Universidade Eduardo Mondlane
Faculdade de Ciências
Departamento de Matemática e Informática

Presenter: Silvino Pedro Cumbane, PhD
(Researcher and Lecturer)

Courses

- **Bachelor:** Geographic Information Sciences
  - GIS, Remote Sensing, Surveying, Photogrammetry, Cartography, Geodesy, Spatial Data Base, etc.
- **Master:** Master in GIS for Sustainable Development
  - GIS, Remote Sensing, Spatial Analysis, Spatial Database, GIS oriented Programming, WebGIS, etc.
- **Short Term Courses:** GIS, Cartography, GNSS Data collection and Processing.

GISciences Lab

- 15 Desktop
- ArcMap License
- Surveying Equipment
- GNSS receiver
- GPS receiver
- …

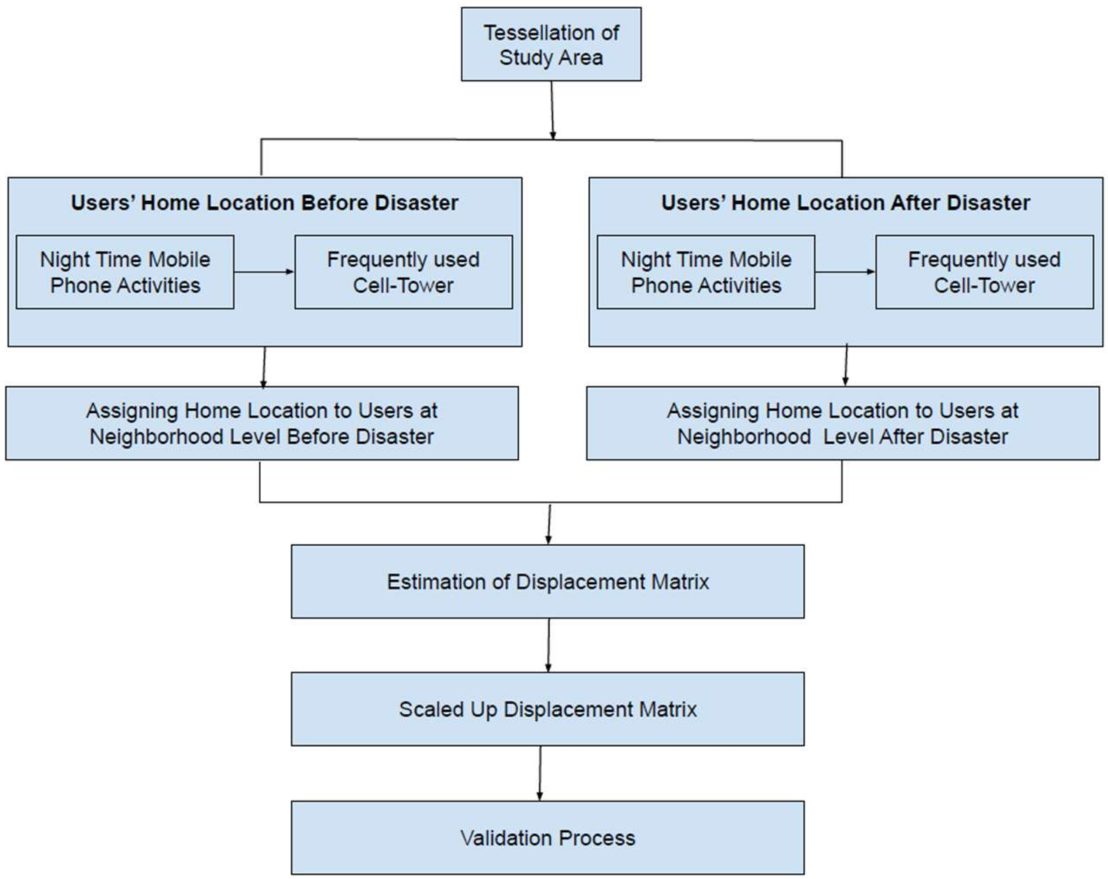
Projects (Data Collection)

- Electrical Assets & Customers Data collection and Processing (supported by EDM) - Ongoing
- Collaborative Research on Mobility Patterns of Vulnerable Populations in Mozambique (supported by Japanese Government, 2020) - Completed
- Collaborative Research on Human Mobility Research. The University of Tokyo and Eduardo Mondlane University: Dynamic Census Project (supported by Melinda and Bill Gates Foundation, 2017) - Completed
- Collaborative Research on Establishment of GNSS Base Station for Joint Research at Eduardo Mondlane University (supported by Center for Spatial Information Science, The University of Tokyo, 2016-2017) - Completed

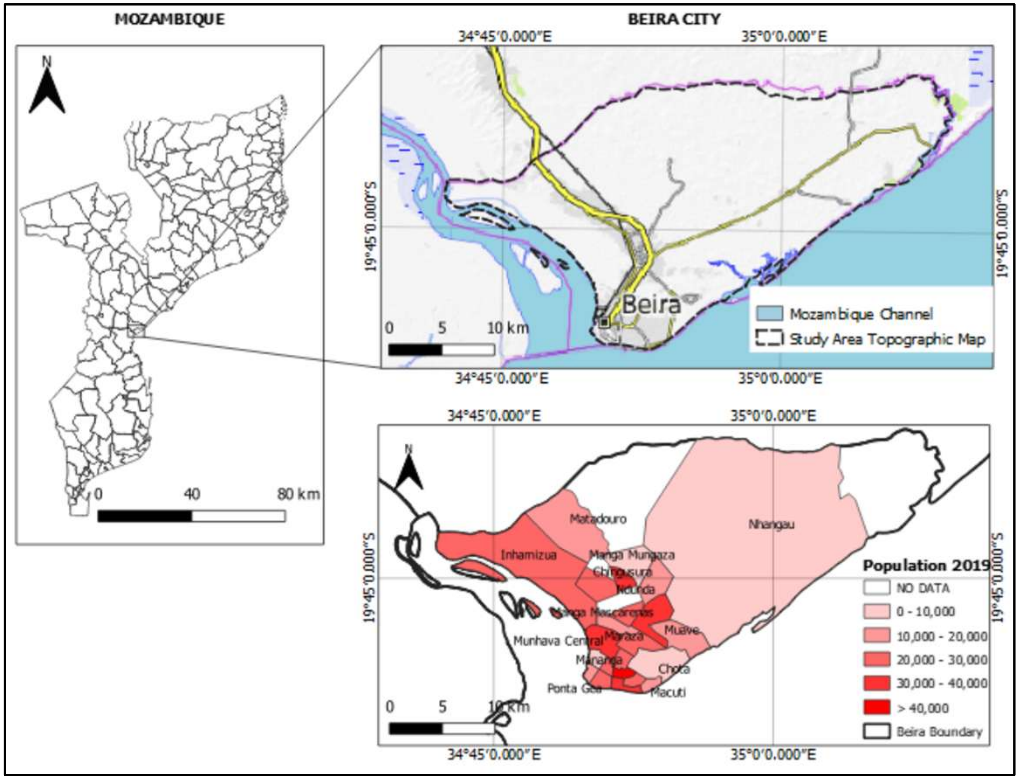
Projects (Applications)

- Using Mobile Phone Data for Disaster Response

Proposed Method



Study Area



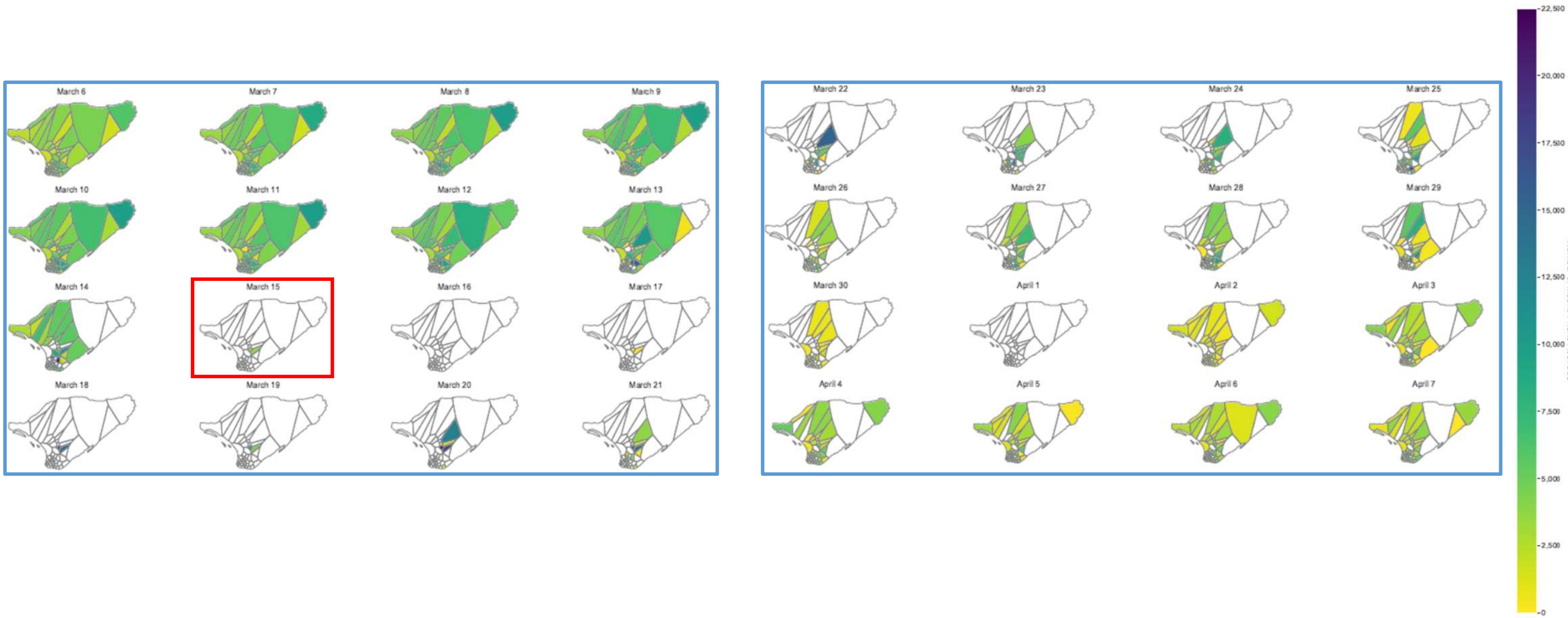
General data

- Area = 631 km2
- Population = 530,604 (2020)
- Neighborhoods = 26

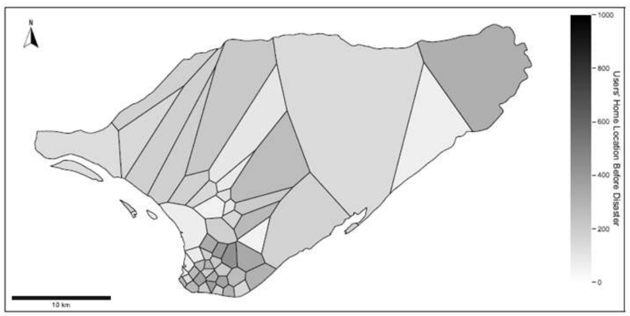
Cyclone Idai in the Central Zone

- Affected 3 million people
- Destroyed around 239,731 homes
- 1,006 deaths (603 in Mozambique, 344 in Zimbabwe and 59 in Malawi)
- 1.77 million hectares of crops destroyed

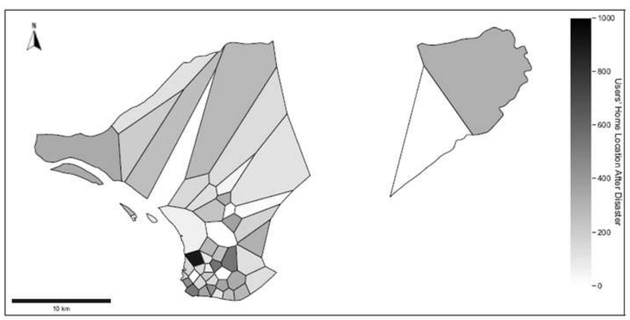
Spatial Distribution of Mobile Phone Activities



Cell-Tower and Neigbourhood-based Home Location

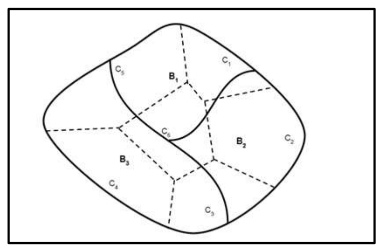


(a)

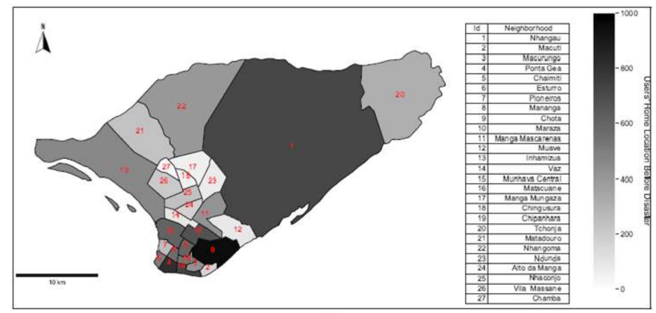


(b)

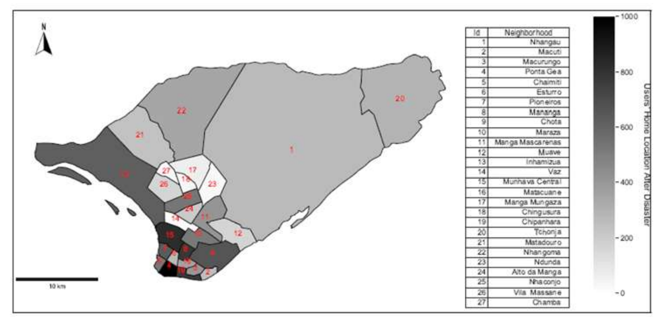
Areal Interpolation Method



$$nu_i = \frac{a_i}{A_t}.u_t$$



(a)

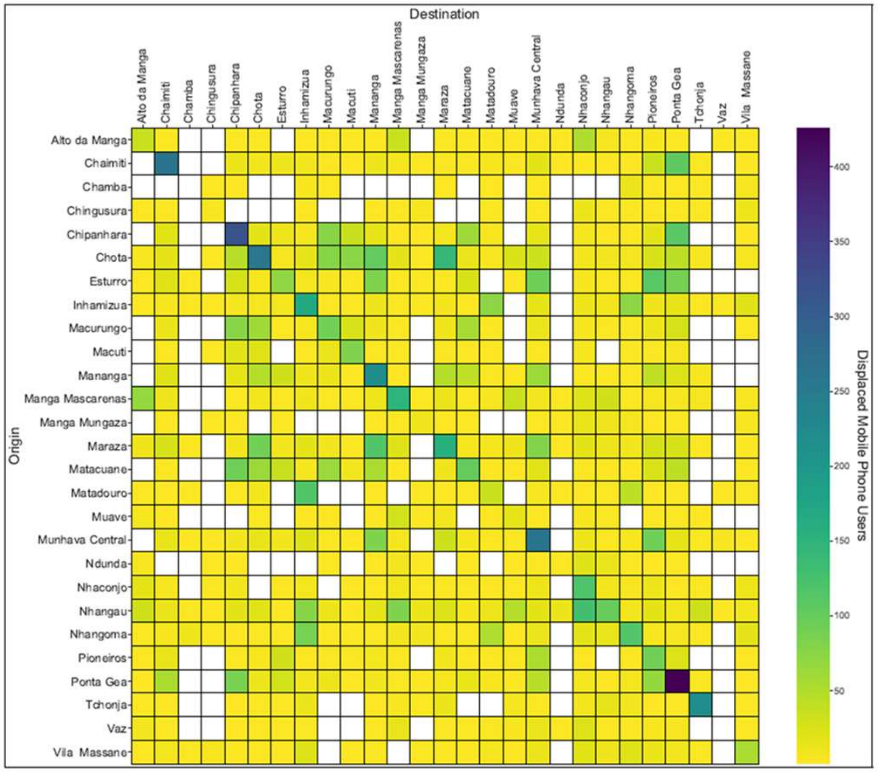


(b)

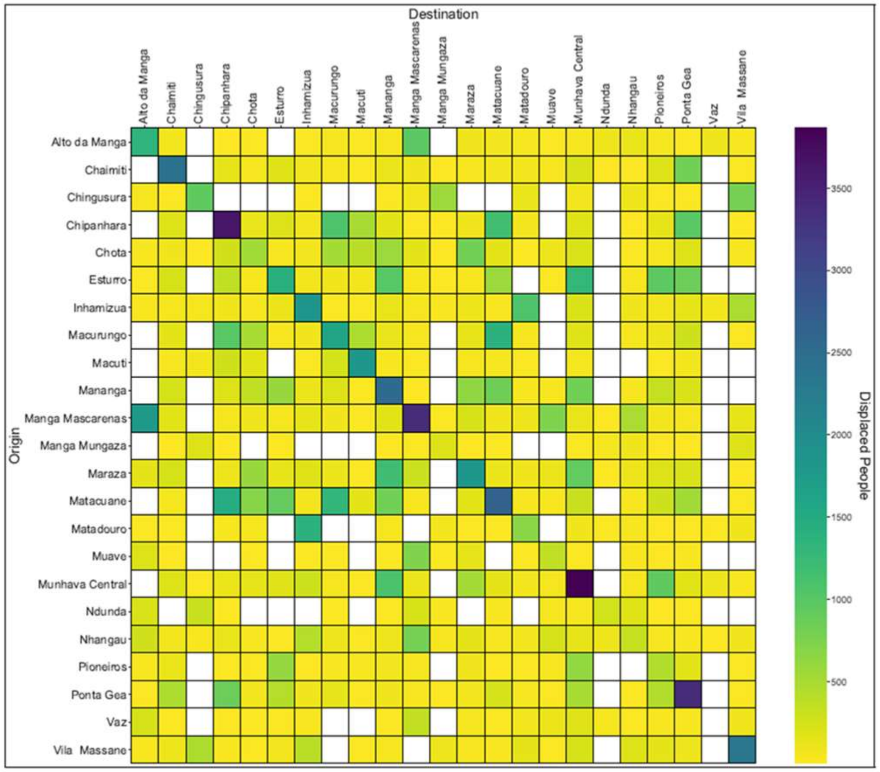
- Increase and decrease of the number of mobile phone users after disaster due to move-in and move-out, respectively.
- In Nhangau neighborhood (1), there is a drastic change in the number of mobile phone users who were assigned to this area.

Displacement Matrix

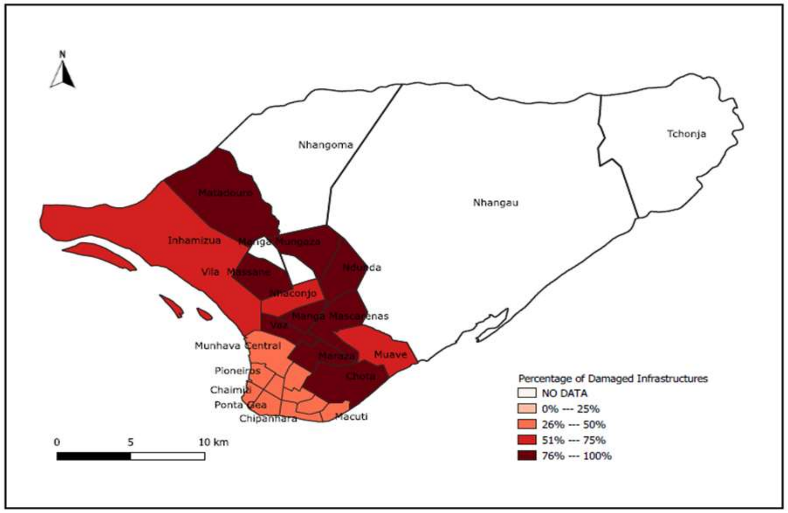
Mobile Phone Users Displacement Matrix

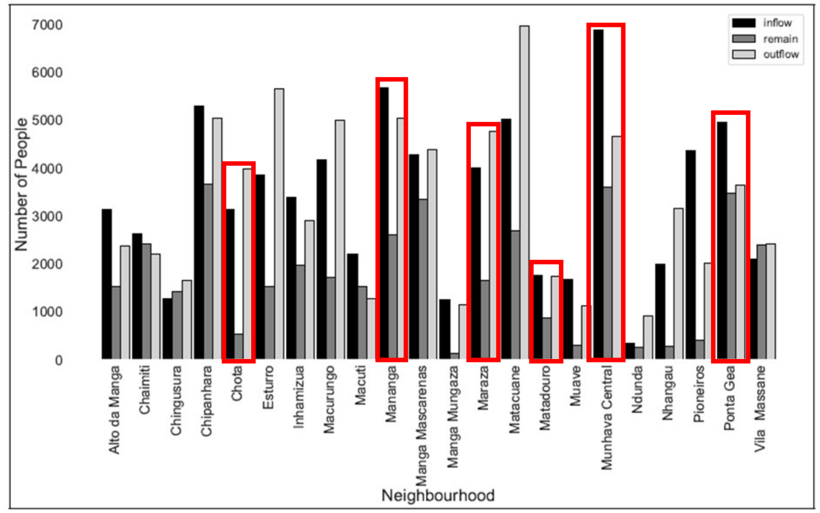


Actual Population Displacement Matrix



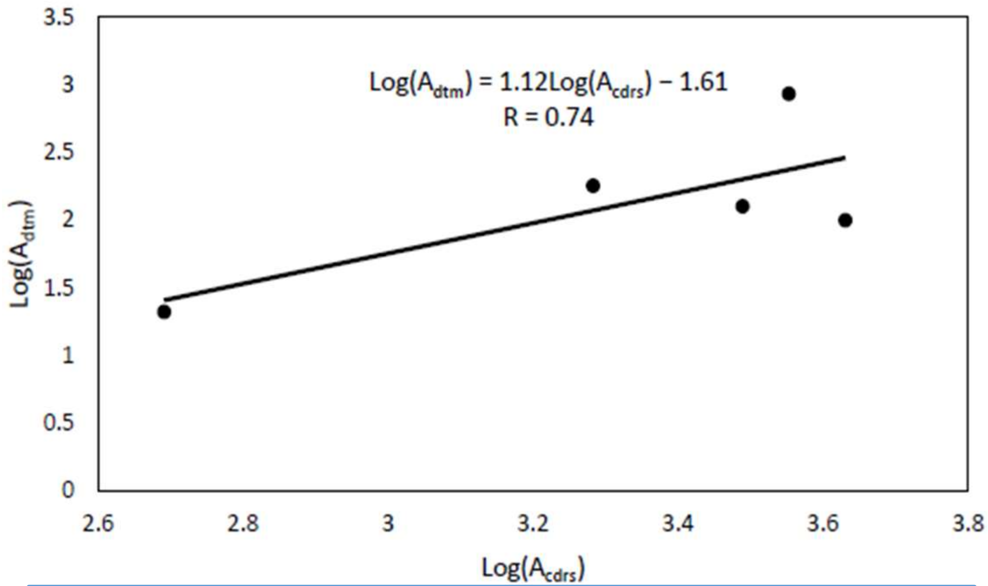
Validation 1 – CDR vs RS Damaged Infrastructures





- Ponta Gea and Munhava were **less affected,** and therefore, have **higher number of inflow** compared to outflow.
- Chota, and Maraza were **highly affected,** and therefore, have **higher number of outflow** compared to outflow.
- Matadouro desagree with the CDR, i.e., highly affected but still have **balanced number of inflow and outflow**.

Validation 2 – CDR vs DTM Survey



If one can trust the model derived from the relationship between arrivals in each area from CDR and DTM, actionable knowledge can be extracted that can be used to predict the demand for disaster support in areas where no information is available.

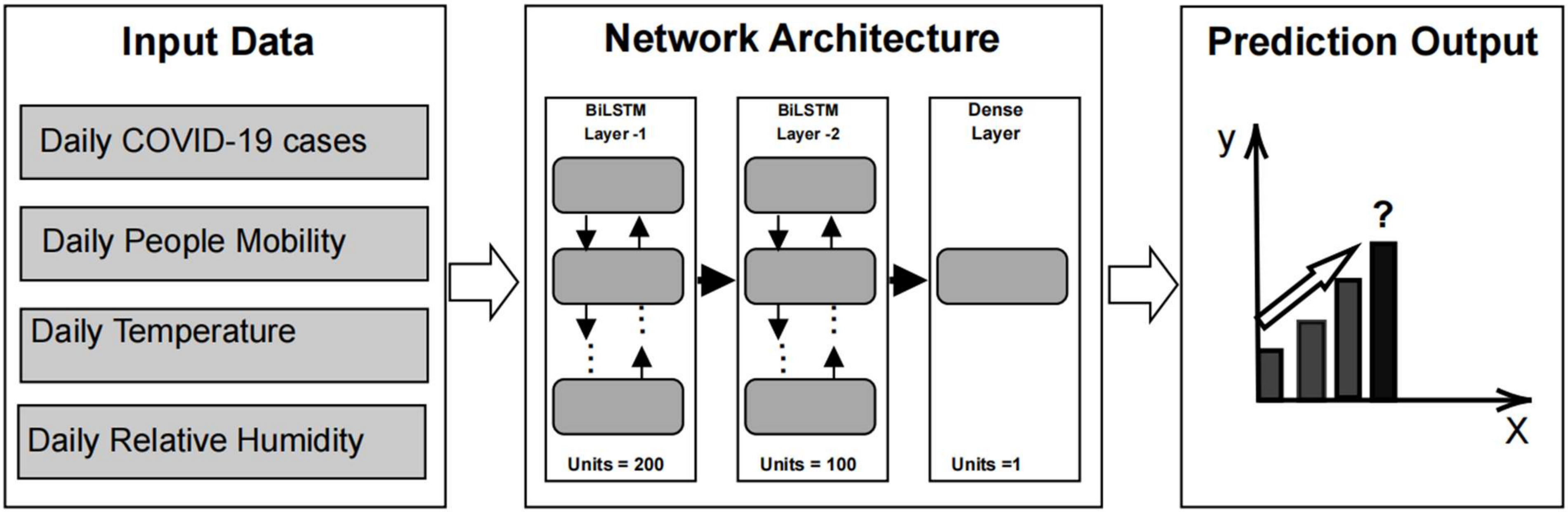
Reasons for drawback

- Limited DTM data sample
- DTM data collected 1 month later
- The survey was conducted only in official shelters
- CDR capture mobility of everyone who used mobile phone before and after disaster while DTM capture mobility of only people that moved to shelters

Projects (Applications)

Deep Learning-based Approach for COVID-19 Response

Network Architecture

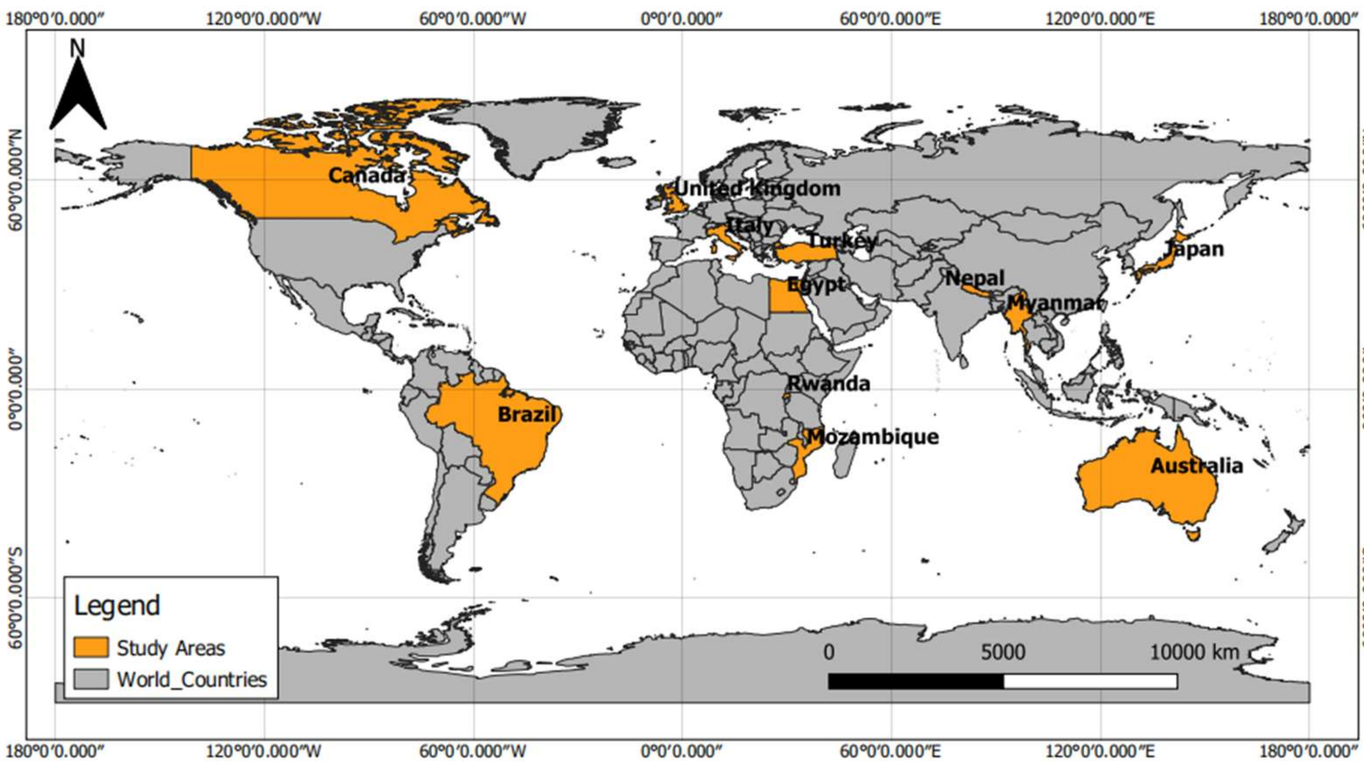


Optimization: Adam Optimizer
Learning rate: 0.1
Loss function: MSE

Training dataset: 80%
Testing dataset: 20%
Batch size: 64

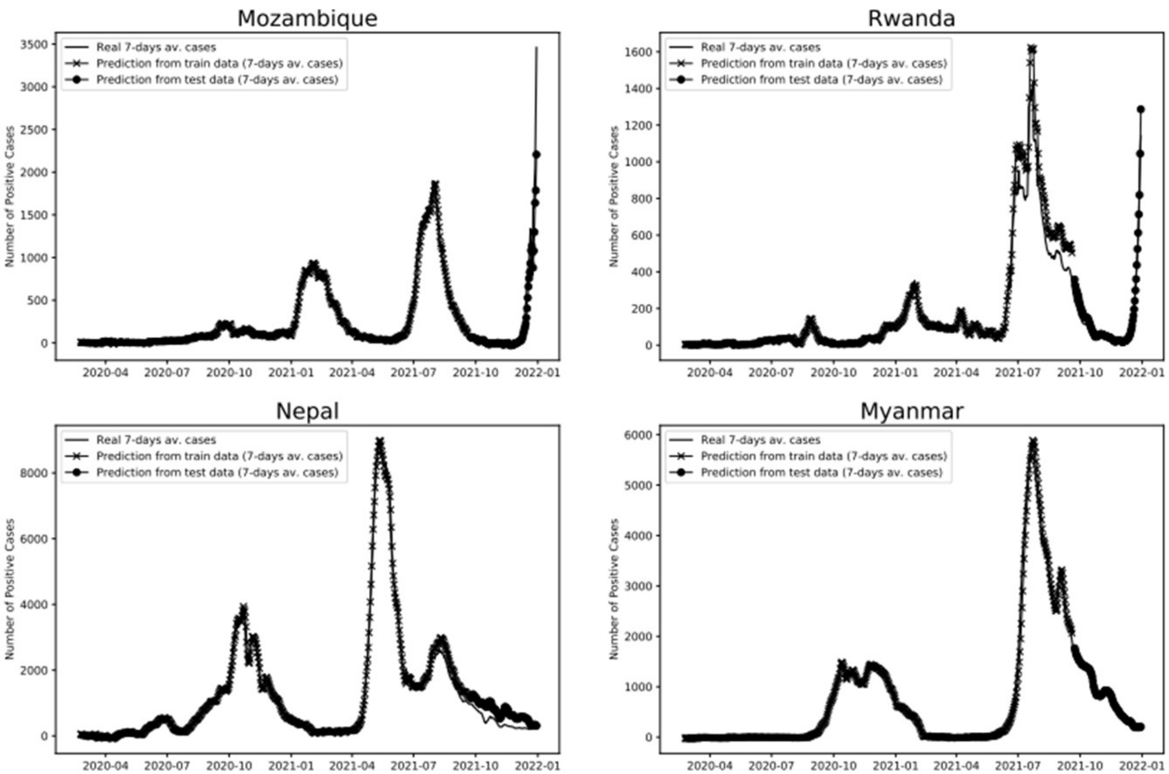
Epoch: 200
Val. Metric: RMSE
Val. Metric: ARE

Study Areas



180°0'0.000"
120°0'0.000"W
60°0'0.000"W
0°0'0.000"
60°0'0.000"E
120°0'0.000"E
180°0'0.000"
60°0'0.000"N
0°0'0.000"
60°0'0.000"S
N
Canada
United Kingdom
Italy
Turkey
Egypt
Nepal
Myanmar
Japan
Brazil
Rwanda
Mozambique
Australia
Legend
Study Areas
World_Countries
0
5000
10000 km

Multilayer BiLSTM vs Multilayer LSTM



Country	RMSE - Proposed BLSTM Model	RMSE-Multi-layer LSTM Model
Mozambique	221.2	326.49
Rwanda	38.53	68.83
Nepal	344.95	447.28
Myanmar	40.45	51.45

Country	Improvement of BiLSTM (%)
Mozambique	47.60
Rwanda	78.64
Nepal	29.67
Myanmar	27.19

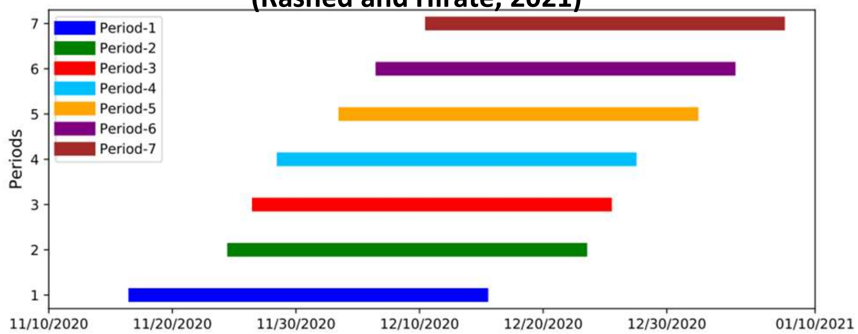
Multi-layer BiLSTM vs Other Models (Country Level)

Country	RMSE BiLSTM	RMSE Multi-layers LSTM	RMSE LSTM	RMSE ARIMA	% improvement BiLSTM
Turkey	140.80	375.39	710.5	1892.3	166.61
Italy	83.67	202.71	229.4	566.8	142.28
Australia	14.05	20.67	148.86	637.01	47.08
Brazil	1125.20	1626.86	10636.22	94302.48	44.58
Canada	177.99	326.27	1641.9	8863.83	83.31
Egypt	252.15	381.26	570.07	2067.55	51.21
Japan	16.97	18.97	236.79	6543.16	11.76
UK	956.57	1131.74	2198.53	24835.98	18.31

The proposed multi-layer BiLSTM model outperformed the second-best model (Multi-layer LSTM) for at least **12%** Japan) and **167%** at most (in Turkey).

Multi-layer BiLSTM vs Other Models (City Level)

(Rashed and Hirate, 2021)



BiLSTM – Proposed Multi-layer BiLSTM model

LSTM – Multi-layer LSTM Model

GCF – Google Cloud Forecast

LSTMall - LSTM-based model considering all the features

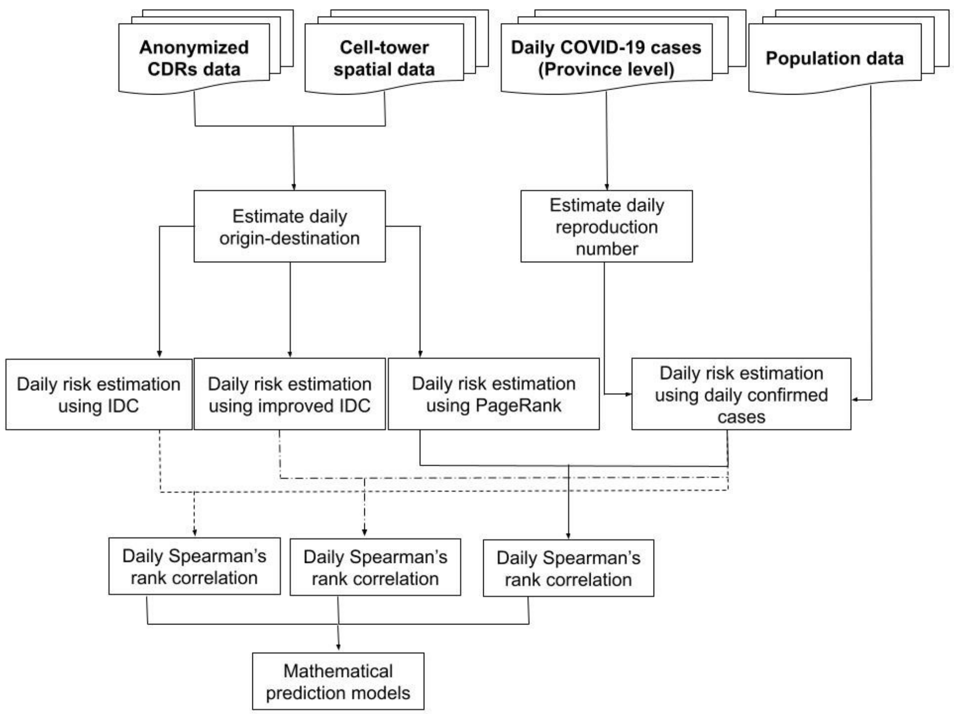
LSTMm - LSTM-based model considering only the mobility data

City	Method	T-1	T-2	T-3	T-4	T-5	T-6	T-7	Aver.
Tokyo	$BiLSTM$	**0.108**	**0.076**	**0.072**	**0.065**	**0.061**	**0.067**	**0.098**	**0.078**
	$LSTM$	0.139	0.100	0.093	0.084	0.071	0.086	0.117	0.098
	GCF	0.289	0.467	0.287	0.330	0.300	0.290	1.943	0.558
	$LSTM_{all}$	0.267	0.169	0.191	0.222	0.229	0.229	0.278	0.226
	$LSTM_m$	0.233	0.221	0.191	0.198	0.210	0.234	0.249	0.219
Aichi	$BiLSTM$	**0.113**	**0.130**	**0.148**	0.172	0.199	0.240	0.298	0.186
	$LSTM$	0.139	0.157	0.180	0.204	0.227	0.277	0.354	0.220
	GCF	0.467	0.301	0.368	0.395	1.102	0.311	0.376	0.474
	$LSTM_{all}$	0.201	0.148	0.189	**0.162**	0.150	**0.157**	**0.187**	**0.171**
	$LSTM_m$	0.184	0.149	0.172	0.174	**0.149**	0.174	0.232	0.176
Osaka	$BiLSTM$	**0.089**	**0.077**	**0.086**	**0.106**	**0.110**	**0.133**	**0.147**	**0.107**
	$LSTM$	0.104	0.092	0.096	0.116	0.112	0.135	0.148	0.114
	GCF	0.318	1.464	1.334	1.477	0.802	0.505	0.533	0.919
	$LSTM_{all}$	0.234	0.148	0.159	0.141	0.130	0.135	0.190	0.162
	$LSTM_m$	0.283	0.171	0.166	0.151	0.132	0.141	0.214	0.180
Hyogo	$BiLSTM$	**0.110**	**0.075**	**0.075**	**0.078**	**0.090**	**0.121**	**0.120**	**0.096**
	$LSTM$	0.129	0.079	0.079	0.081	0.119	0.156	0.125	0.110
	GCF	0.470	1.892	1.740	1.532	0.520	0.425	0.336	0.988
	$LSTM_{all}$	0.282	0.286	0.244	0.229	0.248	0.216	0.336	0.263
	$LSTM_m$	0.330	0.289	0.234	0.248	0.207	0.263	0.360	0.276
Kyoto	$BiLSTM$	**0.178**	**0.135**	**0.140**	**0.144**	**0.234**	**0.193**	**0.127**	**0.164**
	$LSTM$	0.264	0.225	0.229	0.241	0.409	0.312	0.194	0.268
	GCF	0.538	0.749	0.568	0.601	0.758	0.732	0.363	0.616
	$LSTM_{all}$	0.429	0.529	0.564	0.549	0.283	0.342	0.241	0.419
	$LSTM_m$	0.448	0.494	0.519	0.564	0.601	0.615	0.540	0.540
Fukuoka	$BiLSTM$	**0.238**	**0.171**	**0.184**	**0.186**	**0.269**	**0.218**	**0.134**	**0.200**
	$LSTM$	0.221	0.177	0.195	0.204	0.278	0.242	0.166	0.211
	GCF	0.484	0.665	0.506	0.663	0.535	0.443	0.353	0.522
	$LSTM_{all}$	0.399	0.420	0.461	0.428	0.429	0.341	0.245	0.389
	$LSTM_m$	0.519	0.427	0.388	0.358	0.335	0.334	0.375	0.391

Projects (Applications)

Spatio-temporal exposure risk estimation for COVID-19 using social network analysis and mobile phone data

Method



Weighted in-degree centrality, also known as node strength, calculates the sum of the weights of all incoming edges to a node (e.g., province) in a network. It represents the total flow of people into a location.

$$z_i = \sum_{j=1}^{N} w_{ij}$$

Improved in-degree centrality is built upon weighted in-degree centrality by incorporating a tuning parameter k to balance the importance of the number of connections (in-degree) and the strength of those connections (weighted in-degree). Based on simulation results, in this study k is set to 0.8.

$$c_i = d_i * \left(\frac{z_i}{d_i}\right)^k \text{where:} \qquad d_i = \sum_{j=1}^{N} n_{ij}$$

Weighted PageRank goes beyond direct influence and considers the transitive influence of nodes in a network. A location's PageRank score is influenced not only by the number of people visiting it but also by the PageRank scores of the locations from which they originate.

$$PR(P_i) = (1-d) + d \sum_{j \in C(P_i)} PR(P_j) \cdot W_{ji}^{in} \cdot W_{ji}^{out}$$

COVID-19-based Risk Score

The risk score for a given community is derived as the average number of people that are likely to get infected in the next 24 hours by currently infectious individuals divided by the current number of susceptible people

$$r_t = \frac{I(t) * R_t}{P_i}$$

Where:

$I(t)$ - daily new confirmed infectious cases;

R_t - time-varying reproduction number;

P_i – Population of region *i*.

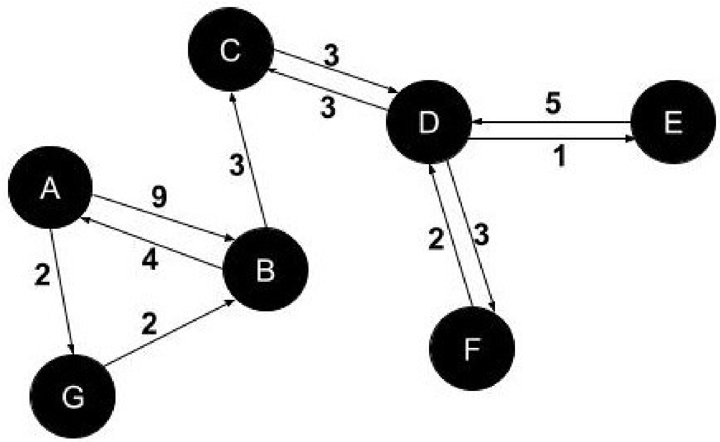
R_t Estimation (*Robert Koch Institute method)*

The effective reproduction number R(t) is calculated as the sum of new reported cases during four consecutive days divided by the sum of new reported cases during four consecutive days prior to the days used in the denominator.

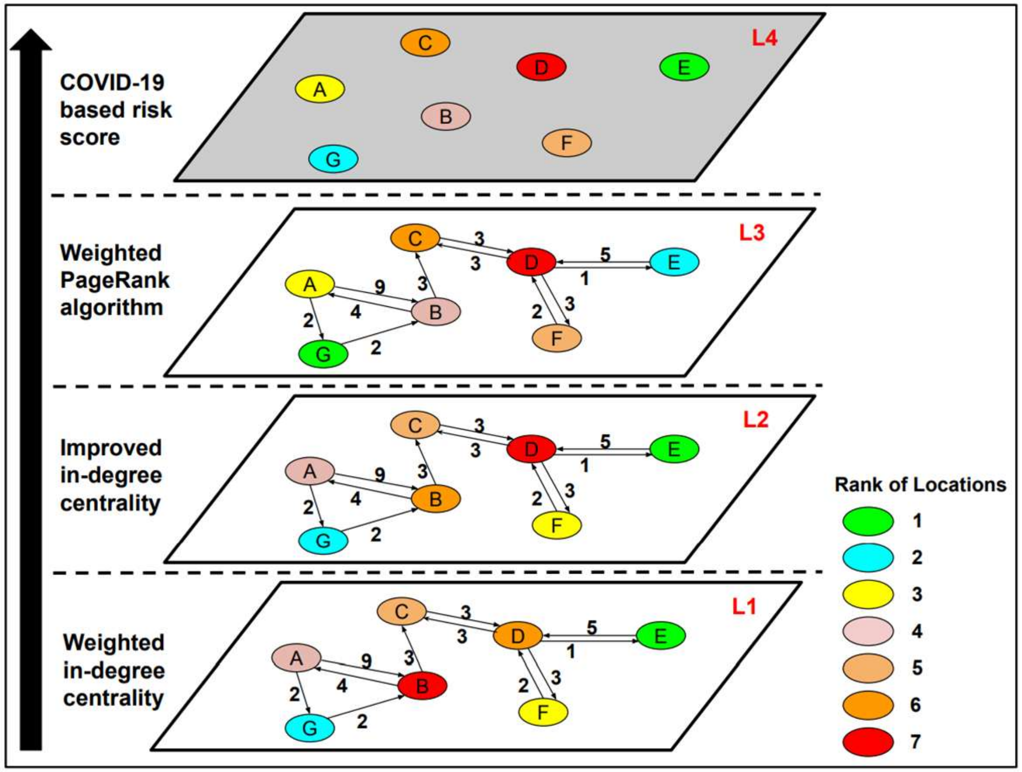
1	2	3	4	5	6	7	8
63	28	41	64	138	239	156	107
			196				640
							R(t) for day 8: 640/196=3.04

R(t) for day 8: 640/196=3.04

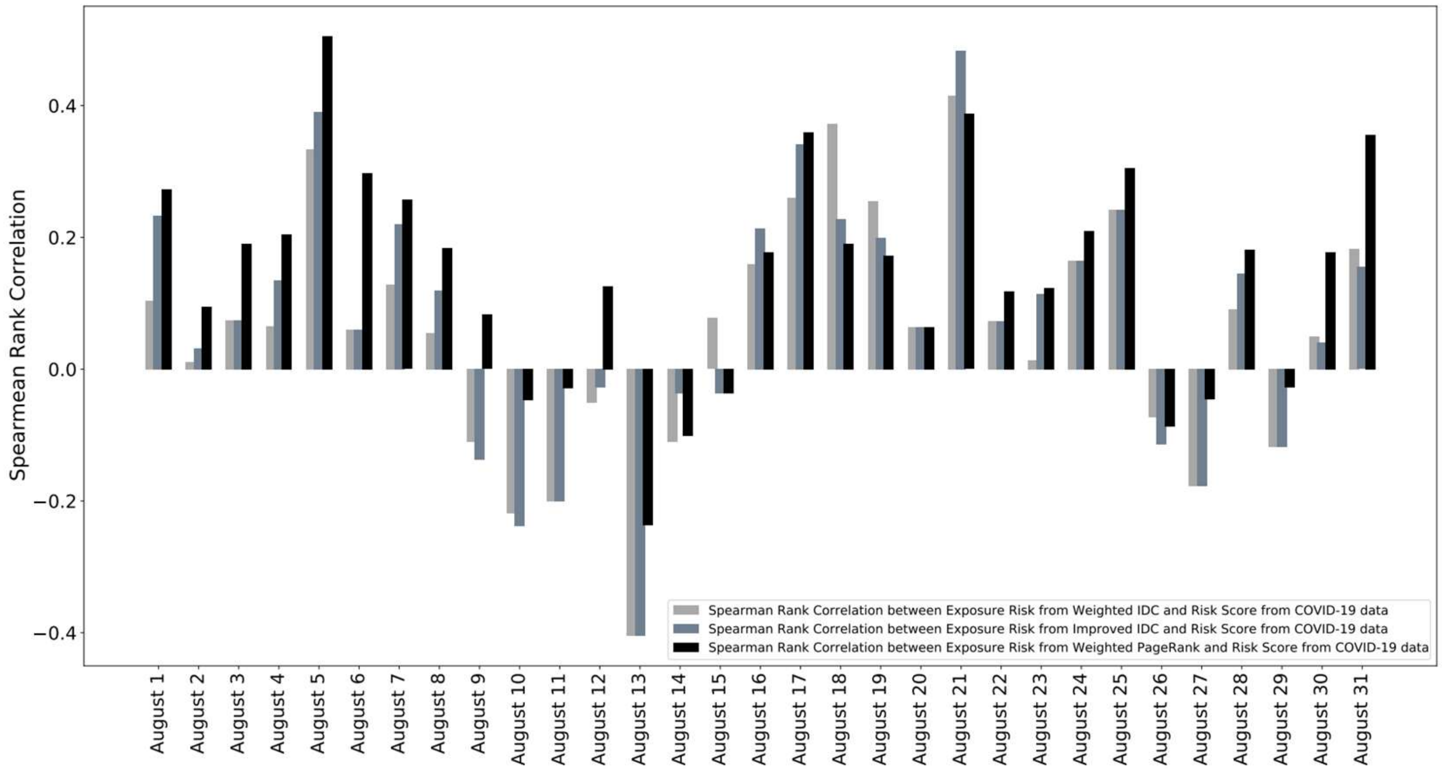
How does the approach work?



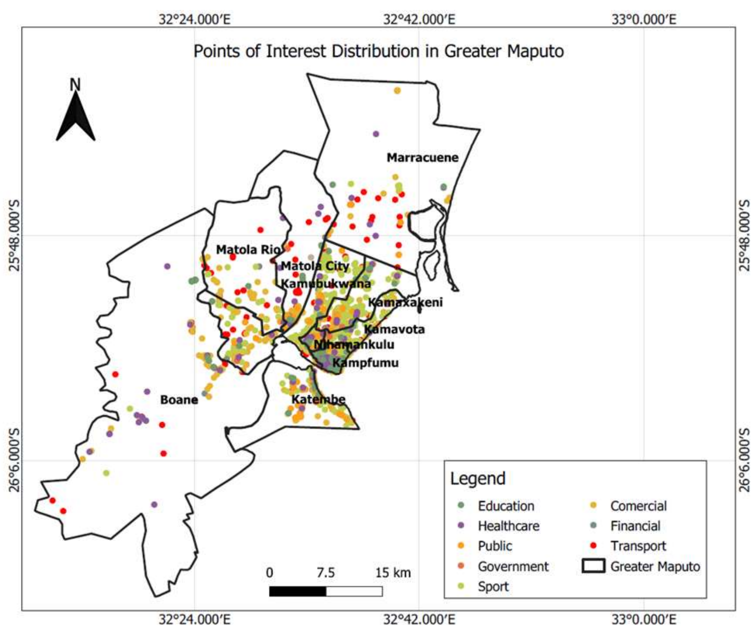
location	widc_rank	i_idc_rank	w_pagerank	widc_rank	i_idc_rank	PRrank
A	0.108	0.106	0.068	4	4	3
B	0.297	0.273	0.096	7	6	4
G	0.054	0.061	0.032	2	2	1
C	0.162	0.168	0.195	5	5	6
D	0.270	0.274	0.381	6	7	7
E	0.027	0.035	0.068	1	1	2
F	0.081	0.084	0.160	3	3	5

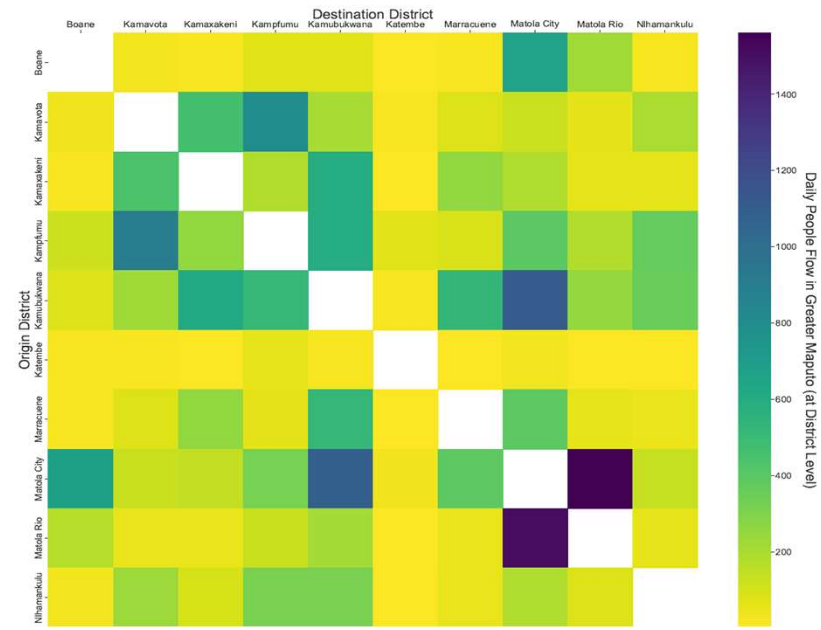


Mobility-based Risk Score vs COVID-19 Risk Score



Greater Maputo Area (GMA)

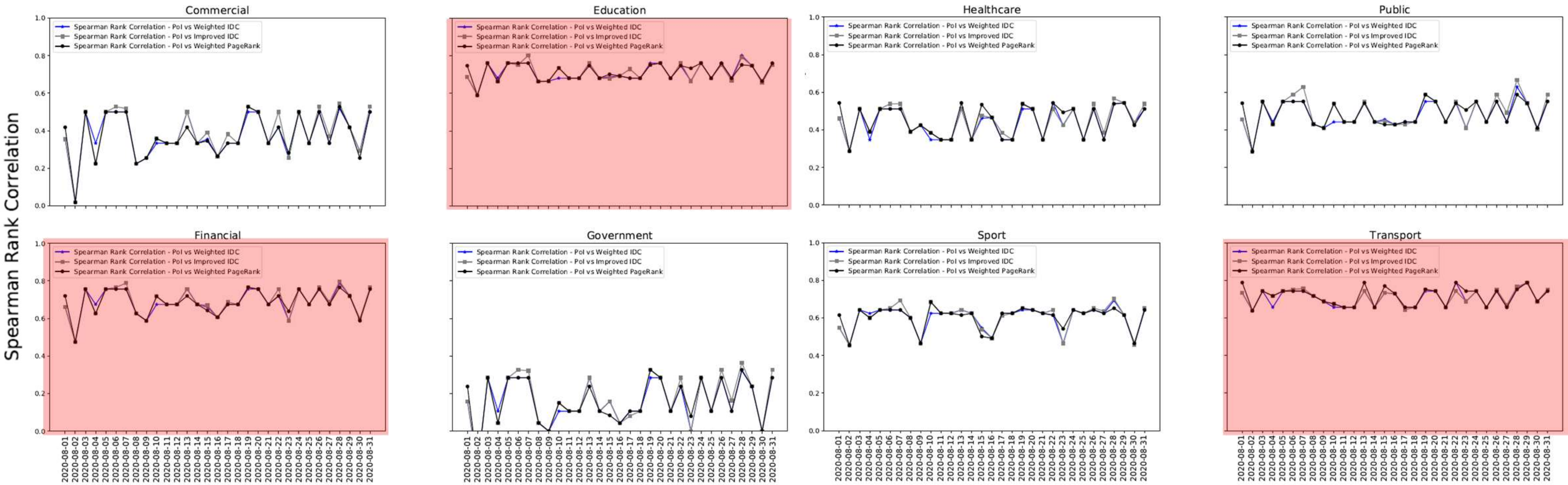




- GMA is located in southern Mozambique.
- Composed of 10 districts.
- It has 2.2 million inhabitants.
- The average population density per hectare is 18 people.

- The highest people flow is registered between Matola city and Matola-rio .
- There are also considerable number of trips between: **Matola city and Boane**, **Kamavota and Kampfumu**, **Kamubukwana and Marracuene**, **Marracuene and Matola ciy**, and **Matola city and Kampfumu**.

Comparison



- The distribution of PoIs related to education, finance, and transport are good indicators of exposure risk across all three centrality measures.
- The lowest correlation is registered when comparing the Government PoI with the daily exposure risk estimated using the three measures.
- The remainder of the PoI have low to high medium correlation with the three centrality measures.

Challenges

- Data Access
  - There is no data privacy/sharing policy in place in Mozambique.
  - The process to get the data is time consuming and many time we cannot delivery the solution timely.
- Collaboration with Public and Private Sector
  - MoU signing between institutions take long time.

ADE is very important partner for data provision for different research projects.

CAP is necessary and urgent.