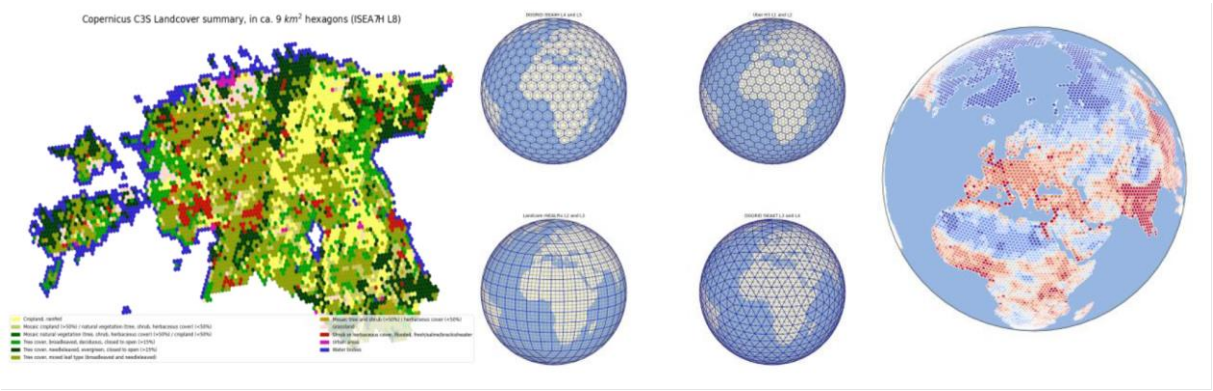
UNIVERSITY OF TARTU
Landscape Geoinformatics Lab

4-year PhD Position in Geoinformatics: A systematic approach to re-gridding and data integration with Discrete Global Grid Systems for machine learning readiness

Global-scale ecological analyses still rely widely on latitude-longitude gridded data and cartographic map projections that inadvertently introduce substantial distortions. To apply machine learning on spatial data or perform spatially distributed modelling, the data needs to be accessible through a uniform interface in standardised format and variables aligned. Moreover, when projecting the spherical Earth onto a plane, which is required for present-day modelling and geospatial analysis methods, area distortion takes place. This is especially critical in large-scale models (at regional and global scales) because the impact of the areal distortion increases with increasing extent of the study area.

The emerging field of Discrete Global Grid Systems (DGGS) offers a method for organizing, storing, and analyzing spatio-temporal data across various resolutions and scales, ranging from near-global dimensions down to microns. Such multi-resolution DGGS facilitate spatial aggregation across scales and when used with the equal-area paradigm, spatial analysis can be replicated consistently anywhere on the Earth's surface, independent of the resolution or scale. The global scope of a DGGS means that algorithms defined for one region of the DGGS reference frame can be deployed without modification anywhere within the DGGS and without having to accommodate different map projections. DGGS is well-suited for efficient storage, visualization, analysis, and modeling of multi-source data in a data cube. However, until newly observed and generated data is not accessible already in a DGGS format, current data needs to be transformed into the DGGS gridded format. This poses the Modifiable Areal Unit Problem (MAUP) - the outcomes of spatial data analysis depend on how areal units are modelled. In the classical situation bias arises from the use of polygons of varying size and shape, such as administrative units, catchment or habitat boundaries, countries, or climatic regions, to model spatially distributed variables like species distribution, soil properties, or to analyse areal changes in land use or forest cover.



The aims of the PhD project:

- Systematically evaluate MAUP effects on transforming raster and vector data of varying resolutions into DGGS grids, particularly hexagonal grids. Utilize spatial statistical indices for methodological assessment.

- Propose, develop, and test methods for integrating diverse data (Earth observation, land cover, hydro/climate, met/ocean) into DGGS data structures at multiple scales and resolutions with minimal distortions and bias, ensuring accuracy in spatial statistics and modeling

DGGS provides an effective and novel method for integrating diverse data sources across scales and boundaries. This project examines DGGS's capabilities as a groundbreaking approach to geospatial data

integration and analysis, enhancing accuracy in spatial statistics and modelling for environmental studies. The research is part of the "Creating Water-Smart Landscapes" project funded by European Research Council and Centre of Excellence for Sustainable Land Use. The PhD student will be working in the Landscape Geoinformatics Lab in the Department of Geography, University of Tartu.

Core tasks include:

- Design systematic experiments to statistically quantify MAUP effects using synthetic and real-world data (such as remote sensing data) across various DGGS topologies and resolutions
- analyze the impact of hierarchical DGGS grid resolutions on multi-source, multi-resolution data integration
- Consider data storage and query formats suitable for GIS and cloud-native geospatial technologies
- Develop best-practices, recommendations for practitioners

Requirements

Master's degree in geography, geoinformatics, quantitative ecology, physics or wider spatial and natural sciences.

The ideal candidate has proven experience in following essential skills:

- spatial analysis and GIS, comfortable with QGIS
- scripting of workflows and data analysis with geospatial Python
- confident in English language spoken and written. See more info about language requirements https://ut.ee/en/sisu/phd-language-requirements English language test must be submitted by 15th of May
- good communication skills and willingness to work in a team
- Willingness to engage with stakeholders and the wider DGGS ecosystem at ESA, Pangeo, Osgeo, OGC, and the Statistical Offices

In addition, following skills would be beneficial:

- understanding of and experience in using machine learning methods (such as Random Forest)
- understanding of landscape processes (physical geography)
- Interest in the open-source geospatial software landscape and community

Desired knowledge

Priority will be given to candidates with experience in spatial data analysis, programming knowledge, especially in Python. For shortlisted candidates we will also aim to assess technical and spoken communication skills in an online interview.

Funding and Health Insurance

The position is fully funded. Full-time PhD student will be on a junior researcher position with gross salary of 1830 EUR (net approx. 1450 EUR) as minimum with expected increase of 5-10% per year. Living costs in Estonia are very reasonable and the allowance can cover your living costs.

All PhD are provided with Estonian national health insurance. Health insurance coverage is available for the full nominal study period of PhD studies (4 years). University is also covering regular health checks and some health improvement (e.g. gym, swimming) costs for the staff members.

Living in Estonia

In Estonia you will be living in a highly connected society, with free wireless Wi-Fi almost everywhere. Many everyday activities are made easier with various IT solutions: register a company with as little as 18 minutes, park your car with phone, register courses online etc. Entrepreneurship and innovative solutions are highly welcomed in Estonia, which has a strong start-up community and has also become known as the Silicon Valley of Europe. Student life in Estonia is full of activities and events. There are many organizations and events that help foreign students to settle into Estonian life and create a social network in the country. For more information about Living in Estonia https://studyinestonia.ee/living

Start of the studies: 2nd of September 2024

To apply

Please submit the following materials via email to evelyn.uuemaa@ut.ee by **20th of March** :

- A one-page letter of motivation explaining how your prior research experience qualifies you for the position and why you would like to work on this topic.
- Copies of both bachelor's and master's diplomas, and diploma supplement (bachelor's and master's transcript/mark sheet, including the description of the grading scale). Applicants graduating in 2024 and having their diplomas issued later than the application deadline, should electronically submit their most recent official Transcript of Records by the required deadline along with the rest of the required documents. Please note that you must have graduated from your master´s studies by 1 August, 2024.
- A two page CV (including publications if available).
- Names and contact details for one references.
- Applicants selected for an interview will be contacted before the end of March 2024.