incorporating fire-induced flows. We develop a two-dimensional hydrodynamic solver that models fire-induced flow as a convective sink that converts the two-dimensional horizontal flow into a vertical flow through the buoyant plume. The resulting equations are the two-dimensional Navier-Stokes equations, but with point-source delta functions appearing in the conservation of mass equation. We develop a projection method to solve these equations and implement them on a GPU architecture. The ultimate goal is to simulate wildfire spread faster than real-time, and with the ability for users to introduce real-time updates in an augmented reality sandbox.



Marzieh (Tara) Khodaei
Ph.D. in Computational Science
Advisor: *Peter Beerli*

PATHTREES: A Python package to explore the tree landscape

Phylogenetic trees are fundamental tools for understanding the evolutionary history of a set of species. Understanding the local neighborhoods of a phylogenetic tree is essential, but since trees are high-dimensional objects, discussing these neighborhoods is difficult. We use different distance methods to explore the phylogenetic tree landscape. Based on the geodesic distance between pairs of trees, we developed a method to generate trees on the shortest path between two arbitrary trees. We are convinced that these paths among trees will improve searching the tree space and therefore find the best trees faster than current methods.

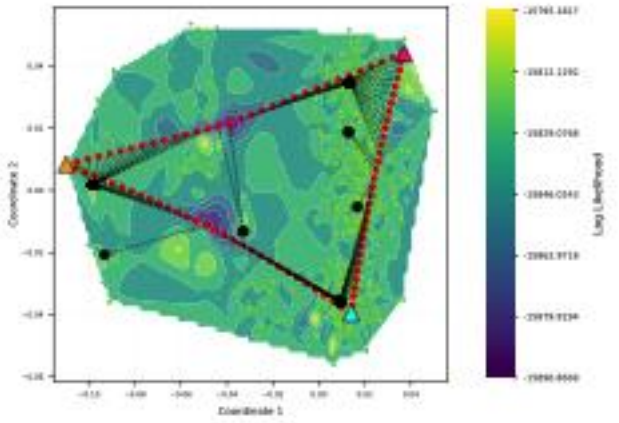
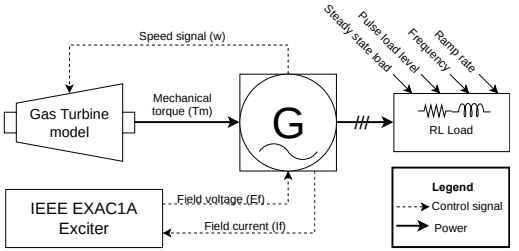


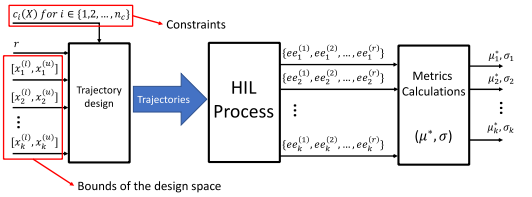
Figure 1: An example of pathtrees (red dots) along the shortest path (geodesic) between three arbitrary trees (colored triangles) in the tree space and their corresponding optimized trees (middle-size black dots). Each pathtree and the corresponding optimized tree are connected with a dashed line. We applied multidimensional scaling (MDS) and cubic spline interpolation to visualize a space of 1000 trees and 90 pathtrees. The log- likelihood was used for the contour color of the surface. Each dot is a tree; the lighter the dot, the higher the likelihood of the tree.

Behshad Mohebali
Ph.D. in Computational Science
Advisor: *Gordon Erlebacher*

Real-time Hardware-in-the-Loop

Real-time Hardware-in-the-Loop (HIL) simulation has been a major step in design, development, and implementation of new technologies in the field of power systems. In many cases, the models of physical systems are elaborate with tens or even hundreds of input parameters and output quantities. Knowing which inputs have the most influence on the output quantities of interest (through factor screening) will help the analysts simplify these models and reduce the computational cost of the evaluation process considerably. The assumption made by common factor screening methods is that the range of variation of each factor does not depend on the value of the other factors. In this work, we try to eliminate this assumption and consider the case where the normalized parameter space is not a perfect hypercube. The presented method is a variation of the Elementary effects (Morris' method 1991) that uses multiple degrees of freedom in the original method to take the space shape into account.





Kevin Mueller
Ph.D. in Computational Science
Advisor: *Anke Meyer-Baese*

Learning data-driven graphs for the classification and prediction of Alzheimer's disease

Numerous deep learning approaches have been proposed to automatically classify Alzheimer's disease (AD) from medical images. However, common approaches such as traditional convolutional neural networks (CNNs), lack interpretability and are prone to overfitting when trained on small datasets. As an alternative, significantly less work has explored applying deep learning approaches to region-based features that are commonly attained from atlas partitions of known regions of interest (ROI). In this work, we combine CNNs with a graph neural network (GNN) to jointly learn a graph of connectivity's between ROIs as a prior for learning meaningful features for AD prediction. We apply our method to the ADNI dataset and show that the edge probabilities alone are sufficient to reach high classification accuracy. In addition, we show how the edge edge probabilities haves built in interpretability properties by visualizing the feature importance for all edges.

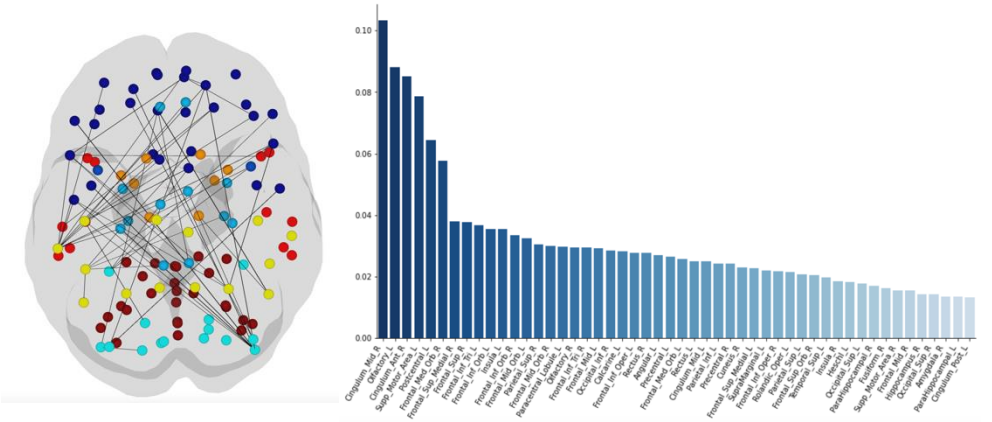


Figure 1, Left: The feature importance for all 6670 edges in the frontal (dark blue), limbic (light blue), occipital (turquoise), parietal (yellow), spinal cord grey matter (orange), temporal (red), and cerebellum (dark red) regions.
Figure 2, Right: Top 50 weighted feature importance node degrees sorted by importance.

Dorianis Perez
Ph.D. in Computational Science
Advisor: *Kevin Speer*

The Development of a Lagrangian Cloud Physics Package in HiGrad for the Simulation of Pyrocumulonimbus (PyroCb) Clouds

In recent years, wildfires of large scale have taken center stage in national media. As these fires continue to intensify each year, we grow desperate to understand the underlying physical phenomena that explain them, highlighting the importance of accurate and comprehensive wildfire models and simulations. These wildfires often produce large thunderstorm clouds, called pyrocumulonimbus clouds (pyroCb), that exacerbate localized weather and cause more damage/loss in the affected communities. PyroCbs are produced when intensely heated air rises due to buoyancy from heat and moisture being released by the fire and condense after saturation due to adiabatic cooling. Aerosol chemistry and physics play a critical role in the formation of pyroCb clouds. Thus, the development of a wildfire model with the capability of including atmospheric aerosol release, cloud microphysical processes, and particulate fallout and/or injection into the stratosphere is essential. 2021The current research aims to build a comprehensive microphysics package for accurate representation of pyroCb formation in HiGrad, LANL's parallel atmospheric hydrodynamics model, using a Lagrangian particle-based approach. This method has proven to be beneficial over an Eulerian scheme for cloud microphysics because it allows explicit simulations of interactions between cloud and aerosol particles using super droplets which represent a collection of aerosol particles and have a position in space, velocity, and size. Since this approach calculates transport and growth of particles individually using ordinary differential equations, numerical diffusion, often a problem in Eulerian schemes, is eliminated. A Lagrangian cloud model has been implemented in HiGrad that includes aerosol activation and droplet growth by condensation/evaporation with a tracking algorithm that tracks aerosol particles of interest, i.e. soot particles, and allows the analysis of the activation and size of these. Aerosols have to achieve a critical radius before condensing water. Thus, in a normal distribution of aerosol particles being lofted into the atmosphere from a large wildfire, smaller aerosols will not activate. We present results of simulations including activated aerosol concentration and size distributions and other relevant statistics.

David Robinson
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

QUIC-URB and QUIC-Fire Extension to Complex Terrain: Development of a Terrain-Following Coordinate System

The first steps in an extension of an existing diagnostic wind model, QUIC-URB, to a terrain-following coordinate system for use with an existing fire model, QUIC-FIRE, are outlined. High resolution, 11-23 m, validation studies using wind measurements over two areas of terrain are illustrated. The first study uses data measured at Askervein Hill whose simple geometry and isolation from surrounding terrain helps reduce buoyancy and momentum effects for comparison with the mass-conserving model. A section of highly complex terrain in the Socorro Mountains where 16 sonic anemometers and vertical profiles from two Lidars measured wind velocities in a 2019 field experiment are used for the second case. These validation studies demonstrate both successes and shortcomings of the model. The model shows good agreement with data in areas of open sloped terrain but lacks in areas where flow separation and thermally driven effects may be present.

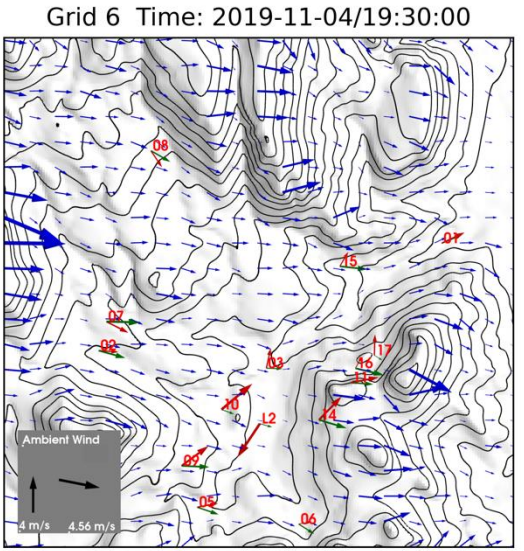
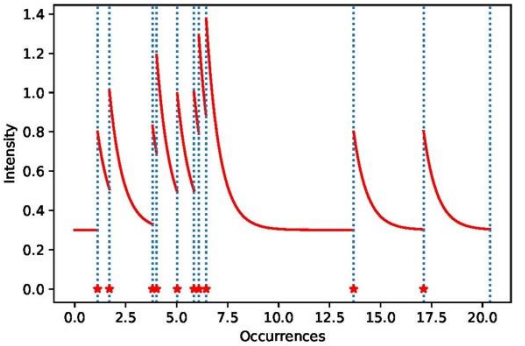


Figure 1, above: Vector plot comparing measurements and simulation results of the terrain-following implementation of QUIC-URB applied to winds in the Soccorro Mountains at 7:30 PM local time. The background wind at 10~m and the scale for the model-computed blue vectors is illustrated in the grey inset. The red vectors represent sensor measurements, labeled by sensor number, and the green vectors are model results interpolated to the sensor locations. Both the red and green vectors are scaled up by a factor of three to make the comparison easier.

Marjan Sadeghi
Ph.D. in Computational Science
Advisor: *Peter Beerli*

A Point Process Approach to Model Coalescent with Recombination

Wuif and Posada (2003) introduced a coalescent model considering recombination hotspots. They considered heterogeneity in recombination rate along the chromosome. They chose the center of recombination hotspots according to some point process and assumed that recombination is happening based on a descending rate from the centers. They used a two-step procedure: using a point process to find the center of the recombination hotspot and then a chosen distribution for the recombination events around this hotspot. Here we propose a new model using Hawkes processes which improves on this; our model locates the recombination events and the hotspots in one step.



Daryn Sagel
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

New Statistical Descriptions of Fire Behavior

When fire propagates through a fuel bed–whether it be a small experimental setup, land managed by prescribed fire, or a vast forested landscape—there is a variety of structures and behaviors that are poorly understood. Two examples are (1) small-scale fire spread through a porous pine straw bed, commonly found in nature, and (2) large-scale interactions between the fire, plume, forest canopy, and atmosphere. I have developed techniques that combine computer vision and graph theory to track such behaviors. Due to the turbulent nature of fire, a statistical approach is used to analyze the data obtained through my method. This analysis reveals behavior of the fire and plume dynamics, heat transport, and fire-atmosphere coupling. I will showcase analysis from an infrared video that captured fire spread in a pine straw plot. Current and future work that considers the fire-atmosphere coupling and heat transport will also be discussed.

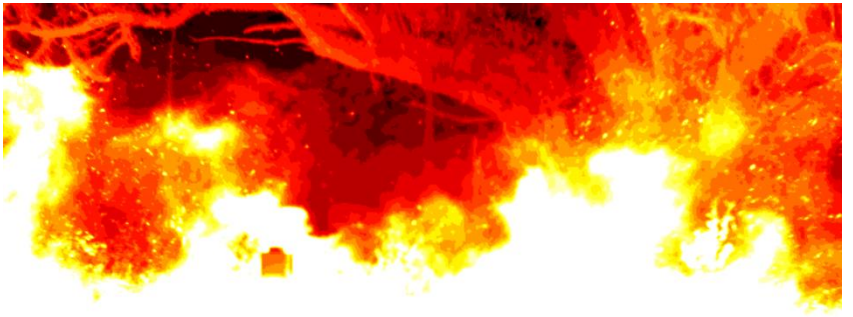


Figure above: Infrared image of a sample fire environment.

Sarthak Sharma
M.S. in Computational Science
Advisor: *Gordon Erlebacher*

Data Visualization and Delay Prediction of Copa Airlines Flight Network

This MS research project aims to: develop a computational tool for visualizing any desired part of the flight network of Copa Airlines; predict flight delays using Machine Learning (e.g., Graph Neural Networks); and if possible, use Artificial Intelligence to automate the flight scheduling process that is currently done manually by Copa Airlines officials. The computational tools used include: NumPy and Pandas for data processing; Plotly, NetworkX, and Matplotlib for data visualization; and TensorFlow and PyTorch for creating, training, and using artificial neural networks. All programming is done in Python language, mostly in Jupyter Notebooks.

An event's delay is defined as the difference between its scheduled and actual times. The event is: late if scheduled time is before actual time (positive delay); early if scheduled time is after actual time (negative delay); and on time if both scheduled and actual times are the same (zero delay). We can have either a classification problem (where we predict the flights' departure and arrival delay categories from six options of minutes delayed: no delay, 1 to 15, 16 to 30, 31 to 45, 46 to 60, 61 or more); or a regression problem (where we predict the minutes of flights' departure and arrival delays). Examining the dataset of past flights (provided by Copa Airlines), we see that, in both departure and arrival, most flights (around 70 % to 80 %) have no delay. Hence, it is an imbalanced dataset.

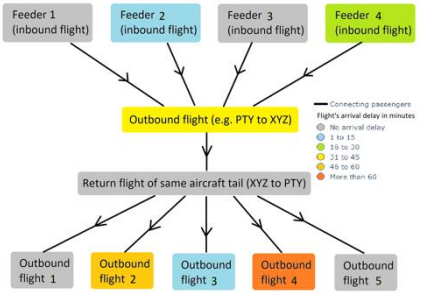


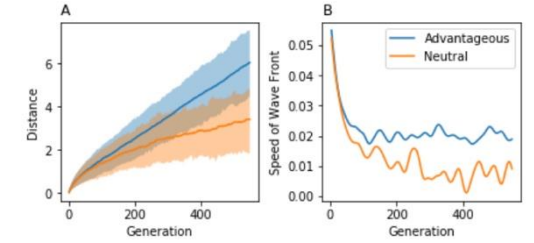
Figure 1: Schematic of Copa Airlines flight network representation used for data visualization. Inbound flights arrive at PTY airport. Outbound flights depart from PTY airport.

Figure 2: Copa Airlines flight network. Almost every flight of Copa Airlines has PTY airport (at Republic of Panama) as either origin or destination.

Kyle Shaw
Ph.D. in Computational Science
Advisor: *Peter Beerli*

Spread of new mutations

The terms population size and population density are often used interchangeably, when in fact they are quite different. When viewed in a spatial landscape, density is defined as the number of individuals within a square unit of distance, while population size is simply the total count of individuals in a population. In panmicitic populations, the effective population size is known to influence the interaction between selection and random drift with selection playing a larger role in large populations while random drift has more influence in smaller populations. However, in a two dimensional spatially defined population, the role of local population density on gene frequencies is still being explored. Using a spatially explicit simulation software we investigate how local population density affects the flow of new mutations through a geographical space; validating with explicit visuals some general mathematical results that advantageous alleles spread at a linear rate, dependent on local population density, selective advantage, and dispersal distance. We also confirm that the rate of spread of neutral alleles decays over time and is independent of local population density. Using population density, selective advantage, and dispersal distribution, we develop mathematical models to predict the speed at which advantageous and neutral alleles travel from their point of origin in a two-dimensional spatial landscape.



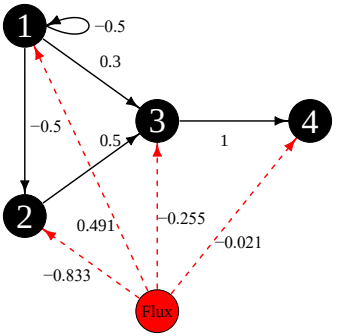
Left: (A) Maximum distance that the new allele is found from the point of origin at each generation. The solid line is the average of 50,000 simulations and the shaded region marks range of 90% of simulations. The advantageous allele has a selective advantage of s = 0.05 and the neutral allele s = 0.0. (B) Speed of the plots on the left. Derivatives were calculated on smoothed data with the centered difference formula.

Philip Solimine
Ph.D. in Computational Science
Advisor: *Anke Meyer-Baese*

Optimal control for networked moments

We study the optimal control of the mean and variance of the network state vector. We develop an algorithm that uses projected gradient descent to optimize the control input placement, subject to constraints on the state that must be achieved at a given time threshold; seeking an input placement which moves the moment at minimum cost.

First, we solve the state-selection problem for a number of variants of the first and second moment, and find solutions related to the eigenvalues of the systems' Gramian matrices. We then nest this state selection into projected gradient descent to design an optimal input augmentation.



Stephen Townsend
Ph.D. in Computational Science
Advisors: *Peter Beerli & Anke Meyer-Baese with acknowledgement to Dennis Slice*

A Fast and Automated Pipeline for Time-Domain Bioacoustic Analysis

Speed, efficiency and automation are extremely important for modern field researchers who are increasingly using computer tools in their work. Bioacoustics has traditionally involved very computationally expensive and very complex techniques which, while highly accurate, can be very time consuming and can require a lot of computational power which may not be easily available to researchers in the field. Another aspect of modern software-based analysis pipelines is the need for end-to-end automation. Most systems require a good amount of scripting, parameter selection and tuning and pre-processing work from the end-user. This work presents a new approach to bioacoustics analysis which allows for full end-to-end automation and a reduction in processing complexity and computational power requirements by using a strictly time-domain and shape-driven approach.

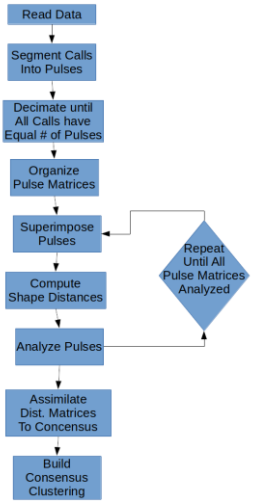


Figure 1: A block-flow diagram of the steps of the pipeline.

1

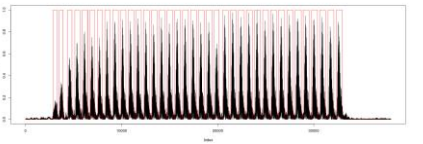


Figure 1: A demonstration of a simplified modification of the Pan-Tompkins Algorithm for extracting pulses from a normalized signal

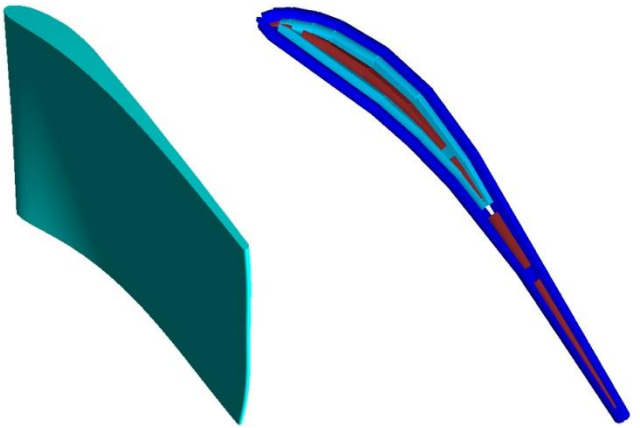
1

Liam White
Ph.D. in Computational Science
Advisor: *Bryan Quaife*

A Novel Method for Deep-Reaching Adaptive Skeleton Path Generation

Within the field of large-scale additive manufacturing (3D printing) ORNL Slicer 2 is a novel slicing engine that is used to convert digital objects into specific printing instructions. To accomplish this, Slicer 2 imports a digital representation of a model and slices it into layers. Toolpaths are generated for each layer according to the user specified settings and exported as G-Code (machine readable manufacturing instructions).

One such type of toolpath is known as a Skeleton. Skeleton path generation is defined as the generation of single bead-width paths in regions where closed contour paths are not possible or ideal. Here we present a novel method for Deep-Reaching Adaptive Skeleton path generation which is comprised of four integral subprocesses: Input Geometry Cleaning, Voronoi Diagram Generation, Output Geometry Cleaning, and Path Adaptation and Consolidation. By adapting and consolidating the generated paths, Skeletons are able to optimally fill their environment, thus eliminating over-filling and minimizing under-filling.



Above: Narrow curved wing (left), Layer 1 slice using 1 Perimeter, 1 Inset, and Deep- Reaching Adaptive Skeletons (right)

Jingze Zhang
Ph.D. in Computational Science
Advisor: *Gordon Erlebacher*

Unsupervised representation learning on written language style

Over the last few years, NLP researchers proposed various machine learning models to understand, analyze, and generate human language. However, the study on language style only limits within style transfer and usually makes a simplifying assumption that defines style into one or several categories, such as sentiment and formality. This study aims to develop a neural network method to learn a comprehensive style representation. The learned embedding performs well on 15 different style classification/regression tasks, comparing to singly trained tasks. More experiments and evaluations are in progress.

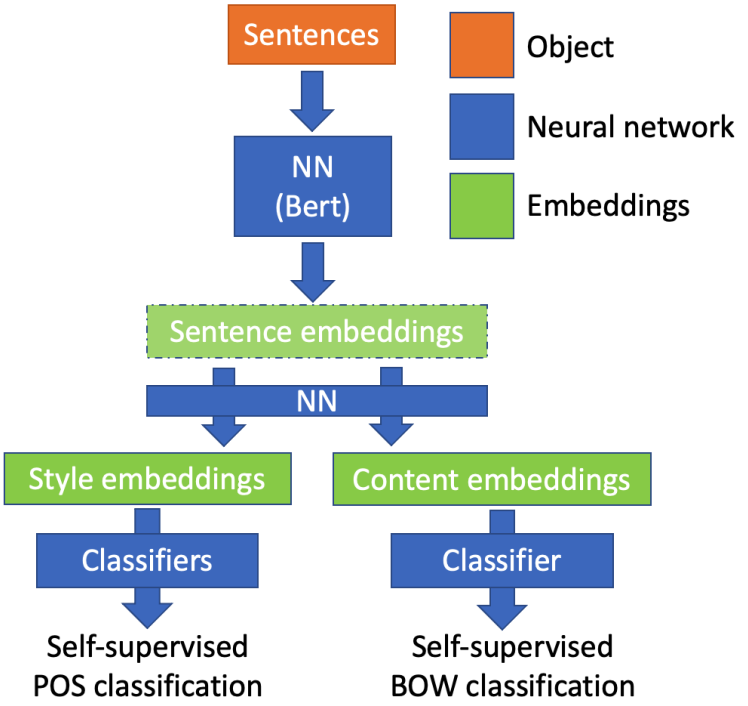
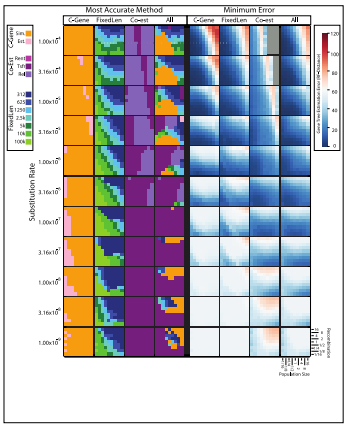


Figure 1: The architecture of the unsupervised style representation learning model.

Kevin Ziegler
Ph.D. in Computational Science
Advisor: *Alan Lemmon*

Estimating Accurate Gene Trees in the presence of Intra-locus Recombination; a Simulation Study

Accurate gene trees are difficult to estimate with traditional methods due to the effects of recombination. New methods that co-estimate gene trees and recombination breakpoints function differently than the traditional maximum likelihood (ML) framework, and therefore have the potential to alleviate inaccuracies caused by recombination. However, the accuracy of gene trees produced by these methods has yet to be evaluated under a broad range of conditions. Using simulations, we studied gene tree accuracy in the presence of intra-locus recombination. Using a previously published model of human population history, we simulate the process of recombination along large sections of a genome to produce DNA sequence alignments. We varied three parameters that influence gene tree accuracy: recombination rate, population size, and substitution rate. We then compare the accuracy of gene trees estimated from different methodologies, including traditional maximum likelihood estimation of a single and concatenated regions,as well as more sophisticated co-estimation methods. Not surprisingly, we found that traditional approaches can only produce accurate gene trees in narrow regions of parameter space, even when gene regions are divided accurately on recombination breakpoints. As the number of analyzed sites is increased, recombination reduces accuracy well before accurate gene trees can be obtained. Some, but not all of the newer co-estimation methods successfully circumvent this tradeoff and have the potential to produce accurate gene trees in broader regions of parameter space. Although these new approaches are typically applied to whole genomes, we found that they can produce accurate gene trees using much smaller genomic regions (e.g. 100kb) as well. These results indicate that by adopting co-estimation methods, systematists may be able to improve gene tree accuracy.



Left: Summary figure in which the left-hand side of the figure displays the best method for different categories using a color key, and the right-hand side displays the accuracy of that method. Categories tested are C-gene, FixedLength, Co-estimation, and All categories together. Unless substitution rate is very low, estimating a good gene tree is possible if an appropriate method is chosen.