The relationship between environmental variables and the occurrence of two BAP flies *Dolichopus laticola* and *D. nigripes* (Diptera: Dolichopodidae) in Norfolk fens

Report (part 1) of second year's work funded by DEFRA

C. Martin Drake

2011

Hymettus Ltd.
Lea-Side, Carron Lane
MIDHURST
West Sussex GU29 9LB
Registered in England No. 5761114
Registered Charity No. 1124636

Introduction

Two species of flies, *Dolichopus laticola* Verrall and *D. nigripes* Fallén (Diptera, Dolichopodidae), were included in the list of priority species in the Biodiversity Action Plan (www.jncc.gov.uk, accessed October 2010). Their distribution is almost confined to the fens of the Norfolk Broads. An initial study in 2010 investigated the relationship of these flies with the habitat features at six fens (Drake 2010). Although analysis indicated a few broad habitat preferences based on broad habitat categories to which each sample was allocated, it was felt that more could be gained from detailed analysis of the data collected at a finer level. This report gives the results of that investigation. The earlier report gives methods and details of the environmental variables that were measured.

Statistical methods

Initial exploration showed that the abundance data for *D. laticola* was far from normally distributed, owing to the large number of zero values and rapidly decreasing abundances from one individual to a maximum of seven. Scatterplots of the abundance of *D. laticola* with each variable showed almost no trends when a smoothing curve was applied, and a Pearson correlation coefficient was always low; the highest values were for wetness ($r = 0.3$) and litter ($r = -0.25$). Further regressions using, for instance, polynomial or logarithmic transformations, also showed insignificant regressions with very small r^2 values. Coplots, where the response variable is plotted against each explanatory variable but subdivided into groups defined by a second environmental variable, showed almost no interaction between the pairs of variables. Also, the distribution of nearly all variables was not normal. Therefore, parametric multiple regression or generalised linear regression could not be applied as the conditions of normality was violated. In fact, a GLM analysis using presence – absence data for *D. laticola*, and forward selection of the variables, including two-way interactions between the continuous variables, showed no variable to be significant. Therefore, even had the model been valid, it appeared highly unlikely to have given a useful result.

Instead, the relationship of each *Dolichopus* species with these was examined using tree models (Zuur *et al*., 2007). These show the relationship of the response variable (*Dolichopus*) with all explanatory variables and give an indication of the relative importance of the variables. They are unaffected by lack of normality or linearity in the data, and are easy to interpret so have advantages over other types of modelling.

Abundance data were investigated using a regression tree and presence-absence data using a classification tree. Both types of tree were investigated since there is an ecological difference between finding a relationship just for presence or absence compared to one that shows preference reflected in larger catches.

As the structural variables had been measured in rather too detailed manner, they were simplified in the expectation that any important effects would still stand out.

The vegetation 'layers' measured on the DAFOR scale were converted to numbers (1 for Rare, 2 for Occasional, etc). This conversion approximated to a log transformation so further transformation to achieve normality seemed inadvisable.

A few variables were excluded for the following reasons:

- Moss and tussock were rarely recorded.
- Ditch vegetation was three nominals, of which ‘short’ and ‘*Juncus’* were rarely recorded, leaving just ‘reed’ which was equivalent of merely having sampled a ditch.
- The three categories of reed, sedge and mixed were amalgamated into one called ‘monocot’.
- Ditch age had to be excluded as non-ditch sites could not be scored on this pseudo-numerical scale.
- The four nominal classes of management were not real measurements but constructs based on a best guess.
- The analysis by habitat features had shown that *D. laticola* avoided carr woodland so, to prevent this appearing as the first split in the tree, the 19 samples from carr were excluded, leaving 162 samples for *D. laticola* and 80 samples from the Bure for *D. nigripes*.

Results

Dolichopus laticola

A classification tree, obtained using presence-absence data, with just two branches was the best fit when the ‘one standard deviation rule’ was applied. However, further useful information was found in further splits even if these could not be fully supported by cross-validation (Figure 1). The notation on the figure is that the condition that defines the split at a node is true for the left-hand branch, and the value at the bottom of the branch is the average for the set of samples in that group; in this case, 1 = present and 0 = absent. Below this average condition is the number of samples for which it is true (left) and false (right). The values for vegetation variables at each node are the average of DAFOR rating once converted to corresponding numbers 1 to 5, and for wetness the average on the scale 1 to 4.

Wetness was responsible for the first split of the tree, and this split can be regarded as statistically significant using the ‘one standard deviation rule’. So in Figure 1, the left-hand branch has a total of 91 samples from sites with an average wetness less than 2.25 (nearer to being ‘soft’ and damp than ‘saturated’). In this left-hand branch of drier sites, *D. laticola* was more often absent in those with low amounts of leaf litter (scored less than ‘frequent’), being found in 25 of 74 (34%) of these low-litter sites and in 10 of 17 (59%) of those with more litter, although this distinction was not supported by the cross-validation test. Of the 71 samples from wetter points on the right-hand branch, *D. laticola* was present in a large proportion (30 of 37 samples, 81%) of the samples with larger amounts of tall herb vegetation (abundant or dominant on the DAFOR scale). Where tall herbs were less frequent, *D. laticola* was less likely to be present in the densest reed or sedge with a DAFOR of dominant. This appeared to show that the fly preferred fairly dense mixed vegetation but not perhaps where reed or sedge were clearly dominant.

The regression tree, derived from the abundance data of *D. laticola*, was also taken beyond the limit set by the pruning rule, and the error versus complexity parameter showed that the tree fitted the model poorly since no split was above the average error (Figure 2). However, this tree suggested that the average density of *D. laticola* was 1.64, which was moderately high in the context of this survey, at sites with at least some scrub, nearly always *Myrica gale* (right-hand branch). Average density was highest at sites with almost no scrub but which were fairly wet and where reed or sedge were not dominant (right-hand side of the main left-hand branch). Dominant monocots clearly led to low densities of *D. laticola*. Densities were also fairly low in drier sites where the soil was less than saturated.

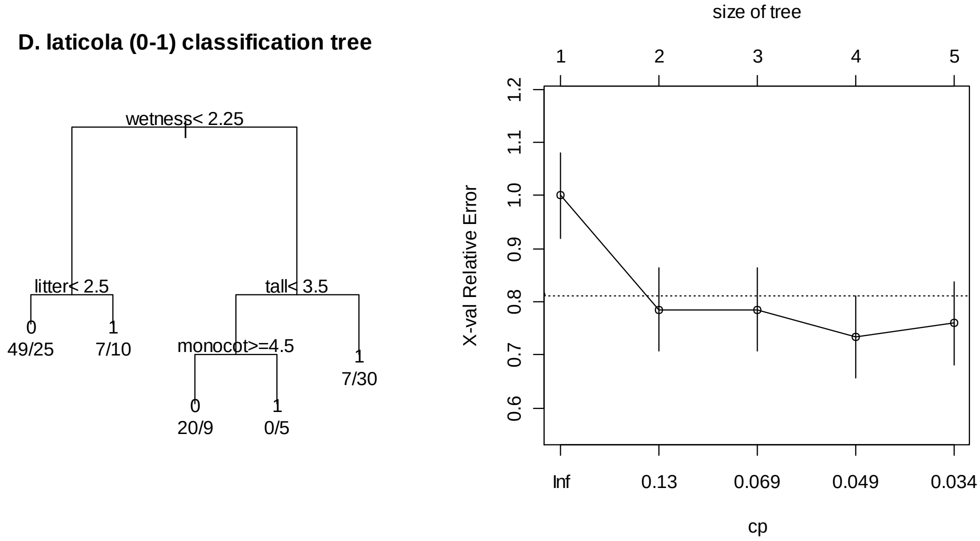

Figure 1. Classification tree for *Dolichopus laticola* using presence-absence data, and the graph of relative error against the complexity parameter.

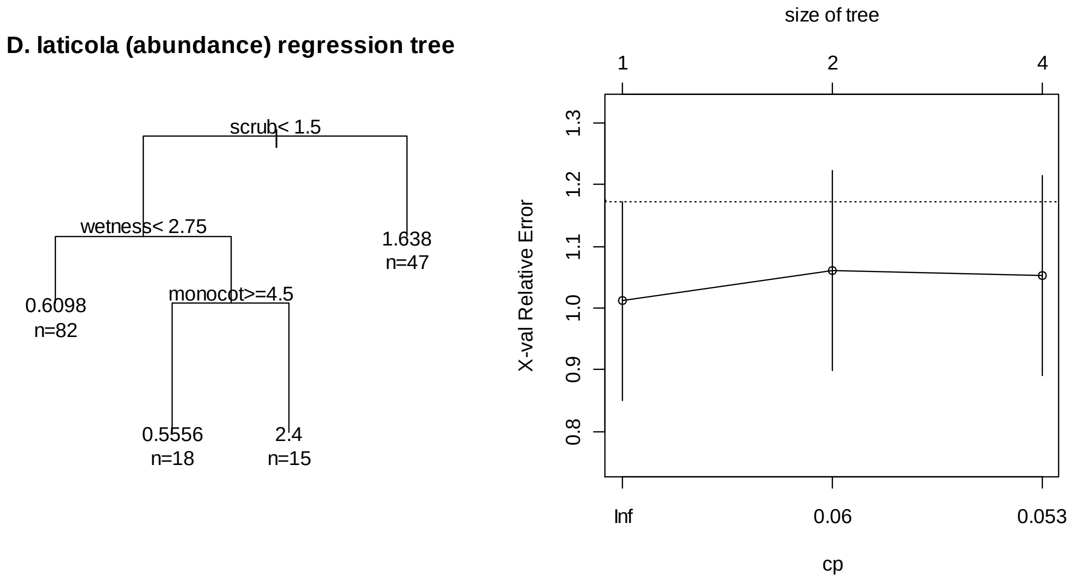

Figure 2. Regression tree for *Dolichopus laticola* using abundance data, and the graph of relative error against the complexity parameter.

Dolichopus nigripes

Neither tree fitted the model well, as indicated by no point in the graphs being above the average error line (Figures 3 and 4). Both had the same initial split, in which *D. nigripes* was rarely recorded (in 7 of 42 samples) where tall reed and sedge were dominant. Where reed and sedge were less than dominant, *D. nigripes* was usually present (in 18 of 23 samples) where there was scarcely any bare peat (0.5 is between rare and occasional on the DAFOR scale) (Figure 3). Its average abundance in these less reed-dominated sites was considerably greater (4.14 flies per sample) where short vegetation was dominant (>4.5 on the numerical DAFOR scale) rather than where short vegetation was only up to ‘abundant’ (Figure 4). As mown paths were frequent at Woodbastwick Fen where *D. nigripes* was most often found, this explains the pre-eminence of short vegetation in this analysis.

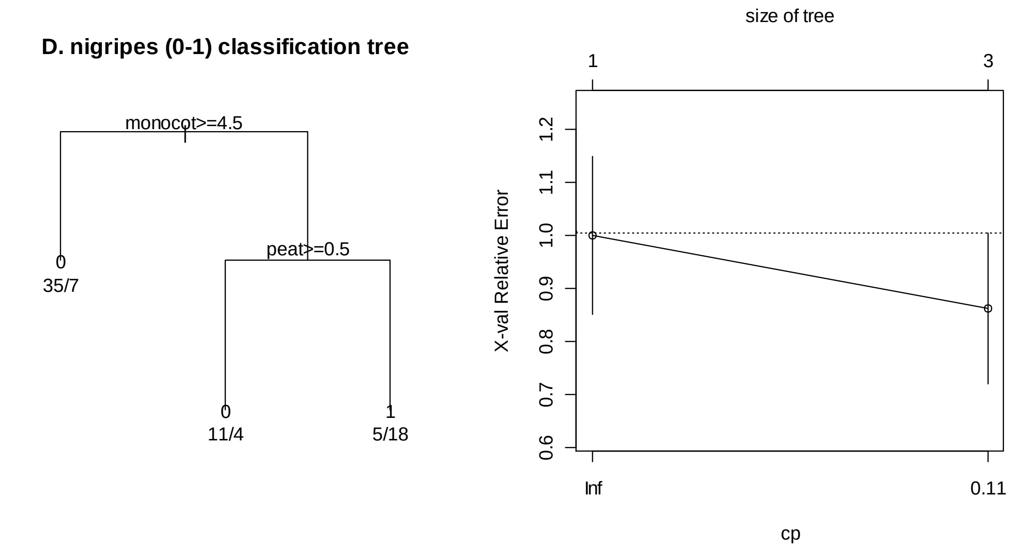

Figure 3. Classification tree for *Dolichopus nigripes* using presence – absence data, and the graph of relative error against the complexity parameter.

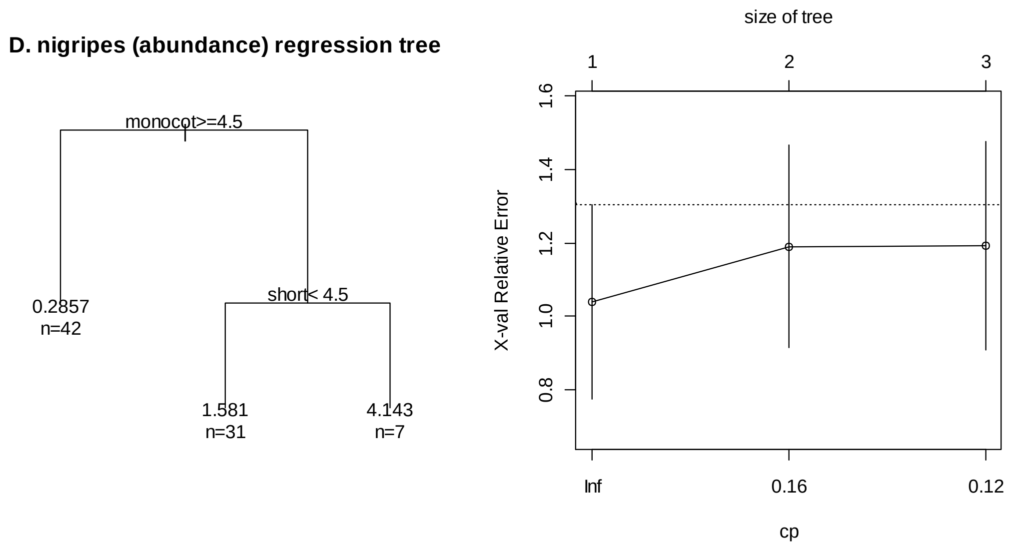

Figure 4. Regression tree for *Dolichopus nigripes* using abundance data, and the graph of relative error against the complexity parameter.

Conclusions

Although the evidence for which factor controlled the local distribution of the two Dolichopus species was weak, the following conclusions were drawn.

- *Dolichopus laticola* preferred vegetation that suggested that it was 'old cut' fen, characterised by a larger proportion of tall herb vegetation and moderate amounts leaf litter. It appeared to avoid places dominated by tall dense reed to the exclusion of tall herb vegetation. This may be more characteristic of frequently cut and wetter commercial reedbed. Higher densities of flies were found in wetter areas, but this result conflicted with an apparent increased frequency of occurrence in slightly drier places (although still damp).
- *Dolichopus nigripes* preferred more open vegetation, in which highest densities were reached, and strongly avoided vegetation dominated by tall reed and sedge. The lack of a relationship with soil wetness may have been to the similarly damp conditions in all the Bure sites.

Acknowledgements

I thank Stephen Freeman (CEH) for discussion about statistical techniques.

References

Drake, C.M. 2010. *The BAP flies Dolichopus laticola and D. nigripes (Diptera, Dolichopodidae) in the Broadland fens of Norfolk*. Unpublished report to Hymettus 1-14.

Zuur, A.F., Ieno, E.N. & Smith, G.M. 2007. *Analysing ecological data*. Springer Science & Business Media, LLC. pp672.