Supplemental Documents for: “High-resolution short-wave infrared imagery provides improved estimates of rock gardening on Rapa Nui (Easter Island)”

Dylan S. Davis

2023-02-16

This is an R Markdown document that contains the R-script used for analysis documented in the manuscript. Following the code, the output is provided for all machine learning algorithms reported in the main text.. Script was designed and run in R v.4.0.2 (R Core Team 2020).

Load Libraries

```
library(RStoolbox)
library(caret)
library(MIAmaxent)
library(randomForest)
library(e1071)
library(raster)
library(spatstat)
library(rgdal)
library(sf)
library(stars)
library(rgeos)
library(dplyr)
library(maptools)
library(sp)
```

Load Datasets

```
#set working directory
setwd("C:/Users/dylan/Downloads/Mosaic_PS_16bit")


#Open VNIR files
B <- raster::raster("MS_8b_rs.tif", band = 1)
G <- raster::raster("MS_8b_rs.tif", band = 2)
Y <- raster::raster("MS_8b_rs.tif", band = 3)
R <- raster::raster("MS_8b_rs.tif", band = 4)
RE <- raster::raster("MS_8b_rs.tif", band = 5)
NIR1 <- raster::raster("MS_8b_rs.tif", band = 6)
NIR2 <- raster::raster("MS_8b_rs.tif", band = 7)
```

```
VNIR_stk <- raster::stack(B, G, Y, R, RE, NIR1, NIR2)

#Set working directory
setwd("C:/Users/dylan/Documents/School_Work/Rapa_Nui")

#Load image files
#SWIR
B1 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 1)
B2 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 2)
B3 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 3)
B4 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 4)
B5 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 5)
B6 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 6)
B7 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 7)
B8 <-
raster::raster("Delivery/Mosaic_8bit/Easter_Island_SWIR_8Band_8bit.tif", band
= 8)

SWIR_stk <- raster::stack(B1, B2, B3, B4, B5, B6, B7, B8)
```

Pansharpen and Merge VNIR and SWIR data

```
#load AOI shapefile
AOI <- rgdal::readOGR("AOI.shp")

#clip and merge SWIR and VNIR files

SWIR_clp <- raster::crop(SWIR_stk, AOI)

VNIR_clp <- raster::crop(VNIR_stk, AOI)

#Pansharpen SWIR using VNIR
WV3_PS <- RStoolbox::panSharpen(SWIR_stk, NIR1, method = "pca")

#Merge VNIR and pansharpened SWIR
```

```
WV3 <- raster::stack(WV3_PS, VNIR_stk)
```

Load Training Data and Classify Images

```
#Load training data
train <- rgdal::readOGR("Training_v7.shp")

##Plot Training Data
olpar <- par(no.readonly = TRUE) # back-up par
colors <- c("yellow", "red", "blue", 'green', 'black')
plot(B1)
plot(train, col = colors[train$Class], pch = 19)

##CLASSIFY SWIR DATASET

## Fit classifier (splitting training into 70\% training data, 30\%
validation data)

MaxEnt_SWIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                           algorithm = "maxent", tuneLength = 1,
trainPartition = 0.8, filename = "MxEnt_v3.tif")

MaxEnt_SWIR #Display model results, accuracy metrics, and confusion matrix

MLC_SWIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                           model = "mlc", tuneLength = 1,
trainPartition = 0.8, filename = "MLC_v3.tif")

MLC_SWIR #Display model results, accuracy metrics, and confusion matrix

RF_SWIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                       model = "rf", tuneLength = 1,
trainPartition = 0.8, filename = "RF_v3.tif")

RF_SWIR #Display model results, accuracy metrics, and confusion matrix


##CLASSIFY VNIR DATASET

## Fit classifier (splitting training into 70\% training data, 30\%
validation data)

MaxEnt_VNIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                             algorithm = "maxent", tuneLength =
```

```
1, trainPartition = 0.8, filename = "MxEnt_v3.tif")

MaxEnt_VNIR #Display model results, accuracy metrics, and confusion matrix

MLC_VNIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                        model = "mlc", tuneLength = 1,
trainPartition = 0.8, filename = "MLC_v3.tif")

MLC_VNIR #Display model results, accuracy metrics, and confusion matrix

RF_VNIR       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                       model = "rf", tuneLength = 1,
trainPartition = 0.8, filename = "RF_v3.tif")

RF_VNIR #Display model results, accuracy metrics, and confusion matrix

##CLASSIFY Pansharpened SWIR DATASET

## Fit classifier (splitting training into 70\% training data, 30\%
validation data)

MaxEnt_PS       <- RStoolbox::superClass(SWIR_stk, trainData = train,
responseCol = "Name",
                                         algorithm = "maxent", tuneLength =
1, trainPartition = 0.8, filename = "MxEnt_v3.tif")

MaxEnt_PS #Display model results, accuracy metrics, and confusion matrix
```

Model Output

SWIR Classifications

Random Forest

```
************ Validation **************

$validation

Confusion Matrix and Statistics

                 Reference
```

Prediction	bare_soil	Bedrock	Forest/Trees	Grass	Mulch	Urban/Developed
bare_soil	409	40	0	1	4	2
Bedrock	168	789	0	21	24	348
Forest/Trees	0	0	1002	0	0	0

Grass	185	6	0	846	0	0
Mulch	22	149	0	0	81	44
Urban/ Developed	218	23	0	0	0	609

Overall Statistics

```
               Accuracy : 0.7485
                 95% CI : (0.7363, 0.7605)
    No Information Rate : 0.2018
    P-Value [Acc > NIR] : < 2.2e-16
                  Kappa : 0.6915
 Mcnemar's Test P-Value : NA
```

Statistics by Class:

	bare_soil	Bedrock	Forest /Trees	Grass	Mulch	Urban/Developed
Sensitivity	0.40818	0.7835	1.0000	0.9747	0.74312	0.6072
Specificity	0.98822	0.8592	1.0000	0.9537	0.95596	0.9396
Pos Pred Value	0.89693	0.5844	1.0000	0.8158	0.27365	0.7165
Neg Pred Value	0.86924	0.9401	1.0000	0.9944	0.99404	0.9049
Prevalence	0.20076	0.2018	0.2008	0.1739	0.02184	0.2010
Detection Rate	0.08195	0.1581	0.2008	0.1695	0.01623	0.1220
Detection Prevalence	0.09136	0.2705	0.2008	0.2078	0.05931	0.1703
Balanced	0.69820	0.8214	1.0000	0.9642	0.84954	0.7734

Accuracy						

Maximum Likelihood

********** Validation **********

$validation

Confusion Matrix and Statistics

	Reference					
Prediction	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
bare_soil	490	213	1	0	4	0
Bedrock	111	566	2	0	79	243
Forest/Trees	0	1	993	0	0	0
Grass	63	36	6	1004	0	0
Mulch	29	60	0	0	297	42
Urban/ Developed	0	132	1	0	0	204

Overall Statistics

```
               Accuracy : 0.7765
                 95% CI : (0.7641, 0.7885)
    No Information Rate : 0.2202
    P-Value [Acc > NIR] : < 2.2e-16
                  Kappa : 0.7248
 Mcnemar's Test P-Value : NA
```

Statistics by Class:

	bare_soil	Class: Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
Sensitivity	0.7071	0.5615	0.9900	1.0000	0.78158	0.41718
Specificity	0.9439	0.8781	.9997	0.9706	0.96879	0.96747
Pos Pred Value	0.6921	0.5654	0.9990	0.9053	0.69393	0.60534
Neg Pred Value	0.9475	0.8764	0.9972	1.0000	0.98000	0.93278
Prevalence	0.1514	0.2202	0.2191	0.2194	0.08302	0.10684
Detection Rate	0.1071	0.1237	0.2170	0.2194	0.06489	0.04457
Detection Prevalence	0.1547	0.2187	0.2172	0.2423	0.09351	0.07363
Balanced Accuracy	0.8255	0.7198	0.9949	0.9853	0.87518	0.69232

Maximum Entropy

********** Validation **********

$validation

Confusion Matrix and Statistics

Reference

Prediction	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
bare_soil	713	12	0	0	5	2
Bedrock	259	630	2	83	91	9
Forest/	0	0	1001	1	0	0

Trees						
Grass	10	3	0	919	0	0
Mulch	7	293	0	0	657	21
Urban/Developed	13	69	0	0	18	295

Overall Statistics

```
               Accuracy : 0.8244
                 95% CI : (0.8137, 0.8347)
    No Information Rate : 0.1969
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7865

 Mcnemar's Test P-Value : NA
```

Statistics by Class:

	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
Sensitivity	0.7116	0.6256	0.9980	0.9163	0.8521	0.90214
Specificity	0.9954	0.8919	0.9998	0.9968	0.9261	0.97911
Pos Pred Value	0.9740	0.5866	0.9990	0.9861	0.6718	0.74684
Neg Pred Value	0.9340	0.9067	0.9995	0.9799	0.9724	0.99322
Prevalence	0.1960	0.1969	0.1962	0.1962	0.1508	0.06395

Detection Rate	0.1394	0.1232	0.1958	0.1797	0.1285	0.05770
Detection Prevalence	0.1432	0.2101	0.1960	0.1823	0.1913	0.07725
Balanced Accuracy	0.8535	0.7587	0.9989	0.9565	0.8891	0.94062

VNIR Classifications

Random Forest

********** Validation **********

$validation

Confusion Matrix and Statistics

Reference

Prediction	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
bare_soil	448	51	13	26	29	19
Bedrock	257	570	27	111	155	127
Forest/ Trees	2	46	713	0	1	29
Grass	12	96	37	819	3	59
Mulch	203	239	7	37	815	15
Urban/ Developed	80	3	0	3	0	748

Overall Statistics

Accuracy : 0.7091

95% CI : (0.6973, 0.7208)

No Information Rate : 0.1733

P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.6505

Mcnemar's Test P-Value : < 2.2e-16

Statistics by Class:

	bare_soil	Bedrock	Trees/ Forest	Grass	Mulch	Urban/ Developed
Sensitivity	0.44711	0.56716	0.8946	0.8223	0.8126	0.7503
Specificity	0.97124	0.85881	0.9844	0.9569	0.8956	0.9821
Pos Pred Value	0.76451	0.45710	0.9014	0.7982	0.6193	0.8969
Neg Pred Value	0.89375	0.90446	0.9832	0.9629	0.9581	0.9499
Prevalence	0.17276	0.17328	0.1374	0.1717	0.1729	0.1719
Detection Rate	0.07724	0.09828	0.1229	0.1412	0.1405	0.1290
Detection Prevalence	0.10103	0.21500	0.1364	0.1769	0.2269	0.1438
Balanced Accuracy	0.70917	0.71299	0.9395	0.8896	0.8541	0.8662

Maximum Likelihood

********** Validation **********

$validation

Confusion Matrix and Statistics

Reference

Prediction	bare_soil	Bedrock	Forest/	Grass	Mulch	Urban/

			Trees			Developed
bare_soil	475	373	14	42	13	4
Bedrock	27	224	35	90	37	739
Forest/Trees	9	51	694	20	0	1
Grass	365	105	43	835	0	32
Mulch	107	228	3	7	953	10
Urban/Developed	17	20	26	0	0	214

Overall Statistics

```
               Accuracy : 0.584
                 95% CI : (0.5712, 0.5967)
    No Information Rate : 0.1725
    P-Value [Acc > NIR] : < 2.2e-16
                  Kappa : 0.5002
 Mcnemar's Test P-Value : < 2.2e-16
```

Statistics by Class:

	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
Sensitivity	0.47500	0.22378	0.8515	0.8400	0.9501	0.21400
Specificity	0.90733	0.80715	0.9838	0.8869	0.9262	0.98691
Pos Pred Value	0.51574	0.19444	0.8955	0.6051	0.7286	0.77256

Neg Pred Value	0.89268	0.83330	0.9760	0.9641	0.9889	0.85802
Prevalence	0.17203	0.17220	0.1402	0.1710	0.1725	0.17203
Detection Rate	0.08171	0.03853	0.1194	0.1436	0.1639	0.03681
Detection Prevalence	0.15844	0.19818	0.1333	0.2374	0.2250	0.04765
Balanced Accuracy	0.69117	0.51546	0.9177	0.8635	0.9382	0.60046

Maximum Entropy

********** Validation **********

$validation

Confusion Matrix and Statistics

Reference

Prediction	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
bare_soil	694	170	17	93	25	2
Bedrock	43	464	31	49	293	28
Forest/ Trees	0	9	751	262	8	20
Grass	59	29	33	592	20	23
Mulch	202	241	1	2	646	9
Urban/ Developed	5	89	2	2	12	907

Overall Statistics

Accuracy : 0.695

95% CI : (0.683, 0.7068)

No Information Rate : 0.1721

```
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.6342

 Mcnemar's Test P-Value : < 2.2e-16
```

Statistics by Class:

	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Develpoed
Sensitivity	0.6919	0.46307	0.8994	0.5920	0.6434	0.9171
Specificity	0.9364	0.90809	0.9402	0.9661	0.9058	0.9773
Pos Pred Value	0.6933	0.51101	0.7152	0.7831	0.5867	0.8918
Neg Pred Value	0.9361	0.89076	0.9824	0.9196	0.9243	0.9830
Prevalence	0.1720	0.17178	0.1432	0.1714	0.1721	0.1696
Detection Rate	0.1190	0.07955	0.1288	0.1015	0.1107	0.1555
Detection Prevalence	0.1716	0.15567	0.1800	0.1296	0.1888	0.1744
Balanced Accuracy	0.8142	0.68558	0.9198	0.7790	0.7746	0.9472

Pansharpened VNIR/SWIR Composite Classification

Maximum Entropy

```
************ Validation **************

$validation

Confusion Matrix and Statistics

                  Reference
```

Prediction	bare_soil	Bedrock	Forest/ Trees	Grass	Mulch	Urban/ Developed
bare_soil	729	11	6	13	11	10
Bedrock	132	653	5	56	343	58
Forest/ Trees	0	2	812	64	0	0
Grass	0	0	33	862	0	0
Mulch	131	278	1	0	461	12
Urban/ Developed	10	59	0	0	5	923

Overall Statistics

```
               Accuracy : 0.7817
                 95% CI : (0.7707, 0.7924)
    No Information Rate : 0.1766
    P-Value [Acc > NIR] : < 2.2e-16
                  Kappa : 0.7378
 Mcnemar's Test P-Value : NA
```

Statistics by Class:

	bare_soil	Bedrock	Trees/ Forest	Grass	Mulch	Urban/ Developed
Sensitivity	0.7275	0.6510	0.9475	0.8663	0.56220	0.9202
Specificity	0.9891	0.8730	0.9863	0.9930	0.91317	0.9842
Pos Pred Value	0.9346	0.5237	0.9248	0.9631	0.52208	0.9258

Neg Pred Value	0.9443	0.9210	0.9906	0.9722	0.92516	0.9829
Prevalence	0.1764	0.1766	0.1509	0.1752	0.14437	0.1766
Detection Rate	0.1283	0.1150	0.1430	0.1518	0.08116	0.1625
Detection Prevalence	0.1373	0.2195	0.1546	0.1576	0.15546	0.1755
Balanced Accuracy	0.8583	0.7620	0.9669	0.9296	0.73768	0.9522