

## MULTIPLE ENDMEMBER SPECTRAL MIXTURE ANALYSIS OF DESIS IMAGE TO IDENTIFY ROOFTOPS IN KIGALI.

J. G. Matabishi<sup>1</sup>, A. Braun<sup>2</sup>, G. Warth<sup>2</sup>.

<sup>1</sup> Julius-Maximilians Universität Würzburg, EAGLE Master program, Oswald-Külpe-Weg 86 97074 Würzburg, Germany – [joy-giovanni.matabishi@stud-mail.uni-wuerzburg.de](mailto:joy-giovanni.matabishi@stud-mail.uni-wuerzburg.de)

<sup>2</sup> Eberhard Karls Universität Tübingen, Department of Geography, Rümelinstr. 19-23, D-72070 Tübingen, Germany – [an.braun@uni-tuebingen.de](mailto:an.braun@uni-tuebingen.de)

<sup>2</sup> Eberhard Karls Universität Tübingen, Department of Geography, Rümelinstr. 19-23, D-72070 Tübingen, Germany – [gebhard.warth@uni-tuebingen.de](mailto:gebhard.warth@uni-tuebingen.de)

### Commission III, WG III/4

**KEY WORDS:** Mesma, DESIS, Remote Sensing, Hyperspectral, Urban planning, Kigali, roof materials.

### ABSTRACT:

The development and increase of multi and hyperspectral sensors in the recent years have significantly improved urban structure analysis and interpretation. The current study is the first to investigate the potential of DESIS hyperspectral images for the detection or identification of urban roof materials. After field campaigns in 2014, 2015 and 2018 to collect ground truth points and rooftops radiometric properties; a linear spectral mixture, implemented using a non-negative least squares (NNLS) regression based on the sequential coordinate-wise algorithm (SCA) was applied on a DESIS image from 2020 of Kigali city to identify the different rooftops material and color. Although results show that most endmembers were predicted with a very low probability, the study proved that the combination of spectral mixture and hyperspectral data such as DESIS have great potential in the detection of rooftops material. The presented study also highlighted a number of challenges resulting from the choice of spectral mixture algorithm and colinearity between materials.

### 1. INTRODUCTION

Human activities, since their beginnings, have left an undeniable footprint on the World; Especially cities are among the most noticeable activities (UN, 2014). Today, almost 60 percent of the world's population live in urban areas, a proportion that is expected to increase to 66 by 2050. This development underlines the need of sustainable urban planning which prepares cities for challenges related to overpopulation, the supply of infrastructures, teaching and medical services, and the creation of housing space (UN, 2014).

On the other hand, the development of cities is determined by economic arrangements, social relations and division, legal constructions, political systems and the history of all these interdependent processes (Taubenböck, 2018).

(Pereira and Oliveira, 2014) showed that these spatial and social patterns allow to describe networks, built space and empty spaces in shape-related, topological and hierarchical terms in two, three and four dimensions. Earth Observation provides us with spatial and spectral characteristics of these urban structures (Taubenböck et al., 2018).

Building materials are an indicator of the socio-economic status but satellite images of very high-resolution (VHR) are needed to achieve the desired quality and detail (Ye et al., 2017). Yet, most VHR imaging satellites only have limited numbers of spectral bands, making the identification of materials impossible. As an alternative, hyperspectral satellites operate with a significantly higher number of bands, but at the cost of spatial resolution.

Therefore, single pixels are characterized by presence of various materials (van der Linden et al., 2018).

Spectral mixture models of hyperspectral data demonstrated their facility and ability to display various surface reflectance types as simple combinations of endmember abundance within an image pixel (Small, 2003).

For a long time hyperspectral data was only used for the detection of minerals than urban surfaces (Chisense, 2012), but their potential in urban areas is to be investigated.

Among the first, (Roessner et al., 2001) tested and presented the potentials of airborne hyperspectral data for detailed classification of urban surface cover types.

(van der Linden et al., 2018) reviewed the progress and increase in number of spaceborne hyperspectral sensors which fostered the global transferability of approaches due to standardization of observation modes and analytical techniques.

The DESIS sensor with 235 spectral bands within a range of 400 to 1000 nanometers (nm) offers best condition to collect data in urban areas as most carbonaceous materials observed in roof materials have distinct spectral characteristics at 514, 633 and 780 nm (Carmona, 2019).

This study is the first to investigate the potential of DESIS images for the detection of urban roof materials. The aim of this study is to demonstrate the capabilities of DESIS data for the spectral unmixing of urban surfaces. We see this as a first step towards

the generation of a locally adapted spectral library of rooftop materials which can be applied to newly acquired images for a steady and standardized monitoring (Braun et al., 2019).

Furthermore, the characteristics of different surface types are investigated regarding their separability and detectability in the image.

## 2. DATA AND METHODS

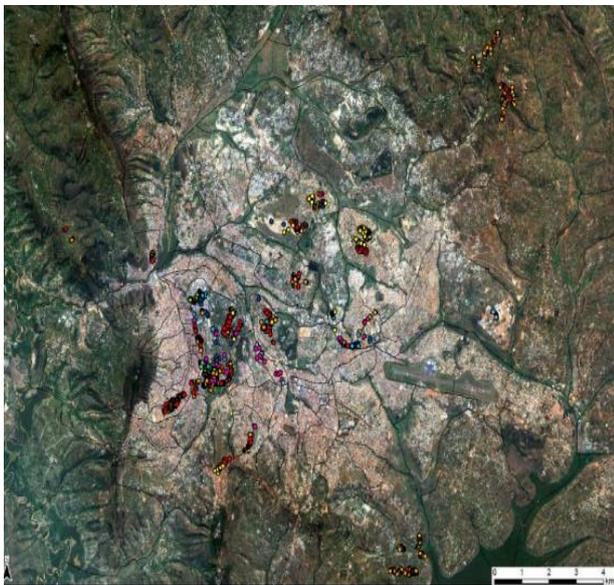
### 2.1 Study Area

Kigali is the largest city and the capital of Rwanda, as well as the country's most important business centre. It is also the main port of entry with a population estimated at 1.135 million residing in an area of 730 km<sup>2</sup> (Rwanda, 2012) with a population growth rate of 3.4 %.

Geographically, Kigali is located in the Central-East African region slightly in South of Equator at 1°56'38"S and 30°3'34"E.

### 2.2 Data

**2.2.1 Building References:** Between November 2014 and October 2015, 1053 reference datasets on building were collected within several field campaigns which included information on GPS coordinates, building type, height, usage, roof type and material, roof color, roof orientation together with information on the neighbourhood of each buildings (Braun et al., 2019). A clustered spatial sampling was conducted in advance to identify areas which represent the full range of socio-economic conditions within the city. Figure 1 displays the distribution of the collected building references.



**Figure 1:** Overview on the city of Kigali. The illustrate the spatial distribution of the collected reference data (1,053 samples).

**2.2.2 Rooftop Spectral Reflectance:** Additional field campaigns were carried out in July 2018. Rooftop samples (n=900) were collected from the main suppliers in the city of Kigali: *Ruliba clay*, *Rwanda Mastersteel Ltd* and *Safintra*. These samples (clay tiles, concrete roofs, plain metal, coated metal and corrugated

metal (clear and rusty)) were analysed under a field spectrometer (Spectral Evolution SR-3500) to retrieve radiometric reflectance values between 350 and 2500 nm at intervals of 1nm. Furthermore, natural artificial surfaces, such as asphalt, vegetation and bare soil were sampled (n = 200) (Braun et al., 2019). A list of materials used for this study is shown in Table 1. A total number of n=x individual samples was used in this study.

**Table 1:** Table of Roof Material names and code

Roof Material	Code name
Clay	X02_clay
Metal Orange	X03_metal_orange
Metal Red	X04_metal_red
Metal Dark Red	X05_metal_dark_red
Metal Blue	X07_metal_blue
Metal Red Rubber	X14_metal_red_rubber
Metal Orange Rubber	X15_metal_orange_rubber
Corrugated	X10_corrugated_1
	X10_corrugated_2
Corrugated Rusty	X11_corrugated_rusty_1
	X11_corrugate_rusty_2
Vegetation (Hasel)	X17_hasel_1
Metal bright	X30_metal_bright

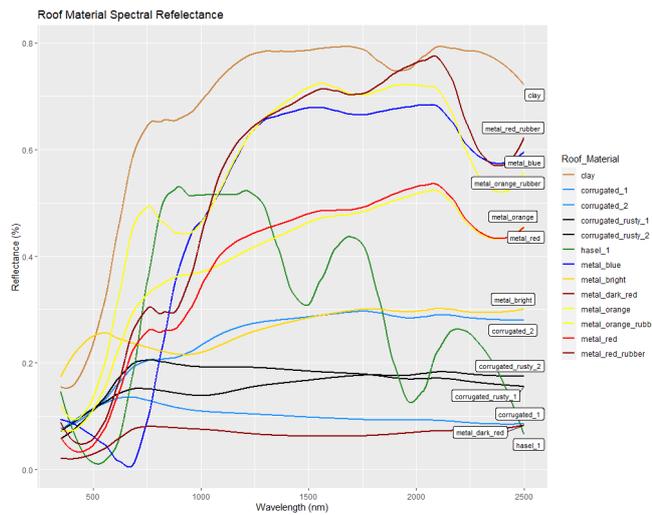
A first analysis of class consistency showed that corrugated materials are not a homogenous surface because of the many stages of material deterioration. To reduce the impact of single measurements, each material was repeatedly observed by a minimum of 4 measurements which were later aggregated to two unseparable classes (see Table 1) using the pavo package in R (Maia et al., 2013). The results are shown in Figure 4. Further details and the similarity of different measurements are presented in Figure 5.



**Figure 2:** 10 Corrugated Metal, partly slight rusty



**Figure 3:** 11 Corrugated Metal, rusty condition.



**Figure 4:** Roof Material Spectral Reflectance in nanometres.

### 2.3 Methods

**2.3.1 Pre-processing:** (Keshava, 2003) describes spectral unmixing as the procedure by which the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or endmembers, and a set of corresponding fractions, or

abundances, that indicate the proportion of each endmember present in the pixel.

Multiple Endmember Spectral Mixture Analysis (*mesma*) is a widely applied tool to retrieve spatially explicit information on urban materials and land cover from both hyperspectral and multispectral data, because it shows strengths in processing and analysing larger number of materials (Roberts et al., 1998).

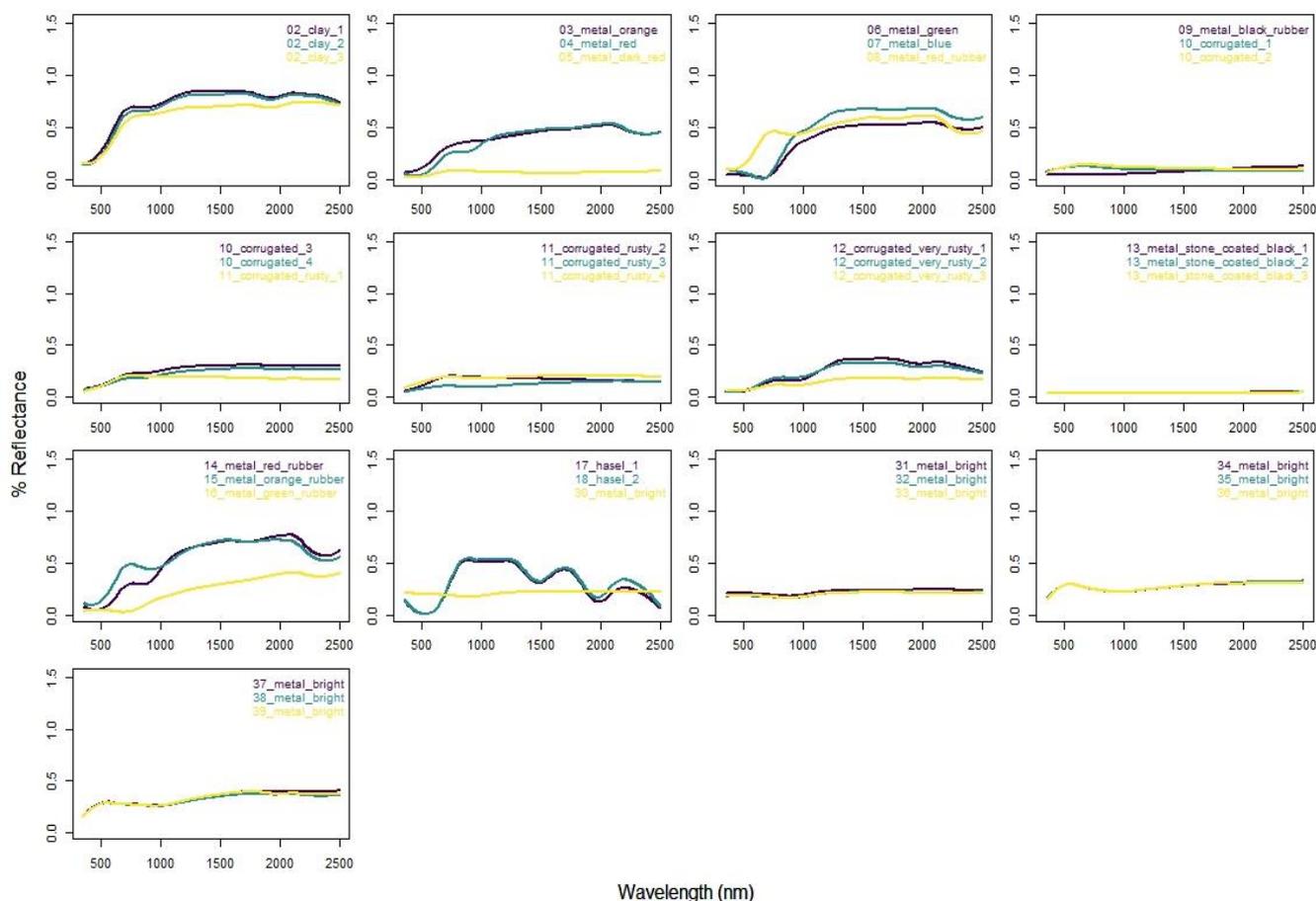
For this project all processing was done in R, with three main packages namely: *hdsar* (Lehnert et al., 2018) and *pavo* (Maia et al., 2013) for pre-processing, *RStoolbox* (Leutner, 2019) with the “*mesma*” library was used to compute the spectral unmixing.

“Pavo” mainly provides functions which allow organizing spectral data by importing raw files, visualize, explore and process their spectral content (Maia et al., 2013). We used it to compare the signatures with respect to their separability and to merge spectra to a joint class signal.

Our raw spectral files were stored in self extraction directive (.sed) files and we used the “*getspec*” function to import all samples. Using “*prospec*” we applied a normalization or smoothing algorithm to the spectra for further analysis.

Plotting all spectral curves together is possible with the “*aggplot*” function but the high number of samples did not allow any detailed visual analysis and observation. We therefore used the “*explorespec*” function which plots groups of curves for a better comparability of spectra. The number of groups and number of curves in one group is selected manually. Figure 5 shows the grouping of spectral curves. As shown, the first three clay samples barely differ, the metal coated black have close to null reflectance and the bright metals separated into three groups because of the high within-class heterogeneity.

This step is crucial as it serves as a guide to the subsequent step of aggregating similar spectra and reducing the number of input for the final *mesma* computation. For this step we used the “*aggspec*” function which combines spectra (by taking the average) according to an index or a vector of identity.



**Figure 5:** Spectral curves in groups of 3

Similarly, the “*hsdar*” package (Lehnert et al., 2018) is also used for pre-processing and processing of hyperspectral data. In this study, we mainly used the “*spectralresampling*” function to resample the spectral resolution of the field spectrometer (1 nm) to the bandwidths of the satellite sensor (DESI), for this is a prerequisite for the *mesma* function.

The function has a built-in list of characteristics of already implemented sensors such as Quickbird, RapidEye and World-View 2. Because DESIS was not within these built-in sensor list, its bandwidth characteristics are passed to the algorithm using a data frame with two columns: first column with lower bounds of channels and second column with upper bounds.

Using the minimum of 402 nm and maximum of 1000 nm with a spectral resolution of 2.55 nm we created the data frame with one column of lower limits and another for upper limits of bands and resampled the ground data to match with the DESIS spectral range.

**2.3.2. Mesma:** “*RStoolbox*” is an R package providing a wide range of tools for remote sensing operations. The available toolset covers the entire image processing workflow, from data import, pre-processing, data analysis, image classification to graphical display (Leutner).

The “*mesma*” library of the “*RStoolbox*” is implemented using a non-negative least squares (NNLS) regression based on the sequential coordinate-wise algorithm (SCA).

The SCA produces a sequence of solutions  $x^{(0)}, x^{(1)}, \dots, x^{(i)}$  which converges to the optimal  $x^*$ . The idea is to optimize in each iteration with respect to a single coordinate while the remaining coordinates are fixed. The optimization with respect to a single coordinate has an analytical solution, thus it can be computed efficiently (Vojtěch et al., 2005).

The NNLS is a statistical approach to fit model parameters to data, assuming that the model parameters are always expressed linearly to those not expressed by the model and that the model parameters can never be negative (Schwalb-Willmann).

This approach of fitting model parameters to data and forcing parameters to be non-negative is helpful for cases where negative values do not correspond to the physical reality, for example when talking about the abundance of certain materials which cannot be below zero (Vojtěch et al., 2005).

To evade from physical absurdity and maintain inherent characteristics such as pixel intensities and chemical concentration, it makes sense to enforce the nonnegativity (Chen and Plemmons). As described below “*mesma*” performs a multiple endmember spectral mixture analysis on a multiband raster image.

The following is an example of the execution of `mesma` within R including its relevant parameters.

**mesma (img, em, method = “NNLS”, iterate = 400, tolerance = 1e-08, ..., verbose)**

- **img** is a RasterBrick or RasterStack, in our case DESIS image.
- **em** is a matrix or data.frame with spectral endmembers. Rows represent a single endmember of a class, columns represent the spectral bands (i.e. columns correspond to number of bands in **img**) Number of rows needs to be > 1.
- **Method [Character]** is the selected unmixing method. In our case “NNLS”
- **Iterate [integer]** this sets maximum iteration per pixel. Processing time could increase the more iterations are made possible. Default is 400.
- **Tolerance [Numeric]** the tolerance limit represents a nearly zero minimal number. Default is 1e-08.
- ... further arguments passed to `writeRaster`
- **Verbose [Logical]** this prints progress messages during execution.

The output is a RasterBrick which contains one band per endmember, with each value representing the estimated presence probability of the endmember per pixel (0 to 1), and a Root Mean Square Error (RMSE) band.

The **iterate** and **tolerance** settings must be considered with precautions as they might affect the sum of estimated probabilities per pixel i.e. the sum varies around 1.

The Mean Square Error (MSE) is used to evaluate the performance in most regression algorithms; but since it has an order of 2 and our dataset and prediction are of order 1, RMSE is the best option to allow direct correlation of data with the error (Kathuria, 2019).

Since its introduction by C.F. Gauss, the RMSE has been proven to be a mathematically more tractable definition to derive the principle of least squares (Ozawa, 2019). Within our analysis, the RMSE determines the difference between values predicted by the model and the values observed.

$$L = \sqrt{\frac{1}{N} [\sum (\hat{Y} - Y)^2]}$$

Where the summation goes from  $n=1$  to  $N$ , and  $N$  is the number of instances in the dataset.

### 3. RESULTS

#### 3.1 Spatial Results

The spatial results of the multiple endmember spectral mixture analysis are shown in Figure 7. Each map shows the abundance probability of an endmember, representing the respective roof material. The extent of the city, as well as its internal differentiation is clearly reflected by the different maps, especially the abundance of clay, as well as metal roofs (orange and corrugated) are strongly related to the different parts of the city. Also, vegetation ( $X_{17}$ ) clearly dominates outside the city, as expected. However, the images also indicate that some

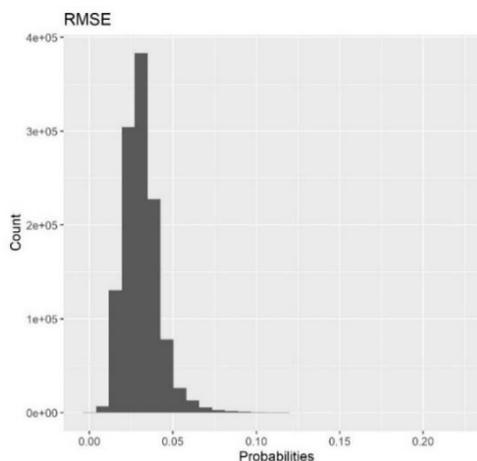
materials are not fully represented by the actual occurrence of building types, but rather biased by other surfaces which were not part of the reference spectra, such as open soils or concrete. This leads to partially overestimations, as observed for blue metals and rusty corrugated metals.

It is also worth to mention that each abundance map is independent and the sum of abundances should not necessarily be equal to 100% per pixel. When we overlaid these layers with a satellite basemap like in the example provided in Figure 7, we realized the corrugated material ( $X_{11\_2}$ ) was overfitting and mapped the entirety of built-up areas. In contrast ( $X_{10\_2}$ ) displayed a more realistic approach with high abundance in areas known for informal settlements; Errors appeared in residential (villa) areas with metal red or metal dark red being considered as corrugated

#### 3.2 Error assessment

Figure 6 shows the overall RMSE calculated based on all 13 abundance maps produced Figure 8. The combination of hyperspectral data with field sample data within the `mesma` approach allowed a clear visual separability of the spectral characteristics of the materials.

The Root Mean Square Error (Figure 6) which measures the difference between values predicted by a model and the values observed, displayed an overall high distribution of errors between probabilities ranging from 0 to 0.05. Accordingly, the predicted probability of a roof material to occur within a pixel has an error margin of around  $\pm 3\%$ .



**Figure 6:** Overall RMSE

#### 3.3 Validation

**3.3.1 Qualitative validation:** For a qualitative validation we decided to overlay the abundance maps with VHR imagery to test for their validity and plausibility.

Figure 8 shows three selected examples of abundance maps at the local scale. They demonstrate that, despite the comparably coarse spatial resolution of the DESIS sensor (30 m), the results can be used to assess roof materials in urban areas, but can contain certain challenges

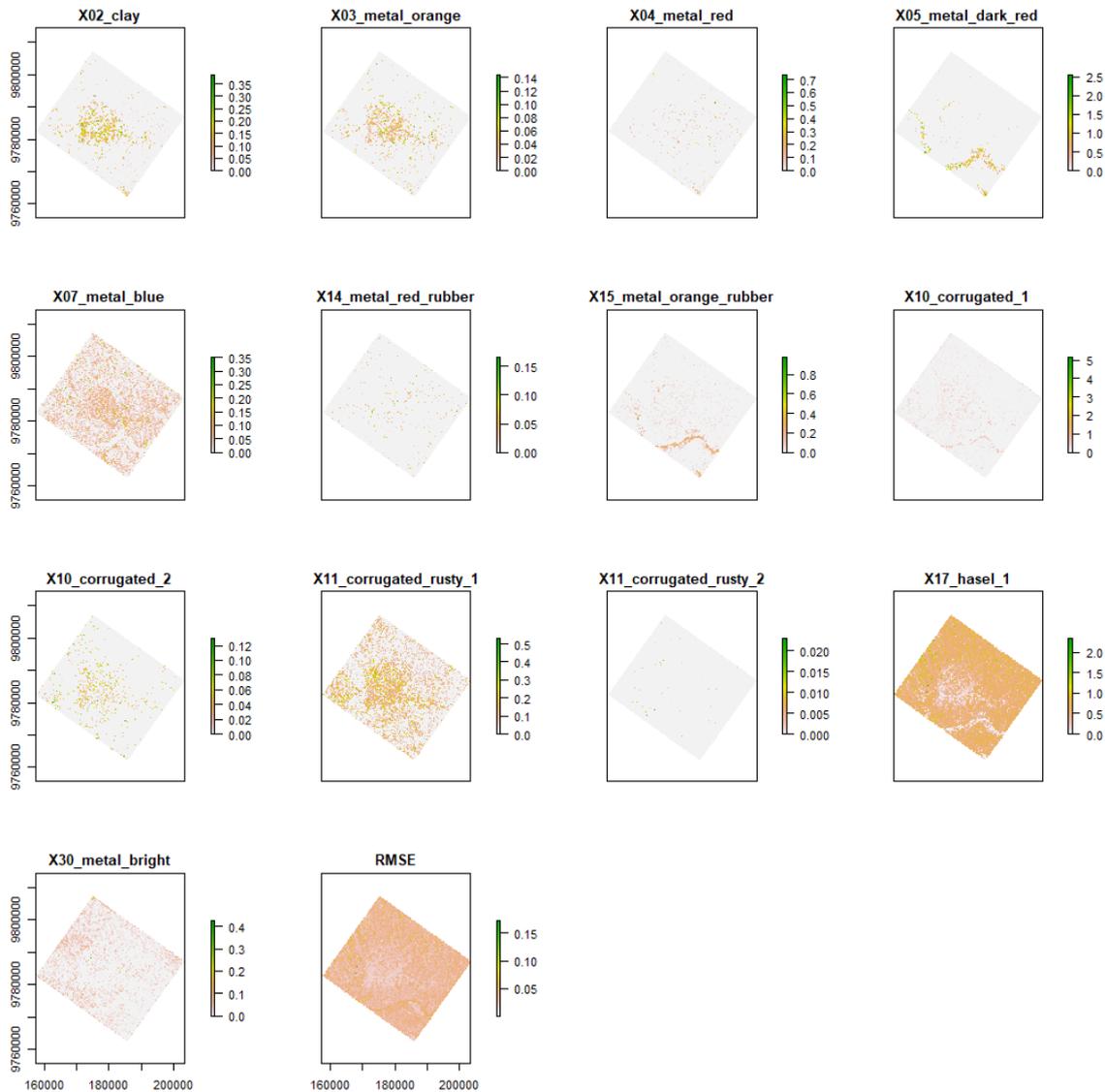
The first example of clay (top) shows the challenges related to spectral unmixing in the city: While the clay roofs of the luxury apartments in the Nyarugenge upgrading areas are clearly visible in the upper centre of the map, the bare areas in the right part are also indicating abundance of clay, but compared to the image we realized this was the Kigali golf course being renovated (Humphreys, 2020).

The second example (middle) shows the abundance of bright metals. These displayed higher abundance only at large buildings especially factories which are located in the Kigali special economic zone. Smaller houses showed moderate abundance but

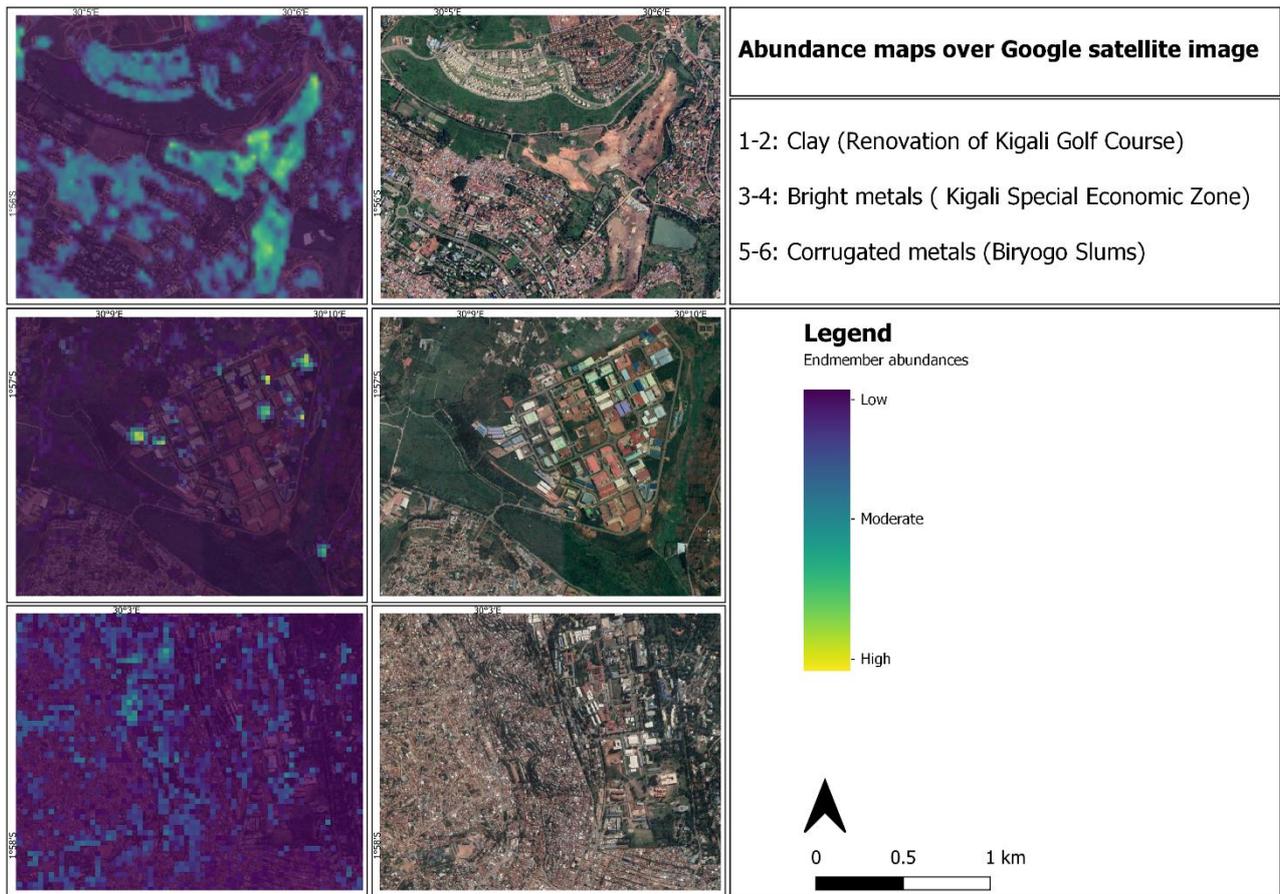
were one of the highlights of the findings as most of these houses are really small and scattered.

Finally, corrugated material demonstrated high abundance in the Biryogo slum very close to the Kigali CBD and made up of old and informal settlements.

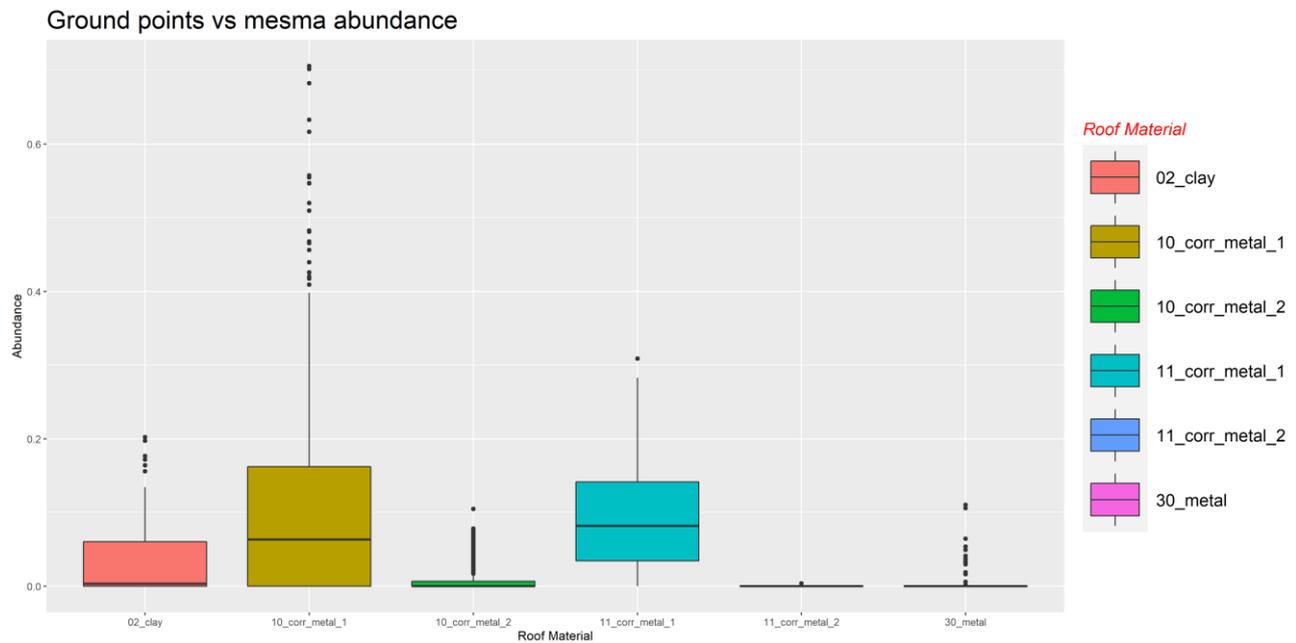
**3.3.2 Quantitative validation :** Using the 1053 ground truth points from the field surveys introduced in section 2.1.1, we extracted raster values from the *mesma* abundance maps to produce boxplots of DESIS abundance maps against the collected data in Kigali. These are shown in Figure 9, highlighting the size of the ground points and the abundance pixel values



**Figure 7:** mesma plots for all the materials and RMSE; Values in the mesma function vary between 0-1.



**Figure 8:** Abundance maps over Google satellite images



**Figure 9:** Comparison of clay, bright metals and corrugated metals with abundance values

From 1053 ground reference points, 99 were clay, 60 were bright metals and 722 were corrugated metal.

Of the 99 reference buildings with clay roofs, 64% were confirmed as such based on their pixel's probability predicted from the DESIS data. Other clay rich areas in Kigali such as open quarries, riverbanks and construction grounds have also shown a high abundance of clay.

Only 9.6% of the pixels which contained the 60 reference buildings with bright metal roofs showed high abundances of bright metals in the *mesma* output

These 9.6% as displayed in figure 9, are slightly greater than zero. This is largely because the small size of the buildings with metal roofs which hold small shares within the pixels of 30 by 30 meters.

With 722 points representing corrugated metal, the first sample of corrugated metal (10) had 66.62 and 30.74 % respectively, while the second sample of corrugated rusty metal (11) had 0.13 and 94.73% respectively.

#### 4. DISCUSSION

The Current study has shown that hyperspectral data from DESIS can be used to identify some not all rooftops material. The study proves the benefits of including context information to reduce misinterpretation of detected materials.

As *mesma* is used to detect materials, the likelihood of confusing materials of similar origin, such as clay rich soil and clay roof tiles, is very high. Therefore, we added building footprint data to successfully reduce misinterpretation of the results. Future projects should consider applying masks for areas such as open land, clay quarries, riverbanks etc, to have the endmember probability prediction limited to built-up areas only.

The *mesma* function from the RStoolbox was faulty due to unprecise iterate and tolerance values with values slightly above designed range of 0-1. We followed up and informed the creator of the function and he has promised to update the function. Meanwhile we recommend future projects to try other available platforms such as the Viper tools software package designed for Envi but can since 2017 be applied in QGIS as a plugin.

Most endmembers were predicted with a very low probability and it is within the same range of 0-0.05 that the RMSE displayed a high difference between the observed and predicted values. Further studies should therefore focus on the required minimum building size and spatial share within the spatial resolution of an image pixel to be reliably predicted as such. This could be targeted by a more comprehensive field survey on roof materials on few selected focus areas.

In the beginning of the project, we observed a high collinearity between some roof materials which affected the model. With the example of metal green which covered a huge part of vegetation. Therefore, more investigation is needed on how the selection and number of classes affects the later results. More statistically robust methods have to be applied to determine the ideal definition and composition of endmembers.

We, therefore, highly recommend performing a collinearity test between the materials in the data preparation phase, to select the

less collineated variables. We also encourage future projects to compare abundance maps and socio-economic data to evaluate the possibility of patterns or relationship between roof material and socio-economic data.

#### ACKNOWLEDGMENTS

I thank A. Braun, G. Warth and the entire Department of Geography at the University of Tübingen chaired by Prof. Dr. Volker Hochschild, for the opportunity to conduct my internship on this project and the valuable inputs and contributions towards the completion of this article.

#### REFERENCES

Braun A, Warth G, Bachofer F, et al., 2019. Identification of roofmaterials in high-resolution multispectral images for urban planning and monitoring. IEEE.

Carmona E, 2019. DESIS Product Specification. Reportno. Report Number|, Date. Place Published|: Institution|.

Chen D and Plemmons JR Nonnegativity constraints in numerical analysis. 31.

Chisense C, 2012. Classification of roof materials using hyperspectral data. *ISPRS -International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXIX-B7*.

Humphreys R, 2020. *Gary Player Design completely revamps Kigali course to 18 holes*. Available at: <https://www.golfcoursearchitecture.net/content/gary-player-design-completely-revamps-kigali-course-to-18-holes> (accessed 21/12/2021).

Kathuria C, 2019. *Regression - Why Mean Square Error?* Available at: <https://towardsdatascience.com/https-medium-com-chayankathuria-regression-why-mean-square-error-a8cad2a1c96f>.

Keshava N, 2003. A survey of spectral unmixing algorithms.

Lehnert W, Lukas, Meyer H, Obermeier A, Wolfgang, et al., 2018. Hyperspectral Data Analysis in R\_the hsdar Package. *Journal of Statistical Software VV(II)*.

Leutner B, *RStoolbox: Tools for Remote Sensing Data Analysis in R*. Available at: <https://bleutner.github.io/RStoolbox/blog.html>.

Leutner B, 2019, Package "RStoolbox". Reportno. Report Number|, Date. Place Published|: Institution|.

Maia R, Eliason CM, Bitton P-P, et al. (2013) pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210x.12069. n/a-n/a.

Ozawa M, 2019. Soundness and completeness of quantum root-mean-square errors. *npj Quantum Information* 5(1).

Pereira A and Oliveira CP, 2014. SUSTAINABILITY OF URBAN FORM -

Application of Salat's Model.

Roberts DA, Gardner M, Church R, et al., 1998. Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment* 65(3): 267-279.

Roessner S, Segl K, Heiden U, et al., 2001. Automated Differentiation of Urban Surfaces Based on Airborne Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 39(7): 8.

Rwanda NIoSo, 2012. Fourth Population and Housing Census - 2012.

Schwalb-Willmann J, 2017. *Spectral unmixing in R using RStoolbox*. Available at: <https://jakob.schwalb-willmann.de/blog/spectral-unmixing-using-rstoolbox/> (accessed 15-12).

Small C, 2003. High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment* 88(1-2): 170-186.

Taubenböck H, 2018. *Remote Sensing for the Analysis of the Global Urbanization*. Julius-Maximilians University of Wuerzburg, Wuerzburg, Germany.

Taubenböck H, Kraff NJ and Wurm M, 2018. The morphology of the Arrival City - A global categorization based on literature surveys and remotely sensed data. *Applied Geography* 92: 150-167.

UN, 2014. 2014 revision of the World Urbanization Prospects.

van der Linden S, Okujeni A, Canters F, et al. (2018) Imaging Spectroscopy of Urban Environments. *Surveys in Geophysics* 40(3): 471-488.

Vojtěch F, Václav H and Mirko N, 2005. Sequential Coordinate-wise Algorithm for the Non-negative Least Squares Problem. 8.

Ye C-m, Cui P, Pirasteh S, et al., 2017. Experimental approach for identifying building surface materials based on hyperspectral remote sensing imagery. *Journal of Zhejiang University-SCIENCE A* 18(12): 984-990.