AQUATIC CONSERVATION: MARINE AND FRESHWATER ECOSYSTEMS
Aquatic Conserv: Mar. Freshw. Ecosyst. **8**: 515–527 (1998)

Principles of using River Habitat Survey to predict the distribution of aquatic species: an example applied to the native white-clawed crayfish Austropotamobius pallipes

MARC NAURA[a,]* and MAGGIE ROBINSON[b]
[a] *The Environment Agency, Richard Fairclough House, Knutsford Road, Warrington WA4 1HG, UK*
[b] *English Nature, Juniper House, Murley Moss, Oxenholme Road, Kendal OA9 7RL, UK*

ABSTRACT

1. The Environment Agency has compiled a comprehensive database of the physical structure of British rivers using the River Habitat Survey (RHS) methodology.
2. A series of models was developed to ascertain whether RHS could be used to predict the environmental requirements of riverine species. *Austropotamobius pallipes* (Lereboullet) was chosen for the model, as it is an endangered species throughout Europe. In Britain, many populations have been eradicated or severely depleted by the crayfish plague (an introduced fungal disease), pollution, competition from non-native species and loss of habitat.
3. A series of models was developed to predict crayfish occurrence according to habitat features recorded on the Environment Agency RHS database, and also from map-derived information. Sites surveyed for crayfish were matched to existing RHS sites. Incidental sightings during Agency routine biological or fisheries monitoring were also included.
4. Logistic regressions on RHS variables yielded two models which showed suitable fit of the data, and high success rates when tested on an independent sample of sites. The first model mainly comprised variables on channel vegetation and on bank and channel structure. Further analyses without channel vegetation yielded a second model with variables more closely related to crayfish habitat. The variables identified as having a positive impact on crayfish presence were overhanging boughs, the presence of boulders, the amount of tree shading and the number of riffles. The variables with a negative impact were exposed tree roots, eroding cliffs, the amount of poached or reinforced banks, gravel/pebble/sand banks and cobble substrate.
5. A discriminant analysis on crayfish presence/absence according to altitude, slope and distance from source transformed for normality, showed that crayfish could be predicted from map-derived information. Results were discussed and consideration was given to the wider applicability and potential uses of the models.

KEY WORDS: River Habitat Survey; distribution; prediction; native crayfish

* Correspondence to: The Environment Agency, Richard Fairclough House, Knutsford Road, Warrington WA4 1HG, UK.

Contract grant sponsor: Institute of Terrestrial Ecology (Monkswood)
Contract grant sponsor: Environment Agency

CCC 1052–7613/98/040515–13$17.50

Received 17 March 1997
Accepted 20 October 1997

INTRODUCTION

The white-clawed crayfish *Austropotamobius pallipes* (Lereboullet) is the only indigenous species of freshwater crayfish to be found in Britain and Ireland (Holdich and Reeve, 1991; Holdich *et al*., 1995; Holdich, 1996). Its distribution is limited primarily by calcium concentration. *A. pallipes* requires 2.8 mg L^{-1} of calcium for calcification of the exoskeleton (Chaisemartin, 1967), but in Britain it rarely occurs in water with a calcium concentration below 5 mg L^{-1} (Jay and Holdich, 1981). Its lotic distribution is therefore limited to calcareous catchments, where erosion provides receiving water-bodies with a high concentration of inorganic ions.

Since the 1970s the species has declined alarmingly in British rivers, due to pollution, drought, engineering works, competition from non-native species, and most critically, from the effects of crayfish plague (Holdich *et al*., 1995). Crayfish plague is a fungal disease caused by the oomycete *Aphanomyces astaci* Schikora (Alderman and Polglase, 1988). The usual vector of transmission is the American signal crayfish, *Pacifastacus leniusculus* Dana, a North American import, introduced in the mid-1970s (Holdich *et al*., 1995). Since the detection of crayfish plague in 1983, the spread of the disease has been rapid and has resulted in annihilation of many native populations. *P. leniusculus* is also a superior competitor for resources and exhibits predation on *A. pallipes* (Alderman, 1993), resulting in yet further decline of the native crayfish population. Mixed populations which have developed have soon become dominated by *P. leniusculus* (Holdich and Domaniewski, 1995).

Three other species of exotic crayfish have also become established in British waters: *Astacus leptodactylus* Eschscholtz or narrow-clawed crayfish; *Procambarus clarkii* Girard or red swamp crayfish; and *Astacus astacus* Linnaeus or noble crayfish (Holdich *et al*., 1995). All three species might also pose problems for survival of *A. pallipes* in mixed populations.

Crayfish plague was imported into Europe in the mid-19th century probably on infected American crayfish, and the disease has now spread throughout much of Europe. Because of its threatened status, *A. pallipes* is listed as an endangered species by the IUCN (Groombridge, 1993). The species is listed in Appendix III of the Bern Convention and in Annexes IIa and Va of the 'Directive 92/43/EEC on the conservation of natural habitats and of wild fauna and flora' (the 'Habitats Directive'). *A. pallipes* is also protected under Schedule 5 of the Wildlife and Countryside Act 1981, with respect to taking from the wild and sale. Holdich (1996) summarizes the biology, distribution and legislation relating to *A. pallipes* in Europe.

In response to the United Nations Convention on Biological Diversity (Palmer, 1994), English Nature proposes to implement the Joint Nature Conservation Council's action plan for the conservation of *A. pallipes* by supporting re-introduction programmes for native crayfish at suitable and strategic sites. For crayfish to be successfully reintroduced, there is a need to establish which environmental parameters are necessary to support a population. This would help in developing effective management plans for Special Areas of Conservation (SACs) and Sites of Special Scientific Interest (SSSIs) containing native crayfish, and could also be used as a management tool for sympathetic river engineering works.

River Habitat Survey (RHS) is a method for describing the habitat diversity of British rivers, and comprises a standard survey method and a database of more than 5000 sites (Raven *et al*., 1997).

This study explores the scope for using RHS information for predicting the occurrence of *A. pallipes*, using data from the literature and from existing databases.

METHODS

River Habitat Survey

The RHS field method is designed to yield reliable information on the physical structure of a 500 m stretch of river in a format suitable for statistical analysis (Fox *et al*., 1998). The survey is organized in two major sections: 'spot-checks' and 'sweep-up'. The spot-checks are a series of ten 1 m wide transects across the channel at 50 m intervals, where bank and channel physical structure, as well as man-made modifications, land uses and vegetation structure, are recorded in a replicable manner. What is not recorded in the spot-checks is included in the 'sweep-up' section, with other habitat components such as trees, flow features, and bank structure. In addition, background map-based information on altitude, slope, distance from source, height of source, solid and drift geology, flow category and water quality class, is also collected.

Sites in every 10 km square of England, Wales, Scotland and Northern Ireland have been sampled. Altogether, more than 5000 sites were visited by trained surveyors in spring and summer 1994, 1995, and 1996. The data were then fed into a database, which constitutes a unique representative inventory of river habitat features.

Crayfish sampling

National Grid References (NGR) of crayfish sightings were obtained from the crayfish database held by the Institute of Terrestrial Ecology in Monkswood, from incidental sightings during Agency biological or fisheries surveys and from records within existing literature (Holdich *et al*., 1995; Rogers and Holdich, 1995a,b; Hutchings, 1996; Summers, 1996). Only sightings occurring after 1990 were considered, so as to minimize the chances that modifications had taken place since the sighting occurred.

Crayfish were recorded as being present on a site if they had been observed, incidentally or not, regardless of the investigation method. On the contrary, crayfish were recorded as absent, only if searching had been actively undertaken using trapping or stone turning. Only sites with a calcium concentration suitable for crayfish were retained. Calcium concentration was derived from the Environment Agency water quality data. When crayfish was recorded as absent, sites on catchments affected by the crayfish plague were discarded.

Using this selection protocol, more than 700 sites were assembled. After matching with RHS sites, 140 sites remained, among them 122 with crayfish, and 18 without. In order to increase the size of the sample without crayfish, a further 10 sites were sampled during the summer of 1996 on the Eden catchment. Four of the sites had already been surveyed for RHS during the summer of 1996, and the remaining six were surveyed before searching for crayfish. Sampling was carried out using stone turning for a period of 60 min over the whole length of the stretch. A distribution map of all 150 sites is shown in Figure 1.

Matching of RHS and crayfish sites

Using 1:50000 Ordnance Survey maps, crayfish sites were paired with existing RHS sites from the national database, and with additional sites collected for various purposes during Agency routine works. Only sites falling on the same watercourse within 5 km of each other were retained for analysis. Crayfish sites and RHS sites were not paired if a tributary occurred between them, or if one of them had already been paired with another site. Where two or more crayfish sightings were within 5 km of an RHS site, only the sighting closest in distance to the RHS site was included in the matching. A sample of 30 sites (20% of the dataset) was randomly selected and set aside to test the model, thus leaving 120 sites for analysis (99 sites with crayfish, and 21 sites without crayfish). The average distance between RHS and crayfish sites was 1.7 km, 40% of the sites fell within 1 km of each other, 67% within 2 km, and 93% within 3.5 km.

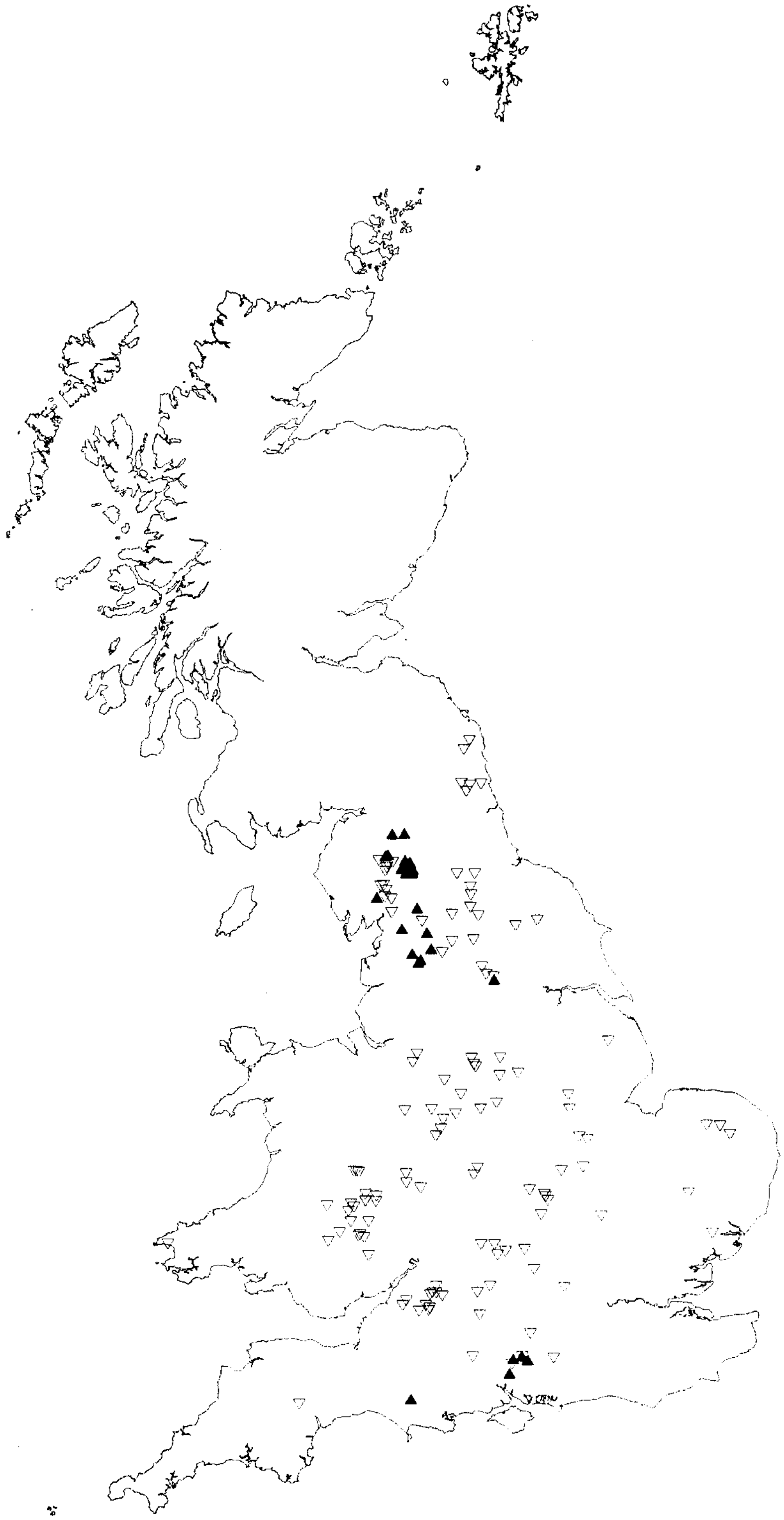

Figure 1. Map showing the distribution of the sites included in the analyses (sites with crayfish, ▽; sites without crayfish, ▲).

Table 1. Rules for the conversion of RHS variables

Conversion rules for submerged vegetation			Conversion rules for modified bank profiles		
Fine leaved	Broad leaved	Total	Left banks	Right banks	Code
Absent	Absent	Absent	Absent	Absent	0
Absent	Present	Present	Absent	Present	1
Present	Present	Present	Present	Present	2
Absent	Extensive	Extensive	Absent	Extensive	3
Present	Extensive	Extensive	Present	Extensive	4
Extensive	Extensive	Extensive	Extensive	Extensive	5

Coding of the variables

Altogether, more than 100 environmental variables were included in the analyses, both from the RHS spot-check and the sweep-up sections.

For each physical attribute, a series of variables representing all the possible options was created. For example, 'channel substrate' was split into eight distinct variables coding for bedrock, boulder, cobble, gravel/pebble, sand, silt, clay and peat densities over the 10 spot-checks.

Spot-check channel vegetation variables were coded differently, as they were not recorded in presence/absence, but on a 'none'/'present'/'extensive' scale. Each vegetation variable was split into three variables corresponding to each point of the scale. In 1994, there was no distinction between submerged plants (fine leaved and broad leaved). In order to be able to include submerged vegetation in the analyses, submerged fine and broad leaved were pooled into one vegetation type, using the rules shown in Table 1.

Sweep-up variables were coded according to their nature. Single variables like 'exposed tree roots' or 'waterfall' were coded as 0 (absent), 1 (present) or 2 (extensive). Variables recorded for both banks, e.g. 'bank profiles' or 'extent of trees', were coded in ways that eased statistical analysis and interpretation. Tree cover was coded on a five point scale, showing increasing tree density. Scores were added for both banks, thus yielding an estimate of the tree cover ranging from 0 (no trees on either bank), to 10 (continuous tree cover on both banks). Natural bank profiles were split into variables in the same manner as spot-check vegetation variables. Modified bank profiles could not be transformed in the same way, because of the rarity of these structures. Instead, information on both banks was amalgamated into one variable following the conversion rules shown in Table 1. Although this transformation minimizes information loss, it may create artificial differences between combinations of cover levels. For example, it assumes the presence of poached banks on both sides to be more important than if it is observed only on one side, which may be wrong in terms of the total length of bank which is actually poached.

Flow type data were not included in the analyses because flow type definitions changed between 1994 and 1995. This made the data impossible to use, unless 1994 data were discarded.

Statistical analyses

On-site prediction

Models predicting crayfish occurrence according to RHS features were derived using logistic regression analysis. All regression analyses were carried out using SPIDA statistical software (version 6), and MINITAB version 11.

Logistic regression is a statistical technique that is used to study the relationships between a binary variable (coded as 0 or 1) and a set of independent variables, which can be continuous, ordinal, nominal, or even binary (Hosmer and Lemeshow, 1989). The logistic regression model takes the form

$$P(Y=1) = 1/(1+\exp(-\beta_0 - \beta_1 X_1 - \beta_2 X_2 - \ \ldots\ - \beta_n X_n))$$

where $P(Y=1)$ is the probability of crayfish being present, and β_{i} are the regression coefficients associated with the independent variables X_{i} (Gebski *et al*., 1992).

Due to the size of the dataset and the computing time involved, the most significant variables related to crayfish occurrence were selected. Sweep-up variables were analysed using chi-square tests for association, whereas spot-checks and continuous variables were selected using non-parametric Mann-Whitney U tests. The level of significance for inclusion in the final analyses was set at $p=0.2$. A list of selected variables is shown in Table 2.

Due to its small size, the set of sites without crayfish could not cover a wide range of structural diversity. As a consequence, some variables showed high levels of association with crayfish, solely due to the poor (or high) representation of the feature in the 'no crayfish' sample. For example, clay is the least represented channel substrate (after peat) in England and Wales. It is only present at less than 10% of the RHS reference sites ($n=4569$). Clay was present at 10% of the sites where crayfish was found, and totally absent from the sites without crayfish. As a consequence, clay was wrongly identified as being highly significant for crayfish presence. Such spurious variables were removed from the analyses after careful investigation of the raw data.

Table 2. List of RHS variables most associated with crayfish, and levels of significance (+/− $p<0.05$, ++/−− $p<0.01$, +++/−−− $p<0.001$). Selection was made using chi-square tests for association and Mann–Whitney U tests

Bank substrate	Channel vegetation (spot-checks)
Boulder/cobble+	No vegetation++
Gravel/pebble/sand−−	Liverworts present
Bank features	Reeds present
Eroding cliff−	Floating leaved present+
Vegetated point bar	Amphibious present+++
Unvegetated side bar	Filamentous algae present++
Bank modifications	Submerged present
Reinforcement (spot-checks)−	Liverworts extensive−−
Reinforced toe−−	Filamentous algae extensive
Poached banks (sweep-up)	
Bank profile	Channel vegetation (sweep-up)
Vertical bank extensive−	Liverworts−−−
Steep bank extensive++	Filamentous algae
	Floating leaved+
Channel substrate	Amphibious+
Bedrock	
Cobbles−−−	Trees and associated features
Channel features	Extent of trees
Exposed boulders (spot-checks)	Overhanging boughs
Exposed boulders (sweep-up)	Exposed tree roots
Channel features	Underwater tree roots
Riffles (number)	Fallen trees++
Channel modifications	Tree debris+
Weirs (number)	Tree shading+
Channel dimension	
Channel wetted width	

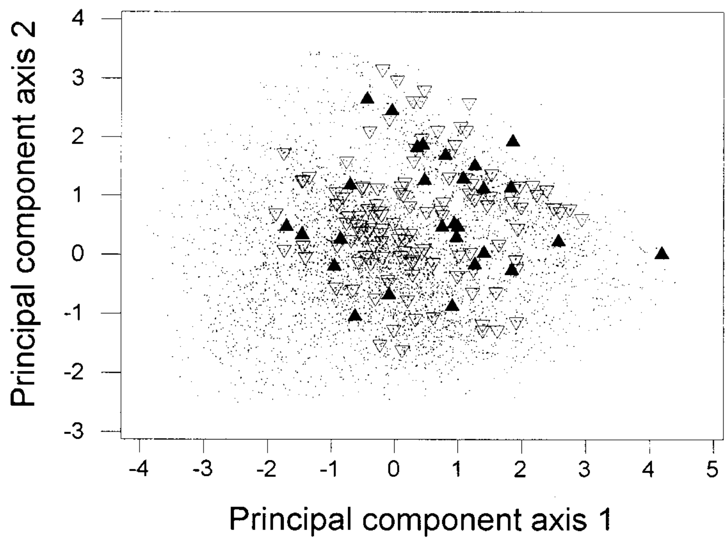

Figure 2. Comparison of the crayfish sample distribution ($n = 150$; sites with crayfish, ▽; sites without crayfish, ▲) with the watercourses population distribution (●, $n = 4569$).

A best subset selection procedure was then carried out, and models were chosen according to their ecological significance and their quality of fit. Goodness-of-fit was tested using Pearson, Deviance and Hosmer–Lemeshow tests. The overall significance of the coefficients was tested with a likelihood ratio G-test (Hosmer and Lemeshow, 1989), while individual score tests were performed on each coefficient.

Models were also tested by estimating their success at predicting crayfish occurrence. Predicted probabilities were computed for each site. Crayfish was recorded as present if the predicted probability exceeded 50% ($P(Y = 1) > 0.5$). The success rate (SR) was then defined as the proportion of correct predictions over the sample. If the model is inefficient at predicting crayfish presence/absence, it will have a success rate of 50%, the equivalent of tossing a coin every time a new site is assessed. Thus, the significance of the model can be assessed by testing whether the success rate observed comes from a (binomial) population with an average success rate of 50% ($p = 0.5$).

Model success at predicting crayfish presence and absence was tested separately, as a model which was effective at predicting both the presence and the absence of crayfish was required.

Map-based predictions

Predictive models were derived using solely map-derived information. Altitude, slope, distance from source and height of source were used in a multiple discriminant analysis to predict the occurrence of crayfish. The variables were first transformed for normality, using Box-Cox power transformation procedure on MINITAB. The model yielding the best fit was chosen and its significance was tested as previously.

To assess how applicable the model was to other rivers in England and Wales, it was necessary to know how well the sample represented the overall river population variability. The 150 sites on which the model was based were compared to 4569 sites from the RHS network. The 4569 sites were plotted on a map produced with principal component analysis, which accounts, in two principal components, for 90% of the variability observed between the sites' altitude, slope, distance from source, and height of source. The 150 sites were then plotted on the same map and their respective coverage examined (Figure 2). The first principal component axis represents increasing altitude and slope, and the second axis increasing energy (Jeffers, 1998).

RESULTS

On-site predictions

After selection, 37 variables were retained for the analyses (Table 2). The first logistic regression model was driven from a best subset selection procedure including all variables. The model contained 10 variables, among them five related to channel vegetation (Table 3). The coefficients for each variable are shown in Table 3, along with their associated standard deviation, odds ratio, and significance. The likelihood ratio G-test, along with the univariate score tests, showed high levels of significance, whereas the goodness-of-fit tests indicated suitable fit of the data. The model success rate (SR) was estimated on the initial sample ($n = 120$) and on the test sample ($n = 30$). The efficiency of the model at predicting crayfish presence and absence separately was also tested on both samples. Altogether, success rates were high for all samples. The model explained 91% of the main sample cases ($p < 0.001$), and 93% of the test sample ($p < 0.001$). The ability of the model at predicting crayfish presence was also high (main sample, SR = 97%, $n = 99$, $p < 0.001$; test sample, SR = 96%, $n = 23$, $p < 0.001$), whereas the prediction of crayfish absence was satisfactory (main sample, SR = 62%, $n = 21$, $p = 0.09$; test sample, SR = 86%, $n = 7$, $p = 0.008$).

The odds ratio (Ψ) is a measure of association between variables: odds ratios above 1.0 indicate positive association between features and crayfish, whereas odds ratios below 1.0 indicate a negative association. The variables most associated with crayfish presence were overhanging boughs, with an odds ratio of 86.72, extensive steep banks ($\Psi = 6.77$), and the presence of amphibious vegetation ($\Psi = 6.26$). On the other hand, the variables associated with crayfish absence were the presence of liverworts ($\Psi = 0.09$), poached banks ($\Psi = 0.31$), and reinforced toe ($\Psi = 0.52$).

A second model was derived using all variables except channel vegetation variables. The results are shown in Table 4. Although the test on all coefficients was highly significant ($G = 70.854$, df = 11,

Table 3. Logistic regression model derived from a best subset selection on all RHS variables

Predictor	Coefficients	St. dev.	Z	p	Odds ratio (Ψ)
Constant	0.1218	0.9927	0.12	0.902	
Cobbles (channel substrate)	−0.3913	0.1509	−2.59	0.009	0.68
Poached banks	−1.1858	0.4123	−2.88	0.004	0.31
Reinforced toe	−0.6457	0.3524	−1.83	0.067	0.52
Steep bank extensive	1.9125	0.9182	2.08	0.037	6.77
Overhanging boughs	4.4630	1.2900	3.46	0.001	86.72
Reeds present	0.3034	0.1599	1.90	0.058	1.35
Amphibious present	1.8342	0.6826	2.69	0.007	6.26
Filamentous present	0.2244	0.1356	1.65	0.098	1.25
Submerged present	−0.3584	0.1939	−1.85	0.065	0.70
Liverworts (sweep-up)	−2.4434	0.9139	−2.67	0.008	0.09

Log-likelihood = −24.562
Test that all slopes are zero: $G = 62.170$, df = 10, p value = 0.000

Goodness-of-fit tests Method	chi-square	df	p value
Pearson	54.688	109	1.000
Deviance	49.124	109	1.000
Hosmer–Lemeshow	3.409	8	0.906

Table 4. Logistic regression model derived from a best subset selection on all RHS variables except channel vegetation

Predictor	Coefficients	St. dev.	Z	p	Odds ratio (Ψ)
Constant	2.516	1.0220	2.46	0.014	
Cobbles (channel substrate)	−0.6858	0.2992	−2.29	0.022	0.50
Exposed boulders (sweep-up)	0.7330	1.2370	0.59	0.553	2.08
Riffles (number)	0.2973	0.1980	1.50	0.133	1.35
Boulder cobble banks	1.7469	0.7317	2.39	0.017	5.74
Gravel/pebble/sand banks	−1.0400	0.5260	−1.98	0.048	0.35
Eroding cliffs	−1.3469	0.8018	−1.68	0.093	0.26
Poached banks	−1.1894	0.5398	−2.20	0.028	0.30
Reinforced banks	−0.8586	0.3427	−2.51	0.012	0.42
Overhanging boughs	3.2440	1.2940	2.51	0.012	25.64
Tree shading	1.6207	0.9078	1.79	0.074	5.06
Exposed tree roots	−2.5080	1.0620	−2.36	0.018	0.08

Log-likelihood = −18.457
Test that all slopes are zero: $G = 70.854$, df = 11, p value = 0.000

Goodness-of-fit tests Method	chi-square	df	p value
Pearson	66.751	106	0.999
Deviance	36.914	106	1.000
Hosmer–Lemeshow	4.638	8	0.795

$p < 0.0001$), tests on individual coefficients showed differences. The presence of exposed boulders and the number of riffles had the lowest levels of significance (exposed boulders, $Z = 0.59$, $p = 0.55$; number of riffles, $Z = 1.50$, $p = 0.13$). Alternative models without these features were tried, but were not satisfactory. The goodness-of-fit tests, with p values ranging from 0.795 to 1, suggests that the model fits the data adequately. Success rates were again high. The model explained 95% of the main sample cases ($p < 0.001$), and 87% of the test sample ($p < 0.001$). Predictions for crayfish presence and absence were highly significant for both samples (crayfish present: main sample, SR = 99%, $n = 99$, $p < 0.001$; test sample, SR = 87%, $n = 23$, $p < 0.001$; crayfish absent: main sample, SR = 80%, $n = 21$, $p < 0.001$; test sample, SR = 86%, $n = 7$, $p = 0.008$). The most influential variables for crayfish presence were overhanging boughs ($\Psi = 25.64$), boulder/cobble as a bank substrate ($\Psi = 5.74$), and the amount of tree shading ($\Psi = 5.06$). The variables most associated with crayfish absence were exposed tree roots ($\Psi = 0.08$), eroding cliffs ($\Psi = 0.26$), poached banks ($\Psi = 0.30$), gravel/pebble/sand bank substrate ($\Psi = 0.35$), and reinforced banks ($\Psi = 0.42$).

Map-based predictions

Using Box-Cox transformation procedure, the following transformations were used for all map-derived variables: altitude transformed = altitude$^{0.337}$; slope transformed = slope$^{0.224}$; distance from source transformed = distance from source$^{0.224}$; height of source transformed = $\log_{10}$(height of source).

All the possible models were iterated and tested for goodness-of-fit. The best model was obtained for a combination of altitude, slope and distance from source (Table 5). The success rate for the main sample was 70% ($p < 0.001$) and 77% on the test sample ($p < 0.001$). The model still performed well when tested on crayfish presence and absence separately (crayfish present: main sample, SR = 71%, $n = 99$, $p < 0.001$; test sample, SR = 78%, $n = 23$, $p < 0.001$; crayfish absent: main sample, SR = 67%, $n = 21$, $p = 0.04$; test sample, SR = 71%, $n = 7$, $p = 0.06$).

DISCUSSION

The study shows the potential for using RHS data to predict the occurrence of freshwater invertebrates such as crayfish. The two models developed illustrate the various uses that can be made of all the variables recorded in the RHS form.

The first model concentrated on environmental parameters which are not known for being directly related to crayfish habitat. Half the variables represented channel vegetation, with the presence of amphibious plants, reeds and filamentous algae being seen as critical for crayfish, whereas liverworts showed a strong negative relationship. It is worth mentioning that only 'presence' of vegetation—as opposed to 'absence' or 'extensive' in RHS terms—was seen as significant. Little work has been done on aquatic vegetation at sites where crayfish are found, though they are known to eat *Fontinalis* moss if no other food sources are available (Brown and Bowler, 1977). One possibility is that channel vegetation is acting as an indicator of crayfish habitat. The particular combination of vegetation types selected by the regression analysis may in fact share some of the habitat requirements of crayfish, or simply occur in the same types of river. Nonetheless, the use of these variables in combination with information on the bank structure and modifications, and on overhanging vegetation and channel substrate, was sufficient to predict successfully the occurrence of crayfish.

The second model was derived from an analysis where all channel vegetation variables were excluded. It selected a more diverse set of parameters, comprising bank and channel attributes, as well as tree features. Three variables related to the presence of trees were highly influential on crayfish occurrence. As in the previous model, overhanging boughs presented the highest positive relationship with crayfish, closely followed by the canopy cover. Trees have been shown to be an important source of food for organisms living in the channel. Reynolds (1979) recorded the preference of *A. pallipes* for soaked leaf litter. Leaf litter also provides a suitable medium for the growth of microbes and fungi, which are another source of food for crayfish (Kaushik and Bird, 1987). Other important inputs of trees are the invertebrates falling from the overhanging leaves into the water (Mason and MacDonald, 1982). Thus, 'overhanging boughs' and 'tree shading' might in fact represent the same ecological dimension: the presence of overhanging vegetation acting as a source of food, either directly, by falling into the channel, or indirectly, by providing a support for species forming part of the crayfish diet.

The impact showed by 'exposed tree roots' is unclear. Underwater tree roots have been mentioned as a place for searching for juveniles of *A. pallipes* (Foster, 1995; Rogers and Holdich, 1995a; Smith *et al.*, 1996), but no correlation between exposed roots and crayfish has so far been shown. Exposed tree roots may in fact indicate erosion (leading to increasing suspended solids concentrations), or provide shelter for crayfish predators, such as mink and otters.

The presence of erosion seemed to be another important dimension affecting crayfish occurrence. Its influence was perceivable through four variables all showing strong negative impacts on crayfish occurrence: eroding cliffs, poached banks, gravel/pebble/sand bank substrate, and reinforced banks.

Table 5. Discriminant analysis on crayfish occurrence according to transformed altitude, slope and distance from source

Linear discriminant functions	Crayfish	
	Absent	Present
Constant	−29.067	−28.054
Altitude	5.900	4.879
Slope	2.605	4.862
Distance from source	15.087	15.657

Although erosion might not affect crayfish directly, it will increase the amount of suspended solids in the water and on the river bed, thus reducing the amount of habitat available to crayfish. Products of erosion, such as silt, have been consistently described as important detrimental influences on crayfish habitat (Hogger, 1988; Summers, 1996). However, silt as a specific variable was not chosen as a significant influence in the logistic regression. Fine substrates were indeed poorly represented in the 'no crayfish' sample, making them difficult to use in the analyses. This is probably due to the way substrates are recorded during the survey. Indeed, only the predominant substrate in terms of its coverage of the river bed is recorded in the survey form. Consequently, substrates which are never predominant on a site are not recorded.

The presence of reinforced banks implies a site where erosion is a problem. The materials used for reinforcement could also have an impact on crayfish. *A. pallipes* would make use of dry-stone walling, but other types of reinforcement, such as concrete and sheet piling, could severely deplete suitable habitats available for the species. At reinforced sites, trees are usually removed prior to the engineering work; according to the model, this would have a detrimental effect on crayfish occurrence.

The negative influence of cattle poaching on crayfish presence confirms subjective observations by the authors or other crayfish surveyors (Summers, 1996). Cattle poaching undoubtedly disturbs the sediment, increasing siltation downstream and possibly causing localized inputs of ammonia, to which *A. pallipes* is susceptible (Foster, 1995). The localized habitat degradation caused by cattle in the river means that the site is likely to be unsuitable for many macroinvertebrates, resulting in a restricted biotic assemblage (Wright *et al*., 1993).

The positive association between exposed boulders and boulder/cobble banks, and *A. pallipes* reflects the importance of availability of refuges of the correct size. This was highlighted by the study of Foster (1993) on the relationship between refuge size and body size in *A. pallipes*. Foster found a minimum stone size (area) for use as a refuge by crayfish, and a significant relationship between the area of a stone being used as a refuge and the carapace length of the crayfish underneath the stone. Foster hypothesized that the crayfish size:stone area ratio may be due to the minimum area required for concealment from predators or from light. This is also likely to be the reason why cobble substrate is identified as negatively associated with *A. pallipes* occurrence. Although crayfish are found in areas of cobbled substrate, this is usually where boulders are also present. When cobbles are recorded as the major substrate on the site, they are often associated with smaller substrates, and it is then likely that the stones available are not large enough to provide adequate refuges.

The number of riffles was defined as a positive variable for crayfish. Riffles are often the preferred environment for other macroinvertebrates, and as an omnivorous feeder, *A. pallipes* is likely to exploit this habitat for feeding.

Crayfish occurrence could also be predicted using map-based information. A discriminant analysis identified altitude, slope, and distance from source as the best variables to predict crayfish occurrence. The model performed well on both samples, and showed some potential for using it as a first method for identifying sites where crayfish are likely to be found using maps or a GIS. The applicability of the model to other rivers in England and Wales was verified by comparing our sample to the overall population of watercourses. It can be seen in Figure 2 that the sample covered the overall variability well, apart from the lower left hand corner, which corresponds to low altitude, low slope, low energy rivers. Large lowland rivers were indeed poorly represented in the sample. This sampling bias is probably due to their size, which makes sampling difficult, and the fact that they tend to be more intensively managed, as they often flow through urban areas.

Sampling is one aspect on which the various models produced could be improved. In particular, more sites are needed presenting a wider range of features, where crayfish were looked for and not found. A more random sampling strategy would also help in describing crayfish habitat more widely. Indeed, most of the sites were chosen by experienced surveyors, who targeted their survey towards sites where crayfish were most likely to be found. Therefore, the sample mainly represents high quality habitats, and may miss out on poorer (but still suitable) sites, or even on unconventional habitats. Also, all sites selected were of high water quality, which might reflect sampling bias towards good quality rivers. The coincidence between RHS sites and crayfish sites could also be improved by surveying more RHS sites, or carrying out crayfish survey on existing RHS sites. However, it would be rather labour intensive, and Sansbury (1994) showed in a study on the River Wyre, that a 500 m RHS stretch represented a much longer section of river.

River Habitat Survey seems to be a practical way of generating hypotheses to determine which factors are influencing crayfish presence, and could be applied to other species found in riverine environments. Its main advantages on other methods is that it is a standard method whose replicability has been tested, with a growing database of sites, and a guidance manual where features are defined and illustrated. Also, the assessment of features over a series of spot-checks minimizes variability between surveyors, a factor which is often not taken into consideration in survey methodologies (Fox *et al*., 1998).

ACKNOWLEDGEMENTS

We wish to thank Dr David Holdich, the Institute of Terrestrial Ecology (Monkswood), and members of staff of the Environment Agency for their help in collating the data. Thanks are also due to Professor John Jeffers, who gave essential advice on statistical analyses, and to Mike Furse, Fiona Duke and Paul Green for valuable comments on the manuscript and the analyses.

REFERENCES

Alderman, D.J. and Polglase, J.L. 1988. 'Pathogens, parasites and commensals', in Holdich, D.M. and Lowery, R.S. (Eds), *Freshwater Crayfish: Biology, Management and Exploitation*, Chapman and Hall, London, 167–212.

Alderman, D.J. 1993. 'Crayfish plague in Britain, the first twelve years', *Freshwater Crayfish*, **9**, 266–272.

Brown, D.J. and Bowler, K. 1977. 'Some aspects of the growth in the British freshwater crayfish *Austropotamobius pallipes* (Lereboullet)', *Freshwater Crayfish*, **4**, 35–42.

Chaisemartin, C. 1967. 'Contribution à l'étude de l'économie calcique chez les astacidae', Thèse de doctoral de Sciences Naturelles, Archives Originales du Centre de Documentation du CNRS, Poitiers.

Foster, J. 1993. 'The relationship between refuge size and body size in the crayfish *Austropotamobius pallipes*', *Freshwater Crayfish*, **9**, 345–349.

Foster, J. 1995. 'Factors influencing the distribution and abundance of crayfish *Austropotamobius pallipes* in Wales and the Marches, UK', *Freshwater Crayfish*, **8**, 78–93.

Fox, P.J.A., Naura, M. and Scarlett, P. 1998. 'An account of the derivation and testing of a standard field method, River Habitat Survey', *Aquatic Conservation: Marine and Freshwater Ecosystems*, **8**, 455–475.

Gebski, V., Leung, O., McNeil, D. and Lunn, D. 1992. *Statistical Package for Interactive Data Analysis (SPIDA), User Manual, version 6*, Statistical Computing Lab., Macquarie University, Australia.

Groombridge, B. (Ed.) 1993. *1994 IUCN Red List of Threatened Animals*, IUCN, Gland.

Hogger, J.B. 1988. 'Ecology, population biology and behaviour', in Holdich, D.M. and Lowery, R.S. (Eds), *Freshwater Crayfish: Biology. Management and Exploitation*, Chapman and Hall, London, 114–144.

Holdich, D.M. and Reeve, I.D. 1991. 'Distribution of freshwater crayfish in the British Isles, with particular reference to crayfish plague, alien introductions and pollution', *Aquatic Conservation: Marine and Freshwater Ecosystems*, **1**, 139–158.

Holdich, D.M. and Domaniewski, J.C.J. 1995. 'Studies on a mixed population of the crayfish *Austropotamobius pallipes* and *Pacifastacus leniusculus* in England', *Freshwater Crayfish*, **10**, 37–45.

Holdich, D.M., Rogers, W.D. and Reader, J.P. 1995. *Crayfish Conservation*, R&D Project 379, National Rivers Authority, Bristol.

Holdich, D.M. 1996. '*Austropotamobius pallipes*', in Van Helddingen, P.J., Willemse, I. and Speight, M.C.D. (Eds), *Background Information on the Invertebrates of the Habitats Directive and the Bern Convention. Part 1—Crustacea. Coleoptera and Lepidoptera. Nature and Environment No 79*, Council of Europe Publishing, Strasbourg, 2–8.

Hosmer, D.W. and Lemeshow, S. 1989. *Applied Logistic Regression*, John Wiley, New York.

Hutchings, A.R. 1996. 'Native crayfish project, River Itchen, Hampshire: interim report on the project', Sparshot College, Hampshire.

Jay, D. and Holdich, D.M. 1981. 'The distribution of *Austropotamobius pallipes* (Lereboullet) in British waters', *Freshwater Biology*, **2**, 121–129.

Jeffers, J. 1998. 'Characterization of river habitats and prediction of habitat features using ordination techniques', *Aquatic Conservation: Marine and Freshwater Ecosystems*, **8**, 529–540.

Kaushik, N.K. and Bird, G.A. 1987. 'Processing of maple, leaf, grass and fern packs and their colonisation by invertebrates in a stream', *Journal of Freshwater Ecology*, **4**, 177–189.

Mason, J.C. and MacDonald, S.M. 1982. 'The input of terrestrial invertebrates from tree canopies to a stream', *Freshwater Biology*, **12**, 305–331.

Palmer, M. 1994. 'Action plan for the conservation of the native crayfish *Austropotamobius pallipes* in the United Kingdom', Report **193**, Joint Nature Conservation Committee, Peterborough.

Raven, P.J., Fox, P., Holmes, N.T.H. and Dawson, F.H. 1997. 'River Habitat Survey: a new system for classifying rivers according to their habitat quality', in Boon, P.J and Howell, D.L. (Eds), *Freshwater Quality: Defining the Indefinable*?, The Stationery Office, Edinburgh, 215–234.

Reynolds, J.D. 1979. 'Ecology of *Austropotamobius pallipes* in Ireland', *Freshwater Crayfish*, **4**, 215–219.

Rogers, W.D. and Holdich, D.M. 1995a. 'Survey of white-clawed crayfish distribution in the tributaries of the Rivers Usk and Wye', Contract Science Report **137**, Countryside Council for Wales, Bangor.

Rogers, W.D. and Holdich, D.M. 1995b. *Crayfish Survey of Selected Rivers of the Kent and Eden Catchments*, English Nature, Peterborough.

Sansbury, J. 1994. 'Survey of the River Wyre Catchment Lancashire testing the River Habitat Survey Methodology', Unpublished MSc dissertation, Lancaster University Institute of Environmental and Biological Sciences, Lancaster.

Smith, G.R.T., Learner, M.A., Slater, F.M. and Foster, J. 1996. 'Habitat features important for the conservation of the native crayfish *Austropotamobius pallipes* in Britain', *Biological Conservation*, **75**, 239–246.

Summers, D.W. 1996. *A Preliminary Investigation into the Distribution and Habitat Use of White-Clawed Crayfish* (*Austropotamobius pallipes*) *in the River Piddle*, The Game Conservancy Trust, Hampshire.

Wright, J.F., Furse, M.T. and Armitage, P.D. 1993. 'RIVPACS—a technique for evaluating the biological quality of rivers in the UK', *European Water Pollution Control*, **3**, 15–26.

