IDENTIFICATION OF CLIMATE CHANGE-RELATED HAZARDS IN INFORMAL COMMUNITIES THROUGH THE APPLICATION OF MACHINE LEARNING TO SATELLITE IMAGES

FEDERICO BAYLE,* NABIL KAWAS, ALEJANDRA MORTARINI,*** CARLOS RUFÍN,*** ALFREDO STEIN,**** LIDIA TORRES,**DANIEL TSAI*****

*Dymaxion Labs, Buenos Aires, Argentina
**Instituto Hondureño de Ciencias de la Tierra, Universidad Nacional Autónoma de Honduras
***Institute for International Urban Development, Cambridge, Mass., USA
University of Manchester, United Kingdom
rufinca@gmail.com

Paper prepared for presentation at the
"2020 WORLD BANK CONFERENCE ON LAND AND POVERTY"
The World Bank – Washington DC, March 16-20, 2020

Abstract

We describe a new Artificial Intelligence (AI) tool, specifically a Machine Learning (ML) tool that can, through an integrated analysis of a large variety of data including satellite or aerial images, provide a rapid and low-cost identification of exposure to climate-related hazards for informal settlements. The tool can identify exposure to these hazards more economically, quickly, frequently, and transparently than current approaches, and can facilitate visual accessibility to otherwise inaccessible areas. With wide potential availability to public officials and the public, it will add value to planning practices, municipalities and communities, leading to more inclusive and equitable cities, and to governance processes that address the joint impacts of rapid informal urban expansion and climate change. Prototyping will be based on data from Tegucigalpa, Honduras. A second stage will expand the tool into a wide-range planning function by adding scenario analysis and enhancing ML capabilities.

Key Words: Informality, climate change, hazards, Machine Learning, planning

Introduction

Informal settlements, housing large numbers of urban residents, are expected to expand disproportionately around the world through rural urban migration over the next decades (World Bank, 2019). At the same time, the impacts of global warming will become increasingly uncertain, creating new risks for the most vulnerable urban residents. Low income households will continue to settle in low income areas deemed unsuitable for housing increasing their exposure to risks, or to encroach on natural resources essential to cities' environmental sustainability (*ibid*.). The combination of these twin forces of informal urbanization and climate change thus create an urgent need for adaptation and mitigation measures to increase the resilience of vulnerable urban communities.

Yet the fast pace of urbanization and uncertainties about climate change mean that decision makers and low-income communities face growing complexity in finding ways to respond to acute shocks and chronic stresses associated with weather variability. Rapidly identifying how informal settlements grow and become increasingly exposed to risks can provide vital information to plan for climate change adaptation as well as to reduce spatial segregation. Moreover, in many quickly growing or expanding medium to small municipalities and communities, there is a lack of resources to monitor the expansion of informal settlements, to conduct and update assessments of exposure and risk, and to plan interventions to increase resilience. Current means to identify hazards is slow and expensive. For example, large-scale geospatial monitoring of urban growth requires expensive resources. It can take a team of human analysts six months to survey 5,000 square kilometers. Moreover, analysts have to inspect satellite, drone and StreetView images manually, losing accuracy and scale.

The use of artificial intelligence (AI) in connection with urbanization is rapidly developing and demonstrating great potential for many purposes. Specifically, the application of Machine Learning (ML) to remote sensing allows the detection and mapping of informal settlements as well as a growing number of other features of geographic space (World Bank, 2019). A further possibility is the use of ML not only for static recognition, but also for prediction through the examination of past patterns and their correlates across a variety of data. With increasingly abundant spatial data, and rapidly falling data capture and analysis costs, advances in ML create the opportunity to drastically cut the time and expense involved in the identification and prediction of climate-related hazards in vulnerable urban communities, such as flooding, deforestation, erosion and mudslides, fires, and urban heat islands.

In what follows, we describe a new tool that harnesses advances in ML and growing data availability to help vulnerable urban communities address rapidly evolving hazards at the intersection of informal urbanization and climate change. Importantly, the tool is not fully predefined and restricted to an elaborate piece of information technology. Technology, specifically ML, provides the basis for the identification of hazards. This information then becomes the key input for a process of collaboration with officials and communities, at both regional and local levels, to develop and implement adaptation and mitigation interventions that can increase the resilience of vulnerable communities in the face of these hazards. Our focus in this paper is on river and precipitation flood-related stresses, such as regular flooding after heavy rains, rather than shocks (e.g. flooding from hurricanes) or other hazards (e.g., heat island effects). Nonetheless, our tool is readily extendable to shocks and to other hazards. Specifically, both flood-related stresses and shocks are driven by settlement in hazardous locations and lack of interconnectivity or flow of water conduits. Moreover, stresses often lead to shocks, for example when repeated flooding weakens the ground and results in landslides.

The value of this tool lies in its potential to: (i) facilitate visual accessibility to climate change-related hazard areas otherwise inaccessible due to budgetary, geographic, or security constraints; (ii) disrupt or change current urban planning processes by reducing their time and cost; (iii) correlate urban expansion with climate change-related hazards, suggest planning scenarios that facilitate transparent formulation of interventions, and allow effective risk reduction; (iv) cover large areas while responding to local settlement, geography, and climate conditions; and (v) allow frequent updates of local data and thus of risk assessments and exposure at a speed that does not exist at present, crucial for disaster prevention and planning at a time of rapid change in climate patterns.

This tool will be useful to practitioners, public officials and communities in cities that are fast expanding their urban footprint through informal settlements due to urban-rural migration, political migration and clime-related migration. The tool can also be available to the public through a user-friendly interface. Public officials can use the tool to make timely informed land policy decisions at a local, regional or national level, thus advancing SDG 11; practitioners can use this tool as a planning instrument to divert or prevent exposure of informal settlements to flooding and other climate related risks and draw local strategies and interventions, furthering SDG 13; and governments that embrace inclusionary strategies can use the tool as a participatory instrument for inclusive planning framework to together produce mapping and other information to make inclusive decisions, as sought by SDG 16.

This tool is particularly suited for medium to small cities, which typically lack the resources for traditional planning projects, and that are experiencing fast uncontrolled growth, while being subject to flash flooding and other climate-related hazards. Moreover, as an analytical platform that can rely on public-domain data and be easily updated and maintained through its open-source configuration, it will create a permanent local capacity and reduce the need for reliance on expensive outside expertise.

In the remainder of the paper, we first offer a brief overview of the current methodologies for flood risk management in vulnerable urban communities. Then, we describe our novel approach, followed by our proposed pilot project. The last section of the paper considers the scaling of the tool and its extension to various kinds of hazards, and its limitations.

Flood risk management in vulnerable urban communities

To the best of our knowledge, no other tool is currently available or at an advanced stage of development that, at a fraction of time and cost relative to current methods, addresses the joint impacts of rapid informal urban expansion and climate change hazards, and thus has the potential to lead to more inclusive and equitable cities and governance processes. Although significant advances are being made in the recognition of different kinds of features of geographic space and the built environment, we are not aware of predictive tools in existence that can deliver the fast results of ML. Current flood hazard mapping approaches appear to rely on very high resolution images and multi-criteria (MCA) analysis (Franci et al., 2016). MCA analysis requires the selection and testing of multiple layers of geo-referenced data, and the selection or computation of relative weights for each type of data in a predictive model. ML, by contrast, allows much faster analysis by means of an algorithmic search for the combination of data with the best predictive power, obviating the need for modeling floods.

For this reason, ML is being applied to an increasing number of phenomena, but not floods as far as we know. Descartes Labs uses ML to predict food production from satellite imagery of land used to grow food crops (Chaturvedi, 2018). Stanford University researchers have developed a ML model to predict poverty on the basis of night lights (*ibid*.). The World Bank, through its Global Program for Resilient Housing, is developing ML-based methods to identify poor-quality housing that can be highly vulnerable to a variety of hazards (World Bank, 2019). This differs from our focus on flooding rather than on housing per se, which could be a complementary type of information in order to determine vulnerabilities and develop interventions.

On the planning side, highly effective participatory processes for the management of local flood hazard stresses have been developed in recent years, offering valuable models for our tool. The Kounkuey Design Initiative (KDI) in Nairobi has implemented innovative methods for extensive community involvement in the identification of hazards and the development of interventions. In Honduras, the Asset Planning for Climate Change Adaptation (APCA) project likewise developed a highly participatory process that also included extensive collaboration with municipal authorities (Stein & Moser, 2018). Both of these projects, entailed time-consuming analyses of local conditions with regard to flood hazards, through extensive community input. In the KDI case, additional flood modeling analysis was carried out by a European engineering firm, BuroHappold Engineering, requiring travel by company engineers to Nairobi as well as detailed surveys of hydraulic structures and river cross-sections by local staff and volunteers (Mulligan et al., n.d.). Our tool can greatly reduce these needs through the use of ML to identify local hazards, although we will still seek community validation of findings as part of the participatory process described below.

Approach

The first component of our tool is the identification of flood hazards via ML analysis of satellite images, at a regional and local scale, with a focus on the risks and vulnerabilities arising from flash flooding in rapidly expanding informal settlements. The starting point for the development of this component is an existing application, developed and successfully applied by our co-author and colleagues at Dymaxion Labs, that uses ML to identify informal settlements through satellite images (see, e.g., UNICEF, 2018; IADB, 2019). We will add topographic and microclimate information, including local precipitation and wind data, and, for the ML stage, the identification of areas that experience flash flooding.[1] The

[1] A recent paper on flood hazard mapping using satellite images and MCA of geo-referenced data used topography (slope), distance to drainage channels, drainage texture, geology, and land cover as predictive factors for flood

learning process will result in the ability to identify potential hazard areas based on patterns of informal occupation, topography, and microclimate.

Dymaxion Labs is one of the pioneers in the use of ML to identify informal settlements, and has developed an open-source tool, AP-Latam,[2] which uses geospatial data to identify informal settlements in Latin America. AP-Latam detects and maps informal settlement patterns such as texture and morphology of roofs, drawing a territorial boundary for every potential zone. It does so by combining satellite images with convolutional networks, a technique applied to the analysis of automatic images or computational vision within the ML field. The program consists of a binary classifier of image tiles that can classify images as containing an informal settlement or not. In order to build a dataset for training and validation, the classifier takes a vector file of polygons of previously-known informal settlements, and takes fixed-size tiles of images by sliding a window across the entire satellite image. For each image tile, it checks if the tile intersects with some polygon and tags it appropriately. Then, to make predictions over new images, it slides a window over the new image and builds a new vector file of polygons of the size of each positively-tagged image tile.

This technique allows the automatic analysis of images on a continuous basis, since the algorithm, through experience and quantity of data, is always learning by itself to improve results.[3] Using the AI tool, Dymaxion Labs and partner organizations identified nearly 4,500 settlements in Argentina, Honduras, and Peru. For example, in seven regions in Argentina, 100 new settlements were identified, from a total of 354 informal settlements surveyed in 2016.

The use of temporal series of satellite images, allowing us to track changes over time, will enhance the predictive power of the ML analysis, as will the addition of data from multiple sites as use of the tool grows over time.

---

hazards (Franci et al., 2016; see also Rahman & Di, 2017). Other variables in use in flood forecasting models include land use or cover characteristics, such as surface runoff, permeability, and roughness (Rahman & Di, 2017).

[2] https://code.iadb.org/en/repository/60/ap-latam

[3] More detailed methodological aspects, shared technical documentation and references can be found on AP-Latam website at https://ap-latam.dymaxionlabs.com/en/publications.

The main output will be in the form of risk "heat maps" showing the varying degrees of flood-flash risks affecting informal settlements.[4]

The second component of the tool is the use of the hazard identification information from the first component to develop, plan, and execute adaptation and mitigation interventions. This process begins in parallel with the identification of hazards, through a scan of the institutional context of the area, and particularly the formal and informal organizations and governance processes (Ngobi & Mulligan, n.d.). The scan will reveal the legal mandates, resources, structure, and behavior of municipal and other levels of government that interact with informal settlements in the area; and it will also identify other organizations, particularly community-based ones, that articulate the interests and voices of the settlements' residents. We will then approach a variety of institutions and organizations identified by means of the scan, in order to organize the rest of the process, leading up to the implementation of agreed interventions.

This process will continue with the analysis of the hazard identification information. For this purpose, we intend to incorporate scenario analysis functionality to the tool. Scenario analysis will allow policy makers and communities to assess future risks and uncertainties, and to examine the effects of various potential interventions, including their interaction with likely urban expansion paths and related land-use and local topographic alterations, which will both affect the evolution of flood hazards and constrain the range of potential interventions. As with hazard visualization, a number of scenario planning instruments are available for our purposes, particularly those developed by the Lincoln Institute of Land Policy.[5]

Scenario analysis will enable the development of implementation options at different geographic scales (regional, city and community) and time frames. These options can range from community-managed interventions to urban growth management plans, or even watershed management policies, among other possibilities. Again, we will make use of new products for the visualization of spatial interventions (e.g. Urban Footprint[6]) to enable a participatory approach for the development and selection of options.

[4] The World Bank's ThinkHazard tool (http://thinkhazard.org/en/) is precisely intended for the purpose of hazard visualization, and can be harnessed for use with our tool. Another example of flood hazard heat maps can be found in Franci et al., 2016.

[5] See http://www.scenarioplanning.io/

[6] https://urbanfootprint.com/

The participatory process will be based on the experience of the Institute for International Urban Development in Arusha, Tanzania, and other locations, the work of Alfredo Stein and Caroline Moser in Tegucigalpa, Honduras, and other project sites (Moser & Stein, 2011; Stein & Moster, 2014; Stein & Moser, 2018), and the experience of the Kounkuey Design Initiative in Nairobi, Kenya (Mulligan et al., 2015). The process will then conclude with training and mentoring of local officials and community leaders and participants not only with regard to the execution and maintenance of the chosen options, but also to the permanent updating of the hazard identification information, in order to detect new or changing hazards, and to remain actively engaged in the management of the flooding hazards affecting the informal settlements.

Pilot project for Proof of Concept

A proof-of-concept demo will be created with the aim of testing the degree of accuracy of ML in the identification of flood hazards, improving the predictive capacity of the ML algorithm, for example through the use of additional or different types of data, and to better understand the level of granularity of the images required for accurate prediction, especially at the community level. We will start with existing ML-generated maps that locate informal settlements, adding to them layers of topographic information and precipitation through a Geographic Information Systems (GIS) platform.

Prototyping will be based on data from Tegucigalpa, Honduras. The proposed test site, Tegucigalpa, is Honduras' capital and main city, and is experiencing rapid urban expansion due to internal migration, with annual population growth rates in excess of 2% and the largest concentration of urban poverty in the country (Stein & Moser, 2018). The city's location in a mountainous area has led to the growth of informal settlements in steep hill slopes that create significant flood stresses during rainy seasons and that have led to increasingly severe shocks during hurricanes and tropical storms since 1998, making landslides the main flood-related hazard in Tegucigalpa (*ibid*.). One of our co-authors led the Asset Planning for Climate Change Adaptation project mentioned above, and is therefore extensively acquainted with the physical, social, and institutional conditions in Tegucigalpa, as well as the development of community-scale interventions to address flood risks. In addition, Dymaxion has already mapped informal settlements in the area, and other initiatives have led to extensive assessment of various climate-related hazards affecting the city. Lastly, to supplement access to these various databases and its extensive local experience, and thus ensure the feasibility of ML implementation through the combination of satellite image analysis and knowledge about local conditions, the project team is also collaborating with the Institute of Earth Sciences (IHCIT) at the Autonomous National University of Honduras (UNAH). IHCIT has an extensive database of local weather data in the country covering approximately the last 30 years, and leads extensive research efforts on the assessment of climate-related hazards in the country.

Project coordination and the development of the tool's policy and planning functionality will be led by the Institute for International Urban Development. I2UD has significant experience in the use of geospatial data related to urban planning and urban informality, including the use of AI techniques for urban density analysis in a 2015 regional plan for Arusha, Tanzania, and in participatory action and planning, as well as field experience in Latin America and the Caribbean (LAC), Sub-Saharan Africa and Middle East and North Africa (MENA).

Scale, scope, and limitations

Once we have established the proof of concept, the tool can be rapidly scaled to other areas. Additional data from other locations will improve the accuracy of hazard detection through additional learning. Data availability may be more limited in other locations, but fortunately there is a fast-growing menu of possibilities to fill any gaps. Remote sensing techniques allow the quantification of precipitation levels and intensity (Rahman & Di, 2017). Other means of aerial data collection include crowdmapping (Panek and Sobotová, 2015), with products like OpenStreetMap[7] and OpenMapKit,[8] and the City Scanner developed by the MIT Senseable City Lab.[9] With regard to shocks, Facebook's Data for Good Disaster Maps[10] can be a valuable source of information.

We also envisage the application of the tool to other kinds of hazards and urban challenges. Identification of flood hazards is likely to point to problems of deforestation, land-use change and erosion that go beyond strict water-related concerns, such as the cutting down of brushwood to make

[7] See openstreetmap.org
[8] http://openmapkit.org/
[9] http://senseable.mit.edu/cityscanner/
[10] https://dataforgood.fb.com/tools/disaster-maps/

charcoal for household use, or the diversion of water for industrial use. ML can also be harnessed to identify fire hazards in rapidly expanding settlements that reach woodland or bush areas in urban peripheries; or urban heat island hazards that can result in high localized temperatures posing heightened health risks for local residents. While the former kinds of hazards should be discernible from aerial images, like flood-related hazards, increased availability of global surface temperature data will enable the latter. The tool can thus help address not one but many if not most of the hazards created by climate change and informal urbanization.

Naturally, our tool has limitations that should be acknowledged. The accuracy of ML-based predictions remains to be shown. Data availability in some locations around the world is certain to pose challenges. Companies like DigitalGlobe and Planet gather comprehensive, high-resolution images of the entire planet, and use ML to identify a variety of features in these images, but this information is only available on commercial terms at high prices (Chaturvedi, 2018). And above all, the development and implementation of interventions will take time, especially if it is to follow a participatory approach, for it requires conveying complex information, articulating a minimum degree of agreement on preferred interventions, harnessing resources to carry out and maintain those interventions, and organizing the permanent updating and monitoring of evolving hazards. Still, we are confident that accuracies can be improved with better data, that data is increasingly available and obtainable at a rapidly declining cost, and local organizational processes can be put in place.

References

Baker, J.L. 2012. *Climate Change, Disaster Risk, and the Urban Poor: Cities Building Resilience for a Changing World.* Washington, DC: World Bank.

Chaturvedi, A.2018. The tremendous potential of Machine Learning in satellite imagery. *Geospatial World*. July 23.

Dewan, A. 2013. *Floods in a Megacity: Geospatial Techniques in Assessing Hazards, Risk and Vulnerability*. Dordrecht: Springer.

Douglas, I., Alam, K., Maghenda, M., Mcdonnell, Y., McLean, L., Campbell, J. 2008. Unjust Waters: Climate Change, Flooding and the Urban Poor in Africa. *Environment and Urbanization* 20 (1): 187–205.

Franci, F., Bitelli, G., Mandanici, E., Hadjimitsis, D., Agapiou, A. 2016. Satellite remote sensing and GIS-based multi-criteria analysis for flood hazard mapping. *Natural Hazards* 83: S31-S51.

Inter-American Development Bank (IADB). 2019. Monitoring informal settlement growth in Manaus, Brazil with drones. https://blogs.iadb.org/ciudades-sostenibles/en/monitoring-informal-settlement-growth-in-manaus-brazil-with-drones/ Accessed February 12, 2020.

Jha, A.K., Bloch, R., Lamond, J. 2012. *Cities and Flooding. A Guide to Integrated Urban Flood Risk Management for the 21st Century.* Washington, DC: World Bank.

Jha, A.K., Miner, T.W., Stanton-Geddes, Z. 2013. *Building Urban Resilience: Principles, Tools, and Practice*. Washington, DC: World Bank.

Kuffer, M., Pfeffer, K., Sliuzas, R. 2016. Slums from Space—15 Years of Slum Mapping Using Remote Sensing. *Remote Sensing* 8, 455: 1-29.

Miguez, M. G., Verol, A. P., Santos, R. B. 2013. Alternative Solutions for Sustainable Urban Drainage Systems Integrating Areas of Irregular Urban Growth with the Formal City. *Applied Mechanics and Materials* 409–410: 996–1003.

Moser, C. Stein, A. 2011. A methodological guideline for implementing Urban Participatory Climate Change Adaptation Appraisals. *Environment and Urbanization* 22(2): 463-486.

Mulligan, J., Harper, J., Ngobi, B. 2015. Consultation and Data Collection Methodology for the Building Urban Flood Resilience in Kibera project. N.p.

Mulligan, J., Venn, N., Gregoriou, R., Ker-Reid, D., Travers, A. n.d. Integrated Flood Risk Management in Slums - Applying 1D Modelling in Kibera, Nairobi. N.p.

Ngobi, B., Mulligan, J. n.d. The Institutional and Regulatory Context for Flood Risk Reduction in Nairobi, Kenya: Final Briefing Paper, 2015-2016. Nairobi: Kounkuey Design Initiative.

Panek, J., Sobotová, L. 2015. Community Mapping in Urban Informal Settlements: Examples from Nairobi, Kenya. *Electronic Journal of Information Systems in Developing Countries* 68(1): 1-13.

Rahman, M.S., Di, L. 2017. The state of the art of spaceborne remote sensing in flood management. *Natural Hazards* 85: 1223-1248.

Stein, A., Moser, C. 2014. Asset planning for climate change adaptation: lessons from Cartagena, Colombia. *Environment and Urbanization* 26(1):166-183.

Stein, A., Moser, C. 2018. Asset Planning for Climate Change Adaptation in Poor Neighborhoods of Tegucigalpa, Honduras. Washington, DC: Inter-American Development Bank.

Turner, B.L., Kasperson, R.E., Matson, P.M., McCarthy, J.J., Corell, R.W., Christensen, L., Eckley, N., Kasperson, J.X., Luers, A., Martello, M.L., Polsky, C., Pulsipher, A., Schiller, A. 2003. A framework for vulnerability analysis in sustainabilty science. Proceedings of the National Academy of Sciences of the USA 100(14): 8074–8079.

UNICEF. 2018. UNICEF Innovation Fund: Dymaxion Labs. unicefstories.org/2018/04/04/venturefunddymaxionlabs/ Accessed June 18, 2019.

World Bank. 2019. Brief: Global Program for Resilient Housing. https://www.worldbank.org/en/topic/disasterriskmanagement/brief/global-program-for-resilient-housing. Accessed on November 11, 2019.