Biodiversity and Conservation Science

An investigation into identifying wild sandalwood using aerial imagery

Ricky van Dongen[1], Katherine Zdunic[1], Hans Blom[2] and Irfan Iqbal[2]

[1]Department of Biodiversity, Conservation and Attractions

[2]Forest Products Commission

January 2020

Department of **Biodiversity, Conservation and Attractions**

GOVERNMENT OF WESTERN AUSTRALIA

Department of Biodiversity, Conservation and Attractions
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: (08) 9219 9000
Fax: (08) 9334 0498

www.dbca.wa.gov.au

This report/document/publication was prepared by Ricky van Dongen

Questions regarding the use of this material should be directed to:
Research Officer
Biodiversity and Conservation Science
Department of Biodiversity, Conservation and Attractions
Locked Bag 104
Bentley Delivery Centre WA 6983
Phone: 9219 9571
Email: Ricky.vanDongen@dbca.wa.gov.au

The recommended reference for this publication is:
Department Biodiversity, Conservation and Attractions, 2019, *Investigating wild sandalwood identification using aerial imagery*, Department of Biodiversity, Conservation and Attractions, Perth.

This document is available in alternative formats on request.

Contents

Appendices

Figures

Tables

Summary

The aim of this project was to assess the feasibility of using aerial imagery to map the occurrence wild sandalwood. While the accuracy of known identified sandalwood stems in the classification was high (>0.8) there are also an excessive number of false positives, with over 1500 sandalwood predicted in an area known to have only 257. The main factor causing the high number of false positives is the small proportion of sandalwood in the population (257 in an area with ~ 9000 plants). This small proportion means that even if a model has a small percentage error the number of false positives will be greater than the number of sandalwood.

In terms of model performance, the best classification was achieved using eCognition 9.5 software and implementing a watershed segmentation, at a scale of 5 on a smoothed 0.25 m green leaf algorithm image. This method appears to optimally smooth within crown variability while maintaining between crown definition. While this was judged to be the optimal model many other options performed comparably.

The 0.25 m imagery pixel size appeared to be an optimal resolution for the task, the 0.05 m data did not produce better results. The normalized surface model (NSM) was rated as the most influential variable in the modelling process. However, a comparison of sandalwood NSM heights and field height measures highlighted that the NSM dataset contains significant errors. These errors are likely to impact the classification accuracy.

1 Introduction

The ability to detect individual wild sandalwood trees would provide a powerful management and regulation tool for the Forest Products Commission (FPC) and Department of Biodiversity, Conservation and Attractions (DBCA). The output of this project will provide guidance in the effectiveness of this type of imagery and analysis to identify sandalwood locations and guide future research for the management of wild sandalwood.

1.1 Study site

The study site is 65 km north east of Kalgoorlie. It is located in the Eastern Murchison sub IBRA region (Thackway and Cresswell 1995). A 24 ha (600 m x 400 m) area was delineated as a research site for wild sandalwood identification. It ensured that i) the research site had at least 100 sandalwood trees; ii) the site fell within a single land system; and iii) it contained ground features that could be identified on aerial imagery acquired by Arbor Carbon in April 2018.

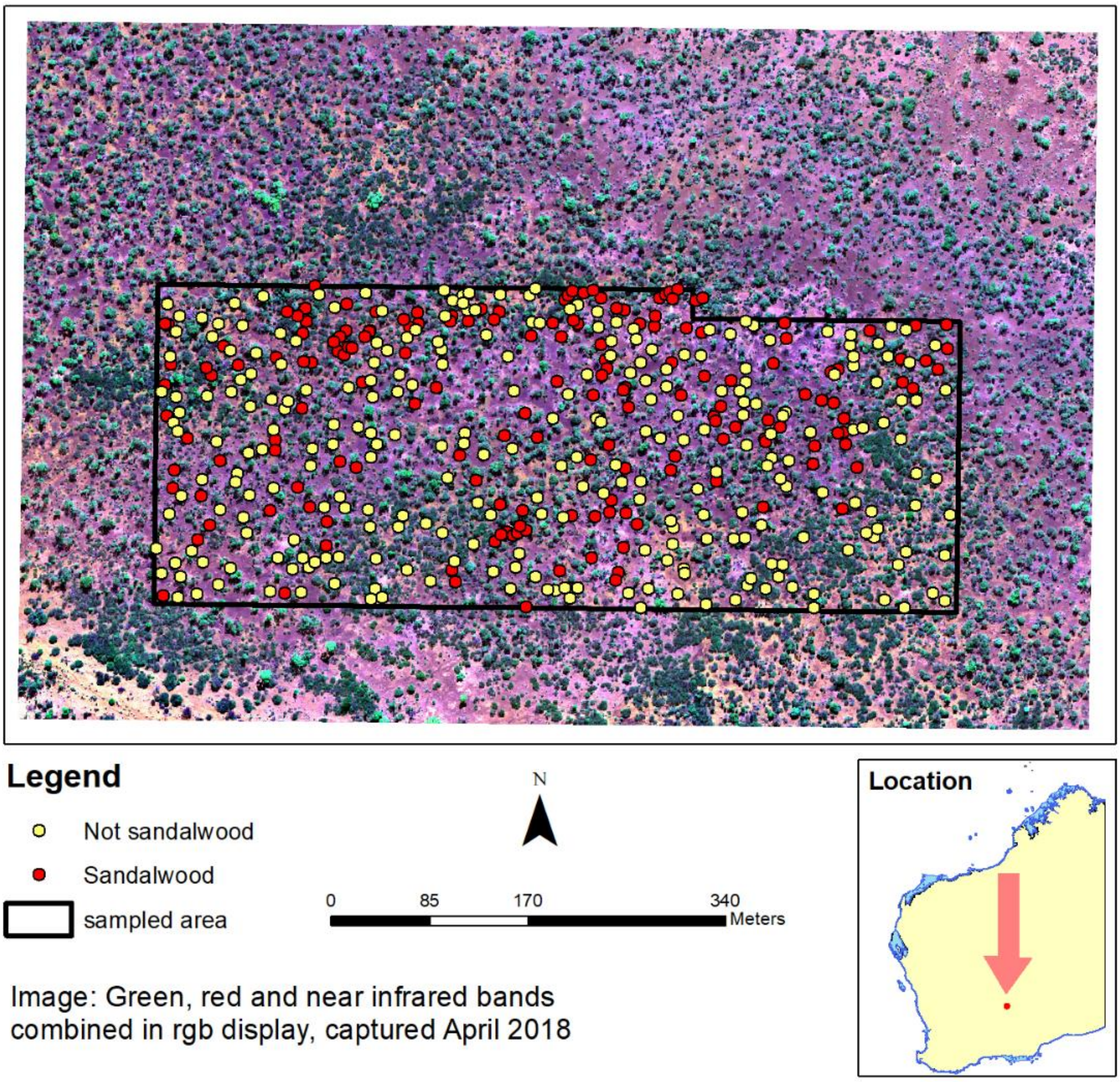


Figure 1: The study site location along with the aerial imagery and sandalwood and "not sandalwood" field data. The black bounding box marks 16 ha, out of 24 ha, within which data for all sandalwood trees were collected.

1.2 Imagery data

The imagery provided to DBCA by FPC consisted of four bands (blue, green, red and near-infrared) at 25 cm resolution and three bands (blue, green and red) at 5 cm resolution. A normalised surface model (NSM) at 25 cm was also provided. The NSM is a raster dataset where each pixel represents the height above ground level. The NSM was calculated from high-resolution (RGB; 3 cm) drone photography captured – with 80% forward overlap and 60% side overlap in two directions (0 and 90 degrees), on 17 September 2019.

1.3 Field data

A 25 m x 25 m grid was laid over the 24 ha research site to facilitate navigation and ensure completeness of data collected in field. Field data was collected between 10 and 12 September 2019. The coordinates of all sandalwood trees were collected using a dGPS unit (Stonex 800/900 with H10 Atlas signal). Due to time constraints, only sandalwood data were collected within 16 ha out of 24 ha (Figure 1). Field data was collected, with sub-meter accuracy, for a total of 257 sandalwood trees. For each tree, an ArcGIS Collector form was used to log stem diameter (mm), crown diameter in north-south and east-west directions (m), height (m) and the proportion to which a sandalwood crown was covered by another species. No data for trees other than sandalwood were collected in the field; the coordinates for non-sandalwood trees (n = 272) were instead determined through desktop analysis by FPC. All points are shown in Figure 1.

A total of 11 corflute panels (50 cm x 50 cm) were distributed across the research site, fixed to wooden pegs and their centre coordinates determined with a dGPS. These panels were used as ground control points (GCPs) to geo-rectify drone photography, which was then used to validate the geometric accuracy of the aerial imagery provided by Arbor Carbon.

2 Methods

2.1 Software

Software used in this analysis included eCognition version 9.5 (https://geospatial.trimble.com/products-and-solutions/ecognition) and R version 3.5.2 (R Core Team 2017). eCognition was used for undertaking image segmentation and calculating segment variables. R was used in image pre-processing and statistical modelling.

2.2 Pre-processing

A median filter was applied to individual bands and two spectral indices. The indices included the green leaf algorithm (GLA) (Macfarlane and Ogden 2012) and the normalised difference vegetation index (NDVI) (Rouse et al. 1974). The purpose of applying the median filter was to smooth out variations within individual canopies. This allows the following iterations of image segmentations to more accurately delineate individual crowns.

The effect of applying a median filter to a GLA image is shown in Figure 2. From visual inspection the 5 x 5 filter appeared to most effectively smooth out within crown variability while maintaining differences between crowns.

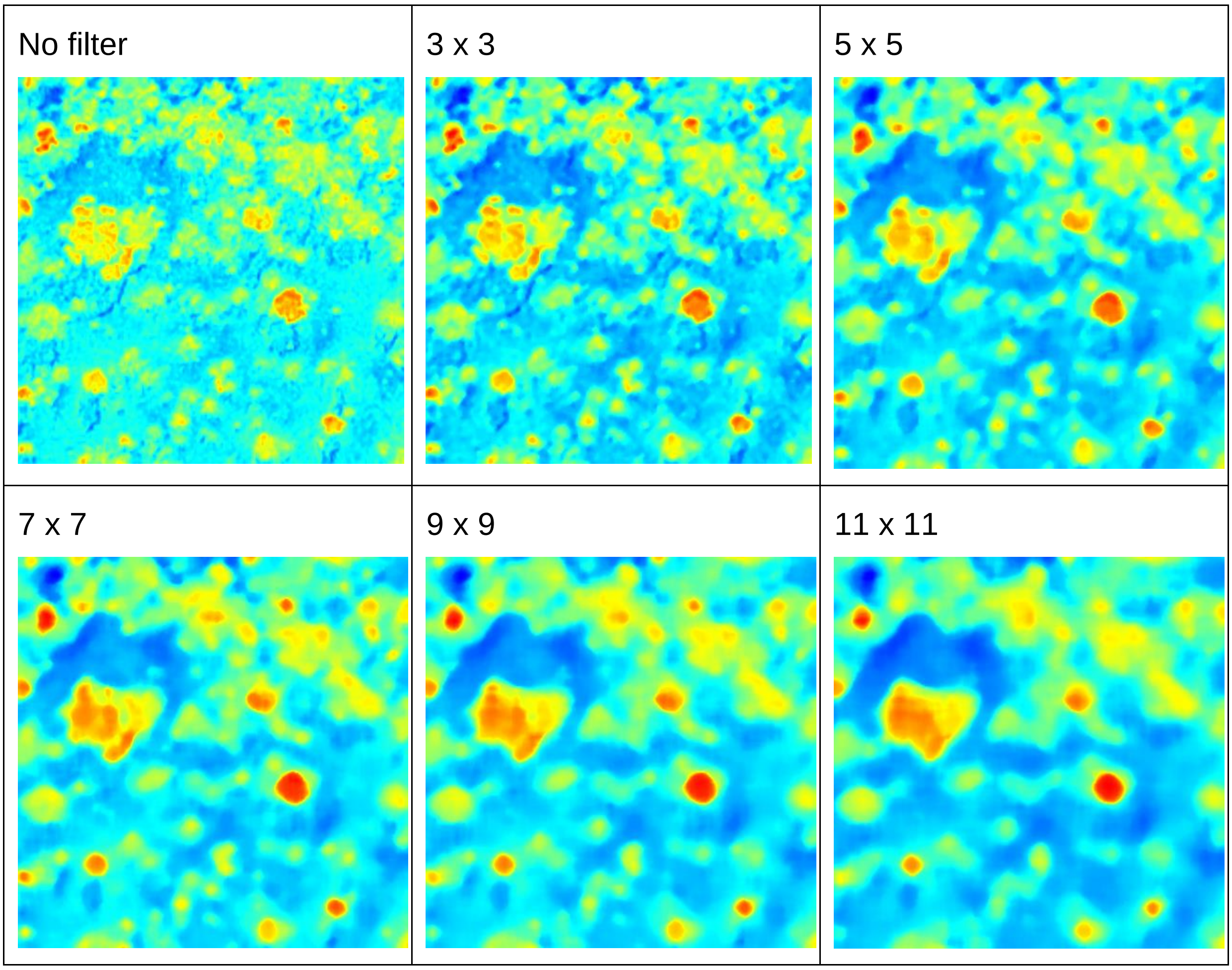


Figure 2: Median filters ranging in size applied to the GLA image.

2.3 Segmentation

Multiresolution and watershed segmentation methods are available in eCognition. These were deemed most likely to be effective at delineating sandalwood tree crowns. Multiresolution segmentation groups pixels with like values, this process can be constrained by size, shape and compactness. The watershed segmentation treats the data like an elevation model where segments attempt to follow “ridge lines”. This can be run in an inverted mode so segment boundaries follow “valleys”. The inverted method was used in this study. The size of watershed segments can be varied using a threshold parameter. Examples of the watershed and multiresolution segmentations are shown in Figure 3.

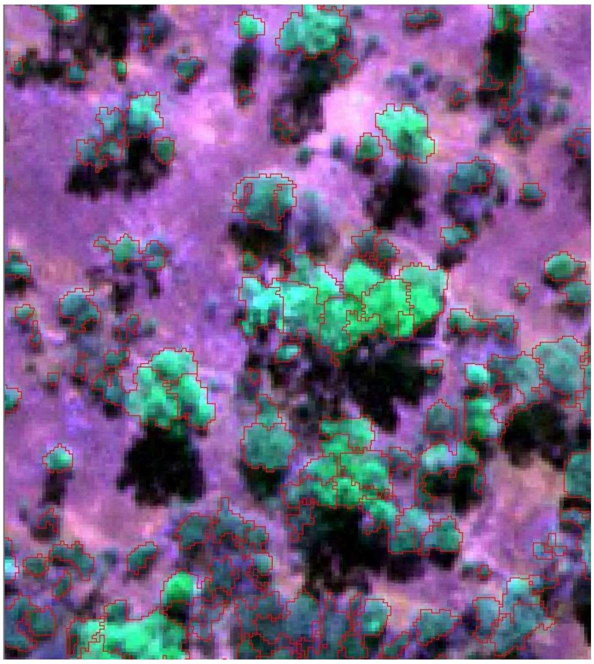
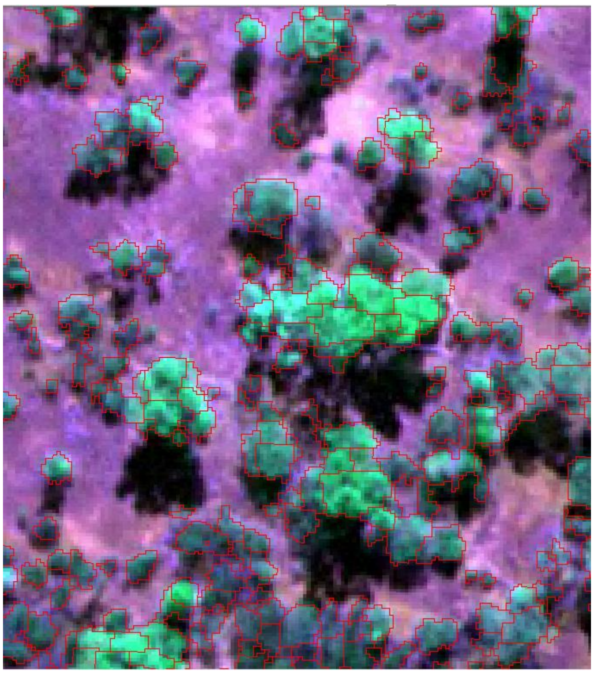
<table>
<tr>
<td>Watershed segmentation
Size threshold: 10
</td>
<td>Multi-resolution segmentation
Scale: 10
Shape: 0.8
Compactness: 0.8
</td>
</tr>
</table>

Figure 3: Examples of watershed and multiresolution segmentation methods.

A range of segmentation options were trialled:

> Trial 1: Using the 25 cm data, watershed segmentation at thresholds 5, 10, 20, 30 and 40 using the GLA, green layer and NDVI with a 5 x 5 median filter as the input segmentation layers.

Trial 2: Using the 25 cm data, multiresolution segmentation at scales of 10 to 100 incrementing every 10 counts.

Trial 3: Using the 5 cm data, multiresolution segmentation at scales of 40, 50 and 60.

Trial 4: Using the 25 cm data, watershed segmentation at thresholds 5, 10, 20, 30 and 40 with the GLA image and training data converted to polygons.

Prior to the watershed segmentations “green” vegetation was separated from bare ground and shadow using the following thresholds: mean NDVI > 0.15 and green brightness > 1100. These values were changed incrementally, and the results visually inspected until a suitable result was achieved.

2.4 Variables

A number of image indices were calculated to try to differentiate sandalwood from non-sandalwood trees. These are summarised in table 1. The GLA and NDVI were calculated in R during pre-processing, the remainder were calculated in eCognition.

Table 1: Indices and individual bands trialled in the image analysis.

Index	Description	Formula
BNDVI	Blue Normalized Difference Vegetation Index	(nir − blue) / (nir +blue)
GNDVI	Green Normalized Difference Vegetation Index	(nir − green) / (nir + green)
NDVI	Normalized Difference Vegetation Index	(nir − red) / (nir + red)
SIPI2	Structure Intensive Pigment Index 2	(nir − green) / (nir - red)
VARI2	Visible Atmospheric Resistant Index 2	(green − red) / (green + red + blue)
RVI	Ratio Vegetation Index	red / nir
DVI	Difference Vegetation Index	nir - red
GLA	Green Leaf Algorithm	((2 * green) – red – blue) / ((2 * green) + red + blue)
brightness	Mean band value	(nir + red + green + blue)/4
Blue	Blue band value	blue
Green	Green band value	green
Red	Red band value	red
Nir	Near infrared band value	nir
NSM	Normalized surface model	NSM

In eCognition size, shape and positional variables can also be calculated. These are shown in Table 2.

Table 2: Size, shape and positional variables calculated in eCognition.

Index	**Description**
Area	Segment area in number of pixels
Radius	Radius of largest ellipse that can be enclosed by the segment
Roundness	Segment roundness
LengthWidth	Segment length / segment width
Border_grn	The proportion of the segment that borders other tree (green) segments
Border_unc	The proportion of the segment that borders unclassified (not green) segments

2.5 Classification

2.5.1 Optimal size and cover variables

To improve the usage of the field data of the size and percentage cover in the classification model an optimisation process is undertaken. The sandalwood trees identified in the field were attributed with canopy diameter (east/west and north/south) and the percentage coverage of other species. To determine the optimal diameter (using the east/west and north/south minimum) and percent cover of other species overlapping the sandalwood canopy the model was run over every combination of these measures.

Minimum diameter values were trialled at 0.25 m increments from 0.25 to 4 m, and percent cover values were trialled at 10 % increments from 10 to 100. As the random forest model produces slightly variable results each time it is run the process was repeated 10 times and the mean kappa value recorded. Kappa is a measure of how well the classifier performs compared to what can be achieved by chance. This process was run in R with the model running a total of 1600 times.

2.5.2 Modelling methodology

To rigorously assess the accuracy of the model 85 sandalwood samples were withheld by FPC for independent an accuracy assessment. The remaining sandalwood samples (n = 171) were filtered to remove those below the optimal size (0.75 m) and with greater than the optimal cover (20 %) (n = 104). The sandalwood samples were attributed as “sw” and combined with the non-sandalwood samples (n = 272), which were attributed as “other”. This “point samples” file was then spatially intersected with image segments created in eCognition. As each segment was attributed with the variables listed in section 2.4, the result was a “samples” file with sandalwood and other species each with list of associated variables.

The samples file was then split using random selection into training (60%) and testing (40%). The random forest model from the ranger package in R (Wright and Ziegler 2017) was run on the training data. The random forest model was set to produce a probabilistic value for each image segment, then the accuracy and number of predicted sandalwood at each accuracy threshold from 20 % to 80 % was calculated. This process was repeated in a loop 100 times and the results averaged and graphed.

To assess the influence of variables in the process with each iteration a “variable importance plot” was calculated. This process was repeated for each segmentation version, as listed in section 2.3.

3 Results

3.1 Analysis of variables

The degree to which sandalwood is different from other species can be explored using boxplots (Figure 4). With the majority of variables there is a large degree of overlap between sandalwood and other species. The variables with distributions that appear to show some separation include border_grn and border_unc (which are inversely related) and Max_nsm. The differentiation in the “border” variables indicates that sandalwood are more likely to be isolated (not touching other plants). The difference in the Max_nsm (height) variable indicates that sandalwood are generally smaller and occupy a smaller range of height values. However, consideration must be given as to whether bias in sampling “other” species may have contributed to these differences.

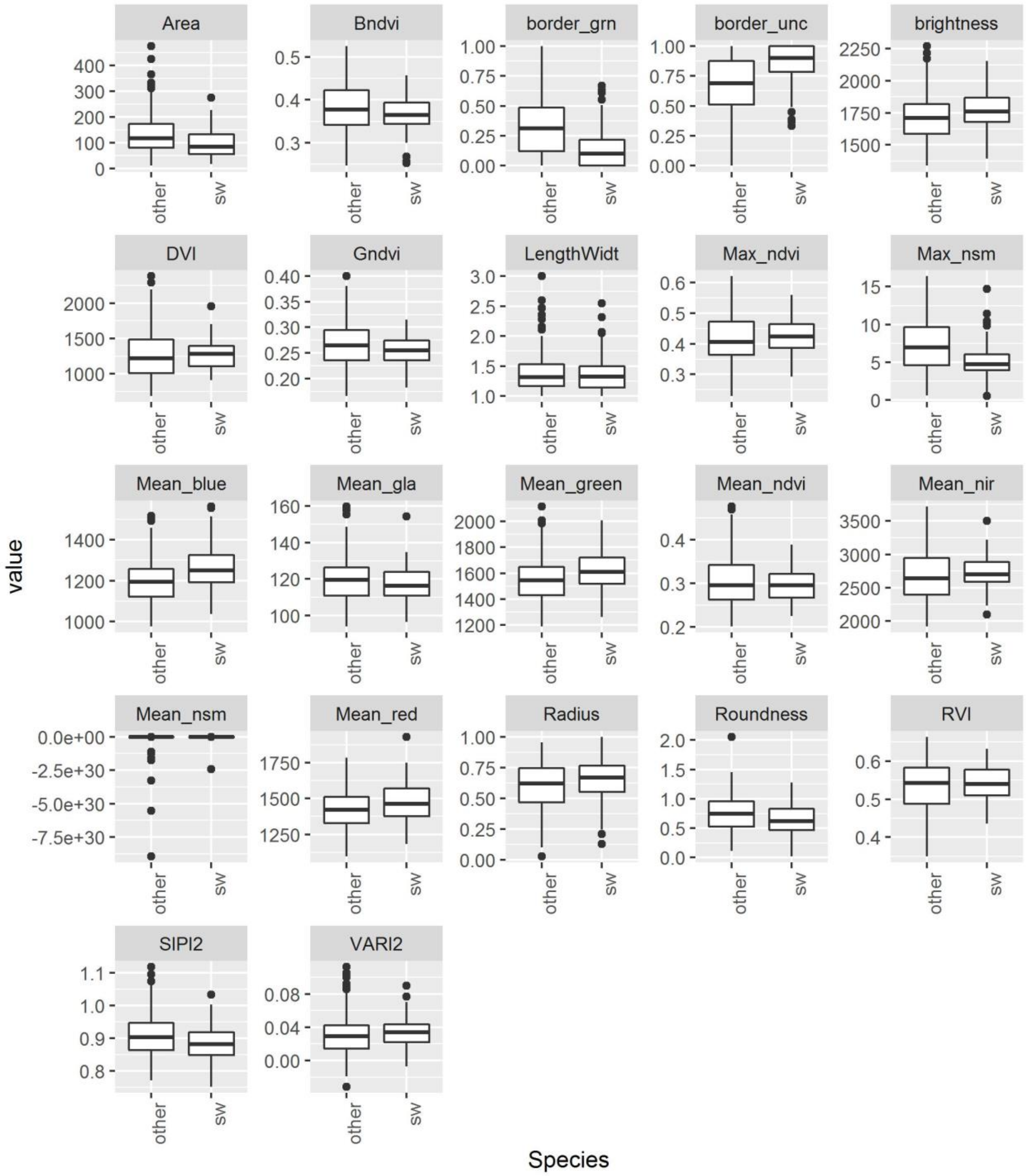


Figure 4: Boxplots of sandalwood (sw) and "other" species for all variables.

The correlation between field measures of height and height values from the NSM (Max_nsm) is examined in Figure 5. The majority of points appear to follow a line close to one to one. However, there are still a large number of points where NSM values are significantly higher than field measures. These errors result in a poor correlation ($r^2 = 0.164$).

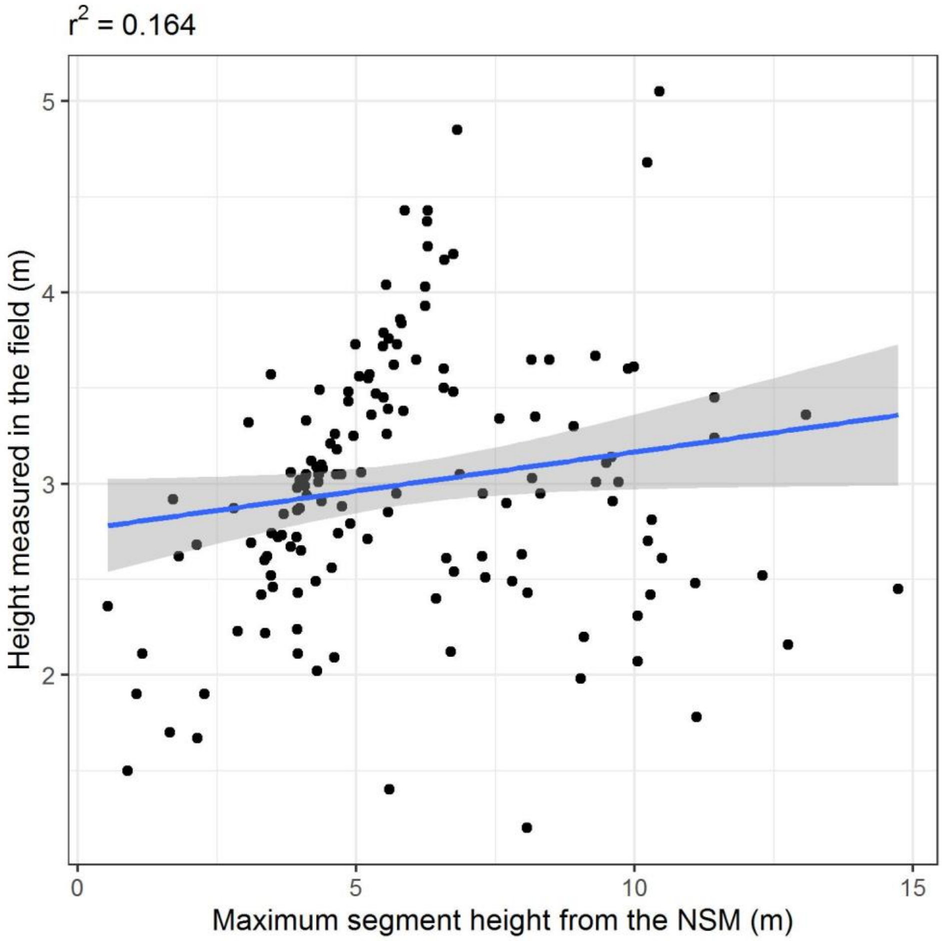


Figure 5: Correlation between height measures in the field and maximum segment height from the NSM.

The correlation between spectral index variables used in the analysis is shown in Figure 6. The colour indicates the direction and tone and size indicate the strength of the correlation. The plot indicates that many of the variables are highly correlated.

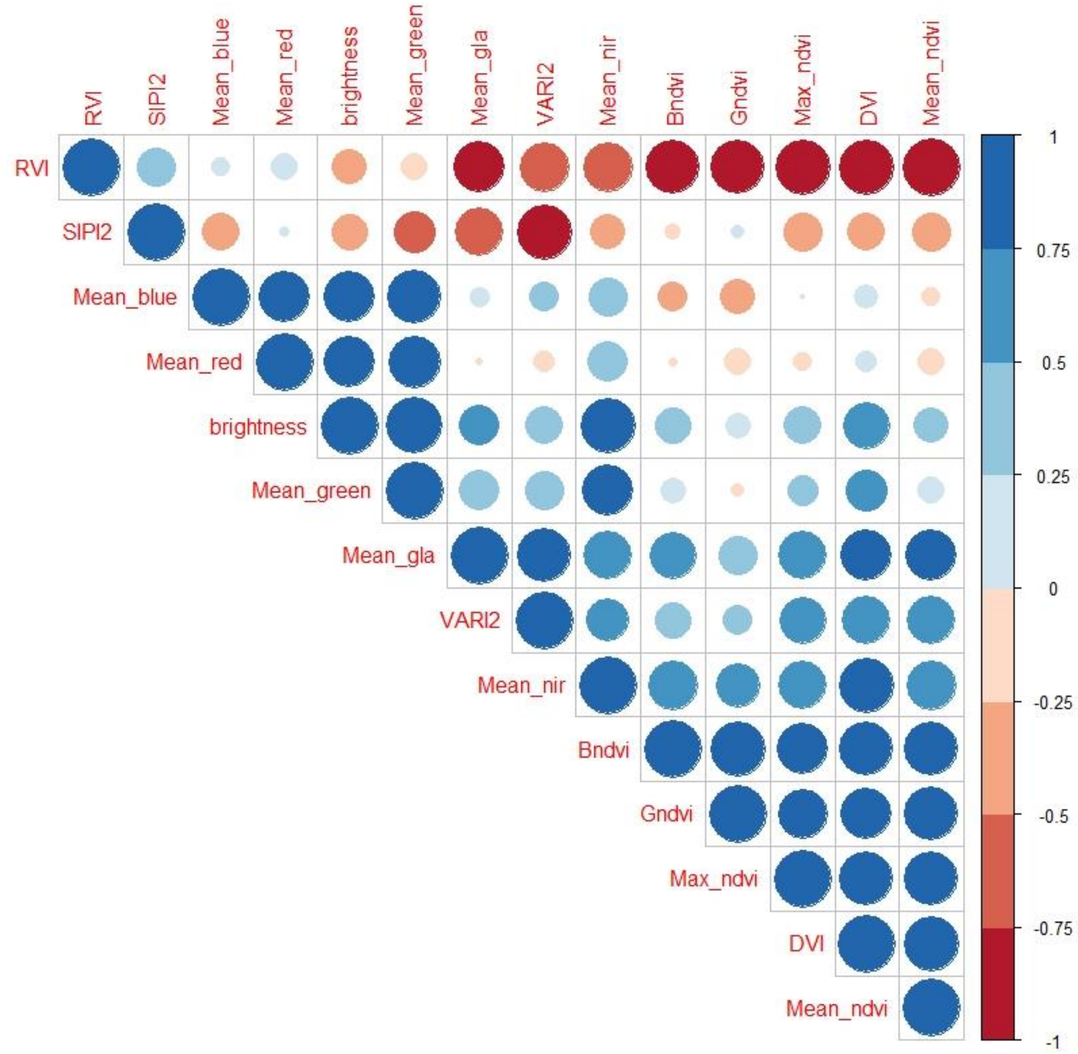


Figure 6: Correlation between spectral index variables.

The correlation between size and position variables used in the analysis is shown in Figure 7. The correlation between most variables is low (< 0.5).

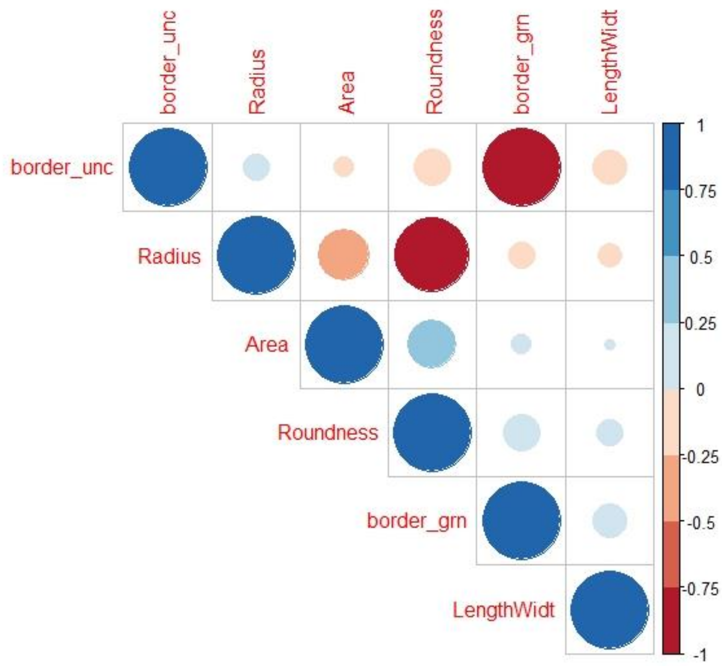


Figure 7: Correlation between size and position variables.

3.2 Optimal size and cover

The results from the assessment of optimal crown diameter and cover are shown in Figure 8. The highest kappa value (0.59) was achieved by only including sandalwood samples with a canopy diameter greater than 0.75 m and cover from another species less than 20 %. When samples that do not meet these criteria are excluded 104 sandalwood samples remain (Figure 9 and Figure 10).

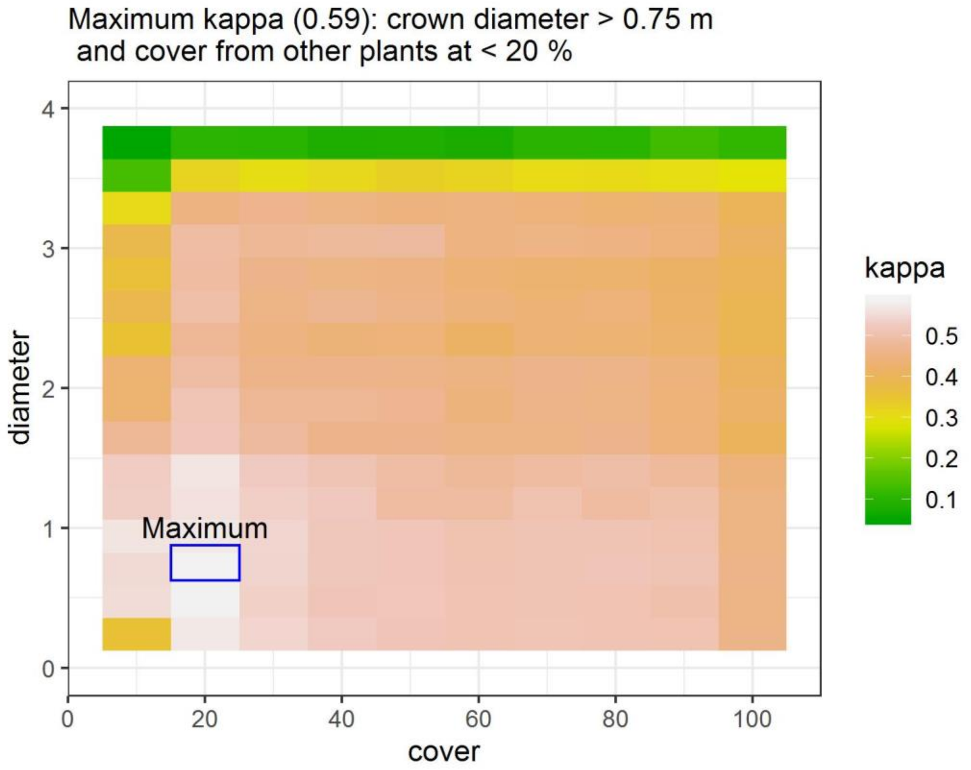


Figure 8: Kappa scores for cover values 10 % to 100 % and crown diameter 0.25 to 4 m.

While it may appear logical that including only the largest sandalwood with no cover would produce the best results, this means only a small number of samples being available for modelling and this can adversely impact the classification result.

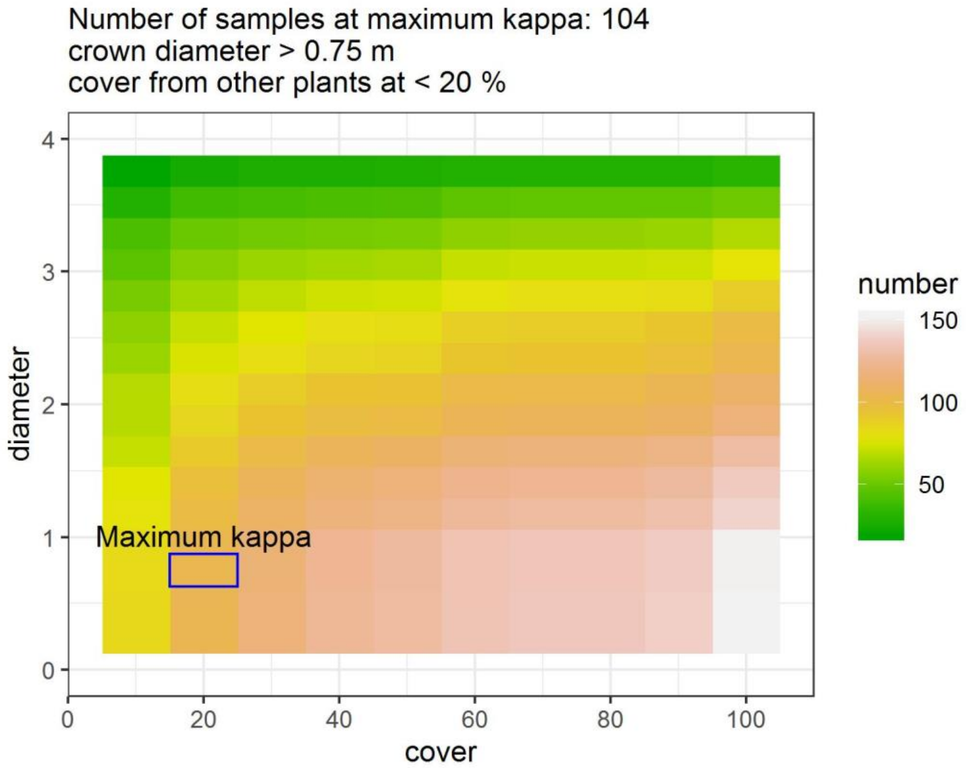


Figure 9: The number of samples available at each canopy diameter/cover percentage variant.

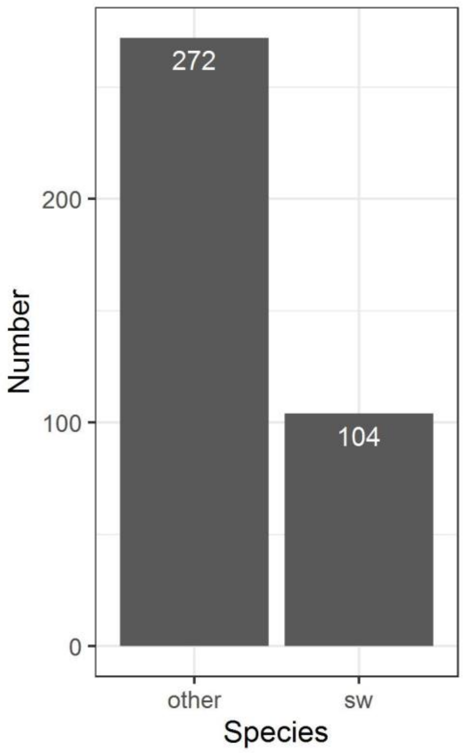


Figure 10: The number of sandalwood (sw) and "other" species samples used for modelling.

3.3 Classification and predictions

Measures of kappa, accuracy, sensitivity and specificity along with the number of predicted sandalwood were used to assess the accuracy of the versions trialled. Kappa is a measure of how well the classifier performs compared to what can be achieved by chance. Sensitivity refers to the proportion of sandalwood samples classified as sandalwood (also known as recall or true positive rate). Specificity refers to the proportion of other species classified as other (also known as true negative rate).

The number of predicted sandalwood is perhaps the best (and most relevant) measure of the models performance. This can be compared to the number of known sandalwood in the study area (n = 257).

The best classification was achieved using the watershed segmentation at a scale of 5, with the GLA index smoothed with a 5 x 5 median filter. The accuracy and prediction numbers for this version are shown in Figure 11. The figure shows how the accuracy measures, and predicted number of sandalwood in the study area, vary as the probability of a plant being a sandalwood is varied. When a plant is classified as being a sandalwood only if it has a probability greater than around 0.42 achieves the highest kappa value but results in 1563 sandalwood being mapped in the study area. Increasing the probability threshold will lower the number of sandalwood predicted but also lower the classification accuracy.

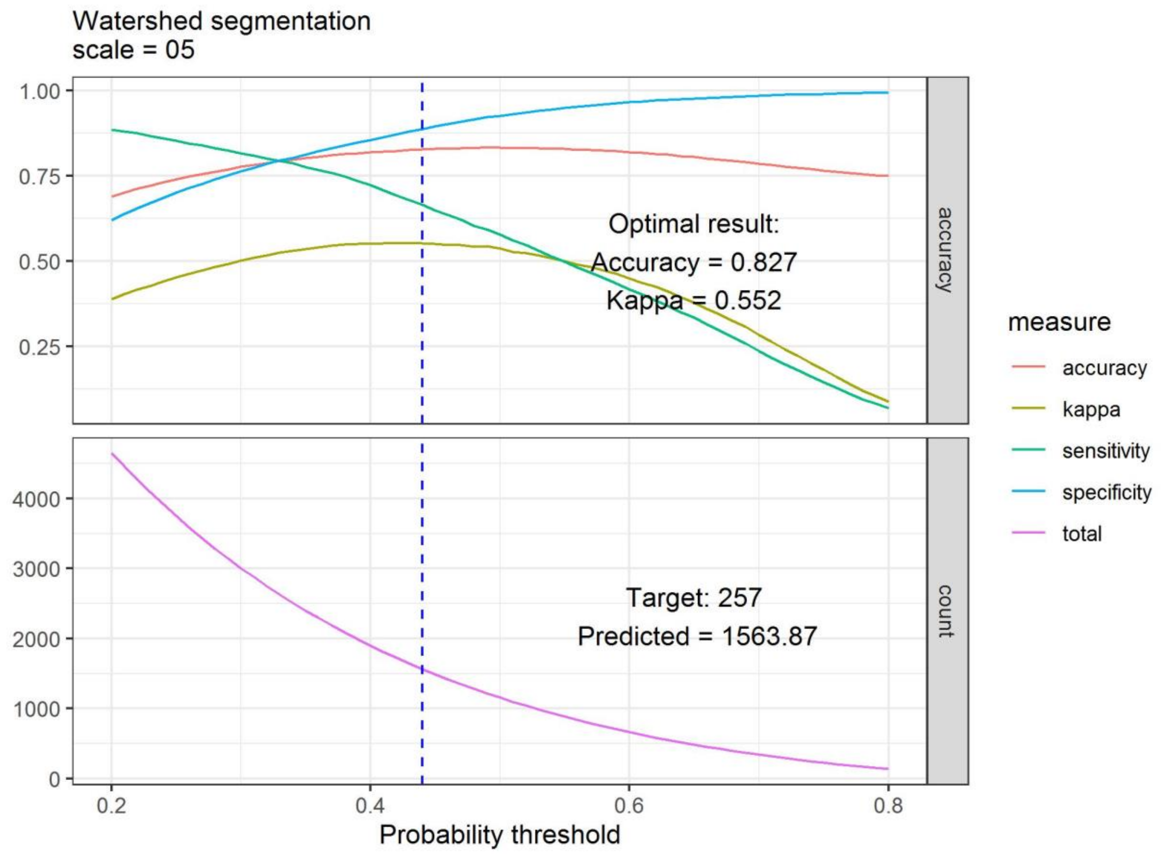


Figure 11:Accuracy and prediction numbers from the watershed segmentation at a scale of 5, using the GLA with a median 5x5 filter.

A variable importance plot is shown in Figure 12. This figure shows the relative influence of variables in the random forest model. Height (Max_nsm) is by far the most influential variable followed by the positional variable border_unc, which is a measure of the proportion of the plant that borders another. The mean of the blue band and DVI are the most influential indices. The shape variables (Roundness, Radius and LengthWidth) have the lowest influence.

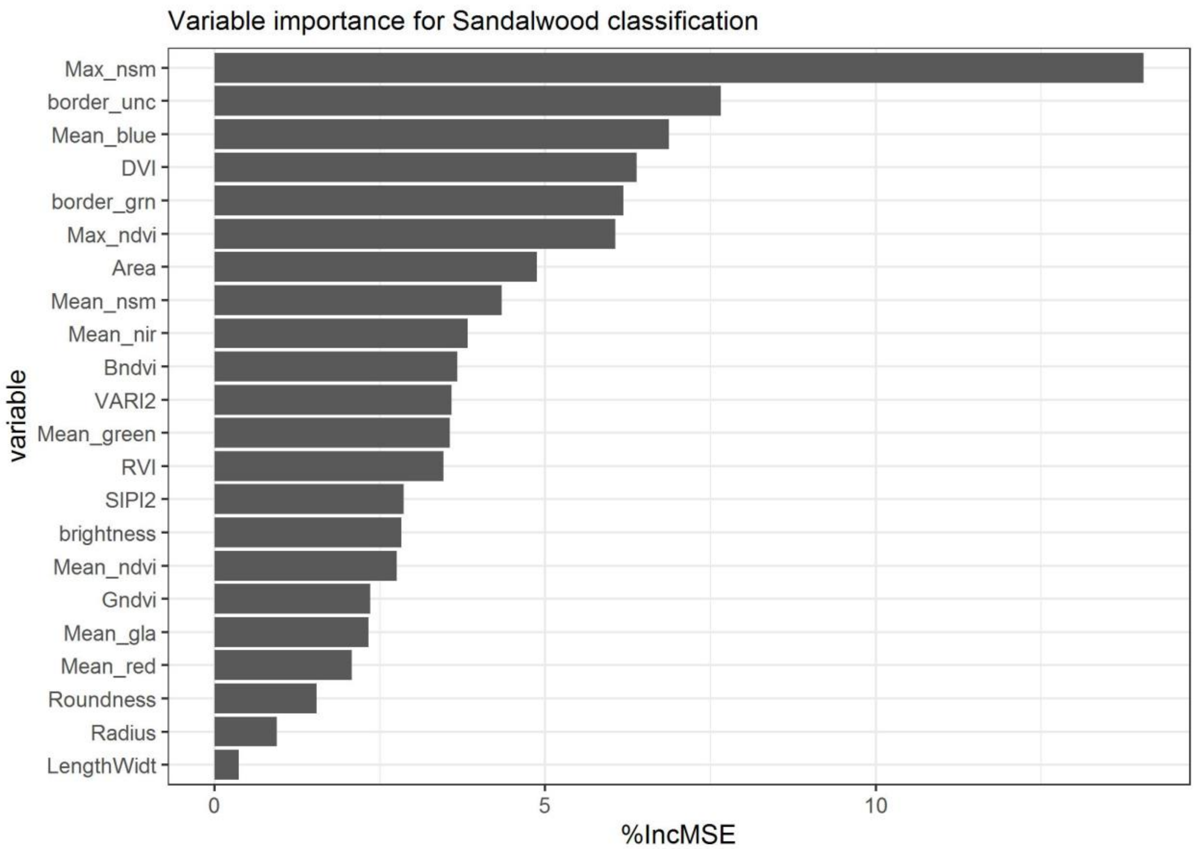


Figure 12: A variable importance plot from the watershed segmentation at a scale of 5, using the GLA with a median 5x5 filter.

Accuracy and prediction numbers for all other variations are included in the appendix. A summary of all version results are shown in Table 3.

Table 3: A summary of results from each segmentation version including accuracy, kappa and predicted sandalwood numbers.

Segmentation	scale	Layer / filter	Accuracy/kappa	N. predicted
Watershed	10	25 cm, green 5 x 5	0.762/0.389	2479
Watershed	20	25 cm green 5 x 5	0.738/0.34	1982
Watershed	30	25 cm green 5 x 5	0.742 /0.342	1515
Watershed	40	25 cm green 5 x 5	0.747 /0.328	1374
***Watershed**	**5**	**25 cm gla 5 x 5**	**0.827/0.552**	**1563**
Watershed	10	25 cm gla 5 x 5	0.84/0.581	1720
Watershed	20	25 cm gla 5 x 5	0.822 /0.566	1982
Watershed	30	25 cm gla 5 x 5	0.81 /0.546	1651
Watershed	40	25 cm gla 5 x 5	0.81 /0.553	1592
Watershed	10	25 cm ndvi 5 x 5	0.787 /0.416	2073
Watershed	20	25 cm ndvi 5 x 5	0.784/0.433	1711
Watershed	30	25 cm ndvi 5 x 5	0.74 /0.341	1713
Watershed	40	25 cm ndvi 5 x 5	0.758 /0.339	1331
Multi-resolution	10	25 cm, gla 5 x 5	0.78/0.48	33879
Multi-resolution	20	25 cm, gla 5 x 5	0.797/0.472	11508
Multi-resolution	30	25 cm, gla 5 x 5	0.78/0.443	8556
Multi-resolution	40	25 cm, gla 5 x 5	0.75/0.365	6816
Multi-resolution	50	25 cm, gla 5 x 5	0.721/0.34	7787
Multi-resolution	60	25 cm, gla 5 x 5	0.719/0.335	7275
Multi-resolution	70	25 cm, gla 5 x 5	0.738/0.298	5809
Multi-resolution	80	25 cm, gla 5 x 5	0.701/0.294	6834
Multi-resolution	90	25 cm, gla 5 x 5	0.704/0.263	5939
Multi-resolution	100	25 cm, gla 5 x 5	0.711/0.268	5790
Watershed	40	5 cm, gla 11 x 11	0.795/0.496	9706
Watershed	50	5 cm, gla, 11 x 11	0.77/0.455	6757
Watershed	60	5 cm, gla, 11 x 11	0.744/0.406	4965
With training data as polygons				
Watershed	**5**	**25 cm gla 5 x 5**	**0.847/0.595**	**2030**
Watershed	**10**	**25 cm gla 5 x 5**	**0.84/0.581**	**1709**
Watershed	**20**	**25 cm gla 5 x 5**	**0.822/0.566**	**1720**
Watershed	**30**	**25 cm gla 5 x 5**	**0.81/0.546**	**1651**
Watershed	**40**	**25 cm gla 5 x 5**	**0.81/0.553**	**1592**

***Best performing model**

3.4 Density surface

Density surfaces were also produced to examine whether zones of sandalwood occurrence could be identified. Density surfaces for actual sandalwood locations and predictions at threshold levels of 0.8 and 0.9 are shown in Figure 13. When only segments with high probabilities of being sandalwood are classified as sandalwood the spatial distributions are somewhat similar to the actual sandalwood distribution.

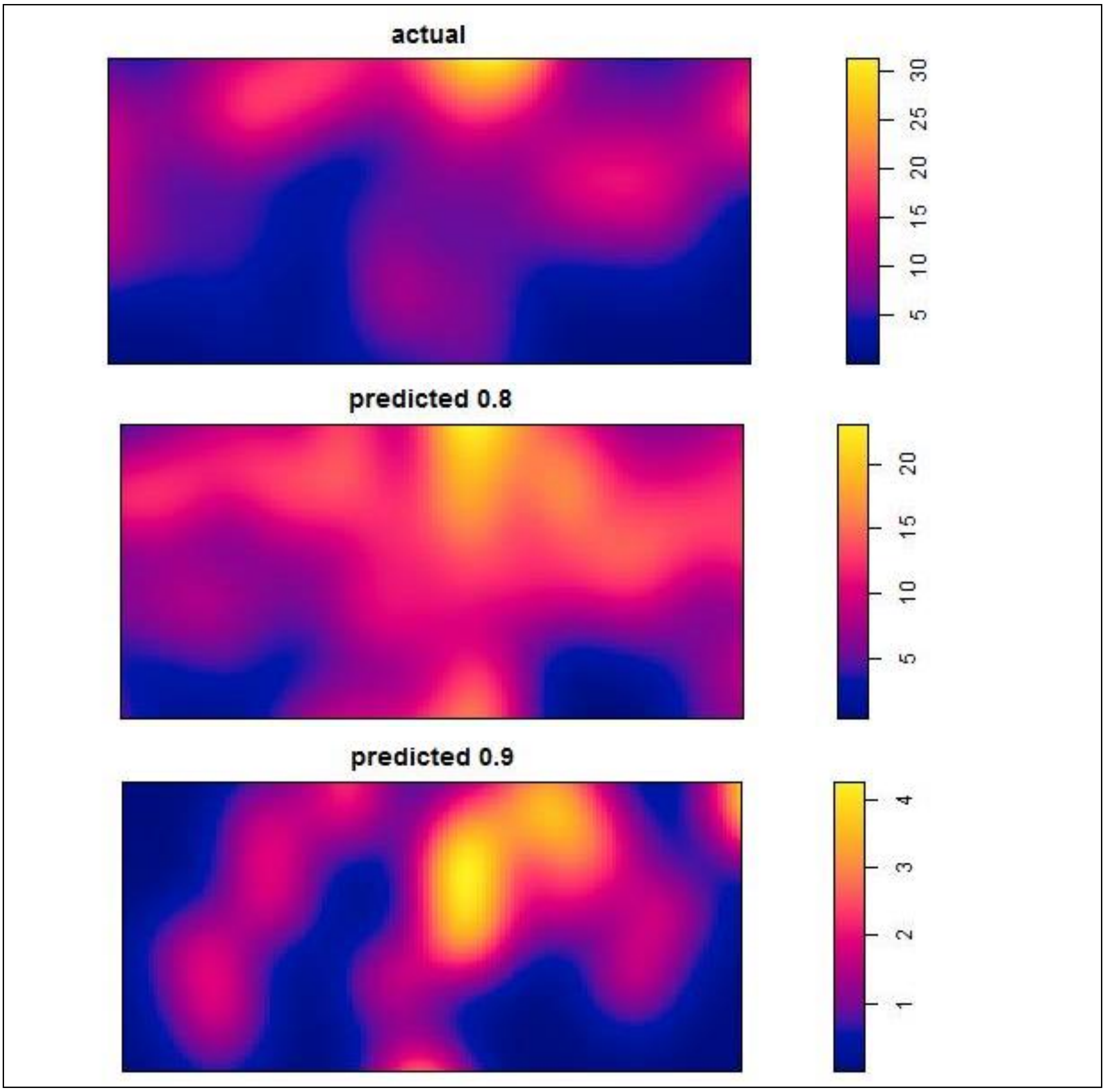


Figure 13: A sandalwood density surface, in stems per hectare, from field data along with predictions at 0.8 and 0.9 probability thresholds.

4 Discussion

Multispectral imagery at 0.25 m and 0.05 m was supplied for the project. The 0.05 m data appeared to be of limited value. Although segmentations based on the 0.05 m data achieved reasonably high classification accuracies (> 0.75), the smaller segmentation size resulted in a high number of predicted sandalwood (n > 6000).

The 0.25 cm data appeared to be a suitable resolution. Smoothing this data with a 5 x 5 median filter seemed affective at removing within crown variability while maintaining between crown definition. The watershed segmentation outperformed the multiresolution segmentation. Filtering applied to the GLA image, combined with the watershed segmentation, at a scale of 5 was the most effective method to delineate (produced small polygon) each individual plant.

Of the input data the consistency of the NSM was most concerning. Error in the NSM is evident in the regression between NSM values and the field measures of height (Figure 5). The error in the NSM is likely to have a significant impact on the achieved accuracy as it is the most influential variable in the random forest model.

Identifying sandalwood in the landscape using digital aerial photography presents a number of challengers. While at first appearance the classification accuracy greater than 80% is a good result. However, the application of this model results in an unacceptable number of false positives, with 1563 sandalwood predicted where only 257 were identified. The reason for the apparent discrepancy between the model accuracy and the large number of false positives is the difference between the proportion of sandalwood in the training data and the proportion of sandalwood in the population. With the training data around a third of the samples are sandalwood, whereas in the population only 257 of around 9000 (~ 2%) are sandalwood. This means that if only a small percentage of other species are incorrectly classified as sandalwood the number of false positives will be greater than the actual number of sandalwood. The low proportion of sandalwood in the landscape is a significant limiting factor to achieving an acceptable model. For any model to be effective it would have to achieve an accuracy close to 100%.

5 References

Macfarlane, Craig, and Gary N. Ogden. 2012. “Automated Estimation of Foliage Cover in Forest Understorey from Digital Nadir Images.” *Methods in Ecology and Evolution* 3 (2): 405–415. https://doi.org/10.1111/j.2041-210X.2011.00151.x.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rouse, J. W., R.H. Hass, J.A. Schell, and D.W. Deering. 1974. “Monitoring Vegetation Systems in the Great Plains with ERTS.” In *Proceedings Third Earth Resources Technology Satellite-1 Symposium*, 3010–17. SP-351. Greenbelt: NASA. https://ntrs.nasa.gov/search.jsp?R=19740022614.

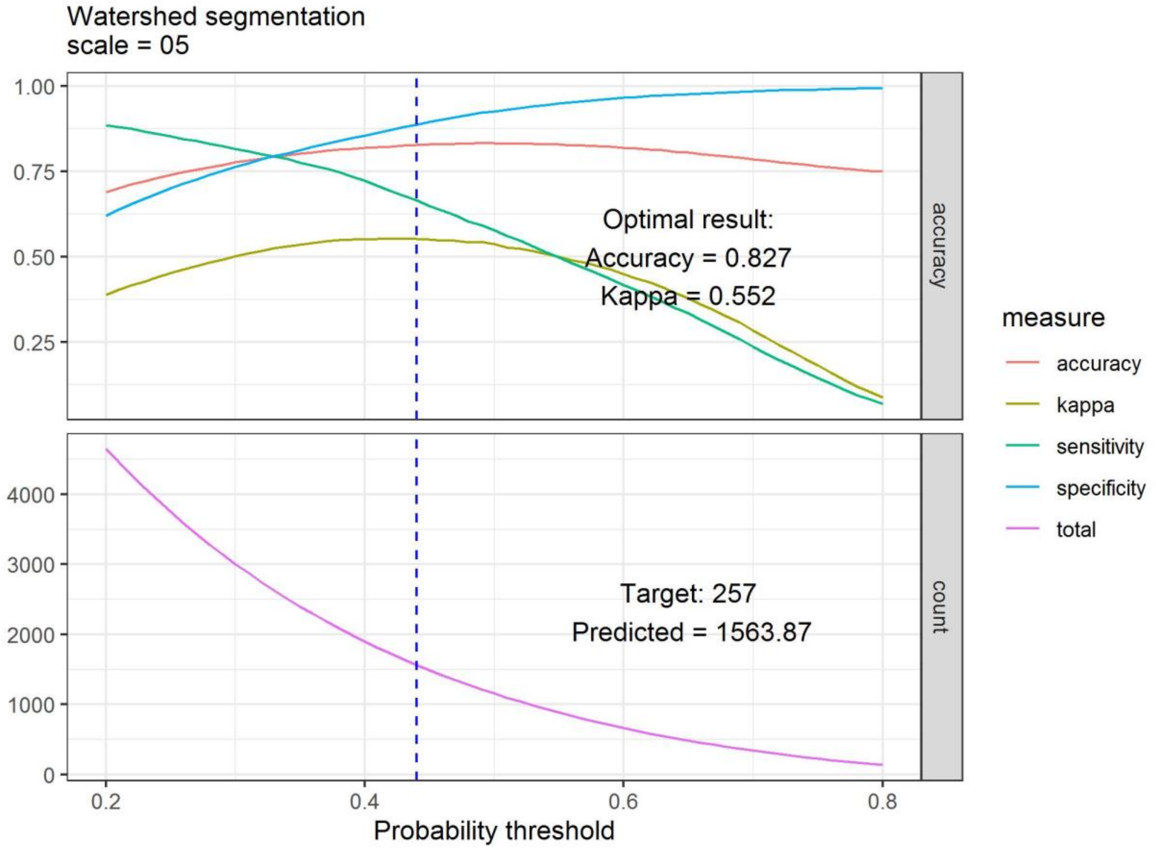
Thackway, Richard, and I. D Cresswell. 1995. *An Interim Biogeographic Regionalisation for Australia: A Framework for Setting Priorities in the National Reserves System Cooperative Program*. Canberra: Australian Nature Conservation Agency, Reserve Systems Unit.

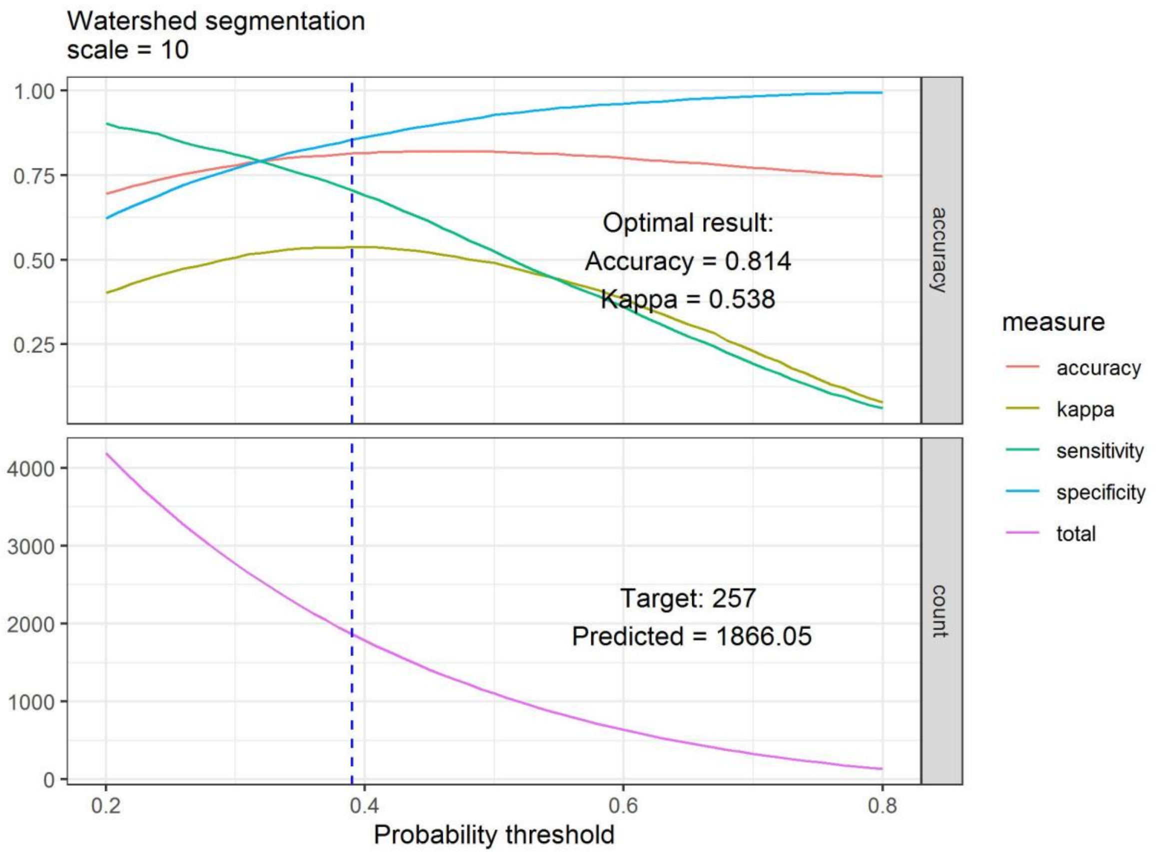
Wright, Marvin N., and Andreas Ziegler. 2017. “Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software* 77 (1): 1–17. https://doi.org/10.18637/jss.v077.i01.

Appendices

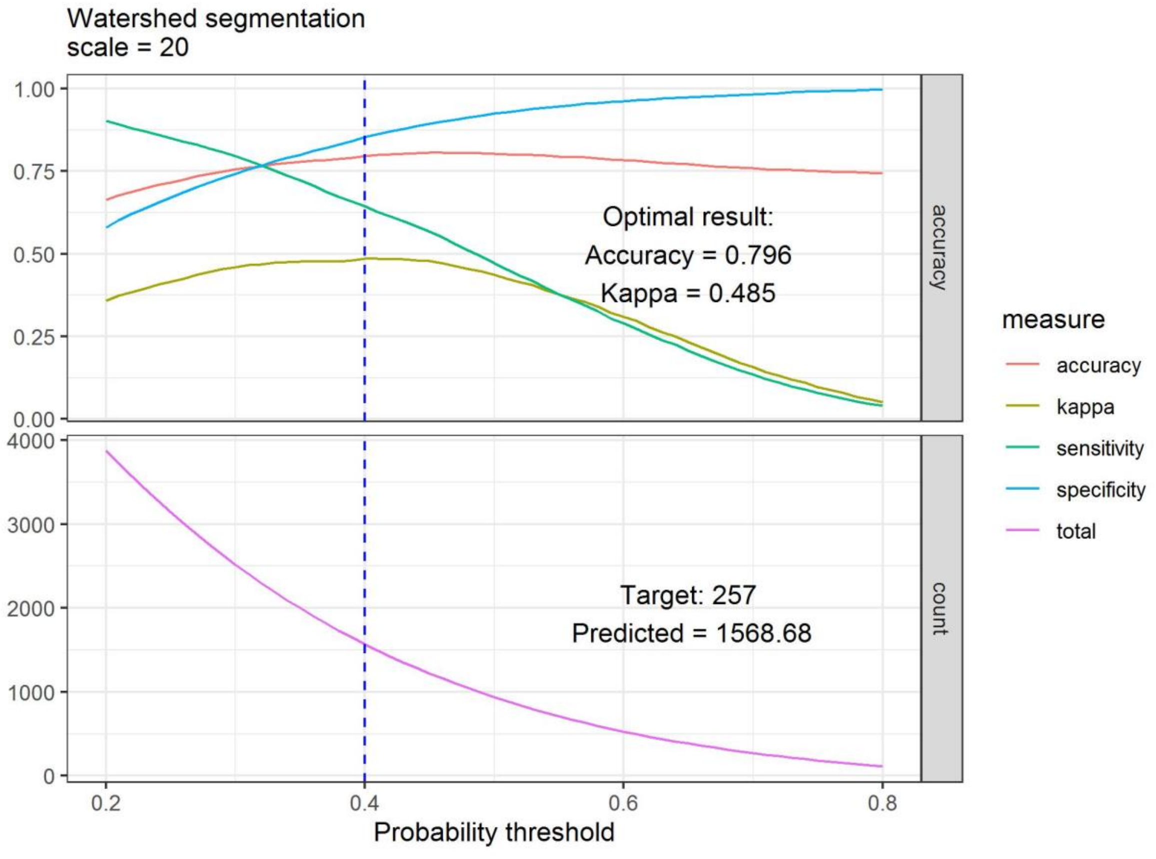
Trial 1

Using the 25 cm data, watershed segmentation at thresholds 5, 10, 20, 30 and 40 using the GLA, green layer and NDVI with a 5 x 5 median filter as the segmentation layers

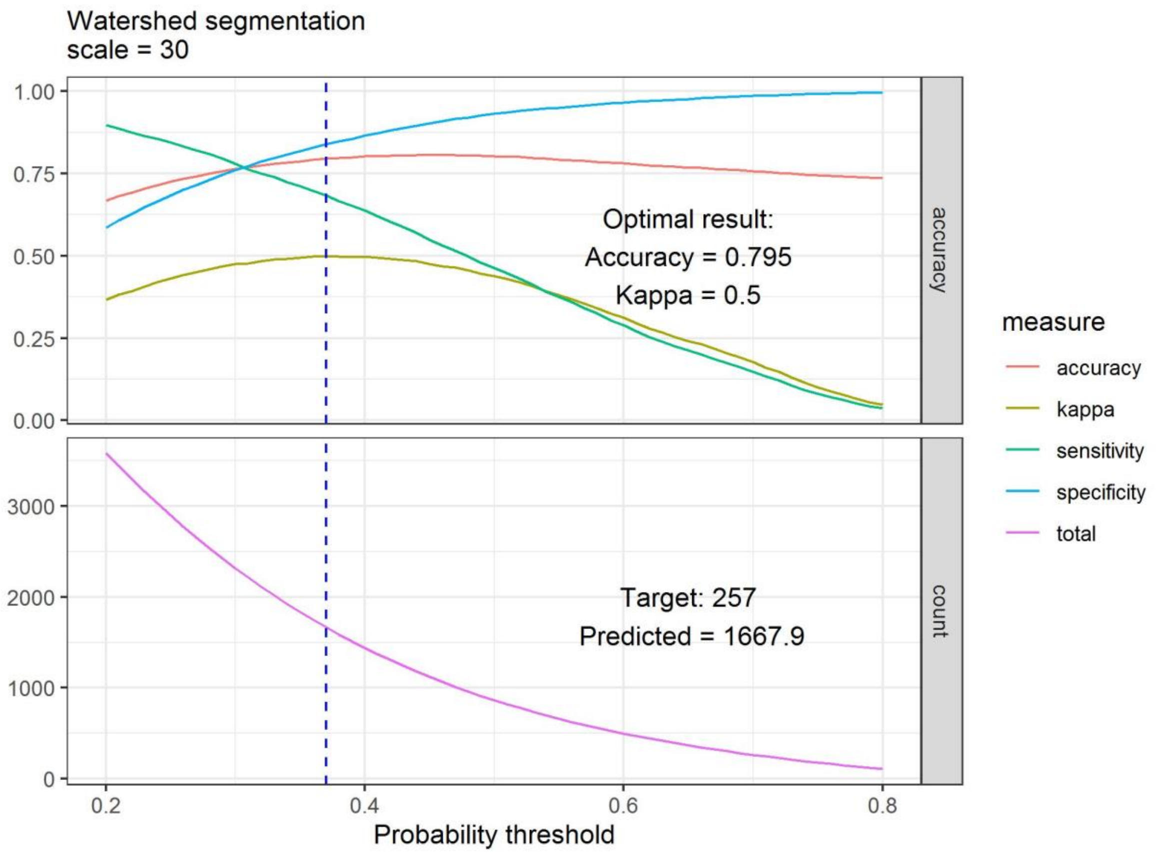


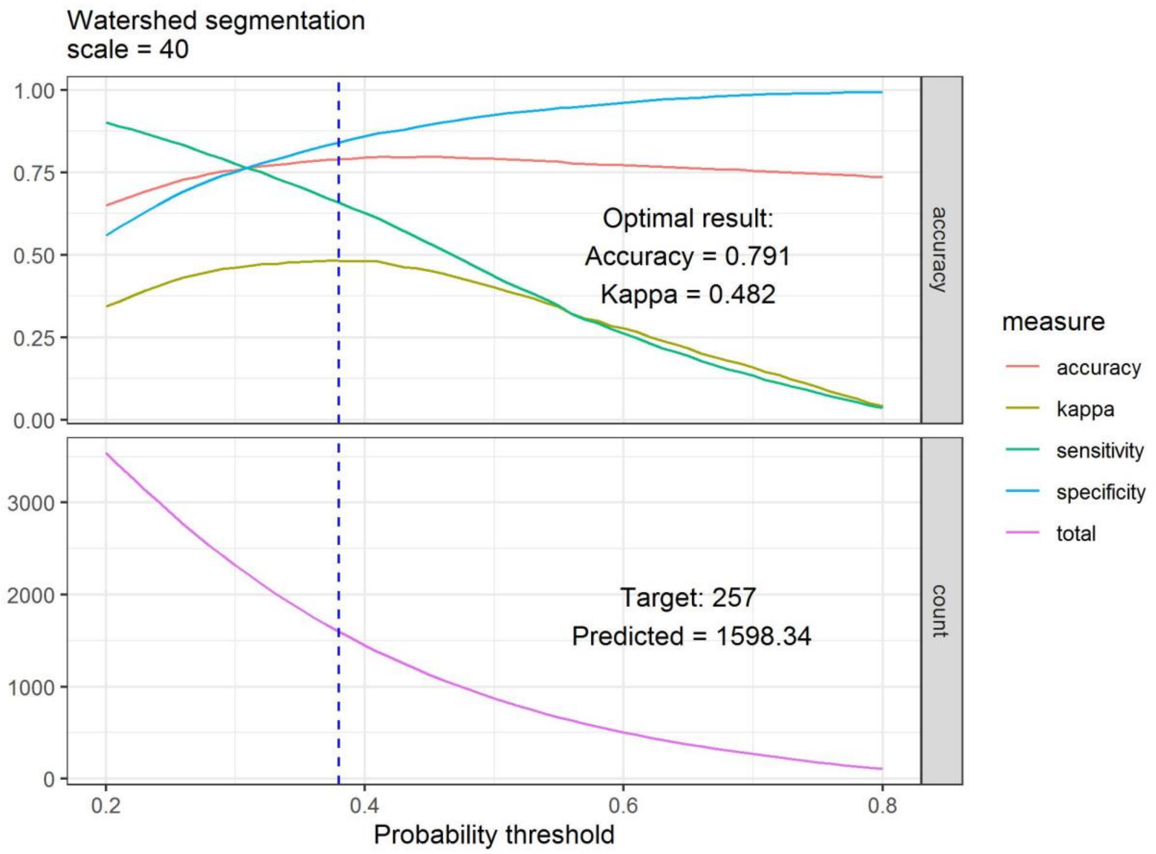


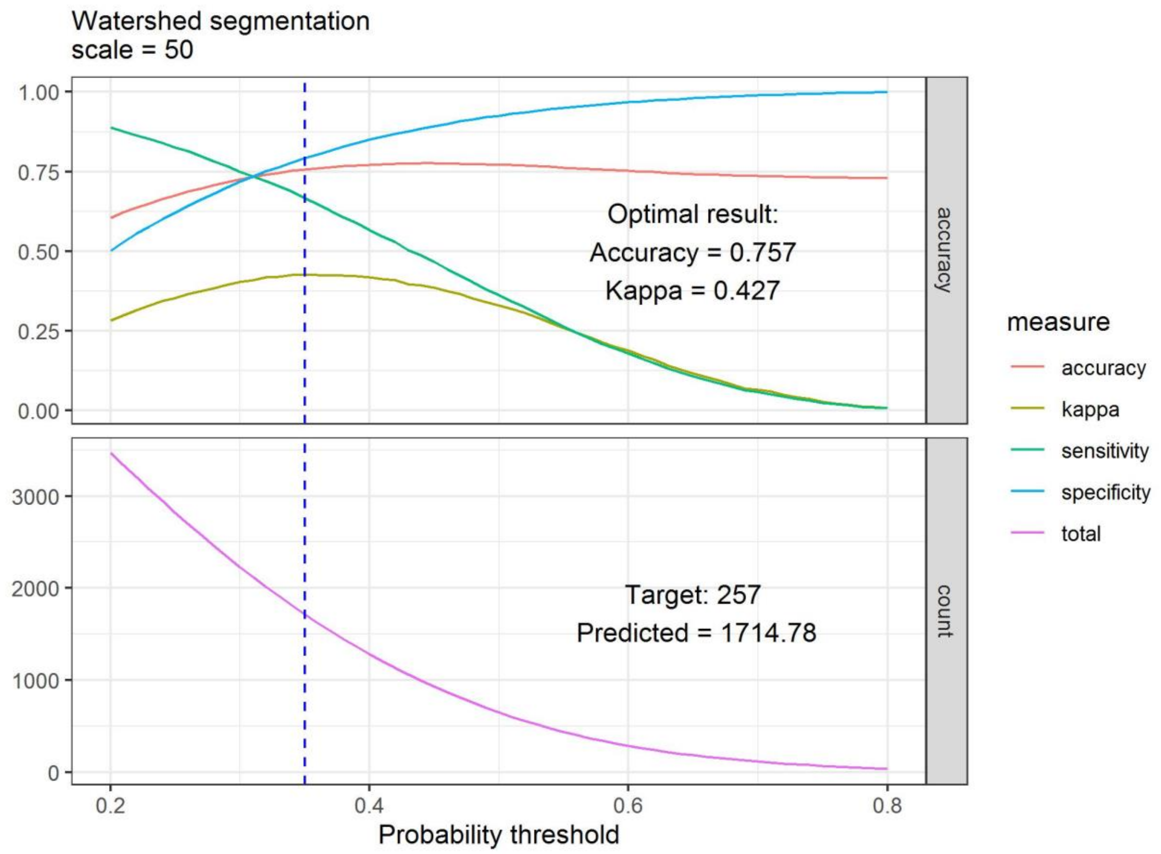
Watershed segmentation
scale = 10
Optimal result:
Accuracy = 0.814
Kappa = 0.538
Target: 257
Predicted = 1866.05
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

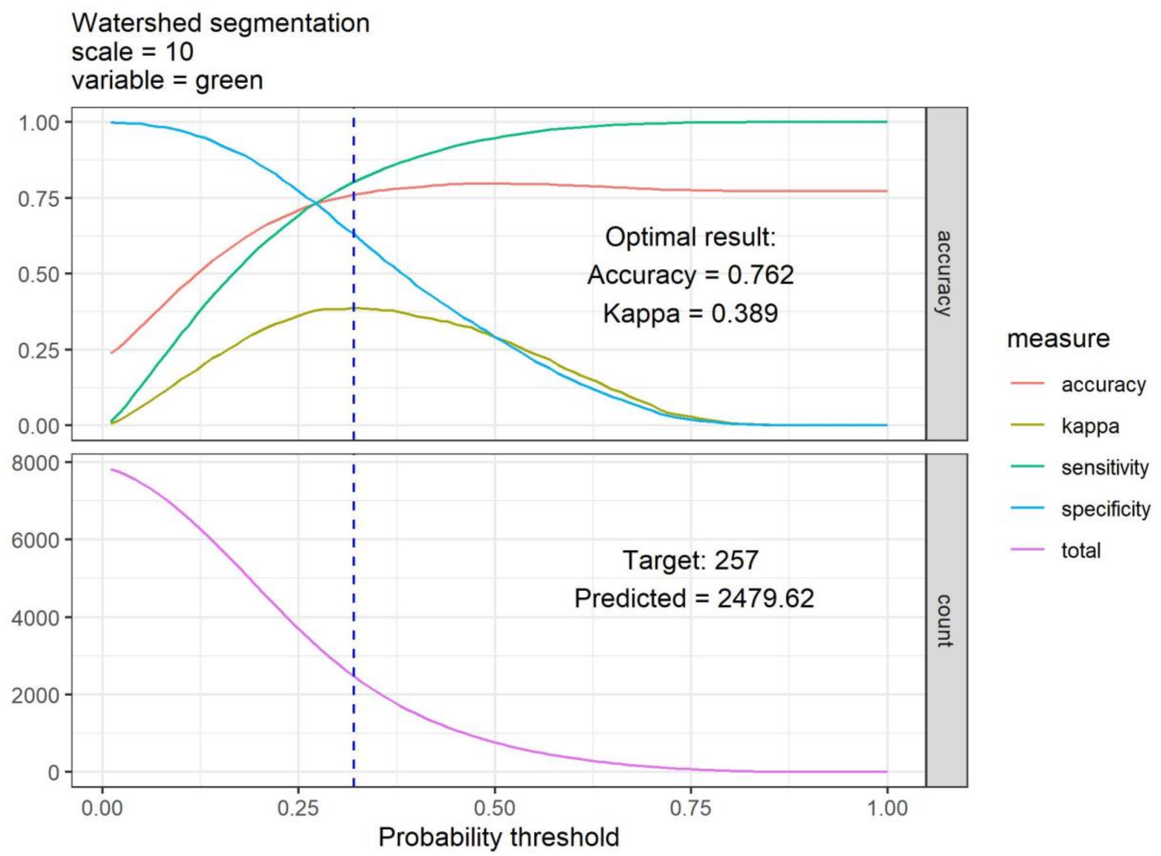


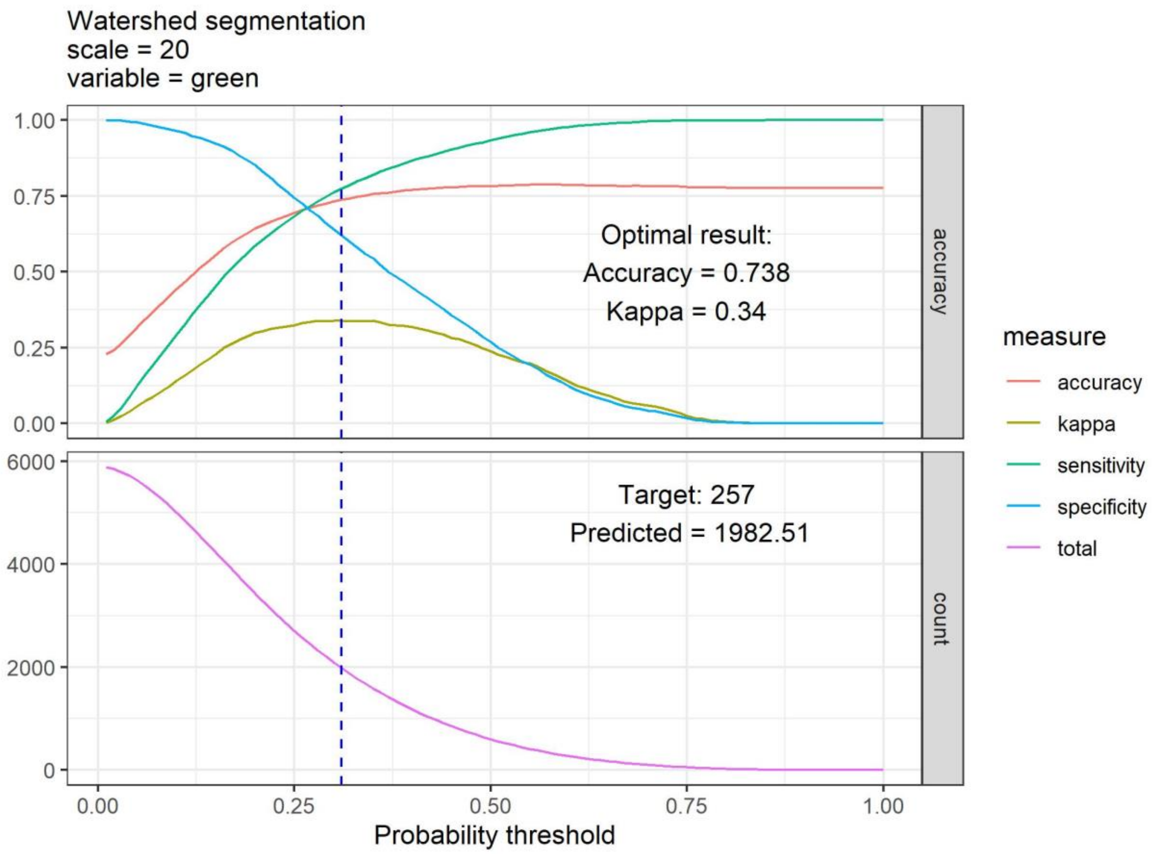
Watershed segmentation
scale = 20
Optimal result:
Accuracy = 0.796
Kappa = 0.485
Target: 257
Predicted = 1568.68
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

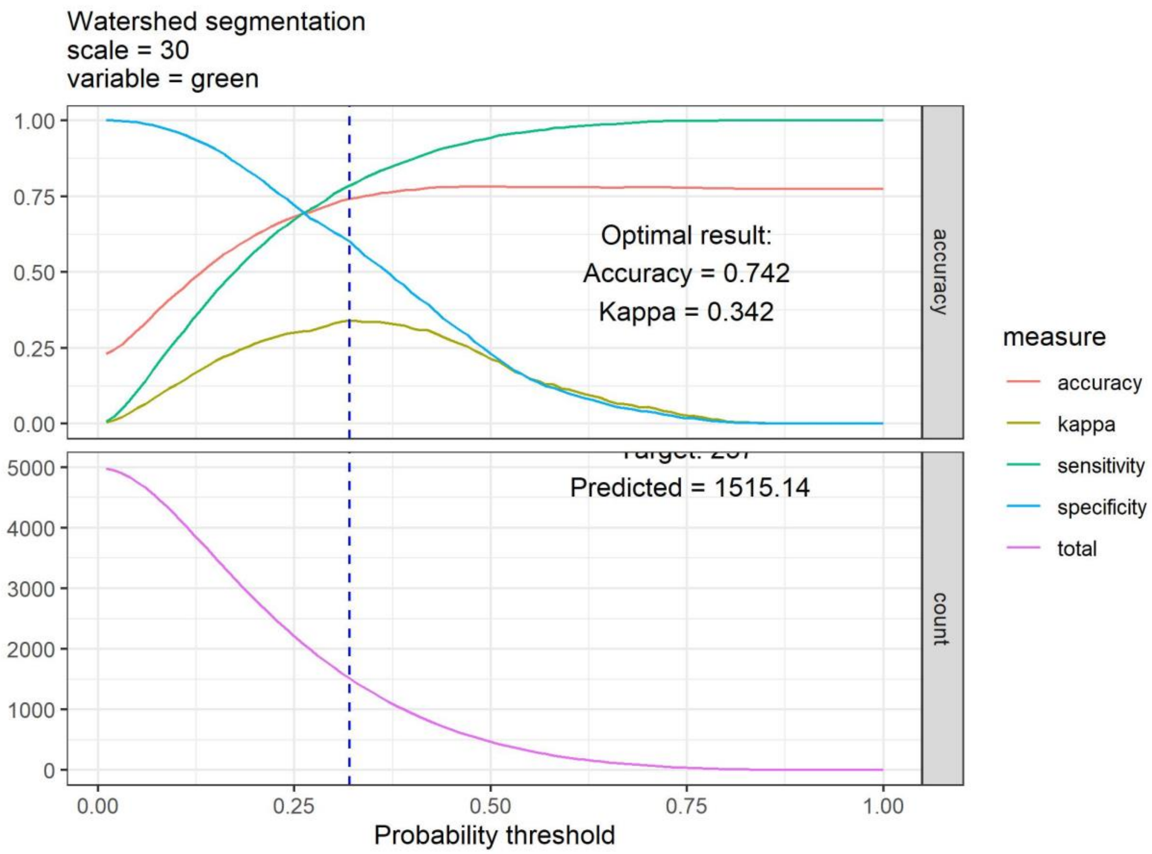

Watershed segmentation
scale = 30
Optimal result:
Accuracy = 0.795
Kappa = 0.5
Target: 257
Predicted = 1667.9
accuracy
count
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total

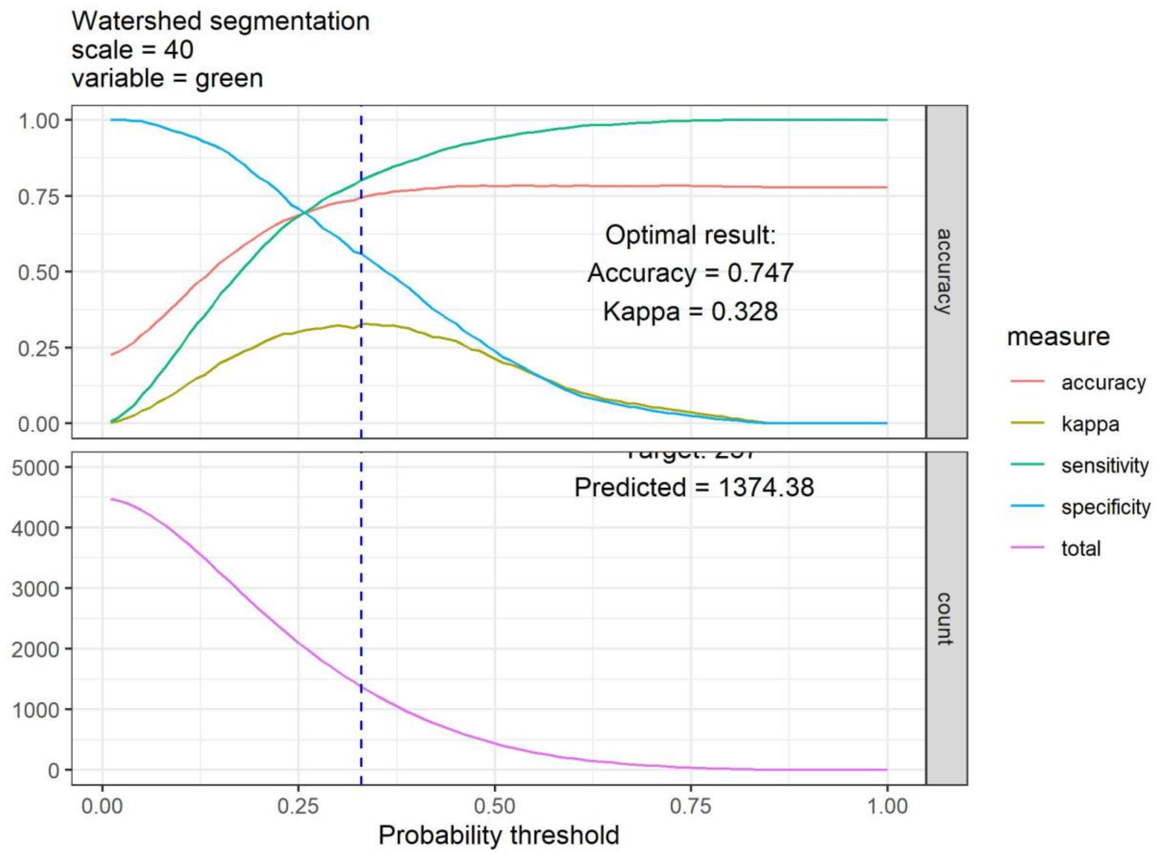

Watershed segmentation
scale = 40
Optimal result:
Accuracy = 0.791
Kappa = 0.482
Target: 257
Predicted = 1598.34
accuracy
count
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total

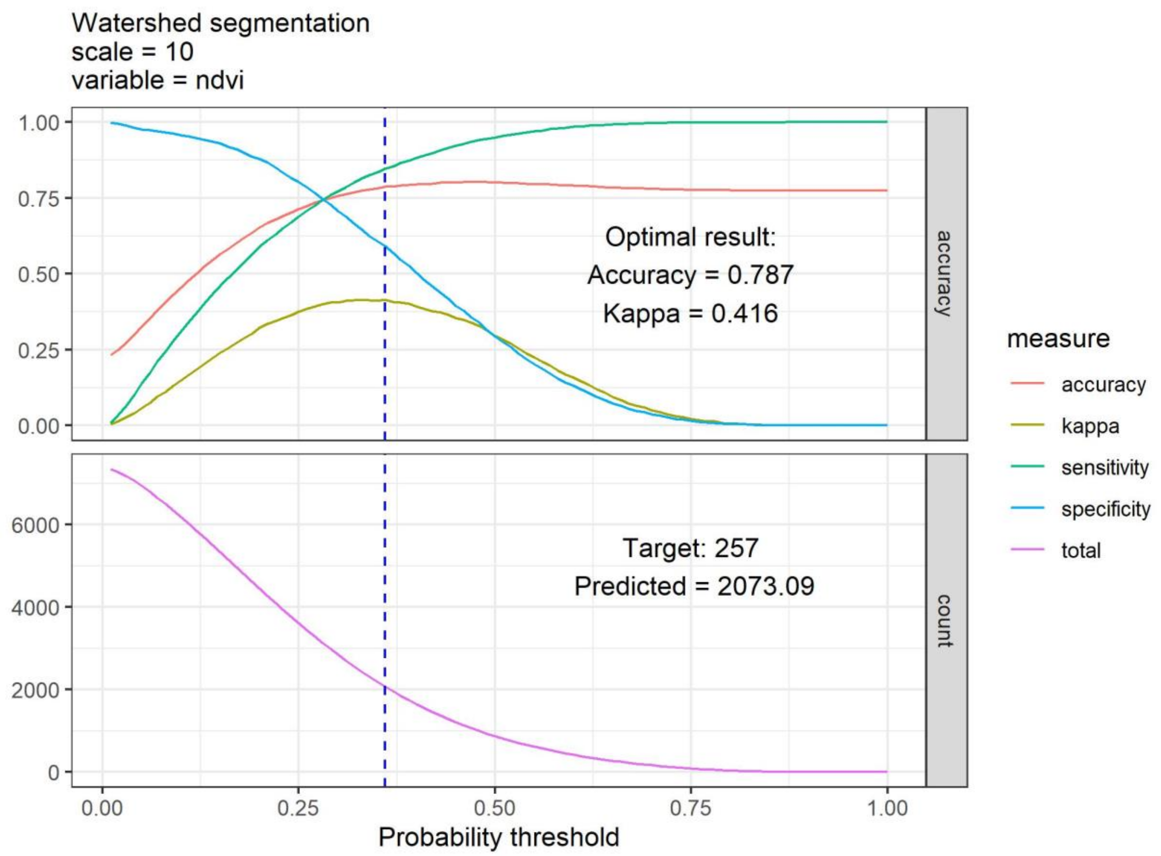

Watershed segmentation
scale = 50
Optimal result:
Accuracy = 0.757
Kappa = 0.427
Target: 257
Predicted = 1714.78
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

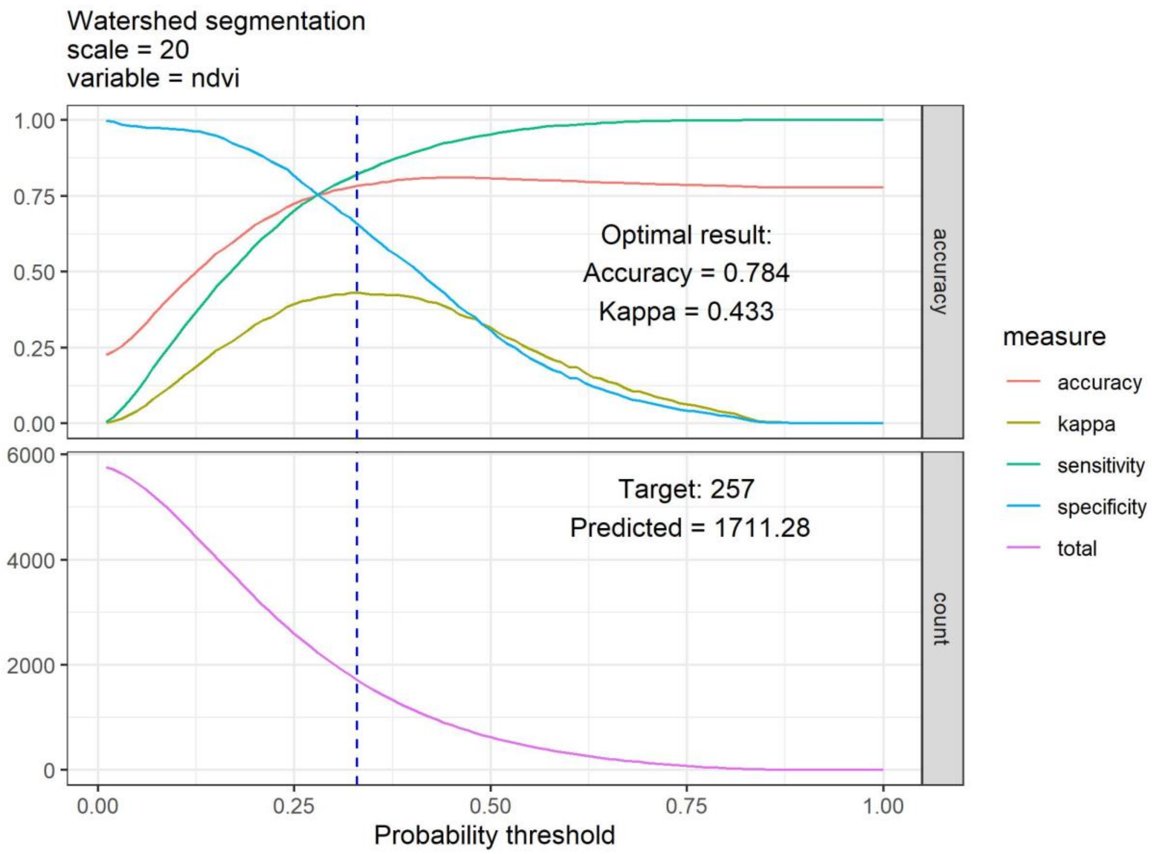

Watershed segmentation
scale = 10
variable = green
Optimal result:
Accuracy = 0.762
Kappa = 0.389
Target: 257
Predicted = 2479.62
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

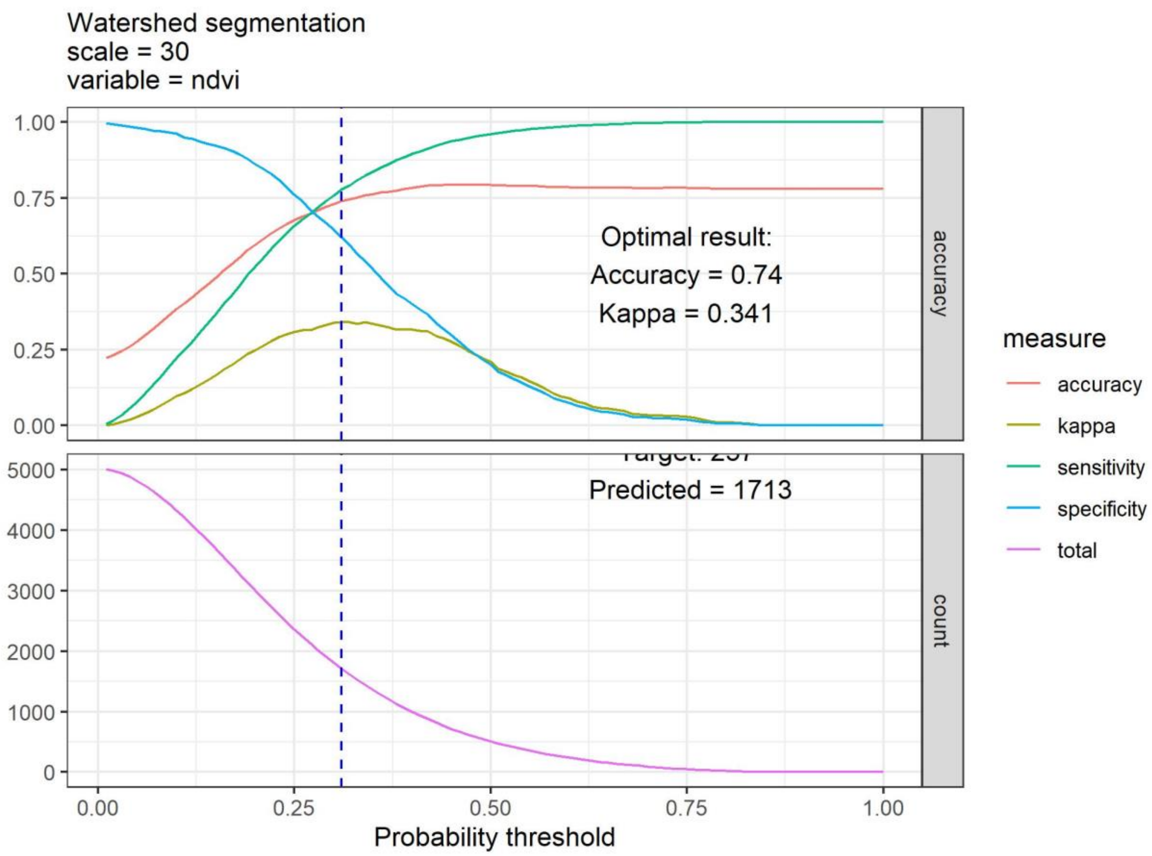

Watershed segmentation
scale = 20
variable = green
Optimal result:
Accuracy = 0.738
Kappa = 0.34
Target: 257
Predicted = 1982.51
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

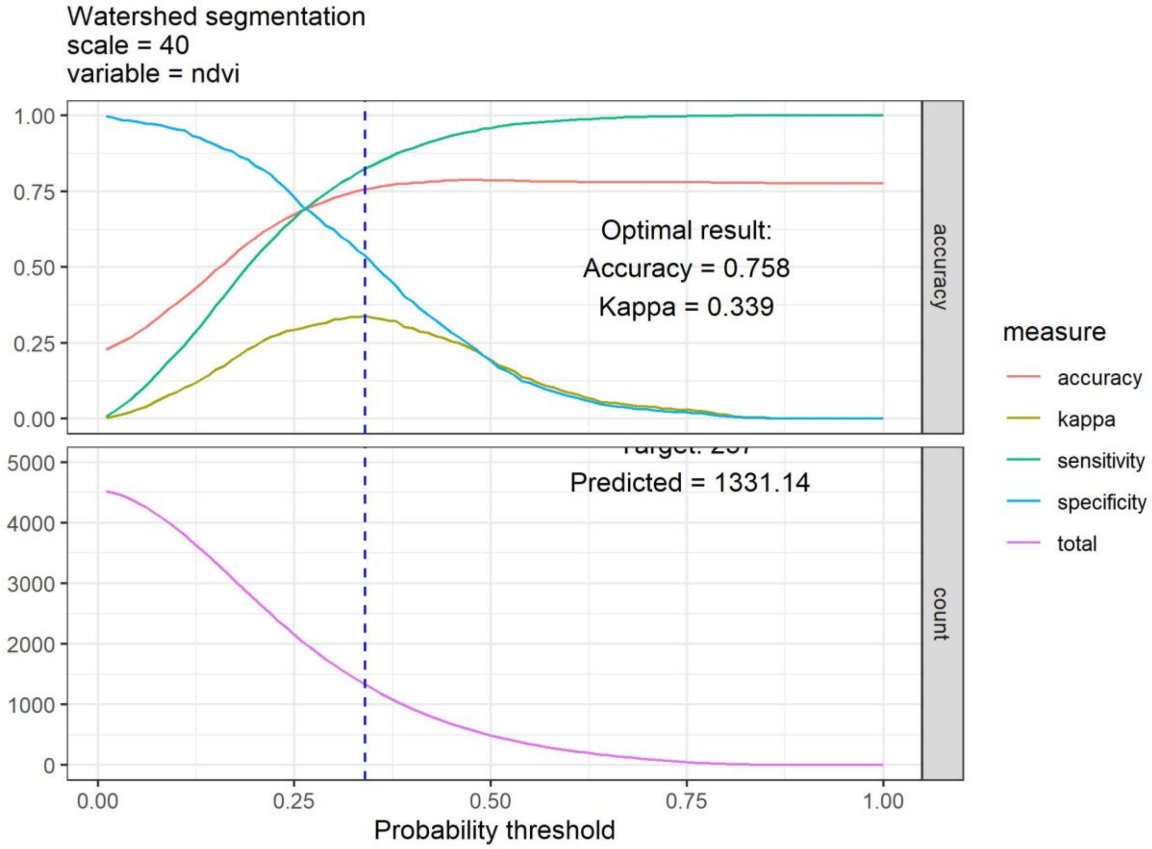

Watershed segmentation
scale = 30
variable = green
Optimal result:
Accuracy = 0.742
Kappa = 0.342
Predicted = 1515.14
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold


Watershed segmentation
scale = 40
variable = green
Optimal result:
Accuracy = 0.747
Kappa = 0.328
Predicted = 1374.38
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold


Watershed segmentation
scale = 10
variable = ndvi
Optimal result:
Accuracy = 0.787
Kappa = 0.416
Target: 257
Predicted = 2073.09
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

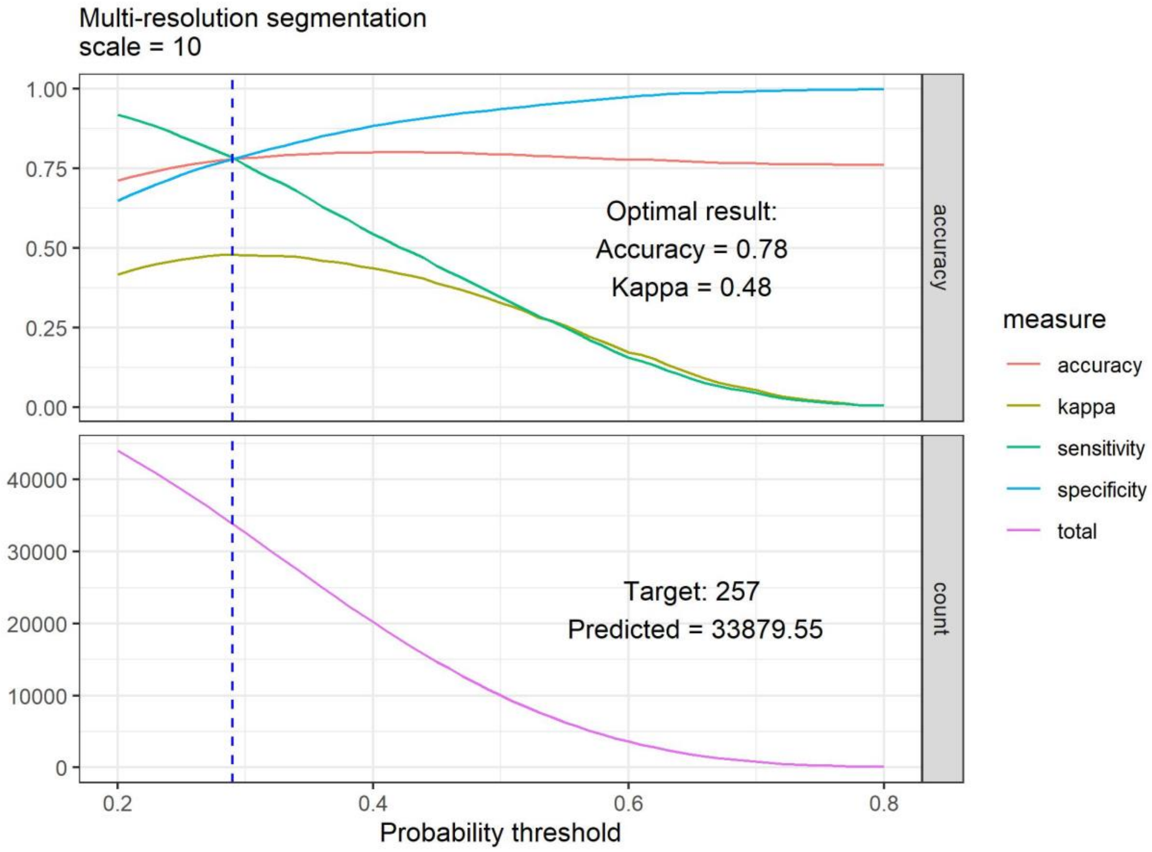

Watershed segmentation
scale = 20
variable = ndvi
Optimal result:
Accuracy = 0.784
Kappa = 0.433
Target: 257
Predicted = 1711.28
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold
1.00
0.75
0.50
0.25
0.00
6000
4000
2000
0

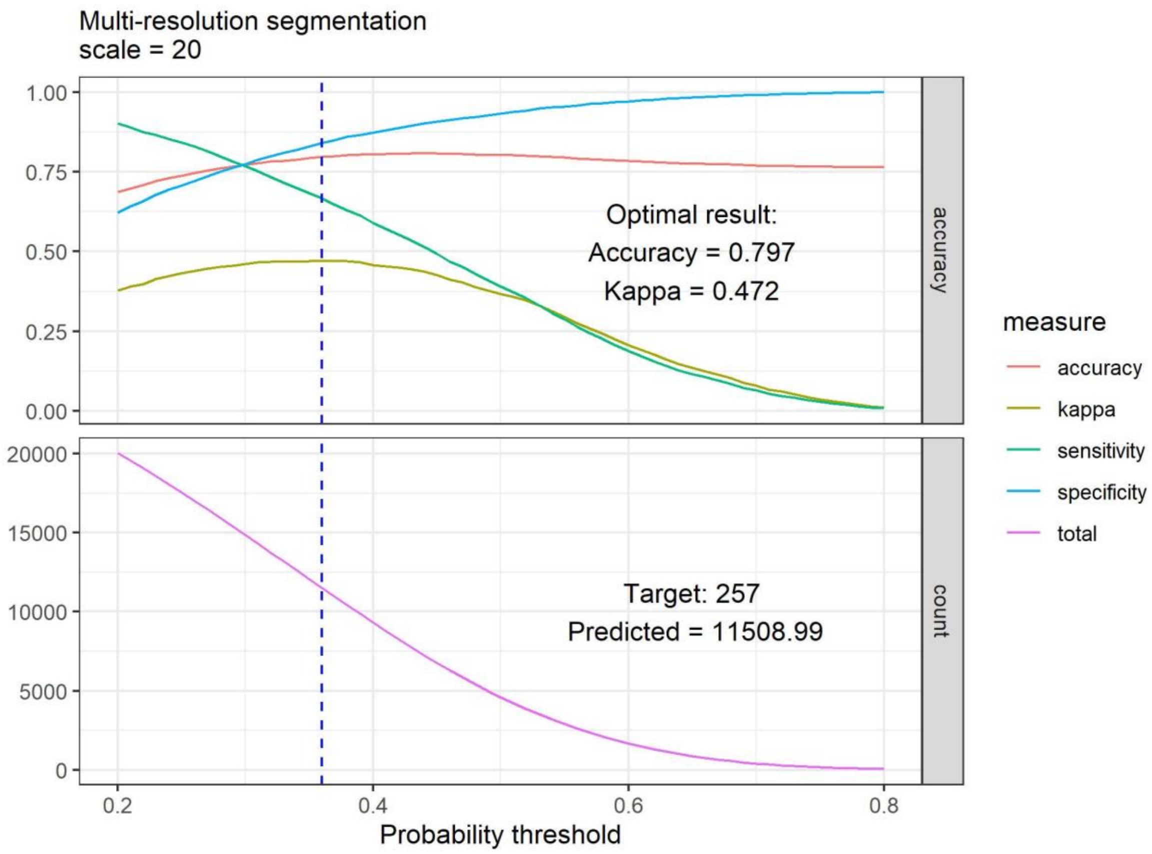

Watershed segmentation
scale = 30
variable = ndvi
Optimal result:
Accuracy = 0.74
Kappa = 0.341
Predicted = 1713
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold
1.00
0.75
0.50
0.25
0.00
5000
4000
3000
2000
1000
0

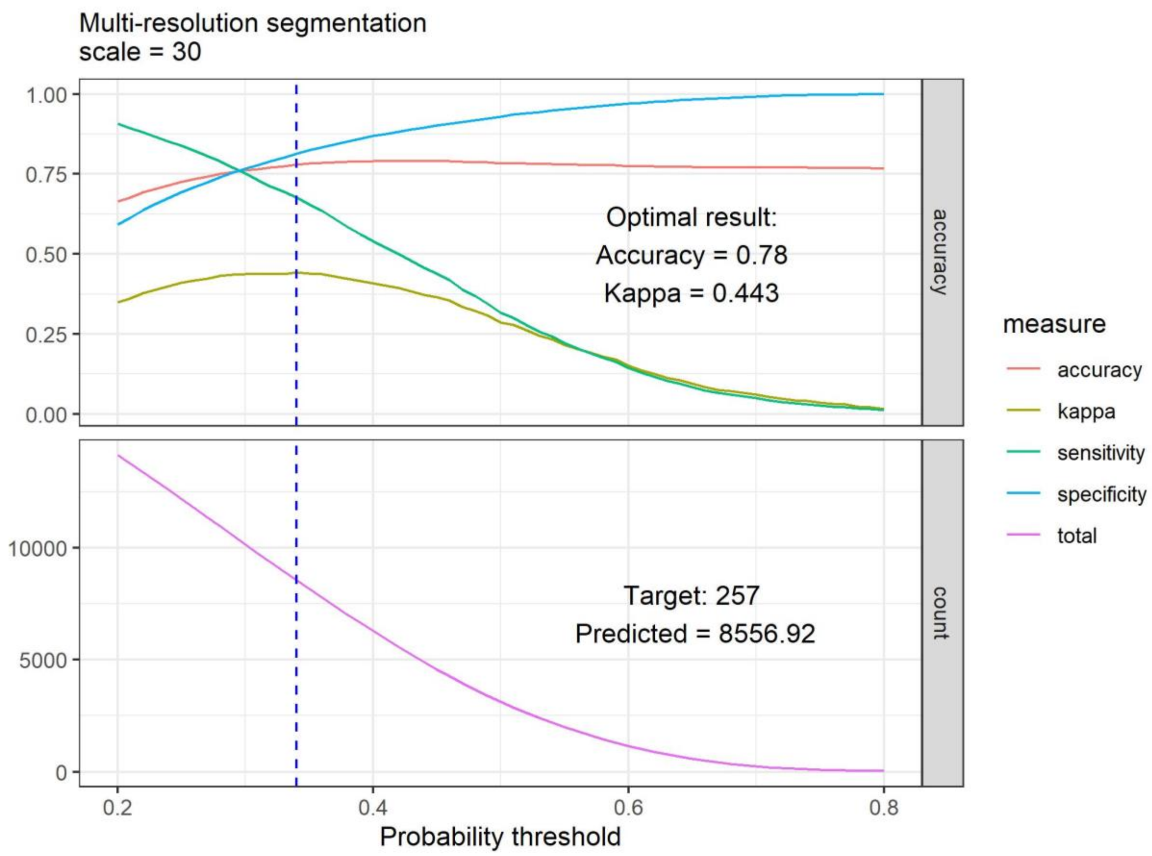

Watershed segmentation
scale = 40
variable = ndvi
Optimal result:
Accuracy = 0.758
Kappa = 0.339
Predicted = 1331.14
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold
1.00
0.75
0.50
0.25
0.00
5000
4000
3000
2000
1000
0

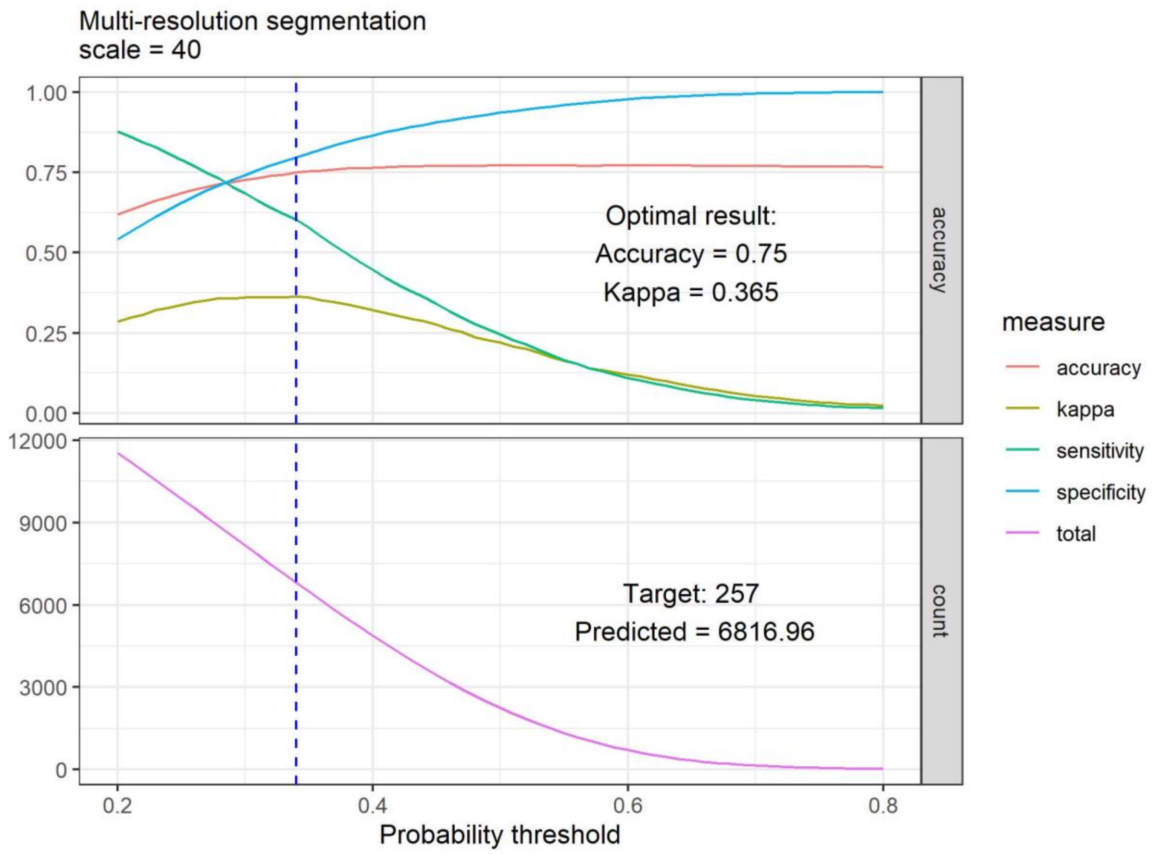
Trial 2

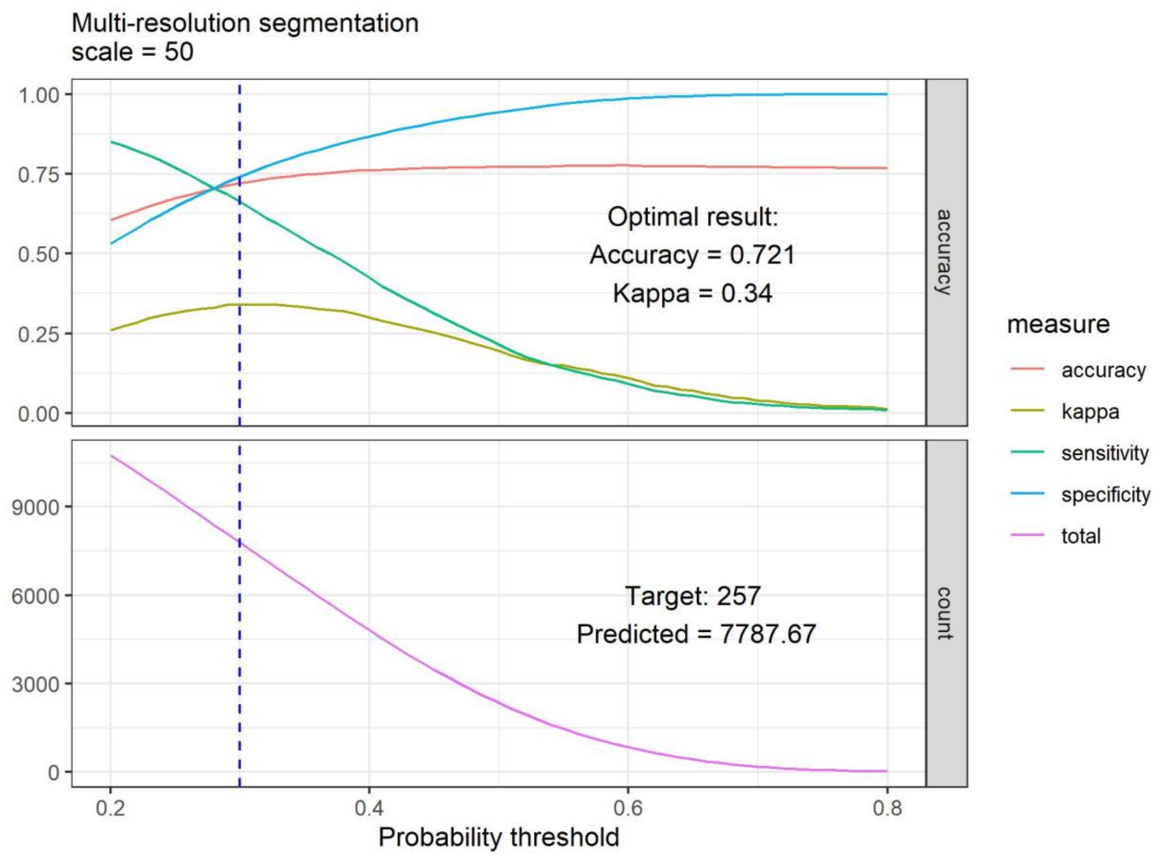
Using the 25 cm data, multiresolution segmentation at scales of 10 to 100 incrementing every 10 counts.

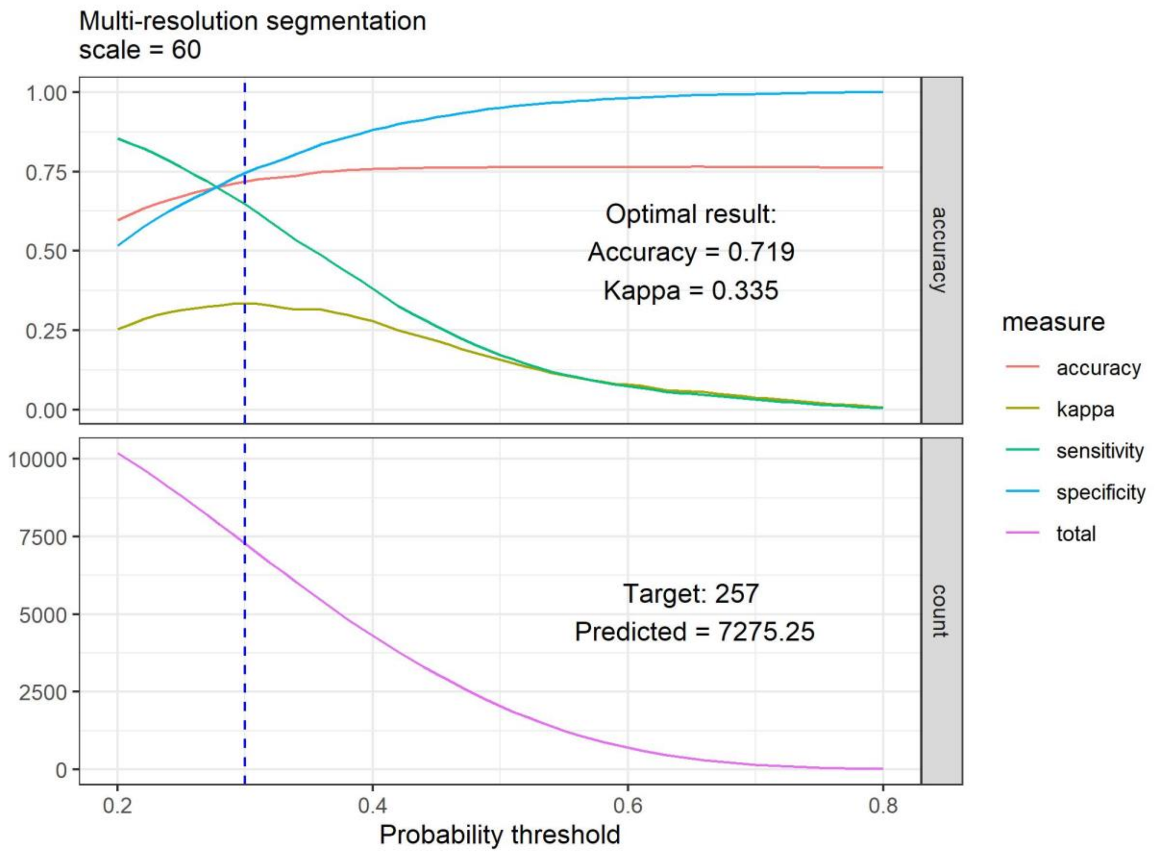



Multi-resolution segmentation
scale = 20
Optimal result:
Accuracy = 0.797
Kappa = 0.472
Target: 257
Predicted = 11508.99
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
1.00
0.75
0.50
0.25
0.00
20000
15000
10000
5000
0
0.2
0.4
0.6
0.8
Probability threshold

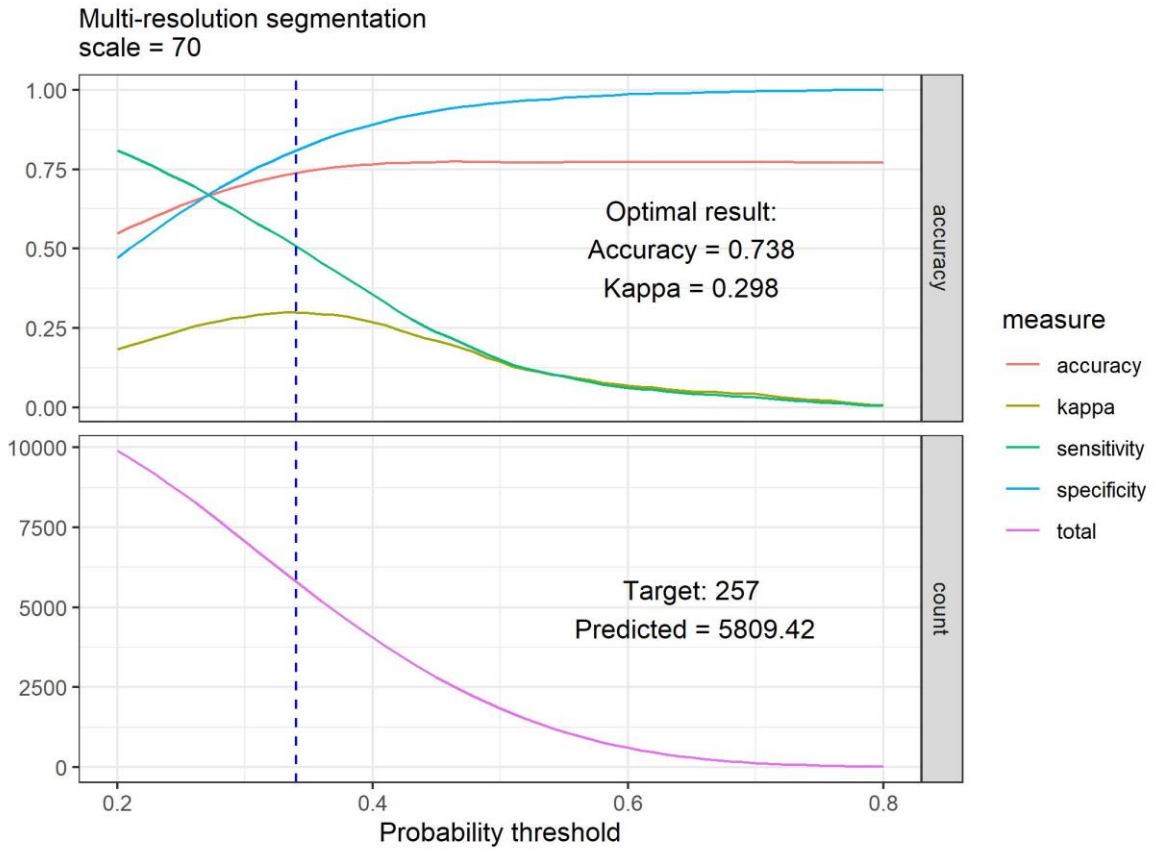

Multi-resolution segmentation
scale = 30
Optimal result:
Accuracy = 0.78
Kappa = 0.443
Target: 257
Predicted = 8556.92
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
1.00
0.75
0.50
0.25
0.00
10000
5000
0
0.2
0.4
0.6
0.8
Probability threshold


Multi-resolution segmentation
scale = 40
1.00
0.75
0.50
0.25
0.00
Optimal result:
Accuracy = 0.75
Kappa = 0.365
accuracy
12000
9000
6000
3000
0
Target: 257
Predicted = 6816.96
count
0.2
0.4
0.6
0.8
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total

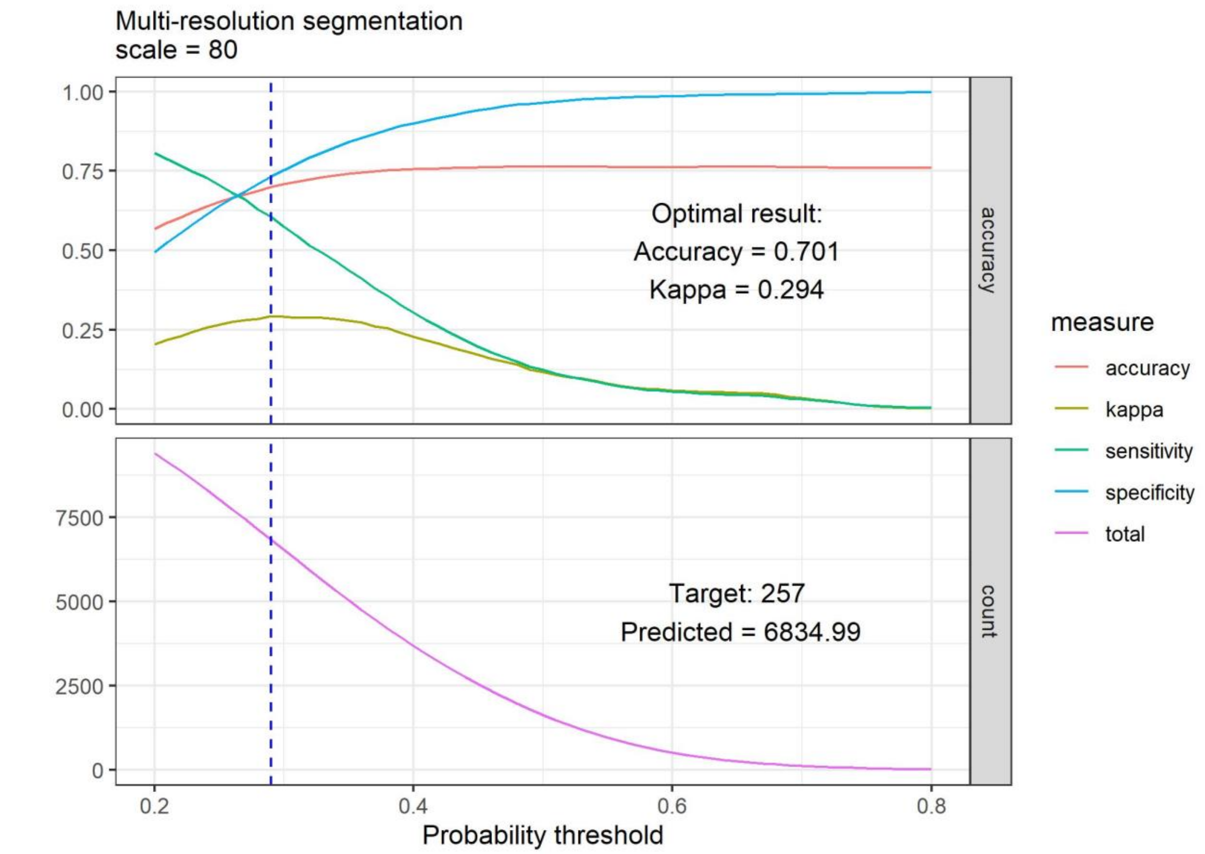

Multi-resolution segmentation
scale = 50
1.00
0.75
0.50
0.25
0.00
Optimal result:
Accuracy = 0.721
Kappa = 0.34
accuracy
9000
6000
3000
0
Target: 257
Predicted = 7787.67
count
0.2
0.4
0.6
0.8
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total

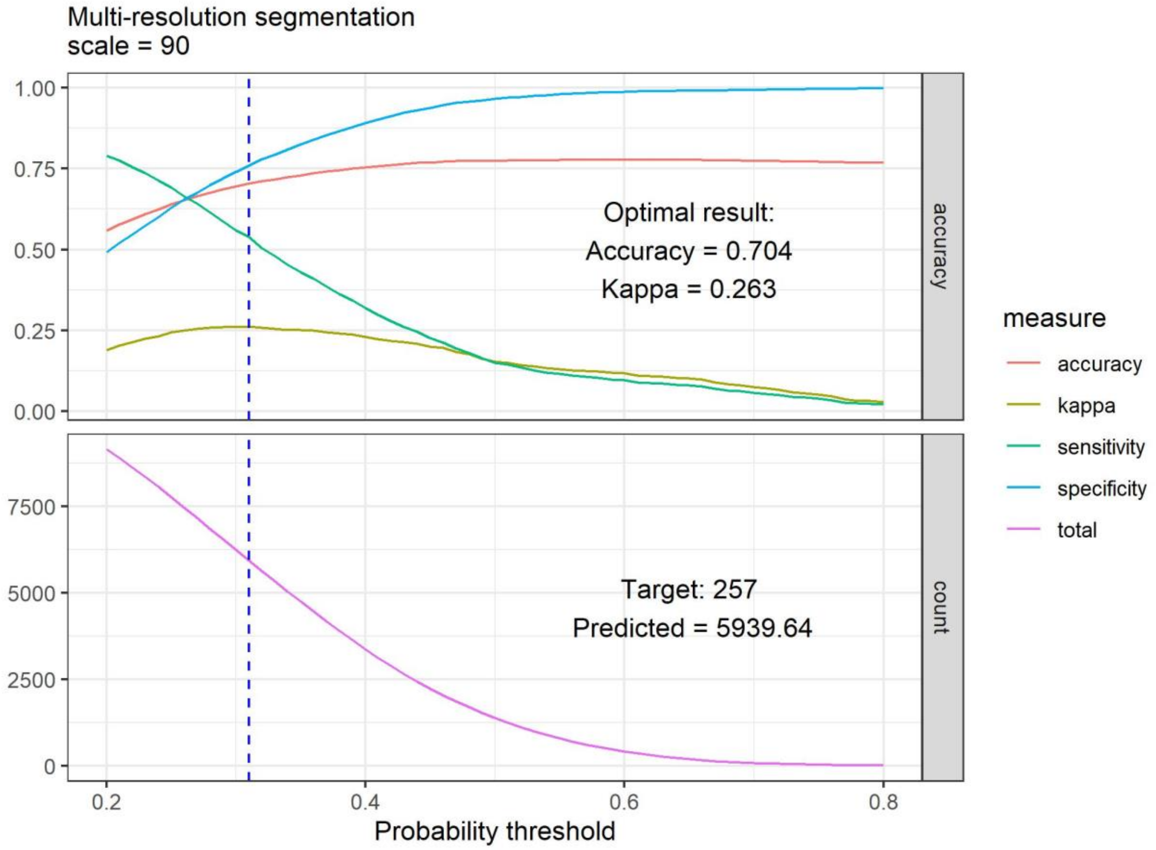


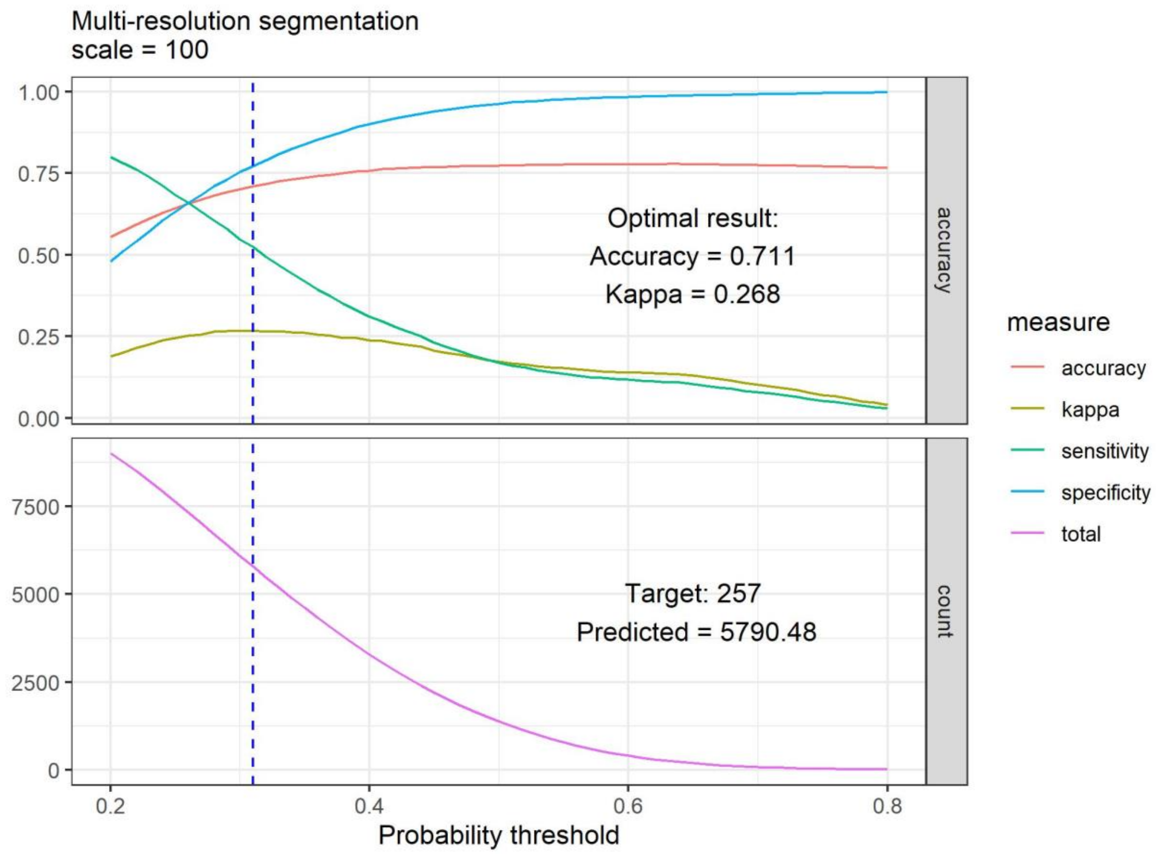
Multi-resolution segmentation
scale = 60
Optimal result:
Accuracy = 0.719
Kappa = 0.335
Target: 257
Predicted = 7275.25
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold



Multi-resolution segmentation
scale = 70
Optimal result:
Accuracy = 0.738
Kappa = 0.298
Target: 257
Predicted = 5809.42
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

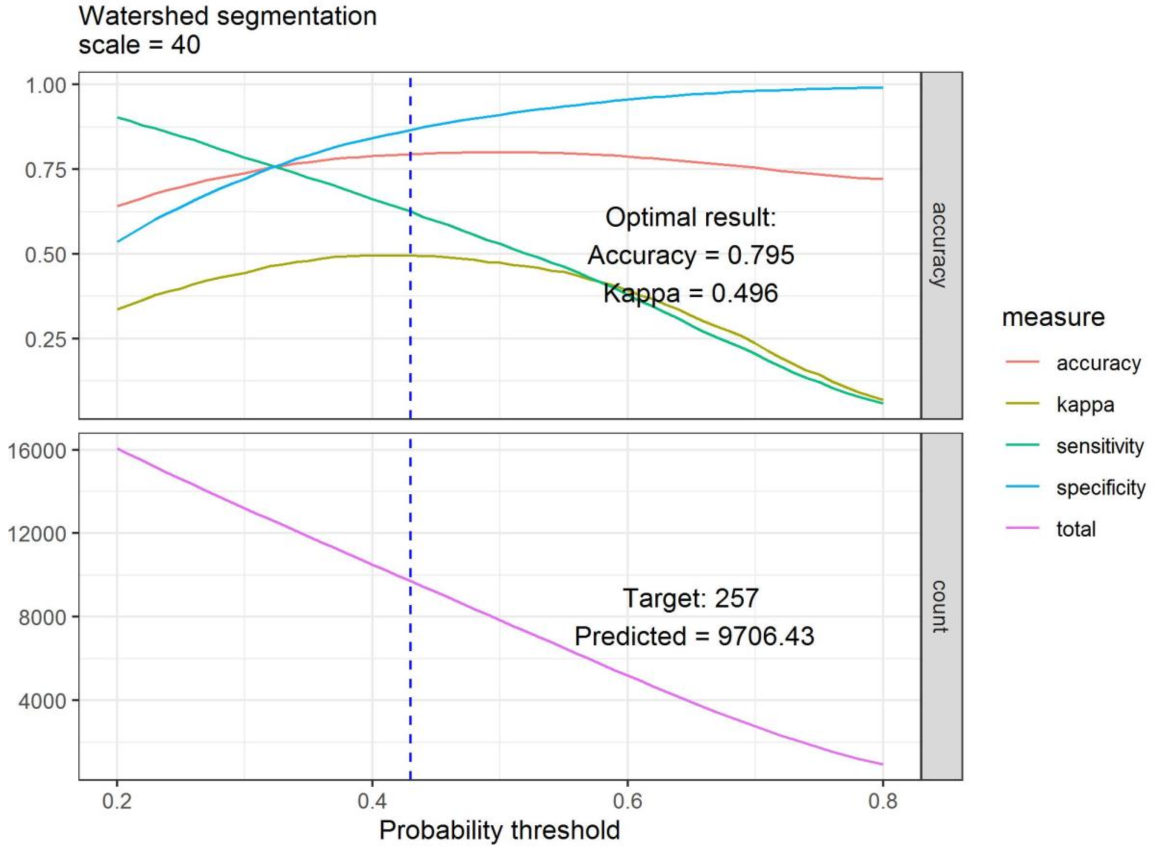

Multi-resolution segmentation
scale = 80
Optimal result:
Accuracy = 0.701
Kappa = 0.294
Target: 257
Predicted = 6834.99
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

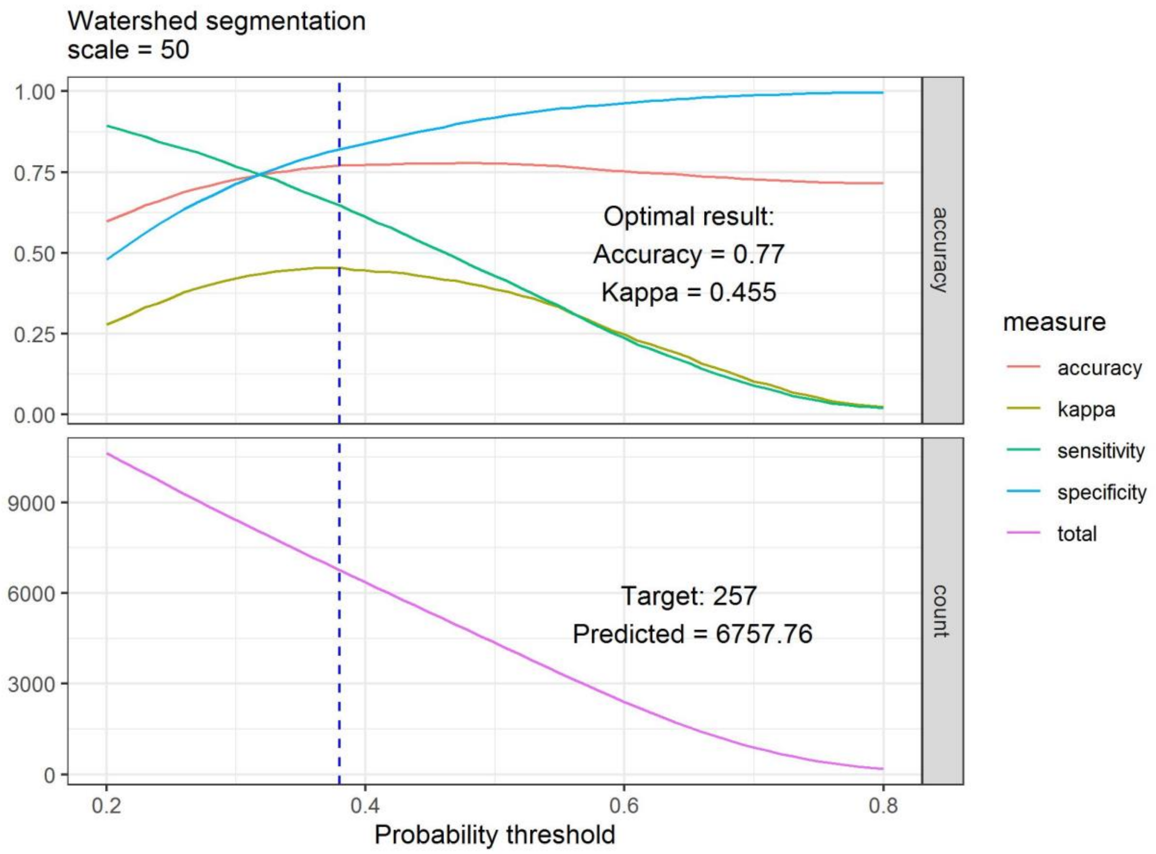

Multi-resolution segmentation
scale = 90
Optimal result:
Accuracy = 0.704
Kappa = 0.263
Target: 257
Predicted = 5939.64
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

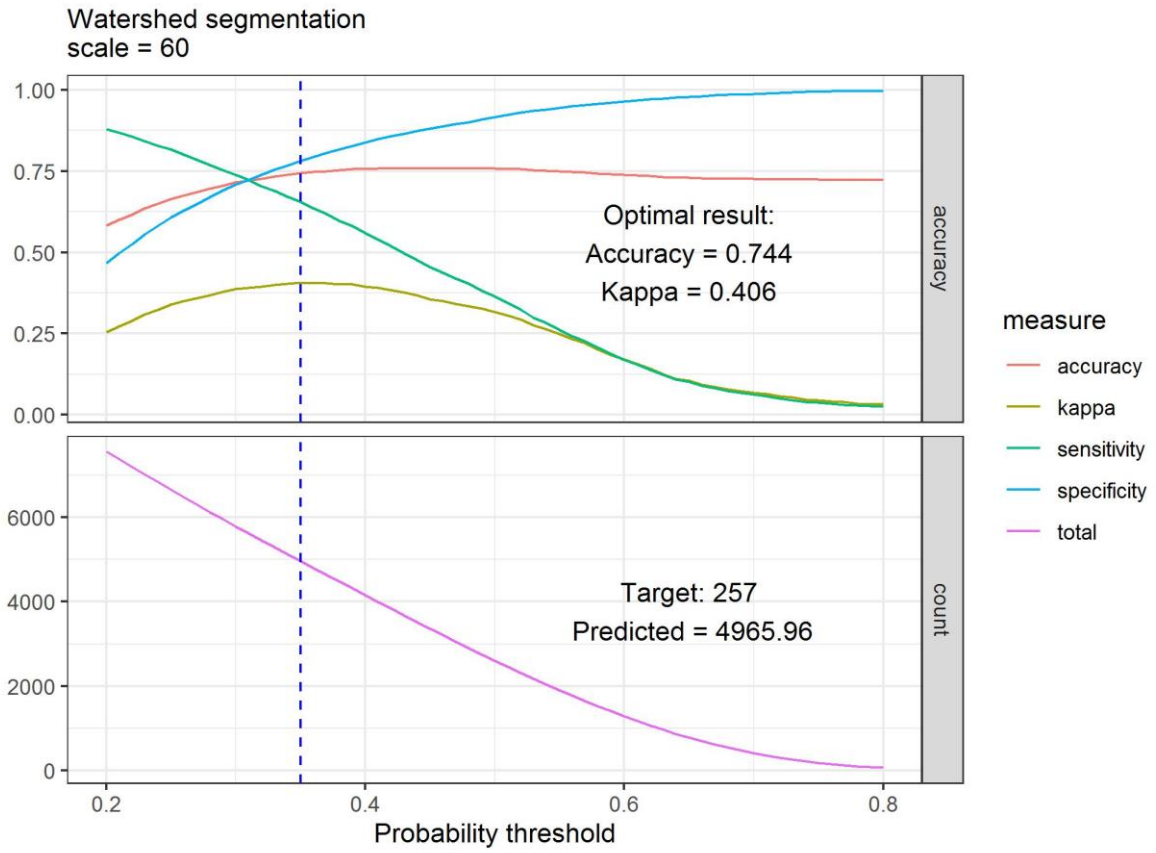

Multi-resolution segmentation
scale = 100
1.00
0.75
0.50
0.25
0.00
Optimal result:
Accuracy = 0.711
Kappa = 0.268
accuracy
measure
accuracy
kappa
sensitivity
specificity
total
7500
5000
2500
0
Target: 257
Predicted = 5790.48
count
0.2
0.4
0.6
0.8
Probability threshold

Trial 3

Using the 5 cm data with the watershed segmentation at scales of 40 to 60 incrementing every 10 counts.

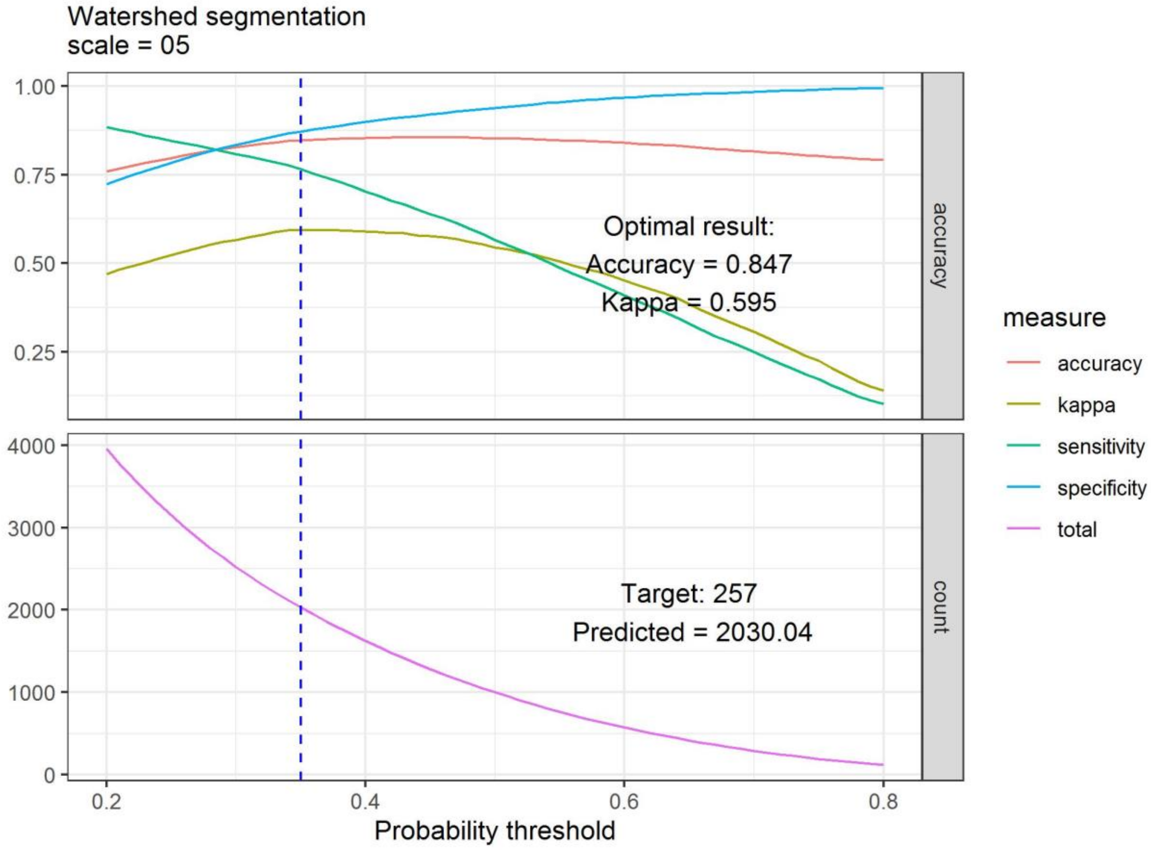


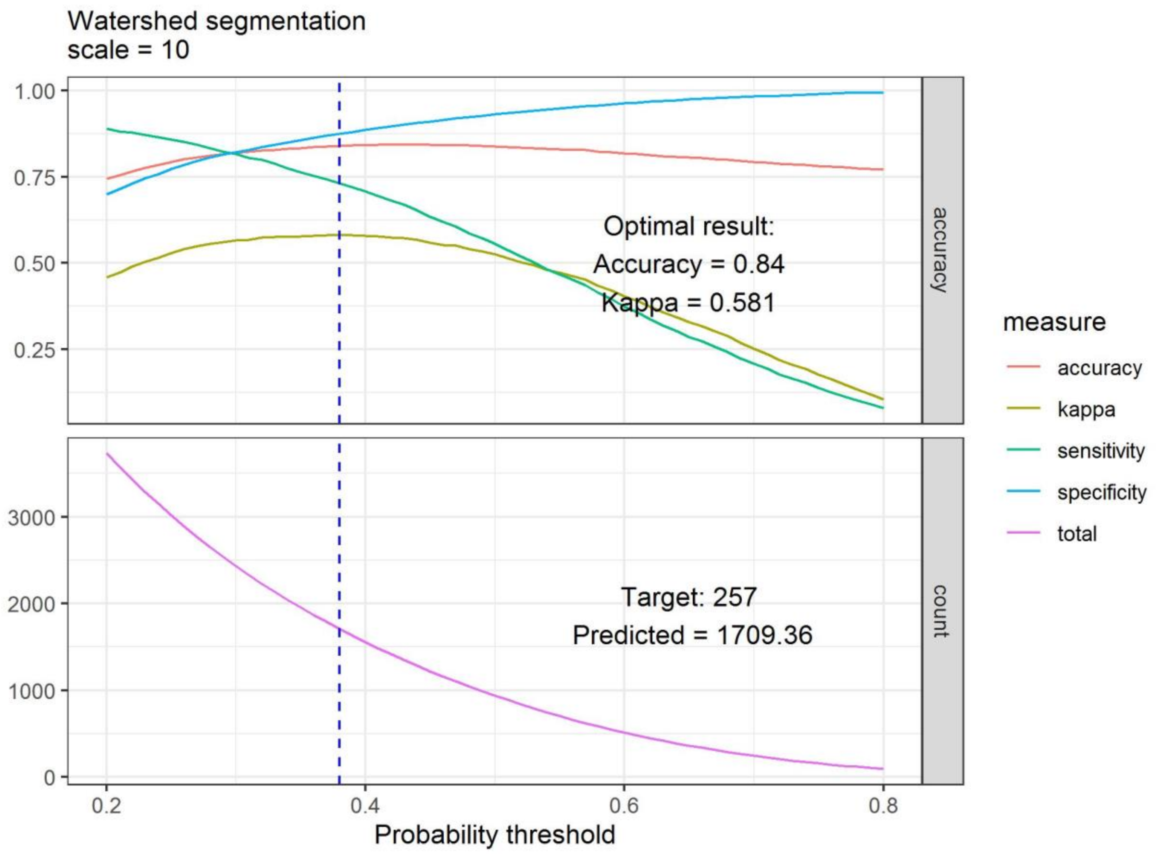

Watershed segmentation
scale = 50
Optimal result:
Accuracy = 0.77
Kappa = 0.455
Target: 257
Predicted = 6757.76
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

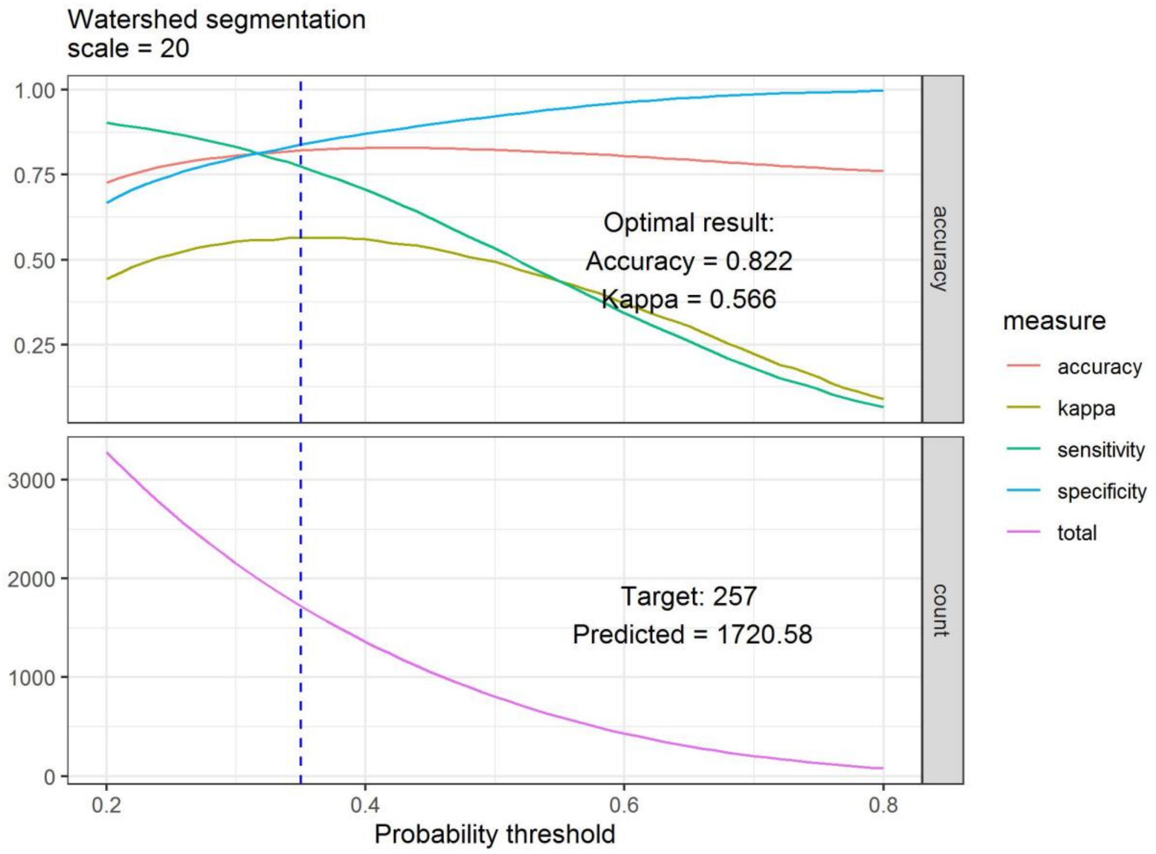

Watershed segmentation
scale = 60
Optimal result:
Accuracy = 0.744
Kappa = 0.406
Target: 257
Predicted = 4965.96
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold

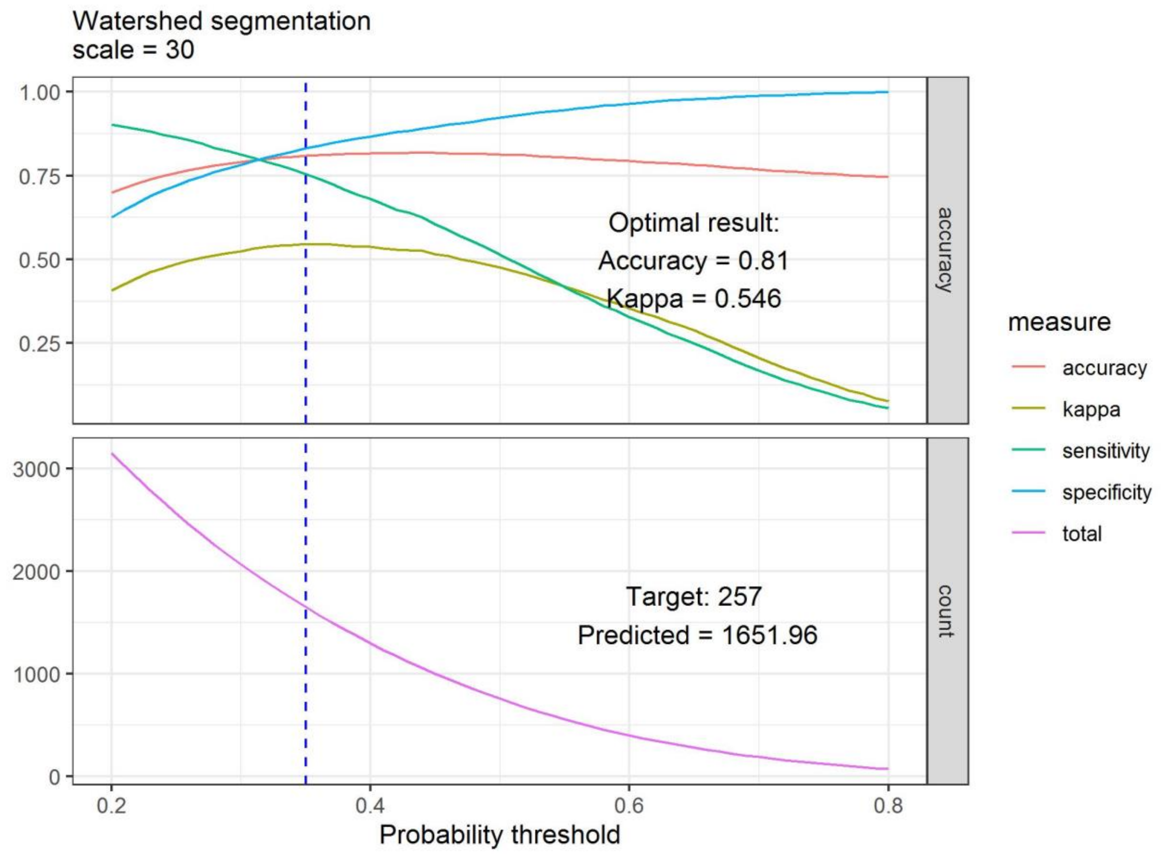
Trial 4

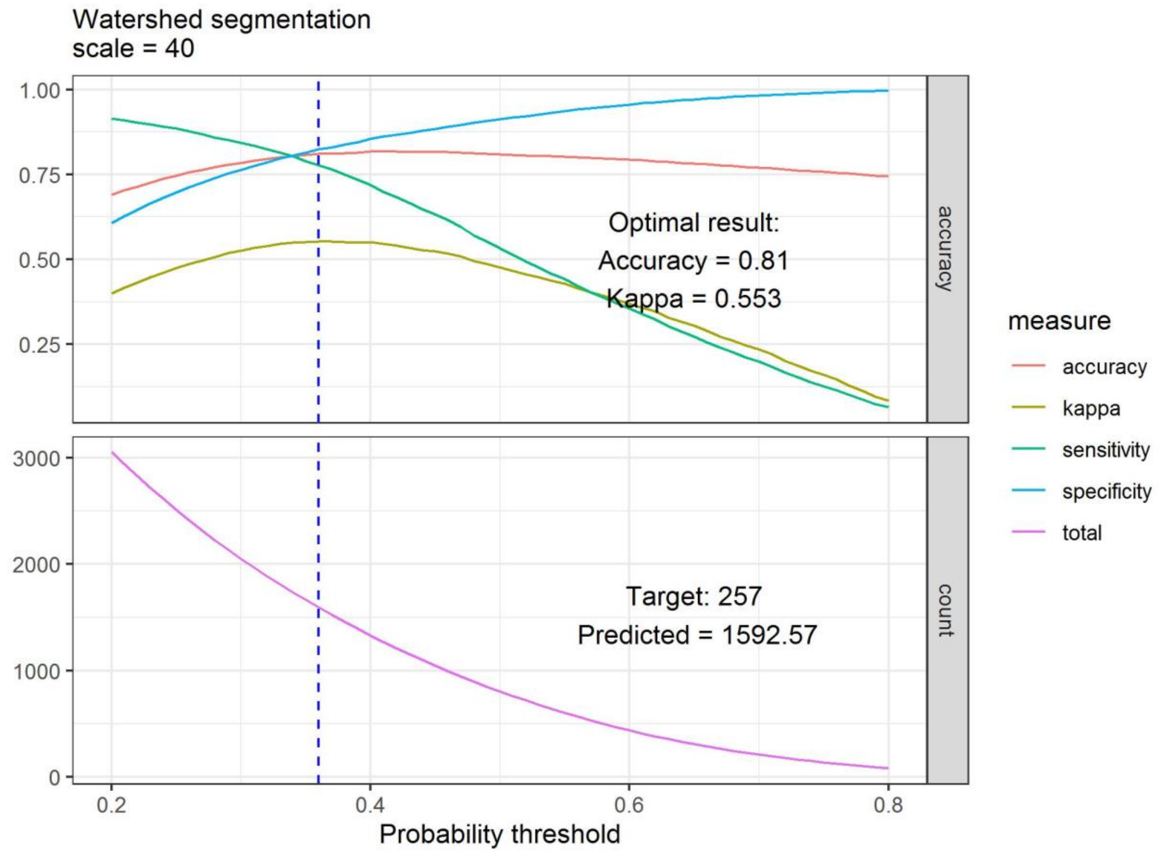
Using the 25 cm data, watershed segmentation at thresholds 5, 10, 20, 30 and 40 using the GLA with training data converted to polygons.

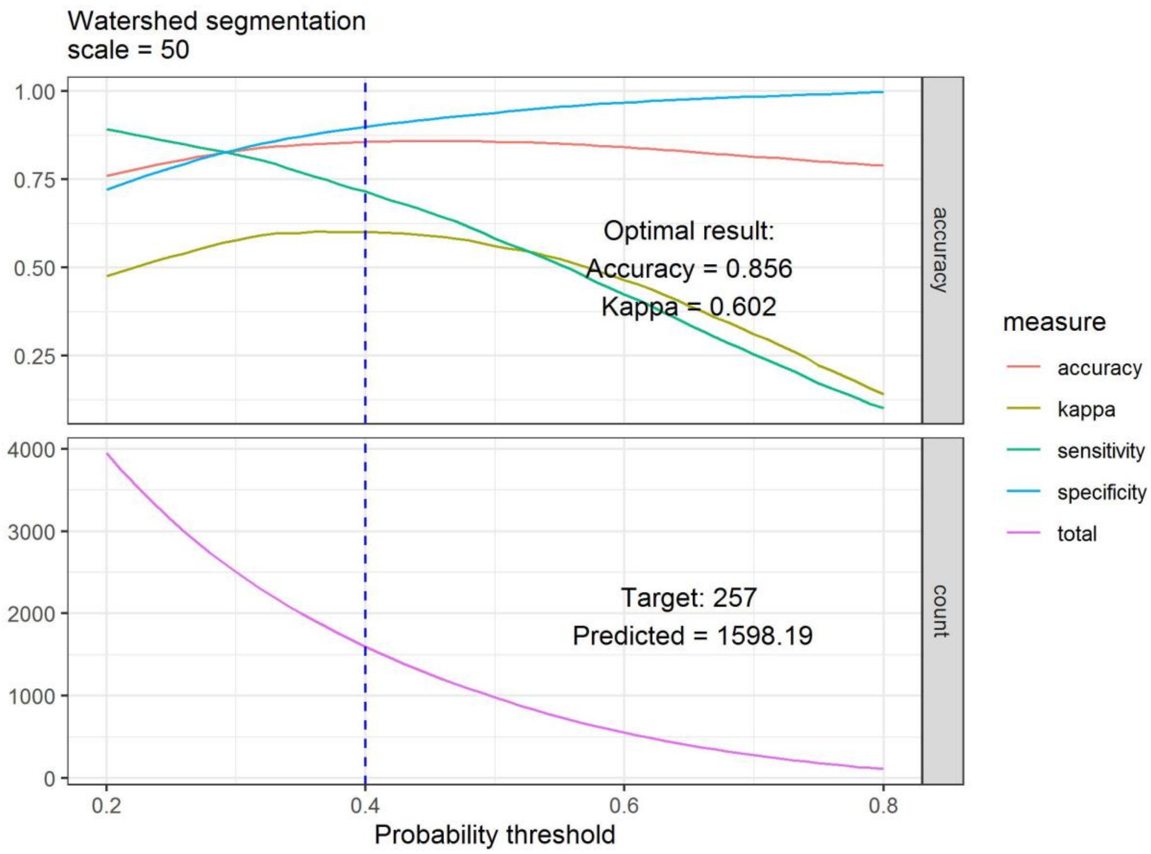



Watershed segmentation
scale = 10
Optimal result:
Accuracy = 0.84
Kappa = 0.581
Target: 257
Predicted = 1709.36
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold
1.00
0.75
0.50
0.25
3000
2000
1000
0
0.2
0.4
0.6
0.8


Watershed segmentation
scale = 20
Optimal result:
Accuracy = 0.822
Kappa = 0.566
Target: 257
Predicted = 1720.58
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total
Probability threshold
1.00
0.75
0.50
0.25
3000
2000
1000
0
0.2
0.4
0.6
0.8


Watershed segmentation
scale = 30
Optimal result:
Accuracy = 0.81
Kappa = 0.546
Target: 257
Predicted = 1651.96
accuracy
count
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total


Watershed segmentation
scale = 40
Optimal result:
Accuracy = 0.81
Kappa = 0.553
Target: 257
Predicted = 1592.57
accuracy
count
Probability threshold
measure
accuracy
kappa
sensitivity
specificity
total


Watershed segmentation
scale = 50
1.00
0.75
0.50
0.25
4000
3000
2000
1000
0
0.2
0.4
0.6
0.8
Probability threshold
Optimal result:
Accuracy = 0.856
Kappa = 0.602
Target: 257
Predicted = 1598.19
accuracy
count
measure
accuracy
kappa
sensitivity
specificity
total