POTENTIAL FOR MEASUREMENT OF THE RURAL ACCESS INDEX IN THE FUTURE

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

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. At present the methodology for the RAI is being refined 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 explores alternative method of RAI measurement for the future, using satellite imagery, mobile phone data and a range of alternative technologies. Some of these are likely to be possible in the short term, but others are looking beyond the horizon and will require a leap in technology to become feasible.

1. INTRODUCTION

The Rural Access Index (RAI) is an indicator developed by the World Bank to measure rural accessibility. It measures the proportion of the rural population who live within 2 km of an all-season road. It was developed with the aim of poverty reduction in mind and focuses on the recognised links between physical isolation and poverty. The RAI has been adopted as SDG indicator 9.1.1 and is currently at Tier II level, which means that is conceptually clear, has an internationally established methodology and standards are available, but data are not regularly produced by countries.

The Research for Community Access Partnership (ReCAP), funded by UKAid, initiated a project to refine the methodology for the RAI to make it more relevant, consistent and sustainable. Phase 1 was completed in 2018 with a comprehensive review of the RAI status (Vincent, 2018), which included a detailed history of the development of the RAI and established the focus for Phase 2. The project is currently in Phase 2 and has consolidated existing and proposed approaches to data collection and refined the RAI process in collaboration with the World Bank and other stakeholders. The refined process should eradicate inconsistencies in data collection, meet international standards and provide a clear framework for data validation.

The methodology at present focuses on geo-spatial technologies to measure RAI. Some other methods of measurement were considered, but these are not yet at the stage of development whereby they can reliably be used to measure RAI. They do however have potential for the future measurement of RAI, and are being considered in this paper for the medium to long term measurement of RAI.
2. BACKGROUND

The RAI was initially measured through a variety of data collection methods, but mainly the interpretation of existing household surveys. The first study in 2006 was carried out for 64 countries and was published by the World Bank (Roberts, 2006). In 2016 a World Bank team developed and tested a new methodology to measure the RAI using spatial techniques and innovative technologies, based on creating three layers in Geographic Information System (GIS) software. These layers included population, road location and road condition, and GIS tools were used to calculate the RAI.

The new methodology was designed to be sustainable, consistent, simple, and operationally relevant, and was considered to be potentially more cost effective and sustainable than the original methods used in 2006. The methodology was tested in eight ReCAP countries in 2016 (limi, 2016) and was successful, although the results from these trials raised some correlation issues with the household survey interpretations from 2006.

The Status Review, conducted in Phase 1, looked at the differences in data collection methods and results between 2006 and 2016 (Vincent, 2018), and Phase 2 was designed to develop, propose and obtain agreement on a harmonised approach to data collection and measurement of the RAI. The 2016 methodology is established and the geo-spatial approach is the main principle for RAI and is the basis upon which it was promoted to Tier II of the SDG scale.

However, the 2016 (limi) report did consider alternative methods of measurement for the future, which included satellite imagery. At the time there was not sufficient evidence to trial this as a methodology, but its potential was recognised. The Status Review (Vincent, 2018) also considered this issue, and suggested the potential for using mobile phone data to determine the all-season status of a road.

This paper will consider these options in more detail, and will also consider a range of other options for RAI measurements into the future.

3. FUTURE OPTIONS FOR MEASURING RAI

The options presented here for measuring RAI in the future are largely based on existing technology that could potentially be innovated and developed to assist in RAI measurement. In most cases the technology is not at a stage of development whereby it can be used to measure RAI at present, but there is potential for it to be changed and developed to do so. This paper considers how these technologies could be used to measure RAI in the medium to long term, beyond five years in the future, and maybe much longer. The transport sector is changing rapidly with electric power, autonomous vehicles, alternative means of transport and the pervasive effects of climate change. It is difficult to predict what will happen in the future, but this paper proposes some potential technologies that could be used for RAI measurement, and may well have other uses in the sector.
3.1. Satellite Imagery

The use of satellite imagery is a possibility to identify all-season roads. The potential to assess and report road condition from satellite imagery has been researched by ReCAP (Workman, 2017) in five countries in Africa. This research found that manual assessment of road condition from imagery is possible to an accuracy of between 65% and 85%, depending on the type of road and the number of condition levels used. The trials used very high resolution satellite imagery, which is relatively expensive to procure, so the process was not developed further for asset management.

![Figure 1: Monitoring of road condition from satellite imagery (Workman, 2017)](image)

Research is still ongoing at TRL to determine the parameters of using road width and surface texture to indicate road condition, with a view to developing a structure that could inform an algorithm or machine learning based approach to condition assessment. However, ultimately this would only provide condition, which would then again have to be interpreted as all-season status.

There would be, however, potential for satellite imagery to estimate all-season status directly. The current accessibility factor aspect of the methodology bases the all-season status on the intention of the road to be all-season, as well as critical road features such as the surface type, the terrain and the climate.

**Terrain:**

If the satellite image is accurately orthorectified using Digital Elevation Models (DEM) it can be used to indicate the nature of the terrain. The important aspects here are the gradient of roads, which could potentially be determined on an individual roads basis if necessary.

**Climate:**

Climate records are well documented around the world and would be available as GIS layers to overlay the networks for each country or each area. The climate is a factor that will contribute towards determining the accessibility factor.
**Surface type:**

The surface type of the road is important because different types of road react differently to traffic, terrain, climate and maintenance inputs. It is expected that roads would be divided into the broad areas of paved and unpaved, as has been implemented in the 2016 methodology (Iimi, 2016). The condition factors within those areas were then set based on IRI or visual condition assessment. Road surface type is relatively easy to determine from very high resolution satellite imagery, including the resolutions that are typically used for open source mapping such as Google Earth and OSM. It should therefore be possible to extract surface type from OSM if it is found to be sufficiently reliable, and alternatively it should also be possible to develop an algorithm to identify surface condition from existing satellite imagery.

**Width and width variation:**

The width of a road and its variation in width also has the potential to indicate the all-season nature of a road. Given that the accessibility factor is based on the intention of the road to be all-season it should be possible using ground truthing to identify typical widths and variation of road widths to correlate with all-season roads. The width and variation of width should be possible to measure automatically from a satellite image, especially if the road centre line is already mapped.

All of these factors together could provide an accurate assessment of the all-season status of a road, using satellite imagery. The drawbacks would be the cost and availability of the imagery. There would also need to be extensive research to determine the parameters of climate, terrain, surface type and width that could indicate an all-season road.

3.2. Mobile Phone data

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 using passive events, i.e. by the connection of phones to cells; it is not necessary to use call data or GPS measurements. This type of monitoring can be cell to cell, or between groups of cells, as shown in the diagram in Figure 2 (Hemmings, 2016). The top line shows cell to cell monitoring, which gives greater resolution than between groups of cells.

![Figure 2: Passive monitoring of mobile networks (from Hemmings, 2016)](image.png)
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.

3.3. Social Media

With the expansion of smartphones, social media has naturally followed and is being used in the most remote areas of the world that have mobile coverage. The remotest areas are, however, likely to have the least coverage of social media, and these are the areas that are most important for measuring RAI because they have less chance of being connected to an all-season road. However, with the expansion of internet connectivity and the probability that broadband will be provided to the remotest areas of the world using drones (Airbus, 2016), the use of social media to monitor access to all-season roads is likely to become a possibility in the near future.

As part of the ReCAP Satellites project a social media app was investigated for reporting road condition and accessibility issues, as can be seen in Figure 3 (Workman, 2017). This could be enhanced further to allow people to report road blockages and issues with accessibility through social media. In the short term this information would need to be supplemented with additional data, but there is potential for it to be viable in the longer term. Figure 3: Potential social media interface

3.4. Data Scraping

This is a technique whereby human-readable data is extracted from a source, which could be a mobile phone, social media, a computer, the internet, a report, online maps, etc. There is research ongoing in Nairobi, Kenya (Kenya Engineer, 2017), to use data scraping to map Matatu routes and identify where traffic collision blackspots develop by using key works in Google or Facebook social media posts. This process can be a rather unsophisticated way to obtain data and includes a certain amount of assumptions. It is often used as a last resort when other sources of data have been exhausted.

However, there is potential to use scraped data to indicate which roads are all-season, as per the prevailing definition. This is unlikely to be a reliable indicator of all-season status for some time until mobile coverage increases in rural areas and the proportion of people who have smartphones increases, but there is potential.
3.5. Internet of Things

The internet of things (IoT) is the network of physical devices, vehicles, buildings and other items that are embedded with electronics, sensors, software and connectivity to the internet or other networks that enable these objects to collect and exchange data (Figure 4). More recently it has been defined as “A global infrastructure for the Information Society, enabling advanced services by interconnecting (physical and virtual) things based on, existing and evolving, interoperable information and communication technologies” by the International Telecommunication Union (ITU) (ITU, 2012).

The IoT allows objects to be sensed or controlled remotely using existing network infrastructure. This creates opportunities for more direct integration of computer-based systems into the physical world. The IoT can also be augmented with sensors and actuators, which can produce the concepts of smart homes, intelligent transportation and smart cities. Each thing is identifiable uniquely through its embedded computing system, but is able to operate within the existing internet infrastructure.

- Firstly a ‘thing’ needs to have a unique identity. This is at the core of computer functionality, in the same way that every website has to have a unique address so that the computer can determine exactly what it is connecting with.

- Secondly the ‘thing’ has to have a means of communicating, in order for it to be connected to the internet. There is a wide range of different ways to connect, including mobile data, GPS, WiFi, etc.

- Thirdly, the thing that really defines the IoT is that a ‘thing’ has to have a sensor. This can be any number of sensors, from an RFID to a GPS sensor, but it will be able to tell you something about the ‘thing’.

- Lastly there has to be a way to control the ‘thing’. We are seeing more and more how devices are being controlled by smart phones, which is the most common way that IoT things are controlled.

The Data/Information/Knowledge/Wisdom (DIKW) pyramid (Figure 5) is often referred to for an explanation of the hierarchical relationships between data and its various stages, through to wisdom (Rowley, J. 2007). The ultimate goal of IoT is to learn from the data collected, which makes the IoT a candidate for assisting in the measurement of RAI.

Initially data is collected and analysed for its core principles, and then separated into its core principles. This allows questions to be asked of it, such as who, what, where? Patterns are typically formed which transform the information into knowledge, which is essentially an understanding of the data, but this knowledge is subjective to the end goal of the user. Ultimately, wisdom can be gained through recognising relationships between the knowledge centres, which can then be used to shape the future. New data that is
gathered will be based on previous wisdom gained, which helps to move society forward and in theory avoid repeating mistakes of the past.

In terms of the RAI, the IoT would essentially be a network of sensors that can be found in vehicles, buildings, cameras and videos, sat-navs, mobile phones and elsewhere that can be coordinated to provide information on traffic movements and hence theoretically help to measure whether roads are all-season. At present the volume of relevant sensors may be too small to gather enough information to measure RAI, but in the future as more modern vehicles replace older ones, and sensors and technologies develop further, this is likely to become a feasible way to measure RAI.

3.6 Big Data and Data Science

Big data is the term used for data sets that are so large or complex that traditional data applications do not have the capacity to analyse and manage them. There are several aspects of big data; it is not just the volume of data that defines it:

- **Volume** – By its definition, the volume of data collected is large, and organisations can receive data from a wide range of sources including transactions, sensors, communications, social media, etc. Storing this data can be an issue, although systems are being developed to manage this.

- **Speed** – Data can be streamed at very high speeds, but this increases the need to deal with the data in a timely way, potentially in real-time.

- **Diversity** – Data is produced in a wide variety of formats, so in order to get the most from diverse data it is necessary to be able to combine that data in different formats.

- **Inconsistency** – The supply of data can be variable, varying daily, weekly or seasonally and can even be triggered by certain events. This makes it challenging to manage, especially if the data is unstructured.

- **Complexity** – Data comes from many different sources with different complexities, providing a challenge to link it together and transform it. There is a need to manage this challenge to prevent losing control of the data coming in.
In addition to analysis and management of the data, other challenges include capture, data cleaning, search, data sharing, data storage, data transfer, visualisation, querying, updating and ensuring that the data maintains its privacy. The term big data is often used to define how a data set is analysed or how the value is extracted from it, but big data can lead to very accurate results. This in turn could lead to more confident decision making and logically better decisions, which can result in greater operational efficiency, cost reduction and reduced risk.

There are numerous applications that have already been identified for big data, but inevitably very many that are yet to be found. Data sets are growing rapidly, partly because they are increasingly gathered by numerous and cheap information-sensing mobile phones and tablets, remote sensing devices, software logs, cameras, microphones, radio frequency identification (RFID) readers and wireless networks, as well as many other devices. In addition many more organisations are making data available through open source outlets, with governments and aid agencies especially adopting open source policies. The world’s technological per capita capacity to store information has roughly doubled every 40 months since the 1980s; (CGD Report, 2014)

It is claimed that as of 2012, 2.5 Exabytes (2.5×10^18) of data is created every day (M Wall, 2014 / IBM, 2013). The ownership of big data is also a question that has not been sufficiently answered.

In terms of measuring the RAI, big data is a potential source of information. In theory it also covers many of the methods already proposed, which will essentially utilise big data in their process. A diagram of the potential for big data in the transport sector is shown in Figure 6.

Data science is described as a multi-disciplinary field of science that can extract knowledge from both structured and unstructured data (often ‘big data’) by using algorithms, systems, processes and scientific methods. It encompasses statistics, data analysis, machine learning and other processes to achieve this. The Harvard Business Review referred to data science as the sexiest job of the 21st century (Davenport, 2012), which propelled the discipline into the public eye. Many Universities now offer Data Science courses. In terms of RAI data science has the potential to make sense of the many different types and formats of data that would be useful for measurement of RAI, in the population, mapping and road condition areas.

3.7. Machine learning and Artificial Intelligence

Machine learning is a subfield of computer science that provides computers with the ability to learn without being explicitly programmed. It has been around for a long time and initially evolved from the study of pattern recognition and computational learning theory in artificial intelligence. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data.
Machine learning is closely related to (and often overlaps with) computational statistics, which also focuses in prediction-making through the use of computers. It has strong ties to mathematical optimisation, which delivers methods, theory and application domains to the field. Machine learning is sometimes linked with data mining, where the latter focuses more on exploratory data analysis and is often known as unsupervised learning.

The applications of this to roads, and specifically to roads in Africa, is in some aspects quite straightforward. The most obvious use is to map roads from a satellite image, where one would expect that AI could work out the places where roads are obscured by trees or other objects and be able to continue to map the road when it becomes visible again. It would also be able to identify significant changes in a road’s character or width or other aspect and still recognise it as the same road.

A further step for AI would be to assess the condition of a road from satellite imagery. This would require many more variables and the computer would need to learn from its experiences in order to be able to assess the condition accurately. Given the variety of options available from earth observation if AI is developed in this area, it should be possible to make it more accurate than human visual assessment.

3.8. Beyond the horizon

The potential of some ‘beyond the horizon’ technologies are explored in this section. These are technologies that are still under development or have not been used in the road sector before, but that could conceivably be innovated to assist in the measurement of the RAI:

- Pollution or air quality monitoring: There are now ways to measure air pollution remotely, using satellites. The Sentinel-5P spacecraft, which was built in Britain using Dutch technology, was launched last year with the purpose of tracking several of the most damaging contaminants in existence, including sulphur dioxide (SO2), nitrogen dioxide (NO2) and carbon monoxide (CO) (Figure 8).
Today’s vehicles create pollution when they access a road, and in many developing countries vehicles create high levels of pollution due to their age and lack of maintenance. There may come a time when vehicle movements can be tracked by monitoring the pollution they leave behind. Of course to provide any information on all-season status of a road this could have to be monitored on almost a daily basis, and the technology to do this is not yet developed.

- Synthetic Aperture Radar (SAR) satellites: This is an active mode of satellite monitoring, whereby radar waves are dispatched from a satellite and their return is monitored by a sensor. This creates a point cloud that can provide information on materials and terrain. As radar is sensitive to the physical structure of the surface it should be possible to determine the pavement characteristics, such as the roughness (Figure 9). This has not been researched so far due to the high cost of SAR imagery, lower resolution and the relatively fewer number of SAR satellites. The use of SAR to contribute towards monitoring of the RAI may be possible in the long term.
Hyperspectral monitoring: Spectral Reflectance is a remote sensing method that can help to identify different earth features or materials by analysing their spectral reflectance patterns or spectral signatures. The signatures can be visualised in spectral reflectance curves as a function of wavelength. The technology works by measuring reflected or emitted radiation from different bodies. Objects with different surface features reflect or absorb radiation from the sun in different ways, so the reflectance properties of an object will depend on the particular material and chemical state it is in at the time (moisture content), its surface roughness and geometric circumstances (angle of sunlight). The most important features are colour, structure and surface texture. The graph in Figure 10 shows the spectral reflectance for three different materials; vegetation, soil and water (ESA, 2019).

In theory this type of monitoring would be able to identify a road that is saturated or under water and therefore not all-season. It could also identify different materials and the condition of those materials, for example aged asphalt that has a high incidence of cracking (Emery, 2014).

Hyperspectral goes one step further by using a higher spectral resolution with wider land coverage, allowing features beyond the visible and NIR range to be detectable. A paper by Shaw (2003) shows how a space-borne image can simultaneously sample multiple spectral wavebands on the ground over a wide area, and how different details can be determined from the spectral signatures. If the resolution can be improved enough to identify road surfaces, this has potential for RAI monitoring.
Smart materials: The medical profession has started to use smart materials to help us identify various types of diseases and to aid healing. Smart materials, also referred to as intelligent or responsive materials, are designed with one or more properties that can be significantly changed in a controlled fashion by external stimuli, such as stress, temperature, moisture, pH, electric or magnetic fields, light, or chemical compounds. There could be potential for smart materials to be used in roads to allow us to remotely monitor the road condition and accessibility. The technology exists to use steel strands in asphalt to induce self-healing, where the steel is heated by induction and softens the asphalt, allowing it to be re-compact and seal the cracks.

Pseudo satellites: This is a recent development of Unmanned Aerial Vehicles (UAV) or drones, that fly at very high altitude, up to 70,000 ft. They are solar powered and can stay aloft for long periods. An example is shown in Figure 12.

The potential uses of pseudo satellites include high resolution photographs and video, LIDAR, RADAR and Broadband Communications, as well as some military information collecting applications such as Electronic Support Measures (ESM) and Electronic Signals Intelligence (ELINT). They can provide very high resolution imagery of the earth, at NIIRS 6 to 8, with the highest resolution so far recorded at 13cm. They can be airborne for several weeks at a time and have the potential to be cheaper and more accessible than regular satellites. With the higher resolution
and the ability to record video, these satellites have the potential to contribute to RAI measurement above and beyond that of existing passive satellite systems, although they are not fully commercially operational yet.

![Satellite Image](image.jpg)

**Figure 12: Airbus DS pseudo-satellite, Zephyr (Workman, 2017)**

4. CONCLUSIONS

This paper has included some potential technologies with the potential to assist in the measurement of RAI in the future. The technologies that are most likely to provide results in the short term are measurement by satellite imagery and mobile phone data. These are already in use and being research within the road sector, and have the potential to be applied to the RAI without too much further development. Other processes that could possibly be implemented in the medium term are data collection by social media, data scraping and the Internet of Things. Aligned with this is the use of big data, data science and machine learning, which are crosscutting technologies across all of the potential RAI measurement processes presented here. The potential of some ‘beyond the horizon’ technologies have also been explored, including Air quality monitoring, SAR, smart materials and pseudo satellites, which could become feasible in the longer term.
REFERENCES