

SPATIOTEMPORALLY IMPROVING THE RURAL ACCESS INDEX – A REMOTE SENSING BASED APPROACH

M. Halbgewachs^{1*}, A. Schneibel¹, M. Gähler¹

¹ German Remote Sensing Data Center (DFD), German Aerospace Center (DLR),
82234 Wessling, Germany – magdalena.halbgewachs@dlr.de

KEY WORDS: RAI, infrastructure, state capacity, SDG, Landsat, Sentinel, OpenStreetMap, Lake Chad Basin.

ABSTRACT:

Many countries, especially in the global south still lack the ability to effectively pursue basic policies, which can lead, in the worst case, to armed conflicts. Access to markets is a key factor for economic growth and an important component in reducing poverty. The SDG 9.1.1 addresses the proportion of the rural population who live within 2km of an all-season road, which can be mapped by the Rural Access Index (RAI), introduced by the World Bank in 2006. This requires the road network of so-called all-season roads, population distribution and rural areas. We developed a fully automated approach, using remote sensing and other open source data to calculate the RAI on an annual basis between 2013 and 2020 for the Lake Chad region. We achieved an overall accuracy between 97.0% and 97.5% in detecting all-season roads using a Random Forest classification. Our method shows similar results to those published by the World Bank. However, our approach provides a higher spatial and temporal resolution measuring the RAI compared to previous studies and is independent of field studies.

1. INTRODUCTION

Countries, which lack the ability to effectively pursue domestic or international policies are often at risk of becoming what is known as a 'fragile state'. To assess this risk, scientists use the concept of state capacity, a measure that is closely connected to the fragile state concept (Besley and Persson, 2010). It measures the ability of a state to e.g. ensure political stability, charge taxes and provide essential health or educational supplies to its population (Woodward, 2004, Müller-Crepon, 2021).

One of these key factors, access to markets, is crucial to economic growth and an important component in reducing poverty. Especially the rural population is often insufficiently connected to a well-developed infrastructure. The Sustainable Development Goal (SDG) 9.1 also accounts for this factor and includes the development of "[...] reliable, sustainable and resilient infrastructure, including regional and transborder infrastructure, to support economic development and human well-being, with a focus on affordable and equitable access for all" (United Nations, 2022). The connection of rural population to infrastructure has been translated into the Rural Access Index (RAI) by Roberts et al., 2006 and estimates the proportion of the rural population with adequate access to the transport system, thus directly mapping SDG indicator 9.1.1 (World Bank Group, 2016). Precisely, the RAI is defined as the proportion of the rural population living within two kilometers (~ 20-25 minutes walking time) of an all-season road. All-season roads are defined as being navigable throughout the year and are only temporarily impassable due to weather conditions. The initial approach to calculate the RAI was based on household surveys (Roberts et al., 2006). In 2015, the World Bank designed a new GIS-based method to provide more accurate and cost-effective results by using specific datasets, i.e. the population distribution, road infrastructure and the condition of these roads (World Bank Group, 2016). According to the World Bank Group, 2016, roads are all-season roads if they are either paved (with an International Roughness Index (IRI) of less than 6

meters/km or in excellent, good or fair condition) or unpaved (with an IRI of less than 13 meters/km or in excellent or good condition). Such an assessment requires an on-site survey of the roads or most often a subjective assessment based on high-resolution imagery, if available. For the RAI calculation, the World Bank proposes the use of different data sets, which are partly publicly available, others restricted, or, depending on the study area, proprietary. The RAI is made partially available for various years at country level, in some cases Administrative Level 1 (admin-1 level) (Workman and McPherson, 2019, World Bank Group, 2019). Other studies calculated the RAI for only one point in time (Akin and Demirel, 2019, World Bank Group, 2016), a specific subregion (Akin and Demirel, 2019), only partly using remote sensing data (Akin and Demirel, 2019, World Bank Group, 2016, Workman and McPherson, 2019), or in a GIS environment (Workman and McPherson, 2021, World Bank Group, 2016).

Our approach consists of using a standardized objective data basis, including remote sensing data, in order to be able to calculate the RAI consistently across countries with a higher temporal and spatial resolution than provided to date. Only open source data is used as part of our framework, which allows researchers and analysts to gather and analyze large amounts of data without incurring significant costs. The OpenStreetMap (OSM) database is used as basis for the road network. Since in 2013 about 72% and in 2020 still about 40% of the roads within our study area do not contain information about the road surface, a Random Forest classification was carried out on Landsat 8 and Sentinel-2 data for each year to identify paved pixels. For this paper, we focus on an annual calculation of the RAI between 2013 and 2020 in the Lake Chad region, which includes parts of Niger, Chad, Cameroon, and Nigeria. As spatial units we use both, admin-1 regions and PRIO-GRID cells therein, a global vector matrix commonly used in social sciences (Tollefsen et al., 2012). In summary, the study specifically addresses the following research questions: (I) Is the differentiation between paved and unpaved road pixels possible by conducting a land cover classification? (II) Can adequate

* Corresponding author

RAI results be obtained using remote sensing and other open source data? (III) Is it possible to provide a higher temporal and spatial resolution of the RAI than previous studies show?

2. MATERIALS AND METHOD

2.1 Study Area

For this paper, we focus on the Lake Chad region, which covers the cross-border region of Niger, Chad, Cameroon, and Nigeria. A total of eleven admin-1 regions are considered, one in Niger, two in Nigeria, two in Cameroon and six in Chad (Figure 1). All of these countries have been affected by the violent terrorism of Boko Haram, which has built its success on dysfunctional socioeconomic and political gaps (United Nations Development Programme, 2022). The conflict has resulted in more than 2.9 million people being displaced and according to UN OCHA, about 5.6 million people are directly facing food insecurity (UN OCHA, 2022). Insufficient basic social services and the already low availability of natural resources are increasingly deteriorating the situation. Due to years of dominating violence and insecurity, thousands of children lack education, estimations made by UN OCHA account for more than 1,000 closed schools (UN OCHA, 2022). The population derives its income mainly from agriculture and fishing (Masaki and Rodríguez-Castelán, 2021). In some cases, the poverty rate in the region is up to three times higher than in the rest of the neighbouring countries reaching values between 31% and 72% (Masaki and Rodríguez-Castelán, 2021).

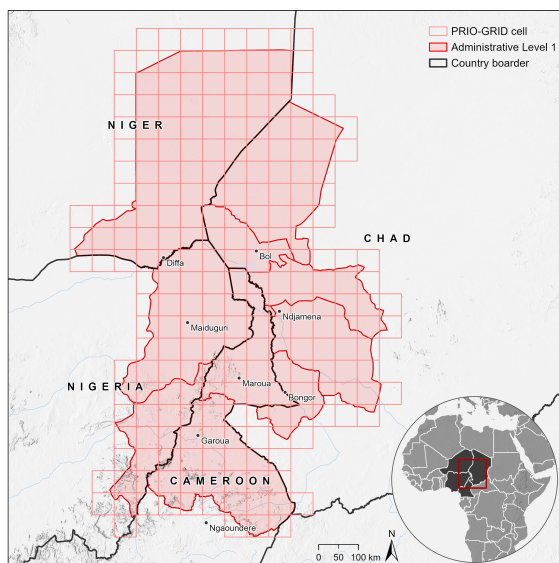


Figure 1. Study Area (data source: Global Administrative Areas, 2012, Tollefsen et al., 2012, OpenStreetMap contributors, 2022; Service Layer Credits: Airbus, USGS, NGA, NASA, CGIAR, NCEAS, NLS, OS, NMA, Geodatastyrelsen, GSA, GSI and the GIS User Community; Esri, HERE, Garmin, FAO, NOAA, USGS)

2.2 Data

Only open source data is used as part of our framework. The gridded population count dataset from WorldPop, available annually since 2000, is taken as population distribution data (WorldPop and CIESIN, 2018). The used mapping approach is a Random Forest-based dasymetric redistribution (Stevens

et al., 2015). This involves matching census-based population counts with the corresponding administrative units and disaggregating them to 3 arc ($\sim 100 \times 100 \text{m}$) grid cells using machine learning techniques that exploit relationships between population densities and a set of spatial covariates (Lloyd et al., 2019). As a second dataset, the Global Human Settlement Layer of the Joint Research Centre (JRC) from 2015 with a spatial resolution of 1km, is used to identify rural regions which is based on satellite and population data (Pesaresi et al., 2016). The dataset from 2015 was chosen because it is the closest to our time period. Only pixels that are not classified as urban centers or urban clusters according to the DEGURBA (Degree of Urbanisation) classification scheme and thus have a population density of less than 300 inhabitants per km^2 are considered in our study (OECD et al., 2021)

The road network is obtained via OpenStreetMap (OSM) and queried for the corresponding, retrospective years through the ohsome API. This API is a data analysis tool, provided by the Heidelberg Institute for Geoinformation Technology (HeiGIT), for working with OSM historical data on a global scale. It allows to query and aggregate the recorded vector data and get insights into the spatial and temporal evolution and associated annotated attributes (Raifer et al., 2019).

Atmospherically corrected surface reflectance from Sentinel-2 MSI and Landsat 8 OLI/TIRS time series data is used for road classification. The data is provided along with a quality assessment (QA) band that indicates the presence of clouds (USGS, 2019, ESA, 2015). In addition, the ESA WorldCover 10m classification and the JRC Global Surface Water Mapping Layer are used to mask water and wetland areas from the satellite mosaics. The ESA WorldCover 10m provides a global land cover classification of 2020, based on Sentinel-1 and Sentinel-2 data. The classification consists of eleven land cover classes, one of which designates permanent water bodies (Zanaga et al., 2021). The JRC Global Surface Water Mapping Layer provides information about spatial and temporal distribution of surface water with a spatial resolution of 30 meters between 1984 and 2021, based on Landsat 5, 7 and 8 data. Both permanent and seasonal water surfaces are recorded (Pekel et al., 2016).

The calculations are performed both on admin-1 level (Global Administrative Areas, 2012), the largest subnational administrative unit of a country, and on PRIO-GRID, which is a global vector matrix of 0.5×0.5 decimal degrees spatial resolution and often used in social science research being insensitive of political boundaries (Tollefsen et al., 2012).

2.3 Methods

2.3.1 Random Forest and Road Classification Within our approach we defined all season roads as roads that are paved to calculate the RAI and determine the condition of the roads. Even though the literature refers to paved roads as all-weather roads rather than all-season roads which is rated to a higher standard (World Bank Group, 2016, Workman and McPherson, 2019), our approach does not rely on time-consuming in situ data that must be collected annually. To acquire this information about the road surface, a yearly Random Forest classification was carried out on Landsat 8 (between 2013 and 2018) and Sentinel-2 (for 2019 and 2020) data, where paved pixels were identified. Figure 2A shows a schematic overview of the workflow.

As first pre-processing step, annual multitemporal composite mosaics are created based on the available imagery. Secondly, cloud masking is performed. The provided QA band contains

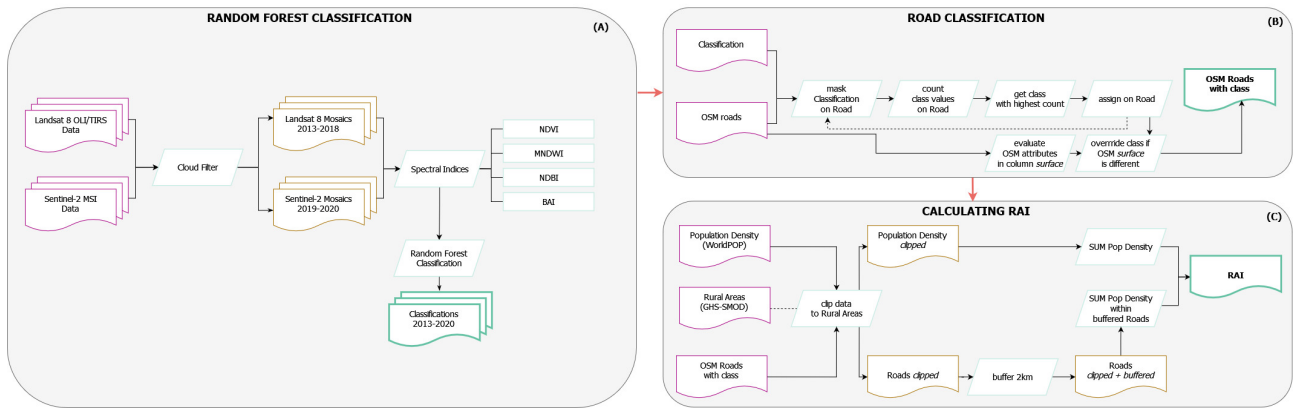


Figure 2. Schematic workflow

bitmasks for clouds and cloud shadows. All pixels containing the corresponding bit value are removed. To avoid misclassification, permanent and seasonal water as well as wetland areas, derived from the ESA WorldCover 10m and Global Surface Water Mapping Layer, are excluded.

In total, 1,648 training samples were collected on the mosaic of the year 2021 for the five classes paved roads, open ground, shrub, vegetation and building. The spectral bands Red, Green, Blue, NIR and SWIR are used for the classification. Additionally, four indices are added to enhance the detection of paved roads and the distinction between the different classes:

The Normalized Difference Vegetation Index (NDVI) is used to highlight the signal from vegetation, which highly reflects in the NIR band (Kamal et al., 2015). The Modified Normalized Difference Water Index (MNDWI) helps to delineate water and artificial surfaces (e.g. asphalt roads). Both surfaces have similar spectral properties and reflect especially in the green wavelength range, the MIR band used improves the distinction between the two (Xu, 2007). The Normalized Difference Built-up Index (NDBI) is particularly useful in detecting built-up areas (Zha et al., 2003). The Burned Area Index (BAI) is additionally used to identify burn scars (Chuvieco et al., 2002). As paved roads also have a lower spectral reflectance in the Red band than vegetation, this index also helps to distinguish asphalt from other land covers (Kamal et al., 2015). During the development of the method it has come to attention that burned areas were often classified as paved roads, which is prevented by incorporating the BAI. The indices are calculated as follows:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

$$MNDWI = \frac{GREEN - MIR}{GREEN + MIR} \quad (2)$$

$$NDBI = \frac{MIR - NIR}{NIR + NIR} \quad (3)$$

$$BAI = \frac{1}{(0.1 - RED)^2 + (0.06 - NIR)^2} \quad (4)$$

After randomizing, the training collection is split into 70% training and 30% testing points. A 10-tree Random Forest classifier is then carried out, resulting in an overall training accuracy of 95.80% and overall testing accuracy of 69.44%.

The classification is used as input for assigning the road surface on the annual OSM road network. A road was classified as being paved by a winner-takes-all approach, i.e. if most of the

road pixels were classified as such. However, if a definite road surface attribute was given by an OSM user (e.g. asphalt, paved, concrete, bitumé, unpaved, sand, earth, dirt, gravel, pebblestone), this was taken, accounting for their expert-based knowledge (Figure 2B).

2.3.2 Calculating the Rural Access Index The calculation of the RAI follows the approach suggested by the World Bank Group, 2019. The RAI represents the percentage of the rural population (RP) living within 2km of an all-season road (RP_{road}):

$$RAI = \frac{RP_{road}}{RP} \quad (5)$$

This is done by determining the sum of the rural population within the Area of Interest (AOI). Subsequently, all roads marked as all-season roads within the rural regions are provided with a buffer of 2km and the population is then summarized within these areas (Figure 2C).

2.3.3 Accuracy Assessment For the accuracy assessment, stratified random sampling was used to generate reference points for 2017 and 2019, evaluating the recorded all-season roads. These two years were chosen to include both Landsat 8 and Sentinel-2 classifications. All-season roads are defined here as those captured by the Random Forest classification and related OSM attributes, as described in chapter 2.3.1. As paved roads represent a rather small proportion in the study site compared to other land surfaces, the stratified random sampling method allows us to increase the proportion of the underrepresented paved class (Olofsson et al., 2014). Our requirement was to have at least 100 samples in each stratum, which could be ensured by a proportional redistribution of the sample size. In total 4,125 samples were distributed randomly throughout the whole study area for the two points in time, representing the strata paved and unpaved. The points were assigned expert-based to be either paved or unpaved using Google Earth high resolution imagery of the respective year.

3. RESULTS

3.1 Rural Access Index

The results are shown in Figure 3 and Figure 5 on admin-1 levels and in Figure 4 on PRIO-GRID cells. On average, the

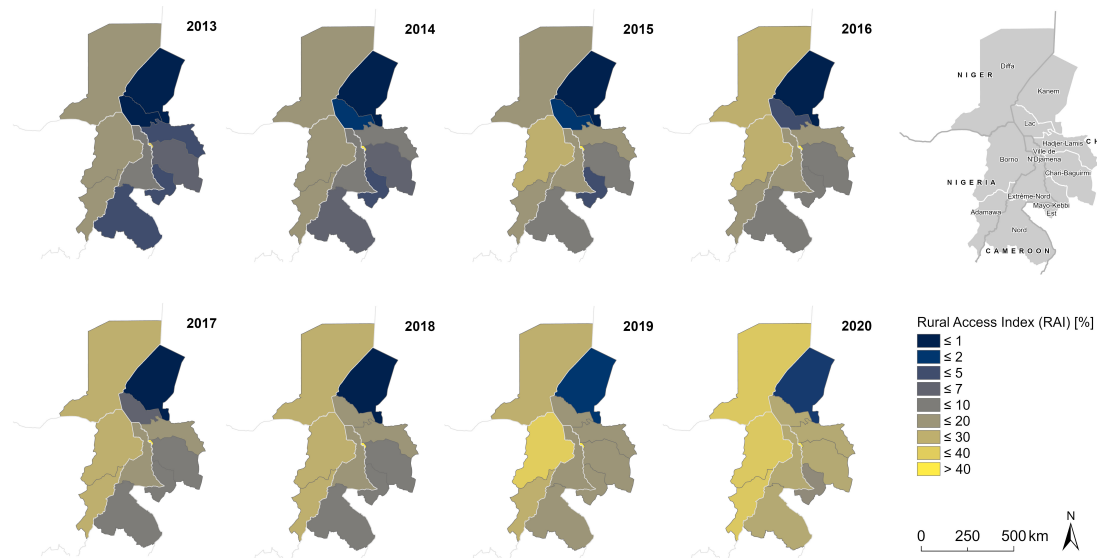


Figure 3. RAI between 2013 and 2020 on admin-1 level (data source: Global Administrative Areas, 2012)

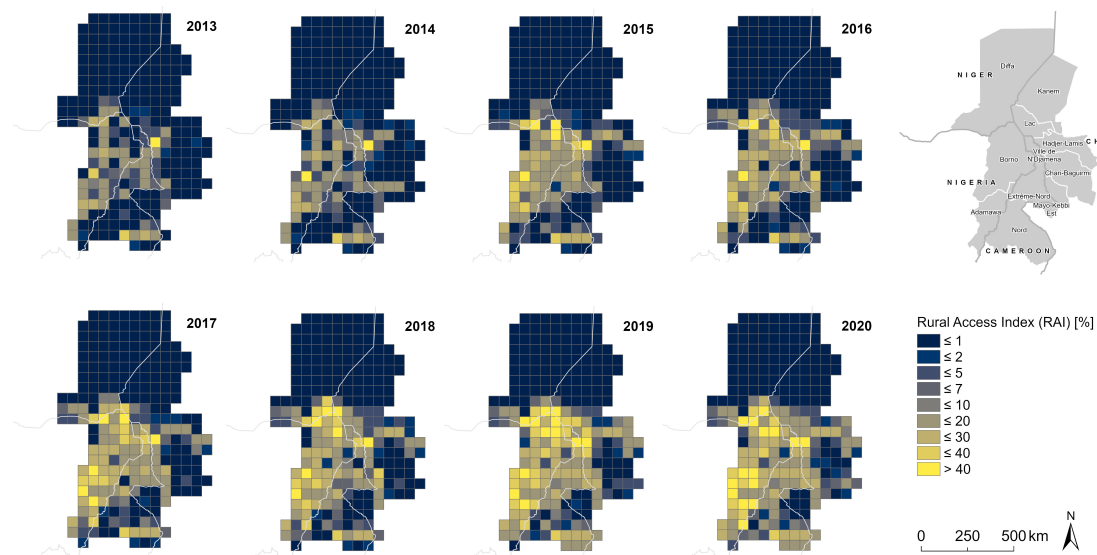


Figure 4. RAI between 2013 and 2020 on PRIO-GRID cell level (data source: Tollefsen et al., 2012, Global Administrative Areas, 2012)

RAI has increased over time and a larger proportion of rural population has gained access to all-season roads. Admin-1 areas with larger towns or cities (such as in Ville de N'Djamena in Chad or Borno in Nigeria) show higher RAI values throughout the whole time series (RAI >10%) than unpopulated regions (RAI <10%). Figure 5 also shows that the RAI in Ville de N'Djamena increased from 35.66% to 81.01% between 2013 and 2020. This admin-1 region has the highest RAI values of all processed admin-1 regions throughout the entire analysis period. This contrasts with the RAI values in Kanem (Chad), where population density is very low and the smallest increase was observed compared to other regions, from 0.00% to 1.53%. This becomes even more noticeable at PRIO-GRID levels where grid cells containing larger settlements (such as Biu in south-western Borno, Nigeria, Guélgendeng in central north Mayo-Kebbi Est, Chad or N'Djamena in Ville de N'Djamena, Chad) stand out clearly (RAI between 28% and 82% over time). The same applies to grid cells holding main road axes to larger nearby cities (such as PRIO-GRID cells in

the southern part of Nord, Cameroon where Ngaoundéré is a nearby city). Similarly, it is notable that higher RAI values at admin-1 level often result from single higher values and a large number of unpopulated areas in PRIO-GRID cells contained therein (e.g. Diffa in Niger). In general, the RAI never exceeds 35% during the analysis period, with the exception of Ville de N'Djamena (Chad).

Statistics can be viewed in more detail in Table 2: A positive trend of 2.33% can be observed for the whole region regarding the rural populations' walkable access to all-season roads. For the entire AOI, the value of RAI is 17.17%. Regions in Cameroon rank the lowest, with a mean value of 11.33%, while nigerian regions rank the highest, with 22.79%. Chad regions show the widest range of RAI values, from 0.00% to 81.01%, with the high values being attributed to Ville de N'Djamena.

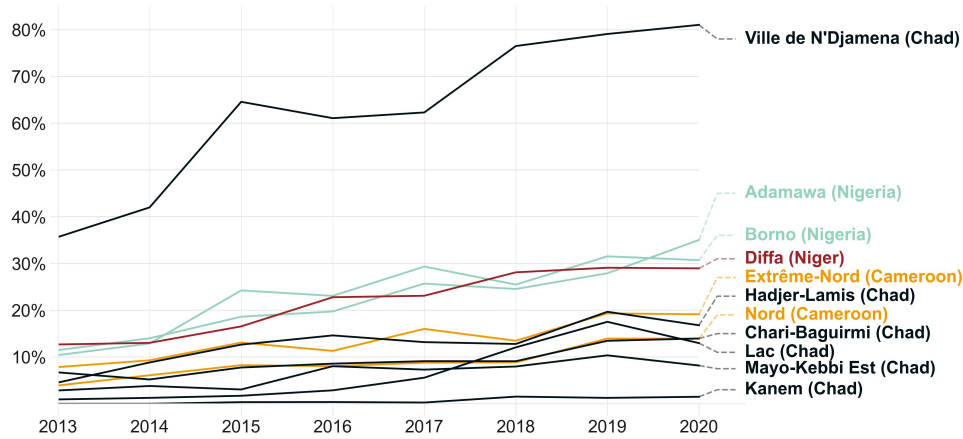


Figure 5. Lineplot of RAI on admin-1 level between 2013 and 2020, colour-coded by country: Chad (blue), Nigeria (turquoise), Niger (red), Cameroon (yellow).

Country (considered regions)	Minimum [%]	Maximum [%]	Mean [%]
Chad (6)	0.00	81.01	16.47
Cameroon (2)	3.91	19.32	11.33
Niger (1)	12.68	29.09	21.79
Nigeria (2)	10.41	35.02	22.79
Study area	0.00	81.01	17.17

Table 2. Country statistics of calculated RAI results.

3.2 Overall Classification Performance

The accuracy assessment in 2017 results in an overall accuracy of 97.0%, and the 2019 verification shows an overall accuracy of 97.5%. The Cohen’s Kappa value is 61.6% for the year 2017 and 66.3% for 2019, with values between 61.0% and 80.0% reflecting high agreement (Landis and Koch, 1977). For the paved class, the precision is 54.0% in both years, which means a slight overestimation on paved roads. The recall shows accuracies of 76.1% and 90.0% for the years 2017 and 2019. The class unpaved roads has a precision of 99.1% and 99.7% for 2017 and 2019 and a recall of 97.7% in both years (see Table 3). By comparing our findings with the World Bank’s results, which have been published for parts of our study area, we find an overall strong agreement. For 2014, we obtain a RAI of 15.5% for the admin-1 level of Borno (Nigeria), and 15.9% for Adamawa (Nigeria), the World Bank values of 14.9% and 12.8% correspond respectively (World Bank Group, 2019).

		2017	2019
Precision	paved	0.540	0.540
	unpaved	0.991	0.997
Recall	paved	0.761	0.900
	unpaved	0.977	0.977
Overall Accuracy		0.970	0.975
Cohen’s Kappa		0.616	0.663

Table 3. Accuracy assessment values for street classifications for 2017 and 2019.

4. DISCUSSION

The main objective of this study was the application of an automated workflow using freely available remote sensing and geo data to calculate the Rural Access Index for admin-1 areas and PRIO-GRID cells for the Lake Chad region between 2013 and 2020. The comparison between the two spatially differing admin-1 level and PRIO-GRID regions has shown that the higher spatial resolution of the PRIO-GRID provides more accurate results with respect to the RAI. Thus, changes can be observed and analyzed more accurately and may serve as valuable input for decision makers.

Based on Landsat 8 and Sentinel-2 mosaics as well as the additional use of four spectral indices over the time period considered, a Random Forest classification was performed to detect paved pixels. Compared to in situ surveys, the remote sensing data used has the advantage of being available free of charge over a long period of time and of allowing an automated classification procedure. The method used is reproducible over different areas of interest, only the training samples would have to be adjusted, depending on the land use of the region. In general, the distinction between paved roads and other land uses (e.g. roofs, water) is difficult because the spectral properties are observed to be often similar (Xu, 2007, Kavzoglu et al., 2009). This could be solved by the initial masking of water areas and bringing in historical OSM road data serving as road database. Due to local wind drifts, asphalt road sections may be covered with sand or vegetation and identified as such. In order to minimize such misclassifications, all OSM roads were assigned the attribute paved, either if the majority of pixels on a road were detected as asphalt or if a corresponding tag was provided in the OSM data itself. Given that OSM is an open user-generated content platform, the quality of the data provided is dependent on the OSM community and their know-how (Neis and Zielstra, 2014). Since the OSM roads serve as data basis for calculating the RAI, a possible mis- or non-editing within the road dataset is inevitable. There can also be large differences in data quality between countries, as well as between rural and urban regions (Neis and Zielstra, 2014). Nevertheless, it is an open source data portal that has seen a strong increase in contributions in recent years (Neis and Zielstra, 2014). It keeps being continuously updated and has been further improved for many years now, thus offering great advantages compared to time-consuming and expensive in situ data collection or surveys, which are often rarely available (Workman and McPherson,

2021). Partial decreases and increases of the RAI within the admin-1 regions can result from different reasons. One factor could be misclassifications resulting from the Random Forest approach, but also due to the used OSM road data. It can happen, for example, that some roads were added incorrectly by OSM users the one year and then deleted in the following year. The RAI calculations depend not only on the input street network, but also on rural areas and population data used. In this study we included rural areas from 2015, derived from the Global Human Settlement Layer of the JRC and annual WorldPop population densities. These data sets represent an approximation to reality by statistical techniques and thus serve as a valuable data basis for the RAI calculations, independent of irregular official government data collections.

Research by the World Bank has shown that on average 57% of the rural population of the International Development Association (IDA) countries in 2003, including the four countries of our study area, have access to the transport network (Roberts et al., 2006). Our results show that the admin-1 regions we analyzed do not fall within this national average, with the exception of Ville de N'Djamena (Chad). These results strengthen the assumptions of the World Development Report that population in poor countries and regions especially have a longer travel time to reach basic services (World Bank Group, 2003). The report also states that children from the poorest 20% of the population from rural areas in Nigeria have to travel more than five times farther to the nearest elementary school than children from the richest 20% and even more than seven times further to the next health facility. In Chad, 80% of the poorest 20% of the population need to travel more than one hour to reach the nearest health facility (World Bank Group, 2003).

5. CONCLUSION AND OUTLOOK

This study analyzes temporal and spatial changes in the rural population in the Lake Chad region between 2013 and 2020 in terms of access to all-season roads. Our approach shows that the calculation of the RAI based on open source data in combination with remote sensing data yields very good results. The availability of time series data from Landsat 8, Sentinel-2 and OSM (in this case road networks) provides a cost-effective and solid basis for detecting asphalt roads using a Random Forest classification. By using the WorldPop population data set, an annual calculation of the RAI is possible. Furthermore, the approach can be fully automatized making it transferrable to other world regions, as long as the ground materials show the same spectral responses as our study area. Otherwise new training samples must be collected. Our approach can effectively contribute to the global calculation of SDG 9.1.1, an important variable for many political, social and natural scientists. The methodology of the RAI can also be extended to other research questions, such as the accessibility of educational institutions or health sites, which allows drawing conclusions about the need for political action.

REFERENCES

Akin, O., Demirel, H., 2019. Measuring Rural Access Index via Spatial Information Science. *2019 9th International Conference on Recent Advances in Space Technologies (RAST)*, IEEE, 461–463.

Besley, T., Persson, T., 2010. State Capacity, Conflict, and Development. *Econometrica*, 78(1), 1–34.

Chuvieco, E., Martín, M. P., Palacios, A., 2002. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, 23(23), 5103–5110.

ESA, 2015. *Sentinel-2 User Handbook*.

Global Administrative Areas, 2012. GADM database of Global Administrative Areas, version 2.0. www.gadm.org.

Kamal, M., Phinn, S., Johansen, K., 2015. Object-Based Approach for Multi-Scale Mangrove Composition Mapping Using Multi-Resolution Image Datasets. *Remote Sensing*, 7(4), 4753–4783.

Kavzoglu, T., Sen, Y. E., Cetin, M., 2009. Mapping urban road infrastructure using remotely sensed images. *International Journal of Remote Sensing*, 30(7), 1759–1769.

Landis, J. R., Koch, G. G., 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159.

Lloyd, C. T., Chamberlain, H., Kerr, D., Yetman, G., Pistoletti, L., Stevens, F. R., Gaughan, A. E., Nieves, J. J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A., Tatem, A. J., 2019. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big earth data*, 3(2), 108–139.

Masaki, T., Rodríguez-Castelán, C., 2021. *Lake Chad Regional Economic Memorandum*. World Bank.

Müller-Crepon, C., 2021. State reach and development in Africa since the 1960s: new data and analysis. *Political Science Research and Methods*, 1–10.

Neis, P., Zielstra, D., 2014. Recent Developments and Future Trends in Volunteered Geographic Information Research: The Case of OpenStreetMap. *Future Internet*, 6(1), 76–106.

OECD, The European Commission, United Nations Human Settlements Programme, The World Bank, 2021. *Applying the Degree of Urbanisation: A Methodological Manual to Define Cities, Towns and Rural Areas for International Comparisons*. Manuals and guidelines / Eurostat, 2021 edition edn, OECD Publishing, Paris.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., Wulder, M. A., 2014. Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, 148, 42–57.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A. S., 2016. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633), 418–422.

Pesaresi, M., Ehrlich, D., Florczyk, A., Freire, S., Julea, A., Kemper, T., Soille, P., Syrris, V., 2016. GHS-BUILT R2015B: GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014).

Raifer, M., Troilo, R., Kowatsch, F., Auer, M., Loos, L., Marx, S., Przybill, K., Fendrich, S., Mocnik, F.-B., Zipf, A., 2019. OSHDB: a framework for spatio-temporal analysis of OpenStreetMap history data. *Open Geospatial Data, Software and Standards*, 4(1).

Roberts, P., KC, S., Rastogi, C., 2006. *Rural Access Index: A Key Development Indicator*. Transport Papers, 10, World Bank, Washington, DC.

Stevens, F. R., Gaughan, A. E., Linard, C., Tatem, A. J., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. *PLoS one*, 10(2), e0107042.

Tollefsen, A. F., Strand, H., Buhaug, H., 2012. PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363–374.

UN OCHA, 2022. Lake Chad Basin: Humanitarian Snapshot: As of 08 December 2022. <https://reliefweb.int/report/chad/lake-chad-basin-humanitarian-snapshot-08-december-2022>.

United Nations, 2022. Goal 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation. <https://sdgs.un.org/goals/goal9>.

United Nations Development Programme, 2022. Conflict Analysis in the Lake Chad Basin 2020-2021: Trends, developments and implications for peace and stability. <https://www.undp.org/africa/publications/conflict-analysis-lake-chad-basin>.

USGS, 2019. *Landsat 8 (L8): Data Users Handbook*. Reston, VA, USA.

Woodward, S. L., 2004. Fragile States: Exploring the Concept. *Peace and Social Justice Meeting*.

Workman, R., McPherson, K., 2019. *Measuring Rural Access Using New Technologies: Supplemental Guidelines*. ReCAP for DFID, London.

Workman, R., McPherson, K., 2021. Measuring rural access for SDG 9.1.1. *Transactions in GIS*, 25(2), 721–734.

World Bank Group, 2003. *World Bank Report: Making services work for poor people*. World development report, 26.2004, Oxford Univ. Press, Oxford.

World Bank Group, 2016. *Measuring Rural Access: Using new technologies*. Washington, DC.

World Bank Group, 2019. *Measuring Rural Access : Update 2017/18*. Washington, DC.

WorldPop, CIRESIN, 2018. Global High Resolution Population Denominators Project: Global 100m Population. <https://hub.worldpop.org/geodata/listing?id=29>.

Xu, H., 2007. Extraction of Urban Built-up Land Features from Landsat Imagery Using a Thematic-oriented Index Combination Technique. *Photogrammetric Engineering & Remote Sensing*, 73(12), 1381–1391.

Zanaga, D., van de Kerchove, R., de Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., Cartus, O., Santoro, M., Fritz, S., Georgieva, I., Lesiv, M., Carter, S., Herold, M., Li, L., Tsendbazar, N.-E., Ramoino, F., Arino, O., 2021. ESA WorldCover 10 m 2020 v100.

Zha, Y., Gao, J., Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3), 583–594.