CAN WE IMPROVE PREDICTIONS OF SPECIES DISTRIBUTIONS THROUGH FINER SPECIES-CLIMATE TEMPORAL MATCHING?

Nicolas Bonzon (nicolas.bonzon@unil.ch)

1st step project, directed by Antoine Guisan supervised by Olivier Brönnimann

Master in Behavior, Evolution and Conservation, University of Lausanne, 2016

In order to implement a fine species-climate temporal matching approach for species distribution models (SDM) and investigate whether this increases the accuracy of predictions, we took daily temperature and precipitation dataset (1981-2015) recently developed at a resolution of 25m for Switzerland (Broennimann et al. in prep). We used these climate data to create SDM of several species of butterflies. We then compared these new models with conventional models using 30-years averaged climate data (1980 to 2010). In order to compare these two types of modelling, we retained the TSS and ROC scores attesting to the precision of the models. Our results show no significant difference depending on the methods used. We believe that this is mainly due to the biases caused by other variables such as the distribution of host plants, strongly linked to climatic conditions and influencing the distribution of butterflies.

INTRODUCTION

In the field of ecology, the analysis of the relationships between species and their environment is a crucial point. The quantification of these relationships allows the creation of species distribution models (SDMs). SDMs work according to the following scheme: several environmental variables explain a response variable: the presence or absence of a species. The use of SDMs also allows to predict the probability of species presence in a new environment. These projections can be done in space or time. (Guisan & Zimmermann, 2000; Schröder, 2008; Elith & Leathwick 2009). In order to use correctly these models in real cases, it is important to be able to evaluate their accuracy. This can be done by comparing these models with existing surveys and then counting the number of times when the model correctly predicts the presence or absence of the species. Based on these results, it is possible to evaluate the accuracy of a model (Fielding & Bell, 1997; Allouche *et al.*, 2006; Liu *et al*., 2009).

The SDM is a very useful tool in the field of conservation. Having an accurate model is important for taking decisions in different stages of a conservation project. Examples include measuring and predicting the impact of an invasive species, defining critical habitats for a threatened species or predicting the shift of a species based on anthropogenic changes

(Guisan *et al.*, 2013). These tools are especially important in the context of global warming, where many species are forced to change habitats or disappear (Chen *et al.*, 2011).

Most of SDMs use environmental variables that are averaged over a long period of time, typically 30 years (Elith *et al.*, 2006). This approach is valid when using environmental variables experiencing little change over time, such as substrate or topography. However, other variables, such as climate, evolve more rapidly. By using these long-term averaged climatic data to explain the distribution of an annual species (such as butterflies), they do not necessarily relate to conditions that organisms actually experience through their lifetimes and that are crucial for their survival and reproduction.

The increased precision of available climatic data allows a new modelling approach with a finer temporal matching between the environmental variables and the observation of a species. This makes it possible to use the real climatic conditions experienced by the organism during the year. Currently, only few studies have used this approach (Reside *et al.*, 2010; Bateman *et al*., 2012; Wernberg *et al.*, 2013; Wisz *et al.,* 2015).

The aim of this study is to compare these two approaches, namely the use of 30-year averages against the use of annual data. We created SDMs for several butterfly's species using climatic variables averaged between 1981 and 2010 and evaluated their accuracy. We also created and evaluated these same SDMs but using daily temperature and precipitation dataset (1981-2015) recently developed at a resolution of 25m for Switzerland (Broennimann *et al*., in prep). In order to compare both kind of modelling, we evaluated their differences based on the obtained results. Moreover, to better describe the possible uses of these more precise data, we used two levels of resolution: i) An annual level for which we used the variables on the whole year or butterflies were observed. ii) A monthly level for which we selected, for each variable, the most crucial months for the survival and reproduction of the species.

We hypothesize that the use of these finer temporal matching variables produces more accurate SDMs than those calculating with 30-years averages variables.

METHODS

STUDY SPECIES

We tested our hypothesis by modelling the distribution of Lepidoptera (butterflies). The main reasons for this choice are their high sensitivity to climate and their rapid responses to climatic variations. Indeed, the fitness of the butterflies is directly influenced by weather mean and extremes (Buckley & Kingsolver, 2012; Karl *et al.*, 2011). Their dispersal capacity (therefore their presence probability at one point) is also influenced by annual weather conditions (Cormont *et al.*, 2010; Kuussaari *et al.*, 2016). It should be noted, however, that butterflies have patterns of lives that vary enormously between species and within a single species depending on the region. That is why we have selected species of butterflies with the same

pattern of life, namely: a single generation per year, no migration and a passage of winter in the form of egg or caterpillar.

Study area

The fields samplings data were collected by the Spatial Ecology Group from University of Lausanne? in the Western Swiss Alps at an elevation from around 650 to 3210m. This area includes five vegetation belt: a foothill belt with deciduous forest, a mountainous belt with mixed forests, a subalpine belt with coniferous forest, an alpine belt with heath, meadow and grassland vegetation and a nival belt with sparse vegetation. This area possesses also very varied environments in terms of slope and exposure as well as the human land use. For more information of the study area and the field sampling method see Pellissier *et al.* (2012).

Climatic variables

Primary climate data are provided by Météo Suisse **(citation?)** and more precise data were derived and/or extrapolated (Broennimann *et al*., in prep). We used average, minimum and maximum temperatures in degrees Celsius (°C), precipitation in mm (or l/m^2), evapotranspiration as well as the growing degree day above 10 ° C.

Modelling framework

We made two comparisons:

The first comparison is between the use of one-year averaged climatic variables versus the use of 30-years averaged climatic variables (Figure 1). The assumption is that with one-year averaged climate variables, the model will be more accurate than with 30-years averaged climatic variables. This comparison helps to understand whether the year in which the species is observed has a greater influence on its distribution than the average of the last 30 years. For this comparison, we modelled the distribution28 species.

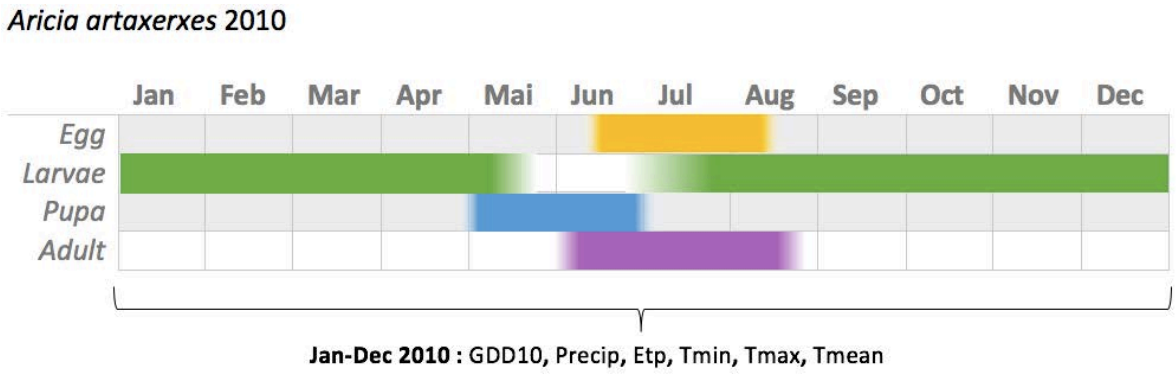


Figure 1: Diagram of life stages of *Aricia artaxerxes* butterfly. For this one-year modelling method, the set of variables is applied throughout the year (compare with figure 2).

The second comparison is between the use of few-month averaged climatic variables according to the periods considered crucial for the survival or reproduction of the species versus the use of the same few-month but averaged on the last 30-years (Figure 2). The assumption is that with variables selected over a few months, the model will be more accurate than the one using the 30-years average data. However, the risk is that by restricting too much the temporal window of these variables according to theoretical data, they no longer correspond to the key periods of the species. For this comparison, we modelled the distribution of 13 species. The key periods for each species were chosen based on observation data in Switzerland available on the pieris.ch website. The choice of explanatory variables is based on the general physiology of butterfly (Jones *et al.*, 1982; Feltwell, 1983; Fischer *et al.*, 2003; Geister *et al.*, 2008).

As illustrated in Figure 2, we used the minimum temperature during the months of January to April, when excessive frosts can destroy the larvae (or eggs). Then we used growing degree days above 10°C (GGD10), average temperature (Tmean) and precipitations (Precip) from the beginning of the caterpillar activity to its chrysalis passage because its development and metabolism are directly related to temperature and precipitation. Moreover, its activity begins only above a certain temperature (here fixed at 10 ° C., but which may vary according to the species). During the pupal stage, we used GDD10 as well as the temperature values (minimum, maximum and mean) which influence the metabolism and potentially the duration of the chrysalis passage. Then during the adult stage, we used precipitations as well as temperatures that influence the possibility of moths to mate and lay eggs. Finally, during the egg period, we used the temperatures and GGD10 of the previous year observation (i.e. 2009 in the example of Figure 2). Indeed, temperatures must probably have a role in the development of the eggs. These choices were made to best match all the species used for the comparisons, however, these criteria vary widely between species.

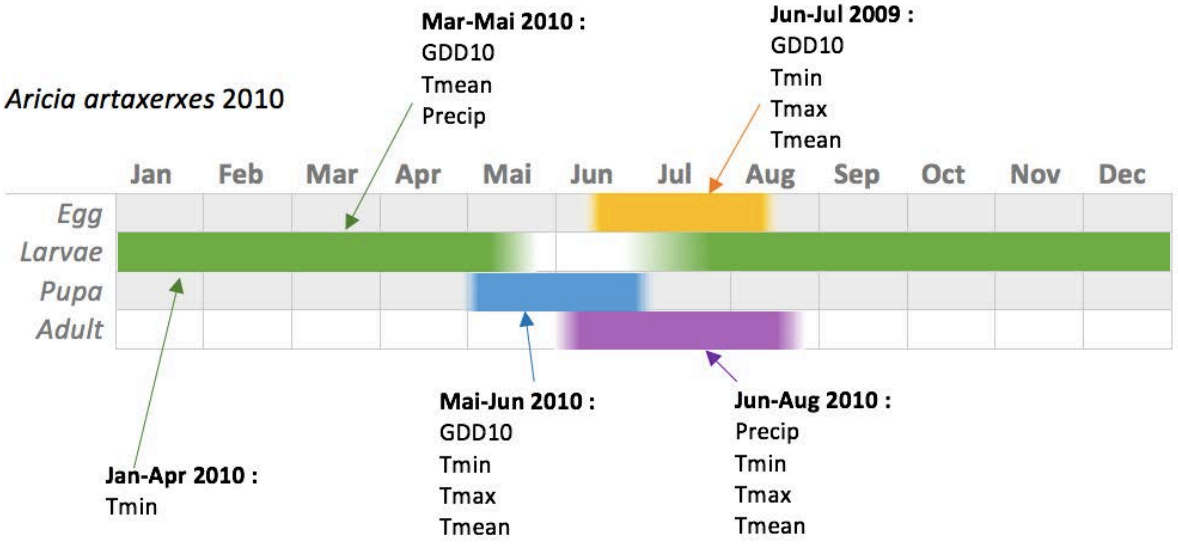


Figure 2: Monthly accurate modelling method. Diagram of the temporal placement of the different variables according to the stages of life of the butterflies. This is an example for species *Aricia artaxerxes* observed in 2010. The same scheme was adapted for the 12 other species used.

MODEL FITTING AND EVALUATING

For all SDMs calculated, the response variable is the presence or absence of the species at each sample points during three years (2009, 2010 and 2014). The climatic variables are selected at the same points to explain the presence or absence of the species. The modelling was calculated on *R* (*R* Development Core Team) using techniques implemented on *BIOMOD2* package (Thuiller *et al.,* 2009). Five algorithms were run 20 times for each SDM. These algorithms are: Generalized Linear Models (McCullagh& Nelder, 1989; Guisan *et al.*, 2002), Boosted Regression Tree (Friedman, Hastie & Tibshirani, 2000 ; Elith *et al.*, 2008), Generalized Additive Models (Hastie & Tibshirani, 1990 ; Guisan *et al.* 2002), Classification Tree Analysis (Gordon *et al.*, 1984 ; Thuiller *et al.*, 2003) and Flexible Discriminant Analysis (Hastie *et al.*, 1994).

Then the results of the 20 runs of each of algorithms are combined into a single ensemble according to the weighted sum of probabilities of each results (Araujo & New, 2007; Thuiller *et al.*, 2009). We conducted a repeated split sample cross validation with 80% of the data for the modelling and 20% for the evaluation. We used four indicators to compare our models. The sensitivity, specificity, Receiver Operating Characteristic (ROC) (Fielding & Bell, 1997) and True Skill Statistics (TSS) (Allouche *et al.*, 2006), see Tables 1 and 2. The sensitivity is the proportion of correct predicted presence while the specificity is the proportion of correct predicted absence. ROC is the relation between all sensitivity values and their equivalent values (1-specificity). The area under the ROC function (AUC) provides a single measure of accuracy comprised between 0.5 (the score not differ for two groups) and 1 (no overlap in the distributions of the group scores). Thus, a score of 0.5 indicates totally random predictions while a score of 1 indicates predictions that are always correct.

		Validation data set	
		Presence	Absence
Model	Presence	*a*	*b*
	Absence	*c*	*d*

Table 1: Error matrix used to test model accuracy. The value *a* represents the number of presence correctly predicted (true positive). The value *b* indicates the absences predicted as presence (omission error). The value *c* indicates the presence predicted as absence (commission error). Finally, the value *d* indicates the absences correctly predicted (true negative).

Table 2: Formulas of the different precision measurements used.

Measure	Formula
Sensitivity	$\frac{a}{a+c}$
Specificity	$\frac{d}{b+d}$
TSS	sensitivity + specificity – 1

RESULTS

Since the results of the first approach (one-year average data versus 30-years average data) do not change the conclusion, the results are presented in supplementary materials. Here are presented the results obtained with the second approach (accurate monthly data against the use of 30-years average data).

Tableau 3: TSS evaluations of each modelling. The last column represents the difference of evaluations (30-years minus monthly).

	SPECIES	TSS 30-YEARS	TSS MONTHLY	DIFFERENCE
1	Aporia_crataegi	0.773	0.773	0
2	Aricia_artaxerxes	0.716	0.689	-0.027
3	Aricia_eumedon	0.816	0.685	-0.131
4	Boloria_napaea	0.871	0.844	-0.027
5	Brenthis_ino	0.929	0.872	-0.057
6	Carterocephalus_palaemon	0.837	0.829	-0.008
7	Colias_phicomone	0.696	0.683	-0.013
8	Lycaena_hippothoe	0.799	0.788	-0.011
9	Pyrgus_serratulae	0.843	0.720	-0.123
10	Argynnis_paphia	0.980	0.971	-0.009
11	Boloria_pales	0.742	0.706	-0.036
12	Hesperia_comma	0.754	0.753	-0.001
13	Polyommatus_eros	0.829	0.768	-0.061
	Mean	**0.81**	**0.78**	**-0.04**

Tableau 4: ROC evaluations of each modelling. The last column represents the difference of evaluations (30-years minus monthly).

	SPECIES	ROC 30-YEARS	ROC MONTHLY	DIFFERENCE
1	Aporia_crataegi	0.943	0.944	0.001
2	Aricia_artaxerxes	0.930	0.933	0.003
3	Aricia_eumedon	0.964	0.918	-0.046
4	Boloria_napaea	0.973	0.972	-0.001
5	Brenthis_ino	0.978	0.972	-0.006
6	Carterocephalus_palaemon	0.969	0.952	-0.017
7	Colias_phicomone	0.914	0.913	-0.001
8	Lycaena_hippothoe	0.965	0.953	-0.012
9	Pyrgus_serratulae	0.958	0.932	-0.026
10	Argynnis_paphia	0.997	0.993	-0.004
11	Boloria_pales	0.953	0.934	-0.019
12	Hesperia_comma	0.936	0.941	0.005
13	Polyommatus_eros	0.972	0.943	-0.029
	Mean	**0.96**	**0.95**	**-0.01**

TSS and ROC evaluations carried out on 13 species of butterflies are not significantly different between models using monthly accurate data and corresponding models using 30-years average data (Welch Two Sample t-test: TSS comparison, p-value = 0.2488 and ROC comparison, p-value = 0.1962). In both cases, the evaluations of accuracy are high, with TSS evaluations between 0.683 and 0.980 and ROC evaluations between 0.918 and 0.997. The sensitivity as well as the specificity do not differ significantly according to the model (Welch Two Sample t-test : Sensitivity, p-value = 0.7581 and specificity, p-value = 0.1074). However, the difference in specificity average for the 13 species is higher (-0.044) than that of sensitivity (0.006) but these values remain very low (Figure 3).

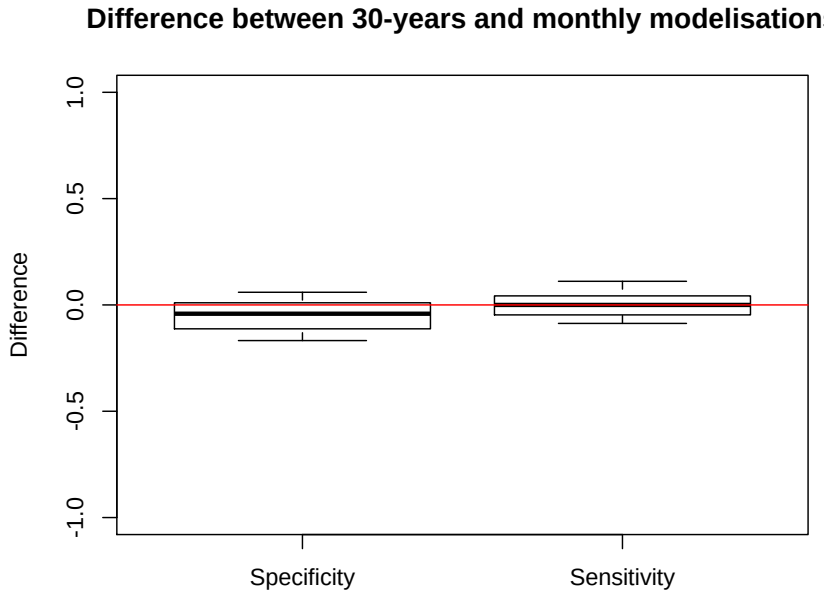


Figure 3 Boxplot of the differences in values between the 30-years models and the month models regarding specificity (left) and sensitivity (right). The differences can range from -1 to 1. In this case, we have -0.04416 for specificity and 0.0065 for sensitivity.

The projection maps for each model are available in supplementary materials. As illustration, only the maps for *Aricia eumedon* are presented here (Figure 4). All maps are available in the supplementary material section. Although the evaluations presented above are not significantly different, we observe differences on projection maps according to the method used. In some cases, a model predicts the absence of the species in a large region while the other model predicts a high probability of presence. These differences in predictions are shown in Figure 4, on the map on the right.

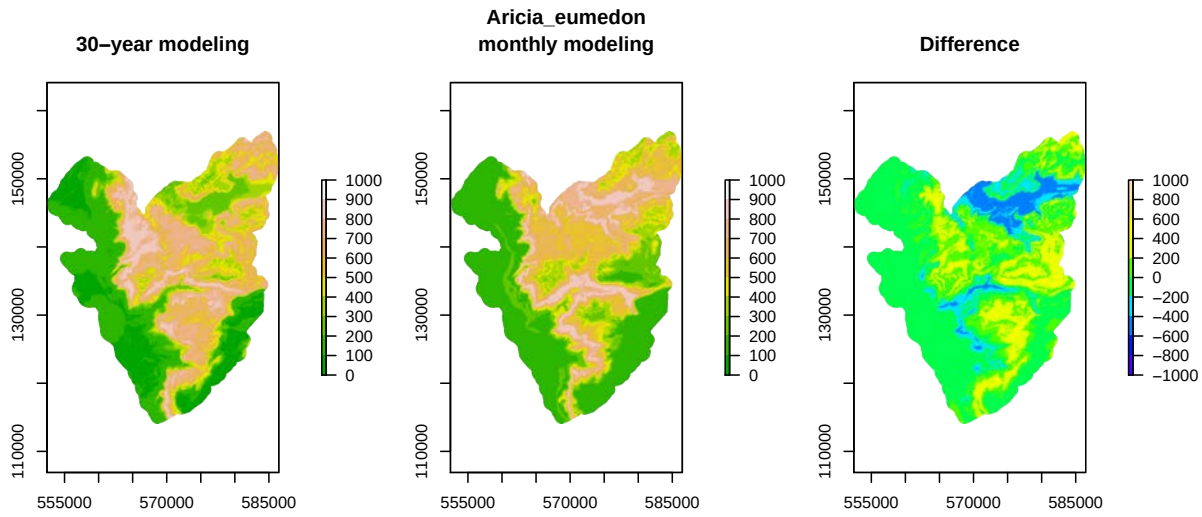


Figure 4: Projection maps of the species *Aricia eumedon*. The map on the left is modeled from climatic data averaged over 30 years while that of the center from data precise at the month (of the year of observation). The map on the right shows the difference between the two models (green = same predictions, more yellow = more presence probabilities predicted by the model 30-years, and more blue = more presence probabilities predicted by the models specific to the month)

DISCUSSION

These results do not confirm the hypothesis that the use of these finer temporal matching variables produces more accurate SDMs than those calculating with 30-years averages variables. We will discuss in detail some elements that may explain these non-expected results.

CLIMATIC VARIABLES VS OTHER VARIABLES

Obviously the presence of a butterfly is not explained only by the climatic variables. Many other variables are involved, such as elevation, exposure, competition with other species, influence of human activity, etc. Since these variables are not taken into account in any models compared, they should not have an impact on the results. However, one other variable could be involved, it is the presence of host plants of the butterfly. Indeed, each species of butterfly is totally dependent on one or more host plants. The butterfly cannot be present if the host plant is not. In this way, the distribution of host plants could be better than climatic variables to explain the distribution of butterflies (Cormont *et al.*, 2013; but Hawkins & Porter, 2003). In addition, plants have slow response to climatic variation. It is therefore possible that 30-years average climatic variables explain better the distribution of host plants than temporally more precise variables (over a year or a few months). Thus, the results would be biased by the influence of the distribution of the host plants.

THE RISK OF FALSE ABSENCES

The number of passages per sampling point was three in the majority of cases. However, it is possible that during the passages all butterfly species present in the area have not been observed due to their high mobility. This can therefore generate false absences. However, although too many false absences reduce the quality of the model, this reduction is valid for all modelling performed. Thus, there is no impact on the results of comparisons between models.

FURTHER INVESTIGATION

In order to better temporally situate climatic variables, it would be interesting to verify the biology of the species studied in the field (instead of using the literature as in this study). As we have seen with butterflies, the periods of the different stages of life vary according to the region. The number of generation and behavior (migration or not) also varies. These are all data to be checked on the field for each population modeled.

In order to deepen the potential uses of a finer temporal matching, it is necessary to test this approach on other datasets. This will allow comparisons of potential improvements across species and regions. The results should vary depending on the mobility of the species as well as its behavioral adaptation to the weather.

We have seen that the projection maps were different according to the approach used. It would be interesting to characterize regions where predictions differ. A finer temporal matching variables approach is maybe more useful in high altitude where the weather varies greatly or in area where anthropogenic perturbations interact.

ACKNOWLEDGMENTS

I would like to thank Olivier Brönnimann and Antoine Guisan for their guidance through this project and for their wise advice. I also would like to thank Heiner Ziegler for providing the data on the pieris.ch website and for answering all my questions. Finally, I thank Yannick Chittaro for taking the time to give me a lot of advice and information about butterflies. Data were kindly provided by Ecospat groups where many people participated in the field gathering as well as the processing of these data, thanks to all of them.

REFERENCES

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223-1232.

Araujo, M. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42-47.

Buckley, L. & Kingsolver, J. (2012) The demographic impacts of shifts in climate means and extremes on alpine butterflies. *Functional Ecology*, **26**, 969-977.

Cormont, A., Malinowska, A., Kostenko, O., Radchuk, V., Hemerik, L., WallisDeVries, M. & Verboom, J. (2010) Effect of local weather on butterfly flight behaviour, movement, and colonization: significance for dispersal under climate change. *Biodiversity and Conservation*, **20**, 483-503.

Cormont, A., Wieger Wamelink, G., Jochem, R., WallisDeVries, M. & Wegman, R. (2013) Host plant-mediated effects of climate change on the occurrence of the Alcon blue butterfly (Phengaris alcon). *Ecological Modelling*, **250**, 329-337.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. & E. Zimmermann, N. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.

Elith, J. & Leathwick, J. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677-697.

Feltwell, J. (1983) Large White Butterfly. The Biology, Biochemistry and Physiology of Pieris Brassicae (Linnaeus). *The Quarterly Review of Biology*, **58**, 99-100.

FIELDING, A. & BELL, J. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.

Geister, T., Lorenz, M., Hoffmann, K. & Fischer, K. (2008) Energetics of embryonic development: effects of temperature on egg and hatchling composition in a butterfly. *Journal of Comparative Physiology* B, **179**, 87-98.

Gordon, A., Breiman, L., Friedman, J., Olshen, R. & Stone, C. (1984) Classification and Regression Trees. *Biometrics*, **40**, 874.

Guisan, A. & Zimmermann, N. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147-186.

Guisan, A., Edwards, T. & Hastie, T. (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89-100.

Guisan, A., Tingley, R., Baumgartner, J., Naujokaitis-Lewis, I., Sutcliffe, P., Tulloch, A., Regan, T., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T., Rhodes, J., Maggini, R., Setterfield, S., Elith, J., Schwartz, M., Wintle, B., Broennimann, O., Austin, M., Ferrier, S., Kearney, M., Possingham, H. & Buckley, Y. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424-1435.

Hastie, T., Tibshirani, R. & Buja, A. (1994) Flexible Discriminant Analysis by Optimal Scoring. *Journal of the American Statistical Association*, **89**, 1255.

Hawkins, B. & Porter, E. (2003) Does Herbivore Diversity Depend on Plant Diversity? The Case of California Butterflies. *The American Naturalist*, **161**, 40-49.

KARL, I., STOKS, R., DE BLOCK, M., JANOWITZ, S. & FISCHER, K. (2011) Temperature extremes and butterfly fitness: conflicting evidence from life history and immune function. *Global Change Biology*, **17**, 676-687.

Kuussaari, M., Rytteri, S., Heikkinen, R., Heliölä, J. & von Bagh, P. (2016) Weather explains high annual variation in butterfly dispersal. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 20160413.

Liu, C., White, M. & Newell, G. (2010) Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography*, **34**, 232-243.

Moritz, C. (1995) Uses of Molecular Phylogenies for Conservation. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **349**, 113-118.

Reside, A., VanDerWal, J., Kutt, A. & Perkins, G. (2010) Weather, Not Climate, Defines Distributions of Vagile Bird Species. *PLoS ONE*, **5**, e13569.

Schröder, B. (2008) Challenges of species distribution modelling belowground. *Journal of Plant Nutrition and Soil Science*, **171**, 325-337.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M. (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.

Thuiller, W., Vayreda, J., Pino, J., Sabate, S., Lavorel, S. & Gracia, C. (2003) Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). *Global Ecology and Biogeography*, **12**, 313-325.

Wernberg, T., Smale, D., Tuya, F., Thomsen, M., Langlois, T., de Bettignies, T., Bennett, S. & Rousseaux, C. (2013) An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. *Nature Climate Change*, **3**, 78-82.

Wisz, M., Broennimann, O., Grønkjær, P., Møller, P., Olsen, S., Swingedouw, D., Hedeholm, R., Nielsen, E., Guisan, A. & Pellissier, L. (2015) Arctic warming will promote Atlantic–Pacific fish interchange. *Nature Climate Change*, **5**, 261-265.

SUPPLEMENTARY MATERIALS

RESULTS OF THE FIRST APPROACH

The first approaches compare one-year average data versus 30-years average data (Figure 1). This comparison does not show any significant difference between the two models (Welch Two Sample t-test). This is so for the TSS (t = 0.78317, df = 53.944, p-value = 0.437), ROC (t = 1.1993, df = 53.872, p-value = 0.2357) as well as the sensitivity (t = -1.132, df = 48.461, p-value = 0.2632). Therefore, the difference in specificity is slightly significant (t = 2.081, df = 53.803, p-value = 0.0422) with an average decrease of 3.86% for the one-year modelling (figure 5).

	SPECIES	TSS 30-YEARS	TSS 1-YEAR	DIFFERENCE
1	Anthocharis_cardamines	0.856	0.787	-0.069
2	Aporia_crataegi	0.749	0.729	-0.02
3	Argynnis_adippe	0.881	0.903	0.022
4	Argynnis_aglaja	0.722	0.663	-0.059
5	Argynnis_niobe	0.737	0.742	0.005
6	Argynnis_paphia	1	0.99	-0.01
7	Aricia_artaxerxes	0.718	0.749	0.031
8	Aricia_eumedon	0.717	0.713	-0.004
9	Boloria_napaea	0.883	0.865	-0.018
10	Boloria_pales	0.763	0.741	-0.022
11	Brenthis_ino	0.899	0.902	0.003
12	Callophrys_rubi	0.846	0.768	-0.078
13	Carterocephalus_palaemon	0.879	0.813	-0.066
14	Colias_phicomone	0.689	0.685	-0.004
15	Cupido_minimus	0.699	0.73	0.031
16	Hamearis_lucina	0.891	0.931	0.04
17	Hesperia_comma	0.771	0.719	-0.052
18	Lycaena_hippothoe	0.854	0.849	-0.005
19	Maculinea_arion	0.723	0.727	0.004
20	Parnassius_apollo	0.863	0.893	0.03
21	Pieris_bryoniae	0.729	0.719	-0.01
22	Pieris_napi	0.643	0.653	0.01
23	Pieris_rapae	0.632	0.662	0.03
24	Polyommatus_eros	0.878	0.784	-0.094
25	Pyrgus_alveus	0.856	0.744	-0.112
26	Pyrgus_serratulae	0.843	0.754	-0.089
27	Thymelicus_lineola	0.759	0.748	-0.011
28	Thymelicus_sylvestris	0.823	0.822	-0.001
	Mean	0.80	0.78	-0.02

	SPECIES	ROC 30-YEARS	ROC 1-YEAR	DIFFERENCE
1	Anthocharis_cardamines	0.971	0.951	-0.02
2	Aporia_crataegi	0.936	0.932	-0.004
3	Argynnis_adippe	0.978	0.981	0.003
4	Argynnis_aglaja	0.933	0.907	-0.026
5	Argynnis_niobe	0.929	0.931	0.002
6	Argynnis_paphia	1	0.995	-0.005
7	Aricia_artaxerxes	0.926	0.926	0
8	Aricia_eumedon	0.943	0.922	-0.021
9	Boloria_napaea	0.981	0.962	-0.019
10	Boloria_pales	0.952	0.955	0.003
11	Brenthis_ino	0.977	0.98	0.003
12	Callophrys_rubi	0.965	0.95	-0.015
13	Carterocephalus_palaemon	0.978	0.962	-0.016
14	Colias_phicomone	0.913	0.912	-0.001
15	Cupido_minimus	0.934	0.939	0.005
16	Hamearis_lucina	0.981	0.983	0.002
17	Hesperia_comma	0.942	0.94	-0.002
18	Lycaena_hippothoe	0.973	0.963	-0.01
19	Maculinea_arion	0.929	0.926	-0.003
20	Parnassius_apollo	0.978	0.966	-0.012
21	Pieris_bryoniae	0.934	0.937	0.003
22	Pieris_napi	0.901	0.901	0
23	Pieris_rapae	0.878	0.89	0.012
24	Polyommatus_eros	0.967	0.959	-0.008
25	Pyrgus_alveus	0.973	0.904	-0.069
26	Pyrgus_serratulae	0.961	0.929	-0.032
27	Thymelicus_lineola	0.951	0.943	-0.008
28	Thymelicus_sylvestris	0.974	0.964	-0.01
	Mean	0.95	0.94	-0.01

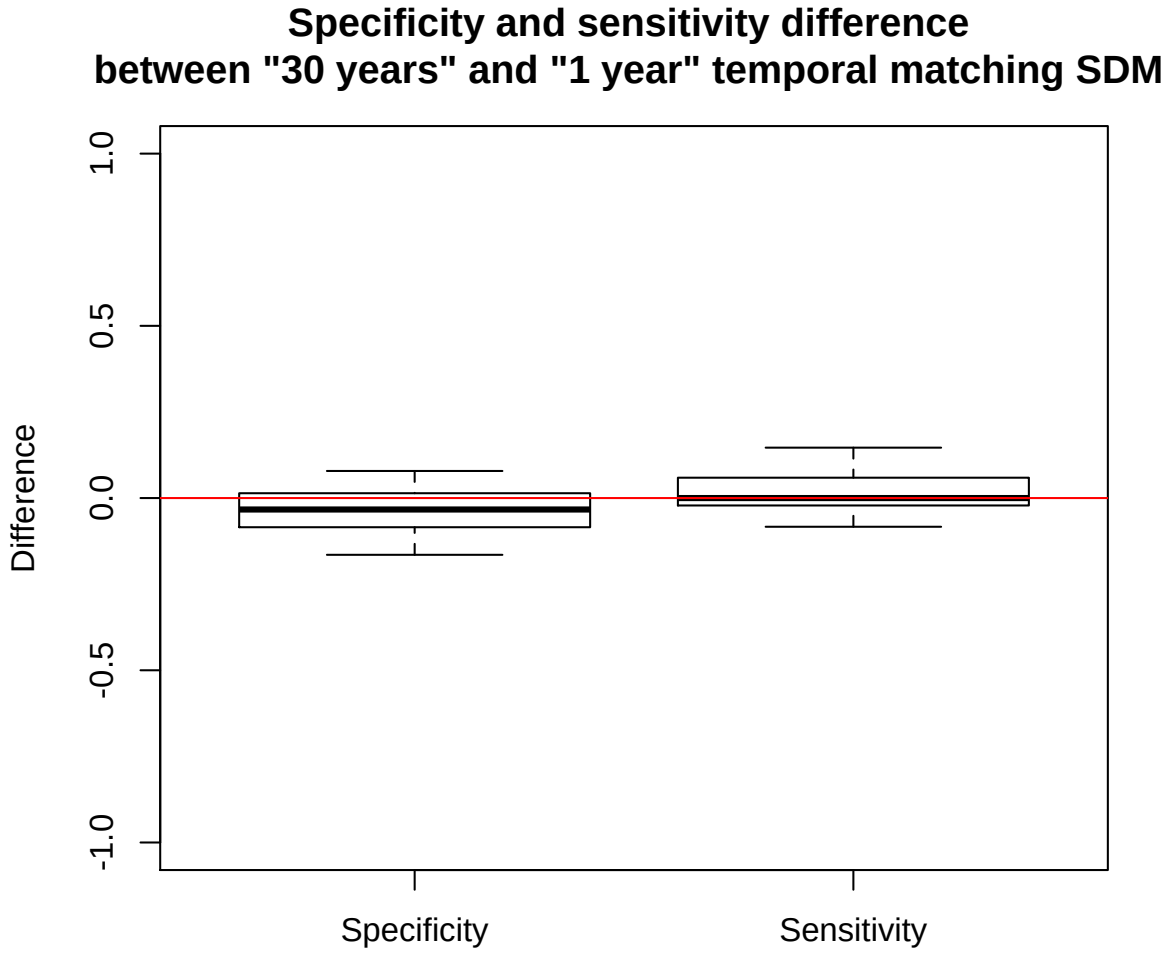
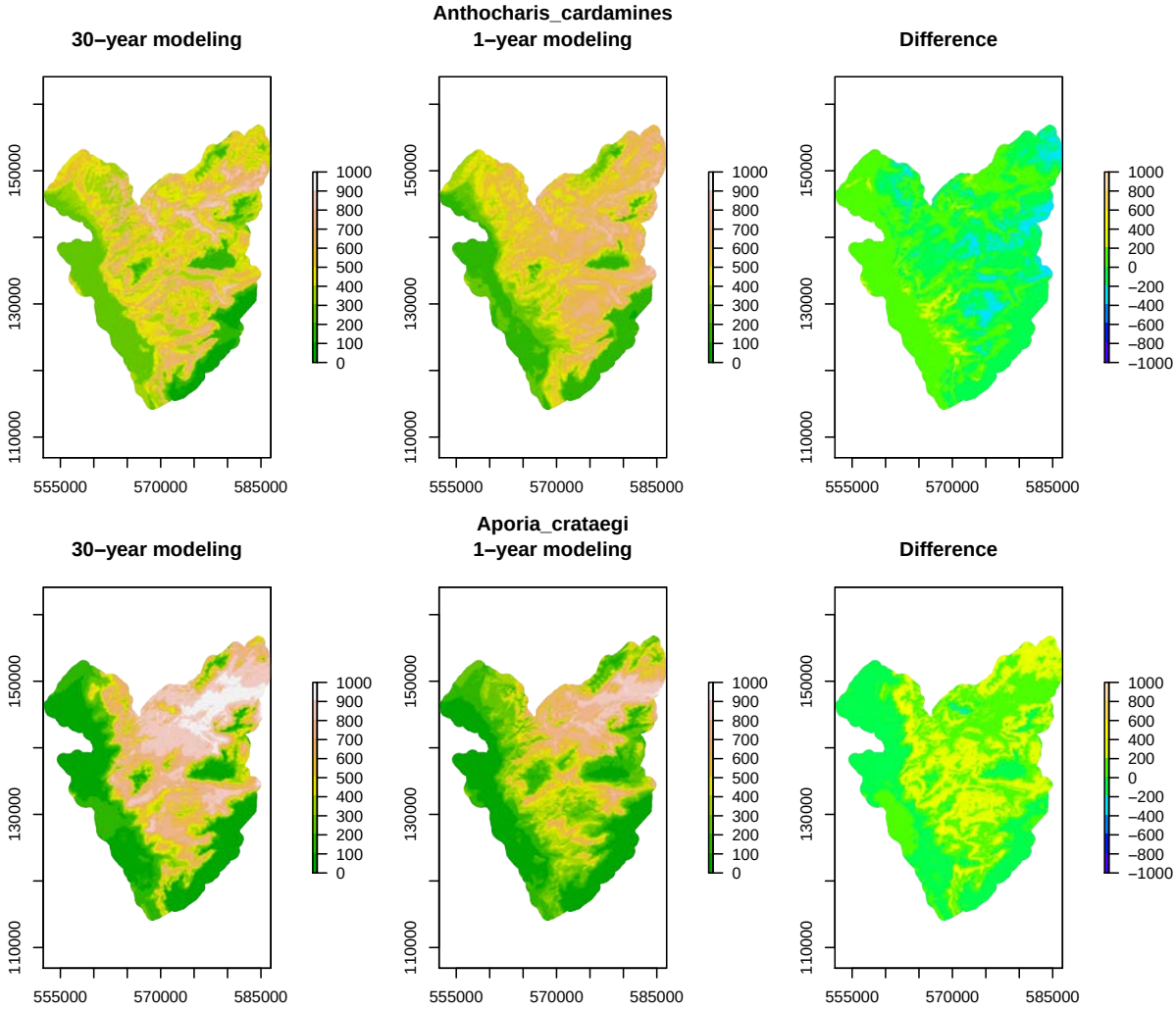
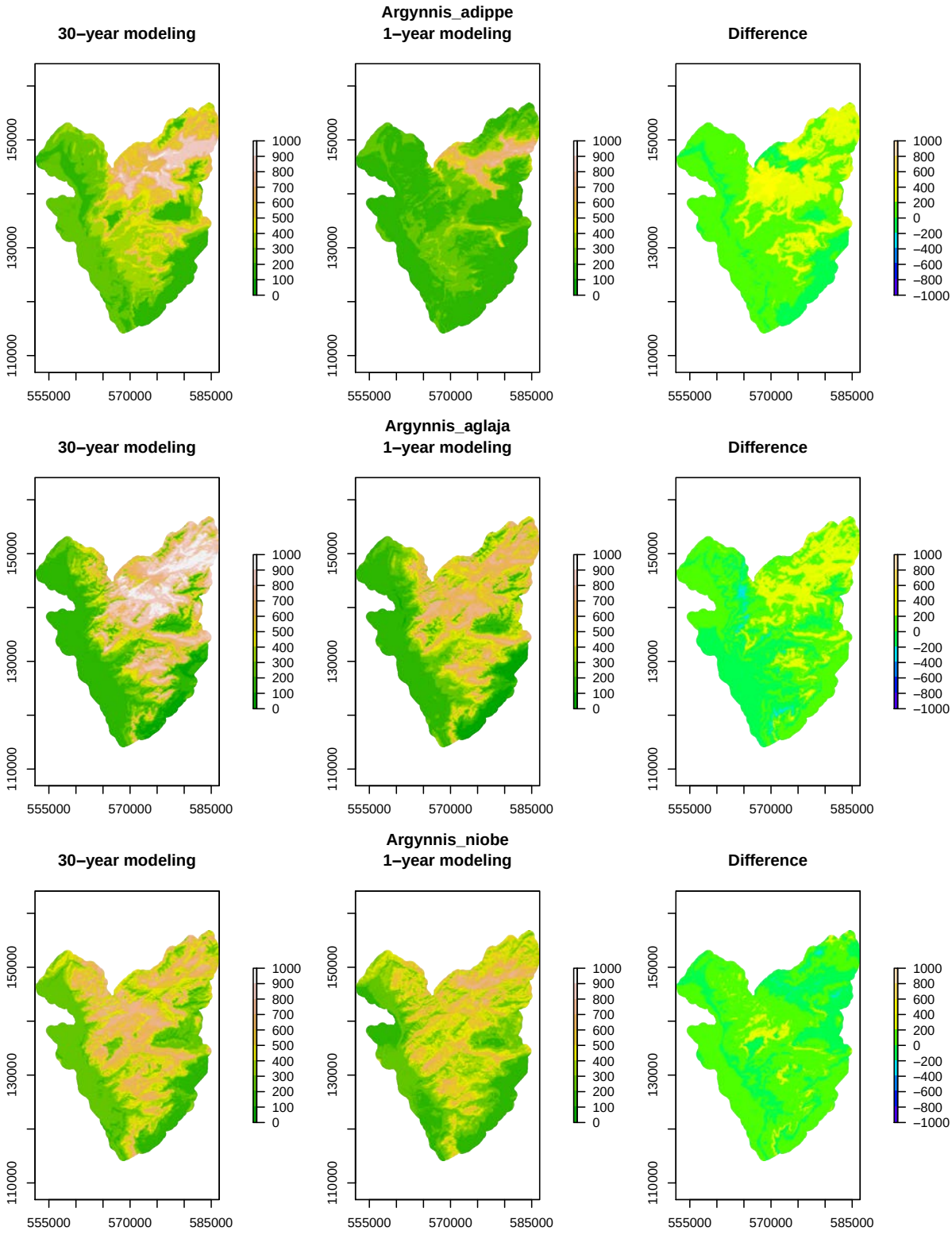


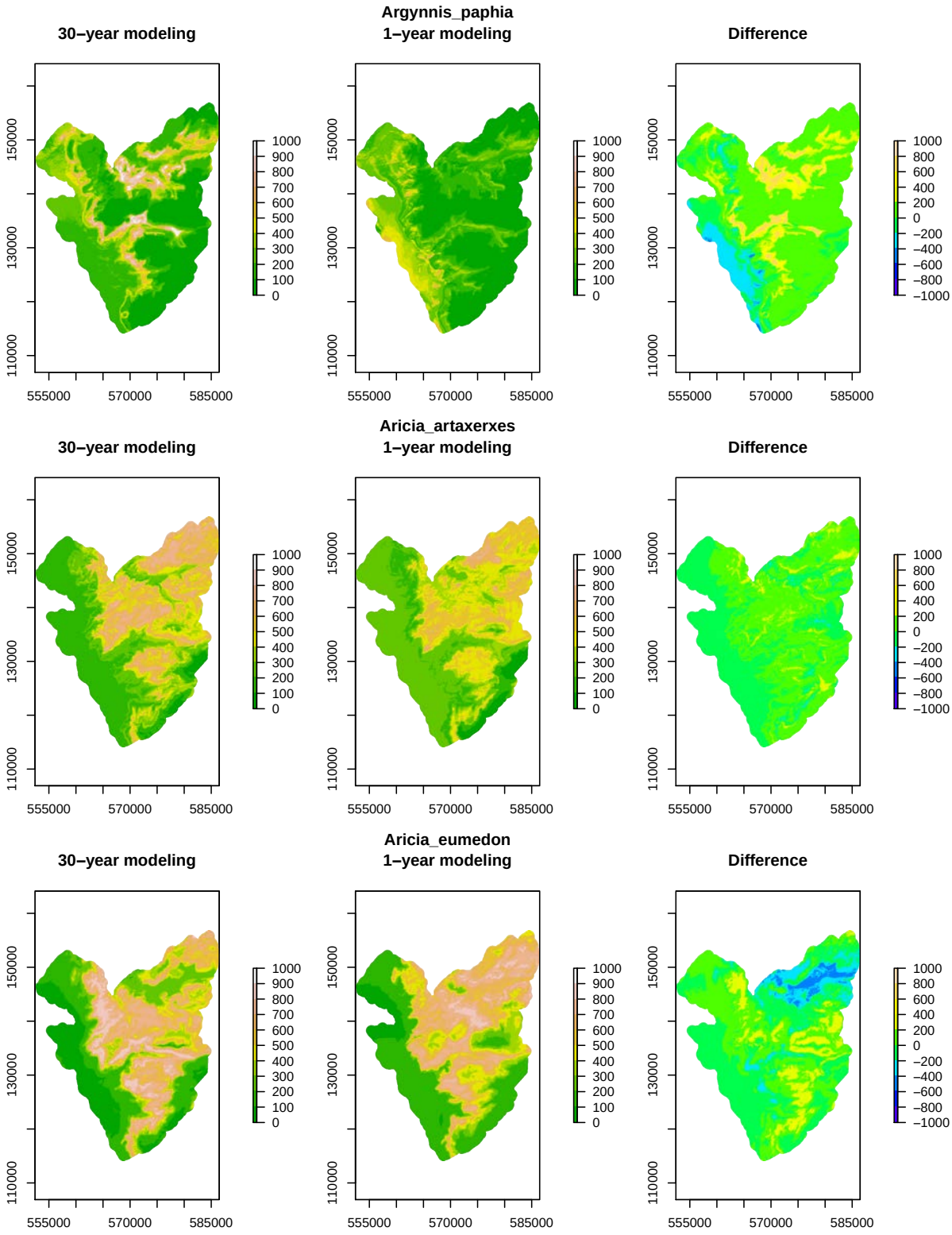
Figure 5 Boxplot of the differences in values between the 30-years models and the one-year models regarding the specificity (left) and the sensitivity (right).

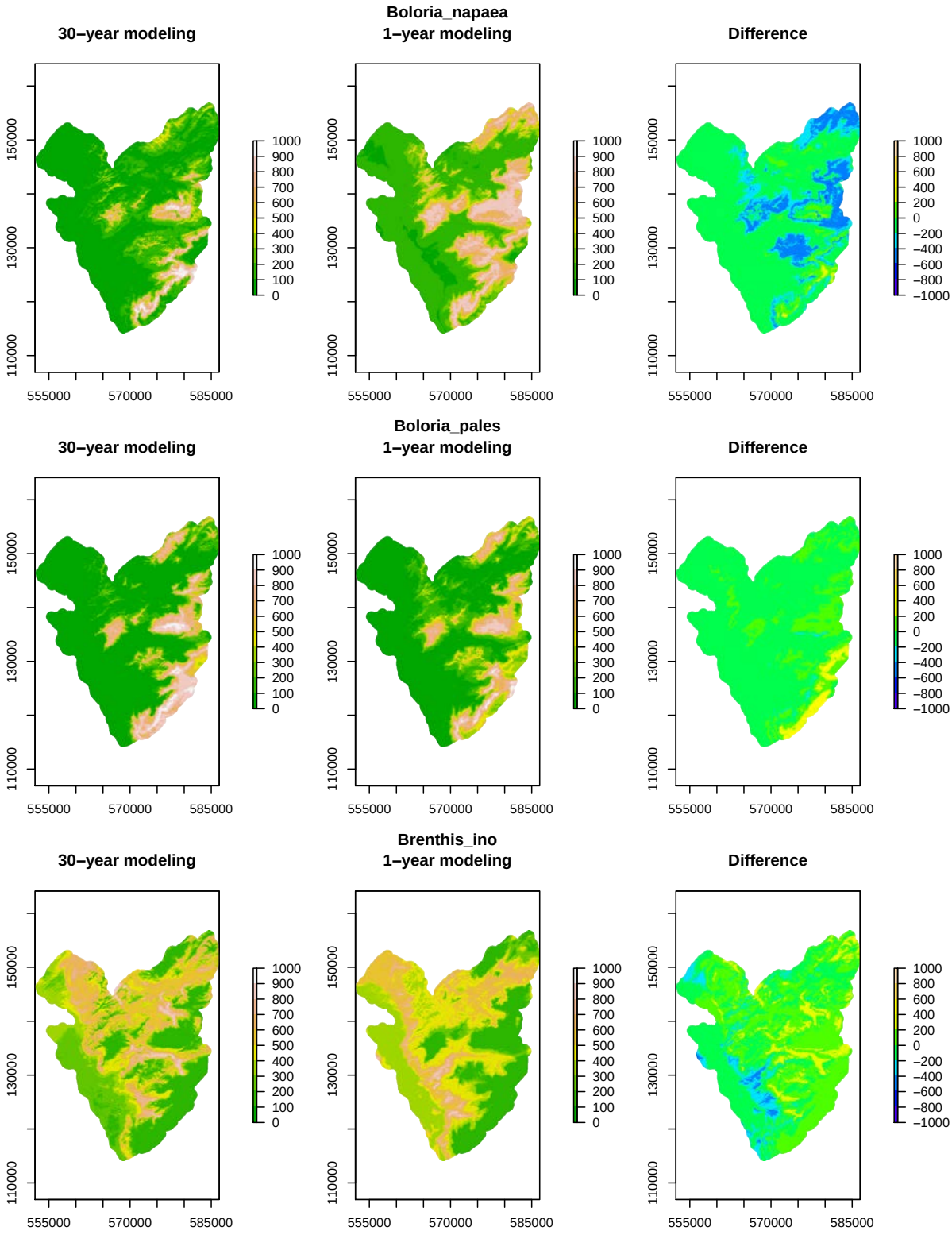
PROJECTIONS MAPS OBTAINED WITH THE FIRST APPROACH

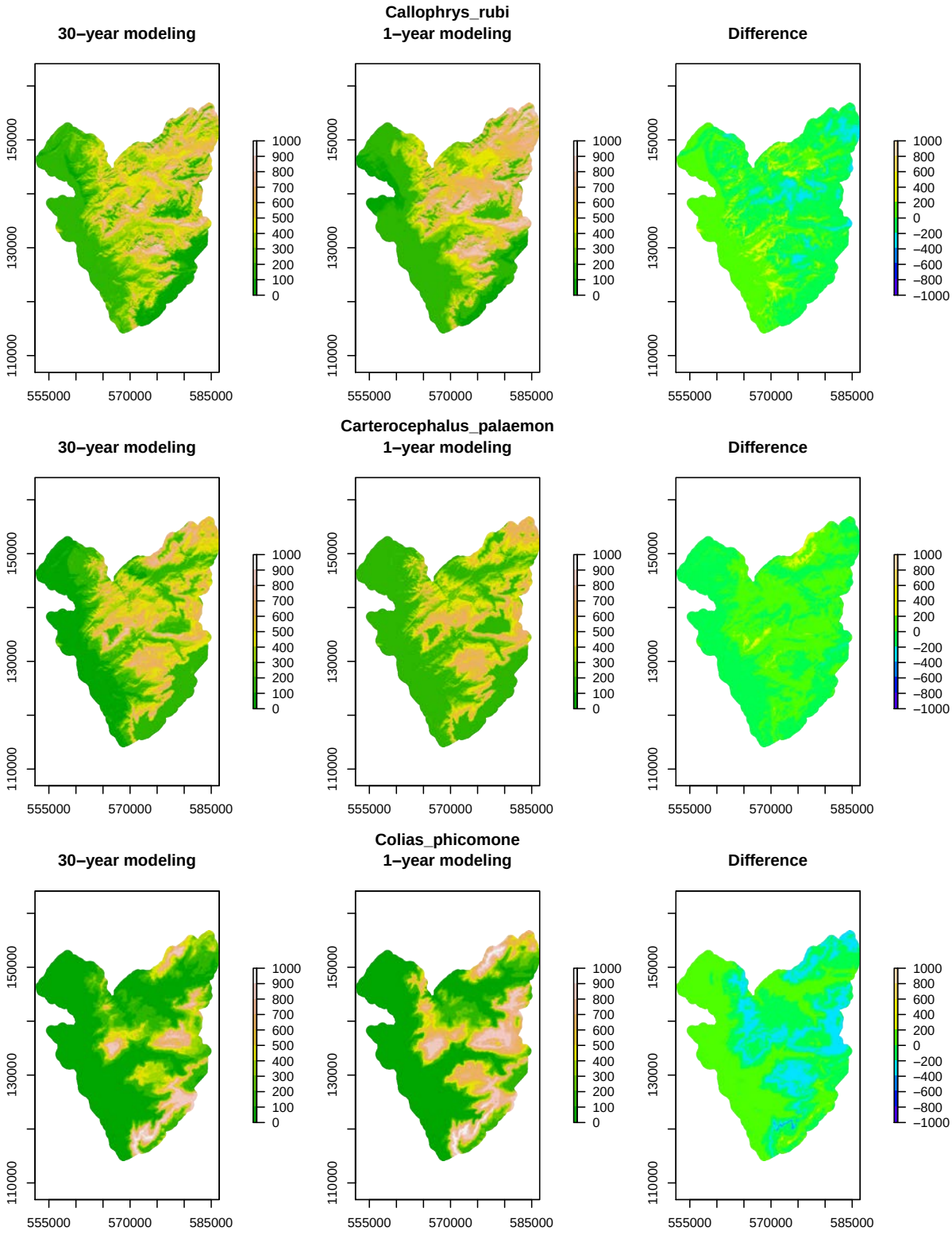
For each species, the projection map on the is based from climatic data averaged over 30 years while that of the center is based from data averaged over 1 year. The right map shows the difference between these two projections (green = same predictions, more yellow = more presence probabilities predicted by the model 30-years, and more blue = more presence probabilities predicted by the models specific to the year).

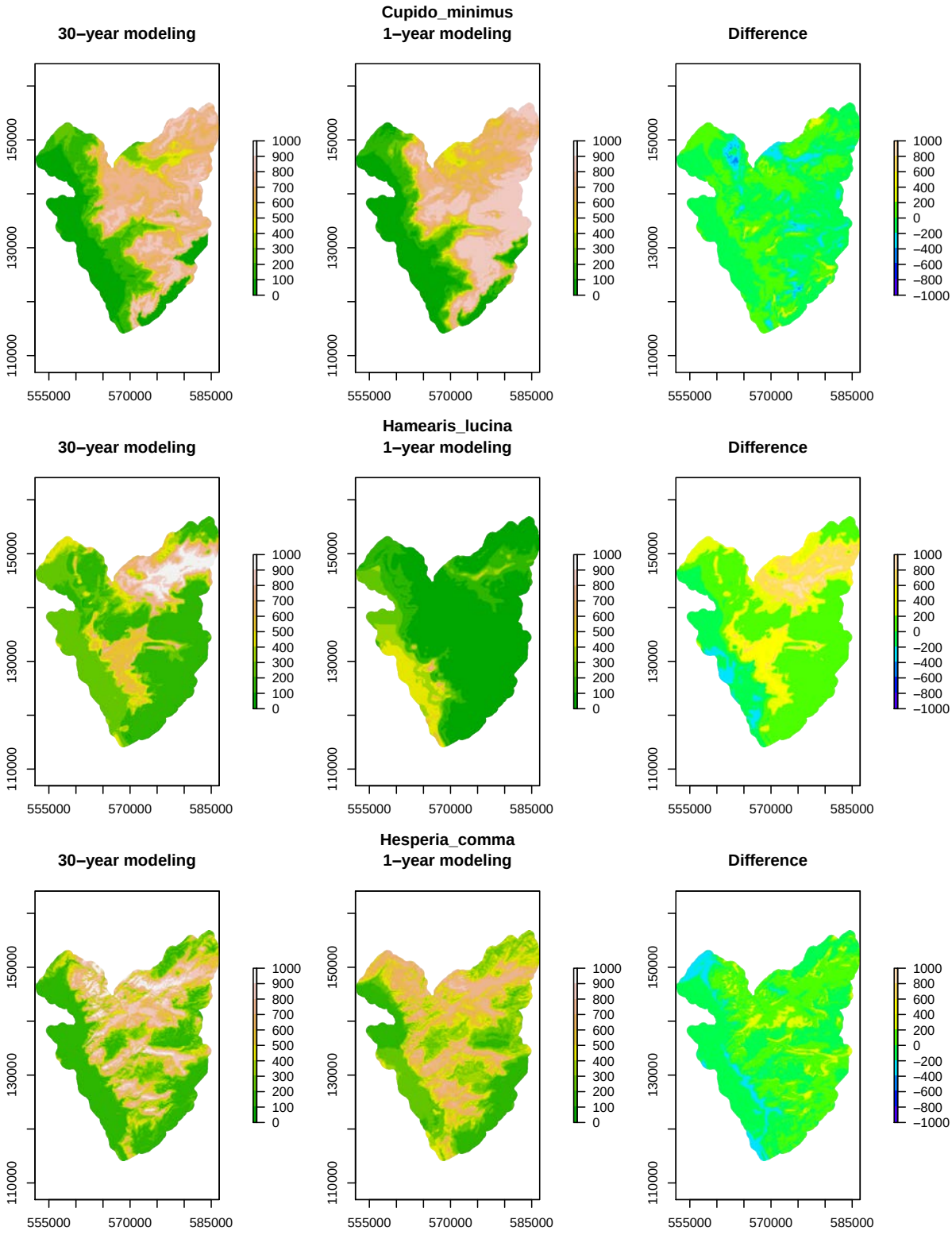


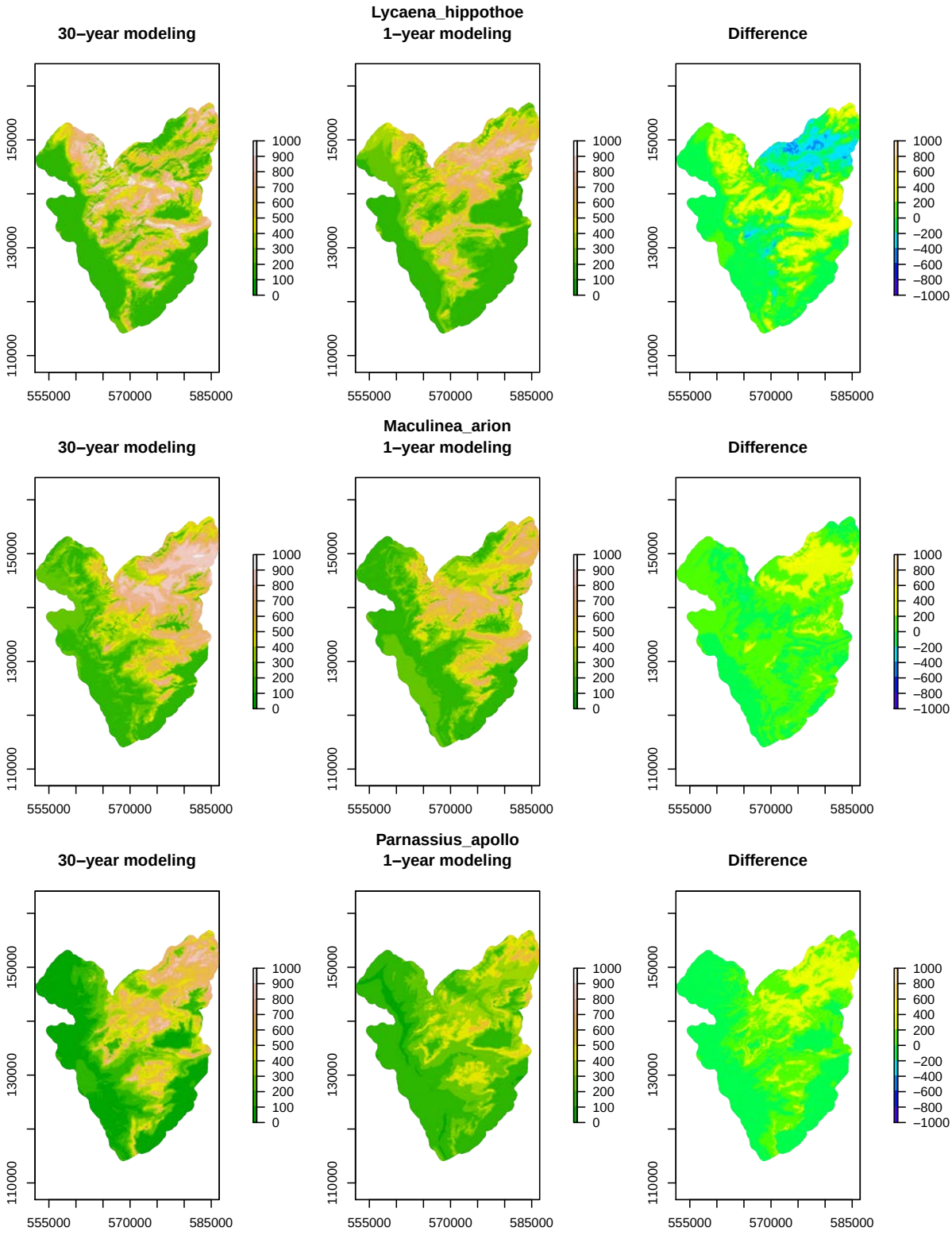

Argynnis_adippe
30–year modeling
1–year modeling
Difference
Argynnis_aglaja
30–year modeling
1–year modeling
Difference
Argynnis_niobe
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
800
600
400
200
0
–200
–400
–600
–800
–1000

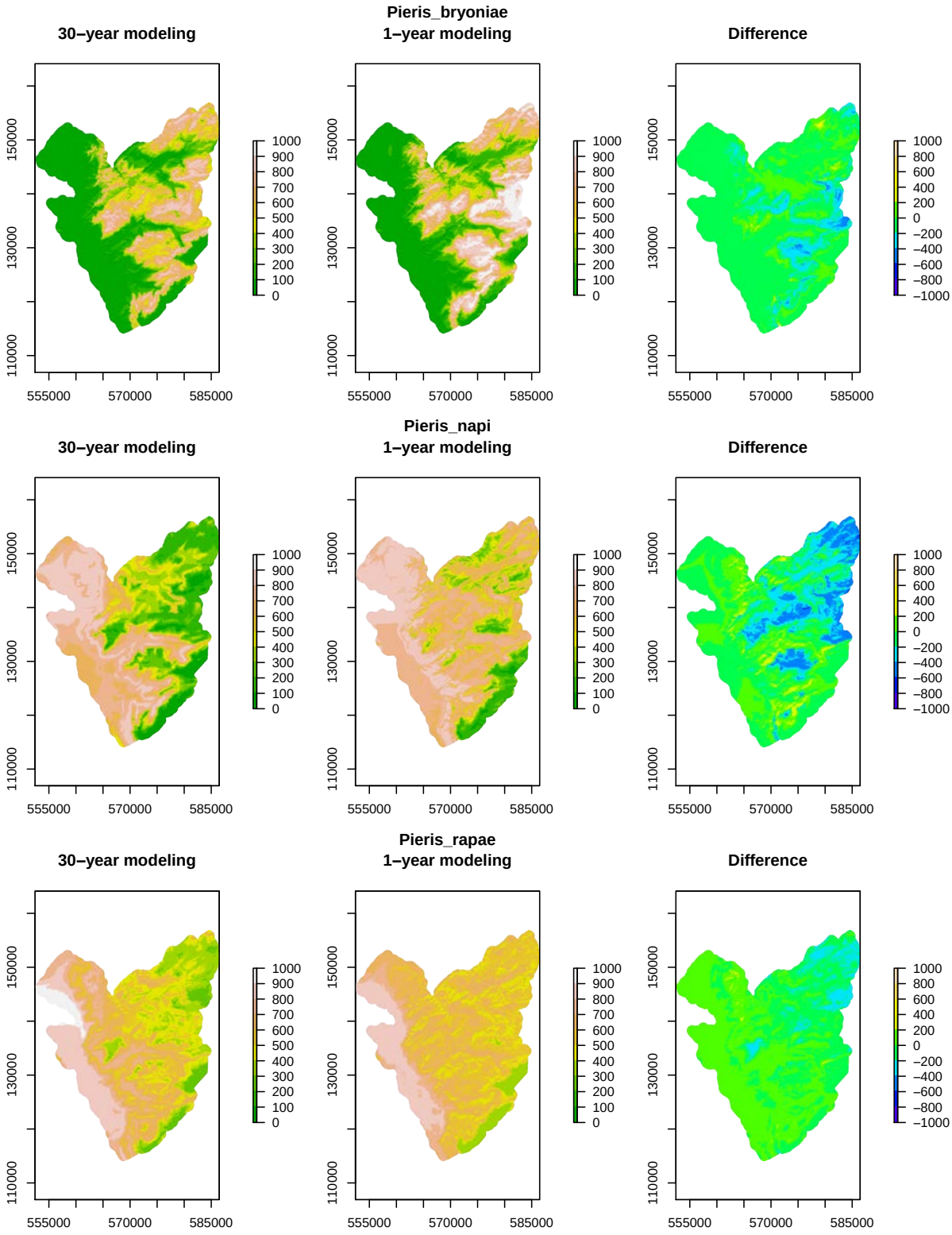

Argynnis_paphia
30–year modeling
1–year modeling
Difference
Aricia_artaxerxes
30–year modeling
1–year modeling
Difference
Aricia_eumedon
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
−200
−400
−600
−800
−1000

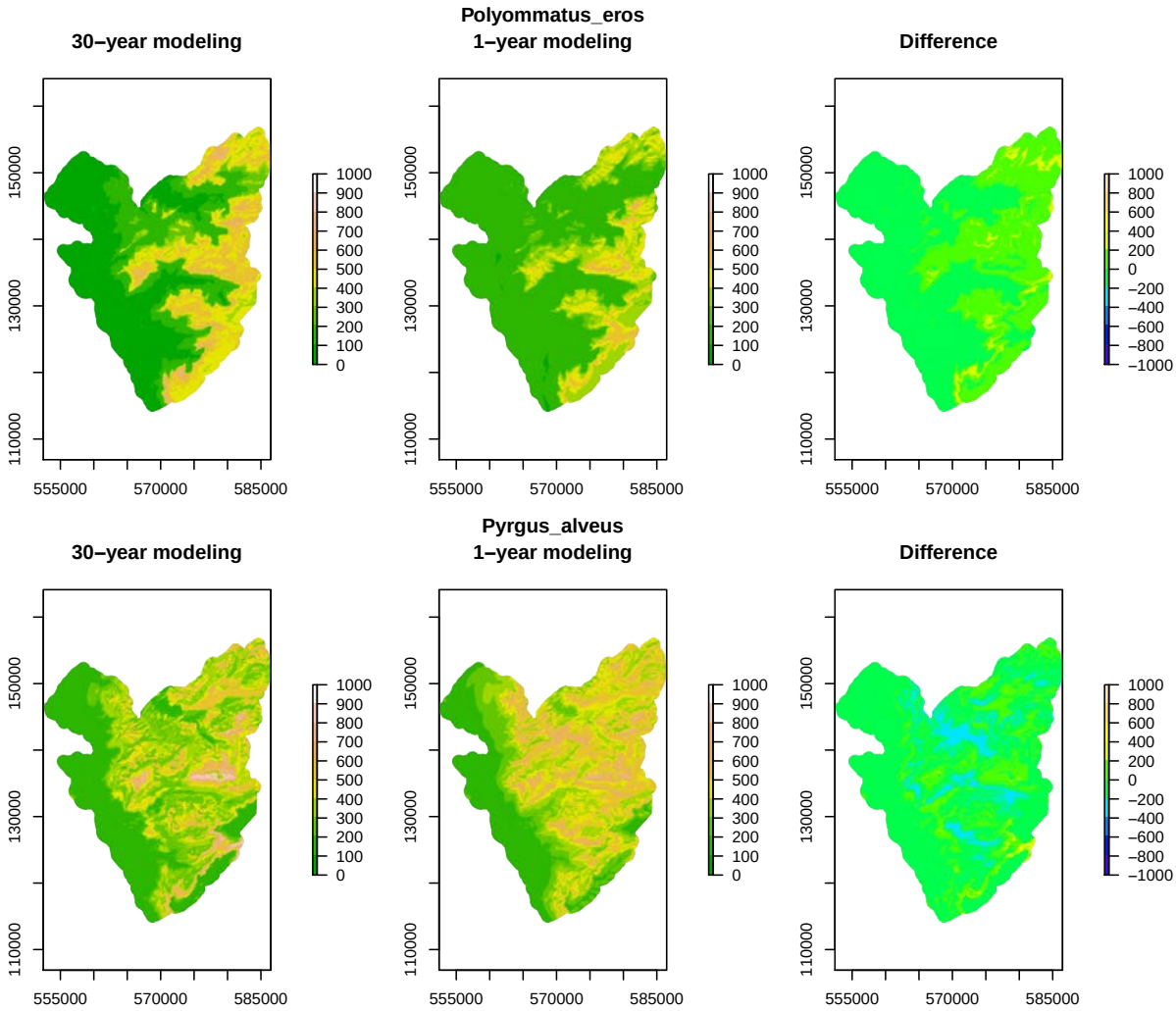

Boloria_napaea
30–year modeling
1–year modeling
Difference
Boloria_pales
30–year modeling
1–year modeling
Difference
Brenthis_ino
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
800
600
400
200
0
–200
–400
–600
–800
–1000

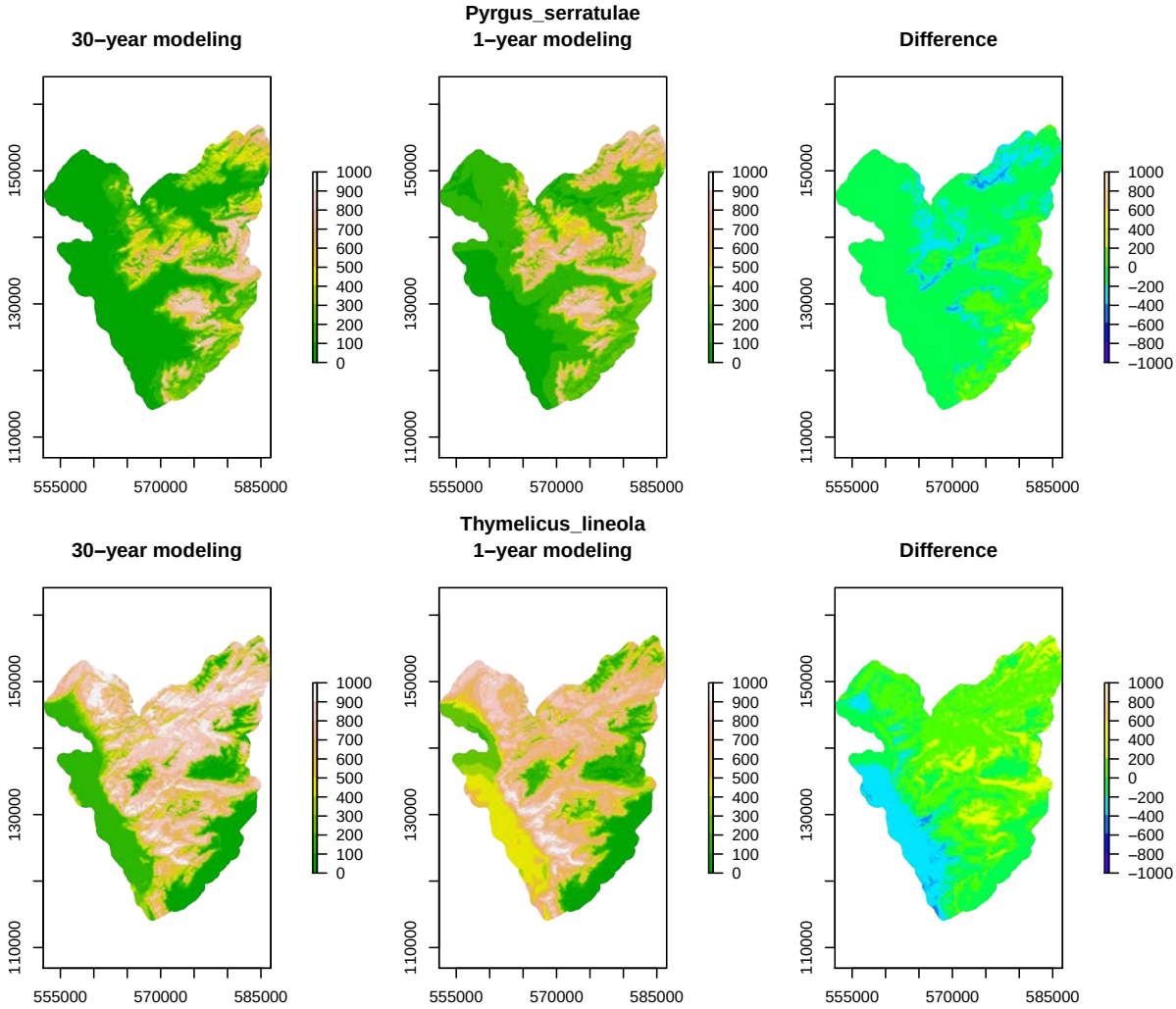

Callophrys_rubi
30–year modeling
1–year modeling
Difference
Carterocephalus_palaemon
30–year modeling
1–year modeling
Difference
Colias_phicomone
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
−200
−400
−600
−800
−1000


Cupido_minimus
30–year modeling
1–year modeling
Difference
Hamearis_lucina
30–year modeling
1–year modeling
Difference
Hesperia_comma
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
800
600
400
200
0
–200
–400
–600
–800
–1000


Lycaena_hippothoe
30–year modeling
1–year modeling
Difference
Maculinea_arion
30–year modeling
1–year modeling
Difference
Parnassius_apollo
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
800
600
400
200
0
–200
–400
–600
–800
–1000

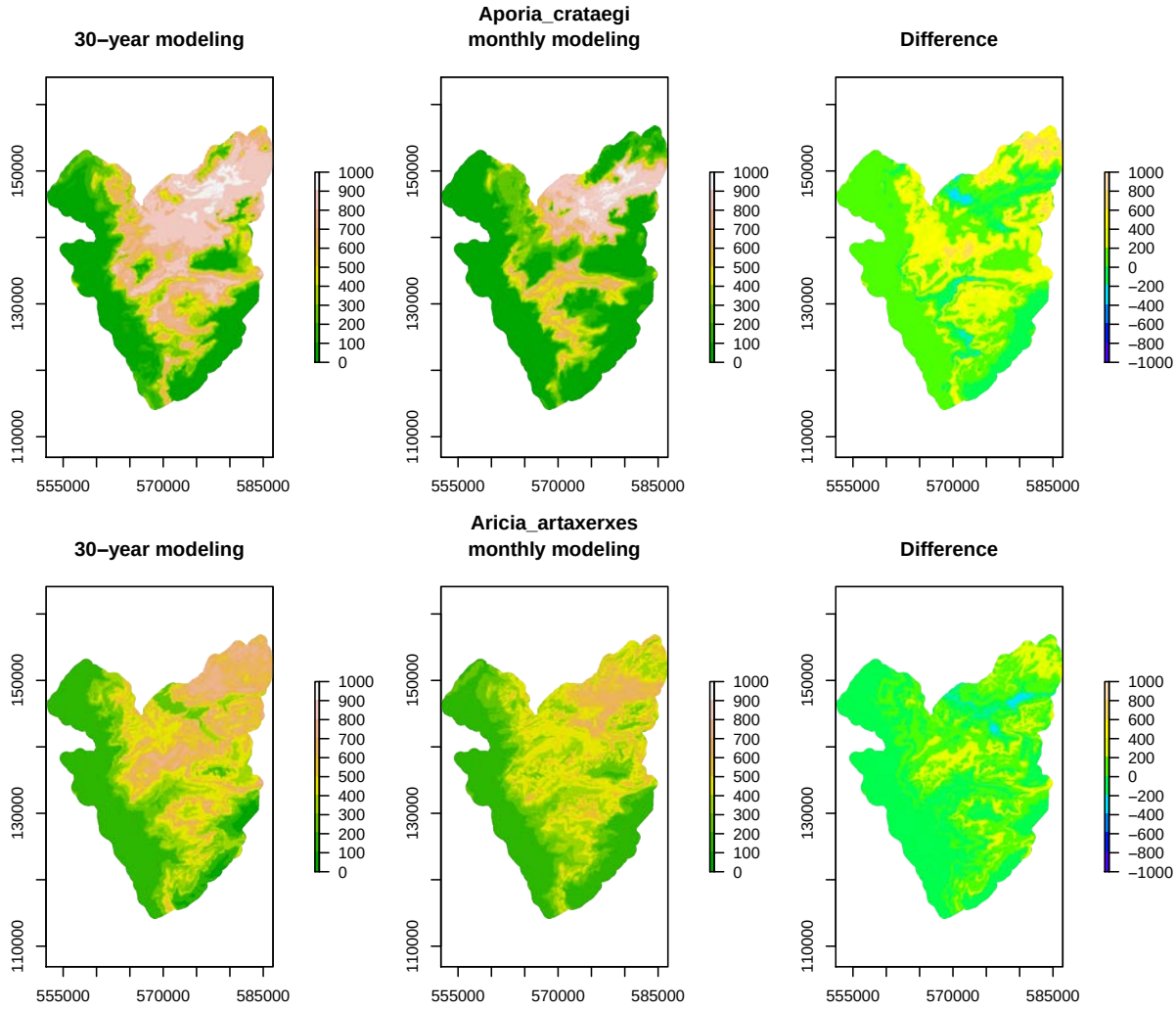

Pieris_bryoniae
30–year modeling
1–year modeling
Difference
Pieris_napi
30–year modeling
1–year modeling
Difference
Pieris_rapae
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
1000
800
600
400
200
0
–200
–400
–600
–800
–1000

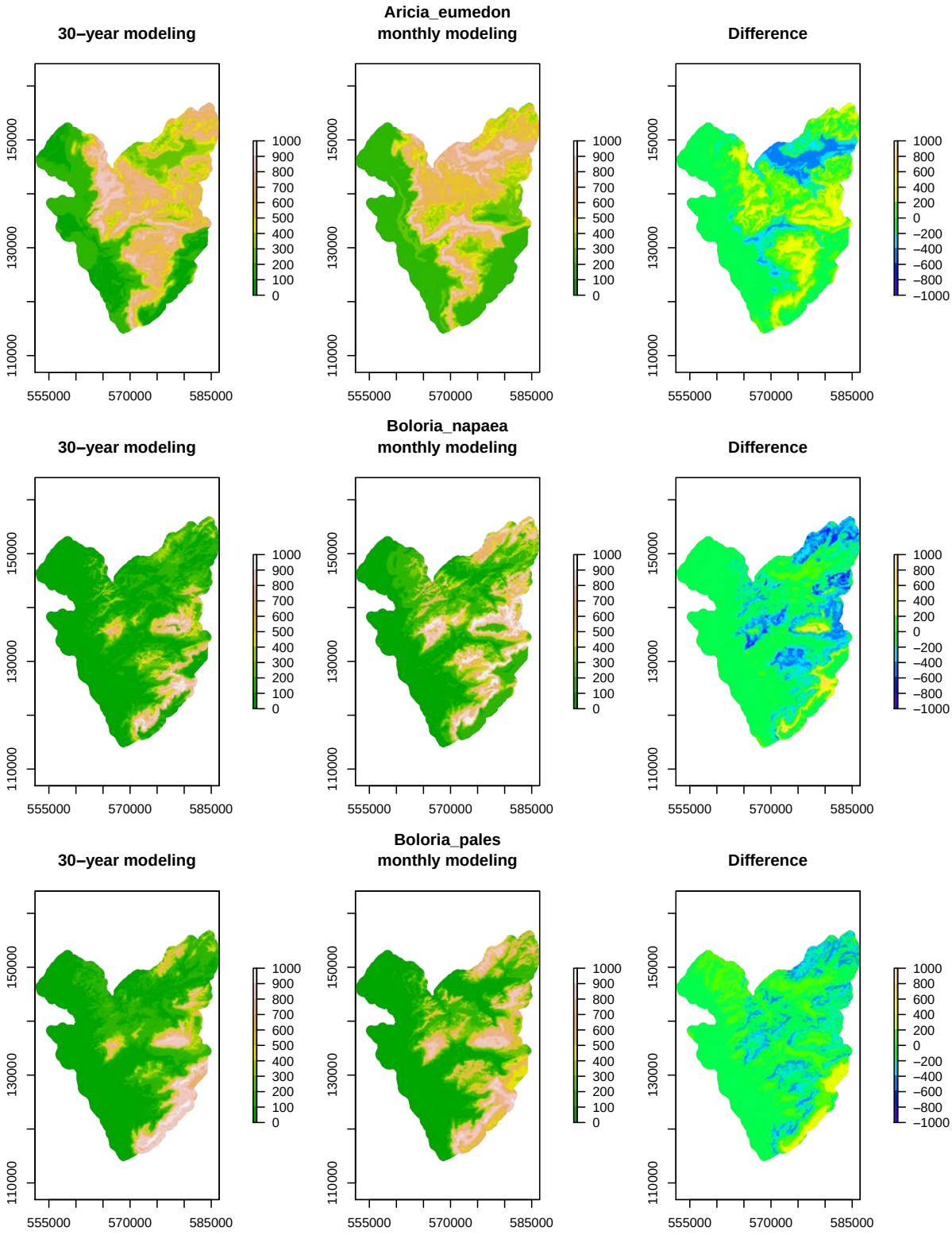

Polyommatus_eros
30–year modeling
1–year modeling
Difference
Pyrgus_alveus
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
–200
–400
–600
–800
–1000

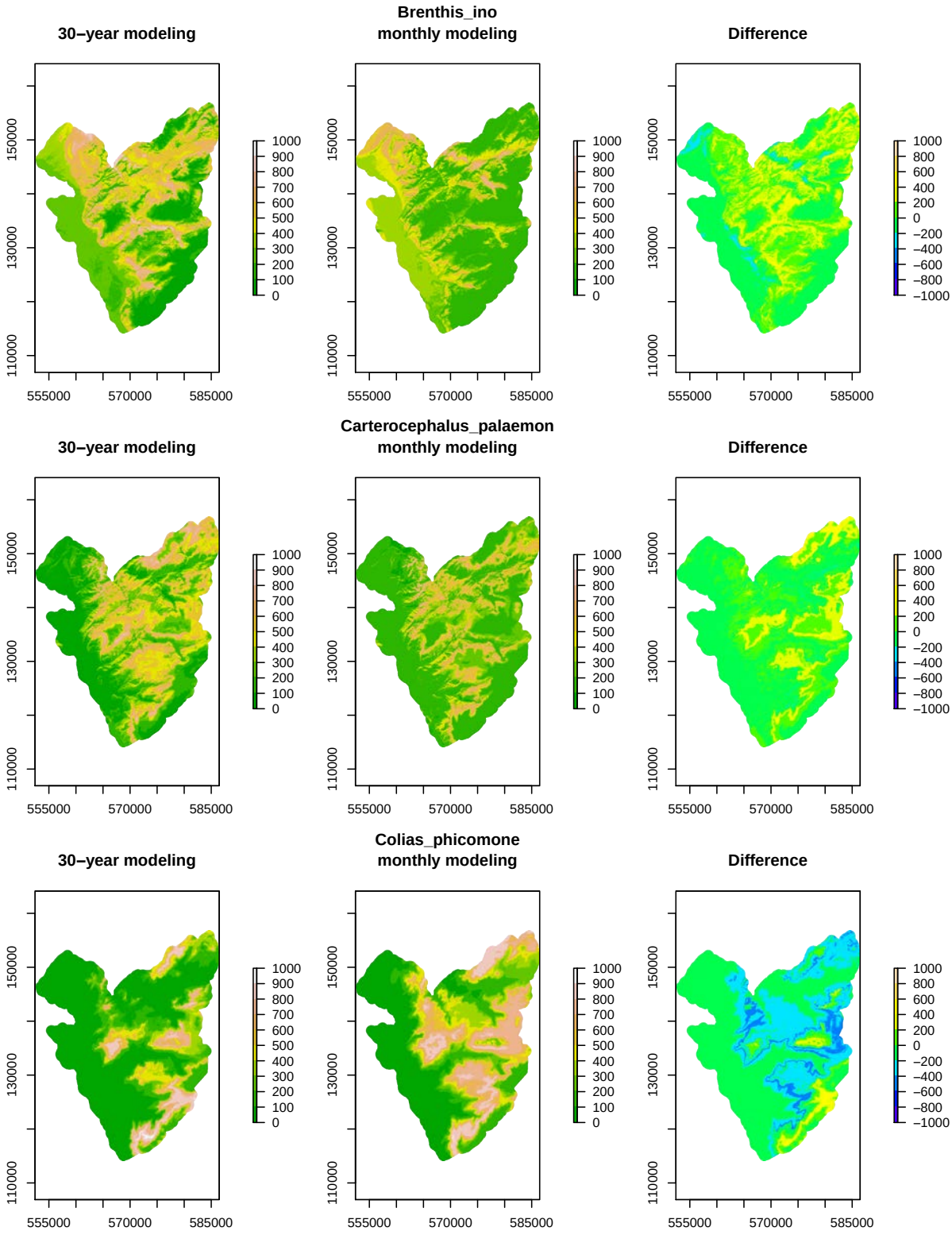

Pyrgus_serratulae
30–year modeling
1–year modeling
Difference
Thymelicus_lineola
30–year modeling
1–year modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
–200
–400
–600
–800
–1000

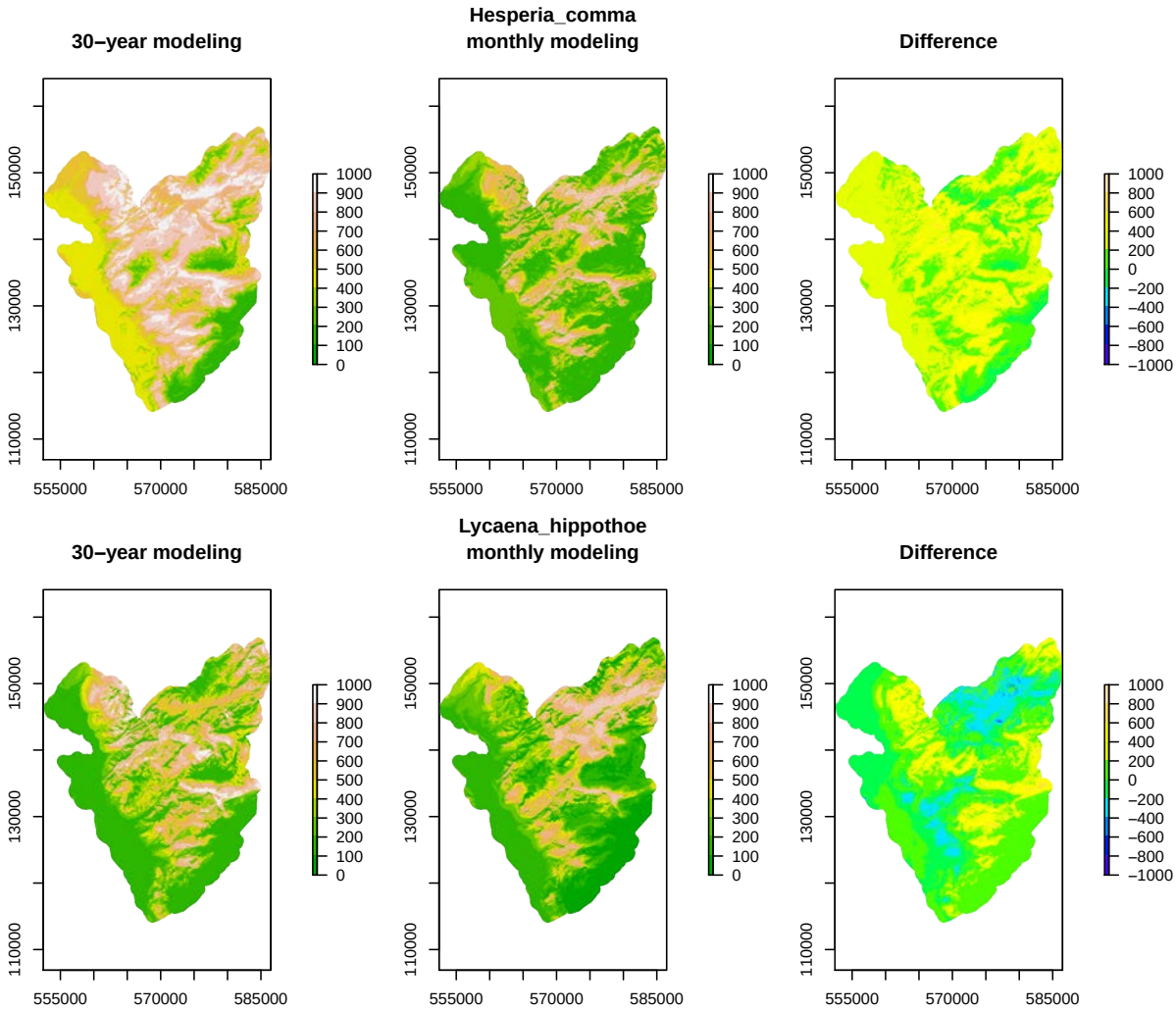
PROJECTIONS MAPS OBTAINED WITH THE SECOND APPROACH

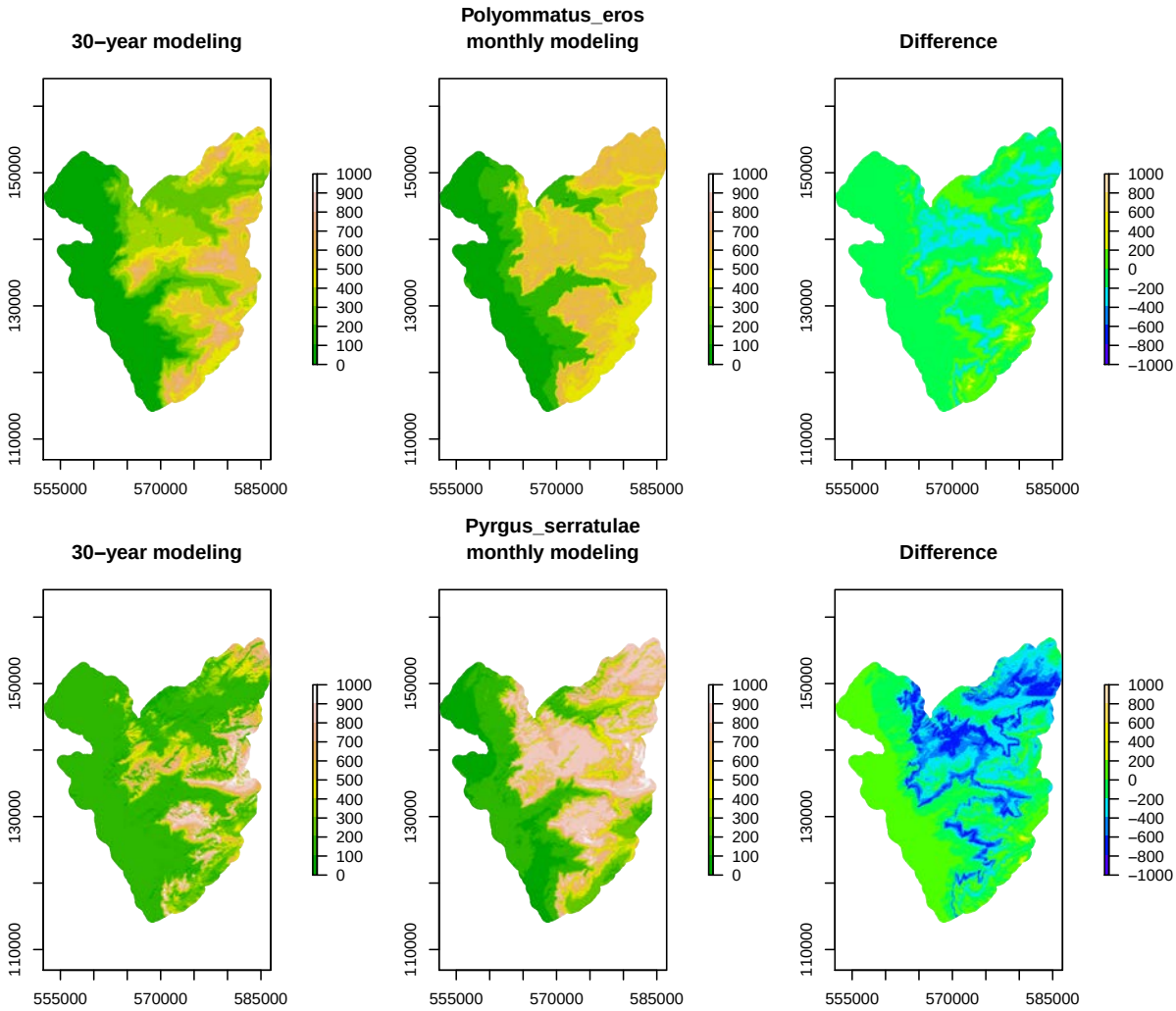
For each species, the projection map on the left is based on climatic data averaged over 30 years while that of the center is based from data precise at the month (of the year of observation). The right map shows the difference between these two projections (green = same predictions, more yellow = more presence probabilities predicted by the model 30-years, and more blue = more presence probabilities predicted by the models specific to the month)




Aricia_eumedon
30–year modeling
monthly modeling
Difference
Boloria_napaea
30–year modeling
monthly modeling
Difference
Boloria_pales
30–year modeling
monthly modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
800
600
400
200
0
–200
–400
–600
–800
–1000


Brenthis_ino
30–year modeling
monthly modeling
Difference
Carterocephalus_palaemon
30–year modeling
monthly modeling
Difference
Colias_phicomone
30–year modeling
monthly modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
–200
–400
–600
–800
–1000


30–year modeling
Hesperia_comma
monthly modeling
Difference
30–year modeling
Lycaena_hippothoe
monthly modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
–200
–400
–600
–800
–1000


Polyommatus_eros
30–year modeling
monthly modeling
Difference
Pyrgus_serratulae
30–year modeling
monthly modeling
Difference
150000
130000
110000
555000
570000
585000
1000
900
800
700
600
500
400
300
200
100
0
–200
–400
–600
–800
–1000