RIVER FLOODPLAIN VEGETATION CLASSIFICATION USING MULTI-TEMPORAL HIGH-RESOLUTION COLOUR INFRARED UAV IMAGERY

W.K. van Iersel[a]*, E.A. Addink[a], M.W. Straatsma[a], H. Middelkoop[a]

[a] Utrecht University, Department of Physical Geography, Faculty of Geosciences, Heidelberglaan 2, 3584 CS Utrecht, the Netherlands - (w.k.vaniersel@uu.nl)

KEY WORDS: river floodplains, land cover classification, multi-temporal data, aerial photography, UAV, DSM

ABSTRACT:

To evaluate floodplain functioning, monitoring of its vegetation is essential. Although airborne imagery is widely applied for this purpose, classification accuracy (CA) remains low for grassland (< 88%) and herbaceous vegetation (<57%) due to the spectral and structural similarity of these vegetation types. Increased availability of Unmanned Aerial Vehicles (UAV) allows low-cost production of high-resolution orthophotos and digital surface models (DSMs). Multi-temporal DSMs and orthophotos may be used as input for an improved classification methodology, using differences in phenological changes between vegetation types. The aim of this study was (1) to evaluate the improvement of the CA when using multi-temporal UAV-derived imagery and (2) to determine which layers of a multi-temporal imagery and derived DSMs yield an optimal balance between CA and acquisition effort. During six field surveys with six to ten weeks intervals over one year, a floodplain section along the lower Rhine, the Netherlands, was recorded with true-colour and false-colour imagery with a UAV. In several segmentation-classification-evaluation loops we determined the most important set of variables and the data layers providing them. Our main conclusions are (1) Multi-temporal data input greatly improve CAs of grassland and herbaceous vegetation classes in floodplains: user's accuracies exceed 90%, and (2) the input data contributing most to these high CAs are NDVI layers from winter, spring and summer, and nDSM layers from winter and end of summer.

1. INTRODUCTION

The different user functions of river floodplains might compete with each other, such as water storage during peak discharge, ecological value, agricultural practices and recreation. They also result in a different land cover of the floodplain, or may have demands concerning land cover characteristics, such as a low hydraulic roughness for flood protection. Ecological restoration in floodplains has generally resulted in more heterogeneous floodplain vegetation (Göthe et al., 2016). Due to vegetation phenology and succession, both ecological and hydraulic characteristics of the floodplain change over time. To evaluate floodplain functioning, monitoring of its vegetation is essential. However, using the current classification methods of airborne surveys classification accuracy (CA) remained low for grassland and herbaceous vegetation (Knotters and Brus, 2013) due to the spectral and structural similarity of these vegetation types (Geerling et al., 2007). For floodplains along the lower Rhine river in the Netherlands, Knotters and Brus (2008)

reported a user's accuracy (UA) of 22-57% for herbaceous vegetation and 74-88% for grassland. Moreover, current vegetation mapping of these floodplains is based on manual identification and classification of vegetation objects on airborne imagery. Until now vegetation classification has mainly been performed using single-moment datasets. The increased availability of Unmanned Aerial Vehicles (UAV) allows low-cost production of high-resolution orthophotos and digital surface models (DSMs). Multi-temporal DSMs and orthophotos may be used as input for an improved classification methodology, using the differences in development over time of object variables as vegetation height, colour and photosynthetic activity (Xie et al., 2008; Yan et al., 2008). The aims of this study were (1) to evaluate the improvement of the CA with multi-temporal data and (2) to determine which layers of a multi-temporal imagery and derived DSMs yield an optimal balance between CA and acquisition effort.
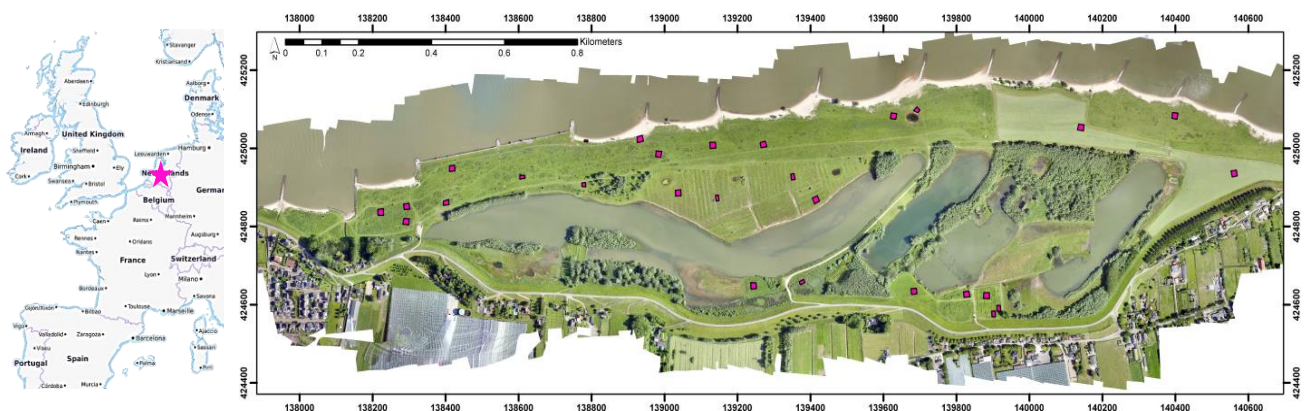

Figure 1. Location of study area in the Netherlands (left). True colour orthophoto of Breemwaard in June with reference plots (magenta).

2. STUDY AREA

We studied the Breemwaard floodplain, sized 116 ha, in the Netherlands. Approximately 30% of the area is used as hayfield and is cut several times per year. Several parts of the floodplain are managed by private owners and used as willow fields, reed fields or for grazing ponies. The remaining part is a nature area, which is managed by the state forestry and grazed by cows and ponies. Certain part of the nature area, which are fenced off for cattle, developed into riparian woodland. This variability in management of the area has resulted in heterogeneous vegetation types and structure, which made it well suited for classification experiments of riparian vegetation.

3. METHODS

3.1 Imagery and DSMs data layers

During six field surveys with six to ten weeks intervals over one year, the entire study area was recorded with true-colour and false-colour imagery at 5 cm resolution from a UAV. Surveys were done in February, April, June, September, November 2015 and January 2016. The UAV images were converted to true and false colour orthophotos and point cloud DSMs of the study area for each survey using the Structure-from-Motion (SfM) method (Lucieer et al., 2014). The false colour orthophotos were used to derive the Normalized Difference Vegetation Index (NDVI) with the near-infrared and blue band. The true and false colour imagery was used to establish respectively RGB DSMs and NIR DSMs. To determine the height of the vegetation above the terrain from the imagery, the DSM of each survey was normalized by abstracting the DSM of February 2015 from it. Thereby, it was implicitly assumed that the February DSM represents the Digital Terrain Model (DTM), which is reasonable for herbaceous vegetation in leaf-off conditions. The DTM of February was denoised by gridding the initial point cloud obtained from the SfM preprocessing to a 0.2 m resolution grid. Average spacing of the initial dense point cloud was 0.08 m. Large water areas were masked in the normalized DSMs (nDSMs), because of unreliably high values in the dense point clouds.

A layer is a raster-based derivative of a UAV campaign, with nine layers per campaign: R-G-B from the true color orthophoto NIR-R-G from the false colour orthophoto, RGB-nDSM, NIR-nDSM, and NDVI. The six campaigs provided a total of 54 layers.

Field reference data for the vegetation classes was collect during the survey in September 2015. Vegetation classes pioneer, grassland, herbaceous and reed vegetation were assigned to field plots of 15m x 15m based on dominant species and vegetation height in the field.

3.2 Segmentation-classification-evaluation loop

The layers (section 3.1) were used as input for the segmentation, leading to segment attributes: the mean and standard deviation of the 54 layers, plus the length-with ratio, giving a total of 109 attributes. These segment attributes were used as explanatory variables in the RF classifier. An overview of the segmentation-classification-evaluation loop is shown in figure 2. To find the most important variables the segmentation was performed with Definiens Developer. Initially all 54 data layers were used as input for segmentation. A scale parameter of 50 showed the best fit in object size with previous results based on multi-temporal NIR imagery only. Shape and compactness were both kept at 0.1 during every loop. The segments overlapping with field plots were classified using the Random Forest classification

(Breiman, 2001). The object variables available for the classification were the length/width ratio of the objects and their mean and standard deviation from all orthophotos, NDVI and nDSM layers for each time step. The classification was trained with 50% of the reference data from the field and validated with the remaining 50%. The 10 most important variables for the RF classification, based on the decreased mean accuracy and gini index, were used to determine which layers were to be used in a new segmentation-classification-evaluation. Only those layers with the 10 most important variables were used to perform the segmentation again. For each segmentation, the scale parameters were adjusted to keep the number of objects constant (deviation max. ±10.000) over the segmentation loops. The new segmentation would again be classified by taking all the same segment features as in the first segmentation into account.

This segmentation-classification-evaluation loop was repeated three times until convergence to a static set of variables was obtained (Figure 2). The final set of variables was each time one of the 10 most important layers in all of the three loops. During the final classification of the study area only the final set of variables was used to segment and classify the study area.
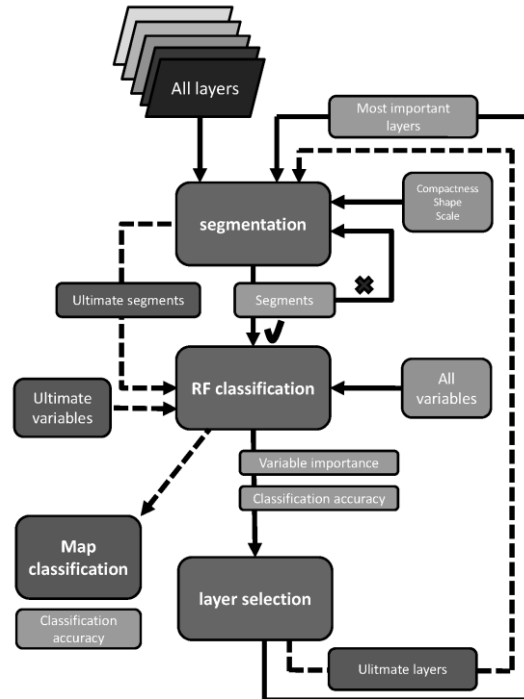

Figure 2. Segmentation-classification-evaluation loops, which were performed three times, plus an additional segmentation and classification with only the ultimate selection of layers.

4. RESULTS

4.1 Multi-temporal segmentation-classification-evaluation

For the segmentation of the first loop with all data layers a scale parameter of 50 was used. In the second loop only nine data layers were used, in the third loop ten data layers and in the fourth loop also nine (table 1). To keep objects of approximately the same size over the loops, the scale parameters could be kept as 50 in the second and third loop, but was adjusted to 45 in the fourth loop.
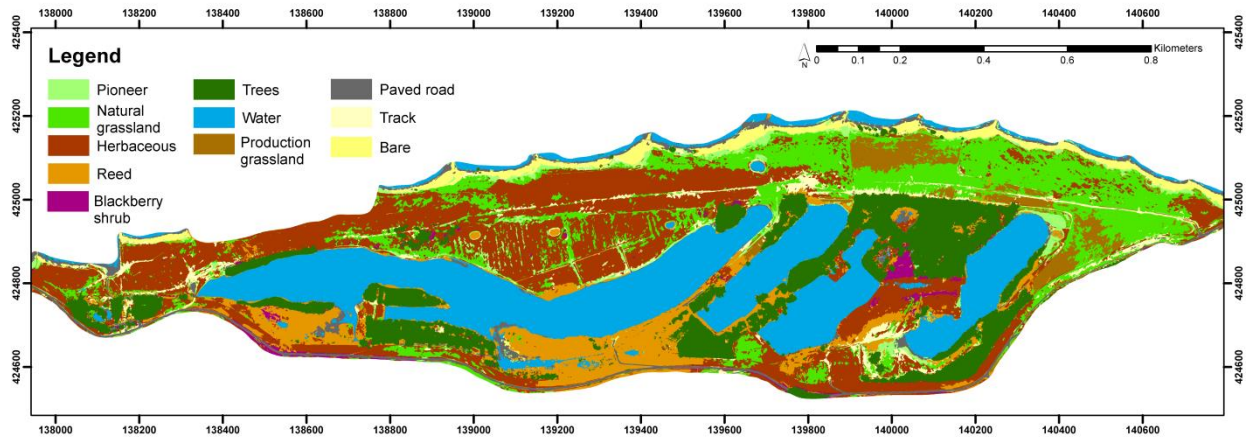
Figure 3. Map of classification of the Breemwaard study area.

Table 1. Most important variables for the four iterations. The position of the final variables is highlighted in grey shading in the first three loops. M=mean, first 4 numbers = year, last two number = month, last part of code is the data layer.

Top 10	Explanatory variables of loop1	Explanatory variables of loop2	Explanatory variables of loop3	Final selection of variables
1	M_201502_nDSM_NIR	M_201601_NDVI	M_201504_NDVI	M_201504_NDVI
2	M_201504_NDVI	M_201506_NDVI	M_201601_NDVI	M_201502_nDSM_NIR
3	M_201506_NDVI	M_201511_RGB_B	M_201509_NIR_G	M_201506_NDVI
4	M_201506_NIR_NIR	M_201504_NDVI	M_201506_NDVI	M_201509_nDSM_NIR
5	M_201509_nDSM_NIR	M_201509_nDSM_NIR	M_201509_nDSM_NIR	M_201509_nDSM_RGB
6	M_201509_NIR_G	M_201509_nDSM_RGB	M_201509_nDSM_RGB	M_201601_NDVI
7	M_201509_NIR_B	M_201502_nDSM_NIR	M_201509_NIR_B	
8	M_201509_nDSM_RGB	M_201601_NIR_NIR	Length/width	
9	M_201601_NDVI	M_201601_RGB_B	M_201502_nDSM_NIR	
10	Length/Width	M_201502_NIR_B	SD_NIR_R	

Each segmentation-classification-evaluation loop had a different set of 10 most important variables (table 1). However, all loops had six variables present top 10 of most important variables for their RF classification, which are highlighted in table 1. The ultimate set of layers is composed of NDVI and nDSM mean values from different time steps. Moreover, the important mean nDSM of February is obtained from false colour imagery and the mean nDSM from November from true and false colour imagery. The important mean NDVI variables are obtained from two moments in spring (April and June) and one in winter (January).

4.1 Ultimate multi-temporal segmentation and classification

For the segmentation with the ultimate data layers (table 1) the scale parameters was adjusted to 37 to maintain comparable object sizes with the other loops. The RF classification using only the ultimate variables layers was applied on all objects resulting from the segmentation with only the ultimate data layers. Even with this data reduction of 94.5% user's accuracies for grassland and herbaceous vegetation are still 92.6% and 92.0% respectively. The map of all classified object in general represent the spatial arrangement of the different vegetation classes well (figure 3). Some classes still get mixed up in the classification, like paved road and water surfaces.

5. DISCUSSION AND CONCLUSIONS

Even though the data reduction is 94.5%, the ultimate variables are still from 5 different time steps. This means that still data acquisition needs to take place at five moments in time to get the desired classification accuracies. On the other hand, five out of six ultimate variables are derived from false colour imagery. Moreover, these ultimate layers cover the extremes of the phenology of the vegetation, since (1) the NDVI changes most from winter to spring and summer, and (2) vegetation height difference is largest between winter and the end of summer.

It also needs to be noted that high correlation can be expected between nDSM_NIR and nDSM_RGB, resulting in an even further data reduction. The contribution of the individual layers has not been checked yet.

For practical reasons it will also be interesting to see the CA when the procedure is repeated with fewer time steps. For this purpose the approach needs to be changed from keeping layers for the next segmentation-classification loop to dropping all layers of non-important variables during the evaluation. Each loop a time step can be removed, until the classification accuracy becomes unacceptably low.

The confusion of the classes water and paved road in the ultimate mapped classification can be explained by their low mean NDVI and low mean height values which also do not change over time.

Our main conclusions are:

- Multi-temporal data input greatly improved CAs of grassland an herbaceous vegetation classes in floodplains: user's accuracies exceed 90%.

- The input data contributing most to this high CAs are multi-temporal NDVI layers from both winter, spring and summer and nDSM layers from winter and end of summer. These layers cover the extremes of the phenology of the vegetation.

- Non-vegetation classes behaving similarly in NDVI and height over time, like water and paved roads, still get confused in the classification.

ACKNOWLEDGEMENTS

REFERENCES

Breiman, L. (2001), Random Forests, Machine Learning 45(1), 5-32.

Geerling, G. W., Labrador-Garcia, M., Clevers, J. G. P. W., Ragas, A. M. J., & Smits, A. J. M. (2007). Classification of floodplain vegetation by data fusion of spectral (CASI) and LiDAR data. International Journal of Remote Sensing, 28(19), pp. 4263-4284.

Göthe, E., Timmermann, A., Januschke, K., & Baattrup-Pedersen, A. (2016). Structural and functional responses of floodplain vegetation to stream ecosystem restoration. Hydrobiologia, 769(1), pp. 1-14

Knotters, M., & Brus, D. J. (2013). Purposive versus random sampling for map validation: a case study on ecotope maps of floodplains in the Netherlands. Ecohydrology, 6(3), 425-434.

Lucieer, A., De Jong, S. M., & Turner, D. (2014). Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. Progress in Physical Geography, 38, pp. 97–116

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. Journal of plant ecology, 1(1), pp. 9-23.

Yan, E., Wang, G., Lin, H., Xia, C., & Sun, H. (2015). Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. International Journal of Remote Sensing, 36(2), 489-512.