

# Bachelor Thesis

Assessment of approximation method for TSP path length on road networks: a simulation study

Koen Stevens, S5302137

2025-04-10

## Introduction

The Traveling Salesman Problem is an important problem in operations research. It is particularly relevant for last-mile carriers and other logistics companies where efficient routing directly impacts cost, time and service quality. Since the number of parcels worldwide has increased between 2013 and 2022 and is expected to keep increasing [Statista (2025)], the need for fast, scalable route planning methods becomes ever more pressing.

The TSP is an NP-hard problem, it is computationally intensive to find the exact solution for large instances. In many real-world scenarios, companies may not need the exact routes for deliveries, but instead require rough, reliable estimates. For instance a last-mile carrier or a food delivery chain expanding to a new area. Then this firm may only know what area they will serve and have some estimates of how many customers they will get. In such cases, they may need quick answers to questions like how many vehicles and personnel they will need.

Efficient approximation methods provide a solution for such practical applications where exact solutions are too computationally intensive to conduct or not feasible due to insufficient data. These methods aim at approximating the expected optimal total travel time or distance, while using minimal data and computational effort.

There is extensive research on such approximation methods and how they perform in the euclidean plane. Consider  $n$  uniformly drawn locations inside some area in  $\mathbb{R}^2$  with area  $A$ . Beardwood, Halton, and Hammersley (1959) prove the relation:

$$L \rightarrow \beta\sqrt{nA}, \quad \text{as } n \rightarrow \infty \tag{1}$$

as an estimation for the length of the shortest TSP path measured by Euclidean distance through these random locations, where  $\beta$  is some proportionality constant. This formula is a very elegant result, approximation and it requires very little data. However, its assumptions, uniform random locations and euclidean space differ from real-world applications, which are defined by complex geographic features, such as road networks.

This research investigates how well this approximation method performs when we consider real road networks. Using OpenStreetMap data, we simulate TSP instances in a selection of different areas in the Netherlands, then solve these for the actual shortest paths using the Lin–Kernighan heuristic. Then we estimate  $\beta$  from equation 1 and analyze the performance of this formula. Additionally, we compare the results for  $\beta$  and the performance across the selected areas, with the goal of understanding the driving factors behind the varying results.

In section 2 we dive deeper in the context and previous research in this field. Then, in section 3 we show the methodology ...

## Literature Review

There is extensive research in approximation methods for the length of a TSP path.

## Methodology

## References

- Beardwood, Jillian, John H Halton, and John Michael Hammersley. 1959. “The Shortest Path Through Many Points.” In *Mathematical Proceedings of the Cambridge Philosophical Society*, 55:299–327. 4. Cambridge University Press.
- Statista. 2025. “Global Parcel Shipping Volume Between 2013 and 2027 (in Billion Parcels)\*.” <https://www.statista.com/statistics/1139910/parcel-shipping-volume-worldwide/>.

## Appendix

```
print("Hello from Python!")
```