RAISING THE PROFILE OF THE RURAL ACCESS INDEX AS A VITAL SDG INDICATOR FOR MEASURING RURAL DEVELOPMENT AND CONNECTIVITY

R Workman
TRL Ltd., Wokingham, UK
rworkman@trl.co.uk

K McPherson
TRL Ltd., Wokingham, UK
kmcpherson@trl.co.uk

A Bradbury
Cardno, London, UK
Annabel.bradbury@cardno.com

ABSTRACT

The Rural Access Index (RAI) was defined in 2005 as the proportion of a rural population living within 2 km of an all-season road. Initial measurements of the RAI, obtained through a variety of data collection methods for 64 countries, were published by the World Bank in 2006. In 2016 the RAI definition was adopted as Sustainable Development Goal (SDG) Indicator 9.1.1 requiring regular update of RAI data for the majority of United Nations (UN) countries. With support from the UKAid funded Research for Community Access Partnership (ReCAP), the World Bank developed updated spatial analysis techniques to measure RAI, and trials were carried out in 8 ReCAP partner countries. The outcome from these trials raised significant correlation issues. A follow-on status review of the RAI was conducted in 2018 with ReCAP funding and a clearer way forward was identified to accelerate progress within its geographical coverage. The consequent research is working to develop, propose and obtain agreement on a harmonised approach to data collection and measurement of RAI, and scale up implementation of the RAI across UN member countries, starting with a trial of the proposed measurement framework in Sub-Saharan Africa and South Asia. This paper will describe how the measurement approach for the RAI has been refined in order to eradicate inconsistencies in data collection, meet international standards, and provide a clear framework for data quality assurance.

1. INTRODUCTION

The Rural Access Index (RAI) is the ‘Proportion of the rural population who live within 2 km of an all-season road’. It was developed because of its relevance to poverty reduction strategies, given the recognised links between physical isolation and poverty. It was expected to provide stronger links to the Millennium Development Goals (MDGs) that preceded the SDGs between 2000 and 2015, and reinforce donor assistance to the sustainable development of beneficiary countries. The measurement of the RAI has been adopted as SDG indicator 9.1.1, and was promoted from a Tier III indicator to Tier II in December 2018. Tier II signifies that the RAI is conceptually clear, has an internationally established methodology and standards are available, although data are not regularly produced by countries. To progress to Tier I, RAI data must be measured regularly for at least 50% of UN countries and of the population in every region where the indicator is relevant.

The Research for Community Access Partnership (ReCAP), funded by UKAid, initiated a project to develop a harmonised approach to data collection and measurement of the RAI that is relevant, consistent and sustainable. The project will facilitate scaling up implementation of the RAI across UN member countries. Phase 1 was completed in 2018, and included a comprehensive review of the status of the RAI to date (Vincent, 2018). It included a detailed history of the development of the RAI and gave several references to key documents.

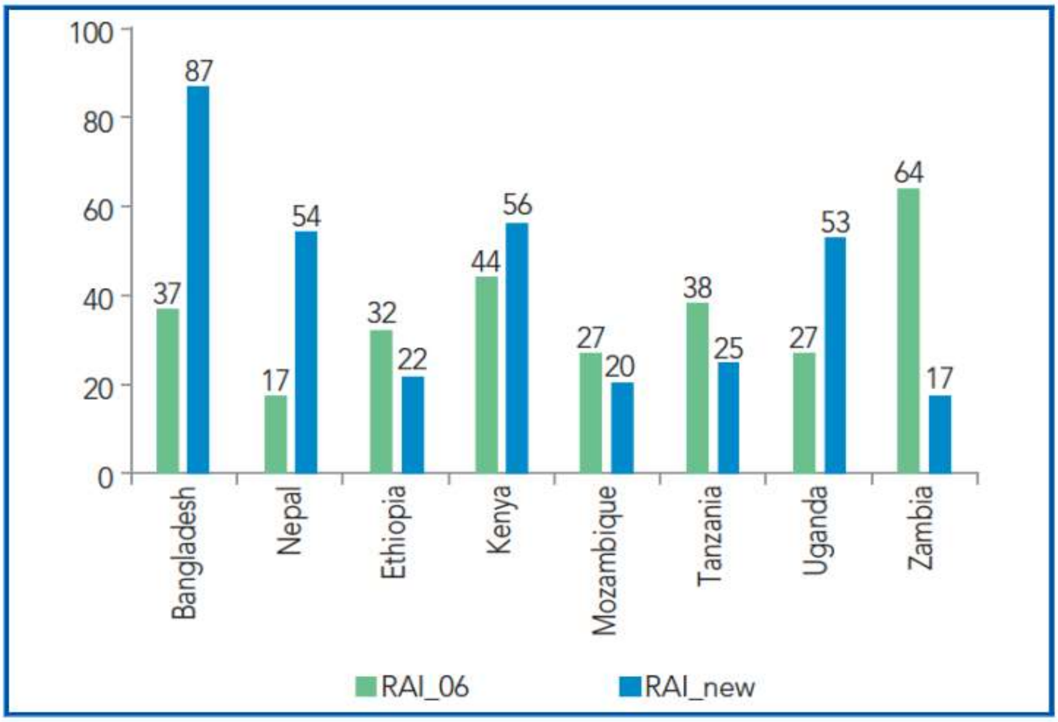
Phase 2 is an opportunity to consolidate existing and proposed approaches to data collection and refine the RAI methodology in collaboration with the World Bank and other stakeholders. The refined methodology should aim to eradicate inconsistencies in data collection, meet international standards and provide a clear framework for data validation. This process is expected to raise the profile of the RAI by putting in place clear data-quality standards and robust analytical methods. This will enhance the RAI's reliability as an indicator (in terms of accuracy and repeatability) and should increase the number of countries routinely collecting RAI data.

2. BACKGROUND

Initial measurements of the RAI, obtained through a variety of data collection methods, including interpretation of existing household surveys and modelling, were established for 64 countries and were published by the World Bank in 2006. In 2016, with funding from UKAid through ReCAP, a World Bank team developed and tested a new methodology to measure the RAI using spatial techniques and innovative technologies. This was based on creating three layers in Geographic Information System (GIS) software, comprising population, road location and road condition, overlaying them and using GIS tools to calculate the RAI.

The new methodology was intended to be sustainable, consistent, simple, and operationally relevant. It was considered to be potentially more cost effective and sustainable than the original methods used in 2006 (Roberts, 2006). The methodology was tested in eight ReCAP countries in 2016, and the outcome from these trials raised significant correlation issues (Iimi, 2016) with some countries varying by up to plus or minus 50%, as can be seen in Figure 1.

The Status Review, conducted in Phase 1, investigated the reasons for discrepancies in the data collection methods and results between 2006 and 2016 (Vincent, 2018), and Phase 2 was borne out of a need to develop, propose and obtain agreement on a harmonised approach to data collection and measurement of the RAI.



Source: World Bank Transport & ICT, 2016

Figure 1: RAI measurements in 2006 and 2016

3. INCONSISTENCIES IN DATA

A number of inconsistencies in data collection and analysis were noted in the Status Review report. Data is analysed using three layers of GIS information, as shown in Figure 2.

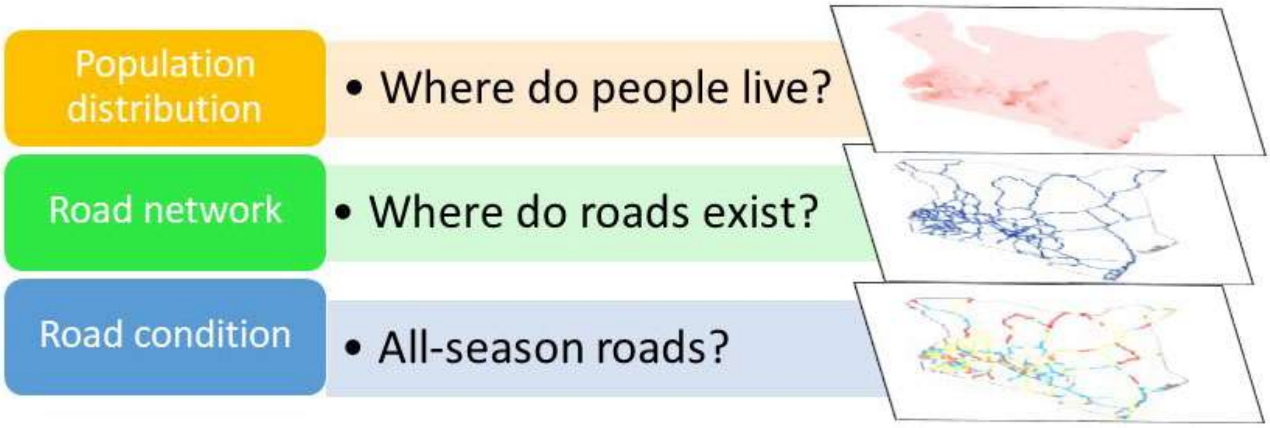


Figure 2: Layers of GIS information

3.1. Inconsistencies in GIS layers

The main issues in the three layers of information are:

3.1.1 *Population layer*

Reliable population data is available globally, free at source and linked to national census data. Of the different available sources, WorldPop (worldpop.org) is generally regarded as being reliable and comprehensive, and provides the highest spatial resolution at 100 m x 100 m. However, these population databases do not provide rural and urban boundaries, which are an essential input to the indicator. Also, there is no uniform global definition of rural and urban: the World Bank has noted that the minimum population required for an urban area can vary from 200 to 50,000, depending on the country (World Bank, 2018). The United Nations Statistics Division (UNSD) advises that, because of national characteristics which distinguish urban from rural, each country should decide which areas are to be classified as urban and which as rural, in accordance with their own circumstances.

More recently, the UN held an Expert Group Meeting on Statistical Methodology for Delineating Cities and Rural Areas in January 2019 (UN, 2019), to instil some consistency into the definitions. It concluded that adjustments and adaptations of the Degree of Urbanization (DegUrba) methodology were needed in a number of participating countries, to better reflect national circumstances. Critical to the development and implementation of the DegUrba methodology is the quality of underlying statistical data on population. The meeting recommended engagement on the DegUrba approach by the Inter Agency and Expert Group on SDG Indicators (IAEG-SDGs) to help with monitoring SDG indicators. For national level reporting of SDGs, national definitions of urban-rural are used. For global reporting and aggregations, however, a harmonised approach is worth considering, for specific SDG targets and for rural-urban disaggregations.

The rural and urban divide can have a significant impact on the RAI. It is also recognised that population migration within countries is a dynamic process and the population location is continually changing, usually between rural and urban with increasing urbanisation.

3.1.2 *Road location layer*

To measure the RAI using geo-spatial techniques it is important to know the full extent of the road network and where it is located. Many developing countries do not have complete GIS coverage of their rural road network, although coverage of the strategic or main road network is usually quite good. There are several sources of mapping available, which are derived from various techniques to map roads. OpenStreetMap (OSM) is generally regarded as the most accurate and comprehensive open data set available, and was used by the World Bank in 2016 to measure RAI. However, although OSM tends to have good coverage in urban areas, it tends to be less comprehensive and accurate in rural areas, which are the most important for measurement of the RAI.

3.1.3 *Road condition layer*

In addition to establishing the road network location, it is necessary to determine which roads are ‘all-season’. Past RAI assessments have interpreted all-season as paved only, and/or have interpreted it from road condition. Few countries formally measure the condition of their rural roads, but in the 2016 study the World Bank team arranged for road roughness to be measured using a smartphone app, namely the RoadLab Pro application. This can provide an approximation of the International Roughness Index (IRI), but it is not regarded as an accurate assessment of “all-season” especially for unpaved roads whose condition can change on a daily basis.

The World Bank team used a simple way to link condition with the all-season parameter. This was simply that all paved roads in good and fair condition would be assumed to be all-season, and all unpaved roads in good condition would also be all-season. This is a quite simplified correlation which, although appropriate for the assessment at that time, is not a particularly accurate or sustainable way to determine all-season roads.

3.2. Other Issues

The RAI Status Review also recognised other issues with RAI measurement, which include the following.

3.2.1 *Country comparisons*

Given the different geography, environments, climates, vegetation, vehicles, data collection technologies and statistical definitions in each country, it will be difficult to make direct comparisons of RAI globally. However, what is probably more valuable is to monitor how individual countries perform in terms of RAI year on year.

3.2.2 *Modes of transport*

The RAI does not adequately reflect recent changes in modes of rural transport. In many countries over the past 10 years or so, Intermediate Means of Transport (IMT) such as motorcycles, three-wheelers and others have become more prevalent. The RAI definition specifically targets four-wheeled vehicles as the prevalent means of transport, which in some cases is not entirely true. Many people living in rural areas are now able to use motorcycles and other IMTs for access to essential services such as hospitals, schools and markets.

3.2.3 *Coordination between local agencies*

It is usually the National Statistical Office (NSO) or equivalent statistical organisation that has the responsibility to report the SDGs. In the case of SDG 9.1.1 this would require close liaison with the national and local roads organisations. This is not a natural partnership in many cases, so there are some reservations about how well this would work in practice, especially when there is little direct support to countries in the form of technical assistance or coordination to measure and report SDG 9.1.1.

4. NEW MEASUREMENT APPROACH

A number of suggestions were made in the Status Review on how to improve data consistency, meet international standards and establish a clear framework for data validation.

The data available in each country will vary. Typically countries collect data for many different purposes, and will collect data in the most efficient and cost effective way to suit their needs. There is no advantage in trying to force a country to collect elaborate and expensive data if it is beyond their means and they see no benefit in collecting it. In order to develop a sustainable system, different types and even different qualities of data should be allowable, so the methodology needs to account for this and provide different ways to assess and adjust the data, to be relevant and useful for the RAI.

The suggestions for data collection and measurement of RAI suggested in Phase 1 have been built on by the present project team and are outlined as follows.

4.1. Population data measurement

WorldPop is generally regarded as the most easily accessible and granular population database at present, although it should be checked against some ground truthing for accuracy. With regards to the definition of urban/rural boundaries, it is recommended that the individual country definitions are retained for purposes of RAI measurement, but a parallel measure of population based on density or settlement size be used in order to allow better comparison between countries.

It is essential that the NSO is able to report figures as 'official', so government approved figures must be used. The ideal scenario is for the NSO and WorldPop to reconcile their figures. It is also desirable to disaggregate the RAI to the district, or equivalent, so that different factors can be applied where the climate or terrain differs significantly.

Figure **3** shows the process by which reconciliation and disaggregation may be achieved. While the NSO may retain population census data at the level of the household (first graphic), it may provide boundaries of enumeration areas along with aggregated population of those areas to WorldPop (second graphic), so that WorldPop can re-run its population models to ensure that its data match those of the NSO at the level of enumeration area (third graphic).

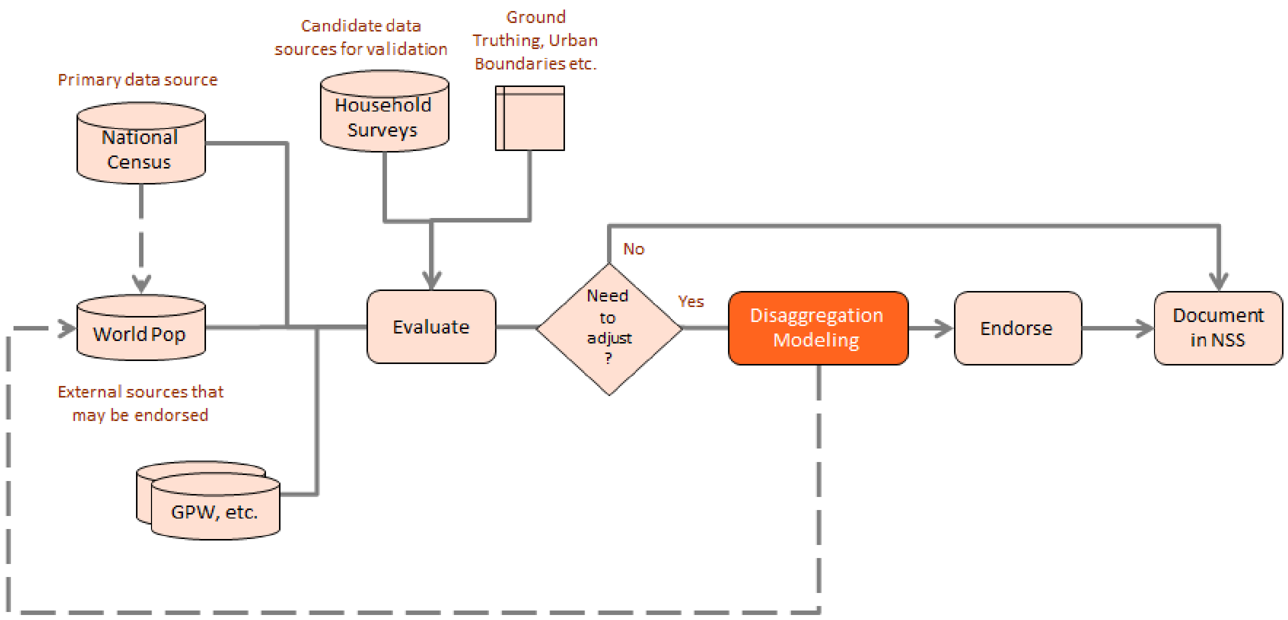


Figure 3: Population reporting

Figure 4 shows the process by which reconciliation and disaggregation may be achieved. Whilst the NSO may retain population census data at the level of the household (first graphic), it may provide boundaries of enumeration areas along with aggregated population of those areas to WorldPop (second graphic), so that WorldPop can re-run its population models to ensure that its data match those of the NSO at the level of enumeration area (third graphic).

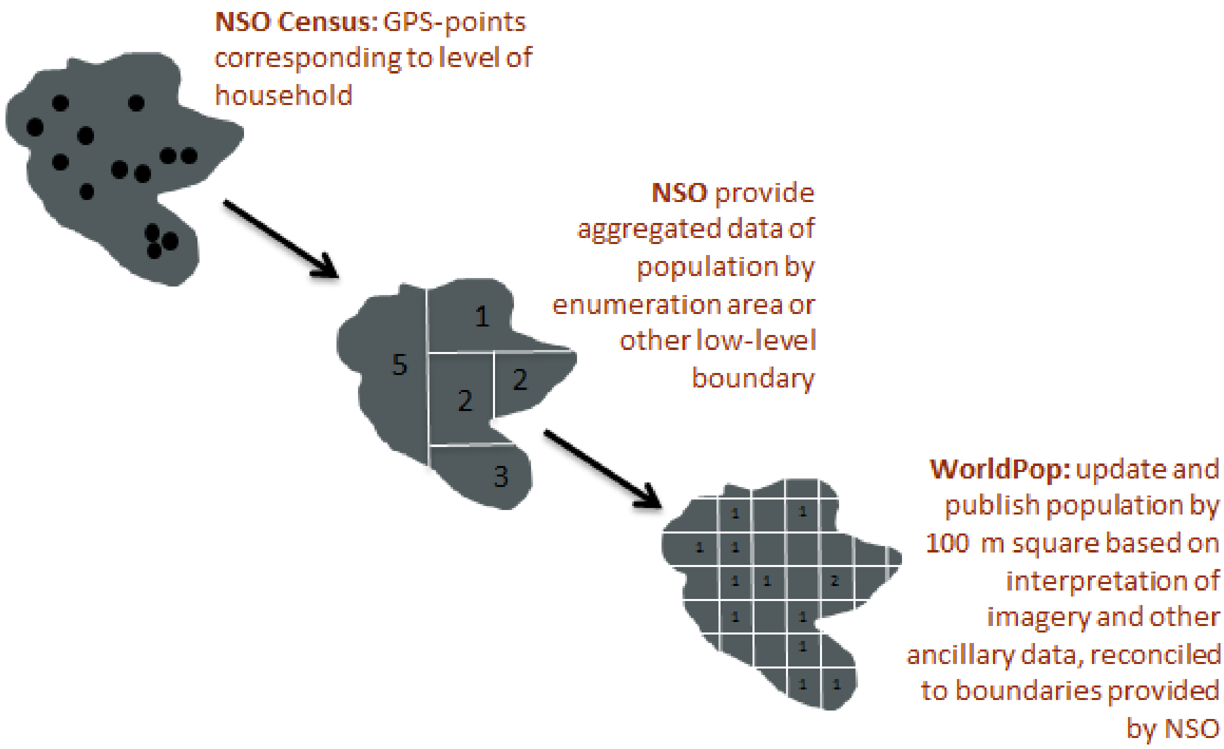


Figure 4: Population disaggregation

4.2. Road location measurement

In order for a country to provide official statistics on RAI it is likely that they are required to use the official road network data as collected and published by their respective roads agencies, even if this may be less complete than other sources, such as OSM. It is recommended that countries use OSM as a tool to both update their own networks, as well as to upload data to OSM, as shown in Figure **5**. In this case OSM would eventually become the most reliable global source of road mapping and would be freely available. This would facilitate the accurate and frequent measurement of RAI for all countries.

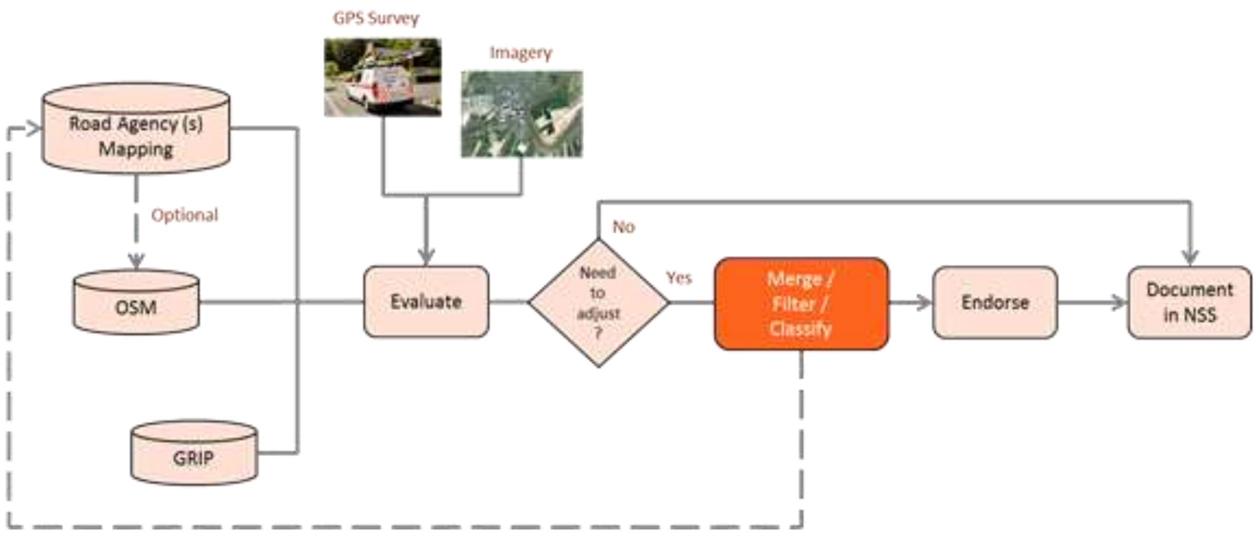


Figure 5: Road location

There is also an argument for adjusting the 2 km distance from network parameter for mountainous countries. In a country like Nepal for example, 2 km on the plains (20-25 minutes' walk) is a very different measure to 2 km in the mountains (up to 2 hours, or more if there are natural obstacles such as rivers or ravines). The team is looking at adjustment factors for particularly tortuous terrain.

4.3. Road condition measurement

The challenge is to determine whether roads are all-season or not. Condition itself is not a requirement for RAI. It is however used at the moment to interpret whether a road is all-season or not, to varying degrees of success. If all-season could be measured directly, or estimated accurately, this would greatly enhance the precise measurement of RAI and improve its sustainability. The team are testing the possibility of having country parameters that would determine all-season road condition directly, for example using rainfall, surface type and topography to determine the likelihood of a road being all-season. This could provide an approximation that would need to be verified, but this check should be possible by interviewing local engineers or technicians who know the local roads. This also avoids the use of condition data that could be inaccurate due to difficulties or inconsistencies in measuring it, or that have been distorted in order to gain additional funding for maintenance on political grounds. It is noted that while OSM can store condition data, very few roads currently have such data stored against them.

4.3.1 *Accessibility Factors*

An alternative means to identify "all-season" roads has been proposed, called Accessibility Factors. This takes a different approach to identifying all-season roads that is broader based, more sustainable and should facilitate international comparison. It does not require ground measurements of road condition to be made.

Accessibility factors are based on variables that can determine the all-season nature of a road, such as road surface, climate and terrain. The accessibility factor can be determined broadly for a country, but it is recommended to be applied at a reasonable level of population distribution, typically at district or county level, or climate zone level. This can then be aggregated up to provide a national figure for RAI.

The accessibility factor determines the likelihood of a road being all-season, or the risk of a road being inaccessible. This is closely aligned with the intent of the original study, i.e. "accessible all year with the prevailing mode of transport", and "… may be temporarily unavailable during inclement weather" (Roberts, 2006).

There are a number of characteristics of a country and its roads that will have a bearing on the accessibility factors, including:

- Road surface type: The road network can be divided into paved roads and unpaved roads. Unpaved roads are more vulnerable to being impassable to traffic than paved roads. They therefore have a higher risk of reduced accessibility, and hence a lower accessibility factor would be applied.
- Climate: Each country is unique in its climate impacts on roads. Climate change is exacerbating the risks of impassability through extreme weather events. Individual events should not affect the ability of a road to be all-season, because individual washouts and damage can be repaired, but the risk to a network will affect the accessibility. Countries with a benign climate will experience a lower level of road closures due to weather and should therefore apply a higher accessibility factor,

whereas countries with tropical or monsoonal rains will experience a higher risk of road closure due to weather and should apply a lower accessibility factor.

- Terrain: Gradient of roads has an effect on their durability and their ability to withstand heavy rainfall. Steep roads are more vulnerable to scour, washout and slipperiness, especially when unpaved. Therefore countries with mountainous areas should apply a lower accessibility factor to those areas, and countries with flat or rolling terrain should apply a higher accessibility factor.

A potential example is shown in Figure **6**, with (a) showing paved roads, and (b) showing unpaved roads.

		Terrain	
		Low Risk (e.g. Flat, Rolling)	High Risk (e.g. Mountainous)
Climate	Low Risk (e.g. Dry Season)	1	1
	High Risk (e.g. Wet Season)	1	0.95

(a) - Accessibility Factors for Paved Network (Malawi)

		Terrain	
		Low Risk (e.g. Flat, Rolling)	High Risk (e.g. Mountainous)
Climate	Low Risk (e.g. Dry Season)	1	0.95
	High Risk (e.g. Wet Season)	0.95	0.90

(b) - Accessibility Factors for Unpaved Network (Malawi)

Figure 6: Accessibility Factors example

Figure 7 shows a worked example of how RAI would be defined and calculated for a simple area with 3 roads.

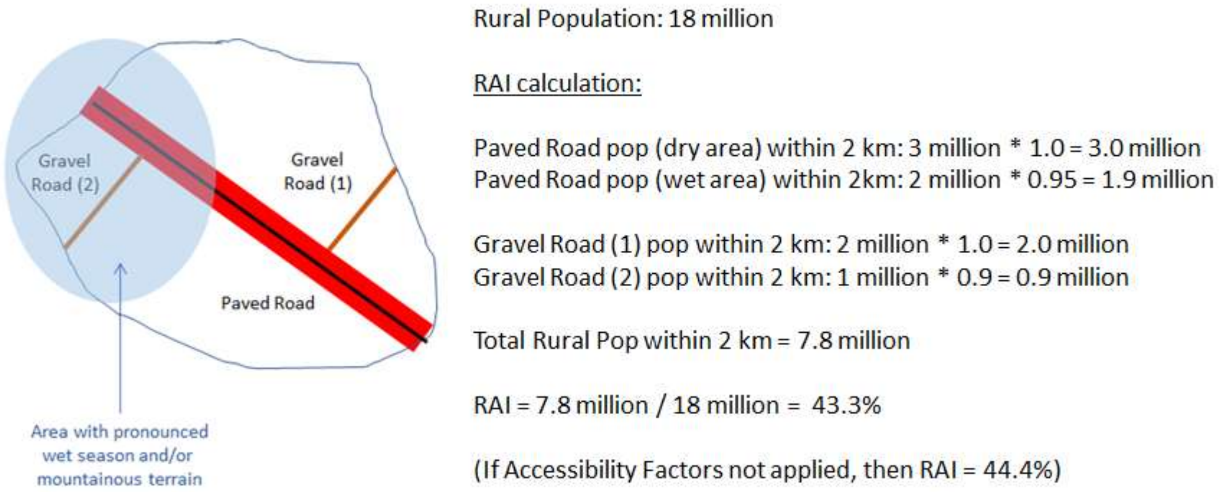


Figure 7: Worked example of accessibility factors

4.4. Check data quality

The quality of the RAI measurement is dependent on the quality of the data used. It is understood that a variety of data quality issues will arise in RAI data collection. The team is working towards a framework or checklist that can be used to assess data quality before it is accepted into the RAI assessment.

4.4.1 *Population*

For population, the WorldPop database should be evaluated, any issues will be in the spatial distribution of population and the methods used to estimate this. It is noted that the WorldPop methodology is open source, and that if there are any issues relating to its estimating methodologies in a given country, then the methods may be refined.

4.4.2 *Road Mapping*

For road mapping there are various sources that can be used to check road organisation data, from OSM to open source maps like Google and Bing. Tools and algorithms are being developed by a number of organisations that can identify changes in the road network from satellite imagery using Machine Learning (ML).

4.4.3 *Road Condition*

For road condition, a summary of data quality has been proposed, as shown in Table 1. This shows four potential levels of quality, based on how accurate or up-to-date the data is likely to be. The highest level (Platinum) is to measure directly whether a road is all-season or not, i.e. determine how many days per year it is closed for. The source of this data could be local records, or local knowledge from local roads organisations. The Gold standard would be interpretation of measured surveys. For this data to be fully checked the quality of the surveys themselves would need to be verified. Silver would comprise interpretation of data from less conventional sources, while the lowest level (Bronze) would consist of secondary sources of data that have not been directly corroborated. This table is in the process of being verified and quantified, so the ranges shown in the accuracy column are only indicative at present.

Table 1 – Condition data quality assessment

Standard	Data	Data Source	Quality Assessment
Platinum	Direct measurement of all-season	Combination of official measured surveys & local knowledge, interviews with local engineers/technicians	90-100%
Gold	Interpretation from established measured surveys	Data from official visual surveys, IRI roughness and other recent surveys, HH surveys	80-90%
Silver	Interpretation from less conventional	Interpretation of speed for condition, satellite or drone	70-80%

	sources	imagery, mobile phone data	
Bronze	Secondary sources of data	Historical data, OSM, Global Roads Inventory Project (GRIP), travel time surveys, origin-destination surveys, etc. which have no corroboration with other sources	<70%

4.5. Correction Factors

When a systematic error is identified, for example if it is known that the extent of the rural road network is under-represented in a GIS analysis, a suitable country or region specific correction factor should be calculated and applied. The method used to calculate any such correction factor must be fully documented, and should be reviewed by a statistician, preferably from the NSO, to confirm that the correction factor is appropriate.

4.6. Calculation of RAI

The calculation of the geo-spatial RAI is essentially carried out in GIS software, using the three different layers that are selected. This calculation is based on the guide to compute the RAI contained in Annex 1 of the 2016 World Bank report (Iimi, 2016). The principle of the calculation is valid, although it may vary with different GIS software packages. The 2016 calculations were made using ArcGIS which is expensive to license, but many countries will use open source software such as QGIS because it is lower cost. A new guide will be produced using examples from QGIS.

Therefore different data collection methods and processes are covered in this methodology, including ways to account for and adjust data that is known not to be 100% accurate or complete.

4.7. Alternative indicators

In some countries it is apparent that rural transport has become dominated by IMTs, such as motorcycles. In such cases measurement of RAI is still required, but the methodology is being refined to allow measurement of an alternative value, based on local interpretation of all-season access to any transport service within 2 km. This would include tracks regularly used by motorcycle taxis.

These alternative values will be accompanied by a detailed explanation of why a secondary value was measured, and will be subject to the same statistical verification as the standard value of RAI. This measurement is expected to provide input to the development of more suitable measurements of rural access in the future, for instance to incorporate inland waterways, which are essential for connecting some rural communities that are inaccessible by road.

It should be noted that RAI is seen as a “network indicator” measuring access to infrastructure. There is considerable research at present on “service indicators” that measure access to health services, education services, banking services etc. These service indicators are seen as complementary to RAI, and both types of indicator may be used in a country’s rural development policies or programs.

4.8. Involve additional partners

At present the World Bank is the custodian of SDG 9.1.1. There are two official partners, the United Nations Economic Commission for Europe (UNECE) and the United Nations Environment Programme (UNEP). Other Multilateral Development Banks (MDB) are also expected to join, such as the Asian Development Bank (ADB) and the African Development Bank (AfDB), with others potentially interested. The aim of the project is to have full coverage globally by MDBs as partners. This will provide more weight and a higher status for the RAI, with partners potentially being able to provide impetus for RAI data collection via projects and other initiatives.

4.9. Establish a measurement and reporting framework

There is no formal measurement and reporting framework for SDG 9.1.1. The team have proposed a process that can be followed by the custodian, partners and countries to collect data, report and store data, as well as presenting it on an appropriate platform. This generic process is shown in Figure 8 and may be used as the basis for incorporating the RAI into a country's statistical system. This process will evolve as more partners become involved and as tools and services such as those available through the UN Global Platform emerge.

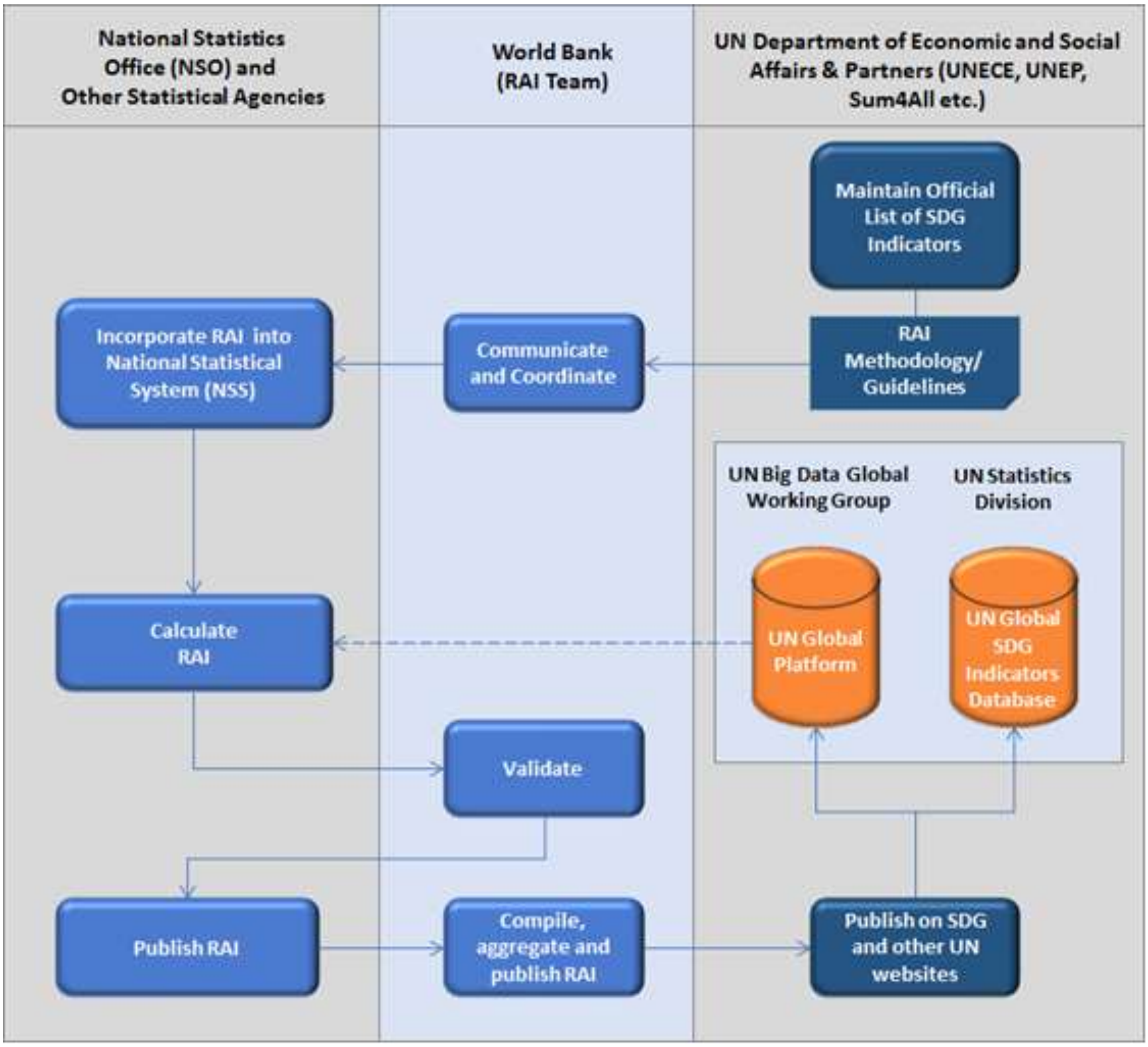


Figure 8: Proposed process for data collection, processing and reporting

4.10. Frequency of updating the RAI

The authors of the original RAI in 2006 recommended a minimum assessment frequency of three years. This was largely based on household survey data, although a GIS based method was envisaged at that time.

With current open source population data being available on an annual basis, the main constraint to improving this frequency will be the updating of road network mapping and road condition data. Many countries carry out annual condition surveys on their primary and secondary network, but even though the tertiary network often has a recommended survey frequency of one or two years, this is seldom achieved. Condition surveys on rural roads which receive little or no maintenance could seem to be somewhat unnecessary to local roads organisations, and this along with a common lack of resources can mean that surveys are delayed for several years.

Under the current methodology it is recommended that the RAI assessment is made when significant new data is available, but with a minimum frequency of every 3 years (extendable to 5 years in exceptional circumstances). This would mean employing different frequencies of RAI measurement in different countries, which, although it makes the monitoring of RAI more complicated, is positive in promoting sustainability. If, as seems likely in the not so distant future, RAI can be calculated automatically, the frequency can be increased.

4.11. Future options for measuring RAI

A number of alternative options have been explored for measuring the RAI in future. Prominent amongst these is the use of mobile phone data. Mobile data is increasingly being used for transport surveys, and to monitor traffic in real time. There is therefore potential for mobile data to indicate whether a road is being used, and how fast the people using the road are travelling. This could be used to estimate when a road is closed and how long for, and therefore whether a road is considered as all-season or not. This can be determined by the connection of phones to cells, it is not necessary to use call data or GPS measurements.

With the expansion of mobile phones in developing countries, and in Africa especially, this possibility is becoming increasingly plausible. The drawbacks would conceivably be the cost of data, which at present is only provided free of charge by one operator, and the potential data protection and permissions required.

The use of satellite imagery is another possibility to identify all-season roads. This potential has been researched by ReCAP (Workman, 2017) in five countries in Africa and found that manual assessment of road condition is possible to an accuracy of between 65% and 85%, depending on the type of road and the number of condition levels used. However, the trials used very high resolution satellite imagery which is relatively expensive to procure, although it would save funds normally spent on traditional driven visual condition surveys, and reduce any associated logistical issues.

There are also various methods that can be employed to determine the International Roughness Index (IRI) value, but it would still need to be interpreted to determine whether the road is all-season or not.

5. CONCLUSIONS

The project is taking a pragmatic approach towards RAI data collection and measurement. The priority is to maintain sustainability so that regular RAI monitoring can continue beyond the project and can be managed largely by the countries themselves, with minimal input from the custodian and other partners. This means that new and onerous data collection regimes should not be imposed, especially given that the RAI by itself is not widely used for development planning in most countries. The refined approach will be trialled in four ReCAP partner countries, Ghana, Malawi, Myanmar and Nepal. These countries were selected because of their varied environments and previous experiences with RAI, with a view towards learning lessons and refining the RAI further.

The new framework will utilise existing and freely available data. It will provide guidance on checking the quality of RAI data, and will produce parallel indications of data where appropriate, in order to provide some consistency between countries and regions. The recommended process for data collection, calculation and presentation of the RAI is designed to maximise impact and raise the profile of the RAI, in order to consolidate it as the only direct SDG indicator of rural access.

REFERENCES

1. Iimi, A. Diehl, A. 2015. A new measure of rural access to transport -Using GIS Data to Inform Decisions and Attainment of the SDGs'. World Bank, Washington DC.
2. Roberts P, Shyam KC, Rastogi C, 2006 'Rural Access Index: A Key Development Indicator', Transport Paper TP-10. World Bank, Washington DC.
3. UN, 2019. United Nations Expert Group Meeting on Statistical Methodology for Delineating Cities and Rural Areas. Available at: https://unstats.un.org/unsd/demographic-social/meetings/2019/newyork-egm-statmeth/
4. Vincent S, 2018. 'Status Review of the Updated Rural Access Index (RAI)-Final Report'. ReCAP, London, UK.
5. Workman R, 2017. 'AfCAP Guideline on the use of high tech solutions for road network inventory and condition analysis in Africa', ReCAP, London, UK.
6. World Bank, 2018. Atlas of Sustainable Development Goals; From World Development Indicators. World Bank, Washington DC.