Title: Spatial-temporal variation in population and household infrastructure exposed to urban flood risks in India, 2000-2014

Authors: Jessie Pinchoff[1*], Mark Montgomery[1,3], Deborah Balk[2], Hasim Engin[2], Natalie Lin[2]

* Corresponding author, jpinchoff@popcouncil.org

[1] Population Council, New York, NY
[2] CUNY Institute for Demographic Research and Marxe School of International and Public Affairs, Baruch College, City University of New York, New York, NY
[3] Stony Brook University, Stony Brook, New York

Acknowledgements

The work was funded, in large part, by the US National Science Foundation award #1416860 to the City University of New York, the Population Council, the National Center for Atmospheric Research (NCAR) and the University of Colorado at Boulder, and from an Andrew Carnegie Fellowship from the Carnegie Corporation of New York to Deborah Balk. We thank Kytt McManus of Columbia University for access to the high-resolution LECZ data and Wangyal Shawa, of Princeton University, for his invaluable knowledge of the Indian spatial data collection.

Abstract

India is at high risk of both inland and coastal flooding, and simultaneously transitioning from rural to urban, exposing millions of people to flood risks. Flooding affects fresh water access and household sanitation, contributing to the spread of water and vector-borne diseases. Flood and population data will be overlaid using R software to measure spatial-temporal exposure. Urbanization, total population, and population by indicators of vulnerability will be quantified using the Global Human Settlement Layer (GHSL) and the 2011 Indian Census. Flooding will be characterized by multiple datasets including the low elevation coastal zone (LECZ), the Dartmouth Flood Observatory flood event catalogue, passive microwave flood detection data, Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) and cyclone and tropical storm tracks datasets. Results will highlight variation in urban population and household infrastructure exposed to flood risks to guide better-targeted disaster preparedness efforts, adaptation policies, and city planning regarding water and safety.

Background

Inland and coastal India are prone to floods, and this risk is increasing as climate change exacerbates the frequency and intensity of flood events.(1) With its current population of over 1 billion persons, India is only considered to be about one-third urban, with the majority of its urbanization to occur in the decades ahead.(1) Urban areas have significant vulnerability to flooding, due to the presence of slums, lack of sanitation and urban planning, and poor housing quality. Floods cause significant damage to infrastructure and threaten the livelihoods of millions of individuals every year.(1, 2) The risks are likely to be especially severe in the cities and towns where income is low and protective infrastructure is lacking.(3) Despite wide recognition of urbanization in India, for a variety of reasons traditional census data fail to capture the full, complex fabric of that transition. This study not only employs conventional demographic data sources, but also makes use of satellite-derived data to more fully understand urbanization. We combine these with multiple measures of flood-risk to establish populations at risk and assess housing infrastructure vulnerabilities.

Both coastal and inland flooding are of concern in India, particularly in urban areas. A recent study estimated that the percentage of urbanized areas in high-frequency flood zones will increase from 52% in 2000 to 88% in 2030, which would make India the most at-risk country in the world in the near future.(2) Flooding can be exacerbated in urban areas that restrict flood-water flow due to land use conversions resulting in high coverage by houses, roads, and pavement.(4) Low-lying cities and towns near the coast will face additional risks from tropical storms, storm surges and coastal flooding.(3) These coastal areas can be identified by their proximity in the low elevation coastal zone (LECZ); the LECZ covers 2% of the world's land area but contains 10% of human population and 13% of the world's urban population.(5) Population density is significantly higher in coastal than in non-coastal areas, with an ongoing trend of coastal urbanization.(3) This higher density is due to combinations of history, geography and economics, for example trade opportunities, that have led most of the world's megacities to be located in the coastal zone, attracting people and driving coastal migration.(5, 6) Recent studies suggest that the average sea levels could rise by 1m or more by 2100, with significant impacts on coastal environments including inundation and flooding, coastal erosion, shoreline relocation, and salt water intrusion, in addition to exposure to larger disasters such as tsunamis. (6-8) Between 2000 and 2060, India could experience a 3-fold increase of its LECZ population from 6.1% to 10.3% of its population, under high growth scenarios.(6)

Urbanization in India and the demand for clean water may also place increasingly competing pressure on the water supply of Indian cities, particularly in slums where development is largely haphazard and unplanned. Urbanization adversely affects hydrological processes leading to a deteriorating water environment, as impervious surface area increases and disrupts the natural water balance, obstructing water channels, and building drains to redirect water flows.(9) This increases runoff and the duration of flooding particularly in urban slums where drainage systems and household access to water infrastructure may be poor. Lack of preparedness in these settings is widespread and can worsen the outcome of flood events, resulting in loss of life, temporary evacuation, epidemics, and damage to the cities infrastructure including transportation and power systems.(4) It is challenging in urban areas to accurately measure or predict the population exposed to flood events, the within population variation in vulnerability by key demographic indicators, and variation in household access to critical water access and infrastructure that may be damaged by floods.

Study Objectives

Our objectives are to describe urban growth and vulnerability over time in relation to flood risks in urban cities across India. This will entail 1) quantifying changes in GHSL data from 1975, 1990, 2000 and 2014 in relation to flood data from multiple sources, to measure change in built up area in relation to flood risk; and 2) overlaying 2011 census data in relation to flood data to quantify the number of persons at risk and also determine any variation in their household level vulnerability to water access shortages.

Methods

All data necessary for analyses have been compiled, processed and cleaned including the two population datasets and four environmental datasets. Table 1 provides data descriptions and citations of sources.

Population Data

First, we will explore urbanization using the satellite-derived Global Human Settlement Layer (GHSL), which provides high-resolution estimates of land covered by structures, that is, the density of built structures.(10) It is commonly used as a proxy for urbanization, depicting 'built up area' at each time point. These data will be used to extrapolate risk exposure into the pre- and post-2011 period.(11) The most recent version of the GHSL data, still in beta-release format, combines Landsat scenes corrected with a newer, higher-resolution inputs from the Sentinel satellite, currently available for the years 1975, 1990, 2000 and 2015.(12)

To disaggregate population and housing data, we will explore the finest within-urban spatial units available for the 2011 Indian census (ward-level). Previous work has focused on settlement-level (one order coarser than wards) representing villages and cities and towns as a whole, but in this analysis we examine the finest resolution spatial data for cities only. Wards are available for 60 of the largest cities in India, including all cities with populations greater than 1 million. Thus, for this analysis we aim to focus on urban areas to explore internal variation in their risk to flooding. Lastly, we explore housing infrastructure and access related to water and sanitation that may be most impacted by floods. The Indian Census includes housing data with information reported on housing quality, latrine type, drinking water access, and other key indicators of household water infrastructure and vulnerability.

Environmental Data

To explore coastal flooding in relation to the population datasets, these will be overlaid with the low elevation coastal zone (LECZ), a widely used coastal flood proxy.(5) Being geographically located in the LECZ is predicted to increase risk of exposure to seaward hazards, but is not itself a measure of flooding events. To complement the LECZ, we draw on coastal and in-land flooding captured in the Dartmouth Flood Observatory (DFO) daily flood catalog and daily records based on passive microwave flood detection (these daily records are summarized for 2000 and 2015 time points). Rainfall data from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) data are explored to quantify weekly or monthly rainfall anomalies in addition to the two flood event datasets; storm tracking data are used to identify high-risk threatening events such as cyclones.

Data Extraction

In order to describe populations at risk from the 2011 Indian Census, these various spatial layers are overlaid and extracted using well-established methods.(5) Environmental data will be extracted to the city ward level for analysis. In addition, CHIRPS and cyclone storm tracks data can be overlaid to pinpoint particularly high flood risk areas and weeks. Once overlaid, estimates of total population exposed will be calculated according to census data on gender, age, literacy, occupation, scheduled caste/scheduled tribe. The overlay will also allow for extraction of household level variables such as type of latrine, drinking water source, housing condition for each ward in order to examine a range of vulnerability characteristics. We will experiment with inequality index or measure, following recent approaches to examine inequality in urban areas in the context of climate change.(13) Similar methods will be used to extract and

overlay GHSL data. Although GHSL lacks specificity on population at risk, it allows us to explore multiple time points of urbanization in a dataset that is comparable over space and time.
The analysis will describe spatial variation in population exposed to flooding and explore the variation in risk factors that may increase vulnerability to this exposure. And, place that risk in the context of locations experiencing urbanization.

Preliminary Results

GHSL and flooding

The proportion of total area built up and the average built up area per time point (1975, 1990, 2000 and 2014) are tabulated and graphed for all 13 coastal states and union territories in India (36 total), by LECZ risk category. Table 2 highlights the results from Maharashtra State, that suggests consistent increases in built-up land over time, with the highest in the high- and medium-risk LECZ categories (from 25.14-28.56% built-up in the high risk <5 meter LECZ, and from 28.33-35.36% in the medium-risk 5-10 meter LECZ). This suggests higher levels of urbanization along the coast, which is consistent as India’s largest city of Mumbai is located on the coast in Maharashtra State. Patterns over space and time are being explored for all coastal states. The maps in Figure 1 highlights Kolkata city and the spatial growth in built-up area between 2000 and 2014 in relation to the LECZ and to DFO flood events (recorded in the month of January 2012, as an example). DFO flood events are being summarized by season and month.

GHSL data can be used to highlight changes in built area exposure to LECZ and flood events over time between 1975 and 2014 but do not measure population exposed, only a proxy for urbanization based on land use.

Census and flooding

The ward level data from the 2011 Indian Census contains population information disaggregated by various demographic and socioeconomic characteristics, allowing for exploration of population exposed to flooding and proximity to the LECZ and variation in vulnerability within wards. Figure 2 highlights the example of Kolkata, indicating the distribution of total population per ward, and the overlap with LECZ and reported floods from January 2012. Although the census data only represent one point in time, they are critical for exploring detailed variation in vulnerability and risk within the urban population at a fine scale not previously reported.

In progress

Next steps will include: 1) exploration of GHSL in relation to flood events over time, and assessment of the best way to aggregate and present flood events; 2) use of CHIRPS and storm tracks data to identify particularly high-risk time points and geographic areas for more in depth analysis; 3) spatial mapping and descriptive analysis of vulnerability variables from the census including demographics and housing; 4) development of an indicator of vulnerability using available population and demographic data.

We hypothesize that over time, the built-up area exposed to floods will increase in both inland and coastal areas, with coastal cities potentially growing at a disproportionately high rate.
We hypothesize that there will be significant spatial variation in vulnerability to flood events specifically regarding household quality and water access.
After taking the steps to quantify urban population and infrastructure exposed, we will explore the policy implications of these findings for urban planning, early warning systems and disaster preparedness.

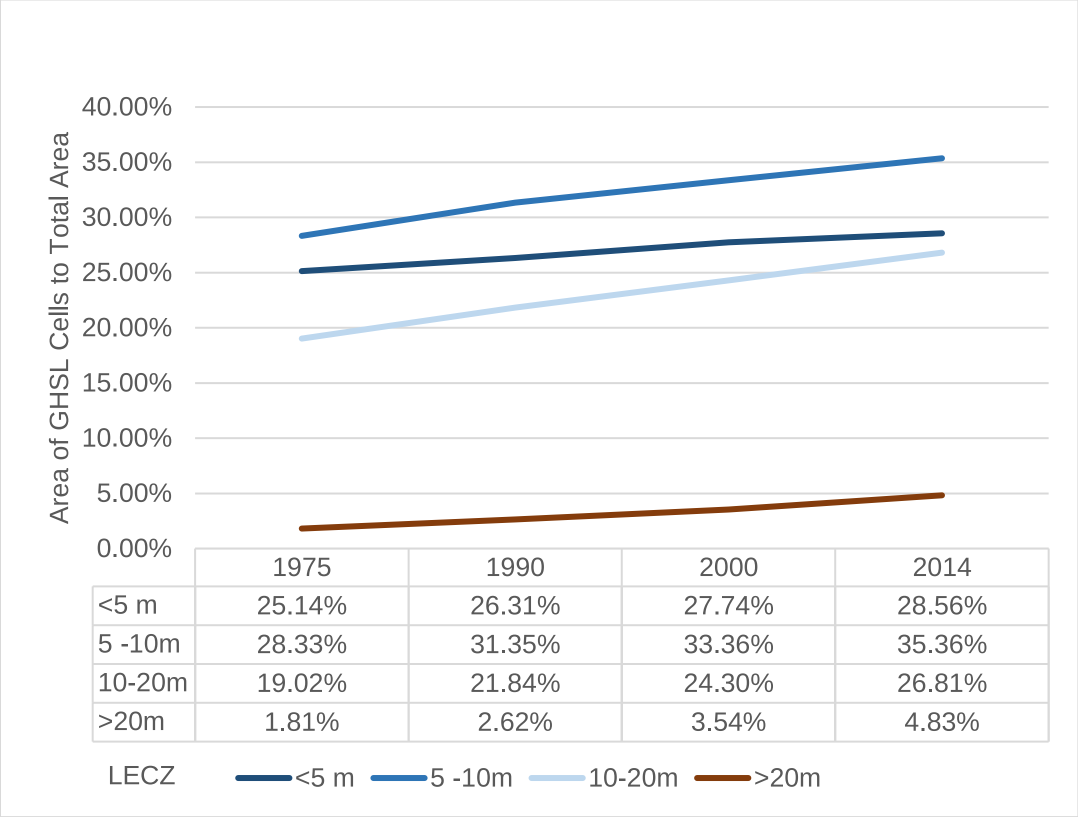
Figures/Tables:

Table 1: Sources, time periods, and measures of datasets used for analysis

Dataset and source	**Time periods**	**Measure (and format)**
Population Data		
Global Human Settlement Layer (GHSL) - http://ghsl.jrc.ec.europa.eu/	1990, 2000, 2015	Gridded data set derived from Landsat imagery. Percentage of 250-meter grid cell covered by structures. Higher built-up percentages likely imply a greater degree of urbanization and economic development (Also available at 38m.)
Settlement-level administrative data corresponding to Indian Census. MLInfoMap (https://www.mlinfomap.com/;http://www.censusindia.gov.in/2011census/)	2011	Tabular and vector data (polygon and points) representing the finest-scale boundaries of administrative data corresponding to the publicly available Indian Census. Approximately 650,000 spatial units.
Environmental Data		
Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) - http://chg.geog.ucsb.edu/data/chirps/	1999-2015	A gridded rainfall time series, combining daily precipitation estimates and weather monitoring station data to quantify precipitation and seasonal drought. Has 0.05 degree resolution.
Dartmouth Flood Observatory - http://floodobservatory.colorado.edu/	1985 – present	Vector data indicating the spatial distribution and frequency of large flood events between 1985 and 2016. A new version of the dataset available 2011-present at finer spatial resolution will be explored.
High-resolution low elevation coastal zone (LECZ) – not yet publicly available; related data product http://sedac.ciesin.columbia.edu/data/set/lecz-urban-rural-population-land-area-estimates-v2/metadataSEDAC	2000	Gridded dataset (90-meters) based on the Shuttle Radar Topography Mission (SRTM) elevation dataset for year 2000. Measures elevation bands per meter between 0-20 meters contiguous to sea-coast.
Passive Microwave Global Flood Detection System data: http://www.gdacs.org/flooddetection/	1997-present	Global Flood Detection System is an operational flood monitoring system using microwave data in near-real time to produce raster datasets of detected flooding anomalies (10x10km pixels)
International best track archive for climate stewardship (IBTrACS) – tropical cyclone data https://www.ncdc.noaa.gov/ibtracs/	1979-present	Collation of multiple sources to construct best-track lines of all unique storms reported to various systems.

Table 2: State level GHSL build up (urbanization) over four time periods (1975, 1990, 2000, 2014) in relation to LECZ categories for coastal districts (Example: Maharashtra State)

LECZ Category	Pop (2011)	Area (km2)	Built up area (KM2)				Built up area / Total Area (%)				Mean Build Up			
			1975	1990	2000	2014	1975	1990	2000	2014	1975	1990	2000	2014
<5 m	5,700,054	1,260	316.81	331.66	349.62	360.00	25.14%	26.31%	27.74%	28.56%	12.36	12.37	13.41	13.47
5 -10m	4,654,752	978	277.19	306.72	326.45	345.96	28.33%	31.35%	33.36%	35.36%	16.42	16.19	16.25	15.83
10-20m	5,645,092	1,807	343.77	394.67	439.20	484.49	19.02%	21.84%	24.30%	26.81%	18.25	17.58	17.32	16.64
>20m	96,217,978	304,051	5,509.93	7,978.82	10,758.30	14,673.54	1.81%	2.62%	3.54%	4.83%	11.44	13.15	14.36	13.82



	1975	1990	2000	2014
<5 m	25.14%	26.31%	27.74%	28.56%
5 -10m	28.33%	31.35%	33.36%	35.36%
10-20m	19.02%	21.84%	24.30%	26.81%
>20m	1.81%	2.62%	3.54%	4.83%

Figure 1: Kolkata district built up area (urbanization) from GHSL in 2000 and 2014 overlaid with LECZ (Panel A, B) and January 2012 flood days (Panel C, D)

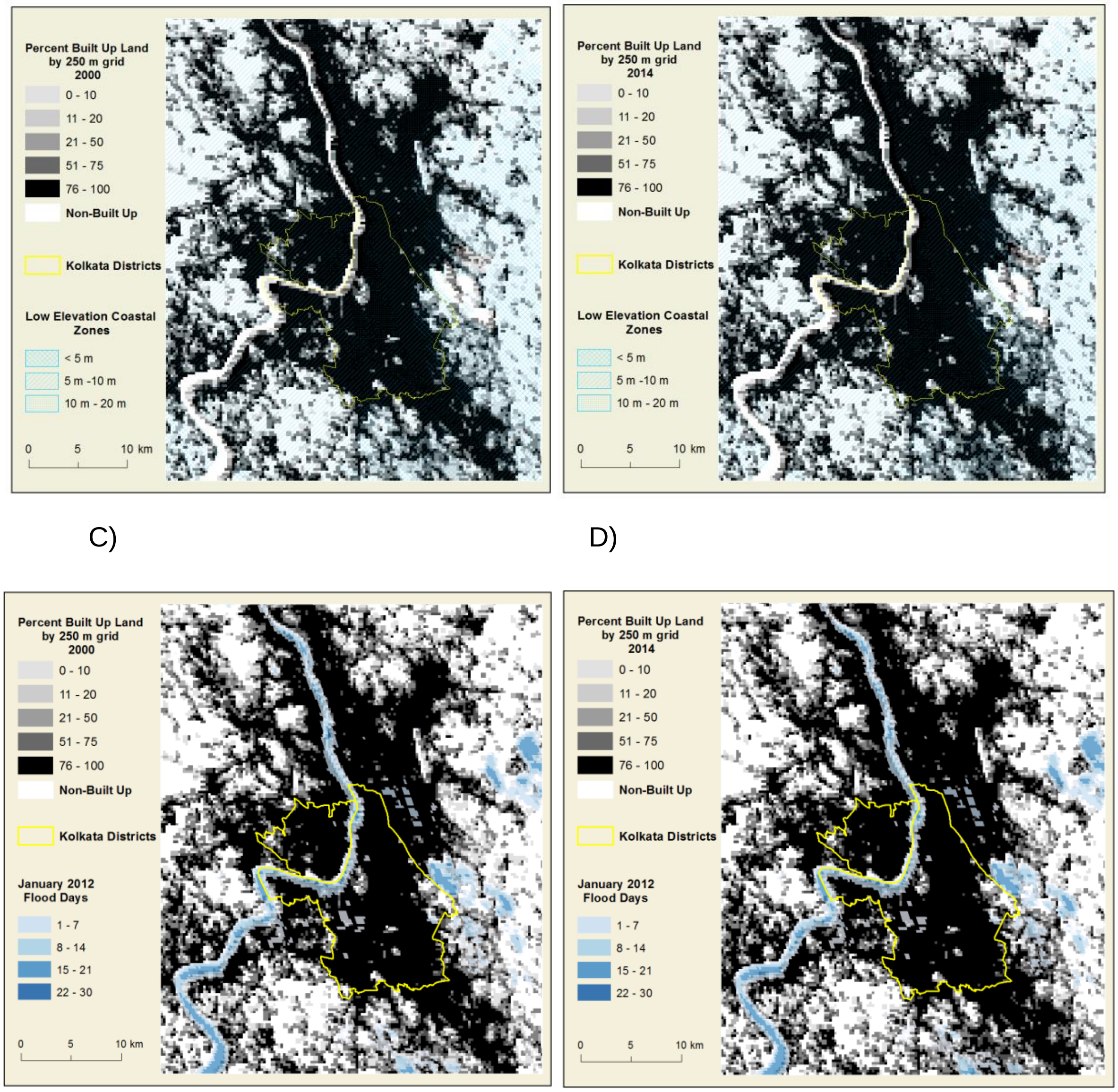
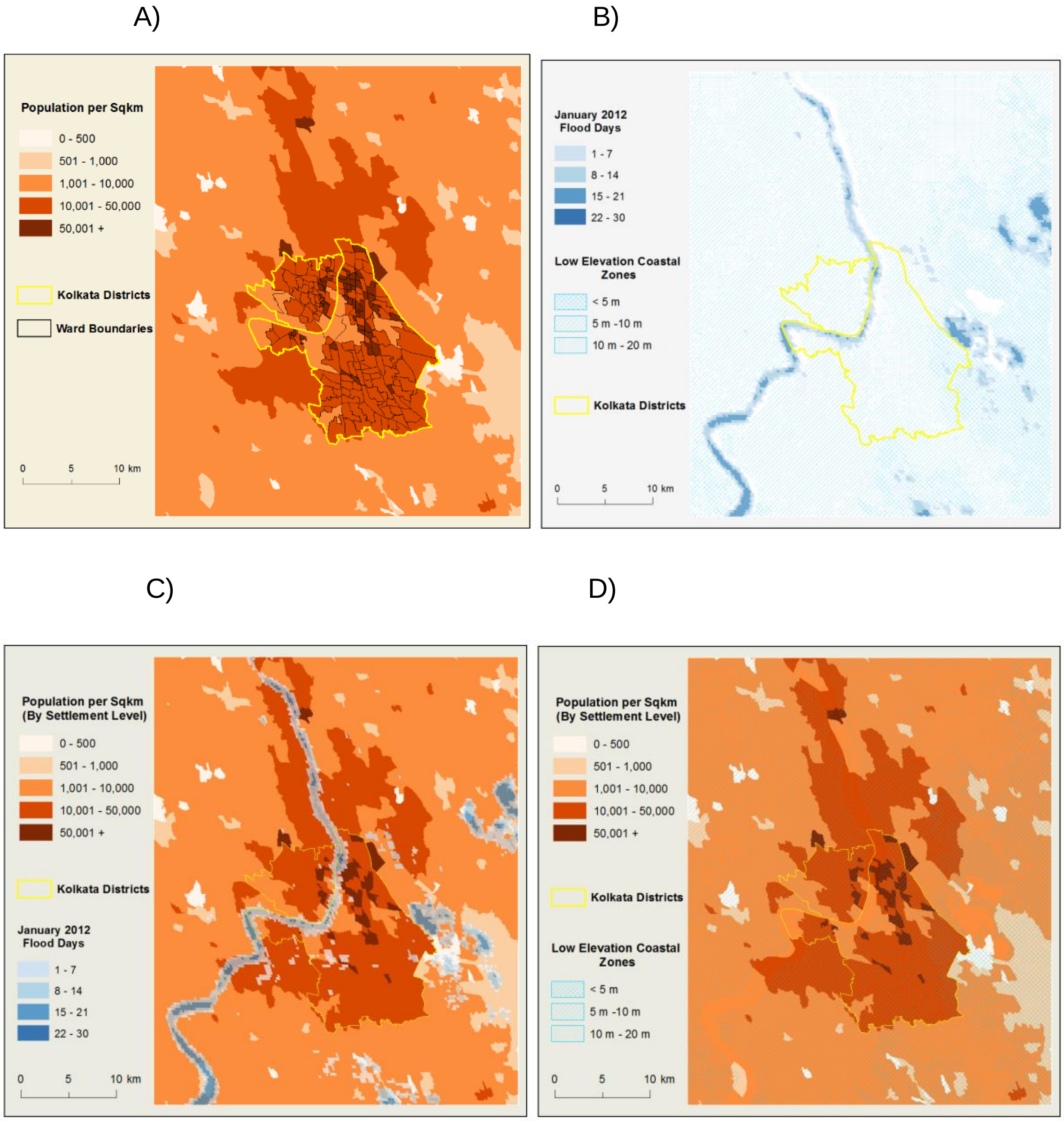

A) B) C) D)

Figure 2: Kolkata Districts: Panel A) ward level population from 2011 census; Panel B) Number of days reporting flooding in January 2012 (DFO reported floods) and LECZ category; Panel C) Overlay of population and flood data; Panel D) Overlay of population and LECZ



References

1. IPCC. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Geneva, Switzerland: IPCC; 2014.
2. Guneralp B, Guneralp I, Liu Y. Changing Global Patterns of Urban Exposure to Flood and Drough Hazards. Global Environmental Change. 2015;31:217-25.
3. Balk D, Montgomery M, McGranahan G, Kim D, Mara V, Todd M, et al. Mapping Urban Settlements and the Risks of Climate Change in Africa, Asia and South America. In: M.Guzmán GM, G. McGranahan, D. Schensul and C. Tacoli., editor. Population Dynamics and Climate Change. New York: UNFPA; London: IIED2009.
4. Sarmah T, Das S. Urban flood mitigation planning for Guwahati: A case of Bharalu basin. 2018.
5. McGranahan G, Balk D, Anderson B. The rising tide: assessing the risks of climate change and human settlements in low elevation coastal zones. Environment and Urbanization. 2007;19:17-37.
6. Neumann B, Vafeidis A, Zimmermann J, Nicholls R. Future coastal population growth and exposure to sea-level rise and coastal flooding - a global assessment. PLoS ONE. 2015;10(3).
7. Nicholls R, Hanson S, Lowe J, Warrick R, Lu X, Long A. Sea-level scenarios for evaluating coastal impacts. Wiley Interdisciplinary Reviews: Climate change. 2014;5:129-50.
8. Nicholls R, Cazenave A. Sea-level rise and its impact on coastal zones. Science. 2010;328:1517-20.
9. Suriya S, Mudgal B. Impact of urbanization on flooding: The Thirusoolam sub watershed – A case study. Journal of Hydrology. 2011;412-413:210-9.
10. Pesaresi M, Ehrlich D, Ferri S, Florczyk A, Freire S, Haag F, et al. Global Human Settlement Analysis for Disaster Risk Reduction. The International Archives of Photogrammetrry, Remote Sensing and Spatial Information Sciences. 2015;XL-7/W3:837.
11. Pesaresi M, Huadong G, Blaes X, Ehrlich D, Ferri S, Gueguen L, et al. A global human settlement layer from optical HR/VHR RS data: concept and first results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2013;6(5):2102-31.
12. Florczyk A, Airaghi D, Ehrlich D, Corban C, Freire S, Kemper T, et al. Community pre-Release of GHS Data Package (GHS CR2018) in support to the GEO Human Planet Initiative. In: Commission E, editor. Luxembourg: Publications Office of the European Union; 2018.
13. Rasch RJ. Assessing urban vulnerability to flood hazard in Brazilian municipalities. Environment and Urbanization. 2015;28(2):145-68.