

Modeling stand mortality using Poisson mixture models with mixed-effects

Xiong-Qing Zhang⁽¹⁾, Yuan-Cai Lei⁽²⁾, Xian-Zhao Liu⁽²⁾

Stand mortality models play an important role in simulating stand dynamic processes. Periodic stand mortality data from permanent plots tend to be dispersed, and frequently contain an excess of zero counts. Such data have commonly been analyzed using the Poisson distribution and Poisson mixture models, such as the zero-inflated Poisson model (ZIP), and the Hurdle Poisson model (HP). Based on mortality data obtained from sixty Chinese pine (*Pinus tabulaeformis*) permanent plots near Beijing, we added the random-effects to the Poisson mixture models. Results showed that the random-effects in the ZIP model was not convergent, and HP mixed-effects model performed better in modeling stand mortality than the Poisson fixed-effects model, the Poisson mixed-effects model, the ZIP fixed-effects model and the HP fixed-effects model. Moreover, the HP model accounts for two sources of dispersion, the first accounting for extra zeros and the second accounting to some extent for the dispersion due by individual heterogeneity in the positive set. We also found that stand mortality was negatively related to stand arithmetic mean diameter and positively related to dominant height.

Keywords: Hurdle Models, Mixed Models, Poisson Models, Stand Mortality, Zero Inflated Model

Introduction

An important component equation of whole stand models is one that predicts, explicitly or implicitly, stand mortality (trees ha⁻¹) over a period of time (Woollons 1998). However, in contrast to other stand variables, it is often difficult to accurately describe patterns of mortality across stand and site conditions (Keane et al. 2001). This may be because tree mortality is a complex process affected by environmental, pathological, and physiological factors (Yang et al. 2003). Due to the uncertainty of the tree mortality process, mortality remains one of the least understood components of natural stand growth processes (Álvarez-González et al. 2004). Over the years, researches have focused on individual tree mortality for analyzing stand mortality: tree mortality was predicted with a logistic model (Zhang et al. 2010), and aggregated for estimating stand

mortality. Unfortunately, collecting detailed tree information to develop a tree mortality model is costly (Monserud & Sterba 1999). What is more, stand mortality predicted from a tree mortality model often suffers from an accumulation of errors and subsequently poor accuracy and precision (Qin & Cao 2006, Zhang et al. 2010). Woollons (1998) used a two-step method to estimate the stand mortality. However, this method does not admit a well-defined stochastic structure that can serve as a basis for inference (Affleck 2006, Zhang et al. 2012).

Mortality is a complicated stochastic process influenced by several stand characteristics and environment factors, which often exist with complex interactions. It is hardly possible to capture all of the observed variability in empirical mortality data. The evolution of the mixed-effects modeling methodology provided a statistical method capable

of explicit modeling stochastic structure is a possible approach for solving this problem (Calama & Montero 2005).

In addition, it has to be considered that a relatively high number of permanent sample plots has no occurrence of mortality, even over periods of several years. Therefore, mortality data are truncated and characteristically exhibit varying degrees of dispersion and skewness in relation to the mean. Moreover, the data often contain an excess number of zero counts. The least squares method implicitly presumes that the data are Gaussian distributed with constant variances, or at least satisfy the Gauss-Markov conditions. If the least squares method is applied to data with a large proportion of zero counts, the estimated results will be biased. Alternatively, if only nonzero mortality observations are used for model development, then mortality will be overestimated (Woollons 1998, Eid & Øyen 2003). In recent years, there has been considerable interest in models for count data, such as manufacturing defects (Lambert 1992), patent applications (Crepon & Duguet 1997), species abundance (Hall 2000), medical consultations (Gurmu 1997), use of recreational facilities (Shonkwiler & Shaw 1996). There are also a few studies that have addressed this issue in forestry (e.g., Affleck 2006, Fortin & DeBlois 2007, Zhang et al. 2012).

Although many forest models accounted for mixed effects such as tree diameter growth models (Kiernan et al. 2008, Subedi & Sharma 2011), height growth models (Fang & Bailey 2001), and others (Garber & Maguire 2003, Uzoh & Oliver 2006), compared with the substantial literature on cross-sectional zero-inflated count data, relatively few studies have considered mixed effects in the count-data models in forestry (Hall 2000, Rodrigues-Motta et al. 2010), due to the complexity and difficulty in obtaining convergence, especially for applications in forestry (Li et al. 2011). The objective of this study was to develop and compare the Poisson mixture models with and without random-effects for predicting stand mortality over a period of time.

Material and methods

Dataset

A systematic sampling of permanent, square-shaped plots (0.067 ha) was carried out by the Inventory Institute of Beijing Forestry with a 5-year re-measurement interval. Overall, sixty plots were sampled in several Chinese pine (*Pinus tabulaeformis* Carr.) plantations situated on upland sites of Beijing (Fig. 1): 44 plots were measured in 1986, 58 plots in 1991, and 60 plots in 1996, 2001. The final dataset consisted of measurements from 162 stands obtained between 1986 and 2001.

□ (1) Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091 (China); (2) Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091 (China)

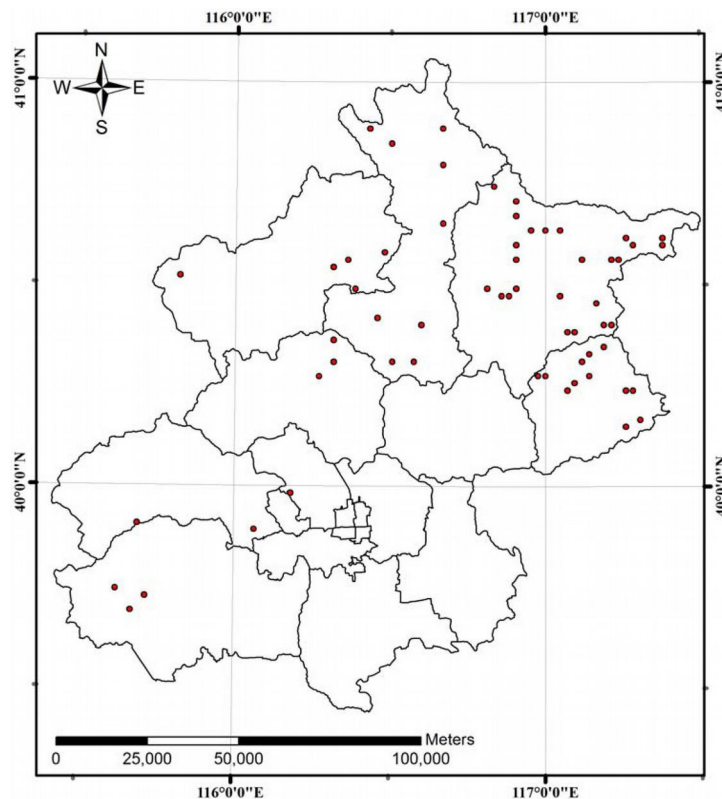
@ Yuan-Cai Lei (ylei@caf.ac.cn)

Received: Apr 26, 2013 - Accepted: Jul 13, 2014

Citation: Zhang X-Q, Lei Y-C, Liu X-Z, 2015. Modeling stand mortality using Poisson mixture models with mixed-effects. iForest 8: 333-338 [online 2014-09-05] URL: <http://www.sisef.it/forest/contents/?id=ifor1022-007>

Communicated by: Renzo Motta

Fig. 1 - Location of the 60 plantations of Chinese pine (*Pinus tabulaeformis* Carr.) studied in this investigation. The black lines on the map represent the boundaries of the Beijing's administrative region.



Summary statistics of stand variables are presented in Tab. 1. The climate in the studied area is warm temperate and semi-humid. Mean annual temperature ranges from 9 to 11 °C, and rainfall varies between 500 and 700 mm.

Variable selection

Independent variables characterized by a meaningful biological interpretation were selected among those describing altitude (*Alt*), site conditions (dominant height, *H*) and stand characteristics (stand age, *A*; stand density, *N*; stand arithmetic mean diameter, *Dm*; relative spacing etc.). Multicollinearity among independent variables was verified by applying the variance inflation factor (VIF) test. According to a common rule-of-thumb, multicollinearity among variables was considered to occur when VIF > 5. In addition,

variables showing $p > 0.05$ after a pairwise Student's *t*-test were discarded from the model.

Poisson model

Count data models are a subset of discrete-response regression models aimed to describe the number of occurrences or counts of an event. Poisson regression is the simplest regression model for count data, and the probability mass function (PMF) is expressed as follows (eqn. 1):

$$P(Y=y) = \frac{\exp(-\lambda)\lambda^y}{y!}$$

where *y* refers to the random variable of stand mortality (dead counts), $y = 0, 1, 2, \dots, n$, and $\lambda > 0$. A Poisson regression model is obtained by relating the mean λ to a vector of independent variables *X*, by $\lambda = \exp(X\beta)$,

Tab. 1 - Summary statistics of stand-level variables. (SD): standard deviation; (SM): stand mortality counts in a plot; (A): stand age; (H): stand dominant height; (N): stand density; (B): stand basal area; (Dm): stand arithmetic mean diameter; (Al): altitude; and (Rs): the relative spacing. The relative spacing was computed as a function of dominant height and stand density: $Rs = (10000/N)^{0.5}/H$.

Variables	Min	Max	Mean	SD
SM (trees plot ⁻¹)	0	38	2.6	6.46
A (years)	12	55	27	8.46
H(m)	2.5	17.4	6.4	2.83
N (trees ha ⁻¹)	132	2164	933	492.24
Dm (cm)	5.8	16.13	9.76	2.4
Rs (trees ha ⁻¹ m ⁻¹)	0.17	3.45	0.71	0.63
Alt (m)	140	1400	449.91	268.09

where β is the vector of regression coefficients to be estimated. A characteristic of the Poisson probability function (eqn. 1) is that the mean and the variance are equal, that is $\text{Var}[Y | X] = E[Y | X] = \lambda$. When data do not fit to the Poisson distribution, it is typically because of overdispersion, i.e., the model's variance exceeds the mean value.

ZIP model

ZIP (zero-inflated Poisson) model is a mixed model combining a Poisson distribution with a point mass at zero. In zero-inflated Poisson model, there are two sources of zeros, deriving both from the point mass and from the count component (Lambert 1992). Actually, two regression equations are created in zero-inflated Poisson models, one predicting whether the count occurs and the other predicting differences on the occurrence of the count (Poisson - Karazsia & van Dulmen 2008). ZIP is a very flexible model for handling data dispersion, and the PMF for ZIP is given by (eqn. 2):

$$P(Y=y) = \begin{cases} p + (1-p)\exp(-\lambda) & y=0 \\ \frac{(1-p)\exp(-\lambda)\lambda^y}{y!} & y>0 \end{cases}$$

where *p* is the probability of a zero count in excess, and (1-*p*) is the probability of belonging to the Poisson component. Logistic regression is commonly used to fit the point mass component: $\text{logit}(p) = \text{Log}(p/1-p) = X\delta$. As for the Poisson component, it is obtained from a vector of independent variables, $\lambda = \exp(X\beta)$, where *X* is a vector of independent variables, δ and β are vectors of the parameters to be estimated.

HP model

HP (Hurdle Poisson) model, originally proposed by Mullahy (1986) in the econometrics field (Cameron & Trivedi 1998), is another model class capable of dealing with excess zero counts. However, HP is slightly different from ZIP model with all zero counts from two different sources, and assumes that zero counts might come from a single statistical process (Liu & Cela 2008).

Similar to ZIP model, HP model is the combination of a logit regression modeling point mass at zero and a truncated Poisson regression modeling count component. Its PMF is expressed as follows (eqn. 3):

$$P(Y=y) = \begin{cases} p & y=0 \\ \frac{(1-p)\exp(-\lambda)\lambda^y}{(1-\exp(-\lambda))y!} & y>0 \end{cases}$$

The point mass component is obtained by logistic regression using $\text{logit}(p) = \text{Log}(p/1-p) = X\delta$, while the Poisson component is obtained from a vector of independent variables, $\lambda = \exp(X\beta)$.

In this study, a plot-level random-effects

parameter was added to the intercept for the Poisson model, and to the intercept of the Poisson component for the estimation of positive mortality counted for ZIP model and HP model. The random-effect parameter was defined as: $u \sim N(0, \nu)$, where ν is the variance.

All parameter vectors can be estimated through the optimization of the likelihood function, that is, by applying the maximum likelihood method (ML). The unstructured covariance structure (Littell et al. 1996) was used for describing the variance-covariance structure of random effects. Parameter estimation was implemented using the SAS/STAT NL MIXED procedure.

Model selection and goodness-of-fit

Performances of the Poisson fixed-effects model, ZIP fixed-effects model, HP fixed-effects model and corresponding mixed-effect models calibrated with the same data set can be compared by using the log-likelihood values $[-2L(\phi, y)]$, the Akaike information criterion (AIC), as well as the Bayesian information criterion (BIC). Smaller values of AIC, BIC and $-2L(\phi, y)$ indicate better performances for the model considered. Both AIC and BIC rely on a penalized maximum log-likelihood value. As the penalty is based on the number of model parameters, the use of these statistics ensures the best trade-off between the goodness-of-fit and the number of parameters included in the model. The penalty is more influential in the BIC, making this criterion more conservative than the AIC (Fortin & DeBlois 2007).

The Vuong test is a popular method to compare two non-nested models for count data (Vuong 1989). If we define (eqn. 4):

$$m_i = \log \left(\frac{P_1(Y_i | X_i)}{P_2(Y_i | X_i)} \right)$$

where $P_j(Y_i | X_i)$ is the predicted probability of the observed count for the i -th case in the j -th model, then Vuong statistic to test the hypothesis $E(m_i) = 0$ is expressed as follows (eqn. 5):

$$V = \frac{\sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n m_i \right)}{\sqrt{\frac{1}{n} \sum_{i=1}^n (m_i - \bar{m})^2}}$$

where n represents the number of count classes. The first model should be preferred when $V > 1.96$, while the second should be preferred in the case of $V < -1.96$ (Liu & Cela 2008).

Because of the use of maximum log-likelihoods, AIC, BIC, and Vuong parameters are relative statistics, as they do not ensure that the fit of the “best” model is good. Hence, residual plots (residual values between predicted and observed probabilities plotted

against the count class j) were used to detect any predictive bias of the models and assess their goodness-of-fit (Lambert 1992, Fortin & DeBlois 2007). The residual r_j , between predicted probabilities and observed probabilities was computed as (eqn. 6):

$$r_j = \sum_{i=1}^n \left[\frac{P(y_i = j)}{n} - \frac{\# y_i = j}{n} \right]$$

where $\#$ represents the frequency of observations y_i in the count class j , n is the number of count classes, and $P(y = j)$ is the predicted probability that an observation belongs to the j -th count class.

Results and discussion

In this study, a relatively high number of the plots considered has no occurrence of mortality. Mortality data are zero-bounded and characteristically exhibit varying degrees of dispersion around the mean and skewness (Fig. 2). According to the VIF test, no multicollinearity was found among the independent variables included (VIF < 5).

During the processes of estimating all the six models considered, we found that the ZIP mixed-effects model was not convergent, so the following results were obtained from the remnant five models.

Results showed that the statistics $-2L(\phi, y)$, AIC and BIC obtained for the HP mixed-effects model were much smaller than those of the other models (Poisson fixed-effects model, Poisson mixed-effects model, ZIP fixed-

effects model and HP fixed-effects model - Tab. 2), suggesting that the HP mixed-effects model provided the “best fit”.

Poisson fixed-effects and Poisson mixed-effects models were found to largely underestimate the zero-class counts, while ZIP fixed-effects, HP fixed-effects and HP mixed-effects models exactly estimated the zero dead counts (Fig. 3). The residuals r_j of the Poisson mixed-effects model were smaller than those obtained for the Poisson fixed-effects model, while the HP mixed-effects model showed lower residuals than those obtained from the HP fixed-effects, ZIP fixed-effects and Poisson mixed-effects model (Fig. 3).

Additionally, the Vuong test-statistic V obtained from the comparison between the Poisson mixed-effects model and the Poisson fixed-effects model was 4.5, while a value $V = 4.8$ was estimated by comparing the HP mixed-effects model with the HP fixed-effects model. This means that models with random effects had better performances than fixed-effects models. We also compared the HP fixed-effects model with the ZIP fixed-effects model, remarking that the former model outperformed the latter ($V = 5.0$).

Overall, the best fit was detected for the HP fixed-effects and the HP mixed-effects model, therefore only these models were used in further analysis. Some variables with $p > 0.05$ after the t -test were removed from the models (Tab. 3). In the positive count component of the models, all variables were

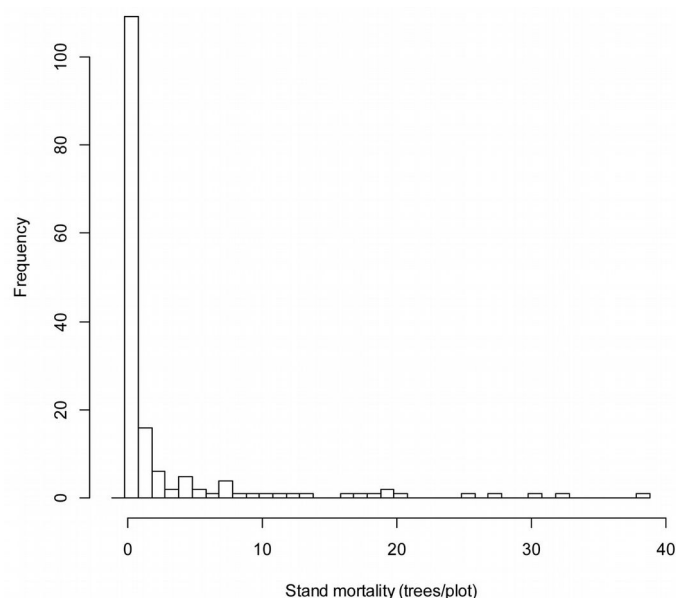


Fig. 2 - Histogram of stand mortality for Chinese pine obtained from the sixty plots selected in the study area.

Tab. 2 - Fit statistics of Poisson fixed-effects model, Poisson mixed-effects model, ZIP fixed-effects, HP fixed-effects model, and HP mixed-effects model.

Statistic	Poisson-fixed	Poisson-mixed	ZIP-fixed	HP-fixed	HP-mixed
$-2L(\phi, y)$	1305.8	658.9	783.1	783	513.8
AIC	1319.8	670.9	795.1	795	527.8
BIC	1341.4	683.8	813.6	813.6	542.8

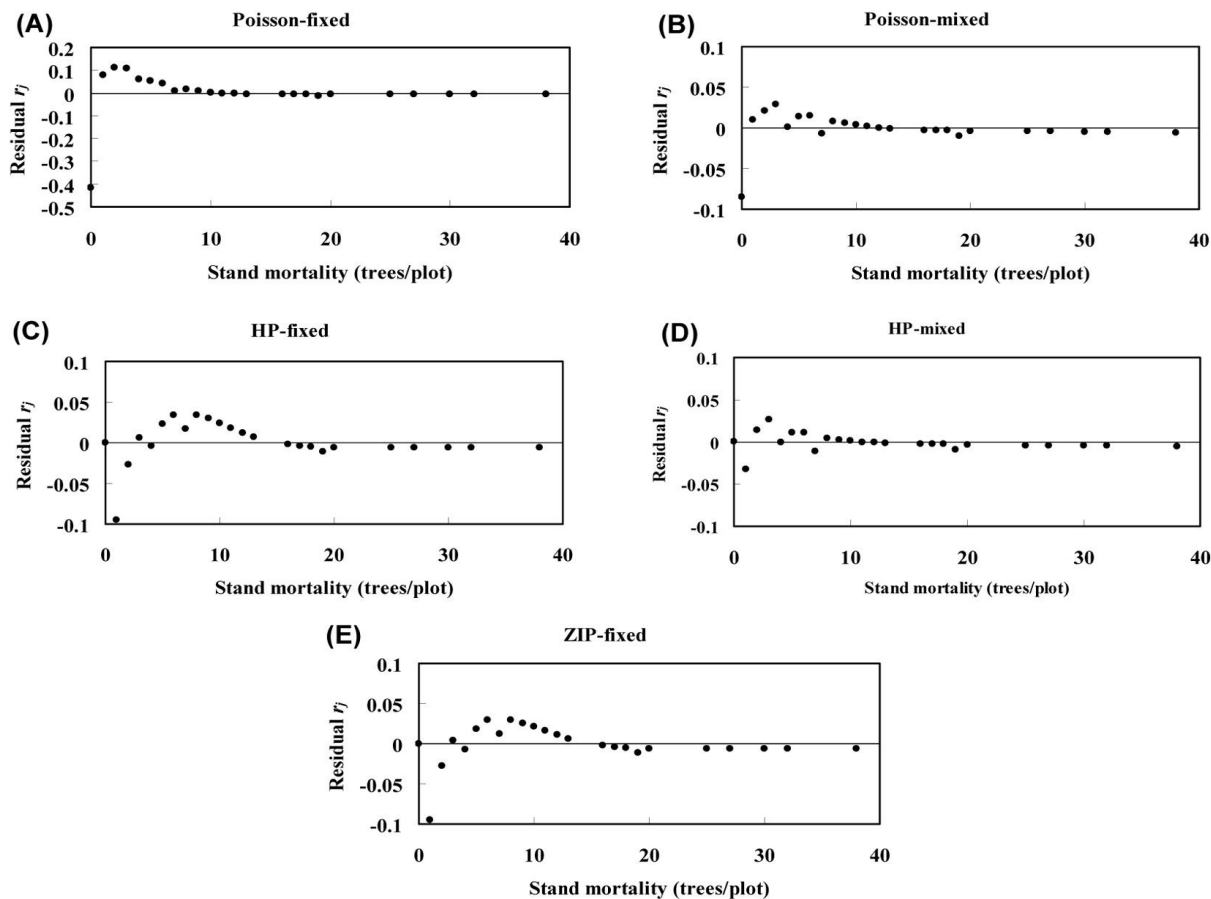


Fig. 3 - Plots of residuals for the Poisson fixed-effects model (A), Poisson mixed-effects model (B), HP fixed-effects model (C), HP mixed-effects model (D), and ZIP fixed-effects model (E). (r_j): residuals between predicted and observed probability as computed by using the eqn. 6.

significant at the 0.05 level (Tab. 3). Such variables have a meaningful biological interpretations in terms of tree mortality. Relative spacing is an indirect measure of the competition within the stand, being related to tree density (number of trees per hectare) and stand dominant height. According to previous studies (Diéguez-Aranda et al. 2005), stand mortality decreases when the relative spacing is larger. Similarly, the effect of stand arithmetic mean diameter (size index) was negative (*i.e.*, the greater Dm , the smaller stand mortality), indicating that stand

mortality is more likely in forests with many small trees compared to forests with larger trees (Juknys et al. 2006). Due to the random effects on plots, altitude was not significant in the mixed models (Poisson mixed-effects model, HP mixed-effects model). This may be because the random parameter explained the effects of altitude. In the zero component of the models, all the parameters listed in Tab. 3 were also significant at the 0.05 level. In this study, site quality (as reflected by the dominant height) was positively related with the probability of mortality. Accordingly,

several studies reported a higher mortality in sites showing higher productivity (Yao et al. 2001, Diéguez-Aranda et al. 2005). On the other hand, Woollons (1998) found that a higher mortality was related to lower site productivity. Therefore, site quality effects should be considered with caution before their inclusion in mortality models.

In a tree mortality context, the use of HP models seems to be the most appropriate when applied to cases with no mortality, as it occurs in a stand before the competition among trees takes place. Once competitive pressures within the stand exceeds a certain threshold, positive mortality occurs consistently and in accordance with a ZIP probability mass function (Affleck 2006).

In our study, the application of both the above models led to the same qualitative results and gave very similar model fitting performances (Tab. 2, Fig. 3). These models have three main advantages. Firstly, by allowing distinct functions for the two components, the predicted stand mortality may be compared with that obtained assuming a Poisson-shaped data distribution. Secondly, the zero component and the positive count component of the two models may be di-

Tab. 3 - Estimations of parameters of the two model components in the HP fixed-effects and the HP mixed-effects models. (SE): standard error.

Model Component	Parameter	HP-fixed			HP-mixed		
		Estimate	SE	p-value	Estimate	SE	p-value
Positive count component	Intercept	4.0177	0.3249	<0.001	6.2386	1.0851	<0.001
	R_s	-1.5462	0.3091	<0.001	-3.3797	1.1974	<0.01
	D_m	-0.1187	0.0244	<0.001	-0.3171	0.0701	<0.001
	v	-	-	-	1.764	0.5998	<0.01
Zero component	Intercept	-3.321	1.1945	<0.01	-3.321	1.1945	<0.01
	H	0.1846	0.0903	<0.05	0.1846	0.0903	<0.05
	R_s	4.7581	1.222	<0.001	4.758	1.222	<0.001

rectly interpreted in terms of the empirical features of the data. Thirdly, they perform better in modeling the data variability observed, therefore ensuring an appropriate use of the information collected (Barry & Welsh 2002, Zhang et al. 2012).

In this study, we used Poisson mixture model with mixed-effects to analyze stand mortality. Random-effects of both the Poisson model and HP model were significant at the 0.05 level. However, the ZIP model was not convergent after the random-effects were incorporated in the intercept of positive count component. Min & Agresti (2005) found that fitting a ZIP random effect model was more complex than fitting a HP random effect model. Li et al. (2011) incorporated the random effects in the count data models for analyzing tree in-growth, finding an improved model performance. Overall, both the Poisson and the HP model were largely improved by the inclusion of random effects, accounting for most of the variation among plots and data sources.

Acknowledgments

This work was carried out within the frame of the following projects: the Chinese State “TWELVE FIVE” Science and Technology Project (No. 2012BAD22B0201), and the “863” program from the Ministry of Science and Technology of China (No. 2012AA12A306). The authors are grateful to two anonymous reviewers for their valuable suggestions and comments on an earlier version of the manuscript.

References

- Affleck DLR (2006). Poisson mixture models for regression analysis of stand-level mortality. *Canadian Journal of Forest Research* 36: 2994-3006. - doi: [10.1139/x06-189](https://doi.org/10.1139/x06-189)
- Álvarez-González JG, Dorado FG, Gonzalez ADR, López-Sánchez CA, Gadow, KV (2004). A two-step mortality model for even-aged stands of *Pinus radiata* D.Don in Galicia (Northwestern Spain). *Annals of Forest Science* 61: 439-448. - doi: [10.1051/forest:2004037](https://doi.org/10.1051/forest:2004037)
- Barry SC, Welsh AH (2002). Generalized additive modeling and zero inflated count data. *Ecological Modelling* 157: 179-188. - doi: [10.1016/S0304-3800\(02\)00194-1](https://doi.org/10.1016/S0304-3800(02)00194-1)
- Calama R, Montero G (2005). Multilevel linear mixed model for tree diameter increment in Stone pine (*Pinus pinea*): a calibrating approach. *Silva Fennica* 39: 37-54. [online] URL: [http://www.researchgate.net/publication/228627412_Multilevel_linear_mixed_model_for_tree_diameter_increment_in_stone_pine_\(Pinus_pinea\)_a_calibrating_approach/file/e0b49525fea47da547.pdf](http://www.researchgate.net/publication/228627412_Multilevel_linear_mixed_model_for_tree_diameter_increment_in_stone_pine_(Pinus_pinea)_a_calibrating_approach/file/e0b49525fea47da547.pdf)
- Cameron AC, Trivedi PK (1998). *Regression analysis of count data*. Cambridge University Press, Cambridge, UK, pp. 123-137.
- Crepon B, Duguet E (1997). Research and development, competition and innovation - pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. *Journal of Econometrics* 79: 355-378. - doi: [10.1016/S0304-4076\(97\)00027-4](https://doi.org/10.1016/S0304-4076(97)00027-4)
- Diéguez-Aranda U, Castedo-Dorado F, Álvarez-González JG, Rodríguez-Soalleiro R (2005). Modelling mortality of Scots pine (*Pinus sylvestris* L.) plantations in the northwest of Spain. *European Journal of Forest Research* 124(2): 143-153. - doi: [10.1007/s10342-004-0043-5](https://doi.org/10.1007/s10342-004-0043-5)
- Eid T, Øyen BH (2003). Models for prediction of mortality in even-aged forest. *Scandinavian Journal of Forest Research* 18: 64-77. - doi: [10.1080/02827581.2003.10383139](https://doi.org/10.1080/02827581.2003.10383139)
- Fang Z, Bailey RL (2001). Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. *Forest Science* 47: 287-300. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2001/00000047/00000003/art00002>
- Fortin M, DeBlois J (2007). Modeling tree recruitment with zero-inflated models: the example of hardwood stands in Southern Quebec, Canada. *Forest Science* 53: 529-539. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2007/00000053/00000004/art00008>
- Garber SM, Maguire DA (2003). Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. *Forest Ecology and Management* 179: 507-522. - doi: [10.1016/S0378-1127\(02\)00528-5](https://doi.org/10.1016/S0378-1127(02)00528-5)
- Gurmu S (1997). Semi-parametric estimation of hurdle regression models with an application to Medicaid utilization. *Journal of Applied Econometrics* 12: 225-242. - doi: [10.1002/\(SICI\)1099-1255\(199705\)12:3<225::AID-JAE445>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1099-1255(199705)12:3<225::AID-JAE445>3.0.CO;2-Y)
- Hall DB (2000). Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics* 56: 1030-1039. - doi: [10.1111/j.0006-341X.2000.01030.x](https://doi.org/10.1111/j.0006-341X.2000.01030.x)
- Juknys R, Vencloviene J, Jurkonis N, Bartkevicius E, Sepietiene J (2006). Relation between individual tree mortality and tree characteristics in a polluted and non-polluted environment. *Environment Monitoring and Assessment* 121: 519-542. - doi: [10.1007/s10661-005-9152-y](https://doi.org/10.1007/s10661-005-9152-y)
- Karazsia BT, van Dulmen MH (2008). Regression models for count data: illustrations using longitudinal predictors of childhood injury. *Journal of Pediatric Psychology* 33: 1076-1084. - doi: [10.1093/jpepsy/jsn055](https://doi.org/10.1093/jpepsy/jsn055)
- Keane RE, Austin J, Field C, Huth A, Lexer MJ, Peters D, Solomon A, Wyckoff P (2001). Tree mortality in gap models: application to climate change. *Climatic Change* 51: 509-540. - doi: [10.1023/A:1012539409854](https://doi.org/10.1023/A:1012539409854)
- Kiernan DH, Bevilacqua E, Nyland RD (2008). Individual-tree diameter growth model for sugar maple trees in uneven-aged northern hardwood stands under selection system. *Forest Ecology and Management* 256: 1579-1586. - doi: [10.1016/j.foreco.2008.06.015](https://doi.org/10.1016/j.foreco.2008.06.015)
- Lambert D (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* 34: 1-14. - doi: [10.2307/1269547](https://doi.org/10.2307/1269547)
- Li R, Weiskittel AR, Kershaw JAJ (2011). Modeling annualized occurrence, frequency, and composition of ingrowth using mixed-effects zero-inflated models and permanent plots in the Acadian forest region of North America. *Canadian Journal of Forest Research* 41: 2077-2089. - doi: [10.1139/x11-117](https://doi.org/10.1139/x11-117)
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996). *SAS system for Mixed Models*. SAS Institute Inc., Cary, NC, USA, pp. 31-63.
- Liu W, Cela J (2008). Count data models in SAS. In: *In Proceedings of the “SAS Global Forum 2008 Conference”*. San Antonio (TX, USA) 16-19 March 2008. SAS Institute, Cary, NC, USA, pp. 1-12.
- Min Y, Agresti A (2005). Random-effects models for repeated measures of zero-inflated count data. *Statistical Modelling* 5: 1-19. - doi: [10.1191/1471082X05st0840a](https://doi.org/10.1191/1471082X05st0840a)
- Monserud RA, Sterba H (1999). Modeling individual tree mortality for Austrian forest species. *Forest Ecology and Management* 113: 109-123. - doi: [10.1016/S0378-1127\(98\)00419-8](https://doi.org/10.1016/S0378-1127(98)00419-8)
- Mullahy J (1986). Specification and testing of some modified count data models. *Journal of Econometrics* 33: 341-365. - doi: [10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3)
- Qin JH, Cao QV (2006). Using disaggregation to link individual-tree and whole-stand growth models. *Canadian Journal of Forest Research* 36 (4): 953-960. - doi: [10.1139/x05-284](https://doi.org/10.1139/x05-284)
- Rodrigues-Motta M, Gianola D, Heringstad B (2010). A mixed effects model for overdispersed zero inflated poisson data with an application in animal breeding. *Journal of Data Science* 8: 379-396. [online] URL: <http://www.jds-online.com/files/JDS-620.pdf>
- Shonkwiler J, Shaw W (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural and Resource Economics* 21: 210-219. [online] URL: <http://www.jstor.org/stable/40986910>
- Subedi N, Sharma M (2011). Individual-tree diameter growth models for black spruce and jack pine plantations in northern Ontario. *Forest Ecology and Management* 261: 2140-2148. - doi: [10.1016/j.foreco.2011.03.010](https://doi.org/10.1016/j.foreco.2011.03.010)
- Uzoh FCC, Oliver WW (2006). Individual tree height increment model for managed even-aged stands of ponderosa Pine throughout the western United States using linear mixed effects models. *Forest Ecology and Management* 221: 147-154. - doi: [10.1016/j.foreco.2005.09.012](https://doi.org/10.1016/j.foreco.2005.09.012)
- Vuong QH (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57: 307-333. - doi: [10.2307/1912557](https://doi.org/10.2307/1912557)
- Woollons RC (1998). Even-aged stand mortality estimation through a two-step regression process. *Forest Ecology and Management* 105: 189-195. - doi: [10.1016/S0378-1127\(97\)00279-X](https://doi.org/10.1016/S0378-1127(97)00279-X)
- Yang Y, Titus SJ, Huang S (2003). Modeling individual tree mortality for white spruce in Al-

- berta. *Ecological Modelling* 163(5): 209-222. - doi: [10.1016/S0304-3800\(03\)00008-5](https://doi.org/10.1016/S0304-3800(03)00008-5)
- Yao X, Titus S, MacDonald SE (2001). A generalized logistic model of individual tree mortality for aspen, white spruce, and lodgepole pine in Alberta mixedwood forests. *Canadian Journal of Forest Research* 31: 283-291. - doi: [10.1139/x00-162](https://doi.org/10.1139/x00-162)
- Zhang X, Lei Y, Cao QV (2010). Compatibility of stand basal area predictions based on forecast combination. *Forest Science* 56 (6): 552-557. [online] URL: <http://www.ingentaconnect.com/content/saf/fs/2010/00000056/00000006/art00004>
- Zhang X, Lei Y, Cai D, Liu F (2012). Predicting tree recruitment with negative binomial mixture models. *Forest Ecology and Management* 270: 209-215. - doi: [10.1016/j.foreco.2012.01.028](https://doi.org/10.1016/j.foreco.2012.01.028)