

A Geotechnical Survey: Predicting Swelling Potential in Nairobi-Thika Corridor Black Cotton Soils Using Site-Specific Mineralogy and Machine Learning

S, o, f, i, a, M, e, n, d, e, s, ,, C, a, r, l, o, s, L, o, p, e, s, ,, J, o, a, q, u, i, m, F, o, r, t, e, s, ,, M, a, r, i, a, A, n, d, r, a, d, e

DOI: <https://doi.org/10.5281/zenodo.18567639>

| Abstract

Expansive black cotton soils present considerable geotechnical challenges for infrastructure across East Africa. The Nairobi-Thika Highway corridor is underlain by these soils, where conventional methods for predicting swelling potential frequently lack the site-specific precision needed for reliable design. This survey research evaluated the efficacy of integrating site-specific mineralogical data with machine learning (ML) techniques to predict the swelling potential of black cotton soils along the Nairobi-Thika expansion corridor. The objective was to assess whether this hybrid approach provides a more accurate and practical alternative to traditional empirical correlations. A comprehensive survey of existing geotechnical investigations, laboratory data, and academic literature related to the corridor's soils was conducted. The methodology focused on analysing documented applications of ML models, such as Random Forest and Artificial Neural Networks, which utilise mineralogical composition alongside standard index properties as input variables for swell prediction. The survey indicates that ML models trained on site-specific mineralogical data consistently outperform traditional empirical methods. A key theme is that models incorporating montmorillonite content show a strong positive correlation with predicted swell strain. However, the availability of comprehensive, high-quality mineralogical datasets for model training was identified as a major limiting factor. Integrating machine learning with

detailed mineralogical analysis presents a promising avenue for improving the accuracy of swell predictions in black cotton soils. This approach could enhance risk mitigation and design optimisation for infrastructure projects in expansive soil regions. Future research should prioritise the systematic collection and open sharing of standardised geotechnical and mineralogical datasets from the region. Practitioners are encouraged to validate ML model predictions with targeted site investigations before full-scale implementation. Expansive soils, black cotton soil, swelling potential, machine learning, mineralogy, geotechnical prediction, Nairobi-Thika corridor. This survey synthesises and critically assesses the emerging paradigm of using site-specific mineralogy and machine learning for swell prediction in a key East African context, highlighting both its potential superiority to conventional methods and the critical data constraints for its regional application.
