Spatial Planning Under Data Paucity: Dasymetric Interpolation of Population, Validated by Google Earth, to Support Health Facility Location Modelling

Idris Mohammed Jega [1], Alexis J. Comber [2], Nicholas J. Tate [3]

[1]*Strategic Space Applications, National Space Research and Development Agency, Abuja, Nigeria*
[2]*School of Geography, University of Leeds, Leeds LS2 9JT, UK*
[3]*Department of Geography, University of Leicester, Leicester LE1 7RH, UK*

Abstract

Small area population estimates are important for facility location-allocation analyses as they provide a spatial distribution of 'demand' against which potential 'supply' locations are evaluated. However, in many parts of the world small area estimates of census data are not available which makes it difficult to validate and constrain interpolated population or demand surfaces. For such cases a number of interpolation methods have been proposed to redistribute summary population census totals over small areas. Binary dasymetric interpolation has been shown to perform well across a range of spatial scales and resolutions supported by ancillary data. To date no published research describes the use of dasymetric approaches as methods for disaggregating population data over small areas in any part of West Africa. This study applies a binary dasymetric approach to generate population surface for Port-Harcourt, Nigeria, an area for which small area population estimates are unavailable. This was validated by visual inspection using Google Earth and found to be 87% correct. The demand surface was then used as input to a location-allocation analysis of health facility provision. The locations of current primary healthcare centres (PHCCs) were evaluated and then alternative, improved spatial arrangement for the facilities that optimised the spatial distribution of supply with that of demand, were suggested. The results show 13 alternative sites for PHCCs to be located would provide almost the same demand coverage as the 17 current locations of PHCCs. The analysis indicated that fewer PHCCs in a different spatial arrangement could satisfy the spatial distribution of demand. The results suggest that such methods can be used to support and inform decision making and spatial planning in countries where only limited socio-economic data exist.

Keywords *small area population estimates, binary dasymetric mapping, health facility location-allocation, spatial planning.*

I. INTRODUCTION

Census data at the small area level such as output areas (OAs) in the UK are unavailable in most parts of the world. This is especially the case in many developing countries. In such places the need to disaggregate coarse population counts to small areas to represent demand populations is important in order to support robust spatial planning [1, 2]. Specifically, calibrated solutions for estimating populations over small areas are needed to provide repeatable and transparent facility location analyses.

Interpolation has at its core the aim of redistributing data from spatially coarse, high level entities (source zones) to spatially detailed lower level ones (target zones) in order to generate models or estimates of the actual spatial distribution of the phenomenon under consideration. Dasymetric interpolation involves the use of ancillary data to mask out areas over which interpolation estimates should not be generated. The source zone is divided into populated and unpopulated regions, and the source zone population is then redistributed to only the populated regions [3-5]. The advantage of the dasymetric approaches is that they eliminate sharp differences at source zone boundaries and have been found to reduce errors associated with within-source-zone uniformity [6]. As noted by Cai et al. [7], the dasymetric model has been found to give better information about the distribution of population in many studies related to environmental justice [8], public service accessibility [6], environmental health [9], creation of land use maps from agricultural census data [10] and crime analysis [11, 12].

Dasymetric interpolation using land cover as ancillary data offers an approach that has consistently been found to out-perform other interpolation approaches [3, 11, 13-15]. Dasymetric mapping has a long history [16, 17] but its basic idea is to constrain areal interpolation to areas for which the phenomenon under investigation is likely to exist. For estimating spatially distributed populations, the constraint is typically to limit the extent of the spatial interpolation to residential areas. It is an approach that provides a clear understanding of population distribution within area being considered [18, 19]. For this reason, different studies have used different ancillary data as constraints including land cover data [3, 4, 18], road networks [20, 21], cadastral data [22], address points [23], and parcel data [24].

The most commonly used ancillary data in dasymetric mapping has been land cover information derived from classified satellite imagery [13, 25]. However, a review of dasymetric mapping literature shows the development of several techniques using different ancillary data types (see Table I) all with the aim of improving interpolation accuracy [4]. Table I shows a selection of previous literature reporting different ancillary data types used in dasymetric mapping. The contents presented in the table are arranged alphabetically by ancillary data types.

Table I Ancillary Data Types Used in Dasymetric Mapping

S/No	Ancillary data	Author(s)/year
1	Address points	Harris and Longley [29]; Zandbergen and Ignizio [23]
2	Aerial photographs	Green [30]
3	Cadastral data	Maantay et al. [22]; Bentley et al. [31]
4	Control zone	Goodchild et al. [32]
5	Housing distribution	Moon and Farmer [33]; Poulsen and Kennedy [11]; Leyk et al. [34]
6	Image pixels	Harvey [35]; Holt et al. [36]
7	Image texture	Chen [37]; Liu et al. [38]
8	Land cover data	Langford [13]; Mennis [4]; Ward et al. [39]
9	Night-time lights	Pozzi et al. [40]; Briggs et al. [41]
10	OS VectorMap District	Langford [13]
11	Parcel data	Tapp [24]
12	Raster pixel maps	Langford [5]
13	Satellite imagery	Langford and Unwin [18]; Wu et al. [42]
14	Three dimensional LiDAR data	Sridharan and Qiu [43]
15	Topographic sheet	Wright [16]
16	Vector GIS	Eicher and Brewer [3]

The aim of this study was to generate demand population values in an area where their actual distribution is unknown. Spatially distributed estimates of the population are important for many types of spatial data analysis. This research applies a dasymetric approach to develop spatially distributed estimates of population in an area for which spatially detailed census summaries are not available. This is problematic as the usual approach in any spatial interpolation is to compare the model outputs (estimates) with some form of validation data. However, if these do not exist then it is instructive to consider the findings of others who have compared different interpolation approaches, across a range of problems, across a range of scales. Many researchers have found that dasymetric interpolation with urban land cover consistently provide better estimates of population than other approaches [3, 11, 13-15, 26-28].

II. BACKGROUND

In Nigeria, population counts are collected for each household and censuses have been held in 1963, 1973, 1991 and 2006. Census data are published only for States and Local Government Areas (LGAs), highly spatially aggregated areas. The aggregation to LGA level, purportedly for confidentiality reasons, poses a serious problem for effective and transparent spatial planning for example in healthcare. To date there is no published research reporting the use of dasymetric interpolation method to estimate aggregate census data over small areas in any part of West Africa. However, population data for small areas have long been more generally found to provide information on local population characteristics that assist in coordinating, monitoring and evaluating service delivery [44]. Reliable, spatially detailed population estimates are essential to support economic development, management decisions, disaster management, and urban and regional planning [45]. Providing information describing the weight and distribution of demand over the region of interest also allows analyses of facility locations and population allocation.

This study applied the dasymetric interpolation method to estimate summary population totals over small areas of unknown distributions in Port-Harcourt, Nigeria. An insight was obtained into the chosen interpolation method via work on Leicester data [46]; although it can be argued that the best performing parameters for an area may not likely provide the best fitting target zone estimates for another different area, but in the absence of alternatives we can only make use of what we know.

Population allocations to small areas where there are rich data available at a variety of spatial scales (e.g. UK) are commonly validated by comparing the estimates with some known values for those areas. However, there is an important question which is: How can a population surface be determined in areas of unknown distributions and with no validation data? This is an important issue as the spatial distribution of demand is a critical input for determining healthcare facility provision to ensure coverage. Cromley et al. [47] and Langford and

Higgs [19] have discussed the significant contribution of the spatial distributions of demand to potential measure of access to health facilities in the absence of small area census data. This background provides the primary motivation for the research reported in this paper.

III MATERIALS AND METHODS

A. Study area

Port-Harcourt is the capital of Rivers State in Nigeria. It is the country's second largest commercial centre. Port-Harcourt covers an area of about 109 square kilometres with a population of 541,115 in 126,010 households [48]. It lies on the coastal plain of eastern Niger Delta along Bonny river located in Nigeria's oil rich Niger Delta [49]. A map of Nigeria highlighting the location of Rivers State is shown in Figure 1a. Figure 1b shows map of Rivers State with Port-Harcourt LGA in dark shade around the middle of the map. Figure 1c shows the boundary map of Port-Harcourt Local Government Area.

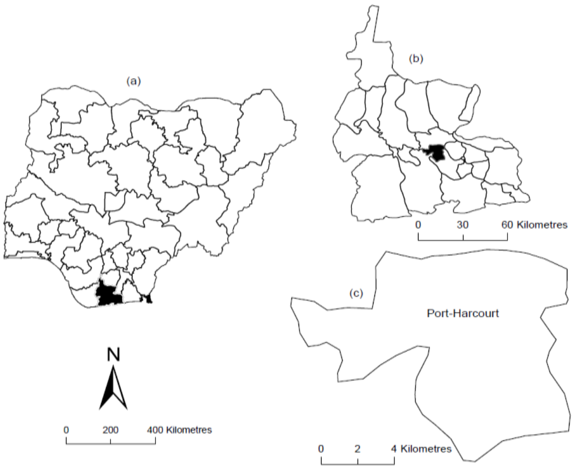

Fig 1: The map of (a) Nigeria showing location of Rivers State; (b) Rivers State showing location of Port-Harcourt; (c) Port-Harcourt City Local Government Area. The digital boundaries are Copyright for Geo-technics Services 2011.

B. Data

Supply and demand data are needed for location-allocation analyses. The locations of current health facilities are known and thus the first part of this paper was to generated estimates of the spatial distribution of demand over small areas of unknown distributions in Port-Harcourt. Table II shows the data acquired for implementing areal interpolation for Port-Harcourt. The table also shows the format and source of the data.

C. The binary dasymetric method

The dasymetric mapping is an areal interpolation technique that incorporates ancillary data sources as control variables in order to identify zones having different population densities [18, 50]. It constrains the disaggregation of population values from source zones to specific target zones, which can be weighted for example by expected residential density [3, 4, 18, 24, 45, 51]. Binary dasymetric approach [18] divides source zones into populated and unpopulated areas and allocates population only to the populated areas. In this study, land cover data was obtained from Landsat7 image (see Table II) to identify built-up and non-built-up areas using a supervised classification technique. This provides one of the key inputs to the binary dasymetric method and the 30m framework was used in the analysis as the support zones, over which target zones would be created. A maximum likelihood classification algorithm was applied and the classified image with the highest accuracy (81.64% in this case) was selected as the input. Accuracy was assessed by comparing 256 random points to reference pixels for which the class was known. Figures 2 and 3 show classified Landsat7 (ETM+) 30m spatial resolution image and a binary mask derived from the classified image respectively.

Table II Data for Port-Harcourt

Data	Format	Source
*Landsat7 (ETM+) 30m spatial resolution (acquired 8th January 2003) with WRS_PATH=188 and WRS_ROW=057	Image	United States Geological Survey (USGS) website (http://www.usgs.gov/)
Population census (2006) with priority tables	Excel	National Bureau of Statistics' website http://www.nigerianstat.gov.ng/
States and LGAs boundary	Shapefile	Geo-technics Services Limited, Port-Harcourt

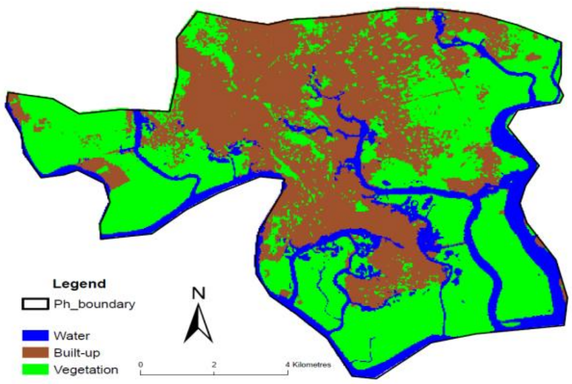

Fig 2: The classified Port-Harcourt image derived from Landsat7 (ETM+) 30m spatial resolution. The digital boundary is Copyright for Geo-technics Services 2011.

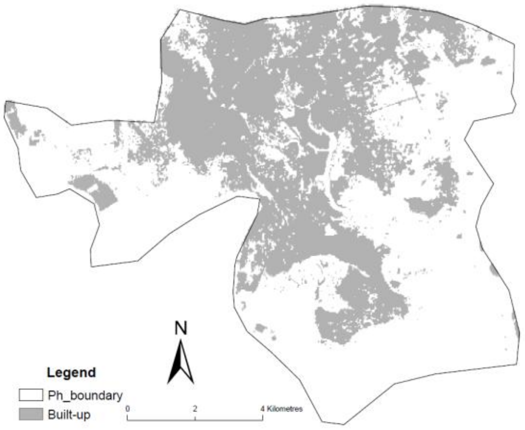

Fig 3: A Binary Mask Derived from Land Cover Data Derived from Classified Landsat7 (ETM+) 30m Spatial Resolution Image Data. The Digital Boundary Is Copyright for Geo-Technics Services 2011.

The binary dasymetric method involves a number of steps applied as part of a PhD [46] looking at Leicester data: (1) a specially overlay of the ancillary data with the source zone; (2) removing areas of overlap; (3) recalculating source zone area; (4) calculating the source zone population density; (5) overlaying the 30m support grid with the areas of intersect in Step 1 above; (6) calculating the areas of the newly overlaid grids; (7) calculating population estimates for each overlaid grid by multiplying the area with the density; (8) overlaying the target zones with the population for each overlaid grid; and (9) obtaining the interpolated population of each target zone by summing all the populations for each overlaid grid within each target zone. A flowchart describing these steps is shown in Figure 4.

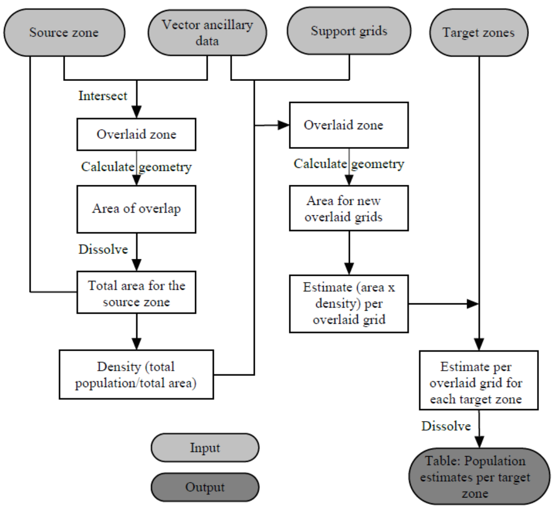

Fig 4: Implementation Steps for the Binary Dasymetric Method (Source: Jega, I.M., 2015)

D. Accuracy assessment

In the absence of field information (survey data) or another data source to validate the modelled population surfaces, a random sample of 200 locations was taken to visually inspect the results in order to assess the accuracy of the population distribution. The randomly placed points were generated to have a shortest distance of 200m between any two random points (see Figure 5). Google Earth 7.1 was used as a reference to validate the surfaces. The error matrix is summarised in Table III. It shows that some of the randomly selected points on the populated surfaces do not correspond to a populated area when viewed using Google Earth and some other random points on the unpopulated surfaces were identified on Google Earth in populated areas. This is likely due to land cover changes between the date of the Google Earth reference (20th December 2013) and the date the Landsat7 image was acquired (as shown in Table II). The effect of using built-up areas to represent populated areas is seen where some of the random points on the populated surface were identified in an industrial area. The summary of reference data in Table III shows sixteen random points identified in populated areas on the interpolated surface were actually in unpopulated areas on Google Earth reference. There were ten random points identified in unpopulated areas on the interpolated surface that correspond to populated areas on Google Earth reference. The accuracy totals in Table IV shows the validation of interpolated surfaces using Google Earth reference was found to be 87% correct.

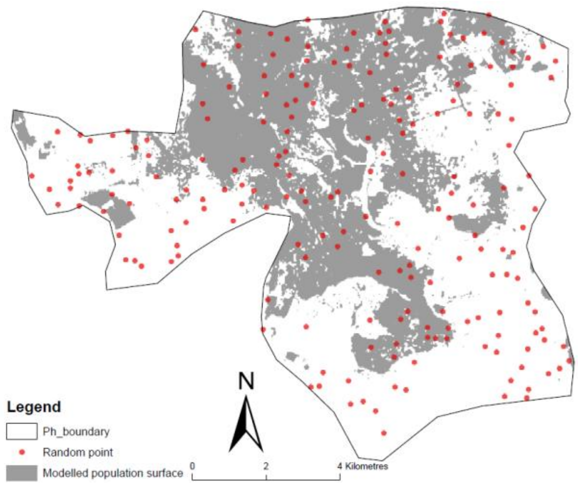

Fig 5: The Demand Surfaces for Port-Harcourt with 200 Random Points Generated Within the Boundary of Port-Harcourt. The Digital Boundary Is Copyright for Geo-Technics Services 2011.

Table III Summary of Reference Data

		Interpolated surface		
		Populated	Unpopulated	Row Total
Google Earth Ref.	Populated	75	10	85
	Unpopulated	16	99	115
	Column Total	91	109	200

Table IV Accuracy Totals

Class Name	Ref. Totals	Class Totals	Number Correct	Producers Accuracy	Users Accuracy
----------	----------	----------	-------	---------	-----
Pop	91	85	75	82.42%	88.24%
Unpop	109	115	99	90.83%	86.07%
Totals	200	200	174	87.00%	87.00%
Overall Classification Accuracy = 87%					
----- End of Accuracy Totals -----					

IV. LOCATION-ALLOCATION ANALYSES

The aim was to evaluate the locations of 17 current PHCCs in Port-Harcourt and then suggest alternative, improved spatial arrangement for the facilities, with an objective of optimising the spatial distribution of supply with the spatial distribution of demand. The p-median problem addresses this objective.

The p-median problem was introduced by Hakimi [52] with an objective function of selecting *p* facilities (among the total facilities) that minimises the total weighted distance travelled (or time) between facilities and demand points [53, 54], thereby maximising accessibility. There are two basic approaches to solve the p-median problem: optimal and heuristic [55]. The most robust of the optimisation techniques is the Lagrangian Relaxation with sub-gradient optimisation developed by Narula et al. [56] but this technique requires too much computation time to solve [55]. A number of heuristic procedures have been developed to help solve the p-median problem. They include: [53, 57-59], Simulated Annealing, GRASP and a combination of two or more procedures to form the hybrid approach [55]. The choice of heuristic procedure to solve the p-median problem in a GIS system must be based on its robustness, speed, simplicity and ease of integrating into existing data structure and software. Teitz and Bart [53] suggested an alternative approach to solve the p-median problem by adding an interchange heuristic to the model. This controls the selection of locations that are more likely to reduce the average weighted distance (or time) from demand to all locations. It is a "proven approach, easy to program, relatively fast, easy to explain, and produces good results" [55].

The objective function of the p-median method is to select the required number of facilities among the total potential facilities that minimises the total weighted distance travelled between facilities and demand points. The Teitz and Bart's heuristic search algorithm requires the number of facilities to locate, the demand population values and the distance matrix. The algorithm first randomly searches for the required number of facilities, then substitute one of the selected with one not selected and tests to verify if the average weighted travelled distance from demand to all locations is minimised. This is repeated until no distance is minimised by the substitution then the heuristic stops and the selected locations are assumed to be the optimal locations.

In this study, Teitz and Bart's heuristic search algorithm was used to solve the p-median problem and to minimise the total demand weighted distance in Port-Harcourt. The inputs into Teitz and Bart's heuristic search algorithm are: the number of facilities to locate, the demand population values generated from dasymetric interpolation, potential supply locations and distance measures between potential supply locations and demand locations. The data are shown in Table V. The potential supply locations (85) were generated using a 500m grid across Port-Harcourt city and grid points that were within 30m of an existing road were selected. The demand population values created using the binary dasymetric interpolation were converted to 56,457 grid points spaced at 30m apart with each point representing estimates of the population redistributed from 2006 census totals for Port-Harcourt. The road distance from each demand point to each current PHCC was calculated and to 85 grid points representing potential new locations of PHCCs.

Table V Data for the Case Study Analyses

Data	Format	Source
Demand population values	Grid points	Created from binary dasymetric interpolation method.
Current PHCCs in Port-Harcourt	Points shapefile	GPS locations directly observed by the analyst during field trip.
Road network data	Shapefile	Geo-technics Services Limited, Port-Harcourt
States and LGAs boundary	Shapefile	Geo-technics Services Limited, Port-Harcourt

The process starts with a configuration of 17 facility locations. This is because there are 17 current PHCCs in Port-Harcourt. The 17 selected locations in the subset are called the *p*-facility locations while the potential locations not in the subset are the candidates. The heuristic search algorithm selects a candidate site and swaps this candidate for each of the current *p*-

facility locations. Any swap that improves weighted distance, replaces the *p*-facility location. The process then continues by selecting another candidate site, testing swaps and replacing the *p*-facility location where there is an improvement in weighted distance. The heuristic stops when no swap between candidate and *p*-facility location improves weighted distance. The spatial distribution of 17 current and 85 potential new locations of PHCCs and the demand points generated from dasymetric interpolations are shown in Figure 6.

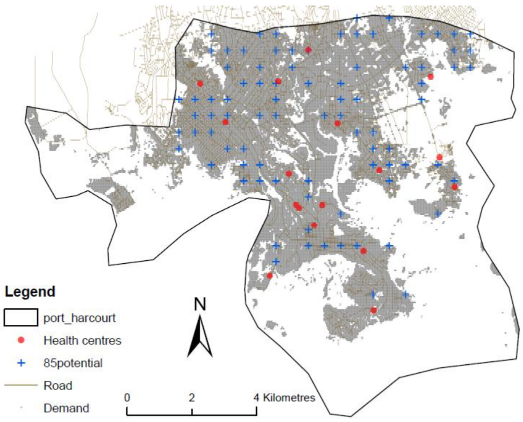

Fig 6: The Spatial Distribution of 17 Current (Red Circles) and 85 Potential (Blue Crosses) Locations of Phccs And The Demand Points Generated From Dasymetric Interpolations, Overlaid on the Road Network. The Digital Boundary is Copyright for Geotechnics Services 2011.

Some assumptions were made by the location-allocation analyses relating to both demand and supply. On the demand side, the assumption is that the majority of the population will attend the nearest facility and that all of the facilities that are available are within the boundary of Port-Harcourt. In reality, PHCCs located outside the boundary of Port-Harcourt may be preferred by some residents especially if they are closer to their places of work, for example. On the supply side, the assumption is all of the facilities have the same capacity and provide the same services. In reality this may not be the case.

Three analyses were run to: (1) assess the locations of 17 current PHCCs using an evaluation function to minimise the demand weighted distance. Sets of locations that minimises the person-distances are assumed to be optimal; (2) suggest new set of potential locations for 17 facilities using the 85 potential locations, again with an evaluation function to minimise demand weighted distance; (3) suggest the number of facilities needed using both current and new locations to satisfy different levels of demand.

V. RESULTS AND DISCUSSIONS

A) Current PHCC locations

Table VI shows the total demand allocated to each of the 17 current locations of PHCCs, the mean distances between each PHCC and each demand within its catchment and the demand weighted distances (Persons/Km). The average mean distance is 1204m. This result is within the international criteria of 2km maximum distance a patient is expected to travel to an urban health centre that provide comprehensive primary care, as outlined by Centre for Health Policy [60] and discussed in [61, 62]. The data in Table VI show the first five PHCCs were allocated about 54% of the total demand with the remaining twelve PHCCs having about 46% of the total demand. Considering the assumption that these PHCCs have the same capacity and provide the same services, it is expected that the demand allocation should be equal for an optimal location. This means the current locations of PHCCs are not optimal. The implication of this result is that PHCCs with high demand allocation will be overstretched while those with low demand allocation will be underutilised. The total demand weighted distance for the current set of PHCC locations is 720,464 Persons/Kilometre. This result provides evidence for informed decision making in spatial planning and policy development.

Table VI Demand Allocated to 17 Current PHCCs in Port-Harcourt

S/no	Health Centres	Demd	Demd (%)	Mean dist.(m)	Pers/Km
1	Okija Street	65001	12.68	1672	108682
2	Azuabie	57179	11.16	2717	155355
3	Amadi-Ama	56645	11.05	1383	78340
4	Orogbum	53543	10.45	1263	67625
5	Mile3 Fsp Clinic	41784	8.15	1343	56116
6	Borikiri	41133	8.03	1306	53720
7	Elekahia	32059	6.26	1151	36900
8	Ozuboko Potts	28393	5.54	946	26860
9	Johnson	24911	4.86	1005	25036
10	Churchill Street	23115	4.51	813	18792
11	Bmh Immun. Post	20683	4.04	1132	23413
12	Marine Base	18337	3.58	1156	21198
13	Bundu Ama	17263	3.37	968	16711
14	Abuloma	12838	2.51	784	10065
15	Bank Road	10203	1.99	1472	15019
16	Okuru-Ama	5273	1.03	897	4730
17	City	4090	0.80	465	1902

Council

B) *Potential locations of PHCCs*

The spatial distributions of the 17 optimal locations selected from 85 potential locations were mapped together with the location of 17 current PHCCs (see Figure 7). The details of the locations and their allocation are shown in Table VII along with the mean distances between each optimal location and the demand within its catchment and their demand weighted distances (Persons/Km). The average mean distance for the 17 optimal locations was found to be 1074m. This indicates a reduction in the average distance travelled when compared with the current locations of PHCCs by 130m. It can be argued that for this set of experiment, the results show little improvement in the current locations of PHCCs but the demand serviced by each of the locations is now much more even and the total demand weighted distance is now 569,084 Persons/Kilometre suggesting that considerable improvement on current locations is possible with large impacts for many people.

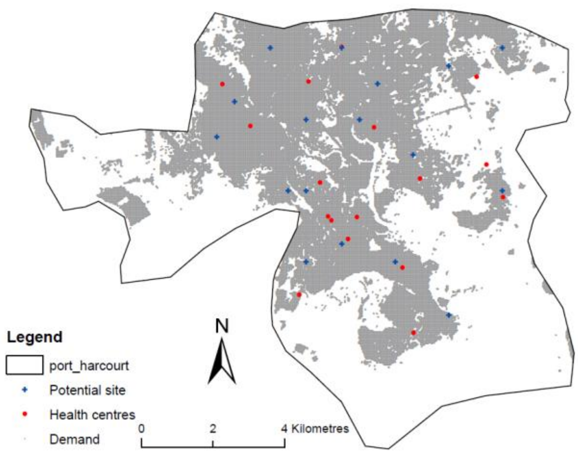

Figure 7 Spatial Distribution of 17 Optimal Locations Selected from 85 Potential Locations (blue crosses) and 17 Current Locations of PHCCs (red circles). The Digital Boundary is Copyright for Geotechnics Services 2011.

C) *Adjusting the Number of Locations*

The analysis was run to select sets of 5 to 16 PHCCs from current locations and to select sets of 5 to 20 locations from 85 potential new locations. In both cases the mean distances to the nearest PHCC (in metres) and the total demand weighted distances were considered. The results in Table VIII show the current distances, modelled distances and the difference between current and modelled distances for specific number of PHCCs used as subsets. The average distances (in metres) to the nearest current locations (current distances) and optimal locations selected from 85 potential new locations (modelled distances) were plotted against the number of PHCCs in a subset (see Figure 8).

Figure 8 shows a graph of average distances (in metres) to the nearest current locations (current distances) and optimal locations selected from 85 potential new locations (modelled distances) plotted against the number of PHCCs in a subset. The graph shows the average travelled distance for both current and modelled distances reduces as the number of locations in the subset increases but that it reduces more with the optimally selected locations. While the distances may be small (as shown in Table VIII) many fewer people have further to travel to their nearest PHCC and access to treatment within a very short time period is often critical in many medical emergencies. The results suggest that 13 optimally sited PHCCs points selected from 85 potential new locations provide the same coverage as the current 17 facilities. The results also suggest considerable cost savings could be achieved without affecting service delivery if fewer but new locations are used.

Table VII Demand Allocated to Potential PHCCs in Port-Harcourt

S/no	Health Centres	Demand	Demand (%)	Mean dist.(m)	Person /Km
1	2	43625	8.51	1704	74337
2	41	43542	8.50	1115	48549
3	10	40839	7.97	1141	46597
4	62	38070	7.43	1256	47816
5	30	37194	7.26	1053	39165
6	68	35797	6.99	1080	38661
7	56	35755	6.98	1001	35791
8	33	34986	6.83	1113	38939
9	71	31854	6.22	1045	33287
10	46	31305	6.11	1038	32495
11	16	23883	4.66	1129	26964
12	8	21508	4.20	784	16862
13	42	21501	4.20	795	17093
14	4	19265	3.76	869	16741
15	15	18668	3.64	1173	21898
16	19	17498	3.41	1008	17638
17	72	17160	3.35	947	16251

This study applied the binary dasymetric interpolation method to generate demand population values over small areas of unknown distributions in Port-Harcourt, Nigeria. The accuracy of the interpolations was assessed using google earth reference and it was found to be 87% correct. The choice of the interpolation method was via work on Leicester data [46] where the actual population distribution is known and the results were validated across a range of spatial scales and resolutions supported by ancillary data. It can be argued that the best performing technique for an area may not likely provide the best fitting target zone estimates for another different area, as has been found by

Zandbergen and Ignizio [23] who compared dasymetric mapping techniques in four US States, and no single technique performed best in more than one state. But the spatially distributed population surfaces generated in this study showed estimates were allocated to only areas identified as built-up and the surfaces have a potentially useful degree of precision based on the results of the accuracy assessment of some random points using google earth reference. The interpolated surfaces also depict the underlying population distribution in Port-Harcourt. The high accuracy of the interpolated surfaces was not surprising because Fisher and Langford [63] examined the sensitivity of the population estimates to error in the classified imagery with the assumption that errors are spatially random, and the results show the dasymetric method to be robust to classification error. A robust technique is more likely to "perform well for different study areas across a range of different conditions" [23].

The demand surfaces were then used to evaluate the locations of current PHCCs and then to suggest alternative spatial arrangements for health centres in Port-Harcourt, Nigeria. The total demand weighted distance was found to be 720,464 Persons/Km and 569,084 Persons/Km for the current set of PHCC locations and for the alternative spatial arrangements for health centres respectively. This result suggests that considerable improvement on current locations is possible with large impacts for many people. Distance plays a vital role in evaluating spatial access to health services and analysis of distance from patients' homes to the nearest health centre is an objective indicator of geographic accessibility to health services [64]. In such situations, the most appropriate method is to measure actual travel distance (or travel time) on a road network [65]. Travel distance to health facilities from homes has been shown to be a barrier to accessing health centres [66, 67].

Table VIII Current and Modelled Average Distances for Facilities

PHCCs	Current (m)	Modelled (m)	Difference (m)
5	2141	2017	124
6	1968	1839	129
7	1837	1716	121
8	1624	1604	20
9	1565	1476	89
10	1488	1348	140
11	1454	1307	147
12	1392	1255	137
13	1324	1206	118
14	1292	1162	130
15	1270	1130	140
16	1235	1105	130
17	1204	1074	130
18	-	1046	1046
19	-	1033	1033
20	-	996	996

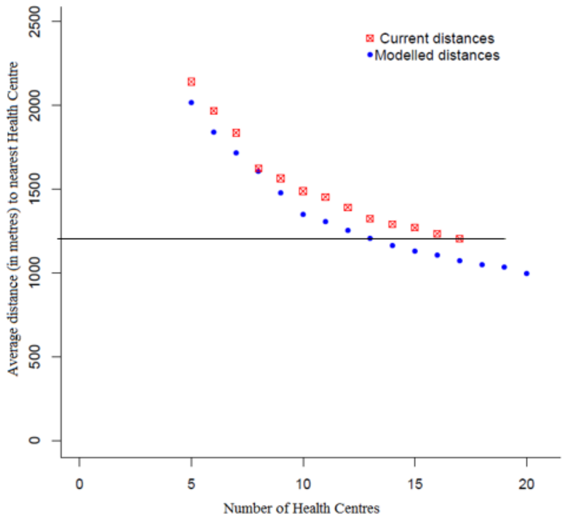

Fig 8 Variation in Average Distance Travelled from Demand To: (⊠) Nearest Location of Current Phccs (Current Distance); And (●) Nearest Potential Location Selected From 85 Potential New Locations.

The analysis indicated that fewer PHCCs in a different spatial arrangement could satisfy the spatial distribution of demand. This is very important as it shows the use of these demand population values to obtain optimal locations of health centres. This has the potential to reduce travel distance to PHCCs and more children are likely to attend routine childhood immunisation that is more likely to reduce outbreak of diseases, thereby, reduce the death of children below the age of five. Many studies have reported reduced childhood vaccination coverage with increased travel distance from homes [68-70]. Some other studies have reported strong association between routine childhood vaccinations and survival among infants in developing countries [71, 72] as these vaccinations guard against life-threatening diseases and reduce the spread of these diseases [73]. Potential beneficiaries are more likely to complete routine childhood immunization if the services are provided within a reasonable travel distance from their homes. The implication of lack of this is that more children remain unvaccinated and are exposed to risk of death from preventable diseases. Therefore, optimal location of PHCCs is of critical importance to provision of adequate care as it is more likely to increase patients' attendance to PHCCs and reduce outbreak of diseases, thereby reducing the death of children below the age of five [68, 74].

VI. CONCLUSION

This study demonstrates the usefulness of applying the binary dasymetric interpolation method to a region where detailed population data are not available, in order to redistribute summary population totals over small areas. The method divides the source zones into populated and unpopulated regions, and uses only the populated regions to calculate the population density and generate population surfaces. These spatially distributed estimates of the population were applied to a case study and used as demand to evaluate location decisions of 17 PHCCs in Port-Harcourt. PHCCs provide first level health care services critical to the survival of children below the age of five. A robust location-allocation method applied in this study minimises the total weighted distance travelled between PHCCs and demand points, thereby more likely to provide optimal accessibility to facility locations. This approach offers a useful improvement in net service accessibility, as it reduces the average travelled distance to its nearest health centre for the most people. These approaches can provide critical evidence to health planners to inform spatial decision-making.

A major limitation of this study is the lack of validation for the population estimates generated for Port-Harcourt. One approach to validate the population surfaces is to compare the results of a section of the study area with estimates obtained using manually digitised data for the selected area. The concept of dasymetric mapping was useful in Port-Harcourt since about one-third of the area within its boundary is covered by water. The method has the potential to allow public health agencies to target interventions by spatially modelling demand and to potentially increase the uptake of things like childhood immunisation to reduce outbreak of diseases and the death of children below the age of five. One of the Millennium Development Goals (MDGs) as described by [75] is to reduce "child mortality rate among children under 5 by two-thirds by the year 2015". They may have done this, they may not, but the methods presented in this work show how such goals could be more easily achieved.

There are a number of areas for future work arising from this research. Future research should consider conducting small area surveys to collect data from individual households or another data source at small area level, in order to validate the large scale demand surfaces. There are many mobile and web-based technologies that would enable that to happen remotely. Second, examine the dynamic pattern of demand rather than operate on the assumption of static demand. This is well known to have temporal variation at many different scales. This study has assumed demand only comprises residential-based night-time population as obtained from census estimates and not proximity to work places which may have affected the results of the analyses. This is because daily activities such as travel to work that characterise peoples' behaviour were not considered when modelling locations for services. It is expected that more detailed census data for Nigeria will soon become available as the office of the Surveyor General of Nigeria and the National Boundary Commission are jointly working on demarcating ward boundaries to allow creation of their digital boundaries. Future research could also employ proxy datasets such as mobile phone records from Nigerian Communications Commission (NCC) to represent the spatial distribution of the population. However, there are a number of concerns with these proxy datasets such as the different time of the day the data was recorded, the data may not represent all population as the children and elderly are not likely to use mobile phones, duplication of records as some residents may have more than one mobile phone etc. Footpath data would be obtained and incorporated into the distance matrix to evaluate access as the present study assume all roads are easily travelled with no traffic concentrations. This is because in reality some residents travel on foot from home to PHCCs, which most times travelling on a footpath is shorter than the road network. The study also considered a decrease in average travel distance, although travel time may be a better choice, and is more likely to improve overall accessibility to PHCCs.

ACKNOWLEDGEMENT

This work was undertaken as part of a PhD funded by the Petroleum Technology Development Fund (PTDF) under the Federal Government of Nigeria [PTDF/E/OSS/PHD/MIJ/316/10]. The authors would like to express gratitude to Geo-technics Services Limited for providing boundary and roads data for Port-Harcourt, Nigeria.

REFERENCES

[1] Cromley, R.G., J. Lin, and D.A. Merwin, Evaluating representation and scale error in the maximal covering location problem using GIS and intelligent areal interpolation. International Journal of Geographical Information Science, 2012. 26(3): p. 495-517.

[2] Tomintz, M.N., et al., Optimising the location of antenatal classes. Midwifery, 2013. 29(1): p. 33-43.

[3] Eicher, C.L. and C.A. Brewer, Dasymetric mapping and areal interpolation: Implementation and evaluation. Cartography and Geographic Information Science, 2001. 28(2): p. 125-138.

[4] Mennis, J., Generating surface models of population using dasymetric mapping∗. The Professional Geographer, 2003. 55(1): p. 31-42.

[5] Langford, M., Rapid facilitation of dasymetric-based population interpolation by means of raster pixel maps. Computers, Environment and Urban Systems, 2007. 31(1): p. 19-32.

[6] Langford, M., et al., Urban population distribution models and service accessibility estimation. Computers, Environment and Urban Systems, 2008. 32(1): p. 66-80.

[7] Cai, Q., et al., Estimating Small‐ Area Populations by Age and Sex Using Spatial Interpolation and Statistical Inference Methods. Transactions in GIS, 2006. 10(4): p. 577-598.

[8] Higgs, G. and M. Langford, GIScience, environmental justice, & estimating populations at risk: The case of landfills in Wales. Applied Geography, 2009. 29(1): p. 63-76.

[9] Maantay, J.A., A.R. Maroko, and H. Porter-Morgan, Research Note—A New Method for Mapping Population and Understanding the Spatial Dynamics of Disease in Urban Areas: Asthma in the Bronx, New York. Urban Geography, 2013. 29(7): p. 724-738.

[10] Comber, A., C. Proctor, and S. Anthony, The creation of a national agricultural land use dataset: combining pycnophylactic interpolation with dasymetric mapping techniques. Transactions in GIS, 2008. 12(6): p. 775-791.

[11] Poulsen, E. and L.W. Kennedy, Using dasymetric mapping for spatially aggregated crime data. Journal of Quantitative Criminology, 2004. 20(3): p. 243-262.

[12] Herrmann, C. and A. Maroko, Crime pattern analysis: exploring Bronx auto thefts using GIS. GIS for the Urban Environment. Redlands, CA: Environmental Systems Research Institute (ESRI), 2006: p. 407-413.

[13] Langford, M., An Evaluation of Small Area Population Estimation Techniques Using Open Access Ancillary Data. Geographical Analysis, 2013. 45(3): p. 324-344.

[14] [14]. Mennis, J. and T. Hultgren, Intelligent dasymetric mapping and its application to areal interpolation. Cartography and Geographic Information Science, 2006. 33(3): p. 179-194.

[15] Martin, D., N.J. Tate, and M. Langford, Refining population surface models: experiments with Northern Ireland census data. Transactions in GIS, 2000. 4(4): p. 343-360.

[16] Wright, J.K., A method of mapping densities of population: With Cape Cod as an example. Geographical Review, 1936. 26(1): p. 103-110.

[17] Petrov, A.N., Setting the Record Straight: On the Russian Origins of Dasymetric Mapping. Cartographica: The International Journal for Geographic Information and Geovisualization, 2008. 43(2): p. 133-136.

[18] Langford, M. and D.J. Unwin, Generating and mapping population density surfaces within a geographical information system. The Cartographic Journal, 1994. 31(1): p. 21-26.

[19] Langford, M. and G. Higgs, Measuring potential access to primary healthcare services: the influence of alternative spatial representations of population. The Professional Geographer, 2006. 58(3): p. 294-306.

[20] Xie, Y., The overlaid network algorithms for areal interpolation problem. Computers, environment and urban systems, 1995. 19(4): p. 287-306.

[21] Mrozinski, R.D. and R.G. Cromley, Singly- and Doubly- Constrained Methods of Areal Interpolation for Vector- based GIS. Transactions in GIS, 1999. 3(3): p. 285-301.

[22] Maantay, J.A., A.R. Maroko, and C. Herrmann, Mapping Population Distribution in the Urban Environment: The Cadastral-based Expert Dasymetric System (CEDS). Cartography and Geographic Information Science, 2007. 34(2): p. 77-102.

[23] Zandbergen, P.A. and D.A. Ignizio, Comparison of dasymetric mapping techniques for small-area population estimates. Cartography and Geographic Information Science, 2010. 37(3): p. 199-214.

[24] Tapp, A.F., Areal interpolation and dasymetric mapping methods using local ancillary data sources. Cartography and Geographic Information Science, 2010. 37(3): p. 215-228.

[25] Kim, H. and X. Yao, Pycnophylactic interpolation revisited: integration with the dasymetric-mapping method. International Journal of Remote Sensing, 2010. 31(21): p. 5657-5671.

[26] Cockings, S., P.F. Fisher, and M. Longford, Parameterization and visualization of the errors in areal interpolation. Geographical Analysis, 1997. 29(4): p. 314-328.

[27] Fisher, P.F. and M. Langford, Modelling the errors in areal interpolation between zonal systems by Monte Carlo simulation. Environment and Planning A, 1995. 27(2): p. 211-224.

[28] Langford, M., P. Fisher, and D. Troughear. Comparative accuracy measurements of the cross areal interpolation of population. in Proceedings of EGIS. 1993.

[29] Harris, R.J. and P.A. Longley, New data and approaches for urban analysis: modelling residential densities. Transactions in GIS, 2000. 4(3): p. 217-234.

[30] Green, N.E., Aerial photographic analysis of residential neighborhoods: An evaluation of data accuracy. Soc. F., 1956. 35: p. 142.

[31] Bentley, G.C., R.G. Cromley, and C. Atkinson-Palombo, The Network Interpolation of Population for Flow Modeling Using Dasymetric Mapping. Geographical Analysis, 2013. 45(3): p. 307-323.

[32] Goodchild, M.F., L. Anselin, and U. Deichmann, A framework for the areal interpolation of socioeconomic data. Environment and Planning A, 1993. 25(3): p. 383-397.

[33] Zola K. Moon, F.L.F., Population Density Surface: A New Approach to an Old Problem. Society & Natural Resources, 2001. 14(1): p. 39-51.

[34] Leyk, S., N.N. Nagle, and B.P. Buttenfield, Maximum Entropy Dasymetric Modeling for Demographic Small Area Estimation. Geographical Analysis, 2013. 45(3): p. 285-306.

[35] Harvey, J.T., 3opulation Estimation Models 3ased on Individual TM Pixels. Photogrammetric Engineering & Remote Sensing, 2002. 68(11): p. 1181-1192.

[36] Holt, J.B., C.P. Lo, and T.W. Hodler, Dasymetric Estimation of Population Density and Areal Interpolation of Census Data. Cartography and Geographic Information Science, 2004. 31(2): p. 103-121.

[37] Chen, K., An approach to linking remotely sensed data and areal census data. International Journal of Remote Sensing, 2002. 23(1): p. 37-48.

[38] Liu, X., K. Clarke, and M. Herold, Population density and image texture. Photogrammetric Engineering & Remote Sensing, 2006. 72(2): p. 187-196.

[39] Ward, D., S.R. Phinn, and A.T. Murray, Monitoring growth in rapidly urbanizing areas using remotely sensed data. The Professional Geographer, 2000. 52(3): p. 371-386.

[40] Pozzi, F., C. Small, and G. Yetman, Modeling the distribution of human population with nighttime satellite imagery and gridded population of the world. Earth Observation Magazine, 2003. 12(4): p. 24-30.

[41] Briggs, D.J., et al., Dasymetric modelling of small-area population distribution using land cover and light emissions data. Remote Sensing of Environment, 2007. 108(4): p. 451-466.

[42] Wu, S.-s., X. Qiu, and L. Wang, Population Estimation Methods in GIS and Remote Sensing: A Review. GIScience & Remote Sensing, 2005. 42(1): p. 80-96.

[43] Sridharan, H. and F. Qiu, A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes. Geographical Analysis, 2013. 45(3): p. 238-258.

[44] Curtis, S. and A. Taket, The development of geographical information systems for locality planning in health care. Area, 1989: p. 391-399.

[45] Mennis, J., Dasymetric mapping for estimating population in small areas. Geography Compass, 2009. 3(2): p. 727-745.

[46] Jega, I.M., Estimating population surfaces in areas where actual distributions are unknown: dasymetric mapping and pycnophylactic interpolation across different spatial scales. 2015, Department of Geography.

[47] Cromley, R.G., A.Y. Ebenstein, and D.M. Hanink, Estimating components of population change from census data for incongruent spatial/ temporal units and attributes. Journal of Spatial Science, 2009. 54(2): p. 89-99.

[48] Commission, N.P., Nigeria: National Population Census Priority tables, N.P. Commission, Editor. 2006: Nigeria.

[49] International, A., Forced eviction in Port-Harcourt, Nigeria, "Just move them". 2010, Amnesty International publication: London.

[50] Cromley, R.G., D.M. Hanink, and G.C. Bentley, A Quantile Regression Approach to Areal Interpolation. Annals of the

Association of American Geographers, 2012. 102(4): p. 763-777.
[51] Langford, M., D.J. Maguire, and D.J. Unwin, The areal interpolation problem: estimating population using remote sensing in a GIS framework. Handling geographical information: Methodology and potential applications, 1991: p. 55-77.
[52] Hakimi, S.L., Optimum locations of switching centers and the absolute centers and medians of a graph. Operations research, 1964. 12(3): p. 450-459.
[53] Teitz, M.B. and P. Bart, Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph. Operations Research, 1968. 16(5): p. 955-961.
[54] ReVelle, C.S. and R.W. Swain, Central facilities location. Geographical analysis, 1970. 2(1): p. 30-42.
[55] Church, R.L. and P. Sorensen, Integrating normative location models into GIS: problems and prospects with thep-median model. Spatial analysis: modelling in a GIS environment, 1996: p. 179-190.
[56] Narula, S.C., U.I. Ogbu, and H.M. Samuelsson, Technical Note—An Algorithm for thep-Median Problem. Operations Research, 1977. 25(4): p. 709-713.
[57] Maranzana, F., On the location of supply points to minimize transport costs. OR, 1964: p. 261-270.
[58] Holland, J.H., Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. 1975: U Michigan Press.
[59] Glover, F., Tabu search: A tutorial. Interfaces, 1990. 20(4): p. 74-94.
[60] Policy, C.f.H., The determination of need norms for health services, Part 2. A summary of norms in other countries, N.H.a.P. Development, Editor. 1993.
[61] Rispel, L., et al., A description and evaluation of primary care services delivered by the Alexandra Health Centre and University Clinic. 1995, Technical Report. Johannesburg: The Centre for Health Policy, Department of Community Health, University of the Witwatersrand.
[62] Doherty, J., L. Rispel, and N. Webb, Developing a plan for primary health care facilities in Soweto, South Africa. Part II: Applying locational criteria. Health policy and planning, 1996. 11(4): p. 394-405.
[63] Fisher, P.F. and M. Langford, Modeling Sensitivity to Accuracy in Classified Imagery: A Study of Areal Interpolation by Dasymetric Mapping∗. The Professional Geographer, 1996. 48(3): p. 299-309.
[64] Seidel, J.E., et al., Location of residence associated with the likelihood of patient visit to the preoperative assessment clinic. BMC Health Serv Res, 2006. 6: p. 13.
[65] Kohli, S., et al., Distance from the Primary Health Center: a GIS method to study geographical access to health care. Journal of Medical Systems, 1995. 19(6): p. 425-436.
[66] Payne, S., N. Jarrett, and D. Jeffs, The impact of travel on cancer patients' experiences of treatment: a literature review. European Journal of Cancer Care, 2000. 9(4): p. 197-203.
[67] Brewer, N., et al., Travel time and distance to health care only partially account for the ethnic inequalities in cervical cancer stage at diagnosis and mortality in New Zealand. Australian and New Zealand Journal of Public Health, 2012. 36(4): p. 335-342.
[68] Jamil, K., et al., The immunization programme in Bangladesh: impressive gains in coverage, but gaps remain. Health Policy and Planning, 1999. 14(1): p. 49-58.
[69] Acharya, L.B. and J. Cleland, Maternal and child health services in rural Nepal: does access or quality matter more? Health Policy and Planning, 2000. 15(2): p. 223-229.
[70] Tanser, F., B. Gijsbertsen, and K. Herbst, Modelling and understanding primary health care accessibility and utilization in rural South Africa: an exploration using a geographical information system. Soc Sci Med, 2006. 63(3): p. 691-705.
[71] Kristensen, I., et al., Routine vaccinations and child survival: follow up study in Guinea-Bissau, West AfricaCommentary: an unexpected finding that needs confirmation or rejection. Bmj, 2000. 321(7274): p. 1435.
[72] Arevshatian, L., et al., An evaluation of infant immunization in Africa: is a transformation in progress? Bulletin of the World Health Organization, 2007. 85(6): p. 449-457.
[73] Gu, W., X. Wang, and S.E. McGregor, Optimization of preventive health care facility locations. International Journal of Health Geographics, 2010. 9(1): p. 1.
[74] Hemat, S., et al., Health-care provision factors associated with child immunization coverage in a city centre and a rural area in Kabul, Afghanistan. Vaccine, 2009. 27(21): p. 2823-9.
[75] Programme, U.N.D., Millennium Development Goal, U.N.D. Programme, Editor. 2000.