Experimental validation of ecological coexistence theory to forecast extinction under rising temperatures.

J. Christopher D. Terry[1*]

1: Department of Biology, University of Oxford, 11A Mansfield Road, Oxford, OX1 3SZ

christopher.terry@biology.ox.ac.uk 0000-0002-0626-9938

Abstract

Global environmental changes are driving ecological community reorganisation at an accelerating rate and managing these impacts requires reliable forecasts. Individual species responses to environmental change are often driven by interactions with other species adding considerable complexity. To meet this challenge, theoretical ecology, specifically coexistence theory, defines precise conditions by which species can or cannot persist alongside competitors. However, although coexistence theory is being increasingly deployed for projections, these frameworks have rarely been subjected to critical validation tests. Here, using a highly replicated multigenerational mesocosm experiment, I directly test if the modern coexistence theory approach can accurately predict time-to-extirpation in the face of rising temperatures within the context of competition from a heat-tolerant species. Competition hastened expiration and the modelled point of coexistence breakdown reasonably aligned with mean observations under both a steady temperature increase or with additional environmental stochasticity. That said, predictive precision was low and coexistence outcome predictions depended on the error model used. Nonetheless, these results support the careful use of coexistence modelling for near-term forecasts.

Main Text

Species are already shifting their distributions across the world as environmental conditions rapidly change[1,2]. Forecasting how the identities of species resident in ecosystems will change is both of profound applied importance[3] and a key test of our ecological understanding of community assembly. Beyond direct physiological responses, the potentially large influence of interspecific competition in determining whether a species can persist under a different environment has been demonstrated in both mesocosm[4,5] and field studies[6–9]. Parsing this complexity of multi-species dynamics to make forecasts is a considerable challenge given the range of possible processes[10], but theoretical ecology has developed streamlined frameworks defining precise conditions for when species are, or are not, expected to coexist[11–14].

Modern coexistence theory focuses on a widely applicable currency: invasion growth rate[11,15], the per-capita population growth rate from low densities into an equilibrium community of competitors. Assuming no positive frequency-dependence (‘Allee effects’[16]), a positive invasion growth rate is taken to imply that a species is able to persist on the basis that it could always recover from low

densities. Where both species fulfil the mathematical criteria[13], they are claimed to be able to coexist in a locality. This approach is frequently framed as an assessment of whether stabilising niche differences (that favour coexistence) can overcome inherent fitness differences (that drive exclusions) between pairs of species[11].

There have been increasing efforts to deploy the extensive body of theory that has developed around the determinants of coexistence for more applied purposes[17,18]. In particular, coexistence theory has been proposed as a tool to understand how environmental change will impact communities in the future, largely based on parametrising models from short-term experiments [9,19–21] under different environmental conditions. Where key model parameters are found to be environmentally dependent, the future environments in which species can (or cannot) coexist can be inferred. An ultimate hope is to be able to make trait or phylogeny-based predictions[20,22,23] of niche and fitness differences in different environments which would potentially unlock the capacity to forecast species responses to climatic changes at scale without the need for individual parameterisation.

However, the extent to which this framework is applicable to contemporary challenges remains unclear. Critically, the underlying mathematical assumptions of modern coexistence theory are widely recognised as being some distance from reality: the approach implicitly assumes infinite time and space horizons[24], as well as continuous population counts without the demographic stochasticity that discrete individuals at small population sizes bring[25]. It remains an unsettled question the extent to which this framework is adequate[26] for real ecological questions and complexity.

To assess the capacity for modern coexistence theory to forecast extirpations (i.e. local extinctions) under changing environments, I conducted a highly replicated *Drosophila* mesocosm experiment where populations were tracked through discrete generations in small mesocosms (Fig 1.). Focal species are drawn from a well-characterised community[27–30] from a montane Australian tropical rainforest that shows pronounced species turnover over an elevational gradient. Previous empirical work using this community has demonstrated roles for both temperature and competition in determining species range boundaries[31,32]. Consequently, I focussed on predicting the survival of populations of *Drosophila pallidifrons*, a highland-distributed species with a comparatively cool thermal optimum, in the context of competition from a lowland-distributed species *Drosophila pandora* with a comparatively warm thermal optimum. Each generation the mean temperature was increased by 0.4°C, which was expected to shift the competitive balance from *D. pallidifrons* to *D. pandora*.

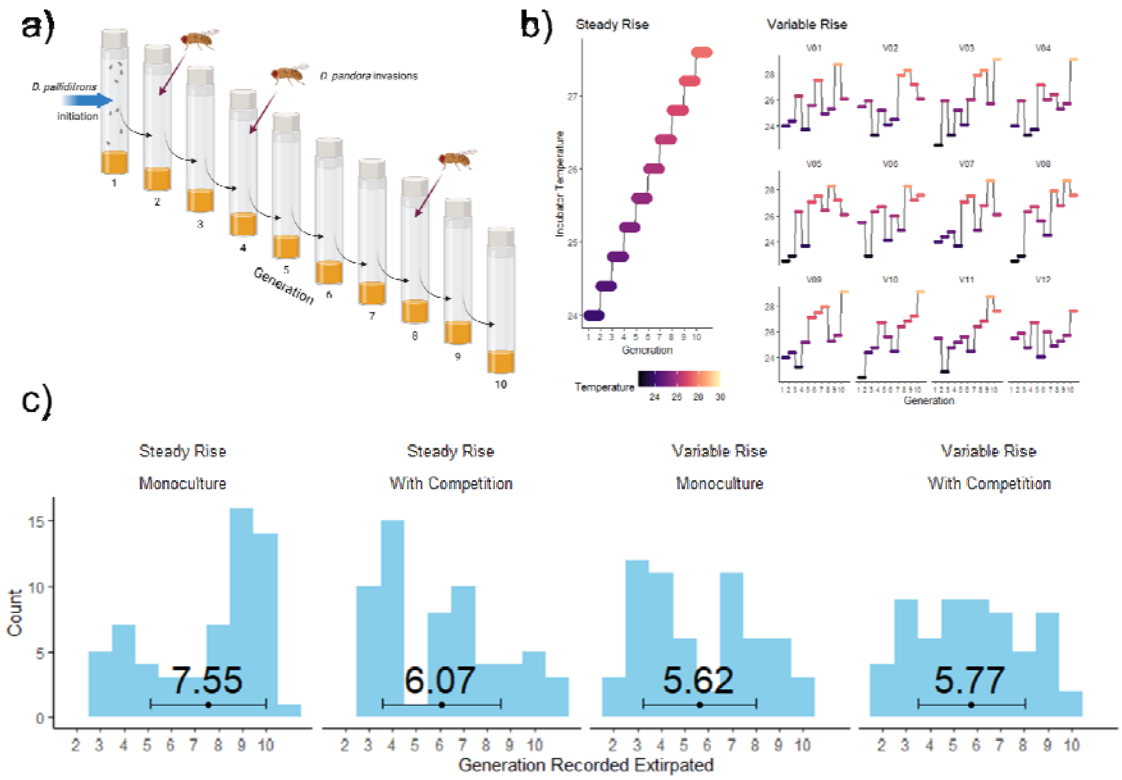

Figure 1. Outline of experiment and time-to-extinction results a) Overall experiment schematic. After initiation of population vials with *D. pallidifrons*, the flies were allowed two days to lay eggs before removal and censusing. After a further 10 days, all emerged flies were transferred to a new vial for two days to start the next discrete generation. In the 'competition' treatment, small numbers of *D. pandora* were added alongside the founders of generation 2, 4 and 8. b) Temperature treatments. In the 'steady rise' scenario, the temperature increases 0.4°C at the start of each generation, while in the variable rise treatment, the temperature imposed on each population follows one of 12 trajectories (V01 – V12) that superimpose variability on top of the steady rise. c) Histograms of points of extirpation, defined as the first generation where zero *D. pallidifrons* individual were censused. Values shown are means and error bars show +/- 1 SD. n=60 per top-level treatment combination (240 total).

Sixty replicates were run for each of two crossed treatment combinations: monoculture or the intermittent introduction of small numbers of *D. pandora*, (Figure 1A), and steady temperature rise or with additional generational-scale thermal variability (either +1.5°C, -1.5°C or 0, with equal chance, Figure 1B). The experiment was run for 10 generations, by which point *D. pallidifrons* had been extirpated in 237 of 240 population lines, and populations were censused each discrete generation.

Treatment combination had a significant overall effect on the generation at which extirpation was observed (Figure 1C, ANOVA: = 8.152, p = 0.0003) but only the slower 'steady rise monoculture' treatment was statistically distinct (Tukey HSD test: p< 0.005 for all comparisons with this treatment, p> 0.7 for all others) with a mean expiration 1.73 (SE: 0.358) generations later. When viewed as concurrent stressors, environmental stochasticity and competition had antagonistic interactive effects[33], in the sense that their combined effect was less than the sum of their individual impacts. The particular trajectory of temperature variability had a significant impact on time-to-

extirpation (ANOVA model comparison of Extirpation~Treatment with Extirpation~Treatment +(1|Trajectory): χ^2 = 6.4029, p = 0.011, n = 120, 12 groups). Trajectories with a lower temperature in the first generation were correlated with earlier expiration, but across the 12 trajectories this was not statistically significant (post-hoc linear regression, p = 0.145, $F_{1,10}$ =2.498, adjusted R^2 = 0.12).

To form the basis of predicting coexistence at different temperatures, the observed transitions were used to fit population dynamics models describing the thermal performance, competition and error distribution for each species. Fitting was based on counts of adult females with a total of 1,223 observed generation-to-generation population transitions for *D. pallidifrons* and 592 for *D. pandora*. Model fitting and selection used a Bayesian framework and followed recently established best practices[34,35] (see methods). The selected population model for both species used a simplified Brière thermal performance function[36] and a Beverton-Holt competition model with temperature-dependent competition exerted by *D. pallidifrons*:

$$N_{i,g+1} \sim N_{i,g} \frac{a\left(T_g - T_{min}\right)\sqrt{T_{max} - T_g}}{1 + \left(\alpha_{0i,PAL} + \alpha_{Ti,PAL} T_g\right) N_{PAL,g} + \alpha_{i,PAN} N_{PAN,g}}$$

where N is the number of individual females of species i in generation g, T_g is the temperature in generation g, a, T_{min} and T_{max} are parameters of the thermal performance function defining the low-density growth rate, and α terms are competition coefficients. Following standard approaches[23,34,37], coefficients from these fitted models were used to infer the population fitness differences and niche overlap at each temperature and categorise the invasion growth rate of each species. With increased temperature, the models predicted a shift from *D. pallidifrons* (the resident, highland species) excluding *D. pandora* (the lowland, invader species), through the zone of likely coexistence, to the exclusion of *D. pallidifrons* by *D. pandora* (Figure 2a).

Using mean parameter estimates, the threshold for predicted exclusion of *D. pallidifrons* is crossed at 26.7°C, although using the full posterior there is support for predicted exclusion somewhat earlier (Figure 2b). Comparing to the observed extinctions, in the steady temperature rise treatment, this threshold is first breached in generation 8, just under two generations (corresponding to a 0.8°C difference) after the mean generation of observed extirpation (6.07).

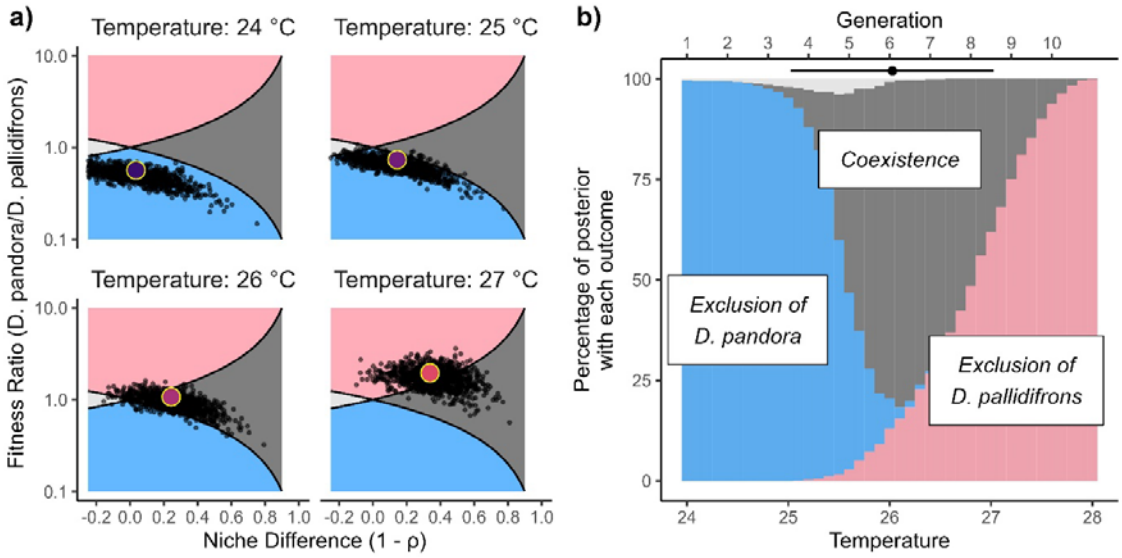

Figure 2: Predictions of coexistence in the steady rise and competition scenario based on model fits. a) Full posterior distribution of the niche difference ($1 - \rho$) and fitness differences calculated based on parameter distributions from the selected models. Coloured dots show results based on average parameters. Small black dots show draws from posterior distribution. Shading indicates boundary on the plane between predicted coexistence outcomes. b) Allocation of posterior support for each of the four possible predicted coexistence outcomes based on inferred niche and fitness differences. Blue corresponds to predicted exclusion of *D. pandora* at low temperatures, dark grey to the region of predicted coexistence, pink to exclusion of *D. pallidifrons* and light grey corresponds to 'priority effects', where each species can exclude the other. Point at top shows mean observed extirpation (+/- 1 SD) in the competition and steady rise treatment.

To test the ability to predict extirpation under the environmental stochasticity treatment, I used a simulation-based extension of modern coexistence theory[12]. This makes asymptotic predictions of the impact of the temperature variability regime on the invasion criterion by averaging over long synthetic time series. Despite several formal assumptions of the framework (such as infinite time horizons) not being met by the experimental scenario examined, the framework is able to make somewhat accurate predictions - the introduction of variability in the competition treatment is predicted to cause exclusion of *D. pallidifrons* at a slightly lower temperature of ~26.2°C. Although the statistical support for earlier extirpation under variability in the experiments is weak, the difference in mean quite closely corresponds to the observed difference in extirpation.

Assessing the value of the accuracy of the model's predictions is inevitably somewhat subjective. An optimistic perspective of the alignment would highlight that an estimate within two generations is close enough to be useful. Given the discordance between the theoretical assumptions of the modern coexistence theory and the empirical reality of the experiment, it is surprising and heartening that the results are aligned at all. Further, the experiment does not include any possibility for *D. pallidifrons* recolonising after a stochastic loss, likely speeding the observed extirpation compared to a more natural situation. A pessimistic view would highlight the substantial temperature difference between

observed and predicted extirpations (akin to several decades under most warming scenarios) and the high uncertainty in the prediction, despite levels of replication greatly exceeding the norms of the field.

Aligning the scales of experiment and natural scenarios is a particular challenge for extinction dynamics. Experiments with more generations and shallower temperature gradients would allow finer resolution, but will be limited to an even more restricted repertoire of model species. Despite being a controlled mesocosm experiment, individual population trajectories were characterised by a high degree of variation (Figure S1). Counts were complete censuses with populations individually sexed and counted so minimal measurement error can be assumed. Population counts were of the order of dozens (mean of non-zero females counts = 19.9, σ =12.7) giving considerable scope for demographic stochasticity[38] stemming from individual level-differences within the populations. Within-treatment 'external' environmental variability was minimised by the experimental design. Nonetheless, there was likely significant emergent environmental heterogeneity when defined as variation correlated amongst individuals in a vial[39]. Individual vials in the same treatment could be highly divergent, with some vials containing apparently stressed communities with late-developing pupae. Capturing this variation is a major, but underacknowledged[34], challenge for applying coexistence theory. The best-fit models incorporated zero-inflated negative-binomial error distributions with temperature dependent parameters, a considerably more complex expression than is frequently assumed[34]. Modelling the system with approaches that did not have the flexibility to represent this variability gave different predictions. For example, using a Poisson error distribution coupled with the best fit competition model described above, the system is predicted to instead move through a region of priority effects to a sharp shift in predicted dominance of *D. pandora* at an earlier temperature (Figure S6).

It must nonetheless be acknowledged that the alignments observed in this experiment likely represent a 'best-case' expectation of accuracy. Beyond the model and parameter uncertainty captured within ecological modelling, natural world forecasts must also grapple with climate change uncertainty (incorporating both human and earth system responses)[40]. The extent of data available in this test is likely to exceed that available for many applied cases. While the forecasts could be considered acceptably accurate, the high heterogeneity among replicates suggests a strong fundamental limits of predictability at local scales. This heterogeneity arises even without incorporating dispersal variation[41] or the true field community complexity, including many more competitive rivals and dynamic interactions with natural enemies such as parasitoids[42]. In natural systems, the *Drosophila* exist in a highly dynamic metacommunity as individual ephemeral food resources become available and are consumed. Such a spatially distributed system will likely mitigate the substantial failure rate of individual patches. Further, within the mesocosms there was no distinction between the conditions under which the species' performance was assessed and those for the long-term experiment. In field conditions it can be hard to capture environmental impacts on all life-stages appropriately, and indeed

each may undergo distinct environmental trajectories[43]. Nonetheless, this study offers support that when carefully applied even quite abstract coexistence models can indeed make informative forecasts for contemporary change beyond the precise bounds of their formal applicability.

As communities reorganise in the face of changing environments, we need trustworthy tools to offer guidance. There has been substantial and ongoing debate about how best to address the gap between the asymptotic theoretical world and the finite and messy reality that we must confront. Ultimately, observation and experimentation are essential to offer guidance and since we do not have the luxury of sufficient time to conduct equivalent direct validation experiments in species with longer lifecycles, mesocosm experiments with high-turnover model species will have a major role to play. Validation tests such as those presented here can contextualise the expected accuracy from models to develop an expectation of the appropriate level of confidence in, and hence application and deployment of, coexistence theory.

Methods

Species

The two focal species were *Drosophila pallidifrons* (*nasuta* species complex) and *Drosophila pandora* (*ananassae* subgroup[44]), both drawn from a well characterised Australian tropical forest community that shows distinct elevational community turnover[27]. Laboratory populations had been maintained at Oxford University since 2018, originally drawn from multiple isofemale lines sampled from rainforest sites in northern Queensland, Australia (permit PWS2016-AU-002018, Australian Department of Energy and Environment). These cultures had been maintained at large population sizes in a 25°C controlled temperature room. While some laboratory adaptation over this period cannot not be ruled out, preceding work using these populations had demonstrated that species had maintained distinct thermal physiologies[32].

Experimental Procedures

Experimental populations were maintained in three incubators of two equivalent models (Sanyo MIR-154 & Sanyo MIR-153) under a 12-12hr light-dark cycle. Temperature and humidity loggers indicated these incubators were able to successfully maintain a controlled environment throughout the experiment.

Each generation was initiated by transferring the founder flies into a standard 25 mm diameter *Drosophila* vial with 5 ml of cornflour–sugar-yeast–agar *Drosophila* medium. The founders were allowed approximately 48 hours to lay eggs before being removed and frozen for later censusing. After 10 days incubation any emerged flies became the founders of the next generation in that line, giving a total tip-to-tip generation time of 12 days. This length was chosen to match the development time across the temperatures used in the experiment. During censusing, all individuals were identified

by species, sexed and counted under a stereo microscope. Flies that had clearly died before freezing were not counted. Each generation 1 (G1) vial was initiated with 3 female and 2 male *D. pallidifrons* counted under light CO_2 anaesthesia that had been raised in 180 ml Drosophila flasks under moderate density at 24°C. 60 replicate vials were concurrently maintained for each of the four treatment combinations: Steady or Variable temperature rises, and Monoculture or Competition biotic conditions.

In the 'Steady' rise temperature treatment, G1 was founded at a constant temperature of 24°C and the temperature was raised by a 0.4°C step at the start of each generation, implying a 4 degree rise across the experiment. For the 'Variable' rise treatment, 12 different variable temperature trajectories were generated by randomly assigning a third of the trajectories to each of lowered (-1.5), mean (=) or raised (+1.5) temperature conditions for each generation. Each trajectory had 5 replicates within each treatment. Temperature values were chosen to span the expected shift in coexistence, while also allowing the use of a fixed generation duration across the experiment (since colder temperatures delay development significantly).

In the competition biotic treatment, *D. pandora* were intermittently added alongside the founders from the preceding generation. At the start of G2, approximately 15 *D. pandora* were added to each treatment vial. At the start of G4 and G8 4 female and 2 male *D. pandora* were added to all lines that still had a *D. pallidifrons* population. These additions were subtracted from the subsequent census to infer the reproductive rate in the previous generation. The experiment was ended at the census for G10 when 237/240 of the lines no longer contained any female *D. pallidifrons.*

Population Data

In total 37,826 *D. pallidifrons* and 20,786 *D. pandora* flies were counted from 1,045 censuses with non-zero counts. Population dynamics were based just on female counts to better account for the reproductive capacity of small populations and since the sex ratio of both species skewed towards females at higher densities (Figure S2, likely due to slower emergence of males and increased male mortality).

Across all temperatures, there were 1,223 *D. pallidifrons* and 592 *D. pandora* observed transitions from non-zero female populations. Because there were so few (6) *D. pallidifrons* data points from temperatures above 28°C, these were excluded to leave a total data set of 1,217 transitions. This allowed the lower bound of T_{max} (see below) to be reduced without causing numerical errors.

Population Model Fitting

A suite of phenomenological models were tested that systematically varied the model form used to represent the fundamental thermal performance curve and the reduction in fecundity caused by competition. These models were fit and compared separately for each species. All models were fit in a

Bayesian framework using the brms[45] version 2.19 R interface to STAN[46] version 2.26.1. Priors on the core growth rate function parameters were chosen to be weakly informative and bounded only to prevent numerical errors. Priors on the competition model parameters were based on existing expectations of the strength of competition[30,32] and to enforce competition (i.e positive α_{ij}) rather than positive interactions. For each model, 4 chains were each run for 4000 total iterations (2000 used as a warm up) on standard settings. Model code in R (*brms*) and STAN is available in the code supplement. Runs were checked for convergence through the Gelman diagnostic $\hat{r}$ and posterior distributions and predictive plots are available in the code supplement.

Each model followed the overall structure:

$$N_{i,t+1} \sim N_{i,t} \times B(T_t) \times C\left(N_{i,t}, N_{j,t},\ T_t\right).$$

Based on existing expectations for *Drosophila* thermal performance curves[47], four phenomenological asymmetric thermal performance curve models $B(T_t)$ were trialled, selected for having a moderate number of parameters and their distinctiveness[48]: 'Taylor-Sexton', 'Atkin', 'Simplified β-type' and 'Simplified Brière'. Only the Simplified Brière thermal performance function[36] was able to consistently converge:

$$B(T) = a \times (T - T_{min}) \times \sqrt{T_{max} - T}.$$

Eight phenomenological competition response models $C\left(N_{i,t}, N_{j,t},\ T_t\right)$ were tested (Table S1), chosen from a standard suite of models that are frequently used to represent competition, and supplemented by models that include temperature-dependent competition coefficients. Linear relationships (i.e. a Lotka-Volterra model) were also tested but converged poorly and were excluded from consideration.

Six error models for count data were tested: Poisson, Negative Binomial, zero-inflated Negative Binomial, temperature-dependent-zero-inflated Negative Binomial, zero-inflated Negative Binomial with competition dependent shape, and zero-inflated Negative Binomial with temperature dependent shape (Table S2).

Models were compared using ELPD-LOO[49], which identifies the model best suited to make predictions on unseen further data (Table S3-4). Where models could not be clearly distinguished on the basis of predictive performance the simpler model was used. Models including temperature dependent competition consistently performed better than models that did not. Models including an additional exponent term had slightly better predictive performance (ΔELPD =6.3, SE = 2.5), but not to the extent to justify the additional analytical complication.

For both species, the selected competition kernel was the Beverton-Holt model of density-dependent competition with the competitive pressure exerted by *D. pallidifrons* negatively impacted by temperature:

$$C\left(N_{PAL,t}, N_{PAN,t},\ T_t\right) = \frac{1}{1 + \left(\alpha_{0i,PAL} + \alpha_{Ti,PAL}T_t\right)N_{PAL,t} + \alpha_{i,PAN}N_{PAN,t}}.$$

Note that the number of intraspecific competitors of species was taken to be $max(0, N-1)$, incorporating the assumption that individuals do not compete with themselves. The α_{ij} terms represent constant competition coefficients, whereas α_{0ij} is an intercept and α_{Tij} a slope term to linearly link a competition coefficient to temperature.

The differences in error functions had a greater impact on ELPD-LOO than choice of competition model (Table S3-4). Error for *D. pallidifrons* was best represented by a zero-inflated negative binomial model, where the extent of zero inflation was temperature dependent (via a logit link function). For *D. pandora*, error was best represented by a negative binomial function where the shape parameter (ϕ) was temperature dependent. See Table S5 and S6 for model parameter values and Figure S4 temperature dependence of emergent parameters.

Coexistence Predictions

The full posterior parameter estimates of the best performing model was propagated through to coexistence predictions. The temperature dependent zero-inflation was incorporated into the overall growth rate term λ by scaling the mean expectation of the number of offspring by the zero inflation: $\lambda(T_t) = B(T_t) \times \left(1 - inv.logit(z_0 + z_T T_t)\right)$.

Following standard approaches for the Beverton-Holt model, the fitness difference $\frac{\kappa_i}{\kappa_j}$ and niche overlap (ρ) of the two species was calculated: $\frac{\kappa_i}{\kappa_j} = \frac{\lambda_i}{\lambda_j}\sqrt{\frac{\alpha_{ji}\alpha_{jj}}{\alpha_{ij}\alpha_{ii}}}$, $\rho = \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$. Coexistence was inferred based on whether the key inequality was fulfilled (Chesson 2000): $\rho < \frac{\kappa_i}{\kappa_j} < \frac{1}{\rho}$.

Prediction of coexistence in variable environments through simulation.

Predictions of coexistence in a variable environment used the simulation-based method of Ellner *et al*[12]. For this, a random sequence of 1000 temperature fluctuations (drawn from +1.5, 0 or -1.5) was generated following the experimental scenario. The best fit population model with mean parameter values was used to simulate the dynamics of *D. pandora* in monoculture, averaging over the final 900 observations to identify a mean carrying capacity value at each central temperature. These carrying capacities and temperature sequences were then used to infer the average population growth rate from rare of *D. pallidifrons* at each central temperature, and hence the temperature at which invasion would no longer be predicted.

Acknowledgments

JCDT was funded by a Leverhulme Trust Early Career Fellowship (ECF-2022-666). The help and support from the Oxford Biology Fly Lab and Community Ecology research groups is gratefully acknowledged.

Data Availability

Raw data and code for all analysis steps is available at github.com/jcdterry/MCT_Fly_Val_public and archived on Zenodo[50]

References

1. Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B. & Thomas, C. D. Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science* **333**, 1024–1026 (2011).

2. Lenoir, J. *et al.* Species better track climate warming in the oceans than on land. *Nature Ecology & Evolution* **4**, 1044–1059 (2020).

3. Pecl, G. T. *et al.* Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017).

4. Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783–786 (1998).

5. Legault, G., Bitters, M. E., Hastings, A. & Melbourne, B. A. Interspecific competition slows range expansion and shapes range boundaries. *Proceedings of the National Academy of Sciences* **117**, 26854–26860 (2020).

6. Suttle, K. B., Thomsen, M. A. & Power, M. E. Species Interactions Reverse Grassland Responses to Changing Climate. *Science* **315**, 640–642 (2007).

7. Descombes, P. *et al.* Novel trophic interactions under climate change promote alpine plant coexistence. *Science* **370**, 1469–1473 (2020).

8. Freeman, B. G., Strimas-Mackey, M. & Miller, E. T. Interspecific competition limits bird species' ranges in tropical mountains. *Science* **377**, 416–420 (2022).

9. Alexander, J. M., Diez, J. M. & Levine, J. M. Novel competitors shape species' responses to climate change. *Nature* **525**, 515–518 (2015).

10. Terry, J. C. D., O'Sullivan, J. D. & Rossberg, A. G. Synthesising the multiple impacts of climatic variability on community responses to climate change. *Ecography* **2022**, e06123 (2022).

11. Chesson, P. Mechanisms of Maintenance of Species Diversity. *Annual Review of Ecology, Evolution, and Systematics* **31**, 343–366 (2000).

12. Ellner, S. P., Snyder, R. E., Adler, P. B. & Hooker, G. An expanded modern coexistence theory for empirical applications. *Ecology Letters* **22**, 3–18 (2019).

13. Barabás, G., D'Andrea, R. & Stump, S. M. Chesson's coexistence theory. *Ecological Monographs* **88**, 277–303 (2018).

14. Armitage, D. W. & Jones, S. E. Coexistence barriers confine the poleward range of a globally distributed plant. *Ecology Letters* **23**, 1838–1848 (2020).

15. Grainger, T. N., Levine, J. M. & Gilbert, B. The Invasion Criterion: A Common Currency for Ecological Research. *Trends in ecology & evolution* **34**, 925–935 (2019).

16. Courchamp, F., Clutton-Brock, T. & Grenfell, B. Inverse density dependence and the Allee effect. *Trends in Ecology & Evolution* **14**, 405–410 (1999).

17. Aoyama, L. *et al.* Application of modern coexistence theory to rare plant restoration provides early indication of restoration trajectories. *Ecological Applications* **32**, e2649 (2022).

18. Alexander, J. M., Diez, J. M., Hart, S. P. & Levine, J. M. When Climate Reshuffles Competitors: A Call for Experimental Macroecology. *Trends in Ecology & Evolution* **31**, 831–841 (2016).

19. Adler, P. B., Dalgleish, H. J. & Ellner, S. P. Forecasting plant community impacts of climate variability and change: when do competitive interactions matter? *Journal of Ecology* **100**, 478–487 (2012).

20. Van Dyke, M. N., Levine, J. M. & Kraft, N. J. B. Small rainfall changes drive substantial changes in plant coexistence. *Nature* **611**, 507–511 (2022).

21. Alexander, J. M. *et al.* Lags in the response of mountain plant communities to climate change. *Global Change Biology* **24**, 563–579 (2018).

22. Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences* **112**, 797–802 (2015).

23. Godoy, O., Kraft, N. J. B. & Levine, J. M. Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters* **17**, 836–844 (2014).

24. Schreiber, S. J., Levine, J. M., Godoy, O., Kraft, N. J. B. & Hart, S. P. Does deterministic coexistence theory matter in a finite world? *Ecology* **104**, e3838 (2023).

25. Pande, J., Fung, T., Chisholm, R. & Shnerb, N. M. Mean growth rate when rare is not a reliable metric for persistence of species. *Ecology Letters* **23**, 282 (2020).

26. Getz, W. M. *et al.* Making ecological models adequate. *Ecology Letters* **21**, 153–166 (2018).

27. Jeffs, C. T. *et al.* Molecular analyses reveal consistent food web structure with elevation in rainforest Drosophila – parasitoid communities. *Ecography* **44**, 403–413 (2021).

28. Kellermann, V. *et al.* Upper thermal limits of Drosophila are linked to species distributions and strongly constrained phylogenetically. *Proceedings of the National Academy of Sciences* **109**, 16228–16233 (2012).

29. Hangartner, S. B., Hoffmann, A. A., Smith, A. & Griffin, P. C. A collection of Australian Drosophila datasets on climate adaptation and species distributions. *Scientific Data* **2**, 150067 (2015).

30. Terry, J. C. D., Chen, J. & Lewis, O. T. Natural enemies have inconsistent impacts on the coexistence of competing species. *Journal of Animal Ecology* **90**, 2277–2288 (2021).

31. O'Brien, E. K., Higgie, M., Reynolds, A., Hoffmann, A. A. & Bridle, J. R. Testing for local adaptation and evolutionary potential along altitudinal gradients in rainforest Drosophila: beyond laboratory estimates. *Global Change Biology* **23**, 1847–1860 (2017).

32. Chen, J. & Lewis, O. T. Limits to species distributions on tropical mountains shift from high temperature to competition as elevation increases. *Ecological Monographs* **94**, e1597 (2024).

33. Orr, J. A. *et al.* Towards a unified study of multiple stressors: divisions and common goals across research disciplines. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20200421 (2020).

34. Terry, J. C. D. & Armitage, D. W. Widespread analytical pitfalls in empirical coexistence studies and a checklist for

improving their statistical robustness. *Methods in Ecology and Evolution* 2023.07.04.547661 (In press) doi:10.1101/2023.07.04.547661.

35. Cervantes-Loreto, A. *et al.* Environmental context, parameter sensitivity, and structural sensitivity impact predictions of annual-plant coexistence. *Ecological Monographs* e1592 (2023) doi:10.1002/ecm.1592.

36. Briere, J.-F., Pracros, P., Le Roux, A.-Y. & Pierre, J.-S. A Novel Rate Model of Temperature-Dependent Development for Arthropods. *Environmental Entomology* **28**, 22–29 (1999).

37. Levine, J. M. & HilleRisLambers, J. The importance of niches for the maintenance of species diversity. *Nature* **461**, 254–257 (2009).

38. Anderson, R. M., Gordon, D. M., Crawley, M. J. & Hassell, M. P. Variability in the abundance of animal and plant species. *Nature* **296**, 245–248 (1982).

39. Shoemaker, L. G. *et al.* Integrating the underlying structure of stochasticity into community ecology. *Ecology* **101**, e02652 (2020).

40. Beaumont, L. J., Hughes, L. & Pitman, A. J. Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters* **11**, 1135–1146 (2008).

41. Melbourne, B. A. & Hastings, A. Highly Variable Spread Rates in Replicated Biological Invasions: Fundamental Limits to Predictability. *Science* **325**, 1536–1539 (2009).

42. Chen, J. & Lewis, O. T. Experimental heatwaves facilitate invasion and alter species interactions and composition in a tropical host-parasitoid community. *Global Change Biology* **29**, 6261–6275 (2023).

43. Sinclair, B. J. *et al.* Can we predict ectotherm responses to climate change using thermal performance curves and body temperatures? *Ecology Letters* **19**, 1372–1385 (2016).

44. McEvey, S. F. & Schiffer, M. New species in the Drosophila ananassae subgroup from northern Australia, New Guinea and the South Pacific (Diptera: Drosophilidae), with historical overview. *Records of the Australian Museum* **67**, 129–161 (2015).

45. Bürkner, P.-C. Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal* **10**, 395–411 (2018).

46. Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. *Journal of Statistical Software* **76**, (2017).

47. Chen, J. & Lewis, O. T. Climate and competition combine to set elevational distributions of tropical rainforest Drosophila. 2022.04.01.486700 Preprint at https://doi.org/10.1101/2022.04.01.486700 (2023).

48. Kontopoulos, D.-G., Sentis, A., Daufresne, M., Dell, A. I. & Pawar, S. No model to rule them all: a systematic comparison of 83 thermal performance curve models across traits and taxonomic groups. 2023.09.08.556856 Preprint at https://doi.org/10.1101/2023.09.08.556856 (2023).

49. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* **27**, 1413–1432 (2017).

50. Terry, J. C. D. Code and raw data to support Terry (2024) Experimental validation of ecological coexistence theory to forecast extinction under rising temperatures. Zenodo https://doi.org/10.5281/zenodo.10679556 (2024).