MEASURING MOBILITY IN AFGHANISTAN USING TRAVEL TIME MODELS
METHODOLOGY NOTE[1]

Walker Bradley and Brian Blankespoor

Measuring access

This note describes the methodology to measure within country mobility with an illustrative example applied to the context of Afghanistan.

The role of transportation in connecting people to markets and services is important in economic development (Yoshida and Deichmann 2009). Using various measures of access, previous studies examined the local impact of transport improvements on urban growth (e.g. Baum-Snow et al. 2015; Jedwab and Storeygard 2017), agriculture (Donaldson and Hornbeck 2016), local GDP (van der Weide et al. 2018) and employment specialization (Blankespoor et al. 2018) (also see Redding and Turner, 2015, and Berg el al. 2017 for surveys).

Traditionally, Ministries and Development organizations measure the number of people that have access to services and markets within a linear distance threshold, e.g. how many residents are within 2 kilometers of a school or an all-weather road?[2] However, considerable variation of travel time at the local level exists where traditional linear measurements do not include local geographic detail such as rivers and topography. For example, in the context of rural Afghanistan, a distance of 200 meters could include a river that separates someone from a school. Thus, this person would need to travel to the nearest bridge, and then from the bridge back to the school.

Recent improvements in the detail/resolution of satellite imagery and in computational methods allow us to address limitations in distance-based measures of access. Building on previous literature (e.g. Uchida and Nelson 2008; Berg et al. 2017; Weiss et al. 2018), we construct a travel time model with a gridded data structure to measure access from origins(s) to destination(s), which requires a mapping in units of time. One advantage of this method is the complete spatial coverage of all land area compared to a network method.[3] We apply the model to Afghanistan and provide two main analytic efforts. First,

---

[2] For example, the Rural Access Index (Roberts et al. 2006) measures the rural population who live within two kilometers (i.e. approximately a walk of 20-25 minutes) of an all-season road as a proportion of the total rural population. Iimi et al. (2016) present an updated method to measure rural access using spatial data and methods for eight countries.

[3] An alternative would be to use a network analysis using vector data. Vector data relies on mathematically defined points (often relative to the center of the Earth) and relating the points linearly in the case of roads. In the case of polygons, the features are also defined linearly, like roads or rivers, but the start point is also the endpoint. We use raster modeling because: 1. Many people travel cross country, i.e. off-road, which would be largely missed using network analyses; 2. The available road network data for Afghanistan is incomplete, and topologically incorrect. Raster-based models allow the authors to compensate for missing data more effectively than if we used vector data. A literature review indicates that raster-based travel calculations tend to estimate travel times about 5% higher than vector-based models, so we factor that into our model.

we use detailed data sources to account for the local geographic context at 10-meter resolution taking advantage of recent computational efficiencies. This is used to model access to various services. Second, we can use modifications to the roads data to model the potential effects of various infrastructure improvements on access, including the placement of bridges and changes to road class.

The note is structured as follows: data, model, an illustrative application, conclusion, and discussion of limitations.

Data

We constructed the model using gridded, i.e. raster data. This data structure provides complete spatial coverage in a matrix of a fixed size and a format that facilitates processing and integration across data sources by aligning the pixel.

In the context of this model, we define all the measurable aspects of a natural land surface to combine this information into an estimated travel time across the pixel: slope; land cover type, e.g. grasslands, swamp, and irrigated farmland; perennial water sources, e.g. rivers, canals, and lakes. We use 25 different classes of land cover data, derived from satellite imagery from 2010 (FAO's *The Islamic Republic of Afghanistan Land Cover Atlas* 2010). Next, we use ESRI Arc Pro software to convert the vector data into raster and assign an estimated travel speed for each class (Table available here). The spatial unit of analysis for the model is 10-meter resolution of the land cover dataset. Following the conversion of speed to time associated with the land cover, we modify the time value based on the slope, which is derived from 30-meter elevation data, using an algorithm created by the United States Department of the Interior based on hiking speeds in a variety of terrains (Van Wagtendock and Benedict 1980). For the foot-mobile model, we add in roads, and conduct a final elevation modification to simulate the slowing down of walking at higher elevations. For mixed mobility, we modify by elevation, and then add the roads using the assumption that when one is traveling on roads they are in a vehicle and are thus not influenced by elevation.

Travel time model

In a more detailed sense the purpose of this model is to convert aspects of these basic layers to units of time. That is to say, how long does it take to cross a given pixel?

We use a time-cost raster, which is a composite of multiple factors estimated in time across a pixel. Then, we use ArcGIS software to solve least-cost paths between origin(s) and destination(s). This model relies on a mix of algorithms and binary conditional statements, the former modifies existing values, and the later takes the higher value of the two estimated travel times. For example, we start with a user-defined speed, which is first modified by the challenges, or lack thereof, associated with a land cover type, then we subsequently modify the resulting speed with slope, to ensure any significant slope slows down the traveler and increases the travel time. This allows us to more accurately capture conflicts between a permissive land cover type, and a restrictive slope, and vice versa. In the case of the conditional statements, we use values associate with roads in the place of the underlying terrain, because the road has modified the terrain, making it easier to traverse.

The model uses an equation describe the effects of slope on travel in a continuous fashion, rather than discrete blocks. Of the available methods, the authors chose van Wagtendonk and Benedict (1980)[4]:

(1) $V = V_0 e^{-ks}$

> Where:
> v = off road foot based velocity over the sloping terrain,
> v_0 = the base speed of travel over flat terrain, 5 km/hr as a maximum.
> s = slope in gradient (meters per meter) and,
> k = a factor which defines the effect of slope on travel speed
> For this general case we assume a constant speed for each land cover type and k = 3.0 and constant for uphill and downhill travel.

The unit of the slope raster is in gradient ΔY/ΔX, which is expressed as a percent instead of degrees[5]. Using the raster calculator function within ArcGIS we convert slope to velocity, using the formula above, assuming 5 km/hr as our initial input of an average walking speed. Subsequently, we convert this velocity to a slope-penalty raster by dividing the same initial speed, 5km/hr, by the slope-restricted speed. This gives us the penalty, on a per-pixel basis, that the slope imparts on the fixed speeds from each of the land cover data classes.

Equation 2 details how the slope modifies the starting land cover time values, which can be found in Annex 1.

$$(2)\ Time\ per\ meter = LCT * (\frac{5km}{hr} / (\frac{5km}{hr} * (e^{-3\left(\frac{S}{100}\right)})))$$

> Where S is the slope of the pixel.[6]

> LCT = is the time, per meter of travel, associated with a given pixel of a land cover type

Continuing our terrain constructions, we construct our perennial water category derived from source data including: lakes, rivers and canals (perennial extracted from dry channels), and similar to the land cover data, we add a new field, calculate the field with the value of seven minutes per 10-meter pixel, and convert them to raster as well. While the seven minutes per pixel value is arbitrary, it shows that, while not impossible to cross, algorithmically, bridges, etc., gain greater relative time to cross the pixel, and areas without them are penalized. Following this step, we use a conditional phrase giving preference to higher time-cost values of the water features over the land cover because the water layers, in effect, cut across the terrain preventing direct access from one point to another. We give the

---

[4] Another common function to measure walking speed is from Tobler (1993), however van Wagtendonk and Benedict developed the model in a similar terrain to Afghanistan. The method is flexible to adopt either method.
[5] For example, on a 45° slope one would travel 100 vertical meters in 100 horizontal meters, thus a 45 °angle is equal to a 100% gradient slope.
[6] The slope is divided by 100 because ArcGIS Slope calculations display percent and not decimal, i.e. 12.3 instead of .123

water layers preference to the land cover, because many ditches and canals were too small to appear in the land cover dataset and yet are significant barriers to travel.

The final layer, involving the land cover and water barriers, makes up what we refer to as the Terrain layer. This terrain layer is not removed like a normal intermediate output, because we wish to retain the ability to add roads data without rebuilding the entire effort. In support of the World Bank's Open Data initiative we use OpenStreetMap as the base for the roads data, and these layers are updated and modified regularly, thus, for ease and shorter computation times, we retain the Terrain layer for the "Roads Update" model.

The roads take precedence over the previous results because they are modifications to the environment, creating an active layer instead of merely receiving the values of the environment/terrain. This layer, like the preceding layers, needs to be converted from vector to raster with different time-costs associated with each class type. Like the previous layers, the roads, too, will have a field added and times for each class of road calculated, which will serve as the basis for the raster conversion. This layer is unique, because it allows the user to modify any road value to simulate changes to the end time-cost. As a result of this chosen order, we gain a forecasting aspect of the model, rather than merely a now-casting, or back-casting approach. Such forecasting enables this model to simulate improvements to various road segments and the subsequent effect on all localized travel, (see graphic 1).
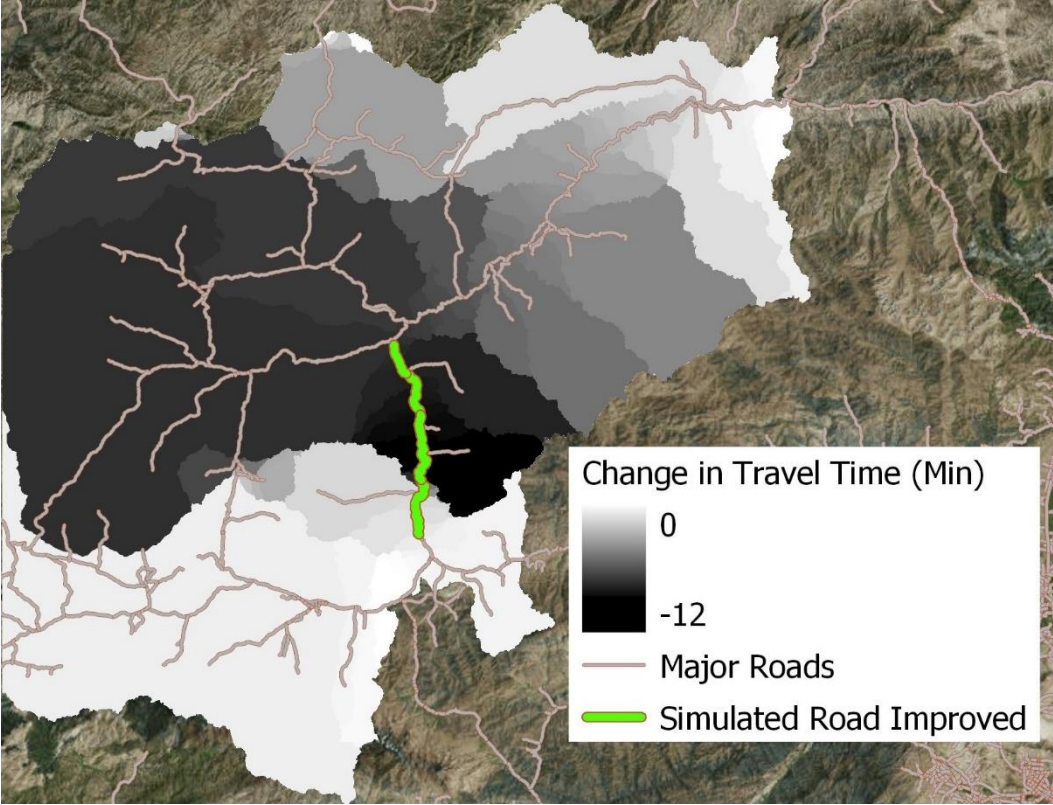
Graphic 1: Simulated Road Improvement

Finally, raw elevation will impact foot travel times, so there will be a multiplier for each cell in the raster above 2000 meters in elevation, simulating the "slowing" down of foot travel using the following equation from (Berg et al. mimeo).[7]

$$(3)\ If\ Elevation \leq 2000;\ TimeCost * 1$$

$$If\ 2000 < Elevation \leq 3000;\ TimeCost * (1 + (\frac{Elevation - 2000}{5000}))$$

$$If\ 3000 < Elevation;\ TimeCost * (0.15 * e^{(.0007*Elevation)})$$

Roads update model:

The largest part of the processing is taken up with the land cover, water layers, and slope. The current model structure relies on OpenStreetMap (OSM) for roads layers, which enables this model to improve as the road network data in OSM improves. This is one of the strongest aspects of this model. Any time the raster surface needs to be updated, the user simply runs the Roads Update tools, which calculates substantially faster.

Travel time model:

We created two versions of the time-cost model and apply it to Afghanistan: a travel time model by foot (including on roads) and a travel time model with mixed-modes of foot and vehicle. The result of the mixed-modes model assumes that each pixel uses the maximum possible speed allowed by the road. Thus, the specified model does not include measures of congestion or the cost of travel between modes.[8] Since travel on roads is assumed to be by car, the slope and elevation modifiers apply only to the general terrain model *before* the inclusion of roads. The foot model assumes that even on roads, the traveler is walking, and thus *is* affected by elevation. Thus, in the foot model, the elevation modifier is applied *after* the inclusion of roads. The final time-cost raster dataset is created for the whole country.[9]

Results

Following the cost-distance process, the output is reclassified to a binary raster, with areas either in the catchment area, or outside it. This is then converted to a polygon. Using the "Summarize Within" function in ArcGIS Professional we count the population within the catchment area from gridded population datasets[10].

---

[7] ESRI Raster Calculator syntax: Con("%Elevation Raster%" <= 2000, "%TimeCostRaster_ft_pre%"*1 , Con(("%Elevation Raster%">2000) & ("%Elevation Raster%"<=3000), ("%TimeCostRaster_ft_pre%"*(1+(("%Elevation Raster%"-2000)/5000))), Con("%Elevation Raster%">3000,"%TimeCostRaster_ft_pre%"*(0.15*Exp(.0007*"%Elevation Raster%")))))

[8] The modeling framework is flexible to include a cost of traffic (e.g. all urban roads have a sub-optimal speed at a percent of the maximum design speed).

[9] However, this input is specific to a set of origins and/or destinations. We suggest to only include areas of the country that are needed for the area of interest in order to minimize computational time. For example, a map with a two-hour walking distance from a single health facility does not require the entire area of the country. First, we calculate the maximum possible distance from each health facility (2 hrs * 5 km/hr = 10 km). Then, we create a 10-km buffer to clip out the corresponding area from the time cost raster. The cost-distance algorithm would run on this smaller area.

[10] Leyk et al. 2019 provide a recent review of large-scale gridded population data products and their fitness for use.

Case Study: Access to health facilities

This model was used to check the quality of health clinic data, to extract the proportion of population covered, and to site possible new locations and to identify gaps. To support this effort, we used the mobility model to build the time-cost raster, then we conducted a cost-distance analysis to capture the area within 2-hour walking from each health facility (See graphic 2). This "catchment" area is then used to extract the population numbers contained therein, which could be taken alone, or aggregated to a more suitable administrative division or other geographic boundaries. Given that the population of Afghanistan is highly concentrated around water sources, time-cost raster model analytic methods provide a far more realistic understanding of the localized travel limitations, and thus, the population within range of health clinics. Clearly, those within walking distance of a clinic, may be *outside*, the district in which that clinic falls. To account for inter-provincial, or inter-district, travel for health services we split the catchment areas at the border so that people in District A being served by a clinic in District B are counted as "covered" in District A. Of course, if the intent is to capture the total number of people within two hours of a clinic, irrespective of their district of residence, we would omit this step.
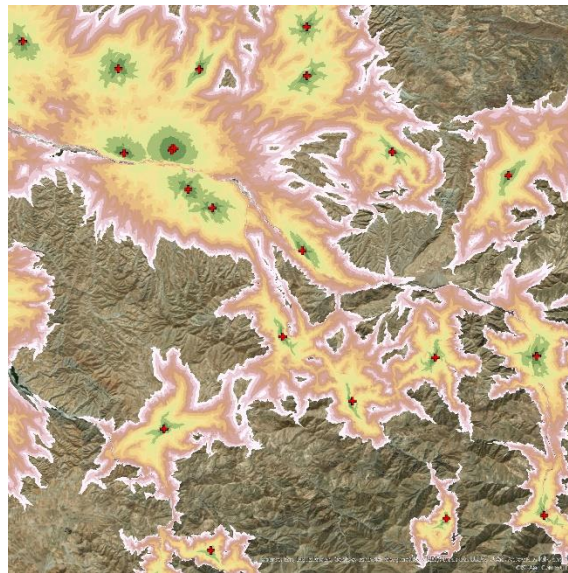
Graphic 2: Areas within two-hours walking of health clinics in Badakhshan Province

Specifically, in Afghanistan, this analysis was used to conduct spot checks on the data provided by the health monitoring and information system, specifically site locations and potential numbers of population served. For example, if a specific clinic states that they gave 1,000 vaccinations, but only 500 people are within the catchment area from the available population data, we can explore the causes for the discrepancy, which may include inaccurate population numbers due to coverage or currency such as increases in displaced persons and/or likewise inaccuracies in the travel time model inputs and estimates.

One clear example of the utility of this novel approach is comparing the populations served using the travel-time method compared to a generic distance buffer. The following graphic, Figure 1, depicts the population extracted from a 10km linear buffer and the population extracted using 2-hours travel-time, which, under optimal conditions, would also equal 10km. This example uses verified locations of schools.
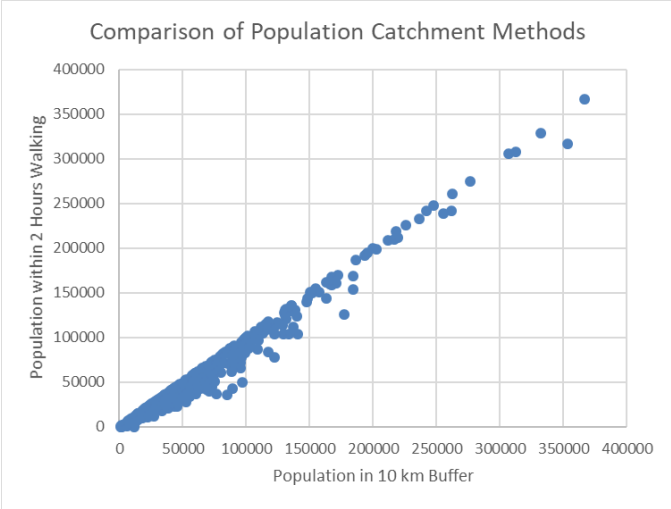
Figure 1: Variations in Population Served

If the linear distance and walking methods had the same served populations, the points on the scatterplot in Figure 1 would resemble a 45° line. On the surface, the linear distance method appears to be approximately correct, but out of 383 observations, more than 160 (40%) have population variations greater than 10% (see points in red in Figure 2).
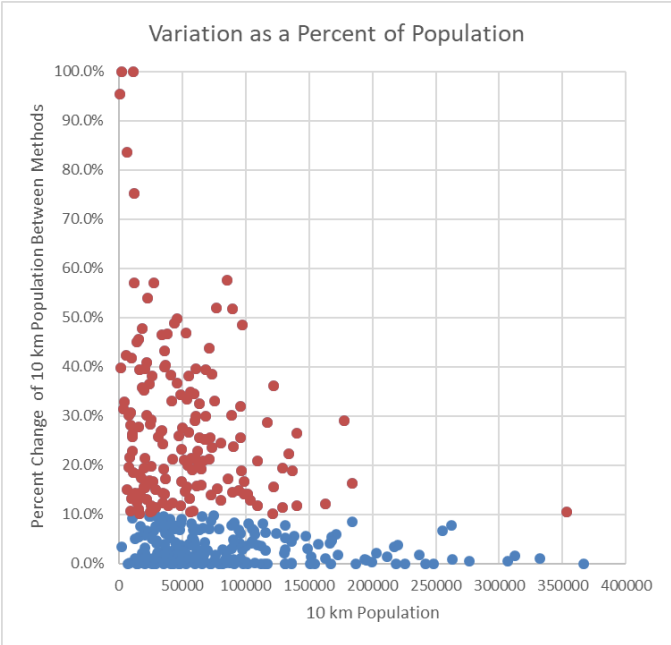

Figure 2: Percent of Population Difference

These variations are highest in the lower-density areas, which is intuitive because the lowest-density areas tend to have the roughest topography. In contrast, the four largest population centers in Afghanistan (Kabul, Herat, Kandahar, and Mazar-i-Sharif), all had virtually no difference between linear distance and walking models.

This analysis is also used to provide objective information in order to provide potential site locations for new facilities where the most people are within a specified travel time, rather than allow more subjective methods to take hold.

Other possible applications also exist. Not only can we consider access to health facilities at the district level, but we can estimate the share of the total population that can be reached within a level of travel time to a district center. Figure 3 illustrates that approximately half of the population is within 2 hours travel time of district center.
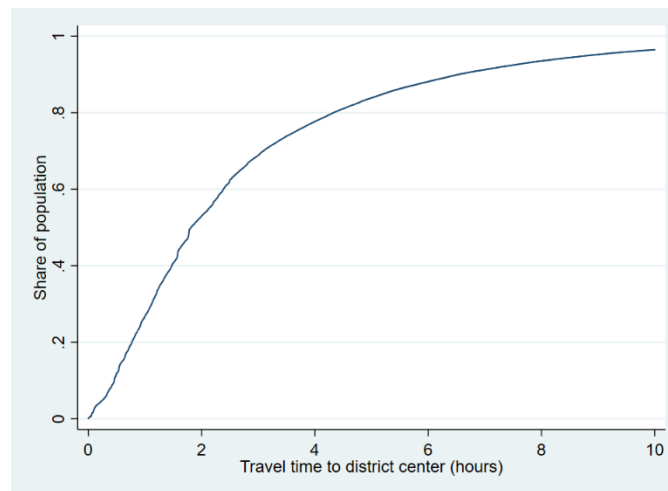

Figure 3: Share of population by travel time to a district center

Conclusion

Measuring physical access to services is important. Adapting from the literature, we document a flexible model to measure access at a comparatively high-resolution even in data-limited contexts. We present the following improvements: the nuance of the model results with detailed data, a geospatial method for calculating access in the presence of incomplete or topologically incorrect road networks, accounting for variations caused by extremely rough topography, and an application to compare and complement existing data of population served by health and education centers and administrative population information.

Known Limitations

All models are merely approximations of the true environment. For the most part in Afghanistan, the outputs cannot be tested in the real world, so we must be fully transparent regarding our assumptions and uncertainties and must take care to use the best possible data and models.

1. There can be concern about mixing data of differing resolutions (e.g. 10-meter landcover modified by 30-meter slope and going with a 10-meter output). We assume uniformity of slope across the entire slope pixel, which clearly deviates from reality. In the long run, we will acquire 10-meter elevation models from the European Space Agency (ESA)'s Sentinel Program to allow for direct pixel matching. We anticipate technology and measurement to continue to evolve and improvement and this is a flexible model to incorporate detailed data as it becomes available.

Ultimately, the time is per meter, so whether calculation is done with a 30-meter input or a 10-meter, the output time will be the same.

2. The slope is direction agnostic, i.e. the model does not know which way someone is traveling when calculating times, thus it is *more* conservative than reality. We did this to ensure the widest possible application for this model. If this is problematic, one could run the model without the slope, and add the slope using a different function in ArcGIS which would allow for directionality (going up versus down) to influence speed.

3. Possible improvements include alternative modes of transport, and the travel times associated therewith. Motorcycles present a unique challenge because they can access the same terrain as foot mobile humans, but with greater speeds, so we could apply an overall "Motorcycle modifier" to the foot version if so desired. Alternatively, and preferably, one could model for motorcycles, horses, etc., by modifying the time aspect rather than the underlying cost data. For example, two hours by foot would be 120 min, + 5% raster error is 126 min (Mulrooney and McGinn, 2017). If one assumes that a motorcycle could do twice the speed, double the time available to 252.

4. The Mixed-mode model does not factor in any penalty for shifting from foot to vehicle travel. Thus, the assumption that someone walks across country and gets into a waiting car clearly does not reflect reality.

References

Baum-Snow, N., Henderson, J. V., Turner, M., Brandt, L., & Zhang, Q. (2015). Transport infrastructure, urban growth and market access in China.

Berg, Blankespoor, Li, & Selod. Global travel time to major cities, circa 2010. Unpublished manuscript under preparation. Washington, D.C.: The World Bank.

Berg, C.N., Deichmann, U., Liu, Y. and Selod, H., (2017). Transport Policies and Development. Journal of Development Studies, 53(4), pp.465-480.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. The Quarterly Journal of Economics, 131(2), pp.799-858.

Jedwab, R. and Storeygard, A., 2017. The average and heterogeneous effects of transportation investments: Evidence from Sub-Saharan Africa 1960-2010 (No. 0822). Department of Economics, Tufts University.

FAO. (2016). The Islamic Republic of Afghanistan Land Cover Atlas 2010. Rome, Italy.

Fisher, Rohan and Jonatan Lassa. (2017). Interactive, open source, travel time scenario modelling: tools to facilitate participation in health service access analysis. International Journal of Health Geographics, 16:13.

Iimi, A., Ahmed, F., Anderson, E.C., Diehl, A.S., Maiyo, L., Peralta-Quirós, T. and Rao, K.S. (2016). New rural access index: main determinants and correlation to poverty. Washington, D.C.: The World Bank.

Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J., and Pesaresi, M. (2019). The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use, Earth Syst. Sci. Data, 11, 1385–1409.

Mulrooney, T., Beratan, K., McGinn, C., & Branch, B. (2017). 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.

Redding, S.J. and Turner, M.A. (2015). Transportation costs and the spatial organization of economic activity. In Handbook of regional and urban economics (Vol. 5, pp. 1339-1398). Elsevier.

Roberts, P., KC, Shyam. and Rastogi, C. (2006). Rural access index: a key development indicator. World Bank.

Tobler, Waldo (February 1993). "Three presentations on geographical analysis and modeling: Non-isotropic geographic modeling speculations on the geometry of geography global spatial analysis"
Uchida, H. and Nelson, A. (2010). Agglomeration index: Towards a new measure of urban concentration (No. 2010, 29). Working paper//World Institute for Development Economics Research.

Van Der Weide, R., Rijkers, B., Blankespoor, B. and Abrahams, A. (2018). Obstacles on the road to Palestinian economic growth. World Bank Policy Research Working Paper 8385.

Van Wagtendock, Jan and James Benedict. (1980). "Travel Time Variation on Backcountry Trails." *Journal of Leisure Research*. Published 2nd Quarter, 99-106.

Weiss, D.J., Nelson, A., Gibson, H.S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N. and Mappin, B., 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature.

Yoshida, N., and Deichmann, U. (2009). Measurement of Accessibility and Its Applications. *Journal of Infrastructure Development*, *1*(1), 1-16.

Annex 1: Land Cover Speed and Times

Land Cover Class	Key	Foot Speed (km/h)	Foot Multiplier	Time (min/m)	Mixed Speed (km/h)	Mixed Multiplier	Time (min/m)
1A	Settlements/Urban Areas	5	1	0.012	30	1	0.002
1B	Non Urban Built up areas	5	1	0.012	30	1	0.002
2A	Fruit Trees	4	0.8	0.015	4	0.8	0.015
2B	Vineyards	3.5	0.7	0.017	3.5	0.7	0.017
3A	Intensively Cultivated Areas	3.5	0.7	0.017	3.5	0.7	0.017
3A1	Irrigated Herbaceous Crop(s)	3.5	0.7	0.017	3.5	0.7	0.017
3B	Marginal Irrigated Crop	3.5	0.7	0.017	3.5	0.7	0.017
3C	Karez System	3.5	0.7	0.017	3.5	0.7	0.017
4A	Rainfed Cultivation in flat areas	3.5	0.7	0.017	3.5	0.7	0.017
4B	Rainfed Cultivation in sloping land	3.5	0.7	0.017	3.5	0.7	0.017
6A	Closed trees	2	0.4	0.030	2	0.4	0.030
6B	Open Trees	4	0.8	0.015	4	0.8	0.015
6B1	Open Trees Undifferentiated	4	0.8	0.015	4	0.8	0.015
6C	Close to Open Shrubland	5	1	0.012	5	1	0.012
7	Rangeland	5	1	0.012	5	1	0.012
8A	Bare Soil OR Rock outcrops	5	1	0.012	5	1	0.012
8B	Sandy Areas	2	0.4	0.030	2	0.4	0.030

8C	Dunes	2	0.4	0.030	2	0.4	0.030
9A	Marsh	2	0.4	0.030	2	0.4	0.030
9B	Seasonally inundated vegetation	2	0.4	0.030	2	0.4	0.030
10A	Artificial & Natural Waterbodies	0.06	0.012	1.000	0.06	0.012	1.000
10B	Seasonal Lakes	0.06	0.012	1.000	0.06	0.012	1.000
11	River	0.06	0.012	1.000	0.06	0.012	1.000
12	River Banks	4	0.8	0.015	4	0.8	0.015
13	Perennial Snow	1	0.2	0.060	1	0.2	0.060

Appendix 1: Experience from initial use

1. Given the processing time of these layers, it is highly advisable to use a buffer to mask out the areas from the overall time-cost raster. For example, for two hours of travel by foot, the model does not need to contain more than 15 kilometers from the start/end points.