ABSTRACT

The pivotal role of remote sensing in Earth observation lies in its capacity to gather comprehensive data about our planet, overcoming the limitations of traditional, in-situ measurements. Remote sensing often necessitates "inference", as sensors cannot directly capture the specific information sought. Algorithms play a crucial role in this process, leveraging data characteristics like reflectance and shape to infer the needed information. The integration of artificial intelligence (AI) with geospatial science, particularly through deep learning, has significantly advanced these capabilities by learning the relationship between observations and human-provided labels, thus enhancing the toolset for inference.

In the development of remote sensing algorithms, emphasis has often been placed on the accuracy directly obtained from the algorithms themselves, highlighting the initial inference capabilities. However, this firsthand accuracy represents only a fraction of the broader narrative within the geospatial mapping workflow, which encompasses data collection, algorithm application, and final validation stages. The error patterns generated by these algorithms are equally critical to their overall accuracy, as understanding these patterns can significantly aid error management and enhance data collection process designs. Such understanding is crucial, especially as the volume of geospatial mapping outputs expands beyond the human capacity for validation, and errors can potentially mislead subsequent analyses if not identified and rectified promptly. This challenge is particularly pronounced in the current era, where deep learning-based methods have become a significant part of geospatial mapping outputs. Their dependence on imperfect human-provided labels, combined with their opaque nature, often results in spatially heterogeneous errors that can undermine the trustworthiness of geospatial mappings. Moreover, such errors can result in spatially biased outcomes, jeopardizing crucial decision-making processes. Despite its critical importance, this aspect has often been overlooked.

In response, this dissertation champions the development of a mapping system that emphasizes transparency and scalability to effectively address the challenge of spatially heterogeneous errors. It advocates for the use of generalizable prior knowledge about targets to offer clear definitions, thereby avoiding complex inference procedures and enhancing the

system's transparency and scalability across extensive areas. The dissertation introduces a transparent and scalable 3D urban mapping system that focuses on the universal physical properties of critical urban features. Further building on this system, a novel, transparent, and scalable framework, 3D Landscape Clustering, is introduced to improve urban climate studies by effectively and reliably incorporating landscape variables. Additionally, the dissertation presents a deep learning method for addressing spatially heterogeneous errors in building mapping, a technique for extracting building elevation attributes, and a novel tool for analyzing urban pathway accessibility. Also, an evaluation study on the spatially biased error in data-driven AI-based mapping is presented.

This dissertation underscores the necessity of extending our focus from mere accuracy to also consider the broader implications of error patterns in geospatial mappings. It demonstrates the value of integrating generalizable target knowledge, as explicitly represented in remote sensing data, into geospatial mapping to tackle the challenges of trustworthiness in AI mapping systems. By developing mapping systems that are open, transparent, and scalable, this work seeks to reduce the effects of spatially heterogeneous errors, thereby improving the reliability and reproducibility of geospatial data analysis across various fields. Moreover, this dissertation introduces methodologies to support urban heat, pathway accessibility, and flood management studies through dependable geospatial systems. These efforts aim to establish a robust and effective geospatial foundation for informed urban planning, efficient resource allocation, and enriched environmental insights, contributing to the development of more sustainable and resilient cities.