SPATIAL AND SPECTRAL SEPARABILITY OF GRASSLANDS IN THE INNER TURKU ARCHIPELAGO USING LANDSAT THEMATIC MAPPER SATELLITE IMAGERY

Simon Amelinckx

Master's thesis
May 2010

Master's Degree Programme in Environmental Sciences
Specialisation: Geoinformatics in Environmental Research

DEPARTMENT OF GEOGRAPHY
UNIVERSITY OF TURKU
FINLAND

UNIVERSITY OF TURKU
Department of Geography
Faculty of Mathematics and Natural Sciences

AMELINCKX, SIMON: SPATIAL AND SPECTRAL SEPARABILITY OF GRASSLANDS IN THE INNER TURKU ARCHIPELAGO USING LANDSAT THEMATIC MAPPER SATELLITE IMAGERY

Master's Thesis, 104 pp., 1 annex.
Environmental Sciences – Geoinformatics in Environmental Research
May 2010

The present study investigates the spatial and spectral discrimination potential for grassland patches in the inner Turku Archipelago using Landsat Thematic Mapper satellite imagery. The spatial discrimination potential was computed through overlay analysis using official grassland parcel data and a hypothetical 30 m resolution satellite image capturing the site. It found that Landsat TM imagery's ability to retrieve pure or near-pure pixels (90% purity or more) from grassland patches smaller than 1 hectare was limited to 13% success, compared to 52% success when upscaling the resolution to 10 x 10 m pixel size. Additionally, the perimeter/area patch metric is proposed as a predictor for the suitability of the spatial resolution of input imagery. Regression analysis showed that there is a strong negative correlation between a patch's perimeter/area ratio and its pure pixel potential.

The study goes on to characterise the spectral response and discrimination potential for the five main grassland types occurring in the study area: recreational grassland, traditional pasture, modern pasture, fodder production grassland and overgrown grassland. This was done through the construction of spectral response curves, a coincident spectral plot and a contingency matrix as well as by calculating the transformed divergence for the spectral signatures, all based on training samples from the TM imagery. Substantial differences in spectral discrimination potential between imagery from the beginning of the growing season and the middle of summer were found. This is because the spectral responses for these five grassland types converge as the peak of the growing season draws nearer. Recreational grassland shows a consistent discrimination advantage over other grassland types, whereas modern pasture is most easily confused. Traditional pasture land, perhaps the most biologically valuable grassland type, can be spectrally discriminated from other grassland types with satisfactory success rates provided early growing season imagery is used.

Keywords: spatial separability, spectral separability, grassland, Landsat TM, remote sensing, Turku Archipelago

Index

1 Introduction

The adoption of modern agriculture has resulted in intensification of agricultural land use, particularly in favourable regions (Stoate *et al*. 2001). Agricultural intensification results in a declining area of non-crop agricultural land, such as grasslands (Matson *et al*. 1997). Grasslands are considered as key habitats of European agricultural environments. They are typically characterised by high species richness and a great number of specialised species (Kivinen 2007). However, studies on species-rich grassland have been limited to a few individual study sites only and no efforts have been made for collecting relevant information on a landscape level (Toivonen and Luoto 2003).

In Finland, the amount of grassland habitats has decreased dramatically during the past hundred years, due to the rapid decline in small-scale dairy farming and associated grazing (Pykälä 2001, Tiainen *et al*. 2004). The decrease in area of semi-natural grassland alone has retreated to 1% of the area covered in the beginning of the 20th century (Vainio *et al*. 2001). Natural grasslands do occur, but generally only in small patches along shorelines, in mires and rocks rocky areas. The majority of Finnish grasslands has resulted from agricultural management (Marttila *et al*. 1999).

Although historically rich in meadows and pasturelands, the Turku Archipelago region underwent a series of large-scale land cover changes after traditional animal husbandry dramatically decreased in the 1950s (Kotiluoto 1998). Subject to the process of agricultural intensification, such practices were no longer considered economically profitable (Lindgren 2000). Despite severe declines in grassland abundance, the Turku Archipelago is still considered among Finland's main concentrations of agricultural biodiversity (Kivinen 2007). This calls for careful habitat monitoring with particular attention to further degradation of the remaining grassland patches.

Developments in remote sensing of the Earth's surface offer an ever-improving toolset for studying biogeographical processes. For studying large areas, satellite images have proven especially useful (Gallego 2005). Their ability to provide consistent and repeatable measurements at a spatial scale appropriate for many processes renders them highly suitable for monitoring changes (Kennedy *et al*. 2007). Using satellite imagery, changes in

composition of grassland have been successfully discriminated (*e.g.* Peterson *et al*. 2002, Price *et al*. 2002, Guo *et al*. 2003, Miatkowski 2004, Psomas *et al*. 2005).

The Thematic Mapper sensor carried by the Landsat 5 satellite is characterised by a medium-quality spatial resolution and low spectral resolution (Rock *et al*. 1993). Additionally, with a 16-day repeat cycle, Landsat 5 cannot readily provide frequent and global coverage as is the case for more recently launched instruments. However, the thematic sensor is praised for its relatively broad spectral coverage (Rock *et al*. 1993). Moreover, the Landsat 5 satellite is the only fully functional Landsat platform currently in orbit (NASA 2009a). Due to its early launch (1.3.1984) and continued operation, its data constitute the longest record of the Earth's continental surfaces as seen from space (USGS 2009a).

In order to successfully discriminate grassland patches, *spatial* separability of these patches is a first requisite. Using medium-resolution satellite imagery, this may not be the case for the Turku Archipelago, which is characterised by small-scale variation of land cover across the landscape (Metsähallitus 2010). In fact, Thematic Mapper imagery has previously been found too coarse for accurately mapping smaller patches of semi-natural grassland in southern Finland (Luoto 2002, Toivonen & Luoto 2003).

A second requisite for satellite technology to be effective at discriminating valuable grassland occurrences in the Turku Archipelago is their *spectral* separability. For land cover classes to be separable, the variance in spectral reflectance between the investigated features needs to be greater than that within investigated features (Schmidt & Skidmore 2001). Although visibly different at ground level, no information is available on the variance in spectral characteristics of different grassland types occurring in the region.

Little research has been done on the ability of Landsat Thematic Mapper imagery to discriminate between different types of grassland occurring in the boreal zone. Moreover, few studies have included landscape matrix characteristics such as patch size and shape complexity in their assessment. In terms of selecting scene capture dates optimal for the discrimination of grassland types, a number of recommendations have been previously formulated (Price *et al*. 2002, Guo *et al*. 2003) but whether they equally apply to the boreal zone has not been confirmed. Especially for large regions such as the Turku Archipelago where close monitoring of the evolution of environmental conditions is crucial if the loss

of biodiversity is to be halted, a long and consistent record of remote sensing imagery may prove useful. Prior to the use of such imagery however, its suitability needs to be ascertained.

The aim of the present research was to assess the separability of grassland occurrences in the inner Turku Archipelago, based on Landsat Thematic Mapper satellite imagery. This included both an assessment of the *spatial* separability of typical grassland patches in the archipelago landscape and the characterisation of their *spectral* response patterns so as to evaluate the potential for a spectral differentiation between the types of grassland present in the region.

In order to attain the aim of the study, three main objectives were defined. The first objective deals with the spatial aspect of the ability of Thematic Mapper 5 imagery to allow for grassland patch shapes to be correctly extracted. With a spatial resolution limited to 30 m, typical archipelago patch dimensions may not be compatible with such coarse satellite imaging. Especially the smallest landscape entities are sensitive to the spatial precision of remote sensing material. By isolating these *vulnerable* grassland patches and subjecting them to a hypothetical 30 x 30 m Thematic Mapper grid overlay, the spatial discrimination of the imagery can be tested. If 30 m resolution imagery allows the retrieval of the smallest patches to a satisfactory level, then it is also suitable for larger patches. However, with some grassland patches being the same size as a single Landsat TM 30 x 30 m grid cell or smaller, a successful spatial discrimination of *all* patches seems highly unlikely. This leads to the question as to whether an increase in spatial resolution of the imagery would also lead to a similar increase in spatial discrimination success. Assuming that Landsat TM 5 satellite imagery does not allow the retrieval of the smaller grassland patches in the archipelago, how much more successful is a more sensitive sensor like SPOT?

The second objective of this study was to determine the ideal time of year for the Thematic Mapper sensor to capture the phenology of the Turku Archipelago if it is to result in the most accurate discrimination between different grassland types. Due to changes in the phenology of plant communities throughout the growing season and beyond, different types of land cover show a converging or diverging spectral response to irradiance over time. Plant phenology depends directly on species composition and growing conditions and is therefore unique for each habitat type or management regime. Since the existence of

different grassland types also pertains to variation in growing conditions – be it because of natural processes or anthropogenic action – some level of variance in phenological patterns can be assumed. The challenge is to define the moment in time when the type-specific variance in grassland phenology is most pronounced, so that different types of grassland can be most easily distinguished.

The third objective was to analyse the spectral responses of different grassland types and their behaviour from maximal to minimal phenological divergence. In this study, grassland is defined as a patch of land characterised by the predominance of gramineous species (i.e. true grasses). The species composition of the studied grassland patches is therefore similar by definition, although differences in management regime or growing conditions may be the cause of slight variations. This implies that there is a risk that the spectral responses of these different grassland patches may not be sufficiently divergent to allow type discrimination based on their spectral characteristics only. As part of the third objective, this study aims to address the question as to how dissimilar or similar the spectral responses of grassland types are at a selection of key phenological moments. Which bands show more pronounced differences than other? Are these differences sufficient to successfully sort grasslands by type using an automated classification procedure?

Ultimately, the present study was intended to provide an answer to the question as to whether Landsat TM imagery is suitable – both in terms of its spatial and spectral resolution – for the discrimination of valuable grasslands from other grasslands that occur in the inner Turku Archipelago. Although Landat TM satellite imagery has been successfully used for such analysis elsewhere, it is unclear whether this can be extended to other climates and other geographical conditions. Would other types of imagery produce better results? In other words, is Landsat Thematic Mapper 5 up for the job?

2 Theoretical background

2.1 Definition of grassland

Many definitions for grassland have been put forward. Given the multitude of grassland types, definitions are typically type-specific. An important distinction can be observed between (semi-)natural grasslands and non-natural grasslands, but important variations in definition also exist within habitat types. The necessary level of generalisation and the aim of the study for which a grassland type is defined often lead to significantly different if not contradictory definitions.

The European Topic Centre on Land Use and Spatial Information (2006) provides an excessive definition for natural grasslands on a European-wide scale. According to this definition, natural grassland is a type of low productivity environment, often situated in areas of rough and uneven ground. It frequently includes rocky areas, briars and heathland. More specifically, natural grasslands are areas where herbaceous vegetation (typified by a maximum height of 150 cm and predominance of gramineous species) covers at least 75% of the surface covered by vegetation. This includes the following:

- saline grasslands grown on temporally wet areas of saline soils,
- humid meadows where sedges, rushes, thistles, nettles cover more than 25% of the parcel,
- natural grasslands with trees and shrubs if they do not cover more than 25% of the surface to be considered,
- high-productive Alpine grasslands far from houses, crops and farming activities,
- herbaceous military training areas,
- grasslands which can be grazed, never sown and not otherwise managed by way of application of fertilizers, pesticides, drainage or reseeding except by burning,
- grasslands with a yearly productivity less than 1 500 units of fodder/ha,
- herbaceous grass covered or composed of non-palatable gramineous species such as *Molinia* spp. and *Brachypodium* spp.,
- derelicted natural grass land where ligneous vegetation covers less than 75% of the area,
- grasslands found on calcareous soils with a high proportion of calcicole species of limestone, chalk Machair or Karst,

- grasslands dotted with bare rock areas which represent less than 25% of the surface.

Grey dunes, swampy grassland and fallow lands are not included.

As is the case for the above definition, minimum or maximum percent patch coverage of grassland vegetation and other land cover types respectively are often included. This is because grass or grassland vegetation tend to continue across patch boundaries along a gradient of decreasing relative abundance, making it difficult to delineate patches. Especially tree crown cover overhead receives attention. Pykälä *et al*. (2005) for example only considered a landscape unit to be a patch of grassland if more than 80% was covered by grassland vegetation while total tree cover had to be less than 20%.

Often, definitions also include specifications regarding the purpose, management practices and botanical characteristics. Luoto *et al*. (2002) for instance, defined semi-natural grassland being grasslands that consists of meadows and traditional semi-natural pastures where modern fertilizers are applied, pre-dominated mainly by indigenous herbs, especially grasses (*Poa* sp. L and *Festuca* sp. L) and sedges (*Carex* sp. L).

Studying Swedish landscapes, Skånes (1996) placed especially strong emphasis on the management factor, stating that grasslands are the highest hierarchical level of all land cover types which are currently influenced by, or still show evidence of, grazing or mowing, characterised by light-demanding herbaceous vegetation. This approach stresses the importance of some sort of growth limitation – either of herbivorous or directly anthropogenic nature – as a requisite for the existence of boreal grasslands. In fact, in Finland natural non-managed grasslands only exist in small quantities in the uplifting coastal zone, adjacent to flooding rivers or lakeshores and on some mountains, cliffs, mires and some larger areas on the fells in Lapland (Jutila 1999, Marttila *et al*. 1999).

Traditionally, northern European grasslands constituted a dynamic transition zone between the arable fields and the outland forests. However, such zones have considerably diminished in importance with arable fields often stretching right up to the forest edge in the present-day rural landscape (Käyhkö and Skånes 2006). With the arrival of fertilisers and more productive grass varieties as well as the need for constructing grasslands for recreational and logistical purposes, grasslands diversified immensely, complicating the application of a universal definition.

2.2 Finnish grassland types

Just as with grasslands worldwide, Fennoscandian grasslands are classified as meadows and pastures according to management and as dry, mesic and wet vegetation types or high or low herb or grass-dominated vegetation types according to the ecological factors such as soil moisture and nutrition. Dry to mesic vegetation types on acidic soil are specific to and common in the Nordic countries (Ihse 1995, Jutila 1999). Annex 1 of the EU Habitats Directive groups valuable natural grasslands such as those in the Rekijoki valley (Pykälä 2003) into a vegetation class called *Fennoscandian lowland species-rich dry to mesic grasslands* and considers them habitat types of community interest whose conservation requires the designation of special areas of conservation (European Council 1992).

In terms of grasslands occurring in Finland, several grassland classifications have been proposed. Pykälä (2003) distinguished between wet grasslands, fen-meadows, wooded meadows, seashores, dry grasslands and mesic grasslands whereas Marttila *et al*. (1999) embraced a differentiation between dry rocky grassland, dry grassland, mesic grassland, riparian grassland and other grasslands (sites characterised by high vegetation diversity, not originating from traditional agricultural practices) when studying semi-natural grasslands in mainland Karelia. In an attempt to discriminate mesic semi-natural grassland in southern Finland, Pykälä later (2004) defined three classes: old (continuously cattle grazed), new (cattle grazing restarted 3-8 years ago) and abandoned pastures (grazing terminated > 10 years ago).

Some have further developed a species-based identification of grassland types. Indicator species for dry rocky grassland for instance include *Sedum* spp., *Potentilla argentea* and *Festuca ovina*. Dry, non-rocky grassland on the other hand typically hosts species such as *Galium verum*, *Campanula rotundifolia* and *Pilosella officinarum* whereas for mesic grassland this is *Ranunculus acris*, *Potentilla erecta*, *Galium boreale*, *Agrostis capillaries* and *Deschampsia cespitosa*. The species composition may also function as an indicator for management practices. Abundance of *Epilobium angustifolium* and *Filipendula ulmaria*, often in combination with shrubs and tree sapling occurrence can be the consequence of site encroachment and low intensity management while *Aegopodium podagraria, Anthriscus sylvestris*, *Urtica dioica*, *Taraxacum* spp. and *Poa pratensis* indicate nutrient enrichment (Marttila *et al*. 1999). Timothy-grass (*Phleum pratense*) points to fodder production practices since it is the most widely grown sown grass species for silage and hay production in the Nordic countries (Höglind *et al*. 2001).

2.3 Grassland occurrence in Finland

As mentioned earlier, most meadows of the boreal zone resulted from agricultural management as natural grasslands only exist in a limited set of environments (Jutila 1999, Marttila *et al*. 1999). In mainland Finland, various types of grassland are typically abundant in regions with versatile agriculture and varied topography (Kivinen 2007).

The most valuable grasslands are traditional rural biotopes. They have an exceedingly high cover in the Åland islands, where traditional agricultural practices have been carried out until recently (Schulman *et al*. 2005). Generally, the Åland islands represent one of the main concentrations of agricultural biodiversity in Finland, together with some southwestern and southern (especially coastal) regions (Kivinen 2007). The valley of the river Rekijoki for instance, is a good example of a valuable concentration of grassland in southern Finland. With 218 ha of mesic grassland, it is the largest area of its kind in Finland (Kontula *et al*. 2000).

According to Statistics Finland (2009), grassland crops were cultivated nationwide on a total of 655 200 ha of arable land in 2008 (approx. 30% of all arable land). Of this figure, 451 600 ha went to silage, 102 500 ha was used for hay-making and 80 600 ha was reserved for pasture. Another 20 600 ha went to the production of grass and clover seeds as well as green fodder. In terms of *rough grazing* (Finnish: *luonnonniitty*) and pasture on a regional level, 1 989 ha occurred in the region of Southwest Finland in 2008. This figure excludes data for the agricultural production of fodder and seeds.

2.4 Decline in grassland abundance

There have been dramatic declines in Finnish grasslands during the last centuries, especially after World War II (Pykälä 2001, Luoto *et al*. 2003, Tiainen *et al*. 2004, Haapanen 2005). During the last century, the total area of semi-natural grasslands shrank to roughly 1% of its original size (Vainio *et al*. 2001). Whereas the total area of meadow, for instance, still amounted to approximately 1.6 million hectares just over 100 years ago (Alanen 1996), less than 22 000 hectares or 1.38% of that remained in 2006 (Statistics Finland 2007). This evolution is mainly due to rapid decrease of small-scale dairy farming and associated grazing (Pykälä 2001, Luoto *et al*. 2003, Tiainen *et al*. 2004) and typically results in conversion to arable field or afforested ground (Haapanen 2005). The latter process is generally referred to as *bush encroachment* or *vegetative overgrowth*.

Similar declines have been observed in neighbouring Sweden, where there has been a reduction of semi-natural grasslands by around 90% from at least 2 million ha at the beginning of the 20th century, to *ca.* 270 000 ha at the end of that same century (Öster *et al*. 2008).

2.5 Grassland species richness

Grasslands, and more specifically meadows and pastures originating from traditional practices, are considered key habitats of European agricultural environments. They are typically characterised by high species richness and a great number of specialised species (Kivinen 2007). Although Finnish grasslands only account for a relatively small stake of all agricultural land across the country, they play a crucial role in enhancing biodiversity of the boreal environment (Pöyry *et al*. 2004).

There are, however, notable differences in species richness of grasslands, largely depending on environmental factors. Water supply, soil type, solar irradiation, fertilisation, presence of grazing and local topography are among the most important. Dry and mesic grasslands are the most diverse and species-rich types of grassland (Pykälä 2003, Cousins 2001) and the same was found to be the case for grasslands with a tree cover of *ca.* 5% (Pykälä *et al*. 2005). Topography on the other hand, has an especially pronounced effect on species richness in northern Europe since it is where many plant species are on their northern distribution limit. Grasslands situated on south- and west-facing slopes (*i.e.* those slopes that receive the highest solar irradiation) are particularly species-rich (Pykälä *et al*. 2005).

2.6 Fluctuations in grassland vegetation

The species composition of natural grassland fluctuates much more in response to fluctuations in the environment than does that of the woody component of scrubland and forests. This is because the herbaceous shoots of grasslands have a much shorter life-span than do woody stems, but also because the degree of phenological influence by weather conditions increases along the forest-scrubland-grassland gradient, especially in respect to precipitation (Coupland 1974, Rabotnov 1974).

The grassland communities that are most responsive to changes in weather are presumably those dominated by annuals, which depend on germination to provide new plants at the

beginning of each growing season. In grasslands dominated by perennials, the life form of dominant grasses determines the degree of change that will take place in the density of plants or shoots (Coupland 1974).

Plant community fluctuations are the dynamics of phytocoenosis that occur from year to year or from one period of time to another as a consequence of peculiar meteorological, hydrological and other conditions important for plant growth. Unlike plant successions, fluctuations are characterised by (a) differently oriented changes in phytocoenosis from year to year or from one period to another, (b) reversibility of changes and (c) stability of floristic composition in typical cases (*i.e.* the absence of invasion of new species) (Rabotnov 1974).

2.7 Vegetative overgrowth

When the accumulation of biomass continues uninterruptedly or when invasive species permanently install themselves in a given patch of grassland, it is considered to undergo vegetative overgrowth rather than to be subject to mere fluctuation dynamics. The process of vegetative overgrowth is often initiated when grassland management practices such as grazing or mowing cease and is most imminent or notable when it affects open or semi-open grasslands. When site management is terminated in these biotopes, the standing crop continues growing at first. A few competitive tall grass and herb species however, become dominant and the amount of trees and bushes increases (Pykälä *et al*. 2005).

There is, however, a notable difference in time span needed for vegetative succession to take effect, depending on environmental conditions. Succession is more rapid in mesic to wet grasslands than in dry grasslands with a thin soil layer (Cousins 2001). Steep slopes may be less sensitive to overgrowth given their unusual microclimatological conditions (Pykälä *et al*. 2005). Other factors such as litter accumulation and the presence of perennial species may alter the rate of biotope conversion considerably. In fact, given the right circumstances, they may prevent the establishment of trees and shrubs altogether. If tree saplings are not present when management is, grasslands may occasionally remain treeless for decades (Pykälä *et al*. 2005).

Not only environmental conditions, but also the type of management regime the grassland is subjected to can heavily influence the way vegetative overgrowth proceeds. Hay fields for

instance, are often subject to intensive agricultural practices that aim to maximise the yield. This translates into ploughing, sowing and fertilising these grounds using heavy machinery. When the cultivation of a hay field ceases, perennial species continue to grow but the colonisation by annual species is almost completely suppressed due to the sown grasses (Jukola-Sulonen 1983). The course of natural succession in the respective field is affected by these grasses for a long time after the hay harvesting activities have ceased (Beckwith 1954).

When natural succession in abandoned grasslands adopts such proportions that it becomes a concern of some kind, it is often referred to as *vegetative overgrowth*, *bush encroachment* or *site enclosure*. In contrast to the use of *vegetative overgrowth* however, the latter two terms draw attention to the non-linear growth process in which taller vegetation grows radially from existing stands of bushes and trees. Typically, a patch of grassland subjected to overgrowth will temporally retain its original low-growing vegetation in central parts while the growth of taller vegetation such as tree saplings is already well under way along its edges.

Determining whether a given patch of grassland can be considered overgrown or not remains to some extent subject to specific site conditions as well as to the appreciation of the observer. Pykälä (2003) for instance, considered mesic semi-natural grasslands to be overgrown when they had been abandoned more than 10 years earlier whereas others have concluded vegetative overgrowth was at work much sooner after abandonment. Despite the absence of fixed quantitative measures for interpreting the overgrowth status, there are a number of site characteristics that may indicate site abandonment. First and foremost, there will be shrub and tree encroachment as well as notable amounts of dead grass (Cousins 2001). In terms of moisture supply, drier land will have become damper (Lindgren 2000), which in turn affects the faunal and floral composition. Many perennials and shade tolerant plants for instance, prefer moist to mesic environments and will therefore outcompete other types of vegetation. Light-demanding species and annuals are typically negatively affected by vegetative overgrowth (von Numers & Korvenpää 2007).

The vegetative overgrowth of grasslands in Finland has received excessive attention by the scientific community and relevant authorities and there is consensus that the issue has affected a substantial number of grasslands throughout the country. Mikkola (1997) considered overgrowth of meadows of such large scale as to propose the term *Bush-Finland*. Despite a number of localised initiatives taken by the responsible authorities, vegetative

overgrowth is an ongoing process. As part of a near-nationwide study of grasslands, Kivinen (2007) observed clear discrepancies between grassland data derived from the SLICES database (NLS 2000) and in-situ conditions and related them directly to ongoing overgrowth.

2.8 Habitat monitoring using remote sensing

Despite ongoing efforts, the recent status of grasslands in Finland is poorly known (Marttila *et al*. 1999). This is where the advance of remote sensing technology can come to aid. Whereas in 1990 Meyer and Werth still claimed that digital remote sensing was not a viable tool for most resource applications, there now seems to be scientific consensus that the technology is a vital tool for environmental assessment. Satellite images are especially useful to study large areas (Gallego 2005) and their ability to provide consistent and repeatable measurements at a spatial scale appropriate for many processes causing change on the land surface renders them highly suitable for monitoring changes on the Earth's surface (Kennedy *et al*. 2007). With particular regard to the analysis of grassland and variations in their composition and spatial characteristics, remote sensing offers promising prospects. Multispectral sensors like Landsat TM have been used for the discrimination of grassland types under different management practices (*e.g.* Peterson *et al*. 2002, Price *et al*. 2002, Guo *et al*. 2003, Miatkowski 2004, Psomas *et al*. 2005). Ever-improving precision and accuracy of the material allows the retrieval of more subtle details.

2.9 Landsat TM imagery

Landsat is an Earth-observing programme administered by both the NASA and the U.S. Geological Survey. It consists of a series of satellite missions, providing specialised digital photographs of Earth since 1972. The only fully functional Landsat platform currently in orbit is the Landsat 5 satellite, launched on 1 March 1984 (NASA 2009a). This implies that the images its Thematic Mapper sensor captures today rely on technology of over 25 years old, which explains why it has only 6 bandpasses in the reflective section of the electromagnetic spectrum while its spatial resolution is limited to 30 m ground sampling distance (GSD). Furthermore, with a 16-day repeat cycle, Landsat 5 cannot readily provide frequent and global coverage (Rock *et al*. 1993). Compared to newer sensor models, Landsat TM imagery may appear outdated and non-competitive.

However, due to its early launch and continued operation, the data from the Landsat spacecraft constitute the longest record of the Earth's continental surfaces as seen from space (USGS 2009a). This renders TM data especially useful for change detection analysis, certainly when considering sensor consistency. These merits are highly respected, even in current-day mapping projects. The Global Land Survey Program (GLS) for instance, still employs Landsat 5 images – among others – for its international *Global Land Survey 2010* project, which aims to build global satellite coverage of Earth by the end of 2011 (GLS 2010).

2.10 Vegetation response to irradiance

Transmission, reflection, absorption, emission and scattering of electromagnetic energy by any particular kind of matter are selective with regard to wavelength, and are specific for that particular kind of matter, depending primarily on its atomic and molecular structure (Colwell 1969). The cells in plants for instance effectively scatter light because of the high contrast in the index of refraction between water-rich cell contents and the intercellular air spaces. The reflectance and transmission of light from plants in the wavelength region 0.5-2.5 μm is influenced by at least three phenomena (Ritari & Saukkola 1985):

1. Firstly, chlorophyll and carotene absorption at visible wavelengths to about 0.7 μm. Blue and red light are absorbed by foliage whereas up to 20 percent of the incident green light is reflected. In autumn, chlorophyll deteriorates, reducing the absorption of incident red energy (Sabins 1997).

2. Secondly, the physiological structure of the plants at NIR (near-infrared) wavelengths to 1.3 μm. Values in this range of the electromagnetic spectrum peak strongly for chlorophyll in healthy vegetation, meaning that most of the NIR energy is reflected. Since most incident light is absorbed by leafy vegetation in TM band 3 (red visible light, 0.63-0.69 μm), this results in a characteristic *red-edge* between TM band 3 and the adjacent TM band 4 (NIR) (Infoterra 2009).

3. Thirdly, there is a strong absorption by the water contained in the plant tissue between mid-infrared (mid-IR) wavelengths of 1.3 and 2.5 μm. Thematic Mapper bands 5 and 7 are therefore indicative of vegetation moisture content and soil moisture. By absorbing

the incident electromagnetic energy in these bands, water lowers DN values, meaning dry material results in relatively higher values (Infoterra 2009). Soil moisture may significantly affect perceived leaf water content from 2.1-2.5 μm if plant cover is low (Jacobsen *et al*. 1995).

In addition to the above-mentioned phenomena, a number of other factors are important. Morphological characteristics have been correlated with reflectance across the 0.5-2.5 μm spectral region (Gates 1970, Goillot 1980). Imaging and illumination angle, the number and configuration of leaves, other parts of plants, shadows and background response all affect the spectral signature of a given set of vegetation as well. Atmospheric effects are known to exert a strong influence, especially in visible bands. Thematic Mapper band 1 (0.45-0.52 μm) for instance may have up to a 70% contribution from sky radiance (Infoterra 2009). However, not all circumstances affect spectral response as strongly. The intensity of radiation reflected by vegetation in the range of visible light for instance, changes only slightly with a change in angle of incidence (Coulson 1966). The effect of all these interferences varies depending on the wavelength range and has to be considered when discriminating different vegetation surfaces by their spectral signature (Goillot 1980, Hildebrandt 1976).

2.11 Spectral separability of grasslands

Despite a range of possible interferences when measuring reflected or emitted radiation, vegetation-specific interaction with incident light can generally be picked up if it is sufficiently different from other types of vegetation. If so, the vegetation type concerned is considered spectrally separable. In other words, to be spectrally separable means that the variance in reflectance between the investigated features is greater than that within investigated features (Schmidt & Skidmore 2001). Separability is therefore a measure for the ease with which patterns can be correctly associated with their classes using statistical pattern classification (Richards & Jia 2006).

Grassland as a spectral class is often treated as a single land cover class. Moreover, for a long time it was believed that no high accuracy distinction was possible between grassland and certain other types of vegetation such as cropland using imagery alone (Anderson *et al*. 1976). However, grasslands do by no means constitute a homogenous class in terms of

spectral response. They generally encompass a great range of vegetation communities for which the spectral response varies depending on differences in agricultural management, soil, climate and other environmental parameters (Askew & Slater 1995). In fact, the British National Vegetation Classification recognises 48 different main communities of grassland in Britain alone (Rodwell 1992), each of which must show slight variations in spectral response.

Several research groups have undertaken to characterise the spectral discrimination potential for different types of grasslands (*i.e. grassland discrimination analysis*), using *in situ* spectrometer measurements, spectral recordings from an airborne platform and satellite imagery. Jacobsen *et al*. (1995) for instance, investigated the discrimination potential for both different types of grasslands using spectral reflectance data obtained from *in situ* measurements. In their study area Mols Bjerge (NE Jutland, Denmark), four types of grassland were discerned based on species diversity and age. When plotting the 10 nm resolution spectral measurements, they found that although some differences in reflectance patterns can be observed, some challenges with regard to grassland type recognition remained. When they subjected the data to a discrimination method combining t-tests and F-tests, satisfying discrimination potential values were obtained for the visible, near-infrared and mid-infrared parts of the electromagnetic spectrum.

Especially when dealing with biogeographical phenomena subject to a strong phenological influence such as grasslands, special attention needs to be paid to temporality. In this light, Psomas *et al*. (2005) focused on deriving optimal points in time during the growing season for discriminating four different Swiss grassland types spectrally. Not surprisingly, they found there is seasonal variation in the hyperspectral recordings they obtained for grasslands. Furthermore, combining recordings from throughout the growing season seemed to offer a much better understanding of the spectral differences between grassland types and increase the possibilities for successful discrimination and classification. Certain grassland type combinations showed better separability earlier in the growing season than others.

Using multitemporal imagery in order to improve the results of grassland discrimination analysis however, is no guarantee for success. Peterson *et al*. (2002) tested single date and multitemporal classification approaches when attempting to discriminate between cool season and warm season grassland cover types in northeastern Kansas but concluded

multidate images did not lead to higher discrimination accuracy than when using a single dataset for mid-summer and its derivatives only. This can be attributed to the particular types of prairie grasslands they studied since they exhibit inherently different phenological developments and biotope-specific spectral characteristics. Little can therefore be generalised regarding the spectral discrimination of grasslands and caution needs to be exerted when relating findings to similar studies in other environmental settings.

Although most grassland discrimination analyses have concentrated on differentiating between grassland *biotopes*, some studies have extended their scope to include discrimination between different grassland *management* types. Especially the tallgrass prairies of the Central Great Plains (North America) have received considerable attention in this sense. Many of the prairie's native and non-native grasslands and associated land use practices have been established to be spectrally distinguishable (Guo *et al*. 2003). In their study, Guo *et al*. obtained an overall accuracy of 70.4% TM image classification accuracy for three common grassland management types: Conservation Reserve Program (CRP), grazed and hayed grasslands. CRP grasslands are grasslands resulting from a soil erosion prevention initiative and are characterised by a mixture of seeded native grass species and a burning and/or mowing management regime. Despite moderately high accuracy results, they found that obtained accuracy levels are highly grassland treatment-dependent and also advised against the generalisation of the potential to discriminate different grassland management, saying that the success of the analysis is highly site- and imagery quality-dependent.

2.12 Non-quantitative assessment of spectral separability

A common way to represent the spectral separability of different types of vegetation cover is by means of a spectral response curve plot. Each individual spectral response curve results from the quantitative measurement of the spectral properties of an object at one or several wavelength intervals (NASA 2009b). Spectral response curves are sometimes also referred to as *spectral signatures* because the reflectance and emittance of incident electromagnetic radiation is specific to the feature or land cover type that receives it. Spectral response curves typically show *absorption bands* and *reflectance peaks* (cf. Fig. 1 below). Absorption bands occur at intervals of the electromagnetic spectrum where electromagnetic radiation is absorbed by the atmosphere or other substances whereas at reflectance peaks, the reflectance or emission of radiation is at its climax (Navalgund 2001).

Spectral response curves are mostly constructed from *reflectance spectrometers*, which record the response at high spectral resolution along large parts of the electromagnetic spectrum.

The analysis of spectral response data can be based on *in situ* collections as well as on remote sensing material. In the latter case, the analyst needs to delineate the feature(s) of interest in the imagery in order to collect representative reflectance/emission data, also known as *training*. For each feature class, the statistical quality of the data extracted from the training sites increases with the total number of training pixels that are included. To attain a representative sample, a minimum of from 10*n* to 100*n* pixels is used in practice, where *n* represents the number of spectral bands (Lillesand and Kiefer 1994). Not only do the spectral means retrieved need to be based on a large enough sample, data need to be collected from training sites throughout the study area. Once the initial training stage is completed, the spectral response data for each feature class and each interval of the electromagnetic spectrum under investigation need to be individually checked for normality. Even though common spectral response distributions can be generally assumed to be normal (Lillesand and Kiefer 1994), non-normality can occur and indicates the presence of a spectrally distinct sub-class or non-representativeness of the sample.

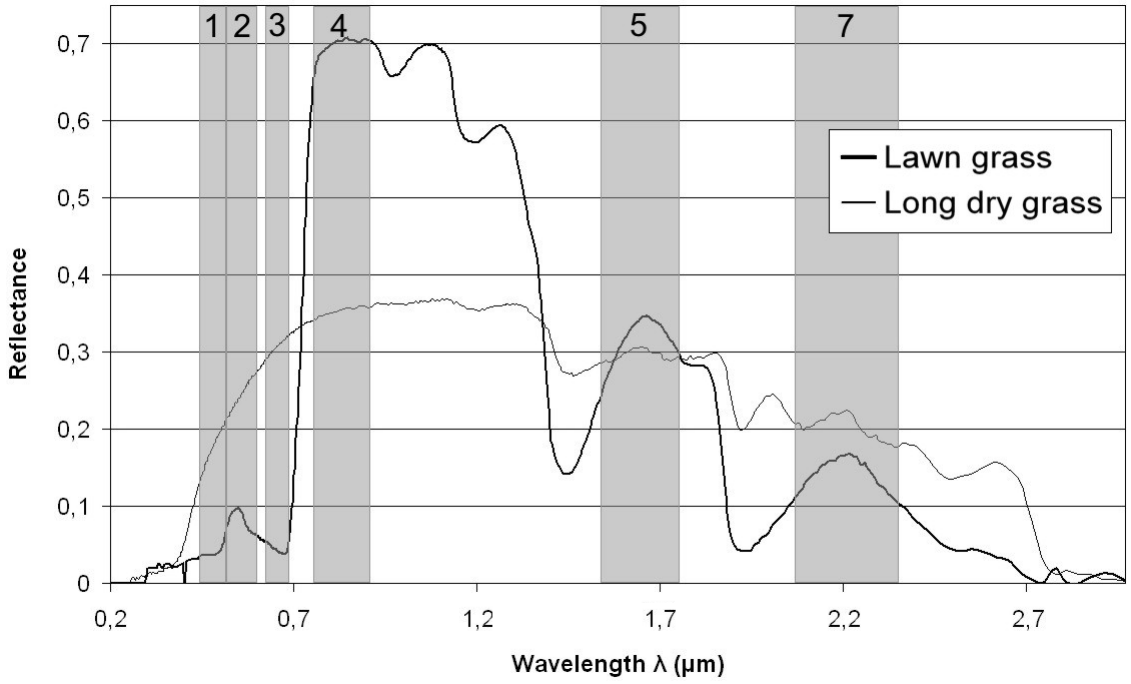


Figure 1. Spectral signatures for green lawn grass and brown dry grass derived from tabular data available through the U.S. Geological Survey's digital spectral library (USGS 2009b). The grey bands in the background represent Landsat TM bands 1 (blue), 2 (green), 3 (red), 4 (NIR), 5 (mid-IR) and 7 (mid-IR).

There are different ways to present a spectral response data. The conventional representation is a 2-dimensional plot in which the recorded reflectance/emittance (y-axis) is plotted along a wavelength gradient (x-axis) (*e.g.* Fig. 1). This method however does not allow for the analyst to visually assess if the data normality condition is complied with. By separating all spectral bands for a set of response data for a single feature of interest and by treating them like frequency distribution charts, bimodality or other deviations from a normal data distribution quickly show. If normality is indeed accounted for, and an effective way is desired to visually check response data derived from different features for spectral distinctness, a so-called *coincident spectral plot* (Lillesand and Kiefer 1994) may be useful. Instead of showing the entire frequency distribution for each feature and for each layer, it merely shows where the mean value and respective standard deviation derivatives of a given response are located along the digital number (DN) spectrum for each band. This allows for the spectral response of several feature classes in several spectral bands to be shown simultaneously.

However, the coincident spectral plot only really tells what categories could not be accurately classified on any *single* band. Features with an apparent overlap in – say – two different spectral bands, may in fact be separable. Scatter diagrams allow the analyst to plot spectral response data for a number of features along two or sometimes even three axes, each representing the DN spectrum for a different spectral band (*e.g.* TM bands 2, 3 and 4). In these diagrams, separability issues only persist for the overlapping areas between different feature *ellipses*. Unfortunately, such visual scatter analysis is limited to 3D vision, so for more advanced analysis involving more than 3 spectral bands the theoretical computation capacity of artificial intelligence needs to be employed.

A more recent trend in visualising spectral separability is to concentrate on statistically significant differences in spectral response only. This accommodates the need to involve more feature classes and a much higher spectral resolution (often up to hyperspectral level). Only if feature pairs are statistically significantly different in their spectral response, are they entered into their respective wavelength class in a frequency distribution chart. This accumulative approach allows for the analyst to appreciate in which parts of the electromagnetic spectrum feature classes are most often significantly different, indicating optimal bands for spectral separation (*e.g.* Schmidt & Skidmore 2001).

2.13 Quantitative assessment of spectral separability

Naturally, spectral separability assessment techniques are not limited to those that produce a means for visual interpretation (graph, diagram, *etc.*). Several methodologies have been proposed. One way to quantitatively express statistical separation between response patterns is *divergence*, a covariance-weighted distance between category means. It is calculated for all response pairs and presented in the form of a matrix (Lillesand & Kiefer 1994). Additionally, one can also conduct a more pragmatic test and quite simply run an unsupervised classification on all training site pixels selected. By arranging the obtained classification success rates in a matrix (so-called *contingency matrix*), the spectral distinctness between the delineated feature classes can be estimated (ERDAS 2005).

Certain distinct landscape features can never be spectrally separated based on a single input image, no matter how carefully training pixels are selected. This problem can however be mitigated in a number of ways, among which being the employment of auxiliary data, visual interpretation, field checks and the use of multitemporal or spatial pattern recognition.

2.14 Spatial separability

Spatial separability refers to the ability of an electronic image to capture the desired level of detail in spatial terms. Since the basic structure of electronic imagery is typically a regular grid of pixels, spatial separability is directly related to grid cell width. A pixel's width, rectangular shape, orientation and regularity impose important restrictions on how data are captured, stored, processed, interpreted and represented. Imposed as an artificial division of the space that is imaged, the pixel is very unlikely to match the contents of that space (Fisher 1997).

Figure 2 illustrates the mathematics behind patch size in relation to obtaining one or more pure pixels. The dark shapes in the centre represent the extent and the shape of a given patch for which the vegetation cover is assumed to be homogeneous. Diagrams A1 and B1 illustrate the minimum dimensions of a circular and a square patch respectively if the number of pure whole pixels obtained when capturing them in an electronic image with resolution P can be one single pixel at most. Obtaining a single pure pixel for a circular patch with area $S_{patch} = (\pi.P^2)/2$ or a square patch with $S_{patch} = P^2$ is possible, but highly unlikely since it requires for the image grid to perfectly coincide with the patch in the

manner illustrated in diagrams A1 and B1 respectively. In the case of Thematic Mapper 30 m resolution imagery, this implies that a square patch needs to be at least 0.09 ha in size whereas for a circular patch the necessary area amounts to 0.14 ha if it is to be theoretically possible to obtain a single pure pixel for these patches.

Diagrams A2 and B2 illustrate the necessary spatial properties of a circular and a square patch respectively if capturing them using electronic imagery is to result in at least one whole pure pixel, regardless of where and how the capturing grid is placed. When the image's pixel size is *P*, the area of a circular patch S_{patch} needs to equal $2\pi.P^2$ whereas the area of the square patch needs to total $8P^2$. For 30 m resolution imagery, the necessary size of the circular and square patch amounts to 0.57 ha and 0.72 ha respectively. All threshold values are recapitulated in Table 1 (p.25).

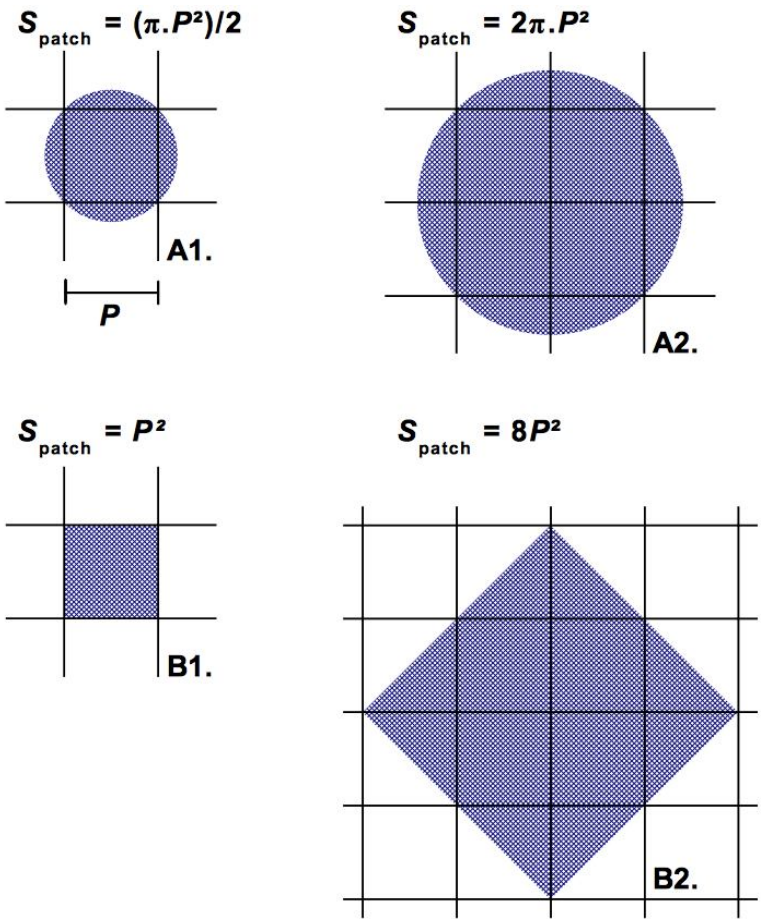


Figure 2. Theoretical area values (*S*) for circle-shaped (A) and square-shaped (B) patches for when at most one pure pixel can be obtained in corresponding remote sensing imagery (A1 and B1) and for when at least one pure pixel will be obtained (A2 and B2). The grids represent the pixel arrangement in electronic imagery with pixel size *P*.

Most patches of grassland in the Turku Archipelago have much more complex shapes that that of a perfect circle or square. Especially grasslands such as traditional pastures are often located in areas where the terrain is too complex to convert into arable land. Coastal meadows and rocky grounds are good examples of this. Such patches of grassland will typically exhibit complex shapes with fuzzy boundaries and thereby raise the difficulty associated with obtaining pure pixels when imaging the area.

Table 1. Summary of threshold area values for circular and square patches if they are to be captured by 30 m resolution remote sensing imagery.

Patch shape	max 1 pure pixel	min 1 pure pixel
Circular	≤ 0.14 ha (ø = 42 m)	≥ 0.57 ha (ø = 85 m)
Square	≤ 0.09 ha (*s* = 30 m)	≥ 0.72 ha (*s* = 85 m)

Whatever resolution of the sensor however, investigators will attempt to extract information which is actually smaller than the size of the pixel (Fisher 1997). Figure 3 below illustrates the spatial separability of two hypothetical grassland patches for the inner Turku Archipelago with typical size and shape complexity. The diagrams on the left depict the same patches after they were captured by 10 m resolution SPOT imagery and 30 m Landsat TM imagery respectively. Both resolutions incur a loss of spatial detail such as the

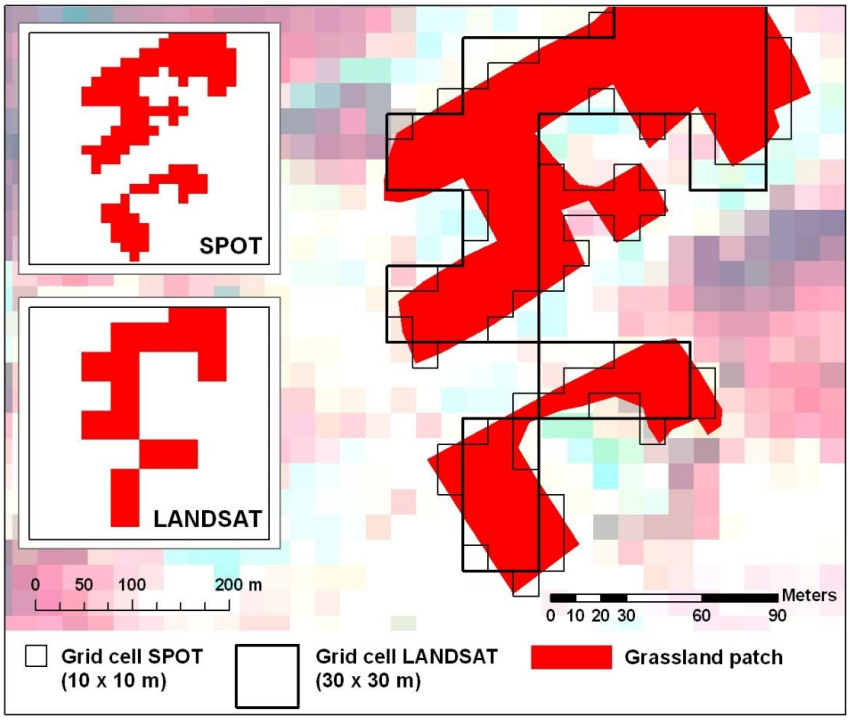


Figure 3. Assuming both occurrences of grassland represented here are recognised as successfully by LANDSAT as by SPOT, their respective representation is highly influenced by resolution restrictions. The larger patch portrayed in the figure above is approximately 1 ha in size.

general orientation of the patches and the presence of small appendicles, especially in the 30 m resolution interpretation.

Spatial resolution and with it spatial separability play an important role in determining the success of a land cover classification exercise, even when high-precision input data are at hand. In their study of shrub encroachment in southern New Mexico, Laliberte *et al*. (2004) noted that shrub and grass cover were underestimated by about 15% in areal coverage when classifying land cover in a relatively small area (150 ha) using aerial photographs and high resolution satellite imagery. The mean resolution of the input data amounted to a mere 0.54 m.

The smaller the cell width, the more spatial variation detail can be distinguished. Ideally, the spatial resolution of an image perfectly suits the detail extraction needs. The study of semi-natural grasslands from satellite imagery implies employing effective methods in isolating grass occurrence in the imagery. This however, may be a challenging task. In an effort to predict the presence of suitable habits for the clouded apollo butterfly (*P. Mnemosyne*) – which relies on traditionally managed rural biotopes such as flower-rich meadows for its existence – Luoto *et al*. (2002) found the classification or semi-natural grassland patches the most challenging part of the satellite image processing simply because they are often too small for mapping TM imagery. They concluded that the combination of these pixel size and pixel shape limitations undoubtedly affected the perceived size and continuity of patches.

3 Study area

3.1 Location

The study area consists of the innermost part of the Turku Archipelago, covering a square area between the municipal centres of Dragsfjärd, Paimio, Nousiainen and the island of Norrskata (Korppoo) with coordinates between 60°7'-60°37'N and 21°45'-22°47'E (Fig. 4). Its dimensions amount to approximately 45 x 45 km, resulting in a surface of about 2 000 km^2 in size. In the present work, the *Turku Archipelago* is defined as the geographical region delimited by the city of Uusikaupunki in the north, the administrative boundary separating the regions of Åland and Southwest Finland in the west and the Peninsula of Hanko in the east. It may also be referred to as *the Åboland Archipelago*. The geographical region named *the Archipelago Sea* and more specifically *the Archipelago Sea of SW Finland* differs from the Turku Archipelago in that it encompasses both the Turku Archipelago and the Åland islands (Frisén *et al*. 2005).

The Turku Archipelago comprises 6 665 islands larger than 1 ha and 15 818 skerries smaller than 1 ha (Granö *et al*. 1999). Most of these islands have emerged during the past few thousand years as a consequence of the process of glacio-isostatic land uplift (Kakkuri 1987), which roughly amounts to 35-40 cm/100 years in this part of Fennoscandia (after Eronen 2005, Frisén *et al*. 2005).

The Finnish archipelagos are typically divided up into zones, based on the differences in vegetation and topography. The Turku Archipelago for instance has an inner, middle and outer zone. Although the physical delimitation of these zones has by no means been fixed, the area under investigation in this study is – for its largest part – located within what is generally referred to as the inner zone.

In a recent report, the Finnish Environment Institute (2006) defined the inner zone of the Turku Archipelago as a stretch of archipelago reaching from northern parts of the city of Uusikaupunki to the eastern side of the Porkkala Peninsula. It is characterised by a notably high land-to-water ratio, large, continental islands with narrow water straits in between them and bays that reach far into the mainland. A more botanical way of approaching archipelago zonation is proposed by Frisén *et al*. (2005), who defined the inner zone as that

part of the archipelago characterised by very rich vegetation and forest that is often low but resilient against storms.

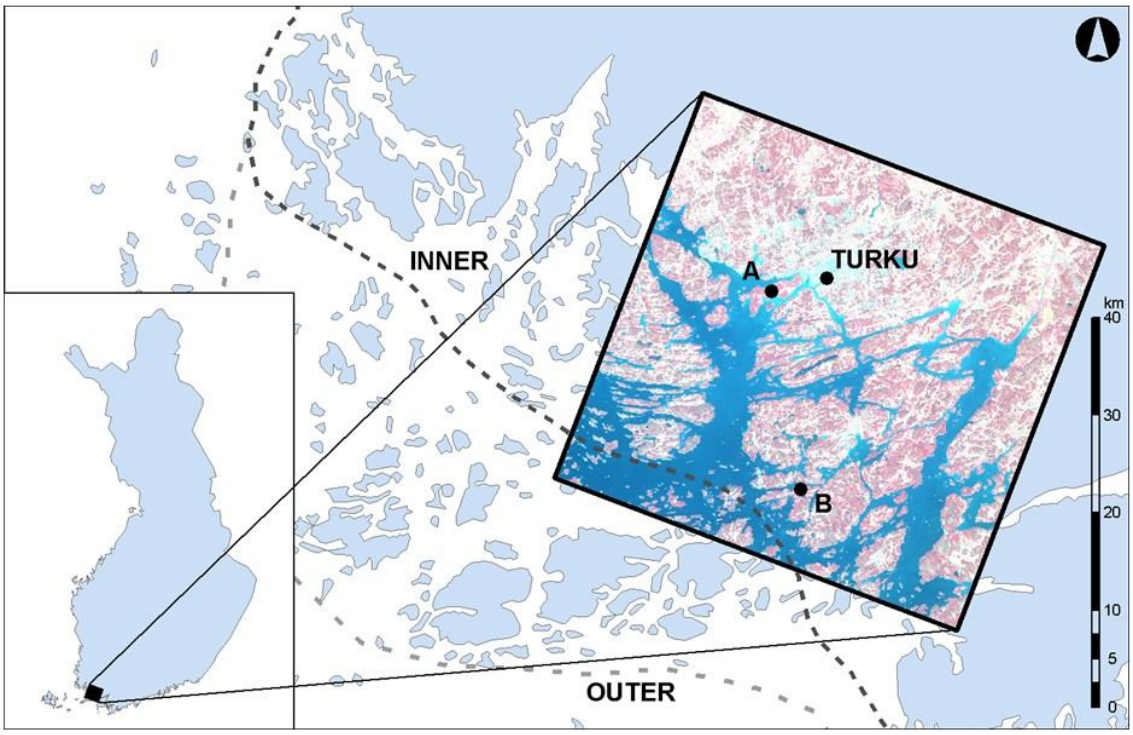


Figure 4. Location of the study site within the Turku Archipelago. The 45 x 45 km square area mainly falls within the inner archipelago zone, as defined by Vuori *et al*. (2006). The boundaries of the inner and outer archipelago zones are indicated here with darker and brighter dashed lines respectively. The marked points refer to the islands of Ruissalo (A) and Lenholm (B).

The Turku Archipelago is virtually non-tidal. However, changes in water level do occur and depend on atmospheric conditions. A combination of low pressures and strong winds for instance, may press the water mass into the extremes of the Baltic Sea basin. Records show that in the northernmost part of the Gulf of Bothnia, this has resulted in water stage variations up to 3.2 m. High-pressure systems during spring are typically associated with low water levels, whereas under stormy low-pressure conditions in autumn and winter, the water level can at times be threateningly high (Frisén *et al*. 2005, Ryhänen 2005).

3.2 Climate and growing season

Finland belongs to the cold temperate climate with cool summers (Tuhkanen 1984). Its southwestern archipelago, the area of interest in this study, is situated at the northernmost extreme of the hemiboreal vegetation zone, bordering the south-boreal zone (Ahti *et al*. 1968). In this area, the coldest month is February with a mean temperature of -4°C and the warmest month is July with a mean temperature of +15°C (Atlas of Finland 1987). Near

the sea and in the islands, temperatures over 30°C are extremely rare (FMI 2009a). The annual mean temperature is about 5.5°C (FMI 2009b).

Maritime influences balance the variation in temperature in coastal areas while rainfall is rather evenly distributed throughout the seasons (Wuolijoki & Iltanen 2005). The mean annual precipitation for the period 1961-90 is between 500 and 650 mm with the precipitation increasing along a directional gradient from west to east across the study area (Wuolijoki & Iltanen 2005). The least rain falls in March. Along the coast, the rainfall gradually increases until September and October and then decreases towards winter and spring (FMI 2009b).

The average duration of snow cover on open ground is between 90 and 120 days per year (Kuusisto 2005) and the length of the growing season (*i.e.* days with mean daily temperature of more than 5 degrees Celsius) is 180 days (Wuolijoki & Iltanen 2005). The thermal growing season starts when the mean daily temperature exceeds 5°C (FMI 2009a). This takes place at the end of April in southern Finland, one month after the beginning of spring. Spring begins in early April in the southwestern archipelago and is marked by a rise in average daily temperature from 0°C to 10°C. Because of a pronounced cooling effect by the cold seawater, spring lasts especially long in the coastal zone, reaching up to 65 days of duration. The onset of the growing season is also dependent on the amount of snow cover. Open areas typically loose their snow cover within two or three weeks of the beginning of spring, but if the snow remains for exceptionally long time, the vegetative growth may be delayed (FMI 2009a).

Summer usually begins in late May in southern Finland and lasts until mid-September and is characterised by an abundance of daylight, reaching up to 19 hours around Midsummer. In southwestern Finland, summer ends around the last week of September, but meteorological conditions that allow vegetation to grow persist until late October or early November. This is when mean daily temperatures in the SW Finnish Archipelago drop below 5°C. The average length of the growing season in the study area is therefore 180 days (FMI 2009a).

Not only temperature, but also the level of received solar irradiation is important for the rate of biomass accumulation. Due to cloud cover, the highest annual radiation levels are usually attained throughout Finland before aphelion, usually in early June. The daily

maximum normally occurs before noon, also due to cloud cover. The annual amount of sunshine in the southwestern maritime region amounts to 1 900 hours (FMI 2009b).

3.3 Site history

Pollen diagrams suggest that the Turku Archipelago is one of the first places in Finland where farming occurred with the earliest traces of arable farming and animal husbandry in the region have been dated to 2 000 – 1 500 BC (Lindgren 2000). Given the continuous isostatic land uplift which has affected the area since the end of the last ice age, the archipelago was located about 100 – 200 kilometres further to the northeast at the time (*i.e.* the latter Stone Age), in an area now part of the mainland. The larger islands today then constituted the outer archipelago and were marked by a land elevation about 20 – 50 meters lower than at present (Lindgren 2000).

Historically, the sort of meadow management typical for the archipelago area included raking dead plant material in spring, haymaking in July and grazing in autumn or occasionally also in spring (Kotiluoto 1998). Meadows were cut for hay on even the most distant islands, and cattle were transported to islands and left to graze freely throughout the summer (von Numers and Korvenpää 2007). Pollarding deciduous trees (*e.g. Fraxinus exelsior*, *Alnus glutinosa*, *Betula* sp.) for fodder took place in autumn (Slotte 1993). Of these practices, the grazing primarily resulted in different types of pastures while hay-making and leaf-fodder cutting maintained open or wooded meadows. This traditional type of animal husbandry developed and maintained the species rich semi-natural meadows and pastures in the Archipelago of SW Finland for centuries (Gardberg 1931).

During the first half of the 20th century, the abundance of grazing ungulates had taken on such proportions, that several botanists that had been studying archipelago flora described the islands as "destroyed by sheep". Large parts of the Archipelago Sea, especially islands in its outer belt, were unquestionably overgrazed at the time (Lindgren 2000).

In the post-war period however, such practices became increasingly uneconomical in the area and there was a large-scale reduction in traditional archipelago animal husbandry. For a few decades, the previously suppressed vegetation of meadows and pastures in the archipelago flourished. Given the extensive absence of biomass control however, the landscape started to become overgrown, faster on damp soils, slower on dry ones. Typical

low meadow flora quickly became overwhelmed by taller plants and grasses, and in time by brushwood and bushes (Lindgren 2000, Kotiluoto 1998).

The poor post-war socio-economic circumstances in the Turku Archipelago caused a large-scale depopulation of the region, which reached its peak in the 1970s. Animal husbandry shrank drastically in response and by the 1980s had virtually entirely disappeared. Since the 1970s, the area witnessed an extensive spread of overgrowth. Coastal meadows with low vegetation for instance, were overrun by different kinds of sedges, but above all by common reed (*Phragmites australis*). No other phenomenon has affected the terrestrial environment in the Finnish archipelago during the latter 20th century as powerfully as that of vegetative overgrowth (Lindgren 2000).

3.4 Grasslands in the archipelago

Although grasslands are generally scarce in southern parts of the Finnish mainland, the Archipelago Sea is one of the few remaining regions where the concentration of grassland is high both in terms of total coverage and share of all agricultural land (Statistics Finland 2007, Kivinen 2007). Grazing activity in the area is however in decline. Whereas in 2006, the administrative region of Southwest Finland (i.e. the region to which the Turku Archipelago belongs) still hosted 2 265 hectares of non-abandoned rough grazing (Finnish: luonnonniitty) and pasture (Statistics Finland 2007), this figure had been reduced to 1 989 ha in 2008 (Statistics Finland 2009). Based on regional figures for 2006, grazed land accounted for 0.21% of the entire land area for Southwest Finland, being inferior in importance only to Åland (4.59%).

In 1992, the Finnish Environment Institute started a nationwide inventory in order to map semi-natural environments and other agricultural biotopes. According to the findings of this large-scale effort, both the highest number and largest area of traditional rural biotopes is to be found in the province of Southwest Finland. Traditional rural biotopes are culturally important natural sites, consisting for the largest part of biotopes resulting from meadow and pasture practices. They encompass certain meadows, wooded meadows, *hakamaat* (wooded pasture with rocky outcrops), forest pastures, heaths and burn-beaten forest. Despite a land area that only accounts for 3.5% of that of the entire country, the region of Southwest Finland hosts over 15% of the nation's traditional biotopes area-wise, totalling 2 962 ha (Vainio *et al*. 2001). Although most of the valuable semi-natural

grasslands in Finland have been mapped for conservation purposes, studies on the distribution of this habitat type as such and its occurrence within a habitat mosaic have concentrated on limited study sites only and no methodology has been presented for the collection of landscape level information of this rare habitat type that often occurs in small patches (Toivonen & Luoto 2003).

According to the Finnish national topographic database (NLS 2007), a total of 1 568 meadows (partly) fall within the study frame and amount in total area to 1 895 ha. Table 2 lists the statistics for area, perimeter and perimeter/area for the patches occurring fully within the study area (n = 1 547). Further analysis shows that over 65% of all patches have a size of less than 1 ha. The vast majority of grassland occurrences have highly irregular shapes, often with long but thin appendages radiating from the centre of the patch body. Figure 5 illustrates the typical distribution, size and shape of grasslands in the inner Turku Archipelago.

Table 2. Spatial statistics for meadows occurring fully within the study area, based on a meadow inventory by the National Land Survey of Finland (NLS).

	Area *A* (ha)	**Perimeter *P* (m)**	***P*/*A* (m/ha)**
Min	0.03	74.5	129.7
Max	31.32	6014.5	4956.6
Mean	1.20	559.3	715.2

The Turku Archipelago hosts an array of different grassland types. Among 11 different types of pasture identified to exist or to have existed in the region alone, 5 large classes of open or mostly open grassland persist today: wooded meadow, coastal meadow, mesic meadow, dry meadow and fen meadow (Lindgren 2000). A more extensive classification is offered by the Ministry of Agriculture, listing 18 different types of grassland occurring in the innermost part of the archipelago for 2007. That is to mention semi-natural grasslands and agricultural grasslands only. Although not as proportionately significant, other types of grassland such as recreational grassland and something that I called *infrastructure grassland* also occur in the Turku Archipelago. Infrastructure grassland is grassland that is purposely created and maintained for the sake of infrastructure functionality such as stretches of mowed land underneath high-voltage power lines and road verges.

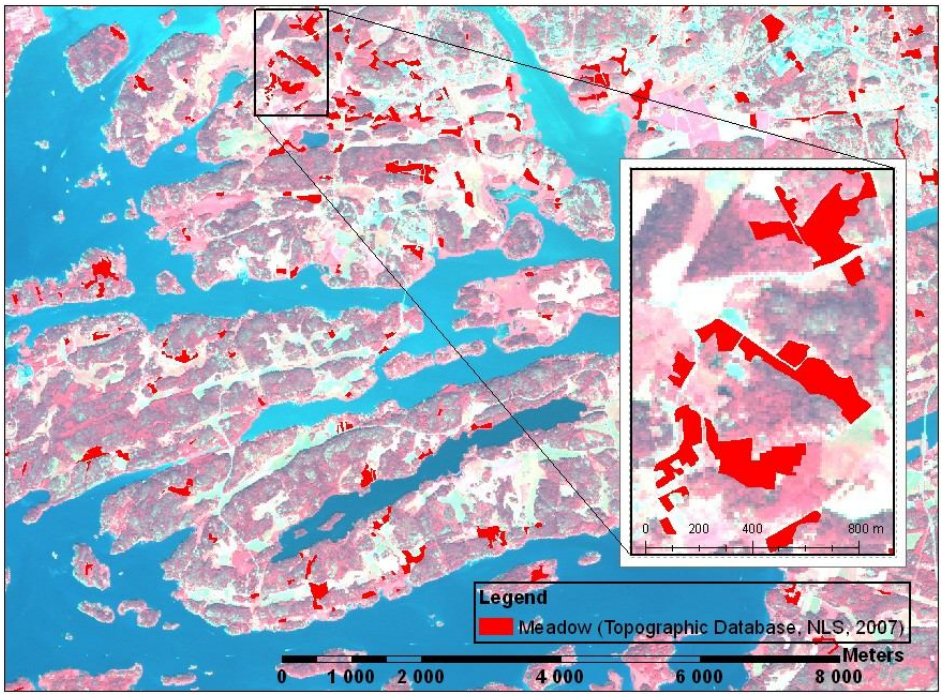


Figure 5. Selected site within the inner Turku Archipelago showing meadows as recorded in the nationwide Topographic Database (NLS 2007). Meadows occur scattered across the landscape at varying density and typically exhibit complex, fine-scale shape geometry. Total patch size varies greatly. Note that the NLS meadow data layer depicts mainly pastureland.

3.5 Field work sites

A number study sites in the study area were selected for closer inspection of the site conditions: the island of Ruissalo near the city of Turku and a set of sites somewhat further out in the archipelago, one of which is the island of Lenholm. Both Ruissalo and Lenholm are of special scientific interest in that they host an array of unusual landscapes and species diversity. In fact, due to natural characteristics and long-term human management, Ruissalo has become home to one of the richest species communities in Finland (Vuorela 2000). Another surprising quality of Ruissalo is its occurrence of stands of Pedunculate Oak (*Quercus robur*), which is atypical in Finland (Käyhkö & Skånes 2006). The island is now managed as a recreation area but with a large proportion in nature reserves and is characterised by a highly patchy landscape, resulting centuries of use as meadows and pasture for grazing, wood production and scattered habitation. This is typical for the hemiboreal coastal archipelago of Southwest Finland (Käyhkö & Skånes 2006).

Lenholm is famous for its oak stands too. In fact, its oak-dominated broad-leaf forest and traditional pastures are considered among the best in Finland (Lindgren 1998). Moreover, the oak forests, dry meadows and grazed coastal meadows comprise the entirety of the

landscape, resulting in an exceptionally high biodiversity. Records for the island date back to the end of the Middle Ages. The area was presumably used as meadow- and pastureland for centuries. Nowadays the southern past of the island is a nature reserve. Although moderate in total size (35.7 ha), about a third of this consists of various types of meadow, pastureland and heath (Lindgren 1998).

The elevated scientific interest associated with both Ruissalo and Lenholm has given rise to numerous publications and datasets (e.g. Krogerus 1921, Söderman & Tenovuo 1960, Kallio 1979, Lindgren 1998, Metsähallitus 2010). This implies historic records, remotely sensed material (aerial photographs and satellite images) and other types of documentation are readily available for these sites. This allows a better understanding of the current *in situ* conditions and how they came about, rendering these islands highly suitable reference sites for this inner archipelago-wide study.

4 Material

4.1 Ministry of Agriculture data

No single data layer depicting all grassland occurrences is available for the SW of Finland. This can be attributed to the multifaceted character of grasslands; they fulfil diverse functions as they serve both recreational and agricultural purposes, which are rarely depicted in the same GIS data layer. Furthermore, grass often also occurs in places that are not subjected to any type of management. Such abandoned sites are often not mapped in relation to their vegetation cover. Using only existing cartographic and other types of land cover/use inventory material, it is therefore impossible to compile a dataset listing *all* grassland occurrences in the study area.

In order to enable studying the spatial separability of grassland patches across the study area nonetheless, a selection of existing grassland spatial data was compiled based on official agricultural land use data from the Finnish Ministry of Agriculture for the years 2006 and 2007 so as to temporally coincide the land use data with the Landsat imagery used for the present study (see below). Agricultural land use data are provided in two parts. One part consists of a nationwide vector dataset depicting the official field parcel boundaries (*peltolohkorekisteri*) and is compiled by the Finnish Agency for Rural Affairs (*Maaseutuvirasto*) based on digitised orthophotos with a pixel size of 0.5-1 m and a spatial accuracy of 2.5 m. The second part consists of the agricultural land use data, which are distributed based on a temporal and spatial selection of the data. The selection of data used for the present study was based on the collection of municipal units intersecting the area of interest square (Kaarina, Masku, Naantali, Parainen, Raisio, Rymättylä, Turku, Kemiö, Lieto, Nauvo, Piikkiö, Rusko, Sauvo, Dragsfjärd, Korppoo, Lemu, Merimasku, Nousiainen, Paimio ja Velkua).

4.2 National Land Survey data

In addition to the Ministry of Agriculture data, a second set of high-detail vector shapes depicting land cover was consulted, namely those available through the *Topographic Database* (NLS 2007). The database incorporates the most accurate positional data about Finnish topography, and in this respect is comparable to maps with a scale of 1:5 000 - 1:10 000. Its

positional accuracy is approximately 5-10 m, depending on the feature class concerned (NLS 2010). Two vector layers of interest form part of the *Terrain1*-area data group: meadows and fields. The meadow vector shapes refer in fact to grazed grasslands (i.e. pasturelands) and do not represent present-day conditions to the highest accuracy. This implies that in terms of the depiction of grasslands, the meadow shapes of the Topographic Database are no more than a mere indication of possible grassland occurrence. The field shapes on the other hand are more accurate.

4.3 USGS Spectral Library data

In order to enable to verify the validity of spectral responses obtained part of the present study, high-resolution laboratory response spectra provided by the U.S. Geological Survey were incorporated. Specialised scientific bodies have documented the spectral response of a wide range of inorganic and organic materials. Some of these data are freely available in graph as well as in tabular format. The U.S. Geological Survey Spectral Library (cf. Clark *et al*. 2007) hosts an elaborate collection of continuous spectral responses for countless mineral types and – to some extent – also biotic material, including individual plants and composite vegetation communities. Of the available biotic spectra, the tabular response data for lawn grass and dry grass were downloaded.

4.4 Remote sensing data

Two types or remote sensing data were consulted for the study: false-colour aerial photographs and Landsat Thematic Mapper satellite imagery. The aerial photographs (scale 1:30 000) were available through the Laboratory of Computer Cartography of the University of Turku and had already undergone high-accuracy mosaicing so as to fully cover the islands of Ruissalo and Lenholm respectively. Several mosaic versions from several years are available, but in order to have the aerial imagery approximate the selected satellite imagery, the most recent mosaics were opted for. These image compilations depict the summer phenology for the years 2002 and 2003 for Ruissalo and 2003 for Lenholm respectively.

The selection of satellite imagery was based on available imagery for 2006 and 2007. Several Landsat image scenes from these two years were purchased by the Department of Geography and offered an appealing research opportunity for a master thesis project. Although initially restricted to a few purchased Landsat scenes, all of the Landsat 4-5 TM

and Landsat 1-5 MSS images became available online for free download through the U.S. Geological Survey's EarthExplorer and Global Visualization Viewer on 8 December 2008 (USGS 2009c). This had important implications for the choice of input imagery in that this constituted an opportunity to assess *all* available Landsat images for 2006-2007 for their grassland discrimination potential. For the Turku Archipelago this meant a total of six Landsat 5 TM scenes characterised by good image quality were available for this study. Table 3 below lists the images concerned.

Table 3. List of all six Landsat TM scenes freely available and suitable for inclusion in the present study.

Date of capture		**Selected for study**
1 May 2007	-	X
2 June 2007		
8 June 2006		
17 July 2006	-	X
5 August 2007		
21 August 2007	-	X

The available TM 5 imagery covers the phenological period between the beginning of May and the end of August. This largely corresponds to the duration of the growing season in the south of Finland. In terms of the potential of the imagery to allow for spectral responses for grassland to be contrasted for various stages during the growing season, the image of 1 May 2007 showed very promising characteristics, even before it was subjected to any type of enhancement. It appeared that during the earliest stage of the growing season, grasslands stand out and can be much more accurately be distinguished that in other images. In addition to an early growing season image, two other images representing key stages of the growing season were opted for: one image from 17 July and one other from 21 August depicting peak greenness and plant senescence respectively. The full details for the selected TM scenes are available in Annex I.

5 Methods

5.1 Fieldwork

Several field survey rounds were organised in 2008 and 2009 (Fig. 6). The island of Ruissalo was visited both during August 2008 and August 2009 while observations for Lenholm as well as for six other islands in southern Parainen were made during September 2009. By visiting Ruissalo at near-anniversary dates (17.8.2008 and 7.8.2009), the comparison of the observations for these two dates was facilitated (Coppin *et al*. 2004). This is to say that any site-specific land cover changes or annual fluctuations in phenology could be more easily noted, leading to a better understanding of biogeographical developments in the archipelago and – ultimately – to a better interpretation of the selected satellite images.

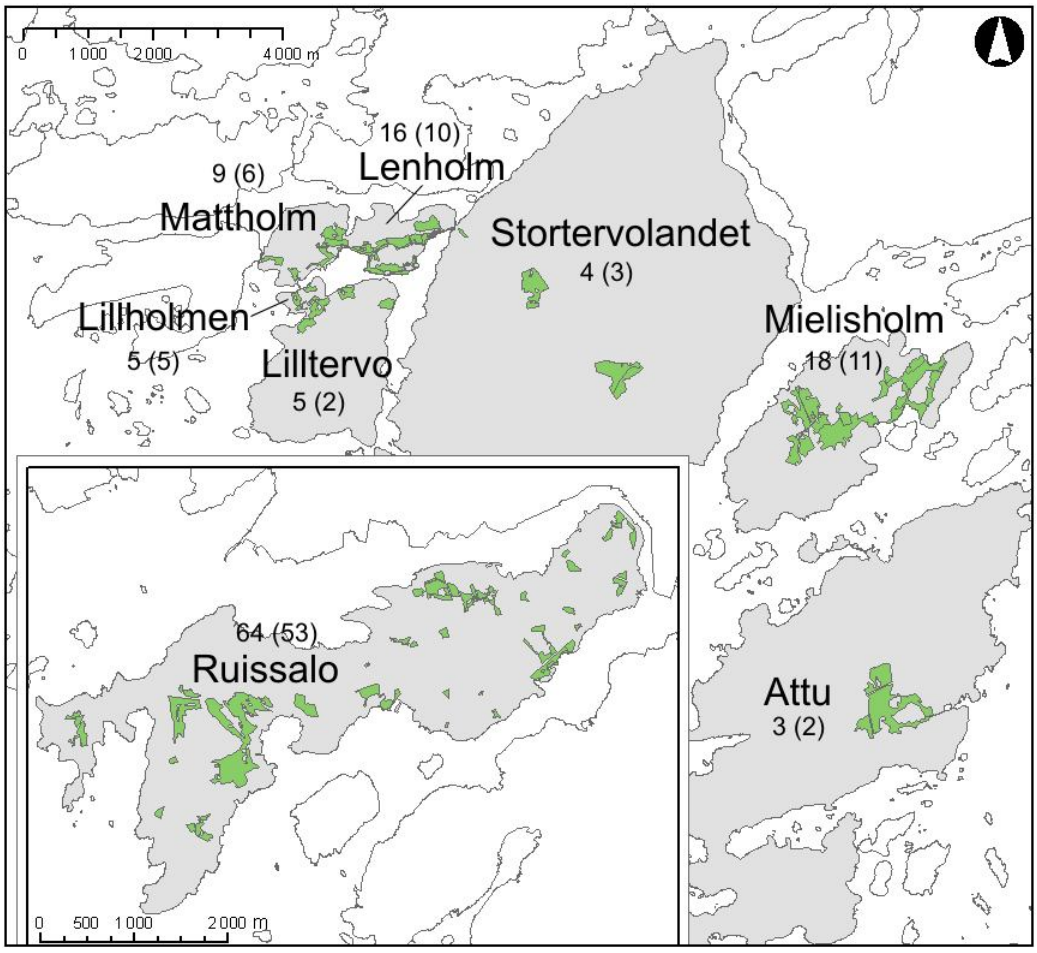


Figure 6. Cartographic representation fieldwork sites in the Turku. The figures below or above the name labels refer to the total number of patch entries into the field observations database for each island, of which the number of those considered grassland is given between brackets.

During field visits, observations were noted for all parcels considered *grassland* as well as for those that could be confused with grassland when interpreting spectral recordings. Grassland was considered to be any open and relatively homogenous grass-dominated type of land cover where grass domination is determined by the relative presence of gramineous species in relation to other occurring types of vegetation. Residential gardens as well as grasslands patches smaller than 0.1 ha were not included, although there are a few exceptions to that rule. Field visits were aided by printed aerial photographs overlain with vector data from the Topographic Database (NLS 2007) depicting field and meadow occurrence to improve ease of navigation, completeness of the observations and patch boundary sketching for later reference. Moreover, it allowed for especially large grassland units the study area to be identified and visited, such as was the case for the islands of Stortervolandet, Mielisholm and Attu. A handheld GPS device was operated in order to collect more precise spatial reference data necessary to process field observations.

After each survey round, all gathered information was entered into a database. Using the orthophotographs for Ruissalo and Lenholm as well as Landsat imagery for the outlying sites, boundaries were drawn around all patches of interest to the study occurrences of interest (both grassland and non-grassland), after which attribute information was attached. For each treated patch, the following information was recorded:

- Location specifications, especially of patch extremes
- Compliance with the definition of grassland
- Vegetation cover (plant species, vegetation proportions)
- Patch openness (occurrence of shrubs and trees)
- Colour (vegetation state, exposed soil or rock masses)
- Presence of herbivory by domestic grazing stock
- Signs of biomass control
- Further remarks

A total of 124 patch entries were registered, 92 of which were grassland. When considering total area and patch shape complexity, this corresponds to an estimated 1 600 - 1 700 pure or near-pure Landsat image pixels depicting grassland. Along the same lines of logic, field survey information was collected for approx. 700 pure or near-pure pixels depicting types of non-grassland land cover potentially relevant to the study. The technique for deriving these Landsat pixel equivalents is explained later, but it is important to understand the size

of the set of pixels that allowed for spectral information to be extracted based on *in situ* observations. Table 4 below presents the statistics for the surveyed grassland patches.

Table 4. Spatial statistics for database patch entries considered grassland.

Variable	Value
Total of entries	92
Minimum size	0.07 ha
Maximum size	32.46 ha
Mean size	2.46 ha

5.2 Compiling a study area-wide grassland vector set

The methodology for the present study consists of two main parts: the assessment of the spatial separability of grasslands and the assessment of their spectral separability. For the first part – the part dealing with the spatial separability of grasslands – a comprehensive grassland vector dataset consistently representing grassland occurrences throughout the study area had to be compiled. The methodological steps involved in this compilation are presented by the left half of Figure 7.

As mentioned earlier, no single complete grassland occurrence vector layer is available. In order to analyse the spatial characteristics of grasslands in the whole of the study area, the field parcel vector data from the Ministry of Agriculture were processed to suit the research aims of this study. After they were linked to the official field cover data for 2007, only parcel vectors fully within the study area were retained. Although this introduces a bias in the statistics on the total coverage of grassland, it insures that the spatial metrics derived from the data describe *actual* grassland patch characteristics rather than those of a mixture of actual grassland patch shapes and patch shapes partially manipulated by spatial subsetting.

Once only whole patches had been retained, a selection was made based on reported field cover for 2007. This was a somewhat challenging task in that the land cover classification scheme of the Ministry of Agriculture embraces an agricultural production-centred classification procedure. In order to produce a grassland vector layer compatible with the aims of this study, the definition of grassland used during the fieldwork was resorted to. Out of all agricultural fields, only those constituting open habitats dominated by gramineous species were considered suitable for inclusion. Whereas an open grazed

meadow for instance undeniably comprises a patch of *grassland*, a field of sown clover for soil properties improvement purposes does not, despite its resemblance to a rich meadow.

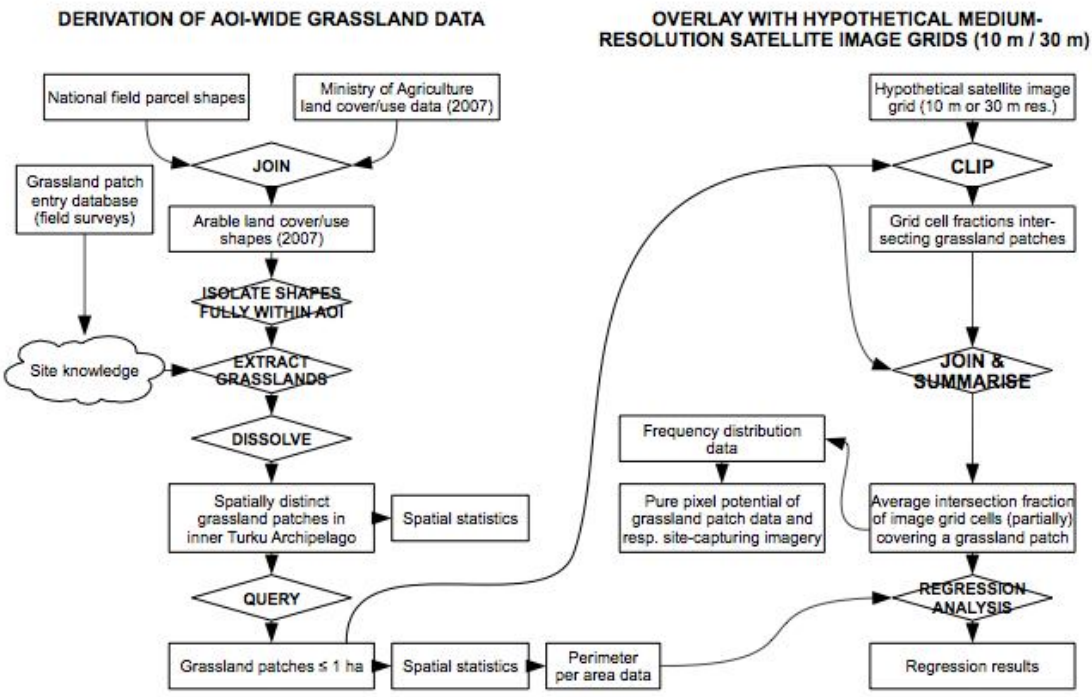


Figure 7. Schematic representation of the methodology for the spatial separability analysis

Table 5 lists the field cover classes that occurred in the study area in 2007 and were considered grassland. This list of field use types was compiled based on the inspection of the semantics of the available type descriptions in relation to known conditions out in the field. Use type 9720 (*managed uncultivated field*) required special consideration in that – based on its textual description – it does not guarantee that field parcels allocated this code are grasslands *per se*. Careful comparison with the collected field observations showed that despite the inclusion of various types of land cover other than grasslands, grassland cover prevails in this class. Moreover, a substantial portion of the sites known to be grassland belonged to this class. This observation is believed to relate to the continued field abandonment typical for the study area whereby sites are still managed for maintenance reasons, but the agricultural production potential is depreciated. By excluding the managed uncultivated fields, the completeness of the study area-wide grassland vector layer would be notably compromised. This was considered a more detrimental effect on the data quality than consciously introducing a number of field vectors possibly constituting a field type other than grassland. Managed, uncultivated fields were therefore included in the final grassland vector set.

Table 5. Classes of the Ministry of Agriculture agricultural field use data considered *grassland*, based on the reported field cover in 2007 for parcels fully located within the study area.

Code	Official description (Finnish)	Interpreted land cover type
6050	Viherlannoitusnurmi	Green manure
6111	1-vuotiset kuivaheinä-,säilörehu-,tuorerehunurmet	Annual dry hay, preserved fodder or fresh fodder
6112	1-vuotiset laidunnurmet	Annual pasture
6121	Monivuot. kuivaheinä-,säilörehu-ja tuorerehunurmet	Multiyear dry hay, preserved fodder or fresh fodder
6122	Monivuotiset laidunnurmet	Multiyear pasture
6123	Monivuotiset siemennurmet	Multiyear grass seed production
6210	Pysyvä kuivah.,säilör., tuorer. (väh 5, alle10 v)	Permanent dry hay, preserved fodder or fresh fodder (5-10 yr)
6220	Pysyvä laidunnurmi (väh 5, alle 10 v)	Permanent pasture (5-10 yr)
6300	Luonnonlaidun ja -niitty	Natural pasture or meadow
6545	Englannin raiheinän siemen, valvottu tuotanto	Controlled seed production of English ryegrass (*Lolium perenne*)
6550	Ruokonadan siemen, valvottu tuotanto	Controlled seed production of Tall fescue (*Festuca arundinacea*)
6562	Timotein siemen, valvottu tuotanto	Controlled seed production of Timothy-grass (*Phleum pratense*)
6565	Nurminadan siemen, valvottu tuotanto	Controlled seed production of Meadow fescue (*Festuca pratensis*)
6710	Hakamaa, avoin	Traditional pasture with rocky outcrops, open
9720	Hoidettu viljelemätön pelto	Managed, uncultivated field
9801	Erityistukisopimusala, pysyvä laidun	Parcel under special agreement, permanent pasture
9810	Suojavyöhykenurmi	Buffer belt sward
9820	Suojakaista	Buffer strip

The resulting grassland patch layer contained 2 239 spatially distinct grassland patches fully located within the study area. Table 6 lists their spatial statistics. Although the average patch size is 2.00 ha, the majority (52.1%) of patches are smaller than a single hectare. Almost 90% of all patches are under 5 ha in size. Patches occur roughly evenly spread throughout the area and generally correspond well to known occurrences of grassland in the test areas (Ruissalo and SW Parainen). A number of shortcomings however could be noted, such as the omission of certain grassland patches for various reasons. Grassland occurrences with a purposely installed recreational use, infrastructure grasslands or grasslands that are left to overgrow do not show in the dataset.

Table 6. Spatial characteristics for all grassland patches that fall entirely within the study area based on a selection of the official field parcel data.

	Area ***A*** **(ha)**	Perimeter ***P*** **(m)**	***P*** **/** ***A*** **(m/ha)**
Minimum	0.03	79.65	99.48
Maximum	45.67	7 350.21	5 368.75
Sum	4 481.89	1 556 379.55	-
Mean	2.00	695.12	633.82

Despite being incomplete to some extent, the grassland patch data layer for the inner Turku Archipelago was considered sufficiently representative of grassland spatial characteristics in the study area. However, in order to place special emphasis on the

grassland patches most vulnerable to spatial resolution limitations of satellite imagery when being captured, only the discrete patches with a size of 1 hectare or less were used for further analysis. This was done by dissolving all adjacent patches after which the area for each self-standing patch unit was calculated. Often, larger grassland units are composed of several smaller units that share one or more boundaries. Although administratively treated like individual units, these smaller units were collectively treated as a continuous patch in the present study. After dissolving adjacent grassland patches, the grassland patch layer held a total of 1 358 patch entries. Their spatial characteristics are presented in Table 7.

Table 7. Spatial characteristics for study area-wide grassland patch vectors after adjacent patches were merged and those resulting patches larger than 1 hectare excluded.

	Area *A* (ha)	**Perimeter *P* (m)**	***P*/*A* (m/ha)**
Minimum	0.03	79.65	395.88
Maximum	1.00	1 581.15	5368.75
Sum	645.09	-	-
Mean	0.48	344.82	877.93
Standard deviation	0.26	151.91	411.60

5.3 Spatial separability assessment

5.3.1 *Differential satellite grid overlay*

Once the grassland patch layer was fit for analysis, it was subjected to the overlay of a hypothetical site-capturing satellite image grid with a cell size of 30 x 30 m, so as to make it correspond to Landsat TM imagery. Figure 8 below illustrates a typical situation of such overlay, drawing attention to the possibilities and the limitations of the spatial fit between the *in situ* patch orientation and dimensions and the attempted sensing from space using a regular grid with a random orientation. Patch B for instance, is characterised by an average grassland intersection fraction of 45.6%. The average fraction of a certain patch is the mean of the intersecting fractions of the grid cells intersecting the patch and is assumed to provide a measure of the suitability of a certain type of imagery to successfully discriminate distinct patches. The higher the average fraction, the more pure pixels are derived from the given patches, the more successfully the imagery discerns these patches in spatial terms. So, in the hypothetical case presented in Figure 8, the average intersection fraction of the 16 grid cells intersecting patch B is 45.6% of the cell area. When comparing patch B to patch C, the latter is of notably smaller size, resulting in a reduced likelihood of pure pixels

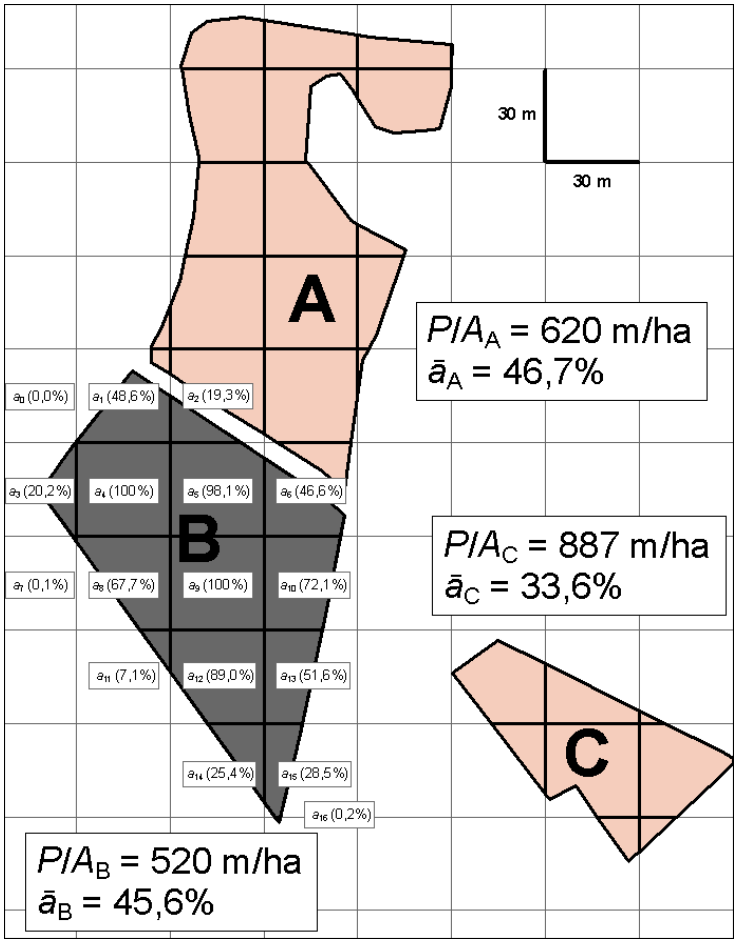


Figure 8. Schematic representation of a capture of typical patches using a 30 x 30 m resolution satellite sensor. For grassland patches A, B and C, spatial statistics are listed and include a measure of shape complexity (perimeter (*P*) per area (*A*)) and the average fraction ($\bar{a}$) of a 900 m^2 pixel intersecting the respective grassland patch.

forming within it and therefore also an average fraction (33.6%) inferior to that of patch C. After testing the capacity of Landsat TM 30 m resolution imagery to spatially discern the grassland patches in the inner Turku Archipelago, a hypothetical improvement in spatialresolution of the sensor was tested. The spatial resolution of the rectangular grid that represents the remote sensing image capturing the site was upscaled from 30 m to 10 m, corresponding to the resolution of coarse SPOT imagery. Upscaling grid cell size to 10 m however resulted in an inability of the software to successfully carry out the instructed computations and called for a reduction in number of input grassland patches. A selection of 74 representative patches occurring along a southwest-to-northeast axis in the centre of the study area allowed for the necessary computations to be carried out within a feasible time span (*ca.* 8 hrs). The spatial characteristics for this reduced data layer are given in Table 8 below. Note that irrelevant data were left out and that the total number of patch entries is 74. The spatial characteristics given in this table are generally similar to those for

the complete vector set of grassland patches smaller than 1 hectare (cf. Table 7 p.43), indicating the selection of grassland occurrences for the 10 m resolution grid overlay is a representative sample.

Table 8. Spatial characteristics for study area-wide grassland patch vectors for the 10 m grid overlay based on official field parcel data from the Finnish Ministry of Agriculture.

	Area *A* (ha)	**Perimeter *P* (m)**	***P*/*A* (m/ha)**
Minimum	0.05	87.29	406.21
Maximum	0.99	693.15	2035.71
Sum	38.68	-	-
Mean	0.52	350.09	838.03
Standard deviation	0.27	122.98	383.30

5.3.2 *Pure pixel potential of grassland patches*

Further, the relationship between perimeter/area ratio and the *pure pixel potential* of the respective patches was investigated using regression analysis. As mentioned earlier, the pure pixel potential of a patch is considered the extent to which the spatial properties of a given patch allow for pure pixels to form within it. Given the complexity and processing-intensiveness of the grid overlay procedure for assessing the spatial separability potential of 10 m and 30 m resolution imagery, a more efficient method for assessing the suitability of imagery is desirable. As the perimeter/area ratio is an established metric for describing shape complexity (Riitters *et al*. 1995), this part of the spatial separability assessment of the present study addresses the question as to how the *P*/*A* ratio of patches can be used to predict the suitability of 30 m resolution satellite imagery for capturing these patches if a sufficient amount of pure or near-pure pixels is to be retrieved.

In order to do this, the grassland feature layer representing grassland patches smaller than 1 hectare (Ministry of Agriculture data) was overlain with a hypothetical 30 m grid representing a satellite image capturing the scene. Once overlain, only grid cells intersecting the grassland patches were retained, populating them with spatial statistics treating the ratio of overlap with the respective grassland patches. This allowed for an average percentage of grid cell overlap to be allocated to each respective grassland patch (cf. Fig. 8). In other words, each grassland patch was supplied with a figure representing average overlap ratio of the grid cells intersecting it partially or fully. Secondly, the perimeter per area ratio was computed for all grassland patches as a measure of how effective patch edge encloses patch area and – consequently – of how complex the patch shape is. These two variable where

then subjected to regression analysis to test whether patch-specific perimeter/area data can be used as a predictor for pure pixel retrieval success.

5.4 Histogram matching of satellite imagery

Although TM imagery consists of seven spectral bands, only six were included in the input data used throughout the present study. The thermal energy band has a coarser resolution (120 m) and is therefore considered of no additional value to the spectral analysis. Whenever there is mention to *all spectral bands* or alike in the remainder of this work, only Thematic Mapper bands 1, 2, 3, 4, 5 and 7 are referred to.

Out of the three TM images (1.5.2007 (T1), 17.7.2006 (T2) and 21.8.2007 (T3)) selected to cover the crucial stages of the growing season (early, middle and late stages resp.), only two images were subjected to further analysis. In the first phase of the spectral study of grasslands in the Turku Archipelago, these three images were rendered comparable by bringing them to a common illumination level while substantial attention was paid to the accentuation of grassland occurrence and composition. Figure 9 presents the methodological steps involved in the selection of image scenes T1, T2 and T3 as well as further methodological steps the imagery underwent part of the spectral separability assessment.

Images T1, T2 and T3 showed strong radiometric variation. This can be attributed to a combination of atmospheric scattering, varying levels of illumination, cloudiness and resulting occurrence of shadow. While illumination is mostly dependent on sun angle, atmospheric scattering results from multiple interactions between light rays and the gases and the particles of the atmosphere (Sabins 1997). To correct for these inconsistencies, both T2 and T3 had to undergo a normalisation procedure relative to image T1. This was done through histogram matching. Histogram matching is a procedure in which the histogram of one image is transformed to match that of another. This technique is useful to compensate for differences in illumination or atmospheric effects for scenes scanned on separate dates (ERDAS 2005).

However, before proceeding with histogram matching, a number of conditions had to be complied with. Relative dark and light features in the image needed to be the same, while

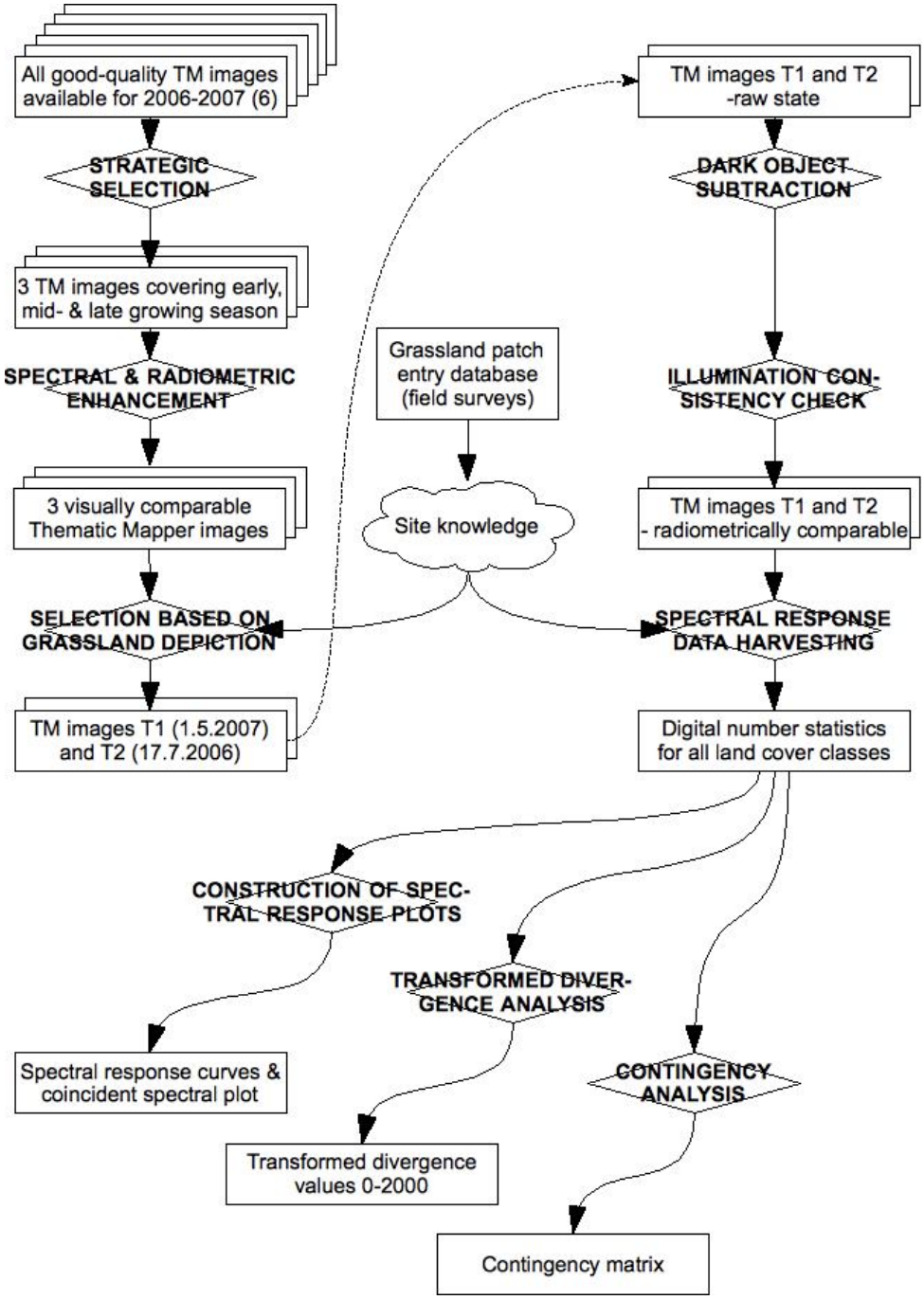


Figure 9. Schematic representation of the methodology for the spectral separability analysis

the relative distributions of different land cover types was to be comparable as well. Furthermore, the general shape of the histogram curves for the input images needed to be similar. Since the study site here comprises a relatively small island for which imagery was collected over a two year time period (2006-2007), these conditions were assumed to be respected. Figure 10 below shows the histograms for band 4 of all three images to illustrate the similarity in overall histogram shape. The bimodal distribution is composed of two

peaks; a taller peak representing water and a lower but wider peak depicting living biomass. This pattern persists throughout the imagery, also for the other TM bands (not shown). Note that the scale of the presented histograms is image-dependent (cf. image caption) and that therefore no observations other than those about the overall distribution shape can be deducted.

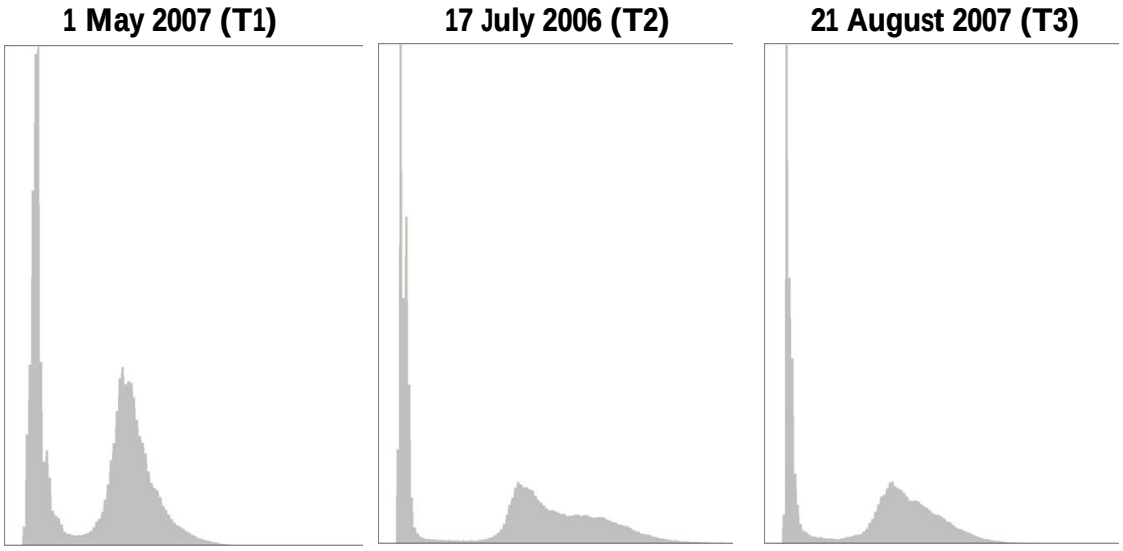


Figure 10. Histograms for band 4 for all three images showing consistent bimodality, indicating that – at least based on reflectance in the near-infrared – the input images are eligible for histogram matching. Note that the scale of the histograms presented here is image-specific (maximum y-values: T1: 861 696; T2: 1 309 856; T3: 1 591 472)

Since there appeared to be good compliance with all conditions for carrying out histogram matching, the procedure was proceeded with. Since any interpretation of phenological changes across the growing season requires good knowledge of site conditions, only the island of Ruissalo was subject to the histogram matching procedure. Image T3 however, required special attention in that a cloud shadow was cast on the western section of the island, blurring the transmission of ground cover data for that area. In order to resolve this, the affected zone was simply isolated and histogram matched separately. Although not seamless, the result of the merged parts reveals more about the biogeographical conditions at the end of August then the original, non-histogram matched image.

It needs to be noted that both 2006 and 2007 were exceptionally warm years. The year 2006 was typified by an exceptionally long, dry and hot summer (FMI 2007), while 2007 was unusual in that record high temperatures were measured in March and that the thermal spring starting a month earlier than on average (FMI 2008). This implies that although

perhaps the phenological conditions for 2006 and 2007 are mutually comparable, care needs to be exerted when extrapolating observations to other years.

5.5 Image nomination through spectral optimisation

Once the three time slices (T1, T2 and T3) had been corrected for variance in illumination and atmospheric conditions, the attention was turned to optimising the representation of grassland occurrences so as to best allow the analyst to visually appreciate the information content of each image in that sense. Image T1 (early May) was selected as reference image since illumination and cloudiness conditions were near-ideal at the time of capture. Furthermore, a quick comparison of the image with gathered field data revealed that there is a strong correlation in grassland occurrence.

Previously, Price *et al*. (2002) showed through applying stepwise discriminant analysis for six grassland types in eastern Kansas that the best band combinations include TM bands 3, 4, 5 or 7 or a combination thereof, depending on the capture dates of the input imagery. Bands 4 and 5 were found to be among the most useful bands for six out of the seven image(s) (combinations) they tested. Parallel to the findings of Price *et al*., bands 4, 5 and 7 were selected for the final visualisation.

In terms of hue-to-band allocation, a visualisation scheme that allows easy interpretation through improved feature recognisability was targeted. When selecting RGB colour combination 547 (*i.e.* red = TM5, green = TM4 and blue = TM7), grasslands appear as light green to light yellow, depending on conditions. Furthermore, coniferous forest appears dark green, which also constitutes a real-life approximation. Bare soil and built-up areas on the other hand show as pink to purple, providing sufficient amounts of contrast with vegetated areas. Figure 11 presents the resulting image scenes.

Out of the three histogram-matched satellite images, early growing season image T1 and peak greenness image T2 were selected. The reason for this choice pertains to the excellent grassland portrayal properties of image T1 and the key phenological conditions that image T2 describes (i.e. peak greenness). After selecting these two images, the spectral response analysis of these two remaining images was proceeded with. Histogram-matching however, is a relatively aggressive technique that *forces* pixel values to alter so as to better suit those in the reference image. In the next part of the study, the main idea was to obtain insight in the

subtle changes between grassland types, over time and over composition or management regime. Although histogram matching corrects for differences in illumination, atmospheric scattering and perhaps even sensor inconsistencies, it has the potential to modify pixel attributes to the extent that they no longer contain the information necessary to deduct conclusions with regard to spectral discrimination.

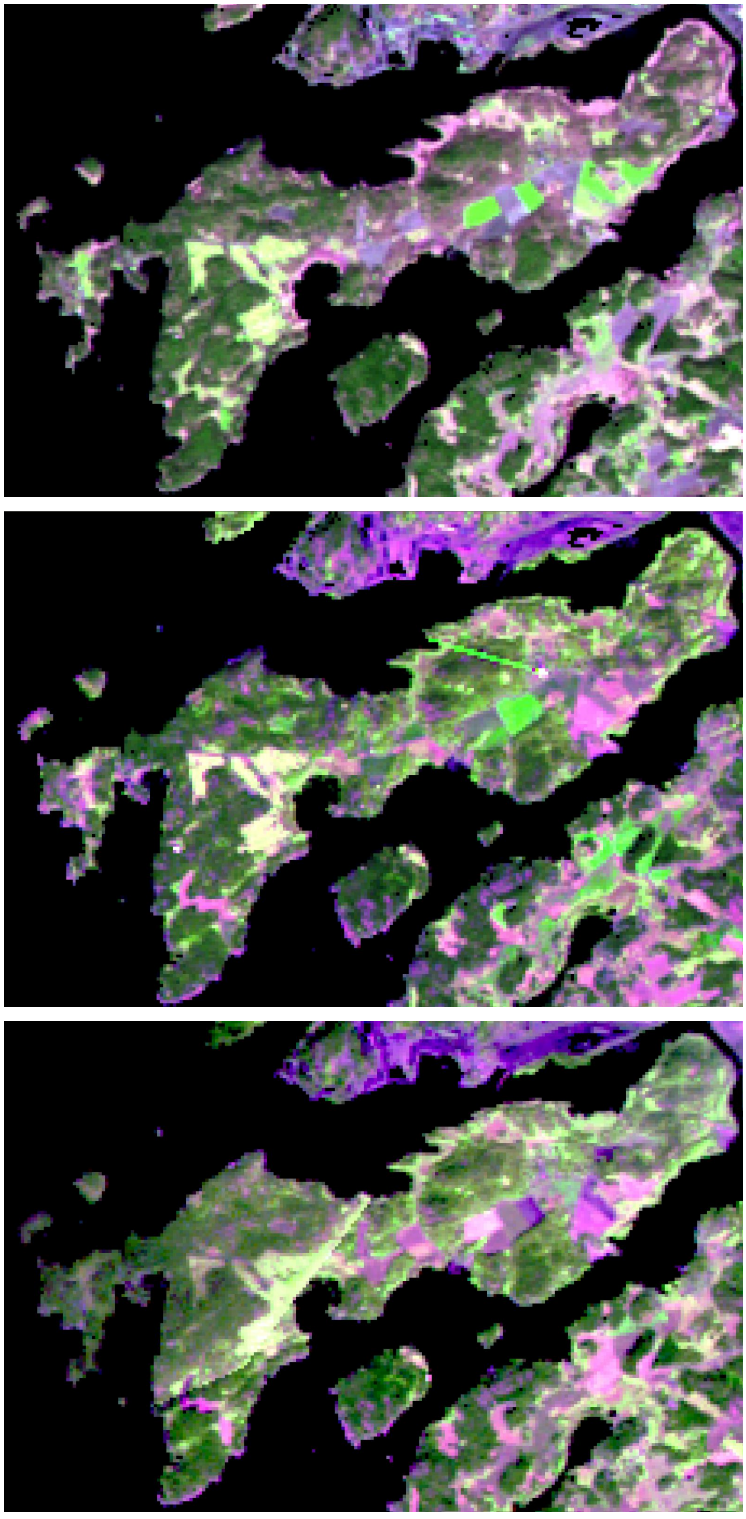


Figure 11. Locally radiometrically normalised images (Ruissalo)

5.6 Radiometric normalisation

In the light of the concerns regarding histogram matching, a different methodology needed to be adapted if radiometric corrections were to be executed to render the two selected images comparable without disrupting the original digital numbers to the extent crucial information is lost. This was done through assessing the need for correcting for atmospheric scattering and for illumination variance separately, after which only the vital adjustments were made to the raw imagery material.

5.6.1 *Simple dark-object subtraction*

As mentioned earlier, the atmosphere often has a degrading effect on the quality of satellite imagery. It does so by scattering, absorbing and refracting light but the effect of scattering is dominant (Slater *et al*.1983). Atmospheric scattering – often called haze – is an additive component to satellite imagery data in since it gives rise to higher reflectance values than would be observed at ground level. The effect is highly wavelength dependent, especially degrading readings for the shorter wavelengths of UV energy and blue light, often referred to as *selective scattering* (Sabins 1997). In Thematic Mapper band 1 (blue light), selective scattering may lead to up to 70% contribution to the pixel values (Infoterra 2009). Additionally, a process of non-selective scattering affects all wavelengths equally when particles larger than the wavelengths of light are present in the atmosphere, such as dust, clouds and fog (Sabins 1997).

Both selective and non-selective atmospheric scattering can be corrected for using the *simple dark-object subtraction method* (Chavez 1988). This technique is especially useful when no data are available on the atmospheric conditions at the time of image capture, since it only requires the information contained by the satellite image itself. It is based on the assumption that of those thousands if not millions of pixels that make up a single satellite image, at least a few should show no reflection at all, be it because of shadow formations or complete absorption of the incident radiation, as is the case for a large clear lake for instance. By subjecting each pixel in the entire image to subtraction by the digital number (DN) that has been attributed to these dark and non-reflective objects, the effect of atmospheric scattering is removed. Since the interference of atmospheric scattering is wavelength-specific, this operation needs to be carried out for each spectral band and each satellite image separately. This of course assumes a constant haze value throughout the

image, which is often not the case. However, it does accomplish a first-order correction, which is better than no correction at all.

5.6.2 *Illumination consistency*

After atmospheric scattering had been removed from every spectral band, true spectral reflectance values could be extracted from the imagery. Since no histogram matching was undergone however, illumination differences between the images might still persist. Compared to the early growing season image (T1), the sun elevation angle at the centre point of the Landsat TM scene for 17 July 2006 (T2) is 5.6° larger (cf. Annex I), implying somewhat elevated illumination levels can be expected.

Before proceeding with the extraction of spectral response values for both images, the illumination consistency was assessed. This can be done by comparing the spectral reflectance values for land cover types that do not undergo phenological changes, such as rocky outcrops or certain types of human-built infrastructure, analogue to the technique of *pseudo-invariant normalisation* (Schott *et al*. 1988).

Schott's relative radiometric normalisation technique assumes the existence of pseudoinvariant objects in the landscape that do not show any significant change in reflectivity between two given dates. These pseudoinvariant objects are best extracted by assessing vegetation cover through image ratioing or vegetation indices. Typically, man-made structures such as roads, urban areas and industrial centres will constitute the most reliable point of reference, since they are unlikely to change over relatively long periods of time. Water bodies on the other hand are best avoided since they absorb most incident energy and therefore can provide little information about difference in illumination.

In order to ascertain limited illumination variability between the two atmospheric scattering-corrected images, digital numbers for total of 4 225 pixels depicting built-up area were extracted from both the early growing season and the mid-summer image. The training pixels used represented areas that consist of large surfaces dominated by concrete, tarmac and other artificial materials such as industrial zones and urban areas with limited presence of vegetative growth. Especially the centre of the city of Turku and the industrial zone formed by the ports of Turku and Naantali provided good ground for pixel extraction. The exact same pixels were used for both images.

5.7 Delineation of grassland classes

In order to distinguish between different classes of grassland, the abundance of both *feature classes* and *spectral* classes needs to be assessed (Lillesand & Kiefer 1994). Several distinct spectral classes may contribute a single feature class, such as a particular type of land use. In order to allow the collection of a sufficiently large set of pixels for each grassland type and to insure compatibility of the classification scheme with both the early growing season and the mid-summer image, grassland occurrences for which field observations were available were initially grouped into four management practice classes only: fodder production grassland, grazed grassland, recreational grassland and overgrown grassland.

Fodder production grassland (FP) is a highly productive type of grassland solely intended for the production of hay, silage or other form of storable biomass suitable for feeding to domestic animals. This type of grassland is generally a highly homogenous type of field cover, often characterised by idealised growth conditions (large open patches, nutrient-rich soils, productive grass species, etc.). Depending on conditions, fodder production fields are typically mown once or twice over the growing season.

Grazed grassland or pasture is quite simply grassland subject to periodical grazing, mostly by cattle, sheep and horses. In most cases, grazed grassland can be considered a type of semi-natural grassland. While extracting digital numbers for the grazed grasslands for the early growing season image (T1), a notable discrepancy in spectral signatures was observed between certain patches. Patches located near the shoreline or close to water bodies showed spectral characteristics distinct from those located more land-inward. Patches near shores are typically marked by limited productivity since the terrain is more easily inundated and less accessible for agricultural machinery to optimise site productivity. Stands of reed, shrub and tree species that thrive in mesic environments such as Black Alder (*Alnus glutinosa*) are common, while rocky outcrops and exposed soil may occur too. This type of grazed grasslands was considered *traditional pasture* (TP), and will be referred to as such in further writing. As opposed to traditional pasture, those grazed grasslands located in more productive environments were considered *modern pasture* (MP).

Aiming to accommodate the two-headed divergence of spectral signatures within the grazed grasslands, the total number of grassland classes participating in the spectral separability analysis was raised to five: fodder production grassland, traditional pasture,

modern pasture, recreational grassland and overgrown grassland. Even though seemingly all-encompassing, this grassland classification scheme if of course a highly simplified interpretation of reality. Each of these five classes is really the sum of a number of sub-classes while other occurring types of grassland such as infrastructure grassland – be it in very small quantities – are not accounted for altogether. However, the final land use and land cover classes are always based on compromises in the classification process (Mikkola *et al*. 1999).

Recreational grassland (RC) groups all grasslands used for the recreational purposes. Such use virtually always implies regular to very regular mowing, keeping the grass at a suitable length for recreational activity practiced on site. Besides an intensive management regime, recreational grasslands are typically landscaped to some extent, with individual trees and other amenity features being actively installed or conserved. Within the study area, the most important recreational grasslands are golf courses, although playgrounds and parklands also occur.

Overgrown grasslands (OG), the final group of grasslands, are those environments in which the occurrence of grass species still dominates the vegetation composition to date, but where clear signs of management abandonment or vegetative encroachment exist. Two scenarios can be discriminated. The first type of overgrown grassland is the result of the cessation of biomass harvesting on established fodder production sites. Typically, these are large fields where a competitive grass species has been sown to maximise production. When no longer mown, dead grasses accumulate and deteriorate growing conditions for young plants. The second type of overgrown grassland consists of much smaller patches, often located near or within forested zones. Given patch size, soil type or other conditions inhibiting the potential for high-productivity agriculture, these grasslands were often used for animal husbandry. Once biomass control ceases, rapid invasion of the site by herbaceous plants, shrubs and tree saplings is incurred by the relative proximity of existing stands of these species.

5.8 Extraction of digital numbers

Once a grassland classification system compatible with conditions in the study area at the times of image capture had been developed, the process of extracting digital numbers for each class from the TM imagery was initiated. Only grassland patches visited during the

field survey rounds were considered potential training sites so as to maximise the confidence level associated with the derived spectral response curves. Each of the patches was carefully examined on its compliance with recorded field observations and its representativeness of the grassland class it belonged to as a whole. When a given patch was found suitable, digital numbers (DN) were harvested from its pixels in bands 1-3 (visible light), 4 (near-infrared), 5 and 7 (mid-infrared) in the radiometrically normalised TM imagery.

The extraction technique was dependent on the image and the spatial characteristics of the grassland patches. The basic idea was to obtain a representative sample for each grassland type. This was done by carefully assessing the suitability of each patch entry and the pixels therein while ensuring a large enough sample was collected in terms of total number of pixels. Because there is a temporal discrepancy between the different input data and field observations, the possibility that site conditions had changed since the time of image capture was kept in mind at all times. Using all available information, an understanding of the typical spectral characteristics of each grassland type was always established before proceeding with digital number extraction. This allowed the identification of sites that showed deviating spectral response, which were consequently left out of the sample in order to ensure purity and representativeness of the spectral response.

Often however, the extraction process was hampered by limited patch size. In order to obtain pure or near-pure spectral response data, mixed edge pixels were avoided. When a patch is only a few pixels in size, this generally implied that the concerned patch had to be left out of the sampling exercise altogether. This problem is not uncommon. Askew and Slate (1995) excluded boundary pixels when extracting pixel values from the sample fields in TM imagery part of their study of northern British grasslands too. With a mean field size in the area amounting to 1.8 ha, only 171 out of 316 fields were large enough for inclusion.

In total, 1 951 pixels delivered spectral response information for image T1 and 2 031 for T2, each class being represented by an average of 325.2 and 338.5 pixels respectively. When creating training statistics to be interpreted by a statistics-based classifier such as the maximum likelihood method, a minimum of from 60 to 600 pixels per training is recommended for six-band imagery (Lillesand and Kiefer 1994). Table 9 below lists the number of pixel entries for each grassland class for each image, showing consistent compliance with this guideline. Spectral data were harvested almost exclusively for

grassland patches visited during field survey rounds in 2008-2009. Only one set of pixels was involved without *a priori* knowledge on the exact site conditions: the Harjattula golf course on the island of Kakskerta. Auxiliary data such as aerial photographs and online documentation confirmed that the concerned site can be considered recreational grassland, similar to the golf course of Ruissalo. Given the size of both golf courses, a wealth of pixels for spectral response analysis was at hand, leading to a somewhat elevated number of total pixel samples for characterising recreational grassland based on both the early growing season and the mid-summer image.

Table 9. Number of source pixels sampled for the construction of spectral response graphs for five grassland types.

Grassland types	**1 May 2007**	**17 July 2006**
Modern pasture	398	438
Traditional pasture	444	425
Fodder production grassland	285	362
Recreational grassland	565	582
Overgrown grassland	259	224

Source patches of pixels belong to the same spectral class were created both using the seed pixel approach (Lillesand and Kiefer 1994) and manual patch delineation using simple geometric shapes, depending on circumstances. Given a number of coordinate system interoperability constraints, no sufficiently geometrically accurate overlay between the satellite images and the grassland patch vectors from the field survey database was achieved. Although restricting at first sight, this limitation forced the analyst to more carefully assess spectral responses in the imagery, giving way to a better grassland response extraction as impure and non-compliant pixels were more readily left out. In many ways, the training effort is both an art and a science (Lillesand and Kiefer 1994).

Not all types of grassland showed the same level of *response extractability*. Overgrown grassland patches for instance are inherently more likely to be small in size, meaning a substantial proportion of the imaging pixels will be of mixed consistency. Furthermore, although constituting one feature class (*i.e.* overgrown grassland), two distinct spectral classes can be discerned within it: the dead grass-dominated type and tree-encroachment type. Despite this duality in consistency, the number of overgrown grassland patches large enough for the extraction of representative digital numbers was too limited to pursue discrimination of these two sub-classes. This implies that the overgrown grassland class is

based on a slightly lower number of pixel DN extractions and that it describes both the more encroachment- and dead grass-dominated types over vegetative overgrowth.

5.9 Construction of spectral response curves

Once sufficiently large and representative spectral response data had been acquired for all five grassland classes, they were plotted in a variety of response graphs. First, *conventional response curves* were constructed using the average reflectance value for each class in each band only. Although showing general trends, conventional response curve does not necessarily provide information on *true spectral separability*, mainly because the different spectral classes typically show different levels of variance in DN. In order to fully assess separability, a combination of both the distance between class means and a measure of standard deviation is required (Richards & Jia 2006). This is why upper and lower standard deviation curves were included in a second stage, producing elongated graph areas collectively referred to as *standard deviation swath graph*. A serious disadvantage with presenting spectral response data using such swatch graphs was however noted in that these graphs reached great levels of complexity, even when simultaneously presenting data for five classes only. Visual assessment was especially challenging when spectral responses were similar or non-distinct. In order to overcome this, a different type of spectral response representation was constructed: a coincident spectral plot (Lillesand Kiefer 1994).

In order to produce continuous spectral response curves based on DNs extracted from low spectral resolution satellite imagery that is Thematic Mapper data, a number of intermediary steps had to be undergone. First, single wavelength values had to be associated with the spectral response data for each TM band. Satellite sensor bands are *sections* of the electromagnetic spectrum rather than discrete wavelength values, rendering it difficult to plot response data for each of these bands along the spectrum with the aim to produce continuous spectral curves. In this context, the extremes for each of the Landsat TM spectral bands were averaged. Table 10 below lists the wavelength sections for each band and the *central* value derived through averaging. Both the conventional spectral plots and the standard deviation swath graphs are based on these average band wavelength values.

Table 10. Thematic Mapper band ranges and derived average values. Source for EMS ranges: U.S. Geological Survey (USGS 2009d).

TM band	EMS range (μm)	Average (μm)
Band 1	0.45-0.52	0.485
Band 2	0.52-0.60	0.560
Band 3	0.63-0.69	0.660
Band 4	0.76-0.90	0.830
Band 5	1.55-1.75	1.650
Band 7	2.08-2.35	2.215

Once distinct wavelength values had been associated with all TM bands, their respective spectral response data were plotted along the electromagnetic spectrum. In order to further enhance the graphic representation, smoothed curves linking spectral response data from all bands for each grassland class were then constructed. Spectral response curves presented in the present study are thus approximations of the truly continuous spectral signature such as recorded by a laboratory spectrometer. Smoothed curves were preferred over linear interpolation for enhanced interpretability reasons since the true spectral response of features moves along the electromagnetic spectral in a polynomial fashion rather than with linear jumps.

As mentioned earlier, three different methods of graphically presenting spectral response were embraced: the conventional spectral response plot with average DN curves only, the standard deviation swath graph with ± 1 standard deviation curves for each average DN curve and the coincident spectral plot. All three graphic representations were constructed for TM image T1 (1.5.2007) whereas the response data for image T2 (17.7.2006) were only presented in the form of a conventional spectral plot.

The coincident spectral plot is really the standard deviation swath graph presented in a vertical manner, but contains a substantially larger amount of information while presenting it in a more readable format. It is constructed by placing a number of elongated plot areas underneath each other, each representing one spectral band. Subsequently, DN value ranges are plotted horizontally for each class and spectral band by means of straight line segments, the length and position of which is depends on the data (Lillesans & Kiefer 1994). The range of the DN values to be presented needs to be set in advance and is typically expressed in terms of one or more standard deviations away from the average response value. The more variance that a set of spectral data are characterised by, the larger the standard deviation value and the larger the respective line segment in the

coincident spectral plot. This technique naturally assumes a normal distribution of the data (Lillesand &Kiefer 1994). If the normality condition is complied with, ± 1 standard deviation of the mean ($\mu \pm 1\sigma$) ranges theoretically includes 68.3% of the data whereas $\mu \pm 2$ standard deviations include 95.4% of the data (Ghahramani 2005). In the present study, a DN range of $\mu \pm 2\sigma$ was considered suitable for presenting the spectral responses of the five grassland classes and the overlaps thereof. With a ± 2 standard deviations range, the vast majority of DNs are covered while values most likely to result from data interpretation errors (*i.e.* those outside the $\mu \pm 2\sigma$ range) are left out.

5.10 Quantitative spectral separability assessment

In addition to observations based on a number of graphical representations of the spectral response of grasslands, the retrieved spectral response data were subjected to two quantitative tests: transformed divergence analysis and contingency analysis. Both tests are based on the covariance and the mean vectors of the signatures and involve the maximum likelihood decision rule, helping to predict the results of a maximum likelihood classification (ERDAS 2005). Maximum likelihood is the most widely used traditional classifier (Luoto *et al*. 2002).

5.10.1 *Transformed divergence*

Divergence is a measure of the separability of a pair of probability distributions that has its basis in their degree of overlap (Richards & Jia 2006). The analysis of the divergence provides information regarding the relative degree to which land cover categories can be classified accurately (Mausel *et al*. 1990). The formula for divergence of spectral classes marked by normal data distribution is as below (Equation 1 p.60). Since this measure is not bound however, a non-linear relationship exists between observed classification accuracy and corresponding class divergence. This may have misleading consequences since it implies that an increase in class separation will always lead to better classification accuracy, even when these classes are already fully distinct. Naturally, this cannot be the case. In this light, a transformation has been applied to saturate the measure of divergence (Swain & Davis 1978, Equation 2 p.60). Calculating transformed divergence, values between 0 - 2 000 can be obtained, where 0 corresponds to non-separable and 2 000 to perfect separability. As a general rule, if the result is greater than 1 900, classes can be separated.

Between 1 700 and 1 900, the separation is fairly good whereas below 1 700, the separation is typically poor (Jensen 1996).

$$D_{ij} = \frac{1}{2}\mathrm{tr}((C_i - C_j)(C_i^{-1} - C_j^{-1})) + \frac{1}{2}\mathrm{tr}((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T) \quad (1)$$

$$TD_{ij} = 2000\left(1 - \exp\left(\frac{-D_{ij}}{8}\right)\right) \quad (2)$$

Where:

i and *j* = the two signatures (classes) being compared

C_i = the covariance matrix of signature *i*

μ_i = the mean vector of signature *i*

tr = the trace function (matrix algebra)

T = the transposition function

5.10.2 *Contingency analysis*

Following the calculation of the transform divergence values for the grassland classes based on a set of training sites, the pixels in these training sites were subjected to contingency analysis. This was done based on a somewhat altered but evenly representative set of training pixels, the statistics for which can be found in Table 11.

Table 11. Number of source pixels sampled for the construction of spectral response graphs for five grassland types.

Grassland type	**1 May 2007**	**17 July 2006**
Modern pasture	191	438
Traditional pasture	490	404
Fodder production grassland	382	362
Recreational grassland	582	582
Overgrown grassland	242	224

Contingency analysis is a preliminary assessment of the success of a supervised classification using the training site pixels and their statistics only. This allows predicting the accuracy of a classification for the entire dataset using the same set of spectral signatures. The principle behind this test is based on the assumption that the pixels of each

training sample are not always so homogeneous that every pixel of a sample is classified into its corresponding class. Even though each sample pixel weights the statistics that determine the classes, pixels may be placed in a different class if their spectral characteristics match those of that other class better (ERDAS 2005). Contingency therefore assesses how easily confused training pixels are despite careful inclusion in one particular spectral class by the analyst, providing a measure of spectral separability.

Supervised classification places the input pixels into the land cover class it considers most likely to correspond. The results of such classification can then be contrasted with the class membership of training pixels assumed when creating the training sites, producing a classification error matrix. Such a contingency matrix (*i.e.* error matrix) allows for a set of descriptive accuracy measures to be derived: overall accuracy, producer's accuracy (omission error) and user's accuracy (commission error). Using a contingency matrix to represent accuracy has been recommended by many researchers and should be adopted as the standard reporting convention (Congalton 1991).

Several options for conducting the contingency analysis classification were available. The main difference related to the use of a parametric or non-parametric rule for classifying the training pixels. The non-parametric rule configuration used a parallelepiped classifier to determine class membership for non-overlapping pixels. In case of overlap, a parametric rule (maximum likelihood) was involved to allocate a class to the respective pixels. For the parametric rule configuration on the other hand, a maximum likelihood algorithm was used for the classification of all input pixels.

Both options were tested and overall classification accuracies were calculated. Contingency matrices were subsequently constructed for both input images (T1 and T2), based on the classification procedure yielding the least overall error. Since the sample size for each training class has the potential to bias the results of the contingency analysis, a normalised sample size of 500 pixels for each grassland type was assumed.

The contingency matrices then gave rise to *mutual confusion matrices*, in which the discrimination success rate associated with each grassland class pair is given. This was achieved by summating the number of misclassified pixels for each class pair and computing their stake of the total number of misclassified pixels. Practically speaking, if the mutual confusion for grassland class A and grassland class B needs to be quantified, then it equals the sum of the pixels belonging to class A but classified as class B and those

belonging to class B but classified as class A. In other words, it is a measure for the total confusion related to a particular pair of classes.

In order to allow the appreciation of the spectral separability of each grassland class as a whole, all misclassified pixels associated with each class were summated, both for early growing season image T1 and mid-summer image T2. Using normalised values rather than percentages (as were calculated for the mutual confusion matrices), the classification success rate for each grassland class as a whole could be related to the temporality of image capture.

6 Results

6.1 Spatial separability assessment

6.1.1 *Differential satellite grid overlay*

The results for the comparison between a 30 m resolution grid and a 10 m resolution grid are presented in Table 12 below. Moving from a 30 m resolution grid (*e.g.* Landsat TM) to a 10 m resolution grid (e.g. SPOT) resulted in pure pixel ratio 8.4 times as high (5.1% to 42.7%). Note that this figure applies to grassland patches ≤ 1 ha only.

Table 12. Results for the resolution upscaling analysis of all grid cells intersecting with grassland patches ≤ 1 ha in the study area.

	30 m res. grid	**10 m res. grid**
Number of intersecting grid cells	18 287	5 566
Number of pure grid cells	932	2 375
Pure pixel ratio	5.1%	42.7%

Figure 12 below offers a more elaborate view on this outcome. The frequency distribution charts for both the grassland patch intersecting grid cells in the 30 m and the 10 m resolution grids show a dramatic shift in pure or near-pure pixel percentage. Whereas only 13% of the 30 x 30 m grid cells overlapped respective grasslands patches for 90% or more, this figure went up to 52% for the 10 x 10 m cells.

In Figure 12, the x-axis represents the fraction of a pixel that overlaps the respective grassland patch. Note that when improving the grid resolution from 30 m to 10 m, the percentage of pixels overlapping patches for more than 90% increases from 13% to 52% respectively. This implies that using 10 m resolution imagery for mapping grassland patches ≤ 1 ha in the archipelago increases the number of near-pure and pure pixels (> 90% purity) obtained by fourfold.

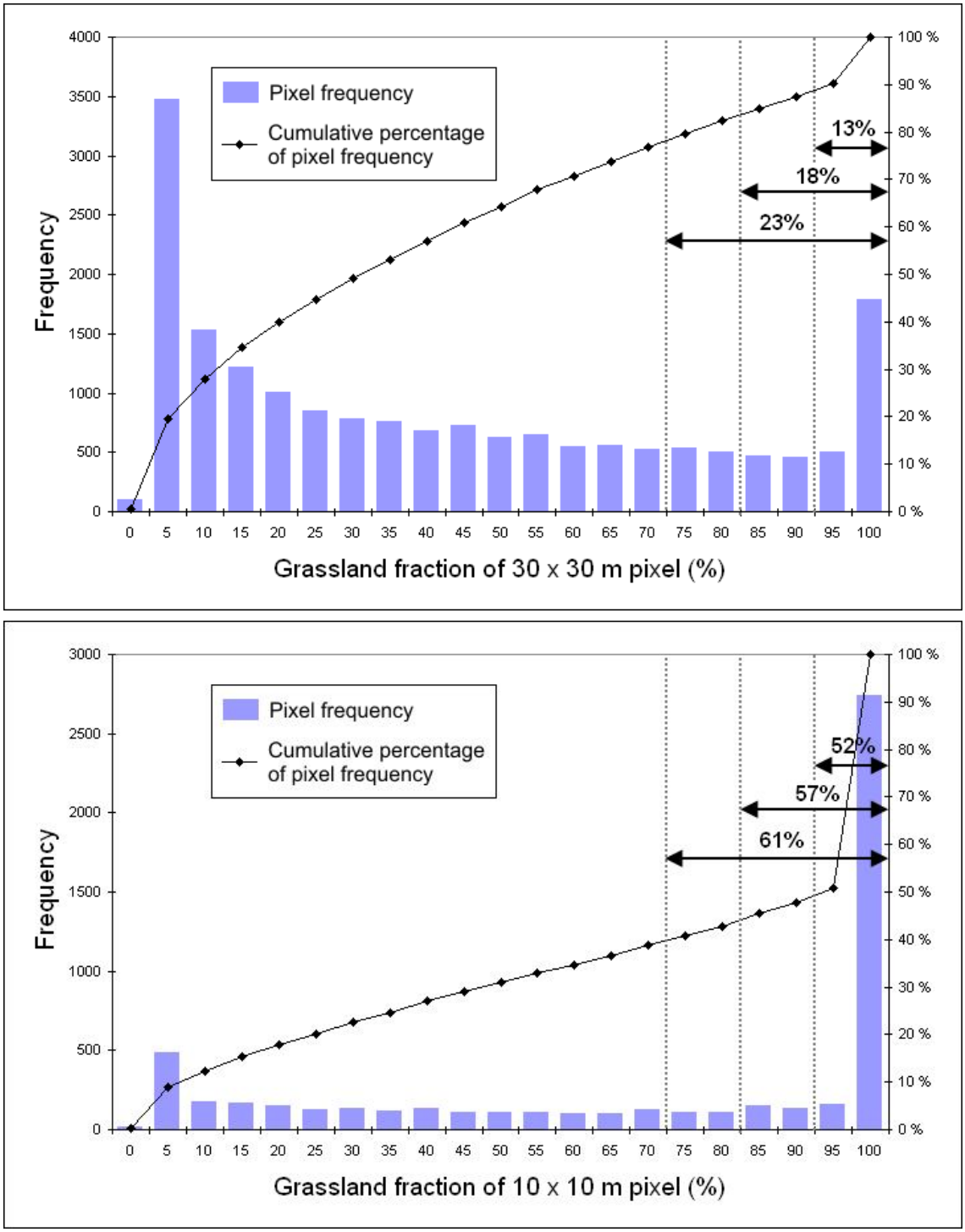


Figure 12. Frequency distribution diagrams for 30 m pixels (upper) and 10 m pixels (lower) intersecting grassland patches ≤ 1 ha.

6.1.2 *Pure pixel potential of grassland patches*

Figure 13 below presents the results of the regression analysis. According to the findings, the majority of grassland patches ≤ 1 ha has a perimeter/area ratio of roughly 500-1000 m/ha, corresponding to an average grid cell intersection of approx. 55% - 30% respectively. As the *P*/A ratio increases, the average grid cell intersection drops

exponentially. This implies that *pure pixel potential* is relatively quickly lost in the process of capturing archipelago grasslands ≤ 1 ha using 30 m resolution imagery as patch shapes become more complex. The majority of smaller archipelago grasslands however is concentrated around the relatively low perimeter per area ratio of 500 m/ha, meaning most patches ≤ 1 ha exhibit good pure pixel potential given their limited size.

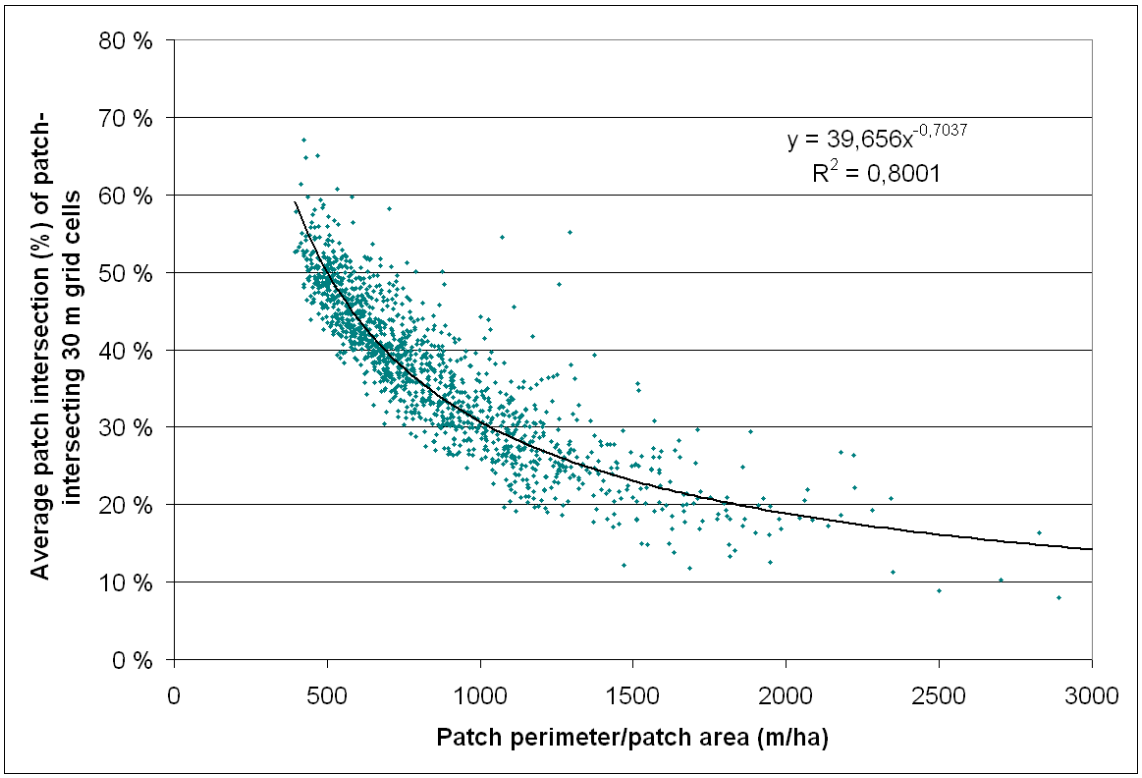
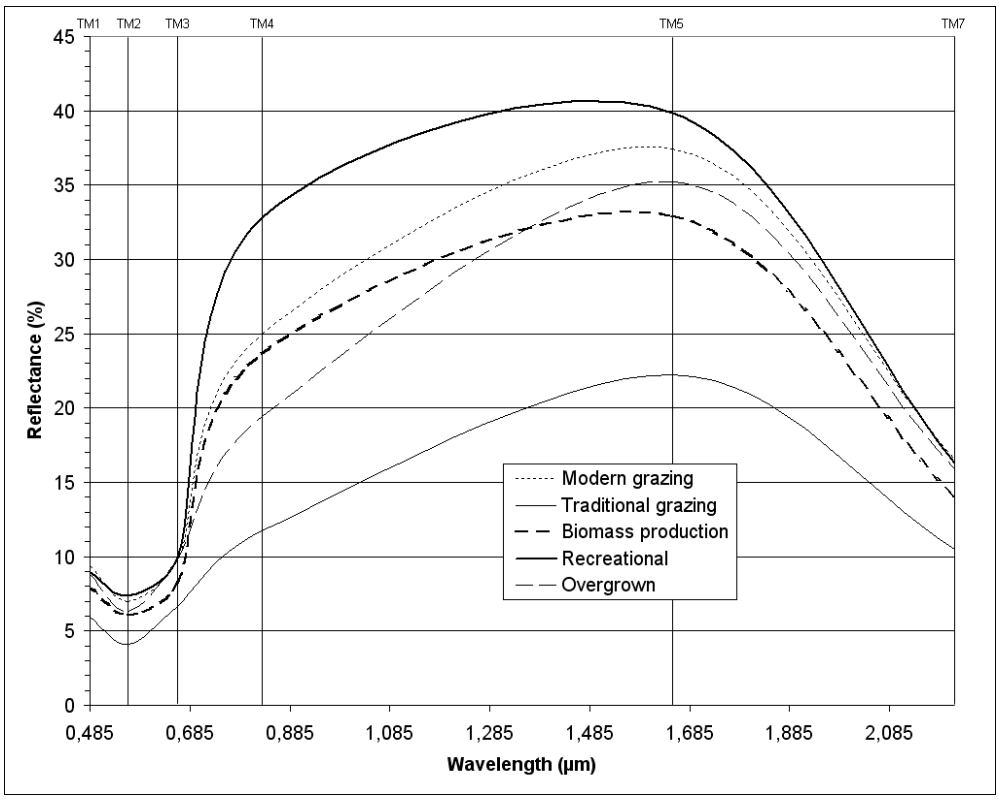


Figure 13. Regression analysis graph testing the relationship between perimeter/area and *pure pixel potential* for grassland patches in the inner Turku Archipelago. Data based on patches smaller than 1 hectare as listed in the official register of field parcels from the Finnish Ministry of Agriculture (n = 1 358). Note that the graphic representation was limited to patch complexity 3 000 m/ha. The highest shape complexity recorded was 5 193 m/ha (y = 9.2%).

6.2 Conventional response curves

6.2.1 *Response curves for T1 and T2*

Figure 14 below presents the spectral responses for five grassland classes based on averaged reflectance values extracted from early growing season image T1 (1.5.2007). These curves need to be interpreted carefully, because they are the result of an interpolation of distinct values for 6 TM bands rather than continuous spectral measurements along the electromagnetic spectrum.



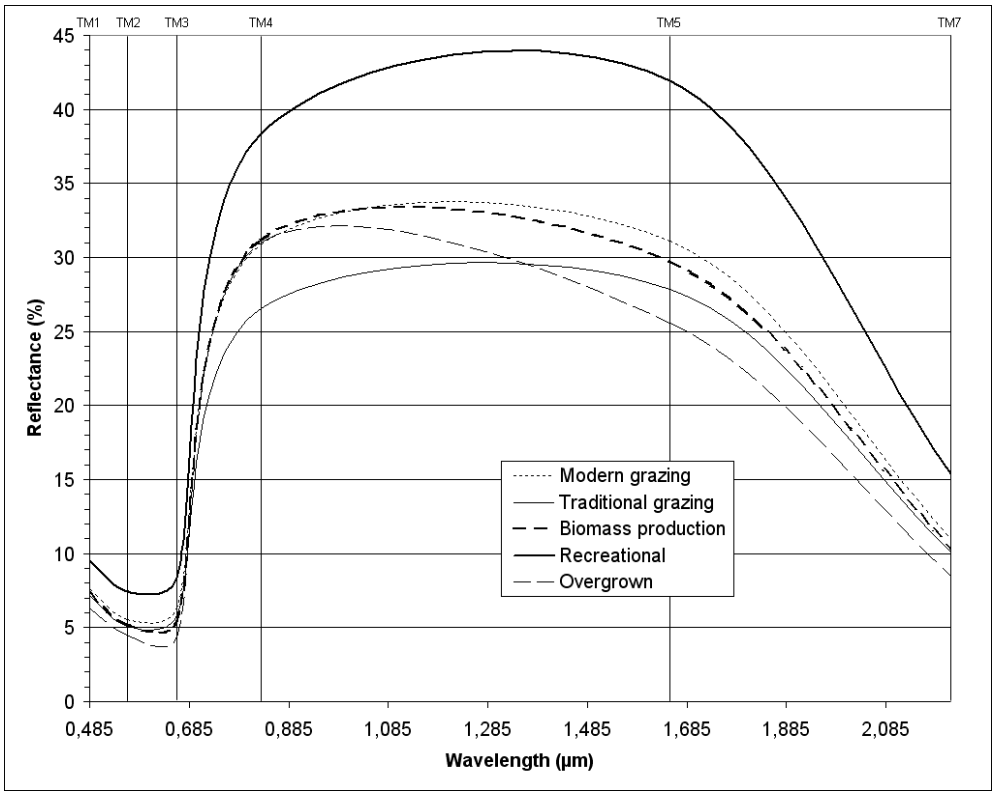


Figure 14. Spectral response curves for five grassland classes constructed from Thematic Mapper imagery for 1.5.2007 (T1) and 17.7.2006 (T2). Note that the vertical lines intersecting the curves represent the average wavelengths at which Landsat TM bands 1-5 and 7 sample spectral response.

Generally, reflectance values remain limited in the visible light bands, reaching a maximum of 10% reflectance. Subsequently, there is a dramatic increase in reflectance for most classes, also known as the *red edge* (Infoterra 2009). The increasing trend is sustained through the near-infrared to mid-infrared band TM 5 where it reaches maximum values between 20%-40% reflectance. Once past TM band 5, values show a rapid and more uniform decline, reaching values between 10%-17% in TM band 7.

Several observations can be made with regard to the spectral separability of these classes. Recreational grassland and traditional pastures show the most significant divergence from the other classes. In fact, the spectral response curve for recreational grassland shows the highest reflectance values all along the investigated segment of the spectrum (485-2 215 nm). Traditional pasture on the other hand consistently remains at the bottom of the collective chart, reaching a maximum radiation reflectance value of 22.2%, approximately half of the equivalent for recreational grassland. In the near-infrared (TM4), the difference between these two classes is even more pronounced with reflectance values for recreational pasture reaching the threefold of those for traditional pasture.

The other grassland classes show a much more mutually comparable evolution, which lies somewhere between those of recreational grasslands and traditional pasture, although more toward the former. Despite some alternations in ranking for these three remaining curves, the maximum difference between the reflectance values for overgrown grassland, fodder production grassland and modern pastures rarely exceeds 5%.

Other than for early growing season, the grassland spectral separability at the height of the growing season was also investigated. A Landsat TM image from 17 July 2006 (T2) served as the basis for the extraction of digital number for grassland pixels. The results of the interpolation of reflectance averages for each TM sensor band (except TM6) are presented in Figure 14. In the this section, I will compare the conventional spectral curve graphs for satellite images T1 and T2 and concentrate on the importance of temporality in analysing spectral separability.

A quick glance at both plots shows that the mutual spectral separability of grassland types is dramatically reduced during the height of the growing season (*i.e.* T2). This is mostly due to a significant convergence of the spectral response curves for traditional pasture, modern pasture, fodder production grassland and overgrown grassland across the electromagnetic

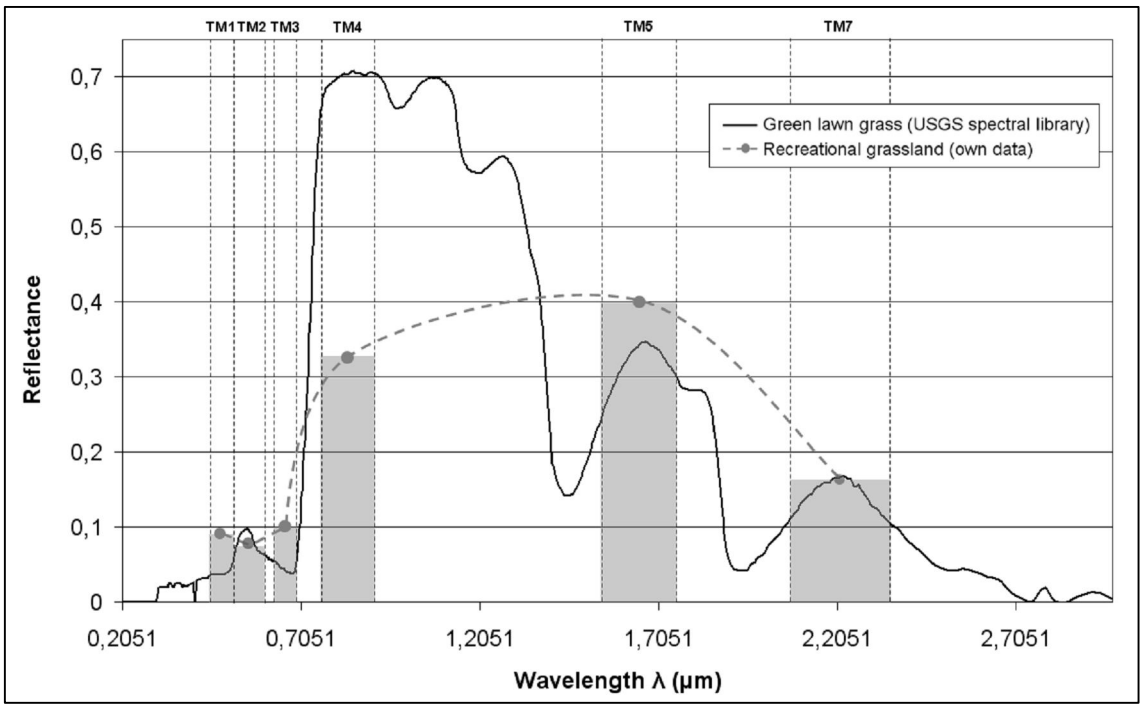
spectrum, be it more pronounced in TM bands 4 (NIR) and 5 (mid-IR). At the beginning of the growing season, the largest distance between the curves for these classes – expressed in percent reflectance – still amounted to 13.2% for TM band 4 and 15.2% for TM band 5. By mid-summer, these figures had been reduced to 4.6% and 5.5% respectively. Although less pronounced, a similar evolution can be observed in the visible light range.

Another important general evolution that the data show is an increase in reflection of infrared radiation while that in TM band 5 (mid-infrared) decreases. The latter trend however only applies to modern pasture, fodder production grassland and overgrown grassland, witnessing an average reflectance reduction of 17.9%. Recreational grassland and traditional pasture show an increase in reflectance in TM band 5. In comparison to the situation in early growing season image T1, the infrared reflectance recorded for all grassland types during mid-summer (T2) increases 51.4% on average.

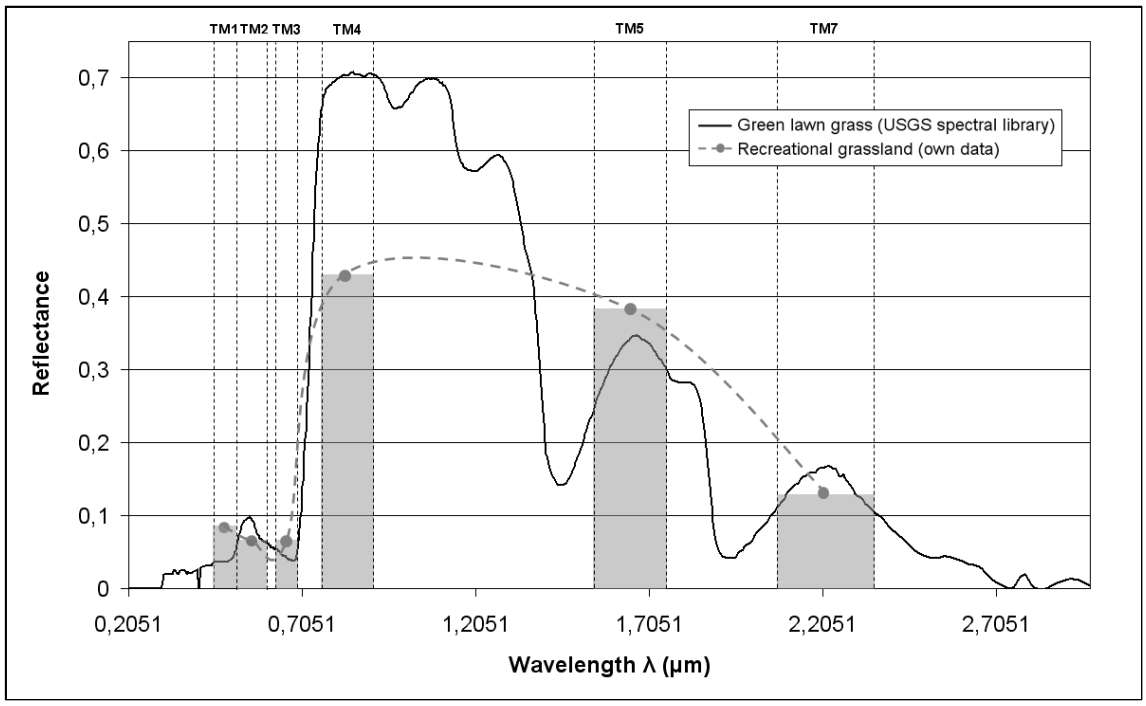
When studying curve similarity, the curves for overgrown grassland, fodder production grassland and modern pasture show much higher mutual similarity than similarity to those of recreational grassland and traditional pasture. The curves for modern pasture and overgrown pasture are perhaps most similar since they show the same overall shape and do not intersect each other. Compared to these latter two, fodder production grassland shows a distinct shift from relatively higher reflection in TM4 (NIR) to a relatively lower reflection in TM5 (mid-IR). Such relative reversal of spectral response may be the key to separating this grassland type from the rest (Lillesand and Kiefer 1994).

6.2.2 *Curve validity verification*

In Figure 15, the validity of these constructed curves can be examined. The continuous black curve represents laboratory measurements for fresh grass from a lawn in Kansas (U.S.) and released by the U.S. Geological Survey whereas the grey bars correspond to



T1



T2

Figure 15. Comparison of laboratory measurements for fresh green grass (Clark *et al*. 2007) and TM reflectance data retrieved from occurrences of recreational grassland as depicted by a Landsat image of 1.5.2007 (T1) and 17.7.2006 (T2). Note that the continuous recreational grassland curves are approximations of the interpolation curves for recreational grassland presented earlier (cf. Fig. 14).

the measured reflectance values for recreational grassland in the inner Archipelago of Turku after the source image (T1) was corrected for atmospheric scattering. At first sight,

these data seem incongruous. Recorded reflectance values appear especially incompatible in the near-infrared (TM4), with TM-based reflectance measurements reaching less than half the value recorded by the laboratory spectrometer. In the visible light range, the evolution of the data from blue over green to red light does not agree with laboratory measurements. In fact, the evolution derived from Landsat TM imagery is the reverse of the laboratory measurements in that more light is reflected in TM bands 1 and 3 relative to band 2. For green vegetation, blue and red light are normally absorbed whereas green light is more likely to be reflected. However, not only does the TM sensor effectively cover the mid-infrared reflectance peaks of fresh green grass, a surprisingly good match with the laboratory readings for these bands (TM5 and TM7) can be observed.

6.3 Spectral response swath graphs

It is important however, to keep in mind that the presented spectral response curves are mere interpolations of average reflectance values. Apparent spectral separability between the response curves in Figure 14 does not imply that grassland pixels can be guaranteed to belong to the spectrally nearest grassland class. As always with land cover classification, there are large overlaps in class range and pixels are classified based on likelihood of class membership. In this light, a *standard deviation swath curve* was produced. Next to the average curves already presented in Figure 14, it also depicts upper and lower curves at one standard deviation from the average curve for each grassland class. This allows for the viewer to form a better understanding of the range within which a theoretical 68.3% of values for each grassland class are situated and consequently the extent of the overlap thereof. This of course requires a normal distribution of the reflectance values for each class (Lillesand and Kiefer 1994). Statistical analysis of the data showed that normality can indeed be assumed.

As can be see in Figure 16, a combination of all five standard deviation swatch curves results in substantial amounts of spectral overlapping. The only grassland class swath curves that show good disjunction from the other grassland classes' are recreational grassland and traditional pasture. A theoretical 68.3% of reflectance values derived from patches of traditional pasture – based on data normality logic - lie almost entirely below those of the other four grassland classes. For recreational grassland, the best spectral separability can be found between bands TM4 (NIR) and TM5 (mid-IR), although more towards the former. The results for other grassland types show little potential for spectral

separation on a per-band basis. Overgrown grassland shows some separation potential between traditional pasture and the other classes in TM4 (NIR) and the same is true for the part of modern pasture values that are to be found just below the lower standard deviation curve for recreational grassland in the same band, although be it to a very limited extent.

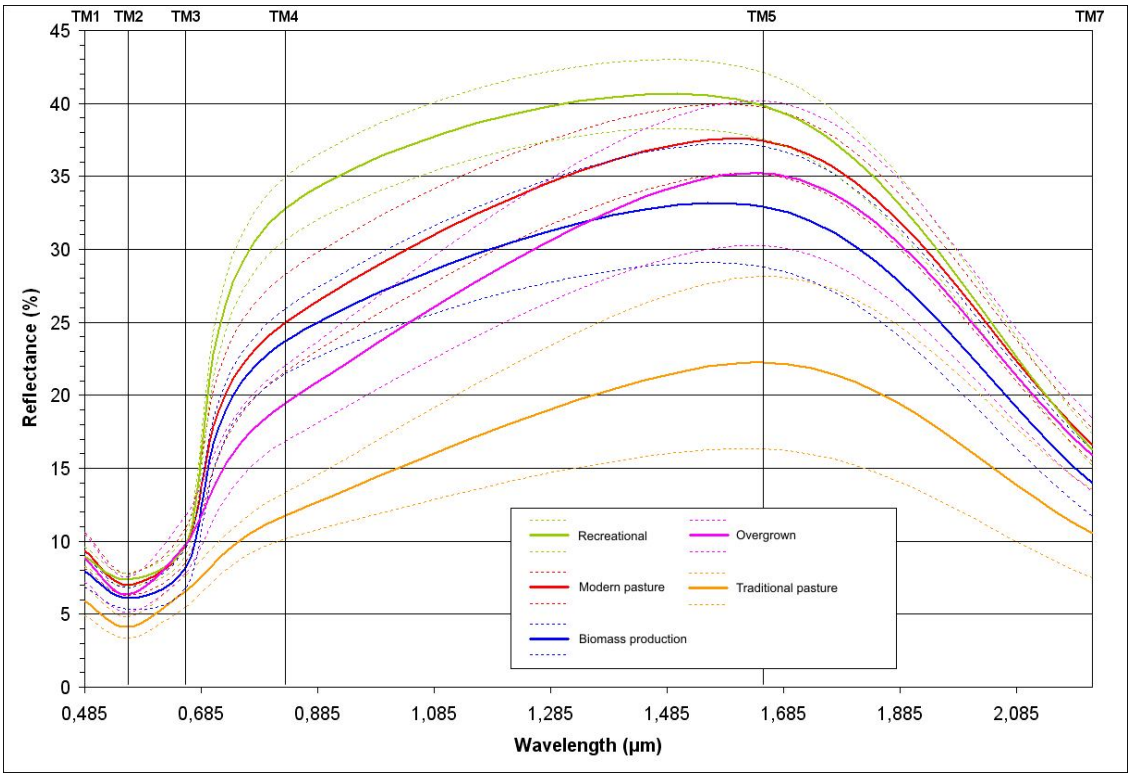


Figure 16. Spectral swath graph for TM image T1 (1.5.2007)

Another observation that producing standard deviation swath graphs allows, is an appreciation of the measure of pixel value variance within each class along a gradient of increasing wavelength. The wider the curve swath at a given wavelength, the more important the variation in pixel values in that spectral band and therefore also the more heterogeneity in the source patch(es). In this respect, traditional pasture and overgrown grassland show the highest variance values, cumulating in band TM5 (mid-infrared) with the difference in upper and lower standard deviation values amounting to 11.8% and 9.9% respectively, compared to an average of 6.5% for the other grassland classes.

6.4 Coincident spectral plot

Figure 17 below is the coincident spectral plot for early growing season TM image T1. For Thematic Mapper spectral bands 1-5 and 7 (TM1-TM5 and TM7), the spectral response for each of five different grassland types is shown. The lower and the upper limits of each

spectral reading correspond to the average reflectance value plus and minus two standard deviations, so each segment shown theoretically contains 95.4% of all values for the respective band and class.

It is impossible to describe all observations that can be made based on the coincident spectral plot presented. Therefore, I will now concentrate on those observations that were impossible or difficult based on the conventional spectral curve graphs (Figure 14), but easily retrieved from the coincident spectral plot. One interesting finding is that overgrown grassland dominates the upper reach of the reflectance values. Only in band 4 (near-infrared) is it bypassed by recreational grassland. Recreational grassland may have higher average reflectance values compared to other grassland types, it has a relatively homogenous cover meaning its overall range is limited. These observations conflict with earlier findings, stating that pixels with the highest reflectance values are most likely to belong to the recreational grassland class.

Traditional pasture on the other hand does indeed exhibit a lower limit inferior to all other classes, but – as Figure 17 shows – only really experiences full spectral separability in band 4 (NIR). Again, this contradicts earlier findings in that it certainly cannot be considered *fully* spectrally distinct as appeared to be the case in the spectral response swath chart. Stress must be exerted however, on the fact that comparison is complicated by the representation of only ± one standard deviation (68.3% of pixels) for the conventional graphs and ± two standard deviations (95.4% of pixels) for the coincident spectral plot. In the latter case however, any conclusions drawn are more likely to be universally applicable since they apply to all but a few of the recorded reflectance values.

Another trend that only really became appreciable studying the coincident spectral plot is the spectral distinctness of traditional and modern pasture. Traditional and modern pasture show a general tendency of minimal overlap in reflectance range, with quasi-complete separation in green light and the near-infrared. Especially in the visible light bands of the spectral swath plot (Figure 16), this spectral separability potential was not as pronounced.

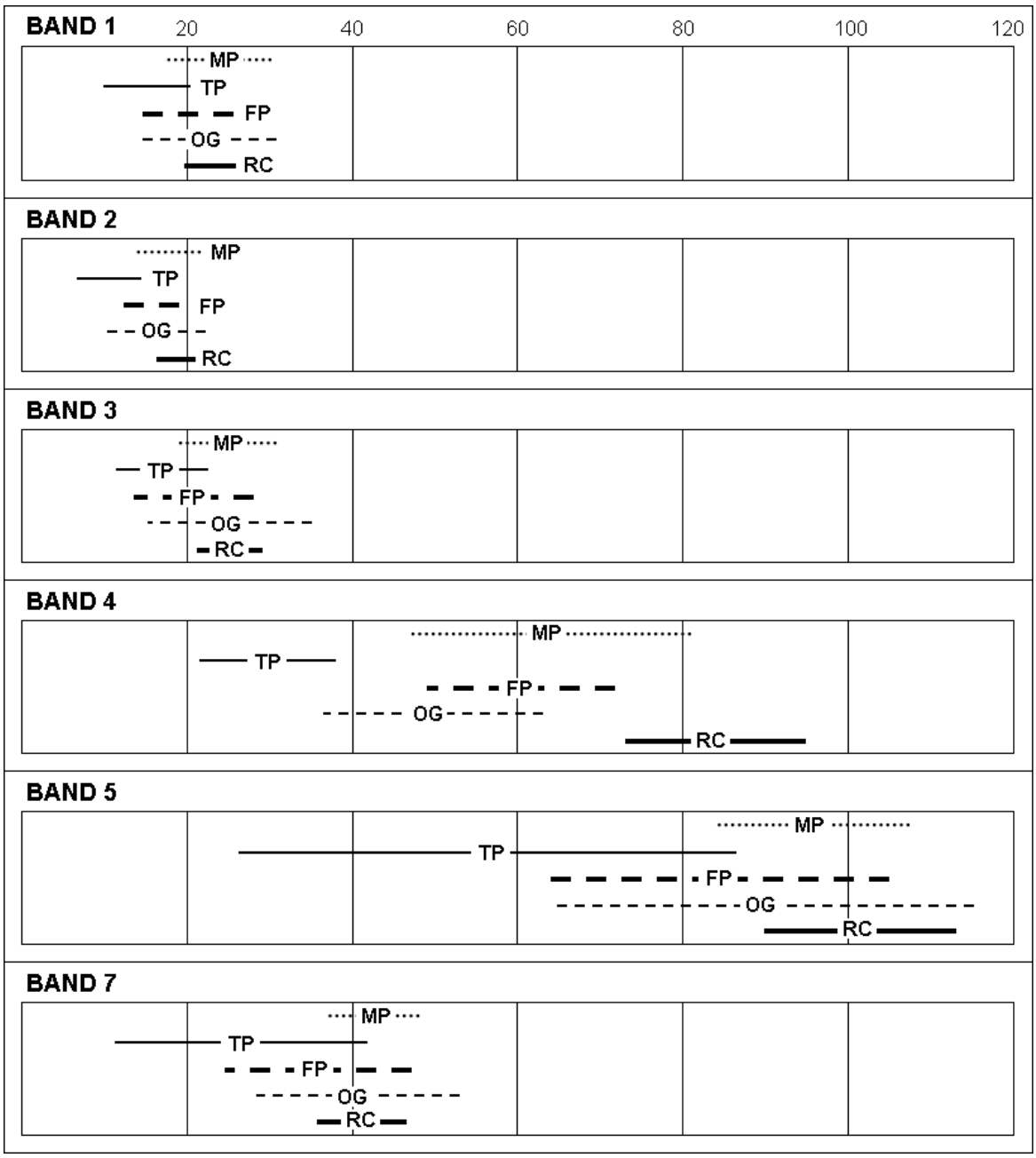


Figure 17. Coincident spectral plot for a selection of grassland pixels derived from Thematic Mapper image T1 (1.5.2007). The letter codes used are abbreviations for the five grassland types: modern pasture (MP), traditional pasture (TP), fodder production (FP), overgrown grassland (OG) and recreational grassland (RC). Note that the values along the x-axis (top) signify the recorded digital number (DN), of which only 0-120 of a maximum value of 255 is shown.

If spectral separability is assessed on a per-band basis, overgrown grasslands and fodder production are the least separable. When counting the mutual overlaps for each band, these two grassland types overlap with all other classes in all bands except for band TM4 (NIR), where they do not overlap with traditional pastures or recreational grassland. By further investigating the extent of overlap in range for each band, overgrown grassland shows more overlapping than fodder production grassland due to a higher variance of its

reflectance values. This implies of all five classes, overgrown grassland is most challenging to identify in electronic imagery depicting grassland occurrences only.

Another observation concerns the spectral differentiation between modern pastureland and fodder production grassland. At first sight, similar responses for both these grassland types would seem logical. The data however indicate that although largely analogous, some interesting discrepancies between these two grassland types do exist. The reflectance range for modern pasture for instance, reaches higher digital numbers in TM bands 4 (NIR), 5 and 7 (mid-IR) than is the case for fodder production grassland. The latter on the other hand, shows a significantly more pronounced variance in reflectance values in TM bands 5 and 7 (mid-IR), but not for TM band 4 (NIR).

6.5 Quantitative spectral separability assessment

6.5.1 *Transformed divergence*

The values for the transformed divergence test are presented in Table 13 below. Values above the dashed diagonal line are those for early growing season (T1) and those below for mid-summer (T2). The expected spectral separability is indicated for each value: good (**), fair (*) and poor ( ) (Jensen 1996) and the highest possible value is 2 000. The results suggest that during early growing season, more grassland types are spectrally separable than during mid-summer. At the beginning of May, four grassland class combinations show good separability and another two show fair separability, whereas at the height of the growing season only three class combinations show fair to good separability.

During the early growing season, good spectral separability is especially associated with recreational grassland. The highest separability value obtained is for the combinations recreational grassland and traditional pasture. Given average reflectance values for these two classes form the extremes of the conventional spectral response curve plot, this is not a surprising result. Another earlier observation that has been confirmed is that – as observed in the spectral response plots – overgrown grasslands, fodder production site and modern pastureland are challenging to spectrally separate. All three combinations of these classes (MP-OG, MP-FP, OG-FP) show poor separability according to the transformed divergence analysis.

Table 13. Transformed divergence values for five grassland types: fodder production (FP), recreation (RC), overgrowth (OG), modern pasture (MP) and traditional pasture (TP).

	FP	RC	OG	MP	TP
FP	-	1 925**	1 429	1 192	1 814*
RC	1 881*	-	1 985**	1 728*	1 998**
OG	1 113	1 947**	-	1 436	909
MP	933	1 653	1 163	-	1 537
TP	984	1 854*	651	665	-

As grassland vegetation develops over the course of spring and early summer, some drastic changes in grassland discrimination potential occur. As mentioned earlier, the number of one-to-one grassland combinations that results in good or fair spectral separation is reduced from 5 (T1) to 3 (T2). Only one set of grasslands is predicted to lead to good spectral separation in the mid-summer image (T2): recreational grassland and overgrown grassland. Just as was the case during early May (T1), traditional pasture and overgrown grassland are least separable (651), although the divergence value for early May was 258 units higher. A generalised decrease in divergence value can be observed for all class

Table 14. Average transformed divergence values for each grassland class.

	1.5 (T1)	17.7 (T2)
FP	1590.0	1227.8
RC	1909.0	1833.8
OG	1439.8	1218.5
MP	1473.3	1103.5
TP	1564.5	1038.5
Average	1595.3	1284.4

combinations. The earlier assumption based on the spectral response plot for T2 (17.7) that traditional pasture, modern pasture, overgrown grassland and fodder production grassland are difficult to separate during mid-summer is confirmed by the transformed divergence scores. In fact, this statement can be extended to include poor spectral separability prospects for the grassland class set recreational grassland-modern pasture. Recreational grassland *can* be discriminated from the other grassland classes with fair or good success rates.

Table 14 presents the averaged transformed divergence values for each grassland class. Based on these data, recreational grassland is indeed the most spectrally separable, both

during early growing season and mid-summer. Contrary to earlier assumptions however, traditional pasture does not come second in the separability ranking but rather third for TM image T1 and last for TM image T2. Based on the spectral response curve plots for early May (T1), the following ranking of separability potential was assumed (from large to small): recreational grassland (1), traditional pasture (2), overgrown grassland (3), modern pasture (4) and fodder production (5). This assumption is entirely dismissed by the average transformed divergence values. The observation derived from the coincident spectral plot (T1) that overgrown grassland is probably the most challenging grassland class to separate from the other *does* still stand. Also, the overall average divergence values for each image are in agreement with the early observation that the TM image for early growing season (T1) allows a significantly better spectral discrimination of grasslands than that for mid-summer (T2).

6.5.2 *Contingency analysis*

The results for the contingency analysis are presented in Table 15 below. It presents the contingency matrices for the non-parametric classification of training pixels derived from early growing season image T1 (Table 15a) and mid-summer image T2 (Table 15b), using normalised sample sizes. The overall classification accuracy levels reached are 75.6% for early May (T1) and 67.6% for mid-summer (T2) whereas the minimum level of accuracy of identifying land use and land cover categories from remote sensing data should be at least 85 percent (Anderson *et al*. 1976).

Returning to the contingency matrix for 1.5.2007 (Table 15a) omission error values ranging from 2.9 - 54.5% were obtained. Recreational grassland constitutes the lower extreme of this range whereas modern pasture stands at the upper limit. The remaining classes (*i.e.* fodder production grassland, overgrown grassland and traditional pasture) show intermediate omission error values of approximately 20-30%. On the commission error side of the contingency matrix, modern pasture and overgrown grassland show the highest error values (50.6% and 41.8% CE resp.) whereas recreational grassland and traditional pasture are associated with the lowest (4.4% and 9.1% CE resp.). Fodder production grassland approaches the 20% error level.

Table 15. Contingency matrices based on the normalised training samples presenting the user accuracy (UA) and commission error (CE) as well as the producer accuracy (PA) and omission error (OE) percentages for the classification of a selection of grassland pixels into the five classes Note that the reference data and the classification data are given in number of pixels.

		REFERENCE DATA							
		FP	**RC**	**OG**	**MP**	**TP**	**Sum**	**UA**	**CE**
CLASS. DATA	**FP**	**405.8**	**0.9**	**10.3**	**104.7**	**31.6**	**553.3**	73.3%	26.7%
	RC	**1.3**	**485.4**	**4.1**	**52.4**	**3.1**	**546.3**	88.9%	11.1%
	OG	**32.7**	**0.0**	**394.6**	**107.3**	**72.4**	**607.1**	65.0%	35.0%
	MP	**44.5**	**13.7**	**45.5**	**227.7**	**17.3**	**348.8**	65.3%	34.7%
	TP	**15.7**	**0.0**	**45.5**	**7.9**	**375.5**	**444.5**	84.5%	15.5%
	Sum	**500.0**	**500.0**	**500.0**	**500.0**	**500.0**	**2 500.0**		
	PA	81.2%	97.1%	78.9%	45.5%	75.1%			
	OE	18.8%	2.9%	21.1%	54.5%	24.9%			

Table 15a. Contingency matrix for TM image T1 (1.5.2007).
Overall classification accuracy: 75.6%

		REFERENCE DATA							
		FP	**RC**	**OG**	**MP**	**TP**	**Sum**	**UA**	**CE**
CLASS. DATA	**FP**	**403.3**	**8.6**	**37.9**	**102.7**	**54.5**	**607.0**	66.4%	33.6%
	RC	**5.5**	**472.5**	**2.2**	**11.4**	**3.7**	**495.4**	95.4%	4.6%
	OG	**22.1**	**3.4**	**343.8**	**107.3**	**108.9**	**585.5**	58.7%	41.3%
	MP	**41.4**	**13.7**	**22.3**	**216.9**	**79.2**	**373.6**	58.1%	41.9%
	TP	**27.6**	**1.7**	**93.8**	**61.6**	**253.7**	**438.4**	57.9%	42.1%
	Sum	**500.0**	**500.0**	**500.0**	**500.0**	**500.0**	**2 500.0**		
	PA	80.7%	94.5%	68.8%	43.4%	50.7%			
	OE	19.3%	5.5%	31.3%	56.6%	49.3%			

Table 15b. Contingency matrix for TM image T2 (17.7.2006).
Overall classification accuracy: 67.6%

In the early growing season image (T1), only the pixels identified as recreational grassland and traditional pasture meet the threshold of 85% classification accuracy (Anderson *et al*. 1976), although fodder production is not far behind (80.1% UA). The modern pastureland and overgrown grassland classes on the other hand produced too high commission error values to be said to be spectrally separable with reasonable success.

As mentioned earlier, the overall classification accuracies obtained for images T1 and T2 are 75.6% and 67.6% respectively. Mid-summer image T2 is therefore characterised by an overall trend of augmented classification error, which can also be observed from the commission and omission error fields given in Table 15b. Although the general ranges of

error percentages remain roughly the same for both measures, significant changes can be noted for overgrown grassland, traditional pasture and perhaps also recreational grassland.

Over the course of spring, the omission error for overgrown grassland pixels increases by 50% (21.1% > 31.3%) while that for traditional pasture doubles (24.9% > 49.3%). On the user accuracy side of the contingency matrix, overgrown grassland classification success remains similar to that in the early growing season image, but traditional pasture does continue to show a downward trend with a commission error almost three times larger. Recreational grassland is the only class that witnesses an improvement in classification accuracy (11.1% > 4.6% CE).

Table 16 below provides information on the separability for each grassland class pair based on the number of mutually misclassified pixels and their contribution to the overall classification success. At the beginning of May (Table 16a), modern pasture and overgrown grassland are least separable from each other, followed by modern pasture and fodder production grassland. Respectively, these combinations account for 25.0% and 24.4% of all misclassified pixels. Recreational grassland and fodder production grassland on the other hand are most mutually separable (0.4%), followed by recreational grassland and traditional pasture (0.5%).

Over the course of spring and early summer, a number of changes in spectral separability occur. At mid-summer (T2), the least separable grassland class pair was that of overgrown grassland and traditional pasture (25%) followed by fodder production and modern pasture (17.8%). The least separable pair during early growing season – modern pasture and overgrown grassland – moved to the fourth ranking (16%). On the other extreme of the range, again the best separable pairs are associated with recreational grassland: recreational grassland and traditional pasture comes, followed by recreational grassland and overgrown grassland, each pair accounting for approximately 0.7% of all classification error.

The most significant changes however, are all associated with the modern pasture class. Whereas the combination traditional pasture – modern pasture amounted to a mere 4.1% in early May, its share of the total classification error increased by 13.3% by mid-summer. The combinations overgrown grassland and modern pasture as well as recreational

Table 16. Mutual confusion matrices for images T1 and T2 showing the number of mutually confused pixels above the diagonal line and the stake of all misclassified pixels for each grassland class combination underneath the diagonal line.

	FP	RC	OG	MP	TP	Sum
FP	-	2.2	43.1	149.2	47.3	**241.8**
RC	0.4%	-	4.1	66.1	3.1	**73.3**
OG	7.0%	0.7%	-	152.8	117.9	**270.7**
MP	24.4%	10.8%	25.0%	-	25.2	**25.2**
TP	7.7%	0.5%	19.3%	4.1%	-	
Sum	**39.6%**	**12.0%**	**44.3%**	**4.1%**		**611.0 (100%)**

Table 16a. Mutual confusion matrix for image T1 (1.5.2007).

	FP	RC	OG	MP	TP	Sum
FP	-	14.1	60.0	144.2	82.1	300.4
RC	1.7%	-	5.7	25.2	5.4	36.3
OG	7.4%	0.7%	-	129.6	202.7	332.3
MP	17.8%	3.1%	16.0%	-	140.9	140.9
TP	10.1%	0.7%	25.0%	17.4%	-	
Sum	**37.1%**	**4.5%**	**41.0%**	**17.4%**		**809.8 (100%)**

Table 16b. Mutual confusion matrix for image T2 (17.7.2007).

grassland and modern pasture on the other hand, show a significant decrease in mutual misclassification (-9% and -7.7% resp.). It most be noted however that these percentages are no absolute change in error procurement but rather relative values, describing the stake of the total number of classification errors for each image. Between images T1 and T2, an increase in total number of misclassified pixels by 8% was observed. This implies that even though a small decrease in error *stake* may be observed over the course of spring and early summer, the *absolute number* of misclassified pixels associated with the same grassland type pair can still be larger when compared to the situation during early growing season.

Modern pasture grasslands are associated with surprisingly low levels of classification accuracy, both during early growing season and mid-summer. A total of 64.3% and 54.3% of all classification errors are associated with modern pastureland for images T1 and T2 respectively. Especially the observed spectral separation difficulty for the combination modern pasture – overgrown grassland during early growing season (25.0%) had not been foreseen. Returning to the spectral coincident plot (cf. Fig. 17), modern pasture does show the largest overlap with overgrown grassland in TM band 4 (NIR) – which has previously proven a crucial band in the discrimination of grassland types – but fodder production grassland certainly shows a much more elevated similarity with overgrown pasture across

the electromagnetic spectrum. Yet the percentage of classification errors associated with the combination overgrown grassland – fodder production grassland only amounts to 7.0% (compared to 25.0% for modern pasture – overgrown grassland**).**

Table 17 below provides an overview of the classification error for each grassland class as a whole, both for TM images T1 and T2. Each error figure represents the number of misclassified pixels associated with the corresponding class. Figures assume a total of 500 training pixels for each class. As mentioned earlier, an important observation relates to the fact that generally more classification error is attributed to the spectral discrimination attempt conducted on the mid-summer image (T2) than on that from the beginning of May (T1). Recreational grassland constitutes the only error to this general trend, showing a decrease in misclassified pixels from 75.5 to 50.4 over the course of spring and early summer. This implies that for the best overall spectral mutual discrimination of grassland types, early growing season imagery is preferred over mid-summer imagery.

Table 17. Summated normalised error values for each grassland class based on the figures produced by the contingency analysis presented in Table 15. The relative ranking in terms of separability is given in brackets.

Grassland type	**T1**	**T2**
Fodder production grassland	241.8 (3)	300.4 (2)
Recreational grassland	75.5 (1)	50.4 (1)
Overgrown grassland	317.9 (4)	398.0 (3)
Modern pasture	393.3 (5)	439.9 (5)
Traditional pasture	193.5 (2)	431.1 (4)

Generally, the spectral separability ranking of the five grassland classes remains constant when moving from image T1 to image T2. Traditional pasture however constitutes the exception to the rule in that is moves from second ranking in early May (T1) to fourth ranking during mid-summer, causing fodder production grassland and overgrown grassland to shift upwards by one place. This suggests that in order to successfully discriminate traditional pastureland from other grassland types, early growing season imagery is mandatory.

Recreational grassland is highly spectrally separable, during early growing season and even more so during mid-summer. Modern pasture on the other hand is least successfully discriminated at both times. This latter finding had not been predicted through the construction of the spectral response graphs or by the transformed divergence analysis.

Other predictions made earlier have also been dismissed. Although averaged divergence values projected traditional pasture to constitute the third most spectrally separable class during early growing season, it ranks second according to the more empirical contingency analysis. Moreover, whereas counting class overlaps in the coincident spectral plot indicated fodder production grassland and overgrown grassland would be the two classes most challenging to discriminate in image T1, the contingency analysis gives rise to the assumption it would be overgrown grassland and modern pasture instead.

Table 18 below provides a more systematic insight in the inconsistencies observed when comparing the results of the transformed divergence and those from the contingency analysis, especially with regard to the classes that are most and least easily mutually confused. The figure suggests that the most significant disagreement between the findings affects the class of modern pasture and that of recreational grassland. Generally speaking, classes are most easily confused with modern pasture or overgrown grassland and least easily with recreational grassland. However, a shift in phenology between early growing season (T1) and mid-summer (T2) does to some extent alter this.

Table 18. Most and least easily mutually confused grassland classes on the contingency analysis (cf. Table 15). The mentions in brackets indicate disagreeing findings for the transformed divergence test.

	Most easily confused		**Least easily confused**	
	T1	**T2**	**T1**	**T2**
FP	MP (OG)	MP	RC	RC
RC	MP	MP	FP (TP)	TP (OG)
OG	MP (TP)	TP	RC	RC
MP	OG (FP)	FP (TP)	TP (RC)	RC
TP	OG	OG	RC	RC

7 Discussion

7.1 Spatial separability

The presented findings raise questions as to the suitability of Landsat TM 30 m resolution imagery for allowing the retrieval of smaller grassland patches in the Turku Archipelago. Since these findings apply to over half of all grassland patches in the study area, they have important implications for the spatial separability of grassland patches in general. Clearly, the vast majority of the hypothetical grid cells capturing grassland patches ≤ 1 ha is not pure or near-pure and therefore inhibits an accurate spatial delineation of the respective grassland occurrences. Especially in the light of biodiversity mapping for the archipelago, where key grassland habitats are often associated with small-scale yet rugged patches (*e.g.* Hinneri & Lehtomaa 1994, Kukkonen 1994) and high levels of landscape heterogeneity (Raatikainen *et al.* 2007; Metsähallitus 2010), important shortcomings can be attributed to the use of Landsat data. From the observations made, medium-resolution imagery such as Landsat scenes with a cell size of 30 m may not be suitable to map grassland patches in the Turku Archipelago to the desired level of accuracy. Given the small size and delicate shapes of grasslands in the region, the spatial resolution should be sufficiently fine, perhaps even down to 10 m (Kivinen 2007).

Grassland patches characterised by low shape complexity however, often belong to the group of least biodiverse grasslands. Simple shapes often indicate subjection to more industrial agricultural practices since modern-day machinery requires space and simple parcel edge shapes in order to manoeuvre with ease. From an ecological viewpoint therefore, those grasslands that require extra attention are often to be found in the lower right quarter of the regression analysis graph (Fig. 13 p.65). Such patches are typified by a notably lower average grid cell intersection (*i.e.* pure pixel potential), the most complex of which are associated with values of less than 20%. This renders it highly unlikely that pure pixels can be obtained from these sites, not only jeopardising the retrieval of its accurate location and extent but also the retrieval of any accurate spectral information. In other words, most of the more complex kind of small grassland patches (≤ 1 ha) occurring in the inner archipelago will be overlooked in sensing efforts using TM imagery. Moreover, since the Ministry of Agriculture dataset does not include many of the smallest patches (overgrown grasslands, road verges, …), an even larger proportion of all grassland patches

occurring in the archipelago can be expected to be unsuitable for Thematic Mapper retrieval.

The spatial separability of grassland patches assumes that only the presence of pure or near-pure grassland pixels can lead to a successful delineation of this land cover class. In practice however, numerous additional techniques are at hand to improve the success rate of that delineation. Object-oriented segmentation methods for instance could lead to more reliable classification results, particularly in small structured landscaped (Luoto *et al.* 2002). Such segmentation acknowledges that part of resolution constraints, at least in high-resolution imagery, can be overcome by looking at the relationship between adjacent pixels. Much information such as texture and shape of individual objects is not visible by looking at individual pixels (Thomas *et al*. 2003). Temporal pattern recognition on the other hand may also be employed. It uses changes in feature response over time as an aid in feature identification. Distinct spectral and spatial changes during the growing season can permit discrimination from multidate imagery that would be impossible given any single date (Lillesand & Kiefer 1994). This approach has shown to be especially effective in relative terms for deciduous forests and semi-natural grasslands (Price *et al*. 2002, Debinksi *et al*. 2000). However, these assumptions may not apply to all available remote sensing technology available. Smaller pixel sizes combined with fewer spectral bands in aerial photography and new high-resolution satellite imagery (IKONOS, QuickBird) can create classification problems rather than solve them due to greater spectral variation within a class and a greater degree of shadow (Laliberte *et al*. 2004).

7.2 Conventional response curves

7.2.1 *Response curves for T1 and T2*

Although conventional spectral response curves are largely the result of a simplification of the original spectral response data, they indicate general trends in the spectral response patterns of grassland types. The results show that there are differences in reflection of electromagnetic radiation depending on grassland type, spectral band and time of year. Despite some differences, the general trends – both during early growing season and mid-summer – comply with vegetation reflectance patterns described in the literature (Sabins 1997, Infoterra 2009). A clear red edge shows when moving from the visual bands to the near-infrared beyond which the absorption of incident radiation by leaf water starts playing

an important role. This observation however, is somewhat more true for the mid-summer image (T2) than for the early growing season image (T1). The reason for this pertains to the fact that vegetation reflection patterns are generally deduced from observations made at peak growth which in this case corresponds to mid-summer.

The most striking differences observed in the conventional spectral signatures are that exhibited by recreational grassland and traditional pasture. These differences could already be assumed beforehand. The most diverging grassland class – recreational grassland – is also the most unusual grassland type involved in the analysis, in that the source patches (i.e. parts of the Ruissalo and Harjattula golf courses) are known to be subject to very intensive management regimes that render them into a very artificial type of grassland. Traditional pasture on the other hand shows a similar dissimilarity to the *general* grassland spectral response because of its peculiar spectral composition. This – again – is not surprising since a distinction was made *a priori* between modern pasture and traditional pasture because the latter clearly constituted a distinct spectral class when visually assessing the imagery from the beginning of May (T1).

Despite some divergence in the spectral response induced by the exceptional classes of recreational grassland and traditional pasture, the other grassland classes (i.e. modern pasture, traditional grassland and fodder production grassland) show rather little spectral divergence, especially toward peak greenness. The perfect convergence of the spectral signatures of these grassland classes in the near-infrared band in image T2 is a good example of this. Since the near-infrared band captures the level of chlorophyll in vegetation, this perfect convergence could perhaps be caused by the presence of a natural limit in photosynthetic activity in grasslands. Although during the early growing season there are still clear differences in chlorophyll content of the foliage due to inherently different conditions in the studied grasslands (wetness, soil fertility, *etc.*), by peak greenness the vegetation might have enough time to compensate for these differences and photosynthesises at the maximum possible rate, which is reflected by equal radiation reflection readings in the NIR band.

Although the curves for overgrown grassland, fodder production grassland and modern pasture strongly converge across the spectrum, there may still be a way to successfully determine grassland classes of spectral responses from these three habitat types, using complex computations offered by artificial intelligence. These complex computations can

identify multispectral (and therefore multidimensional) correlations that the human mind cannot. Such relative reversal of spectral response may be the key to separating this grassland type from the rest (Lillesand and Kiefer 1994).

7.2.2 *Curve validity verification*

The spectral response curves produced by this study may not perfectly fit laboratory measurements, but there is a number of constraints at work when comparing these findings. Naturally, comparing a batch of blades of fresh grass with the spectral signature of a golf course retrieved through low spectral resolution satellite imagery may not be possible given the drastic differences in data quality. Not only may there be dissimilarity in the spectral response properties of the grass species as such, recreational grasslands and – in this case – golf courses are more than merely a collection of grass blades. Although carefully selected during the reflectance value retrieval stage, recreational grassland pixels are likely to include individual trees, small pools or sand bunkers, all of which are sub-pixel scale features and therefore likely to interfere with the extraction of pure grassland data for such environments. Furthermore, given that recreational grasslands are typically kept short to very short may lead to reduced reflection of near-infrared energy, which is correlated to amount of biomass. Even if the grass was allowed to grow tall, photosynthetic activity at the very start of the growing season (1.5.) is too little to give rise to high reflectance values in the near-infrared.

7.3 Spectral response swath graphs

Given the spectral distinctness of traditional pasture and recreational grassland persists in the spectral response swath graphs, the observation regarding the exceptionality of these two classes made based on the conventional response plot for early growing season image T1 is confirmed. When 68.3% of the digital numbers derived from recreational grassland and traditional pasture have a value outside the ± 1 standard deviation range of other spectral response swaths, the classification success of pixels from these two classes can be presumed to be reasonable high. Thematic Mapper bands 4 and 5 appear to be the optimal bands for discriminating between different types grassland. This observation corresponds to that by Price *et al*. (2002) for the discrimination between six grassland types in eastern Kansas. It shows that particularly chlorophyll content and leaf water are important factors to analyse when attempting to discern different types of grassland.

Overgrown grassland and traditional pasture can be expected to show more-than-average variance since they are inherently characterised by a high vegetation and other cover diversity. Overgrown grasslands will typically show a mixture of fresh and dry grasses as well as more woody plants and organic debris that has accumulated over time. Traditional pasture on the other hand is typically found in locations not suitable for modern pasture to be implemented. This is commonly due to regular flooding or high levels of wetness as well as the presence of rocky outcrops and stands of shrubs, reeds and trees. Although these factors are generally common, the exact mixture and relative share of each of them will always be somewhat different, resulting in a more diversified spectral response of the grassland habitats they occur in.

7.4 Coincident spectral plot

Although fundamentally very close to the spectral response swath plot, the coincident spectral plot revealed a few traits in the data that were difficult if not impossible to derive form the spectral response curves. This is partially due to the manner in which the response data are presented by the coincident spectral plot, but also the change in chosen data range from ± 1 standard deviation from the mean spectral response to ± 2 standard deviations. In the latter case, virtually all (95.4%) response data are represented, giving rise to a shift in observations regarding the spectral separability of some of the grassland types. For instance, although recreational grassland was assumed to comprise the top range of all response values, the ± 2 standard deviations coincident spectral plot showed that overgrown grassland and not recreational grassland hosts the highest reflectance values. By showing *quasi* all response data, it also becomes clear which spectral classes are characterised by high variance in DN value and which ones are more homogenous; the data ranges for overgrown grassland and traditional pasture now clearly overlap other classes to a much more substantial extent than previously thought. This makes sense given the typical compositional diversity of these classes, as was discussed earlier.

Looking at particular grassland class pairs in more detail, the data indicate that modern pasture is typified by higher reflection rates than those for fodder production. This is likely to be the result of contrasting management regimes. Fields intended for fodder production will be subjected to mowing once or twice throughout the growing season, preferably not

until the grass has grown tall but certainly before general senescence. In practice, this often implies mowing these fields for the first or the second time during the late growing season, minimising and homogenising vegetative growth during the process. When the grass starts growing again at the beginning of the growing season, little or no dead material is present on site and overall biomass is limited, keeping the reflectance of near-infrared radiation low. In modern pastures on the other hand, biomass remaining at the end of the growing season is not necessarily cleared.

This possibly means that grass re-establishes itself more easily as soon as spring arrives while some dead plant material remains. When dead, dry material remains in the pastures after winter, the effect of moisture content on reflections in the mid-infrared field is weakened, producing higher values. This corresponds to the observations made. Vegetation water content in fodder production grassland on the other hand, is typically elevated because tall-growing grass species are preferred. With a leafs up to 45 cm long and 1 cm wide, Timothy-grass (*Phleum pratense*) – the most widely sown grass species for silage and hay production in the Nordic countries (Höglind et al. 2001) – is likely to absorb more mid-infrared radiation than the more woody species found in pastureland.

7.5 Quantitative spectral separability assessment

7.5.1 *Transformed divergence*

There are some severe discourses between predictions based on spectral response plots and statistical separability analysis results. Although the most notable spectral discrimination potential and evolutions were correctly identified, many of more subtle differences in spectral response depicted by spectral response plots do not provide ground for conclusive remarks regarding separability. Despite the full spectral separability of traditional pasture in TM band 4 described earlier for instance, low transformed divergence values were obtained when comparing it to overgrown grassland. Another finding contrary to the interpretation of the coincident spectral plot is that traditional and modern pasturelands are not spectrally distinct. With a transformed divergence value of 1 537, only poor discrimination success can be expected (Jensen 1996). Counting the mutual overlaps for grassland classes in determining the least spectrally separable combination such as was done for the coincident spectral plot, has also proven unreliable. Although from the overlap counting exercise it seemed overgrown grassland and fodder production grassland were the least separable

grassland types, two other grassland combinations receive even lower divergence scores (*i.e.* TP-OG and MP-FP).

7.5.2 *Contingency analysis*

The overall classification accuracy levels attained by the contingency analysis of images T1 (75.6%) and T2 (67.6%) are similar to those found by Guo *et al*. (2003) when they attempted to distinguish between three common types of grassland management occurring in north-eastern Kansas: pasture, hayed grasslands and rehabilitated grassland. However, they do not reach the recommended 85% success rate recommended by Anderson *et al*. (1976). In fact, aiming to distinguish between all five major types of grassland occurring in the Turku Archipelago, such level of accuracy cannot be obtained using early May or mid-summer imagery. It appears unlikely any single-date Landsat satellite imagery allows supervised classification accuracy levels superior to 85%, unless perhaps, when it is subjected to a higher order of pre-processing, such as when computing vegetation indices.

Exclusively concerned with the mutual spectral separation potential of five grassland classes, the present study does not treat the separability of grasslands from non-grassland land cover categories. It is however important to mention that severely diminished classification accuracy levels are to be expected if these five grassland classes were to be separated in imagery containing spectrally similar but botanically discrete types of land cover such as reed beds or deciduous forest. When conducting supervised classification of an IRS satellite image, Vescovo & Gianelle (2006) for instance found that mixed pixels in which both grassland and tree stands occur are more likely to be classified as forest than as grassland, implying an underestimation of the abundance of grassland. Category confusion issues arising when involving landscape features other than grassland most certainly lowers the grassland classification success rate.

Assuming the input of spectral response signatures from grasslands only, the results of the contingency analysis indicate that the mutual discrimination of grassland types is a challenging task when using mid-summer imagery. Only recreational grassland (95.4% UA) achieves a classification success rate higher than 85%. This can be attributed to the convergence of the spectral responses. As noted earlier, any discrimination potential traditional pasture may have during early growing season is lost as the growing season progresses toward peak greenness. As grassland patches reach their maximum level of

photosynthetic activity, spectral responses show great similarity in the near-infrared band, inhibiting spectral separation potential. Recreational grassland is the only grassland class that manages to exceed the others in terms of NIR energy reflection and therefore becomes even more discernable than during early growing season.

Most of the results of the contingency analysis can be logically explained. The observed degradation of the potential for the spectral separation of modern and traditional pasture between images T1 and T2 for instance, can be attributed to an equalisation of vegetative growth and environmental conditions. Whereas still typified by flooding, exposed soil and hampered plant growth during early growing season, traditional pasturelands accumulate sufficient amounts of biomass to largely resemble modern pastures at the height of the growing season. Traditional pasture and overgrown grassland as well as fodder production grassland and modern pasture are two grassland class combinations that show high levels of classification error. This can be explained by general similarity in the composition of these classes and derived spectral response. Traditional pasture and overgrown grasslands for instance are both characterised by patch heterogeneity, exhibiting features such as individual stands of trees and shrubs. With, modern pasture and fodder production grassland on the other hand, the focus is directed toward optimising grass biomass accumulation through idealised growing conditions (soil quality, water supply, nutrient availability, *etc.*). Despite clear differences in site purpose (pasture, fodder production, no purpose), similarity in site conditions through incidental correspondence of their spectral response partially inhibits management-based discrimination of grasslands from TM imagery.

7.6 Improving spectral separability

7.6.1 *Hyperspectral remote sensing*

For a considerable amount of time since the onset of spaceborne imaging, sensing systems were limited to multispectral devices. These are devices collecting data for carefully chosen discontinuous wavebands only. Aware of the restrictive nature of this type of data acquisition, attempts to measure electromagnetic energy across the spectrum started in the early 1980's. However, it would take until well into the 1990's before the developments in electronics, computing and software required to obtain and process hyperspectral imaging data reached a larger segment of the Earth observation community (Goetz 2009).

Hyperspectral data differ from multispectral data in that they are a collection of measurements across the electromagnetic spectrum (cf. Fig. 18). These quasi-continuous data typically have a spectral resolution of 0.005-0.01 μm (Tsai and Philpot 1998), allowing a much more precise selection of wavebands relevant to a given analysis exercise. Imagery with such high spectral resolution often reveals reveal small spectral anomalies, allowing spectral differentiation that would not be possible based on multispectral scanner systems such as TM (Schmidt & Skidmore 2001).

With the current availability of hyperspectral data, new techniques are being developed in order to exhaust the potential of these data to its fullest extent. Tsai and Philpot (1998) acknowledged that techniques developed in the field of spectroscopy have potential relevance to the analysis of hyperspectral data too as they realised that treating hyperspectral data as truly continuous – just like is done in spectroscopy – allows access to information that is often suppressed by standard analysis methods. There are however, a number of risks involved in directly adapting spectroscopy techniques in that conditions such as illumination and sample purity are substantially different when comparing lab experiments to air- or spaceborne sensing of surfaces on Earth.

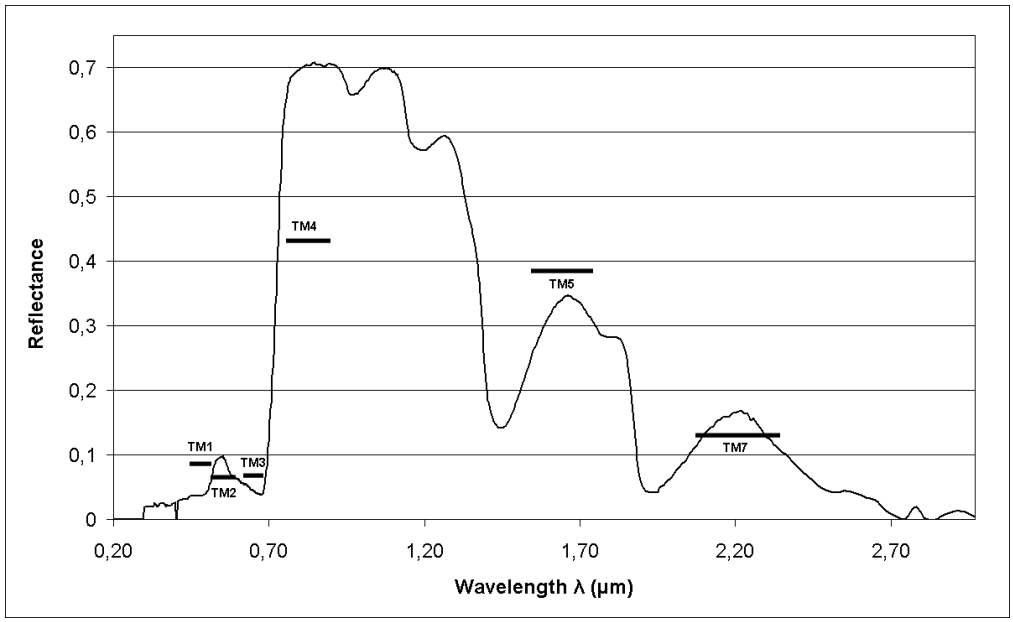


Figure 18. Multispectral Thematic Mapper data (bold segments) versus hyperspectral data (continuous curve) for fresh grass. Note the difference in coverage of the electromagnetic spectrum and the spectral coarseness of TM data. TM band 6 (10.40-12.50 μm) not shown.

Aiming to mitigate these risks, Tsai and Philpot assessed analysis techniques used in spectroscopy and found that one particular technique called derivative analysis is

particularly promising for use with remote sensing data. Second order or higher order derivatives of reflectance curves are relatively insensitive to variation in illumination intensity whether caused by changes in sun angle, cloud cover or topography. Furthermore, derivatives should also be relatively less sensitive to spectral variations in sunlight and skylight at the spectral sampling interval typical of hyperspectral systems.

Given the availability of hyperspectral data, derivative analysis may allow to distinguish between different types for grassland where classic multispectral analysis fails. In a study on grassland vegetation in Inner Mongolia, Yamano *et al*. (2003) successfully differentiated between four grassland species typical for the area using high-order derivative analysis of hyperspectral reflectance data acquired at 5 km above ground. Fourth order derivative analysis allowed the characterisation of key spectral features. Reflectance spectra around 670 nm and 720 nm allowed the most effective discrimination between the studied grass species.

It is possible that the inclusion of a derivative analysis based on hyperspectral data would reveal subtle differences in grassland type that would be difficult if not impossible to discriminate using multispectral sensor data. Vegetation characterisation using derivative analysis however does assume a high level of vegetation homogeneity. Despite continuous improvements in spatial resolution, spectral responses from certain types of grassland that can be found in the Turku Archipelago will always show impurities. Overgrown patches of grassland and traditional pastures with rocky outcrops and bush stands are examples of inherently yet non-consistently heterogeneous land cover types. Derivative analysis of hyperspectral remote sensing may therefore only offer a solution to certain grassland types.

7.6.2 *Using vegetation indices*

The discrimination between grasslands can be improved by computing a vegetation index (VI) for the input imagery instead of working with raw sensor band combinations. From previous studies, it appears that in order to further efforts to improve the spectral discrimination of grasslands, the variance in level of *greenness* of these grasslands may be the key.

Price *et al*. (2002) applied this assumption to six grassland types in eastern Kansas, each of which is subjected to a different management regime. They found that out of nine

vegetation indices, a *Tasselled Cap Greenness Vegetation Index* was found to deliver the best results. Moreover, the grassland type discrimination analysis based on the vegetation index outperformed the ability of raw TM bands by 10%. However, no significant improvement was found when raw TM bands were combined with the VIs. Of the TM bands, TM4 (NIR) was found to be most effective single band at discriminating the 6 grassland types. When bands of multiple dates are included, the discrimination accuracy improved. The inclusion of too many bands however lead to a decrease in discrimination accuracy.

The green herbage ratio (GR) is an important biophysical parameter. It is the equivalent of biomass/(biomass + necromass) and indicates photosynthetic activity of vegetation components, pedoclimatic conditions and phenological state of vegetation, rendering it a useful indicator for type of management regime (Vescovo & Gianelle 2006). A study by Vescovo & Gianelle 2006 showed that out of 10 most commonly used vegetation indices based on satellite imagery (using the red, green and infrared bands), only one (i.e. Green-NDVI) showed a significant correlation with the observed green herbage ratio. The green-NDVI or green normalised difference vegetation index is calculated using the green spectral channel instead of the red one (Gitelson *et al*. 1996).

8 Conclusion

When looking at minimum patch size for guaranteeing the capture of at least one pure pixel in remote sensing imagery, surprisingly large values are at play. A circular patch needs to be 0.57 ha in size when imaged by a 30 m resolution sensor. For a square patch this amounts to 0.72 ha. These geometrical shapes are however highly theoretical and unlikely to correspond to real-life patch shapes. This implies that most patches may have to exceed 1 ha in size before they allow for a single pure pixel to be retrieved from 30 m resolution imagery.

Over half of all grassland patches in the study are smaller than 1 ha. Patches with a size of less than 1 hectare are challenging to capture using Landsat TM. When upscaling the spatial resolution from 30 m to 10 m, a substantial improvement in the retrieval of pure of near-pure pixels was achieved. A hypothetical 30 m resolution grid produced a success rate of 13%, whereas the same test with a spatial resolution of 10 m made this rate jump to 52%. This implies that input imagery with a spatial resolution of 10 m or better can be considered a minimum requirement for grassland patch discrimination efforts in the inner Turku Archipelago.

The P/A metric can be used as a predictor of the *pure pixel potential* of grassland patches and other landscape units with similar spatial characteristics. An inverse variation function has been presented for the prediction of the *average fraction* of patches, which is directly related to their pure pixel potential. Using this function, patch vector sets can be rapidly assessed for their compatibility with 30 m resolution imagery. This may be useful if good quality vector data are at hand while there is uncertainty about the required image quality for producing accurate land cover classification results.

The ideal phenological stage for the discrimination of boreal grasslands was found to be during early growing season. When the growing season starts, grasslands are easily discerned from other habitat types while the spectral variance between the different types of grassland is at its peak. The patch type-specific variance is largely due to differences in patch composition that show before fast vegetative growth is initiated by warmer weather conditions. As time progresses toward the peak greenness moment, the spectral reflectance patterns of different grassland habitats converge and much of the discrimination potential

is lost. The observed convergence was especially marked for in near-infrared band, indicating that the response of foliage chlorophyll equalises over time. This may be an indication of the existence of a limit in photosynthetic activity of grasslands. An important exception to this rule is recreational grassland, which is characterised by notably higher reflection levels across the reflective part of the electromagnetic spectrum.

Out of the applied techniques, the coincident spectral plot provided the best means for visual appreciation of the spectral discrimination potential of the grasslands, while the contingency analysis delivered the most reliable quantitative results. Although each technique has its merits, severe discrepancies were found when comparing the predicted spectral separability of the different grassland types. The transformed divergence analysis greatly underestimated the ability of the classification algorithm while the conventional spectral response curves proved to be an oversimplification of the truth. Since there can be high levels of variance in recorded digital numbers for certain grassland classes, it is recommendable to always include all observations within two standard deviations of the response mean.

In terms of spectral separability of the studied grassland types, recreational grassland consistently showed the best discrimination potential. Other grasslands showed varying levels of separability over time. Traditional pasture and modern pasture for instance, can be easily discerned from one another during early growing season, but have spectral responses that are very similar at mid-summer. Although the results show that the spectral responses of some grasslands can be discriminated with an acceptable level of accuracy, very limited success was noted when attempting to distinguish between those from traditional pasture and overgrown grassland on one hand and those from fodder production grassland and modern pasture on the other. This implies that it is not possible to conduct a management-based discrimination of grasslands using Thematic Mapper imagery. Despite this limitation, overall classification success rates superior to 75% can be reached when attempting to discern different grassland types from each other using early growing season imagery.

In order to improve the success rate of grassland discrimination from Thematic Mapper imagery in complex and fine-scale environments such as the Turku Archipelago, vegetation indices should be involved. Studies have shown that grasslands have a *greenness* factor that – when incorporated in the analysis – offers additional information about the vegetation composition or management regime of grasslands. With the development of higher-detail

imagery however, substantial levels of supplementary discrimination potential are available. Especially hyperspectral remote sensing data and the derivative analysis thereof have proven effective at distinguishing different types of grassland, even up to species level.

9 List of abbreviations

CE	Commission error
CLC	Corine Land Cover
CRP	Conservation Reserve Program
DN	Digital number
EMS	Electromagnetic spectrum
ERDAS	Earth Resource Data Analysis System
FMI	Finnish Meteorological Institute
GLS	Global Land Survey Program
GSD	Ground Sampling Distance
NASA	U.S. National Aeronautics and Space Administration
NLS	Finnish National Land Survey
OE	Omission error
PA	Producer's accuracy
SYKE	Finnish Environment Institute
TM	Thematic Mapper
UA	User's accuracy
USGS	United States Geological Survey
VI	Vegetation index

10 Annexes

Annex I - Technical specifications for TM scenes 1.5.2007, 17.7.2006 and 21.8.2007

	1 May 2007 (T1)	17 July 2006 (T2)	21 August 2007 (T3)
Full scene image (RGB 321)			
Satellite	Landsat 5	Landsat 5	Landsat 5
Sensor	Thematic Mapper	Thematic Mapper	Thematic Mapper
Bands	1,2,3,4 (multispectral) 5,7 (mid-infrared) 6 (thermal infrared)	1,2,3,4 (multispectral) 5,7 (mid-infrared) 6 (thermal infrared)	1,2,3,4 (multispectral) 5,7 (mid-infrared) 6 (thermal infrared)
Pixel size	30 m (bands 1, 2, 3, 4, 5, 7) 120 m (band 6)	30 m (bands 1, 2, 3, 4, 5, 7) 120 m (band 6)	30 m (bands 1, 2, 3, 4, 5, 7) 120 m (band 6)
Format	TIF	TIF	TIF
Processing status	Georeferenced	Georeferenced	Georeferenced
Capture date & time	1 May 2007 09:36:40	17 July 2006 09:35:06	21 August 2007 09:35:18
Path/Row	190/18	190/18	190/18
Provider	U.S. Geological Survey, obtained through EarthExplorer	U.S. Geological Survey, obtained through EarthExplorer	U.S. Geological Survey, obtained through EarthExplorer
Solar azimuth / Sun elevation	162.1 ° / 43.9 °	157.1 ° / 49.5 °	160.5 ° / 40.8 °
Atmospheric conditions within study area boundaries	Few scattered cloud occurrences, some haze to in western parts	Cloud- and haze-free	Some cloud disturbance in central and northeastern parts

11 References

Ahti T, Hämet-Ahti L, Jalas J (1962) Vegetation zones and their sections in northwestern Europe. *Annales Botanici Fennici* 5:169-211.

Alanen A (1996) Maaseudun mansikkapaikat – muistojako vain? *Luonnon Tutkija* 100:197-208 (in Finnish).

Anderson JR, Hardy ET, Roach JT, Witmer RE (1976) A land use and land cover classification system for use with remote sensor data. *Geological Survey Professional Paper* 964.

Askew D, Slater J (1995) Assessing the Nature Conservation Value of Grasslands by Remote Sensing in Geoscience and Remote Sensing Symposium IGARSS '95'. *Quantitative Remote Sensing for Science and Applications* 2:1233-1235.

Atlas of Finland (1987). *Climate (Ilmasto), folio 131*. National Board of Survey & Geographical Society of Finland. Helsinki. 31 pp.

Beckwith, SL (1954) Ecological succession on abandoned farmlands and its relationship to wildlife management. *Ecological Monographs* 24:349-376.

Chavez PS (1988) An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* 24:459-479.

Clark RN, Swayze GA, Wise R, Livo E, Hoefen T, Kokaly R, Sutley SJ (2007) USGS digital spectral library splib06a. *Digital Data Series 231*. For more information refer to http://speclab.cr.usgs.gov/spectral.lib06.

Colwell RN (1966) Uses and limitations of multispectral remote sensing. *Proceedings of the Fourth Symposium on Remote Sensing of Environment*, revised edition. University of Michigan, 71-100 pp.

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37:35–46.

Coppin P, Jonckheere I, Nackaerts K, Muys B and Lambin E (2004) Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25:1565-1596.

Coulson KL (1966). Effect of reflection properties of natural surfaces in aerial reconnaissance. *Applied Optics* 5:905-917.

Coupland RT (1974) Fluctuations North American grassland vegetation. In: Reinhold T (ed) *Handbook of vegetation science*, Part VIII. Dr. W. Junk, The Hague, 233-241 pp.

Cousins SAO (2001) Analysis of land-cover transitions based on 17th and 18th century cadastral maps and aerial photographs. *Landscape Ecology* 16:41-54.

Debinski DM, Jakubauskas ME, Kindscher K (2000) Montane meadows as indicators of environmental change. *Environmental Monitoring and Assessment* 64:213-225.

Earth Resources Data Analysis System (2005) *ERDAS Field Guide*™, Leica Geosystems Geospatial Imaging.

European Council (1992) Habitats Directive - Council directive 92/43/EEC on the conservation of natural habitats and of wild fauna and flora. Publications Office of the European Union, L 206:7-50.

Eronen, M (2005) Land uplift: Virgin land from the sea. In Seppälä M (ed) *The physical geography of Fennoscandia*. Oxford University Press, New York, 463 pp.

European Topic Centre on Land Use and Spatial Information (2006),http://terrestrial.eionet.europa.eu/CLC2000/classes/Pictures?CLCcategory=3/3.2/3.2.1&CLCtitle=Natural%20grasslands, accessed on 12.03.2008.

Finnish Environment Institute (2006) Pintavesien tyypittely. *The Finnish Environment* 807:11-21 (in Finnish with English summary).

Finnish Meterological Institute (2007), http://www.fmi.fi/weather/climate_13.html, accessed on 16.11.2009.

Finnish Meterological Institute (2008), http://www.fmi.fi/weather/climate_14.html, accessed on 16.11.2009.

Finnish Meteorological Institute (2009a), http://www.fmi.fi/weather/climate_4.html, accessed on 2.9.2009.

Finnish Meteorological Institute (2009b), http://www.fmi.fi/weather/climate_3.html, accessed on 2.9.2009.

Fisher P (1997) The pixel: a snare and a delusion. *International Journal of Remote Sensing* 18:679-685.

Frisén R, Johansson CE and Suominen V (2005) Archipelagos in the Baltic Sea. In Seppälä M (ed) *The physical geography of Fennoscandia*. Oxford University Press, New York, 463 pp.

Gallego FJ (2005) Stratified sampling of satellite images with a systematic grid of points. *ISPRS Journal of Photogrammetry and Remote Sensing* 59:369-376.

Gardberg J (1931) Samfällda näringsfång i havsbandet. Skrifter utgivna av Svenska Litteratursällkapet i Finland 217:99-152 (in Swedish).

Gates DM (1970) Physical and physiological properties of plants. In: National Academy of Sciences. *Remote sensing with special reference to agriculture and forestry*. Washington D.C., 224-252 pp.

Ghahramani S (2005) *Fundamentals of probability, with stochastic processes*, 3rd edition. Prentice Hall, Upper Saddle River, New Jersey, 644 pp.

Gitelson AA, Kaufman YJ, Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment* 58:289-298.

Global Land Survey (2010), http://gls.umd.edu/index.html, accessed on 13.1.2010.

Goetz A (2009) Three decades of hyperspectral remote sensing of the Earth: A personal view. Remote Sensing of Environment 113:S5-S16.

Goillot G (1980) Significance of spectral reflectance for natural surfaces. In: Fraysse G (ed) *Remote sensing application in agriculture and hydrology*. A.A. Balkema, Rotterdam, 53-68 pp.

Granö O, Roto M and Laurila L (1999) Environment and land use in the shore zone of the coast of Finland. *Turun yliopiston maantieteen laitoksen julkaisuja* 160, 76 pp.

Guo X, Price KP, Stiles J (2003) Grasslands discriminant analysis using Landsat TM single and multitemporal data. *Photogrammetric Engineering & Remote Sensing* 69:1255-1262.

Hildebrandt G (1976) Die spektralen Reflexionseigenschaften der Vegetation. In: *Proceedings of the Symposium on Remote Sensing in Forestry*. IUFRO, Oslo, pp. 9-23 (in German).

Hinneri S (1994) Kedot ja ketokasvit – lounaissuomalainen näkökulma. *Lutukka* 10: 35-40 (in Finnish with English summary).

Hinneri S, Lehtomaa L (1994) Ketokasvien ekologiasta lounaisrannikolla ja -saaristossa. *Lutukka* 10:41-50 (in Finnish with English summary).

Haapanen A (2005) Biodiversity conservation. In Seppälä M (ed) *The physical geography of Fennoscandia*. Oxford University Press, New York, 463 pp.

Höglind M, Schapendonk AHCM, Van Oijen M (2001) Timothy growth in Scandinavia: combining quantitative information and simulation modelling. *New Phytologist* 151: 355–367

Ihse M (1995) Swedish agricultural landscapes – patterns and changes during the last 50 years, studied by aerial photos. *Landscape and Urban Planning* 31:21-37.

Infoterra (2009), http://www.infoterra.co.uk/data_sat_banddesc.php, accessed on 20.10.2009.

Jacobsen A, Broge NH, Hansen BU (1995) Monitoring wheat fields and grasslands using spectral reflectance data. *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR)*, Nov. 26 to Dec. 1, 1995. Melbourne, Australia.

Jensen JR (1996) *Introductory digital image processing: a remote sensing perspective*, 2nd edition. Prentice Hall, Upper Saddle River (NJ), 316 pp.

Jukola-Sulonen EL (1983) Vegetation succession of abandoned hay fields in central Finland – a quantitative approach. *Communicationes Instituti Forestalis Fenniae* 112:1-85.

Jutila HM (1999) Vegetation and seed bank of grazed and ungrazed Baltic coastal meadows in SW Finland. PhD thesis. Turku University, Turku.

Kakkuri J (1987) Character if the Fennoscandian land uplift in the 20th century. *Geological Survey of Finland Special Paper* 2: 15-20.

Kallio P (1979) *Ruissalo: luontoa ja kulttuuria (Ruissalo: nature and culture)*. Otava, Helsinki, 224 pp. (in Finnish).

Käyhkö N, Skånes H (2006) Change trajectories and key biotopes – Assessing landscape dynamics and sustainability. *Landscape and Urban Planning* 75:300-321.

Kennedy RE, Cohen WB, Schroeder TA (2007) Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment* 110:370–386.

Kivinen, S (2007) Local and regional determinants of biodiversity patterns in boreal agricultural landscapes. PhD thesis. Turku University, Turku.

Kontula T, Lehtomaa L, Pykälä J (2000) Land-use history, vegetation and flora in Rekijoki valley, Somero, SW Finland. *Suomen Ympäristö* 306:1-91 (in Finnish with English summary).

Kotiluoto R (1998) Vegetation changes in restored semi-natural meadows in the Turku Archipelago of SW Finland. *Plant Ecology* 136:53-67.

Krogerus R (1921) Ruissalon tammimetsien kovakuoriaiseläimistö. *Luonnon Ystävä* 4:79-83.

Kukkonen I (1994) Poutajoen laakso - ajatuksia maiseman ja kasvillisuuden historiasta. *Lutukka* 10:51-57 (in Finnish with English summary).

Kuusisto E (2005) Snow as a geographic element. In Seppälä M (ed) *The physical geography of Fennoscandia*. Oxford University Press, New York, 463 pp.

Laliberte AS, Rango A, Havstad KM, Paris JF, Beck RF, McNeely R, Gonzalez AL (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* 93:198-210.

Lillesand TM, Kiefer RW (1994) *Remote sensing and image interpretation*, 3rd edition. John Wiley & Sons Inc., New York, 750 pp.

Lindgren L (1998) Lenholmin luonnonsuojelu- ja lehtojensuojelualueen hoito- ja käyttösuunnitelma (Management plan for Lenholm's broad-leaf forest and nature reserve). *Metsähallituksen luonnonsuojelujulkaisuja Sarja B* 44:1-136 pp. (in Finnish).

Lindgren L (2000) *Island pastures*. Metsähallitus and Edita Ltd., Helsinki, 203 pp.

Luoto M, Kuussaari M and Toivonen T (2002) Modelling butterfly distribution based on remote sensing data. *Journal of Biogeography* 29:1027-1037.

Luoto M, Rekolainen S, Aakkula J, Pykälä J (2003) Loss of plant species richness and habitat connectivity in grasslands associated with agricultural change in Finland. *Ambio* 32:447-452.

Marttila O, Jantunen J, Saarinen K (1999) The status of semi-natural grasslands in the province of South Karelia, SE Finland. *Annales Botanici Fennici* 36:181-186.

Matson PA, Parton, WJ, Power AG, Swift MJ (1997) Agricultural intensification and ecosystem properties. *Science* 277:504-509.

Mausel PW, Kramber WJ, Lee JK (1990) Optimum band selection for supervised classification of multispectral data. *Photogrammetric Engineering and Remote Sensing* 56:55-60.

Metsähallitus (2010) *Lenholmin luonnonsuojelualue*, http://www.luontoon.fi/page.asp?Section=703, accessed on 2.5.2010 (in Finnish).

Metsähallitus (2010) Natura 2000 site fact file for the Archipelago Sea, accessed through http://www.natura.org/sites_fi_archi.html on 11.2.2010.

Meyer M, Werth L (1990) Satellite data: management panacea of potential problem? *Journal of Forestry* 88:10-13.

Miatkowski Z (2004) Identification of grassland transformations due to water regime changes caused by opencast mining applying remote sensing data. In: Goossens R (ed) *Remote sensing in transition*. Millpress, Rotterdam, 405-408 pp.

Mikkola K (1997) Population trends of Finnish Lepidoptera during 1961-1996. *Entomologica Fennica* 8:121-143.

NASA (2009a), http://landsat.gsfc.nasa.gov/about/history.html, accessed on 13.1.2010.

NASA (2009b), Landsat-7 Science Data User's Handbook, http://landsathandbook.gsfc.nasa.gov/handbook/handbook_toc.html, accessed on 14.1.2010.

National Land Survey of Finland (2000) SLICES land cover dataset, Helsinki. For more information refer to www.slices.nls.fi (in Finnish).

National Land Survey of Finland (2007) Topographic Database. The National Land Survey, Helsinki, Finland.

National Land Survey of Finland (2010) Technical description for the Topographic Database, accessed through http://www.maanmittauslaitos.fi/en/default.asp?id=488 on 8.5.2010.

Navalgund RR (2001) Remote sensing, 1. Basics and applications. *Resonance* 6:51-60.

Öster A, Persson K, Eriksson O (2008) Validation of plant diversity indicators in semi-natural grasslands. *Agriculture, Ecosystems & Environment* 125:65-72.

Peterson DL, Price KP, Martinko EA (2002) Discriminating between cool season and warm season grassland cover types in northeastern Kansas. *International Journal of Remote Sensing* 23:5015-5030.

Pöyry J, Lindgren S, Salminen J, Kuussaari M (2004) Restoration of butterfly and moth communities in semi-natural grasslands by cattle grazing. *Ecological Applications* 14:1656-1670.

Psomas A, Zimmermann E, Kneubühler M, Kellenberger T, Itten K. (2005) Seasonal variability in spectral reflectance for discriminating grasslands along a dry-mesic gradient in Switzerland. In: Zagajewski B, Sobczak M, Wrzesien M (eds) *Proceedings of 4th EARSeL Workshop on Imaging Spectroscopy: new quality in environmental studies*. EARSeL, Warsaw, Poland, 48 pp.

Price PP, Guo X, Stiles JM (2002) Optimal Landsat TM band combinations and vegetation indices for discrimination of six grassland types in eastern Kansas. *International Journal of Remote Sensing* 23:5031-5042.

Pykälä (2001) Maintaining biodiversity through traditional animal husbandry. *The Finnish Environment* 495:1-202 (in Finnish with English summary).

Pykälä, J (2003) Effects of restoration with cattle grazing on plant species composition and richness of semi-natural grasslands, *Biodiversity and Conservation* 12, 2211-2226.

Pykälä J (2004) Cattle grazing increases plant species richness of most species trait groups in mesic semi-natural grasslands. *Plant Ecology* 175:217–226.

Pykälä J, Luoto M, Heikkinen RK and Kontula T (2005) Plant species richness and persistence or rare plants in abandoned semi-natural grasslands in northern Europe. *Basic and Applied Ecology* 6:25-33.

Raatikainen KM, Heikkinen RK, Pykälä J (2007) Impacts of local and regional factors on vegetation of boreal semi-natural grasslands. *Plant Ecology* 189:155-173.

Rabotnov TA (1974) Differences between fluctuations and successions – examples in grassland phytocoenoses of the U.S.S.R. In: Reinhold T (ed) *Handbook of vegetation science*, Part VIII. Dr. W. Junk, The Hague, 19-24 pp.

Richards JA, Jia X (2006) *Remote sensing digital image analysis: an introduction*, 4th edition. Springer-Verlag, Berlin, 476 pp.

Ritari A, Saukkola P (1985) Spectral reflectance as an indicator of ground vegetation and soil properties in northern Finland. *Communicationes Instituti Forestalis Fenniae* 132:1-37.

Riitters KH, O'Neill RV, Hunsaker CT, Wickham JD, Yankee DH, Timmins SP, Jones KB and Jackson BL (1995) A factor analysis of landscape pattern and structure metrics. *Landscape Ecology* 10:23-39.

Rock BN, DL Skole, Choudhury BJ (1993) Monitoring vegetation change using satellite data. In: Solomon AM, Shugart HH (eds) *Vegetation dynamics and global change*. Chapman and Hall, New York, 153-167 pp.

Rodwell JS (ed) (1992) *British plant communities volume 3 – Grasslands and montane communities*. Cambridge University Press, Cambridge.

Ryhänen EL (2005) Itämeri – ainokaisemme. In: Wuolijoki A, Iltanen J (eds) *Aino – suuri suomen kartasto*. Genimap, Helsinki/Porvoo, 22-23 pp (in Finnish).

Sabins FF (1997) *Remote sensing: principles and interpretation*. W. H. Freeman & Co, New York, 449 pp.

Schott JR, Salvaggio C, Volchok W (1988) Radiometric scene normalization using pseudoinvariant features. *Remote Sensing of Environment* 26:1-6.

Schmidt KS, Skidmore AK (2001) Exploring spectral discrimination of grass species in African rangelands. *International Journal of Remote Sensing* 22:3421-3434.

Schulman A, Heliölä J, Kuussaari M (eds) (2005) Farmland biodiversity on the Åland islands and assessment of the effects of the agri-environmental support scheme. *The Finnish Environment* 734:1-210.

Skånes H (1996) Landscape change and grassland dynamics – retrospective studies based on aerial photographs and old cadastral maps during 200 years in south Sweden. Phd thesis. Department of Physical Geography, University of Stockholm, Stockholm.

Slater PN, Doyle FJ, Fritz NL, Welch R (1983) Photographic systems for remote sensing. *American Society of Photogrammetry Second Edition of Manual of Remote Sensing* 1: 231-291.

Slotte H (1993) Hamslingsträd på Åland. (Pollards in Åland, SW Finland). *Svensk Botanisk Tidskrift* 87:283–304 (in Swedish with English abstract).

Söderman N, Tenovuo R (1960) *Ruissalon linnut (Birds of Ruissalo)*. Werner Söderström Osakeyhtiö, Porvoo, 116 pp. (in Finnish).

Stoate C, Boatman ND, Borralho RJ, Carvalho CR, Snoo GR, Eden P (2001) Ecological impacts of arable intensification in Europe. *Journal of Environmental Management* 63:337-365.

Statistics Finland (2007) *Statistical yearbook of Finland* 2007, volume 102. Statistics Finland, Helsinki.

Statistics Finland (2009) *Statistical yearbook of Finland* 2009, volume 104. Statistics Finland, Helsinki.

Swain FH, Davis SM (eds) (1978) *Remote sensing: the quantitative approach*. McGraw – Hill, New York, 396 pp.

Thomas N, Hendrix C and Congalton RG (2003) A comparison of urban mapping methods using high-resolution digital imagery. *Photogrammetric Engineering and Remote Sensing*, 69:963-972.

Tiainen J, Kuussari, M, Laurila IP, Toivonen T (eds) (2004) *Elämää pellossa – Suomen maatalousympäristön monimuotoisuus*. Edita Publishing Oy, Helsinki, 366 pp. (in Finnish).

Toivonen T, Luoto M (2003) Landsat TM images in mapping semi-natural grasslands and analysing of habitat pattern in an agricultural landscape in Southwest Finland. *Fennia* 181: 49-67.

Tsai F, Philpot W (1998) Derivative analysis of hyperspectral data. Remote sensing of Environment 66:41-51.

Tuhkanen S (1984) A circumboreal system of climatic-phytogeographical regions. *Acta Botanica Fennica* 127:1-50.

United States Geological Survey (2009a) *Landsat Update*, volume 3 issue 1, accessed through http://landsat.usgs.gov/about_LU_Vol_3_Issue_1.php, USGS on 23.5.2009.

United States Geological Survey (2009b) *U.S. Geological Survey's digital spectral library*, accessed through http://speclab.cr.usgs.gov/spectral.lib06 on 23.5.2009.

United States Geological Survey (2009c) *What's New section of the USGS Global Visualization Viewer*, accessed through http://glovis.usgs.gov/WhatsNew.shtml on 15.1.2009. For more information refer to http://gloves.usgs.gov or http://earthexplorer.usgs.gov.

United States Geological Survey (2009d), http://earthshots.usgs.gov/Help-GardenCity/RGB-NRG, 20.10.2009.

Vainio M, Kekäläinen H, Alanen A, Pykälä J (2001). *Traditional rural biotopes in Finland – final report of the nationwide inventory*. Finnish Environment Institute, Helsinki, 163 pp. (in Finnish with English summary).

Vescovo L, Gianelle D (2006) Mapping the green herbage ratio of grasslands using both aerial and satellite-derived spectral reflectance. *Agriculture, Ecosystems and Environment* 115:141-149.

Von Numers M, Korvenpää T (2007) 20th century vegetation changes in an island archipelago, SW Finland. *Ecography* 30:789-800.

Vuorela, N (2000) Can data combinations help to explain the existence of diverse landscapes? *Fennia* 178:55–80.

Vuori K, Bäck S, Hellsten S, Karjalainen SM, Kauppila P, Lax HG, Lepistö L, Londesborough S, Mitikka S, Niemelä P, Niemi J, Perus J, Pietiläinen OP, Pilke A, Riihimäki J, Rissanen J, Tammi J, Tolonen K, Vehanen T, Vuoristo H, Westberg V (2006) *The basis for typology and ecological classification of water bodies in Finland*. Finnish Environment Institute, Helsinki, 807 pp (in Finnish).

Wuolijoki A, Iltanen J (eds) (2005) Aino – Suuri suomen kartasto. Genimap, Helsinki/Porvoo, 377 pp. (in Finnish).

Yamano H, Chen J, Tamura M (2003) Hyperspectral identification of grassland vegetation in Xilinhot, Inner Mongolia, China. International Journal of Remote Sensing 24:3171-3178.