Animal Conservation (1998) **1**, 253–262 Printed in the United Kingdom

Unsupervised recognition of individual tigers and snow leopards from their footprints

P. Riordan

Behavioural and Environmental Biology, Department of Biological Sciences, Manchester Metropolitan University, Manchester M1 5GD, UK

(Received 2 January 1998; accepted 27 May 1998)

Abstract

This study presents the testing of two unsupervised classification methods for their ability to accurately identify unknown individual tigers, *Panthera tigris*, and snow leopards, *Panthera uncia*, from their footprints. A neural-network based method, the Kohonen self-organizing map (SOM), and a Bayesian method, AutoClass, were assessed using hind footprints taken from captive animals under standardized conditions. AutoClass successfully discriminated individuals of both species from their footprints. Classification accuracy was greatest for tigers, with more misclassification of individuals occurring for snow leopards. Examination of variable influence on class formations failed to identify consistently influential measurements for either species. The self-organizing map did not provide accurate classification of individuals for either species. Results were not substantially improved by altering map dimensions nor by using principal components derived from the original data. The interpretation of resulting classifications and the importance of using such techniques in the study of wild animal populations are discussed. The need for further testing in the field is highlighted.

INTRODUCTION

Many ecological field studies have used sign survey techniques for gaining information about animal populations. Such methods utilize track and dropping counts as indices of animal activity (e.g. Van Dyke, Brocke & Shaw, 1986; Smallwood & Fitzhugh, 1995; Zielinski & Stauffer, 1996), habitat use (Petrak, 1990; Putman, 1990) and population density (Koster & Hart, 1988; Smallwood & Fitzhugh, 1993; Komers & Brotherton, 1997). These methods have advantages in that they are relatively inexpensive, logistically straightforward and they do not require direct contact with the animal or animals in question (Putman, 1984; Clevenger, 1993). When compared with more direct methods, such as mark–recapture or radio-tracking, the data gained from sign surveys may not be as rigorous (Servin, Rau & Delibes, 1987).

The ability to identify which individuals are responsible for sets of footprints would allow indirect tracking methods to be used for gaining in-depth ecological information, such as home-range estimates. Animals with conspicuous footprint characteristics, such as missing digits or obvious scars, may be discriminated by eye, but such animals may be rare in the population of interest. Smallwood & Fitzhugh (1993) showed that for mountain lion the tracks of known individuals could be distinguished using discriminant function analysis (DFA). One problem with this is that in most field studies there is no prior information about the animals responsible for sets of tracks. In such circumstances supervised classification methods such as DFA or non-hierarchical cluster analysis are of little use, since class assignments (i.e. individual identities) are required before analysis. Many surveys of free-ranging tigers have attempted to manually allocate footprints to individuals, by comparing and sorting tracings taken in the field (Panwar, 1979). Such methods have received criticism, because of their subjectivity (Karanth, 1987). The requirement has thus been highlighted that, for tigers, a reliable systematic means of identifying individuals from their footprints should be developed (Nowell & Jackson, 1996). Such a method could also be applied to the study of a range of other animal species.

Unsupervised classification methods automatically extract inherent clusters within datasets, without prior labelling of cases within the data. This gives a potential means by which footprints from unknown animals

All correspondence to: P. Riordan. Tel: +44 (0)161 247 6325; E-mail: P.Riordan@mmu.ac.uk.

could be grouped into individuals, without subjective bias. To assess unsupervised classification as a potential tool in animal ecology, two methods were tested using footprints from captive tigers and snow leopards. These were the Kohonen self-organizing map (SOM): a neural-network based system developed at Helsinki University of Technology (Kohonen, Kangas & Laaksonen, 1992); and AutoClass: a Bayesian method developed at NASA (Stutz & Cheeseman, 1994). Software for both methods are freely available for research purposes over the internet via anonymous ftp (SOM: *ftp://cochlea.hut.fi/pub/som_pak* [130.233.168.48]; AutoClass: *ftp://csr.uta.edu/pub/autoclass*).

The Kohonen self-organizing map is a neural-network based clustering method, involving the projection of a multivariate input space onto a two-dimensional array (Kohonen *et al*., 1992). Each instance within the multivariate input space is mapped to all nodes on the two-dimensional array. The array is made to resonate in accordance with the input data, resulting in one or more regions of excitation, corresponding to clusters within the input space. More detailed technical accounts of this method are given by Kohonen (1989) and Kohonen *et al*. (1992). SOMs have been successfully applied to a number of classification problems in astronomy (Murtagh & Hernandez-Pajares, 1995), biochemistry (Reibnegger, Weiss & Wachter, 1993) and spatial patterning (Varfis & Versino, 1992).

AutoClass is an unsupervised classification system based upon Bayesian theory. Bayes' rule is a mathematical method of describing probabilities given prior expectations and defines how probabilities alter with evidence. AutoClass searches the input data space repeatedly, automatically selecting the most probable classifications. The best classifications optimally trade-off predictive accuracy against class complexity, avoiding over-fitting the data (Cheeseman & Stutz, 1995). Cases are allocated to classes probabilistically, rather than employing cut-off criteria. Class membership probabilities sum to 1, resulting in a 'fuzzy' classification. Input data are expressed as statistical models, which may be either independent or allow for covariance between variables. A full account of Bayesian theory and it's application in AutoClass is given by Hanson, Stutz & Cheeseman (1991). AutoClass has been successfully demonstrated as a classification tool in a number of fields including astronomy, biochemistry and remote sensing (Cheeseman & Stutz, 1995).

METHODS

The two classification methods were tested using 10 tigers (*Panthera tigris*) and 6 snow leopards (*Panthera uncia*), all held at Port Lympne Zoo. Of the tigers, 7 were Bengal tigers (*P. t. tigris*) and 3 were Siberian tigers (*P. t. altaica*) (Table 1).

Footprints were taken from a consistent substrate of fine builder's sand. The sand was placed on the floor of the enclosures and spread to an approximately even depth of 2 cm. The study animals moved over the sand patches one at a time enabling each track set to be assigned to an individual. Footprints were collected from sets only if the animals were travelling at a normal walking or trotting pace. Only footprints made by the hind feet were collected. This was because, during normal walking or trotting, the hind footprints often register over those of the fore feet, resulting in fewer intact fore footprints.

Table 1. Footprint collections made at Port Lympne Zoo from tigers and snow leopards using acetate and photographic methods

Name	Sex	Age	Left hind foot		Right hind foot	
		(years)	Acetate	Photo	Acetate	Photo
Tigers:						
Chambal	M	11	10	10	10	11
Harami	M	8	10		10	
Nari	F	17	10	10	10	10
Pindi	F	9	10		11	
Spotty	M	4	10		11	
Stripey	M	4	10		11	
Thana	F	9	10		11	
Vosthok[a]	M	11	10	10	10	
Zeyan[a]	F	11	10	10	10	10
Zeyna[a]	F	11	10	10	10	10
Snow Leopards:						
Alf	M	6	10		11	
Aria	F	6	10		10	
Atheni	F	6	10		10	
Atheni-Moon	F	3	10		10	
Messalina	F	16	10		10	
Sitar	F	14	10		10	

[a]Subspecies *P. t. altaica*.

Two methods of footprint acquisition were used: acetate tracing (Panwar, 1979) and photographs. With the acetate method, acetate sheets were placed onto a picture frame, which could be positioned over the prints without disturbing them. The outline of the print was traced onto the acetate using an indelible pen. The method for doing this was to kneel over the print and look directly down onto it at all times whilst tracing. Photographs were taken using a standard 35 mm camera, with a 50 mm lens, pointing directly down onto the print. A flash gun was positioned at a low angle to the prints, to give depth to the photographs. Scale rules were placed next to the prints for calibration. A minimum of 10 footprints were acquired from each hind foot for both methods. All animals were sampled using acetate tracing and a subset of tigers were photographed for comparison and as a test of the consistency of the acetate method.

A total of 27 morphometric measurements were taken from each print: 18 linear and 9 angular measures (Fig. 1). Measurements were taken from tracings of acetates and photographs, which allowed annotations and reference lines to be added without defacing the original samples. The units of linear measurement were millimetres, with a precision of 0.5 mm. Angular variables were measured in angular degrees, with 0.5° (i.e. 30′) level of precision.

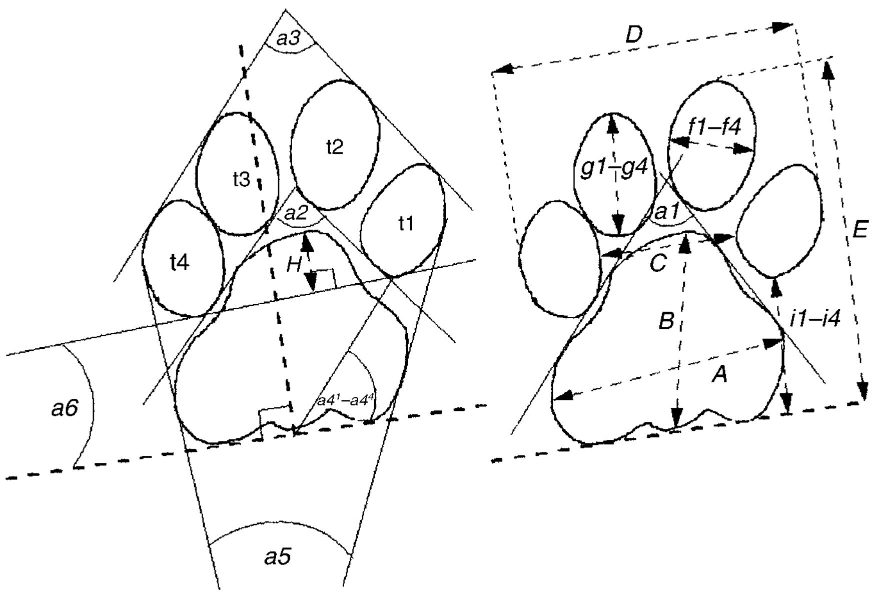

Fig. 1. Morphometric measurements taken from footprints: *A*, main pad width; *B*, main pad length; *C*, inner toe distance; *D*, total width; *E*, total length; *f1* to *f4*, toe widths; *g1* to *g4*, toe lengths; *H* main pad top to toe baseline; *i1* to *i4*, distance from toes to baseline; *a1*, main pad top angle; *a2*, inner toe spread angle; *a3*, outer toe angle; $a4^1$ to $a4^4$, right-hand angle from baseline to each toe; *a5*, outer toes to main pad angle; *a6*, baseline to toe baseline angle. t1 to t4, toe 1 to toe 4.

Analysis

Data for each hind foot were analysed separately. It was assumed that footprints taken from track sets in the field could be accurately identified as being made by either left or right hind feet. It was also assumed that an experienced tracker or field biologist could differentiate between footprints made by fore and hind feet.

Unsurprisingly, the amount of covariance between linear measures was relatively high, since all variables were presumably related to overall footprint size. For tiger left hind feet (LHF) the average Pearson correlation coefficient (r) between linear measures was 0.615 (n = 150), where n is the number of records. Angular variables were also correlated with one another (tiger LHF: average r = 0.229; n = 150), though the covariance between linear and angular variables was relatively low (tiger LHF: average r = 0.021; n = 150). This pattern of variable correlations was also shown by tiger and snow leopard RHF footprints. Neural networks, such as SOM, are reportedly relatively insensitive to covariance between input variables (Ripley, 1993). However, sensitivity to covariance is likely to be dependent on the nature of the data and, in some cases, classification efficiency may be improved using a principal components analysis (PCA) prior to entry into the SOM (Murtagh & Hernandez-Pajares, 1995). With AutoClass, linear and angular variables were entered within separate multinormal ‘multi_normal_cn’ models, which were reportedly able to allow for these covariances (Stutz & Cheeseman, 1994).

To assess any effects of covariance on the two classifiers, data dimensionality was reduced into independent variables using PCA. Ten principal components were generated from the 27 original variables in each case. These derived variables consistently accounted for at least 95% of the variability of the original datasets. Entry of the principal components into the SOM required no modification from that required for the original data. With AutoClass, principal components were entered as a ‘single_normal_cn’ model, which assumes independence between variables.

The probable overall size difference between Siberian and Bengal tiger footprints was not considered problematic for the analysis. However, in addition to analysing all tigers together, Bengal tiger footprints were analysed as a separate subset. Data from both footprint acquisition methods, acetates and photographs, were initially pooled prior to entry into both classifiers. Cases were labelled, and thus the relative class allocations of footprints derived from each method could be observed within the classifications.

For the SOM, maps of dimensions 3×3, 3×5, 10×10, 10×15, 20×20, 20×25 and 50×50 were used to assess the influence of map size on classification accuracy. Rectangular, rather than square maps, may be more efficient for classification (Kohonen *et al.*, 1992), thus both types were assessed.

RESULTS

Self-organizing map tested using tiger footprints

Self-organizing maps (SOMs) derived from all tiger footprints combined failed to sufficiently cluster individuals into distinct groups. The resultant 50×50 map for left hind footprints is shown in Fig. 2. Increasing the dimensions of the map appeared to allow more cluster separation, however there was no apparent difference between rectangular and square maps. The prints from

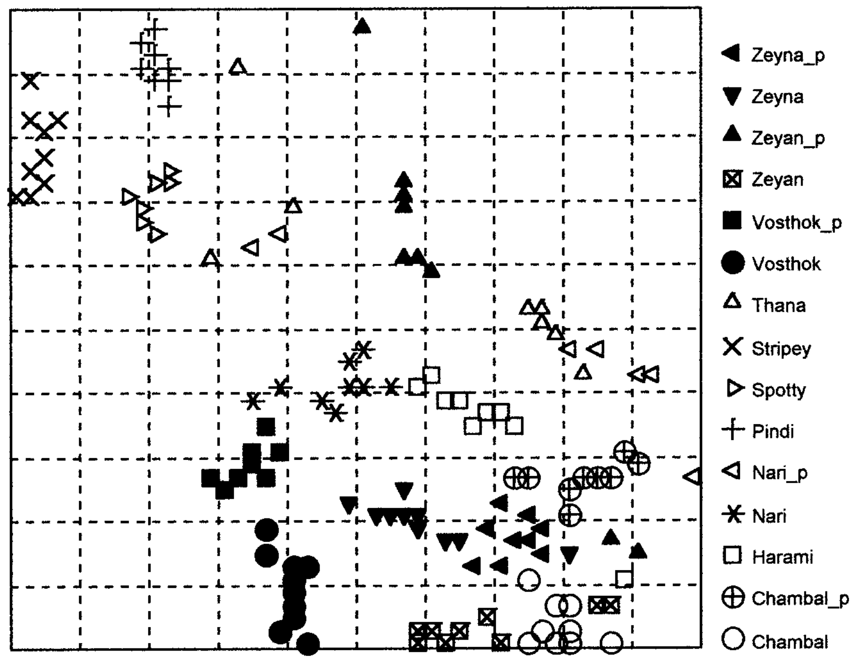

Fig. 2. Self-organizing map (50×50) for left hind footprints of all tigers combined. The '_p' suffix on tiger identifiers indicates measurements were derived from photographs of the footprints

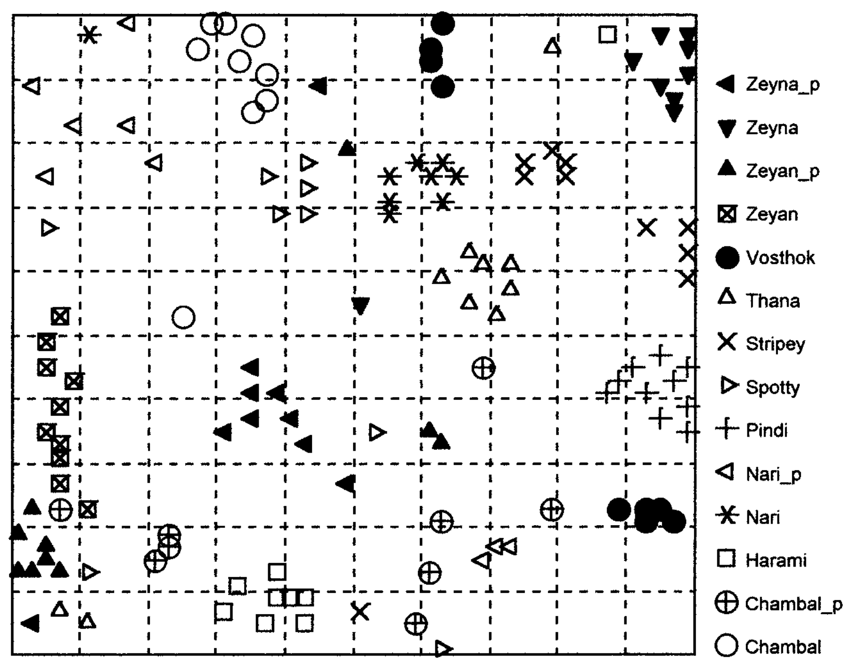

Fig. 3. Self-organizing map (50×50) for all tigers, left hand footprints, constructed from 10 principal components, accounting for 96.4% of the original variability. The '_p' suffix on tiger identifers indicates measurements were derived from photographs of the footprints

individual animals were mapped within closer proximity to one another, though the overall separation from other individuals was poor.

Measurements of footprints made from acetates and photographs did not occur within close proximity for all individuals. Footprints from both methods for Vosthok, Zeyna, Nari and Chambal were distributed in relatively close proximity, though not sufficiently to be able to consider the two methods as equivalent. Similarly unsuccessful results were also obtained for SOMs derived from the right hind footprints of all tigers combined.

The separation of individuals was slightly worse when Siberian tigers were excluded from the analyses of both feet and no improvement in classification was gained by

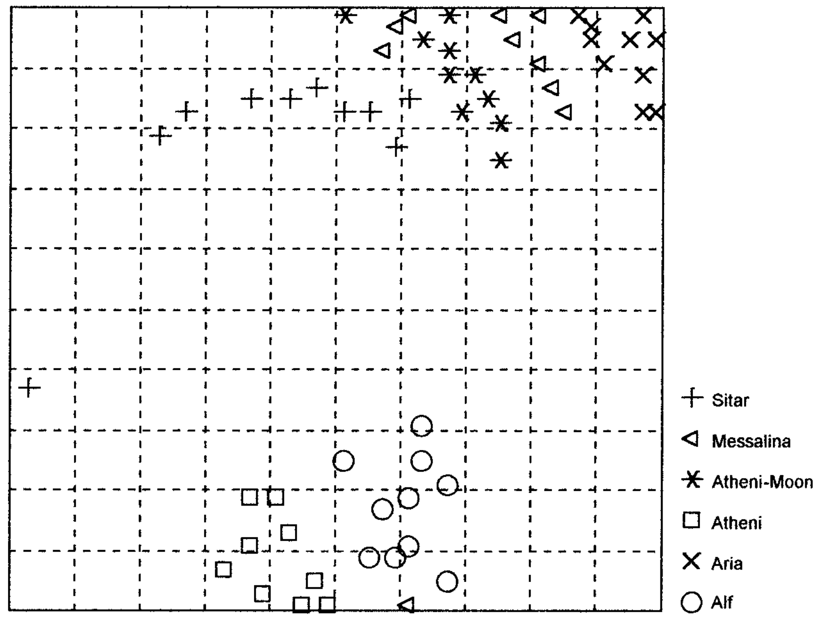

Fig. 4. Self-organizing map (50×50) derived from snow leopard left hind footprints.

using variables derived solely from acetates. The use of principal components slightly improved individual separation for both left (Fig. 3) and right footprints. Ten principal components were used, accounting for 96.4% of the variability in the original data.

The increase in separation of individuals shown in Fig. 3 was offset by an increase in overall scatter within the map. Individuals such as Chambal, Pindi, Zeyna and Harami show close aggregation in their point distributions. However, there are many outlying points which would confuse the classification were this to come from field data of unknown individual composition.

Self-organizing map tested using snow leopard footprints

Results from snow leopard footprints were similar to those from tiger prints. No discrimination of individual snow leopards could be gained from SOMs. Larger maps produced greater separation between individuals, but there was no apparent benefit from using rectangular, rather than square, maps. As with tigers, cases for individuals occurred within relatively close proximity, but true separation from other individual groupings could not be attained (Fig. 4).

Table 2. AutoClass classification of left footprints for all tigers combined

	Class															
Animal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
Chambal	10															10
Chambal_p		10														10
Harami			10													10
Nari				10												10
Nari_p					10											10
Pindi						10										10
Spotty							10									10
Stripey								10								10
Thana									10							10
Vosthok										10						10
Vosthok_p											10					10
Zeyan												9		1		10
Zeyan_p													10			10
Zeyna														10		10
Zeyna_p															10	10
Total	10	10	10	10	10	10	10	10	10	10	10	9	10	11	10	150

Class assignments are shown with the numbers of cases occurring from each individual. The '_p' suffix on tiger identifiers indicates measurements were derived from photographs of the footprints.

Table 3. AutoClass classification of right footprints for all tigers combined

Animal	Class 0	1	2	3	4	5	6	7	8	9	10	11	12	13	Total
Chambal	10														10
Chambal_p		11													11
Harami			10												10
Nari	1			9											10
Nari_p					10										10
Pindi						11									11
Spotty						1	10								11
Stripey								11							11
Thana									11						11
Vosthok										10					10
Zeyan											10				10
Zeyan_p												10			10
Zeyna													10		10
Zeyna_p														10	10
Total:	11	11	10	9	10	12	10	11	11	10	10	10	10	10	145

Class assignments are shown with the numbers of cases occurring from each individual.

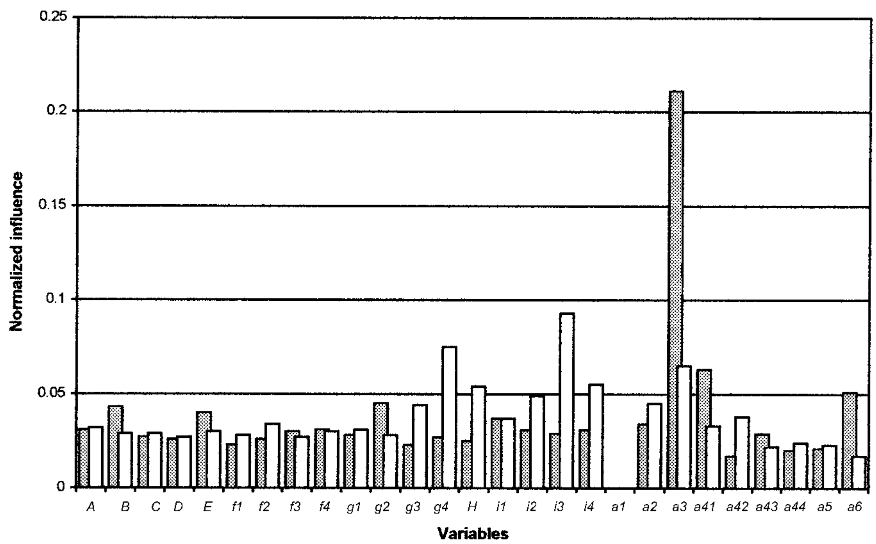

Fig. 5. Normalized influences of each variable within the AutoClass classifications of left (▩) and right (□) hind footprints of all tigers combined.

With snow leopards, using principal components added nothing to the power of the classification. As with tigers, PCA appeared to increase noise within the map, but with snow leopards there was no additional individual separation. The effect was to further confuse any potential separations.

AutoClass tested using tiger footprints

AutoClass successfully discriminated tigers into individual classes using left (Table 2) and right (Table 3) footprints.

The classification using left hind footprints misclassified only one case of Zeyan into the class predominated by Zeyna. All cases had probabilities of class membership exceeding 0.95.

The misclassification rate for the right hind footprints was slightly higher than that for the left foot. Two cases were misrepresented: one case from Spotty was misclassified into the class of Pindi and one case from Nari was misclassified as Chambal. Probabilities of class membership were consistently high, with $P > 0.94$. For those individuals whose footprints were represented by both acetates and photographs, data from each method were classified as separate class instances.

Variable *a3*, the outer toe angle, appears to be most influential in the classification of left hind footprints (Fig. 5), with no other variables contributing appreciably. Closer examination of the outputs from AutoClass reveal that *a3* was only influential in determining classes 11 and 12 (Zeyan and Zeyan_p). No isolated variable or subset of variables was consistently influential in the overall classification. In general, angular measures were the most important for the classification.

In the classification of right hind footprints, variables *i3* (toe3 to baseline) and *g4* (toe4 height) are most influential (Fig. 5). As with classification of left prints, vari-

Table 4. AutoClass classification of left footprints for snow leopards

	Class						
Animal	1	2	3	4	5	6	Total
Alf	10						10
Aria		10					10
Atheni			10				10
Atheni-Moon				10			10
Messalina					10		10
Sitar	1					9	10
Total	11	10	10	10	10	9	60

Table 5. AutoClass classification of right footprints for snow leopards

	Class					
Animal	1	2	3	4	5	Total
Alf	11					11
Aria		10				10
Atheni			10			10
Atheni-Moon			3	7		10
Messalina			2	8		10
Sitar					10	10
Total	11	10	15	15	10	61

able influences were not consistent for all classes. Classes 3 (Nari) and 6 (Spotty) were strongly influenced by *i3* and only class 9 (Vosthok) was influenced by *g4*. Classes 3 (Nari) and 10 (Zeyan) were also significantly influenced by *i4* (toe4 to baseline) although the overall influence of *a3* across the entire classification appears greater (Fig. 5). Overall, the linear measures relating to toe distance from the main pad (*i1*–*i4*) were most significant for the classification.

The inclusion of the Siberian tigers in the classification did not have a detrimental effect. Omitting them from the classification still resulted in accurate discrimination of the remaining Bengal tigers.

Using 10 principal components within AutoClass, as with the SOM method, appeared to confuse the classifications of both LHF and RHF. Misclassification between individuals prevented accurate individual identification.

AutoClass tested using snow leopard footprints

As with tigers, AutoClass accurately discriminated individual snow leopards from their left hind footprints (Table 4). All but one instance of Sitar were clustered into discrete classes. Probabilities of case inclusion into each class were consistently greater than 0.97, with the exception of Sitar's occurrence within class 1, where $P(c1) = 0.57$ and probability of membership to class 6, $P(c6) = 0.42$.

AutoClass classification of snow leopard right hind footprints was less powerful than that for the left foot (Table 5). The major class assignments of Atheni-Moon and Messalina co-occurred in class 4 with instances of both animals conflicting with Atheni in class 3. All probabilities of class membership exceeded 0.95.

For the classification of the left foot, variables *a6* (baseline to toe baseline angle) and *a5* (outer toes to main pad angle) had greatest overall importance (Fig. 6). Variable *a6* was particularly important in defining class 2 (Aria) and *a5* was influential for class 3 (Atheni) and class 2. Angular variables and the linear *ix* (distance from toes to baseline) are most important for the overall classification of individuals from the left footprints.

In classifying individuals based on right footprints, variables *a3* (outer toe angle), *H* (main pad apex to toe baseline) and *a2* (inner toe spread angle) were most influential. Variable *a3* was important in defining class 5 (Sitar) and

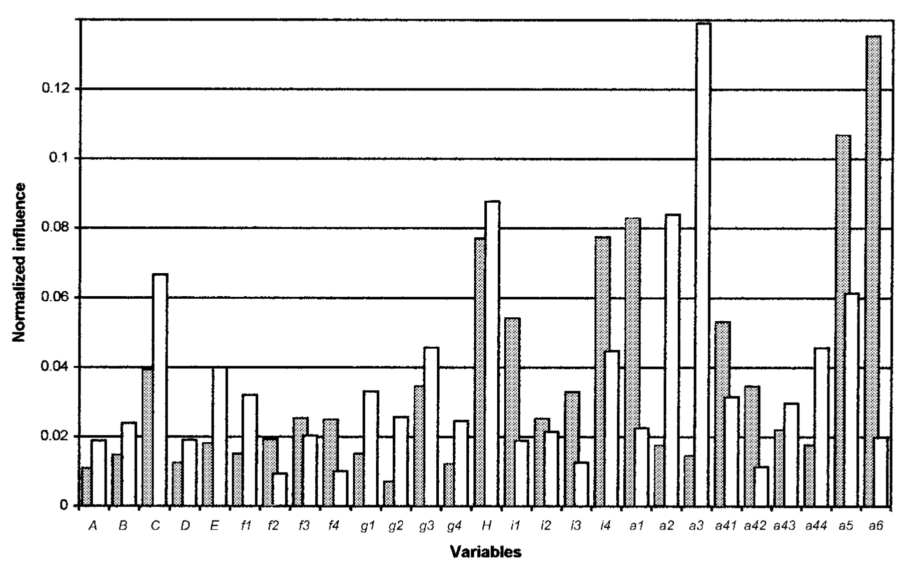

Fig. 6. Normalized influences of each variable within the AutoClass classifications of left (▩) and right (□) hind footprints of snow leopards.

variable *H* was dominant in class 2 (Aria). Variable *a2* was influential over all classes, except for class 2.

As with tigers, entering principal components into AutoClass as separate, independent models created a number of misclassifications. This prevented accurate classification of individuals from both LHF and RHF.

DISCUSSION

In discussing the results of the two classification methods tested, the question must be asked as to whether they would discriminate individuals from footprints taken in the field, with no prior knowledge of the animals concerned. Given the results of this test, the Kohonen self-organizing map would not. AutoClass, however, may be able to discriminate wild animals in a field study.

The reasons for the disparity in performance between these two classification methods, given the same data, presumably lies in their different algorithmic approaches to unsupervised learning. Both systems have been successfully demonstrated in a number of applications and so one must suppose that the nature of the data presented here was not suited to SOM, whereas AutoClass was able to perform well. This may have been in part due to covariance, since PCA gave slight improvement to the SOM results, though not sufficient to allow clear individual discrimination. Within the data set, there is a high degree of heteroscedasticity between individuals for each variable. For example, for acetate tracings of tiger left footprints, values of the coefficient of variation (V), as calculated by $V = s \times 100/\bar{x}$ (Zar, 1984), ranged from 0.21 ($V_{\mathrm{Stripey},\, a4^3}$) to 64.48 ($V_{\mathrm{Harami},\, a6}$). Not surprisingly, V was greatest for the smaller angles, *a5* and *a6*, where relative precision was lowest. This pattern was consistent for data from tiger RHF as well as for snow leopards. Variability was also not consistent for each individual, with some animals showing marked variable dispersion. The combination of covariance and high variability within individual footprint data produced relatively noisy data.

The SOM, as a neural-based classifier, attempts a non-linear projection of the input data probability density functions. Neural networks seek a global approximation of such functions, i.e. a generalization relating all footprints to the individual animals responsible. As data variance increases, function approximation tends toward more local solutions, i.e. overfitting (Geman, Bienenstock & Doursat, 1992; Lawrence, Tsoi & Back, 1996). The degree of noise within the data may inhibit the SOM from finding global function generality, thus preventing the formation of an accurate classification.

AutoClass describes the functional form of the probability functions as class models, incorporating a number of input variables. Models in this case were defined according to covariance, with correlated variables occurring within the same model. Classes are then formed from the particular set of parameter values and their associated model, resulting in a probabilistic classification (Stutz & Cheeseman, 1994). Data noise is thus distributed among the class models, diluting its effects and thus allowing a workable classification. This is demonstrated by the failure of AutoClass to construct an accurate classification from derived principal components. The removal of covariances required separate class models for each principal component with no dilution of the effects of data variability. This may give rise to local minima in the probability density functions, in a similar fashion to those experienced by neural networks.

Given that this study was performed by a single researcher, under standardized conditions, data gained from footprints taken as part of a field study may show marked differences to those presented here. Differences in substrate, slope and animal speeds may all have influence on intra-individual data variability to some degree, potentially influencing classification accuracy. Where acetate tracings are taken, inconsistencies between field-workers may arise where more than one person is collecting data. All of these factors need addressing prior to using this method in a field situation.

To some extent, these issues can be controlled for. Animal speeds can be approximated from the relative positions of footprints in a set of tracks (Liebenberg, 1990). Analyses can thus be performed using footprints made by animals travelling at the same speed, although the available dataset may be reduced. The effects of substrate and slope differences would require more investigation. Smallwood & Fitzhugh (1993) showed that, for mountain lion, slope was significant in affecting toe spread and main pad width, whilst having the smallest effect on total print length and main pad width. AutoClass was little influenced by either main pad width (variable *A*) or total print length (*E*) and so slope effects may be detrimental to individual classification. Field study sites possessing differing substrates or studies performed in both wet and dry conditions also require caution. Further study may reveal additional footprint variables that are not as sensitive to the effects of substrate or slope.

In many field situations, it may be impractical to apply this method to prints occurring on natural substrates. Under such circumstances this method could still be employed for mark–recapture census studies, using sand-traps, or another suitable artificial substrate, to collect footprints from defined locations (S. Alibhai, pers. comm.; Karanth, 1995).

Misclassifications of footprints from a field study will not be obvious, since learning is unsupervised and thus individual identities are not known prior to analysis. Several misclassifications occurred in the results from AutoClass presented here. For snow leopard RHF, to take the highest misclassification rate, these results could be interpreted as five animals occurring in the study with 15 prints from the first two animals, 11 from the third and 10 from the last two. The actual situation was that six animals produced the prints, with 11 footprints from the first animal and 10 from all others. In a carefully designed field study, one would aim to acquire as many footprints from a set of tracks as possible, since all tracks in a set are known to have been made by one animal.

Footprints that are classified into classes that do not contain the majority of prints from the same set can be ignored, since they must be misclassifications. Small classes, containing complete sets of prints, such as Nari in the LHF classification of Bengal tigers, must be treated cautiously. By disregarding small classes as unrepresentative cases, one risks omitting what may be an additional animal. Such situations may occur at the edges of the study site or may be the result of a dispersing animal, passing through the area. Careful examination of the spatio-temporal distribution of such prints may give clues as to their reliability.

As the number of animals included in the classification increases, the variability within any particular measurement, across all individuals, is likely to diminish. This may in turn lead to a reduction in overall classification accuracy. All measurements will be bounded by maxima and minima, which will be species typical. Improvements in classifications involving a potentially large number of individuals may be gained by weighting those measurements that retain the greatest variance.

In presenting this study, I have established that it is possible to accurately identify unknown individual tigers and snow leopards from their footprints using unsupervised learning techniques. I have demonstrated that with further research, this method could be used in field work to provide a reliable classification of individual animals from a study area, which may be particularly important in tiger conservation (Nowell & Jackson, 1996). It must be emphasized that in a field situation, a classification using AutoClass can not rely solely on the footprints presented to the classifier. Additional information is likely to be important in estimating the reliability of the resulting classifications. This method need not be restricted to carnivore research. Any animal with sufficiently complex footprints could potentially be classified using AutoClass, though preliminary studies would be needed. Footprint classification may not be suitable for smaller animals, since the precision with which track variability can be expressed will decrease with print size. One could envisage the situation where measurement error exceeds the inter-individual differences in footprint dimensions. Classifications of a variety of animals of differing size may indicate where this threshold lies. Irrespective of these issues, I believe that this technique could be a significant tool in animal ecology and conservation.

Acknowledgements

For commentary and advice throughout this project I thank Dawn Burnham, Alan Fielding and Barry Stevens-Wood. For reading and commenting on this manuscript I also thank Rory Putman. I thank the staff at Port Lympne Zoo, in particular Adrian Harland, Simon Nash, Neville Buck and Mike Lockyer, who provided essential logistical support. Thanks also go to Mick Hoult at MMU for his technical assistance. Financial support for data collection was provided by Manchester Metropolitan University.

REFERENCES

Cheeseman, P. & Stutz, J. (1995). Bayesian classification (AutoClass): theory and results. In *Advances in knowledge discovery and data mining*. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. & Uthurusamy, R. (Eds). Menlo Park, CA: AAAI Press.

Clevenger, A. P. (1993). Sign survey as an important tool in carnivore conservation research and management programmes. In *Seminar on the management of small populations of threatened mammals. Convention on the conservation of European wildlife and natural habitats (Bern convention):* 36–46. Strasbourg, France: Council of Europe.

Geman, S., Bienenstock, E. & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Comput*. **4 (1)**: 1–58.

Hanson, R., Stutz, J. & Cheeseman, (1991). *Bayesian Classification Theory*. Technical Report FIA-90-12-7-01, NASA Ames Research Center, Artificial Intelligence Branch.

Karanth, K. U. (1987). Tigers in India: a critical review of field censuses. In *Tigers of the world – the biology, biopolitics, management and conservation of an endangered species:* 118–132. Tilson, R. L. & Seal, U. S. (Eds.). Park Ridge, NJ: Noyes Publications.

Karanth, K. U. (1995). Estimating tiger *Panthera tigris* populations from camera–trap data using capture-recapture models. *Biol. Conserv*. **71**: 333–338.

Kohonen, T. (1989). *Self-organization and associative memory*. Berlin: Springer-Verlag.

Kohonen, T., Kangas, J. & Laaksonen, J. (1992*). SOM_PAK: the self-organizing map program package, version 1.0*. Helsinki, Finland: Helsinki University of Technology.

Komers, P. E. & Brotherton, P. N. M. (1997). Dung pellets used to identify the distribution and density of dik-dik. *Afr. J. Ecol*. **35 (2)**: 124–132.

Koster, S. H. & Hart, J. A. (1988). Methods of estimating ungulate populations in tropical forests. *Afr. J. Ecol*. **26**: 117–126.

Lawrence, S., Tsoi, A. C. & Back, A. D. (1996) Function approximation with neural networks and local methods: bias, variance and smoothness. In *Australian conference on neural networks, ACNN '96:* 16–21. Bartlett, P., Burkitt, A. & Williamson, R. (Eds.). Canberra: Australian National University.

Liebenberg, L. (1990) *A field guide to the animal tracks of Southern Africa*. Cape Town: David Philip.

Murtagh, F. & Hernandez-Pajares, M. (1995) The Kohonen self-organizing map method: an assessment. *J. Classif*. **12 (2)**: 1–21.

Nowell, K. & Jackson, P. (Eds) (1996). *Wild cats: status survey and conservation action plan*. Gland, Switzerland: IUCN.

Panwar, H. S. (1979) A note on tiger census techniques based on pugmark tracing. *Tigerpaper* **6**: 16–18.

Petrak, M. (1990). Habitat use as assessed by vegetation survey. In *Methods for the study of large mammals in forest ecosystems:* 22–31. Groot Bruinderink, G. W. T. A. & van Wieren, S. E. (Eds). Arnhem, The Netherlands: Research Institute for Nature Management.

Putman, R. J. (1984). Facts from faeces. *Mamm. Rev*. **14**: 79–97.

Putman, R. J. (1990). Patterns of habitat use: an examination of the available methods. In *Methods for the study of large mammals in forest ecosystems:* 22–31. Groot Bruinderink, G. W. T. A. & van Wieren, S. E. (Eds). Arnhem, The Netherlands: Research Institute for Nature Management.

Reibnegger, G., Weiss, G. & Wachter, H. (1993) Self-organizing neural networks as a means of cluster analysis in clinical chemistry. *Eur. J. Clin. Chem. Clin. Biochem*. **31**: 311–316.

Ripley, B. D. (1993) Statistical aspects of neural networks. In *Networks and chaos – statistical and probabilistic aspects:* 40–123. Barndorff-Nielsen, O. E., Jensen, J. L. & Kendall, W. S. (Eds). London: Chapman and Hall.

Servin, J. I., Rau, J. R. & Delibes, M. (1987). Use of radio tracking to improve the estimation by track counts of the relative abundance of red fox. *Acta Theriol*. **32 (30)**: 489–492.
Smallwood, K. S. & Fitzhugh, E. L. (1993). A rigorous technique for identifying individual mountain lions (*Felis concolor*) by their tracks. *Biol. Conserv*. **65**: 51–59.
Smallwood, K. S. & Fitzhugh, E. L. (1995). A track count for estimating mountain lion, *Felis concolor*, Californian population trend. *Biol. Conserv*. **71**: 251–259.
Stutz, J. & Cheeseman, P. (1994). AutoClass – a Bayesian approach to classification. In *Maximum entropy and Bayesian methods, Cambridge, 1994:* Skilling, J. & Sibisi, S. (Eds). Dordrecht, The Netherlands: Kluwer Academic Publishers.
Van Dyke, F. G., Brocke, R. H. & Shaw, H. G. (1986). Use of road track counts as indices of mountain lion presence. *J. Wildl. Mgmt*. **50 (1)**: 102–109.
Varfis, A. & Versino, C. (1992). Clustering of socio-economic data with Kohonen maps. *Neural Net. Wld* **2**: 813–833.
Zar, J. H. (1984) *Biostatistical analysis*. Upper Saddle River, NJ: Prentice-Hall International, Inc.
Zielinski, W. J. & Stauffer, H. B. (1996). Monitoring *Martes* populations in California – survey design and power analysis. *Ecol. Applic*. **6 (4)**: 1254–1267.