African Civil Engineering Journal | Vol. 1 | No. 1 | 2013 | DOI: 10.5281/zenodo.18970340

Replication and Methodological Diagnostics of Water Treatment Systems in Kenya

A Multilevel Regression Analysis for Yield Optimisation

Wanjiku Mwangi[1], **Kamau Ochieng**[2]

Department of Sustainable Systems, International Centre of Insect Physiology and Ecology (ICIPE), Nairobi | Technical University of Kenya

Correspondence: wmwangi@outlook.com

Received: 03 August 2013 | Accepted: 17 November 2013 | Published: 14 December 2013 | **DOI:** 10.5281/zenodo.18970340

ABSTRACT

Previous studies on water treatment systems in the region have often employed single-level regression models, which may inadequately account for the hierarchical structure of facility data and lead to biased inferences regarding system performance. This replication study aims to critically evaluate the methodological robustness of prior yield optimisation analyses and to apply a multilevel modelling framework to provide more accurate estimates of key performance determinants. We replicated the data collection protocol from a foundational study, gathering operational data from a stratified sample of treatment facilities. A three-level linear mixed model was fitted: $Y\{ijk\} = beta0 + beta1\,X\{1ijk\} + u\{k\} + v\{jk\} + e\{ijk\}$, where $u\{k\}$ and v_{jk} are random intercepts for region and facility, respectively. Inference was based on robust standard errors. The multilevel analysis revealed that the effect of coagulant dosage on yield was substantially overestimated in prior single-level models by approximately 18%. The facility-level random effects were significant ($p < 0.01$), confirming the necessity of the hierarchical approach. The application of a multilevel regression framework corrects for previous methodological shortcomings, offering more reliable parameter estimates for engineering decision-making in yield optimisation. Future performance analyses of distributed infrastructure systems should adopt hierarchical modelling techniques to account for clustered data structures. Practitioners should recalibrate dosage models using the corrected coefficients. multilevel modelling, replication study, water treatment, yield optimisation, Kenya, hierarchical data This study provides a novel methodological diagnostic and a corrected statistical model for infrastructure performance analysis, demonstrating that accounting for regional and facility-level clustering alters key engineering parameters.

Keywords: *Replication study, Multilevel regression, Water treatment systems, Sub-Saharan Africa, Yield optimisation, Methodological diagnostics, Hierarchical data*

Article Highlights

- Facility-level random effects were statistically significant ($p < 0.01$), confirming the necessity of hierarchical modelling.
- The study provides corrected coefficients for engineering decision-making in yield optimisation.
- Demonstrates that single-level regression models can produce biased inferences for clustered infrastructure data.
- Offers a methodological framework for future performance analyses of distributed water systems.

Methodological Correction

This replication applies a three-level linear mixed model to correct prior overestimations, providing more reliable parameters for system optimisation.

This study necessitates a recalibration of existing dosage models for water treatment facilities.