Enabling the Ghanaian National Statistics Office to produce official statistics derived from mobile operator data with applications across the public sector

Romain Goldenberg[1]*, Thomas Smallwood[1], Cathy Riley[1], James Harrison[1], Richard Attandoh[1], Omar Seidu[2] and Veronique Lefebvre[1]

[1] Flowminder Foundation, [2] Ghana Statistical Service

* Corresponding author. email: romain.goldenberg@flowminder.org

Abstract

Human mobility has important implications for decision-makers working across a broad range of sectors, including those working in government. Anonymised and aggregated mobile operator data provide near-real time insights into human mobility at high spatial and temporal resolution across a whole country and, after correcting for representation biases using survey data, can support decision-makers across a broad range of applications. However, there are a number of barriers to the use of mobile operator data by national statistics offices (NSOs) and other government agencies, including gaining access to data and the technical capacity to effectively process and analyse the data while preserving the individual privacy of subscribers.

To integrate mobile operator data into official statistics, and support branches of the government in using novel data sources for improved decision-making, Ghana Statistical Service (GSS), Vodafone Ghana and Flowminder entered a unique and long-standing partnership - the Data for Good partnership - for the wellbeing of all in Ghana. Within this partnership, Flowminder has been strengthening the capacity of GSS to process and analyse mobile phone data to produce routine mobility statistics and support other parts of the Government, independently. Flowminder is using its tools and expertise in the secure and privacy-preserving processing and analysis of call detail record (CDR) data to to set up a data and analytical pipeline for GSS to eventually run independently, receiving daily pseudonymised CDRs and extracting and bias-correcting mobility indicators of migration and travel. These indicators will then be integrated into decision-making in ministries, departments and agencies (MDAs) across the Ghanaian government.

In order to support GSS and other MDAs, the Data for Good partnership has focussed on delivering CDR-derived indicators for several applications including official mobility statistics, dynamic public health indicators, and dynamic disaster risk mapping which we describe below.

(1) As a result of this joint effort, a standardised mobility product has been created, primarily aimed at comprehending overall mobility patterns in the country and for key administrative areas of interest. It is intended to be generated at regular intervals (6 month periods) and in an automated manner, to ensure consistent and timely updates on a collection of recent mobility trends, such as migrations or pendular movements (Fig. 1).

(2) Flowminder and GSS are also collaborating with Ghana Health Service (GHS) and the Ministry of Health to identify public health metrics which may be improved by incorporating dynamic estimates of population density. Many public health metrics, including metrics of infectious disease such as disease incidence and prevalence, are calculated using the population of an area as a denominator (i.e. are calculated per capita). However, the traditional data sources most often used to estimate population size, such as censuses or surveys, only provide a static snapshot of the population, which ignores seasonal variations and migration trends, and therefore may be inaccurate and outdated. Flowminder is currently producing four experimental health metrics, covering infectious and non-infectious disease incidence, vaccination coverage and resource allocation, with monthly district-level population estimates.

(3) The National Disaster Management Organisation (NADMO) has requested support from Flowminder and GSS to help improve disaster preparedness by incorporating dynamic population estimates into risk analyses. Exposure, or the number of people in an area who may be exposed to a given disaster, is an important dimension of risk. However, exposure can vary as people move in or out of an area over both the short- and long-term (e.g. commuting and migration, respectively). Flowminder has produced preliminary dynamic flooding risk maps capturing the change in district-level flooding risk with hourly resolution for different days of the week, covering the whole of Ghana. The analyses demonstrate that CDRs can effectively be used to estimate the large variations in the number of people who may be exposed to flooding, and therefore the risk, in some districts depending on the time of day and day of the week due notably to commuting movements (Fig. 2). This variation in risk can have important implications for disaster preparedness. Future work will aim to ensure CDR-derived hourly metrics are corrected for biases

related to phone usage variations and for representation biases so that the resulting risk maps can be used operationally.

The Data for Good partnership is a ground-breaking example of a sustainable system for the use of CDR data by a National Statistics Office and other MDAs to support decision-makers across multiple sectors in a middle-income country. Furthermore, partners are continuing to develop the applications described here and Flowminder is also testing additional applications, such as combining CDR data and other geospatial data for predictive modelling of poverty indicators.

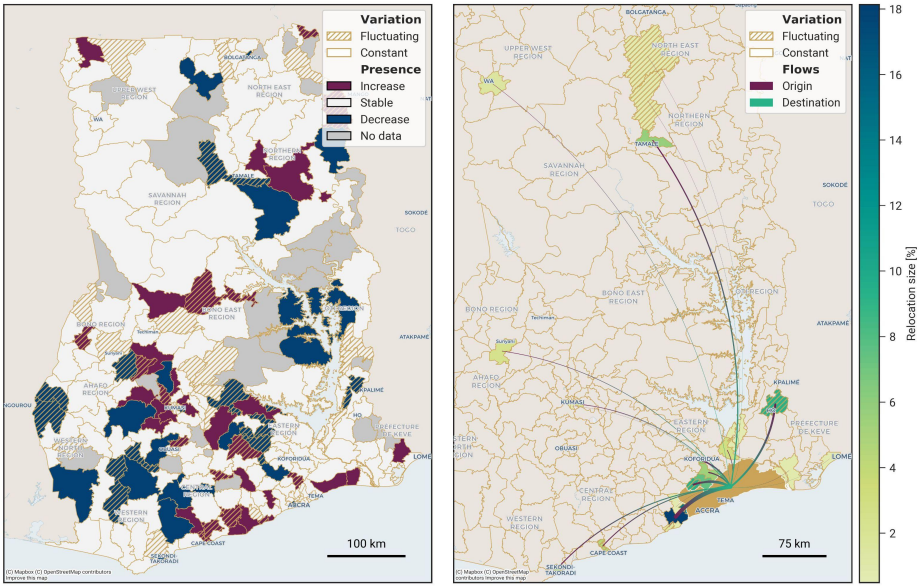

Figure 1. *Trends in residents counts for each district in Ghana (resulting from internal migration in the country) are shown on the left-hand side. On the right-hand side is an example of such relocations, showing migrations to the Greater Accra region. Both figures are averages over the period 2020 to 2022.*

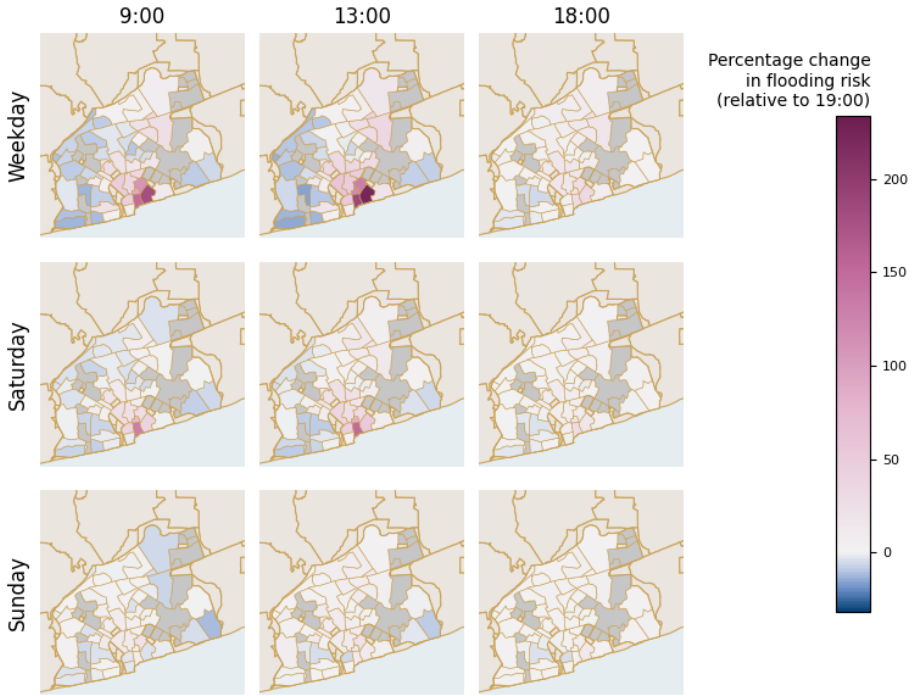

Figure 2. *Percentage change in flooding risk in central Accra at different times of day and different days of the week, relative to the risk at 19:00.*