Multivariate clustering of reindeer herding districts in Sweden according to range prerequisites for reindeer husbandry

Henrik Lundqvist & Öje Danell

Reindeer Husbandry Unit, Swedish University of Agricultural Sciences, Box 7023, S-750 07 Uppsala, Sweden (corresponding author: henrik.lundqvist@rene.slu.se).

Abstract: The 51 reindeer herding districts in Sweden vary in productivity and prerequisites for reindeer herding. In this study we characterize and group reindeer herding districts based on relevant factors affecting reindeer productivity, i.e. topography, vegetation, forage value, habitat fragmentation and reachability, as well as season lengths, snow fall, ice-crust probability, and insect harassment, totally quantified in 15 variables. The herding districts were grouped into seven main groups and three single outliers through cluster analyses. The largest group, consisting of 14 herding districts, was further divided into four subgroups. The range properties of herding districts and groups of districts were characterized through principal component analyses. By comparisons of the suggested grouping of herding districts with existing administrative divisions, these appeared not to coincide. A new division of herding districts into six administrative sets of districts was suggested in order to improve administrative planning and management of the reindeer herding industry. The results also give possibilities for projections of alterations caused by an upcoming global climate change. Large scale investigations using geographical information systems (GIS) and meteorological data would be helpful for administrative purposes, both nationally and internationally, as science-based decision tools in legislative, economical, ecological and structural assessments.

Key words: Reindeer husbandry, multivariate techniques, PCA, cluster analysis, seasonal ranges, *Rangifer tarandus tarandus*.

Rangifer, 27 (2): 107-119

Introduction

The Swedish reindeer herding area (Fig. 1) covers about half of Sweden's total area. It stretches from the herding districts Idre in the south to Könkämä in the north, extending from approximately 12ºE to 24ºE and 62ºN to 69ºN. The total area is 227 000 km^2, of which 221 000 km^2 is within Sweden (see County Administration Boards, 2000). The reindeer herding area is divided into 51 reindeer herding districts, some of which overlap. The reindeer herding districts are managed by reindeer herding communities, to which reindeer owners are associated as members. In 2003 there were 940 enterprises with 4596 reindeer owners (Swedish Board of Agriculture, 2005). The herding districts are composed of ranges and resources used during all seasons, but vary considerably in conditions for reindeer husbandry as well as productivity (Statistics Sweden, 1999; Swedish Board of Agriculture, 2005). The herding often includes extended migrations between summer and winter ranges, as the original migration routes extended from the Baltic Sea coast of Sweden to the North Sea coast of Norway. The Swedish-Norwegian border today blocks these routes and the grazing rights across the border are decided upon the agreement of the Swedish-Norwegian reindeer grazing commission (Svensk-Norska Renbeteskommissionen, 2001).

The herding districts are grouped into mountain, forest and concession herding districts. Generally the mountain herding districts are found all along the Scandic mountain range, the forest herding districts in the central east and the concession herding districts in the northeast. The mountain herding districts have typically outstretched shapes, with summer ranges in the mountain area and winter ranges in the inland and coastal areas towards the southeast. The southern mountain districts, however, have smaller ranges and more limited east-west extensions (County Administration Boards, 2000). In the forest and concession herding districts, the seasonal division of the ranges is less distinct and some herding districts do not have separate seasonal lands. The difference between these two groups is mainly administrative and legal. In the concession herding districts, non-Sami also have grazing rights and own reindeer but the concession for reindeer herding is given to Sami reindeer herders.

The bottlenecks of production are obviously not the same for all herding districts and are hitherto only partly and verbally or qualitatively described, for example in official investigations (e.g., Svensk-Norska Renbeteskommissionen, 2001). These differences in conditions can be associated with a large number of interacting factors, many of which are properties of the ranges (Lundqvist *et al.*, 2007). A parsimonious and operational summary of the conditions could be to develop a zonation of the reindeer herding area relevant on the district level, based on the multivariate background of conditions, similar to e.g. Fairbanks (2000). Currently the reindeer herding districts are only categorised by herding logistics and administrative/legal aspects such as the previously mentioned mountain, forest and concession herding districts, or according to county borders, rather than according to ecological conditions. The county boards usually base their decisions, e.g. maximum herd sizes allowed, on historical facts and figures, but less on a truly assessed reasoning.

Our objective was to examine, in a multivariate perspective, the quantitative variables suggested to

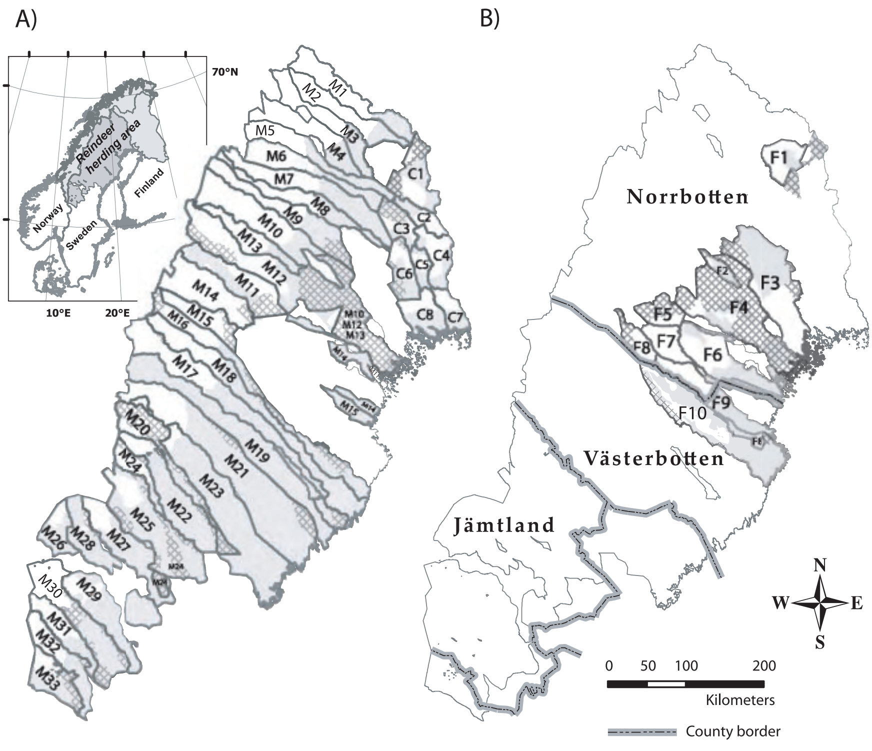

Fig. 1. The Fennoscandian reindeer herding area (shaded area in the small map) and the 51 reindeer herding districts in Sweden. The concession herding districts (C) and mountain herding districts (M) are shown in map A) and the forest herding districts (F), as well as the administrative county borders (Norrbotten, Västerbotten and Jämtland), are shown in map B). The winter ranges are shaded in the large maps. Some seasonal ranges are overlapping (grey cross-striped) and are usually shared by two neighbouring herding districts. Reindeer herding district F4 shares its whole-year ranges with the winter ranges used by districts M10, M12 and M13. See Appendix for names of the herding districts.

Table 1. Overview of the 15 variables used for clustering the reindeer herding districts.

Variable	Abbrev	Range[1)]	Method	Note[2)]	Data source
Elevation	Elev	SSA	Mean value of elevation grid		GSD[3)] – The Terrain Elev. DB (National Land Survey of Sweden, 2002a)
North facing slopes	SlopeN	SSA	Proportion of slopes with northward direction (330-30°)	Aspect function in ArcGIS	
Ruggedness	RugSSA RugW	SSA, W	Mean ruggedness index – raw data	(Riley *et al.*, 1999)	“
Growth season length	GrowthSea	SSA	Linearly interpolated mean, of days/year with average temp > 5°C	30-year average	National Atlas of Sweden (National Land Survey of Sweden, 2002b)
Snow season length	SnowSea	W	Linearly interpolated mean, of days/year with snow cover	“	“
Snow precipitation	SnowPrec	W	Linearly interpolated mean of products of annual precipitation and proportion of snow precipitation	“	“
Ice crust probability	IceSSA IceW	SSA, W	Ordinary cokriging interpolation, average ice-crust formation incidence (days/year)	10-year average (Table 2)	SMHI[4)] – data: temperature and precipitation
Harassing insect abundance probability	MI12h	SSA	Ordinary cokriging interpolated average of accumulated - Mörschel Index (days/year)	“	SMHI[4)] – data: temperature and wind
Forest with lichen	ForestL	W	Area of forests containing lichen		GSD[3)] – Land Cover Data (National Land Survey of Sweden, 2004)
Uncut forest	UncutFor	W	Percentage uncut forest		“
Winter forage	ForageW	W	Proportion of winter forage	Classified reindeer forage (Table 3)	REN2000[5)] (County Administration Boards, 2000)
Road density	LinStrWgh	W	Weighted total lengths of linear structures	Weights in Table 4	GSD[3)] – Roads from the Blue Map (National Land Survey of Sweden, 1998)
Reachability	ReachEd	SSA	3 classes of forage, linear structure weights as LinStrWgh, with edge effects	Lundqvist *et al.* (2007) Vegetation and road class. (Table 3 & 4)	REN2000[5)] and GSD[3)] – Roads from the Blue Map

[1)] SSA = Spring/summer/autumn ranges, W = winter ranges.
[2)] Slope and interpolation methods are found in ArcGIS and are grid functions using the moving window procedure.
[3)] GSD – Geographical Data Sweden.
[4)] SMHI – Swedish Meteorological and Hydrological Institute (2005).
[5)] REN2000 – Reindeer Husbandry Database.

affect reindeer productivity on a herding district scale using GIS layers on resources, topography and infrastructure, as well as relevant meteorological data covering area studied. Our approach was to determine the long-term characteristics, categorise the herding districts, and cluster similar districts into an applicable number of groups. Furthermore, we identified their characteristics by describing the variable compositions of the resulting groups and outliers, and compared the resulting groups with present categorization and administrative divisions. Here we are also trying to assess the current pattern on its validity using scientific principles.

Material and methods

The analyses were based on a previously developed detailed mapping (10 km × 10 km square raster with 1958 cells) of the entire reindeer herding area for 15 variables (Table 1, Lundqvist *et al.*, 2007) expected to affect productivity in reindeer husbandry. These variables included climate and weather aspects, topographic features, vegetation and fragmentation by infrastructures. The 15 variables were extracted from an original list of 37 variables capturing factors, which were suggested to be highly relevant in literature on *Rangifer* ecology. A detailed account of data sources and variable derivation is found in the reference above.

The total range for each reindeer herding district was divided into spring-summer-autumn (SSA) and winter (W) ranges. The extensions of the seasonal ranges of the reindeer herding districts were extracted from the Reindeer Husbandry Database - REN2000 (County Administration Boards, 2000). The mapped variables were used for calculation of mean variable values for the SSA (7 variables, see Table 1) and W (8 variables) ranges in each herding district. Extraction of variables from the data grids for the seasonal ranges of the herding districts was done with the zonal function in ArcGIS™ (ESRI, 2005).

The topographical variables included mean elevation, north-facing slope ratio and mean ruggedness. Water bodies larger than 0.25 ha were excluded to avoid influences of water surfaces on the topography indices. The mean elevation and the north-facing slope ratio were derived only for SSA areas, as the ecological impacts on reindeer of these variables on winter ranges were not proven in literature.

Growth season length was defined as the number of days per year with average temperature over 5 C° and applied on the SSA ranges. Snow season length was defined as the number of days with snow cover. Yearly snow precipitation was calculated as the product of the annual average precipitation and the proportion of annual snow precipitation in the meteorological data.

The ice crust formation probability was based on minimum, average and maximum temperatures as well as precipitation at 58 meteorological stations, and was derived from daily weather data from 1993-2002 according to weather events described in Table 2. The average number of ice crust days per year and weather station were interpolated using ordinary cokriging and applied on both winter and SSA areas, since ice-crust can be formed as early as late autumn, when reindeer still may be in SSA areas, and remain in the spring areas until the snow has melted.

Table 2. Description of weather events used for calculation of ice-crust formation probability. Excerpt from (Lundqvist *et al.*, 2007).

Event	Condition	Argument
1	3 days (average temp > 1 °C) followed by 2 days (average temp < -1 °C)	Snow melts and get packed with increasing density when temperature above 0 °C. Since snow has high temperature capacitivity due to its high water content, high temperatures are needed or for a longer time. The same goes for low temperatures. Probably a deeper and thicker ice crust is formed during this event.
2	Day when max temp > 2 °C min temp < -2 °C max temp – min temp > 10 °C	When irradiation and albedo are high, usually during clear skies, the surface of the snow melts and freezes which causes ice crust formation. Probably not too deep in the snow cover.
3	Day when max temp > 2 °C min temp < -2 °C precipitation > 3 mm	Wet precipitation, which freezes directly or shortly after reaching the ground, can create an ice crust, which also can lead to firn layers in the snow pack.
4	Day when average temp > 1 °C precipitation > 3 mm followed by day with average temp < -1 °C	As above but in a longer time perspective.

A cumulative warble fly activity index was computed from temperature and wind data (see Mörschel, 1999) recorded every third hour at 58 meteorological stations during the snow-free season from 1993-2002, and interpolated using ordinary cokriging.

Uncut forest ratio was defined as the ratio of uncut forest area in relation to the total forest area. For grading winter forage quality we used the reindeer forage classifications of REN2000 (County Administration Boards, 2000; Table 3). The reachability values (Lundqvist, 2007) combined available green forage with effects of fragmentation. These were estimated for the SSA ranges with a sample point distance of 10 km and a 50 m resolution using the SSA forage quality classifications in Table 3 and the relative fragmenting weights of the linear structures in Table 4. Depth of edge effects on forage patch edges was set to 150 m. Reachability values were not calculated for W ranges due to less predictable effects of roads on foraging during winter. Instead, a weighted road density index was developed where the road classes obtained from the roadmap were used, assuming that vicinities of roads were avoided in the herding because of traffic hazards and risk of spreading the herd. The weightings per unit road length were those shown in Table 4.

Statistical analyses

Our analyses were based on means derived for each seasonal range. Normalization of the data was done using power transformations (Box & Cox, 1964) with the exponents rounded to the nearest 0.5. The 15 variables and 51 observations (herding districts) were analysed with principal component analysis (PCA) and cluster analysis (CA) using the SAS (SAS Institute Inc., 2005) and SIMCA-P (Umetrics, 2004) statistical software. We used PCA to identify in which sense herding districts (and subsequently the groups) differed from one another. The number of PCs extracted was according to the commonly suggested Kaiser–Guttman criterion and Cattell Scree plot (e.g., Hair *et al.*, 1998; McGarigal *et al.*, 2000).

To identify groups in of the 51 herding districts we used four different CA methods: Ward's method, single linkage, average linkage, and complete linkage (e.g., Johnson & Wichern, 1998), all with standardized Euclidean distances. The average linkage method was used as the main CA method, and single and complete linkages as well as Ward's method were used to support the identification of strong and weak groups without losing the presence of outliers and to control the bias of "noisy" data. Single linkage method connects the closest observations between clusters in the multidimensional space and hence focuses on dissimilarities, which results in many outliers. Complete linkage, on the other hand, connects the furthest observations of every cluster with focus on similarities and results in strong groups. Average linkage method, which connects the average position of every cluster, and Ward's method, which strengthens similarities and weakens dissimilarities, were therefore given the highest attention.

We determined the number of clusters by dendrograms and the cubic clustering criterion (CCC) (Sarle, 1983) derived from the CA linkage methods. To compare the internal disparity of the distinct groups, the root-mean-square standard deviation (RMSSD) of each

Table 3. Classification and weighting of vegetation types as reindeer green forage (from REN2000: The County Administration Boards, 2000).

Vegetation class	Forage quality	Forage value
Winter		
Coniferous and birch forest – heath type w. lichens Dry/Extremely dry heath	Very good	1
Other	None	0
Spring, summer and autumn		
Wet rich marshland/bog Dry rich marshland/bog Meadow with herbs	Very good	3
Coniferous forest - rich heath type Birch forest - rich heath type Fresh heath Grass heath Scrubland/Willow	Good	2
Wet poor marshland/bog Dry poor marshland/bog Marshland/bog with trees Coniferous and birch forest – heath type w. lichens Dry/Extremely dry heath	Less good	1
Other	None	0

Table 4. Friction values for linear structures used in the reachability variable and weights per unit of structure length used in the linear structure index.

Linear structures	Friction values	Weight values
Railroads	100	20
Highways – Nat. roads	100	20
Public roads	50	10
Private roads	10	5
Trails	2	2
Streets	100	20

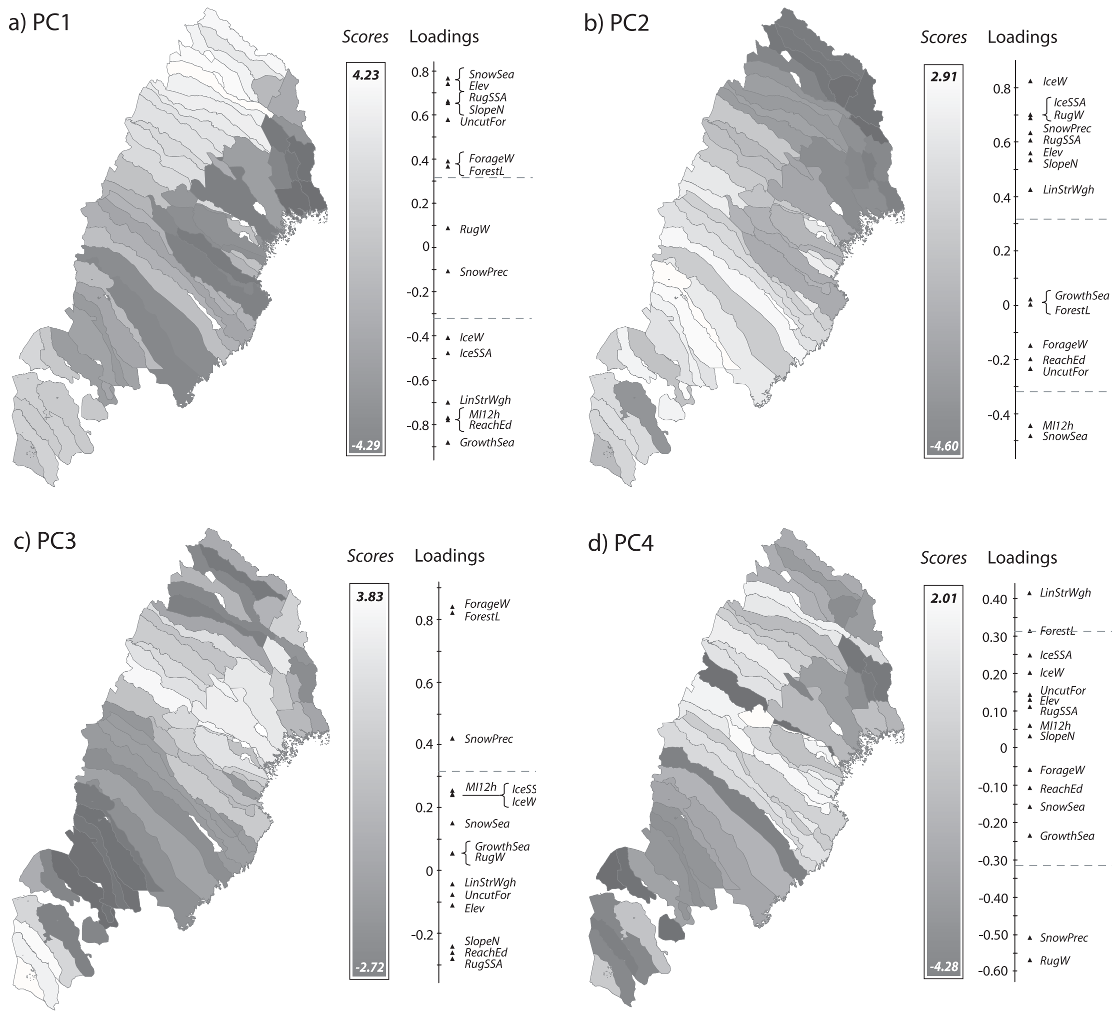

Fig. 2a-d. Geographical distribution of PC scores of the reindeer herding districts based on the 15 retained variables in the in the first four principal components. Maximum and minimum score values of the PCs are in italics inside the legends. The loadings of the 15 variables in each PC are shown on the vertical axis. The grey dotted lines represent the factor loading value (±0.32), which accounts for 10% of the variance in the component (e.g., McGarigal *et al.*, 2000). Loadings with absolute values exceeding 0.13 are considered statistically significant.

cluster was calculated, where low RMSSD values indicate more homogeneous herding districts within a group. To distinguish the characteristics of each resulting group, each group's PCA score contribution was compared with the average score of all observations for each variable.

Results

Principal Component Analysis

The correlation matrix of the 15 variables is shown in Table 5. Four PCs were extracted according to the Kaiser–Guttman criterion, as the first four PCs exceeded unity, and the Cattell scree plot, which showed an "elbow" at the fourth component. These four PCs explained 37%, 25%, 13% and 7.5% of the compound variance of the 15 variables between herding districts, in total 83%.

Figs. 2a–d show the geographical distributions of the scores together with the loadings of the first four PCs. The first PC reveals, quite expectedly, a dominant east-west gradient with growth and snow season lengths negatively correlated to each other. The topographical and winter forage variables, together with uncut forests, determined the scores together with snow season lengths in the northwest as well as the southwest. High values of reachability, roads, insect harassment, ice-crust and growth season length were

Table 5. Correlation matrix of reindeer herding district mean values of the 15 variables included in the analyses. Variable abbreviations are explained in Table 1.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Elev	1														
2	SlopeN	0,75	1													
3	RugSSA	0,87	0,73	1												
4	RugW	0,30	0,34	0,39	1											
5	GrowthSea	-0,64	-0,61	-0,57	0,03	1										
6	SnowSea	0,24	0,24	0,17	-0,18	-0,66	1									
7	SnowPrec	0,14	0,18	0,11	0,64	0,25	-0,25	1								
8	IceSSA	0,01	0,04	0,05	0,32	0,42	-0,61	0,47	1							
9	IceW	0,14	0,13	0,15	0,44	0,33	-0,69	0,56	0,93	1						
10	MI12h	-0,89	-0,75	-0,80	-0,35	0,63	-0,29	-0,09	0,22	0,07	1					
11	ForestL	0,23	0,05	0,04	-0,08	-0,36	0,34	0,10	0,06	0,10	-0,08	1				
12	UncutFor	0,21	0,31	0,24	-0,11	-0,59	0,57	-0,21	-0,31	-0,41	-0,27	0,13	1			
13	ForageW	0,10	-0,03	-0,08	-0,04	-0,27	0,50	0,19	-0,20	-0,17	-0,04	0,79	0,14	1		
14	LinStrWgh	-0,25	-0,21	-0,15	0,04	0,52	-0,85	0,14	0,65	0,65	0,35	-0,17	-0,39	-0,38	1	
15	ReachEd	-0,61	-0,43	-0,55	-0,12	0,64	-0,51	-0,07	0,17	0,15	0,61	-0,45	-0,44	-0,46	0,42	1

apparent, particularly in the southeast and east. The second PC identified a north-south latitudinal gradient where high ice-crust formation probability, snow precipitation and topographical variables were positively combined and apparent in the central and southern areas. Snow season length and insect harassment were negatively related to these and apparent in the north and northeast. The third PC was clearly dominated by winter forage and lichen-rich forests, together with snow precipitation. The forest herding districts and the southernmost mountain herding districts had high scores in this PC and were also spatially fairly assembled. On the other hand, the districts M20-M25 together with M29 (see Fig. 1 for codes) and some of the northern mountain herding districts showed the opposite tendency in these variables. The fourth PC showed no geographical gradient and the pattern seems random. Here ruggedness in the winter ranges together with snowfall and roads together with lichen-rich forest achieved the most extreme loadings. In this PC, several districts showed clearly different scores relative to their neighbours.

Fig. 2 shows that adjacent herding districts generally were similar in their characteristics. Furthermore, assembled and small districts stand out as having more extreme scores. Oblong and extended districts achieve more intermediate scores in many variables. In PC 2, we found M29 to be unique for its geographical position. In PC 3, the districts C3, M11 and M31-33 differ markedly from their neighbouring districts. In PC 4, C3-5, M11, M19, M26 and M28 stood out at the negative side of the score scale, in contrast to F5, F8 and F9, which stood out at the positive side of the scale.

Cluster analyses

The average linkage clustering method was the selected main clustering method (Fig. 3). The other linkage methods agreed in large extent with the average linkage method, with small differences in the order the clusters were distinguished and in the number and placement of outliers. With the average and complete linkage methods, the CCC suggested a clustering into ten groups and Ward's method only eight groups. Another clustering into six groups with the average linkage method was also possible according to CCC, but not as significant as the ten-group division. The CCC on single linkage suggested one group only, probably due to the chaining effect of single linkage (Everitt & Dunn, 2001). Only one outlier (M29) fell out in the eight group clustering suggested by Ward's method. Fig. 4 shows the geographical distributions of the herding district groups.

As indicated in Fig. 3, groups I–V and VI–IX formed two main clusters and group X was an outlying group consisting of herding district M26. In the first large cluster (groups I–V), the concession districts together with the northernmost forest district were easily distinguished as group I and II, and the remaining forest districts formed group III, in which districts F2 and F3 could be considered distinct. The district C2 in group I was considered as an outlier with the average linkage method but not with the other linkage methods. Group IV, consisting of 14 centrally located mountain herding districts, was suggested to be further divided into four subgroups (IVa – IVd) in order to increase the resolution of this large group. Subgroups IVa, IVb and IVc consisted of adjacent herding districts. The most distinct subgroup (IVd) was district M28, which could be considered as an outlier within group IV. The district M29 (V) was an outlier in this first large cluster. The other large cluster (group VI – IX) contained the northernmost (VI and VII) and southern mountain herding districts (VIII). District M11 (IX) was an outlier in the second large cluster.

The average score's contributions of each group in the first four PCs in the PCA are shown in Table 6. To describe the homogeneity within each group, the table also lists the RMSSD of each group from the cluster analysis. Characteristics of the groups are shown in Table 7.

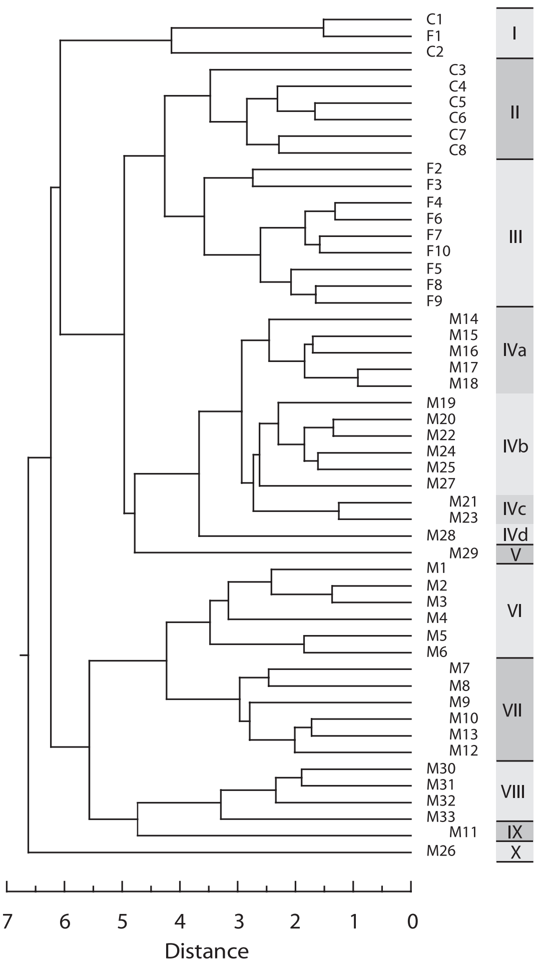

Fig. 3. Dendrogram of CA of the 51 herding districts using the average linkage method with normalized variables and standardized Euclidean distances. The group division is based on the cubic clustering criterion (CCC) by Sarle (1983), which suggested a division into ten groups.

Discussion

Combining GIS with spatial information on habitat resources offers ample potential for science-based investigations, e.g. on sustainability and management of natural resources. The spatial information and results can be used for administrative purposes, both nationally and internationally, and be used for science-based decision tools in legislative, economical, ecological and structural planning and assessments. This is called for more than ever as we are facing a global climate change and its impact on societal, economical and ecological systems as well as natural resource based industries such as tourism, silviculture, agriculture and pastoralism, which are increasingly competing with each other.

Classification and clustering habitats or herding districts, based on factors affecting the animal's biol-

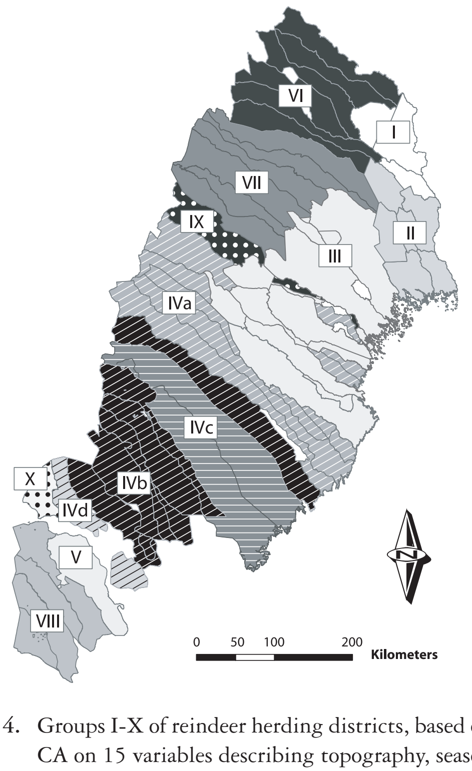

Fig. 4. Groups I-X of reindeer herding districts, based on CA on 15 variables describing topography, season lengths, snow precipitation, ice crust probability, harassing insect activity, vegetation and forage, forestry and range fragmentation. Group IV (striped on map) was further divided into four subgroups due to its large size. Note the overlapping areas shown in Fig. 1.

Table 6. Score contribution in the PCA of each group relative to total average, and root-mean-square standard deviation (RMSSD) of each cluster in the average linkage CA. The score contributions of the PCA are the differences, in scaled units, between the average of the group and the average of all herding districts. Bold and bold shaded numbers highlight differences larger than 1.5 and 3, respectively. For abbreviations, see Table 1.

	Groups												
	I	**II**	**III**	**IVa**	**IVb**	**IVc**	**IVd**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**
No. of districts	3	6	9	5	6	2	1	1	6	6	4	1	1
Variables	*PCA Score contribution*												
Elev	-1,44	**-1,98**	-0,64	0,53	0,39	-0,07	0,32	0,87	0,67	0,84	0,62	0,47	0,04
SlopeN	-0,71	**-1,92**	-0,29	0,38	0,88	-0,14	0,42	-0,35	0,53	0,81	-0,05	0,05	0,37
RugSSA	**-2,02**	-1,31	-0,76	0,58	0,85	-0,31	1,07	0,74	0,61	0,88	0,22	0,51	-0,08
RugW	**-2,35**	-0,34	-0,33	0,21	0,97	0,94	**2,88**	-1,31	-0,45	-0,17	0,35	**3,36**	**5,24**
GrowthSea	0,06	1,42	0,71	-0,22	0,07	0,60	0,33	0,56	**-1,87**	**-1,64**	0,11	-0,86	0,70
SnowSea	0,86	-0,69	-0,49	-0,61	-0,58	-1,05	-0,43	0,47	**1,65**	0,97	-0,07	**1,60**	0,32
SnowPrec	**-1,87**	-0,11	0,27	0,34	0,38	-0,18	**1,85**	-1,41	-1,19	0,00	1,26	3,57	5,72
IceSSA	**-2,17**	-0,44	0,83	0,80	0,74	1,32	-0,56	-0,13	**-1,63**	-0,07	0,07	0,51	-0,74
IceW	**-2,65**	-0,50	0,47	1,10	0,99	**1,56**	-0,13	-0,96	**-1,73**	-0,04	0,55	0,21	-0,73
MI12h	0,84	**1,85**	0,81	-0,16	-0,53	0,12	-0,72	-0,49	-1,09	-0,63	-0,63	0,06	-0,47
ForestL	0,01	-0,41	0,52	0,17	**-1,99**	-0,61	**-2,31**	**-1,97**	0,02	1,08	**3,20**	1,24	**-1,81**
UncutFor	-0,39	-0,43	-0,03	-0,09	-0,26	-0,53	-0,09	-0,01	1,01	1,14	-0,40	0,35	0,58
ForageW	0,17	-0,15	0,27	-0,33	**-2,48**	-1,03	-0,65	-0,86	0,10	0,81	**3,51**	1,46	0,05
LinStrWgh	-0,58	0,48	0,91	1,00	0,21	0,54	0,22	-0,43	-1,35	-0,67	-0,16	**-1,95**	**-1,77**
ReachEd	0,47	1,22	0,14	0,10	0,36	0,72	-0,44	-0,12	-1,37	-0,94	-0,58	-0,85	-0,38
RMSSD (CA)	0.63	0.53	0.52	0.32	0.40	0.23	-	-	0.57	0.50	0.51	-	-
					------------- 0.51 --------------								

ogy, are informative when making relevant comparisons between areas. Identification of possibilities and constraints for reindeer herding when grouping herding districts also gives us better understanding of limitations of ranges and differences between the herding districts. The variables included in the study were all considered important for the reindeer and their productivity (see Lundqvist *et al.*, 2007), and are qualitative since quantitative variables, such as district area, do not tell us about the quality of ranges. Districts have often been developed towards oblong extensions to contain sufficient and ecologically different seasonal ranges. With these, the migratory behaviour of the reindeer reinforced with active herding are fundamental measures in the utilisation of the districts' resources and overcome limitations. The heterogeneous characteristics of the districts also provide flexibility, which is essential in the mitigation of year-to-year variation in conditions.

PCA and CA are common techniques but assume linearity. We can imagine that if we examine a geographical variable such as elevation vs. reindeer productivity we would find an optimum at which the productivity is maximised. Since we don't include causality and only investigate the herding area based on similarities and dissimilarities to group the herding districts we assumed that linear methods would be satisfactory. The biological relevancy of the resulting groups should however be considered. This could perhaps be done using methods such as artificial neural network (e.g., Stern, 1996) to assess "the curse of dimensionality" which always are present in multivariate analyses, especially with few observations such as the number of reindeer herding districts. Other methods of linking observations and defining clusters could be e.g. Random Forest (Breiman, 2001) or Classification and Regression Trees (CART, e.g., Breiman *et al.*, 1984; De'ath, 2002), but are not assessed in this study.

Since the observations contained high and diverse variances, the PCA suggested several linear dimensions to decompose the variance of the data, as shown in Fig. 2. Here 84% of the total variance was however explained by only four linear dimensions instead of the 15 observed variables, which is accomplished by the fact that the variables are correlated. If all variables would be totally uncorrelated, PCA would not be able to reduce the dimensionality other than by chance. The main difference between this study and variable reduction by Lundqvist *et al.* (2007) was

Table 7. Most significant characteristics of the identified groups of reindeer herding districts. Predicted impacts are indicated with + (positive impact), - (negative impact) and ? (uncertain impact), based on ecological knowledge on reindeer in literature.

		Characteristics	
Group	Herding districts	Spring, summer, autumn areas	Winter areas
I	C1-C2 F2	? Fairly low elevation - Low ruggedness + Low ice crust probability	- Low ruggedness + Fairly low snow precipitation + Low ice crust probability + Fairly low road density
II	C3-C8	? Very low elevation ? Few north-facing slopes - Low ruggedness + Long growth season - Very high insect activity + Fairly high reachability	
III	F2-F10	? Fairly low elevation - Fairly low ruggedness + Fairly long growth season - Fairly high ice crust probability - Fairly high insect activity	- Fairly high road density
IVa	M14-M18	- Fairly high ice crust probability	- Fairly high ice crust probability - Fairly high road density
IVb	M19-M20 M22 M24-M25 M27	? Fair amount north-facing slopes + Fairly high ruggedness	+ Fairly high ruggedness - Fairly high ice crust probability - Very low amount of lichen-rich forest - Very low forage value
IVc	M21 M23	- Fairly high ice crust probability	+ Fairly high ruggedness + Fairly short snow season - Fairly high ice crust probability - Fairly low forage value
IVd	M28	+ Fairly high ruggedness	+ Very high ruggedness - Fairly high snow precipitation - Low amount of lichen-rich forest
V	M29	? Fairly high elevation	- Fairly low ruggedness + Fairly low snow precipitation + Fairly low ice crust probability - Low amount of lichen-rich forest
VI	M1-M6	- Very short growth season + Low ice crust probability + Fairly low insect activity - Fairly low reachability	- Long snow season + Fairly low snow precipitation + Low ice crust probability + Low forestry intensity + Fairly low road density
VII	M7-M10 M12-M13	? Fairly high elevation ? Fair amount north-facing slopes + Fairly high ruggedness - Short growth season - Fairly low reachability	- Fairly long snow season + Fairly high amount of lichen-rich forest + Fairly low forestry intensity + Fairly high forage value + Fairly low road density
VIII	M30-M33		- Fairly high snow precipitation + Very high amount of lichen-rich forest + Very high forage value
IX	M11	- Fairly short growth season	+ Very high ruggedness - Fairly long snow season - Very high snow precipitation + Low road density + Fairly high amount of lichen-rich forest + Fairly high forage value
X	M26		+ Very high ruggedness - Very high snow precipitation - Low amount of lichen-rich forest + Low road density

the scale of the data, i.e. the size, form and number of observations; Lundqvist *et al.* used 1958 equally sized squares while this study was applied on 51 herding districts. Applying a study using seasonal ranges of herding districts, the means of the variables for such large areas become more close to the average of the total area, hence shrinking extreme values toward common averages. This would explain the difference with lower values in loadings of each PC in this study compared to Lundqvist *et al.*

In CA like this, the number of evident clusters could be a matter of discussion, especially as the districts are more or less related in a continuum of different dimensions. Tools to define the number of clusters, such as the CCC by Sarle (1983), may give guidance, but the final cluster quantity should be decided according to the field of application and practical considerations, i.e. in this case an appropriate number of district groups for administrative purposes. A division into too many or too few clusters could fail to achieve a useful and relevant division for this purpose. Divisions into many clusters can result in a large number of outliers while divisions into few clusters may force disparate observations together. A compromise between the number of groups and group sizes is important. However, foremost the biology of the animal studied and the data drive the classification, and the classification should be assessed as such. Therefore, the average linkage method was chosen over the others to find and define outliers. Other alternative groupings according to the complete linkage method was a division of group III into two clusters, where F2 and F3 were clustered together with C3 of group II and M14 of group IVa joined group VIII. All these herding districts were outliers or loosely joined with unclear groups in single linkage. This suggests a closer examination of these herding districts if the underlying variables were to be altered. Groups consisting of several herding districts naturally achieve a score contribution average closer to the total average than single outliers do, so score contributions of small groups should not be overrated. However, to distinguish weaker clusters and their characteristics within the large group IV, a further division into subgroups was made.

The main pattern of the herding district clusters showed similarities to the original classification of mountain, forest and concession herding districts. The mountain herding districts are, however, rather diverse and were grouped in seven clusters from north to south. The northernmost mountain herding districts (group VI, VII and IX) and the southernmost one (group VIII) were similar, assembled around alpine areas and distinct compared to other groups. The central mountain herding districts, mostly consisting of the districts of group IV, were more like the forest and concession herding districts. The concession herding districts were divided into a southern group (II) and a northern group (I), to which the northernmost forest district (F1) was joined. The outliers (V, IX, X and IVd) were geographically located in the borderland between the two major clusters and surrounding the large group IV. M11 (group IX) and M29 (group V) in particular could be described as intermediates of these two large clusters.

The administrative division of herding districts inherits the category division (C, F and M), but the mountain herding districts are currently further divided between the three counties Norrbotten, Västerbotten and Jämtland (Fig. 1b). Norrbotten, the northernmost county, includes group VI, VII, IX and the two northernmost districts of group IVa. Västerbotten, the central county, includes the remaining districts of group IVa, group IVc and the northernmost district (M19) of group IVb. Jämtland includes the remaining districts of group IVb, IVd, V, VIII and X. By this means, we conclude that the mountain herding districts of Norrbotten and Jämtland counties are heterogeneous while those in the county of Västerbotten are more similar to each other and to adjacent districts outside the county. Thus, a county-wise division does not seem optimal according to reindeer prerequisites and characteristics of the herding districts.

Lundqvist *et al.* (2007) suggested a tentative zonation of ranges based on their properties using 1958 10 km × 10 km observations, independent of the herding district pattern. When comparing the reindeer herding groups with the zonation of the rangeland done by Lundqvist *et al.*, groups I and II are mainly within zone C and E, respectively. Group IV is mainly within zones A and C, group X within zones A and B, and group III within zones D and E. All other groups stretched across at least three zones, which implied that these groups consisted of more heterogeneous landscapes and therefore are expected to have more average scores in several variables. These groups also consisted of mountain herding districts that were more outstretched in their geographical extension and consequently included many different biotopes. Differences found between adjacent herding districts could therefore also be due to their different shapes and distributions of seasonal ranges. Small herding districts are often subject to distinctiveness due to their limited extension and marginal possibilities of having backup areas of different types, which would level out their perhaps extreme characteristics.

The results of this study suggest a new, but a more homogenous, division of the herding districts that are based on the relevant characteristics for reindeer, which could improve planning and decision-making for the industry and its administration. For administrative purposes and to include the outliers in larger sets, we hence suggest a northern set consisting of group VI and a northwestern set consisting of groups VII and IX, a northeastern set consisting of groups I and II and a central eastern set consisting of group III. Finally, we suggest a central western set consisting of group IV and a southern alpine set consisting of groups V, VII and X. This should be done keeping in mind the characteristics of each district and group, and the understanding of the group affiliation in case of a change in prerequisites and conditions. The group affiliations of the districts may change due to future changes in climate, vegetation, production strategies, and further complications resulting from land use conflicts with landowners, new legislations and other competing industries. Therefore, updated characterisations of the herding area should be done periodically, and each time with best available techniques, e.g. updated statistical algorithms and latest high-resolution data.

To further improve the landscape characterization and investigate the causality of the variables, productivity information such as carcass weights, meat production, meat quality and reproductive success need to be combined with the resource data. Such data are however quite variable in itself and need to be collected in a consistent manner. On the other hand, this would give the opportunity to investigate the effect of the temporal variables, which were handled only as long-term means in this study and the previous one by Lundqvist *et al.* (2007). Moreover, such data should integrate not only the prerequisites and conditions of the ranges, but also the effects of predators and traffic accidents, as well as management measures such as animal densities, slaughter strategies, herd demography, supplementary feeding and parasite treatments. Including such data would allow for a truly holistic approach assessing cumulative impacts and sustainability. The complexity of this was however to large to be included in this study, but will be treated in a forthcoming paper (Lundqvist *et al.,* in prep). The data and results in this paper and Lundqvist *et al.* (2007) also form a basis for projections of future changes in resources and condition for reindeer husbandry in a climate change perspective.

Acknowledgements

We thank Lennart Norell for many valuable discussions during the course of this study. The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS) is acknowledged for financial support.

References

Box, G. E. P. & Cox, D. R. 1964. An analysis of transformations. – *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 26: 211-243.

Breiman, L. 2001. Random Forests. – *Machine Learning* 45: 5-32.

Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. 1984. *Classification and Regression Trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, USA.

County Administration Boards. 2000. Ren2000 [ArcIMS] - The Reindeer Husbandry Database. The reindeer husbandry units of the counties BD, AC and Z of Sweden. [Online] http://www.Ren2000.se

De'ath, G. 2002. Multivariate Regression Trees: A New Technique for Modeling Species-Environment Relationships. – *Ecology* 83: 1105-1117.

ESRI. 2005. ArcGIS v. 9.2 (ArcInfo software). Redlands, USA.

Everitt, B. S. & Dunn, G. 2001. *Applied Multivariate Data Analysis*. Arnold, a member of the Hodder Headline Group, London. 342pp.

Fairbanks, D. H. K. 2000. Physio-climatic classification of South Africa's woodland biome. – *Plant Ecology* 149: 71-89.

Hair, J. F. J., Anderson, R. E., Tatham, R. L. & Black, W. C. 1998. *Multivariate Data Analysis*. Prentice-Hall, Inc., Upper Saddle River, New Jersey. 730pp.

Johnson, R. A. & Wichern, D. W. 1998. *Applied Multivariate Statistical Analysis.* Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA. 816pp.

Lundqvist, H. 2007. Ecological cost-benefit modelling of herbivore habitat quality degradation due to range fragmentation. Accepted manuscript. – *Transaction in GIS*.

Lundqvist, H., Norell, L. & Danell, Ö. 2007. Multivariate characterisation of environmental conditions for reindeer husbandry in Sweden. – *Rangifer* 27: 3-21.

McGarigal, K., Cushman, S. & Stafford, S. 2000. *Multivariate Statistics for Wildlife and Ecology Research*. Springer-Verlag New York, Inc., New York. 283pp.

Mörschel, F. M. 1999. Use of climatic data to model the presence of Oestrid flies in caribou herds. – *Journal of Wildlife Management* 63: 588-593.

National Land Survey of Sweden. 1998. Geographical Sweden Data - Roads from the Blue Map. Metria.

National Land Survey of Sweden. 2002a. Geographical Sweden Data - The Terrain Elevation Databank. Metria.

National Land Survey of Sweden. 2002b. National Atlas of Sweden. Metria. [Online] http://www.sna.se/webbatlas/index.html

National Land Survey of Sweden. 2004. Geographical Sweden Data - Land cover data. Metria.

Riley, S. J., DeGloria, S. D. & Elliot, R. 1999. A terrain ruggedness index that quantifies topographic heterogeneity. – *Intermountain Journal of Sciences* 5: 23-27.

Sarle, W. S. 1983. Cubic Clustering Criterion – Technical Report A-108: SAS Institute Inc., Cary, NC (US).

SAS Institute Inc. 2005. SAS v. 9.1 software. Cary, NC, USA.

SMHI. 2005. Meteorological data - temperature, wind, cloudiness, precipitation, snow cover. Swedish Meteorological and Hydrological Institute.

Statistics Sweden. 1999. *Swedish Reindeer Husbandry*. In Swedish: *Svensk rennäring*. Statistiska centralbyrån, Svenska Samernas Riksförbund, Jordbruksverket, Sveriges Lantbruksuniversitet. 149pp.

Stern, H. S. 1996. Neural networks in applied statistics. – *Technometrics* 38: 205-214.

Swedish Board of Agriculture. 2005. *The reindeer husbandry in numbers 2004*. In Swedish: *Rennäringen i siffror 2004*. [Online] http://www.sjv.se/amnesomraden/djurveterinar/rennaring

Svensk-Norska Renbeteskommissionen. 2001. Report / published by The Swedish-Norwegian Reindeer Grazing Commission. In Swedish: Betänkande / avgivet av Svensk-norska renbeteskommissionen av år 1997. – *In*: Svensk-norska renbeteskommissionen (ed.). Ministry for Foreign Affairs, p. 173.

Umetrics. 2004. SIMCA-P v. 10.5 software. Umeå, Sweden.

Manuscript received 21 May, 2007
accepted 20 August, 2007

Appendix

Names of coded reindeer herding districts.

Concession herding districts		Forest herding districts		Mountain herding districts		Mountain herding districts (cont.)	
C1	Muonio	F1	Vittangi	M1	Könkämä	M18	Ran
C2	Sattajärvi	F2	Serri	M2	Lainiovuoma	M19	Vapsten
C3	Tärendö	F3	Gällivare	M3	Saarivuoma	M20	Frostvikens norra
C4	Korju	F4	Udtja	M4	Talma	M21	Vilhelmina norra
C5	Pirttijärvi	F5	Ståkke	M5	Gabna	M22	Ohredahke
C6	Ängeså	F6	Östra Kikkejaure	M6	Laevas	M23	Vilhelmina södra
C7	Liehittäjä	F7	Västra Kikkejaure	M7	Girjas	M24	Raedtievaerie
C8	Kalix	F8	Maskaure	M8	Baste	M25	Jiingevaerie
		F9	Mausjaure	M9	Sörkaitum	M26	Kall
		F10	Malå	M10	Sirges	M27	Jovnevaerie
				M11	Luokta-Mávas	M28	Njaarke
				M12	Tuorpon	M29	Tåssåsen
				M13	Jåhkågasska	M30	Handölsdalen
				M14	Semisjaur-Njarg	M31	Mittådalen
				M15	Svaipa	M32	Ruvhten
				M16	Gran	M33	Idre Nya
				M17	Ubmeje		

Multivariat gruppering av svenska samebyar baserat på renbetesmarkernas grundförutsättningar

Abstract in Swedish / Sammanfattning: Svenska renskötselområdet består av 51 samebyar som varierar i produktivitet och förutsättningar för renskötsel. Vi analyserade variationen mellan samebyar med avseende på 15 variabler som beskriver topografi, vegetation, betesvärde, fragmentering av betesmarker, klimat, skareförekomst och aktivitet av parasiterande insekter och vi föreslår en indelning av samebyarna i tio grupper. Den största gruppen, som bestod av 14 samebyar, delades vidare in i 4 undergrupper. Klusteranalyser med 4 olika linkage-varianter användes till att gruppera samebyarna. Principalkomponentsanalys användes för att kartlägga undersökta variabler och de resulterande samebygruppernas karaktär. Samebygrupperna följde inte länsgränser och tre samebyar föll ut som enskilda grupper. Denna undersökning ger underlag för jämförelser mellan samebyar med beaktande av likheter och olikheter i fråga om produktivitet och funktionella särdrag istället för länsgränser och historik. Vi föreslår en ny administrativ indelning i sex områden som skulle kunna fungera som ett alternativt underlag för planering och beslut som rör produktionsaspekter i rennäringen. Resultaten ger också underlag för förutsägelser av förändringar i samebyars produktionsförutsättningar till följd av klimatförändringar.