Online Appendix: Mixed-Effects Additive Transformation Models

Bálint Tamási

balint.tamasi@uzh.ch

Torsten Hothorn

Torsten.Hothorn@R-project.org

5/4/2022

Abstract

This vignette serves as an online appendix for the manuscript "Mixed-Effects Additive Transformation Models". It presents four example analyses that use mixed-effects additive transformation models to reanalyze ecological phenomena from recently published studies.

1 The rat carrion decomposition experiment

Our first example presents the reanalysis of the carrion decomposition data by Englmeier et al. (2022). In the original study, the authors analyzed the environmental factors that affect the decomposition process of small rodent carrion using data from an experiment, in which they placed rat carcasses in different environments and recorded the time until complete decomposition. As we describe it in the main article, the outcome variable (time until complete decomposition) is *interval censored* due to the discrete follow-up. The survival curves for specimen with and without the presence of insects are shown in Figure 1
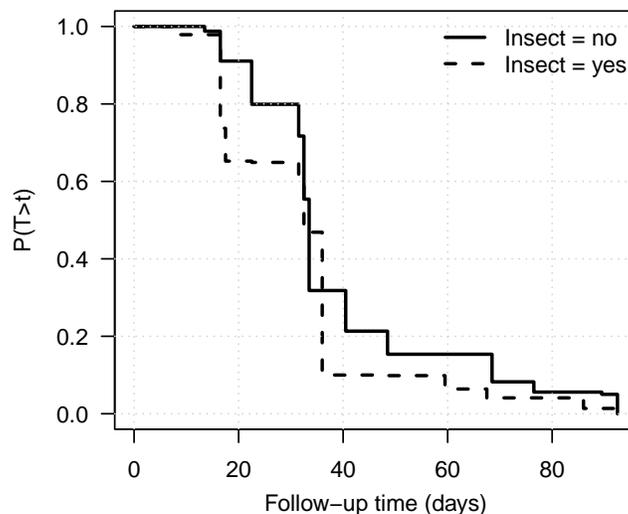

Figure 1: Non-parametric survival probability estimates associated with the decomposition times in groups where insect access was and was not allowed in the carrion decomposition experiment.

The main environmental variables we include in our model are the indicator for the presence of insects (`Insects`), local (`Habitat`) and regional (`Landscape`) land use types, The average temperature on each experimental plot was measured with thermologgers (`Temperature`), and the elevational gradient was used as a surrogate for the long-term macroclimate (`Elevation100`). The plot-level unobserved sources of variability were modeled by including random intercepts (grouping variable: `PlotID`). We estimate the following proportional-hazards mixed-effect additive transformation model for the decomposition intervals

```
R> dcmp <- CoxphME(Time ~ Insects + Habitat + Landscape
+                   + s(Temperature, k = 20) + s(Elevation100, k = 20)
+                   + (1 | PlotID), data = carrion,
+                   log_first = TRUE, order = 6)
R> summary(dcmp)


Additive Mixed-Effects Parametric Cox Regression Model


Formula: Time ~ Insects + Habitat + Landscape + s(Temperature, k = 20) +
    s(Elevation100, k = 20) + (1 | PlotID)


Fitted to dataset carrion


Fixed effects parameters:
=========================


                      Estimate Std. Error z value Pr(>|z|)
Insectsyes              1.0562     0.2493    4.24  2.3e-05 ***
Habitatarable field    -0.8986     0.4947   -1.82    0.069 .
Habitatmeadow          -0.1898     0.4665   -0.41    0.684
Habitatsettlement      -0.8492     0.5699   -1.49    0.136
Landscapeagriculture   -0.3964     0.4332   -0.92    0.360
Landscapeurban          0.0337     0.4373    0.08    0.938
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Smooth shift terms:
===================


                edf
s(Temperature)    1
s(Elevation100)   1


Random effects:
===============


Grouping factor: PlotID (144 levels)
Standard deviation:
(Intercept)
```

```
        1.73
```

```
Log-likelihood: -314 (npar = 18)
```

The smooth terms can be evaluated and plotted with

```
R> plot(smooth_terms(dcmp))
```

As the results in Figure 2 show, the effects of temperature and elevation look fairly linear.
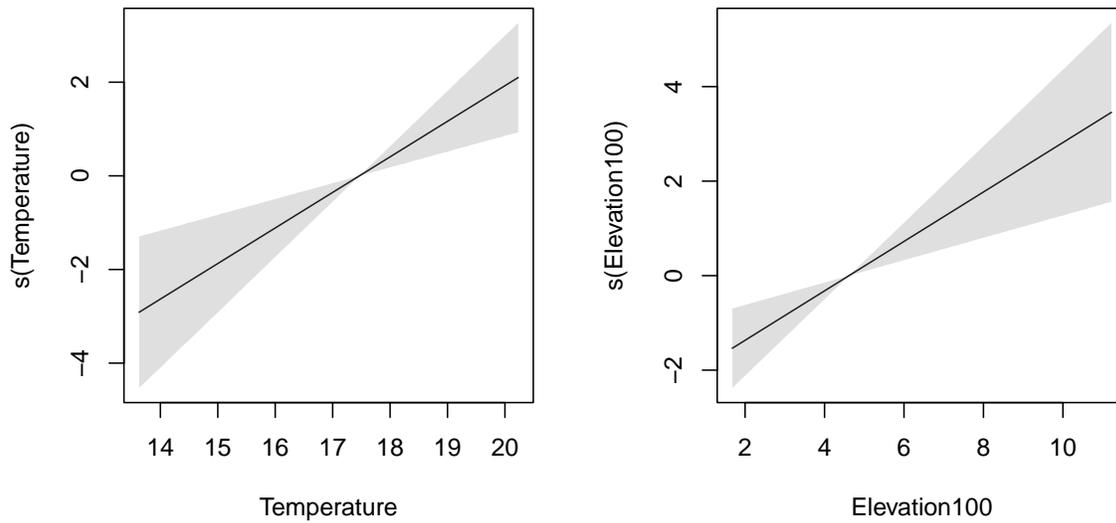

Figure 2: Smooth effects of the continuous variables in the model of the decomposition times.

As a general check of the appropriateness of the estimated model, we can evaluate the *marginal* distribution function or, as it is more commonly done in survival analysis, the marginal survivor function of the outcomes at the observations in our dataset by integrating over the random effects numerically.

$$\widehat{S}(t_i \mid \boldsymbol{x}_i) = \widehat{\mathbb{P}}(T_i > t_i \mid \boldsymbol{x}_i) = \int_{-\infty}^{+\infty} \widehat{\mathbb{P}}(T_i > t_i \mid \boldsymbol{x}_i, \gamma)\phi(\gamma)\, \mathsf{d}\gamma.$$

$\widehat{S}(t_i \mid \boldsymbol{x}_i)$ denotes the fitted survivor function, which is straightforward to calculate from the mixed-effects additive transformation model, due to its fully parametric approach to approximate the outcome distribution. By evaluating $-\log(\widehat{S}(t_i \mid \boldsymbol{x}_i))$ at the observations in our dataset, we get the Cox-Snell residuals (Klein and Moeschberger, 2003, Chapter 11), which are unit exponentially distributed under the correct model.

Interval-censored outcomes pose a technical difficulty in assessing the Cox-Snell residuals. In this situation we can either evaluate the marginal survivor function at the upper and lower bounds of the censoring intervals and assess the interval-censored version of the residuals, or we can apply the adjustment proposed by Farrington (2000), which replaces these intervals with expected values under unit exponential distribution. To assess the distribution of the residuals, we estimate the cumulative hazard

function of the Cox-Snell residuals, which should be close to a straight line with unit slope through the origin under the correct model. Figure 3 presents the distributions of the Cox-Snell residuals using the two approaches to interval censoring. Neither of these plots signals serious departures from the unit exponential distribution, which confirms the appropriateness of our regression model. Because the marginal Cox-Snell residuals are not independent in our case, these plots only provide a crude visual check of the model fits.
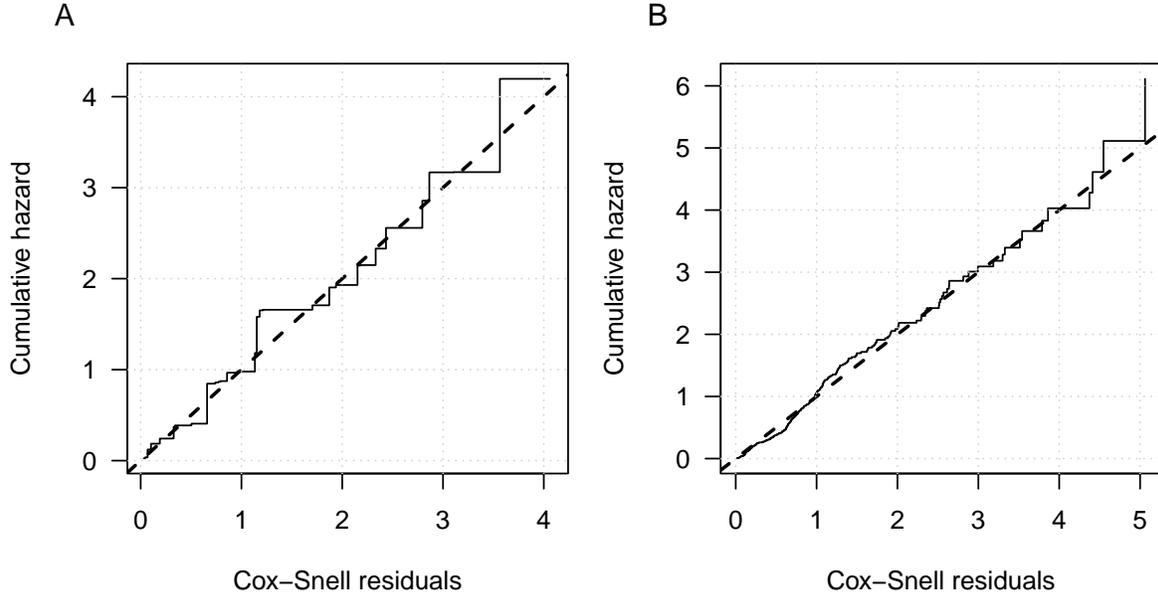
A
B

Figure 3: Cox-Snell residuals of the carrion decomposition model. *Panel A*: Treating the residuals as interval-censored and estimating the cumulative hazard function using the Turnbull non-parametric maximum likelihood estimator. *Panel B*: Using the adjustment proposed by Farrington (2000). The dashed lines correspond to the unit exponential distribution.

We can relax the assumption of proportional hazards by allowing for time-dependent covariate effects. A transformation model with time-dependent effects for the `Insects` indicator can be estimated as

```
R> dcmp2 <- CoxphME(Time | Insects ~ Habitat + Landscape
+                   + s(Temperature, k = 20) + s(Elevation100, k = 20)
+                   + (1 | PlotID), data = carrion,
+                   log_first = TRUE, order = 6)
```

In Figure 4, we compare the effect estimates of the presence of insects (on the log-hazard scale) from the proportional hazards and non-proportional hazards (time-varying effects) models. According to these results, the proportional hazards assumption seems plausible.

2   *E. coli* concentrations in streams with different grazing periods

Hulvey et al. (2021) compare the concentration levels of *Escherichia coli* bacteria (most probable number, MPN) in streams under three different rotational grazing regimes. In the additive mixed model specifications they estimated, within-year variability was modeled, as functions of the day of
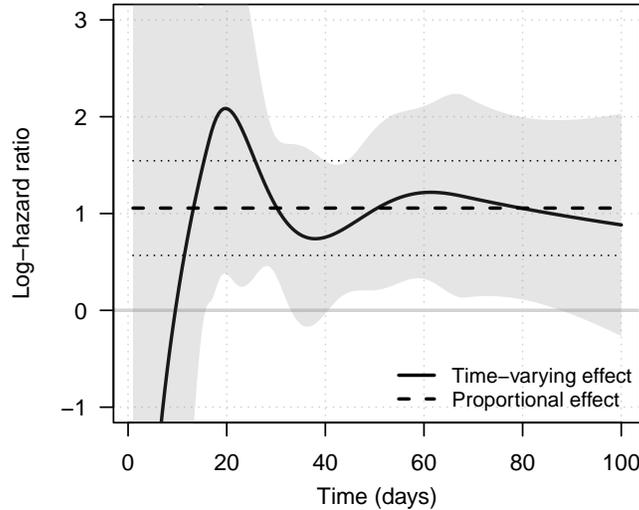
Figure 4: The effect of the presence of insects on the decomposition process from the proportional-hazards and non-proportional hazards models.

year (DOY), with cubic regression splines and between-year and location-level variability were captured by random intercepts of pasture-specific year effects and separate stream effects. Note that although the cyclic version of the cubic regression splines (`bs = 'cc'` in `mgcv` and `tramME`) (Tamasi, 2022) would be more appropriate for modeling the within-year trend, the original article used `bs = 'cr'` and hence we also stick with this basis in our reanalysis.

As a first step, we replicate the results of all model variants that Hulvey et al. (2021) investigated in the original article with the R package `gamm4` (Wood and Scheipl, 2020). In a second step, we fit the *same* models using the software implementation of additive mixed transformation models in package `tramME`, that is, using a linear transformation function (function `tramME::LmME`). We do expect identical results in steps one and two, although the two implementations rely on two completely distinct code bases. Thus, these results are only interesting from a quality assurance point of view. In the last step, we relax the normal distributional assumption by allowing a nonlinear transformation function (`tramME::BoxCoxME` function) and evaluate how the model fits change. We are primarily interested in potential changes of the model interpretation induced by a shift from a normal to a distribution-free model.

As Table 1 shows, we managed to reproduce the `gamm4` results with `tramME`. Moreover, relaxing the distributional assumption of the normal linear model resulted in stronger model fits in terms of in-sample log-likelihood values.

```
R> ## specifications w/o random effects
R> mf <- c(log10(ecoli_MPN) ~ treatment + cattle +
+              s(DOY, bs = 'cr', by = treatment),
+           log10(ecoli_MPN) ~ treatment + cattle + s(DOY, bs = 'cr'),
+           log10(ecoli_MPN) ~ treatment + s(DOY, bs = 'cr', by = treatment),
+           log10(ecoli_MPN) ~ cattle + s(DOY, bs = 'cr'),
+           log10(ecoli_MPN) ~ treatment + s(DOY, bs = 'cr'),
+           log10(ecoli_MPN) ~ s(DOY, bs = 'cr'))
R> names(mf) <- paste("Model", c(1:5, "Null"))
```

```
R> ecoli_res <- data.frame(matrix(NA, nrow = length(mf), ncol = 3))
R> colnames(ecoli_res) <- c("gamm", "LmME", "BoxCoxME")
R> rownames(ecoli_res) <- names(mf)
R> for (i in seq_along(mf)) {
+    m_gamm <- gamm4(mf[[i]], data = ecoli,
+                    random = ~ (1 | year:stream:pasture) + (1 | stream),
+                    REML = FALSE)
+    ecoli_res$gamm[i] <- logLik(m_gamm$mer)
+    mf2 <- update(mf[[i]], . ~ . + (1 | year:stream:pasture) + (1 | stream))
+    m_LmME <- LmME(mf2, data = ecoli)
+    if (m_LmME$opt$convergence == 0) ecoli_res$LmME[i] <- logLik(m_LmME)
+    m_BCME <- BoxCoxME(mf2, data = ecoli)
+    if (m_BCME$opt$convergence == 0) ecoli_res$BoxCoxME[i] <- logLik(m_BCME)
+ }
```

Table 1: Log-likelihood values of the fitted models presented by Hulvey et al. (2021, *GAMM*), replicated as mixed-effects additive transformation models assuming conditional normality (*Additive normal transformation model*) and extended as flexible (non-normal) mixed-effects additive transformation models (*Additive non-normal transformation model*).

	GAMM	Additive normal transformation model	Additive non-normal transformation model
Model 1	-339.23	-339.23	-320.94
Model 2	-343.66	-343.66	-324.54
Model 3	-368.33	-368.33	-349.10
Model 4	-347.70	-347.70	-328.25
Model 5	-367.15	-367.15	-347.27
Model Null	-373.76	-373.76	-353.50

Let us focus on the most complicated specification, Model 1,

```
R> update(mf[[1]], . ~ . + (1 | year:stream:pasture) + (1 | stream))

log10(ecoli_MPN) ~ treatment + cattle + s(DOY, bs = "cr", by = treatment) +
    (1 | year:stream:pasture) + (1 | stream)
```

and compare the effect estimates from the normal model to its non-parametric counterpart. But first, notice that by changing the transformation from $h(y) = \vartheta_0 + \vartheta_1 y$ to $h(y) = \mathbf{a}(y)^\top \boldsymbol{\vartheta}$, we change the scale on which the coefficients and the smooth terms are interpreted. In the normal additive mixed model, the coefficient of a fixed effect captures the change in the expectation of the outcome when increasing the respective predictor by one unit (keeping everything else unchanged). In the non-normal transformation model with Φ as the inverse link, the coefficients capture similar effects but on a latent scale defined by the transformation $h(Y)$.

To cast the effect estimates from the two models to a common scale, we can calculate the *probabilistic indices* (PI, Thas et al., 2012). To simplify the notation, first, we will now focus on the simple, fixed

effects-only case with a single predictor:

$$\mathbb{P}(Y \leq y \mid \boldsymbol{X} = x) = \Phi\left(h(y) - \beta x\right)$$

The PI is the probability that one outcome $(Y^\star)$ is larger than the other (Y), given the same covariate values $(\boldsymbol{X})$ except for one, which is larger with one unit $(\boldsymbol{X}^\star)$. In our simplified example, this means

$$
\begin{aligned}
\mathbb{P}\left(Y < Y^\star \mid \boldsymbol{X} = x, \boldsymbol{X}^\star = x + 1\right) &= \mathbb{P}\left(h(Y) < h(Y^\star) \mid \boldsymbol{X} = x, \boldsymbol{X}^\star = x + 1\right) \\
&= \mathbb{P}\left(\frac{h(Y) - h(Y^\star) + \beta}{\sqrt{2}} < \frac{\beta}{\sqrt{2}}\right) \\
&= \Phi\left(\frac{\beta}{\sqrt{2}}\right).
\end{aligned}
$$

The third line follows from the fact that, in a transformation model with $\Phi(\cdot)$ as the inverse link, $h(Y)$ and $h(Y^\star)$ are independent, normally distributed random variables with unit variance and a mean difference of β. Notice that the PI does not depend on the transformation function. When random effects are present in the model, the PI is conditional on the cluster.

In the case of transformation models with non-linear additive terms the probabilistic index is a function of the covariate. In the simplest form of an additive transformation model with probit link, we have

$$\mathbb{P}(Y \leq y \mid \boldsymbol{X} = x) = \Phi\left(h(y) - f(x)\right)$$

and the PI is

$$PI(x) = \mathbb{P}\left(Y < Y^\star \mid \boldsymbol{X} = x, \boldsymbol{X}^\star = x + 1\right) = \Phi\left(\frac{f(x + 1) - f(x)}{\sqrt{2}}\right).$$

By transforming the effect estimates to the probability scale, Figure 5 compares the smooth terms from the normal and non-normal versions of Model 1, while the first two blocks of Table 2 contrasts the fixed effects estimates. The results are very close to each other, which suggests that the original log-normal model is actually appropriate. As a built-in visual normality check, we can compare the fitted transformation functions of the normal and non-normal transformation models. The linear function corresponds to a conditional normal distribution in Figure 6. This result further confirms the appropriateness of the normal additive model in this specific example.
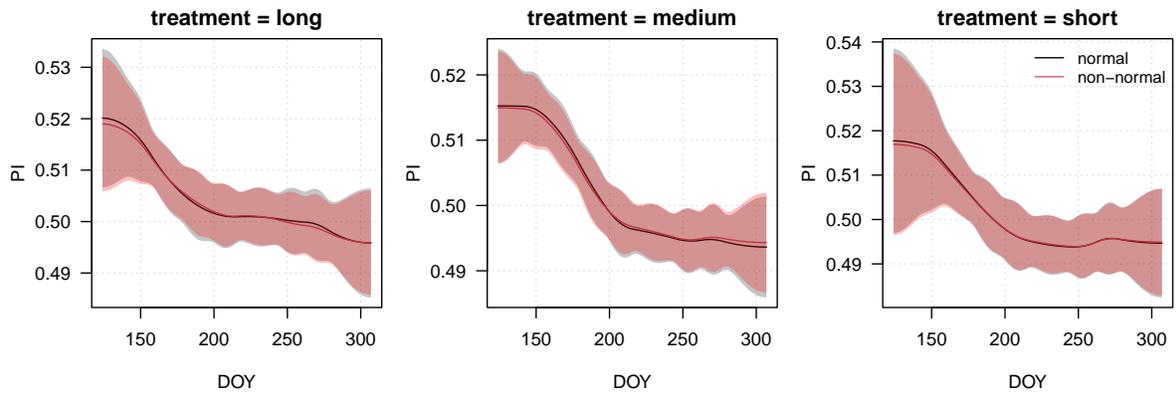

Figure 5: The comparison of the smooth terms from the normal and non-normal (probit link) mixed-effects additive transformation models (specification Model 1).

The outcome variable (MPN per 100 ml) was measured with the Quanti-Tray System, which can detect *E. coli* concentrations up to a maximum of 2,419.6 MPN without dilution. This means that
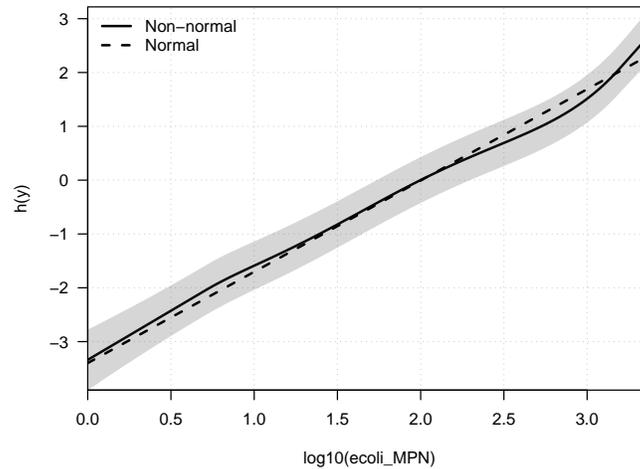
Figure 6: Baseline transformation functions from the normal and non-normal mixed-effects additive transformation models.

there is an effective upper detection limit on the outcome, i.e., the 25 observations with the value of 2,419.6 are *right censored*. The authors of the original article mention this fact, but they do not take into account in the subsequent analyses. Because random censoring can be easily handled in `tramME`, we will rerun the model taking the upper limit into account.

```
R> fm1c <- update(fm1, Surv(log10(ecoli_MPN), event = ecoli_MPN < 2419.6) ~ .)
R> ecoli_m1_cens <- BoxCoxME(fm1c, data = ecoli)
R> summary(ecoli_m1_cens)


Non-Normal (Box-Cox-Type) Linear Additive Mixed-Effects Regression Model

Formula: Surv(log10(ecoli_MPN), event = ecoli_MPN < 2419.6) ~ treatment +
    cattle + s(DOY, bs = "cr", by = treatment) + (1 | year:stream:pasture) +
    (1 | stream)

Fitted to dataset ecoli

Fixed effects parameters:
=========================


                 Estimate Std. Error z value Pr(>|z|)
treatmentmedium   -0.680      0.230    -2.95   0.0032 **
treatmentshort    -0.772      0.317    -2.44   0.0148 *
cattlePresent      1.108      0.149     7.42  1.2e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Smooth shift terms:
===================
```

8

```
                            edf
s(DOY):treatmentlong    4.38
s(DOY):treatmentmedium  4.55
s(DOY):treatmentshort   4.30


Random effects:
===============


Grouping factor: year:stream:pasture (32 levels)
Standard deviation:
(Intercept)
      0.431

Grouping factor: stream (12 levels)
Standard deviation:
(Intercept)
   0.000204



Log-likelihood: -358 (npar = 18)
```

The fitted non-linear terms are compared to the original (normal linear) estimates in Figure 7 and the fixed effects are presented in the third block of Table 2.

Table 2: Estimates of the parametric fixed-effects terms on the *probability scale* (PI: probabilistic index) from the normal, non-normal and non-normal (with censoring taken into account) models, respectively.

	Normal		Non-normal		Non-normal, censored	
	PI	95% CI	PI	95% CI	PI	95% CI
treatment = medium	0.32	0.22—0.44	0.33	0.22—0.45	0.32	0.21—0.44
treatment = short	0.29	0.16—0.45	0.30	0.17—0.46	0.29	0.16—0.46
cattle = present	0.79	0.72—0.84	0.78	0.72—0.84	0.78	0.72—0.84

Because the transformation model approximates the conditional distribution of the outcome, in theory, we do not even have to take the base 10 logarithm of the *E. coli* most probable numbers (MPN) on the left-hand side of the model formula. `tramME` should be able to approximate the *most likely transformation*.

```
R> f_nontr <- update(fm1, Surv(ecoli_MPN, event = ecoli_MPN < 2419.6) ~ .)
R> ecoli_nontr <- BoxCoxME(f_nontr, data = ecoli, log_first = TRUE)
R> summary(ecoli_nontr)


Non-Normal (Box-Cox-Type) Linear Additive Mixed-Effects Regression Model


Formula: Surv(ecoli_MPN, event = ecoli_MPN < 2419.6) ~ treatment + cattle +
```
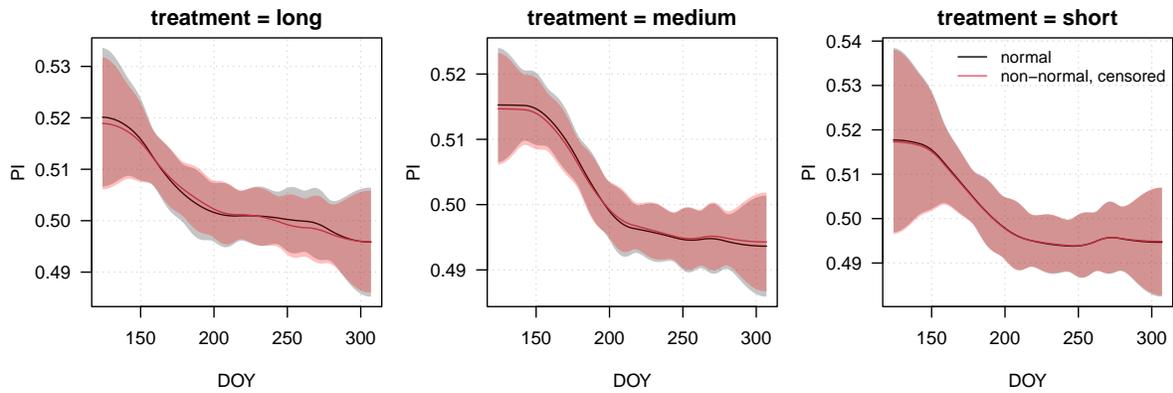
Figure 7: The comparison of the smooth terms from the original model (normal linear) and the non-normal (probit link) extension where censoring is also taken into account.

```
    s(DOY, bs = "cr", by = treatment) + (1 | year:stream:pasture) +
    (1 | stream)


Fitted to dataset ecoli


Fixed effects parameters:
=========================


              Estimate Std. Error z value Pr(>|z|)
treatmentmedium   -0.680      0.230   -2.95   0.0032 **
treatmentshort    -0.772      0.317   -2.44   0.0148 *
cattlePresent      1.108      0.149    7.42  1.2e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Smooth shift terms:
===================


                         edf
s(DOY):treatmentlong    4.38
s(DOY):treatmentmedium  4.55
s(DOY):treatmentshort   4.30


Random effects:
===============


Grouping factor: year:stream:pasture (32 levels)
Standard deviation:
(Intercept)
      0.431


Grouping factor: stream (12 levels)
```

```
Standard deviation:
(Intercept)
   0.000394



Log-likelihood: -2027 (npar = 18)
```

Notice that we set `log_first = TRUE` in the function call, to take the natural logarithm of the outcome before setting up the Bernstein bases. This usually helps the approximation in the case of positive right-skewed outcomes. With this, we basically estimate the same model as the original, but with the natural logarithm instead of base-ten. Because of this difference, the log-likelihood values are also different, but the fixed effects and variance components parameter estimates, as well as the smooth terms are essentially the same as in the case of the model `ecoli_m1_cens`.

In summary, after bringing the estimates to the same scale, the results of the additive mixed effects model did not change much in this specific example by switching to the transformation model approach. The originally applied base 10 logarithm falls very close to the fitted "most likely transformation", i.e., taking the logarithm of the outcome was sufficient to achieve (close) conditional normality. This could be verified through comparing the baseline transformation functions of the normal and non-normal models, which can also serve as a visual check on conditional normality. Moreover, the number of censored outcomes was relatively small in the sample, so taking the censoring properly into account did not result in large differences, either. However, as the example demonstrated, transformation models are flexible enough to accommodate these properties of the response of interest (non-normality and censoring) automatically, without the need to apply ad hoc transformations or to implement new estimation procedures. In this sense, `tramME::BoxCoxME` provides a simple way of checking the impact of the more restrictive assumptions hard-wired in `gamm4::gamm4` on model interpretation and of handling censoring properly in the estimation procedure.

3   Sea urchin removal experiment

Andrew and Underwood (1993) analyzed the percentage cover of filamentous algae under four sea urchin removal treatments (Control/33%/66%/Removal). The algae colonization was measured on five quadrants located on several larger patches, so there is a clear grouped structure in the data. Douma and Weedon (2019) reanalyzed the data as a demonstration for the usage of mixed-effects models for zero-inflated beta regression models. Here we fit mixed-effects transformation models to the data, and compare the results to zero-inflated mixed-model estimates obtained from `glmmTMB` (Brooks et al., 2017, Magnusson et al., 2021). Figure 8 presents the empirical cumulative distribution functions of the outcome under the four treatments. Note the large number of zeros, especially in the control group.

First, we fit a zero-inflated beta regression model with random intercepts for the patches. The probability of observing zero values is allowed to vary with the treatment.

```
R> urchin_zib <- glmmTMB(pALGAE ~ TREAT + (1 | PATCH), ziformula = ~ TREAT,
+                        data = andrew, family = beta_family())
R> summary(urchin_zib)

 Family: beta  ( logit )
Formula:          pALGAE ~ TREAT + (1 | PATCH)
```
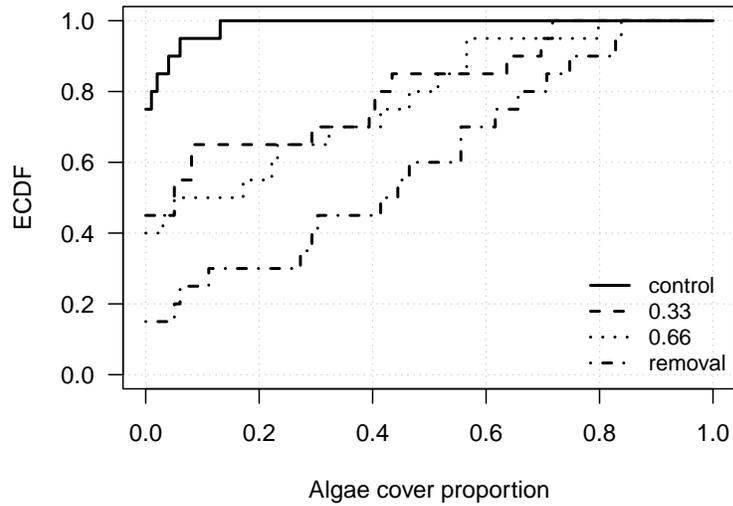
11

Figure 8: Empirical CDFs of the algae cover proportions under the four treatments.

```
Zero inflation:            ~TREAT
Data: andrew


     AIC      BIC   logLik deviance df.resid
    87.2    111.0    -33.6     67.2        70


Random effects:


Conditional model:
 Groups Name        Variance Std.Dev.
 PATCH  (Intercept) 0.124    0.352
Number of obs: 80, groups:  PATCH, 16


Dispersion parameter for beta family (): 4.06


Conditional model:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.060       0.530   -3.89   0.0001 ***
TREAT0.33      1.280       0.614    2.08   0.0372 *
TREAT0.66      1.374       0.602    2.28   0.0223 *
TREATremoval   1.783       0.585    3.05   0.0023 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Zero-inflation model:
            Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept)      1.099       0.516     2.13  0.03338 *
TREAT0.33       -1.299       0.685    -1.90  0.05772 .
TREAT0.66       -1.504       0.689    -2.18  0.02908 *
TREATremoval    -2.833       0.812    -3.49  0.00048 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As an alternative to the traditional beta regression approach, we estimate a mixed-effects continuous outcome logistic regression.

```
R> urchin_tram <- ColrME(
+     Surv(pALGAE, pALGAE > 0, type = "left") ~ TREAT + (1 | PATCH),
+     bounds = c(-0.1, 1), support = c(-0.1, 1), data = andrew,
+     order = 6)
R> summary(urchin_tram)


Mixed-Effects Continuous Outcome Logistic Regression Model

Formula: Surv(pALGAE, pALGAE > 0, type = "left") ~ TREAT + (1 | PATCH)

Fitted to dataset andrew

Fixed effects parameters:
==========================


             Estimate Std. Error z value Pr(>|z|)
TREAT0.33       -2.04       1.31    -1.56   0.1178
TREAT0.66       -2.49       1.31    -1.90   0.0571 .
TREATremoval    -4.10       1.34    -3.06   0.0022 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random effects:
===============


Grouping factor: PATCH (16 levels)
Standard deviation:
(Intercept)
      1.48



Log-likelihood: -26.3 (npar = 11)
```

To allow for a jump in the conditional CDF of the outcome, we expand its bound and treat the zero observations as left-censored. This way, we can place a point mass on zero, i.e., introduce a jump at 0 (see Figure 9).

Because the zero-inflated beta model is a mixture of two models, the interpretation of its results
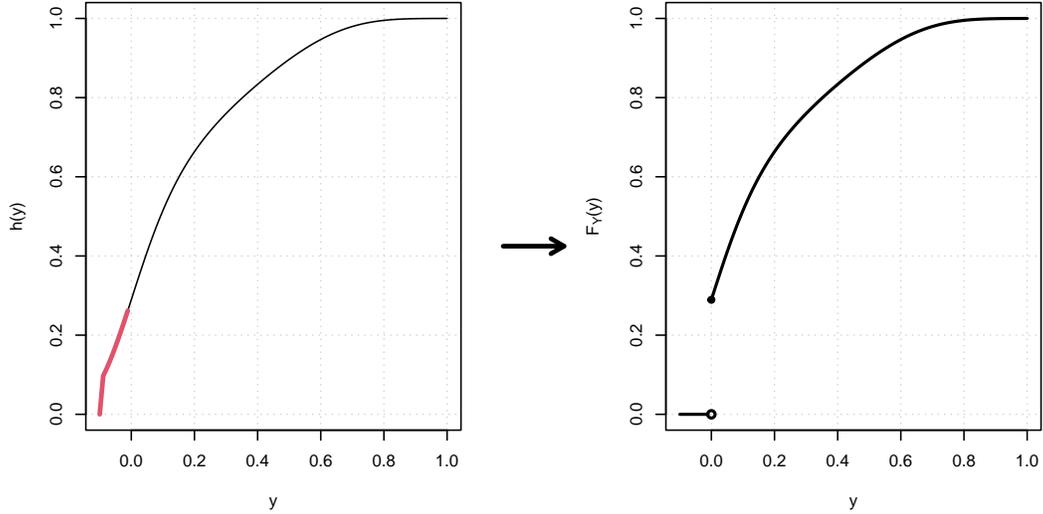
Figure 9: Visual demonstration of how a discrete jump is introduced in the CDF by extending the support and treating the edge cases as censored.

is cumbersome. It is not clear which parameters, or combinations of parameters, one needs to inspect to contrast the effects of the various treatments. Moreover, extra steps are needed to calculate the marginal effects of the covariates. In contrast, the mixed-effects transformation model only contains a single set of fixed effects parameters and their interpretation is straightforward: For example, the odds of observing higher proportions of algae cover under the 33% removal treatment is about $\exp(-\widehat{\beta}_{0.33}) = 7.71$ times higher compared to the control group.

To assess the fits of the two models we can marginalize the conditional distributions by integrating over the random effects numerically, and compare against the ECDFs. As Figure 10 shows, both model overestimate the dispersion in the control group.
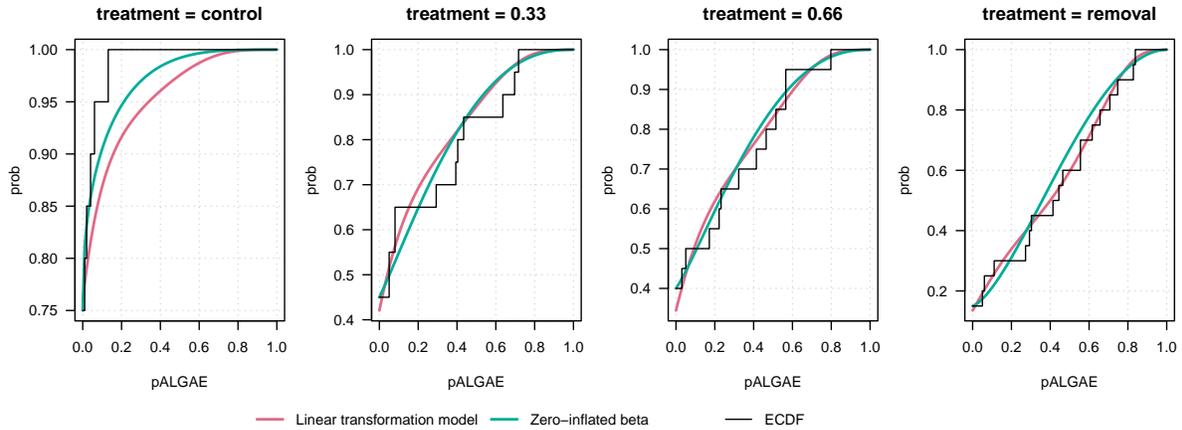

Figure 10: Fitted marginal distributions of algae cover proportion from the zero-inflated beta regression and the mixed-effects transformation model, respectively. The step functions show the empirical cumulative distribution functions in the four treatment groups.

Systematic differences in the outcome variability in the treatment groups occur in many situations (Douma and Weedon, 2019). By modeling the dispersion separately, we can incorporate such differences in the beta regression model.

```
R> urchin_zib_disp <- glmmTMB(pALGAE ~ TREAT + (1 | PATCH),
+                            ziformula = ~ TREAT, dispformula = ~ TREAT,
+                            data = andrew, family = beta_family())
R> summary(urchin_zib_disp)

 Family: beta  ( logit )
Formula:          pALGAE ~ TREAT + (1 | PATCH)
Zero inflation:          ~TREAT
Dispersion:              ~TREAT
Data: andrew


     AIC      BIC   logLik deviance df.resid
    87.9    118.8    -30.9     61.9       67


Random effects:

Conditional model:
 Groups Name        Variance Std.Dev.
 PATCH  (Intercept) 0.198    0.445
Number of obs: 80, groups:  PATCH, 16


Conditional model:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.908      0.420   -6.92  4.5e-12 ***
TREAT0.33      2.158      0.587    3.68  0.00023 ***
TREAT0.66      2.213      0.559    3.96  7.6e-05 ***
TREATremoval   2.595      0.523    4.96  7.0e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Zero-inflation model:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)    1.099      0.516    2.13  0.03338 *
TREAT0.33     -1.299      0.685   -1.90  0.05772 .
TREAT0.66     -1.504      0.689   -2.18  0.02908 *
TREATremoval  -2.833      0.812   -3.49  0.00048 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Dispersion model:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)    3.612      0.849    4.26  2.1e-05 ***
TREAT0.33     -2.424      0.925   -2.62   0.0087 **
TREAT0.66     -2.279      0.921   -2.47   0.0134 *
TREATremoval  -2.036      0.870   -2.34   0.0193 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the mixed-effects linear transformation model, we stratify to the treatment group to allow for separate transformation functions.

```
R> urchin_tram_strat <- ColrME(
+     Surv(pALGAE, pALGAE > 0, type = "left") | 0 + TREAT ~ 1 + (1 | PATCH),
+     bounds = c(-0.1, 1), support = c(-0.1, 1), data = andrew,
+     order = 6, control = optim_control(iter.max = 1e3, eval.max = 1e3,
+                                        rel.tol = 1e-9))
R> summary(urchin_tram_strat)


Stratified Mixed-Effects Continuous Outcome Logistic Regression Model

Formula: Surv(pALGAE, pALGAE > 0, type = "left") | 0 + TREAT ~ 1 + (1 |
    PATCH)

Fitted to dataset andrew

Fixed effects parameters:
==========================

No estimated shift coefficients.

Random effects:
===============

Grouping factor: PATCH (16 levels)
Standard deviation:
(Intercept)
       1.51


Log-likelihood: -22.9 (npar = 29)
```

As Figure 11 illustrates, the two models fit the data much better. However, the cost of this flexibility is that we cannot reduce the group comparisons to inference on a small set of parameters anymore.

Figures 10 and 11 demonstrate the flexibility of the distribution-free approach of transformation models compared to the parametric alternative. This is also reflected in the log-likelihood values (Table 3).

In summary, although the shift-scale beta regression model is not a special case of a transformation model and one thus cannot expect identical results with a specific parameterisation of `tramME::ColrME`, the simpler transformation model (with one instead of two linear predictors) produced a better model fit (when comparing the in-sample log-likelihoods).

4   Mosquito control trial

Juarez et al. (2021) presented the results of a cluster randomized crossover trial that assessed the efficacy of Autocidal Gravid Ovitrap (AGO) as a tool for against the mosquito species *Aedes aegypti*.
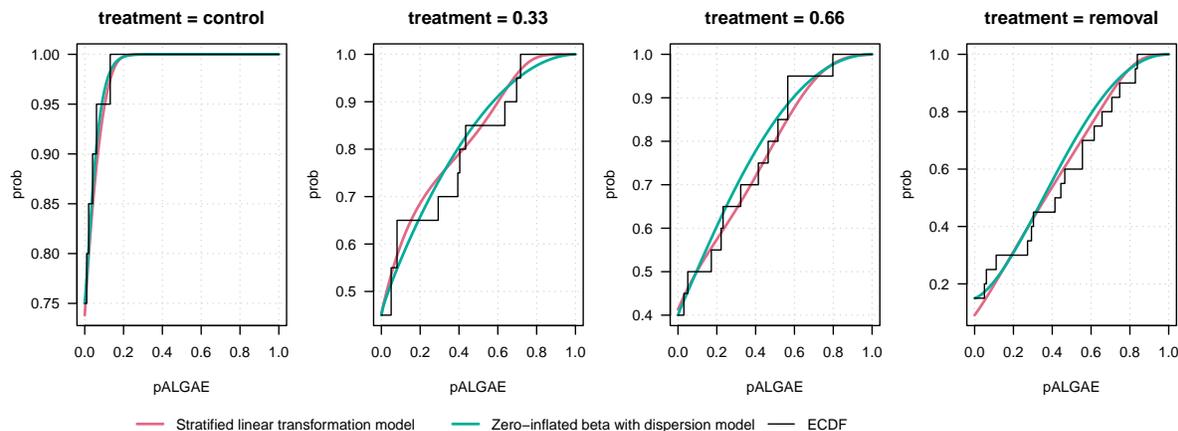
Figure 11: Fitted marginal distributions of algae cover proportion from the zero-inflated beta regression with dispersion model and the stratified mixed-effects transformation model, respectively. The step functions show the empirical cumulative distribution functions in the four treatment groups.

Table 3: Log-likelihood values of the four model specifications for the sea urchin removal experiment.

	$\log \mathcal{L}$
Zero-inflated beta w/o dispersion model	-33.60
Linear transformation model	-26.27
Zero-inflated beta w/ dispersion model	-30.93
Stratified linear transformation model	-22.86

The outcome of interest was the number of female mosquitoes collected on glue boards that were placed either inside or outside of the selected houses in various neighborhoods. Within-year patterns in mosquito counts as well as the coverage of the treatment in different areas were modeled with non-linear smooths, while unobserved household and community level effects were captured by nested random effects. The original article presented the results of a conditional Poisson and a negative binomial model. We reproduce these results with `gamm4`, and also estimate a mixed-effects additive transformation model for count data with "expit" inverse link function. Detailed exposition of count transformation models is given by Siegfried and Hothorn (2020). For fitting such a model, we will use the following custom-made 'CotramME' model class implementing the likelihood for count data via interval censoring (Siegfried and Hothorn, 2020), which is currently not part of the `tramME` package.

```r
R> ## additive count transformation model
R> CotramME <- function(formula, data,
+                       method = c("logit", "cloglog", "loglog", "probit"),
+                       log_first = TRUE, plus_one = log_first, prob = 0.9,
+                       ...) {
+     method <- match.arg(method)
+     rv <- all.vars(formula)[1]
+     stopifnot(is.integer(data[[rv]]), all(data[[rv]] >= 0))
```

17

```
+      data[[rv]] <- data[[rv]] + as.integer(plus_one)
+      sup <- c(-0.5 + log_first, quantile(data[[rv]], prob = prob))
+      bou <- c(-0.9 + log_first, Inf)
+      data[[rv]] <- as.Surv(R(data[[rv]], bounds = bou))
+      fc <- match.call()
+      fc[[1L]] <- switch(method, logit = quote(ColrME), cloglog = quote(CoxphME),
+                      loglog = quote(LehmannME), probit = quote(BoxCoxME))
+      fc$method <- NULL
+      fc$plus_one <- NULL
+      fc$prob <- NULL
+      fc$log_first <- log_first
+      fc$bounds <- bou
+      fc$support <- sup
+      fc$data <- data
+      out <- eval(fc, parent.frame())
+      out$call$data <- match.call()$data
+      class(out) <- c("CotramME", class(out))
+      out
+  }
R> mosquito_tram <- CotramME(AEAfemale ~ Year + Income*Placement
+      + s(Week) + s(CovRate200) + (1|HouseID)
+      + (1|Community), offset = -log(daystrapping), data = AGO,
+      method = "logit", order = 5, log_first = TRUE, prob = 0.9)
```

Table 4 compares the log-likelihood values of the three model versions. In terms of in-sample model fit, as measured by the log-likelihood value, both the negative binomial and the transformation model perform much better than the Poisson GAMM. The results suggest slight improvement in the model fit when we relax the conditional distribution assumption of the negative binomial GAMM and follow the distribution-free transformation model approach.

Table 4: Log-likelihood values of the fitted Poisson and negative binomial GAMMs reproduced from Juarez et al. (2021) along with the log-likelihood of an additive transformation model for count data.

	Log-likelihood
Poisson GAMM	-6875.73
Negative binomial GAMM	-4883.26
Additive count transformation model	-4873.07

We will now concentrate on comparing the estimates from the negative binomial and the count transformation models. Note that the scales on which the parameters are interpreted are different in the two models: While in the negative binomial model, the parametric and smooth terms affect the log of the conditional mean of the outcome, in the transformation model with "logit" link (i.e., "expit" inverse link), they are interpreted on the log-odds scale. Unlike in the example application of Section 2, we cannot easily transform the negative binomial parameters to the probability scale. Although the magnitudes of the effect estimates of the two models are not directly comparable, their directions, significance and the general shapes of the smooths can be contrasted.

Figure 12 compares the smooth estimates of the GAMM from `gamm4` and the transformation model from `tramME`. Although the within-year time patterns (`s(Week)`) from the two models are almost identical (on different scales), the difference of the smooth estimates of the coverage rate (`s(CovRate200)`) is marked. The general shapes of the smooths are similar, but the negative binomial GAMM penalizes it more, which is also reflected in the EDFs: 2.96 and 17.49 for the negative binomial and count transformation models, respectively.
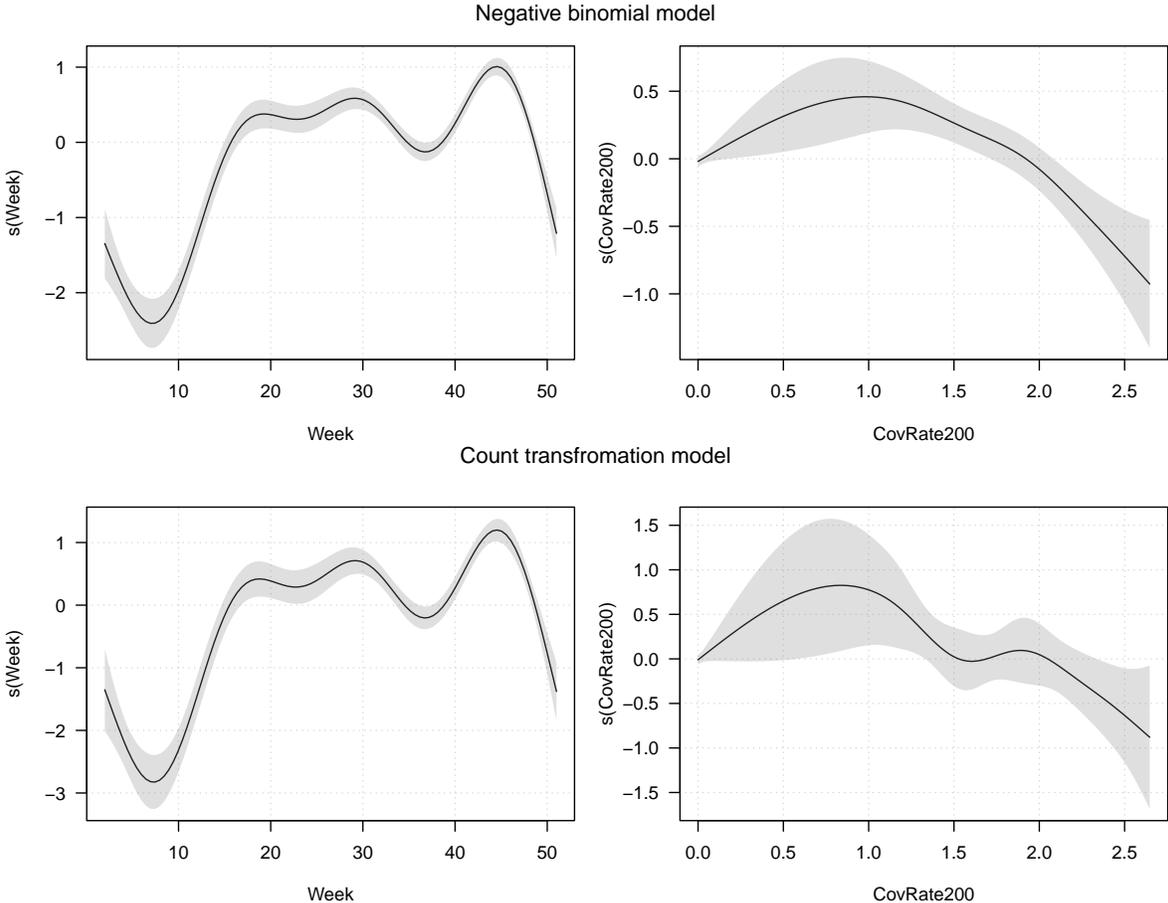

Figure 12: Smooth terms from the negative binomial and transformation models of the *A. aegypti* counts. The dashed lines and the grey areas denote the 95% confidence intervals

Because the parametric and smooth terms of the two models are defined on different scales, the magnitudes of the effect estimates are not directly comparable. As Table 5 shows, the directions of the effects match and neither model finds evidence that the main effect of middle income is different from zero.

Again, the models compared for this example are not nested and it is therefore hard to compare them directly. The transformation model leads to a similar model interpretation as the model based on the negative-binomial distribution. Model uncertainty was larger in the transformation model, at least for the nonlinear effect of `CovRate200`, and thus one might wonder if the stricter distributional assumption lead to overconfident model interpretation.

Table 5: Point estimates and 95% confidence intervals of the parametric fixed effects terms from the negative binomial and count transformation models of the mosquito control data by Juarez et al. (2021). Note that the scale of the parameters are different and the effect sizes are not directly comparable.

	Negative binomial		Count transformation	
	$\widehat{\beta}$	95% CI	$\widehat{\beta}$	95% CI
Year = 2018	-0.20	-0.34 —-0.06	-0.35	-0.55 —-0.15
Income = middle	-0.78	-1.69 — 0.13	-0.83	-2.02 — 0.36
Placement = out	2.37	2.22 — 2.52	3.01	2.79 — 3.24
Income = middle & Placement = out	0.38	0.13 — 0.64	0.51	0.16 — 0.86

References

Neil L. Andrew and Antony J. Underwood. Density-dependent foraging in the sea urchin Centrostephanus rodgersii on shallow subtidal reefs in New South Wales, Australia. *Marine Ecology Progress Series*, 99:89–98, 1993. doi:10.3354/meps099089.

Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Mächler, and Benjamin M. Bolker. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2): 378–400, 2017. doi:10.32614/RJ-2017-066.

Jacob C. Douma and James 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. doi:10.1111/2041-210X.13234.

Jana Englmeier, Oliver Mitesser, M. Eric Benbow, Torsten Hothorn, Christian von Hoermann, Caryl Benjamin, Ute Fricke, Cristina Ganuza, Maria Haensel, Sarah Redlich, Rebekka Riebl, Sandra Rojas Botero, Thomas Rummler, Ingolf Steffan-Dewenter, Elisa Stengel, Cynthia Tobisch, Johannes Uhler, Lars Uphus, Jie Zhang, and Jörg Müller. Diverse effects of climate, land use, and insects on dung and carrion decomposition. *Ecosystems*, 2022. Accepted: 03.04.2022.

C. Paddy Farrington. Residuals for proportional hazards models with interval-censored survival data. *Biometrics*, 56(2):473–482, 2000. doi:10.1111/j.0006-341X.2000.00473.x.

Kristin B. Hulvey, Cassie D. Mellon, and Andrew R. Kleinhesselink. Rotational grazing can mitigate ecosystem service trade-offs between livestock production and water quality in semi-arid rangelands. *Journal of Applied Ecology*, 58(10):2113–2123, 2021. doi:10.1111/1365-2664.13954.

Jose G. Juarez, Luis F. Chaves, Selene M. Garcia-Luna, Estelle Martin, Ismael Badillo-Vargas, Matthew C. I. Medeiros, and Gabriel L. Hamer. Variable coverage in an Autocidal Gravid Ovitrap intervention impacts efficacy of Aedes aegypti control. *Journal of Applied Ecology*, 58(10): 2075–2086, 2021. doi:10.1111/1365-2664.13951.

John P. Klein and Melvin L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Statistics for biology and health. Springer, New York, 2nd edition, 2003.

Arni Magnusson, Hans Skaug, Anders Nielsen, Casper Berg, Kasper Kristensen, Martin Maechler, Koen van Bentham, Ben Bolker, and Mollie Brooks. *glmmTMB: Generalized Linear Mixed Models using Template Model Builder*, 2021. URL https://CRAN.R-project.org/package=glmmTMB. R package version 1.1.2.3.

Sandra Siegfried and Torsten Hothorn. Count transformation models. *Methods in Ecology and Evolution*, 11(7):818–827, 2020. doi:10.1111/2041-210X.13383.

Balint Tamasi. *tramME: Transformation Models with Mixed Effects*, 2022. URL http://ctm.R-forge.R-project.org. R package version 1.0.1.

Olivier Thas, Jan De Neve, Lieven Clement, and Jean-Pierre Ottoy. Probabilistic index models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(4):623–671, 2012. doi:10.1111/j.1467-9868.2011.01020.x.

Simon Wood and Fabian Scheipl. *gamm4: Generalized Additive Mixed Models using 'mgcv' and 'lme4'*, 2020. URL https://CRAN.R-project.org/package=gamm4. R package version 0.2-6.

```
R> sessionInfo()

R version 4.1.3 (2022-03-10)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 20.04.4 LTS

Matrix products: default
BLAS:   /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_GB.UTF-8        LC_COLLATE=C
 [5] LC_MONETARY=en_GB.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_GB.UTF-8       LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_GB.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] gamm4_0.2-6    lme4_1.1-28    Matrix_1.4-0    xtable_1.8-4
 [5] glmmTMB_1.1.2.3 mgcv_1.8-34    nlme_3.1-152    survival_3.2-13
 [9] tramME_1.0.1   tram_0.6-4     mlt_1.4-0       basefun_1.1-2
[13] variables_1.1-1

loaded via a namespace (and not attached):
 [1] Rcpp_1.0.6          highr_0.8          nloptr_1.2.2.2
 [4] TMB_1.8.0           compiler_4.1.3     tools_4.1.3
 [7] boot_1.3-27         evaluate_0.14      lattice_0.20-45
[10] polynom_1.4-0       mvtnorm_1.1-1      xfun_0.23
[13] coda_0.19-4         stringr_1.4.0      BB_2019.10-1
[16] knitr_1.36          grid_4.1.3         orthopolynom_1.0-5
[19] multcomp_1.4-17     minqa_1.2.4        TH.data_1.1-0
[22] alabama_2015.3-1    Formula_1.2-4      magrittr_2.0.1
[25] emmeans_1.7.2       codetools_0.2-18   splines_4.1.3
[28] MASS_7.3-54         numDeriv_2016.8-1.1 quadprog_1.5-8
[31] sandwich_3.0-1      estimability_1.3   stringi_1.5.3
[34] coneproj_1.14       zoo_1.8-9
```