Assessing the utility of open-source geo-spatial data for AI-driven estimation of fire spread risk in informal settlements

Citation for published version:
Panas, D, Seth, S, Watmough, G & Rush, D 2022, Assessing the utility of open-source geo-spatial data for AI-driven estimation of fire spread risk in informal settlements. in *41st EARSeL Symposium/6th EARSeL Workshop on Developing Countries: Symposium Book of Abstracts.* Proceedings of the EARSeL Symposium, European Association of Remote Sensing Laboratories (EARSeL), pp. 174-176, 41st EARSeL Symposium, Paphos, Cyprus, 13/09/22. <https://cyprus2022.earsel.org/assets/BookOfAbstracts.pdf>

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
41st EARSeL Symposium/6th EARSeL Workshop on Developing Countries

Assessing the Utility of Open-Source Geo-Spatial Data for AI-Driven Estimation of Fire Spread Risk in Informal Settlements 41st EARSeL Symposium 2022

EARSeL Cyprus 2022
Abstract
Corresponding Author:
sohan.seth@ed.ac.uk

DAGMARA PANAS[1,*], SOHAN SETH[1,*], GARY WATMOUGH[3,*], DAVID RUSH[4,*]

[1]UNIVERSITY OF EDINBURGH, SCHOOL OF INFORMATICS, UK
[2] UNIVERSITY OF EDINBURGH, SCHOOL OF GEOSCIENCES, UK
[3] UNIVERSITY OF EDINBURGH, SCHOOL OF ENGINEERING, UK

[*] THIS WORK WAS SUPPORTED BY WELLCOME TRUST

Keywords (5): open-source, remote sensing, AI, fire, resilience, informal settlements

The challenge

Populations housed in informal settlements are rapidly growing, and due to their informal nature, they are typically lacking in basic safety features. For example, dense layouts and unsupervised building practices result in increased vulnerability to fire spread. Developing an open-source, data-driven AI model of assessing the risk of fire spread at any location within informal settlements using remote sensing and geospatial data could aid risk mitigation and contribute to Sustainable Development Targets 11.1 and 11.5 We present an analysis of suitability of publicly available data for this purpose. Specifically, we assessed the coverage and volume of geo-mapped fire events, required as target for training the AI model; and the quality of building footprint data, from Google Open Buildings (abbrev. OB), that is crucial for deriving relevant risk factors such as density of dwellings.

Methodology

To understand the quality of fire event data we have performed basic exploratory analysis, examining the completeness, coverage, and resolution in both spatial and temporal domains of two existing datasets from Cape Town, South Africa from 2020. We also visualised the distributions of event sizes to verify whether these comprised a representative sample.

To understand the quality of publicly available building footprint data we have performed both an exploratory analysis, and a comparative analysis against an independently obtained ground truth dataset (obtained from VHR aerial photography and LIDAR via manually-corrected automatic detection) in the region of interest. The comparative analysis can be further subdivided in two parts: first, we computed the values of several previously identified aggregate settlement risk metrics (e.g. average distance to nearest neighbour) and computed correlation and goodness-of-fit statistics between OB and the ground truth dataset (across the 291 known settlements in the region of interest); second, we performed a left-intersecting spatial join (joining data wherever an OB footprint overlaps with a ground truth footprint), followed by a comparison of distributions of individual footprint characteristics across all the settlements, and an assessment of observed overlaps between individual footprints. We have focussed specifically on OB as the most comprehensive and mature product (e.g., recently released Microsoft footprints do not cover South Africa where we have ground truth data).

Results

The first dataset, a tabular dataset provided by the city of Cape Town, is heavily biased towards single dwellings, which would be a serious obstacle in extrapolating fire spread risks. The second dataset, a geo-mapped spatial dataset co-curated by the city of Cape Town and University of Edinburgh, presents excellent spatial precision and size distribution, but consists of only 46 events, which limits its use.

While the footprint data provided freely by Google represents an unprecedented opportunity for humanitarian efforts, it is not sufficient in and of itself to support fire spread risk modelling. The spatial coverage is excellent, nominally the entirety of sub-Saharan Africa. However, the temporal coverage is limited to a single point in time. As such, it would therefore not be suitable for modelling, unless back-dated versions in regions of interest were available. Furthermore, a comparison against reference data indicated that the quality of detections within informal settlements would most likely be insufficient for accurate representation of layout characteristics informative of fire risk assessment (Jaccard Index of 0.722). At individual dwelling level, it appears that Open Buildings are skewed towards small and scattered, sometimes overlapping structures. At settlement level, average inter-building distance and landscape density tend to be under-estimated, while edge density, patch density and fractal dimension tend to be over-estimated (see Figure).

Outlook for the future

A previous feasibility study (L. Gibson et al., 2019, in Fire Safety Journal) has demonstrated the utility of remote sensing in geo-mapping of fire events when the approximate location and time are known. While both socioeconomic factors and satellite imagery availability can prevent capturing every single fire event and their respective burn scars, we believe in large part the lack of data is due to the lack of necessary systems and tools to use the data in a meaningful way.

We hypothesise that the quality gap of footprint data for our specific purpose is due to the broad applicability of the data product, and in particular the fact that it had been developed to work well across both rural and urban settings, and with limited labelled data compared to the diversity of architecture throughout sub-Saharan Africa. An algorithm specifically targeted towards informal settlement footprints is achievable.

We propose that with a larger set of geo-mapped fire events, and more accurate footprint dataset, fire spread risk modelling is feasible, as suggested by previous analysis of the relationship between fire severity and settlement footprint characteristics and weather conditions (L. Gibson et al., 2022, in Computers, Environment and Urban Systems)

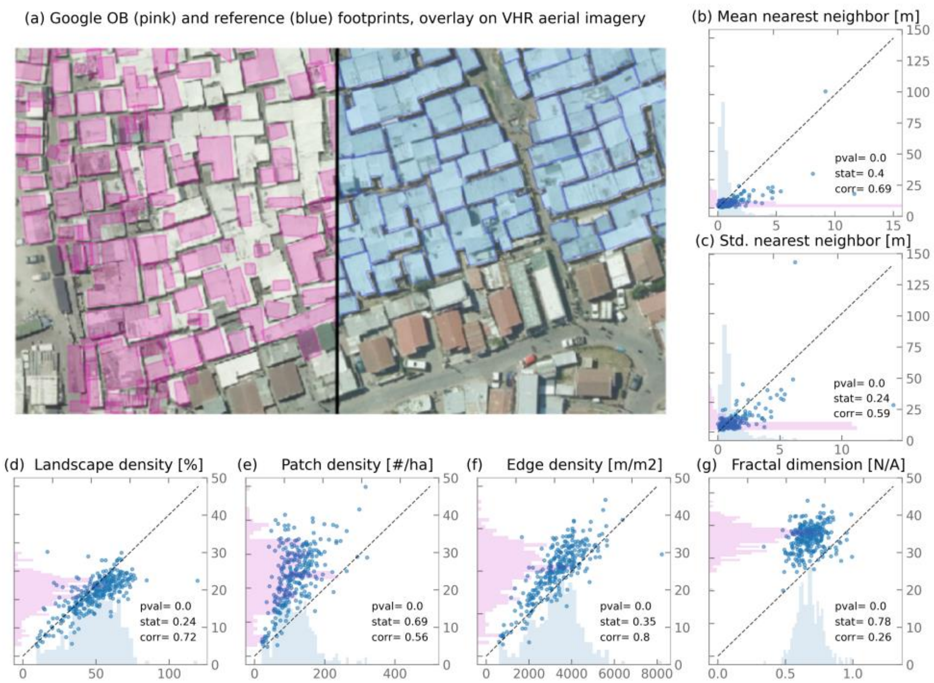

Figure Comparison of Google OB and ground-truth footprints. Note that for panels b-g the x/y axes show risk metric values, the secondary x/y axes (right/top) show histogram counts.