

Statistical and Machine Learning Approaches for Municipal Population Projection in Gauteng, South Africa

Chukwuma Okolie, Mawande Ngidi, Zukisa Sogoni, Lesego Tshuwa and Samukele Ngema (South Africa)

Key words: Geoinformation/GI; GIM; GSDD; Implementation of plans; Spatial planning; Geospatial data; population projection; population growth; statistics; machine learning.

SUMMARY

Population data and population projections are crucial for planning all around the world, especially in urban areas as cities are expected to ensure basic service provision to communities. This is especially important for African cities which are experiencing rapid population growth and experiencing significant resource constraints. Despite the importance of accurate population projections, there is no consensus on the best approach for its modelling. While statistical time-series models have been widely used in demographic projections, machine learning approaches have shown promise in recent studies. The relative performance of both techniques is still an ongoing debate in the literature. We address this research gap by comparing the accuracy of ten (10) statistical and machine learning models (with default hyperparameters) in five municipalities of Gauteng Province, South Africa (Johannesburg, Tshwane, Ekurhuleni, Sedibeng and West Rand). The models were trained and tested using authoritative mid-year population estimates from 2002 to 2024, and then future projections were generated for 2025 – 2030. The best performing models were Random Forest (Johannesburg), ETS (Tshwane), Extra Trees (Ekurhuleni), and Exponential Smoothing (Sedibeng and West Rand). Our findings show that trends in historical population data are reasonably well captured by both approaches, with a few exceptions. There is no best model for all scenarios, and machine learning projections should be treated as complementary, scenario-based estimates rather than deterministic forecasts in long-term demographic analysis. Further application of these modelling approaches has the potential to stir up the urban governance paradigm by giving wider access to data-driven decision making in local municipalities.

Statistical and Machine Learning Approaches for Municipal Population Projection in Gauteng, South Africa (13900)
Chukwuma Okolie, Mawande Ngidi, Zukisa Sogoni, Lesego Tshuwa and Samukele Ngema (South Africa)

FIG Congress 2026
The Future We Want - The SDGs and Beyond
Cape Town, South Africa, 24–29 May 2026