ABSTRACT

This study aims at finding successful methods for mapping hydraulic resistance of vegetation in a river floodplain using imaging spectroscopy and field data. The river discharge capacity is affected by the presence of different vegetation types and their hydraulic resistance, as its value affects the velocity of the water regime.

The concept of plant functional types (PFT) has been introduced to cluster species with similar morphology and phenology. Field measurements indicate a relationship between type of PFT and hydraulic resistance. In this study we investigated if hydraulic resistance of vegetation in a river floodplain can be estimated by combining a PFT classification with biomass maps derived from imaging spectroscopy data.

Mixture Tuned Matched Filtering (MTMF) was used for the classification of PFTs in the study area as this method could deal with the heterogeneous characteristics of the study site. MTMF performed better than the classical LSU classification with a 0.62% of accuracy at the collected sampling points.

For the biomass estimation, statistical methods using vegetation indices were applied. Due to the heterogeneity of the area, the methodology applied was PFT-dependent. The biomass best-correlated vegetation indices were the Soil Adjusted Vegetation Index (SAVI) and the Weighted Difference Vegetation Index (WDVI). Statistical approaches gave successful results ($r^2$ ranging between 0.4 and 0.7) but did not yield a generally applicable empirical equation.

As a conclusion, the methodology presented in this study can be used to estimate hydraulic resistance in a river floodplain.

INTRODUCTION

The management of natural landscapes requires knowledge and understanding of their biophysical characteristics e.g. biomass quantity. Productivity, species richness and biodiversity are valuable inputs for monitoring of changes by environmental managers (i). Moreover, the biomass distribution frequency, together with the existing maps of vegetation types can be used as inputs for modeling issues dependent on vegetation in river flood models (ii).

Under the influence of changing climate, the flooding risk along the river Rhine is expected to increase in the coming decades. As a result river managers develop alternative flood risk management strategies and measures to anticipate higher peak discharges in the future. In the Netherlands, river floodplains are developed in heterogeneous vegetated areas (e.g., grass, shrub and softwood forest) in which several land use functions are combined, including flood protection, nature rehabilitation and recreation (iii).
In a floodplain area, river managers are required to assess the influence of riverside natural vegetation on the discharge capacity of the river. River management can be improved by studying the presence of the different vegetation types and their hydraulic resistance. Its value affects the velocity of the water regime, with vegetation structure playing a role through biophysical quantities such as height, diameter and density of stems, height of shoot branches, density of branches and the surface area of leaves.

In order to make a decision regarding management measures and biomass reduction measures, spatial explicit information on the vegetation biomass in river floodplains in the Netherlands is needed. The last decades, Remote Sensing has offered new approaches, providing a spatially continuous data set reducing interpolation techniques using expensive and time consuming field campaigns. Furthermore, remote sensing enables the use of study areas on inaccessible environments (iv).

In this study we investigated if hydraulic resistance of vegetation in a river floodplain can be estimated by combining a Plant Functional Type’s classification and biomass maps derived from imaging spectroscopy data. The study area is a floodplain located along a branch of river Rhine in The Netherlands.

**METHODS**

The study area

The floodplain Millingerwaard is located to the east of Nijmegen along the river Waal, one of the main branches of the river Rhine in the Netherlands. The floodplain belongs to the nature reserve Gelderse Poort, a large floodplain area of 1200 ha along the river Rhine, close to the Dutch-German border.

The area is subject to several nature rehabilitation projects, which means that individual floodplains are taken out of agricultural production and are allowed to undergo their natural succession. This has resulted for the Millingerwaard in a heterogeneous landscape forming a mosaic pattern of different succession stages (pioneer, grassland, shrubs, softwood forest). However, the discharge capacity of the river should be above the critical safety levels during flooding events (v).

The concept of plant functional types (PFT) has been introduced to group clusters of species with similar morphology and phenology. For this study, vegetation in the floodplain was divided into four main PFTs: “Grazed Grasslands”, “Mixed Herbs”, “Shrubs” and “Forest”.

Data acquisition and pre-processing

A field campaign was conducted in the study area in July 2004 in which Hyperspectral imagery was acquired at the same time field data were gathered for 21 sample points.

Imaging spectroscopy data for the Millingerwaard were acquired on the 28th of July with the Hymap airborne instrument. A complete spectrum over the range of 450-2480 nm was recorded with a bandwidth of 15-20 nm by 4 spectrographic modules. Each module provided 32 spectral channels giving a total of 128 spectral measurements for each pixel. However, the delivered data contained 126 bands because the first and last bands of the first spectrometer were deleted during pre-processing. The spatial resolution of the acquired image was 5 m. The flight line was oriented close to the solar principal plane to minimize directional effects. Hymap images were geo-atmospherically processed to obtain geo-coded top-of-canopy reflectance data. Visibility was estimated by combining sun photometer measurements with Modtran4 radiative transfer simulation. Further explanation of the campaign can be found in (v).

A subsample of the acquired image was used in this study. A mask operation was applied to select the vegetated areas based on a NDVI minimum threshold of 0.2 and a maximum reflectance at the wavelength of 665 nm (Band 13) of 7.4%.

At the same time, a ground campaign was organized to describe the vegetation present at 21 locations in the floodplain according to the method of Braun-Blanquet (vi). In each location, a plot
x 2 meters was selected with a relatively homogeneous vegetation cover. Three biomass samples of subplots measuring 0.5 x 0.5 m were taken per plot at 0.5 cm above the ground level. The collected material was stored in paper bags; air-dried, first for 5 days at room temperature in open bags, and subsequently for 24 h at 70°C, and weighed to obtain the dry biomass (g). Dry biomass values of each subplot were stored in weight/area units. The dry biomass per plot was calculated by means of the average between the dry biomass of the corresponding subplots. Final values were obtained in g/m².

The PFT method was based on the vegetation description that was made during the field work mentioned above. The parameters used for the classification were “height of the vegetation”, “management” and “main species composition” (Table 1).

Table 1: Range of vegetation height consider for the PFTs “grazed-grasslands” and “Mixed-herbs”, management under which they are held and main characteristic species per PFT existing in the study area.

<table>
<thead>
<tr>
<th>Height ranges (cm)</th>
<th>Management</th>
<th>Dominant species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grazed-Grasslands</td>
<td>Grazing</td>
<td><em>Trifolium repens</em>, <em>Potentilla reptans</em>, <em>Cynodon dactylon</em></td>
</tr>
<tr>
<td>Mixed-Herbs</td>
<td>No</td>
<td><em>Rubus caesius</em>, <em>Calamagrostis epigejos</em>, <em>Urtica dioica</em></td>
</tr>
</tbody>
</table>

The samples without height information were included in one group or the other based on information from photographs taken during the field work. Moreover, managed and non-managed plots were also separated to get a more accurate relationship between biomass and the different vegetation indices. The management can influence the age of the vegetation. In this case, the managed areas corresponded to the samples considered Grazed-grassland.

Classification of Plant Functional Types

The Mixture Tuned Matched Filtering (MTMF) method was chosen to make the PFT’s classification in the Millingerwaard floodplain because other authors (vii, viii) had used this classification method in previous studies for areas with relatively high vegetation heterogeneity (e.g. prairie sites mixed with woody areas).

MTMF (xix) is a partial unmixing algorithm that is capable of determining target abundance within a pixel. It has in common with the Linear Spectral Unmixing (LSU) that the constraints about the sum of the classes per pixel equal to 1, and the use of values is constrained between 0 and 1. Moreover, none of the two methods in ENVI software can be applied fully constrained. On the other hand, using MTMF, information about all the present end-members spectra is not needed, what makes MTMF different from LSU and more similar to Matched Filtered models (xix). The end-member collection to make the classification was conducted from the field samples pixels in the case of grasslands and mixed herbs; and manually in the map for the rest of the end-members (soil, shrubs, forest and agricultural crops).

Not all the Hymap bands were used in the classification process. Before applying MTMF, a previous transformation called Maximum Noise Fraction (MNF) was carried out. From MNF transformation, the reliability of the eigenvectors created is obtained. To avoid the use of noisy eigenvectors, eigenvalues greater than 5 were not considered in the classification. Then, the classification was performed with 8 eigenvectors from the MNF file.

Because of the heterogeneity of the area and the overlap of spectral signatures of different PFTs, the classification did not performed in a totally constraint way. Masks were applied to avoid values of the output classes below 0 and over 1 per pixel.
A test of the error was carried out, checking if non-possible values (above 1 and below 0) were present over the image, or if they were PFT related. As the affected pixels were present all over the area and in a low percentage, the error was attributed to the difficulty of the classification.

**Dry biomass estimation**

Several studies have shown that biomass can be estimated from remote sensing data by means of statistical methods based on vegetation indices (x-xx). In this study, four vegetation indices were selected to be later related with measured vegetation biomass in Millingerwaard. The four indices derived from the 2004 Hymap image were the Normalized Difference Vegetation Index (NDVI) (xxi), the Weighted Difference Vegetation Index (WDVI) (xxii), the Soil Adjusted Vegetation Index (SAVI) (xxiii), and the Simple Ratio (RSR) (xxiv).

The four vegetation indices were obtained for each sample location and compared to the dry biomass weight from the aboveground vegetation collected in the same location in the field. Some uncertainties were found as the location of each sample point was not the central point of a pixel. The maximum deviation between the sampling coordinates and the respective pixel centre is 2.91 meters.

The process to evaluate the relationship between the four vegetation indices and biomass was done using a cross-validation technique in which an equation is got from the VI-biomass relationship of 20 samples, while validating with the 21st sample. This process is done per PFT with all the samples. The Root Mean Square Error of the Cross Validation (RMSECV) and the Root Mean Square Error of Prediction (RMSEP) were calculated.

\[
RMSEP = \sqrt{\frac{\sum (r - r')^2}{n}}
\]

Where:
- \(r\) = Measured biomass in the field.
- \(r'\) = Corresponding biomass according to the regression line.
- \(n\) = Number of biomass samples included in the analysis.

The RMSECV analyses the difference in between the value of a sample and the predicted value at that point depending on the rest of the samples. Then, the validation is independent in this case from the field sampling in each point.

\[
RMSECV = \sqrt{\frac{\sum (r - r')^2}{n-1}}
\]

Where:
- \(r\) = Measured biomass in the field.
- \(r'\) = Corresponding biomass according to the regression line calculated from the rest of the measurements from the field.
- \(n\) = Number of biomass samples included in the analysis.

The regression analysis was conducted for three sample sets: grazed-grassland (n=7), mixed-herbs (n=14) and the whole dataset (n=21). Linear (\(y = ax + b\)) and exponential (\(y = a \cdot e^{bx}\)) or logarithmic (\(y = a \cdot \ln(x) + b\)) relationships were considered as in previous studies it is demonstrated these are the ones explaining the relationship between vegetation indices and quantitative variables (xxi).

The biomass map was obtained with the results of the regression analysis by means of applying the best relationship proportionally per coverage of PFT per pixel. The general equation used was:
\[ B = \left[ \text{Best relation}(GR) \times (\%\text{grasslands}) \right] + \left[ \text{Best relation}(MH) \times (\%\text{mixed-herbs}) \right] \]

Where:

- B = Dry biomass
- Best relation(GR) = Best regression result between grasslands biomass and the different VI
- Best relation(MH) = Best regression result between Mixed-herbs biomass and the different VI.
- %grasslands = Grasslands coverage percentage per pixel.
- %mixed-herbs = Mixed-herbs coverage percentage per pixel.

**RESULTS**

The resulting map after applying the MTMF classification technique for the existing PFT in Millingerwaard is presented in Figure 1. When comparing the results with the data from the field, 50% of grasslands and 79% of mixed-herbs sample points were well classified. The results were satisfactory when comparing with what it could be observed in the field and aerial pictures.

The maps derived from the calculations of the four vegetation indices are shown in Figure 2 and the minimum and maximum values of each vegetation index in the Millingerwaard are summarized in Table 2.

In general, the vegetation index maps (Figure 2) show that agricultural areas are the ones in the north with a rectangular shape, showing the highest values (lightest). The forest, on the other hand, has the lowest values as can be seen in the eastern part of the area. The western area is where more grasslands and mixed-herbs are present; they have high NDVI, WDVI and SAVI. In the case of the RSR, nevertheless, the values are similar to the ones of the forest.

**Figure 1: Plant Functional Types classification in the floodplain Millingerwaard applying Mixed-Tuned Matched Filtering classification technique.**

**Table 2: Minimum, maximum, mean and standard deviation for the values obtained for NDVI, WDVI, SAVI and RSR in the study area.**

<table>
<thead>
<tr>
<th>Index</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>0</td>
<td>0.8913</td>
<td>0.0498</td>
<td>0.1837</td>
</tr>
<tr>
<td>WDVI</td>
<td>-4.147</td>
<td>5.5687</td>
<td>0.6289</td>
<td>1.3438</td>
</tr>
<tr>
<td>SAVI</td>
<td>0</td>
<td>0.1572</td>
<td>0.0059</td>
<td>0.0221</td>
</tr>
<tr>
<td>RSR</td>
<td>0</td>
<td>11.426</td>
<td>0.3534</td>
<td>1.3592</td>
</tr>
</tbody>
</table>
After correlating the values of the vegetation indices of the sample pixels with the measured biomass values of the sampling locations, VI-biomass relationships were derived. The results of the correlation analysis with linear regression are presented in Table 3. The values for exponential and logarithmic regression can be found in Table 4.

The best result found for grazed-grasslands was the exponential relationship with the SAVI index. The biomass of the mixed-herbs, nevertheless, was better correlated by a logarithmic relationship with WDVI index (Table 4). This means that soil background effects are important in the total reflectance of the area, as these two indices are the ones that minimize such background effects.

Table 3: Results summary of the linear regression applied between the VI: NDVI, WDVI, SAVI and RSR and biomass. The coefficients a and b correspond to the linear equation \( y = ax + b \), \( r^2 \) = determination coefficient, RMSEP= Root Mean Squared Error of Prediction and RMSECV= Root Mean Squared Error of Cross-Validation.

<table>
<thead>
<tr>
<th>VI</th>
<th>PFT</th>
<th>n</th>
<th>LINEAR REGRESSION</th>
<th>RMSEP</th>
<th>RMSECV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>a</td>
<td>b</td>
<td>( r^2 )</td>
</tr>
<tr>
<td>NDVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>-0.00001</td>
<td>0.7796</td>
<td>0.0577</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>0.0003</td>
<td>0.5338</td>
<td>0.4321</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>0.0002</td>
<td>0.6638</td>
<td>0.1381</td>
</tr>
<tr>
<td>WDVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>-0.0026</td>
<td>39.414</td>
<td>0.2452</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>0.0035</td>
<td>10.125</td>
<td>0.559</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>0.0014</td>
<td>24.814</td>
<td>0.127</td>
</tr>
<tr>
<td>SAVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>0.00007</td>
<td>0.1231</td>
<td>0.3702</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>0.00008</td>
<td>0.0558</td>
<td>0.5713</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>0.00006</td>
<td>0.0889</td>
<td>0.1091</td>
</tr>
<tr>
<td>RSR</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>-0.0023</td>
<td>41.702</td>
<td>0.0706</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>0.0033</td>
<td>19.737</td>
<td>0.2338</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>0.0019</td>
<td>28.722</td>
<td>0.0998</td>
</tr>
</tbody>
</table>
Huete (1985) stated that the sensitivity of vegetation indices to the soil background is greatest with intermediate levels of vegetation cover (xxv). At very high densities there is not enough soil signal emerging from the canopy to be of significance (xxiii). In this study, although the percentage cover is high, the dominant species in the area are short, therefore the soil reflectance has a higher influence than in the case of higher vegetation species as in forested areas. This fact can explain why, in this case, soil-sensitive indices like WDVI and SAVI are better correlated with biomass than the other indices studied.

Moreover, Clevers (1989) and other authors later explained how vegetation indices such as NDVI reached a saturation level in their relation with quantitative biophysical variables like biomass. This means they become insensitive to the increase of such variables at a certain point (xxii, xxvi).

The biomass map (Figure 3) was calculated using for the Grazed-grasslands the logarithmic relationship SAVI-Biomass, and for the Mixed-herbs, the exponential relationship WDVI-Biomass was applied. The formula applied per pixel to get the final map of biomass was:

\[
B = \left[ -616.18 \times Ln(SAVI - 1086.2) \times \%\text{grasslands} \right] + \left[ 215.51 \times e^{0.3107 \times WDVI} \times \%\text{mixed - herbs} \right]
\]

This map shows the biomass corresponding to the grassland and herbaceous vegetation in the area. The lightest parts are where the biomass of grasslands and mixed-herbs plant functional types is higher and the darkest parts show were the biomass of the formers is low. Indeed, the darkest areas in the map correspond to the forested area, because the analysed plant functional types are less abundant.

In the corresponding frequency distribution (Figure 3), we can see the median occurrence corresponds to a dry biomass of 300 g/m².

### Table 4: Results summary of the exponential/logarithmic regression applied between the VI: NDVI, WDVI, SAVI and RSR and biomass. The coefficients a and b correspond to \( y = a \cdot e^{bx} \) in the case of exponential relation and \( y = a \cdot \ln(x) + b \) in the case of logarithmic relation, \( r^2 \) = determination coefficient, RMSEP= Root Mean Squared Error of Prediction and RMSECV= Root Mean Squared Error of Cross-Validation.

<table>
<thead>
<tr>
<th>VI</th>
<th>PFT</th>
<th>n</th>
<th>TYPE</th>
<th>EXP-LOG REGRESSION</th>
<th>RMSEP</th>
<th>RMSECV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>( a ) ( b ) ( r^2 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>EXP</td>
<td>0.7704 -0.0002 0.0438</td>
<td>171.82</td>
<td>341.78</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>LOG</td>
<td>0.2367 -0.729 0.62</td>
<td>163.05</td>
<td>179.62</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>LOG</td>
<td>0.058 0.392 0.1063</td>
<td>242.21</td>
<td>264.98</td>
</tr>
<tr>
<td>WDVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>EXP</td>
<td>39.019 -0.0008 0.249</td>
<td>153.11</td>
<td>215.37</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>LOG</td>
<td>22.679 -11.236 0.7046</td>
<td>150.02</td>
<td>173.77</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>LOG</td>
<td>0.398 0.7674 0.0608</td>
<td>245.55</td>
<td>283.13</td>
</tr>
<tr>
<td>SAVI</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>EXP</td>
<td>0.1237 -0.0007 0.4057</td>
<td>135.4</td>
<td>177.03</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>LOG</td>
<td>0.0477 -0.2008 0.6879</td>
<td>149.5</td>
<td>175.88</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>LOG</td>
<td>0.0068 0.0618 0.0386</td>
<td>248.36</td>
<td>287.21</td>
</tr>
<tr>
<td>RSR</td>
<td>Grazed-grasslands</td>
<td>7</td>
<td>LOG</td>
<td>0.4854 61.628 0.038</td>
<td>170.97</td>
<td>368.96</td>
</tr>
<tr>
<td></td>
<td>Mixed-herbs</td>
<td>14</td>
<td>LOG</td>
<td>24.504 -11.589 0.3965</td>
<td>198.32</td>
<td>223.74</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>21</td>
<td>LOG</td>
<td>0.7565 -0.0774 0.0939</td>
<td>248.08</td>
<td>271.13</td>
</tr>
</tbody>
</table>
CONCLUSIONS

Results shown in this study demonstrate a methodology to estimate hydraulic resistance in a river floodplain, considering that hydraulic resistance depends on PFT and vegetation biomass.

River management methods can be helped by using an accurate PFT classification map and a vegetation biomass monitoring program. Both can be assessed using imaging spectroscopy as monitoring tools.

Statistical approaches estimating biomass in a heterogeneous area are highly dependent on the presence of different plant functional types. The vegetation indices do not have comparable values between different vegetation types, and then there is no single general relationship applicable to the whole area. Consequently, in heterogeneous areas, statistical approaches have to be applied PFT-dependent.

Hydraulic resistance, as dependant on PFT can be spatially monitored. Moreover, in the cases in which a plant functional type’s map is available, the biomass distribution may give information about the density of such PFT. Higher vegetation densities (vegetation density understood as biomass/vegetation volume) correspond to a higher hydraulic resistance.

Furthermore, when floodplains belong to natural areas where the ecological value is tried to be preserved, biomass mapping is a tool to quantify the environmental impact of the river management process (e.g. side channels).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Belgian Science Policy Office for providing the HyMap dataset and Staatsbosbeheer for their permission to use the Millingerwaard as test area.
REFERENCES


