Technical report: making Bioscore distribution models based on Butterfly Monitoring Transects

Technical report: making Bioscore distribution models based on Butterfly Monitoring Transects

Authors and affiliations

Chris van Swaay, De Vlinderstichting/Dutch Butterfly Conservation, Wageningen, Netherlands,

Tom Brereton, Butterfly Conservation, East Lulworth, Dorset, United Kingdom

Bruce Carlisle, Northumbria University, Newcastle-upon-Tyne / Operation Wallacea, UK / Fundatia ADEPT Transilvania, Romania

Claude Dopagne, Natagriwal asbl, Gembloux, Belgium

Ruth Escobés, ZERYNTHIA Society, Spain

Reinart Feldmann, Helmholtz Centre for Environmental Research - UFZ, Halle/Leipzig, Germany

José María Fernández-García, Hazi Foundation, Spain

Benoît Fontaine, Muséum National d'Histoire Naturelle, Paris, France

Ana Gracianteparaluceta, Hazi Foundation, Spain

Alexander Harpke, Helmholtz Centre for Environmental Research - UFZ, Halle/Leipzig, Germany

Janne Heliölä, Finnish Environment Inst., Natural Environment Centre, Helsinki, Finland

Elisabeth Kühn, Helmholtz Centre for Environmental Research - UFZ, Halle/Leipzig, Germany

Andreas Lang, Büro Lang, Germany

Dirk Maes, Research Institute for Nature and Forest (INBO), Brussels, Belgium

Xavier Mestdagh, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg

Yeray Monasterio León, Asociación Española para la Protección de las Mariposas y su Medio (ZERYNTHIA Association), Spain

Miguel López Munguira, Universidad Autónoma de Madrid, Spain

Tomás Murray, National Biodiversity Data Centre, Carriganore, Co. Waterford, Ireland

Martin Musche, Helmholtz Centre for Environmental Research - UFZ, Halle/Leipzig, Germany

Erki Õunap, University of Tartu, Estonia

Lars Pettersson, Swedish Butterfly Monitoring Scheme, University of Lund, Lund, Sweden

Julien Piqueray, Natagriwal asbl, Gembloux, Belgium

Tobias Roth, Hintermann & Weber AG, Reinach, Switzerland

David Roy, NERC Centre for Ecology and Hydrology, Crowmarsh Gifford, United Kingdom

Reto Schmucki, NERC Centre for Ecology and Hydrology, Crowmarsh Gifford, United Kingdom

Josef Settele, Helmholtz Centre for Environmental Research - UFZ, Halle/Leipzig, Germany

Constantí Stefanescu, Butterfly Monitoring Scheme, Museu de Ciències Naturals de Granollers, Spain

Giedrius Švitra, Ukmerge, Lithuania

Anu Tiitsaar, University of Tartu, Estonia

Rudi Verovnik, University of Ljubljana, Ljubljana, Slovenia

Content

Introduction

This technical report describes the collection of data for an assessment of the multivariate relationships between a group of parameters and the distribution of butterflies over Butterfly Monitoring Transects. The results can be used for the development of future scenarios to investigate the effects on biodiversity, and butterflies in particular.

Policy makers need tools to evaluate the effects of policy measures on the environment. Tools for evaluating effects of environmental policies on Europe's biodiversity are scares. In 2005 the Netherlands Environmental Assessment Agency (PBL) developed, together with several institutes, a first set of such tools in BioScore 1 (www.bioscore.eu). Since then PBL has used these tools in scenario studies. However, BioScore 1.0 wasn't fit for all policy related questions and needed extension towards additions pressures and drivers. Furthermore the dose-response functions between environmental pressures and biodiversity in BioScore 1.0 were primarily based on expert judgment.
BioScore 2 extended the models of BioScore 1 with field data and connected them closer to the needs of policy makers in Europe. The model, developed in close cooperation with Alterra, should make it possible to study the effects of future spatial environmental scenarios based on anticipated land-use changes, policies and strategies with environmental impacts, such as green infrastructure strategy, Natura 2000, restoration and rewilding projects, the Common Agricultural Policy, Nitrate Directive (affecting Nitrogen deposition) and Water Framework Directive (affecting water quality and sources of soil pollution).
The present study is an extension to Bioscore 2. It studies the multivariate relationships between a group of parameters and the distribution of butterflies over Butterfly Monitoring Transects.

This technical report only describes the collection of data and a short validation of the outcomes.

Method

Butterfly Monitoring Data

At present over 22 countries in Europe engage in Butterfly Monitoring Schemes (BMS), that are based on repeated visits to fixed transects. All together regular counts are made on more than 5000 transects in Europe (Van Swaay et al., 2016). The method is described and discussed in Van Swaay et al. (2008) (including quality control) and ultimately are based on Pollard (1977).

For this assessment data was gathered from the following BMS's (for co-ordinators check page 4): eBMS (www.butterfly-monitoring.net), which includes the data from Belgium (Flanders), Germany, Spain (outside Catalonia and Basque Country), Catalonia, Finland, France, Ireland, Luxembourg, Netherlands, Sweden and the United Kingdom.
Separate requests for data were made to Slovenia, Switserland, Lituania, Basque Country, Estonia, Belgium (Wallonie) and Romania.

Data was only collected for the period 2000-2015. Overall 5313 sites scattered over Europe have been used (Figure 1), concentrated on the EU and Switzerland.
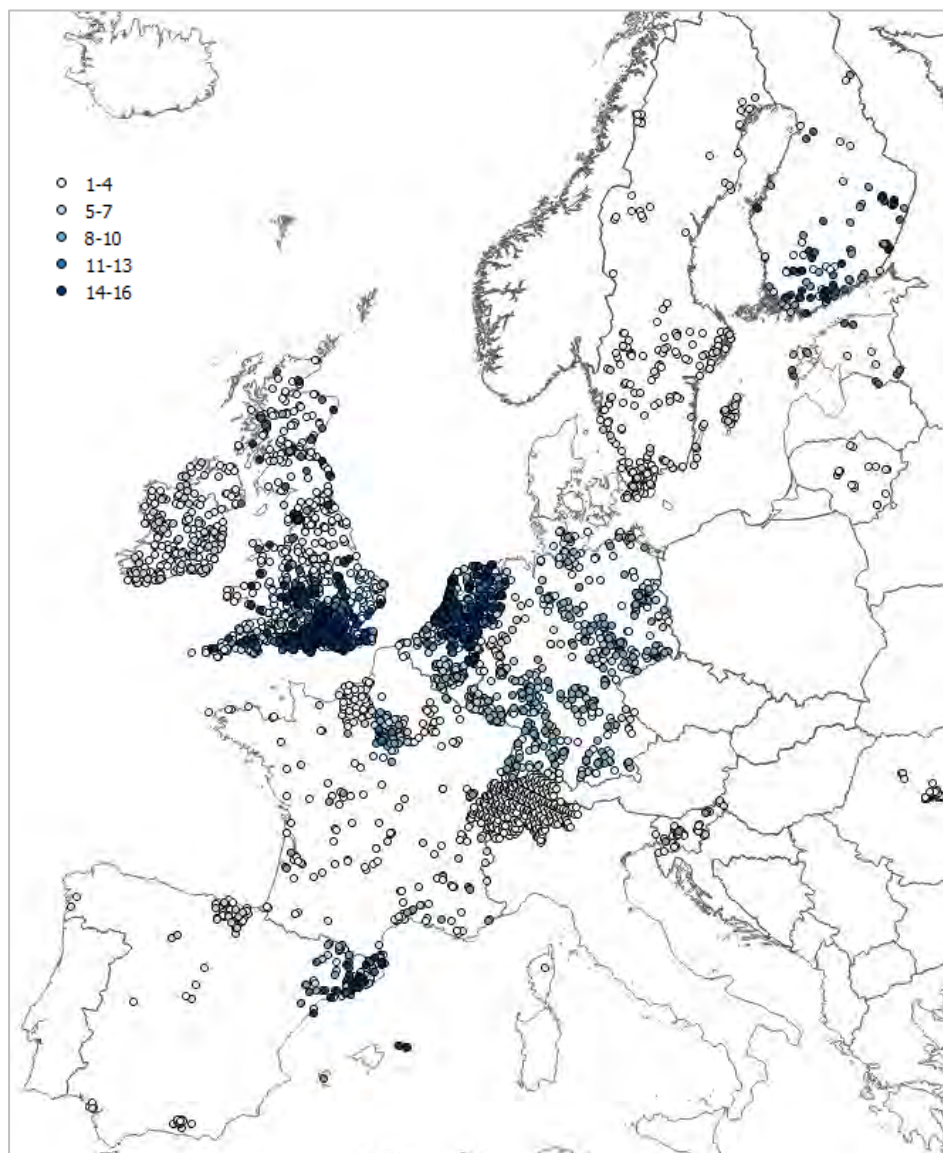

Figure 1: Location of the 5313 Butterfly Monitoring Transects used for this assessment. The colour indicates the number of years between 2000 and 2015 used for this report.

The transects are not evenly distributed across Europe. Most transects are located in the Netherlands and the United Kingdom. Van Swaay et al. (2016) give an overview of the characteristics of the Butterfly Monitoring Schemes on the following points:

- Starting year
- Area represented (w=whole country, r=region)
- Average transect length
- Number of transects per year 2013-2015
- Number of counts on a transect per year
- Counts by volunteers or professionals
- Method to choose sites (free, by co-ordinator, grid or random)
- Representativeness for agricultural grassland
- If nature reserves are overrepresented

The location of each transect has been supplied by the co-ordinators, usually it is the centroid of the sections. As many transects are longer than 1 km, this means that part of the butterflies will have been reported at a distance of several hundreds of meters from the centroid location.
All locations have been transformed to ETRS89 / LAEA Europe - EPSG:3035.

Not all transects have been counted in every year (Table 1). The number of transects in 2015 will have been higher, but not all BMS's have forwarded their data to eBMS at the time of the assessment (October 2017).

Table 1: Number of transects counted per BMS per year as used for this assessment.

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Belgium-Flanders	13	13	12	13	16	19	18	17	21	22	21	7	7	9	8	
Belgium-Wallonie							8	25		53	40	36	36	35	40	38
Switzerland				85	93	97	94	90	93	96	94	88	94	96	97	94
Germany							300	360	365	397	358	398	369	379	339	334
Spain (excl. Basque Country and Catalonia)												2	18	19	35	29
Basque Country											13	21	25	21	28	34
Catalonia	31	42	40	45	51	51	65	69	69	65	69	67	66	63	64	63
Estonia					6	8	8	10	10	10	10	10	11	11	13	13
Finland	39	49	49	45	41	51	63	64	64	63	70	68	67	65	63	44
France						13	90	95	96	118	122	96	140	148	158	128
Ireland											138	135	139	126	118	
Lithuania										14	9	4	2			
Luxembourg											30	30	29	29	30	32
Netherlands	288	292	299	417	425	443	441	422	417	417	428	451	438	487	477	425
Romania_1														10	10	10
Romania_2															92	93
Sweden											61	100	133	159	173	
Slovenia								21	16	12	10	8	12	13	14	9
United Kingdom	354	342	406	466	444	457	517	514	524	551	548	526	494	1082	1221	1202
Total	**725**	**738**	**806**	**1071**	**1076**	**1139**	**1604**	**1687**	**1675**	**1818**	**2021**	**2047**	**2080**	**2752**	**2980**	**2548**

Species

All species registered on the 5313 transects between 1990 and 2015 have been assessed based on the taxonomy of Van Swaay et al. (2010), the same as in the Fauna Europaea in November 2017.
In the eBMS some countries use aggregate species. These are species which can be hard to identify in the field. In this analysis we have combined the data from all countries into the aggregate species over all BMS's following Table 2.

Table 2: Aggregated species.

Species	Name of aggregate
Argynnis niobe/aglaja/adippe	*Argynnis_compl*
Thymelicus lineola/sylvestris	*Thymelicus_compl*
Pieris napi/rapae	*Pieris_compl*
Colias hyale/alfacariensis	*Colias_compl*
Leptidea sinapis aggr.	*Leptidea_compl*
Melitaea aurelia/athalia/britomartis/parthenoides	*Melitaea_compl*
Plebeius argus/argyrognomon/idas	*Plebejus_compl*
Aricia agestis/artaxerxes	*Aricia_compl*
Pontia daplidice/edusa	*Pontia_compl*
Pyrgus malvae/malvoides	*Pyrgus_compl*

Apart for these ten aggregate species, data was available for 258 species. However 56 species of these were only observed on less than ten transects. The ten species found on most transects are listed in Table 3.

Table 3: Top ten of the species found on most transects.

Species	Number of transects
Pieris_compl	4596
Pieris brassicae	4531
Maniola jurtina	4485
Vanessa atalanta	4457
Aglais urticae	4426
Aglais io	4395
Polyommatus icarus	4280
Vanessa cardui	3969
Pararge aegeria	3926
Anthocharis cardamines	3864

Selection of transects

Transects with missing values for one or more environmental variables were excluded. In order to limit spatial bias and pseudo-replication, we then randomly selected for each species per 1 km grid cell only one transect of one year where the species was present.

Environmental variables

The selected environmental variables cover climate, soil, land cover and nitrogen deposition (Table 4Table 4: Environmental variables included in the model, including the source of the data and the resolution of the data.**Fout! Verwijzingsbron niet gevonden.**). The climatic variables cover various aspects of continentality and seasonality, represented by a set of bioclimatic variables: annual mean temperature (bio1), temperature seasonality (bio4), minimum temperature of the coldest month (bio6), annual precipitation (bio12), and last precipitation of the warmest and coldest quarter (bio18 and bio19). These variables were retrieved from the CHELSA data set by Karger et al. (2017), based on the monthly mean, minimum and maximum values averaged for the time period of 1979-2013. Soil variables included organic carbon, silt, sand and clay content of the top soil, bulk density, volume of coarse fragments, pH and cation exchange capacity of the top soil (Hengl et al., 2017). Land cover was represented by the non-urban land cover types (classification level 2) of the Corine land cover map for the year 2000 (Hazeu et al., 2008). The land cover class of permanent crops, where less than 4000 vegetation plots were recorded, was aggregated with arable land (Table 4). Nitrogen deposition in the year 2013 was added as a proxy for nitrogen input in the soil (Fagerli et al., 2015).

Table 4: Environmental variables included in the model, including the source of the data and the resolution of the data.

Variables	Data source	resolution	Reference
Winter precipitation (bio18)	Chelsa	30 arc sec (~1^2 km)	(Karger *et al.*, 2017)
Summer precipitation (bio19)	Chelsa	30 arc sec(~1 km)	(Karger *et al.*, 2017)
mean minimum temperature of the coldest month (bio6)	Chelsa	30 arc sec (~1 km)	(Karger *et al.*, 2017)
Mean annual temperature (bio1)	Chelsa	30 arc sec (~1 km)	(Karger *et al.*, 2017)
Annual precipitation (bio12)	Chelsa	30 arc sec (~1 km)	(Karger *et al.*, 2017)
Temperature seasonality (bio4)	Chelsa	30 arc sec (~1 km)	(Karger *et al.*, 2017)
Arable land and permanent crops (CLC-codes 211 - 223)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Pastures (CLC-code 231)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Heterogeneous agricultural areas (CLC-code 241 - 244)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Scrub and/or herbaceous vegetation associations (CLC-code 321 - 324)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Forests (CLC-code 311 - 313)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Open spaces with little or no vegetation (CLC-code 331 - 335)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
Inland wetlands (CLC-code 411 - 412)	Corine land cover	100 meter	(Hazeu *et al.*, 2008)
organic carbon content in the top soil	Soil Grids	1 km	(Hengl *et al.*, 2017)
clay content in the top soil	Soil Grids	1 km	(Hengl *et al.*, 2017)
silt content in the top soil	Soil Grids	1 km	(Hengl *et al.*, 2017)
Sand content in the top soil	Soil Grids	1 km	(Hengl *et al.*, 2017)
Bulk density (fine earth) in kg / cubic-meter	Soil Grids	1 km	(Hengl *et al.*, 2017)
Coarse fragments volumetric in %	Soil Grids	1 km	(Hengl *et al.*, 2017)
Cation exchange capacity of soil in cmolc/kg	Soil Grids	1 km	(Hengl *et al.*, 2017)
Soil pH x 10 in H2O	Soil Grids	1 km	(Hengl *et al.*, 2017)
Nitrogen deposition	EMEP	0.1 degrees (~11 km)	(Fagerli *et al.*, 2015)

All environmental variables were resampled to a 1 km resolution using the mean value. For the land cover variables we employed the fraction of each type within the 1 km grid cell. To decrease correlation between the variables, only variables with a variance inflation factor (VIF) below 10 were selected for inclusion in the SDMs (Zuur et al., 2009). This eliminated sand content in the top soil, mean minimum temperature of the coldest month (bio6) and annual precipitation (bio12).

Fitting the SDMs
Absences were randomly selected, such that for each species there was not more than one transect per 1 km grid cell where the species was absent, conform the selection procedure of the presence values. Subsequently, two sub-sets of absences were compiled, one with all randomly selected absence records and one with 1,000 or an number equal to the number of presences when this is greater than 1,000. The SDMs were fitted with the BIOMOD2 package (version 3.3-7; Thuiller et al., 2016), using default settings. A GLM, GAM and BRT were applied to fit a model. These frequently used algorithms (e.g. Araújo et al., 2011; Thuiller et al., 2014; Dullinger et al., 2017) range from simplistic to complex (Merow et al., 2014). The GLM and GAM were fitted with the set containing all randomly selected absence records and the BRT were fitted with the other set (Barbet-Massin, 2012). An ensemble model was built with each of the three algorithms weighted based on the TSS values of the fitted models (TSS : True kill statistic (Hanssen and Kuipers discriminant, Peirce's skill score)). Algorithms with a TSS < 0.5 were excluded. The fitted models were used to project distribution maps for Europe with the probability of occurrence per algorithm of each species. For this purpose the environmental variables were aggregated to a 5km resolution, taking the mean value, in order to reduce calculation time. This resulted in one distribution map with probability of occurrence of each species. This map was transformed into a binary presence-absence map by using a threshold which maximizes the TSS (Liu et al., 2013).

Validation
A validation of the results was performed for the 100 species for which atlas distribution data was available from Bioscore 2 (Van Swaay et al., 2014), originating from the LepiDiv database (UFZ, Leipzig-Halle), partly based on the 'Distribution Atlas of Butterflies in Europe' (Kudrna et al., 2011). To this end the binary presence-absence maps were aggregated to the same resolution as the atlas distribution data (50km) by taking the maximum value within the 50km grid cell. The overlap and differences between the aggregated presence-absence maps and the atlas distribution maps was assessed for each species.

Results

For 71 species the distribution as resulting from the models could be compared with atlas data from the LepiDiv database.

The results of the models in this report will be first visualised by an example species *Anthocharis euphenoides* (Figure 2).

Figure 2: Female (left) and male of Anthocharis euphenoides*.*

Occurrence on the transects

Figure 3 shows the distribution of the 85 transects where *Anthocharis euphenoides* was reported in at least one year. The species is restricted to the Iberian Peninsula and SE France.
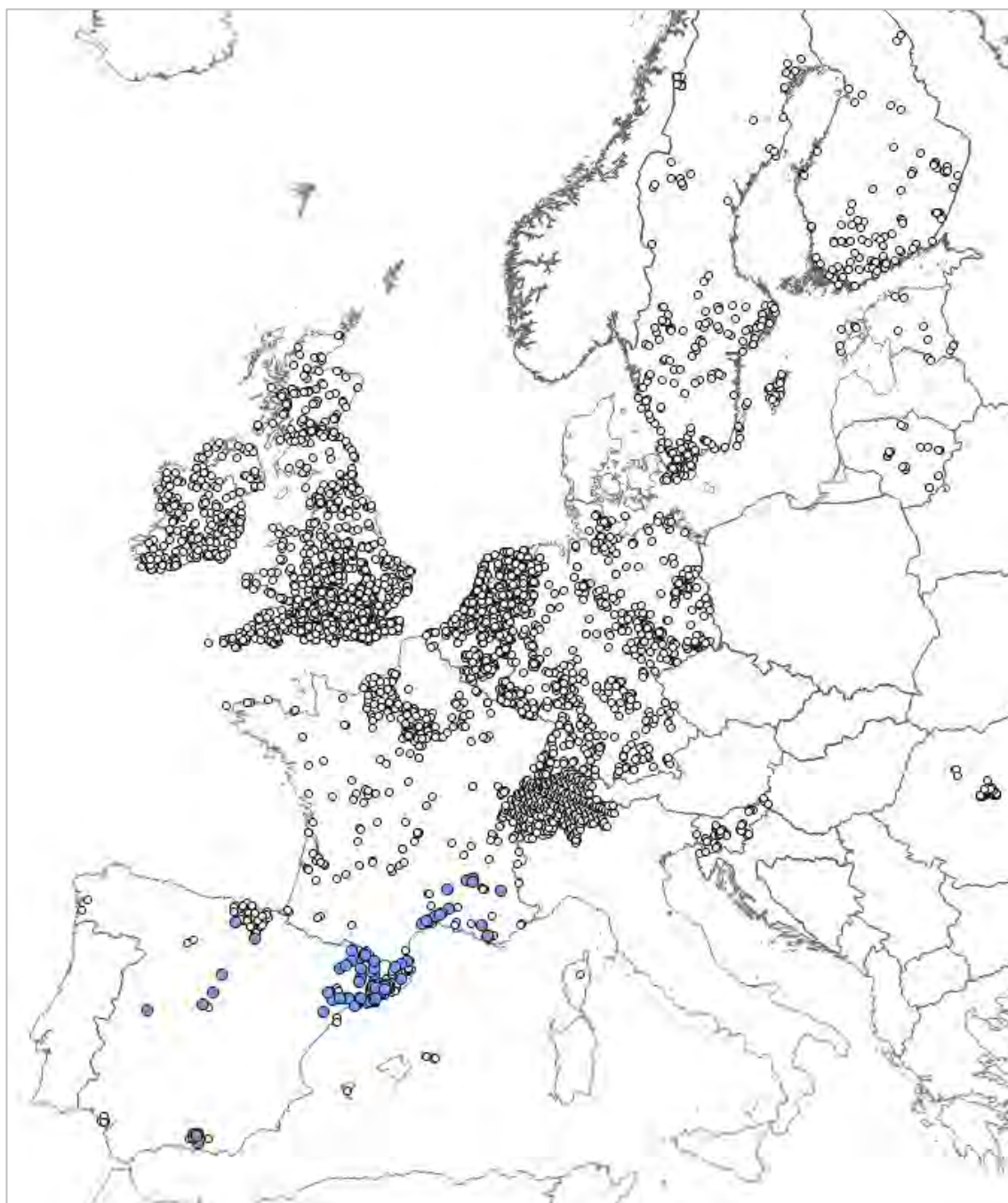

Figure 3: Location of all transects where Anthocharis euphenoides *has been reported between 2000 and 2015.*

Model output

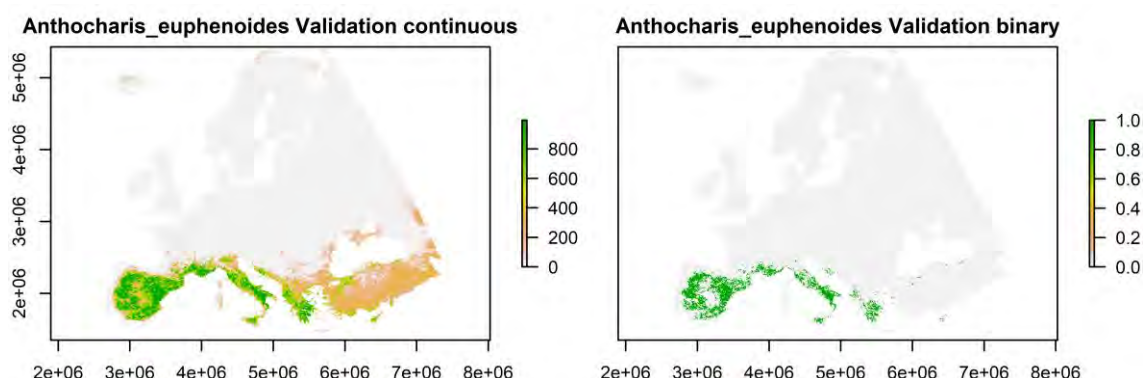Figure 4 shows the result of the model.

Figure 4: Model output for Anthocharis euphenoides*. Left: continuous scale, where dark green areas have the highest occurrence. Right: Binary map after application of cut-off value.*

Errors

In this brief validation we will check for the two main types of errors:

- a type I error (a "false positive" finding), and
- a type II error (a "false negative" finding).

The maps with the atlas data and the result of the Binary map from the model (as presented in Figure 4) are presented in Figure 5. All data is transformed to ETRS89 / LAEA Europe - EPSG:3035 50km squares.
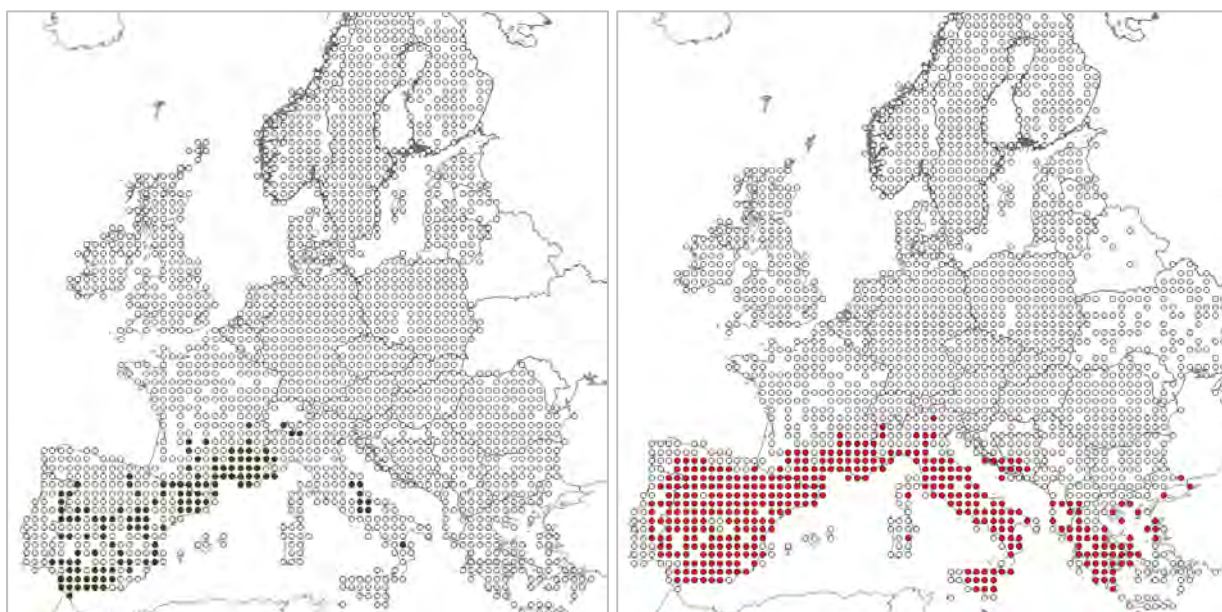
Figure 5: Distribution of Anthocharis euphenoides *from atlas data (left) and as a result of the binary map from the Bioscore model (right).*

From the 133 squares where the species is reported in the atlas data, the model finds most of them (116 squares, 87%), only 17 are missed. However the model reports the species on 388 squares, and on 272 of them (70%) the species is missing in the atlas data (error type I: false positives).

The error I type (false positives) for this species is high with 70% false positives. For most other species this is lower (Table 5). The mean of false positives is 49%, the median value is 48%.

The error II type (false negatives) is very small for *Anthocharis euphenoides*: only 17 out of 1872 negative squares were false negatives (1%). This is low compared to the other species (Table 6). The mean of false negatives is 11%, the median value is 9%.

Table 5: Number of false positive squares (error I type) per species.

Species	false positive squares	total number of squares	Error type I: Percentage false positives
Plebejus idas	23	30	77
Polyommatus escheri	375	502	75
Heteropterus morpheus	950	1286	74
Anthocharis euphenoides	272	388	70
Arethusana arethusa	502	730	69
Carterocephalus silvicolus	519	765	68
Euphydryas aurinia	974	1483	66
Euphydryas maturna	138	216	64
Satyrium pruni	830	1314	63
Phengaris arion	752	1223	61
Cupido argiades	902	1471	61
Phengaris alcon	419	687	61
Brenthis daphne	624	1031	61
Phengaris teleius	277	466	59
Polyommatus amandus	969	1662	58
Spialia sertorius	657	1127	58
Melitaea diamina	564	980	58
Aricia eumedon	560	999	56
Apatura ilia	750	1347	56
Melitaea phoebe	815	1486	55
Apatura iris	832	1520	55
Carterocephalus palaemon	758	1388	55
Libythea celtis	156	288	54
Satyrium spini	601	1124	53
Hamearis lucina	557	1049	53
Limenitis camilla	580	1110	52
Coenonympha glycerion	733	1407	52
Coenonympha tullia	616	1189	52
Melitaea cinxia	803	1577	51
Brintesia circe	515	1030	50
Melitaea didyma	760	1520	50
Thecla betulae	747	1505	50
Glaucopsyche alexis	738	1489	50
Satyrium ilicis	564	1138	50
Lycaena dispar	516	1055	49
Hesperia comma	823	1703	48
Charaxes jasius	163	344	47
Hipparchia semele	740	1569	47
Polyommatus coridon	491	1052	47
Cupido minimus	635	1372	46
Aporia crataegi	855	1851	46
Lycaena hippothoe	592	1284	46
Lampides boeticus	404	884	46
Brenthis ino	681	1500	45
Boloria dia	542	1214	45
Limenitis reducta	334	764	44
Polyommatus bellargus	554	1269	44
Lycaena tityrus	615	1415	43
Plebejus optilete	310	729	43
Thymelicus acteon	418	983	43
Coenonympha arcania	680	1602	42
Hipparchia statilinus	273	646	42

Species	false positive squares	total number of squares	Error type I: Percentage false positives
Cyaniris semiargus	691	1738	40
Nymphalis antiopa	606	1564	39
Carcharodus alceae	447	1161	39
Erynnis tages	574	1504	38
Lasiommata maera	630	1652	38
Lycaena virgaureae	475	1298	37
Erebia ligea	331	911	36
Boloria euphrosyne	549	1533	36
Argynnis paphia	642	1911	34
Pyronia tithonus	252	756	33
Phengaris nausithous	53	160	33
Iphiclides podalirius	435	1318	33
Boloria aquilonaris	153	469	33
Boloria selene	500	1544	32
Gonepteryx cleopatra	187	586	32
Araschnia levana	331	1108	30
Melanargia galathea	360	1293	28
Issoria lathonia	415	1520	27
Papilio machaon	398	1867	21
Grand total	**38487**	**81656**	**47**

Table 6: Number of false negative squares (error II type) per species.

Species	false negative squares	Total number of squares	Error type II: Percentage false negatives
Plebejus idas	945	2230	42
Papilio machaon	138	393	35
Issoria lathonia	238	740	32
Cupido minimus	271	888	31
Erynnis tages	225	756	30
Boloria selene	197	716	28
Nymphalis antiopa	191	696	27
Cyaniris semiargus	137	522	26
Hipparchia semele	174	691	25
Aporia crataegi	97	409	24
Euphydryas aurinia	180	777	23
Boloria euphrosyne	163	727	22
Lasiommata maera	132	608	22
Carcharodus alceae	176	1099	16
Polyommatus coridon	192	1208	16
Satyrium ilicis	156	1122	14
Melitaea cinxia	89	683	13
Phengaris arion	134	1037	13
Limenitis camilla	147	1150	13
Hesperia comma	70	557	13
Pyronia tithonus	181	1504	12
Argynnis paphia	40	349	11
Boloria aquilonaris	201	1791	11
Lycaena virgaureae	102	962	11
Lycaena tityrus	89	845	11
Thymelicus acteon	134	1277	10
Satyrium spini	119	1136	10
Iphiclides podalirius	97	942	10
Polyommatus bellargus	99	991	10

Species	false negative squares	Total number of squares	Error type II: Percentage false negatives
Melitaea diamina	127	1280	10
Apatura ilia	90	913	10
Brenthis ino	74	760	10
Phengaris alcon	151	1573	10
Lycaena hippothoe	92	976	9
Boloria dia	98	1046	9
Thecla betulae	68	755	9
Melanargia galathea	86	967	9
Aricia eumedon	112	1261	9
Hipparchia statilinus	143	1614	9
Carterocephalus palaemon	74	872	8
Hamearis lucina	97	1211	8
Glaucopsyche alexis	61	771	8
Plebejus optilete	121	1531	8
Apatura iris	58	740	8
Phengaris nausithous	149	2100	7
Euphydryas maturna	142	2044	7
Libythea celtis	136	1972	7
Phengaris teleius	111	1794	6
Coenonympha tullia	65	1071	6
Melitaea didyma	44	740	6
Coenonympha glycerion	50	853	6
Polyommatus amandus	33	598	6
Limenitis reducta	80	1496	5
Coenonympha arcania	34	658	5
Arethusana arethusa	78	1530	5
Satyrium pruni	48	946	5
Erebia ligea	63	1349	5
Carterocephalus silvicolus	68	1495	5
Melitaea phoebe	35	774	5
Brintesia circe	48	1230	4
Araschnia levana	44	1152	4
Spialia sertorius	42	1133	4
Brenthis daphne	45	1229	4
Lycaena dispar	40	1205	3
Heteropterus morpheus	30	974	3
Polyommatus escheri	54	1758	3
Lampides boeticus	36	1376	3
Charaxes jasius	50	1916	3
Gonepteryx cleopatra	31	1674	2
Cupido argiades	13	789	2
Anthocharis euphenoides	17	1872	1
Grand total	8152	78804	10

In some regions the number of correct predictions from the model is much higher than in other regions (Figure 6). The model predicts especially well in the regions with long running BMS's scattered in many habitats, like the United Kingdom, Finland and Slovenia. In Spain outside Basque country and Catalonia, the model predictions are quite bad. Probably this is caused by most of the transects being in National parks, as e.g. the Sierra Nevada. This makes the models not functioning very well outside those parks. Furthermore the model performs especially bad on the Balkans, where there are almost no transects (Figure 1).
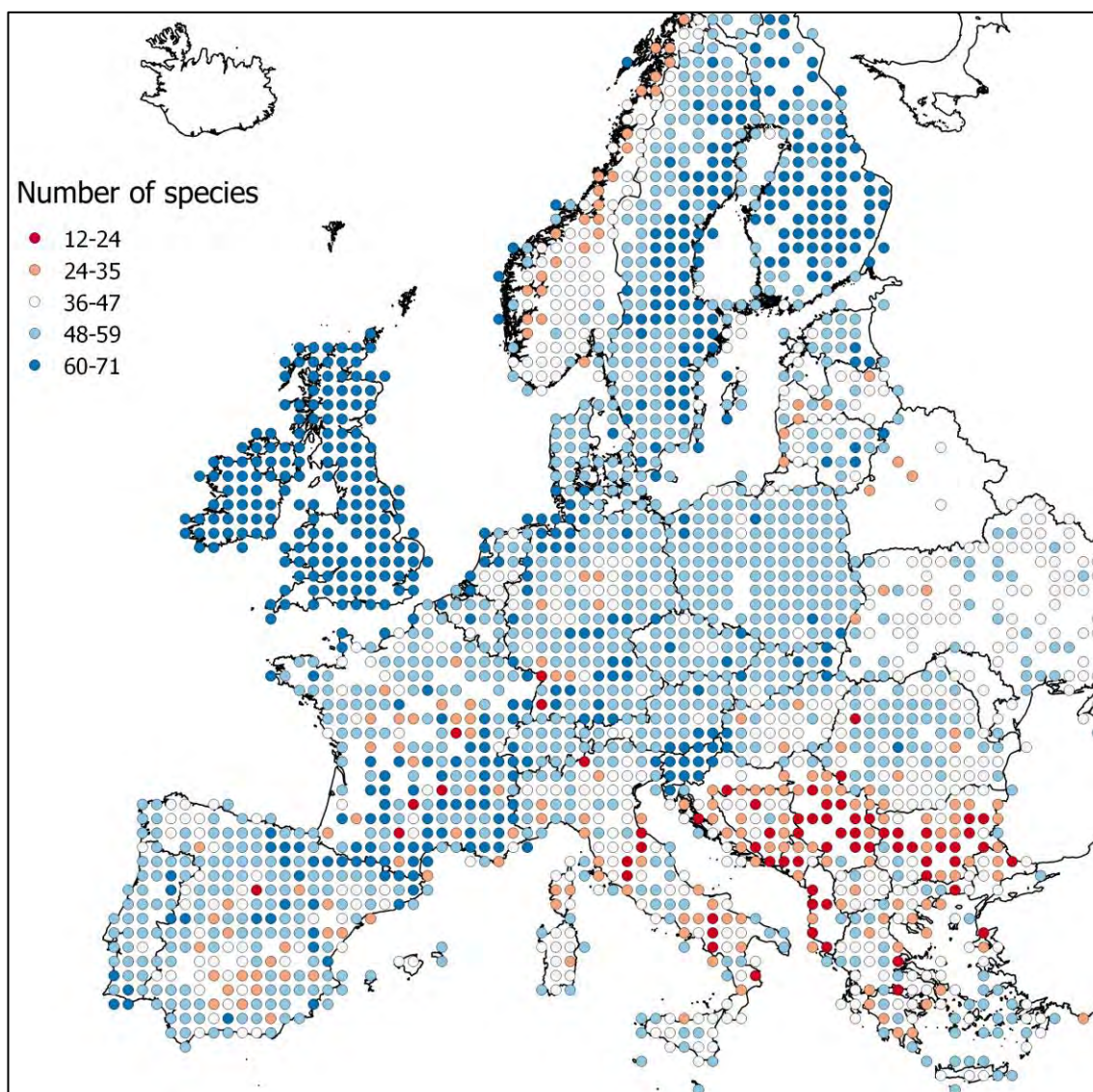

Figure 6: Number of species for which the Bioscore model predicted either a correct presence or a correct absence.

Conclusions

- This projects shows that the results from Butterfly Monitoring Schemes (BMS's) can be used to extrapolate to distribution maps on a European scale.
- The mean of error I type (false positives) is 49%, the median value is 48%, with a range from 21 to 77%. This means the models are in general too optimistic, especially for relatively rare and localized species. It should be noted however that the positive squares were upgraded from positive 5 km squares. Only one positive 5 km square was enough to make the whole 50 km square positive. Butterflies in general need enough habitat and often a metapopulation. The error I type might reduce, if a threshold would be set for the number of positive 5 km squares needed to make a 50 km square positive.
- The mean of error II type (false negatives) is 11%, the median value is 9%, with a range from 1 to 42%.
- The model predicts especially well in the regions with long running BMS's scattered in many habitats, like the United Kingdom, Finland and Slovenia.
- The model predictions are not very good in areas without transects, as in the Balkans and in Italy. Norway does have a monitoring scheme, which was not included in this analysis. This might improve the results in that country considerably.
- Doing the same validation analysis with the results of Bioscore 2 (models based on atlas data; Van Swaay et al., 2014) will show which of the two approaches is better.

Literature

Araújo, M.B., Alagador, D., Cabeza, M., Nogués-Bravo, D. & Thuiller, W. (2011) Climate change threatens European conservation areas. Ecology Letters, **14**, 484-492.

Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods in Ecology and Evolution, 3, 327-338.

Dullinger, I., Wessely, J., Bossdorf, O., Dawson, W., Essl, F., Gattringer, A., Klonner, G., Kreft, H., Kuttner, M., Moser, D., Pergl, J., Pyšek, P., Thuiller, W., van Kleunen, M., Weigelt, P., Winter, M. & Dullinger, S. (2017) Climate change will increase the naturalization risk from garden plants in Europe. Global Ecology and Biogeography, **26**, 43-53.

EEA (2014) EUNIS habitat classification. Available at: https://www.eea.europa.eu/themes/biodiversity/eunis/eunis-habitat-classification (accessed 2017).

Fagerli, H., Tsyro, S., Simpson, D., Schulz, M., M., G., Jonson, J.E., Benedictow, A., Wind, P., Nyíri, Á., Steensen, B.M., Valiyaveetil, S., Aas, W., Hjellbrekke, A., Solberg, S., Stebel, K., T., H., Tørseth, K., Yttri, K.E., Mareckova, K., Wankmüller, R., Pinterits, M., Ullrich, B., Posch, M., Denier van der Gon, H., Alastuey, A. & Theys, N. (2015) Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. In. Norwegian Meteorological Institute

Hazeu, G.W., Mücher, C.A., Kienast, F. & Zimmerman, N.E. (2008) Land cover maps for environmental modeling at multiple scales. In. EU Funded Project FP6 2006 GOCE 036866 for the ECOCHANGE project, Wageningen.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A. & al., e. (2017) SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE, **12**, e0169748.

Karger, D.N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R.W., Zimmermann, N.E., Linder, H.P. & Kessler, M. (2017) Climatologies at high resolution for the earth's land surface areas. **4**, 170122.

Kudrna, O.; Harpke, A.; Lux, K.; Pennerstorfer, J.; Schweiger, O; Settele, J.; Wiemers, M. (2011): Distribution atlas of butterflies in Europe. Gesellschaft für Schmetterlingschutz, Halle, Germany

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. Journal of Biogeography, 40, 778–789.

Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E. & Elith, J. (2014) What do we gain from simplicity versus complexity in species distribution models? Ecography, **37**, 1267-1281.

Pollard, E. (1977): A method for assessing changes in the abundance of butterflies. Biological Conservation 12 (2), 115-134

Thuiller, W., Georges, D., Engler, R., Breiner, F., Georges, D. & Thuiller, W. (2016) Package 'biomod2'.

Thuiller, W., Guéguen, M., Georges, D., Bonet, R., Chalmandrier, L., Garraud, L., Renaud, J., Roquet, C., Van Es, J., Zimmermann, N.E. & Lavergne, S. (2014) Are different facets of plant diversity well protected against climate and land cover changes? A test study in the French Alps. Ecography, **37**, 1254-1266.

Swaay, C.A.M. van, Nowicki, P., Settele, J. & Strien, A.J. van (2008): Butterfly monitoring in Europe : methods, applications and perspectives. Biodiversity and Conservation **17 (14)**, 3455-3469

Swaay, C. van, Cuttelod, A., Collins, S., Maes, D., López Munguira, M., Šašic, M., Settele, J., Verovnik, R., Verstrael, T., Warren, M., Wiemers, M., Wynhoff, I. (compilers) (2010): European Red List of butterflies. IUCN Red List of Threatened Species - Regional Assessment - Office for Official Publications of the European Communities, Luxembourg

Van Swaay, C.A.M., Van Strien, A.J., Aghababyan, K., Åström, S., Botham, M., Brereton, T., Carlisle, B., Chambers, P., Collins, S.,Dopagne, C., Escobés, R., Feldmann, R., Fernández-García, J.M., Fontaine, B., Goloshchapova, S., Gracianteparaluceta, A., Harpke, A., Heliölä, J., Khanamirian, G., Komac, B., Kühn, E., Lang, A., Leopold, P., Maes, D., Mestdagh, X., Monasterio, Y., Munguira, M.L., Murray, T., Musche, M., Õunap, E., Pettersson, L.B., Piqueray, J., Popoff, S., Prokofev, I., Roth, T., Roy, D., Schmucki, R., Settele, J., Stefanescu, C., Švitra, G., Teixeira, S.M., Tiitsaar, A.,

Verovnik, R., Warren, M.S. (2016): The European Butterfly Indicator for Grassland species 1990-2015. Report VS2016.19, De Vlinderstichting. Wageningen

Van Swaay, C.A.M., Schweiger, O., Settele, J., Kühn, E., Harpke, A., Wiemers, M., Musche, M., Pe'er, G., Stefanescu, C., Fontaine, B., Julliard, R., Roy, D., Brereton, T., Petterson, L.B., Kuussaar, M., Heliola, J., Schmucki, R. (2014): Bioscore 2 : Potential distribution of butterflies and assessment of environmental pressures in Europe. Rapport VS2014.03 - De Vlinderstichting, Wageningen, Butterfly Conservation Europe

Zuur, A., Ieno, E.N. & Walker, N. (2009) Mixed Effects Models and Extensions in Ecology with R. Springer New York.