Swarm Information Foraging to Develop Forest Inventories

Dhruva Arun
SEAP Intern
Distributed Autonomous Systems Section
Navy Center for Applied Research in AI
Naval Research Laboratory
Washington DC, United States
Email: dhruva.ranger@gmail.com

Loy McGuire
SSEP Intern
Distributed Autonomous Systems Section
Navy Center for Applied Research in AI
Naval Research Laboratory
Washington DC, United States
Email: loy.mcguire@nrl.navy.mil

Donald Sofge
Section Head
Distributed Autonomous Systems
Navy Center for Applied Research in AI
Naval Research Laboratory
Washington DC, United States
Email: donald.sofge@nrl.navy.mil

Abstract— **This study addresses the challenge of accurate wildfire prediction through comprehensive forest attribute databases. Essential Diameter at Breast Height (DBH) metrics for tree assessment are difficult to obtain at scale. We propose a swarm-based methodology for DBH value selection and collection. Comparing tree selection value algorithms, we find Gaussian Regression as a UAV decision making algorithm excels in minimizing information uncertainty and accurately predicting DBH values. This research showcases the potential of Gaussian Regression in information foraging problems as well as the effectiveness of utilizing swarms for DBH collection and predictions.**

I. Introduction

Wildfires are fast, destructive, and erratic, rendering them a threat to both civilians and infrastructure. The trajectory of a wildfire depends on several factors, including tree species, tree heights, and tree trunk diameters. Accurately predicting wildfires requires a database of all of these tree attributes for forests across the globe. Many organizations are working towards developing such an inventory with various strategies and resources. [3]

Diameter at Breast Height (DBH) is a crucial metric for assessing tree growth, health, and overall ecosystem dynamics. It is challenging to sample DBH values at a large scale. However, research suggests that DBH values can be predicted with the tree's height and the density of trees in the area. By carefully selecting a set of trees with varying heights and area densities to sample within a forest stand, the rest of the tree diameters can be accurately predicted [1].

Swarm intelligence has proven to be a powerful tool in information foraging. [4] The objective of this paper is to design a swarm methodology that selects and collects DBH values which most effectively identifies a trend function capable of predicting the DBH of the rest of the trees in the forest stand.

II. Setup

A. Simulation Environment

The simulation environment is based on a dataset collected by USDA at the Manitou Experimental Forest, containing tree coordinates, heights, and diameters. The database is parsed and a spatial grid of cells containing either trees or open areas is constructed. Each tree is assigned a density value based on the number of and distance to nearby trees using computations from a density kernel.

A 3x3 “home” area is designated at the center of the forest, where all UAVs are deployed from and return to. Each UAV gets deployed to a random cell on the grid, assuring spatial coverage across the forest. The decisions each UAV makes runs as a separate task on a separate thread, thus establishing a distributed swarm network. Each UAV has access to limited fuel and must return to the home area before the fuel runs out. The fuel depletes proportionally to the distance the drone travels, and the drone spends fuel to measure the DBH value of a tree. After returning home, a UAV may be redeployed and its fuel will be refilled. This restricts the distance the UAV can travel and the number of samples it can gather, increasing the need for intelligent selectivity in the simulation.

The swarm is limited to 5 UAVs, each of which are deployed once, resulting in a small amount of collected data points, further increasing the importance of intelligent selectivity.

Throughout the simulation, each UAV only has access to nearby forest information within a specified radius. This reduces the amount of computation required in the simulation, thus allowing the simulation to scale up.

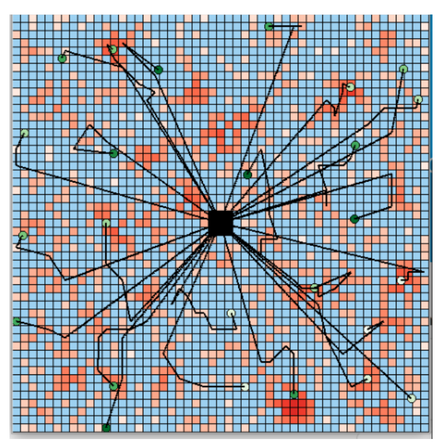
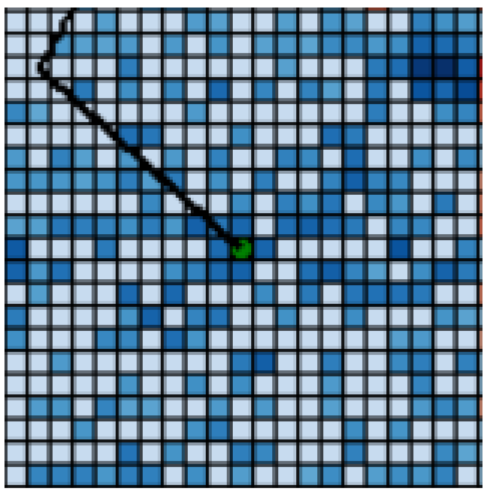


Fig. 1. - This image shows a simulation larger than the environment used in this paper. The shade of red represents the tree density.

B. Potential Field Overlay

The UAV tree selection process makes use of potential fields, which are grids with different values, or potentials, in each cell. The information field holds the information value of each nearby cell, which is computed using the algorithms described in the next section. The proximity field is pre-defined, favoring closer cells over further cells. The two potential fields are combined with a weighted sum, and the UAV chooses the cell with the greatest value, represented as the darkest cell in Figure 2.



Overlaying potential fields allows for changes in UAV decision-making to be easily implemented if additional constraints are placed or different forms of information are introduced.

Fig. 2. Potential Field

C. Performance Metric

The performance metric of the simulation measures the usefulness of the collected DBH data points in predicting the rest of the forest, thus valuing intelligent selectivity. Gaussian Regression is used to compute a trend function relating density and height to tree diameter by fitting the sampled data points [2]. The difference between the predicted trend function and the actual DBH values is then evaluated using mean squared error. The MSE difference is the sole metric of the simulation.

III. Information Value Algorithms

This section describes different methods used to compute the information value of a given cell, which is the main influence of UAV tree selection.

A. Random Information Value

This algorithm is a control for a baseline comparison. The information value of each nearby cell is completely randomized.

B. Information Value from Threshold

This algorithm does not consider the DBH value of the sampled tree, and only aims to sample trees across the range of tree heights and densities. A 2D grid with an X axis of density and Y axis of height keeps track of the information value of

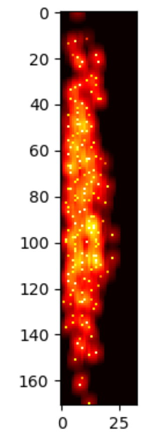


Fig. 3: Threshold

every (height, density) coordinate. The information of a coordinate is solely dependent on the previously collected coordinates.

Each sampled tree is plotted on the threshold as a (height, density) coordinate, signifying the coordinate now contains very little additional information, along with a small surrounding radius to show that the information value of nearby coordinates has also reduced. This is because to build an accurate trend function, a wide spread of tree heights and densities must be sampled. Gathering many data points with similar heights and densities does not provide any additional insight about the overall trend of the forest stand.

C. Information Value from Gaussian Regression

Gaussian Regressions are unique in that they not only determine the curve of best fit for a set of points, but also the uncertainty of the prediction of each data point. Building a good trend function requires minimizing the amount of uncertainty for the curve prediction. Thus, utilizing Gaussian Regression to choose data points with minimal uncertainty directly works towards building a good trend function. The information value of a tree can be quantified simply as the amount of uncertainty of the Gaussian Regression at the respective (height, density, DBH) coordinate in 3D space.

An initial attempt at Gaussian Regression of randomly selected points showed that there was very little correlation between density and DBH, high correlation between height and DBH. It also yielded a result with a constant and low uncertainty throughout the graph (Fig. 4), suggesting the model required tuning to yield accurate uncertainty data.

Since density did not have much of an impact on DBH values, it could be ignored during the tuning process and tuning could be done in 2D (height and DBH).

Different hyper parameters for the smoothness, length scale, noise level, and kernel width were tested with various number of training dataset sizes to yield smooth results with larger uncertainties occurring in areas that have fewer data points or in areas with data points that significantly diverge from the general trend of the previously collected data points. Figure 5 shows the Gaussian Regression output with the final hyper parameters for various training dataset sizes.

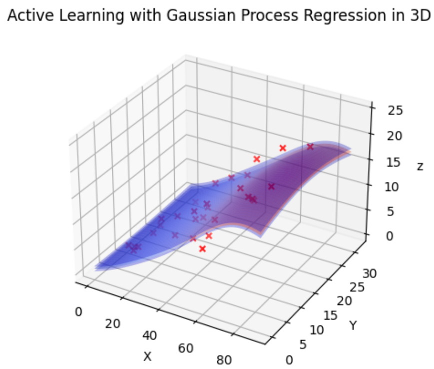


Fig. 4. – Initial Gaussian, X Height, Y Density, Z DBH (Note the lack of correlation between Y and Z)

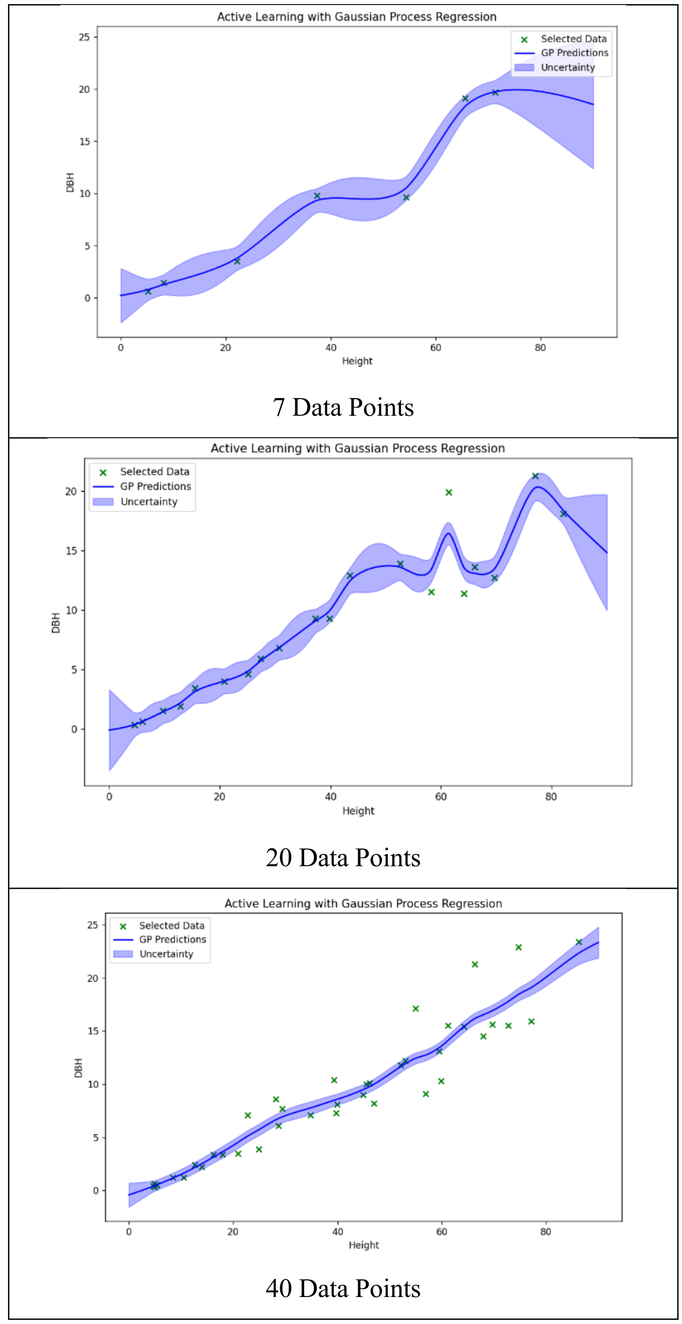


Fig. 5. – 2D Gaussian Tuning

IV. Results

A. Data

Thirty (30) simulations were run with each of the information value algorithms, and the performance metric (MSE of DBH pred. and actual) was measured. Figure 6 exhibits box plots of the output metric values for each algorithm.

Gaussian Regression as an information value algorithm performed better than the threshold algorithm. Both performed significantly better than the random information assignment algorithm.

B. Discussion

This study has shown that Gaussian Process Regression is very effective as a decision-making algorithm for UAVs to minimize uncertainty in a trend function. GPRs can be applied to other information foraging problems that attempt to determine trend functions between various variables.

Currently, GPRs in swarms may not be scalable as the GPR repeatedly fits the whole set of selected points for each new point selection. Developing a strategy to optimize this process, such as only performing regression on segments of the selected dataset based on the new trees available, will allow for the scaling of this foraging algorithm to large forest areas with large amounts of UAVs.

Density has no correlation with DBH values, despite our intuition. It is possible that the density-DBH correlation is too complex to be discovered by GPRs, so further research exploring this topic may yield interesting results.

Introducing more variables, such as species and location, may yield trend functions with smaller mean squared errors and will allow swarms to travel across various stands, instead of being limited to a single stand per GPR.

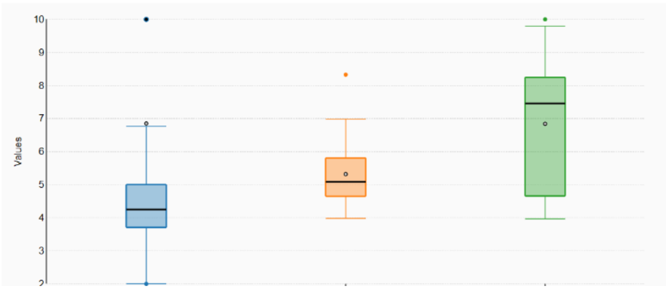


Fig. 6. – Box Plots for MSE in GPR, Threshold, and Random Information Value Algorithms respectively

Figure 7 exhibits the GPR fit on a set of selected points. The 2D tuning process proved to be effective in increasing the variance of uncertainty across the 3Dgraph, thus increasing intelligent selectivity.

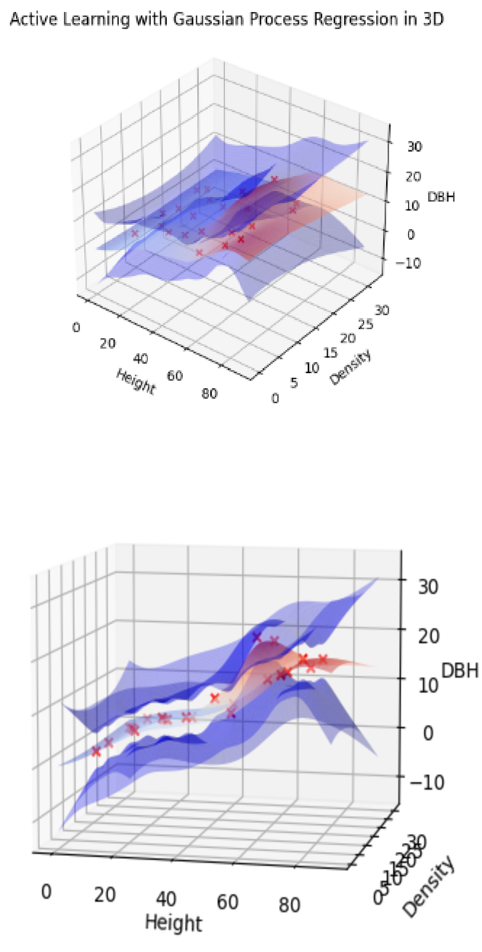


Fig. 7. – GPR 3D prediction graph

REFERENCES

[1] Swayze, N. C., Tinkham, W. T., Vogeler, J. C., & Hudak, A. T. (2021). Influence of flight parameters on UAS-based monitoring of tree height, diameter, and density. *Remote Sensing of Environment*, *263*, 112540. https://doi.org/10.1016/j.rse.2021.112540

[2] Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, *85*, 1–16. https://doi.org/10.1016/j.jmp.2018.03.001

[3] *Forsite Consultants Ltd - Forest Management Specialists, Western Canada*. (n.d.). Www.forsite.ca. Retrieved August 18, 2023, from https://www.forsite.ca/

[4] Bernstein, J. (2011). *Project Swarm How technological advances are inspired by swarming animals*. https://jeremybernste.in/projects/swarm/printerfriendly.pdf