Martes in Carnivore Communities, pages 177-202
M. Santos-Reis, J. D. S. Birks, E. C. O'Doherty, and G. Proulx, editors, 2006
Alpha Wildlife Publications, Sherwood Park, Alberta, Canada

A QUALITY-SCORING SYSTEM FOR USING SIGHTINGS DATA TO ASSESS PINE MARTEN DISTRIBUTION AT LOW DENSITIES

SIMON POULTON
BioEcoSS Ltd, 1 Granary Steps, Bridgnorth, Shropshire WV16 4BL, UK.

JOHNNY D. S. BIRKS[1]
The Vincent Wildlife Trust, 3 & 4 Bronsil Courtyard, Eastnor, Ledbury, Herefordshire HR8 1EP, UK

JOHN E. MESSENGER
The Vincent Wildlife Trust, 3 & 4 Bronsil Courtyard, Eastnor, Ledbury, Herefordshire HR8 1EP, UK

DON J. JEFFERIES.
The Vincent Wildlife Trust, 3 & 4 Bronsil Courtyard, Eastnor, Ledbury, Herefordshire HR8 1EP, UK

Abstract: Rare carnivores live at low densities and detection problems make cost-effective census programmes over wide areas difficult. We developed a system based on structured interviews to record details of *ad hoc* sightings of pine martens (*Martes martes*) reported from England and Wales by casual observers in response to targeted publicity. Each of 525 sightings made during 1996-2003 was subjectively quality-scored following interview. We used computer-intensive methods to evaluate the system and build confidence in its ability to assess pine marten distribution. Objective scores were calculated using weighted averages produced by a genetic algorithm. Mantel tests were used to confirm highly significant intrinsic spatial clustering and extrinsic association with earlier, independent data on marten distribution. These relationships held even when the distribution of publicity was accounted for. Both subjective and objective high-scoring sightings exhibited greater intrinsic clustering than low-scoring ones; only objectively scored sightings showed a similar trend in extrinsic association with known marten distribution. A major benefit of quality-scoring following structured interviews is the capacity to account for possible species misidentification by observers. The system is highly cost-effective: over 8 yrs it absorbed <200 staff-days; this compares with an estimate of 370 yrs of equivalent professional survey time required to gather the same number of sightings. We used high-scoring sightings to produce a predictive map of pine marten distribution that could not be explained in terms of the occurrence of other species that might be misidentified as martens. It revealed 6 concentrations, 5

[1] Corresponding author; johnnybirks@vwt.org.uk

of which corresponded with marten populations identified by earlier studies. We conclude that this approach can be used reliably to produce information on distribution.

Introduction

Carnivores of high conservation interest tend to occur at low densities over much of their range, which creates difficulties for biologists planning census and monitoring programmes (Macdonald et al. 1998, Gese 2001). Low density results in low contact rates, so professional sign surveys or live-trapping programmes may not generate reliable distribution data over wide areas at reasonable cost. Further, legal protection associated with high conservation value means that returns from lethal trapping cease to be available and details of specimens killed unintentionally during predator control are typically suppressed (Jefferies and Critchley 1994). These conditions hinder biologists' attempts to detect pine martens (*Martes martes)* in some areas. For example, in parts of Britain there are sparse and persistent populations, some occupying sub-optimal habitat, that are extremely difficult to detect with standard survey methods (Strachan et al. 1996, Messenger et al. 1997, Messenger and Birks 2000, Birks et al. 2004). Detecting martens under these conditions represents an unusual challenge beyond the experience of most *Martes* biologists accustomed to working on well-established populations. In such circumstances, alternative detection methods are required that are cost-effective and reliable.

Compilation of reported sightings was an accepted way of detecting pine martens and assessing distribution (e.g. Howes 1984, Strachan et al. 1996) prior to the development of more systematic census techniques (Velander 1983, Bright and Harris 1994, McDonald et al. 1994). However, the value of collating sightings by casual observers has been questioned recently on the grounds that such records cannot be verified (McDonald et al. 1994). These concerns are lent some credence by a decline in the numbers of dead martens (i.e. specimens) recovered during the 20th century (Strachan et al. 1996) that previously provided corroboration for sightings data. However, Strachan et al. (1996) suggest this decline may reflect the changing perceptions and behaviour of gamekeepers, making them less likely to report pine martens they had killed, due to the species' acknowledged rarity in the mid-1900s and its partial and total legal protection in Britain, in 1982 and 1988 respectively.

Recent studies of rare carnivores have demonstrated that valuable distribution data can be derived from sighting reports obtained from interviews and questionnaires (e.g. Easterbee et al. 1991, Palma et al. 1999, review by Gese 2001). Conditions in southern Britain are well suited to the application of a pine marten recording system based on sightings reported by casual observers. Average human population density in England and Wales is 198 people/km^2 (Office for National Statistics 2003), making it the fourth most densely populated country in the pine marten's range and double the European average (United Nations 2002). Moreover, public interest in wildlife, wilderness and the countryside is high in Britain, as indicated by the 6 million members of environmental organisations (Office for National Statistics 2003) and the 92.5 million visitor days per year to national parks in England and Wales (ANPA 2001). Finally and importantly for a sightings-based system, pine martens are not strictly nocturnal. Daytime activity is common in summer (Zalewski 2000) when most human visits to the countryside occur. Under these conditions occasional daytime encounters between people and pine martens may be expected to occur, even where the latter species is rare, and a proportion of sightings are likely to be reported in response to appropriate publicity appeals.

Where other affordable detection techniques are likely to fail, *ad hoc* sightings cannot be ignored as a potential source of information on pine marten presence. However, in the absence of frequent corroborative evidence, caution must be applied when using sightings from casual observers to determine pine marten distribution (Strachan et al. 1996). This is especially important where the target species is rare and observers are unfamiliar with its appearance, leading to the risk of misidentification and erroneous reports. Therefore, a discriminatory system is required that enables biologists to make use of reliable sightings with confidence, while minimising the influence of erroneous reports. In this paper we describe the development, operation and evaluation of such a system.

The overall aim of this study was to evaluate the use of *ad hoc* sightings of pine martens by members of the public to assess the species' distribution in England and Wales. Specifically, the objectives were to collate reports of sightings using a structured questionnaire-interview to evaluate each sighting and assign it a quality-score based on available information, and to evaluate the pattern of quality-scored sightings using an intrinsic measure and by extrinsic comparison with independent data on pine marten distribution.

Methods

Publicity

Among the general public the pine marten is commonly perceived as extinct in England and Wales, and even among naturalists, this perception was not unusual during the 1990s (Messenger et al. 1997). Indeed, very few mammalogists have ever seen an English or Welsh pine marten in the field. Furthermore, no organisation had actively sought records of pine martens widely in England or Wales since work in the late 1980s by Strachan et al. (1996). Under these circumstances a new appeal for information was required in order to encourage people to report sightings to an appropriate authority.

During the second half of the 1990s, The Vincent Wildlife Trust (VWT) produced and distributed copies of posters and a leaflet appealing for recent evidence of pine martens in England and Wales. This appeal was promoted further through illustrated talks, articles in magazines and newspapers and radio and television interviews. To maximize the value of this publicity we attempted to target it both geographically and by interest group. We broadly targeted northern England and Wales because these included the main areas of known pine marten range (Strachan et al. 1996, Messenger et al. 1997). Within this geographical distribution we targeted those people most likely to encounter, recognise and report martens through postal distribution of publicity material to organisations concerned with wildlife, farming, forestry and access to the countryside. The destination of all targeted publicity was recorded for use in subsequent analyses. In addition to this controlled publicity effort, the appeal for pine marten sightings was promoted more widely at regional and national levels by announcements on television, radio and in newspapers.

The questionnaire interview

Some people reporting sightings of pine martens might be mistaken in their judgement, so it is unsafe to accept such reports uncritically. Therefore, a sightings-based recording system requires a mechanism for discriminating between those reports that probably relate to pine martens and those that are more likely to relate to other species. We designed a structured questionnaire (Appendix A), based on 21 questions, that could be completed by an experienced interviewer in approximately 5 minutes during an informal conversation conducted either over the telephone or in person. Most interviews (87%) were carried out by 2 of the authors (JM and JB); the remainder were conducted by a small number of experienced naturalists who were familiar with pine martens and who had been trained to use the interview technique. The main objective of the interview was to gather information in ways that would enable us to determine the extent to which a reported sighting of a pine marten could be considered reliable. In addition to recording essential details such as the date, time and location of a sighting, non-leading questions were designed to gather information on the following 3 main aspects that we believe influence the reliability of each reported sighting:

Conditions and context of the sighting

We asked for information on how clearly the animal was observed. This included such details as the length of time the animal was in view, the distance between the observer and the animal, the prevailing light and weather conditions and whether the animal was viewed through binoculars.

Knowledge and experience of the observer

During the interview we attempted to establish the level of an observer's knowledge and experience of wildlife in general and pine martens in particular. We asked observers to explain their reason for believing the animal they saw was a pine marten and whether they were confident that the animal they had observed was not one of the other species that might be mistaken for a pine marten (e.g. polecat *Mustela putorius*, feral polecat-ferret *M. furo*, fox *Vulpes vulpes*, stoat *M. erminea,* weasel *M. nivalis*, mink *M. vison*). After completing the interview and assessing an observer's knowledge, experience and ability to describe the animal reported, each was categorised as follows: where applicable observers were identified as either 'amateur naturalist' or 'countryside professional' (e.g. gamekeeper, forestry worker, farmer); the remainder were classified according to their skill levels as 'other skilled observer', 'semi-skilled observer' or 'unskilled observer'.

Description of the animal

We used non-leading questions to encourage observers to provide as full and accurate a description as possible of the animal they had seen. Questions requested details of the animal's size, shape, colouration, behaviour and movement pattern. In particular we noted any reports of extreme agility, such as animals climbing trees or other vertical surfaces with ease.

Subjective quality-scoring

Having gathered and assessed all relevant information about the reported sighting, the interviewer assigned it a quality score on a scale of 1-10. A score of 0 was assigned to sightings where insufficient information was available to reach a judgement about quality. Sightings that were definitely not of pine martens (for example, where a photograph or specimen was available confirming that another species was involved) were assigned a score of 1. Conversely, where a photograph or specimen confirmed that the sighting was of a pine marten, a quality score of 10 was assigned. In all other cases quality scores of 2-9 were assigned subjectively according to the level of confidence with which a sighting could be defended as a pine marten on the basis of the information gathered.

The description of the animal and its behaviour was a dominant factor in deciding which score to assign to a sighting. For example, a full description of a pine marten, including accurate statements about body size, shape, colouration and markings would be given a higher score than an inaccurate or incomplete description. Where an accurate description included behaviour that was typical of martens, such as climbing a vertical surface, this would further elevate the score. The score was also influenced by the observer's experience as a naturalist and the conditions under which the sighting was made. For example, a good description of a pine marten reported by an experienced naturalist familiar with the species would generally score higher than the same description offered by a less experienced observer. Sightings of animals recorded in good daylight at close range and over several seconds tended to score higher than fleeting glimpses of animals seen at some distance in poor light.

As a general rule, interviewers assigned quality scores of ≥6 where they believed the animal was probably a pine marten on the basis of information gathered during the interview. Quality scores of <6 were assigned to sightings where interviewers believed the animal was probably not a pine marten or where the information provided did not enable them to defend it as such. Therefore, some lower scoring sightings may quite possibly have related to pine martens, but these records could not be defended with confidence because of a paucity of supportive information. Quality scores were not normally divulged to the observers.

Data analyses

Interviewer Effects – During the interview, the interviewer inevitably became aware of the geographical location of the sighting. To investigate whether this had an effect on the subsequent score, we carried out a subsidiary experiment on the 2 main interviewers (JB and JM). Each selected 20 completed sighting questionnaires from those they had carried out, 5 each from 4 regions: southern Wales, northern Wales, northern England and the rest of England. The questionnaires were photocopied with all their results visible, but with all geographical information removed. The other interviewer was then asked to score the 20 questionnaires without any "geographical pre-knowledge" available to him. The data were then analysed with two-way ANOVA, including the interaction term Interviewer x Region.

Calculation of Objective Scores – In an attempt to remove the subjectivity of the scoring system, an "Expert System" was devised. It was based on the subjective scores, but provided an objective, mathematical process for calculating scores based on the individual questions in the questionnaire. The objective scores were calculated as simple weighted averages of the individual elements of the questionnaire in the following way. Eight of the 21 questions gave categorical answers that were converted to 45 binary variables. These were combined with the 13 "Yes/No" questions to give a total of 58 binary variables. Each variable was assigned a weighting (an integer from 0 to 10) and the objective score for each sighting record could then be calculated:

$$\text{Objective Score} = \text{int}\left(\sum_{i=58}^{i=1}(W_i \times V_i) \Big/ 21\right)$$

where i = the 58 binary variables, W = the weighting for variable i, and V = the value for variable i in that questionnaire.

Clearly, as V could only take values of 0 or 1, this was equivalent to averaging the weightings for the original 21 questions. To derive the weightings, a genetic algorithm was developed using the rationale in Forsyth (1987). Starting with uniform weightings of 5 for all the binary variables, the algorithm "bred" a "population" of sets of weightings over 10,000 generations. In each generation, a "breeding" population of 30 sets of weightings survived from the previous generation. From these, the algorithm randomly selected weightings for each variable, as well as generating a number of new weightings entirely randomly, and combined them into another 30 sets of weightings. This population of 60 sets of weightings was then tested individually against a loss-function that minimised the differences between the objective and subjective scores:

$$= abs\left(\sum_{i=525}^{i=1}(OS_i - SS_i)\right)$$

where i = each of the completed questionnaires, OS = the Objective Score and SS = the Subjective Score for the ith questionnaire.

In each generation, the 30 sets with the smallest values for the loss-function survived to constitute the breeding population in the next generation; otherwise they were eliminated. Over the full cycle of 10,000 generations a population of sets of weightings was created that best approximated to the subjective scores. In this way, the Expert System combined the knowledge and experience used by all the interviewers in assigning scores subjectively, but removed the individual biases.

Data collation

All sighting records and independent data were geo-referenced to 10-km Ordnance Survey grid squares, because this was the minimum common resolution. Records were also tallied to 20-km, 50-km and 100-km grid squares and tested for normality. All 4 resolutions were highly non-normal due to the number of squares with 0 records. However, as no parametric testing was carried out on the geo-referenced data, 50 km was chosen subjectively as the resolution that gave the most useful combination of sample size with minimal 0 records. The basic dataset was, therefore, the tally of sightings in the 92 50-km squares covering England and Wales, hereafter referred to as the Primary Units. All publicity records were similarly collated and tallied into the Primary Units. Maps were produced using Dmap (Morton 1998).

The analysis of spatial distribution of questionnaire records

Two types of tests for the spatial distribution of the questionnaire sightings were used. Both were based on the Mantel test (Mantel 1967; *see also* Manley 1991) and used random permutations to provide significance levels for the test statistic. The first was a test for spatial clustering. It can be considered an intrinsic test as it simply identifies significant patterns in the geographical distribution of the sightings themselves. This was used to test the null-hypothesis that there was no clustering of sightings, i.e. that they were being reported in a random manner from across England and Wales.

Conversely, the second type of test can be defined as extrinsic, because it compared the spatial distribution of questionnaire sightings with an earlier (1979-1988), independent set of 146 pine marten records compiled by D. J. Jefferies (unpublished). This dataset was part of a long series (1800-1988) of 861 records used by Strachan et al. (1996) to assess changes in the distribution of pine martens in England and Wales prior to the current study. These records were gathered in ways that differed significantly from the approach adopted in the present study; instead of a widespread appeal among the general public, records were gathered predominantly through informal approaches to local naturalists, through letters to museums and other relevant organisations and via a trawl of naturalists' publications. Because of this different approach and the absence of any chronological overlap, we regard this earlier dataset as independent of the questionnaire sightings gathered by the VWT. For the extrinsic tests we chose to use the most recent decade of the earlier series in order to minimise the possibility that shifts in pine marten distribution over time might explain differences between the spatial distribution of our questionnaire sightings and the earlier dataset. The 146 records were tallied into the Primary Units and the tests were based on the null-hypothesis that there was no spatial association between the distributions of the VWT and D. J. Jefferies datasets.

Mantel tests are essentially correlations between distance matrices. One of these is always a true Euclidean distance matrix, representing the real distance in space from each case to every other. In this analysis, the basic dataset comprised tallies of sightings in the 92 Primary Units in England and Wales. This resulted in a 92 x 92 cell matrix with 4,186 ([92 x 91]/2) unique elements. Euclidean distances were calculated from the centroids of the Primary Units where they were entirely inland. But where they encompassed the coast or the border with Scotland, the centroids were calculated from the geometric centre of the constituent 10-km squares. Clearly, this is an approximation, but is adequate for the purposes of this analysis. The "test" matrix was calculated in exactly the same way, except that instead of Euclidean distances, the matrix comprised distances in "sightings" space. For example, Primary Unit NT50 had 12 sightings and Unit NU00 had 16, so the "distance" in this case was 4.

The full Mantel test calculates the correlation coefficient between these 2 distance matrices. This approach has been used for the intrinsic test with the simple hypothesis of spatial clustering. However, the more powerful partial Mantel Test calculates the partial correlation coefficient between 2 distance matrices whilst accounting for the effects of a third (or more) matrix. This has been used in the intrinsic test accounting for the effects of publicity and in all the extrinsic tests (Table 1).

Table 1. Summary of the distance matrices used in the various Mantel tests in this analysis.

Type of test	***Test***	***Distance matrices***	
		Permuted	***"Held Constant"***
Intrinsic	Spatial clustering of sightings	Sightings	Space
	Spatial clustering of sightings, accounting for publicity	Sightings	Space and Publicity
Extrinsic	Association between sightings and known distribution	Sightings	Known distribution and Space
	Association between sightings and known distribution, accounting for publicity	Sightings	Known distribution, Space and Publicity

Because the data in distance matrices violate the assumptions of normality required for parametric testing, randomisation techniques were used to obtain significance levels for the test statistics. In all analyses, the permutations were carried out on the sightings data (note that it is only necessary to permute one variable, even if a partial Mantel test is utilised). For each permutation, all candidate sighting records were randomly assigned to one of the 1,714 10-km grid squares in England and Wales. These were then tallied to 50-km squares and the sightings distance matrix recalculated. Depending on the test, either the full or partial correlation coefficient was calculated and stored. This process was repeated 9,999 times and the distribution of the permuted test statistics was compared with the actual test statistic derived from non-permuted data.

Estimation of equivalent professional time in the field

In order to assess the cost-effectiveness of using sightings reported by casual observers, we estimated the equivalent time required for a professional surveyor to gather the same number of sightings. This involved making the assumption that professional surveyors would have similar marten-sighting opportunities to the many diverse and widely-distributed casual observers that responded to the survey. We further assumed that 3 categories of casual observer differed in respect of the hours per week and weeks per year that they spent in the field, and in terms of how many years they had been aware of the sightings survey (these values are defined in columns 3 to 5 in Table 2). Simple multiplication of these values by the respective number of casual observers in each category produced estimates of the total 'observer time' in hours and years that produced the sightings upon which this survey was based.

Table 2. Estimates of the equivalent 'observer time' in the field that generated the 525 sightings upon which this survey was based.

Occupation or skill level	***n***	***Hours per week***	***Weeks per year***	***Years aware of scheme***	***Equivalent professional time***	
					Hours	***Years***
Countryside professional	79	30	45	4	426,600	267
Amateur naturalist	112	10	50	2	112,000	70
Skilled / semi-skilled / unskilled	215	5	50	1	53,750	34
TOTAL	**406**				**592,350**	**370**

Results

For the purposes of this analysis, 694 reports of pine martens were originally considered. These comprised all sightings or specimens obtained from 1990 to 2003 for which a subjective score had been assigned at interview.

Interviewer effects

No effects were found to be significant. However, both interviewers showed a small tendency to score sightings from north Wales (a known pine marten core area) slightly higher when they knew the location than when scoring blind. Although this was not significant, we point it out simply to caution against psychological bias creeping into the scoring process. It is important to recognise, however, that the ANOVA tests assumed true randomness of sighting quality between interviewers and this may not have been the case.

Subjective versus objective scores

The relationship between subjective and objective scores showed a very strong pattern over time. For sightings before 1996 there was virtually no relationship between subjective and objective scores on an annual basis (Figure 1a). However, from 1996 onwards, the relationships were very strong (Figure 1b). One explanation for this effect is that because the survey started in 1996, the interval between interview and original sighting increased linearly for pre-1996 sightings. This had a small but significant effect ($P = 0.041$), on the differences, so we excluded the 169 records from the pre-1996 period, leaving 525 for this analysis.

Spatial distribution of the questionnaire sightings

The distribution of sightings was highly clustered (Mantel test for spatial aggregation; $P < 0.0001$) (Figure 2). The 2 main centres were in Wales and the far north of England, although there appeared to be another smaller focus in the Peak District of central England. South-east of a line from the Humber to the Severn estuaries there were only scattered records, although these were quite widespread. Of the 92 squares, 60 contained at least one sighting.

Accounting for the distribution of publicity

The 942 targeted publicity incidents were distributed largely in Wales and northern England (Figure 3). However, all parts of the country were covered, with 84% of the Primary Units having some publicity within them (those that had no publicity were largely coastal and had very small areas of land within them.) To investigate whether the clustered distribution of sightings was simply a response to the clustered distribution of publicity, a partial Mantel test was carried out including a "Publicity" distance matrix. Even accounting for publicity there was still a highly significant ($P < 10^{-5}$) spatial clustering of sightings. This can be displayed by adjusting the tallies shown in Figure 2 by standardised (mean and standard deviation of 1) publicity tallies (Figure 4). The main effect of accounting for publicity was to decrease the tallies in Wales and, because of the low publicity, to increase sporadically the tallies in eastern England. Tallies in northern England remained almost unaffected.

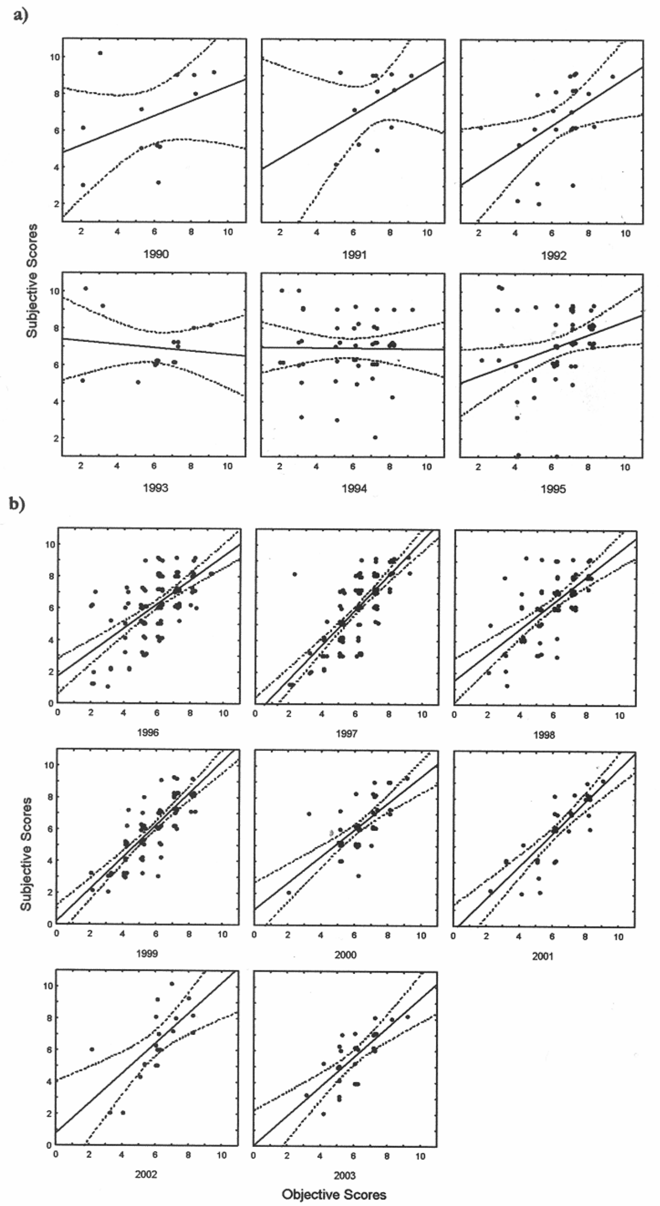


Figure 1. The relationship between subjective and objective scores for a) pre-1996 records and b) 1996 onwards. A simple linear regression with 95% confidence intervals on the regression line is shown. Data points have been jittered slightly to facilitate their display.

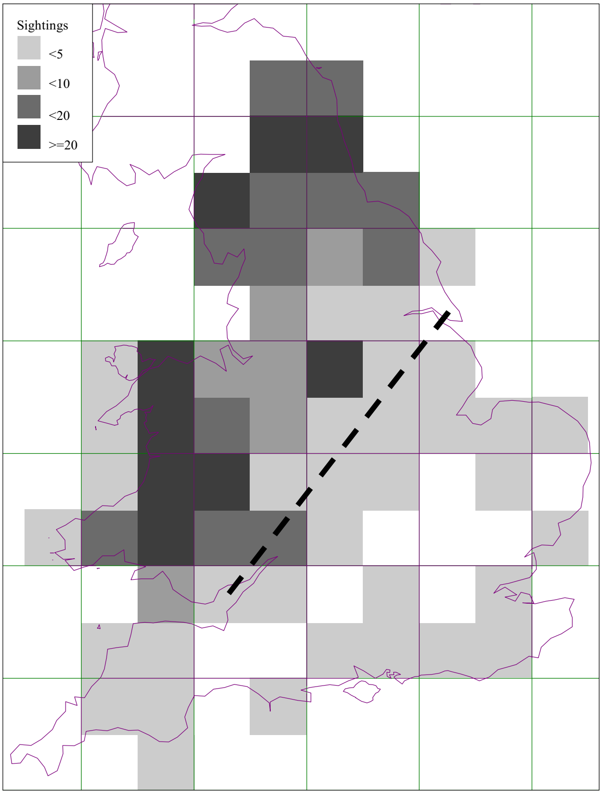


Figure 2. The distribution of sightings within the Primary Units (50-km Ordnance Survey grid squares). The dashed line joins the Humber and Severn estuaries.

Comparison of sightings with previously known marten distribution

The 146 records from D. J. Jefferies (unpublished) were tallied into the same Primary Units as the main dataset (Figure 5). A partial Mantel test of questionnaire sightings against the previously known distribution, accounting for the spatial aggregation already shown (Figure 2), was found to be very highly significant ($P < 0.0001$). Furthermore, even when the distribution of publicity was accounted for in the partial Mantel test, there was still a very highly significant spatial association between the questionnaire sightings and the previously known distribution of martens ($P < 10^{-5}$).

Intrinsic and Extrinsic tests on individual questionnaire scores

The purpose of scoring the questionnaire sightings was to attach an indication of quality to each sighting record. If this process works, we can hypothesise that the degree of clustering and the association between the VWT and D. J. Jefferies (unpublished) datasets will be greater for more highly scored sightings. Therefore, the sightings were sub-divided into six groups based, firstly, on their subjective (Figure 6) and, secondly, on their objective scores (Figure 7) (due to the small number of sightings with extreme scores, those scoring < 4 and > 7 were separately aggregated). The two types of Mantel test used in the previous analyses were then run on each score group in turn. These generated a partial correlation coefficient for each score, separately for each type of score. These were then used as the response variables in simple linear regressions against the scores.

Using the subjective scores (Figure 8a) with an intrinsic test, there was a significant positive relationship ($P < 0.020$). This indicated that the degree of clustering was greater for higher scoring sightings. Conversely, there was no significant relationship for the extrinsic tests ($P < 0.108$). For the objective scores (Figure 8b), both types of test were significant ($P < 0.026$ and $P < 0.023$ respectively). In particular, the extrinsic test indicated that there was a stronger spatial association between the high-scoring questionnaire sightings and previously known marten distribution than there was for the low-scoring sightings.

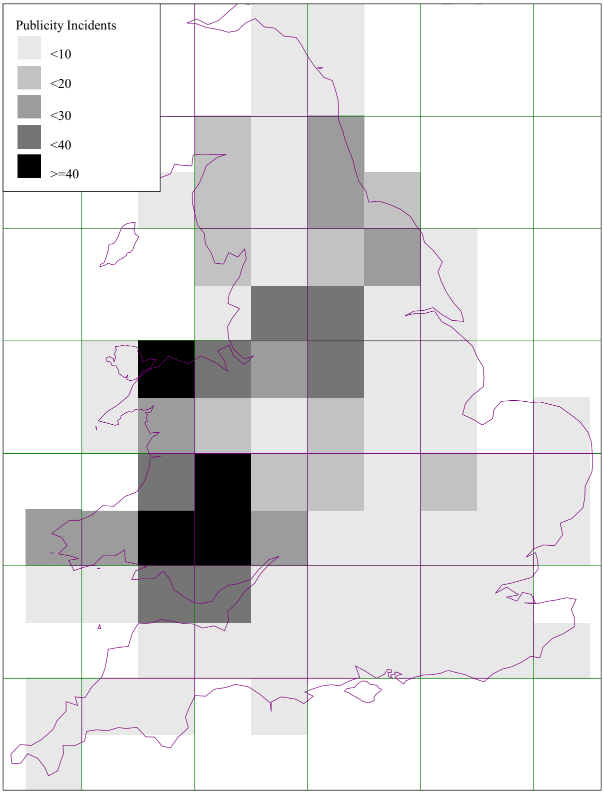


Figure 3. The distribution of publicity incidents tallied into Primary Units.

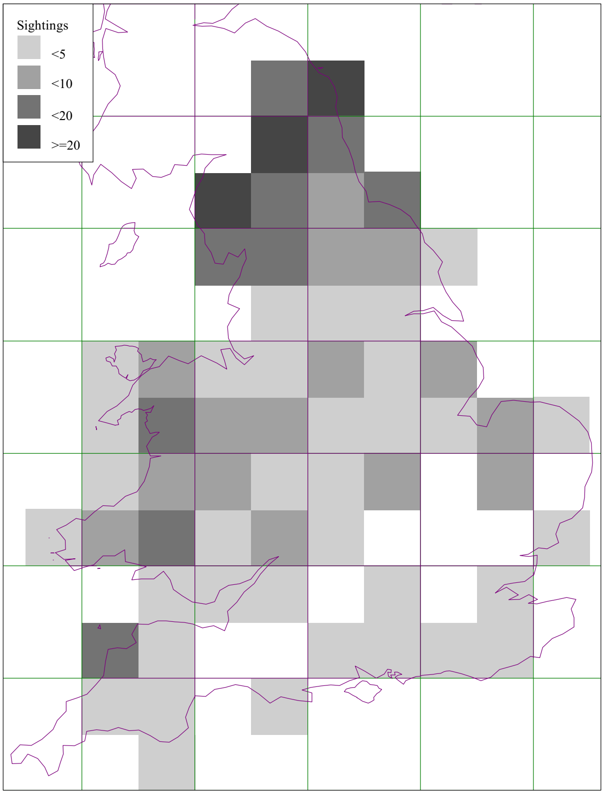


Figure 4. The distribution of sightings allowing for publicity.

Discussion

This study shows how casual sightings by members of the public can be used to derive information on the distribution of a rare carnivore. Although individual sighting reports are unverifiable and, therefore, disputable, the pattern of sightings recorded over several years has great value to wildlife policy and conservation. Even in countries where the sympatric and abundant presence of stone martens (*Martes foina*) raises questions about misidentification (not an issue in this study), data on very sparse pine marten populations have been gathered partly via oral inquiries and postal questionnaires (e.g. Álvares and Brito 2006, Matos and Santos-Reis 2006). The benefits of engaging the public in wide scale surveys of relatively common species have long been recognised by bird study groups (e.g. Royal Society for the Protection of Birds 2005). Less common species require a more cautious approach, yet the benefits may be greater. In the present study, the decision to involve members of the public and amateur naturalists as stakeholders in the survey has raised awareness of the existence of sparse pine marten populations in many positive ways (e.g. Fletcher 2004, Hawkins 2000).

Confidence in the use of sightings data depends upon careful collection of essential details via structured interviews, robust quality-scoring and rigorous testing. Some caveats should be explored because of the role of subjectivity and other characteristics of the system.

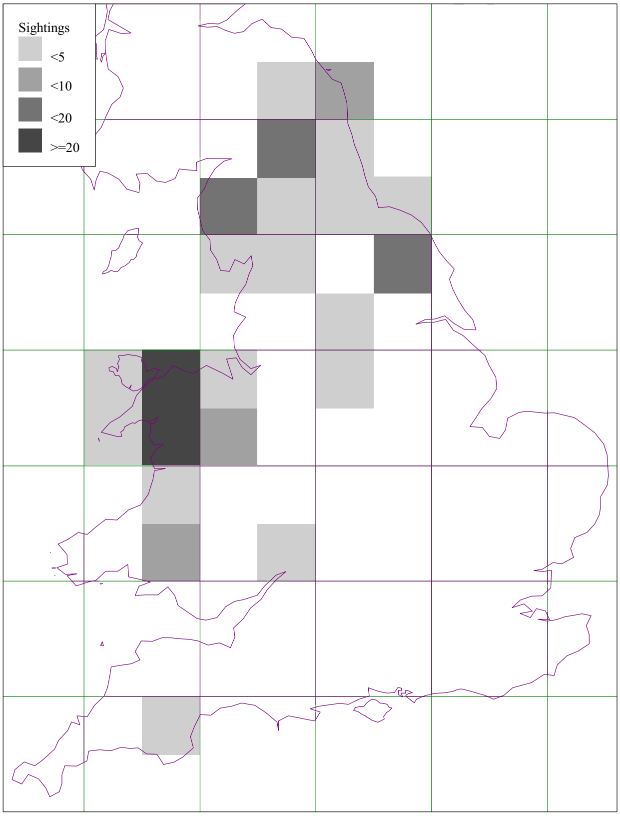


Figure 5. The distribution of 146 pine marten records from D. J. Jefferies (unpublished) tallied into Primary Units (50-km Ordnance Survey grid squares).

Figure 6. The distribution of sightings, tallied at 50 km resolution, broken down into subjective score groups.

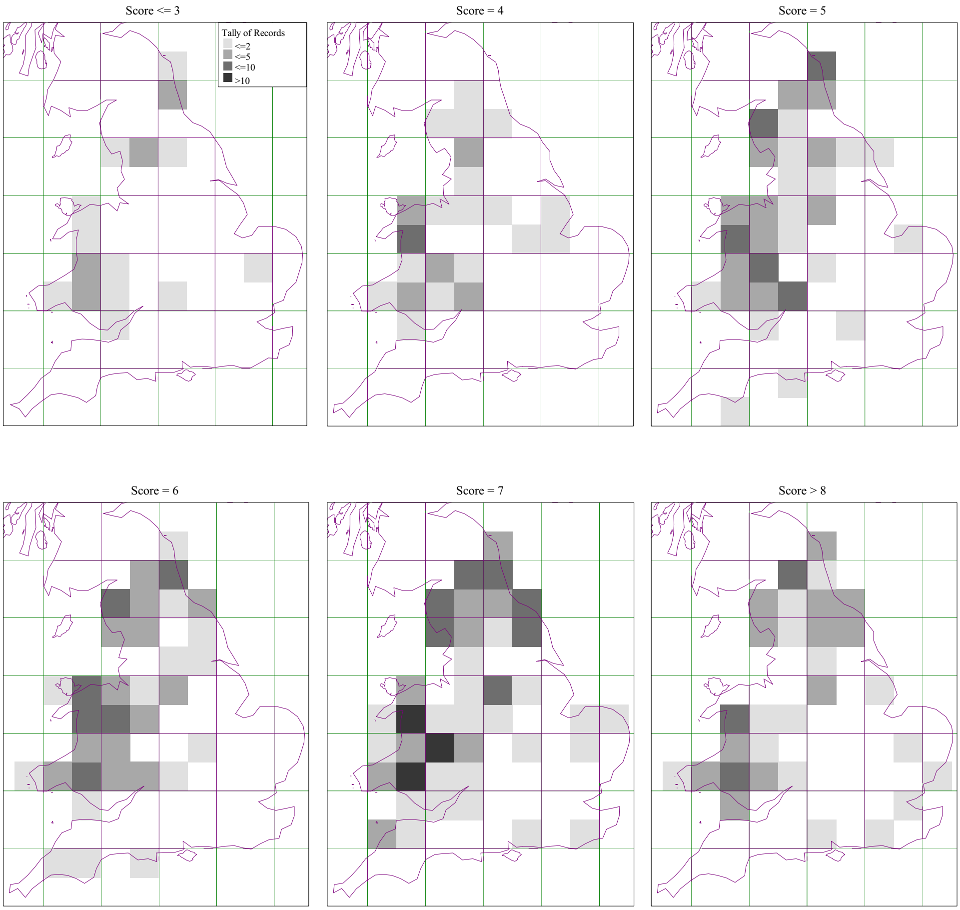
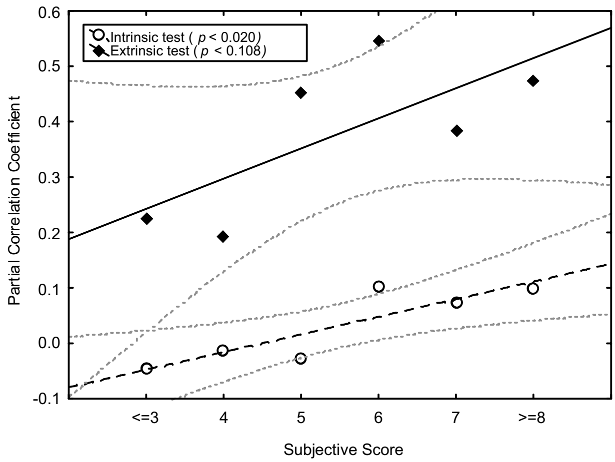


Figure 7. The distribution of sightings, tallied at 50 km resolution, broken down into objective score groups.

a)



b)

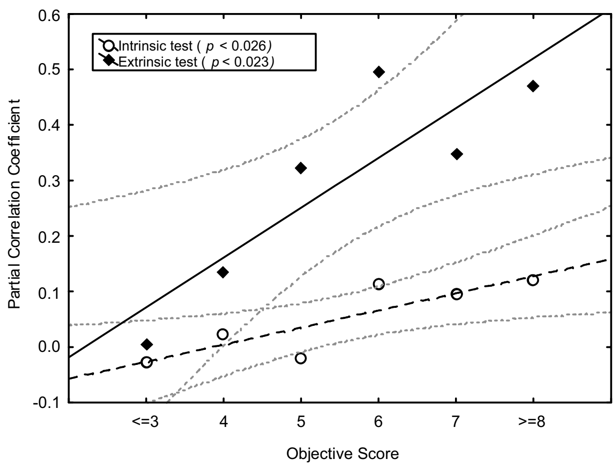


Figure 8. The relationships between intrinsic and extrinsic test statistics and questionnaire scores for a) subjective and b) objective scoring systems. Linear regression lines with 95% confidence intervals are given.

Publicity and respondent effects

A report of a marten sighting normally enters this recording system only when all of the following conditions are met: a) a person believes he/she may have seen a pine marten recently, b) that person is aware of requests to report such sightings (usually from exposure to publicity) and has access to the necessary contact details, and c) that person decides to contact the VWT and supply details when interviewed. At each of these stages a proportion of 'failures' is bound to occur. For example, martens may be seen by people who do not recognise them as such, by people who are unaware of the survey, or by people who do not bother to report their sightings. A likely consequence of such failures is that many human-marten encounters are never reported, especially when a well-targeted publicity effort is not maintained.

In this survey, widespread publicity in national or local newspapers tended to generate large volumes of low-scoring sightings reported by inexperienced observers who contacted the VWT speculatively rather than because they were confident they had seen pine martens. This supported our decision to target publicity through relevant organisations because less time would be spent processing records of limited value. Conversely, when interviewing enthusiastic amateur naturalists, we were conscious of the possibility that reports might be embellished, using prior knowledge of pine marten appearance and behaviour, in order to enhance the credibility of a record.

Another human factor that influenced the quality-scoring process was the differing ability of observers to describe their sightings articulately. In some cases, sightings that seemed to the interviewer to relate to pine martens could not be given high quality scores because of the very limited descriptions that were provided.

Finally, a significant respondent effect is unfamiliarity with pine martens among people in England and Wales and the associated tendency to misidentify other species as martens. An important role of the interviewers in this study, therefore, was recognising such mistakes and scoring them accordingly. Misidentifications probably comprised a large proportion of the 32% of sightings that scored < 6 on the quality scoring scale. Our confidence in the system's capacity to distinguish misidentifications is enhanced, firstly, by the strong association of high scoring records with the previously known pine marten distribution and, secondly, because the wild British mammals most likely to be misidentified as pine martens (red fox, polecat, feral mink and stoat) are characterised by patterns of both distribution (Corbet and Harris 1991) and predicted abundance (Macdonald et al. 1998) that differ markedly from those of the pine marten (Strachan et al. 1996).

Human activity effects

Beyond the effects of publicity, the distribution of reported marten sightings may be influenced by geographical patterns of human activity. If pine martens are commonest in areas or habitats visited least by people, then the distribution of reported sightings may exaggerate the importance of peripheral areas and habitats. Such an effect was noted by Macdonald et al. (1998) in their comparison of the distributions of red squirrel *Sciurus vulgaris* records and areas of suitable squirrel habitat in northern

England. Because people rarely visited the best areas of squirrel habitat (dense, remote woodland) and because squirrels are more visible in open or fragmented habitats, the authors concluded that the distribution of sightings reflected peripheral habitats with less well-established squirrel populations, and that a proportion of sightings related to dispersing individuals rather than established populations. The same bias is likely to influence the local distribution of pine marten sightings in the present study, so caution is needed when choosing the resolution at which the pattern of records is examined. However, we feel that such small-scale variations are obscured at the resolution of 50-km that was used for these analyses

The genetic algorithm

If the weighted averaging system for calculating the scores is considered to be a simple deterministic model, then the genetic algorithm can be treated simply as a process for parameterising the model. Although the algorithm was run for 10,000 cycles, 2 points regarding the resultant weightings should be made. Firstly, although the rate of improvement in the sets of weightings was highest in the first few hundred cycles (as is common with genetic algorithms) the "best" set of weightings was generated on cycle 9,781. This indicates that "good" weightings were still being found and more cycles might have been valuable. Secondly, the population of high-scoring rules often had quite different weightings for a single variable (some of which were counter-intuitive) even when the overall score for the sets was very similar. This indicates that many of the variables were redundant in the calculation of the score, but had not yet been "disabled" with an average weighting.

By setting the loss-function of the genetic algorithm to minimise the difference between the objective and subjective scores, it may appear that we are simply trying to mimic the subjective scores. However, this generates an Expert System that can assign a score to a questionnaire based on the expertise of a number of interviewers gained over several years. In this way, large surveys may be able to utilise less expert interviewers, as long as they record the answers to all the questions.

The merits of using objective scores

One of the main insights from producing the objective scoring system was that it highlighted the discrepancy between objective and subjective scores before 1996. There are 2 main inferences to draw here. Firstly, after the survey had started in 1996, sightings could have been made by members of the public who were already aware of the survey and might have been more vigilant, especially in noting details of the sighting. Conversely, before 1996 no members of the public could have been aware of the survey so, by definition, their sightings must have been *ad hoc*. The second point is that before 1996 the interval between sighting and interview increased by about 350 days per year; so the average interval between a sighting made in 1990 and the interview was about 2,200 days, whereas the interval for a sighting made in 1995 was only about 500 days. In contrast, after 1996 the average interval remained relatively constant at around 200 days. This factor was significant in explaining some of the score discrepancies in early sightings, and suggests that surveys of this sort should use only "current" sightings and not *post hoc* records. Indeed, there may be merit in setting a minimum time interval between sighting and interview.

Another advantage of using objective scores is that apparently anomalous subjective scores are smoothed out by the weighted averaging process. There may be cases where an expert interviewer can glean knowledge intuitively about the quality of a sighting that is not manifest directly in the answers to questions. However, this has the danger already discussed that previous knowledge of the location, or even of the person reporting the sighting, could influence the scoring in a biased way.

The main benefit of using the intrinsic tests on the distribution of sightings is that they require no assumptions about the actual distribution of pine martens. The assumption is simply that martens have a much more restricted range than the other species with which they could be confused, such as domestic cats and dogs, or foxes, mink, stoats or polecats. It follows from this that a higher degree of spatial clustering (i.e. less ubiquity) found for the higher scores suggests that the scoring system is identifying those sightings having a greater likelihood of being pine martens.

Cost-effectiveness of using casual observers

The cost-effectiveness of this system is emphasised by comparing the actual staff time involved in coordinating the present survey with an estimate of the time required for a professional surveyor to gather comparable information on pine marten presence. The survey was conducted on a part-time basis (total estimated staff-days = 192) over 8 years by 2 officers of the VWT. From their offices they collated 525 sightings made by 406 casual observers who were classified into 3 groups (Table 2). We have made some general assumptions about the time the different groups spent in the field, and the length of time they were likely to have been aware of the survey, to calculate the equivalent professional time. This equates to an estimated 592,350 hrs, or 370 yrs of a single professional surveyor's time. Clearly, this amount of effort would be entirely prohibitive for any conservation or research organisation.

Predicting pine marten distribution

Our confidence in the system encourages us to use higher scoring sightings to produce a predictive map of likely pine marten distribution (Figure 9). It is based on records with objective scores ≥ 6 and shows 6 areas of concentration, 5 of which correspond closely with the twentieth century populations identified by Strachan et al. (1996). The sixth area of concentration, first identified in southern Wales by Morgan (1992) and confirmed by the present survey, has only become well known as a focus of pine marten records since Strachan et al.'s (1996) collation. A further difference between the results of our study and Strachan et al.'s (1996) work is the greater geographical extent of reported sightings between the main concentrations apparent in our sample. This may indicate an expansion of the pine marten's range in England and Wales for the first time in a century. Significantly, its timing could correlate with the provision of full legal protection for the species by its inclusion on schedule 5 of the Wildlife and Countryside Act 1981 in March 1988.

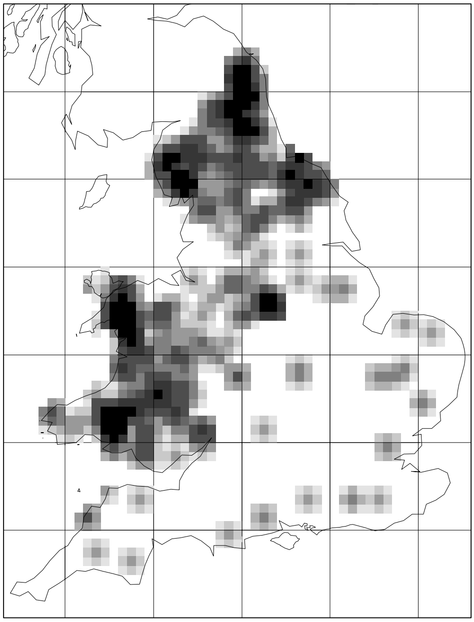


Figure 9. The distribution of reported marten sightings in England and Wales between 1996 and 2003 (n = 357) that scored ≥ 6 on the objective scale after interview. Sightings are tallied at 10-km square resolution, weighted (by score) and smoothed. Darker areas have higher concentrations of high scoring records.

In southern and eastern England, away from the main concentrations shown in Figure 9, are scattered marten records that require some explanation. These may be a consequence of the occasional escape or release into the wild of captive martens. However, historical marten presence offers another explanation. D. J. Jefferies (unpublished) collected 57 records (both specimens and sightings) of martens from 19 counties in southern and eastern England earlier in the 20th century. These records fell into three persistent groupings: 1. Welsh borders, 2. South West and south coast of England, 3. East coast from Lincolnshire to Essex (Strachan et al. 1996). Furthermore, division of the records into 2 periods, 1900-1945 and 1946-1988, showed the same 3 groupings in the same locations in both periods (Strachan et al. 1996). Their absence from the records of the decade 1979-1988 used for comparison in the present study (except for 2 records from Devon; see Figure 5) may have been due to the short period and so small number of records necessarily selected because of the requirement for a close chronological juxtaposition to the present study for statistical comparison. The presence of a third representation in 1996-2003 of the same 3 area groupings in southern England, as found in 1900-1945 and 1946-1988, provides additional evidence for the persistence of marten presence here.

Corroboration of sightings data

Reliance upon sightings data should not inhibit other approaches to gathering indisputable evidence of the presence of pine martens. Collection of photographs, specimens or other sources of DNA are valuable ways of confirming marten presence and corroborating sightings data. For example, pine marten presence was confirmed post-1990 by recovered specimens (Jefferies and Critchley 1994, Birks et al. 1997, Kyle et al. 2003), DNA from pine marten faeces (Davison et al. 2002) and a photograph of a live individual (Crawley and Birks 2004) within or close to 5 of the 6 sightings concentrations identified in Figure 9. Furthermore, since 2000, pine marten specimens have been recovered from 2 of the 3 persistent area groupings of marten sightings in southern England identified by Strachan et al. (1996) and discussed in the paragraph above (Forrest et al. 2002, Birks et al. 2005).

Conclusions

We have demonstrated that a structured questionnaire system can be applied effectively, over a period of years by experienced interviewers, to record the essential elements of sightings of pine martens where the species is scarce. This approach is highly cost-effective when compared with estimates for professional surveys required to gather equivalent information.

The distribution of high-scoring sightings shows a highly significant spatial clustering that is independent of the distribution of the original publicity. This confirms that such sightings are not being reported in a random manner that might indicate sightings of more ubiquitous species. Furthermore, this distribution is very highly associated with independent data on pine martens in England and Wales. The use of subjective scores assigned at interview, or objective scores calculated from the elements of the questionnaire is also highly beneficial. Questionnaires with high scores, objective scores in particular, have significantly higher degrees of spatial clustering and closer association with known marten distribution. This indicates that sightings by members of the public of scarce mammal species can be used to derive distribution information provided they are part of a well-designed survey, with carefully controlled publicity and adequate quality-scoring.

Finally, we assert that the significant association between the post-1995 marten sightings from this study and an earlier, independent dataset, and the evidence of possible range expansion since the period 1979-1988, confirms that the pine marten maintains self-sustaining populations in southern Britain.

Acknowledgements

We thank all who have helped The Vincent Wildlife Trust with this survey. Special thanks are due to Tony Braithwaite for his pursuit of marten records in Wales and for his helpful comments on a draft of this paper. We thank Hilary Macmillan for her help in preparing this paper.

Literature cited

Álvares, F., and J. C. Brito. 2006. The pine marten in North-western Portugal: habitat requirements and predicting areas of occurrence. Pages 29-45 *in* M. Santos-Reis, J. D. S. Birks, E.C. O'Doherty and G. Proulx, editors. *Martes* in carnivore communities. Alpha Wildlife Publications, Alberta, Canada.

ANPA 2001. The National Park Authority: A Guide for Members. Countryside Agency, Cheltenham, UK.

Birks, J. D. S., J. E. Messenger, and A. Davison. 1997. A 1994 pine marten *Martes martes* (L.) record for Lancashire, including a preliminary genetic analysis. Naturalist 122: 13-18.

Birks, J. D. S., J. E. Messenger, A. C. Braithwaite, A. Davison, R. C. Brookes, and C. Strachan. 2004. Are scat surveys a reliable method for assessing distribution and population status of pine martens? Pages 235 - 252 *in* D. J. Harrison, A. K. Fuller, and G. Proulx, editors. Marten and fishers (*Martes*) in human-altered environments: an international perspective. Springer Science+Business Media, New York, New York, USA.

Birks, J., J. Crooks, and M. Noble. 2005. A 2003 pine marten record for Hampshire. HMG News: Newsletter of the Hampshire Mammal Group 6: p. 9.

Bright, P. W., and S. Harris. 1994. Reintroduction of the pine marten: feasibility study. English Nature Contract Report F72-11-10. University of Bristol, Bristol, UK.

Corbet, G.B., and S. Harris, editors. 1991. The Handbook of British mammals. Blackwell. Oxford, UK.

Crawley, D., and J. Birks. 2004. Pine marten confirmed in Staffordshire! Mammal News 138: 13.

Davison, A., J. D. S. Birks, R. C. Brookes, A. C. Braithwaite, and J. E. Messenger. 2002. On the origin of faeces: morphological versus molecular methods for surveying rare carnivores from their scats. Journal of Zoology (London) 257: 141-143.

Easterbee, N., L.V. Hepburn, and D.J. Jefferies. 1991. Survey of the status and distribution of the wildcat in Scotland, 1983-1987. Nature Conservancy Council for Scotland Edinburgh, UK.

Fletcher, T. 2004. On the trail of the lonesome pine marten. Dalesman, December 2004: 62-64.

Forrest, G., M. Needham and J. Birks. 2002. A 2002 pine marten record for Worcestershire? Worcestershire Record 13: 23-24.

Forsyth, R. 1987. PC / Beagle – User Guide. Warm Boot Ltd., Nottingham, UK.

Gese, E. M. 2001. Monitoring of terrestrial carnivore populations. Pages 372-396 *in* J. L. Gittleman, S. M. Funk, D. W. Macdonald, and R. K. Wayne, editors. Carnivore conservation. Cambridge University Press, Cambridge, UK.

Hawkins, J. 2000. Search for shy pine dweller. The Westmorland Gazette, 4th February 2000.

Howes, C.A. 1984. Changes in the status of some Yorkshire Mammals 1600-1980. M. Phil. thesis, University of Bradford, Bradford, UK.

Jefferies, D. J., and C. H. Critchley. 1994. A new pine marten *Martes martes* (L.) record for the North Yorkshire Moors: skull dimensions and confirmation of species. Naturalist 119: 145-150.

Kyle, C. J., A. Davison, and C. Strobeck. 2003. Genetic structure of European pine martens (*Martes martes*), and evidence for introgression with *M. americana* in England. Conservation Genetics 4: 179-188.

Macdonald, D. W., G. Mace, and S. Rushton. 1998. Proposals for future monitoring of British mammals. Department of the Environment, Transport and the Regions, London, UK.

Manley, B. F. J. 1991. Randomisation and Monte Carlo methods in biology. Chapman and Hall, London, UK.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209–20.

Matos, H., and M. Santos-Reis. 2006. Distribution and status of the Pine marten *Martes martes* in Portugal. Pages 47-61 *in* M. Santos-Reis, J. D. S. Birks, E. C. O´Doherty and G. Proulx, editors. *Martes* in carnivore communities. Alpha Wildlife Publications, Alberta, Canada.

McDonald, R., P.W. Bright, and S. Harris. 1994. Baseline survey of pine martens in Wales. Unpublished Report (FC 73-01-91) by Bristol University for the Countryside Council for Wales, Bristol, UK.

Messenger, J. E., J. D. S. Birks, and D. J. Jefferies. 1997. What is the status of the pine marten in England and Wales? British Wildlife 8**:** 273-279.

Messenger, J. E. and J. D. S. Birks. 2000. Monitoring the very rare: pine marten populations in England and Wales. Pages 217-230 *in* H. I. Griffiths, editor. Mustelids in a modern world. Management and conservation aspects of small carnivore: human interactions. Backhuys, Leiden, Netherlands.

Morgan, I. K. 1992. Interim notes on the status of the pine marten in south-west and mid-Wales. Llanelli Naturalists Newsletter, Winter 1992-93: 11-22.

Morton, A. 1998. Dmap for Windows, Version 6.5c. Berkshire, UK

Office for National Statistics. 2003. Census 2001: National Report for England and Wales. The Stationery Office. London, UK.

Palma, L., P. Beja, and M. Rodrigues. 1999. The use of sighting data to analyse Iberian lynx habitat and distribution. Journal of Applied Ecology 36: 812-824.

Royal Society for the Protection of Birds. 2005. The Big Garden Birdwatch. http://www.rspb.org.uk/birdwatch/index.asp

Strachan, R., D. J. Jefferies, and P. R. F. Chanin. 1996. Pine marten survey of England and Wales 1987-1988. Joint Nature Conservation Committee, Peterborough, UK.

United Nations. 2002. Demographic yearbook 2000. United Nations, New York, USA.

Velander, K. A. 1983. Pine marten survey of Scotland, England and Wales 1980-1982. The Vincent Wildlife Trust, London, UK.

Zalewski, A. 2000. Factors affecting the duration of activity by pine martens (*Martes martes*) in the Białowieża National Park, Poland. Journal of Zoology (London) 251: 439-447.

Appendix A. Details of questionnaire and weightings assigned by the genetic algorithm.

Question	*Answer*	*Weighting*	*Question*	*Answer*	*Weighting*
Occupation or skill level of recorder	Amateur naturalist	9	Time of day	Day	4
	Other skilled observer	3		Night	2
	Unskilled observer	0		Twilight	4
	Countryside professional	9	Weather conditions	Clear	2
	Semi-skilled observer	5		Fog/mist	6
Record type	Carcass	2		Rain/snow	6
	Sighting	3	Reason for thinking the sighting was a pine marten	Book/specimen/ article	4
	Trapped	4		Poster	0
Duration of sighting	Fleeting glimpse	3		Television	5
	2 - 4 seconds	2		Previous experience	10
	5 - 9 seconds	3		Process of exclusion	1
	10 - 29 seconds	4		Expert advice	3
	30 seconds - 1 minute	6	Description of size	very small	1
	>1 minute	5		small	0
Distance to sighting	< 5 m	2		OK	8
	5 - 9m	4		large	3
	10 - 24m	2		very large	1
	25 - 49m	5	Features that were described correctly	Head	5
	50 - 100m	2		Body	1
	> 100m	6		Tail	4
Activity during sighting	Climbing	5		Legs	9
	Cycling	7		Colour	5
	Driving	4		Bib	8
	Horse riding	4		Agility	8
	Sitting/standing	4		Gait	7
	Walking	4	Confident the sighting was not a …	Polecat / ferret	6
	Walking dog	4		Stoat / weasel	1
	Working	7		Fox	2
Sighting	Off ground	10		Mink	7