A framework for habitat monitoring and climate change modelling: construction and validation of the Environmental Stratification of Estonia

Citation for published version:
Villoslada, M, Bunce, RGH, Sepp, K, Jongman, RHG, Metzger, MJ, Kull, T, Raet, J, Kuusemets, V, Kull, A & Leito, A 2016, 'A framework for habitat monitoring and climate change modelling: construction and validation of the Environmental Stratification of Estonia', *Regional Environmental Change*. https://doi.org/10.1007/s10113-016-1002-7

Digital Object Identifier (DOI):
10.1007/s10113-016-1002-7

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Regional Environmental Change

A framework for habitat monitoring and climate change modelling: construction and validation of the Environmental Stratification of Estonia.

Miguel Villoslada[1], Robert G.H. Bunce[1], Kalev Sepp[1], Robert H.G. Jongman[2], Marc J. Metzger[3], Tiiu Kull[1], Janar Raet[1], Valdo Kuusemets[1], Ain Kull[4] and Aivar Leito[1]

(1) Estonian University of Life Sciences, Kreutzwaldi 5, 51014 Tartu, Estonia
(2) Alterra, Wageningen UR, Droevendaalsesteeg 3, 6708 PB Wageningen, The Netherlands
(3) School of GeoSciences, The University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK
(4) University of Tartu, Vanemuise 46, 51014 Tartu, Estonia

Robert G.H. Bunce
robert.bunce@emu.ee

Kalev Sepp
kalev.sepp@emu.ee

Robert H.G. Jongman
rob.jongman@wur.nl

Marc J. Metzger
marc.metzger@ed.ac.uk

Tiiu Kull
tiiu.kull@emu.ee

Janar Raet
janar.raet@emu.ee

Valdo Kuusemets
valdo.kuusemets@emu.ee

Ain Kull
ain.kull@ut.ee

Aivar Leito
aivar.leito@emu.ee

Corresponding author: Miguel Villoslada
E-mail: mpecina@emu.ee
Phone: (+372) 56890255
Address: Kreutzwaldi 5, room 2C18, 51014 Tartu, Estonia

Abstract

Environmental stratifications provide the framework for efficient surveillance and monitoring of biodiversity and ecological resources, as well as modelling exercises. An obstacle for agricultural landscape monitoring in Estonia has been the lack of a framework for the objective selection of monitoring sites. This paper describes the construction and testing of the Environmental Stratification of Estonia (ESE). Principal components analysis (PCA) was used to select the variables that capture the most amount of variation. Seven climate variables and topography were selected and subsequently subjected to the ISODATA clustering routine in order to produce relatively homogeneous environmental strata. The ESE contains eight strata, which have been described in terms of soil, land cover and climatic parameters. In order to assess the reliability of the stratification procedure for the selection of monitoring sites, the ESE was compared with the previous map of Landscape Regions of Estonia and correlated with five environmental datasets. All correlations were significant. The stratification has therefore already been used to extend the current series of samples in agricultural landscapes into a more statistically robust series of monitoring sites. The potential for applying climate change scenarios to assess the shifts in the strata and associated ecological impacts is also examined.

Key words: climate, geomorphology, clustering algorithm, monitoring, stratified random sampling

Length of the manuscript

Number of words: 5434 words

Figures and tables: 13

Introduction

Environmental stratification is the process that applies multivariate statistical analysis to divide the environmental gradients of a given region into relatively homogeneous units, which can then be used as a framework for sampling both socio-economic and ecological features. Tried-and-tested statistical procedures are used to ensure that the environmental strata are independent of personal bias (Metzger et al. 2005). Commonly, climatic and topographic parameters are used as input variables in the clustering procedure. The resulting environmental strata are relatively homogeneous in terms of the climatic and environmental variables (Klijn and de Haes 1994). These units help in the interpretation of climatic and environmental patterns and thus lead to a better understanding of underlying ecological processes (Jongman et al. 2006).

At present, environmental stratifications have been developed at several levels: global (Metzger et al. 2012, Metzger et al. 2013), continental e.g. Europe (Metzger et al. 2005; Jongman et al. 2006), national e.g. Great Britain (Bunce et al. 1996), Northern Ireland (Cooper 2000), Spain (Elena-Roselló 1997; Regato et al. 1999), Norway (Bakkestuen et al. 2008), Sweden (Ståhl et al. 2011) and the Czech Republic (Fňukalová and Romportl 2014) and regional e.g. Bunce & Smith (1978). The original methodology was published in 1975 (Bunce et al. 1975) and has undergone progressive development since then, as described by Sheail and Bunce (2003). Environmental stratification has primarily been applied in strategic ecological survey projects by using the strata to select statistically representative random samples for surveillance and subsequent monitoring of biodiversity. Environmental stratification has also been used for climate change modelling (Metzger et al. 2008).

National level stratifications have usually been carried out regularly in regions characterized by considerable environmental variability. The main aim of this paper is to demonstrate that it is also feasible to implement environmental stratifications in regions or countries without pronounced topographic and climatic variability. The data required to construct regional level environmental stratifications are usually in more detail than those used at the continental level. This paper describes the construction of the Environmental Stratification of Estonia (ESE) and therefore provides an example of regional level classification. An important step in this process was to explore the data required to cluster the environmental variability of the country into interpretable strata. The suitability of the ESE for modelling possible future ecological changes according to climate change scenarios is also discussed.

Estonia covers 45227 km^2 in the Baltic region of north-eastern Europe between Finland, Russia and Latvia, as shown in Fig. 1. According to the European Environmental Stratification (EnS) (Metzger et al. 2005), Western

Estonia belongs to two classes of the Nemoral Zone, whereas Eastern Estonia is situated in the least cold of the eight classes of the Boreal Zone. Therefore, although a small country, Estonia is located on the boundary between two of the largest EnS classes. According to the Intergovernmental Panel on Climate Change (IPCC) climate change scenarios (Nakicenovic et al. 2000), the border between these zones is likely to shift by 2050. In addition, the Atlantic North Zone may extend into Western Estonia by 2080, as modelled by Metzger et al (2008). The EnS partitions climate variation in Europe, but is not suitable for modelling changes in smaller regions due to the insufficient regional detail of the climate datasets used and because the number of strata produced is not adequate for capturing local environmental gradients. The availability of detailed physiographic and climatic datasets in Estonia facilitates the construction of finer divisions at the national level, as compared to the coarser resolution of the European Zones previously described by Metzger et al. (2005). It has already been recognised that subdivisions of the EnS are needed for local studies. For example, Jongman et al. (2006) subdivided one of the European Environmental Strata in Portugal on the basis of soil types. The existing classification of Estonian landscapes (Arold 2005) was based on the interpretation of geomorphological and soil patterns. The boundaries between the landscape units were descriptive and defined according to expert knowledge, whereas an objective regional classification is required as a framework for landscape and biodiversity monitoring strategies based on a stratified random sampling design. Statistical clustering of the environmental variability into homogeneous units allows deriving reliable estimates on biodiversity, habitats and land cover (Jongman et al. 2006). In this regard, Estonia lacks a robust statistical framework for the selection of biodiversity and vegetation sampling and monitoring plots. The Environmental Stratification of Estonia (ESE) provides the structure needed for such assessment and monitoring strategy, thus the statistical validity of these strata is also examined in this paper.

The present study was initiated in the frame of a multidisciplinary project within the Estonian University of Life Sciences concerning national ecotones and boundaries. A key module in the project is the assessment of the impact of climate change on vegetation and habitats. The aim of the present study is therefore to describe the construction and validation of the Environmental Stratification of Estonia (ESE), which will be used as a basis for the selection of representative sampling sites for recording data on habitats and vegetation. Moreover, the ESE will provide the statistical framework required to upgrade the current Agricultural Landscape Monitoring programme in Estonia. The collected data will then be used in modelling the potential impacts of climate change on the stock and change of biodiversity (Berry et al. 2003; Thuiller et al. 2008). Modelling exercises will also include determining the shifts in the distribution of the strata under different climate change scenarios. The ESE will also be used as a framework to determine the provision of ecosystem services throughout Estonia.

Materials and Methods

Based on previous experience, it was initially decided to examine the potential use of climatic, geomorphological and soil data as input variables to generate the ESE. The data flow was organized in successive steps, as shown in Fig.2. The input variables used in the stratification (Table 1) were selected based on the conceptual model described by Klijn and de Haes (1994), Bunce et al. (1996) and Metzger et al. (2005). The concept is based on a regression model between the environmental strata and the observed ecological parameters. In the functional hierarchy described by Klijn and de Haes (1994), lower components (e.g. vegetation) are dependent on parameters at a higher level (e.g. climate and geomorphology). This hierarchical framework has been recognized by other authors (Godron 1994; Breckle and Walter 2002; Ferrier 2002). At the landscape scale, the variability of environmental conditions is relatively high and the interrelationships between factors that determine this heterogeneity are complex. However, ecosystem patterns and habitat distributions can be analysed using this model even at the national scale.

Climate data

The climate data were interpolated from 26 Estonian meteorological stations, covering a period of 30 years. The data were obtained over the period 1971-2000, which is used nationally as the official period for climate reporting and analysis. In addition, the recording methodology at weather stations in Estonia has been standardized only from 1971 onward. The daily observations at meteorological stations were provided by the Estonian Weather Service. Latvian, Russian and Finnish weather stations were also included in the climate dataset to expand the coverage of the environmental model and provide a more accurate interpretation of the climate in border regions. The climate variables corresponding to the Latvian, Russian and Finnish weather stations were obtained from the E-OBS dataset (Haylock et al. 2008). In order to avoid high correlations and give equal weight to the climate variables, Principal Components Analysis (PCA) was used to generate a subset from a climate dataset composed of 16 parameters (King and Jackson 1999). PCA is a variable reduction procedure that extracts independent components from a large set of variables. PCA identifies the variables that capture the most amount of variation, as well as those that are redundant (Jolliffe 1972; Krzanowski 1987; McCabe 1984). A threshold of 90% of variance explained was used to select the first four components. Subsequently, the two variables with the highest positive and negative loadings were selected from each component. A total of seven climate variables were selected from the initial dataset, as shown in Table 1. This variable selection method has been previously used by Saxon et al. (2005) to generate homogenous climate domains of the continental sector of the United States of America.

The environmental stratification clustering process requires gridded raster layers as input variables. Therefore, the climate data obtained from the weather stations were interpolated into 1x1km raster climate surfaces using the Spline function in ArcGIS 10.1. As a result, seven climate raster grids were produced (Fig. 3). An analogous interpolation procedure has been used by Hijmans et al. (2005) and New et al. (2002).

Geomorphology data

The influence of geographical factors in the distribution and coverage of plant species, even in lowland regions such as Estonia has already been described (Kull et al. 2002; Palo et al. 2008). In order to provide sufficiently detailed information at the local scale in the stratification, geomorphological data were also included by incorporating a digital elevation model, derived from the Estonian LIDAR database. Mean elevation data were calculated within each 1x 1km climate grid cell.

Soil data

At the initial stage of the modelling process, two soil databases were considered for analysis: the Soil Map of Estonia (1:10.000) and the European Soil Database (1:1.000.000) (European Commission 2004). The Soil Map of Estonia proved impractical because of inconsistencies in the definitions of the classes. Before any data could be used, extensive pre-processing would have been required in order to ensure that the classes were consistent throughout the country. In contrast, the coarse resolution of the European Soil Database (ESDB) does not capture the necessary detail required at the regional scale. Moreover, soil information is expressed as categorical classes, which are not compatible with the climate and geomorphology variables expressed as continuous gridded raster layers. Although a transformation of categorical soil data into a continuous grid is possible, the large amount of soil classes in combination with the coarse resolution of the ESDB unbalanced the clustering process, and lead to certain strata being defined exclusively by a unique soil class. The soil data were therefore not included.

The input variables are measured in different units, and some also have large variances which can in turn, have an undesired effect on the resulting clusters. The variables were therefore standardized to zero mean and unit standard deviation.

The variables were subsequently subjected to the ISODATA clustering algorithm to generate the environmental strata. This procedure has been used in comparable studies by Metzger et al. (2005) and Tou and Gonzalez (1974). ISODATA is an iterative algorithm that uses minimum Euclidean distances between each pixel and the closest cluster in the multi-dimensional feature space of the selected variables. The process starts with arbitrary means being assigned to a pre-defined number of clusters. Each raster cell is then assigned to the cluster of which the mean is the closest. The process repeats itself, each raster cell being progressively assigned to the closest cluster in the multidimensional space until no more grid cells are reassigned. The Runtime software program ArcGIS 10.1 was used to perform the analysis. As stated by Memarsadeghi et al. (2007), the main advantage of ISODATA over other clustering procedures is the ability of the algorithm to split large diffuse clusters and to merge small clusters whose centres are closer than a certain threshold. The clustering operation reduces the overall environmental variation into groups with comparable variation around a mean. The number of strata is arbitrary, but each stratum is distinctive and interpretable in terms of its environmental characteristics. The number of strata at which the clustering procedure was stopped was eight. This was considered an interpretable division of Estonia: while reflecting the well-known division between East and West (Lippmaa 1935), the main geomorphological features and contrast between Upland and Lowland regions is appropriately captured by eight strata. In addition, it was observed that the ISODATA algorithm failed to produce clusters when the number was set at ten and above. This could be explained by the fact that the algorithm was not able to create distinguishable clusters above a certain limit. Given the size of Estonia and the main aim of the present study, eight strata is thus considered a practical number for scientific and policy objectives, as well as an adequate reflection of the variation in the environment of the country.

At the last stage of the process, isolated pixels and scattered regions smaller than 15 km^2 were assigned to the closest stratum. In relatively flat regions such as Estonia, these scattered pixels are statistical artefacts of the clustering algorithm, rather than real local features, and the spatial integrity of classes was therefore considered to be the overriding factor (Metzger et al. 2005).

In order to determine the reliability of the stratification, it was necessary to compare this method with independent classifications and to assess the correlations between the clustered input variables and the underlying environmental gradients using the datasets shown in Table 2. The Estonian Landscape Regions classification (Arold 2005) is derived from soil and geomorphological characteristics using expert judgement. Although the classes are not based on statistical reproducible criteria, it is useful to compare their distribution patterns with the ESE to examine the extent of agreement. The Fuzzy Kappa statistic (Hagen 2003) was therefore calculated using Map Comparison Kit v 3.0 (Visser 2004). The objective of the Fuzzy Kappa statistic is to assess the degree of agreement between maps of different classes and the comparisons should thus be treated as similarity coefficients rather than as measures of correlation (Bunce et al. 2002; Klotz et al. 2016).

In order to test the relationships between the ESE and the underlying environmental gradients, regressions were calculated between a number of datasets and the ESE. According to the hierarchical framework previously described, correlations should be present between higher, independent components and the dependent variables. As shown in Table 2, the components derived from the variables used for clustering were correlated with five environmental datasets. Bunce et al. (1996) described classical regression as the most appropriate model to assess the abovementioned correlation. The complex of underlying environmental factors used to create the classes and the selected environmental datasets (Table 2) are the independent and dependent variables respectively. This procedure, referred to as orthogonal regression, has previously been used (Metzger et al. 2005) to assess the validity of the European classification.

Regressions cannot be directly calculated between nominal variables such as Corine land cover and the ESE clustering variables. In order to calculate a multivariate proxy of the land cover classes, the percentages of each class within each stratum were calculated. The first principal component of the land cover percentages was then extracted and correlated with the mean first principal component of the clustering variables within each stratum. Although being directly influenced by human activity, the broad land cover distribution was expected to show a strong relationship with the environmental gradients captured by the ESE. The same procedure was applied to the European Soil Database soil types.

The distribution of plant species was also expected to show significant correlations with the ESE. Species that show well defined distribution patterns in Estonia were chosen for the analysis. The whole flora could not be used because the majority of species that are present throughout the country would provide much background noise in the analysis. Consequently, 26 species were selected as representatives of the Estonian flora, recorded from the 6’x10’ grid used for the Atlas of the Estonian Flora (Kukk and Kull 2005). Binary distribution data were then analyzed by Canonical Correspondence Analysis (CCA) (Ter Braak 1986). The distribution data were fed into the statistical analysis software Canoco 5 to obtain CCA first axis scores for each grid square (Ter Braak and Šmilauer 2012). The mean CCA scores within each stratum were subsequently calculated and then correlated with the mean PCA first axis scores of the stratification for each stratum.

For continuous variables such as topsoil organic carbon, the regression was calculated between the mean score of the first principal component of the classification variables and the mean value of the response variable within each stratum. Cover Management factor is also a continuous variable, therefore the same procedure was applied. Cover Management factor is one of the five factors included in the Revised Universal Soil Loss Equation (RUSLE) and it accounts for the effects of land cover, crops and crop management practices in soil loss (Panagos et al. 2015).

The next stage in the approach is to select a set of randomly located survey sites for sampling biodiversity and landscape monitoring from within each stratum. The procedure used has been described by Metzger et al. (2013) and Carvalho et al. (2015). The design of the sampling framework as well as the number of sampling sites required depends upon the population or the area of habitat or land cover type being sampled. In the present study, the use of the ESE as a framework for monitoring is exemplified by its application in the Agricultural Landscape Monitoring (ALM) programme in Estonia. The number of agricultural landscape monitoring sites in Estonia currently being surveyed as a basis for modelling is only 22, which is statistically unreliable. However, the aim is to increase the number of sites for a long term ALM programme in order to obtain estimates on agricultural land use change and landscape metrics. Stratified random sampling was chosen as the most suitable strategy for the objects of the ALM (Peterseil et al. 2004; Ståhl et al. 2011). The first step of the sampling design was the definition of the target population, which was restricted to all 1 km squares containing agricultural land. Subsequently an agricultural raster mask layer was extracted from the Estonian Basic Map and intersected with a grid of 1 km square resolution specifically created for this process. In order to set the minimum required sampling size, the coefficient of variation of agricultural area within the 1 km squares was defined as the quality constraint (Brus et al. 2011). For any required coefficient of variation, the minimum amount of total sampling units in a stratified random sampling design is defined by Eq. 1 below (de Gruijter et al. 2006):

$$n_{req} = \frac{1}{V_{max}} \left(\sum_{h=1}^{L} N_h S_h(y) \right)^2$$

where n_{req} is the required total sample size, V_{max} is the maximum sampling variance of the total area, N_h is the number of 1km squares in stratum h, S_h is the spatial standard deviation of y within stratum h and y is the land cover class or habitat being sampled. V_{max} is obtained by multiplying the required coefficient of variation by the total area of the population being sampled.

Results

The distribution map of the strata of the ESE is shown in Fig. 4. The boundary between the Nemoral and Boreal Zones (Metzger et al. 2005) is almost precisely reproduced at the border between classes 1-3 and 4-8. This result confirms the significance and stability of this boundary, since the ESE and the European Environmental Stratification were generated from different climate and topography datasets.

Names have been ascribed to the classes, which together with summary information are shown in Table 3. In order to better understand the characteristics of the environmental strata, a brief description of each stratum based on geomorphology, soils and land cover is presented in Table 4.

The ESE was compared with the landscape classification of Arold (2005) and the Fuzzy Kappa comparison yielded a kappa statistic of 0.415, interpreted as "moderate strength of agreement" (Landis and Koch 1977), which is indicative of similarities between the classifications.

The correlations between the selected environmental datasets and the ESE were significant at the 0.05 level. The scatter diagram for the relationship of the Cover Management factor mean value within each stratum with the first PCA axis of the environmental variables is shown in Fig. 5a and Table 2, with an r-value of .77. A summary of the percentage of each soil type within each stratum is provided in Fig. 6

The results of the analysis of the 26 species are shown in Fig. 5b and Table 2, with an r-value of .76. This analysis confirms the role of the principal environmental gradients as determinant factors in the distribution patterns of the flora of Estonia, therefore validating statistically the reliability of the environmental stratification procedure.

Fig. 5c and Table 2 show a correlation of the overall land cover pattern and the underlying environmental gradients with an r-value of .67. A summary of the percentage of each Corine Land Cover class within each stratum is provided in Fig. 7.

Based on the proposed sampling design and an initial coefficient of variation set at 0.1, a total minimum number of 40 1 km monitoring squares was obtained as a basis for monitoring. The minimum required amount of monitoring sites per stratum was therefore set at five, which were assigned to the smallest stratum (Northern Lowlands). The allocation of monitoring sites was subsequently weighted according to stratum size, as described by Haines-Young et al. (2000). This represents the most effective method for reducing the final standard errors of any parameters for which estimates are required. The result was a total sample size of 100 monitoring sites and a final coefficient of variation of 0.06. Further samples can be added later according to the objectives of a given project (Haines-Young et al 2000). Increasing the number of samples does not usually change the total figures but reduces the standard errors (Mateus 2004; Jongman et al 2006).

Table 5 shows the current number of agricultural landscape monitoring sites in Estonia and the number of additional monitoring sites needed per stratum. Further steps in the construction of the stratified random sampling design involve subdividing the strata into equal-area polygons according to the number of sampling sites required per stratum as described by Metzger at al. (2013). Subsequently, a sampling 1 km square will be placed at random within each equal-area polygon. A similar sampling methodology has been successfully implemented in Portugal (Carvalho et al. 2015).

Discussion

The methodology presented in this paper produced eight unbiased environmental classes for Estonia that are based on explicit criteria and explain its environmental variability. The division between eastern and western environmental regions in Estonia (Lippmaa 1935; Laasimer 1965) has been confirmed in the ESE and many of the observed distribution limits of plant species occur along this border between Nemoral and Boreal strata, which is likely to shift under climate change scenarios (Metzger et al. 2008). Consequently Estonia is an optimal location for modelling the impacts of climate change.

The ESE differs from the previous environmental classification (Arold 2005) in having explicit statistical criteria for defining the classes and is therefore independent of personal judgement. The comparison with the classification of Estonian Landscapes confirms the fact that, although based on contrasting conceptual frameworks and datasets, the ESE and the Estonian Landscape Regions reflect the same general environmental patterns. Jones and Bunce (1985) and Metzger et al. (2005) reached the same conclusion with respect to the validity of statistical classification versus intuitive and expert knowledge based procedures, proving the benefits of statistically robust stratification methods. More recently Carvalho et al. (2015) confirmed the value of the approach described in the present paper.

Several regressions were calculated between the ESE and environmental datasets. The distribution patterns of the environmental strata are related to the known distribution of individual plant species, two of which are shown in Fig. 8. *Myrica gale* is a north-western Atlantic species and a major contributor to vegetation cover in the bogs of

Western Britain and Norway. In contrast, *Chamaedaphne calyculata* is a species with affinities with continental conditions, which replaces *Myrica gale,* to some extent, occupying a similar role within bog habitats. Regarding land cover, many factors, such as socio-economic changes, major political decisions and cultural background, have affected its distribution patterns (Mander and Palang 1994; Fuchs et al. 2013). However, the results shown in Fig. 5c demonstrate that the overall pattern is still correlated with the underlying environmental gradients.

The number of strata that is required should be determined according to the objectives of individual projects. Bunce et al. (1996) discussed the use of complex stopping rules, such as testing the variance between the classes and concluded that, in order to obtain statistically reliable results, the most appropriate procedure is to define a minimum size of group that is appropriate for the objectives of the particular project. The divisions made for large regions such as Europe will inevitably be much coarser than for a small country such as Estonia but this does not detract from their value in selection of representative sites, which will be based on the variation within the given domain. The statistical procedure inside the clustering algorithm ensures that the environmental gradients within a given region and the corresponding variation in the data are appropriately clustered in the resulting strata. Additional divisions within a large region can be made for specific objectives. An example of subdividing classes is given by Jongman et al. (2006), who partitioned three EnS strata in the Alpine South zone into six substrata according to altitude in order to capture the full complexity of the Alpine zone from valley floors to summits. Because climate data are continuously variable, there is rarely any natural cut-off point, as is often the case in the analysis of plant taxonomic data.

The climate data used in the construction of the ESE were the best available at the time of analysis. However, when more detailed data becomes available, it could be used to update the stratification in order to improve the definition of boundaries between classes. When the boundaries are shifted, a reassessment of the existing sample and an assessment of the need of additional 1 km monitoring squares are needed. Barr (2011) provides a complete overview on how to proceed when monitoring sites are reallocated. However, Bunce et al. (1996) showed that, in practice, only minor variations are observed through re-classification. In addition, any inefficacies in the strata will be incorporated in the standard errors of the field estimates (Metzger et al. 2005).

In order to make informed decisions, reliable monitoring data derived from statistically robust sampling designs is required (Ortega et al. 2011). In this regard, a main shortcoming of the Agricultural Landscape Monitoring methodology in Estonia has been the lack of a framework for the objective selection of monitoring sites for stock and change of vegetation and habitats. The sampling efficiency is maximized when the population is stratified according to the environmental gradients that define the site characteristics (Jongman et al. 2006). As described in this study, the ESE has been used as a framework to optimize the ALM programme in Estonia in order to obtain more reliable estimates of spatiotemporal trends of land use. The number of monitoring sites was determined based on the coefficient of variation of agricultural land within the 1 km squares in Estonia. As stated by Jongman et al. (2006), improvements in the sampling effort can be made in later stages, once exploratory data has been collected. The results obtained from the representative set of sampling sites can subsequently be extrapolated into national or regional estimates (Bunce et al. 1996; Haines-Young et al. 2000).

Another key objective of the ESE is providing the framework for modelling exercises. Several modelling exercises have been previously performed using environmental stratifications as a framework. For example, Petit et al. (2001) assessed the consequences of environmental change for biodiversity in each of the EnS. On the other hand, Bugter et al. (2011) examined the likelihood of exotic species to survive according to temperature zones defined by the Global Environmental Stratification (Metzger et al. 2012) and climate change models. Leito et al. (2015) used the European stratification as a framework to assess the effects of climate change in wintering and stopover sites of the Eurasian crane (*Grus grus*).

A current on-going project in ecotones and boundaries in Estonia has recently implemented the ESE to examine the potential effects of climate change on habitats and groups of species. In this regard, Liivamägi et al. (2013) showed the changes in the distribution of the Clouded Apollo butterfly (*Parnassius mnemosyne*). The distribution of this species is strongly limited to classes four, five, seven and eight of the ESE, and further work is needed to model the changes in the distribution limits based on climate change models. Moreover, the ESE is currently being used for vegetation and habitat recording from dispersed random squares based on the procedure described in this paper and previously defined by Metzger et al. (2013) and Carvalho et al. (2015). Further applications include the assessment of the provision of certain ecosystem services, utilizing the environmental strata as units for stratified random sampling.

The European environmental stratification has previously been used to evaluate the potential impact of climate change in the provision of ecosystem services (Metzger et al. 2006). The climate change scenarios created for Estonia show a range of results, depending on the General Circulation Models and IPCC storylines adopted. However, mean increases of 10-20% in annual precipitation and a mean warming by 2.3–4.5°C are projected by the end of the 21st Century (Jaagus and Mändla 2014). The implications for biodiversity and ecosystem services provision in Estonia have yet to be determined but the ESE will make such analyses possible. Consequently, further research will involve coupling climate change models (climate change simulations have already been calculated in Estonia by Luhamaa et al. (2014) with the Environmental Stratification of Estonia. This approach has already been tested by Metzger et al. (2008). By incorporating climate change predictions in the stratification as input data, the future distribution of the strata can be quantified in terms of the direction and extent of change. The results of such analyses will in due course enable the estimation of the potential changes in ecological resources and the provision of ecosystem services in Estonia.

Conclusions

The Environmental Stratification of Estonia (ESE) was constructed using climate and geomorphological data and applying standard statistical procedures. The classification has been tested and correlated with environmental data sets, demonstrating that the strata are representative of the principal underlying environmental gradients. Because the strata are determined statistically and independently of personal judgement, the ESE provides the framework for optimizing the existing Agricultural Landscape Monitoring programme in Estonia, in order to obtain statistically robust figures. Furthermore, the ESE will provide the background for modelling the effects of climate change on habitats, species distribution and the provision of ecosystem services.

References

Arold I (2005) Estonian landscapes. University of Tartu, Tartu

Bakkestuen V, Erikstad L, Halvorsen R (2008) Step-less models for regional environmental variation in Norway. J Biogeogr 35: 1906–1922. doi: 10.1111/j.1365-2699.2008.01941.x

Barr CJ (2011) The Sampling Strategy for Countryside Survey (up to 2007). CEH Lancaster, Lancaster

Barredo JI, San Miguel J, Caudullo G, Busetto L (2012) A European map of living forest biomass and carbon stock. Executive report. Joint research centre of the European Commission, Ispra. doi: 10.2788/780

Berry PM, Dawson TP, Harrison PA, Pearson R, Butt N (2003) The sensitivity and vulnerability of terrestrial habitats and species in Britain and Ireland to climate change. J Nat Conserv 11: 15-23. doi: 10.1078/1617-1381-00030

Breckle SW, Walter H (2002) Walter's vegetation of the earth: the ecological systems of the geo-biosphere, 4th edn. Springer, Berlin

Brus DJ, Knotters M, Metzger MJ, Walvoort DJJ (2011) Towards a European-wide sampling design for statistical monitoring of common habitats. Alterra Report 2213, Wageningen

Bugter RFJ, Ottburg FGWA, Roessink I, Jansman HAH, Van der Grift EA, Griffioen AJ (2011) Invasion of the turtles? Exotic turtles in the Netherlands: a risk assessment. Alterra Report 2186. Alterra Wageningen UR. http://www.wageningenur.nl/en/Publication-details.htm?publicationId=publication-way-343039323632. Accessed 17 March 2015

Bunce RGH, Barr CJ, Clarke RT, Howard DC, Lane AMJ (1996) Land classification for strategic ecological survey. J Environ Manage 47: 37–60. doi: 10.1006/jema.1996.0034

Bunce RGH, Carey PD, Elena-Rosello R, Orr J, Watkins J, Fuller R (2002) A comparison of different biogeographical classificatios of Europe, Great Britain and Spain. J Environ Manage 65: 12-134. doi: 10.1006/jema.2002.0533

Bunce RGH, Morrel SK, Stel HE (1975) The application of multivariate analysis to regional survey. J Environ Manage 3: 151–165

Carvalho SB, Gonçalves J, Guisan A, Honrado JP (2015) Systematic site selection for multispecies monitoring networks. J Appl Eco. doi: 10.1111/1365-2664.12505

Cooper A (2000) Land cover monitoring in Northern Ireland. In: Rushton BS (ed) Biodiversity: the Irish dimension. Royal Irish Academy, Dublin, pp 122–131

de Gruijter JJ, Brus DJ, Bierkens MFP, Knotters M (2006) Sampling for Natural Resource Monitoring. Springer, Berlin.

EEA (2000) Corine land cover technical guide – addendum 2000. Report no. 40. European Environment Agency, Copenhagen

Elena-Rosselló R (1997) Biogeoclimatic classification of the Spanish regions of Iberian Peninsula and Balearic islands. Ministry of Agriculture, Fisheries and Food, Madrid (in Spanish)

European Commission (2004) The European Soil Database distribution version 2.0 (CD-ROM) European Commission and the European Soil Bureau Network. http://eusoils.jrc.ec.europa.eu/ESDB_Archive/ESDBv2/index.htm. Accessed 17 March 2015

Ferrier S (2002) Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning: Where to from Here? Syst. Biol. 51(2): 331-363. doi: 10.1080/10635150252899806

Fuchs R, Herold M, Verburg PH, Clevers JGPW (2013) A high-resolution and harmonized model approach for reconstructing and analysing historic land changes in Europe. Biogeosciences 10: 1543–1559. doi:10.5194/bg-10-1543-2013

Fňukalová E, Romportl D (2014) A typology of natural landscapes of Central Europe. AUC Geographica 49 (2): 57-63. doi: 10.14712/23361980.2014.15

Godron M (1994) The natural hierarchy of ecological systems. In: Klijn F (ed) Ecosystem classification for environmental management. Kluwer Academic Publishers, Dortdrecht, pp 69-83

Hagen A (2003) Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci 17(3): 235–249. doi:10.1080/13658810210157822

Haines-Young RH, Barr CJ, Black HIJ, Briggs DJ, Bunce RGH, Clarke RT, Cooper A, Dawson FH, Firbank LG, Fuller R, Furse MT, Gillespie MK, Hill R, Hornung M, Howard DC, McCann T, Morecroft MD, Petit S, Sier ARJ, Smart SM, Smith GM, Stott AP, Stuart R, Watkins JW (2000) Accounting for nature: assessing habitats in the UK countryside. DETR, London

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation. J Geophys Res (Atmospheres) 113, D20119. doi:10.1029/2008JD10201

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965-1978. doi:10.1002/joc.1276

Jaagus J, Kaupo M (2015) Climate change scenarios for Estonia based on climate models from the IPCC Fourth Assessment Report. Est J Earth Sci 63 (3): 166-180. doi: 10.3176/earth.2014.15

Jolliffe IT (1972) Discarding Variables in a Principal Component Analysis I: Artificial Data. J Roy Stat Soc C-App 21(2): 160-173. doi: 10.2307/2346488

Jones HE, Bunce RGH (1985) A preliminary classification of the climate of Europe from temperature and precipitation records. J Environ Manage 20: 17- 29

Jones RJA, Hiederer R, Rusco E, Loveland PJ, Montanarella L (2004) The map of organic carbon in topsoils in Europe, Version 1.2, September 2003: Explanation of Special Publication. Ispra 2004 No.72 (S.P.I.04.72). Office for Official Publications of the European Communities. http://139.191.1.96/ESDB_Archive/octop/OCtopMapBkLet76.pdf. Accessed 17 March 2015

Jongman RHG, Bunce RGH, Metzger MJ, Mücher CA, Howard DC, Mateus VL (2006) A statistical Environmental Stratification of Europe: objectives and applications. Land Ecol 21: 409–419. doi: 10.1007/s10980-005-6428-0

King JR, Jackson DA (1999) Variable selection in large environmental data sets using principal components analysis. Environmetrics 10: 67-77. doi: 10.1002/(SICI)1099-095X(199901/02)10:1%3C67::AID-ENV336%3E3.0.CO;2-0

Klijn F, de Haes HAU (1994) A hierarchical approach to ecosystems and its implications for ecological land classification. Land Ecol 9 (2): 89–104. doi: 10.1007/BF00124376

Klotz M, Kemper T, Geiß C, Esch T, Taubenböck H, (2016) How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. Remote Sens Environ 178: 191-212. doi: 10.1016/j.rse.2016.03.001

Krzanowski WJ (1987) Selection of Variables to preserve Multivariate Data Structure, using Principal Components. J Roy Stat Soc C-App 36(1): 22-33. doi: 10.2307/2347842

Kukk T, Kull T (2005) Atlas of the Estonian flora. Estonian University of Life Sciences, Institute of Agricultural and Environmental Sciences, Tartu (in Estonian)

Kull T, Kukk T, Leht M, Krall H, Kukk Ü, Kull K, Kuusk V (2002) Distribution trends of rare vascular plant species in Estonia. Biodivers Conserv 11: 171-196. doi: 10.1023/A:1014564521151

Laasimer L (1965) Plant cover of Estonian S.S.R. Valgus, Tallinn (in Estonian)

Landis JR, Koch GC (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159–74. doi: 10.2307/2529310

Leito A, Bunce RGH, Külvik M, Ojaste I, Raet J, Villoslada M, Leivits M, Kull A, Kuusemets V, Kull T, Metzger MJ, Sepp K (2015) The potential impacts of changes in ecological networks, land use and climate on the Eurasian crane population in Estonia. Land Ecol 30(5): 887–904. doi: 10.1007/s10980-015-0161-0

Liivamägi A, Kuusemets V, Luig J, Kask K (2013) Changes in the distribution of Clouded Apollo *Parnassius mnemosyne* (Lepidoptera: Papilionidae) in Estonia. Entomol Fennica 24(3): 186–192

Lippmaa T (1935) Main features of Estonian geobotany. University of Tartu, Tartu (in Estonian).

Luhamaa A, Kallis A, Mändla K, Männik A, Pedusaar T, Rosin K (2014) Eesti tuleviku kliima stsenaariumid aastani 2100. Estonian Environment Agency

Mander Ü, Palang H (1994) Changes of landscape structure in Estonia during the Soviet period. GeoJournal 33(1): 45 - 54. doi: 10.1007/BF00810135

Mateus VL (2004) Countryside Land Cover: Strategic Sample Survey. Universidade de Évora, Évora

McCabe GP (1984) Principal variables. Technometrics 26(2): 137-144. doi: 10.1080/00401706.1984.10487939

Memarsadeghi N, Mount DM, Netanyahu NS, le Moigne J (2007) A fast implementation of the Isodata clustering algorithm. Int J Comput Geom Appl 17(1): 71-103

Metzger MJ, Brus DJ, Bunce RGH, Carey PD, Gonçalves J, Honrado JP, Jongman RHG, Trabucco A, Zomer R (2013) Environmental stratifications as the basis for national, European and global ecological monitoring. Ecol Indic 33: 26–35. doi: 10.1016/j.ecolind.2012.11.009

Metzger MJ, Bunce RGH, Jongman RHG, Mücher CA, Watkins JW (2005) A climatic stratification of the environment of Europe. Global Ecol Biogeogr 14: 549–563. doi: 10.1111/j.1466-822X.2005.00190.x

Metzger MJ, Bunce RGH, Jongman RHG, Sayre R, Trabucco A, Zomer R (2012) A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring. Global Ecol Biogeogr 22(5): 630–638. doi: 10.1111/geb.12022

Metzger MJ, Bunce RGH, Leemans R, Viner D (2008) Projected environmental shifts under climate change: European trends and regional impacts. Environ Conserv 35(1): 64–75. doi: 10.1017/S0376892908004529

Metzger MJ, Rounsevell MDA, Leemans R, Schröter D (2006) The vulnerability of ecosystem services to land use change. Agr Ecosyst Environ 114: 69–85. doi: 10.1016/j.agee.2005.11.025

Mücher CA, Hennekens SM, Bunce RGH, Schaminée JHJ, Schaepman ME (2009) Modelling the spatial distribution of Natura 2000 habitats across Europe. Land Urb Plan 92 (2): 148-159. doi:10.1016/j.landurbplan.2009.04.003

New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. Climate Res 21: 1–25. doi: 10.3354/cr021001

Ortega M, Metzger MJ, Bunce RGH, Wrbka T, Allard A, Jongman RHG, Elena-Roselló R (2011) The potential for integratio of environmental data from regional stratifications into a European monitoring framework. J Environ Plann Man 55 (1): 39-57. doi: 10.1080/09640568.2011.575698

Palo A, Linder M, Truu J, Mander Ü (2008) The influence of biophysical factors and former land use on forest floristic variability on Saaremaa and Muhu islands, Estonia. J Nat Conserv 16: 123-134. doi: 10.1016/j.jnc.2008.08.001

Panagos P, Borrelli P, Meusburger ., Alewell C, Lugato E, Montanarella L (2015) Estimating the soil erosion cover-management factor at European scale. Land Use Policy 48(c): 38-50. doi:10.1016/j.landusepol.2015.05.021

Peterseil J, Wrbka T, Plutzar C, Schmitzberger I, Kiss A, Szerencsits E, Reiter K, Schneider W, Suppan F, Beissmann H (2004) Evaluating the ecological sustainability of Austrian agricultural landscapes-the SINUS approach. Land Use Pol 21: 307-320. doi: 10.1016/j.landusepol.2003.10.011

Petit S, Firbank L, Wyatt B, Howard D (2001) MIRABEL: models for integrated review and assessment of biodiversity in European landscapes. Ambio 30: 81–88. doi: 10.1579/0044-7447-30.2.81

Regato P, Castejón M, Tella G, Giménez S, Barrera I, Elena-Rosselló R (1999) Cambios recientes en los paisajes de los sistemas forestales de España. Invest Agr. Sist Recur For 1: 383-398.

Saxon E, Baker B, Hargrove W, Hoffman F, Zganjar C (2005) Mapping environments at risk under different global climate change scenarios. Ecol Lett 8: 53–60. doi: 10.1111/j.1461-0248.2004.00694.x

Sheail J, Bunce RGH (2003) The development and scientific principles of an environmental classification for strategic ecological survey in Great Britain. Environ Conserv 30: 147–159. doi: 10.1017/S0376892903000134

Ståhl G, Allard A, Esseen PA, Glimskär A, Ringvall A, Svensson J, Sture Sundquist S, Christensen P, Gallegos Torell Å, Högström M, Lagerqvist K, Marklund L, Nilsson B, Inghe O (2011) National Inventory of Landscapes in Sweden (NILS) - Scope, design, and experiences from establishing a multi-scale biodiversity monitoring system. Environ Monit Assess 173: 579-595. doi: 10.1007/s10661-010-1406-7

Ter Braak CJF (1986) Canonical Correspondence Analysis: A New Eigenvector Technique for Multivariate Direct Gradients Analysis. Ecology 67(5): 1167-1179. doi: 10.2307/1938672

Ter Braak CJF, Šmilauer P (2012) CANOCO reference manual and user's guide: software for ordination (version 5.0). Microcomputer Power, Ithaca, New York

Thuiller W, Albert C, Araujo MB, Berry PM, Cabeza M, Guisan A, Hickler T, Midgley GF, Paterson J, Schurr FM, Sykes MT, Zimmermann NE (2008) Predicting global change impacts on plant species' distributions: Future challenges. Perspect Pant Ecol 9: 137-152. doi: doi:10.1016/j.ppees.2007.09.004

Tou JT, Gonzalez RC (1974) Pattern recognition principles. Addison-Wesley Publishing Company, Massachusetts

Van der Sluis T, Pedroli B, Kristensen SBP, Lavinia Cosor G, Pavlis E (2015) Changing land use intensity in Europe – Recent processes in selected case studies. Land Use Policy. doi: 10.1016/j.landusepol.2014.12.005

Visser H (2004) The map comparison kit: methods, software and applications. RIVM report 550002005/2004. RIVM (National Institute for Public Health and the Environment, the Netherlands). http://www.pbl.nl/sites/default/files/cms/publicaties/550002005.pdf. Accessed 17 March 2015.

List of figures

Fig. 1 Map of Estonia in relation to surrounding countries with the primary cities and lakes

Fig. 2 Flow chart describing the main stages in the construction of the environmental stratification of a region. PCA = Principal Components Analysis. ISODATA = multivariate method for classification of component values for sites

Fig. 3 Selected climate variables and elevation used in the ISODATA procedure to generate the environmental strata. The gridded climate raster layers were interpolated from weather stations using the Spline function in ArcGIS 10.1. The climate maps represent monthly averages for the period 1971-2000

Fig. 4. Distribution and short descriptive names for the eight Environmental Strata in Estonia

Fig. 5 Orthogonal regression plots between the mean values for each stratum of the first principal component of Principal Components Analysis (PCA) of the variables used for constructing the Environmental Stratification of Estonia (ESE) and comparable values from three independent data sets. (a) Mean values of the Cover Management factor within each stratum compared with the ESE. (b) Mean values for the Canonical Correspondence Analysis (CCA) scores for 26 vascular plants for each stratum compared with the ESE. (c) Mean values of the PCA values for Corine Land Cover classes for each stratum compared with the ESE.

Fig. 6 Distribution of the European Soil Database soil types (European Commission 2004) within each stratum of the ESE. Strata names are given in Fig. 4 and table 3

Fig. 7 Distribution of Corine land cover 2006 classes (aggregated at the second level) within each stratum of the ESE. Strata names are given in Fig. 4 and table 3

Fig. 8 Distribution of two species of vascular plants in Estonia as representatives of other species in the country. *Chamaedaphne calyculata* (a) and *Myrica gale* (b) overlaid on the strata shown in Fig. 3. The distribution of these species has been previously published (Kukk and Kull 2005)