Location–scale models in ecology and evolution: heteroscedasticity in continuous, count and proportion data

Shinichi Nakagawa [1,2*], Santiago Ortega [1], Elena Gazzea [3], Malgorzata Lagisz [1, 2], Anna Lenz [1], Erick Lundgren [1], and Ayumi Mizuno [1*]

[1]Collaboration for Open Science and Synthesis in Ecology and Evolution (COSSEE), Department of Biological Sciences, Faculty of Science, University of Alberta, Edmonton, Canada

[2]Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia

[3]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Legnaro, Italy

[*]Corresponding author: Shinichi Nakagawa (snakagaw@ualberta.ca) and Ayumi Mizuno (amizuno@ualberta.ca)

November 3, 2025

Running headline: Location-scale model in ecology and evolution

The authors declare no conflict of interest: The authors declare no conflict of interest.

Author Statement: Conceptualization: Shinichi Nakagawa, Santiago Ortega, Elena Gazzea, Malgorzata Lagisz, Anna Lenz, Erick Lundgren, Ayumi Mizuno, Data curation: Santiago Ortega, Erick Lundgren, Ayumi Mizuno, Formal analysis: Shinichi Nakagawa, Santiago Ortega, Elena Gazzea, Malgorzata Lagisz, Anna Lenz, Erick Lundgren, Ayumi Mizuno, Funding Acquisition: Shinichi Nakagawa, Malgorzata Lagisz, Investigation: Shinichi Nakagawa, Santiago Ortega, Elena Gazzea, Malgorzata Lagisz, Anna Lenz, Erick

Lundgren, Ayumi Mizuno, Methodology: Shinichi Nakagawa, Santiago Ortega, Elena Gazzea, Malgorzata Lagisz, Anna Lenz, Erick Lundgren, Ayumi Mizuno, Project administration: Shinichi Nakagawa, Ayumi Mizuno, Supervision: Shinichi Nakagawa, Validation: Santiago Ortega, Erick Lundgren, Ayumi Mizuno, Elena Gazzea, Anna Lenz, Visualization: Ayumi Mizuno, Writing - Original Draft: Shinichi Nakagawa, Santiago Ortega, Erick Lundgren, Ayumi Mizuno, Writing - Review & Editing: Shinichi Nakagawa, Santiago Ortega, Elena Gazzea, Malgorzata Lagisz, Anna Lenz, Erick Lundgren, Ayumi Mizuno.

data availability: Data available via https://doi.org/10.5281/zenodo.17515632 Mizuno et al. (2025).

A step-by-step online tutorial is also available at https://ayumi-495.github.io/Eco_location-scale_model.

Acknowledgment: This research was made possible with the Canada Excellence Research Chairs, Government of Canada (CERC-2022-00074) and ARC (Australian Research Council) Discovery Project grant DP230101248. EG received funding from the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.4 – D.D. 1032 17/06/2022, CN00000022) within Agritech National Research Center.

Abstract

1. Biological data often violate the assumption of constant variance, yet such heteroscedasticity can reflect meaningful biological processes such as plasticity, canalization, or stress responses. Despite this, most models treat variance as statistical noise. Here, we reintroduce location–scale regression as a general framework that jointly models the mean (location) and variance (scale) components of a response. We describe three hierarchical extensions: 1) fixed-effects, 2) mixed-effects, and 3) double-hierarchical models, which allow researchers to formally test variance structures alongside mean effects, enhancing biological interpretation.

2. This framework is highly flexible and can extend beyond Gaussian assumptions to accommodate real-world data. The framework accommodates overdispersed, underdispersed, and zero-inflated count data through the use of negative binomial and Conway–Maxwell–Poisson distributions, and bounded proportion data through beta-binomial and beta regressions. Submodels can also be incorporated to account for structural zeros and ones when boundary outcomes are common. These extensions allow researchers to capture ecological processes such as presence–absence, success rates, and bounded response rates.

3. Using worked examples from published evolutionary and behavioral ecological studies, we illustrate how location–scale models can uncover biologically meaningful variance patterns that are overlooked in models focused solely on means. For instance, we show how food supplementation, hatching order, and predation risk influence not only average trait values but also their variability. Each example corresponds to one of the model types and is implemented using widely used R packages such as `glmmTMB` and `brms`. All examples are accompanied by a freely accessible, step-by-step online tutorial, thereby lowering technical barriers and fostering broader adoption of location–scale modeling in ecological and evolutionary research.

4. Finally, we propose a practical workflow for model selection and diagnostics and highlight recent extensions of the framework. These include multi-response models, meta-analytic models, phylogenetic comparative models, and models including shape parameters such as skewness. Treating variance as a biologically informative response opens new avenues for us to explore the evolutionary, ecological, and environmental processes that shape biological systems across diverse contexts.

Keywords— Bayesian statistics, distributional regression, GLMM, homoscedasticity, linear modeling, mixed-effects models, over-dispersion, zero-inflation

1 Introduction

Ecologists and evolutionary biologists strive to explain and account for variation in nature; this is usually done by statistically modeling target traits or measurements with hypothesized causal factors (e.g., a particular environmental factor accounts for 8% of the variance). In contrast, they rarely test whether variation changes across an environmental gradient or between groups (Cleasby and Nakagawa, 2011). Although ecological data often exhibit non-constant variance, this variation is commonly considered a mere nuisance that violates the model's assumption of homogeneity (i.e., homoscedasticity). In reality, patterns in variance, or heteroscedasticity, can signal ecological, evolutionary, and environmental processes. For example, environmental stress (e.g., temperature increases) can not only change the mean but can also generate more variance in organismal responses (e.g., Buckley and Huey, 2016; O'Dea et al., 2016). On the other hand, plasticity, such as learning, can canalize variability because most individuals uniformly reach the behavioral optimum (e.g., Baldwin, 1896; Crispo, 2007).

More than a decade ago, Cleasby and Nakagawa (2011) surveyed and reported that over 95% of published studies in behavioral ecology ignored heteroscedasticity. Such neglect can yield incorrect standard errors (SE) of regression coefficients (e.g., Type I error) and, critically, overlook biological insights in dispersion patterns. Therefore, they recommended two practical solutions. First, they suggested the use of heteroscedasticity-consistent ("sandwich") estimators of SE, which resolve the statistical issues such as inflated Type I error (Hayes and Cai, 2007). Second, one can model different residual variances for different groups or across a continuous predictor (i.e., heteroscedasticity). This approach, however, does not directly provide inferential statistics – whether changes in variance are statistically significant or not. In their paper, Cleasby and Nakagawa (2011) neglected the third option: location–scale regression modeling, which provides statistical inference on both mean (*location*) and variance (*scale*, also known as *dispersion*) and thus resolves all issues at once. Statistically, location–scale models remove bias in SE and test statistics under heteroscedasticity (Carroll and Ruppert, 1988; Zuur et al., 2009b). Biologically, these models can reveal when and how both mean and variance respond to environmental and other drivers.

The most flexible forms of location-scale models are double-hierarchical with random effects in both mean (*location*) and variance (*scale*)(Lee and Nelder, 1996, 2006; Rönnegård and Lee, 2013). However, these models are computationally complex and require Bayesian implementation, which may have hindered wider adoption. However, simpler location-scale models, which can only include random effects in the location part, are straightforward to implement in widely used statistical software. For example, these variants can be implemented readily in `glmmTMB` (Brooks et al., 2017) with minimal additional coding.

Therefore, we aim to reintroduce the utility of location–scale regression models. To facilitate broader use, we focus on two simpler, practical formulations sufficient for many applications. In the following sections, we first introduce location–scale models with only fixed effects on both mean (*location*) and variance (*scale*) (Model 1). Next, we extend these to include random effects on the *location* part (Model 2) and, for completeness, describe the double-hierarchical framework with random effects on both *location* and *scale* (Model 3). We then expand these models (mainly Model 2) to non-Gaussian responses, namely count and proportion data; although such data are common, modeling over-dispersion of count and proportion seems to be rare in ecology, evolution, and environmental sciences (cf., Bolker et al., 2009). These non-Gaussian location-scale models can handle zero-inflation, and we refer to the issues of under-dispersion and one-inflation. We provide a range of examples illustrating biological insights obtained from location-scale models with both frequentist and Bayesian implementations using `glmmTMB` and `brms` (Bürkner, 2017), respectively (see the online tutorial: link). We also suggest a practical workflow to guide model selection. Finally, we discuss broader applications of location-scale models (e.g., meta-analytic location-scale models; Nakagawa et al., 2025a) and related advanced models, which are potentially even more flexible and biologically informative (Rigby and Stasinopoulos, 2005).

2 From simple to location-scale regression (Model 1)

2.1 Model and motivation

We begin with the familiar simple regression models (only with fixed effects), where we assume constant residual variance as well as data independence:

$$y_i = \beta_0 + \sum_{k=1}^{K} \beta_k \, x_{ik} \; + \; e_i, \tag{1}$$

$$e_i \sim \mathcal{N}\big(0,\ \sigma^2\big), \tag{2}$$

where y_i is the response for observation i, x_{ik} $(k = 1, \ldots, K)$ are the fixed covariates (predictors), $\{\beta_0, \beta_1, \ldots, \beta_K\}$ are the regression coefficients, and the residual e_i is normally (Gaussian) distributed with mean zero and variance σ^2. Note that the predictor x_{ik} can be either a continuous or categorical variable. More accurately, for the latter case, when a categorical predictor has H levels, it becomes $H - 1$ 'dummy' variables or predictors. That is, a categorical variable becomes $(H - 1)$ binary variables in the model, and corresponding regression coefficients represent contrasts (differences) between a reference level (the intercept β_0) and another level.

Equivalently, we can write the model in its distributional form:

$$y_i \sim \mathcal{N}\big(\mu_i,\ \sigma^2\big), \tag{3}$$

$$\mu_i = \beta_0 + \sum_{k=1}^{K} \beta_k \, x_{ik}, \tag{4}$$

where μ_i denotes the expected value of y_i given the covariates, and σ^2 remains the constant variance.

This basic regression treats any heteroscedasticity as a nuisance. To turn it into biological/ecological signals, we allow the residual standard deviation to vary with predictors. The location–scale regression then comprises two linked submodels which can be written as (Model 1; Jorgensen, 1997; Lee et al., 2006; Cleasby et al., 2015):

$$y_i \sim \mathcal{N}\big(\mu_i,\ \sigma_i^2\big), \tag{5}$$

$$\mu_i = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ik}, \quad \text{(location submodel)} \tag{6}$$

$$\ln(\sigma_i) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ik}, \quad \text{(scale submodel)} \tag{7}$$

where μ_i is its expectation, modeled by the location submodel coefficients $\beta^{(l)}$ and covariates x_{ik}, and σ_i is the residual standard deviation, modeled on the log (ln) scale by the scale submodel coefficients $\beta^{(s)}$ and the same covariates.

This fixed effects location-scale regression, by linking predictors to both the mean and the ln(standard deviation), allows us to test if an environmental gradient or experimental treatment shifts not just the average response, but also its individual variability. In other words, if a predictor (x_{ik}) influences the mean, its corresponding regression coefficient ($\beta_k^{(l)} \neq 0$) will be non-zero (significant). If a predictor influences variance, the associated regression coefficient for the scale component, represented as ($\beta_p^{(s)} \neq 0$), will also be non-zero. This suggests that the heterogeneity in the data varies in relation to that predictor, a phenomenon referred to as heteroscedasticity. Translating variance signals into regression coefficients formalizes heterogeneity analysis and makes it accessible to researchers already familiar with interpreting regression coefficients for the location part (Fig. 1).

The syntax for writing location-scale models in `R` builds off familiar modeling syntax in `R`. To fit a (location-only) regression model on the relationship between the location (mean) of `y` by `x`, we would write the following.

```
library(glmmTMB)
location_model <- glmmTMB(y~x, data = dt)
```

To explicitly model the ‘scale’ as well as ‘location’, we simply add the same formula (without the response variable) to the `dispformula` argument.

```
location_scale_model <- glmmTMB(y~x, dispformula = ~ x, data = dt)
```

This second model returns two regression tables, one (referred to as the `Conditional` by `glmmTMB`) describes the relationship between `x` and mean `y`, while the second table (referred to as `Dispersion`) describes the relationship between `x` and the variance of `y`.

2.2 Illustrative example

In the following illustrative examples, we report representative model results using the R packages `glmmTMB` and/or `brms`, selected based on model type and functionality. Full model specifications, code, and detailed explanations of datasets and interpretations are available in our online tutorial (link), where we also explain how to interpret regression coefficients on the log scale in terms of percentage change in details.

We reanalyzed whether early-life food supplementation had sex-specific effects on body size variability, using adult tarsus length as an indicator, in a wild population of house sparrows (*Passer domesticus*) on Lundy Island, England (Cleasby and Nakagawa, 2011). The model's location component showed no significant effect of sex, treatment, or their interaction on mean adult tarsus length. However, the scale (dispersion) component revealed a significant negative interaction between sex and treatment (`glmmTMB`: $\beta^{(s)}_{\text{[interaction]}} = -0.95$, 95% CI $[-1.66, -0.24]$), which corresponds to a 61.3% lower residual SD for supplemented males relative to baseline (non-supplemented females) (% change in SD $= 100[\exp(\beta^{(s)}) - 1]$; CI -81.0% to -21.3%) and 58.2% lower than non-supplemented males ($\beta^{(s)}_{\text{[male-female]}} + \beta^{(s)}_{\text{[interaction]}} = -0.87$; $100[\exp(-0.87) - 1] = -58.2$). Neither treatment nor sex alone significantly influenced variance. This suggests early-life food supplementation can canalize trait development, leading to more uniform adult male morphology under favorable nutritional conditions.

3 Adding random effects in the location part only (Model 2)

3.1 Model and motivation

Ecological and environmental datasets often violate both the homoscedasticity and non-independence assumptions. The latter is common due to clustered or grouped data, such as multiple measurements *per* site or individual. Consequently, 'mixed-effects' models are widely used in ecology and evolution, as they incorporate both fixed and random effects to model these clustering and grouping structures (Bolker et al., 2009; Nakagawa and Schielzeth, 2013).

Introducing a random effect (intercept in the location submodel) allows each group j to have a group-specific mean, while keeping the scale model fixed-effects only. Such models can be written as (Model 2; Jorgensen, 1997; Lee et al., 2006; Cleasby et al., 2015):

$$y_{ij} \sim \mathcal{N}\big(\mu_{ij},\ \sigma_{ij}^2\big), \tag{8}$$

$$\mu_{ij} = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{9}$$

$$\ln(\sigma_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{10}$$

where the random intercept $u_j^{(l)}$ is distributed as $u_j^{(l)} \sim \mathcal{N}(0,\ \sigma_u^2)$. Here y_{ij} is the i-th response in group j, μ_{ij} its expected value including the group-specific shift $u_j^{(l)}$, and σ_{ij} the residual standard deviation driven by the scale covariates alone. This (mixed-effects) location–scale model tests whether predictors affect both the mean across and within groups, while allowing groups to differ in their overall mean level.

It should be noted that Model 2's location submodel has the simplest random effect structure and, in practice, this submodel may need to have more than one random effect (intercept) and random slopes. Indeed, such models with multiple random effects may be the rule rather than an exception in ecological and evolutionary data (e.g., site and year, or individuals nested in sites Schielzeth and Nakagawa, 2013).

3.2 Illustrative example

We re-examined the difference in fledging scaled mass index (SMI), i.e., mass corrected by body size, between first- and second-hatched blue-footed booby (*Sula nebouxii*) chicks (Drummond et al., 2025). This Gaussian location–scale model included nest identity ($\sigma_{\text{Nest ID}(l)}$) and hatching year ($\sigma_{\text{hatching.year}(l)}$) as random effects in the location submodel, and hatching order in both submodels. We found a mean ln(SMI) difference between first- and second-hatched chicks (`brms`: $\beta^{(l)}_{[\text{first–second}]} = -0.02$, 95% CI $[-0.02, -0.01]$), which on the response scale corresponds to a $\sim$ 2.0% lower mean SMI for second-hatched ($\exp(\beta^{(l)}) = 0.980$; CI $\approx$ 0.980–0.990; % change $\approx -2.0\%$ to -1.0%). Moreover, second-hatched chicks exhibited greater ln(SMI) variability compared to their first-hatched counterparts (`brms`: $\beta^{(s)}_{[\text{first–second}]} = 0.13$, 95% CI $[0.08, 0.18]$), implying a +13.9% residual SD (CI +8.3% to +19.7%). Random effects in the location component also showed that average ln(SMI) differed between nests (`brms`: $\sigma_{\text{Nest ID}} = 0.05$, 95% CI $[0.04, 0.05]$; multiplicative spread $\exp(0.05) = 1.051$, i.e., $\sim$ +5.1%) and hatching years (`brms`: $\sigma_{\text{hatching.year}} = 0.10$, 95% CI $[0.07, 0.14]$; $\exp(0.10) = 1.105$, i.e., $\sim$ +10.5%). These results suggest that second-hatched chicks not only have a slightly lower average ln(SMI) but also exhibit greater variability in their SMI compared to first-hatched chicks.

4 Double-hierarchical model (Model 3)

4.1 Model and motivation

Model 2 naturally begs a question: why do not add random effects in the scale part? Indeed, "double-hierarchical" models were the first to arrive in ecology and evolution nearly a decade ago (e.g., Westneat et al., 2013). The double-hierarchical formulation jointly models how each group j shifts its mean and its standard deviation on the natural logarithm scale (Model 3; Lee and Nelder, 1996, 2006; Cleasby et al., 2015; O'Dea et al., 2022):

$$y_{ij} \sim \mathcal{N}\big(\mu_{ij},\ \sigma_{ij}^2\big), \tag{11}$$

$$\mu_{ij} = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{12}$$

$$\ln(\sigma_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk} + u_j^{(s)}, \quad \text{(scale submodel)} \tag{13}$$

with the bivariate random-effect vector $\big(u_j^{(l)},\, u_j^{(s)}\big)^\top$ following

$$\begin{pmatrix} u_j^{(l)} \\ u_j^{(s)} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0},\ \begin{pmatrix} \sigma_{u(l)}^2 & \rho_u\, \sigma_{u(l)}\, \sigma_{u(s)} \\ \rho_u\, \sigma_{u(l)}\, \sigma_{u(s)} & \sigma_{u(s)}^2 \end{pmatrix}\right). \tag{14}$$

Here, each group j has its own intercept in the mean ($u_j^{(l)}$) and in the ln-standard deviation ($u_j^{(s)}$), with their covariance governed by ρ_u. A positive ρ_u implies that groups with higher means also exhibit greater variability, whereas a negative ρ_u indicates that high-mean groups are more tightly canalized. This full double-hierarchical model thus allows simultaneous inference on fixed effects and group-level mean–variance associations. An extension of this model with a random slope in both location and scale parts in the context of uni- and multi-variate cases is well described in O'Dea et al. (2022). For example, when the cluster u_j represents individuals (y_{ij} is repeated behavioral measures of an individual), the parameter ρ_u is referred to as the personality-predictability association. This is because $\sigma_{u(l)}^2$ reflects between-individual differences in mean behavior (personality), while $\sigma_{u(s)}^2$ captures differences in behavioral variance (predictability). For instance, a positive correlation would indicate that more aggressive individuals are also more unpredictable in the intensity of their aggression at one time point.

As described, our focus in this article is to highlight Model 2 (and Model 1). Therefore, even if one is interested in $\sigma_{u(s)}^2$ and ρ_u, one should start with Model 2 as a robust baseline. One can fit Model 3, and compare Models 2 and 3 using information criteria or likelihood-ratio tests, if sample size permits (more than 10 repeats or observations per group may be required to model $\sigma_{u(s)}^2$ reliably; O'Dea et al., 2022); indeed, a simple simulation reveals that one requires 20 observations to get unbiased variance estimates (see the online tutorial (link)). Such a modeling strategy leverages the stability of Model 2 while allowing the richer inferences of Model 3 when data permit (for more on model selection, see Section 7).

4.2 Illustrative example

Building upon the previous example of fledging scaled mass index (SMI) (Drummond et al., 2025), we fitted a double-hierarchical Gaussian location–scale model. This extended Model 2 by incorporating nest identity ($\sigma_{\text{Nest ID}}$) as a correlated random effect in both the location and scale submodels. This allowed us to assess how average ln(SMI) and its variability differed across nests, and if these nest-specific variations were related. Average ln(SMI) differed between nests (`brms`: $\sigma_{\text{Nest ID}(l)} = 0.05$, 95% CI $[0.04, 0.05]$; $\exp(0.05) = 1.051$, $\sim +5.1\%$), and some nests showed greater ln(SMI) variability (`brms`: $\sigma_{\text{Nest ID}(s)} = 0.36$, 95% CI $[0.32, 0.40]$; $\exp(0.36) = 1.433$, i.e., +43.3% SD; CI $\approx +37.7\%$ to $+49.2\%$). Notably, a negative correlation between location and scale random effects within nests (`brms`: $\rho_{\text{Nest ID}} = -0.46$, 95% CI $[-0.58, -0.33]$) indicated that nests with higher average ln(SMI) tended to exhibit lower variability; a +1 SD increase in a nest's location effect is associated with an expected change of $\rho_u \sigma_{u(s)} \approx -0.166$ on ln(SD), i.e., $\approx -15.3\%$ SD $\big(\exp(-0.166) - 1\big)$, with a rough range of $\sim -20.7\%$ to -10.0% across the CI limits. Fixed effects for hatching order remained consistent with our previous model, further supporting that second-hatched chicks have slightly lower mean ln(SMI) and greater variability.

5 Beyond Gaussian I: over-dispersed count data

In this and the next section, we turn from Gaussian responses to non-Gaussian data common in the natural world. Our focus is deliberately selective: we concentrate on count and proportion responses, omitting ordinal outcomes despite their feasibility with location-scale models (e.g., Martin et al., 2017). For these two response variable types, we develop three practical formulations for researchers. Because structural zeros (and ones for proportions) are common in ecological and environmental datasets, some count and proportion models include zero- or zero/one-inflation components (submodels). To keep the description clear, we present each model with the single random-intercept structure for the location, introduced in Model 2, though Models 1 and 3 forms are also applicable.

5.1 Negative-binomial location–scale model

Many ecological questions involve integer counts: fledglings per nest, insect colony size, or the number of eco- or endo-parasites. While Poisson regression is the usual starting point, real data rarely meet its assumption that mean equals variance (i.e., $E[y] = Var[y]$). Indeed, as many researchers know, count data often exhibit over-dispersion ($E[y] < Var[y]$). Negative-binomial regression offers a solution because the negative-binomial (NB) distribution (family) has an extra parameter to model this over-dispersion (Stoklosa et al., 2022).

A negative-binomial location scale model – in the form of Model 2 (a random effect only in the location part) – can be written as (Jorgensen, 1997; Lee and Nelder, 1996, 2006):

$$y_{ij} \sim \mathrm{NB}\big(\mu_{ij},\, \theta_{ij}\big), \tag{15}$$

$$\ln(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{16}$$

$$\ln(\theta_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{17}$$

where y_{ij} is the count for observation i in group j, μ_{ij} is its mean, linked via a log (ln) link to fixed covariates x_{ijk} and a group-level random intercept $u_j^{(l)}$, θ_{ij} is the dispersion parameter, linked on the ln scale to the same covariates but no random effect, $u_j^{(l)} \sim \mathcal{N}(0, \sigma_u^2)$ captures group-level shifts in the mean, and the log links ensure $\mu_{ij}, \theta_{ij} > 0$. The parameter θ_{ij} is analogous to the Gaussian dispersion parameter σ_{ij} but is quite different; it calibrates over-dispersion, and a larger value of θ_{ij} represents less variation. This role becomes clear when one sees the formula for variance for the negative-binomial distribution. $\mathrm{Var}(Y_{ij}) \;=\; \mu_{ij} + \mu_{ij}^2/\theta_{ij}$ so that as $\theta_{ij} \to \infty$, the term μ_{ij}^2/θ_{ij} vanishes and the distribution approaches the Poisson mean-variance expectation ($E[y] = Var[y]$); conversely, smaller θ_{ij} produces increasingly strong over-dispersion relative to the Poisson expectation. It should be noted that there are alternative parametrisations of negative-binomial models, for example, in `glmmTMB`.

5.2 Zero-inflated negative-binomial location-scale model

Ecological and evolutionary applications frequently encounter count data with both an excess of true absences alongside over-dispersed counts (cf., Zuur et al., 2009a). For example, surveying soil invertebrates across patchy habitats might yield samples with zero individuals (structural zeros) and others with wildly varying densities. Similarly, parasite counts in wildlife often include hosts with no infection and others with heavy infections (Taylor et al., 2017, e.g.,). To model these dual processes while allowing for distinct underlying distributions across populations or sites, we embed a single random intercept in the location submodel of a zero-inflated negative-binomial location-scale framework:

$$y_{ij} \sim \begin{cases} 0, & \text{with probability } \pi_{ij}, \\ \\ \mathrm{NB}\big(\mu_{ij},\, \theta_{ij}\big), & \text{with probability } 1 - \pi_{ij}, \end{cases} \tag{18}$$

$$\mathrm{logit}(\pi_{ij}) = \beta_0^{(0)} + \sum_{k=1}^{K} \beta_k^{(0)}\, x_{ijk}, \quad \text{(zero-inflation submodel)} \tag{19}$$

$$\ln(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{20}$$

$$\ln(\theta_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{21}$$

where y_{ij} is the count for observation i in group j. The zero-inflation submodel predicts the probability π_{ij} of a guaranteed zero via a logit link and fixed covariates x_{ijk}. Here, $\beta_0^{(0)}$ is the baseline log-odds of an excess zero when all covariates $x_{ijk} = 0$, and each $\beta_k^{(0)}$ represents the change in log-odds of a guaranteed zero per unit increase in covariate x_{ijk}. A positive $\beta_k^{(0)}$ thus indicates that higher values of x_k increase the probability of structural absence, whereas a negative $\beta_k^{(0)}$ decreases it. The location submodel predicts $\mu_{ij} > 0$ via a log link, including the group-specific random intercept $u_j^{(l)} \sim \mathcal{N}(0, \sigma_u^2)$, which captures unobserved differences among groups. The scale submodel with fixed covariates alone governs the dispersion parameter $\theta_{ij} > 0$, so larger θ_{ij} yields variance closer to the mean, as described above.

This model formulation allows researchers to simultaneously investigate how habitat characteristics and evolutionary history influence (1) the chance of encountering no individuals at all, (2) the expected abundance when presence occurs, and (3) the degree of overdispersion beyond the Poisson expectation. Notably, Stoklosa et al. (2022), in their review of negative-binomial modeling, advocate for negative-binomial models as a default for count data in ecology and biodiversity, given their near-ubiquitous over-dispersion.

5.3 Conway–Maxwell–Poisson location–scale model

Under-dispersion ($Var(Y) < E[Y]$) is probably less common but potentially important in ecological and environmental datasets. For example, stabilizing selection and biological ceiling (floor) effects could canalize count data (in this case, the ceiling effect means that values cannot go over a certain upper biological limit, while the floor effect means a lower limit). The Conway–Maxwell–Poisson (CMP) family (distribution) spans under- and over-dispersion with a parameter ν (variance drops as $\nu\uparrow$) (Sellers and Shmueli, 2010):

$$y_{ij} \sim \mathrm{CMP}(\mu_{ij}, \nu_{ij}), \tag{22}$$

$$\ln(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{23}$$

$$\ln(\nu_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{24}$$

where y_{ij} is the count for observation i in group j; $\mu_{ij} > 0$ is the CMP "rate" (mean, often denoted as λ), on the log scale linked to predictors x_{ijk} and a random intercept $u_j^{(l)} \sim \mathcal{N}(0, \sigma_u^2)$, and $\mathrm{Var}(Y_{ij}) \approx \mu_{ij}^{\nu_{ij}^{-1}}\, \nu_{ij}^{-1}$: $\nu > 0$ represents under-dispersion, $\nu = 1$ recovers the Poisson $\mathrm{Var}(Y) = E[Y]$ yields over-dispersion, and $\nu > 1$ under-dispersion.

By fitting this mixed-effects location–scale CMP model, ecologists and environmental biologists can probe not only how drivers such as resource availability, temperature stress, or habitat fragmentation shift the average count of organisms or events, but also whether these same forces tighten or loosen the Poisson expectation on variability. Notably, Brooks et al. (2019) points out the dual ability of CMP to deal with both overdispersion and underdispersion. Moreover, they introduce zero-inflated CMP models (ZICMP) using `glmmTMB` (Brooks et al., 2019). As we mentioned earlier, its capability to model underdispersion is important, because this cannot be done by negative-binomial models. For example, under strong stabilizing selection on clutch size, many bird species have evolved canalized brood counts, often producing almost exactly the same number of eggs each year, a pattern of under-dispersion captured by $\nu > 1$ (e.g., Boyce and Perrins, 1987; Liou et al., 1993; Santos and Nakagawa, 2013).

5.4 Illustrative example

We reanalyzed visual preference in Estrildid finches by measuring gaze frequency to dot stimuli under food-supplied and food-deprived conditions (Mizuno and Soma, 2023). To account for overdispersed count data, we used a negative-binomial location–scale model (corresponding to Model 2), with species and individual (nested within species) as random effects in the location component. Birds gazed significantly less at dots when food was supplied (`glmmTMB`: $\beta^{(l)}_{\text{deprived–supplied}} = -0.85$, 95% CI $[-1.08, -0.61]$), which corresponds to a rate ratio of $\exp(\beta^{(l)}) = 0.427$ (i.e., -57.3% mean; CI -66.0% to -45.7%). The scale component revealed greater individual-level variability under deprivation, indicated by a negative effect on θ ($\beta^{(s)}_{\text{deprived–supplied}} = -0.66$, 95% CI $[-1.15, -0.18]$), giving a θ-ratio of $\exp(\beta^{(s)}) = 0.517$ (i.e., -48.3% precision; CI -68.3% to -16.5%), noting that lower θ implies more scatter than the Poisson expectation. Species-level variation in average gaze frequency (SD $= 0.55$, 95% CI $[0.31, 0.99]$) exceeded within-species individual variation (SD $= 0.34$, 95% CI $[0.17, 0.68]$). Thus, food deprivation increased average gazing, while availability reduced gazing but amplified individual variability.

6 Beyond Gaussian II: over-dispersed proportion data

Proportions come in two flavors. Discrete (binomial) proportions arise as “successes out of trials”, for example, the number of germinated seeds out of 20, the tally of infected hosts in a sample, the abundance of a certain taxon in microbial communities. They are naturally modeled with binomial regression (e.g., Bolker et al., 2009; Zuur et al., 2009b). Continuous proportions, in contrast, are already measured as rates on the unit interval, $[0, 1]$ – leaf-area loss, percent cover, the fraction of time an animal spends foraging. Continuous proportions are usually analyzed with beta regression, which takes values between 0 and 1 (Ferrari and Cribari-Neto, 2004; Douma and Weedon, 2019).

Boundary values (i.e., 0 and 1) complicate matters differently for the two types of proportion. Because the binomial distribution already includes zero and n (the number of ‘trials’), discrete counts can generate observed proportions of exactly 0 or 1; yet in practice, true absences (e.g., empty traps and seeds that could never germinate) often occur more frequently than a binomial distribution can allow (cf., Warton, 2005). A zero-inflation component, therefore, captures a separate “structural-zero” process. In contrast, structural ones (a one-inflation component) are seldom, if ever, needed because excess of perfect ‘successes’ are unlikely to occur in nature (e.g., Zuur et al., 2009b). Beta regression models, by construction, exclude the boundaries of the unit interval, so when continuous proportions include any zeros or ones – for example, sprayed plots with 0 % damage, or quadrats that are completely vegetated – both zeros and ones must be modeled via zero- and one-inflation submodels respectively (Ospina and Ferrari, 2012). Yet, we note that there exist methods to rescale contentious proportion data to eliminate zeros and ones, especially when these values are rare (e.g., lemon squeezer transformation; Smithson and Verkuilen, 2006). Nevertheless, whenever possible, it is advantageous to model zeros and ones explicitly, because these values can be due to some ecological or evolutionary processes. Bearing this in mind, we introduce three location-scale models for proportion data below.

6.1 Beta-binomial location–scale model

For discrete proportions (e.g., seedling emergence, infection prevalence), one usually starts modeling by assuming a binomial distribution:

$$y_{ij} \sim \text{Binomial}\big(n_{ij},\, \mu_{ij}\big), \tag{25}$$

where y_{ij} is the number of successes out of n_{ij} trials in group j and $\mu_{ij} \in (0, 1)$ is the underlying success probability (often denoted p). Yet, a binomial distribution ‘fixes’ the variance at $n_{ij}\mu_{ij}(1 - \mu_{ij})$ (i.e., the binomial-variance expectation) and therefore cannot accommodate the extra-binomial dispersion that is common in field data.

However, if we assume that the success probability itself varies among observational units according to a beta distribution, $\mu_{ij} \sim \text{Beta}(\alpha_{ij}, \beta_{ij})$, we can combine these two distributions to yield a beta–binomial distribution:

$$y_{ij} \sim \text{Beta-binomial}\big(n_{ij},\ \mu_{ij},\ \phi_{ij}\big), \tag{26}$$

where the beta distribution's parameters are reparameterized as $\alpha_{ij} = \mu_{ij}\,\phi_{ij}$ and $\beta_{ij} = (1-\mu_{ij})\,\phi_{ij}$. Here $\phi_{ij} > 0$ is a precision (inverse-dispersion or inverse-variance) term. For the resulting beta–binomial the variance is $\text{Var}(y_{ij}) = n_{ij}\,\mu_{ij}(1-\mu_{ij})\,((n_{ij}+\phi_{ij})/(1+\phi_{ij}))$. When $\phi_{ij} \to \infty$, the fraction $(n_{ij}+\phi_{ij})/(1+\phi_{ij})$ to 1; the variance collapses to the binomial-variance expectation $n_{ij}\,\mu_{ij}(1-\mu_{ij})$ and there is no over-dispersion. Therefore, ϕ has the same role as the θ over-dispersion parameter in the negative binomial distribution. Given this property of a beta-binomial distribution, we can let predictors explain both the mean success probability and the amount of extra dispersion, while allowing for group-level shifts in the mean (Jorgensen, 1997; Lee and Nelder, 1996, 2006):

$$\text{logit}(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\,x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{27}$$

$$\ln(\phi_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\,x_{ijk}. \quad \text{(scale submodel)} \tag{28}$$

In the location submodel, the random intercept $u_j^{(l)} \sim \mathcal{N}(0, \sigma_u^2)$ captures baseline differences among sites or populations. The scale submodel links the ln-precision to the same (or different) covariates, so predictors can inflate ($\phi_{ij} \downarrow$) or dampen ($\phi_{ij} \uparrow$) the variation beyond the binomial-variance expectation. Relatedly, Martin et al. (2020) introduced the use of the beta-binomial location-scale model to quantify the relative abundance of a specific taxon in microbial communities (genetic sequencing of microbiome samples results in discrete proportion data). They indeed emphasized the importance of its ability to model dispersion.

6.2 Zero-inflated beta–binomial location–scale model

In many ecological discrete proportion data (e.g., seedling emergence, infection prevalence), counts of "successes" out of n_{ij} trials show both structural zeros (true absences) and extra-binomial scatter. A zero-inflated beta–binomial location–scale model accommodates: 1) a point-mass at zero, 2) group-level shifts in the mean, and 3) over-dispersion beyond the binomial expectation, all within a single framework:

$$y_{ij} \sim \begin{cases} 0, & \text{with probability } \pi_{ij}, \\ \text{Beta-binomial}\big(n_{ij},\, \mu_{ij},\, \phi_{ij}\big), & \text{with probability } 1 - \pi_{ij}, \end{cases} \tag{29}$$

$$\operatorname{logit}(\pi_{ij}) = \beta_0^{(0)} + \sum_{k=1}^{K} \beta_k^{(0)}\, x_{ijk}, \quad \text{(zero-inflation submodel)} \tag{30}$$

$$\operatorname{logit}(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{31}$$

$$\ln(\phi_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{32}$$

Here y_{ij} is the number of successes in n_{ij} trials for observation i in group j. The zero-inflation submodel predicts the probability π_{ij} of a "structural" zero via a logit link and covariates x_{ijk}. Conditional on non-zero counts, the beta–binomial component arises by assuming the success probability itself follows $\text{Beta}(\,\mu_{ij}\,\phi_{ij},\,(1-\mu_{ij})\,\phi_{ij}\,)$. The location submodel – with its random intercept $u_j^{(l)}$ – captures baseline differences among sites or populations, while the scale submodel lets covariates modulate the precision ϕ_{ij}.

Similar to Martin et al. (2020), Hu et al. (2018) proposed zero-inflated beta-binomial models for microbiome data. While not full location-scale models, their examples underscore the importance of modeling zeros in such data.

6.3 Zero-and-one-inflated beta location–scale model

Continuous proportions often include exact zeros or ones (e.g., complete absence or saturation), which standard beta regressions cannot accommodate. Zero-and-one-inflated beta models resolve this by mixing three submodels to estimate coefficients for point mass at 0, point mass at 1, and the beta-distributed interior (Ospina and Ferrari, 2012). This approach models the occurrence of boundary outcomes and the variability of intermediate proportions in a single, interpretable framework, without dropping or adjusting boundary data:

$$y_{ij} \sim \begin{cases} 0, & \text{with probability } \pi_{0,ij}, \\ 1, & \text{with probability } \pi_{1,ij}, \\ \text{Beta}\big(\mu_{ij}\,\phi_{ij},\ (1-\mu_{ij})\,\phi_{ij}\big), & \text{with probability } 1-\pi_{0,ij}-\pi_{1,ij}, \end{cases} \tag{33}$$

$$\text{logit}(\pi_{0,ij}) = \beta_0^{(0)} + \sum_{k=1}^{K} \beta_k^{(0)}\, x_{ijk}, \quad \text{(zero-inflation submodel)} \tag{34}$$

$$\text{logit}(\pi_{1,ij}) = \beta_0^{(1)} + \sum_{k=1}^{K} \beta_k^{(1)}\, x_{ijk}, \quad \text{(one-inflation submodel)} \tag{35}$$

$$\text{logit}(\mu_{ij}) = \beta_0^{(l)} + \sum_{k=1}^{K} \beta_k^{(l)}\, x_{ijk} + u_j^{(l)}, \quad \text{(location submodel)} \tag{36}$$

$$\ln(\phi_{ij}) = \beta_0^{(s)} + \sum_{k=1}^{K} \beta_k^{(s)}\, x_{ijk}, \quad \text{(scale submodel)} \tag{37}$$

Here $\pi_{0,ij}$ and $\pi_{1,ij}$ are the structural-zero and structural-one probabilities; μ_{ij} and ϕ_{ij} govern the continuous beta component; and $u_j^{(l)}$ is the lone random intercept in the location submodel, allowing group j to differ in its baseline mean proportion. The parameters $\beta_0^{(0)}$ and $\beta_k^{(0)}$ set the log-odds of an exact zero, while $\beta_0^{(1)}$ and $\beta_k^{(1)}$ set the log-odds of an exact one; each as a function of covariates. The variance of the beta-distributed interior is $\text{Var}(y_{ij}) = \mu_{ij}(1-\mu_{ij})/(1+\phi_{ij})$.

When $\phi_{ij} \to \infty$ the dispersion shrinks to zero and data distribution concentrates around its mean, whereas as ϕ_{ij} approaches zero, the variance approaches its maximum $\mu_{ij}(1-\mu_{ij})$. Thus, lower ϕ_{ij} inflates and higher ϕ_{ij} deflates variability around the mean, and the scale submodel lets predictors modulate dispersion separately from the mean process. Note that when data does not include zeros and ones, one can remove corresponding submodels (i.e., beta location-scale models).

Burke et al. (2023) used a zero-inflated beta location-scale model – without one-inflation as their dataset did not have ones — to examine patterns and drivers of coral diseases (measured by percentage areas of diseased corals) in a meta-analytic context (see Section 8). They found that when sea surface temperature increases, not only did the mean percentage of coral disease increase, but so did its variability, and, surprisingly, the observations of zero-percent disease, too.

6.4 Illustrative example

Lundgren et al. (2022) investigated whether mountain lion predation reduced feral donkey impacts on desert wetlands. We re-analysed some of these data with a beta location–scale model. We included zero–one inflation and conditional one-inflation submodels to account for exact 0 and 1 values. The model revealed that on average, the log-odds of the mean percentage of trampled bare ground were lower in areas with high predation risk (`brms`: $\beta^{(l)}_{\text{predation–no predation}} = -1.22$, 95% CI $[-2.27, -0.71]$), corresponding to an odds ratio of $\exp(\beta^{(l)}) = 0.295$ (i.e., -70.5%; CI -89.7% to -50.9%). The scale component showed that log-precision (phi) was lower at sites with predation (ϕ: $\beta^{(s)}_{\text{predation–no predation}} = -1.07$, 95% CI $[-2.01, -0.04]$), implying a ϕ-ratio of $\exp(\beta^{(s)}) = 0.343$ (i.e., -65.7% precision; CI -86.6% to -3.9%) and therefore more variation in trampling in areas with predation. See our tutorial (link) for the R code and interpretation of zero and one-inflated submodels.

7 Proposed workflow and diagnostics

Before any plotting or fitting, we recommend identifying whether the biological question concerns (i) changes in the mean alone, (ii) changes in variance (e.g., canalization, predictability, plasticity), or (iii) both. If variance is central to inference, one should start with a location–scale specification so that dispersion is modeled, estimated, and interpreted from the outset (e.g., Cleasby and Nakagawa, 2011; Nakagawa et al., 2025b). If variance is plausibly constant, a location-only baseline may be reasonable; however, one should verify this with targeted diagnostics (below) before concluding that homoscedasticity holds.

Then, one could plot the raw response against each predictor to look for fans/funnels and group-wise spread for categorical predictors (heteroscedasticity cues). For non-Gaussian or transformed responses, one could also display the data (or fitted means) on the model's link scale (log, logit) to align visualization with the inferential scale and avoid misreading curvature or boundary effects. When helpful, one might pair response-scale and link-scale panels in figures to aid interpretation. When a location-only baseline is fitted (Gaussian regression or mixed model with random intercepts for clustering), standard Q–Q plots are used to diagnose Gaussian residual assumptions, but they are not suitable for discrete responses. One should use randomized quantile residuals to obtain uniform residual checks for any GLM/GLMM family; these quickly reveal dispersion misfit, zero/one inflation, and other distributional problems (Dunn and Smyth, 1996). Such residuals can be calculated by the R package `DHARMa` (Hartig, 2022).

These graphics could indicate whether a location-only model is defensible or whether a scale submodel is needed (see Fig. 2). If diagnostics or the question motivate it, one should specify a scale submodel for the residual SD (Gaussian), the dispersion/precision (negative binomial θ, beta-binomial ϕ), or other variance-link parameters (e.g., CMP ν). With sufficient replication per cluster (roughly >5 observations), random effects can be introduced in the scale part and, where scientifically motivated, a correlation between location and scale random effects can be modeled (double-hierarchical formulation). If possible, one should treat such expansions as hypothesis-driven, not purely data-driven, although there are cases where pure exploration is warranted.

Information criteria can triage candidate models, but should not replace subject-matter reasoning. One could use AIC (Akaike information criterion) for frequentist fits and WAIC (widely applicable information criterion) or LOO-CV (leave-one-out cross validation) for Bayesian fits (Akaike, 2003; Anderson and Burnham, 2004; Vehtari et al., 2017). Then, one should inspect whether retained models answer biological questions of interest. For Bayesian implementations, one should routinely check any convergence/mixing issues before interpreting parameters via, for example, Gelman-Rubin statistics ($\hat{R} \approx 1$) and effective sample sizes; posterior predictive checks can be performed to see predictions (mis-)match observed data (Gelman and Rubin, 1992; Vehtari et al., 2017; Gelman et al., 2020). Furthermore, regardless of the approaches (frequentist/Bayesian), one could go as far as performing simulation-based validation by generating data under the fitted generative process at the same or comparable sample size, refitting, and reporting bias and coverage for key model parameters. These and relevant procedures are increasingly available and supported by recent statistical tools (Säilynoja et al., 2025; Modrák et al., 2025; Monnahan et al., 2017; Allegue et al., 2017).

Importantly, we can report location effects on their natural or link scale, whichever makes biological sense. For scale submodels, it is easy to interpret if we report percentage change in SD or dispersion for a unit change in a predictor, e.g., % change in SD $= 100\,[\exp(\beta^{(s)}) - 1]$; if $\beta^{(s)}$ is a contrast between two groups, it represents % change in SD from one group to the other, as we have done in our examples above. In this way, we could emphasise magnitudes and uncertainty for biological interpretation rather than solely relying on interval overlap with zero, i.e., statistical significance (Nakagawa and Cuthill, 2007).

8 Further extensions and future perspectives

Location–scale thinking invites a broader re-imagination of data analysis. To assist this, we describe four extensions that expand the analytical capability to understand variability and quantify heteroscedasticity. First, ecological and environmental traits/measurements rarely act in isolation. Multivariate location–scale models analyze suites of traits simultaneously, estimating covariances not only among means but also among variances, and even mean–variance cross-links among traits. Such models can test, for instance, whether life-history 'syndromes' involve coordinated changes in both average values and trait predictability, or whether plasticity in one dimension buffers variability in another (O'Dea et al., 2022).

Second, Blowes (2024) and Nakagawa et al. (2025a) have introduced and highlighted that bringing location–scale thinking into ecological and evolutionary meta-analysis would allow evidence syntheses to ask when and why heterogeneity among effect sizes change along environmental gradients and methodological differences. Meta-analytic location–scale models treat heterogeneity, which dominates ecological and evolutionary meta-analyses (Senior et al., 2016), as a parameter to be explained rather than tolerated. As such, these models can uncover hidden structure in the “noise” of published effect sizes. Indeed, using several datasets from community ecology, Blowes (2024) showed that location-scale meta-regression can significantly improve model fit compared to location-only meta-regression.

Third, Halliwell (2025) and Nakagawa et al. (2025b) have introduced phylogenetic location-scale models, emphasizing that variance itself can evolve and should be a part of macro-evolutionary and community-ecological investigation. Embedding phylogenetic covariance structures in both mean and variance sub-models opens new terrain for comparative biology. A phylogenetic location–scale model can reveal whether evolutionary shifts in trait means are accompanied by shifts in trait variability, and whether certain clades are consistently more (or less) variable than expected. By quantifying “phylogenetic heritability” for variance and means, researchers gain a fuller picture of evolutionary constraints, innovations and trade-offs.

Fourth, responses not only have location and scale but also have ‘shape’. Extending the framework to include a shape component (e.g., skewness, kurtosis or heavy tails) would ask how entire distributions shift under ecological, evolutionary and environmental change (Stemkovski et al., 2023; Cornwell and Ackerly, 2009). ‘Location–scale–shape’ models are already feasible in generalized additive or flexible Bayesian settings (Rigby and Stasinopoulos, 2005; Corrales and Cepeda-Cuervo, 2022; Stasinopoulos and Rigby, 2008; Umlauf et al., 2021). Such models promise insights into the frequency of extreme events, asymmetric risks, stabilizing selection, and bet-hedging strategies (Pick et al., 2022; Starrfelt and Kokko, 2012; Pollo et al., 2025; Anderson et al., 2017).

Collectively, these extensions remind us that mean responses are only the tip of the statistical iceberg. Embracing location, scale and (eventually) shape as joint products of ecological and evolutionary processes will deepen our understanding of how organisms and ecosystems respond to an increasingly variable world.

9 Conclusions

Location-scale models provide a powerful lens through which ecologists and evolutionary biologists can interpret different types of data (i.e., continuous, count and proportion data). Building on the call from Cleasby and Nakagawa (2011) to treat heteroscedasticity as a biological clue and process, these approaches offer both conceptual and practical tools for richer inference. As datasets grow larger and more complex, studying variance as well as the mean should be standard practice in our analytical workflow in ecology, evolution, and environmental sciences. Let’s re-imagine heterogeneity.

References

H. Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 2003.

H. Allegue, Y. G. Araya-Ajoy, N. J. Dingemanse, N. A. Dochtermann, L. Z. Garamszegi, S. Nakagawa, D. Reale, H. Schielzeth, and D. F. Westneat. Statistical quantification of individual differences (squid): an educational and statistical tool for understanding multilevel phenotypic data in linear mixed models. *Methods in Ecology and Evolution*, 8(2):257–267, 2017.

D. Anderson and K. Burnham. Model selection and multi-model inference. *Second. NY: Springer-Verlag*, 63(2020): 10, 2004.

S. C. Anderson, T. A. Branch, A. B. Cooper, and N. K. Dulvy. Black-swan events in animal populations. *Proceedings of the National Academy of Sciences*, 114(12):3252–3257, 2017.

J. M. Baldwin. A new factor in evolution. *American Naturalist*, 30(354):441–451, 1896.

S. A. Blowes. Known unknowns and model selection in ecological evidence synthesis. *bioRxiv*, page https://doi.org/10.1101/2024.12.18.629303, 2024.

B. M. Bolker, M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in ecology & evolution*, 24 (3):127–135, 2009.

M. S. Boyce and C. Perrins. Optimizing great tit clutch size in a fluctuating environment. *Ecology*, 68(1):142–153, 1987.

M. E. Brooks, K. Kristensen, K. J. Van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modelling. *The R Journal*, 9(2):378–400, 2017. doi: 10.32614/RJ-2017-066.

M. E. Brooks, K. Kristensen, M. R. Darrigo, P. Rubim, M. Uriarte, E. Bruna, and B. M. Bolker. Statistical modeling of patterns in annual reproductive rates. *Ecology*, 100(7):e02706, 2019.

L. B. Buckley and R. B. Huey. How extreme temperatures impact organisms and the evolution of their thermal tolerance. *Integrative and comparative biology*, 56(1):98–109, 2016.

S. Burke, P. Pottier, M. Lagisz, E. L. Macartney, T. Ainsworth, S. M. Drobniak, and S. Nakagawa. The impact of rising temperatures on the prevalence of coral diseases and its predictability: A global meta-analysis. *Ecology letters*, 26(8):1466–1481, 2023.

P.-C. Bürkner. brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80: 1–28, 2017.

R. J. Carroll and D. Ruppert. *Transformation and Weighting in Regression*. Chapman and Hall, New York, 1988.

I. R. Cleasby and S. Nakagawa. Neglected biological patterns in the residuals: a behavioural ecologist's guide to co-operating with heteroscedasticity. *Behavioral Ecology and Sociobiology*, 65:2361–2372, 2011.

I. R. Cleasby, S. Nakagawa, and H. Schielzeth. Quantifying the predictability of behaviour: statistical approaches for the study of between-individual variation in the within-individual variance. *Methods in Ecology and Evolution*, 6 (1):27–37, 2015.

W. K. Cornwell and D. D. Ackerly. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal california. *Ecological monographs*, 79(1):109–126, 2009.

M. L. Corrales and E. Cepeda-Cuervo. Bayesian modeling of location, scale, and shape parameters in skew-normal regression models. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 15(1):98–111, 2022.

E. Crispo. The baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, 61(11):2469–2479, 2007. doi: 10.1111/j.1558-5646.2007.00177.x.

J. C. Douma and J. T. Weedon. Analysing continuous proportions in ecology and evolution: A practical introduction to beta and dirichlet regression. *Methods in Ecology and Evolution*, 10(9):1412–1430, 2019.

H. Drummond, C. Rodríguez, and S. Ortega. Long-term insights into who benefits from brood reduction. *Behavioral Ecology*, 36(4):araf050, 2025.

P. K. Dunn and G. K. Smyth. Randomized quantile residuals. *Journal of Computational and graphical statistics*, 5 (3):236–244, 1996.

S. Ferrari and F. Cribari-Neto. Beta regression for modelling rates and proportions. *Journal of applied statistics*, 31 (7):799–815, 2004.

A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4): 457–472, 1992.

A. Gelman, A. Vehtari, D. Simpson, C. C. Margossian, B. Carpenter, Y. Yao, L. Kennedy, J. Gabry, P.-C. Bürkner, and M. Modrák. Bayesian workflow. *arXiv preprint arXiv:2011.01808*, 2020.

B. Halliwell. Rethinking niche conservatism with phylogenetic location-scale models. *bioRxiv*, pages 2025–03, 2025.

F. Hartig. Dharma: residual diagnostics for hierarchical (multi-level/mixed) regression models. r package version 0.4. 6, 2022.

A. F. Hayes and L. Cai. Using heteroskedasticity-consistent standard error estimators in ols regression: An introduction and software implementation. *Behavior research methods*, 39:709–722, 2007.

T. Hu, P. Gallins, and Y.-H. Zhou. A zero-inflated beta-binomial model for microbiome data analysis. *Stat*, 7(1): e185, 2018.

B. Jorgensen. *The theory of dispersion models.* CRC Press, 1997.

Y. Lee and J. A. Nelder. Hierarchical generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(4):619–656, 1996.

Y. Lee and J. A. Nelder. Double hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society Series C: Applied Statistics*, 55(2):139–185, 2006.

Y. Lee, J. A. Nelder, and Y. Pawitan. *Generalized linear models with random effects: unified analysis via H-likelihood.* Chapman & Hall/CRC, 2006.

L. W. Liou, T. Price, M. S. Boyce, and C. M. Perrins. Fluctuating environments and clutch size evolution in great tits. *The American Naturalist*, 141(3):507–516, 1993.

E. J. Lundgren, D. Ramp, O. S. Middleton, E. I. Wooster, E. Kusch, M. Balisi, W. J. Ripple, C. D. Hasselerharm, J. N. Sanchez, M. Mills, et al. A novel trophic cascade between cougars and feral donkeys shapes desert wetlands. *Journal of Animal Ecology*, 91(12):2348–2357, 2022.

B. D. Martin, D. Witten, and A. D. Willis. Modeling microbial abundances and dysbiosis with beta-binomial regression. *The annals of applied statistics*, 14(1):94, 2020.

J. G. Martin, E. Pirotta, M. B. Petelle, and D. T. Blumstein. Genetic basis of between-individual and within-individual variance of docility. *Journal of Evolutionary Biology*, 30(4):796–805, 2017.

A. Mizuno and M. Soma. Pre-existing visual preference for white dot patterns in estrildid finches: a comparative study of a multi-species experiment. *Royal Society Open Science*, 10(10):231057, 2023.

A. Mizuno, S. Nakagawa, and co-authors. Data and code for ”location–scale models in ecology and evolution: heteroscedasticity in continuous, count, and proportion data”, 2025. URL `https://doi.org/10.5281/zenodo.17515632`.

M. Modrák, A. H. Moon, S. Kim, P. Bürkner, N. Huurre, K. Faltejsková, A. Gelman, and A. Vehtari. Simulation-based calibration checking for bayesian computation: The choice of test quantities shapes sensitivity. *Bayesian Analysis*, 20(2):461–488, 2025.

C. C. Monnahan, J. T. Thorson, and T. A. Branch. Faster estimation of bayesian models in ecology using hamiltonian monte carlo. *Methods in Ecology and Evolution*, 8(3):339–348, 2017.

S. Nakagawa and I. C. Cuthill. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews*, 82(4):591–605, 2007.

S. Nakagawa and H. Schielzeth. A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in ecology and evolution*, 4(2):133–142, 2013.

S. Nakagawa, A. Mizuno, K. Morrison, L. Ricolfi, C. Williams, S. M. Drobniak, M. Lagisz, and Y. Yang. Location-scale meta-analysis and meta-regression as a tool to capture large-scale changes in biological and methodological heterogeneity: A spotlight on heteroscedasticity. *Global Change Biology*, 31(5):e70204, 2025a.

S. Nakagawa, A. Mizuno, C. Williams, M. Lagisz, Y. Yang, and S. M. Drobniak. Quantifying macro-evolutionary patterns of trait mean and variance with phylogenetic location-scale models. *Methods in Ecology and Evolution*, 2025b. doi: https://doi.org/10.32942/X2XS7K.

R. E. O'Dea, D. W. Noble, and S. Nakagawa. Unifying individual differences in personality, predictability and plasticity: a practical guide. *Methods in Ecology and Evolution*, 13(2):278–293, 2022.

R. Ospina and S. L. Ferrari. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis*, 56(6):1609–1623, 2012.

R. E. O'Dea, D. W. Noble, S. L. Johnson, D. Hesselson, and S. Nakagawa. The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. *Environmental epigenetics*, 2(1):dvv014, 2016.

J. L. Pick, H. E. Lemon, C. E. Thomson, and J. D. Hadfield. Decomposing phenotypic skew and its effects on the predicted response to strong selection. *Nature Ecology & Evolution*, 6(6):774–785, 2022.

P. Pollo, S. M. Drobniak, H. Haselimashhadi, M. Lagisz, A. Mizuno, D. W. Noble, L. A. Wilson, and S. Nakagawa. Beyond sex differences in mean: meta-analysis of differences in skewness, kurtosis, and correlation. *EcoEvoRxiv*, 2025. doi: https://doi.org/10.32942/X20K9W. EcoEvoRxiv preprint.

R. A. Rigby and D. M. Stasinopoulos. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 54(3):507–554, 2005.

L. Rönnegård and Y. Lee. Exploring the potential of hierarchical generalized linear models in animal breeding and genetics. *Journal of Animal Breeding & Genetics*, 130(6), 2013.

T. Säilynoja, M. Schmitt, P.-C. Bürkner, and A. Vehtari. Posterior sbc: Simulation-based calibration checking conditional on data. *arXiv preprint arXiv:2502.03279*, 2025.

E. S. Santos and S. Nakagawa. Breeding biology and variable mating system of a population of introduced dunnocks (prunella modularis) in new zealand. *PLoS One*, 8(7):e69329, 2013.

H. Schielzeth and S. Nakagawa. Nested by design: model fitting and interpretation in a mixed model era. *Methods in Ecology and Evolution*, 4(1):14–24, 2013.

K. F. Sellers and G. Shmueli. A flexible regression model for count data. *The Annals of Applied Statistics*, pages 943–961, 2010.

A. M. Senior, C. E. Grueber, T. Kamiya, M. Lagisz, K. O'Dwyer, E. S. A. Santos, and S. Nakagawa. Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology*, 97(12):3293–3299, 2016. doi: https://doi.org/10.1002/ecy.1591. URL `https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecy.1591`.

M. Smithson and J. Verkuilen. A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*, 11(1):54, 2006.

J. Starrfelt and H. Kokko. Bet-hedging—a triple trade-off between means, variances and correlations. *Biological Reviews*, 87(3):742–755, 2012.

D. M. Stasinopoulos and R. A. Rigby. Generalized additive models for location scale and shape (gamlss) in r. *Journal of Statistical Software*, 23:1–46, 2008.

M. Stemkovski, R. G. Dickson, S. R. Griffin, B. D. Inouye, D. W. Inouye, G. L. Pardee, N. Underwood, and R. E. Irwin. Skewness in bee and flower phenological distributions, 2023.

J. Stoklosa, R. V. Blakey, and F. K. Hui. An overview of modern applications of negative binomial modelling in ecology and biodiversity. *Diversity*, 14(5):320, 2022.

R. A. Taylor, S.-J. Park, and P. S. Grewal. Nematode spatial distribution and the frequency of zeros in samples. *Nematology*, 19(3):263–270, 2017.

N. Umlauf, N. Klein, T. Simon, and A. Zeileis. bamlss: a lego toolbox for flexible bayesian regression (and beyond). *Journal of Statistical Software*, 100:1–53, 2021.

A. Vehtari, A. Gelman, and J. Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27:1413–1432, 2017.

D. I. Warton. Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics: The official journal of the International Environmetrics Society*, 16 (3):275–289, 2005.

D. F. Westneat, M. Schofield, and J. Wright. Parental behavior exhibits among-individual variance, plasticity, and heterogeneous residual variance. *Behavioral Ecology*, 24(3):598–604, 2013.

A. F. Zuur, E. N. Ieno, N. Walker, A. A. Saveliev, G. M. Smith, A. F. Zuur, E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. Zero-truncated and zero-inflated models for count data. *Mixed effects models and extensions in ecology with R*, pages 261–293, 2009a.

A. F. Zuur, E. N. Ieno, N. J. Walker, A. A. Saveliev, G. M. Smith, et al. *Mixed effects models and extensions in ecology with R*, volume 574. Springer, 2009b.

Figure captions

Figure 1. Homoscedasticity and heteroscedasticity patterns in common data distributions. Examples for (a) continuous, (b) count, and (c) proportion data. Top panels show homoscedasticity; bottom panels show heteroscedasticity. (a) Continuous responses demonstrate how continuous and categorical predictors can exhibit constant or varying variance across. (b) Count data inherently links the mean and variance, if counts follow Poisson distributions ($E[y] < Var[y]$). Thus, variance increases with expected value. The bottom panel, though visually uniform, represents the heteroscedasticity with larger dispersion at higher means. (c) Proportion data (proportion 0 - 1) shows heteroscedasticity (bottom) as inflated frequencies at the boundaries (0 - 1), reflecting overdispersion. This is often modeled by a beta-binomial distribution, where the success probability varies across observations.

Figure 2. Practical workflow for detecting and modeling heteroscedasticity with location-scale models. This diagram outlines a step-by-step guide for applying location–scale models to identify and interpret non-constant variance in continuous, count, or proportion data. The workflow progresses from initial data visualization and distribution identification (steps 1 and 2) to fitting a location-only baseline model and conducting residual diagnostics for variance patterns (steps 3 and 4). If heteroscedasticity is detected, a location-scale model is fitted (step 5) and compared against other possible models (e.g., ones with fewer or more fixed effects or random effects) using information criteria such as AIC (frequentist) or WAIC/LOO (Bayesian) (step 6). Finally, clearly report both mean and variance effects as final results. Note that the table in step 5 summarizes key variance-related parameters (e.g., σ^2, θ, ϕ, ν) and their corresponding interpretations (for more details, see the main text).