# CASA dissertation: Literature Review

Student Number: 22186878

Word count (excluding tables and figures): aim for 1200-1500 words.

## Table of Contents

### Introduction

1. World Bank Water Project: Sri Lanka
2. Justification of extending project to India
3. Research Questions

### Literature Review

1. Agricultural Dependent Populations
   1. How is this defined in the literature?
   2. Why is this term useful? What is the value of understanding this population?
2. Indian context
   1. States/districts/villages formation
   2. National-level agricultural dependence, statistics
   3. Available census data; opportunities and limitations
3. Spatial disaggregation
   1. What does this term mean?
   2. World gridded population
   3. Examples of spatial disaggregation methods; their strengths and weaknesses.
   4. Binary dasymetric mapping: the main method to be used
   5. Considerations: MAUP, etc.
4. Scaling Geocomputation \*\*TBD if this section will be relevant\*\*
   1. Challenges/Limitations of computation at large spatial scales
   2. Methods for overcoming this

## Introduction

Understanding where people live, and the social and economic characteristics of those populations, is core to providing adequate, efficient, and targeted services and investment.

Justification

This study is a novel addition to the field as it extends upon existing methodologies used to estimate total population and applies this to the estimation of the agricultural dependent population. Additionally, the case study of India is designed to assess feasibility and performance at a large spatial scale, comparative to partner research testing proof-of-concept in districts of Sri Lanka (unpublished). Understanding particularly the distribution of agricultural population in a region will provide a more accurate estimate of local demand on water resources.

Stretch out ‘scale’ component further; key aspect of study

Text

## Literature Review

### Preamble

Introduction for the literature review; summarise the findings and topics that will be explored – ADP, India, and Spatial Disaggregation Methodology.

Text

### Agricultural Dependent Populations

Text

### Indian Context

Paragraph topic: Intro; What is the current situation in India for ADP?

India, the subject of this study, is one of the world’s largest countries by area, the third-largest economy, and is expected to become the most populous country by the end of 2023 (United Nations in India, 2022).

Paragraph topic: Administrative formation.

India is divided into 28 states and 8 union territories, each of which are further subdivided into districts and smaller administrative divisions.

Paragraph topic: census data – what is available? Opportunities/Limitations?

The most recent Census of India was conducted in 2011.

### Spatial Disaggregation

Paragraph topic: Define spatial disaggregation; Why is it important?

Spatial disaggregation is a broad term which applies to the process of transforming data from a set of source zones into target zones, such as a raster grid, at a finer level of spatial resolution. There is considerable interest in the process across both academic literature and in policy, particularly applied to estimating population at fine spatial scales, as this has important implications for service planning and delivery (Deichmann, 1996), disaster preparation and response (Schneiderbauer and Ehrlich, 2005), monitoring international development goals (United Nations, 2022) and the implementation of public health interventions (Viel and Tran, 2009), among others.

Topic: World gridded population estimates; history, status

On a global scale, the spatial disaggregation of administrative census data has been used to develop world gridded population estimates, providing regularly sized comparable population estimates across regions. Early iterations of this, such as the Gridded Population of the World (GPW) version 1 (Tobler *et al.*, 1997), have inspired a variety of contemporary global grid models, each utilising a specialised methodology and with particular strengths and limitations, bolstered by advances in computational power and the availability of high quality census and earth observation data (Wardrop *et al.*, 2018). [Discuss models from Leyk review]. Table 1 provides a summary of current global grid models and their key characteristics.

Table : Selected World Population Grid Datasets, adapted from Leyk et al. (2019)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Source** | **Method** | **Spatial Resolution** | **Ancillary data layers** |
| Gridded Population of the World (GPW) | CIESINa | Areal weighting | 1km | Water bodies |
| Global Human Settlement Layer – Population (GHS-POP) | JRCb and CIESINa | Dasymetric | 250m | Built structures |
| WorldPop | University of Southampton | Statistical/ Dasymetric | 100m | Roads, Land cover, Built structures, Urban areas, Night-time lights, Infrastructure, Climate, Topography, Elevation, Water bodies |
| LandScan Global | ORNLc | Smart interpolation | 30 arcsec | Roads, Land cover, Built structures, Urban areas, Infrastructure, Climate, Topography, Elevation, Water bodies |
| World Population Estimate (WPE) | Esri | Dasymetric redistribution | 150m | Roads, Land cover, Urban areas, Water bodies |

a Centre for International Earth Science Information Network; b Joint Research Centre of the European Commission; c Oak Ridge National Laboratory.

Topic: Spatial disaggregation continued – dasymetric, pycnophylactic

The simplest

Topic: Spatial disaggregation methods: Top down vs Bottom up

Alternative to top-down spatial disaggregation, ‘bottom-up’ approaches for small area estimation can also be used to produce gridded population estimates. These methods are designed for estimating population distribution in the absence of high-level source data such as a national or regional census, or when such data is out-of-date or known to be inaccurate (Wardrop *et al.*, 2018). However, bottom-up estimation requires the completion of tailored geo-located field surveys, and are generally viewed as complementary to traditional census enumeration in specific data-poor settings (Leyk *et al.*, 2019).

Topic: Gaps in spatial disaggregation literature

Despite extensive research and methodological development in the field of population disaggregation, there is a considerable gap in studies that extend these methods to additional demographic or socioeconomic characteristics beyond population count or density. Szarka and Biljecki (2022) present a method to predict age structure of disaggregated populations, combining a set of machine learning models with real estate data for a case study area in Singapore.

Topic: Considerations – MAUP; resolution used; temporal differences between census and ancillary data

Like in most aspects of spatial analysis, the spatial resolution or scale used in a spatial disaggregation can markedly influence the outcomes and interpretation. This challenge is defined as the modifiable areal unit problem (MAUP), where the same analysis performed on data aggregated at different levels or in different zones will produce different results (Wong, 2009).

## References

Deichmann, U. (1996) *A Review of Spatial Population Database Design and Modeling*. Santa Barbara, CA: National Centre for Geographic Information and Analysis. Available at: https://escholarship.org/uc/item/6g190671 (Accessed: 28 February 2023).

Leyk, S. *et al.* (2019) ‘The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use’, *Earth System Science Data*, 11(3), pp. 1385–1409. Available at: https://doi.org/10.5194/essd-11-1385-2019.

Schneiderbauer, S. and Ehrlich, D. (2005) ‘Population Density Estimations for Disaster Management: Case Study Rural Zimbabwe’, in P. van Oosterom, S. Zlatanova, and E.M. Fendel (eds) *Geo-information for Disaster Management*. Berlin, Heidelberg: Springer, pp. 901–921. Available at: https://doi.org/10.1007/3-540-27468-5\_64.

Szarka, N. and Biljecki, F. (2022) ‘Population estimation beyond counts—Inferring demographic characteristics’, *PLOS ONE*, 17(4), p. e0266484. Available at: https://doi.org/10.1371/journal.pone.0266484.

Tobler, W. *et al.* (1997) ‘World population in a grid of spherical quadrilaterals’, *International Journal of Population Geography*, 3(3), pp. 203–225. Available at: https://doi.org/10.1002/(SICI)1099-1220(199709)3:3<203::AID-IJPG68>3.0.CO;2-C.

United Nations (2022) *The Sustainable Development Goals Report 2022*. New York, NY: United Nations. Available at: https://unstats.un.org/sdgs/report/2022/.

United Nations in India (2022) *UN India Annual Report 2021*. New Delhi, India. Available at: https://india.un.org/en/195240-un-india-annual-report-2021, https://india.un.org/en/195240-un-india-annual-report-2021 (Accessed: 30 May 2023).

Viel, J.-F. and Tran, A. (2009) ‘Estimating Denominators: Satellite-Based Population Estimates at a Fine Spatial Resolution in a European Urban Area’, *Epidemiology*, 20(2), pp. 214–222.

Wardrop, N.A. *et al.* (2018) ‘Spatially disaggregated population estimates in the absence of national population and housing census data’, *Proceedings of the National Academy of Sciences*, 115(14), pp. 3529–3537. Available at: https://doi.org/10.1073/pnas.1715305115.

Wong, D.W. (2009) ‘Modifiable Areal Unit Problem’, in R. Kitchin and N. Thrift (eds) *International Encyclopedia of Human Geography*. Oxford: Elsevier, pp. 169–174. Available at: https://doi.org/10.1016/B978-008044910-4.00475-2.