Land-Use Microsimulation Model for Livelihood Diversification after the 2010 Merapi Volcano Eruptions

Martin Garcia-Fry[1*], Osamu Murao[2], Syamsul Bachri[3], Luis A. Moya[4]

Affiliations
1. Department of Architecture and Building Science, Graduate School of Engineering, Tohoku University, Sendai, Japan.
2. International Research Institute of Disaster Science, Tohoku University, Sendai, Japan.
3. Department of Geography, Faculty of Social Science, State University of Malang, Yogyakarta, Indonesia.
4. Japan-Peru Center for Earthquake Engineering Research and Disaster Mitigation, National University of Engineering, Lima, Peru.

*Corresponding author: garcia.fry.martin.p2@dc.tohoku.ac.jp

Data sample and models in Mendeley Data, v.3: http://dx.doi.org/10.17632/t9p23k3pyn.3

Supplementary Material

This file operates each step of the Operational Land Use and Transport Microsimulation (OLUTM) model corresponding to Chapter 5 of the manuscript. The compact modules described herein use seven data inputs to process data which is then used in the next modules, thus integrating steps of the model within the process. The model interface has three main environments and three data support systems to output datasets in visual representations. The transport component of the model simulates travel accounting for activity-based travel scheduling. The model's first integration of data holds the Geographic and the Socioeconomic Environments. While a different level of integration occurs in the Settlement Environment. Finally, the resulting scenario is packed with numerical data and mapping.

5.1 Geographic Environment

5.1.1 Territorial Analysis

The last administrative census tract available for the Cangkringan sub-district was in 2017;
2010-2017 (Growth Rate: 0,5%) found in [Dataset 1].

Data Input 01
Area: 48 sq/m (squake-meter)
Growth Rate (2010-2017) census: 0,5%
Starting Population Count (2016): 29,321 persons
Population for Evaluation (2017): 29,456 persons

Growth Projection for 2018:
29,456 – 29,321 = 135 (0,46%)
0.0046*(29,456) =135.62
29,456+135.62 = 29,591.62
Total: 29,592 persons*
*Note: This value was confirmed by projections made by the Indonesian Statistical Census Bureau (2010-2035) [Dataset 2]. **https://slemankab.bps.go.id/statictable/2019/07/09/520/distribusi-dan-kepadatan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2018.html**

Growth Projection for 2019:
(2018) 29,592 – (2017) 29,456 = 136 (0,46%)
0.0046 x (29,592) = 147.96
29,592 +148= 29,740
Total: 29,740 persons
Density: 619 / sq. km (square-kilometer)

The latter procedure is summarized with the following equations (i) and (ii) for n number of years;

$$n = (y - y_s) - 1 \quad (1)$$

$$\int(n) = \left[\left(\frac{(P_l - P_s)}{P_s}\right) P_l\right] + P_l \quad (2)$$

Where:
y =Target year (desired future population count);
y_s =Starting year (the last administrative census count for a region, district, or sub-district)
P_l =Last population count;
P_s =Starting population count;

Growth Projection (2030):
n = (2030-2019) + 1
n=10
$f(2030)^{10}$: [($\frac{(29,740 - 29,592)}{29,592}$) x 29,740] + 29,740 = 29,888.7, iterate this function 10 times.

$f(2030)^{10}$=31,302 persons

5.1.2 Density Redistribution

Cangkringan Sub-District
-Growth Rate: 0,5% (see growth estimations sub-section 5.1):
-Area: 48 km^2
-Population 2019: 29,740
-Density in 2019: 619 / km^2
-Population 2030: 31,302 persons
-Density in 2030: 652 / km^2

Ancillary Data

Emergency planning for disaster-affected areas requires population distribution counts for density that is hard to come by in remote locations, especially after a disaster or human displacement event. We used planet-OSM building polygon data, an open-source product that provides world-wide geographic datasets (**https://planet.openstreetmap.org/**) [Dataset 3] as ancillary data for a dasymetric redistribution of the smallest census-based enumeration regarded as a trustworthy source published annually by the Cangkringan sub-district.

Dasymetric Redistribution

Density weights were calculated for a sq./km spatial resolution of the administrative area using QGIS 3.4, Google Earth Engine, and EarthSat GeoCover Land Cover Thematic Mapper from MDA Federal Inc. adapted from [1, 2] to provide us with a spectrum of global land cover designations within our target district.

A Target Density Weight (TDW) can be dasymetrically calculated by gathering building footprint areas for all raster cells comprising a census area using a population count less than 10 years old (step 5.1.2; 2010-2017 count), by using equation (iii):

$$TDW_1 = \frac{P_1\,(ADV_1)}{Vp_1} \quad \text{(iii)}$$

Where the TDW (first weight) of a sq./km spatial resolution (raster cell) is a product of the administrative-unit level population count P_1 (for the year in question) and the Ancillary Data Value (ADV) of that cell. This is sub-divided by the output value Vp_1 corresponding to the total built footprint area in the administrative unit representing the population count. This calculation is applied to all raster cells.

A selective sampling approach is herein described and adapted from [3] to normalize GeoCover land cover designations with a local disaggregated value. Samples provide a mean density and a target mean estimation, with a standard deviation of that mean used to classify a range of values to validate accuracy of each target density weight for each normalized class, as shown in Table S1 and associated to Fig. S1(A) and S1(B).

Table S1: Land Cover Selective Sampling for Weight 1 (W_1)

Land Cover (1 sq. / km res.) WGS 84 / Pseudo-Mercator	**Sample Number**	**Sample Mean (SM)**	**Sample SD**	**Estimation Method**	**Target Mean (TM)**	**Target SD**	**Density Target Range**
Terrestrial Vegetation - Tree Cover				Preset	0	0	0 - 1
Bare Areas	3	4,069	5,191	Sample	55	70,6	2 - 82
Cultivated and Managed Lands	4	8,058	3,766	Sample	110	51,3	83 - 253
Rural Settlement	5	29,238	2,974	Sample	398	40,5	254 - 650
Semi-Urban Settlement	4	66,382	6,368	Sample	903	86,7	651 - 1038
Urban Settlement	6	86,285	27,640	Sample	1174	376,2	1,039+

*Note: "Sample Number" is the number of sampled source zones. "Sample Mean" is the mean data density (for built area per sq. / km) of the sampled source zones. "Sample SD" is the standard deviation of the samples collected. "Estimation Method" is the method used to estimate the target mean density. "Target Mean" is the mean data density of target zones associated to Sample Means (the dasymetric map density weight). "Target SD" is the Standard deviation of the Target Mean data density providing a comprehensive "Density Target Range" of values for each classification.

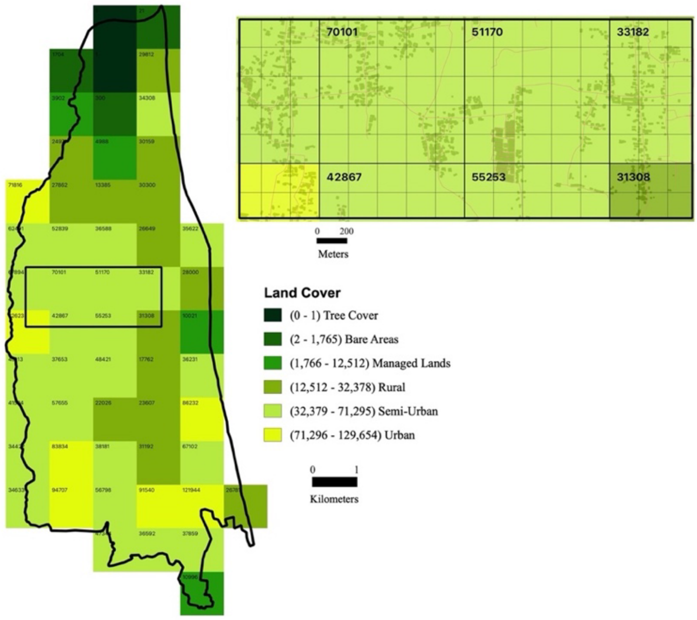


Fig. S1A: Land cover classifications normalized to building footprint areas (in sq./m) after performing a selective sampling approach. Classifications are adapted from [1], except "urban" settlements, applied using a Google Earth Engine designation and planet-OSM ancillary data. The inset box defines the interest area where resettlement is intended and illustrates ancillary data used for TDWs.

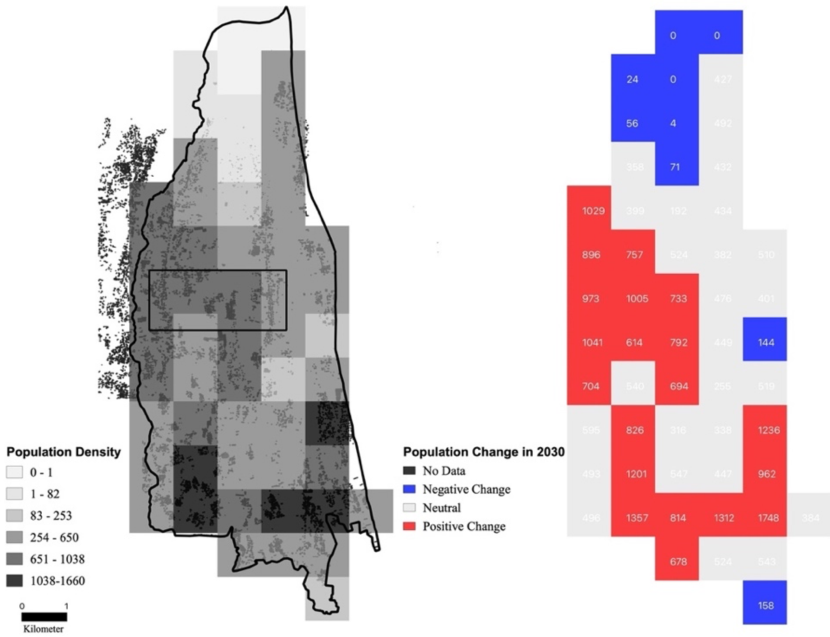


Fig. S1B: Dasymetric mapping of density weights for the Cangkringan sub-district on a square-kilometer spatial resolution with regions and presets of zero density for "tree cover" areas. Note that class intervals are the same as for Fig. 1(A) and color schemes are inversely proportional.

Population Change in 2030 illustrates the number of people with a "negative change" equal to less than one person per year growth rate while "Neutral" refers to a 1-3 person per year growth rate. Inscribed data is the population count for each cell.

Target Area Density Weights (2019) and (2030)

-Area: 4.17 sq. km
-Population Growth: 0,5%

To find the population count for the inset box in Fig. S1(A) and S1(B), density weights obtained in section 5.2.2 reveal a population density count for each of the eight raster cells that make up the target area. To estimate a density weight that is accurately bounded by our target area, a fine-level (200 sq./m) resolution is obtained from sub-dividing each "parent" cell into 25 raster cells (Fig. S1A). A Target Density Weight (TDW_2) is determined with areal interpolation of land cover population density weights (TDW_1) and fine-resolution building footprint areas to estimate land cover class entities normalized to a higher spatial resolution using equation (iv):

$$TDW_2 = \frac{TDW_1\,(Tw_2)}{VW_1} \qquad \text{(iv)}$$

Where, TDW_2 is a product of TDW_1 and the value for ancillary building footprint data in a higher spatial resolution Tw_2, is divided by the total building footprint value VW_1 of its "parent cell".

Once target densities are obtained for TDW_2, the classification is normalized to determine land cover classifications adequate to this resolution and corresponding to the target area in the district. The Target Mean (TM) from each "parent cell's" class in Table 1 becomes the Sample Mean (SM) for a land cover class in Table S2. This states that each class in Table S2 is normalized with TDW_2, computed using a TDW_1, corresponding to the SM and the SD in Table S1. The new set of land cover density weights are used to estimate a Density Target Range of values for a TM value and the subsequent value of that column. The process is illustrated in Fig. S2.

Table S2: Land Cover Selective Sampling for Target Density Weight 2 (W_2)

Land Cover (200-m res.) WGS 84 / Pseudo-Mercator	**Sample Number**	**Sample Mean (SM)**	**Sample SD**	**Estimation Method**	**Target Mean (TM)**	**Target SD**	**Density Target Range**
Terrestrial Vegetation - Tree Cover				Preset	0	0	0 - 1
Bare Areas	4	55	70,6	Sample	7	1,1	2 - 16
Cultivated and Managed Lands	4	110	51,3	Sample	26	6,3	17 - 34
Rural Settlement	4	398	40,5	Sample	42	6,9	35 - 54
Semi-Urban Settlement	4	903	86,7	Sample	67	13,8	55 - 127
Urban Settlement	5	1174	376,2	Sample	188	95,8	128+

*Note: Classes are sampled using density weights (from Table 1) and built area per 0.2 sq./km as ancillary data input for refined spatial resolution density weights. "Sample Number" is the number of sampled source zones. "Sample Mean" is the mean population density (for a single type of land cover) of the sampled source zones. "Sample SD" is the standard deviation of the samples collected in (Table 1). "Estimation Method" is the method used to estimate the target mean density weight. "Target Mean" is the mean data density of target zones in a finer-spatial resolution associated to the Sample Mean (the dasymetric map density weight). "Target SD" is the Standard deviation of the Target Mean (TM) data density providing a comprehensive "Target Range" of data values for each class.

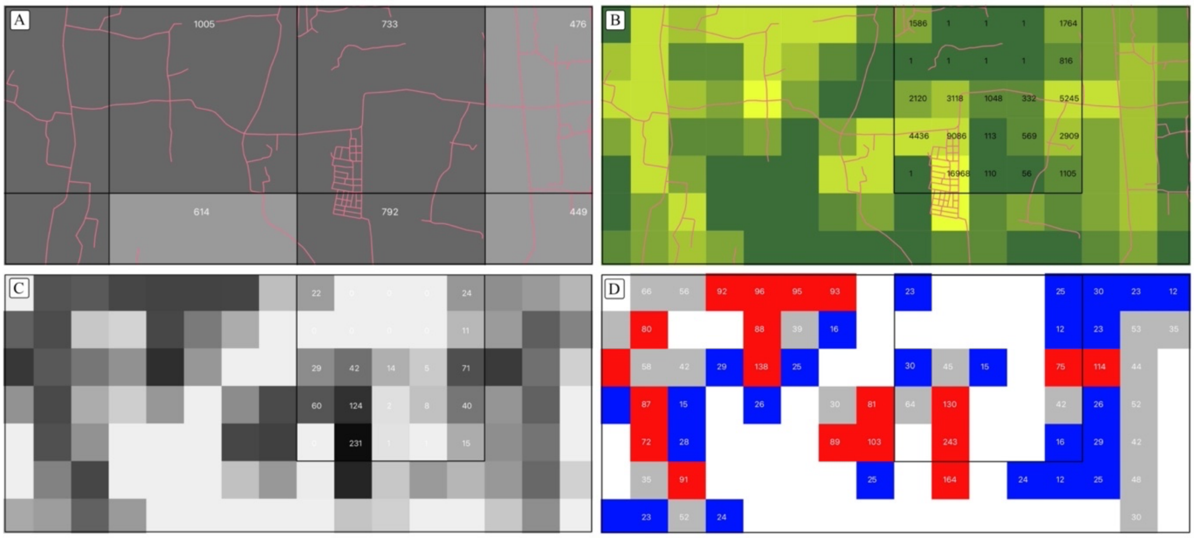


Fig. S2: Maps of the inset box shown in Fig.1 (A) and (B). Population density weight (1) by raster tract (A); normalized land cover with ancillary building data (B); higher-resolution population density weight (2) by dasymetric redistribution (C); population change between 2019 and 2030 (D). Class intervals and color codes are as shown in previous figures.

Target Area Population Count (2019): 3,409 persons (sum of all TDW_2)
Target Area Population Count (2030): 3,590 (with a 0,5% growth rate)
Target Area Population Growth Forecast (2019-2030): 181 persons
Target Area Density (2019): 817.5 persons / km^2
Target Area Density (2030): 860.9 persons / km^2

Target Area Population Density Range

A population range can be estimated to have a more accurate value of the population count for urban, semi-urban, or rural settlement land cover areas that compile two or more raster cells, visualized in Fig. S3. The values for TDW_2 represent the median value of a population density range using the following equation (v);

$$D_{min} = TDW_2 - \left[\frac{(TDM)(TDW_2)}{SMW_2}\right] \qquad \text{(v)}$$

Where, the minimum value for a density range D_{min} is the product between the mean of all TDW_2 in a "parent cell" (TDM) and the TDW_2 in question, sub-divided by the Sample Mean (SMW_2) of its corresponding land cover type using Table S2. The resulting value is subtracted from the target density weight TDW_2 providing the minimum value (or added to the TDW to obtain the maximum value of that range).

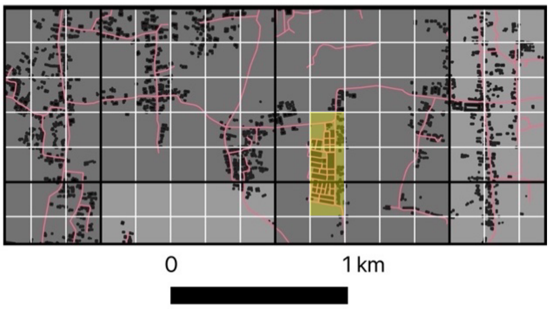


Fig. S3: Map of inset box in Fig. S1(B) corresponding to the census tract dasymetric mapping. Note that Pagerjurang settlement is located in the extent of 3 high-resolution (200-meter) raster cells within two semi-urban "parent" class. Color intervals are as shown in Fig. S1(B).

Accuracy Assessment

The following calculation is performed to validate the aforementioned dasymetric distribution sequence with an official high-spatial resolution census count in [Dataset 4] and [Dataset 5] for Pagerjurang in Table S3.

1. "Parent" cell id: 731
$TDW_2 = 124$ (id. 804)
TDM= 28 (the mean value for all density weights inside a "parent" raster cell)
$SMW_2 = 920$ (Sample Mean density for Semi-Urban settlements in Table 2)

$$D_{min} = 124 - \left[\frac{(124)x(28)}{920}\right]$$
= 120
$$D_{max} = 124 + \left[\frac{(124)x(28)}{920}\right]$$
= 128

2. "Parent" cell id: 731
$TDW_2 = 231$ (id. 805)
$D_{min} = 224$
$D_{max} = 238$

3. “Parent” cell id: 736
$TDW_2 = 156$ (id. 806)
TDM= 21
$SMW_2 = 920$

$D_{min} = 152$
$D_{max} = 160$

4. Pagerjurang Population Estimation (±17 persons)

$\sum D_{min} = 496; \quad \sum TDW_2 = 511 \,; \sum D_{max} = 526$

Population Count for 2019 in Table S3 is: 528.3 persons

Table S3: Pagerjurang Population Count for 2014 and 2017 with a growth rate of 0,3%

2014	2015	2016	2017	2018	2019
495	501.66	508.32	515	521.64	528.3
Note: The value for the year 2019 is calculated by applying the 0,3% growth rate from the 2014 and 2017 official census counts sourced from: Monografi Desa Kepuharjo 2017 [Dataset 4], and Data Sekunder Laporan Profil Desa Kepuharjo 2015 [Dataset 5].					

Therefore, the upper value of the threshold (526) is 2 persons away from the actual value (528) in 2019, illustrating robustness of the model with an accuracy of 96.78% ± 3.2%.

Population by Age Group (Adults: 18 years or older)

Target Area Population Growth Forecast (2019-2030): 181 persons

Population according to Age Group (2019) [Dataset 6]:
0 – 19 years: 29%
20+ years: 71%
Adults = 128.5

5.1.3 Node Weight Estimation

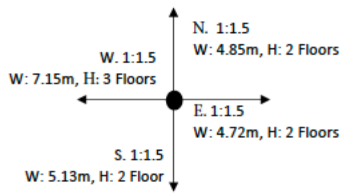
Input Data 2: Household Preferences from housing-preference-driven demand in typical urban structures.
Social Neighborhoods: 71%
Surroundings: 12%
Centrality Near Job: 10%
Education Facility: 2%
Crime Rate: 0%
Green Areas Supply: 0%
NA: 5%

(71%+12%+10%) = 93%
0.93 x 128.5 = 119.5

Total New Adults: 119.5

Node 1
Category: Local Node
Influence Radius: 300 meters

Road Type:
A. Vertical axis: Local Road (width (m): 4.85 north, 5.13 south) 2 Lanes / Ratio 1.5:1
B. Horizontal axis (left side): Collector Road (width (m): 7.15 west) 2 Lanes / Ratio 1.5:1
C. Horizontal axis (right side): Local Road (width (m): 4.72 east) 1 Lane / Ratio 1.5:1



Node 2
Category: Local
Influence Radius: 300 meters

Road Type:
A. Vertical axis: Collector Road (width (m): 4.62 north, 5.10 south) 2 Lanes / Ratio 1:1.5
B. Horizontal axis: Local Road (width (m): 5.60 west, 6.00 east) 1 Lane / Ratio 1:1.5

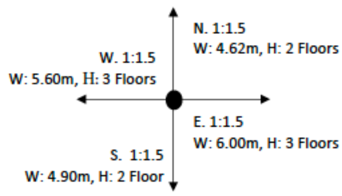


Fig. S4: Road width and enclosure ratios for comparison between typologies of equal categories. Note that each road segment is treated independently to achieves the highest building density output.

Weights
Node 1: 21.85 meters (51%)
Node 2: 21.12 meters (49%)
Total sum: 42.97 (100%)

Total Adults: 119.5
119.5 x 0.51 = 61 adults in Node 1
119.5 x 0.49 = 58.5 adults in Node 2

5.1.4 Housing Demand

Input Data 3: Household composition of Adults per Household (APH):
1-Adult HHs: 52%
2-Adult HHs: 32%
3-Adult (or more) HHs: 16%

Node 1
61 x 0.52 = 31.72 → 32 HHs
61 x 0.32 = 19.52 → 10 HHs
61 x 0.16 = 9.76 → 3 HHs
Total: 45 HHs

Node 1 – 51% (0.51)
0.51 x 88 = 45 HHs

Node 1
32 (1 adult) x 70m^2 = 2,240 m^2
10 (2 adult) x 100m^2 = 1000 m^2
3 (+3 adult) x 120m^2 = 360 m^2
Total Area: 3,600m^2

Node 2
58.5 x 0.52 = 30.42 → 31 HHs
58.5 x 0.32 = 18.72 → 9 HHs
58.5 x 0.16 = 9.36 → 3 HHs
Total: 43 HHs

Node 2 – 49% (0.49)
0.49 x 88 = 43 HH

Node 2
31 (1 adult) x 70m^2 = 2,170 m^2
9 (2 adult) x 100m^2 = 900 m^2
3 (+3 adult) x 120m^2 = 360 m^2
Total: 3,430 m^2

Housing Demand Area: 7,030 m^2

The following Table S4 corresponds to the data collected using Google Street View for the current (2019) building heights. Road widths together with enclosure ratios were used to increment space for each lot with access to roads pertaining to a node. Finally, the net-built capacity was calculated in surface area (m^2).

Table S4: Built Density Thresholds. Enclosure ratios provide a maximum height parameter for building growth in roads leading to the junction of each node. This table shows the net-built capacity of buildings with access to these roads by 2030.

Road	Hierarchy	Node 1 Direction	Width (m)	Lanes (unit)	Enclosure Ratio (H:W) 2019	Enclosure Ratio (H:W) 2030	Lots with direct street access (m2) 2019	Building Levels (m)	Growth / Level (m2) 2030	Net Growth / Level (m2) 2030	Max. Built Area (m2) 2030
1	Local	North	4.85	2	0.6:1	1.5:1	4,292	L1: +0.00			
							356	L2: +2.85	3885	3885	7,770
							-	L3: +5.7.0	763	1170	2,289
2	Local	South	5.13	2	0.5:1	1.5:1	3,798	L1: +0.00			
							-	L2: +2.85	3,723	3,723	7,446
							-	L3: +5.70	75	150	225
3	Collector	West	7.15	2	0.4:1	1.5:1	3,835	L1: +0.00			
							296	L2: +2.85	681	681	1,362
							-	L3: +5.70	3450	6604	10,054
4	Local	East	4.72	1	0.6:1	1.5:1	2,428	L1: +0.00			
							269	L2: +2.85	783	783	1,028
							-	L3: +5.70	1,914	3,559	6,011
							15,274		15,274	20,555	36,185

Road	Hierarchy	Node 2 Direction	Width (m)	Lanes (unit)	Enclosure Ratio (H:W) 2019	Enclosure Ratio (H:W) 2030	HHs with direct street access (m^2) 2019	Building Levels (m)	Growth / Level (m^2) 2030	Net Growth / Level (m^2) 2030	Max. Built Area (m^2) 2030
1	Collector	North	4.62	2	0.6:1	1.5:1	1,934	L1: +0.00			
							261	L2: +2.85	2,195	1,934	4,390
							-	L3: +5.70			
2	Collector	South	5.1	2	0.6:1	1.5:1	3,351	L1: +0.00			
							90	L2: +2.85	2,877	2,877	5,754
							-	L3: +5.70	564	1,038	1,692
3	Local	West	5.6	1	0.5:1	1.5:1	1,511	L1: +0.00	255		
							452	L2: +2.85			
								L3: +5.70	1,705	2,958	4,985
4	Local	East	6	1	0.5:1	1.5:1	933	L1: +0.00	177		
							264	L2: +2.85			
							-	L3: +5.70	1,023	1,782	3,063
							8,796		8,796	10,589	19,884

Note: the mean inter-story height is calculated as 2.85 m. This value corresponds to the mean first level height of buildings surveyed using the ruler tool and standard finishing materials in Google Street View.

5.2 Socioeconomic Environment

5.2.1 Land Use Designation Module

Data Input 3: User-defined activities compatible with the structuring activities of a node.

Growth Areas by 2030

Category:
Node (1): Local-Institutional Service Node Node (2): Local-Educational Service Node

Proposed Primary Activities:
(1) Technical Research Institute (primary "Institutional" node compatible with "educational" land use) [4].
*Area: 4,060m^2

(2) Cultural Language Institute (primary "Educational" node compatible with "educational" land use) [5].
*Area: 4,027 m^2

(3) Local Medical/Veterinary Center (primary "Educational" node compatible with "service" land use) [6, 7].
*Area: 2,300 m^2

Total: 10,387 m^2

5.2.2 Employment Distribution

Table S5: Data collected for the OLS regression based on sectoral areas of the local employment structure in the Cangkringan district (see Mendeley Repository). Data source is openly available in Bahasa language. (**https://kependudukan.jogjaprov.go.id/statistik/penduduk/statusperkawinan/9/0/17/04/34.clear**).

Year	T1	NT	T	A	Pop	NT*T	NT*A	T*A	(xi-xmean)^2	(yi-beta(xi))	(yi-ymean)^2	yhat	(yhat-ymean)^2	Error	Error^2

2015	18	64	20	-106	7979	1280	-6784	-2120	21.78	49.94	1021.87	35.31	10.76	28.69	822.88
2015	10	93	24	-238	10694	2232	-22134	-5712	75.11	76.13	3716.93	38.13	37.12	54.87	3011.14
2015	0	22	30	-78	4026	660	-1716	-2340	215.11	0.91	100.67	42.34	106.32	-20.34	413.89
2015	1	25	-1	-36	3334	-25	-900	36	266.78	25.70	49.47	20.55	131.85	4.45	19.80
2015	9	36	24	-55	4950	864	-1980	-1320	75.11	19.13	15.73	38.13	37.12	-2.13	4.52
2016	21	17	8	-62	7864	136	-1054	-496	53.78	11.38	226.00	26.88	26.58	-9.88	97.57
2016	14	20	12	-81	10726	240	-1620	-972	11.11	11.56	144.80	29.69	5.49	-9.69	93.89
2016	0	38	1	-90	4018	38	-3420	-90	205.44	37.30	35.60	21.96	101.54	16.04	257.39
2016	4	1	5	-9	3379	5	-9	-45	106.78	-2.52	963.07	24.77	52.77	-23.77	564.96
2016	9	34	16	-23	5041	544	-782	-368	0.44	22.75	3.87	32.50	0.22	1.50	2.24
2017	18	20	27	-123	7777	540	-2460	-3321	136.11	1.02	144.80	40.24	67.27	-20.24	409.47
2017	15	10	23	-90	10607	230	-900	-2070	58.78	-6.17	485.47	37.42	29.05	-27.42	752.03
2017	0	3	2	-36	3948	6	-108	-72	177.78	1.59	842.93	22.66	87.86	-19.66	386.50
2017	4	21	-1	-23	3383	-21	-483	23	266.78	21.70	121.73	20.55	131.85	0.45	0.20
2017	9	-5	17	-20	5058	-85	100	-340	2.78	-16.95	1371.47	33.21	1.37	-38.21	1459.62
2018	23	44	36	-50	7811	1584	-2200	-1800	427.11	18.69	143.20	46.56	211.09	-2.56	6.57
2018	17	98	34	-82	10713	3332	-8036	-2788	348.44	74.10	4351.60	45.16	172.21	52.84	2792.45
2018	1	22	13	-11	4000	286	-242	-143	5.44	12.86	100.67	30.39	2.69	-8.39	70.44
2018	3	11	18	-18	3396	198	-198	-324	7.11	-1.65	442.40	33.91	3.51	-22.91	524.78
2018	10	52	30	-12	5186	1560	-624	-360	215.11	30.91	398.67	42.34	106.32	9.66	93.23
2019	25	59	20	-52	7829	1180	-3068	-1040	21.78	44.94	727.20	35.31	10.76	23.69	561.02
2019	17	33	20	-69	10755	660	-2277	-1380	21.78	18.94	0.93	35.31	10.76	-2.31	5.36
2019	1	50	-3	-19	4045	-150	-950	57	336.11	52.11	322.80	19.14	166.12	30.86	952.05
2019	3	17	9	-18	3453	153	-306	-162	40.11	10.67	226.00	27.58	19.82	-10.58	111.95
2019	11	10	23	-35	5227	230	-350	-805	58.78	-6.17	485.47	37.42	29.05	-27.42	752.03
2020	27	18	21	-41	7898	378	-738	-861	32.11	3.24	196.93	36.02	15.87	-18.02	324.62
2020	21	97	9	-108	10842	873	-10476	-972	40.11	90.67	4220.67	27.58	19.82	69.42	4819.02
2020	2	26	3	-52	4080	78	-1352	-156	152.11	23.89	36.40	23.36	75.18	2.64	6.96
2020	3	8	16	34	3489	128	272	544	0.44	-3.25	577.60	32.50	0.22	-24.50	600.35
2020	11	17	4	-33	5332	68	-561	-132	128.44	14.19	226.00	24.07	63.48	-7.07	49.92

Source: Semester II Employment Data, Secretariat of Governance Bureau DIY [Dataset 7].

Table S6: Jobs Placed and Removed

Sector	Sub-district	Shift-Share Instrument District share	Industrial Mix	Sub-District share	Total Change	**2015-2020**	Placed	Removed	LM	**2020 - 2025**	Removed	LM	**2025 - 2030**	Removed	Total Placed	Total Removed
Tradable Skilled Labor *(15m2/person) **Education (55 m2 / 30 persons)	A	-1.3	114.0	-0.6	112											
	W	-1.1	90.9	7.1	97											
	G	-0.5	42.9	-30.4	12											
	K	-0.3	27.0	10.3	37	18								-88		
	U	-0.9	77.3	13.6	90		44		1.51	109		1.51	164			
	Total	-4.1	352.1	0.0	348	17	120		1.96	268		1.96	437	-88	1,116	-88
Non Tradable Employees *(10m2/person)	A	-3.3	155.1	-16.8	135											
	W	-4.8	229.2	-22.4	202											
	G	-0.9	40.8	68.1	108											
	K	-1.3	59.9	-0.6	58	28			2.95	214		2.95	632			
	U	-2.1	99.3	-28.2	69											
	Total	-12.4	584.4	0.0	572	28				214			632		874	0
Agriculture	A	-5.5	-224.0	-98.5	-328											
	W	-8.0	-322.0	-100.0	-430											
	G	-5.6	-227.6	25.2	-208											
	K	-3.7	-151.5	121.3	-34	-16			2.60		-42	2.60		-110		
	U	-4.2	-170.8	52.0	-123				0.58		-9	0.58		-5		
	Total	-27.1	-1095.9	0.0	-1123	-16					-52			-115	0	-183
Tradable State Enterprise *(10m2/person)	A	-0.1	6.5	-0.4	6											
	W	0.0	4.3	2.7	7											
	G	0.0	0.3	0.7	1											
	K	0.0	1.2	-2.2	-1	-0.5			2.95		-1.5	2.95		-4.4		
	U	0.0	2.8	-0.8	2											
	Total	-0.2	15.2	0.0	0						-1.5			-4.4	0	-6
Non Tradable Entrepreneur *(12 m2/person)	A	-2.0	32.2	-48.2	-18											
	W	-2.4	38.2	25.2	61									-36		
	G	-0.8	13.2	12.6	25											
	K	-0.9	14.8	-12.9	1											
	U	-1.0	15.8	23.2	38							2.95				
	Total	-7.2	114.2	0.0	107	21		0	2.95	62	0		183	-36	266	-36
Non Tradable Medical ***(9 nurses / 1000 persons) ***(3 nurses / doctor)	A	-0.1	14.4	0.7	15											
	W	-0.1	10.5	1.6	12											
	G	0.0	1.3	2.7	4	*These were placed under skilled										
	K	0.0	0.0		0											
	U	0.0	3.9	-4.9	-1											
	Total	-0.2	30.2	0.0	30											
	Total	**-51**	**-30.2**	**0**	**-51**	**67**	**120**	**-17**	**-**	**653**	**-53**	**-**	**1417**	**-243**	**2257**	**-313**

Note: *The Employment Density Guide 3rd Edition, 2015, **Area Guidelines for Mainstream Schools. 2014, ***International Health Facility Guidelines Part C: Version 4, 2015. Administrative entities correspond to the Cangkringan district's sub-divisions Argomulyo Village (A); Wukirsari (W); Glagaharjo (G); Kepuharjo (K); and, Umbulharjo (U). Local Multipliers (LM) are applied to two observances (2020-2025) and (2025-2030).

Density Guide 3rd Edition, 2015

Commerce: 15 m^2/job

Small Service: 10m^2/job
Industry: 47 m^2/job
Manufacturing: 36 m^2/job
Small retail (HBE): 12 m^2/job

General teaching 55 m^2/job (30-person classroom)
Medical center 9 nurses/1000 persons
Medical center 3 nurses/1 doctor

5.2.3 Employment Allocation

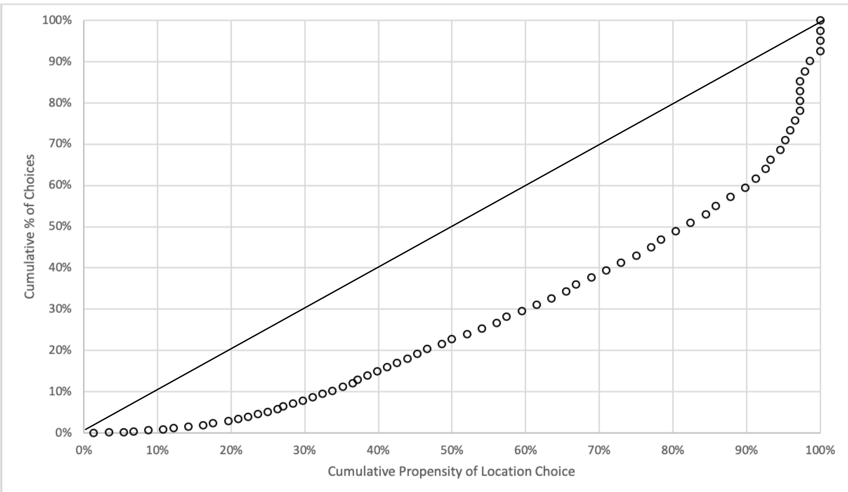
Regression Statistics	
Multiple R	0.56425996
R Square	0.318389302
Adjusted R Square	0.277696126
Standard Error	0.766674846
Observations	72

F-Test of Overall Significance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	15.79022172	3.947555429	7.781854432	3.266E-05
Residual	67	33.98755606	0.507276956		
Total	71	49.77777778			

Residual square-sums are too high and the f-test result indicates a low value with high significance. However, in Fig. S5, the efficiency dictates prediction of job location choices with 91% efficiency.

Fig S5: Lorenz-curve Predictive Efficiency.



*Note: 80% of location choices were made with 50% of the total set of choices.

Likelihood Ratio Test	
Vacancy	-11.84
Built Density	-12.30
Building Capacity	-11.84
Land Value coef.	-11.85

OBS	72
Probability Log-Likelihood	0.79
df	4

Note: An overall 79% probability of prediction. Predictors are all above the differential chi-square value 4. Therefore, all are significant predictors for job location choice.

Pool of Vacancy Areas by Land Use

Node 1	Built Capacity (m^2)	Jobs
Structuring Activity	4027	60
Residential Capacity	5687	
HBE Capacity	5224	435
Services	3912	391
Commerce	5524	368
Light Industry	208	6

Node 2	
Structuring Activity	6327
Residential	7817
HBE	2379
Vacant Lots	4909
Services	309
Commerce	2720
Light Industry	370

Aggregate Built Capacity
Total built capacity node 1: 24,582 m^2
Total built capacity node 2 (Structuring Activity excluded): 24,831 m^2
Total: 49,414 m^2

Aggregate Employment Demand / Jobs
Tradable: 1,022 jobs
Non-Tradable: 692 jobs
Entrepreneurs: 230 jobs
Total: 1,944 jobs

Tradable demand:1,022 – (164 Services +549 commerce +422 services) = 113 jobs vacancy available.
Entrepreneur demand: 230 – (633 HBEs) = 403 job vacancies available*
*Note: become rental space for non-tradable jobs below.
Non-tradable demand: 692 – (113+403+16 Light Industry) = 160 jobs left without space for allocation.
91% allocated

Non-Allocated (non-tradable): 160 jobs
*These were assumed to take hold of vacant lots in node 2 with residential use, thus becoming HBE vacancies.

Vacant Residential Lots node 2: 4,909 m^2
Vacant Residential Area used for 160 HBE jobs: 1,920 m^2
160 x 12 m^2 = 1,920 m^2
Vacant Area left: 2,989 m^2

Land use change (residential → mixed-use):
7,603 m2 (HBE Capacity) + 1,920 m^2 (Vacant Lots used for HBEs) = 9,523 m^2

5.2.4 Spillover Management

Regression to predict the probability of a site being selected. The R^2 value indicates that 50-53% of the choices can be predicted with the model. However, the f-test with a significance higher than 99% provides an overall fitted measure of robustness.

Regression Statistics	
Multiple R	0.728863266
R Square	0.53124166

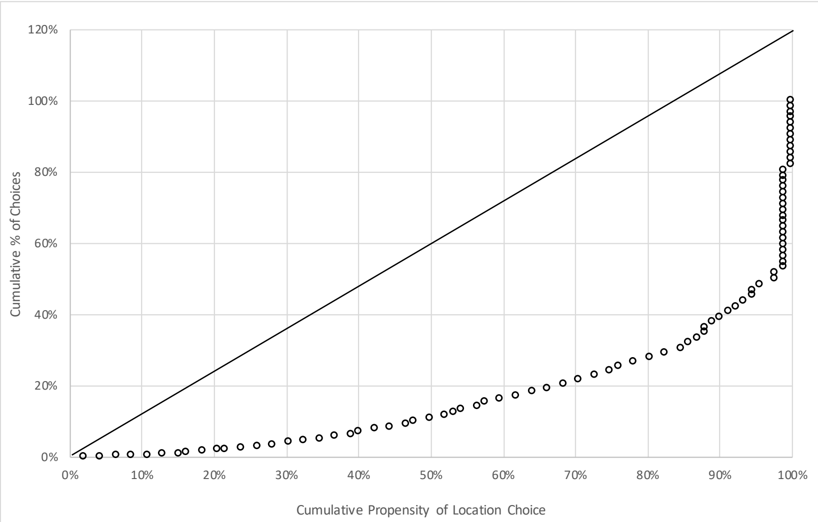
Adjusted R Square	0.5083754
Standard Error	0.902066226

F-Test of Overall Significance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	75.61950252	18.9048756	23.2325553	7.35327E-13
Residual	82	66.72532507	0.81372348		
Total	86	142.3448276			

Note: High f-value with very significant rating (99%). However, residual values were high.

Fig. S6: Lorenz-curve distribution for predictive efficiency.



Note: 90% of location choices were predicted with 40% of total choices. Predictive efficiency = 81%.

Likelihood Ratio Test

Intercept	-24.85
Vacancy	-20.37
Built Density	-22.27
Building Capacity	-24.17
Land Value coef.	-23.10
OBS	87
Probability log-Likelihood	0.77
df	4

Note: resulting values are very high with an overall prediction probability of 77%. Positive or negative signs are not considered.

Floating Population - Residential Demand
Node 1
In-Migration for the 1,260 jobs allocated
1,260 x 0.16 = 202 jobs → 202 HHs of 3-adults = 606 adults
1,260 x 0.32 = 403 jobs → 403 HHs of 2-adults = 806 adults
1,260 x 0.52 = 655 jobs → 655 HHs of 1-adult = 655 adults

Estimation of Adults that do not work (16% of Kepuharjo sub-district [Dataset 7])
3-adult HHs + 2-adult HHs
(933 adults + 1244 adults) x 0.16 = 348 adults in total do not work in the TAZ.

16% → X
(16+32=48) → 100%
X=33%, Y=77%
348 x 0.33 = 115 (3-adult HHs)
348 x 0.77 = 233 (2-adult HHs)

3-Adult HHs Node 1
115 adults x 0.51 (Node Weight Estimations in sub-section 5.4.1) =58.65 adults
59 adults that do not work → 59 HHs x 2 adults = 118 jobs + (59 adults not working)
Therefore, 59 HHs have 2 working adults and 59 adults that do not work.
202 jobs – 118 jobs = 84 jobs / 3 adults = 28 HHs with 3-adults working.

Total: 28 HHs (3 working adults) + 59 HHs (2 working adults) = 87 HHs, 202 jobs, 261 adults.

2-Adult HHs Node 1
233 x 0.51 (Node Weight Estimations in sub-section 5.4.1) =118.83 adults
119 adults that do not work → 119 HHs x 1 adult = 119 jobs + (119 adults not working)
Therefore, 119 HHs have 1 working adult and 119 adults do not work.
403 jobs – 119 jobs = 284 jobs / 2 adults = 142 HHs with 2-adults working.

Total: 119 HHs (1 working adult) + 142 HHs (2 working adults) = 261 HHs, 403 jobs, 522 adults.

1-Adult HHs in Node 1
1,260 x 0.52 = 655 jobs → 655 HHs of 1-adult = 655 adults (all work)

Total Residential Demand for Node 1 (1,012 HHs, 1,456 adults)
87 (3-adult HHs) x 120 m2 = 10,440 m^2
261 (2-adult HHs) x 100 m2 = 26,100 m^2
655 (1-adult HH) x 70 m2 = 45,850 m^2
Sub-Total: 82,390 m^2
Additional Housing Demand (sub-section 5.1.4): 3,600 m^2
Total Demand: 85,990 m^2
Residential vacancy: 10,911 m^2
Deficit: 75,079 m^2

Node 2
In-Migration for 684 jobs allocated (+160 HBE jobs)
684 x 0.16 = 109 jobs → 109 HHs of 3-adults = 327 adults
684 x 0.32 = 219 jobs → 219 HHs of 2-adults = 438 adults
684 x 0.52 = 356 jobs → 356 HHs of 1-adult = 356 adults

3-Adult HHs Node 2
115 x 0.49 (Node Weight Estimations in sub-section 5.4.1) =56.35 HHs
56 adults that do not work → 56 HHs x 2 adults = 112 jobs + (56 adults not working)
Therefore, 54 HHs have 2 working adults and 54 adults that do not work.
109 jobs – 108 jobs = 1 jobs / 3 adults = 0 HHs with 3-adults working.

Total: 1 HHs (3 working adults) + 53 HHs (2 working adults) = 54 HHs, 109 jobs, 162 adults.

2-Adult HHs Node 2
233 x 0.49 (Node Weight Estimations in sub-section 5.4.1) =114.17 HHs
114 adults that do not work → 114 HHs x 1 adult = 114 jobs + (114 adults not working)
Therefore, 114 HHs have 1 working adult and 114 adults do not work.
219 jobs – 114 jobs = 105 jobs / 2 adults = 52.5 HHs with 2-adults working.

Total: 115 HHs (1 working adult) + 52 HHs (2 working adults) = 167 HHs, 219 jobs, 334 adults.

1-Adult HHs in Node 2

Total: 684 x 0.52 = 356 jobs → 356 HHs of 1-adult = 356 adults (all work)

Total Residential Demand for Node 2 (573 HHs, 844 adults)
54 (3-adult HHs) x 120 m2 = 6,480 m^2
167 (2-adult HHs) x 100 m2 = 16,700 m^2
356 (1-adult HH) x 70 m2 = 24,920 m^2
Sub-Total: 48,100 m^2
Additional Housing Demand (sub-section 5.1.4): 3,430 m2
Total Demand: 51,530 m^2
Residential vacancy: 15,105 m^2
Deficit: 36,425 m^2

Residential Demand in TAZ (aggregate values)
87 + 54 = 141 (3-adult HHs) x 120 m2 = 16,920 m^2
261 + 167 = 428 (2-adult HHs) x 100 m2 = 42,800 m^2
655 + 356 = 1,011 (1-adult HH) x 70 m2 = 70,770 m^2
Sub-Total: 130,490 m^2
Additional Housing Demand (sub-section 5.1.4): 7,030 m2
Total Demand: 137,520 m^2
Residential vacancy: 26,016 m^2
Deficit: 111,504 m^2

Household Allocations in Nodes
From the estimations above: 301 HHs

The Horizontal Boundary road between node 1 and node 2 is applied with a (H:W) enclosure ratio applicable to its width.

Service: 943 m^2
Commerce: 2,544 m^2
Industry: 207 m^2
Residential: 2,088 m^2
Mixed-Use Residential: 3,570 m^2
Total residential: 5,658 m^2
Total non-residential: 3,694 m^2
Total: 9,352 m^2

The vertical boundary road is applied with a (H:W) enclosure ratio and delivers the following growth areas.

Residential: 1,232 m^2
Mixed-Use Residential: 1,388 m^2
Total residential: 2,620 m^2

Non-Allocated in nodal Areas (HH Type, Income USD: 100>HH income>100, HH APH ratios) = Max. of 1,284 HHs using the Household Allocation Model (in Mendeley Data).
*Note: Max. = Equilibrium. 100 USD is an approximation to the mean income in our data sample (78 USD).

(3-Adult, >100 USD, 0.16) = 141 HHs (549 p), >100 → 100% allocated
(3-Adult, < 100USD, 0.16) = 0% allocated
(2-Adult, >100 USD, 0.32) = 377 HHs (644 p) > 100 → 100% allocated
(2-Adult, < 100USD, 0.32) = 51 HHs (88 p) <100 → 90 HHs → 100% allocated (+39 vacancies left)
(1-Adult, >100 USD, 0.52) = 607 HHs, (356 p) >100 → 100% allocated
(1-Adult, < 100USD, 0.52) = 404 HHs (238 p) <100 → 146 HHs → (+/-45 vacancies) →219 HHs (80%)

Non-Allocated in Nodes = 1,284 HHs
Non-Allocated at all = 254 HHs / 86% of 1-adult HHs and 14% of 3-adult HHs.

5.3 Settlement Environment

5.3.1 Household Clustering

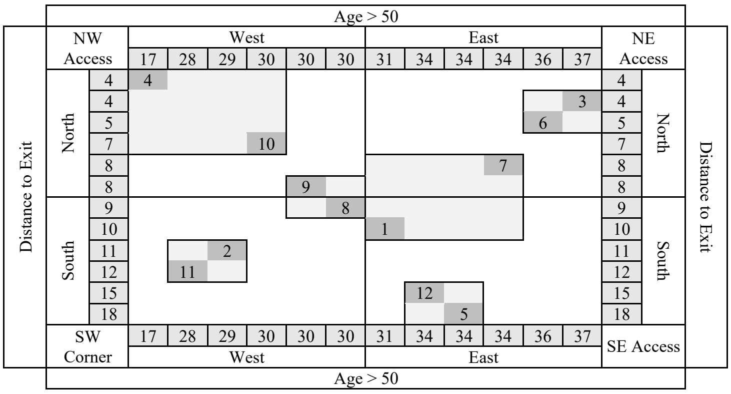
Table S5: Geo-referenced Attribute Values by Cluster

X-Axis

Cluster	4	11	2	10	9	1	12	5	7	6	3
Land Use	Educati...	Comme...	Comme...	Industry	Services	Comme...	Comme...	Comme...	Industry	Industry	Services
Age>50	17	28	29	30	30	31	34	34	34	36	37

Y-Axis

Cluster	Land-Use	Distance to Exit
4	Education	4
3	Services	4
6	Small Industry	5
10	Small Industry	7
7	Small Industry	8
9	Services	8
8	Commerce	9
1	Commerce	10
2	Commerce	11
11	Commerce	12
12	Commerce	15
5	Commerce	18



5.3.2 Street Network

Block Types

Three block types are analyzed to satisfy 26 HHs (1 cluster) or 52 HHs (2 clusters). Lots are equal to 120 m^2 in order to satisfy household sizes of more than 2 adults. However, 30% is established to be the minimum footprint area for gardens and 250 m^2 recreational areas are integrated to the three standard block types in Table S6.

Table S6: Standard Block Types

Neighborhood Cell Estimation				
Unit		1 Small Cell	1 Big Cell	Combined Cell
Unit Area		80 x 80	100 x 100	80 x 100
Land Use	-	1 cluster	2 cluster	2 clusters
Persons		50 persons	100 persons	100 persons
Households		26 HH	52 HH	52 HH
Plots	1 Plot	26 Plots	52 Plots	52 Plots
Plot Area	120	3,120	6,240	6,240
Unbuilt Area	35	910	1,820	1,820
Built Area	85	2,460	4,420	4,420
Total Unbuilt Area		3,940	5,330	3,580
Public Space		3,030	3,760	1,760
Density m2 / person		78 / Ha	100 / Ha	125 / Ha
Recreational Areas	2.5 / person	250	250	200
Agro-Community	35 / HH	910	1820	1820
Roads (10%)		454	564	234

*Note: values are in (m; m^2; m^3) unless indicated differently. Interior lot open space was designated to small agricultural plots, which in combination, are very useful to the community. Roads take 10% of each block area.

Selective Sampling of the Accessibility Index

The sampling method is performed to quantify the accessibility index for settlement simulation using Table S7. The following was evaluated: (i) Intersection density in a 0.25 (km^2) target area; (ii) road hierarchies and connectivity between opposite sides of a network; (iii) number of access points to the network; and, a permeability designation for the target area based on the results as either (1) High, (2) Limited, (3) Low, and (4) none.

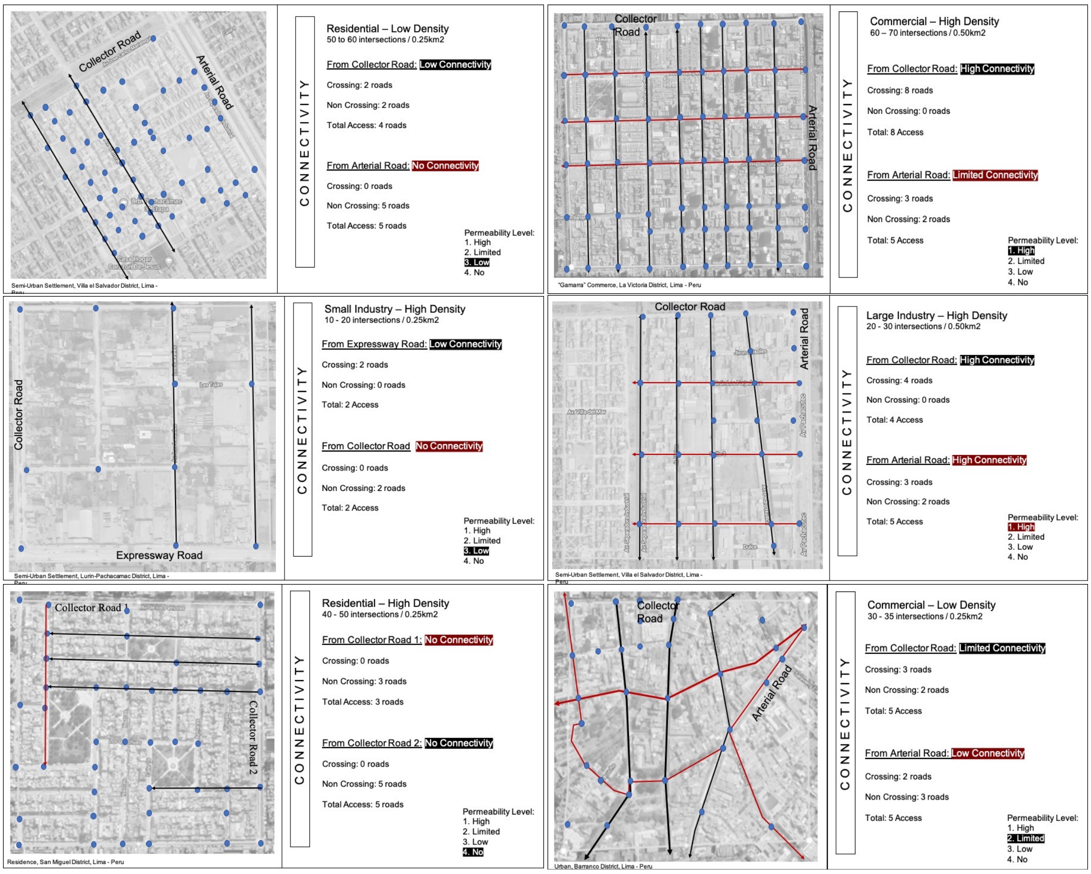


Fig. S9: Selective Sampling method for six different settlement areas in Peru, evaluated using Google Earth to quantify accessibility measures and density indices in [12] for: high- and low-density residential and commercial areas, office areas, and small- and large-industry densities, classified with two urban designations (i.e., semi-urban and urban). The number of intersections, block sizes, block proportions, and connectivity are variables sampled and they convey standard topological attributes of street networks for this study. Further sampling might be required to replicate this data for other locations around the world.

Table S7: Results from selective sampling of spatial characteristics in Lima city

Categories		Connectivity		Intersections		Permeability	
Zone	Density	Boundary Roads	Block Size (m)	Semi Urban	Urban	Road Crossing	Level
Residential	High	Local / Collector / Sub-Arterial	80 / 150 / 200	30 – 40	40 – 50	0	Zero
	Low	Local / Collector	80 / 100	40 – 50	50 – 60	2	Low
Commercial	High	Collector / Sub-Arterial	100 / 150	20 – 30	30 – 40	3	High
	Low	Local / Collector	80 / 100	30 – 35	30 – 40	2	Limited
Office	High	Siub-Arterial / Arterial	200 / 150	10 – 25	20 – 35	2	High
Small Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	10 – 20.	15 – 25	2	Low
Large Industry	High	Local / Collector / Sub-Arterial	80 / 100 / 300	15 – 20	20 – 30	3	High

5.3.3 Land Use Change

Here, we report the full statistic results for the first and second iteration of the predictive variables with our data sample. Then, the land use multinomial logit model is provided with full regression scores for the maximized log-likelihood of conversion choices. Our trained data and optimized logit scores are available online in our Mendeley data repository (Find the link at the top of this document).

First Iteration (Only those with primary occupations other than farming, services, or education.)

Regression Statistics	
Multiple R	0.96
R Square	0.93
Adjusted R Square	0.92
Standard Error	0.16
Observations	41

F-Test of Overall Significance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	12.21687472	2.036145786	76.48961888	2.70502E-18
Residual	34	0.905076502	0.026619897		
Total	40	13.12195122			

Regression Scores

Variables	*Coefficients*	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.569	0.24	-2.407	0.022
HH comp	0.029	0.04	0.675	0.505
Occupation	0.789	0.05	17.476	0.000
Risk	0.031	0.04	0.840	0.407
income	0.001	0.00	0.994	0.327
age	-0.001	0.00	-0.618	0.541
education	-0.013	0.02	-0.798	0.430

**Note: P-value significance is low due to the number of observations.*

Second Iteration (Only those with commercial and farming as their primary occupation. Assigned values to Commercial = 2 and Farming = 1).

Regression Statistics	
Multiple R	0.79
R Square	0.63
Adjusted R Square	0.56
Standard Error	0.20
Observations	41

F-Test of Overall Significance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	2.269614294	0.378269049	9.596855821	3.48989E-06
Residual	34	1.340141803	0.039415935		
Total	40	3.609756098			

Regression Scores

Variables	*Coefficients*	*Standard Error*	*t Stat*	*P-value*
Intercept	0.844	0.267	3.161	0.003

HH comp	-0.101	0.052	-1.941	0.061
Occupation	0.290	0.066	4.38	0.000
Risk	-0.034	0.046	-0.745	0.461
income	-0.002	0.001	-1.943	0.060
age	0.003	0.003	1.174	0.249
education	-0.017	0.020	-0.823	0.416

*Note: Higher P-values correspond to a trained dataset using the abovementioned values.

Table S8. Analysis of Predictive Efficiency

Variables	*First Iteration* **1**	*Min* **2**	*Max* **3**	*Second Iteration* **4**	*Min* **5**	*Max* **6**
Intercept	-0.569** (0.24)	-1.21	0.08	0.844***(0.267)	0.12	1.57
HH composition	0.029 (0.04)	-0.09	0.15	-0.101* (0.05)	-0.24	0.04
Occupation	0.789*** (0.05)	0.67	0.91	0.29*** (0.06)	0.11	0.47
Segmentations	0.031 (0.04)	-0.07	0.13	-0.034 (0.04)	-0.16	0.09
Income	0.001 (0.00)	0.00	0.00	-0.002* (0.001)	0.00	0.00
Age	-0.001 (0.00)	-0.01	0.01	0.003 (0.003)	0.00	0.01
Education	-0.013 (0.02)	-0.06	0.03	-0.017 (0.02)	-0.07	0.04

***significant at 99%, **significant at 95%, *significant at 90%. Note that education with negative coefficient values indicate a negative correlation, while income and HH composition represent accurate predictors of land use change. F-significance values for the first and second iteration account for 76.49 and 9.59 with R2 values of 93% and 63%, respectively. Given that the first iteration is statistically significant, occupation (constant variable), income, and risk were considered.

Multinomial Logistic Regression with Training Data

Regression Statistics	
Multiple R	0.893
R Square	0.797
Adjusted R Square	0.789
Standard Error	0.281
Observations	139

F-Test of Overall Significance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	41.26	8.25	104.3	2.83989E-44
Residual	133	10.52	0.079		
Total	138	51.78			

Note the high F value with upper 99% significance for the overall regression model. The sum of squares residual accounts for approximately 20% of the total, which is adequately low.

Regression Scores

Variables	*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Min*	*Max*
Intercept	0.227	0.161	1.406	0.162	-0.19	0.65
Segmentation	0.484	0.042	11.464	0.000	0.37	0.59
Occupation	0.226	0.070	3.223	0.002	0.04	0.41
Education	0.075	0.024	3.134	0.002	0.01	0.14
HH Composition	-0.291	0.045	-6.408	0.000	-0.41	-0.17

Income	0.010	0.002	5.801	0.000	0.01	0.01

All P-values are significant predictors excluding the y-intercept with standard errors in the low 5 – 10%. Additionally, the instrument reveals that HH composition has a negative correlation indicating that land use change may not account for only large families. Income and Risk-taking lifestyle segmentations have the largest predictive outcome. This is followed by Education and then Occupation. All variables are relevant, indicating the possibility of various indicators of land use change among sociodemographic attributes.

Multinomial logistic regression model with trained data

Variables	Maximized Coef.	Probability	Log-Likelihood	Likelihood Ratio Test
Segmentation	33.682	1.00	-1.75	-6.75
Occupation	1.585	0.83	-0.37	-5.37
Education	-0.380	0.41	-0.37	-5.37
HH Composition	-13.602	0.00	-0.37	-5.37
Income	0.414	0.60	-0.37	-5.37
Mean	3.69	0.57	-0.65	-5.65
Standard Deviation	15.75	0.39	0.62	0.62

Risk is a potentially very prominent predictor of land use change followed by occupation and Income. Together, the mean probability is low, but the standard deviation of the log-likelihood is also low. Meaning that education and HH composition were not able to suggest a predictive potential in the change of land use. This is because HHs in our target area mainly small, and do not represent a major indication of livelihood change, while education may seem appropriate to measure livelihood diversification, in reality, it does not stimulate livelihood changes in the region, yet. It is however, well understood that education is a major driver of specialization, but not necessarily an occupational multiplicity in rural areas.

5.4 Transport Component

5.4.1 Travel within the settlement

25 HHs are randomly selected and origins and destinations are calibrated in the model. Travel schedules are revised for individuals with destinations in a 400-meter radius from their lots. Route directness is calculated between each selected lot and their corresponding entrance/exit point.

Change of Job into Node 2

4 vendors from the 36 adults who earn a job in node 2 are found within the 25 randomly selected HHs and their destination is changed to node 2.

Change of Job into Market

4 vendors are found within the selected 25 HHs and their destination is changed to the local market. This is based on couples with a same occupational land-use, which sum up to 30% of settlement population. The decision is also based on their individual itineraries (if they represent a part of the 30% of the population who have a side job as tourist vendors).

Pedestrians

Walking only to very close destinations (Market and Schools near the site)
Non-essential tours in these itineraries reflect stops in-between tour origin and final destination found to provide a measure for non-essential travel in neighborhood areas.

5.4.2 Travel in Total

Congestion
Density and transit volume calculations can be seen in Table S9, where congestion levels substantially rise along the Dairy Farming Route by 2030.

Volume of vehicles per minute = (Density + Through Transit) / (Time Threshold for peak hour estimation) / (mean road width)

*Note: Time threshold deviation at 6:00±30min and 16:00±30min.
**Assumption: subjects with a same routine depart within a 30-minute interval.

Node 1 (Via "Jl. Raya Merapi" to "Jl..Kaliurang")

Density
2019 → 319 (OSM land use "residential" data) = 319 persons
2030 → 1,084 persons (Floating Population) + 66 (growth rate) + 319 = 1,469 persons

Through Transit
2019 → 6:00am: 27.42 persons (survey data x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].
2019 → 16:00pm: 27.42 persons

2030 → 6:00am: 200 persons (survey data x 10.5 → factor of new employment locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].
2030 → 16:00pm: 200 persons

Congestion Node 1
Road mean width: 7.15m + 4.72m / 2 = 5.93m (Section 5.3)

(1) V (2019) = [Density (319) + Through Transit (27.42) / 30 min] / 5.93m = 1.95 persons / meter-wide / min. at 6:00 and 16:00

(2) V (2030) = [(1,084+200) / 30 min] / 5.93m = 7.22 persons / meter-wide / min. at 6:00 and 16:00

Node 2 (Via "Jl. Petung Merapi")

Density
2019 → 140 (OSM land use "residential" data) = 140 persons
2030 → 1,073 persons (Floating Population) + 65 (growth rate) + 140 (OSM polygon "residential" data) = 1,278 persons

Through Transit
2019 → 6:00: 283 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].
2019 → 16:00: 210 persons (survey data sample)

2030 → 6:00: 273 persons (survey data sample x 10.5 → factor of New Employment Locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].
2030 → 16:00: 189 persons

Congestion Node 2
Road mean width: 5.60m + 4.62m / 2 = 5.11m (section 5.3)

V (2019) = [(140+283) / 30 min] / 5.11m = 2.76 persons / meter-wide / min. at 6:00

V (2019) = [(140+210) / 30 min] / 5.11m = 2.28 persons / meter-wide / min. at 16:00

V (2030) = [(1,073+273) / 30 min] / 5.11 = 8.78 persons / meter-wide / min. at 6:00

V (2030) = [(1,073+189) / 30 min] / 5.11 = 8.23 persons / meter-wide / min. at 16:00

Node 3 (Via "Jl. Petung Merapi")

Density
2019 → 501 (OSM polygon "residential" area) = 501 persons
2030 → 1000~ persons (Floating Population) + 250 (growth rate) + 501 (OSM polygon "residential" data) = 1,751 persons

Through Transit (same as Node 2 – 95 people staying at Node 2)
2019 → 6:00 : 283 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].
2019 → 16:00 : 210 persons (survey data sample)

2030 → 6:00 : 178 persons (survey data sample x 10.5 → factor of New Employment Locations statistically applied to Pagerjurang Population (605 persons) in 2030 [Dataset 4 growth rate].
2030 → 16:00 : 94 persons

Congestion Node 3
Road Width = 7.49m (Table 1-Manuscript)

V (2019) = [(501+283) / 30 min] / 7.49m = 3.49 persons / meter-wide / min. at 6:00

V (2019) = [(501+210) / 30 min] / 7.49m = 3.16 persons / meter-wide / min. at 16:00

V (2030) = [(1,751+178) / 30 min] / 7.49 = 8.58 persons / meter-wide / min. at 6:00

V (2030) = [(1,751+94) / 30 min] / 7.49 = 8.21 persons / meter-wide / min. at 16:00

Node 4 (Via "Jl. Wukirsari")

Density
2019 → 159 (Dasymetric Redistribution Data) = 159 persons
2030 → 1000~ persons (Floating Population) + 79 (growth rate) + 159 = 1,238 persons

Through Transit
2019 → 6:00am: 36.56 persons (survey data sample x 9.42 → factor of travel agendas statistically applied to Pagerjurang Population (528 persons) in 2019 [Dataset 4].
2019 → 16:00pm: 54.84 persons (survey data sample)

2030 → 6:00am: 164.5 persons (survey data sample x 10.5 →(Factor of new employment locations statistically applied to Pagerjurang's population (605 persons) in 2030 [Dataset 4].
(128 (census count) + 36.56 (survey data sample))
2030 → 16:00pm: 191 persons (54.84+128)

Congestion Node 4
Road Width = 4.80m

V (2019) = [(159+36.56) / 30 min] / 4.80m = 1.36 persons / meter-wide / min. at 6:00am

V (2019) = [(159+54.84) / 30 min] / 4.80m = 1.48 persons / meter-wide / min. at 16:00pm

V (2030) = [(1,238+164.5) / 30 min] / 4.80 = 9.74 persons / meter-wide / min. at 6:00am

V (2030) = [(1,238+191) / 30 min] / 4.80 = 9.92 persons / meter-wide / min. at 16:00pm

Table S9: Results for Congestion along the Dairy Farming Route by 2030.

Travel Route Via Kaliadem Raya (persons / 15 min)					
Year		Dist (km)	Speed (km/h)	Congestion (6:00)	Congestion (17:00)
2019	2019	1.2	10.77	41	34.2
	2019	1.28	42	0	0
	2019	0.68	10.77	52	47.4
	2019	3.05	42	0	0
2030	2030	0.72	42	0	0
	2030	0.6	10.77	132	123.45
	2030	0.49	42	0	0
	2030	0.68	10.77	129	123.15
	2030	3.05	42	0	0
Travel Route Via Wukirsari (persons / 15 min)					
2019	2019	0.96	42	0	0

	2019	1.44	10.77	26.4	22.2
	2019	1.56	42	0	0
	2019	0.69	10.77	0	0
2030	2030	0.96	42	0	0
	2030	2.95	10.77	146	149
	2030	0.69	10.77	0	0

Table S10: Comparative Urban Form Measures for 2019 and 2030

		2019	2030	Total
	Density Index (person / sq.km)	818	1,589	194%
1	Census Population	3,409	3,590	181
2	Density (person / sq./km)	818	861	43
3	Census Growth HHs (0.5%)	1,774	1,868	94
4	Population (census growth rate)	3,409	3,590	105%
5	Population (with intervention)	3,409	6,193	181%
6	Transit Volume Node 2 (6:00 am)	2.76	8.78	169%
7	Transit Volume Node 3 (6:00 am)	3.49	8.58	159%
8	Transit Volume Node 4 (6:00 am)	1.36	9.74	186%
	Land Use Index (neighborhood shopping in a 400 m radius from main entrance in sq./m)	1,994	6,424	169%
9	Employment Demand		1,944	154%
10	Employment Offer		1,260	100%
11	Entrepreneurs (HBE land use change)		1,023	100%
12	Census Housing Demand (sq./m)		7,030	5.2%
13	Total Housing Demand (sq./m)		130,490	94.8%
14	Edge Housing Demand (Spillover)		27,876	20.3%
15	Housing Supply (sq./m)		15,674	56%
16	Housing Deficit (sq./m)		12,202	44%
17	Total Development (sq./m)		61,386	83.4%
	Accessibility Index	"No"	"Low"	Higher
18	Block Length (m)	30, 55, 85	26, 33, 65	Lower
19	Access Points	14	13	-7%
20	Intersections	57	45	-17%
21	Connectivity	0	1	Higher

6. Public Transportation Assessment

6.1 Estimation

Note: A vehicle with petrol fuel and gas was opted over an electric vehicle considering additional expenses.

A convenient route is traced on GIS using node hierarchies and sub-centers visualized in Fig. S10. A Hyundai H-1, 2.4 Wagon GLS with a consumption level of 10.2 L /100 km and 241 g/km emission rate of CO_2 is selected for semi-urban terrain. The vehicle has a capacity of nine passengers but only six are considered for real-time operation between 5:00am and 5:00pm, Monday through Saturdays. Subjects going to Ngudi Makmur, Desa Petung, or Kaliadem take this mode from node 2. They are selected by age, distance to bus stop, and HH expenditure comparisons without exceeding 30% of all subject candidates as the minimum number considered for sustainable enterprises. The latter assertion results in 13 subjects in total (32%). The choice modelled considers all previous requirements if their daily fuel consumption rate is equal or more than $0.17 daily. It is worth mentioning that the price of a complete tour to Kaliadem (the farthest destination) is $0.12 using public transportation as opposed to $0.17 using a motorcycle. Operation costs are $57.65 and net income adds to $75.00 / month working at 8-route cycles daily with 19.84 VKT per cycle. The results account for simulated HHs (8.3% of the settlement) and gathers a 57% total reduction of VKT, 33% reduction for travel within the settlement, and 34% in fuel savings equal to $1.75 per day as well as 32% in energy savings and 41% in CO_2 emissions.

Vehicle Capacity: 9 Passengers
Motor: 10.2L = 100 km,
1L = 9.80 km,
Petrol: 1L = $0.49,
1 km = $0.05 + 30% Income = $0.065
Cost/km: $0.065

**Assumption: 6 Persons/km (average)

Table S11: Calculation and Estimation for Travel Fares and Fuel Consumption

H1 2.4 Wagon GLS (Petrol $0.49/L, 9 Passenger Capacity)					Calculation with 6 Persons per/Km	
Origin	**Destination**	**Total VKT (one way)**	**Total VKT (Km)**	**Consumption**	**Total Consumption**	**Travel Cost (+30% Income)**
Node 2	Kaliadem	5.28	10.56	0.05 /Km	$ 0.69	$ 0.120
Node 2	Desa Petung	2.48	4.96	0.05 /Km	$ 0.32	$ 0.053
Node 2	Ngudi Makmur	4.525	9.05	0.05 /Km	$ 0.59	$ 0.098
10.2L = 100 Km, 1L = 9.80Km, $0.05 = 1Km					https://www.group1hyundai.co.za/hyundai-h1-bus/	

Destination Kaliadem: 10.56 km x $0.065 = 0.686/6 = $0.12 (Total Fare)
Destination Desa Petung: 4.96 km x $0.065 = 0.322/6 = $0.053 (Total Fare)
Destination Ngudi Makmur: 9.05 km x $0.065 = 0.588/6 = $0.098 (Total Fare)

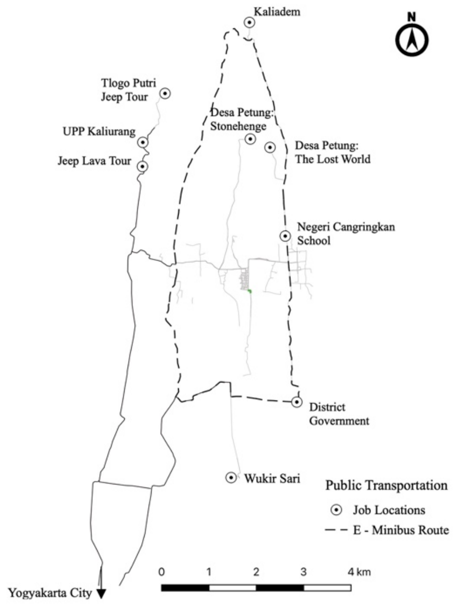
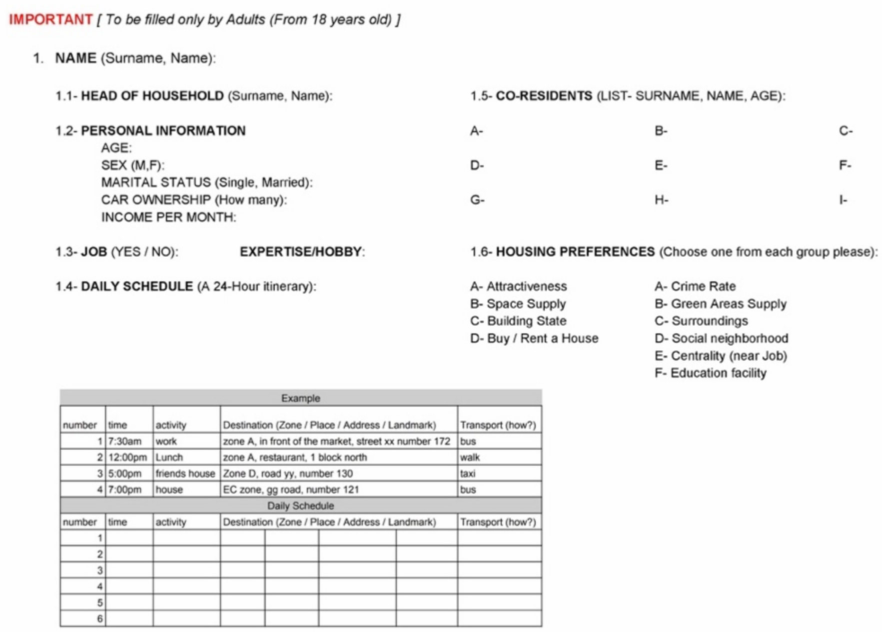


Fig. S10: Tentative Public Transportation Route and Convenient Bus Stops

7. Travel Origins and Household Mapping

Fig. S11: Household numbers in order of interviews used for the first travel simulation (Left). Households randomly selected for the second travel simulation (Right).

8. Survey Questionnaire



IMPORTANT [To be filled only by Adults (From 18 years old)]

1. **NAME** (Surname, Name):

1.1- **HEAD OF HOUSEHOLD** (Surname, Name):

1.2- **PERSONAL INFORMATION**
AGE:
SEX (M,F):
MARITAL STATUS (Single, Married):
CAR OWNERSHIP (How many):
INCOME PER MONTH:

1.3- **JOB** (YES / NO): **EXPERTISE/HOBBY**:

1.4- **DAILY SCHEDULE** (A 24-Hour itinerary):

1.5- **CO-RESIDENTS** (LIST- SURNAME, NAME, AGE):

A- B- C-
D- E- F-
G- H- I-

1.6- **HOUSING PREFERENCES** (Choose one from each group please):

A- Attractiveness
B- Space Supply
C- Building State
D- Buy / Rent a House

A- Crime Rate
B- Green Areas Supply
C- Surroundings
D- Social neighborhood
E- Centrality (near Job)
F- Education facility

Example

number	time	activity	Destination (Zone / Place / Address / Landmark)	Transport (how?)
1	7:30am	work	zone A, in front of the market, street xx number 172	bus
2	12:00pm	Lunch	zone A, restaurant, 1 block north	walk
3	5:00pm	friends house	Zone D, road yy, number 130	taxi
4	7:00pm	house	EC zone, gg road, number 121	bus

Daily Schedule

number	time	activity	Destination (Zone / Place / Address / Landmark)	Transport (how?)
1				
2				
3				
4				
5				
6				

Fig. S12: Census, Housing Preferences, Activity-Based Schedules (Page 01)

2. **LIFESTYLES** (ARELLANO)

2.1- **What is your main occupation now?**

(A) dependent worker (I work for a company I like to buy the latest technology products and I am not the owner)
(B) independent worker (I have my own business or company)
(C) my home (I dedicate myself to my home most of the time)
(D) Student (I study most of the time)
(E) eventful (I work from time to time)
(F) Pension
(G) unemployed (more than 6 months)
(H) live from rents / receive donations

2.2- **What is the last degree of study you reached?**

(A) primary school
(B) incomplete middle school
(C) complete middle school
(D) incomplete high school
(E) complete high school
(F) short career (less than 2 years)
(G) higher technical career incomplete
(H) higher technical career complete (more than 3 years)
(I) superior university career incomplete
(J) superior university career complete
(K) Master Degree (2 years)
(L) Doctoral degree (PhD)
(M) Post bachelor courses (specialized courses)
(N) none (didnt study)

2.3- **Now, how do you perceive ?**

your personal laboral situation:	Personal economic situation:
(A) 0 (very bad)	**(A)** 0 (very bad)
(B) 1	**(B)** 1
(C) 2	**(C)** 2
(D) 3	**(E)** Selalu
(E) 4	**(E)** 4
(F) 5	**(F)** 5
(G) 6	**(G)** 6
(H) 7 (very good)	**(H)** 7 (very good)

2.4- **How much do you agree or disagree with the following phrases?**

I like to buy the latest technologies (...)
(A) 3 (strongly disagree)
(B) 2 (very much disagree)
(C) 1 (A little disagree)
(D) 1 (A little agree)
(E) 2 (very much agree)
(F) 3 (strongly agree)

I usually buy products at lower prices (...)
(A) 3 (strongly disagree)
(B) 2 (very much disagree)
(C) 1 (A little disagree)
(D) 1 (A little agree)
(E) 2 (very much agree)
(F) 3 (strongly agree)

2.5- **According to the next phrases : How often do you?**

Keep up with the latest fashion (...)
(A) never
(B) A few times
(C) Sometimes yes, sometimes no
(D) Almost always
(E) Always

Blame your personal situation to the government
(A) never
(B) A few times
(C) Sometimes yes, sometimes no
(D) Almost always
(E) Always

Participate in promotions or in offers
(A) never **(D)** Almost always
(B) A few times **(E)** Always
(C) Sometimes yes, sometimes no

Fig. S13: Lifestyle Questionnaire (in Spanish: **www.arellano.pe**) (Page 02)
*Note: Other suitable segmentation questionnaires for developing countries can be used.

References

[1] MDA Federal Inc. (1995–2007) EarthSat GeoCover LC Overview. Rockville, MD: MacDonald, Dettwiler and Associates Ltd. http://www.mdafederal.com/geocover/geocoverlc/gclcoverview In: Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013), "High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015". PLoS ONE 8 (2), p. 4. Apr. 19, 2021. http://doi:10.1371/journal.pone.0055882

[2] Nirav N. P., et al., (2015), "Multitemporal settlement and population mapping from Landsat using Google Earth Engine, International Journal of Applied Earth Observation and Geoinformation, Vol. 35, Part B, pp. 199-208. https://doi.org/10.1016/j.jag.2014.09.005

[3] Mennis, J., and Hultgren, T., (2006), "Intelligent dasymetric mapping and its application to areal interpolation." Cartography and Geographic Information Science, vol. 33, no. 3, p. 186. Accessed 24 Apr. 2021. https://www.tandfonline.com/doi/abs/10.1559/152304006779077309

[4] Gaughan AE, et al., (2013), "High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015", PLOS ONE, vol. 8, Issue 2, pp.1-11.https://doi.org/10.1371/journal.pone.0055882

[5] Stevens FR, et al., (2015), "Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data", PLOS ONE, vol. 10, Issue 2, pp.1-22. https://doi.org/10.1371/journal.pone.0107042

[6] Linard C, et al., (2011), "Assessing the use of global land cover data for guiding large area population distribution modelling", GeoJournal, vol.76, Issue 5, pp.525-538. https://doi.org10.1007/s10708-010-9364-8

[7] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). "Google Earth Engine: Planetary-scale geospatial analysis for everyone", Remote Sensing of Environment. Accessed: May 05, 2021.

[8] Garcia–Fry, M., Murao, O., (2020), "A Post–disaster Forecast Analysis for Relocation Settlements Using a Combined Model Sequence", 17th World Conference on Earthquake Engineering, N.C000108.

[9] Alonso, W. (1964), "Location and Land Use. Toward a General Theory of Land Rent", Cambridge: Harvard University Press. Accessed: August, 2019. http://dx.doi.org/10.4159/harvard.9780674730854

[10] Krizek, K. J. (2003), "Operationalizing Neighborhood Accessibility for Land Use–Travel Behavior Research and Regional Modeling" *Journal of Planning Education and Research*, Vol. 22, No. 3, pp. 270–287. **doi.org/10.1177/0739456X02250315**

[11] "Hyundai H1 Bus – Price, Specs, Fuel Consumption", Hyundai Vehicles at Group 1 Motors, Accessed: May, 2020. https://www.group1hyundai.co.za/hyundai–h1–bus/

Dataset Citations

1. Indonesia, Statistics Sleman Regency, “Population and Population Growth Rate by Subdistrict in Sleman Regency, 2010, 2016, and 2017”, Badan Pusat Statistik Kabupaten Sleman, 12, Oct. 2018. Accessed: May 19, 2020. **https://slemankab.bps.go.id/statictable/2018/10/12/363/jumlah–penduduk–dan–laju–pertumbuhan–penduduk–menurut–kecamatan–di–kabupaten–sleman–2010–2016–dan–2017.html**

2. Indonesia, Statistics Sleman Regency, “Population Distribution and Density by Subdistrict in Sleman Regency, 2018”, Badan Pusat Statistik Kabupaten Sleman, 09, Jul. 2019. Accessed: Apr. 19, 2021. **https://slemankab.bps.go.id/statictable/2019/07/09/520/distribusi-dan-kepadatan-penduduk-menurut-kecamatan-di-kabupaten-sleman-2018.html**

3. OpenStreetMap (2021) OpenStreetMap The Free Wiki World Map. Open-Street Map contributors, CC BY-SA. Accessed April 15, 2021. **http://www.openstreetmap.org/**

4. Monografi Desa Kepuharjo, 2017 in: Ruswantoro, R., 2018. Evaluation of the Impact of Government Policies on Relocation Post Eruption of Merapi 2010, [Doctoral Dissertation, Sekolah Tingghi Pengambunan Masyarakat Desa "APMD"] APMD Repository, p.44 – Table 5. Access on Apr. 20, 2021. **http://repo.apmd.ac.id/618/1/repo%20Tesis%20Ruwantoro.pdf**

5. Data Sekunder Laporan Profil Desa Kepuharjo, 2015 in: Wibowo, P., D., 2016. Kebutuhan Pelayanan Dasar Desa Di Daerah Rawan Bencana. PT Sulaksana Watinsa Indonesia, p. 46 – Table 8. Access Apr. 20, 2021. https://www.academia.edu/40312242/Kebutuhan_Pelayanan_Dasar_Desa_Di_Daerah_Rawan_Bencana

6. Indonesia, Yogyakarta DIY, (2019), “Total Population of Cangkringan district, Sleman district, DI Yogyakarta, Semester II-2019 According to Age Group”, *Data on the Consolidation and Clearance of Population Databases by the Directorate General of Population for Civil Registration of the Ministry of Home Affairs, processed by the DIY Secretariat's Administration Bureau. Accessed Apr. 26, 2021.* ***https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/13/5/17/04/34.ez***

7. Garcia-Fry, Martin (2021), “OLUTM Model Dataset”, Mendeley Data, V3, doi: https://data.mendeley.com/datasets/t9p23k3pyn/3