

HYPOTHESES

Integrating dynamic environmental predictors and species occurrences: Toward true dynamic species distribution models

Pietro Milanesi¹  | Francesca Della Rocca²  | Robert A. Robinson^{1,3} 

¹Swiss Ornithological Institute, Sempach, Switzerland

²Department of Earth and Environmental Sciences, University of Pavia, Pavia, Italy

³British Trust for Ornithology, Thetford, UK

Correspondence

Pietro Milanesi, Swiss Ornithological Institute, Seerose 1, 6204 Sempach, Switzerland.

Email: pietro.milanesi@vogelwarte.ch

Abstract

While biological distributions are not static and change/evolve through space and time, nonstationarity of climatic and land-use conditions is frequently neglected in species distribution models. Even recent techniques accounting for spatiotemporal variation of species occurrence basically consider the environmental predictors as static; specifically, in most studies using species distribution models, predictor values are averaged over a 50- or 30-year time period. This could lead to a strong bias due to monthly/annual variation between the climatic conditions in which species' locations were recorded and those used to develop species distribution models or even a complete mismatch if locations have been recorded more recently. Moreover, the impact of land-use change has only recently begun to be fully explored in species distribution models, but again without considering year-specific values. Excluding dynamic climate and land-use predictors could provide misleading estimation of species distribution. In recent years, however, open-access spatially explicit databases that provide high-resolution monthly and annual variation in climate (for the period 1901–2016) and land-use (for the period 1992–2015) conditions at a global scale have become available. Combining species locations collected in a given month of a given year with the relative climatic and land-use predictors derived from these datasets would thus lead to the development of true dynamic species distribution models (D-SDMs), improving predictive accuracy and avoiding mismatch between species locations and predictor variables. Thus, we strongly encourage modelers to develop D-SDMs using month- and year-specific climatic data as well as year-specific land-use data that match the period in which species data were collected.

KEYWORDS

climate change, dynamic predictors, land-use change, nonstationarity, R, spatiotemporal model

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

In the last decades, species distribution models (SDMs), which relate species occurrence locations with environmental predictors to estimate the probability of species occurrence to unsurveyed sites or to unsurveyed times (e.g., to predict response to climate change), experienced vigorous development at the intersection of ecology, biogeography, applied statistics, and computer science (e.g., Guillera-Aroita, 2017; Kéry, 2011). SDMs work under the assumption that species are in equilibrium with their environment; however, it is increasingly accepted that this is unrealistic in many cases, leading to the emergence of “temporal ecology” to complement the more established study of “spatial ecology” (Ryo, Aguilar-Trigueros, Pinek, Muller, & Rillig, 2019). This recognizes that the distributions of wild species are not static, but change/evolve through time and space (as do the environmental conditions in which they occur), in a hierarchical way, with longer-term interannual dynamics overlaying shorter-term intraannual dynamics. Recently developed approaches, such as spatiotemporal exploratory models (STEM; Fink et al., 2010) and dynamic occupancy models (also known as multi-season occupancy models; Kéry & Chandler, 2012), have accounted for this spatiotemporal variation of species occurrence. Specifically, STEM is based on an ensemble or a mixture of static SDMs applied at a spatiotemporally restricted extent and then averaged over the whole extent (to account for local spatial and temporal patterns, reducing misleading extrapolation to distant regions), while dynamic occupancy models describe the occurrence at each site and the colonization and extinction probabilities from the previous time step.

However, even in these recent techniques, as well as in the more traditional static SDMs, the environmental predictors considered to estimate species occurrence are basically static. For instance, most of the studies applying SDMs attempt to predict species' distribution under future climate change scenarios (Titeux et al., 2016), but through static climatic predictors; specifically, most of these studies use climatic data that are averaged over a 50- (1950–2000, Hijmans, Cameron, Parra, Jones, & Jarvis, 2005) or 30-year period (1970–2000; Fick & Hijmans, 2017). Thus, there could be a strong bias due to monthly/annual variation between the climatic conditions in which mobile species' locations were recorded and those used to develop SDMs, or even a complete mismatch in case of species locations collected after the year 2000. This problem becomes more acute with the increasing prevalence of climatic extremes (e.g., the summer of 2003 in Europe; Jentsch & Beierkuhnlein, 2008), either models using averaged baseline data are increasingly unrepresentative, or, if such events are included in the baseline, the mean values may be unduly skewed (or at least the variability about them, which is often unaccounted for, is increased).

Currently, there is wide consensus that both climate and land-use change are among the most important threats to biodiversity and ecosystem services worldwide (IPBES, 2018, 2019; Maxwell, Fuller, Brooks, & Watson, 2016; Scheffers et al., 2016), and often interact (Oliver & Morecroft, 2014). However, the impact of land-use change has rarely been explored in SDMs to date (Titeux et al., 2016)

and the few studies (Milanesi, Breiner, Puopolo, & Holderegger, 2017; Newbold et al., 2015; Radinger et al., 2016) that have considered it have used static predictors (e.g., neglecting the effect of the destruction and modification of natural habitats). Actually, while the use of climate change projections is a commonplace, land-use change projections have more rarely been used in species forecasting (Bateman et al., 2013). This may be because they are not as widely available as climatic predictions are; incorporating these will be of critical importance since the interactions among multiple drivers of global change have recently been identified as a major cause of uncertainty in climate change attribution projection (Oliver & Morecroft, 2014; Parmesan et al., 2013), in part because multiple environmental pressures may have a greater joint impact than when operating in isolation (Ostberg, Schaphoff, Lucht, & Gerten, 2015). Thus, in the absence of integrative multi-driver approaches, limited understanding of how interactions among drivers affect species distribution will be likely to hamper reliable projections (Titeux et al., 2016).

Although in some cases the distribution and/or diversity of species in different seasons may be better explained by a single predictor rather than multiple seasonally specific ones (e.g., Lennon, Greenwood, & Turner, 2000), SDMs that do not account for non-stationarity of climatic and land-use predictors may suffer from at least three problems: (a) The probability of species' occurrence at particular locations could be inaccurate due to the (temporally) averaged values in the predictors; (b) estimated slopes of predictor relationships could be biased; (c) some predictors may be wrongly identified as determinants of species' occurrence and/or may mask the real determinants, resulting, for example, in misleading inference or the wrong identification of areas of conservation importance. For example, species distribution may be negatively related to summer temperatures, but positively related to those in winter (Kawamura, Yamaura, Senzaki, Ueta, & Nakamura, 2019), patterns that would be masked by simply using annual values.

Treating the environment as static has, in part, been enforced by a lack of relevant data; however, researchers have recently developed open-access spatially explicit databases providing monthly and annual variation of climate and land-use conditions at global scale. For example, Karger et al. (2018) produced CHELSAcruts, a monthly climate (e.g., precipitation, maximum, and minimum temperature) global dataset at ~1 km spatial resolution for the period 1901–2016 (<http://chelsa-climate.org/chelsacruts/>), while Abatzoglou, Dobrowski, Parks, and Hegewisch (2018) produced TerraClimate, a monthly climate (e.g., precipitation, temperature, and wind speed) and climatic water balance (e.g., actual and potential evapotranspiration, soil moisture) global dataset at ~4 km spatial resolution for the period 1958–2015 (currently updated to the year 2017; https://climate.northwestknowledge.net/TERRA CLIMATE/index_directDownloads.php). On the other side, the European Space Agency (ESA – European Space Agency, 2017) have recently developed annual global land cover time series from 1992 to 2015 (at ~0.3 km resolution), including two levels of details and a total of 37 land-use category (e.g., different types of

croplands, forests, shrublands, and grasslands as well as urban areas, water bodies, and glaciers, etc.; <https://www.esa-landcover-cci.org/?q=node/175>).

These datasets have the potential to provide important inputs for ecological studies requiring high spatial resolution and time-varying climate and land-use data, at both local and global scales. Specifically, considering occurrences of a species (sensitive to both climate and land-use change), collected during the period 1992–2015, it is currently possible to model its distribution relating species' locations collected in a given month of a given year, matching climatic and land-use conditions. These dynamic species distribution models (D-SDMs) would be especially useful to model the occurrence of species displaying different patterns of distribution through seasons (e.g., migratory birds, insects, plants, etc.) and/or years (e.g., any species undergoing changes in its distribution).

To demonstrate how predicted patterns of species occurrence could vary between static and dynamic distribution models, we simulated a set of species recording locations ($n = 3,000$) in central Europe (5° – 11° E; 45° – 48° N) such as might be obtained from a citizen science survey. We used the ensemble predictions of four algorithms, namely boosted regression trees (BRT; Friedman, 2001), generalized additive models (GAM; Hastie & Tibshirani, 1990), generalized linear models (GLM; McCullagh & Nelder, 1989), and random forests (RF; Breiman, 2001) in R (R Development Core Team, 2013) although D-SDMs could be

developed using any appropriate algorithm. Similarly, while one can use any of the widely used validation statistics (Lecocq, Harpke, Rasmont, & Schweiger, 2019), to compare the predictive accuracy of both static and D-SDMs, in our simulations we considered the area under the receiver operating characteristic curve (AUC) and the true skills statistic (TSS). AUC ranges between 0 and 1 (worse than a random model and best discriminating model, respectively) while TSS varies between -1 and 1 (higher values indicate a good predictive accuracy, while 0 indicates random prediction). By using a random subsample of 90% of the locations to calibrate the models and the remaining 10% to evaluate them (Thuiller, Lafourcade, Engler, & Araújo, 2009), we carried out 10-fold cross-validations to test the predictive accuracy of both static and D-SDMs (see average values below).

In this example, pooling virtual species locations (e.g., collected in the period 2010–2015) and ignoring the temporal mismatch between them and the predictor variables (from the WorldClim 2 database; Fick & Hijmans, 2017) leads to biased estimation of species distribution (overestimated in this case; AUC = 0.893; TSS = 0.809). On the other side, we developed D-SDMs splitting virtual species locations and accounting for year-specific predictors (from the CHELSAcruts database; Karger et al., 2018) which showed, together with interannual fluctuations, a relatively limited average species distribution (Figure 1; AUC = 0.968; TSS = 0.898). One of the main drawbacks of this approach is that for each time period (i.e., year in our example) the number of species locations could be much lower

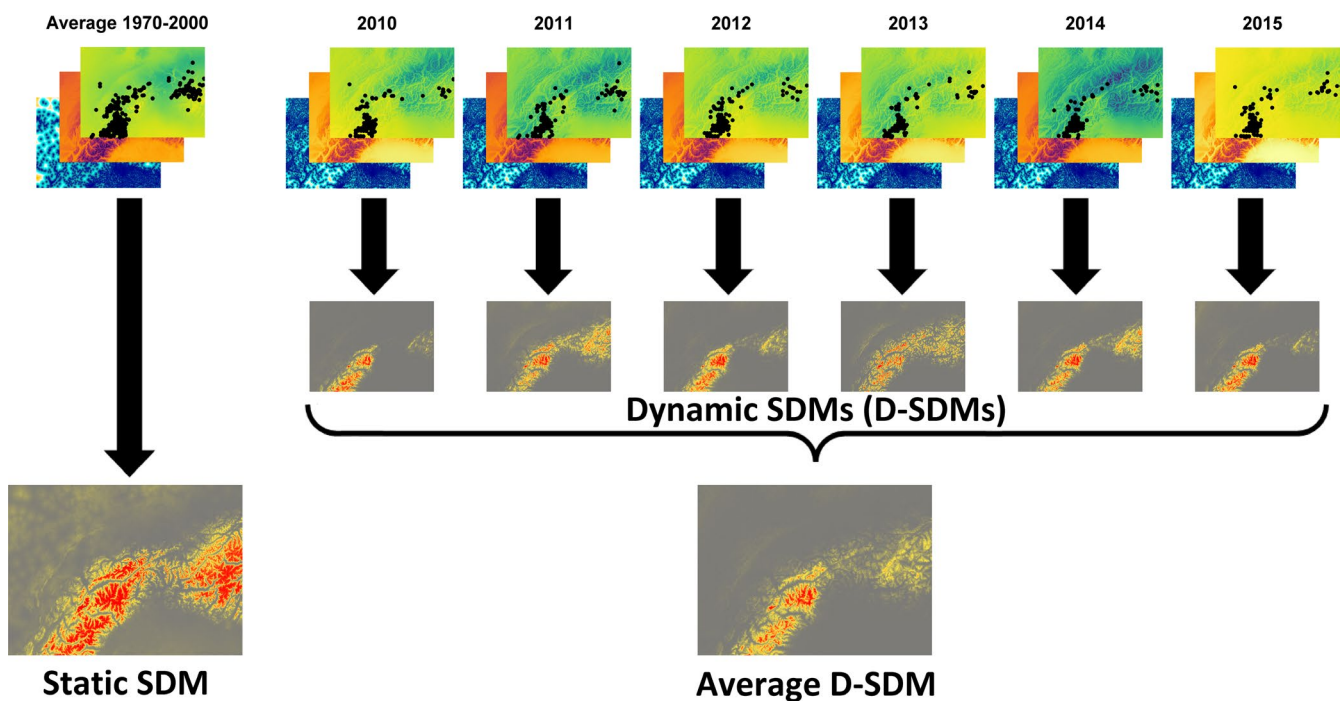


FIGURE 1 Differences between static and dynamic species distribution models. Predictor variables (in this case Bio 3—Isothermality, Bio 10—Mean temperature of the warmest quarter, and Bio 17—Precipitation of the driest quarter) came from WorldClim 2 database (Average 1970–2000; left) and CHELSAcruts (annually for 2010–2015, right), together with relative virtual species locations (black dots) collected in the period 2010–2015, are shown on the first line. Resulting maps of dynamic (annual) species distribution models are shown on the second line while those of static and averaged dynamic species distribution models are shown on the third line. Red-gray scale indicates high–low probability of occurrence

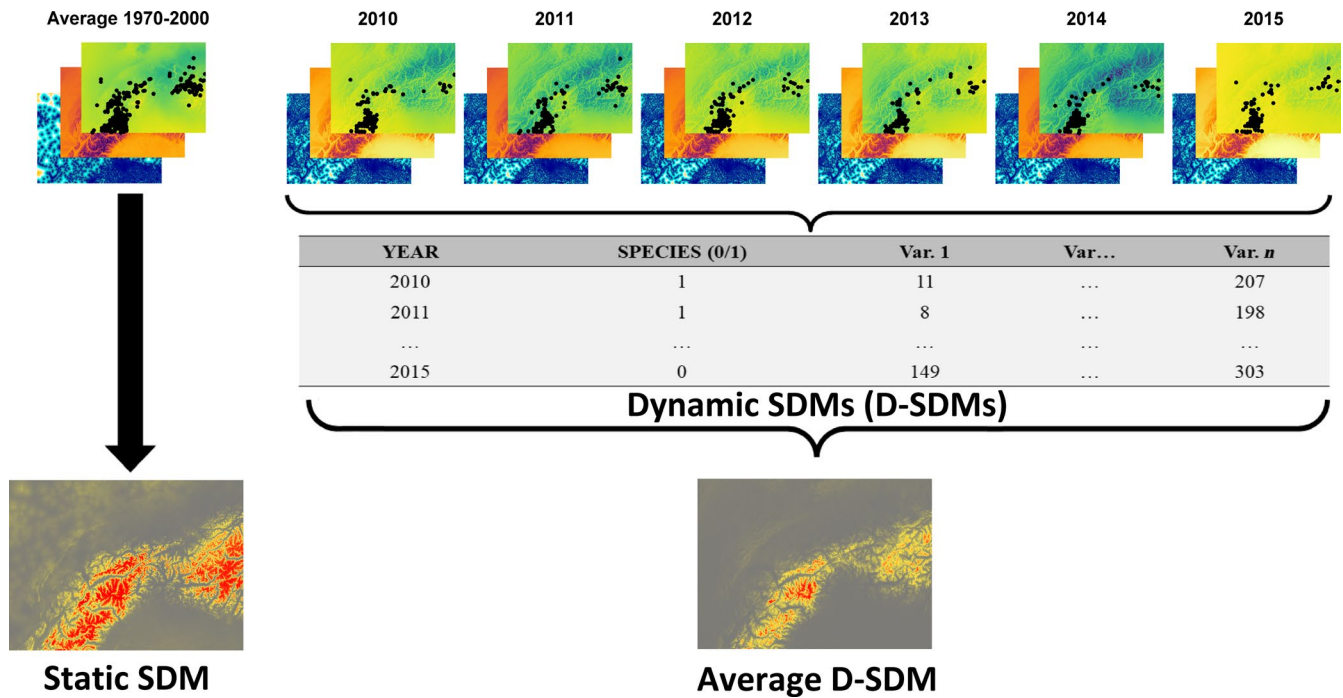


FIGURE 2 Differences between static and averaged dynamic species distribution models. Predictor variables and static model as in Figure 1, for the dynamic model year-specific predictor values are pooled for 2010–2015 into a single database, shown on the second line. Resulting map of static and averaged dynamic species distribution models is shown on the third line. Red-gray scale indicates high–low probability of occurrence

than those of static SDMs (which pool data) leading, potentially, to a higher influence of unusual observations (which may be diluted by pooling) and possible model overfitting (i.e., model fits the calibration data too closely, in environmental space, but fails to predict independent evaluation data accurately; Radosavljevic & Anderson, 2014). To overcome this, we pooled virtual species locations but associated each location to its relative year-specific predictor values and thus developed a single SDM to predict virtual species distribution under different year-specific conditions, which again showed a more restricted average species distribution of D-SDMs compared with the static SDMs (Figure 2; AUC = 0.943; TSS = 0.871). As well as reducing the likelihood overfitting, this approach also has the benefit of accounting for interannual variation (e.g., species adaptability to climatic variations) in the prediction of species distribution under different year-specific conditions and thus also improves estimation during climatic anomalies (which have become larger and more frequent in the last decade due to climate change; Ummenhofer & Meehl, 2017). The recently developed “ensemble of small models” approach (Breiner, Guisan, Bergamini, Nobis, & Anderson, 2015) provides an alternative way to reduce overfitting.

Thus, D-SDMs can show higher predictive accuracy compared with static SDMs and could also improve the outcomes of STEM and dynamic occupancy models, providing more robust trends of species occurrence. For these reasons, we strongly encourage modelers to develop D-SDMs in order to provide more accurate estimates of species distribution. Such approaches are likely to be particularly valuable in more climatically variable regions and habitats (e.g.,

Reside, Wal, Kutt, & Perkins, 2010) and may become more important as climatic variability is predicted to increase (Kharin, Zwiers, Zhang, & Hegerl, 2007). However, species data quality (i.e., availability of information about the year and month in which species locations were collected), as well as the lack of open-access databases for nonterrestrial environments (e.g., pH concentration for freshwater biotopes), can strongly limit the application of D-SDMs and thus further improvements of both species data quality and accessibility of additional spatiotemporal environmental data are needed (e.g., Wetzel et al., 2018).

Finally, we urge researchers to make use of newly available datasets coming online to include both climate and land-use dynamic predictors, as the strength of impacts on biodiversity will likely depend on the interaction between these (Oliver & Morecroft, 2014). This will ensure, perhaps in combination with recently developed methods for generating high-resolution climate projections (Maclean, 2019), that future decision-making, such as prioritizing areas for conservation (e.g., Jones, Watson, Possingham, & Klein, 2016), more robustly anticipates the response of biodiversity to future climate and land-use changes.

ACKNOWLEDGMENTS

The comments of the Editor-in Chief Gareth Jenkins and those of two anonymous reviewers greatly helped improving this paper.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

PM conceived and designed the overall study, PM and RAR wrote the manuscript, FDR and PM conducted the statistical analyses and designed the figures. FDR prepared the figures. All three authors contributed substantially to revisions of the paper.

DATA AVAILABILITY STATEMENT

Climatic GIS layers used for analysis in this paper are freely available at <http://chelsa-climate.org/chelsacruts/> and <http://worldclim.org/version2>.

ORCID

Pietro Milanese  <https://orcid.org/0000-0002-1878-9762>

Francesca Della Rocca  <https://orcid.org/0000-0002-7617-9276>

Robert A. Robinson  <https://orcid.org/0000-0003-0504-9906>

REFERENCES

- Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, 5, 170191. <https://doi.org/10.1038/sdata.2017.191>
- Bateman, I. J., Harwood, A. R., Mace, G. M., Watson, R. T., Abson, D. J., Andrews, B., ... Termansen, M. (2013). Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science*, 341, 45–50. <https://doi.org/10.1126/science.1234379>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiner, F. T., Guisan, A., Bergamini, A., Nobis, M. P., & Anderson, B. (2015). Overcoming limitations of modelling rare species by using ensembles of small models. *Methods in Ecology and Evolution*, 6, 1210–1218. <https://doi.org/10.1111/2041-210X.12403>
- ESA – European Space Agency. (2017) 300 m annual global land cover time series from 1992 to 2015. Retrieved from <https://www.esa-landcover-cci.org/?q=node/175>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37, 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fink, D., Hochachka, W. M., Zuckerberg, B., Winkler, D. W., Shaby, B., Munson, M. A., ... Kelling, S. (2010). Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications*, 20, 2131–2147. <https://doi.org/10.1890/09-1340.1>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29, 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Guillera-Aroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: Advances, challenges and opportunities. *Ecography*, 40, 281–295. <https://doi.org/10.1111/ecog.02445>
- Hastie, T., & Tibshirani, R. (1990). *Generalized additive models*. London, UK: Wiley Online Library.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978. <https://doi.org/10.1002/joc.1276>
- IPBES (2018). *The regional assessment report on biodiversity and ecosystem services for Europe and Central Asia*. Bonn, Germany: Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.
- IPBES (2019). *Global Assessment Report on Biodiversity and Ecosystem Services*. Bonn, Germany: Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.
- Jentsch, A., & Beierkuhnlein, C. (2008). Research frontiers in climate change: Effects of extreme meteorological events on ecosystems. *Comptes Rendus Geoscience*, 340, 621–628. <https://doi.org/10.1016/j.crte.2008.07.002>
- Jones, K. R., Watson, J. E., Possingham, H. P., & Klein, C. J. (2016). Incorporating climate change into spatial conservation prioritisation: A review. *Biological Conservation*, 194, 121–130. <https://doi.org/10.1016/j.biocon.2015.12.008>
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M. (2018). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 170122. <https://doi.org/10.1038/sdata.2017.122>
- Kawamura, K., Yamaura, Y., Senzaki, M., Ueta, M., & Nakamura, F. (2019). Seasonality in spatial distribution: Climate and land use have contrasting effects on the species richness of breeding and wintering birds. *Ecology and Evolution*, 9, 7549–7561. <https://doi.org/10.1002/ece3.5286>
- Kéry, M. (2011). Towards the modelling of true species distributions. *Journal of Biogeography*, 38, 617–618. <https://doi.org/10.1111/j.1365-2699.2011.02487.x>
- Kéry, M., & Chandler, R. (2012). Dynamic occupancy models in unmarked. Retrieved from <http://cran.r-project.org/web/packages/unmarked/vignettes/colectx.pdf> (Accessed 20 April 2015).
- Kharin, V. V., Zwiers, F. W., Zhang, X., & Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, 20, 1419–1444. <https://doi.org/10.1175/JCLI4066.1>
- Lecocq, T., Harpke, A., Rasmont, P., & Schweiger, O. (2019). Integrating intraspecific differentiation in species distribution models: Consequences on projections of current and future climatically suitable areas of species. *Diversity and Distributions*, 25, 1088–1100. <https://doi.org/10.1111/ddi.12916>
- Lennon, J. J., Greenwood, J. J. D., & Turner, J. R. G. (2000). Bird diversity and environmental gradients in Britain: A test of the species–energy hypothesis. *Journal of Animal Ecology*, 69, 581–598. <https://doi.org/10.1046/j.1365-2656.2000.00418.x>
- Macleod, I. M. D. (2019). Predicting future climate at high spatial and temporal resolution. *Global Change Biology*, 25, 14876. <https://doi.org/10.1111/gcb.14876>
- Maxwell, S. L., Fuller, R. A., Brooks, T. M., & Watson, J. E. M. (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature*, 536, 143–145. <https://doi.org/10.1038/536143a>
- McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. London, UK: Chapman and Hall.
- Milanesi, P., Breiner, F. T., Puopolo, F., & Holderegger, R. (2017). European human-dominated landscapes provide ample space for the recolonization of large carnivore populations under future land change scenarios. *Ecography*, 40, 1359–1368. <https://doi.org/10.1111/ecog.02223>
- Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., ... Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45–50. <https://doi.org/10.1038/nature14324>
- Oliver, T. H., & Morecroft, M. D. (2014). Interactions between climate change and land use change on biodiversity: Attribution problems, risks, and opportunities. *Wiley Interdisciplinary Reviews: Climate Change*, 5, 317–335. <https://doi.org/10.1002/wcc.271>
- Ostberg, S., Schaphoff, S., Lucht, W., & Gerten, D. (2015). Three centuries of dual pressure from land use and climate change on the biosphere. *Environmental Research Letters*, 10, 44011. <https://doi.org/10.1088/1748-9326/10/4/044011>
- Parmesan, C., Burrows, M. T., Duarte, C. M., Poloczanska, E. S., Richardson, A. J., Schoeman, D. S., & Singer, M. C. (2013). Beyond climate change attribution in conservation and ecological research. *Ecology Letters*, 16, 58–71. <https://doi.org/10.1111/ele.12098>

- R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Radinger, J., Hölker, F., Horký, P., Slavík, O., Dendoncker, N., & Wolter, C. (2016). Synergistic and antagonistic interactions of future land use and climate change on river fish assemblages. *Global Change Biology*, 22, 1505–1522. <https://doi.org/10.1111/gcb.13183>
- Radosavljevic, A., & Anderson, R. P. (2014). Making better MAXENT models of species distributions: Complexity, overfitting and evaluation. *Journal of Biogeography*, 41, 629–643. <https://doi.org/10.1111/jbi.12227>
- Reside, A. E., van der Wal, J. J., Kutt, A. S., & Perkins, G. C. (2010). Weather, not climate, defines distributions of vagile bird species. *PLoS ONE*, 5, 13569. <https://doi.org/10.1371/journal.pone.0013569>
- Ryo, M., Aguilar-Trigueros, C. A., Pinek, L., Muller, L. A., & Rillig, M. C. (2019). Basic principles of temporal dynamics. *Trends in Ecology & Evolution*, 34, 723–733. <https://doi.org/10.1016/j.tree.2019.03.007>
- Scheffers, B. R., De Meester, L., Bridge, T. C. L., Hoffmann, A. A., Pandolfi, J. M., Corlett, R. T., ... Watson, J. E. M. (2016). The broad footprint of climate change from genes to biomes to people. *Science*, 354, aaf7671. <https://doi.org/10.1126/science.aaf7671>
- Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*, 32, 369–373.
- Titeux, N., Henle, K., Mihoub, J., Regos, A., Geijzendorffer, I. R., Cramer, W., ... Brotons, L. (2016). Biodiversity scenarios neglect future land-use changes. *Global Change Biology*, 22, 2505–2515. <https://doi.org/10.1111/gcb.13272>
- Ummenhofer, C. C., & Meehl, G. A. (2017). Extreme weather and climate events with ecological relevance: A review. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372, 20160135.
- Wetzel, F. T., Bingham, H. C., Groom, Q., Haase, P., Kõljalg, U., Kuhlmann, M., ... Häuser, C. L. (2018). Unlocking biodiversity data: Prioritization and filling the gaps in biodiversity observation data in Europe. *Biological Conservation*, 221, 78–85. <https://doi.org/10.1016/j.biocon.2017.12.024>

How to cite this article: Milanesi P, Della Rocca F, Robinson RA. Integrating dynamic environmental predictors and species occurrences: Toward true dynamic species distribution models. *Ecol Evol*. 2020;10:1087–1092. <https://doi.org/10.1002/ece3.5938>