

Ecography

ECOG-00535

Lawson, C. R., Bennie, J., Hodgson, J. A., Thomas, C. D. and Wilson, R. J. 2014. Topographic microclimates drive microhabitat associations at the range margin of a butterfly. – *Ecography* doi: 10.1111/ecog.00535

Supplementary material

Supporting Information: Lawson C.R., Bennie J.J., Hodgson J.A., Thomas C.D., & Wilson R.J. “Topographic microclimates drive microhabitat associations at the range margin of a butterfly”

Contents

Appendix 1: microhabitat association models including an interaction between temperature and host plant cover 2

Appendix 2: details on empirical validation of microclimate model 3

Appendix 3: posterior distributions of microhabitat association model parameters 4

Appendix 4: Z-tests of significance of parameters in the *full* model 5

Appendix 5: outlier analyses 6

Appendix 6: model predictions and data for each study site 8

Appendix 7: discussion and analysis of habitat availability effects 10

Appendix 1: microhabitat association models including an interaction between temperature and host plant cover

Using the same dataset as presented in the main paper, we tested whether host plant cover had different effects on egg-laying in sites with different temperatures. We fitted a model additional to those presented in the main text, termed the *hfull* model. *hfull* contained all terms present in the *full* model, plus an interaction between host plant cover and temperature. This tests the hypothesis that the strength of association with host plant cover changes with site-level temperature.

Including the interaction between temperature and host plant cover worsened the fit of the full model (Table A1). Thus, there was little support for the hypothesis that the effects of host plant cover on egg-laying differed at different temperatures.

Model	Terms										K	AIC	δ AIC	AICWt	
	main effects					interactions									
	quadrat	site-level				bare ground	host plant		temperature						
	b_{ij}	h_{ij}	\bar{b}_j	\bar{h}_j	T_j	$b_{ij}\bar{b}_j$	$b_{ij}\bar{h}_j$	$h_{ij}\bar{b}_j$	$h_{ij}\bar{h}_j$	$b_{ij}T_j$	$h_{ij}T_j$				
full	+	+	+	+	+	+	+	+	+	+		12	1178.9	0.00	0.7
<i>hfull</i>	+	+	+	+	+	+	+	+	+	+	+	13	1180.8	1.9	0.3

Table A1: Model selection table including host plant-temperature interaction, showing empirical support for different models of *H. comma* microhabitat associations, based on sampled egg locations. Constituent terms are indicated by plus signs (+) with empty spaces indicating terms that were not included; see Table 1 in the main text for term descriptions. K=number of parameters; AIC=Akaike’s Information Criteria Score; δ AIC=difference in AIC between current and highest-ranked model; AICWt =AIC weights.

Appendix 2: details on empirical validation of microclimate model

The microclimate model used in this study was first developed and parameterised using measured ambient temperature, short and long-wave radiation and wind speed at a weather station to estimate air temperature near the surface of short *Festuca*-grazed chalk grassland on slopes, using equations to adjust solar radiation due to cloud cover and solar position, and an empirically-parameterised relationship between sward temperature, net radiation, wind speed and ambient (2 metre) temperature (Bennie *et al.*, 2008). Under these circumstances, and excluding situations where temperature inversions formed within the valley system used to test the model during winter and clear nights, the mean absolute error (bias) of hourly mean sward temperature at independent testing sites was between 0 and 1 °C and root mean squared errors (RMSE) were less than 2 °C (Bennie *et al.*, 2008). Since in the context of this study the aim is to model daytime temperatures during the summer flight season of *H. comma*, temperature inversions are highly unlikely to occur. In order to extend predictions across regions to sites without *in situ* weather stations, methods were developed to down-scale regional climate data from hourly data from a weather station in the centre of the study region, by adjusting for consistency with a 5 km resolution climate data set (Bennie *et al.*, 2013). The accuracy of the predicted temperatures therefore also depends on the accuracy of the 5km gridded dataset, the representativeness of the regional weather station in capturing the daily variation in cloud cover and temperature, and any errors incorporated in the downscaling procedure.

To assess the accuracy of the model when used to predict temperature in *H. comma* habitat during the flight season using data from a remote weather station, we measured air temperature at 10 cm height using temperature loggers with an external probe (Tinytag Talk 2, Gemini Dataloggers, Cambridge) on contrasting north- and south-facing slopes of a valley system containing typical *H. comma* habitat (Malling Down; 50.83°N 0.02°E) for 7 days during the August 2010 flight season. On sunny days, the maximum temperature differences between sites during this period were up to 12 °C. The mean absolute error in hourly daytime temperature was 0.7 °C for the north-facing slope and 0.8 °C for the south-facing slope; RMSE was 1.2 for the north-facing slope and 2.3 for the south-facing slope. The mean error in the predicted daily number of hours exceeding 25 °C was 0.21 hours on the north-facing slope and 0.32 hours on the south-facing slope. We conclude that the model is able to capture the main landscape-scale gradients in near-surface air temperature due to differences in exposure to solar radiation between sites.

Appendix 3: posterior distributions of microhabitat association model parameters

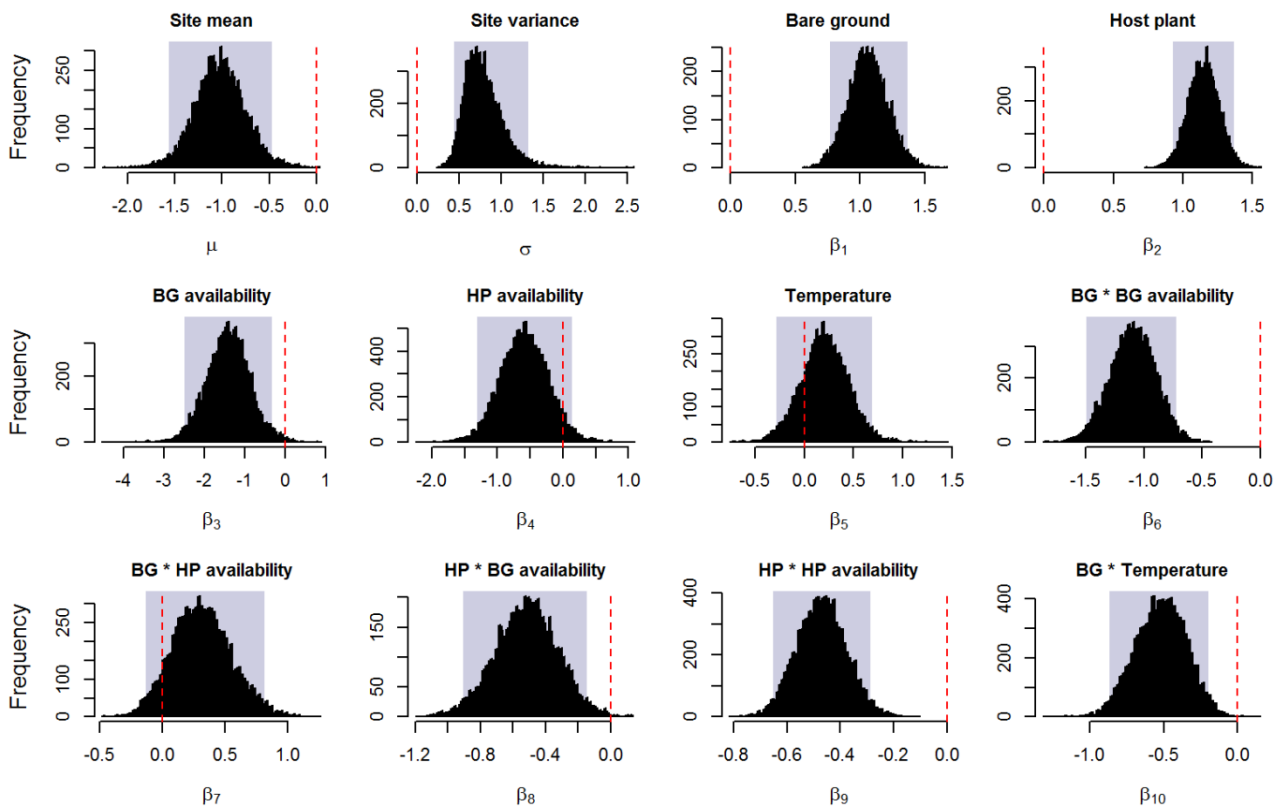


Figure A1: Histograms showing posterior distributions of parameters from MCMC draws. The red line highlights the parameter value of zero, whilst the blue box shows the 95% credible interval for each parameter.

Appendix 4: Z-tests of significance of parameters in the *full* model

To ensure that our conclusions about the importance of variables in the *full* model were robust to the statistical method of analysis used, we assessed the significance of individual parameters in the *full* model using Z-tests of significance, as implemented using the *summary* function in the *lme4* R package. The results are shown below.

Coefficient	Estimate	SE	Z	p
Site mean	-1.03	0.20	-5.14	-
Bare ground	1.00	0.14	6.94	-
Host plant	1.13	0.11	10.35	-
BG availability	-1.36	0.40	-3.37	-
HP availability	-0.60	0.29	-2.06	-
Temperature	0.20	0.18	1.15	-
BG * BG availability	-1.03	0.19	-5.39	>0.00001
BG * HP availability	0.27	0.24	1.13	0.26
HP * BG availability	-0.49	0.19	-2.61	0.0090
HP * HP availability	-0.45	0.09	-5.08	>0.00001
BG * Temperature	-0.45	0.16	-2.87	0.0042

Table A2: Z-tests on individual parameters from the *full* model. P-values are shown to 2 significant figures or abbreviated to >0.00001. P-values for lower-order (non-interaction terms) are not shown, since they do not test relevant hypotheses (i.e. “does the effect of the variable differ from zero at 0% bare ground cover, 0% host plant cover, and/or 0 hours above 25°C?”).

Appendix 5: outlier analyses

To ensure that our findings were not solely driven by data from sites with unusually high bare ground or host plant cover, we re-fitted all models described in Table 1 of the main paper excluding (a) the site with the highest bare ground cover (n=70 quadrats) and (b) the site with the highest host plant cover (n=212 quadrats). The results are summarised below.

Model	Terms										K	AIC	δ AIC	AICWt
	main effects					interactions								
	quadrat		site-level			bare ground		host plant		temperature				
b_{ij}	h_{ij}	\bar{b}_j	\bar{h}_j	T_j	$b_{ij}\bar{b}_j$	$b_{ij}\bar{h}_j$	$h_{ij}\bar{b}_j$	$h_{ij}\bar{h}_j$	$b_{ij}T_j$					
<i>(a) Site with highest bare ground cover removed</i>														
full	+	+	+	+	+	+	+	+	+	+	12	1064.0	0.0	0.93
inhabitat	+	+	+	+		+	+	+	+		10	1069.2	5.2	0.07
inthost	+	+		+				+	+		8	1081.1	17.1	0.00
intbare	+	+	+			+	+				8	1089.3	25.3	0.00
temp	+	+			+					+	6	1117.2	53.2	0.00
main	+	+									4	1126.8	62.8	0.00
null											2	1331.0	267.0	0.00
<i>(b) Site with highest host plant cover removed</i>														
full	+	+	+	+	+	+	+	+	+	+	12	1082.4	0.0	0.62
inhabitat	+	+	+	+		+	+	+	+		10	1083.4	1.0	0.37
intbare	+	+	+			+	+				8	1091.7	9.3	0.01
inthost	+	+		+				+	+		8	1128.8	46.4	0.00
temp	+	+			+					+	6	1138.7	56.3	0.00
main	+	+									4	1154.0	71.6	0.00
null											2	1376.4	294.0	0.00

Table A3: Empirical support for different models of *H. comma* microhabitat associations. Constituent terms are indicated by plus signs (+) with empty spaces indicating terms that were not included. K=number of parameters; AIC=Akaike's Information Criteria Score; δ AIC=difference in AIC between current and highest-ranked model; AICWt =AIC weights.

Overall, our findings are reasonably robust to the removal of these sites. In both cases, the rank order of the models is the same, with the exception that the *inthost* model (containing the interactions with host plant availability) becomes a better explanation of the data than the *intbare* model (containing the interactions with bare ground availability) when the site with the highest bare ground cover removed. However, both of these models are special cases of the *inhabitat* model, which includes interactions with both bare ground and host plant cover, and continues to rank more than 8 AIC units above both of these simpler models in both analyses; this suggests that both bare ground and host plant availability continue to provide important explanatory power when the outlying sites are removed.

The removal of the highest bare ground (a) and highest host plant cover (b) sites does not affect the conclusion that eggs are associated with lower bare ground in sites with warmer microclimates: in both cases, the *full* model remains the best explanation of the data. However, as might be expected, the removal of datapoints decreases the AIC difference between the *full* and *inhabitat* model in each case ($\delta\text{AIC} = 5.2$ when the site with the highest bare ground cover is removed; $\delta\text{AIC} = 1.0$ when the site with the highest host plant cover is removed)

Appendix 6: model predictions and data for each study site

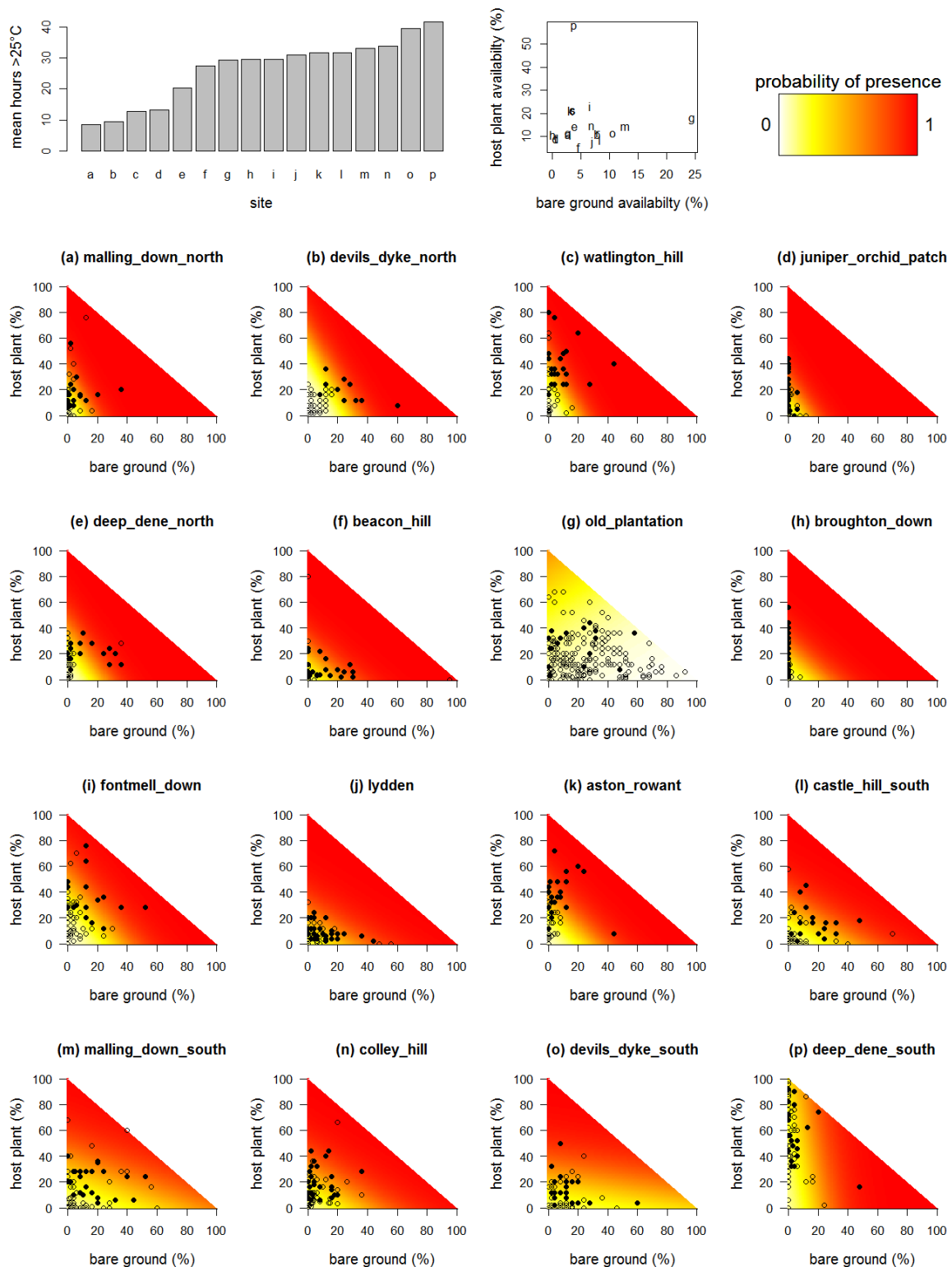


Figure A2: Predictions and data for each of the 16 sites in the study (labelled a-p). Predictions are based on the *full* model. Sites are ordered by temperature, from the least to the most hours above 25°C; the temperatures for each site are shown in the barplot at the top of the figure. The mean

percentage cover of bare ground and host plant cover at each site is shown in the letters scatterplot. Probability of presence predictions range from 0 (white) to 1 (red). Quadrats with eggs present are shown with filled circles, and quadrats with no eggs present are shown as open circles. The axes on individual plots indicate quadrat-level bare ground and host plant (*Festuca ovina*) cover. Note that the scale on the graph showing site-level habitat availability values (top row, middle column) is different to the scale for the quadrat-level bare ground values. The low probability of presence at the Old Plantation site (panel g) results from the small number of presences sampled at that site (n=17), whilst the apparently weak association with bare ground (weak colour gradient moving from left to right across the bare ground axis) results from the high availability of bare ground at this site (see letters scatterplot and Discussion in main text).

Appendix 7: discussion and analysis of habitat availability effects

We used our models to assess the extent to which accounting for differences in habitat availability among sites altered our conclusions regarding the effects of temperature on microhabitat selection. To do this, we compared predictions from the *full* model, which included habitat availability effects, with the *temp* model, which only included the effects of temperature; if the models made similar predictions about the effects of temperature on microhabitat associations, this would indicate that our results would have been similar even if we had not included habitat availability effects. In fact, the strength of relationship between bare ground association and site-level climates was similar in both the *full* and *temp* models (Fig. A3). The lack of discrepancy between the median predictions of these models can be explained by weak statistical association between climate and bare ground or *F. ovina* availability at a site level (Fig. A4), which meant that the “naïve” *temp* model did not mistakenly attribute habitat availability effects to climate effects. However, the *full* model predicts a slightly stronger association with bare ground across all temperatures (Fig. A3), and accounting for habitat availability increased between-population variance in microhabitat associations, resulting in wider population prediction intervals around the median bare ground association (Fig. A3).

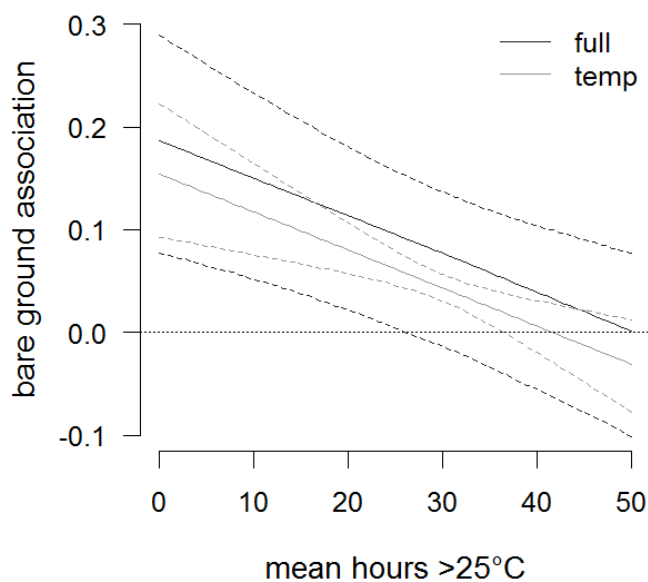


Figure A3: Comparison of predicted changes in bare ground association between the *full* model, which included habitat availability effects, and the *temp* model, which did not. Solid lines show median prediction; dashed lines show 95% population prediction intervals. The median predictions of both models are similar, but including habitat availability effects increases the expected variance in bare ground selection between sites (*full* model has larger credible intervals).

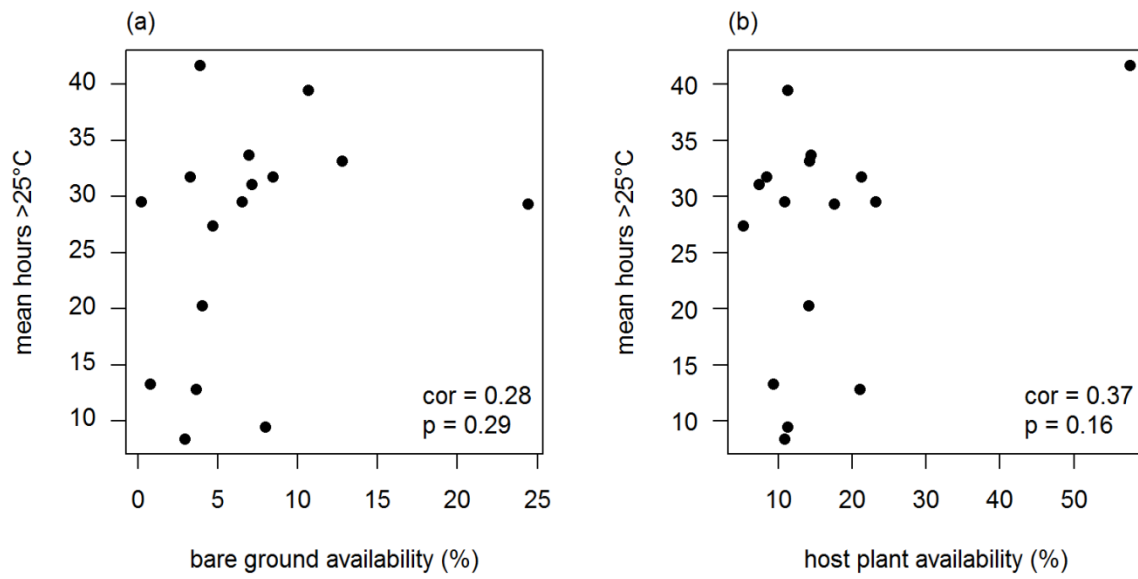


Figure A4: Correlations between habitat availability and climate. Panel (a) shows the correlation between site-level bare ground availability and modelled microclimate; panel (b) shows the correlation between site-level host plant ground availability and modelled microclimate. Results of Pearson’s correlation tests are given in the bottom right of each figure; cor=estimated strength of correlation, p=p-value. Although estimated correlation coefficients are reasonably large in both cases, this is due to outliers (a: 24% bare ground cover at “Old Plantation” site; b: 58% host plant cover at “Deep Dene South” site), and there is no significant correlation between modelled microclimate and either habitat availability variable.