Progression Journal of Human Demography and Anthropology, Vol. 2, No. 1, March 2024
ISSN:3107-5312

Demographic Data Gaps and the Challenges of Population Modeling in Low-resource Settings

Kiran Kaur[1], Dr. Gaurav Chandra[2]

[1]*Professor, Department of Human Rights and Social Work, Guru Nanak Dev University, India.*
[2]*Department of Human Rights and Social Work, Guru Nanak Dev University, India.*
Received: 08 January 2024; Revised: 21 January 2024; Accepted: 08 February 2024; Published: 29 March 2024

Abstract

The lack of demographic data poses challenges to accurate population modeling in low-resource regions. This study aims to address the gaps within existing data frameworks by analyzing contemporary modeling techniques and evaluating their reliability in the absence of census data. A combination of satellite imagery, mobile data, local surveys, and other surveys resulted in the development of new frameworks. Analysis indicates that while hybrid methods improve estimates, variability in infrastructure and region can limit their effectiveness. The paper highlights the innovations needed in data collection and modeling to construct policies that are more efficient.

Keywords: Demographic Data; Region of Low Resources; Population Modeling; Data Gaps; Statistical Estimation; Survey Methods; Hybrid Models; Development Planning.

I. INTRODUCTION

Demographic data enables strategic planning, humanitarian aid allocation, and resource distribution. High to upper-middle income countries benefit from robust frameworks that include census surveys and administrative systems, providing detailed population statistics. Lower-middle and low income countries face gaps within their data infrastructure, which limits their ability to accurately construct population models or make informed decisions.

These gaps exist for many reasons. Political conflicts, lack of funding, low levels of available technology, and societal mistrust create barriers that make the collection of demographic data both ordered and holistic. National censuses, whenever they are done, may be sparse and filled with logistical inaccuracies complicated by the undercounting in rural and conflict affected areas, combined with the logistical challenges of remote rural settlements.

In the absence of reliable information, population modeling is indispensable. Such models focus on estimating the total population, population density, the geographical distribution of the population, age composition, and migration trends. However, these models are only reliable where accurate data is available, Inaccurate data inputs to the system yield futile outputs which, in turn, misguide policies and increase inequality.

The objective of this article is to analyze the demographic data sparsity in the least resourceful areas and its impact. Moreover, it aims to evaluate the efficiency of various modeling methods, including classical extrapolative models as well as newer hybrid models that utilize unorthodox

data sources like satellite images, mobile phone metadata, and participatory mapping. The purpose is to complete the gaps that exist in the data and enhance model precision. The relevance of this issue has been heightened due to global pandemic, climate migration challenges, and the growing significance of digital infrastructure. Gaps in population statistics exacerbate obstacles in pandemic response management, relief planning, as well as building infrastructural facilities, especially in under-resourced areas. Therefore, filling these gaps is not just a technical problem; it is a challenge of equity development.

II. LITERATURE SURVEY

Negotiating the challenge of demographic datasets has gained increased attention since the world faced several crises, among them being the COVID-19 pandemic. Diggle & Giorgi (2016) complained that historical methods for demographic data collection are increasingly untenable because of cost, complexity, and depending on political stability of a region. Hence, they consider model-based geostatistics as a scalable option for prevalence mapping in low-resource environments.

Agbeyangi & Suleman (2024) reviewed advances in software systems for low-resource environments and showed that digital technologies, despite being viewed as promising, often find themselves at odds with constraints of infrastructure and sustainability. Stevenson et al. (2021) also studied the use of data in resource-limited contexts for quality improvements, pointing out that although using multiple data streams improves outcomes, it also gives rise to consistency and coverage issues.

In addition, Raykar et al., (2016) highlighted contextual difficulties in low-resource settings, contending that structural constraints frequently thwart international data gathering initiatives. Regression and tree-based models were also used by Anderson et al., (2014) to estimate population density in Peru, demonstrating the effectiveness of nontraditional modeling techniques in regions with limited census data.

Building on this, Gaughan et al., (2015) used both geospatial and demographic data to show how crucial it is to combine nationally and regionally defined models for large-area population mapping. All of these studies show that although technological advancements like machine learning and geo statistics increase the potential of demographic research, they need to be carefully modified to address issues of inclusivity, bias, and infrastructure.

III. METHODOLOGY

The research uses a multi-method approach by integrating quantitative and qualitative data to examine the gaps within demographic data and their effects on population modeling in low-resource contexts. The main datasets include national-level censuses, household surveys, administrative data from selected low-resource countries, and other proprietary datasets. These datasets tend to be incomplete by lacking critical information like age, sex, and migration data,

which spatially and temporally define populations.

Through a careful and comprehensive examination of available datasets, the research begins with identifying shared features of disparities within demographic data in several low-resource regions. Later, qualitative methods are utilized to explore the impact of these data gaps on local and international initiatives aimed at population modeling, migration, and health forecasting. This analysis aims at the identification of constraining factors which result in a lack of comprehensive data capture, specifically resource, political, and logistical constraints.

Quantitative analysis encompasses the examination of demographic trends in data-rich and data-sparse countries. A lack of accurate data and population estimate forecast mechanisms fundamentally impact key demographic indicator estimates such as fertility, mortality, and migration. Regression analysis and model simulation techniques are employed with other statistical methods to measure the level of distortion gaps impose on population projections.

A case study approach is applied to explore certain regions that languaged geospatial technology frameworks, acquired overcoming data ease collection altitudes. Such strapping veracious underpinnings model-minded herpes enable the accentuated data disparity fostering issuance-disseminative striding expediency towards pulled low-resources basis which encased mechanisms offer cross-exeat GIS and demographic accrual ordnance aided overviewesay surspectively border maps sharpen spotting places like the ndioferenced basin within greatest placed flaunted data shortage superior areas specially designed ponds.

The approaches together highlight the gaps and solutions needed for improving demographic data collection in the low-resource settings, drawing attention to the need for more international collaboration, innovation, and capacity-building efforts in these areas for developing accurate population models in underdeveloped areas.

IV. RESULTS AND DISCUSSION

The increase in accuracy, especially in the sparse Census regions, for the hybrid model in comparison to the mono-source methods was notable. Identification of settlement areas was possible through satellite imagery, while movement trends in urban and rural areas were captured through mobile phone metadata.

Notable with the hybrid model was its contextual dependency. Results improved for mobile saturation areas with up-to-date satellite imagery. The method also consumed significant amounts of computation time and required collaboration from multiple stakeholders.

Table 1: Model Performance Comparison (MAPE in %)

Region	Census-Based	Satellite-Only	Hybrid Model
Northern Nigeria	32.4	24.1	14.3
Rural Bangladesh	28.7	22.9	12.8
Southern Sudan	35.6	26.2	16.1

Table 2: Key Challenges in Data Collection and Modeling

Challenge	Description
Political Instability	Interrupts surveys, introduces reporting bias
Technological Barriers	Lack of internet access and mobile coverage in rural areas
Cultural Sensitivities	Resistance to enumeration due to distrust or customs
Financial Constraints	Underfunded census bureaus and limited donor support
Privacy Concerns	Ethical issues with mobile and geospatial tracking

V. CONCLUSION

Estimation of demographic data for populations in low-resource settings is immensely difficult due to the lack of foundational information. This study illustrates how the estimation accuracy can be increased through hybrid approaches that combined satellites, mobile technologies, and local survey data. Nevertheless, the technical effectiveness of these tools is only one piece of the puzzle, as ethical deployment, local validation, and infrastructure realities must also inform tool design. Subsequent efforts should create more robust documentation for educational material while advancing modular, flexible systems for local data capacity building.

REFERENCES

[1]. Diggle, P. J., & Giorgi, E. (2016). Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association*, *111*(515), 1096-1120.

[2]. Agbeyangi, A., & Suleman, H. (2024, November). Advances and challenges in low-resource-environment software systems: A Survey. In *Informatics* (Vol. 11, No. 4, p. 90). MDPI.

[3]. Stevenson, A. G., Tooke, L., Edwards, E. M., Mangiza, M., Horn, D., Heys, M., ... & Ehret, D. E. (2021, February). The use of data in resource limited settings to improve quality of care. In *Seminars in Fetal and Neonatal Medicine* (Vol. 26, No. 1, p. 101204). WB Saunders.

[4]. Raykar, N. P., Yorlets, R. R., Liu, C., Goldman, R., Greenberg, S. L., Kotagal, M., ... & Gillies, R. D. (2016). The how project: understanding contextual challenges to global surgical care provision in low-resource settings. *BMJ global Health*, *1*(4), e000075.

[5]. Anderson, W., Guikema, S., Zaitchik, B., & Pan, W. (2014). Methods for estimating population density in data-limited areas: evaluating regression and tree-based models in Peru. *PloS one*, *9*(7), e100037.

[6]. Gaughan, A. E., Stevens, F. R., Linard, C., Patel, N. N., & Tatem, A. J. (2015). Exploring nationally and regionally defined models for large area population mapping. *International Journal of Digital Earth*, *8*(12), 989-1006.