Getting started with the `glmmADMB` package

Ben Bolker, Hans Skaug, Arni Magnusson, Anders Nielsen

November 25, 2011

1 Introduction/quick start

`glmmADMB` is a package, built on the open source AD Model Builder nonlinear
fitting engine, for fitting generalized linear mixed models and extensions.

- response distributions: Poisson, binomial, negative binomial (NB1 and
  NB2 parameterizations), Gamma, Beta; Gaussian coming soon

- link functions: log, logit, probit, cloglog, inverse, identity

- zero-inflation (models with a constant zero-inflation value only)

- single, nested, or crossed random effects

- offsets

- post-fit MCMC chain for characterizing uncertainty

As of version 0.6.5, the package has been greatly revised to allow a wider
range of response and link functions and to allow models with multiple random
effects. For now, the resulting package is slower than the old (single-random-
effect version), but we hope to increase its speed in the future.

In order to use `glmmADMB` effectively you should already be reasonably fa-
miliar with generalized linear mixed models (GLMMs), which in turn requires
familiarity with (i) generalized linear models (e.g. the special cases of logistic,
binomial, and Poisson regression) and (ii) 'modern' mixed models (those work-
ing via maximization of the marginal likelihood rather than by manipulating
sums of squares).

In order to fit a model in `glmmADMB` you need to:

- specify a model for the fixed effects, in the standard R (Wilkinson-Rogers)
  formula notation (see `?formula` or Section 11.1 of the Introduction to R.
  Formulae can also include *offsets*.

- specify a model for the random effects, in the notation that is common to
  the `nlme` and `lme4` packages. Random effects are specified as `e|g`, where
  `e` is an effect and `g` is a grouping factor (which must be a factor variable,

1

or a nesting of/interaction among factor variables). For example, the formula would be `1|block` for a random-intercept model or `time|block` for a model with random variation in slopes through time across groups specified by `block`. A model of nested random effects (block within site) would be `1|site/block`; a model of crossed random effects (block and year) would be `(1|block)+(1|year)`.

Random effects can be specified either in a separate `random` argument (as in `nlme`) or as part of the model formula (as in `lme4`).

- choose the error distribution by specifying the family (as a string: e.g. `"poisson"` or `"binomial"`)

- specify a link function (as a string: e.g. `"logit"` or `"log"`.

- optionally specify that zero-inflation is present `zeroInflation=TRUE`. In the current version, zero-inflation can only be specified as a single constant term across the entire model — i.e. it cannot vary across groups or with covariates.

2 Owls data

These data, taken from [2] and ultimately from [1], quantify the number of negotiations among owlets (owl chicks) in different nests *prior* to the arrival of a provisioning parent as a function of food treatment (deprived or satiated), the sex of the parent, and arrival time. The total number of calls from the nest is recorded, along with the total brood size, which is used as an offset to allow the use of a Poisson response.

Since the same nests are measured repeatedly, the nest is used as a random effect. The model can be expressed as a zero-inflated generalized linear mixed model (ZIGLMM).

First we draw some pictures (Figures 1, 2).

Load the `glmmADMB` package to get access to the `Owls` data set; load the `ggplot2` graphics package.

```
> library(glmmADMB)
> library(ggplot2)
```

Various small manipulations of the data set: (1) reorder nests by mean negotiations per chick, for plotting purposes; (2) add log brood size variable (for offset); (3) rename response variable.

```
> Owls <- transform(Owls,
  Nest=reorder(Nest,NegPerChick),
  logBroodSize=log(BroodSize),
  NCalls=SiblingNegotiation)
```

Now fit some models:
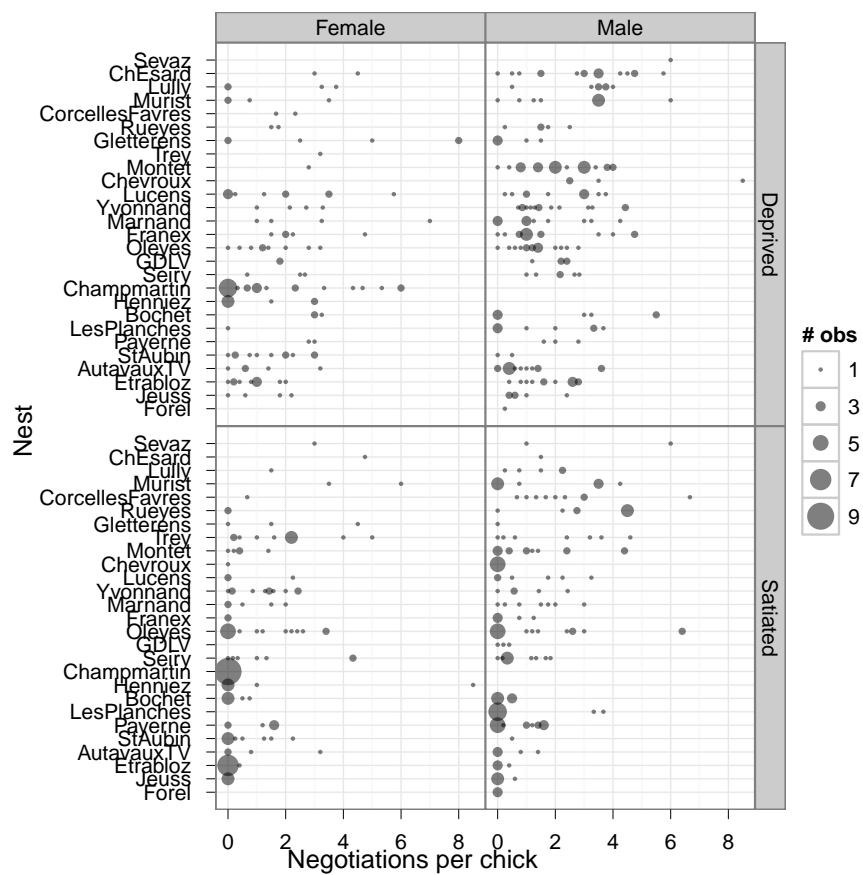The basic `glmmadmb` fit — a zero-inflated Poisson model.

Figure 1: Basic view of owl data (arrival time not shown).
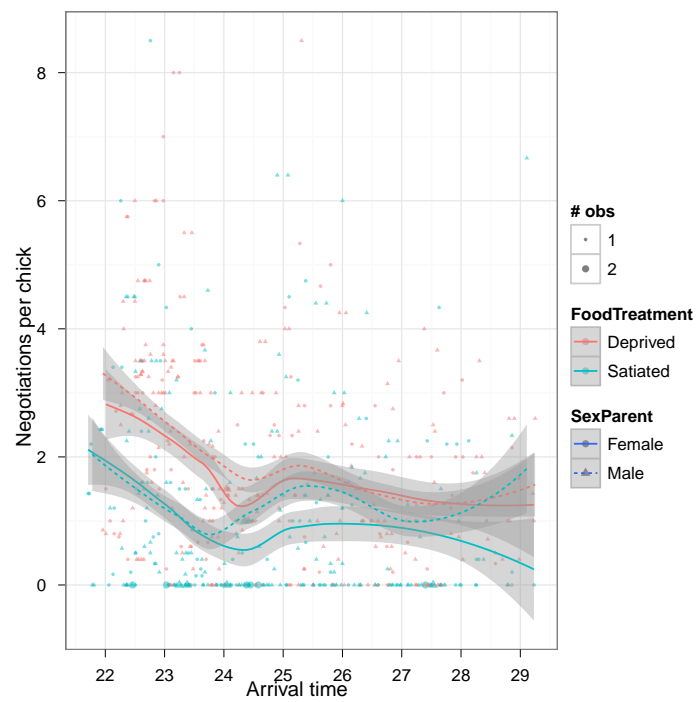
Figure 2: Basic view of owl data, #2 (nest identity not shown)

```
> fit_zipoiss <- glmmadmb(NCalls~(FoodTreatment+ArrivalTime)*SexParent+
                              offset(logBroodSize)+(1|Nest),
                              data=Owls,
                              zeroInflation=TRUE,
                              family="poisson")

> summary(fit_zipoiss)

Call:
glmmadmb(formula = NCalls ~ (FoodTreatment + ArrivalTime) * SexParent +
    offset(logBroodSize) + (1 | Nest), data = Owls, family = "poisson",
    zeroInflation = TRUE)


Coefficients:
                                     Estimate Std. Error z value Pr(>|z|)
(Intercept)                            2.8562     0.3871    7.38  1.6e-13 ***
FoodTreatmentSatiated                 -0.3314     0.0635   -5.22  1.8e-07 ***
ArrivalTime                           -0.0807     0.0156   -5.18  2.3e-07 ***
SexParentMale                          0.2882     0.3575    0.81     0.42
FoodTreatmentSatiated:SexParentMale    0.0740     0.0761    0.97     0.33
ArrivalTime:SexParentMale             -0.0150     0.0143   -1.05     0.29
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Number of observations: total=599, Nest=27
Random effect variance(s):
$Nest
            (Intercept)
(Intercept)     0.14001


Zero-inflation: 0.25833  (std. err.:  0.018107 )


Log-likelihood: -1985.3
```
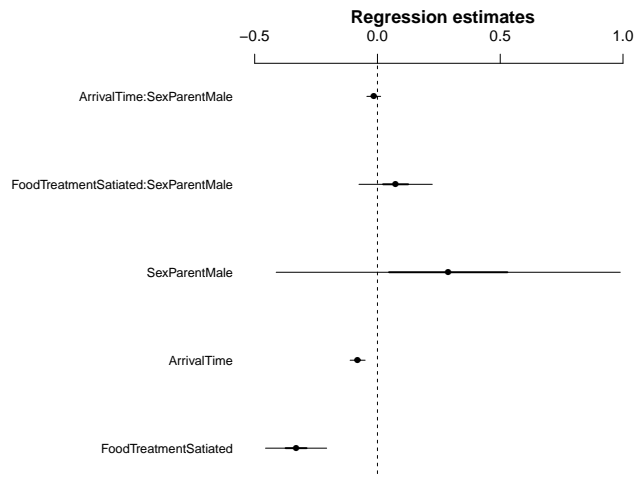
The `coefplot2` package knows about `glmmadmb` fits:

```
> library(coefplot2)
> coefplot2(fit_zipoiss)
```

Regression estimates

We can also try a standard zero-inflated negative binomial model; the default is the "NB2" parameterization (variance $= \mu(1 + \mu/k)$).

```
> fit_zinbinom <- glmmadmb(NCalls~(FoodTreatment+ArrivalTime)*SexParent+
                  offset(logBroodSize)+(1|Nest),
                  data=Owls,
                  zeroInflation=TRUE,
                  family="nbinom")
```

Alternatively, use an "NB1" fit (variance $= \phi\mu$).

```
> fit_zinbinom1 <- glmmadmb(NCalls~(FoodTreatment+ArrivalTime)*SexParent+
                                 offset(logBroodSize)+(1|Nest),
                                 data=Owls,
                                 zeroInflation=TRUE,
                                 family="nbinom1")
```

Relax the assumption that total number of calls is strictly proportional to brood size (i.e. using log(brood size) as an offset):

```
> fit_zinbinom1_bs <- glmmadmb(NCalls~(FoodTreatment+ArrivalTime)*SexParent+
                                 BroodSize+(1|Nest),
                                 data=Owls,
                                 zeroInflation=TRUE,
                                 family="nbinom1")
```
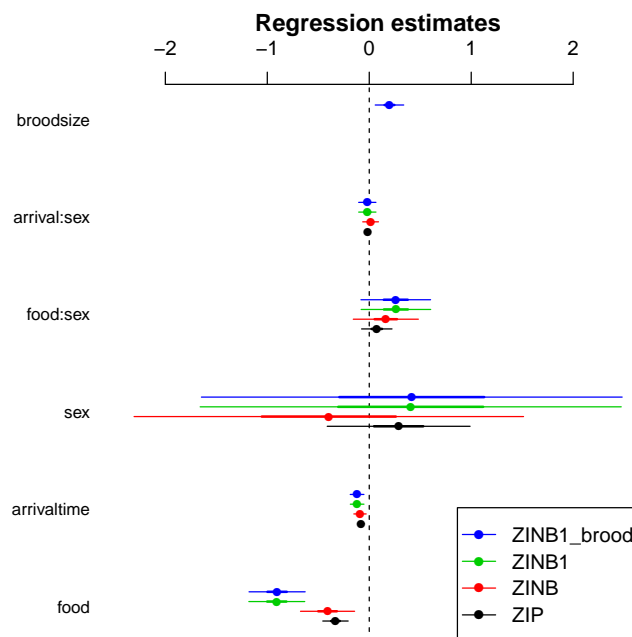
Every change we have made so far improves the fit — changing distributions improves it enormously, while changing the role of brood size makes only a modest (-1 AIC unit) difference:

```
> library(bbmle)
> AICtab(fit_zipoiss,fit_zinbinom,fit_zinbinom1,fit_zinbinom1_bs)
```

```
              dAIC  df
fit_zinbinom1_bs   0.0 10
fit_zinbinom1      1.2  9
fit_zinbinom      68.7  9
fit_zipoiss      637.0  8
```

Compare the parameter estimates:

```
> vn <- c("food","arrivaltime","sex","food:sex","arrival:sex","broodsize")
> coefplot2(list(ZIP=fit_zipoiss,
                 ZINB=fit_zinbinom,
                 ZINB1=fit_zinbinom1,
                 ZINB1_brood=fit_zinbinom1_bs),
           varnames=vn,
           legend=TRUE)
```

Regression estimates

broodsize

arrival:sex

food:sex

sex

arrivaltime

- ZINB1_brood
- ZINB1
- ZINB
- ZIP

food

The standard set of accessors is available:

coef extract (fixed-effect) coefficients

fixef a synonym for coef, for consistency with nlme/lme4

ranef extract random effect coefficients ("BLUPs" or "conditional modes")

residuals extract (Pearson) residuals

`fitted` fitted values

`predict` predicted values (*based only on fixed effects, not on random effects*), possibly with standard errors (*based only on uncertainty of fixed effects*), possibly for new data

`logLik` extract log-likelihood

`AIC` extract AIC

`summary` print summary

`stdEr` extract standard errors of coefficients

`vcov` extract estimated variance-covariance matrix of coefficients

`VarCorr` extract variance-covariance matrices of random effects

`confint` extract confidence intervals of fixed-effect coefficients

Missing: specifying starting values; MCMC

References

[1] A. Roulin and L. Bersier. Nestling barn owls beg more intensely in the presence of their mother than in the presence of their father. *Animal Behaviour*, 74:1099–1106, 2007.

[2] Alain F. Zuur, Elena N. Ieno, Neil J. Walker, Anatoly A. Saveliev, and Graham M. Smith. *Mixed Effects Models and Extensions in Ecology with R*. Springer, 1 edition, March 2009.