



Supplementary Materials for

Mapping *Opuntia stricta* in the Arid and Semi-Arid Environment of Kenya Using Sentinel-2 Imagery and Ensemble Machine Learning Classifiers

James M. Muthoka ^{1,*}, Edward E. Salakpi ², Edward Ouko ³, Zhuang-Fang Yi ⁴, Alexander S. Antonarakis ¹ and Pedram Rowhani ¹

- ¹ Department of Geography, University of Sussex, Brighton BN1 9QJ, UK; A.Antonarakis@sussex.ac.uk (A.S.A.); p.rowhani@sussex.ac.uk (P.R.)
- ² Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK; E.Salakpi@sussex.ac.uk
- ³ Regional Centre for Mapping of Resources for Development Technical services, Nairobi 00618, Kenya; eouko@rcmrd.org
- ⁴ Development Seed, Washington, DC 20001, USA; nana@developmentseed.org
- * Correspondence: j.muthoka@sussex.ac.uk

Contents of this file

Supplementary Table S1 - vegetation indices

The support information provided in this file presents the common formulas used to derive vegetation indices from optical Sentinel-2 images. Details regarding each of the presented spectral and topographic indices can be found in the main papers referenced.

<i>Table S1.</i> Formulas of	vegetation and	l topographic indices in this study.	
------------------------------	----------------	--------------------------------------	--

Name	Equation	Reference
RVI	$RVI = \frac{B_{NIR}}{B_{RED}}$	[70]
PVI	$PVI = \left(\frac{1}{\sqrt{a^2} + 1}\right)(B_{NIR} - ar - b)$	[71]
NDVI	$NDVI = \frac{B_{NIR} - B_{RED}}{B_{NIR} + B_{RED}}$	[72]
IPVI	$IPVI = \frac{B_{NIR}}{B_{RED} + B_{NIR}}$	[73]
ARVI	$ARVI = \frac{B_{NIR} - B_{RED} - y(B_{RED} - B_{BLUE})}{B_{NIR} - B_{RED} - y(B_{RED} - B_{BLUE})}$	[74]
	MSAVI	
MSAVI	$-\frac{2B_{NIR} + 1\sqrt{(2B_{NIR} + 1)^2 - 8(B_{NIR} - B_{RED})}}{2B_{NIR} + 1\sqrt{(2B_{NIR} + 1)^2 - 8(B_{NIR} - B_{RED})}}$	[75]
	- 2	

Note: BNIR. BRED, BGREEN, BBLUE correspond to Near Infrared, Red, Green and Blue Sentinel 2 bands, y = quotient from components of the atmospheric reflectance in blue and red channels, a = 1.17, b = 3.37.

Supplementary Table S2 – Topographic indices

The support information provided in this file presents the common formulas used to derive topographic index from a Digital Elevation Model (DEM) data. Details regarding each of the presented spectral and topographic indices can be found in the main papers referenced.

Table S2. Formulas of topographic indices in this study.

Name	Equation		Refe	rence		
Slope (S)	[05]					
Aspect (A)	$A = 270^\circ + \operatorname{arctanf}_y f_x - 90^\circ f_x f_x $	[66]				
TWI	$TWI = In\left(\frac{\alpha}{\tan\beta}\right)$	[86]				
Where f_x and f_y rep	resent gradients, N-S and W-E direc-					
tions, respectively, r	is the grid resolution, $z_i (1 \le i \le 9)$ is the					
elevation at a 3 × 3 ma	trix with $i = 5$ as centre point (see fig-					
	7	8	9			
	4	5	6			
$f_x = (z_8 - $	$(z_2)/2r, f_y = (z_6 - z_4)/2r$	rÎ.				
T: 11 · · ·			▶	3		
Finally, α is cat	chment area and β slope angle.	зX	3 X 3 matrix			
		77)ac			

Note: No further processing was undertaken on the topographic indices derived.

Supplementary Table S3 – Results of the Shapiro-Wilk normality test The support information provided in this file section presents the Shapiro-Wilk test for the land cover training based on Sentinel-2 spectral bands.

Table S3. Results of the Shapiro-Wilk normality test.

Land Covers	Bare C	Ground	d Fores		Grass		Opuntia stricta		Shrubs		Water	
Sentinel-2 bands	W-stat	p-value	W-stat	p-value	W-stat	p-value	W-stat	p-value	W-stat	p-value	W-stat	p-value
Blue	0.975	0.275	0.959	0.220	0.992	0.947	0.971	0.400	0.987	0.768	0.948	0.560
Green	0.973	0.228	0.964	0.304	0.981	0.414	0.952	0.086	0.979	0.344	0.958	0.607
Red	0.986	0.751	0.947	0.101	0.965	0.054	0.987	0.911	0.966	0.078	0.976	0.703
RedEdge1	0.967	0.114	0.972	0.496	0.981	0.384	0.920	0.008	0.971	0.144	0.994	0.850
RedEdge2	0.966	0.098	0.973	0.517	0.977	0.247	0.964	0.232	0.982	0.472	0.997	0.893
RedEdge3	0.983	0.583	0.961	0.246	0.972	0.141	0.979	0.647	0.968	0.102	0.998	0.907
NIR	0.944	0.010	0.955	0.160	0.967	0.069	0.937	0.027	0.986	0.705	0.997	0.899
NIR1	0.972	0.192	0.954	0.147	0.964	0.052	0.977	0.570	0.974	0.207	0.994	0.854
SWIR1	0.970	0.162	0.953	0.141	0.965	0.055	0.958	0.145	0.981	0.446	0.970	0.668
SWIR2	0.988	0.841	0.953	0.145	0.989	0.807	0.977	0.564	0.987	0.768	0.996	0.875
W-stat = Shapiro statistic					p-	<i>p</i> -value = significance (<0.05) are shown in bold						

Supplementary Figure S1 – Model 2.a and Model 2b feature importance analysis Supplementary Figure S1 – Feature importance analysis

The support information provided in this file presents the feature importance analysis for model 2.a and model 2.b.



Figure S1. model 2.a and model 2.b feature importance analysis.