UNIVERSITY *of* York

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

WILEY   ECOLOGICAL APPLICATIONS

Integrating field and satellite data for spatially-explicit inference on the density of threatened

arboreal primates

Journal:	*Ecological Applications*
Manuscript ID	EAP16-0224.R1
Wiley - Manuscript type:	Articles
Date Submitted by the Author:	n/a
Complete List of Authors:	Cavada, Nathalie; Universita degli Studi di Trento, DICAM; MUSE - Museo delle scienze, Tropical biodiversity section Ciolli, Marco; Universita degli Studi di Trento, DICAM Rocchini, Duccio; Fondazione Edmund Mach, Research and innovation centre, Biodiversity and Molecular Ecology Department Barelli, Claudia; MUSE - Museo delle scienze, Tropical biodiversity section Marshall, Andrew; University of York, CIRCLE, Environment Department; Flamingo Land Ltd. Rovero, Francesco; MUSE - Museo delle Scienze, Tropical biodiversity section
Substantive Area:	Statistics and Modeling < Theory < Substantive Area, Spatial Statistics and Spatial Modeling < Statistics and Modeling < Theory < Substantive Area, Endangered Species < Management < Substantive Area, Reserves/Protected Areas < Management < Substantive Area, Remote Sensing < Methodology < Substantive Area, Conservation < Landscape < Substantive Area
Organism:	Mammals < Vertebrates < Animals, Primates < Mammals < Vertebrates < Animals, Plants
Habitat:	Tropical Zone < Terrestrial < Habitat, Rain Forest < Tropical Zone < Terrestrial < Habitat
Geographic Area:	Africa < Geographic Area, East Africa < Africa < Geographic Area
Additional Keywords:	abundance, basal area, GIS, Landsat, primates, remote sensing, spatially explicit models, tropical forest, Udzungwa
Abstract:	Spatially explicit models of animal abundance are a critical tool to inform conservation planning and management. However, they require the availability of spatially diffuse environmental predictors of abundance, which may be challenging especially in complex and heterogeneous habitats. This is particularly the case for tropical mammals, such as non-human primates, that depend on multi-layered and species-rich tree canopy coverage, which is usually measured through a limited sample of

ground plots. We developed an approach that calibrates remote-sensing imagery to ground measurements of tree density to derive basal area, in turn used as a predictor of primate density based on published models. We applied generalized linear models (GLM) to relate 9.8 ha ground samples of tree basal area to various metrics extracted from Landsat 8 imagery. We tested the potential of this approach for spatial inference of animal density by comparing the density predictions for an endangered colobus monkey, to previous estimates from field transect counts, measured basal area, and other predictors of abundance. The best GLM had high accuracy and showed no significant difference between predicted and observed values of basal area. Our species distribution model yielded predicted primate densities that matched those based on field measurements. Results show the potential of using open-access and global remote sensing data to derive an important predictor of animal abundance in tropical forests and in turn to make spatially explicit inference on animal density. This approach has important, inherent applications as it greatly magnifies the relevance of abundance modeling for informing conservation. This is especially true for threatened species living in heterogeneous habitats where spatial patterns of abundance, in relation to habitat and/or human disturbance factors, are often complex and, management decisions - such as improving forest protection - may need to be focused on priority areas.

Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.

DataS1.csv

1   **Integrating field and satellite data for spatially-explicit inference on the density of threatened**

2   **arboreal primates**

3

4   Nathalie Cavada[1,2], Marco Ciolli[1], Duccio Rocchini[3], Claudia Barelli[2], Andrew R. Marshall[4,5],

5   Francesco Rovero[2,6]

6

7   [1]DICAM Department of Civil, Environmental and Mechanical Engineering, University of Trento,

8   Trento, Italy.

9   [2]Tropical Biodiversity Section, MUSE - Museo delle Scienze, Trento, Italy.

10 [3]Biodiversity and Molecular Ecology Department, Research and Innovation Centre – Fondazione

11 Edmund Mach, San Michele all'Adige, Italy.

12 [4]CIRCLE, Environment Department, University of York, York, United Kingdom.

13 [5]Flamingo Land Ltd., Kirby Misperton, North Yorkshire, York, United Kingdom.

14 [6]Udzungwa Ecological Monitoring Centre, Udzungwa Mountains National Park, Mang'ula,

15 Tanzania.

16

17 Correspondence: Nathalie Cavada, DICAM Department of Civil, Environmental and Mechanical

18 Engineering, University of Trento, Via Mesiano 77, 38123 Trento, Italy.

19 Phone: +393404122810; E-mail: nathalie.cavada@unitn.it

20

1

23   **Abstract**

24   Spatially explicit models of animal abundance are a critical tool to inform conservation planning

25   and management. However, they require the availability of spatially diffuse environmental

26   predictors of abundance, which may be challenging especially in complex and heterogeneous

27   habitats. This is particularly the case for tropical mammals, such as non-human primates, that

28   depend on multi-layered and species-rich tree canopy coverage, which is usually measured through

29   a limited sample of ground plots. We developed an approach that calibrates remote-sensing imagery

30   to ground measurements of tree density to derive basal area, in turn used as a predictor of primate

31   density based on published models. We applied generalized linear models (GLM) to relate 9.8 ha

32   ground samples of tree basal area to various metrics extracted from Landsat 8 imagery. We tested

33   the potential of this approach for spatial inference of animal density by comparing the density

34   predictions for an endangered colobus monkey, to previous estimates from field transect counts,

35   measured basal area, and other predictors of abundance. The best GLM had high accuracy and

36   showed no significant difference between predicted and observed values of basal area. Our species

37   distribution model yielded predicted primate densities that matched those based on field

38   measurements. Results show the potential of using open-access and global remote sensing data to

39   derive an important predictor of animal abundance in tropical forests and in turn to make spatially

40   explicit inference on animal density. This approach has important, inherent applications as it greatly

41   magnifies the relevance of abundance modeling for informing conservation. This is especially true

42   for threatened species living in heterogeneous habitats where spatial patterns of abundance, in

43   relation to habitat and/or human disturbance factors, are often complex and, management decisions

44   - such as improving forest protection - may need to be focused on priority areas.

45

46   **Introduction**

47   Species abundance estimation and the identification of factors predicting its variation is a pervasive

48   goal in ecology and conservation biology and it is gaining increasing attention through the emergent

49   potential of spatially explicit modeling (Guisan and Zimmermann 2000, Guisan and Thuiller 2005,

2

50   Wulder and Franklin 2006, Anadón et al. 2010). This is particularly true for threatened species

51   living in heterogeneous landscapes, where habitat structure and human disturbance vary according

52   to complex spatial patterns. In these contexts, inference on abundance becomes truly informative

53   only when it accounts for such heterogeneity (Arroyo-Rodríguez and Fahrig 2014). Human-

54   modified landscapes are also expanding in tropical areas, where forest fragmentation, degradation

55   and defaunation strongly affect species viability (Balmford and Whitten 2003, Arroyo-Rodríguez

56   and Fahrig 2014). However, because of limited and substandard data, spatially explicit models are

57   less exploited in tropical areas compared to temperate ones (Cayuela et al. 2009). Thus, integrating

58   the use of field data with remote sensing data represents an advantageous approach to ensure data

59   quality for spatial modeling in these areas (Wilkie and Finn 1996, Proisy et al. 2007).

60

61   Remote sensing data (especially Landsat) have been used to investigate several ecological

62   questions, mainly related to land cover change, carbon storage and habitat mapping (Schroeder et

63   al. 2011, Legaard et al. 2015, Mayes et al. 2015, Twongyirwe et al. 2015). However, the resolution

64   and quality of Landsat data do not always adequately represent environmental components that are

65   most important for target species, such as vegetation structure, because optical satellite imagery is

66   not three-dimensional (Hall et al. 1995, Duncanson et al. 2010). Therefore, methods are needed to

67   characterize features of the forest structure that are relevant to target species, particularly for

68   inaccessible areas where Landsat images represent the only feasible option.

69

70   In this study, we aimed to derive arboreal primate density from remote sensing estimates of 'tree

71   stem basal area'. Basal area is typically related to canopy cover (Alexander 1971, Farr et al. 1989,

72   Smith et al. 1992), but the two measures are not directly interchangeable (Cade 1997). In particular,

73   mean basal area specifically measures the contribution of each tree to biomass and hence identifies

74   forest structure, succession stage and disturbance. Accordingly, it is a common measure of habitat

75   quality for predicting animal abundance (Braithwaite et al. 1989, Medley 1993, Umapathy and

76   Kumar 2000). This is especially true for non-human primates (Mbora and Meikle 2004, Cristóbal-

3

77   Azkarate et al. 2005, Anderson et al. 2007, Struhsaker and Rovero 2007) which are globally

78   threatened and in urgent need of conservation actions (Schipper et al. 2008, Schwitzer et al. 2015).

79   Our specific objectives were to: a) model measured basal area against a combination of different

80   metrics and indices derived from Landsat imagery; b) test the performance of the best-performing

81   model to predict values of basal area outside of the sampled areas; c) use the results to derive a

82   spatial map of population density of the endangered (IUCN 2015) Udzungwa red colobus monkey

83   (*Procolobus gordonorum*), based on previously published density-basal area model; d) compare the

84   modeled primate density to previous predictions from field measurements; e) further refine these

85   estimates using environmental and human predictors.

86

87   **Materials and Methods**

88   **Study area**

89   The Udzungwa Mountains are located in the south-central part of Tanzania and represent the largest

90   mountain bloc in the Eastern Arc Mountains, covering an area larger than 19,000 km^2 (Platts et al.

91   2011). Closed forest blocs, ranging in size from 12 to over 500 km^2 (Marshall et al. 2010), are

92   interspersed with drier habitats. We focused our study on the forest of Mwanihana, one of the largest

93   forest blocs (150.6 km^2) and under the protection of the Udzungwa Mountain National Park

94   (UMNP) since 1992. Highly variable habitat types are distributed along the altitudinal gradient of

95   the forest ranging from 351 to 2,263 m a.s.l. Deciduous forest is found in the lowland, with semi-

96   deciduous and evergreen forests covering the sub-montane and montane areas, while *Hagenia* and

97   bamboo-dominated forest characterize the upper montane level (Lovett et al. 2006). Woody

98   vegetation density increases with elevation, with the largest trees found at mid elevation, probably a

99   result of human disturbance and tree respiration costs (Marshall et al. 2012).

100

101   **Vegetation data**

102   We derived field data for tree stems ≥10cm DBH (Diameter at Breast Height; 1.3m) from three

103   sources (Fig. 1): (1) From the Tropical Ecology Assessment and Monitoring Network (TEAM)

104   (http://www.teamnetwork.org/, dataset ID 0327011905 4443), comprising six vegetation plots of

105   100 × 100m on a horizontal plane (i.e. adjusted for slope), following a standardized protocol

106   (TEAM Network 2011); (2) 153 vegetation plots of 25 × 25m, sampled along line transects

107   uniformly distributed in the forest (from Barelli et al. 2015); (3) 33 new randomly placed vegetation

108   plots of 25 × 25m, sampled in June-July 2015, stratified according to the predominant habitat

109   gradient from disturbed lowland deciduous to mature montane evergreen forest. All newly-sampled

110   plots were placed in the centre of Landsat pixels for concordance with our remote-sensing imagery.

111   A summary of the vegetation data sets is provided in Data S1.

112   We obtained a single, cloud free, L8 OLI/TIRS Landsat image (Landsat scene ID

113   LC81670652014299LGN00, courtesy of the U.S. Geological Survey), acquired October 26, 2014.

114

115   **Primate density data**

116   Density data on the Udzungwa red colobus from across the study area were obtained from an earlier

117   study (Cavada et al. 2016). This study used environmental covariates from the 153 plots established

118   by Barelli et al. (2015) and distance sampling along line transects, to estimate colobus density

119   across the study area. Transect data were modeled as a hierarchical coupled logistic regression,

120   assuming a Poisson distribution for the animal abundance at a transect level. The detection process

121   of the distance sampling was modeled according to a multinomial distribution, assuming a

122   monotonical decrease of the detection probability with the increasing distance of the animal groups

123   from the observer. The influence of a series of environmental and human disturbance covariates was

124   evaluated and incorporated on both the abundance and detection steps in the model. Final density

125   estimates at the plot level were derived from environmental correlates that included mean basal

126   area, elevation and distance from disturbance (i.e. forest edge), that were found to significantly

127   affect the abundance and detectability of the red colobus in the study area.

128

129   **Analysis**

130   **Landsat metrics and vegetation indices**

131   To model basal area we first derived various Landsat metrics (Table 1). This began with a Principal

132   Component Analysis (PCA) to extract uncorrelated information from the different spectral bands

133   provided by the Operational Land Imager (OLI) sensor of the Landsat 8 satellite. After applying

134   PCA we further compressed the spectral data applying the Tasseled Cap Transformation (TCT) to

135   represent forest structure (Cohen et al. 1995). We also used a GRASS module (Neteler et al. 2012),

136   modified to derive vegetation-related spectral indices, combining specific bands of the Landsat 8

137   satellite images (Data S2). Such indices enhance the signal related to vegetation, while minimizing

138   background edaphic, solar and atmospheric effects (Jackson and Huete 1991).

139

140   **Model building**

141   To relate field sampled values of basal area to the metrics calculated from the Landsat images, we

142   used all newly-sampled plots, plus a subsample of the TEAM and Barelli et al. (2015) plots. The

143   subsample plots were those showing at least 75% overlap with Landsat pixels (N=115). In each plot

144   we calculated the basal area (BA, m^2) for each sampled tree (DBH ≥10cm) as BA=π*(DBH/2)^2. We

145   then derived the mean basal area (MBA) for each plot, for use as the response variable (following

146   Barelli et al. (2015) and Cavada et al. (2016)).

147

148   We used generalized linear modeling (GLM) to investigate the relationship between the MBA- field

149   sampled values and the Landsat metrics and indices. Prior to building the models we checked for

150   the presence of collinearity among predictor variables to remove those providing identical

151   information. We thus calculated Variance Inflation Factor (VIF), using a cut off value of 10

152   (Marquardt 1970, Hair et al. 2006, Kennedy 2008) and we retained the uncorrelated predictors P1,

153   P2, RGI, RR, SLAVI. From an Empirical Cumulative Distribution Function (ECDF) of the response

154   variable, we decided to use an inverse Gaussian error distribution for the GLM with an inverse

155   squared link function (Fig. 2).

156

6

157 We built models using all the possible combinations of the retained Landsat predictors and we used

158 the Akaike Information Criterion (AIC) to rank the candidate models. We considered those models

159 showing ΔAIC<2 as equivalent (Anderson and Burnham 2002) and defined an average model by

160 determining Akaike weights (w_i) for each of the best models, using the packages 'AICcmodavg'

161 (Mazerolle 2015) and 'MUMin' (Barton 2014) in R version 3.2.1 (R Core Team 2015). For

162 validating the model we randomly split the MBA dataset into two subsets, one for model fitting

163 with 75% of the data (N=109) and one with the remaining 25% of the data (N=37). We then used

164 bootstrapping to verify the goodness of fit of the selected average model: we simulated 1,000

165 datasets from the subset derived for model fitting (i.e the one considering 75% of the data) and then

166 defined a function that returned the fit-statistic Pearson χ^2. We validated the model by checking the

167 distribution of the residuals for the validation subset. We evaluated model bias by comparing both

168 observed and predicted values, to a null model of mean residual prediction equal to zero, using

169 Wilcoxon's signed rank test (for α=0.05).

170

171 **Predictions: MBA values and RC density**

172 To predict density values for groups of red colobus across the entire Mwanihana forest, we first

173 derived spatially diffused values for MBA from our best fitting averaged model, giving an MBA

174 value for each Landsat pixel in the entire study area. We removed those values of MBA that

175 appeared as outliers in the derived dataset (i.e. >0.5m^2). We believed these outliers were found for

176 those pixels where our model was not able to derive realistic MBA values, inside those areas close

177 to forest borders as well as in areas located at high elevation (above 1800 m), where trees are sparse

178 and are replaced by other vegetation (Lovett et al. 2006).

179

180 Besides MBA, previous modeling of red colobus group density was most effective using elevation

181 (negative sign) and distance from disturbance/forest edge (negative sign) (Cavada et al. 2016). We

182 therefore calculated spatially diffused values for these variables from a Digital Elevation Model

183 (DEM) and from a shapefile of the forest edge, respectively. We then used a published hierarchical

7

184   model (Cavada et al. 2016) to predict primate density across the Mwanihana forest using these two

185   variables and spatially diffused values for MBA derived from our model.

186   Finally, we verified the accuracy of our approach by comparing the predicted primate density to

187   density estimates in Cavada et al. (2016) for those plots in Barelli et al. (2015) (N=65) that were

188   excluded while building the MBA model (see 'Model building' above). These density estimates

189   were plot-specific values derived from the hierarchical analysis described above, and hence were

190   effectively the only field based and site-specific density estimates that could be used for such

191   validation. We compared observed and predicted values using OP regression (Piñeiro et al. 2008)

192   and we compared the slope and the intercept of the fitted model with the 1:1 line.

193

194   **Results**

195   After selecting the plots suitable for the analysis, we retained 61 plots from Barelli et al. (2015) and

196   54 TEAM sub-plots. Adding these to the 33 newly sampled plots, we obtained an overall dataset of

197   148 plots and their corresponding sampled MBA values. We built models using all the possible

198   combinations of the metrics and indices calculated from the Landsat images, including a null model.

199   We retained six competing models of MBA (Table 2) that were averaged for predictions. The

200   resulting average model retained the first and the second components of the PCA and the indices

201   RGI, RR and SLAVI (Table 3). This model showed adequate fit based on the bootstrap P value

202   based on the Chi-square statistic (P=0.66) and no significant difference between observed and

203   predicted MBA values (W=602, P=0.92). The MBA model failed to derive plausible values in those

204   areas located at high altitudes as well as close to the forest edge (Fig. 3). We obtained a spatially-

205   explicit map of estimated density of red colobus groups across the whole study area, as influenced

206   by the covariates MBA (predicted from our model and with a positive effect), elevation and distance

207   from disturbance (i.e from the forest edge), both with a negative effect, according to the hierarchical

208   model defined in Cavada et al. (2016) (Fig. 4).

8

209   The OP regression yielded a R^2 of 0.84 attesting the accuracy of the predicted red colobus group

210   density values as derived by using the spatially diffused values for MBA obtained from the GLM

211   analysis (Fig. 5).

212

213   **Discussion**

214   We have successfully predicted and mapped the spatial density of an endangered primate, hence

215   showing how modeling ecologically-relevant predictors of abundance can improve predictions on

216   species distribution (Franklin 1995), across a broad spatial extent. The species' density pattern

217   highlighted in our map is consistent with results in previous studies that were based solely on

218   ground data and hence with limited spatial inference (Struhsaker and Rovero 2007, Barelli et al.

219   2015, Cavada et al. 2016).

220

221   Our best supported models showed high accuracy in predicting MBA values, making it a reliable

222   tool for inference beyond the ground measurement sites, with a good level of confidence and

223   precision. MBA is a highly relevant descriptor of the canopy structure as well as a significant

224   covariate that has emerged in different studies as influential for predominantly arboreal primates

225   (Struhsaker and Rovero 2007, Cavada et al. 2016). As a parameter quantifying forest cover, MBA is

226   also a recognized proxy for habitat degradation and fragmentation (Urquiza-Haas et al. 2007). The

227   best fit model we derived from GLM retained the first two components of the PCA. This fitted the

228   acknowledged evidence that Landsat products are able to discriminate forested habitats, through the

229   information provided by specific spectral channels (Blair and Baumgardner 1977, Jakubauskas

230   1996, Eklundh et al. 2001, Cohen and Goward 2004), in terms of the differential reflectance emitted

231   by the higher strata of the canopy. The information provided by the Landsat sensors can highlight

232   specific vegetation components (Thenkabail et al. 2000, Almeida and De Souza Filo 2004); in fact

233   the bands of the visible spectrum and of the Short-wave Infrared (SWIR) can be correlated with

234   several forest structures, including basal area (Muukkonen and Heiskanen 2005, 2007, Hall et al.

235   2006). The relationship with MBA shown by the first PCA component of our model might be due to

236   a large presence of trees with great basal area and tall canopy, causing pronounced shadowing

237   which translates in a lower reflectance.

238

239   Among the vegetation indices retained by the models, RGI can be interpreted as a proxy of the

240   forest phenology by the time when the Landsat image was acquired. Since such an index provides

241   information on the ratio of red to green reflectance, the positive effect we found on MBA could be

242   due to the contribution the index generally gives in evaluating the size of the tree crowns, which is

243   related to the basal area extent. During that period, a high amount of trees shows indeed a

244   breakdown of green pigments and leaves fade from green to yellow and red (Motohka et al. 2010).

245   The positive effect we found for RR was also confirmed by other studies that found a correlation

246   between the visible and the SWIR band of the Landsat with several physical structures of the forest

247   canopy, including basal area (Muukkonen and Heiskanen 2005, Hall et al. 2006, Tonolli et al.

248   2011). In addition, the positive relationship we found between MBA and SLAVI index is not

249   surprising given that the index accounts for the sensitivity of the mid-infrared wavelength to the

250   structure of the canopy, especially for heterogeneous forest compositions (Lymburner et al. 2000).

251

252   As the main goal of our study, we used the predicted and spatially diffused values of MBA to derive

253   a map of the Udzungwa red colobus density. This matched, at a wider and spatially-diffuse scale,

254   the density estimates found in prior studies (Barelli et al. 2015; Cavada et al. 2016). In particular, it

255   confirmed the red colobus's preference for lower-elevation forest that are close to its edge, variably

256   disturbed and covered with regenerating vegetation, that is recognized as an important food source

257   for the species (Barelli et al. 2015). Densities decreased where MBA values increased, i.e. in the

258   interior and old growth forest parts and at higher elevation. This in turn indicates resilience of the

259   animal to anthropogenic disturbance and again the preference shown by the species for forest edges.

260   Such a counter intuitive density trend, is clearly visualized in the spatially explicit map we obtained.

261   This provides novel indications for the protection of forest areas that are located at the interface

262   with intense anthropogenic activity.

263

264   We have confirmed that the use of remote sensing represents a robust tool to improve model

265   performance and to reduce the costs of data collection (He et al. 2015), which implies bypassing the

266   sample size limits associated with field measurements. We stress the importance of carefully

267   evaluating the process regarding the selection of adequate satellite images, given the sensitivity for

268   seasonality shown by some vegetation indices. High resolution images should certainly be preferred

269   when deriving remote-sensing based predictor variables that can be essential to improve predictive

270   species modeling. Nonetheless, the quality of such images can often be poor, due to cloud coverage

271   that hides the underlying canopy, i.e. the carried amount of information is lower than the spectral

272   noise (Woodcock and Strahler 1987, Ricotta et al. 1999). This phenomenon consistently arises in

273   images of tropical mountain forests, since clouds accumulate relatively more in dense forest cover

274   areas due to evapotranspiration (Nagendra and Rocchini 2008). Still, we demonstrated that since

275   high resolution products in some cases cannot be used, medium resolution images like Landsat

276   proved to be an excellent source of data for applications both in the study of tropical forest structure

277   and to develop reliable species distribution models. However, caution is recommended regarding

278   the generalization of our approach, which is mainly relevant to comparable study systems in terms

279   of both habitat and target species characteristics.

280

281   **Conclusions**

282   Spatially explicit, predictive models of animal abundance can offer a powerful insight on the

283   species status and distribution, helping to identify those sites where urgent intervention is needed in

284   terms of protection and conservation. Overcoming the lack of high resolution and high quality

285   remote sensing products as well as of spatially diffused covariates of abundance is essential, as it

286   can firmly boost the usefulness of species distribution models. By focusing on the endangered

287   Udzungwa red colobus, we showed the potential of this approach to derive accurate spatially

288   diffused estimates of animal density and distribution. This approach is particularly suitable for

289   species for which data availability is incomplete and spatial coverage is heterogeneous, affecting the

290   capacity of developing site-specific conservation and restoration programs where urgent forest and

291   species protection is needed.

292

293   **Acknowledgements**

308

309

310

311

312

313

314

315

316

317   **References**

318   Alexander, R. R. 1971. Crown Competition Factor (CCF) for Engelmann Spruce in the Central
319   Rocky Mountains. Ed 1971 Forest Service, U.S. Dept. of Agriculture, Rocky Mountain Forest and
320   Range Experiment Station.

321   Almeida, T. I. R., and C. De Souza Filo. 2004. Principal component analysis applied to feature-
322   oriented band ratios of hyperspectral data: A tool for vegetation studies. International Journal
323   of Remote Sensing 25:5005–5023.

324   Anadón, J. D., A. Giménez, and R. Ballestar. 2010. Linking local ecological knowledge and habitat
325   modelling to predict absolute species abundance on large scales. Biodiversity and
326   Conservation 19:1443–1454.

327   Anderson, D. R., and K. P. Burnham. 2002. Avoiding pitfalls when using information-theoretic
328   methods. The Journal of Wildlife Management 66:912–918.

329   Anderson, J., G. Cowlishaw, and J. M. Rowcliffe. 2007. Effects of forest fragmentation on the
330   abundance of Colobus angolensis palliatus in Kenya's coastal forests. International Journal of
331   Primatology 28:637–655.

332   Araldi, A., C. Barelli, K. Hodges, and F. Rovero. 2014. Density estimation of the endangered
333   Udzungwa red colobus (Procolobus gordonorum) and other arboreal primates in the Udzungwa
334   Mountains using systematic Distance Sampling. International Journal of Primatology 35:941–
335   956.

336   Arroyo-Rodríguez, V., and L. Fahrig. 2014. Why is a landscape perspective important in studies of
337   primates? American Journal of Primatology 909:901–909.

338   Balmford, A., and T. Whitten. 2003. Who should pay for tropical conservation, and how could the
339   costs be met? Oryx 37:238–250.

340   Barelli, C., R. Mundry, A. Araldi, K. Hodges, D. Rocchini, and F. Rovero. 2015. Modelling primate
341   abundance in complex landscapes: a case study from the Udzungwa Mountains of Tanzania.
342   International Journal of Primatology 36:209-226.

343   Barton, K. 2014. Multi-model inference. R package MuMIn version 1.10.5, 46.

344   Blair, B. O., and M. F. Baumgardner. 1977. Detection of the green and brown wave in hardwood
345   canopy covers using multidate, multispectral data from LANDSAT-1. Agronomy Journal 69:
346   808-811.

347   Braithwaite, L. W., M. P. Austin, M. Clayton, J. Turner, and A. O. Nicholls. 1989. On predicting the
348   presence of birds in Eucalyptus forest types. Biological Conservation 50:33–50.

349   Brown, L., J. M. Chen, S. G. Leblanc, and J. Cihlar. 2000. A shortwave infrared modification to the
350   Simple Ratio for LAI retrieval in boreal forests: an image and model analysis. Remote Sensing
351   of Environment 71:16–25.

352   Cade, B. S. 1997. Comparison of tree basal area and canopy cover in habitat models: subalpine
353   forest. Journal of Wildlife Management 61:326–335.

354   Cavada, N., C. Barelli, M. Ciolli, and F. Rovero. 2016. Primates in human-modified and fragmented
355   landscapes: the conservation relevance of modelling habitat and disturbance factors in density
356   estimation. Plos One 11:e0148289.

357   Cayuela, L., D. Golicher, A. Newton, H. Kolb, F. S. de Alburquerque, E. J. M. M. Arets, J. R. M.
358   Alkemade, and A. M. Pérez. 2009. Species distribution modelling in the tropics: problems,
359   potentialities, and the role of biological data for effective species conservation. Tropical
360   Conservation Science 2:319–352.

361   Chen, J. M. 1996. Evaluation of vegetation indices and a modified Simple Ratio for boreal
362   applications. Canadian Journal of Remote Sensing 22:1–21.

363   Cohen, W. B. and S. N. Goward. 2004. Landsat's role in ecological applications of Remote Sensing.
364   BioScience 54:535–545.

365   Cohen, W. B., T. A. Spies, and M. Fiorella 1995. Estimating the age and structure of forests in a
366   multi-ownership landscape of western Oregon, U.S.A. International Journal of Remote Sensing
367   16:721–746.

368   Coops, N. C., M. Johnson, M. A. Wulder, and J. C. White 2006. Assessment of QuickBird high
369   spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation.
370   Remote Sensing of Environment 103:67–80.

371   Cristóbal-Azkarate, J., J. J. Veà, N. Asensio, and E. Rodríguez-Luna. 2005. Biogeographical and
372   floristic predictors of the presence and abundance of mantled howlers (Alouatta palliata
373   mexicana) in rainforest fragments at Los Tuxtlas, Mexico. American Journal of Primatology
374   67:209–222.

375   Duncanson, L. I., K. O. Niemann, and M. A. Wulder. 2010. Integration of GLAS and Landsat TM
376   data for aboveground biomass estimation. Canadian Journal of Remote Sensing 36:129–141.

377   Eklundh, L., L. Harrie, and A. Kuusk. 2001. Investigating relationships between Landsat ETM+
378   sensor data and leaf area index in a boreal conifer forest. Remote Sensing of Environment
379   78:239–251.

380   Farr, W. A., D. J. DeMars, and J. E. Dealy. 1989. Height and crown width related to diameter for
381   open-grown western hemlock and Sitka spruce. Canadian Journal of Forest Research 19:
382   1203–1207.

383   Franklin, J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in
384   relation to environmental gradients. Progress in Physical Geography 19:474–499.

385   Gamon, J. A., and J. S. Surfus. 1999. Assessing leaf pigment content and activity with a
386   reflectometer. New Phytologist 143:105–117.

387   Gitelson, A. A., Y. J. Kaufman, and M. N. Merzlyak. 1996. Use of a green channel in remote
388   sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment 58:289–298.

389   Guisan, A. and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat
390   models. Ecology Letters 8:993–1009.

391   Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology.
392   Ecological Modelling 135:147–186.

393 Hair, J. F., B. Black, B. Babin, B., R. E. Anderson, and R. L. Tatham. 2006. Multivariate Data
394   Analysis 6th ed. Prentice Hall.

395 Hall, F. G., Y. E. Shimabukuro, and K. Huemmrich. 1995. Remote sensing of forest biophysical
396   structure using mixture decomposition and geometric reflectance models. Ecological
397   Applications 5:993–1013.

398 Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modelling forest stand structure
399   attributes using Landsat ETM+ data: application to mapping of aboveground biomass and
400   stand volume. Forest Ecology and Management 225:378–390.

401 Hardinsky, M. A., V. Klemas, and R. M. Smart. 1983. The influence of soil salinity, growth form
402   and leaf moisture on the spectral radiance of Spartina alterniflora canopies. Photogrammetric
403   Engineering and Remote Sensing 48:77-84.

404 He, K. S., B. A. Bradley, A. F. Cord, D. Rocchini, M. N. Tuanmu, S. Schmidtlein, W. Turner, M.
405   Wegmann, and N. Pettorelli. 2015. Will remote sensing shape the next generation of species
406   distribution models? Remote Sensing in Ecology and Conservation 1:4-18.

407 Hédl, R., M. Svátek, M. Dančák, A. W. Rodzay, A. B. Salleh, and A. S. Kamariah. 2009. A new
408   technique for inventory of permanent plots in tropical forests: a case study from lowland
409   dipterocarp forest in Kuala Belalong, Brunei darussalam. Blumea: Journal of Plant Taxonomy
410   and Plant Geography 54:124–130.

411 IUCN. 2015. The IUCN Red List of Threatened Species. Version 2015-4. <www.iucnredlist.org>.
412   Downloaded on 08 March 2016.

413 Jackson, R. D., and A. R. Huete. 1991. Interpreting vegetation indices. Preventive Veterinary
414   Medicine 11:185–200.

415 Jakubauskas, M. E. 1996. Thematic mapper characterization of lodgepole pine seral stages in
416   Yellowstone National Park, USA. Remote Sensing of Environment 56:118–132.

417 Jordan, C. F. 1969. Derivation of leaf-area index from quality of light on the forest floor. Ecology
418   50:663–666.

419 Kennedy, P. 2008. A Guide to Econometrics 6th ed. Wiley-Blackwell.

420 Legaard, K. R., S. A. Sader, and E. M. Simons-Legaard. 2015. Evaluating the impact of abrupt
421   changes in forest policy and management practices on landscape dynamics: analysis of a
422   Landsat image time series in the Atlantic Northern Forest. Plos One 10:e0130428.

423 Lovett, J. C., A. R. Marshall, and J. Carr. 2006. Changes in tropical forest vegetation along an
424   altitudinal gradient in the Udzungwa Mountains National Park, Tanzania. African Journal of
425   Ecology 44:478-490.

426 Lymburner, L., P. Beggs, and C. Jacobson. 2000. Estimation of canopy-average surface-specific leaf
427   area using Landsat TM data. Photogrammetric Engineering & Remote Sensing 66:183–191.

428 Marquardt, D. W. 1970. Generalized inverses, Ridge regression, biased linear estimation, and
429   nonlinear estimation. Technometrics 12:591–612.

430 Marshall, A. R., H. I. O. Jørgensbye, F. Rovero, P. J. Platts, P. C. L. White, and J. C. Lovett. 2010.
431   The species-area relationship and confounding variables in a threatened monkey community.
432   American Journal of Primatology 72:325–336.

433 Marshall, A. R., S. Willcock, P. J. Platts, J. C. Lovett, A. Balmford, N. D. Burgess, J. E. Latham, P.
434   K. T. Munishi, R. Salter, D. D. Shirima, and S. L. Lewis. 2012. Measuring and modelling
435   above-ground carbon and tree allometry along a tropical elevation gradient. Biological
436   Conservation 154:20–33.

437 Mayes, M. T., J. F. Mustard, and J. M. Melillo. 2015. Forest cover change in Miombo Woodlands:
438   modelling land cover of African dry tropical forests with linear spectral mixture analysis.
439   Remote Sensing of Environment 165:203–215.

440 Mazerolle, M. J. 2015. R package " AICcmodavg " Model selection and multimodel inference
441   based on (Q)AIC(c).

442 Mbora, D. N. M., and D. B. Meikle. 2004. Forest fragmentation and the distribution, abundance and
443   conservation of the Tana river red colobus (Procolobus rufomitratus). Biological Conservation
444   118:67–77.

445 Medley, K. E. 1993. Primate conservation along the Tana River, Kenya: an examination of the
446   forest habitat. Conservation biology 7:109–121.

447 Motohka, T., K. N. Nasahara, H. Oguma, and S. Tsuchida. 2010. Applicability of Green-Red
448   Vegetation index for Remote Sensing of vegetation phenology. Remote Sensing 2:2369–2387.

449 Muukkonen, P., and J. Heiskanen. 2005. Estimating biomass for boreal forests using ASTER
450   satellite data combined with standwise forest inventory data. Remote Sensing of Environment
451   99:434–447.

452 Muukkonen, P., and J. Heiskanen. 2007. Biomass estimation over a large area based on standwise
453   forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon
454   inventories. Remote Sensing of Environment 107:617–624.

455 Nagendra, H., and D. Rocchini. 2008. High resolution satellite imagery for tropical biodiversity
456   studies: The devil is in the detail. Biodiversity and Conservation 17:3431–3442.

457 Nemani, R., L. Pierce, S. Running, and L. Band. 1993. Forest ecosystem processes at the watershed
458   scale: sensitivity to remotely-sensed Leaf Area Index estimates. International Journal of
459   Remote Sensing 14:2519–2534.

460 Ota, T., O. Ahmed, S. Franklin, M. Wulder, T. Kajisa, N. Mizoue, S. Yoshida, G. Takao, Y. Hirata,
461   N. Furuya, T. Sano, S. Heng, and M. Vuthy. 2014. Estimation of airborne Lidar-derived
462   tropical forest canopy height using Landsat time series in Cambodia. Remote Sensing
463   6:10750–10772.

464 Piñeiro, G., S. Perelman, J. P. Guerschman, and J. M. Paruelo. 2008. How to evaluate models:
465   observed vs. predicted or predicted vs. observed? Ecological Modelling 216:316–322.

466 Platts, P. J., N. D. Burgess, R. E. Gereau, J. C. Lovett, A. R. Marshall, C. J. Mc Clean, P. K. E.
467   Pellikka, R. D. Swetnam, and R. Marchant. 2011. Delimiting tropical mountain ecoregions for
468   conservation. Environmental Conservation 38:312–324.

469   Proisy, C., P. Couteron, and F. Fromard. 2007. Predicting and mapping mangrove biomass from
470   canopy grain analysis using Fourier-based textural ordination of IKONOS images. Remote
471   Sensing of Environment 109:379–392.

472   R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for
473   Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

474   Ricotta, C., G. C. Avena, and F. Volpe. 1999. The influence of principal component analysis on the
475   spatial structure of a multispectral dataset. International Journal of Remote Sensing 20:3367–
476   3376.

477   Rouse, J. W., R. H. Haas, J. A. Schell, W. D. Deering, and J. C. Harlan. 1974. Monitoring the vernal
478   advancement and retrogradation (greenwave effect) of natural vegetation. Technical report,
479   NASA, United States.

480   Schipper, J., et al. 2008. The status of the world's land and marine mammals: diversity, threat, and
481   knowledge. Science 322:225–230.

482   Schroeder, T. A., M. A. Wulder, S. P. Healey, and G. G. Moisen. 2011. Mapping wildfire and
483   clearcut harvest disturbances in boreal forests with Landsat time series data. Remote Sensing
484   of Environment 115:1421–1433.

485   Schwitzer, C., R. A. Mittermeier, A. B. Rylands, F. Chiozza, E. A. Williamson, J. Wallis, and
486   Cotton, A., editors. 2015. Primates in Peril: The World's 25 Most Endangered Primates 2014-
487   2016. IUCN SSC Primate Specialist Group (PSG), International Primatological Society (IPS),
488   Conservation International (CI), and Bristol Zoological Society, Arlington, VA. iv+93pp.

489   Smith, W. R., R. M. Farrar Jr., P. A. Murphy, J. L. Yeiser, R. S. Meldahl, and J. S. Kush. 1992.
490   Crown and basal area relationships of open-grown southern pines for modelling competition
491   and growth. Canadian Journal of Forest Research 22:341–347.

492   Steel, R. I. 2012. The effects of habitat parameters on the behavior, ecology, and conservation of the
493   Udzungwa red colobus monkey (Procolobus gordonorum). Graduate School of Duke
494   University PhD thesis, Department of Biological Anthropology and Anatomy, Duke
495   University.

496   Struhsaker, T. T., and F. Rovero. 2007. Vegetative predictors of primate abundance: utility and
497   limitations of a fine-scale analysis. American Journal of Primatology 69:1242–1256.

498   TEAM Network. 2011. Terrestrial vertebrate protocol implementation manual, v. 3.1. Tropical
499   Ecology, Assessment and Monitoring Network, Center for applied biodiversity science,
500   Conservation International, Arlington, VA, USA.

501   Thenkabail, P. S., R. B. Smith, and E. De Pauw. 2000. Hyperspectral vegetation indices and their
502   relationships with agricultural crop characteristics. Remote Sensing of Environment 71:158–
503   182.

504   Tonolli, S., M. Dalponte, M. Neteler, M. Rodeghiero, L. Vescovo, and D. Gianelle. 2011. Fusion of
505   airborne LiDAR and satellite multispectral data for the estimation of timber volume in the
506   Southern Alps. Remote Sensing of Environment 115:2486–2498.

507 Twongyirwe, R., M. Bithell, K. S. Richards, and W. G. Rees. 2015. Land use policy three decades of
508   forest cover change in Uganda's Northern Albertine Rift Landscape. Land Use Policy 49:
509   236–251.

510 Umapathy, G., and A. Kumar. 2000. The occurrence of arboreal mammals in the rain forest
511   fragments in the Anamalai Hills, South India. Biological Conservation 92:311–319.

512 Urquiza-Haas, T., P. M. Dolman, and C. A. Peres. 2007. Regional scale variation in forest structure
513   and biomass in the Yucatan Peninsula, Mexico: effects of forest disturbance. Forest Ecology
514   and Management 247:80–90.

515 Vescovo, L., and D. Gianelle. 2008. Using the MIR bands in vegetation indices for the estimation of
516   grassland biophysical parameters from satellite remote sensing in the Alps region of Trentino
517   (Italy). Advances in Space Research 41:1764–1772.

518 Wilkie, D. S., and J. T. Finn. 1996. Remote sensing imagery for natural resources monitoring: a
519   guide for first-time users. Columbia University Press.

520 Woodcock, C. E., and A. H. Strahler. 1987. The factor of scale in remote sensing. Remote Sensing
521   of Environment 21:311–332.

522 Wulder, M. A., and S. E. Franklin. 2006. Understanding forest disturbance and spatial pattern,
523   Remote Sensing and GIS approaches. CRC Press.

524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548

549   **Tables**

550

551   **Table 1.** Vegetation indices extracted from a Landsat 8 image for comparison to ground sampled

552   measures of mean basal area (MBA).

Index	Algorithm	Description	References
Simple Ratio (**SR**)	$SR = \rho_{nir}/\rho_{red}$	Index related to changes in the amount of green vegetation; reduces the effect of atmosphere and topography.	(Jordan 1969)
Corrected Simple Ratio (**SRC**)	$SRC = SR\ (1-((\rho_{mir} - \rho_{mir\ min})/(\rho_{mir\ max} - \rho_{mir\ min}))$	Linearizes the relationships with parameters, accounting for MIR band.	(Brown et al. 2000)
Normalized Difference Vegetation Index (**NDVI**)	$NDVI = (\rho_{nir} - \rho_{red})/(\rho_{nir} + \rho_{red})$	Estimates the amount of vegetation, it assumes values that are normalized for the amount of incident radiation.	(Rouse et al. 1974)
Corrected Normalized Difference Vegetation Index (**NDVIC**)	$NDVIC = NDVI\ (1-((\rho_{mir} - \rho_{mir\ min})/(\rho_{mir\ max} - \rho_{mir\ min})$	Linearizes the relationships with parameters, accounting for MIR band	(Nemani et al. 1993)
Modified Simple Ratio (**MSR**)	$MSR = (\rho_{nir}/\rho_{red} - 1)/((\rho_{nir}/\rho_{red})^{1/2} + 1)$	Linearizes the	(Chen 1996)

19

		relationship between the index and biophysical parameters	
Reflectance Ratio (**RR**)	$RR = \rho_{mir}/ \rho_{red}$	Substitutes NIR band in SR with MIR band, which is more sensitive in distinguishing complex and stratified forest structures	(Tonolli et al. 2011)
Normalized Difference Water Index (**NDWI**)	$NDWI = (\rho_{nir} - \rho_{mir})/(\rho_{nir} + \rho_{mir})$	Sensitive to vegetation water	(Hardinsky et al. 1983)
Specific Leaf Area Vegetation Index (**SLAVI**)	$SLAVI = \rho_{nir}/(\rho_{red} + \rho_{mir})$	Estimates Specific Leaf Area	(Lymburner et al. 2000)
Red Green Ratio (**RGR**)	$RGR = \rho_{red}/\rho_{green}$	Sensitive to different foliar pigments	(Gamon and Surfus 1999)
Red Green Index (**RGI**)	$RGI = (\rho_{green} - \rho_{red})/(\rho_{green} + \rho_{red})$	Normalization of RGR results	(Coops et al. 2006)
Green Normalized Difference Vegetation Index (**GNDVI**)	$GNDVI = (\rho_{nir} - \rho_{green})/(\rho_{nir} + \rho_{green})$	Estimates the amount of green vegetation, exploiting the green channel, sensitive to chlorophyll	(Gitelson et al. 1996)
Normalized Canopy Index (**NCI**)	$NCI = (\rho_{mir} - \rho_{green})/(\rho_{mir} + \rho_{green})$	Linearizes the relationships with parameters, accounting for MIR	(Vescovo and Gianelle 2008)

20

		and green bands	
Tasseled Cap Angle (**TCA**)	TCA = arctan(TCG/TCB)	Index based on the angle formed by brightness (TCB) and greenness (TCG) in the vegetation plane, calculated from TCT (Tasseled Cap Transformation)	(Powell et al. 2010)

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

21

572   **Table 2.** Akaike Information Criterion (AIC) value for high ranked models (ΔAIC<2) of mean basal

573   area (MBA) modeled as a function of predictors derived from a Landsat 8 image.

Model	AIC	ΔAIC
MBA~P1+RGI	-620.70	0
MBA~P1+RGI+RR	-619.89	0.81
MBA~P1+SLAVI	-619.46	1.24
MBA~P1	-619.097	1.607
MBA~P1+P2+RGI	-619.096	1.609
MBA~P1+RR+SLAVI	-618.98	1.72

574   P1=First component of the Principal Component Analysis; P2= Second component of the Principal

575   Component Analysis; RGI=Red Green Index; RR=Red Ratio; SLAVI=Specific Leaf Area

576   Vegetation Index.

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592 **Table 3.** Estimates and standard errors for the parameters retained in the averaged model for mean

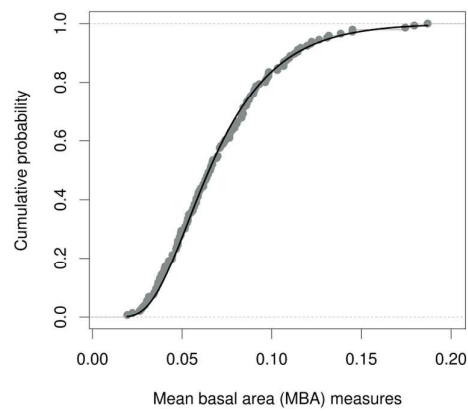593 basal area (MBA) modeled as a function of metrics and indices extracted from a Landsat 8 image.

Model-averaged coefficients	Estimate	SE	p
P1	-37.92	19.61	0.05
RGI	31.71	15.43	0.04
RR	19.40	16.45	0.2
SLAVI	27.09	16.18	0.09
P2	18.15	24.64	0.4

594 P1=First component of the Principal Component Analysis; P2= Second component of the Principal

595 Component Analysis; RGI=Red Green Index; RR=Red Ratio; SLAVI=Specific Leaf Area

596 Vegetation Index.

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

23

613   **Figures**

614

615

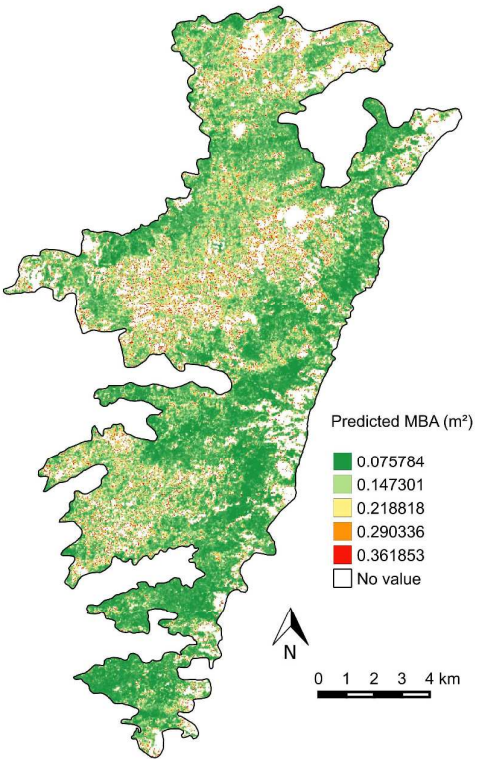616

617

618

619
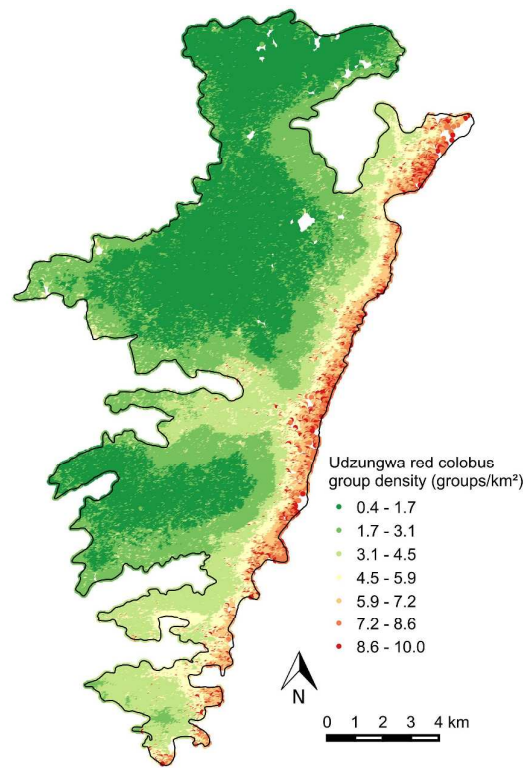
620

621

622

623

624

625

626   **Fig. 1.** Map of Mwanihana forest in the Udzungwa Mountains of Tanzania showing the distribution

627   of three vegetation plots data-sets used to derive basal area.
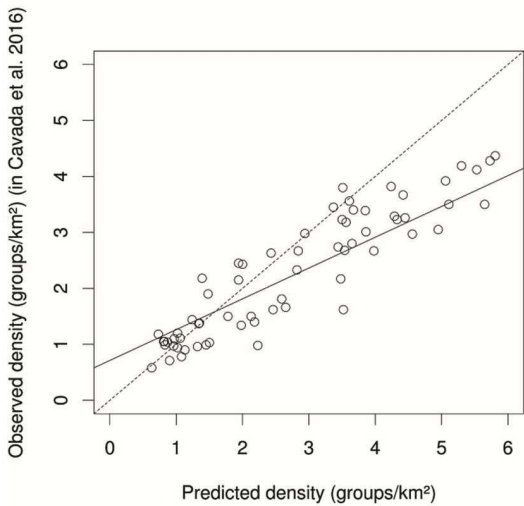
628

629

630

631

632

633

634

635

636　**Fig. 2.** Empirical cumulative distribution function of ground sampled measures of mean basal area

637　(MBA, gray dots) collected at tree plots in Mwanihana forest, Udzungwa Mountains, Tanzania. The

638　black line shows the fit of the theoretical inverse Gaussian distribution.

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

Fig. 3. Predicted values of mean basal area (MBA) across Mwanihana forest using the average

model of ground sampled values versus Landsat 8 metrics. White areas show pixels where the

model failed to predict plausible values of MBA (i.e. $<0.5m^2$).

670

671

672

673

674

675

676

677

678

679

680

681

Fig. 4. Predicted Udzungwa red colobus group density in Mwanihana forest using a species density

model (Cavada et al. 2016) derived from remotely sensed mean basal area.

684

685

686

687

688

689

690

691

692

Fig. 5. Linear regression (dotted line) of observed versus predicted values of Udzungwa red colobus

density (groups/km^2) among test vegetation plots (N=66). A 1:1 relationship is indicated by the

solid line.

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

28

711   **Supporting information**

712   **Data S1.** Summary of the dataset regarding the field sampled vegetation

713   **Data S2.** Code for the GRASS 7.0 module that was implemented to derive a series of vegetation

714   indices, combining specific bands of a Landsat 8 image.
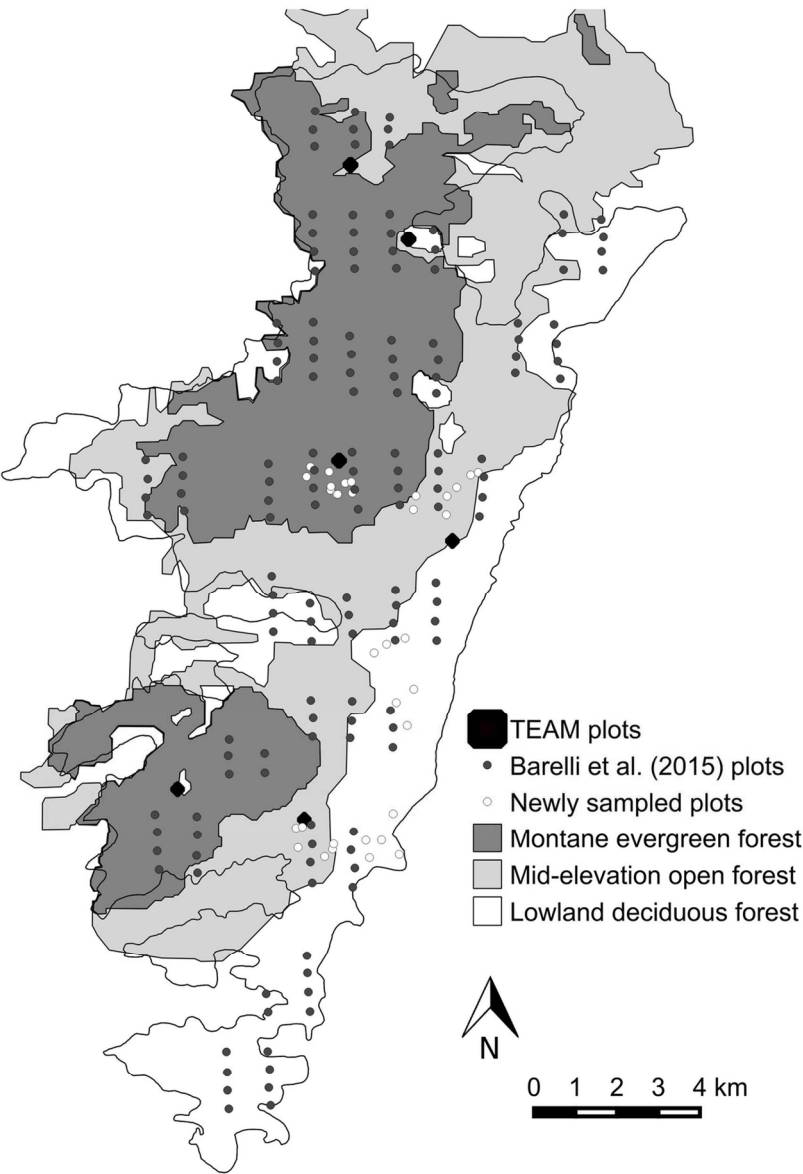
Fig. 1. Map of Mwanihana forest in the Udzungwa Mountains of Tanzania showing the distribution of three vegetation plots data-sets used to derive basal area.
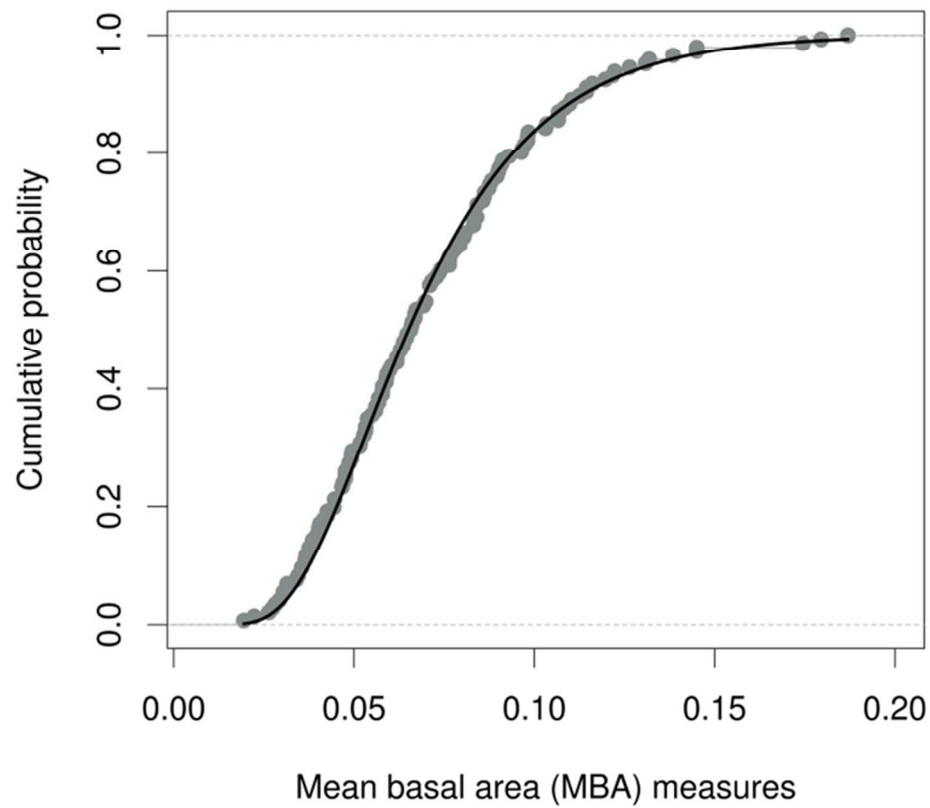
107x152mm (300 x 300 DPI)

Fig. 2. Empirical cumulative distribution function of ground sampled measures of mean basal area (MBA, gray dots) collected at tree plots in Mwanihana forest, Udzungwa Mountains, Tanzania. The black line shows the fit of the theoretical inverse Gaussian distribution.
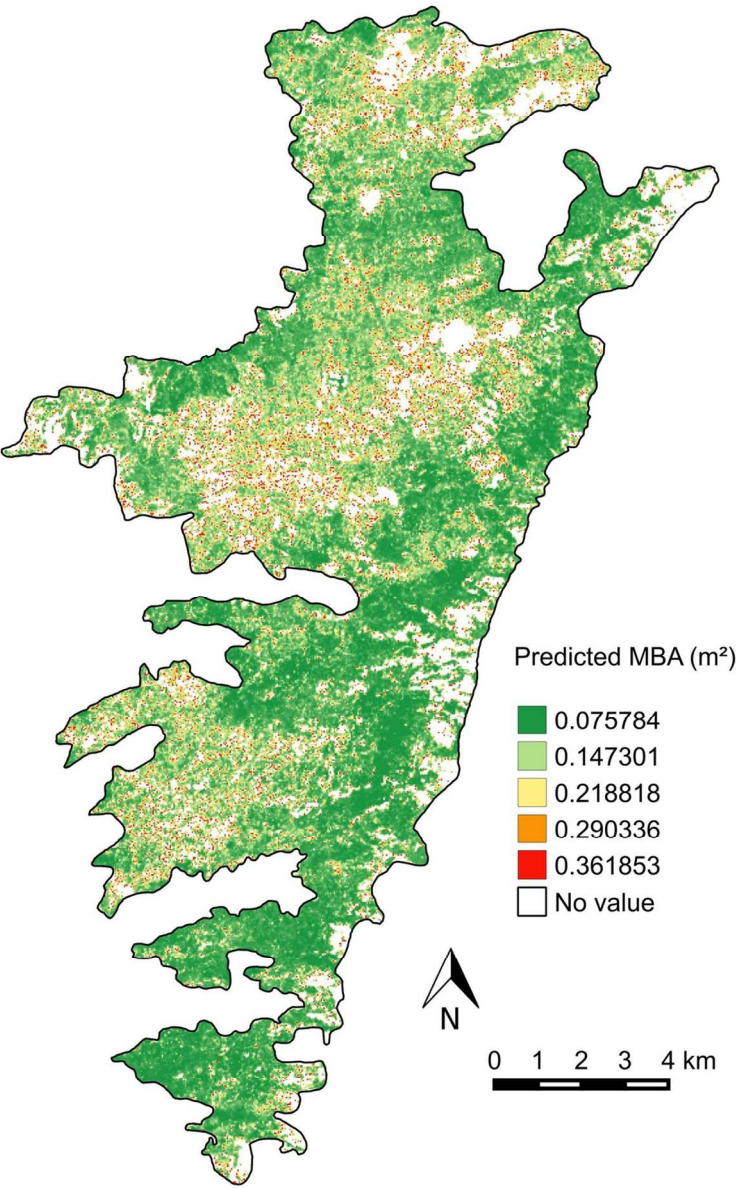
70x70mm (300 x 300 DPI)

Fig. 3. Predicted values of mean basal area (MBA) across Mwanihana forest using the average model of ground sampled values versus Landsat 8 metrics. White areas show pixels where the model failed to predict plausible values of MBA (i.e. <0.5m2).
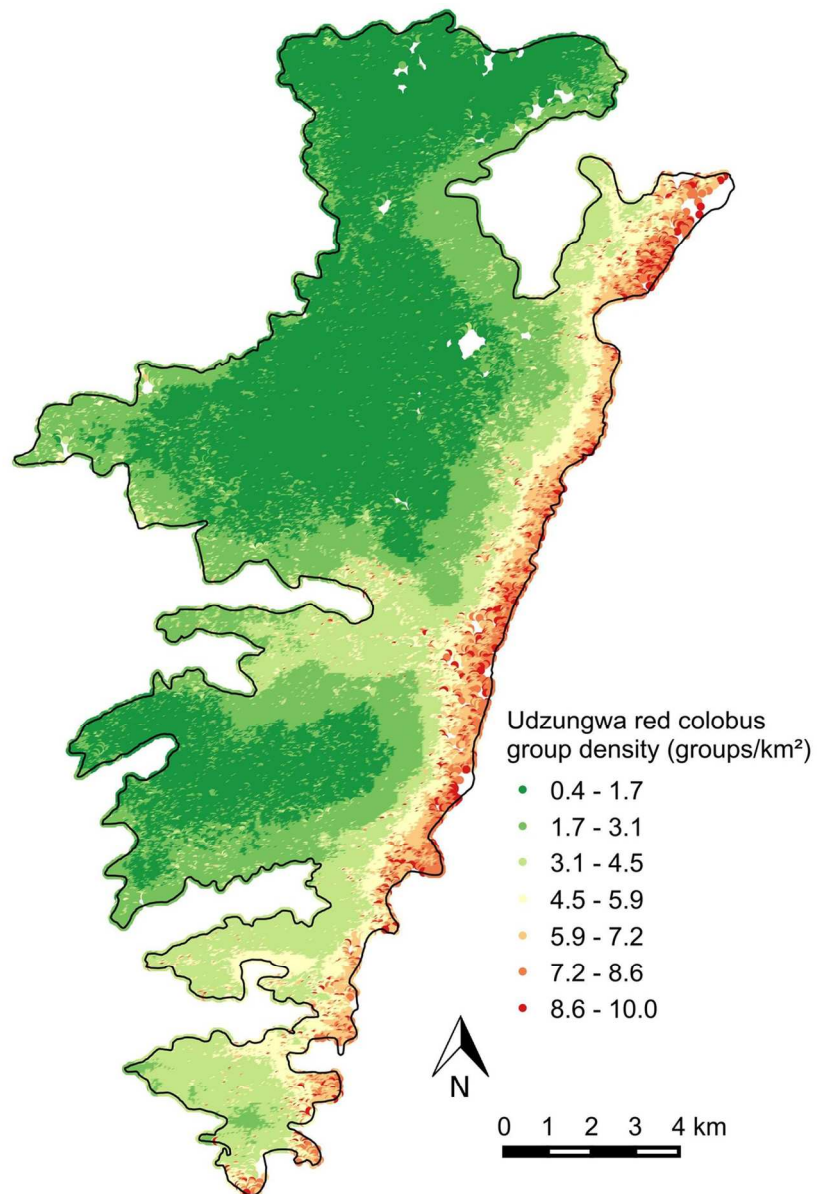
107x152mm (300 x 300 DPI)

Fig. 4. Predicted Udzungwa red colobus group density in Mwanihana forest using a species density model (Cavada et al. 2016) derived from remotely sensed mean basal area.
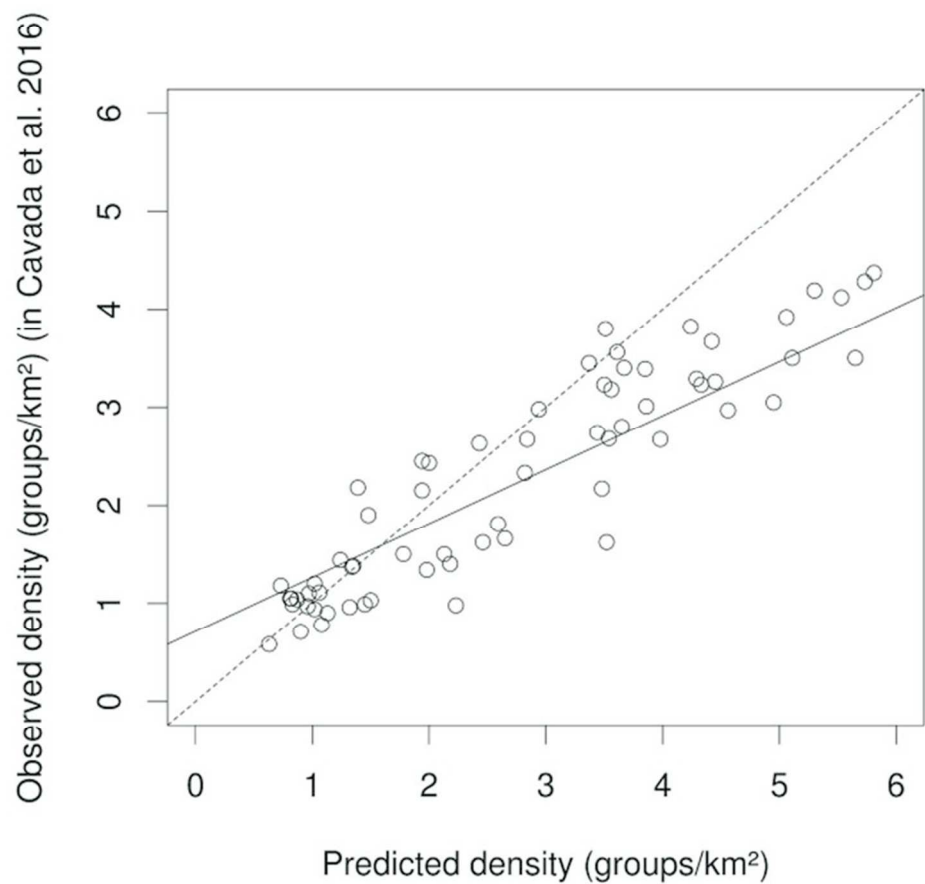
107x152mm (300 x 300 DPI)

Fig. 5. Linear regression (dotted line) of observed versus predicted values of Udzungwa red colobus density (groups/km2) among test vegetation plots (N=66). A 1:1 relationship is indicated by the solid line.

76x76mm (300 x 300 DPI)

S1 Metadata

Data set: ID for the data set source

DBH: Diameter at breast height, measured for all the tree stems having diameter >=10cm

Basal area: BA=π*(DBH/2)²

Climber: visually estimated coverage of climbers on trees as proportion of volume of the canopy, using 5 classes (0,25,50,75,100%).

Canopy: visually estimated extent of canopy cover,using 5 classes (0,25,50,75,100%)

```
#!/usr/bin/env python
#%module
#% description: Calculates vegetation indices for Landsat TM/ETM+/OLI spectral
bands
#% keywords: landsat, vegetation, indices, spectral, bands
#%end

#%option
#% key: band_prefix
#% type: string
#% gisprompt: old,cell,raster
#% description: Base name of input raster bands or a raster band map
#% required: yes
#%end
#%option
#% key: indices_prefix
#% type: string
#% description: Prefix for output raster indices maps
#% answer: spectral
#% required : yes
#%end
#%flag
#%  key: t
#%  description: Use bands for LANDSAT-4,5,7 (TM/ETM+)
#%END
#%flag
#%  key: o
#%  description: Use bands for LANDSAT-8 (OLI)
#%END
#%flag
#%  key: c
#%  description: Calculates also Cap Tassellation Indices
#%END
#%option
#% key: tc_prefix
#% type: string
#% gisprompt: old,cell,raster
#% description: If c flac: base name of input Tasselled Cap or a Tasselled Cap
map
#% required: no
#%end
#%Option
#% key: sensor
#% type: string
#% required: yes
#% multiple: no
#% options: LANDSAT-4;5;7 (TM/ETM+),LANDSAT-8 (OLI)
#% description: Use bands for sensor
#% answer: LANDSAT-8 (OLI)
#%End


import os, sys, shutil
import os.path, re
import grass.script as g


def main():

    #r.mapcalc float coercing with integer input
```

```
#(dn_B6-dn_B4)/(dn_B6+dn_B4)
#1.0*(dn_B6-dn_B4)/(dn_B6+dn_B4)
#(1.0*dn_B6-1.0*dn_B4)/(1.0*dn_B6+1.0*dn_B4)
#(float(dn_B6)-float(dn_B4))/(float(dn_B6)+float(dn_B4))

# define indices formulas

# RR: SWIR/Red reflectance ratio
rr_expr = '%(outpref)s_rr =1.0* %(mir)s / %(r)s'

# SR: Simple ratio NIR/Red reflectance ratio (Jordan, 1969)
sr_expr = '%(outpref)s_sr =1.0* %(nir)s / %(r)s'

# SRc: Corrected Simple Ratio (Brown et al. 2000)
src_expr = '%(outpref)s_src =1.0* $sr *(1-((%(mir)s -
%(minmir)s)/(%(maxmir)s - %(minmir)s)))'

# MSR: Modified Simple Ratio (Chen, 1996)
msr_expr = '%(outpref)s_msr =1.0* (%(nir)s / %(r)s -1)/(sqrt(%(nir)s /
%(r)s)+1)'

# RGR: Red Green Ratio (Gamon and Surfus)
rgr_expr = '%(outpref)s_rgr =1.0* %(r)s / %(g)s'

# RGI: Red Green Index (Coops et al.)
rgi_expr = '%(outpref)s_rgi =1.0* (%(g)s - %(r)s)/(%(g)s + %(r)s)'

# NDVI: Normalized Difference Vegetation Index (Rouse et al., 1974)
ndvi_expr = '%(outpref)s_ndvi =1.0* (%(nir)s - %(r)s)/(%(nir)s + %(r)s)'

# NDVIc: Corrected NDVI (Nemani et al., 1993)
ndvic_expr = '%(outpref)s_ndvic =1.0* $ndvi *(1-((%(mir)s -
%(minmir)s)/(%(maxmir)s - %(minmir)s)))'

# GNDVIgreen: NGreen Normalized Difference Vegetation Index (Gitelson et
al., 1996)
gndvi_expr = '%(outpref)s_gndvi =1.0* (%(nir)s - %(g)s)/(%(nir)s + %(g)s)'

# NDWI: Normalized Difference Water Index (Gao, 1996)
ndwi_expr = '%(outpref)s_ndwi =1.0* (%(nir)s - %(mir)s)/(%(nir)s +
%(mir)s)'

# SLAVI: Specific Leaf Area Vegetation Index (Lymburner et al., 2000)
slavi_expr = '%(outpref)s_slavi =1.0* %(nir)s /(%(r)s + %(mir)s)'

# NCI: Normalized Canopy Index (Vescovo & Gianelle, 2008)
nci_expr = '%(outpref)s_nci =1.0* (%(mir)s - %(g)s)/(%(mir)s + %(g)s)'

# NBR: Normalized Burn Ratio -> NOT IMPLEMENTED
# fire/burn index, use TM7/OLI_SWIR2

# TCA: Tasselled Cap Angle (Powell et al., 2010; Gomez et al., 2011)
tca_expr = '%(outpref)s_tca =1.0* atan(%(gr)s / %(br)s)' #deg angle

# ln(-We)
lnmwe_expr = '%(outpref)s_lnmwe =1.0* log(-%(we)s)'


# MAIN
landname= options['band_prefix'] #'toare_B'
```

```
indicespref= options['indices_prefix'] #'spectral'

#remove path before names and anything aftre the last point (ext)
#landpref=os.path.splitext(os.path.basename(landname))[0]

#remove ending numer from basename (purge path and @mapset)
#BASH: echo $(basename $landname) | sed 's/[0-9]*$//'
landpref=re.sub('[0-9]*$', '',os.path.basename(landname.split('@')[0]))

# define bands maps
if flags['o']:
    #landsat8
    g.message("OLI sensor")
    blue=landpref+'2'
    green=landpref+'3'
    red=landpref+'4'
    ninfrar=landpref+'5'
    minfrar=landpref+'7' #SWIR1
elif flags['t']:
    #landsat7
    g.message("TM/ETM+ sensor")
    blue=landpref+'1'
    green=landpref+'2'
    red=landpref+'3'
    ninfrar=landpref+'4'
    minfrar=landpref+'5'
else:
    #landsat8
    g.message("Warning: no sensor specified, defaout OLI used")
    blue=landpref+'2'
    green=landpref+'3'
    red=landpref+'4'
    ninfrar=landpref+'5'
    minfrar=landpref+'7' #SWIR1

#set region on a band map (are all equal)
g.run_command('g.region', rast = minfrar)

# mir max and min
min_mir = g.raster_info(minfrar)['min']
max_mir = g.raster_info(minfrar)['max']

bands= {
    "outpref" : indicespref,
    "b" : blue,
    "g" : green,
    "r" : red,
    "nir" : ninfrar,
    "mir" : minfrar,
    "minmir" : min_mir,
    "maxmir" : max_mir,
}

# compute indices with GRASS mapcalc
g.message("Calculating vegetation indices")
g.mapcalc(rr_expr % bands, overwrite = True)
g.mapcalc(sr_expr % bands, overwrite = True)
g.mapcalc(src_expr % bands, sr=indicespref+'_sr', overwrite = True)
g.mapcalc(msr_expr % bands, overwrite = True)
g.mapcalc(rgr_expr % bands, overwrite = True)
```

```
    g.mapcalc(rgi_expr % bands, overwrite = True)
    g.mapcalc(ndvi_expr % bands, overwrite = True)
    g.mapcalc(ndvic_expr % bands, ndvi=indicespref+'_ndvi', overwrite = True)
    g.mapcalc(gndvi_expr % bands, overwrite = True)
    g.mapcalc(ndwi_expr % bands, overwrite = True)
    g.mapcalc(slavi_expr % bands, overwrite = True)
    g.mapcalc(nci_expr % bands, overwrite = True)

    if flags['c']:
        tcname= options['tc_prefix']
        if tcname=="":
            g.message("Warning: no TC prefix, defaout 'tct8_C.' used")
            tcpref='tct8_C.'
        else:
            tcpref=re.sub('[0-9]*$',
'',os.path.basename(tcname.split('@')[0]))

        comp= {
            "outpref" : indicespref,
            "br" : tcpref+'1',
            "gr" : tcpref+'2',
            "we" : tcpref+'3',
        }

        g.message("Calculating Cap Tassellation indices")
        g.mapcalc(tca_expr % comp, overwrite = True)
        #g.mapcalc(lnmwe_expr % comp, overwrite = True) #null() 4 We>0

    return 0
    #End main


if __name__ == "__main__":
    options, flags = g.parser()
    sys.exit(main())
```