

Abstract

Horizontal directional drilling (HDD) can induce ground settlement that poses a threat to roads, railways, and buried utilities. This study reviews and compares four approaches for estimating ground settlement: empirical, analytical, numerical, and machine learning methods. Eight representative studies are evaluated, encompassing clay, sand, layered ground, HDD, pipe jacking, and shield tunnelling. Each study is assessed using six criteria: (1) closeness to measured settlement, (2) number of input parameters, (3) overall complexity, (4) time required for calculation, (5) principal strengths, and (6) key methodological limitations. Evidence from published field measurements indicates that empirical formulas can reproduce settlement behaviour within the soils and project conditions for which they were calibrated, but their validity outside those conditions may be limited or uncertain. Analytical methods provide explicit relationships among ground loss, soil properties, geometry, and settlement-trough behaviour, but rely on idealized assumptions regarding soil behaviour, geometry, and boundary conditions. Finite-element and machine learning models demonstrated strong agreement with measurements in studies that included field validation. However, direct accuracy rankings among the reviewed approaches cannot be made reliably because the methods were applied to different projects and evaluated using different datasets and performance metrics. The study concludes that no single method is universally optimal. Instead, a combined strategy is recommended: simple empirical or geometry-based formulas may be used for preliminary HDD screening, while calibrated analytical or numerical models are more appropriate for critical crossings with strict settlement limits. Machine learning methods may provide additional predictive capability when sufficient high-quality, project-specific monitoring data are available.



Read the full paper
Scan to open the article.

Ground Settlement Estimation in HDD: Lessons from Empirical, Analytical, Numerical, and Machine Learning Models

Aitegin Alymzhanov

Oxford Journal of Student Scholarship, Vol. 10, No. 4 (August 2026)

Introduction

The problem. HDD drills a curved hole under roads or rivers, widens it, and pulls a pipe through. Soil around the hole can shift, making the surface sink (settlement).

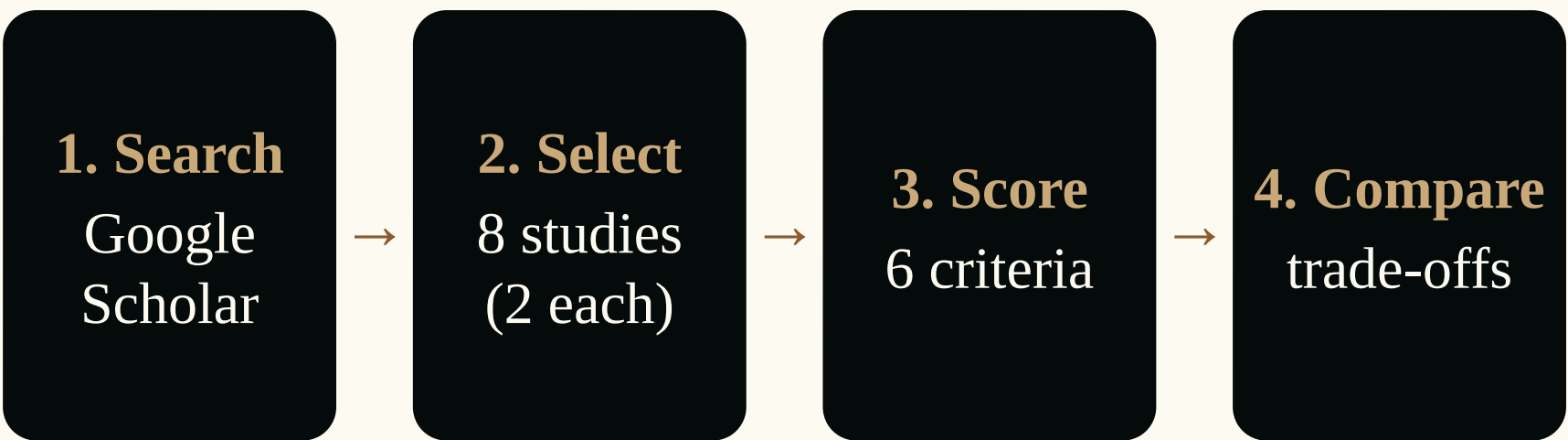
Four ways to predict it:

- **Empirical:** formulas fitted to past field data
- **Analytical:** equations derived from soil physics
- **Numerical:** computer simulations of the soil
- **Machine learning:** models that learn from monitoring data

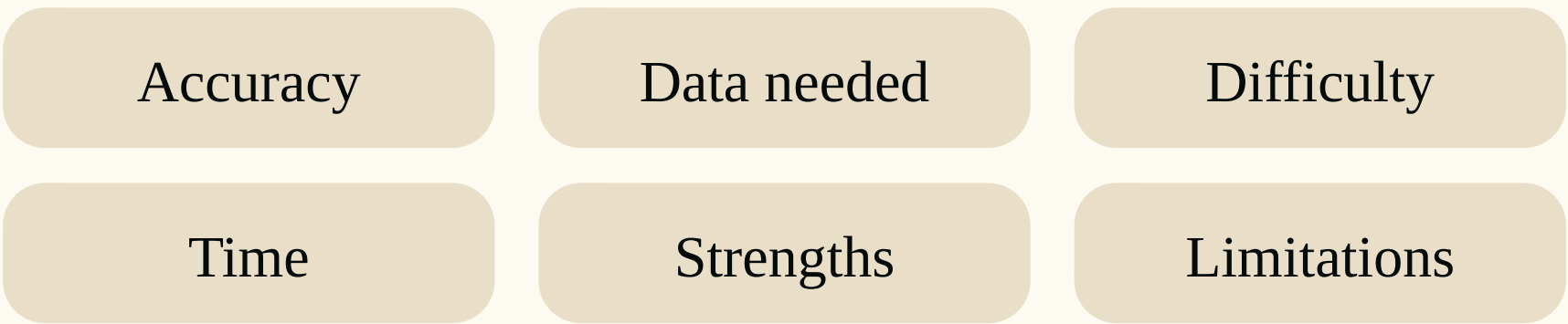
Question. Which method works best for HDD, and when?

Methods

Literature review in four steps:



The 6 criteria:



Some studies cover tunnelling, which also causes ground loss.

Limitations

- Only 8 studies, because many journals were not accessible.
- Most studies gave no exact error %, so some ratings are judgment calls.
- Each method was tested on a different project, so this shows trends, not which method is best.

Results

- Numerical and machine learning methods are more complex and need much more data and time (Figure 1).
- Where they were checked against real sites, numerical and machine learning models matched closely.

Checked against real measurements?

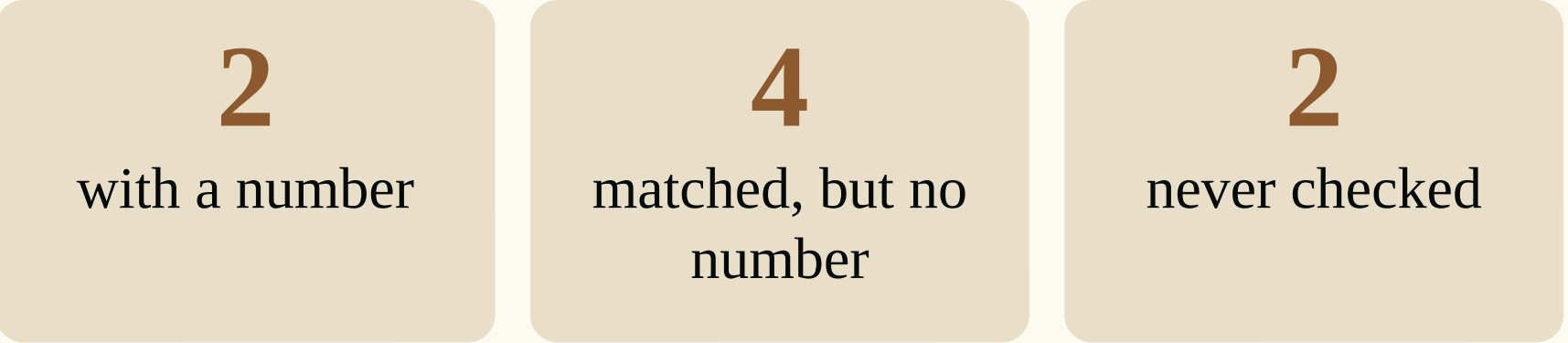
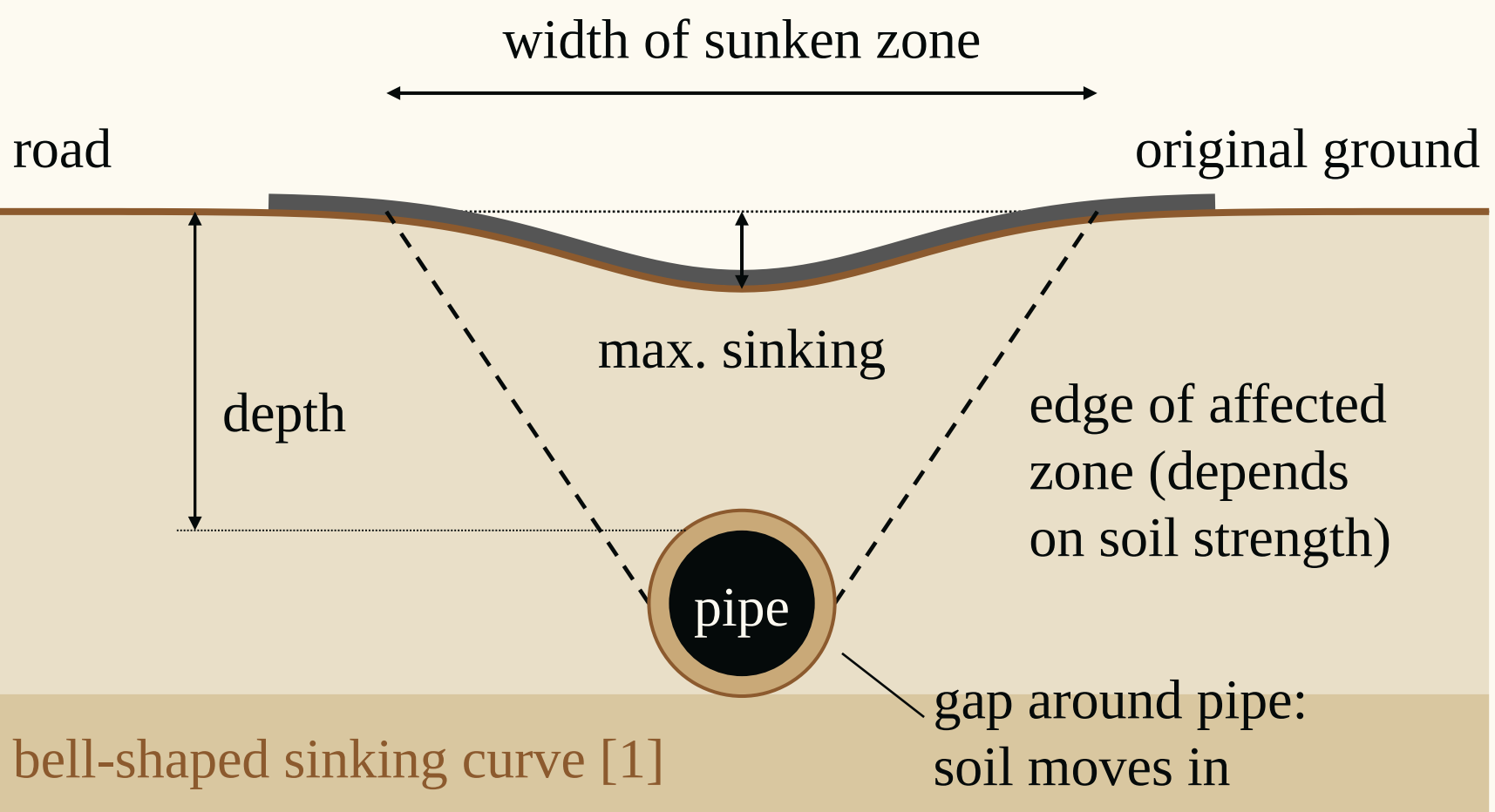


Figure 1. How the eight studies compare

Study	Data needed	Difficulty	Time	Compared with real data?
Empirical				
Mair [1]	Low	Medium	Fast	Yes, no %
Wang [2]	Medium	Medium	Fast	Yes, no %
Analytical				
Wallin [3]	Very low	Low	Very fast	No
Loganathan [4]	Medium	Medium	Fast	Yes, 5 tunnels
Numerical				
Ahmed [5]	High	High	Slow	No
Xu [6]	High	High	Slow	Yes: 11% off
Machine learning				
Shang [7]	High	High	Medium	Yes, own site
Ayari [8]	High	High	Medium	Yes: R² > 0.92

Darker = more demanding (relative ratings).

Figure 2. How the ground sinks above HDD (simplified)



Discussion

- **Empirical:** simple and accurate in the soil it was built from, but may not work in other soils.
- **Analytical:** clear equations, but they assume ideal soil, so accuracy depends on how close the real ground is.
- **Numerical:** handles layered soil and complex shapes, but needs many soil inputs, time, and software skill.
- **Machine learning:** accurate on the project it learned from, but must be retrained for a new site.

Conclusion

No method was best in every case. Choose based on the situation:

Need a quick first estimate?
Use simple empirical or geometry-based formulas.

Strict limit on how much the ground can sink?
Use calibrated analytical equations or a numerical simulation.

Enough good-quality data from the site?
Machine learning can be added.

Future work: test all four methods on the same HDD project to find out which is truly most accurate.

References

[1] Mair et al. (1993). Géotechnique, 43(2).
[2] Wang et al. (2016). KSCE J. Civil Eng., 20(7).
[3] Wallin et al. (2008). NASTT No-Dig Conference.
[4] Loganathan & Poulos (1998). J. Geotech. Geoenviron. Eng., 124(9).
[5] Ahmed et al. (2023). Underground Space, 8.
[6] Xu et al. (2024). Sustainability, 16(1).
[7] Shang et al. (2024). Scientific Reports, 14.
[8] Ayari & Derbal (2025). Algiers Metro ANN study [unpublished].