Journal of Environmental Management (2002) **64**, 13–23
doi:10.1006/jema.2001.0481, available online at http://www.idealibrary.com on IDEAL®

Multivariate ordination approach for identification of sub-regional homogeneities in Gujarat, western India

A. M. Dixit* **and C. P. Geevan**

Gujarat Institute of Desert Ecology, 106, GSFC Township, Fertilizer Nagar, Vadodara - 391 750, Gujarat, India

Received 6 November 2000; accepted 8 June 2001

Identification of homogeneous invariant geographical units assumes considerable importance in the context of current attempts at developing environmental indices to monitor changes. Data on the edaphic and climatic parameters for the state of Gujarat – a predominantly arid to semi-arid province in western India – has been used to elaborate an approach that employs principal component analysis (PCA) as the basis for unbiased clustering of similar geographic units at sub-regional scales. PCA identified different sets of edaphic and climatic variables responsible for separating one unit from another. With the exception of a few cases, soil variables exhibited greater influence on the separation of units than climate. Response of some agro-ecological variables (total cultivable area, forestland and normalized difference vegetation index) to both original and ordinated variables in each of the identified edaphic-climatic units (ECU) has been evaluated using step-wise multiple regressions. The edaphic-climatic parameters were good predictors of these variables, although ordination did not necessarily enhance the predictive power of the independent variables. Instead, in many cases original variables showed better correlations. The response of these variables, chosen on the basis of varying levels of dynamism (slow to seasonal fluctuations), showed that independent of the degree of dynamism, the dependence of these variables was distinct in each ECU. This provided a reality check on the agro-ecological and bio-climatic differentiation among the ECUs.

© 2002 Academic Press

Keywords: edaphic, climatic zones, PCA, western India.

Introduction

The interplay of a variety of independent and mutually dependent factors – predominantly of climate regime and soil – determines the biogeographic and ecological characteristics of different regions and the assemblages of smaller entities that constitute larger zones. The different ecological zones are distinguished by different climates, landforms, soils, vegetation and land-use (Schultz, 1995). The major efforts at identifying zones with similar climate-vegetation relationship are based on correlations between the two (Holdrige, 1947; Brisse and Grandjouan, 1978). Agro-climatic zones are usually characterised by a climate regime suitable for a range of crops and cultivars (FAO, 1983) while an ecological region is marked by distinct ecological responses to a climatic regime as manifested in the type of vegetation cover, faunal and floral assemblages.

At present there are several classifications for India such as biogeographic regions (Rodgers and Panwar, 1988), climatic (Carter, 1954) and agro-ecological (Murthy and Pandey, 1978; Subramaniam, 1983; ARPU, 1998; Sehgal *et al*., 1992). While some of these are based on qualitative criteria, others are based on a system of assigning scores and, to an extent, arbitrary weights to some attributes that are important for a particular purpose such as agricultural development. To a very large extent they have remained mostly descriptive and qualitative. In addition, most zoning efforts ignore the high correlation of parameters with each

* Corresponding author. Email: arunmdixit@hotmail.com; cpgeevan@yahoo.com

0301–4797/02/010013+11 $35.00/0 © 2002 Academic Press

other. Quantitative methods such as multivariate analysis, which reduce both correlations and subjectivity, have rarely been employed in India.

Approach

The identification of homogeneous geographic units (HGUs) could, perhaps, be regarded as a necessary pre-requisite for monitoring environmental change at the regional or provincial level. Such units could be based, at the first level of generalisation, on edaphic and climatic factors. Changes within the HGUs become comparable over time provided they are defined using attributes that may be expected to remain nearly invariant over relatively long time scales (>25 years). The need for such invariant land units assumes considerable importance in the context of current attempts at developing environmental indices to monitor change and compare the environmental risk faced by different regions. Bearing in mind the different data sources available in India for evaluating environmental change, we adopted three selection criteria: (a) the availability of time series data in the existing system of data collection and reporting, (b) units which are neither too small nor too large and (c) the existence of benchmark data on relatively invariant attributes such as edaphic and climatic factors.

Districts – the administrative units often used in the zoning schemes in India – are too large and highly heterogeneous from the point of view of environmental characteristics to be of much use for the grouping of similar units. However, approaches that are based entirely on natural boundaries suffer from paucity of time series data. The *taluka* or *tehsil* is an administrative unit between district and village level for which a large amount of time series data is available to monitor land-use and other environmental changes. It has also been observed that although administrative divisions are not directly based on natural divisions, they do to some extent reflect natural topographic and climatic variation (Nagendra and Gadgil, 1998).

Considering the need to have long-term constancy in the parameters used for identification of HGUs, edaphic and climatic factors are possibly the best candidates and are often used in many classification schemes. Barring climate change and/or large cataclysmic events, these are rarely expected to show year-to-year variations. In this work, we have identified HGUs in the state of Gujarat in western India by applying principal component analysis (PCA). The usefulness of multivariate techniques for defining HGU has been demonstrated by previous workers. For instance, Bunce *et al*. (1975) used a reciprocal averaging ordination method for land classification in the Lake District (UK) while Retuerto and Carballeira (1991) employed multivariate methods for defining phytoclimatic units in Spain. The HGU we have devised is based on edaphic and climate parameters and is referred to henceforth as the Edaphic-Climatic Unit (ECU). We have also explored the response of some agro-ecological variables in each ECU. These variables are chosen on the basis of their scale of dynamism: (a) total cultivable area, which changes slowly over time, (b) forest area, that changes in the short to medium term and (c) vegetation index, that varies seasonally.

Gujarat

The State of Gujarat extends over an area of 195 984 km^2, between 20°1 to 24°7 N and 68°4 to 74°4 E (see Figure 2). The province was recently reorganized into 25 districts consisting of 223 *talukas*. However, the data used in the present paper correspond to the former administrative boundaries of 19 districts and 184 *talukas* for which time series data are available. Atmospheric conditions are influenced mostly by the monsoon climate regime, physiography, insularity and the neighbouring Thar Desert to the North. Most areas have sub-humid and semi-arid climates, merging in the arid zone in the north and north-west. Large parts of the province are drought-prone. Distribution of temperature and relative humidity is more or less uniform. Summer maximum varies from 37°C in the coastal region to 42°C in the interior. The mean winter minimum ranges between 10°C to 8°C (Anon, 1987). The variation in annual rainfall is very high in the western arid region, decreasing gradually towards the south-east. According to the benchmark soil survey conducted by the National Bureau of Soil Survey and Land Use Planning, soils of Gujarat belong to 5 orders, 11 sub-orders, 20 great-groups and 45 sub-groups (NBSS & LUP, 1994).

Methodology

Data used

The data sets associated with nearly 1400 polygons of different soil types was extracted from 1:500 000 scale maps of NBSS & LUP (1994) and overlaid with the administrative boundaries of 184 *talukas* of Gujarat State using Idrisi32 and Cartalinx GIS software (Clark Labs, Clark University: www.clarklabs.org). Saline desert areas like the Rann of Kachchh were excluded from the analysis. We estimated the proportional area coverage of each *taluka* under various classes of eight edaphic parameters. Appropriate scores were assigned to different classes of each parameter. Except for soil texture and drainage, the scores correspond to the mid-point of the class interval (Table 1). The Weighted Sum of Scores (WSS) for each parameter was obtained by summing the scores weighted by proportional area covered.

The mean annual rainfall (MAR) for each *taluka* was derived from data for 98 years (1901–1998). Two more rainfall-related parameters – coefficient of variation (CV) and mean number of rainy days – were estimated from 30 years of data (1968–1998). Temperature is an important parameter for delineation of ECU. However, for the region, the spatial variation in temperature regime is low with the isotherms for annual mean varying between approximately 26°C and 28°C (Anon, 1987). Since the distance between isotherms ranges from 100 km to over 400 km, temperature gradients are small. Therefore, we have not included temperature parameters in the analysis. In total, therefore, 11 edaphic and climate parameters were used. All values were transformed to a common scale of 0 to 100 (see Rubin, 1998).

Delineation of ECU

We carried out clustering using orthogonal, composite edapho-climatic axes produced by ordination. Principal Component Analysis (PCA) modules from SPSS (SPSS Inc., Chicago) were employed. PCA axes are normally un-correlated with each other and contain independent information (Digby and Gower, 1981; Gauch, 1982; Gower, 1986). Using the respective loadings, we determined the scores for different Principal Component (PC) axes for each *taluka* as explained by Kent and Coker (1992). These scores were then used for stratification of different homogeneous units using Ward's minimum variance method (Ward, 1963) of hierarchical clustering, which is regarded as an 'optimal' method of similarity analysis. The methodology is schematically represented in Figure 1.

Multiple regression analysis

In order to explore the influence of different ECUs on the agro-ecological characteristics, we performed stepwise multiple regression analysis using both the original and ordinated PC axes

Table 1. Soil parameters, classes and assigned scores

Parameters	Class	Score assigned	Parameters	Class	Score assigned
Calcareousness	Nil	1	Texture	Sandy	1
$CaCO_3$ (%)	<5	3		Loamy	2
	5–15	10		Clayey	3
	>15	15			
Soil erosion	<5	3	pH	6·5–7·5	7
(t ha^{-1} $year^{-1}$)	5–10	7		7·5–8·5	8
	10–20	15		8·5–9·5	9
	20–40	30		>9·5	10
	>40	40			
Depth (cm)	<10	5	Slope (%)	0–1	1
	10–25	12		1–3	2
	25–50	37		3–8	6
	50–75	62		8–15	12
	75–100	87		15–30	22
	>100	100			
Drainage	Extremely poor	1	Salinity	Nil	1
	Very poor	2	(dS m^{-1})	2–4	3
	Poor	3		4–8	6
	Imperfect	4		8–15	12
	Moderately good	5		15–25	20
	Good	6		25–50	37
	Somewhat excessive	7		>50	50
	Excessive	8			

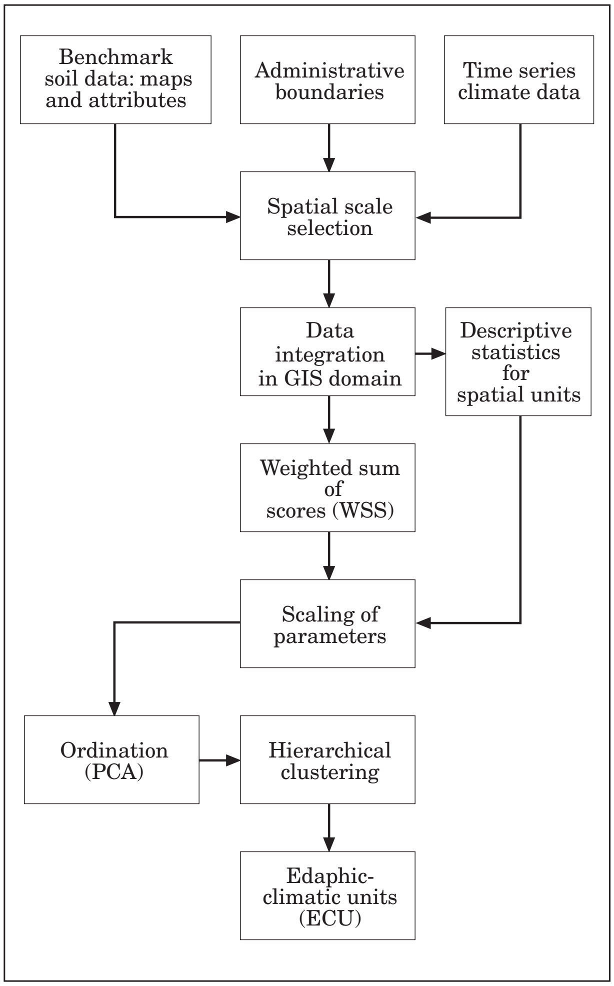

Figure 1. Schematic diagram depicting the methodology used. Homogeneous geographic units of appropriate scale were identified using ordination and hierarchical clustering techniques.

as independent variables. By applying stepwise multiple regression it is possible to identify the most important parameters that influence dependent agro-ecological variables, without *a priori* knowledge about the possible determining parameters. The ordinated variables reduce the correlations between different soil and climatic parameters and, theoretically reduce the redundancy and noise in the data (Gauch, 1982). However, in addition to the usual caveats on the need for caution when interpreting projections, extra attention is needed with correlations (Gower, 1986). We have kept in mind the warning by Gower (1986) that: (a) correlations can be expected to have well-defined meanings only when the data is likely to be approximately multi-normal, (b) this is rarely the case in reality and (c) even exact non-linear relationships may not yield high correlations. Besides, it is sometimes very difficult to interpret the results in the real world due to the abstractness of each PC axis. Therefore, for clarity, we also used the original variables in the regression analysis. Cardillo *et al.* (1999) has illustrated the utility of employing both the original and ordinated variables in similar situations.

Three agro-ecological variables – Total Cultivable Area (TCA, net sown area plus current fallow, which is the agricultural land lying fallow in the current year), area under forest cover (FCA) and Normalized Difference Vegetation Index (NDVI) – were selected as dependent variables. The TCA and FCA data were compiled from land-use statistics. NDVI is a widely used index of vegetation condition derived from satellite based remote sensing data (Lillesand and Kiefer, 1987). We generated the NDVI map of the greenest season for vegetation in a normal rainfall year (i.e. November 1998) using digital data from the Wide Field Sensor (WiFS) of the Indian Remote Sensing satellite, IRS-1C employing Idrisi32 software. We also used the benchmark 1:1 million scale vegetation map prepared in the 1960s by the French Institute of Pondichery (Gaussen *et al.*, 1992) to understand the past distribution pattern of different natural vegetation types in each ECU. Based on proportional area coverage in each ECU, frequencies of different vegetation types were ranked into five categories: dominant, abundant, frequent, occasional and rare.

Results

The PCA showed that out of a possible 11 axes, the first six cumulatively explained about 92% of the variation, indicating the importance of many different edapho-climatic factors. We retained these six components for further analysis. Loadings of different soil and climatic parameters onto six axes are shown in Table 2. It is clear that, except for soil texture and drainage, which are the main contributors to the third PC, all the other parameters had a high degree of influence on the first two. The loadings of the first two axes also indicated strong correlations between mean rainfall and number of rainy days, soil pH and calcareousness, and erosion and slope.

Edaphic-climatic units (ECU)

The clustering procedure produced ten ECUs at the first level of segregation (Figure 2). The number of

Table 2. Loadings of edaphic and climatic parameters onto the first six PCA ordination axes

Parameters	PC1	PC2	PC3	PC4	PC5	PC6
Calcareousness (CLC)	0·77	−0·14	0·18	0·15	0·39	0·08
Erosion (ERS)	−0·75	−0·53	−0·09	−0·24	0·11	−0·07
pH	0·81	−0·13	0·30	0·24	0·16	−0·22
Drainage (DRG)	0·05	−0·26	−0·78	0·45	0·24	−0·08
Salinity (SLN)	0·64	0·26	0·20	−0·51	0·32	−0·23
Depth (DPT)	0·40	0·75	−0·37	0·07	0·07	−0·004
Texture (TXT)	−0·26	−0·09	0·82	0·41	−0·03	0·01
Slope (SLP)	−0·67	−0·55	0·02	−0·08	0·37	−0·09
Mean Annual Rainfall (MAR)	−0·73	0·52	0·12	0·08	0·31	0·19
Mean Annual Rainy Days (RDY)	−0·75	0·57	0·09	0·04	0·23	0·14
Coefficient of Variation in Annual Rainfall (CVR)	0·62	−0·37	−0·03	−0·13	0·08	0·65
Cumulative Variance Explained (%)	39·8	58·5	73·1	80·3	86·2	91·7

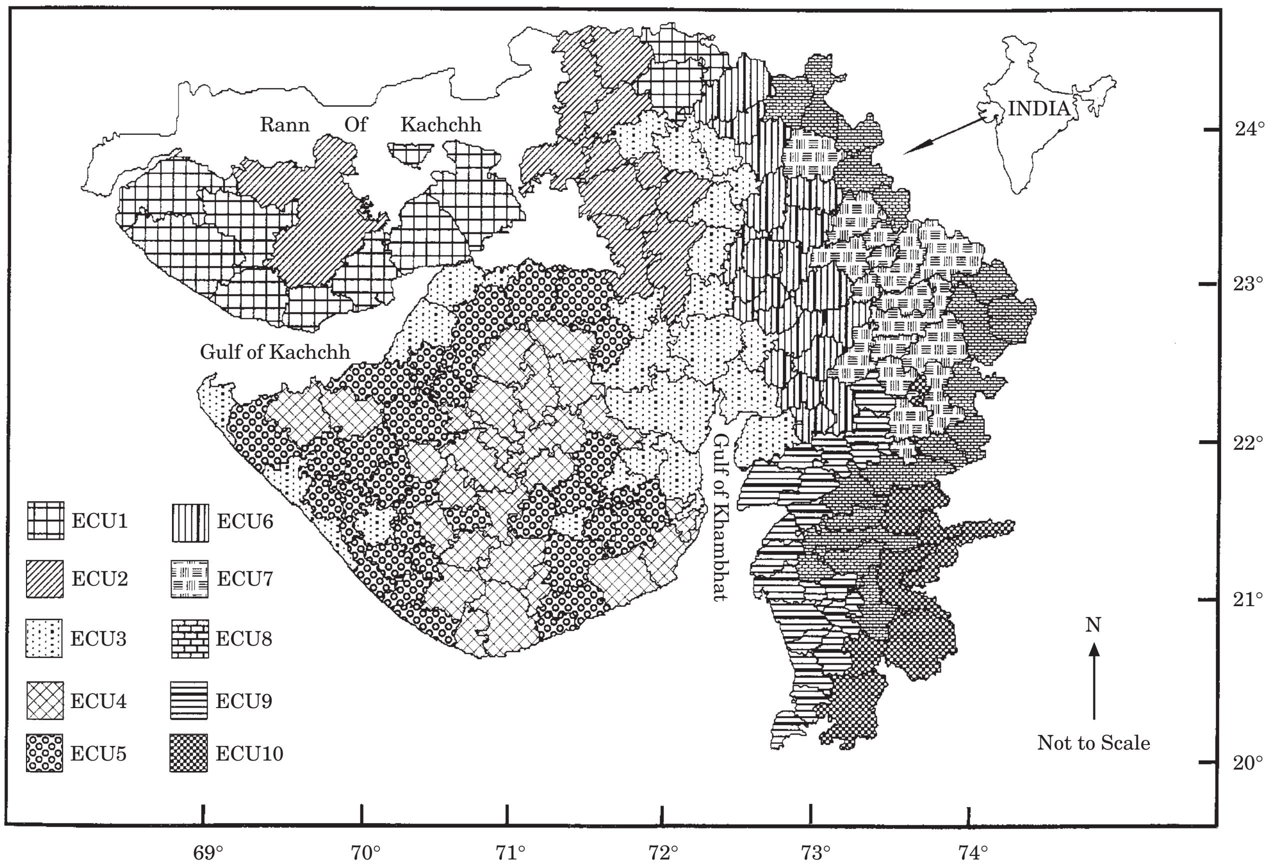

Figure 2. Gujarat State-distribution map of Edaphic Climatic Units. The smallest spatial unit shown is the *Taluka*, an administrative division.

talukas falling into different ECUs varied between 9 and 31, while the total geographical area under each ECU was between 7960 km^2 and 24 684 km^2 (Table 3). Further, in order to characterize and identify more influential parameters in each ECU, we ran PCA separately for each unit. The ordination of different parameters is illustrated using biplots of the first two axes (Figure 3).

Except for soil pH, almost all the other parameters had a significant role in delineating ECU1. While calcareousness, erosion and salinity were the major controlling factors in ECU2, soil texture, slope and MAR were also important. ECU2 consists mostly of flat areas with moderately heavy textured soils. The major part of ECU3 represents areas characterized by a high inherent salinity, which had surfaced in the recent geological past, due to regression of the sea after the Holocene period (Merh, 1992). The PCA revealed that MAR, soil drainage, erosion and salinity were

Table 3. Salient features of different Edaphic-Climatic Units

ECU	Area* (km^2)	% of State	No. of Taluka	MAR (mm±SE)	Mean TCA (%±SE)	Mean FCA (%±SE)	Mean** NDVI
1	17641	9·0	10	393·4±28·3	51·4±6·0	7·64±2·2	0·044
2	16716	8·5	11	482·7±27·8	73·3±5·9	2·48±1·3	0·066
3	22163	11·3	23	590·8±20·2	73·5±2·9	0·70±0·2	0·091
4	23018	11·8	26	619·3±24·4	67·7±2·3	2·96±1·0	0·108
5	24684	12·6	31	638·3±19·2	72·2±1·6	2·21±0·8	0·091
6	14805	7·6	21	794·8±20·6	75·8±2·9	1·25±0·8	0·200
7	9241	4·7	13	912·4±29·1	69·1±2·4	6·82±1·1	0·225
8	15238	7·8	18	951·3±44·4	54·8±3·1	18·71±2·1	0·205
9	10598	5·4	22	1205·8±79·5	77·3±2·0	1·78±0·6	0·203
10	7960	4·1	9	1482·2±125·9	36·1±4·2	38·39±8·8	0·181

*Excluding saline desertic area of Rann of Kachchh.
**SE in NDVI, in most cases, is about 0·01 except in ECU10 where it is 0·05.

more important in defining this unit. While soil erosion, slope, texture and pH were the crucial parameters in defining ECU4, calcareousness and slope were the two major controlling parameters in ECU5. ECU6 forms a transitional zone between the moist and arid regions (Figure 2). MAR, coefficient of variation in annual rainfall (CVR), soil depth, pH and erosion were the major parameters influencing this ECU. Soil pH, soil depth, slope and CVR were the major parameters influencing ECU7. Calcareousness, salinity and slope were the most influential parameters determining ECU8, while soil drainage, salinity and MAR were the most important parameters influencing ECU9. Estuaries of two major river systems – Narmada and Tapi – are present in ECU9, which also experiences a high degree of seawater ingress in the coastal tracts. ECU10 receives the maximum rainfall and is predominantly a hilly tract with relatively steep slopes and faces a high degree of erosion. With the exception of calcareousness, soil depth and CVR, all other parameters played an important role in defining this unit. While most of the ECUs form more or less contiguous clusters, ECUs 3, 5 and 8 consist of well separated clusters bound by edaphic and climatic similarities.

Agro-ecological responses

It is instructive to evaluate the responses from both agricultural and ecological variables within each ECU. To demonstrate this, we selected TCA, FCA and NDVI as dependent variables. We also assessed, in qualitative terms, the distribution of major vegetation types in different ECUs.

Total cultivable area

The TCA represents in a general sense the total land available for cultivation, which is not expected to undergo major year-to-year changes, unless there are sudden technological shifts in agricultural practices. It has reached the upper limits possible under current farming technology. It exhibits considerable spatial variation across ECUs (Table 3) and has the potential of being used as one of the proxy variables for monitoring the degradation of agro-ecosystems. We have, therefore, examined the variation in proportion of TCA in the different ECUs and the degree of dependence this variable has on both the original and ordinated parameter set.

Multiple regression analysis showed that both the parameter sets were good predictors of TCA, although the degree varied in different ECUs (Table 4). Ordinated variables were better predictors in ECU1, ECU9 and ECU10, while in ECU2, ECU3 and ECU4 there was not much difference between the two sets. Original parameters, however, were the better predictors in the remaining ECUs. The PC4 was the most useful predictor for TCA, since it appeared in the regression models of 7 out of 10 ECUs. Interestingly, except in ECU6, the transition zone between humid and arid regions, MAR does not appear in the regression models for TCA. Since TCA is primarily a function of soil conditions, there is no need to read too much into the rather weak relationship with MAR.

Forest cover

Since 1987 the Forest Survey of India has been carried out using satellite-based biennial surveys of FCA. These provide an accurate picture of the

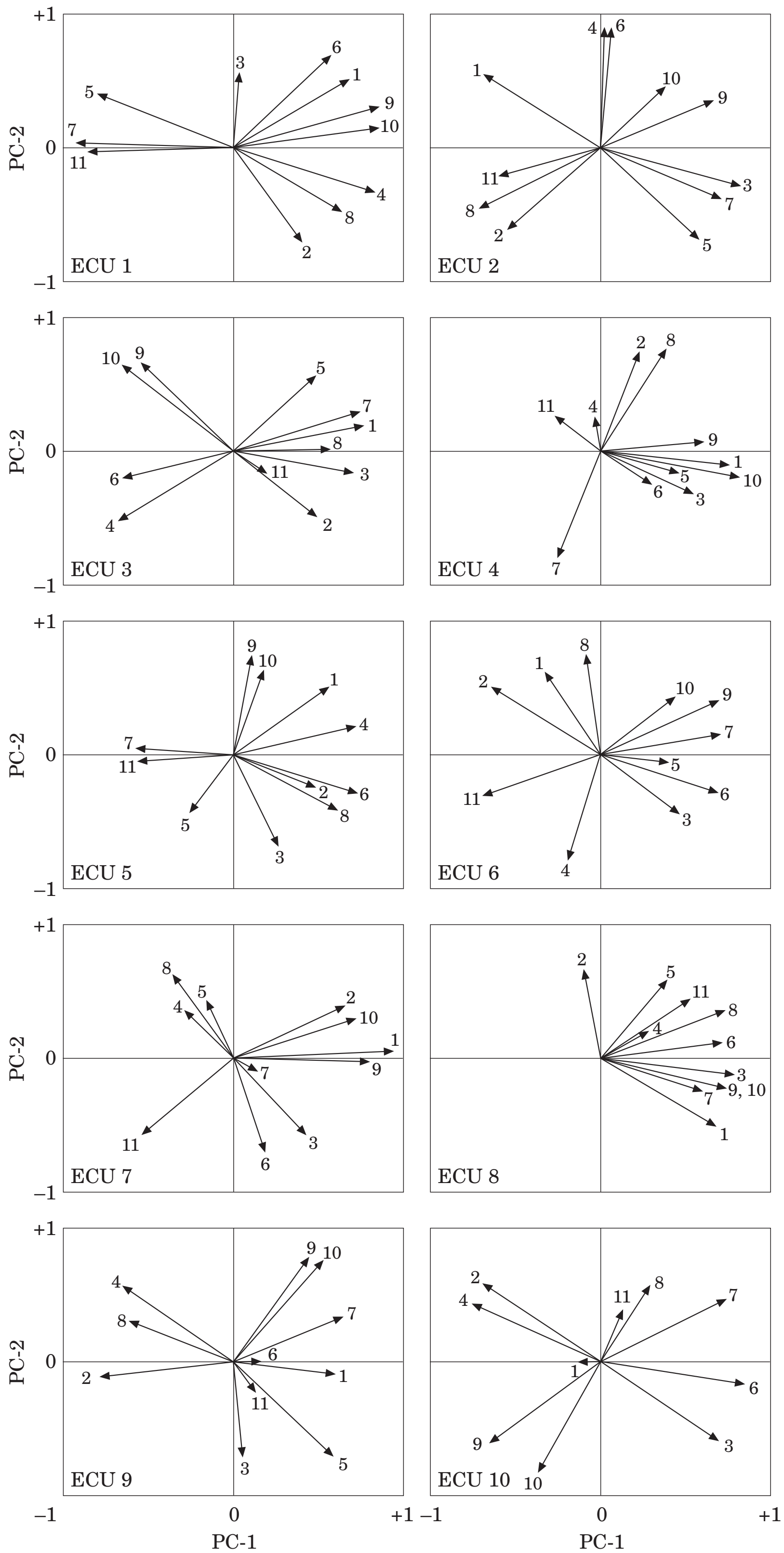

Figure 3. Biplots of first two Principal Component Axes of the edaphic-climatic parameters for ECU1 to ECU10. Key to the variables: 1 – Soil calcareousness, 2 – Soil erosion, 3 – Soil pH, 4 – Soil drainage, 5 – Soil salinity, 6 – Soil depth, 7 – Soil texture, 8 – Surface slope, 9 – Mean annual rainfall, 10 – Avg. annual rainy days, 11 – Coefficient of Variation in annual rainfall.

Table 4. Regression of agro-ecological parameters with original and PCA-ordinated variables in different ECUs

ECU	TCA				FCA				NDVI			
	Original variable		Ordinated variable		Original variable		Ordinated variable		Original variable		Ordinated variable	
	Predictor	R^2	Predictor	R^2	Predictor	R^2	Predictor	R^2	Predictor	R^2	Predictor	R^2
1	TXT (−)	0·86	PC1 (−) PC4 (+)	0·96	CVR (+) pH (+)	0·83	PC1 (+) PC4 (−)	0·86	MAR (+) pH (−)	0·91	PC4 (+)	0·67
2	ERS (−) DPT (+)	0·85	PC2 (+) PC4 (+)	0·84	ERS (+) SLP (−)	0·86	PC4 (−)	0·48	MAR (+)	0·71	PC3 (+) PC4 (+)	0·59
3	DRG (+)	0·79	PC3 (−) PC4 (+) PC5 (+)	0·75	CVR (+) ERS (−) pH (+) DRG (−)	0·72	PC1 (+)	0·51	MAR (+) DRG (+)	0·75	PC3 (−)	0·57
4	RDY (−) pH (+) SLP (−)	0·5	PC1 (+) PC6 (−)	0·49	RDY (+)	0·42	PC5 (+)	0·17	MAR (+)	0·39	PC5 (+)	0·19
5	DRG (+)	0·57	PC3 (−) PC4 (+)	0·39	MAR (+) DRG (−)	0·34	PC1 (−) PC3 (+)	0·29	CVR (−)	0·24	PC6 (−)	0·25
6	MAR (+) RDY (−)	0·34	PC1 (+)	0·04*	DPT (−)	0·68	PC4 (−)	0·32	SLN (−)	0·16*	PC1 (−) PC2 (−) PC4 (+)	0·31*
7	SLP (−)	0·69	PC2 (+)	0·22*	SLP (+)	0·53	PC4 (−)	0·22*	CLC (+)	0·52	PC5 (+)	0·42
8	CLC (−) DPT (+) TXT (+)	0·89	PC4 (+)	0·76	RDY (+) DRG (−)	0·74	PC3 (+) PC4 (−) PC5 (+)	0·82	ERS (+) DRG (−) SLP (+)	0·73	PC4 (−) PC6 (+)	0·98
9	SLN (−)	0·56	PC2 (−) PC4 (+)	0·63	pH (−)	0·14*	PC2 (+)	0·07*	DRG (+)	0·74	PC3 (−) PC4 (+)	0·64
10	CVR (+) SLP (−)	0·64	PC3 (−) PC4 (+) PC5 (−) PC6 (+)	0·95	DRG (−) TXT (−)	0·89	PC5 (+)	0·40	MAR (−) DPT (+) TXT (−)	0·96	PC6 (+)	0·45
Entire state	pH (+) DRG (+) DPT (+) CVR (−) SLN (−) SLP (−) TXT (+)	0·53	PC1 (+) PC2 (+) PC4 (+) PC3 (−) PC5 (−) PC6 (−)	0·55	MAR (+) SLN (+) DRG (−) DPT (−) SLP (+) TXT (−)	0·61	PC1 (+) PC4 (+) PC5 (−) PC6 (−)	0·58	MAR (+) CVR (−) CLC (−) SLN (−) DRG (−) DPT (+) TXT (−)	0·65	PC1 (−) PC2 (+) PC3 (−)	0·57

Sign in parenthesis shows direction of slope. *Indicates $P > 0·05$.

actual area at district level as opposed to the reported area given out by the state revenue department in annual land-use statistics at the *taluka* level. However, there is often a significant mismatch between the actual and reported areas. Among the districts, the discrepancy ranges from a factor of 0·14 to 1·41 of the reported area (Dixit *et al*., 2001). In order to obtain a better estimate of FCA closer to the actual forest cover, we adjusted the *taluka*-level 'reported' area by multiplying it with the respective district-level correction factors. The variation in the FCA for each ECU obtained in this manner is presented in Table 3. Multiple regression analysis revealed that, in general, both the original and ordinated variables were good predictors, except in ECU9 (Table 4). The PCs had higher correlation in ECU8 and original variables in the rest of ECUs, while both types of variables were equally good in ECU1 and ECU5.

NDVI

NDVI reflects the overall picture of vegetation in the area and has strong seasonal dynamics. The use of a single 'greenest' season and normal rainfall year data in this study is merely to explore the variations in such an index across the ECUs. Among the ECUs, lower NDVI values were recorded in the arid ECU1 to ECU3 and higher values in the more humid ECU7 to ECU10 (Table 3). The original variables were, in general, better predictors of NDVI than the PC axes, except in ECU6 and ECU8, where the ordinated variables showed higher correlations. Of the original variables, MAR and soil drainage occurred more frequently (Table 4). Similarly, PC4 appeared in the equations for five ECUs. It may be of interest that MAR was common to the arid and semi-arid regions (ECU1 to ECU4), while slope, erosion, soil depth and texture were more important in the humid regions (ECU8 to ECU10).

Distribution of vegetation types

Since natural vegetation and its distribution pattern is significantly influenced by edaphic and climatic regimes, we examined this aspect vis-à-vis ECUs using the benchmark vegetation map of Gaussen *et al*. (1992), which described vegetation in the 1960s through the notion of 'vegetation series'. The vegetation series denotes the likely 'near climax' or 'plesioclimax' vegetation type that can be supported by the edaphic and climatic conditions of the area. Each series consists of different physiognomic forms that have emerged in response to biotic interferences, resulting in 21 major vegetation types from five vegetation series and 10 physiognomic types.

Vegetation types differed in each ECU (Table 5). While the *Salvadora–Prosopis* series was restricted mainly to ECU1 and ECU2, the *Acacia–Capparis* series and its associated physiognomic forms were widely distributed from ECU2 to ECU7. In these ECUs, the most dominating vegetation types were shrub savannah and scattered shrubs. While restricted distribution of the different physiognomic forms of *Anogeissus–Terminalia* series were recorded only in ECU5, the *Anogeissus–Terminalia–Tectona* series was found widely in ECU7 to ECU10. Moist-deciduous forest of the *Tectona–Terminalia–Adina–Anogeissus* series was, however, reported only from ECU8 and ECU10. The patterns in vegetation distribution further highlight the peculiarities of each ECU.

Discussion

This study demonstrates the effective use of ordination techniques for unbiased identification of homogeneous geographic units, based on edaphic and climatic parameters. These units have several structural and functional advantages over existing zoning schemes. Structurally, more than one ECU is likely to be part of a single spatially contiguous zone, whereas even one ECU may consist of geographically separate clusters bound by edaphic-climatic similarity. The relevance of ECU is clearly related to the scale that one is interested in. At larger scale (regional/country/continental), the zones may suffice, but at intermediate and smaller scales, the ECU seems more appropriate.

Functionally, unlike the existing zoning schemes, these ECUs serve the broader purpose of tracking the changes of any indicator, irrespective of whether they are agricultural, ecological, or environmental. This is possible because they have been developed in an unbiased manner without assigning any *a priori* weights either for or against any development or management goals (e.g. agricultural development or biodiversity conservation). The lack of variation of ECUs over reasonably long time scales also helps highlight environmental changes within each unit in the short to intermediate term, particularly in the context of current attempts at developing environmental indices. The

Table 5. Distribution of major vegetation types in different ECUs

#	Vegetation types*	ECU									
		1	2	3	4	5	6	7	8	9	10
	Mangroves	O		R							
	Salvadora–Prosopis series										
1	Discontinuous thorny thickets	A									
2	Scattered shrubs	D	F								
	Acacia-Capparis series										
3	Discontinuous thorny thickets				O	O	O				
4	Scattered shrubs			O	D	D	D	O			
5	Shrub savanna		A	D	F		F	D			
	Anogeissus–Terminalia series										
6	Savannah woodland					R					
7	Tree savannah					R					
8	Shrub savannah					R					
9	Scattered shrubs					R					
	Anogeissus–Terminalia-Tectona series										
10	Deciduous open forest							F	D		F
11	Savannah woodland				O			O	A		
12	Tree savannah				O				F		O
13	Shrub savannah							O	F	O	
14	Discontinuous thorny thickets								O	F	
15	Scattered shrubs						O	R		A	
	Tectona–Terminalia–Adina–Anogeissus series										
16	Moist deciduous forest								O		D
17	Open forest										F
18	Scattered shrubs									R	
19	Shifting cultivation										A

D=Dominant; A=Abundant; F=Frequent; O=Occasional; R=Rare.
*Based on Gaussen *et al*. (1992).

ECUs can, therefore, be very useful in exercises such as regional environmental assessments for long-term planning or for monitoring desertification at the regional or provincial level.

The relatively static land units show varying degrees of correlation with edapho-climatic parameters. There was greater variation of these in certain ECUs, like 1, 2 and 10. As anticipated from known rainfall patterns, MAR exhibited a clear gradient from ECU1 to ECU10 (Table 3). It is clear that there is a certain degree of variation in the importance of, and correlation between, the parameters within different ECUs, corresponding to the overall environmental heterogeneity among them (Figure 3). The inter dependence of a few agro-ecological variables with the original and ordinated edaphic-climatic variables illustrates the distinctive features of each ECU. The response of these variables, chosen on the basis of varying levels of dynamism (slow to seasonal fluctuations), shows that independent of the degree of dynamism, their independence is different in each ECU. Such behaviour provides a reality check on any agro-ecological and bio-climatic differentiation.

PCA has therefore provided a sound approach for objectively clustering similar units. Ordination has had a less powerful effect as a predictor for agro-ecological responses, as it is both enhanced and diminished by ordination of edaphic and climatic parameters. The increased or decreased predictive power of these parameters is also reported in other studies (Danell *et al*., 1996, Macdonald *et al*., 1996, Cardillo *et al*., 1999). Clearly, the advantage of ordinated variables in predicting and accounting for the influence of correlated parameters needs to be established separately for each agro-ecological variable and ECU.

In conclusion, the approach was found to be provide valuable insights into a region like Gujarat where the gradients in climate regime are gradual, while edaphic conditions are marked by abrupt discontinuities and large heterogeneities. Soil parameters were found to be highly relevant, first in delineation of separate regions and later in terms of the dependence of agro-ecological variables on them. The approach may be significantly improved if better edaphic and climatic data (in terms of parameters and spatial resolution) become available.

Acknowledgements

This work was carried out partially under the studies commissioned by the Gujarat Ecology Commission during the preparation of the State Environmental Action Program for Gujarat (World Bank aided India: Environmental Capacity Building – Technical Assistance Project). We wish to thank Prof. Kirit Parikh for useful discussions and C.S.Silori for sharing his insights on land-use data. We also wish to thank two anonymous referees for their valuable comments on the original manuscript.

References

Anon (1987). *Planning atlas of Gujarat. Resource Profile*. Directorate of Economics and Statistics. Government of Gujarat, Gandhinagar.

ARPU (1998). *Agro-climatic Regional Planning: recent developments*. ARPU working paper no. 10. ARPU, Planning Commission, Ahmedabad.

Brisse, H. and Grandjouan, G. (1978). *Elements de phytoclimatologie numerique. Principles de la classification climatique des plantes*. Actes 3ème Colloque de l'Association Informatique et Biosphere, Paris, pp. 55–184.

Bunce, R. G. H., Morrell, S. K. and Stel, H. E. (1975). The application of multivariate analysis to regional survey. *Journal of Environmental Management* **3**, 151–165.

Cardillo, M., Macdonald, D. W. and Rushton, S. P. (1999). Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. *Landscape Ecology* **14**, 423–435.

Carter, D. B. (1954). *Climates of Africa and India according to Thornthwaite's (1948) classification*. Johns Hopkins University Publication in Climatology.

Danell, K., Lundberg, P. and Niemela, P. (1996). Species richness in mammalian herbivores: patterns in the boreal zone. *Ecography* **19**, 404–409.

Digby, P. G. N. and Gower, J. C. (1981). Ordination between- and within- groups applied to soil classification. In *Down to Earth Statistics: Solutions looking for geological problems* (D. F. Merriam, ed.), pp. 63–75. Syracuse University Geological Contribution.

Dixit, A. M., Geevan, C. P. and Silori, C. S. (2001). Status of natural terrestrial vegetation in Gujarat – a reassessment. *Indian Forester* **127**(5), 533–546.

FAO (1983). *Guidelines: Land Evaluation for Rainfed Agriculture*. Soils Bull. 52, Food and Agriculture Organisation (FAO), Rome, p. 237.

Gauch, H. G. (1982). *Multivariate analysis in community ecology*. Cambridge: Cambridge University Press.

Gaussen, H., Legris, P., Blasco, F., Meher-Homji, V. M. and Troy, J. P. (1992). *International Map of the Vegetation and of Environmental Conditions – Notes on the sheet Kathiawar*. Institut Francais de Pondichery.

Gower, J. C. (1986). Introduction to ordination techniques. In *Developments in Numerical Ecology* (P. Legendre and L. Legendre, eds), pp. 3–64. Berlin: Springer-Verlag.

Holdrige, L. R. (1947). Determination of world plant formation from simple climatic data. *Science* **105**, 367–368.

Kent, M. and Coker, P. (1992). *Vegetation Description and Analysis: A Practical Approach*. London: Belhaven Press.

Lillesand, T. M. and Kiefer, R. W. (1987). *Remote Sensing and Image Interpretation*. New York: Wiley.

Macdonald, D. W., Michelmore, F. and Bacon, P. J. (1996). Predicting badger sett numbers: evaluating methods in East Sussex. *Journal of Biogeography* **23**, 649–655.

Merh, S. S. (1992). Quaternary sea level changes along Indian coast. *Proceedings of Indian National Science Academy*, **58A**, 461–472.

Murthy, R. S. and Pandey, S. (1978). Delineation of Agro-Ecological Regions of India. 11th Congress of ISSS, Edmonton, Canada, June 19–27.

Nagendra, H. and Gadgil, M. (1998). Linking regional and landscape scales for assessing biodiversity: a case study from western ghats. *Current Science* **75**, 264–271.

NBSS & LUP (1994). *Soils of Gujarat for Optimising Land Use*. National Bureau of Soil Survey & Land Use Planning, Nagpur.

Retuerto, R. and Carballeira, A. (1991). Defining phytoclimatic units in Galicia, Spain by means of multivariate methods. *Journal of Vegetation Science* **2**, 699–710.

Rodgers, W. A. and Panwar, H. S. (1988). *Planning a wildlife protected area network in India*. Dehra Dun: Wildlife Institute of India.

Rubin, F. (1998). The physical environment and major plant communities of the Tankwa-Karoo National Park. *Koedoe* **41**, 61–94.

Schultz, J. (1995). *The Ecozones of the World. The Ecological Divisions of the Geosphere*. New York: Springer-Verlag.

Sehgal, J., Mandal, D. K., Mandal, C. and Vadivelu, S. (1992). *Agro-Ecological Regions of India*. Technical Bulletin No. 24 (Second edition). National Bureau of Soil Survey & Land Use Planning, Nagpur.

Subramaniam, A. R. (1983). Agro-ecological zones of India. *Archives of Meteorology, Geophysics and Bioclimates Series* **B32**, 329–333.

Ward, J. H. (1963). Hierarchical grouping to optimise an objective function. *American Statistical Association Journal* **58**, 236–244.