Impact of multidimensional parameters on the area of applicability (AOA) for machine learning based classifications

The impact of training design and predictor variables on the prediction and AOA

Andreas Schönberg B.Sc.[1]

[1] Master student at the department of Geography, Philipps University of Marburg, Germany

Abstract
Machine learning algorithms like random forest have become commonly used in spatial predictions. A major problem with the results of those approaches is the estimation of the performance outside the training areas. Cross validated model accuracy values could easily lead to overoptimistic results. While target-orientated approaches can help to gain more realistic performance values one major problem still is the estimation of the applicability for unknown space. The area of applicability (AOA) approach developed by MEYER & PEBESMA (2020) can be used to test the applicability of a model for an area of interest. It is based on the relationship of the predictor values in each cell compared to those used in the training. Therefore the AOA of a model is highly depending on the used data. I will test the impact of different training designs and the use of artificial layer (spectral indices and filter) on the resulting predictions and the AOA. The aim of this study is to investigate the importance of the accuracy of the set training areas for small scale classifications on high resolution RGB images and the use of artificial layers compared to RGB bands only.

Keywords: Area of applicability (AOA), machine learning, random forest, remote sensing, spatial mapping, spatial filter, spectral indices.

1. Introduction

Spatial predictions based on machine learning approaches recently have become a buzz word in remote sensing science (FOX ET AL 2017; MEYER ET AL 2016; CUTLER ET AL 2007). In spatial classifications for example models can be trained from small training areas or even single pixels to predict the classes for much greater areas of interest (AOI). A common used algorithm is random forest (BREIMAN 2001) which is known for relatively high stability and robustness. Random forest is a non-parametric ensemble approach and suffers a direct quantification of prediction error (COULSTON ET AL 2016). A popular way for estimating the performance of a model is the k-fold cross validation (CV) approach (KUHN & JOHNSON 2013) but could lead to overoptimistic results in spatial predictions (MEYER ET AL 2016). The leave location out cross validation (LLOCV) approach (MEYER ET AL 2016) can improve the estimation of model performance by leaving out data based on locations instead of random pixels. See MEYER ET AL (2016) for a more detailed description for target-orientated validation strategies.

Even with high performance models using the LLOCV the prediction could be less useful due to other relations of the data outside the training areas. The area of applicability (AOA) approach developed by MEYER & PEBESMA (2020) provides a way to estimate the spatial applicability of the prediction for unknown space by comparing the euclidian distances of the predictors in each cell with those used for the training. For a detailed description of the AOA approach see MEYER & PEBESMA (2020). The AOA therefore highly depends on the data and the relation of its dimensions. Especially for low spectral resolution data sources like RGB images the performance of the model can be increased by giving the algorithm more value information due to data processing (HUNT ET AL 2005; MEYER ET AL 2019). There are several strategies to generate sets of predictors (Raster Stacks) containing additional information like the computation of artificial layers like filter and indices (HUNT ET AL 2005; MEYER ET AL 2019; FIERENS & ROSIN 1994) and principal component analysis (PCA) (ESTORNELL ET AL 2013). Machine learning algorithms in general tend to be sensibility to over-fitting due to over-sized data sets, high correlating data or layers which contains continuous data (like coordinates) (MEYER ET AL 2016). So a dimensional reduction is highly recommended. A forward feature selection (FFS) (MEYER ET AL 2019) can be used to determine the best fitting layers but although can result in long processing times. Further the design of the training area will have an impact on the results (ZHAO ET AL 2020; JIN ET AL 2014). Common workflows for classifications based on random forest machine learning approaches (MEYER ET AL 2019) are used on great scales.

In this study I will test the impact of the training design on the prediction and AOA for classifications of small scale entities on high resolution RGB images at the upper alpine tree line. Those areas typically are represented by small trees surrounded by grass and open soil. With the high resolution small stones can be detected lying within the grass areas as well as small spots of grass on rocky areas. I hypothesize that increasing sizes of training areas will result in decreasing performances for the predictions. Greater training areas would tend to represent more than one class. With the AOA highly depending on the values and its distribution I assume that more data could further result in lesser AOA. Therefore I hypothesize that increasing sizes for the training areas lead to decreasing AOA. Furthermore it could be difficult to set training areas accurate enough on small scale entities in high resolution images. Therefore I will test the sensitivity of the workflow by using two different sets of training points. Besides the original RGB data I will compute a Raster Stack of several artificial layers to compare the performance and processing time. The Raster Stack will include spatial filter, spectral indices and PCA. For dimensional reduction I will drop highly correlation layers and use a FFS to select the best fitting layers. I assume that the performance will be improved compared to the use of only RGB bands but would take more time to proceed. To validate the results I will use three different study sites and several sizes of training areas for both sets of training points.

2. Data and Methods

For this study I will use three high resolution RGB images from a study area located at the upper alpine treeline. I will use two diffenret sets of training points with buffers for increasing sizes of training areas. For the models I will use RGB bands as predictor variables as well as a Raster Stack containing artificial layer. The data processing and modelling is performed in R version 4.0.4 (R CORE TEAM 2020). The used image, data and R-scripts can be retrieved from https://github.com/SchoenbergA/AOA-Impact.

2.1 Study Area and Training Design

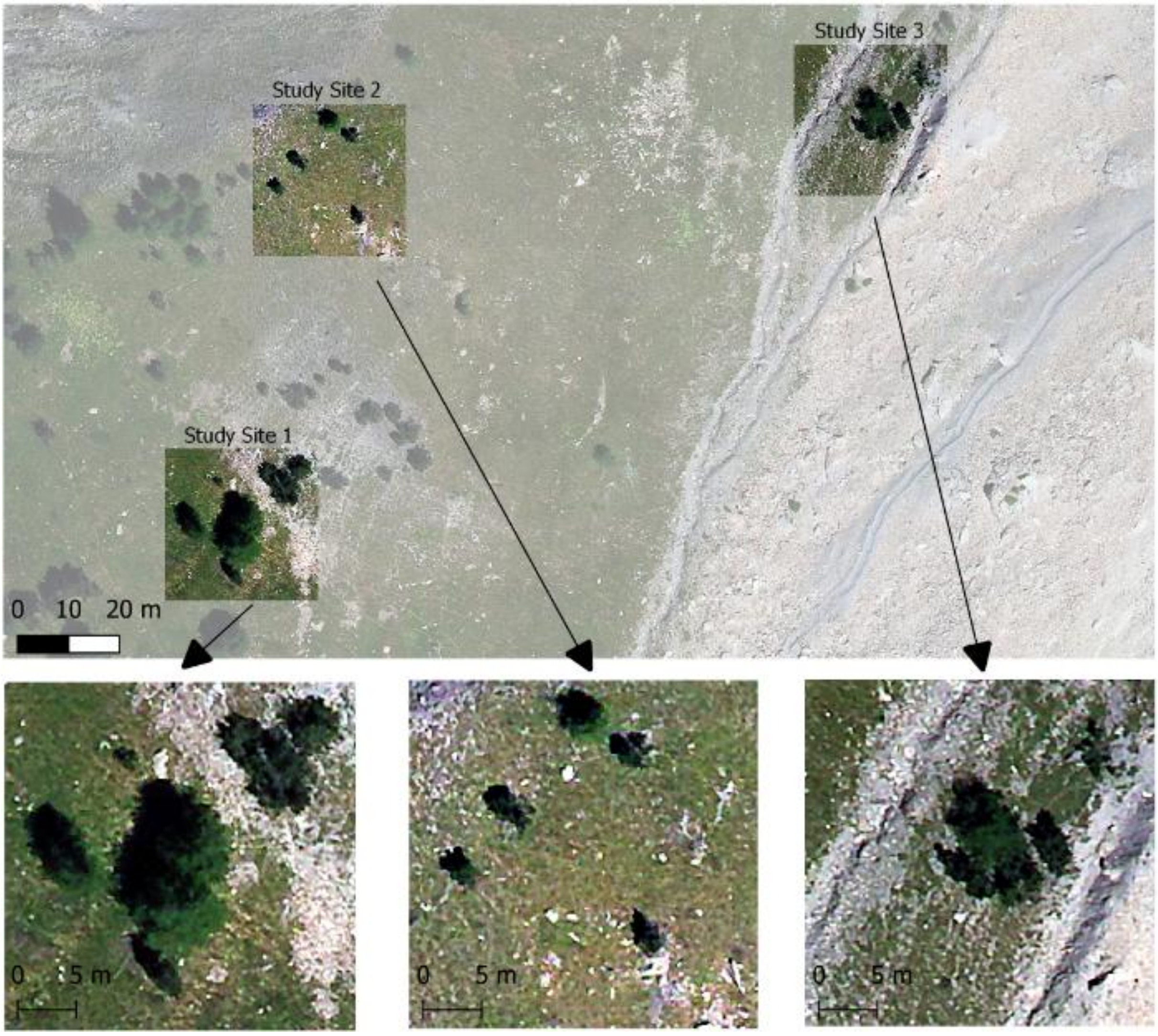

Fig. 1*: The study area and study sites.*

To investigate my hypothesis I will use high resolution RGB aerial images with 0.15 meter resolution of an upper treeline (**Fig.1**). The study area is located in the Lautaret Valley in the Frence Alps with rocky areas covered by spots of grass, areas of grass with stones and small trees. For the classification I will use the three classes: trees (t), grass (g) and soil (s) for the three study sites with 30x30 meter extent representing different compositions of the classes. The first study site is covered by relatively large trees with crown radiuses of about 5 meter. The second study site is mostly covered by grass with lesser open soil and small single trees while the third study site represents a small and greater group of trees in close neighborhood. While soil is separated by gray to silver colors from the other classes trees and grass share dark to brighter greens. Due to the spectral similarity of tree and grass the scenes are perfect to test the impact of the training design on the classification. For a more standardized and comparable testing I will use two different sets of training points for each study site (**Fig.2**). The points are set intuitive on positions representing the average behavior of the classes. To implement the LLOCV approach I will cut the scene into four sectors (NE, SE, SW, NW) and set three training points for each class in each sector. To validate the resulting prediction I will use a response layer for the class of trees. The response layer is digitalized by hand representing the visual estimated areas of trees in the study area.

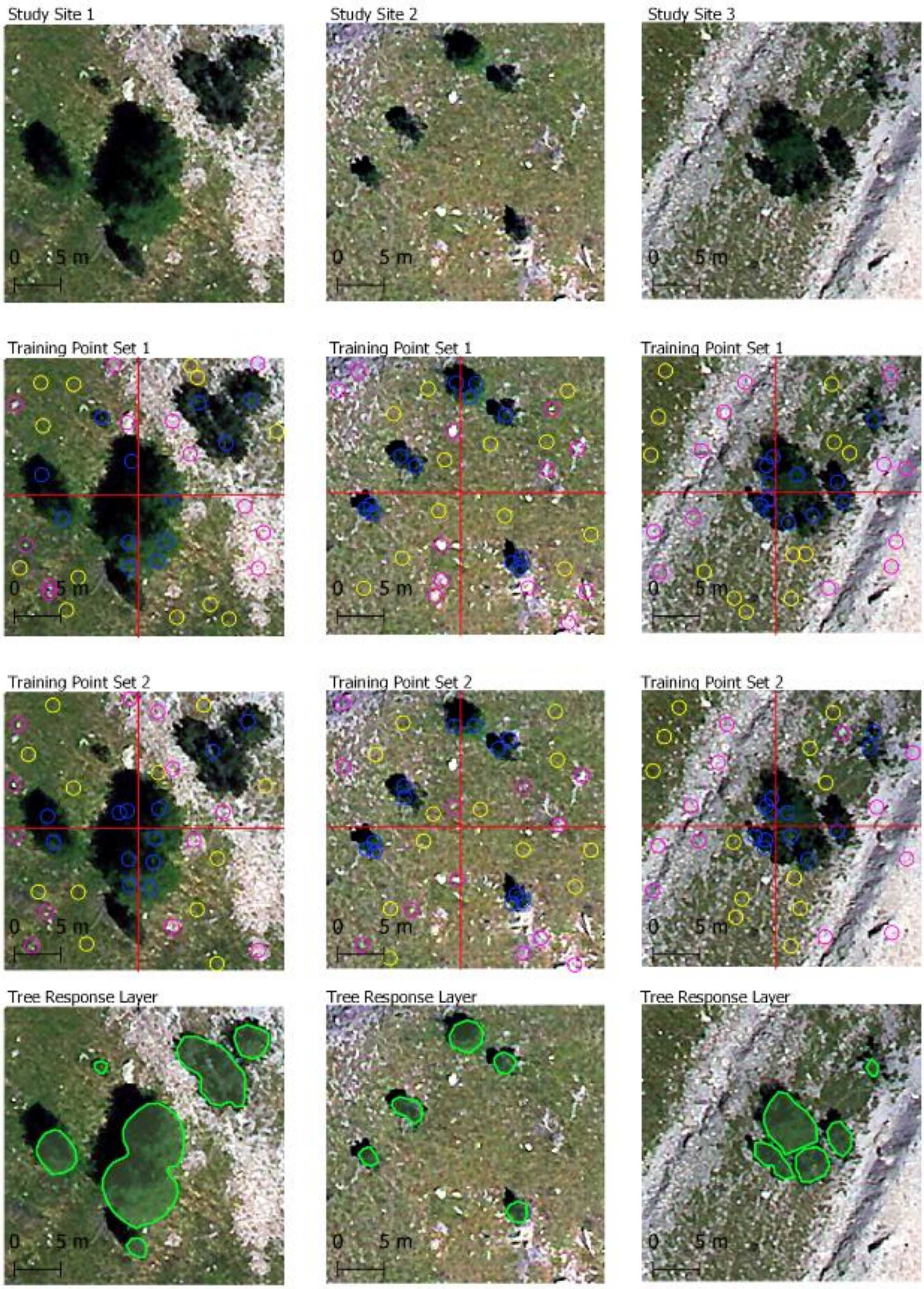

Fig. 2: *Training points for grass (yellow), soil (magenta), trees (blue) and response layer for the study sites.*

2.2 Artificial Layer Computation

Besides the original RGB bands I will use a Raster Stack with artificial layers for each study site. First I will use a principal component analysis (PCA) which is commonly used for data processing (Estornell et al 2013; Desmer et al 2013; Cao et al 2003). To generate additional information for the machine learning algorithm first I will first use the 'rasterPCA' function provided by the 'Rstoolbox' package (Leutner at al 2019) to compute the first PCA for the RGB bands. Due to the high resolution the cell values could be highly different even in close neighborhood. So I assume that filters would probably increase the performance of a classification by smoothing (Fierens & Rosin 1994) the values in close neighborhood of the training areas as well as for the prediction. The filters will be assigned to the computed PCA layer with common spatial filters based on the 'Raster::focal' function (Hijmans 2020) and sobel filter for edge detection (**Tab. 1**). The filter requires a moving window defining the amount of cells in neighborhood which the filter function is assigned to. With the extents of the study sites of 30x30 meter and the 0.15 meter resolution I will use moving windows of 3x3 and 5x5. Further indices are commonly used in remote sensing approaches (Ray et al 2004; Hunt et al 2005; Meyer et al 2019) to detect vegetation and could help to detect and separate non-vegetation (soil) (Ray et al 2004) and vegetation (grass/trees) due to the sensitivity to chlorophyll (Hunt et al 2005; Hunt et al 2013). For further information about the spectral RGB indices see: The IDB Project (2020); Ray et al (2004); Hunt et al (2013). For the artificial Raster Stack I will use common RGB indices (**Tab. 2**) along with the RGB bands, the first PCA as well as the filter layers based on the PCA.

There would be a high amount of correlating layers in the Raster Stacks due to the mechanism of computation which could lead to over-fitting when used for the machine learning. Further the FFS would take long time to proceed. To reduce the dimensions in the Raster Stacks I will test for correlation and delete highly correlating layers. Further some layers especially the filtered indices could result in highly homogeneous layers containing only a few values. Those layers could lead to explaining the dependencies of the training area very easily for the algorithm and lead to false predictions. To handle those layers I will further test the distribution of the data values and drop layers by thresholds. For the correlation test I will drop all layer with a cor value of >= 0.9 and <= -0.9 and further drop homogeneous layers which have >= 90% of data values in <=10% of the data range.

Table 1: *Filter functions used for the Raster Stack.*

Name	Calculation
Sum	sum of all cells in a MovingWindow
Minimum	minimum value of all cells in a MovingWindow
Maximum	maximum value of all cells in a MovingWindow
Mean	mean value of all cells in a MovingWindow
Standard deviation	standard deviation of all cells in a MovingWindow
Modal	most frequent value of all cells in a MovingWindow
Sobel	sobel edge detection filter in horizontal and vertical directions
Sobel hrz	sobel edge detection filter in horizontal direction only
Sobel vct	sobel edge detection filter in vertical direction only

Table 2: *Indices used for the Raster Stack.*

Name	Tag	Calculation
Visible Vegetation Index	VVI	(1 - abs((red - 30) / (red + 30))) * (1 - abs((green - 50) / (green + 50))) *(1 - abs((blue - 1) / (blue + 1)))
Visible Atmospherically Resistant Index	VARI	(green-red)/(green+red-blue)
Normalized Difference Turbidity Index	NDTI	(red-green)/(red+green)
Redness index	RI	(red**2/(blue*green**3)
Soil Colour Index	CI	(red-green)/(red+green)
Brightness Index	BI	sqrt((red**2+green**2+blue*2)/3)
Spectra Slope Saturation Index	SI	(red-blue)/(red+blue)
Primary Colours Hue Index	HI	(2*red-green-blue)/(green-blue)
Triangular Greenness Index	TGI	(-0.5*(190*(red - green)- 120*(red - blue))
Green Leaf Index	GLI	(2*green-red-blue)/(2*green+red+blue)
Normalized Green Red Difference Index	NGRDI	(green-red)/(green+red)

2.3 Test Method

To investigate my hypothesis I will use increasing training areas in cycle form for each study site with both sets of training points and further using both RGB and the artificial layer Raster Stack as predictors. The training areas will be computed by using spatial buffers with given radiuses around the training points. Due to the resolution and the extent of the image I will use multiplies of 0.15 meter from 0.15 up to 0.9 meters. I suggest wider areas would easily catch values for more than one class which would lead to false results.

For the computation of the model and the respective AOA calculation I will use a workflow mainly based on 'CAST' (MEYER 2020). First the training data is extracted for the desired training design. Next the LLOCV is prepared using 'CAST::CreateSpacetimeFolds' (MEYER 2020) by defining indices for the data to be left out based on the four sectors of the study site. For the machine learning algorithm the 'Random forest' (BREIMAN 2001) is used.

Further I will use a forward feature selection (FFS) for the artificial layer Raster Stack which will select the predictors with highest importance. The models are trained first for every possible pair of two layers and the best model is selected. Based on this model the layers are iteratively increased to test for an improvement in model performance until none of the remaining layers would further increase the current best model (MEYER 2020). The resulting model based on the selected variables will be used to predict the classification for the study area. The RGB bands will be used without an FFS. Finally the AOA is calculated using 'CAST::aoa' (MEYER 2020) . For the evaluation of the resulting performance the 'accuracy' and 'kappa' will be used as well as the percentage of the AOA. Further the runtime for the process is saved for comparison reasons.

For the validation the respective response layer is compared to the areas which are classified by the algorithm as class tree. Estimating the response layer to be accurate the performance of the prediction will be estimated by the amount of cells classified as trees which overlap with the response layer ('hitrate') and those outside the response areas ('missrate'). The performance of the prediction will be estimated by the 'accuracy' and 'kappa' values as well as the response values of 'hitrate' and 'missrate'. The AOA is primarily evaluated by its total percentage but further visually to check for significant spatial pattern.

3. Results

In total I tested three study sites with six sizes of buffers for two different sets of training points for both RGB bands (**Tab.3**) and the artificial layer Raster Stack (**Tab.4**). Overall the forward feature selection for the artificial layer Raster Stacks took up 28 times longer in mean compared to the RGB bands without FFS. For most smaller sizes of 0,15 meter and 0,30 meter designs the time upkeep was between 13 to 15 times longer while the greater design in several cases took 20 to 30 times to proceed with some cases around 40 up to 50 times.

The model accuracy in general is slightly higher using the FFS on the artificial layer Raster Stack with lowest values around 0.80 and maximum values of 1.00 compared to the use of RGB bands with respective values from around 0.70 up to 0.90 with highest values around 0.95. Overall the AOA reaches high values of around 95% to 100% coverage with lowest values of around 80%. In general the AOA is slightly greater using the RGB bands only. For the validation scores most predictions using the artificial layer Raster Stacks reach around 99% for 'hitrate' which is overall higher than using the RGB bands. Further the 'missrate' in most cases reaches values around 40% for study site 1 and around 60% to 70% for the other study sites with no significant differences for the used set of predictors.

Both sets of training points have comparable values in performance and validation scores for each study site and both predictor sets. Except for the second set of training points used for the first study area the accuracy and AOA is decreasing with increasing size of the training areas. Further the validation scores area decreasing with increasing sizes of training areas for both predictor sets except the 'hitrate' using the artificial layer Raster Stack which is constant around 99%.

Considering high accuracy and AOA for the performance values along with high 'hitrates' and low 'missrates' for the validation score the best results for each study site using RGB bands can be received using the smallest training areas (**Fig. 3**) while for the artificial layer Raster Stack the best results occur with different sizes of training areas for each study site (**Fig.4**).

Comparing the resulting predictions shows that the use of artificial layers leads to a far clearer separation of the classes while using the RGB bands results in more single cell which are assigned to different classes. All resulting images and tables can be retrieved from https://github.com/SchoenbergA/AOA-Impact/Results.

Table 3: *Results for the RGB bands ('rgb') for each study site (1, 2, 3) using respective training areas with set of training points (tp1, tp2) and radius in cycle form.*

Stk / design	accuracy	kappa	AOA	Hitrate	missrate	runtime
rgb1_tp1_c15	0.93	0.86	1.00	0.92	0.45	5.77
rgb1_tp1_c30	0.91	0.84	0.99	0.94	0.45	5.80
rgb1_tp1_c45	0.87	0.77	0.99	0.92	0.44	7.29
rgb1_tp1_c60	0.84	0.70	0.99	0.92	0.41	9.73
rgb1_tp1_c75	0.83	0.69	0.99	0.94	0.42	13.25
rgb1_tp1_c90	0.80	0.65	0.98	0.94	0.43	25.22
rgb1_tp2_c15	0.82	0.57	0.99	0.90	0.44	4.93
rgb1_tp2_c30	0.85	0.62	0.97	0.86	0.41	6.18
rgb1_tp2_c45	0.78	0.51	0.95	0.89	0.39	7.77
rgb1_tp2_c60	0.79	0.49	0.94	0.87	0.38	9.54
rgb1_tp2_c75	0.77	0.44	0.95	0.86	0.37	15.18
rgb1_tp2_c90	0.75	0.41	0.95	0.87	0.37	25.97
rgb2_tp1_c15	0.96	0.89	0.94	0.92	0.70	5.62
rgb2_tp1_c30	0.95	0.72	0.99	0.94	0.62	6.21
rgb2_tp1_c45	0.93	0.68	0.99	0.95	0.62	7.89
rgb2_tp1_c60	0.88	0.60	0.99	0.96	0.61	12.28
rgb2_tp1_c75	0.83	0.52	0.99	0.97	0.63	13.98
rgb2_tp1_c90	0.76	0.44	0.99	0.98	0.67	17.52
rgb2_tp2_c15	0.91	0.78	0.99	0.95	0.71	5.85
rgb2_tp2_c30	0.86	0.56	0.99	0.96	0.71	7.34
rgb2_tp2_c45	0.80	0.46	0.98	0.96	0.66	9.12
rgb2_tp2_c60	0.75	0.42	0.98	0.94	0.67	12.12
rgb2_tp2_c75	0.69	0.34	0.98	0.95	0.69	14.89
rgb2_tp2_c90	0.64	0.29	0.97	0.96	0.72	19.36
rgb3_tp1_c15	0.91	0.71	0.90	0.97	0.63	5.35
rgb3_tp1_c30	0.93	0.71	0.96	0.97	0.63	6.06
rgb3_tp1_c45	0.94	0.72	0.96	0.98	0.58	8.80
rgb3_tp1_c60	0.95	0.73	0.96	0.98	0.57	8.66
rgb3_tp1_c75	0.94	0.70	0.96	0.98	0.57	10.42
rgb3_tp1_c90	0.92	0.68	0.96	0.98	0.59	12.97
rgb3_tp2_c15	0.81	0.44	0.99	0.97	0.54	5.48
rgb3_tp2_c30	0.90	0.45	0.98	0.97	0.52	6.26
rgb3_tp2_c45	0.93	0.48	0.99	0.98	0.54	6.66
rgb3_tp2_c60	0.90	0.46	0.98	0.98	0.58	8.75
rgb3_tp2_c75	0.88	0.43	0.97	0.98	0.62	11.17
rgb3_tp2_c90	0.85	0.40	0.97	0.98	0.62	14.29

***Table 4**: Results for the artificial layer Raster Stack ('FFS') for each study site (1, 2, 3) using respective training areas with set of training points (tp1, tp2) and radius in cycle form. Factor RGB: The runtime compared to RGB bands (as factor).*

Stk / design	accuracy	kappa	AOA	Hitrate	Missrate	Runtime	factor RGB
FFS1_tp1_c15	1.00	1.00	0.88	0.92	0.40	75.84	13.13
FFS1_tp1_c30	0.99	0.97	0.95	0.94	0.40	111.57	19.23
FFS1_tp1_c45	0.96	0.92	0.94	0.95	0.44	185.31	25.41
FFS1_tp1_c60	0.94	0.89	0.93	0.94	0.39	391.76	40.28
FFS1_tp1_c75	0.91	0.83	0.93	0.95	0.39	599.03	45.22
FFS1_tp1_c90	0.89	0.78	0.93	0.96	0.40	810.44	32.14
FFS1_tp2_c15	0.89	0.78	0.93	0.92	0.46	63.81	12.95
FFS1_tp2_c30	0.88	0.74	0.95	0.89	0.37	83.11	13.44
FFS1_tp2_c45	0.93	0.72	0.94	0.91	0.36	174.25	22.43
FFS1_tp2_c60	0.84	0.64	0.92	0.85	0.37	249.33	26.13
FFS1_tp2_c75	0.82	0.60	0.92	0.89	0.35	553.58	36.46
FFS1_tp2_c90	0.81	0.54	0.91	0.81	0.36	594.89	22.91
FFS2_tp1_c15	1.00	1.00	0.96	0.94	0.46	74.49	13.25
FFS2_tp1_c30	1.00	1.00	0.96	0.97	0.48	96.06	15.48
FFS2_tp1_c45	1.00	0.99	0.95	0.98	0.53	183.62	23.28
FFS2_tp1_c60	0.98	0.96	0.95	0.99	0.55	322.44	26.25
FFS2_tp1_c75	0.96	0.73	0.95	1.00	0.59	604.88	43.27
FFS2_tp1_c90	0.93	0.69	0.96	1.00	0.63	820.39	46.84
FFS2_tp2_c15	0.97	0.90	0.96	0.99	0.58	76.33	13.06
FFS2_tp2_c30	0.98	0.95	0.94	0.99	0.56	143.47	19.56
FFS2_tp2_c45	0.96	0.89	0.94	0.99	0.62	207.16	22.72
FFS2_tp2_c60	0.94	0.69	0.94	1.00	0.62	399.93	32.99
FFS2_tp2_c75	0.89	0.62	0.94	1.00	0.62	571.87	38.40
FFS2_tp2_c90	0.85	0.56	0.94	1.00	0.63	972.02	50.22
FFS3_tp1_c15	1.00	1.00	0.95	0.98	0.35	76.63	14.32
FFS3_tp1_c30	1.00	1.00	0.94	0.99	0.36	92.92	15.33
FFS3_tp1_c45	1.00	0.99	0.81	0.99	0.70	158.30	17.99
FFS3_tp1_c60	0.98	0.77	0.81	1.00	0.64	325.47	37.60
FFS3_tp1_c75	0.98	0.76	0.81	0.99	0.64	529.84	50.85
FFS3_tp1_c90	0.97	0.74	0.87	1.00	0.59	492.86	38.00
FFS3_tp2_c15	0.92	0.72	0.94	0.99	0.73	78.02	14.23
FFS3_tp2_c30	0.96	0.87	0.91	0.99	0.48	122.70	19.59
FFS3_tp2_c45	0.94	0.67	0.89	0.99	0.61	180.34	27.08
FFS3_tp2_c60	0.91	0.56	0.90	1.00	0.60	318.98	36.46
FFS3_tp2_c75	0.92	0.48	0.91	1.00	0.56	435.71	39.01
FFS3_tp2_c90	0.89	0.43	0.92	1.00	0.65	538.08	37.66

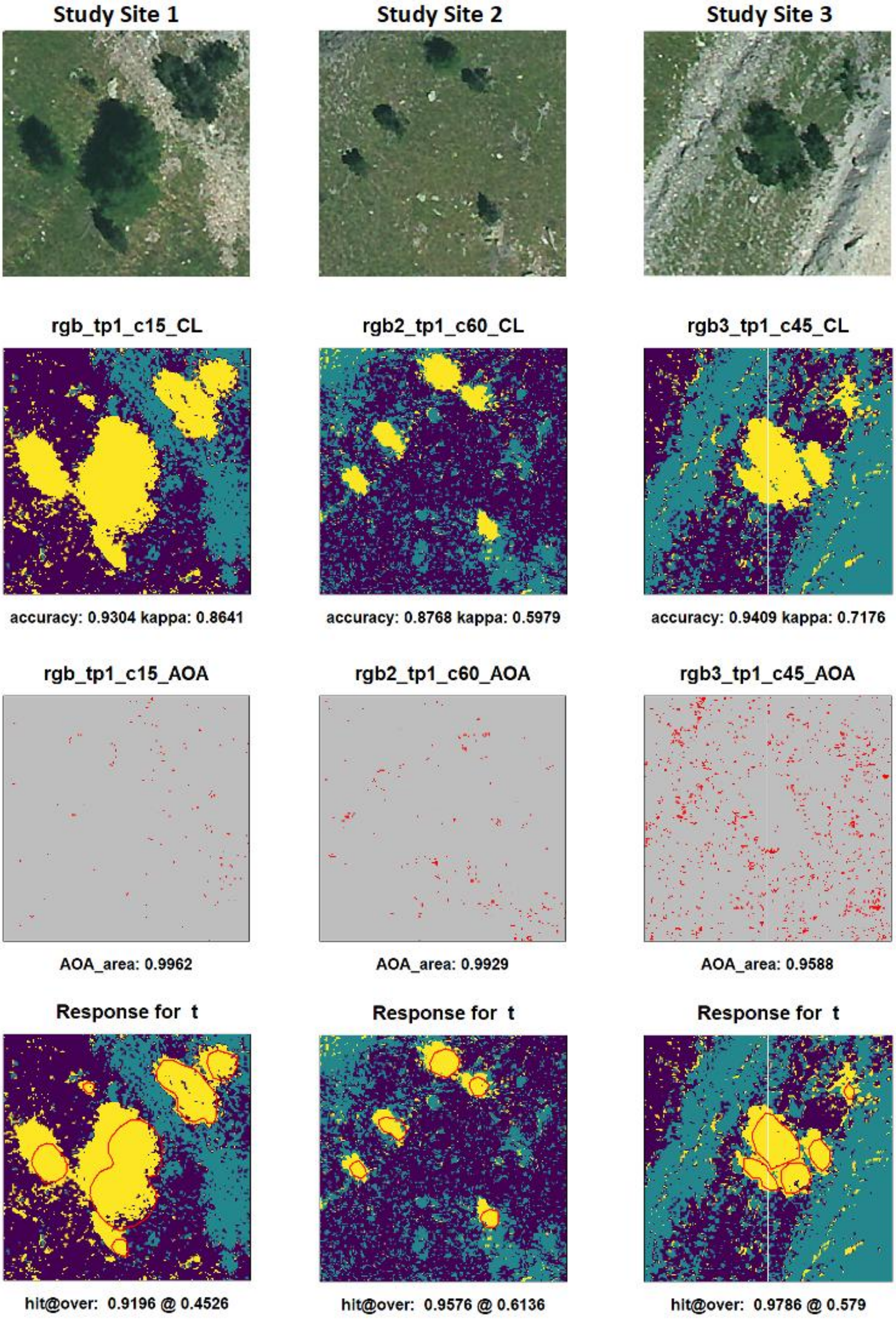

Figure 3: *Best resulting predicition, AOA and respective response for each study site based on RGB bands. Prediction classes: Trees (yellow), grass (dark blue) and soil (mint); AOA: inside (grey), outside (red); Response: response tree polygons (red line).*

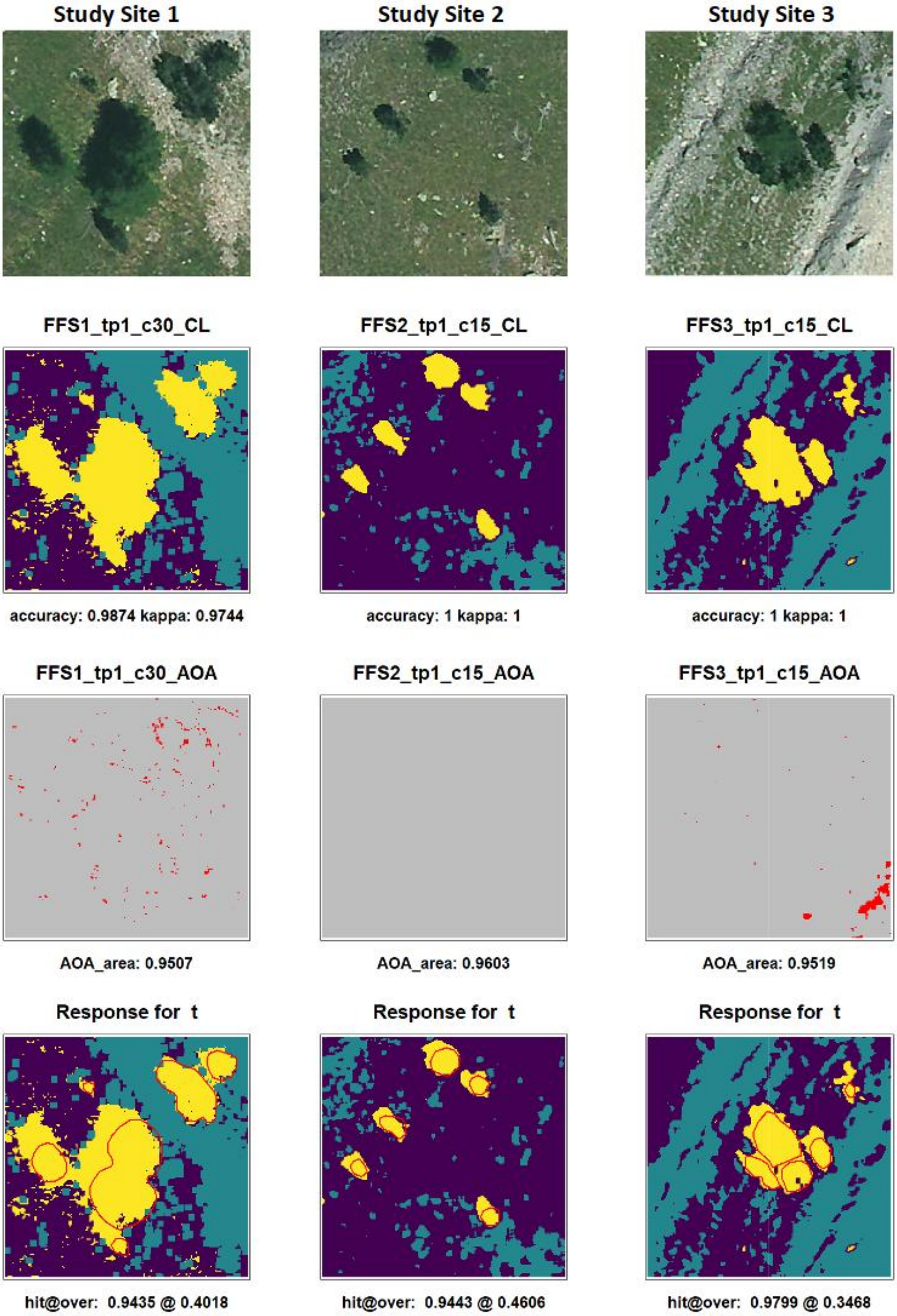

Figure 4: *Best resulting predicition, AOA and respective response for each study site based on artifically layers Raster Stack. Prediction classes: Trees (yellow), grass (dark blue) and soil (mint); AOA: inside (grey), outside (red); Response: response tree polygons (red line).*

4. Discussion

I hypothesized that increasing sizes of training areas would lead to decreasing performance for the prediction and lesser AOA. I can conclude that in most cases the accuracy is decreasing significantly and the AOA is slightly decreasing.

Further my aim for this study was to test if different settings of training points would have an impact on the results. I can conclude that both sets of training points lead to comparable values for each study site for both predictor sets. The second set of training points overall has slightly lower value compared to the first set. I suggest that the manual setting of training points is highly subjective and the first intuitive setting leads to more accurate results because for the second set alternative positions where used.

Furthermore I hypothesized that the use of artificial layers combined with a FFS would lead to higher performances but take far more time to proceed. With the results I can conclude that the performance is moderately increased using the artificial layers while the processing time is highly increased. In general the moderate improvement combined with the only slightly better validation scores and the far longer processing time seems to be less useful compared to the use of only RGB bands which lead to comparable high performances but take significantly less time to proceed. Further the use of the artificial layer Raster Stacks lead to a clearer separation of the classes. I suggest this effect is caused by the used filter functions. This probably leads to lesser accurate results because smallest spots of grass or stone would not be predicted correctly but on the other hand with RGB only there are many single cells assigned to class tree spread all over the images. I assume that the results for RGB only should be considerated with caution if the aim of the classification is to detect very small spots.

Despite of the used predictor set and training design there is a high amount of cells classified as trees outside the response layer leading to overall high values for the 'missrate' around 30% to 40%. Those cells mainly are distributed in the areas of the treeshadows. Due to the missing of a class for treeshadows the cells are assigned to the class for trees. Therefore the 'missrate' probably would be significantly lesser if treeshadows would be trained in an additionally class.

5. Conclusion

With the results in total I can conclude that for small scale classifications on high resolution RGB images small sizes of training areas lead to better results while the subjective setting of the training areas has lesser impact on the results. The use of artificial layers computed with the method in this study can improve the performance compared with only using the RGB bands but highly increases the processing time. Finally there is no significant impact on the AOA between using only RGB bands or the artificial layers while the size of the training areas have only slightly an impact on the AOA.

References

BREIMAN, L. (2001): Random forests. Mach. Learn. 45, 5–32

CAO, W., LI, B. & Y. ZHANG (2003): A remote sensing image fusion method based on PCA transform and wavelet packet transform. International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003. Nanjing, China.

COULSTON, J., BLINN, C., THOMAS, V. & R. WYNNE (2016): Approximating Prediction Uncertainty for Random Forest Regression Models Photogrammetric Engineering & Remote Sensing Volume 82, Issue 3, March 2016, Pages 189-197

CUTLER, D., EDWARDS JR, T., BEARD, K., CUTLER, A., HESS, K., GIBSON, J. & J. LAWLER (2007): Random Forests for Classification in Ecology. Ecology 88(11). 2783–2792.

DEMSAR, U., HARRIS, P., BRUNSDON,C.,FOTHERINGHAM, A. & S. MCLOONE (2013): Principal Component Analysis on Spatial Data: An Overview. In: Annals of the Association of American Geographers 103(1). 106-128.

ESTORNELL, J. MARTI-GAVILA, J. SEBASTIA, M. & J. MENGUAL (2013): Principal component analysis applied to remote sensing Modelling in Science Education and Learning Volumen 6 (2). No. 7.

FIERENS, F. & P. ROSIN (1994): Filtering Remote Sensing Data in the spatial and feature domains. SPIEVol.2315. Imageand SignalProcessingfor RemoteSensing. 472-482.

FOX, E., HILL, R., LEIBOWITZ, S., OLSEN, A., THORNBRUGH, D. & M. WEBER (2017): Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. Environ Monit Assess 189: 316.

HIJMANS, R. (2020): raster: Geographic Data Analysis and Modeling. R package version 3.3-13. https://CRAN.R-project.org/package=raster.

HUNT E.R., CAVIGELLI M., DAUGHTRY C. , MCMURTREY III J. & C. WALTHALL (2005): Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. Precision Agriculture 6. 359–378.

HUNT, E.R., DORAISWAMY, P.C., MCMURTREY, J.E., DAUGHTRY, C.S., PERRY, E.M. & B. AKHMEDOV (2013): A visible band index for remote sensing leaf chlorophyll content at the canopy scale. Int. J. Appl. Earth Obs. Geoinf. 21, 103–112.

JIN, H., STEHMAN,S. & G. MOUNTRAKISA (2014): Assessing the impact of training sample selection on accuracy of an urban classification: a case study in Denver. International Journal of Remote Sensing. Vol. 35, No. 6, 2067–2081.

KUHN, M. & K. JOHNSON (2013): Applied Predictive Modeling, 1st ed. Springer, New York.

KUHN M. (2020): caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret.

LEUTNER B. , HORNING N. AND J. SCHWALB-WILLMANN (2019): RStoolbox: Tools for Remote Sensing Data Analysis. R package version 0.2.6. https://CRAN.R-project.org/package=RStoolbox.

MEYER, H. & E. PEBESMA (2020): Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models. arXiv:2005.07939.

MEYER, H., REUDENBACH, C., WÖLLAUER,S. & T. NAUSS (2019): Importance of spatial predictor variable selection in machine learning applications – Moving from data reproduction to spatial prediction. Ecological Modelling 411.108815.

MEYER, H., REUDENBACH, C., HENGL, T., KATURJI, M. & T. NAUSS (2018): Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environ. Model. Softw. 101, 1–9.

MEYER, H. (2020): CAST: 'caret' Applications for Spatial-Temporal Models. R package version 0.4.2. https://CRAN.R-project.org/package=CAST.

R CORE TEAM (2020): R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna. Austria. URL https://www.R-project.org/.

RAY, S., SINGH, J., DAS, G. & S. PANIGRAHY (2004): Use of High Resolution Remote Sensing Data for Generating Sitespecific Soil Management Plan Conference: The Int. Arch. Photogram. Rem. Sens. & Spatial Inform. Syst At: Istanbul Volume: 35 (B7)

THE IDB PROJECT (2020): Index Database https://www.indexdatabase.de/

ZHAO, D., JIAO, Y., WANG, J., DING, Y. LIU, Z., LIU, C., QIU, Y., ZHANG, J., XU, Q. & C. WU (2020): Comparative performance assessment of landslide susceptibility models with presence-only, presence-absence, and pseudo-absence data. Journal of Mountain Science 17(12).