**Modelling Capture Probability from Large Electrofishing Datasets.**

Colin P. Millar, Robert J. Fryer, Karen J. Millidine and Iain A. Malcolm.

# Abstract

* Fisheries management and conservation decisions are often based on estimates of juvenile salmonid abundance obtained from electrofishing data.
* Accurate and un-biased estimates of abundance require estimates of capture probability.
* Capture probability and abundance can both vary with habitat and over time.
* Capture probability can be estimated using multi-pass electrofishing data.
* Hierarchical Bayesian models have been developed to model capture probability (and jointly fish density) but these approaches do not scale well to larger datasets. Computing power (big n) time to fit therefore model selection and exploration is prohibitively time consuming [Have they been used in cases of quantitative fishing only or in cases of mixed 1 and 3 passes?].
* Consequently, the use of these models appears to be restricted to smaller case studies and recent development in capture probability modelling do not seem to have been taken up by the applied fisheries community. Also smaller case studies can be easier to model as there is consistent uniform data.
* In larger scale models (WFD, DFO), however, the most common approaches are to either assume constant capture probability or estimate it on a sample by sample basis using a Zippin estimator.
* This paper presents an approach for estimating capture probability based on classical conditional likelihood methods which allow capture probability to be modelled in terms of linear and non-linear (non-parametric) relationships with covariates that affect capture probability
* Capture probability models were fitted to multi-pass electrofishing data for Scotland over the years 1996 to 2014. GIS covariates were included as proxies for habitat thereby removing the requirement for site specific habitat data. I Information on data source (organisation) allowed for variability in equipment, procedures and personnel.
* Sampling effects were the most important, followed by fish life stage, species and sampling date. Habitat characteristics such as width and gradient affected capture probability to a lesser extent.
* The results emphasise the risks of spatial bias in density estimates with consequences for management decision making [when inappropriate assumptions are made for capture probability?].
* Furthermore, we show modelling capture probability improves density estimates over two commonly used approaches in applied management: constant capture probability and site-wise estimates of capture probability.

Keywords: Atlantic Salmon, Brown Trout, Lifestage, [modelling words], Capture Probability, Electrofishing.

# 1. Introduction

Fish abundance and distribution are important indicators of population heath, and estimating density from electrofishing data is a common problem in fishery management and conservation. Accurate estimates of abundance from electrofishing data require robust estimates of capture probability, however, capture probability can be difficult to estimate on a site by site basis when abundance is low (ref). Capture probability varies with habitat characteristics (Bayley and Dowling, 1990) and sampling procedure (Deuwalter and Fischer, 2007), which are also known to affect abundance (ref), hence ignoring differences in capture probability can produce biased estimates of abundance (Price and Petterson, 2010). Thus, obtaining accurate estimates of electrofishing capture probability and how it varies is an important issue for fishery management and conservation, particularly so when dealing with stocks at low abundance.

Estimating and modelling capture probability for single sites has a long history (Seber, 1982; Huggins and Yip, 1997; Huggins and Hwang, 2011). Recently models extending to multiple sites have been developed (Wyatt, 2004, Rivot et al., 2008; and Conroy et al., 2008) which jointly model capture probability and abundance using hierarchical Bayesian models (HBMs). Although HBMs provide a powerful tool for ecological modelling (Cressie et al, 2009), they are computationally intensive, and hence do so at the expense of model fitting and setup time. So, *despite such advances there is still a place for computationally simpler approaches which can be used routinely, especially when a range of candidate models are under consideration (Diggle and Ribeiro, 2007)* and large dataset are under consideration. In this paper a conditional likelihood approach (Huggins and Hwang, 2011) is developed that is a simplification over HBMs but provides savings in fitting time and allows for a wide range of models to be applied relatively easily.

To illustrate the benefits of a conditional likelihood approach an analysis of capture probability is presented for a large scale national dataset covering 2749 discrete sites and 6049 site visits in Scotland over the years 1996 to 2013. Covariates that require minimal site specific data collection are used to allow the incorporation of data from a range of sources with different sampling and recording procedures. Among the factors known to affect capture probability are velocity (Bayley and Dowling, 1990; Price and Peterson 2010), cross sectional area (Price and Peterson, 2010), fish length (Price and Peterson, 2010; etc.), wood density, site width (Hedger et al, 2013), total fish captured (Hedger et al, 2013; Pregler 2015), temperature, conductivity, undercut bank (Pregler 2015, Rodtka, 2015). Some of these features can be derived directly from a GIS (e.g., width), while GIS derived gradient, for example, can be considered to approximate velocity; other examples are landuse for wood density, upstream catchment for discharge.

The organisation of the paper is as follows. A modelling framework allowing capture probability to vary from site to site is developed, where covariates enter the model through a linear predictor. The application to a large scale data set follows, in which a best approximating model for capture probability for Scotland is selected. Finally, because it is still common in applied management to either ignore the effects of variable capture probability, or to estimate capture probability separately for each sample, simple density estimates from these approaches applied to the Scottish data are compared to the modelled capture probabilities, demonstrating the potential impacts for bias and/or precision when using un-modelled capture probability.

* Capture probability estimation requires multipass fishing.
* The argument is simply about practicality of local scale data collection for large datasets and the lack of common local standards for Scotland (although have them for other countries e.g. Habscore) and the qualitative nature of local assessments.

# 2. Method

In this section a model for capture probability is developed for sites, assuming that capture probability pj varies from site to site. It is also shown that the techniques of generalised linear models can be used to test for outliers and assess the fit of the model.

A model for capture efficiency *pj* at *j* = 1, .., *n* samples, can be derived from the standard model for electrofishing dating back to Moran (1956). The following is based on Millar (1992), but the use of conditional likelihood can be seen in a range of similar problems where there is incomplete detection (for a recent review see Huggins and Wang, 2007). Let yij be the counts of fish in the *i*th pass of the *j*th sample, with associated abundance *Nj* and capture efficiency *pj*.

,  (1)

This model assumes 1) a closed population, 2) all fish are equally and independently catchable and 3) the capture efficiency is the same for each fishing pass. This results in a multinomial distribution with *Nj* trials and cell probabilities *pj* (1- *pj*)*i*-1, *i* = 1, …, *Sj*. Because interest is in estimating capture efficiency *pj*, the parameters *Nj* can be considered nuisance parameters. For fixed *pj*,

,  (2)

is sufficient for *Nj*. Conditioning on *Tj* eliminates the dependence on *Nj* and the conditional distribution of *yij*, *j* = 1, …, *k*, is multinomial with *Tj*, trials and cell probabilities

, ,  (3)

The denominator in (3) is the probability of that a fish was caught over all, hence (3) is the probability of capture conditional on the observed fish. The full likelihood for *pj* can be written succinctly as

, (4)

where

, (5)

It is assumed that the capture efficiency for each sample can be expressed as a linear function of covariates on the logistic scale

 (6)

Although the use of linear models may sound limiting, the above equation can include: factors, interactions, fixed degrees of freedom splines and spatial models. This will be illustrated in the application (Section 3) below. Models are fitted by maximising the conditional likelihood (4) using a robust optimisation tool provided by the STAN library (ref) for R (ref).

If fish are not equally and independently catchable, then the counts *yij* may be overdispersed with respect to the multinomial probabilities defined by (3). Over dispersion can exists within a site either because a) fish are not behaving independently, or b) the capture probability changes with fishing pass. Within-sample overdispersion can be assessed by comparing the deviance components for each sample to a chi-sq distribution with *Sj* – 2 degrees of freedom. If samples are found to be significantly overdispersed this implies violations of the model assumptions, and these samples should be removed. Overdispersion can occur between sites either due to lack of model fit or non-independence between samples. Between site overdispersion can be assessed by comparing the deviance of a saturated model for *pj*, *Dsat*, to that of some large model, *Dfull*, the biggest that would be considered for a given set of covariates. Deviance residuals can be constructed from the estimated capture efficiencies from the saturated model  and those estimated from the full model 

 (7)

Where *dj* is the component of *Dsat* - *Dfull* due to the *j* th sample. These residuals can be used to check for lack of fit. If there is no evidence for lack of fit and overdispersion is present the approach suggested here is to estimate the overdispersion using

 (8)

and apply this in model selection. This estimate provides a means of approximately addressing between site overdispersion through using a scaled log likelihood in model selection criteria, for example the AIC (Akaike, 1973) adjusted for overdispersion is

 (9)

And similarly for BIC (Schwarz, 1978)

 (10)

The next section shows the practical application of the conditional likelihood modelling approach applied to a large scale national dataset. This both demonstrates the utility of the model described and provides an opportunity to present an approach for dealing with and making predictions for large scale datasets.

# 3. Application

## Data

Figure 1 shows the locations of xx electrofishing samples in Scotland between 1980 and 2014. The data consist of 22000 samples taken from xx distinct sites covering 208 catchments. Data was obtained from Fisheries trusts and boards, who sample at local scales and Marine Scotland Science (MSS) and the Scottish Environment Protection Agency (SEPA) who sample at a Scotland level (Figures 1 and 2). Electrofishing samples consisted of between 2 and 6 passes, however the majority of samples had 3 (84%) or 2 (14%) passes. Counts of salmon and trout were collected at each pass with life stage (fry or parr) being estimated in the field. Common recording included location of sample site, date, and fished area.

A feature of this particular large scale datasets is a lack of common standards and the qualitative nature of local assessments relating to habitat. This was addressed here by generating covariates retrospectively in a consistent manner. Several types of covariates considered: habitat, sampling, spatial and temporal. In the case of habitat variables these were generated using a geographic information system (GIS).

The habitat covariates considered are proxies for processes considered to affect capture probability. These are: *altitude*, upstream catchment area (*UCA*), distance to sea (*DS*), *gradient*, landuse and channel width (*width*). Landuse measures the proportion coverage of seven landuse types*:* *marsh*, *urban*, mixed woodland (*mixed*), deciduous woodland (*deciduous*), conifer woodland (*conifer*) and *other*. Altitude, for example, affects river temperature and gradient influences channel velocity. All habitat covariates are treated as continuous variables. The organisation that collected the data was used a proxy for the effects of different sampling procedures. *Organisation* was treated as a categorical variable with xx levels. To account for spatial variability not incorporated in habitat variables sites were attributed to regions defined by SEPA hydrometric areas (HA). SEPA HAs represent spatially coherent catchment groupings of roughly similar areas (Figure 1), that contain between 1 (e.g., Tay) and x (e.g., west coast) catchments reflecting regional differences in catchment size. *HA* is treated as a categorical variable with x levels. Temporal effects were *year* which is treated as a factor, and day of the year (*doy*) which was treated as a continuous variable. Figure 3 shows a scatterplot of the continuous covariates including latitude and longitude to indicate spatial coverage.

## Model selection

Prior to model selection a test for within sample overdispersion was condicted which identified xx samples as outliers, these sites showed clear evidence for variable capture efficiency, an example being a sequence of counts ***y****j* = (95, 80, 2). Subsequently, a test for between sample overdispersion was conducted based on a full model

GIVE FULL MODEL (11)

This estimated the overdispersion to be 2.5 and analysis of residuals showed this was not related to lack of fit. Hence all further model selection and inference was adjusted to account for overdispersion.

Model selection was based on a forwards and backwards stepwise selection procedure starting from a model with no covariates. Continuous covariates were considered first as linear effects then as smoothers with fixed 3 degrees of freedom. By fixing the degrees of freedom it was possible to fit these effects as linear effects (see appendix for details). Hydrometric area was fitted as a categorical variable were neighbouring regions are correlated. This was achieved through the use of a smoothed spatial effect with 12 degrees of freedom that provided a reasonable compromise between complexity and model fit. Again this can be fitted as a linear effect by fixing the degrees of freedom (Appendix 1).

Adjusted BIC was used due to provide more parsimonious models given the large number of observations. The importance of covariates to the overall model fit was assessed by the change in BIC by dropping one covariate at a time from the final model.

The model selected for the capture efficiency of salmon fry was

logit p ~ Organisation + s(DoY) + Year + regional(HA) + DS + Width + s(Gradient) (12)

The maximum likelihood fit of the above model is summarised in Table 1. The model shows no lack of fit if the overdispersion is fixed at 2.5. Figure 4, 5 show the fitted effects of organisation, hydrometric area, Figure 6 the remaining effects, with temporal variation in deviance residuals by organisation shown in Figure 7.

* List effects in order of importance
* Plot effects and describe them Figures 5, 4 and 6
* Plot residuals - temporal variability by organisation Figure 7.
* Brief understanding: this is consistent with understanding of… [this can include the interpretation to focus the discussion on the modelling]
* What was found and how it relates to literature. Is it consistent with previous studies?

## Performance of different estimates of capture probability

Generally there are three approaches for dealing with capture efficiency *p* in analyses of electrofishing data: assume p is constant, 2) estimate *p* for each site and 3) model *p* as done here. In order to assess the relative performance of these approaches the effect on estimates of density are shown. For a given estimate of capture efficiency for the *j* th sample, , the maximum likelihood estimate of abundance is

 (13)

The influence on estimates of abundance from alternative estimates of capture probability  can be measured by the ratio

 (14)

which is less (greater) than 1 when the alternative estimate results in an underestimate (overestimate) with respect to the modelling approach. The variance of the density estimates can also be computed and compared. Figure 8 shows the effect of assuming a common capture probability – describe spatial trends.

Figure 9 shows the loss in precision when capture probability is estimated separately for each site – describe trends.

# 4. Discussion

## General discussion/recommendations

* We present an approach for estimating and modelling electrofishing capture probability for large datasets that provides a balance of complexity and computational efficiency.

## Method

* A wide variety of models can be fitted as linear terms when the degrees of freedom are selected prior to analysis, for example, interactions smoothers and … refer to package.
* Work is underway to … fitting penalised splines then AIC or GCV can be used to estimate the smoothing parameters. It is important when using AIC to estimate the appropriate degrees of freedom of the smoothing terms which reduce as penalisation increases.
* Although HBMs are a very useful tool in ecological modelling (Cressie), they can be difficult to extend to large datasets such as the Scotland wide data being analysed here. Previous HBM developments include: Conroy et al. (2008), Wyatt, Rivot. However all of these approaches also consider density either as a random effect or in terms of covariates and in order to model capture probability end up with a hierarchical Bayesian model. Why is this approach better – model selection and exploration. Although people don’t report time taken fit models. SNIFFER reported it took 7 days for model fitting, whereas the approach reported here takes 15mins for the entire model selection procedure.

## Generating covariates retrospectively

* Discuss how GIS was at generating proxies for the reported influencers on catchability.
* Issues with spatial confounding. Spatial terms can be thought of as capture un-modelled variation – advisable to attempt to find covariates that can describe the likely cause of the variation possibilities include. Other spatial model to be used as covariates: SRTMN.
* equipment use:
  + Organisation is a mixture of effects. Something like: A variety of sampling methods were used across the organisations who supplied data: with and without stop nets, backpack electrofishing, bank based equipment and generators. Unfortunately information on sampling equipment was not routinely or reliably recorded across data sources and as such could not be formally included in the analysis.
  + In a more detailed study it might be possible to model these effects separately (see refs )
* Talk about correlation between covariates – talk about differences between distance to sea, altitude gradient. It is a potential issue (compute correlations and quote maximum correlation).

## Management consequences

* An issue identified is the lack of uptake of recent developments for modelling capture probability into the applied fisheries community. This paper highlights Management consequences when assuming constant capture probability / site by site capture prob.
* Although single pass electrofishing allows for greater spatial coverage, it comes with implication of bias and precision.
* The use of a model for capture probability would allow the incorporation of single pass fishing data (in conjunction with sufficiently representative multipass data) into quantitative advice.
* Although the model can be used to estimate p for new sites, it is still essential to maintain a level of multipass fishing to permit valid use of single pass sampling. Critical given year and organisation effects, especially where it is unknown what causes the effects of organisation! Also important to record information on sampling approaches and equipment and personnel. New databases allow this but not possible for older data. Effects of various components should be investigated.

# 5. Acknowledgments

The authors thank the Scottish Fisheries Co-ordination Centre (SFCC) and associated fishery trusts that provided data: Annan, Argyll, Ayr, Clyde, Conon, Deveron, Esk, Findhorn, Forth, Galloway, Kyle of Sutherland, Lochaber, Ness and Beauly, Outer Hebrides, Dee, Spey, Tay, Tweed, West Sutherland, Wester Ross, Naver and Nith. The data contributions of SEPA, Alan Youngson and MSS staff are also gratefully acknowledged. We thank Sean Dugan for liaising with trust members and exporting data from the SFCC database.

# References

Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M. and Milner, N.J. 2003. Habitat Requirements of Atlantic Salmon and Brown Trout in Rivers and Streams. *Fisheries Research*, **62** (2): 143–70.

Bohlin, T., Pettersson J. and Degerman E. 2001. Population Density of Migratory and Resident Brown Trout (*Salmo Trutta*) in Relation to Altitude: Evidence for a Migration Cost. *Journal of Animal Ecology*, **70** (1): 112–21.

Borgstrøm, R. and Skaala Ø. 1993. Size-Dependent Catchability of Brown Trout and Atlantic Salmon Parr by Electrofishing in a Low Conductivity Stream. *Nordic Journal of Freshwater Research*, **68**: 14–21.

Cressie, N., Frey, J., Harch, B. and Smith, M. 2006. Spatial Prediction on a River Network. *Journal of Agricultural, Biological, and Environmental Statistics*, **11** (2): 127–50.

Deschênes, J. and Rodríguez M.A. 2007. Hierarchical Analysis of Relationships between Brook Trout (*Salvelinus Fontinalis*) Density and Stream Habitat Features. *Canadian Journal of Fisheries and Aquatic Sciences*, **64** (5): 777–85.

Fahrmeir, L., Kneib, T., Lang, S. and Marx, B. 2013. Regression. Berlin, Springer.

Fausch, K.D., Hawkes, C.L. and Parsons, M.G. 1988. Models That Predict Standing Crop of Stream Fish from Habitat Variables: 1950-85. *Gen. Tech. Rep. PNW-GTR-213*. U.S. Department of Agriculture.

Godfrey, J.D. 2005. Site Condition Monitoring of Atlantic Salmon SACs. Report by the SFCC to Scottish Natural Heritage, Contract F02AC608, 274 pp.

Kennedy, G.J.A. and Strange, C.D. 1981. Efficiency of Electric Fishing for Salmonids in Relation to River Width. Aquaculture Research, **12**: 55–60.

Lanka, R.P., Hubert, W.A. and Wesche, T.A. 1987. Relations of Geomorphology to Stream Habitat and Trout Standing Stock in Small Rocky Mountain Streams. *Transactions of the American Fisheries Society*, **116** (1): 21–28.

Marsh, T. J. and Hannaford, J. (Eds). 2008. UK Hydrometric Register. Hydrological data UK series. Centre for Ecology & Hydrology. 210 pp.

Moran, P.A.P. 1951. A Mathematical Theory of Animal Trapping. *Biometrika*, **38**: 307–11.

Niemelä, E., Julkunen, M. and Erkinaro, J. 2000. Quantitative Electrofishing for Juvenile Salmon Densities: Assessment of the Catchability during a Long-Term Monitoring Programme. *Fisheries Research*, **48** (1): 15–22.

Otis, D.L., Burnham, K.P., White, G.C. and Anderson, D.R. 1978. Statistical Inference from Capture Data on Closed Animal Populations. *Wildlife Monographs*, **62**: 3–135.

Peterson, E.E., Ver Hoef, J.M., Isaak, D.J., Falke, J.A., Fortin, M.J., Jordan, C.E., McNyset, K., *et al.* 2013*.* Modelling Dendritic Ecological Networks in Space: An Integrated Network Perspective. *Ecology Letters*, **16** (5): 707–19.

Prévost, E., Parent, E., Crozier, W., Davidson, I., Dumas, J., Gudbergsson, G., Hindar, K., McGinnity, P., MacLean, J. and Sættem, L.M. 2003. Setting Biological Reference Points for Atlantic Salmon Stocks: Transfer of Information from Data-Rich to Sparse-Data Situations by Bayesian Hierarchical Modelling. *ICES Journal of Marine Science: Journal Du Conseil*, **60** (6): 1177–93.

Rivot, E., Prévost, E., Cuzol, A., Baglinière, J.-L. and Parent, E. 2008. Hierarchical Bayesian Modelling with Habitat and Time Covariates for Estimating Riverine Fish Population Size by Successive Removal Method. *Canadian Journal of Fisheries and Aquatic Sciences*, **65** (1): 117–33.

Rosenfeld, J., Porter, M. and Parkinson, E. 2000. Habitat Factors Affecting the Abundance and Distribution of Juvenile Cutthroat Trout (*Oncorhynchus Clarki*) and Coho Salmon (*Oncorhynchus Kisutch*). *Canadian Journal of Fisheries and Aquatic Sciences*, **57** (4): 766–74.

Rue, H. and Held, L. 2005. Gaussian Markov Random Fields. Monographs on Statistics and Applied Probability, 104. Boca Raton, Chapman and Hall / CRC press.

Schwarz, G. 1978. Estimating the dimension of a model. *The annals of statistics*, **6** (2): 461-464.

SNIFFER. 2011. River Fish Classification Tool: Science Work, Phase 3 Report, Final. Project WFD68c. Scottish and Northern Ireland Forum for Environmental Research.

Wood, S.N. 2006.  Generalized additive models: an introduction with R. CRC press.

Wood, S.N. 2011. Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73** (1): 3–36

Wyatt, R.J. 2002. Estimating Riverine Fish Population Size from Single- and Multiple-Pass Removal Sampling Using a Hierarchical Model. *Canadian Journal of Fisheries and Aquatic Sciences*, **59** (4): 695–706.

Wyatt, R.J. 2003. Mapping the Abundance of Riverine Fish Populations: Integrating Hierarchical Bayesian Models with a Geographic Information System (GIS). *Canadian Journal of Fisheries and Aquatic Sciences*, **60** (8): 997–1006.

Wyatt, R.J. 2005. River Fish Habitat Inventory Phase 2 : Methodology Development for Juvenile Salmonids. Science Report SC980006. Environment Agency.

Wyatt, R.J. and Barnard, S. 1997. The transportation of the maximum gain salmon spawning target from the River Bush (N.I.) to England and Wales. *R & D Technical Report* *W65*. Environment Agency.

Wyatt, R.J., Sedgwick, R. and Burrough, R. 2007a. A Statistical Approach to the Assessment of Coarse Fish Populations. Science Report SC030214. Environment Agency.

Wyatt, R.J., Sedgwick, R. and Simcox, H. 2007b. River Fish Habitat Inventory Phase 3: Multi-Species Models. Science Report SC040028. Environment Agency.

Zippin, C. 1956. An Evaluation of the Removal Method of Estimating Animal Populations. *Biometrics*, **12** (2): 163–89.

# Tables

|  |  |  |
| --- | --- | --- |
| **Covariate** | **Degrees of Freedom** | **Change in BIC** |
| Organisation | 23 | 889.2 |
| s(DoY) | 2 | 438.1 |
| Year | 16 | 207.4 |
| regional(HA) | 7 | 176.2 |
| DS | 1 | 57.3 |
| Width | 1 | 27.7 |
| s(Gradient) | 2 | 0.6 |

Table 1 The relative importance of explanatory covariates in the capture probability model as indicated by changes in BIC where single terms were removed from the final model. The degrees of freedom shows the reduction in parameters associated with removing the term.

# Figures

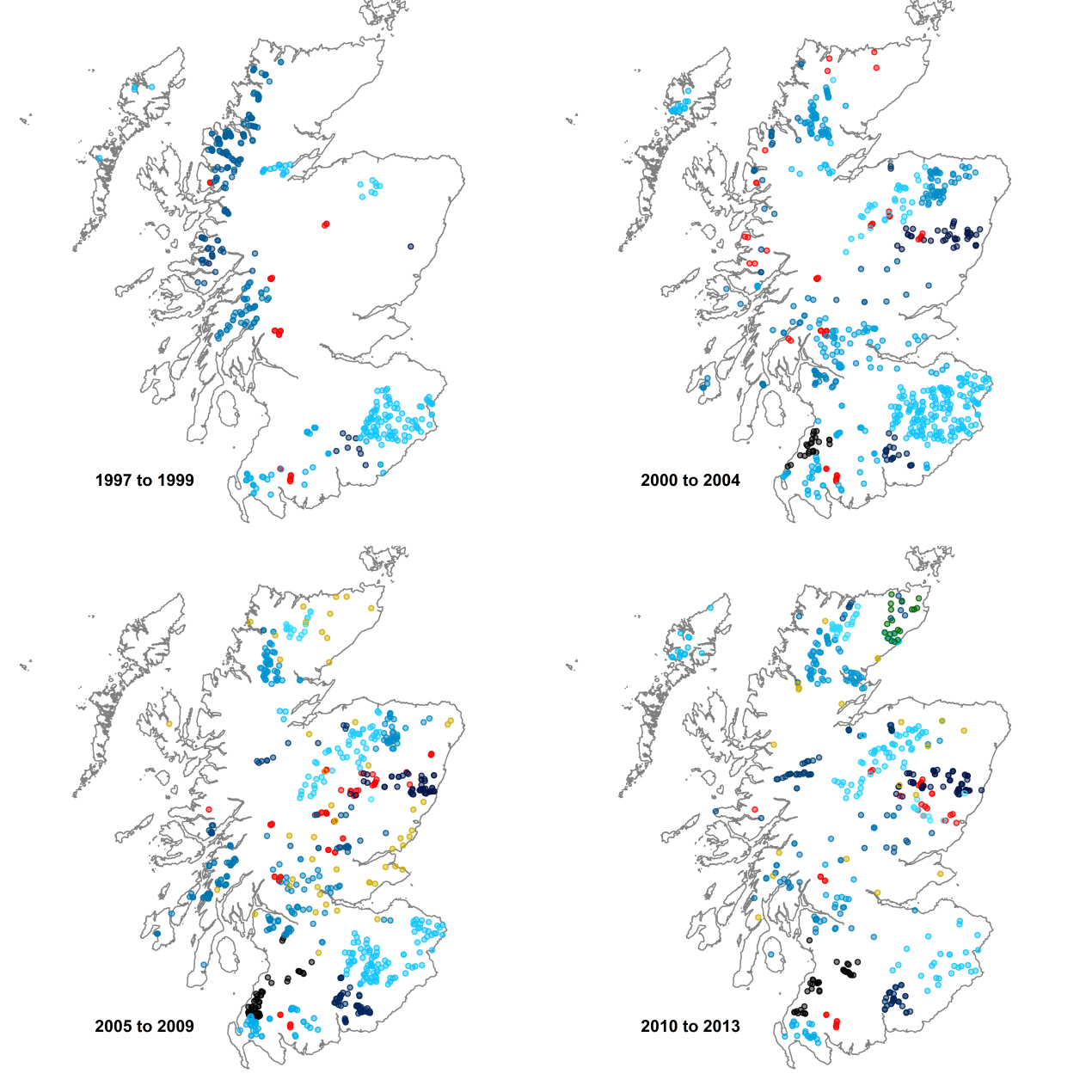


Figure 1 Spatial and temporal data coverage (un-stocked sites with multi-pass electrofishing, below impassable barriers) between 1997 and 2013. Prior to 1997 there are too few data from too constrained an area for useful large scale model fitting. Data are colour coded by source: MSS (red), SEPA (yellow), Other (green), SFCC (each trust is represented by a different shade of blue).

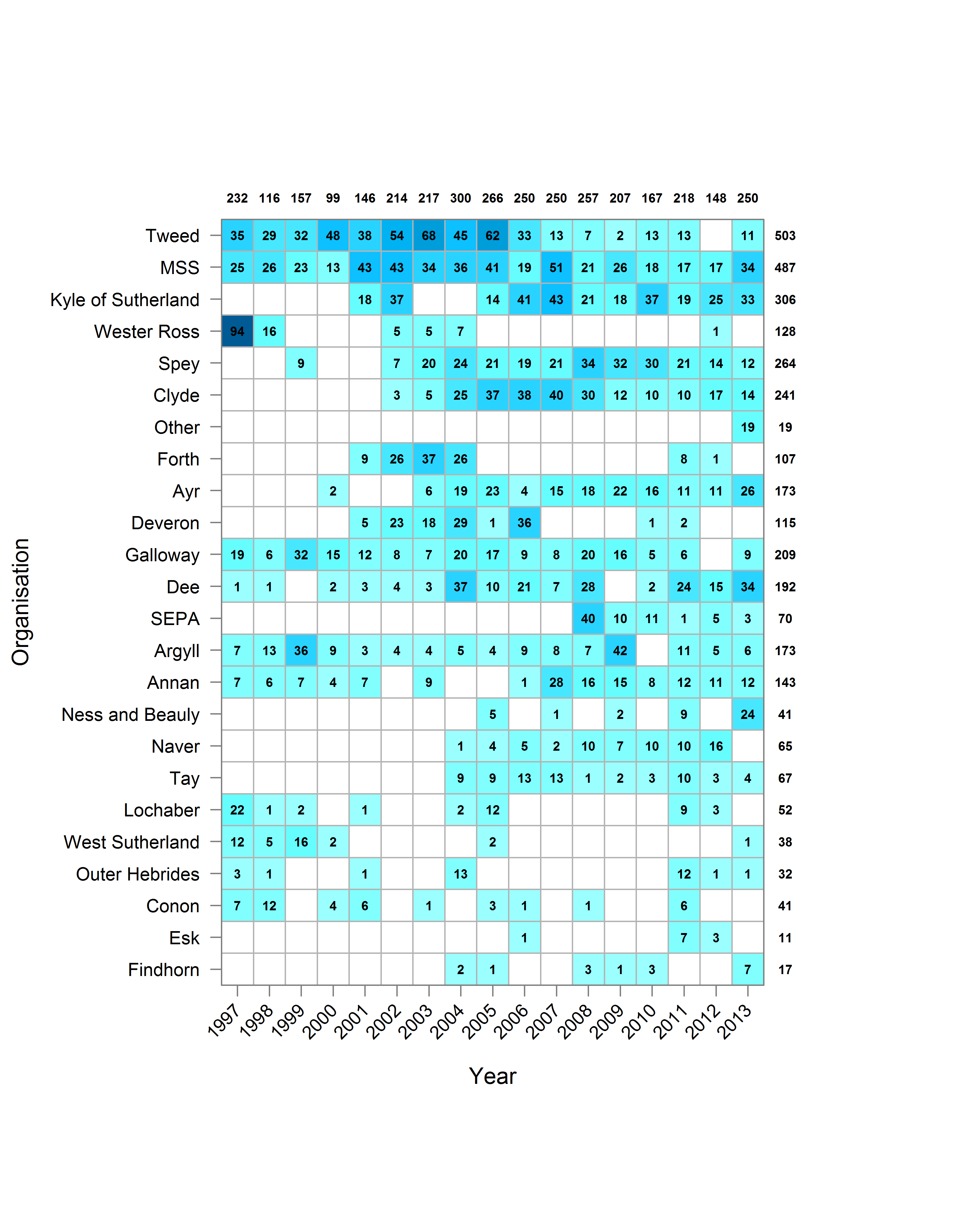
****

Figure 2 Schematic showing the temporal coverage of data by provider. Organisations are ordered by the mean annual number of site visits. Total site visits by Year and by Organisation are given in the margins.

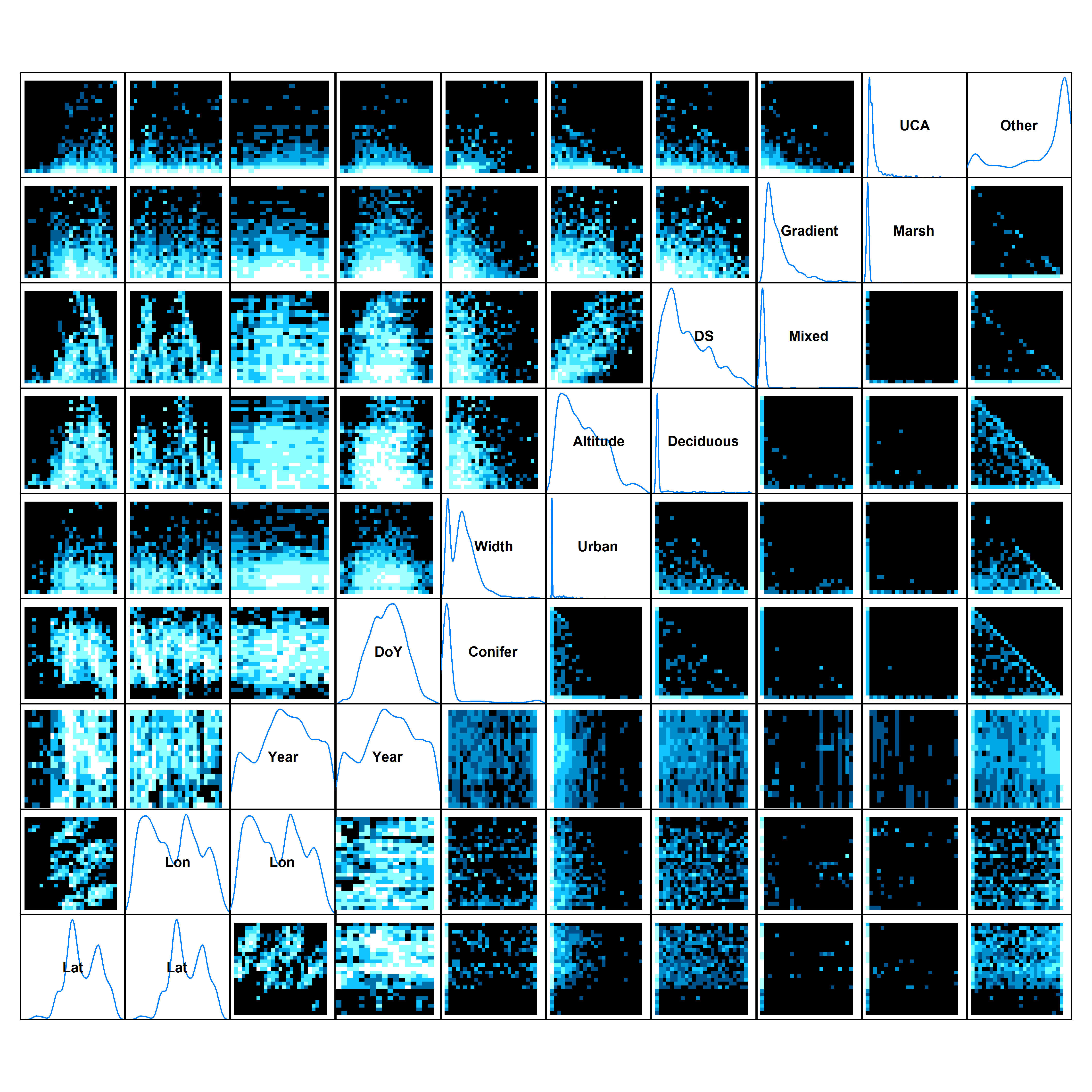


Figure 3 Density plots showing the distribution of available data in relation to combinations of environmental covariates (white: lots of data, blue: few data, black no data). Latitude (Lat) and Longitude (Lon) are included to provide an indication of spatial coverage although these were not included in model fitting.

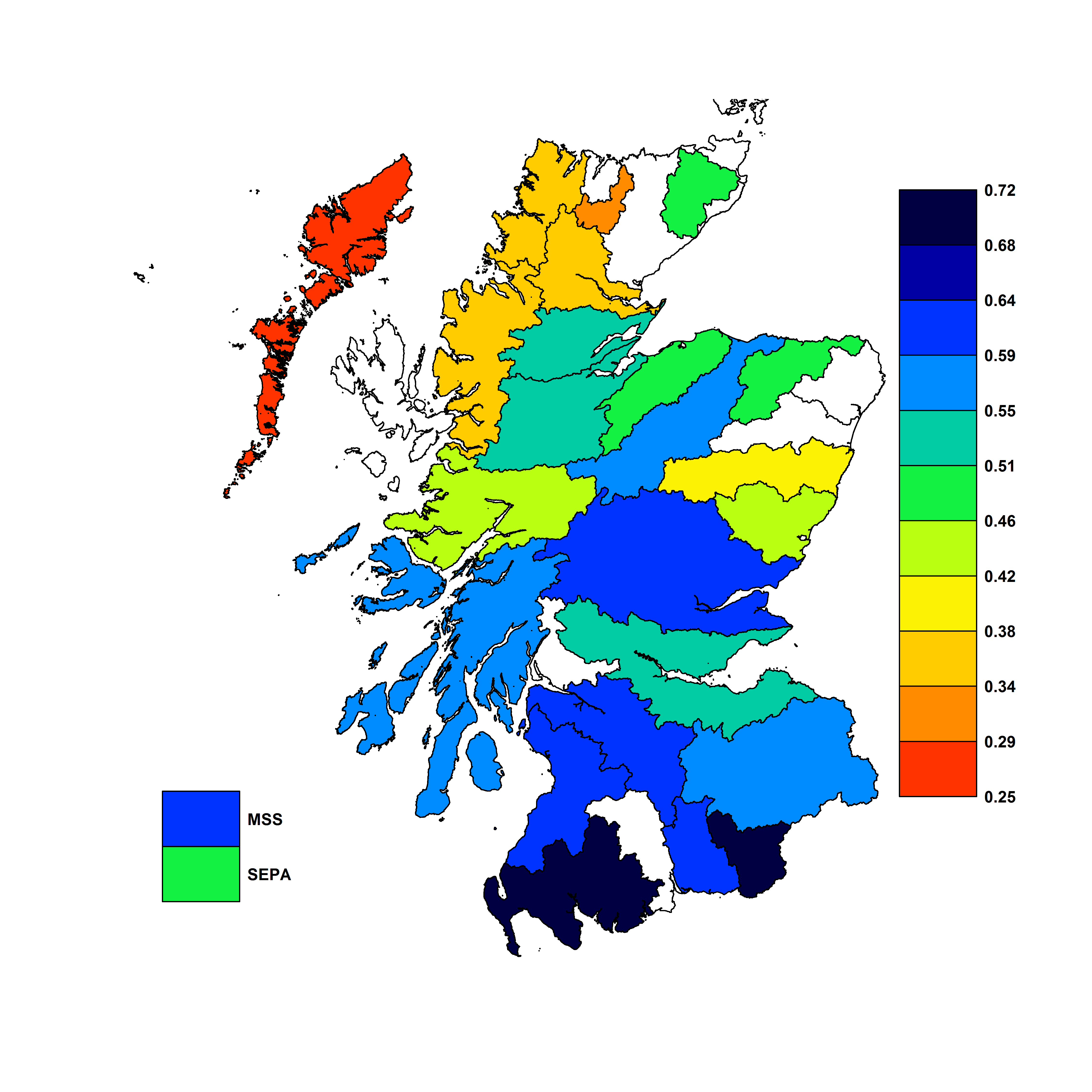


Figure 4 Estimates of capture probability by data provider (Organisation). Estimates are plotted in relation to the geographic area of responsibility (e.g., Trust Boundaries) of each organisation, although data coverage may be broader. Note that the Galloway Fisheries Trust covers the Galloway region and the Border Esk to the East. White areas indicate either no multi-pass fishing data or no data provider present in that region. In the case of the River Don, permission to use available data came too late to include in this report. MSS and SEPA estimates are indicated to the side of the map given their wide ranging data coverage. Estimates are conditioned on HA Tay, Year 1996 and median values for all remaining covariates. Missing estimates reflect a lack of data availability at the time of model fitting. Map based on digital spatial data licensed from Centre for Ecology and Hydrology, © NERC.

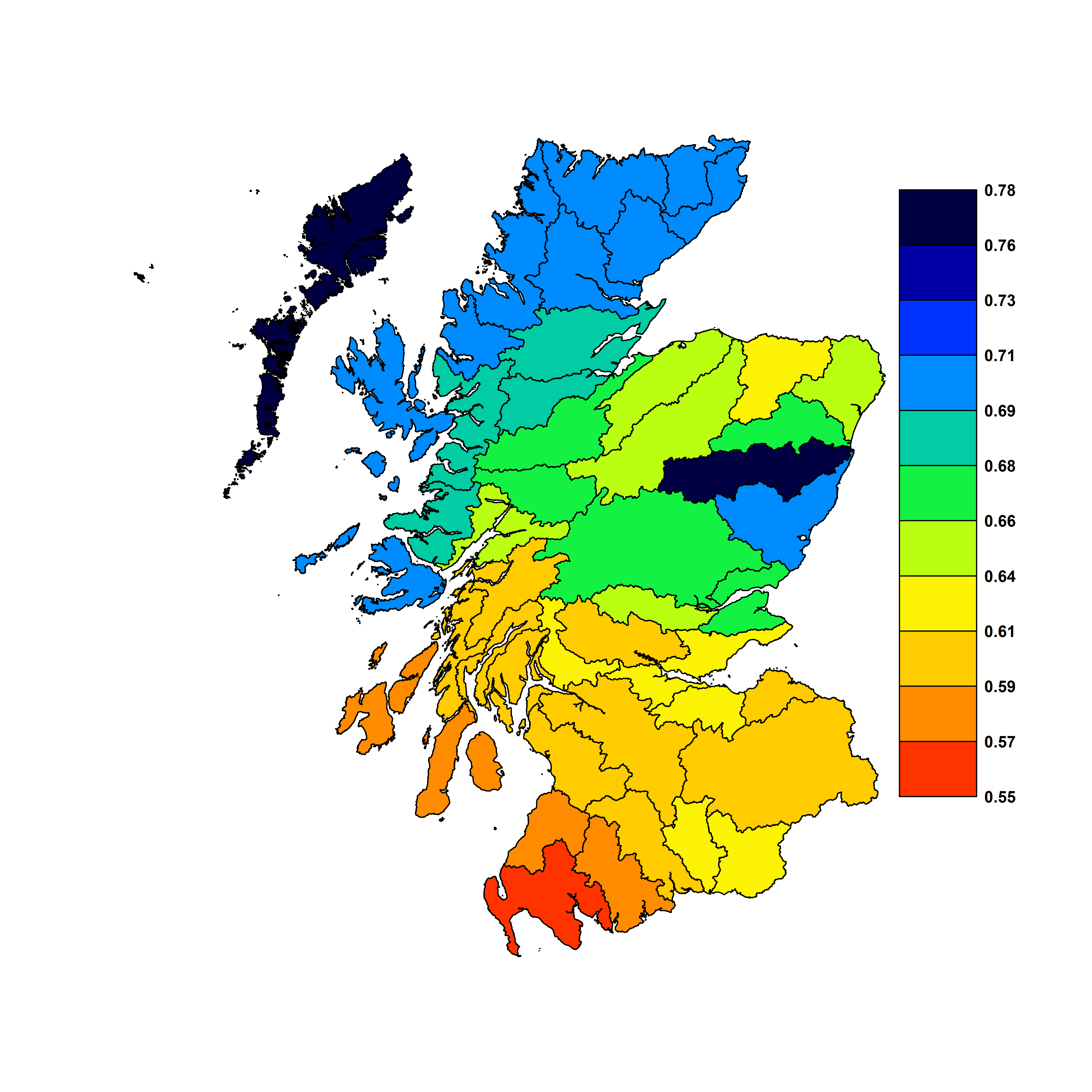


Figure 5 Estimates of spatial variability in capture probability (HA). Estimates are conditioned on Year 1996, Organisation MSS, and median values for all remaining covariates. Map based on digital spatial data licensed from Centre for Ecology and Hydrology, © NERC.

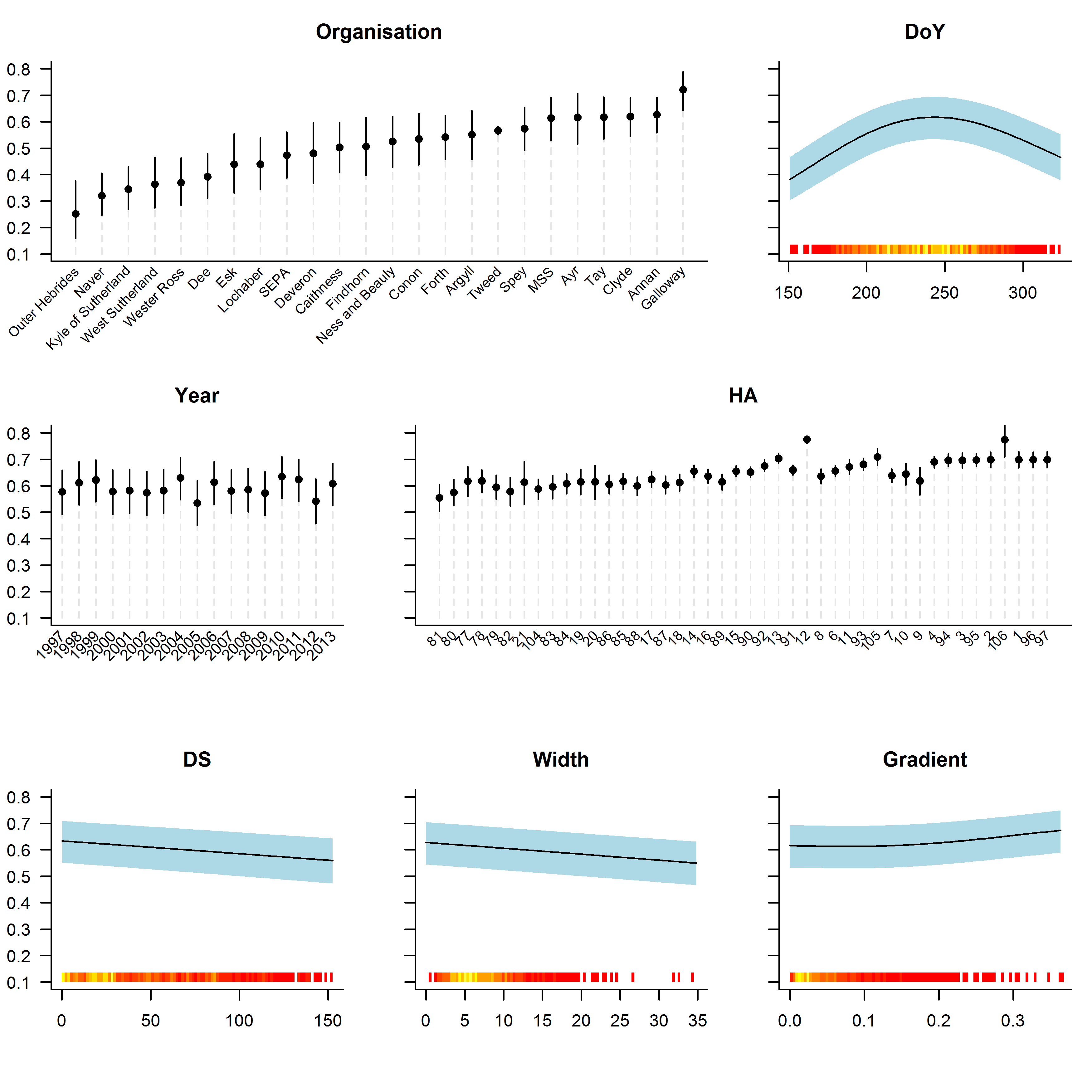


Figure 6 Relationships between capture probability and covariates. Plots are conditioned on HA Tay, Organisation MSS, Year 1996, and median values for remaining covariates. Organisation names have been abbreviated. HA values are ordered from South to North. 95% pointwise confidence intervals are shown as shaded blue areas or vertical bars. A ‘rug’ indicates the distribution of available data on the x-axis (red: few values, yellow: many values).

Figure 7 Residuals by organisation.

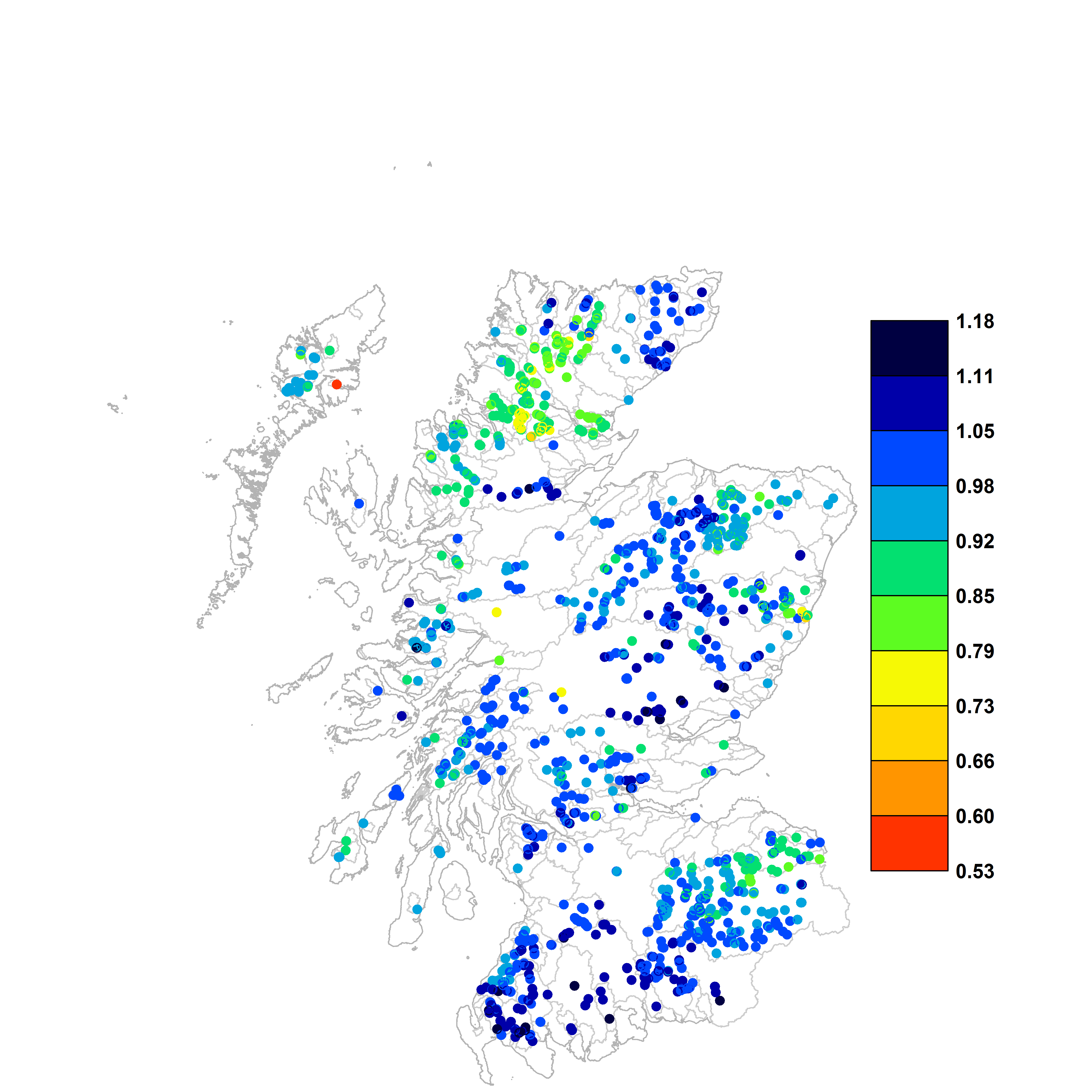


Figure 8 Map showing proportional differences in density estimated using a constant mean capture probability (0.53) and modelled capture probabilities. Higher values indicate higher modelled densities relative to constant capture probability. Map based on digital spatial data licensed from Centre for Ecology and Hydrology, © NERC.

# Appendix 1

The derivation of reduced rank spatial effects.

NOTES:

Include reference to software.