

Geodesic vs. Euclidean Dense Retrievers: Robustness to Adversarial and Noisy Queries in BEIR-NL

Assignee Research

June 3, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: How do geodesic distance-based dense retrievers perform compared to Euclidean-based models in terms of robustness to adversarial perturbations or noisy queries in the BEIR-NL benchmark. 7 claims were extracted from source literature; 6 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Effects of high-quality elevation data and explanatory variables on the accuracy of flood inundation mapping via Height Above Nearest Drainage. Research question: How do geodesic distance-based dense retrievers perform compared to Euclidean-based models in terms of robustness to adversarial perturbations or noisy queries in the BEIR-NL benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.7/10.

3 Results

4 papers retrieved. 7 claims extracted; 6 independently verified. Quality review score: 7.7/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
Using high-quality and high-spatial-resolution digital elevation maps (DEMs) from the United States Geological Survey's	✓	0.52
Varying the spatial resolution of the 3DEP DEMs at 3, 5, 10, 15, and 20 meters showed no significant overall effect on F	✓	0.33
Further analysis at coarser resolutions of 60 and 90 meters revealed a significant degradation in FIM skill, highlightin	✓	0.32
The experiments demonstrated a significant burden in terms of the computational time required to produce HAND and relate	✓	0.29
A multiple linear regression model was fit to help explain catchment-scale variations in the four metrics employed.	✓	0.23
The lack of reservoir flooding or inundation upstream of river retention systems was a significant factor in the analysi	✓	0.26
Interagency Flood Risk Management data was used for validation.	×	0.12

References

- <https://doi.org/10.3390/ijgi13040138>
- <https://doi.org/10.5194/hess-28-1287-2024>
- <https://doi.org/10.1016/j.jss.2025.112436>