

Supplementary Material: Appendices

1. Appendix 1: Tools for organism-based ecology: data, traits, OEUs, and models

1.1. Data

To empirically test OE hypotheses, one needs information on the properties of organisms of the community under study, or at least a representative sample of those organisms. In field studies, the number of organisms in a sample may be very high and the assessment of the properties of all these organisms a tremendous task.

Since the end of the last century, a huge amount of new ecological and nature conservation data has become available in datasets like Bac*Dive*, CESTES, EzTaxon, IUCN Red lists, GBIF, PREDICTS, SPI-Birds and TRY. These datasets are quickly expanding. Most of the information is on species and species populations. Fortunately, some information on the properties of individual organisms is also available (e.g., GBIF has more than 1.39 billion records of the occurrence of organisms [https://www.gbif.org/; assessment March 19th, 2020]; see also the SPI-Birds Network and Database [https://nioo.knaw.nl/en/spi-birds#quicktabs-qt_hole_nesting_network=0] and the TRY database [https://www.try-db.org/TryWeb/Home.php] [115]).

A great prospect for OE lies in new observation techniques that collect information on the properties of multiple organisms *in situ* simultaneously. For example, flow cytometry, single-cell ICP-MS, single-cell metabolics or nanosims can be used to count and measure properties of micro-organisms [116]; for vertebrates and invertebrates, camera (including smartphone cameras in citizen science projects) and microphone traps coupled to data science methods can be used to identify properties; and for plants, (airborne) hyperspectral imaging and LiDAR collecting information on chemical, morphological and structural traits is available [117]. Each of these techniques collect information on the properties of organisms and may, therefore, give a boost to OE [46].

Statistical analyses of these datasets of properties of organisms and features of location and timing are not fundamentally different from analyzing taxon-based datasets. However, because these datasets could be huge, new statistical analyzing techniques for 'big data' based on machine learning could be helpful [24,118,119].

1.2. Traits

It is obvious that organisms may differ greatly in their trait values, and that this is ecologically very relevant. The issue is studied in Trait-based Ecology (TE)

[19,32,120,121,122,123,124,125,126,127,128]. Especially in plant ecology there is a long tradition of TE, also presented under the name 'functional diversity' [32,63,129,130,131,132,133]. However, in many TE studies, the trait values are not actually measured, but gathered from existing datasets or taxonomic descriptions, and then the means and variances of the trait values are attributed to the taxa studied (e.g., [32,134,135,136]). In those cases, the taxon is the research unit, not the organism, and trait variance within the taxon is ignored [19,62,65].

Other studies combine trait-based and organism-based approaches [137]. Much of this research focusses on the ecological relevance of intraspecific variability (e.g., [31,55,65,138,139]). Clutton-Brock and Sheldon [140] reviewed long-term studies of individually recognizable birds and mammals and concluded that these studies play an important role in ecology and evolutionary biology. Recent studies using agent-based modelling (ABM) cover a diversity of issues, varying from the assessment of wildlife diseases [141] to eradication of invasive small mammals [77], and coexistence of small species [59]. ABM has been used to predict the effect of changes in ground water salinity due to sea level rising on individual adaption of trees [33,47]. Shifts in isotope ratios have been successfully used as a proxy for changes in traits of individual trees to study the effect of changes in salinity of ground water on interactions between plant communities [47,48].

A key question is which traits to include in OE [65]. The list of potential traits is long [19,32,142,143,144]. However, some traits may not be relevant for all organisms and traits may have high collinearity, which could veil causality. A limited list of standardized traits would be very helpful. It would enable meta-analyses of data collected around the world. Selected traits should a) be present in all organisms, b) be essential for the survival and reproduction, i.e., for the fitness of the organism, and c) express the effect that the organism has on the fitness of other organisms. In other words, they should be essential for the performance of the organism and its relationship to its biotic and abiotic environment [145,146]. And preferably they should be weakly correlated. So, we are looking for 'Independent Essential Organism Traits'. 'Body size' and 'metabolism rate' are proven examples of such traits [37,80]. In addition, we consider traits that express resource uptake and waste output, reproduction investment, moment of appearance in and disappearance from an area (including birth and death), and movement through the area. We challenge ecologists to try to define a comprehensive list of Independent Essential Organism Traits and their measurement.

1.3. OEUs

Assigning organisms as research and measurement units may have far-reaching consequences for study designs in ecology. Assessing the effect of abiotic factors on highly divers organisms may be challenging. And all potential interactions between organisms within a community will be an even more frighteningly huge number of interactions than the interactions between species.

For the study of interactions between organisms, ways of grouping the organisms are needed. Organisms are in their striving to survive or reproduce limited by their genotype and experience. We can try to capture the genotype and experience of an organism in a set of measurable properties of organisms given our research questions, i.e., in the ecological relevant features of their phenotypes. For that, we revitalize the 'Operational Ecological Unit' (OEU) [57] under a slightly new definition, viz as a set of organisms that are ecologically alike, i.e., all organisms of an OEU have properties that lay in the same predefined trait value range for a predefined, limited set of traits. For example, the autotrophs that live in fresh water could be divided in a range of OEUs based on body size. Or the arthropods of a grassland could be assigned to OEUs according to their food preferences, such as herbivores, omnivores, predators, and detrivores.

In ecology, 'functional' classifications are widely used and known as 'guild', 'functional group', 'Plant Functional Type (PFT)', 'trait syndromes', or the like [19,115,147,148]. However, these functional groups are usually groups of species, not of individual organisms. As a consequence,

variance in a relevant trait may be large within such groups [19,148]. For example, the simple growth form PFT defines 'herbs', 'bushes' and 'trees', which seem to coincide with plant height groups. But a seedling of an oak is regarded as a very small tree, not as an herb, attributing to large variance in plant height within the PFT. To stress the importance of grouping organisms instead of species, we propose OEU to be used in OE.

While studying the interactions between all organisms of a community is impractical, and usually not feasible, hypotheses of the form "OEU **O**, i.e., organisms with the set of trait values **O**, interact in such and such way with OEU **P**, i.e., organisms with the set of trait values **P**" do allow testing.

Note that OEUs are pragmatic sets that are defined for answering spatiotemporal specific research questions, i.e., for empirically testing predictions. They can be any ecological relevant subsets of a community. They can bring together organisms that belong to different species, but also subsets of species, as long as they have more or less the same properties. For example, males and females can be different OEUs, as can larvae and adults, or organisms of different body size.

It is crucial to actually be able to assess the trait values of the organisms in field studies [19,65]. For analyzing field data, OEUs offer the same statistical possibilities and limitations as the often used 'Operational Taxonomic Units' (OTUs). OEUs can replace species in METE, so that the ASNE-model would become an AONE-model, predicting how metabolic energy (E) is distributed over individuals (N), and how individuals are distributed over OEUs (O) and area (A) [98].

In order to reduce computing time, in computer models the use of 'super-individuals' has been proposed, which are in many respects equivalent to OEUs [87,149]. But OEUs can also be used for empirical studies, because they are sets that have emergent properties which can differ between spatiotemporal subsets of the same OEU and which can be assessed, depending on the research question. In fact, OEUs can be regarded as having the same analytic role in OE as species have in OSCE. The size of an OEU could be the number of individuals, but also its total biomass. At any time, an OEU may have a different spatial range, so that an OEU could colonize an area or become extinct. Also, the distribution of individuals over values of both defining and non-defining traits may be different between locations or may change over time. For example, the ratio of day to night active large mammal predators may change as a result of human activities [150].

1.4. Models

The long tradition of using models for predicting and simulating communities will help to further develop OE (e.g., [46,151]). In ecology, models do not provide proof of general patterns themselves because, although they contain parameters that relate dependent and independent variables, the parameters are not well-established constants as in physics. The models are descriptive and their parameters need to be estimated from empirical data by model fitting. This usually results in parameter values that are highly contingent [15,152,153]. But the models can be used for generating patterns, of which the generality can then be studied [34].

Based on our reasoning, we think that a distinction between two types of models is relevant here: species-based models (SBM) and agent-based models (ABM, including the closely related spatial-explicit models).

SBM are rooted in Lotka-Volterra models and calculate the development of communities based on abstract population parameters like birth rates, death rates, and conversion efficiencies, which makes them essentially species-based. The models can be extended for location and time specific networks, but then tend to become overly complex (e.g. [154]).

In contrast, ABM, which are also called individual-based models, simulate the development of populations and communities based on the properties of individual organisms, locations and time [34,35,36,41,59,72,139,155]. By focusing on ecologically relevant entities (for this reason these models are called 'agent-based'), ABMs aim to simulate the actual processes that take place in an ecosystem [46]. To incorporate the feedbacks and uncertainty that are present in real ecological systems, a trend has been to simulate how proximate decisions are made based on the individual's current state and short-term predictions. DeAngelis and Diaz [33] discuss the different modelling approaches and prospects. Zakharova et al. [137] plead for combining trait-based modelling and ABM in order to simplify the parameterization of models.

In that regard, it is interesting that in some of the functional–structural plant models (FSPMs) [156], building blocks, or metamers, are used to describe the individual plant. These models have been around for a couple of decades, and could be a starting point for a type of OE at least in plant ecology. The agents in this case are modules (metamers) within the plant and algorithms are used to model repetitive branching patterns by which the plants grow, with different morphologies for different species. These methods arose out of interest in the development of 3-D spatial structure of individual woody and herbaceous plants and in the visualization of their structures. Because the development of plant structure interacts with light capture and material flows, these can be integrated into a general methodology that combines geometric structure with function. i.e., FSPM's, in which plant architecture, including roots, is explicitly modelled in 3-D and linked to process-based modelling of functional aspects, such a light capture, and carbon and nutrient flows (e.g., [157,158]). This approach, while focusing on the internal structure and function of plants, is also used at the scale of multi-plant combinations to provide better understanding of the details of plant interactions. Overlapping of objectives was seen in the survey of several types of applications, which showed that population-level ABMs and FSPMs are addressing some of the same issues of primary production, biodiversity, spatial heterogeneity and pest problems, although at different spatial scales. For example, although FSPMs do not deal with population growth through seed dispersal, they do simulate fruit and thus fecundity. It is possible that the time is right to make use of the complementarity of the approaches to facilitate overall progress in ABMs and OE. One sign is that modelling methodologies such as Pattern-Oriented Modelling [88] and Overview, Design and Details [89] are being adopted in some FSPM papers (e.g., [159]).

ABMs are in a developmental stage [36]. At the moment, they usually have numerous parameters: the diversity of species is extended to the diversity of properties of organisms, so that species are still part of ABMs (e.g., [59]). So, they often do not yet result in generalities over different kind of organisms. But we expect that, because ABMs try to simulate processes, they are able to show real patterns [160,161]. These patterns and the knowledge of the structure of the ABM can then be used to develop a theory that predicts these patterns.

Next, hypotheses are to be deduced from such a theory that can be tested with empirical data. The testing of the Neutral Theory of Biodiversity is a good example of this procedure [162,163]. We also think that, in the end, ABMs will need fewer parameters than SBMs, because we expect that the

number of relevant traits that describe organisms will be much lower than the number of species [89] (Appendix 2). Ultimately this may lead to accurate predictions of communities [25,160,164].

2. Appendix 2: The temporal change of North American migration birds studied by Organism-based Ecology

Recently, Weeks et al. [55] showed that over a period of 40 years, body size of North American migratory bird species decreased and wing length increased, probably in response to climate change. They focused on differences between species, but the pattern seemed consistent across species. We wanted to know whether OE was able to show us this general pattern of body size increase and wing length decrease, even if we had no knowledge on the species of the individual birds. For answering this question, we reanalyzed the dataset of Weeks and others in two ways: the traditional use of Linear Mixed Models (LMM) and the use of Maximum Entropy Theory of Ecology (METE).

LMM estimate the parameters of a linear, additive equation that describes the relationship between dependent and independent variables by fitting an equation optimal to observations under the assumption that the deviations of the observation from the equation, the residuals, are normally distributed. The equation predicts the values of the dependent variable and the parameters represent the effect sizes. The models are called mixed because they correct for unbalanced sampling by including random effect variables.

METE is not an alternative for LMM, but a way of generating unbiased neutral null-models with a limited number of constraints based on the first principle that in a system that is in equilibrium, particles will be distributed such that maximum entropy will be reached. In ecology, the individual organisms are regarded as the particles and the system is a community. The METE models predict the distribution of organisms when the community is in equilibrium, i.e., when the distribution of the organisms changes only due to random movements, and the organisms are neutral, i.e., they have no properties that affect each other's distribution. In that case, the system is accurately described by its constraints. Therefore, deviations of observation from the predictions indicate that the system is not sufficiently described by its constraints. These deviations can then be analyzed to find out whether factors push the community out of equilibrium, organisms have non-neutral properties, or the community has spatiotemporal properties that are not defined in the constraints. Here, we will analyze the deviations with an LMM, but better ways of analyzing may be available. Also, new ways of analyzing may need to be developed.

In the following analyses, we explore these different approaches for analyzing ecological data.

2.1. Traditional use of Linear Mixed Models

In order to study the change over time of the tarsus length, a proxy of body size, and the wing length of migration birds, we applied four Linear Mixed Models to the observations of Weeks and others, two including information on the species to which the organisms belong and two without species information. We included sex and age class of the birds as possible independent variables. By doing so we acknowledged that sex and age define four Operational Ecological Units (OEUs): female hatchlings, male hatchlings, female post-hatchlings, and male post-hatchlings.

The estimated effect size of Year on the log-transformed tarsus and wing length showed that the models without species as factor were able to show an effect of Year on tarsus and wing length, although their conditional and marginal R^2 's were low (about 1% resp. 0.6% for tarsus length and 3% resp. 2.4% for wing length; Table A2.1). Low R^2 's should not surprise us, knowing that the data are the outcome of an extremely noisy process: the fatal collision of a migrating birds with buildings. What should surprise us is that models including species showed very high conditional and low marginal R^2 , and lowest log-likelihood of the models including Year, which means that the models without Year are actually the best models. The very high conditional R^2 indicate overfitting. This problem is absent in the models without taxonomic information. So, ignoring species results in reliable models. Graphs of the raw data of the change of length suggested that the rate of length change altered somewhere in the years 2000-2005 (Fig. A2.1). Tarsus length seemed to decrease over the years, but this decrease accelerated most strongly in birds in their hatching year, probably after 2003. Wing length seemed to increase, but only after 2000 and only in birds after their hatching year and had, therefore, experienced at least one migration.

Tarsus length	Year	Cond. R ²	Marg. R ²	#Df	LogLik	Df	χ^2	p-value	
Model incl. taxon	incl.	0.980	0.003	13	119213	-6	41.52	<0.001	***
	excl.	0.979	0.002	7	119234				
Model excl. taxon	incl.	0.010	0.006	10	23228	-4	15.76	0.003	**
	excl.	0.010	0.005	6	23220				
Wing length									
Model incl. taxon	incl.	0.985	0.013	13	137416	-6 315.98	<0.001 **	***	
	excl.	0.988	0.010	7	137574				
Model excl. taxon	incl.	0.030	0.024	10	15528	-4	175.45	< 0.001	***
	excl.	0.028	0.019	6	15441				

Table A2.1: Effect of year on tarsus length (n=58,497) and wing length (n=62,628) in LMM models.

The grey area shows the Likelihood Ratio Test that include taxonomic information as a random effect variable:

Model incl. taxon and year: $\ln(\text{Length}) \sim 1 + \text{Year} * \text{Age} * \text{Sex} + (1 | \text{Taxon}) + (1 | \text{factor}(\text{Year}));$

Model incl. taxon, excl. year: $ln(Length) \sim 1 + Age * Sex + (1 | Taxon) + (1 | factor(Year))$.

The white area shows the Likelihood Ratio Test without taxonomic information:

Model excl. taxon and year: $\ln(\text{Length}) \sim 1 + \text{Year}^*\text{Age} * \text{Sex} + (1 | \text{factor}(\text{Year}));$

Model excl. taxon, excl. year: $\ln(\text{Length}) \sim 1 + \text{Age }^* \text{Sex} + (1 \mid \text{factor}(\text{Year}))$.

The best model, with or without Year as fixed effect variable, is indicated in bold in the Year column. In all cases the fixed variables include Age, Sex, and all the interactions, while the factor Year, i.e., year as a categorical variable, is a random effect variable. Models were estimated using the function *lmer()* of the package *lme4* (version 1.1-27) [165]; R² estimations with *performance::r2* of the package *performance* (version 0.7.2) [166]; LRT was done with *lrtest()* of the package *lmtest* (version 0.9-38) [167], all in R version 4.0.3 [56].



Figure A2.1: Smoothed nonlinear regression lines of the raw relationship between Year and **a**) Tarsus and **b**) Wing length in mm for birds in their hatching year and after their hatching year. Lines were constructed with the default settings of the function *scatterplot()* of the package *car* (version 3.0-10) [168] in R version 4.0.3 [56].

For this reason, we decided to include a new variable, Period, that separates the time before 2000 from 2000 onwards, in our model. The results showed that the rate of change was indeed different before and after 2000 (Table A2.2). The model for ln(Wing length) predicted that wing length did not change in hatchlings, but increased up to 2000 in older birds and decreased slightly after 2000 (Fig. 2 in main text). So, wing length change is only visible in birds that migrated at least once. We might be tempted to conclude that wing length change is due to alterations during migration, although also non-migrating birds may show changing wing length [55,169].

Table A2.2: Effect sizes of the fixed variables, including their interactions, on tarsus and wing length, as estimated by the function *lmer()* of the package *lme4* [165] in R version 4.0.3 [56].

ln(Tarsus length)

Cond. R² Marg. R² df LogLik AIC

	0.010	0.007	58461	23236.4	-46436.8	
	Estimate	Std. Error	df	t-value	Pr(> t)	
(Intercept)	6.04900	1.40500	82.77	4.306	0.000	***
Age	1.24700	1.28900	58000	0.967	0.333	
Sex	-0.79320	1.36500	58460	-0.581	0.561	
Year	-0.00144	0.00070	82.68	-2.064	0.042	*
Period	0.07118	1.69900	91.61	0.042	0.967	
Age:Sex	-1.13100	1.80700	58460	-0.626	0.532	
Age:Year	-0.00063	0.00064	58000	-0.981	0.327	
Sex:Year	0.00040	0.00068	58460	0.59	0.555	
Age:Period	-3.56900	1.71200	57620	-2.085	0.037	*
Sex:Period	-1.49600	1.67200	58470	-0.895	0.371	
Year:Period	-0.00005	0.00085	91.64	-0.056	0.956	
Age:Sex:Year	0.00057	0.00090	58460	0.634	0.526	
Age:Sex:Period	5.17600	2.38300	58470	2.172	0.030	*
Age:Year:Period	0.00179	0.00086	57620	2.094	0.036	*
Sex:Year:Period	0.00076	0.00083	58470	0.907	0.365	
Age:Sex:Year:Period	-0.00260	0.00119	58470	-2.185	0.029	*
ln(Wing length)						
	Cond. R ²	Marg. R ²	df	LogLik	AIC	
	0.032	0.027	62610	15571.4	-31106.8	
	Estimate	Std. Error	df	t-value	Pr(> t)	

(Intercept)	-2.81300	1.74800	74.87	-1.609	0.112	
Age	7.00500	1.43900	62430	4.868	0.000	***
Sex	-1.63500	1.50900	62610	-1.083	0.279	
Year	0.00353	0.00087	74.79	4.053	0.000	***
Period	8.80800	2.11300	83.44	4.169	0.000	***
Age:Sex	-0.49860	2.01100	62610	-0.248	0.804	
Age:Year	-0.00350	0.00072	62430	-4.882	0.000	***
Sex:Year	0.00084	0.00075	62610	1.12	0.263	
Age:Period	-12.08000	1.93700	62170	-6.236	0.000	***
Sex:Period	0.74120	1.87400	62620	0.396	0.692	
Year:Period	-0.00441	0.00106	83.48	-4.182	0.000	***
Age:Sex:Year	0.00025	0.00100	62610	0.248	0.804	
Age:Sex:Period	4.04700	2.68900	62620	1.505	0.132	
Age:Year:Period	0.00606	0.00097	62170	6.256	0.000	***
Sex:Year:Period	-0.00037	0.00094	62620	-0.393	0.694	
Age:Sex:Year:Period	-0.00203	0.00134	62620	-1.513	0.130	

Our reanalysis confirmed that we did not need information on the species to which the birds belonged to establish the general pattern of the change in tarsus and wing length over time. As a matter of fact, the models without information on species were more reliable. Moreover, we were able to show that the rate of change differed before and after 2000 between age groups, a pattern that could not be discovered when including species as factor in the analysis.

2.2. Maximum Entropy Theory of Ecology

As a tentative example of prospects of the application of METE, we also analyzed the metabolic rate of the individuals, its change over the years, and the effect of age, sex, and body size on this change. We don't want to pretend that we were able to fully apply a METE-analysis, we just wanted to show that such an analysis may result in supplementary information that could deepen our knowledge. The analyses were done with the R package *meteR* (version 1.2) [170]).

The data of Weeks and others [55] provide the biomass of each migration bird found. This offers the possibility to study the metabolic rate of the birds, because of Kleiber's general relationship: metabolic rate = $mass^{0.75}$ [98]. The change of metabolic rate over time showed about the same pattern as that of tarsus length (Fig. A2.1a; Fig. A2.2). This was to be expected since tarsus length is a proxy for body size. But the LMM analysis showed different results: of the main effects, not Year, but Age and Sex are significant (Table A2.2; Table A2.3).



Figure A2.2: Smoothed nonlinear regression lines of the raw relationship between year and standardized metabolic rate of hatchlings (magenta) and older birds (blue). Lines were constructed with the default settings of the function *scatterplot()* of the package *car* (version 3.0-10) [168] in R version 4.0.3 [56].

Table A2.3: Effect sizes of the fixed variables, including their interactions, on standardized metabolic rate, as estimated by the function *lmer()* of the package *lme4* [165] in R version 4.0.3 [56].

Standardized Metabolic rate								
	Cond. R ²	Marg. R ²	df	LogLik	AIC			
	0.022	0.011	52372	-73355.8	146747.5			
	Estimate	Std. Error	df	t-value	Pr(> t)			
(Intercept)	13.5400	13.0000	29.3	1.04	0.306			

Age	17.2300	8.2590	52390.0	2.09	0.037 *
Sex	-19.5600	8.9270	52340.0	-2.19	0.028 *
Year	-0.0067	0.0065	29.3	-1.04	0.306
Period	-18.0800	15.6000	31.6	-1.16	0.255
Age:Sex	18.3100	11.7100	52340.0	1.56	0.118
Age:Year	-0.0086	0.0041	52390.0	-2.10	0.036 *
Sex:Year	0.0098	0.0044	52340.0	2.21	0.027 *
Age:Period	-23.4000	11.1400	52340.0	-2.10	0.036 *
Sex:Period	15.5100	11.0200	52340.0	1.41	0.159
Year:Period	0.0090	0.0078	31.6	1.16	0.256
Age:Sex:Year	-0.0091	0.0058	52340.0	-1.56	0.118
Age:Sex:Period	-10.3500	15.5900	52350.0	-0.66	0.507
Age:Year:Period	0.0117	0.0056	52340.0	2.10	0.036 *
Sex:Year:Period	-0.0077	0.0055	52340.0	-1.40	0.163
Age:Sex:Year:Period	0.0051	0.0078	52350.0	0.65	0.515

However, when the expected distribution of metabolic rate over all birds was calculated according to METE, it was clear that the actual distribution deviated strongly from the one expected, with metabolic rates higher than expected in all birds with metabolic rates under 5 $g^{0.75}$ (Fig. A2.3).



Figure A2.3: Distribution of metabolic rate. Red line is the expected distribution based on METE, black circles are the metabolic rates found. Distribution of observed metabolic rates deviates significantly from expected METE distribution (tested using the *logLikZ()* function; z=949.996 [170]).

When we analyzed the effect of Age, Sex, Year, and Period on the deviation of actual metabolic rate from the predicted one using LMM, the marginal R² was not better than in the traditional LMM's (Table A2.4). The results showed that the intercept was significantly higher than the expected zero. Further, most remarkable, the same main factors and interactions were significant as when we analyzed the effect of these independent variables on wing length, but with inversed regression coefficients (compare Table A2.2 and Table A2.4). The model predicted in hatchlings that deviations of metabolic rate slightly increased over time, as did wing lengths, but in older birds that deviations decreased before 2000, but increased after 2000, while wing lengths changed *vice versa* (Fig. 2 and Fig. 3a). Could the inverse relationship between wing length and deviations of expected metabolic rate distribution mean that in birds, wing length is the factor that makes birds metabolically non-neutral, i.e., that wing length affects metabolic rate independently of body size?

Table A2.4: Effect sizes of the fixed variables, including their interactions, on the residuals of the METE predicted metabolic rate, as estimated by the function *lmer()* of the package *lme4* [165] in R version 4.0.3 [56].

Residuals METE Metabolic rate								
	Cond. R ²	Marg. R ²	df	LogLik	AIC			
	0.024	0.019	52372	3532.3	-7028.6			

	Estimate	Std. Error	df	t-value	Pr(> t)	
(Intercept)	7.4170	2.3650	55.1	3.14	0.003	**
Age	-5.0520	1.9030	52340.0	-2.65	0.008	**
Sex	3.1600	2.0580	52350.0	1.54	0.125	
Year	-0.0034	0.0012	55.0	-2.92	0.005	**
Period	-9.6260	2.8670	60.8	-3.36	0.001	**
Age:Sex	-4.3570	2.6990	52360.0	-1.61	0.106	
Age:Year	0.0025	0.0009	52340.0	2.66	0.008	**
Sex:Year	-0.0016	0.0010	52350.0	-1.55	0.120	
Age:Period	7.4070	2.5670	52140.0	2.89	0.004	**
Sex:Period	-0.4315	2.5390	52360.0	-0.17	0.865	
Year:Period	0.0048	0.0014	60.8	3.37	0.001	**
Age:Sex:Year	0.0022	0.0013	52360.0	1.61	0.107	
Age:Sex:Period	-0.3099	3.5940	52370.0	-0.09	0.931	
Age:Year:Period	-0.0037	0.0013	52130.0	-2.89	0.004	**
Sex:Year:Period	0.0002	0.0013	52360.0	0.15	0.880	
Age:Sex:Year:Period	0.0002	0.0018	52370.0	0.11	0.916	

Residuals METE Metabolic rate, including body size

	Cond. R ²	Marg. R ² df		LogLik	AIC	
	0.060	0.055	52356	4530.1	-8992.1	
	Estimate	Std. Error	df	t-value	Pr(> t)	
(Intercept)	9.7270	2.3350	78.4	4.166	0.000 ***	•

Age	-6.4370	2.0250	52360.0	-3.179	0.001	**
Sex	4.3400	2.1590	52370.0	2.010	0.044	*
Year	-0.0046	0.0012	78.3	-3.951	0.000	***
Period	-12.8600	2.8270	85.5	-4.548	0.000	***
Size	-0.8878	4.4120	52390.0	-0.201	0.841	
Age:Sex	-5.5460	2.8550	52370.0	-1.942	0.052	•
Age:Year	0.0032	0.0010	52360.0	3.181	0.001	**
Sex:Year	-0.0022	0.0011	52370.0	-2.032	0.042	*
Age:Period	9.3460	2.6880	52250.0	3.477	0.001	***
Sex:Period	-1.7330	2.6330	52370.0	-0.658	0.510	
Year:Period	0.0064	0.0014	85.5	4.566	0.000	***
Age:Size	8.3610	5.3980	52380.0	1.549	0.121	
Sex:Size	-2.7230	6.3110	52370.0	-0.431	0.666	
Year:Size	0.0005	0.0022	52390.0	0.223	0.824	
Period:Size	7.7210	5.9800	52300.0	1.291	0.197	
Age:Sex:Year	0.0028	0.0014	52370.0	1.939	0.053	
Age:Sex:Period	0.6012	3.7440	52370.0	0.161	0.872	
Age:Year:Period	-0.0047	0.0013	52250.0	-3.479	0.001	***
Sex:Year:Period	0.0008	0.0013	52370.0	0.642	0.521	
Age:Sex:Size	6.2780	7.9090	52370.0	0.794	0.427	
Age:Year:Size	-0.0042	0.0027	52380.0	-1.548	0.122	
Sex:Year:Size	0.0014	0.0031	52370.0	0.442	0.658	

Age:Period:Size	-19.7800	8.0520	52390.0	-2.456	0.014	*
Sex:Period:Size	4.5710	8.8310	52390.0	0.518	0.605	
Year:Period:Size	-0.0039	0.0030	52300.0	-1.295	0.195	
Age:Sex:Year:Period	-0.0003	0.0019	52370.0	-0.141	0.888	
Age:Sex:Year:Size	-0.0031	0.0039	52370.0	-0.793	0.428	
Age:Sex:Period:Size	-7.1440	12.0900	52370.0	-0.591	0.555	
Age:Year:Period:Size	0.0099	0.0040	52390.0	2.458	0.014	*
Sex:Year:Period:Size	-0.0023	0.0044	52390.0	-0.519	0.603	
Age:Sex:Year:Period:Size	0.0036	0.0061	52370.0	0.589	0.556	

However, LMM's may not be fit for application here, because of the specific frequency distribution of the deviations. In this case, the deviations showed a distinct pattern, viz, that metabolic rates were higher than expected in *all* small birds, which cannot be acknowledged in LMM's. What we could do, though, was to include another independent variable in our LMM, a variable that distinguished small (metabolic rate $< 5 \text{ g}^{0.75}$) from large (metabolic rate $\geq 5 \text{ g}^{0.75}$) birds. This resulted in a much better fitting LMM ($\Delta \text{ AIC} = 1963.5$; marginal R² more than doubled from 0.019 to 0.055; Table A2.4). The predictions of this model showed that the large birds dominate the picture when body size was ignored, but actually only the hatchlings of large birds did not change in deviations of METE predictions in the period till 2000, and the small, older birds decreased after 2000 (Fig. 3). We leave the ecological interpretation of these results to those that are more familiar with the effects of climate change on North American birds. We would like to stress here that the application of METE showed us that body size, as well as age and sex, is a highly relevant trait for such an ecological interpretation.

2.3. Conclusions

The application of OE on the temporal change of North-American migrating birds that collided to buildings showed us that:

- ignoring species resulted in reliable models and showed that hatchlings and older birds changed differently in the period before than that after the year 2000
- analyzing how metabolic rate deviated from that expected of METE showed that small birds deviated differently from large birds; including size class as independent variable resulted in a better fitting model and thus in higher predictability

- specific statistical techniques addressing the analysis of deviations from METE predictions may be needed.

2.4. References

References [1]-[114] are in the reference list of the main text.

- 115. Kattge J, Díaz S, Lavorel S, Prentice IC, Leadley P, Bönisch G et al. TRY a global database of plant traits. Glob Change Biol (2011) 17:2905–2935.
- 116. Martiny JBH, Jones SE, Lennon JT, Martiny AC. Microbiomes in light of traits: a phylogenetic perspective. Science (2015) 350: DOI 10.1126/science.aac9323.
- 117. D'Urban Jackson T, Williams GJ, Walker-Springett G, Davies AJ. Three-dimensional digital mapping of ecosystems: a new era in spatial ecology. Proc. R. Soc. B (2020) 287:20192383.
- 118. Peters DPC, Havstad KM, Cushing J, Tweedie G, Fuentes O, Villanueva-Rosales N. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. Ecosphere (2014) 5:67.
- 119. Farley SS, Dawson A, Goring SJ, Williams JW. Situating ecology as a big-data science: current advances, challenges, and solutions. BioScience (2018) 68:563–576.
- 120. Łomnicki A. Individual-based models and the individual-based approach to population ecology. Ecological Modelling (1999) 115:191–198.
- 121. Bolnick DI, Yang LH, Fordyce JA, Davis JM, Svanbäck R. Measuring individual level resource specialization. Ecology (2002) 83:2936-2941.
- 122. Bolnick DI, Svanbäck R, Fordyce JA, Yang LH, Davis JM, Hulsey CD et al. The ecology of individuals: incidence and implications of individual specialization. The American Naturalist (2003) 161:1-28.
- 123. Carrer M. Individualistic and time-varying tree-ring growth to climate sensitivity. PLOS ONE (2011) 6:e22813.
- 124. Araújo MS, Bolnick DI, Layman CA. The ecological causes of individual specialization. Ecology Letters (2011) 14:948-958.
- 125. Violle C, Enquist BJ, McGill BJ, Jiang L, Albert CH, Hulshof C et al. The return of the variance: intraspecific variability in community ecology. TREE (2011) 27:244-252.
- 126. Schirmer A, Herde A, Eccard JA, Dammhahn M. Individuals in space: personality-dependent space use, movement and microhabitat use facilitate individual spatial niche specialization. Oecologia (2019) 189:647–660.
- 127. Westneat DF, Potts LJ, Sasser KL, Shaffer JD. Causes and consequences of phenotypic plasticity in complex environments. TREE (2019) 34:555-568.

- 128. De Roos AM. Dynamic population stage structure due to juvenile–adult asymmetry stabilizes complex ecological communities. PNAS (2021) 118:e2023709118.
- 129. McGill BJ, Enquist BJ, Weiher E, Westoby M. Rebuilding community ecology from functional traits. TREE (2006) 21:178-185.
- 130. Kearney M, Porter WP. Ecologists have already started rebuilding community ecology from functional traits. TREE (2006) 21:481-482.
- 131. Violle C, Navas M-L, Vile D, Kazakou E, Fortunel C, Hummel I et al. Let the concept of trait be functional! Oikos (2007) 116:882-892.
- 132. Mouillot D, Graham NAJ, Villéger S, Mason NWH, Bellwood DR. A functional approach reveals community responses to disturbances. TREE (2013) 28:167-177.
- 133. Meunier CL, Boersma M, El-Sabaawi R, Halvorson HM, Herstoff EM, van de Waal DB et al. From elements to function: toward unifying ecological stoichiometry and trait-based ecology. Front. Environ. Sci. (2017) 5:18.
- 134. Ieromina O, Musters CJM, Bodegom PM, Peijnenburg WJGM, Vijver MG. Trait modality distribution of aquatic macrofauna communities as explained by pesticides and water chemistry. Ecotoxicology (2016) 25:1170-1180.
- 135. Moreno-Martínez A, Camps-Valls G, Kattge J, Robinson N, Reichstein M, van Bodegom P et al. A methodology to derive global maps of leaf traits using remote sensing and climate data. Remote Sensing of Environment (2018) 218:69–88.
- 136. McLean NM, van der Jeugd NP, van Turnhout CAM, Lefcheck JS, van de Pol M. Reduced avian body condition due to global warming has little reproductive or population consequences. Oikos (2020) 129:714-730.
- 137. Zakharova L, Meyer KM, Seifan M. Trait-based modelling in ecology: A review of two decades of research. Ecological Modelling (2019) 407:108703.
- Melian CJ, Vilas C, Baldó F, González-Ortegón E, Drake P, Williams RJ. Eco-evolutionary dynamics of individual-based food webs. Advances in Ecological Research (2011) 45: 225-268.
- 139. Villegas-Ríos D, Claudet J, Freitas C, Moland E, Huneide Thorbjørnsen S, Alonso-Fernández A et al. Time at risk: Individual spatial behaviour drives effectiveness of marine protected areas and fitness. Biological Conservation (2021) 263:109333.
- 140. Clutton-Brock T, Sheldon BC. Individuals and populations: the role of long-term, individualbased studies of animals in ecology and evolutionary biology. TREE (2010) 25:562–573.
- 141. Belsare AV, Gompper ME, Keller B, Sumners J, Hansen L, Millspaugh JJ. An agent-based framework for improving wildlife disease surveillance: A case study of chronic wasting disease in Missouri white-tailed deer. Ecological Modelling (2020) 417:108919.

- 142. Moretti M, Dias ATC, de Bello F, Altermatt F, Chown SL, Azcárate FM et al. Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. Functional Ecology (2017) 31:558-567.
- 143. Musters CJM, van Bodegom PM 2018. Analysis of species attributes to determine dominant environmental drivers, illustrated by species decline in the Netherlands since the 1950s. Biological Conservation (2018) 219:68-77.
- 144. Gallagher RV, Butt N, Carthey AJR, Tulloch A, Bland L, Clulow S et al. A guide to using species trait data in conservation. One Earth (2021) 4:927-936.
- 145. Arnold SJ. Morphology, performance and fitness. Amer Zool (1983) 23:347-361.
- 146. Higham TE, Ferry LA, Schmitz L, Irschick DJ, Starko S, Anderson PSL et al. Linking ecomechanical models and functional traits to understand phenotypic diversity. TREE (2021) 36:860-873.
- 147. Simberloff D, Dayan T. The guild concept and the structure of ecological communities. Annu. Rev. Ecol. Syst. (2004) 22:115-143.
- 148. Kattge J, Bönisch G, Díaz S, Lavorel S, Prentice IC, Leadley P et al. TRY plant trait database enhanced coverage and open access. Glob Change Biol (2020) 26:119–188.
- 149. Scheffer M, Baveco JM, DeAngelis DL, Rose KA, van Nes EH. Super-individuals a simple solution for modelling large populations on an individual basis. Ecological Modelling (1995) 80:161-170.
- 150. Lesilau F, Fonck M, Gatta M, Musyoki C, van 'tZelfde MA, Persoon GA et al. Effectiveness of a LED flashlight technique in reducing livestock depredation by lions (Panthera leo) around Nairobi National Park, Kenya. PLOS ONE (2018): <u>https://doi.org/10.1371/journal.pone.0190898</u>
- 151. Jørgensen SE. Overview of the model types available for development of ecological models. Ecological modelling (2008) 215:3-9.
- 152. Sagoff M. Are there general causal forces in ecology? Synthese (2016) 193:3003-3024.
- 153. Donhauser J. Informative ecological models without ecological forces. Synthese (2020) 197:2721–2743.
- 154. Varughese MM. A framework for modelling ecological communities and their interactions with the environment. Ecological Complexity (2011) 8:105–112.
- 155. DeAngelis DL, Yurek S. Spatially explicit modeling in ecology: a review. Ecosystems (2017) 20:284–300.
- 156. Godin C, Sinoquet H. Functional–structural plant modelling. New Phytologist (2005) 166:705–708.

- 157. Landsberg JJ, Waring RW. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecology and Management (1997) 95:209–228.
- 158. Faverjon F, Escobar-Gutiérrez A, Litrico I, Julier B, Louarn G. A generic individual-based model can predict yield, nitrogen content, and species abundance in experimental grassland communities. Journal of Experimental Botany (2019) 70:2491–2504.
- 159. Garin G, Fournier C, Andrieu B, Houlès V, Robert C, Pradal C. A modelling framework to simulate foliar fungal epidemics using functional–structural plant models. Annals of Botany (2014) 114:795–812.
- 160.Wiegand T, Jeltsch F, Hanski I, Grimm V. Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. Oikos (2003) 100:209-222.
- 161. Sakschewski B, von Bloh W, Boit A, Rammig A, Kattge J, Poorter L et al. Resilience of Amazon forests emerges from plant trait diversity. Nature Clim Change (2015) 6:1032–1036
- 162. Hubbell SP. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton, USA: Princeton University Press (2001).
- 163. Rosindell J, Hubbell SP, Etienne RS. The unified neutral theory of biodiversity and biogeography at age ten. TREE (2011) 26:340-348.
- 164. Breckling B, Middelhoff U, Reuter H. Individual-based models as tools for ecological theory and application: Understanding the emergence of organizational properties in ecological systems. Ecological Modelling (2006) 194:102-113.
- 165. Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using Ime4. Journal of Statistical Software (2015) 67: 1.
- 166. Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software (2021) 6: 3139.
- Zeileis A, Hothorn T. (2002). Diagnostic Checking in Regression Relationships. R News 2: 7-10.
- 168. Fox J, Weisberg S. An R Companion to Applied Regression, Third Edition. Thousand Oaks, USA: Sage Publications (2019).
- 169. Jirinec V, Burner RC, Amaral BR, Bierregaard Jr. RO, Fernández-Arellano G, Hernández-Palma A et al. Morphological consequences of climate change for resident birds in intact Amazonian rainforest. Science Advance (2021) 7:eabk1743.
- 170. Rominger AJ, Merow C. meteR: an R package for testing the maximum entropy theory of ecology. Methods in Ecology and Evolution (2017) 8:241–247.