DEVELOPMENT OF AN AUTOMATIC CLASSIFICATION ALGORITHM FOR DIFFERENTIATION OF VEGETATION STATUS IN SEMI-ARID ENVIRONMENTS USING MULTI-TEMPORAL HYMAP DATA

Daniel Spengler[1,2], Sabine Chabrillat[1], Paula Escribano[3], Martin Bachmann[4] and Hermann Kaufmann[1]

1. GeoForschungsZentrum (GFZ) Potsdam, Section 1.4: Remote Sensing, Telegrafenberg, 14473 Potsdam, Germany
2. Email: daniel@gfz-potsdam.de
3. University of Almeria, Department of Soil science, Almeria 04120, Spain
4. German Aerospace Centre (DLR-DFD) Oberpfaffenhofen, 82234 Weßling, Germany

ABSTRACT

Changes in the vegetation status are important indicators for land degradation processes in semi-arid ecosystems. The multi-temporal analysis of hyperspectral remote sensing data builds the basis for understanding vegetation specific changes. The overall goal of this study is the development of an automatic classification algorithm based on the decision tree system. Green vegetation, dry senescent vegetation, and soils are successively classified through spectral absorption features in the VNIR-SWIR range, with different vitality degree both for green and dry vegetation. The decision tree differs between pure dry vegetation pixels and soil-influenced pixels mixed with dry vegetation. The result is a classification map including different vitality states of the vegetation. By aggregation of all vegetation classes, the total fraction of vegetation cover can be calculated. The results show that vegetation cover is estimated with a high precision and independently from phenological situation.

Hyperspectral airborne HyMap data of the Natural Park Cabo de Gata Níjar are used for the development of the algorithm. Human activity in the parks preserved semi-arid Mediterranean ecosystem leads to several changes in landscaper patterns and vegetation cover. Since the past ~40 years increases tall grass steppes and arid garrigues increased while tall arid scrublands and crop fields decreased. Since the Natural Park becomes under protection in 1987, these former typical vegetation recovered. Presently, major risk is observed from a potential effect of exponential tourism and agricultural growth.

INTRODUCTION

Arid and semi-arid lands cover approximately one third of the continental surface of the Earth. Around 10 to 20% of these arid ecosystems are affected by desertification. Nearly one-third of the world population lives in this region where desertification is one of the most serious environmental problems (i) For the detection of the desertification potential hazard and monitoring of surface changes the expansion of some biophysically indicators can be used. Remote sensing data and in particular hyperspectral remote sensing can dramatically increase the accuracy of indicators for land degradation and desertification.

The work is framed in the project INCAMAR of the University Almeria that integrates multispectral and hyperspectral information for the selection and characterization of change indicators in arid ecosystems. The indicators will be the basis for a monitoring program in the Natural Park Cabo de Gata-Níjar. The vegetation status is one of the most important indicators. The vegetation cover and vegetation vitality are key variables that can be monitored by hyperspectral remote sensing. The multi-temporal analysis of such data builds the basis for understanding vegetation specific changes. In our study we focus on the development of a classification algorithm that determines

the vegetation cover independent from phenological status of the vegetation and in addition quanti-fies the vegetation vitality.

STUDY AREA

General

The study area is situated in the Cabo de Gata–Níjar Natural Park located in the Almeria province, SE Spain (Figure 1). It represents a fragile semi-arid ecosystem with a variety of characteristic vegetation types that is highly sensitive to land degradation processes. This marine–terrestrial Park was declared at Spain's first Natural Park in 1987. It covers 37570 ha of land and 12126 ha of marine coastline with altitudes ranging between 562 m high and 60 m under sea level. The climate is Mediterranean semi-arid with average annual rainfall around 250 mm. It drops predominantly in the months from October to March (ii). The mean annual temperature is 17,8°C. The park cuts into two big features: the Cabo de Gata massif (Sierra del Cabo de Gata), and the coastal plains (Ba-hia de Almeria). The Sierra del Cabo de Gata is a section of the Betic Cordilleras, that are the big-gest mountainous massif with volcanic origin in Europe. This volcanic rock of the calco-alcaline type formed during the Tertiary Era in four periods of hydrothermal activity between ca. 15 and 8 million years ago. (iii) Presently the park is partly used for farming. About 3500 people are living in the park.

The volcanic origin, semi-arid climate and the marine terrestrial character produce a region with high ecological importance. The park represents one of the few semi-deserts as the case may be veldt regions in Europe. The two main soil types are calcareous or volcanic. The vegetal landscape is composed of various types of scrubland, ranging from tall arid brushes with species like the European Fan Palm (Chamaerops humilis L. -considered the only native palm in the European continent), phlomix, to tall grass steppes of Stipa tenacissima L., arid garrigues (genista, ar-temisia), arid scrubs (thymus), opuntia (cactus), and crop fields. The maximum vitality of the vege-tation occurs in spring while in the course of the years it changes differently depending on growing location, e.g. distance from the sea, slope exposition.
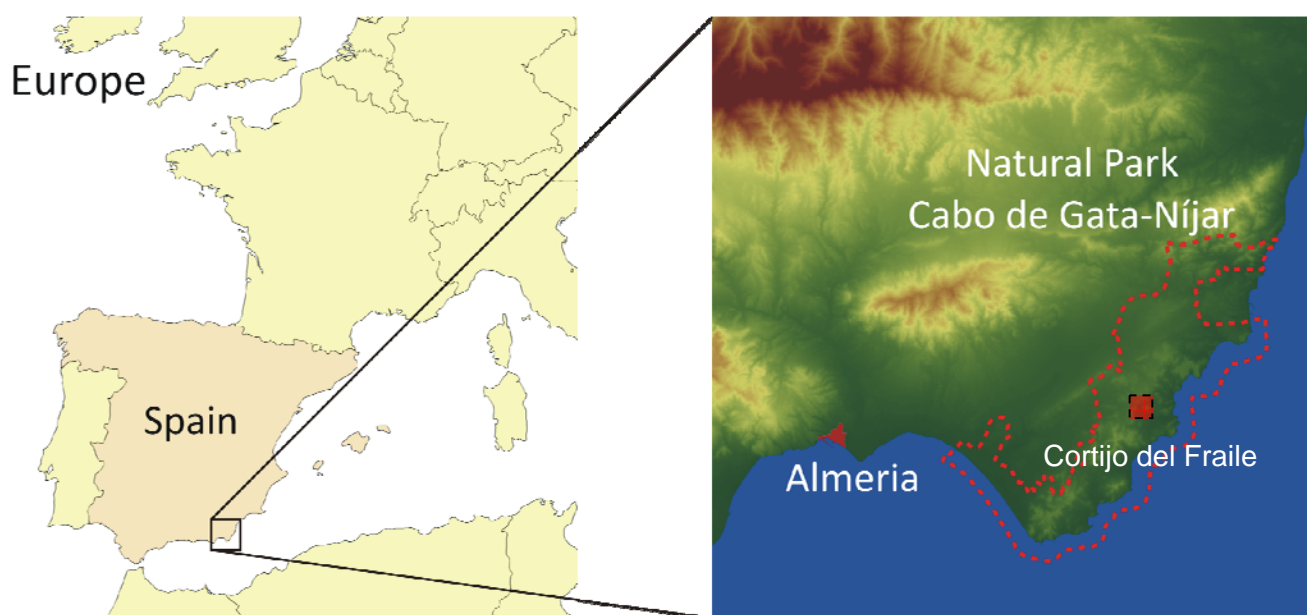

Figure 1: Location of the Cabo de Gata-Níjar Natural Park
The square locates the area of interest, covered by HyMap data in 2004

Test site

In the Natural Park one test site was chosen, that includes many of the natural conditions of the study site. The test site Cortijo del Fraile is situated in the agricultural part of the park. In this area three types of land use are found. Wide parts are used as farmland for vegetable production (Figure 2a). Through irrigation these areas have very good growing conditions. A second complex is build up from natural vegetation that is not man-made influenced. Here you find the typically Stipa – Chamaerops plant society. These areas are mainly at the upper parts of the hills (Figure 2b). The lower parts of the hills were former used as farmland or pastures. The clear difference with the natural surfaces can be seen as it results in the nearly totally lack of Chamaerops humilis in the human-influenced areas. Also these areas have a sparse vegetation cover.
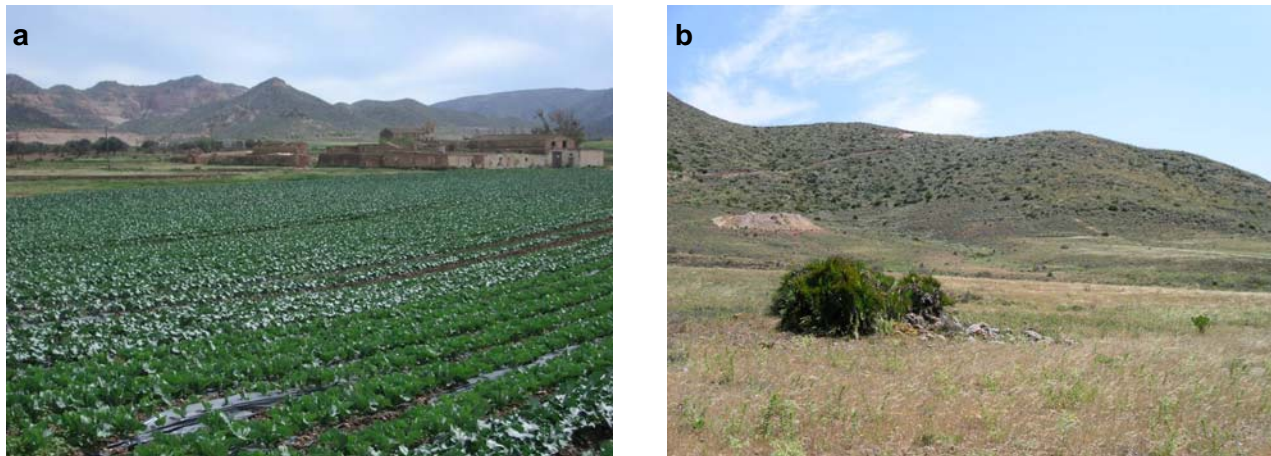

Figure 2: a) Landuse at the test site Cortijo del Fraile; b) natural vegetation uphill, grassland down-hill (former agricultural use)

DATA

Hyperspectral image data

Multi-temporal HyMap hyperspectral images were acquired in spring and summer 2004 (May 18th 2004 and August 13th 2004), during the DLR / HyVista HyEurope campaigns. Between these two acquisition times the vegetation totally changes from vital to dry status. This enables the analysis of the phenological changes of the vegetation cover and vitality. The HyMap sensor is an airborne imaging spectrometer covering the spectral range 450-2500 nm with an average bandwidth of 16nm in 128 bands. Additional Information are shown in Table 1. The radiometric and geometric correction was done by the DLR with ATCOR4 and ORTHO.

Table 1: 2004 HyMap airborne campaigns inn Cabo de Gata-Níjar

Acquisition date	Time	Flight height	Pixel size	Flight lines	Spectral coverage	Flight direction
18 May 2004	9:45	2242	5 m	5 E-W	0.45-2.45 μm	E – W
13 August 2004	9:15	2315	5 m	5 E-W	0.45-2.45 μm	E – W

Field Spectra

During the project field spectral data were acquired. It is clearly visible that the spectral signature of vegetation changes from vital to dry status as shown for the example of Stipa tenacissima in Figure 3. The dry vegetation spectra do not show the green peak and the red edge. Therefore appears a strong absorption feature around 2100nm resulting from cellulose and lignin content.
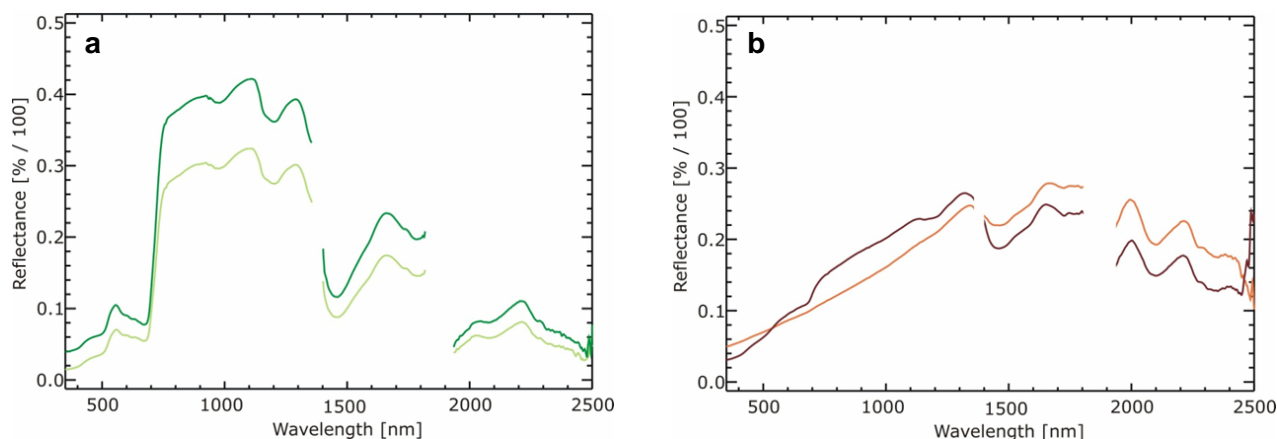
Figure 3: Spectral reflectance of Stipa tenacissima in a) vital (May 2004) and b) dry (July 2003) status

METHODS

Decision Tree Classification

The Decision Tree Classification has high potentials for the acquisition of land cover. It allows to combine different datasets and algorithms for the classification. Unlike conventional classification methods like the Maximum Likelihood Classification or the Minimum Distance Operation the Decision Tree Classifier has not been such extensively tested (iv).

To differentiate between three main classes vital vegetation, dry vegetation and non vegetated materials, the use of a decision tree classifier is very suggestive. Through this the implementation of different classification parameters is possible. The combination of different operators is necessary because an optimal detection of the vegetation status call for the inclusion of different wavelength.

The dataset will be classified stepwise in repeated decision steps. It allows that different spectral attributes can be observed and the class separation occurs with defined limits.

Classification Algorithm

The green vegetation is classified by the Normalized Difference Vegetation Index (NDVI). It minimizes the topographic effects (v), is independent from the surface character and offers a high sensibility to photosynthetic active vegetation. The scale of the NDVI varies from -1.0 up to 1.0 (vi, vii). The NDVI is less accurate in regions with low vegetation cover, because of the soil signal that influenced the remote sensing signal directly. To solve this difficulty, there are improved variations of the NDVI like the Soil-Adjusted Vegetation Index (SAVI) or the Optimized Soil Adjusted Vegetation Index (OSAVI) that include a soil factor (viii). If the vegetation cover is constant these indexes give better results than the NDVI (ix; x ; xi). The vegetation cover in the study area is very variable, so that of a soil factor is not meaningful and the NDVI is used instead as best classifier for photosynthetically active vegetation.

The dry vegetation has three characteristic absorption features with maximum values around 1730 nm, 2100 nm and 2300 nm. These absorption features are caused by the overlapping of the absorption from the cell components lignin, starch and sugar (1730nm), cellulose and starch (2100 nm), cellulose, starch and sugar (2300nm) (xii, xiii; xiv). Thereby the three absorption features differ very strongly. Only the 2100nm dry vegetation band is clearly visible in the HyMap data, and cannot be confused with e.g. soil absorption features that could result from the presence of carbonate soils in the area (2300nm). The Cellulose Absorption Index (CAI) (xv), calculates the relative absorption depth at the 2100 nm band. For the classification of dry vegetation a modified form of this CAI is used. Through normalization (nCAI) the index has a range of scale between -1.0 and 1.0.
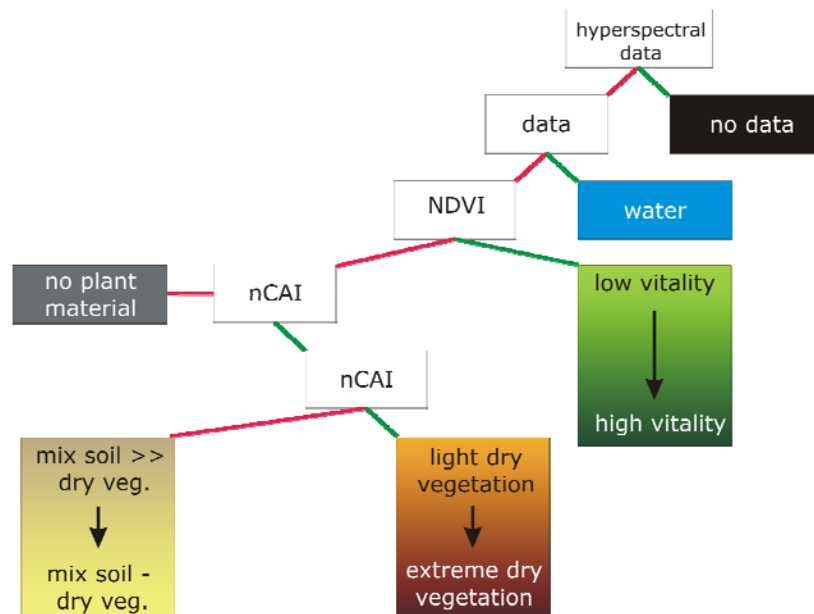
Figure 4: Structure of the decision tree classification

The decision tree (see Figure 4) is built up so that in a first decision point the pixels of the dataset are classified in data and no data areas. Hereon follows the classification of water. Now comes the Decision between dry and green vegetation based on NDVI values. If NDVI is greater than 0.3 the pixels are classified again by their NDVI as green vegetation, from light to very high vitality. If the NDVI is lower than 0.3 all the pixels go into the next decision note that differentiates the non-plant material from the dry vegetation influenced pixels. If there is no absorption around the 2100nm band, no cellulose is detected, and the pixels are classified as non-plant material.

RESULTS

The classification result of the test site Cortijo del Fraile shows the mixture of natural vegetation development and anthropogenic changes. The areas covered by very green vegetation (NDVI > 0,6) in the May 2004 dataset (Figure 5a) are agricultural fields with vegetables. It is visible that wide ranges of the agricultural areas are already harvested. These areas are classified as non-plant material. The natural vegetation is always drier than in the agricultural fields and at some slopes is classified in different dry vegetation classes. This is due to a larger distance to the sea that is associated with lower air moisture. Strong differences in the vitality depending on the slopes are visible. The northern exposed hillsides have a much higher vitality than the southern exposed slopes, resulting from the lower sun irradiation on the northern parts. High vitality of the natural vegetation is especially visible near the water lines. The small valleys are better water supplied and so have a higher concentration of Chamaerops humilis. In comparison to the classification of the August data (Figure 5b) the fresh vegetation in these areas is strongly reduced. But in some parts the presence of still green Chamaerops humilis is visible. Nearly all pixels classified as vital vegetation in August result from the existence of the palm tree. Generally the vegetation in August becomes extremely dry. A difference of vitality due to slope expositions is then not anymore visible. The pixels formerly classified as vital vegetation are now mainly classified as dry vegetation or a mix of dry vegetation and soil. It is noticeable that some areas that are classified as non-plant material in May are classified in August as a mixture between soil and dry vegetation. This could be a result of the lower accuracy of the NDVI in areas with sparse vegetation cover. Here exist a potential of optimization for the classification tree.
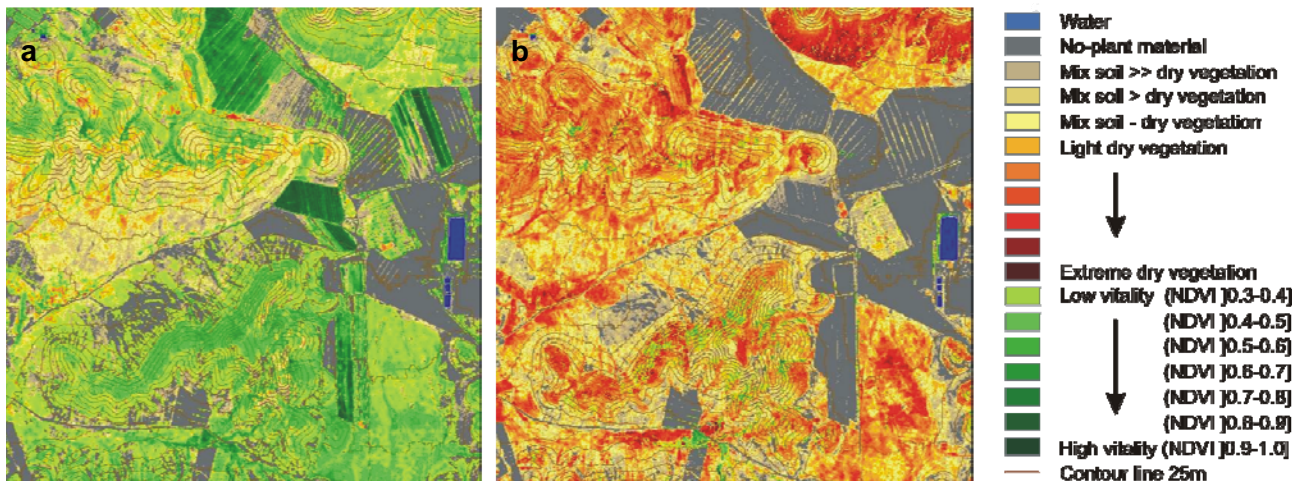
Figure 5: Classification result for the test site Cortijo del Fraile based on the HyMap datasets from a) May 18th 2004 and b) August 13th 2004.

CONCLUSIONS

An automatic algorithm classifying different degree of green vegetation, dry vegetation, and soil has been developed, and works well in the study area. The change of the vegetation vitality fits well with multitemporal field observations. Through the use of the classification tree the vegetation cover can be detected independently of the season with a high accuracy. In the regions of very sparse vegetation cover, some algorithm limitations are observed. A possible reason for that is the high dependence of the NDVI with the soil signal. Here lies the highest potential of optimization for the classification tree.

REFERENCES

i   Ustin, S.L., S. Jacquemoud, A. Palacios-Orueta, L.Li & M.L. Whiting, 2005: Remote sensing based assessments of biophysical indicators for land degradation and desertification for atmospheric and space research. In: RGLDD e-Proceeding, UN and European Commission.

ii   Aranda, V. & C. Oyonarte, 2005. Effect of vegetation with different evolution degree on soil organic matter in semi-arid environment (Cabo de Gata-Nijar Natural Park, SE Spain). Journal of Arid Environment, 62: 631-647.

iii   Chabrillat, S., H. Kaufmann, A. Palacios-Orueta, P. Escribano & A. Mueller, 2004. Development of land degradation spectral indices in a semiarid Mediterranean ecosystem. In: Proceedings of SPIE, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IV - Vol. 5574, 235-243.

iv   Friedl, M.A. & C.E. Brodley,1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment, 61: 399-409.

v   Holben, B.N. & C.O. Justice, 1981. An examination of spectral band rationing to reduce the topographic effect on remotely sensed data. International Journal of Remote Sensing. 2: 115-121.

vi   Myneni, R.B., B.D. Ganapol & G. Asrar, 1992. Remote sensing of vegetation canopy photosynthetic and stomatal conductance efficiencies. Remote Sensing of Environment, 42: 217-238.

vii   Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8: 127-150.

viii   Gilabert, M.A., J. Gonzales-Piqueras, F.J. Garcia-Haro & J. Melia, 2002. A generalized soil-adjusted vegetation index. Remote Sensing of Environment, 82: 303-310.

ix   Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25: 295-309.

x   Baret, F. & G. Guyot, 1991. Potentials and limits for vegetation indices for LAI and PAR assessment. Remote Sensing of Environment, 35: 161-173.

xi   Rondeaux, G., M. Steven & F. Baret, 1996. Optimization of soil-adjusted vegetation indices. Remote Sensing of Environment, 55: 95-107.

xii   Gao, B.-C. & A.F.H. Goetz, 1994. Extraction of dry leaf spectral features from reflectance spectra of green vegetation. Remote Sensing of Environment, 47: 369-374.

xiii   Curran, P.J.,1989. Remote sensing of foliar chemistry. Remote Sensing of Environment, 26: 271-278.

xiv   Kokaly, R.F. & R.N. Clark, 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. Remote Sensing of Environment, 67: 267-287.

xv   Daughtry, C.S.T., E.R. Hunt & J.E. McMurtrey III, 2004. Assessing crop residue cover using shortwave infrared reflectance. Remote Sensing of the Environment, 90:126-134.