Geodata-driven approaches to financial inclusion – Addressing the challenge of proximity

Fibæk, Casper Samsø; Laufer, Hanna; Keßler, Carsten; Jokar Arsanjani, Jamal

Published in:
International Journal of Applied Earth Observation and Geoinformation

DOI (link to publication from Publisher):
10.1016/j.jag.2021.102325

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

International Journal of Applied Earth Observations and Geoinformation 99 (2021) 102325

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Geodata-driven approaches to financial inclusion – Addressing the challenge of proximity

Casper Samsø Fibæk [a,b,*], Hanna Laufer [c], Carsten Keßler [a], Jamal Jokar Arsanjani [a]

[a] Department of Planning, Aalborg University, Copenhagen, A.C. Meyers Vænge 15, 2450 Copenhagen SV, Denmark
[b] NIRAS Consulting, Sortemosevej 19, 3450 Allerød, Denmark
[c] Oxford Policy Management, Clarendon House, Level 3, 52 Cornmarket Street, Oxford OX1 3HJ, United Kingdom

ARTICLE INFO

Keywords:
Financial inclusion
Sustainable development goals
Decision support system
Earth observation
Copernicus programme

ABSTRACT

Financial Inclusion is, in many ways, a spatial planning issue: Where do financial institutions provide services, how far do customers travel to access mobile money, which services are available where and how is agent cash-flow handled? Utilising geodata can contribute significantly to measuring financial access and thus assist in improving Financial Inclusion by expanding the reach of services and locating areas of economic exclusion. This study presents a new Spatial Decision Support System, with a frontend embedded directly within a spreadsheet interface, that enables measuring and planning financial access through geospatial analysis and Earth Observation derived products. The purpose is to complement existing Financial Inclusion measures, which rely significantly on large-scale representative household surveys to quantify financial access and opportunity to proxy quality of inclusion. The Decision Support System relies on Earth Observation and Public Participatory GIS, which enables a decoupling from the census cycle and global reach. Our findings indicate that a geospatial approach to measuring and making decisions regarding the location of financial access points can positively affect both tracking and delivering Financial Inclusion and reducing the urban–rural service cliff-edge. Our proposed geospatial methodology is useful for decision-makers in two ways: a) It allows the measurement of the large-scale geospatial reach of financial services – useful for decision-makers, planners, politicians, national statistical offices, and NGOs in charge of tracking progress towards the Sustainable Development Goals. b) It helps with planning and optimising services for local financial entities such as mobile money agents, brick and mortar bank branches, and less formal saving mechanisms such as saving clubs. The Spatial Decision Support System is currently used by several Financial Service Providers in Ghana and undergoing implementation for one in North-western Tanzania.

1. Introduction

Earth Observation (EO), as a data source and discipline, contributes in many significant ways to the realisation of the 2030 Sustainable Development Goals (SDGs) (Andries et al., 2019). One of the less explored areas of applying EO is Financial Inclusion (FI), a theme for reaching several SDGs (UK Space Agency, 2020). Eight of the seventeen goals feature elements of FI. Goal one, *end poverty*, strives to ensure equal rights to economic resources and access to financial services (United Nations, 2017). EO enables global and up to data mapping of urban structures, infrastructure, population, pollution, and much more. Combining EO's temporal and spatial resolution with high-quality local geospatial data sources makes it possible to create a highly scalable Spatial Decision Support System (SDSS) (see Table 1).

Access to and sustained use of financial services is, among other factors, a function of social and physical distance of existing or potential customers to branches and agents of Financial Services Providers (FSPs) (Forster et al., 2013). For the efforts towards FI to be successful, financial access must encompass rural and remote areas and under-served urban communities (Mahendra, 2006). To reach these areas, FSPs must find sustainable business models for expanding catchment areas (Peachey & Mutiso, 2019). The creation of business models that support FI is possible; when FSPs invest in the adoption of new technologies, innovation, and data-driven business models, such as applying geospatial data to support route to market models - getting financial services to the end-users - and agency banking (Demirguc-Kunt et al., 2018).

* Corresponding author at: Lyneborggade 8, 2300 Copenhagen S, Denmark.
E-mail addresses: cfi@plan.aau.dk, cfi@niras.dk (C. Samsø Fibæk).

https://doi.org/10.1016/j.jag.2021.102325
Received 14 December 2020; Received in revised form 11 March 2021; Accepted 13 March 2021
Available online 30 March 2021

Table 1
An overview of the location of participating FSPs.

Country	Amount of FSPs	Employees in total
Ghana	4	6
Tanzania	5	9

Agency banking is a function whereby a banking institution vets and hires a third party to offer financial services on their behalf. Services carried about by agents can include cash withdrawal and deposits, bills payments, and mobile banking services (Josephat, 2016). Optimising and tracking the location of mobile money agents is important to improve FI.

The collection of official statistical data occurs at a temporal and spatial resolution that is generally unsuitable for day-to-day decision-making tools for FSPs' with agency banking strategies or customer profiling. In Sub-Saharan Africa, there is often insufficient authoritative national data available, leading to low data resolution, inconsistent and spatially and temporally incomplete data (Steele et al., 2017). The data insufficiency is partially due to the reliance on housing and population census data, which by convention is conducted decennially (UN-Stats, 2008). A high temporal resolution is vital for accurate FI indicators, as urbanisation and demographic changes are rapidly altering land-use patterns (Shi, Jiang, and Yao, 2018).

By supplementing authoritative national data with EO data and public open data, it is possible to reduce the data shortfall on FI measurements. Insufficient data can prevent policymakers, regulators and FSPs from addressing FI in an informed manner.

The creation of a new geodata-driven SDSS serves two distinct purposes:

a) Enabling National Statistical Offices (NSOs) and Non-Governmental Organisations' (NGOs) tracking of FI to address intercensal questions such as: How much of the population has physical access to financial services? Where are the under-served communities located, and what characterises these communities?
b) Facilitating agency banking for FSPs by easing customer segmentation, network optimisation, and locating new business opportunities, thereby supporting route to market and business case discovery, which is necessary to ensure economic incentives for the FSPs to reduce financial exclusion.

This paper presents a study on how joining the FI and EO domains makes FI an addressable challenge. The presented system is the first web Application Programming Interface (API), and spreadsheet-based SDSS explicitly created for supporting FI to the best of the authors' knowledge.

Most FSPs will add additional proprietary data to their version of the SDSS that is not available to the public or other FSPs. Despite a significant amount of this additional data being proprietary, the free data and the SDSS are beneficial for third parties and non-technical and non-financial users. Urban planners and climate researchers can use the SDSS and the available data to improve their modelling.

The case study area is Ghana – West Africa, shown in Fig. 1. The country was selected as it contains a mixture of settlement types and landscapes. Its climate is tropical, based on the Koppen-Geiger climate classification; monsoon in the South-West, and savannah elsewhere (Beck et al., 2018). The diverse landscape provides a good test-area for Land Cover and Land Use classification methods as the mixed terrain offer the opportunity to test classification methods in multiple climate zones. Ghana consists of four geographical plains: The Coastal Plain in the south, The home of the Ashanti kingdom, Ashanti-Kwahu with extensive forests in the mid-west, the Volta-basin in the east and the northern savannah plains. Besides having both arid and vegetated climates, Ghana is a mid-sized African country with the fastest-growing mobile money market in Africa (Geiger et al., 2019) and at the forefront of mobile banking.

The remainder of this paper is structured as follows: Section 2 presents background information illustrating the study's relevancy. Section 3 outlines the data sources used in the decision support system and presents the methodology behind the Earth Observation-based classification of settlement types in the study area and, finally, the development of the frontend of the SDSS. Section 4 presents the results and feedback taken from the rollout and testing in Ghana, followed by a discussion and conclusions in Section 5.

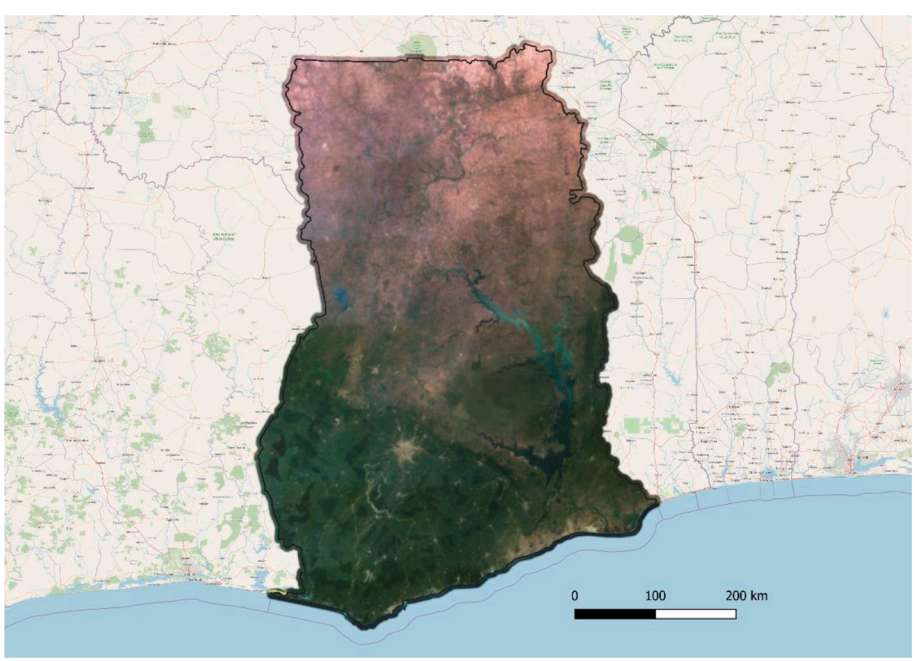


Fig. 1. A mosaic of Sentinel 2 images taken in Q1 2020 covering Ghana.

2. Background

Financial Inclusion (FI) concerns equal access to financial services that are fit for purpose. In practice, this means ensuring access to savings mechanisms and other financial products at an affordable cost, particularly to previously under-served groups, such as low-income households and rural communities. Aspects of FI are measurable in many distinct ways, such as the physical distance to agents, the number of formalised saving clubs, the number of savings accounts or local branches of financial services providers, and financial literacy, use and capability surveys. Improving FI is, in part, the process of extending the reach of the formal financial sector and reducing social exclusion within the economy (Varghese and Viswanathan, 2018).

The multi-faceted, complex nature of FI makes it challenging to measure in a meaningful fashion. Counting the number of households or individuals who have opened a savings account at an aggregate level is insufficient. Up to date data will often only be available to the data collecting institutions, which are unwilling to share sensitive data. If the data is collected through surveys or during censuses, the spatial and temporal resolution is limited. There can be large, localised differences in access to financial services within a community or city, and detaching the analysis from aggregate areas enable a targeted approach to FI. An underserved community does not have physical access to financial services. In Ghana, mobile banking is prevalent, and mobile money agents and cell phone coverage are important indicators of Inclusion.

We propose a complementary manner of measuring and qualifying FI through comprehensive mapping of population density, settlements types, cell phone coverage and distance to financial services, mobile money agents, or brick and mortar institutions. This study presents a way of adding these and other relevant datasets to the traditional ways of measuring Inclusion, such as counting savings accounts.

2.1. Proximity analysis and financial inclusion

There is an increasing interest in the geospatial aspects of FI (UK Space Agency, 2020). One of the key takeaways is a clear gap between financial services made available for urban, peri-urban, and rural citizens; this is referred to as the 'proximity gap' or the 'proximity challenge'. The availability of formal financial services drops off quickly with distance from urbanised cores (Peachey & Mutiso, 2019).

Studies by the World Bank (Forster et al., 2013) reveal that it is important to take social and physical distance into account when measuring proximity. People are less likely to cross social barriers to access financial services. It is, therefore, necessary to further qualify physical distances with contextual information, for example, the socioeconomic profiles of neighbourhoods that access points are in, to get an accurate sense of the level of Inclusion.

Forster et al. (2013) used 2010 census data and a mapping of access points for financial services across Kenya, Tanzania, and Uganda to investigate the relationship between financial access and use. They found that a poor rural family is willing to accept the costs of walking half a day to collect a monthly transfer 5 – 10 km away if that money is a significant amount of weekly household spending. Depositing a spare dollar into the formal financial system is uneconomic if it is necessary to walk for more than an hour, despite a waiver of fees. Based on a 5 km radius, 60–75% of peri-urban/rural Kenyans and Ugandans and 20% for Tanzanians had access to local person-to-person money transfers. Based on a 2 km radius, local reach is reduced to 25 per cent for Uganda, 50 per cent for Kenya, and less than 10 per cent for Tanzanian rural savers (Forster et al., 2013; Peachey & Mutiso, 2019). These previous studies relied on circular buffers applied to manually estimated urban cores.

2.2. Mapping and remote sensing

While Geographic Information Systems have previously been used in the financial sector to estimate FI, usually at the administrative boundary level, there is no formal spatial definition of FI.

Earth Observation (EO) is the act of collecting spatial data on planets by using remote sensing technologies. In this paper, EO refers to data collected from space-borne optical and active sensors. EO data is integrated into the SDSS, as it enables global reach and timely open data. It is generally a challenging data source to process, but experiences from workshops show that the presented tooling processes and presents the EO data, so that little to no prior knowledge of EO is necessary to make meaningful decisions. EO makes it possible to do global maps on diverse topics such as settlement structures, population estimates, climate tracking, poverty and pollution (Paganini et al., 2018). The SDSS includes EO derived maps of settlement structures, population estimates and night lights.

Besides measuring proximity and segmenting settlement types, it is possible to apply EO and mixed data approaches to infer socioeconomic status. The World Bank has been researching methods of estimating income using satellite imagery and cell tower data (Heitman & Buri, 2019). Improving these methods in data-poor areas could further FI by providing remote credit scoring and thereby ease access to loans. Such an approach would be substantially dependent on local tenure security enforcement, as highlighted by Hernando de Soto (De Soto, 2001).

Geospatial mapping is especially beneficial for FSPs working with agency banking. Data collection allows optimisation of agent networks by location, profitability, and reach. Global data sources that enable these types of analysis are available through the WorldPop project headed by the University of Southampton (AJ. Tatem, 2017). The SDSS presented in this study incorporates the WorldPop global demographics data. Combining satellite data, socioeconomic data, and public participatory data enables financial service providers and other parties to perform customer and market segmentation.

3. Design and methodology

This section describes the data sources and methodology behind creating the segmentation of settlement types provided in the Spatial Decision Support System (SDSS), the spreadsheet-based frontend and the web-based Application Programming Interface (API) that underpins it.

The general design of the SDSS builds upon continuous feedback from 15 technical employees, representing nine different FSPs in Ghana and Tanzania, about their needs and wants for a system to support Financial Inclusion and the proximity challenges. An overview of the participating Financial Service Providers (FSPs) location is shown in Fig. 1. During the interviews, it became evident that all the participants were confident using spreadsheet technology to facilitate decision making and that none of the participants had prior experience with GIS software. The different FSPs had different proximity challenges, unique to their services, that they were trying to solve that each tie into Financial Inclusion (FI). These issues include optimising agent networks, saving group collections, brick and mortar bank locations, connecting marketplaces and locating new business opportunities (*Proximity Meeting, Dar es Salaam, 2019)*.

After a presentation on the possibilities of satellite data and other open data sets, the FSPs highlighted some of the questions they would want the decision support system to assist in answering:

- How many people live where?
- What is the current financial infrastructure?
- Where are different types of neighbourhoods located?
- What is the socioeconomic status of the neighbourhoods?
- How long does it take to go to the nearest market or FSP?
- What is the optimal placement of new services and agents?
- Where are the underserved communities?
- What is the cell phone coverage like in this area?
- What are the conditions of the roads in this area?

The following sections will describe the SDSS and consider the above questions to make the SDSS useful in identifying business cases and furthering FI.

3.1. Architecture

The design of the SDSS enables FSPs to integrate the system directly into their proprietary applications and existing workflows. If they do not wish to - or do not have access to - an existing application, they can use the supplied spreadsheet-based interface as their frontend. The system relies on an open REST web API accessible to any applications over the internet by providing the correct credentials (L. Richardson & S. Ruby, 2008). Fig. 2 outlines the SDSS architecture. By designing the architecture around an open API, it enables the use of different internal frontends. As such, the FSPs can develop proprietary frontends such as a mobile app that uses some or all the API functions to gather information about travel times to nearest services, how many people are in the area and create area profiles indicating the composition of the general catchment area. Separating the SDSS into frontend(s), a database and an accompanying API is a common and well-tested architecture for decision support systems and web-based applications (R. Fielding & R. Taylor, 2002).

EO and geographical data processing is done locally before being uploaded to a cloud service provider where the API, database, and static files are hosted in separate services.

3.2. Data sources

The primary data sources come from the European Commission's and European Space Agency's Copernicus Programme via the Sentinel Satellites shown in Table 2. This study focuses on the use of free and open data sources and technologies. Besides Sentinel data, OpenStreetMap (OSM) is used. OSM is a collaborative effort to map the world, where the data collected is free and openly available (Keßler, 2015). Nightlight data from the Visible Infrared Imaging Radiometer Suite (VIIRS) (Hillger et al., 2013) is integrated into the SDSS. Since October 2019, up to date, VIIRS data is no longer free available due to lack of sponsorship. Table 3 show which statistics were calculated for each data source.

Besides the continuously updated geospatial data shown in table two, data from large-scale FI datasets such as FinScope (Honohan & King, 2012) and Findex (Demirguc-Kunt et al., 2018) is also being integrated into the tool as auxiliary data. Large-scale FI and house surveys target a large geographical scale. They are not of a spatial and temporal resolution that facilitate day-to-day decision making on the placement of

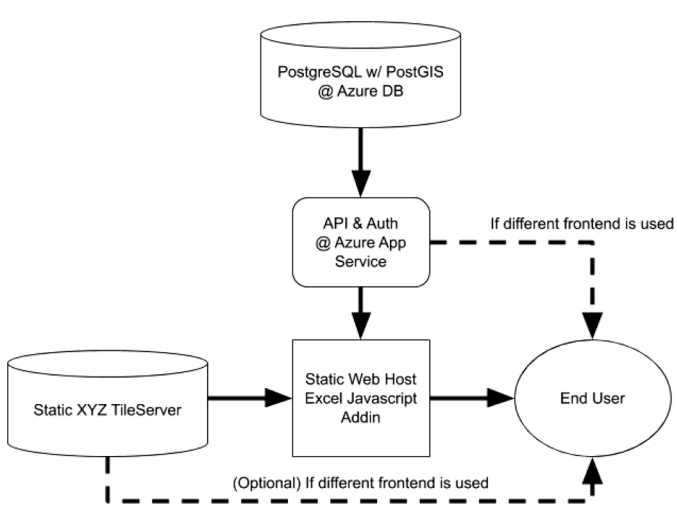


Fig. 2. Infrastructure diagram for the Decision Support System.

Table 2
An overview of the geospatial data sources used.

Name	Use	Source
Sentinel 1	Classifying settlement types	
Sentinel 2	Vegetation indices Texture analysis Overview maps Classifying settlement types	
VIIRS	Classifying settlement types Night Light Time Series	payneinstitute.mines.edu/eog
OpenStreetMap	Isochrones Points of Interests (Banks, Markets)	openstreetmap.org
WorldPop	Population density estimates	worldpop.org
SRTM	Classifying settlement types Isochrone penalties	usgs.gov/centers/eros/data-tools
FSPs Proprietary data	Informal saving groups ("Susu") Mobile Money Agents	FSPs
OpenCellID	Cell Phone Coverage Map	OpencellId.org and GSMA

Table 3
Zonal Statistics calculated. These are the data sources and zonal statistics included in the design of the model to classify types of settlement areas.

Name	Statistics
Sentinel 2 Red - Standardized	Mean, Standard Deviation
Sentinel 2 NIR – Standardised	Mean, Standard Deviation
Sentinel 2 Red - Texture (mad)	Mean, Standard Deviation
Sentinel 2 NIR – Texture (mad)	Mean, Standard Deviation
Average nightlights	Mean
NDVI	Mean, Standard Deviation
Backscatter (γ)	Mean, Standard Deviation
Coherence (coh)	Mean, Standard Deviation
Coh * γ	Mean, Standard Deviation
Terrain (slope)	Mean, Standard Deviation

agents and services by the FSPs. The Bank of Ghana provides a list of all registered financial institutions. The information about the registered institutions is continuously added to OSM, where the SDSS extracts institutional data.

During joint consultations with FSPs, discussions have taken place on sharing data on mobile money agents, brick and mortar bank branches and informal savings clubs. There is a reluctance to share these data sources with competitors and the general public. Still, there is a willingness to add some data back into the system and OpenStreetMap. The data the FSPs are willing to share comprises the placement of bank branches, cell phone coverage and road conditions. While the FSPs expressed a willingness to share and contribute data, to the authors' knowledge, no data has been shared or contributed to OpenStreetMap and OpenCellID so far.

3.2.1. Sentinel 1

Sentinel 1 is a constellation of two satellites, each carrying a C-band Synthetic Aperture Radar (SAR) instrument (Torres et al., 2012). The constellation has a revisit time of approximately six days over the case study area. In the primary operative mode (IW), the derived Ground Range Detected (GRD) and terrain corrected data has a spatial resolution of 5 m by 20 m, which is commonly processed to either 10 m by 10 m or 20 m by 20 m spatial resolution. The radar is capable of all-weather data capture, giving measurements of the phase and amplitude of the returned signal, which are useful for classifying the surface area (Torres et al., 2012).

The radar satellites are particularly suited to identify urban areas due to a phenomenon known as the 'double bounce' effect or reflection (Brunner et al., 2008). The effect is exhibited by structures facing the satellite, as the signal from the antenna bounces from satellite to a surface to the ground and returns to the satellite — the effect results in increased signal strength around structures (Sun et al., 2019). Fig. 3a

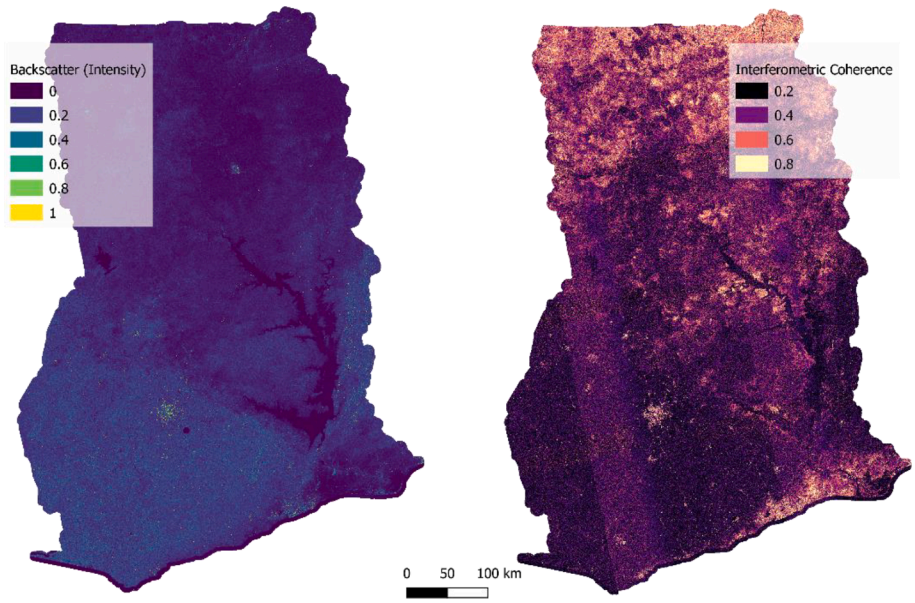


Fig. 3. (a) (Left). A merged (median merged) time series of backscatter (VV-γ) radar data over Ghana. The bright yellow spots are examples of the backscatter effect. (b) (Right). A measurement of Interferometric Coherence (VV) between the 4th of January 2020 and the 10th of January 2020 over Ghana. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shows the increased signal intensity, evident around the large cities Kumasi and Accra.

Not only human-made structures have a high backscatter value. It is possible to measure radar interferometric coherence over a time series to distinguish settlements from agriculture and forest to map urban areas. Fig. 3b shows a measurement of coherence over Ghana. Permanent structures have a consistent signal - high coherence - over time when viewed from the same orbital track, under similar weather conditions, and within a reasonable timeframe (baselines), whereas vegetation and waterbodies exhibit low coherence. The difference in signal coherence makes it possible to exclude most vegetation and water using this method. Coherence is calculated using image pairs of the complex SLC data product (Koppel et al., 2015). Fig. 4b shows the product of backscatter and coherence, as seen over Kumasi, the second-largest city in Ghana. The figure clearly shows the effects of coherence and the double bounce effect on urban mapping.

3.2.2. Sentinel 2

Sentinel 2 is a constellation of two satellites carrying multispectral imaging instruments that capture imagery at many different wavelengths at spatial resolutions ranging from 10 m to 60 m. The satellites have a revisit time of approximately six days over Ghana; however, as clouds obscure the instruments' view of the Earth, images are not always usable. Mosaics of several images from a time-series were created to

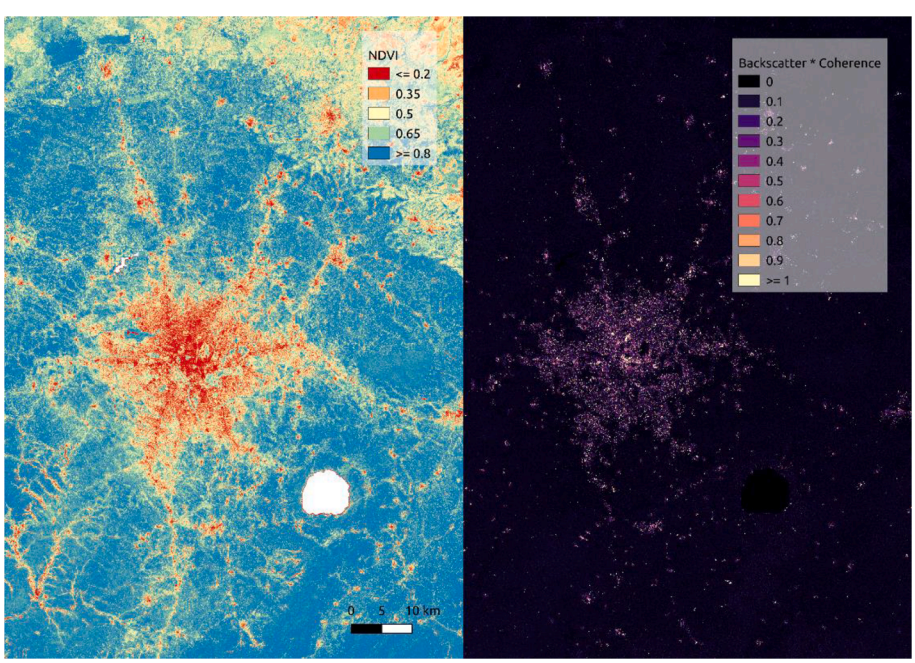


Fig. 4. (a) (Left). NDVI – Normalized Differential Vegetation Index for Kumasi in Ghana Q1 2020. (b) (Right). Backscatter (VV) multiplied by coherence (VV) for Kumasi.

compensate for cloud coverage. The methods used to generate the mosaics are custom toolchains accessible through the project repository https://github.com/casperfibaek/buteo. The mosaicking process applied combines efforts from both Sen2mosaic (Bowers, S. 2020) and Sen2Three (Müller-Wilm, U. 2020) by doing a weighted optimisation of temporal and radiometric quality. The classification of settlement types uses the red and infrared bands and a local variance texture analysis through median absolute deviation filters and circular kernels. NDVI for Ghana was also calculated (Rouse et al., 1973).

3.2.3. Visible infrared imaging radiometer suite (VIIRS)

Visible night-time lights from the VIIRS sensor onboard the Suomi-NPP satellite are included in the SDSS and classification of settlement types. Time-series on the development of radiated nightlight over time is made available in the SDSS. These time-series allow the end-user to investigate emitted nightlight over for an area as a proxy for development. Average night-time lights for Q2 2019 – Q1 2020 were used with outliers removed and non-lights set to zero using the "vcm-orm-ntl" products and averaging the months. The layer was resampled and aligned to match the 10x10m resolution of the sentinel imagery (Hillger et al., 2013), as is shown in Fig. 5a.

3.2.4. Terrain data (SRTM)

A slope map of the study area was created based on the 1-arc second digital terrain model derived from the NASA Shuttle Radar Topography Mission. The classification of settlement types utilises the slope derived from the terrain data. The SRTM layer was filtered using a median 5x5 circular kernel filter to reduce noise and is shown in Fig. 5b.

3.2.5. Vector data (OSM)

OSM data is used to supplement the satellite data in the SDSS. However, the OSM dataset is incomplete is some parts of the country, especially regarding metadata for road segments, which indicates whether a segment is a walking path or a primary road. The incompleteness meant that the data was not used as part of the settlement type classification. Examples of how the OSM data was analysed and included in the system can be seen in Fig. 5 below (see Fig. 6).

3.3. EO data – Segmentation of settlement types

At the inception workshop, FSPs requested a way of measuring the reach and extent of services to rural customers. The classification of settlement types is a key layer for the SDSS. The methodology and reproducibility of the analysis mean that it is possible to continuously update the classifications and measure the urban/rural divide. The classification of settled areas consists of a fusion of several data sources and classification steps. The classification separates rural, suburban, urban, and dense urban areas.

The creation of the classification has four distinct phases:

- Phase I: General segmentation
- Phase II: Masking settled areas (unsupervised classification)
- Phase III: Classification of settled types (supervised classification)
- Phase IV: Postprocessing of continuous areas

The segmentation and classification were done using TensorFlow and Orfeo Toolbox and a custom implementation of common EO processes.

3.3.1. Phase I. Segmenting Ghana

The initial step was to segment the case study area into distinct homogenous clusters of appropriate size. The Large-Scale Means Shift (LSMS) segmentation algorithm, as implemented in the Orfeo Toolbox, was used for the segmentation (Grizonnet et al., 2017). For the segmentation, clear visual boundaries between urban neighbourhoods were prioritised, which meant that Sentinel 2 data was used. LSMS works by iteratively clustering values based on proximity to the input bands' radiometric mean and variance. It is a memory-efficient approach of segmenting data, which is a critical feature in segmenting EO data at the country scale (Michel et al., 2014).

Urban areas stand out particularly well in the red band, and vegetation stands out in the NIR band. These two bands were used for the segmentation. The data sources were standardised to the standard score (Kreyszig, 1979), after masking all major water bodies using OpenStreetMap data, as such:

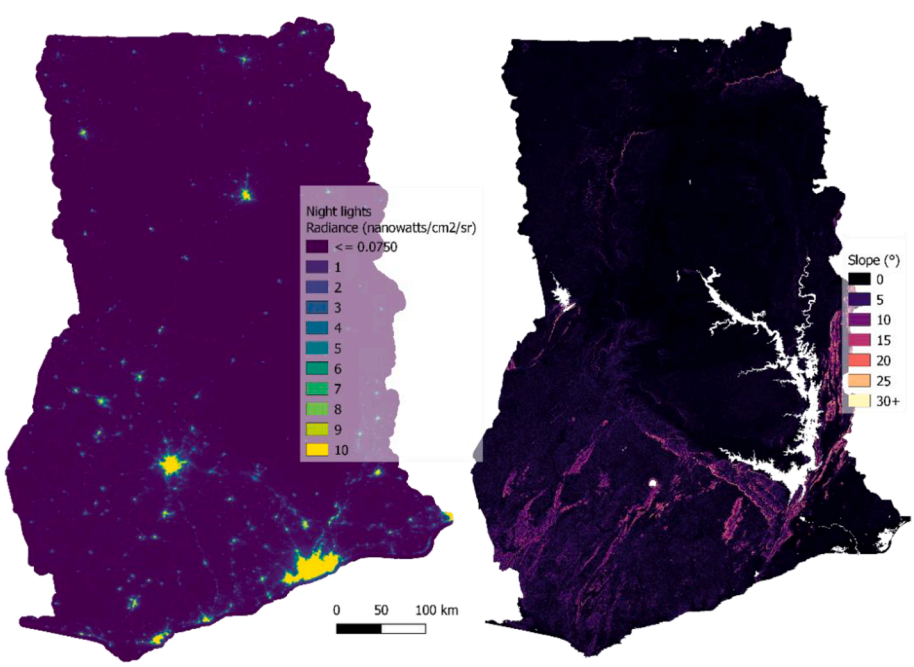


Fig. 5. (a) (Left). Average night-time lights over Ghana in 2019–2020. Accra International Airport (Kotoka) and industrial areas are visible. Data in radiance per nanowatts/cm2/sr. (b) (Right). Slope derived from filtered SRTM data over Ghana.

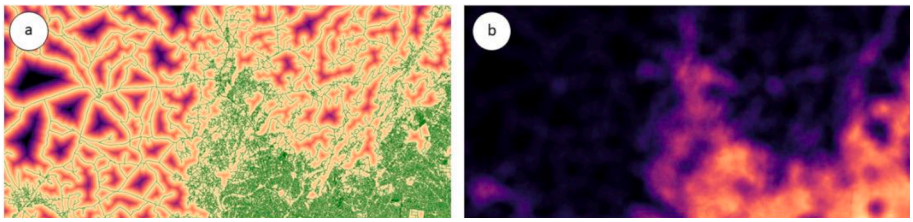


Fig. 6. (a) (Left). Road proximity in the Northern suburbs of Accra. Roads are shown in green. Bright yellow is zero distance to a road, and black is 2 km or more. (b) (Right) Road density of the same area as Fig. 6a. Bright yellow is the amount of 10x10m roads segments within a 1 km distance; black is zero. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$z = \frac{x - \overline{x}}{\sigma}$$

The segmentation algorithms' parameters were iteratively tuned, and the best parameters were chosen by visually selecting the segmentation with the highest degree of separability between the classes at small urban scales.

The chosen parameters were:

- Pixel neighbourhood size: 7 - Corner connectivity enabled.
- Spectral range: 0.15 - Standard score with waterbodies removed.
- Minimum segment size: 200 pixels.

The parameters resulted in small and clearly defined segments. The calculation required 8.6 Gb storage and a significant processing time of 1.5 days on a high-end 2019 DELL XPS 15 laptop with eight CPU cores and 64 GB of RAM. The level of processing time could be unsuitable for many larger countries or less powerful computers. It is suggested to use a larger min-size and spatial neighbourhood if applying the approach to a more extensive study area. The segmentation is shown in Fig. 7 (see Figs. 8 and 9).

After creating the segmentation, the following zonal statistics were calculated for each of the segments:

The statistics were calculated using the zonal_statistics.py script in the linked repository.

After calculating the statistics, the perimeter, area and the *iso*-perimetric quotient (IPQ) measure of compactness for each segment were calculated as defined in (R.Osserman, 1978). The IPQ is calculated as:

$$ipq = \frac{(4^*\pi^*area)}{perimeter^2}$$

Shape measurements were included to analyse whether the segmentation created unique shapes for specific settlement types; this did not turn out to be the case.

3.3.2. Phase II. Masking settled areas

After creating the initial segmentation using the normalised Sentinel 2 data, an unsupervised classification approach was used to eliminate non-settled areas, mainly water bodies, large forests, agriculture, shrubland and plains. Everything removed during this phase was classified as rural.

The unsupervised classification was created using k-means unsupervised classification (Forgy, 1965). The method works by separating the input data into clusters, where each data point is given the class of the nearest centroid or cluster center. The classification provided good results when using a high number of clusters (15+) and fit into the memory of the processing computer, which was not possible for other, more advanced, unsupervised classification methods. Table 4 shows the datasets used in the unsupervised classification:

Twenty-five clusters were used, and all the classes, elements were randomly inspected for habitation. Clusters with zero total settlements were excluded from the final layer. The classes within the settled areas serve as an initial gauge of the separability of the datasets.

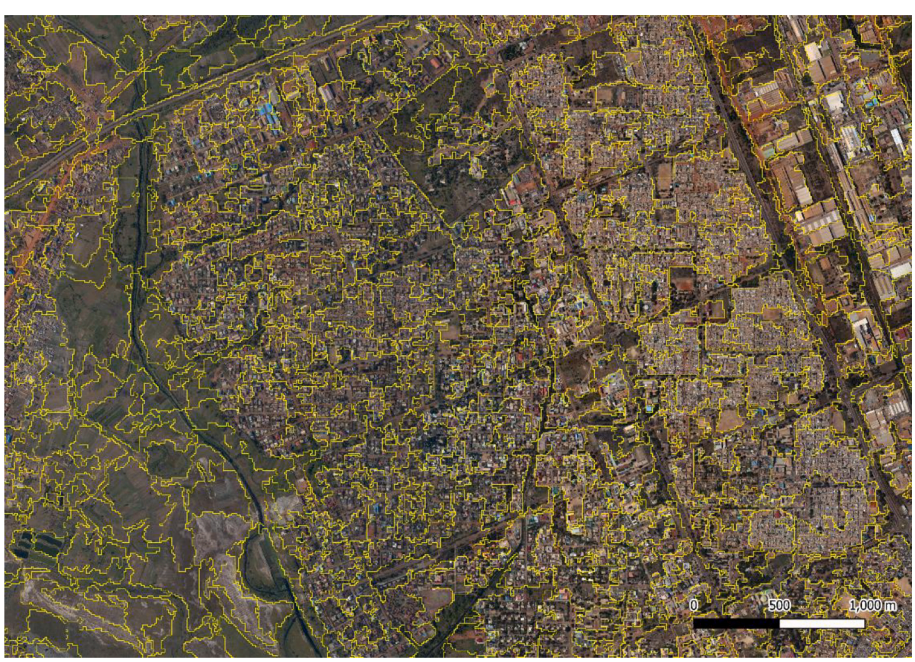


Fig. 7. Example area from Tema, Greater Accra of the LSM-segmentation. The background image is from Google Earth, CNES / Airbus, Maxar Technologies.

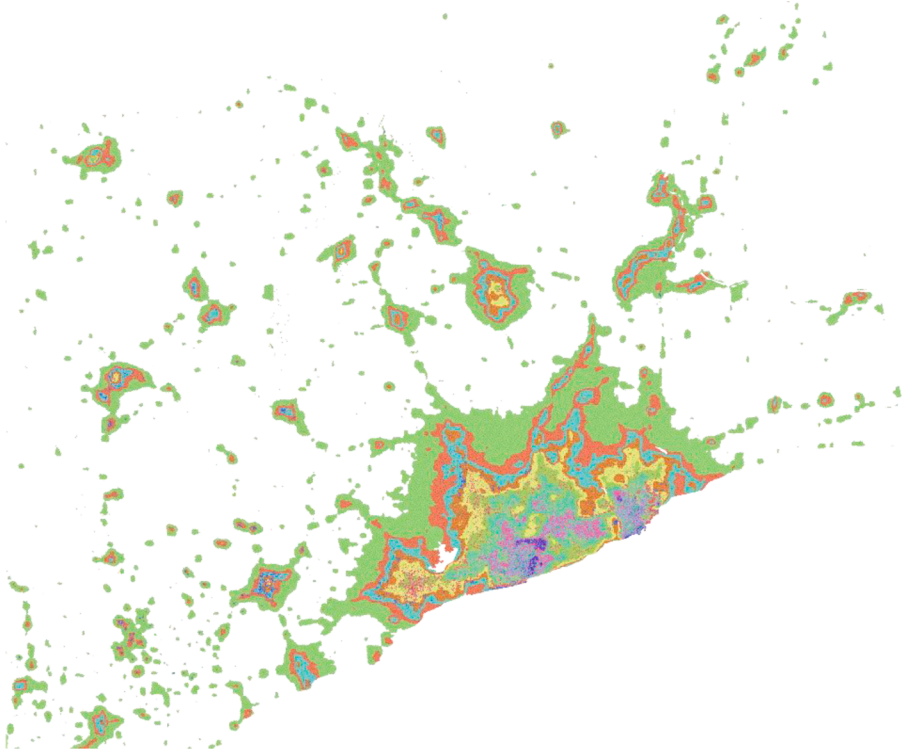


Fig. 8. K-means unsupervised classification of the Greater Accra Region showing the separation between settled and unsettled lands. Each colour represents a separate class.

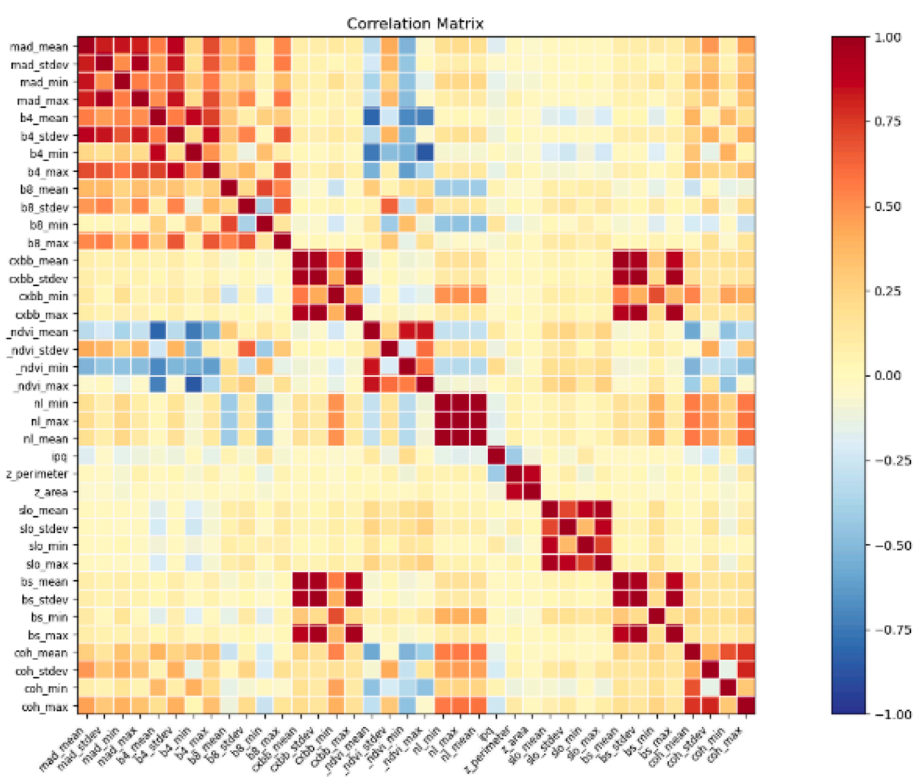


Fig. 9. Correlation matrix between model features.

3.3.3. *Phase III. Classifying settled types*

The data processed in the previous steps were used to separate Ghana into four classes: Rural, Suburban, Urban, Dense-Urban. The four initial classes were defined in the following fashion:

Rural: Very few to no people live here. Agricultural plots, water bodies, small villages with few inhabitants outside of cities.

Suburban: Low-density settlements.

Urban: Medium density settlements. Predominantly single-story

Table 4
The two layers used for performing the k-means unsupervised classification. The first layer is the product of coherence and backscatter intensity, as shown in Fig. 4b.

Name	Statistics
Sentinel 1 – Coh * γ	Mean, Standard Deviation, Min, Max
VIIRS – Nightlights	Mean, Standard Deviation, Min, Max

housing.

Dense-urban: Very dense settlements, possible multi-story housing.

Training data

The training data was collected based on expert knowledge coming from relevant individuals' knowledgeable about Ghana and has been perceived in-situ by Abednego Darko from Tema, Greater Accra and Mbinya Mutiso, an independent Financial Inclusion expert. Polygons representing the four classes were drawn in QGIS on top of high-resolution satellite imagery. From these polygons, segments were selected, which overlapped with at least 75% of their area within the polygons. Table 5 shows the number of training polygons, their geographical size, and the number of segments from the LSMS segmentation located within the respective polygons (see Tables 6–8).

Feature selection

The most important features were selected using Scikit-learn to reduce the classification problem's dimensionality (Pedregosa et al., 2011). Feature ranking with recursive feature elimination and cross-validated selection (RFECV) was used along with the feature importance of a random forest model (A. Géron, 2019).

Furthermore, the correlation matrix was plotted and investigated to eliminate features with significant correlation.

For highly correlated features, only one data source was selected. We observe that interferometric coherence over Ghana is positively correlated with nightlights and negatively to NDVI.

Deep learning model

As the training data is still unbalanced, oversampling was done using the Synthetic Minority Over-Sampling Technique (SMOTE) to create synthetic training data, thereby balancing out the training set. SMOTE creates synthetic data points by selecting k-nearest neighbours to a point, creates a point at a random distance between those neighbours, and then randomly selects one of the minority class neighbours and creates a point between those two points. Eight neighbours were used for the oversampling. (Chawla et al., 2002).

A multilayer perceptron neural network was trained using TensorFlow to perform the classification. The model consists of five layers with a dropout of 25% (Srivastava et al., 2014) in between dense layers to reduce overfitting (Kingma & Ba, 2014). The He normal weight initialisation was used (He et al., 2015) and the ReLU activation function (Glorot et al., 2011). The input layer has 64 nodes halving for each layer to 4 nodes for the output layer. The output layer uses the softmax activation function. The optimiser is the Adam algorithm implemented in TensorFlow and described in Kingma and Ba (2014).

Accuracy assessment

The accuracy assessment was done using SciKit-learn, and the overall accuracy of the model is 93.8% for the four classes with a training/test split of 25%.

Table 5
Size and count of the training material used to train the model.

Class	Training polygons	Training Size	Training segments
Rural	110	287.6 km^2	*5000
Suburban	182	150.7 km^2	3403
Urban	174	66.1 km^2	1365
Urban - Dense	117	32.5 km^2	595

* The rural class contained a significant number of segments (342,031) due to the rural polygons' covering large areas of land. A subset of 5,000 of these segments was selected using a random stratified approach to reduce the initial training set imbalance.

Table 6
Feature selection results.B4 refers to Sentinel 2 red band and B8 to Sentinel 2 NIR band.

Feature	RFECV included	RF Importance	RF included	Included in final
B4 MAD mean	Yes	2.4%	Yes	Yes
B4 MAD standard deviation	No	1.0%	Yes	No
B4 MAD minimum	Yes	1.1%	Yes	No
B4 MAD maximum	No	0.6%	No	No
B4 mean	No	0.6%	No	Yes
B4 standard deviation	Yes	0.9%	No	No
B4 minimum	No	0.5%	No	No
B4 maximum	No	0.3%	No	No
B8 mean	Yes	4.0%	Yes	Yes
B8 standard deviation	No	0.7%	No	No
B8 minimum	No	0.7%	No	No
B8 maximum	No	0.7%	No	No
Coh * Back mean	Yes	3.9%	Yes	No
Coh * Back standard deviation	No	5.2%	Yes	No
Coh * Back minimum	Yes	2.8%	Yes	No
Coh * Back maximum	No	6.1%	Yes	No
NDVI mean	Yes	3.3%	Yes	No
NDVI standard deviation	Yes	1.6%	Yes	No
NDVI minimum	Yes	0.9%	No	No
NDVI maximum	Yes	1.0%	Yes	No
Nightlights mean	Yes	12.5%	Yes	No
Nightlights minimum	Yes	14.7%	Yes	Yes
Nightlights maximum	Yes	13.4%	Yes	No
IPQ	No	0.2%	No	No
Perimeter	No	0.2%	No	No
Area	No	0.2%	No	No
Slope mean	No	0.3%	No	No
Slope standard deviation	No	0.3%	No	No
Slope minimum	No	0.2%	No	No
Slope maximum	No	0.4%	No	No
Backscatter mean	Yes	2.6%	Yes	Yes
Backscatter standard deviation	No	4.0%	Yes	Yes
Backscatter minimum	Yes	0.5%	No	No
Backscatter maximum	No	4.0%	Yes	No
Coherence mean	Yes	3.2%	Yes	Yes
Coherence standard deviation	Yes	1.9%	Yes	Yes
Coherence minimum	No	0.5%	No	No
Coherence maximum	Yes	2.6%	Yes	No

The classification achieves high accuracy, especially in rural areas. The least accurate class is dense-urban which has partial overlap with urban. Classes of urban settlements are challenging to differentiate, and the results reflect our assumptions of the classes' accuracy. The model could achieve better results using higher-resolution input data, such as aerial imagery or commercial satellite constellations. The results of the classification are shown in Fig. 10.

3.3.4. Phase IV. Postprocessing

Selecting meaningful classes for settlement types and defining 'what is urban' is a difficult task that is being tackled by UN-Habitat and other organisations working on formal definitions (UN-Habitat, 2018). EUROSTAT uses the following three categories:

Table 7
Confusion matrix for the classification of urban area types.

	Rural	Suburban	Urban	Dense - Urban
Rural	1240	12	2	0
Suburban	15	808	29	2
Urban	0	33	243	46
Dense - Urban	0	1	21	139

Table 8
Classification report for the classification of urban area types.

Class	Precision	Recall	F1-Score	Support
Rural	0.99	0.99	0.99	1254
Suburban	0.95	0.95	0.95	832
Urban	0.82	0.75	0.79	322
Dense - Urban	0.74	0.86	0.80	161

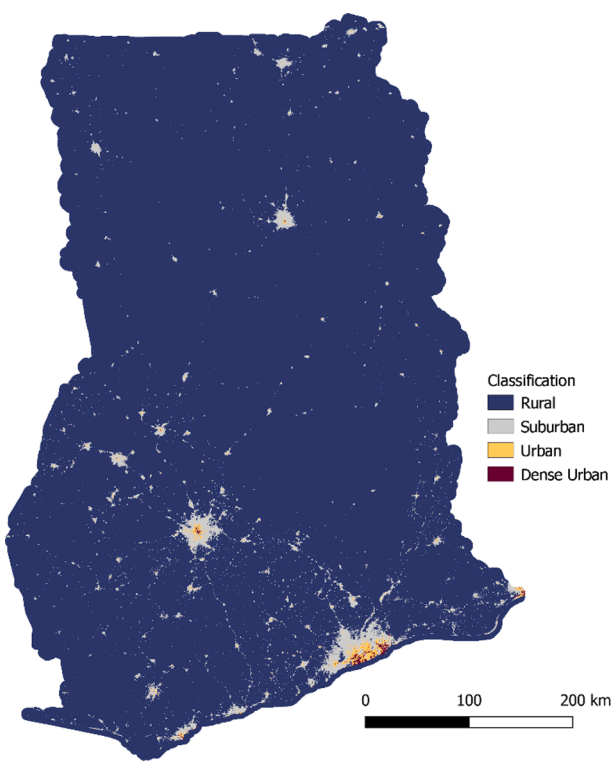


Fig. 10. The resulting urban classification for Ghana.

- Cities (Densely populated)
- Towns and suburbs (Intermediate density)
- Rural Areas (Thinly populated)

The method works on a 1 km^2 grid of Europe and classifies zones depending on population density and a morphological postprocessing step. The used postprocessing steps seek to emulate some of the postprocessing approaches of EUROSTAT's DEGURBA (Degree of Urbanisation). It merges adjacent cells based on a class's density in a 1 km neighbourhood (Dijkstra & Poelman, 2014). Each of the four classes from phase III were rasterised. The raster values were 1 for the presence of the class and 0 otherwise. These four layers were then stacked, and a circular edge-weighted kernel function was used to calculate the 1 km radius sum of each class. The class with the highest sum for each pixel represents the final pixel (see Fig. 11).

3.4. Application programming interface

The Web-API is designed as a RESTful API (R. Fielding, 2000), which provides endpoints for each of the functions available in the spreadsheet add-in. The API is written in NodeJS using the ExpressJS framework and bearer tokens for user control (S. Tilkov & S. Vinoski, 2010). It runs in a Microsoft Azure environment supported by a hosted PostgreSQL – PostGIS database holding the data and handling the API queries (Obe & Hsu, 2011). The API is agnostic regarding the technology used to fetch information, and as such, can be tied into any other frontend such as a mobile phone app, Excel, PowerBI or Tableau.

DSS technologies, a fintech operating a mobile savings platform in Ghana's informal financial sector, are integrating the API into their application. The integration allows the operator to inspect their agents' location and enrich their data visually and do customer segmentation. The integration gives the operator information into their potential reach, accessibility and cash-flow optimisation, and the type of neighbourhoods their agents predominantly operate. The points of interests, such as banks and markets, population density estimates, and land cover classifications, allow the operator to create recommendations for the agents that might improve their outreach and update. The recommendations can be used to direct agents towards areas of untapped potential or underserved communities.

3.5. Spreadsheet spatial decision support system

Spreadsheet software remains a common planning, management and control tool in the financial sector, and interviews with analysts and product managers at FSPs in Ghana and Tanzania revealed that basing the spatial decision support system around the spreadsheet format that the FSPs were familiar with would significantly ease the adaptation of the system.

An important aspect of Excel and similar spreadsheet software, such as LibreOffice Calc, is that they facilitate end-user development on top of the spreadsheet system and the functionality it makes available (Lieberman et al., 2006). Spreadsheet software has illustratively been referred to as the beach of 'no-code' ocean. End-User developed applications enable FSPs to tailor the decision support system to their specific needs and present it in a format familiar to the project and product teams in financial institutions (McDaid et al., 2011). The system allows locating proximity gaps, further to understand access challenges for different customers in different locations and situates different levels of

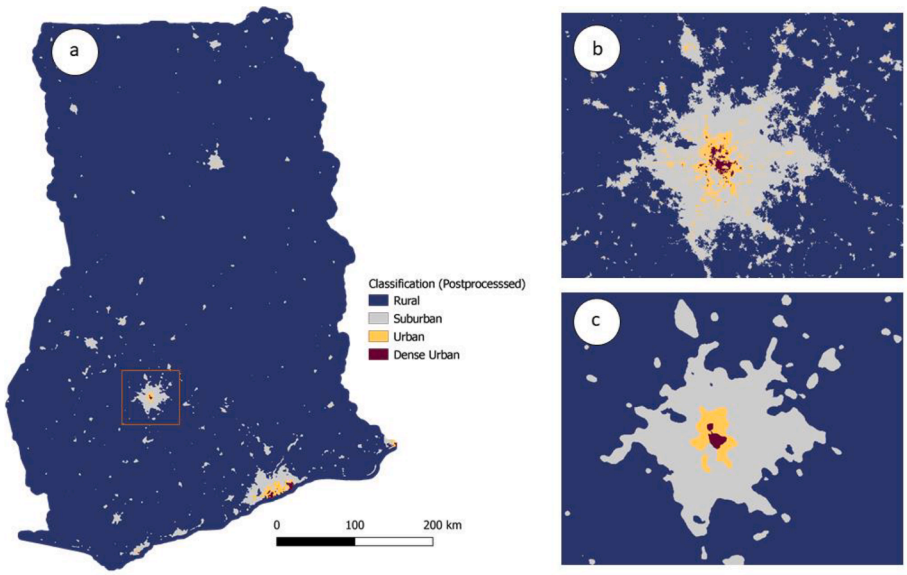


Fig. 11. (a) Simplified urban classification for the Greater Accra Area. (b) Kumasi before simplification, (c) Kumasi after simplification.

Financial Inclusion in different spatial contexts. Much of the actively used data by the financial service providers are kept in spreadsheets. The new functions allow the users to easily combine geospatially derived data from the Web API with internal data.

The decision support system combines the classification of rural/suburban/urban and dense-urban with population density data (Lloyd et al., 2019), Cell phone coverage through OpenCellID, distance and density of points of interests (e.g., banks, ATMs marketplaces) and transportation analysis through isochrones calculated on the OSM road networks. The classification can assist in market segmentation efforts that allow FSPs to tailor offers and deliver services where there is demand (McGill and Klobas, 2005) and NGOs to track progress towards the SDGs. OSM data provides information on financial institutions, public transport, marketplaces, road networks and other points of interests.

These datasets are made available through queries illustrated in Fig. 12. Each function takes an address, a place name or a coordinate pair as a parameter (see Fig. 13).

Coordinates for all spatial functions can be supplied either as latitude and longitude, google plus codes, what3words or Ghana Postal Address, also known as GhanaPostGPS or Ghana Digital Address. The Excel JavaScript API enables autocompletion and help-text on hover to support the end-user in properly using the functions and investigating other functionality relevant to their end-user application.

Previous research such as (Forster et al., 2013) relies on buffers for proximity analysis. The SDSS uses isochrones for all the spatially aware functions. Isochrones are a representation of time instead of spatial distances. That means that a polygon could represent how far a person can walk in ten minutes (Efentakis et al., 2013). Isochrones are possible to use for all functions that rely on a distance parameter. These functions take time as a parameter instead of distance and make it possible to do queries such as '*How many people live within 5 min of an Ahantaman Rural Bank branch?*' and combine queries to calculate which settlement type are the people predominately located. A time-distance calculation made through the API for Takoradi, Southwest Ghana, is shown in Fig. 14.

The Excel frontend uses a two-panel layout with the mapping functionally shown in the task pane to the right of the main cell window. The task pane is detachable and can be dragged anywhere within the Excel window. The Layer Control allows the users to add multiple geometries from the cells and show them in custom colours on the map. The map panel and the cells communicate through update buttons located at the bottom of the map. Clicking added points on the map enables the editing of attributes, which can be synced back to the cells. The map has a separate layer control for base maps in the top right corner, where the satellite data used in the analysis is available. The available layers include RGB Sentinel 2 mosaics of the wet and dry season, various Sentinel 1 layers and the settlement type classifications before and after postprocessing the classification. The interface is shown in Fig. 15.

4. Ensuring accessibility for decision makers

Nine different Financial Service Providers (FSPs) in Ghana and Tanzania assisted in the development of the system, as shown in Fig. 1. The collaboration helps ensure that the Spatial Decision Support System (SDSS) applies to concrete Financial Inclusion (FI) issues facing the individual FSPs. Following an inception workshop and start-up session in Dar Es Salaam Q4 2019, a second workshop presenting the prototype system was held mostly online and in Takoradi - Ghana in Q3 2020 with three FSPs from Ghana and one from Tanzania. At the second workshop, meetings were held individually and focused on defining user stories and training in setting up and using the system. A third workshop on integrating the SDSS was planned for Q1 2021 but has been postponed due to coronavirus restrictions.

The subgroup of FSPs has been applying the prototype system for diverse tasks that help them achieve their FI goals. Their stated goals fall into one of the below categories.

- How can we track and manage agent networks?
- How can we estimate financial catchment areas?
- Where is the untapped business potential?
- How do we use the tool to map informal saving clubs ("Susu")?
- How can we calculate time distances between sub-suppliers, markets and customers?

At the training workshop in Ghana, these goals were discussed and how the SDSS could be applied to help. Most of the goals are directly addressable within the SDSS, such as estimating catchment areas, geotagging, and assessing business potential. However, some of the requests, such as how to track agent locations, are not feasible to implement directly into the tool, as they require access to proprietary data or data that has not been gathered yet. If the FSPs gather the data, it is possible to add the data into the SDSS.

The feedback for the SDSS has been positive, especially regarding the familiar interface of the spreadsheet interface and the intuitiveness of the added map pane and custom functions. Despite none of the

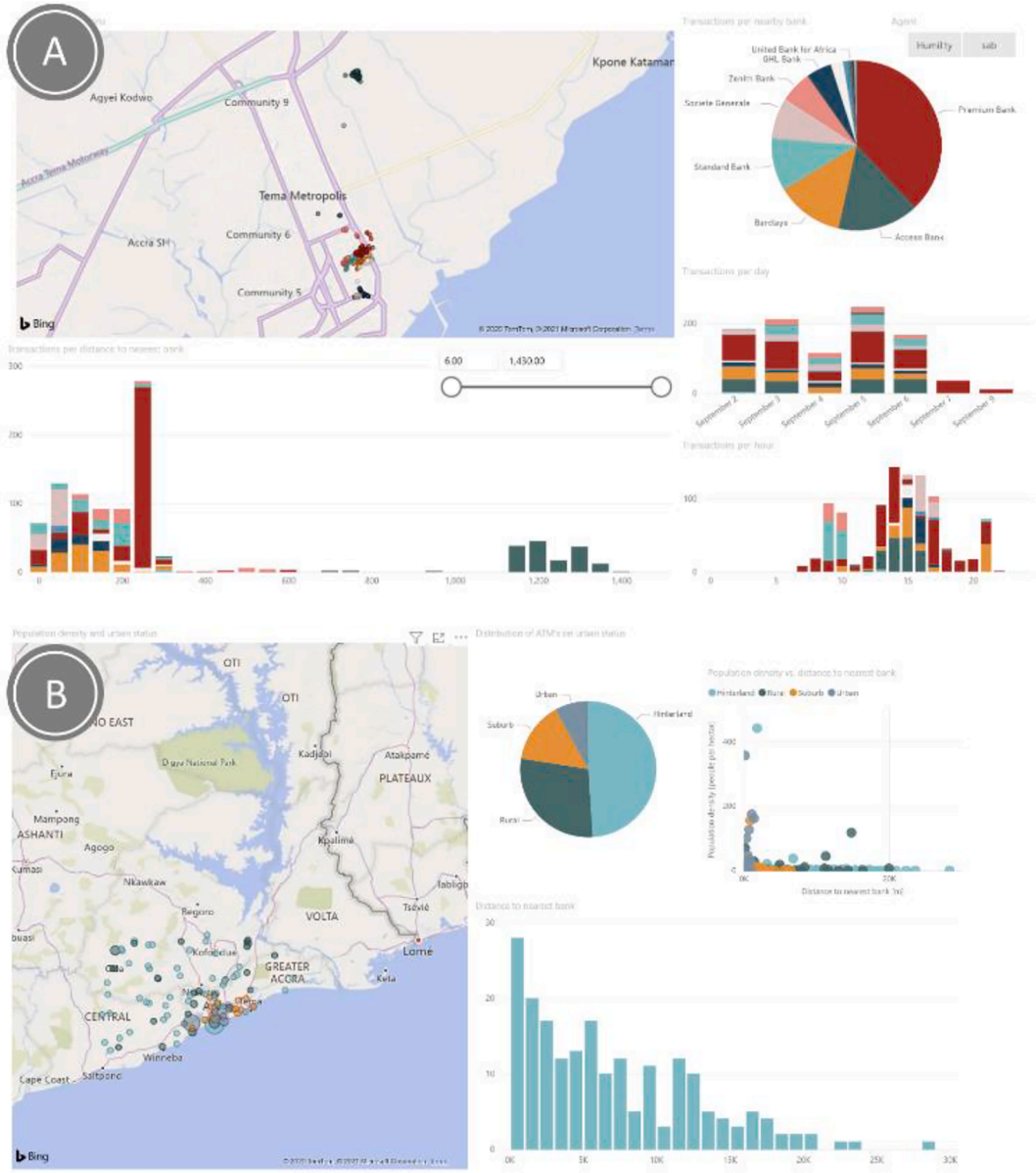


Fig. 12. (a) An example PowerBI dashboard showing the average distances to banks for communities. (b) The dashboard initially used by the FSP DSS in Tema. Uses the Web-API to add information on the nearest bank, population density and urban classification to an agent location.

Fig. 13. Example of functions available in the DSS.

participants having prior knowledge of using geographical or EO data beyond Google Maps, the combination of isochrones and geospatial functions proved intuitive. All participants were able to do basic geotagging and catchment area analysis after a brief one-hour training session. The API's documentation was deemed insufficient by the technical staff of the FSPs, and an updated version of the documentation has since been developed. Most requests from the FSPs centred around adding additional datasets to the SDSS and implementing a recommendation engine for agent location optimisation.

The specific data layers requested:

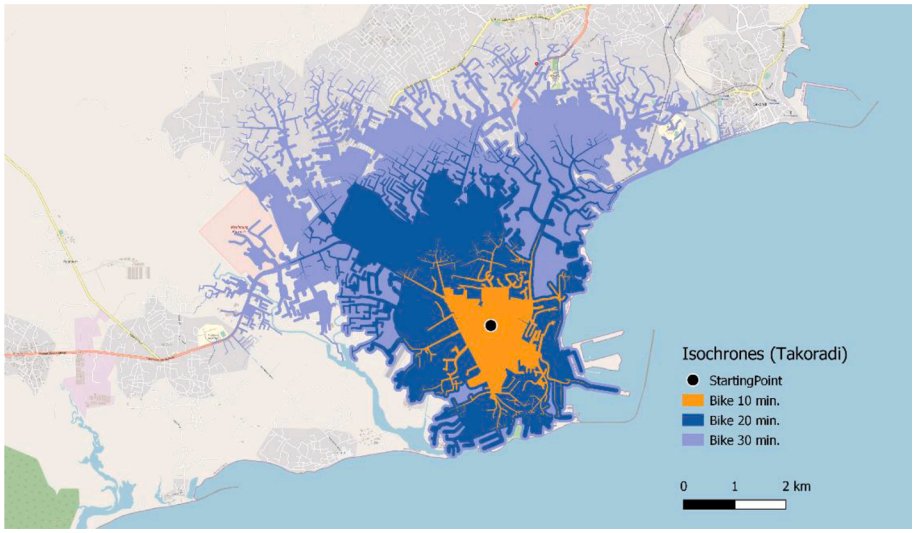
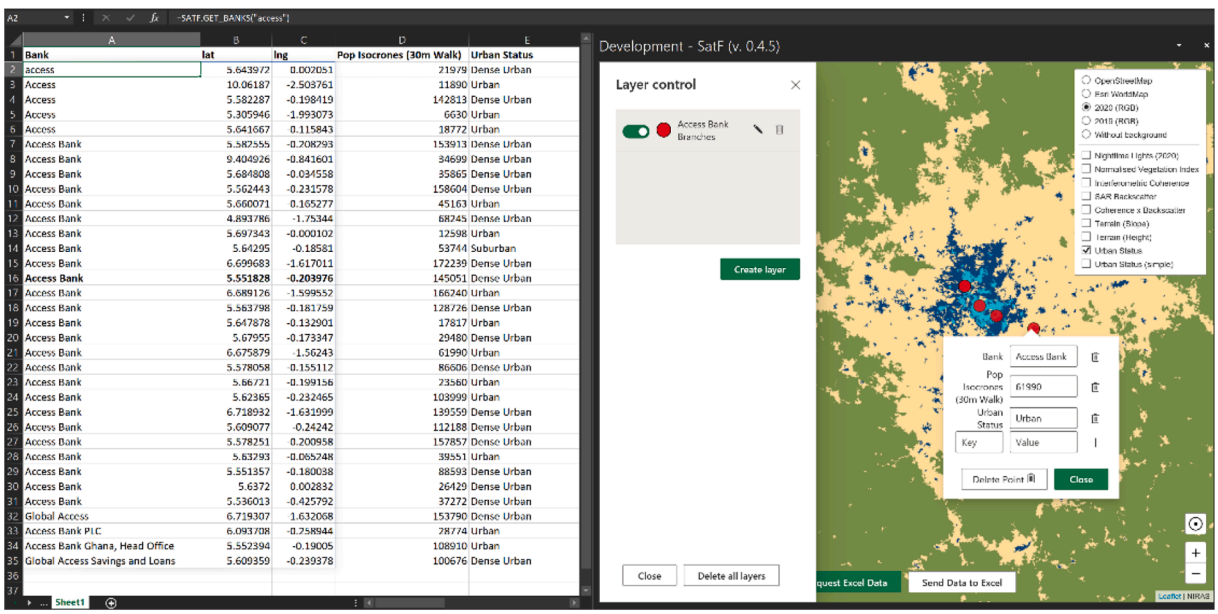


Fig. 14. Biking isochrones calculated using OpenStreetMap, PostGIS and pgRouting.



Bank	lat	lng	Pop Isocrones (30m Walk)	Urban Status
access	5.643972	0.002051	21979	Dense Urban
Access	10.06187	-2.503761	11890	Urban
Access	5.582287	-0.198419	142813	Dense Urban
Access	5.305946	-1.993073	6630	Urban
Access	5.641667	0.115843	18772	Urban
Access Bank	5.582555	-0.208293	153913	Dense Urban
Access Bank	9.404926	-0.841601	34699	Dense Urban
Access Bank	5.684808	-0.034558	35865	Dense Urban
Access Bank	5.562443	-0.231578	158604	Dense Urban
Access Bank	5.660071	0.165277	45163	Urban
Access Bank	4.893786	-1.75344	68245	Dense Urban
Access Bank	5.697343	-0.000102	12598	Urban
Access Bank	5.64295	-0.18581	53744	Suburban
Access Bank	6.699683	-1.617011	172239	Dense Urban
Access Bank	**5.551828**	**-0.203976**	145051	Dense Urban
Access Bank	6.689126	-1.599552	166240	Urban
Access Bank	5.563798	-0.181759	128726	Dense Urban
Access Bank	5.647878	-0.132901	17817	Urban
Access Bank	5.67955	-0.173347	29480	Dense Urban
Access Bank	6.675879	-1.56243	61990	Urban
Access Bank	5.578058	-0.155112	86606	Dense Urban
Access Bank	5.66721	-0.199156	23560	Urban
Access Bank	5.62365	-0.232465	103999	Urban
Access Bank	6.718932	-1.631999	139559	Dense Urban
Access Bank	5.609077	-0.24242	112188	Dense Urban
Access Bank	5.578251	0.200958	157857	Dense Urban
Access Bank	5.63293	-0.065248	39551	Urban
Access Bank	5.551357	-0.180038	88593	Dense Urban
Access Bank	5.6372	0.002832	26429	Dense Urban
Access Bank	5.536013	-0.425792	37272	Dense Urban
Global Access	6.719307	1.632068	153790	Dense Urban
Access Bank PLC	6.093708	-0.258944	28774	Urban
Access Bank Ghana, Head Office	5.552394	-0.19005	108910	Urban
Global Access Savings and Loans	5.609359	-0.239378	100676	Dense Urban

Fig. 15. The Excel interface of the SDSS. The spreadsheet is shown on the left and the added mapping and illustration functionality to the right.

- Higher-resolution population density data
- Wealth estimates
- Cell phone coverage
- Agent networks and saving clubs (proprietary data)
- Timeseries of NDVI data for an agricultural focused FSP
- Timeseries of Nightlights to estimate local development

Since the second workshop, cell phone coverage and time series of nightlight data has added to the SDSS. A meeting was held on sharing data regarding agent networks and savings clubs, but the FSPs had significant reservations about sharing sensitive data.

Wealth estimates is a topic of on-going research, as shown in (Yeh et al., 2020) and policy papers by the World Bank (Engstrom et al., 2017). High-resolution poverty estimates are available from WorldPop for Tanzania, but not for Ghana. The Tanzanian estimates are based on data from 2013 (Tatem et al., 2013) and therefore have not been added to the SDSS.

Up-to-date agricultural maps and harvest times have not been included in the SDSS as it was deemed economically infeasible. Providing or creating higher resolution population data for Ghana and parts of Tanzania is being investigated. Ghana conducted a census in 2020, and Tanzania will do so in 2022.

5. Discussion and conclusion

The presented Spatial Decision Support System (SDSS) enables Financial Service Providers (FSPs) and FinTech's to develop end-user applications and add geospatial context to their analysis. Earth Observation (EO) is a useful tool to support decision-makers in their efforts to further Financial Inclusion (FI), to support a supply-led and sustainable expansion of financial access, and track progress towards the SDGs. Using the SDSS, FSPs can devise more targeted management of agent locations or field officer routes for individual savings collection mechanisms, leading to a reduction in the cost of expanding financial access to previously under-served areas. Integrating EO and geospatial tools directly with every-day planning tools of FSP personnel is a promising method for improving decision making at all levels in the organisations.

Several workshops with FSPs were held in Ghana and Tanzania to

introduce the SDSS tool and increase its uptake and usability. The most pressing question from the FSPs was how to enable the tools to automatically suggest locations for new bank branches and mobile money agents. Semi-automatic or automatic network suggestions were out of the scope of the research presented in this paper, but it is a promising avenue for further research.

The SDSS can help the FSPs build business cases to expand their services to previously under-served communities, both in urban and rural areas. While it is not feasible to build brick and mortar banks in deeply rural areas, the SDSS can locate mobile money agents and mobile banks periodically. Without a system to show where people live, how they live and how they are connected to the infrastructure, it is challenging to make a business case necessary to promote FI.

The classification approach presented in the research achieves an accuracy of 93% in separating settlement types. The classes are density focused but area dependent: what will be considered a dense urban in Ghana might not be considered dense in Nigeria. A new classification using a common global definition is needed to make the SDSS globally scalable. The use of coherence and backscatter in conjunction with night light data proved to be a promising approach to classify and quantify the density of human-made structures. The corner reflectors highlighted by the Sentinel 1 data seem biased towards detached housing, which is more prevalent in Kumasi than Accra. Accra has a higher density of structures, but the backscatter signal is stronger in most of Kumasi. There appears to be a decline in backscatter signal in very densely populated neighbourhoods. Sentinel 1 data is only available from one orbital direction over Africa. Research into the difference access to both directions would make on estimating human-made structural volume would be beneficial for improving the method. A model of global human-made structural volume could also be used as a worldwide baseline in urban definitions. The SDSS could be enhanced further by the Inclusion of pollution data as different settlement types are likely to exhibit different pollution patterns, which would ease the model class separation.

Inference of socioeconomic variables, such as wealth, using a combination of data sources, including Earth Observation, would be highly useful to make global decision support tools in support of SDGs. The research shows that it is possible to create a SDSS for FI that makes proximity and financial Inclusion an addressable challenge and enables the creation of business cases to reach the previously under-served while managing the last-mile delivery of financial services.

We propose that further research is conducted into the design of SDSSs for supporting FI that incorporates a variety of open data sources and inference from machine learning. The methodology for classifying settlement types should be expanded further to include subclasses for the urban areas that are highly relevant for a financial institution, such as organised vs self-organised, and the creation of wealth estimates derived from open data. Furthermore, research into the design of the SDSS and the creation of structural volume estimates from sentinel 1 should be investigated.

Funding

Danish Innovation Fund, Niras/Alectia Foundation

CRediT authorship contribution statement

Casper Samsø Fibæk: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Writing - original draft. **Hanna Laufer:** Investigation, Writing - review & editing. **Carsten Keßler:** Supervision, Writing (Review and editing). **Jamal Jokar Arsanjani:** Supervision, Writing (Review and editing).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Andries, A., Morse, S., Murphy, R., Lynch, J., Woolliams, E., Fonweban, J., 2019. Translation of Earth Observation data into sustainable development indicators: An analytical framework. Sustainable Dev. 27, 366–376.

Beck, H., Zimmermann, N., McVicar, T., Vergopolan, N., Berg, A., Wood, E., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5 (1).

Bowers, S., 2020. Sen2Mosaic, BitBucket, https://bitbucket.org/sambowers/sen2mosaic/src/master/sen2mosaic/.

Brunner, D., Bruzzone, L., Ferro, A., Fortuny, J., Lemoine, G., 2008. Analysis of the double bounce scattering mechanism of buildings in VHR SAR data. In Proc. SPIE Conf. Image Signal Process. Remote Sens. XIV (Vol. 7109, pp. 71090Q-1).

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artificial Intelligence Res. 16, 321–357.

De Soto, H., 2001. The mystery of capital. Finance and Development 38 (1), 29–33.

Demirguc-Kunt, A., Klapper, L., Singer, D., Ansar, S., Hess, J., 2018. The Global Findex Database 2017: Measuring Financial Inclusion and the fintech revolution. The World Bank.

Dijkstra, L., Poelman, H., 2014. A harmonised definition of cities and rural areas: the new degree of urbanisation. European Commission (No. 01). Regional Policy Working Papers.

Efentakis, A., Grivas, N., Lamprianidis, G., Magenschab, G., Pfoser, D., 2013. November. **Isochrones, traffic and DEMOgraphics**. In Proceedings of the 21st.

Engstrom, Ryan, Hersh, Jonathan, Newhouse, David, 2017. Poverty from space: using high-resolution satellite imagery for estimating economic well-being.

Fielding, R.T., Taylor, R.N., 2000. Architectural styles and the design of network-based software architectures, Vol. 7. University of California, Irvine, Irvine.

Fielding, Roy T., Taylor, Richard N., 2002. Principled design of the modern web architecture. ACM Transactions on Internet Technology (TOIT) 2 (2), 115–150.

Forgy, E., 1965. Cluster Analysis of Multivariate data: efficiency vs. Interpretability of Classifications. Biometrics 21, 768–769.

Forster, D., Peachey, S., Stahl, S., 2013. Mapping Proximity Bringing Products and Services Close Enough to the Poor to be Meaningfully Usable and Still Keep Them Sustainable for WSBI Partner Banks. https://www.wsbi-esbg.org/SiteCollectionDocuments/Proximity Paper EN (Long Version).pdf.

Geiger, M., Kwakye, K., Vicente, C., Wiafe, B., Boakye, A., Yaa, N., 2019. Fourth Ghana Economic Update: Enhancing Financial Inclusion - Africa Region (English). Ghana Economic Update; no. 4. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/395721560318628665/Fourth-Ghana-Economic-Update-Enhancing-Financial-Inclusion-Africa-Region.

Géron, A., 2019. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media.

Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. In: In Proceedings of the fourteenth international conference on artificial intelligence and statistics, pp. 315–323.

Grizonnet, M., Michel, J., Poughon, V., Inglada, J., Savinaud, M., Cresson, R., 2017. Orfeo ToolBox: open source processing of remote sensing images. Open Geospatial Data, Software and Standards 2 (1), 1–8.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision, pp. 1026–1034.

Heitmann, S., Buri, S. Poverty Estimation with Satellite Imagery at Neighborhood Levels. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/318541559339293802/Poverty-Estimation-with-Satellite-Imagery-at-Neighborhood-Levels-Results-and-Lessons-for-Financial-Inclusion-from-Ghana-and-Uganda.

Hillger, D., Kopp, T., Lee, T., Lindsey, D., Seaman, C., Miller, S., Rink, T., 2013. First-light imagery from Suomi NPP VIIRS. Bull. Am. Meteorol. Soc. 94 (7), 1019–1029.

Honohan, Patrick, King, Michael, 2012. Cause and effect of financial access: cross-country evidence from the Finscope surveys. Banking the World: Empirical Foundations of Financial Inclusion 45–84.

Josephat, L., 2016. The role of agency banking in promoting financial inclusion: descriptive analytical evidence from Tanzania. Eur. J. Business Manage. 8 (33), 231–340.

Keßler, C., 2015. OpenStreetMap. In: Shekhar, S., Xiong, H., Zhou, X. (Eds.), Encyclopedia of GIS. Springer, Cham. https://doi.org/10.1007/978-3-319-23519-6_1654-1.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimisation. arXiv preprint arXiv:1412.6980.

Koppel, K., Zalite, K., Sisas, A., Voormansik, K., Praks, J., Noorma, M., 2015. Sentinel-1 for urban area monitoring—Analysing local-area statistics and interferometric coherence methods for buildings' detection. In 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 1175-1178). IEEE.

Kreyszig, E., 1979. Applied mathematics. John Wiley & Sons, Hoboken, NJ.

Lieberman, H., Paternò, F., Klann, M., Wulf, V., 2006. End-user development: An emerging paradigm. In: End User Development. Springer, Dordrecht, pp. 1–8.

Lloyd, C.T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F.R., Sinha, P., 2019. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. Big Earth Data 3 (2), 108–139.

Mahendra, S., 2006. Financial inclusion: issues and challenges. Econ Polit. Wkly 41 (41), 4310–4313.

McDaid, K., MacRuairi, R., Clynch, N., Logue, K., Clancy, C., Hayes, S., 2011. Spreadsheets in financial departments: An automated analysis of 65,000 spreadsheets using the luminous technology. arXiv preprint arXiv:1111.6866.

McGill, T.J., Klobas, J.E., 2005. The role of spreadsheet knowledge in user-developed application success. Decis. Support Syst. 39 (3), 355–369.
Michel, Julien, Youssefi, David, Grizonnet, Manuel, 2014. Stable mean-shift algorithm and its application to the segmentation of arbitrarily large remote sensing images. IEEE Trans. Geosci. Remote Sens. 53 (2), 952–964.
Müller-Wilm, U., 2020. Sen2Three. GitHub Repository, https://github.com/senbox-org/sen2proc/tree/master/SEN2THREE.
Obe, R., Hsu, L., 2011. PostGIS in action. GEOInformatics 14 (8), 30.
Osserman, R., 1978. The isoperimetric inequality. Bull. Am. Math. Soc. 84 (6), 1182–1238.
Paganini, M., et al., 2018. Satellite earth observations in support of the sustainable development goals. The CEOS Earth Observation Handbook.
Peachey, S., Mutiso, M., 2019. Moving proximity from critical issue to addressable challenge: Possible approaches and tools. https://www.opml.co.uk/files/Publications/a0600-savings-at-the-frontier/proximity-scoping-main-report-satf.pdf?noredirect=1.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learning Res. 12, 2825–2830.
Richardson, L., Ruby, S., 2008. RESTful web services. O'Reilly Media, Inc.
Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. 'Monitoring vegetation systems in the Great Plains with ERTS'. Third ERTS Symposium, NASA SP-351 I, 309–317.
Shi, G., Jiang, N., Yao, L., 2018. Land use and cover change during the rapid economic growth period from 1990 to 2010: A case study of Shanghai. Sustainability 10 (2), 426.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learning Res. 15 (1), 1929–1958.
Steele, J.E., Sundsøy, P.R., Pezzulo, C., Alegana, V.A., Bird, T.J., Blumenstock, J., Hadiuzzaman, K.N., 2017. Mapping poverty using mobile phone and satellite data. J. R. Soc. Interface 14 (127), 20160690.
Sun, Z., Xu, R., Du, W., Wang, L., Lu, D., 2019. High-resolution urban land mapping in China from sentinel 1A/2 imagery based on Google Earth Engine. Remote Sens. 11 (7), 752.
Tatem, A.J., 2017. WorldPop, open data for spatial demography. Sci. Data 4 (1), 1–4.
Tatem, A., Gething, P., Bhatt, S., 2013. Pilot high resolution poverty maps. Southampton and Oxford University.
Tilkov, S., Vinoski, S., 2010. Node. js: Using JavaScript to build high-performance network programs. IEEE Internet Comput. 14 (6), 80–83.
Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Traver, I.N., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24.
UK Space Agency, 2020. Space for Finance in Developing Countries. Caribou Space, Niamh Barry. https://www.spacefordevelopment.org/library/space-for-finance-in-developing-countries/.
Unhabitat.org. 2018. Harmonisation of Urban Definitions Key To Monitoring Implementation Of SDGs And The NUA | UN-Habitat. https://unhabitat.org/harmonization-of-urban-definitions-key-to-monitoring-implementation-of-sdgs-and-the-nua [Accessed 11 May 2020].
United Nations, Statistics Division, Department of Economic and Social Affairs, Statistics Division, 2008. Principles and Recommendations for Population and Housing Censuses Revision II. ST/ESA/STAT/SER.M/67/Rev.2.
United Nations, 2017. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. A/RES/71/313. In press.
Varghese, G., Viswanathan, L., 2018. Financial inclusion: opportunities, issues and challenges. Theor. Econ. Lett. 8 (11), 1935–1942.
Yeh, C., et al., 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nat. Commun. https://doi.org/10.1038/s41467-020-16185-w.