Data Science Internship at the University of Göttingen, Germany

Intern needed for statistical analyses and data visualization on a wildlife research project. Spend a summer working with PhD students from different fields and learn how computer science, statistics, and biology together improve our understanding of nature. You are going to help us explore data on small mammal behavior, both by means of explorative visualizations and statistical models. To this end, you will be working on exciting programming tasks using modern computer languages and tools for statistics and data science. But that's not all: Every good statistician needs to understand their data and how it is collected. For this reason, you are going to join our biologists on their field trips and learn about small mammal live-trapping methods, animal handling and identification, tissue collection, behavioral observations, and much more. Our field sites are spread across northern Germany, giving you the opportunity to see a wide variety of local landscapes, and the city of Göttingen, where our university is based, is one of the most vibrant, diverse, and beautiful small cities in Germany.

Research Topic

In this project, we are working on the development and application of modern statistical methodology, particularly of flexible regression models, to different types of small mammal data. The data is collected on eight study sites, each with five plots representing different mixtures of tree species. We are trying to understand and model the differences in small mammal abundance, community composition, and behavior between different forest types. Our focus is on European Beech (Fagus sylvatica) and Norway Spruce (Picea abies), which are native to Europe, as well as Douglas Fir (Pseudotsuga menziesii), a non-native conifer from North America. We are investigating both pure stands of all of these species and mixed stands of beech and each conifer.

To assess small mammal abundance and community composition, we conduct mark-recapture surveys. All five plots at each site are surveyed concurrently for four nights, twice per summer season. The traps are set every 10 m in an $8 \times 8$ configuration, and checked three times per day. Individuals captured are identified to species, measured, and injected with a unique PIT tag for re-identification. In addition, camera trapping of small mammals is conducted following each mark-recapture survey to identify the presence of trap shy species. We also conduct seed preference trials to assess the impact of small mammals on forest regeneration. All of these experiments give rise to rich and exciting datasets, on which you can work during the internship. Depending on your skills and interests, some potential research questions could be:

- Based on the mark-recapture data, which species are trapped more frequently and which species tend to avoid traps? What is the probability to recapture an individual after it has first been trapped? Does this probability differ between juveniles and adults or between forest types? These problems can be tackled with different kinds of regression models.
- In contrast to the mark-recapture experiment, the camera traps can be visited multiple times and by different individuals. We would like to find out which traps are most popular and why. Do different species have different preferences? This dataset is ideal for visualizations like heatmaps, animations of trap visits over time, or even interactive data exploration tools.
Another advantage of camera trapping is the possibility to track the movement of individual animals in a comparatively high temporal resolution, as illustrated in the figure on the first page. This raises the question whether there are patterns in the movement of individual animals. How large is their home range? Do different species have differently sized home ranges? And do they remember specific traps and return there systematically?

Tasks and Requirements for the Intern

You are going to assist us with statistical analyses and data visualization as well as with data collection in the field. This work includes among other things:

- **Implementation and application of statistical models** in Python
- **Data and model visualization** with Python or R
- Data processing and cleaning as needed
- **Capture and handling of small mammals** (mark-recapture) and genetic sampling
- Observation of small mammal behavior for seed preference trials
- Installation and checks of camera traps for biodiversity surveys

The desired skills are:

- Solid **programming skills**, preferably in Python and/or R
- Solid **understanding of basic mathematics**, particularly of linear algebra and calculus
- Preferably some experience with scientific computing and the NumPy software package
- Preferably some understanding of statistics, although this is not strictly required because we are going to guide you on the implementation of the statistical models

As you are going to spend part of your internship in the field, an interest in wildlife and nature is required. The days in the field can be long and require a flexible schedule, so a sense of humor, positive attitude, patience, and self-motivation are essential. The successful applicant will have demonstrable academic experience, work well independently and with others, communicate effectively in a variety of situations, demonstrate the willingness and ability to live in shared housing, enjoy working outdoors, and be conscientious about safety. Applicants must be physically fit enough to move quickly across the landscape in variable weather conditions, sometimes while carrying heavy loads.

About RTG 2300 and the University of Göttingen

The research training group (RTG) 2300 “Enrichment of European beech forests with conifers: impacts of functional traits on ecosystem functioning” is investigating mixed forest stands composed of native European beech, Norway spruce, and non-native Douglas fir. It is hypothesized that the mixtures show higher functional diversity than pure European beech stands and that the effects of admixed conifers on ecosystem functioning are more pronounced in mixtures with Douglas fir. To test this hypothesis, PhD students and a postdoc work in 11 closely linked interdisciplinary subprojects exploring various functional traits of the three tree species and associated organisms.

The University of Göttingen is highly reputed throughout the world of academia and research, and would not be the same without the numerous international partnerships and exchange opportunities, international students, and academics on campus from more than 130 countries. There is hardly any other German town of comparable size with so many institutes of higher education and notable academic facilities. Almost 50,000 people here are engaged in research and teaching on a daily basis. Göttingen is situated in the very center of Germany at the intersection of several high-speed rail lines, and is well connected to the rest of Germany and Europe, so your stay will give you the opportunity to explore other exciting places as well.

Contact Data

Hannes Riebl, Chair of Statistics, University of Göttingen
[hriebl@uni-goettingen.de](mailto:hriebl@uni-goettingen.de) +49 551 - 39 27237
**Additional Information**

Though the internship will take place in Germany, the working language of this research group is English, so while some knowledge of German is helpful it is not required.

The internship provides a travel subsidy to cover the flight to Germany and back, plus 750 Euros per month stipend to cover living costs (housing and food). The program also includes a group excursion for all interns to Heidelberg that is paid for by the DAAD.

This internship is open to current undergraduate students attending universities in the United States, Canada, Ireland, and the UK. Students should have completed at least two years of undergraduate study before the start of the internship, and must retain their student status after they complete the internship, i.e. you must be enrolled for the fall 2020 term, and you cannot apply if you will graduate before fall 2020.

Note that the earliest start date is May 15th, 2020, but the start date is flexible depending on the academic schedule of the successful applicant. Applicants who are available for a full three months are preferred.

To apply, visit the DAAD RISE program website at: [https://www.daad.de/rise/en/rise-germany/find-an-internship/application-portal/](https://www.daad.de/rise/en/rise-germany/find-an-internship/application-portal/)

The reference code for this internship (which you can use to find this specific offer at the DAAD RISE program website after registering) is: Goettingen_CS_BI_3746

For further information, contact Hannes Riebl: hriebl@uni-goettingen.de