Monitoring environmental supporting conditions of a raised bog using remote sensing techniques

Saheba Bhatnagar, Bidisha Ghosh, Shane Regan, Owen Naughton, Paul Johnston, and Laurence Gill

Trinity College Dublin, Department of Civil, Structural and Environmental Eng, Dublin, Ireland

**Correspondence:** Saheba Bhatnagar (sbhatnag@tcd.ie)

Abstract. Conventional methods of monitoring wetlands and detecting changes over time can be time-consuming and costly. Inaccessibility and remoteness of many wetlands is also a limiting factor. Hence, there is a growing recognition of remote sensing techniques as a viable and cost-effective alternative to field-based ecosystem monitoring. Wetlands encompass a diverse array of habitats, for example, fens, bogs, marshes, and swamps. In this study, we concentrate on a natural wetland – Clara Bog, Co. Offaly, a raised bog situated in the Irish midlands. The aim of the study is to identify and monitor the environmental conditions of the bog using remote sensing techniques. Environmental conditions in this study refer to the vegetation composition of the bog and whether it is in an intact (peat-forming) or degraded state. It can be described using vegetation, the presence of water (soil moisture) and topography. Vegetation indices (VIs) derived from satellite data have been widely used to assess variations in properties of vegetation. This study uses mid-resolution data from Sentinel-2 MSI, Landsat 8 OLI for VI analysis. An initial study to delineate the boundary of the bog using the combination of edge detection and segmentation techniques namely, entropy filtering, canny edge detection, and graph-cut segmentation is performed. Once the bog boundary is defined, spectra of the delineated area are studied. VIs like NDVI, ARVI, SAVI, NDWI, derived using Sentinel-2 MSI and Landsat 8 OLI are analysed. A digital elevation model (DEM) was also used for better classification. All of these characteristics (features) serve as a basis for classifying the bog into broad vegetation communities (termed "ecotopes") that indicate the quality of raised bog habitat. This analysis is validated using field derived ecotopes. The results show that, by using spectral information and vegetation index clustering, an additional linkage can be established between spectral RS signatures and wetland ecotopes. Hence, the benefit of the study is in understanding ecosystem (bog) environmental conditions and in defining appropriate metrics by which changes in the conditions can be monitored.

1 Introduction

A bog is a type of wetland which primarily depends on rainfall for water and nutrients. Bogs can be categorised as blanket bog and raised bog. Raised bogs are discrete, raised, dome-shaped masses of peat occupying former lakes or shallow depressions in the landscape (Fossitt, 2000). They occur throughout the midlands of Ireland (Felicity Hayes-McCoy, 2017) and in this study, we focus on one of the largest raised bogs in Ireland, Clara Bog, Co. Offaly. Monitoring wetland structure and function typically requires recurrent site visits, which can be prohibitively labour intensive, costly and time-consuming. Monitoring is often unfeasible due to the poor accessibility, and is thus, only practical on relatively small areas (Adam et al., 2010). To acquire frequent measurements and timely information remote sensing (RS) is a cost-effective tool. Remote sensing provides invaluable information to characterize and measure the conditions of wetlands and their functioning.

The current state of art primarily focuses on mapping different types of the wetlands (Mahdavi et al., 2017). The mapping is done using different wavelengths and spectral response of the objects. Satellite imagery-derived vegetation indices can be effectively used for assessing the vegetation status of an ecosystem. Vegetation communities present within an ecosystem are defined as ecotopes. The Normalized Difference Vegetation Index (NDVI), Soil Adjusted
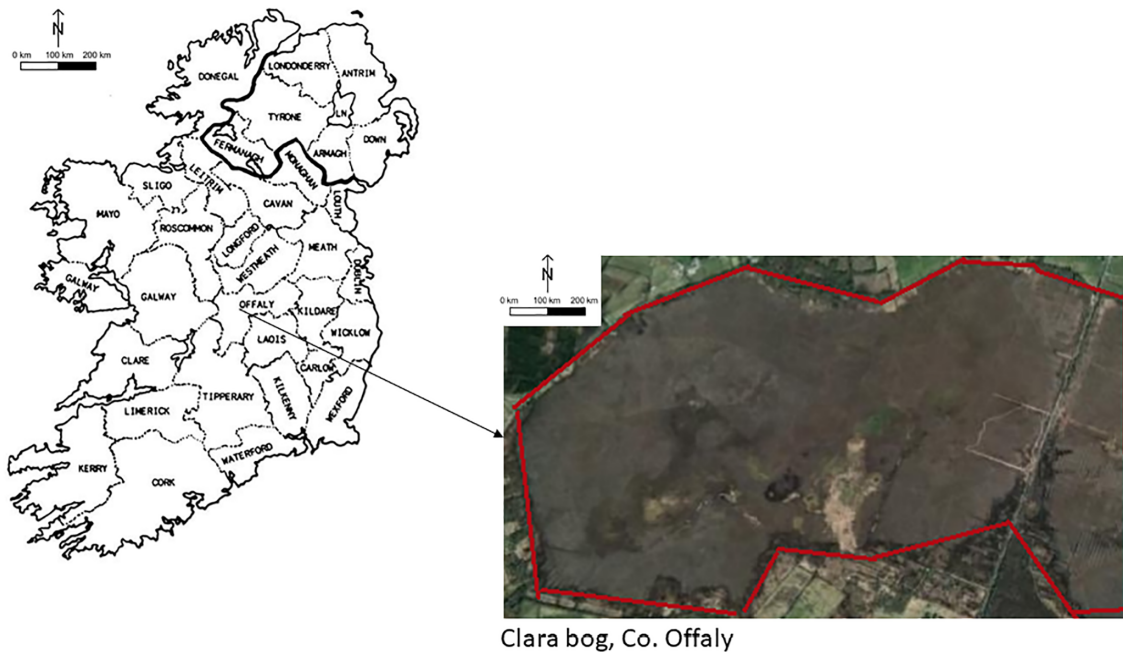
Figure 1. Clara Bog, Co. Offaly.

Vegetation Index (SAVI), Atmospherically Resistant Vegetation Index (ARVI) are the most effective vegetation indices stated in the literature (Wiegand et al., 1991). Soil moisture is not directly derivable from the optical bands. The Normalized Difference Water Index (NDWI) using near-infrared (NIR) and short-wave infrared (SWIR) bands give an indication of wetness of the surface inferred as soil moisture. These indices can be used to provide a clearer picture of vegetation and water extent in an area.

Topography plays a vital role in analysing an ecosystem. It gives an accurate idea of elevation difference present between various plant communities. Light detection and ranging (LiDAR) provides with point cloud information which can be used to deduce the topography of an area. LiDAR systems can be terrestrial, airborne, or space-borne. Normally, terrestrial systems are used for 3-D reconstruction, whereas air and space systems are utilized for remote sensing and wide territory mapping. In this study, an airborne LiDAR derived DEM is used for analysing the topographic extent of the bog.

For analysis of the RS data, various machine learning tools have proven to be useful (Lu and Weng, 2007). There are many state of the art segmentation and classification algorithms available. It is necessary to make full use of the advantages of different algorithms on the basis of multi-feature fusion, so as to achieve better segmentation effect (Yuheng and Hao, 2017). Hence, in this study a combination of segmentation algorithms is deployed. Classification accuracy is tested using various classifiers namely, SVM, Bagged Tree and Subspace KNN. SVM can be tuned using the value of op-

timization parameter, kernel used and hence, overfitting can be avoided. Bagged Tree, Subspace KNN are ensemble classifiers. The idea behind ensemble classifiers is to learn from a set of classifiers rather than a single classifier. The final result is either the average of the result from all the classifiers or is obtained using majority voting. An ensemble learner is more robust and less manpower is required for tuning the parameters. Here, a comparative study on the performance of the classifiers is carried out using freely available Landsat 8 OLI and Sentinel 2 MSI data for monitoring ecological condition and mapping ecotopes present inside the bog.

2   Materials and methodology

2.1   Study region and datasets

The site selected for this study is one of the largest bogs in Ireland, Clara Bog (Fig. 1), covering approximately 840 ha of which 443.36 ha is uncut high bog with the remaining 393.18 ha mostly cutover bog (About Clara Bog, 2018).

In the Clara bog, 9 broad categories of ecotopes have been defined namely, Sub-marginal, Sub-central, Marginal, Central, Inactive flush, Active flush, Open water, Face bank and Bog woodland.

The health of the bog is indicated by its ability to form peat. The formation of peat is depicted by central, subcentral and active flush. Ecotopes like marginal, submarginal are indicative of peatland degradation (Fernandez Valverde, 2012).

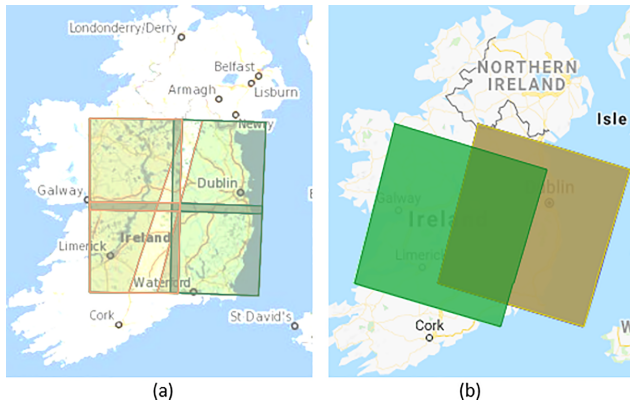For the best description of the bogs following open source data is used:

Figure 2. Footprints over Clara, Co. Offaly, Ireland **(a)** Sentinel-2, L2A, Tile_Id – 29UNV (https://scihub.copernicus.eu, last access: 5 March 2018) **(b)** Landsat-8, OLI+TIRS, Path 207, Row 023 (https://earthexplorer.usgs.gov/, last access: 19 February 2018)

1. Sentinel-2 Multispectral Instrument Level 2A (S2-MSIL2A).

   S2MSIL2A has bottom-of-atmosphere (BOA) reflectance in cartographic geometry. The granules also called tiles, are $100 \times 100\,\mathrm{km}^2$ ortho-images in UTM/WGS84 projection. The L2A-BOA product is atmospherically corrected and ready to use (Gatti and Bertolini, 2013) and is accessed from https://scihub.copernicus.eu/ (last access: 5 March 2018). The area at test lies under tile id – T29UNV. Sentinel-2 has total of 12 bands out of which 9 bands are used for analysis in this study (Band 2–8A, Band 11).

2. Landsat 8 Combined (LC08).

   Landsat 8 carries two push-broom instruments: The Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) in UTM/WGS84 projection. Atmospherically corrected, ready to use data is accessed from http://earthexplorer.usgs.gov/ (last access: 19 February 2018). The area at test lies under path 207, row 23. In this study, 12 bands are used for analysis (Band 1–11 + Pixel quality assessment, QA) (Zanter, 2016).

The latitde-longitude extent of the Clara bog is $53°19'47''$ N, $7°39'34''$ W; $53°18'55''$ N, $7°37'24''$ W. Images used are acquired by S2 and L8 on the same date (20 June 2017). Image from Sentinel-2 is resampled to 10 m (appropriate bands) and image from Landsat 8 is resampled to 30 m. The footprints for both the satellites can be seen in Fig. 2.

2.2 Methodology

The methodology used is described in the following flowchart (Fig. 3). First, it is necessary to delineate the ecosystem from the surrounding area in order to minimise
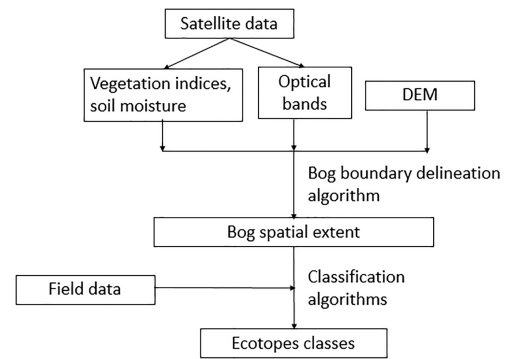

Figure 3. Methodology Flowchart.

the effect of outliers; this is achieved using segmentation algorithms. The delineated ecosystem is further divided into vegetation communities or ecotopes using a set of ensemble classifiers, namely Bagged Tree and Subspace KNN along with SVM.

2.3 Vegetation Indices and Soil Moisture

The vegetation indices used in this study are (Vegetation Indices, 2018):

1. Normalized Difference Vegetation Index

   $$NDVI = (NIR - Red) / (NIR + Red) \qquad (1)$$

   NDVI indicates the amount of vegetation, distinguishes vegetation from the soil, minimizes topographic effects, etc.

2. Soil Adjusted Vegetation Index

   $$SAVI = [(NIR - Red) / (NIR + Red + L)] \cdot (1 + L) \quad (2)$$

   where L is a soil correction factor.

3. Atmospherically Resisted Vegetation Index

   $$ARVI = (NIR - RB) / (NIR + RB) \qquad (3)$$

   where RB is a combination of the reflectance in the Blue (B) and Red (R) channels $RB = R - \gamma(B - R)$ and γ depends on the aerosol type.

4. Soil Moisture: Normalized Difference Water Index

   $$NDWI = (NIR - SWIR) / (NIR + SWIR) \qquad (4)$$

   Values range from -1, very low moisture level, to 1 very high moisture level.

Hence, there are a total of 5 extra layers i.e., NDVI, SAVI, ARVI, NDWI, and DEM along with satellite bands which are fed into the algorithm as input characteristics. Therefore, for Sentinel-2 data set there is a total of 14 layers and for Landsat 8 there is a total of 17 layers.
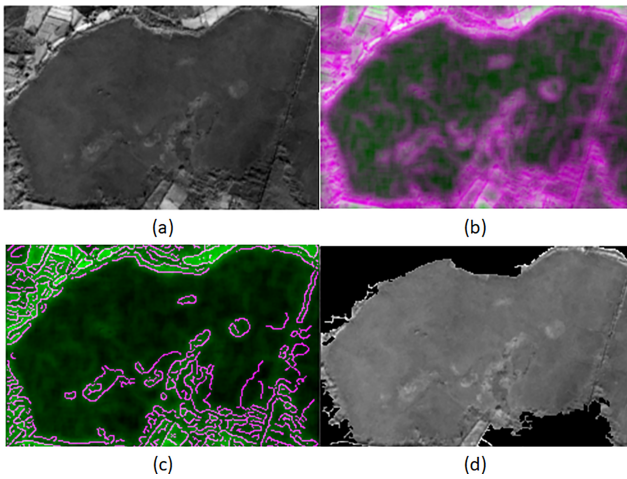
Figure 4. (a) Original Clara Image **(b)** Entropy Filter Image **(c)** Canny Edge Image **(d)** Boundary Delineated Image.

2.4 Bog Boundary Delineation

Delineation of the wetland extent was carried out using three algorithms in conjunction:

1. Entropy Filtering: measures the relative change in entropy for detection of edges. All the areas with potential objects are thus highlighted.

2. Canny Edge Detection: Initially, the intensity gradient is measured according to which background pixels are removed, i.e. only thin lines depicting edges remain. The algorithm uses 2 thresholds (upper and lower) to accept the pixels as the edge.

3. Graph Cut Segmentation: Divides every pixel into foreground (source) and background (sink) based on probability and neighbourhood information. Two major steps then follow, first is to construct the graph and the second is to produce min cut (i.e. max flow).

Since the ecosystem and surrounding areas contain distinct vegetation, the NDVI image is used as the base image (Bi) for defining the bog boundary for Sentinel data, and band 8 (panchromatic) for Landsat 8 dataset.

Algorithm 1 Bog Boundary Delineation

EBi ← Entropy_filter(Bi)
CBi ← CannyEdge_Detect(Bi, EBi)
GCi ← Graph_Cut(Bi, CBi)

Graph Cut image (GCi) is the delineated image. From GCi, spurious regions are removed on the basis of thresholding (Fig. 4).

2.5 Ecotope Identification and Classification

In this study, we have explored the applicability of pixel based, supervised classification on a raised bog. The classifier is first trained on a subset of available data, and then tested on a new location (test) to predict the classes (ecotopes) present. Here, we are doing a direct transfer of pixel-based knowledge from the training area to the testing area. The following classifiers are used for this purpose:

1. SVM – These are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis. Given a set of training data, an SVM training algorithm builds a model that assigns new data to one of the two categories, making it a non-probabilistic binary linear classifier (Cortes and Vapnik, 1995).

2. Bagged Tree – Ensemble, supervised classifier. It approaches to combine several machine learning techniques into one predictive model in order to decrease the variance hence, tuning the prediction into an expected outcome (Ensembles, 2018).

3. Subspace KNN – Similar to Bagging, subspace KNN is an ensemble method to reduce the correlation between estimators (Ho, 1998).

2.6 Validation

In this study, five major classes are considered namely Sub-marginal, Sub-central, Marginal, Central and Active flush as these are the key ecological classes indicating bog condition.

The result achieved is verified using field derived ecotope map (Fig. 5). Initially, a classification model is created using training data. For each of the classifier and data set, Model Accuracy (MA), Transfer Accuracy (TA) and Kappa Coefficient (kappa) is measured. The model accuracy (MA) is measured using 5-fold validation of training data. The transfer accuracy (TA) is the test accuracy when the model is applied to the testing data.

Ground truth is divided into:

– Case 1: Training (50 %) and Testing (50 %)

– Case 2: Training (70 %) and Testing (30 %)

3 Results and discussion

The results obtained using the aforementioned algorithms are validated using field derived ecotopes. The spatial location of the ecotopes is not weather dependent, compared to satellite imagery and corresponding vegetation indices which change with respect to weather and other environmental conditions. This study primarily highlights the condition of the raised bog during the summer season. TA signifies the scope of
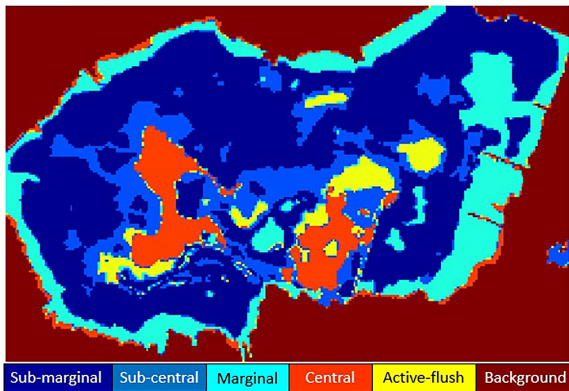
Figure 5. Ground truth with 5 ecotopes.

Table 1. Accuracies for all the cases.

	MA %	TA %	kappa
Sentinel-2 Case 1:			
SVM	73.96	32.40	0.1302
Bagged Tree	83.63	53.46	0.3424
Subspace KNN	81.65	52.95	0.3370
Sentinel-2 Case 2:			
SVM	71.71	36.36	0.2050
Bagged Tree	83.38	65.23	0.5013
Subspace KNN	62.02	36.50	0.1922
Landsat-8 Case 1:			
SVM	77.82	26.76	0.1305
Bagged Tree	87.27	53.87	0.3883
Subspace KNN	87.22	51.97	0.3635
Landsat-8 Case 2:			
SVM	72.92	29.06	0.1711
Bagged Tree	85.55	52.69	0.3539
Subspace KNN	85.00	52.71	0.3549

transferring the knowledge gained from the first half (Train) to identify ecotopes in the second half (Test).

Table 1 states the accuracy achieved using the aforementioned methodology:

3.1 Discussion

In this study, we have studied the ecological conditions of a raised bog using data from Sentinel-2, Landsat-8, their vegetation derivatives, and DEM. The key points in the result are discussed below:

1. The use of a single algorithm for boundary delineation leads to the formation of smaller, non-connected objects and hence, the bog is not delineated properly. Using the entropy filter, canny edge detection, and Graph Cut in

Table 2. Number of pixels per class – S2.

Name of Ecotope	Number of pixels
Submarginal	13 771
Subcentral	4633
Marginal	4913
Central	1192
Active Flush	2443
Background	12 248

conjunction proved to be an effective way of delineating a complex structure from a middle-resolution image.

2. SVM achieved the highest accuracy in the delineation of the submarginal ecotope, but was not viable for marginal, active flush or background. The classifier has confused between marginal, active flush and background, giving 0 % class accuracy in both datasets (Fig. 6). SVM has a major drawback of tuning the parameters. In this study the parameters were kept constant for both datasets, which increased the chances of overfitting, hence, the low accuracy.

3. Ensemble classifiers (BT, SKNN) show similar results due to the fact that the ensemble methods are generally consistent (in terms of their effect on accuracy) (Figs. 7, 8) (Maclin and Opitz, 2011).

4. The test (OA) accuracy (transfer) is also highly dependent on the number of training pixels (Table 2).

   Total pixels in the image (Sentinel 2 MSI; $160 \cdot 245) = 39\,200$. The Submarginal ecotope is most correctly classified by all the classifiers followed by Marginal. Since other ecotopes are present in much lower quantities (compared to submarginal and background), they are not identified correctly.

5. Similarity between the signatures of the classes:

   Jeffries-Matusita (JM) distance is a widely used method for feature selection in multiclass problems (Swain and Davis, 1978). The values of JM distance (Table 3) between the ROI pixels depicts low spectral separability between the ecotopes. These values were measured using 100 points from each ecotope-pair. A lower value means low separability and higher value shows higher spectral separability between ecotope pair. A higher value of JM distance is desirable for better identification of classes (Whelley et al., 2014).
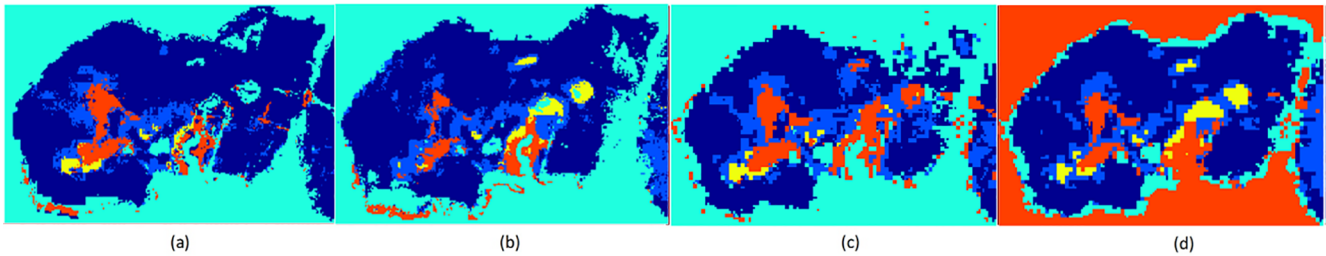
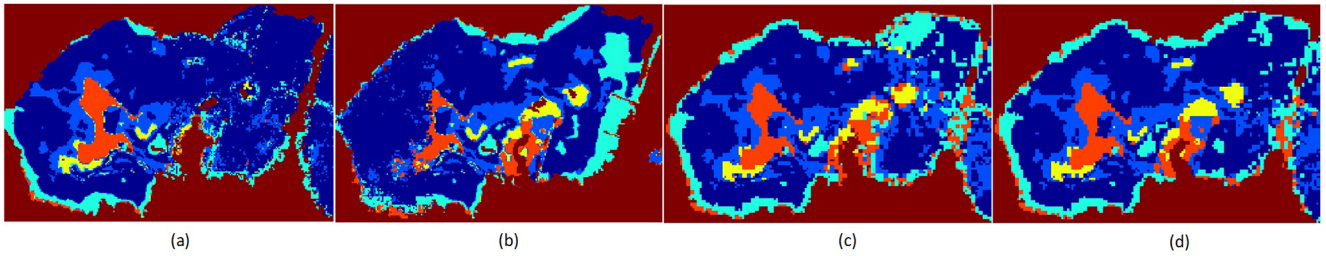**Figure 6.** SVM Classified Image: **(a)** Sentinel-2 Case 1 **(b)** Sentinel-2 Case 2 **(c)** Landsat-8 Case 1 **(d)** Landsat-8 Case 2.

Figure 7. Bagged Tree Classified Image: **(a)** Sentinel-2 Case 1 **(b)** Sentinel-2 Case 2 **(c)** Landsat-8 Case 1 **(d)** Landsat-8 Case 2.
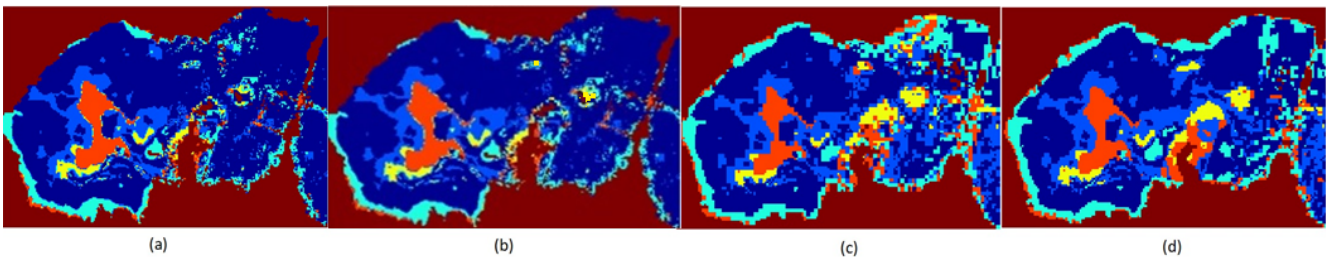

Figure 8. Subspace KNN Classified Image: **(a)** Sentinel-2 Case 1 **(b)** Sentinel-2 Case 2 **(c)** Landsat-8 Case 1 **(d)** Landsat-8 Case 2.

Table 3. JM distance between ecotopes

Ecotope Pair	JM distance
Submarginal and Subcentral	0.22
Submarginal and Marginal	0.85
Subcentral and Central	0.67
Central and Activeflush	1.48

4　Conclusions

In this study, we studied the application of mid-resolution satellite data for classification of a raised bog. The study was carried out using data from two satellites, Sentinel-2 and Landsat-8. The data used was resampled to 10 m for appropriate Sentinel-2 bands and 30 m for Landsat-8. The final classification accuracy was similar for both the satellite, unaffected by the resolution of the images. The study initially describes a competent way of boundary delineation using a series of edge detection techniques. Vegetation indices along with soil moisture and DEM information are used as features to train the classification algorithms. Bagged tree (BT) classifier proves to be the best classifier for classification of the raised bog providing better accuracy than SVM or SKNN. This is due to the nature of the ensemble classifier to reduce variance and avoid overfitting. Transfer of knowledge directly from train location to test location is not achieved effectively due to limitations in data-resolution and the amount of input training pixels. The study suggests that transfer of knowledge is effective between similar ecosystems when there is a distinct difference in the distribution of various ecotopes and pixels can be unmixed.

References

About Clara Bog: http://raisedbogs.ie/about-clara-bog/, last access: 28 March 2018.

Adam, E., Mutanga, O., and Rugege, D.: Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review, Wetl. Ecol. Manag., 18, 281–296, https://doi.org/10.1007/s11273-009-9169-z, 2010.

Cortes, C. and Vapnik, V.: Support-vector networks, Machine Learning, 20, 273–297, https://doi.org/10.1007/BF00994018, 1995.

Ensembles: available at:https://martin-thoma.com/ensembles/, last access: 8 March 2018.

Felicity Hayes-McCoy, W. J.: Dingle and its Hinterland: People, Places and Heritage, The Collins Press, 224 pp., 2017.

Fernandez Valverde, F., Crowley, W., and Wilson, S.: Raised Bog Monitoring Project 2011, in: Volume 1: Main Report, Irish Wildlife Manuals No. 62, National Parks and Wildlife Service, Department of Arts, Heritage and the Gaeltacht, Dublin, Ireland, 2012.

Fossitt, J. A.: A Guide to Habitats in Ireland, The Heritage Council, Co Wicklow, Ireland, 2000.

Gatti, A. and Bertolini, A.: Sentinel-2 products specification document, available at: https://earth.esa.int/documents/247904/685211/Sentinel-2+Products+Specification+Document (last access: 5 March 2018), 2013.

Ho, T. K.: Nearest neighbors in random subspaces, in: Advances in Pattern Recognition, edited by: Amin, A., Dori, D., Pudil, P., and Freeman, H., 640–648, Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.

Lu, D. and Weng, Q.: A survey of image classification methods and techniques for improving classification performance, Int. J. Remote Sens., 28, 823–870, https://doi.org/10.1080/01431160600746456, 2007.

Maclin, R. and Opitz, D. W.: Popular Ensemble Methods: An Empirical Study, CoRR, abs/1106.0257, http://arxiv.org/abs/1106.0257 (last access: 16 March 2018), 2011.

Mahdavi, S., Salehi, B., Granger, J., Amani, M., Brisco, B., and Huang, W.: Remote sensing for wetland classification: a comprehensive review, GISci. Remote Sens., 55, 623–658, https://doi.org/10.1080/15481603.2017.1419602, 2017.

Swain, P. H. and Davis, S. M.: Remote sensing: the quantitative approach, IEEE T. Pattern Anal. Mach. Intel., 1, 713–714, 19784.

Vegetation Indices: available at: http://web.pdx.edu/~nauna/resources/8-2012_lecture1-vegetationindicies.pdf, last access: 6 March 2018.

Wiegand, C. L., Richardson, A. J., Escobar, D. E., and Gerbermann, A. H.: Vegetation indices in crop assessments, Remote Sens. Environ., 35, 105–119, 1991.

Whelley, P. L., Glaze, L. S., Calder, E. S., and Harding, D. J.: LiDAR-Derived Surface Roughness Texture Mapping: Application to Mount St. Helens Pumice Plain Deposit Analysis, IEEE T. Geosci. Remote, 52, 426–438, https://doi.org/10.1109/TGRS.2013.2241443, 2014.

Yuheng, S. and Hao, Y.: Image Segmentation Algorithms Overview, CoRR, abs/1707.02051, http://arxiv.org/abs/1707.02051 (last access: 19 March 2018), 2017.

Zanter, K.: Landsat 8 (L8) data users handbook, Landsat Science Official Website, available at: https://landsat.usgs.gov/landsat-8-l8-data-users-handbook (last access: 19 February 2018). 2016.