



OPEN

DATA DESCRIPTOR

# Using open-source data to construct 20 metre resolution maps of children's travel time to the nearest health facility

Gary R. Watmough<sup>1,2,3</sup>✉, Magnus Hagdorn<sup>2</sup>, Jodie Brumhead<sup>1,2</sup>, Sohan Seth<sup>1,4</sup>, Enrique Delamónica<sup>1,6</sup>, Charlotte Haddon<sup>1,2</sup> & William C. Smith<sup>1,5</sup>

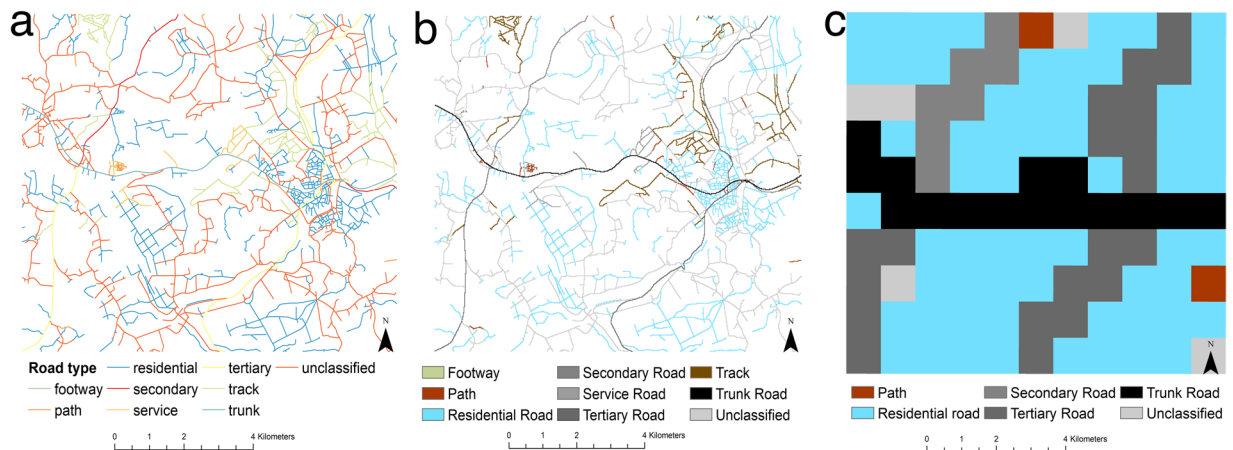
Physical access to health facilities is an important factor in determining treatment seeking behaviour and has implications for targets within the Sustainable Development Goals, including the right to health. The increased availability of high-resolution land cover and road data from satellite imagery offers opportunities for fine-grained estimations of physical access which can support delivery planning through the provision of more realistic estimates of travel times. The data presented here is of travel time to health facilities in Uganda, Zimbabwe, Tanzania, and Mozambique. Travel times have been calculated for different facility types in each country such as Dispensaries, Health Centres, Clinics and Hospitals. Cost allocation surfaces and travel times are provided for child walking speeds but can be altered easily to account for adult walking speeds and motorised transport. With a focus on Uganda, we describe the data and method and provide the travel maps, software and intermediate datasets for Uganda, Tanzania, Zimbabwe and Mozambique.

## Background & Summary

Access to services such as schools, health centres, markets and larger regional towns are important for individual's wellbeing, health and livelihoods in rural areas of developing countries. Increased travel times to key services can result in decreased utilisation<sup>1,2</sup> and in Uganda 44.0% of women in rural areas reported that distance to health centre contributed to deciding not to access healthcare for themselves<sup>3</sup>. Access to health centres is a key component of national preparedness for infectious disease planning<sup>4</sup> and a component of several Sustainable Development Goals (SDGs). Therefore, understanding access to key services at sub-national levels is required for improved delivery. Here we present new data of 20-m spatial resolution travel time estimates to health facilities for Uganda, Zimbabwe, Tanzania and Mozambique. We describe here the data characteristics and method for Uganda and provide access to the data for Uganda, Zimbabwe, Tanzania and Mozambique. It is our intention to provide the travel time maps for all sub-Saharan countries periodically through the Edinburgh data share portal <https://datashare.ed.ac.uk/handle/10283/3898>.

Utilisation of health facilities can be characterised by a complex mixture of travel time, financial costs involved in attending and travelling to a location, education of parents, quality of previous care and other political and societal barriers sometimes preventing access<sup>5,6</sup>. Here we are focused only on travel time as a physical measure of access. Travel time to key services such as schools, clinics and markets is seen as a key component of rural poverty and a major barrier preventing the realization of the right to health among children and, thus, impact child poverty (SDG 1.2.2). However, little empirical evidence currently exists on the levels to which time to travel to health centre contributes to health and poverty outcomes of populations. Therefore, these data provide an opportunity to establish high-resolution estimates of travel time to explore the impact that it has on socioeconomic outcomes such as poverty.

<sup>1</sup>Data for Children Collaborative with UNICEF, Edinburgh, United Kingdom. <sup>2</sup>School of GeoSciences, University of Edinburgh, Edinburgh, United Kingdom. <sup>3</sup>Global Academy of Agriculture and Food Systems, University of Edinburgh, Edinburgh, United Kingdom. <sup>4</sup>School of Informatics, University of Edinburgh, Edinburgh, United Kingdom. <sup>5</sup>Moray House School of Education and Sport, University of Edinburgh, Edinburgh, United Kingdom. <sup>6</sup>Division of Data, Analytics, Planning, and Monitoring, UNICEF, New York, United States. ✉e-mail: [gary.watmough@ed.ac.uk](mailto:gary.watmough@ed.ac.uk)



**Fig. 1** The roads polygon file (a) used in the analysis was a combination of Open Street Map and MapwithAI road datasets. The sinuosity of the roads was maintained when using a 20 m raster grid (b) compared to a 1 km raster grid (c) which is used in other studies.

Measuring distance to health services using straight line or Euclidian distance methods overestimate geographical access or underestimate travel time to services as it does not account for transport routes and barriers to travel such as waterbodies<sup>7–11</sup>. Least cost path estimates are often used to estimate travel time across surfaces<sup>12–14</sup> as they allow for the mode of transport to be considered<sup>4</sup>. Using a least cost path approach Weiss *et al.* 2020<sup>14</sup> estimated that approximately 43% of the World's population could not reach a health facility within 1 hour of walking. Travel time was based on data with a 1 km spatial resolution produced by Weiss *et al.*<sup>15</sup>. The size of cell used in cost surfaces impacts the results with coarser resolution cells leading to the sinuosity of roads being simplified, roads being overrepresented compared with other land cover and the over connection of roads across a landscape (Fig. 1). We developed an estimate of travel time to health centres following published methods<sup>15,16</sup> using 20 m spatial resolution data. We chose 20 m as it is the finest spatial resolution land cover data available globally at the time of analysis (from Sentinel-2 <http://2016africallandcover20m.esrin.esa.int/>) and thus is repeatable and transferable. The intention of this study and data is to provide a fine resolution estimation of travel time to service access points to be used by governments and humanitarian agencies such as UNICEF to explore development issues for example, exploring the links between access to health facilities and poverty or planning vaccination rollouts. To the authors' knowledge we have created the finest spatial resolution travel time data currently available which also uses a more detailed roads dataset from the Mapwith AI project (<https://github.com/facebookmicrosites/Open-Mapping-At-Facebook/wiki/Available-Countries>) not previously used in travel time map creation.

## Methods

The least cost path method used here can be broken down into two key phases: the first involves creating a 'cost' allocation surface which can also be referred to as an effort or friction surface<sup>14</sup>. This represents the effort to travel across a particular pixel and can thus consider the variation in effort (cost) required to travel across different land surfaces. The second phase uses this cost allocation (or effort surface) in a least cost path analysis to estimate travel time from every pixel to the nearest destination location (in this case health centres). This is done using Dijkstra's algorithm<sup>17</sup> to create Dijkstra trees which find the shortest path from one point to another<sup>12</sup>. It is implemented here using the graph module of scikit-image in python<sup>18</sup>.

**Creating cost allocation surfaces.** The cost allocation or effort surface was assembled using three primary input datasets on land surface characteristics that help or hinder travel speeds: land cover, roads and topography. Initial processing and analysis were undertaken on the individual datasets before being combined into a final cost allocation surface which depicts travel times from health centres across the country. The following sections detail the methodology and technical details of the cost surface creation as well as details on the choices made and research undertaken to feed into the travel time estimates.

**Transport mode choice.** There is little published information on how patients in Uganda access health services. However, in Kenya over 80% of patients were found to access health services by walking<sup>19</sup>. To measure travel-time we focused on pedestrian access because non-motorised transport is used for 50% of journeys in Uganda<sup>20</sup> and at least 50% of households in rural Uganda do not own any form of transport asset<sup>21</sup>. Further, children are most often the responsibility of women but most means of transport in Uganda are owned by men<sup>21,22</sup>. Some women do utilise bicycles, but issues of cost and gender relations make it easier for men to own and use bicycles and there are also negative cultural traditions which inhibit women's use of transport<sup>20</sup>.

Public transport mostly exists in Uganda through boda bodas (motorcycle taxis) or matutus (shared minibuses). It is likely that when accessing health services further away or when in need of particular medical

Sentinel-2 Land Cover Categories	CCI Land cover Code	Walking Speed (km/h)	Walking speed weighted for children (22% reduction)
Trees	1	1.50	1.17
Shrubs	2	1.50	1.17
Grassland	3	3.00	2.34
Cropland	4	3.00	2.34
Often Flooded	5	N/A	N/A
Sparse Vegetation	6	3.00	2.34
Bare Areas <sup>1</sup>	7	1.15	0.89
Built-Up	8	1.50	1.17
Open Water	10	1.00	0.78
No Data	200	N/A	N/A

**Table 1.** Sentinel-2 land cover categories and associated walking speeds assigned with and without children.

<sup>1</sup>Areas identified as 'Bare Areas' within the Sentinel-2 LULC 2016 for Uganda were primarily dry riverbed or sandbanks within rivers.

attention that motorised transport of some form would be used. However, we did not at the time of development have access to transport speed information. Thus, we only consider walking/pedestrian travel in this study.

**Child walking speed definitions.** Published walking speeds specific to road surface and land cover type for modelling access to services for Niger<sup>23</sup> and the Horn of Africa<sup>24</sup> were used as the baseline for speed estimations in Uganda. We used the maximum walking speed of 5 km/hr (1.39 m/s) that is commonly used to be representative of adult walking speeds<sup>19,23,24</sup>. Speeds ranged from 1.0 km/hr to 3.0 km/hr for non-road and path pixels such as grassland, cropland (Table 1) and from 3.5 km/hr to 5.0 km/hr for paths and road (Table 2). Speeds were fastest on main roads such as highways, primary roads and trunk roads. Pedestrians often walk along the side of these main roads (Fig. 2e) in many countries including Uganda, Mozambique and Zimbabwe as they are easier to traverse than unsealed roads, footpaths and other land cover types. Therefore, we preferentially pushed pedestrians to walk along these roads.

In the current work we have focused on child walking speeds since our work in Uganda seeks to understand the impact of access to health centres on multidimensional childhood poverty. Therefore, we reduced walking speeds to account for adults accompanied by children. Bouterse and Wall-Scheffler<sup>22</sup> observed that unloaded adults travelling alone had an average speed of 1.001 m/s which decreased to 0.773 m/s when they had to carry a child or to 0.785 m/s when older children were present that did not require carrying (they were not counted as a load but were accompanying the adult). Thus, the approximate decrease in travel speed was 22% and we reduced the speeds proposed by<sup>23</sup> by 22% in the child travel time maps. The speed reduction can be easily switched off or adapted in the software<sup>25</sup>.

**Data acquisition.** Data sources were selected based upon resolution and coverage with fine spatial resolution data preferred. As it was intended for the research to be applicable over multiple countries in which UNICEF is working as well as being user friendly for the wider scientific community, data sources were selected which had at least continental Africa coverage. The three data sets required for the travel-time maps are described below and in Supplementary Table A.

- Land cover – Land cover was derived from the Sentinel 2 CCI Land Use Land Cover (LULC) 2016 (20 m resolution) data product <http://2016africallandcover20m.esrin.esa.int/><sup>26–28</sup>, (Supplementary Table A). Other products with a greater number of land cover classes were available but only for a limited number of countries. The larger number of land cover classes was of little use as walking speeds were only available from the literature for a limited number of land cover types.
- Elevation – Shuttle radar topography mission (SRTM) Digital Elevation Model (DEM) no-void filled data at 30 metres resolution<sup>29</sup>, (Supplementary Table A) is available globally. Finer resolution datasets were available in some places, however, the use of the SRTM 30 m data allows for standardisation of the elevation data globally.
- Roads – OpenStreetMap (OSM) roads data <https://www.openstreetmap.org> (Supplementary Table A) is available for most countries globally<sup>30</sup>. Open street map is known to have inconsistent coverage due it being created using volunteered geographic information<sup>31</sup>. Thus, it was merged with the OSM mapwith.ai project (<https://github.com/facebookmicrosites/Open-Mapping-At-Facebook/wiki/Available-Countries>) data which used AI and high-resolution imagery to identify roads missing in OSM data to produce a more complete road dataset. The standard OSM road data and the mapwith.ai road data are available globally.

**Rasterizing roads.** The roads data from OpenStreetMap<sup>30</sup> is crowdsourced geographic data uploaded by volunteers<sup>31</sup>. This means completeness varies spatially both across and within countries, with Uganda being estimated to be 70% complete<sup>31</sup>. The *mapwith.ai* project by Facebook in conjunction with OpenStreetMap, have

OSM Roads Feature Class	Count (Total)	Notes (Ministry of Works and Transport - The Republic of Uganda, 2012; Ramm, 2019; OpenStreetMap, 2020)	Walking Speed (km/h)
motorway	105	A restricted access major divided highway, normally with 2 or more lanes plus emergency hard shoulder.	5.0
motorway_link	20	Connecting slip road/ramp	5.0
Trunk & Trunk_link	370 + 91	A main road with a motorway-like layout with multiple lanes which is restricted to motorised vehicles. Unlike motorways, trunk roads might have crossings or traffic lights. Their surface is always tarmacked.	5.0
Primary & primary_link	973 + 57	National roads connect the most important cities/towns in a country. These may be tarmacked and show centre markings.	5.0
Secondary & secondary_link	1991 + 47	Major transportation routes connecting cities and large towns. May be tarmacked but often not.	4.5
Tertiary & Tertiary_link	5460 + 62	These are busy through roads that link smaller towns and larger villages. Most often unpaved, but wide enough to allow two cars to pass safely.	4.5
residential	483527	Roads lined with housing in urban or village areas and where roads do not serve a through connection function.	4.0
unclassified	179029	Minor collector roads that link settlements. These roads are usually unpaved and are only wide enough for one vehicle. Primarily in rural areas and outside of inhabited places, though unclassified roads can be used to link suburbs in a city or town.	3.5
track	80217	This tag is usually used for roads providing access to agricultural or forestry facilities. In Africa, roads within National Parks mostly qualify as tracks too.	3.5
track_grade1-5	644	Solid, usually paved or sealed	3.5
service	14557	Driveways, entrances, private roads, service roads for industry etc	3.5
Bridleway & cycleway	38	Visual inspection shows no specific features to identify as bridleway	3.5
living_street	325	Streets where pedestrians have priority	4.0
pedestrian	121	Pedestrian only streets	4.0
path	153656	Paths are usually impassable for motorised vehicles but may be passable by motorcycles.	3.5
footway	8857	Pedestrian only, separated from parallel highway for vehicles. Can often be obstructed by traders and motor vehicles.	3.5
steps	176	Pedestrian only by nature	3.5
unknown	1507	Unknown type	3.5

**Table 2.** Open Street Map (OSM) road categories provided in the data downloads from OSM and Map with AI. Associated walking speeds assigned in the CPAS software are listed. These are based on previous speeds assigned in other studies. The speeds can be changed easily in the CPAS software to reflect local variations or the availability of motorised transport.

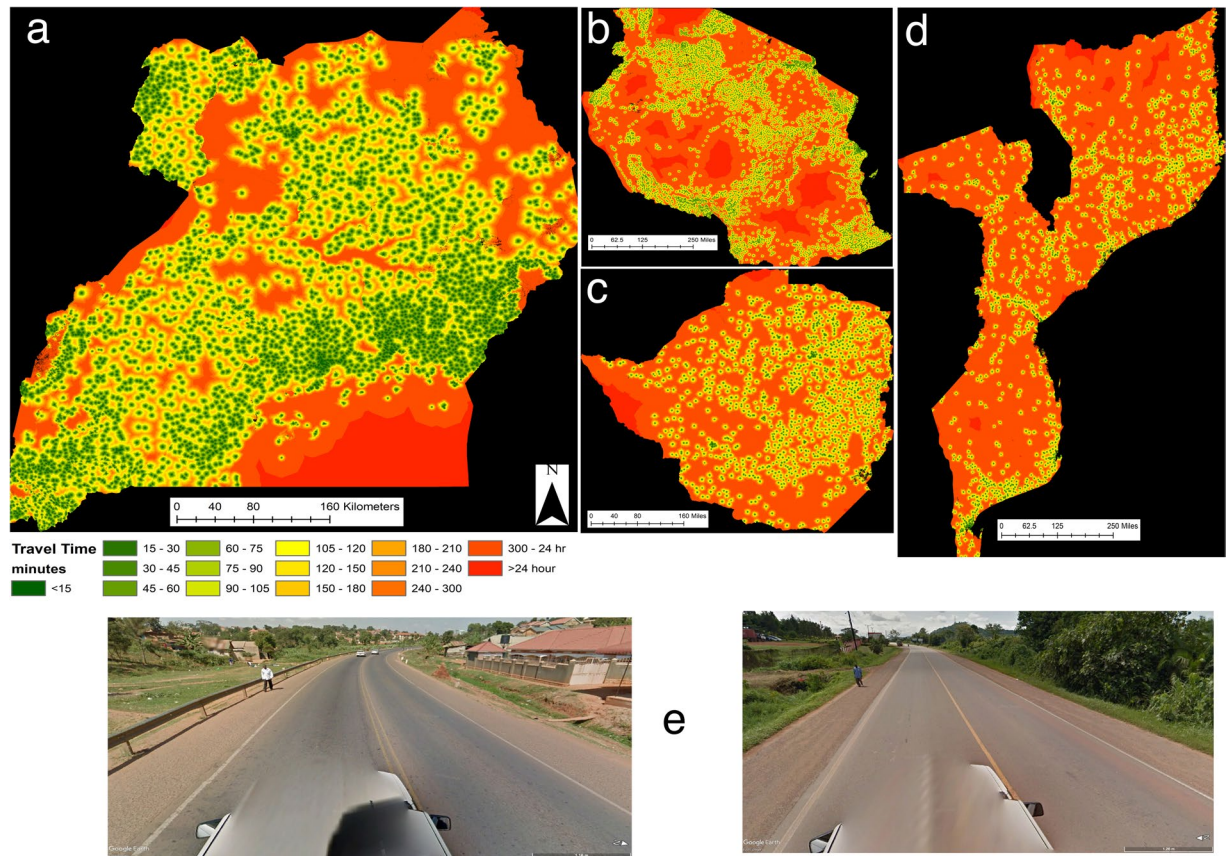
applied artificial intelligence alongside human validation to add missing roads and footpaths to the OSM dataset. The roads in both OSM and *Mapwith.AI* datasets were merged using ArcMap 10.7.1<sup>32</sup> 'merge' function in the ArcToolBox. The merged data required a new 'tag' attribute to be created which contained the road type and is the only editing a user is required to perform prior to running the code on a new location.

There are up to 29 types of road and path categories in the merged road data and walking speeds for each were derived from speeds used in existing studies (Table 2). Roads categorised as 'motorway' or 'trunk' are restricted to motorized vehicles according to OSM metadata. However, pedestrians were seen on the sides of some of these roads in Google Maps Street View (Fig. 2e). Therefore, walking speeds were allocated for all road types including motorways (Table 2). Roads are generally considered to be faster for travelling, but also most likely to be used by those walking as they allow easier navigation, therefore the walking speeds assigned to roads were faster than the walking speeds assigned to different landcover types.

The roads data were converted from a polyline shapefile (vector) to a raster surface so that they could be combined with the land cover and elevation data. Vector to raster conversion of roads can introduce issues related to the Modified Areal Unit Problem<sup>33</sup> as linear data is aggregated into the grid format of the raster surface<sup>34</sup>. When using a coarser cell size, it is more likely that multiple roads will be present inside the same cell. For example, when using a 1 km cell size Delamater *et al.*<sup>16</sup> found that road coverage was overestimated which would result in faster travel time estimations. The overestimated road coverage derives from simplifying the landcover within each 1 km pixel from multiple ground features (different types of road, paths and other landcovers) to a single feature. Often the fastest road type is prioritised when rasterizing multiple polygons into a single grid cell as was performed in Weiss *et al.*<sup>15</sup> and in our approach. This overestimation of road coverage will result in overestimation of access or underestimation of travel time to services. Further simplifications such as over connections of roads that result from smoothing the complex edges of polygons and reducing the sinuosity of the roads will also contribute to the overestimation of access or underestimation of travel times.

A 20 m cell size was used for the OSM roads in Uganda to match the CCI land cover data. While 20 m is still wider than many roads it is likely that the impact of time taken to reach the road from within a 20 m cell would be lower than larger cell sizes. Overlapping roads were prioritised by speed with the fastest roads given priority but this can be turned off in the software<sup>25</sup>. However, due to the fine resolution cell size of 20 m, the raster surface remained representative of the road network with little over connection between roads (Fig. 1) this merged data set is published through the Edinburgh Data Share portal for Uganda<sup>35</sup>, Mozambique<sup>36</sup>, Zimbabwe<sup>37</sup> and Tanzania<sup>38</sup>.





**Fig. 2** Travel time to any health centre in (a) Uganda, (b) Tanzania (c) Zimbabwe and (d) Mozambique calculated using the CPAS method and produced on a 20 m raster grid. These data are available for download (see Supplementary Table B). Google street view images showing pedestrians using the sides of major roads and highways in Uganda (e), included to demonstrate that a walking speed of 5 km/hr is appropriate for highways.

The cost surface is a scalar field but roads are directional. Rasterising a road implies that a cell can be crossed equally from left to right as bottom to top. At coarser resolution this may be a problem which could be dealt with by using an anisotropic cost surface. Using a 20 m resolution approach it is likely that anisotropic rasters are not required because the pixels adjacent to a road would have a higher cost and therefore the cost surface would preferentially push people in the correct direction.

**Converting land cover data to travel speeds.** The nine land cover categories available in the CCI (Trees; Shrubs; Grassland; Cropland; Often flooded; sparse vegetation; bare area; built-up; open water) were mapped to walking speeds from Blanford *et al.*<sup>23</sup> (Table 1). Open-water was initially coded as NA to create a barrier to travel as it would result in pixels with values of infinity in the least cost path allocation. A secondary cost surface model was created with a travel speed of 1.0 km per hour for the ‘Open Water’ category (the water speed can be changed in the configuration file of the software). This is because many of the island populations in Uganda are unlikely to be able to access services and, in particular healthcare services, without travelling over water. The reason for doing so is that when we use the cost allocation surfaces to estimate travel time we are able to prevent travel across water unless it is absolutely necessary. We define absolutely necessary as those living on islands where there are no health facilities present on the island and therefore, the inhabitants would have to use water to travel. This prevents travel times being calculated across water bodies where there may be no ferry service and forces the model to use roads. These speeds can be altered in the software easily and so other users can experiment with these.

**Weighting speeds using topography.** Methods of estimating the impact of gradient on walking speeds such as Tobler’s Formula or the Naismith-Langmuir rule are often used when modelling access to services<sup>39</sup>. Recent adaptations to these have been developed to account for slopes up to 30% or approximately 17°<sup>39</sup> and slopes up to 100% approximately 45°<sup>39</sup>. Therefore, to calculate the impact of gradient on walking speed we used Eq. 1 from Irmischer and Clarke<sup>40</sup>.

$$\text{Speed (m/s)} = 0.11 + e^{\frac{-(S+5)^2}{2 \times 30^2}} \quad (1)$$

where, S is the slope in percent.

The impact on walking speeds is calculated by computing the ratio of the walking speed for a slope and a flat surface. Slopes over 100% (over 45° gradient) were considered impassable. An assumption was made that those travelling to healthcare services were making a return trip and this was accounted for by taking a mean of the speed for a positive slope and negative slope of the same slope value<sup>39</sup>. The impact of negative and positive slopes is not equal (Eq. 2).

$$\text{Speed impact} = \frac{0.5 \times \left( \left( 0.11 + e^{\frac{-(+S+5)^2}{2 \times 30^2}} \right) + \left( 0.11 + e^{\frac{-(-S+5)^2}{2 \times 30^2}} \right) \right)}{0.11 + e^{\frac{-(0+5)^2}{2 \times 30^2}}} \quad (2)$$

**Assembling the cost allocation surface.** Two cost allocation surfaces are created using the above methods and are included in the data downloads for each country (Supplementary Table B). The first defines open water as a barrier to travel and so the speed allocated to this landcover is 'NA'. The second defines open water with an associated speed. To create a walking speed array, first the road walking speeds were used and then missing values were filled with landcover walking speed values. This walking speed array was multiplied by the slope impact grid. The speed for each cell was converted from kilometers per hour to meters per second. Finally, the time (in seconds) to walk across each cell was calculated using Eq. 3 where distance is equal to cell size (20 m). These travel time or friction surfaces as they are also known are published and made available through the Edinburgh Data Share Portal<sup>41–44</sup>.

$$\text{Travel time over cell} = \text{cell size} / (\text{speed} * \text{speed impact} * (1 - \text{child impact})) \quad (3)$$

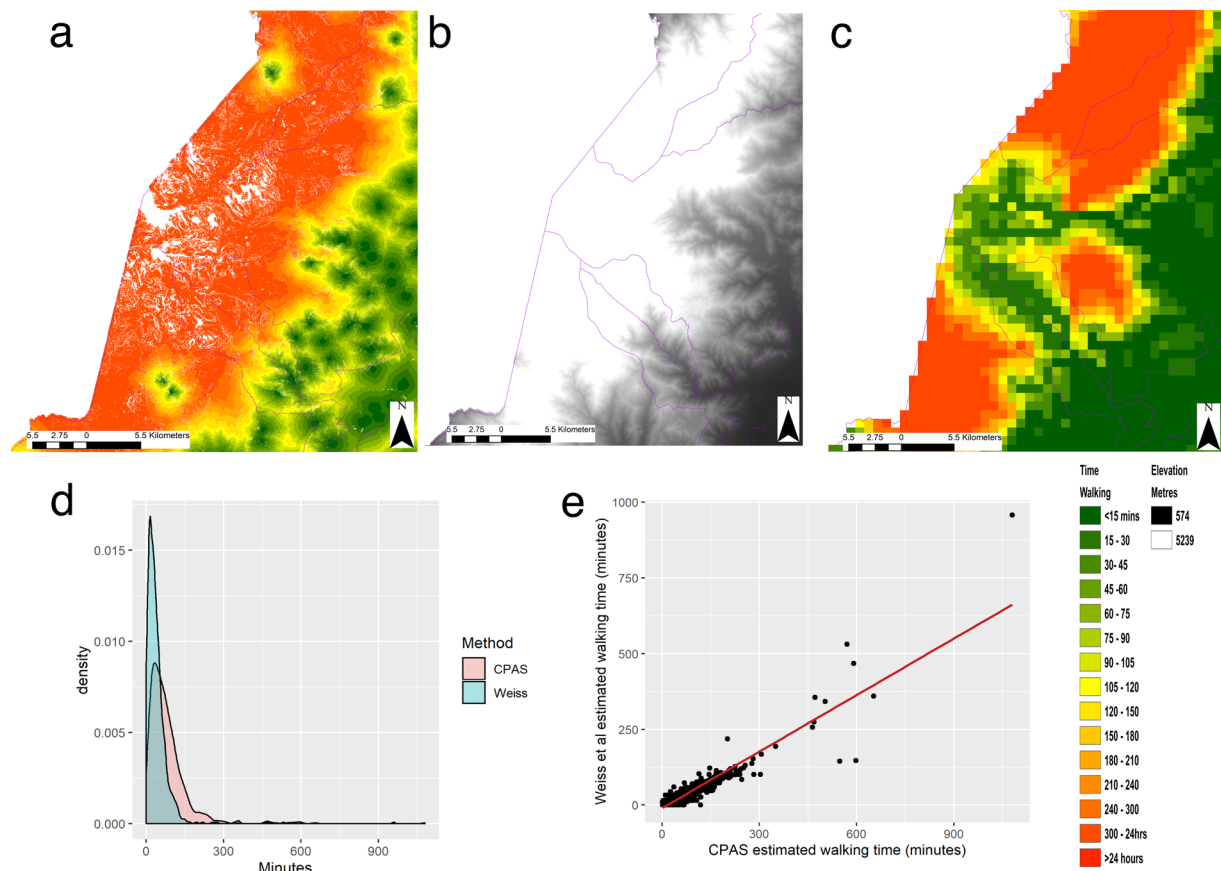
**Least cost path analysis for estimating travel time to health facilities.** A World Health Organization (WHO) validated spatial database of health facilities in sub-Saharan Africa<sup>45</sup> was used to indicate the location of health services in Uganda. The data were clipped to Uganda using GADM administrative boundary level-0 data. Health facilities with invalid locations such as those outside Uganda or located on water are identified and removed in the software<sup>25</sup> and a report is generated in csv format. Healthcare facilities which were located within cost surface cells that had been deemed impassable were identified and if adjacent cells were passable were randomly moved to one of these or if no adjacent cells were passable then the location was removed from the analysis, these facilities are also identified and included in a report produced by the software<sup>25</sup>.

The cost allocation surface and health facilities locations are combined using the least cost path method to calculate travel time from each cell in the array to the nearest health facility. The code first uses the cost allocation surface with water having an NA for travel time. Where travel time to healthcare services was measured as infinite, this indicated cells that had to traverse water in order to reach a health facility. In these cases, the travel time was measured using the water passable cost allocation surface in our example this meant the water had a travel speed of 1 km/hr applied (Table 1).

**Type of health facilities.** We provide several travel time maps for each country one of which is the travel time to the closest facility of any type (Figs. 2a, 3). Since this is a simplification of the reality of accessing health facilities, we split the health facility data by type and calculate travel times to each. For each country there are different facility types in the Maina *et al.*<sup>45</sup> data thus we describe below briefly which facility types have been used in Uganda, Tanzania, Zimbabwe and Mozambique.

The health facility types available in the Uganda data were: clinics, health centres (Level II, III and IV), hospitals and regional referral hospital. We used the Uganda Hospital and Health Centre Survey Census<sup>46</sup> to identify the services available at each facility type. Level II centres were excluded as they provide only basic medical care, whereas level III and IV facilities provide maternity care, basic lab services and inpatient care. Level IV facilities are the only facility type to have qualified doctors<sup>46</sup>. Hospitals represent access to surgical care and specialist health services. Therefore, in Uganda we have produced the following maps of travel time to: (1) any type of facility<sup>47</sup> (Fig. 2a); (2) level III health centre<sup>48</sup>; (3) level IV health centres<sup>49</sup>; and; (4) level IV health centres and hospitals<sup>50</sup>. Regional and national referral hospitals were excluded as to access these facilities a patient must be referred from another health facility (Level IV). Clinics were excluded because they are undefined in the census document, although we did not calculate travel time to Level II facilities these are the most frequent and therefore are covered in the map of travel time to any facility.

The health facility types available in the Zimbabwe data were; clinics, rural health clinics, rural hospital, district hospital, provincial hospital and central hospital. We used the USAID supported Zimbabwe Health System Assessment<sup>51</sup> to identify what type of services were available at each facility type. Rural health clinics and hospitals provide basic preventative medicine, maternity care and curative services. District hospitals provide primary health care services and can treat emergencies. Therefore, in Zimbabwe we have produced the following travel time maps to: (1) any health facility (Fig. 2c)<sup>52</sup> which includes all the health facilities available in Zimbabwe from the Maina *et al.* (2019) data; (2) rural health centres<sup>53</sup>; and (3) rural and district Hospitals<sup>54</sup>. Clinics were excluded from some of the maps as, despite several clinic types being mentioned in the literature, it was not possible to differentiate these types in the data. However, since they were the most common facility type in Zimbabwe they are included in the maps to any facility type.



**Fig. 3** (a) Travel time to any clinic in minutes calculated using the CPAS 20 m method. (b) Travel time to clinics and hospitals using the method from Weiss *et al.*<sup>16</sup> and a 1 km grid (c) Elevation data derived from the Shuttle radar topography mission (SRTM) showing mount Stanley. (d) Density plots showing the distribution of travel times for 682 Demographic and Health Survey Clusters across Uganda for both CPAS and Weiss *et al.* methods. (e) Scatterplot showing the monotonic relationship between the travel times estimated by the CPAS and Weiss *et al.* methods. Note in 3E the point at the top-left is not an outlier, it is a DHS cluster located on an Island with no health centre recorded in the WHO data. The white pixels in (a) indicate no data and result from CPAS removing any pixels with a slope  $>45$  degrees. Using a 1 km grid as in (c) results in aggregation of slope angles and therefore there are unlikely to be any pixels with slopes  $>45$  degrees. This results in the 1 km grid approach producing faster travel time estimates in the steepest cells which in reality are likely to be avoided by pedestrians<sup>65–69</sup>.

The health facility types in Tanzania were; dispensary, health centre, hospital, district hospital, designated district hospital, referral hospital and national hospital. We used the Health Sector Strategic Plan July 2015 – Jun 2020<sup>55</sup> to identify the type of services available at each facility type. Dispensaries provide the most basic of services and do not provide inpatient care. Health centres can admit patients and sometimes provide surgical services. Hospitals and district hospitals provide medical and basic surgical services. Referral, regional referral and national hospitals provide specialist medical care and medical training. Therefore, for Tanzania we have provided the following travel time maps to the following types of health facility: (1) dispensary (Fig. 2b)<sup>56</sup>; (2) health centres<sup>57</sup> and (3) hospitals and district hospitals<sup>58</sup>.

For Mozambique we provide a single travel time map to any health facility (Fig. 2d) listed in the Maina *et al.*<sup>45</sup> database<sup>59</sup>. At the current time it was not possible to identify the services offered at each type of clinic but these will be updated in the near future.

### Data Processing Software

- Python 3.
- Python - required packages: fuzzywuzzy; numpy; pandas; osgeo; geopandas; skimage; xarray; rioarray; Fiona; rasterio – these get installed automatically using the setup script or can be installed into a conda environment using the environment file.
- ArcMap 10.7.1



## Data Records

All data generated in this study are available at Edinburgh Data Share<sup>35–38,41–44,47–50,52–54,56–59</sup>. Supplementary Table B provides the descriptions of the data shared on Edinburgh Data Share and DOIs for each of the datasets as part of the submission. These data include the children's travel time to any health centre, the children's travel time to specific health centre types (3 additional maps in Uganda and 2 additional maps in Tanzania and Zimbabwe). We also provide the cost allocation/effort surfaces for children and the health facility locations (Supplementary Table B). The external data sources used for generating the cost surface are listed in Supplementary Table A.

## Technical Validation

To validate the travel times from the CPAS software we have compared the travel times from CPAS with those estimated by Weiss *et al.*<sup>14</sup> at 1 km spatial resolution which have been validated against GoogleMaps API data in the past<sup>60</sup>. We used 682 village clusters in the Ugandan 2016 Demographic and Health Survey (DHS)<sup>3</sup>. We overlaid village clusters onto both cost surfaces and used the 'extract to point' function in ArcMap 10.7.1<sup>32</sup>. We created summary statistics, density plots and ran a Spearman's rank correlation to compare the travel times estimated by the two approaches. A scatterplot (Fig. 3e) shows that the two travel time estimations are monotonically related which is further supported by the Spearman's rank correlation between the two walking speeds ( $\rho$  0.89 and  $p < 0.01$ ). The density plots (Fig. 3d) indicate that the CPAS method estimates a wider distribution of travel times than the Weiss *et al.* method. The walking times estimated in both the Weiss *et al.*<sup>14</sup> 1 km and the CPAS 20 m approaches are skewed towards lower values (Fig. 3d) which is expected as there will be more health centres located near to populated areas. Those clusters with higher estimated travel times are either located on Islands in Lake Victoria or in isolated mountain regions of Uganda.

Comparing the distribution of travel times and correlation indicates that there is strong agreement between the two methods as to which clusters have the greatest access to health facilities. The differences lie in the absolute values or the total minutes estimated for walking to the nearest health centre. Summary statistics show that the CPAS walking times have no zero estimated walking times (Table 3). This is because at 20 m resolution it is highly unlikely that a village/cluster will fall within the same grid cell as a health centre. In comparison, the Weiss *et al.*<sup>14</sup> estimations have 72 clusters estimated to have zero minutes travel time (Table 3). This is due to the aggregation of the landscape using a 1 km resolution grid and the smoothing involved in this approach, as zero minutes in Weiss *et al.*<sup>14</sup> could translate into a travel distance of 0–1000 m whereas in the CPAS approach a 0-minute travel time would only occur if the village cluster was 20 m or less from the health centre.

The median and mean estimated travel times (Table 4) are roughly twice as large in the CPAS travel times than the Weiss *et al.* travel times (Table 3). This further support conclusions from Delamater *et al.*<sup>16</sup> that a larger grid overestimates road coverage and therefore underestimates travel times. This is due to the simplification of road sinuosity in larger grids that is occurring in the Weiss *et al.*<sup>14</sup> 1 km estimations allowing roads and particularly the fastest roads to be overrepresented over-connected compared with other land cover types (Fig. 1b,c).

The overall differences between the two approaches are considerable when considering how many village clusters are reported as being within 30 minutes of a health centre. According to the Weiss *et al.*<sup>14</sup> method roughly half (344/682) of DHS clusters are within a 30-minute walk of the nearest health centre. But if using the CPAS method this drops to a quarter of DHS clusters (159/682). It is expected that, given the improvement in sinuosity of roads (Fig. 1b), the addition of roads and footpaths previously missed in OSM data and the finer resolution (Fig. 3), the CPAS method is more accurately characterising travel time by pedestrians. This could have large implications for public health planning in Uganda. If using the finer resolution approach in CPAS and targeting a 30-minute travel time then potentially the 1 km grids produced by Weiss *et al.*<sup>14</sup> are overestimating access or underestimating the time it takes for people to walk to a health centre. Further work is needed to fully validate these travel speeds using ground-based surveys which were not possible during this study due to travel restrictions imposed by the COVID-19 pandemic.

The validation detailed above did not make use of 'ground-truth' data. We examined the Uganda DHS<sup>3</sup> data as it sometimes contains information on travel times to health facilities. However, the DHS data for Uganda in 2016 did not contain this information. Conducting our own survey was not possible due to budget restrictions and the COVID-19 pandemic preventing UNICEFs ground team performing validations. Further, there is uncertainty in which health centre households are travelling to. Our model as well as almost all other similar approaches assume travel to the nearest facility. But we know from previous field work in other countries that households choose to go to different health facilities for a variety of reasons and often these are not the nearest facilities to the village but along a route that the individual takes when carrying out other activities for example. Thus, there is limited opportunity for validation at this time.

All of the input data that we have used has been independently validated. The walking speeds we have used have been published in peer-reviewed journal articles. Each of the component data sets have been validated or examined for completeness by others in the literature.

- The ESA CCI 2016 Land Cover data for Africa is classed as a prototype dataset and has been independently validated<sup>61</sup>. The validation found that the land cover accuracy was highly variable across Africa with an average overall accuracy of 65%. In Uganda the accuracy ranged between 46% and 65% due to misclassification errors between grassland and cropland. This overall accuracy is low, however, the highly detailed roads data available from OSM and mapwith.ai meant that the land cover map was less important for overall estimates of travel speeds. Furthermore, the land cover can be improved in the future or can be replaced in the code with a user defined Sentinel-2 land cover classification.



	Min	1 <sup>st</sup> quartile	Median	Mean	3 <sup>rd</sup> quartile	Max
1 km Weiss <i>et al.</i> <sup>16</sup>	0	15	30	40.85	50.50	958
20 m CPAS	1	33.03	62.76	81.55	102.15	1080

**Table 3.** Summary statistics for the walking only travel times from the 682 clusters to the nearest health centre calculated using the 1 km Weiss *et al.*<sup>16</sup> and 20 m CPAS methods. Calculated using 682 clusters from the 2016 Ugandan Demographic and Health Survey (DHS).

	Weiss <i>et al.</i> <sup>16</sup> (1 km)	CPAS (20 m)
0 minutes	72	0
< = 15 (including 0)	227	50
>0 & <= 15	155	50
< = 30	344	159
>0 & <= 30	189	109

**Table 4.** The number of village clusters that were within given walking times of the nearest health centres for the 1 km Weiss *et al.*<sup>15</sup> and the 20 m CPAS method. Calculated using 682 clusters from the 2016 Ugandan Demographic and Health Survey (DHS).

- The OSM data is continually validated during the digitisation process by OSM. Road completeness varies spatially both across and within countries, with Uganda being estimated to be 70% complete in 2017<sup>31</sup>. However, this completeness was increased by merging the roads with the mapwith.ai project footpaths and roads.
- The mapwith.ai data was manually validated prior to release. However, there were no published estimates of the completeness of roads and footpaths across Uganda at the time of this writing.
- The health facility location data were validated before being published<sup>45</sup>.

### Usage Notes

Currently, all the input datasets use geographic coordinates which get converted into metres using a constant factor so that a cartesian coordinate system can be used to compute the gradient. The code assumes a uniform factor to convert degrees latitude and degrees longitude into metres. This approximation works near the equator but may be problematic further away from the equator. However, the gradient is used to modify the cost which is an approximate model so it will have a small impact on the final cost allocation and travel time estimations. The code is open and licensed in such a way that users could adapt it to accept different projected coordinate systems should they see fit.

The cost allocation surfaces provided can be used in GIS analysis to estimate time to travel to user selected services. The user would need some form of location information for a particular service such as health centre, market, school, city or all-weather road for example. Thus, these surfaces can be used by others for further analysis when considering access to services across different countries.

No editing is required of the ESA CCI land cover data prior to running the code, it just needs to be downloaded and clipped to the country boundaries. Users need to ensure that the landcovercosts.csv file has the correct land cover class codes and required speeds specified. These land cover classes are the same as those in the ESA CCI land cover product so if a user opts to use a different land cover dataset the land cover class values will need to be changed.

The 20 m spatial resolution used in this analysis does mean that some files are very large, preventing the software being converted into an online web platform currently. High-performance computing is required for larger countries such as Tanzania and Mozambique. If users have a low bandwidth and cannot run the software they can make reasonable requests to the corresponding author to calculate travel times to particular facility locations of their choosing. The files are large due to being 20 m spatial resolution:

- Uganda – 6.7 gb
- Tanzania – 26 gb
- Zimbabwe – 11.6gb
- Mozambique – 37.9 gb

**System requirements.** The code works out of the box on an Ubuntu 20.04 system. For analysis using data for Uganda the code required 50GB of memory. For analysis in larger countries the code required 128GB and took approximately 3 hours to run for Zimbabwe, 4.5 hours for Tanzania and 6 hours for Mozambique.

The CPAS software works out of the box in that only standard Python packages and dependencies are required for it to run. These packages can be installed into either a conda or standard python virtual environment. We have run the software on Linux however, it is not restricted to this OS and we have tested in LTS ubuntu, but it will work on windows/mac/other linux with a suitable python installation. ArcMap was used for some of the preprocessing, but this could be done in QGIS or GRASS or R if a user does not have access to ESRC ArcMap Licenses. It is possible to distribute the CPAS software as a docker container with a user-friendly graphical frontend similar to AccessMod5. However, due to of the high-resolution of the datasets, significant computational resources are required to compute them.

## Code availability

The software/code works out of the box on a current Ubuntu system using Python, and all source data were derived from open and free datasets available for the whole of Sub-Saharan Africa. The Python software<sup>25</sup>, scripts and a virtual environment are provided with the data and requires minimal input from the user.

The software is controlled by a configuration file. The GitHub repository contains an example configuration file, `creation.cfg`. In this file the following should be changed:

- Path to input data sets
- Names of input data sets should these have been changed.
- Path to the location where the output files should be saved
- Names of the output files.

The user can also choose to change the following settings in the configuration file:

- Walking speed reduction to account for children. As standard it is set to 0.78 to generate a 22% reduction in the travel speeds. If a user wishes to generate a cost allocation surface for adult walking speeds they should change the factor to 1.0.
- Water speed can be changed in this file. It is set to NA in the `landcover.csv` and should remain as such. The water speed can also be changed in the configuration file and will appear in the water passable cost allocation surface.

Additional steps required prior to running the code:

- The roads from the standard OSM roads download and `mapwith.ai` download should be merged. We used the merge function in ArcMap 10.7.1<sup>32</sup>. The merged road data require a new 'tag' attribute to be created which contains the road type (name).
- The user should check the `roadcosts.csv` file to ensure the road names match those in the shapefile.
- The user can change the travel speeds assigned to each road type by editing the `roadcosts.csv` file.
- The user can change the travel speeds assigned to each land cover type by editing the `landcovercosts.csv` file. If users want to use a different land cover map they can do so by ensuring that the land cover types in the `landcovercosts.csv` match those in the chosen land cover data.

All of the analysis was conducted using Python-3 apart from merging the OSM roads with the MapwithAi roads. The code has been developed as an all-in-one software which can be downloaded from Zenodo<sup>25</sup> and GitHub [<https://github.com/ChildPovetyAccesstoServices/cpas/tree/v1.0>] and is licensed under the GNU General Public License v3.0 only, meaning changes to the code are permitted as long as they are distributed under the same license.

The code includes an example configuration file `creation.cfg`. This is the only file the user needs to alter to repeat the process. Within this file the user specifies the location of the input data and the names and paths for the output files. The user can also specify the weighting if they wish to consider children's travel speed, if not they can set the reduction factor to 0 which will generate adult cost allocation surfaces. The code runs in two steps (1) calculate the cost allocation surfaces (2) use the least cost path analysis to estimate the travel time from every pixel to the nearest health centre.

All rasters are read and written using the rasterio python package (<https://github.com/mapbox/rasterio>) together with xarray<sup>62</sup>. The steps and python packages used for pre-processing the roads data:

- Walking speed values are assigned to an array based on land cover type.
- A CSV file of OSM road categories and associated walking speeds was used to input the assigned walking speeds into a pandas Data Frame<sup>63</sup>.
- A shapefile containing roads was rasterised using the walking speeds assigned to each type of road. The fastest walking speeds are used for pixels that contain roads of different types. There were a few steps to this which are detailed below:
  - a. The road types of the CSV file were matched to the road types of the shapefile using a fuzzy match to ensure a walking speed was assigned to every road type in the shapefile.
  - b. The road types were grouped by walking speed. Each group of roads was processed separately.
  - c. Roads from the shapefile were filtered by road type and rasterised using the `rasterize` function of the `rasterio.features` module. The walking speed is used as value for the rasterised roads. The `all_touched` option was used so that all grid squares in which road segments existed were counted as road segments in the new raster. This was essential so the rasterization of the roads retained the connections.
  - d. The resulting walking speed surfaces were merged taking the maximum value at each pixel for all road types.

Pre-processing steps of the SRTM data required:

- Re-sampled to 20 m resolution, using bilinear interpolation Slope is calculated using the formula

$$slope = \sqrt{\left(\frac{dH}{dx}\right)^2 + \left(\frac{dH}{dy}\right)^2} * \frac{100}{111120}$$

The factor 100 converts the gradient to a percentage slope and the factor 111120 converts degree latitude/longitude to metres. This approximation is only valid close to the equator which is the case in this study.

As Eq. 1 does not account for gradients over 100%, these were removed and replaced with a NA value.

The cost allocation surface and health facilities locations were combined to calculate travel time from each cell in the array to the nearest health facility using the graph.MCP\_Geometric() and find\_costs() methods from the SciPy package<sup>64</sup>.

Received: 13 September 2021; Accepted: 25 March 2022;

Published online: 17 May 2022

## References

- Gabrysch, S. & Campbell M.R. Still too far to walk: Literature review of the determinants of delivery service use, *BMC Pregnancy and Childbirth* **9** (2009).
- Tansley, G. *et al.* The correlation between poverty and access to essential surgical care in Ghana: A geospatial analysis. *World Journal of Surgery* **41**, 639–643 (2017).
- Uganda Bureau of Statistics (UBOS) and ICF. *Uganda Demographic and Health Survey 2016*. Kampala, Uganda and Rockville, Maryland, USA: UBOS and ICF <https://dhsprogram.com/pubs/pdf/FR333/FR333.pdf> (2018).
- Hulland, E. N. *et al.* Travel time to health facilities in areas of outbreak potential: Maps for guiding local preparedness and response. *BMC Medicine* **17**, 1–16 (2019).
- Ahinkorah, B. O. *et al.* Examining barriers to healthcare access and utilization of antenatal care services: evidence from demographic health surveys in sub-Saharan Africa. *BMC Health Serv Res* **21**, 125 (2021).
- Wong, K. L. M., Benova, L. & Campbell, O. N. R. A look back on how far to walk: Systematic review and meta-analysis of physical access to skilled care for childbirth in Sub-Saharan Africa. *PLoS ONE* **12**, e0184432 (2017).
- Rutherford, M. E., Mulholland, K. & Hill, P. C. How access to health care relates to under-five mortality in sub-Saharan Africa: Systematic review. *Tropical Medicine and International Health* **15**, 508–519 (2010).
- Perry, B. & Gesler, W. Physical access to primary health care in Andean Bolivia. *Social Science and Medicine* **50**, 1177–1188 (2000).
- Costa, L. S. *et al.* Accessibility of selected hospitals and medical procedures by means of aerial and transit network-based measures. *Health Services Management Research* **16**, 136–140 (2003).
- Gething, P. W. *et al.* Empirical modelling of government health service use by children with fevers in Kenya. *Acta Tropica* **91**, 227–237 (2004).
- Ouma, P. O. *et al.* Access to emergency hospital care provided by the public sector in sub-Saharan Africa in 2015: a geocoded inventory and spatial analysis. *The Lancet Global Health*. **6**, e342–e350 (2018).
- Etherington, T. R. Least-Cost Modelling and Landscape Ecology: Concepts, Applications, and Opportunities. *Current Landscape Ecology Reports* **1**, 40–53 (2016).
- Gething, P. W. *et al.* Geographical access to care at birth in Ghana: A barrier to safe motherhood, *BMC Public Health*. **12** (2012).
- Weiss, D. J. *et al.* Global maps of travel time to healthcare facilities. *Nature Medicine*. **26**, 1835–1838 (2020).
- Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**(7688), 333–336 (2018).
- Delamater, P. L. *et al.* Measuring geographic access to health care: raster and network-based methods. *International Journal of Health Geographics* **11**, 1–18 (2012).
- Dijkstra, E. W. A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271 (1959).
- van der Walt, S. *et al.* scikit-image: Image processing in Python. *PeerJ* **2**, e453, <https://doi.org/10.7717/peerj.453> (2014).
- Noor, A. M. *et al.* Modelling distances travelled to government health services in Kenya. *Tropical Medicine and International Health* **11**, 188–196 (2006).
- Ministry of Works and Transport. *Draft Non-Motorised Transport Policy, Ministry of Works and Transport*, Republic of Uganda [https://wedocs.unep.org/bitstream/handle/20.500.11822/25414/uganda\\_nmtpolicy.pdf?sequence=1&isAllowed=y](https://wedocs.unep.org/bitstream/handle/20.500.11822/25414/uganda_nmtpolicy.pdf?sequence=1&isAllowed=y) (2012).
- Uganda Bureau of Statistics. *The National Population and Housing Census 2014 – Main Report*, Uganda Bureau of Statistics, Kampala, Uganda, <https://unstats.un.org/unsd/demographic/sources/census/wphc/UGA-2016-05-23.pdf> (2016).
- Bouterse, L. & Wall-Scheffler, C. Children are not like other loads: A cross-cultural perspective on the influence of burdens and companionship on human walking. *PeerJ* **9**, 1–18 (2018).
- Blanford, J. I. *et al.* It's a long, long walk: accessibility to hospitals, maternity and integrated health centers in Niger. *International Journal of Health Geographics* **11**, 1–15 (2012).
- Pozzi, F. and Robinson, T. Accessibility Mapping in the Horn of Africa: Applications for Livestock Policy, *IGAD LPI Working Paper*, 11–08, 42 <https://core.ac.uk/download/pdf/132642458.pdf> (2008).
- Hagdorn, M. SChildPovetyAccesstoServices/cpas: initial release Zenodo. <https://doi.org/10.5281/zenodo.4638563> (2021).
- Shuttle Radar Topography Mission (SRTM). Digital Elevation Model (DEM) for Zimbabwe – 30 m Mosaic RCMRD [http://geoportal.rcmrd.org/layers/servir%3Azimbabwe\\_srtm30metres](http://geoportal.rcmrd.org/layers/servir%3Azimbabwe_srtm30metres) (2021).
- Shuttle Radar Topography Mission (SRTM). Digital Elevation Model (DEM) for Mozambique – 30 m Mosaic RCMRD [http://geoportal.rcmrd.org/layers/servir%3Amozambique\\_srtm30meters](http://geoportal.rcmrd.org/layers/servir%3Amozambique_srtm30meters) (2021).
- ESA CCI Land Cover Mozambique. Sentinel-2 Prototype Landcover 20 m map of Africa 2016, *European Space Agency* <http://2016africallandcover20m.esrin.esa.int/> (2021).
- USGS. Digital Elevation Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global Non-Void Filled *United States Geological Survey* <https://doi.org/10.5066/F7K072R7> (2020).
- Ramm, F. OpenStreetMap Data in Layered GIS Format, *Geofabrik*; 1–20 <https://www.geofabrik.de/data/geofabrik-osm-gis-standard-0.7.pdf> (2019).
- Barrington-Leigh, C. & Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS ONE* **12**, 1–20 (2017).
- ESRI. ArcGIS Desktop: Release 10.7. *Environmental Systems Research Institute*. Redlands, CA (2019).
- Dong, D.W. Modifiable Areal Unit Problem (MAUP) *International Encyclopaedia of Human Geography*, 169–174 (2009).
- Mulrooney, T. *et al.* A comparison of raster-based travel time surfaces against vector-based network calculations as applied in the study of rural food deserts. *Applied Geography*. **78**, 12–21 (2017).
- Watmough, G. Uganda Road Data for Travel Time Maps 2016–2020. *Edinburgh DataShare* <https://doi.org/10.7488/ds/3074> (2021).
- Watmough, G. Mozambique Road Data for Travel Time Maps 2016–2020. *Edinburgh DataShare* <https://doi.org/10.7488/ds/3075> (2021).
- Watmough, G. Zimbabwe Road Data for Travel Time Maps 2016–2020 *Edinburgh DataShare*. <https://doi.org/10.7488/ds/3080> (2021).
- Watmough, G. Tanzania Road Data for Travel Time Maps 2016–2020. *Edinburgh DataShare* <https://doi.org/10.7488/ds/3077> (2021).
- Campbell, M. J. *et al.* Using crowdsourced fitness tracker data to model the relationship between slope and travel rates. *Applied Geography*. **106**, 93–107 (2019).
- Irmischer, I. J. & Clarke, K. C. Measuring and modelling the speed of human navigation. *Cartography and Geographic Information Science*. **45**, 177–186 (2018).
- Watmough, G. R., Hagdorn, M. & Brumhead, J. Uganda 20 m cost allocation surfaces for the travel time maps, 2016–2020. *Edinburgh Data Share* <https://doi.org/10.7488/ds/3073> (2021).
- Watmough, G. R., Hagdorn, M. & Brumhead, J. Tanzania friction surface 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3089> (2021).

43. Watmough, G. R., Hagdorn, M. & Brumhead, J. Mozambique friction surface 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3086> (2021).
44. Watmough, G. R., Hagdorn, M. & Brumhead, J. Zimbabwe friction surface 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3088> (2021).
45. Maina, J. *et al.* A spatial database of health facilities managed by the public health sector in sub Saharan Africa. *Scientific data* **6**, 134 (2019).
46. Uganda Ministry of Health *Uganda Hospital and Health Centre IV Survey Census*, World Health Organization, African Development Bank and Republic of Uganda Ministry of Health, p224 [https://www.who.int/healthinfo/systems/SARA\\_H\\_UGA\\_Results\\_2014.pdf](https://www.who.int/healthinfo/systems/SARA_H_UGA_Results_2014.pdf) (2014).
47. Watmough, G. R., Hagdorn, M. & Brumhead, J. Uganda Travel Time to nearest Level III health centre Map Output 2016–2020. *Edinburgh Data Share* <https://doi.org/10.7488/ds/3085> (2021).
48. Watmough, G. R., Hagdorn, M. & Brumhead, J. Uganda Travel Time to nearest Level IV health centre Map Output 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3078> (2021).
49. Watmough, G. R., Hagdorn, M., Brumhead, J. Uganda Travel Time to nearest Level IV health centre & Hospital Map Output 2016–2020 <https://doi.org/10.7488/ds/3082> (2021h)
50. Watmough, G. R., Hagdorn, M. & Brumhead, J. Uganda Travel Time to any health facility Map Output. 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3084> (2021).
51. Osika, J., *et al.* *Zimbabwe Health System Assessment* USAID, Bethesda, MD: Health Systems 20/20 Project, Abt Associates Inc. [https://www.hfgproject.org/wp-content/uploads/2015/02/Zimbabwe\\_Health\\_System\\_Assessment20101.pdf](https://www.hfgproject.org/wp-content/uploads/2015/02/Zimbabwe_Health_System_Assessment20101.pdf) (2010).
52. Watmough, G. R., Hagdorn, M. & Brumhead, J. Zimbabwe Travel Time to any health facility Map Output, 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3083> (2021).
53. Watmough, G. R., Hagdorn, M. & Brumhead, J. Zimbabwe Travel Time to the nearest health clinic Map Output, 2016–2020. *Edinburgh Data Share* <https://doi.org/10.7488/ds/3079> (2021).
54. Watmough, G. R., Hagdorn, M. & Brumhead, J. Zimbabwe Travel Time to nearest hospital Map Output 2016–2020. *Edinburgh Data Share* <https://doi.org/10.7488/ds/3081> (2021).
55. Tanzania Ministry of Health *Health Sector Strategic Plan HSSV IV: Reaching all Households with Quality care*, Tanzania Ministry of Health [http://www.tz.dpg.or.tz/fileadmin/documents/dpg\\_internal/dpg\\_working\\_groups\\_clusters/cluster\\_2/health/Key\\_Sector\\_Documents/Induction\\_Pack/Final\\_HSSP\\_IV\\_Vs1.0\\_260815.pdf](http://www.tz.dpg.or.tz/fileadmin/documents/dpg_internal/dpg_working_groups_clusters/cluster_2/health/Key_Sector_Documents/Induction_Pack/Final_HSSP_IV_Vs1.0_260815.pdf) (2015).
56. Watmough, G. R., Hagdorn, M. & Brumhead, J. Tanzania travel time to nearest dispensary 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3091> (2021).
57. Watmough, G. R., Hagdorn, M. & Brumhead, J. Tanzania travel time to nearest health centre 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3092> (2021).
58. Watmough, G. R., Hagdorn, M. & Brumhead, J. Tanzania travel time to nearest hospital 2016–2020 [dataset]. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3093> (2021).
59. Watmough, G. R., Hagdorn, M. & Brumhead, J. Mozambique travel time to any health facility 2016–2020. *Edinburgh Data Share*. <https://doi.org/10.7488/ds/3087> (2021).
60. Nelson, A. *et al.* A suit of global accessibility indicators. *Scientific Data* **6**, 266 (2019).
61. Lesiv, M. *et al.* Evaluation of ESA CCI prototype land cover map at 20 m, IIASA Working Paper IIASA, Laxenburg, Austria, WP-17-021 <http://pure.iiasa.ac.at/id/eprint/14979/> (2017).
62. Hoyer, S. & Hamman, J. Xarray: N-D labelled Arrays and Datasets in Python. *Journal of Open Research Software*. **5**, 1 (2017).
63. McKinney, W. Data structures for Statistical Computing in Python, *Proceedings of the 9th Python in Science Conference*, 51–56 <http://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf> (2010).
64. SciPy 1.0 Contributors. 'SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python', *Nature Methods*, **17**, 261–272 (2020).
65. ESA CCI Land Cover Uganda. Sentinel-2 Prototype Landcover 20 m map of Africa 2016, *RCMRD Geoportal* [http://geoportal.rcmr.org/layers/servir%3Auganda\\_sentinel2\\_julc2016](http://geoportal.rcmr.org/layers/servir%3Auganda_sentinel2_julc2016) (2021).
66. Shuttle Radar Topography Mission (SRTM). Digital Elevation Model (DEM) for Uganda – 30 m Mosaic. *RCMRD Geoportal* [http://geoportal.rcmr.org/layers/servir%3Auganda\\_srtm30meters](http://geoportal.rcmr.org/layers/servir%3Auganda_srtm30meters) (2021).
67. ESA CCI Land Cover Tanzania. Sentinel-2 Prototype Landcover 20 m map of Africa 2016, *RCMRD geoportal* [http://geoportal.rcmr.org/layers/servir%3Atanzania\\_sentinel2\\_julc2016](http://geoportal.rcmr.org/layers/servir%3Atanzania_sentinel2_julc2016) (2021).
68. Shuttle Radar Topography Mission (SRTM). Digital Elevation Model (DEM) for Tanzania – 30 m Mosaic *RCMRD Geoportal* [http://geoportal.rcmr.org/layers/servir%3Atanzania\\_srtm30meters](http://geoportal.rcmr.org/layers/servir%3Atanzania_srtm30meters) (2021).
69. ESA CCI Land Cover Zimbabwe. Sentinel-2 Prototype Landcover 20 m map of Africa 2016, *RCMRD geoportal* [http://geoportal.rcmr.org/layers/servir%3Azimbabwe\\_sentinel2\\_julc2016](http://geoportal.rcmr.org/layers/servir%3Azimbabwe_sentinel2_julc2016) (2021).

## Acknowledgements

We thank the constructive comments of the anonymous reviewers and editor for helping to improve this manuscript. This work was funded by the Data for Children Collaborative with UNICEF under the Child Poverty Access to Services project (<https://www.dataforchildrencollaborative.com/>) a collaboration between the Scottish Government, The DataLab, UNICEF and the University of Edinburgh. This work came about as part of the Data for Children Collaborative (DfCC) with UNICEF. DfCC uses a co-constructive process to bring together expertise across sectors to address questions targeted at improving the lives of children across the globe using innovative data science techniques. The research was compliant with ethics approvals in the school of geosciences at the University of Edinburgh and the DataLab's ethical assessment. We thank Alex Hutchison and Alex Fassio for support in designing the project, on-going project management and Ethical advice. We thank Charlotte Marcinko & Peter Hargreaves for testing the code and providing feedback on the user instructions and Bob Sanders and team for uploading and checking the data files in Edinburgh DataShare facility. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (<http://www.ecdf.ed.ac.uk/>).

## Author contributions

G.R.W. undertook data analysis and validation, tested the code and transferred the code to other countries, drafted the manuscript, performed data visualisation, managed the work and conceived the study with contributions from E.D., J.B., S.S. and W.C.S. M.H. contributed python code, developed the software. J.B. developed an earlier version of the code and validation and processed the data. C.H. contributed to the data analysis, performed data visualisation and edited the manuscript. S.S. undertook data analysis, tested the software and edited the manuscript. E.D. edited the manuscript. W.C.S. edited the manuscript. All authors read and approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.



### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-022-01274-w>.

**Correspondence** and requests for materials should be addressed to G.R.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022